South Dakota State University

Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange

Electronic Theses and Dissertations

2020

Interaction of Streptomyces Closely Associated with Bacillus in Nitrogen-limiting Conditions

Aminya Baniya
South Dakota State University

Follow this and additional works at: https://openprairie.sdstate.edu/etd

Part of the Bacteriology Commons, and the Biology Commons

Recommended Citation

Baniya, Aminya, "Interaction of Streptomyces Closely Associated with Bacillus in Nitrogen-limiting Conditions" (2020). *Electronic Theses and Dissertations*. 4068.
https://openprairie.sdstate.edu/etd/4068

This Thesis - Open Access is brought to you for free and open access by Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Open PRAIRIE: Open Public Research Access Institutional Repository and Information Exchange. For more information, please contact michael.biondo@sdstate.edu.

INTERACTION OF STREPTOMYCES CLOSELY ASSOCIATED WITH BACILLUS IN NITROGEN-LIMITING CONDITIONS

BY

AMINA BANIYA

A thesis submitted in partial fulfillment of the requirements for the

Master of Science

Major in Biological Sciences

Specialization in Microbiology

South Dakota State University

2020

THESIS ACCEPTANCE PAGE

Amina Baniya

This thesis is approved as a creditable and independent investigation by a candidate for the master's degree and is acceptable for meeting the thesis requirements for this degree. Acceptance of this does not imply that the conclusions reached by the candidate are necessarily the conclusions of the major department.

Volker Brozel
Advisor Date

Heike Bücking, Ph.D.
Department Head Date

Dean, Graduate School Date

ACKNOWLEDGEMENTS

To my inspiration, my Mom Tika Baniya: Blessed to have you. Thank you very much!

I would like to express my deep and sincere gratitude and heartfelt appreciation to my advisor Dr. Volker Brözel, for providing me invaluable guidance throughout this research. I am thankful for his encouragement and support, without which this would not be successful.

I would like to thank all my Committee for giving me feedbacks and encouraging me to achieve my goal. My sincere thanks to Dr. Micheal Hildreth and Dr. Joy Scaria for their support, assistance, and encouragement.

I would like to pay my special regards to my lab mates, Amrit Koirala, Bikram Das, Nabilah Alshibli, York Dakota and Armaan Kaur Sandhu. I am also grateful to all the faculties and staffs of Department of Biology and Microbiology who have directly or indirectly helped me to accomplish my goal.

Last but not the least, a very special gratitude and grateful thanking to my family members and friends for their constant encouragement and moral support.

Thank you all for invaluable support and encouragement.

TABLE OF CONTENTS

ABBREVIATIONS viii

LIST OF FIGURES x

LIST OF TABLES xiii

ABSTRACT xiv

1. Literature Review 1

1.1 Introduction 1

1.1.1 Microbe-Microbe Interactions 1

1.1.2 Syntrophy/ Nutrient exchange 1

1.1.3 Endophytes/Endosymbionts 2

1.2 Synergy in *Streptomyces* species 6

1.2.1 *Streptomyces* as symbiont 8

1.2.2 *Bacillus*-*Streptomyces* interaction 10

1.2.3 Co-cultures aka Hitchhikers 11

1.3 Bacterial growth under nitrogen deficient conditions 12

1.3.1 Biological nitrogen fixation 12

1.3.2 Free-living diazotrophs 13

1.3.3 Nitrogenase enzyme 14

1.3.4 Methods to determine nitrogen fixing capability 14

1.4 Methods to study bacterial interactions 16

1.4.1 Separation techniques 16

1.4.2 MALDI-imaging mass spectrometry 18

1.4.3 Quantitative qPCR 18

1.4.4 Transformation with reporter genes 19

1.4.5 Fluorescence *In Situ* Hybridization (FISH): 23

1.5 Whole Genome Sequence Analysis: 23

1.6 Conclusion 27

2. Chapter I: Characterization of Bacillus that co-occur with Streptomyces isolate 28

2.1 Introduction 28

2.2 Materials and Methods 30

2.2.1 Sources of cultures 30

2.2.2 Approaches to obtain *Bacillus* into single culture 32

2.2.3 Biochemical evidence for nitrogen fixation 33

2.2.4 DNA extraction and PCR amplification of 16S rRNA and *nifH* 34

2.2.5 Phylogeny of *Bacillus* isolates 36

2.3 Results 36

2.3.1 Isolation of *Bacillus* 37

2.3.2 Growth of isolated *Bacillus* on NFA 38

2.3.3 Biochemical evidence for Nitrogen fixation 39

2.3.4 Amplification of *nifH* by PCR 42

2.3.5 Phylogeny of *Bacillus* 44

2.4 Discussion 45

3. Chapter 2: Characterizing interaction between Bacillus and Streptomyces..48

3.1 Introduction 48

3.2 Materials and Methods 51

3.2.1 Source of cultures 51

3.2.2 Growth of *Streptomyces* in NFA 52

3.2.3 Quantification of ratio of *Bacillus* to *Streptomyces* 52

3.2.4 Approaches to obtain *Streptomyces* into single culture 54

3.2.5 Transformation of *Bacillus* with mCherry (RFP) 56

3.3 Results 58

3.3.1 Growth of *Streptomyces* in NFA 58

3.3.2 Prevalence of *Bacillus* associated with *Streptomyces* during growth 59

3.3.3 Separation of *Streptomyces* from *Bacillus* 65

3.3.4 Transformation of *Bacillus* isolates with mCherry 68

3.3.5 Co-culture of transformed *Bacillus* 34B with *Streptomyces* isolates 71

3.4 Discussion 71

4. Chapter 3: Whole Genome Sequence Analysis of Bacillus and Streptomyces ……..…….75

4.1 Introduction 75

4.2 Materials and Methods 77

4.2.1 Culture preparation 77

4.2.2 Assembly and annotation 77

4.2.3 Phylogeny of *Streptomyces* 78

4.2.4 KEGG annotation and reconstruction of pathways 79

4.2.5 Analysis of genes using Kaiju 79

4.3 Results 80

4.3.1 Genome assembly and annotation 80

4.3.2 Phylogeny of *Streptomyces* 82

4.3.3 Functional annotation 83

4.3.4 Analysis of genes using Kaiju 88

4.3.5 Bio-synthetic gene clusters of *Streptomyces* 90

4.4 Discussion 99

5. Conclusions and Questions for future study 101

5.1 Conclusion 101

5.2 Questions for future study 102

REFERENCES 103

APPENDIX 121

ABBREVIATIONS

α: Alpha

ATCC: American Type Culture Collection

β: Beta

^{o}C: Degrees Celsius

CO_2: Carbon dioxide

Ct: Cycle threshold

ARA: Acetylene Reduction Assay

BLAST: Basic Local Alignment Search Tool

BNF: Biological Nitrogen Fixation

δ: Delta

DNA: Deoxyribonucleic acid

dNTP: Deoxyribonucleic triphosphate

FISH: Fluorescence In Situ Hybridization

GC: Gas Chromatography

GFP: Green Fluorescent Protein

HGT: Horizontal Gene Transfer

KEGG: Kyoto Encyclopedia of Genes and Genomes

KO id: KEGG Orthology id

M: Molar

mM: Millimolar

μ: Micro

MALDI: Matrix-assisted laser desorption ionization

MEMs: Maximum Exact Matches

MLST: Multilocus Sequence Typing

Min: Minute

MS: Mass Spectrometry

NCBI: National Centre for Biotechnology

NFA: Nitrogen Free Agar

NGS: Next Genration Sequencing

ONT: Oxford Nanopore Technologies

OTU: Operational Taxonomic Unit

PCR: Polymerase Chain Reaction

PEG: Polyethyl glycol

R2A agar: Reasoner's agar

RFP: Red Fluorescent Protein

RNA: Ribonucleic acid

rRNA: Ribosomal ribonucleic acid

SDSU: South Dakota State University

Sec: Second

TTC: 2,3,5 Triphenyl Tetrazolium chloride

THOR: The hitchhikers of the rhizosphere

USGS: U.S. Geological Survey

V: Volt

qPCR: Quantitative Polymerase Chain Reaction

LIST OF FIGURES

Fig 1: Possible basic steps involved during the formation of endosymbiosis and organellogenesis. 5

Fig 2: Life cycle of Streptomyces beginning with spore germination, formation of substrate mycelium and aerial hyphae to spore formation 8

Fig 3: Interaction of *Streptomyces* spp. with eukaryotic organisms showing both positive and negative effects 10

Fig 4: Steps involved in amplification of target DNA sequences in qPCR reaction using SYBR Green reporter dye (A) and TaqMan reporter dye (B) 19

Fig 5: Steps illustrating reference-guided de novo assembly 25

Fig 6: Three *Streptomyces* isolates growing on NFA (top) and R2A (bottom) for a week. 31

Fig 7: Slide press method to isolate free *Bacillus* from *Streptomyces* colonies grown on R2A. Mature Streptomyces colonies were placed onto sterile microscope slides 32

Fig 8: Replica-cultures of glass slide having smear of *Streptomyces* 321I, 34I and 40I pressed onto nitrogen-rich (R2A) and nitrogen-limited (NFA) media 38

Fig 9: *Bacillus* isolates 321B, 34B and 40B growing on NFA after 7 d of incubation and on R2A after 3 days of incubation at 28^{o}C. 39

Fig 10: Acetylene reduction assay of *Bacillus* isolates 321B, 34B and 40B, and *S. typhimurium* as negative control cultured on nitrogen free agar in sealed glass tubes for 5d, and supplemented with acetylene. 40

Fig 11: Growth of isolates (321B, 34B and 40B) in semi-solid NFA media visualized by reduction of TTC to red pigment after incubation for 5 d at 28^{o}C 41

Fig 12: Gel image showing bands of *nifH* PCR with positive control: *H. seropedicae* (1), negative control: water (2), isolates 321B (3), 34B (4) and 40B (5). 42

Fig 13: Alignment of sequence obtained by amplifying *nifH* by PCR of *Bacillus* 321B using BLASTN. 43

Fig 14: Chromatogram of *nifH* obtained by Sanger sequencing of *Bacillus* 321B 43

Fig 15: Maximum likelihood phylogenetic tree using concatenated sequences of *rpoA*, *recA*, *gyrA* and 16S rRNA 45

Fig 16: Three *Streptomyces* isolates growing on NFA (top) and R2A (bottom) for 7d 51

Fig 17: Microscopy of NFA and R2A-grown isolate 321I (a), 34I (b) and 40I (c) over time using bright field microscopy at 1000X 61

Fig 18: Optimization of PCR annealing temperatures for Firmicutes-specific primers (a) and Actinomycetes-specific primers (b) 62

Fig 19: The ratio of Firmicutes to Actinomyces 16S rRNA gene threshold values for isolates 321I (a), 34I (b) and 40I (c) cultured on NFA and R2A 64

Fig 20: DNA extracted from filter separated *Streptomyces* amplified by Firmicutes-specific (a) and Actinomycetes-specific primers (b) 66

Fig 21: Zones of inhibition formed by Chloramphenicol discs on lawn culture of *Bacillus* isolates 321B, 34B and 40B (left) and *Streptomyces* cultures 321I, 34I and 40I (right). 67

Fig 22: *Streptomyces* colonies on R2A after serial dilution of spores of 321I (a) and 40I (b) 68

Fig 23: Growth curves of *Bacillus* and *B. cereus* ATCC 1457a isolates in R2A broth at 37^{o}C with continuous shaking 69

Fig 24: Gel image showing plasmid bands obtained from plasmid extraction from transformed *Bacillus* isolate 34B.....70

Fig 25: Bioinformatics analysis workflow for de-novo genome assembly by Genewiz...78

Fig 27: Phylogenetic relationship using five-gene (*recA*, *atpD*, *gyrB*, *rpoB* and 16S rRNA) concatenated sequences.....83

Fig 28: Taxonomy classification of reference genomes *B. subtilis* 168 (a) and *B. subtilis* UBBS-14 (b), and *Bacillus* 321B (c) and 40B (d) paired raw reads using Kaiju.....90

LIST OF TABLES

Table 1: Primers used for PCR amplification of 16S rRNA and *nifH* genes. 35

Table 2: Sampling times for microscopy of cultures grown on R2A and NFA. 52

Table 3: Phylum-specific 16S rRNA gene primers used. 53

Table 4: Growth of *Streptomyces* species (NRRL) in NFA. 59

Table 5: Sequencing quality overview. 80

Table 6: Genome assembly showing the contigs and numbers of genes. 81

Table 7: Analysis of reconstruct KEGG pathways using KO ids of *Bacillus* 40B, *Streptomyces* 40 and *Streptomyces* 40 filtered*. 84

Table 8: Bio-synthetic gene clusters of *Streptomyces* 40I determined by anti-SMASH software version 5.1.2. 90

Table 9: Bio-synthetic gene clusters of *Streptomyces* 321I determined by anti-SMASH software version 5.1.2. 94

ABSTRACT

INTERACTION OF STREPTOMYCES CLOSELY ASSOCIATED WITH BACILLUS IN NITROGEN-LIMITING CONDITIONS

AMINA BANIYA

2020

Biological Nitrogen Fixation is the process of conversion of atmospheric dinitrogen into ammonia which is performed by symbiotic as well as free-living diazotrophic bacteria. A previous project to isolate free-living nitrogen fixing bacteria from a natural prairie yielded *Streptomyces* which grew on nitrogen free medium. Genome sequencing of three isolates revealed the presence of a second 16S rRNA gene sequence aligning with *Bacillus*, however repeated streaking only yielded *Streptomyces* colonies. The aim of this research was to characterize the interaction of three *Streptomyces* isolates that grow under nitrogen-free conditions, and appeared to be associated with *Bacillus*. The separation of *Bacillus* was achieved due to formation of heat resistant endospores, but isolation of *Streptomyces* into purity could not be achieved despite multiple approaches. Several separation techniques such as the size occlusion method, Chloramphenicol treatment of resistant *Streptomyces*, and growth of *Streptomyces* spores indicated that *Bacillus* is very closely and tightly associated with *Streptomyces,* that it seems to either have an endophytic relationship or stay as hitchhiker. The analysis of nitrogen fixing ability of *Bacillus* was done using acetylene reduction assay, $^{15}N_2$ isotope incorporation assay, growth in nitrogen free soft agar, and amplification of *nifH* by PCR. The weakly positive acetylene reduction result suggested

weak ability to fix nitrogen, however the genome did not contain any *nif* genes. We studied the growth of *Bacillus* and *Streptomyces* in co-culture in nitrogen-rich (R2A) and nitrogen-limiting (NFA) conditions. Microscopic investigation showed the presence of free rods on R2A after extended incubation, but absence of free rods on NFA, suggesting that *Bacillus* could be localized inside *Streptomyces* filaments, with some release upon ageing of the cultures. However, the quantification of 16S rRNA of both *Bacillus* and *Streptomyces* by qPCR indicated a constant ratio of *Bacillus* to *Streptomyces*, even in very young cultures. The genomes of *Bacillus* (321B and 40B) and *Streptomyces* (321I and 40I) were used to analyze the predicted metabolic pathways, with various pathways indicating species-species interaction. Degradation of aromatic compound pathways were more numerous in *Streptomyces* than *Bacillus*. Metabolic pathways such as RNA degradation, amino acids degradation were found to be more prominent in *Bacillus*. *Streptomyces* had a Vancomycin group antibiotics biosynthesis pathway, whereas a Vancomycin resistance mechanism was present in *Bacillus*. A four gene phylogenetic tree of *Bacillus* 321B and 40B showed that they are very closely related to each other and fall in the same cluster as *B. subtilis*. A five gene phylogenetic tree of *Streptomyces* 321 showed its closeness to *S. phaeochromogenes*, and 40 was closer to *S. umbrinus*. This study has opened areas for further investigation on the physical location of the interaction that helps to determine whether *Bacillus* occurs inside *Streptomyces* filaments as an endophyte or is only present as close association.

1. Literature Review

1.1 Introduction

Insights into properties and behavior of microorganisms have been derived primarily through single or pure-culture approaches. This reductionist approach has yielded much valuable understanding, but has inadvertently made us think of microbial species as self-sufficient islands unaffected by others. This review seeks to highlight the interactions among microbes.

1.1.1 Microbe-Microbe Interactions

Microorganisms are seldom found as single species populations, rather as communities comprising various different species leading to interaction within the same or different species (Stubbendieck, Vargas-Bautista *et al.*, 2016). Interactions between two bacteria, two fungi, bacteria and fungi, fungi and plants, bacteria and plants, and among bacteria, fungi and plants are widely found (Braga, Dourado *et al.*, 2016). These interactions can be mutualistic or parasitic, often involving physiochemical and ecological changes due to metabolite exchange, signaling, as well as genetic exchange (Pacheco & Segre, 2019). Results of these interactions may lead to either production of secondary metabolites, siderophores, quorum sensing, biofilm production, or cellular transduction signaling that generally benefits the partners involved in the interaction.

1.1.2 Syntrophy/ Nutrient exchange

Syntrophy (syn, together; trophy, feeding) is a metabolic collaboration between species. It is used by scientists studying eukaryotic origins, in the Symbiogenesis theory,

that posits the origin of eukaryotes by metabolic syntrophy of an archaeal cell and a bacterial cell (Lang & Burger, 2012). The best known Symbiogenesis theory entails an endosymbiotic α-Proteobacterial partner capable of producing hydrogen (oxidative phosphorylation) that was hosted by a hydrogen-consuming archaeal cell (methanogen). There is no firm evidence that mutual syntrophy may lead to the evolution of endosymbiosis or symbiogenesis, as examples of intermediate forms such as prokaryote-prokaryote endosymbioses have not been reported (Zachar & Boza, 2020).

In a complex community found in soil, relatively few microbes proliferate while the majority are checked through various processes. Intricate communities have syntrophic interactions, biomass recycling and exchange of biogeochemical cycles. Syntrophic interactions among heterotrophs and autotrophs were demonstrated using genomics-informed proteomic data in an acetate-amended, sediment-hosted subsurface microbial community (Handley, VerBerkmoes *et al.*, 2013). High acetate concentrations stimulated anaerobic heterotrophs to reduce sulfate, fix nitrogen and oxidize acetate to CO_2, which was beneficial for autotrophs that could utilize the byproducts to proliferate in that environment.

1.1.3 Endophytes/Endosymbionts

Diverse bacteria or fungi are found living within plant tissue establishing mutual/beneficial relationships, which can be for at least parts of their life cycle. “Any organism that grows within plant tissue is termed an endophyte”, the first definition of an endophyte given by De Barry (1866). This has been explored well due to the importance of the interaction of endophytes with various plants. Although attempts to discover the origin of endophytic organisms in different plant species have been made, information on

the specific interaction, strategy of existence and transmission is provided by genomic data such as correlation of bacterial lifestyle and its genomic size (Gouda, Das *et al.*, 2016). Endophytes are able to communicate with the host plant cells, using various chemical messengers that have receptors on both sides of the interacting partners, affecting host genetic and phenotypic expression such as immune response against pathogens, increase in resistance to biotic and abiotic stress, and production of plant secondary metabolites (Khare, Mishra *et al.*, 2018). Members of Actinobacteria, Proteobacteria, and Firmicutes are the most common documented endophytes (Golinska, Wypij *et al.*, 2015). Archaeal taxa of *Thaumarchaeota*, *Crenarchaeota*, and *Euryarchaeota* was detected as endophytes, but their significant role in the interaction is not clear (Müller, Berg *et al.*, 2015).

Insect and microbe endosymbioses are another widely known phenomenon where the insect may depend completely on the respective microbe for some essential nutrient or development signal. Insects have evolved with special host cells known as bacteriocytes that harbor endosymbionts. For example, the endosymbiont *Wolbachia* provides antiviral resistance to its host (mosquitoes), an effort that has been used to control arbovirus replication and transmission to vertebrates (Rainey, Shah *et al.*, 2014).

Endosymbioses among eukaryotes and between prokaryotes and eukaryotes are omnipresent and, thus are well documented (López-García, Eme *et al.*, 2017; Nowack & Melkonian, 2010), whereas, endosymbioses among prokaryotes are rare (Zachar & Boza, 2020). Zachar and Boza (2020) have listed the prokaryotic examples showing endosymbiosis known to date; *Tremblaya princeps-Morganella endobia*, tick mitochondrion-*Midichloria mitochondrii* bacterium, dinoflagellate plastid (Cyanobacteia)

in *Woloszynskia pascheri* and *Pleurocaspa minor* (Cynanobacteria)-Bacterium. The first three examples represent multilevel endosymbioses of one prokaryote inside another, located in a eukaryotic cell. The fourth example is the only one reported to have one-one prokaryotic endosymbiosis, however, there is limited data to determine whether the endosymbiosis is stable or some occasional infection, based on a single report (Wujek, 1979). There is a huge question on the mechanism of endosymbiosis and organellogenesis, starting either due to syntrophic interaction, infection or phagocytosis. Evolution of an interaction that leads to obligate endosymbiosis from facultative and finally, obligate co-dependency with nuclear integration and protein import is necessary. Hypothetically, possible stages in endosymbiosis are outlined in Fig 1.

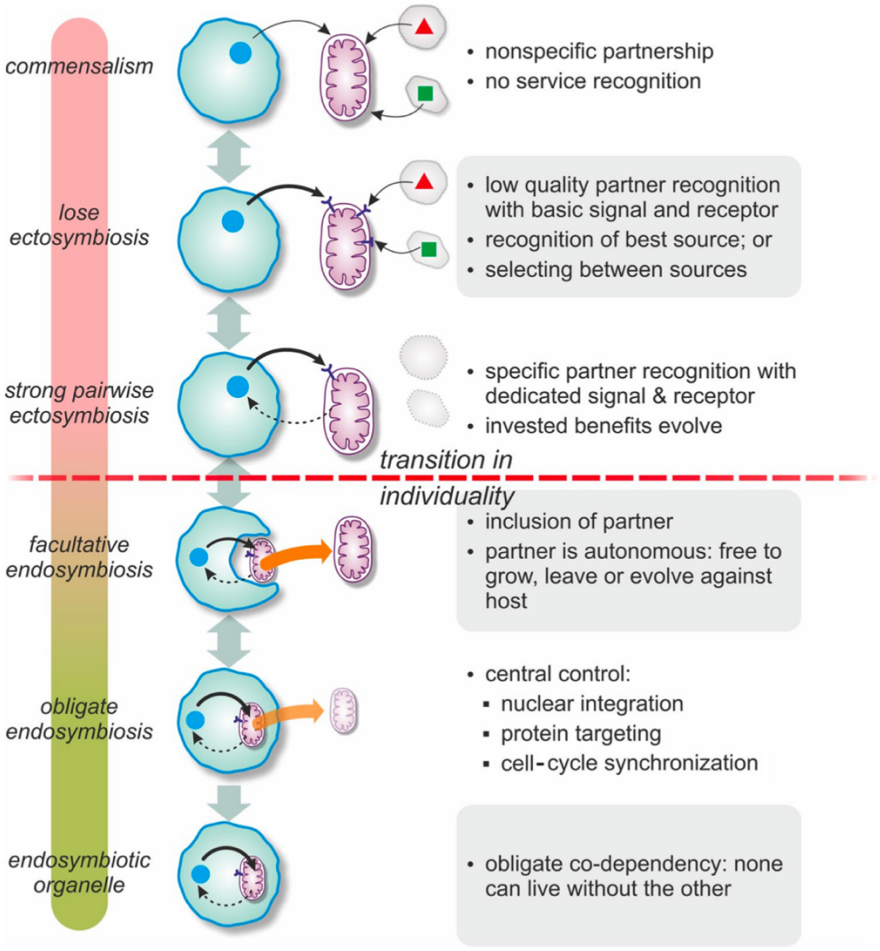


Fig 1: Possible basic steps involved during the formation of endosymbiosis and organellogenesis. Geometric shapes represent various benefits (e.g., metabolites), solid black arrows represent the source and flow of the various benefits, dashed arrows indicate investments, and colored arrows indicate the option to leave the host. (From (Zachar & Boza, 2020)).

A mealybug species, *Planococcus citri*, has a unique nested endosymbiotic system that harbors *Tremblaya princeps* (β-proteobacterium) in its bacteriocytes, which in turn harbors many *Moranella endobia* (gamma-proteobacterium) intracellularly, (Lopez-Madrigal, Balmand *et al.*, 2013). Genome sequencing and analysis of *M. endobia*

and *T. princeps* has revealed that *M. endobia* is responsible for cellular components and energy synthesis, and *T. princeps* utilizes those as precursors for amino acid synthesis (Lopez-Madrigal, Latorre *et al.*, 2013). The genomes of both the bacterial partners have been very reduced such that neither has complete metabolic pathways, but is dependent on the partner. The genome of *T. princeps* has 120 protein-coding genes, one of the smallest known, and has lost functional genes relying, on its host for bacterial structure, and on *M. endobia* for other functional genes. The decoupling of *M. endobia* and *T. princeps* densities has revealed the different regulatory mechanisms present in this kind of nested intracellular symbiosis (Parkinson, Gobin *et al.*, 2016). To date this is the only well-known case of a bacterial endosymbiont occurring in another bacterium, both obligately dependent on their eukaryotic host. Reports of other examples are scant.

1.2 Synergy in *Streptomyces* species

Soil is not a uniform system, rather a heterogenous system, due to the complex composition of its living entity. All kinds of microbes including fungi, worms and insects dwell in the soil, making it a hub for studying interactions between each member (Braga, Dourado *et al.*, 2016). The genus *Streptomyces* of the Actinobacteria phylum is widely prevalent in soil and has been extensively studied, primarily due to its antibiotic producing ability (Challis & Hopwood, 2003). Actinobacteria form one of the largest phyla and are ancient among lineages of bacteria (Ventura, Canchaya *et al.*, 2007). *Streptomyces* is a gram-positive, mycelium- and spore-forming soil bacterium with large genome and multiple gene clusters for secondary metabolite production such as antibiotics, and surfactants. Although *Streptomyces* have been studied for around 80 years, the full potential for metabolite production has not been achieved. Many gene

clusters that appear to encode synthesis of complex metabolites are not expressed in single culture. More recent research has explored the interaction of *Streptomyces* with other soil microbes to obtain novel secondary metabolites. Recent studies on cocultivation of different soil microbes with *Streptomyces* have been able to produce novel metabolites. For example, *Streptomyces* sp. CJ-5 co-cultured with *Tsukamurella pulmonis* TP-B0596 led to isolation of three novel butanolide chojalactones (Hoshino, Wakimoto *et al.*, 2015).

Besides their antibiotic producing ability, *Streptomyces* are also extraordinary in their complex developmental life cycle (Fig 2). They have four distinct stages of development; spore germination, vegetative or substrate mycelium, aerial hyphae formation, and sporulation. A new mode of *Streptomyces* growth known as exploratory growth has been identified, which was induced upon physical association with yeast as well as some volatile signals (Jones, Ho *et al.*, 2017). Explorer cells reveal a novel growth strategy obtained due to inter-species interaction. Jones (2017) demonstrated exploratory growth that is induced by fungal interactions, nutrient depletion and volatile signals. Exploratory cells are distinguished from vegetative mycelium by their non-branching conformation that rapidly traverse biotic and abiotic surfaces.

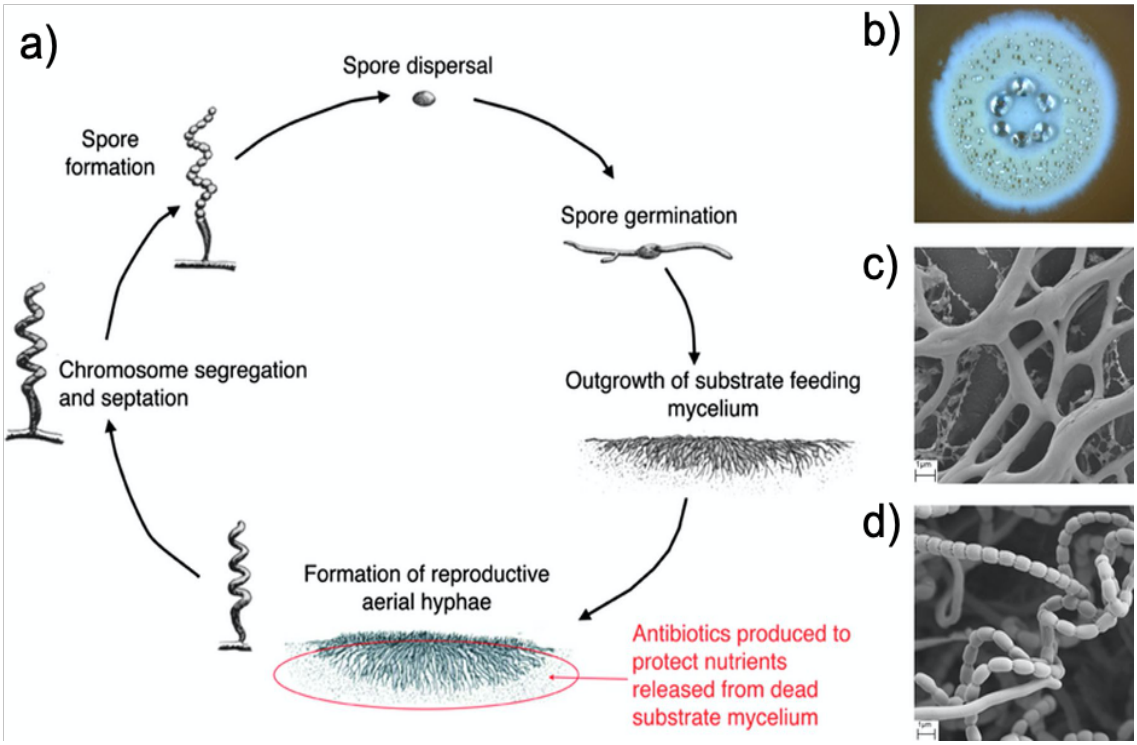


Fig 2: Life cycle of Streptomyces beginning with spore germination, formation of substrate mycelium and aerial hyphae to spore formation (a). The image at right show a colony of Streptomyces spp. at 40X magnification (b), scanning electron photographs with 400X magnification of substrate mycelium (c) and aerial hyphae with spore chains (d) of Streptomyces coelicolor grown on soya flour mannitol agar (From (Seipke, Kaltenpoth et al., 2012)).

1.2.1 *Streptomyces* as symbiont

Streptomyces do not exist as free-living bacteria only, but some are found to form symbioses with higher organisms such as plants and invertebrates (Seipke, Kaltenpoth *et al.*, 2012) (Fig 3). Ants, beetles and wasps are symbiotically associated with *Streptomyces* that provide defense against pathogenic bacteria, and in return get food and

protection from the host (Kaltenpoth, 2009). Research done on isolation of *Streptomyces* from marine sponge has always yielded *Streptomyces* species as the dominant actinobacteria, and suggests that they are stably associated with each other (Bull, Stach *et al.*, 2005). *Philanthus* sps. (Beewolf solitary digger wasps) larvae use multidrug cocktails of antibiotics, secreted by endosymbiont *Streptomyces* growing inside the antennal glands of female wasps, to combat bacterial and fungal infection from the surrounding (Kroiss, Kaltenpoth *et al.*, 2010). Several other associations between *Streptomyces* and diverse eukaryotes have been discovered during searches for novel antibiotics, such as the attine ant system, southern pine beetles, allomerus ants, as well as in the guts of various species, including termites, beetles, millipedes, wood lice and earthworms (Seipke, Kaltenpoth *et al.*, 2012).

