

Institut für Biochemie und Biologie



*Development of Concepts for the Genomic  
Mining of Novel Secondary Metabolites in  
Symbiotic Cyanobacteria*

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum naturalium  
(Dr. rer. nat.)

im Fach Mikrobiologie

eingereicht an der  
Mathematisch-Naturwissenschaftlichen Fakultät  
der Universität Potsdam

von

M.Sc. Daniel Dehm

Geboren am 1.6.1990 in Würzburg

This work is licensed under a Creative Commons License:  
Attribution-NonCommercial-NoDerivatives 4.0 International.  
This does not apply to quoted content from other authors.  
To view a copy of this license visit  
<https://creativecommons.org/licenses/by-nc-nd/4.0/>

### **Referees**

**Prof. Dr. Elke Dittmann**, University of Potsdam: Institute of Biochemistry and Biology,  
Department for Microbiology

**Prof. Dr. Helge B. Bode**, University of Frankfurt am Main: Institute for Molecular Bio  
Science, Department for Molecular Biotechnology

**Dr. Dennis J. Nürnberg**, Freie Universität Berlin: Institute of Experimental Physics,  
Department for Biophysics and Photosynthesis

**Thesis defense held on 2.10.2020.**

Published online on the  
Publication Server of the University of Potsdam:  
<https://doi.org/10.25932/publishup-47834>  
<https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-478342>

## Deutsche Zusammenfassung

Naturstoffe sind seit der goldenen Ära der Antibiotika von immer größerem Interesse, sowohl für die Grundlagenforschung als auch die Angewandten Wissenschaften, da sie die Hauptquelle für neuartige Pharmazeutika mit starken antibiotischen, anti-entzündlichen und Antitumor-Aktivitäten darstellen. Neben den technologischen Fortschritten im Bereich der Hochdurchsatz Genomsequenzierung und dem verbesserten Verständnis des modularen Aufbaus der Biosynthesewege von Sekundärmetaboliten, kam es auch zu einem Wechsel vom labor-gestützten Screening aktiver Zellextrakte hin zum Algorithmen-basierten *in silico* Screening nach neuen Naturstoff-Biosyntheseclustern. Obwohl die steigende Zahl verfügbarer Genomsequenzen zeigte, dass nicht-ribosomale Peptid-Synthetasen (NRPS), Polyketid-Synthasen (PKS), und ribosomal synthetisierte und posttranslational modifizierte Peptide (RiPPs) ubiquitär in allen Sparten des Lebens gefunden werden können, so zeigen einige Phyla wie Actinobakterien oder Cyanobakterien eine besonders hohe Dichte an Sekundärmetabolitclustern.

Der fakultativ symbiotische, N<sub>2</sub>-fixierende Modellorganismus *N. punctiforme* PCC73102 ist ein terrestrisches typ-IV Cyanobakterium, welches nicht nur einen besonders hohen Anteil seines Genoms der Produktion von Sekundärmetaboliten widmet, sondern zusätzlich noch genetisch modifizierbar ist. Eine AntiSMASH Analyse des Genoms zeigte, dass *N. punctiforme* insgesamt sechzehn potentielle Sekundärmetabolitcluster besitzt, von denen aber bis heute nur zweien ein spezifisches Produkt zugewiesen werden konnte. Das macht *N. punctiforme* zu einem perfekten Testorganismus für die Entwicklung eines neuartigen kombinatorischen Genomic Mining Ansatzes zur Detektion von bislang unbeschriebenen Naturstoffen.

Der neuartige Ansatz, der im Rahmen dieser Studie entwickelt wurde, stellt eine Kombination aus Genomic Mining, unabhängigen Monitoring-Techniken sowie modifizierten Kultivierungsbedingungen dar und führte nicht nur zu neuen Erkenntnissen im Bereich cyanobakterieller Naturstoffsynthese, sondern letztlich auch zur Entdeckung eines neuen, von *N. punctiforme* produzierten, Naturstoffs. Die Herstellung und Untersuchung einer Reporterstamm Bibliothek, bestehend aus je einem CFP-produzierenden Transkriptionsreporter für jedes der sechzehn Sekundärmetabolitcluster von *N. punctiforme*, zeigte, dass im Gegensatz zur Erwartung nicht alle Biosynthesecluster für die man kein Produkt nachweisen kann auch nicht exprimiert werden. Stattdessen konnten klar definierbare Expressionsmuster

beschrieben werden, was deutlich machte, dass die Naturstoffproduktion einer engen Regulation unterliegt und nur ein kleiner Teil der Biosynthesecluster unter Standardbedingungen tatsächlich still sind. Darüber hinaus führte die Erhöhung der Lichtintensität sowie der Kohlenstoffdioxid-Verfügbarkeit zusammen mit der Kultivierung von *N. punctiforme* zu extrem hohen Zelldichten zu einer starken Erhöhung der gesamten metabolischen Aktivität des Organismus. Nähere Untersuchungen der Zellextrakte dieser hoch-dichte Kultivierungen führten letztlich zur Entdeckung einer neuartigen Gruppe von Microviridinen mit verlängerter Peptidsequenz, welche Microviridin N3-N9 genannt wurden. Sowohl die Kultivierung der Transkriptionsreporter als auch die RTqPCR-basierte Untersuchung der Transkriptionslevel der verschiedenen Biosynthesecluster zeigten, dass die hoch-Zelldichte Kultivierung von *N. punctiforme* zu einer Aktivierung von 50% der vorhandenen Sekundärmetabolitcluster führt. Im Gegensatz zu dieser sehr breitgefächerten Aktivierung, führt die Co-Kultivierung von *N. punctiforme* in chemischen oder physischen Kontakt zu einer N-gehungerten Wirtspflanze (*Blasia pusilla*) zu einer sehr spezifischen Aktivierung der RiPP4 und RiPP3 Biosynthesecluster. Obwohl dieser Effekt mittels verschiedener unabhängiger Methoden bestätigt werden konnte und trotz intensiver Analysebemühungen, konnte jedoch keinem der beiden Cluster ein Produkt zugeordnet werden.

Diese Studie stellt die erste weitreichende, systematische Analyse eines cyanobakteriellen Sekundärmetaboloms durch einen kombinatorischen Ansatz aus Genomic Mining und unabhängigen Monitoring-Techniken dar und kann als neue strategische Herangehensweise für die Untersuchung anderer Organismen hinsichtlich ihrer Sekundärmetabolit-Produktion dienen. Obwohl es bereits gut beschriebene einzelne Sekundärmetabolite gibt, wie beispielweise den Zelldifferenzierungsfaktor PatS in *Anabaena* sp. PCC7120, so ist der Grad an Regulation der in dieser Studie gezeigt werden konnte bislang beispiellos und die Entschlüsselung dieser Mechanismen könnte die Entdeckung neuer Naturstoffe stark beschleunigen. Daneben lassen die Ergebnisse aber auch darauf schließen, dass die Induktion der Biosynthesewege nicht das eigentliche Problem darstellt, sondern vielmehr die verlässliche Detektion deren Produkte. Die Erarbeitung neuer Analytik-Strategien könnte somit auch einen deutlichen Einfluss auf die Geschwindigkeit der Entdeckung neuer Naturstoffe haben.

## Abstract

Since the golden era of antibiotics natural products are of ever growing interest to both basic research and applied sciences as they are the main source of new bioactive compounds delivering lead structures for new pharmaceuticals with potent antibiotic, anti-inflammatory or anti-cancer activities. Alongside the technological advances in high-throughput genome sequencing and the better understanding of the general organization of those modular biosynthetic assembly lines of secondary metabolites, there was also a shift from wet-lab screening of active cell extracts towards algorithm-based *in silico* screening for new natural product biosynthesis gene clusters (BGCs). Although the increasing availability of full genome sequences revealed that such non-ribosomal peptide synthetases (NRPS), polyketide synthases (PKS) and ribosomally synthesized and post-translationally modified peptides (RiPPs) can be found in all three kingdoms of life, certain phyla like actinobacteria and cyanobacteria show a very high density of these secondary metabolite BGCs.

The facultative symbiotic, N<sub>2</sub>-fixing model organism *N. punctiforme* PCC73102 is a terrestrial type IV cyanobacterium that not only dedicates a very large fraction of its genome to secondary metabolite production but is also amenable to genetic modification. AntiSMASH analysis of the genome showed that there are sixteen potential secondary metabolite BGCs encoded in *N. punctiforme*, but until now there were only two compounds assigned to their respective BGC leaving the remaining fourteen orphan. This makes the organism a perfect subject for the establishment of a novel combinatorial genomic mining approach for the detection of new natural products.

In the course of this study a combinatorial approach of genomic mining, independent monitoring techniques and alteration of cultivation conditions lead to new insights in cyanobacterial natural product biosynthesis and ultimately to the description of a novel compound produced by *N. punctiforme*. With the generation and investigation of a reporter strain library consisting of CFP-producing transcriptional reporter mutants for every predicted secondary metabolite BGC of *N. punctiforme*, it could be shown that natural product expression is in fact not silent for all those BGCs where no compound can be detected. Instead several distinct expression patterns could be described highlighting that secondary metabolite production is under tight regulation and only a minor fraction of these BGCs is in fact silent under standard laboratory conditions. Furthermore, increasing light intensity and carbon dioxide availability and cultivating *N. punctiforme* to

very high cell densities had a tremendous impact on the overall metabolic activity of the organism. Investigation of high density cultivated cell extracts ultimately lead to the detection of a so far undescribed set of microviridins with unusual extended peptide sequences named Microviridin N3 – N9. Both cultivation of the transcriptional reporter mutants as well as RTqPCR-based detection of secondary metabolite BGC transcription levels revealed that in fact 50% of *N. punctiforme*'s natural product BGCs are upregulated under high cell density conditions. In contrast to this very broad response, co-cultivation of *N. punctiforme* in chemical or physical contact with a N-deprived host plant (*Blasia pusilla*) lead to a very specific upregulation of two natural product BGCs, namely RIPP3 and RIPP4. Although this response could be confirmed by various independent monitoring techniques and heavy analytical efforts were spent, no compound could be assigned to either of these BGCs.

This study is the first in-depth systematic investigation of a cyanobacterial secondary metabolome by a combinatorial approach of genome mining and independent monitoring techniques that can serve as a new strategic approach to gain further insight into natural product synthesis of various organisms. Although there are single well described examples of secondary metabolites like the cell differentiation factor PatS in *Anabaena* sp. strain PCC 7120, the level and extent of regulation observed in this study is unprecedented and understanding of these mechanisms might be the key to streamline natural product discovery. However, the results of this study also highlight that induction of secondary metabolite BGCs is not the real challenge. Instead the new insights point towards analytical issues being a severe hurdle and finding reliable strategies to overcome these problems might as well drive natural product discovery.

## Table of Content

<b>Deutsche Zusammenfassung .....</b>	<b>I</b>
<b>Abstract .....</b>	<b>III</b>
<b>Table of Content .....</b>	<b>V</b>
<b>1 Introduction .....</b>	<b>1</b>
1.1 Secondary Metabolites and Chemical Diversity of Natural Products.....	1
1.1.1 Nonribosomal Peptides and Polyketides .....	1
1.1.2 Ribosomally Synthesized and Post-translationally Modified Peptides.....	3
1.1.3 Chemical Diversity of Natural Products .....	5
1.2 Tools for Natural Product Discovery .....	6
1.2.1 Classical Screening Approaches .....	7
1.2.2 Biotechnological Approaches .....	8
1.2.3 Genomic Mining .....	8
1.3 Cyanobacteria are a Rich Source of High Value Compounds.....	10
1.3.1 Photoautotrophic bacteria.....	10
1.3.2 Cyanobacterial natural products.....	10
1.3.3 The Genus <i>Nostoc</i> and the Model Organism <i>N. punctiforme</i> PCC73102.....	11
1.3.4 Facultative Symbiotic Lifestyle of <i>N. punctiforme</i> PCC73102 .....	14
1.4 Aims of this Study .....	16
<b>2 Material and Methods .....</b>	<b>17</b>
2.1 Material .....	17
2.1.1 Organisms.....	17
2.1.2 Genome sequences, DNA oligos and Primers.....	17
2.1.3 Chemicals .....	18
2.1.4 Media and solutions .....	19
2.1.5 Commercial Kits.....	20
2.1.6 Commercial Enzymes and Auxiliary Products .....	20
2.1.7 Cloning Vectors.....	21
2.1.8 Membranes, Filters, Cartridges and Cuvettes.....	21
2.1.9 Technical Devices .....	21
2.2 Methods .....	23
2.2.1 Culture Maintenance .....	23
2.2.1.1 Cultivation of <i>Nostoc punctiforme</i> PCC73102.....	23
2.2.1.2 High Density cultivation of <i>N. punctiforme</i> .....	23
2.2.1.3 Cryopreservation of <i>N. punctiforme</i> .....	23
2.2.1.4 Cultivation of <i>Escherichia coli</i> .....	24
2.2.1.5 Cryopreservation of <i>E. coli</i> .....	24

2.2.1.6	Cultivation of <i>Blasia pusilla</i> under alternating nitrogen availability .....	24
2.2.2	Molecular biology and cloning techniques for <i>Escherichia coli</i> .....	24
2.2.2.1	Preparative DNA amplification for cloning (Fusion PCR) .....	24
2.2.2.2	Qualitative DNA amplification for genotyping (colony PCR).....	25
2.2.2.3	Agarose gel electrophoresis for DNA .....	26
2.2.2.4	DNA extraction from agarose gels.....	26
2.2.2.5	Preparative plasmid isolation.....	26
2.2.2.6	Digestion of DNA with restriction endonucleases .....	27
2.2.2.7	Spectrophotometric quantification of nucleic acid concentrations .....	27
2.2.2.8	<i>In vitro</i> DNA assembly via homologous recombination .....	27
2.2.2.9	<i>In vivo</i> DNA assembly via homologous recombination (FullRecET).....	28
2.2.2.10	Chemical transformation of <i>Escherichia coli</i> .....	29
2.2.3	Molecular biology and cloning techniques for <i>Nostoc punctiforme</i> PCC73102....	29
2.2.3.1	Isolation of <i>N. punctiforme</i> genomic DNA.....	29
2.2.3.2	Plasmid transfer for the generation of CFP-reporter mutant strains of <i>N. punctiforme</i> .....	30
2.2.3.3	Tri-parental conjugation for knock-out/in mutagenesis of <i>N. punctiforme</i> .....	30
2.2.3.4	Selection process during <i>N. punctiforme</i> mutagenesis .....	32
2.2.3.5	DNA isolation of potential <i>N. punctiforme</i> mutants for genotyping PCR.....	32
2.2.4	Transcriptomics.....	33
2.2.4.1	RNA isolation of <i>E. coli</i> .....	33
2.2.4.2	RNA isolation of <i>N. punctiforme</i> .....	33
2.2.4.3	RNA Quality assessment after Isolation (gDNA check, MAN system, DNase digestion) .....	34
2.2.4.4	cDNA synthesis for real-time qPCR.....	35
2.2.4.5	Semi-quantitative gene expression analysis using qPCR .....	36
2.2.4.6	Estimation of qPCR primer efficiencies.....	37
2.2.4.7	Calculation of fold-difference using $\Delta\Delta C_t$ method (Comparative $C_t$ Method) .....	37
2.2.4.8	<i>N. punctiforme</i> transcriptome analysis.....	38
2.2.5	Proteomics .....	39
2.2.5.1	Aqueous protein extraction of <i>E. coli</i> expression cultures.....	39
2.2.5.2	Aqueous protein extraction of <i>N. punctiforme</i> .....	39
2.2.5.3	SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE) .....	39
2.2.5.4	Coomassie staining .....	40
2.2.5.5	High Performance Liquid Chromatography (HPLC) sample preparation .....	40
2.2.5.6	HPLC qualitative analysis and preparative fractionation of complex protein extracts .....	40
2.2.5.7	MALDI-TOF sample preparation.....	41
2.2.5.8	MALDI-TOF mass spectrometry of HPLC peak fractions.....	41



