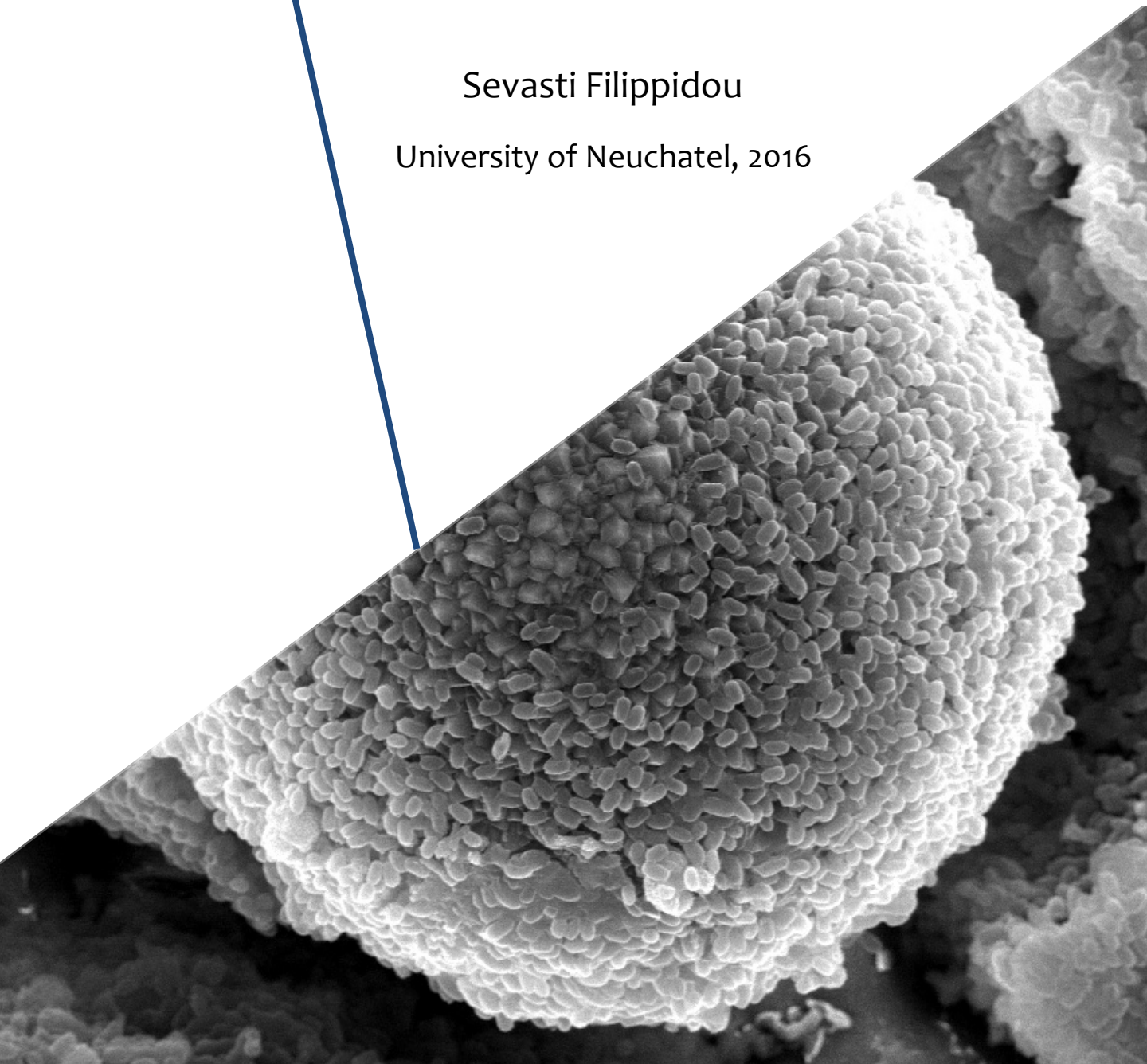
Sporulation Capability and Metabolic Mechanisms of Endospore-Forming Firmicutes under Conditions Limiting for Growth and Survival

Sevasti Filippidou

University of Neuchatel, 2016



Sporulation Capability and Metabolic Mechanisms of Endospore-Forming Firmicutes under Conditions Limiting for Growth and Survival

A dissertation submitted to the University of Neuchatel

for the degree of

Docteure ès Sciences

by Sevasti Filippidou, MSc Molecular Genetics

Accepted by the Jury:

Prof. Pilar Junier, thesis director, University of Neuchatel

Prof. Maarten Voordow, University of Neuchatel

Prof. Melanie Blokesch, EPFL, Lausanne

Dr. David Russel Johnson, Eawag, Dübendorf

Dr. Paul Herron, University of Strathclyde, Glasgow, UK

Defended the 11th April 2016

University of Neuchatel

UNIVERSITÉ DE NEUCHÂTEL

Faculté des sciences
Secrétariat-décanat de Faculté
Rue Emile-Argand 11
2000 Neuchâtel - Suisse
Tél: + 41 (0)32 718 2100
E-mail: secretariat.sciences@unine.ch

IMPRIMATUR POUR THESE DE DOCTORAT

La Faculté des sciences de l'Université de Neuchâtel autorise l'impression de la présente thèse soutenue par

Madame Sevasti Filippidou

Titre:

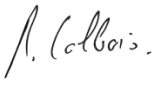
"Sporulation capability and metabolic mechanisms of endospore-forming Firmicutes under conditions limiting for growth and survival"

sur le rapport des membres du jury composé comme suit:

- Prof. Pilar Junier, directrice de thèse, Université de Neuchâtel, Suisse
- Prof. ass. Maarten Voordouw, Université de Neuchâtel, Suisse
- Prof. Melanie Blokesch, EPF Lausanne, Suisse
- Dr David R. Johnson, EAWAG, Dübendorf, Suisse
- Dr Paul Herron, University of Strathclyde, Glasgow, UK

Neuchâtel, le 28 avril 2016

Le Doyen, Prof. B. Colbois



This work is dedicated to

The memory of *Marilena Vourkou*, for her inspiration to life,

Entropia, that has shaped me to what I am,

Dimitris Serafis, without whom I wouldn't have reached that far.

Acknowledgements

Firstly, I would like to thank my professor Pilar Junier for being a very supportive tutor and a real mentor throughout all these years. Pilar has taught me, guided me, encouraged me, and gave me opportunities I never imagined I could have. She is constantly showing how academic science should be and inspires her team, being a reassurance for brighter days.

I would also like to thank Mrs. Nicole Jeanneret for having introduced me to the lab and who has been a good-hearted advisor on techniques, laboratory practices and French since day one.Many thanks also go to all my lab collaborators, and more specifically the 'spores' team, Dr. Matthieu Bueche, Dr. Ludovic Roussel-Delif, Dr. Sathiya Ganesan, Mr. Loic Sauvain, Mrs. Marion Jaussi and Mr. Christophe Paul for their help, contribution and pleasant atmosphere in the group. In particular, Dr. Tina Wunderlin has generously coached, advised, supported, and shared with me in any possible way. I thank her for being there to encourage me since I could never imagine for a better colleague, office mate and friend. I am most grateful to Dr. Thomas Junier for almost all bioinformatic analysis and for patiently spending his time teaching me phylogenetic tools and rules of scientific writing. Moreover, I would like to thank Dr. Radu Slobodeanu for his statistics advice for the article presented in chapter three. Finally, I am greatful to all our collaborators, and especially the team of Dr. Patrick S. Chain, without whose valuable contributions, this work would never be completed.

Thanks also go to all the students who have worked with me, during their learning stage: Mrs. Amandine Pillonel, Mr. Andrej Al-Dourobi, Mrs. Ilona Palmieri, technicians; Mr. Patricio Munoz-Ruffat, undergraduate student and Mr. Fabio Palmieri, master student. I hope their learning procedure was as constructive as mine during our collaboration.

Special thanks to Mrs. Anna Voulgari for being the perfect ecological eye during sampling campaigns, and who, along with Mrs. Vassiliki Mpelogianni, Dr. Tina Wunderlin and Dr. Nikos Mitakidis have critically read my general introduction. Mrs. Monica Albini has patiently helped me during the last hours before submission and I thank her for that. Dr. Vincent Hervé has translated my summary into "good" French, for which I am most grateful.

I thank my thesis jury for accepting to evaluate this work and provide feedback.

Special thanks go to what I call "an army of people" that helped and supported me throughout these years and which include my family and dearest friends. Although some of them are already acknowledged, all of them have always been available to support and encourage me in any possible manner.

Cover photo: Endospore-forming Firmicutes forming a sphere with copper carbonate crystals. Scanning Electron Microscopy. Photo taken by S. Filippidou. LAMUN 2012

Summary

Life, as we know it, has physical and chemical limits. More precisely, physical and chemical environmental parameters set the thresholds for reproduction, metabolism and survival for the organisms known to date. Environmental conditions unsuitable for survival and development are the rule rather than the exception in most habitats. Microorganisms have developed various strategies to withstand environmental conditions that limit active growth. A microbial group that displays a large array of strategies to resist adversity is endospore-forming Firmicutes. These strategies range from the formation of resting states (endospores), to biofilms and metabolic adaptation. These strategies are a costly biological investment and therefore might affect their success.

Endospore-forming Firmicutes have been isolated from various environments, including extreme habitats. Paradoxically, in diversity studies they are either absent from these environments or they represent a small fraction of the microbial community. Using a profile analysis of the *spoOA* gene, unique to endospore-forming Firmicutes, we have shown that they cannot be found in publically available metagenomic datasets, despite the fact that many of these datasets correspond to well-known habitats for endospore-forming Firmicutes. We have shown that this bias is likely due to the fact that commonly used methodological approaches are inefficient. Moreover, we have shown experimentally that an improved DNA extraction can improve detection in amplicon sequencing; however, this was not the case for shotgun classification. Although this group is known to colonize every habitat on Earth, endospore-firming Firmicutes are not *prevalent* in all habitats. Their energy-demanding survival strategies become an actual benefit only when multiple physical and chemical limits of life are present. Extremity favors the presence of the survival strategies deployed by endospore-forming Firmicutes.

More specifically, we have shown that extreme environmental conditions are an important factor for the survival strategy of sporulation to evolve. Two novel discoveries support this suggestion. Firstly, we have shown that extremity is a driving force for sporulation in a species that was not known to sporulate, *Serratia ureilytica*. Secondly, we demonstrated that the ability to sporulate is *not* lost when there is environmental pressure. That was the case of *Kurthia* spp. a previously known asporogenic genus. The common ancestor of *Kurthia*, and the endospore-forming Firmicutes was able to produce spores, however sporulation is considered to have been lost within the lineage of *Kurthia*. The genomic analysis and the microscopic observation of a *Kurthia* sp. strain isolated from a geothermal reservoir reveal that the sporulation pathway has not been lost, and that *Kurthia* is not an asporogenic but rather a *cryptosporulating* genus.

The survival strategy of sporulation makes endospore-forming Firmicutes capable of tolerating adverse conditions and thriving in extreme environments. A novel species that has a particular ecological niche in geothermal reservoirs was discovered; *Anoxybacillus geothermalis* was revived under laboratory conditions and is hypothesized to have remained inactive in the reservoir since the Permian age. More isolates belonging to the same species were also discovered in different geothermal reservoirs.

Extreme environments do not only allow endospore-forming Firmicutes to deploy their sporulation strategy but also their high metabolic diversity. Manganese oxidation and copper reduction of endospore-forming Firmicutes in natural, uncontaminated environments, were studied leading to the conclusion that metal tolerance is a widespread phenomenon in unrelated aerobic endospore-forming Firmicutes from natural uncontaminated environments. Finally, in saline habitats, both metabolic strategies are deployed, resulting in an impressive diversity of endospore-forming Firmicutes.
Biochemical, genomic, ecological and environmental data are pieces to fill in the puzzle of adaptations of endospore-forming Firmicutes in extreme habitats.

Keywords: Microbial Ecology, Firmicutes, endospore-forming Firmicutes, sporulation, *Kurthia*, *Serratia*, *Anoxybacillus geothermalis*, Manganese Oxidation, Copper Tolerance, Halophiles.

Résumé

La vie, telle que nous la connaissons, possède des limites physiques et chimiques. Plus précisément, des paramètres environnementaux physiques et chimiques limitent la reproduction, le métabolisme et la survie des organismes vivants décrits à ce jour. Des conditions environnementales défavorables à la survie et au développement biologique sont la règle plutôt que l'exception dans la plupart des habitats. Les microorganismes ont cependant développé différentes stratégies pour résister aux conditions environnementales qui limitent leur croissance. Parmi ces microorganismes, le groupe des Firmicutes endosporulantes présente de nombreuses stratégies pour résister à ces contraintes environnementales, allant de l'état de dormance (endospore), à la formation de biomembranes en passant par l'adaptation métabolique. Ces stratégies sont un investissement biologique coûteux pour les organismes et par conséquent, peuvent influencer leur succès.

Les Firmicutes endosporulantes ont été isolées dans des environnements variés, y compris dans des habitats extrêmes. Paradoxalement, dans les études de diversité elles sont, soit absentes de ces environnements, soit elles représentent une faible fraction de la communauté microbienne. A l'aide d'une analyse de profil du gène *spoOA*, un gène spécifique des Firmicutes endosporulantes, nous avons montré que ces dernières ne pouvaient pas être détectées dans les métagénomes publiques disponibles, et ce malgré le fait que plusieurs de ces métagénomes étaient issus d'habitats connus pour abriter des Firmicutes endosporulantes. Nous avons mis en évidence que ce biais était dû au fait que les approches méthodologiques couramment utilisées étaient inefficaces. De plus, nous avons démontré expérimentalement qu'une extraction d'ADN optimisée permettait d'améliorer la détection de Firmicutes endosporulantes par séquençage d'amplicon. Cependant ce n'était pas le cas pour le séquençage *shotgun*. Bien que ce groupe bactérien soit connu pour sa capacité à coloniser tous les habitats sur Terre, les Firmicutes endosporulantes ne sont pas fréquemment retrouvées dans tous les habitats. Leurs stratégies de survie étant énergétiquement coûteuses, elles ne deviennent un avantage uniquement lorsque plusieurs contraintes physiques et chimiques sont présentes. Les conditions extrêmes favorisent la présence des stratégies de survie déployées par les Firmicutes endosporulantes.

Plus spécifiquement, nous avons démontré que des conditions environnementales extrêmes étaient un facteur important pour l'apparition de la sporulation comme stratégie de survie. Deux nouvelles découvertes supportent cette hypothèse. Premièrement, nous avons montré que les conditions extrêmes étaient une force motrice de la sporulation pour une espèce qui n'était pas connue pour sporuler, *Serratia ureilytica*. Deuxièmement, nous avons démontré que la capacité de sporuler n'était pas perdue lorsqu'il y avait une pression environnementale. C'était le cas de *Kurthia* spp., un genre décrit jusqu'à présent comme asporogénique. L'ancêtre commun de *Kurthia*, une Firmicutes endosporulante, était capable de produire des spores, cependant cette capacité était considérée comme perdue dans la lignée de *Kurthia*. L'analyse génomique ainsi que des observations microscopiques d'une souche de *Kurthia* sp. isolée d'un

réservoir géothermique ont révélé que la voie de sporulation n'avait pas été perdue et que *Kurthia* n'est pas un genre asporogénique mais un genre cryptosporulant.
La stratégie de survie de sporulation rend les Firmicutes endosporulantes capables de tolérer des conditions défavorables et de prospérer dans des environnements extrêmes. Une nouvelle espèce, *Anoxybacillus geothermalis*, a été découverte dans une niche écologique particulière, les réservoirs géothermiques. Cette souche a été remise en culture au laboratoire et nous avons émis l'hypothèse que cette souche était inactive dans le réservoir depuis le Permien. D'autres isolats appartenant à la même espèce ont également été découverts dans différents réservoirs géothermiques.
Les environnements extrêmes permettent aux Firmicutes endosporulantes d'utiliser leur stratégie de sporulation mais également de tirer profit de leur grande diversité métabolique. L'analyse des processus d'oxydation du manganèse et de réduction du cuivre par les Firmicutes endosporulantes dans des environnements naturels et non contaminés a révélé que la tolérance aux métaux est un phénomène largement répandu dans les environnements non contaminés, y compris parmi des Firmicutes endosporulantes aérobies ne présentant pas de lien de parenté. Enfin, dans les milieux salins, ces deux stratégies de survie sont utilisées et génèrent une impressionnante diversité de Firmicutes endosporulantes.
En conclusion, l'utilisation de données biochimiques, génomiques, écologiques et environnementales a permis de mieux comprendre l'adaptation des Firmicutes endosporulantes aux environnements extrêmes.
Mots-clés: Ecologie Microbienne, Firmicutes, Firmicutes formatrices des endospores, sporulation, *Kurthia*, *Serratia*, *Anoxybacillus geothermalis*, Oxidation du Manganese, Resistance au Cuivre, Halophiles.

Table of Contents

1. General Introduction

1.1. Thesis outline

This thesis studies two survival strategies that drive the ecology of endospore-forming Firmicutes (EFF): their sporulation capability and their diverse metabolic activity. Although both abilities are encoded in the genetic material of EFF, they are mostly expressed under conditions limiting growth and survival.

In the first part of this dissertation, sporulation, observed in four bacterial taxa so far, is explored from an ecological point of view, trying to understand the costs and benefits of this survival strategy. A primary hypothesis is formulated: since sporulation is a beneficial strategy, spore-forming bacteria should survive under unfavorable conditions and prevail everywhere. The ubiquity of EFF, however, has until now not been confirmed, based on previous literature.

Chapter two explores whether the under-detection of EFF is due to methodological biases. This hypothesis was evaluated by testing the detection of EFF in metagenomic datasets. Two molecular markers unique to this bacterial group (*spo0A* and *gpr*) were screened for in 73 datasets, and found to be absent, with the exception of *spo0A* present in three mammalian gut microbiomes. An improved DNA extraction method resulting in the detection of a large diversity of endospore-formers in amplicon sequencing of the 16S rRNA and *spo0A* genes, once applied to a sample from a geothermal spring, was insufficient to overcome the limitations for detecting EFF in whole-genome metagenomic analysis. The results showed limitations in sequencing depth, coverage and annotation.

The methodological biases, however, are not sufficient to explain the non-detection of EFF in every environment. In chapter three, we argue that there is a pattern for the prevalence of EFF under limiting conditions. Indeed, we have observed that the co-occurrence of limiting environmental conditions increases the relative abundance of EFF to the whole bacterial community in mineral springs.

We have shown that environmental conditions, and more specifically, multiple limiting factors, play a role in the prevalence of EFF in the environment. Whether these factors can also be a driving force for the evolution of sporulation in species that are not known to sporulate, is discussed in chapters four and five. *Serratia ureilytica*, a hospital-acquired pathogen, was isolated in a geothermal spring and found to produce spore-like resistant structures. Biochemical analysis showed similarities to EFF spores and genome analysis revealed a potential lateral gene transfer of crucial sporulation genes from EFF to this strain. This discovery is described in chapter four. In chapter five, the discovery of another species that is not known to sporulate, *Kurthia* sp. str. 11kri321 is described. Its ecology and distribution as well as the genomic imprints of sporulation are shown.

The second part of this thesis contains numerous examples of metabolic capabilities found in EFF that make them capable of tolerating adverse conditions and thriving in extreme environments.

Chapter six includes the discovery of a novel *Anoxybacillus* species, *Anoxybacillus geothermalis*, which was isolated from a geothermal reservoir. This strain was revived under laboratory conditions and is hypothesized to have remained inactive in the

reservoir since the Permian age. More isolates belonging to the same species were also discovered in different geothermal reservoirs.

Chapter seven focuses on other genomic imprints found in the genomes of EFF strains isolated from natural environments. These genomes have been sequenced for the first time and have been published as novel genome announcements.

Manganese oxidation and copper reduction of EFF in natural, uncontaminated environments, was studied and is presented in chapter eight. Our results lead to the conclusion that metal tolerance is a widespread phenomenon in unrelated aerobic EFF from natural uncontaminated environments.

Chapter nine discusses the limits of salt tolerance in aerobic and anaerobic EFF isolated from marine environments or salt lakes.

Finally, chapter ten is a general synthesis of this work and gives perspectives on future research. The ecological importance of sporulation under limiting conditions is also discussed. All conclusions drawn from these examples are summarized in chapter ten.

Chapter eleven summarizes all the collaborations that have been made in parallel to this thesis. Chapter twelve contains an article submitted to the Bulletin de la Société des Sciences Naturelles de Neuchâtel, addressed to the Swiss broad public interested in nature, and describes the microbial community structure of the natural mineral springs of Ponts-de-Martel, located in the Jura of Neuchatel, from an ecological perspective.

1.2. Background

1.2.1. What is Ecology?

Pascal Acot, in the introduction of his *Histoire de l'Ecologie* [1] in 1988, described that since the beginning of mankind, man has observed the relationships between plants and animals, and even tried to use these relationships to favor him, like in the case of fishing using a worm as bait. However, he states that the role of an ecologist is not simply to observe but also to study the relationships among organisms and their environment and to reveal patterns that drive ecosystems and formulate laws to describe these patterns. This study, along with the proposed laws that govern habitats, is defined as *Ecology*. This definition is not far from those proposed already since 1866 by Ernst Haeckel ("the comprehensive science of the complex relationship of organisms to the environment [...] described by Darwin as the necessary conditions for existence") [2], by Odum, in 1963 as the "study of the structure and function of nature" [3], and finally by Krebs, in 1972 as the "study of the interactions that determine the distribution and abundance of organisms" [4].

Mostly in the last two centuries, laws, principles and patterns about abundance, distribution and function of ecosystems have been proposed, at first for plants and then extrapolated or adapted to the animal kingdom. In 1970, Pierre Dansereau described ecological patterns and collected them in 27 laws [5], a selection of which are summarized herein. Any given environment cannot offer the optimum conditions for all of the functions for a given species (*law of the inoptimum*). However, a species can survive and grow in a habitat, depending on its limits of tolerance with regards to every

environmental factor (*law of tolerance*). Irrespective of the tolerance limits of an organism, stress will eventually occur and define the species' distribution (*law of climatic stress*). As a consequence, a species can spread in a geographical area, to a larger or lesser extent, due to variations in environmental factors (*law of valance*), that in their turn create overlapping ecological niches, which allow a gradual change in community composition and structure (*law of the continuum*). However, these gradients, community or environmental, cannot be considered as static: they steepen or smoothen at various times and places. Species with fitness potential cease the opportunity to prevail (*law of cornering*). Organisms share resources in a way that allows a greater portion to the most efficient (*law of competition and cooperation*). Species tend to survive in their habitat, even when the latter is altered (*law of persistence*). The habitat (i.e. the nature and population structure of the communities, the geography and the environmental factors) defines the ecological success of a species (*law of evolutionary opportunity*). Species abundance is controlled by the scarcest resource, rather than by the most abundant (*law of factorial control or law of the minimum*).

An unoccupied environment should be first inhabited by organisms with high tolerance and generally with low requirements (*law of ecesis*). The first settlers alter the environment, creating favorable conditions for other invaders, who, if fitter, may displace first settlers (*law of succession*). Succession should reach an equilibrium (*law of regional climax*). During climate change, elimination of some species or abundance reduction may occur through migration (*law of association segregation*). Migration is influenced not only by climate change, but also by population pressure (*law of migration*).

The local distribution of a plant depends on the long-range geographical distribution, since short- and long-scale environments are determined similarly (*law of geoecological distribution*). Drastic events put a selective pressure on organisms: therefore, it is more likely that differentiation occurs during these events than in other periods (*law of geological alternation)* [5].

Rather than reviewing all 27 laws, in this introduction, a selected few on abundance and diversity are brought into focus.

Liebig's law of the minimum [6] states that the abundance of species or biomass growth of an organism are controlled by the scarcest resource, rather than by the most copious. In more detail, for a specific organism in a given ecosystem, it should always be that a particular nutrient, or another environmental factor that influences growth, should determine the biomass or abundance of this organism in the ecosystem. Experimentally, the addition of this limited factor ought to increase the abundance of the species, provided that all the other requirements for growth are met. This law, however, overlooks biological alternation of environments, therefore of the provided nutrients, and is difficult to be applied in ecosystems, where more than one factor is scarce. It also focuses on the specific pair "organism-given environmental factor".

A fundamental ecological principle, Shelford's law of tolerance, relates the role of environmental factors to plant and animal niche differentiation [2]. This law postulates that "[an organism] is absent or found in minimal numbers only […] should a(n)

[environmental] condition vary outside the limits tolerated by the animal" [7]. Such conditions, or factors, can be 'operationally significant' for the distribution of an individual, but they have also been extended to explain the distribution of a population or even a community [8–10].

If abundance and diversity are taken into account, based on this ecological law, a general theoretical distribution model with maximum diversity and abundance towards the middle range of each environmental factor can be predicted. This model suggests that the distribution is unimodal. For some environmental factors, this is indeed the case: extremity acts the same in both directions, for example in the case of temperature or pH. For other environmental factors, however, the distribution is rather monotonic, as in the case of salinity. In both cases, therefore, a monotonic relationship can be presented as shown in Figure 1.1.

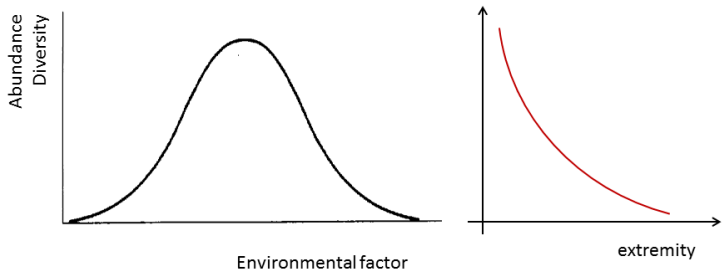


Figure 1.1. Theoretical curves representing abundance and diversity of species. At extreme values of a given environmental factor, both abundance and diversity are lower. This bell-shaped distribution curve can be summarized to a monotonal one, when conditions become in general adverse.

1.2.2. Microbial Ecology: from microscopes to metagenomes

The roots of all ecological laws described so far lie in the observations that ecologists made in nature based on plants and animals specimens, and then tested experimentally. The same principle of observation and experimental testing would not be possible for the microbial world, if it were not for the discovery of the light microscope. Indeed, Antonie van Leeuwenhoek, an amateur microscope maker in the 17th century, was the first to observe and describe in detail almost all the unicellular microorganisms that are known today: protozoa, algae, unicellular fungi, and bacteria [11]. These microscopic organisms were named *animalcules* [12], and soon a debate started concerning their origin. On the one hand, the theory of spontaneous generation was already well established. On the other hand, the idea that animalcules were transferred through air in boiled infusions, started being tested experimentally (Spallanzani). The debate lasted for more than a century until 1861, when Louis Pasteur performed a series of experiments to prove that firstly animalcules exist in air (by air-sampling and microscopic observation); secondly, that such an air-sample provokes microbial growth in sterile medium; and finally that a sterilized infusion remains sterile if placed in his renowned bent-neck flasks, thus not in contact with air. Once this neck breaks, the sterile medium is contaminated and

important animalcule growth is observed [12]. However, Pasteur did not win the battle until a decade later when Tyndall's experiments on spontaneous generation supported Pasteur's findings and led to the discovery of the sterilization method of discontinuous heating, known as *tyndallisation*. In parallel, the foundations of many biological processes of microorganisms were demonstrated, such as fermentation by yeasts and microbial anaerobic growth. In 1876, Robert Koch in his experiments on *Bacillus anthracis* showed that the causative agents of infectious diseases are bacteria. The same experiments were performed independently by Pasteur's group and supported Koch's findings.

A big step towards the establishment of Microbial Ecology was taken with the discoveries that microorganisms are geochemical agents and that they participate actively in geochemical cycles, as those of carbon, sulfur and nitrogen. This was mainly the work of M.W. Beijerinck and S. Winogradsky, who also developed the technique of enrichment cultures. This application is used until today, and for many decades almost all microbial ecology research has been based on this technique. Enrichment cultures for example depict natural selection at a microscale. Also, the successful creation of pure cultures has resulted in the isolation of many new species and their physiological description. In 1985, however, it was first demonstrated that microbes, which can be cultured, represent approximately 1% of the total diversity in soil [13]. This was called "the great plate anomaly" [14,15] and methodologies to overcome this issue started to emerge. The use of genes as indicators of biodiversity led to the direct detection of those genes in various environments. A milestone in the direct detection of genes was the description of the 16S rRNA gene as a phylogenetic marker by Carl Woese [13]. The 16S rRNA gene is a universal gene encoding the rRNA constitutive to the small subunit of the ribosome for bacteria. This gene contains both well-conserved and variable regions. Since the discovery of the 16S rRNA gene marker and its application to environmental samples, the number of potentially novel species has steadily increased. This phylogenetic marker, along with other markers such as universal genes (*rpoB*, *recA*) or group-specific genes (*nif*) has paved the way to a more detailed description of microbial communities in various environments.

For the last two decades, novel methodologies have emerged, resulting into the production of large gene sequence datasets and collection of associated environmental parameters in different habitats. This "era of metagenomics", as it is more commonly named [16,17] has been facilitated by an exponential technological advancement, in favor of microbiology.

Although technological advances have provided scientists with the tools to describe the microbial world and answer the question "who", still, the questions "why" and "how" need to be addressed.In 2007, an article published in *Nature* by *Prosser et al.*, explained the importance of ecological theory in microbial ecology, the need to propose hypotheses, laws and patterns based on observations then reject, confirm or discuss them based on experiment [18]. This article also highlights the problem of many contemporary microbiological research articles that often resemble more to the fisherman's than the ecologist's conclusions in Pascal Acot's example.

Whether plant and animal ecological laws are applicable to the microbial world is still to be determined. There is no doubt, however, that this question is an interesting challenge, since microbes set the limits beyond of what has been known for macro-organisms: limits of tolerance and distribution, as well as survival strategies, are already discovered concepts, but need to be mapped against ecological laws. This is no easy task; it is still very difficult to conclude whether theories that are already established in the macro-organismal world, can be extrapolated to microbes. The difficulties lie mostly in the fact that microbes have set new grounds on the species concept and have extended the limits of life, the rates of dispersal, and the laws of reproduction.

Bacteria and archaea can be differentiated from other organisms based on several fundamental biological processes. First, prokaryotic microorganisms are morphologically simple: bacterial or archaeal cells come only in a few shapes and forms. This morphological simplicity does not allow distinction between species, as is often the case for eukaryotic organisms. Second, a new generation in the prokaryotic world can be produced as quickly as within twenty minutes. This short reproduction time, along with differences between the bacterial/archaeal and the eukaryotic replication machinery, influences dramatically the mutation rate and, as a consequence, the evolutionary context of microbes. Third, in bacteria and archaea there is only asexual reproduction. In addition, horizontal gene transfer among closely related but also between distinct species is a phenomenon mostly observed in prokaryotes.

Besides differences in fundamental biological processes, there are differences in the ecology, i.e. their interactions with the environment. Firstly, bacteria and archaea have high dispersal rates; this observation led to the suggestion that there is no geographical divergence for these microbes: they can be distributed everywhere and the environmental factors alone would drive the selection of each habitat. This concept, that "everything is everywhere, but the environment selects" was first described by M.W. Beijerinck in the early 20th century and persists today. Second, their metabolisms are more diversified than those found in plants and animals. Anaerobic growth, anoxygenic photosynthesis, chemolithotrophy are only few examples of metabolisms restricted to bacteria and archaea. Third, microorganisms have expanded the limits of life, as we know them. They are able to survive temperatures as high as 120^{o}C, pH as low as 1 and as high as 13, high pressure and UV radiation. Survival to these conditions needs special adaptations rarely found in the macro-organismal world.

Microbes, thus, can colonize previously well-known habitats in potentially different patterns than those that drive the colonization by plants and animals.

1.2.3 Extreme habitats

Extreme is in the eyes of the beholder, stated Rothschild and Mancinelli about life in extreme environments [19]. Most of the time we tend to follow anthropocentric definitions of normality and extremity; this results in a mesophilic and extremophilic categorization of the (microbial) world. A number of physicochemical parameters, such as temperature (10 to 45°C), pH (5.5 to 8), atmospheric pressure (~1 atm), salinity (up to

0.3M salt content), humidity (water activity (a_w)> 0.8) and concentrations of elements or chemical compounds (minimum inhibitory concentrations), are used to describe mesophilic conditions. In that context, any conditions departing from such given norms are by default defined as extreme. Habitats that present these conditions are also characterized as extreme (for life). In the case of multiple extreme conditions co-occurrence in natural ecosystems, habitats are then defined as 'multiple extreme'.

For this dissertation, the extreme habitats of geothermal reservoirs, geothermal natural springs, brines, desserts and highly contaminated environments with trace metals are considered.

Geothermal sites are terrestrial and hydrothermal (oceanic) hot springs or underground reservoirs that consist of highly pressurized water heated up from the earth mantle, in which minerals from the earth crust are dissolved. These environments house a large variety of microorganisms. These microbes not only survive in the presence of CO_2 (exclusively), inorganic constituents and with other minimal requirements, but they also exhibit high diversity [20,21]. As conditions in geothermal sites differ, so does the microbial metabolic activity and diversity.

In a community study published in 2013 at the Lower Geyser Basin, Yellowstone, Schubotz *et al.* suggested that differences between communities in proximate but different sites occur because of variances in microbial metabolism and because of the exogenous carbon input into these systems [20]. In other words, different cellular functions serving various microbial needs occur in different environments. Indeed, 16S rRNA gene sequencing and metagenomic analysis showed abundant Crenarchaeota and Aquificales, Thermales and Thermotogales in all sites, however the community composition changed significantly when sampling progressively downstream of the natural spring's pool (decreasing temperature).

Environmental stress factors play an important role in the formation of community patterns. Distinct community types have been observed according to four environmental factors (pH, temperature, dissolved sulfide and sulfur, and physiographic site description) [22]. In addition, one single environmental factor is not enough to determine taxon segregation in a study of two alkaline hot springs that share temperature conditions but not water chemistry, as proposed by Weltzer and colleagues. They also introduced the idea that thermal gradients can create habitat gradients because of different temperature fitness of various taxa, resulting in differences in species richness and dissimilarity among sites [23]. Finally, studying the community composition and associated metabolism can also provide very informative insights. A study of Arctic and Antarctic microbial mat communities suggested that communities dominated by different phyla may vary in the response to stress at a genetic level [24]. In a later study, Takacs-Vesbach *et al.* discussed that different environments can favor different characteristics, even among members of the same family [25]. Indeed, evolutionary pressure in a given environment is what triggers these intra-genera and intra-family differentiations.

Saline environments are also considered extreme, since the concentrations of Na^+, K^+, and Cl^- exceed those of mesophilic environments and cells tend to dehydrate due to

osmosis. Such saline environments are often natural, for example salt lakes, seashore evaporations or estuaries, but they can also be artificial brines [26]. Microorganisms capable of surviving under these conditions of total salinity 10 times higher than seawater [26] require specific cellular adaptations to withstand dehydration. These adaptations differ in regards to whether these microorganisms are salt-requiring (halophiles) or salt-tolerant (halotolerant). Halophilic microorganisms withstand osmotic dehydration of their cytoplasm by accumulating potassium and chloride. High concentrations of potassium and chloride however can be toxic to proteins (denaturation) thus, extended proteomic modifications have been observed in these organisms. These proteins are now more acidic than those found in mesophilic organisms and denature in low salt concentrations [27]. For this reason, these organisms require salt for growth. Salt-tolerant organisms use a different strategy. When halotolerant bacteria are found in an environment where the external NaCl concentration is higher than the cytoplasm, water osmoses out of the cell, resulting in cell shrinkage. This decrease in cellular volume triggers the synthesis of compatible solutes and as a result, the cellular volume returns to the acceptable size for growth [26,27].

Environmental contamination by toxic metals, such as mercury, lead, copper, cadmium, zinc, manganese, nickel, cobalt, silver, gold, uranium, and thorium, can also have a significant impact on microbial populations. Metal contamination at toxic levels has a strong influence on biological processes, since metals are involved in practically every metabolic path, as ligands to enzymes, and consequently every index of microbial metabolic activity (respiration, methanogenesis and nitrogen fixation, among others) can be adversely affected by elevated concentrations of toxic metals [26,28–30]. Therefore, microorganisms that thrive in metal-contaminated environments have developed a variety of strategies for their survival including detoxifying mechanisms such as bioaccumulation, biotransformation, biomineralization or biosorption [31–35].

Another important environmental factor is pH. Enzymes (and proteins in general) are very sensitive to changes in pH; however cells are able to control the cytoplasmic pH in order to avoid enzyme degradation and consequent death. The range of pH that most macro-organisms can tolerate is from pH 5.5 to pH 8. However, acidic and alkaline environments exist, for example when soluble sulfur compounds are present or under high sodium carbonate concentrations, a pH as low as one and as high as eleven can be measured, respectively [26,36]. pH is an environmental factor that is influenced by other environmental parameters, such as temperature and salt and metal concentrations. Thus, thermophilic, halophilic or metal-contaminated environments are frequently multiply extreme, since pH is an associated factor.

1.2.4. Survival strategies in the microbial world

In order to avoid extinction during unfavorable conditions, microbial communities have developed diverse survival strategies. One of the most common is the formation of biofilms. By using minerals as a substrate, bacteria of different species can form biofilms in which they survive and multiply. They are able to exchange plasmids, and genes that

code for traits that allow them to adapt and evolve. Biofilms consist of "bacterial neighborhoods" in which the community deals with the problem of limited nutrients in the "economical" way of using common resources. This strategy is considered to be an altruistic behavior among bacterial individuals in a microbial community, contributing to the survival, fitness and evolution of species [37].

Apart from community survival strategies, bacteria often develop individual characteristics in times of adversity. They can develop morphological plasticity [38], they can switch to low catabolic activity or to a metabolic stand-by, or exhibit post-transcriptional modifications to optimize their fitness [39].

Dormancy is a well-studied survival strategy, which simultaneously fulfills both categories described above, community and individual survival. It is defined as "a state whereby metabolism and normal progression of life activities and development are dramatically reduced or brought to a halt" [40]. It is an individual characteristic, since not all bacteria possess genes that allow them to switch from a vegetative to a dormant state. However, dormancy has also been described as a community strategy: many hypotheses have been advanced, suggesting that there is an "altruistic behavior" among dormant cells before, during and after entering the state of dormancy [41–43].

1.3 Sporulation as a survival strategy in adverse conditions

1.3.1. Sporulation

Sporulation is a dormancy state. It is described as a condition of low metabolic activity and shrinking in size (in most cases) [44]. Four taxa are known to produce spore or spore-like forms, structures that are more resistant to environmental stress than the equivalent vegetative forms. Spores are found within Firmicutes, Actinobacteria, Cyanobacteria, and in a genus of Proteobacteria, *Myxococcus* spp. Among these phyla, not all members can enter sporulation. A series of genes are necessary for this process. Absence or malfunction of one or more genes results in the failure of sporulation. Indeed, there are bacteria, which are not known as endospore-formers, but that do encode some (but not all) sporulation genes in their genome. These bacteria cannot produce spores and for that reason they are called *asporogenic*. Sporulating bacteria undergo an intricate sequence of cell differentiation events leading to the formation of spores. These developmental processes can be used as a model of the evolution of structural and cellular functions. The spores, or spore-like forms, that are produced by bacteria are significantly different between taxa, in terms of structure, sporulation procedure and resistance to limiting environmental conditions. Although the sporulation procedure differs significantly among these taxa, it can be described as a general 4-step process. First, cells detect, in most cases, unfavorable conditions in their environment. Second, commitment to sporulation follows, as an irreversible step. Third, a dormant structure, mostly called a *spore*, is then produced either as a result of a special cell division or a modification of the vegetative cell structure. The spore structure is lighter, denser and more resistant than the vegetative mother cell they derived from, and less metabolically active. In all four taxa, this structure does not allow growth and replication, it does, however, permit

survival and dispersal. Finally, a mature spore is produced, ready to germinate once conditions are favorable again.

1.3.2. Sporulation in endospore-forming Firmicutes

Sporulation in EFF can be schematically described in five steps [44]. The initiation of sporulation takes place when the bacterial cell senses adverse conditions. Before changing morphologically, the vegetative cell programs the initiation of sporulation by activating the early sporulation genes. These are the 0/I stages. At stage II, the vegetative cell enters a particular cell division state, during which it creates a septum that is located at the one side of the cell and not in the middle, as in normal cell division. This special cell division and subsequent engulfing results in the production of a *forespore* inside the original vegetative cell, during stage III. However, before entering stage IV, this forespore transforms into a protoplast that has a double membrane layer. Between the two membranes the spore wall and cortex is formed. The spore coat is also produced to surround the forespore. These structures, which are not present in a vegetative cell, are all formed during stage IV. During stage V, the vegetative cell is lysed and the endospore is released [44].

The spore itself can alter its size, depending on humidity, without exiting dormancy [45]. During this dynamic state of low or no metabolic activity, the spore can survive for a very long time and once conditions are favorable again, it initiates germination to return to the vegetative state. Indeed, there are many publications claiming revival of endospore-forming bacteria, which have been in a dormant state for millions of years. An impressive example of such revival is the isolation of a 250 million years old *Bacillus* from a salt crystal [46].

Not all bacteria can enter sporulation. A series of genes are necessary for this process [47]. All endospore-forming bacteria, known so far, belong to the phylum Firmicutes (Gram-positive, low G+C content bacteria), although not all Firmicutes form these structures [48].

1.3.3. Sporulation in Cyanobacteria

Cyanobacteria are a very diverse phylum, in terms of metabolism and morphology. Some filamentous cyanobacteria are able to produce differentiated cell types that are capable of nitrogen fixation, called *heterocysts* [49]. These filamentous forms of cyanobacteria also form a resting cell, called an *akinete*. These forms are produced after light or phosphate deprivation, but also at low temperatures, low K^+ availability, and when their C:N ratio is fluctuating [40]. Akinetes are often larger than the vegetative form, and have a certain metabolic activity, although lower than the vegetative cell or the heterocyst [40]. Through the process of akinete maturation, akinetes are subjected to different metabolic stages. Young akinetes are able to fix CO_2, however, metabolic rates decrease progressively, and mature akinetes do not possess any functional photosystem and lack chlorophyll. A few reaction centers do exist, however, in the mature akinete, so that photosynthesis can be activated quickly during germination. Akinetes contain double the

quantity of DNA than their mother vegetative cell, and up to a ten-fold of the protein content [40].

Not all Cyanobacteria produce akinetes. Some genera are able to produce other spore-like structures, such as exospores, baeocytes, and hormocytes [50].

1.3.4. Sporulation in Actinomycetes

Sporulation in Actinomycetes is a complex procedure, tightly related to germination, cell division and colony growth. Among Actinomycetes, the genus *Streptomyces* is a model for the description of the cell germination and sporulation cycle. Firstly, a *Streptomyces* spore germinates, producing one or more hyphae, which branch to form a vegetative mycelium [51]. That mycelium becomes more complex during exponential growth; apical growth and branching occur during this step. Upon nutrient limitation, sporulation-programmed hyphae are formed and multiple cell divisions take place [52]. Septal peptidoglycan synthesis takes place, and peptidoglycan is also formed between the spores. The final step includes lysis of the peptidoglycan in between the spores and release of the mature spore [11].

A family of activators, the SALP proteins (SsgA-like proteins), controls sporulation in *Streptomyces*. They are exclusively found in Actinomycetes. The SsgA protein accumulates in the cell during mycelium growth, and upon reaching a crucial concentration it up-regulates a series of genes involved in aerial growth and spore formation resulting in the activation of sporulation-specific cell division. SsgB is responsible for the cessation of the aerial tip growth. SsgD regulates the formation of the peptidoglycan cell wall, while SsgG ensures that all sporulation septa are formed simultaneously. Finally, SsgE and SsgF are responsible for the correct autolysis of the peptidoglycan between spores during the maturation step [53].

1.3.5. Sporulation in *Myxococcus*

Sporulation in the genus *Myxococcus* is mostly studied in *Myxococcus xanthus*. Two different sporulation pathways are described; the first is a response to nutrient deprivation and the second one is induced by an increase of glycerol concentration in the microenvironment of the bacterium. The starvation-induced sporulation takes place in nature, while the chemically (glycerol and other organic chemical compounds) induced pathway has been observed in the laboratory [54].

In the first sporulation pathway, upon detection of limiting nutrients in the environment, a re-modeling of the rod-shaped cell, rather than a cell division, takes place resulting in a spherical structure [55]. The whole developmental program from a vegetative cell to a resting spore lasts approximately 72 hours and is a complex procedure. Unlike endospore-formation in Firmicutes that occurs individually, sporulation in *Myxococcus* is a multicellular process that involves intra-cellular communication and sacrifice of the majority of the cells for the survival of the minority [56]. Upon nutrient deprivation, an average of 10^5 cells aggregate into a mound, called a *fruiting body*. At the completion of aggregation, spore differentiation is induced in the mound. Most of the cells

(approximately 80%) lyse. The peripheral cells maintain their rod shape and do not sporulate. A series of inter- and intra-cellular signals induces sporulation in the rest of the cells [57]. During this last step, the bacterial genome is duplicated. In the second case, the glycerol-induced sporulation, fewer cells are involved and all of them differentiate into spores within eight hours. Spore coats are thinner and each spore contains multiple genome copies [58].

A series of proteins are specific to sporulation and some of them are homologous to proteins responsible for sporulation in Firmicutes, such as the CbgA and FdgA that are homologs to SpoVR, and the ActA and ActB proteins that are homologs with CsgA in Firmicutes. Finally, the *nfs* operon encodes a series of proteins responsible for the production of viable spores [59].

1.3.6. Costs and benefits of sporulation

Whether sporulation is a costly or beneficial survival strategy is considered based on the fitness of a sporulating species in a given environment. According to Levin's definition of fitness, it is "the extent to which an individual contributes its genes to future generations relative to other individuals in the same population; i.e. the individual's relative reproductive success" [60]. This is a general definition that applies to all organisms. When it comes to bacteria, however, it is not only the number of offspring (generation time and growth rate) that is taken into account but also the resistance and survival under given conditions [61].

As previously presented, sporulation in any of the above-mentioned taxa, is a complex developmental procedure that is controlled by a series of factors unique to spore-formation. These extra genes that encode for these factors are an addition to the bacterial chromosome. To maintain a larger chromosome is energetically costly. Thus, under constant optimal conditions, spore-forming bacteria tend to lose their sporulation genes, and as a consequence, their sporulation capability, in order to maximize fitness. When environmental conditions are adverse, in order to enter sporulation, the bacterial cell needs to decide whether entering sporulation would be beneficial or not. During sporulation, extra proteins are synthetized to produce a specialized structure. This procedure is very energetically demanding and commitment to sporulation is therefore a decision that the cell takes as a last resort. Under catastrophic conditions, sporulation has a clear survival advantage, as it may be the only way to avoid cellular death.

1.4. Survival by adaptation

In order to withstand adverse conditions, microbes alter their physiology to adapt accordingly, and consequently thrive in extreme habitats.

In hotspots, such as geothermal reservoirs, physiological processes are generally less efficient than in mesophilic conditions. At high temperatures, microbes face an irreversible breakdown of their biomolecules and a disruptive high fluidity of their plasma membrane [62]. Consequently, thermophilic and hyperthermophilic organisms evolved thermostable proteins and enzymes and their cell membranes have a different

composition. The stability of their proteins is guaranteed through extra chemical bonds (S-S bridges, H-bonds, metal bindings) [63], by hydrophobic amino acids, by producing multiple small subunits for the formation of an enzyme, and by the presence of chaperone proteins. Stability of proteins by extra chemical bonds results in rigid enzymes that seem to be less productive catalysts than mesophilic enzymes. Thermophiles seem to use thermal energy to overcome this reduction. A second adaptation is the accumulation of chemicals, such as amines and polyamines, which increase the stability of NADH, ATP and amino acids [50]. A final adaptation concerns the integrity of the plasma membrane, as at high temperatures, the fluidity of the plasma membrane increases. The stability of the membrane is guaranteed through an increase in the length of carbon chains and an extension of the branching of the phospholipids, an accumulation of saturated phospholipids, and through changes in the heads of the phospholipids [50, 64].

At low temperatures, microorganisms face the low availability of liquid water, and the damages to cellular integrity caused by the formation of ice crystals. The cellular response to cold is a change of the membrane composition; as expected, the opposite modifications of those taking place in thermophiles are made: unsaturated and shorter fatty acids are incorporated into the cellular membrane, while branching between lipids is also limited [65]. These fatty acids make the membrane more fluid because they introduce gaps that push apart the components of the membrane. At the interior of the cell, they contain antifreeze agents (mainly sugars) and small acidic proteins, which prevent ice formation [62]. Moreover, their proteins tend to have more α-helices than β-folded sheets, to allow flexibility [63]. Finally, psychrophiles are found to have a unique cold-stable translational system [50].

In highly acidic or highly alkaline habitats, microbes enable processes for gaining energy and carrying out chemical reactions inside the cell in order to regulate the intracellular pH to neutral. This is mainly accomplished through pumping protons out or into the cell, respectively. In an acidic environment, a K^+/H^+ antiporter is pumping K^+ in and H^+ out to make the cytoplasm alkaline. In a basic environment, a Na^+/H^+ antiporter is pumping H^+ in and Na^+ out to produce the opposite effect [50].

In highly saline as well as dry habitats, microorganisms struggle with high osmotic pressure and low water availability. Moreover, the membrane integrity is threatened by disruption due to dryness or salinity. In both saline and dry habitats, microbes control the water loss from the cell by producing compatible solutes [66]. These molecules are polar, water-soluble and are capable of stabilizing proteins [50]. Examples of organic compatible solutes in bacteria are glycine betaine, ectoine, and trehalose. Osmoprotection can be accomplished by accumulation of potassium chloride. Cl^- and K^+ are transported separately into the cytoplasm. KCl is formed to counterbalance the high concentration of NaCl that is found outside the membrane [67].

Osmotic pressure is only one challenge that prokaryotes face. There is also high atmospheric pressure, under which the tight packing of molecules and the loss of fluidity of the cellular membrane result in impaired cellular functions. Microorganisms have developed mechanisms of alternative gene expression in order to produce molecules

that enhance the uptake of nutrients and a change in the membrane structure, by incorporating unsaturated fatty acids to guarantee fluidity.

Microorganisms are in contact with light in most habitats, and for phototrophs light is their energy source. However, ultraviolet (UV) light and ionizing radiation are extremely harmful because the cells only perform a limited repair of damaged DNA. Bacteria can survive only low exposure to this radiation since they can only repair limited DNA damage. Some extremophilic bacteria can survive very high levels of radiation (thousand times higher than other cells) due to advanced DNA repair mechanisms, and due to their specialized cell membrane and cell wall. Moreover, these organisms often have multiple copies of their genomes at stationary phase. However, this resistance is highly costly: DNA damage repair and the replication of a 9-fold genome both demand high amounts of energy [50].

Finally, the chemical composition of the environment plays an important role in microbial survival. Low nutrient availability or presence of harmful compounds result in a fatal reduction of biomolecule synthesis and enzyme productivity. Slow metabolic activity and production of extracellular -often polysaccharides- mucus or an impermeable cell wall, in the case of harmful chemicals, are the main adaptations for microorganisms under such conditions [62].

Needless to say that impaired cellular functions under the above-mentioned extreme conditions are confronted by entering dormancy, as well. At the edges of extremity, this could be the last resort solution.

1.4.1 Adaptation under extreme conditions at a molecular level

From a genetic point of view, the above-mentioned adaptations to extreme environments are imprinted as molecular modifications; they are either genomic imprints, or post-transcriptional and post-translational modifications. The metabolic adaptations that are imprinted in the genomes of microbes and the modifications at an intracellular level are summarized herein.

In thermophilic and hyperthermophilic bacteria and archaea a series of modifications at the genomic level are necessary for the cells to withstand high temperatures. Firstly, changes at the amino acid level are observed; proteins of thermophiles contain mostly alanine, threonine, arginine and glutamic acid residues, while amino acids that enable flexibility of the protein are rare [68]. Moreover, the overall codon usage and nucleotide content, especially concerning rRNA and tRNA, vary significantly between mesophiles and thermophiles [69]. Secondly, multiple chaperone genes are found in the genomes of thermophiles [50, 70]. Thirdly, genes that can be used as thermophily-specific biomarkers have been defined and are related to the supercoiling of the circular DNA [71], DNA repair and transcription regulation [70]. Genes that encode for proteins related to metal detoxification have also been identified [50, 70]. Finally, differences at the level of gene expression are also observed between mesophiles and thermophiles, especially concerning genes for amino acid synthetases and the regulation of these genes [72].

To date, two major genomic imprints are known for psychrophiles, related to low temperature adaptation. On one hand, there are the general genomic characteristics, such as high GC content in specific genes that encode for RNAs, elongation factors, and RNA polymerases [73] and the specific amino acid composition: hydrophobic and charged residues are characteristics of psychrophilic proteins [74]. On the other hand, genes that encode for small acidic proteins (cold acclimation proteins) are psychrophily-specific markers [75].

In acidophilic, acid-tolerant, alkalophilic and alkalo-tolerant prokaryotes, a series of universal genes are differentially expressed compared to mesophilic (neutrophilic) bacteria [50]. These genes encode for proteins that are mainly transporters and antiporters, but they also encode for enzymes, such as the glutamine decarboxylase (in acidophiles) and the cyclodextrin glycosyltransferase (in the case of alkalophiles) [50].

Halophiles may have two types of genetic halophilic markers in their genomes. On one hand, characteristic acidic proteins are the typical markers for halophilic bacteria, although this is not the case for all halophilic species [76]. On the other hand, there are genes that are related to the uptake of osmoprotectants. An example of such genes is the *kdp* operon that encodes and regulates a kinase, as part of a two-component system of signal transduction. Another example is the transporters or symporters responsible for the uptake of compatible solutes. The transcription of all these genes is regulated by specific proteins that, themselves, are osmoregulated [50].

In prokaryotes that tolerate high ultraviolet or ionizing radiation exposure, typical repair systems are in place. Genes that encode for specific enzymes, which are related to the removal of thymine dimers, the SOS repair, or photoactivation, are some of the genomic markers of UV resistance. Genes that are involved in these processes are *uvrA, uvrB* and *uvrC* that encode for proteins that have excision activities. Additionally, there is the gene that encodes for the photolyase enzyme. Finally, a high level of RecA, which binds to the thymine dimer and influences the polymerase to proceed without stopping, is also an example for the T-T dimer excision. LexA is a typical marker for the SOS repair mechanism. Finally, superoxide reductases, responsible for the removal of oxygen radicals, are highly active in UV/ionizing radiation-resistant cells [50].

Metal tolerance is often related to EPS production or to reduction or oxidation of metals. Moreover, there are transport-related mechanisms of resistance [62]. All these processes are controlled by reductases, oxidases or transporters.

1.5 Research Objectives

The general purpose of this research was to understand the role of sporulation, not only under limiting laboratory conditions but also, more importantly, in natural extreme environments. More specifically the research objectives of this study were:

- To understand whether the under-representation of EFF in metagenomic datasets is due to detection, annotation or other methodological issue.
- To reveal a pattern for prevalence of EFF in extreme environments.
- To understand the sporulation pathway and reveal the lateral gene transfer of sporulation genes in *Serratia ureilytica* str. Lr5/4.
- To describe the sporulation pathway in *Kurthia* sp. str. 11kri321, a previously asporogenic species.
- To describe a novel species, *Anoxybacillus geothermalis*, and to understand its niche specialization to geothermal reservoirs.
- To provide full genomes of 15 spore-forming species.
- To find genes of extremity in genomes of strains isolated in extreme environments, using comparative genomics.
- To correlate oxidation of manganese, copper resistance and salinity tolerance in EFF isolates from natural ecosystems.

1.6 References

[1] Acot P. Histoire de l'écologie. Paris: Presses Universitaires de France - PUF; 1988.

[2] Odum EP, Odum HT. Fundamentals of ecology. Saunders; 1959.

[3] Odum EP. Ecology. [New York]: [Holt, Rinehart and Winston]; 1963.

[4] Krebs CJ. Ecology: The experimental analysis of distribution and abundance. New York: Harper and Row; 1972.

[5] Dansereau P. Ecological Impact and Human Ecology. Future Environ. N. Am. The Natural History Press, division of Doubleday & Company, New York: F. Fraser Darling and John P. Milton; 1966.

[6] Danger M, Daufresne T, Lucas F, Pissard S, Lacroix G. Does Liebig's law of the minimum scale up from species to communities? Oikos 2008; 117: 1741–51. doi:10.1111/j.1600-0706.2008.16793.x.

[7] Shelford VE. Some Concepts of Bioecology. Ecology 1931; 12: 455–67. doi: 10.2307/1928991.

[8] Pörtner HO. Climate variations and the physiological basis of temperature dependent biogeography: systemic to molecular hierarchy of thermal tolerance in animals. Comp Biochem Physiol A Mol Integr Physiol 2002; 132: 739–61.

[9] Calow P, Sibly RM. A Physiological Basis of Population Processes: Ecotoxicological Implications. Funct Ecol 1990; 4: 283–8. doi: 10.2307/2389587.

[10] Woodward FI. Climate and Plant Distribution. Cambridge University Press; 1987.

[11] Madigan M, Martinko J, Stahl D, Clark D. Brock Biology of Microorganisms (13th Edition). Benjamin Cummings; 2010.

[12] Stanier E' 'R Y. General Microbiology. 5rev ed Edition. Palgrave Macmillan; 1987.

[13] Woese CR. Bacterial evolution. Microbiol Rev 1987; 51:221–71.

[14] J T Staley, Konopka and A. Measurement of *in situ* Activities of Nonphotosynthetic Microorganisms in Aquatic and Terrestrial Habitats. Annu Rev Microbiol 1985; 39: 321–46. doi:10.1146/annurev.mi.39.100185.001541.

[15] Ward DM. A natural species concept for prokaryotes. Curr Opin Microbiol 1998; 1: 271–7.

[16] Oremland RS, Capone DG, Stolz JF, Fuhrman J. Whither or wither geomicrobiology in the era of "community metagenomics." Nat Rev Microbiol 2005; 3: 572–8. doi: 10.1038/nrmicro1182.

[17] Handelsman J. Metagenomics: Application of Genomics to Uncultured Microorganisms. Microbiol Mol Biol Rev 2004; 68: 669–85. doi:10.1128/MMBR.68.4.669-685.2004.

[18] Prosser JI, Bohannan BJM, Curtis TP, Ellis RJ, Firestone MK, Freckleton RP, et al. The role of ecological theory in microbial ecology. Nat Rev Microbiol 2007; 5: 384–92. doi: 10.1038/nrmicro1643.

[19] Rothschild LJ, Mancinelli RL. Life in extreme environments. Nature 2001; 409: 1092–101. doi: 10.1038/35059215.

[20] Schubotz F, Meyer-Dombard DR, Bradley AS, Fredricks HF, Hinrichs K-U, Shock EL, et al. Spatial and temporal variability of biomarkers and microbial diversity reveal

metabolic and community flexibility in Streamer Biofilm Communities in the Lower Geyser Basin, Yellowstone National Park. Geobiology 2013; 11: 549–69. doi:10.1111/gbi.12051.

[21] Inskeep WP, Herrgard MJ, Jay ZJ, Rusch DB, Tringe SG, Parenteau MN, et al. Community structure and function of high-temperature chlorophototrophic microbial mats inhabiting diverse geothermal environments. Front Microb Physiol Metab 2013; 4: 106. doi:10.3389/fmicb.2013.00106.

[22] Inskeep WP, Jay ZJ, Herrgard MJ, Kozubal MA, Rusch DB, Tringe SG, et al. Phylogenetic and functional analysis of metagenome sequence from high-temperature archaeal habitats demonstrate linkages between metabolic potential and geochemistry. Front Microb Physiol Metab 2013;4:95. doi:10.3389/fmicb.2013.00095.

[23] Weltzer ML, Miller SR. Division-specific differences in bacterial community assembly of two alkaline hot spring ecosystems from Yellowstone National Park. Microb Ecol 2013; 65: 537–40. doi: 10.1007/s00248-013-0214-3.

[24] Varin T, Lovejoy C, Jungblut AD, Vincent WF, Corbeil J. Metagenomic analysis of stress genes in microbial mat communities from Antarctica and the High Arctic. Appl Environ Microbiol 2012; 78: 549–59. doi:10.1128/AEM.06354-11.

[25] Takacs-Vesbach C, Inskeep WP, Jay ZJ, Herrgard MJ, Rusch DB, Tringe SG, et al. Metagenome sequence analysis of filamentous microbial communities obtained from geochemically distinct geothermal channels reveals specialization of three Aquificales lineages. Front Microbiol 2013; 4: 84–84. doi:10.3389/fmicb.2013.00084.

[26] Edwards C. Microbiology of Extreme Environments. New York: Mcgraw-Hill; 1990.

[27] Oren A. Microbial life at high salt concentrations: phylogenetic and metabolic diversity. Saline Syst 2008; 4: 2. doi: 10.1186/1746-1448-4-2.

[28] Chaudri AM, Allain CMG, Barbosa-Jefferson VL, Nicholson FA, Chambers BJ, McGrath SP. A study of the impacts of Zn and Cu on two rhizobial species in soils of a long-term field experiment. Plant Soil 2000; 221: 167–79. doi: 10.1023/A:1004735705492.

[29] Pereira SIA, Lima AIG, Figueira EM de AP. Heavy metal toxicity in *Rhizobium leguminosarum* biovar *viciae* isolated from soils subjected to different sources of heavy-metal contamination: Effects on protein expression. Appl Soil Ecol 2006; 33: 286–93. doi:10.1016/j.apsoil.2005.10.002.

[30] Robinson T, McMullan G, Marchant R, Nigam P. Remediation of dyes in textile effluent: a critical review on current treatment technologies with a proposed alternative. Bioresour Technol 2001; 77: 247–55. doi: 10.1016/S0960-8524(00)00080-8.

[31] Ghosh A, Saha PD. Optimization of copper bioremediation by *Stenotrophomonas maltophilia* PD2. J Environ Chem Eng 2013; 1: 159–63. doi:10.1016/j.jece.2013.04.012.

[32] Pradhan AA, Levine AD. Microbial biosorption of copper and lead from aqueous systems. Sci Total Environ 1995; 170: 209–20. doi: 10.1016/0048-9697(95)04709-7.

[33] Wang J, Chen C. Biosorbents for heavy metals removal and their future. Biotechnol Adv 2009; 27: 195–226. doi:10.1016/j.biotechadv.2008.11.002.

[34] Zhou W, Zhang H, Ma Y, Zhou J, Zhang Y. Bio-removal of cadmium by growing deep-sea bacterium *Pseudoalteromonas* sp. SCSE709-6. Extremophiles 2013; 17: 723–31. doi: 10.1007/s00792-013-0554-4.

[35] Zouboulis A., Loukidou M., Matis K. Biosorption of toxic metals from aqueous solutions by bacteria strains isolated from metal-polluted soils. Process Biochem 2004;39:909–16. doi:10.1016/S0032-9592(03)00200-0.

[36] Sorokin DY, Berben T, Melton ED, Overmars L, Vavourakis CD, Muyzer G. Microbial diversity and biogeochemical cycling in soda lakes. Extrem Life Extreme Cond 2014;18:791–809. doi:10.1007/s00792-014-0670-9.

[37] Kreft JU. Biofilms promote altruism. Microbiol-Sgm 2004;150:2751–60. doi:10.1099/mic.0.26829-0.

[38] Justice SS, Hunstad DA, Cegelski L, Hultgren SJ. Morphological plasticity as a bacterial survival strategy. Nat Rev Microbiol 2008;6:162–8. doi:10.1038/nrmicro1820.

[39] Zhao H, Msadek T, Zapf J, Madhusudan, Hoch JA, Varughese KI. DNA complexed structure of the key transcription factor initiating development in sporulating bacteria. Structure 2002;10:1041–50. doi:10.1016/S0969-2126(02)00803-1.

[40] Lubzens E, Cerda J, Clark M. Dormancy and Resistance in Harsh Environments. Springer Science & Business Media; 2010.

[41] Lennon JT, Jones SE. Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat Rev Microbiol 2011;9:119–30. doi:10.1038/nrmicro2504.

[42] Buerger S, Spoering A, Gavrish E, Leslin C, Ling L, Epstein SS. Microbial scout hypothesis, stochastic exit from dormancy, and the nature of slow growers. Appl Environ Microbiol 2012;78:3221–8. doi:10.1128/AEM.07307-11.

[43] Wolf DM, Vazirani VV, Arkin AP. Diversity in times of adversity: probabilistic strategies in microbial survival games. J Theor Biol 2005;234:227–53. doi:10.1016/j.jtbi.2004.11.020.

[44] Driks A. Overview: Development in bacteria: spore formation in *Bacillus subtilis*. Cell Mol Life Sci CMLS 2002;59:389–91.

[45] Driks A. The dynamic spore. Proc Natl Acad Sci U S A 2003;100:3007–9. doi:10.1073/pnas.0730807100.

[46] Vreeland RH, Rosenzweig WD, Powers DW. Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal. Nature 2000;407:897–900. doi:10.1038/35038060.

[47] Galperin MY. Genome Diversity of Spore-Forming Firmicutes. Microbiol Spectr 2013;1:TBS – 0015–2012. doi:10.1128/microbiolspectrum.TBS-0015-2012.

[48] Onyenwoke RU, Brill JA, Farahi K, Wiegel J. Sporulation genes in members of the low G+C Gram-type-positive phylogenetic branch (Firmicutes). Arch Microbiol 2004;182:182–92. doi:10.1007/s00203-004-0696-y.

[49] Stal LJ. Physiological ecology of cyanobacteria in microbial mats and other communities. New Phytol 1995;131:1–32. doi:10.1111/j.1469-8137.1995.tb03051.x.

[50] Barton L. Structural and Functional Relationships in Prokaryotes. Springer Science & Business Media; 2005.

[51] Shepherd MD, Kharel MK, Bosserman MA, Rohr J. Laboratory Maintenance of Streptomyces species. Curr Protoc Microbiol 2010;CHAPTER:Unit – 10E.1. doi:10.1002/9780471729259.mc10e01s18.

[52] Brun Y, Shimkets LJ. Prokaryotic Development. ASM Press; 2000.

[53] Traag BA, van Wezel GP. The SsgA-like proteins in Actinomycetes: small proteins up to a big task. Antonie Van Leeuwenhoek 2008;94:85–97. doi:10.1007/s10482-008-9225-3.

[54] Komano T, Inouye S, Inouye M. Patterns of protein production in *Myxococcus xanthus* during spore formation induced by glycerol, dimethyl sulfoxide, and phenethyl alcohol. J Bacteriol 1980;144:1076–82.

[55] O'Connor KA, Zusman DR. Induction of beta-lactamase influences the course of development in *Myxococcus xanthus*. J Bacteriol 1999;181:6319–31.

[56] Jelsbak L, Søgaard-Andersen L. Cell behavior and cell-cell communication during fruiting body morphogenesis in *Myxococcus xanthus*. J Microbiol Methods 2003;55:829–39.

[57] Kaiser D. Signaling in myxobacteria. Annu Rev Microbiol 2004;58:75–98. doi:10.1146/annurev.micro.58.030603.123620.

[58] Müller F-D, Treuner-Lange A, Heider J, Huntley SM, Higgs PI. Global transcriptome analysis of spore formation in *Myxococcus xanthus* reveals a locus necessary for cell differentiation. BMC Genomics 2010;11:264. doi:10.1186/1471-2164-11-264.

[59] Dahl JL, Tengra FK, Dutton D, Yan J, Andacht TM, Coyne L, et al. Identification of Major Sporulation Proteins of *Myxococcus xanthus* Using a Proteomic Approach. J Bacteriol 2007;189:3187–97. doi:10.1128/JB.01846-06.

[60] Levin SA, Carpenter SR, Godfray HCJ, Kinzig AP, Loreau M, Losos JB, et al. The Princeton Guide to Ecology. Princeton University Press; 2009.

[61] Pope CF, McHugh TD, Gillespie SH. Methods to determine fitness in bacteria. Methods Mol Biol Clifton NJ 2010;642:113–21. doi:10.1007/978-1-60327-279-7_9.

[62] Edwards C. Microbiology of extreme environments. McGraw-Hill; 1990.

[63] Madigan MT, Martinko JM, Bender KS, Buckley DH, Stahl DA. Brock Biology of Microorganisms, Global Edition: UEL. Pearson Education Limited; 2015.

[64] Stetter KO. Extremophiles and their adaptation to hot environments. FEBS Lett 1999;452:22–5. doi:10.1016/S0014-5793(99)00663-8.

[65] Cavicchioli R, Siddiqui KS, Andrews D, Sowers KR. Low-temperature extremophiles and their applications. Curr Opin Biotechnol 2002;13:253–61.

[66] Oren A. Microbial life at high salt concentrations: phylogenetic and metabolic diversity. Saline Syst 2008;4:2. doi:10.1186/1746-1448-4-2.

[67] Kivistö AT, Karp MT. Halophilic anaerobic fermentative bacteria. J Biotechnol 2011;152:114–24. doi:10.1016/j.jbiotec.2010.08.014.

[68] Puigbò P, Pasamontes A, Garcia-Vallve S. Gaining and losing the thermophilic adaptation in prokaryotes. Trends Genet TIG 2008;24:10–4. doi:10.1016/j.tig.2007.10.005.

[69] Singer GAC, Hickey DA. Thermophilic prokaryotes have characteristic patterns of codon usage, amino acid composition and nucleotide content. Gene 2003;317:39–47.

[70] Makarova KS, Wolf YI, Koonin EV. Potential genomic determinants of hyperthermophily. Trends Genet TIG 2003;19:172–6. doi:10.1016/S0168-9525(03)00047-7.

[71] Forterre P. A hot story from comparative genomics: reverse gyrase is the only hyperthermophile-specific protein. Trends Genet TIG 2002;18:236–7.

[72] Song H, Wang H, Gigot D, Dimova D, Sakanyan V, Glansdorff N, et al. Transcription regulation in thermophilic bacteria: high resolution contact probing of *Bacillus stearothermophilus* and *Thermotoga neapolitana* arginine repressor-operator interactions. J Mol Biol 2002;315:255–74. doi:10.1006/jmbi.2001.5236.

[73] De Maayer P, Anderson D, Cary C, Cowan DA. Some like it cold: understanding the survival strategies of psychrophiles. EMBO Rep 2014;15:508–17. doi:10.1002/embr.201338170.

[74] Ayala-del-Río HL, Chain PS, Grzymski JJ, Ponder MA, Ivanova N, Bergholz PW, et al. The Genome Sequence of *Psychrobacter arcticus* 273-4, a Psychroactive Siberian Permafrost Bacterium, Reveals Mechanisms for Adaptation to Low-Temperature Growth. Appl Environ Microbiol 2010;76:2304–12. doi:10.1128/AEM.02101-09.

[75] Leblanc L, Leboeuf C, Leroi F, Hartke A, Auffray Y. Comparison Between NaCl Tolerance Response and Acclimation to Cold Temperature in *Shewanella putrefaciens*. Curr Microbiol 2003;46:0157–62. doi:10.1007/s00284-002-3837-z.

[76] Bardavid RE, Oren A. The amino acid composition of proteins from anaerobic halophilic bacteria of the order Halanaerobiales. Extremophiles 2012;16:567–72. doi:10.1007/s00792-012-0455-y.

Chapter 2

Under-detection of endospore-forming Firmicutes in metagenomic data

Sevasti Filippidou[1], Thomas Junier[1,2], Tina Wunderlin[1§], Chien-Chi Lo[3], Po-E Li[3], Patrick S. Chain[3], Pilar Junier[1]

[1]Laboratory of Microbiology, Institute of Biology, University of Neuchatel, CH-2000, Neuchâtel, Switzerland

[2]Vital-IT group, Swiss Institute of Bioinformatics, CH-1015, Lausanne, Switzerland

[3]Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, 87545, USA

This chapter is adapted from the article: Filippidou S., Junier T., Wunderlin T., Lo C.-C., Li P.E., Chain P.S., Junier P. : Under-detection of endospore-forming Firmicutes in metagenomic data. Computational and Structural BiotechnologyJournal.2015 Apr; 13: 299-306. doi:10.1016/j.csbj.2015.04.002.

Any reference to the data or the context of this article should cite the above-mentioned publications.

Abstract

Microbial diversity studies based on metagenomic sequencing have greatly enhanced our knowledge of the microbial world. However, one caveat is the fact that not all microorganisms are equally well detected, questioning the universality of this approach. *Firmicutes* are known to be a dominant bacterial group. Several *Firmicutes* species are endospore formers and this property makes them hardy in potentially harsh conditions, and thus likely to be present in a wide variety of environments, even as residents and not functional players. While metagenomic libraries can be expected to contain endospore formers, endospores are known to be resilient to many traditional methods of DNA isolation and thus potentially undetectable. In this study we evaluated the representation of endospore-forming *Firmicutes* in 73 published metagenomic datasets using two molecular markers unique to this bacterial group (*spoOA* and *gpr*). Both markers were notably absent in well-known habitats of *Firmicutes* such as soil, with *spoOA* found only in three mammalian gut microbiomes. A tailored DNA extraction method resulted in the detection of a large diversity of endospore-formers in amplicon sequencing of the 16S rRNA and *spoOA* genes. However, shotgun classification was still poor with only a minor fraction of the community assigned to *Firmicutes*. Thus, removing a specific bias in a molecular workflow improves detection in amplicon sequencing, but it was insufficient to overcome the limitations for detecting endospore-forming *Firmicutes* in whole-genome metagenomics. In conclusion, this study highlights the importance of understanding the specific methodological biases that can contribute to improve the universality of metagenomic approaches.

Key words: endospores, *gpr,* metagenomics, profile analysis, *spoOA*

2. 1. Introduction

Metagenomic studies have emerged as promising methods for the collective study of microbial communities directly extracted from environmental samples [1–3]. These approaches have been successfully applied to a variety of environments and have helped to unveil new functional pathways and metabolic processes within the microbial world [4–8].

Biases, however, can occur at all the steps involved in a metagenomic workflow. They can be associated to the specific type of environment [9, 10], the DNA yields obtained [11], the DNA extraction method [12], the amplification (for example in amplicon sequencing), but also in the sequencing and the analysis of the sequences, as well. These limitations have been highlighted in the recent literature and result in problems such as low coverage of the less abundant taxa (the so-called "depth bias" for example in the detection of ribosomal genes [13]), low reproducibility of results [14] and underrepresentation of certain taxa, as discussed herein. In order to overcome these limitations, single-cell genomics or novel approaches in culture-dependent methodologies, such as culturomics [15, 16] which, in their turn, have their own limitations.

Even though methodological bias of metagenomic diversity surveys associated to particular types of environments such as soil has been demonstrated experimentally [9,10], the specific coverage of individual microbial groups within the community is still unknown. One example of a bacterial group that can be used to test coverage bias in metagenomic datasets is endospore-forming *Firmicutes*. Even though culturing of microorganisms is largely acknowledged to be biased, according to previous research based on culture collections as well as whole-genome sequencing, *Firmicutes* is the second-most abundant bacterial phylum [17]. Endospore formers live in a wide range of environments on Earth's surface and subsurface [18, 19]. The hardy outer cortex of endospores and the small acid-soluble proteins stabilizing their DNA [20–22], allow these bacteria to be distributed into every habitat on Earth [23]. However, a phylogenetic assessment of the microbial communities in four metagenomic datasets has revealed surprisingly few endospore formers [24]. This might appear surprising considering their ubiquity, but endospores are known to withstand many traditional methods of DNA isolation and are thus potentially undetectable in a sample. Recently, a DNA extraction method for the extraction of resistant structures such as endospores has been developed by our group [12]. This DNA extraction method was combined with amplicon sequencing of the gene coding for the master regulator for the initiation of sporulation (*spo0A* gene) to demonstrate an improved detection of endospore-forming Firmicutes in sediment samples [12]. Our group has developed further methods to separate endospores from vegetative cells, which has open the possibility to carry out genomic studies only focused on endospores [12,25]. These two studies demonstrate by amplicon sequencing that the diversity of endospore-forming *Firmicutes* is far from revealed. However, the effectiveness of the improved DNA extraction method for whole-genome metagenomic studies is unknown.

The aim of this study was to measure the level of detection of endospore formers in metagenomic studies carried out so far, and to evaluate the effect of an improved DNA extraction method on the detectability of this group. To do this, we initially searched for functional gene markers of endospore formation in metagenomic datasets using profiles. We then applied a modified DNA extraction method that is tailored to release DNA from resistant structures such as endospores [12] in a selected environmental sample. Amplicon sequencing of the 16S rRNA and *spoOA* genes were performed on the sample in order to assess the relative abundance and phylogenetic diversity of *Firmicutes*. This was complemented by shotgun sequencing and classification of the metagenome reads. Our results indicate that endospore-forming *Firmicutes* are overlooked in environmental diversity surveys using traditional whole metagenomic approaches.

2.2. Material and methods

2.2.1. Genome sequence retrieval

Complete and draft genome sequences of endospore-forming *Firmicutes* were downloaded from the Comprehensive Microbial Resource (CMR, 24.0 data release, cmr.jcvi.org) and Integrated Microbial Genomes (IMG, 3.0, img.jgi.doe.gov) websites. Protein and nucleotide sequences of spore-related genes were obtained by search for role category/function *sporulation and germination* (CMR) and *sporulating* (IMG). Additional information on all retrieved genomes was obtained from the GenBank database (www.ncbi.nlm.nih.gov/genome).

2.2.2. Detection of orthologous sporulation genes common to all endospore formers

Orthologous groups were delineated based on best reciprocal BLASTp hits [26]. BLASTp was used to align each sequence in the set against all sequences except those of the same species (thus avoiding paralogs). The best hit in each species was retained, and sequence pairs, that were each other's best match, were defined as best reciprocal hits (BRHs). Putative orthologous groups were defined using the algorithm used by OrthoDB [27].OrthoDB has data on Fungi, Metazoa, and Bacteria. An early version of the BRHCLUS program (unpublished at the time) was obtained from its author, Dr. Tegenfeldt (pers. comm.) and run according to the author's instructions. The program is now available from http://orthodb.org/. To our knowledge, its utility does not depend on the clade it is used for - OrthoDB uses the same clustering program for all data in its scope.

2.2.3. Profile construction and validation

The genomic sequences were filtered in such a way as to keep only one (randomly chosen) sequence per genus, thus reducing taxonomic sampling bias. Multiple alignments of SpooA and Gpr were produced with MAFFT [28]. Gribskov-style sequence profiles were constructed with EMBOSS's prophet program [29]. The profiles' score cutoffs were determined by searching with EMBOSS's prophet program against the original SpooA (resp. Gpr) sequence set as a positive control, and against shuffled versions of the same as negative set.

2.2.4. Metagenomic datasets retrieval

The metagenome datasets (supplementary Table 2.1) were downloaded from IMG, GOLD (genomesonline.org), or the metagenomes subset of the WGS section of EMBL (ebi.ac.uk/genomes/wgs.html). These datasets included all the metagenomic studies available at EMBL when the profile analysis was performed. Only sequences or contigs of > 800 bp, which are slightly shorter than the full-length sporulation genes, were kept for analysis.

2.2.5. Environmental sampling, DNA extraction and quantitative PCR

The sample was collected at Nea Apollonia (NAP) geothermal spring (N 40° 39,191 E 22° 56,707), Greece, in June 2011. Geothermal reservoir was reached through a 120m drilling pipe, used mostly for pumping 80° C water for bathing purposes. Biofilm from the pipe interior was collected and frozen within 2h of collection. Upon arrival at the laboratory, a tailored DNA extraction method previously described [12] was applied to the sample. More precisely, DNA was extracted using the FastDNA Spin Kit for Soil (MP Biomedicals, California), using a modified protocol in order to ensure that DNA was not only extracted from vegetative cells but also from spores and other cells difficult to lyse. These modifications were (a) a separation of the biomass from the soil, using a Na-hexa-meta-phosphate solution and (b) a sequential bead-beating step (three times) to ensure mechanical disruption of cells. In total, 10ug of high molecular RNA-free DNA was obtained.

Moreover, 16S rRNA gene and *spoOA* gene copy numbers were calculated using a quantitative PCR assay, as previously described [30].

2.2.6. Amplicon sequencing of the 16S rRNA and *spoOA* genes

In order to verify the presence and relative abundance of endospore formers, 454 pyrosequencing of a fragment of the 16S rRNA and *spoOA* genes was firstly applied to the sample NAP. Sequencing was done using the services of Eurofins MWG Operon (Ebersberg, Germany). For 16S rRNA amplicon sequencing, fragments of approximately 500 bp were retrieved using primers Eub8f (5'-AGAGTTTGATCCTGGCTCAG-3') and Eub519r (5'-GTATTACCGCGGCTGCTGG-3'), as previously described [31]. 16S rRNA gene raw sequence data was analyzed with QIIME [32], using the pipeline for de novo OTU picking. OTUs were identified using a threshold of 97% sequence similarity. The sequences were then clustered into putative OTUs with the pick_otus.py program from the QIIME package using the Uclust method [32]. The single sequence picked by the program as a representative of each OTU was used to build a phylogeny.

For the *spoOA* amplicon sequencing, a 602 bp sequence of the *spoOA* gene was amplified using the degenerated primer *spoOA*166f (5'-GATATHATYATGCCDCATYT-3') and *spoOA*748r (5'-GCNACCATHGCRATRAAYTC-3') [12]. 42151 sequences were received from the sample. Sequences were then filtered according to Phred [33] quality score (minimum of 30) and sequences of length shorter than 600 bp were removed. Remaining sequences were translated to their amino acid sequence; resulting full-length ORFs were then matched against the SpoOAprofile, in order to confirm that the primers actually amplified the *spoOA* sequences.

Phylogenies were constructed from Phylip-formatted alignments with PhyML [34], using default parameters. The trees were re-rooted, condensed according to protocol, and displayed with the Newick Utilities. Each branch represents a cluster of OTUs of > 97% sequence similarity. Identification of the closest relatives of the environmental sequences was done by protein BLAST [26] with the translated protein sequences using a reference database of 581 Spo0A protein sequences from the InterPro site [35].

All metagenomic sequences were submitted to GenBank. The 16S rRNA amplicon sequencing data can be retrieved under the BioProject ID PRJNA267761 and BioSample ID SAMN03198953 and the *spo0A* amplicon sequencing data under the BioProject ID PRJNA276803 and Biosample ID SAMN03392534.

2.2.7. Metagenomic sequencing

Once a high prevalence of putative endospore-formers was confirmed in the 16S rRNA pyrosequencing data (41% of total bacterial community), whole-metagenome sequencing of NAP was performed on a full plate of a GS FLX platform, followed by *de novo* assembly using the services of GATC- biotech (Konstanz, Germany). The whole metagenome dataset can be retrieved from GenBank under the BioProject ID PRJNA271123 and BioSample ID SAMN03273062.

2.2.8. Metagenome data annotation

Several tools were used to produce the read-based metagenomic analysis of NAP metagenome dataset. GOTTCHA [36] was run using BWA [37] against 4 databases consisting of Phylum, Genus, Species and Strain-level unique signatures. MetaPhlAn v1.7.7 [38] was run using BowTie2 [39] with default parameters against its clade-specific maker genes database. Kraken was run with its reduced taxonomic-specific 31-mer database (mini-database). BWA v0.7.4-r385 that was used as a stand-alone tool was run locally using BWA-backtrack algorithm to map reads against a custom database of bacterial, archaeal and viral complete genomes retrieved from NCBI RefSeq database [40]. The mapped reads were subsequently assigned organisms by mapping the GI numbers of aligned references to NCBI taxonomic ID and rolled up to higher ranks. mOTUs v1.0 [41] was run with the database composed of 10 universal marker genes and LMAT v1.2.1 [42] was run with the pre-computed reference search database (kML.18mer.16bit.reduced.db) with default parameters. Since BWA (standalone), Kraken and LMAT only reported read counts of taxonomies, the relative abundances were represented by the portion of total classified reads in these tools (Table 1). While each tool tries to identify similarities among the reads and the databases used, each tool is centered around a different algorithmic approach to solve this complex challenge, using either a unique search algorithm, a uniquely designed database, or both. The interpretation of the results from each tool should thus be taken within its own context. For example, mOTUs and MetaPhlAn use pre-selected marker genes to perform the analysis, however different marker genes are used and different methods are used to identify reads that are similar to these marker genes. Kraken and LMAT both use subsequences within reads (k-mers) and match k-mers observed within the reads with those observed within known reference genomes.