The relationship between plants and *Streptomyces* vary from benign saprophytes to beneficial endosymbionts to pathogens (Seipke, Kaltenpoth *et al.*, 2012). Endophytic *Streptomyces* are different than pathogenic ones in that they persist for longer periods of time inside the plant without causing disease and lack virulence factors too (Coombs & Franco, 2003). Production of plant hormones like auxin and nutrient assimilation is enhanced by endophytic *Streptomyces*, for example, *S. lydicus* colonizes pea which increases the root nodulation frequency by *Rhizobium* spp., contributing to increase in iron and molybdenum assimilation and hence, robust growth of the plant is achieved (Tokala, Strap *et al.*, 2002).

Scientists have been arguing that the secondary metabolites produced by *Streptomyces* are not only for inhibition of other microbes but also for modulating the transcription and translation of beneficial genes in target microbes (Yim, Wang *et al.*,

2007). Naturally occurring small molecules modulate transcription in the target cells at very lower concentrations, affecting 5-10% of all transcripts (Goh, Yim *et al.*, 2002), due to their interaction with receptors or RNA polymerase (Yim, Wang *et al.*, 2007).

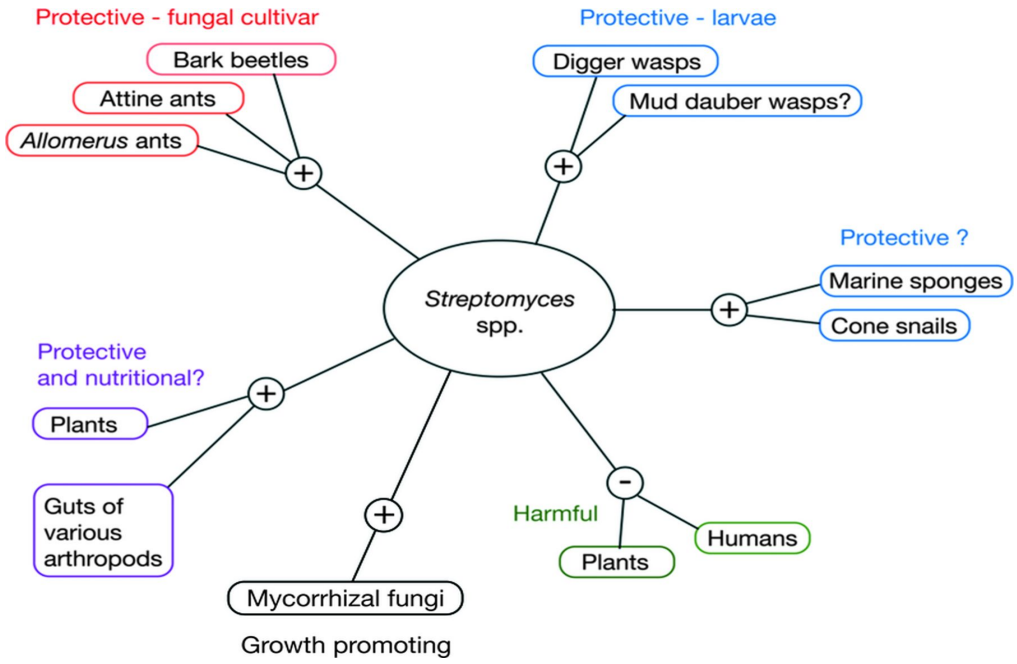


Fig 3: Interaction of *Streptomyces* spp. with eukaryotic organisms showing both positive and negative effects. (From (Seipke, Kaltenpoth *et al.*, 2012)).

1.2.2 *Bacillus-Streptomyces* interaction

Bacillus are gram positive rods forming tough endospores, reported to be found in soil as well as the gastrointestinal tract. Both *Bacillus* and *Streptomyces* are soil dwelling microbes and there is high probability of encounter with each other in this environment. *Bacillus-Streptomyces* interactions have been reported in various studies that show competition for survival in the nutritionally poor soil habitat and its use to enhance antibiotic secretion for industrial purposes. *S. coelicolor* in the presence of *B. subtilis*

produced undencylprodiogiosin, a pigmented antibiotic, which has useful anticancer property (K. J. Luti & F. Mavituna, 2011). Researchers were able to discover the new metabolite dentigerumycin E with antiproliferative and antimetastatic activities from a coculture of marine *Streptomyces* and *Bacillus* strains that were isolated from an intertidal mudflat. However, the production was not detected in single culture (Shin, Byun *et al.*, 2018), indicating a synergistic effect.

There are examples where *Bacillus* hinders in the development of *Streptomyces* when exposed to secondary metabolites produced by *Streptomyces*. Expression of *ytnP* by *B. subtilis* in the presence of streptomycin produced by *S. griseus* was able to inhibit the streptomycin production pathway and aerial mycelium development in *S. griseus* (Schneider, Yepes *et al.*, 2012). The effects of the presence of a second bacterial species on a well-characterized pathway of bacterial development has been studied recently (Hoefler, Gorzelnik *et al.*, 2012). Secondary metabolite surfactant produced by *B. subtilis* can affect the developmental pathway of *S. coelicolor* with no aerial hyphae production but without having detrimental effects on vegetative growth.

1.2.3 Co-cultures aka Hitchhikers

Hitchhikers are bacteria that are not visible bacterial in colonies incubated for long times (up to 4 weeks) but are carried along with the colony forming bacteria (Lozano, Bravo *et al.*, 2017). It was reported that *Pseudomonas koreensis CI12* was co-isolated as hitchhiker with *B. cereus* cultures from soybean roots (Lozano, Bravo *et al.*, 2017). *B. cereus* when isolated as model rhizosphere Firmicute was later characterized to have 20 other hitchhikers that play roles in induction of dendritic expansion of *B. cereus*, and robust production of biofilms when compared to individual growth (Lozano, Bravo *et*

al., 2019). Lozano et al. have introduced a model microbiome, THOR, that dissects community behavior at the genetic level and the members of the community are the hitchhikers of the rhizosphere.

1.3 Bacterial growth under nitrogen deficient conditions

Microbes in the soil have direct impact on the production of plants. They are involved in the carbon cycle, decomposition, nitrogen cycle, phosphorous cycle as well as regulation of plant diversity (van der Heijen, 2008). Although the atmosphere has 80% stable dinitrogen, plants can utilize either ammonium or nitrate forms of nitrogen (Maathuis, 2009). The major physiological functions of nitrogen are to be incorporated as amino groups in amino acids, form ring structures of purines and pyrimidines of nucleotides, along with acting as co-enzymes, photosynthetic pigments, secondary metabolites and polyamines. Nitrogen deficiency has led to much loss in net crop yield throughout the world (van der Heijen, 2008). According to a U.S. Geological survey fact sheet, about 85% of ammonia produced is used in fertilizers (USGS fact sheet 2019), the majority of which is produced via industrial process (Haber-Bosch process) which uses a high amount of energy.

1.3.1 Biological nitrogen fixation

The original alternative to industrially fixed nitrogen, biological nitrogen fixation (BNF), occurs naturally. BNF is the process of conversion of atmospheric dinitrogen into ammonia which is done by symbiotic as well as free-living diazotrophs (Galloway, 1998). The ability to reduce atmospheric dinitrogen to ammonia (Reed, Cleveland *et al.*, 2011) is unique among prokaryotes, including green Sulphur bacteria, *Firmicutes*,

Cyanobacteria, *Alphaproteobacteria*, *Betaproteobacteria*, *Deltaproteobacteria,* and *Gammaproteobacteria* (Dixon & Kahn, 2004). Symbiotic BNF that occurs in close association between plants such as legumes and nitrogen fixing bacteria is well-studied. Close association of bacterial genera like *Azospirillum* is often found with non-leguminous crops such as sugarcane, rice, maize, wheat, and helps with enhancement of growth and production of such crops due to supply of fixed nitrogen (Saikia & Jain, 2007). Grasslands such as natural prairies do not have such leguminous plant in such amount that could contribute to the nitrogen needs to sustain the ecosystem. Free-living diazotrophs have not been well studied. Nitrogen fixation in agricultural land with crops such as legumes and rhizobia symbiosis can be determined with some degree of confidence, but such determinations are hard to make in natural ecosystems such as grasslands having mostly free-living diazotrophs (Cleveland, Townsend *et al.*, 1999).

1.3.2 Free-living diazotrophs

Free living nitrogen fixation evolved before symbiotic nitrogen fixation (Boyd & Peters, 2013). Free-living diazotrophs are more diverse due to the diverse and variable conditions in which they thrive (Dos Santos, Fang *et al.*, 2012; Raymond, Siefert *et al.*, 2004). Although they are found in various environments, the majority of nitrogen fixation occurs in the rhizosphere due to the availability of carbon from root exudates. Environmental factors such as high carbon availability, less diversity of carbon source and compounds, low oxygen concentration, low potential of *nifH* downregulation by combined nitrogen (Reed, Cleveland *et al.*, 2011), high phosphorous supply and accessibility and high micronutrient supply and accessibility can be considered as ideal conditions for maximum nitrogen fixation (Smercina, Evans *et al.*, 2019). *Azospirillum* is

one of the best characterized free-living diazotrophs that is closely associated with plants, having capacity to infect plants. It fixes nitrogen in microaerobic and nitrogen limiting conditions (Steenhoudt & Vanderleyden, 2000).

1.3.3 Nitrogenase enzyme

Diazotrophs express a unique enzyme complex, nitrogenase, that is responsible for nitrogen fixation. There is a high diversity of diazotrophs, however, only three known forms of nitrogenase enzyme (Nif, Vnf and Anf) are involved in BNF (Smercina, Evans *et al.*, 2019). Nitrogenase is assembled of two major components, a heterotetramer MoFe protein (NifD-K), which uses the electrons provided to reduce N_2 to NH_3, and homodimer Fe-only protein (NifH), which has reducing power and supplies electrons from $NADH+H^+$. *nifH*, *nifD* and *nifK* encode the nitrogenase subunits, while there are other *nif* genes involved in assembly and incorporation of iron and molybdenum atoms into the nitrogenase subunits. Among them, NifB is an extremely unstable and oxygen sensitive component and has been used to study the diversity and phylogeny of naturally occurring diazotrophs (Arragain, Jimenez-Vicente *et al.*, 2017). For Vnf and Anf complexes, the D-K heterotetramer has Fe or Fe-V complexes in place of Mo-Fe.

1.3.4 Methods to determine nitrogen fixing capability

Assays like $^{15}N_2$ incorporation for direct, and acetylene reduction assay for indirect assessment (Unkovich & Baldock, 2008), and *nifH* gene amplification (Gaby & Buckley, 2012) are commonly used to determine the nitrogen fixing capability in either soil samples or in pure cultures of bacteria.

1.3.4.1 Biochemical assays

The stable isotope $^{15}N_2$ is used in the confirmation of endophytic biological nitrogen fixation activity in non-legumes (Chalk, 2016). Samples incubated with $^{15}N_2$ are then dried and the ^{15}N-incorporated into biomass quantified by mass-spectrometry (Fry, 2007). It is often used in quantitative estimation of the proportion of nitrogen obtained from the atmosphere. Nitrogenase enzyme is responsible for reduction of dinitrogen, and has promiscuous affinity for acetylene which gets reduced to ethylene (Dilworth, 1966). Both acetylene and ethylene are quantified by gas chromatography. Thus, acetylene reduction assay is used to determine the activity of nitrogenase enzyme due to its sensitivity and lower cost.

Utilization of tetrazolium chloride (TTC) to test cell respiratory activity through the bacterial cell's reductase activity has long been used. Water-soluble clear TTC gets reduced to formazan, an insoluble red compound, and gets trapped inside the cell, which was used in the assessment of *Lactobacilli* reducing capacity (Saide & Gilliland, 2005). The indirect method for determining nitrogen fixation can be done by assessing visible growth in a nitrogen-defficient soft agar containing TTC. The appearance of red zones indicates bacterial activity. Nitrogenase enzyme is sensitive to oxygen and this method can also be employed to determine the growth of diazotrophs with respect to oxygen sensitivity, by measuring how deep under the surface bacterial activity occurs.

1.3.4.2 Genetic assay

A criterion of presence of a minimum set of six genes (*NifHDK* and *NifENB*) has been established as a computational prediction for nitrogen fixation, as these code for structural and biosynthetic components (Dos Santos, Fang *et al.*, 2012). However, the full

minimal set of genes may vary within different nitrogen fixing taxa due to various sets of genes that are necessary to complement *nif* genes. However, a minimal cluster of *nif* genes, consisting of an operon of *nifB*, *nifH*, *nifD*, *nifK*, *nifE*, *nifN*, *nifX*, *hesA* and *nifV* under the control of sigma (70)-dependent promoter of *nifB*, was transferred to *E. coli* and was able to synthesize active nitrogenase (Wang, Zhang *et al.*, 2013). Although there are various genes involved in nitrogenase enzyme function and structure, *nifH* is the one marker gene used in primary identification of nitrogen-fixing Bacteria and *Archaea*. Primer designing for *nifH* gene amplification to be used in PCR for identification of diazotrophs has not yielded in any ideal primer sets. For example, fifteen out of 51 universal primers could hit 90% or more of the *nifH* sequences publicly available (Gaby & Buckley, 2012).

1.4 Methods to study bacterial interactions

Interaction between the individual units can be studied using various technologies either at the physical or metabolic level. The first step to study interactions, such as *Bacillus-Streptomyces* co-culture, is to separate these two species into their pure cultures, followed by studying their individual characteristics, followed by co-culture characteristics.

1.4.1 Separation techniques

Sub-culturing single colonies multiple times on agar surfaces to get pure cultures of microorganisms is used regularly in microbiological studies. However, the *Streptomyces-Bacillus* co-cultures in this study were not amenable to this approach. Alternative separation approaches can be devised. For example, *Bacillus* produces

endospores that can survive extreme conditions such as high heat, due to its structure that consists of exosporium, layered coats, outer membrane, cortex, inner membrane and core (Setlow, 2006). *Streptomyces* produces exospores which are not tough enough to resist wet heat above 65^{o}C as they have peptidoglycan-based spore walls usually 1.5 to 2 times thicker than vegetative cells (Kalakoutskii & Agre, 1976), unlike endospores with many layers. Using the difference in spore stability of two species, heat can be used to separate *Bacillus* from *Streptomyces*, provided that some of the *Bacillus* cells form mature endospores. Each spore is responsible to give rise to a single colony upon growth in suitable media. The principle behind spore germination can be used to separate or remove *Streptomyces* from *Bacillus* after serially diluting heat-treated suspensions and plating onto culture medium.

The isolation of *Streptomyces* from the nature having diverse microbes has always been complicated due to its slower growth rate. Two basic approach was used previously to isolate *Streptomyces*, formulating media that was preferably utilized by *Streptomyces* only, and adding antibiotics to inhibit other microbes. The characteristics of formation of branched hyphae was utilized to devise a new technique for isolation of actinomycetes from its natural habitat. Cellulose ester membrane filter of pore size 0.22- to 0.45-µm was overlaid in an agar medium and inoculation of mixture of microbes was done on top of filter membrane. This restricted growth of non-actinomycete, whereas, formation of mycelia penetrated the filter pores to let the actinomycetes grow (Hirsch & Christensen, 1983).

Streptomyces spores has been used for reliable long-term storage of strains, conjugation protocols, and various genetic and strain production purposes. *Streptomyces*

cultures are grown in solid media for 4-6 days, and spores are rubbed from the surface using spreader or loop. Cotton filters packed in 50 ml conical screw top vial are used to filter spores from vegetative hyphae of *Streptomyces* (Shepherd, Kharel *et al.*, 2010).

1.4.2 MALDI-imaging mass spectrometry

Studying translation of genomic information to proteins and is termed proteomics. The use of matrix-assisted laser desorption ionization (MALDI) as mass spectrometry imaging is not limited to recording the mass spectrum of particular molecules but has entered the field of tissue-based research to analyze tissue specimens. Various analytes such as proteins, peptides, lipids, drugs and their metabolites can be analyzed to show correlation of molecular information in these kinds of research (Singhal, Kumar *et al.*, 2015). This approach can be used to visualize the molecular components exchanged during the interaction of bacterial species. This can help to infer the necessity of the interaction. Using this technique, a study was done on *Streptomyces coelicolor*, to view secreted chemicals when interacting with five different species of actinomycetes, and the result suggested that the response of *S. coelicolor* in each interaction was unique (Traxler, Watrous *et al.*, 2013). Also, spectral networking can be used to analyze the compounds secreted by two species in isolation and in co-culture.

1.4.3 Quantitative qPCR

Quantitative qPCR is based on the principle a specific or non-specific detection chemistry, that allows the quantification of the amplified product (Torres-Corral, Fernandez-Alvarez *et al.*, 2019). The amount detected at a certain point of the run is directly related to the initial amount of target in the sample. The most common applications of quantitative PCR are gene expression analysis, pathogen

detection/quantification and microRNA quantification (Kim, Lim *et al.*, 2013). Analysis takes the Ct (cycle number) value, at the point when the signal is detected above the background and the amplification is in exponential phase. Two types of reporter dyes are commonly used; TaqMan and SYBR Green (Fig 4).

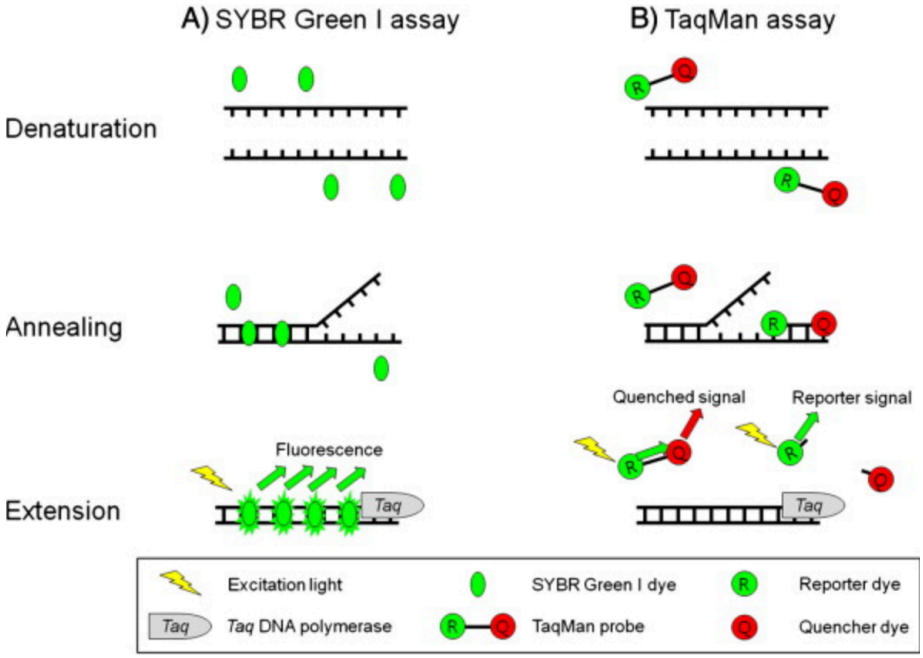


Fig 4: Steps involved in amplification of target DNA sequences in qPCR reaction using SYBR Green reporter dye (A) and TaqMan reporter dye (B) (From (Kim, Lim *et al.*, 2013)).

1.4.4 Transformation with reporter genes

Reporter genes are regulatory sequences which provides a measurable signal as output when introduced into a biological system (Wood, 1995). Plasmids encoding various fluorescent proteins are available and can be customized to make it suitable for particular bacterial species by adjusting the parameters such as codon optimization and adjusting methylation pattern, along with the incorporation of strong promoters that are

suitable for strong expression in the desired host (Moore, MacDonald *et al.*, 2018). Plasmids are one of several components having significant role in shuffling genes among the members of Bacteria, Archaea and Eukarya through horizontal gene transfer (Sorensen, Bailey *et al.*, 2005).

The easy expression of green fluorescent protein (GFP) in heterologous hosts, without requirement of jellyfish-specific cofactors, made it popular for studying protein localization and gene expression in both prokaryotes and eukaryote (Chalfie, Tu *et al.*, 1994). Wild-type GFP is excited with UV or blue light at 396 nm or 475 nm and emits green fluorescence at 508 nm (Liu, Germaine *et al.*, 2010). Six new fluorescence proteins were discovered from *Anthozoa*, two of those had spectral characteristics different from GFP, with one emitting at yellow and another at red wavelengths (Matz, Fradkov *et al.*, 1999). Several other reporter genes that are frequently used include *lacZ*, *lucFF, luxAB*, and *luxCDABE*. The bioluminescence gene lux codes for luciferase, a light that can be labeled and has been successfully used as a reporter for pollution detection (Heitzer, Malachowsky *et al.*, 1994).

High-resolution confocal microscopy has been employed for determining autofluorescence spectrum of naturally occurring pigments and other molecules in the microbes under study. Assessing the autofluorescence can help a researcher to choose the correct spectrum of fluorescence plasmid for visualizing the interaction in wild vs transformed cells. Colocalization determines whether two or more fluorescent signals are in very close positions in the observed specimen (Zinchuk, Zinchuk *et al.*, 2007). Colocalization using confocal microscopy is widely used to investigate the interaction

between a specific drug and the tissue. Such visualization techniques can make the direct observation of physical interaction between the species in the co-culture possible.

Transformation can be challenging while working with wild strains. Various methods have been established such as using polyethylene glycol, electroporation, using protoplasts and chemically competent cells. Laboratory strains can be transformed readily, whereas, wild strains isolated from the environment can be laborious (Nijland, Burgess *et al.*, 2010).

1.4.4.1 Electroporation

Electroporation is used to enhance transformation rates and has been widely used in bacterial cells as well as eukaryotic cells. Improved transformation efficiency using electroporation was achieved by optimizing parameters such as growth phase, ionic strength, concentration and size of plasmid into *B. subtilis* PB1424 (10^4 transformants/μg plasmid) (Brigidi, Derossi *et al.*, 1990). Growth medium is also crucial for obtaining efficient transformation and may vary with the bacterial species being transformed. *B. subtilis* ZK showed greater transformation efficiency in LBSP and lower in NCM medium, whereas in contrast, *B. amyloliquefaciens* was transformed with higher yield in NCM medium (Zhang, Ding *et al.*, 2015). This indicates that the ionic composition of the medium should be considered. Mid-exponential to late-exponential phase of growth is generally suitable for transformation as those cells are actively dividing. Wall weakening agents such as glycine, DL-threonine, Tween 80, and ampicillin are used which helps in increasing the permeability of the cell wall.

1.4.4.2 Transformation of protoplasts:

Transformation of *B. subtilis* protoplasts with plasmid DNA, and regeneration of the transformed protoplasts in selective media has been achieved using the simple media c-R5 and DM3 (Puyet, Sandoval *et al.*, 1987). Lysozyme is used to remove the peptidoglycan layer of the bacterial wall. The suspension has to be isotonic to the bacterial cytosol such that protoplasts are stable. Assessment of protoplasts can be done by phase microscopy. Generally, regeneration of protoplasts may take 10-14 days. Sometimes, the protocol has to be optimized according to the nature of the strain that has to be transformed that includes factors such as agar density and tonicity of the medium. For example, an alkaliphilic *Bacillus* strain was transformed using Puyet *et al.* methods, however, protoplast regeneration was done using 5% agar (hard agar regeneration medium) (Gao, Xue *et al.*, 2011).

1.4.4.3 Transformation by natural competence:

Natural competence is a phenomenon observed in *B. subtilis*, where a fraction of cells are able to uptake DNA/plasmids from their environment *via* DNA uptake (Boonstra, Vesel *et al.*, 2018). Presence of competence machinery (at least 13 competence proteins), activated by the transcription regulator ComK makes the process of natural competence possible and takes place by firstly transferring DNA through the cell wall (ComGA), binding dsDNA to membrane-bound ComEA protein and cleavage of DNA by NucA endonuclease into shorter pieces (Kaufenstein, van der Laan *et al.*, 2011).

1.4.5 Fluorescence *In Situ* Hybridization (FISH):

Artificially manufactured DNA probes labelled with a fluorescent dye are used to hybridize specific rRNA targets (DeLong, Wickham *et al.*, 1989). This technique can be used to determine the colocalization of *Bacillus* and *Streptomyces* in a co-culture by labelled their 16S rRNA with probes bound to different fluorescent dyes. However, effective permeabilization for passage of probes and reagents has been the major drawback to utilize this technology in this study, as effective reagent equilibration is very crucial for the process. It has been found that effectiveness of FISH varies a lot and variability depends mostly on methodological factors (Bouvier & Del Giorgio, 2003).

1.5 Whole Genome Sequence Analysis:

The idea of knowing each nucleotide present in the genome of a bacterium is overwhelming. A genome comprises single or double stranded nucleic acids that store the complete genetic information of an organism or a cell. Due to the advancing technology, we have been able to sequence the whole genome of desired species, even up to single cell (Tang, Huang *et al.*, 2019). However, determining the precise sequence with minimum error has been a huge challenge, and requires effective technology with increased accuracy, throughput and sequencing speed. Sequencers generates reads, consisting sequences with range of length, that are far shorter than the size of the genome (Giani, Gallo *et al.*, 2020).

Three major revolutions have happened in the first two decades of bacterial genome sequencing; whole-genome shotgun sequencing, high-throughput sequencing and single-molecule long-read sequencing (Loman & Pallen, 2015). There are many techniques available for sequencing of whole genomes, each having their own bright side

and shadowed side. Next Generation Sequencing (NGS) overtook the famous Sanger sequencing due to its features such as *in vitro* construction of the sequencing library, *in vitro* clonal amplification, array-based sequencing (multiplexing) and solid phase immobilization of DNA (beads and glass substrate) (Cao, Fanning *et al.*, 2017). The *de novo* genome assembly is the process of deducing the complete genome by overlapping the reads. As the time advances, lots of genome sequencing are been done, which uses reference genome to guide resequencing in the same species (Fig 5).

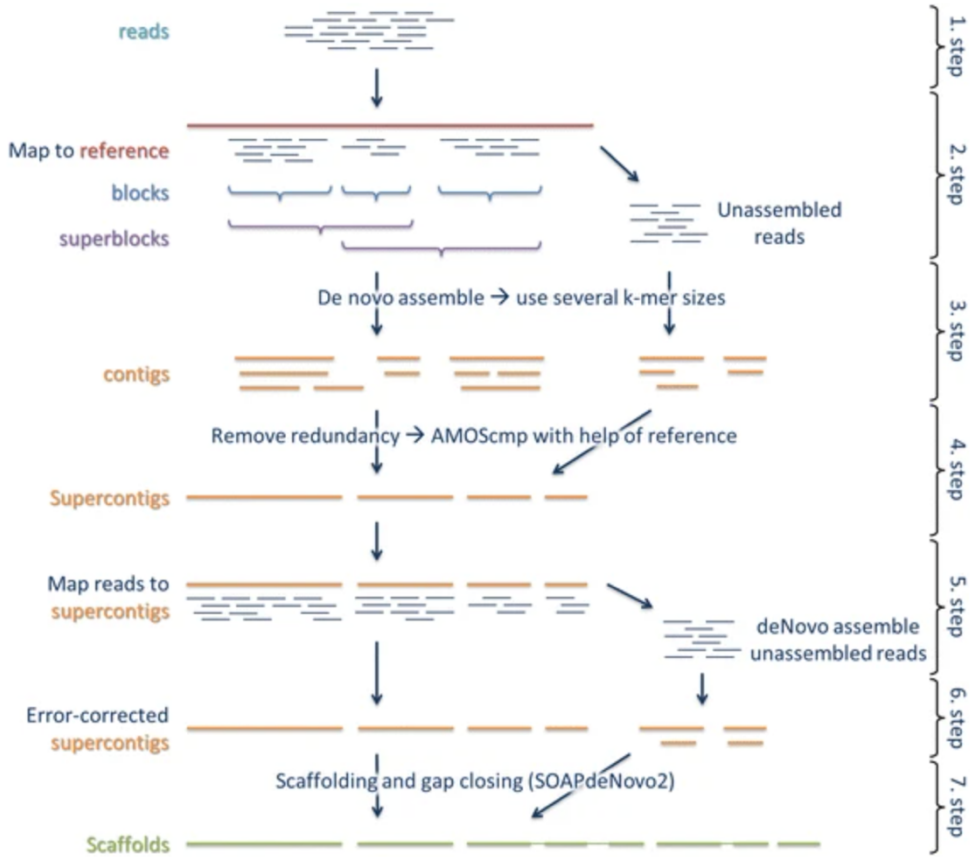


Fig 5: Steps illustrating reference-guided de novo assembly. Trimming of raw reads (Step 1) followed by mapping against reference (Step 2) gives blocks with continuous reads. Superblocks are result of combination of blocks. All unmapped reads and superblocks ae separately *de novo* assembled (Step 3) giving rise to contigs, which is merged into supercontigs (Step 4). Reads are mapped back to supercontigs and additional supercontigs are assembled with unmapped reads (Step 5). Error correction (Step 6) and gap closing (Step 7) are final steps. (From (Lischer & Shimizu, 2017))

Illumina DNA sequencing, the currently dominant technology, generates large numbers of short reads generally used to produce accurate genome assemblies. The major differences between technologies that generate long reads and short reads is sequencing

error. Illumina reads are short (500 bp or less) and less error prone but produce fragmented genome assemblies. Long reads up to tens of thousands can be generated using Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT). These days both the short read and long reads are combined for better assembly of genome with minimum sequencing error using modules such as a Unicycler (Wick, Judd *et al.*, 2017). Assembly of shotgun sequence reads into longer contiguous genomic sequences can be done by two approaches; using *de novo* assembly where sequence reads are compared and overlapped to yield longer sequences, and using reference-based approach where each read is mapped to the reference genome (Ng & Kirkness, 2010).

Assembled whole genome sequences can be used to investigate the possible solutions to research questions. Genotypic approaches to investigate the taxonomy of species have been possible. The 16S rRNA gene sequence is widely used to construct phylogenetic trees in the bacterial and archaeal kingdoms. Sometimes, the 16S rRNA genes are too conserved and not enough to determine the relationship within closely related species. For example, taxonomic relationship of reference *Streptomyces* at species level was better resolved using Multilocus Sequence Typing (MLST) than 16S rRNA method (Guo, Zheng *et al.*, 2008). MLST entails the sequencing of several conserved genes of all isolates, concatenating them and treating the resulting sequence as one gene for phylogeny. Bioinformatics tools such as KEGG (Kyoto Encyclopedia of Genes and Genomes) can be used to analyze genome sequences by characterizing gene functions and reconstructing KEGG pathways (Minoru Kanehisa, Sato *et al.*, 2016). KEGG offers knowledge base for linking genomes to biological functions, however it requires its own annotated genes (KO ids) to perform pathway analysis. Taxonomy classifier, Kaiju, is a

novel program that classifies sequence by comparing to reference database of microbial proteins (Menzel, Ng *et al.*, 2016). It uses maximum exact matches (MEMs) on the protein level, in contrast to traditional local alignment, which makes the program sensitive and fast.