2.2.6	Study specific methods .....	41
2.2.6.1	Confocal fluorescence microscopy of <i>N. punctiforme</i> / reporter strains using an LSM 710 system .....	41
2.2.6.2	Generation of concentrated <i>B. pusilla</i> exudate for chemical interaction studies ..	42
2.2.6.3	Chemical interaction of <i>N. punctiforme</i> wild-type or reporter strains with <i>B. pusilla</i> using concentrated host plant exudate .....	42
2.2.6.4	Co-cultivation of <i>N. punctiforme</i> wild-type or reporter strains with <i>B. pusilla</i> .....	42
2.2.6.5	Heterologous expression of PKS5 expression construct in <i>E. coli</i> .....	43
<b>3</b>	<b>Results .....</b>	<b>44</b>
3.1	Generation of a <i>N. punctiforme</i> secondary metabolite gene cluster CFP reporter strain library .....	44
3.1.1	Identification of potential natural product biosynthesis gene clusters and respective promoter regions .....	44
3.1.2	Construction of the cyanobacterial shuttle plasmids and <i>N. punctiforme</i> reporter strain mutagenesis .....	48
3.1.3	Transcriptome analysis of a growing <i>N. punctiforme</i> culture .....	52
3.1.4	Evaluation of reporter strain CFP signal patterns in comparison to the RNA deep-seq data .....	57
3.1.5	Cell type specific expression of CFP in <i>N. punctiforme</i> reporter mutant strains ..	64
3.2	Influencing <i>N. punctiforme</i> 's secondary metabolite production .....	66
3.2.1	High density cultivation of <i>N. punctiforme</i> .....	66
3.2.2	Cyanobiont – host plant interaction studies of <i>N. punctiforme</i> and <i>Blasia pusilla</i>	74
3.2.2.1	<i>N. punctiforme</i> – <i>B. pusilla</i> chemical interaction .....	74
3.2.2.2	<i>N. punctiforme</i> – <i>B. pusilla</i> physical interaction .....	77
3.2.2.3	<i>N. punctiforme</i> CFP reporter mutant strains in chemical and physical interaction with <i>B. pusilla</i> .....	80
3.2.2.4	HPLC and mass spectrometric analysis of <i>N. punctiforme</i> wild-type cultures in physical contact with <i>B. pusilla</i> .....	82
3.3	Mutagenesis of <i>N. punctiforme</i> 's NRPS-PKS hybrid BGC PKS4 .....	87
3.4	Heterologous expression of <i>N. punctiforme</i> 's PKS5 BGC in <i>E. coli</i> .....	90
<b>4</b>	<b>Discussion .....</b>	<b>93</b>
4.1	<i>N. punctiforme</i> PCC73102: a suitable target for cyanobacterial natural product research .....	93
4.2	Genomic mining is a potent tool for natural product research .....	94
4.3	The majority of <i>N. punctiforme</i> secondary metabolite BGCs is not silent under standard laboratory conditions .....	96
4.4	Alternative genomic mining strategies .....	100
4.5	Symbiotic interactions trigger major genetic reprogramming in <i>N. punctiforme</i> ..	102
4.6	Inaccurate predictions and analytical limits are the main bottlenecks of natural product discovery .....	104

<b>5</b>	<b>References</b> .....	<b>106</b>
<b>6</b>	<b>Appendix</b> .....	<b>113</b>
6.1	List of Figures.....	113
6.2	List of Tables.....	118
6.3	Table of DNA Oligos.....	118
6.4	Transcriptomics Diagrams including Gene Annotations.....	121
	<b>Publications and Conference Contributions</b> .....	<b>IX</b>
	<b>Acknowledgements</b> .....	<b>XI</b>
	<b>Eigenständigkeitserklärung</b> .....	<b>XII</b>

# 1 Introduction

## 1.1 Secondary Metabolites and Chemical Diversity of Natural Products

Small molecule natural products (SMNPs) are produced by a biological source and harbor a broad spectrum of biological activities and pharmaceutical properties (1, 2). Among others there are, pigments, siderophores, toxins and cell differentiation factors but also bioactive compounds with clinical use such as immunosuppressants, cytostatics, anticancer agents and antibiotics (3–6). Natural products can be divided into several groups. Four major categories are terpenoids, polyketides, nonribosomal peptides and ribosomally synthesized and post-translationally modified peptides (RiPPs) (7–9).

### 1.1.1 Nonribosomal Peptides and Polyketides

Polyketides and nonribosomal peptides are two large families of natural products that are produced by large polyketide synthase (PKS) and nonribosomal peptide synthetase (NRPS) enzyme complexes, respectively (7). In a recent survey Wang et al. identified over 3000 NRPS and PKS gene clusters from 911 organisms demonstrating the widespread distribution of those natural product biosynthesis gene clusters (BGCs) in all three domains of life (10).

Although all PKSs synthesize polyketides by sequential condensation of acyl-coenzyme A (CoA) substrates they are currently subdivided into three distinct classes based on the organization and mode of action of their catalytic activities. Type I PKSs are large multi-domain enzymes that have a modular structure. Each module facilitates substrate recognition, activation and elongation of the growing polyketide chain by one unit in a non-iterative way. Type II PKS systems are multienzyme complexes that harbor just a single set of catalytic domains that iteratively extend the polyketide by one unit. Both type I and type II PKS systems depend on acyl carrier proteins (ACP) covalently binding the cofactor phosphopantetheine that accepts activated substrates via formation of thioester bonds (11). However, in contrast to type II systems type III PKS enzymes directly interact with acyl-CoA substrates in an iterative yet ACP-independent manner (10, 12). A schematic representation a PKS assembly line is depicted in Figure 1.

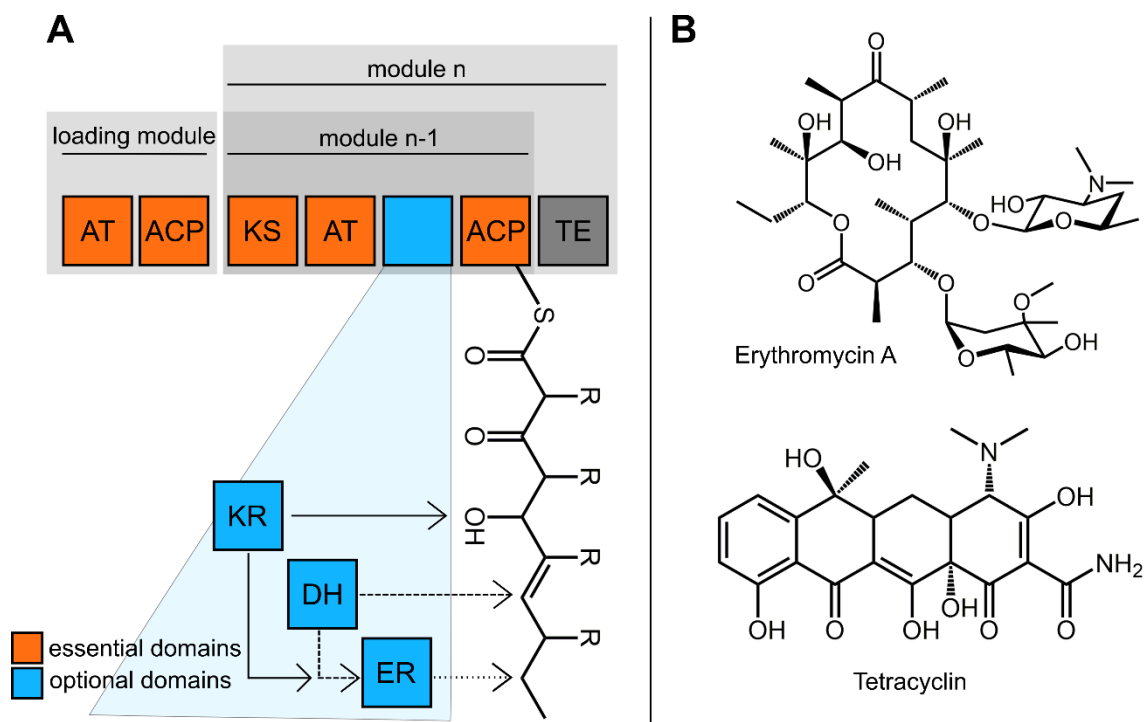


Figure 1 **General organization of a PKS assembly line.** A: Schematic display of a PKS assembly line. AT: Acyltransferase; ACP: Acyl carrier protein; KS: Ketosynthase; KR: Keto-reductase; DH: Dehydratase; ER: Enoyl reductase; TE: Thioesterase. Adopted from Kehr, et al., 2011. *Beilstein J Org Chem.* 7:1622–1635. B: Examples of chemical structures of PKS derived compounds (13).

Just like type I PKS systems NRPSs are large multi-domain enzyme complexes that are organized in modules (14). The smallest functional unit consists of an adenylation domain, peptidyl carrier domain and condensation domain but instead of acyl-CoA substrates NRPS enzymes recognize and activate various proteinogenic and nonproteinogenic amino acid substrates (4, 15). An adenyated substrate is transferred to a phosphopantetheinylated peptidyl carrier protein (PCP) and subsequently a new peptide bond is formed by the catalytic activity of the condensation domain (4). A schematic overview of the enzymes participating in an NRPS assembly line is depicted in Figure 2.

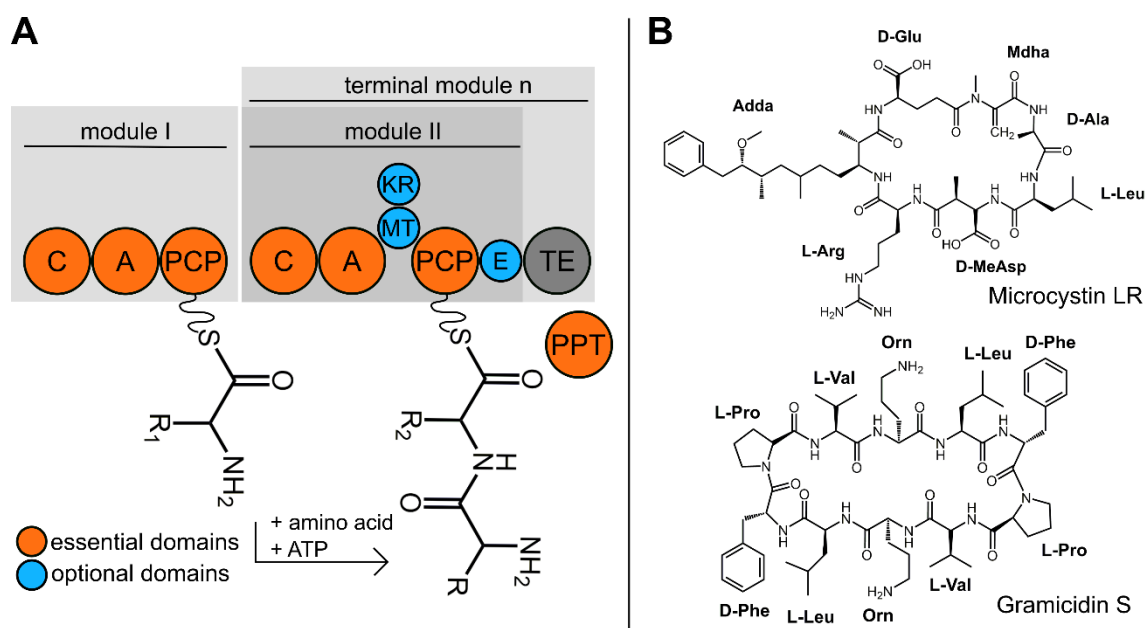


Figure 2 **General organization of a NRPS assembly line.** A: Schematic display of a NRPS assembly line. C: Condensation domain; A: Adenylation domain; PCP: Peptidyl carrier protein; KR: Ketoreductase; MT: Methyltransferase; E: Epimerase; TE: Thioesterase; PPT: Phosphopantetheinyl Transferase. Adopted from Kehr, et al., 2011. *Beilstein J Org Chem.* 7:1622–1635. B: Examples of chemical structures of NRPS derived compounds (16, 17).

Both NRPS and type I PKS systems are commonly referred to as biosynthetic assembly lines because of their highly ordered synthesis process and the fact that intermediate products remain covalently tethered to the enzyme complex at all times (14). Although NRPSs and type I PKSs share remarkable structural and catalytical similarities the NRPS systems are not subdivided into different classes. Despite the classical description of NRPS systems there are examples of stand-alone PCP proteins (18) as well as adenylation and condensation proteins (19). Moreover, in the above mentioned broad scoped survey it was demonstrated that an unanticipated large number (34,4 %) of the identified gene clusters in fact harbor both PKS and NRPS core domains and thus have to be classified as hybrid PKS-NRPS clusters (10).

### 1.1.2 Ribosomally Synthesized and Post-translationally Modified Peptides

Ribosomally derived natural products are a more recently described group of secondary metabolites that were discovered due to the broad genome sequencing efforts in the first decade of the 21<sup>st</sup> century (9). Gene clusters of this chemically diverse group of small yet heavily modified peptides occur all over the three domains of life and although their synthesis is bound to the restraints of the ribosomal peptide synthesis machinery the intensive post-translational processing results in a functional group density and overall stability that easily surpasses natural ribosomal peptides (20, 21). Among others some

important subfamilies of those secondary metabolites are lantibiotics, bacteriocins, cyanobactins, microcins, lasso peptides and microviridins. Despite the already described numerous classes of ribosomal natural products the biosynthesis of nearly all types of RiPPs share common features (9). Most ribosomal natural products are initially produced as larger precursor peptides that are modified after translation. These precursors consist of a C-terminal core peptide that later forms the mature product and a N-terminal leader peptide that has a variety of proposed roles including tailoring enzyme recruitment, secretion signaling and protection from degradation and that gets cleaved off upon core peptide maturation (21). Sometimes also C-terminal attachments to the core peptide occur that are referred to as recognition sequences and are responsible for cyclization as well as N-terminal extensions to the leader region that are referred to as signal sequences and often occur in higher organism to direct the precursor peptide to specific compartments of the cell (9, 22, 23). Another common feature in bacteria producing RiPPs is that genes for precursor peptides, tailoring enzymes, secretion and resistance often cluster together in the genomes facilitating the detection of RiPP clusters in sequencing data (20). A schematic overview of a RiPP gene cluster organization including the precursor peptide sections is depicted in Figure 3.

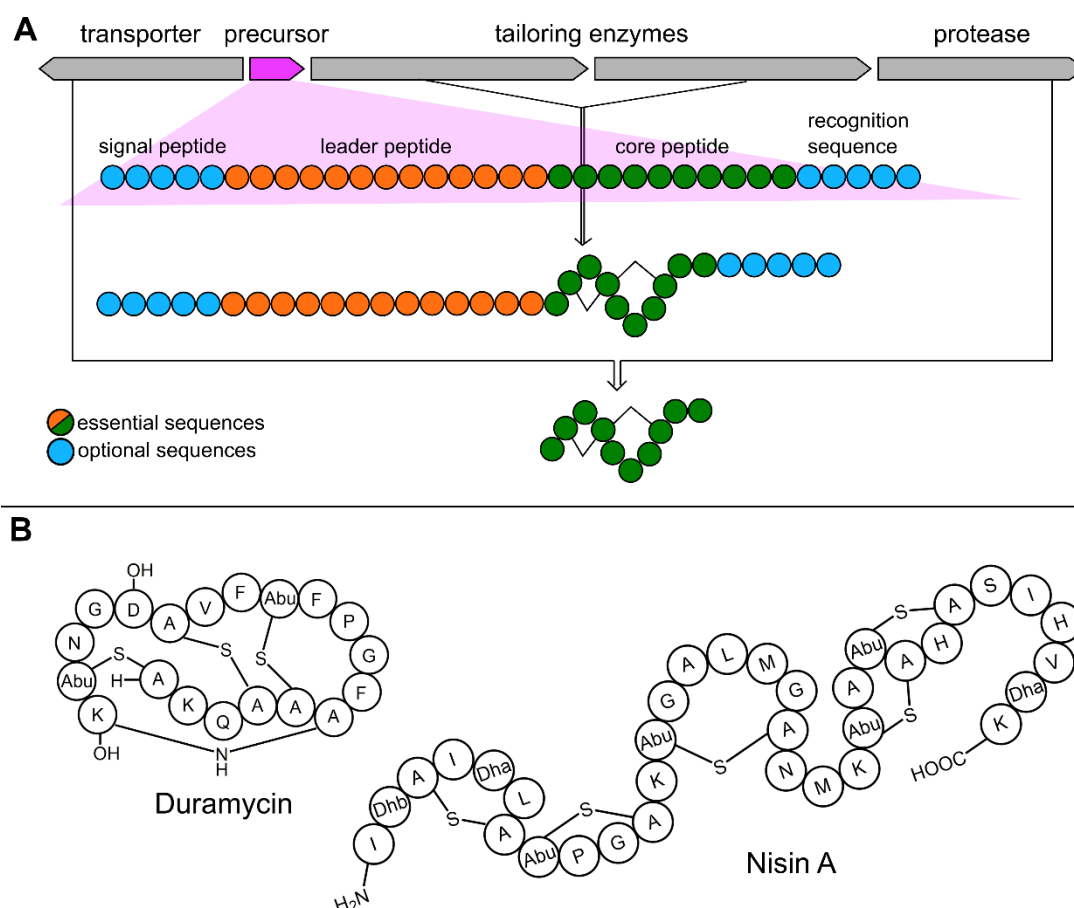


Figure 3 **Schematic representation of a RiPP gene cluster and precursor peptide features. A: Process of RiPP maturation.** Adopted from Kehr, et al., 2011. *Beilstein J Org Chem.* 7:1622–1635. B: Simplified presentation of two mature RiPP products (9).

### 1.1.3 Chemical Diversity of Natural Products

Although the above described classes of secondary metabolites are all built up by rather simple building blocks like acyl-CoAs (polyketides) or proteinogenic (RiPPs and NRPS) and nonproteinogenic (NRPS) amino acids, these substances show a remarkable degree of chemical complexity often paired with high functional group densities and restricted conformations that facilitate high target specificity (6, 14).

When taking secondary metabolites of ribosomal origin into account it is striking that although this category of natural products encompasses a large variety of compound classes, each having a defining structural motif, the number of enzymes involved in the maturation process is relatively low (9). On the one hand certain post-translational modifications of different RiPP classes are carried out by common enzymes, like the Ser/Thr dehydratases involved in the biosynthesis of lanthipeptides, proteusins, thiopeptides and linearazole-containing peptides, driving the susceptibility to natural evolution (24). On the other hand, other modifications like macrocyclization, disulfide

bridges or N- and C-terminal alterations seem to have evolved convergently several times among different RiPP classes. A third aspect promoting the chemical diversity of RiPPs is the permissive nature of many post-translationally modifying enzymes in terms of core peptide sequence mutations (25). The leader peptide focused biosynthesis ultimately opens up the pathways to natural evolution (9).