Meanwhile BWA is a read-mapping tool that we use against the RefSeq database to report matching reads.

Prediction tool	Top 5 Phyla		Frequency	Relative %
16S RNA gene amplicon pyrosequencing (QIIME))	1	*Firmicutes*	41.70	41.70%
	2	Proteobacteria	26.14	26.14%
	3	Bacteroidetes	10.55	10.55%
	4	Planctomycetes	5.35	5.35%
	5	Chlorobi	3.88	3.88%
Kraken (mini database)	1	Proteobacteria	16644	82.71%
	2	Actinobacteria	1744	8.67%
	3	*Firmicutes*	322	1.60%
	4	Bacteroidetes	298	1.48%
	5	Cyanobacteria	192	0.95%
MetaPhlAn	1	Proteobacteria	82.01061	82.01%
	2	Chloroflexi	9.24158	9.24%
	3	Actinobacteria	2.32449	2.32%
	4	Bacteroidetes	2.08071	2.08%
	5	Acidobacteria	1.54098	1.54%
BWA	1	Proteobacteria	452	75.21%
	2	*Firmicutes*	32	5.32%
	3	Thaumarchaeota	28	4.66%
	4	Actinobacteria	26	4.33%
	5	Bacteroidetes	17	2.83%
LMAT	1	Ascomycota	425	35.68%
	2	Cyanobacteria	385	32.33%
	3	Proteobacteria	190	15.95%
	4	Thaumarchaeota	145	12.17%
	5	Basidiomycota	20	1.68%

Table 2.1. Prevalence of Firmicutes in 16S rRNA gene amplicon sequencing and shotgun metagenomic sequencing applied to the NAP sample. Different prediction tools were used to establish the five most frequent Phyla in the samples. With the exception of the 16S rRNA gene amplicon sequencing, the relative percentage indicated corresponded to the fraction of the sequences that could be classified and not to the frequency of any of the groups for the total reads generated after sequencing.

2.3. Results and discussion

2.3.1. Selection of functional markers for endospore-formation

We recently identified functional marker genes involved in endospore formation in endospore-forming *Firmicutes* [12]. Bidirectional BLAST of the genes annotated as part of the cellular function of sporulation allowed to select six highly conserved orthologous genes as part of the endospore-forming *Firmicutes* proteome. Among those, Spo0A and Gpr, were selected for the construction of profiles based on their consistent phylogenetic reconstruction with the 16S rRNA gene phylogeny. These two genes represent significant stages of the endospore-formation process, namely the commitment to enter sporulation (Spo0A) and the proteolytic activity on acid-soluble spore proteins (SASPs) during germination (Gpr) [43]. In recent studies analyzing the minimal set of endospore-formation genes required by endospore-formers had indicated that *spo0A* is indeed one of the most conserved genes almost exclusively found among this bacterial group [44–46]. In the case of *gpr*, it has been shown that it belongs to a category of genes present in *Bacillus* and *Clostridium* without any known orthologue in Gram-negative Proteobacteria or Cyanobacteria [21].

2.3.2. Profile analysis of sporulation genes in metagenomes

Profiles of Spo0A and Gpr were constructed and compared to metagenomic datasets to find sequences of high similarity with *spo0A* and *gpr*. Profiles are models of conserved sequences built from an alignment and are more sensitive than BLAST or other pair-wise comparisons especially for protein searches [47]. The sequence profiles were generated based on 14 aligned sequences. They were validated on genomes of known endospore-forming and non-sporulating bacteria (Figure 1.1A). A single positive hit was found in the genome of each endospore-forming bacterium, while no hits were found in the negative controls. This result also allowed determining a score cut-off for Spo0A (2000) and Gpr (2500) profiles to distinguish between positive and negative hits. Using this cut-off value one orthologous sequence of each of the two genes could be detected in a further 59 genomes of endospore-forming bacteria (Figure 1.1B) reported in the genomic databases of the Comprehensive Microbial Resource (CMR) and Integrated Microbial Genomes (IMG) (Supplementary Table 1).

The profile analysis was then used to detect Spo0A or Gpr in publicly available environmental metagenomes. For this, 73 microbial metagenomic datasets (Supplementary Table 2) from a total of 25 publications or direct submissions were retrieved. The datasets consisted of 6220494 sequences of average length of 957 bp and represented different environments, including marine, fresh- and ground-waters, acid mine drainage, compost, hypersaline environments, hot springs, soils, sludge, food and organism-associated environments (ant fungus garden, coral, fish and human gut).

The profile analysis revealed only three sequences with a score above the cutoff of the Spo0A profile in all metagenomic datasets (Figure 2.2A). All three metagenomes (AAQL, BAAY, BAAZ) originated from human gut [48,49], in which *Firmicutes* are known to be one of the dominant bacterial groups [50,51]. For the *gpr* gene profile (Figure 2.2B), no sequences were found with a similarity score above the cutoff value. These results are

surprising considering that some of these metagenomes were sampled in environments with high abundance of endospore-forming *Firmicutes* (e.g. gut or soil; [52,53]).
These results showed that these two genes from endospore-forming *Firmicutes* are underrepresented in metagenomes. This had been alluded to earlier by von Meringet *al.*, [24], and is now confirmed here.

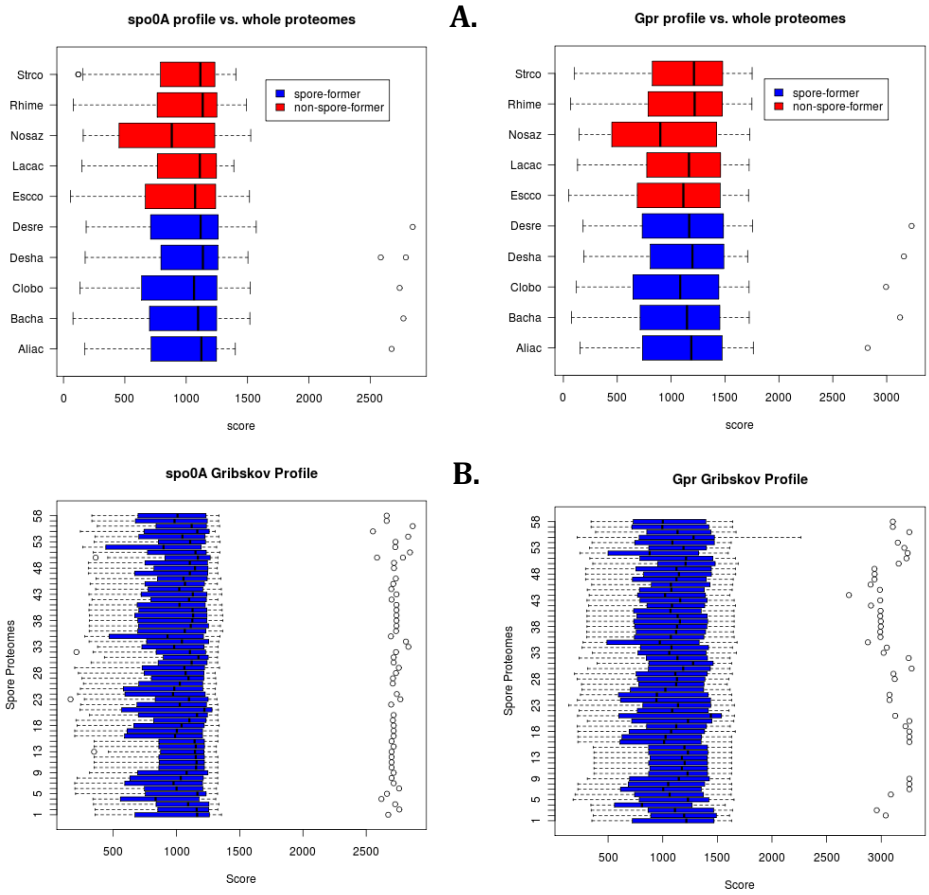


Figure 2.1 A. Validation of the profiles created for the genes *spoo*A and gpr compared to a selection of genomes of endospore-forming *Firmicutes* (blue bars) and non-spore-forming genomes (red bars). In endospore-forming *Firmicutes* a single hit with a score above 2000 (*spoo*A) and 2500 (gpr) distinguish between positive and negative hits. Strco= *Streptomyces coelicolor*; Rhime= *Rhizobium melliloti*; Nosaz: *Nostocazollae*; Lacac= *Lactobacillus acidophilus*; Escco= *Escherichia coli*; Desre= *Desulfotomaculum reducens*; Desha= *Desulfitobacterium hafniense*; Clobo= *Clostridium botulinum*; Bacha= *Bacillus halodurans*; Aliac= *Alicyclobacillus acidocaldarius*. B. The same analysis was repeated using all 59 endospore-forming genomes retrieved from IMG and CMR databases (see supplementary Table 1).

A methodological bias during the DNA extraction of resistant structures such as bacterial endospores has been suggested as the origin of an underrepresentation of microbial groups producing these structures [24]. Indeed, independently of the methodological approach taken (i.e. whole genome shotgun analysis, activity- or sequence-driven screening), the first and most crucial step in any metagenomic project is the extraction of nucleic acids. The isolated DNA should be representative of all cells in the sample and of sufficient quality and amount for subsequent sequencing [54]. Clearly, not all microbial species are equally amenable to the DNA extraction methods used today [9,10], especially considering the diversity of morphological and physiological states in which microbes can be found in environmental samples. Therefore, complementary

information, in particular concerning the method used for DNA extraction of the metagenomes was thus considered. The described DNA extraction methods (Supplementary Table 2.2) consisted of enzymatic or chemical protocols (18 datasets) or mechanical procedures of cell lysis (8 datasets). Sequences associated to *Firmicutes* are reported for some of the analyzed metagenome projects regardless of the DNA extraction protocol. For example, sequences of clostridia (30 %) and bacilli (1 %) were reported in the wallaby gut extracted enzymatically [55]. Also, in the compost metagenome extracted by bead beating, more than 13 % of sequences were reported as members of endospore-formers *Bacillus* spp. or *Paenibacillus* spp. [56]. Our profile analyses however, do not show positive hits for Spo0A or Gpr in either of these metagenomes. Whether this is due to the extraction method applied, to the depth of sequencing or to other specific bias is hard to establish. We have developed a tailored DNA extraction method that allows a better assessment of the abundance and diversity of endospore-formers in environmental samples for amplicon sequencing [12, 57]. Therefore, we next evaluated if using this extraction protocol in an environmental sample could improve the detection of endospore-formers in a metagenome.

2.3.3. Amplicon sequencing of an environmental sample with high prevalence of endospore-forming Firmicutes

We performed amplicon sequencing from a sample in which high prevalence of endospore-forming *Firmicutes* was suspected from the ratio of 16S rRNA (bacterial) and *spo0A* (endospore-formers) gene numbers measured by quantitative PCR [58]. This ratio was obtained from DNA extracted using our modified protocol. Sequencing of the 16S rRNA and *spo0A* gene amplicons was conducted and revealed not only a high prevalence of endospore-forming *Firmicutes*, but also a high diversity of endospore formers (Figure 2.3). In the amplicon sequencing of the 16S rRNA gene, *Firmicutes* accounted for 41,70% of the total bacterial community. The abundance of 16S rRNA amplicons corresponding to *Firmicutes* was nearly double the amount of Proteobacteria, which was the second most abundant bacterial Phylum (26.14%).

Among the endospore-formers observed in the pyrosequencing results, the genera *Clostridium* and *Desulfosporosinus* dominated the community in the sample, indicating a clear dominance of anaerobic endospore-formers [59] as could be expected considering the temperature and other environmental conditions at this geothermal spring. Amplicons affiliated to *Clostridium* and *Desulfosporosinus* were also dominant in the Spo0A amplicon sequencing, which also showed the dominance of anaerobic endospore-formers. Even though Spo0A sequences related to aerobic endospore-formers (e.g. *Geobacillus* and *Bacillus*) were also obtained, the classification of the Spo0A from aerobic endospore-formers was ambiguous as shown by the existence of, for example, clades related to *Anoxybacillus* but placed at different positions in the phylogeny (Figure 2.3C). In fact, only recently environmental *spo0A* sequences have started to be obtained [12], and the phylogenetic assignment needs to be refined.

2.3.3. Metagenomic sequencing

In addition to pyrosequencing, the same sample was also subjected to metagenomic sequencing. It is worth mentioning that in whole-genome metagenomics a PCR amplification biases does not apply and thus we did not necessarily expect to find the same groups or the same frequency detected in the amplicon sequencing. However, the results of the qPCR quantification and the amplicon sequencing were taken as an indication of the prevalence of *Firmicutes* in this specific environmental sample. The NAP dataset consisted of a total of 481810 sequences of average length of 330 bp. When the Spo0A and Gpr profile analysis were conducted on this metagenome, none of the two genes were detected. However, looking only at two specific genes could be an issue, since those could be, for various reasons, underrepresented in the sequences. Therefore, an extended search for reads that could be assigned to *Firmicutes* using different prediction tools on the assembled metagenome was also carried out.

Relative abundances from classified reads were considered to establish the five most prevalent phyla present in the sample (Table 2.1).

Firmicutes appear in the top five Phyla only for two of the four prediction tools used. In the case of Kraken, *Firmicutes* reads corresponded to 1.60% of the classified data, being the third most abundant phylum (the most abundant one was Proteobacteria with 82.71%). BWA predicted 5.32% of the classified sequences as to belong to *Firmicutes* (second most abundant phylum after Proteobacteria with 75.21%). *Firmicutes* were not listed after classification with MetaPhlAn and LMAT. Likewise, when reconstruction of full bacterial genomes was attempted for the NAP metagenome using MetaPhlAn, none of the top 5 microorganisms was assigned to *Firmicutes* (data not shown).

Thus, even though amplicon sequencing revealed a large fraction of the community as belonging to *Firmicutes*, this was not observed in the shotgun metagenome. There are several possible explanations for these results. One of those is the fact that the ribosomal (rrn) operon is normally found in several copies and thus the representation of a microbial community based on 16S rRNA gene sequencing is skewed. Furthermore, the average number of rrn operon copies depends on the group of bacteria. An average value of 7.01 copies of 16S rRNA genes was found for the phylum *Firmicutes* in the rrnDB [60], which implies that this group can be overrepresented in 16S rRNA gene amplicon libraries. In addition, it should be noted that for all the tools used, classification was poor and only a very small fraction of the sequences could be actually assigned to a particular taxonomic group. Therefore, the lack of detection of *Firmicutes* could be due to the current limitations of the analysis tools. In fact, recent sequencing technologies generate such large quantities of data as to bring along a new set of challenges in data analysis, the so-called bioinformatics bottleneck [61]. On the level of interpretation of metagenomic data there is still an important amount of unexplored information available from the results, simply because the advances in sequencing technologies are greater than the complementary progress in annotation, data inventory and standardization of metadata [14].

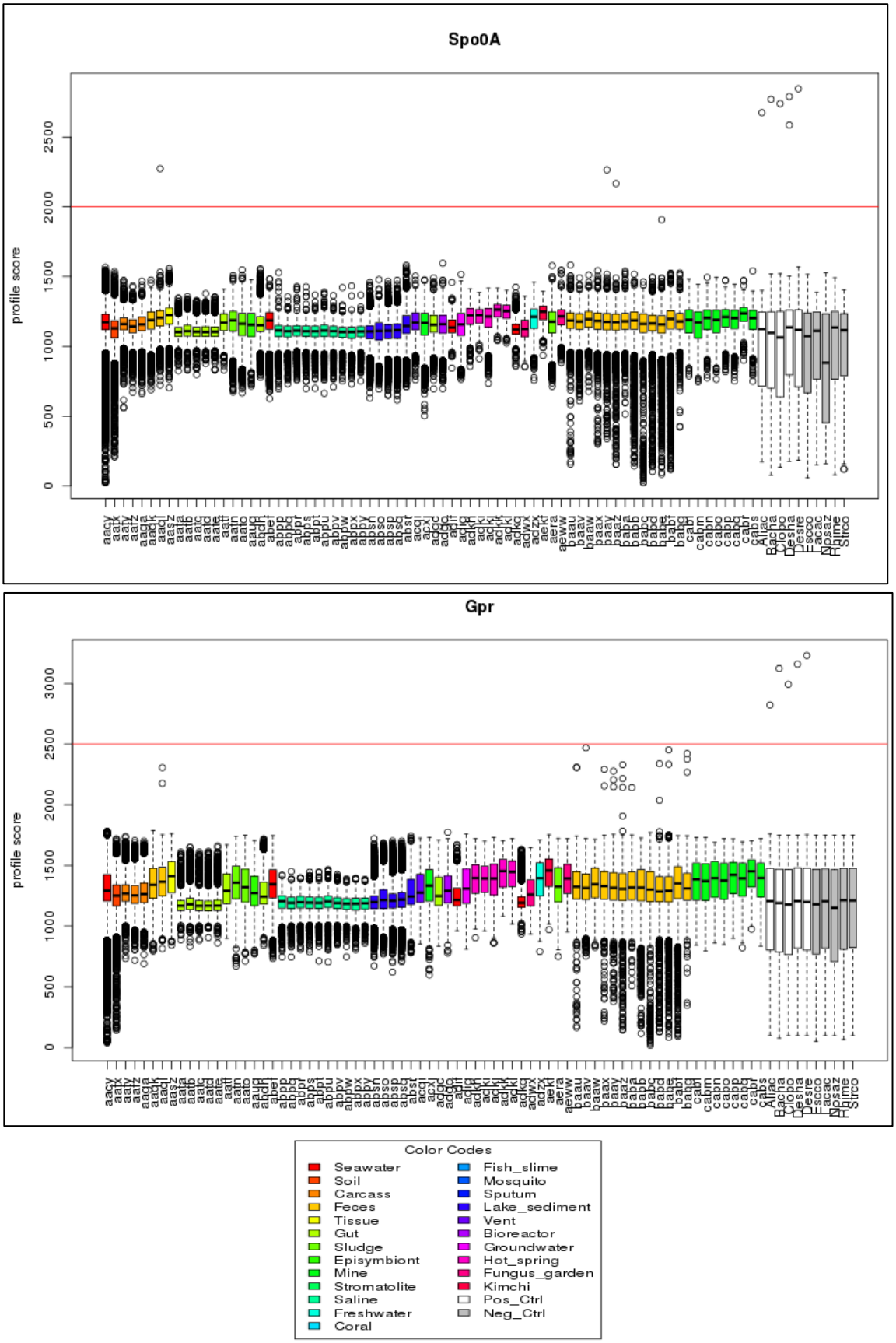
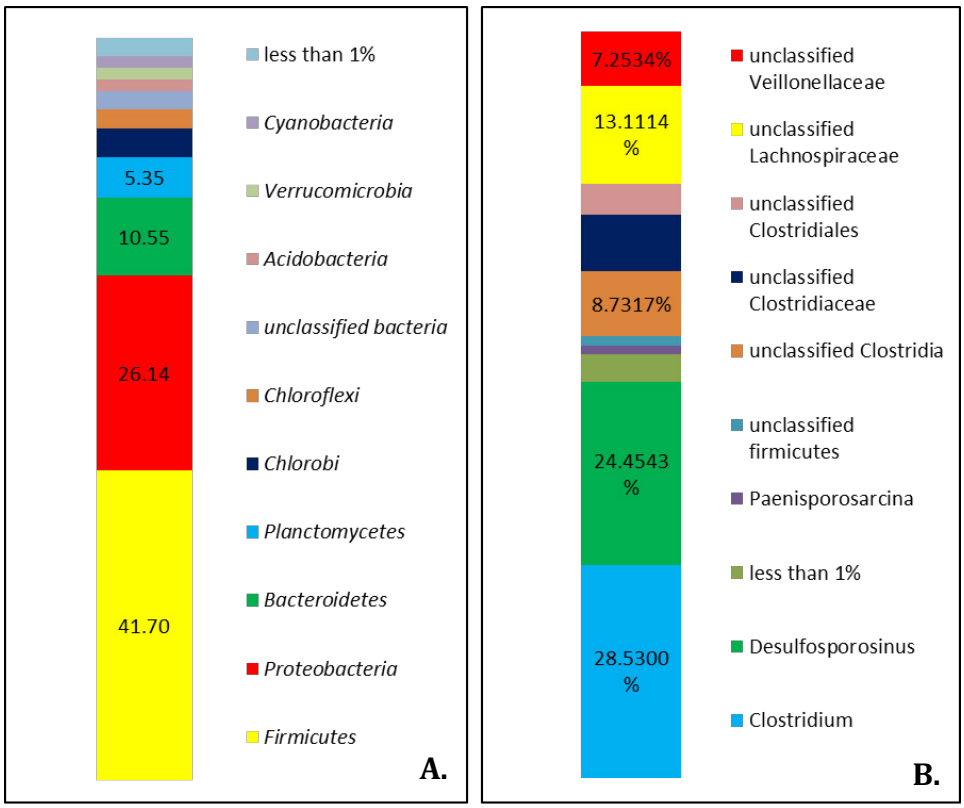


Figure 2.2 Profile similarity hits for Spo0A and Gpr protein profiles in metagenomes from different origins. The color code identifying different environments is presented under the results. The genomes included in profile testing (see Figure 2.1A) were also included in the analysis and are presented in white (endospore-formers) and grey (non-spore formers).



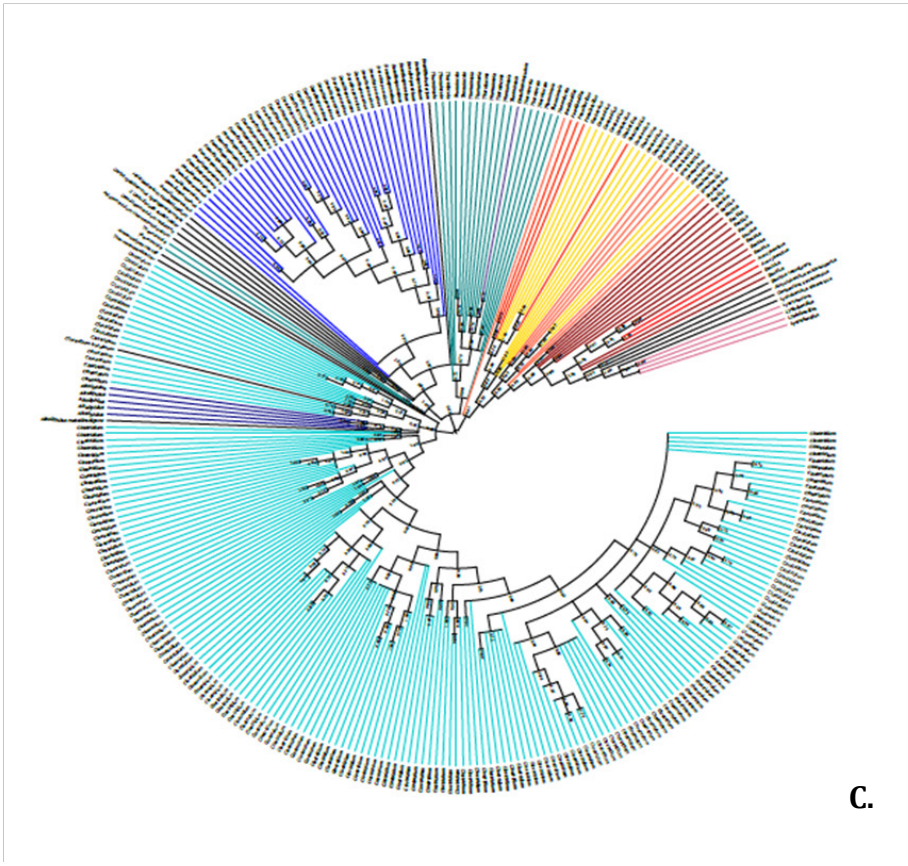


Figure 2.3. Analysis of pyrosequencing results obtained from 16S rRNA gene and *spoA* amplicons, from an environmental sample with high prevalence of endospore-forming *Firmicutes* (Nea Apollonia, NAP). (A) Total 16S rRNA gene community composition to the phylum level. (B) Firmicutes fraction of the total community (16S rRNA gene) to the genus level. (C). Cladogram representing the community composition of *Firmicutes* using the *spoA* gene. Sequences color coded by genus.

2.4. Conclusions

Since Staley and Konopka introduced the "great plate count anomaly" [62,63], revealing that only a small fraction of the microbial community can be cultured in the laboratory, one of the great challenges in environmental microbiology is the understanding of the diversity and metabolic capabilities of microbes in a culture-independent manner. That bias was partly overcome by moving into the direction of directly extracting genetic material from environmental samples. However, our results reveal that for specific microbial groups, we are still in a phase in which, similar to a percentage of the community being *not culturable* in culture-based approaches, a fraction of the genomes of the community might be considered as *not detectable* for culture-independent approaches. Nonetheless, profiling of the taxonomic and phylogenetic composition of microbial communities is at the heart of many metagenomic studies, and it is an obligatory step to draw conclusions on the role of microorganisms in the environment based on metagenomics. Our results suggest that in the case of endospore-forming *Firmicutes*, classification by various methods still lags behind. However, starting from samples such as NAP, in which evidence for high frequency of this bacterial group exists, could be the first step towards developing improved methods of classification and phylogenetic assignment of metagenomic data.

2.5. Acknowledgments

This work was supported by the Swiss National Science Foundation grant No. 31003A-132358/1 and 31003A_152972, from Fondation Pierre Mercier pour la science and from REGARD for equality of women in science.

2.6. References

[1] Suenaga H. Targeted metagenomics: a high-resolution metagenomics approach for specific gene clusters in complex microbial communities. Environ Microbiol 2012;14:13–22. doi:10.1111/j.1462-2920.2011.02438.x.

[2] Warnecke F, Hugenholtz P. Building on basic metagenomics with complementary technologies. Genome Biol 2007;8:231. doi:10.1186/gb-2007-8-12-231.

[3] Xu J. Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. MolEcol 2006;15:1713–31. doi:10.1111/j.1365-294X.2006.02882.x.

[4] Béjà O, Suzuki MT, Heidelberg JF, Nelson WC, Preston CM, Hamada T, et al. Unsuspected diversity among marine aerobic anoxygenic phototrophs. Nature 2002; 415:630–3. doi: 10.1038/415630a.

[5] Béjà O, Aravind L, Koonin EV, Suzuki MT, Hadd A, Nguyen LP, et al. Bacterial rhodopsin: evidence for a new type of phototrophy in the sea. Science 2000; 289:1902–6.

[6] Ram RJ, Verberkmoes NC, Thelen MP, Tyson GW, Baker BJ, Blake RC, et al. Community proteomics of a natural microbial biofilm. Science 2005; 308:1915–20. doi:10.1126/science. 1109070.

[7] Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA, et al. Environmental genome shotgun sequencing of the Sargasso Sea. Science 2004;304:66–74. doi:10.1126/science.1093857.

[8] Voget S, Leggewie C, Uesbeck A, Raasch C, Jaeger K-E, Streit WR. Prospecting for novel biocatalysts in a soil metagenome. Appl Environ Microbiol 2003; 69:6235–42.

[9] Delmont TO, Robe P, Cecillon S, Clark IM, Constancias F, Simonet P, et al. Accessing the soil metagenome for studies of microbial diversity. Appl Environ Microbiol 2011; 77:1315–24. doi:10.1128/AEM.01526-10.

[10] Lombard N, Prestat E, van Elsas JD, Simonet P. Soil-specific limitations for access and analysis of soil microbial communities by metagenomics. FEMS MicrobiolEcol 2011; 78:31–49. doi:10.1111/j.1574-6941.2011.01140.x.

[11] Pinard R, de Winter A, Sarkis GJ, Gerstein MB, Tartaro KR, Plant RN, et al. Assessment of whole genome amplification-induced bias through high-throughput, massively parallel whole genome sequencing. BMC Genomics 2006; 7:216. doi: 10.1186/1471-2164-7-216.

[12] Wunderlin T, Junier T, Roussel-Delif L, Jeanneret N, Junier P. Stage 0 sporulation gene A as a molecular marker to study diversity of endospore-forming Firmicutes. Environ Microbiol Rep 2013;5:911–24. doi:10.1111/1758-2229.12094.

[13] Batmalle CS, Chiang H-I, Zhang K, Lomas MW, Martiny AC. Development and Bias Assessment of a Method for Targeted Metagenomic Sequencing of Marine Cyanobacteria. Appl Environ Microbiol 2014; 80:1116–25. doi:10.1128/AEM.02834-13.

[14] Yilmaz P, Gilbert JA, Knight R, Amaral-Zettler L, Karsch-Mizrachi I, Cochrane G, et al. The genomic standards consortium: bringing standards to life for microbial ecology. ISME J 2011;5:1565–7. doi:10.1038/ismej.2011.39.

[15] Lagier J-C, Armougom F, Million M, Hugon P, Pagnier I, Robert C, et al. Microbial culturomics: paradigm shift in the human gut microbiome study. ClinMicrobiol Infect 2012;18:1185–93. doi:10.1111/1469-0691.12023.

[16] Lagier J-C, Hugon P, Khelaifia S, Fournier P-E, Scola BL, Raoult D. The Rebirth of Culture in Microbiology through the Example of Culturomics To Study Human Gut Microbiota. ClinMicrobiol Rev 2015;28:237–64. doi:10.1128/CMR.00014-14.

[17] Hugenholtz P. Exploring prokaryotic diversity in the genomic era. Genome Biol 2002;3: reviews 0003.1–reviews0003.8.

[18] Nicholson WL, Munakata N, Horneck G, Melosh HJ, Setlow P. Resistance of Bacillus endospores to extreme terrestrial and extraterrestrial environments. Microbiol Mol Biol Rev MMBR 2000;64:548–72.

[19] Nicholson WL. Roles of *Bacillus* endospores in the environment. Cell Mol Life Sci CMLS 2002;59:410–6.

[20] Driks A. Overview: Development in bacteria: spore formation in *Bacillus subtilis*. Cell Mol Life Sci CMLS 2002;59:389–91.

[21] Onyenwoke RU, Brill JA, Farahi K, Wiegel J. Sporulation genes in members of the low G+C Gram-type-positive phylogenetic branch (Firmicutes). Arch Microbiol 2004; 182:182–92. doi: 10.1007/s00203-004-0696-y.

[22] Yudkin MD, Clarkson J. Differential gene expression in genetically identical sister cells: the initiation of sporulation in *Bacillus subtilis*. MolMicrobiol 2005;56:578–89. doi:10.1111/j.1365-2958.2005.04594.x.

[23] Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, et al. Microbial biogeography: putting microorganisms on the map. Nat Rev Microbiol 2006; 4:102–12. doi: 10.1038/nrmicro1341.

[24] Von Mering C, Hugenholtz P, Raes J, Tringe SG, Doerks T, Jensen LJ, et al. Quantitative phylogenetic assessment of microbial communities in diverse environments. Science 2007; 315:1126–30. doi: 10.1126/science.1133420.

[25] Wunderlin T, Junier T, Roussel-Delif L, Jeanneret N, Junier P. Endospore-enriched sequencing approach reveals unprecedented diversity of Firmicutes in sediments. EnvironMicrobiol Rep 2014;6:631–9. doi:10.1111/1758-2229.12179.

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

[27] Kriventseva EV, Rahman N, Espinosa O, Zdobnov EM. OrthoDB: the hierarchical catalog of eukaryotic orthologues. Nucleic Acids Res 2008; 36:D271–5. doi:10.1093/nar/gkm845.

[28] Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 2002;30:3059–66.

[29] Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet TIG 2000;16:276–7.

[30] Bueche M, Wunderlin T, Roussel-Delif L, Junier T, Sauvain L, Jeanneret N, et al. Quantification of endospore-forming firmicutes by quantitative PCR with the functional gene *spo0A*. Appl Environ Microbiol 2013;79:5302–12. doi:10.1128/AEM.01376-13.

[31] Li H, Zhang Y, Li D, Xu H, Chen G, Zhang C. Comparisons of different hypervariable regions of *rrs* genes for fingerprinting of microbial communities in paddy soils. Soil BiolBiochem 2009;41:954–68. doi:10.1016/j.soilbio.2008.10.030.

[32] Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010;7:335–6. doi:10.1038/nmeth.f.303.

[33] Ewing B, Green P. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res 1998; 8:186–94.

[34] Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. SystBiol 2003; 52:696–704.

[35] Mulder NJ, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, et al. InterPro: an integrated documentation resource for protein families, domains and functional sites. Brief Bioinform 2002;3:225–35.

[36] Freitas TAK, Li P-E, Scholz MB, Chain PSG. Accurate read-based metagenome characterization using a hierarchical suite of unique signatures. Nucleic Acids Res 2015:gkv180. doi:10.1093/nar/gkv180.

[37] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. BioinformaOxfEngl 2009; 25:1754–60. doi:10.1093/bioinformatics/btp324.

[38] Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. Nat Methods 2012; 9:811–4. doi:10.1038/nmeth.2066.

[39] Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012; 9:357–9. doi:10.1038/nmeth.1923.

[40] Pruitt KD, Brown GR, Hiatt SM, Thibaud-Nissen F, Astashyn A, Ermolaeva O, et al. RefSeq: an update on mammalian reference sequences. Nucleic Acids Res 2014; 42:D756–63. doi:10.1093/nar/gkt1114.

[41] Sunagawa S, Mende DR, Zeller G, Izquierdo-Carrasco F, Berger SA, Kultima JR, et al. Metagenomic species profiling using universal phylogenetic marker genes. NatMethods 2013; 10:1196–9. doi:10.1038/nmeth.2693.

[42] Ames SK, Hysom DA, Gardner SN, Lloyd GS, Gokhale MB, Allen JE. Scalable metagenomic taxonomy classification using a reference genome database. BioinformaOxfEngl 2013; 29:2253–60. doi:10.1093/bioinformatics/btt389.

[43] Stragier P, Losick R. Molecular genetics of sporulation in Bacillus subtilis. AnnuRev Genet 1996; 30:297–241. doi:10.1146/annurev.genet.30.1.297.

[44] Abecasis AB, Serrano M, Alves R, Quintais L, Pereira-Leal JB, Henriques AO. A genomic signature and the identification of new sporulation genes. J Bacteriol 2013; 195:2101–15. doi:10.1128/JB.02110-12.

[45] Traag BA, Pugliese A, Eisen JA, Losick R. Gene conservation among endospore-forming bacteria reveals additional sporulation genes in Bacillus subtilis. J Bacteriol 2013; 195:253–60. doi:10.1128/JB.01778-12.

[46] Galperin MY, Mekhedov SL, Puigbo P, Smirnov S, Wolf YI, Rigden DJ. Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. Environ Microbiol 2012; 14:2870–90. doi:10.1111/j.1462-2920.2012.02841.x.

[47] Gribskov M, McLachlan AD, Eisenberg D. Profile analysis: detection of distantly related proteins. Proc Natl AcadSci U S A 1987; 84:4355–8.

[48] Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, et al. Metagenomic analysis of the human distal gut microbiome. Science 2006;312:1355–9. doi:10.1126/science.1124234.

[49] Kurokawa K, Itoh T, Kuwahara T, Oshima K, Toh H, Toyoda A, et al. Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. DNA Res Int J Rapid Publ Rep Genes Genomes 2007; 14:169–81. doi:10.1093/dnares/dsm018.

[50] Zoetendal EG, Vaughan EE, De Vos WM. A microbial world within us. MolMicrobiol 2006; 59:1639–50. doi:10.1111/j.1365-2958.2006.05056.x.

[51] Suzuki TA, Worobey M. Geographical variation of human gut microbial composition. Biol Lett 2014; 10:20131037. doi:10.1098/rsbl.2013.1037.

[52] Felske ADM, Tzeneva V, Heyrman J, Langeveld MA, Akkermans ADL, De Vos P. Isolation and biodiversity of hitherto undescribed soil bacteria related to *Bacillus niacini.* MicrobEcol 2004; 48:111–9. doi: 10.1007/s00248-003-2025-4.

[53] Hoyles L, Honda H, Logan NA, Halket G, La Ragione RM, McCartney AL. Recognition of greater diversity of *Bacillus* species and related bacteria in human faeces. Res Microbiol 2012; 163:3–13. doi:10.1016/j.resmic.2011.10.004.

[54] Thomas T, Gilbert J, Meyer F. Metagenomics - a guide from sampling to data analysis. MicrobInformExp 2012;2:3. doi: 10.1186/2042-5783-2-3.

[55] Pope PB, Denman SE, Jones M, Tringe SG, Barry K, Malfatti SA, et al. Adaptation to herbivory by the Tammar wallaby includes bacterial and glycoside hydrolase profiles different from other herbivores. Proc Natl AcadSci U S A 2010; 107:14793–8. doi:10.1073/pnas.1005297107.

[56] Allgaier M, Reddy A, Park JI, Ivanova N, D'haeseleer P, Lowry S, et al. Targeted discovery of glycoside hydrolases from a switchgrass-adapted compost community. PloS One 2010; 5:e8812. doi:10.1371/journal.pone.0008812.

[57] Wunderlin T, Junier T, Roussel-Delif L, Junier P. Profile analyses of sporulation genes reveal underrepresentation of endospore- forming bacteria in metagenomes. ISME14, Copenhagen, Denmark: 2012.

[58] Filippidou S, Bueche M, Wunderlin T, Junier T, Roussel-Delif L, Jeanneret N, et al. Survival strategy meets classic ecological theory: the case of diversity and abundance of endospore-forming Firmicutes in extreme environments. *in preparation.*

[59] Schleifer KH. Classification of Bacteria and Archaea: past, present and future. Syst Appl Microbiol 2009;32:533–42. doi:10.1016/j.syapm.2009.09.002.

[60] Lee ZM-P, Bussema C, Schmidt TM. rrnDB: documenting the number of rRNA and tRNA genes in bacteria and archaea. Nucleic Acids Res 2009; 37:D489–93. doi:10.1093/nar/gkn689.

[61] Scholz MB, Lo C-C, Chain PSG. Next generation sequencing and bioinformatic bottlenecks: the current state of metagenomic data analysis. Curr Opin Biotechnol 2012; 23:9–15. doi:10.1016/j.copbio.2011.11.013.

[62] Staley JT, Konopka A. Measurement of *in situ* activities of non-photosynthetic microorganisms in aquatic and terrestrial habitats. Annu Rev Microbiol 1985; 39:321–46. doi:10.1146/annurev.mi.39.100185.001541.

[63] Amann RI, Ludwig W, Schleifer KH. Phylogenetic identification and in situ detection of individual microbial cells without cultivation.MicrobiolRev 1995; 59:143–69.

2.7. Supplementary material

Supplementary Table 2.1. Complete genome sequences from 57endospore-forming bacteria used in this study. All genomes were retrieved from the CMR database.

Name	NCBI Taxon ID	Classification	Sequencing status	Isolation	Genes	CDS	rRNA	GC Perc	Bases	Database
Alicyclobacillus acidocaldarius subsp. acidocaldarius DSM 446	521098	Bacilli	Finished	"Yellowstone National Park, acid hot spring"	3235	3150	18	0.62	3205686	IMG
Alkaliphilus metalliredigens QYMF	293826	Clostridia	Finished	Borax leachate ponds	5016	4801	31	0.37	4929566	CMR
Alkaliphilus oremlandii OhILAs	350688	Clostridia	Finished	Freshwater USA	2951	2836	26	0.36	3123558	CMR
Anaerostipes caccae DSM 14662	411490		Draft	Human feces collected in UK	3845		26	0.44	3605636	IMG
Anoxybacillus flavithermus WK1	491915	Bacilli	Finished	waste water drain at the Wairakei geothermal power station in New Zealand	2933	2832	24	0.42	2846746	IMG
Bacillus amyloliquefaciens FZB42	326423	Bacilli	Finished	Soil	3814	3696	30	0.46	3918589	CMR
Bacillus anthracis A0174	486622	Bacilli	Draft	Canada	5306	5198	29	0.35	5291908	IMG
Bacillus anthracis A0193	486619	Bacilli	Draft	Cow Bovine isolate obtained in South Dakota, USA	5416	5309	28	0.35	5392880	IMG
Bacillus anthracis A0248	592021	Bacilli	Finished	human isolated from USAMRIID, Ohio	5418	5291	33	0.35	5503926	CMR
Bacillus anthracis A0389	486623	Bacilli	Draft	"Bekasi, Indonesia"	5416	5296	33	0.35	5420403	IMG
Bacillus anthracis A0442	486621	Bacilli	Draft	Human isolate obtained in the Kruger National Park, South Africa"	5380	5256	29	0.35	5374836	IMG
Bacillus anthracis A0465	486620	Bacilli	Draft	Bovine isolate obtained from France	5420	5300	29	0.35	5407149	IMG
Bacillus anthracis A0488	486624	Bacilli	Draft	Infected cattle in the United Kingdom in 1935	5404	5288	33	0.35	5392168	IMG
Bacillus anthracis Ames	198094	Bacilli	Finished	Soil	5545	5311	33	0.35	5227293	CMR
Bacillus anthracis CDC 684	568206	Bacilli	Finished	Soil	6031	5902	33	0.35	5506763	CMR
Bacillus anthracis Sterne	260799	Bacilli	Finished	Soil	5521	5287	33	0.35	5228663	CMR
Bacillus cereus 03BB102	572264	Bacilli	Finished	Human blood	5767	5621	42	0.35	5449308	CMR
Bacillus cereus 10987	405917	Bacilli	Finished	Canada during a study of cheese spoilage	5601	5601		0.35	5432652	CMR
Bacillus cereus cytotoxis NVH 391-98	315749	Bacilli	Finished	Soil	4250	4015	39	0.36	4094159	CMR
Bacillus cereus Zk	288681	Bacilli	Finished	wab of a dead zebra carcass in Namibia	5134	5323	39	0.35	5300915	CMR
Bacillus clausii KSM-K16	66692	Bacilli	Finished	Soil	4108	4108	22	0.44	4303871	CMR
Bacillus coagulans 36D1	345219	Bacilli	Permanent Draft	Intestinal tract	3461	3347	30	0.46	3551306	IMG
Bacillus halodurans C-125	272558	Bacilli	Finished	Freshwater, Soil	4239	4066	25	0.44	4202352	IMG
Bacillus licheniformis ATCC 14580 (Goettingen)	279010	Bacilli	Finished	Soil	4356	4196	21	0.46	4222645	CMR
Bacillus licheniformis ATCC 14580 (Novozymes)	279010	Bacilli	Finished	Soil	4420	4196	21	0.46	4222334	CMR
Bacillus mycoides DSM 2048	526997	Bacilli	Draft	Soil	5747	5658	6	0.35	5541906	IMG
Bacillus pseudomycoides DSM 12442	527000	Bacilli	Draft	Soil	5941	5851	4	0.35	5752014	IMG
Bacillus pumilus SAFR-032	315750	Bacilli	Finished	Soil	3823	3729	21	0.41	3704465	CMR
Bacillus subtilis 168	224308	Bacilli	Finished	X-ray irradiated strain in Marburg in 1947	4298	4106	30	0.44	4214630	CMR
Bacillus thuringiensis Al Hakam	412694	Bacilli	Finished	Severe human tissue necrosis	5050	4798	42	0.35	5313030	CMR
Bacillus thuringiensis sv konkukian 97-27	281309	Bacilli	Finished	Severe human tissue necrosis	5452	5197	42	0.35	5314794	CMR
Bacillus weihenstephanensis KBAB4	315730	Bacilli	Finished	Soil	5983	5831	42	0.35	5872743	CMR
Candidatus Desulforudis audaxviator MP104C	477974	Clostridia	Finished	Fracture water from a borehole at a depth of 2.8 km in a South African gold mine	2293	2239	6	0.61	2349476	CMR
Carboxydothermus hydrogenoformans Z-2901	246194		Finished	hot swamp from Kunashir Island, Russia.	2707	2645	12	0.42	2401520	CMR
Clostridium beijerinckii NCIMB 8052	290402	Clostridia	Finished	Freshwater, Soil	5290	5100	43	0.3	6000632	CMR
Clostridium botulinum A ATCC 19397	441770	Clostridia	Finished	Laboratory strain probably from foodborne botulism cases in the western US	3750	3596	24	0.28	3863450	CMR
Clostridium botulinum A Hall	441771	Clostridia	Finished	Harvard University in 1947	3622	3462	24	0.28	3760560	CMR
Clostridium botulinum A2 Kyoto-F	536232	Clostridia	Finished	infant botulism in Kyoto, Japan in 1978	3978	3878	20	0.28	4155278	CMR
Clostridium botulinum A3 Loch Maree	498214	Clostridia	Finished	Duck liver paste during a botulism outbreak at a hotel in the Scottish highlands in 1922.	4092	3984	27	0.28	4259691	CMR
Clostridium botulinum B Eklund 17B	508765	Clostridia	Finished	Marine sediments taken off the coast of Washington, USA	3639	3527	34	0.27	3847969	CMR
Clostridium botulinum B1 Okra	498213	Clostridia	Finished	Foodborne botulism incident in the US	3961	3852	27	0.28	4107013	CMR
Clostridium botulinum Ba4 657	515621	Clostridia	Finished	Infant botulism case in Texas in 1976	4206	4097	27	0.28	4257769	CMR
Clostridium botulinum F Langeland	441772	Clostridia	Finished	Home-prepared liver paste involved in an outbreak of foodborne botulism on the island of Langeland in Denmark in 1958	3863	3705	27	0.28	4012918	CMR
Clostridium difficile 630	272563	Clostridia	Finished	clincal isolate Switzerland	3983	3777	32	0.29	4298133	CMR
Clostridium kluyveri DSM 555	431943	Clostridia	Finished	Mud of a canal in Delft, The Netherlands	4073	3913	20	0.32	4023800	CMR
Clostridium novyi NT	386415	Clostridia	Finished	Soil	2485	2325	29	0.29	2547720	CMR
Clostridium perfringens 13	195102	Clostridia	Finished	Soil isolate	2905	2723	30	0.29	3085740	CMR
Clostridium perfringens ATCC 13124	195103	Clostridia	Finished	Soil	3066	2895	24	0.28	3256683	CMR
Clostridium perfringens SM101	289380	Clostridia	Finished	Soil	2748	2578	30	0.28	2921996	CMR
Desulfitobacterium hafniense Y51	138119	Clostridia	Finished	Soil contaminated with tetrachloroethene in Japan	5208	5060	18	0.47	5727534	CMR
Geobacillus kaustophilus HTA426	235909	Bacilli	Finished	Deep Sea, Marine	3714	3540	27	0.52	3592666	CMR
Geobacillus thermodenitrificans NG80-2	420246	Bacilli	Finished	Oil reservoir formation water taken at a depth of 2000 m	3642	3471	30	0.49	3608012	CMR
Heliobacterium modesticaldum Ice1	498761	Clostridia	Finished	hot spring microbial mats and volcanic soil in Iceland	3142	3001	30	0.57	3075407	CMR
Lysinibacillus sphaericus C3-41	444177	Bacilli	Finished	Mosquito breeding site in China in 1987	4887	4771	31	0.37	4817463	CMR
Pelotomaculum thermopropionicum SI	370438	Clostridia	Finished	Thermophilic upflow anaerobic sludge blanket reactor	3018	2920	6	0.53	30	
Thermoanaerobacter pseudethanolicus ATCC 33223	340099	Clostridia	Finished	Thermal springs in Yellowstone National Park	2363	2291	16	0.35	23	
Thermoanaerobacter sp. X514	399726	Clostridia	Finished	Anaerobic enrichment culture from a deep subsurface sample (2000 m below the surface) taken from a core hole at the Piceance Basin, Colorado, USA	2467	2397	16	0.35	24	

Supplementary Table 2.2. Metagenomic datasets used to test the presence of endospore-forming bacteria in environmental samples by profile analyses with Spo0A and Gpr. Grouping types of DNA information was used for the color code in Figure 2.1. For all direct submissions (Reference column) the cell lysis method was not indicated in the submission. NA= the information on the percentage of Firmicutes in the dataset is not available.

Accession Code	Grouping types of DNA	Cell lysis method	Positive hit Spo0A	Firmicutes abundance (%)	Reference
ADGO	Compost	Bead-beating		*Bacillus* sp. 7%, *Bacillus* sp. 5%, *Paenibacillus* sp. 1%,	(Allgaier et al 2010)
AEWW	Ant fungus garden	NA		NA	Aylward et al, Direct submission
CABL, CABM, CABN, CABO, CABP, CABQ, CABR, CABS	Mine drainage	Vortex		NA	(Bertin et al 2011)
ACQI	Hydrothermal vent	Mortar		NA	(Brazelton and Baross 2009)
ACXJ	Mine drainage	Enzyme, chemical		NA	(Dick et al 2009)
ADZX	Freshwater	Enzyme, chemical		Firmicutes 5%	(Ferrer et al 2011)
AAQK, AAQL	Gut	Bead-beating	yes	16S: Firmicutes > 85%	(Gill et al 2006)
AAUQ	Epibiont	Enzyme, chemical		NA	(Grzymski et al 2008)
ADIG	Groundwater	Mortar, freeze-thaw		*Bacilli* 0.12%	(Hemme et al 2010)
ADKH, ADKI, ADKJ, ADKK, ADKL	Hot Springs	NA		NA	Inskeep et al, 2010, Direct submission
AEKF	Food	Bead-beating		NA	(Jung et al 2011)
ABSN, ABSO, ABSP, ABSQ, ABSR	Freshwater	Enzyme, chemical		NA	(Kalyuzhnaya et al 2008)
ABEF	Marine	Enzyme, chemical		NA	(Konstantinidis et al 2009)
ABPP, ABPQ, ABPR, ABPS, ABPT, ABPU, ABPV, ABPW, ABPX, ABPY	Hypersaline mat	Vortex		NA	(Kunin et al 2008)
BAAU, BAAV, BAAW, BAAX, BAAY, BAAZ, BABA, BABB, BABC, BABD, BABE, BABF, BABG	Gut	Enzyme, chemical	yes	NA	(Kurokawa et al 2007)
ADIF, ADKQ	Marine	NA		NA	Lucas et al, 2010, Direct submission
ADGC	Gut	Enzyme, chemical		*Clostridia* 30%, *Bacilli* <1%	(Pope et al 2010)
AERA	Sludge	NA		NA	Purohit et al, 2011, Direct submission
ADWX	Ant fungus garden	Physical		NA	(Suen et al 2010)
AAFX, AAFY, AAFZ, AAGA	Soil, Whale fall	Enzyme, chemical		NA	(Tringe et al 2005)
AATA, AATB, AATC, AATD, AATE, AATF	Gut	Physical		NA	(Turnbaugh et al 2006)
AACY	Marine	Enzyme, chemical		0.5% Firmicutes (clones)	(Venter et al 2004)
ABDH	Gut	Enzyme, chemical		7% Firmicutes (16S based) of total of 216 OTUs	(Warnecke et al 2007)
AASZ	Oligochaete symbionts	Enzyme, chemical		NA	(Woyke et al 2006)
AATN, AATO	Sludge	Bead-beating		NA	(Martin et al 2006)

2. 8. References on Supplementary Material

Allgaier M, Reddy A, Park JI, Ivanova N, D'Haeseleer P, Lowry S et al (2010). Targeted Discovery of Glycoside Hydrolases from a Switchgrass-Adapted Compost Community. Plos One5.

Bertin PN, Heinrich-Salmeron A, Pelletier E, Goulhen-Chollet F, Arsene-Ploetze F, Gallien S et al (2011). Metabolic diversity among main microorganisms inside an arsenic-rich ecosystem revealed by meta- and proteo-genomics. ISME J5: 1735-1747.

Brazelton WJ, Baross JA (2009). Abundant transposases encoded by the metagenome of a hydrothermal chimney biofilm. Isme Journal3: 1420-1424.

Challacombe JF, Altherr MR, Xie G, Bhotika SS, Brown N, Bruce D et al (2007). The complete genome sequence of *Bacillus thuringiensis* Al Hakam. J Bacteriol189: 3680-3681.

Chen XH, Koumoutsi A, Scholz R, Eisenreich A, Schneider K, Heinemeyer I et al (2007). Comparative analysis of the complete genome sequence of the plant growth-promoting bacterium *Bacillus amyloliquefaciens* FZB42. Nat Biotechnol25: 1007-1014.

Chivian D, Brodie EL, Alm EJ, Culley DE, Dehal PS, Desantis TZ et al (2008). Environmental genomics reveals a single-species ecosystem deep within Earth. Science322: 275-278.

Dick GJ, Andersson AF, Baker BJ, Simmons SL, Yelton AP, Banfield JF (2009). Community-wide analysis of microbial genome sequence signatures. Genome Biology 10.

Edgar RC (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res32: 1792-1797.

Feng L, Wang W, Cheng J, Ren Y, Zhao G, Gao C et al (2007). Genome and proteome of long-chain alkane degrading *Geobacillus thermodenitrificans* NG80-2 isolated from a deep-subsurface oil reservoir. Proc Natl Acad Sci U S A104: 5602-5607.

Ferrer M, Guazzaroni ME, Richter M, Garcia-Salamanca A, Yarza P, Suarez-Suarez A et al (2011). Taxonomic and Functional Metagenomic Profiling of the Microbial Community in the Anoxic Sediment of a Sub-saline Shallow Lake (Laguna de Carrizo, Central Spain). Microbial Ecology62: 824-837.

Gill SR, Pop M, DeBoy RT, Eckburg PB, Turnbaugh PJ, Samuel BS et al (2006). Metagenomic Analysis of the Human Distal Gut Microbiome. Science312: 1355-1359.

Gioia J, Yerrapragada S, Qin X, Jiang H, Igboeli OC, Muzny D et al (2007). Paradoxical DNA repair and peroxide resistance gene conservation in *Bacillus pumilus* SAFR-032. PLoS One2: e928.

Grzymski JJ, Murray AE, Campbell BJ, Kaplarevic M, Gao GR, Lee C et al (2008). Metagenome analysis of an extreme microbial symbiosis reveals eurythermal adaptation and metabolic flexibility. P Natl Acad Sci USA105: 17516-17521.

Han CS, Xie G, Challacombe JF, Altherr MR, Bhotika SS, Brown N et al (2006). Pathogenomic sequence analysis of *Bacillus cereus* and *Bacillus thuringiensis* isolates closely related to *Bacillus anthracis*. J Bacteriol188: 3382-3390.

Hemme CL, Deng Y, Gentry TJ, Fields MW, Wu LY, Barua S et al (2010). Metagenomic insights into evolution of a heavy metal-contaminated groundwater microbial community. Isme Journal4: 660-672.

Hu X, Fan W, Han B, Liu H, Zheng D, Li Q et al (2008). Complete genome sequence of the mosquitocidal bacterium Bacillus sphaericus C3-41 and comparison with those of closely related Bacillus species. J Bacteriol190: 2892-2902.
Jung JY, Lee SH, Kim JM, Park MS, Bae JW, Hahn Y et al (2011). Metagenomic analysis of kimchi, a traditional Korean fermented food. Appl Environ Microbiol77: 2264-2274.
Junier P, Junier T, Podell S, Sims DR, Detter JC, Lykidis A et al (2010). The genome of the Gram-positive metal- and sulfate-reducing bacterium *Desulfotomaculum reducens* strain MI-1. Environ Microbiol.
Junier T, Zdobnov EM (2010). The Newick Utilities: High-throughput Phylogenetic tree Processing in the UNIX Shell. Bioinformatics.
Kalyuzhnaya MG, Lapidus A, Ivanova N, Copeland AC, McHardy AC, Szeto E et al (2008). High-resolution metagenomics targets specific functional types in complex microbial communities. Nat Biotech26: 1029-1034.
Katoh K, Kuma K, Toh H, Miyata T (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res33: 511-518.
Konstantinidis KT, Braff J, Karl DM, DeLong EF (2009). Comparative Metagenomic Analysis of a Microbial Community Residing at a Depth of 4,000 Meters at Station ALOHA in the North Pacific Subtropical Gyre. Applied and Environmental Microbiology75: 5345-5355.
Kunin V, Raes J, Harris JK, Spear JR, Walker JJ, Ivanova N et al (2008). Millimeter-scale genetic gradients and community-level molecular convergence in a hypersaline microbial mat. Mol Syst Biol4.
Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V et al (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature390: 249-256.
Kurokawa K, Itoh T, Kuwahara T, Oshima K, Toh H, Toyoda A et al (2007). Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. DNA Research14: 169-181.
Martin HG, Ivanova N, Kunin V, Warnecke F, Barry KW, McHardy AC et al (2006). Metagenomic analysis of two enhanced biological phosphorus removal (EBPR) sludge communities. Nat Biotech24: 1263-1269.
Pope PB, Denman SE, Jones M, Tringe SG, Barry K, Malfatti SA et al (2010). Adaptation to herbivory by the Tammar wallaby includes bacterial and glycoside hydrolase profiles different from other herbivores. P Natl Acad Sci USA107: 14793-14798.
Rey MW, Ramaiya P, Nelson BA, Brody-Karpin SD, Zaretsky EJ, Tang M et al (2004). Complete genome sequence of the industrial bacterium *Bacillus licheniformis* and comparisons with closely related *Bacillus* species. Genome Biol5: R77.
Rice P, Longden I, Bleasby A (2000). EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet16: 276-277.
Sattley WM, Madigan MT, Swingley WD, Cheung PC, Clocksin KM, Conrad AL et al (2008). The genome of *Heliobacterium modesticaldum*, a phototrophic representative of the Firmicutes containing the simplest photosynthetic apparatus. J Bacteriol190: 4687-4696.

Sebaihia M, Wren BW, Mullany P, Fairweather NF, Minton N, Stabler R et al (2006). The multidrug-resistant human pathogen *Clostridium difficile* has a highly mobile, mosaic genome. Nature Genetics38: 779-786.

Seedorf H, Fricke WF, Veith B, Bruggemann H, Liesegang H, Strittimatter A et al (2008). The genome of *Clostridium kluyveri*, a strict anaerobe with unique metabolic features. Proceedings of the National Academy of Sciences of the United States of America105: 2128-2133.

Shimizu T, Ohtani K, Hirakawa H, Ohshima K, Yamashita A, Shiba T et al (2002). Complete genome sequence of *Clostridium perfringens*, an anaerobic flesh-eater. Proc Natl Acad Sci U S A99: 996-1001.

Suen G, Scott JJ, Aylward FO, Adams SM, Tringe SG, Pinto-Tomas AA et al (2010). An Insect Herbivore Microbiome with High Plant Biomass-Degrading Capacity. Plos Genetics 6.

Tringe SG, von Mering C, Kobayashi A, Salamov AA, Chen K, Chang HW et al (2005). Comparative Metagenomics of Microbial Communities. Science308: 554-557.

Turnbaugh PJ, Ley RE, Mahowald MA, Magrini V, Mardis ER, Gordon JI (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. Nature444: 1027-1031.

Venter JC, Remington K, Heidelberg JF, Halpern AL, Rusch D, Eisen JA et al (2004). Environmental Genome Shotgun Sequencing of the Sargasso Sea. Science304: 66-74.

Warnecke F, Luginbuhl P, Ivanova N, Ghassemian M, Richardson TH, Stege JT et al (2007). Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature450: 560-U517.

Woyke T, Teeling H, Ivanova NN, Huntemann M, Richter M, Gloeckner FO et al (2006). Symbiosis insights through metagenomic analysis of a microbial consortium. Nature443: 950-955.

Wu M, Ren Q, Durkin AS, Daugherty SC, Brinkac LM, Dodson RJ et al (2005). Life in hot carbon monoxide: the complete genome sequence of *Carboxydothermus hydrogenoformans* Z-2901. PLoS Genet1: e65.

Chapter 3

Extreme conditions favor the prevalence of endospore-forming Firmicutes in natural springs

Sevasti Filippidou[1], Tina Wunderlin[1], Thomas Junier[1,2], Nicole Jeanneret[1], Cristina Dorador[3], Veronica Molina[4], David R. Johnson[5], Pilar Junier[1]
Affiliations
[1]Laboratory of Microbiology, University of Neuchatel, Emile Argand 11, Neuchatel, CH-2000, Switzerland.
[2]Vital-IT group, Swiss Institute of Bioinformatics, Lausanne, CH-1015, Switzerland.
[3]Laboratorio de Complejidad Microbiana y Ecología Funcional; Departamento de Biotecnología; Facultad de Ciencias del Mar y Recursos Biológicos, Universidad de Antofagasta; CL-, 1270190, Antofagasta, Chile.
[4]Departamento de Biología, Facultad de Ciencias Naturales y Exactas. Universidad de Playa Ancha, CL-2360002, Valparaíso, Chile.
[5]Department of Environmental Microbiology, Swiss Federal Institute of Aquatic Science and Technology (Eawag), Dübendorf, CH-8600, Switzerland

This chapter is to be submitted for publication at *Frontiers in Microbiology*

Abstract

Environmental conditions unsuitable for survival and development are the rule rather than the exception in most habitats. Microorganisms have developed various strategies to withstand environmental conditions that limit active growth. Endospore-forming Firmicutes (EFF) deploy a myriad of survival strategies in order to withstand adverse conditions. Like many bacterial groups, they can form biofilms and detect nutrient scarcity through chemotaxis. Moreover, within this paraphyletic group of Firmicutes, ecophysiological optima are diverse. A unique response to adversity in this group is that when all other survival mechanisms have failed, they can enter sporulation in order to survive. These strategies are energetically demanding and therefore might affect the biological success of EFF. Therefore, we hypothesize that abundance and diversity of EFF should be enhanced in those environments in which the payoff of this combined set of survival strategies offsets its cost. Geothermal and mineral springs and drillings are examples of environments in which the steep physicochemical gradients might render diversified survival strategies a favorable strategy. In order to address this hypothesis, we have collected 71 samples from geothermal and mineral environments characterized by none, single or multiple limiting environmental factors (temperature, pH, UV radiation and presence of minerals) that can be expected to enhance the payoff of diversified survival strategies. To measure EFF success we quantified their relative abundance compared to total bacteria. This quantification showed that only the co-existence of multiple limiting environmental factors increases the relative abundance of EFF. This likely reflects the diversity of survival strategies deployed by EFF. This is supported by community composition analyses based on the 16S rRNA gene amplicon sequencing from samples representing environments with none, single or multiple limiting factors. Community composition reflected the high metabolic and functional diversity of EFF, which allows for active growth even under specific limiting conditions, while the plasticity of survival strategies such as sporulation was also observed. These results suggest that because of their greater survival rate, EFF display a unique distribution pattern that might be replicated by other microorganisms with diversified survival adaptations.

Keywords: Endospores; Firmicutes; qPCR; *spo0A*; 16S rRNA gene; extreme environments

3.1. Introduction

Representatives of the Phylum Firmicutes have been known since the dawn of microbiology. Robert Koch when studying Anthrax, described the first Firmicutes species: *Bacillus anthracis* [1]. Since then, this microbial group has never ceased to amaze microbiologists because of the survival strategies and the large functional diversity displayed by its members. Probably the best studied of these strategies is the formation of a *spore*, a structure that contains and protects the genetic material of the bacterium. Spore formation is a trait found in the phyla of Firmicutes, Cyanobacteria, Actinobacteria, and some orders of Proteobacteria. However, among all spore-like structures, only endospores produced by species belonging to Firmicutes (endospore-forming Firmicutes or EFF) can survive wet heat, as a result of both their hardy outer coat [2] and the dipicolinic acid contained in their inner core [3–5]. Endospores are said to be the most resistant cellular structures on Earth [6]. As endospores remain viable for long periods, EFF can bear long-distance transportation, leading to a higher dispersal rate compared to other bacteria [7, 8].

Although endosporulation is considered to be the ultimate survival mechanism [9], EFF can also deploy other strategies to improve survival and resistance, including motility, chemotaxis, DNA uptake, transformation, or biofilm formation [10]. In addition, EFF display a large functional diversity and are, therefore, involved in a variety of ecosystem functions [11]. Like many bacterial groups, a multitude of metabolisms can be found among Firmicutes, including aerobic and anaerobic respiration (e.g. sulfate-reduction), acetogenesis, fermentation or phototrophy [12–14]. Likewise, different species of Firmicutes display diverse ecological optima. Improved survival due to sporulation combined with metabolic diversity, offers EFF a key ecological advantage among bacteria.

The unique advantages of survival and functional plasticity have been shown by many studies leading to the isolation of EFF in a multitude of environments. EFF have been found in mesophilic environments, such as the mammalian gut [15] but also under more limiting conditions such as those found in psychrophilic, thermophilic, alkaline, acidic, and saline environments [7]. For many of these habitats, EFF might be found in the state of endospores [16–18], as is the case of the isolation of thermophilic species in artic sediments [19 ,20]. However, in other cases their metabolic and functional diversity [7] can favor the ubiquity of active cells. For example, thermophilic EFF species have been previously isolated and described from diverse geothermal environments [21, 22].

Despite the apparent ubiquitous distribution of EFF, molecular ecology studies have, paradoxically, failed to detect this group. In most environmental genomic studies this group has been under-detected [23, 24]. There are at least two potential explanations to this. On the one hand, a methodological bias against molecular detection of Firmicutes could explain these results and we have shown that a tailored molecular method allows for a better assessment of the abundance and diversity of EFF in environmental samples [25]. On the other hand, the poor detection by molecular methods might accurately reflect the relative distribution of EFF. In this case, it can be possible that the high

energetic demands of their survival strategies limit the distribution of this group under non-limiting conditions. This would suggest the biological cost of increased survival as a better explanation of the distribution patterns of EFF in the environment. However, this has not been previously studied. In this study, we test this latter hypothesis to explain the distribution patterns of EFF. For this the relative abundance and diversity of EFF was compared to that of total bacteria in different mineral springs. We have investigated 71 samples from geothermal springs and drillings and from natural mineral springs. These samples were categorized into three groups based on whether they were characterized by the presence of multiple, single or no environmental factors that can potentially limit active growth. Finally, we tested the effect of these limiting conditions on the relative abundance and diversity of EFF.

3.2. Materials and methods

3.2.1. Sampling and environmental factors

Samples were collected from 24 sites in Chile, Colombia, Germany, Greece, France, and Switzerland between 2011 and 2013 (Figure 3.1). Four environmental factors were included in this study: temperature, pH, UV radiation, and mineral characterization of the springs. At all sampling sites, temperature and pH have been measured. For the mineral characteristics of each spring, previously published data were used [26, 27].

Figure 3.1. Sampling sites worldwide. (A) Map showing the countries in Europe from which samples were collected, as well as pictures of the geothermal springs and reservoirs (Switzerland: Ponts-de-Martel; France: Soultz-sous-forets; Germany: Gross Schoenebeck and Bruschal; Greece: North: Nea Apollonia, Eleftheres, Agia Paraskevi, South: Milos B. (B) Map showing the sampling points in Colombia (Los Volcanes) and Chile (Lirima, El Tatio, Aguas Calientes), along with pictures from the sampling sites.

Based on literature definitions of extremity (Supplementary table 3.1), these environmental parameters were then transformed to a qualitative index of potentially limiting environmental factors (Supplementary table 3.2), attributed "0" for normal conditions and "1" for limiting.

3.2.2. DNA extraction

Water samples were filtered through 0.22 µm membranes to collect biomass. Soil, sediment and biofilm samples were subjected to indirect DNA extraction as previously

described [28]. DNA was extracted using the FastDNA Spin Kit for Soil (MP Biomedicals, California), using a modified protocol, in order to ensure that DNA was not only extracted from vegetative cells but also from spores and other cells difficult to lyse [28]. This modified protocol included the separation of biomass from soil and sediment using a 1% (w/v) hexa-meta-phosphate solution followed by the collection of the biomass for an indirect DNA extraction. The first step of the extraction protocol was modified to include three sequential bead-beating steps. After each step, the sample was treated independently, allowing for DNA extraction of easy-to-lyse cells (first bead-beating round), harder cell wall cells or other structures (second bead-beating), and resistant structures such as spores (second and third bead-beating). This method ensures that on one hand, DNA from all types of cells is extracted and on the other hand that the harsh bead-beating treatment does not degrade DNA already released in the previous rounds, compromising the representativeness of certain bacterial groups in downstream analysis. DNA concentration was measured with a Qubit Fluorimeter using a dsDNA HS Assay Kit (Invitrogen, California). The concentration of all samples was adjusted by dilution to 2ng/µl.

3.2.3. Quantitative PCR assays

Prior to quantification, the integrity of 16S rRNA and *spo0A* (transcriptional factor responsible for the initiation of sporulation) genes was verified by PCR amplification of the complete genes (approximately 1500bp for 16S rRNA gene and 600bp for *spo0A*). PCR amplification of the 16S rRNA gene was performed using the primer set GM3F and GM4R, according to Muyzer *et al.* [29], while for *spo0A* amplification a set of specific primers (*spo0A*166f and *spo0A*748r) was used as described previously [25]. In order to calculate the relative abundance of EFF in the bacterial communities of each sample, two quantitative PCR (qPCR) assays were used. First, for quantifying total bacteria, qPCR amplification of the V3 hyper-variable region from the 16S rRNA gene was carried out. The primers used were 338f (5'-ACTCCTACGGGAGGCAGCAG-3') and 520r (5'-ATTACCGCGGCTGCTGG-3') [29,30] and amplification was carried out under conditions previously described [31]. For the quantification of EFF, a primer pair targeting the *spo0A* gene was used as previously described [31]. Reactions were carried out in a final volume of 10 µl with 5 µl Rotor-Gene SYBR green PCR master mix (Qiagen, Germany), 1 µM and 0.45 µM of primers *spo0A*655f and *spo0A*834r respectively and 4 ng of DNA template. Amplification conditions were previously described [31]. All reactions were performed in triplicates. The standard curves for quantification of both V3 16S rRNA gene and *spo0A* gene were prepared from 10-fold dilutions (10^8 to 10^2 copies/µl) of a plasmid in which the V3 16S rRNA and *spo0A* gene of *B. subtilis* was inserted, respectively. The TOPO TA cloning kit (Invitrogen, California) was used to produce this plasmid in One Shot TOP10F' chemically competent *E. coli* cells (Invitrogen, California), following the manufacturer's guidelines. Plasmid DNA was extracted with the Wizard Plus SV Miniprep DNA purification system (Promega, Wisconsin) following the manufacturer's instructions. DNA quantification was carried out with a Qubit 2.0 fluorimeter and assay kit (Invitrogen,

California) and the number of gene copies was calculated based on this quantification. The relative abundance of EFF in the samples was calculated as the ratio of the values obtained from the qPCR assays (*spo0A*/ 16S rRNA gene copies).

3.2.4. Amplicon pyrosequencing and analysis

In order to study the diversity of the bacterial communities and more specifically the diversity of EFF, eight samples were selected for 454 pyrosequencing of the 16S rRNA gene using the services of Eurofins MWG Operon (Germany). Fragments of approximately 500 bp were retrieved using primers Eub8f (5'-AGAGTTTGATCCTGGCTCAG-3') and Eub519r (5'-GTATTACCGCGGCTGCTGG-3'), as previously described [32]. Raw sequence data was analyzed with QIIME [33], using the pipeline for *de novo* OTU picking and diversity analyses from 454 data suggested in QIIME tutorials. Amplicon sequencing resulted in a total of 117,542 sequence reads after quality filtering. Sequences were de-noised with the split_library.py function implemented in QIIME, and verified for chimeras using USEARCH version 6.1 with the reference database used by in the version 1.8.0 of QIIME. 2365 chimeric sequences were detected and removed from further analysis. To the rest of the trimmed and processed sequences, alignment was performed through the RDP website (https://rdp.cme.msu.edu/tutorials/aligner/RDPtutorial_ALIGNER.html) using Infernal Aligner [34]. OTUs were identified using a threshold of 97% sequence similarity with USEARCH version 6.1. Alpha diversity within the samples was calculated in rarefied subsets sequences to have equal sequence coverage (7302 sequences per sample) following the tutorial suggested by QIIME for 454-sequencing analysis. The parameters retained for the analysis were Richness, Shannon and Simpson diversity indices, and the percentage of the ratio OTUs/chao1 (coverage). The same analyses were performed after selecting solely the sequences assigned to the phylum Firmicutes. In this case, alpha diversity was calculated based on 1188 sequences (equal sequence coverage).

For the same set of eight samples, *spo0A* amplicon pyrosequencing was also performed using the services of Eurofins MWG Operon (Germany). A 602 bp sequence of the *spo0A* gene was amplified with the primers spoA166f and spoA748r, as previously described [28]. For quality filtering, the nucleotide sequences were translated to their amino acid sequences, based on ORF detection. The amino acid sequences were then aligned and compared to a Gribskov-style protein profile of Spo0A [35] that was built based on 27 known Spo0A sequences. Filtration was applied as a function of the profile score and profile alignment length, which separates noise or negatives hits from true positives *spo0A* sequences. The nucleotide sequences were clustered into operational taxonomic units (OTUs) at 97% sequence identity using Uclust [36]. The centroid (representative sequence) of each OTU was classified using MLgsc, a general sequence classifier adapted for protein and customized to Spo0A [37].

All metagenomic sequences were submitted to Sequence Read Archive of the National Center for Biotechnology Information (NCBI) under BioProject IDs PRJNA267761 and PRJNA276803.

3.2.5. Statistical analysis

Statistical analyses were performed using R, version 3.0.2 [38], Rstudio, version 0.98.1049, and BiodiversityR [39]. Correlations between diversity and environmental limiting factors were estimated using both Pearson's and Spearman's methods, however since our data are not normally distributed and also taken or transformed into ordinal scale, Spearman's tests were considered as more appropriate and therefore applied to this dataset. Since we focused on inferring the effect of each environmental parameter on the relative abundance (*spoOA*/ 16S rRNA gene counts), a generalized additive model (GAM) was used to represent graphically the dependence of the relative abundance to the environmental factors. Moreover, in order to determine the significance of the difference in the ratio obtained for three decision nodes pre-defined from the data (multiple, single and no limiting factor), the statistical significance was evaluated using a Kruskal-Wallis post-hoc tests according to Nemenyi using the PMCMR (Pairwise Multiple Comparison of Mean Ranks Package) library in R, after verifying this model's assumptions. This test was chosen, after performing a Shapiro-Wilk test (Multiple p-value = 5.376e-09, Single p-value = 3.102e-12, No Factors p-value = 1.116e-05), which showed no normal distribution invalidating a classical ANOVA approach. All scripts used for the statistical analysis are provided as supplementary material.

3.3. Results

3.3.1. Characterization of the natural springs

In this study, 71 sampling points in 21 locations were investigated (Figure 3.1). The samples were collected from geothermal springs and drillings (Chile: Salar de Aguas Calientes, Lirima wetland, El Tatio geysers; Greece: Eleftheres, Krinides, Lagadas, Milos, Nea Apollonia, Nigrita, Potamia, Pozar, Thermia, Traianoupoli; Germany: Gross Schoenebeck and Bruschal power plants; France: Soultz-sous-forets power plant; and Colombia: Los Volcanes) and mineral springs (Greece: Agia Paraskevi, Aggistro, Pikrolimni; Switzerland: Ponts-de-Martel).

The sites studied exhibited diverse environmental characteristics concerning temperature, depth (i.e. pressure), UV radiation, low and high pH, and mineral compounds present (Supplementary table 3.2). The samples were grouped into three categories before genomic measurements by transforming the quantitative measurements of limiting environmental factors into a qualitative limiting index. For the transformation we used as reference physicochemical parameters and ranges considered to describe mesophilic conditions (Supplementary table 3.1). This included temperature (20 to 55°C), pH (5.5 to 8.5), atmospheric pressure (~1 atm), exposure to UV radiation and concentration of cell-toxic chemical compounds. This categorization is far from perfect as the relative importance of individual factors might be different for different species and this is ignored when giving the same weight to each factor studied here. However, this scoring was selected because a large number of our samples are far from this mesophilic range, and thus using another approach (e.g. quintiles or percentiles) would result in the underestimation of limiting conditions in a traditional sense. Three categories of limiting

environmental factors (multiple, single or none) were defined. In total, 20, 30 and 21 samples from multiple, single and non-limiting environments were obtained.

A large variation in DNA yield was observed based on the initial sample type (i.e. soil, sand, mud, sediment, biofilm, microbial mat, water; Supplementary figure 3.1). This was particularly noticeable in water samples that had significantly lower biomass than soils, sediments, biofilms or microbial mats. Total bacterial abundance was determined by quantifying the 16S rRNA gene copy numbers, while EFF abundance was measured using the specific *spoOA* transcriptional factor gene. The 16S rRNA gene is found in multiple copies per bacterial cell [40], while *spoOA* gene is a single-copy gene [31, 41]. However, for comparison between samples, no normalization based on gene copy numbers was made and the relative abundance of EFF was calculated based only on the qPCR counts. Relative abundance (*spoOA* copy numbers/16S rRNA gene copy numbers) ranged from <0.0001 % to 100 %, with an average abundance of 6.79 % (Supplementary table 3.3).

3.3.2. Single environmental factors do not influence EFF relative abundance

Temperature, pH and UV radiation are among factors that have been suggested to influence the abundance and diversity of microbial communities in extreme sites [42,43]. Considering the variation in the number of samples within different categories (i.e. larger number of thermophilic versus mesophilic environments), a generalized additive model (GAM) was used to analyze the role of temperature and pH on relative abundance of EFF (Supplementary figure 3.2). No significant correlation between relative EFF abundance and temperature or pH was obtained (Temperature R^2=-0.233, *p*-value=0.272; pH R^2= -0.144, *p*-value=0.232).

3.3.3. EFF are prevalent in multiple-limiting environmental factors

We next analyzed if the co-existence of multiple limiting factors could affect the relative abundance of EFF. Based on the multiple-single-null grouping, EFF had a statistically significant higher relative abundance (median ratio =20.43%) in sites with multiple limiting environmental factors, compared to environments with single (median ratio =1.56 %) or null (ratio median=1.27%) factor (Figure 3.2). Post-hoc tests showed that the difference between "multiple" and "single" factor groups was statistically significant (*p*-value=0.0007), and so was between "multiple" and "null" groups (*p*-value=0.006).

3.3.4. Diversity of EFF in environmental samples

In order to evaluate the effect of limiting environmental factors on diversity, we analyzed the bacterial community composition of eight samples representing the three categories shown in Figure 3.2. Firstly, we used data obtained from 16S rRNA gene amplicon pyrosequencing (Table 3.1). A total of 117,542 high quality sequence reads were obtained with 8,050–29,335 sequences per sample (mean 14,693). Sequence clustering at 97% identity resulted in 17,596 operational taxonomic units (OTUs) in the data set. Overall, richness in samples from sites with varying limiting environmental factors (none to multiple) remained stable and did not decrease with an increase in limiting conditions. In

the analysis of the community composition, Firmicutes represented a significant fraction of the community in some samples (Supplementary figure 3.3) regardless of the environmental conditions of the site. However, it is important to indicate that the identification of endospore-formers based on the 16S rRNA gene is not entirely possible because of the patchiness of the distribution of sporulation as a trait within related Firmicutes clades.

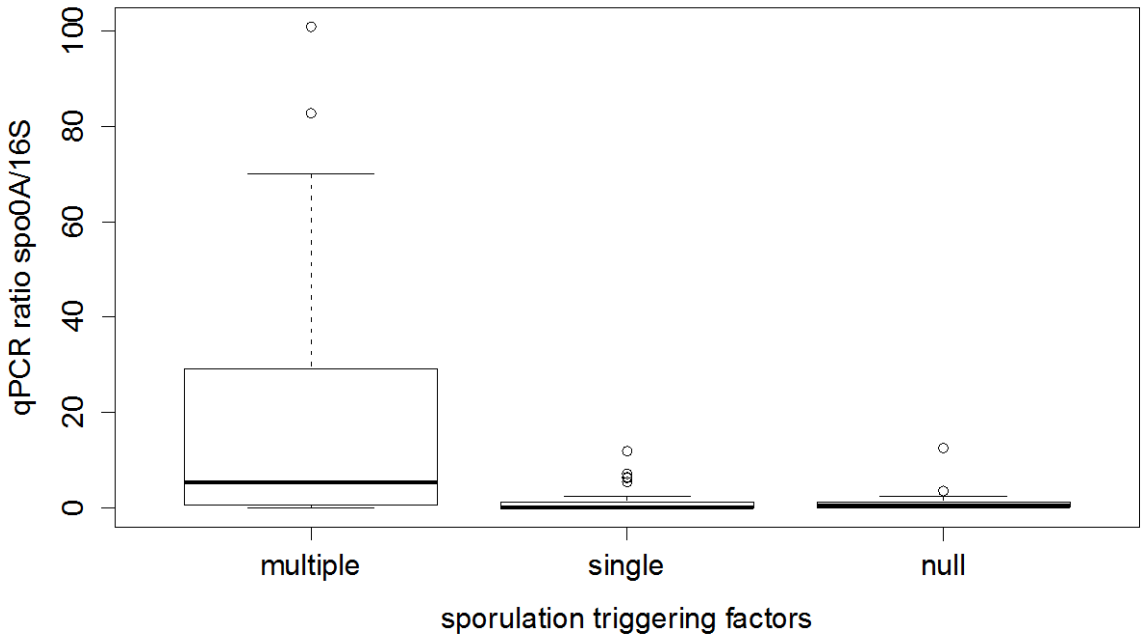


Figure 3.2 Effect of multiple, single and no limiting factors on relative abundance of EFF. The boxplots represent the qPCR ratios of *spo0A* gene/16S rRNA gene, grouped by decision node.

In spite of this, Firmicutes richness remained also stable regardless of the presence of limiting factors. This was equally true for Shannon and Simpson indices. Considering community, the sample with the highest proportion of Firmicutes (4NAP), which is influenced by temperature, uranium and alkaline pH, Firmicutes represented 37.3% of the community, followed by Proteobacteria (12.69%) and Bacteroidetes (6.86%). In the samples with single or no limiting factor, Firmicutes represented less than 10% of the community.

The endospore-forming group of Firmicutes was also studied. Amplicon sequencing of the *spo0A* gene resulted in a total of 14,362quality reads with an average length of 491 bp. These reads were clustered into 3392 OTUs, applying a 97% identity threshold. The OTUs were assigned to eight different genera (Supplementary table 3.4). In the community composition analysis, three genera were common to all the samples: *Clostridium*, *Bacillus* and *Anaerostipes. Clostridium* was the most prevalent genus, with the exception of 25KAM6 that was dominated by *Bacillus* (Figure 3.3).

3.4. Discussion

It has been suggested that because of their dispersal potential and diverse metabolic capabilities, endospore-forming Firmicutes (EFF) are one of the most ubiquitous microbial

groups [7, 8]. However, there is no conclusive experimental evidence of this. On the contrary, many molecular ecology diversity surveys have failed to detect this group [23, 24]. Our results offer, for the first time, experimental evidence regarding the ubiquity of this group of Firmicutes in mineral springs. We detected a variable abundance of EFF in the majority of environments studied. However, for most of the sites, EFF represent only a small fraction of the total bacteria observed in the samples, probably explaining the difficulties of detecting this rare component of the microbial community.