1.6 Conclusion

The aim of the research in this thesis was to characterize the interaction of three *Streptomyces* isolates that grow under nitrogen-free conditions, and are associated with *Bacillus*. The separation of *Bacillus* was achieved due to formation of endospores, but isolation of *Streptomyces* into purity could not be achieved despite multiple approaches. Several separation techniques have indicated that *Bacillus* is very closely and tightly associated with *Streptomyces,* that it seems to either have an endophytic relationship or stay as hitchhiker. Genome sequences can be helpful to determine the possible ways of the interaction between the two partners, by analyzing the pathways that may interlink between the two.

2. Chapter I: Characterization of *Bacillus* that co-occur with *Streptomyces* isolates

2.1 Introduction

In natural environments, bacteria are seldom found as single species populations, rather as communities comprising various different species leading to interactions within or among species (Stubbendieck, Vargas-Bautista *et al.*, 2016). Interactions between two bacteria, two fungi, bacteria and fungi, fungi and plants, bacteria and plants, and among bacteria, fungi and plants are widely found (Braga, Dourado *et al.*, 2016). Insights into properties and behavior of bacteria have been derived primarily through single or pure-culture approaches. Bacterial cultures appearing as pure-culture may carry several other species along, that have been interacting closely in the natural environment, often known as hitchhikers (Lozano, Bravo *et al.*, 2017). Such hitchhikers can be challenging to obtain into single culture.

Endophytic associations of bacteria inside eukaryotic host cells are well characterized, with bacterial endosymbionts found in fungi (Bonfante & Desirò, 2017), insects (Kaltenpoth, 2009) and plants (Khare, Mishra *et al.*, 2018). To date there is only one well characterized example of a bacterial endophyte occurring in another prokaryote-*Moranella endobia* (γ-Proteobacteria) occurring inside *Tremblaya princeps*, a ß-Proteobacteria. *T. princeps* occurs in bacteriocyte cells of the mealybug *Planococcus citri* and has resisted culture to date (Zachar & Boza, 2020). The genomes of both *T. princeps* and *M. endobia* are highly reduced, with the species relying on each other and the eukaryotic host for core metabolic requirements (Lopez-Madrigal, Balmand *et al.*, 2013;

Lopez-Madrigal, Latorre *et al.*, 2013). There are no reliable reports of bacterial endophytes hosted in other prokaryotes.

Syntrophic interaction found in intricate microbial communities are the result of the environmental structure and its demands. An example of syntrophic interactions among heterotrophs and autotrophs was demonstrated using genomics-informed proteomic data in an acetate-amended, sediment-hosted subsurface microbial community (Handley, VerBerkmoes *et al.*, 2013). Soils of prairies have deficient nitrogen supply, as they lack symbiotic nitrogen fixation due to sparsity of legumes. In such conditions, free living diazotrophs have a huge role to play in replenishing fixed nitrogen to the ecosystem. Unique interactions may occur in such conditions between both nitrogen-fixing and fixed nitrogen consuming individuals.

Biological Nitrogen Fixation (BNF) is the process of conversion of atmospheric dinitrogen into ammonia which is done by symbiotic as well as free-living diazotrophs (Galloway, 1998). The ability to reduce atmospheric dinitrogen to ammonia (Reed, Cleveland *et al.*, 2011) is unique among prokaryotes, including green Sulphur bacteria, *Firmicutes*, *Cyanobacteria*, *Alphaproteobacteria*, *Betaproteobacteria*, *Deltaproteobacteria* and *Gammaproteobacteria* (Dixon & Kahn, 2004). Grasslands such as natural prairies have few leguminous plant, where symbiotic nitrogen fixation occurs, and in such amount that could contribute to the total fixed nitrogen for sustaining of that ecosystem. Hence, the prevalence of free-living diazotrophs in the rhizosphere is expected. Diazotrophs express a unique enzyme complex, nitrogenase, that is responsible for nitrogen fixation. Advancement in integrated technology has made it possible, to characterize and identify, biochemical and genetic properties of nitrogenase complex.

Biochemical assays such as $^{15}N_2$ incorporation for direct and acetylene reduction assay for indirect assessment (Unkovich & Baldock, 2008), and genetic assay using *nifH* gene amplification (Gaby & Buckley, 2012) are well-established for determination of nitrogen fixation.

A recent project to isolate free-living nitrogen fixing bacteria from a natural prairie at Sioux Prairie, Colman SD, yielded multiple *Streptomyces*. These grew both on nitrogen-free agar (NFA) and in nitrogen free minimal broth in sealed containers to exclude atmospheric ammonia. Genome sequencing of three isolates revealed the presence of a second 16S rRNA gene sequence aligning with *Bacillus*. Despite aggressive restreaking across both nitrogen-free and rich agar, no *Bacillus* colonies were obtained. Microscopy of NFA-grown cultures did not reveal regular rod-shaped Gram positive cells (Alshibli, 2018).

In this study, we isolated and characterized *Bacillus* from cocultures with *Streptomyces* that were obtained on nitrogen-free agar. We have analyzed its nitrogen fixing ability using acetylene reduction assay, $^{15}N_2$ isotope incorporation assay and amplification of *nifH* by PCR. The phylogenetic relationship with other known species of *Bacillus* genus has been determined.

2.2 Materials and Methods

2.2.1 Sources of cultures

Cultures were originally isolated from Sioux Prairie of South Dakota 231st Street (44.2'8" N 96° 46' 0" W 1720 ft Elevation) onto nitrogen free agar (NFA) by a previous student, Nabhilah Alshibli. Isolates growing in nitrogen free media 'NFA' (Dahal,

NandaKafle *et al.*, 2017) with addition of four carbon sources (glucose, mannitol, arabinose and malic acid) and noble agar (BD Difco™), and incubation under reduced oxygen using gas-pack (Gaspack EZ Campy, BD) were sub-cultured over R2A (BD Difco™) to obtain pure cultures. The total nitrogen was determined by Southeast Environmental Research Center (Florida International University), which showed the water to prepare NFA had 2.57 µmol/L, liquid NFM had 6.42 µmol/L total nitrogen, and no nitrogen was detected in noble agar (detection limit).

Out of 110 isolates displaying *Streptomyces* colony morphology, the 3 fastest growing isolates were chosen and named 321I, 34I and 40I (Fig 6). While they appeared as uniform colonies of *Streptomyces* during subculture, whole genome sequencing revealed the presence of 16S rRNA genes homologous to *Bacillus* (Alshibli, 2017). *Herbaspirillum seropedicae* ATCC 35892 was used as positive control for free-living nitrogen fixation throughout.

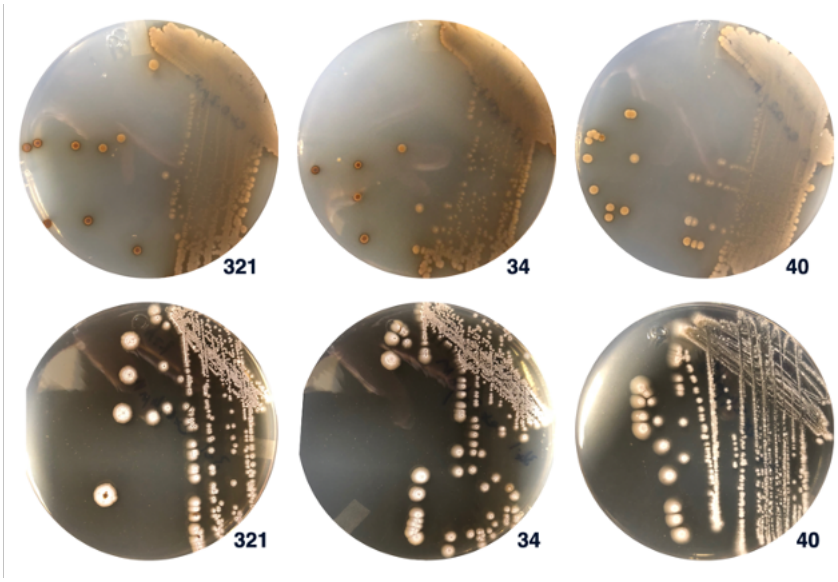


Fig 6: Three *Streptomyces* isolates growing on NFA (top) and R2A (bottom) for a week.

2.2.2 Approaches to obtain *Bacillus* into single culture

Genetic evidence suggesting presence of *Bacillus* co-occurring with the apparently pure *Streptomyces* prompted attempts to obtain this into single culture.

2.2.2.1 Selection of endospores using heat

Isolates were grown in R2A broth (0.5 g/l yeast extract, 0.5 g/l protease peptone no. 3, 0.5 g/l casamino acids, 0.5 g/l glucose, 0.5 g/l soluble starch, 0.3 g/l sodium pyruvate, 0.3 g/l dipotassium phosphate and 0.05 g/l magnesium sulfate (Reasoner & Geldreich, 1985)) for 3 d and pellets were obtained by centrifugation at 5428 X g for 7 min in 1.5 ml tubes. Pellets were resuspended by vortexing in R2A broth and incubated at 90ºC for 30 min. The content was vortexed and then spread onto R2A plates and incubated at 28ºC for 2 d. Single colonies were sub-cultured on R2A.

2.2.2.2 Glass slide press method to select free *Bacillus* colonies

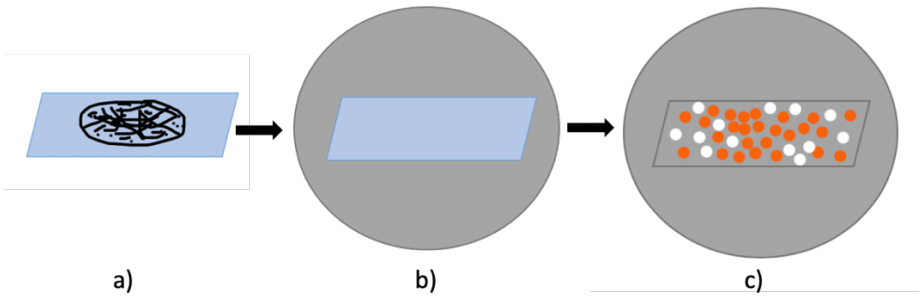


Fig 7: Slide press method to isolate free *Bacillus* from *Streptomyces* colonies grown on R2A. Mature Streptomyces colonies were placed onto sterile microscope slides (a). Slides were inverted onto R2A, pressed down (b) and removed (c).

The principle of preparing a smear for microscopy was used to isolate free *Bacillus* from 7 d old *Streptomyces* colonies on R2A. Water (10µl) was aliquoted onto a

sterile glass slide and a single colony of *Streptomyces* was pressed onto glass slide to make smear (Fig 7a). Using sterile forceps, the glass slide with smear was inverted onto a R2A plate and pressed gently to transfer the bacterial cells to agar (Fig 7b) and removed. Plates were incubated at 28^{o}C for 3 d to obtain *Streptomyces* and *Bacillus* colonies (Figure 7c). Growth appearing free from Streptomyces was sub-cultured to purity on R2A.

2.2.3 Biochemical evidence for nitrogen fixation

2.2.3.1 Acetylene reduction assay

NFA agar slants in glass tubes were inoculated in three replicates with *Bacillus* isolates (321B, 34B and 40B), along with *H. seropedicae* as positive control and *Salmonella typhimurium* as negative control. Incubation was done at 28^{o}C for 5 d, and the headspace of the sealed glass tubes was supplied with 1 ml of acetylene gas, and incubated for 5 h. The amount of acetylene reduced into ethylene was measured using a gas chromatograph (GC-MS) (Agilent Technologies), using ethylene gas for calibration.

2.2.3.2 $^{15}N_2$ incorporation assay

Nitrogen free medium (20ml) was inoculated with the *Bacillus* isolates (321B, 34B and 40B), *H. seropedicae* and *S. typhimurium* in serum bottles with tight rubber caps. Air in the headspace was replaced by 5ml $^{15}N_2$ and incubated at 28^{o}C for 4 weeks. Bacterial cells were collected by centrifugation at 5428 X g for 10 min. Cells were lyophilized using a speed vac (Thermo Scientific Savant ISS110 SpeedVac) overnight at low temperature setting.

2.2.3.3 Oxygen sensitivity test

TTC (2,3,5 triphenyl tetrazolium chloride) (0.5 g/l) was filter sterilized and added to soft NFA with 0.2% noble agar. *Bacillus* (321B, 34B and 40B) and *H. seropedicae* were grown on R2A for 2 d at 28^{o}C. Single colonies were picked, washed in sterile distilled water 3 times, and suspended in 300µl of water. In a glass tube, 10ml of soft NFA was inoculated with 100µl of bacterial culture and incubated at 28^{o}C for 5 d. All cultures were prepared in triplicate. Red coloration indicated reduction of TTC, indicative of bacterial respiratory activity.

2.2.4 DNA extraction and PCR amplification of 16S rRNA and *nifH*

For DNA extraction, cultures were grown in 50ml R2A broth at 37^{o}C overnight in a water bath shaker at 180 rpm. DNA was extracted from 5ml of cultures using the DNeasy UltraClean Microbial kit (Qiagen) and stored at -20^{o}C. Primers used for 16S rRNA and *nifH* PCR amplification are listed in Table 1. Amplification of V1-V3 region of 16S rRNA gene was done by PCR reaction. Reaction volume (30µl) was used that comprised of 0.2µl (5000U/ml) Taq polymerase (NewEngland BioLabs), 3µl (10X) PCR buffer, 0.6µl dNTPs (each 10mM), 0.6µl (10mM) each primer and 2.4µl (25mM) $MgCl_2$. PCR conditions were, initial denaturation at 94^{o}C for 5 min, 30 cycles at 94^{o}C for 1 min, 54^{o}C for 1 min, and 72^{o}C for 1 min, and a final elongation at 72^{o}C for 10 min.

Table 1: Primers used for PCR amplification of 16S rRNA and *nifH* genes.

S. N.	Primer Name	Sequence	Reference
1.	16S rRNA	27F: AGA GTT TGA TCM TGG CTC AG 518R: GTA TTA CCG CGG CTG CTG G	(Weisburg, Barns *et al.*, 1991) (Muyzer, Dewaal *et al.*, 1993)
2.	*nifH*	Pol F: TGC GAY CCS AAR GCB GAC TC Pol R: ATS GCC ATC ATY TCR CCG GA	(Poly, Ranjard *et al.*, 2001)

The partial *nifH* genes were amplified using primers Pol F and Pol R (Table 2) and reaction mixture was same to 16S rRNA PCR amplification as mentioned above. The PCR conditions for *nifH* were, initial denaturation at 94^{o}C for 5 min, 30 cycles at 94^{o}C for 1 min, 55^{o}C for 1 min, and 72^{o}C for 30 s, and a final elongation at 72^{o}C for 5 min.

Agarose gel electrophoresis (1.2% agarose in TBE buffer, 60V, 60min) was used to confirm the amplification in each PCR with 100 bp ladder (NewEngland BioLabs) for reference of band size. Sanger sequencing (GeneScript) was done to determine the sequences of PCR products. Quality of sequences (chromatogram file) were analyzed using MEGAX (Kumar, Stecher *et al.*, 2018) and homology was analyzed using blastn (Altschul, Gish *et al.*, 1990).

2.2.5 Phylogeny of *Bacillus* isolates

Nucleotide sequences of four genes *groEL*, *gyrA*, *polC* and 16S rRNA for reference species were extracted from their genomes on GenBank as described by Dunlap (2019), and concatenated. Each reference strain was selected from each distinct branch of the phylogeny tree constructed by Dunlap (2019), which represents the identification of the correct taxonomy of *Bacillus* species based on genome sequences. Alignments and construction of phylogenetic trees were performed using MEGAX software (Kumar, Stecher *et al.*, 2018). Muscle was used to align multiple sequences that contained four genes concatenated end to end. A Maximum likelihood tree was generated using the general time reversible model. Bootstrapping (1000X) was used to interpret the probabilities that respective groups of species are present in the true phylogeny.

2.3 Results

The three apparent pure cultures (321I, 34I and 40I) consistently formed what appeared as pure cultures characteristics *Streptomyces* on NFA and R2A. *nifH* showed homology to *Herbaspirillum*. The acetylene reduction assay was weakly positive, which indicated the ability of the isolates to fix nitrogen. Whole genome sequencing of those *Streptomyces* isolates showed the presence of *Bacillus* 16S rRNA gene, indicating these isolates are not single cultures, but possibly cocultures of *Streptomyces* with *Bacillus* as hitchhiker (Lozano, Bravo *et al.*, 2017).

2.3.1 Isolation of *Bacillus*

Intensive sub-culturing by streaking of *Streptomyces* isolates (321I, 34I and 40I) on both R2A and NFA yielded only *Streptomyces* colonies. No visible colony of *Bacillus* was obtained. Gram staining of young *Streptomyces* did not show free *Bacillus* while a week-old culture showed the presence of free rods (*Bacillus*).

2.3.1.1 Isolation by heat treatment

Endospores produced by *Bacillus* were able to withstand the application of 90^{o}C moist heat for 30 min, while *Streptomyces* spores and mycelium were not. All three isolates subjected to heat, gave *Bacillus* colonies (321B, 34B and 40B) on R2A, with no outgrowth of *Streptomyces* overtime.

2.3.1.2 Isolation by slide press method

Application of physical force to make smears of *Streptomyces* culture on a glass slide had shown the presence of free rods under light microscopy. The smear on the glass slide when pressed onto R2A and NFA agar yielded both confluent and separate *Bacillus* and *Streptomyces* colonies (Fig 8). *Bacillus* colonies had whitish coloration with soft morphology while *Streptomyces* colonies were yellowish in color with hard morphology. There were a larger number of *Bacillus* colonies on R2A plates than compared to NFA plates as seen visually.

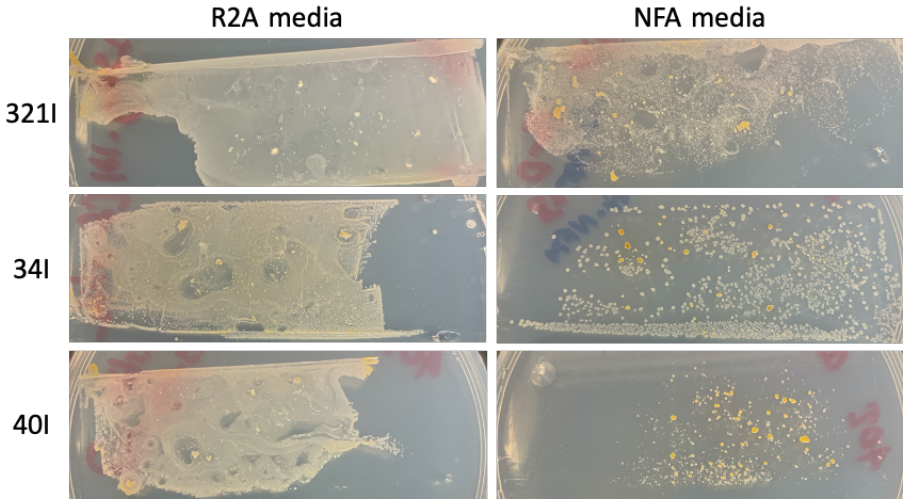


Fig 8: Replica-cultures of glass slide having smear of *Streptomyces* 321I, 34I and 40I pressed onto nitrogen-rich (R2A) and nitrogen-limited (NFA) media. Whitish colonies represent *Bacillus* and yellowish colonies represent *Streptomyces*.

2.3.2 Growth of isolated *Bacillus* on NFA

Isolated *Bacillus* cultures (321B, 34B and 40B) grew on NFA media when incubated at 28°C for 7 days (Fig 9). Single colonies could be obtained which were smaller than grown on R2A (Fig 9).

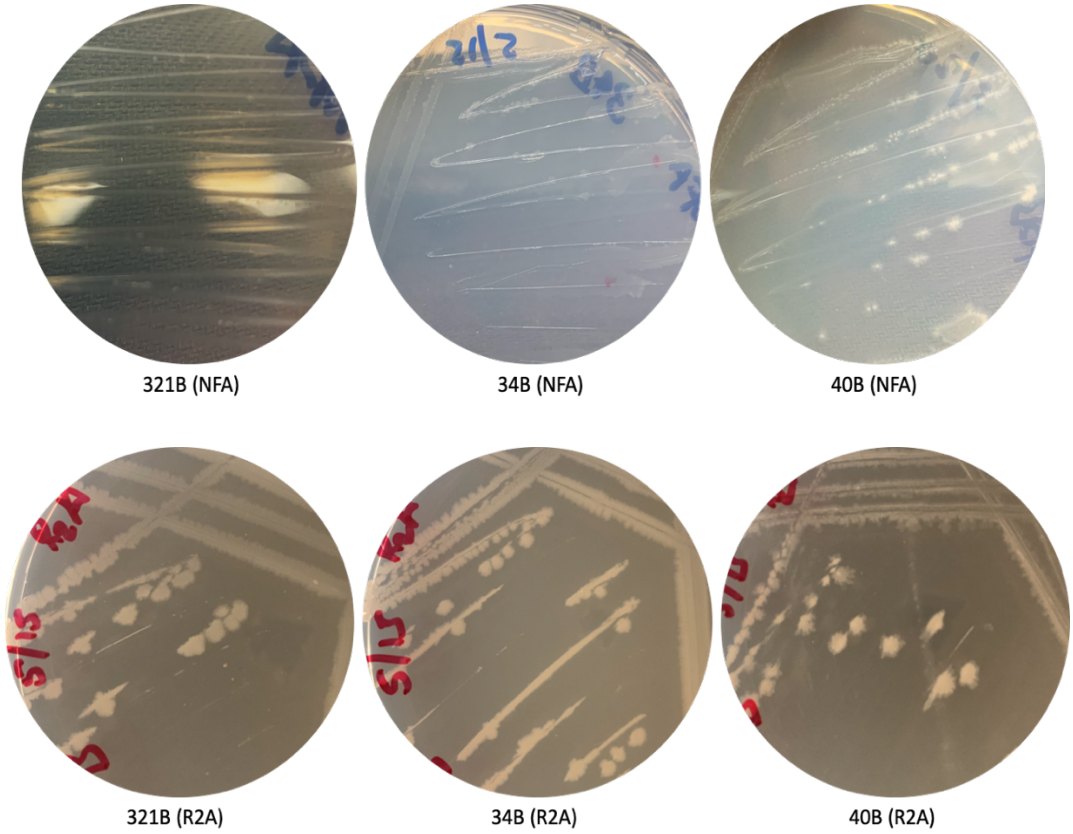


Fig 9: *Bacillus* isolates 321B, 34B and 40B growing on NFA after 7 d of incubation and on R2A after 3 days of incubation at 28ºC.

2.3.3 Biochemical evidence for Nitrogen fixation

Growth of *Bacillus* isolates on NFA suggested the ability of these isolates to fix molecular nitrogen.

2.3.3.1 Acetylene reduction assay

All three *Bacillus* isolates showed ability to reduce acetylene to ethylene. Bacterial nitrogenase has long been known to reduce compounds similar in structure to dinitrogen, including acetylene to ethylene (Stewart, Fitzgerald *et al.*, 1967). Ethylene peaks were observed from headspaces of all three cultures (Fig 10).

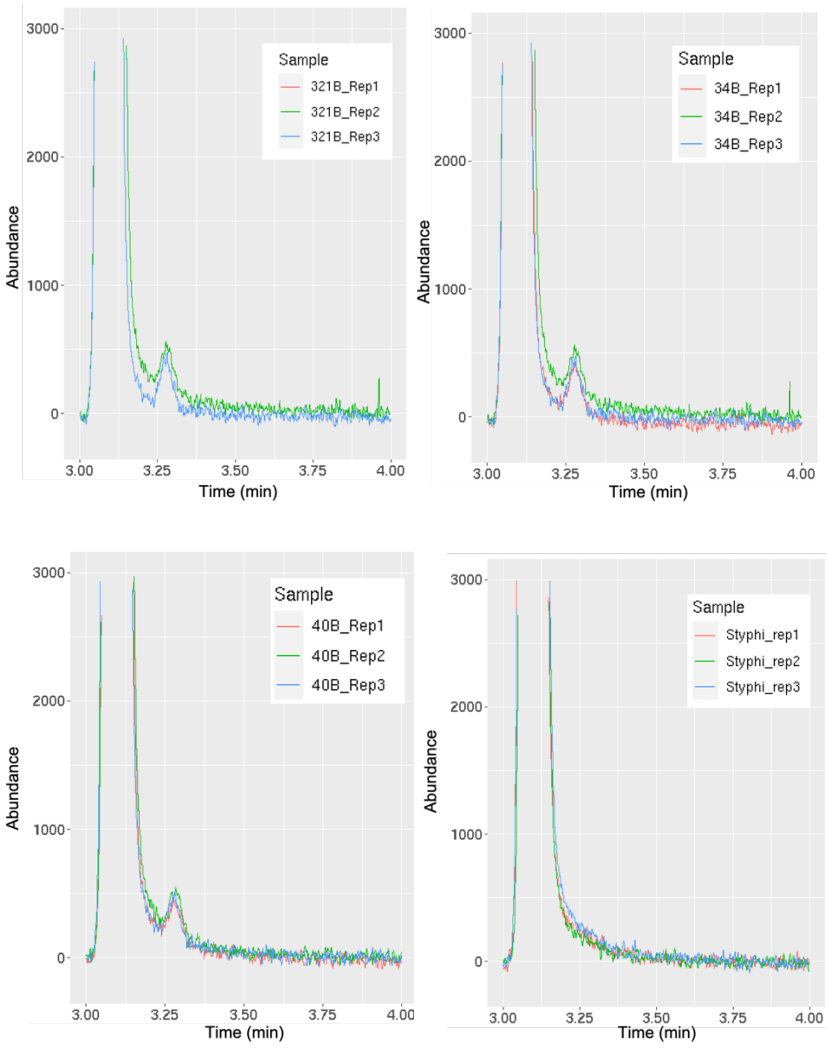


Fig 10: Acetylene reduction assay of *Bacillus* isolates 321B, 34B and 40B, and *S. typhimurium* as negative control cultured on nitrogen free agar in sealed glass tubes for 5d, and supplemented with acetylene.

2.3.3.2 $^{15}N_2$ incorporation assay

Data was not obtained from the Stable Isotope Facility at UC Davis by the time of final thesis submission.

2.3.3.3 Oxygen sensitivity test

Inoculation of the three *Bacillus* and the positive control *H. seropedicae* in soft NFA did not yield visible growth. Respiring bacteria have the ability to reduce TTC to a red pigment due to their reducing activity. This characteristic can be used to localize growth of microbes that do not produce pigments. The growth of *H. seropedicae*, 321B and 40B was observed in semi-solid NFA supplemented with TTC (Fig 11). *H. seropedicae* made a dark ring just below the surface, whereas isolates 321B and 40B made comparatively thicker rings. Isolate 34B did not produce red pigmentation, suggesting it is not able to grow in nitrogen-free soft agar. Oxygen availability by diffusion is inversely related to the depth of soft agar. Growth of *H. seropedicae*, 321B and 40B just below the surface suggests the growth of these bacteria is sensitive to atmospheric oxygen concentration yet requires oxygen for growth. Nitrogenase is an oxygen sensitive enzyme and its activity is best under low oxygen condition (Arragain, Jimenez-Vicente *et al.*, 2017).

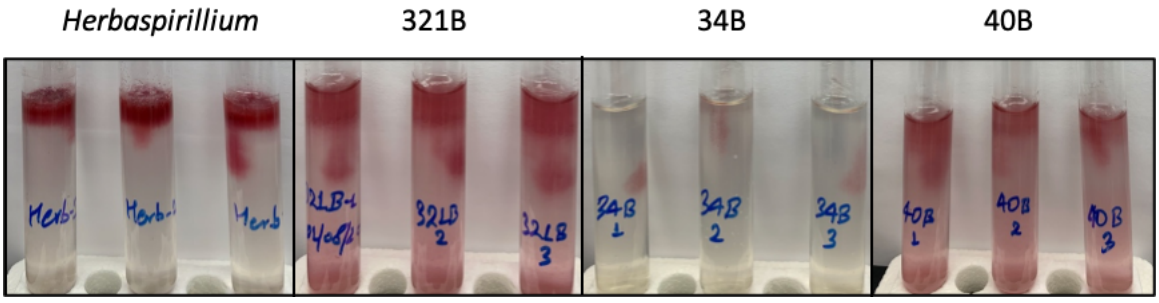


Fig 11: Growth of isolates (321B, 34B and 40B) in semi-solid NFA media visualized by reduction of TTC to red pigment after incubation for 5 d at 28ºC. *H. seropedicae* was used as positive control for nitrogen fixation.

2.3.4 Amplification of *nifH* by PCR

PCR amplification using primers PolF and PolR yielded a band for all three *Bacillus* isolates (321B, 34B and 40B) (Fig 12). The PCR products gave sequences that were homologous to *Herbaspirillum nifH* gene, the alignment is shown in Fig 13. However, the chromatogram of the sequence suggests the poor quality of sequences and did not contain well-resolved peaks (Fig 14). The chromatogram showed that the sequence quality had heterozygous peaks, and more than half of the sequences had broader and less well-resolved peaks. This result questions the validity of the PCR and whether the cultures contained a *nifH* gene. *nifH* could not be found in the whole genome of *Bacillus* 321B and 40B, however, the presence of sequence size similar to *H. seropedicae* while resolving the gel, suggests presence of a possible third bacterium in the interaction. However, *H. seropedicae* is not heat resistant and its 16S rRNA gene was not detected in the whole genome.

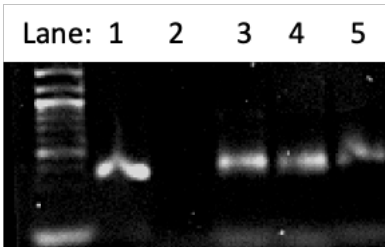


Fig 12: Gel image showing bands of *nifH* PCR with positive control: *H. seropedicae* (1), negative control: water (2), isolates 321B (3), 34B (4) and 40B (5).