In contrast to RiPPs the chemical diversity of the modular type I PKS and NRPS biosynthesis products is based on other highly effective principles. In addition to the minimal set of modules necessary for chain elongation there are often auxiliary modules incorporated into the different steps of a given assembly line. These additional catalytic domains in terms of PKS systems can be among others ketoreductases, enoyl reductases, dehydratases, methyltransferases as well as repetitions and combinations of those (14, 26). Analogously many NRPS machineries contain domains for heterocyclization, formylation, oxidation and methylation (26). Furthermore, there can be epimerization domains to produce nonproteinogenic D-amino acid substrates for incorporation into the growing peptide chain or specific condensation domains that catalyze intramolecular cyclization (14, 27). Besides different ways of resolving the thioester bond during chain release of NRPS and PKS systems, products are also often modified after they are detached from the enzyme complex. Such post-assembly modifications can be glycosylations, acylations and oxidations (27, 28). On top of the diversity that is introduced via the action of specific tailoring enzymes or modules a sizable fraction of nonribosomal peptide or polyketide gene clusters actually harbor both NRPS and PKS genes and are thus considered NRPS-PKS-hybrid clusters, furthermore increasing product versatility (29).

## 1.2 Tools for Natural Product Discovery

Since the discovery of the first antibiotics there is an ever-growing demand on new drug lead structures that are mainly derived from newly discovered secondary metabolites. The golden age of antibiotics provided us with effective weapons against the main reason for human morbidity and mortality, microbial infections (30). Despite the increased of human overall lifespan this also gave rise to unanticipated new life-threatening problems that humanity now desperately seeks to overcome (31). These issues include the evolution of natural multi-drug resistances, making it harder and harder to treat certain infections, but also the probability to develop varying kinds of cancer increased alongside with the increased lifespan (32, 33). However, the pace and mode of natural product discovery also changed over the past decades, switching from more classical activity



driven approaches towards more effective computational algorithm-based prediction approaches (34).

### **1.2.1 Classical Screening Approaches**

In a recent review Katz and Baltz divided the history of natural product discovery into three overlapping timespans beginning with the golden era of antibiotics (1940s-70s), over to an intermediate time driven by new insights (1970s-2000s) and finally leading into the modern period mainly fueled by inexpensive high-throughput next generation sequencing techniques (2000s – today) (35).

With the discovery of penicillin but also actinomycin and streptomycin, the early years of natural product discovery were mainly characterized by intense screening efforts of bacterial and fungal extracts for activities against pathogenic bacteria. The technique used is called phenotypic screening or bioactive-guided screening and typically involves the acquisition of microorganisms, fermentation, compound isolation and testing (35, 36). For identification of promising candidates, no prior knowledge of the mode of action is necessary since crude cell extracts are tested for their inhibitory activity against a target cell line. After active extracts were identified intense efforts were spent to isolate and purify the active compounds. Although this method is laborious, between 1940 and 1970 many natural products with pharmaceutical activities were identified that way and ultimately approved as drugs (35).

Two major issues of the initial period of NP discovery were that spatial restriction limited the number of microorganisms that could be cultivated and thus screened at the same time and that the manual screening approach lead to a high rediscovery rate of already known compounds. The second period (1970s-2000s) was thus mainly driven by new screening strategies and improvements in NMR and MS based structure determination to decrease rediscovery rates. However, the basic principle of phenotypic screening remained unchanged. On the other hand, new developments in synthetic biology and advancements in related technologies led to a better understanding of the mode of action of many NPs. Based on this knowledge target-based screening approaches could be developed that allow for the detection of specific activities against, for example, certain cancer types or specific pathogens (37).

Driven by both the new insights gathered during the first two periods as well as the technological advancements in terms of DNA sequencing lead to the most recent time

period of natural product discovery (2000s and beyond) (35). Today most efforts spent are either based on bioinformatics or synthetic biology approaches or a combination of both.

### **1.2.2 Biotechnological Approaches**

Based on the discoveries during the golden era of antibiotics new technologies utilizing recombinant DNA in *Escherichia coli* were developed (35). The identification of different antibiotics and the connection to their corresponding biosynthesis genes was the main driving force for the development of these techniques in *E. coli* as all vectors used for cloning depend on the use of genes that convey resistances to the afore mentioned substances (35, 38). With increasing amount of genome sequencing data, it became obvious that such natural product biosynthesis genes are often organized in clusters several tens of kilobasepairs in size and that new cloning methods are needed for precise targeting, high efficiency and high capacity transfer of target DNA sequences. Such sophisticated cloning techniques quickly came up and are often based on homologous recombination either *in vitro* like Gibson assembly (39), *in vivo* like transformation-associated recombination (TAR) (40) or a combination of both in case of exonuclease *in vitro* assembly combined with RecET (ExoCET) (41). These orthogonal approaches are of special interest for natural product research since 90% of the bacteria are unculturable under standard laboratory conditions and the transfer of their interesting NP BGCs to a suitable lab strain host makes the clusters available for further investigation and also alteration (42). However, the emergence of increasing amounts of sequencing data of varying sources also gave rise to a whole new approach of putative natural product BGC identification nowadays referred to as genomic mining.

### **1.2.3 Genomic Mining**

The main driving forces of all computational natural product detection approaches were the technical improvements in DNA sequencing both in terms of costs and throughput. When the first full genome sequences of actinomycetes were available it became clear that the majority of natural product gene clusters encoded in the genomes are in fact silent and only a minor fraction is actively produced. This insight led to the basic idea of genome mining which is the prediction and isolation of natural products based on genetic information with no molecular structure known (43).

The most classical way of genome mining is the search for genes encoding core biosynthetic enzymes of secondary metabolites. This approach exploits the fact that

despite the structural and mechanistical diversity of many natural products the core biosynthetic enzymes are in fact highly conserved. This is also true for NRPS, PKS and RiPP based natural products. Although this classical homology based search can in principle be conducted manually by the utilization of powerful comparison tools like BLAST (44) or HMMer (45) there were bioinformatic suits developed that allow for automated cluster mining. Prominent examples are BAGEL CLUSEAN and antiSMASH and although all these tools combine carefully selected search profiles based on sets of highly conserved biosynthetic enzymes with manually curated search rules and are under constant development, antiSMASH is the most comprehensive tool available (43, 46).

Despite these classical genome mining approaches being a great way of identifying new or cryptic BGCs in both culturable and unculturable organisms, this method is highly reliant on already described BGCs with known enzymes and strict rules for cluster organization. This leads to the problem that especially new compound classes can often not be associated with their respective BGC. To tackle that issue different types of information were combined in so called comparative genome mining approaches connecting metabolomics data, like MS/MS fragment databases and molecular networking, with genome mining information in order to identify so far undescribed natural products (43). An example of such an approach is also included in the antiSMASH suit which is called the ClusterFinder algorithm (47). This probabilistic model is trained on sets of Pfam domains of already described BGCs and sets of Pfam domains of random genomics regions to enable the prediction of clusters within a given query sequence (47). Other strategies to detect new BGCs utilize phylogenetic information (48, 49), search for the presence of self-resistance mechanisms (50) or conserved regulatory elements (51).

Although the improvements in the field of sequencing had a strong impact on almost every biological discipline, especially natural product research took advantage of the quickly expanding genome data that became available. Applying the above described tools there is no shortage of yet undescribed secondary metabolite gene clusters and rediscovery rates of already described compounds could be diminished with ease streamlining the discovery process of new metabolites.

## **1.3 Cyanobacteria are a Rich Source of High Value Compounds**

### **1.3.1 Photoautotrophic bacteria**

The gram-negative prokaryotes, also known as blue-green algae, are an ancient group of slow-growing bacteria capable of conducting oxygenic photosynthesis with chlorophyll  $\alpha$  as the main photoactive pigment and fossil records dating back up to 3.5 billion years (52, 53). As primary producers they play key roles in global carbon and nitrogen cycles (29). They occur in various different morphotypes including unicellular, multicellular, filamentous, surface attached, colony- and mat forming species and are adapted to various ecological habitats ranging from terrestrial biotopes like deserts and forests to aquatic biotopes like marine, freshwater or brackish water systems (54). These morphologies together with phylogenetic traits were used to classify cyanobacteria into five sections, (I) the unicellular *Chroococcales* dividing by binary fission, (II) the unicellular *Pleurocapsales* producing baeocytes via multiple fission, (III) the *Oscillatoriales*, forming filaments with only vegetative cells dividing in one plane, (IV) the *Nostocales* forming vegetative filaments that divide in one plane but also differentiate into nitrogen fixing heterocysts, motile hormogonia and spore-like akinetes and (V) the *Stigonematales* forming branching filaments dividing in several planes that can also differentiate into heterocysts, hormogonia and akinetes (55).

### **1.3.2 Cyanobacterial natural products**

Cyanobacteria not only have diversified into various morphotypes but they also developed a tremendous arsenal of bioactive natural products including some very potent toxins like the hepatotoxic microcystins or the neurotoxic anatoxins. These potent toxins together with the ability to form large blooms rendered some cyanobacteria a severe health threat to both animals and humans worldwide and made them target of numerous studies (56). Because of the focus on toxins there is readily a lot of knowledge gathered about cyanobacterial secondary metabolites and thus not only harmful natural products were identified but also compounds with beneficial bioactivities including antibiotic, antifungal, anti-inflammatory and anti-cancer activities (57).

A comprehensive comparative genome mining study conducted by Shih et al. 2013 connected the diversity of morphotypes with the secondary metabolite production capacity. This study revealed that especially species with complex morphologies, e.g. filamentous or heterocystous types, frequently encode for large numbers of NRPS/PKS and RiPP BGCs (29). The toxicity issue of bloom forming cyanobacteria caused a heavy

focus of studies on marine and freshwater species. This overrepresentation lead to a bias in the availability of cyanobacterial genetic information. However, in an overarching study about cyanobacterial natural products Demay et al. 2019 demonstrated that a substantial fraction of natural product compound families are found among genera living in neither marine nor freshwater environments with the most capable genus in this section being terrestrial *Nostoc* (56). Furthermore, they pointed out that, taking all habitats together, the family Nostocaceae is the second most potent secondary metabolite producer merely surpassed by the Oscillatoriaceae (56).

### 1.3.3 The Genus *Nostoc* and the Model Organism *N. punctiforme* PCC73102

Species of the genus *Nostoc* belong to the order of *Nostocales* and are filamentous cyanobacteria of the section IV.

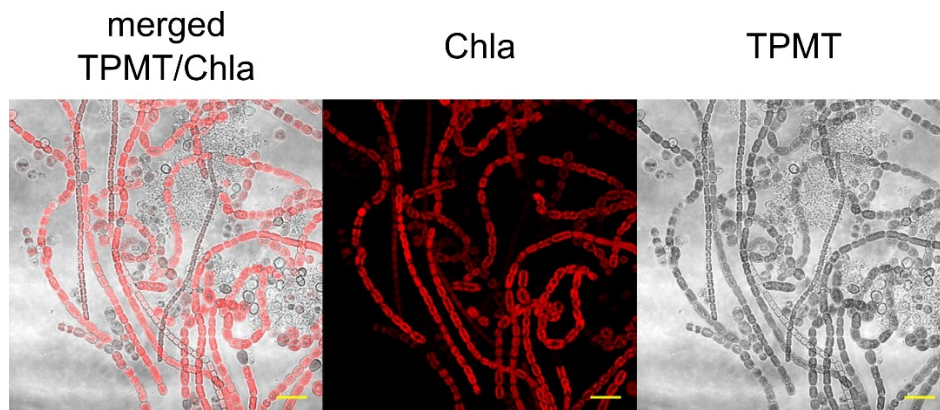


Figure 4 Fluorescence-microscopic picture of *N. punctiforme* PCC73102.

The model organism of this genus is *N. punctiforme* Pasteur Culture Collection (PCC) 73102 or American Type Culture Collection (ATCC) 29133 is a terrestrial species colonizing a wide range of ecological niches. As member of section IV cyanobacteria *N. punctiforme* PCC73102 (from here on *N. punctiforme*) has a complex life cycle and is capable of differentiating multiple distinct cell types in response to environmental signals (58) (Figure 4).

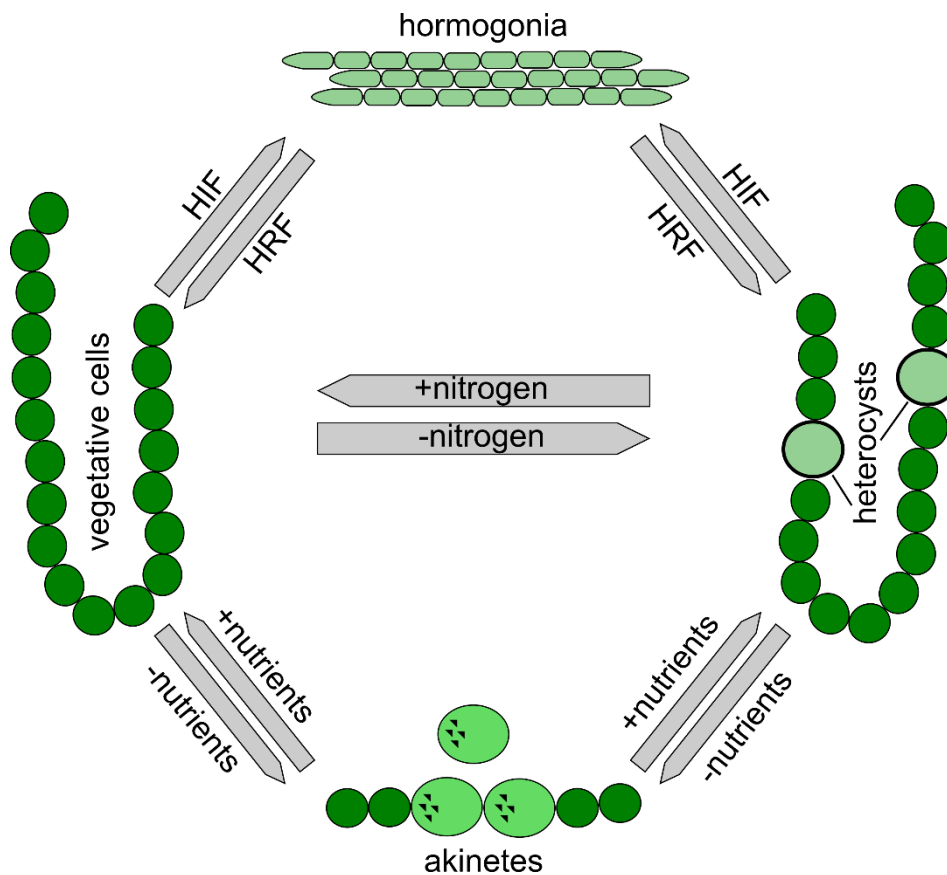


Figure 5 **Schematic representation of the various cell types *N. punctiforme* can differentiate including differentiation triggering conditions.** HIF: Hormogonia inducing factors; HRF: Hormogonia reducing factors. Adapted from Liaimer, et al., 2015. PNAS. 112:1862–1867.

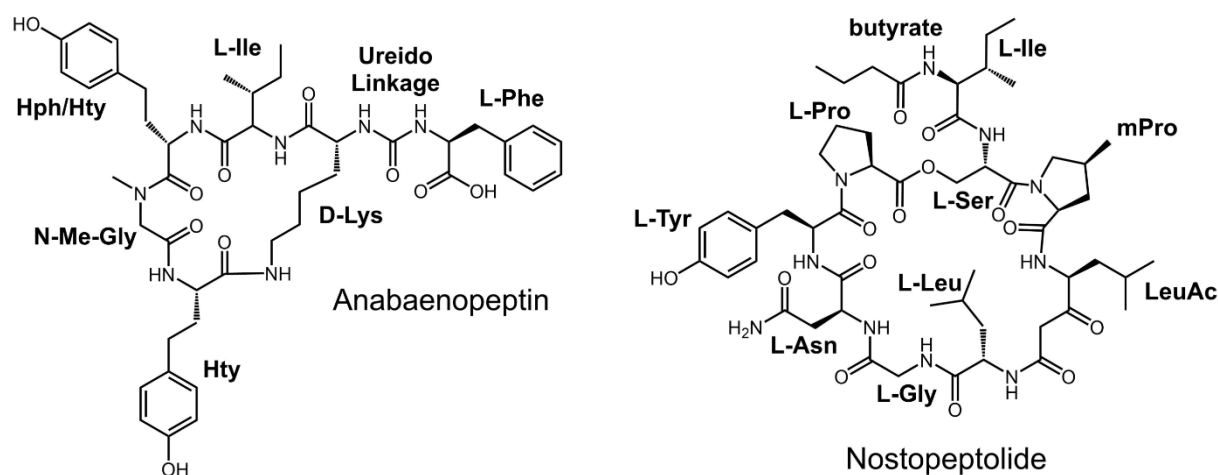
One of these signals is the absence of combined nitrogen that causes certain vegetative cells within a filament to undergo the terminal differentiation into nitrogen fixing heterocysts. The oxygen sensitive reduction of  $N_2$  to  $NH_4^+$  catalyzed by a nitrogenase seems incompatible with oxygenic photosynthesis, however heterocysts create a microoxic environment and are photosynthetically inactive to enable the fixation of atmospheric nitrogen. In contrast to the differentiation into heterocysts the development of vegetative cells into spore-like akinetes is transient and takes place when nutrition levels drop (59). These cells are more resistant to cold and desiccation and can re-differentiate into vegetative cells when the environmental conditions become more favorable (60). Another transient developmental alternative for vegetative cells are hormogonia. These filaments are shorter and consist of smaller cells compared to vegetative cells, they lack heterocysts but instead they are motile and have a short-distance dispersal function (58). While akinete differentiation only occurs among heterocyst-forming cyanobacteria, hormogonia are found in both heterocyst-forming and non-forming species (59). A schematic representation of *N. punctiforme*'s various cell types is depicted in Figure 5.