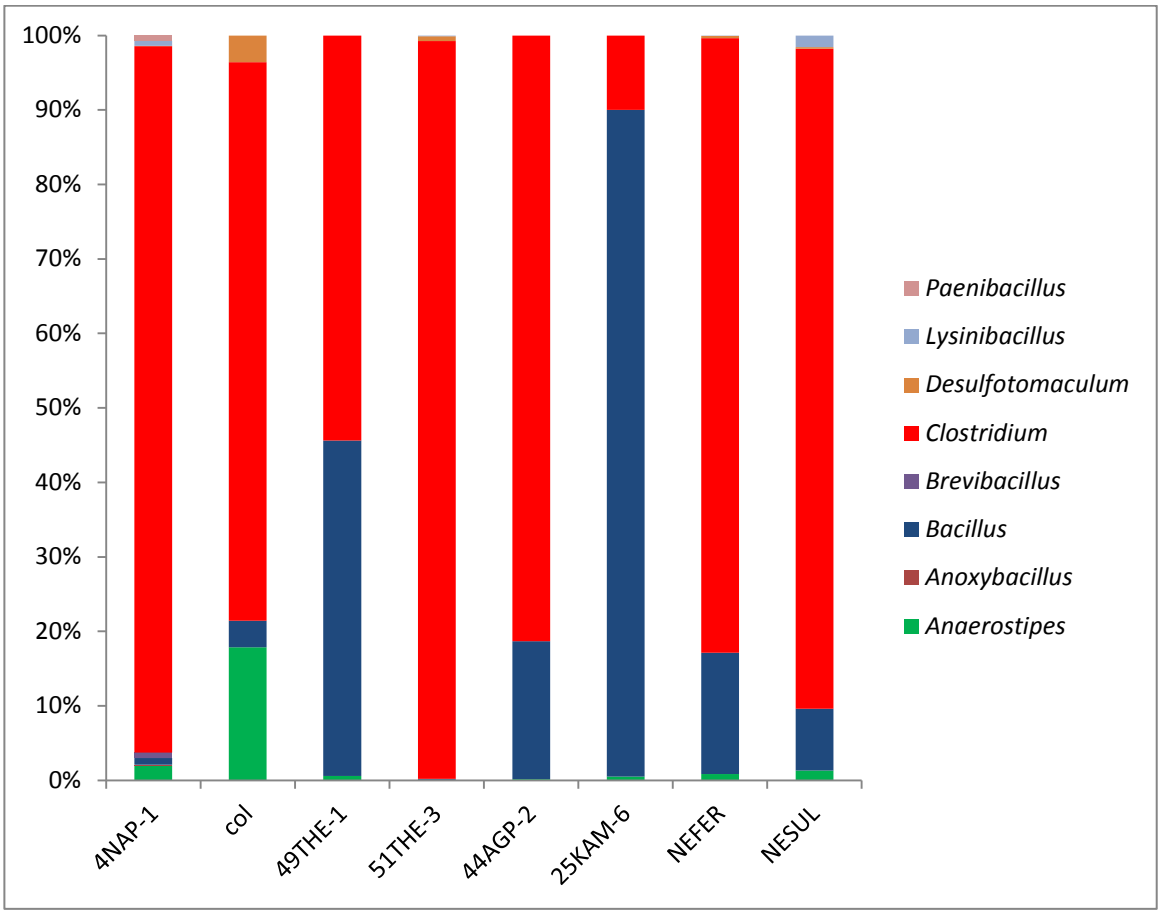


Figure 3.3 Community composition based on *spoA* gene sequencing data. The diversity of the endospore-forming Firmicutes community per sampling site. OTUs detected were classified to known endospore-forming genera.

An ecological explanation for the poor representation of EFF relative to the total bacterial community can be found when analyzing the tradeoffs of the survival strategies deployed by EFF. This is particularly clear for endospore-formation, a notably energy-demanding process [44] that leads to the formation of a resistance structure. Although the energy requirements of other survival strategies deployed by EFF are not known, one can consider that the biological cost of their myriad of survival strategies limits the distribution of this group under non-limiting conditions. This was the hypothesis tested in this study. However, the definition of a factor that limits microbial growth is not an easy task as *extreme is in the eyes of the beholder* [45]. Upper limits to life have been suggested [46, 47], as well as upper limits of habitable ecosystems [48]. In geothermal

and mineral springs a combination of steep physicochemical gradients might offer EFF an advantageous niche. Hotspots of geothermal springs exhibit temperatures exceeding 100°C and are often highly acidic [43]. However, a gradient of temperature, pH and dissolved minerals is formed [49] creating a gradient of habitats and niches because of the different temperature fitness of various taxa.

The results obtained here support our hypothesis suggesting that limiting environmental conditions favor the relative abundance of EFF. However, surprisingly, this was true only for the co-existence of various limiting factors. Theoretically, a single limiting factor should suffice to reduce total bacterial abundance and consequently lead to a relative enrichment of endospore-formers. However, our data show that there is no significant difference between environments with no limiting factors and others where a single limiting factor is present. This indicates that even though each individual limiting factor studied here is reported to reduce microbial abundance in general [10, 18, 50–54], this by itself does not increase the relative abundance of EFF. Our data also demonstrate that abundance, species richness and diversity do not depend on the limiting factor. It is the co-existence of limiting factors that drives the increase in prevalence of EFF in the environments studied here.

Studies analyzing the role of environmental factors in distribution patterns of microbial communities in environments with limiting environmental conditions have started to emerge. For instance, in geothermal environments, recent publications have shown that temperature [42], and to a lesser extent pH [43], dictate the prevalence of specific bacterial and archaeal groups. In other environments, such as salt flats (salars), salinity is believed to control microbial distribution [55]. These environmental factors determine the distribution of individuals, but they can also explain the distribution of a population or even a community [56–58]. Although one could argue that not all environmental factors determine to the same extent the ecological niche of a species -which is most likely the case in nature-, a general theoretical unimodal distribution model with maximum abundance towards the middle range of individual environmental factors has been predicted based on Shelford's law of tolerance, according to which "[an organism] is absent or found in minimal numbers only [...] should a [environmental] condition vary outside the limits tolerated by the animal" [59]. So far, the same model has been applied for diversity. It has often been discussed that in the case of microbial communities, abundance and diversity also decrease towards extremity [55, 60] and thus, it can be assumed that the distribution of microbial communities follows Shelford's law. This has been supported by patterns of species distribution across altitudinal gradients for different taxa [61, 62]. Based on our results, EFF abundance and diversity do not follow this theoretical model, at least in the case of geothermal and mineral springs. In fact, if strategies such as dormancy are in place, limiting environmental conditions may play a subtler role on community structure because total community composition may differ largely from the active populations. It has been proposed that EFF persist in the environment primarily in a spore state rather than in as vegetative cells [16]. This would explain detecting the psychrophilic species *Clostridium bowmanii* [63] in the 16S rRNA

gene data of a hot water spring. Likewise, the strictly anaerobic *Clostridium* spp. dominated community composition in all the samples based on *spo0A* sequencing, although only a small fraction of the samples corresponded to anaerobic environments. Although additional experiments to measure the numbers of spores in each environment are needed to evaluate the role of sporulation, the results obtained suggest that in the environment, multiple factors would be operationally significant to act as sporulation triggers. In fact, the environmental triggers of sporulation are still unclear, and even among closely related strains, there is no unique trigger [64–67].

In contrast, other groups appear to reflect the limiting conditions of the environment and the selection for active vegetative communities. For example, in a hot water spring contaminated by uranium, sequences of the uranium-reducing *Desulfosporosinus* were prevalent [68, 69]. Likewise, in a sulfuric hot spring, the genus *Alicyclobacillus* composed of sulfur-oxidizing thermotolerant bacteria was prevalent [70, 71]. Finally, in springs used for bathing purposes (Colombia, Greece), species belonging to *Lachnospiraceae* were found in abundance, an observation that agrees with the human origin of this family [72].

In conclusion, the results of relative abundance and diversity obtained here suggest that commonly used patterns of bacterial distribution cannot be applied to EFF. More precisely, for EFF "everything is everywhere but the environment selects" is an over-simplistic explanation of their universality and that does not consider the complex role of spores, enhanced geographical dispersal, or functional and metabolic diversity and adaptations of vegetative cells. In addition, the fact that we observed an enrichment of EFF with multiple limiting factors but not with single ones suggests that the effects of the latter are not additive, but rather multiplicative. This fact might have a profound effect in our ability to predict EFF distribution in natural environments. Finally, the complexity of EFF ecology, fascinating though it can be *per se*, is significant to the study of the ecology of other microbial groups too, as it can provide insights on the application of ecological theory to the microbial world.

3.5. Acknowledgements

For the sampling campaigns we would like to thank Dr. Simona Regenspurg, Dr. Detlev Rettenmaier, Dr. Albert Genter, Dr. Ludovic Roussel-Delif, Dr. Theodoros Coutles, Prof. Alexandra Ioannidou, Prof. Georgios Vergemezis, and Mrs. Anna Voulgari. For the sampling campaign in Chile, we would wish to thank CONAF for access to Salar de Huasco. For the analysis concerning the presence of uranium, we thank Prof. Ioannis Paschalidis. For assistance in the statistical analysis, we would like to thank Dr. Matthieu Bueche and Dr. Radu Alexandru Slobodeanu. We acknowledge funding from the Swiss National Funding projects 31003A_132358/1 and 31003A_152972, from Fondation Pierre Mercier pour la science and from REGARD for equality of women in science. The genetic information downloaded from GenBank may be considered to be part of the genetic patrimony of Switzerland, Greece, Germany, France, Colombia, and Chile respectively, the countries from which the samples were obtained. Users of this information agree to: 1) acknowledge Switzerland, Greece, Germany, France, Colombia, and Chile as the countries

of origin in any country where the genetic information is presented and 2) contact the CBD focal point and the ABS focal point identified in the CBD website http://www.cbd.int/information/nfp.shtml if they intend to use the genetic information for commercial purposes. PJ designed the research. SF and TW held the experiments and statistical analysis. TJ performed bioinformatics analysis. SF, TW, NJ, CD, VM, and PJ held the sampling. SF and PJ wrote the manuscript. All authors contributed to the interpretation of data and to the critical revision of the manuscript. The research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. Individuals responsible for funding this work did not participate in study design, data collection and interpretation, or the decision to submit this work for publication. Thus, the authors declare no conflict of interest.

3.6. References

[1] Blevins SM, Bronze MS. Robert Koch and the "golden age" of bacteriology. Int J Infect Dis 2010;14:e744–51. doi:10.1016/j.ijid.2009.12.003.

[2] Driks A. Maximum shields: the assembly and function of the bacterial spore coat. Trends Microbiol 2002;10:251–4. doi:10.1016/S0966-842X(02)02373-9.

[3] Driks A. Overview: Development in bacteria: spore formation in *Bacillus subtilis*. Cell Mol Life Sci CMLS 2002;59:389–91.

[4] Onyenwoke RU, Brill JA, Farahi K, Wiegel J. Sporulation genes in members of the low G+C Gram-type-positive phylogenetic branch (Firmicutes). Arch Microbiol 2004;182:182–92. doi:10.1007/s00203-004-0696-y.

[5] Yudkin MD, Clarkson J. Differential gene expression in genetically identical sister cells: the initiation of sporulation in *Bacillus subtilis*. Mol Microbiol 2005;56:578–89. doi:10.1111/j.1365-2958.2005.04594.x.

[6] Abecasis AB, Serrano M, Alves R, Quintais L, Pereira-Leal JB, Henriques AO. A genomic signature and the identification of new sporulation genes. J Bacteriol 2013:JB.02110–2. doi:10.1128/JB.02110-12.

[7] Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, et al. Microbial biogeography: putting microorganisms on the map. Nat Rev Microbiol 2006;4:102–12. doi:10.1038/nrmicro1341.

[8] Roberts MS, Cohan FM. Recombination and Migration Rates in Natural Populations of *Bacillus subtilis* and *Bacillus mojavensis*. Evolution 1995;49:1081–94. doi:10.2307/2410433.

[9] Stephens C. Bacterial sporulation: A question of commitment? Curr Biol 1998;8:R45–8. doi:10.1016/S0960-9822(98)70031-4.

[10] Errington J. *Bacillus subtilis* sporulation: regulation of gene expression and control of morphogenesis. Microbiol Rev 1993;57:1–33.

[11] Mandic-Mulec I, Prosser JI. Diversity of Endospore-forming Bacteria in Soil: Characterization and Driving Mechanisms. In: Logan NA, Vos P, editors. Endospore-Form. Soil Bact., Springer Berlin Heidelberg; 2011, p. 31–59.

[12] Galperin MY. Genome Diversity of Spore-Forming Firmicutes. Microbiol Spectr 2013;1:TBS – 0015–2012. doi:10.1128/microbiolspectrum.TBS-0015-2012.

[13] Logan NA. *Bacillus* and relatives in foodborne illness. J Appl Microbiol 2012;112:417–29. doi:10.1111/j.1365-2672.2011.05204.x.

[14] Béjà O, Suzuki MT, Heidelberg JF, Nelson WC, Preston CM, Hamada T, et al. Unsuspected diversity among marine aerobic anoxygenic phototrophs. Nature 2002; 415:630–3. doi:10.1038/415630a.

[15] Jandhyala SM, Talukdar R, Subramanyam C, Vuyyuru H, Sasikala M, Reddy DN. Role of the normal gut microbiota. World J Gastroenterol WJG 2015;21:8787–803. doi:10.3748/wjg.v21.i29.8787.

[16] Zeigler DR. The *Geobacillus* paradox: why is a thermophilic bacterial genus so prevalent on a mesophilic planet? Microbiology 2013;160:1–11. doi:10.1099/mic.0.071696-0.

[17] Fajardo-Cavazos P, Nicholson W. *Bacillus* Endospores Isolated from Granite: Close Molecular Relationships to Globally Distributed *Bacillus* spp. from Endolithic and Extreme Environments. Appl Environ Microbiol 2006;72:2856–63. doi:10.1128/AEM.72.4.2856-2863.2006.

[18] Nicholson WL, Fajardo-Cavazos P, Rebeil R, Slieman TA, Riesenman PJ, Law JF, et al. Bacterial endospores and their significance in stress resistance. Antonie Van Leeuwenhoek 2002;81:27–32. doi:10.1023/A:1020561122764.

[19] Hubert C, Loy A, Nickel M, Arnosti C, Baranyi C, Bruechert V, et al. A Constant Flux of Diverse Thermophilic Bacteria into the Cold Arctic Seabed. Science 2009;325:1541–4. doi:10.1126/science.1174012.

[20] Hubert C, Arnosti C, Brüchert V, Loy A, Vandieken V, Jørgensen BB. Thermophilic anaerobes in Arctic marine sediments induced to mineralize complex organic matter at high temperature. Environ Microbiol 2010;12:1089–104. doi:10.1111/j.1462-2920.2010.02161.x.

[21] Atanassova M, Derekova A, Mandeva R, Sjøholm C, Kambourova M. *Anoxybacillus bogrovensis* sp. nov., a novel thermophilic bacterium isolated from a hot spring in Dolni Bogrov, Bulgaria. Int J Syst Evol Microbiol 2008;58:2359–62. doi:10.1099/ijs.0.65745-0.

[22] Derekova A, Sjoholm C, Mandeva R, Kambourova M. *Anoxybacillus rupiensis* sp nov., a novel thermophilic bacterium isolated from Rupi basin (Bulgaria). Extremophiles 2007;11:577–83. doi:10.1007/s00792-007-0071-4.

[23] Filippidou S, Junier T, Wunderlin T, Lo C-C, Li P-E, Chain PS, et al. Under-detection of endospore-forming Firmicutes in metagenomic data. Comput Struct Biotechnol J 2015;13:299–306. doi:10.1016/j.csbj.2015.04.002.

[24] Mering C von, Hugenholtz P, Raes J, Tringe SG, Doerks T, Jensen LJ, et al. Quantitative Phylogenetic Assessment of Microbial Communities in Diverse Environments. Science 2007;315:1126–30. doi:10.1126/science.1133420.

[25] Wunderlin T, Junier T, Roussel-Delif L, Jeanneret N, Junier P. Stage 0 sporulation gene A as a molecular marker to study diversity of endospore-forming Firmicutes. Environ Microbiol Rep 2013;5:911–24. doi:10.1111/1758-2229.12094.

[26] Fytikas M, Innocenti F, Kolios N, Manetti P, Mazzuoli R, Poli G, et al. Volcanology and petrology of volcanic products from the island of Milos and neighbouring islets. J Volcanol Geotherm Res 1986;28:297–317. doi:10.1016/0377-0273(86)90028-4.

[27] Teschner M, Faber E, Poggenburg J, Vougioukalakis GE, Hatziyannis G. Continuous, direct gas-geochemical monitoring in hydrothermal vents: Installation and long-term operation on Nisyros Island (Greece). Pure Appl Geophys 2007;164:2549–71. doi:10.1007/s00024-007-0285-5.

[28] Wunderlin T, Junier T, Roussel-Delif L, Jeanneret N, Junier P. Stage 0 sporulation gene A as a molecular marker to study diversity of endospore-forming Firmicutes. Environ Microbiol Rep 2013:n/a – n/a. doi:10.1111/1758-2229.12094.

[29] Muyzer G, Teske A, Wirsen C, Jannasch H. Phylogenetic-Relationships of *Thiomicrospira* Species and their Identification in Deep-Sea Hydrothermal Vent Samples

by Denaturing Gradient Gel-Electrophoresis of 16S rDNA Fragments. Arch Microbiol 1995;164:165–72. doi:10.1007/BF02529967.

[30] Bakke I, De Schryver P, Boon N, Vadstein O. PCR-based community structure studies of bacteria associated with eukaryotic organisms: a simple PCR strategy to avoid co-amplification of eukaryotic DNA. J Microbiol Methods 2011;84:349–51. doi:10.1016/j.mimet.2010.12.015.

[31] Bueche M, Wunderlin T, Roussel-Delif L, Junier T, Sauvain L, Jeanneret N, et al. Quantification of endospore-forming firmicutes by quantitative PCR with the functional gene *spo0A*. Appl Environ Microbiol 2013;79:5302–12. doi:10.1128/AEM.01376-13.

[32] Li H, Zhang Y, Li D, Xu H, Chen G, Zhang C. Comparisons of different hypervariable regions of rrs genes for fingerprinting of microbial communities in paddy soils. Soil Biol Biochem 2009; 41:954–68. doi:10.1016/j.soilbio.2008.10.030.

[33] Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010;7:335–6. doi:10.1038/nmeth.f.303.

[34] Nawrocki EP, Eddy SR. Query-dependent banding (QDB) for faster RNA similarity searches. PLoS Comput Biol 2007;3:e56. doi:10.1371/journal.pcbi.0030056.

[35] Gribskov M, McLachlan AD, Eisenberg D. Profile analysis: detection of distantly related proteins. Proc Natl Acad Sci U S A 1987;84:4355–8.

[36] Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinforma Oxf Engl 2010;26:2460–1. doi:10.1093/bioinformatics/btq461.

[37] Junier T, Hervé V, Wunderlin T, Junier P. MLgsc: A Maximum-Likelihood General Sequence Classifier. PloS One 2015; 10:e0129384. doi:10.1371/journal.pone.0129384.

[38] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: 2013.

[39] Kindt R, Coe R. Tree diversity analysis. A manual and software for common statistical methods for ecological and biodiversity studies. Nairobi (Kenya): World Agroforestry Centre (ICRAF); 2005.

[40] Farrelly V, Rainey FA, Stackebrandt E. Effect of genome size and rrn gene copy number on PCR amplification of 16S rRNA genes from a mixture of bacterial species. Appl Environ Microbiol 1995; 61:2798–801.

[41] Galperin MY, Mekhedov SL, Puigbo P, Smirnov S, Wolf YI, Rigden DJ. Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. Environ Microbiol 2012;14:2870–90. doi:10.1111/j.1462-2920.2012.02841.x.

[42] Sharp CE, Brady AL, Sharp GH, Grasby SE, Stott MB, Dunfield PF. Humboldt's spa: microbial diversity is controlled by temperature in geothermal environments. ISME J 2014. doi:10.1038/ismej.2013.237.

[43] Hou W, Wang S, Dong H, Jiang H, Briggs BR, Peacock JP, et al. A Comprehensive Census of Microbial Diversity in Hot Springs of Tengchong, Yunnan Province China Using 16S rRNA Gene Pyrosequencing. PLoS ONE 2013;8:e53350. doi:10.1371/journal.pone.0053350.

[44] Höfler C, Heckmann J, Fritsch A, Popp P, Gebhard S, Fritz G, et al. Cannibalism stress response in Bacillus subtilis. Microbiol Read Engl 2016;162:164–76. doi:10.1099/mic.0.000176.

[45] Rothschild LJ, Mancinelli RL. Life in extreme environments. Nature 2001;409:1092–101. doi:10.1038/35059215.

[46] Cockell C s., Bush T, Bryce C, Direito S, Fox-Powell M, Harrison J p., et al. Habitability: A Review. Astrobiology 2016;16:89–117. doi:10.1089/ast.2015.1295.

[47] Harrison JP, Gheeraert N, Tsigelnitskiy D, Cockell CS. The limits for life under multiple extremes. Trends Microbiol 2013;21:204–12. doi:10.1016/j.tim.2013.01.006.

[48] Cockell CS. Astrobiology: Understanding Life in the Universe. John Wiley & Sons; 2015.

[49] Weltzer ML, Miller SR. Division-specific differences in bacterial community assembly of two alkaline hot spring ecosystems from Yellowstone National Park. Microb Ecol 2013;65:537–40. doi:10.1007/s00248-013-0214-3.

[50] Müller AL, de Rezende JR, Hubert CRJ, Kjeldsen KU, Lagkouvardos I, Berry D, et al. Endospores of thermophilic bacteria as tracers of microbial dispersal by ocean currents. ISME J 2014;8:1153–65. doi:10.1038/ismej.2013.225.

[51] Sanchez-Salas J-L, Setlow B, Zhang P, Li Y, Setlow P. Maturation of Released Spores Is Necessary for Acquisition of Full Spore Heat Resistance during *Bacillus subtilis* Sporulation. Appl Environ Microbiol 2011; 77:6746–54. doi:10.1128/AEM.05031-11.

[52] Setlow B, Loshon C a., Genest P c., Cowan A, Setlow C, Setlow P. Mechanisms of killing spores of *Bacillus subtilis* by acid, alkali and ethanol. J Appl Microbiol 2002;92:362–75. doi:10.1046/j.1365-2672.2002.01540.x.

[53] McDougald D, Gong L, Srinivasan S, Hild E, Thompson L, Takayama K, et al. Defences against oxidative stress during starvation in bacteria. Antonie Van Leeuwenhoek 2002;81:3–13. doi:10.1023/A:1020540503200.

[54] Losick R, Stragier P. Crisscross regulation of cell-type-specific gene expression during development in *B. subtilis*. Nature 1992;355:601–4. doi:10.1038/355601a0.

[55] Oren A. The bioenergetic basis for the decrease in metabolic diversity at increasing salt concentrations: implications for the functioning of salt lake ecosystems. Hydrobiologia 2001; 466:61–72. doi:10.1023/A:1014557116838.

[56] Pörtner HO. Climate variations and the physiological basis of temperature dependent biogeography: systemic to molecular hierarchy of thermal tolerance in animals. Comp Biochem Physiol A Mol Integr Physiol 2002;132:739–61.

[57] Calow P, Sibly RM. A Physiological Basis of Population Processes: Ecotoxicological Implications. Funct Ecol 1990;4:283–8. doi:10.2307/2389587.

[58] Woodward FI. Climate and Plant Distribution. Cambridge University Press; 1987.

[59] Shelford VE. Some Concepts of Bioecology. Ecology 1931; 12:455–67. doi:10.2307/1928991.

[60] Fierer N, Jackson RB. The diversity and biogeography of soil bacterial communities. Proc Natl Acad Sci U S A 2006;103:626–31. doi:10.1073/pnas.0507535103.

[61] Bryant JA, Lamanna C, Morlon H, Kerkhoff AJ, Enquist BJ, Green JL. Microbes on mountainsides: Contrasting elevational patterns of bacterial and plant diversity. Proc Natl Acad Sci U S A 2008;105:11505–11. doi:10.1073/pnas.0801920105.

[62] Lozupone CA, Knight R. Global patterns in bacterial diversity. Proc Natl Acad Sci 2007;104:11436–40. doi:10.1073/pnas.0611525104.

[63] Spring S, Merkhoffer B, Weiss N, Kroppenstedt RM, Hippe H, Stackebrandt E. Characterization of novel psychrophilic clostridia from an Antarctic microbial mat: description of *Clostridium frigoris* sp. nov., *Clostridium lacusfryxellense* sp. nov., *Clostridium bowmanii* sp. nov. and *Clostridium psychrophilum* sp. nov. and reclassification of *Clostridium laramiense* as *Clostridium estertheticum* subsp. *laramiense* subsp. nov. Int J Syst Evol Microbiol 2003;53:1019–29. doi:10.1099/ijs.0.02554-0.

[64] Duc MTL, Dekas A, Osman S, Moissl C, Newcombe D, Venkateswaran K. Isolation and Characterization of Bacteria Capable of Tolerating the Extreme Conditions of Clean Room Environments. Appl Environ Microbiol 2007;73:2600–11. doi:10.1128/AEM.03007-06.

[65] Mohapatra BR, Duc MTL. Detecting the dormant: a review of recent advances in molecular techniques for assessing the viability of bacterial endospores. Appl Microbiol Biotechnol 2013;97:7963–75. doi:10.1007/s00253-013-5115-3.

[66] Mohapatra BR, La Duc MT. Rapid detection of viable *Bacillus pumilus* SAFR-032 encapsulated spores using novel propidium monoazide-linked fluorescence *in situ* hybridization. J Microbiol Methods 2012;90:15–9. doi:10.1016/j.mimet.2012.04.006.

[67] Setlow P. Spores of *Bacillus subtilis*: their resistance to and killing by radiation, heat and chemicals. J Appl Microbiol 2006;101:514–25. doi:10.1111/j.1365-2672.2005.02736.x.

[68] Suzuki Y, Kelly SD, Kemner KM, Banfield JF. Enzymatic U(VI) reduction by Desulfosporosinus species. Radiochim ActaInternational J Chem Asp Nucl Sci Technol 2009;92:11–6. doi:10.1524/ract.92.1.11.25404.

[69] Suzuki Y, Kelly SD, Kemner KM, Banfield JF. Radionuclide contamination: Nanometre-size products of uranium bioreduction. Nature 2002;419:134–134. doi:10.1038/419134a.

[70] Guo X, You X-Y, Liu L-J, Zhang J-Y, Liu S-J, Jiang C-Y. *Alicyclobacillus aeris* sp. nov., a novel ferrous- and sulfur-oxidizing bacterium isolated from a copper mine. Int J Syst Evol Microbiol 2009;59:2415–20. doi:10.1099/ijs.0.008870-0.

[71] Mavromatis K, Sikorski J, Lapidus A, Glavina Del Rio T, Copeland A, Tice H, et al. Complete genome sequence of *Alicyclobacillus acidocaldarius* type strain (104-IAT). Stand Genomic Sci 2010;2:9–18. doi:10.4056/sigs.591104.

[72] Meehan CJ, Beiko RG. A Phylogenomic View of Ecological Specialization in the Lachnospiraceae, a Family of Digestive Tract-associated Bacteria. Genome Biol Evol 2014:evu050. doi:10.1093/gbe/evu050.

3.7. Supplementary material

Supplementary Table 3.1. Samples selected for diversity analysis. In the Q-reads column the number in parenthesis corresponds to chimeras removed from further analysis. Values for the overall community correspond to samples rarefied to 7302 reads. Values for Firmicutes sub-community correspond to samples rarefied to 994 reads. (N/A= Not available)

Code	Location	Source	Decision node	*spoA*/16S rRNA Ratio %	Q-Reads	Shannon index total	Shannon index Firmicutes	Simpson index total	Simpson index Firmicutes	Richness total - obs. OTUs (not rarefied)	OTUs/ Chao1 % total	Richness Firmicutes (not rarefied)	OTUs/ Chao1 % Firmicutes
4nap-1	Nea Apollonia (GR)	Biofilm from drilling pipe	Multiple	100	17348 (436)	9.700	7.472	0.996	0.985	2490 (4452)	0.332	403 (1661)	0.284
Col	Los volcanos (CO)	Biofilm	Single	11.90	21886 (368)	5.600	7.264	0.889	0.985	900 (1877)	0.317	357 (1160)	0.302
49the-1	Thermia (GR)	Sediment	Single	1.18	11709 (121)	7.980	N/A	0.983	N/A	1353 (1770)	0.466	N/A (86)	N/A
51the-3	Thermia (GR)	Biofilm	Single	7.15	29335 (550)	6.524	7.057	0.932	0.969	1216 (3020)	0.351	400 (1290)	0.264
44agp-2	Agia Paraskevi (GR)	Precipitate	No factor	0.06	10616 (14)	6.647	N/A	0.955	N/A	888 (1122)	0.380	N/A (97)	N/A
25kam-6	Kanava, Milos (GR)	Marine water	No factor	0.41	9201 (31)	4.822	N/A	0.863	N/A	521 (596)	0.420	N/A (1)	N/A
NeFer	Ponts-de-Martel, iron (CH)	Water and sediment	No factor	0.11	9397 (102)	9.781	5.895	0.995	0.917	2493 (2912)	0.410	264 (287)	0.443
Nesul	Ponts-de-Martel, sulfur (CH)	Water and sediment	No factor	0.03	8050 (97)	7.976	N/A	0.978	N/A	1742 (1847)	0.374	N/A (194)	N/A

Supplementary Table 3.2. Theoretical information on abiotic environmental factors that influence niche differentiation. Ranges and description of limiting conditions are restricted to Bacteria and Archaea. The role of those factors in sporulation of endospore-forming Firmicutes is also indicated. *In the column "Habitat" only those used for sampling in this study are mentioned. Numerous other environments can have the below mentioned characteristics but are not indicated here.

Factor	Subdivision	Definition	Range	Sporulation	Habitat*
Temperature	Psychrophilic	As a result of Earth's tilt towards the sun and altitude (Molles, 2005)	-20 to 12°C (Madigan *et al.*, 2010; Barton & Northup, 2011)	Yes (Müller *et al.*, 2014)	Lake bottom and sediments
	Thermophilic	"As a result of volcanic activity or movement of the Earth's crust at tectonically active sites" (Edwards, 1990)	55 to 130°C (Edwards, 1990; Kashefi & Lovley, 2003)	Yes (Nicholson *et al.*, 2002; Sanchez-Salas *et al.*, 2011)	Hot springs, Geothermal drillings
pH	Acidic	"Acid environments that provide the thermodynamic potential for cell maintenance and growth" (Edwards, 1990)	1 to 4 (Edwards, 1990; Madigan *et al.*, 2010)	Yes (Setlow *et al.*, 2002)	Hot springs, Mineral springs, Geothermal drillings
	Alkaline	"Proton concentration low, but also low availability of other metabolites and ions that precipitate (insoluble)" (Edwards, 1990)	8.5 to 12 (Madigan *et al.*, 2010)	Yes (Setlow *et al.*, 2002)	Hot springs, Mineral springs, Salars
Light[&]	UV light	Fluxes of solar ultraviolet radiation may be a pervasive and potentially damaging influence for both aquatic and terrestrial ecosystems (e.g. Karentz, 1994; Zagarese and Williamson, 1994; Diffey, 1991)	(UVR, 280-400nm)	Yes (Sanchez-Salas *et al.*, 2011; Nicholson *et al.*, 2000)	Atakama Dessert
Salinity	Halophilic	As a result of high rates of evaporation resulting in desiccation (Edwards, 1990)	6-32% (w/v) (Madigan *et al.*, 2010)	Yes (Nicholson et al, 2000), for desiccation)	Salars
Pressure	barophilic	high hydrostatic pressures encountered in deep sea and reservoir habitats (Barton & Northup, 2011)	500 to 1000 atm (Madigan *et al.*, 2010)	Higher than 200 atm (Nicholson *et al.*, 2000)	Geothermal drillings

Supplementary table 3.3. Sampling sites, environmental factors and extremity index

Location	Sample ID	Temperature	pH	temperature	pH	Radiation	Extremity
Aggistro	13agg1	36.46	7	0	0	0	0
Aggistro	14agg2	37	8.08	0	0	0	0
Agia Paraskevi	43AGP-1	33.59	7.6	0	0	0	0
Agia Paraskevi	44AGP-2	33.59	7.6	0	0	0	0
Agia Paraskevi	45AGP-3	35.07	7.6	0	0	1	1
Agia Paraskevi	46AGP-4	35.07	7.6	0	0	1	1
Agia Paraskevi	47AGP-5	35.07	7.6	0	0	1	1
Agia Paraskevi	48AGP-6	35.07	7.6	0	0	1	1
Aguas Calientes	Ac3b	22	6	0	0	0	0
Aguas Calientes	Ac5b	25	7.58	0	0	0	0
Aguas Calientes	Ac6b	47	7.47	0	0	0	0
Aguas Calientes	Ac3a	22	6	0	0	1	1
Aguas Calientes	Ac5a	25	7.58	0	0	1	1
Aguas Calientes	Ac6a	47	7.47	0	0	1	1
Aguas Calientes	Ac4b	20	5	0	1	0	1
Aguas Calientes	Ac4a	20	5	0	1	1	2
Bruschal	Br2	122	5.4	1	1	0	2
Bruschal	Br3	122	5.4	1	1	0	2
Bruschal	Br6	122	5.4	1	1	0	2
Bruschal	Br7	15	5.4	1	1	0	2
Charos Adama	28cam1	60	7.8	1	0	0	1
Colombia	Col	60	6.8	1	0	0	1
Eleftheres	12ele1	41	8.04	0	0	0	0
Kanava, Milos	20kam-1	70	7.5	1	0	1	2
Kanava, Milos	21kam-2	70	7.5	1	0	1	2
Kanava, Milos	22kam-3	70	7.5	1	0	1	2
Kanava, Milos	23kam-4	100	6.61	1	0	1	2
Kanava, Milos	24kam-5	60	6.47	1	0	1	2
Kanava, Milos	25kam-6	60	7	1	0	1	2
Krinides	10kri2	25	7.99	0	0	0	0
Krinides	9kri1	29.7	9	0	1	1	2
liogerma	31alm1	35	7.2	0	0	0	0
liogerma	32alm2	35	7.2	0	0	0	0
Lirima	Lr10	45	8.04	0	0	1	1
Lirima	Lr11	45	8.04	0	0	1	1
Lirima	Lr12	45	8.04	0	0	1	1

Lirima	Lr13	45	8.04	0	0	1	1
Lirima	Lr2	54	7.74	0	0	1	1
Lirima	Lr3	54	7.74	0	0	1	1
Lirima	Lr5	45	7	0	0	1	1
Lirima	Lr6	51	7.6	0	0	1	1
Lirima	Lr7	51	7.6	0	0	1	1
Lirima	Lr1	56	7.48	1	0	1	2
Lirima	Lr8	55	7.56	1	0	1	2
Lirima	Lr9	55	7	1	0	1	2
Nea Apollonia	5nap-2	80	8.2	1	0	0	1
Nea Apollonia	4nap-1	80	8.88	1	1	0	2
Nigrita	39NIG-1	43	8.2	0	0	0	0
Nigrita	40NIG-2	43	8.2	0	0	0	0
Nigrita	41NIG-3	43	8.2	0	0	0	0
Palaiochori	17pam2	37	6.47	0	0	0	0
Palaiochori	16pam1	80	2.9	1	0	1	2
Palaiochori	18pam3	80	6.74	1	0	1	2
Pikrolimni	2pik1	32	9.21	0	1	1	2
Pikrolimni	3pik2	32	9.86	0	1	1	2
Ponts-de-Martel, Iron	NeFer	15	6.94	1	0	0	1
Ponts-de-Martel, Sulfur	NeSulf	15	7.83	1	0	0	1
Potamia	8pot-1	72	8.56	1	1	0	2
Pozar	36POZ-2	40	8.37	0	0	0	0
Pozar	37POZ-3	40	8.37	0	0	0	0
Provatas	15prm1	42	6.13	0	0	0	0
Soultz	S3	144	6.2	1	0	0	1
Thermia	52the-4	20	7.6	0	0	0	0
Thermia	49the-1	57	7.6	1	0	0	1
Thermia	50the-2	60	7.6	1	0	0	1
Thermia	51the-3	60	7.6	1	0	0	1
Traianoupoli	6tra1	41	7.56	0	0	0	0
Traianoupoli	7tra2	41	7.34	0	0	0	0
Tria Pigadia	27tpm2	35	6.77	0	0	0	0
Zefuria plain	33zpm1	80	7	1	0	0	1
Zefuria plain	34zpm2	80	7	1	0	0	1

Supplementary table 3.4. Samples and normalized counts for 16S rRNA and *spoOA* gene copy numbers. The ratio between 16S rRNA and *spoOA* counts is presented in the last column (Ratio %).

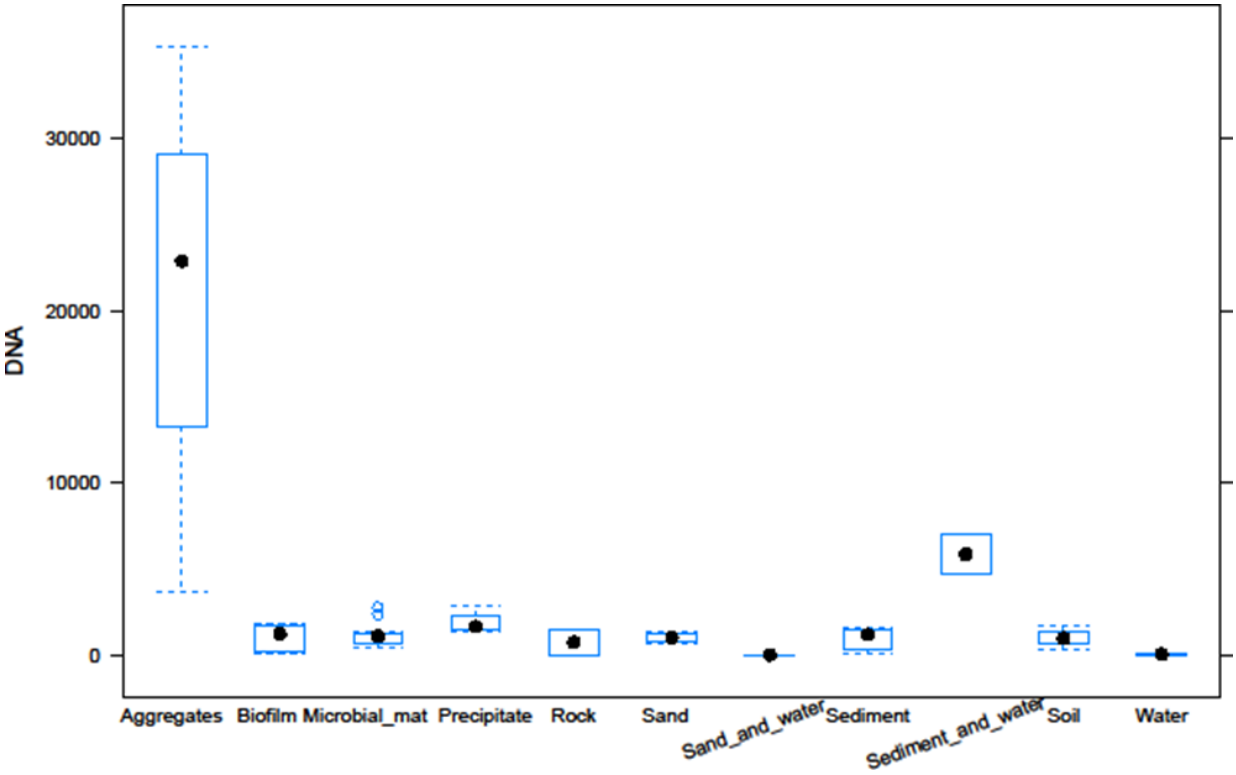
Location	Sample ID	16S rRNA copy numbers	*spoOA* copy numbers	Ratio %
Agia Paraskevi	43AGP-1	9.88E+06	5.31E+02	0.0054
Krinides	10kri2	1.80E+04	1.76E+00	0.0098
Aggistro	13agg1	1.56E+06	1.72E+02	0.0111
liogerma	32alm2	2.18E+06	2.50E+02	0.0115
Pozar	36POZ-2	2.86E+06	3.38E+02	0.0118
Tria Pigadia	27tpm2	5.20E+06	7.05E+02	0.0136
Agia Paraskevi	44AGP-2	6.65E+06	3.91E+03	0.0589
Pozar	37POZ-3	9.64E+05	1.59E+03	0.1646
Nigrita	39NIG-1	3.82E+06	1.69E+04	0.4420
Provatas	15prm1	7.07E+05	3.22E+03	0.4558
Nigrita	40NIG-2	4.26E+06	2.02E+04	0.4730
Traianoupoli	6tra1	5.62E+06	4.66E+04	0.8298
Aguas Calientes	Ac3a	4.61E+05	4.69E+03	1.0180
Thermia	49the-1	1.08E+06	1.27E+04	1.1759
Aguas Calientes	Ac3b	4.41E+05	5.54E+03	1.2552
Nigrita	41NIG-3	5.85E+05	9.39E+03	1.6042
Aggistro	14agg2	3.13E+07	7.47E+05	2.3899
liogerma	31alm1	3.68E+02	1.31E+01	3.5598
Aguas Calientes	Ac4a	1.05E+05	5.29E+03	5.0517
Traianoupoli	7tra2	3.60E+06	4.52E+05	12.5544
Eleftheres	12ele1	6.22E+06	8.92E+05	14.3316
Palaiochori	16pam1	1.40E+03	9.79E+02	70.0587
Lirima	Lr1	5.46E+06	5.08E+01	0.0009
Lirima	Lr10	9.00E+06	2.19E+02	0.0024
Lirima	Lr2	3.90E+06	2.28E+02	0.0058
Lirima	Lr13	4.09E+06	2.93E+02	0.0072
Lirima	Lr3	7.29E+06	1.07E+03	0.0147
Ponts-de-Martel, Sulfur	NeSulf	7.27E+07	1.98E+04	0.0272
Zefuria plain	33zpm1	8.09E+05	3.33E+02	0.0411
Agia Paraskevi	48AGP-6	1.80E+08	7.47E+04	0.0415
Zefuria plain	34zpm2	1.24E+06	8.44E+02	0.0683
Lirima	Lr12	4.19E+06	3.04E+03	0.0726
Lirima	Lr11	2.21E+06	1.81E+03	0.0819
Soultz	S3	4.48E+07	3.85E+04	0.0859
Ponts-de-Martel, Iron	NeFer	1.09E+07	1.20E+04	0.1095

Charos Adama	28cam1	9.23E+02	1.26E+00	0.1365
Lirima	Lr6	1.13E+07	2.60E+04	0.2299
Thermia	52the-4	4.28E+05	2.17E+03	0.5061
Aguas Calientes	Ac5b	1.39E+05	7.96E+02	0.5718
Agia Paraskevi	46AGP-4	2.76E+07	1.72E+05	0.6229
Lirima	Lr5	7.17E+06	5.68E+04	0.7922
Aguas Calientes	Ac4b	1.85E+05	1.55E+03	0.8387
Aguas Calientes	Ac5a	6.15E+04	1.20E+03	1.9559
Aguas Calientes	Ac6a	2.05E+05	5.11E+03	2.4948
Thermia	50the-2	1.49E+06	8.11E+04	5.4430
Agia Paraskevi	45AGP-3	4.00E+06	2.51E+05	6.2771
Agia Paraskevi	47AGP-5	3.93E+06	2.53E+05	6.4408
Thermia	51the-3	5.19E+07	3.71E+06	7.1484
Nea Apollonia	5nap-2	9.50E+05	6.83E+04	7.1895
Colombia	Col	3.89E+07	4.63E+06	11.9023
Lirima	Lr9	4.36E+06	8.36E+00	0.0002
Lirima	Lr7	8.02E+06	2.13E+02	0.0027
Kanava, Milos	22kam-3	2.25E+06	1.09E+02	0.0049
Lirima	Lr8	3.66E+06	7.28E+02	0.0199
Kanava, Milos	20kam-1	8.03E+05	6.32E+02	0.0787
Palaiochori	17pam2	5.26E+06	7.35E+03	0.1397
Kanava, Milos	21kam-2	1.83E+06	2.83E+03	0.1546
Pikrolimni	3pik2	1.27E+06	3.13E+03	0.2464
Kanava, Milos	25kam-6	8.44E+04	3.46E+02	0.4100
Bruschal	Br7	6.09E+07	3.59E+05	0.5895
Aguas Calientes	Ac6b	1.90E+05	3.38E+03	1.7795
Bruschal	Br3	9.34E+04	2.18E+03	2.3340
Pikrolimni	2pik1	3.46E+06	1.03E+05	2.9829
Bruschal	Br2	2.50E+03	1.17E+02	4.6800
Bruschal	Br6	4.00E+03	2.32E+02	5.8000
Potamia	8pot-1	2.43E+06	3.46E+05	14.2222
Krinides	9kri1	2.15E+03	4.27E+02	19.8666
Palaiochori	18pam3	9.61E+02	3.71E+02	38.5976
Kanava, Milos	24kam-5	2.00E+06	7.72E+05	38.6000
Kanava, Milos	23kam-4	5.03E+05	4.16E+05	82.7038
Nea Apollonia	4nap-1	1.86E+06	1.88E+06	100.8602

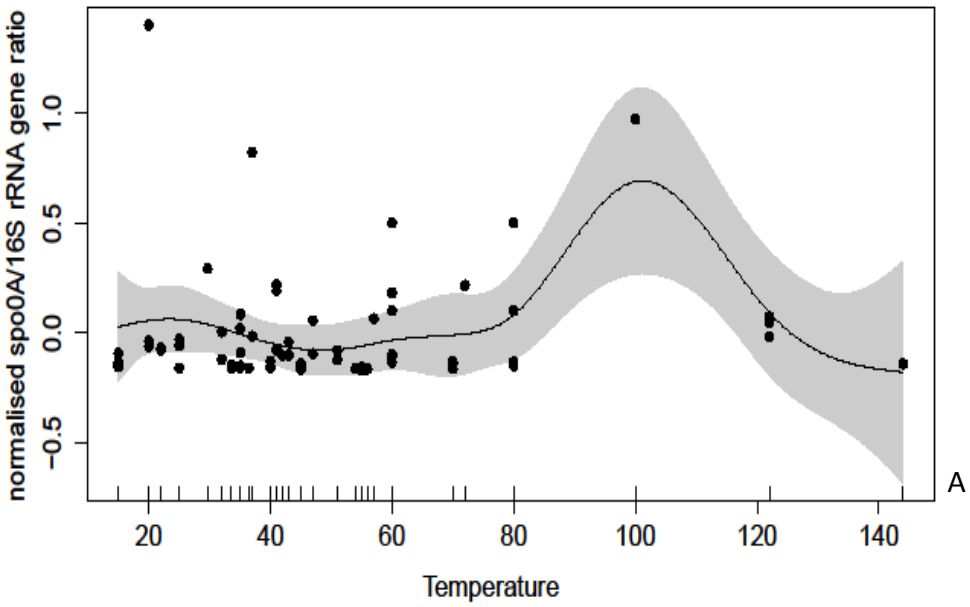
Supplementary table 3.5 Diversity of endospore-forming Firmicutes. Numbers correspond to OTUs per sample and per genus.

Taxonomy	4NAP-1	col	49THE-1	51THE-3	44AGP-2	25KAM-6	NEFER	NESUL
Anaerostipes	76	5	16	2	4	2	21	6
Anoxybacillus	3	0	0	0	0	0	0	0
Bacillus	33	1	1237	2	494	340	391	37
Brevibacillus	27	0	0	0	0	0	0	0
Clostridium	3544	21	1493	1939	2165	38	1983	395
Desulfotomaculum	2	1	0	13	0	0	7	1
Lysinibacillus	25	0	0	2	0	0	2	7
Paenibacillus	27	0	0	0	0	0	0	0

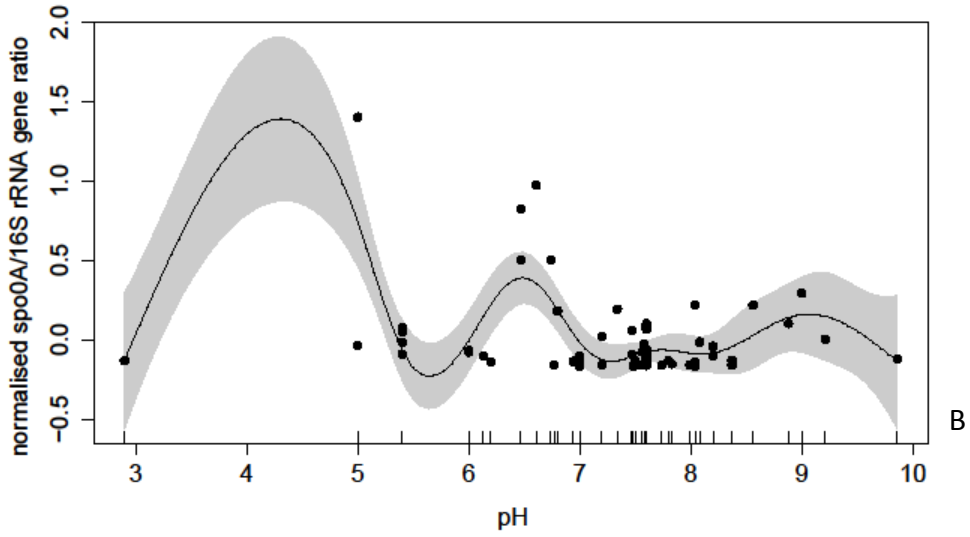
Supplementary Figure 3.1. DNA isolated in ng per type of sample.



Supplementary Figure 3.2. Correlation of relative abundance of EFF and selected environmental factors. Generalized additive models depicting the relative abundance of Endospore-forming Firmicutes –EFF- (16S rRNA/*spo0A* ratio; y-axis) to the in-situ measurements of temperature (A), pH (B). No significant correlation between relative abundance of EFF and these two factors is observed.

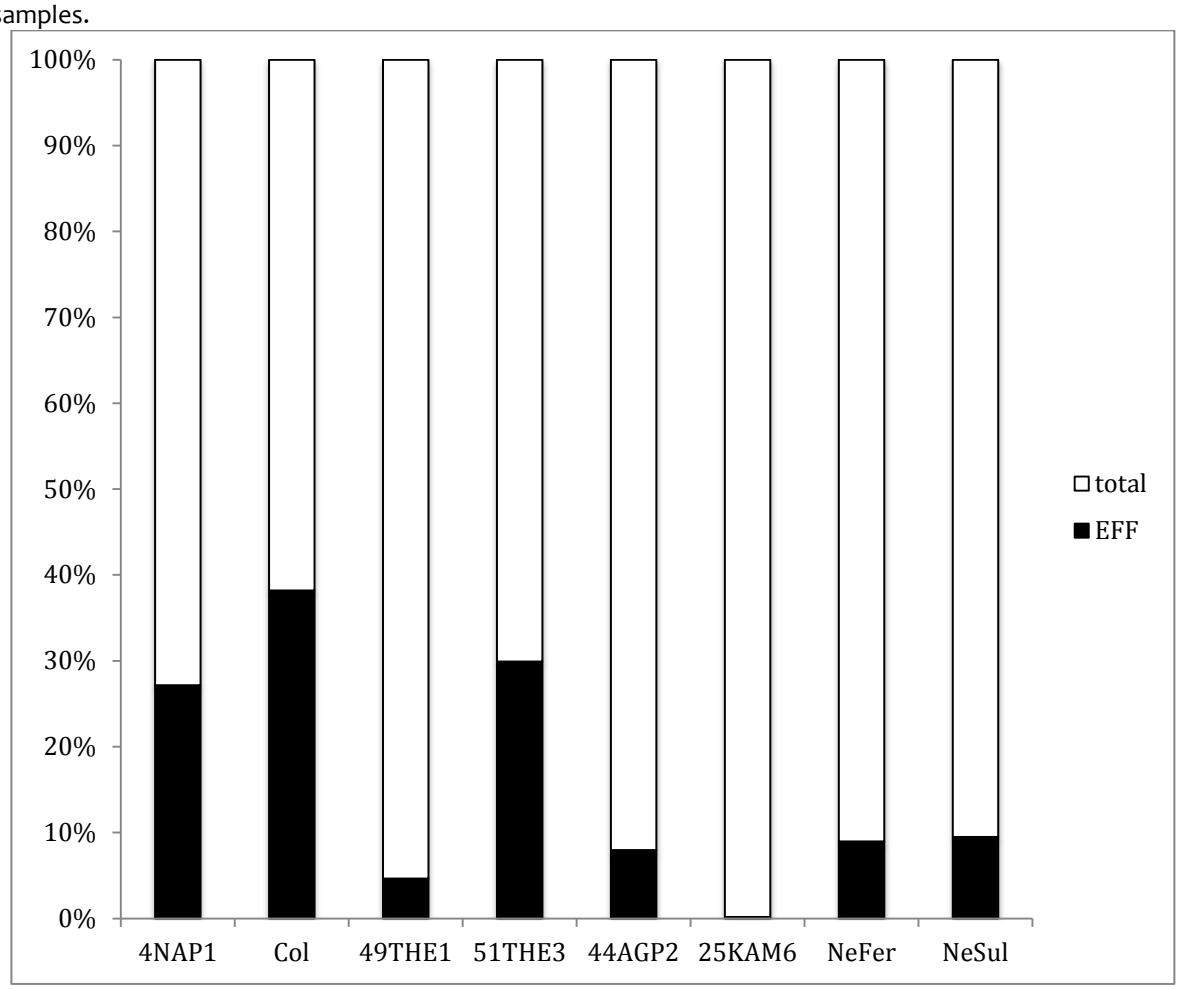


A



B

Supplementary Figure 3.3. Abundance of Firmicutes compared to the rest of the bacterial community in 8 samples.



```
#The following libraries were run upon initiation of the analysis
library(lattice)
library(car)
library(multcomp)
library(gplots)
library(effects)
library(MASS)

library(nlme)

################################

# Quantitative data

################################
library(RODBC)

dat = read.table("fil_quanty.txt", sep="\t", header=TRUE)
head(dat)
tail(dat)
summary(dat)
dim(dat)
names(dat)
dat$ratio <- dat$ratio/100
dim(dat[dat$ratio<=1,])-dim(dat)
dat[dat$ratio > 1,]            # the value bigger than 1
dat[15,]$ratio <- 1  # replace the value bigger than 1
dat$ratio.tr <- asin(sqrt(dat$ratio)) # transformed data
hist(dat$ratio.tr, col="mistyrose")
dat0 <- dat[,c(8:10,12)]  # only variables of interest
# Visualize data
pairs(dat0, panel = panel.smooth,

    cex = 1.5, pch = 24, bg = "light blue",

    diag.panel = panel.hist, cex.labels = 2, font.labels = 2)
# Tree method
require(party)
plot(ctree(ratio.tr ~ ph+HTM+Temp, data=dat))
plot(ctree(DNA ~ ph+HTM+Temp, data=na.omit(dat[,c(4,8:10)])))
# Some mean/vars in each group
bwplot(ratio.tr ~ Location, data=dat)
bwplot(ratio.tr ~ Source,   data=dat)
# Generalized additive model
```

```
require(mgcv)
plot(gam(ratio.tr ~ s(Temp), data=dat), residuals=T,  pch=16, pages=1, shade=T)
plot(gam(ratio.tr ~ s(ph), data=dat),   residuals=T,  pch=16, pages=1, shade=T)
plot(gam(ratio.tr ~ s(HTM), data=dat),  residuals=T,  pch=16, pages=1, shade=T)

# Full correlation matrix

cor(dat0, method="spearman", use="pairwise.complete.obs")

# Corelation ratio-Temperature

 # Scatterplot

 x <- dat$Temp

 y <- dat$ratio

 sequence <- order(x)

 plot(x, y, pch=16, col=2, ylab="ratio", xlab="Temp")

       lines(x[sequence],y[sequence], col=3)
dim(na.omit(cbind.data.frame(dat$ratio.tr,  dat$Temp)))[1]    # number of pairwise complete obs
cor(dat$ratio.tr, dat$Temp, use="pairwise.complete.obs")^2  # R-squared
cor.test(dat$ratio.tr, dat$Temp)
cor.test(dat$ratio.tr, dat$Temp, method = "spearman", continuity=F)
cor.test(dat$ratio.tr, dat$Temp, method = "kendall",  continuity=F)

cor.test(dat[dat$Temp >=12 & dat$Temp <=60 & dat$ratio <=0.004,]$ratio.tr,

 dat[dat$Temp >=12 & dat$Temp <=60 & dat$ratio <=0.004,]$Temp,

     method = "spearman", continuity=F)  # correlation on the subset of "normal temperature"


# Correlation log(DNA)-Temperature
hist(log(dat$DNA+1), col="mistyrose")
cor.test(log(dat$DNA+1), dat$Temp, method = "spearman")

 # Scatterplot

 y <- log(dat$DNA+1)

 x <- dat$Temp
```

```
 sequence <- order(x)

plot(x, y, pch=16, col=2, ylab="DNA log transformed", xlab="Temp")

lines(x[sequence],y[sequence], col=3)
 # Correlation ratio-HTM

 # Scatterplot

 x <- dat$HTM

y <- dat$ratio.tr

sequence <- order(x)

 plot(x, y, pch=16, col=2, ylab="ratio", xlab="HTM")

          lines(x[sequence],y[sequence], col=3)
dim(na.omit(cbind.data.frame(dat$ratio.tr, dat$HTM)))[1] # number of pairwise complete
obs
cor(dat$ratio.tr, dat$HTM, use="pairwise.complete.obs")^2  # R-squared
cor.test(dat$ratio.tr, dat$HTM)
cor.test(dat$ratio.tr, dat$HTM, method = "spearman", continuity=F)
cor.test(dat$ratio.tr, dat$HTM, method = "kendall",  continuity=F)

lmo <- lm(ratio.tr ~ HTM, data=dat[-c(3,5),])  #  3,5 are outliers

summary(lmo)
     # Check model assumptions
      par(mfrow=c(2,2));plot(lmo); par(mfrow=c(1,1))
plot(dat$HTM, dat$ratio.tr, col="red"); abline(lmo, col="blue")

# Corelation ratio-pH

 # Scatterplot

 x <- dat$ph

 y <- dat$ratio.tr

 sequence <- order(x)

 plot(x, y, pch=16, col=2, ylab="ratio", xlab="ph")
```

```
 lines(x[sequence],y[sequence], col=3)
dim(na.omit(cbind.data.frame(dat$ratio.tr, dat$ph)))[1]  # number of pairwise
complete obs
cor(dat$ratio.tr, dat$ph, use = "pairwise.complete.obs")^2  # R-squared
cor.test(dat$ratio.tr, dat$ph)
cor.test(dat$ratio.tr, dat$ph,  method = "spearman", continuity=F)
cor.test(dat$ratio.tr, dat$ph,  method = "kendall",  continuity=F)

##############################

# Qualitative data

##############################
library(RODBC)
#datfile = "fil_quali.xlsx"                          # get the data path
#getxlsbook = odbcConnectExcel2007(datfile)                  # link the excel data book
#dat2 = sqlFetch(getxlsbook, "Sheet2")                       # get the data from the data
sheet
#odbcCloseAll()
dat2=read.table("fil_quali.txt", sep="\t", header=TRUE)
head(dat2)
summary(dat2)
dim(dat2)
names(dat2)
dat2 <- dat2[,-c(11,12,14)]
names(dat2)[7] <- "ratio"
dat2$ratio <- dat2$ratio/100
round(range(dat2$ratio),2)
dat2[118,]$ratio <- 1
hist(dat2$ratio, col="mistyrose")

dat2$ratio.tr <- asin(sqrt(dat2$ratio)) # transformed data
hist(dat2$ratio.tr, col="mistyrose")

# Transform Nofact in a factor (ordered)
dat2$Nofact <-as.factor(dat2$Nofact)
is.ordered(dat2$Nofact)
levels(dat2$Nofact)

dat2$Nofact <- as.ordered(dat2$Nofact) # test used to verify orthogonal polynomial
contrasts
#levels(dat2$Nofact)[5] <- "4"
```

```
#levels(dat2$Nofact)[2] <- "2"

table(dat2$Nofact)

# Visualize data as Grouped Bar Plot

barplot(as.matrix(rbind(tapply(dat2$tot, dat2$Nofact, mean),tapply(dat2$mycro, dat2$Nofact, mean))),

 main="Microorganism abundance",

 xlab="Nb of extreme factors", col=c("darkblue","red"),

 legend = c("total", "my organism"), beside=TRUE)

barplot(as.matrix(rbind(tapply(log(dat2$tot+1), dat2$Nofact, mean),tapply(log(dat2$mycro+1), dat2$Nofact, mean))),

 main="Log-transformed Microorganism abundance",

 xlab="Nb of extreme factors", col=c("darkblue","red"),

 legend = c("total", "my organism"), beside=TRUE)

plot(levels(dat2$Nofact), tapply(dat2$tot, dat2$Nofact, mean), type="b", col = "darkblue")

plot(levels(dat2$Nofact), tapply(dat2$mycro, dat2$Nofact, mean), type="b", col= "red", add=TRUE)

plot(levels(dat2$Nofact), tapply(dat2$ratio, dat2$Nofact, mean), type="b", col= "orange", add=TRUE)

# Tree method

require(party)
plot(ctree(ratio.tr~ Nofact + HTM, na.omit(dat2[,c(10,12:13)])))

# Some mean/vars in each group

bwplot(log(tot+1) ~ Nofact, data=dat2[-c(84,89,90,103),])  # we excluded some extreme values from the plot
```

```
bwplot(ratio.tr~Nofact, data=dat2[-c(84,89,90,103),])      # we excluded some extreme values from the plot


bwplot(ratio.tr~Location, data=dat2)

bwplot(ratio.tr~Source, data=dat2)

# Classical ANOVA

model <- lm(ratio.tr ~ Nofact, data=dat2[dat2$Nofact!="8",])

summary(model)

# Check model assumptions

      par(mfrow=c(2,2));plot(model); par(mfrow=c(1,1)) # not OK!! model is not appropriate

model1 <- gls(ratio.tr ~ Nofact, weights = varIdent(form=~1|Nofact), data=dat2[-c(40,52,53),])
summary(mmodel3)

 # Check model assumptions

 qqnorm(model1, abline=c(0,1), id = 0.05)

 plot(model1)

plot(effect("Nofact", model1)) # you can use this plot even if the model is not quite OK


# Non-parametric tests of Nofact effect
require(agricolae)
comparison <- kruskal(dat2$ratio.tr, dat2$Nofact, group=TRUE, main="Relative abundance")
comparison
comparison1 <- kruskal(dat2$tot, dat2$Nofact, group=TRUE, main="Relative abundance")
comparison1
# Model for total abundace of microorganisms
mmodel3t <- gls(log(tot+1) ~ Nofact, weights = varIdent(form=~1|Nofact), data=dat2)
```

```
summary(mmodel3t)

# Check model assumptions

qqnorm(mmodel3t, abline=c(0,1), id = 0.05)

plot(mmodel3t)

plot(effect("Nofact", mmodel3t))

#a model that remove the variability between locations
mmodel <- lme(log(tot+1) ~ Nofact, random = ~ 1|Location, data=dat2)
summary(mmodel)
plot(mmodel)
mmodel1 <-update(mmodel, weights = varIdent(form=~1|Nofact) ) summary(mmodel1)
anova(mmodel,mmodel1) # improvement

# Check model assumptions

plot(mmodel1, resid(., type = "p") ~ fitted(.), id = 0.05)

qqnorm(mmodel1, abline=c(0,1), id = 0.05, col="blue")

qqnorm(mmodel1, ~ ranef(.))

plot(mmodel1, Location~resid(.), abline = 0 )

require(lattice)
multiple.txt <- read.table("multiple2.txt", header=T)
head(multiple.txt)
single.txt <- read.table("single2.txt", header=T)
head(single.txt)
no.txt <- read.table("no2.txt", header=T)
head(no.txt)
summary(multiple.txt)
summary(single.txt)
summary(no.txt)
shapiro.test(multiple.txt$Multiple) # not normal
shapiro.test(single.txt$Single)    # not normal
shapiro.test(no.txt$No_Factor)     # not normal

varia <- c(multiple.txt$Multiple, single.txt$Single, no.txt$No_Factor)
treat <- c(rep("Multiple", length(multiple.txt$Multiple)),
rep("Single", length(single.txt$Single)),
```

```
    plot(mmodel1, resid(., type = "p") ~ fitted(.)|Location, id = 0.05)
rep("NoFact", length(no.txt$No_Factor)))
all <- cbind.data.frame(varia, treat)
summary(all)
boxplot(c(multiple.txt, single.txt, no.txt), ylab="qPCR ratio spo0A/16S", xlab = "sporulation triggering factors" )
bwplot(varia ~ treat, ylab="qPCR ratio spo0A/16S", xlab = "sporulation triggering factors")
(Means <- tapply(all$varia, all$treat, mean))
(Vars <- tapply(all$varia, all$treat, var))
plot(Means, Vars)
##############################################
 # Non-parametric analysis
##############################################
# 1st version
require(agricolae)
cmp <- kruskal(all$varia, all$treat, group=T)
cmp
# 2nd version
require(PMCMR)
kruskal.test(all$varia~all$treat)
posthoc.kruskal.nemenyi.test(x=all$varia, g=all$treat, method="Tukey")
```

Chapter 4

Another path to survival: discovery of sporulation in *Serratia ureilytica*

Sevasti Filippidou[1], Thomas Junier[2], Tina Wunderlin[1], Wafa M. Kooli[1], Nicole Jeanneret[1], Veronica Molina[3], Reto Lienhard[4], Jorge Spangerberg[5], Shannon Johnson[6], Patrick Chain[6], Cristina Dorador[7], Pilar Junier[1]

[1]Laboratory of Microbiology, University of Neuchatel, Emile Argand 11, Neuchatel, CH-2000, Switzerland.

[2]Vital-IT Group, Swiss Institute of Bioinformatics, Lausanne, CH-1015, Switzerland

[3]Departamento de Biología, Facultad de Ciencias Naturales y Exactas. Universidad de Playa Ancha, CL-2360002,Valparaíso, Chile.

[4]ADMED Laboratoires, site de La Chaux-de-Fonds, ADMED Microbiologie, Boucle de Cydalise 16, CH-2300 La Chaux-de-Fonds Switzerland.

[5]Institute of Earth Surface Dynamics, University of Lausanne, Building Géopolis, Lausanne, CH-1015, Switzerland.

[6]Los Alamos National Laboratory, Los Alamos, New Mexico, 87545, USA.

[7]Laboratorio de Complejidad Microbiana y Ecología Funcional; Departamento de Biotecnología; Facultad de Ciencias del Mar y Recursos Biológicos, Universidad de Antofagasta; Antofagasta, CL-1270190, Chile.

This chapter is to be submitted for publication at *Nature Microbiology*

Abstract

Spore-formation is a sophisticated survival strategy to withstand adverse environmental conditions. It is described as a condition of low metabolic activity resulting into a highly resistant specialized cellular form. Spore-like structures have been found in only four taxa: Firmicutes, Actinomycetes, Cyanobacteria, and *Myxococcus*. Here, we show that a γ-proteobacterium, *Serratia ureilytica* strain Lr5/4,is able to produce spores that resemble those of endospore-forming Firmicutes in some structural and functional aspects. Sporulation in strain Lr5/4 is triggered by nutrient starvation, thermal shock and UV radiation. Like those of Firmicutes, Lr5/4 spores are resistant to heat, UV radiation, and desiccation. Genomic analysis of this strain shows evidence of unidirectional horizontal gene transfer from thermophile endospore-forming Firmicutes. The transfer involves elements of all stages of this morphophysiological differentiation process. These findings challenge the uniqueness of characteristics assigned to endospores and underline the role of horizontal gene transfer as a mechanism of acquiring new functions among unrelated microorganisms.

4.1. Introduction

In order to avoid extinction under unfavorable conditions, microbial communities have developed diverse survival strategies. These strategies can be beneficial for the entire community, such as in the case of biofilm formation [1]; or favor the survival of individuals, such as processes involving morphological plasticity [2] or post-transcriptional modifications optimizing fitness [3]. Dormancy is a well-studied survival strategy, which corresponds simultaneously to both categories: community and individual survival. Dormancy is an individual characteristic, since not all bacteria possess genes that enable them to switch from vegetative to dormant state. However, it has been recently described as a community strategy as well, proposing an "altruistic behavior" among dormant cells before, during and after entering the state of dormancy [4–6]. Sporulation is one of such a dormant state. It is described as a condition of low metabolic activity and shrinking in size [7]. To date, it is known to occur in four bacterial phyla, Firmicutes (producing endospores), Actinomycetes (spores), Cyanobacteria (akinetes), and in the *Myxococcus* genus, belonging to δ-Proteobacteria (spores) [8].

Sporulation is not only a sophisticated survival strategy, but also a prime example of an elaborate developmental process in bacteria [7]. Sporulating bacteria undergo an intricate sequence of cell differentiation events leading to the formation of spores. A series of genes is necessary for sporulation to occur and the absence or malfunction of one or more of those results in failure to sporulate [9]. Indeed, there are bacteria that do encode some sporulation genes in their genome but cannot produce spores, and are therefore called asporogenic. Although sporulation differs significantly among taxa, it can be described as a general 4-step process. In the first step cells detect unfavorable conditions in their environment. Afterwards, commitment to sporulation follows, as an irreversible step. A dormant structure, mostly called *a spore*, is then produced either as a result of a special cell division or a modification of the vegetative cell structure. Finally, a mature spore is produced, ready to germinate once conditions are favorable once again.

Among the different spore-forming taxa, spores formed by Firmicutes are unique in their formation and their resistance to wet heat. They are formed by an asymmetrical division within a "mother" cell and are thus named *endospores*. Recent genomic studies among endospore-forming Firmicutes have shown that there is a minimum set of 60 genes required during the entire sporulation pathway [10, 11]. The knowledge of the genetic determinants of endospore-formation has allowed detecting endospore-formation in previously thought asporogenic species (i.e. *Carboxydothermus hydrogenoformans* [12]), as well as among uncultured bacteria (e.g. Candidatus *Desulforudis audaxviator*, the only known example of a single-species ecosystem [13]).

4.2. Materials and methods

4.2.1. Sampling and isolation

The microbial mat was collected and stored using sterile material. After sampling, the sample was preserved at 4°C until enrichment. For aerobic enrichment, one gram of homogenized mat was inoculated in nutrient broth (NB) (Biolife, Italy) and in modified

marine broth (MB) [14], pH adjusted to 7.2, and incubated at room temperature for 7 days under aerobic conditions. 100 µl of the enrichment culture was plated on nutrient agar (NA) (Biolife, Italy) or on modified marine agar (MA), respectively, and incubated for 24 h at room temperature (RT) under aerobic conditions. Pure strains were isolated after repeated serial dilutions on NA or MA medium at RT for 24 h. Strains were then tested for purity and observed by microscopy. Pure isolates were cryopreserved in 30% (v/v) of glycerol at -80°C.

4.2.2. Physiological characterization of Lr5/4

Cell growth was monitored at different temperatures (4, 15, 25, 30, 37, 45, 55 and 60°C) for up to 72 h, measuring optical density at 600nm (OD600) with a spectrophotometer Genesys 10S UV-Vis Thermoscientific (Thermo Fisher Scientific Inc., USA). In addition, growth was measured at different pH (2 to 13) and at different NaCl concentrations (1, 2, 3, 3.5, 4, 5, 5.5, 6, 6.5, 8, 9, 10, 11, 12% w/v) with a microplate reader at 590nm (UVM 340 Microplate Reader, ASYS, Hiteck, UK) for up to 72 h. Finally, a growth curve was established under optimum conditions (Data not shown). The need for oxygen during growth was verified using the method of thioglycollate medium [15].

Phospholipid extraction from the whole cells was done using a single-phase (methanol/dichloromethane/phosphate buffer; 2:1:0.8; v/v/v) extraction procedure [16]. Fatty acid methyl esters (FAMES) were prepared from the extracted acyl lipid mixture by acid-catalyzed methanolysis [17]. The chemical characterization of the lipids was performed with an Agilent gas chromatograph 6890 coupled to an Agilent 5973 quadrupole mass selective detector (GC-MS; Palo Alto, USA). For the analyses of FAME fractions the system was equipped with an Agilent free fatty acids phase (FFAP) fused silica capillary column (50 m length, 0.20 mm i.d.) coated with nitroterephthalic acid modified polyethylene glycol stationary phase (film thickness 0.33 µm). A sample aliquot was injected splitless at a temperature of 200°C. Helium was used as carrier gas (1 mL/min flow rate). After an initial period of 2 min at 100°C, the column was heated to 240°C at 5°C/min followed by an isothermal period of 30 min. The MS was operated in the electron impact mode at 70 eV, source temperature of 250°C, emission current of 1 mA and multiple-ion detection with a mass range from 50 to 600 amu. Compound identifications were based on comparison of standards, GC retention time, and mass spectrometric fragmentation patterns.

4.2.3. Molecular characterization

For G+C content analysis, the services of DSMZ, Germany were used. Cells were disrupted by French pressing and DNA was purified on hydroxyapatite as previously described [18]. The DNA was hydrolyzed with P1 nuclease and the nucleotides were dephosphorylized with bovine alkaline phosphatase [19]. The resulting deoxyribonucleosides were analyzed by HPLC [20]. Lambda-DNA and three DNAs with published genome sequences representing a G+C range of 43-72 mol% were used as standards. G+C values were

calculated from the ratio of deoxyguanosine and thymidine according to the method of Mesbah *et al.* [19].
DNA-DNA hybridization was carried out at DSMZ, Germany. Cells were disrupted by using a Constant Systems TS 0.75 KW (IUL Instruments, Germany) and the DNA in the crude lysate was purified by chromatography on hydroxyapatite as previously described [18]. DNA-DNA hybridization was carried out as described previously [21,22] using a model Cary 100 Bio UV/VIS-spectrophotometer equipped with a Peltier-thermostatted 6x6 multicell changer and a temperature controller with in-situ temperature probe (Varian). Strain Lr5/4 was tested against *S. marcescens* (DSMZ 30121) and *S. ureilytica* (DSMZ 16952).
MALDI-TOF mass spectra were acquired on a Bruker Microflex RLF mass spectrometer (Bruker Daltonics, Bremen, Germany). Spectral mass resolutions and signal-to-noise ratios were determined with the software Flex Analysis 3.3.65 (Bruker Daltonics).
For PCR amplifications, genomic DNA was extracted using the InnuPREP Bacteria DNA kit (Analytik Jena, Germany), according to the manufacturer's instructions. DNA was quantified fluorometrically with a Qubit® dsDNA HS Assay Kit with a Qubit®2.0 Fluorimeter (Invitrogen Ltd., Paisley, UK) according to the manufacturer's instructions. PCR amplification of the 16S rRNA gene was performed using the primer set GM3F and GM4R, according to Muyzer *et al.* [23]. The PCR products were purified with a MultiScreen PCRμ96 Filter Plate (Millipore), according to the manufacturer's instructions and sequenced using the services of GATC Biotech (Germany). To obtain the full-16S rRNA gene sequence, Lr5/4 PCR products were sequenced in addition with the primers 907r, 520r and 926f primers [23,24]. The 16S rRNA gene sequence was compared in GenBank by using the BLASTn tool [25]. After alignment, a series of sequences that had an identity match over 98% were selected, along with 16S rRNA gene sequences from cultured representative strains belonging to different species of the genus *Serratia* and a 16S rRNA gene sequence of a *Geobacillus* sp. (serving as an outgroup in order to root the tree). These sequences were used to build-up a cladogram using the online tools of phylogeny.fr website [26]. Multiple alignment has been made using MUSCLE and the phylogenetic tree was constructed using PhyML. To determine the confidence values for individual branches, 100 bootstrap replications were done for each generated tree. In order to test whether this *Serratia* strain has the gene that encodes the transcriptional factor responsible for the initiation of sporulation in endospore-forming Firmicutes (*spo0A* gene), a set of specific primers (*spo0A*166f and *spo0A*748r) was used as described previously [27].

4.2.4. Morphological characterization

Colonies of strain Lr5/4 were obtained and observed after overnight growth on NA. Gram staining was performed on an overnight solid culture using the Hucker staining method [28]. Spores were observed with a contrast-phase microscope (Leica DM R, magnification 1000x). Vegetative cells and spores were also observed by Scanning and Transmission Electron Microscopy (SEM and TEM). Vegetative cells were observed from a fresh 24 h culture inoculated in NB. Spores were collected from a 3-month old solid culture left at RT

to sporulate. Both preparations have been fixed in 2.5% glutaraldehyde in a cacodylate buffer (0.1M; pH7.4) for 2h at room temperature and then overnight at 4°C. They were then washed by gentle immersion in cacodylate buffer (0.2M; pH7.4) post fixed with 1% OsO_4 in the same buffer, and carefully washed with the above buffer. For SEM, the samples were dehydrated in 15 to 100% of ethanol solution and finally fixed on Poly-L-Lysine slides and coated with a 23nm gold layer in a BaltecSCD005 sputter apparatus. The samples were observed with a Philips XL30 SEM at acceleration voltages of 10-20kV. For TEM, the samples were dehydrated in 15–100% acetone and embedded in Spurr's resin. Serial sections were made with a Reichert Ultracut-S microtome, mounted on copper grids, double stained with uranyl acetate and lead citrate, and observed with a Philips CM 100 TEM at 60 or 80 kV.

4.2.5. Verification of heat-resistance

A heat resistance test for the spores of strain Lr5/4 was performed as previously described [29], along with *S. marcesens* (DSMZ 30121), *S. ureilytica* (DSMZ 16952), and *Bacillus subtilis* (Neu1294). For the heat resistance test, cultures were left to sporulate between 10 and 90 days. Suspensions of these cultures in tryptic soy broth (TSB) medium were heat-shocked at 70 and 75°C for 20 min and then re-cultured under optimal conditions. Growth was macroscopically verified after 24 and 48h.

4.2.6. DPA measurement

The presence of dipocolinic acid (DPA) in the spores was assessed on 3-month old spore preparation of Lr5/4 (wet weight= 20mg), according to a previously published method [30]. Fluorescence was measured within a Perkin-Elmer LS50B fluorimeter. The excitation wavelength was set at 272 nm with a slit width of 2.5 nm. Emission was measured at 545 nm (slit width 2.5 nm). The device was set in the phosphorescence mode (equivalent to time-resolved fluorescence). The delay between emission and measurement was set at 50 µs. Measurements were performed every 20 ms. The integration of signal was performed over a duration of 1.2 ms. Values recovered for each measurement corresponded to the mean of the relative fluorescence unit (RFU) values given by the instrument within the 30 s following sample introduction in the device. Finally, to transform RFU units into DPA concentrations, a 10-point standard curve was established using increasing concentrations of DPA from 0.5 µM up to 10 µM.

4.2.7. Genome sequencing and analysis

Genomic DNA was extracted from an 12h culture using the Genomic-tip 20/G Kit (Qiagen GmbH, Germany). Sequencing was performed with Pac Bio RS II system based on single molecule, real-time (SMRT) technology (Pacific Biosciences, California). Genome annotation was performed using the NCBI Prokaryotic Genome Annotation Pipeline (PGAAP) and visualized with Artemis Genome Browser and Annotation Tool [31]. Genome annotation utilized an Ergatis based [32] workflow. The circular genome with the additional features was created using DNAplotter [34].

4.2.8. Horizontal gene transfer analysis

The two *Serratia* proteomes were scanned for orthologues of Firmicutes core sporulation genes (at the protein sequence level) with BLASTP [35] using default parameters and an E-value cutoff of 1E-20. This resulted in a set of orthology groups, each group pertaining to one sporulation gene. Within each group, *Serratia* is represented by exactly one sequence, and each Firmicutes species is represented by at most one sequence - this reflects the fact that some sporulation genes are not found in all species. When a Firmicutes species had several matches to *Serratia*, only the best one (in terms of E-value) was retained. Orthology groups with fewer than four species were discarded, as the lateral gene transfer detection method (see below) requires trees of at least four species. The sequences in each orthology group were aligned with Mafft [36], using settings for high-accuracy, local alignments (`linsi`). The phylogeny of each orthology group was computed with PhyML [37], using the alignments computed at the previous step. PhyML parameters were as follows: LG substitution model, 16 substitution rate classes, best search moves, estimation of the proportion of invariants and alpha parameter, and optimization of topologies, substitution rate parameters, and branch lengths. The jobs were run in parallel using GNU `parallel` [38]. The phylogenies were then re-rooted on *Serratia* and converted to cladograms using the Newick Utilities [39]. The resulting trees, as well as a phylogeny of Firmicutes and two *Serratia* genomes, were subjected to analysis by HiDe [40] in order to detect lateral gene transfer highways.

4.3. Results

4.3.1. An unusual spore-former isolated in a thermal spring

Lirima is a geothermal site located in Andean highlands in Northern Chile (3.997m, 19°51.118W, 68°54.402S). During a field campaign in April 2011, we sampled microbial mats in different geothermal environments (temperature ranging from 30°C to over 90°C). Our aim was to study aerobic spore-forming bacteria and therefore a temperature mismatch was used during cultivation to favor the selection of dormant cells. Spore-formers were screened by microscopy after letting the isolates to sporulate by nutriment deprivation. The spore-forming strain Lr5/4 was isolated from a double layer microbial mat in a hot stream of 54°C (Figure 4.1).

Strain Lr5/4 was isolated alongside four other strains that were identified as belonging to *Bacillus* spp. The characterization of strain Lr5/4 showed that its oval cells (0,5x1-3µm) stained Gram-negative, in contrast to Bacilli, which are Gram-positive bacteria. Growth tests showed that strain Lr5/4 is a mesophile (temperature growth range between 10 and 37°C) with an optimum at 25°C, which is far from the temperature measured in the sampling site (56°C). It is also a moderate halotolerant strain (growth up to 8% NaCl), able to develop over a vast range of pH (from 3 to 11; optimal 5 to 6), and a facultative anaerobe.

Sequencing and analysis of the full 16S rRNA gene demonstrated that strain Lr5/4 belongs to the genus *Serratia*, with more than 97% identity to *Serratia marcescens* and *Serratia*

ureilytica. To establish the phylogenetic affiliation of this species, a series of additional analysis were performed. MALDI-TOF characterization suggested that strain Lr5/4 belonged to *S. marcescens* (Supplementary Figure 4.1). However, DNA/DNA hybridization between Lr5/4 and its closest relatives suggest that it belongs to *S. ureilytica* species, with above 99% DNA/DNA relatedness to *S. ureilytica* DSM-16952. Nonetheless, DNA/DNA relatedness was also found to be above 70% in the case of hybridization with *S. marcescens* subsp. *marcescens* (DSM-30121), which is the recommended threshold for bacterial species delimitation [41]. This observation is intriguing as it has been previously shown that DNA/DNA relatedness between the two reference strains (*S. ureilytica* and *S. marcescens* subsp. *marcescens*) is only 43.7% [42]. Analyses of total fatty acids of Lr5/4, *S. marcescens* subsp. *marcescens,* and *S. ureilytica* were also performed (Supplementary Table 4.1). This comparison showed that palmitic acid is the major fatty acid in the three strains and that Lr5/4 is more closely related to *S. ureilytica* based on the whole fatty acid composition, although the fatty acid profile of *S. ureilytica* type strain correlates better with *S. marcescens* (0.86) than with *S. ureilytica* Lr5/4 (0.77) (Supplementary Figure 4.2). Considering all biochemical and molecular identification, we concluded that strain Lr5/4 represents a novel strain of *S. ureilytica*.

Figure 4.1. Geothermal pond at Lirima, Chile

4.3.2. Sporulation in *S. ureilytica* Lr5/4

Spores are a response to unfavorable environmental conditions. Although, many factors can hinder bacterial survival, nutrient deprivation is so far the only common trigger to sporulation known to trigger sporulation in all bacterial spore-formers (Supplementary Table 4.2). In the case of Lr5/4, spores developed spontaneously in response to starvation (after 3 months). The roles of other known triggers of sporulation that could be relevant

in the geothermal spring from which we isolated the strain (UV radiation, desiccation, thermal shock, salinity, toxic metal concentration, pressure and chemicals) were also measured. Spores of Lr5/4 were produced in response to UV radiation and after thermal shock (Supplementary Table 4.3). In the case of salinity and elevated metal concentrations, cells of Lr5/4 were largely tolerant. The same is true for desiccation, and extreme temperatures applied individually (down to -80°C and up to 100°C), which did not affect growth of vegetative cells.