Herbaspirillum seropedicae strain RFNB26 nitrogenase iron protein (nifH) gene, partial cds

Sequence ID: FJ829465.1 **Length:** 372 **Number of Matches:** 1

Range 1: 70 to 371 GenBank Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Identities	Gaps	Strand
263 bits(142)	5e-66	256/310(83%)	11/310(3%)	Plus/Plus

```
Query  52   ATACTCTTCGTGGGGTG-AATCCCAAGGGTGACTCGACCCGCCTGATCCTGCACGCCAAG  110
            || ||| |||| || ||  |||||||||| ||||||||||||||||||||||||||||||
Sbjct  70   ATCCTCATCGTCGGCTGCGATCCCAAGGCTGACTCGACCCGCCTGATCCTGCACGCCAAG  129

Query  111  G-CCAGGACACCATCCTGTCGCTGGCCGCCGATGCCGGGTCGGGGGAGGATCTGGAGTTG  169
            | |||||||||||||||||||||||||||||||||||||||||| ||||||||||||| ||
Sbjct  130  GCCCAGGACACCATCCTGTCGCTGGCCGCCGATGCCGGGTCGGTGGAGGATCTGGAGCTG  189

Query  170  GAAAACGTCATGAGGATCGGCTCCAAAAACATCGGTTGCGTCAAATccgggggcccgaaa  229
            ||| ||||||||| |||||||| | | ||||||| |||||||| ||||||| ||||||  ||
Sbjct  190  GAAGACGTCATGAAGATCGGCTACCAGAACATCCGTTGCGTCGAATCCGGCGGCCCGGAA  249

Query  230  cgggggccgggggggcggccgcggccccccggggt-atccccccccgATCAATTTTCTTCGAA  288
            | |||| | |||| | || || | ||    || || ||| ||  ||||||| || || ||||
Sbjct  250  CCGGGC-GTGGG-CTGC-GCTGGCCGT-GGTGTGATCACCT-CGATCAACTTCCT-CGAA  303

Query  289  GAAAGAGGGCGCCTTATAAGGACCCCGATAATTTGTCCAACGAGGGGGTGGGCCAAGTCG  348
            ||| |||||||||| ||| | ||| ||||| |  ||||| |||| | | ||||| | ||||
Sbjct  304  GAA-GAGGGCGCC-TATGATGACACCGATTACGTGTCCTACGACGTGCTGGGCGACGTCG  361

Query  349  CCCGCGGCGG  358
             | ||||||| 
Sbjct  362  TCTGCGGCGG  371
```

Fig 13: Alignment of sequence obtained by amplifying *nifH* by PCR of *Bacillus* 321B using BLASTN.

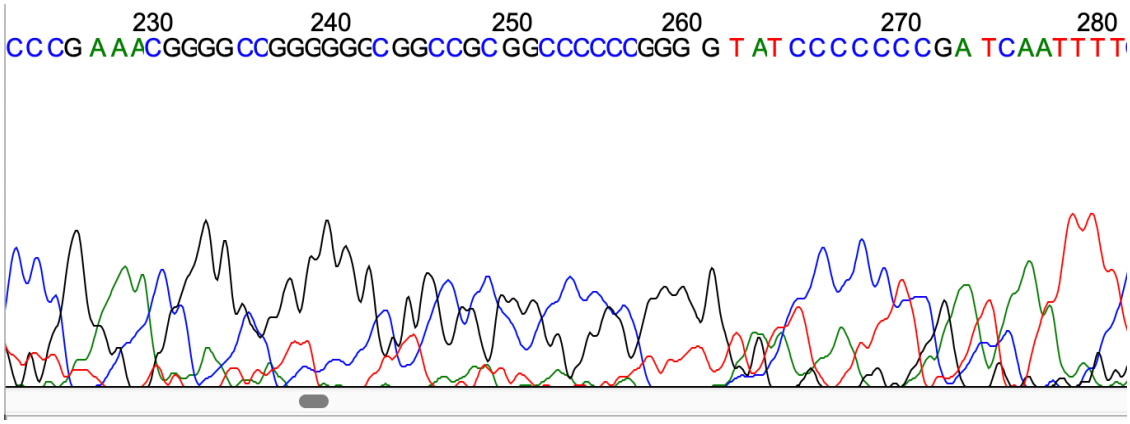


Fig 14: Chromatogram of *nifH* obtained by Sanger sequencing of *Bacillus* 321B. Tracer file was used to visualize the sequence using MEGAX.

2.3.5 Phylogeny of *Bacillus*

Using 16S rRNA gene sequences obtained from whole genome sequencing of 321B and 40B by Illumina, the phylogenetic tree lacked resolution, with several reference sequences clustering very close together. This lead us to opt for a multilocus sequence typing (MLST) approach. Four protein coding genes *rpoA*, *recA*, *gyrA* and 16S rRNA of the reference species were chosen based on the recommendations of (Dunlap, 2019) as the fifth suggested gene *abcD* could not be obtained for all strains. Sequences of the two isolates were obtained from the genome sequence, and sequences for reference strains were obtained from GenBank. The four gene MLST phylogeny of *Bacillus* isolates (321B and 40B) is given in Fig 15. Use of four gene sequences concatenated end to end gave a clearer view among the branches of the phylogenetic tree. Isolate 321B and 40B are very closely related to each other, and they are closely related to *Bacillus subtilis* 168.

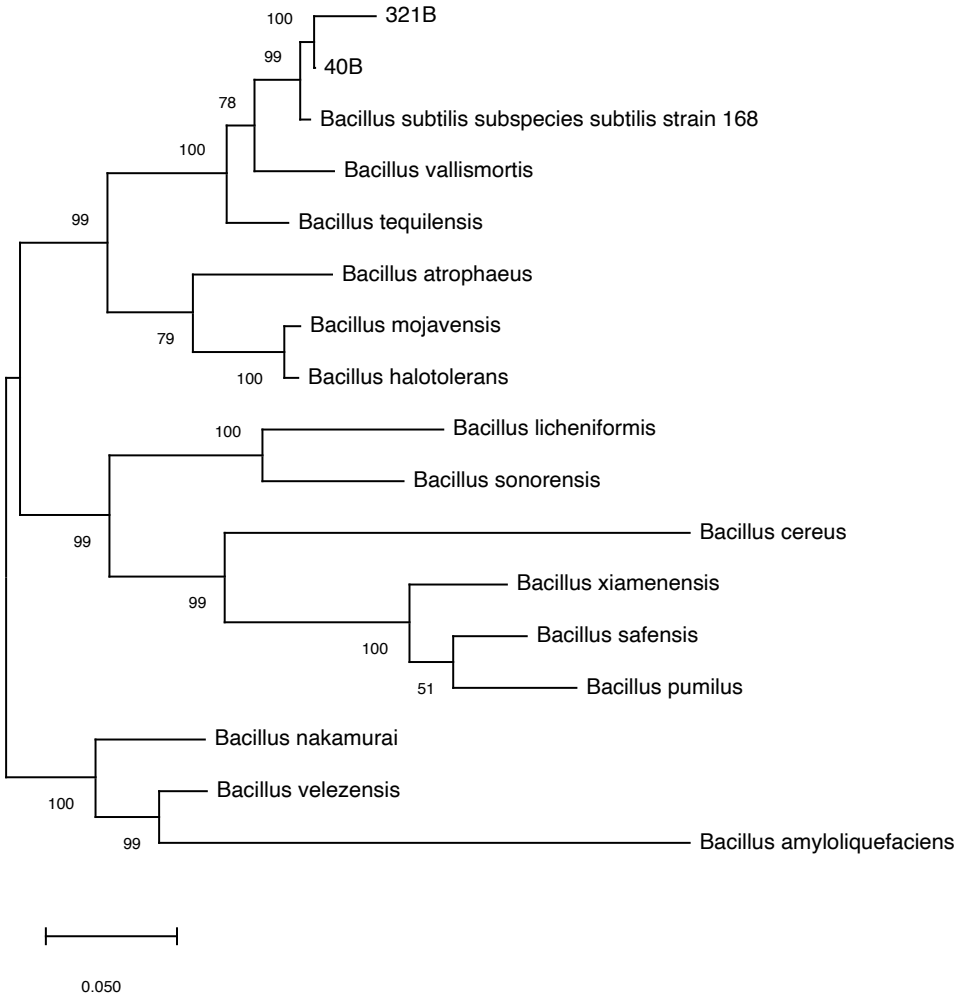


Fig 15: Maximum likelihood phylogenetic tree using concatenated sequences of *rpoA*, *recA*, *gyrA* and 16S rRNA. Numbers represent bootstrap values (%) of 1000 analyses.

2.4 Discussion

In this study, we isolated *Bacillus* from the *Streptomyces* coculture using the heat withstanding property of endospores produced by *Bacillus*. While very much is known about endospores produced by *Bacillus* (Setlow, 2006), there is only scant information on tolerance of exospores produced by *Streptomyces* (Bobek, Šmídová *et al.*, 2017). Based

on the difference of spore structure and composition, we could employ high heat (90^{o}C) to selectively kill all vegetative cells and *Streptomyces* exospores, allowing out-growth of *Bacillus* from its endospores. (Heat-fixed and stained smear of *Streptomyces* under microscopy contained rods that were spread on the glass slide) Physical force applied while making smears of old cultures was able to separate the apparent close association of *Streptomyces* and *Bacillus*, yielding pure *Bacillus* cultures.

nifH is the one marker gene used in primary identification of nitrogen-fixing Bacteria and *Archaea*. Primer designing for *nifH* gene amplification to be used in PCR for identification of diazotrophs has not yielded in any ideal primer sets, for example, 15 out of 51 universal primers could hit 90% or more of the *nifH* sequences publicly available (Gaby & Buckley, 2012). *nifH* PCR of *Bacillus* 321B, 34B and 40B was mostly positive and sanger sequencing gave sequences that were homologous to the *Herbaspirilium nifH* gene. Horizontal gene transfer could be the possible reason for this homology. However, the chromatogram showed that the sequence quality had heterozygous peaks, and more than half of the sequences had broader and less well-resolved peaks. This suggests that the amplification using primers PolF/PolR did not yield good products, and there is no strong support for presence of *nifH* gene in *Bacillus* 321B, 34B and 40B. Yet 321B and 40B grew on NFA and in soft NFA as visualized by TTC reduction. Intriguingly growth in soft NFA occurred a few mm below the surface, indicating oxygen sensitivity under N-free conditions. The isolates grew to the top of soft NFA supplemental with organic nitrogen source. These results are consistent with dependence on an oxygen-sensitive nitrogenase. Acetylene reduction also supported the presence of nitrogenase activity.

There is limitation in the use of 16S rRNA gene, for determination of relatedness among species of the same genus along with lineages of strains which are crucial in the field of global epidemiology. Multilocus sequence typing (MLST) uses multiple housekeeping genes to determine the relatedness among the species or strains with improved resolution (Guo, Zheng *et al.*, 2008; Maiden, Bygraves *et al.*, 1998). Using four genes *rpoA*, *recA*, *gyrA* and 16S rRNA genes based on the recommendation of (Dunlap, 2019), we found 321B and 40B is closely related to *Bacillus subtilis* 168. There are many reports of *Bacillus*, including *B. subtilis* growing on N-free medium, but no biochemical or genetic evidence for nitrogenase activity has been published. *B. subtilis* 168 did not grow to form colonies on NFA in our lab.

The mechanism underlying the ability of *Bacillus* 321B, 34B and 40B to grow on NFA is unresolved. *Rhodococcus erythropolis* grows on nitrogen free medium supported not by nitrogenase but by a high-affinity ammonium transporter AmtB (Yoshida, Inaba *et al.*, 2014). The *Streptomyces* from which the *Bacillus* were obtained grew in N-free liquid medium in sealed container with the ammonium scavenger clinoptilolite (Alshibli, 2018). This speaks against high affinity uptake of low levels of atmospheric ammonia. The two Bacillus isolates sequenced did have ammonium transporter genes, but their affinity for ammonia is not known. Future work should focus on identifying genes expressed only in absence of combined nitrogen to identify components of the mechanism underlying growth in absence of nitrogen.

3. Chapter 2: Characterizing interaction between *Bacillus* and *Streptomyces*

3.1 Introduction

Soil is not a uniform system, rather a heterogenous and dynamic system, due to the composition of its living entity. All kinds of microbes including fungi, worms and insects dwell in the soil, making it a hub for studying interactions between each member (Braga, Dourado *et al.*, 2016). The genus *Streptomyces* of the Actinobacteria phylum is widely prevalent in soil and has been extensively studied, primarily due to its antibiotic producing ability (Challis & Hopwood, 2003). Actinobacteria form one of the largest phyla and are ancient among lineages of bacteria (Ventura, Canchaya *et al.*, 2007). *Streptomyces* is a gram-positive, mycelium- and spore-forming soil bacterium with large genome and multiple gene clusters for secondary metabolite production such as antibiotics, and surfactants. Interaction of *Streptomyces* with other species has been extensively used in the biotech industry to obtain novel secondary metabolites (Hoshino, Wakimoto *et al.*, 2015). *Streptomyces* exists both free-living, and in symbiotic association mostly with eukaryotes. Most of the symbiotic associations of *Streptomyces* are due to its secondary metabolites that provide defense against pathogenic bacteria (Kaltenpoth, 2009; Kroiss, Kaltenpoth *et al.*, 2010; Seipke, Kaltenpoth *et al.*, 2012). The role of secondary metabolites in the target organism is achieved by its ability to modulate transcription due to interaction with receptors or RNA polymerase (Yim, Wang *et al.*, 2007).

Both *Bacillus* and *Streptomyces* are soil dwelling microbes and there is high probability of encounter with each other in this environment. *Bacillus*-*Streptomyces* interactions have been reported in research that evaluates this interaction to enhance antibiotic production (K. J. K. Luti & F. Mavituna, 2011; Shin, Byun *et al.*, 2018), while in some cases, *Bacillus* is found to hinder in the development of *Streptomyces* (Hoefler, Gorzelnik *et al.*, 2012; Schneider, Yepes *et al.*, 2012). Hitchhikers are bacteria that are not visible in colonies incubated for long times (up to 4 weeks) but are carried along with the colony forming bacteria (Lozano, Bravo *et al.*, 2017). *B. cereus* when isolated as model rhizosphere Firmicute was later characterized to have 20 other hitchhikers that play roles in induction of dendritic expansion of *B. cereus*, and robust production of biofilms when compared to individual growth (Lozano, Bravo *et al.*, 2019). *Streptomyces* have a complex developmental life cycle (Fig 2), with four stages of development; spore germination, vegetative or substrate mycelium, aerial hyphae formation, and sporulation (Jones, Ho *et al.*, 2017). *Streptomyces* filaments contain multiple chromosomes per cell which makes it difficult to determine the number of ribosomal copies in each cell.

Endosymbioses among eukaryotes and between prokaryotes and eukaryotes are omnipresent and, thus are well documented (López-García, Eme *et al.*, 2017; Nowack & Melkonian, 2010), whereas, endosymbioses among prokaryotes are rare (Zachar & Boza, 2020). Zachar and Boza (2020) have listed the prokaryotic examples showing endosymbiosis known to date; *Tremblaya princeps*-*Morganella endobia*, tick mitochondrion-*Midichloria mitochondrii* bacterium, dinoflagellate plastid (Cyanobacteia) in *Woloszynskia pascheri* and *Pleurocaspa minor* (Cynanobacteria)-Bacterium. The first three examples represent multilevel endosymbioses of one prokaryote inside another,

located in a eukaryotic cell. The fourth example is the only one reported to have one-one prokaryotic endosymbiosis, however, there is limited data to determine whether the endosymbiosis is stable or some occasional infection, based on a single report (Wujek, 1979). There are no reports on endosymbiosis including *Streptomyces* or any other bacteria.

Study of the interaction between two or more species can be achieved by studying metabolites exchanged using technology such as MALDI-imaging mass spectrometry (Singhal, Kumar *et al.*, 2015). A study on *S. coelicolor* was done to view secreted chemicals when interacting with five different species of actinomycetes and the result suggested that the response in each interaction was unique (Traxler, Watrous *et al.*, 2013). Visualization techniques using fluorescently labelled plasmids to determine colocalization between two or more fluorescent signals in very close positions in the observed specimen (Zinchuk, Zinchuk *et al.*, 2007) is another useful technique to study the interaction.

The aim of this study was to characterize the interactions between *Streptomyces* and *Bacillus* found as a co-culture in nitrogen limiting conditions. We have studied the ratio of growth of *Bacillus* to *Streptomyces* in nitrogen-rich and nitrogen-limiting conditions. We tried several approaches to separate *Streptomyces* from *Bacillus*, but the results indicate the very close association between the two species.

3.2 Materials and Methods

3.2.1 Source of cultures

Cultures were originally isolated from Sioux Prairie of South Dakota 231st Street (44.2'8" N 96° 46' 0" W 1720 ft Elevation) onto nitrogen free agar (NFA) by a previous student, Nabhilah Alshibli. Isolates growing in nitrogen free media 'NFA' (Dahal, NandaKafle *et al.*, 2017) with addition of four carbon sources (glucose, mannitol, arabinose and malic acid) and noble agar (BD Difco), and incubation under reduced oxygen using gas-pack (Gaspack EZ Campy, BD) were sub-cultured over R2A (BD Difco™) to obtain pure cultures. Out of 110 isolates displaying *Streptomyces* colony morphology, the 3 fastest growing isolates were chosen and named 321I, 34I and 40I (Fig 16). While they appeared as uniform colonies of *Streptomyces* during subculture, whole genome sequencing revealed the presence of 16S rRNA genes homologous to *Bacillus* (Alshibli, 2017).

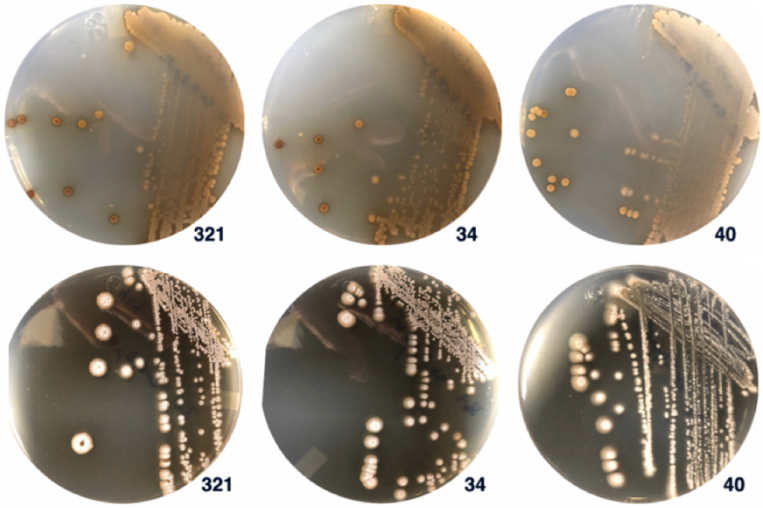


Fig 16: Three *Streptomyces* isolates growing on NFA (top) and R2A (bottom) for 7d.

3.2.2 Growth of *Streptomyces* in NFA

Streptomyces 321I, 34I, 40I and reference species from ARS Culture Collection (NRRL) were grown on nitrogen free agar at 28°C for 7 d.

3.2.3 Quantification of ratio of *Bacillus* to *Streptomyces*

3.2.3.1 Visualization using light microscopy

Isolates were streaked on R2A and NFA and incubated at 28°C for 12 d and viewed microscopically as outlined in Table 2. Cultures was spread and heat-fixed on glass slides and Gram-stained. Cultures were visualized using light microscopy (Olympus BX53 Upright Compound Microscope) at 1000X.

Table 2: Sampling times for microscopy of cultures grown on R2A and NFA.

Culture in R2A media (28°C)	Culture in NFA media (28°C)
Day 1, 3, 5 and 10	Day 3, 5, 7 and 12

3.2.3.2 Study of Growth Pattern of *Streptomyces* and *Bacillus* in co-culture using qPCR

Streptomyces isolates (321I, 34I and 40I) were cultured on R2A and NFA plates at 28°C for different time points as outlined in Table 3. *Streptomyces* isolates were scraped from agar using a sterile loop. To improve the yield of the DNA extraction protocol, the Membrane Binding reagent (100-300 µl) from the gel clean kit (Promega) was used to melt the agar that came with the isolates, and was placed in a heat-block at 55°C for 5 min. The melted agar along with residual reagent was washed using autoclaved nuclease free water 3 times. The cell pellets were resuspended in 300 µl

power bead solution (DNeasy UltraClean Microbial kit, Qiagen) and the cells were ground manually using a tissue grinder pestle in a microcentrifuge tube (1.5 ml) followed by incubation at 55ºC for 5 min. The mixture was vortexed in power bead tubes for 10 min and the remaining steps followed the manufacturers protocol. DNA concentration was determined using a Qubit Fluorometer (Invitrogen) with the Qubit™ dsDNA BR Assay kit (Invitrogen).

Conventional PCR using Actinomycetes-specific primers and Firmicutes specific primers (Table 3) was done at annealing temperatures 57ºC, 58ºC and 59ºC to determine the optimum annealing temperature to be used in real time PCR amplification. The reaction mixture composition was as outlined in section 2.2.4 for the amplification of 16S rRNA.

Table 3: Phylum-specific 16S rRNA gene primers used.

Target	Sequence	Reference
Actinomycetes-specific	S-P-Acti-0927-a-S-17: GGRCCCGCACAAGCGGC S-P-Acti-1154-a-S-19: GADACYGCCGGGGTY AACT	(Pfeiffer, Pastar *et al.*, 2014)
Firmicutes-specific	S-P-Firm-0525-a-A-18: CAGCAGTAGGGAATCTTC S-P-Firm-0525-a-A-18: ACCTACGTATTACCGCGG	

DNA extracted (DNeasy UltraClean Microbial kit, Qiagen) at different time points and from two different media (Table 2) was used to quantify the 16S rRNA genes using Actinomycetes and Firmicutes-specific primers (Table 3) using real-time PCR. SYBR Green Master mix (Bio-Rad) contained premixed dNTPs, DNA polymerase and $MgCl_2$. A reaction volume of 10 µl was used, containing 5 µl SYBR green master mix, 0.25µl (0.5µM) forward and reverse primers, 2.5 µl nuclease free water, and 2µl template DNA. Reaction conditions were set for 45 cycles at 95°C denaturation for 45 s, 59°C annealing for 1 min and 72°C extension for 30 s. The default melting temperature was used. Real-time amplification was performed using a Quant Studio 6 (Applied Biosystems). *Bacillus* 321B and *Arthrobacter aurescens* were used as controls for Firmicutes and Actinomycetes respectively.

3.2.4 Approaches to obtain *Streptomyces* into single culture

3.2.4.1 Size-occlusion by membrane filter

Lawn cultures of *Streptomyces* (321I, 34I and 40I) were prepared on R2A, covered with 0.2 µm and 0.4 µm (GTTP, Isopore™ Membrane Filters) filter membranes and incubated at 28°C for 48 h. To draw young *Streptomyces* hyphae, 20 µl of sterile distilled water was pipetted on top of the filter membrane and pipetted in and out for a few times without allowing the pipettor tip to touch the membrane, and plated on R2A. This was repeated at 43 h, 49 h, 65 h and 73 h of incubation. DNA of cultures obtained was extracted (DNeasy UltraClean Microbial kit, Qiagen) and subjected to PCR reaction using Firmicutes-specific primers (Table 3) to check for the presence of *Bacillus*.

3.2.4.2 Chloramphenicol resistance of *Streptomyces*

Streptomyces isolates (321I, 34I and 40I) and *Bacillus* isolates (321B, 34B and 40B) were grown as lawn cultures on R2A plates and antibiotic discs (Rifampin 5µg, Penicillin 10 IU/IE/UI, Sulfisoxazole 0.25mg, Chloramphenicol 30µg, Streptomycin 10µg, Tetracycline 30µg, Ampicillin 10µg and Vancomycin 30µg) Oxoid were placed to observe the zone of inhibition to determine the resistance of *Streptomyces* and inhibition of pure *Bacillus*. A Chloramphenicol disc was placed at the center of a lawn culture of *Streptomyces* on R2A and incubated for 4 d. The zone of inhibition formed on *Bacillus* plates by Chloramphenicol disc was chosen as standard area around the disc. *Streptomyces* colonies forming within that zone were sub-cultured and incubated at 28ºC for 5 d. Visualization under light microscope was done following gram staining.

3.2.4.3 Collection of *Streptomyces* spores

Week-old *Streptomyces* cultures on R2A plates were taken and overlayed with 3 ml of distilled water. Sterile cotton balls were stuffed only in the mid-region of 15 ml conical tube. The suspension obtained from the culture surface, was aliquoted over the cotton and centrifuged at 5428 X g for 10 min. Cotton balls were removed and 100 µl of the spore suspension was aliquoted onto 1.5 ml tubes. Spore suspension were serially diluted and plated on R2A followed by incubation at 28ºC for 5 d. Visualization of the presence of *Bacillus* was done using light microscopy using gram staining, and separately malachite green to visualize endospores.

3.2.5 Transformation of *Bacillus* with mCherry (RFP)

3.2.5.1 Determination of Autofluorescence spectrum of *Streptomyces* and *Bacillus* and selection of mCherry plasmid

Wet smears of each sample were prepared in 10 µl of water and covered with cover slips sealed using clear nail varnish. Lambda-scanning was done using a confocal microscope (Olympus FV1200 Scanning Confocal Microscope) at 405 nm excitation spectrum for *Bacillus* and *Streptomyces*. Plasmid pTU1_A_pdh_RiboJ_mCherry_Bba_B0015 was purchased from Addgene. Plasmid from *E. coli* DH5α was extracted using a plasmid extraction kit (QIAprep Spin Miniprep Kit, Qiagen). Competent *E. coli* (dam-/dcm-) C2925I (New England Biolabs) were used to generate unmethylated mCherry plasmid using the protocol given by New England Biolabs.

3.2.5.2 Transformation of *Bacillus* with mCherry

Single colonies of *Bacillus* 321B, 34B and 40B and *B. cereus* were inoculated in 50ml LB broth and incubated overnight at 30ºC in a water-bath shaker at 200 rpm. Overnight culture (1ml) was transferred to 50 ml LB broth and incubated for 2 h. At every 15 min interval, absorbance at 600 nm was taken until 6 h. Graphs were plotted to obtain the growth curve.

3.2.5.2.1 Electroporation

Bacillus cultures 321B, 34B, 40B, and *B. cereus* were cultured overnight in LB broth with continuous shaking at 30ºC. Broth culture (0.5 ml) was used as inoculum (OD 0.01 at 600 nm) and incubated in a water bath shaker maintained at 30ºC for 150 min to let the culture reach mid-exponential phase (OD ~0.5A). Electroporation buffer (250mM

sucrose, 1mM HEPES, 1mM $MgCl_2$, 10% Glycerol) was prepared, set to pH 7 and kept at 4°C. Cells were washed 5 times in chilled electroporation buffer and finally concentrated 500-fold. Cells (50µl) and 10µl plasmid were added to 1mm electroporation cuvettes and subjected to electroporation using a Gene Pulser Xcell Electroporation Systems (Bio-rad). An electric field of 15KVcm^{-1}, resistance of 200 and capacitance of 25µF was applied to the cuvettes, immediately followed by addition of 1ml of LB broth. It was incubated at 37°C for 90 min and plated on chloramphenicol LB plates (Brigidi, Derossi *et al.*, 1990). Optimization was pursued by varying the voltage (20KVcm^{-1}), different concentrations of plasmid (100ng-700ng), methylated and unmethylated plasmid and using half-strength LB for incubation.

3.2.5.2.2 Natural competence

Bacillus cultures 321B, 34B, 40B, and *B. subtilis* 168 were grown in competence medium (Boonstra, Vesel *et al.*, 2018) with continuous shaking at 30°C. Cultures were harvested at 3, 4.5, 6 and 7.5 h for transformation. Bacterial cells were incubated with plasmid in ice-bath for 3 minutes. LB was added to the tubes and incubated for 90 min at 30°C and plated onto tetracycline plates.

3.2.5.2.3 Preparation of *Bacillus* protoplast for transformation

Bacillus isolates 321B, 34B and 40B were grown in Penassay broth (1.5g/l Beef extract, 1.5g/l yeast extract, 5g/l peptone, 1g/l dextrose, 3.5g/l sodium chloride, 3.68g/l dipotassium phosphate, 1.32g/l monopotassium phosphate, pH 7.2) overnight. Culture for protoplast preparation was prepared by inoculating 1ml overnight culture into 50ml of Penassay broth at 37°C with rigorous shaking until the OD (550nm) reached 1. Cells were harvested by centrifugation at 5482 X g, resuspended in 5ml of SMMP media

(Puyet, Sandoval *et al.*, 1987) and aliquoted to 100ml flask along with 2mg/ml of lysozyme (Fisher BioReagents) followed by incubation at 37^{o}C for 1 h. Protoplast formation was monitored by observing under microscope using phase contrast setting. Protoplasts were collected by centrifugation at 2000 g for 10 min and washed with SMMP to remove lysozyme. Protoplast (0.2ml) suspension in SMMP was aliquoted to 1.5 ml tubes and 0.4ml of 40% PEG 6000 was added. Plasmid of varying concentration from 100ng to 400ng was added to the tubes. The tubes were incubated on ice for 3 min, when 1.2 ml of SMMP was added and protoplasts were pelleted by centrifugation at 2000 X g for 10 min. Pellets were re-suspended in 400 ml SMMP, and incubated at 37^{o}C for 90 min, followed by plating onto tetracycline regeneration media DM3 and R2YE (5µg/ml tetracycline). Transformation of *Bacillus* was confirmed by fluorescence microscopy and plasmid extraction.

3.2.5.3 Co-culture of Transformed *Bacillus* and *Streptomyces*

Transformed *Bacillus* (34B) was inoculated into 50ml one-day old *Streptomyces* culture (321I, 34I and 40I) with and without tetracyline.

3.3 Results

3.3.1 Growth of *Streptomyces* in NFA

Streptomyces 321I, 34I and 40I showed good growth in NFA when incubated at 28^{o}C for 7 d. *S. ederensis* (B-8146) and *S. pheochromogenes* (B-1248) showed weak growth in NFA (Table 4).

Table 4: Growth of *Streptomyces* species (NRRL) in NFA.

NRRL No.	**Strain Description**	**Growth in NFA (7 d) at 28°C**
B-8146	*Streptomyces ederensis*	Weak growth*
ISP-5182	*Streptomyces phaeoluteigriseus*	No growth
B-1248	*Streptomyces pheochromogenes subspecies pheochromogenes*	Weak growth*
B-1239	*Streptomyces aurantiacus*	No growth
B-16637	*Streptomyces coelicolor subspecies coelicolor*	No growth
B-3804	*Streptomyces termitum*	No growth
ISP-5361	*Streptomyces anulatus*	No growth

* Colonies growing are smaller than *Streptomyces* 321I, 34I and 40I.