*N. punctiforme* PCC73102 has a total genome size of 9.06 Mb split between a bacterial chromosome (8.23 Mb) and five plasmids with an average GC content of 41.34 % (Table 1). The complete genome was submitted to the US DOE Joint Genome Institute in 2008 and released together with the corresponding publication by Ekman and co-workers in 2013 (61).

**Table 1 Genomic elements of *N. punctiforme* PCC73102**

Type	Name	Size [Mb]	RefSeq	INSDC
Chromosome	-	8.23	NC_010628.1	CP001037.1
Plasmid	pNPUN01	0.35	NC_010631.1	CP001038.1
Plasmid	pNPUN02	0.25	NC_010632.1	CP001039.1
Plasmid	pNPUN03	0.12	NC_010630.1	CP001040.1
Plasmid	pNPUN04	0.07	NC_010633.1	CP001041.1
Plasmid	pNPUN05	0.03	NC_010629.1	CP001042.

A large fraction of *N. punctiforme* genome is devoted to secondary metabolite production of non-ribosomal peptides, polyketides and post-translationally modified peptides (62). However, until now only two secondary metabolites could be associated with already described cyanobacterial natural products and assigned to their respective BGC, the non-ribosomal peptide nostopeptolide produced by the gene products Npun\_F2181–2188 (63) and the non-ribosomal peptide anabaenopeptin produced by the gene products Npun\_F2459–2465 (64) (Figure 6). Unfortunately, the remaining secondary metabolite BGCs remain orphan, i.e. they could not yet be assigned to a specific product.



**Figure 6 Chemical structures of already described natural products produced by *N. punctiforme*.**

#### **1.3.4 Facultative Symbiotic Lifestyle of *N. punctiforme* PCC73102**

Cyanobacteria are known for undergoing symbiotic interactions with a variety of different micro- and macro-eukaryotes (54). Besides its complex life cycle and huge genome, *N. punctiforme* is also capable of transitioning to a facultative symbiotic lifestyle with a variety of vascular and non-vascular plants and also fungi (65) (Figure 7, A1-D1). Strikingly, the intensity of the interaction of the symbiotic *Nostoc* (cyanobiont) with its host is also varying from host to host. Ranging from rather loose extracellular, epiphytic interactions inside or outside of specialized cavities the symbiotic interaction can also be intracellular in symbiotic tissue or compartments inside the host (65) (Figure 7, A2-D2). Certain environmental factors can induce the establishment of a cyanobiont-plant symbiosis with one of the strongest being a nitrogen limited host plant. Inside the host *N. punctiforme* differentiates heterocysts more frequently and supplements the host with combined nitrogen (65). The symbiosis is mutualistic since not only the plant host is provided with reduced nitrogen, but also the cyanobiont is provided with nutrients and also physical protection while inside the host (66).



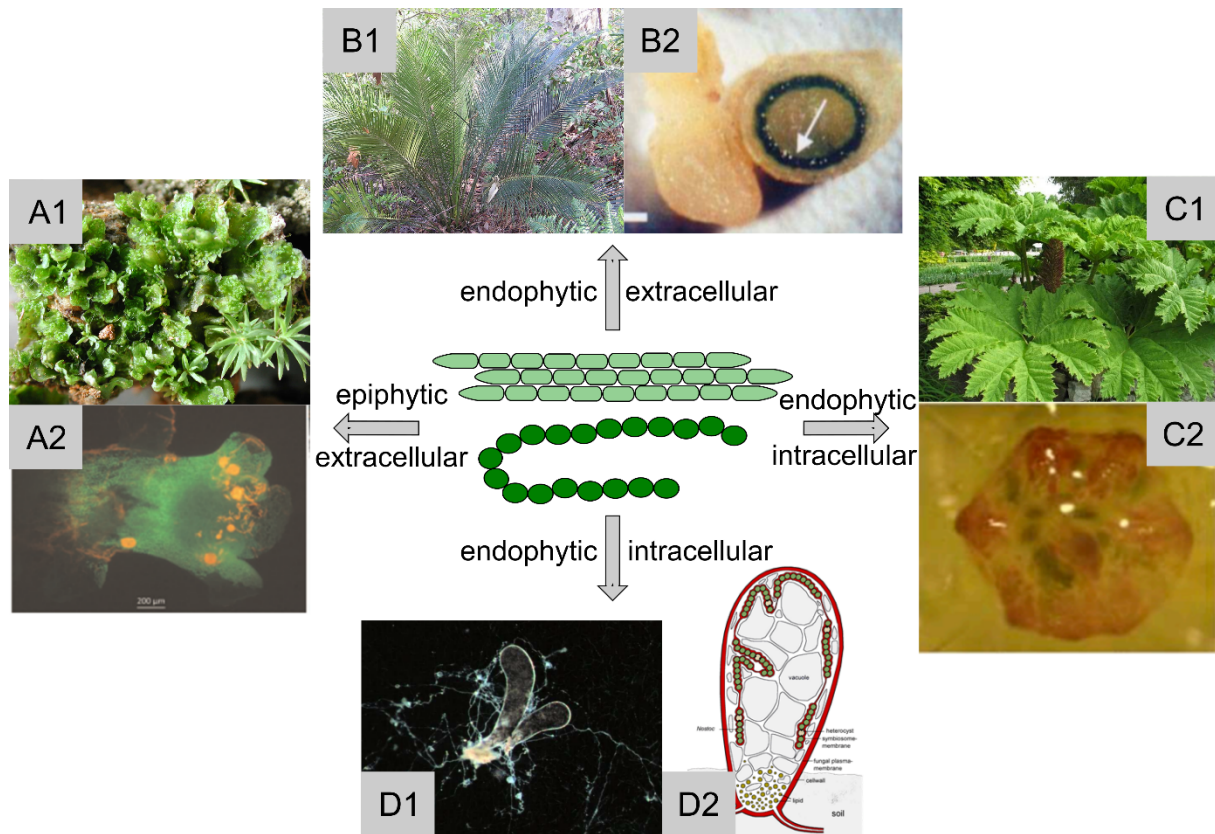


Figure 7 **Overview of various potential symbiotic hosts of *N. punctiforme* including the different modes of symbiotic interaction established with the respective host.** A1: *Blasia pusilla* (source: photo courtesy of Anton Liaimer), A2: Fluorescence micrograph of *Blasia pusilla* colonized by *N. punctiforme* (source: Warshan et al. 2018, Mol. Biol. Evol. 35(5):1160–1175 ); B1: *Macrozamia communis* (source: Photo courtesy of Missouri Botanical Garden), B2: Coralloid root section with highlighted symbiotic tissue (source: Rai et al., 2000, New Phytol. (2000), 147, 449–481); C1: *Gunnera manicata* (source: Photo courtesy of Missouri Botanical Garden), C2: *Gunnera manicata* stem gland colonized by *N. punctiforme* (source: Warshan et al. 2018, Mol. Biol. Evol. 35(5):1160–1175); D1: Micrograph of *Geosiphon pyriformis* showing mycelial structures and bladder bodies (source: Wolf, E. , 2002, Dissertation), D2: Schematic of a bladder body with intracellular cyanobacterial filaments (source: Courtesy of Schüßler, A, [http://www.geosiphon.de/geosiphon\\_home.html](http://www.geosiphon.de/geosiphon_home.html), opened 01.09.2018).

Terrestrial plant symbioses of *N. punctiforme* include Bryophyta like *Blasia pusilla*, a liverwort developing specialized cavities that are inhabited by the cyanobiont, Pteridophyta, like the waterfern *Azolla filiculoides* that is epiphytically colonized by the cyanobiont, Gymnosperms, like the cycad *Cycas revoluta* that develops specialized symbiotic compartments inside its coralloid roots and Angiosperms, like *Gunnera manicata* where *Nostoc* invades its stem glands (67, 68). All these species are ancient lineages pointing towards a long history of co-evolution with *N. punctiforme* (69). Besides these plant interactions there is one example of *Nostoc*-fungi symbiosis known with the fungi *Geosiphon pyriformis* that upon contact with the cyanobiont develops enlarged structures called bladders that encapsulate cyanobacterium (70). Except the *Nostoc*-*Azolla* interaction all the symbiotic interactions are facultative and both host and cyanobiont can exist separately (65).

The mere observation of an overproduction of heterocysts during symbiosis with a nitrogen starved host plant point towards a major genetic reprogramming that place when *N. punctiforme* is undergoing a symbiotic interaction. Furthermore it was already shown that the presence of a potential host can have an impact on the production of nostopeptolide, one of the described non-ribosomal natural products of *N. punctiforme* (71).

#### **1.4 Aims of this Study**

Irresponsible employment of antibiotics and an overall increasing life span connected with increased cancer risks and other diseases lead to an ever-increasing demand on new bioactive compounds exhibiting new modes of action. One prolific source of such compounds are natural products and advances in sequencing techniques and bioinformatics revealed that especially cyanobacteria show a high density of secondary metabolite BGCs.

Thus, the first goal of this study was the development of *N. punctiforme* PCC73102 as a cyanobacterial model strain that can be employed to realize new strategies and approaches for the genomic mining of new natural products in cyanobacteria. Utilizing the large amount of natural product BGCs in its genome its amenability towards various genetic modifications should be assessed.

Based on the findings of Liaimer et al. (71) , the second part of the study dealt with the question to which extent various biotic interactions might have an impact on *N. punctiforme*'s secondary metabolite production and whether these effects can be utilized to specifically induce previously silent natural product BGCs and thus lead to the detection of new compounds.

Assessing *N. punctiforme* as a model system for strategic natural product research and expanding the insight into potential impact factors of cyanobacterial secondary metabolism were the main driving force of this study.

## 2 Material and Methods

### 2.1 Material

#### 2.1.1 Organisms

Table 2 List of organisms used in the current study.

Organism	Source
<i>Nostoc punctiforme</i> PCC73102	Pasteur Culture Collection, France
<i>Blasia pusilla</i>	Provided by Dr. Anton Liaimer, Department of Arctic and Marine Biology, University of Tromsø, Norway
<i>Escherichia coli</i> XL1-Blue chemically competent cells for cloning procedures	Agilent, USA
<i>Escherichia coli</i> GB05RedTrfA strain for RecET direct cloning method (72).	Provided by Dr. Jun Fu, Genomics, Biotechnology Center, Technical University of Dresden, Germany
<i>Escherichia coli</i> GB05-MtaA having a pPant transferase MtaA from myxobacterium <i>Stigmatella aurantiaca</i> DW4/3-1 randomly transposed into the chromosome	Provided by Dr. Jun Fu, Genomics, Biotechnology Center, Technical University of Dresden, Germany
<i>Escherichia coli</i> BL21 (DE3) chemically competent cells for protein expression	Merck, Germany
<i>Escherichia coli</i> J53_RP4 conjugation helper strain (73),	Provided by Prof. Dr. Jeffrey Elhai's lab, Center for the Study of Biological Complexity, Virginia Commonwealth University, USA (74)

#### 2.1.2 Genome sequences, DNA oligos and Primers

The genome sequence for *N. punctiforme* PCC73102 was submitted to the US DOE Joint Genome Institute (JGI) in 2008 and published by Ekman and co-workers in 2013 (61). The accession numbers of the main bacterial chromosome and the five plasmids are listed in

Table 3.

**Table 3 Accession numbers of all genetic elements of *N. punctiforme* PCC73102 used for this study.**

Type	Name	Size [Mb]	RefSeq	INSDC
Chromosome	-	8.23	NC_010628.1	CP001037.1
Plasmid	pNPUN01	0.35	NC_010631.1	CP001038.1
Plasmid	pNPUN02	0.25	NC_010632.1	CP001039.1
Plasmid	pNPUN03	0.12	NC_010630.1	CP001040.1
Plasmid	pNPUN04	0.07	NC_010633.1	CP001041.1
Plasmid	pNPUN05	0.03	NC_010629.1	CP001042.

The annotation version of the main bacterial chromosome of *N. punctiforme* PCC73102 used for this study was updated May 12<sup>th</sup> 2016 and had a total detected gene count of 7 442 and a coding gene count of 6 973.

Noteworthy, the initial gene annotation conducted by Ekman and co-workers, following a Npun\_R/Fnnn pattern, was later replaced with a new annotation, following an NPUN\_RSnnnnn pattern, by the JGI upon release of the sequence. However, for this study the original Npun\_R/Fnnn annotation was used if possible.

All DNA primers were manually deduced for optimized PCR yields. A full overview of all DNA oligos and primers used in this study is listed in appendix I Table 7.

### 2.1.3 Chemicals

Acetonitrile	VWR International (Germany)
Agarose	Biozym (Germany)
Ampicillin	Carl Roth (Germany)
L-Arabinose	Merck (Germany)
Bacto Agar	BD
Chloramphenicol	Carl Roth (Germany)
Chloroform	Carl Roth (Germany)
Diethyl pyrocarbonate (DEPC)	Carl Roth (Germany)
Dimethyl sulfoxide (DMSO)	Carl Roth (Germany)
Ethylenediamine tetraacetic acid disodium salt dihydrate (EDTA)	Carl Roth (Germany)
Ethanol	VWR International (Germany)
Ethidium bromide	Sigma-Aldrich (Germany)
Hydrochloric acid 37%	VWR International (Germany)
Isopropanol	VWR International (Germany)
Kanamycin	Carl Roth (Germany)
Lithium chloride	Carl Roth (Germany)
Methanol HPLC gradient grade	VWR International (Germany)

Phenol, TRIS-HCl equilibrated pH 8.0	Carl Roth (Germany)
Phenol:Chloroform:Isoamyl alcohol 25:24:1, TRIS-HCl equilibrated pH8.0	Carl Roth (Germany)
L-Rhamnose	Carl Roth (Germany)
Sodium nitrate	Carl Roth (Germany)
Streptomycin	SERVA Electrophoresis (Germany)
Tetracyclin hydrochlorid	Carl Roth (Germany)
Trifluoroacetic acid	Sigma-Aldrich (Germany)
TRIS	Carl Roth (Germany)
TRIzol reagent	Thermo Scientific (USA)

All chemicals not listed in detail were purchased at Carl Roth (Germany).

## 2.1.4 Media and solutions

### Media

BG11 <sub>0</sub> medium	Nitrogen-free variant of BG11 as described by Rippka <i>et al.</i> 1979 (75)
BG11 <sub>0</sub> agar	BG11 <sub>0</sub> media supplemented with 0.7% Bacto Agar (BD, Switzerland)
<i>Blasia pusilla</i> regeneration medium	5-times diluted BG11 supplemented with 6.25 mM NaNO <sub>3</sub>
<i>Blasia pusilla</i> N-starvation medium	5-times diluted BG11 <sub>0</sub>
LB medium	Carl Roth (Germany)
LB agar	Carl Roth (Germany)
SOC medium	3.6 g/l glucose, 0.18 g/l KCl, 2.0 g/l MgCl <sub>2</sub> ·6H <sub>2</sub> O, 0.5 g/l NaCl, 20 g/l tryptone, 5.0 g/l yeast extract, pH 7.0

### Buffers and solutions

6xDNA loading dye	10 mM TRIS-HCl pH 7.6, 0.03% bromphenol blue, 0.03% xylene cyanol FF, 60% glycerine, 60 mM EDTA
1x TAE buffer	40 mM TRIS, 0.12% (v/v) acetic acid, 1 mM EDTA pH 8.0
RNA fixation solution	Ethanol + 5% (v/v) TRIS-HCl equilibrated phenol pH 8.0
Homogenization buffer for RNA extraction	100 mM TRIS pH 8.5, 5 mM EDTA pH 8.0, 100 mM NaCl, 0.5% SDS
TE buffer for genomic DNA isolation	10 mM TRIS, 1 mM EDTA, pH 8.0
CTAB solution for genomic DNA isolation	10% (w/v) CTAB + 0.7 M NaCl
5x SDS-PAGE loading dye	0.1% bromophenol blue, 10% (v/v) glycerine, 5% (v/v) 2-mercaptoethanol, 62.5 mM Tris-HCl pH 6.8, 2% (w/v) SDS
SDS-PAGE running buffer	14.42 g glycine, 1 g SDS, 3.03 g tris, ad to 1 L ddH <sub>2</sub> O
Coomassie blue staining solution	5% Al <sub>2</sub> (SO <sub>4</sub> ) <sub>3</sub> ·16H <sub>2</sub> O, 10% ethanol 96%,