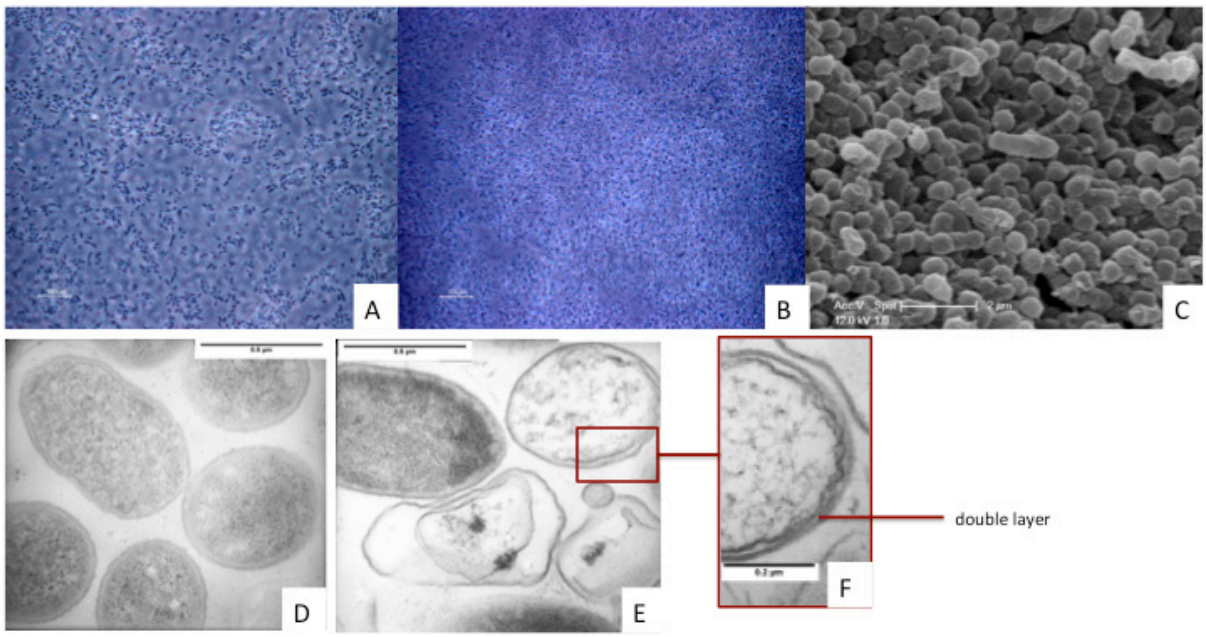


Figure 4.2. Microscopic observation of *Serratia ureilytica* strain Lr5/4: (A) Vegetative cells at contrast-phase microscope (1000x magnification) after 24h growth; (B) Free spores along with vegetative cells in 3 months old culture at contrast-phase microscope (1000x magnification); (C) Vegetative cells and spores at SEM; (D) Thin dissection of vegetative cells at TEM (21500x). (E) Thin dissection of a spore –S-, a vegetative cell –Vc- and phantom –Pc- cells at TEM (21500x). (F) A magnification of the outer layers of Serratia is shown at TEM. Two distinct layers are visible, the inner structure of the spore, however, does not resemble the concentrical structure of Firmicutes spores.

The morphological characterization of spores in strain Lr5/4 was further conducted using electron microscopy. In scanning electron microscopy, vegetative cells have a bacillus-like shape of about 2.0x0.5 μm size, while once formed, spores are round spheres of 0.5 μm in diameter (Figure 4.2c). In transmission electron microscopy the interior of vegetative cells is stained after osmium fixation while spores presented a clearer core surrounded by two thin outer layers (Figure 4.2e-f). Although the exact composition of these two layers is not yet defined, they clearly prevent the penetration of the OsO_4, used for fixation. The later penetrated easily the cell wall and membrane of the vegetative form, which became dark grey (Figure 4.2d). OsO_4 is used for fixation of spore structures of Firmicutes, Actinomycetes [43], and Myxococcus [44]. The ability of OsO_4 to penetrate and stain the coat in Bacilli have been previously shown to depend on the coat proteins that participate in the assembly and formation of the inner and outer layers of the spore coat [45]. An increase in the thickness of the coat renders more difficult the penetration of OsO_4, resulting into a bad resolution of the inner coat and the spore protoplasm [46]. This has

also been shown in germination experiments: at spore state, the spore protoplasm is brighter and becomes darker through the germination process, since the vegetative cell wall and membrane are easier to penetrate than the spore coat [47]. This suggests that on spite of its simplicity, the outer layers in the spores of strain Lr5/4 offer sufficient impermeability to this staining agent. An additional outer structure, such as an exosporium, which is typical of some endospore-forming Firmicutes, was not observed in the case of *S. ureilytica* Lr5/4.

4.3.2. Resistance of the spores

Spores are an effective survival strategy that allows the genetic material of the microbial cell to remain intact for long periods of time (sometimes claimed to be up to $250x10^6$ years [48]). Spores from Firmicutes are considered the toughest biological structures on the planet [49], and are different from other spore-like structures because of their resistance to wet heat [50]. In order to evaluate the effectiveness of Lr5/4 spores as a survival strategy, spores from 90- and 10-day old cultures of strain Lr5/4 were challenged with multiple stressors. Spores of Lr5/4 were able to germinate and develop after a heat-shock at 70 and 75°C, 4h UV radiation exposure, desiccation, and extreme temperatures. This was comparable to the survival of spores from the endospore-forming *Bacillus subtilis* (Supplementary Table 4.4). However, non-spore forming strains of the genus *Serratia* (*S. marcescens* DSMZ 30121 and *S. ureilytica* DSMZ 16952) and *Escherichia coli* did not survive these resistance tests. When spores of Lr5/4 and *B. subtilis* were subjected to multiple stressors applied simultaneously (starvation, desiccation and UV radiation), only Firmicutes spores could be revived, while Lr5/4 spores did not resist.
Although for most of the stressors studied the precise molecular mechanism involved in resistance is unknown, in the case of wet heat the protection of DNA is conferred by the accumulation of dipicolonic acid (DPA) in the spore core of endospore-forming Firmicutes [50]. The presence of DPA in the spores [30] was assessed on serial dilutions of a 3-month old spores preparation of Lr5/4 (wet weight= 20mg).

4.3.3. A framework for the emergence of sporulation in *Serratia*

Several evolutionary scenarios can be proposed in order to explain the unexpected detection of spores in a reportedly non-spore forming bacterial group. On the one hand, the selective pressure from the environment could lead *Serratia ureilytica* str. Lr5/4 to independently evolved this trait. On the other hand, spore-formation can be the result of gene transfer across unrelated bacterial genera, a phenomenon that has been proposed as occurring commonly among hyperthermophiles at geothermal environments [51]. We have found a previous report suggesting the production of spores in *Serratia marcescens* subsp. *sakuensis* in response to heat-shock [29]. This strain was isolated from a wastewater treatment tank, alongside several strains belonging to the genus *Bacillus*, which was also the case of strain Lr5/4. Although the authors proposed the scenarios indicated above (sporulation as an undetected trait or gene transfer from endospore-

forming Firmicutes co-inhabiting the tank) as likely explanations for the production of spores, none of these scenarios was studied further.

Interestingly, it has been shown that sporulation genes can be found in genomes of non-sporulating species [9]. However, these species do not have the ability to sporulate as they have only acquired part of the required machinery, while the formation of a fully functional spore, is an exquisitely complex process requiring not less than 60 genes acting at different stages [11]. Whether the sporulation of *S. ureilytica* strain Lr5/4 is or not a Firmicutes-like process needed to be addressed. Therefore, we investigated the genetic imprints of sporulation in the genome of strain Lr5/4.

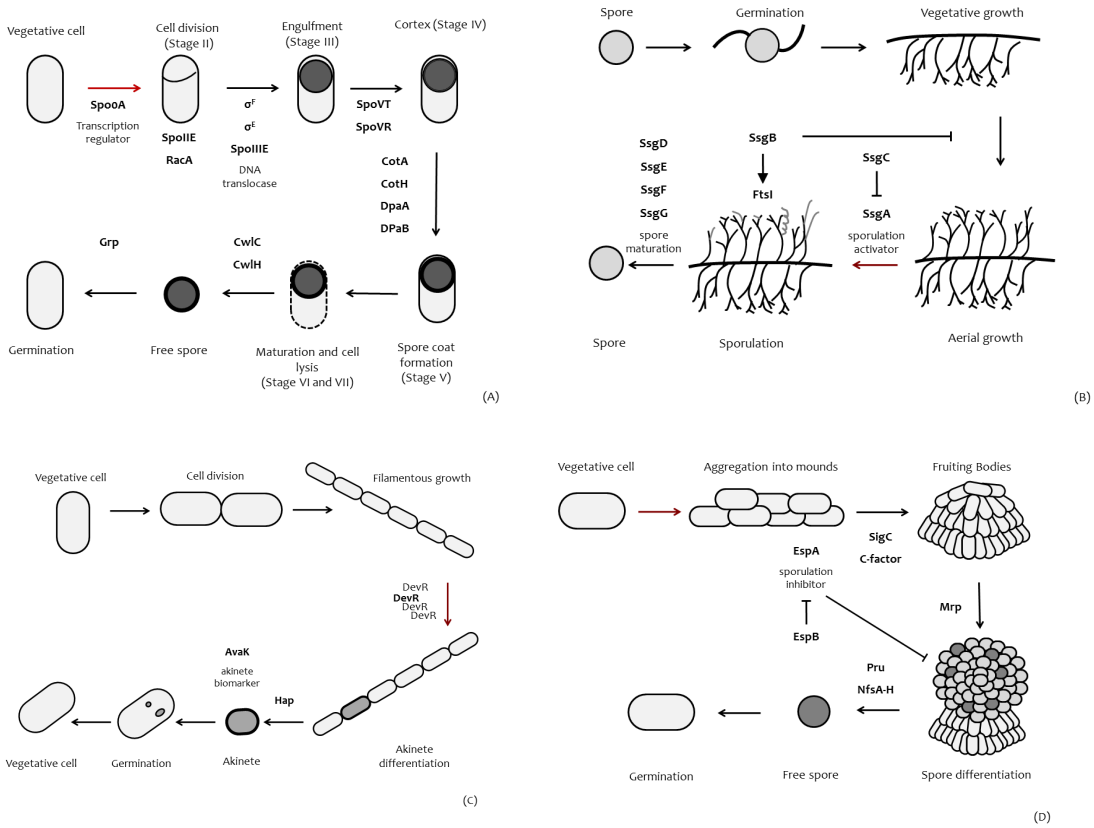


Figure 4.3. Sporulation process in four known taxa and major genes involved in sporulation pathways. Red arrows indicate commitment to sporulation. (A) Firmicutes; (B) Actinomycetes; (C) Cyanobacteria; (D) *Myxococcus* (sporulation due to starvation).

4.3.4. Genomic imprints of sporulation

The full genome of strain Lr5/4 was obtained and annotated. It was estimated to be of 5.39 Mbp in size with 59.2% G+C content. The complete genome sequence contained 5,056 genes, 22 rRNAs (5S, 16S, and 23S), 88 tRNAs, and 1 noncoding RNA (ncRNA) predicted. Based on the current knowledge of the developmental program of sporulation in known spore-forming bacterial groups (Figure 4.3), four sporulation proteome databases were built for each one of the pathways known in the different bacterial groups (Supplementary Table 4.5).The presence of homologues to proteins related to sporulation pathways of Firmicutes, Actinomycetes, Cyanobacteria and *Myxococcus* spp. was then evaluated in the genome of strain Lr5/4.Significant gene similarity was only detected for the spore proteome of Firmicutes (81 genes from 58 different species).

Moreover, the manual verification of these similarities showed that these genes correspond to the different stages of endospore-formation related to asymmetrical cell division and translocation of DNA, core and cortex assembly, and coat formation (Supplementary Table 4.6). These genes along with the GC content plot of the *S. ureilytica* str. Lr5/4genome are shown in Figure 4.4.

The genetic information retrieved from the genome of strain Lr5/4 revealed the presence of 26 spore-related proteins. This number is lower than the proposed set of 60 genes essential to sporulation in Bacilli and Clostridia [52]. However, parts of the molecular pathway of sporulation can be reconstructed from the genomic evidence. In Bacilli, detection of adversity in the microenvironment of the bacterial cell is principally monitored by the Kin protein family, as part of a two-component system [53]. A homolog to KinA was detected in the genome of *Serratia* Lr5/4, but also in the genome of *Serratia marcescens* that was also analyzed. KinA phosphorylates Spo0F, responsible for the regulation of Spo0A and SigH [54]. Obg, a GTP-binding protein involved in the initiation of sporulation [55], was also present in both proteomes. An homolog to the DNA transport protein SpoIIIE, which is involved in translocation of DNA to the forming spore in *B. subtilis*, was present in the *S. ureilytica* str. Lr5/4 proteome, but not in the proteome of *S. marcescens*. During the stage of asymmetrical cell division, another gene, *spoIIIJ* encoding for a membrane protein translocase, was also detected in *S. ureilytica* str. Lr5/4 genome. Among the genes believed to be involved in the formation of the spore cortex in *B. subtilis* [52], only *spoVD* and *spoVE* were identified in both genomes. CotA, related to coat assembly is the only protein of the Cot family detected in both proteomes. Finally, a protein related to germination, GdH, was also detected in both proteomes.

The operon for the formation of DPA synthetase subunits A and B (*spoVFAB*) was found in *S. ureilytica* str. Lr5/4 genome but not in the genome of *S. marcescens.* The genomic information is not only in agreement with the biochemical measurements indicating the presence of DPA in *S. ureilytica* str. Lr5/4 spores, but also with the data showing the resistance to wet heat in the spores of Lr5/4.

In order to verify whether these genes could have been acquired through horizontal gene transfer, protein trees were built for the putative homologs and these trees were compared with a phylogenetic tree of species using the software HiDe, which allows to calculate the likelihood and direction of gene transfer based on the topology of the trees produced. The results showed evidence of unidirectional gene transfer from thermophilic Clostridia to Lr5/4 (Supplementary Figure 4.3).

4.4. Discussion

The discovery of a novel spore-forming strain is presented in this study. Few other similar discoveries have been previously made, as in the case of *Rhodobacter johrii* [56], a novel species belonging to the α-Proteobacteria that produced minuscule spores refracting light in the contrast-phase microscope and staining with malachite green. Another example is that of *Serratia marcescens* subsp. *sakuensis*, isolated from a wastewater tank, found to produce concentric spores, similar to those of Bacilli [57].

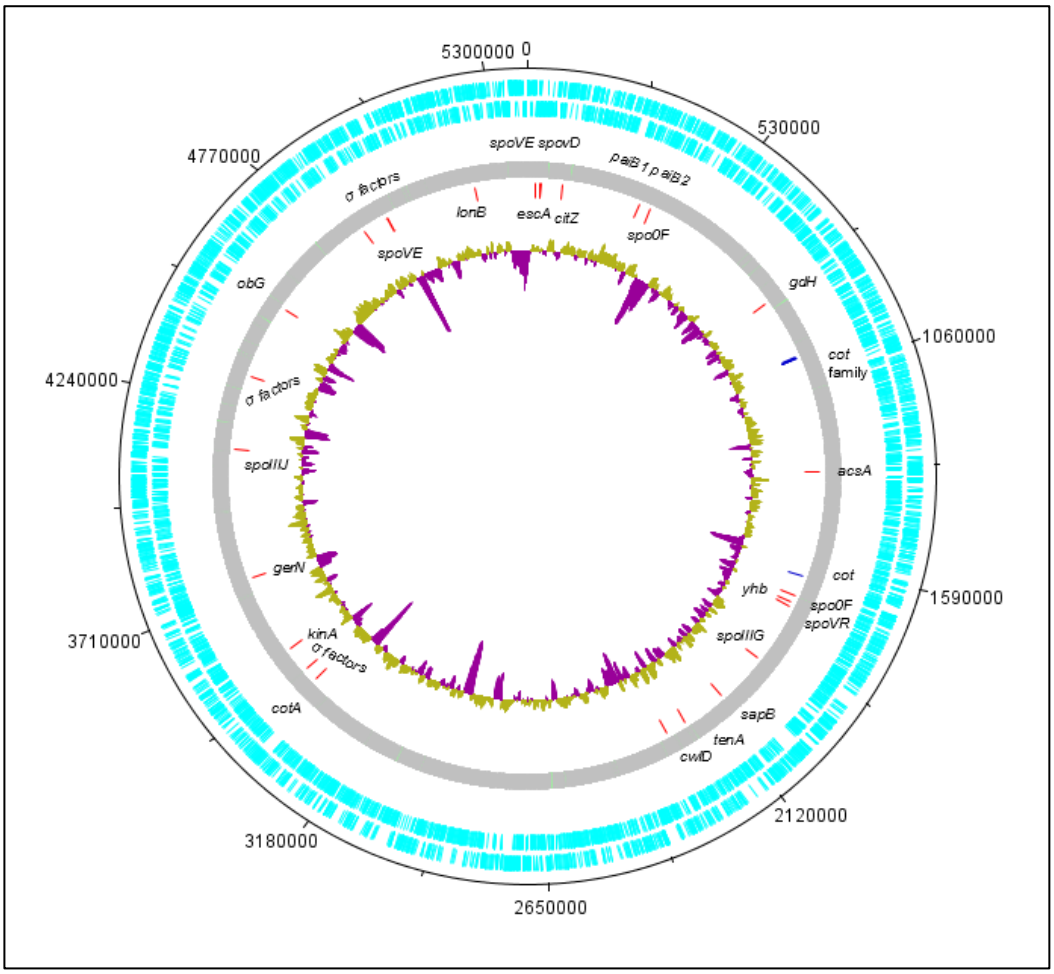


Figure 4.4. Sporulation genes suggested to be transferred from Firmicutes are spread throughout the chromosome of *Serratia ureilytica* str. Lr5/4. In light blue, all CDS identified in the genome. In red, homologs identified by tBLASTn. In dark blue homologs identified by automated annotation. The inner cycle represents C+C content. In the case of *S. ureilytica* str. Lr5/4, G+C content does not seem to be related to the genes that were transferred.

In both cases, microscopic evidence has been provided; however, the genetic mechanism responsible for sporulation had not been studied. A third example, which has been at the center of a debate, is the assumption that Mycobacteria are able to produce endospores similar to those of Firmicutes [58]. Although data on mRNA expression of some key sporulation-related genes was provided, a posterior study on the genomes of Mycobacteria revealed the absence of the core set of sporulation genes [59]. Moreover, spore production in the Mycobacteria strains was not reproducible [59]. Sporulation outside the four known taxa has always been viewed with skepticism. On the one hand, microscopic evidence may not be sufficient for determining if a strain is a spore former. On the other, presence of sporulation genes alone cannot be considered as conclusive evidence of sporulation capability, since many bacterial species possess some sporulation genes but are apparently asporogenic [60]. Herein, we provide evidence for the formation of a resistance structure in *S. ureilytica* str. Lr5/4 that is produced after nutrient starvation, heat-shock and UV radiation. This structure does not resemble the complex concentric structure of Bacilli or Clostridia spores, but its complexity is sufficient to make

it impermeable to common stains and to provide resistance to various environmental insults. This structure is significantly different in terms of size, shape and resistance to the vegetative cells it derived from, thereby fulfilling the definition of a spore. Its resistance to some environmental factors is comparable to that of Firmicutes.

Endospore formation, a trait that is unique to Firmicutes, is believed to have emerged in the common ancestor of Clostridia and Bacilli about 2.3 billion years ago. From this ancestor, both endospore-formers and asporogenic Firmicutes are supposed to have diverged into the diversity of groups found nowadays [61]. The reason for the irregular distribution pattern of endospores across Firmicutes is the energetic cost of this strategy. Spore formation represents a formidable investment of time and energy and it is defined as a survival pathway of last resort [62]. In fact, losing the ability to form spores is considered to be favored in stable environments in which this complex and energetically demanding process will not bring any significant ecological advantage [62]. Furthermore, given the large number of genes that are essential to ensure the formation of endospores [63], it has previously appeared improbable that sporulation genes could have been "shared" between different types of microorganisms. The overlap of the sporulation pathway for *S. ureilytica* strain Lr5/4 to the genetic components found in Firmicutes is thus surprising. The production of spores within the vegetative cell of *S. ureilytica* str. Lr5/4,(i.e. *end*ospores)has so far not been observed in our cultures. Alternative scenarios for the production of spore-like resistant structures may be in place. The hypothesis that the strain uses part of its cell division machinery together with some genes with a potentially foreign origin to shrink in size and resist environmental conditions is more plausible. The step-by-step process, as well as the gene expression in each step of this procedure still needs to be characterized. However, a spore-like structure as the one described in this study for *S. ureilytica* str. Lr5/4, although simpler than those of Firmicutes, appears to be equally suitable as a response for survival against extreme and fluctuating environmental conditions.

4.5. Acknowledgments

The whole-genome project of Lr5/4 has been deposited at GenBank under the BioProject IDs PRJNA224116 and PRJNA260750. For the sampling campaign we would like to thank Dr. Ludovic Russel-Delif. We acknowledge funding from the Swiss National Science Foundation project 31003A_152972, and from Fondation Pierre Mercier pour la science. SF: performed physiological analysis and identification, wrote the manuscript; TJ: performed bioinformatics analysis; NJ: performed isolation and participated in the physiological characterization; TW: participated to the sampling and helped in the isolation; WMK: helped for the collection of physiological data; VM: participated to the sampling; RL: performed the MALDI-TOF analysis; JS: performed the lipid analysis: SJ: performed the automatic genome annotation; PSC: supervised the automatic genome annotation; CD: coordinated access to the sampling site and participated to the *in situ* collection of data and sampling; PJ: designed the study, participated to the data analysis and wrote the manuscript.

4.6. References

[1] Kreft JU. Biofilms promote altruism. Microbiol-Sgm 2004; 150:2751–60. doi:10.1099/mic.0.26829-0.

[2] Justice SS, Hunstad DA, Cegelski L, Hultgren SJ. Morphological plasticity as a bacterial survival strategy. Nat Rev Microbiol 2008; 6:162–8. doi:10.1038/nrmicro1820.

[3] Zhao H, Msadek T, Zapf J, Madhusudan, Hoch JA, Varughese KI. DNA complexed structure of the key transcription factor initiating development in sporulating bacteria. Structure 2002; 10:1041–50. doi:10.1016/S0969-2126(02)00803-1.

[4] Lennon JT, Jones SE. Microbial seed banks: the ecological and evolutionary implications of dormancy. Nat Rev Microbiol 2011; 9:119–30. doi:10.1038/nrmicro2504.

[5] Buerger S, Spoering A, Gavrish E, Leslin C, Ling L, Epstein SS. Microbial scout hypothesis, stochastic exit from dormancy, and the nature of slow growers. Appl Environ Microbiol 2012;78:3221–8. doi:10.1128/AEM.07307-11.

[6] Wolf DM, Vazirani VV, Arkin AP. Diversity in times of adversity: probabilistic strategies in microbial survival games. J Theor Biol 2005;234:227–53. doi:10.1016/j.jtbi.2004.11.020.

[7] Driks A. Overview: Development in bacteria: spore formation in *Bacillus subtilis*. Cell Mol Life Sci CMLS 2002;59:389–91.

[8] Barton L. Structural and Functional Relationships in Prokaryotes. Springer Science & Business Media; 2005.

[9] Onyenwoke RU, Brill JA, Farahi K, Wiegel J. Sporulation genes in members of the low G+C Gram-type-positive phylogenetic branch (Firmicutes). Arch Microbiol 2004; 182:182–92. doi:10.1007/s00203-004-0696-y.

[10] Traag BA, Pugliese A, Eisen JA, Losick R. Gene conservation among endospore-forming bacteria reveals additional sporulation genes in *Bacillus subtilis*. J Bacteriol 2013;195:253–60. doi:10.1128/JB.01778-12.

[11] Galperin MY, Mekhedov SL, Puigbo P, Smirnov S, Wolf YI, Rigden DJ. Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. Environ Microbiol 2012;14:2870–90. doi:10.1111/j.1462-2920.2012.02841.x.

[12] Wu M, Ren Q, Durkin AS, Daugherty SC, Brinkac LM, Dodson RJ, et al. Life in Hot Carbon Monoxide: The Complete Genome Sequence of *Carboxydothermus hydrogenoformans* Z-2901. PLoS Genet 2005;1. doi:10.1371/journal.pgen.0010065.

[13] Chivian D, Brodie EL, Alm EJ, Culley DE, Dehal PS, DeSantis TZ, et al. Environmental Genomics Reveals a Single-Species Ecosystem Deep Within Earth. Science 2008; 322:275–8. doi:10.1126/science.1155495.

[14] Filippidou S, Jaussi M, Jeanneret N, Roussel-Delif L, Wunderlin T, Vieth-Hillebrand A, et al. Description of *Anoxybacillus geothermalis* sp. nov. , an aerobic endospore-forming bacterium isolated from mineral deposits in the Groß Schönebeck geothermal station, Germany. *Submitted at IJSEM.*

[15] Brewer JH. Clear liquid mediums for the "aerobic" cultivation of anaerobes. J Am Med Assoc 1940;115:598–600.

[16] Guckert JB, Antworth CP, Nichols PD, White DC. Phospholipid, ester-linked fatty acid profiles as reproducible assays for changes in prokaryotic community structure of estuarine sediments. FEMS Microbiol Lett 1985;31:147–58. doi:10.1111/j.1574-6968.1985.tb01143.x.

[17] Ichihara K 'ichi, Fukubayashi Y. Preparation of fatty acid methyl esters for gas-liquid chromatography. J Lipid Res 2010;51:635–40. doi:10.1194/jlr.D001065.

[18] Cashion P, Holderfranklin M, McCully J, Franklin M. Rapid Method for Base Ratio Determination of Bacterial DNA. Anal Biochem 1977;81:461–6. doi:10.1016/0003-2697(77)90720-5.

[19] Mesbah M, Premachandran U, Whitman WB. Precise Measurement of the G+C Content of Deoxyribonucleic Acid by High-Performance Liquid Chromatography. Int J Syst Bacteriol 1989;39:159–67. doi:10.1099/00207713-39-2-159.

[20] Tamaoka J, Komagata K. Determination of DNA base composition by reversed-phase high-performance liquid chromatography. FEMS Microbiol Lett 1984;25:125–8. doi:10.1111/j.1574-6968.1984.tb01388.x.

[21] DeLey J, Cattoir H, Reynaert, A. Quantitative Measurement of DNA Hybridization from Denaturation Rates. Eur J Biochem 1970;12:133 – &.

[22] Huss VA, Festl H, Schleifer KH. Studies on the spectrophotometric determination of DNA hybridization from renaturation rates. Syst Appl Microbiol 1983;4:184–92. doi:10.1016/S0723-2020(83)80048-4.

[23] Muyzer G, Teske A, Wirsen C, Jannasch H. Phylogenetic-Relationships of *Thiomicrospira* Species and their Identification in Deep-Sea Hydrothermal Vent Samples by Denaturing Gradient Gel-Electrophoresis of 16S rDNA Fragments. Arch Microbiol 1995;164:165–72. doi:10.1007/BF02529967.

[24] Forney L, Daae F, Torsvik V. Distribution of bacterioplankton in meromictic Lake Saelenvannet, as determined by denaturing gradient gel electrophoresis of PCR-amplified gene fragments coding for 16S rRNA. Appl Environ Microbiol 1997;63:3367–73.

[25] Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25:3389–402. doi:10.1093/nar/25.17.3389.

[26] Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 2008;36:W465–9. doi:10.1093/nar/gkn180.

[27] Wunderlin T, Junier T, Roussel-Delif L, Jeanneret N, Junier P. Stage 0 sporulation gene A as a molecular marker to study diversity of endospore-forming Firmicutes. Environ Microbiol Rep 2013;5:911–24. doi:10.1111/1758-2229.12094.

[28] Gerhardt P. Methods for general and molecular bacteriology. Washington, D.C.: American Society for Microbiology; 1994.

[29] Ajithkumar B, Ajithkumar VP, Iriye R, Doi Y, Sakai T. Spore-forming *Serratia marcescens* subsp. *sakuensis* subsp. nov., isolated from a domestic wastewater treatment tank. Int J Syst Evol Microbiol 2003;53:253–8.

[30] Brandes Ammann A, K&#246, Lle L, Brandl H. Detection of Bacterial Endospores in Soil by Terbium Fluorescence. Int J Microbiol 2011;2011. doi:10.1155/2011/435281.

[31, 33]Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. Bioinformatics 2005;21:3422–3. doi:10.1093/bioinformatics/bti553.

[32, 34] Hemmerich C, Buechlein A, Podicheti R, Revanna KV, Dong Q. An Ergatis-based prokaryotic genome annotation web server. Bioinforma Oxf Engl 2010;26:1122–4. doi:10.1093/bioinformatics/btq090.

[35] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403–10. doi:10.1016/S0022-2836(05)80360-2.

[36] Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 2002;30:3059–66.

[37] Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol 2010;59:307–21. doi:10.1093/sysbio/syq010.

[38] GNU Parallel: The Command-Line Power Tool | USENIX n.d. https://www.usenix.org/publications/login/february-2011-volume-36-number-1/gnu-parallel-command-line-power-tool (accessed February 27, 2016).

[39] Junier T, Zdobnov EM. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinforma Oxf Engl 2010;26:1669–70. doi:10.1093/bioinformatics/btq243.

[40] Bansal MS, Banay G, Harlow TJ, Gogarten JP, Shamir R. Systematic inference of highways of horizontal gene transfer in prokaryotes. Bioinforma Oxf Engl 2013;29:571–9. doi:10.1093/bioinformatics/btt021.

[41] Wayne L, Brenner D, Colwell R, Grimont P, Kandler O, Krichevsky M, et al. Report of the Ad-Hoc-Committee on Reconciliation of Approaches to Bacterial Systematics. Int J Syst Bacteriol 1987;37:463–4.

[42] Bhadra B, Roy P, Chakraborty R. Serratia ureilytica sp. nov., a novel urea-utilizing species. Int J Syst Evol Microbiol 2005;55:2155–8. doi:10.1099/ijs.0.63674-0.

[43] Lechevalier HA, Lechevalier MP, Holbert PE. Electron Microscopic Observation of the Sporangial Structure of Strains of Actinoplanaceae. J Bacteriol 1966;92:1228–35.

[44] Müller S, Strack SN, Ryan SE, Kearns DB, Kirby JR. Predation by *Myxococcus xanthus* Induces *Bacillus subtilis* To Form Spore-Filled Megastructures. Appl Environ Microbiol 2015;81:203–10. doi:10.1128/AEM.02448-14.

[45] Henriques AO, Melsen LR, Moran CP. Involvement of Superoxide Dismutase in Spore Coat Assembly in *Bacillus subtilis*. J Bacteriol 1998;180:2285–91.

[46] Stevenson KE, Vaughn RH, Crisan EV. Fixation of Mature Spores of *Clostridium botulinum*. J Bacteriol 1972;109:1295–7.

[47] Hoeniger JFM, Headley CL. Cytology of Spore Germination in *Clostridium pectinovorum*. J Bacteriol 1968; 96:1835–47.

[48] Vreeland RH, Rosenzweig WD, Powers DW. Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal. Nature 2000;407:897–900. doi:10.1038/35038060.

[49] Nicholson WL, Munakata N, Horneck G, Melosh HJ, Setlow P. Resistance of *Bacillus* endospores to extreme terrestrial and extraterrestrial environments. Microbiol Mol Biol Rev MMBR 2000;64:548–72.

[50] Setlow B, Atluri S, Kitchel R, Koziol-Dube K, Setlow P. Role of dipicolinic acid in resistance and stability of spores of *Bacillus subtilis* with or without DNA-protective alpha/beta-type small acid-soluble proteins. J Bacteriol 2006;188:3740–7. doi:10.1128/JB.00212-06.

[51] van Wolferen M, Ajon M, Driessen AJM, Albers S-V. How hyperthermophiles adapt to change their lives: DNA exchange in extreme conditions. Extrem Life Extreme Cond 2013;17:545–63. doi:10.1007/s00792-013-0552-6.

[52] Galperin MY, Mekhedov SL, Puigbo P, Smirnov S, Wolf YI, Rigden DJ. Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. Environ Microbiol 2012;14:2870–90. doi:10.1111/j.1462-2920.2012.02841.x.

[53] Stephenson K, Hoch JA. Evolution of signalling in the sporulation phosphorelay. Mol Microbiol 2002;46:297–304. doi:10.1046/j.1365-2958.2002.03186.x.

[54] Predich M, Nair G, Smith I. *Bacillus subtilis* early sporulation genes kinA, spo0F, and *spo0A* are transcribed by the RNA polymerase containing sigma H. J Bacteriol 1992;174:2771–8.

[55] Vidwans SJ, Ireton K, Grossman AD. Possible role for the essential GTP-binding protein Obg in regulating the initiation of sporulation in *Bacillus subtilis*. J Bacteriol 1995;177:3308–11.

[56] Girija KR, Sasikala C, Ramana CV, Spröer C, Takaichi S, Thiel V, et al. *Rhodobacter johrii* sp. nov., an endospore-producing cryptic species isolated from semi-arid tropical soils. Int J Syst Evol Microbiol 2010;60:2099–107. doi:10.1099/ijs.0.011718-0.

[57] Ajithkumar B, Ajithkumar VP, Iriye R, Doi Y, Sakai T. Spore-forming *Serratia marcescens* subsp. *sakuensis* subsp. nov., isolated from a domestic wastewater treatment tank. Int J Syst Evol Microbiol 2003;53:253–8. doi:10.1099/ijs.0.02158-0.

[58] Ghosh J, Larsson P, Singh B, Pettersson BMF, Islam NM, Sarkar SN, et al. Sporulation in mycobacteria. Proc Natl Acad Sci 2009;106:10781–6. doi:10.1073/pnas.0904104106.

[59] Traag BA, Driks A, Stragier P, Bitter W, Broussard G, Hatfull G, et al. Do mycobacteria produce endospores? Proc Natl Acad Sci 2010;107:878–81. doi:10.1073/pnas.0911299107.

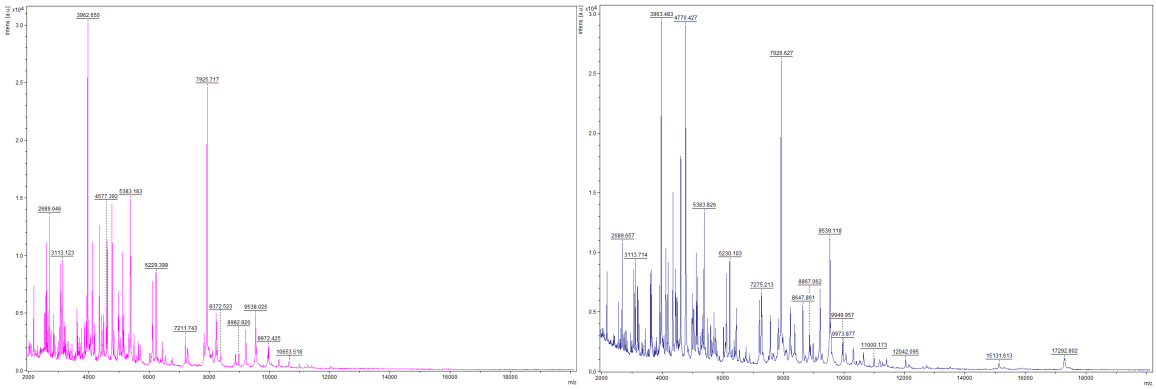
[60] Onyenwoke RU, Brill JA, Farahi K, Wiegel J. Sporulation genes in members of the low G+C Gram-type-positive phylogenetic branch (Firmicutes). Arch Microbiol 2004;182:182–92. doi:10.1007/s00203-004-0696-y.

[61] Galperin MY. Genome Diversity of Spore-Forming Firmicutes. Microbiol Spectr 2013;1:TBS – 0015–2012. doi:10.1128/microbiolspectrum.TBS-0015-2012.

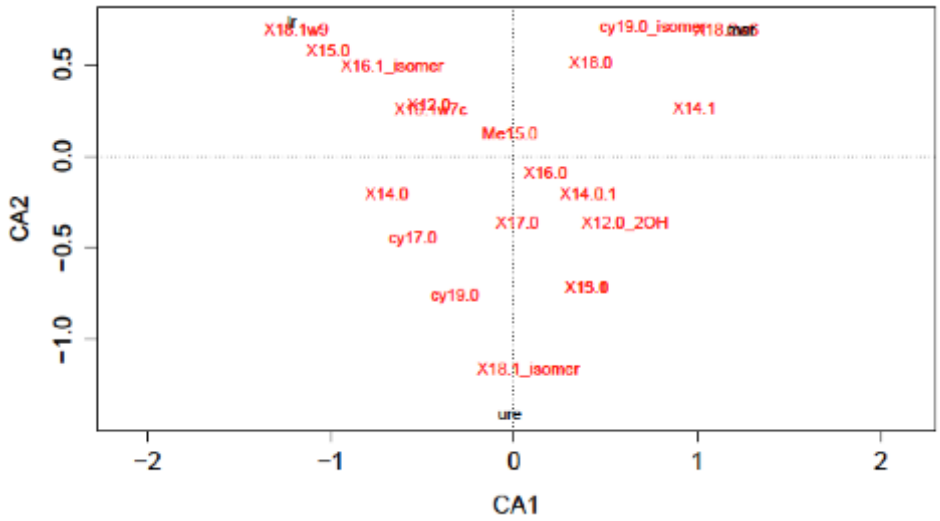
[62] de Hoon MJ, Eichenberger P, Vitkup D. Hierarchical evolution of the bacterial sporulation network. Curr Biol CB 2010;20:R735–45. doi:10.1016/j.cub.2010.06.031.
[63] Galperin MY. Genome Diversity of Spore-Forming Firmicutes. Microbiol Spectr 2013;1:TBS – 0015–2012. doi:10.1128/microbiolspectrum.TBS-0015-2012.

4.7 Supplementary Material

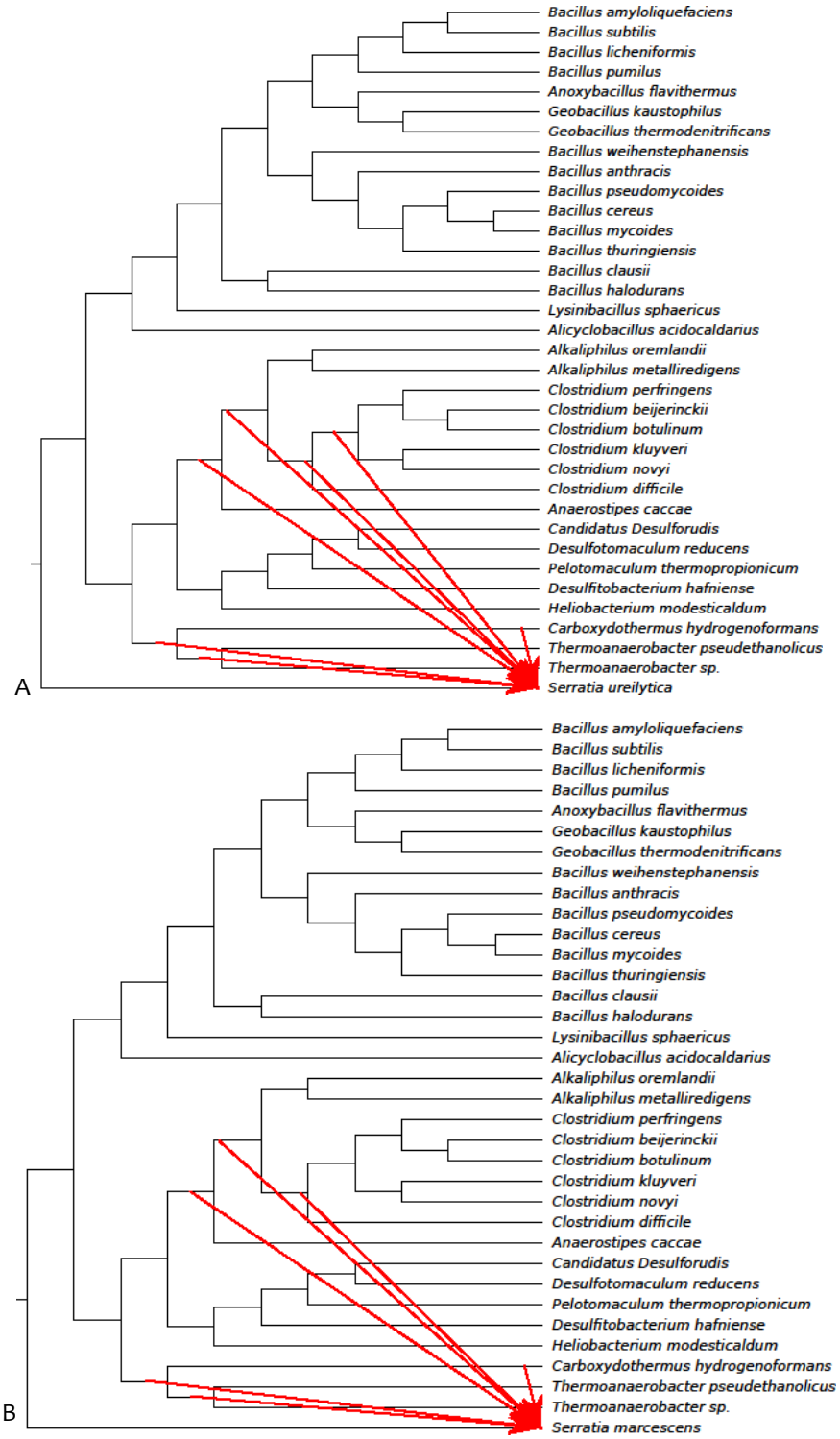
Supplementary Figure 4.1. MALDI-TOF mass spectrometry analysis for the identification of Serratia sp. strain Lr5/4. A, Direct analysis B, Analysis of the strain extraction



Supplementary Figure 4.2. CCA analysis of the fatty acid profiles of 1, *S. ureilytica* Lr5/4; 2, *S. ureilytica* DSMZ16952; 3, *S. marcescens* subsp. *marcescens* DSMZ30121



Supplementary Figure 4.3. Horizontal gene transfer to (A) *Serratia* sp. strain Lr5/4, (B) *S.marcescens*. Details on how the trees were constructed, what red arrows show.



Supplementary Table 4.1. Fatty acid analysis of *S. ureilytica* Lr5/4 and closely related strains *S. ureilytica* DSMZ16952 and *S. marcescens* subsp. *marcescens* DSMZ30121

Fatty acid	Systematic (common) name	*S. ureilytica* Lr5/4	*S. ureilytica* DSMZ16952	*S. marcescens* subsp. *marcescens* DSMZ30121
12:00	dodecanoic (lauric) acid	0,2	0,04	0,07
12:02	dodecadienoic acid	-	-	-
13:00	tridecanoic acid	0	0,07	0,02
OH-14:0	hydroxy-tetradecanoic acid	-	-	-
14:00	tetradecanoic (myristic) acid	0,79	3,27	3,46
14:1n5c	cis-7-tetradecenoic (myristoleic) acid	1,25	0,03	0,08
15:00	pentadecanoic acid	0,75	1,19	0,59
15:1n5	cis-10 pentadecenoic acid	0	0,06	0,04
OH-16:0	hydroxyhexaadecanoic acid	-	-	-
16:00	hexadecanoic (palmitic) acid	34,7	53,7	55
Me-15:0	14-methylpenadecanoic (isopalmitic) acid	0,19	0,14	0,18
16:1n7c	cis-9-hexadecenoic (palmitoleic) acid	22,6	8,5	8
16:1n10c	cis-7 (?)-hexadecenoic acid	0,67	0,1	0,16
17:00	heptadecanoic (margaric) acid	0,69	1,47	0,75
17:01	heptadecenoic acid	-	-	-
cy17:0	2-hexyl-cyclopentanoic acid	11,9	14,7	8,3
18:00	octadecanoic (stearic) acid	3,82	1,2	1,5
18:1n7c	cis-11 -octadecenoic acid	0,29	0	0,06
18:1n9c	cis-9-octadecenoic (oleic) acid	21	13,3	19,8
18:2n6C	cis-9-12-octadecadienoic (linoleic) acid	0,66	0	0,08
cy19:0	2-octyl-cyclopropaneoctanoic acid	0,44	2,26	1,85
19:01	nonadecenoic acid	-	-	-

Supplementary Table 4.2. Summary of sporulation triggering factors for the four spore-producing taxa and *S. ureilytica* Lr5/4

	Firmicutes	Cyanobacteria	Actinobacteria	*Myxococcus*	*S. ureilytica* Lr5/4
Starvation (C)	yes	yes	yes	yes	yes
pH	yes	N.D.	no	no	no
UV	yes	N.D.	no	no	yes
Temperature	yes	low	no	no	yes
Salinity	yes	no	no	no	no
Metal concentration	yes	no	no	no	no
Pressure	yes	no	no	no	no
Chemicals	yes	N.D.	N.D.	yes	no

Supplementary Table 4.3. Resistance of *S. ureilytica* Lr5/4 spores to environmental factors in comparison to *Bacillus subtilis*.

Factors	*Bacillus subtilis*	*S. ureilytica* Lr5/4
UV	(at 120 min)	(at 240 min)

Temperature °C	100	90
Desiccation	12h	72h

Supplementary Table 4.4. Environmental factors that have been tested for the survival of *S. ureilytica* Lr5/4, *Bacillus subtilis, Escherichia coli, S. ureilytica* DSMZ16952 and *S. marcescens* subsp. *marcescens* DSMZ30121

Factors	Experimental design	*S. ureilytica* Lr5/4	*Bacillus subtilis*	*Escherichia coli*	*S. ureilytica* DSMZ16952	*S. marcescens* subsp. *marcescens* DSMZ30121
Starvation (C)	-solid culture for 90 days	yes	yes	no	no	no
UV	-UV radiation for 5, 10, 20, 30, 60, 120, 180 min and 12 h	yes	yes	killed after 30 min	N.D.	N.D.
		(at 12h)	(at 120 min)		N.D.	N.D.
Temperature	-heat-shock test	yes	yes	no	no	no
	-after growth, incubation for 24 h at	no	N.D.	killed	N.D.	N.D.
	-80, -20, 4, 60, 80, 1000C	N.D.	N.D.	< 4 and >x	N.D.	N.D.
	-autoclaving	no	no	no	N.D.	no
Desiccation	-24h cultures for 72h in a desiccator	vegetative cells	yes	no	N.D.	N.D.
Chemicals	-after growth, incubation in a medium containing chemical substances	yes	yes	N.D.	N.D.	N.D.

Supplementary Table 4.5. Sporulation specific genes in every sporulation pathway. These genes were used for the construction of the four sporulation datasets.

protein	function
Cyanobacteria	
ArgC	N-acetyl-gamma-glutamyl-phosphate reductase
HrmA	
hetR	Heterocyst differentiation control protein
DevR	Response regulator receiver protein DevR
HepA	Heterocyst differentiation ATP-binding protein HepA
AvaK	Akinete-marker protein
Myxococcus	
ScgA	Developmental C-signal
SigC	RNA polymerase sigma-C factor
Tps	Development-specific protein
Pru	Protein U
MrpC	transcriptional regulator
MrpB	Sigma-54 dependent DNA-binding response regulator

MrpA	Sensor histidine kinase
EspB	membrane protein
EspA	histidine kinase
Actinomycetes	
SsgA	spore maturation
SsgE	spore maturation
SsgD	spore maturation
SsgG	spore maturation
Firmicutes	
AbrB	regulation of gene expression during the transition from growth to stationary phase
AcsA	utilization of acetate, fatty acids
ald	alanine utilization
AmiC	?
BclA	?
BofC	control of processing of pro-SigK by SpoIVFB
calY	?
CcdA	cytochrome c synthesis
cda	
cgeB	maturation of the outermost layer of the spore
CitB	TCA cycle
CitZ	TCA cycle
CotA	resistance of the spore
CotB	resistance of the spore
CotC	resistance of the spore
CotD	resistance of the spore
CotE	assembly of the outer spore coat
CotF	resistance of the spore
CotH	protection of CotU and CotC in the mother cell
CotI	spore envelope
CotJA	polypeptide composition of the spore coat
CotJB	polypeptide composition of the spore coat
CotJC	polypeptide composition of the spore coat
CotK	protection of spore DNA
CotM	resistance of the spore
CotN	biofilm formation
CotO	controls assembly of the coat layers and coat surface topography
CotP	resistance of the spore
CotPE	?
CotS	resistance of the spore
CotSA	resistance of the spore
CotT	resistance of the spore
CotV	resistance of the spore
CotW	resistance of the spore
CotX	spore coat assembly
CotY	spore coat assembly
CotZ	spore crust assembly
CoxA	resistance of the spore
CptPC	?

CsgA	spore germination
cspA	?
CtaA	heme biosynthesis
CtaF	cytochrome-c oxidase
CwlC	mother cell lysis
CwlD	spore cortex peptidoglycan synthesis
CwlJ	spore germination
DeoR	regulation of deoxyribonucleotide utilization
Dfp	?
EcsA	regulation of the secretion apparatus and of intra-membrane proteolysis
EcsC	?
EscB	regulation of the secretion apparatus and of intra-membrane proteolysis
FliA/WhiG	?
FtsI	?
GdH	germination
GerAA	germination response to L-alanine
GerAB	germination response to L-alanine
GerAC	germination response to L-alanine
GerBA	germination response to the combination of glucose, fructose, and KCl
GerBB	germination
GerBC	germination
GerC	menaquinone biosynthesis
GerD	germination
GerE	regulation of SigK-dependent gene expression
GerHA	?
GerHB	?
GerHC	?
GerIA	?
GerIB	?
GerIC	?
GerKA	germination response to the combination of glucose,fructose, aspartate, and KCl
GerKB	germination
GerKC	germination
GerLA	?
GerLB	?
GerLC	?
GerM	germination (cortex hydrolysis) and sporulation
GerN	?
GerPA	germination
GerPB	germination
GerPC	germination
GerPD	germination
GerPE	germination
GerPF	germination
GerQ	germination

GerSA	?
GerSB	?
GerSC	?
GerT	germination
GerW	germination protein
GerXA	?
GerXB	?
GerXC	?
GerYA	?
GerYB	?
GerYC	?
GlcU	?
Gpr	germination protease degradation of SASPs
KapD	inhibitor of the KinA pathway to sporulation
kbaA	effector of KinB activity
KinA	two-component sensor kinase
KinB	two-component sensor kinase
KinD	two-component sensor kinase
Lgt	lipomodification of lipoproteins
LonA	protein quality control, control of swarming motility
LonB	protein quality control
LonC	?
MecA	control of ComK degradation, regulation of competence
NapA	?
Noc	control of cell division
ObG	spo0B-associated GTP-binding protein involved in initiation of sporulation, robosome assembly
OxaA	membrane insertion of proteins and protein secretion
PaiA	control of intracellular polyamine concentrations
PaiB	regulation of sporulation, degradative enzyme and motility genes
ParA	forespore chromosome partitioning /negative regulation of sporulation initiation
PdaA	spore cortex peptidoglycan synthesis
PdaB	spore cortex formation
PepSY	?
PerM	?
RfbX	?
RpoD	RNA polymerase major sigma factor SigA
RpoE	RNA polymerase delta subunit
RsbW	control of SigB activity
RsfA	control of expression of SigF-dependent genes
SafA	spore coat formation
SapB	?
SasA3	small acid soluble proteins
SasH	small acid soluble proteins
SasO	small acid soluble proteins
SasP1	small acid soluble proteins
SasP2	small acid soluble proteins

Sasα/β	small acid soluble proteins
Sasγ	small acid soluble proteins
SCLE	?
Sda	serine utilization
SdpB	maturation of the SdpC toxin
SeaA	involved in spore envelope assembly
SgiI	?
SigE	transcription of sporulation genes (early mother cell)
sigF	transcription of sporulation genes (early forespore)
SigG	transcription of sporulation genes (late forespore)
SigH	transcription of early stationary phase genes
SigI	control of a class of heat shock genes
sigK	late mother cell-specific gene expression
SinI	control of biofilm formation
SleB	degration of the spore cortex, germination
SleC	?
SpIB	protection of spore DNA against photodamage
SpI	protection of spore DNA against photodamage
SpmA	spore maturation protein (spore core dehydratation)
SpmB	spore maturation protein (spore core dehydratation)
SpooA	master regulator for the initiation of sporulation
SpooB	signal transduction
SpooE	SpooA-P phosphatase
SpooF	regulates *spooA* and sigH
SpooJ	chromosome positioning before assymetric division, centromer binding
SpooM	?
SpoIIAA	control of sigF activity (anti-anti-sigF)
SpoIIAB	septation, phosphorylation and inactivation of SpoIIAA
SpoIIB	facilitator of septal dissolution
SpoIIC	dissolution of the septal cell wall
SpoIID	dissolution of the septal cell wall
SpoIIE	control of SigF activity required for normal formation of the asymmetric septum
SpoIIGA	maturation of SigE
SpoIIIAA	activation of SigG
SpoIIIAB	activation of SigG
SpoIIIAC	activation of SigG
SpoIIIAD	activation of SigG
SpoIIIAE	?
SpoIIIAF	activation of SigG
SpoIIIAG	activation of SigG
SpoIIIAH	activation of SigG, forespore encasement by the spore coat
SpoIIID	regulation of mother cell gene expression
SpoIIIE	DNA translocase FtsK/SpoIIIE. Chromosomal partitioning and orientation
SpoIIIF	membrane insertion of proteins and protein secretion
SpoIIIJ	membrane insertion of proteins and protein secretion

SpoIIM	dissolution of the septal cell wall
SpoIIP	dissolution of the septal cell wall
SpoIIQ	forespore encasement by the spore coat
SpoIIR	control of SigE activation
SpoIISA	programmed cell death
SpoIVA	spore cortex formation and coat assembly
SpoIVB	control of SigK activation
SpoIVFA	control of SigK activation
SpoIVFB	processing of pro-sigma-K to active SigK
SpoIVH	spore cortex formation
SpoVAA	spore maturation
SpoVAB	spore maturation
SpoVAC	spore maturation
SpoVAD	likely germination protein
SpoVAEA	spore germination
SpoVAEB	spore germination
SpoVAF	spore maturation
SpoVC	spore coat formation
SpoVD	spore morphogenesis
SpoVE	spore cortex peptidoglycan synthesis
SpoVFA	dpa synthetase subunit A
SpoVFB	dpa synthetase subunit B
SpoVG	cell division, control of sporulation initiation
SpoVID	spore coat assembly
SpoVK	spore maturation
SpoVM	spore cortex and coat synthesis
SpoVR	spore cortex synthesis
SpoVS	spore coat assembly, spore core dehydratation
SpoVT	regulation of forespore gene expression
SpsF	spore coat polysaccharide synthesis
SspB	protection of spore DNA
SspE	protection of spore DNA
SspF	protection of spore DNA
SspH	protection of spore DNA
SspI	protection of spore DNA
SspJ	protection of spore DNA
SspK	protection of spore DNA
SspL	protection of spore DNA
SspN	protection of spore DNA
SspO	protection of spore DNA
Sspγ	protection of spore DNA
TenA	thiaminase II
TenI	thiazole tautomerase
Tgl	introduction of crosslinks in the spore coat protein GerQ
Tlp	?
YaaH	survival of ethanol stress, protection of the spore
YaaT	regulation of sporulation initiation
YabG	modification of spore coat proteins
YabP	sporulation at a late stage

YabQ	sporulation at a late stage
YbaN	spore cortex formation
YcdA	swarming motility
YdbB	?
YfhP	?
YfkD	?
YfkQ	part of the YfkQ-YfkR-YfkT germinant receptor
YhbH	?
YhcM	?
YhcN	?
YheC	?
YitO	?
YitS	?
YjcC	?
YknT	spore coat protein
YkuD	cell wall biosynthesis
YkvP	?
YkvU	germination at high pressure
YkzQ	?
YlaJ	?
YlbF	regulation of biofilm formation and the phosphorelay
YlbJ	?
YlmC/YmxH	?
YlxY	?
YmaG	spore coat protein
YndD	part of the YndD-YndE-YndF germinant receptor
YndE	part of the YndD-YndE-YndF germinant receptor
YndF	part of the YndD-YndE-YndF germinant receptor
YodI	?
YpeB	assembly of SleB
YpjB	?
YqfC	?
YqfD	?
YqfQ	?
YsxE	inner spore coat protein
YtaF	?
YtcC	lipopolysaccharide biosynthesis
YtrH	?
YtrI	spore cortex formation
YtvI	?
YtxC	?
YunB	release of the forespore into the mother cell cytoplasm
YutH	spore coat protein
YyaC	?

Supplementary Table 4.6 Sporulation genes present in the genomes of *S. ureilytica* Lr5/4 and *S. marcescens.*

S. ureilytica Lr5/4	*S. marcescens*
amic	*ywcF*
citZ	*yhbH*
cotA	*spoVR*
cwlD	*spoVE*
ecsA	*spoVD*
gdH	*spoIIIJ*
gerN	*spooF*
kinA	*spooA*
lonB	*obg*
lonC	*lonC*
obg	*lonB*
oxaA	*kinA*
paiB1-2	*gerN*
sapB	*gdH*
spooF	*ftsI*
spoIIIE	*ecsA*
spoIIIG	*cwlD*
spoIIIJ	*citZ*
spoVD	*amic*
spoVE	*acsA*
spoVR	
tenA	
yhbH	

Chapter 5

Cryptosporulation in *Kurthia* spp. forces a rethinking of asporogenesis in Firmicutes

Sevasti Filippidou[a], Thomas Junier[a,b], Tina Wunderlin[a], Nicole Jeanneret[a], Ilona Palmieri[a], Jochen Blom[c], Shannon Johnson[d], Patrick Chain[d], Pilar Junier[a]#
Laboratory of Microbiology, Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland[a];
Vital-IT group, Swiss Institute of Bioinformatics, Lausanne, Switzerland[b]; Bioinformatics and Systems Biology, Justus-Liebig-University Giessen, Giessen, Germany[c];
Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA[d]

This chapter is to be submitted for publication at *mBio*

Abstract

Sporulation is a complex morphophysiological process resulting in a structure more resistant than the vegetative form of a bacterial cell. In Firmicutes, this structure is produced within the mother cell and is thus called an *endospore*. Endospore formation is thought to have evolved in the common ancestor of Firmicutes. However, sporulation has apparently been lost in some extant lineages. We isolated strain 11kri321, a representative of the genus *Kurthia*, from an oligotrophic geothermal reservoir. While *Kurthia* is considered asporogenic, the physiological and genomic analyses demonstrate that strain 11kri321 is an endospore-former. The genomic reconstruction of the sporulation pathway shows elements typical of sporulation in Bacilli, including signaling for sporulation onset. However, key genes were missing, including those in engulfment or dipicolinic acid synthesis. Accordingly, *Kurthia* sp. str. 11kri321 spores lack dipicolinic acid, which is a likely response to the alkaline environment from which the strain was isolated. Based on the analysis of strain 11kri321, evidence for sporulation was investigated in all publicly available *Kurthia* genomes. Genes involved in signaling, cell division and spore coat formation were detected in all *Kurthia* species. These results suggest that *Kurthia* is of an endospore-forming Firmicute lineage. The genetic background of sporulation in this genus deviates strongly from the known pathways in Firmicutes and even within Bacilli, suggesting a revision of the minimal set of genes required for sporulation. Based on our findings we propose the term *cryptosporulant* to refer to Firmicutes for which a detailed genomic and physiological characterization of sporulating capability is missing.

5.1. Introduction

Sporulation is a morphophysiological response to unfavorable environmental conditions, involving a sophisticated genetic pathway. This response is triggered by nutrient starvation (1), extreme heat, temperature shocks [1–3], high salinity, desiccation [4], ultraviolet radiation [3,4], pressure [4] and extreme pH [5]. These triggers are well studied and have been tested experimentally in the laboratory. However, what drives sporulation in natural environments does not always agree with *in vitro* tests. This process has been identified within only four bacterial phyla, and it results in a structure that is more resistant than the vegetative cell [6].

In Firmicutes, the process of sporulation is called endosporulation, because it is the result of an asymmetrical cell division that leads to the formation of a mature spore within a mother cell [7]. Endosporulation is thought to have emerged in the ancestor of Firmicutes, but appears to have been lost or become inactive in many extant descendants [8]. When sporulation is no longer possible but the genomic remnants of sporulation are still present in the genome, the phenotype of the organism is called *asporogenic*. The advent of genome sequencing technologies has opened the door to re-investigation of supposed asporogenic species. For example, the analysis of the genome of *Carboxydothermus hydrogenoformans* suggested that this species could produce spores, contrary to the previous physiological knowledge. This allowed researchers to experimentally demonstrate its sporulation capability [9]. This finding suggests that similar discoveries are awaiting further genomic and physiologic studies in other lineages of Firmicutes. Nowadays, a series of approximately 60 genes are proposed to encode the minimal core genetic suite for sporulation [10]. The inactivation or loss of some of these genes generates an asporogenic phenotype [11], although some alternative regulatory pathways to spore formation do exist [10]. This minimal set of genes has been key in the investigation of unusual endospore-formers such as segmented filamentous bacteria, which despite their small genomes possess the predicted set of core sporulation genes [12,13].

Although the mechanisms underlying the maintenance of endospore formation are still poorly understood, one can hypothesize that given the high energetic cost and genetic complexity of spore-production, under constant favorable conditions for growth, bacteria might lose sporulation genes and, as a result, the capability to produce spores [11]. Under unfavorable conditions, however, sporulation is a beneficial trait that improves survival and dispersal. Hence, habitats with unstable environmental conditions should harbor a larger diversity of endospore-formers than habitats with constant environmental conditions. Geothermal environments are a good example of the former because steep chemical and physical gradients are characteristic of this type of environment. As part of a large effort to study endospore-forming Firmicutes in geothermal environments, our laboratory initiated multiple enrichments to isolate aerobic and anaerobic strains. The strains were screened for their capability to form spores, without any prior bias regarding their phylogenetic affiliation. In this way, strain 11kri321 was isolated from the geothermal spring of Krinides, Kavala, Greece. The strain

was found to belong to genus *Kurthia*, which so far is classified as a non-spore-forming genus among Firmicutes. This genus was discovered in 1883 [14] and according to the list of prokaryotes [15] it now consists of five recognized species. Four full genomes of *Kurthia* spp. are publicly available [16] and analyzed herein along with the genome of *Kurthia* sp. str. 11kri321. In this study we explore the physiological and genomic evidence demonstrating sporulation as a trait in *Kurthia*. Furthermore, we discuss the implications of defining a truly asporogenic lifestyle for a lineage within Firmicutes, in opposition to those groups for which so far we lack a detailed genomic and physiological characterization.

5.2. Materials and methods

5.2.1. Sample collection and isolation

The water sample from the bore-pipe of the geothermal spring was collected in a sterile 1L bottle and filtered through a 0.22 μm nitrocellulose membrane (Millipore, USA). The membrane was transported to the laboratory on ice and stored at 4 °C for bacterial enrichment into 10 mL of Nutrient Broth (Biolife, Italy). The enriched culture was then plated on Nutrient Agar (NA) and single colonies were obtained. Each colony was plated repeatedly to attain pure aerobic bacterial isolates. Colony morphology was observed after 12 h of growth. The capability to form spores was observed after starvation for 15 days using phase-contrast microscopy (Leica DM R, magnification 1000x). A differential staining for endospores and vegetative cells has been performed using malachite green and safranine, as previously described [17].

Cell growth was monitored at different temperatures (4, 15, 25, 35, 45, 50, 55, and 60°C) over 4 days in nutrient broth medium. To determine the pH range in which *Kurthia* sp. str. 11kri321 grows, nutrient broth medium at pH 4 to 13 was prepared (intervals of 0.5 pH unit), and growth was monitored at optimum growth temperate (25°C), over 4 days. All tests were performed in triplicates.

5.2.2. Strain identification

Genomic DNA was extracted using the InnuPREP Bacteria DNA kit (Analytik Jena, Germany), according to the manufacturer's instructions. For amplification of 16S rRNA gene, the primers GM3F, GM4R and Eub9_27, Eub1542 were used as previously described [18,19]. The gene that encodes the transcriptional factor responsible for the sporulation initiation in endospore-forming bacteria (*spo0A* gene) was also amplified with the specific set of primers spo0A166f and spo0A748r, as described previously [20]. The PCR products were purified with a MultiScreen PCRμ96 Filter Plate (Millipore, USA), according to the manufacturer's instructions, and sequenced using the services of Microsynth AG (Switzerland). The 16S rRNA gene was identified using the online services of EzTaxon, against EzTaxon's cultured isolates database [21]. The sequence was submitted to GenBank under accession number KJ722471.

5.2.3. Phylogenetic analysis

16S rRNA gene sequences (>1200bp) of Firmicutes were retrieved from RDP (http://rdp.cme.msu.edu/) and aligned using the default parameters of MAFFT [22]. A maximum likelihood phylogenetic tree was built using PhyML [23] and then graphics using the Newick utilities [24].

5.2.4. gDNA extraction and sequencing

Genomic DNA was extracted from an overnight culture using the Genomic-tip 20/G kit (Qiagen GmbH, Germany). Sequencing was performed with the PacBio RS II system based on single molecule, real-time (SMRT) technology (Pacific Biosciences, California). The draft genome of the genome of *Kurthia* sp. str. 11kri321 presents a unique contig of 2964527 bases, and a G+C content of 36.7%. Genome annotation was performed using an Ergatis-based [25] workflow with minor manual curation and visualized with the Artemis Genome Browser and Annotation Tool [26]. A total of 2893 coding sequences (CDSs), 82 tRNAs, and 27 rRNAs (9 copies of 16S, 23S and 5S rRNA genes) were predicted. This whole-genome project has been deposited at GenBank under the Bioproject PRJNA301103, and the Biosample ID SAMN04235798.

5.2.5. Retrieval of sporulation genes sequences

Complete and draft sequences of spore-forming Firmicutes were downloaded from Comprehensive Microbial Resource (CMR) and Integrated Microbial Genome (IMG) websites. Search for spore-related genes was based on gene function category sporulation (CMR; sporulating category in IMG). The CMR version was 24.0 data release and the IMG version was 3.0. In addition to the protein sequence, nucleotide sequences including a 50-bp flanking region at both 5'- and 3'- ends were downloaded. Additional information on all retrieved genomes was obtained from the GenBank database.

5.2.6. Sequence data analysis

The *Kurthia* sp. str. 11kri321 genome sequence was scanned for orthologs of the Firmicute core sporulation genes (as protein sequences) with TBLASTN [27], using default parameters and an E-value cutoff of 1E-11. A TBLASTN run on the shuffled protein sequences as a negative control set showed no hit with an E-value lower than 4e-4. The hits were ordered by position on the *Kurthia* sp. str.11kri321 genome and inspected manually. The above procedure did not detect orthologs of SpoVFA and SpoVFB, therefore we attempted to detect those by pairwise dynamic-programming alignment. The protein sequences of SpoVFA and SpoVFB were each compared to all sequences of the *Kurthia* proteome using Needleman and Wunsch's [28] algorithm, as implemented by EMBOSS's `needle` program [29]. No hits were found. The publicly-available *Kurthia* genomes *Kurthia huakuii* LAM0618, *Kurthia massiliensis*, and *Kurthia* sp. JC8E were scanned for sporulation gene orthologs as described above.

The sequences of the *Kurthia* genomes analyzed herein were retrieved from GenBank under accession numbers: *Kurthia huakuii* LAM0618: NZ_AYTB00000000.1, *Kurthia massiliensis*: NZ_CAEU00000000.1, and *Kurthia* sp. JC8E: NZ_CAEW00000000.1.

5.2.7. DPA measurement

The presence of dipocolinic acid (DPA) in the spores was assessed, according to a previously published method [30]. Fluorescence was measured with a Perkin-Elmer LS50B fluorometer. The excitation wavelength was set at 272 nm with a slit width of 2.5 nm. Emission was measured at 545 nm (slit width 2.5 nm). The device was set in the phosphorescence mode (equivalent to time-resolved fluorescence). The delay between emission and measurement was set at 50 μs. Measurements were performed every 20 ms. The integration of the signal was performed over a duration of 1.2 ms. Values recovered for each measurement corresponded to the mean of the relative fluorescence unit (RFU) values given by the instrument within the 30 s following sample introduction in the device. Finally, to transform RFU units into DPA concentrations, a 10-point standard curve was established using increasing concentrations of DPA from 0.5 μM up to 10 μM

5.3. Results and discussion

5.3.1. Sampling and characterization of the isolate

The geothermal reservoir of Krinides (N 41° 00.642' E 024° 15.371'), near Philippoi, is situated in the Rhodope Massif (Eastern Macedonia, Greece) [17]. In ancient times, the exposed geothermal spring was used for bathing and recreation. The geothermal reservoir serves the same purposes nowadays, however due to tectonic activity over the centuries, the spring is no longer exposed and access to the geothermal water is facilitated through two borings. At the time of sampling, the water temperature at the output of the bore-pipe was 29.1 °C, pH was 9 and conductivity 415 μS/cm. Water from the outflow as well as biofilms were collected. Strain 11kri321 was isolated from the biofilm and was characterized as a Gram-positive, pigment-producing, spore-forming bacterium. Spores of strain 11kri321 were observed to refract light under the phase-contrast microscope and could be stained with malachite green (Figure 5.1).

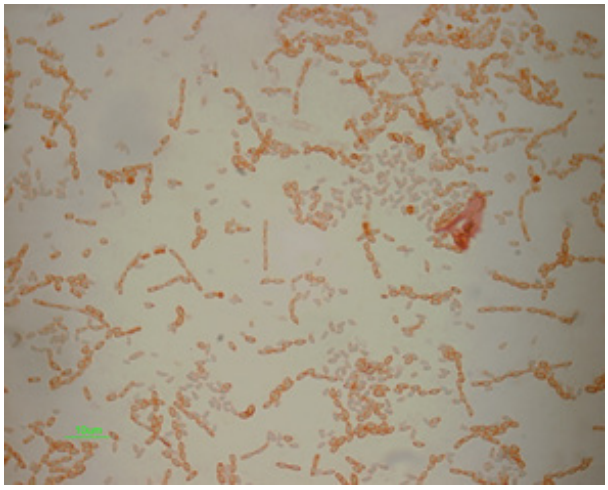


Figure 5.1. Malachite green staining of Kurthia strain 11kri321. Vegetative cells are stained in red and spores are stained in green.

The initial classification of strain 11kri321, based on the comparison of the full 16S rRNA gene sequence with known bacterial species, suggested that it belongs to the genus

Kurthia with 99.65% similarity to *Kurthia gibsonii* NCIMB 8758^T. Affiliation to this genus is also supported by average amino acid identity (AAI) analysis performed with other available *Kurthia* genomes. The AAI values for the comparison between strain 11kri321 are 68.88 % with *Kurthia massiliensis*, 68.50% with *Kurthia huakuii* and 68.58% with *Kurthia* sp. JC8E, which are all above the suggested threshold of 60% used for the definition of a bacterial genus [32]. The phylogenetic placement of strain 11kri321 was verified by comparing all sequences of Firmicutes available from the RDP database. This confirmed that strain 11kri321 forms a coherent cluster with other *Kurthia* spp. as part of the Bacilli (Figure 5.2).

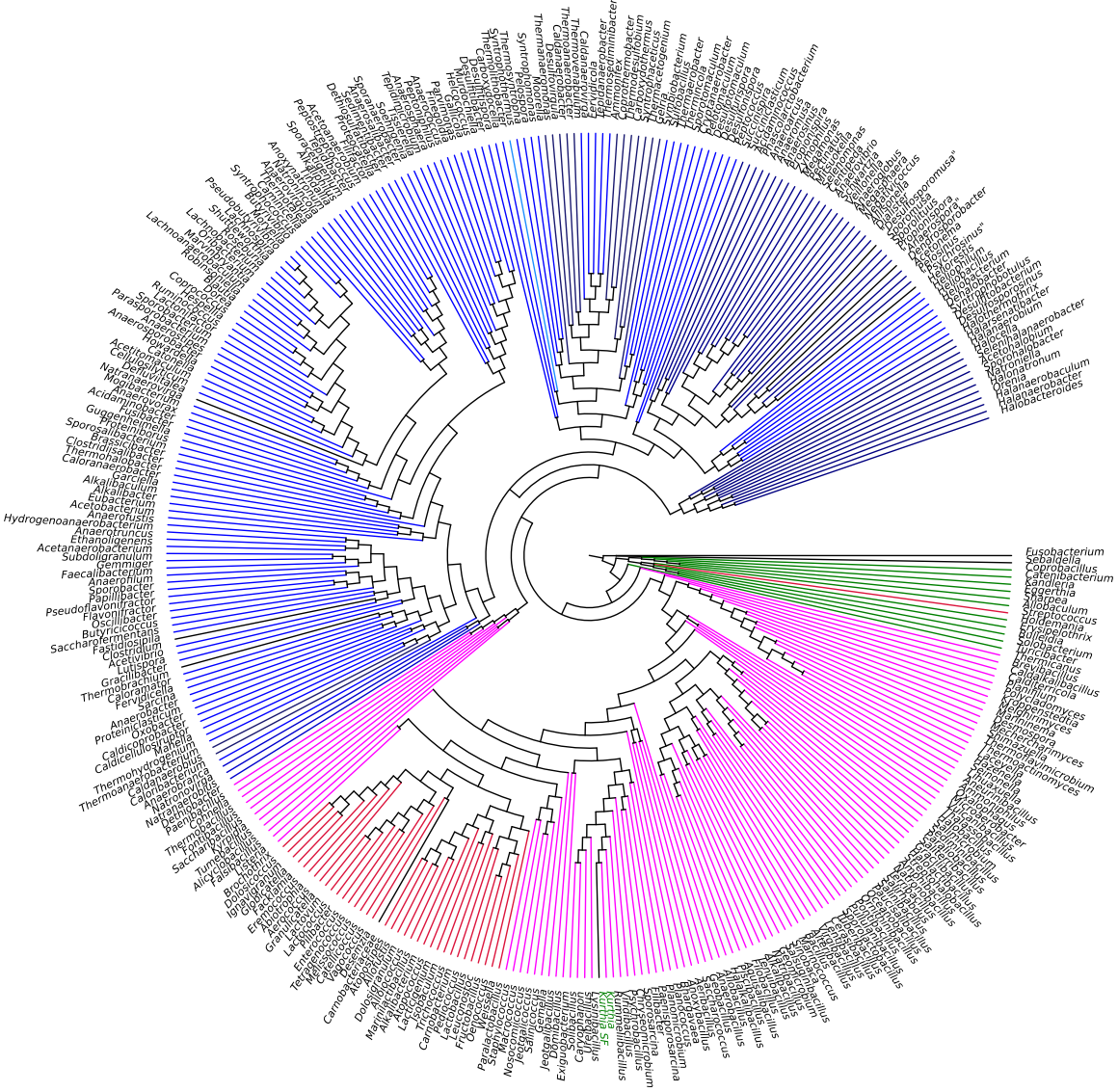


Figure 5.2. Maximum-likelihood tree showing the positioning of the *Kurthia* genus among endospore-forming Firmicutes

In addition, we performed a physiological characterization of the strain, in particular in relation to pH and temperature optima, in order to compare those with the *in situ* conditions. The *in situ* pH of 9 is at the limit of tolerance for all previously described *Kurthia* spp.[33] and the *in situ* temperature (29°C) is the optimum reported so far for growth for this genus [33]. Strain 11kri321 grows at a pH between 5.5 and 11.5. Its

temperature optimum is 25°C, however it can grow in temperatures between 20 and 45 °C. All these observations indicate that strain 11kri321 could have been in a vegetative state in the biofilm. However, upon isolation, spore production was induced *in vitro* by nutrient starvation. Since the genus *Kurthia* is to date considered as non-spore-forming [34], a more detailed investigation of the sporulation capability in strain 11kri321 was conducted.

5.3.2. Genomic imprints of sporulation in *Kurthia* sp. strain 11kri321

The 2.9 Mbp genome of *Kurthia* sp. str. 11kri321 was screened for the presence of genes previously shown to be part of the sporulation pathway in endospore-forming Firmicutes (Supplementary Table 1). In total, 43 sporulation genes were detected in the genome of *Kurthia* sp. str. 11kri321. The genes were assigned to the different stages of the endospore-formation pathway (Figure 5.3). Overall, the gene content is not consistent with the proposed minimum of 60 genes required as a genetic background for sporulation [10]. A closer inspection of the genes shows that only a small fraction of it (18 genes) belongs to those genes reported as conserved in all spore-forming bacilli and clostridia. Even a smaller number (14 genes) of the genes found in *Kurthia* sp. str. 11kri321, correspond to the 44 genes that are considered to be essential for sporulation in B*acillus subtilis* [10].

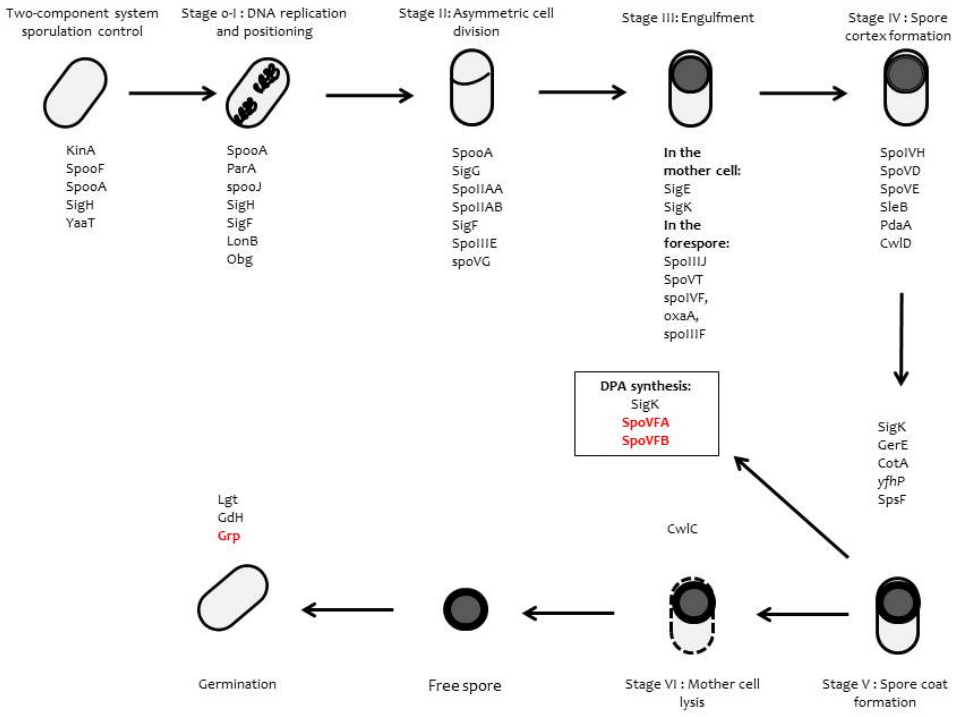


Figure 5.3. The sporulation pathway in Firmicutes. Genes highlighted for each step were present in the *Kurthia* genome. Genes in red are considered essential for sporulation but were not detected in the genomes analysed.

Among the elements involved in sporulation that are conserved in all bacilli and clostridia, the mechanism responsible for the decision to enter this energy-intensive process

appears to be well preserved [35]. Accordingly, the main regulators of the decision-making phase, SigH, the stationary phase sigma factor, as well as Spo0A, the master transcriptional regulator of sporulation are highly conserved among endospore-forming Firmicutes [36,37]. Homologs to these two genes were found in the genome of *Kurthia* sp. str. 11kri321 (Figure 5.3).

Although many proteins that are known to be essential for sporulation in bacilli are also conserved in clostridia, there are substantial differences in the sporulation process between these two clades. One such difference is the way by which Spo0A is activated. In clostridia, it has been observed that Spo0A is phosphorylated directly [38,39], while in bacilli the proteins Spo0B (phosphotransferase) and Spo0F (response regulator) are involved in a process known as phosphorelay [40]. Spo0F was identified in the genome of *Kurthia* sp. str. 11kri321 (FigURE 5. 3). The s*po0B* gene was not detected; however, the *obg* gene, which encodes for a Spo0B-associated GTP-binding protein, was found, suggesting that a phosphotransferase should be present. In fact, by lowering the similarity threshold of detection of sporulation genes, we detected several proteins with conserved Spo0B domains, but those need to be validated experimentally. In bacilli, a family of sporulation-specific sensor histidine kinases (KinA, KinB, KinC, Kin D) that are part of a two-component system is responsible for the activation of Spo0A [41]. Not all members of this family are essential for signaling [42]. In the genome of *Kurthia* sp. str. 11kri321 a gene encoding for KinA was detected, suggesting a signaling pathway similar to that of other bacilli.

In addition to Spo0A, four highly conserved [43] major sigma factors (SigF, SigE, SigG and SigK) are essential for sporulation [44].The genes for these four regulators were found in the genome of *Kurthia* sp. str. 11kri321, indicating that the directing elements of the differential gene expression occurring in the mother cell as well as the endospore are present in our strain. In addition to the sigma factors themselves, regulators of their expression such as SpoIIAA and SpoIIAB, as well as the SpoIIIABCD operon are conserved [45,46]. The first two were found in the genome of our strain, but the latter was missing.

Another key stage of endospore-formation is the packaging of the entire genome within the forespore. It has been postulated that after chromosome replication and migration of the replication origin to the cell poles, only 30% of the original portion of the chromosome is trapped within the forespore in the asymmetrical cell division leading to endospore formation [47] . The DNA transport protein SpoIIIE, which is found in bacilli and clostridia, carries out the translocation of the remaining fraction of the genome. This protein was also identified in the genome of *Kurthia* sp. str. 11kri321.

In the engulfment stage, several alternative mechanisms have been proposed [48]. A molecular zipping model is a feature of sporulation in bacilli, however not all known proteins involved in the process in *B. subtilis* appear in clostridia. None of the genes known to be involved in the zipping process were identified in the genome of *Kurthia* sp. str. 11kri321, and thus an alternative mechanism should operate at this stage of endospore-formation in our strain.

Although the sporulation-specific sigma G factor was found in *Kurthia* sp. str. 11kri321, genes with homology to known small acid-soluble proteins (SASPs) could not be identified. These proteins, which are known to bind DNA and participate in its protection against heat, UV radiation and other damaging agents, represent up to 20% of the total spore proteins found in *B. subtilis* [49–52]. However, the formation of viable spores does not require a great diversity of SASPs [53,54].

Among the genes believed to be involved in the formation of the spore cortex in *B. subtilis,* only *spoVD* and *spoVE* were identified in the genome of *Kurthia* sp. str. 11kri321. In addition to these two genes, a large set of genes involved in cortex formation has been found to be conserved in spore-forming Firmicutes, suggesting a remarkable conservation of the spore cortex biosynthesis [10]. However, according to our analysis, *Kurthia* might have a different mechanism for the formation of the cortex during the maturation of the spore. Likewise, a much smaller set of proteins clearly related to the formation of the spore coat (only CotA), were identified in the genome of strain 11kri321. However, this might not be surprising as the distribution of spore coat proteins appears to reflect a differential adaptation of the organism to specific niches [10], and in consequence, are expected to be highly variable.

Interestingly, the majority of spore coat proteins appear to have higher similarity to homologs in *Clostridium*, rather than in bacilli. This is in sharp contrast to the identification of all the major regulatory elements of the onset of sporulation by a phosphorelay mechanism, which is so far identified in bacilli.

The presence of a conserved set of genes in endospore-forming Firmicutes with dramatically different lifestyles and even in endospore-forming Firmicutes with small and probably streamlined genomes, such as those of segmented filamentous bacteria [12,13,55], suggests that this set of conserved genes represents a close approximation to the true minimal set of sporulation genes [10]. However, many of the potential sporulation protein functions remain unknown and our physiological and genomic analysis suggests that alternative pathways lacking many of the standard elements could lead to endospore-formation in specific lineages of endospore-forming Firmicutes such as *Kurthia*.

5.3.4. Presence of dipicolinic acid inside the spores of strain 11kri321

An important molecule produced in the mother cell of endospore-forming Firmicutes and introduced to the spore at the later stages of sporulation is dipicolinic acid (DPA). DPA is used as a biomarker for detecting endospore-formers in environmental samples [30], and thus we investigated its presence in the spores of *Kurthia* sp. str. 11kri321.