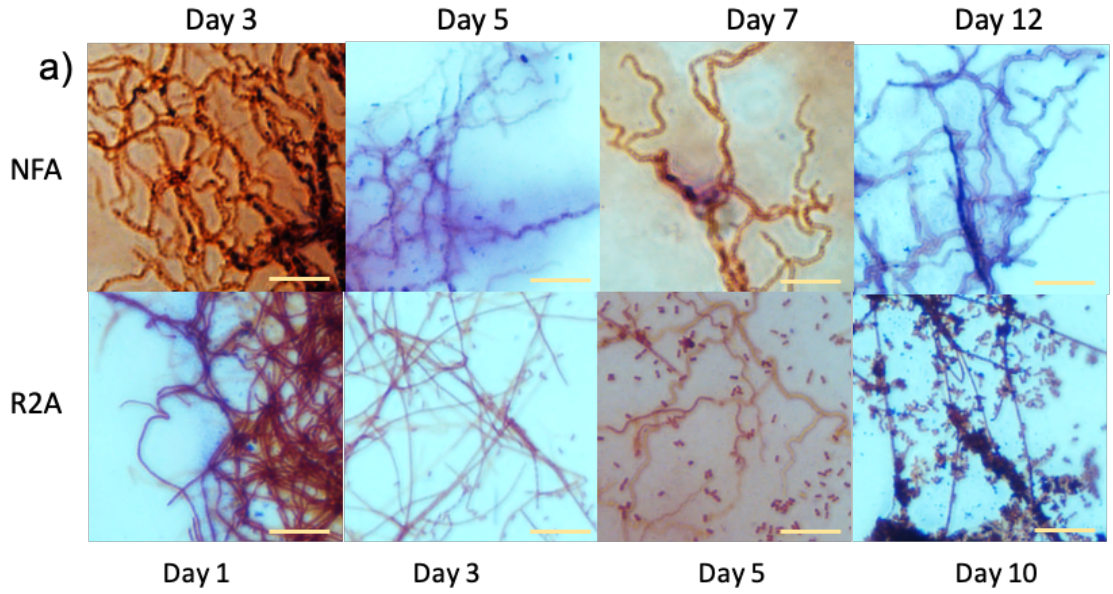
3.3.2 Prevalence of *Bacillus* associated with *Streptomyces* during growth

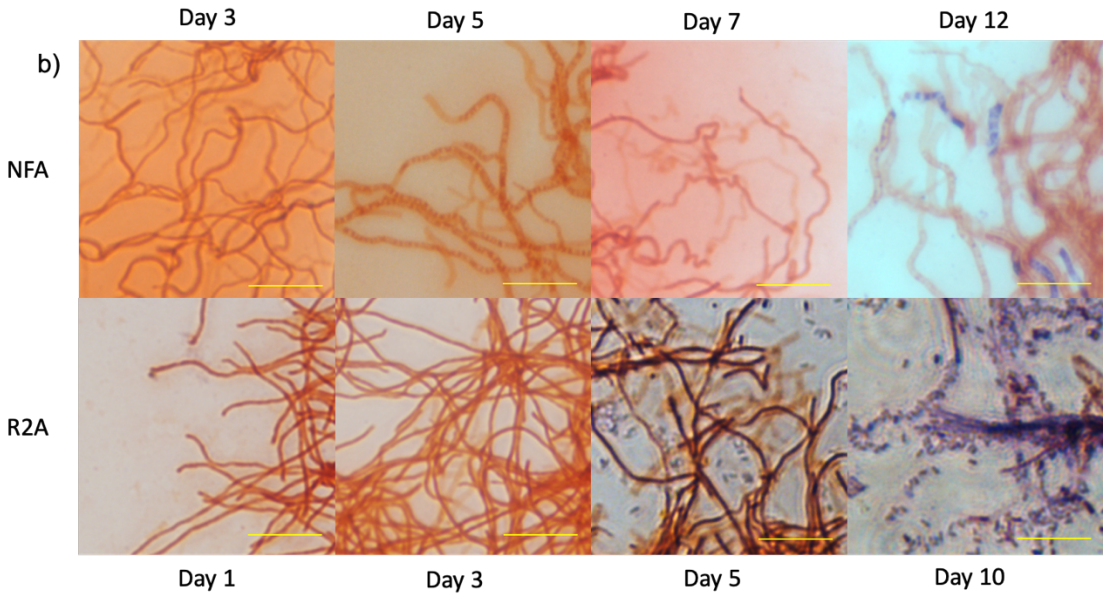
Genetic evidence for *Bacillus* associated with apparently pure *Streptomyces* cultures, combined with isolation of *Bacillus* prompted an investigation of distribution overtime. *Streptomyces* were cultured on NFA and R2A and analyzed at various time points (Table 3).

3.3.2.1 Microscopic investigation

Each isolate (321I, 34I and 40I) was taken from NFA and R2A at different times (Table 3), gram stained, and images captured using an inverted light microscope (1000X). On NFA media, there were no visible rod like structures seen up to d 12 (Fig 17), with the exception of isolate 34I, where a few rods were seen in one trial on all days, but never again in multiple repeated attempts. *Streptomyces* hyphae structures appeared thicker on NFA than R2A and were alternatively densely stained at certain parts of the

filaments. In contrast, free rods were visible from d 5 when cultured on nitrogen-replete R2A. These results suggested that *Bacillus* were absent or present in very low numbers during the earlier period of growth. Alternatively, *Bacillus* could be localized inside *Streptomyces* filaments, with some released upon ageing of the cultures.





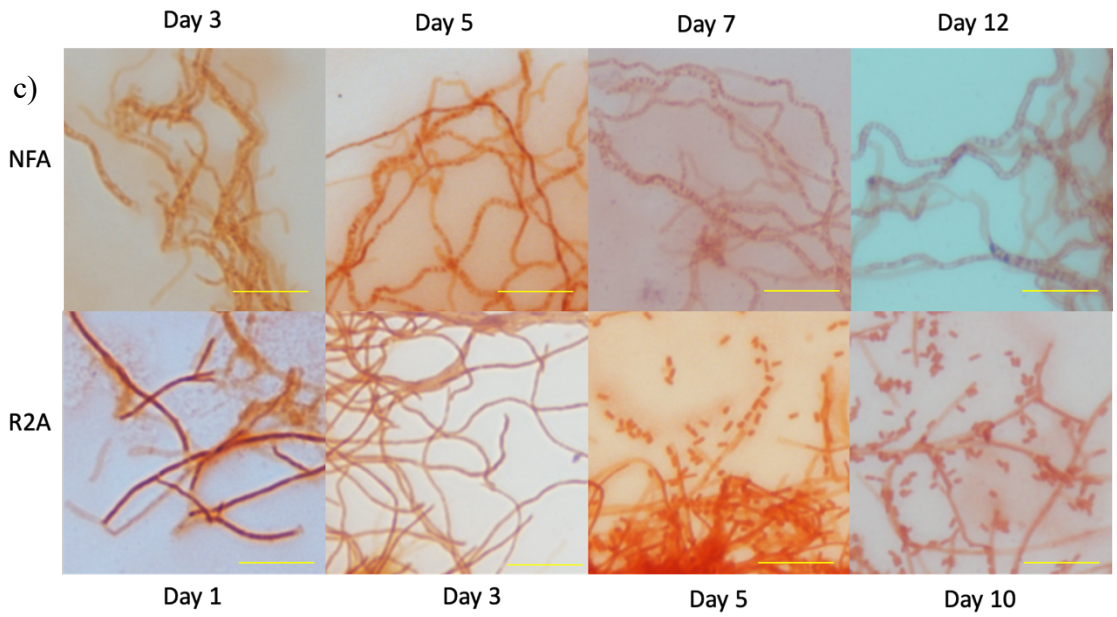


Fig 17: Microscopy of NFA and R2A-grown isolate 321I (a), 34I (b) and 40I (c) over time using bright field microscopy at 1000X. Heat-fixed samples were Gram stained. Yellow bar represents size bar of 10µm.

3.3.2.2 Analysis of ratio of *Bacillus* to *Streptomyces* by qPCR

To determine possible changes in the ratio of *Bacillus* to *Streptomyces* over time, 16S rRNA gene copies were quantified by qPCR. DNA was quantified using Qubit and the concentration was adjusted to between 10-20 ng/µl. The optimum annealing temperatures for Actinomycetes-specific and Firmicutes-specific primers were obtained by evaluating three different temperatures (57ºC, 58ºC and 59ºC). The PCR products were resolved on a gel, showing that all temperatures were suitable for both the Firmicutes-specific primers, whereas there was no amplification of *Firmicutes* (*Bacillus* 34B) using Actinomycetes-specific primers at 59ºC (Fig 18). Thus, 59ºC was chosen as annealing temperature to be used in the qPCR amplification reactions. The Actinomycetes-specific primers amplified only the actinomycete *Arthobacter aurescens*.

We could not use *Streptomyces* isolates as control as these all contained *Bacillus*. These results showed that annealing at 59°C would not result in cross-amplification of *Bacillus* by Actinomycete primers, or *Streptomyces* by Firmicute primers.

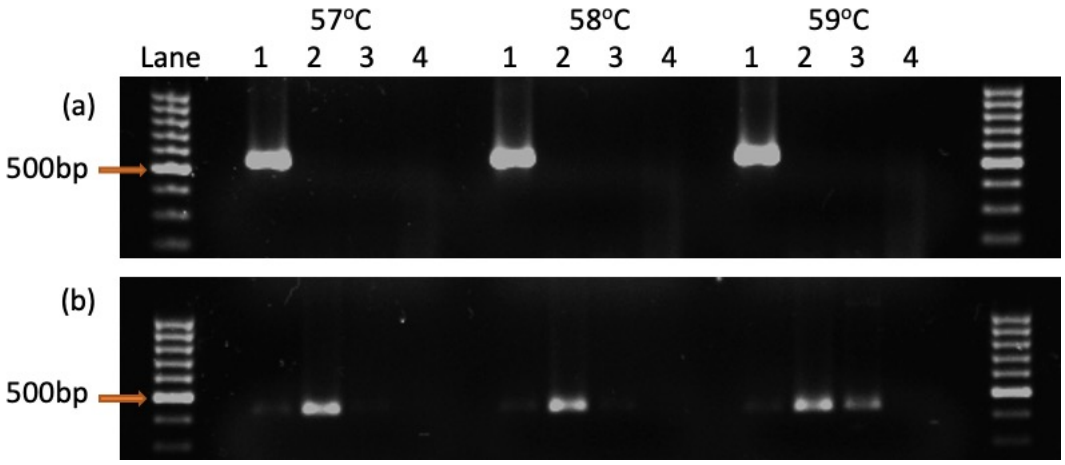


Fig 18: Optimization of PCR annealing temperatures for Firmicutes-specific primers (a) and Actinomycetes-specific primers (b). Lane 1: 34B, Lane 2: *Arthrobacter aurescens*, Lane 3: *E. coli*, Lane 4: Water

The amplification by qPCR of 16S rRNA using Firmicutes-specific and Actinomycetes-specific primers was done using DNA templates extracted at four different time points from NFA and R2A (Table 1). Three technical replicates and three biological replicates were done for each isolate and each time point. The ratio of Firmicutes to Actinomycetes is given in Appendix 1. Actinomycetes- and Firmicutes-specific primers were used to amplify 16S rRNA by qPCR of *Bacillus* 321B, 34B, 40B, and *A. aurescens*. There was difference in the cycle threshold when using two different primers, that indicates the primers are specific and there is relatively little non-specific reaction. The ratio in copy numbers of 16S rRNA genes of Firmicutes (*Bacillus*) to Actinomycetes (*Streptomyces*) was close to 1 at all time points in both media (Fig 19).

We were not able to determine the genomic ribosomal operon copy number from genome sequences conclusively. Also, *Streptomyces* filaments contain multiple chromosomes per cell so that copy number cannot be converted to number of cells. These results suggest that the *Streptomyces* and *Bacillus* grew in a fixed proportion overtime. As *Bacillus* were not visible as distinct cells earlier during growth on agar, these 16S rRNA gene copy number results suggest *Bacillus* occurs either inside or tightly associated with *Streptomyces* filaments.

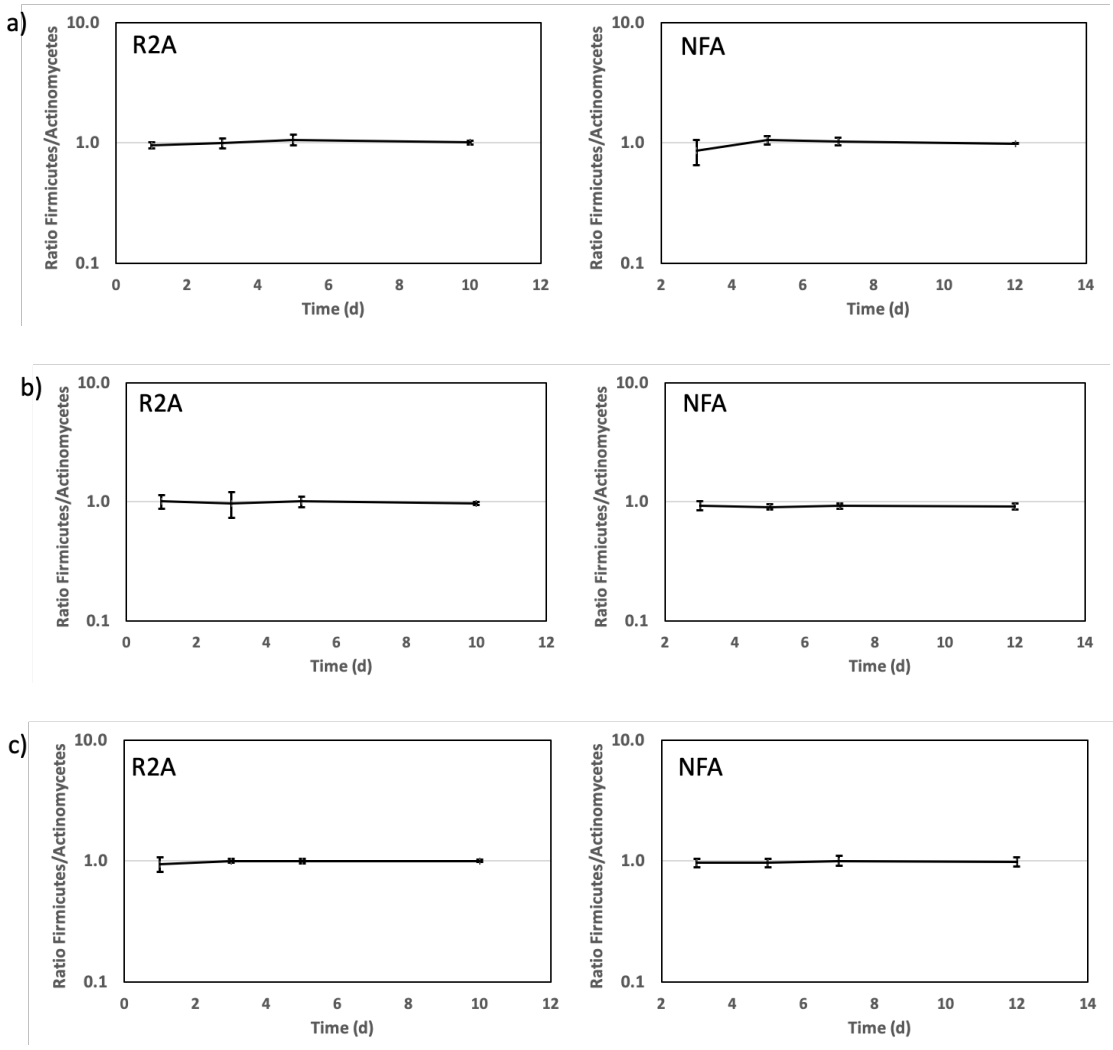


Fig 19: The ratio of Firmicutes to Actinomyces 16S rRNA gene threshold values for isolates 321I (a), 34I (b) and 40I (c) cultured on NFA and R2A. Error bars represent one standard deviation of the mean obtained from three technical replicates of each of three biological replicates.

3.3.3 Separation of *Streptomyces* from *Bacillus*

Bacillus could be isolated into separate culture, however several approaches to separate *Streptomyces* from *Bacillus* were not successful. Each method chosen to separate *Streptomyces* from *Bacillus* suggests peculiar characteristics of this relationship.

3.3.3.1 Separation by size occlusion

Previous reports indicated that *Streptomyces* can be isolated from other bacteria due to their thin filament diameter. By overlaying membrane filters with pore size smaller than average bacterial diameters, *Streptomyces* filaments grew through pores, leaving other species behind (Hirsch & Christensen, 1983). Filter membranes of pore size 0.2 and 0.4 μm were used to let *Streptomyces* hyphae emerge while growing. The tips of putatively emerged *Streptomyces* hyphae were suspended gently using water droplets and plated on to fresh R2A plates. Samples were taken at various time points to extract the emerged *Streptomyces*. DNA extracted from the sub-cultured *Streptomyces* was subjected to amplification by Firmicutes- and Actinomycetes-specific primers by PCR. Both primer sets yielded amplicons (Fig 20) indicating the continued presence of *Bacillus* in the *Streptomyces* obtained by the filter membrane technique. Pure cultures of the three *Bacillus* could not be isolated from the tops of membrane filters onto agar. It is possible that *Streptomyces* growth caused the membrane material to crack, allowing *Bacillus* to move through. No visible membrane damage was seen. These results suggest that the *Bacillus* passed through membranes inside of the *Streptomyces*.

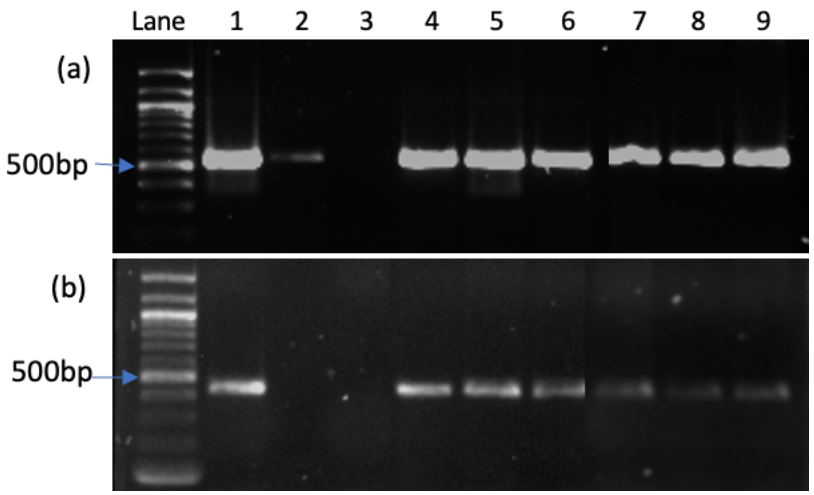


Fig 20: DNA extracted from filter separated *Streptomyces* amplified by Firmicutes-specific (a) and Actinomycetes-specific primers (b). Lane 1: *Bacillus* 34B for (a) and *A. aurescens* for (b), 2: *E. coli*, 3: water control, 4: 321S(0.2µm), 5: 321S(0.4µm), 6: 34S(0.2µm), 7: 34S(0.4µm), 8: 40S(0.2µm), 9: 40S(0.4µm)

3.3.3.2 Utilizing the Chloramphenicol resistance trait of *Streptomyces* isolates

The sensitivity of *Bacillus* isolates and *Streptomyces* cultures to various antibiotics was tested. Chloramphenicol was able to inhibit *Bacillus* while *Streptomyces* was able to grow. *Streptomyces* colonies growing in the region of zone of inhibition for *Bacillus* (Fig 21) were sub-cultured onto R2A. The *Streptomyces* colonies were gram stained after 5 d of growth and inspected microscopically. The presence of *Bacillus* was confirmed due to the visibility of rods. This indicated that the *Bacillus* strains were protected from Chloramphenicol in the presence of the *Streptomyces*, suggesting a sheltering role of the *Streptomyces*.

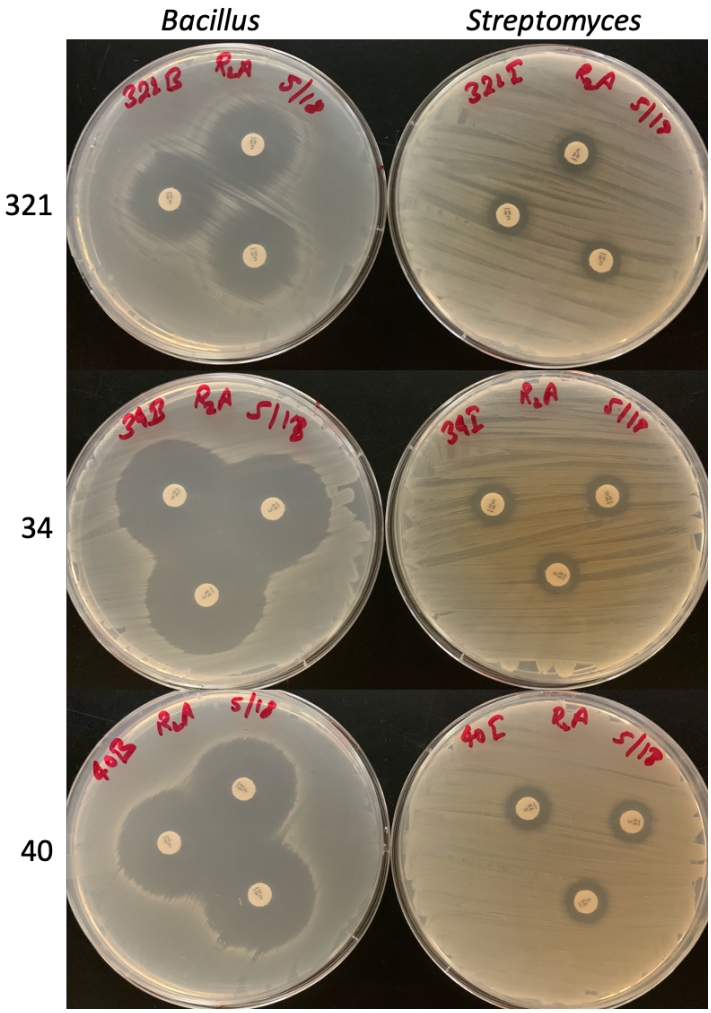


Fig 21: Zones of inhibition formed by Chloramphenicol discs on lawn culture of *Bacillus* isolates 321B, 34B and 40B (left) and *Streptomyces* cultures 321I, 34I and 40I (right).

3.3.3.3 Isolation of spores of *Streptomyces*

While *Streptomyces* hyphae may be long and contain multiple chromosomes, their spores are haploid due to a tightly controlled partitioning system (Ditkowski, Holmes *et al.*, 2013; Jakimowicz & van Wezel, 2012). We tested whether spores of our *Streptomyces* would occlude potential endophytic *Bacillus*. Spores of *Streptomyces* were visualized using Malachite green and Safranin as counterstain. Microscopy of isolates 321I and 34I revealed the presence of small circular spores that were stained lightly with

safranin, suggesting the spores of *Streptomyces* do not retain malachite green as do endospores of *Bacillus*. Isolate 40I had a few blue patches and lots of lightly stained *Streptomyces* spores. There was complete absence of vegetative cells as observed under the microscope.

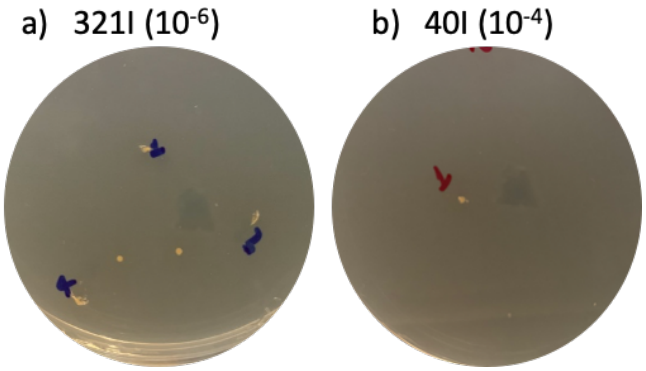


Fig 22: *Streptomyces* colonies on R2A after serial dilution of spores of 321I (a) and 40I (b).

Serial dilutions of spores were plated on R2A and up to 10 colonies were observed per plate (Fig 22) at dilution 10^{-6} for 321I and 10^{-4} for 40I. Spore collection, dilution, plating and subculture was performed 2 times for each isolate. Microscopy of gram stained *Streptomyces* colonies of each isolate on R2A was done at 7d. Presence of rods were observed in each case, indicating the presence of *Bacillus* and therefore its passage via *Streptomyces* spores (Fig 23).

3.3.4 Transformation of *Bacillus* isolates with mCherry

To facilitate visualization of *Bacillus* in *Streptomyces* cultures, we sought to tag these by a fluorescent reporter. *Streptomyces* and *Bacillus* isolates were used for lambda scanning using confocal microscopy, to determine wavelengths outside of

autofluorescence. Autofluorescence of *Streptomyces* was observed at around 509nm which falls in the region of the GFP emission spectrum. Thus, mCherry with excitation spectrum of green wavelength (540-590 nm) and emission spectrum of red wavelength (550-650 nm) was selected for transformation of *Bacillus* isolates.

The transformation protocol required the collection of exponential phase of bacterial cells, so the growth curves of *Bacillus* isolates 321B, 34B, 40B and *B. cereus* were determined (Fig 23).

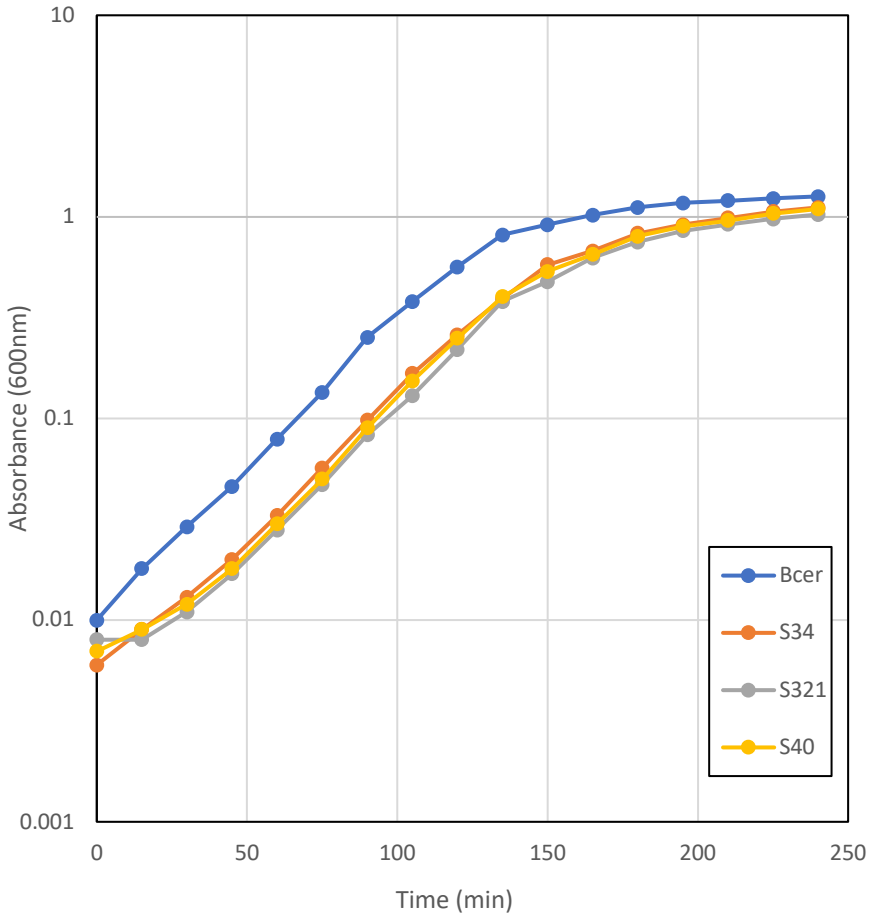


Fig 23: Growth curves of *Bacillus* and *B. cereus* ATCC 1457a isolates in R2A broth at 37ºC with continuous shaking. Absorbance was measured at 600nm.

Several methods employed to transform *Bacillus* isolates 321B and 40B were not successful. Using electroporation, *Bacillus cereus* was transformed while other *Bacillus* isolates were not transformed using various parameters such as methylated and unmethylated (dcm^-/dam^-) plasmid, different plasmid concentration (100-700 ng) and varying electric field (15-20 $KVcm^{-1}$).

Bacillus isolates 321B and 40B belonged to the same branch as *B. subtilis* in the phylogenetic tree (Fig 15). Natural competence generally present in the *B. subtilis* group was employed to transform the *B. subtilis* 168, and *Bacillus* isolates 321B, 34B and 40B. Only *B. subtilis* 168 was transformed with mCherry using competence media as confirmed by plasmid extraction, while fluorescence microscopy showed the fluorescence in the spores only.

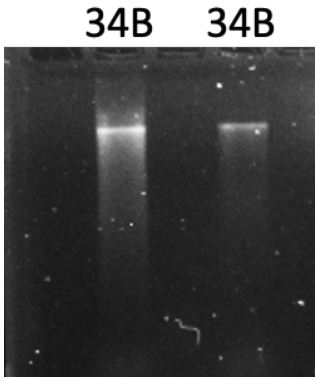


Fig 24: Gel image showing plasmid bands obtained from plasmid extraction from transformed *Bacillus* isolate 34B.

Protoplasts of *Bacillus* 321B, 34B, and 40B were prepared using incubation with lysozyme for one hour. Protoplast formation was confirmed by visualization under phase contrast microscopy (Appendix 2) and all cells appeared circular with no rods visible. 34B protoplasts was transformed with mCherry. Transformed protoplasts regenerated by

sub-culturing in half-strength LB and then on normal LB plates. Transformation of isolate 34B was confirmed by plasmid extraction (Fig 24) but could not be confirmed by fluorescence microscopy. Confirmation of transformed isolate 34B to be original isolate was done by gram staining which gave gram positive rods and Firmicutes-specific PCR amplification gave positive band on gel.

3.3.5 Co-culture of transformed *Bacillus* 34B with *Streptomyces* isolates

Several co-cultures of transformed *Bacillus* 34B with *Streptomyces* isolates 321I, 34I and, 40I were performed with and without tetracycline in the R2A broth. Three out of 4 times, there was growth of *Streptomyces* only in absence of tetracycline and there were no visible clumps of *Streptomyces* formed in presence of tetracycline. Importantly, no free *Bacillus* was seen in the broth culture although it grew when inoculated without *Streptomyces*. This indicated either that growth of free *Bacillus* was inhibited by *Streptomyces* activity, or that *Bacillus* were adsorbed to *Streptomyces* cell surface. Supernatant of *Streptomyces* cultures spotted onto a lawn of *Bacillus* did not show a zone of inhibition.

3.4 Discussion

We determined the prevalence of *Bacillus* associated with *Streptomyces* during growth using microscopy and quantitative amplification of 16S rRNA by qPCR. There is no known dye to differentially bind to each species, hence, to visualize the bacteria we used Gram-stain, although both species are Gram-positive. This allowed us to visualize *Streptomyces* as hyphal and *Bacillus* as rods structures. The presence of free rods on R2A

at later stage, but absence of free rods on NFA suggests that *Bacillus* could be localized inside *Streptomyces* filaments, with some released upon ageing of the cultures. Quantification of 16S rRNA of both *Bacillus* and *Streptomyces* by qPCR indicated a constant ratio of *Bacillus* to *Streptomyces*. *Streptomyces* filaments contain multiple chromosome per cell so that copy number cannot be converted to number of cells. We were not able to determine the genomic ribosomal operon copy number from genome sequences conclusively. Primer specificity is crucial in quantitative analysis. Both Firmicutes- and Actinomycetes-specific primers were specific enough to rule-out false positive. Non-specific binding by primers can be minimized by using optimum concentration of template. Melt curve analysis of quantification of DNA obtained from NFA, having lower concentration, always yielded a single peak, whereas from R2A with higher DNA concentration, there were two peaks in several trials. However, qPCR products when resolved on a gel gave a single band (Appendix 2). The genetic evidence indicates presence of *Bacillus* genomes in *Streptomyces* cultures throughout while distinct rods or their endospores could not be seen in young cultures. This suggests the *Bacillus* are associated tightly with the *Streptomyces*.