	0.02% (w/v) Coomassie brilliant blue, 2% <i>o</i> -phosphoric acid 85%
Coomassie blue destaining solution	10% (v/v) ethanol 96%, 2% <i>o</i> -phosphoric acid 85%
HCCA matrix stock solution	10 mg HCCA in 1 mL 84% acetonitrile/13% ethanol/ 3% TFA 0.1%
HCCA matrix working solution	35 $\mu$ L stock solution + 65 $\mu$ L 84% acetonitrile/13% ethanol/ 3% TFA 0.1%
Carbonate-buffer (5 mbar CO <sub>2</sub> partial pressure)	1:1 (v/v) 3M HKCO <sub>3</sub> , 3M K <sub>2</sub> CO <sub>3</sub>
Carbonate-buffer (32 mbar CO <sub>2</sub> partial pressure)	4:1 (v/v) 3M HKCO <sub>3</sub> , 3M K <sub>2</sub> CO <sub>3</sub>
Carbonate-buffer (90 mbar CO <sub>2</sub> partial pressure)	9:1 (v/v) 3M HKCO <sub>3</sub> , 3M K <sub>2</sub> CO <sub>3</sub>

### 2.1.5 Commercial Kits

GeneJET Gel Extraction Kit	Thermo Scientific (USA)
GeneJET Plasmid Miniprep Kit	Thermo Scientific (USA)
GeneElute Plant Genomic DNA Miniprep Kit	Sigma-Aldrich (Germany)
RNeasy Mini Kit	Qiagen (Germany)
Blue S'Green qPCR Kit	Biozym (Germany)

### 2.1.6 Commercial Enzymes and Auxiliary Products

#### Enzymes

S7 Fusion Polymerase	Mobidiag (Finland)
Maxima Hot Start Green 2x PCR Master Mix	Thermo Scientific (USA)
Maxima Reverse Transcriptase	Thermo Scientific (USA)
NEBuilder HiFi DNA Assembly Master Mix	NEB (USA)
FastDigest Restriction Enzymes	Thermo Scientific (USA)
Proteinase K	Boehringer (Germany)
Lysozyme (chicken egg white)	Sigma-Aldrich (Germany)
RNase-Free DNase Set	Qiagen (Germany)

#### Auxiliary Products

GeneRuler 1 kb DNA Ladder	Thermo Scientific (USA)
GeneRuler Low Range DNA Ladder	Thermo Scientific (USA)
PageRuler Plus Prestained Protein Ladder	Thermo Scientific (USA)
dNTP Set 10mM	Thermo Scientific (USA)
RiboLock RNase Inhibitor	Thermo Scientific (USA)
Random Hexamer Primer (100 $\mu$ M ; 0.2 $\mu$ g/ $\mu$ l)	Thermo Scientific (USA)

### 2.1.7 Cloning Vectors

pRL1049	Cyanobacterial shuttlevector, provided by Prof. Dr. John C. Meeks lab, University of California, USA.
pRL271	Cyanobacterial gene deletion/insertion vector, provided by Prof. Dr. John C. Meeks lab, University of California, USA.
pBR322-amp-tetR-tetO-hyg-ccdB	RecET cloning vector (72), provided by Dr. Jun Fu, Genomics, Biotechnology Center, Technical University of Dresden, Germany
p15A-cm-tetR-tetO-hyg-ccdB	RecET cloning vector (72), provided by Dr. Jun Fu, Genomics, Biotechnology Center, Technical University of Dresden, Germany
pECFP-C1	Source of eCFP-CDS, Takara Bio Europe (France)

### 2.1.8 Membranes, Filters, Cartridges and Cuvettes

Acrodisc 4 mm Syringe Filter	Pall (Germany)
FP 30/0.2 CA-S Syringe Filter	GE Healthcare (USA)
FP 30/0.45 CA-S Syringe Filter	GE Healthcare (USA)
Sep-Pak Plus C18 cartridge	Waters (Germany)
Rotilabo-inserts 100, borosilic. gl.	Carl Roth (Germany)
Septa silicone/PTFE-coated	Carl Roth (Germany)
Electroporation cuvette (1 mm)	PeqLab (Germany)
Electroporation cuvette (2 mm)	PeqLab (Germany)
HATF Transfer Membranes, 85mm, 0.45µm pore size	Merck (Germany)

### 2.1.9 Technical Devices

#### Specialized Equipment

Thermal cycler	Mastercycler Nexus Gradient	Eppendorf (Germany)
qPCR thermal cycler	Lightcycler 96	Roche Diagnostics (Germany)
DNA/RNA quantification	NanoDrop 2000 Spectrophotometer	Thermo Scientific (USA)
UV-VIS Spectrophotometer	UV-VIS Spectrophotometer UV-1800	Shimadzu Europe (Germany)
Fluorescence microscope	EVOS FL Imaging System	Thermo Scientific (USA)
Fluorescence microscope	LSM 700 Confocal Laser Scanning Microscope with ZEN Blue Software	ZEISS (Germany)
Cyanobacterial cell cultivator	High-Density Cell Cultivator System	CellDeg (Germany)
Gel documentation system	Molecular Imager ChemiDoc XRS+ with ImageLab Software	Bio-Rad (Germany)
Electroporator	MicroPulser	Bio-Rad (Germany)
Sonicator	Ultrasonic homogenizer Sonoplus HD	Bandelin (Germany)
HPLC system	Auto injector SIL-20AC	Shimadzu Europe (Germany)
	Column oven CTO-10AS	

	Communications bus module CBM-20A	
	Degasser DGU-20A5	
	Diode array detector SPD-M20A	
	Fraction collector FRC 10A	
	Liquid chromatograph LC-10A	
	SymmetryShield RP18 column (3.5 µm, 4.6 mm × 100 mm)	
	SymmetryShield Sentry Guard column (3.5 µm, 3.9 mm × 20 mm)	
MALDI-TOF mass spectrometer	Microflex LRF with FlexAnalysis Software and mMass mass spectrometry tool	Bruker Corporation (USA)

## General Equipment

Centrifuge 1.5 – 2 ml	Perfect Spin 24	PeqLab (Germany)
Centrifuge 1.5 – 2 ml	Micro Star 17R	VWR International (Germany)
Centrifuge 15 – 50 ml	Sigma 6-16K with Sigma 12169-H rotor	Sigma (Germany)
Centrifuge 500 ml	Sigma 6-16KS with Sigma 12500- H rotor	Sigma (Germany)
MTP centrifuge	Centrifuge 5403 with 16M 2-MT rotor	Eppendorf (Germany)
Thermal shaker	Thriller	PeqLab (Germany)
Vacuum concentrator	Vacuum concentrator Concentrator plus	Eppendorf (Germany)
Vacuum concentrator	RVC 2-25 CDplus CT 02-50 Cold Trap	Christ (Germany)
Ultrapure water system	Milli-Q Reference A+	Merck (Germany)
Electrophoresis power supply unit	PowerPac Basic power supply	Bio-Rad (Germany)
Agarose gel electrophoresis system	PerfectBlue gel system	PeqLab (Germany)
PAGE system	Mini-PROTEAN Tetra Vertical Electrophoresis Cell and Hand Cast System	Bio-Rad (Germany)
Vortex mixer	Vortex-Genie 2 2 Shaker	Scientific Industries (USA)



## 2.2 Methods

### 2.2.1 Culture Maintenance

#### 2.2.1.1 Cultivation of *Nostoc punctiforme* PCC73102

*Nostoc punctiforme* PCC73102 (*N. punctiforme*) was cultured under various diazotrophic conditions. For regular liquid cultivation under standard laboratory conditions the cells were cultivated in BG11<sub>0</sub> medium (BG11 without nitrogen as described by Rippka *et al.* 1979) with a continuous white light irradiation of 30  $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$  at 21 °C. High light condition differs from regular liquid condition as the cultures were exposed to 107  $\mu\text{E}\cdot\text{m}^{-2}\cdot\text{s}^{-1}$ . For batch liquid cultivation under high CO<sub>2</sub> condition a method described by Pörs *et al.* 2010 was adapted (76). Therefore, cultures under standard conditions were supplemented with low-density polyethylene membrane bags filled with a 3M carbonate buffer yielding a 32 mbar CO<sub>2</sub> partial pressure.

Cultivation of *N. punctiforme* CFP reporter mutant strains liquid cultures was identical to standard conditions except that the BG11<sub>0</sub> medium was supplemented with 2  $\mu\text{g}/\text{ml}$  streptomycin (BG11<sub>0</sub>-Strep2) to keep up the selection pressure and ensure plasmid propagation within the reporter culture.

For solid phase cultivation *N. punctiforme* was spread on HATF transfer membranes and kept on BG11<sub>0</sub> medium supplemented with 0.7% Bacto-agar in petri dishes on otherwise unchanged standard low light conditions. Reporter strains were kept on BG11<sub>0</sub>-Strep2 supplemented with 0.7% Bacto-agar.

#### 2.2.1.2 High Density cultivation of *N. punctiforme*

In contrast to the conventional cultivation techniques for *N. punctiforme* described above the organism and its reporter mutant strains were also cultivated to high cell densities in a specialized apparatus provided by CellDeg (Germany). The general method is described by Guljamow *et al.* 2016 and depending on the HD culture vessel used a volume of either 100 ml or 10 ml cells was cultivated (77).

#### 2.2.1.3 Cryopreservation of *N. punctiforme*

For long time storage and preservation of *N. punctiforme* wild-type and mutant strains cryo-cultures were generated and stored at -80 °C. Therefore, a respective conventional culture was grown to high density for 20 d under standard low light conditions. After a healthy and dense culture was achieved 13 ml culture were harvested by centrifugation at 3000 x g for 5 min. The culture supernatant was discarded and the cell pellet was

resuspended in 3 ml BG11<sub>0</sub> supplemented with 5% DMSO. Each 1 ml of the cell resuspension was transferred to a 1.5 ml reaction vessel and stored at -80 °C until further use. In order to revitalize a frozen culture, the cells are thawed on ice, immediately diluted in BG11<sub>0</sub> medium and then kept under standard low light conditions until cell growth is observed.

#### 2.2.1.4 Cultivation of *Escherichia coli*

There were various *Escherichia coli* (*E. coli*) strains cultivated in this study. The cells were cultivated either in liquid LB medium or on LB-agar (1.5% agar) in petri dishes. For isolation of plasmid DNA small liquid cultures of 4 ml LB medium were inoculated and incubated overnight (16h) at 220 rpm and 37 °C. Corresponding to the selection marker used on the vector to be isolated the LB medium was supplemented with either 20 µg/ml streptomycin, 100 µg/ml ampicillin, 5 µg/ml tetracycline or 12.5 µg/ml chloramphenicol.

#### 2.2.1.5 Cryopreservation of *E. coli*

For longtime storage and preservation of the newly assembled plasmids in *E. coli* cryo-cultures were produced and stored at -80 °C. Therefore, an overnight culture was produced and 750 µl of the overnight culture were mixed with 750 µl LB medium supplemented with 50% glycerin in a 2.0 ml reaction vessel and mixed thoroughly. Afterwards the prepared cryo-culture was stored at -80 °C until further use.

#### 2.2.1.6 Cultivation of *Blasia pusilla* under alternating nitrogen availability

*Blasia pusilla* (*B. pusilla*) is maintained in 250 ml Erlenmeyer flasks filled with 100 ml regeneration medium (1/5 BG11 supplemented with 6.25 mM NaNO<sub>3</sub>). To mimic nitrogen starvation conditions during cultivation, a prerequisite for the establishment of a cyanobacteria-hostplant-symbiosis, the moss cultures were alternatively cultivated for 5 weeks in N-starvation medium (1/5 BG11<sub>0</sub>) and regeneration medium.

### **2.2.2 Molecular biology and cloning techniques for *Escherichia coli***

#### 2.2.2.1 Preparative DNA amplification for cloning (Fusion PCR)

For the use in synthetic biology applications like *in vitro* and *in vivo* DNA assembly high yields of precise target DNA fragments are needed. In order to produce the desired amounts a proof-reading, high fidelity DNA polymerase (Fusion S7, Mobidiag) was used for amplifying template sequences of plasmid or genomic origin.

The following reaction setup was used:

Component	Amount
Template DNA	1 ng (plasmid DNA), 20 ng (genomic DNA)
10 $\mu$ M forward primer	1.5 $\mu$ l (300 nM)
10 $\mu$ M reverse primer	1.5 $\mu$ l (300 nM)
10 mM dNTPs	1 $\mu$ l (200 $\mu$ M)
5x HF-Green buffer	10 $\mu$ l
Fusion S7 polymerase (2 U/ $\mu$ l)	0.5 $\mu$ l (0.02 U/ $\mu$ l)
ddH <sub>2</sub> O	Ad 50 $\mu$ l

If complex templates or templates with high GC-content were used the reaction buffer was substituted by 5x GC-Green buffer for optimized yields.

The PCR program used was set up as follows:

Temperature	Time	Cycles
98 °C	2 min	1
98 °C	10 secs	35
58 °C	20 secs	35
72 °C	1 min / 2000 bp target length	35
72 °C	5 min	1

After PCR reaction the reaction mixture was separated and purified on an agarose gel.

#### 2.2.2.2 Qualitative DNA amplification for genotyping (colony PCR)

During a cloning approach in *E. coli* for a first screening for potential positive mutants a technique referred to as colony PCR (cPCR) was used. This method uses whole cells as templates in PCR reactions for a quick screening for presence of potentially correctly assembled plasmids.

The cPCR setup consists of following components:

Component	Amount
Template	1 <i>E. coli</i> colony
10 $\mu$ M forward primer	0.2 $\mu$ l
10 $\mu$ M reverse primer	0.2 $\mu$ l
Maxima Hot Start Green 2x PCR Master Mix	5 $\mu$ l
ddH <sub>2</sub> O	4.6 $\mu$ l

The PCR program used was set up as follows:

Temperature	Time	Cycles
98 °C	5 min	1
98 °C	10 secs	30
58 °C	20 secs	30
72 °C	1 min / 1000 bp target length	30
72 °C	1 min	1

For a qualitative analysis of the PCR results the reaction mixtures were then separated on an agarose gel.

#### 2.2.2.3 Agarose gel electrophoresis for DNA

In order to separate DNA fragments by size gel electrophoresis was performed. Using a PerfectBlue gel systems (PeqLab, Germany) and depending on the expected DNA fragment size range 1-2% (w/v) agarose containing gels, where longer DNA fragments required the lower agarose concentration. The agarose was dissolved in TAE buffer by boiling the suspension in a microwave. Upon casting the hot agarose suspension in the electrophoresis chambers 0.05 µg/ml ethidium bromide were added to allow for DNA band detection under UV light irradiation. If not already included in the reaction buffer DNA samples were supplemented with DNA loading dye to allow for samples to settle inside the gel pockets and to visualize electrophoresis progress during the run. Submerged in TAE buffer the DNA electrophoresis was performed at constant 120 V for 20 - 45 min depending on gel size and agarose content. After the run the gel was analyzed and documented using the ChemiDoc XRS+ system together with the ImageLab software. For DNA fragment preparation small gel pieces containing the desired DNA fragments were sliced out and afterwards extracted from the agarose gel.

#### 2.2.2.4 DNA extraction from agarose gels

Preparative DNA extraction out of agarose gel slices was performed using the GeneJET Gel Extraction Kit (Thermo Scientific, USA) according to manufacturer's instructions with the sole exception that during the last step the elution buffer was substituted with ddH<sub>2</sub>O to reduce the salt concentration in the final DNA solution.

#### 2.2.2.5 Preparative plasmid isolation

To isolate large amounts of plasmid DNA from *E. coli* mutant strain overnight cultures the GeneJET Plasmid Miniprep Kit (Thermo Scientific, USA) was used according to

manufacturer's instructions with the sole exception that during the last step the elution buffer was substituted with ddH<sub>2</sub>O to reduce the salt concentration in the final DNA solution.

#### 2.2.2.6 Digestion of DNA with restriction endonucleases

To prepare vectors for *in vitro* or *in vivo* homologous recombination it was necessary to linearize them and by that expose the recombination sites. This was achieved by digestion of the circular DNA with restriction endonucleases. Using the FastDigest enzymes (Thermo Scientific, USA) according to the manufacturer's instruction a typical digestion reaction was performed in 20 µl and incubated at 37 °C for 2 h. After the reaction the digested DNA was purified via gel electrophoresis followed by gel extraction.

#### 2.2.2.7 Spectrophotometric quantification of nucleic acid concentrations

The concentration of RNA or DNA in aqueous solution was measured by the determination of the UV light absorption at 260 nm using the NanoDrop 2000 photometer (Thermo Scientific, USA) and 1.5 µl of a given sample. In addition to the DNA/RNA concentration, impurities such as phenols and proteins are detected via the simultaneous determination of the UV light absorption at 230 nm and 280 nm.

#### 2.2.2.8 *In vitro* DNA assembly via homologous recombination

For efficient and precise assembly of cloning vectors out of several DNA fragments sharing end homologies a technique first described by Gibson et al. 2009 was utilized (39). This method was commercialized by NEB via a product line named NEBuilder HiFi DNA Assembly Master Mix (HiFi Builder) and is based on three consecutive enzymatic reactions that take place in a single reaction vessel. Two DNA fragments that are to be assembled (typically one being a vector backbone carrying an origin of replication and a resistance marker and the other being a gene of interest or just insert) need to be linear and sharing 25 bp homology on both ends. In the first step of the assembly reaction 3' single stranded overlaps are exposed by the action of a 5' exonuclease. In the second step fragments anneal based on the sequence homology of their respective strand ends and the polymerase in the reaction mixture is filling in any nucleotide gaps. In the last step a ligase is closing the phosphate backbone and by that stabilizing the assembly. A fraction of the reaction mixture containing the newly assembled vector can then be transformed into *E. coli*.