DPA is responsible for accumulation of minerals, especially calcium ions, in the spore core. The presence of calcium and DPA creates a more stable spore core that is shown to guarantee resistance to wet heat [56]. Interestingly, the two genes that encode the DPA synthetase subunits A and B (*spoVFA*, *spoVFB*) were not detected in the genome of *Kurthia* sp. str. 11kri321 and accordingly, we failed to measure any DPA from spore preparations. It has been previously argued that unstable *B. subtilis* spores that do not

contain DPA fail to complete sporulation [56,57]. However, mutants lacking the *spoVF* operon could produce stable spores, though not resistant to wet heat, under the condition that either *ger3* or *sleB* genes would also be absent [57,58]. These studies have shown that DPA does not influence the resistance of the spore to desiccation and dry heat, yet surprisingly, their resistance to UV radiation increases [58]. DPA-free spores tend to have lower core density but are remain resistant and viable [59]. *Kurthia* sp. str. 11kri321 also lacks the gene *ger3*, which suggests that stable spores are produced and are probably better adapted to environmental insults such as high UV radiation. Moreover, the strain was isolated from an alkaline environment. An effect of increase in pH has been previously shown to lead to the release of DPA from *Alicyclobacillus acidoterrestris* spores [60]. Therefore, it is likely that an adaptation to an extremely alkaline pH, such as is tolerated by *Kurthia* sp. str. 11kri321, has selected for the evolution of stable DPA-free spores in this species.

5.3.5. Comparative analysis of all *Kurthia* genomes available

Whether sporulation is a trait observed solely in strain 11kri321 or if it is a widespread phenomenon among *Kurthia* was investigated. Three draft genomes of *Kurthia* strains were also available. Although the absence of specific genes might be taken with caution in draft genomes, these genomes were screened for sporulation genes and their comparison to the sporulation gene set of *Kurthia* sp. str. 11kri321 is shown in Table 5.1. *K. huakuii* str. LAM0618, *Kurthia* sp. str. JC8E and *K. massiliensis* str. JC30 all had a genome size comparable to that of strain 11kri321. A total of 43, 50, and 69 sporulation genes were found in their genomes, respectively. These genomes are, however, unfinished, so there may be some bias in the final number of genes involved.

Sporulation stages	*Kurthia* sp. 11kri321	*Kurthia huakuii* LAM0618	*Kurthia* sp. JC8E	*Kurthia massiliensis*
Growth to stationary phase	*abrB*	*abrB*	*abrB*	*abrB*
Two-component system sporulation control	*kinA*		*kinA*	
	sigH, spo0A, spo0F, yaaT	*sigH, spo0A, yaaT*	*sigH, spo0A, spo0F, yaaT*	*sigH, spo0A, spo0F, spo0B, yaaT, spo0E*
Stage 0-I : DNA replication and positioning	*parA, spo0J*	*parA, spo0J*	*parA, spo0J*	*parA, spo0J*
	sigF, lonB	*sigF, lonB, lonA*	*sigF, lonB, lonA*	*spoIIE, sigF, lonB, lonA*
	obg	*ylbF*	*obg*	*obg*
		spo0M		*spo0M*
Stage II: Asymmetric cell division	*sigG, spoIIAA, spoIIAB*	*sigG, spoIIAA, spoIIAB*	*sigG, spoIIAA, spoIIAB*	*sigG, spoIIAA, spoIIAB*
	spoIIIE	*spoIIIE, Noc*	*spoIIIE*	*spoIIIE*
	spoVG			*spoVG*
				spoIIB, spoIIC, spoIID,
				spoIIGA, spoIIM
		vc		*spoIIIAA, spoIIIAB, spoIIIAC, spoIIIAD,*
				spoIIIAF, spoIIIAG,

				spoIIIAH
Stage III : Engulfment of forespore	*sigE, spoIVF, sigK*	*sigE, sigK*	*sigE, sigK*	*sigE, sigK, spoIIID*
	spoIIIJ	*spoIIIJ*	*spoIIIJ*	*spoIIIJ*
	oxaA, spoIIIF	*oxaA, spoIIIF*	*oxaA, spoIIIF*	*oxaA, spoIIIF*
Stage IV : Spore cortex formation	*spoIVH, spoVD, spoVE*	*spoVD, spoVE*	*spoIVH, spoVD, spoVE*	*spoIVH, spoVD, spoVE*
	sleB		*sleB*	*sleB*
		ybaN	*ybaN*	*ybaN*
	pdaA	*pdaB*	*pdaA, pdaB*	*pdaA*
	cwlD	*cwlD*	*cwlD*	*cwlD*
Stage V : Spore coat formation	*gerE, cotA, yfhP*	*gerE, cotA*	*gerE, cotSA, cotA, yfhP*	*gerE, cotA*
				spoVAA
	spsF			
		spoVK	*spoVK*	*spoVK*
				cotJC
Germination	*lgt*		*lgt*	*lgt*
	gdH	*gdH*	*gdH*	*gdH*
		gerN	*gerN*	*gerN*
Negative signalling of sporulation	*citZ*	*citZ*	*citZ*	*citZ*
Other essential transporters or factors	*ecsA*	*escB*	*ecsA*	
	acsA		*acsA*	
	ylbF		*ylbF*	*ylbF*
	FliA	*FliA, rpoE*	*FliA, FtsI*	*ald, FliA, rpoE*
	ccdA	*ccdA*	*ccdA*	*ccdA*
	tenI	*tenA*	*tenI*	*tenA, tenI*
Non essential genes	*sapB*	*sapB*	*sapB*	*sapB*
Unknown	*yitS,*	*cda, yfhP, yitS, ylxY*	*cda, napA,yitS,*	*glcU, napA, spoIIIAE, yfhP, yitS,*
		lonC	*PerM*	*PerM, rfbX*
genome size (Mbp)	2,9	3.55	2.98	3.23
total number of sporulation proteins	43	43	50	69

Table 5.1. Sporulation genes detected in the analysed genomes of *Kurthia* spp.

Among the genes with a clear function, a striking difference between the genes identified in *Kurthia massiliensis* and the other strains of *Kurthia* is found at the stage of decision-making to commit to sporulation. While homologs to SigH, SpoOA and SpoOF were found in all *Kurthia* genomes, SpoOB, which has a role in signal transduction on the *yaaT* gene was only found in *K. massiliensis*. The entire *spoIIIABCD* operon was detected in the genome of *K. massiliensis*, but was absent from the other three *Kurthia* genomes. The other sporulation genes with a known function appear to be conserved within *Kurthia*.

Although the ecology of *Kurthia* has not been analyzed in detail, representatives of this genus have been isolated from diverse environments such as stool (*K. massiliensis* [61] and *K. senegalensis* [62]), biogas slurry (*K. huakuii* [14]), patient tissue (*K. gibsonii*),

cigarettes [63] and methanogenic bacterial complexes [33]. In metagenomic studies, the genus *Kurthia* has been prevalent in snail gastrointestinal tracts [64], restaurant kitchen cutting boards [65] and soy sauce fermentation processes [66,67]. In many of these habitats, sporulation can be a trait linked to survival and therefore a process that should be under selective pressure for conservation. Whether other strains of *Kurthia* are able to produce spores or whether the number and nature of the genes observed is sufficient for a complete and successful sporulation procedure, still needs to be verified physiologically.

5.3.5. Cryptosporulation and asporogenesis

Endospore-formation is a defining and differentiating character of Firmicutes. It might have been a key element explaining the success and radiation of such a diversified bacterial phylum. Understanding the way in which endospore-formation, a seemingly ancestral characteristic, can be lost to give rise to non-spore-forming Firmicutes is essential in the study of the ecology and evolution of this clade. The feasibility of a comparative genomic approach to understand the genetic mechanisms of endospore-formation has greatly contributed to our current knowledge of the distribution of endospore-formation genes within and outside Firmicutes [8,10,11]. However, even the task of separating species with sequenced genomes on spore-formers versus non-spore-formers is daunting. The discovery of sporulation in *Kurthia,* as well as the finding of a significant deviation in this genus from the accepted core of sporulation-related genes, is an opportunity to extend our understanding of what constitutes a minimal set of sporulation-specific genes [68].

These results also question the validity of describing organisms as asporogenic when sporulation has not been observed, to distinguish them from those in which it is a truly absent trait. Considering that almost any natural habitat is subjected to significant variations in environmental parameters, it is difficult to conceive of a condition in which the complete loss of sporulation would be an advantage for Firmicutes. Accordingly, we propose to use the term "cryptosporogenic" to designate those groups for which sporulation has not been so far observed, and also for which a detailed genomic analysis is not yet possible to reliably define their capability to produce or not spores. Concerning *Kurthia* sp. str. 11kri321 a detailed study of the developmental stages of sporulation and germination still needs to be performed. Moreover, the genetic determinants of the successful production of viable spores also have to be further characterized. This discovery, however, paves the path for further investigation of cryptosporogenic or asporogenic Firmicutes.

5.4. Acknowledgements

Individuals responsible for funding this work did not participate in study design, data collection and interpretation, or the decision to submit this work for publication. The authors declare no conflict of interest. The genetic information downloaded from Genbank may be considered to be part of the genetic patrimony of Greece, the country from which the sample was obtained. Users of this information agree to: 1) acknowledge

Greece as the country of origin in any country where the genetic information is presented and 2) contact the CBD focal point and the ABS focal point identified in the CBD website http://www.cbd.int/information/nfp.shtml if they intend to use the genetic information for commercial purposes. We acknowledge funding from the Swiss National Funding projects 31003A_132358/1 and 31003A_152972, from Fondation Pierre Mercier pour la science and from REGARD for equality of women in science.

5.5. References

[1] Müller AL, de Rezende JR, Hubert CRJ, Kjeldsen KU, Lagkouvardos I, Berry D, et al. Endospores of thermophilic bacteria as tracers of microbial dispersal by ocean currents. ISME J 2014;8:1153–65. doi:10.1038/ismej.2013.225.

[2] Nicholson WL, Fajardo-Cavazos P, Rebeil R, Slieman TA, Riesenman PJ, Law JF, et al. Bacterial endospores and their significance in stress resistance. Antonie Van Leeuwenhoek 2002;81:27–32. doi:10.1023/A:1020561122764.

[3] Sanchez-Salas J-L, Setlow B, Zhang P, Li Y, Setlow P. Maturation of Released Spores Is Necessary for Acquisition of Full Spore Heat Resistance during Bacillus subtilis Sporulation ▿. Appl Environ Microbiol 2011;77:6746–54. doi:10.1128/AEM.05031-11.

[4] Nicholson WL, Munakata N, Horneck G, Melosh HJ, Setlow P. Resistance of Bacillus endospores to extreme terrestrial and extraterrestrial environments. Microbiol Mol Biol Rev 2000;64:548–+. doi:10.1128/MMBR.64.3.548-572.2000.

[5] Setlow B, Loshon C a., Genest P c., Cowan A e., Setlow C, Setlow P. Mechanisms of killing spores of Bacillus subtilis by acid, alkali and ethanol. J Appl Microbiol 2002;92:362–75. doi:10.1046/j.1365-2672.2002.01540.x.

[6] Brun Y, Shimkets LJ. Prokaryotic Development. ASM Press; 2000.

[7] Driks A. The dynamic spore. Proc Natl Acad Sci U S A 2003;100:3007–9. doi:10.1073/pnas.0730807100.

[8] Galperin MY. Genome Diversity of Spore-Forming Firmicutes. Microbiol Spectr 2013;1:TBS-0015-2012. doi:10.1128/microbiolspectrum.TBS-0015-2012.

[9] Wu M, Ren Q, Durkin AS, Daugherty SC, Brinkac LM, Dodson RJ, et al. Life in Hot Carbon Monoxide: The Complete Genome Sequence of Carboxydothermus hydrogenoformans Z-2901. PLoS Genet 2005;1:e65. doi:10.1371/journal.pgen.0010065.

[10] Galperin MY, Mekhedov SL, Puigbo P, Smirnov S, Wolf YI, Rigden DJ. Genomic determinants of sporulation in Bacilli and Clostridia: towards the minimal set of sporulation-specific genes. Environ Microbiol 2012;14:2870–90. doi:10.1111/j.1462-2920.2012.02841.x.

[11] Onyenwoke RU, Brill JA, Farahi K, Wiegel J. Sporulation genes in members of the low G+C Gram-type-positive phylogenetic branch (Firmicutes). Arch Microbiol 2004;182:182–92. doi:10.1007/s00203-004-0696-y.

[12] Kuwahara T, Ogura Y, Oshima K, Kurokawa K, Ooka T, Hirakawa H, et al. The lifestyle of the segmented filamentous bacterium: a non-culturable gut-associated immunostimulating microbe inferred by whole-genome sequencing. DNA Res Int J Rapid Publ Rep Genes Genomes 2011;18:291–303. doi:10.1093/dnares/dsr022.

[13] Sczesnak A, Segata N, Qin X, Gevers D, Petrosino JF, Huttenhower C, et al. The genome of th17 cell-inducing segmented filamentous bacteria reveals extensive auxotrophy and adaptations to the intestinal environment. Cell Host Microbe 2011;10:260–72. doi:10.1016/j.chom.2011.08.005.

[14] Ruan Z, Wang Y, Song J, Jiang S, Wang H, Li Y, et al. Kurthia huakuii sp. nov., isolated from biogas slurry, and emended description of the genus Kurthia. Int J Syst Evol Microbiol 2014;64:518–21. doi:10.1099/ijs.0.056044-0.

[15] Kurthia n.d. http://www.bacterio.net/kurthia.html (accessed January 28, 2016).
[16] Kurthia - Genome - NCBI n.d. http://www.ncbi.nlm.nih.gov/genome/?term=Kurthia (accessed January 28, 2016).
[17] Schaeffer AB, Fulton MD. A SIMPLIFIED METHOD OF STAINING ENDOSPORES. Science 1933;77:194. doi:10.1126/science.77.1990.194.
[18] Liesack W, Weyland H, Stackebrandt E. Potential Risks of Gene Amplification by PCR as Determined by 16s rDNA Analysis of a Mixed-Culture of Strict Barophilic Bacteria. Microb Ecol 1991;21:191–8. doi:10.1007/BF02539153.
[19] Muyzer G, Teske A, Wirsen C, Jannasch H. Phylogenetic-Relationships of Thiomicrospira Species and their Identification in Deep-Sea Hydrothermal Vent Samples by Denaturing Gradient Gel-Electrophoresis of 16S rDNA Fragments. Arch Microbiol 1995;164:165–72. doi:10.1007/BF02529967.
[20] Wunderlin T, Junier T, Roussel-Delif L, Jeanneret N, Junier P. Stage 0 sporulation gene A as a molecular marker to study diversity of endospore-forming Firmicutes: Culture-independent study of endospore-forming Firmicutes. Environ Microbiol Rep 2013;5:911–24. doi:10.1111/1758-2229.12094.
[21] Kim O-S, Cho Y-J, Lee K, Yoon S-H, Kim M, Na H, et al. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. Int J Syst Evol Microbiol 2012;62:716–21. doi:10.1099/ijs.0.038075-0.
[22] Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 2002;30:3059–66. doi:10.1093/nar/gkf436.
[23] Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol 2010;59:307–21. doi:10.1093/sysbio/syq010.
[24] Junier T, Zdobnov EM. The Newick Utilities: High-throughput Phylogenetic tree Processing in the Unix Shell. Bioinformatics 2010:btq243. doi:10.1093/bioinformatics/btq243.
[25] Hemmerich C, Buechlein A, Podicheti R, Revanna KV, Dong Q. An Ergatis-based prokaryotic genome annotation web server. Bioinforma Oxf Engl 2010;26:1122–4. doi:10.1093/bioinformatics/btq090.
[26] Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. Bioinformatics 2005;21:3422–3. doi:10.1093/bioinformatics/bti553.
[27] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol 1990;215:403–10. doi:10.1016/S0022-2836(05)80360-2.
[28] Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 1970;48:443–53. doi:10.1016/0022-2836(70)90057-4.
[29] Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet TIG 2000;16:276–7.

[30] Brandes Ammann A, K&#246, Lle L, Brandl H. Detection of Bacterial Endospores in Soil by Terbium Fluorescence. Int J Microbiol 2011;2011. doi:10.1155/2011/435281.

[31] Tranos MD, Eleftheriadis GE, Kilias AA. Philippi granitoid as a proxy for the Oligocene and Miocene crustal deformation in the Rhodope Massif (Eastern Macedonia, Greece). Geotecton Res 2009;96:69–85. doi:10.1127/1864-5658/09/96-0069.

[32] Rodriguez-R LM, Konstantinidis KT. Bypassing cultivation to identify bacterial species. Issues 2010.

[33] Ruan Z, Zhou S, Jiang S, Sun L, Zhai Y, Wang Y, et al. Isolation and characterization of a novel cinosulfuron degrading Kurthia sp. from a methanogenic microbial consortium. Bioresour Technol 2013;147:477–83. doi:10.1016/j.biortech.2013.08.017.

[34] Stackebrandt E, Keddie RM, Jones D. The Genus Kurthia. In: Dr MDP, Falkow S, Rosenberg E, Schleifer K-H, Stackebrandt E, editors. The Prokaryotes, Springer US; 2006, p. 519–29.

[35] Angert ER, Hutchison EA, Miller DA. Sporulation in Bacteria: Beyond the Standard Model. Microbiol Spectr 2014;2. doi:10.1128/microbiolspec.TBS-0013-2012.

[36] Britton RA, Eichenberger P, Gonzalez-Pastor JE, Fawcett P, Monson R, Losick R, et al. Genome-wide analysis of the stationary-phase sigma factor (sigma-H) regulon of Bacillus subtilis. J Bacteriol 2002;184:4881–90.

[37] Hilbert DW, Piggot PJ. Compartmentalization of gene expression during Bacillus subtilis spore formation. Microbiol Mol Biol Rev MMBR 2004;68:234–62. doi:10.1128/MMBR.68.2.234-262.2004.

[38] Wörner K, Szurmant H, Chiang C, Hoch JA. Phosphorylation and functional analysis of the sporulation initiation factor Spo0A from Clostridium botulinum. Mol Microbiol 2006;59:1000–12. doi:10.1111/j.1365-2958.2005.04988.x.

[39] Steiner E, Dago AE, Young DI, Heap JT, Minton NP, Hoch JA, et al. Multiple orphan histidine kinases interact directly with Spo0A to control the initiation of endospore formation in Clostridium acetobutylicum. Mol Microbiol 2011;80:641–54. doi:10.1111/j.1365-2958.2011.07608.x.

[40] Hoch JA, Varughese KI. Keeping Signals Straight in Phosphorelay Signal Transduction. J Bacteriol 2001;183:4941–9. doi:10.1128/JB.183.17.4941-4949.2001.

[41] Stephenson K, Hoch JA. Evolution of signalling in the sporulation phosphorelay. Mol Microbiol 2002;46:297–304. doi:10.1046/j.1365-2958.2002.03186.x.

[42] LeDeaux JR, Yu N, Grossman AD. Different roles for KinA, KinB, and KinC in the initiation of sporulation in Bacillus subtilis. J Bacteriol 1995;177:861–3.

[43] de Hoon MJ, Eichenberger P, Vitkup D. Hierarchical evolution of the bacterial sporulation network. Curr Biol CB 2010;20:R735–45. doi:10.1016/j.cub.2010.06.031.

[44] Traag BA, Pugliese A, Eisen JA, Losick R. Gene Conservation among Endospore-Forming Bacteria Reveals Additional Sporulation Genes in Bacillus subtilis. J Bacteriol 2013;195:253–60. doi:10.1128/JB.01778-12.

[45] Duncan L, Alper S, Losick R. SpoIIAA governs the release of the cell-type specific transcription factor sigma F from its anti-sigma factor SpoIIAB. J Mol Biol 1996;260:147–64. doi:10.1006/jmbi.1996.0389.

[46] Illing N, Errington J. The spoIIIA operon of Bacillus subtilis defines a new temporal class of mother-cell-specific sporulation genes under the control of the sigma E form of RNA polymerase. Mol Microbiol 1991;5:1927–40.
[47] Encyclopedia of Microbiology. Academic Press; 2009.
[48] Serrano M, Crawshaw AD, Dembek M, Monteiro JM, Pereira FC, Pinho MG, et al. The SpoIIQ-SpoIIIAH complex of Clostridium difficile controls forespore engulfment and late stages of gene expression and spore morphogenesis. Mol Microbiol 2016;100:204–28. doi:10.1111/mmi.13311.
[49] Setlow B, Setlow P. Decreased UV light resistance of spores of Bacillus subtilis strains deficient in pyrimidine dimer repair and small, acid-soluble spore proteins. Appl Environ Microbiol 1988;54:1275–6.
[50] Setlow B, Setlow P. Absence of transient elevated UV resistance during germination of Bacillus subtilis spores lacking small, acid-soluble spore proteins alpha and beta. J Bacteriol 1988;170:2858–9.
[51] Setlow P. I will survive: DNA protection in bacterial spores. Trends Microbiol 2007;15:172–80. doi:10.1016/j.tim.2007.02.004.
[52] Driks A. Overview: Development in bacteria: spore formation in Bacillus subtilis. Cell Mol Life Sci CMLS 2002;59:389–91.
[53] Moeller R, Setlow P, Horneck G, Berger T, Reitz G, Rettberg P, et al. Roles of the major, small, acid-soluble spore proteins and spore-specific and universal DNA repair mechanisms in resistance of Bacillus subtilis spores to ionizing radiation from X rays and high-energy charged-particle bombardment. J Bacteriol 2008;190:1134–40. doi:10.1128/JB.01644-07.
[54] Raju D, Waters M, Setlow P, Sarker MR. Investigating the role of small, acid-soluble spore proteins (SASPs) in the resistance of Clostridium perfringens spores to heat. BMC Microbiol 2006;6:50. doi:10.1186/1471-2180-6-50.
[55] Pamp SJ, Harrington ED, Quake SR, Relman DA, Blainey PC. Single-cell sequencing provides clues about the host interactions of segmented filamentous bacteria (SFB). Genome Res 2012;22:1107–19. doi:10.1101/gr.131482.111.
[56] Daniel RA, Errington J. Cloning, DNA Sequence, Functional Analysis and Transcriptional Regulation of the Genes Encoding Dipicolinic Acid Synthetase Required for Sporulation in Bacillus subtilis. J Mol Biol 1993;232:468–83. doi:10.1006/jmbi.1993.1403.
[57] Setlow B, Atluri S, Kitchel R, Koziol-Dube K, Setlow P. Role of Dipicolinic Acid in Resistance and Stability of Spores of Bacillus subtilis with or without DNA-Protective α/β-Type Small Acid-Soluble Proteins. J Bacteriol 2006;188:3740–7. doi:10.1128/JB.00212-06.
[58] Paidhungat M, Setlow B, Driks A, Setlow P. Characterization of Spores of Bacillus subtilis Which Lack Dipicolinic Acid. J Bacteriol 2000;182:5505–12.
[59] Tovar-Rojo F, Chander M, Setlow B, Setlow P. The Products of the spoVA Operon Are Involved in Dipicolinic Acid Uptake into Developing Spores of Bacillus subtilis. J Bacteriol 2002;184:584–7. doi:10.1128/JB.184.2.584-587.2002.

[60] Bevilacqua A, Ciuffreda E, Sinigaglia M, Corbo MR. Spore inactivation and DPA release in Alicyclobacillus acidoterrestris under different stress conditions. Food Microbiol 2015;46:299–306. doi:10.1016/j.fm.2014.08.017.

[61] Roux V, El Karkouri K, Lagier J-C, Robert C, Raoult D. Non-contiguous finished genome sequence and description of Kurthia massiliensis sp. nov. Stand Genomic Sci 2012;7:221–32. doi:10.4056/sigs.3206554.

[62] Roux V, Lagier J-C, Gorlas A, Robert C, Raoult D. Non-contiguous finished genome sequence and description of Kurthia senegalensis sp. nov. Stand Genomic Sci 2014;9:1319–30. doi:10.4056/sigs.5078947.

[63] Rooney AP, Swezey JL, Wicklow DT, McAtee MJ. Bacterial species diversity in cigarettes linked to an investigation of severe pneumonitis in U.S. Military personnel deployed in operation iraqi freedom. Curr Microbiol 2005;51:46–52. doi:10.1007/s00284-005-4491-z.

[64] Pawar KD, Banskar S, Rane SD, Charan SS, Kulkarni GJ, Sawant SS, et al. Bacterial diversity in different regions of gastrointestinal tract of Giant African snail (Achatina fulica). MicrobiologyOpen 2012;1:415–26. doi:10.1002/mbo3.38.

[65] Abdul-Mutalib N-A, Amin Nordin S, Osman M, Ishida N, Tashiro K, Sakai K, et al. Pyrosequencing analysis of microbial community and food-borne bacteria on restaurant cutting boards collected in Seri Kembangan, Malaysia, and their correlation with grades of food premises. Int J Food Microbiol 2015;200:57–65. doi:10.1016/j.ijfoodmicro.2015.01.022.

[66] Wei C-L, Chao S-H, Tsai W-B, Lee P-S, Tsau N-H, Chen J-S, et al. Analysis of bacterial diversity during the fermentation of inyu, a high-temperature fermented soy sauce, using nested PCR-denaturing gradient gel electrophoresis and the plate count method. Food Microbiol 2013;33:252–61. doi:10.1016/j.fm.2012.10.001.

[67] Wei Q, Wang H, Chen Z, Lv Z, Xie Y, Lu F. Profiling of dynamic changes in the microbial community during the soy sauce fermentation process. Appl Microbiol Biotechnol 2013;97:9111–9. doi:10.1007/s00253-013-5146-9.

[68] Kroos L. The Bacillus and Myxococcus Developmental Networks and Their Transcriptional Regulators. Annu Rev Genet 2007;41:13–39. doi:10.1146/annurev.genet.41.110306.130400.

Chapter 6

Anoxybacillus geothermalis sp. nov., a facultative anaerobic endospore-forming bacterium isolated from mineral deposits in a geothermal station

Sevasti Filippidou[1], Marion Jaussi[1], Thomas Junier[1], Tina Wunderlin[1], Nicole Jeanneret[1], Fabio Palmieri[1], Ilona Palmieri[1], Ludovic Roussel-Delif[1], Andrea Vieth-Hillebrand[2], Alexandra Vetter[2], Patrick S. Chain[3], Simona Regenspurg[2], Pilar Junier[1]

1. Laboratory of Microbiology, Institute of Biology, University of Neuchatel, Emile-Argand 11, CH-2000, Switzerland

2. Helmholtz Centre Potsdam, German Research Centre for Geosciences GFZ, Telegrafenberg, D-14473 Potsdam, Germany

3. Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, 87545, USA.

This chapter is adapted from the article: Filippidou S., Jaussi M., Junier T., Wunderlin T., Jeanneret N., Palmieri F., Palmieri I., Roussel-Delif L.,Vieth-Hillebrand A., Vetter A., Chain P.S., Regenspurg S., Junier P. *Anoxybacillus geothermalis* sp. nov., a facultative anaerobic endospore-forming bacterium isolated from mineral deposits in a geothermal station, accepted for publication in *IJSEM*

Abstract

A novel endospore-forming bacterium designated strain GSsed3^T was isolated from deposits clogging aboveground filters from the geothermal power plant of Gross Schoenebeck in Northern Germany. The novel isolate was Gram-staining-positive, facultative anaerobe, catalase-positive and oxidase-positive. Optimum growth occurred at 60 °C, 0.5 % (w/v) NaCl and pH 7-8. Analysis of the 16S rRNA sequence similarity indicated that strain GSsed3^T belongs to the genus *Anoxybacillus*, and showed 99.8 % sequence similarity to *Anoxybacillus rupiensis* R270^T, 98.2 % similarity to *Anoxybacillus tepidamans* GS5-97^T, 97.9 % similarity to *Anoxybacillus voinovskiensis* TH13^T, 97.7 % similarity to *Anoxybacillus caldiproteolyticus* DSM 15730^T, and 97.6 % similarity to *Anoxybacillus amylolyticus* MR3C^T. DNA-DNA hybridization (DDH) indicated only 16 % relatedness to *A. rupiensis* DSM 17127^T. Furthermore, DDH estimation based on genome analysis indicated only 19.9 % overall nucleotide similarity to *A. amylolyticus* DSM 15939^T. The major respiratory menaquinone was MK-8. The polar lipid profile consisted of phosphatidylethanolamine, phosphatidylglycerol, diphosphatidylglycerol, one unknown phosphoglycolipid and one unknown phospholipid. The predominant cellular fatty acids were iso-$C_{15:0}$, iso-$C_{17:0}$, $C_{16:0}$, iso-$C_{16:0}$, and anteiso-$C_{17:0}$. The peptidoglycan type was A1γ *meso*-Dpm-direct. The genomic DNA G+C content of the strain was 46.9 mol%. The phenotypic, genotypic and chemotaxonomic characterization indicated that strain GSsed3^T differs from related species of the genus. Therefore, strain GSsed3^T is considered to be a novel species of the genus *Anoxybacillus*, for which the name *Anoxybacillus geothermalis* sp. nov. is proposed. The type strain of *Anoxybacillus geothermalis* is GSsed3^T (= CCOS808^T = ATCC BAA2555^T).

6. 1. Introduction and results

The genus *Anoxybacillus* belongs to the family *Bacillaceae* and is related to *Geobacillus*, which explains why, for instance, a former representative of the *Geobacillus*, *Geobacillus tepidamans* [1], is now classified as *Anoxybacillus tepidamans* [2]. The name given to the genus suggests that species assigned to it thrive under anoxic conditions, and, indeed, the first species classified as *Anoxybacillus* were aerotolerant anaerobes [3]. Nevertheless, at the time of writing from the 22 *Anoxybacillus* species, 16 are facultative anaerobes [1,4–15], five are strict aerobes [2,16–19] and one is an aerotolerant anaerobe [3]. Species that belong to *Anoxybacillus* share common characteristics, mainly concerning their major fatty acids (iso-$C_{17:0}$ and iso-$C_{15:0}$), the rod shape of the vegetative cells and the terminal/ subterminal position of their endospores. The G+C content of this genus differs considerably among species, from 37.8 % up to 57 % mol. Besides the phenotypic similarities, species that belong to *Anoxybacillus* share common ecological characteristics, too. All of them are thermophiles or moderate thermophiles, with the majority of the species isolated so far, originated from thermal springs [4–7,9,10,14–19]. This study describes the discovery of a novel *Anoxybacillus* species isolated from a geothermal reservoir.

Samples were taken from deposits at the entrance filters of the geothermal research facility of Gross Schoenebeck. The deposits were obtained from coarse filter bags (10-20 µm pore-size) located right after the geothermal fluid transport to the surface and degassed. The *in situ* geothermal laboratory of Gross Schoenebeck is situated in the North German Basin (52°54'13.15" N, 13°36'5.43" E). The geothermal fluid is produced from a reservoir (Permian sandstone and volcanic rocks) at about 4,200 m depth. At this depth, temperature is around 150 °C. The anoxic hot fluid contains high salt concentration (total dissolved solids > 260 g/L) and measured pH was close to 5.5 at 25 °C. However, at this saline condition, the measured pH value has to be corrected by about one pH unit, shifting the pH to the neutral range [20]. The mineral phases identified as precipitates in filter residues, were barite ($BaSO_4$), laurionite (PbOHCl), halite (NaCl), magnetite (Fe_3O_4), and occasionally quartz (SiO_2), and native copper [21]. Isolation was carried out by inoculating one gram of deposit sample into 5 mL of modified Difco™ D2216 Marine Broth. This medium was modified by using 5 g of tryptone instead of peptone, omitting the addition of potassium bromide and adjusting the pH to 5.2 with HCl, and it will be referred to as modified D2216 medium. Samples were inoculated at 37 and 60 °C under aerobic conditions. Growth was only observed at 60 °C. After enrichments and several purifications on modified D2216 Marine Agar with 2 % (w/v) agar instead of 1.5 % (w/v), strain GSsed3^T was isolated. The strain was conserved at -80 °C as a 30% (w/v) glycerol suspension. In addition, the same procedure of enrichment and isolation was repeated with samples originating from water and a biofilm growing on the pipes of the collection point originated from the geothermal facility of Bruchsal. The municipal Bruchsal geothermal power plant is located in the Southern part of Germany, near the German-French boarder (49°07'31.23" N, 08°34'8.00" E). Its main purpose is to generate electricity; however water is also used for heating purposes. The water is pumped from a

depth of 2.500 meters. The pH of the water varies from 5.4 to 5.8 and the pressure inside the drilling pipes is 22 bar. Temperature at the deepest drilling point is around 132°C, however there is a 10 °C loss in temperature as water reaches the surface. Before sampling, the plant had been circulating for 24 hours, pumping 25L of water per second. Gases are dissolved in the water, with high CO_2 concentration. After enrichment and purification as indicated above, strains B2M1, B7M1 and B7M2 were also obtained. In this study, a polyphasic taxonomy approach combining genotypic, chemotaxonomic and phenotypic characteristics [22] was conducted to determine the precise taxonomic position of a novel Gram-staining-positive endospore-forming bacterium (GSsed3^T) and its related strains (B2M1, B7M1 and B7M2).

Genomic DNA was extracted using the InnuPREP Bacteria DNA kit (Analytik Jena, Germany), according to the manufacturer's instructions. DNA was quantified fluorometrically with a Qubit® dsDNA HS Assay Kit and Qubit®2.0 Fluorometer (Invitrogen Ltd., UK). To obtain the nearly full-length 16S rRNA gene sequence, PCR amplification was performed using the bacterial universal primer set GM3F and GM4R [23]. The PCR product was purified with a MultiScreen PCRμ96 Filter Plate (Millipore, USA) and sequenced. To obtain a full sequence of the amplicon, PCR products were sequenced in addition with the primers 907r, 926f [23] and 518r [24]. Sequencing was conducting using the services of Microsynth AG (Switzerland) and GATC Biotech (Germany).The partial sequences generated were assembled using the online EMBOSS tools revseq and merger and the consensus sequence was corrected manually for errors. A sequence of 1548 bp for GSsed3^T (B2M1 = 1247 bp; B7M1 =1487 bp; and B7M2 = 1487 bp) was obtained. Screening of phylogenetic neighbours of strain GSsed3^T was carried out using EzTaxon-e [25], taking into account 16S rRNA gene sequences from cultured isolates. Sequences of all *Anoxybacillus* species along with *Aeribacillus pallidus* strain GS3372 (Filippidou et al, 2015), as an outgroup, were obtained and used to build phylogenetic trees using the online tools of phylogeny.fr website [26]. The tree topology was verified using four independent methods for the reconstruction of phylogenetic trees, Neighbour-Joining, Maximum Likelihood, Maximum Parsimony and Bayesian inference [26–34]. Trees were processed (re-rooting, extracting topology, and plotting) with the Newick Utilities (Junier and Zdobnov 2010). The EzTaxon-e identification of the 16S rRNA gene sequence of strain GSsed3^T showed 99.8 % similarity to *Anoxybacillus rupiensis* R270^T, 98.2 % similarity to *Anoxybacillus tepidamans* GS5-97^T, 97.9 % similarity to *Anoxybacillus voinovskiensis* TH13^T, 97.7 % similarity to *Anoxybacillus caldiproteolyticus* DSM 15730^T, and 97.6 % similarity to *Anoxybacillus amylolyticus* MR3C^T. However, considering that EzTaxon-e uses a Basic Local Alignment Search Tool for identification and that this approach maximizes similarity for partial regions of the alignment, sequence identity for the five strains with similarity values over 97 % was verified using the pair-wise algorithm Needleman-Wunsch [35]. The results are summarized in Supplementary Table 1. Based on pair-wise alignments, the 16S rRNA gene of strain GSsed3^T has 97.6 and 97.2 % identity to *A. rupiensis* R270^T and *A. amylolyticus* MR3C^T, respectively. A high degree of concordance was observed between the clustering within the trees (Figure 6.1). The data

show that strain GSsed3^T, as well as strains B2M1, B7M1 and B7M2, form a well-supported subline within *Anoxybacillus* that is distinct from all other species within the genus. Furthermore, the branching structure supports the classification of the four isolates as a potentially new species (Figure 6.1 and Supplementary Figure 6.1).

In order to further support the classification of strain GSsed3^T as a novel *Anoxybacillus* species, a DNA-DNA hybridisation test between GSsed3^T and *A. rupiensis* DSM 17127^T was carried out using the services of DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen, Germany). Cells were disrupted by using a Constant Systems TS 0.75 KW (IUL Instruments, Germany) and the DNA in the crude lysate was purified by chromatography on hydroxyapatite [36]. DNA-DNA hybridization was performed as described previously [37,38] using a model Cary 100 Bio UV/VIS-spectrophotometer equipped with a Peltier-thermost6x6 multicell changer and a temperature controller with *in-situ* temperature probe (Varian, USA). According to the DNA-DNA hybridization results, strain GSsed3^T did not belong to the species *A. rupiensis* (16 % DNA-DNA relatedness), when the recommendation of a threshold value of 70 % DNA-DNA relatedness is considered for the definition of bacterial species as suggested by the ad hoc committee [39]. Likewise, the comparison of the whole genome sequences of strain GSsed3^T with *A. amylolyticus* DSM 15939^T using GBDP2_BLASTPLUS [40] indicated a DDH estimate (GLM-based) of 19.9 % (with an error of 2.3 %). These results were below 70 % (the species DDH cutoff value), which also indicated that strain GSsed3^T did not belong to the species *A. amylolyticus*. Finally, in order to demonstrate that the related strains (B2M1 and B7M1) indeed belong to the same species, *A. geothermalis*, we have performed a whole genome sequencing, using the PacBio technology, and average nucleotide identity (ANI) [41] was calculated between strains GSsed3^T and B2M1 ,strains GSsed3^T and B7M2, and strains B2M1 and B7M1. The analysis showed an ANI of 99.99 % (standard deviation 0.26 %), 99.99 % (standard deviation 0.27 %), and 100 % (standard deviation 0.04 %), respectively.

On the basis of the results obtained, GSsed3^T is considered a novel species of the genus *Anoxybacillus*. The novel isolate was characterized by polyphasic taxonomy and a range of phenotypic and molecular characteristics were determined as recommended by the minimal standards for describing new taxa of aerobic, endospore-forming bacteria [42]. Moreover, as recommended by the minimal standards, reference strains obtained from the DSMZ, *A. caldiproteolyticus* DSM 15730^T, *A. rupiensis* DSM 17127^T, *A. tepidamans* DSM 16325^T, *A. voinovskiensis* DSM 17075^T and *A. amylolyticus* DSM 15939^T, were included in this comparative study.

Cell morphology, average cell size at 24 h in modified D2216 liquid medium, and endospore formation were determined using light and phase-contrast microscopy (Leica DM R, magnification 1000x). Gram staining was performed on an overnight solid culture using the Hucker staining method [43]. Mobility was tested according to two different methods [44,45]. The composition of the swimming media were slightly different but contained both 0.3% of Agar (g/w). The solid media were dried for 1 h under laminar flow. Colonies from an overnight solid culture were inoculated by thrusting a straight needle into the center of the petri dishes and incubated for 24 h at 55 °C. Isolate GSsed3T was

motile as the colony spread into the medium in comparison to a non-motile strain (*Pseudomonas putida* ΔfliM). Cells of GSsed3^T were rods and Gram-staining-positive.

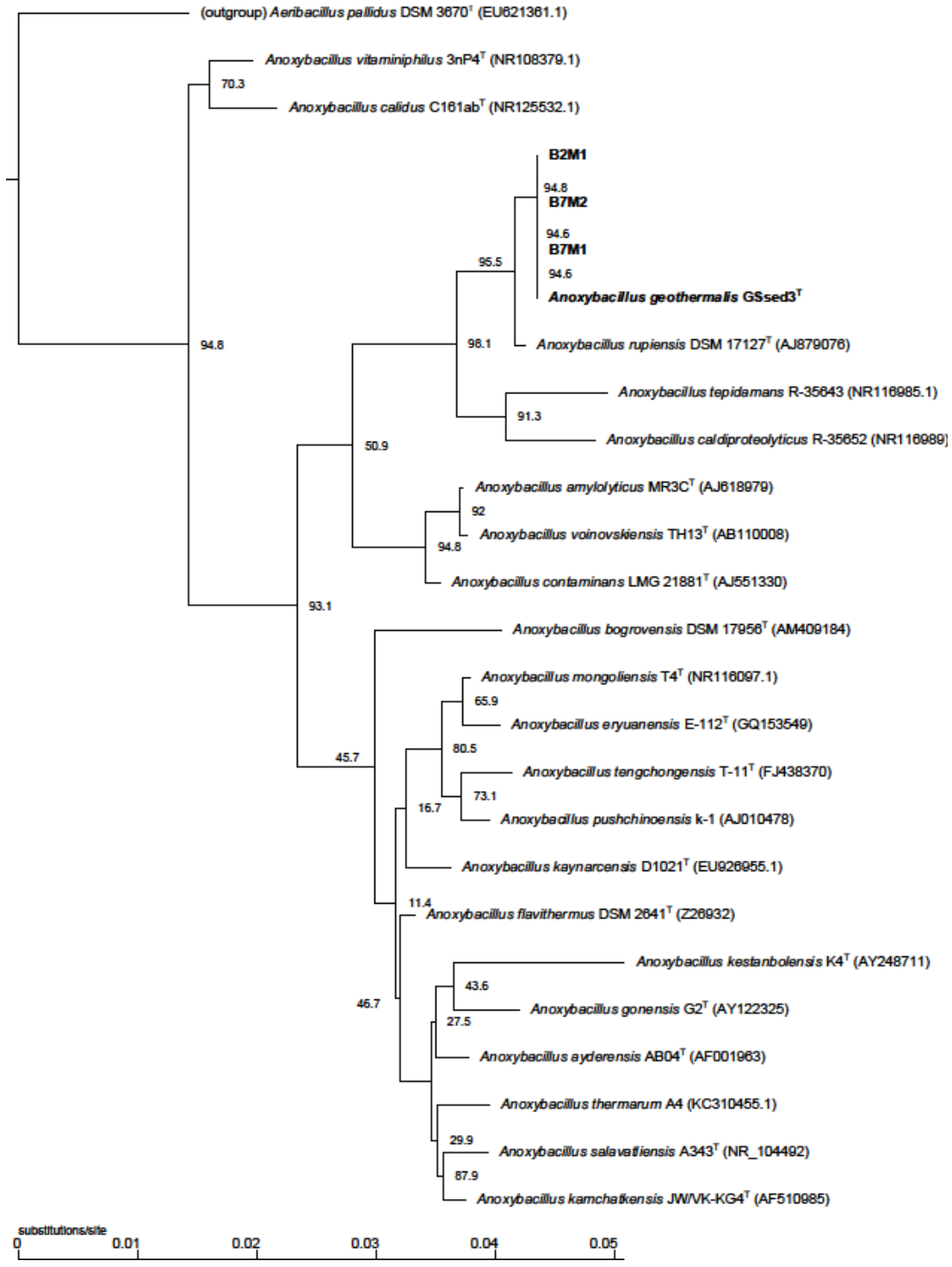


Figure 6.1. Neighbor-Joining phylogenetic tree based on the 16S rRNA gene sequences. The tree shows the relationship between strain GSsed3^T, B2M1, B7M1 and B7M2 and others members of *Anoxybacillus* genus. *Aeribacillus pallidus* was used as outgroup. The phylogenetic tree was built online (phylogeny.fr). Multiple alignment has been made using MUSCLE and the phylogenetic tree has been constructed using BioNJ. Bootstrap values based on 1000 replications have been calculated and expressed as percentages. Bar shows substitutions per nucleotide position. The NCBI accession number of each sequence is given in parenthesis.

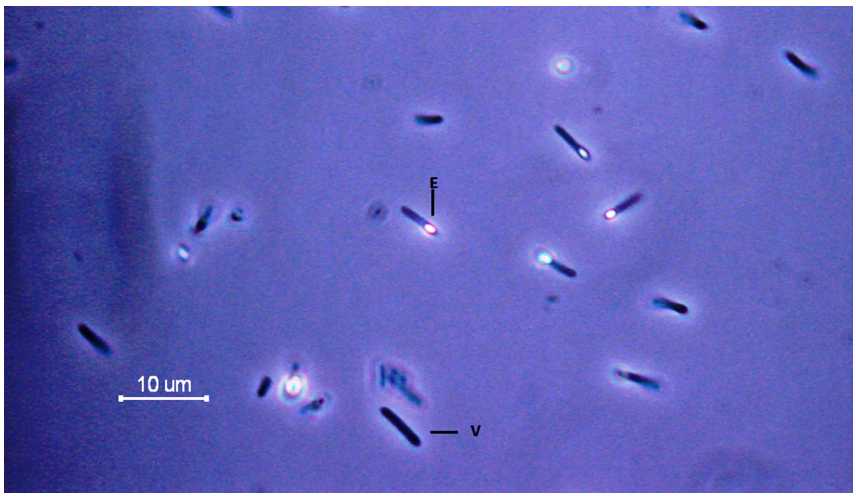


Figure 6.2. Contrast-phase microscopy of *Anoxybacillusgeothermalis* strain GSsed3, (Leica DM R, Nikon camera Digital Eclipse DXM 1200, Nikon ACT-1 software). (V) Vegetative cells after 24 h growth. (E) Endospores

Subterminal to terminal endospores were observed in slightly swollen sporangia (Figure 6.2).

The strain formed brown circular colonies sometimes in spindle form when grown in modified D2216 medium. They had a diameter of 0.8-1.1 mm and showed spreading after 24 h of growth at 55 °C on modified D2216 solid medium. GSsed3^T was motile as shown by the two methods.

Cell growth was monitored at different temperatures (25, 30, 35 40, 45, 50, 55, 60, 65, 70, and 80 °C), measuring optical density at 600 nm (OD_{600}) with a spectrophotometer Genesys 10S UV-Vis (Thermoscientific, UK). Salt tolerance was tested over 7 days in tryptic soy broth medium prepared in-house using 15 g/l tryptone, 5 g/l soytone and 15 g/l agar, pH 7.0, supplemented with various concentrations of NaCl (0, 0.5, 1, 2, 3, 5, 7, 10 and 20 % w/v). Cell growth was evaluated at OD_{600}. To determine the pH suitable for growth, cells were inoculated in TSB media adjusted to pH 3 to 11 (intervals of 0.5 pH unit), using acetate (for pH 3–5.5), phosphate (for pH 6–7.5) or glycine–NaOH (for pH 8.5–11) buffers at concentration 0.05 M, as previously suggested by Derekova et al, 2007. The need for oxygen during growth was verified using the method of thioglycollate medium [46]. According to the thioglycollate test, strains GSsed3, B2M1, B7M1 and B7M2 are facultative anaerobes. All growth experiments were performed in triplicates. Regarding the optimal growth conditions, strain GSsed3^T, as well as the related strains isolated from other geothermal reservoir (strains B2M1, B7M1, B7M2), displayed a growth range similar to other species of *Anoxybacillus* (Table 6.1).

All the biochemical tests used for the differentiation of strain GSsed3^T were conducted in parallel to closely related species with validly published names, as well as for the other strains isolated from the other geothermal reservoir The growth range of the isolateswas 40-65 °C, with a temperature optimum at 55 °C. Strain GSsed3^{T}and its related strains could tolerate up to 3 % w/v of NaCl, with an optimum at 0.5 % of NaCl. Growth was detected from pH 5 to 9.5. Optimum growth was observed at pH 7 to 8.Anaerobic growth was observed in contrast to *A. rupiensis*DSM 17127^T, which is a strict aerobe [16] (Table 6.1).

Nitrate reduction, casein hydrolysis, starch hydrolysis, and oxidase were performed as described previously [47]. For the casein hydrolysis assay, the milk agar medium was modified to the following composition: peptone 0.2 % (w/v), glucose 0.2 % (w/v), $Na_2HPO_4 \bullet 2H_2O$ 0.2 % (w/v), NaCl 1 %, agar 2 %, and skim-milk 10 %. Catalase, citrate utilization, gelatin hydrolysis, urea hydrolysis and aesculin hydrolysis were performed by methods described previously [43]. Mannitol test and indole production were tested as described previously [48]. Bacterial species used as positive and negative controls for each test are provided in supplementary table 6.2. All tests were inoculated with colonies from an overnight solid culture. Incubation was performed at the optimal growth temperature of each individual strain. With the exception of the hydrolysis of gelatine and aesculin and the reduction of nitrate, all the strains from this study (GSsed3^T, B2M1, B7M1, B7M2), were consistent in the results of the tests performed. All the strains were positive for catalase and oxidase, as well as for the hydrolysis of casein and starch. Strain GSsed3^T was negative in regards to nitrate reduction, negative to hydrolysis of gelatine and positive to the hydrolysis of aesculine, while the strains B2M1, B7M1, B7M2 were the opposite for these three tests. The most striking difference between strains GSsed3^T, B2M1, B7M1, B7M2 and *A. rupiensis* DSM 17127^T is the negative reaction for oxidase in the latter. Concerning other *Anoxybacillus* species, strain GSsed3^T shares the capability to perform the hydrolysis of aesculin with *A. caldiproteolyticus* DSM 15730^T and *A. tepidamans* DSM 16325^T.

Utilization of different carbon sources was assessed with two methods. First, the API 20NE system (bioMérieux, France) was performed according to the manufacturer's protocol. The API strips were incubated at 55 °C in sterile glass Petri dishes containing sterile MilliQ water to prevent evaporation. Complementary analyses for utilization of different compounds as sole carbon and energy source were performed using D-xylan, D-pectine, D-(-)-salicine, dulcitol, D-(+)-cellibiose, inuline, olive oil, Na-acetate, Na-propionate, D-(+)-mannose, D-(-)-fructose, D-(+)-galactose, D-glucose, D-(+)-lactose, D-(+)-melibiose, myo-inositol, ribitol, D-(+)-raffinose, D-(-)-ribose, L-(+)-rhamnose, D-(+)-sucrose, D-sorbitol, D-xylose, glycerol, Tween 60, Tween 80, L-arabinose, D-mannitol, D-maltose, potassium glyconate, adipic acid, malic acid, citrate, N-acetyl-glucosamide, phenyl acetic acid, capric acid, trisodium citrate, L-phenylalanine, L-tyrosine, xanthine and hypoxanthine were performed in modified Adkins basal medium (Adkins et al. 1992), containing (per litre of distilled water): 0.8 g of NaCl; 1.0 g NH_4Cl; 0.1 g KCl; 0.1 g K_2HPO_4; 0.2 g $MgSO_4 \bullet 7H_2O$; 0.02 g $CaCl_2 \bullet 2H_2O$; 0.2 g yeast extract (Merck, Germany); 10 g TES (N-(Tris(hydroxymethyl)methyl)-2-aminoethanesulfonic acid); 0.2 g $CaCO_3$ (precipitated chalk); Finally, 5 mL of a trace metal solution and 10 mL of a vitamin solution [49] were added. The pH of the medium was adjusted to 7. Carbohydrates and vitamin solutions were filter-sterilized (0.22 µm pore-size). Sugar solutions were added at 5 g/L in 10 mL of basal medium [50]. 100 µl of modified D2216 agar culture diluted in physiological water (9 g/L NaCl) (OD_{600}= 0.1) were inoculated and incubated under agitation at 55 °C. Two successive inoculations were undertaken to confirm assimilation (100 µl of culture in 10 mL of medium). Cultures were observed over 7 days. Assimilation was considered positive

when turbidity of the culture was different from the negative control [51]. With the exception of assimilation of D-(+)-cellibiose and D-(-)-fructose that were positive, and the negative assimilation of D-xylan, L-tyrosine, D-(+)-melibiose, Tween 80, xanthine and hypoxanthine, the utilization of other carbon sources varied among the four strains studied here (Table 6.1).

Characteristics	1	2	3	4	5	6	7	8	9
Temperature range (°C)	40-65	40-65	40-65	40-65	35-67	39-67	45-65	37-70	40-65
Optimal temperature (°C)	60	60	60	60	55	55	61	60	55
NaCl range (%w/v)	0-3	0-3	0-3	0-3	up to 1	up to 2	up to 2	up to 0.5	up to 3
pH range	5.0-9.5	5.0-9.5	5.0-9.5	5.0-9.5	5.5–8.5	6-9	5-6.5	5-9	7-8
Optimal pH	7.0-8.0	7.0-8.0	7.0-8.0	7.0-8.0	6.0-6.5	7.0-8.0	5.6	6.5-7	7-8
Oxygen Requirements	F.A	F.A	F.A	F.A	S.A	F.A	F.A	S.A	F.A
Catalase	+	+	+	+	+	+	+	+	+
Oxidase	+	+	+	+	-	-	-	+	+
Hydrolysis of Casein	+	+	+	+	+	-	-	+	-
Hydrolysis of Gelatin	-	+	+	+	-	-	-	+	-
Hydrolysis of Starch	+	+	+	+	+	+	+	+	-
Hydrolysis of Esculin	+	-	-	-	-	+		+	
Nitrate reduction to nitrite	-	-	-	-	-	+	+	-	+
Xylan	-	-	-	-	-	-	-	+	-
Pectine	+	+	-	+	-	+	+	+	+
Salicine	+	+	+	-	+	+	-	+	+
L-Phénylalanine	-	+	-	-	+	+	-	+	-
D(+)Raffinose	+	-	-	-	-	+	-	-	-
D(+)Sucrose	+	+	-	+	+	+	+	+	+
L-tyrosine	-	-	-	-	-	-	-	v	+
Melibiose	-	-	-	-	-	+	-	-	-
Inuline	-	+	-	-	-	+	-	+	+
Dulcitol	+	-	-	-	+	+	-	-	+
Cellibiose	+	+	+	+	+	+	-	+	+
Olive oil	+	-	-	-	-	-	-	-	-
Na-acetate	-	+	+	+	+	+	+	+	+
Na-propionate	-	+	-	-	+	+	-	-	-
Mannose	-	+	+	+	+	+	-	+	+
Sorbitol	-	+	+	+	-	-	-	-	-
Myo-inositol	-	+	+	+	+	+	-	-	-
Tween 80	-	-	-	-	-	-	+	+	+
Tween 60	-	-	+	-	-	+	-	+	+
Fructose	+	+	+	+	+	+	+	+	+
Ribitol	-	+	-	-	-	+	-	+	+
Xanthine	-	-	-	-	-	-	-	-	-
Hypoxanthine	-	-	-	-	-	-	-	-	-

G+C content mol %	46.9	xx	N.D.	N.D	41.7	43.2/42.4	43.5	40.2	43.9

Table 6.2. Differential characteristics of *Anoxybacillus geothermalis* sp. nov.strains GSsed3^T, B2M1, B7M1 and B7M2 and related species from the genus *Anoxybacillus.* 1.GSsed3^T; 2. B2M1; 3. B7M1; 4. B7M2; 5. *A. rupiensis* DSM 17127^T; 6. *A. tepidamans* DSM 16325^T; 7. *A. amylolyticus* DSM 15939^T; 8. *A. caldiproteolyticus* DSM 15730^T; 9. *A. voinovskiensis* DSM 17075^T. Characteristics scored as: +, positive; -, negative; N.D., not determined; S.A., strict aerobe; F.A., facultative anaerobe.

Assimilation of D-(-)-fructose and a negative result for the assimilation of xanthine and hypoxanthine were common traits to all the *Anoxybacillus* strains studied. In contrast, olive oil assimilation by GSsed3^T was a unique trait found in the characterization.

GC content was estimated to be 46.9 mol% based on the draft sequence of the full genome of strain GSsed3^T. Genome sequence and annotation of this strain has been submitted to Genbank under accession number YCG00000000.1 [52]. The GC content of GSsed3^T was slightly higher than the values reported for the reference strains (Table 6.1).

Analysis of fatty acid composition, respiratory menaquinones, peptidoglycan structure, and polar lipids were performed using the identification services of DSMZ (Germany). Fatty acid composition comparison among strain GSsed3^T, *A. rupiensis*DSM 17127^T, and *A. tepidamans*DSM 16325^{T}is shown in Table 6.2. The major fatty acids found in GSsed3^T corresponded to iso-$C_{15:0}$ (35.3 %) and iso-$C_{17:0}$ (26.4 %), confirming the affiliation of the strain to the genus *Anoxybacillus*. However, in contrast to *A. rupiensis* DSM 17127^T, and *A. tepidamans* DSM 16325^T, the fatty acids $C_{16:0}$ and iso-$C_{16:0}$ represented a significant fraction of the fatty acids of the strain (14.5 and 10.1 %, respectively). In contrast, for *A. rupiensis* DSM 17127^T these fatty acids represent 5.4 and 2.0 %, only. The significant contribution of these two fatty acids ($C_{16:0}$ and iso-$C_{16:0}$) is particularly noteworthy, as these fatty acids are rarely found representing 25 % of the total fatty acid content in other *Anoxybacillus* species [2,6,18]. The respiratory menaquinones found were MK-8 (92 %) and MK-7 (5 %). These appear also to be characteristic of strain GSsed3^T, compared to other *Anoxybacillus* species for which the information is available and that had as major respiratory menaquinone MK-7 [2,6,7,17,53]. The total hydrolysate (4N HCl, 16 h at 100°C) of the peptidoglycan contained the aminoacids *meso*-diaminopimelic acid (*meso*-Dpm), alanine (Ala) and glutamic acid (Glu). The partial hydrolysate (4 N HCl, 0.75 h at 100 °C) contained (in addition to the aminoacids) the peptides (L-Ala--D-glu and Dpm--D-ala). The peptidoglycan type of strain GSsed3^T is A1γ *meso*-Dpm-direct. The amount of peptidoglycan in the cell wall appeared to be rather low because meso-Dpm was also low in the hydrolysate (4 N HCl, 16 h at 100 °C) of whole cells. This observation was also reported for *Anoxybacillus calidus* [7]. Analysis of the peptidoglycan structure of other *Anoxybacillus* species needs to be performed, in order to conclude whether the cell walls of Anoxybacillus spp., in general, contain low amounts of peptidoglycan.

The analysis of polar lipids, with abundant DPG (Supplementary Figure 6.2), shows a polar lipid profile of the strain GSsed3^T similar to *A.caldiproteolyticus* as shown in the supplementary material of Coorevits *et al.* (2012). On the basis of the results presented in this study, we considered that strain GSsed3^T represents a novel species of the genus *Anoxybacillus*, for which we proposed the name *Anoxybacillusgeothermalis* sp. nov.

	A. geothermalis	*A. rupiensis*	*A. tepidamans*
iso-C14:0	0.3	0	0.6
C14:0	1.0	0.3	4.1
iso-C15:0	35.3	52.8	44.3
anteiso-C15:0	1.6	1.6	6.6
C15:0	0	0.3	0
iso3OH-C14:0	0.8	0	0
iso-C16:0	10.1	2.0	3.2
C16:0	14.5	5.4	15.1
anteiso-C16:0	0	0	0
iso-C17:0	26.4	33.6	15.0
anteiso-C17:0	8.4	3.9	6.1
iso-C18:0	0.4	0	0.6
C18:0	0.9	0	0

Table 6.2. Phospholipid fatty acid (PLFA) composition of GSsed3T cultured at 55°C and its closely related strains Anoxybacillusrupiensis DSM 17127T and Anoxybacillustepidamans DSM 16325T. Fatty acid analysis was performed under the same conditions using the analytical services of DSMZ, Germany.

6.2. Description of *Anoxybacillus geothermalis* sp. nov.

Anoxybacillus geothermalis [ge.o.ther.ma'lis. Gr. n. ge- earth, -thermalis of thermal properties or origin, N.L. masc. adj. geothermalis from hot earth, from geothermal site].

Gram-stain-positive, motile rod of 0.8-1 x 2.5-2.7 µm in size. Elliptic terminal or subterminal endospores are observed in slightly swollen sporangia. Colonies on modified D2216 medium are brownish, smooth, opaque and often spreading. The diameter if the colonies is 0.8-1.1 mm after 24 h of growth at 55 °C on modified D2216 solid medium. Growth occurs at 40-65°C (optimum 55 °C), at pH 5.0-9.5 (optimum 7.0-8.0) and with NaCl (0-3 %, w/v; optimum 0.5 %). It is a facultative anaerobe, catalase and oxidase positive, hydrolyses casein, starch and aesculine, but not gelatin. Nitrate is not reduced to nitrite. Negative for indole production. It ferments mannitol. Assimilates cellobiose, galactose, glucose, ribose, D-xylose, glycerol, L-arabinose, D-mannose, D-mannitol, D-maltose, potassium glyconate, sodium adipate, sodium malate, sucrose and pectin. It assimilates fructose and aerobically produces gas, but does not change the pH of the culture medium. It does not assimilate citrate, lactose, myoinositol, rhamnose, sorbitol, Tween 80, N-acetyl-glucosamide, phenyl acetic acid, capric acid, trisodium citrate, salicin, xylan, raffinose, ribitol, inulin. Phenylalanine is not deaminated, tyrosine is not degraded. The major cellular fatty acids are: iso-$C_{15:0}$, iso-$C_{17:0}$, $C_{16:0}$ and iso-$C_{16:0}$. The major respiratory menaquinones are MK-8 and MK-7. The peptidoglycan type is A1γ meso-Dpm-direct. The polar lipid profile consisted of phosphatidylethanolamine, phosphatidylglycerol, diphosphatidylglycerol, one unknown phosphoglycolipid and one unknown phospholipid.

The type strain GSsed3^T (= CCOS808^T = ATCC BAA2555^T) was isolated from deposits from filter in Gross Schoenebeck power plant, Germany. The DNA content of the type strain is 46.9 mol%.

The GenBank accession number for the 16S rRNA gene sequence of *Anoxybacillus geothermalis* GSsed3^T is KJ722458.

6.3. Acknowledgments

Funding from the Swiss National Science Foundation (Grants31003A_132358/1 and 31003A_152972) and the Foundation Pierre Mercier pour la Science are acknowledged. DNA-DNA hybridization and the analysis of respiratory quinoneswere carried out by the identification service of the DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen, Germany). Sequencing was conducting using the services of Microsynth AG (Switzerland) and GATC Biotech (Germany). Special thanks to Mr. Andrej Al-Dourobi and Ms. Amandine Pillonel for their excellent technical assistance and to Ms. Elvira Feldbusch for help during the sampling. Thanks also to Mr. Kai Mangelsdorf for his help in discussion of the PFLA results.

6.4. References

[1] Schaffer C, Franck WL, Scheberl A, Kosma P, McDermott TR, Messner P. Classification of isolates from locations in Austria and Yellowstone National Park as *Geobacillus tepidamans* sp nov. Int J Syst Evol Microbiol 2004;54:2361–8. doi:10.1099/ijs.0.63227-0.

[2] Coorevits A, Dinsdale AE, Halket G, Lebbe L, De Vos P, Van Landschoot A, et al. Taxonomic revision of the genus *Geobacillus*: emendation of *Geobacillus, G. stearothermophilus, G. jurassicus, G. toebii, G. thermodenitrificans* and *G. thermoglucosidans* (nom. corrig., formerly "*thermoglucosidasius*"); transfer of *Bacillus thermantarcticus* to the genus as *G. thermantarcticus* comb. nov.; proposal of *Caldibacillus debilis* gen. nov., comb. nov.; transfer of *G. tepidamans* to *Anoxybacillus* as *A. tepidamans* comb. nov.; and proposal of *Anoxybacillus caldiproteolyticus* sp. nov. Int J Syst Evol Microbiol 2012;62:1470–85. doi:10.1099/ijs.0.030346-0.

[3] Pikuta E, Lysenko A, Chuvilskaya N, Mendrock U, Hippe H, Suzina N, et al. *Anoxybacillus pushchinensis* gen. nov., sp. nov., a novel anaerobic, alkaliphilic, moderately thermophilic bacterium from manure, and description of *Anoxybacillus flavitherms* comb. nov. Int J Syst Evol Microbiol 2000;50 Pt 6:2109–17.

[4] Atanassova M, Derekova A, Mandeva R, Sjøholm C, Kambourova M. *Anoxybacillus bogrovensis* sp. nov., a novel thermophilic bacterium isolated from a hot spring in Dolni Bogrov, Bulgaria. Int J Syst Evol Microbiol 2008;58:2359–62. doi:10.1099/ijs.0.65745-0.

[5] Belduz AO, Dulger S, Demirbag Z. *Anoxybacillus gonensis* sp. nov., a moderately thermophilic, xylose-utilizing, endospore-forming bacterium. Int J Syst Evol Microbiol 2003;53:1315–20.

[6] Cihan AC, Ozcan B, Cokmus C. *Anoxybacillus salavatliensis* sp. nov., an alpha-glucosidase producing, thermophilic bacterium isolated from Salavatli, Turkey. J Basic Microbiol 2011;51:136–46. doi:10.1002/jobm.201000115.

[7] Cihan AC, Cokmus C, Koc M, Ozcan B. *Anoxybacillus calidus* sp. nov., a thermophilic bacterium isolated from soil near a thermal power plant. Int J Syst Evol Microbiol 2014;64:211–9. doi:10.1099/ijs.0.056549-0.

[8] De Clerck E, Rodríguez-Díaz M, Vanhoutte T, Heyrman J, Logan NA, De Vos P. *Anoxybacillus contaminans* sp. nov. and *Bacillus gelatini* sp. nov., isolated from contaminated gelatin batches. Int J Syst Evol Microbiol 2004;54:941–6.

[9] Dai J, Liu Y, Lei Y, Gao Y, Han F, Xiao Y, et al. A new subspecies of *Anoxybacillus flavithermus* ssp. *yunnanensis* ssp. nov. with very high ethanol tolerance. FEMS Microbiol Lett 2011;320:72–8. doi:10.1111/j.1574-6968.2011.02294.x.

[10] Dulger S, Demirbag Z, Belduz AO. *Anoxybacillus ayderensis* sp. nov. and *Anoxybacillus kestanbolensis* sp. nov. Int J Syst Evol Microbiol 2004;54:1499–503. doi:10.1099/ijs.0.02863-0.

[11] Kevbrin VV, Zengler K, Lysenko AM, Wiegel J. *Anoxybacillus kamchatkensis* sp. nov., a novel thermophilic facultative aerobic bacterium with a broad pH optimum from the Geyser valley, Kamchatka. Extremophiles 2005;9:391–8. doi:10.1007/s00792-005-0479-7.

[12] Namsaraev ZB, Babasanova OB, Dunaevsky YE, Akimov VN, Barkhutova DD, Gorlenko VM, et al. *Anoxybacillus mongoliensis* sp. nov., a novel thermophilic proteinase producing bacterium isolated from alkaline hot spring, central Mongolia. Mikrobiologiia 2010;79:516–23.

[13] Poli A, Esposito E, Lama L, Orlando P, Nicolaus G, de Appolonia F, et al. *Anoxybacillus amylolyticus* sp. nov., a thermophilic amylase producing bacterium isolated from Mount Rittmann (Antarctica). Syst Appl Microbiol 2006;29:300–7. doi:10.1016/j.syapm.2005.10.003.

[14] Yumoto I, Hirota K, Kawahara T, Nodasaka Y, Okuyama H, Matsuyama H, et al. *Anoxybacillus voinovskiensis* sp. nov., a moderately thermophilic bacterium from a hot spring in Kamchatka. Int J Syst Evol Microbiol 2004;54:1239–42. doi:10.1099/ijs.0.02889-0.

[15] Zhang C-M, Huang X-W, Pan W-Z, Zhang J, Wei K-B, Klenk H-P, et al. *Anoxybacillus tengchongensis* sp. nov. and *Anoxybacillus eryuanensis* sp. nov., facultatively anaerobic, alkalitolerant bacteria from hot springs. Int J Syst Evol Microbiol 2011;61:118–22. doi:10.1099/ijs.0.020834-0.

[16] Derekova A, Sjoholm C, Mandeva R, Kambourova M. *Anoxybacillus rupiensis* sp Nov., a novel thermophilic bacterium isolated from Rupi basin (Bulgaria). Extremophiles 2007;11:577–83. doi:10.1007/s00792-007-0071-4.

[17] Inan K, Belduz AO, Canakci S. *Anoxybacillus kaynarcensis* sp. nov., a moderately thermophilic, xylanase producing bacterium. J Basic Microbiol 2013;53:410–9. doi:10.1002/jobm.201100638.

[18] Poli A, Romano I, Cordella P, Orlando P, Nicolaus B, Ceschi Berrini C. *Anoxybacillus thermarum* sp. nov., a novel thermophilic bacterium isolated from thermal mud in Euganean hot springs, Abano Terme, Italy. Extrem Life Extreme Cond 2009;13:867–74. doi:10.1007/s00792-009-0274-y.

[19] Zhang X-Q, Zhang Z-L, Wu N, Zhu X-F, Wu M. *Anoxybacillus vitaminiphilus* sp. nov., a strictly aerobic and moderately thermophilic bacterium isolated from a hot spring. Int J Syst Evol Microbiol 2013;63:4064–71. doi:10.1099/ijs.0.050096-0.

[20] Feldbusch E, Regenspurg S, Banks J, Milsch H, Saadat A. Alteration of fluid properties during the initial operation of a geothermal plant: results from in situ measurements in Groß Schönebeck. Environ Earth Sci 2013;70:3447–58. doi:10.1007/s12665-013-2409-9.

[21] Regenspurg S, Helmholtz EF, Saadat A. Corrosion Processes at the Geothermal Site Groß Schönebeck (North German Basin). Corros 2013 2013.

[22] Vandamme P, Pot B, Gillis M, DeVos P, Kersters K, Swings J. Polyphasic taxonomy, a consensus approach to bacterial systematics. Microbiol Rev 1996;60:407–+.

[23] Muyzer G, Teske A, Wirsen C, Jannasch H. Phylogenetic-Relationships of *Thiomicrospira* Species and their Identification in Deep-Sea Hydrothermal Vent Samples by Denaturing Gradient Gel-Electrophoresis of 16S rDNA Fragments. Arch Microbiol 1995;164:165–72. doi:10.1007/BF02529967.

[24] Ovreås L, Forney L, Daae FL, Torsvik V. Distribution of bacterioplankton in meromictic Lake Saelenvannet, as determined by denaturing gradient gel electrophoresis

of PCR-amplified gene fragments coding for 16S rRNA. Appl Environ Microbiol 1997;63:3367–73.
[25] Kim O-S, Cho Y-J, Lee K, Yoon S-H, Kim M, Na H, et al. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. Int J Syst Evol Microbiol 2012;62:716–21.
[26] Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res 2008;36:W465–9. doi:10.1093/nar/gkn180.
[27] Anisimova M, Gascuel O. Approximate likelihood-ratio test for branches: A fast, accurate, and powerful alternative. Syst Biol 2006; 55:539–52. doi: 10.1080/10635150600755453.
[28] Castresana J. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol Biol Evol 2000;17:540–52.
[29] Elias I, Lagergren J. Fast computation of distance estimators. BMC Bioinformatics 2007;8:89. doi:10.1186/1471-2105-8-89.
[30] Gascuel O. BIONJ: An improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol 1997;14:685–95.
[31] Goboloff P, Farris S, Nixon K. TNT (Tree analysis using New Technology) ver. 1.1 2000. Publ AuthorsTucumán Argent n.d.
[32] Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol 2010;59:307–21. doi:10.1093/sysbio/syq010.
[33] Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic trees. Bioinforma Oxf Engl 2001;17:754–5.
[34] Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinforma Oxf Engl 2003;19:1572–4.
[35] Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol 1970;48:443–53. doi:10.1016/0022-2836(70)90057-4.
[36] Cashion P, Holderfranklin M, McCully J, Franklin M. Rapid Method for Base Ratio Determination of Bacterial DNA. Anal Biochem 1977;81:461–6. doi:10.1016/0003-2697(77)90720-5.
[37] DeLey J, Cattoir H, Reynaert, A. Quantitative Measurement of DNA Hybridization from Renaturation Rates. Eur J Biochem 1970;12:133-.
[38] Huss V, Festl H, Schleifer K. Studies on the Spectrophotometric Determination of DNA Hybridization from Renaturation Rates. Syst Appl Microbiol 1983;4:184–92.
[39] Wayne L, Brenner D, Colwell R, Grimont P, Kandler O, Krichevsky M, et al. Report of the Ad-Hoc-Committee on Reconciliation of Approaches to Bacterial Systematics. Int J Syst Bacteriol 1987;37:463–4.
[40] Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M. Genome sequence-based species delimitation with confidence intervals and improved distance functions. BMC Bioinformatics 2013;14:60. doi:10.1186/1471-2105-14-60.

[41] Rodriguez-R LM, Konstantinidis KT. Bypassing cultivation to identify bacterial species. Issues 2010.
[42] Logan NA, Berge O, Bishop AH, Busse H-J, De Vos P, Fritze D, et al. Proposed minimal standards for describing new taxa of aerobic, endospore-forming bacteria. Int J Syst Evol Microbiol 2009;59:2114–21. doi:10.1099/ijs.0.013649-0.
[43] Gerhardt P. Methods for general and molecular bacteriology. Washington, D.C.: American Society for Microbiology; 1994.
[44] Kohlmeier S, Smits THM, Ford RM, Keel C, Harms H, Wick LY. Taking the fungal highway: Mobilization of pollutant-degrading bacteria by fungi. Environ Sci Technol 2005;39:4640–6. doi:10.1021/es047979z.
[45] Rashid MH, Kornberg A. Inorganic polyphosphate is needed for swimming, swarming, and twitching motilities of *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A 2000;97:4885–90. doi:10.1073/pnas.060030097.
[46] Brewer JH. Clear liquid mediums for the "aerobic" cultivation of anaerobes. J Am Med Assoc 1940;115:598–600.
[47] Harrigan FW, McCance ME. Laboratory Methods in Microbiology. Academic Press.; 1987.
[48] Joffin J-N, Leyral G. Microbiologie technique.: Tome 1, Dictionnaire des techniques, 3ème édition. Centre régional de documentation pédagogique d'Aquitaine; 2001.
[49] Tanner R. Monitoring Sulfate-Reducing Bacteria - Comparison of Enumeration Media. J Microbiol Methods 1989;10:83–90. doi:10.1016/0167-7012(89)90004-3.
[50] Nazina TN, Tourova TP, Poltaraus AB, Novikova EV, Grigoryan AA, Ivanova AE, et al. Taxonomic study of aerobic thermophilic bacilli: descriptions of *Geobacillus subterraneus* gen. nov., sp nov and *Geobacillus uzenensis* sp nov from petroleum reservoirs and transfer of *Bacillus stearothermophilus, Bacillus thermocatenulatus, Bacillus thermoleovorans, Bacillus kaustophilus, Bacillus thermoglucosidasius* and *Bacillus thermodenitrificans* to *Geobacillus* as the new combinations *G-stearothermophilus, G-thermocatenulatus, G-thermoleovorans, G-kaustophilus, G-thermoglucosidasius* and *G-thermodenitrificans*. Int J Syst Evol Microbiol 2001;51:433–46.
[51] Zhao H-P, Wu Q-S, Wang L, Zhao X-T, Gao H-W. Degradation of phenanthrene by bacterial strain isolated from soil in oil refinery fields in Shanghai China. J Hazard Mater 2009;164:863–9. doi:10.1016/j.jhazmat.2008.08.098.
[52] Filippidou S, Jaussi M, Junier T, Wunderlin T, Roussel-Delif L, Jeanneret N, et al. Genome Sequence of *Anoxybacillus geothermalis* Strain GSsed3, a Novel Thermophilic Endospore-Forming Species. Genome Announc 2015;3. doi:10.1128/genomeA.00575-15.
[53] Gul-Guven R, Guven K, Poli A, Nicolaus B. *Anoxybacillus kamchatkensis* subsp. *asaccharedens* subsp. nov., a thermophilic bacterium isolated from a hot spring in Batman. J Gen Appl Microbiol 2008;54:327–34.

6.5. Supplementary Material

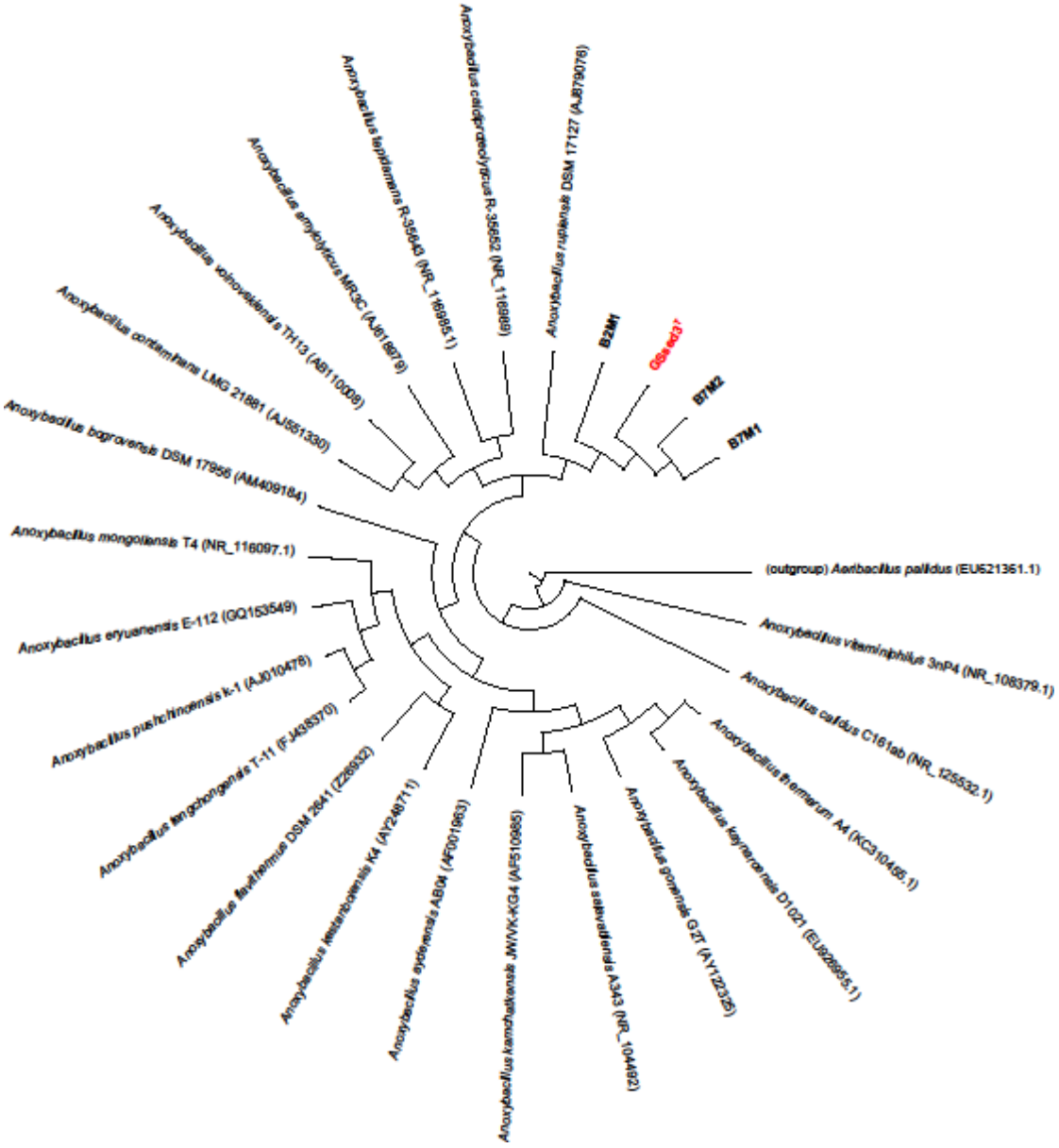
Supplementary Table 6.1. Comparison of similarities between the novel described strains GSsed3^T, B2M1, B7M1, and B7M2 and all *Anoxybacillus* species using local alignment (BLASTn) and pair-wise alignment (Needleman-Wunsch) of the 16S rRNA gene. Values are provided in % of similarity.

Strains	GSsed3^T		B2M1		B7M1		B7M2	
	BlastN	NW	BlastN	NW	BlastN	NW	BlastN	NW
A. rupiensis	99,6	97,6	99,76	81	99,8	96,8	99,8	96,8
A. tepidamans	98,14	95,5	98,48	81	98,2	96	98,2	96
A. amylolyticus	97,6	97,2	97,83	79	97,58	94	97,58	94
A. caldiproteolyticus	97,48	94,8	98,39	81	97,7	95,3	97,7	95,3
A. contaminans	96,75	95,8	97,59	78,9	96,76	93,2	96,76	93,2
A. calidus	96,75	90,4	97,75	84,2	96,88	94,1	96,88	94,1
A. voinovskiensis	96,62	94	98,14	81	97,88	94,6	97,88	94,6
A. vitaminiphilus	96,56	94,1	97,59	80,5	96,63	94,8	96,63	94,8
A. kamchatkensis	96,2	95,6	97,03	78	95,83	92,1	95,83	92,1
A. flavithermus	96,01	94,2	97,35	79,4	95,81	93,2	95,81	93,2
A. ayderensis	95,99	87,4	97,16	84,3	95,88	85,7	95,88	85,7
A. salavatliensis	95,92	86	96,95	86,2	95,76	89,5	95,76	89,5
A. thermarum	95,78	87,2	97,02	80,9	95,29	89,3	95,29	89,3
A. mongoliensis	95,61	89,1	96,95	83,1	95,5	92,7	95,5	92,7
A. bogrovensis	95,61	81,7	97,14	77,9	95,37	85	95,37	85
A. kaynarcensis	95,42	86,9	97	81,2	94,81	89	94,81	89
A. gonensis	95,32	84,6	96,49	80,6	95,2	88	95,2	88
A. tengchongensis	95,2	92,9	96,32	79,2	95,15	93,3	95,15	93,3
A. eryuanensis	95,07	92,5	96,62	78,8	95,14	92,8	95,14	92,8
A. pushchinoensis	94,65	81,6	96,75	83,7	94,97	84,9	94,97	84,9
A. kestanbolensis	94,52	84,2	96	80,4	95	87,6	95	87,6

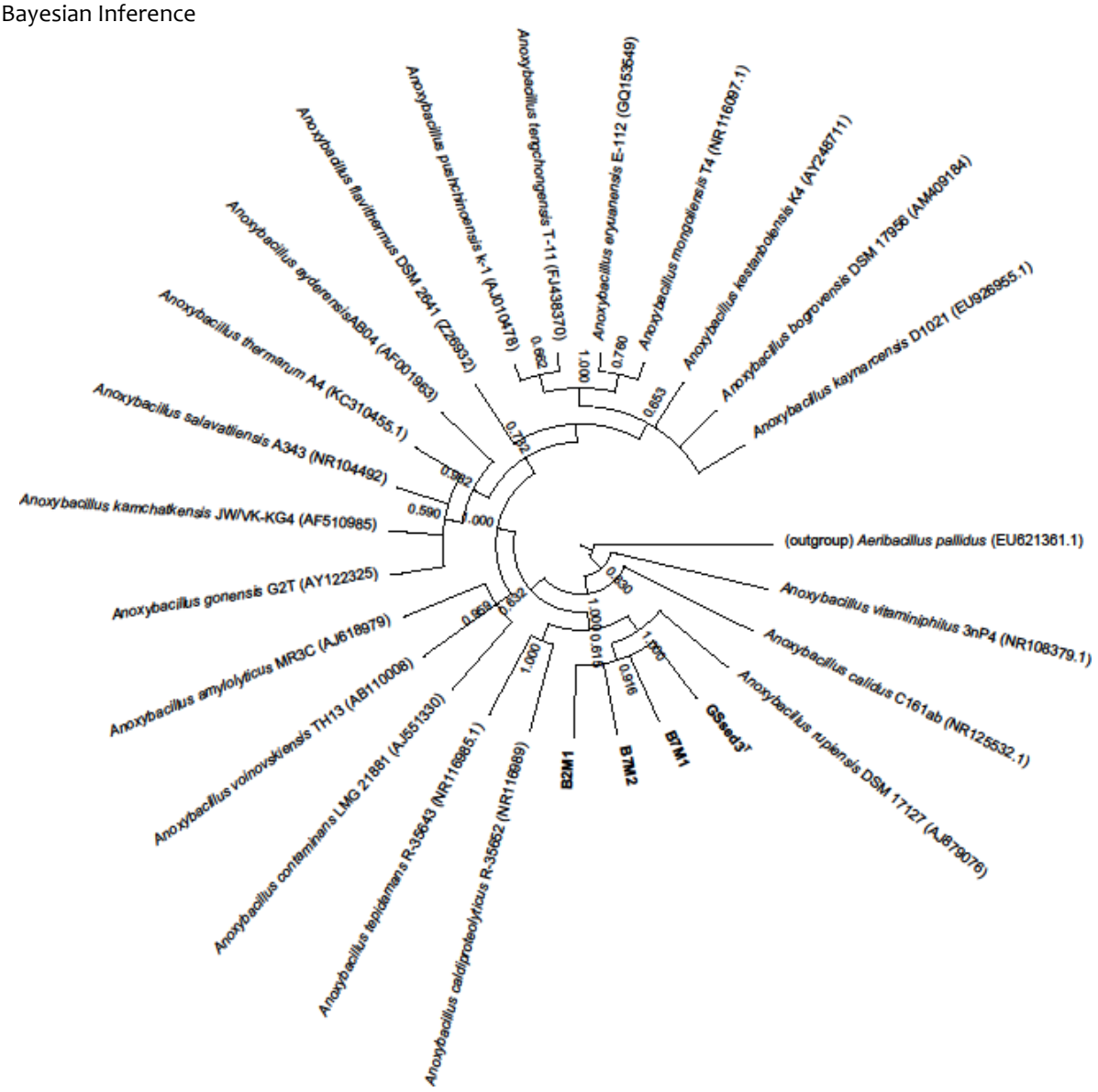
Supplementary Table 6.2. Positive and Negative controls used for all biochemical tests performed.