Several approaches to separate *Streptomyces* from *Bacillus* were not successful. By overlapping membrane filters with pore size smaller than average bacterial diameters, *Streptomyces* filaments grew through pores, leaving other species behind (Hirsch & Christensen, 1983). Pure culture of the three *Bacillus* could not be isolated from the tops of membrane filters, indicating that free *Bacillus* do not pass across the membrane. It is possible that *Streptomyces* growth through the membrane caused the material to crack, allowing *Bacillus* to move through. No visible membrane damage was seen. These results

suggest that the *Bacillus* passed through the membranes inside of the *Streptomyces*. Chloramphenicol sensitive *Bacillus* was detected even after subjecting *Streptomyces* to Chloramphenicol, suggesting that *Bacillus* is protected from the antibiotic due to very close association with Chloramphenicol resistant *Streptomyces*. Spore germination of *Streptomyces* contained *Bacillus* after incubation for 5 d, suggesting spores of *Streptomyces* carries *Bacillus* spores along, however, Malachite staining of *Streptomyces* spores did not show any sign of endospores.

Laboratory strains can be transformed more readily, while wild strains isolated from the environment often need laborious methods such as phage transduction, electroporation or protoplast transformation (Nijland, Burgess *et al.*, 2010). Protoplast transformation of *Bacillus* 34B was achieved, which was confirmed by plasmid extraction. However, detectable fluorescence was not expressed by vegetative cells of *Bacillus*.

Co-culture of *Streptomyces* with transformed *Bacillus* 34B in broth did not show visible free growth of transformed *Bacillus*. This result suggests that *Streptomyces* growth inhibited growth of *Bacillus*, which could be due to production of an antimicrobial metabolite, but could not be confirmed. Alternatively, the *Bacillus* were adsorbed or entrapped by the *Streptomyces*.

These results point to *Bacillus* as an endophyte of *Streptomyces* that is released by older cultures, allowing isolation. As tagging of *Bacillus* was unsuccessful, presence either inside or associated with *Streptomyces* will need to be investigated by fluorescent *in-situ* hybridization using taxon-specific probes. As *Bacillus* could not be removed from *Streptomyces*, even through spore formation, it is tempting to speculate that the

Streptomyces relies on *Bacillus* for critical functions. Genome sequencing of the two species should serve to indicate possible dependence of *Streptomyces* through loss of certain core genes.

4. Chapter 3: Whole Genome Sequence Analysis of *Bacillus* and *Streptomyces*

4.1 Introduction

The interaction between two bacteria in a coculture is often symbiotic in nature. In a previous study (Chapter 1), *Bacillus* was isolated from *Streptomyces* as coculture, and grew well in nitrogen-free medium. There is no species of *Streptomyces* known to fix atmospheric nitrogen. The study characterized the *Bacillus* that grew well on nitrogen free medium, with assays showing the presence of nitrogenase activity. The occurrence of *Bacillus* in the colony of *Streptomyces* was found to be very closely associated, such that the results pointed to *Bacillus* as an endophyte of *Streptomyces* (Chapter 2). To speculate the degree and type of interaction occurring between these two species, genetic information can be very helpful.

A symbiotic partnership/association implies sharing or exchange of metabolites, amino acids, siderophores, as well as signaling molecules (Pacheco & Segre, 2019). Genome sequencing of *T. princeps* and *M. endobia* endosymbiosis inside the mealybug *Planococcus citri* has revealed that there is extensive gene loss over time due to the interaction. *M. endobia* is responsible for complete synthesis of ATP, nucleotides and cellular envelope whose synthesis is completely absent in *T. princeps* (Lopez-Madrigal, Latorre *et al.*, 2013). This endosymbiosis involves dependence to the point where neither species can grow on its own, or together outside of its eukaryotic host. Less interdependent symbioses may involve less gene loss.

Whole genome sequencing of the two species that occur together like *Bacillus* and *Streptomyces* (Chapter 1 and 2), cannot be done separately. Illumina DNA sequencing,

the currently dominant technology, generates large numbers of short reads generally used to produce accurate genome assemblies. The major differences between technologies that generate long reads and short reads is sequencing error. Illumina reads are short (500 bp or less) and less error prone but produce fragmented genome assemblies. Long reads up to tens of thousands can be generated using Pacific Biosciences (PacBio) and Oxford Nanopore Technologies (ONT). The DNA from the coculture can be sequenced, with both the short read and long reads combined using modules such as Unicycler (Wick, Judd *et al.*, 2017). However, the major challenge is in filtering the genes that belong to each species. Bioinformatics tools have been developed that make it easier to map the whole genome sequence to existing metabolic pathways. Blast helps to determine the homology of the gene sequences using the information stored in its database. Searching for conserved regions in each gene and determining homology can be used to filter the genes belonging to each strain. Although computationally difficult, phylogeny of each gene may help in filtering out the genes. The use of the whole *Bacillus* genome to filter out *Streptomyces* genes can be done.

KEGG offers a knowledge base for linking genomes to biological functions, however it requires its own annotated genes (KO ids) to perform pathway analysis (Minoru Kanehisa, Sato *et al.*, 2016). Genotypic approaches to investigate the taxonomy of species have been possible, and the 16S rRNA gene sequence is widely used to construct phylogenetic trees in the bacterial and archaeal kingdom. Sometimes, the 16S rRNA genes are too conserved, not varying enough to determine the relationship within closely related species. For example, the taxonomic relationship of reference

Streptomyces at species level was better resolved using Multilocus Sequence Typing (MLST) than 16S rRNA method (Guo, Zheng *et al.*, 2008).

In this study we sequenced the whole genomes of *Bacillus* and *Streptomyces* that were isolated from nitrogen-limited condition. We have analyzed the predicted metabolic pathways and found prevalence of various pathways that may be part of the interaction. We have determined the phylogeny of *Streptomyces* 321 and 40 that shows the closest relatives using 5 gene MLST to be *S. pheachromogenes* and *S. umbrinus*.

4.2 Materials and Methods

4.2.1 Culture preparation

Out of 3 isolates, 321I and 40I were selected. Isolates were incubated in 250 ml flasks with 50 ml R2A broth for 3 d at 30^{o}C. The broth cultures were harvested by centrifugation at 5428 X g for 10 min, yielding a pellet of about 200µl. Pure culture *Bacillus* was cultured in the same way but overnight. The pellets were sent for whole genome sequencing (Genewiz Next Generation Sequencing) using Illumina sequencing technology.

4.2.2 Assembly and annotation

Paired raw reads (fastq format) and *de-novo* assembled genomes were obtained in fasta file format (Genewiz Next Generation Sequencing). Bioinformatics analysis was done (Fig 25), which gave files with filtered *Streptomyces* contigs for each sample, annotation tables of the *de-novo* assembled genome contigs from Diamond BLASTp (Buchfink, Xie *et al.*, 2015), and open reading frames (nucleotide and protein sequences) (Madeira, Park *et al.*, 2019). Paired raw reads were again used to construct *de-novo*

assembly using modules (FastQC, BBduk, and SPAdes) available in SDSU Linux cluster, Roaring thunder. Quality check and trimming was done using FastQC ("FastQC," 2015) and BBduk (http://jgi.doe.gov/data-and-tools/bb-tools/). *De-novo* genome assembly was done using SPAdes (version 3.10) to generate contigs (Bankevich, Nurk *et al.*, 2012). Prokka annotation tool (Seemann, 2014) and RAST (Aziz, Bartels *et al.*, 2008) was used to annotate the genes.

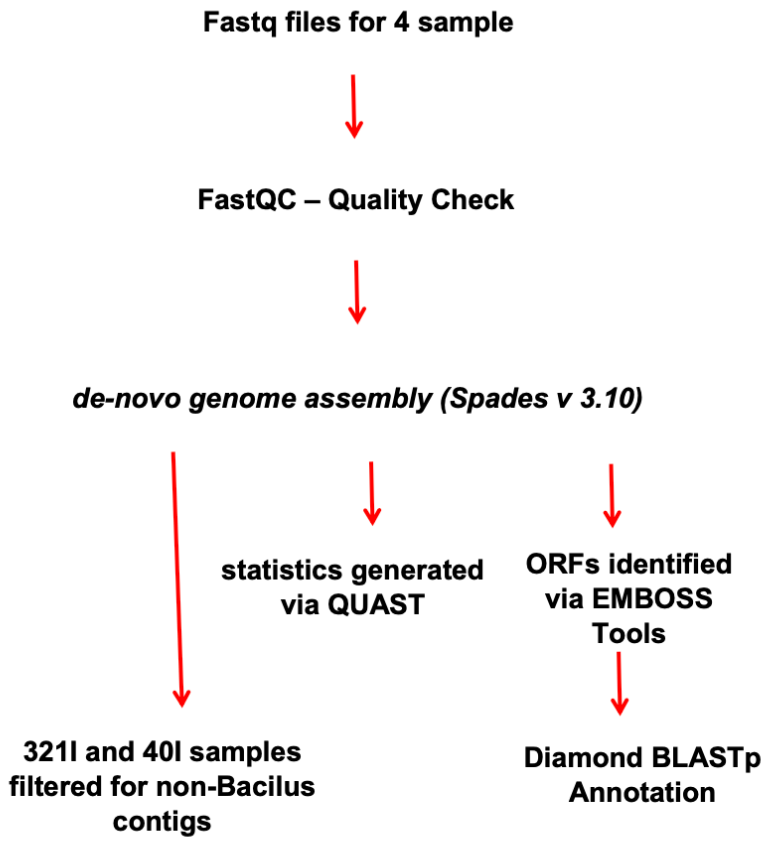


Fig 25: Bioinformatics analysis workflow for de-novo genome assembly by Genewiz.

4.2.3 Phylogeny of *Streptomyces*

Nucleotide sequences of five genes, *recA*, *atpD*, *gyrB*, *rpoB* and 16S rRNA for reference species were extracted from their genomes on GenBank as described by Guo

(2008). The reference species were selected, approximately one from each cluster, from the phylogenetic tree based on six-gene concatenated sequences reported by Guo (2008). Other reference species were the closest that occupied the same cluster of a 16S rRNA phylogenetic tree. Alignments and construction of phylogenetic trees were constructed using MEGAX software (Kumar, Stecher *et al.*, 2018). Muscle was used to align multiple sequences that contained four genes concatenated end to end. A Maximum likelihood tree was generated using the general time reversible model. Bootstrapping (1000X) was used to interpret the probabilities that respective groups of species are present in the true phylogeny.

4.2.4 KEGG annotation and reconstruction of pathways

Open reading frames with protein sequences delivered by Genewiz for *Bacillus* and *Streptomyces* were used. *Bacillus*-filtered by Genewiz *Streptomyces* contigs were used to determine open reading frames using the EMBOSS Sixpack translation tool (Madeira, Park *et al.*, 2019). KO ids were generated using BlastKOALA (Minoru Kanehisa, Sato *et al.*, 2016). To reconstruct pathways using KO ids, KEGG mapper was used (M. Kanehisa & Sato, 2020).

4.2.5 Analysis of genes using Kaiju

Kaiju (Menzel, Ng *et al.*, 2016) web serve was used to upload paired reads of *Bacillus* 321B. Completely assembled RefSeq genomes of Bacteria, Archaea and Viruses were used as reference database. The program parameters were selection to filter low complexity protein query sequences, Greedy run mode that allowed mismatches with minimum match score 75 and 5 allowed mismatches, and minimum match length was 11.

For the reference, the *B. subtilis* 168 whole genome was taken from the NCBI genome database.

4.3 Results

Short paired ends read were generated using Illumina sequencing technology. The sequencing quality overviews are given in table 5.

Table 5: Sequencing quality overview.

Sample ID	**Barcode Sequence**	**Number of Reads**	**Yield (Mbases)**	**Mean Quality Score**	**%Bases >=30**
321B	GCGTAGTA+ACTCTAGG	99,179,297	29,754	37.06	87.6
40B	GCGTAGTA+CTTAATAG	76,399,128	22,920	37.16	87.98

4.3.1 Genome assembly and annotation

Raw reads were trimmed using the BBduk module with around 10% bases removal, and 5966 reads removal for isolate 321B, and 4898 reads removal for 40B. Trimmed reads had good quality scores across all bases with value higher than 22 for both forward and reverse reads of both isolates. The mean GC distribution was observed to be at 45% for *Bacillus* isolates (Appendix 3a). SPAdes, version 3 was used to assemble the reads whose output was used to annotate the genes using Prokka and RAST. Total number of contigs and total number of genes are given in Table 6. We found two 16S rRNA from 321B and 40B with more than 1000 bp.

Table 6: Genome assembly showing the contigs and numbers of genes.

Assembly	**321B**	**321I**	**321I_non-*Bacillus***	**40B**	**40I**	**40I_non-*Bacillus***
# contigs	22	613	222	22	1,036	158
# contigs (>= 1000 bp)	22	613	222	22	1,036	158
# contigs (>= 5000 bp)	16	194	171	16	149	130
# contigs (>= 10000 bp)	14	165	143	14	136	126
# contigs (>= 25000 bp)	10	117	105	10	104	94
# contigs (>= 50000 bp)	9	66	64	9	75	66
Largest contig	1,059,411	332,669	332,669	1,059,411	366,079	366,079
Total length	4,022,784	11,980,001	10,641,397	4,023,147	12,367,140	9,991,625
Total length (>= 1000 bp)	4,022,784	11,980,001	10,641,397	4,023,147	12,367,140	9,991,625
Total length (>= 5000 bp)	4,011,927	11,372,479	10,519,903	4,010,826	11,049,096	9,927,840
Total length (>= 10000 bp)	3,996,615	11,156,807	10,309,634	3,995,514	10,952,594	9,891,450
Total length (>= 25000 bp)	3,942,625	10,386,920	9,710,956	3,941,440	10,455,482	9,394,338
Total length (>= 50000 bp)	3,906,934	8,528,705	8,145,916	3,905,749	9,373,875	8,355,067

N50	962,120	104,283	121,146	961,052	109,468	128,152
N75	429,485	44,290	54,187	429,368	50,309	62,852
L50	2	33	28	2	32	23
L75	5	76	61	5	74	51
GC (%)	43.74	69.18	70.2	43.74	67.66	70.26
N's	280	1017	585	281	681	586
# N's per 100 kbp	6.96	8.49	5.5	6.98	5.51	5.86
Genes	4278	11457	10041	4276	12159	9221

4.3.2 Phylogeny of *Streptomyces*

A phylogenetic tree using only the 16S rRNA gene sequences did not resolve various reference sequences well, with up to five clustered together. The four protein coding genes *recA*, *atpD*, *gyrB*, *rpoB* and the 16S rRNA gene of all reference species were taken from the GenBank database. The sequences were concatenated for phylogenetic analysis using MEGAX. 321S falls in the same cluster as *S. phaeochromogenes*, while 40S is more closely related to *S. umbrinus.*

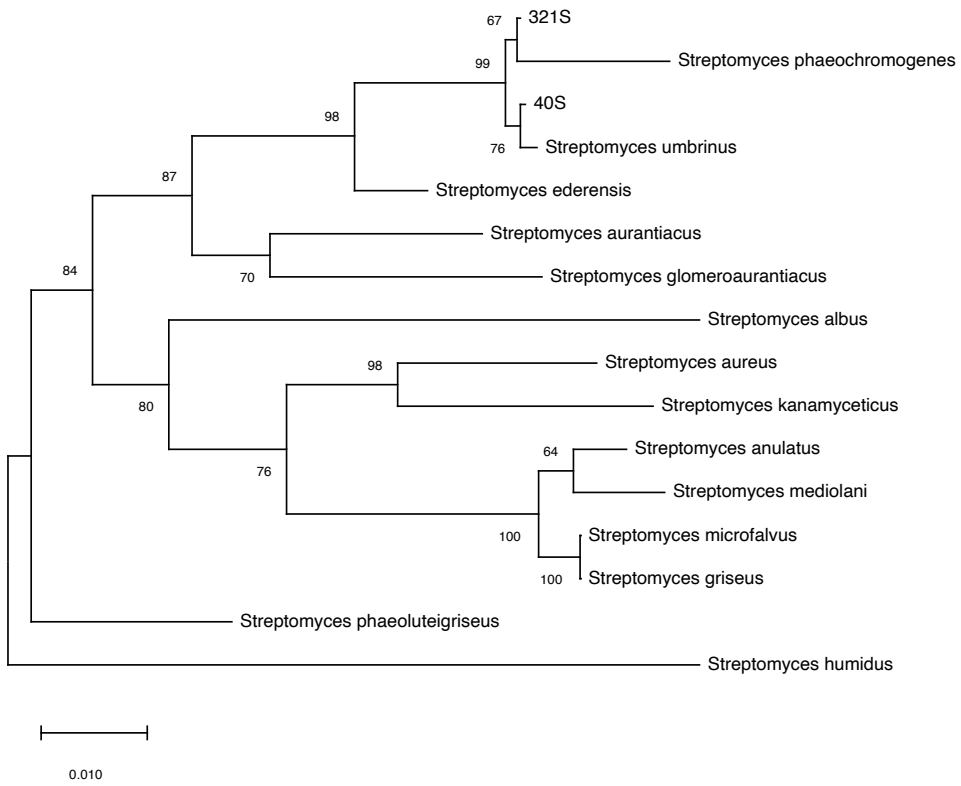


Fig 27: Phylogenetic relationship using five-gene (*recA*, *atpD*, *gyrB*, *rpoB* and 16S rRNA) concatenated sequences. The tree was constructed with maximum likelihood method using *S. humidus* as outgroup. Numbers in each node represent percentage of repeatability in 1000 bootstrap runs.

4.3.3 Functional annotation

Open reading frames generated using EMBOSS tools provided by Genewiz were used for generation of KO ids of *Bacillus* 40B and *Streptomyces* 40I. For *Streptomyces* filtered contigs, open reading frames were generated using EMBOSS Sixpack translation tool. KO ids were used to reconstruct KEGG pathways that yielded 208 pathways for *Bacillus* 40B, 190 pathways for *Streptomyces* filtered, and 179 pathways for *Streptomyces* (mixed culture). Pathways of 40B and 40I are listed in Table 7.

Bacillus 40B has a greater number of genes present in most of the degradation pathways than *Streptomyces* 40I, except Benzoate, aromatic compounds, glycan, chlorocyclohexane and chlorobenzene degradation pathways. For antibiotic synthesis pathways, carbapenem biosynthesis is entirely present in *Bacillus* only and Vancomycin group antibiotics synthesis is present only in *Streptomyces*. *Bacillus* has a greater number of Vancomycin and ß-lactam resistance genes in its genome. This suggests *Bacillus* has acquired resistance genes to combat the antibiotics that are produced by *Streptomyces* in their close association.

Table 7: Analysis of reconstruct KEGG pathways using KO ids of *Bacillus* 40B, *Streptomyces* 40 and *Streptomyces* 40 filtered*.

	Pathway	***Bacillus* 40B**	***Streptomyces* & *Bacillus***	***Streptomyces* filtered**
	Degradation pathways:			
1	Valine, leucine and isoleucine degradation	15	7	8
2	Lysine degradation	6	7	7
3	RNA degradation	10	3	4
4	Fatty acid degradation	8	9	10
5	Benzoate degradation	5	13	16
6	Caprolactam degradation	-	3	3
7	Chloroalkane and chloroalkene degradation	4	5	6
8	Aminobenzoate degradation	4	2	3

9	Degradation of aromatic compounds	4	14	21
10	Synthesis and degradation of ketone bodies	3	1	1
11	Geraniol degradation	2	1	1
12	Naphthalene degradation	1	5	4
13	Limonene and piene degradation	1	2	2
14	Ethylbenzene degradation	1	-	-
15	Xylene degradation	1	1	4
16	Styrene degradation	1	3	3
17	Atrazine degradation	1	2	5
18	Glycosaminoglycan degradation	1	2	3
19	Other glycan degradation	1	7	6
20	Chlorocyclohexane and chlorobenzene degradation	2	5	6
21	Nitrotoluene degradation	-	1	2
22	Toluene degradation	-	1	2
23	Fluorobenzoate degradation	-	1	2
24	Dioxin degradation	-	-	4
25	Polycyclic aromatic hydrocarbon degradation	-	-	1
	Antibiotic synthesis			
1	Streptomycin biosynthesis	4	8	9

2	Monobactam biosynthesis	4	2	4
3	Novobiocin biosynthesis	3	2	2
4	Penicillin and cephalosporin biosynthesis	2	2	2
5	Carbapenem biosynthesis	2	-	-
6	Neomycin, kanamycin and gentamicin biosynthesis	1	1	1
7	Ansamycins biosynthesis	1	-	1
8	Vancomycin group antibiotics biosynthesis	-	4	4
9	Biosynthesis of enediyne antibiotics	-	-	1
	Antibiotics Resistance:			
1	ß-lactam resistance	9	2	3
2	Vancomycin resistance	5	1	2
	Pigments:			
1	Prodigiosin biosynthesis	3	2	2
2	Phenazine biosynthesis	1	4	4
	Nitrate/Nitrite reduction:			
1	Nitrate reductase (nirB, narG, narH)	3	-	-
2	Assimilatory nitrate reductase catalytic unit (nasA)	1	-	-

3	Carbonic anhydrase (cynT)	1	1	1
4	MFS transporter (NRT)	1	-	-
5	Nitronate monooxygenase (ncd2)	-	1	1
	Protein export:			
1	secA	1	-	-
2	secY	1	-	-
3	SRP54	1	-	-
4	ftsY	1	-	-
5	tatC	1	1	1
6	yidC	1	1	1
7	secDF	1	1	1
8	tatA	-	1	1
9	secG	-	1	1
	Interaction (plant pathogen):			
1	Glycerol kinase	1	1	1
2	Elongation factor Tu (tuf)	1	-	-
3	Flagellin	1	-	-
4	Molecular chaperon HtpG	1	-	-
	Exosome:	25	23	23

	Competence protein (ComQ):	1	-	-
	Porphyrin and **chlorophyll metabolism:**	13	12	15
	Quorum Sensing	26	21	24

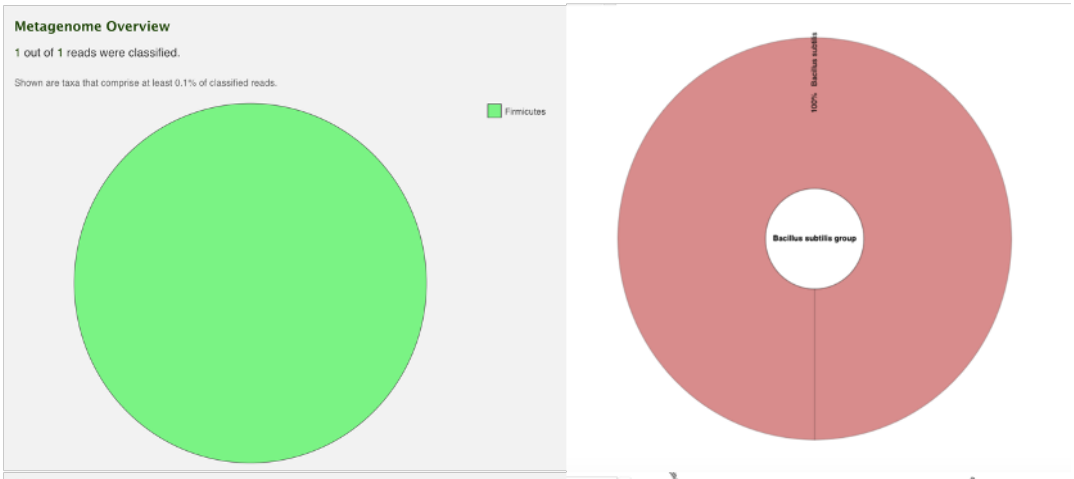
*Streptomyces filtered KO ids were generated using EMBOSS Sixpack translation tool.

4.3.4 Analysis of genes using Kaiju

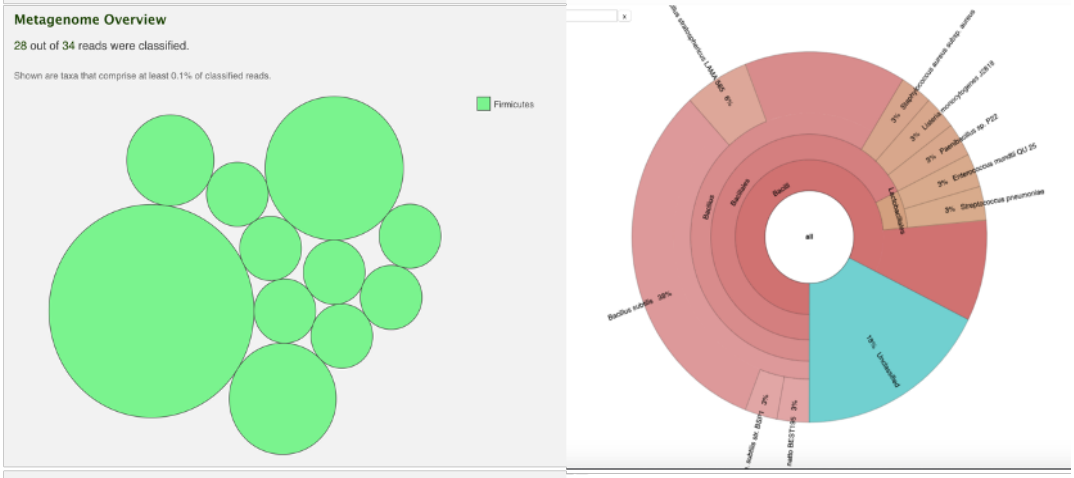
Whole genome sequences of the reference genomes *B. subtilis* 168 and *B. subtilis* UBBS-14, and *Bacillus* 321B, and 40B from this study were used to classify sequences to taxa using Kaiju. Kaiju classifies the genome sequences using protein database and gave a result with the number of taxa that are classified in a bubble plot and Krona chart (Fig 27). Reference genomes were classified as 100% Firmicutes showing all the proteins are true representative of that phylum. The sequences of *Bacillus* 321B and 40B from this study were 85% Firmicutes, and remaining genes were allocated to taxa classified as Actinobacteria, Proteobacteria, Terrabacteria and other bacteria. This indicates that the whole genome sequences of *Bacillus* 321B and 40B have either acquired genes of other classes of bacteria through horizontal gene transfer or carry those bacteria in a minority form along with them.

Metagenome Overview | Krona Chart

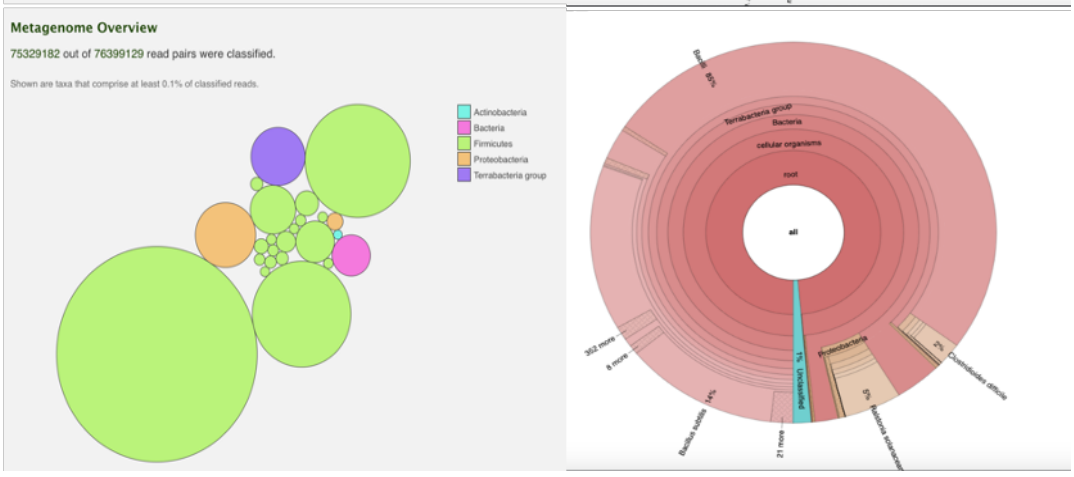
a)



b)



c)



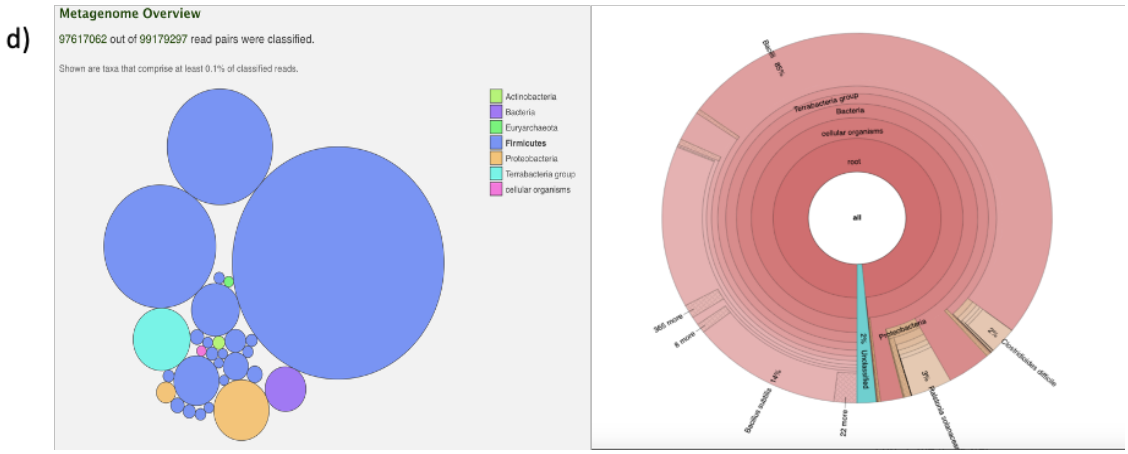


Fig 28: Taxonomy classification of reference genomes *B. subtilis* 168 (a) and *B. subtilis* UBBS-14 (b), and *Bacillus* 321B (c) and 40B (d) paired raw reads using Kaiju. Metagenome overview gives classification of read pairs, each bubble indicating a different class of bacteria, and Krona charts give in-browser interactive visualization of each taxon abundance. *Bacillus* 321B and 40B was predicted to contain genes associated with *Clostridioles difficile*, *Ralstonia solanaceae* and Terrabacteria group.

4.3.5 Bio-synthetic gene clusters of *Streptomyces*

The online software of anti-SMASH was used to predict bio-synthetic gene clusters of *Streptomyces* 40I and 321I and is given in Table 8 and 9.

Table 8: Bio-synthetic gene clusters of *Streptomyces* 40I determined by anti-SMASH software version 5.1.2.