The assembly reaction consists of the following components:

Component	Amount
Insert DNA fragment (linear)	x ng (depending on fragment size)
Backbone DNA fragment (linear)	50 ng
2x NEBuilder HiFi DNA Assembly Master Mix	5 $\mu$ l
ddH <sub>2</sub> O	Ad 10 $\mu$ l

The amount of insert DNA used for assembly is calculated based on the amount of linearized vector backbone used in the reaction:

$$\text{Insert [ng]} = \text{Insert [bp]} / \text{backbone [bp]} * 50 \text{ ng} * \text{excess-factor}$$

The excess-factor ranges from 10 to 2 based on the size difference between the insert and the backbone, where for bigger size differences larger excess-factors are recommended.

5  $\mu$ l of the reaction mixture were then chemically transformed into a suitable *E. coli* cloning strain (typically *E. coli* XL1-Blue).

#### 2.2.2.9 *In vivo* DNA assembly via homologous recombination (FullRecET)

To assemble very large synthetic vectors, typically for heterologous expression approaches of large secondary metabolite gene clusters, a method published by Wang *et al.* 2016 was adapted (72). The cloning technique presented is based on the use of either the  $\lambda$  phage full length RecET recombination system or the lambda phage Red $\alpha\beta$  recombination system. The former facilitates highly efficient homologous recombination of two linear DNA molecules (LLHR) and the latter one enables homologous recombination between a linear and a circular DNA molecule (LCHR). Both systems are combined in a single genetically engineered *E. coli* strain that has the Red $\alpha\beta$  system integrated in its genome and the RecET system provided on a low copy plasmid inducible with l-arabinose and l-rhamnose, respectively. The *in vivo* assembly was strictly performed as described in the step-by-step protocol provided by the authors. In order to clone large clusters directly into one of the provided expression vectors the authors recommend to prepare mildly digested genomic DNA of the target cluster-harboring organism and use it for LLHR using the RecET system. However, this approach requires large amounts of genomic DNA and a suitable restriction enzyme so instead purified PCR products of the target sequence were used for the assembly but the ratio of linear expression vector (entry vector) and target sequence (insert) was kept unchanged.

#### 2.2.2.10 Chemical transformation of *Escherichia coli*

For stable and long-term storage newly designed plasmids were transformed into suitable *E. coli* cloning strains. Utilizing the CaCl<sub>2</sub> chemically induced competence *E. coli* cells of the XL1-Blue genotype (Agilent, USA) were prepared (78). A volume of 50 µl chemically competent *E. coli* XL1-Blue cells were mixed with 5 µl of an NEBuilder assembly reaction mixture and incubated on ice for 15 min. The transformation mixture was then subjected to a heat shock in a water bath at 42°C for 45 sec. Afterwards 450 µl prewarmed (37 °C) SOC medium was added and the cells were incubated for 1 h at 37 °C and 220 rpm. Then the cells were spread on LB agar supplemented with an appropriate antibiotic and incubated at 37 °C overnight. Potential positive colonies were then checked for the presence of the newly introduced plasmid using colony PCR.

### **2.2.3 Molecular biology and cloning techniques for *Nostoc punctiforme* PCC73102**

#### 2.2.3.1 Isolation of *N. punctiforme* genomic DNA

For the isolation of high molecular weight *N. punctiforme* genomic DNA 15 ml of a densely grown culture were harvested in a 15 ml falcon tube. Cells were mildly centrifuged at 3000 x g for 5 min at RT. Afterwards the cell pellet was rinsed with 5 M NaCl twice for the removal of polysaccharides. Then the cell pellet is resuspended in 2.0 ml TE-buffer and a final concentration of 2 mg/ml lysozyme was added followed by a 1 h incubation at 37 °C with periodic inversion during incubation. After that 500 µl 0.5 M EDTA, 1.0 ml 2 mg/ml proteinase K solution and 100 µl 20% SDS solution were added and the mixture gently inverted and again incubated for 1 h at 37 °C with periodic inversion during incubation. Then 1/6 volumes 5 M NaCl were added and mixed. Afterwards 1/8 volumes CTAB solution were added, mixed and incubated for 10 min at 65 °C. Cell debris were separated by centrifugation at 13 000 x g for 5 min at RT. Without disrupting the cell debris disk that may form on the surface the extraction mixture was transferred to a fresh 15 ml falcon tube. The solution was then extracted with 1 volume of chloroform and phases were separated by centrifugation at 13 000 x g for 3 min at RT. Then the aqueous phase was transferred to a new tube and the DNA was precipitated with 2 volumes of 95% ethanol for 30 min at -20 °C. Afterwards the DNA was separated by centrifugation at 13 000 x g for 5 min at RT, the supernatant was discarded and the DNA pellet resuspended in 500 µl TE-buffer and transferred to a 2.0 ml reaction tube. Then the DNA solution was extracted two times with phenol:chloroform:isoamyl alcohol until the interface was clear of

precipitations. The final extraction was performed with 100% chloroform. The aqueous phase was again transferred to a new 2.0 ml reaction tube and the DNA was precipitated with 2 volumes 95% ethanol and 1/10 volume 3 M sodium acetate. The DNA was pelleted by centrifugation at 13 000 x g for 5 min at RT. Then the DNA pellet was rinsed with 1 ml 70% ethanol and again pelleted by centrifugation at 13 000 x g for 5 min at RT. Afterwards the supernatant was discarded and the DNA pellet was dried for 5 – 10 min at 60 °C. The dry pellet of genomic DNA was then resuspended in 100 µl TE-buffer, 1 µl RNase T1 (1000 U/ml, Thermo Scientific, USA) added and the DNA content was determined at the spectro-photometer.

#### 2.2.3.2 Plasmid transfer for the generation of CFP-reporter mutant strains of *N. punctiforme*

In order to incorporate the preassembled promotor-CFP fusion construct carrying plasmids into *N. punctiforme* an electroporation was performed. Therefore 100 ml of a *N. punctiforme* wild-type culture was prepared and grown for four weeks under standard low light conditions. After reaching a high cell density the culture was harvested into 50 ml conical centrifuge tubes (falcon tubes) and centrifuged at 3000 x g for 5 min at room temperature (RT). After discarding the supernatant, the cell pellet was rinsed with 10 ml cold, sterile ddH<sub>2</sub>O. The washing step was repeated four times. After the last centrifugation step each pellet was resuspended in 3.5 ml cold, sterile ddH<sub>2</sub>O. To prepare the electroporation mixture 400 µl of the concentrated cell suspension were mixed with a total of 10 µg of the desired plasmid and put on ice until immediately before transfer into the electroporation cuvette. The mixture was then electroporated at 600 ohms, 1.6 kEV and 25 mF in a 2 mm electroporation cuvette with an expected electroporation pulse duration of 3.5 – 5 ms. After electroporation the cell suspension was spread on HATF membranes placed on BG11<sub>0</sub>-agar petri dishes and the selection process for *N. punctiforme* mutagenesis was started. Typically, each plasmid electroporation was prepared in quadruplicates and an additional negative control without plasmid to enable monitoring of background die-off.

#### 2.2.3.3 Tri-parental conjugation for knock-out/in mutagenesis of *N. punctiforme*

A second method to introduce genetic material in *N. punctiforme* is a technique referred to as tri-parental conjugation. Conjugation based approaches are preferable over electroporation when aiming for gene replacements or integration of genetic material. The method was first described by Wolk et al. 1984 and adapted for *N. punctiforme* by



Cohen et al. 1994 (74, 79). For the conjugation of genetic material three strains are needed, (I) an *E. coli* helper strain carrying the mobilizer and helper plasmid (*E. coli* J53\_RP4), (II) an *E. coli* donor strain carrying the plasmid that will be transferred (*E. coli* XL1-Blue cloning strain) and (III) *N. punctiforme* as the receiving strain.

The protocol used in this study was as follows. For the preparation of *N. punctiforme* for the conjugation procedure 30 ml of an exponentially growing culture were transferred into a 50 ml falcon tube and centrifuged at 1000 x g for 10 min at RT. Then the cells were rinsed with 30 ml BG11<sub>0</sub> medium and centrifuged again. The pellet was again resuspended in 30 ml BG11<sub>0</sub> medium and then the cells were sonicated for 18 sec with 20% amplitude in a 6 s/3 s on/off cycle. Afterwards the culture was centrifuged at 1000 x g for 10 min at RT and then the pellet was once again resuspended in fresh BG11<sub>0</sub> medium and the culture was placed for regeneration under standard low light conditions overnight.

For the preparation of *E. coli* for each donor strain and the helper strain 4 ml LB-medium supplemented with appropriate antibiotics were inoculated and incubated overnight at 37 °C and 220 rpm. At the day of conjugation for each strain prepared 30 ml of prewarmed (37 °C) LB-medium supplemented with appropriate antibiotics were inoculated, set to OD<sub>600 nm</sub> = 0.1 and then the cultures were incubated for approx. 2 h at 37 °C 220 rpm until an OD<sub>600 nm</sub> = 0.8 was reached. Then the cultures were harvested at 3000 x g for 10 min at RT and rinsed with 15 ml LB-medium without antibiotics. To set the cultures to an OD<sub>600 nm</sub> = 10.0 the cell pellets were then resuspended in 2.4 ml LB-medium without antibiotics. Then a pre-mating mixture was prepared by mixing of 250 µl *E. coli* J53\_RP4 helper strain with 250 µl of one of the *E. coli* XL-1Blue donor strains. This cell mixture was then incubated for 1 h in darkness at 30 °C without shaking.

For the final *E. coli-N. punctiforme* conjugation the regenerated *N. punctiforme* culture was centrifuged at 3000 x g for 10 min at RT and the pellet was subsequently resuspended in 5 ml BG11<sub>0</sub>-medium. Conjugation agar plates (BG11<sub>0</sub>-medium supplemented with 0.7% Bacto-agar and 5% LB-medium) were prepared and covered with HATF membranes. Then 500 µl pre-mating *E. coli* mixture and 500 µl concentrated *N. punctiforme* suspension were mixed in a 1.5 ml reaction tube and subsequently pelleted at 5000 x g for 30 sec at RT. The supernatant was discarded leaving behind 200 µl to resuspend the *E. coli-N. punctiforme* mixture. The cell mixture was then gently spread on the HATF membrane and the conjugation agar plates were incubated at low light conditions for two

days prior to the membranes being transferred to regular BG11<sub>0</sub>-agar plates. After this transfer the regular selection process for *N. punctiforme* mutagenesis began.

#### 2.2.3.4 Selection process during *N. punctiforme* mutagenesis

After electroporation or conjugation of a previously assembled plasmid the treated *N. punctiforme* culture was spread on a HATF membrane to be easily transferrable between different BG11<sub>0</sub>-agar containing petri dishes. The selection process began with a 5 days incubation of the treated culture on BG11<sub>0</sub>-agar without any antibiotics under standard low light conditions. When growth of the treated culture was observed the HATF membranes were transferred on BG11<sub>0</sub>-agar supplemented with 50% of the regular antibiotic concentration (typically 1 µg/ml streptomycin). Depending on the condition of the *N. punctiforme* wild-type culture used for the mutagenesis approach, the wild-type background of a treated culture died-off within 7 to 14 days with first signs of cell lysis being observed after 2 to 3 days (intense coloring of the HATF membrane due to pigment release). After 14 days the cultures were transferred on BG11<sub>0</sub>-agar plates supplemented with the full antibiotic concentration (2 µg/ml streptomycin) and continuously incubated under standard low light conditions. Usually four to six weeks after initial plasmid transfer potential plasmid carrying *N. punctiforme* mutant colonies arose on the filter membrane. Another two weeks after those colonies were macroscopically visible, they were large enough to be further processed. In order to gently transfer the mutant culture from solid phase growth to liquid growth the colonies were excised and the membrane slice carrying the potential *N. punctiforme* mutants were transferred to an Erlenmeyer flask containing 15 ml BG11<sub>0</sub>-medium supplemented with 2 µg/ml streptomycin. The culture was then incubated under standard low light conditions until growth was observed. For the last step of the selection process a fraction of the liquid potential mutant culture was subjected to DNA isolation to enable the confirmation of the mutant genotype by PCR.

#### 2.2.3.5 DNA isolation of potential *N. punctiforme* mutants for genotyping PCR

To isolate sufficient amounts of plasmid or genomic DNA of sparse potential *N. punctiforme* mutant cultures the GeneElute Plant Genomic DNA Miniprep Kit (Sigma-Aldrich, Germany) was used. The procedure was performed according to the manufacturer's instructions with the exception that the initial cell disruption step was skipped and that the provided elution buffer was substituted for ddH<sub>2</sub>O. Afterwards the DNA concentration was determined at the spectro-photometer and 1 µl of the isolate was

used undiluted as a template for colony PCR to confirm the genotype of the selected *N. punctiforme* mutant culture.

## 2.2.4 Transcriptomics

### 2.2.4.1 RNA isolation of *E. coli*

In order to isolate RNA of *E. coli* liquid cultures a protocol based on the TRIzol reagent (Thermo Scientific, USA) was used. Of a given *E. coli* liquid culture 2 ml were harvested by centrifugation at 12 000 x g for 1 min at RT in a 2.0 ml reaction tube. After discarding the supernatant, the cell pellet was resuspended in 1 ml TRIzol reagent. The suspension was incubated at 65 °C for 20 min at 1400 rpm in a thermo shaker. Then 200 µl chloroform were added and the mixture was thoroughly shaken at 1400 rpm, RT for 2 min. For phase separation the mixture was then centrifuged at 11 000 x g at 4 °C for 10 min. The aqueous phase was then transferred to a fresh 1.5 ml reaction tube and 0.7 volume isopropanol were added and mixed by inversion. To precipitate the RNA the mixture was then incubated at -20 °C for 1 h and afterwards centrifuged for 10 min at 11 000 x g at 4 °C. The RNA pellet was then rinsed with 1 ml 70% ethanol and afterwards the solution was centrifuged at 7 500 x g at 4 °C for 5 min. The supernatant was then carefully discarded using a suitable pipette tip and the pellet was dried by incubation at 55 °C for 15 min. The RNA was then resuspended in 50 µl RNase-free ddH<sub>2</sub>O and incubated at 55°C for 10 min to support the resuspension. At the end of the procedure the RNA content was determined spectroscopically using the NanoDrop2000 (Thermo Scientific, USA). The RNA isolates were then subjected to RNA quality assessment prior to further use.

### 2.2.4.2 RNA isolation of *N. punctiforme*

For the isolation of RNA of *N. punctiforme* liquid cultures a protocol aimed for plant RNA isolation was adapted. In the first step of the protocol a hot phenol mix consisting of 250 µl phenol (pH 8.0), 5 µl β-mercaptoethanol and 500 µl homogenization buffer was prepared and pre-heated at 60 °C in a thermal shaker while the cyanobacterial cell pellet stored at -80 °C was thawed on ice. Then the cell pellet was fully resuspended in the hot phenol mixture and incubated at 60 °C, 1400 rpm for 15 min. Afterwards 250 µl chloroform were added and the suspension was again incubated at 60 °C, 1400 rpm for 15 min. To separate the organic and aqueous phase the suspension was then centrifuged at 16 000 x g, RT for 15 min. Afterwards the aqueous layer was transferred into a fresh 1.5 ml reaction tube and 550 µl phenol:chloroform:isoamyl alcohol were added. The mixture was then incubated at 60 °C, 1400 rpm for 10 min and spun at 14 000 x g, RT for 10 min. The

aqueous layer was again transferred to a fresh 1.5 ml reaction tube and 50 µl 3 M sodium acetate were added and mixed by inversion. Then 400 µl isopropanol were added, mixed by inversion and the mixture was then incubated at -80 °C for 15 min to precipitate the nucleic acids. Afterwards the solution was centrifuged at 14 000 x g, 4°C for 30 min. The supernatant was carefully removed, and the pellet was air dried by incubation of the open reaction tube under the fume hood for 15 min. The pellet was resuspended in 500 µl RNase-free ddH<sub>2</sub>O and 500 µl 4 M lithium chloride were added. The solution was then incubated at -20 °C for 30 min to precipitate the RNA followed by centrifugation at 14 000 x g, 4 °C for 30 min. After discarding the supernatant, the pellet was rinsed with 1 ml 80% ethanol and spun at 14 000 x g, 4°C for 5 min. Then the supernatant was carefully removed by pipetting and the pellet was fully air dried by incubation of the open reaction tube under the fume hood for 15 min. Afterwards the pellet was resuspended in 40 µl RNase-free ddH<sub>2</sub>O and incubated at 37 °C, 850 rpm for 30 min in a thermal shaker to support the resuspension process. At the end of the procedure the RNA content was determined spectroscopically using the NanoDrop2000 (Thermo Scientific, USA). The RNA isolates were then subjected to RNA quality assessment prior to further use.

#### 2.2.4.3 RNA Quality assessment after Isolation (gDNA check, MAN system, DNase digestion)

To assure the quality of extracted RNA at first a control PCR was performed according to the cPCR setup (as described under 2.2.2.2) except that instead of an *E. coli* colony 1 µl of the RNA raw isolate was used as a template. The control PCR was performed to detect the presence of any genomic DNA contaminations in the RNA isolate and thus to determine whether a follow-up DNase digestion using the RNeasy mini Kit (QIAGEN, Germany) was necessary. For this purpose, primers were used that target the housekeeping gene *rnpB* within *N. punctiforme*'s genome and amplify a 110 bp sequence fragment of it. This gene later on also served as the reference gene for semi-quantitative real-time PCR analysis.