Biochemical test	Positive	Negative
Catalase	*Kocuriarhizophila*	*Enterococcus faecalis*
Oxidase	*Pseudomonas putida*	*Acinetobacter beijerinckii*
Hydrolysis of Casein	*Massiliatimonae*	*Enterococcus faecalis*
Hydrolysis of Gelatin	API test performed	
Hydrolysis of Starch	*Serratiaureilytica*	
Hydrolysis of Esculin	*Enterococcus faecalis*	*Escherichia coli*
Nitrate reduction to nitrite	API test performed	

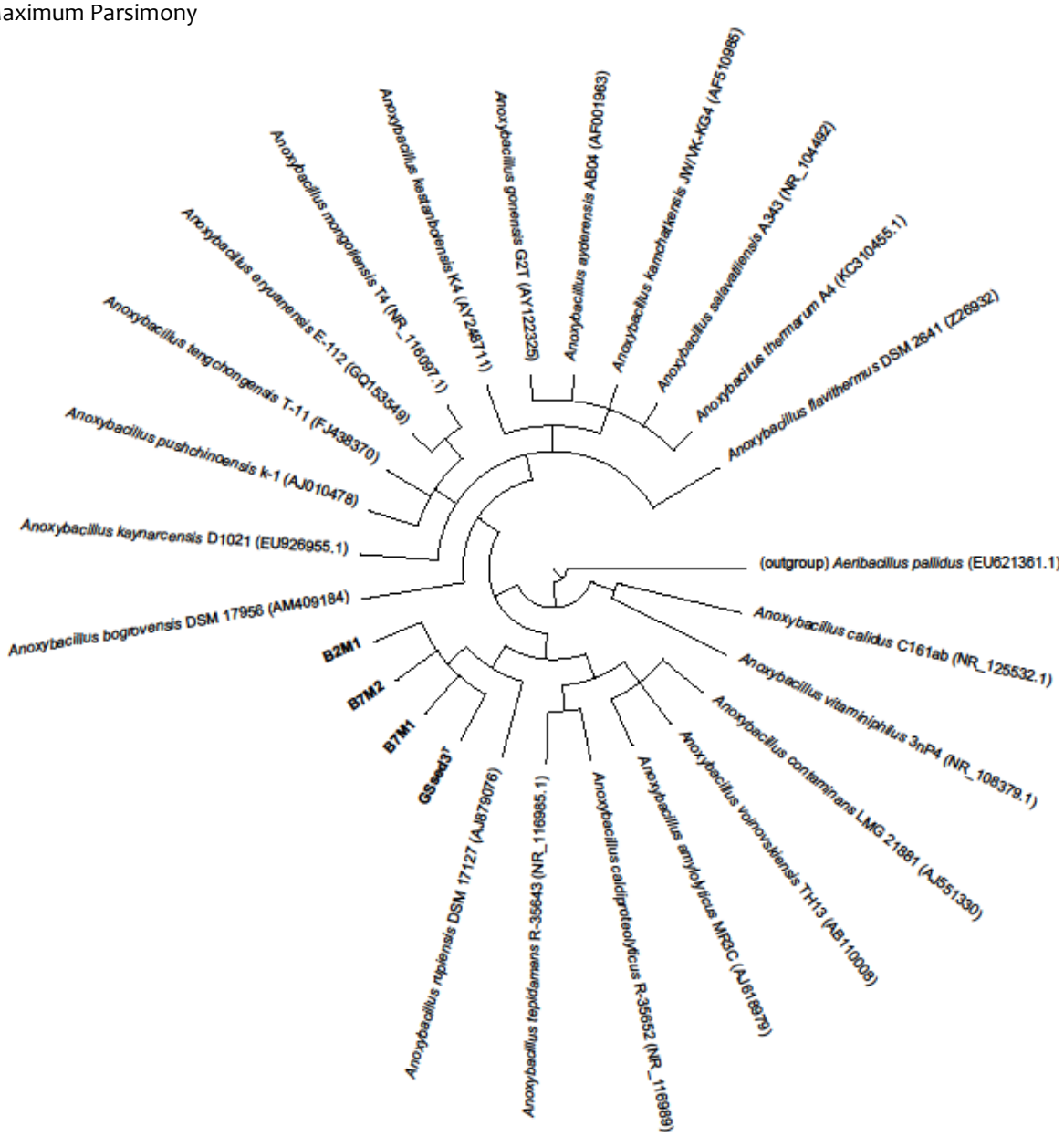
Supplementary Figure 6.1A. Phylogenetic trees of all *Anoxybacillus* species calculated using 4 different algorithms, all of which support the hypothesis that the newly isolated and described strains GSsed3^T, B2M1, B7M1 and B7M2 form a well-supported clade which is distinct to the other *Anoxybacillus* species. A. Maximum Likelihood



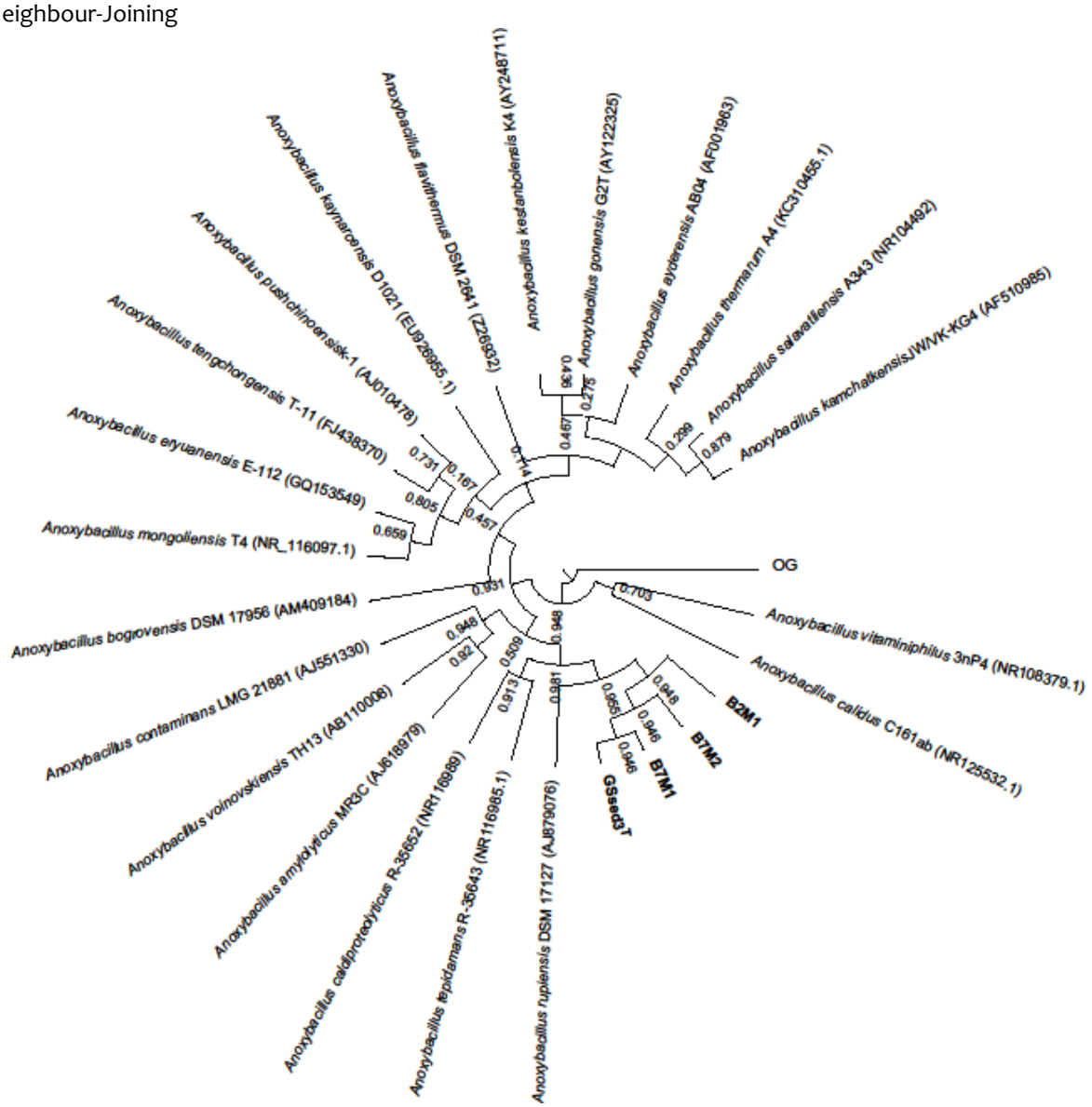
Supplementary Figure 6.1B. Phylogenetic trees of all *Anoxybacillus* species calculated using 4 different algorithms, all of which support the hypothesis that the newly isolated and described strains GSsed3^{T}, B2M1, B7M1 and B7M2 form a well-supported clade which is distinct to the other *Anoxybacillus* species. B. Bayesian Inference



Supplementary Figure 6.1C. Phylogenetic trees of all *Anoxybacillus* species calculated using 4 different algorithms, all of which support the hypothesis that the newly isolated and described strains GSsed3^{T}, B2M1, B7M1 and B7M2 form a well-supported clade which is distinct to the other *Anoxybacillus* species. C. Maximum Parsimony

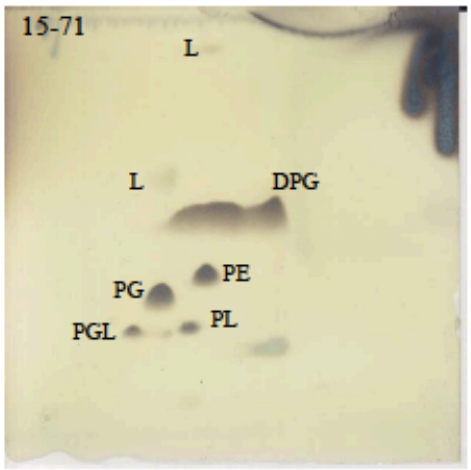


Supplementary Figure 6.1D. Phylogenetic trees of all *Anoxybacillus* species calculated using 4 different algorithms, all of which support the hypothesis that the newly isolated and described strains GSsed3^{T}, B2M1, B7M1 and B7M2 form a well-supported clade which is distinct to the other *Anoxybacillus* species. D. Neighbour-Joining



Supplementary Figure 6.2. Polar lipid report for the strain GSsed3^T

DSMZ Identification Services polar lipid report



L = Lipid

PL = Phospholipid

PE = Phosphatidylethanolamine

PG = Phosphatidylglycerol

DPG = Diphosphatidylglycerol

PGL = Phosphoglycolipid

Chapter 7
Comparative genomics of endospore-forming bacteria isolates for the discovery of potential biomarkers of extremity

This chapter is an overall presentation of the isolates from our worldwide campaign, whose genomes were fully sequenced and annotated. All these genomes are now publically available and three of them have been published in the journal Genome Announcements.

- Filippidou S., Wunderlin T., Junier T., Jeanneret N., Johnson S., McMurry K., Gleasner C.D., Lo C.-C., Li P.E., Vuyisich M., Chain P.S., Junier P. Genome sequence of *Bacillus alveayuensis* strain 24KAM51, a halotolerant thermophile isolated from a hydrothermal vent. Genome Announcements. 2015 Jul; 3(4): e00982-15. doi:10.1128/genomeA.00982-15.
- Filippidou S., Jaussi M., Junier T., Wunderlin T., Jeanneret N., Regenspurg S., Li P.E., Lo C.-C., Johnson S., McMurry K., Gleasner C.D., Vuyisich M., Chain P.S., Junier P. Genome sequence of *Aeribacillus pall*idus strain GS3372, an endospore-forming bacterium isolated in a deep geothermal reservoir. Genome Announcements. 2015 Jul; 3(4): e00981-15. doi:10.1128/genomeA.00981-15.
- Filippidou S., Jaussi M., Junier T., Wunderlin T., Roussel-Delif L., Jeanneret N., Vieth-Hillebrand A., Vetter A., Regenspurg S., Johnson S., McMurry K., Gleasner C.D., Lo C.-C., Li P.E., Vuyisich M., Chain P.S., Junier P. Genomesequence of *Anoxybacillus geothermalis* strain GSsed3, a novelthermophilic endospore-formingspecies. Genome Announcements. 2015 May; 3(3): e00575-15. doi:10.1128/genomeA.00575-15.

Any reference to the data or the context of this chapter should cite the above-mentioned publications.

7.1. Published genomes

7.1.1. Genome sequence of *Bacillus alveayuensis* strain 24KAM51, a halotolerant thermophile isolated from a hydrothermal vent

Bacillus alveayuensis strain 24KAM51 was isolated from a marine hydrothermal vent in Milos, Greece. Its genome depicts interesting features of halotolerance and resistance to heavy metals. *Bacillus alveayuensis* is a thermophilic endospore-forming bacterium. The type strain MT1 was isolated from deep-sea sediment [1]. However, to date, no genome is available for this species. Strain 24KAM51 was isolated from a hydrothermal vent on the coastal line of Alykes beach in Milos, Greece (36°42′353″ N, 24°28′197″ E, depth 1.5 m). It is able to grow at a temperature range of 30 to 80°C, with an optimum at 60°C. It can also tolerate acidic and alkaline conditions (pH growth range from 3 to 10), with an optimum at 7. Finally, it grows with up to 13% (wt/vol) NaCl. Based on 16S rRNA sequence identity, strain 24KAM51 is closely related to *B. alveayuensis* type strain TM1 (99% identity). A series of physiological tests also showed similarity to TM1. The genome of 24KAM51 was sequenced and annotated in order to contribute to a better understanding of the thermophilic and halotolerant lifestyle of extremophilic bacilli. To date, this is the only genome of *B. alveayuensis* publicly available.

Genomic DNA was extracted from an overnight culture using the QIAamp DNA minikit (Qiagen GmbH, Germany). For the draft genome of *B. alveayuensis* strain 24KAM51, an Illumina [2] short-insert (300 ± 70 bp) library was constructed and sequenced generating 43,458,058 reads, totaling 4.39 Mbp. The Illumina draft data were assembled with Velvet, version 1.2.08 [3]. The estimated size of the genome is 6.7 Mbp, which provides 409× coverage of the genome. Genome annotation was performed using an Ergatis-based [4] workflow with minor manual curation and visualized with the Artemis genome browser and annotation tool [5]. The G+C content is 38.1%. The complete genome sequence contained 6,597 genes, 10 rRNAs (5S, 16S, and 23S), 88 tRNAs, and 2 noncoding RNAs (ncRNAs) predicted. The genome of 24KAM51 was annotated, and its proteome revealed the presence of 198 genes related to the sporulation and germination pathways and 10 genes related to dipicolinic acid synthesis. Fifty-three proteins related to flagellar motility were found in the proteome of *B. alveayuensis* strain 24KAM51.

Many proteins are related to halotolerance in bacilli [6], some of which are present in multiple copies in the 24KAM51 proteome. More precisely, genes for Na^+/H^+ antiporters, proline/Na^+ symporters, and glycine/betaine ABC transporters were found. Although isolated from a natural hot spring, its genome contains genes related to copper (copper-binding proteins and multicopper oxidase), manganese (manganese transporters and permeases), cadmium (cadmium transporter), zinc (zinc metalloproteases, proteases, and transporters), and arsenic (arsenic resistance protein, ArsB) resistance. Moreover, it possesses genes encoding the NarH and NarZ proteins (nitrate reduction), as well as DsrE (sulfur reduction).

This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no.JYCE00000000. The version described here is JYCE00000000.1. This work was financially supported by the Swiss National Science Foundationproject

31003A_152972 and by Fondation Pierre Mercier pour la Science. The genetic information downloaded from GenBank is considered to be part of the genetic patrimony of Greece, the country from which the sample was obtained. Users of this information agree to: 1) acknowledge Greece as the country of origin in any country where the genetic information is presented and 2) contact the CBD focal point and the ABS focal point identified at the CBD website http://www.cbd.int/information/nfp.shtml if they intend to use the genetic information for commercial purposes.

7.1.2. Genome sequence of *Aeribacillus pallidus* strain GS3372, an endospore-forming bacterium isolated in a deep geothermal reservoir

The genome of strain GS3372 is the first publicly available strain of *Aeribacillus pallidus*. This endospore-forming thermophilic strain was isolated from a deep geothermal reservoir. The availability of this genome can contribute to the clarification of the taxonomy of the closely related *Anoxybacillus*, *Geobacillus*, and *Aeribacillus* genera. *Aeribacillus pallidus* is the only species of the novel *Aeribacillus* genus[7]. This species was previously classified as *Bacillus pallidus* [8]and later as *Geobacillus pallidus* [9]. The type strain (strain H12) was isolated from thermally treated sewage, and its full genome is not publicly available. More strains of this species have been isolated from hot springs[10,11], production water from an oil reservoir [12], and crude-contaminated soil [13]. Strain GS3372 was isolated multiple times from fluid samples (50 to 70°C) from the deep geothermal reservoir of Gross Schoenebeck, in the North German Basin (52°54′13.15″ N, 13°36′5.43″ E). Analysis of the 16S rRNA gene sequence (99% identity) and DNA-DNA hybridization (79.5% similarity) indicated that GS3372 is a novel strain of *A. pallidus*.

Genomic DNA was extracted from an overnight culture using the QIAamp DNA minikit (Qiagen GmbH, Germany). For this genome, an Illumina short-insert paired-end library was constructed and sequenced, which generated 4.744 Mbp of data, of which 2.41 Mbp is included in the final assembly (482× genome coverage). The data were assembled with Velvet, version 1.2.08 [3], to an estimated size of 4.9 Mbp with a 57.4% G+C content. Genome annotation utilized an Ergatis-based [4] workflow with minor manual curation then visualized with the Artemis genome browser and annotation tool [5]. The complete genome sequence contained 5,015 genes, 9 rRNAs (5S, 16S, and 23S), 69 tRNAs, and 4 noncoding RNAs (ncRNAs) predicted. The proteome of GS3372 revealed the presence of 96 genes related to the sporulation and germination pathways.

The GS3372 genome contains copper oxidase and manganese catalase genes, as well as genes encoding proteins related to arsenic (ArsB) and copper (CopZ) resistance. While the *A. pallidus* type strain is non-motile, strain GS3372 is motile, and accordingly, the loci *flhA* and *flgG* related to flagellar synthesis are present in its genome. Moreover, the ability of this strain to assimilate a large number of carbon sources is depicted in the genome by the presence of transporter, permease, and isomerase genes for glucose, D-xylose, glycerol, ribose, and mannose, all of which are supported by biochemical characterization (S. Filippidou and P. Junier, unpublished data). Finally, YflT which is related to heat-induced thermotolerance [14], is found in its proteome. These observations act in

accordance with the ecology of the strain, which was isolated in a thermophilic and oligotrophic environment. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no.JYCD00000000. The version described here is JYCD00000000.1. This work was supported by the Swiss National Science Foundation project 31003A_152972 and by Fondation Pierre Mercier pour la Science.

7.1.3. Genome Sequence of *Anoxybacillus geothermalis* strain GSsed3, a novel thermophilic endospore-forming species

Anoxybacillus geothermalis strain GSsed3 is an endospore-forming thermophilic bacterium isolated from filter deposits in a geothermal site. This novel species has a larger genome size (7.2 Mb) than that of any other *Anoxybacillus* species, and it possesses genes that support its phenotypic metabolic characterization and suggest an intriguing link to metals. From 2013 to date, 10 genome sequences of a total of 23 species that belong to the *Anoxybacillus* genus have been announced. GSsed3 (also known as ATCC BAA-2555) is the type strain of the novel species *Anoxybacillus geothermalis*, isolated in 2011 from above-ground filter deposits of the geothermal site in Gross Schönebeck, in the North German Basin (52°54′13.15″N 13°36′5.43″E) (S. Filippidou, M. Jaussi, N. Jeanneret, L. Roussel-Delif, T. Wunderlin, A. Vieth-Hillebrand, A. Vetter, P. Chain, S. Regenspurg, and P. Junier, unpublished data). All *Anoxybacillus* species are moderately thermophilic, and most of them have been isolated from geothermal sites. According to its 16S rRNA gene sequence, *A. geothermalis* strain GSsed3 is closely related to *Anoxybacillus rupiensis* [15], whose full genome sequence is not yet available. The genome of GSsed3 was sequenced and annotated, and its metabolic capabilities were revealed.

Genomic DNA was extracted from an overnight culture using the QIAamp DNA minikit (Qiagen GmbH, Germany). The draft genome of A. geothermalis strain GSsed3 was generated by the Los Alamos National Laboratory (LANL) Genome Science Group using Illumina [2] technology. For this genome, an Illumina short-insert paired-end library was constructed and sequenced, generating 21,691,466 reads totaling 2,191 Mbp. The Illumina draft data were assembled with Velvet version 1.2.08 [3]. The estimated size of the genome is 7.2 Mb, and the final assembly is based on 1,330 Mbp of Illumina draft data, which provide 185× coverage of the genome. The genomic DNA G+C content is estimated to be 46.8 mol%. Genome annotation was performed using an Ergatis-based [4] workflow with minor manual curation and visualized with the Artemis genome browser and annotation tool [5].

The complete genome sequence contained 7,003 genes, 7 rRNAs (5S, 16S, and 23S), 72 tRNAs, and 2 noncoding RNAs (ncRNAs) predicted. This strain possesses a large genome compared to those of other *Anoxybacillus* species, whose genomes do not exceed 3.3 Mb. In addition to its size, the genome contains more than double the number of genes in other published genomes of this genus [16–22]. Among the annotated genes, 143 genes encoding proteins related to sporulation and nine genes related to dipicolinic acid synthesis were found. GSsed3 was biochemically tested and found to hydrolyze urea and reduce nitrates to nitrites. Genome annotation supports these findings, since it revealed

the presence of the loci *ureE* and *ureC* for urease activity [23] and two copies of a putative nitrate reductase. It assimilates mannitol as an alternative carbon source to glucose. Six proteins encoding mannitol transporters and dehydrogenases were found. Similarly, ribose ABC transporter permeases have been found, which explains the assimilation of ribose. Interestingly, this *Anoxybacillus* species is found to weakly assimilate xylan, and its genome contains genes for 1,4-β-xylanase, a typical characteristic of the genus *Geobacillus* [24] rarely observed in *Anoxybacillus* (Filippidou et al., unpublished data). Finally, binding proteins and transporters of copper, manganese, cadmium, iron, and zinc were found, as well as an arsenic resistance protein (ArsB). Copper oxidase and manganese catalase genes are also present. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. JYCG00000000. The version described here is JYCG00000000.1.

7.2. Publically available, non-published genomes

In Table 7.1, we describe the strains that were isolated during our sampling campaign and were fully sequenced. In Table 7.2, a physiological characterization of those strains is described.

These isolates were selected based on their metabolic capabilities and their physiological characteristics. They can be categorized into thermophilic and mesophilic stains, alkalophilic, neutrophilic and acidophilic, halotolerant or not, and finally they can be separated into groups based on their capability of metal tolerance. Their genomes were screened for previously described biomarkers of "extremophily". The advantage compared to the publically available genomes is that all properties can be tested biochemically in our laboratory.

Isolate code	Identification	Genome Size (Mb)	Number of proteins	Sequencing platform and coverage
GSsed3	*Anoxybacillus geothermalis*	6.98	6,561	Illumina, 185X
B2M1	*Anoxybacillus geothermalis*	N.D.	N.D.	PacBio, 238X
B7M1	*Anoxybacillus geothermalis*	N.D.	N.D.	PacBio, 214X
Et10/1	*Bacillus thurigiensis*	9.84	N.D.	Illumina, 218X
GS3372	*Aeribacillus pallidus*	4.99	4,719	Illumina, 482X
24KAM51	*Bacillus alveayuensis*	6.7	6,327	Illumina, 409X
Et2/3	*Geobacillus kaustophilus*	3.51	3,393	Illumina, 480X
Et7/4	*Geobacillus kaustophilus*	3.68	3,368	Illumina, 464X
Lr3/2	*Bacillus thurigiensis*	5.57	5,640	Illumina, 317X
11kri323	*Bacillus mycoides*	6.5	6,119	Illumina, 358X
105NE	*Bacillus cereus*	5.55	5,444	Illumina, 438X

Lr7/2	*Bacillus thurigiensis*	5.61	5,478	Illumina, 349X
Lr 4/2	*Bacillus cereus*	6.23	6,139	Illumina, 305X
5NAP23	*Bacillus licheniformis*	4.17	4,064	Illumina, 341X

Table 7.1. Genomic information of the sequenced genomes.

Strain	Growth					Carbon sourse	
	Tmax	Tmin	pHmax	pHmin	Salinity	best	alternatives
GSSed3	65	40	10	5	3%	Glucose	ribose, glycerol,galactose xylose
GS3372	65	30	10	5	5%	Glucose	xylose ribose cellobiose
11kri323	55	4	11	5	5.50%	Glucose	maltose, malic acid, NAG, GNT, trisodium citrate
24kam51	80	30	10	3	13%	Glucose	arabinose mannose GNTmalic acid PAC
5nap23	55	between 10 and 4	11	4	19%	Glucose	manitol arabinose maltose NAG malic acid trisodium citrate
105NE	55	4	11	5	5%	Glucose	maltose, malic acid, N-acetyl-glucosamine
Et2/3	80	30	8	3	11%	Glucose	manitol arabinose maltose NAG malic acid trisodium citrate
Et7/4	70	25	8	3	11%	Glucose	manitol arabinose maltose NAG malic acid
Et10/1	55	25	10	4	17%	Glucose	manitol arabinose maltose NAG malic acid trisodium citrate
Lr3/2	55	4	10	5	8%	Glucose	maltose, malic acid, N-acetyl-glucosamine
Lr4/2	55	between 10 and 4	10	5	11%	Glucose	maltose, malic acid, NAG, GNT
Lr7/2	55	15	11	5	5.50%	Glucose	maltose, N-acetyl-glucosamine

Table 7.2 Physiological characterization of the sequenced strains.

7.4. References

[1] Bae SS, Lee J-H, Kim S-J. *Bacillus alveayuensis* sp. nov., a thermophilic bacterium isolated from deep-sea sediments of the Ayu Trough. Int J SystEvolMicrobiol 2005;55:1211–5. doi:10.1099/ijs.0.63424-0.

[2] Bennett S. Solexa Ltd. Pharmacogenomics 2004;5:433–8. doi:10.1517/14622416.5.4.433.

[3] Zerbino DR, Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 2008;18:821–9. doi:10.1101/gr.074492.107.

[4] Hemmerich C, Buechlein A, Podicheti R, Revanna KV, Dong Q. An Ergatis-based prokaryotic genome annotation web server. BioinformaOxfEngl 2010;26:1122–4. doi:10.1093/bioinformatics/btq090.

[5] Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. Bioinformatics 2005;21:3422–3. doi:10.1093/bioinformatics/bti553.

[6] Takami H, Takaki Y, Uchiyama I. Genome sequence of *Oceanobacillus iheyensis* isolated from the Iheya Ridge and its unexpected adaptive capabilities to extreme environments. Nucleic Acids Res 2002;30:3927–35.

[7] Minana-Galbis D, Pinzon DL, Gaspar Loren J, Manresa A, Oliart-Ros RM. Reclassification of *Geobacillus pallidus* (Scholz et al. 1988) Banat et al. 2004 as *Aeribacillus pallidus* gen. nov., comb. nov. Int J SystEvolMicrobiol 2010;60:1600–4. doi:10.1099/ijs.0.003699-0.

[8] Scholz T, Demharter W, Hensel R, Kandler O. *Bacillus pallidus* sp. nov, A New Thermophilic Species From Sewage. SystApplMicrobiol 1987;9:91–6.

[9] Banat IM, Marchant R, Rahman TJ. *Geobacillus debilis* sp. nov., a novel obligately thermophilic bacterium isolated from a cool soil environment, and reassignment of *Bacillus pallidus* to *Geobacillus pallidus* comb. nov.Int J SystEvolMicrobiol 2004;54:2197–201. doi:10.1099/ijs.0.63231-0.

[10] Yasawong M, Areekit S, Pakpitchareon A, Santiwatanakul S, Chansiri K. Characterization of thermophilic halotolerant *Aeribacillus pallidus* TD1 from Tao Dam Hot Spring, Thailand. Int J MolSci 2011;12:5294–303. doi:10.3390/ijms12085294.

[11] Pinzon-Martinez DL, Rodriguez-Gomez C, Minana-Galbis D, Carrillo-Chavez JA, Valerio-Alfaro G, Oliart-Ros R. Thermophilic bacteria from Mexican thermal environments: isolation and potential applications. Environ Technol 2010;31:957–66. doi:10.1080/09593331003758797.

[12] Chamkha M, Mnif S, Sayadi S. Isolation of a thermophilic and halophilic tyrosol-degrading *Geobacillus* from a Tunisian high-temperature oil field. FemsMicrobiol Lett 2008;283:23–9. doi:10.1111/j.1574-6968.2008.01136.x.

[13] Zheng C, He J, Wang Y, Wang M, Huang Z. Hydrocarbon degradation and bioemulsifier production by thermophilic *Geobacillus pallidus* strains. BioresourTechnol 2011;102:9155–61. doi:10.1016/j.biortech.2011.06.074.

[14] Periago PM, van Schaik W, Abee T, Wouters JA. Identification of proteins involved in the heat stress response of *Bacillus cereus* ATCC 14579. Appl Environ Microbiol 2002;68:3486–95.

[15] Derekova A, Sjoholm C, Mandeva R, Kambourova M. *Anoxybacillus rupiensis* sp. nov., a novel thermophilic bacterium isolated from Rupi basin (Bulgaria). Extremophiles 2007;11:577–83. doi:10.1007/s00792-007-0071-4.

[16] Rozanov AS, Bryanskaya AV, Kotenko AV, Malup TK, Peltek SE. Draft Genome Sequence of *Anoxybacillus flavithermus* Strain 25, Isolated from the Garga Hot Spring in the Barguzin Valley, Baikal Region, Russian Federation. Genome Announc 2014;2:e01258–14. doi:10.1128/genomeA.01258-14.

[17] Caspers MPM, Boekhorst J, Abee T, Siezen RJ, Kort R. Complete Genome Sequence of *Anoxybacillus flavithermus* TNO-09.006, a Thermophilic Sporeformer Associated with a Dairy-Processing Environment. Genome Announc 2013;1:e00010–3. doi:10.1128/genomeA.00010-13.

[18] Patel BKC. Draft Genome Sequence of *Anoxybacillus* Strain BCO1, Isolated from a Thermophilic Microbial Mat Colonizing the Outflow of a Bore Well of the Great Artesian Basin of Australia. Genome Announc 2015;3 :e01547–14. doi:10.1128/genomeA.01547-14.

[19] Tatusova T, Ciufo S, Fedorov B, O'Neill K, Tolstoy I. RefSeq microbial genomes database: new representation and annotation strategy. Nucleic Acids Res 2015;43:3872. doi:10.1093/nar/gkv278.

[20] Goh KM, Gan HM, Chan K-G, Chan GF, Shahar S, Chong CS, et al. Analysis of *Anoxybacillus* genomes from the aspects of lifestyle adaptations, prophage diversity, and carbohydrate metabolism. PloS One 2014;9:e90549. doi:10.1371/journal.pone.0090549.

[21] Kahar UM, Chan K-G, Salleh MM, Hii SM, Goh KM. A high molecular-mass *Anoxybacillus* sp. SK3-4 amylopullulanase: characterization and its relationship in carbohydrate utilization. Int J MolSci 2013;14:11302–18. doi:10.3390/ijms140611302.

[22] Matsutani M, Shirakihara Y, Imada K, Yakushi T, Matsushita K. Draft Genome Sequence of a Thermophilic Member of the Bacillaceae, *Anoxybacillus flavithermus* Strain Kn10, Isolated from the Kan-nawa Hot Spring in Japan. Genome Announc 2013;1. doi:10.1128/genomeA.00311-13.

[23] Kakinuma Y, Iida H, Sekizuka T, Taneike I, Takamiya S, Moore JE, et al. Molecular characterisation of urease genes from urease-positive thermophilic campylobacters (UPTC). Br J Biomed Sci 2008;65:148–52.

[24] Zeigler DR. The *Geobacillus* paradox: why is a thermophilic bacterial genus so prevalent on a mesophilic planet? Microbiology 2013;160:1–11. doi:10.1099/mic.0.071696-0.

Chapter 8

Manganese-II oxidation and copper-II resistance in endospore forming Firmicutes isolated from uncontaminated environmental sites

Ganesan Sathiyanarayanan[1&], Sevasti Filippidou[1&], Thomas Junier[1,2], Patricio Muñoz Rufatt[1], Nicole Jeanneret[1], Tina Wunderlin[1], Nathalie Sieber[1], Cristina Dorador[3], Pilar Junier[1]*

[1]Laboratory of Microbiology, Institute of Biology, University of Neuchâtel, Emile-Argand 11, CH-2000 Neuchâtel, Switzerland

[2]Vital-IT group, Swiss Institute of Bioinformatics, CH-1000 Lausanne, Switzerland

[3]Laboratorio de Complejidad Microbiana y Ecología Funcional; Departamento de Biotecnología; Facultad de Ciencias del Mar y Recursos Biológicos, Universidad de Antofagasta; CL-, 1270190, Antofagasta, Chile

[&]These authors contributed equally to this work.

This article was published at AIMS Environmental Sciences Journal in April 2016. Any reference to the data or the context of this article should cite:

Ganesan S, Filippidou S, Junier T, Muñoz Rufatt P, Jeanneret N, et al. Manganese-II oxidation and Copper-II resistance in endospore forming Firmicutes isolated from uncontaminated environmental sites. AIMS Environ Sci 2016; 3: 220–38. doi: 10.3934/environsci.2016.2.220.

Abstract

The accumulation of metals in natural environments is a growing concern of modern societies since they constitute persistent, non-degradable contaminants. Microorganisms are involved in redox processes and participate to the biogeochemical cycling of metals. Some endospore-forming Firmicutes (EFF) are known to oxidize and reduce specific metals and have been isolated from metal-contaminated sites. However, whether EFF isolated from uncontaminated sites have the same capabilities has not been thoroughly studied. In this study, we measured manganese oxidation and copper resistance of aerobic EFF from uncontaminated sites. For the purposes of this study we have sampled 22 natural habitats and isolated 109 EFF strains. Manganese oxidation and copper resistance were evaluated by growth tests as well as by molecular biology. Overall, manganese oxidation and tolerance to over 2 mM copper was widespread among the isolates (more than 44% of the isolates exhibited Mn(II)-oxidizing activity through visible Birnessite formation and 9.1% tolerate over 2 mM copper). The co-occurrence of these properties in the isolates was also studied. Manganese oxidation and tolerance to copper were not consistently found amongphylogenetically related isolates. Additional analysis correlating the physicochemical parameters measured on the sampling sites and the metabolic capabilities of the isolates showed a positive correlation between *in situ* alkaline conditions and the ability of the strains to perform manganese oxidation. Likewise, a negative correlation between temperature in the habitat and copper tolerance of the strains was observed. Our results lead to the conclusion that metal tolerance is a wide spread phenomenon in unrelated aerobic EFF from natural uncontaminated environments.

Keywords: Endospore-forming Firmicutes, Copper resistance, Manganese oxidation, metal contamination, natural habitats.

8.1. Introduction

The release of metals from various industries [1, 2], raises particular concern because they constitute persistent environmental contaminants that cannot be degraded or destroyed. At high concentrations, metals are toxic to living cells and accumulate throughout the food chain, leading to serious ecological and health issues [3]. Despite of their toxicity, living organisms require some specific metals (e.g.Ni, Cu, Co, Cr, Fe, Ca, Zn, K, Mn, and Mg) at low concentrations as micronutrients and they play a vital role in metalloenzymes as cofactors [4].

Environmental contamination by metals can have a significant impact on the indigenous microbial populations. Nearly every index of microbial metabolic activity (respiration, methanogenesis and nitrogen fixation, among others) is adversely affected by elevated concentrations of toxic metals [5]. Consequently, microorganisms that thrive in metal-contaminated environments have developed a variety of strategies for their survival including detoxifying mechanisms such as bioaccumulation, biotransformation, biomineralization or biosorption. Those mechanisms are of interest for *ex situ* or *in situ* bioremediation technologies [3, 6–9].

Manganese (II) oxidation is of particular relevance in environmental studies because Mn oxides are among the strongest naturally occurring oxidizing agents in the environment and can play a rolein numerous redox reactions controlling the distribution of other trace and contaminant elements [10].Despite the evidence indicating a large phylogenetic diversity among bacteria able to oxidize Mn (II), theevolutionary origin and function of this metabolic process are still under debate [10]. One hypothesis considers the possibility that by coating themselves with Mn oxides, Mn (II)-oxidizing bacteria create a self-made protection layer from environmental insults such as UV radiation, predation, viral attack or even metal toxicity [10]. The latter possibility would suggest that Mn (II)-oxidization could be a good predictor of the ability of an organism to resist elevated concentrations of toxic metals. However, this has not been tested experimentally.

An intriguing physiological link might exist between manganese oxidation and copper tolerance. One of the enzymes identified so far as having an integral role in Mn (II)-oxidation are multicopper oxidases (MCOs) [11, 12]. These enzymes appear also take part in a variety of cellular functions including copper homeostasis [13, 14]. Copper homeostasis is a complex process mediated by various genetic determinants [19], involving sequestration, uptake, and efflux [13]. Hence, Mn (II) oxidation and copper tolerance/resistance may be interrelated phenomena, but this has not been previously studied.

In terms of the diversity of microorganisms able to tolerate metals, tolerance has been mainly investigated in bacteria that are constantly under the selective pressure of high metal contents. This explains why, in environments enriched with specific metals, metal-resistant bacteria are predominant [15–18]. In these communities, endospore-forming Firmicutes (mainly Bacilli and Clostridia) are reported as a major portion of the total bacterial communities [15, 19, 20]. In the case of endospore-forming Firmicutes (EFF), their endospores are resistant to physical and chemical shock, enabling persistence under

stressing conditions for many years without losing germination potential [21]. In addition to their high prevalence in contaminated sites and their survival capability, EFF are known to participate in the biogeochemical cycling of metals, facilitating oxidation/reduction processes [11,22]. For instance, *Bacillus* sp. strain SG1 and *Desulfotomaculumreducens* strain MI-1 have been involved in manganese (II) oxidation [23] and uranium (VI) reduction [24], respectively. The widespread ability to oxidize manganese found among aerobic endospore-forming Firmicutes makes of this group an ideal candidate to study the hypothesis of a correlation between Mn (II)-oxidation and tolerance to toxic metals, and more specifically to copper.In this study, a series of aerobic endospore-forming Firmicutes were isolated from a diverse set of natural environments and their capability to oxidize Mn(II) was evaluated experimentally. In parallel, tolerance to increasing concentrations of copper was measured. The results were analyzed in order to establish a link between both metabolic process and the phylogenetic identity of the strains.

8.2. Materials and methods

8.2.1. Site description and sample collection

Various environmental samples such as soil, water, sediments and biofilms were collected from Chile (CL), Colombia (CO), Greece (GR), France (FR), Germany (DE) and Switzerland (CH) from March 2010 to April 2012. The samples were collected in sterile 50-ml Falcon tubes and transported to the laboratory on ice and stored at 4 ˚C for bacterial enrichment and isolation. The sampling sites corresponded to environmental niches that have not been contaminated by industrial deposition of metals. These environments and their geochemical characteristics are summarized in Supplementary Table 8.1.

8.2.2. Enrichment and isolation

For enrichment, one gram of collected sample was inoculated into 10 mL of Nutrient Broth (Biolife) and 10 mL of modified DifcoTM Marine Broth 2216 (D2216*****, 5 g of tryptone instead of peptone, no potassium bromide, pH adjusted 5.5, 6.8, 7.4 with 2N HCl). The inoculated media were incubated at specific temperatures corresponding to those of the sampling sites under aerobic conditions for 72 h. The enriched cultures were then plated on Nutrient Agar (NA) and modified Marine Agar (MA) (modified Marine Broth with 2% agar). Single colonies were obtained and each colony was plated repeatedly to attain pure aerobic bacterial isolates. Pure cultures were stored at 4 ˚C and cryopreserved in 25% (v/v) glycerol at -80 ˚C. Morphological and colony features were described according to the minimal standards for describing new taxa of aerobic endospore-forming bacteria [25]. Colony morphology was observed after 12 h of growth. Cell morphology, average cell size at 24 h in liquid medium and endospore formation were determined using phase-contrast microscopy (Leica DM R, magnification 1000x). EFF were selected among other isolates based on their capability to form spores after starvation for 15 days.

8.2.3. Identification of the isolated strains

Genomic DNA from each strain was extracted using the InnuPREP Bacteria DNA kit (Analytik Jena, Germany), according to the manufacturer's instructions. For amplification of 16S rRNA gene, the primers GM3F, GM4R and Eub9_27, Eub1542 were used as previously described [26, 27]. Presence of the gene that encodes the transcriptional factor responsible for the initiation of sporulation in endospore-forming bacteria (*spo0A* gene) was verified with the specific set of primers *spo0A*166f and *spo0A*748r, as described previously [28]. The PCR products were purified with a MultiScreen PCRμ96 Filter Plate (Millipore, USA), according to the manufacturer's instructions and sequenced using the services of Microsynth AG (Switzerland) and GATC Biotech (Germany). Nearly full-length 16S rRNA sequences were obtained by sequencing the PCR products in addition with the primers 907r, 926f primers and 518r [27,29]. The 16S rRNA gene were identified using the services of EzTaxon, against EzTaxon's cultured isolates database [30]. The sequences were submitted to GenBank under accession numbers KJ722422-KJ722533.

8.2.4. Manganese (II)-oxidation

Aerobic endospore-forming Firmicutes were screened for their capability to perform Mn (II) oxidation. Mn (II) oxidation was initially detected by the formation of brown Mn oxides on sporulated colonies after 10 days of incubation on solid K medium plates [22]. Mn (II) oxidation was further confirmed by Birnessite formation due to the oxidation of Mn (II) to crystalline Mn (IV) oxides in *Leptothrix* liquid medium [31]. Endospore-forming Firmicutes were also screened for the presence of a Mn oxidation gene that encodes a multicopper oxidase (MCOs) also responsible for copper resistance (*mnxG*), using primers adapted for *Bacillus* spp. [22]. The PCR products were purified on a MultiScreen PCRμ96 Filter Plate (Millipore, USA), according to the manufacturer's instructions and sequenced in order to confirm the identity of the PCR products as *mnxG* gene. Sequencing was performed using the services of GATC Biotech (Germany).

8.2.5. Copper (II) resistant/tolerance

The screening of Cu (II) resistant strains was carried out by using a modified form of Nutrient Agar and DifcoTM Marine Agar 2216* amended with 0.5 mM of $CuSO_4.5H_2O$ (II) (Sigma). The bacterial strains were inoculated onto agar plates and were incubated at each isolate's optimal temperature for 3-5 days. The level of Cu (II) resistance of the bacterial strains, that already grew at 0.5 mM Cu(II), was determined by plating the colonies on agar plates with increasing concentrations of $CuSO_4.5H_2O$ (II) from 0.75, 1.0, 1.5 up to 3.0 mM. Further, high Cu (II) tolerant strains were selected to test their Minimum Inhibitory Concentration (MIC). MIC was determined in nutrient or marine broth amended with 0.1, 0.25, 0.5, 0.75, 1.0, 1.5, 2.0 to 8 mM Cu (II). Strains were incubated at optimal temperatures for at least 1 week, verifying bacterial growth by optical density at 600 nm every 24 h. MIC was estimated as the first dilution which completely inhibits bacterial growth in nutrient or marine medium. The concentration of initial and final Cu (II) in the MIC cultures was quantitatively measured using a

spectrophotometric copper assay [32]. The final pH was also measured at the end of the incubation.

8.2.6. Statistical analysis

Statistical analyses were performed using Rstudio, version 0.98.1049. Correlations between diversity and environmental limiting factors were estimated using both Pearson's and Spearman's methods, however since our data are not normally distributed and also taken or transformed into ordinal scale, Spearman's tests were considered as more appropriate and therefore applied to this dataset.

8.2.7. Phylogenetic analysis of 16S rRNA, Mn (II) oxidation and Cu (II) resistance

Sequences of the 16S rRNA gene corresponding to the 109 strains analyzed, were aligned using MAFFT [33, 34]. A maximum likelihood phylogenetic tree was built using PhyML [35] with default parameters, and then an ornament tree was created using Newick utilities [36].

8.3. Results

8.3.1. Isolation and identification of endospore-forming Firmicutes

Aerobic endospore-forming Firmicutes of 22 non-polluted natural environmental sites from six countries (CL, DE, FR, CO, CH, GR) were characterized. A total of 338 strains representing distinct morphologies (including two phase variants) were selected from the different selective conditions and purified on NA and MA agar plates. All presumptive endospore-formers were selected based on their sporulating activity, the presence of *spo0A* gene and their 16S rRNA gene sequence. Those corresponded to 109endospore-forming strains (Table 8.1).

Sampling Site	Genera	Number of Isolates
Aggistro (GR)	*Bacillus*	1
	Lysinibacillus	1
AgiaParaskevi (GR)	*Exiguobacterium*	4
	Sporosarcina	1
AguasCalientes (CL)	*Bacillus*	1
Bruschal (DE)	*Anoxybacillus*	7
	Aeribacillus	1
	Brevibacillus	2
El Tatio (CL)	*Anoxybacillus*	2
	Bacillus	8
	Geobacillus	3
Eleftheres (GR)	*Bacillus*	1
Huasco (CL)	*Bacillus*	1
Krinides (GR)	*Bacillus*	2
	Kurthia	1
Lagadas (GR)	*Bacillus*	1
Las Piedras (CL)	*Bacillus*	4

Lirima (CL)	*Anoxybacillus*	8
	Bacillus	13
	Lysinibacillus	1
Milos (GR)	*Aeribacillus*	3
	Bacillus	9
	Geobacillus	1
Nea Apollonia (GR)	*Bacillus*	5
Neuchatel (CH)	*Bacillus*	2
	Exiguobacterium	2
	Lysinibacillus	1
Nigrita (GR)	*Aneurinibacillus*	2
	Bacillus	5
Pikrolimni (GR)	*Bacillus*	2
Potamia (GR)	*Anoxybacillus*	1
Pozar (GR)	*Bacillus*	2
	Lysinibacillus	1
Soulz-sous-forets (FR)	*Bacillus*	3
	Geobacillus	1
Thermia (GR)	*Anoxybacillus*	1
	Geobacillus	2
Traianoupoli (GR)	*Bacillus*	2
Yungai (CL)	*Bacillus*	2

Table 8.1. Description of genera isolated per sampling site and number of isolates per genus.

8.3.2. Mn (II)-oxidation

Two types of screening medium (K medium and Birnessite formation) were used to evaluate Mn(II)-oxidation in the 109 endospore-forming strains. Only 18.35% of strains were positive in the screening on K medium at 10 μM of Mn(II). In contrast, more than 44% of the isolates exhibited Mn(II)-oxidizing activity through visible Birnessite formation in *Leptothrix* medium at 1 mMMn(II).

In order to determine if MCOs are involved in Mn(II) oxidation in our collection of aerobic endospore-forming Firmicutes, the strains were screened for amplification of the copper-binding regions of *mnxG*, which are expected to be highly conserved due to their functional roles. A ~900bp product of *mnxG*was successfully amplified and verified by sequencing from 27.52% of the strains. Notably, only 11% of strains (12 strains) were positive for all three types of screening tests (Supplementary Table 8.2), those included mainly strains belonging to *Bacillus* (8 strains), although strains belonging to *Paenibacillus* (1 strain), *Geobacillus* (2 strains) and *Exiguobacterium* (1 strain) were also detected. The same percentage of strains was positive for the amplification of *mnxG* but did not oxidize Mn(II) under the conditions of the culture assays. Among this second group, three *Anoxybacillus*, a *Geobacillus*, an *Aneurinibacillus* and an *Exiguobacterium* were detected besides *Bacillus*.

8.3.3. Cu (II) resistance/tolerance

The isolated endospore-forming Bacilli were then tested for copper resistance with different concentrations of Cu (II). From the total collection, 27 strains were not able to grow in presence of copper, seven of them could tolerate up to 0.5 mM of Cu (II), 25 up to 0.75 mM, 36 up to 1mM, eight up to 1.5mM, two up to 2mM, one up to 2.5mM, and three up to 3mM. Ten strains, however, were able to grow in presence of Cu (II) in the solid medium up to 2 mM of Cu (II) (Supplementary Table 8.3 & 8.4) and four highly resistant strains tolerated up to 3 mM of copper. Minimum inhibitory concentration (MIC) of Cu (II) on these resistant isolates was estimated in liquid medium amended with Cu (II) concentrations from 0.1 to 8 mM. MIC tests were performed in liquid since on solid media there is a potential diffusion bias. Two of these strains (Et 9/2 and Lr 5/4) showed the highest MIC (5 mM) for Cu (II).

8.3.4. 16S rRNA phylogeny and co-existence of Cu (II) resistance and Mn (II)-oxidation

The phylogenetic analysis based on the 16S rRNA gene of the 109 isolates was related to the physicochemical parameters measured *in situ* and to the biochemical tests on manganese oxidation and copper tolerance, as well as to the presence of a manganese oxidation biomarker (*mnxG*) and is summarized in Figure 8.1. This analysis revealed some interesting observations about the ecology of aerobic endospore-forming Firmicutes. First, the phylogeny of these strains is not related to their geographical distribution. The same species can be found in sampling sites that are very distant. Their distribution, however, is sometimes related to environmental conditions and to the similarity between habitats. For example, *Anoxybacillus rupiensis* and *Bacillus smithii* are found to colonize a specific ecological niche, that of geothermal reservoirs. *Exiguobacteriummexicanum* isolates all originated from a sulfur geothermal spring. Other species can be found in different geographical locations but also in different environmental conditions, for example *Bacillus cereus* was found in 11 habitats.

A second observation made is that geographic location and manganese oxidation or copper resistance are not necessarily correlated, unless similar physicochemical parameters co-occur, for example isolates from Lirima, Chile, may or may not oxidize manganese depending on similarities in the temperature and pH measured at the sampling site. Likewise, these isolates exhibit different copper resistance maxima.

Among *Anoxybacillus* genus, absence of manganese oxidation was observed (with the exception of strain Lr10/3), however, some *Anoxybacillus* isolates had the *mnxG* gene. On the contrary, manganese oxidation was observed in all*Lysinibacillus* species. This observation was not the case for other genera, such as *Geobacillus* and *Bacillus*, for which manganese oxidation, as well as presence of *mnxG* gene, depend on pH and not on phylogeny.

Among *Exiguobacterium,* copper tolerance seems to be species specific, as *Exiguobacteriumaurantiacum* does not tolerate copper, while all strains that belong to *E. mexicanum* tolerate up to 1.5 mM of copper. Another species-secific observation is that

isolates that belong to *Bacillus niacini* show a high copper tolerance. Finally, *Bacillus amylolyquefaciens* strains tolerate up to 1.5 mM of copper and oxidize manganese, although *mnxG*could not be amplified.

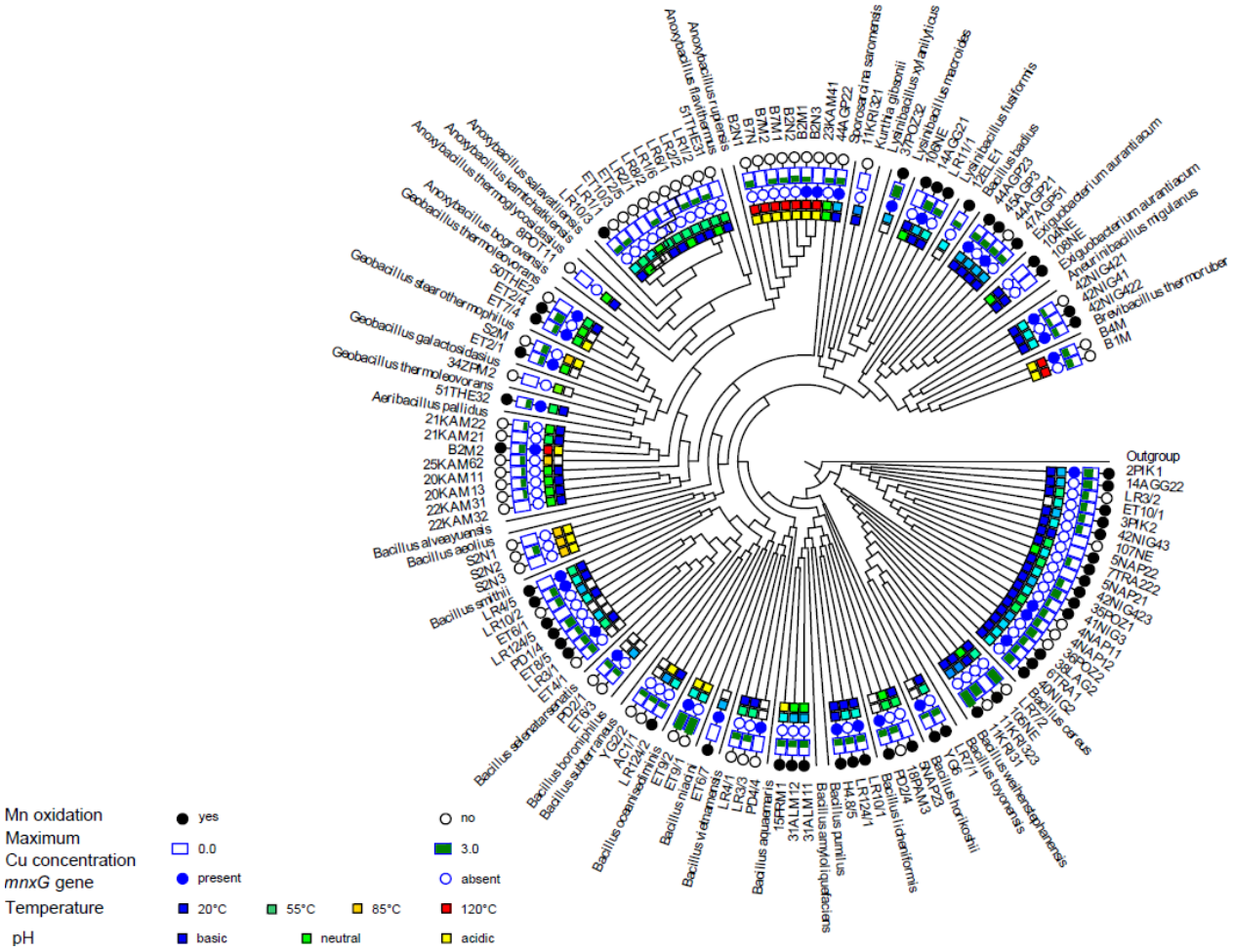


Figure 8.1. Maximum likelihood phylogenetic tree describing the phylogenetic relationship between isolates (code) and reference strains (species name indicated). On each branch of the tree, four parameters are indicated. Shape and color code for each parameter is described in the figure legend.

8.4. Discussion

In this study Mn(II) oxidation and copper tolerance were evaluated for a collection of aerobic endospore-forming Firmicutes isolated from a diverse set of natural uncontaminated sites environments. Despite the fact that research in metal tolerance is normally conducted in contaminated environments, uncontaminated sites might harbor a large diversity of microorganisms displaying different mechanisms of metal homeostasis.

As far as manganese oxidation is concerned, based on the results for Birnessite production, it can be inferred that uncontaminated environmental sites harbor an abundant population of Mn(II)-oxidizing aerobic endospore-forming Firmicutes. Mn(II)-oxidation is a metabolic trait found in a diverse set of unrelated microorganisms including representatives of the Alphaproteobacteria, Betaproteobacteria, Gammaproteobacteria, as well as high GC Gram-positive Actinobacteria and low GC Gram-positive Firmicutes [37]. Mn(II)-oxidizers have been assigned to various phylogenetic lineages within *Bacillus* demonstrating a large diversity of species bearing this trait even within this confined

genus [11,22]. This was also the case for our large collection of unrelated *Bacillus* strains. Moreover,our results show that Mn(II)-oxidation is a trait widely spread in other aerobic endospore-forming Firmicutes genera, since we have observed Mn(II)-oxidation in strains affiliated to other genera such as *Aeribacillus, Geobacillus, Anoxybacillus, Lysinibacillus, Exiguobacterium*and *Brevibacillus*.Since, Mn(II)-oxidation creates biogenic Mn oxides that have a high absorption capacity for metal cations and an ability to oxidize numerous inorganic and organic compounds, understanding the mechanisms of microbial mediated Mn(II)-oxidation process study here could be significant in both biogeochemical and biotechnological contexts [38,39].For many of the strains, amplification of a common genetic marker ofmulticopper oxidases (MCOs), which are so far identified as a key element of Mn(II)-oxidation (Dick et al. 2008; van Waasbergen *et al.* 1996) did not yield results, and thus further studies into the existence of alternative MCOs or other mechanisms of oxidation for the studied strains are still required.

In the case of resistance/tolerance of microorganisms to copper, overall, resistance decreased with increasing Cu (II) concentrations. Our results are consistent with a previous study that indicates that the number of resistant bacteria decreased at higher Cu (II) concentration [15].

Manganese oxidation and copper tolerance are specific in some genera, such as *Anoxybacillus* and *Lysinibacillus*, but this is not the case for all aerobic endospore-forming Firmicutes tested herein. This observation is also true for some species; however, others, like *Bacillus cereus,* do not exhibit a specific pattern concerning either manganese oxidation and copper resistance or geographical distribution and niche specialization. It is worth mentioning that a strain related to *Kurthiagibsonii* strain 11Kri-321 from the collection was positive for the screening with the two culture media but did not produce a PCR product for the *mnxG* gene with the primers evaluated here.

Our findings show that manganese oxidation in nature is not specific to bacterial species or site location. This is equally true for copper tolerance. A positive correlation between pH measured *in situ* and manganese oxidation was observed (R^2= 0.351, p-value= $9.13e^{-05}$) as well as a negative correlation between temperature and Cu concentration (R^2= 0.218, p-value= 0.027). These findings do not contradict the previously well-demonstrated relationship between biosorption and physicochemical parameters [40–42]. The results also suggests that endospore-formers have natural resistance mechanisms for toxic metals and metal resistance is a wide spread phenomenon in endospore-forming Firmicutes. The potential for 182ioaccumulation/biosorption of toxic metals already suggested in the literature [43], opens up the possibility for use of thesespore-formers in biological treatment processes, applied to effluents or sites contaminated with a wide range of toxic metals.

8.5. Conclusions

Aerobic endospore-forming Firmicutes isolated from natural, uncontaminated environments were found to oxidize manganese and resist up to 3mM of copper. These capabilities were related to physicochemical parameters, however, some genera and

some species showed concrete patterns of manganese oxidation and copper resistance. Not all strains able to oxidize manganesewere copper tolerant, and vice-versa, and thus theexistence of one capability cannot be assumed to be a prerequisite for the presence of the other. However, our initial hypothesis that manganese oxidation and copper tolerance are widespread phenomena among aerobic endospore-forming bacilli is well confirmed.

8.6. Acknowledgments

We acknowledge funding from the Swiss National Funding projects 31003A_132358/1 and 31003A_152972, from Fondation Pierre Mercier pour la science and from REGARD for equality of women in science. We would like to thank the Swiss Confederation Fellowship Program for Foreign Researchers for funding S. Ganesan and S. Filippidou.

The genetic information downloaded from Genebank may be considered to be part of the genetic patrimony of Switzerland, Germany, France, Greece, Colombia and Chile respectively, the countries from which the samples were obtained. Users of this information agree to: 1) acknowledge Switzerland, Germany, France, Greece, Colombia and Chile as the countries of origin in any country where the genetic information is presented and 2) contact the CBD focal point and the ABS focal point identified in the CBD website http://www.cbd.int/information/nfp.shtml if they intend to use the genetic information for commercial purposes.

8.7. References

[1] Fu F, Wang Q. Removal of heavy metal ions from wastewaters: A review. J Environ Manage 2011;92:407–18. doi:10.1016/j.jenvman.2010.11.011.

[2] Simsek I, Karatas M, Basturk E. Cu(II) removal from aqueous solution by ureolytic mixed culture (UMC). Colloids Surf B Biointerfaces 2013;102:479–83. doi:10.1016/j.colsurfb.2012.09.027.

[3] Zouboulis A., Loukidou M., Matis K. Biosorption of toxic metals from aqueous solutions by bacteria strains isolated from metal-polluted soils. Process Biochem 2004;39:909–16. doi:10.1016/S0032-9592(03)00200-0.

[4] Berg JM, Tymoczko JL, Stryer L, Berg JM, Tymoczko JL, Stryer L. Biochemistry. 5th ed. W H Freeman; 2002.

[5] Gadd GM, Griffiths AJ. Microorganisms and heavy metal toxicity. Microb Ecol 1977;4:303–17. doi:10.1007/BF02013274.

[6] Ghosh A, Saha PD. Optimization of copper bioremediation by *Stenotrophomonas maltophilia* PD2. J Environ Chem Eng 2013;1:159–63. doi:10.1016/j.jece.2013.04.012.

[7] Pradhan AA, Levine AD. Microbial biosorption of copper and lead from aqueous systems. Sci Total Environ 1995;170:209–20. doi:10.1016/0048-9697(95)04709-7.

[8] Wang J, Chen C. Biosorbents for heavy metals removal and their future. Biotechnol Adv 2009;27:195–226. doi:10.1016/j.biotechadv.2008.11.002.

[9] Zhou W, Zhang H, Ma Y, Zhou J, Zhang Y. Bio-removal of cadmium by growing deep-sea bacterium *Pseudoalteromonas* sp. SCSE709-6. Extremophiles 2013;17:723–31. doi:10.1007/s00792-013-0554-4.

[10] Tebo BM, Johnson HA, McCarthy JK, Templeton AS. Geomicrobiology of manganese(II) oxidation. Trends Microbiol 2005;13:421–8. doi:10.1016/j.tim.2005.07.009.

[11] Dick GJ, Torpey JW, Beveridge TJ, Tebo BM. Direct Identification of a Bacterial Manganese(II) Oxidase, the Multicopper Oxidase MnxG, from Spores of Several Different Marine *Bacillus* Species. Appl Environ Microbiol 2008;74:1527–34. doi:10.1128/AEM.01240-07.

[12] Waasbergen LG van, Hildebrand M, Tebo BM. Identification and characterization of a gene cluster involved in manganese oxidation by spores of the marine *Bacillus* sp. strain SG-1. J Bacteriol 1996;178:3517–30.

[13] Rensing C, Grass G. Escherichia coli mechanisms of copper homeostasis in a changing environment. FEMS Microbiol Rev 2003;27:197–213.

[14] Taylor AB, Stoj CS, Ziegler L, Kosman DJ, Hart PJ. The copper-iron connection in biology: Structure of the metallo-oxidase Fet3p. Proc Natl Acad Sci 2005;102:15459–64. doi:10.1073/pnas.0506227102.

[15] Altimira F, Yáñez C, Bravo G, González M, Rojas LA, Seeger M. Characterization of copper-resistant bacteria and bacterial communities from copper-polluted agricultural soils of central Chile. BMC Microbiol 2012;12:193. doi:10.1186/1471-2180-12-193.

[16] Andreazza R, Pieniz S, Wolf L, Lee M-K, Camargo FAO, Okeke BC. Characterization of copper bioreduction and biosorption by a highly copper resistant bacterium isolated

from copper-contaminated vineyard soil. Sci Total Environ 2010;408:1501–7. doi:10.1016/j.scitotenv.2009.12.017.

[17] Chen G, Chen X, Yang Y, Hay AG, Yu X, Chen Y. Sorption and Distribution of Copper in Unsaturated *Pseudomonas putida* CZ1 Biofilms as Determined by X-Ray Fluorescence Microscopy. Appl Environ Microbiol 2011;77:4719–27. doi:10.1128/AEM.00125-11.

[18] Fan L-M, Ma Z-Q, Liang J-Q, Li H-F, Wang E-T, Wei G-H. Characterization of a copper-resistant symbiotic bacterium isolated from Medicago lupulina growing in mine tailings. Bioresour Technol 2011;102:703–9. doi:10.1016/j.biortech.2010.08.046.

[19] Besaury L, Bodilis J, Delgas F, Andrade S, De la Iglesia R, Ouddane B, et al. Abundance and diversity of copper resistance genes cusA and copA in microbial communities in relation to the impact of copper on Chilean marine sediments. Mar Pollut Bull 2013;67:16–25. doi:10.1016/j.marpolbul.2012.12.007.

[20] Bueche M, Wunderlin T, Roussel-Delif L, Junier T, Sauvain L, Jeanneret N, et al. Quantification of endospore-forming firmicutes by quantitative PCR with the functional gene *spo0A*. Appl Environ Microbiol 2013;79:5302–12. doi:10.1128/AEM.01376-13.

[21] Vreeland RH, Rosenzweig WD, Powers DW. Isolation of a 250 million-year-old halotolerant bacterium from a primary salt crystal. Nature 2000;407:897–900. doi:10.1038/35038060.

[22] Francis CA, Tebo BM. Enzymatic Manganese(II) oxidation by metabolically dormant spores of diverse *Bacillus* species. Appl Environ Microbiol 2002;68:874–80. doi:10.1128/AEM.68.2.874-880.2002.

[23] Nealson KH, Myers CR. Microbial reduction of manganese and iron: new approaches to carbon cycling. Appl Environ Microbiol 1992;58:439–43.

[24] Junier P, Frutschi M, Wigginton NS, Schofield EJ, Bargar JR, Bernier-Latmani R. Metal reduction by spores of *Desulfotomaculum reducens*. Environ Microbiol 2009;11:3007–17. doi:10.1111/j.1462-2920.2009.02003.x.

[25] Logan NA, Berge O, Bishop AH, Busse H-J, De Vos P, Fritze D, et al. Proposed minimal standards for describing new taxa of aerobic, endospore-forming bacteria. Int J Syst Evol Microbiol 2009;59:2114–21. doi:10.1099/ijs.0.013649-0.

[26] Liesack W, Weyland H, Stackebrandt E. Potential Risks of Gene Amplification by PCR as Determined by 16s rDNA Analysis of a Mixed-Culture of Strict Barophilic Bacteria. Microb Ecol 1991;21:191–8. doi:10.1007/BF02539153.

[27] Muyzer G, Teske A, Wirsen C, Jannasch H. Phylogenetic-Relationships of *Thiomicrospira* Species and their Identification in Deep-Sea Hydrothermal Vent Samples by Denaturing Gradient Gel-Electrophoresis of 16S rDNA Fragments. Arch Microbiol 1995;164:165–72. doi:10.1007/BF02529967.

[28] Wunderlin T, Junier T, Roussel-Delif L, Jeanneret N, Junier P. Stage 0 sporulation gene A as a molecular marker to study diversity of endospore-forming Firmicutes: Culture-independent study of endospore-forming Firmicutes. Environ Microbiol Rep 2013;5:911–24. doi:10.1111/1758-2229.12094.

[29] Ovreås L, Forney L, Daae FL, Torsvik V. Distribution of bacterioplankton in meromictic Lake Saelenvannet, as determined by denaturing gradient gel electrophoresis

of PCR-amplified gene fragments coding for 16S rRNA. Appl Environ Microbiol 1997;63:3367–73.
[30] Kim O-S, Cho Y-J, Lee K, Yoon S-H, Kim M, Na H, et al. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. Int J Syst Evol Microbiol 2012;62:716–21. doi:10.1099/ijs.0.038075-0.
[31] Pena J, Kwon KD, Refson K, Bargar JR, Sposito G. Mechanisms of nickel sorption by a bacteriogenic birnessite. Geochim Cosmochim Acta 2010;74:3076–89. doi:10.1016/j.gca.2010.02.035.
[32] Ahmed N, Nawaz A, Badar U. Screening of Copper Tolerant Bacterial Strains and Their Potential to Remove Copper from the Environment. Bull Environ Contam Toxicol 2005;74:219–26. doi:10.1007/s00128-004-0573-z.
[33] Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 2002;30:3059–66. doi:10.1093/nar/gkf436.
[34] Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol 2013;30:772–80. doi:10.1093/molbev/mst010.
[35] Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol 2010;59:307–21. doi:10.1093/sysbio/syq010.
[36] Junier T, Zdobnov EM. The Newick Utilities: High-throughput Phylogenetic tree Processing in the Unix Shell. Bioinformatics 2010:btq243. doi:10.1093/bioinformatics/btq243.
[37] Tebo BM, Bargar JR, Clement BG, Dick GJ, Murray KJ, Parker D, et al. Biogenic manganese oxides: Properties and Mechanisms of Formation. Annu Rev Earth Planet Sci 2004;32:287–328. doi:10.1146/annurev.earth.32.101802.120213.
[38] Geszvain K, Butterfield C, Davis RE, Madison AS, Lee S-W, Parker DL, et al. The molecular biogeochemistry of manganese(II) oxidation. Biochem Soc Trans 2012;40:1244–8. doi:10.1042/BST20120229.
[39] Soldatova AV, Butterfield C, Oyerinde OF, Tebo BM, Spiro TG. Multicopper oxidase involvement in both Mn(II) and Mn(III) oxidation during bacterial formation of MnO2. J Biol Inorg Chem 2012;17:1151–8. doi:10.1007/s00775-012-0928-6.
[40] V.Sridevi MS. A Comparative Study on Determination of PhysicoChemical parameters of Biosorption of Lead (II) by *Aspergillus niger* , NCIM 616 using AAS & ICPMS. IOSR J Pharm IOSRPHR 2012;2:60–4. doi:10.9790/3013-26506064.
[41] Babarinde A, Babalola JO, Adegoke J, Osundeko AO, Olasehinde S, Omodehin A, et al. Biosorption of Ni(II), Cr(III), and Co(II) from Solutions Using *Acalypha hispida* Leaf: Kinetics, Equilibrium, and Thermodynamics. J Chem 2012;2013:e460635. doi:10.1155/2013/460635.
[42] Komy ZR, Abdelraheem WH, Ismail NM. Biosorption of Cu2+ by Eichhornia crassipes: Physicochemical characterization, biosorption modeling and mechanism. J King Saud Univ - Sci 2013;25:47–56. doi:10.1016/j.jksus.2012.04.002.

[43] Philip L, Iyengar L, Venkobachar C. Site of Interaction of Copper on Bacillus Polymyxa. Water Air Soil Pollut 2000;119:11–21. doi:10.1023/A:1005123624184.

8.8. Supplementary Material

Supplementary Table 8.1. Sampling sites and their description

Sampling site	Country	Site description	Temperature (°C)	pH
Neuchatel	Switzerland (CH)	Mineral iron and sulfur springs	15	6.94-7.13
Soultz-sous-forets	France (FR)	Geothermal power plant	100	5.2
Bruschal	Germany (DE)	Geothermal power plant	122	5.4
Salar de Yungai	Chile (CL)	Salt crusts	n/a	n/a
AguasCalientes	Chile (CL)	Geothermal natural spring	32	6
Laguna de las Piedras	Chile (CL)	Atakama Lake	n/a	n/a
El Tatio	Chile (CL)	Geysers	34-70	5.5
Salar de Huasco	Chile (CL)	Lagoon of thermal or cold inlets	15.8	8.6
Lirima	Chile (CL)	Thermal site with springs and streams	45-55	7.48-8.04
Lagadas	Greece (GR)	Geothermal natural spring	38	8
Pikrolimni	Greece (GR)	Lake	35	9.21-9.86
NeaAppolonia	Greece (GR)	Geothermal natural spring	59	8.2-8.8
Milos	Greece (GR)	Volcanic island	35-100	6.13-7.5
AgiaParaskevi	Greece (GR)	Mineral sulfur spring	35	7.6
Thermia	Greece (GR)	Geothermal natural spring	60	7.6
Nigrita	Greece (GR)	Geothermal natural spring	43	8.2
Pozar	Greece (GR)	Geothermal natural spring	35	7.7-8.3
Aggistro	Greece (GR)	Geothermal natural spring	37	8.1
Eleftheres	Greece (GR)	Geothermal natural spring	41	n/a
Krinides	Greece (GR)	Mineral spring	30	7.99
Potamia	Greece (GR)	Geothermal natural spring	70	8.55
Traianoupoli	Greece (GR)	Geothermal natural spring	41	7.34-7.56

Supplementary Table 8.2. Strains isolated per sampling site (Origin) and identified against EzTaxon's cultured database (Closest relative). Manganese oxidation capability on two different media, presence of mnGx gene and copper tolerance at different concentrations is also presented. Co-presence of manganese oxidation capability and copper tolerance is shown in column Mn-Cu.

Nr	Strains	Origin	Closest relative	Manganese (II) Oxidation			Copper (II) Tolerance							Mn -Cu
				K medium	Birnessite	MnOx gene	Copper supplemented medium (Cu II concentration in mM)							
							0.5 mM	0.75 mM	1.0 mM	1.5 mM	2.0 mM	2.5 mM	3.0 mM	
1	107NE	Neuchatel	*Bacillus cereus*	-	-	-	-	-	-	-	-	-	-	-
2	21KAM21	Kalamos-Milos	*Aeribacillus pallidus*	-	-	-	-	-	-	-	-	-	-	-
3	22KAM32	Kalamos-Milos	*Bacillus alveayuensis*	-	-	-	-	-	-	-	-	-	-	-
4	34ZPM2	Zefiria Milos	*Geobacillusgalactosidasius*	-	-	-	-	-	-	-	-	-	-	-
5	51THE31	Thermia	*Anoxybacillusflavithermus*	-	-	-	-	-	-	-	-	-	-	-
6	8POT11	Potarmia	*Anoxybacillusflavithermus*	-	-	-	-	-	-	-	-	-	-	-
7	B2N3	Bruschal	*Anoxybacillusrupiensis*	-	-	-	-	-	-	-	-	-	-	-
8	B7N	Bruschal	*Anoxybacillusrupiensis*	-	-	-	-	-	-	-	-	-	-	-
9	Et10/3	El Tatio	*Anoxybacillussalavatliensis*	-	-	-	-	-	-	-	-	-	-	-
10	Et2/5	El Tatio	*Anoxybacillusbogrovensis*	-	-	-	-	-	-	-	-	-	-	-
11	Lr1/6	Lirima	*Anoxybacillussalavatliensis*	-	-	-	-	-	-	-	-	-	-	-
12	Lr4/5	Lirima	*Bacillus boroniphilus*	-	-	-	-	-	-	-	-	-	-	-
13	Lr8/2	Lirima	*Anoxybacillussalavatliensis*	-	-	-	-	-	-	-	-	-	-	-
14	S2N1	Soultz	*Bacillus smithi*	-	-	-	-	-	-	-	-	-	-	-
15	S2N3	Soultz	*Bacillus smithi*	-	-	-	-	-	-	-	-	-	-	-
16	104NE	Neuchatel	*Exiguobacteriumaurantiacum*	+	-	-	-	-	-	-	-	-	-	-
17	105NE	Neuchatel	*Bacillus thuringiensis*	-	+	-	-	-	-	-	-	-	-	-
18	106NE	Neuchatel	*Lysinibacillusmacroides*	-	+	-	-	-	-	-	-	-	-	-
19	20KAM11	Kalamos-Milos	*Bacillus alveayuensis*	-	-	-	+	-	-	-	-	-	-	-
20	21KAM22	Kalamos-Milos	*Aeribacillus pallidus*	-	-	-	+	-	-	-	-	-	-	-
21	22KAM31	Kalamos-Milos	*Bacillus alveayuensis*	-	-	-	+	-	-	-	-	-	-	-
22	5NAP21	Nea Apollonia	*Bacillus cereus*	-	+	-	-	-	-	-	-	-	-	-

23	5NAP22	Nea Apollonia	*Bacillus cereus*	-	+	-	-	-	-	-	-	-	-	-
24	5NAP23	Nea Apollonia	*Bacillus licheniformis*	-	-	+	-	-	-	-	-	-	-	-
25	Lr124/5	Lirima	*Bacillus jeotgali*	+	-	-	-	-	-	-	-	-	-	-
26	Pd4/4	Las Piedras	*Bacillus aquimaris*	-	-	+	-	-	-	-	-	-	-	-
27	108NE	Neuchatel	*Exiguobacteriummexicanum*	+	+	-	-	-	-	-	-	-	-	-
28	11KRI321	Krinides	*Kurthiagibsonii*	+	+	-	-	-	-	-	-	-	-	-
29	20KAM13	Kalamos-Milos	*Bacillus alveayuensis*	-	-	-	+	+	-	-	-	-	-	-
30	23KAM41	Kalamos-Milos	*Bacillus aeolius*	-	-	-	+	+	-	-	-	-	-	-
31	25KAM62	Kalamos-Milos	*Aeribacillus pallidus*	-	-	-	+	+	-	-	-	-	-	-
32	41NIG3	Nigrita	*Bacillus cereus*	-	-	+	+	-	-	-	-	-	-	-
33	42NIG41	Nigrita	*Aneurinibacillusmigulanus*	-	+	-	+	-	-	-	-	-	-	+
34	42NIG421	Nigrita	*Aneurinibacillusmigulanus*	-	+	-	+	-	-	-	-	-	-	+
35	50THE2	Thermia	*Geobacillusthermoleovorans*	-	+	+	-	-	-	-	-	-	-	-
36	Ac1/1	AguasCalientes	*Bacillus oceanisediminis*	-	-	-	+	+	-	-	-	-	-	-
37	B7M1	Bruschal	*Anoxybacillusrupiensis*	-	-	-	+	+	-	-	-	-	-	-
38	B7M2	Bruschal	*Anoxybacillusrupiensis*	-	-	-	+	+	-	-	-	-	-	-
39	Et6/3	El Tatio	*Bacillus boroniphilus*	-	-	-	+	+	-	-	-	-	-	-
40	Lr1/1	Lirima	*Anoxybacillussalavatliensis*	-	-	-	+	+	-	-	-	-	-	-
41	Lr1/2	Lirima	*Anoxybacillusthermarum*	-	-	-	+	+	-	-	-	-	-	-
42	Lr2/1	Lirima	*Anoxybacilluskamchatkensis*	-	-	-	+	+	-	-	-	-	-	-
43	Lr2/2	Lirima	*Anoxybacillusthermarum*	-	-	-	+	+	-	-	-	-	-	-
44	Lr6/1	Lirima	*Anoxybacillusbogrovensis*	-	-	-	+	+	-	-	-	-	-	-
45	Yg2/2	Yungai	*Bacillus oceanisediminis*	-	-	-	+	+	-	-	-	-	-	-
46	12ELE1	Eleftheres	*Bacillus badius*	-	+	-	+	+	-	-	-	-	-	+
47	18PAM3	Palaiochori	*Bacillus sp.*	-	-	-	+	+	+	-	-	-	-	-
48	51THE32	Thermia	*Geobacillusthermoglucosidasius*	-	-	+	+	+	-	-	-	-	-	-
49	7TRA222	Traianoupoli	*Bacillus cereus*	-	+	-	+	+	-	-	-	-	-	+
50	B1M	Bruschal	*Brevibacillusthermoruber*	-	-	-	+	+	+	-	-	-	-	-
51	B2M1	Bruschal	*Anoxybacillusrupiensis*	-	-	+	+	+	-	-	-	-	-	-

52	B4M	Bruschal	*Brevibacillusthermoruber*	-	-	-	+	+	+	-	-	-	-	-
53	Et4/1	El Tatio	*Bacillus thioparans*	-	-	-	+	+	+	-	-	-	-	-
54	Et6/1	El Tatio	*Bacillus thioparans*	-	-	-	+	+	+	-	-	-	-	-
55	Et6/7	El Tatio	*Bacillus vietnamensis*	+	+	+	-	-	-	-	-	-	-	-
56	Et8/5	El Tatio	*Bacillus jeotgali*	+	-	-	+	+	-	-	-	-	-	+
57	Lr124/2	Lirima	*Bacillus oceanisediminis*	-	+	-	+	+	-	-	-	-	-	-
58	Lr3/1	Lirima	*Bacillus thioparans*	-	-	-	+	+	+	-	-	-	-	-
59	Lr3/2	Lirima	*Bacillus cereus*	-	-	-	+	+	+	-	-	-	-	-
60	Pd2/1	Las Piedras	*Bacillus boroniphilus*	-	-	+	+	+	-	-	-	-	-	-
61	S2M	Soultz	*Geobacillusstearothermophilus*	-	-	-	+	+	+	-	-	-	-	-
62	S2N2	Soultz	*Bacillus smithi*	-	-	-	+	+	+	-	-	-	-	-
63	15PRM1	Provtas-Milos	*Bacillus subtilis*	-	+	-	+	+	+	-	-	-	-	+
64	31ALM11	Milos Liogerma	*Bacillus amylolyquefaciens*	-	+	-	+	+	+	-	-	-	-	+
65	31ALM12	Milos Liogerma	*Bacillus amylolyquefaciens*	-	+	-	+	+	+	-	-	-	-	+
66	38LAG2	Lagadas	*Bacillus cereus*	-	+	-	+	+	+	-	-	-	-	+
67	3PIK2	Pikrolimni	*Bacillus cereus*	-	+	-	+	+	+	-	-	-	-	+
68	42nig43	Nigrita	*Bacillus cereus*	-	+	-	+	+	+	-	-	-	-	+
69	44AGP21	AgiaParaskevi	*Exiguobacteriummexicanum*	-	-	+	+	+	+	-	-	-	-	-
70	44AGP22	AgiaParaskevi	*Sporosarcinasaromensis*	-	+	-	+	+	+	-	-	-	-	+
71	47AGP51	AgiaParaskevi	*Exiguobacteriumaurantiacum*	-	-	+	+	+	+	-	-	-	-	-
72	4NAP11	Nea Apollonia	*Bacillus cereus*	-	+	-	+	+	+	-	-	-	-	+
73	4NAP12	Nea Apollonia	*Bacillus thuringiensis*	-	-	+	+	+	+	-	-	-	-	-
74	B2M2	Bruschal	*Aeribacillus pallidus*	-	-	+	+	+	+	-	-	-	-	-
75	B2N1	Bruschal	*Anoxybacillusrupiensis*	-	-	+	+	+	+	-	-	-	-	-
76	Et10/1	El Tatio	*Bacillus thuringiensis*	-	+	-	+	+	+	-	-	-	-	+
77	Lr10/1	Lirima	*Bacillus licheniformis*	-	+	-	+	+	+	-	-	-	-	+
78	Lr10/3	Lirima	*Anoxybacillusthermarum*	+	-	+	+	+	-	-	-	-	-	+
79	Lr11/1	Lirima	*Lysinibacillus fusiformis*	-	+	-	+	+	+	-	-	-	-	+
80	Lr124/1	Lirima	*Bacillus licheniformis*	-	+	-	+	+	+	-	-	-	-	+

81	Lr7/1	Lirima	*Bacillus cereus*	-	+	-	+	+	+	-	-	-	-	+
82	Yg6	Yungai	*Bacillus horikoshii*	+	+	+	+	-	-	-	-	-	-	+
83	11KRI323	Krinides	*Bacillus toyonensis*	-	-	-	+	+	+	+	+	-	-	-
84	14AGG22	Aggistro	*Bacillus cereus*	-	+	-	+	+	+	+	-	-	-	+
85	35POZ1	Pozar	*Bacillus cereus*	-	+	-	+	+	+	+	-	-	-	+
86	37POZ32	Pozar	*Lysinibacillusxylanilyticus*	-	-	-	+	+	+	+	+	-	-	-
87	40NIG2	Nigrita	*Bacillus thuringiensis*	-	+	-	+	+	+	+	-	-	-	+
88	42NIG422	Nigrita	*Bacillus sp.*	+	+	+	+	+	-	-	-	-	-	+
89	42NIG423	Nigrita	*Bacillus cereus*	+	+	-	+	+	+	-	-	-	-	+
90	44AGP23	AgiaParaskevi	*Exiguobacterium sp.*	+	+	-	+	+	+	-	-	-	-	+
91	6TRA1	Traianoupoli	*Bacillus sp.*	+	+	+	+	+	-	-	-	-	-	+
92	B2N2	Bruschal	*Anoxybacillusrupiensis*	-	+	+	+	+	+	-	-	-	-	+
93	Et2/1	El Tatio	*Geobacillus sp.*	+	+	+	+	+	-	-	-	-	-	+
94	Et2/4	El Tatio	*Geobacillusthermoparaffinivorans*	-	+	-	+	+	+	+	-	-	-	+
95	Et7/4	El Tatio	*Geobacillusstearothermophilus*	+	+	+	+	+	-	-	-	-	-	+
96	Lr10/2	Lirima	*Bacillus subterraneus*	-	+	+	+	+	+	-	-	-	-	+
97	Pd1/4	Las Piedras	*Bacillus selenatarsenatis*	-	+	-	+	+	+	+	-	-	-	+
98	14AGG21	Aggistro	*Lysinibacillus sp.*	-	+	+	+	+	+	+	-	-	-	+
99	2PIK1	Pikrolimni	*Bacillus cereus*	-	+	+	+	+	+	+	-	-	-	+
100	45AGP3	AgiaParaskevi	*Exiguobacterium sp.*	+	+	+	+	+	+	-	-	-	-	+
101	H4.8/5	Huasco	*Bacillus pumilus*	+	+	+	+	+	+	-	-	-	-	+
102	Lr3/3	Lirima	*Bacillus aquimaris*	+	+	+	+	+	+	-	-	-	-	+
103	Lr4/1	Lirima	*Bacillus aquimaris*	+	+	+	+	+	+	-	-	-	-	+
104	Lr7/2	Lirima	*Bacillus thuringiensis*	-	-	-	+	+	+	+	+	+	-	-
105	Pd2/4	Las Piedras	*Bacillus licheniformis*	+	+	+	+	+	+	-	-	-	-	+
106	11KRI31	Krinides	*Bacillus weihenstephanensis*	-	-	-	+	+	+	+	+	+	+	-
107	36POZ2	Pozar	*Bacillus cereus*	+	+	+	+	+	+	+	-	-	-	+
108	Et9/1	El Tatio	*Bacillus niacini*	-	-	-	+	+	+	+	+	+	+	-
109	Et9/2	El Tatio	*Bacillus niacini*	-	-	+	+	+	+	+	+	+	+	+

Supplementary Table 8.3. Cu (II) biosorption and pH change during the growth of high Cu (II) tolerant endospore-forming Firmicutes. MIC= Minimum inhibitory concentration.