Region	**Type**	**From**	**To**	**Most similar known cluster**		**Similarity**
Region 1.1	NRPS	80,370	131,272	coelichelin	NRP	100%

Region 1.2	NRPS	235,280	292,520	streptovaricin	Polyketide	19%
Region 2.1	NRPS-like,phenazine	1	34,181	pyocyanine	Other	71%
Region 2.2	lassopeptide,T1PKS	312,923	346,501	chalcomycin A	Polyketide	39%
Region 4.1	terpene,T1PKS	98,085	160,903	4-hexadecanoyl-3-hydroxy-2-(hydroxymethyl)-2H-furan-5-one	Polyketide	45%
Region 6.1	T2PKS	56,321	128,869	alnumycin A / alnumycin B / alnumycin C / alnumycin P / prealnumycin / thalnumycin A / thalnumycin B / K1115A / 1,6-dihydro-8-	Polyketide:Type II	100%

				propylanthraquinone		
Region 7.1	terpene	1	17,526	ebelactone	Polyketide	5%
Region 7.2	phosphoglycolipid	23,592	52,308	teichomycin	Other	77%
Region 8.1	T1PKS	200,175	232,370	herboxidiene	Polyketide	6%
Region 9.1	NRPS,nucleoside	167,191	220,494	rimosamide	NRP	21%
Region 12.1	terpene,NRPS,T1PKS	72,797	143,394	vazabitide A	NRP	32%
Region 16.1	T3PKS	109,877	150,941	herboxidiene	Polyketide	8%
Region 19.1	melanin	53,430	63,912	melanin	Other	60%
Region 20.1	T3PKS	97,442	138,527	alkylresorcinol	Polyketide	100%
Region 22.1	ectoine	96,349	106,753	ectoine	Other	100%
Region 23.1	T1PKS,terpene	5,092	49,231	cyphomycin	Polyketide	2%

Region 23.2	terpene	88,564	109,478	herboxidiene	Polyketide	4%
Region 24.1	NRPS,T1PKS	1	61,252	melithiazol A	NRP + Polyketide:Modular type I	26%
Region 35.1	NRPS-like	64,681	99,846	stenothricin	NRP:Cyclic depsipeptide	13%
Region 44.1	bacteriocin,lanthipeptide	49,260	77,652	informatipeptin	RiPP:Lanthipeptide	100%
Region 56.1	melanin	1	5,271	melanin	Other	28%
Region 61.1	NRPS	1	41,846	SCO-2138	RiPP	78%
Region 63.1	siderophore	12,046	23,842	desferrioxamin B / desferrioxamine E	Other	83%
Region 67.1	terpene	18,971	45,638	hopene	Terpene	92%
Region 80.1	NRPS-like,T1PKS	18,937	46,689	nogalamycin	Polyketide	40%
Region 82.1	terpene	32,067	44,538	geosmin	Terpene	100%

Region 8 9.1	T1PKS	3	38,793	vicenistatin	Polyketide	80%
Region 1 10.1	T1PKS	1	19,933	pladienolide B	Polyketide	37%
Region 1 15.1	T1PKS	1	16,239	ambruticin	Polyketide	11%
Region 1 22.1	T1PKS	1	14,393	tylactone	Polyketide	27%
Region 1 36.1	T1PKS	1	10,127	maklamicin	Polyketide	13%
Region 1 91.1	lanthipeptide	1	2,636	staphylococcin C55 α / staphylococcin C55 β	RiPP:Lanthipeptide	20%

Table 9: Bio-synthetic gene clusters of *Streptomyces* 321I determined by anti-SMASH software version 5.1.2.

Region	Type	From	To	Most similar known cluster		Similarity
Region 1.1	NRPS-like,phenazine	1	33,849	endophenazine A / endophenazine B	Other:Phenazine	27%

Region 1.2	lassopeptide,T1PKS	299,231	332,669	chalcomycin A	Polyketide	39%
Region 2.1	T1PKS	296,206	325,323	herboxidiene	Polyketide	5%
Region 3.1	terpene	1	16,973	ebelactone	Polyketide	5%
Region 3.2	phosphoglycolipid	23,046	51,762	teichomycin	Other	77%
Region 5.1	T2PKS	107,499	180,047	alnumycin A / alnumycin B / alnumycin C / alnumycin P / prealnumycin / thalnumycin A / thalnumycin B / K1115A / 1,6-dihydro-8-propylanthraquinone	Polyketide:Type II	100%
Region 6.1	NRPS	80,269	131,188	coelichelin	NRP	100%
Region 6.2	NRPS	236,163	274,356	acyldepsipeptide 1	NRP	15%

Region 7.2	T1PKS,terpene	103,808	166,608	4-hexadecanoyl-3-hydroxy-2-(hydroxymethyl)-2H-furan-5-one	Polyketide	45%
Region 8.1	terpene	42,629	60,467	herboxidiene	Polyketide	4%
Region 8.2	T1PKS,terpene	101,145	145,284	cyphomycin	Polyketide	2%
Region 8.3	bacteriocin,lanthipeptide	199,663	228,066	informatipeptin	RiPP:Lanthipeptide	100%
Region 10.1	NRPS,nucleoside	162,451	215,754	rimosamide	NRP	21%
Region 15.1	T3PKS	27,615	68,700	alkylresorcinol	Polyketide	100%
Region 16.1	ectoine	96,100	106,504	ectoine	Other	100%
Region 18.1	terpene,NRPS,T1PKS	51,451	122,245	vazabitide A	NRP	32%
Region 22.1	melanin	45,096	55,476	melanin	Other	28%
Region 26.1	NRPS-like	1	35,432	stenothricin	NRP:Cyclic depsipeptide	13%

Region 27.1	T1PKS,NRPS	70,023	122,283	thienodolin	Other	14%
Region 30.1	melanin	52,530	63,012	melanin	Other	60%
Region 31.1	lanthipeptide	76,801	103,132	incednine	Polyketide	4%
Region 45.1	T3PKS	27,835	68,848	herboxidiene	Polyketide	8%
Region 54.1	betalactone	1	34,384	JBIR-76 / JBIR-77	Polyketide	8%
Region 55.1	siderophore	21,066	32,862	desferrioxamin B / desferrioxamine E	Other	83%
Region 57.1	NRPS	20,797	65,696	SCO-2138	RiPP	78%
Region 58.1	terpene	9,099	35,766	hopene	Terpene	92%
Region 72.1	NRPS-like,T1PKS	1	27,761	nogalamycin	Polyketide	40%
Region 74.1	terpene	30,239	46,065	geosmin	Terpene	100%

Region 76.1	NRPS,T1PKS	1	44,290	thiazostatin / watasemycin A / watasemycin B / 2-hydroxyphenylthiazoline enantiopyochelin / isopyochelin	NRP	20%
Region 87.1	T1PKS	3	39,180	vicenistatin	Polyketide	80%
Region 134.1	T1PKS	1	17,547	mediomiycin A	Polyketide	36%
Region 153.1	T1PKS	1	12,462	tylactone	Polyketide	27%
Region 164.1	T1PKS	1	10,625	angolamycin	Polyketide	100%
Region 165.1	T1PKS	1	10,590	nanchangmycin	Polyketide	21%
Region 221.1	other	1	2,187	bacilysin	Other	28%

4.4 Discussion

The 16S rRNA gene phylogeny did not resolve some closely related species as the gene is very conserved in this genus. Taxonomic relationship of reference *Streptomyces* at species level was more helpful using Multilocus Sequence Typing (MLST) than 16S rRNA method (Guo et al., 2008). Concatenation of *recA*, *atpD*, *gyrB*, *rpoB* and 16S rRNA resolved the closely related species. *Streptomyces* 321S was found to be closely related to *S. phaeochromogenes* and 40S was related to *S. umbrinus*, both closely related to each other. Using MLST thus yielded a better picture of the phylogenetic relationship of *Streptomyces* isolates with reference species.

Illumina DNA sequencing generates large number of short reads, which are used to produce accurate genome assemblies. However, due to large number of shorter reads, it is seldom possible to assemble the whole genome in one contig. A limitation of *de novo* assembly is that it produces shorter assemblies than the reference genome (Alkan, Sajjadian *et al.*, 2011). The exact number of 16S rRNA copy number could not be determined from these genomes. Only two 16S rRNA copies with length more than 1000 bp were found for *Bacillus* 321B and 40B, and one 16S rRNA copy was found for *Streptomyces* 321I and 40I. DNA sequencing producing long reads such as Oxford nanopore should be used to assemble the genome more accurately.

Bioinformatics tools are developed that make it easier to map the whole genome sequence to existing metabolic pathways. KEGG offers a knowledge base for linking genomes to biological functions, however it requires its own annotated genes (KO ids) to perform pathway analysis (Minoru Kanehisa et al., 2016). KEGG reconstruct pathway result gives the number of genes present of that specific pathway. Presence of genes for

certain pathways does not necessarily mean the expression of those genes. Degradation of aromatic compound pathway was found to be dominant in *Streptomyces* when compared to *Bacillus*. Metabolic pathways such as RNA degradation, amino acids degradation were found to be more prominent in *Bacillus*. *Streptomyces* has a Vancomycin group antibiotics biosynthesis pathway, whereas a Vancomycin resistance pathway was present in *Bacillus*. Analysis of these pathways in *Bacillus* and *Streptomyces* suggest the adaptation of these when present in close association.

Classification of reads using Kaiju indicated that *Bacillus* 321B has genes that were either transferred through horizontal gene transfer or carries other taxa along as Hitchhikers (Lozano, Bravo *et al.*, 2017). No 16S rRNA genes of taxa other than *Bacillus* were detected in the genome, sequences indicating horizontal gene transfer. Possible reasons for horizontal gene transfer of taxon Actinobacteria could be due to close association of *Bacillus* and *Streptomyces* with other soil bacteria over time. Resequencing of 40I and 321I by long-read technology should help to more clearly resolve the genomes of the *Streptomyces* and *Bacillus* partners.

5. Conclusions and Questions for future study

5.1 Conclusion

In this study, we have determined the close association of *Bacillus* with *Streptomyces* isolated from nitrogen-free conditions. Isolations of free *Bacillus* from the cocultures was done by applying high temperature to isolate high-heat resistant endospores. *Bacillus* was able to grow on the nitrogen-limited media NFA, and characterization of its nitrogen fixing ability was determined using the acetylene reduction assay, $^{15}N_2$ incorporation assay and amplification of *nifH* by PCR.

We determined the phylogeny of *Bacillus* and *Streptomyces* using MLST. Four gene phylogenetic tree of *Bacillus* 321B and 40B showed that they are very closely related to each other and fall in the same cluster as *B. subtilis* 168. A five gene phylogenetic tree of Streptomyces 321 showed its closeness to *S. phaeochromogenes*, and 40 was closer to *S. umbrinus*.

The Bacillus appear tightly associated with the *Streptomyces* partners. Repeated subculturing by classic streaking never yielded single-culture *Bacillus*. Young cultures appeared to contain only *Streptomyces* as no *Bacillus* (rod-shaped Gram positive cells) could be found by microscopy, while older cultures on rich medium did harbor free-standing *Bacillus*. Yet *Bacillus* were present in young *Streptomyces* cultures in a 1:1 ratio as evidence by qPCR of 16S rRNA gene pools. Several approaches to separate *Streptomyces* from *Bacillus* were not successful, and every approach indicated that the interaction is a close association. There is a possibility of endosymbiosis of *Bacillus* inside *Streptomyces* filaments.

Whole genome sequences of *Bacillus* and *Streptomyces* did not show presence of *nifH* genes, yet they grow on NFA. Metabolic pathway analysis of *Bacillus* and *Streptomyces* indicates that the interaction might have led to adaptation to each other over time.

5.2 Questions for future study

The studies done in this thesis have given a platform to further investigate more on the interaction of *Bacillus* with *Streptomyces*. Major areas that need to be investigated in future include;

1. What is the physical location of the interaction? Does *Bacillus* occur inside *Streptomyces* filaments as an endophyte or just tightly associated?
2. Are there any other genes involved in nitrogen fixation other than the Nif, Vnf, Anf family of nitrogenase enzymes?
3. Does the *Streptomyces* lack essential genes whose functions are provided by the *Bacillus* partner?

REFERENCES

Alkan, C., Sajjadian, S., & Eichler, E. E. (2011). Limitations of next-generation genome sequence assembly. *Nature methods, 8*(1), 61-65. doi:10.1038/nmeth.1527

Alshibli, N. (2018). Diversity of free-living nitrogen-fixing bacteria in soil of Sioux Prairie of South Dakota.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *J Mol Biol, 215*(3), 403-410. doi:10.1016/S0022-2836(05)80360-2

Arragain, S., Jimenez-Vicente, E., Scandurra, A. A., Buren, S., Rubio, L. M., & Echavarri-Erasun, C. (2017). Diversity and Functional Analysis of the FeMo-Cofactor Maturase NifB. *Front Plant Sci, 8*, 1947. doi:10.3389/fpls.2017.01947

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., . . . Zagnitko, O. (2008). The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics, 9*(1), 75. doi:10.1186/1471-2164-9-75

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., . . . Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology : a journal of computational molecular cell biology, 19*(5), 455-477. doi:10.1089/cmb.2012.0021

Bobek, J., Šmídová, K., & Čihák, M. (2017). A Waking Review: Old and Novel Insights into the Spore Germination in Streptomyces. *Frontiers in Microbiology, 8*(2205). doi:10.3389/fmicb.2017.02205

Bonfante, P., & Desirò, A. (2017). Who lives in a fungus? The diversity, origins and functions of fungal endobacteria living in Mucoromycota. *The ISME Journal, 11*(8), 1727-1735. doi:10.1038/ismej.2017.21

Boonstra, M., Vesel, N., & Kuipers, O. P. (2018). Fluorescently Labeled DNA Interacts with Competence and Recombination Proteins and Is Integrated and Expressed Following Natural Transformation of Bacillus subtilis. *Mbio, 9*(5). doi:10.1128/mBio.01161-18

Bouvier, T., & Del Giorgio, P. A. (2003). Factors influencing the detection of bacterial cells using fluorescence in situ hybridization (FISH): A quantitative review of published reports. *FEMS Microbiol Ecol, 44*(1), 3-15. doi:10.1016/s0168-6496(02)00461-0

Boyd, E., & Peters, J. (2013). New insights into the evolutionary history of biological nitrogen fixation. *Frontiers in Microbiology, 4*(201). doi:10.3389/fmicb.2013.00201

Braga, R. M., Dourado, M. N., & Araujo, W. L. (2016). Microbial interactions: ecology in a molecular perspective. *Braz J Microbiol, 47 Suppl 1*, 86-98. doi:10.1016/j.bjm.2016.10.005

Brigidi, P., Derossi, E., Bertarini, M. L., Riccardi, G., & Matteuzzi, D. (1990). Genetic-Transformation of Intact-Cells of Bacillus-Subtilis by Electroporation. *Fems Microbiology Letters, 67*(1-2), 135-138. doi:DOI 10.1111/j.1574-6968.1990.tb13850.x

Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature methods, 12*(1), 59-60. doi:10.1038/nmeth.3176

Bull, A. T., Stach, J. E. M., Ward, A. C., & Goodfellow, M. (2005). Marine actinobacteria: perspectives, challenges, future directions. *Antonie Van Leeuwenhoek, 87*(1), 65-79. doi:10.1007/s10482-004-6562-8

Cao, Y., Fanning, S., Proos, S., Jordan, K., & Srikumar, S. (2017). A Review on the Applications of Next Generation Sequencing Technologies as Applied to Food-Related Microbiome Studies. *Frontiers in Microbiology, 8*(1829). doi:10.3389/fmicb.2017.01829

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., & Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. *Science, 263*(5148), 802-805. doi:10.1126/science.8303295

Chalk, P. M. (2016). The strategic role of 15N in quantifying the contribution of endophytic N2 fixation to the N nutrition of non-legumes. *Symbiosis, 69*(2), 63-80. doi:10.1007/s13199-016-0397-8

Challis, G. L., & Hopwood, D. A. (2003). Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by Streptomyces species. *Proceedings of the National Academy of Sciences, 100*(suppl 2), 14555-14561. doi:10.1073/pnas.1934677100

Cleveland, C. C., Townsend, A. R., Schimel, D. S., Fisher, H., Howarth, R. W., Hedin, L. O., . . . Wasson, M. F. (1999). Global patterns of terrestrial biological nitrogen (N2) fixation in natural ecosystems. *Global Biogeochemical Cycles, 13*(2), 623-645. doi:10.1029/1999gb900014

Coombs, J. T., & Franco, C. M. M. (2003). Isolation and Identification of Actinobacteria from Surface-Sterilized Wheat Roots. *Applied and Environmental Microbiology, 69*(9), 5603-5608. doi:10.1128/aem.69.9.5603-5608.2003

Dahal, B., NandaKafle, G., Perkins, L., & Brozel, V. S. (2017). Diversity of free-Living nitrogen fixing Streptomyces in soils of the badlands of South Dakota. *Microbiol Res, 195*, 31-39. doi:10.1016/j.micres.2016.11.004

DeLong, E. F., Wickham, G. S., & Pace, N. R. (1989). Phylogenetic stains: ribosomal RNA-based probes for the identification of single cells. *Science, 243*(4896), 1360-1363. doi:10.1126/science.2466341

Dilworth, M. J. (1966). Acetylene reduction by nitrogen-fixing preparations from Clostridium pasteurianum. *Biochimica et Biophysica Acta (BBA) - General Subjects, 127*(2), 285-294. doi:https://doi.org/10.1016/0304-4165(66)90383-7

Ditkowski, B., Holmes, N., Rydzak, J., Donczew, M., Bezulska, M., Ginda, K., . . . Jakimowicz, D. (2013). Dynamic interplay of ParA with the polarity protein, Scy, coordinates the growth with chromosome segregation in Streptomyces coelicolor. *Open Biol, 3*(3), 130006. doi:10.1098/rsob.130006

Dixon, R., & Kahn, D. (2004). Genetic regulation of biological nitrogen fixation. *Nat Rev Microbiol, 2*(8), 621-631. doi:10.1038/nrmicro954

Dos Santos, P. C., Fang, Z., Mason, S. W., Setubal, J. C., & Dixon, R. (2012). Distribution of nitrogen fixation and nitrogenase-like sequences amongst microbial genomes. *BMC Genomics, 13*. doi:Artn 162

10.1186/1471-2164-13-162

Dunlap, C. A. (2019). Taxonomy of registered Bacillus spp. strains used as plant pathogen antagonists. *Biological Control, 134*, 82-86. doi:10.1016/j.biocontrol.2019.04.011

FastQC. (2015). In.

Fry, B. (2007). *Stable Isotope Ecology*: Springer New York.

Gaby, J. C., & Buckley, D. H. (2012). A Comprehensive Evaluation of PCR Primers to Amplify the nifH Gene of Nitrogenase. *Plos One, 7*(7), e42149. doi:10.1371/journal.pone.0042149

Galloway, J. N. (1998). The global nitrogen cycle: changes and consequences. *Environmental Pollution, 102*, 15-24. doi:Doi 10.1016/S0269-7491(98)80010-9

Gao, C., Xue, Y., & Ma, Y. (2011). Protoplast transformation of recalcitrant alkaliphilic Bacillus sp. with methylated plasmid DNA and a developed hard agar regeneration medium. *Plos One, 6*(11), e28148. doi:10.1371/journal.pone.0028148

Giani, A. M., Gallo, G. R., Gianfranceschi, L., & Formenti, G. (2020). Long walk to genomics: History and current approaches to genome sequencing and assembly. *Computational and Structural Biotechnology Journal, 18*, 9-19. doi:https://doi.org/10.1016/j.csbj.2019.11.002

Goh, E. B., Yim, G., Tsui, W., McClure, J., Surette, M. G., & Davies, J. (2002). Transcriptional modulation of bacterial gene expression by subinhibitory concentrations of antibiotics. *Proc Natl Acad Sci U S A, 99*(26), 17025-17030. doi:10.1073/pnas.252607699

Golinska, P., Wypij, M., Agarkar, G., Rathod, D., Dahm, H., & Rai, M. (2015). Endophytic actinobacteria of medicinal plants: diversity and bioactivity. *Antonie Van Leeuwenhoek, 108*(2), 267-289. doi:10.1007/s10482-015-0502-7

Gouda, S., Das, G., Sen, S. K., Shin, H. S., & Patra, J. K. (2016). Endophytes: A Treasure House of Bioactive Compounds of Medicinal Importance. *Frontiers in Microbiology, 7*, 1538. doi:10.3389/fmicb.2016.01538

Guo, Y., Zheng, W., Rong, X., & Huang, Y. (2008). A multilocus phylogeny of the Streptomyces griseus 16S rRNA gene clade: use of multilocus sequence analysis for streptomycete systematics. *Int J Syst Evol Microbiol, 58*(Pt 1), 149-159. doi:10.1099/ijs.0.65224-0

Handley, K. M., VerBerkmoes, N. C., Steefel, C. I., Williams, K. H., Sharon, I., Miller, C. S., . . . Banfield, J. F. (2013). Biostimulation induces syntrophic interactions that impact C, S and N cycling in a sediment microbial community. *ISME J, 7*(4), 800-816. doi:10.1038/ismej.2012.148

Heitzer, A., Malachowsky, K., Thonnard, J. E., Bienkowski, P. R., White, D. C., & Sayler, G. S. (1994). Optical biosensor for environmental on-line monitoring of naphthalene and salicylate bioavailability with an immobilized bioluminescent catabolic reporter bacterium. *Applied and Environmental Microbiology, 60*(5), 1487-1494. Retrieved from https://pubmed.ncbi.nlm.nih.gov/8017932

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC201507/

Hirsch, C. F., & Christensen, D. L. (1983). Novel method for selective isolation of actinomycetes. *Appl Environ Microbiol, 46*(4), 925-929. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/6639036

Hoefler, B. C., Gorzelnik, K. V., Yang, J. Y., Hendricks, N., Dorrestein, P. C., & Straight, P. D. (2012). Enzymatic resistance to the lipopeptide surfactin as identified through imaging mass spectrometry of bacterial competition. *Proc Natl Acad Sci U S A, 109*(32), 13082-13087. doi:10.1073/pnas.1205586109

Hoshino, S., Wakimoto, T., Onaka, H., & Abe, I. (2015). Chojalactones A-C, cytotoxic butanolides isolated from Streptomyces sp. cultivated with mycolic acid containing bacterium. *Org Lett, 17*(6), 1501-1504. doi:10.1021/acs.orglett.5b00385

Jakimowicz, D., & van Wezel, G. P. (2012). Cell division and DNA segregation in Streptomyces: how to build a septum in the middle of nowhere? *Mol Microbiol, 85*(3), 393-404. doi:10.1111/j.1365-2958.2012.08107.x

Jones, S. E., Ho, L., Rees, C. A., Hill, J. E., Nodwell, J. R., & Elliot, M. A. (2017). Streptomyces exploration is triggered by fungal interactions and volatile signals. *Elife, 6*. doi:10.7554/eLife.21738

Kalakoutskii, L. V., & Agre, N. S. (1976). Comparative aspects of development and differentiation in actinomycetes. *Bacteriological reviews, 40*(2), 469-524. Retrieved from https://pubmed.ncbi.nlm.nih.gov/786257

https://www.ncbi.nlm.nih.gov/pmc/articles/PMC413963/

Kaltenpoth, M. (2009). Actinobacteria as mutualists: general healthcare for insects? *Trends in Microbiology, 17*(12), 529-535. doi:https://doi.org/10.1016/j.tim.2009.09.006

Kanehisa, M., & Sato, Y. (2020). KEGG Mapper for inferring cellular functions from protein sequences. *Protein Sci, 29*(1), 28-35. doi:10.1002/pro.3711

Kanehisa, M., Sato, Y., & Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *Journal of Molecular Biology, 428*(4), 726-731. doi:https://doi.org/10.1016/j.jmb.2015.11.006

Kaufenstein, M., van der Laan, M., & Graumann, P. L. (2011). The Three-Layered DNA Uptake Machinery at the Cell Pole in Competent Bacillus subtilis Cells Is a Stable Complex. *Journal of Bacteriology, 193*(7), 1633. doi:10.1128/JB.01128-10

Khare, E., Mishra, J., & Arora, N. K. (2018). Multifaceted Interactions Between Endophytes and Plant: Developments and Prospects. *Frontiers in Microbiology, 9*, 2732. doi:10.3389/fmicb.2018.02732

Kim, J., Lim, J., & Lee, C. (2013). Quantitative real-time PCR approaches for microbial community studies in wastewater treatment systems: Applications and considerations. *Biotechnology Advances, 31*(8), 1358-1373. doi:https://doi.org/10.1016/j.biotechadv.2013.05.010

Kroiss, J., Kaltenpoth, M., Schneider, B., Schwinger, M.-G., Hertweck, C., Maddula, R. K., . . . Svatoš, A. (2010). Symbiotic streptomycetes provide antibiotic combination prophylaxis for wasp offspring. *Nature Chemical Biology, 6*(4), 261-263. doi:10.1038/nchembio.331

Kumar, S., Stecher, G., Li, M., Knyaz, C., & Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol, 35*(6), 1547-1549. doi:10.1093/molbev/msy096

Lang, B. F., & Burger, G. (2012). Chapter One - Mitochondrial and Eukaryotic Origins: A Critical Review. In L. Maréchal-Drouard (Ed.), *Advances in Botanical Research* (Vol. 63, pp. 1-20): Academic Press.

Lischer, H. E. L., & Shimizu, K. K. (2017). Reference-guided de novo assembly approach improves genome reconstruction for related species. *BMC Bioinformatics, 18*(1), 474. doi:10.1186/s12859-017-1911-6

Liu, X., Germaine, K. J., Ryan, D., & Dowling, D. N. (2010). Whole-cell fluorescent biosensors for bioavailability and biodegradation of polychlorinated biphenyls. *Sensors (Basel), 10*(2), 1377-1398. doi:10.3390/s100201377

Loman, N. J., & Pallen, M. J. (2015). Twenty years of bacterial genome sequencing. *Nat Rev Microbiol, 13*(12), 787-794. doi:10.1038/nrmicro3565

López-García, P., Eme, L., & Moreira, D. (2017). Symbiosis in eukaryotic evolution. *Journal of Theoretical Biology, 434*, 20-33. doi:https://doi.org/10.1016/j.jtbi.2017.02.031

Lopez-Madrigal, S., Balmand, S., Latorre, A., Heddi, A., Moya, A., & Gil, R. (2013). How Does Tremblaya princeps Get Essential Proteins from Its Nested Partner Moranella endobia in the Mealybug Planoccocus citri? *Plos One, 8*(10). doi:ARTN e77307

10.1371/journal.pone.0077307

Lopez-Madrigal, S., Latorre, A., Porcar, M., Moya, A., & Gil, R. (2013). Mealybugs nested endosymbiosis: going into the 'matryoshka' system in Planococcus citri in depth. *Bmc Microbiology, 13*. doi:Artn 74

10.1186/1471-2180-13-74

Lozano, G. L., Bravo, J. I., Garavito Diago, M. F., Park, H. B., Hurley, A., Peterson, S. B., . . . Handelsman, J. (2019). Introducing THOR, a Model Microbiome for Genetic Dissection of Community Behavior. *Mbio, 10*(2). doi:10.1128/mBio.02846-18

Lozano, G. L., Bravo, J. I., & Handelsman, J. (2017). Draft Genome Sequence of Pseudomonas koreensis CI12, a Bacillus cereus "Hitchhiker" from the Soybean Rhizosphere. *Genome Announc, 5*(26). doi:10.1128/genomeA.00570-17

Luti, K. J., & Mavituna, F. (2011). Elicitation of Streptomyces coelicolor with dead cells of Bacillus subtilis and Staphylococcus aureus in a bioreactor increases production of undecylprodigiosin. *Appl Microbiol Biotechnol, 90*(2), 461-466. doi:10.1007/s00253-010-3032-2

Luti, K. J. K., & Mavituna, F. (2011). Streptomyces coelicolor increases the production of undecylprodigiosin when interacted with Bacillus subtilis. *Biotechnology Letters, 33*(1), 113-118. doi:10.1007/s10529-010-0401-y

Maathuis, F. J. M. (2009). Physiological functions of mineral macronutrients. *Current Opinion in Plant Biology, 12*(3), 250-258. doi:10.1016/j.pbi.2009.04.003

Madeira, F., Park, Y. M., Lee, J., Buso, N., Gur, T., Madhusoodanan, N., . . . Lopez, R. (2019). The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Research, 47*(W1), W636-W641. doi:10.1093/nar/gkz268

Maiden, M. C., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., . . . Spratt, B. G. (1998). Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A, 95*(6), 3140-3145. doi:10.1073/pnas.95.6.3140

Matz, M. V., Fradkov, A. F., Labas, Y. A., Savitsky, A. P., Zaraisky, A. G., Markelov, M. L., & Lukyanov, S. A. (1999). Fluorescent proteins from nonbioluminescent Anthozoa species. *Nat Biotechnol, 17*(10), 969-973. doi:10.1038/13657

Menzel, P., Ng, K. L., & Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. *Nature Communications, 7*(1), 11257. doi:10.1038/ncomms11257

Moore, S. J., MacDonald, J. T., Wienecke, S., Ishwarbhai, A., Tsipa, A., Aw, R., . . . Freemont, P. S. (2018). Rapid acquisition and model-based analysis of cell-free transcription-translation reactions from nonmodel bacteria. *Proc Natl Acad Sci U S A, 115*(19), E4340-E4349. doi:10.1073/pnas.1715806115

Müller, H., Berg, C., Landa, B. B., Auerbach, A., Moissl-Eichinger, C., & Berg, G. (2015). Plant genotype-specific archaeal and bacterial endophytes but similar Bacillus antagonists colonize Mediterranean olive trees. *Frontiers in Microbiology, 6*, 138-138. doi:10.3389/fmicb.2015.00138

Muyzer, G., Dewaal, E. C., & Uitterlinden, A. G. (1993). Profiling of Complex Microbial-Populations by Denaturing Gradient Gel-Electrophoresis Analysis of Polymerase Chain Reaction-Amplified Genes-Coding for 16s Ribosomal-Rna. *Applied and Environmental Microbiology, 59*(3), 695-700. doi:Doi 10.1128/Aem.59.3.695-700.1993

Ng, P. C., & Kirkness, E. F. (2010). Whole Genome Sequencing. In M. R. Barnes & G. Breen (Eds.), *Genetic Variation: Methods and Protocols* (pp. 215-226). Totowa, NJ: Humana Press.