The control PCR was set up as follows:

<b>Component</b>	<b>Amount</b>
Template	1 µl RNA isolate / 1µl <i>N. punctiforme</i> gDNA isolate (positive control) / 1 µl ddH <sub>2</sub> O (negative control)
10 µM RT_rnpB_fw	0.2 µl
10 µM RT_rnpB_rv	0.2 µl
Maxima Hot Start Green 2x PCR Master Mix	5 µl
ddH <sub>2</sub> O	3.6 µl

The PCR program used for the control PCR:

Temperature	Time	Cycles
98 °C	2 min	1
98 °C	10 secs	30
58 °C	15 secs	30
72 °C	15 secs	30
72 °C	1 min	1

The result of the control PCR was analyzed via agarose gel electrophoresis and if traces of DNA were observed within the RNA isolate a follow-up on-column DNase digestion using the RNeasy mini Kit (QIAGEN, Germany) was performed. The RNA clean-up was performed as described by the manufacturer's instructions including the bench top on-column DNase-digestion with the exception that for the final elution step of the protocol the elution buffer was substituted for RNase-free ddH<sub>2</sub>O.

After the RNA clean-up the above described control PCR was repeated with 1 µl of the processed RNA sample and again the results were analyzed by agarose gel electrophoresis. If no DNA contamination was observed the final concentration of the RNA solution was determined spectroscopically using the NanoDrop 2000 (Thermo Scientific, USA). The DNA-free RNA samples were then transcribed into cDNA to either perform regular PCR analysis or semi-quantitative real-time PCR analysis.

#### 2.2.4.4 cDNA synthesis for real-time qPCR

For the preparation of RNA samples for PCR or qPCR analysis it was necessary to retranscribe the RNA into cDNA. This was done using the Maxima Reverse Transcriptase (Thermo Scientific, USA). The reaction mixture was set-up as follows:

Component	Amount
DNA-free RNA isolate	1.5 µg RNA sample
Random Hexamer Primer (100 µM; 0.2 µg/µl)	1 µl (100 pmol)
dNTP Mix (10 mM)	1 µl (0.5 mM)
RNase free ddH <sub>2</sub> O	Ad to 14.5 µl
5x RT buffer	4 µl
RiboLock RNase Inhibitor	0.5 µl
Maxima Reverse Transcriptase	1 µl

The reaction mixture was then incubated as described at 25 °C for 10 min followed by 50 °C for 50 min and finally 85 °C for 5 min in a PCR cycler. Expecting a full cDNA conversion rate, the final concentration of cDNA in the reaction mixture was 75 ng/µl.

To confirm the quality of the cDNA again a control PCR was performed comparable to the RNA control PCR (described under 2.2.4.3) but instead of RNA 1 µl of the cDNA reaction mixture was used as a template. The positive PCR as analyzed by agarose gel electrophoresis confirmed the quality of the transcribed cDNA. The cDNA was ready to use for semi-quantitative real-time PCR analysis.

#### 2.2.4.5 Semi-quantitative gene expression analysis using qPCR

To determine relative transcription level of selected genes Sybr-green chemistry based semi-quantitative real-time PCR utilizing a Lightcycler 96 (Roche Diagnostics, Germany) qPCR cyler and the Blue S'Green qPCR Kit (Biozym, Germany) was performed. The qPCR mixture was prepared based on the manufacturer's recommendations:

<b>Component</b>	<b>Amount</b>
cDNA template (0.25 ng/µl)	4 µl (1ng)
10 µM forward primer	0.5 µl (500 nM)
10 µM reverse primer	0.5 µl (500 nM)
2x Blue S'Green qPCR Mastermix	5 µl

The 96 well plate was first loaded with cDNA solution, afterwards the enzyme mastermix was premixed with the corresponding primer solutions and then added to the respective cDNA fraction on the multi-well plate. After all components were added the plate was sealed with an adhesive optical film (BZO Seal Film, Biozym, Germany) to reduce sample evaporation and contamination. According to the manufacturer the dye-containing enzyme mix was added in the dark and light exposure of the final multi-well plate was minimized prior to running the qPCR analysis. Each sample was prepared as three technical replicates to reduce uncertainties induced via pipetting errors, unequal evaporation or other factors.

The qPCR program used was composed of the following steps:

	<b>Temperature</b>	<b>Duration</b>	<b>Cycle</b>
Initial denaturation	95 °C	2 min	1
Amplification	95 °C	5 sec	45
	60 °C	20 sec	45
Melting analysis	95 °C	10 sec	1
	65 °C	60 sec	1
	97 °C	1 sec	1
	37 °C	5 sec	1

The results of the semi-quantitative real-time PCR were analyzed using the provided LightCycler 96 SW 1.1 software. The calculated Ct-values for gene expression levels were then further used to calculate the log-fold changes using the  $\Delta\Delta C_t$  method.

#### 2.2.4.6 Estimation of qPCR primer efficiencies

A crucial prerequisite for the correct calculation of log-fold changes using the  $\Delta\Delta C_t$  method is that the primers used in the respective qPCR all have the same amplification efficiency of close to 100%. In order to determine the qPCR primer efficiencies a series of cDNA dilutions were run in an otherwise regular semi-quantitative real-time PCR analysis. Each pair of primers aimed to be used for transcriptomic analysis was tested against 0.1 ng, 1 ng, 10 ng and 100 ng of a given cDNA to enable determination of amplification efficiency. The qPCR was performed as described under 0 but instead of 1 ng cDNA per reaction the above-mentioned cDNA amounts were used. The Ct-values were determined using the Lightcycler96 SW1.1 software and plotted against  $\log_{10}([cDNA])$ . Then the slope  $m$  of the described linear regression curve was calculated. In the final step the amplification efficiency was calculated as follows:

$$\text{Primer efficiency [\%]} = (10^{(-1/m)} - 1) \times 100$$

Each sample was run in triplicates to reduce the effects of external error sources. Primers were accepted for use in semi-quantitative real-time PCR when determined efficiencies were close to 100%.

#### 2.2.4.7 Calculation of fold-difference using $\Delta\Delta C_t$ method (Comparative $C_t$ Method)

In order to better visualize the transcriptional differences between a sample and a control group the determined Ct-values were used to calculate fold-differences using the  $\Delta\Delta C_t$  method. In the first step for every target gene within the sample group or the control group the  $\Delta C_t$  values were calculated as follows:

$$\Delta C_{t \text{ gene1}} = C_{t \text{ target gene}} - C_{t \text{ reference gene}}$$

While the standard deviation for each  $\Delta C_t$  value was calculated as follows:

$$s = (s_1^2 + s_2^2)^{1/2}, \text{ where } s = \text{standard deviation and } X^{1/2} \text{ being the square root of } X$$

For this study RnpB served as a reference gene in every semi-quantitative real-time PCR analysis.

In the second step the differences for each genes  $\Delta C_t$  value when comparing control and sample group was calculated as follows:

$$\Delta\Delta C_{t \text{ gene1}} = \Delta C_{t \text{ gene1, sample group}} - \Delta C_{t \text{ gene1, control group}}, \text{ with } s(\Delta\Delta C_{t \text{ gene1}}) = s(\Delta C_{t \text{ gene1}})$$

In the final step the fold-differences, that are typically expressed as a range by incorporation of the standard deviation of the  $\Delta\Delta C_t$  value into the fold-difference calculation, were calculated as follows:

$$\text{Minimal fold-difference}_{\text{gene1}} = 2^{(-\Delta\Delta C_{t \text{ gene1}} + s(\Delta\Delta C_{t \text{ gene1}}))}, \text{ min}_{\text{gene1}}$$

$$\text{Maximal fold-difference}_{\text{gene1}} = 2^{(-\Delta\Delta C_{t \text{ gene1}} - s(\Delta\Delta C_{t \text{ gene1}}))}, \text{ max}_{\text{gene1}}$$

$$\text{fold-difference}_{\text{gene1}} = (\text{max}_{\text{gene1}} + \text{min}_{\text{gene1}}) / 2 \pm (\text{max}_{\text{gene1}} - (\text{max}_{\text{gene1}} + \text{min}_{\text{gene1}}))$$

When in the context of comparing a sample group with a control group positive  $\Delta\Delta C_t$  values were calculated the resulting fold-differences ranged between  $1 > x > 0$ . Fold-differences lower than 1.0 indicate a lower RNA abundance in the sample group compared to the control group and thus a downregulation of the transcription of the investigated gene in the sample group.

#### 2.2.4.8 *N. punctiforme* transcriptome analysis

For a full transcriptome analysis of a growing *N. punctiforme* culture an RNA deep-sequencing approach was conducted. In preparation of the transcriptome analysis a *N. punctiforme* culture was harvested by centrifugation at 3000 x g, RT for 10 min. The cell pellet was resuspended in a smaller volume to yield a concentrated culture. The concentrated culture was subsequently spread evenly on six HATF filter membranes on BG11<sub>0</sub>-Agar containing petri dishes. After growth to medium cell densities under standard low light conditions the HATF filters were transferred to fresh BG11<sub>0</sub>-Agar plates.

The cells were harvested after 0, 1, 3, 5, 7 and 9 days by scrapping them off the HATF filter and transfer into 1 ml TRIzol (Thermo Scientific, USA). Thereby one HATF filter was harvested for every sampling timepoint. The Samples were then incubated at 65 °C at 1400 rpm for 20 min. Afterwards the aqueous phase was recovered by two extractions with chloroform. The resulting RNA containing supernatant was then purified from any contaminating traces of DNA by using the RNeasy clean-up kit including on-column DNase digestion (QIAGEN, Germany) according to manufacturer's instructions.



The RNA samples were then transferred to the DNA sequencing section of the OIST (Okinawa, Japan) where final sample preparation steps, further processing and data analysis were conducted as described within the methods section of Dehm *et al.* 2019 (80).

## **2.2.5 Proteomics**

### **2.2.5.1 Aqueous protein extraction of *E. coli* expression cultures**

For the extraction of proteins from *E. coli* expression cultures the cells were harvested by centrifugation at 10 000 x g, 4 °C for 10 min. The culture supernatant was transferred to a 50 ml falcon tube and subjected to HPLC sample preparation (2.2.5.5). The cell pellet was rinsed with 10 ml 1 M TRIS-HCl pH 7.5 and then centrifuged at 10 000 x g, 4 °C for 10 min. Afterwards the pellet was resuspended in 5 ml ddH<sub>2</sub>O. Then the cell suspension was sonicated for 1.5 min at 90 % amplitude running a 3s/1s on/off cycle. After sonication the cell debris were separated by centrifugation at 13 000 x g, 4 °C for 10 min. In the final step the cell lysis supernatant was transferred to a fresh 15 ml falcon tube and subjected to HPLC sample preparation (2.2.5.5). If desired small aliquots of culture supernatant sample and cell lysate sample were taken to perform SDS-PAGE analysis.

### **2.2.5.2 Aqueous protein extraction of *N. punctiforme***

For the extraction of proteins from *N. punctiforme* liquid cultures the cell suspension was harvested by mild centrifugation at 3000 x g, RT for 10 min. The culture supernatant was transferred to a fresh 50 ml falcon tube and subjected to HPLC sample preparation (2.2.5.5). The cell pellet was resuspended in 10 ml ddH<sub>2</sub>O and subsequently sonicated for 10 min at 65 % amplitude running a 2s/2s on/off cycle. The cell debris were separated by centrifugation at 13 000 x g, 4 °C for 10 min. The cell lysis supernatant was transferred to a fresh 15 ml falcon tube and subjected to HPLC sample preparation (2.2.5.5). If desired small aliquots of culture supernatant sample and cell lysate sample were taken to perform SDS-PAGE analysis.

### **2.2.5.3 SDS Polyacrylamide Gel Electrophoresis (SDS-PAGE)**

Protein extraction samples were electrophoretically separated utilizing the method described by Laemmli in 1970 (81). Separation gels contained 12.5 % acrylamide, while stacking gels contained 4 % acrylamide. During gel preparation ammonium persulfate (APS) and tetramethylethylenediamine (TEMED) were added to final concentrations of 0.5 % and 0.05 %, respectively. To prepare the protein samples for SDS-PAGE they were

mixed with 0.25 volumes of 5xSDS-PAGE loading dye and incubated at 95 °C for 5 min. For the actual electrophoresis the gel was completely submerged in SDS-PAGE running buffer in a Mini-PROTEAN Tetra Vertical Electrophoresis Cell (Bio-Rad, Germany) and run at constant currents of 30 mA per gel. In order to analyze the gels were stained with Coomassie after electrophoresis.

#### 2.2.5.4 Coomassie staining

To visualize proteins separated via SDS-PAGE the gels were incubated in Coomassie blue staining solution for at least 1h on a horizontal shaker. To remove excessive Coomassie the stained gels were then incubated in Coomassie blue destaining solution for 2 h on a horizontal shaker, while the destaining solution was exchanged once after 1 h of incubation. After sufficient destaining the gels were detected using the Molecular Imager ChemiDoc XRS+ together with the ImageLab Software (Bio-Rad, Germany).

#### 2.2.5.5 High Performance Liquid Chromatography (HPLC) sample preparation

To prepare culture supernatant samples or aqueous cell extracts for HPLC analysis sample solutions were selectively enriched based on compound hydrophobicity using Sep-PAK C18 plus cartridges (Waters, Germany). According to manufacturer's instructions the cartridges were prepared for sample binding by rinsing with 2 ml methanol followed by 2 ml ddH<sub>2</sub>O. Afterwards culture supernatant or cell lysis samples were loaded on the silica matrix. Then the column was rinsed with 2 ml 5% methanol and finally eluted into a fresh 2.0 ml reaction tube using 2 ml methanol. To further increase the compound concentration the samples were then fully evaporated at 1 mbar, 30 °C and 1400 rpm using a vacuum concentrator RVC 2-25 CDplus (Christ, Germany). Then the samples were resuspended in 120 µl methanol and incubated for 10 min in an ultrasonic water bath. Afterwards the final methanol concentration of 60% was set by adding 80 µl ddH<sub>2</sub>O to the resuspended samples. To separate any insoluble parts the samples were then centrifuged at 16 000 x g, 4 °C for 20 min and afterwards filtrated using a syringe and Acrodisc 4 mm Syringe Filters (Pall, Germany). Finally, 60 µl of the filtered sample solution were transferred into Rotilabo-inserts 100 (Carl-Roth, Germany) and put into HPLC glass vials closed with a PTFE-coated silicone septum.

#### 2.2.5.6 HPLC qualitative analysis and preparative fractionation of complex protein extracts

Complex protein samples were analyzed by HPLC according to the compounds hydrophobicity using a reverse-phase C18 column and the above-mentioned Shimadzu

Europe HPLC system (2.1.9). Prior to sample injection the C18 column was equilibrated to 2 % acetonitrile. For qualitative analysis after sample injection (5 to 50  $\mu$ l) a linear acetonitrile gradient was applied starting off at 2 % acetonitrile, reaching 60 % acetonitrile at 40 min, ramping up to 100 % acetonitrile at 42 min, keeping 100 % until 45 min and then ramping down to 2 % acetonitrile at 50 min followed by constant 2 % acetonitrile until 55 min mark. In preparation for MALDI-TOF analysis protein samples were also fractionated. Therefore, the same acetonitrile gradient was applied but additionally the automated fraction collector was individually programmed to collect selected peaks as 500  $\mu$ l fractions during HPLC analysis. If a given peak could not be collected within a single 500  $\mu$ l fraction the fractions were combined after the fractionation run.

#### 2.2.5.7 MALDI-TOF sample preparation

To prepare a peak fraction for further analysis via MALDI-TOF mass spectrometry the sample was fully evaporated at 1mbar, 30 °C and 1400 rpm using the vacuum concentrator RVC 2-25 CDplus (Christ, Germany). Afterwards the dried sample was resuspended in 8  $\mu$ l 0.1 % trifluoroacetic acid (TFA) and 2  $\mu$ l acetonitrile.

#### 2.2.5.8 MALDI-TOF mass spectrometry of HPLC peak fractions

Processed sample were analyzed at a Bruker Microflex LRF equipped with a nitrogen laser ( $\lambda = 337$  nm). For mass spectrometry 0.3  $\mu$ l sample solution were spotted on the MALDI target and covered with 0.3  $\mu$ l HCCA matrix solution. Sample spots were air dried for 5 min and the target was then inserted into the mass spectrometer. Using the FlexAnalysis (Bruker, USA) and a detection range of 400 – 4000 Da for every sample approximately 1000 laser shots were averaged into a single spectrum. The gathered data was then analyzed with the open source software mMass.

### **2.2.6 Study specific methods**

#### 2.2.6.1 Confocal fluorescence microscopy of *N. punctiforme* / reporter strains using an LSM 710 system

Confocal fluorescence microscopy was conducted on a Zeiss LSM 780 Axio Observer Z1 equipped with a diode laser (405 nm), as well as a HeNe laser (633 nm) and an AxioCam digital microscope camera. For image acquisition a PlanApo 1.4/63 $\times$  oil immersion objective and filter presets for eCFP and chlorophyll  $\alpha$  detection (excitation at 405/633 nm, detection at 450–550 nm and 650–725 nm, respectively) were used. For device operation, image acquisition, and processing, the ZEN software was used.