Cu II mM	Initial Cu II (mg/L)	Initial pH	Cu II tolerant strains & their pH change, final Cu II; Control pH 7.2±0.2 (0.0 mM Cu II)																			
			Et 9/1		Et 9/2		Lr 5/4		Lr 7/2		9kri1		10kri2		11kri31		11kri323		11kri324		37poz32	
			Final Cu II	Final pH	Final Cu II	Final pH	Final Cu II	Final pH	Final Cu II	Final pH	Final Cu II	Final pH	Final Cu II	Final pH	Final Cu II	Final pH	Final Cu II	Final pH	Final Cu II	Final pH	Final Cu II	Final pH
0.1	24.97	7.08	0.099	7.35	0.072	7.44	0.066	8.25	0.099	7.14	0.070	7.11	0.380	7.44	0.068	4.91	0.077	7.66	0.086	8.38	0.118	8.57
0.25	62.425	7.02	0.260	7.10	0.312	7.57	0.355	8.29	0.197	7.83	0.136	7.46	0.251	7.78	0.258	4.97	0.154	5.09	0.263	8.04	0.367	8.42
0.5	124.85	6.90	0.298	8.22	0.506	8.22	0.543	8.51	0.479	7.54	0.556	7.14	0.603	7.65	0.561	5.05	0.340	7.43	0.492	7.96	0.602	8.58
0.75	187.275	6.80	0.448	6.68	0.746	7.42	0.747	8.21	0.679	7.87	0.828	7.02	0.985	7.55	0.749	5.10	0.582	7.94	0.611	7.28	0.875	8.46
1.0	249.7	6.70	0.979	6.79	0.692	7.19	1.011	8.27	0.889	7.26	1.082	7.28	1.112	7.30	0.969	5.07	0.703	8.28	0.749	7.34	0.686	7.98
1.5	374.55	6.45	0.675	7.91	0.993	6.31	0.735	7.86	0.765	7.12	1.167	6.98	1.048	7.54	1.055	5.18	0.851	5.24	0.647	7.31	0.461	8.55
2.0	499.4	6.22	0.975	6.25	0.757	7.25	0.579	7.49	0.842	5.68	1.144	6.98	0.652	5.67	0.757	5.01	0.731	5.40	0.727	7.49	0.588	5.86
2.5	624.25	6.02	0.816	6.87	0.623	7.42	0.323	7.88	0.948	5.70	1.155	6.74	0.929	5.29	0.649	5.06	0.689	5.40	0.824	7.13	0.829	6.90
3.0	749.1	5.78	0.854	5.22	0.796	6.0	0.857	5.06	0.778	5.47	1.152	5.65	0.639	4.77	0.673	4.97	0.463	6.10	0.616	6.23	MIC	
3.5	873.95	5.55	0.746	5.47	0.774	4.75	0.717	4.73	MIC		1.157	5.59	MIC		MIC		MIC		MIC			
4.0	998.8	5.35	MIC		0.811	5.09	0.726	4.75			MIC											
5.0	1248.5	4.97			MIC		MIC															

Supplementary Table 8.4. Growth in high Cu (II) tolerant endospore-forming bacteria. Growth observed (+).

Strains	Growth with metal (Cu II mM)		MIC-Cu II (mM)
	Solid media	Liquid media	
Et 9/1	3.0++	3.5+	4.0
Et 9/2	3.0+	4.0+	5.0
Lr 5/4	2.0+	4.0+	5.0
Lr 7/2	2.5+	3.0 ++	3.5
9kri 1	2.5+	3.5+	4.0
10kri 2	2.0+	3.0++	3.5
11kri 31	3.0+	3.0++	3.5
11kri323	2.0+	3.0 ++	3.5
11kri324	3.0+	3.0++	3.5
37poz32	2.0+	2.5+	3.0

Chapter 9

Active versus Dormant: Adaptations of Firmicutes to Thrive in Saline Environments

Sevasti Filippidou[1*], Thomas Junier[1,2*], Tina Wunderlin[1], Nicole Jeanneret[1], Ilona Palmieri[1], Cristina Dorador[3], Shannon Johnson[4], Patrick Chain[4], Pilar Junier[1]

[1]Laboratory of Microbiology, Institute of biology, University of Neuchatel, Neuchatel, Switzerland
[2]Vital-IT group, Swiss Institute of Bioinformatics, Lausanne, Switzerland
[3]Departamento de Acuicultura, Facultad de Recursos del Mar, Universidad de Antofagasta, Antofagasta, Chile
[4]Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA

Abstract

How do endospore-forming Firmicutes (EFF) survive in saline and hypersaline environments? The order of Halanaerobiales are halophilic Clostridia and can be found in saline environments according to culture-based and metagenomic diversity studies. It has been previously shown that they mainly accumulate KCl, instead of an organic molecule, as compatible solute for osmotic balance. This metabolic adaptation, interesting as it may be, is only one part of the story. In addition to an active mechanism of salt tolerance, EFF are also known for the production of endospores, specialized survival structures that could allow tolerance to high salt concentrations, although under a dormant metabolic state. In this study we analyzed the interplay of these two strategies: halophily/ halotolerance and dormancy, to understand the distribution and diversity of EFF in halophilic environments. In the case of halotolerance, extensive enrichment and isolation were conducted from halophilic sites. Eighty aerobic and anaerobic endospore-forming strains were isolated from nine saline habitats worldwide including marine environments, brines, salt crusts, and saline geothermal springs. These isolates cover a large taxomomic range of Bacilli and Clostridia. We verified whether these isolates are halotolerant or halophilic, and their ability to accumulate KCl. These results have been correlated to their phylogeny. In addition, the genetic mechanisms of tolerance were investigated in a selection of these strains, whose genome was fully sequenced. In halophilic Clostridia, the typical acidic signatures of the 'halophilic' proteins are absent. The genomic imprints in halophilic and halotolerant Bacilli are described herein for the first time. To evaluate the role of sporulation, we analyzed the *in situ* diversity of EFF in halophilic sites using a molecular marker specific for endospore formation. Twenty-nine samples collected along salinity gradients from different sites at the Atakama dessert, Chile, were investigated. Finally, the diversity *in situ* was compared to the knowledge gained with the cultures. Our findings suggest that in saline habitats, both survival strategies are deployed, resulting into an impressive diversity of EFF. Biochemical, genomic, ecological and environmental data are pieces to fill in the puzzle of halophilic adaptations in saline habitats.

9.1. Introduction

Elevated salt concentrations are a major challenge for living organisms. The water availability in habitats with high salt concentrations is low [1], the osmotic pressure is high [2] and the presence of oxygen is limited [3]. Additionally, most of these habitats are exposed to high doses of ultraviolet or ionizing radiation [4]. However, environments with high salt concentrations are diverse and widely spread and are far from being barren of life [4–9]. Microorganisms able to cope with high salt concentrations have developed diverse survival strategies. The production of so-called compatible solutes allows the organisms to deal with the increase in the extracellular osmotic pressure in environments with high salt content. Some of these compounds are organic osmotic solutes such as glycerol, ectoine, and glycine betaine, among others [3]. Cells that use this strategy exclude as much salt as possible from their cytoplasm and because of the low interference of the solute with the activity of intracellular enzymes; few adaptations of the proteome of the cell are required [3]. In consequence, many of these organisms are often halotolerant and can develop in a wide range of salt concentrations. In contrast to the first strategy, in the second case cells accumulate large amounts of potassium and chloride. This adaptation leads to an extensive modification of the intracellular enzymatic machinery in order to maintain enzymatic activity to near-saturation salt concentrations. This second group of organisms is apparently less diversified, in spite of the lower energetic cost of this strategy, and normally the organisms displaying accumulation of potassium and chloride has a more restricted range of salt requirements.

Both of these strategies are found in diverse groups of archaea, bacteria and eukaryotes. Among bacteria, one of the rare groups in which organisms deploying both the above-mentioned strategies is the phylum Firmicutes. The anaerobic fermentative Halanaerobiales are reported to use KCl rather than organic solutes to osmotically balance their cytoplasm [3].

In addition to an active mechanism of salt tolerance, some members of the Firmicutes are also known for their production of endospores (endospore-forming Firmicutes or EFF). Endospores are specialized structures that could allow tolerance to high salt concentrations, although under a dormant metabolic state [10].

In this study we analyzed the interplay of these two strategies: halotolerance and dormancy, to understand the distribution and diversity of EFF in halophilic environments. In the case of halotolerance, extensive enrichment and isolation were conducted from halophilic sites, followed by a physiological characterization of the role of salt for the development of the isolates. In addition, the genetic mechanisms of tolerance were investigated in a selection of sequenced strains. To evaluate the second strategy, we analyzed the *in situ* diversity of EFF in halophilic sites using a molecular marker specific for endospore formation. Finally, the diversity *in situ* was compared to the knowledge gained with the cultures.

9.2. Materials and Methods

9.2.1. Site description and sample collection

Nine saline habitats worldwide including marine environments, brines, salt crusts, and saline geothermal springs from Chile (CL), Greece (GR), and Germany (DE) were selected for this study (Figure 9.1). The samples were collected in sterile 50-ml Falcon tubes and transported to the laboratory on ice and stored at 4 ˚C for further processing.

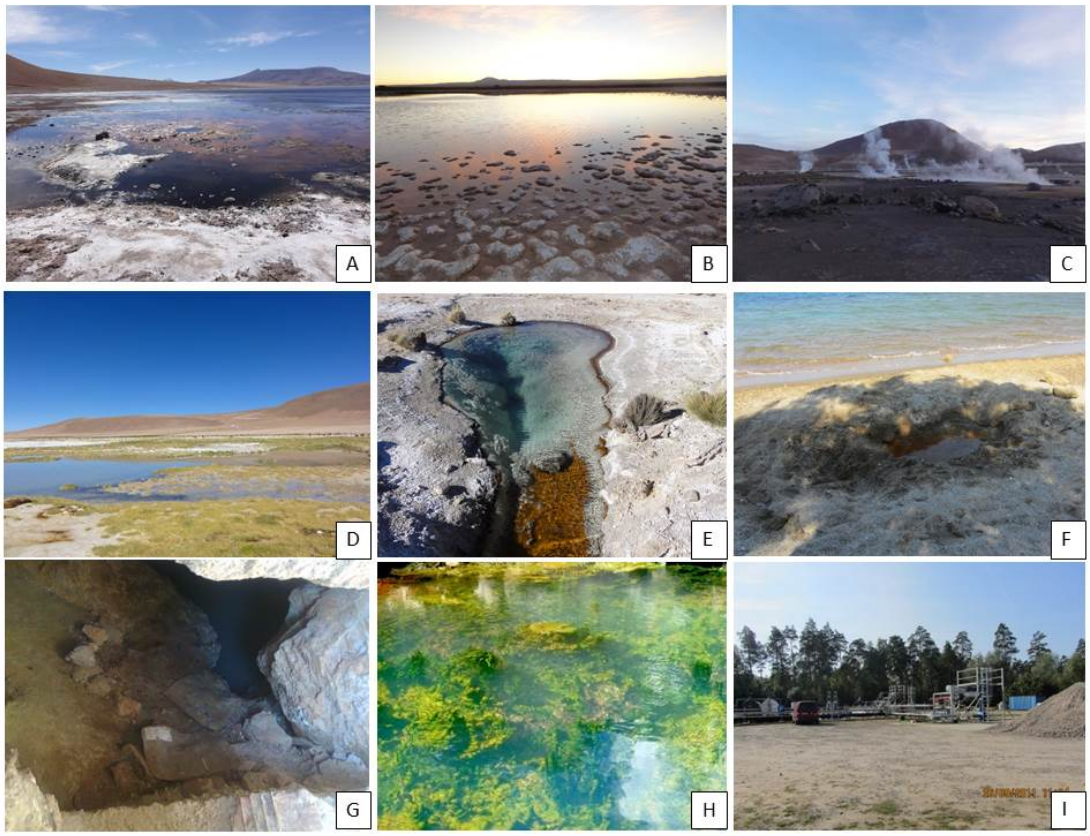


Figure 9.1. Nine saline sites included in this study. (A) Laguna de Aguas Calientes (Ac); (B) Laguna de las Piedras (Pd) in Salar de Atacama ; (C) El Tatio (Et); (D) Huasco (H); Lirima (Lr); Kanava, Milos (KAM); Agia Paraskevi (AGP); Los volcanes (col)

9.2.2. Enrichment and isolation

For enrichment, one gram of collected sample was inoculated into 10 mL of modified DifcoTM Marine Broth 2216 (D2216*, 5 g of tryptone instead of peptone, no potassium bromide, pH adjusted 5.5, 6.8, 7.4 with 2N HCl). The inoculated media were incubated at specific temperatures corresponding to those of the sampling sites under aerobic and anaerobic conditions for 72 h. The enriched cultures were then plated on modified Marine Agar (MA) (modified Marine Broth with 2% agar). Single colonies were obtained and each colony was plated repeatedly to attain pure aerobic bacterial isolates. Pure cultures were stored at 4 ˚C and cryopreserved in 25% (v/v) glycerol at -80 ˚C. EFF were selected among other isolates based on their capability to form spores after starvation for 15 days.

9.2.3. Identification of the isolated strains

Genomic DNA from each strain was extracted using the InnuPREP Bacteria DNA kit (Analytik Jena, Germany), according to the manufacturer's instructions. For amplification

of 16S rRNA gene, the primers GM3F, GM4R and Eub9_27, Eub1542 were used as previously described [11,12]. The PCR products were purified with a MultiScreen PCRμ96 Filter Plate (Millipore, USA), according to the manufacturer's instructions and sequenced using the services of Microsynth AG (Switzerland) and GATC Biotech (Germany). Nearly full-length 16S rRNA sequences were obtained by sequencing the PCR products in addition with the primers 907r, 926f primers and 518r [12,13]. The 16S rRNA gene were identified using the services of EzTaxon, against EzTaxon's cultured isolates database [14]. The sequences were submitted to GenBank under accession numbers KJ722422-KJ722533.

9.2.4. Phylogenetic analysis of 16S rRNA gene and biochemical data

Sequences of the 16S rRNA gene corresponding to the 109 strains analyzed, were aligned using MAFFT [15,16]. A maximum likelihood phylogenetic tree was built using PhyML [17] with default parameters, and then an ornament tree was created using Newick utilities [18].

9.2.5. Whole genome sequencing

Seven aerobic halotolerant strains (Et10/1, Lr3/2, Lr4/2, 24KAM51, Et7/4, GsSed3, and Gs3372) and an aerobic halophilic strain (et2/3) were selected from whole genome sequencing. Genomic DNA was extracted from an overnight culture using the QIAamp DNA minikit (Qiagen GmbH, Germany). An Illumina short-insert paired-end library constructed and sequenced[19]. The data was assembled with Velvet, version 1.2.08 [20] Genome annotation utilized an Ergatis based [21] workflow with minor manual curation then visualized with the Artemis genome browser and annotation tool [22]. The complete genome sequence content information of the seven stains is described in Table 9.1. The genomes were submitted to GenBank under BioProject numbers PRJNA 260734, 260737, and 260740-260745. Annotated genomes of the strains were screened for the presence of genes involved in osmoprotection.

9.2.6. Biofilm DNA extraction

Samples were subjected to indirect DNA extraction as previously described [23]. DNA was extracted using the FastDNA Spin Kit for Soil (MP Biomedicals, California), using a modified protocol, in order to ensure that DNA was not only extracted from vegetative cells but also from spores and other cells difficult to lyse [23]. DNA concentration was measured with a Qubit Fluorometer using a dsDNA HS Assay Kit (Invitrogen, California). The concentration of all samples was adjusted by dilution to 2ng/μl.

9.2.7. Amplicon Pyrosequencing and Analysis

Amplicon pyrosequencing of *spoOA* gene was performed using the services of Eurofins MWG Operon (Germany). A 602 bp sequence of the *spoOA* gene was amplified with the primers spoA166f and spoA748r, as previously described [23]. For quality filtering, the nucleotide sequences were translated to their amino acid sequences, based on ORF

detection. The amino acid sequences were then aligned and compared to a Gribskov-style protein profile of Spo0A [24] that was built based on 27 known Spo0A sequences. Filtration was applied as a function of the profile score and profile alignment length, which separates noise or negatives hits from true positives *spo0A* sequences. The nucleotide sequences were clustered into operational taxonomic units (OTU) at 97% sequence identity using uclust [25]. The centroid (representative sequence) of each OTU was classified using MLgsc, a general sequence classifier adapted for protein and customized to Spo0A [26].

Isolate code	Identification	Genome Size (Mb)	Number of proteins	Sequencing platform and coverage
GSsed3	*Anoxybacillus geothermalis*	6.98	6,561	Illumina, 185X
GS3372	*Aeribacillus pallidus*	4.99	4,719	Illumina, 482X
24KAM51	*Bacillus alveayuensis*	6.7	6,327	Illumina, 409X
Et2/3	*Geobacillus kaustophilus*	3.51	3,393	Illumina, 480X
Et7/4	*Geobacillus kaustophilus*	3.68	3,368	Illumina, 464X
Lr3/2	*Bacillus thurigiensis*	5.57	5,640	Illumina, 317X
Lr7/2	*Bacillus thurigiensis*	5.61	5,478	Illumina, 349X
Lr 4/2	*Bacillus cereus*	6.23	6,139	Illumina, 305X

Table 9.1 Halophilic and halotolerant aerobic endospore-forming Firmicutes were selected for whole genome sequencing and annotation.

9.2.8. Statistical analysis

Statistical analyses were performed using R, version 3.0.2 [27], Rstudio, version 0.98.1049, and BiodiversityR [28].

9.3. Results and Discussion

9.3.1. Characterization of Isolates

In total, 18 anaerobic and 55 aerobic endospore-forming Firmicutes were isolated from nine saline habitats (Supplementary Table 9.1). Twenty-four of them were thermophilic bacteria. Their identification showed that they mostly belong to the genera *Aeribacillus, Anoxybacillus, Bacillus, Clostridium, Exiguobacterium, Geobacillus, Lysinibacillus, Oceanirhabdus, Salimesophilobacter, Sporosarcina, Tepidibacter* and *Vallitalea.* Fifteen isolates could not be identified, since the sequencing of their 1600bp PCR product was interrupted shortly after initiation, giving sequences of maximum 80bp. Sequencing of these PCR products will be attempted using primers for the positions 338, 520, 907 of the 16S rRNA gene. These shorter sequences will be merged, resulting into the maximum 16S rRNA gene partial sequence possible.

Both aerobic and anaerobic isolates have been screen for their maximum limit of NaCl tolerance, as well as their minimum NaCl requirement for growth. Based on the results of this screening, ten strains tolerate up to 2.5M NaCl, 24 up to 1.65M, 1 up to 1.28, 33 up to 0.8M, one up to 0.68M, one up to 0.59M and one up to 0.51M, and two up to 0.33M. Surprisingly, 55 strains of our collection were able to grow in media with 0M NaCl

concentration. For all strains, their KCl content will be measured using a biochemical assay, at minimum and maximum concentrations of NaCl that they can tolerate. As positive control, the strain *Halanaerobium lacusrosei* was selected because it is a well described halophile that belongs to Firmicutes and is known to accumulate KCl at high saline concentrations. Their NaCl tolerance and requirement remains to be correlated to their phylogeny.

9.3.2. Genome Analysis

From our strain collection, seven isolates were selected for whole-genome sequencing. These genomes were screened for the presence of genes reported to be related to a halophilic life-style. These strains and the copy number of each gene are presented in Table 9.2. Comparison to a non-halotolerant *E. coli* stain has also been performed. Our results show that genes encoding for glycine/betaine ABC transporters, the osmotically inducible protein C, sodium: proton antiporters, the potassium transporter Trk, and the Kdp kinase are present in the genomes of halotolerant endospore-forming Firmicutes, but absent in the genome of *E. coli* strain. Although halophilic and halotolerant archaea and bacteria have been studied for decades, limited knowledge on which genes are involved in halophily is available. A genomic marker of halophily is the amino acid distribution [29]; however, it has been shown that halophilic clostridia do not necessarily show a more acidic amino acid content than mesophilic bacteria [30]. Whether this is also the case for aerobic Firmicutes needs to be tested. Therefore, an amino acid profile of these halophilic and halotolerant strains should be compared to that of mesophilic strains. Moreover, besides those genes that are already detected, further adaptations to extremely saline environments should be imprinted in the genomes of halophilic Firmicutes. Clusters of orthologous groups (COGs) of proteins from halophilic and halotolerant strains should be constructed in order to reveal further genomic adaptations to salinity. This approach has been previously applied to reveal potential genomic markers of thermophily [31].

code	identification	glycine/ betaine ABC transporter	osmotically inducible protein C	sodium: proton antiporter	potassium transporter Trk	kdp kinase
Et2/3	*G. kaustophilus*	3	0	3	1	0
Et7/4	*G. kaustophilus*	3	0	3	1	0
Lr3/2	*B. thuringiensis*	8	1	5	2	1
Lr4/2	*B.cereus*	10	1	3	3	1
24kam51	*B. alveayensis*	8	1	9	2	0
GSsed3	*A. geothermalis*	5	1	5	2	0
GS3372	*A. pallidus*	4	1	3	1	1
control	*E.coli*	0	0	3 Ca/Na:H antiporters	0	0

Table 9.2. Genes related to halophily that were identified in the genomes of the sequenced isolates.

9.3.3. Environmental Samples

A total of 44652 high quality *spoA* sequences were obtained with 28 - 4091 sequences per sample (average 1540). Sequence clustering at 97% similarity level resulted in 755 operational taxonomic units (OTUs) in the data set. The results of the analysis are summarized in Table 9.3.

The analysis of the genomic diversity of endospore-forming Firmicutes revealed that the most abundant genus was *Clostridium*, followed by *Bacillus* (Figure 9.2).

Sample Code	OTUs	Sequences	Singletons
Ac2	103	1448	33
Ac3a	244	1270	96
Ac3b	269	1119	107
Ac4a	290	1075	122
Ac4b	548	2647	215
Ac5a	351	1528	167
Ac5b	137	1477	44
Ac6a	40	65	27
Ac6b	99	183	62
Ac7	332	1983	166
col	18	28	13
Et7	82	1172	25
H10a	623	4091	263
H10b	305	1040	163
H11a	755	3362	337
H11b	714	2483	335
H1a	413	1525	216
H1b	622	1993	283
H2a	361	911	206
H2b	364	1219	186
H4a	398	1528	182
H4b	337	1410	150
H5	581	2447	256
H7a	259	1328	104
H9	289	1258	139
Pd1	450	2438	186
Pd2	295	1739	131
Pd3	50	63	42
Pd4	273	1822	93

Table 9.3 Description of the EFF communities, listed based on the number of OTUs detected in each sample.

That observation is true for all samples collected from saline environments. Sample col, which was collected from a non-saline geothermal spring in Colombia and is used for

comparison in this dataset, did not show a similar distribution of genera. In sample col, the most abundant species was *Clostridium*, followed by *Desulfotomaculum*. Overall, the most abundant groups of bacteria in these samples were anaerobic endospore- forming Firmicutes, an observation which is in accordance with the description of saline ecosystems. Indeed, these ecosystems show lower oxygen availability therefore mostly anaerobic bacteria thrive in these habitats [3].

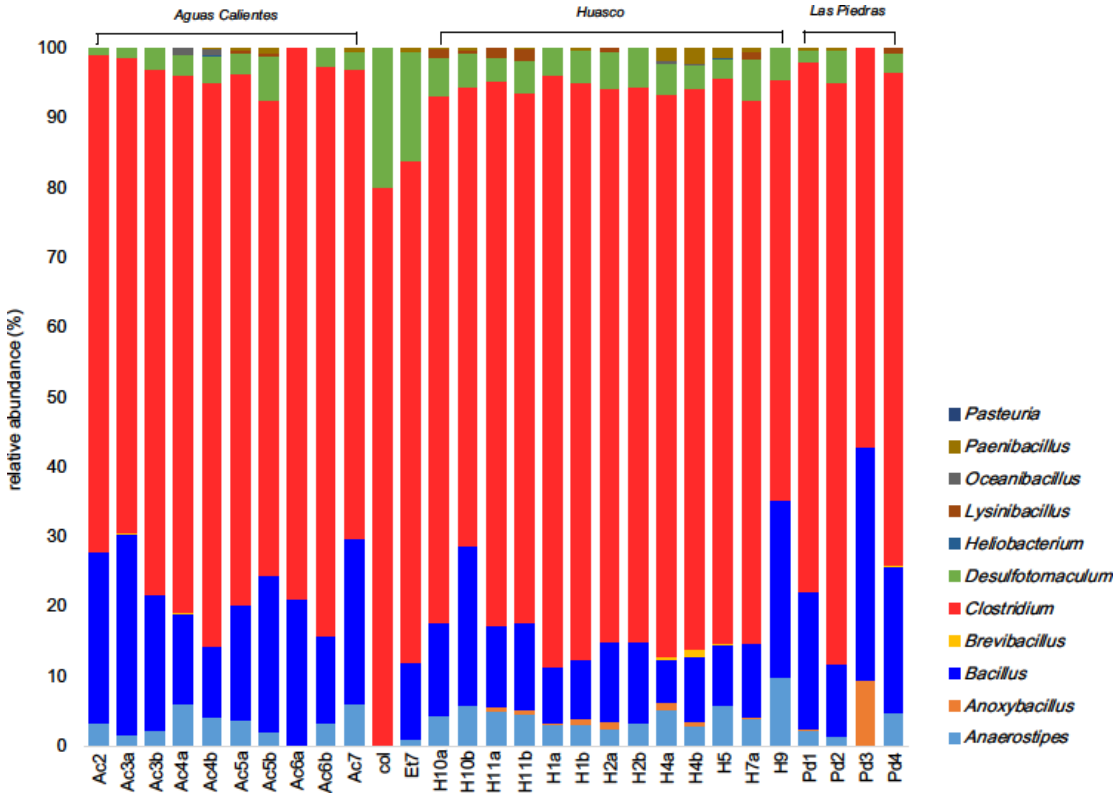


Figure 9.2. Community composition based on *spoA* gene sequencing data. The diversity of the endospore-forming Firmicutes community per sampling site. OTUs detected were classified to known endospore-forming genera. *Clostridium* sp. is prevalent in all samples.

The endospore-forming Firmicutes found in these datasets from saline environments do not correspond to previously known halophilic Firmicutes. Possible scenarios can be proposed. Firstly, halophilic and halotolerant Firmicutes are not only restricted to the order Halanaerobiales, they can also be found within the genera *Clostridium*, *Anoxybacillus*, *Lysinibacillus* and *Bacillus*, an observation that is equally supported by our environmental samples and the data from the isolates. A second scenario is that many of these bacteria, especially the aerobic Firmicutes should have been at a spore state in order to withstand not only the elevated salinity but also other extreme environmental parameters, such as UV radiation, high temperature, low water, and oxygen availability that co-occurred in these sites. A third scenario is related to the carbon and energy source availability of these sites. It has been previously shown that saline habitats harbor a high metabolic diversity of bacteria and archaea, since carbon sources alternative to glucose, electron donors different than H^+ and electron acceptors other than O_2 can be found in these environments, mostly as a result of the cell lysis of other microorganisms

that do not withstand saline conditions [2]. Firmicutes exhibit a high metabolic diversity, thus they could easily adapt to such environments [10].

Both survival strategies are deployed, resulting into an impressive diversity of EFF. Biochemical, genomic, ecological and environmental data are pieces to fill in the puzzle of halophilic adaptations in saline habitats.

9.4. References

[1] Cockell C s., Bush T, Bryce C, Direito S, Fox-Powell M, Harrison J p., et al. Habitability: A Review. Astrobiology 2016;16:89–117. doi:10.1089/ast.2015.1295.

[2] Barton L. Structural and Functional Relationships in Prokaryotes. Springer Science & Business Media; 2005.

[3] Oren A. Microbial life at high salt concentrations: phylogenetic and metabolic diversity. Saline Syst 2008;4:2. doi:10.1186/1746-1448-4-2.

[4] Demergasso C, Escudero L, Casamayor EO, Chong G, Balagué V, Pedrós-Alió C. Novelty and spatio–temporal heterogeneity in the bacterial diversity of hypersaline Lake Tebenquiche (Salar de Atacama). Extremophiles 2008;12:491–504. doi:10.1007/s00792-008-0153-y.

[5] Dong H, Zhang G, Jiang H, Yu B, Chapman LR, Lucas CR, et al. Microbial Diversity in Sediments of Saline Qinghai Lake, China: Linking Geochemical Controls to Microbial Ecology. Microb Ecol 2006;51:65–82. doi:10.1007/s00248-005-0228-6.

[6] Oren A. Diversity of halophilic microorganisms: environments, phylogeny, physiology, and applications. J Ind Microbiol Biotechnol 2002;28:56–63. doi:10.1038/sj/jim/7000176.

[7] Foti MJ, Sorokin DY, Zacharova EE, Pimenov NV, Kuenen JG, Muyzer G. Bacterial diversity and activity along a salinity gradient in soda lakes of the Kulunda Steppe (Altai, Russia). Extrem Life Extreme Cond 2008;12:133–45. doi:10.1007/s00792-007-0117-7.

[8] Abdallah MB, Karray F, Mhiri N, Mei N, Quéméneur M, Cayol J-L, et al. Prokaryotic diversity in a Tunisian hypersaline lake, Chott El Jerid. Extrem Life Extreme Cond 2016;20:125–38. doi:10.1007/s00792-015-0805-7.

[9] Dillon JG, Carlin M, Gutierrez A, Nguyen V, McLain N. Patterns of microbial diversity along a salinity gradient in the Guerrero Negro solar saltern, Baja CA Sur, Mexico. Front Microbiol 2013;4:399. doi:10.3389/fmicb.2013.00399.

[10] Oren A. The Order Halanaerobiales, and the Families Halanaerobiaceae and Halobacteroidaceae. In: Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F, editors. The Prokaryotes, Springer Berlin Heidelberg; 2014, p. 153–77.

[11] Liesack W, Weyland H, Stackebrandt E. Potential Risks of Gene Amplification by PCR as Determined by 16s rDNA Analysis of a Mixed-Culture of Strict Barophilic Bacteria. Microb Ecol 1991;21:191–8. doi:10.1007/BF02539153.

[12] Muyzer G, Teske A, Wirsen C, Jannasch H. Phylogenetic-Relationships of *Thiomicrospira* Species and their Identification in Deep-Sea Hydrothermal Vent Samples by Denaturing Gradient Gel-Electrophoresis of 16S rDNA Fragments. Arch Microbiol 1995;164:165–72. doi:10.1007/BF02529967.

[13] Ovreås L, Forney L, Daae FL, Torsvik V. Distribution of bacterioplankton in meromictic Lake Saelenvannet, as determined by denaturing gradient gel electrophoresis of PCR-amplified gene fragments coding for 16S rRNA. Appl Environ Microbiol 1997;63:3367–73.

[14] Kim O-S, Cho Y-J, Lee K, Yoon S-H, Kim M, Na H, et al. Introducing EzTaxon-e: a prokaryotic 16S rRNA gene sequence database with phylotypes that represent uncultured species. Int J Syst Evol Microbiol 2012;62:716–21. doi:10.1099/ijs.0.038075-0.

[15] Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 2002;30:3059–66. doi:10.1093/nar/gkf436.

[16] Katoh K, Standley DM. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. Mol Biol Evol 2013;30:772–80. doi:10.1093/molbev/mst010.

[17] Guindon S, Dufayard J-F, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol 2010;59:307–21. doi:10.1093/sysbio/syq010.

[18] Junier T, Zdobnov EM. The Newick Utilities: High-throughput Phylogenetic tree Processing in the Unix Shell. Bioinformatics 2010:btq243. doi:10.1093/bioinformatics/btq243.

[19] Bennett S. Solexa Ltd. Pharmacogenomics 2004;5:433–8. doi:10.1517/14622416.5.4.433.

[20] Zerbino DR, Birney E. Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 2008;18:821–9. doi:10.1101/gr.074492.107.

[21] Hemmerich C, Buechlein A, Podicheti R, Revanna KV, Dong Q. An Ergatis-based prokaryotic genome annotation web server. Bioinforma Oxf Engl 2010;26:1122–4. doi:10.1093/bioinformatics/btq090.

[22] Carver TJ, Rutherford KM, Berriman M, Rajandream M-A, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. Bioinformatics 2005;21:3422–3. doi:10.1093/bioinformatics/bti553.

[23] Wunderlin T, Junier T, Roussel-Delif L, Jeanneret N, Junier P. Stage 0 sporulation gene A as a molecular marker to study diversity of endospore-forming Firmicutes. Environ Microbiol Rep 2013:n/a – n/a. doi:10.1111/1758-2229.12094.

[24] Gribskov M, McLachlan AD, Eisenberg D. Profile analysis: detection of distantly related proteins. Proc Natl Acad Sci U S A 1987;84:4355–8.

[25] Edgar RC. Search and clustering orders of magnitude faster than BLAST. Bioinforma Oxf Engl 2010;26:2460–1. doi:10.1093/bioinformatics/btq461.

[26] Junier T, Hervé V, Wunderlin T, Junier P. MLgsc: A Maximum-Likelihood General Sequence Classifier. PloS One 2015;10:e0129384. doi:10.1371/journal.pone.0129384.

[27] R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: 2013.

[28] Kindt R, Coe R. Tree diversity analysis. A manual and software for common statistical methods for ecological and biodiversity studies. Nairobi (Kenya): World Agroforestry Centre (ICRAF); 2005.

[29] Fukuchi S, Yoshimune K, Wakayama M, Moriguchi M, Nishikawa K. Unique amino acid composition of proteins in halophilic bacteria. J Mol Biol 2003;327:347–57.

[30] Elevi Bardavid R, Oren A. The amino acid composition of proteins from anaerobic halophilic bacteria of the order Halanaerobiales. Extrem Life Extreme Cond 2012;16:567–72. doi:10.1007/s00792-012-0455-y.

[31] Makarova KS, Wolf YI, Koonin EV. Potential genomic determinants of hyperthermophily. Trends Genet TIG 2003;19:172–6. doi:10.1016/S0168-9525(03)00047-7.

9.5. Supplementary Material

Table 9.1 Description of the isolates included in this study.

Isolate Code	Identification	pH	Optimal T	Oxygen req.	NaCl req.	NaCl tolerance	Categorisation
3372	*Aeribacillus pallidus*	5.2	60	yes	0	1.28	halotolerant
21kam-2 (2)	*Aeribacillus pallidus*	7.4	60	yes	0	0.8	halotolerant
25kam-6 (2)	*Aeribacillus pallidus*	7.4	60	yes	0	0.8	halotolerant
21kam-2 (1)	*Aeribacillus pallidus*	7.4	60	yes	0.33	0.8	moderate halophile
Lr6/1	*Anoxybacillus bogrovensis*	7.4	50	yes	0	0.8	halotolerant
GS-sed3	*Anoxybacillus geothermalis*	5.2	60	yes	0	0.8	halotolerant
Lr2/1	*Anoxybacillus kamchatkensis*	7.4	55	yes	0	0.8	halotolerant
Lr1/1	*Anoxybacillus salavatliensis*	7.4	55	yes	0	1.65	halotolerant
Lr1/2	*Anoxybacillus thermarum*	7.4	55	yes	0	0.8	halotolerant
Lr10/3	*Anoxybacillus thermarum*	7.4	45	yes	0	0.8	halotolerant
Lr2/2	*Anoxybacillus thermarum*	7.4	55	yes	0	0.8	halotolerant
23kam-4 (1)	*Bacillus aeolius*	7.4	60	yes	0	0.8	halotolerant
20kam-1 (3)	*Bacillus alveayuensis*	7.4	60	yes	0	1.65	halotolerant
22kam-3 (1)	*Bacillus alveayuensis*	7.4	60	yes	0	1.65	halotolerant
22kam-3 (2)	*Bacillus alveayuensis*	7.4	60	yes	0	1.65	halotolerant
24kam-5 (1)	*Bacillus alveayuensis*	7.4	60	yes	0	1.65	halotolerant
Lr3/3	*Bacillus aquimaris*	7.4	RT	yes	0	2.5	halotolerant
Lr4/1	*Bacillus aquimaris*	7.4	RT	yes	0	2.5	halotolerant
Et6/3	*Bacillus boroniphilus*	7.4	37	yes	0	1.65	halotolerant
Pd2/1	*Bacillus boroniphilus*	7.4	RT	yes	0	1.65	halotolerant
Lr3/2	*Bacillus cereus*	7.4	RT	yes	0	0.8	halotolerant
Lr7/1	*Bacillus cereus*	7.4	RT	yes	0	0.8	halotolerant
Lr4/2	*Bacillus cereus*	7.4	RT	yes	0	0.8	halotolerant
Yg6	*Bacillus horikoshii*	7.4	RT	yes	0.33	1.65	halophile
Et8/5	*Bacillus jeotgali*	7.4	37	yes	0	1.65	halotolerant
Lr124/5	*Bacillus jeotgali*	7.4	45	yes	0	1.65	halotolerant
Lr10/1	*Bacillus licheniformis*	7.4	45	yes	0	2.5	halotolerant

Lr124/1	*Bacillus licheniformis*	7.4	45	yes	0	2.5	halotolerant
Pd2/4	*Bacillus licheniformis*	7.4	RT	yes	0	1.65	halotolerant
Ac1/1	*Bacillus oceanisediminis*	7.4	RT	yes	0	1.65	halotolerant
Yg2/2	*Bacillus oceanisediminis*	7.4	RT	yes	0	1.65	halotolerant
18pam-3	*Bacillus sp.*	7.4	37	yes	0	2.5	halotolerant
Lr10/2	*Bacillus subterraneus*	7.4	45	yes	0	1.65	halotolerant
Et6/1	*Bacillus thioparans*	7.4	37	yes	0	2.5	halotolerant
Lr3/1	*Bacillus thioparans*	7.4	RT	yes	0	1.65	halotolerant
Et10/1	*Bacillus thuringiensis*	7.4	45	yes	0	0.8	halotolerant
Et6-ana1	*Clostridium sp.*	7.6	37	no	0.33	0.8	moderate halophile
Et8-ana1	*Clostridium sp.*	7.6	37	no	0.33	0.8	moderate halophile
Lr7-ana1	*Clostridium sporogenes*	7.6	RT	no	0	0.51	halotolerant
47agp-5 (1)	*Exiguobacterium aurantiacum*	7.4	37	yes	0	1.65	halotolerant
44agp-2 (1)	*Exiguobacterium mexicanum*	7.4	37	yes	0	2.5	halotolerant
44agp-2 (3)	*Exiguobacterium sp.*	7.4	37	yes	0	2.5	halotolerant
45agp-3	*Exiguobacterium sp.*	7.4	37	yes	0	1.65	halotolerant
Et2/3	*Geobacillus kaustophilus*	7.4	70	yes	0.09	0.33	marine
Et2/1	*Geobacillus sp.*	7.4	70	yes	0.09	0.33	marine
Et7/4	*Geobacillus stearothermophilus*	5.5	65	yes	0	0.8	halotolerant
48AGP6-ana4	*Geosporobacter sp.*	7.6	37	no	0	0.8	halotolerant
Lr11/1	*Lysinibacillus fusiformis*	7.4	RT	yes	0	0.8	halotolerant
18PAM3-ana3	*Oceanirhabdus sediminicola*	7.6	37	no	0.33	1.65	halophile
27TPM2-ana2	*Oceanirhabdus sediminicola*	7.6	37	no	0.09	0.8	moderate halophile
19PAM4-ana5-2	*Salimesophilobacter sp.*	7.6	37	no	0.33	0.8	moderate halophile
44agp-2 (2)	*Sporosarcina saromensis*	7.4	37	yes	0	0.8	halotolerant
AC8-ana1	*Tepidibacter mesophilus*	7.6	RT	no	0.33	1.65	halophile
17PAM2-ana3-1	*Tepidibacter mesophilus*	7.6	37	no	0.33	0.68	moderate halophile
26TPM1-ana3	*Tepidibacter mesophilus*	7.6	37	no	0.33	0.8	moderate halophile
47AGP5-ana3	*Tepidibacter sp.*	7.6	37	no	0.09	0.8	moderate halophile
44AGP2-ana2	*Vallitalea sp.*	7.6	37	no	0.09	0.8	moderate halophile

Chapter 10

Synthesis

The main aim of this thesis was to understand the survival of endospore-forming Firmicutes in extreme environments either by spore-formation or by metabolic adaptation. In this chapter, the major conclusions and perspectives are presented.

10.1. Under-detection of endospore-forming Firmicutes in metagenomics datasets

In the first research chapter, our aim was to verify whether endospore-forming Firmicutes can be detected in metagenomics datasets. Specific findings were:

- Although endospore-forming Firmicutes have been isolated from various environments, they are under-detected by profile analysis of sporulation genes in metagenomes.
- Endospore formers were absent even from those habitats known to harbor them.
- A tailored DNA extraction method was applied to the sample; qPCR quantification of endospore-forming Firmicutes in comparison to the total bacterial community showed a 100% of endospore-formers in a sample collected from a hot, moderately radioactive spring. This extraction method improved detection in amplicon sequencing (40%).
- Ameliorated DNA extraction did not improve shotgun classification.
- Endospore-formers represent an undetectable community fraction by metagenomic approaches.
- Improved methods of classification and phylogenetic assignment of metagenomic data need to be developed.

In the first published article of this work we have demonstrated how biases in DNA sequencing can have impacts on conclusions made from culture-independent DNA studies. Endospore-forming Firmicutes, as well as maybe other bacterial groups, are un-detectable in metagenomics datasets and this may be due to biases in annotation or in the depth of sequencing required.

Approaches to evaluate the existing methodology, as well as studies to overcome the biases described up to date, take into account the nature of each microbial group that is into focus, the environment that is studied, and the ecological question that is to be answered. An example of such an approach is a study published in 2014, developing methods to better study a specific group of bacteria (Cyanobacteria) in a challenging environment (oceans) from a biogeographical point of view [1].

10.2. Extreme conditions favor the prevalence of Endospore-forming Firmicutes in natural springs

In the second research chapter, we depart from the principles that on one hand, sporulation is an energetically costly survival strategy, however when conditions are harsh, this strategy has a selective advantage and that on the other hand, Firmicutes is one of the most diverse phyla known, containing species with variable metabolic adaptation properties. Thus, we formed the hypothesis that the more extreme the

environment, the more prevalent endospore-forming Firmicutes should be. The main findings were:

- We provide experimental evidence, for the first time, that endospore-forming Firmicutes are an ubiquitous microbial group.
- Endospore-forming Firmicutes do not prevail in every environment. On the contrary, they represent a small fraction of the total bacterial community, since the survival strategies they deploy are energy-demanding to maintain, which gives them an ecological disadvantage when those strategies are "not needed". This was observed in mineral springs, where environmental factors were within mesophilic limits.
- In geothermal sites where a single environmental factor was exceeding mesophilic limits, the relative abundance of EFF did not increase, although each individual limiting factor studied here was reported to reduce microbial abundance in general.
- In poly-extreme environments, the previously described disadvantage turns into a survival advantage. In these environments, endospore-forming Firmicutes are prevalent and represent an important fraction of the bacterial community. This was shown in geothermal springs and drillings were a combination of extreme environmental parameters (such as temperature, pressure, and pH) were in place.
- Analysis of bacterial diversity showed the presence of endospore-forming species that could persist in harsh environments in spore-state. However, extremophilic endospore-forming Firmicutes were also found, suggesting that they may inhabit these sites in vegetative form.
- Ecological theories proposed for microbial distribution cannot be applied to endospore-forming Firmicutes, and this may be the case for other microbial groups that express unique adaptation characteristics.

We showed for the first time that a bacterial group with enhanced survival strategies exhibits a unique distribution pattern, behave as a 'system within a system' and can be influenced by their environment in a different way than the whole bacterial community. This may be the case for other bacterial groups that develop complex survival strategies such as the formation of biofilms. Further studies on the distribution of unique microbial groups should reveal pattern for their survival in relation to their environments.

10.3. Extremity: a driving force for sporulation in species that are not known to sporulate

We have shown that under coexistence of unfavorable conditions, sporulation is a beneficial characteristic for survival and dispersal. In the following two research chapters we argued that extremity can also be a driving force for sporulation in species that are not known to sporulate. Specific findings were:

- A *Serratia ureilytica* strain (Lr5/4) was isolated in a geothermal spring at Lirima, Chile. This strain could not possibly survive the *in situ* environmental parameters, because they were exceeding its limits of tolerance.

- This strain was found to produce viable and resistant spores. The presence of spores was shown microscopically. Spores are produced spontaneously after starvation but also after temperature fluctuations.
- The spores of *S. ureilytica* strain Lr5/4 are resistant to UV radiation, high temperature and desiccation. This observation is in accordance with the environmental characteristics of the environment from which it was isolated.
- The molecular mechanism of sporulation was investigated in comparison to other four known sporulation pathways (Firmicutes, Actinomycetes, Cyanobacteria and *Myxococcus*).
- Homologs to sporulation proteins in Firmicutes were detected throughout the bacterial chromosome of strain Lr5/4.
- A horizontal gene transfer from thermophilic anaerobic Firmicutes to this strain was observed.
- We isolated strain 11kri321, a representative of the genus *Kurthia*, from an oligotrophic geothermal reservoir. The genus *Kurthia* is considered asporogenic; however, this strain produced endospores.
- The molecular pathway of sporulation in *Kurthia* strain 11kri321 was revealed. Genes previously considered as essential for sporulation in *Bacillus subtilis* were observed in the genome of this strain.
- These genes were also detected in three other publically available genomes of the genus *Kurthia*.
- We introduce the term "cryptosporulation" to describe lineages of Firmicutes that have not previously been observed to sporulate and lack a genome analysis for sporulation genes.

In the microbial world there are well-documented examples of transfer of function through horizontal gene transfer. Some of those functions correspond to key steps in the biogeochemical cycling of elements as important as sulfur and nitrogen. Based on phylogenetic analysis, the gene encoding the protein catalyzing the last step during sulfate reduction is a candidate of horizontal gene transfer leading to an acquired function [1]. Adaptation through horizontal gene transfer appears not to be exclusive to Bacteria and Achaea. The completion of the genome of *Galdieria sulphuraria*, one of the few eukaryotic unicellular organisms that can thrive in geothermal sites, has revealed that this alga acquired genes from prokaryotes, providing a remarkable metabolic versatility and more importantly, the ability to survive in its hostile environment (e.g. transfer of genes to detoxify mercury and arsenic) [2]. Moreover, it appears that the acquisition of genes is environment-specific, as the sequencing of the genome of a representative of the sister taxon *Galdieria phlegrea*, which is adapted to dry habitats near fumaroles such as fissures between rocks or cryptoendolithic environments, revealed extensive gene loss and re-adaptation through multiple events of gene transfer from bacteria [3]. Even though the idea of transfer of genetic material as a driving force of evolution in complex cells is controversial, horizontal gene transfer has been brought to the fore very recently when studying the processes leading to the emergence of complex eukaryotic cells.

Transfer of genes by viruses or, in the case of unicellular eukayrotes, by ingestion and digestion of prey are mechanisms facilitating horizontal gene transfer [4]. Although a consensus supporting the transfer of genes through endosymbiotic gene transfer is starting to emerge [5], this shows the importance of considering events of function transfer in the evolution of living organisms.

It is clear that the environment plays a key role on favoring the extent to which processes such as horizontal gene transfer can occur. Cohabitation and the potential benefit of acquiring functions are key elements that might tilt the odds of horizontal gene transfer as a successful strategy [6]. This might be the case of geothermal environments in which the variety of alternative chemoautotrophic and organotrophic metabolisms in closely related microorganisms (based on 16S rDNA) suggests that horizontal gene transfer may be common. Physical proximity and mutualistic or syntrophic relationships among Bacteria and Archaea would likely be important factors aiding DNA exchange between the domains. Therefore, these types of environments can be the ideal "living" laboratories to explore the role of horizontal gene transfer in the evolution of bacteria. Sporulation is a complex procedure that demands a large number of genes to be successfully completed. Which is the minimum set of these genes that can be transferred and how is yet to be described.

As indicated in chapter four, sporulation in *Serratia* appears to be common to at least two strains (*S. ureilytica* Lr5/4 and *S. marcescens* subsp. *sakuensis* [7]). The genome of the latter stain needs to be sequenced in order to identify candidate genes for the sporulation pathway that could have been transferred by horizontal gene transfer. Its analysis may reveal common genes between the two *Serratia* species that may improve our knowledge of sporulation in *Serratia*, but also may indicate the sporulation genes that are more easily transferable from Firmicutes to this genus.

Moreover, the whole sporulation pathway should be shown visually (microscopically) as well as molecularly. In order to do so, synchronisation of the bacterial cells should be attempted, at different spore-developmental stages, as has been shown for *Bacillus* [8]. RNA sequencing at each stage as well as at vegetative and spore state should be held. The environmental factors that trigger sporulation in this genus should also be studied in more detail, concerning their influence on the *Serratia* Lr5/4 or *S. marcescens* subsp. *sakuensis* pathway.

10.4. *Anoxybacillus geothermalis*: a novel species that thrives in geothermal reservoirs

In this research chapter we described the discovery of a novel *Anoxybacillus* species. Highlights of this study were:

- This species was isolated from an oligotrophic, 4 km deep, high-pressure and high-temperature geothermal reservoir.
- The water in the reservoir is dated to the Permian age. Since the *in situ* conditions are unfavorable for growth, it is proposed that *A. geothermalis* strain GSsed3 was in a spore state in the reservoir. The reservoir is a closed system, therefore it is

hypothesised that this strain remained in spore state since the Permian age and revived back in the laboratory.

- The same species was detected in two other geothermal reservoirs, with similar characteristics; therefore we propose that this species has a particular ecological niche in geothermal reservoirs.

An impressive observation in this novel species is the size of its genome. Compared to other *Anoxybacillus* species whose genome is already available, the genome of this strain is almost double in size and it also has a large number of predicted proteins. Considering the possibility that this strain has remained at spore state since the Permian age, it is an appropriate candidate for the study of loss or gain of genes throughout the evolutionary history of this genus.

10.5. Spore-forming Isolates for the description of potential biomarkers of extremity

This research chapter focuses on fourteen isolates from our collection of spore-forming bacteria, whose chromosome has been fully sequenced. Major findings of this work are:

- The genome sequences of three bacilli were announced for the first time and are now publically available: *Bacillus alveayuensis* strain 24KAM51, *Aeribacillus pallidus* strain GS3372, and *Anoxybacillus geothermalis* strain GSsed3.
- All fourteen strains can be categorized into thermophilic and mesophilic stains, alkalophilic, neutrophilic and acidophilic, halotolerant or not, and finally they can be separated into groups based on their capability of metal tolerance. Their genomes can be screened for previously described biomarkers of "extremophily".

Although the physiological adaptations of microorganisms to face extremity have been well described over the last decades [9,10], the imprints of these modifications on the genomes of these organisms are not fully described. This description should allow a better understanding of the ecological adaptations, as well as the transfer of extremophilic characteristics among microorganisms. The description of the genes restricted to adaptations in specific environments should enable the discovery of biological markers of thermophilic, psychrophilic, acidophilic, alkalophilic lifestyles in order to better understand the species geographical distribution. Moreover, such descriptions could also provide insights of cellular mechanisms that have evolved in eukaryotic cells, as in the case of the reverse gyrase, which is unique to hyperthermophiles. The discovery of this gene resulted into a better characterization of the enzyme it encodes. Its deactivation and denaturation result into cellular death of the microbe. It is speculated that the steps of microbial cell death are similar to those of eukaryotic apoptosis [11]. Finally, further biotechnological applications could be developed by studying specific extremophilic markers [12]. Nowadays, advances in biotechnology focus on the study of novel proteins in extremophiles, discovered by studying genomes [13–16].

10.6. Manganese-II oxidation and Copper-II resistance in endospore forming Firmicutes isolated from uncontaminated environmental sites

In this study, Mn (II) oxidation and copper tolerance were evaluated for a collection of aerobic endospore-forming Firmicutes isolated from a diverse set of natural uncontaminated sites environments. The main findings were:

- Uncontaminated environmental sites harbor an abundant population of manganese oxidizing as well as copper tolerant aerobic endospore-forming Firmicutes.
- Manganese oxidation and copper tolerance are specific in some genera, such as *Anoxybacillus* and *Lysinibacillus*.
- Manganese oxidation and copper tolerance in nature are not specific to bacterial species or site location.
- Endospore-formers have natural resistance mechanisms for toxic metals and metal resistance is a wide spread phenomenon in endospore-forming Firmicutes.

Manganese oxidation and copper resistance are mechanisms controlled by a series of genes. The former, has been well described in endospore-forming Firmicutes vegetative cells [17] but also spores [18]. Copper resistance at a genomic level, however, is well studied in microbial groups [19], resulting into the design of primers for the direct detection of copper resistant genes (*cop* family) in the environment [20]. The detection of these genes among endospore-forming Firmicutes remains to be studied.

In this study we have shown that manganese oxidation and copper resistance are two characteristics that do not necessarily co-occur in endospore-forming Firmicutes in uncontaminated environments. However, the role of copper-binding manganese oxidases [21] needs to be better understood in contaminated environments since it has been shown that this enzyme is capable of manganese oxidation and thus removal from the environment [22]. Finally, whether the observation that endospore-forming Firmicutes resistance to manganese and copper in uncontaminated environments can be extrapolated to other heavy/trace metals need to be further examined.

10.7. Active versus Dormant: Adaptations of Firmicutes to Thrive in Saline Environments

In this chapter, we have investigated the tolerance of saline environments in endospore-forming Firmicutes. This bacterial group survives in those habitats either by adaptations to withstand osmotic pressure and low water and oxygen availability, or by endospore-formation. Preliminary findings were:

- Aerobic endospore-forming Firmicutes are able to tolerate high NaCl concentrations (up to 2.5M).
- Based on the up to date identification of our collection, anaerobic endospore-forming halotolerant Firmicutes belong to the genera *Tepidibacter, Vallitalea, Oceanirhabdus, Salimesophilobacter, and Clostridium.* Further identification may reveal anaerobic Firmicutes that belong to the order Halanaerobiales.

- Previously described genomic imprints of halophily were found in the genomes of seven aerobic endospore-forming Firmicutes. These genes encode for glycine/betaine ABC transporters, the osmotically inducible protein C, sodium: proton antiporters, the potassium transporter Trk, and the Kdp kinase.
- Diversity of endospore-forming Firmicutes in saline habitats showed that the most prevalent genera are *Clostridium*, *Bacillus* and *Desulfotomaculum*.

Clusters of orthologous groups (COGs) of proteins from halophilic and halotolerant strains should be constructed in order to reveal further genomic adaptations to salinity. The amino acid profile of these strains should be further investigated in order to verify whether their proteins are more acidic, as it has been previously shown [23]. Finally, in order to verify whether endospore-forming Firmicutes are at a vegetative or dormant state in these habitats, an investigation, including a previously described method for the separation of spore structures from vegetative cells [24], should be conducted for the samples isolated from these habitats.

10.8. References

[1] Klein M, Friedrich M, Roger AJ, Hugenholtz P, Fishbain S, Abicht H, et al. Multiple lateral transfers of dissimilatory sulfite reductase genes between major lineages of sulfate-reducing prokaryotes. J Bacteriol 2001;183:6028–35. doi:10.1128/JB.183.20.6028-6035.2001.

[2] Schönknecht G, Chen W-H, Ternes CM, Barbier GG, Shrestha RP, Stanke M, et al. Gene transfer from bacteria and archaea facilitated evolution of an extremophilic eukaryote. Science 2013; 339:1207–10. doi:10.1126/science.1231707.

[3] Qiu H, Price DC, Weber APM, Reeb V, Yang EC, Lee JM, et al. Adaptation through horizontal gene transfer in the cryptoendolithic red alga *Galdieria phlegrea*. Curr Biol CB 2013;23:R865–6. doi:10.1016/j.cub.2013.08.046.

[4] Archibald JM. Evolution: Gene transfer in complex cells. Nature 2015;524:423–4. doi:10.1038/nature15205.

[5] Ku C, Nelson-Sathi S, Roettger M, Sousa FL, Lockhart PJ, Bryant D, et al. Endosymbiotic origin and differential loss of eukaryotic genes. Nature 2015;524:427–32. doi:10.1038/nature14963.

[6] Arber W. Horizontal Gene Transfer among Bacteria and Its Role in Biological Evolution. Life Basel Switz 2014;4:217–24. doi:10.3390/life4020217.

[7] Ajithkumar B, Ajithkumar VP, Iriye R, Doi Y, Sakai T. Spore-forming *Serratia marcescens* subsp. *sakuensis* subsp. nov., isolated from a domestic wastewater treatment tank. Int J Syst Evol Microbiol 2003;53:253–8. doi:10.1099/ijs.0.02158-0.

[8] Ambriz-Aviña V, Contreras-Garduño JA, Pedraza-Reyes M. Applications of flow cytometry to characterize bacterial physiological responses. BioMed Res Int 2014;2014:461941. doi:10.1155/2014/461941.

[9] Barton L. Structural and Functional Relationships in Prokaryotes. Springer Science & Business Media; 2005.

[10] Cockell CS. Astrobiology: Understanding Life in the Universe. John Wiley & Sons; 2015.

[11] Perugino G, Valenti A, D'Amaro A, Rossi M, Ciaramella M. Reverse gyrase and genome stability in hyperthermophilic organisms. Biochem Soc Trans 2009;37:69–73. doi:10.1042/BST0370069.

[12] Littlechild JA. Archaeal Enzymes and Applications in Industrial Biocatalysts. Archaea Vanc BC 2015;2015:147671. doi:10.1155/2015/147671.

[13] Karan R, Capes MD, DasSarma S. Function and biotechnology of extremophilic enzymes in low water activity. Aquat Biosyst 2012;8:4. doi:10.1186/2046-9063-8-4.

[14] Cavicchioli R, Charlton T, Ertan H, Omar SM, Siddiqui KS, Williams TJ. Biotechnological uses of enzymes from psychrophiles. Microb Biotechnol 2011;4:449–60. doi:10.1111/j.1751-7915.2011.00258.x.

[15] Rashid M, Stingl U. Contemporary molecular tools in microbial ecology and their application to advancing biotechnology. Biotechnol Adv 2015;33:1755–73. doi:10.1016/j.biotechadv.2015.09.005.

[16] Dalmaso GZL, Ferreira D, Vermelho AB. Marine extremophiles: a source of hydrolases for biotechnological applications. Mar Drugs 2015;13:1925–65. doi:10.3390/md13041925.

[17] Dick GJ, Torpey JW, Beveridge TJ, Tebo BM. Direct Identification of a Bacterial Manganese(II) Oxidase, the Multicopper Oxidase MnxG, from Spores of Several Different Marine *Bacillus* Species. Appl Environ Microbiol 2008;74:1527–34. doi:10.1128/AEM.01240-07.

[18] Su J, Bao P, Bai T, Deng L, Wu H, Liu F, et al. CotA, a multicopper oxidase from *Bacillus pumilus* WH4, exhibits manganese-oxidase activity. PloS One 2013;8:e60573. doi:10.1371/journal.pone.0060573.

[19] Behlau F, Hong JC, Jones JB, Graham JH. Evidence for acquisition of copper resistance genes from different sources in citrus-associated xanthomonads. Phytopathology 2013;103:409–18. doi:10.1094/PHYTO-06-12-0134-R.

[20] Sarkar B. Heavy Metals In The Environment. CRC Press; 2002.

[21] Butterfield CN, Tao L, Chacón KN, Spiro TG, Blackburn NJ, Casey WH, et al. Multicopper manganese oxidase accessory proteins bind Cu and heme. Biochim Biophys Acta 2015;1854:1853–9. doi:10.1016/j.bbapap.2015.08.012.

[22] Su J, Deng L, Huang L, Guo S, Liu F, He J. Catalytic oxidation of manganese(II) by multicopper oxidase CueO and characterization of the biogenic Mn oxide. Water Res 2014;56:304–13. doi:10.1016/j.watres.2014.03.013.

[23] Fukuchi S, Yoshimune K, Wakayama M, Moriguchi M, Nishikawa K. Unique amino acid composition of proteins in halophilic bacteria. J Mol Biol 2003;327:347–57.

[24] Wunderlin T, Junier T, Paul C, Jeanneret N, Junier P. Physical Isolation of Endospores from Environmental Samples by Targeted Lysis of Vegetative Cells. J Vis Exp JoVE 2016. doi:10.3791/53411.

Chapter 11

Parallel collaborations

11.1 Exploiting the fungal highway: development of a novel tool for the in situ isolation of bacteria migrating along fungal mycelium

Simon A[1], Bindschedler S[2], Job D[1], Wick LY[2], Filippidou S[1], Kooli WM[1], Verrecchia EP[3], Junier P[1]

Author information

[1]Laboratory of Microbiology, Institute of Biology, University of Neuchâtel, 2000 Neuchâtel, Switzerland.

[2]Department of Environmental Microbiology, Helmholtz Centre for Environmental Research - UFZ, 04318 Leipzig, Germany.

[3]Biogeosciences laboratory, Institute of Earth Surface Dynamics, University of Lausanne, 1015 Lausanne, Switzerland.

Abstract

Fungi and bacteria form various associations that are central to numerous environmental processes. In the so-called fungal highway, bacteria disperse along fungal mycelium. We developed a novel tool for the *in situ* isolation of bacteria moving along fungal hyphae as well as for the recovery of fungi potentially involved in dispersal, both of which are attracted towards a target culture medium. We present the validation and the results of the first *in situ* test. Couples of fungi and bacteria were isolated from soil. Amongst the enriched organisms, we identified several species of fast-growing fungi (*Fusarium* sp. and *Chaetomium* sp.), as well as various potentially associated bacterial groups, including *Variovorax soli, Olivibacter soli, Acinetobacter calcoaceticus*, and several species of the genera *Stenotrophomonas, Achromobacter* and *Ochrobactrum*. Migration of bacteria along fungal hyphae across a discontinuous medium was confirmed in most of the cases. Although the majority of the bacteria for which migration was confirmed were also positive for flagellar motility, not all motile bacteria dispersed using their potential fungal partner. In addition, the importance of hydrophobicity of the fungal mycelial surface was confirmed. Future applications of the columns include targeting different types of microorganisms and their interactions, either by enrichment or by state of the art molecular biological methods.

Published in FEMS Microbiology Ecology in 2015

11.2 Genome Sequence of Kosakonia radicincitans Strain YD4, a Plant Growth-Promoting Rhizobacterium Isolated from Yerba Mate (Ilex paraguariensis St. Hill.)

Bergottini VM[1], Filippidou S[1], Junier T[2], Johnson S[3], Chain PS[3], Otegui MB[4], Zapata PD[4], Junier P[1]

Author information

[1]Laboratory of Microbiology, Neuchâtel, Switzerland.

[2]Vital-IT Group, Swiss Institute of Bioinformatics, Lausanne, Switzerland.

[3]Los Alamos National Laboratory, Los Alamos, New Mexico, USA.

[4]Instituto de Biotecnología, Misiones, Argentina.

Abstract

Kosakonia radicincitans strain YD4 is a rhizospheric isolate from yerba mate (*Ilex paraguariensis* St. Hill.) with plant growth-promoting effects on this crop. Genes involved in different plant growth-promoting activities are present in this genome, suggesting its potential as a bioinoculant for yerba mate.

Published in Genome Announcements in 2015.

11.3 Gains of bacterial flagellar motility in a fungal world

Pion M, Bshary R, Bindschedler S, Filippidou S, Wick LY, Job D, Junier P.
Author information
Laboratory of Microbiology, Institute of Biology, University of Neuchâtel, Neuchâtel, Switzerland.

Abstract

The maintenance of energetically costly flagella by bacteria in non-water-saturated media, such as soil, still presents an evolutionary conundrum. Potential explanations have focused on rare flooding events allowing dispersal. Such scenarios, however, overlook bacterial dispersal along mycelia as a possible transport mechanism in soils. The hypothesis tested in this study is that dispersal along fungal hyphae may lead to an increase in the fitness of flagellated bacteria and thus offer an alternative explanation for the maintenance of flagella even in unsaturated soils. Dispersal along fungal hyphae was shown for a diverse array of motile bacteria. To measure the fitness effect of dispersal, additional experiments were conducted in a model system mimicking limited dispersal, using *Pseudomonas putida* KT2440 and its non-flagellated (ΔfliM) isogenic mutant in the absence or presence of *Morchella crassipes* mycelia. In the absence of the fungus, flagellar motility was beneficial solely under conditions of water saturation allowing dispersal, while under conditions limiting dispersal, the nonflagellated mutant exhibited a higher level of fitness than the wild-type strain. In contrast, in the presence of a mycelial network under conditions limiting dispersal, the flagellated strain was able to disperse using the mycelial network and had a higher level of fitness than the mutant. On the basis of these results, we propose that the benefit of mycelium-associated dispersal helps explain the persistence of flagellar motility in non-water-saturated environments.

Published in Applied and Environmental Microbiology in 2013.

11.4 Emergent spatial patterns allow the coexistence of competing bacterial species in a homogeneous environment

Filippidou S[1], Pion M[1], Bshary R[2], and Junier P[1].
Author information
[1]Laboratory of Microbiology, Institute of Biology, University of Neuchâtel, Switzerland
[2]Laboratory of Eco-ethology, Institute of Biology, University of Neuchâtel, Switzerland

Abstract

Natural systems harboring co-existing species with niche overlap need to be reconciled with evolutionary theory, as natural selection should cause the extinction of less competitive species. Self-structuring patterns allowing species-coexistence are predicted in non-transitive communities and host-parasitoids systems. In these models, as well as in neutral models of biodiversity, key assumptions involve the scale of processes such as dispersal and interaction occurring at a local scale.

Up to now, little experimental evidence exists to test the influence of these processes. In this study, we experimentally evaluate the role of dispersal on species-coexistence for two-species combinations of bacterial strains with unequal fitness (*Pseudomonas putida* KT2440 (WT), its antibiotic-resistant mutant *P. putida* UWC1, *Pseudomonas azelaica* HBP1, *Cupriavidus necator* JMP289, and *Escherichia coli* K12). In addition, we considered the role of cell density, media composition, and features such as flagella production, swarming, swimming, growth rate, or cell size and shape, factors that can affect dispersal.

Here we show that local rather than global dispersal may allow the emergence of spatial patterns between two competing bacterial strains in a homogeneous environment. None of the physical parameters we measured consistently correlated with stable versus instable coexistence, suggesting that multiple factors, affect the formation of spatial patterns and hence coexistence. Such outcomes are robust with respect to changes in cell density or media composition but vanish if mobility is enhanced. Coexistence was linked to self-structuring spatial patterns, observed at the colony growth level and this is dependent on the level of the interaction (local versus global); this observation could not be predicted from species-specific growth patterns. Moreover, the spatial structures were robust regardless the changes in the initial ratio of the strains in the inoculum (from 1:1 to 10:1 or 1:10), agar type, and substrate composition. Our results suggest a multifactorial mechanism leading to species co-existence.

Article in preparation

Chapter 12

Bacterial diversity in the sulfur and iron springs of Ponts-de-Martel, Neuchâtel

Sevasti Filippidou, Saskia Bindschedler, Nicole Jeanneret, Matthieu Bueche & Pilar Junier
Laboratory of Microbiology, University of Neuchâtel, Rue Emile-Argand 11, 2000 Neuchatel.

This chapter is adapted from the article: Filippidou S., Bindschedler S., Jeanneret N., Bueche M., Junier P. Bacterial diversity in the sulfur and iron springs of Ponts-de-Martel, Neuchâtel, submitted for publication at *Bulletin neuchatelois des sciences naturelles*

Abstract

Two interesting models to the effect of microbial activity on a landscape can be found near the village of Les Ponts-de-Martel, Neuchâtel. The emergence of anaerobic water highly enriched in hydrogen sulfide and ferrous iron offers a unique opportunity for the development of microorganisms extracting their energy from these inorganic chemicals. Two entire microbial ecosystems develop thereafter, fuelled by the activity of these microorganisms. These ecosystems have now been studied by genomic sequencing of bacterial molecular markers and the diversity of the bacterial communities in both springs is described herein. For the sulfur spring the most abundant identified bacterial groupscorresponded to photo- and chemotrophic sulfur-oxidizing bacteria, as well as sulfate-reducing bacteria. In the iron spring a few genera known to be involved in ironoxidation and iron reduction could be identified. However, the relationship between the diversity of bacteria and their potential role in the iron cycle was not as clear as it was for the bacterial community observed in the sulfur spring. Overall the results presented here shedlight on the microbial processes occurring in these remarkable microbial ecosystems and might prompt further interest from the general populationand for future generations of scientists to study in detail the relationship between microbial diversity and iron and sulfur cycling.

12.1. Introduction

Microorganisms play a key role in the global cycling of most elements in our planet. Two very important elements are sulfur and iron. While the former is an essential macronutrient for life (required in concentrations of g per liter for the growth of microorganisms), the second is in many cases considered as a micronutrient (only required in small concentrations). However, thelow bioavailability of iron makes it a common limiting factor for the development of life. Besides their nutritional role, sulfur and iron can be used in what is called *dissimilatory* metabolism. In this case, these elements intervene in the production of energy by the cell, but are not incorporated into cellular biomass.

Chemically, sulfur can be found in nature at different oxidation-reduction states that range from sulfides and organic reduced sulfur (-2) to sulfate (+6). Most of the time, elemental sulfur (0) is an intermediate in the oxidation-reduction of sulfur. The oxidation-reduction cycle of iron is relatively simpler, as this element can be only found in two forms: ferrous (+2) or ferric (+3). Chemical or biological agents contribute to the transformation of sulfur or iron from one state to another. A simplified depiction of the cycles of sulfur and iron is shown in Figure 12.1. In the figure, it can be observed that two other major elements play a key role in the fate of sulfur and iron: carbon and oxygen. Normally, under aerobic conditions, the reduced forms of either sulfur or iron are rapidly oxidized by a purely chemical process and therefore rarely found in sufficient concentrations to allow the development of microorganisms able to utilize these elements as an energy source. Very specific geophysical settings thus allow the development of environmentsenriched in these elements. This was the starting point for a theoretical manuscript published in 1992 by Michel Aragno trying to explain the existence of one of these unique ecosystems near the village of Les Ponts-de-Martel [1]. Michel Aragno presented a series of processes that combine the activity of microorganisms and the geology of the site to explain the elevated concentrations of hydrogen sulfide emerging at the site. Although the origin of the iron spring was not discussed in that manuscript, one can imagine that equivalent processesdealing with iron metabolism are responsible for this phenomenon.

The Laboratory of Microbiology of the University of Neuchâtel has been using these sites as anopen-air laboratory to present to successive generations of students in geomicrobiologythe reality of the processes that shape our planet and interconnect biology and geology. As part of this work, we have measured proxies to microbial activity *in situ*, and more recently, we have completed acomprehensive analysis of the diversity of the bacterial communities in water and sediments in both sites.The results of these analyses as well as a discussion of their meaning are presented in this contribution, which shed additional light on the microbial processes occurring in these microbial ecosystems. We hope that this contribution might prompt further interest from the population to visit and marvel at these two prime examples of microbial activity.

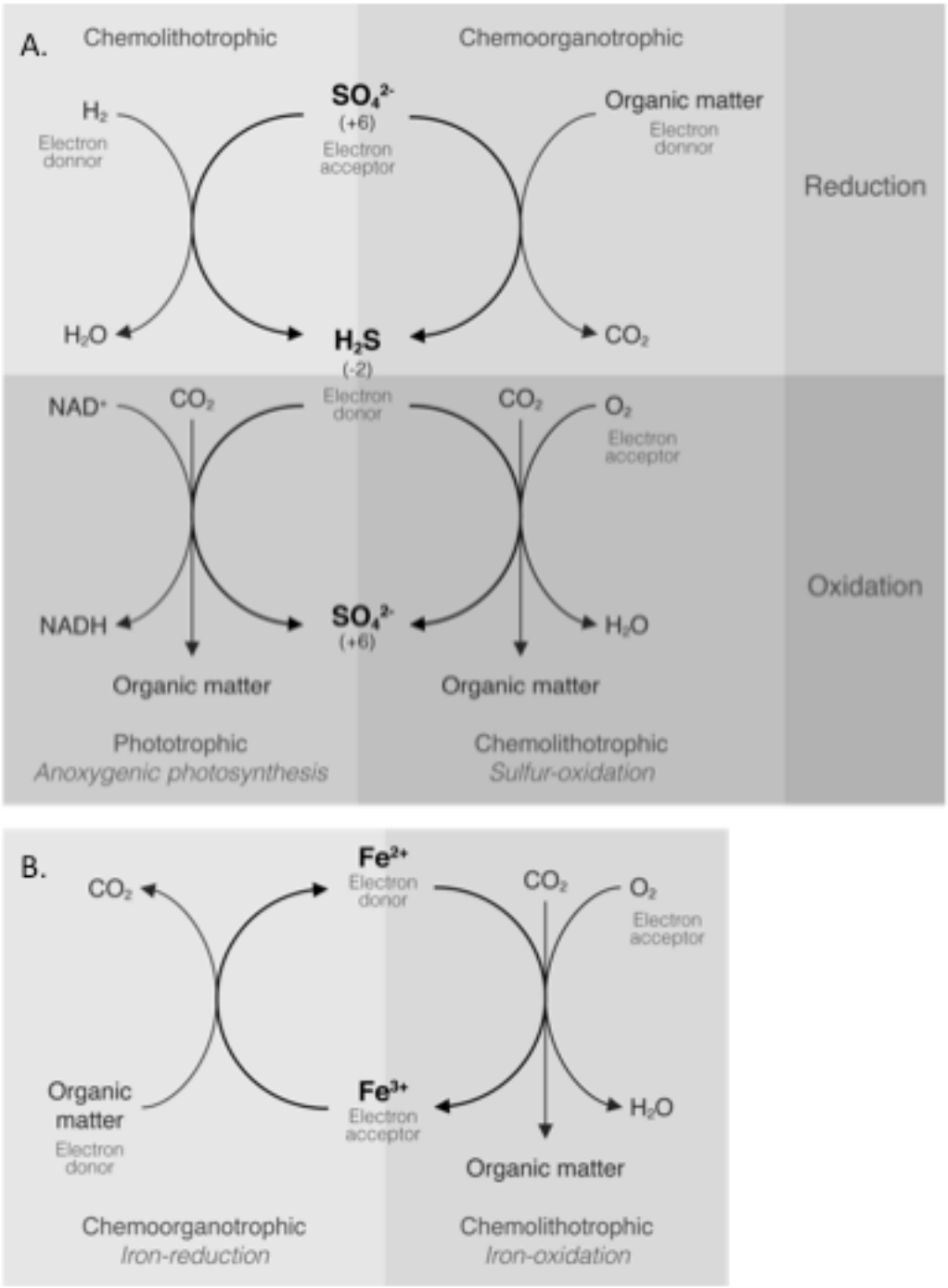


Figure 12.1. Schematic representation of key microbial steps of the sulfur (A) and iron (B) cycles potentially occurring at the sulfur and iron springs of Ponts-de-Martel, Neuchâtel.

12.2. Materials and methods

12.2.1. Description of the site

The sulfur and iron mineral springs are located in the Vallée des Ponts, Neuchâtel (47N, 6.73E) at an altitude between 1000 and 1300m. Concerning the sulfur spring, a characteristic smell, as well as grey deposits are observed. Below the water surface a pink mat is found, covering another black one. This characteristic stratification of the microbial mat layers is also a result of the oxygen gradient formed progressively in depth, depicting the ecology of the site.

As far as the iron spring is concerned, a reddish-brownish mat prevails although greenish mats are also found. In addition to this, foamy raft-like structures with an orange-brownish color are floating at the surface of the water.

12.2.2. Physicochemical measurements made in situ

Temperature, pH, conductivity and oxygen of the mineral water were measured using a HACH HQ40d portable meter (HACH, Loveland CO, USA).Hydrogen sulfide, total iron and ferrous iron were measured using a DR890 HACH colorimeter (HACH, USA).

12.2.3. Microelectrode analysis

For the sulfur and iron springs, samples of a biofilms found at the source of the water spring were taken and brought back to the laboratory. The biofilms were placed in a container with deionized water an oxygen was bubbled in the water to homogenize the concentrations of oxygen in the water-mat interphase.The measurements were made every 500 µm from the underlying water, water-mat interphase and the mat itself. Microsensor probes for oxygen and hydrogen sulfide were purchased from Unisense (Unisense, Denmark). The profiles were collected in a picoammeter PA2000 and analyzed using the SensorTrace Basic 2.0 software (Unisense, Denmark).

12.2.4. Samples for molecular studies

Samples of water, biofilm and microbial mats were collected in 1 L sterile bottles, then stored at -20°C for molecular methods.

12.2.5. DNA extraction, sequencing and analysis

Soil, sediment and biofilm samples were subjected to an indirect DNA extraction as previously described [2]. Water from the samples was also filtered through 0.22 µm membranes to collect biomass. DNA was extracted using the FastDNA Spin Kit for Soil (MP Biomedicals, California), using a modified protocol, in order to ensure that DNA was not only extracted from vegetative cells but also from spores and other cells difficult to lyse [2]. DNA concentration was measured with a Qubit Fluorometer using a dsDNA HS Assay Kit (Invitrogen, California).

In order to study the diversity of the bacterial communities 454 amplicon sequencing of the 16S rRNA gene was performed using the services of Eurofins MWG Operon (Germany). Fragments of approximately 500 bp were retrieved using primers Eub8f (5'-AGAGTTTGATCCTGGCTCAG-3') and Eub519r (5' GTATTACCGCGGCTGCTGG-3'), as previously described [3]. Raw sequence data was analyzed with QIIME [4], using the pipeline for de novo OTU picking and diversity analyses from 454 data suggested in QIIME tutorials. Amplicon sequencing resulted in 8050and 9397sequence reads after quality filtering for the sulfur and iron spring samples respectively. Sequences were de-noised with the split_library.py function implemented in QIIME, and check for chimera using USEARCH version 6.1 with the reference database used by in the version 1.8.0 of QIIME. To the rest of the trimmed and processed sequences, alignment was performed through the RDP website (https://rdp.cme.msu.edu/tutorials/aligner/RDPtutorial_ALIGNER.html)

using Infernal Aligner [5]. OTUs were identified using a threshold of 97% sequence similarity with USEARCH version 6.1. Alpha diversity within the samples was calculated in rarefied subsets sequences to have equal sequence coverage following the tutorial suggested by QIIME for 454-sequencing analysis. The parameters retained for the analysis were Richness, Shannon and Simpson diversity indices, and the percentage of the ratio OTUs/chao1 (coverage).

12.3. Results and discussion

12.3.1. The sulfur spring (NeSul)

The sulfur spring in Les Ponts-de-Martel is an example of a sulfuretum, a peculiar type of environment usually associated with a spring source of sulfide-rich waters (Figure 12.2). These waters are initially anoxic and contain sulfide, but oxygenate rapidly once the stream enters in contact with the atmosphere at the release point. However, the slow pouring of the water allows the establishment of an oxygen gradient towards the bottom of the stream.

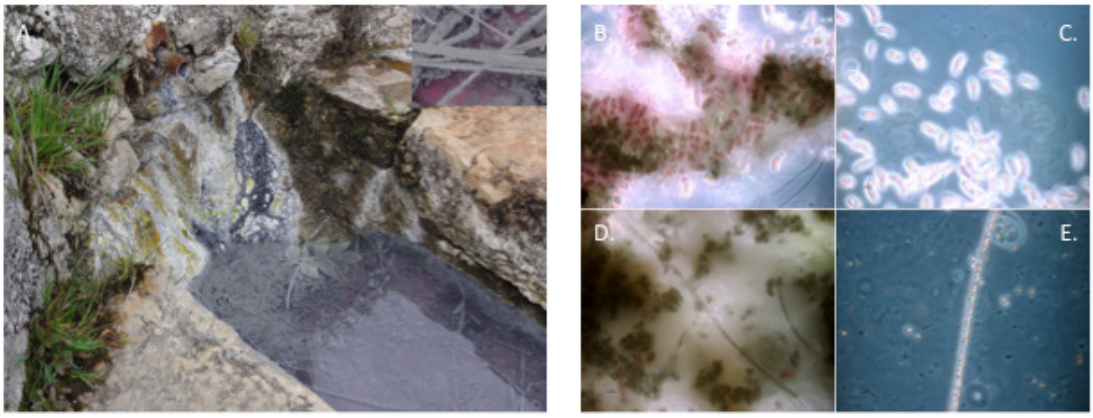


Figure 12.2. Macro and microphotographic images of the sulfur spring of Ponts-de-Martel. A. Source and the surroundings, with a close-up image to the purple mat in the bottom of the source as an insert on the right. B-E. microphotographs of phototrophic (B-C) and colorless (D-E) sulfur-oxidizing bacteria. In both cases the presence of granules of elemental sulfur are clearly distinguishable inside the cells. Images taken during the filed excursion and laboratory work with the Biogeosciences students in May 2011.