Nijland, R., Burgess, J. G., Errington, J., & Veening, J.-W. (2010). Transformation of environmental Bacillus subtilis isolates by transiently inducing genetic competence. *Plos One, 5*(3), e9724-e9724. doi:10.1371/journal.pone.0009724

Nowack, E. C. M., & Melkonian, M. (2010). Endosymbiotic associations within protists. *Philosophical Transactions of the Royal Society B: Biological Sciences, 365*(1541), 699-712. doi:doi:10.1098/rstb.2009.0188

Pacheco, A. R., & Segre, D. (2019). A multidimensional perspective on microbial interactions. *Fems Microbiology Letters, 366*(11). doi:10.1093/femsle/fnz125

Parkinson, J. F., Gobin, B., & Hughes, W. O. (2016). Heritability of symbiont density reveals distinct regulatory mechanisms in a tripartite symbiosis. *Ecology and Evolution, 6*(7), 2053-2060. doi:10.1002/ece3.2005

Pfeiffer, S., Pastar, M., Mitter, B., Lippert, K., Hackl, E., Lojan, P., . . . Sessitsch, A. (2014). Improved group-specific primers based on the full SILVA 16S rRNA gene reference database. *Environ Microbiol, 16*(8), 2389-2407. doi:10.1111/1462-2920.12350

Poly, F., Ranjard, L., Nazaret, S., Gourbiere, F., & Monrozier, L. J. (2001). Comparison of nifH gene pools in soils and soil microenvironments with contrasting properties. *Applied and Environmental Microbiology, 67*(5), 2255-2262. doi:Doi 10.1128/Aem.67.5.2255-2262.2001

Puyet, A., Sandoval, H., Lopez, P., Aguilar, A., Martin, J. F., & Espinosa, M. (1987). A Simple Medium for Rapid Regeneration of Bacillus-Subtilis Protoplasts Transformed with Plasmid DNA. *Fems Microbiology Letters, 40*(1), 1-5. Retrieved from <Go to ISI>://WOS:A1987F545800001

Rainey, S. M., Shah, P., Kohl, A., & Dietrich, I. (2014). Understanding the Wolbachia-mediated inhibition of arboviruses in mosquitoes: progress and challenges. *Journal of General Virology, 95*(3), 517-530. doi:https://doi.org/10.1099/vir.0.057422-0

Raymond, J., Siefert, J. L., Staples, C. R., & Blankenship, R. E. (2004). The Natural History of Nitrogen Fixation. *Molecular Biology and Evolution, 21*(3), 541-554. doi:10.1093/molbev/msh047

Reasoner, D. J., & Geldreich, E. E. (1985). A new medium for the enumeration and subculture of bacteria from potable water. *Appl Environ Microbiol, 49*(1), 1-7.

Reed, S. C., Cleveland, C. C., & Townsend, A. R. (2011). Functional Ecology of Free-Living Nitrogen Fixation: A Contemporary Perspective. *Annual Review of Ecology, Evolution, and Systematics, 42*(1), 489-512. doi:10.1146/annurev-ecolsys-102710-145034

Saide, J. A. O., & Gilliland, S. E. (2005). Antioxidative Activity of Lactobacilli Measured by Oxygen Radical Absorbance Capacity*. *Journal of Dairy Science, 88*(4), 1352-1357. doi:https://doi.org/10.3168/jds.S0022-0302(05)72801-0

Saikia, S. P., & Jain, V. (2007). Biological nitrogen fixation with non-legumes: An achievable target or a dogma? *Current Science, 92*(3), 317-322. Retrieved from <Go to ISI>://WOS:000244429400017

Schneider, J., Yepes, A., Garcia-Betancur, J. C., Westedt, I., Mielich, B., & López, D. (2012). Streptomycin-induced expression in Bacillus subtilis of YtnP, a lactonase-homologous protein that inhibits development and streptomycin production in Streptomyces griseus. *Applied and Environmental Microbiology, 78*(2), 599-603. doi:10.1128/AEM.06992-11

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics, 30*(14), 2068-2069. doi:10.1093/bioinformatics/btu153

Seipke, R. F., Kaltenpoth, M., & Hutchings, M. I. (2012). Streptomyces as symbionts: an emerging and widespread theme? *FEMS Microbiology Reviews, 36*(4), 862-876. doi:10.1111/j.1574-6976.2011.00313.x

Setlow, P. (2006). Spores of Bacillus subtilis: their resistance to and killing by radiation, heat and chemicals. *Journal of Applied Microbiology, 101*(3), 514-525. doi:10.1111/j.1365-2672.2005.02736.x

Shepherd, M. D., Kharel, M. K., Bosserman, M. A., & Rohr, J. (2010). Laboratory maintenance of Streptomyces species. *Current protocols in microbiology, Chapter 10*, Unit-10E.11. doi:10.1002/9780471729259.mc10e01s18

Shin, D., Byun, W. S., Moon, K., Kwon, Y., Bae, M., Um, S., . . . Oh, D. C. (2018). Coculture of Marine Streptomyces sp With Bacillus sp Produces a New Piperazic Acid-Bearing Cyclic Peptide. *Frontiers in Chemistry, 6*. doi:ARTN 498

10.3389/fchem.2018.00498

Singhal, N., Kumar, M., Kanaujia, P. K., & Virdi, J. S. (2015). MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis. *Frontiers in Microbiology, 6*, 791. doi:10.3389/fmicb.2015.00791

Smercina, D. N., Evans, S. E., Friesen, M. L., & Tiemann, L. K. (2019). To Fix or Not To Fix: Controls on Free-Living Nitrogen Fixation in the Rhizosphere (vol 85, e02546-18, 2019). *Applied and Environmental Microbiology, 85*(22). doi:ARTN e02103-19

10.1128/AEM.02103-19

Sorensen, S. J., Bailey, M., Hansen, L. H., Kroer, N., & Wuertz, S. (2005). Studying plasmid horizontal transfer in situ: a critical review. *Nat Rev Microbiol, 3*(9), 700-710. doi:10.1038/nrmicro1232

Steenhoudt, O., & Vanderleyden, J. (2000). Azospirillum, a free-living nitrogen-fixing bacterium closely associated with grasses: genetic, biochemical and ecological aspects. *FEMS Microbiol Rev, 24*(4), 487-506. doi:10.1111/j.1574-6976.2000.tb00552.x

Stewart, W. D., Fitzgerald, G. P., & Burris, R. H. (1967). In situ studies on N2 fixation using the acetylene reduction technique. *Proc Natl Acad Sci U S A, 58*(5), 2071-2078. doi:10.1073/pnas.58.5.2071

Stubbendieck, R. M., Vargas-Bautista, C., & Straight, P. D. (2016). Bacterial Communities: Interactions to Scale. *Frontiers in Microbiology, 7*, 1234. doi:10.3389/fmicb.2016.01234

Tang, X., Huang, Y., Lei, J., Luo, H., & Zhu, X. (2019). The single-cell sequencing: new developments and medical applications. *Cell & Bioscience, 9*(1), 53. doi:10.1186/s13578-019-0314-y

Tokala, R. K., Strap, J. L., Jung, C. M., Crawford, D. L., Salove, M. H., Deobald, L. A., . . . Morra, M. J. (2002). Novel Plant-Microbe Rhizosphere Interaction Involving Streptomyces lydicus WYEC108 and the Pea Plant (Pisum sativum). *Applied and Environmental Microbiology, 68*(5), 2161-2171. doi:10.1128/aem.68.5.2161-2171.2002

Torres-Corral, Y., Fernandez-Alvarez, C., & Santos, Y. (2019). High-throughput identification and quantification of Vagococcus salmoninarum by SYBR Green I-based

real-time PCR combined with melting curve analysis. *J Fish Dis, 42*(10), 1359-1368. doi:10.1111/jfd.13053

Traxler, M. F., Watrous, J. D., Alexandrov, T., Dorrestein, P. C., & Kolter, R. (2013). Interspecies Interactions Stimulate Diversification of the Streptomyces coelicolor Secreted Metabolome. *Mbio, 4*(4). doi:ARTN e00459-13

10.1128/mBio.00459-13

Unkovich, M., & Baldock, J. (2008). Measurement of asymbiotic N2 fixation in Australian agriculture. *Soil Biology and Biochemistry, 40*(12), 2915-2921. doi:https://doi.org/10.1016/j.soilbio.2008.08.021

van der Heijen, M. G. A. (2008). The unseen majority: Soil microbes as drivers of plant diversity and productivity in terrestrial ecosystems (vol 11, pg 296, 2008). *Ecology Letters, 11*(6), 651-651. doi:10.1111/j.1461-0248.2008.01199.x

Ventura, M., Canchaya, C., Tauch, A., Chandra, G., Fitzgerald, G. F., Chater, K. F., & van Sinderen, D. (2007). Genomics of Actinobacteria: tracing the evolutionary history of an ancient phylum. *Microbiol Mol Biol Rev, 71*(3), 495-548. doi:10.1128/mmbr.00005-07

Wang, L., Zhang, L., Liu, Z., Zhao, D., Liu, X., Zhang, B., . . . Li, J. (2013). A minimal nitrogen fixation gene cluster from Paenibacillus sp. WLY78 enables expression of active nitrogenase in Escherichia coli. *PLoS Genet, 9*(10), e1003865. doi:10.1371/journal.pgen.1003865

Weisburg, W. G., Barns, S. M., Pelletier, D. A., & Lane, D. J. (1991). 16s Ribosomal DNA Amplification for Phylogenetic Study. *Journal of Bacteriology, 173*(2), 697-703. doi:Doi 10.1128/Jb.173.2.697-703.1991

Wick, R. R., Judd, L. M., Gorrie, C. L., & Holt, K. E. (2017). Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol, 13*(6), e1005595. doi:10.1371/journal.pcbi.1005595

Wood, K. V. (1995). Marker proteins for gene expression. *Curr Opin Biotechnol, 6*(1), 50-58. doi:10.1016/0958-1669(95)80009-3

Wujek, D. E. (1979). Intracellular Bacteria in the Blue-Green Alga Pleurocapsa minor. *Transactions of the American Microscopical Society, 98*(1), 143-145. doi:10.2307/3225953

Yim, G., Wang, H. H., & Davies, J. (2007). Antibiotics as signalling molecules. *Philos Trans R Soc Lond B Biol Sci, 362*(1483), 1195-1200. doi:10.1098/rstb.2007.2044

Yoshida, N., Inaba, S., & Takagi, H. (2014). Utilization of atmospheric ammonia by an extremely oligotrophic bacterium, Rhodococcus erythropolis N9T-4. *J Biosci Bioeng, 117*(1), 28-32. doi:10.1016/j.jbiosc.2013.06.005

Zachar, I., & Boza, G. (2020). Endosymbiosis before eukaryotes: mitochondrial establishment in protoeukaryotes. *Cellular and Molecular Life Sciences*. doi:10.1007/s00018-020-03462-6

Zhang, Z., Ding, Z. T., Shu, D., Luo, D., & Tan, H. (2015). Development of an efficient electroporation method for iturin A-producing Bacillus subtilis ZK. *Int J Mol Sci, 16*(4), 7334-7351. doi:10.3390/ijms16047334

Zinchuk, V., Zinchuk, O., & Okada, T. (2007). Quantitative colocalization analysis of multicolor confocal immunofluorescence microscopy images: pushing pixels to explore biological phenomena. *Acta Histochem Cytochem, 40*(4), 101-111. doi:10.1267/ahc.07002

APPENDIX

1. Appendix 1a: Ratio of Firmicutes to Actinomycetes Ct value of 321I.

	321 Days R2A	Firmi	Actino	Ratio	Avg R2A	321I Days NFA	Firmi	Actino	Ratio	Avg NFA
Rep 1	1	14.370	15.012	0.957		3	14.733	15.906	0.926	
	1	13.531	14.136	0.957		3	13.336	17.184	0.776	
	1	13.782	16.026	0.860		3	13.726	15.321	0.896	
Rep 2	1	14.640	14.695	0.996		3	15.608	14.878	1.049	
	1	14.656	14.904	0.983		3	15.537	15.170	1.024	
	1	15.043	14.899	1.010		3	14.244	29.022	0.491	
Rep 3	1	13.085	13.449	0.973		3	13.559	15.274	0.888	
	1	13.475	14.008	0.962		3	13.539	15.033	0.901	
	1	12.816	13.890	0.923	0.958	3	13.276	15.114	0.878	0.870
Rep 1	3	13.907	14.039	0.991		5	17.576	15.255	1.152	
	3	15.066	14.882	1.012		5	16.327	14.592	1.119	
	3	17.476	15.068	1.160		5	17.049	14.923	1.142	
Rep 2	3	14.733	15.601	0.944		5	14.992	14.979	1.001	
	3	14.766	16.776	0.880		5	14.727	15.351	0.959	
	3	15.045	14.764	1.019		5	14.891	15.093	0.987	
Rep 3	3	13.254	13.838	0.958		5	12.978	14.641	0.886	
	3	13.495	15.352	0.879		5	12.820	14.715	0.871	
	3	13.047	13.258	0.984	0.981	5	13.066	14.622	0.894	1.001
Rep 1	5	14.541	14.370	1.012		7	16.535	14.129	1.170	
	5	18.389	15.071	1.220		7	13.721	13.834	0.992	
	5	16.384	13.902	1.179		7	14.673	14.056	1.044	
Rep 2	5	16.643	16.392	1.015		7	24.809	24.918	0.996	
	5	16.055	16.250	0.988		7	24.815	25.030	0.991	
	5	16.424	16.693	0.984		7	24.512	25.305	0.969	
Rep 3	5	22.225	22.782	0.976		7	13.277	15.064	0.881	
	5	22.081	22.963	0.962		7	13.963	14.988	0.932	
	5	22.212	23.004	0.966	1.033	7	13.243	14.948	0.886	0.985
Rep 1	10	15.365	16.189	0.949		12	14.034	14.357	0.978	
	10	16.001	15.162	1.055		12	13.880	14.090	0.985	
	10	16.497	15.655	1.054		12	14.441	14.316	1.009	
Rep 2	10	15.157	15.083	1.005		12	27.978	28.330	0.988	
	10	15.078	15.106	0.998		12	27.822	28.118	0.989	

	10	15.125	15.208	0.995		12	28.474	28.827	0.988	
Rep 3	10	12.128	13.624	0.890		12	12.496	14.918	0.838	
	10	12.025	13.960	0.861		12	12.588	14.696	0.857	
	10	12.267	14.228	0.862	0.963	12	12.755	14.970	0.852	0.942

2. Appendix 1b: Ratio of Firmicutes to Actinomycetes Ct value of 34I.

	34I Days R2A	Firmi	Actino	Ratio	Avg R2A	321I Days NFA	Firmi	Actino	Ratio	Avg NFA
Rep 1	1	12.890	13.348	0.966		3	14.326	14.431	0.993	
	1	13.265	13.278	0.999		3	14.384	14.300	1.006	
	1	14.250	13.837	1.030		3	15.831	14.505	1.091	
Rep 2	1	12.110	10.568	1.146		3	13.626	15.401	0.885	
	1	12.684	10.907	1.163		3	14.138	16.102	0.878	
	1	13.014	11.117	1.171		3	14.385	15.616	0.921	
Rep 3	1	12.468	13.435	0.928		3	13.202	14.993	0.881	
	1	12.526	15.971	0.784		3	13.064	15.610	0.837	
	1	13.162	14.563	0.904	1.010	3	13.417	15.401	0.871	0.929
Rep 1	3	13.361	13.810	0.967		5	13.641	14.220	0.959	
	3	13.437	13.877	0.968		5	14.164	14.322	0.989	
	3	13.407	13.872	0.966		5	13.664	15.238	0.897	
Rep 2	3	13.037	11.320	1.152		5	13.994	15.654	0.894	
	3	13.712	10.829	1.266		5	14.237	15.455	0.921	
	3	14.413	11.124	1.296		5	14.146	15.701	0.901	
Rep 3	3	11.697	16.307	0.717		5	12.619	14.484	0.871	
	3	11.430	16.456	0.695		5	12.842	14.865	0.864	
	3	11.319	16.390	0.691	0.969	5	12.807	14.869	0.861	0.906
Rep 1	5	12.674	13.917	0.911		7	13.986	14.472	0.966	
	5	12.734	13.764	0.925		7	14.395	14.206	1.013	
	5	13.019	13.469	0.967		7	14.199	14.683	0.967	
Rep 2	5	14.253	13.208	1.079		7	13.834	15.300	0.904	
	5	15.202	12.640	1.203		7	13.683	15.295	0.895	
	5	14.727	13.047	1.129		7	13.957	15.398	0.906	
Rep 3	5	12.459	12.637	0.986		7	13.064	14.874	0.878	
	5	12.246	13.114	0.934		7	13.561	15.177	0.894	
	5	12.115	12.887	0.940	1.008	7	13.505	15.067	0.896	0.924
Rep 1	10	13.531	13.615	0.994		12	14.226	14.014	1.015	
	10	13.302	13.635	0.976		12	14.166	14.276	0.992	
	10	13.487	13.657	0.988		12	13.622	14.381	0.947	
Rep 2	10	11.366	11.433	0.994		12	13.187	14.724	0.896	

	10	11.048	11.309	0.977		12	12.842	14.854	0.865	
	10	11.210	11.224	0.999		12	13.339	14.888	0.896	
Rep 3	10	12.268	13.176	0.931		12	13.166	14.847	0.887	
	10	12.966	13.094	0.990		12	13.129	14.965	0.877	
	10	12.339	13.543	0.911	0.973	12	13.276	15.135	0.877	0.917

3. Appendix 1c: Ratio of Firmicutes to Actinomycetes Ct value of 40I.

	321 Days R2A	Firmi	Actino	Ratio	Avg R2A	321I Days NFA	Firmi	Actino	Ratio	Avg NFA
Rep 1	1	10.838	10.388	1.043		3	11.017	10.612	1.038	
	1	10.724	10.390	1.032		3	10.679	10.725	0.996	
	1	10.791	10.673	1.011		3	10.923	10.741	1.017	
Rep 2	1	10.838	10.388	1.043		3	11.017	10.612	1.038	
	1	10.724	10.390	1.032		3	10.679	10.725	0.996	
	1	10.791	10.673	1.011		3	10.923	10.741	1.017	
Rep 3	1	12.388	13.799	0.898		3	13.014	15.341	0.848	
	1	12.237	15.630	0.783		3	12.961	14.840	0.873	
	1	12.523	18.512	0.677	0.948	3	12.955	14.975	0.865	0.965
Rep 1	3	11.047	10.627	1.040		5	11.455	11.138	1.028	
	3	11.138	10.731	1.038		5	11.397	10.990	1.037	
	3	11.097	10.967	1.012		5	11.183	11.163	1.002	
Rep 2	3	11.047	10.627	1.040		5	11.455	11.138	1.028	
	3	11.138	10.731	1.038		5	11.397	10.990	1.037	
	3	11.097	10.967	1.012		5	11.183	11.163	1.002	
Rep 3	3	12.693	13.676	0.928		5	13.307	15.867	0.839	
	3	13.314	13.962	0.954		5	13.575	15.672	0.866	
	3	13.319	13.508	0.986	1.005	5	13.751	15.557	0.884	0.969
Rep 1	5	18.098	17.245	1.049		7	11.150	10.509	1.061	
	5	17.812	17.471	1.020		7	11.086	10.420	1.064	
	5	17.724	17.513	1.012		7	11.130	10.346	1.076	
Rep 2	5	18.098	17.245	1.049		7	11.150	10.509	1.061	
	5	17.812	17.471	1.020		7	11.086	10.420	1.064	
	5	17.724	17.513	1.012		7	11.130	10.346	1.076	
Rep 3	5	12.287	12.992	0.946		7	13.045	14.821	0.880	
	5	12.086	13.009	0.929		7	13.009	14.804	0.879	
	5	12.225	12.848	0.951	0.999	7	13.021	14.849	0.877	1.004
Rep 1	10	10.688	10.298	1.038		12	11.109	10.602	1.048	
	10	10.690	10.540	1.014		12	10.965	10.544	1.040	
	10	10.481	10.690	0.980		12	10.920	10.572	1.033	
Rep 2	10	10.688	10.298	1.038		12	11.109	10.602	1.048	

	10	10.690	10.540	1.014		12	10.965	10.544	1.040	
	10	10.481	10.690	0.980		12	10.920	10.572	1.033	
Rep 3	10	13.172	13.113	1.004		12	13.230	15.324	0.863	
	10	12.757	13.006	0.981		12	13.251	15.085	0.878	
	10	12.926	12.973	0.996	1.005	12	13.531	14.979	0.903	0.987

4. Appendix 2: Formation of Protoplast with lysozyme treatment

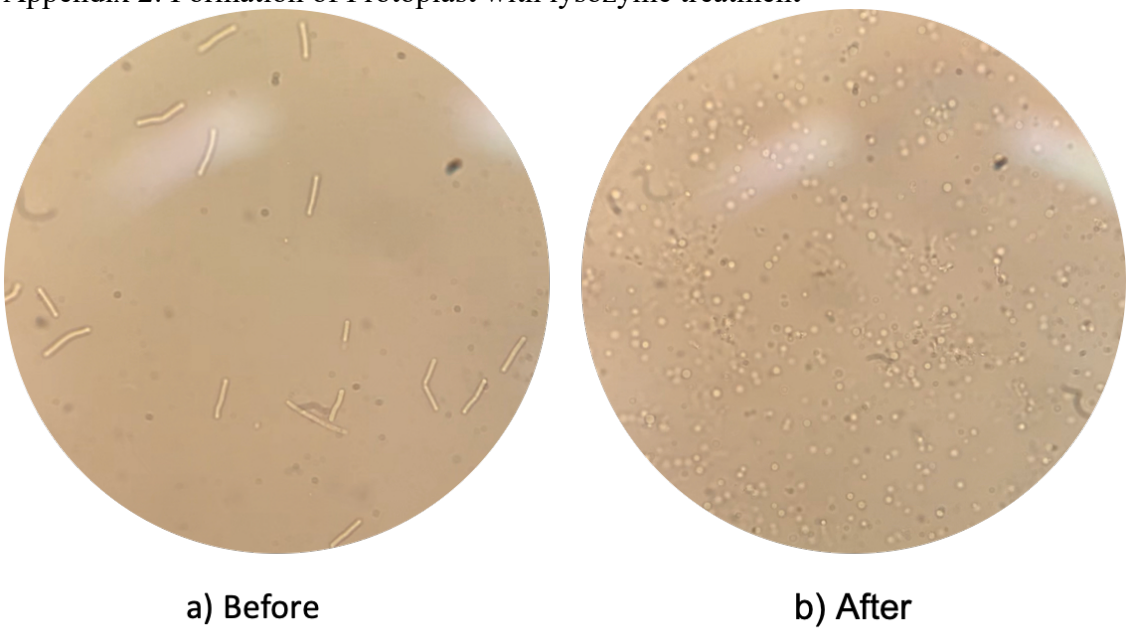


Fig 1: Bacillus cells under phase contrast microscopy before (a) and after (b) treatment with lysozyme. Protoplast formation is represented in (b).

5. Appendix 3: Quality score across all bases and Mean GC distribution of *Bacillus* 321Band 40B

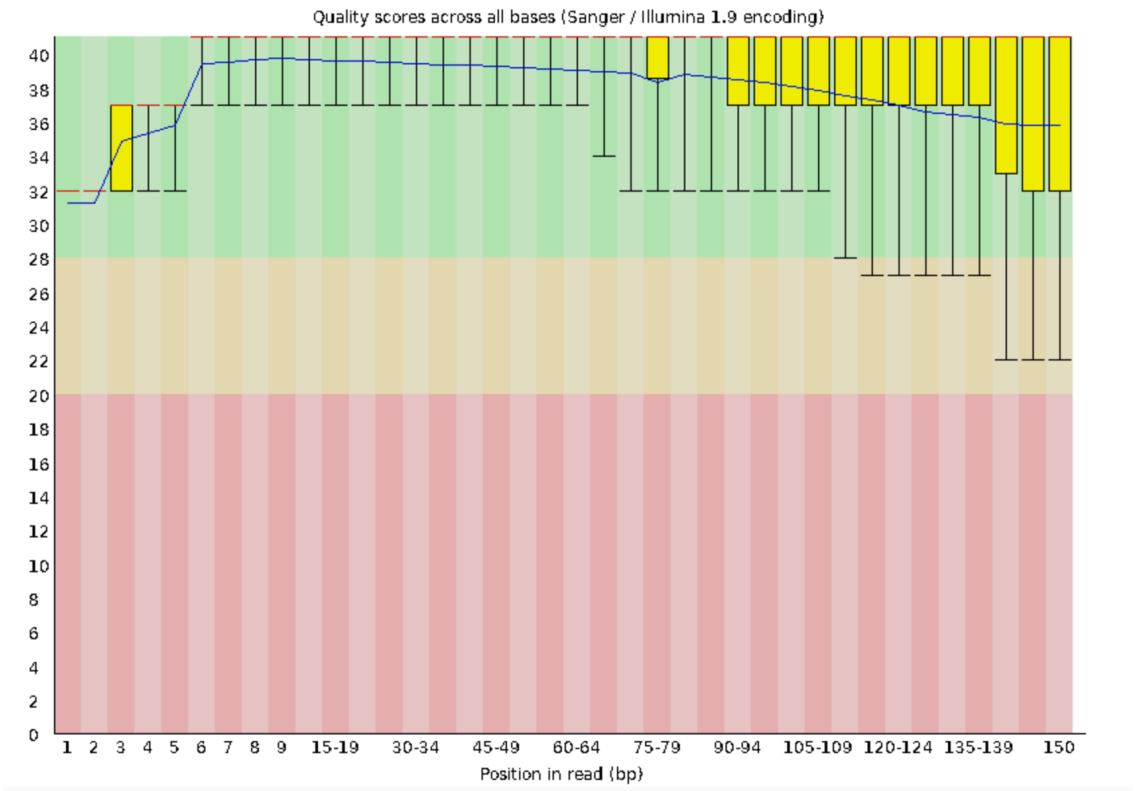


Fig 2: Quality score across all bases of 321B (R1)

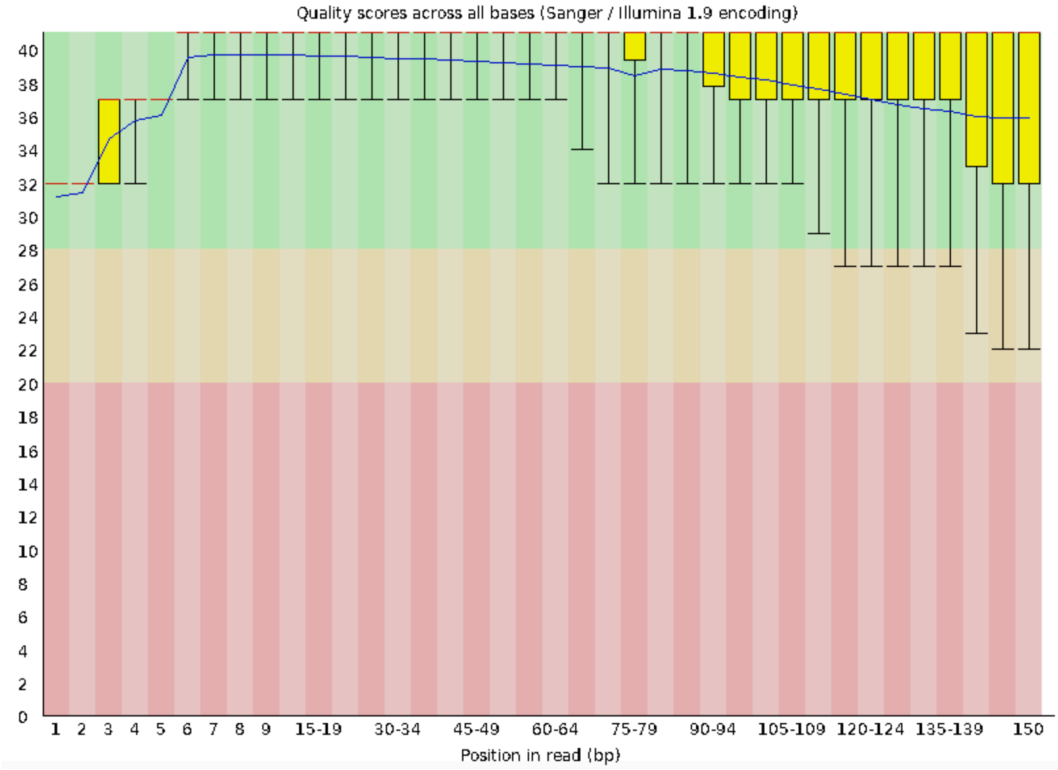
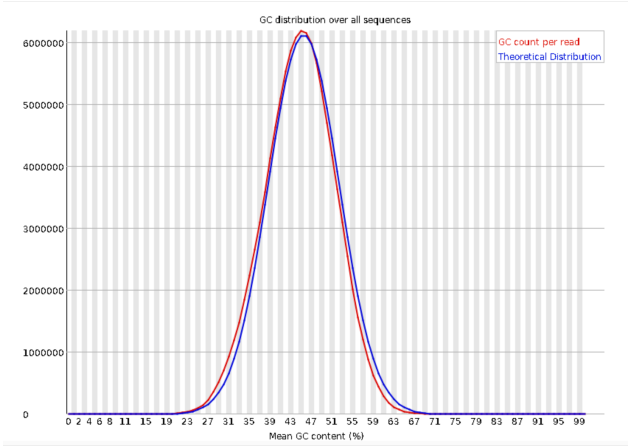


Fig 3: Quality score across all bases of 40B (R1)



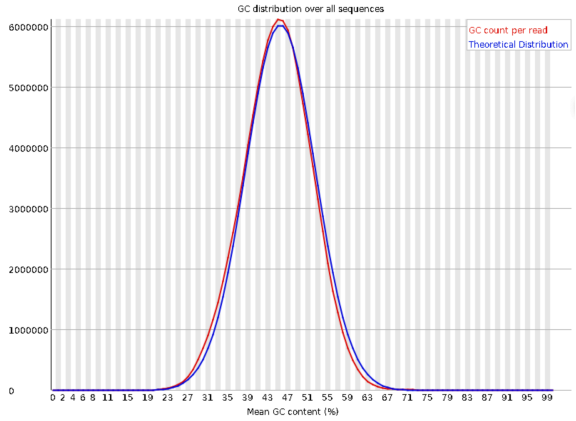


Fig 4: Mean GC distribution of 321B R1 and R2 trimmed reads.

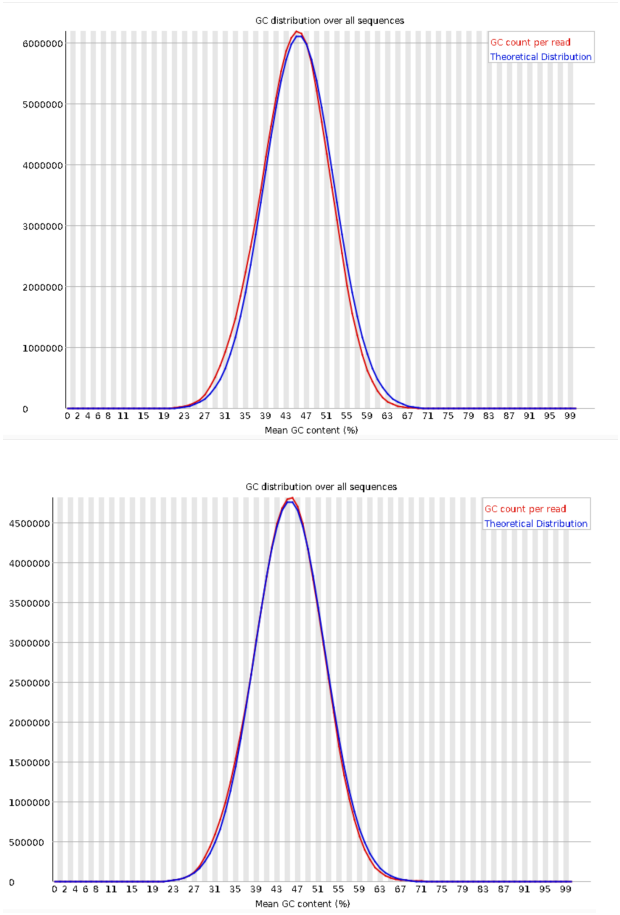


Fig 5: Mean GC distribution of 40B R1 and R2 trimmed reads.