Under these conditions three elements co-exist: light, oxygen and sulfide, and both chemical and biological oxidation of sulfide can occur either associated to photosynthesis or to chemosynthesis [6]. Visual signs of the suitable environment for sulfur bacteria can be easily recognized at the site of Les Ponts-de-Martel. First, the characteristic smell of sulfide indicates the availability of this element. Second, white patches of elemental sulfur resulting from the oxidation of sulfide in contact with the air (chemosynthetic oxidation; Figure 12.2A) can be observed covering certain areas of the stream and the surrounding rocks. Likewise, the formation of colored microbial mats is an indication of photosynthetic sulfur oxidation (Figure 12.2A). The microscopic observation of these different elements clearly indicates the presence of both phototrophic (Figure 12.2B-C) and colorless (Figure 12.2D-E) sulfur-oxidizing bacteria, recognizable by the deposition of elemental sulfur granules and distinguishable by the presence of pigments in the case of the former.

Since a wide variety of microorganisms are able to oxidize, reduce or disproportionate sulfur species, the microbial community structure of sulfur-rich habitats is clearly influenced by the prevalent environmental conditions at the site [7]. In the case of Les Ponts-de-Martel, the sulfur spring is characterized by a near to neutral pH (7.83) anda low temperature (15 °C) at the moment of sampling, which in this case corresponded to the beginning of the spring (Table 12.1).

The emerging concentration of hydrogen sulfide is high (266 mg/L, equivalent to 7.8 mM) and maintains low conditions of dissolved oxygen and iron (Table 12.1). From the source to the stream there is a gradient of temperature, conductivity and oxygen (Table 12.2) that establishes the conditions for the development of specific microbial communities.

Sample	NeSul	NeFer
Site	Ponts-de-Martel (NE)	Ponts-de-Martel (NE)
Description	Sulfur mineral spring at Ponts-de-Martel valley, Neuchatel, Switzerland	Iron mineral spring at Ponts-de-Martel valley, Neuchatel, Switzerland
GPS	47N, 6.73E	47N, 6.73E
Temperature	15°C	15°C
pH	7.83	6.94
Conductivity (µs/cm)	700	679
Dissolved O_2 (mg/l)	2.28	1.96
O_2 (µmole) (mg/l)	44	N.D.
H_2S (mg/l)	266	N.D.
Fe^{+2} (mg/l)	0.04	2.03
Total Fe (mg/l)	0.04	2.62

Table 12.1. Site description and parameters measured *in situ*

The evolution of hydrogen sulfide and oxygen were monitored for one of the microbial mats developing at the spring under laboratory conditions (Figure 12.3). This analysis shows how hydrogen sulfide concentrations are high at the bottom of the mat (sulfidogenic conditions), but at the region in which oxygen becomes available, sulfide is rapidly oxidized and disappears from solution. This process occurs at different scales and places at the site.

Distance from the spring	pH	O_2 (mg/L)	conductivity (µS/sec)	Temperature (^oC)
0.5 cm	7.83	1.53	656	12.4
30 cm	7.75	1.46	645	12.5
30cm, deeper	7.09	0.15	680	12.5
100 cm	7.76	0.98	680	13.6

100cm, deeper	7.7	0.44	680	13.6
120 cm	7.68	0.58	680	14.2
250cm	6.9	0.13	692	16.5
Inside the reed, black mat	7.25	0.21	712	19.2
Entrance of stream	8	8.8	543	20.3

Table 12.2. Measurements of pH, O2 concentration, conductivity and temperature at a distance gradient from the sulfur spring to the stream

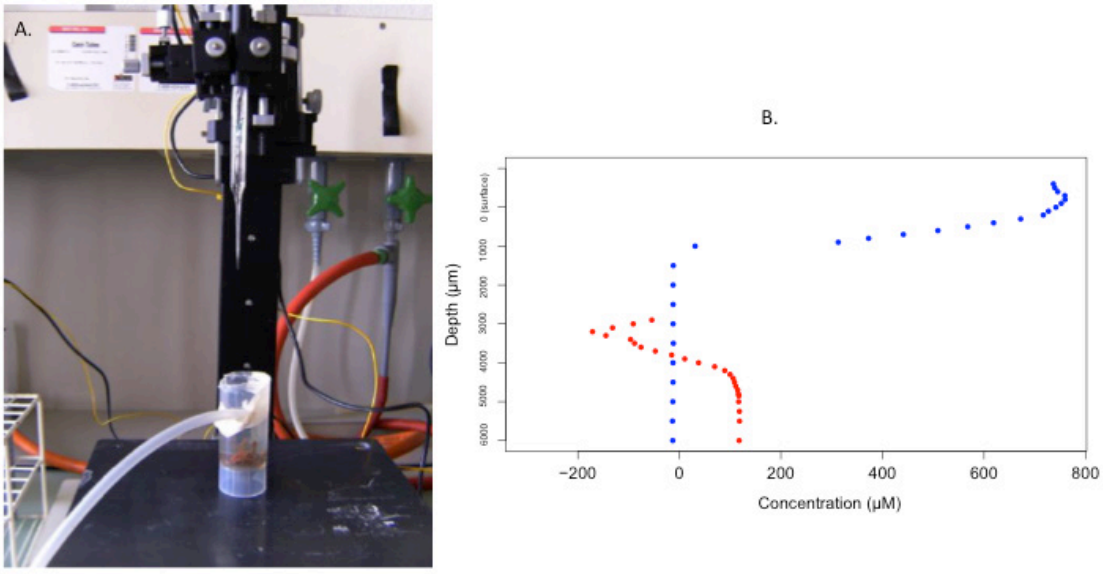


Figure 12.3. Measurements of the concentrations of hydrogen sulfide and oxygen made at a microscale for one of the microbial mats forming at the sulfur spring of Ponts-de-Martel. A. Images of the experimental set-up to of the microelectrode device. B. Gradient of hydrogen sulfide (in red) and oxygen (in blue) measured in the mat. Image A reproduced with the permission of the authors (L. Sauvain and F. Schindelholz).

After the sequencing of the bacterial communities from the sulfur spring, which included a pooled sample of water and different mats, a total of 7954 high-quality sequences were retained for the analysis. The sequences were grouped in 514 operational taxonomic units (OTUs). A large majority of those represented OTUs corresponding to a very small number of sequences (Figure 12.4). Only 15 OTUs were represented by more than 79 sequences, thus corresponding to more than 1% of the community (relative abundance of individualOTUs). The total cumulative frequency of these OTUs represented 67% of the total number of sequences obtained. These OTUs could be identified to different taxonomic levels and their relative abundance is shown in Figure 12.4B.

The most abundant group was identified as belonging to the family Helicobacteraceae, which is known to include genera involved in the sulfur cycle such as *Sulfurimonas* and *Sulforovum* [8]. For some of the culturable representatives of these genera, a facultative mode of energy metabolism (either reducing or oxidizing sulfur compounds) has been shown [9]. One of the few groups that could be identified to species level corresponded to *Sulfuricurvumkujiense*, which is the type species of the genus *Sulfuricurvum*(also belonging to the family Helicobacteraceae), a genus that is of interest because of its

capability to utilize various reduced sulfur compounds such as elemental sulfur, sulfide and thiosulfate as electron donors, particularly in crude oil and oil sands [10].The second most abundant group was identified only to family level as Chlorobiaceae. This family groups numerous anoxygenicphototrophic green sulfur bacteria or GSB [6]. Bacteria related to other GSB groups were also represented by the sequences affiliated to Ignavibacteriaceae, although the cultured representatives of this family are colorless and unable to grow phototrophically [11].

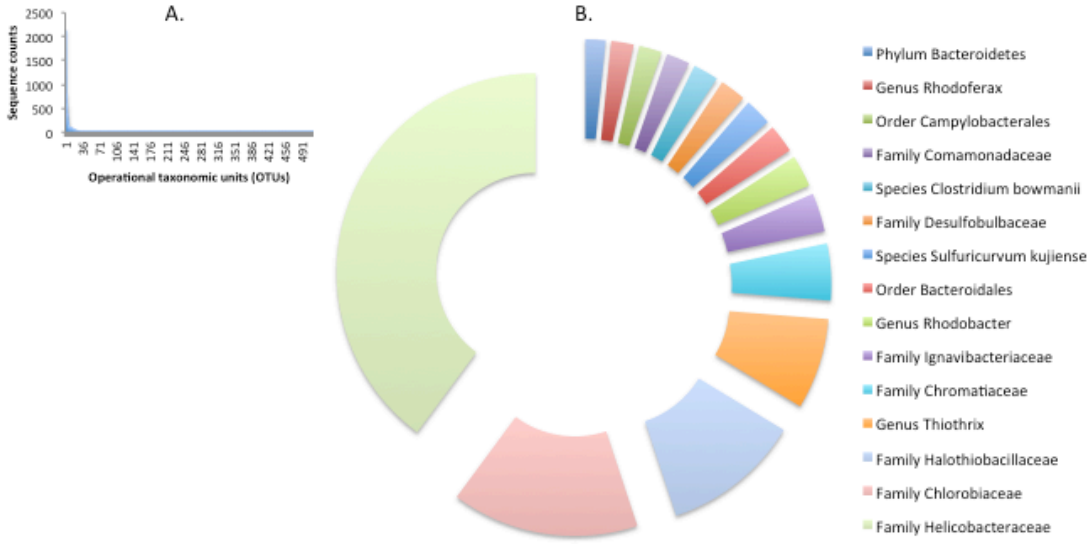


Figure 12.4. Bacterial community composition in the sulfur spring of Ponts-de-Martel analyzed using high-throughput amplicon sequencing of the 16S rRNA gene. A. Distribution of sequence counts for the different taxonomic units (OTUs) obtained from the sample. B. Schematic representation of the relative abundance of the 15 most abundant OTUs (over 1% relative abundance).

The third most abundant OTUs was classified as belonging to the family Halothiobacillaceae, which is so far represented by halophilic obligate aerobic sulfur-oxidizing bacteria, some of which have been isolated from contaminated sites [12]. One of the groups that could be classified to genus level corresponded to *Thiothrix*, which are known gliding filamentouscolorless sulfur-oxidizing bacteria [6] that corresponded to some of the morphotypesobserved microscopically at the site (Figure 12.2D-E) and are characterized by their intracellular deposits of elemental sulfur. A second group of anoxygenicphototrophic sulfur oxidizers was also detected to the family level as Chromatiaceae. This family represents one of the families of purple sulfur bacteria (PSB) belonging to the delta-proteobacteria, which are well knownin freshwater environments [6]. In addition to PSB, other phototrophic sulfur oxidizing bacteria were represented by sequences identified as belonging to the genus *Rhodobacter* [13]. In addition to the diversity of sulfur-oxidizing bacteria identified, sulfate-reducing bacteria were represented by sequences classified as belonging to the family Desulfobulbaceae, which containssulfate-reducers found in diverse environments [14,15].

12.3.2. The iron spring (NeFer)

The transport of ferrous iron into an oxygenated environment leads to the spontaneous reaction with dissolved oxygen at circumneutral pH and to the rapid abiotic precipitation of ferric hydroxides [16]. This gives rise to systems such as the iron spring found near the sulfur spring in Ponts-de-Martel (Figure 12.5A).

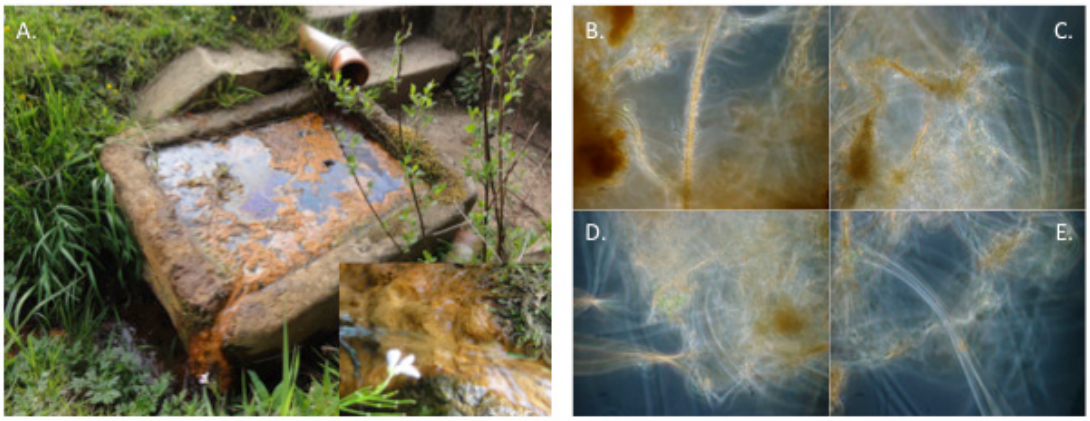


Figure 12.5. Macro and microphotographic images of the iron spring of Ponts-de-Martel. A. Source and the surroundings, with a close-up image to the iron precipitates at the exit of the source as an insert on the right. B-E. microphotographs of the different morphologies of iron hydroxides precipitating in the mat. Images taken during the filed excursion and laboratory work with the Biogeosciences students in May 2011.

In this process, bacteria can act as a passive charged surface that could serve as nucleation center for the formation of the mineral. There are several examples of this type of iron-depositing bacteria in aquatic systems that can bind and precipitate ferric iron as encrusted sheaths (Figure 12.5B-E). Oxidation of ferrous iron allows the growth of bacteria under very specific environmental conditions. For example, the capability of bacteria to oxidize iron at extremely low pH (1-5 to 3.5) is a well-characterized microbial metabolism occurring in acid mine drainages. At circumneutral pH, microaerophilic conditions are required tofavor biological iron reduction over purely abiotic oxidation [16]. Conditions similar to the latter (neutral pH and low oxygen availability) are found in the Les Ponts-de-Martel spring (Table 12.1; Figure 12.6). The measurements of oxygen in microbial mats in this spring show the establishment of microaerophilic conditions in the upper 4.5 mm of the mat.

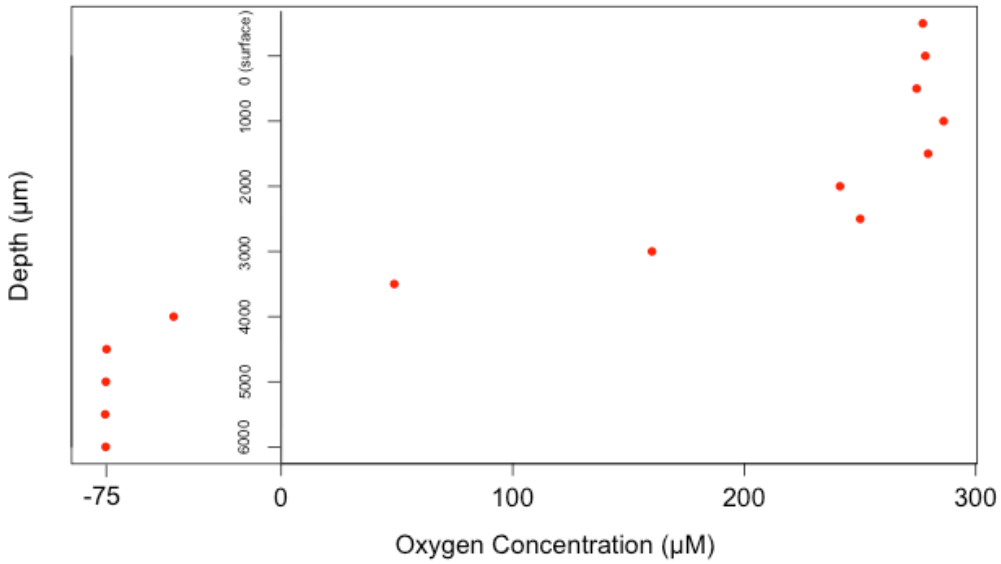


Figure 12.6. Measurements of the concentrations of oxygen made at a microscale for one of the microbial mats forming at the iron spring of Ponts-de-Martel.

The analysis of the microbial communities from water and microbial matsin the iron spring was based in a total of 9320 high-quality sequences, which were grouped into 884 OTUs. As in the case of the sulfur spring, the vast majority of those represented OTUs corresponding to a very small number of sequences (Figure 12.7). Only 20 OTUs were represented by more than 97 sequences, thus corresponding to more than 1% of the community (relative abundance of individual OTUs). The total cumulative frequency of these OTUs represented 56% of the total number of sequences obtained. These OTUs could be identified to different taxonomic levels and their relative abundance is shown in Figure 12.7B.

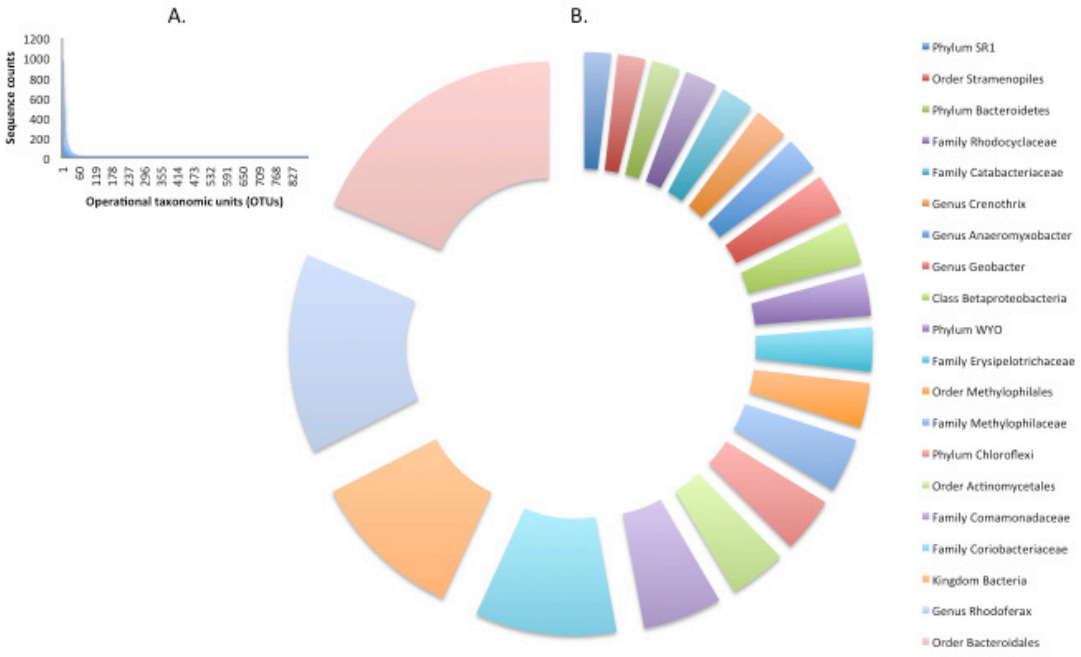


Figure 12.7. Bacterial community composition in the iron spring of Ponts-de-Martel analyzed using high-throughput amplicon sequencing of the 16S rRNA gene. A. Distribution of sequence counts for the different taxonomic units (OTUs) obtained from the sample. B. Schematic representation of the relative abundance of the 20 most abundant OTUs (over 1% relative abundance).

On the one hand, for many of the OTUs identified in the iron spring, a direct relationship with iron metabolism is hard to establish because of the poor phylogenetic resolution obtained (e.g. OTUs identified to Kingdom, Phylum or even Order level). On the other hand, several well-knowngenera of either iron oxidizers such as*Crenothrix*spp [16] and iron reducers such as the genera *Rhodoferax* [17], *Geobacter* [18] and *Anaeromyxobacter* [19] were identified. Other groups, such asthe OTUs related to the family Methylophilaceae have been previously observed in iron-rich waters [20]. Surprisingly, the most abundant group, which was identified to the order level of Bacteriodales, has been recently reported in relationship to the nitrogen cycle in a sulfur-rich system [21], but it has not been reported for its activity in association to the cycling of iron.

12.3.3. Overall conclusion

It has been more than 20 years since Michel Aragno presented a theoretical view to the formation of the sulfur (and potentially iron) springs in Les Ponts-de-Martel. During all this time, these ecosystemshave provided an ideal open-air laboratory in which to study the role of microorganisms on the cycling of two very important elements for life: sulfur and iron. The data presented here, whichcombine observation, *in situ* measurements and molecular analysis, offer a more complete view of the potential processes that explain the functioning of these ecosystems and set the basis for advanced studies focusing for example on the enrichment of specific metabolisms/microorganisms or the demonstration of a coupling between functional activity and the phylogenetic diversity of bacteria found at each one of the sites.

12.4. Acknowledgements

We would like to thank the 2012 and 2015 students of the specialized module in geomicrobiology of the master in Biogeosciences from the Universities of Neuchâtel and Lausanne for their help in collecting field data. We would like to acknowledge funding by the Swiss National Science Foundation under the grant 31003A_152972. We also thank Thomas Junier for his critical reading of the manuscript.

12.5. References

[1] Aragno M. La source sulfureuse de la vallée des Ponts: aspects microbiologiques. Bull Société Neuchâtel Sci Nat 1992.

[2] Wunderlin T, Junier T, Roussel-Delif L, Jeanneret N, Junier P. Stage 0 sporulation gene A as a molecular marker to study diversity of endospore-forming Firmicutes. Environ Microbiol Rep 2013;5:911–24. doi:10.1111/1758-2229.12094.

[3] Li H, Zhang Y, Li D, Xu H, Chen G, Zhang C. Comparisons of different hypervariable regions of rrs genes for fingerprinting of microbial communities in paddy soils. Soil BiolBiochem 2009;41:954–68. doi:10.1016/j.soilbio.2008.10.030.

[4] Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, et al. QIIME allows analysis of high-throughput community sequencing data. Nat Methods 2010;7:335–6. doi:10.1038/nmeth.f.303.

[5] Nawrocki EP, Eddy SR. Query-dependent banding (QDB) for faster RNA similarity searches. PLoSComputBiol 2007;3:e56. doi:10.1371/journal.pcbi.0030056.

[6] Likens GE. Plankton of inland waters.Academic Press; 2010.

[7] Elshahed MS, Senko JM, Najar FZ, Kenton SM, Roe BA, Dewers TA, et al. Bacterial diversity and sulfur cycling in a mesophilic sulfide-rich spring. Appl Environ Microbiol 2003;69:5609–21.

[8] Jorgensen SL, Hannisdal B, Lanzén A, Baumberger T, Flesland K, Fonseca R, et al. Correlating microbial community profiles with geochemical data in highly stratified sediments from the Arctic Mid-Ocean Ridge. Proc Natl AcadSci U S A 2012;109:E2846–55. doi:10.1073/pnas.1207574109.

[9] Yamamoto M, Nakagawa S, Shimamura S, Takai K, Horikoshi K. Molecular characterization of inorganic sulfur-compound metabolism in the deep-sea epsilonproteobacteriumSulfurovum sp. NBC37-1. Environ Microbiol 2010;12:1144–53. doi:10.1111/j.1462-2920.2010.02155.x.

[10] Han C, Kotsyurbenko O, Chertkov O, Held B, Lapidus A, Nolan M, et al. Complete genome sequence of the sulfur compounds oxidizing chemolithoautotroph *Sulfuricurvum kujiense* type strain (YK-1(T)). Stand Genomic Sci 2012;6:94–103. doi:10.4056/sigs.2456004.

[11] Iino T, Mori K, Uchino Y, Nakagawa T, Harayama S, Suzuki K-I. *Ignavibacterium album* gen. nov., sp. nov., a moderately thermophilic anaerobic bacterium isolated from microbial mats at a terrestrial hot spring and proposal of Ignavibacteria classis nov., for a novel lineage at the periphery of green sulfur bacteria. Int J SystEvolMicrobiol 2010;60:1376–82. doi:10.1099/ijs.0.012484-0.

[12] Shi J, Lin H, Yuan X, Zhao Y. Isolation and characterization of a novel sulfuroxidizing chemolithoautotroph *Halothiobacillus* from Pb polluted paddy soil. Afr J Biotechnol 2013;10:4121–6. doi:10.4314/ajb.v10i20.

[13] Friedrich CG, Rother D, Bardischewsky F, Quentmeier A, Fischer J. Oxidation of Reduced Inorganic Sulfur Compounds by Bacteria: Emergence of a Common Mechanism? Appl Environ Microbiol 2001;67:2873–82. doi:10.1128/AEM.67.7.2873-2882.2001.

[14] Asami H, Aida M, Watanabe K. Accelerated sulfur cycle in coastal marine sediment beneath areas of intensive shellfish aquaculture. Appl Environ Microbiol 2005;71:2925–33. doi:10.1128/AEM.71.6.2925-2933.2005.

[15] Green-Saxena A, Feyzullayev A, Hubert CRJ, Kallmeyer J, Krueger M, Sauer P, et al. Active sulfur cycling by diverse mesophilic and thermophilic microorganisms in terrestrial mud volcanoes of Azerbaijan. Environ Microbiol 2012;14:3271–86. doi:10.1111/1462-2920.12015.

[16] Konhauser KO. Diversity of bacterial iron mineralization. Earth-Sci Rev 1998;43:91–121. doi:10.1016/S0012-8252(97)00036-6.

[17] Risso C, Sun J, Zhuang K, Mahadevan R, DeBoy R, Ismail W, et al. Genome-scale comparison and constraint-based metabolic reconstruction of the facultative anaerobic Fe(III)-reducer Rhodoferaxferrireducens. BMC Genomics 2009;10:447. doi:10.1186/1471-2164-10-447.

[18] Childers SE, Ciufo S, Lovley DR. Geobactermetallireducens accesses insoluble Fe(III) oxide by chemotaxis. Nature 2002;416:767–9. doi:10.1038/416767a.

[19] He Q, Sanford RA. Characterization of Fe(III) reduction by chlororespiringAnaeromyxobacterdehalogenans. Appl Environ Microbiol 2003;69:2712–8.

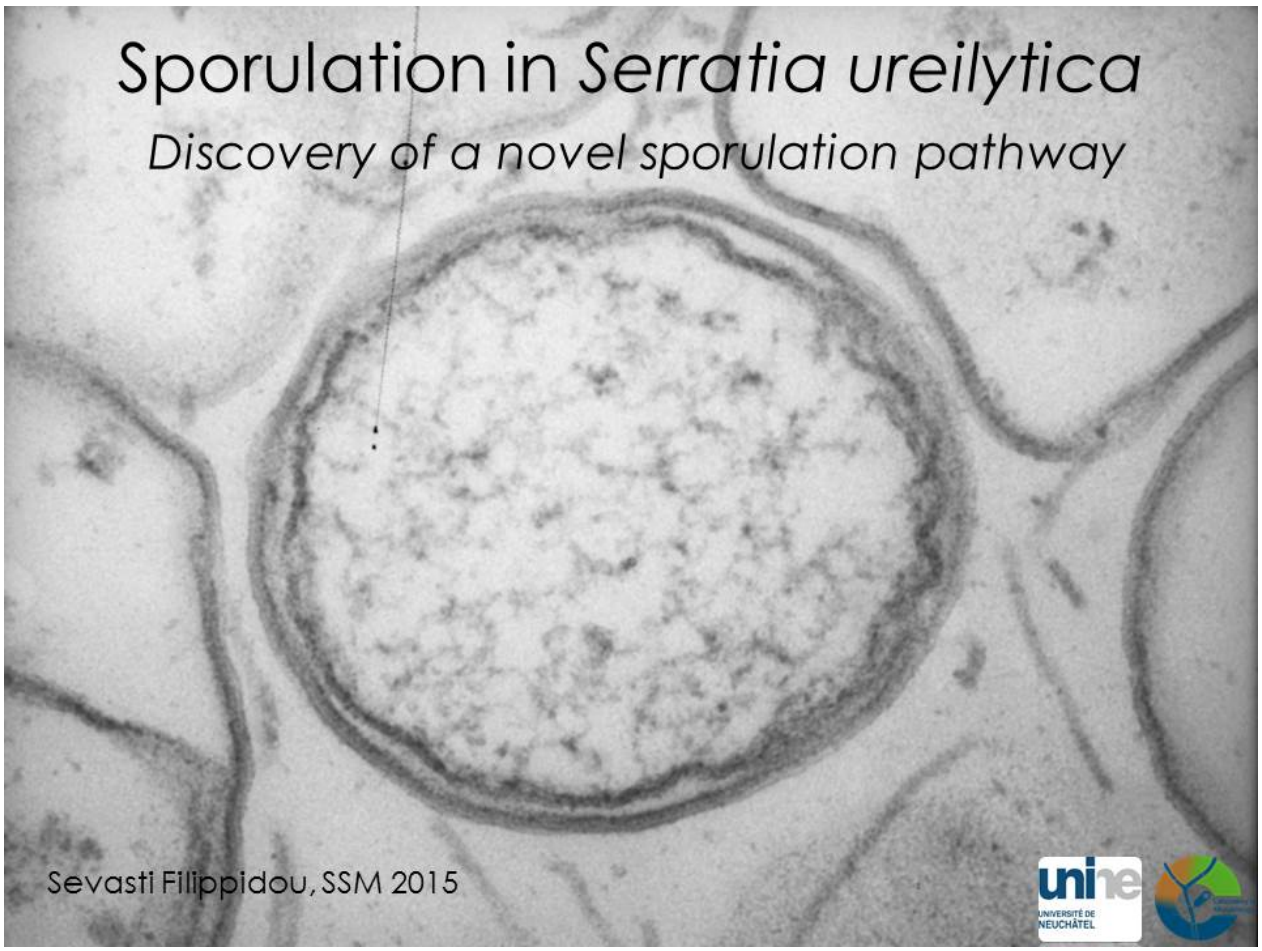
[20] Kato S, Chan C, Itoh T, Ohkuma M. Functional gene analysis of freshwater iron-rich flocs at circumneutral pH and isolation of a stalk-forming microaerophilic iron-oxidizing bacterium. Appl Environ Microbiol 2013;79:5283–90. doi:10.1128/AEM.03840-12.

[21] Llorens-Marès T, Yooseph S, Goll J, Hoffman J, Vila-Costa M, Borrego CM, et al. Connecting biodiversity and potential functional role in modern euxinic environments by microbial metagenomics. ISME J 2015;9:1648–61. doi:10.1038/ismej.2014.254.

Annex 1

Oral presentations at conferences

1. Joint annual meeting SSI - SSHH - SSM - SSTMP, 2015, Lugano, Switzerland.
What triggers sporulation in Serratiaureilytica Lr5/4?
S. Filippidou[1], T. Junier[1], N. Jeanneret[1], P. Junier[1]
Affiliations:
1. Laboratory of Microbiology, Institute of Biology, University of Neuchatel, Emile- Argand 11, CH2000, Neuchatel, Switzerland
Abstract:
Sporulation is known to be triggered by nutrient starvation. Spore or spore-like structures are only found in four bacterial phyla: Actinobacteria, Cyanobacteria, Proteobacteria and Firmicutes. These structures provide resistance to adverse conditions. The ability to form spores is not, however, a widely spread characteristic and it is restricted to only some orders within those phyla. For example, amongst Proteobacteria, solely δ-proteobacteria can produce spore-like fruiting bodies, or so we knew.
A novel γ-proteobacterium, *Serratiaureilytica* str. Lr5/4, was found to produce spores that not only resemble structurally to those produced by endospore-forming Firmicutes, but also provide heat-resistance. Upon the discovery of a new sporulating bacterium, our aim was to investigate what triggers its sporulation, whether its spores are resistant to various extreme conditions and finally, what the genetic imprints of the sporulation triggering factors and the resistance are. In order to address these questions, *S. ureilytica* Lr5/4 vegetative cells were exposed to nutrient starvation, heat and freezing shocks, high salinity, desiccation and UV radiation. In the same environmental conditions, spores were tested for their resistance. These tests were performed in comparison to *Bacillus* sp. for which sporulation triggers and spore resistance are well studied. Moreover, sequencing and annotation of its full genome was performed in order to describe the featured genes devoted to sporulation and resistance. Comparative genomics were used to investigate similar as well as differentiating genomic features of sporulation between *S. ureilytica* Lr5/4 and all sporulating phyla.
The present study demonstrates a novel mechanism for the formation of the described spores of *S. ureilytica* Lr5/4. Moreover, it provides insights in under which environmental conditions survival strategies are expressed.


Sporulation in *Serratia ureilytica*
Discovery of a novel sporulation pathway
Sevasti Filippidou, SSM 2015
unine
UNIVERSITÉ DE NEUCHÂTEL

2. 15th ISME Meeting, 2014, Seoul, South Korea.

Survival strategies meet classical ecological theories: the case of endospore-forming Firmicutes in extreme environments.

S. Filippidou[1], M.Bueche[1], T. Wunderlin[1], T. Junier[1,2], L. Roussel- Delif[1], N. Jeanneret[1], C. Dorador[3], V. Molina[4], A. Ioannidou[5], G. Vargemezis[6], D. R. Johnson[7], P. Junier[1]

Affiliations:

Laboratory of Microbiology, Institute of biology, University of Neuchatel, Neuchatel, Switzerland

Vital-IT group, Swiss Institute of Bioinformatics, Lausanne, Switzerland

Departamento de Acuicultura, Facultad de Recursos del Mar, Universidad de Antofagasta, Antofagasta, Chile

Departamento de Biología, Facultad de Ciencias Naturales y Exactas. Universidad de Playa Ancha, Valparaíso, Chile.

Laboratory of Nuclear Physics, Physics Department, Aristotle University of Thessaloniki, Greece

Laboratory of Applied Geophysics, Aristotle University of Thessaloniki, Greece

Department of Environmental Microbiology, Swiss Federal Institute of Aquatic Science and Technology (Eawag), Dubendorf, Switzerland

Abstract:

Metagenomic data are not only helpful in descriptive studies of diverse ecosystems but have recently been used to reveal relationships between environmental data and patterns of microbial dispersal in extreme environments. For instance, in geothermal environments, recent studies have shown that temperature, and to a less extend pH, dictates prevalence of specific bacterial and archaeal groups. However, this raises the question of how (and if) this trend can be extrapolated to specific microbial groups within the communities.

Endospore-forming bacteria, while in spore state, can survive adverse conditions and disperse over large distances. Thus, they can be found in different environments: from psychrophilic to thermophilic ones, from natural springs to highly contaminated ecosystems and from acidic to alkaline substrates. Based on their ubiquitous distribution, we hypothesize that the co-existence of multiple stressing environmental parameters, rather than temperature or pH alone, favors the relative abundance of endospore-forming bacteria and influences their diversity in extreme environments. These parameters are temperature, pH, salinity, high concentration of heavy or trace metals, UV radiation, and limited carbon and energy sources.

In order to address this hypothesis we have held a worldwide sampling campaign, including 23 sites with diverse environmental characteristics such as high temperature, salinity, depth, altitude, low and high pH, and contamination with heavy and trace metals. In total, 126 samples have been collected and analyzed. DNA was extracted using an extraction method that enables the extraction of DNA from hard to break microbial cell walls or spore coats. 16S rRNA gene (total bacteria) and sporulation transcriptional factor *spo0A* gene (endospore-specific) copy numbers were then estimated by quantitative PCR

(qPCR). Moreover, community composition based on 16S rRNA and *spoOA* gene amplicon sequencing, was analysed in 12 samples representing environments with multiple, single or no stressing factors.

The quantification shows that endospore formers are indeed favored when stressing factors, and more significantly a combination of those, are present in the environment. Our data demonstrate that in sites where two or more environmental stressors are present, endospore-forming bacteria are highly prevalent (ratio median = 26.46%). When there is a single stressor, endospore-forming bacteria are favored but not high in abundance (ratio median = 2.79%), while in non-extreme environments they make up only a small fraction of the microbial community (ratio median = 0.079%).

In comparison to the whole-community, in which bacterial diversity is controlled by a single environmental stressor, (e.g. temperature in geothermal sites or trace metal concentration in polluted sediments), at the scale of endospore- forming bacteria, diversity is influenced by a combination of stressors. These observations suggest that, in extreme environments endospore-formers, behave as a 'system within a system' and are influenced by their environment in a different way than the whole bacterial community.

3. 5th Swiss Microbial Ecology Meeting, 2013, Murten, Switzerland.

Microbial Communities in Geothermal Sites. Are endospore-forming bacteria favoured?

S.Filippidou[1], T. Wunderlin[1], M. Jaussi[1], L. Roussel-Delif[1], N. Jeanneret[1], M. Bueche[1], T. Junier[1,2], P. Junier[1]

Affiliations:

Laboratory of Microbiology, Institute of biology, University of Neuchatel, Neuchatel, Switzerland

Vital-IT group, Swiss Institute of Bioinformatics, Lausanne, Switzerland

Abstract:

Geothermal activity is observed in sites where geothermal energy is produced and stored, due to seismic and volcanic activity. Although these environments host microbial life, conditions for life in general would be characterized as unfavorable: high temperatures, acidic or alkaline pH, lack of nutrients, unfavorable oxygen or hydrogen levels, among others. As a response to these harsh conditions microorganisms have developed various survival strategies. We hypothesized that geothermal sites favor sporulation as a survival strategy and thus endospore-forming bacteria are abundant in microbial communities that inhabit such sites.

In order to evaluate this hypothesis, six geothermal sampling sites were considered: Lirima (Chile), a hot spring (Colombia) and Potamia, Thermia, N. Appolonia and Kanava natural hot springs (Greece). In addition, three geothermal stations in Gross-Shoenenbeck, Bruschal (Germany) and Soultz-sous-forets (France) were included. These sites are all habitats with temperature over 60^{o}C.

DNA was extracted directly from sediments or fluids collected at these sites, using a method developed in our laboratory. A screening of the 16S rRNA gene and of the *spo0A* gene was carried out, in order to verify the presence of bacteria and specifically endospore-forming bacteria. Moreover, a qPCR method, also developed in our laboratory, was performed for both genes and a ratio was calculated in order to determine roughly percentage of endospore-forming bacteria in the whole bacterial communities. Finally, for description of endospore-formers in these communities, 454-pyrosequencing has been performed and comparison between the nine different sites has been made.

Endospore-forming bacteria are an important part in microbial communities that inhabit high-temperature environments. Environmental factors play a crucial role in the selection and abundance of these bacteria. Metagenomic studies based on functional genes could contribute to generate more information on the relationship between these bacteria and their environment.

Microbial Communities in Geothermal Sites. Are Endospore- forming Bacteria Favoured?

S. Filippidou, T. Wunderlin, M. Jaussi, L. Roussel-Delif, N. Jeanneret, M. Bueche, T. Junier, P. Junier

5th SME Meeting, 2013, Murten

Poster Presentations at Conferences

1. SfAM Winter Meeting: Psychrophiles and Extremophiles, 2016, London, UK

Survival off-limits: the key to success of endospore-forming Firmicutes

S. Filippidou[1], M.Bueche[1], T. Wunderlin[1], T. Junier[1,2], L. Roussel- Delif[1], N. Jeanneret[1], C. Dorador[3], V. Molina[4], A. Ioannidou[5], G. Vargemezis[6], D. R. Johnson[7], P. Junier[1]

Affiliations**:**

Laboratory of Microbiology, Institute of biology, University of Neuchatel, Neuchatel, Switzerland

Vital-IT group, Swiss Institute of Bioinformatics, Lausanne, Switzerland

Departamento de Acuicultura, Facultad de Recursos del Mar, Universidad de Antofagasta, Antofagasta, Chile

Departamento de Biología, Facultad de Ciencias Naturales y Exactas. Universidad de Playa Ancha, Valparaíso, Chile.

Laboratory of Nuclear Physics, Physics Department, Aristotle University of Thessaloniki, Greece

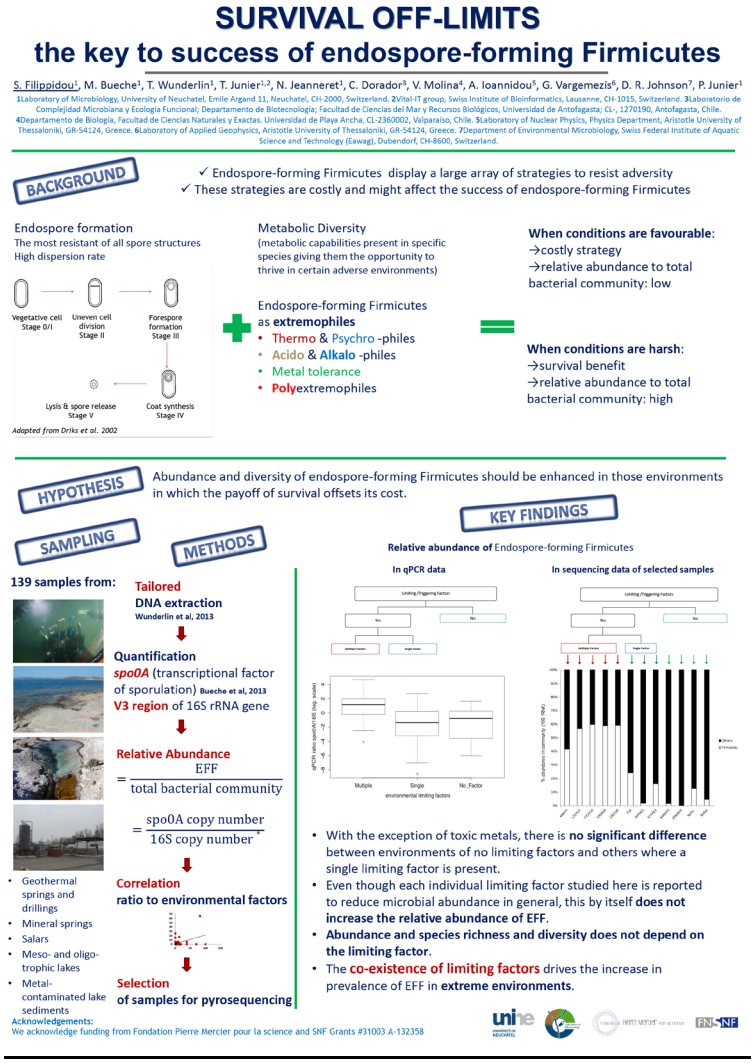
Laboratory of Applied Geophysics, Aristotle University of Thessaloniki, Greece

Department of Environmental Microbiology, Swiss Federal Institute of Aquatic Science and Technology (Eawag), Dubendorf, Switzerland

Abstract:

Aims: Environmental conditions unsuitable for survival and development are the rule rather than the exception in most habitats. Microorganisms have developed various strategies to withstand environmental conditions that limit active growth. A group that displays a large array of strategies to resist adversity is endospore-forming Firmicutes (EFF). These strategies range from the formation of resting states (endospores), to biofilms and metabolic adaptation. These strategies are a costly biological investment and therefore might affect the success of EFF. Therefore, we hypothesize that abundance and diversity of EFF should be enhanced in those environments in which the payoff of survival offsets its cost. Methods and results: In order to address this hypothesis, we have collected 139 samples worldwide from environments characterized by no, single or multiple limiting environmental factors that should enhanced the value of survival. To measure EFF success we quantified their relative abundance compared to total bacteria, using qPCR, and their diversity using 16S rRNA and *spo0A* (sporulation transcriptional factor) pyrosequencing. The quantification showed that only the co-existence of multiple limiting environmental factors increases the relative abundance of EFF. This likely reflects the diversity of survival strategies deployed by EFF. This is supported by community composition analyses based on 16S rRNA amplicon sequencing from 12 samples representing environments with no, single or multiple limiting factors. Diversity reflected the high EFF functional diversity as an active response to limiting conditions, while the plasticity of survival strategies such as sporulation was also observed.Conclusions: These results suggest that because of their enhanced survival, EFF display a unique distribution pattern that might be replicated by other microorganisms with diversified survival

adaptations.Significance of study: This is the first study showing that EFF prevail in environments with multiple limiting factors.



2. FEMS, 2015, Maasticht, Netherlands

When microbial survival strategies never stop to wonder: sporulation outside Firmicutes.

S. Filippidou[1], T. Junier[1,2], T. Wunderlin[1], M. Bueche[1], W.M. Kooli[1], N. Jeanneret[1], L. Roussel-Delif[1], V. Molina[4], R. Lienhard[4], J. Spangerberg[6], S Johnson[5], P.S. Chain[5], C. Dorador[3], P. Junier[1].

Affiliations:

1. Laboratory of Microbiology, Institute of Biology, University of Neuchatel, Emile- Argand 11, CH2000, Neuchatel, Switzerland

2. Vital-IT group, Swiss Institute of Bioinformatics, Lausanne, Switzerland

3. Departamento de Acuicultura, Facultad de Recursos del Mar, Universidad de Antofagasta, Antofagasta, Chile

4. ADMED Laboratoires, site de La Chaux-de-Fonds, Chasseral 20, CH2300, La Chaux-de-Fonds, Switzerland

5. Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, 87545, USA

6. Institute of Earth Surface Dynamics, University of Lausanne, Building Géopolis, CH-1015 Lausanne, Switzerland.

Abstract:

Background: Spore or spore-like structures are only found in four bacterial phyla: Actinobacteria, Cyanobacteria, Proteobacteria and Firmicutes. These structures provide resistance to adverse conditions. The ability to form spores is not, however, a widely spread characteristic and it is restricted to only some orders within those phyla. For example amongst Proteobacteria, solely δ-proteobacteria can produce spore-like fruiting bodies, or so we knew. A novel γ-proteobacterium, *Serratiaureilytica* str. Lr5/4, was found to produce spores that not only resemble structurally to those produced by endospore-forming Firmicutes, but also provide heat-resistance. Objectives: The aim of this study is to describe this novel strain and its Firmicute-like spores and to reveal the molecular pathway of this sporulation procedure in comparison to those that are already known. Methods: Physiological, biochemical, carbon source assimilation and antibiotic resistance tests were performed. Morphology of vegetative cells and spores was described by phase contrast microscopy, SEM, and TEM. Moreover, spores of Lr5/4 were revived after heat-shock tests and shown to contain dipicolinic acid (DPA). These two characteristics were so far unique to the heat-resistant endospores found in Firmicutes. Sequencing and annotation of its full genome has been performed in order to reveal the relationship of spore formation in Lr5/4 to other known sporulation pathways.Conclusions: It has been previously proposed that the properties of spore formation in non-sporulating species were due to molecular gene transfer. However, the present study rejects this scenario and demonstrates a novel mechanism for the formation of the described spores of *S. ureilytica*.

Framework for a novel sporulation pathway

When microbial survival strategies never stop to wonder: sporulation outside Firmicutes

Sevasti Filippidou[1], Thomas Junier[1], Andrej Al-Dourobi[1], Nicole Jeanneret[1], Ludovic Roussel Delif[1], Tina Wunderlin[1], Matthieu Bueche[1], Wafa Kooli[1], Veronica Molina[2], Christina Dorador[3], Reto Lienhard[4], Shannon Johnson[5], Patrick S. Chain[5], Pilar Junier[1]

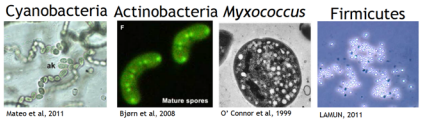
1. Laboratory of Microbiology, Institute of Biology, University of Neuchatel, Emile- Argand 11, CH2000, Neuchatel, Switzerland. 2. Departamento de Biología, Facultad de Ciencias Naturales y Exactas. Universidad de Playa Ancha, Valparaíso, Chile. 3. Departamento de Acuicultura, Facultad de Recursos del Mar, Universidad de Antofagasta, Antofagasta, Chile. 4. ADMED Laboratoires, site de La Chaux-de-Fonds, Chasseral 20, CH2300, La Chaux-de-Fonds, Switzerland. 5. Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, 87545, USA

Key Findings

- ✓ A *Serratia ureilytica* strain has been found to form **heat-resistant spores**
- ✓ A novel sporulation pathway is proposed

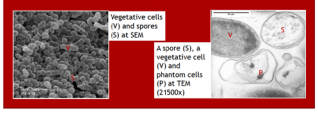
Four bacterial taxa are known to sporulate

Cyanobacteria | Actinobacteria | *Myxococcus* | Firmicutes

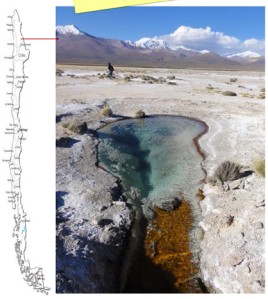


Mateo et al, 2011 | Bjørn et al, 2008 | O' Connor et al, 1999 | LAMUN, 2011

A fifth one was discovered



The story of a discovery

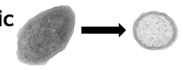


Serratia Lr5/4 was isolated from a **microbial mat** in the **geothermal springs** of Lirima, Chile

	Firmicutes	*Serratia*	Lr5/4
Spores	✓	×	✓
DPA	✓	×	✓
spo0A	✓	×	×
Heat Resistance	✓	×	✓

Strain Lr5/4 shows *Serratia* AND firmicute-like properties

From microscopic observation...



Spores are morphologically distinct from vegetative cells

...to physiology...

sporulation triggering factors

	Firmicutes	Cyanobacteria	Actinobacteria	*Myxococcus*	Lr5/4
Starvation (C)	✓	✓	✓	✓	✓
pH	✓	N.D.	×	×	×
UV	✓	N.D.	×	×	✓
Temperature	✓	low	×	×	✓
Salinity Metal concentration	✓	×	×	×	×
Pressure	✓	×	×	×	×
Chemicals	✓	N.D.	N.D.	✓	×

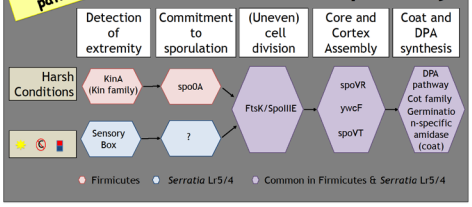
resistance of spores

Factors	*Bacillus subtilis*	Lr5/4
UV	✓ (at 120 min)	✓ (at 240 min)
Temperature	✓	✓
Desiccation	✓	✓

✓ Spores of Lr5/4 are as resistant as those of Firmicutes
✓ Firmicutes and Lr5/4 respond to a larger spectrum of environmental challenges

A firmicute-like sporulation pathway

...and to the molecular pathway!



Acknowledgements
We acknowledge funding from the following institutions

3. Joint annual meeting SSI - SSHH - SSM - SSTMP, 2014, Fribourg, Switzerland

Breaking bacterial stereotypes: the case of a spore-former member of a non-sporulating genus.

S. Filippidou[1], N. Jeanneret[1], L. Roussel-Delif[1], T. Wunderlin[1], T. Junier[1,2], V. Molina[4], C. Dorador[3], R. Lienhard[4], P. Junier[1]. **1st place at Best Poster Award**.

Affiliations:

1. Laboratory of Microbiology, Institute of Biology, University of Neuchatel, Emile- Argand 11, CH2000, Neuchatel, Switzerland

2. Vital-IT group, Swiss Institute of Bioinformatics, Lausanne, Switzerland

3. Departamento de Acuicultura, Facultad de Recursos del Mar, Universidad de Antofagasta, Antofagasta, Chile

4. ADMED Laboratoires, site de La Chaux-de-Fonds, Chasseral 20, CH2300, La Chaux-de-Fonds, Switzerland

Abstract

Sporulation in Firmicutes is a genetically complex procedure, controlled by at least 60 genes, that results in a sophisticated structure, enabling bacteria to survive unfavorable conditions. Other bacterial groups (eg. actinobacteria) can form spores that enable their dispersal but not their survival. Although transfer of sporulation genes from *Bacillus* spp. to other species has been reported, the recipient species cannot enter sporulation because multiple genes are required. *Serratia* spp., mostly known as nosocomial pathogen, has a few representatives isolated in environmental samples. Such an isolate has been recently reported to sporulate, as a result of possible gene transfer from *Bacillus* spp. isolated in the same environment (Ajithkumar *et al.,* 2003). A spore-forming strain (Lr5/4T) was isolated from a microbial mat in the geothermal springs of Lirima, Chile, during sampling for endospore-forming Firmicutes in stagnant ponds, springs and hot streams. 61 strains were isolated, 66% of which were identified as aerobic endospore-forming Firmicutes. However, strain Lr5/4 clearly belongs to *Serratia* genus, on the basis of 16S rRNA gene sequence similarity, MALDI-TOF, and DNA/DNA hybridization, making it the second sporulating strain ever reported from a non-sporulating genus. Physiological, biochemical, carbon source assimilation and antibiotic resistance tests were performed. Morphology of vegetative cells and spores was described by phase contrast microscopy, SEM, and TEM. Moreover, spores of Lr5/4 were revived after heat-shock tests and contained dipicolinic acid, two characteristics unique to Firmicute endospores. The presence of common sporulation genes was also investigated. *Serratia ureilytica* str.Lr5/4 exhibits Firmicute-endospore properties, while the sporulation pathway genes are absent. This excludes a potential gene transfer from bacilli, and suggests a novel type of sporulation that enables both dispersal and survival. Sequencing and annotation of its full genome will provide information not only on potential new sporulation pathways but will also give insights on the diversity of endospore-formers in the environment.

BREAKING STEREOTYPES

The case of a γ -proteobacterium that sporulates

S.Filippidou[1], N.Jeanneret[1], L.Roussel-Delif[1], T.Wunderlin[1], T.Junier[1], V.Molina[2], C.Dorador[3], R.Lienhard[4], P.Junier[1]

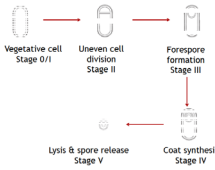
1. Laboratory of Microbiology, Institute of Biology, University of Neuchatel, Emile- Argand 11, CH2000, Neuchatel, Switzerland. 2. Departamento de Biología, Facultad de Ciencias Naturales y Exactas. Universidad de Playa Ancha, Valparaíso, Chile. 3. Departamento de Acuicultura, Facultad de Recursos del Mar, Universidad de Antofagasta, Antofagasta, Chile. 4. ADMED Laboratoires, site de La Chaux-de-Fonds, Chasseral 20, CH2300, La Chaux-de-Fonds, Switzerland

Key Findings

- ✓ A *Serratia ureilytica* strain, isolated from Lirima hot springs, Chile, has been found to form **heat-resistant spores.**
- ✓ The sporulation **pathway** remains **unknown**

What is known...

Firmicutes
Endospore formation



Adapted from Driks et al. 2002

>60 genes required to produce an endospore

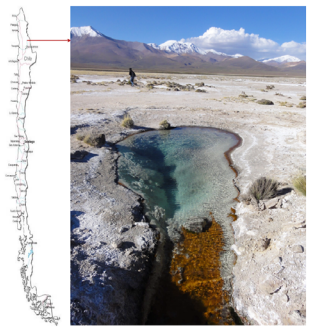
Dispersal and **heat resistance** due to dipicolinic acid **(DPA)**

Vs

Serratia

- Mostly known as **nosocomial pathogen**
- Range of temperature **tolerence: 10-45°C**
- Few species isolated in natural environments

1. Strain Lr5/4 was isolated from a **microbial mat** in the **geothermal springs** of Lirima, Chile

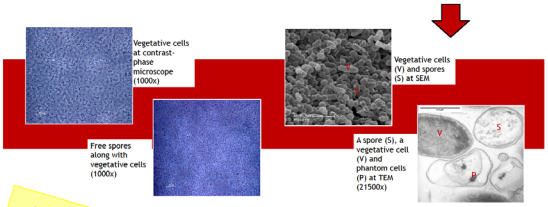


....and how this discovery breaks the stereotypes!

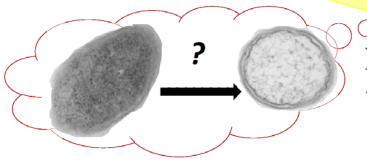
2. Strain Lr5/4 shows *Serratia* AND firmicute-like properties

	Firmicutes	*Serratia*	Lr5/4
Spores	✓	×	✓
DPA	✓	×	✓
spo0A	✓	×	×
Heat Resistance	✓	×	✓

3. **Spores** of *S. ureilytica* Lr5/4 have been observed by microscopy.



What is NOT known...



The sporulation pathway remains unknown.
The **ongoing genome sequencing** of the strain will reveal the mechanism of spore-formation in this unique strain.

Acknowledgements:
We would like to sincerely thank Mrs Wafa Kooli, Mrs Michele Vlimant for SEM and TEM and Dr. Matthieu Bueche for DPA analysis. We acknowledge funding from Fondation Pierre Mercier pour la science and SNF Grandt#31003 A-132358

4. 14th ISME Meeting, 2012, Copenhagen, Denmark

Diversity of Aerobic Endospore-forming Bacteria in Geothermal Sites in Greece.

S. Filippidou[1], N. Jeanneret[1], L. Roussel-Delif[1], T. Wunderlin[1], T. Junier[1,2], P. Junier[1]

Affiliations:

1. Laboratory of Microbiology, Institute of Biology, University of Neuchatel, Emile- Argand 11, CH2000, Neuchatel, Switzerland

2. Vital-IT group, Swiss Institute of Bioinformatics, Lausanne, Switzerland

Abstract

The Greek territory is well recognized for its geothermal activity. Due to seismic and volcanic activity, natural hot springs occur frequently, distributed throughout the continental and island territory. Such environments constitute a diverse assemblage of microbial ecosystems. Temperature variations, from 15° C to 80° C, pH range from 3 to 9 and mineral composition at diverse natural sources have been described. Our aim was to study the distribution of aerobic endospore - forming bacteria at these sites, for which no data was available before. Sampling was conducted in12 natural hot springs in Northern Greece and on the volcanic island of Milos (south-west). Sampling sites were particularly diverse ranging from sites with luxurious vegetation and rivers to caves, inactive volcanoes and underwater geothermal sources. Moreover, human activity was observed in some of these sites, as they are used for recreational thermal baths. In total, 52 sediment, sand, soil and water samples were collected, of which 16 were retrieved from marine environments (underwater hot springs).

In order to study the microbial diversity and distribution by site, the methodological approaches were: Enrichment and isolation of bacteria on marine and nutrient agar; microscopic observation; DNA extraction; PCR amplification of 16S rRNA gene and a molecular marker for endospore-formation(the *spo0A* gene); sequencing and phylogeny.

In total, 80 strains have been isolated, from which 60 were identified as endospore-forming bacteria. In all 60 isolates, the *spo0A* gene was amplified and its sequence confirmed. These strains belong to Bacillus, Geobacillus, Anoxybacillus, Aneurinibacillus, Lysinibacillus, genera. DNA has been extracted directly from volcanic fumaroles and sediments, but no cultured strains could be obtained from these sites so far.

Phylogeny revealed the clustering of strains from distant sites, showing a temperature-based distribution pattern. Isolates from a high temperature drilling (85°C, Zefyria, Milos) clustered together with strains from high altitude natural springs (Thermia, Drama), where the only common characteristic was high temperature (above 60° C). Moreover, 10 Geobacilli strains, from distant natural sources with diverse environmental characteristics and temperature range between 35-40°C, also clustered together. Strains isolated from a marine environment (high salinity cave, Agia Paraskevi, Chalkidiki) cluster with an isolate from a thick mud lake (Pikrolimni, Kilkis). Temperature on both sites was between 25 and 35° C. Human activity does not seem to be a contributing factor concerning distribution.

Overall, additional information from other molecular techniques, such as DGGE and full genome sequencing of our isolates, could provide additional information to understand the factors determining the biogeographical distribution of endospore-formers.

Diversity of Aerobic Endospore-forming Bacteria in Geothermal Sites in Greece

S. Filippidou, N. Jeanneret, L. Roussel-Delif, T. Wunderlin, T. Junier, P. Junier
Laboratory of Microbiology, Univesity of Neuchatel, Switzerland
Contact: sevasti.filippidou@unine.ch

Introduction

Natural hot springs are distributed throughout the continental and island territory of Greece. Such environments constitute a diverse assemblage of microbial ecosystems with temperature variations, from 15o C to 80o C, pH range from 3 to 9 and diverse mineral composition. Our **aim** was to study the distribution of aerobic endospore - forming bacteria at these sites, for which no data was available before.

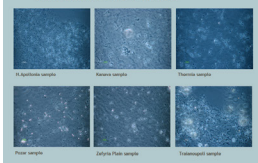
Sampling Points

Methods

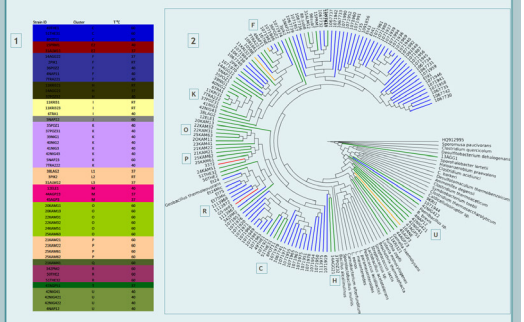
Strain Isolation:
Enrichment, isolation on Nutrient Agar or Marine Agar, microscopic observation, DNA extraction, PCR amplification of the 16S rRNA and *spo0A* genes, sequencing, phylogeny.

DNA Extraction:
Removing biomass from sediment using Ultra-Thurrax and DNA extraction and amplification of 16S rRNA and *spo0A* genes

Enrichment cultures, 72h



Phylogeny reveals a temperature-based distribution pattern



1. Table of Clusters, showing the strain ID code and their culture temperature. Strain names with different letter codes come from different sites.
2. A Maximum-likelihood tree of the 16S rRNA gene sequence from the isolates. Greek isolates in green, Chilean in blue, German in red, Swiss in yellow. (16S rRNA gene)

Conclusions

- Sites with geothermal activity, characterized by extreme conditions (high temperature, salinity, extreme pH values, oligotrophic) harbor a high diversity of endospore-forming bacteria.
- A temperature-based distribution pattern is shown: Isolates from diverse environments (marine, lake, high altitude, different pH range etc) cluster together where the only common characteristic is temperature.
- The majority of Greek isolates from different geothermal sites cluster together. However, in some cases, Greek strains are more closely related to strains from geothermal sites in Germany, Switzerland and Chile.
- Well- studied but also newly described species are present at all these sites.
- Human activity does not seem to be a contributing factor concerning distribution.

Acknowledgments:

We would like to thank the Swiss Government and the Fondation Pierre Mercier pour la Science for the financial support for S. Filippidou.

5. Joint annual meeting SSI - SSHH - SSM - SSTMP, 2012, St-Gallen, Switzerland

Diversity and Distribution of aerobic Endospore-forming Bacteria at Geothermal Sites in Greece.

S. Filippidou[1], T. Wunderlin[1], N. Jeanneret[1], L. Roussel-Delif[1], T. Junier[1,2], P. Junier[1]

Affiliations:

1. Laboratory of Microbiology, Institute of Biology, University of Neuchatel, Emile- Argand 11, CH2000, Neuchatel, Switzerland

2. Vital-IT group, Swiss Institute of Bioinformatics, Lausanne, Switzerland

Abstract

This study aims at revealing the distribution patterns of endospore-forming bacteria at geothermal sites in Greece. The knowledge of their distribution and diversity could be the base for more detail studies into the metabolic diversity and function of these microorganisms. Furthermore, correlation with samples from geothermal sites worldwide is also studied.

Sampling was held at 12 sites of geothermal activity in the continental and island territory of Greece. The methodological approaches consisted on enrichment and isolation of bacteria, microscopic observation, DNA extraction, PCR amplification of the 16S rRNA and *spo0A* genes, sequencing and phylogeny of the isolates [1,2,3]. In addition, direct sediment DNA extraction and amplification of 16S rRNA and *spo0A* genes were also carried out [4].

In total, 52 soil, sand, sediment and water samples were collected. 80 strains have been isolated, from which 60 were identified as spore-forming bacteria (75%). The *spo0A* gene was amplified in all 60 isolates. Based on sequence similarity, these strains are related to the genus Bacillus, Geobacillus, Anoxybacillus, Aneurinibacillus, Lysinibacillus, Kurthia and Exiguobacterium. DNA was also extracted directly from sediments. A successful 16S rRNA gene amplification was obtained for 46 samples (88%). From those, the *spo0A* gene was detected in 18 samples, if appropriate extraction and amplification methods are applied.

Sites with geothermal activity, characterized by extreme conditions (high temperature, salinity, pH, luck of nutrients) harbor a wide diversity of endospore-forming bacteria. Additionally, phylogeny revealed that isolates from different geothermal sites could cluster together, as well as with strains isolated from geothermal sites in Germany and Chile. Well- studied but also newly described species are present at all these sites.

Diversity and Distribution of Aerobic Endospore-forming Bacteria from Geothermal Sites in Greece

S. Filippidou [1], T. Wunderlin [1], N. Jeanneret [1], L. Roussel-Delif [1], T. Junier [1], P. Junier [1]
[1]Laboratory of Microbiology, Institute of Biology, University of Neuchatel, Switzerland
Contact: sevasti.filippidou@unine.ch

Aims

- Reveal the distribution pattern of endospore-forming bacteria from geothermal sites in Greece
- Compare the results with endospore-formers from other natural and exploited geothermal sites

Sampling Points

11 continental sites and 1 island territory

Methods

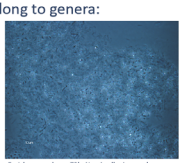
Strain Isolation:
Enrichment, isolation on Nutrient Agar or Marine Agar, microscopic observation, DNA extraction, PCR amplification of the 16S rRNA and *spo0A* genes, sequencing, phylogeny [1,2,3]

DNA Extraction:
Removing biomass from sediment using Ultra-Thurrax and DNA extraction and amplification of 16S rRNA and *spo0A* genes

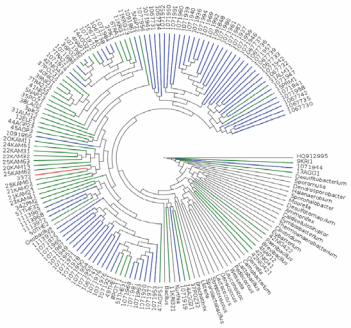
- *spo0A* gene can be detected, if appropriate extraction and amplification methods are applied.

Results

- Strains isolated from these sites belong to genera:
  - *Bacillus*
  - *Geobacillus*
  - *Anoxybacillus*
  - *Aneurinibacillus*
  - *Lysinibacillus*
  - *Kurthia*
  - *Exiguobacterium*


Enrichment culture 72h, Nea Apollonia sample

- 75% of all aerobic isolates (60 out of 80) are endospore-formers
- Direct extraction: 16S rRNA gene amplification in 88% (46 out of 52) of samples and *spo0A* gene amplification in 57,6% of samples (30 out of 52)



Phylogeny of strains: A Maximum-likelihood tree of the 16S rRNA gene sequence from the isolates. Greek isolates in green, Chilean in blue, German in red, Swiss in yellow. (16S rRNA gene)

Conclusions

✧ Sites with geothermal activity, characterized by extreme conditions (high temperature, salinity, extreme pH values, oligotrophic) harbor a high diversity of endospore-forming bacteria.
✧ The majority of Greek isolates from different geothermal sites cluster together. However, in some cases, Greek strains are more closely related to strains from geothermal sites in Germany, Switzerland and Chile.
✧ Well- studied but also newly described species are present at all these sites.

References:
1. Bonjour F. et al. Microb Ecol 1988: 16: 331.
2. Muyzer G. et al. AEM 1993: 59: 695.
3. Brosius J. et al. Proc. Natl. Acad. Sci. U S A. 1978: 75: 4801

Acknowledgements:
We would like to thank the Swiss Government and the Fondation Pierre Mercier pour la Science for the financial support for S. Filippidou.

Sevasti Filippidou

Laboratory of Microbiology, University of Neuchatel
Emile Argand 11• Neuchatel, Switzerland CH2000
Tel: 0041327182264 • e-mail : sevasti.filippidou@unine.ch
• web page: http://www2.unine.ch/lamun/lang/en_US/collaborateurs_actuels_2/filippidou

Research Interests

I am mostly interested in genetics, ecology and distribution of microorganisms in various environments. What mostly attracts me is the understanding of the environmental effect on microbial diversity and metabolism. Moreover, the development of the required tools to study these mechanisms is also in the centre of my interests. My thesis subject concerns the study of diversity and metabolism of endospore-forming bacteria in extreme environments and it is part of the project on the diversity of endospore-formers in natural communities developed at the Laboratory of Microbiology, University of Neuchatel.

Education

PhD in Sciences **May 2012-April 2016**

PhD thesis: Sporulation and metabolism of endospore-forming Firmicutes under conditions limiting for growth and survival.
Laboratory of Microbiology, UNINE, under the supervision of Prof. Pilar Junier.

Master in Molecular Genetics **September 2009-July 2011**

Master thesis: Phenotype and Genotype Characterization of *Enterococcus* spp. from environmental samples and correlation with enterococci spp. from clinical samples
Environmental Microbiology Unit, University of Patras, under the supervision of Prof. Apostolos Vantarakis

Diploma in Molecular Biology and Genetics **September 2003-September 2007**

Graduation thesis: Isolation and identification of enteroviruses, adenoviruses and Hepatitis A virus in the inlet and outlet of the wastewater treatment plant of the municipality of Alexandroupolis
Laboratory of Hygiene and Public Health, Democritus University of Thrace, under the supervision of Prof. Apostolos Vantarakis

Research experience

Swiss Confederation Fellowship **September 2011-May 2012**

Diversity and Distribution of Spore-forming Bacteria from Geothermal Sites in Greece. Laboratory of Microbiology, UNINE, under the supervision of Prof. Pilar Junier.

Research Project **May 2008-December 2008**

Study and Evaluation of Methodologies concerning Early Detection of HPV Infection, for Successful Obligatory Vaccination. 2nd Department of Gynaecology and Obstetrics, University General Hospital "Ippokration", Medical School, Aristotle University of Thessaloniki I, under the supervision of Prof. Aristotelis Loufopoulos

Scholarships and Awards

Travel Grant **March 2014**

Scholarship of UNINE doctoral program, awarded for the participation and oral presentation at ISME 15 Meeting, Seoul, South Korea.

"Egalité" Grant **March 2013**

Scholarship of UNINE for equality of women in science, awarded for sampling campaign in the volcanic arc of southern Aegean in August-September 2013

Swiss Confederation Fellowship **April 2011**

Swiss Confederation Government Scholarship for research (one-year project). Awarded for the research proposal entitled Diversity and Distribution of Spore forming Bacteria at Geothermal Sites in Greece.

Master Degree awarded with grade "Excellent" **November 2011**

Publications

Peer-reviewed publications

1. Filippidou S., Jaussi M., Junier T., Wunderlin T., Jeanneret N., Palmieri F., Palmieri I., Roussel-Delif L., Vieth-Hillebrand A., Vetter A., Chain P.S., Regenspurg S., Junier P. *Anoxybacillus geothermalis* sp. nov., a facultative anaerobic endospore-forming bacterium isolated from mineral deposits in a geothermal station, accepted in IJSEM

2. Ganesan S. & FilippidouS., Junier T., Muñoz Rufatt P., Jeanneret N., Wunderlin T., Sieber N., Dorador C Junier P. Manganese-II oxidation and Copper-II resistance in endospore forming Firmicutes isolated from uncontaminated environmental sites, AIMS Environmental Science, 2016 Mar; 3(2): 220238. doi:10.3934/environsci.2016.2.220

3. Simon, A., Bindschedler, S., Job, D., Wick, L.Y., Filippidou, S., Kooli, W. M., Verrecchia, E.P. & P. Junier. Exploiting the fungal highway: Development of a novel tool for the in situ isolation of bacteria migrating along fungal mycelium. FEMS Microbiology Ecology. 2015 Nov; 91(11). doi: 10.1093/femsec/fiv116.

4. Filippidou S., Wunderlin T., Junier T., Jeanneret N., Johnson S., McMurry K., Gleasner C.D., Lo C.-C., Li P.E., Vuyisich M., Chain P.S., Junier P. Genome sequence of *Bacillus alveayuensis* strain 24KAM51, a halotolerant thermophile isolated from a hydrothermal vent. Genome Announcements. 2015 Jul; 3(4): e00982-15. doi:10.1128/genomeA.00982-15.

5. Filippidou S., Jaussi M., Junier T., Wunderlin T., Jeanneret N., Regenspurg S., Li P.E., Lo C.-C., Johnson S., McMurry K., Gleasner C.D., Vuyisich M., Chain P.S., Junier P. Genome sequence of *Aeribacillus pall*idus strain GS3372, an endospore-forming bacterium isolated in a deep geothermal reservoir. Genome Announcements. 2015 Jul; 3(4): e00981-15. doi:10.1128/genomeA.00981-15.

6. Papadimitriou-Olivgeris M., Filippidou S., Drougka E., Fligou F., Kolonitsiou F., Dodou V., Marangos M., Anastassiou E.D., Vantarakis A., Spiliopoulou I. Biofilm synthesis and presence of virulence factors among enterococci isolated from patients and water samples. Journal of Medical Microbiology. 2015 Nov; 64(11):1270-6. doi: 10.1099/jmm.0.000151.

7. Filippidou S., Jaussi M., Junier T., Wunderlin T., Roussel-Delif L., Jeanneret N., Vieth-Hillebrand A., Vetter A., Regenspurg S., Johnson S., McMurry K., Gleasner C.D., Lo C.-C., Li P.E., Vuyisich M., Chain P.S., Junier P. Genome sequence of *Anoxybacillus geothermalis* strain GSsed3, a novel thermophilic endospore-forming species. Genome Announcements. 2015 May; 3(3): e00575-15. doi:10.1128/genomeA.00575-15.

8. Filippidou S., Junier T., Wunderlin T., Lo C.-C., Li P.E., Chain P.S., Junier P. : Under-detection of endospore-forming Firmicutes in metagenomic data. Computational and Structural Biotechnology Journal. 2015 Apr; 13: 299-306. doi:10.1016/j.csbj.2015.04.002.

9. Bergottini V. M., Filippidou S., Junier T., Johnson S., Chain P.S., Otegui M.B., Zapata P.D., Junier P.: Genome Sequence of *Kosakonia radicincitans* Strain YD4, a Plant Growth-Promoting Rhizobacterium Isolated from Yerba Mate (*Ilex paraguariensis* St. Hill.), Genome Announcements. 2015 Mar; 3(2): e00239-15. doi: 10.1128/genomeA.00239-15.

10. Pion M., Bshary R., Bindschedler S., Filippidou S., Wick L.Y., Job D., Junier P. Gains of bacterial flagellar motility in a fungal world. Applied and Environmental Microbiology. 2013 Nov; 79(22):6862-7. doi: 10.1128/AEM.01393-13.

11. Filippidou S., Vantarakis A. Enterococci in the environment. Epidemiology and public health. Review, Iatriki. 2011, 100(2):113—126.

12. Kokkinos P., Ziros P., Meri D., Filippidou S., Kolla S., Galanis A., Vantarakis A. Environmental surveillance. An additional/alternative approach for virological surveillance in Greece? International Journal of Environmental Research and Public Health. 2011 Jun; 8(6): 1914-22. doi: 10.3390/ijerph8061914.

13. Kokkinos P., Ziros P., Filippidou S., Mpampounakis I., Vantarakis A. Molecular characterization of hepatitis A virus isolates from environmental and clinical samples in Greece. Virology Journal. 2010 Sep; 16, 7:235. doi: 10.1186/1743-422X-7-235.

14. Kokkinos P., Filippidou S., Karlou K., Vantarakis A. Molecular typing of enteroviruses, adenoviruses and hepatitis A viruses in untreated and treated sewage of a biological treatment plant in Greece. Food and Environmental Virology, 2010 Jun; 2(2): 89-96. 10.1007/s12560-010-9036-3

Under review/ Submitted

1. Filippidou S., Bindschedler S., Jeanneret N., Bueche M., Junier P. Bacterial diversity in the sulfur and iron springs of Ponts-de-Martel, Neuchâtel, submitted at *Bulletin neuchatelois des sciences naturelles*

In preparation

1. Filippidou S., Bueche M., Wunderlin T., Junier T., Jeanneret N., Dorador C., Molina V., Ioannidou A., Vargemezis G., Johnson D.R., Junier P. Survival off-limits: the key to success of endospore-forming Firmicutes, to be submitted at *Frontiers in Microbiology*

2. Filippidou S., Junier T., Jeanneret N., Palmieri I., Johnson S., Chain P.S., Junier P. Sporulation in Kurthia sp. str. 11kri323: proposed molecular pathway and spore description, to be submitted to *mBio*

3. Filippidou S., Junier T., Wunderlin T., Bueche M., Kooli W.M., Jeanneret N., Molina V., Lienhard R., Spangerberg J., Johnson S., Chain P.S., Dorador C., Junier P. Survival in bacteria: discovery of horizontal gene transfer leading to endospore-formation outside Firmicutes, to be submitted to *Nature Microbiology*

4. Filippidou S., Junier T., Wunderlin T., Jeanneret N., Palmieri I., Dorador C., Johnson S., Chain P.S., Junier P. Combined strategies to survival and adaptation of Firmicutes in saline environments, to be submitted to *ISMEJ*

Oral Presentations

Extremity: A Driving Force for Sporulation in Species that are Not Known to Sporulate. 7th European Spores Conference, 2016, London, UK.

What triggers sporulation in *Serratia ureilytica* Lr5/4? Joint annual meeting SSI - SSHH - SSM - SSTMP, 2015, Lugano, Switzerland.

Survival strategies meet classical ecological theories: the case of endospore-forming Firmicutes in extreme environments. 15th ISME Meeting, 2014, Seoul, South Korea.

Microbial Communities in Geothermal Sites. Are endospore-forming bacteria favoured? 5th Swiss Microbial Ecology Meeting, 2013, Murten, Switzerland.

Molecular correlation of HAV isolates from environmental and clinical samples. Panhellenic Congress of Biosciences 2010, Athens, Greece.

Poster Presentations

Filippidou S., Bueche M., Wunderlin T., Junier T., Jeanneret N., Dorador C., Molina V., Ioannidou A., Vargemezis G., Johnson D.R., Junier P. Survival off-limits: the key to success of endospore-forming Firmicutes, SfAM Winter Meeting: Psychrophiles and Extremophiles, 2016, London, UK

Filippidou S., Junier T., Wunderlin T., Bueche M., Kooli W.M., Jeanneret N., Roussel-Delif L., Molina V., Lienhard R., Spangerberg J., Johnson S., Chain P.S., Dorador C., Junier P. When microbial survival strategies never stop to wonder: sporulation outside Firmicutes. FEMS, 2015, Maasticht, Netherlands.

Filippidou S., Jeanneret N., Roussel-Delif L., Wunderlin T., Junier T., Molina V., Dorador C., Lienhard R., Junier P. Breaking bacterial stereotypes: the case of a spore-former member of a non-sporulating genus. Joint annual meeting SSI - SSHH - SSM - SSTMP, 2014, Fribourg, Switzerland. **1st place at Best Poster Award**.

Filippidou S., Jeanneret N., Roussel-Delif L., Wunderlin T., Junier T., Junier P. Diversity of Aerobic Endospore-forming Bacteria in Geothermal Sites in Greece. 14th ISME Meeting, 2012, Copenhagen, Denmark.

Filippidou S., Wunderlin T., Jeanneret N., Roussel-Delif L., Junier T., Junier P. Diversity and Distribution of Aerobic Endospore-forming Bacteria from Geothermal Sites in Greece. Joint annual meeting SSI - SSHH - SSM - SSTMP, 2012, St-Gallen, Switzerland

Filippidou S., Parasidis T., Alexandropoulou I., Papapetropoulou M., Vantarakis A. Screening of enteric virus types in raw and treated sewage. 1st Symposium COST 929 Current developments in food and environmental virology, 2008, Pisa, Italy

Filippidou S., Parasidis T., Alexandropoulou I., Stavrou E., Karlou K., Vantarakis A. Detection of Infectious Pathogenic Viruses in Untreated and Treated Wastewater Samples from an Urbanised Area. 13th International Congress on Infectious Diseases, 2008, Kuala Lumpur, Malaysia.

Participation in Training Courses

COST course. Metagenomics: biodiversity data analysis, July 2012, Madrid, Spain

CostAction 929 Training Course on Food and Environmental Virology, May 2010, Patras, Greece

Teaching Experience

Supervision

Master thesis:

Fabio Palmieri, Biologist. March 2015-April 2016. Master thesis: Extremophiles in volcanic environments. UNINE

Projects:

Ilona Palmieri, technician, November 2015 to date. Projects: 1. Definition of limits of salt tolerance in halophilic and halotolerant aerobic and anaerobic endospore-forming firmicutes. 2. Emergent spatial patterns allow the coexistence of competing bacterial species in a homogeneous environment

Andrej Al'Dourobi, technician. February 2013- August 2013. Project: Characterisation and species description of a spore-forming *S. ureilytica* strain Lr5/4. UNINE

Patricio Munos Raffat, undergraduate student, University of Antofagasta, Chile. January 2013- April 2013. Project: Screening of 200 *Bacillus spp* and *Geobacillus spp* for manganese oxidation properties with culture and molecular methods. UNINE

Amandine Pillonel, technician. September 2012- February 2013. Project: Physiological and biochemical characterisation of 12 *Bacillus spp* and *Geobacillus spp* strains and their biomineralisation properties. UNINE

Stages:

Scherler Laeticia, bachelor student in Biology. February 2016 to date: Project: Study of chemical compounds produced by halotolerant and halophilic endospore-forming Firmicutes in extreme saline conditions. UNINE

Sophie Poirier, high-school student. April 2015, December 2015: Maturity Project and Exam: Sporulation triggered by UV radiation and by starvation in *S. ureilytica*, *M. timonae* and *B. subtilis*. UNINE-Lycée Jean Piaget

Wida Brumand, high-school student. April 2015, December 2015: Maturity Project and Exam: Sporulation triggered by psychrophilic and thermophilic conditions, by absence of oxygen, and by desiccation in *S. ureilytica*, *M. timonae* and *B. subtilis*. UNINE-Lycée Jean Piaget

Ilona Palmieri, high-school student. January 2015: Project: Carbon source utilisation of *Anoxybacillus geothermalis*. UNINE

Lucrezia Albertini, bachelor student in Biology. November 2014- December 2014: Project: Physiological and Biochemical characterization of *Massilia timonae*. UNINE

Lectures

Spring semester 2015 (Substitution for Prof. Pilar Junier):

- Seminars in Microbial Ecology, Biogeosciences Master, UNINE, UNIL
- Seminars in Ecology, Master in Biology, UNINE

Winter semester 2014 (Substitution for Prof. Pilar Junier):

- General Bacteriology, Bachelor 2nd year, UNINE
- Molecular Biology, Bachelor 3rd year, UNINE

Practical courses

Spring Semester 2015: Problem based learning in fungal-bacterial interactions, Bachelor 3rd year, UNINE

Winter semester 2014: Bacteriology practical courses, Bachelor 2nd year, UNINE

Molecular Biology practical course, Bachelor 3rd year, UNINE

Spring semester 2014: Problem based learning in water quality, Bachelor 3rd year, UNINE

Winter semester 2013: Bacteriology practical courses, Bachelor 2nd year, UNINE

Spring semester 2013: Problem based learning in water quality, Bachelor 3rd year, UNINE

Winter semester 2012: Bacteriology practical courses, Bachelor 2nd year, UNINE

Training courses

"Water microbiology analysis methods", practical and theoretical seminar, 6-8 December 2010, Univ. of Patras

Service-related activities

Reviewer in the following journals:

PLOS One

Canadian Journal of Microbiology

Organisation

May 2014: Organisation of the Annual PhD meeting, "Ethics in Science", UNINE.

Memberships

Member of the International Society for Microbial Ecology (ISME)

Member of the Society for Applied Microbiology (SfAM)

Member of the Swiss Society of Microbiologists (SSM)