#### 2.2.6.2 Generation of concentrated *B. pusilla* exudate for chemical interaction studies

To mimic the presence of a potential host plant for *N. punctiforme* to colonize conditioned *B. pusilla* medium was produced and concentrated for media addition studies. Therefore, *B. pusilla* was cultivated in 100 ml *Blasia pusilla* N-starvation medium (1/5 BG11<sub>0</sub>) under standard low light conditions for 5 weeks. After that the culture supernatant was transferred into two 50 ml falcon tubes and subsequently concentrated 100-fold down to a volume of 500 µl at 1 mbar, 30°C and 1400 rpm for 15 h using a vacuum concentrator RVC 2-25 CDplus (Christ, Germany). The *B. pusilla* cultures were alternatively cultivated in *Blasia pusilla* N-starvation medium and in *Blasia pusilla* regeneration medium (1/5 BG11 supplemented with 6.25 mM NaNO<sub>3</sub>) to induce and resolve N-deprivation.

#### 2.2.6.3 Chemical interaction of *N. punctiforme* wild-type or reporter strains with *B. pusilla* using concentrated host plant exudate

To specifically investigate the influence of compounds released into the culture medium by N-deprived *B. pusilla* chemical interaction assays were conducted. Therefore, concentrated *B. pusilla* exudate was added to a given sample culture of *N. punctiforme* wild-type or reporter strain equal to 50 % total culture volume of unconcentrated exudate, e.g. 7.5 µl 100x concentrated *B. pusilla* exudate added to a 15 ml *N. punctiforme* culture. Sample cultures and control cultures were then incubated for varying timespans under standard low light conditions. Depending on the aim of study after each sampling point for transcriptomic analysis RNA extraction was performed, for HPLC analysis aqueous protein extraction followed and for *N. punctiforme* reporter-strain-based assays confocal fluorescence microscopy was conducted.

#### 2.2.6.4 Co-cultivation of *N. punctiforme* wild-type or reporter strains with *B. pusilla*

In order to investigate the responses of *N. punctiforme* when exposed to a N-deprived potential host co-cultivation assays of *N. punctiforme* wild-type or reporter strains together with *B. pusilla* were conducted. In preparation for that *B. pusilla* was starved for nitrogen by cultivation in 1/5 BG11<sub>0</sub> under standard low light conditions for 5 weeks. A day before a co-cultivation assay the N-deprived moss was sliced into squares of approx. 1 cm<sup>2</sup> (approx. 1 g wet weight) and incubated overnight in fresh 1/5 BG11<sub>0</sub> under standard low light conditions. For the preparation of *N. punctiforme* wild-type or reporter strains for co-cultivation 100 ml BG11<sub>0</sub> supplemented with or without streptomycin were inoculated with 3 ml of a dense culture of the selected wild-type or mutant strain and

subsequently incubated under standard low light conditions for 5 weeks. For the co-cultivation culture 15 ml of the prepared *N. punctiforme* culture and one *B. pusilla* slice were merged in a 25 ml Erlenmeyer flask. These samples were prepared in triplicates together with a no-moss control culture serving as a reference. The co-cultures were subsequently sampled at varying timepoints. Depending on the aim of study after each sampling point for transcriptomic analysis RNA extraction was performed, for HPLC analysis aqueous protein extraction followed and for *N. punctiforme* reporter-strain-based assays confocal fluorescence microscopy was conducted.

#### 2.2.6.5 Heterologous expression of PKS5 expression construct in *E. coli*

For the heterologous expression of the *N. punctiforme* secondary metabolite cluster PKS5 in *E. coli* the corresponding genes were cloned into an inducible expression plasmid using the FullRecET method (as described under 2.2.2.9). In preparation for the heterologous expression the assembled expression plasmid pBR322\_C21\_F1-2-3 was transformed into *E. coli* GB05-MtaA. Then overnight cultures of *E. coli* GB05-MtaA and *E. coli* GB05-MtaA\_pBR322\_C21-F1-2-3 were prepared in 30 ml LB and 30 ml LB supplemented with 100 µg/ml ampicillin, respectively. The overnight cultures were incubated for 16 h at 37 °C and 230 rpm. On the following day expression cultures as well as a no-plasmid-control (*E. coli* GB05-MtaA) and a no-induction-control (*E. coli* GB05-MtaA\_pBR322-C21-F1-2-3) were inoculated in the respective LB medium from the overnight cultures and set to an OD<sub>600 nm</sub> of 0.1. Then the cultures were incubated at 30 °C, 230 rpm for approx. 3 h until they reached OD<sub>600 nm</sub> of 0.4. Then the no-plasmid-control and the expression cultures were induced by adding a final concentration of 0.5 µg/ml tetracycline followed by incubation at RT, 230 rpm for 4 d. After that the cultures were harvested and proteins were extracted for HPLC and SDS-PAGE analysis.

## 3 Results

### 3.1 Generation of a *N. punctiforme* secondary metabolite gene cluster CFP reporter strain library

#### 3.1.1 Identification of potential natural product biosynthesis gene clusters and respective promoter regions

Cyanobacteria of the genus *Nostoc* are known to harbor a high degree of small molecule natural product biosynthesis gene clusters (BGCs) of the NRPS, PKS and RiPP classes (77). In the course of this study *N. punctiforme* PCC73102 was examined since despite it being an established model system for the investigation of terrestrial filamentous symbiotic cyanobacteria there is only limited information available about its secondary metabolism. *N. punctiforme* is one of the few filamentous cyanobacterial strains that combines the genetic potential to be a rich producer of secondary metabolites with an amenability to genetic modifications making it a perfect candidate organism for the scope of this study. To assess the genetic potential of *N. punctiforme* its genome (BioProject: PRJNA216; GenBank assembly accession: GCA\_000020025.1, 2008), was mined for natural product BGCs using the AntiSMASH (version 3.0.5 and later 4.0) web server based tool (46, 82, 83). Since for the majority of identified potential natural product BGCs the product is still elusive, those BGCs were labeled with arbitrary working names.

Together with two NRPS gene clusters that were already reported in the literature, the nostopeptolide and anabaenopeptin BGCs, a total of four NRPS BGCs were identified (Figure 8) (64, 71). The two cryptic NRPS gene clusters are referred to as NRPS1 and NRPS2. Other than NRPS1, NRPS2 is located on one of the native plasmids occurring in *N. punctiforme*.

## NRPS BGCs

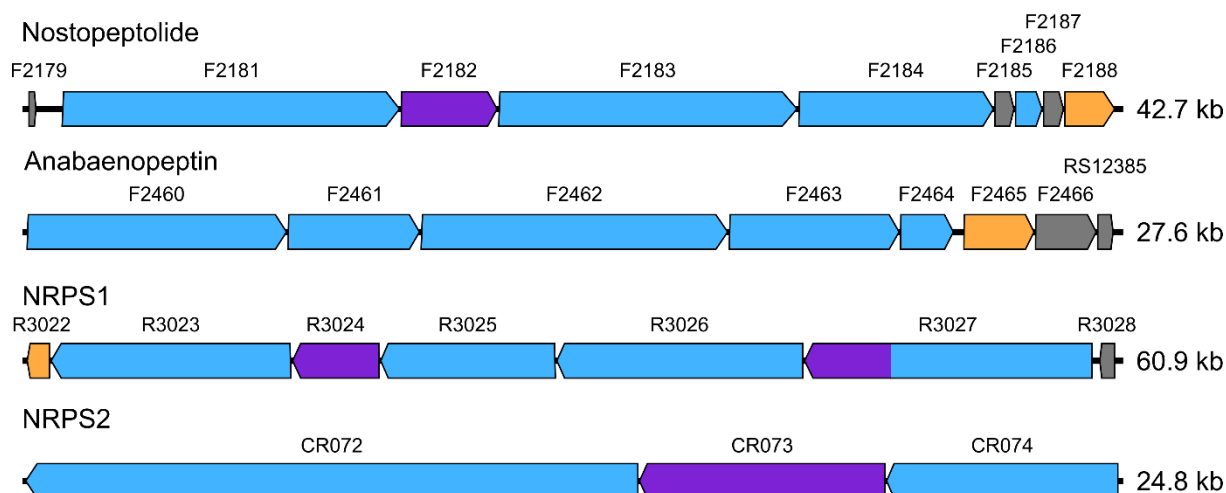


Figure 8 **Potential NRPS BGCs identified in the *N. punctiforme* genome using AntiSMASH.** Genes colored in light blue highlight NRPS genes, purple highlights PKS genes, yellow highlights transport related genes and grey indicates putative genes of unknown function. Gene annotation resemble the “NPun”-type annotations conducted and published by the authors of the original genome sequencing data in 2008, later detected and added genes follow the “RS”-type annotation nomenclature.

There were also three BGCs identified that mainly consist of PKS genes referred to as PKS1, PKS2 and PKS3 (Figure 9).

## PKS BGCs

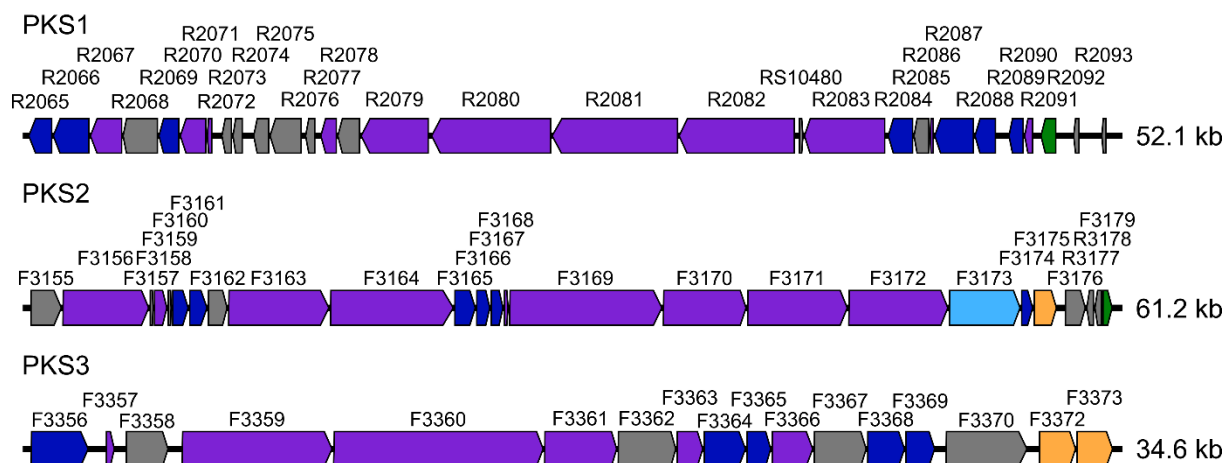


Figure 9 **Potential PKS BGCs identified in the *N. punctiforme* genome using AntiSMASH.** Genes colored in light blue highlight NRPS genes, purple highlights PKS genes, dark blue highlights other biosynthetic genes, yellow highlights transport related genes, green highlights regulatory genes and grey indicates putative genes of unknown function. Gene annotation resemble the “NPun”-type annotations conducted and published by the authors of the original genome sequencing data in 2008, later detected and added genes follow the “RS”-type annotation nomenclature.

Additionally, there were two hybrid BGCs identified encoding both PKS and NRPS genes (Figure 10). These BGCs are referred to as PKS4 and PKS5. PKS4 is the largest potential secondary metabolite BGC detected with 72.0 kb.

## NRPS-PKS hybrid BGCs

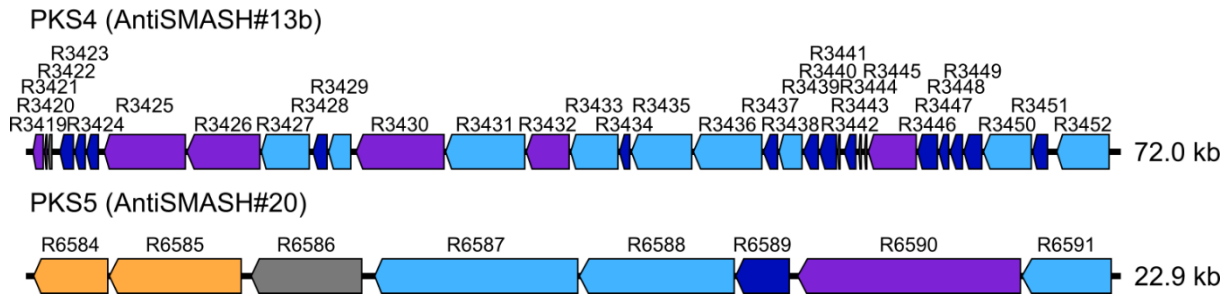


Figure 10 **Potential NRPS-PKS hybrid BGCs identified in the *N. punctiforme* genome using AntiSMASH.** Genes colored in light blue highlight NRPS genes, purple highlights PKS genes, dark blue highlights other biosynthetic genes, yellow highlights transport related genes and grey indicates putative genes of unknown function. Gene annotation resemble the “NPun”-type annotations conducted and published by the authors of the original genome sequencing data in 2008, later detected and added genes follow the “RS”-type annotation nomenclature.

Finally, there were seven RiPP-type BGCs identified referred to as RiPP1a, RiPP1b, RiPP3, RiPP4, RiPP5, RiPP6 and Mvd (microviridin-like) (Figure 11). Two of these gene clusters, RiPP5 and RiPP6, are localized on different native plasmids of *N. punctiforme*.

## RiPP BGCs

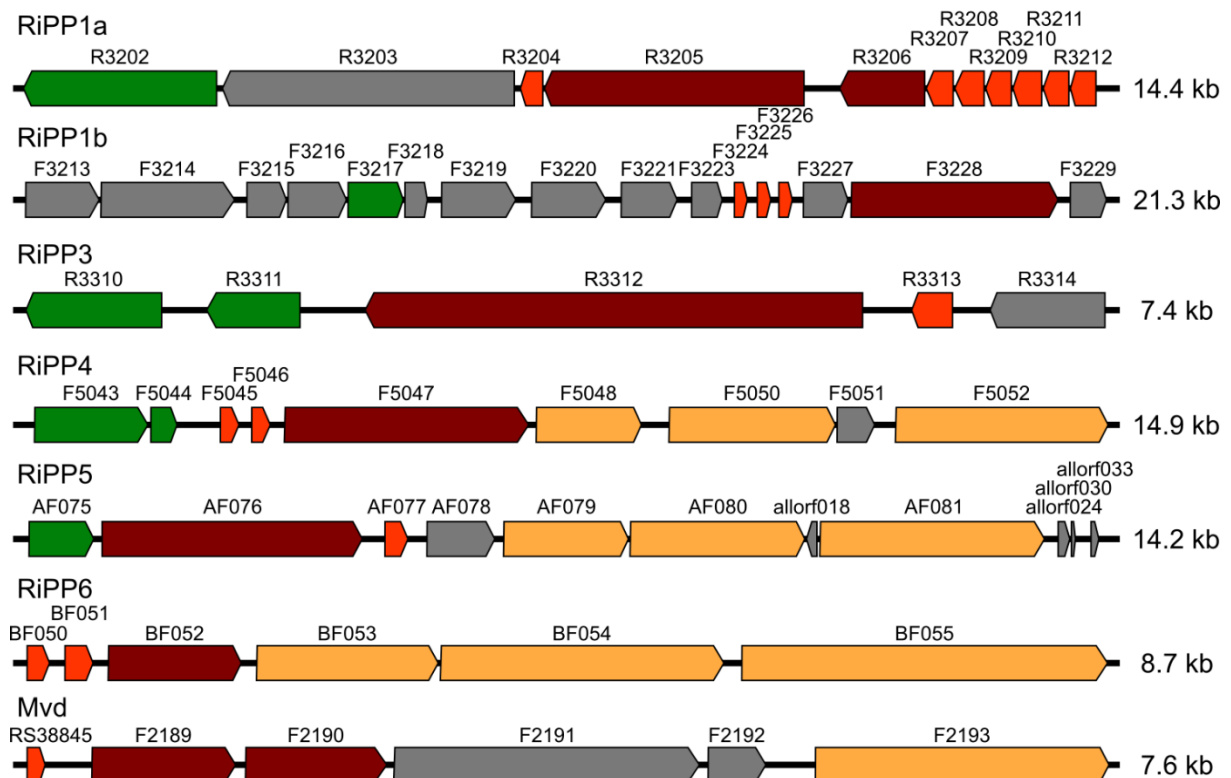


Figure 11 **Potential RiPP BGCs identified in the *N. punctiforme* genome using AntiSMASH.** Genes colored in light red highlight RiPP precursor genes, dark red highlights RiPP tailoring genes, yellow highlights transport related genes, green highlights regulatory genes and grey indicates putative genes of unknown function. Gene annotation resemble the “NPun”-type annotations conducted and published by the authors of the original genome sequencing data in 2008, later detected and added genes follow the “RS”-type annotation nomenclature.



In summary, sixteen natural products BGCs could be identified within *N. punctiforme*'s 9.1 Mb genome with thirteen BGCs being located on the bacterial chromosome and three being located each on a different of the native plasmids. With a total size of 487,3 kb *N. punctiforme* devotes 5.35 % of its total genome to the synthesis of potential natural products. To enable the monitoring of transcriptional activity of those potential natural product BGCs reporter strains should be produced that give a fluorescence signal when a representative gene of a given cluster is transcribed.

For the generation of these fluorescence transcriptional reporter strains for each of the natural product BGCs cyanobacterial shuttle plasmids were designed that encode a CFP coding sequences (CDS) fused to a 5' untranslated region (UTR) upstream of a representative gene of a given BGC of interest (Figure 12).

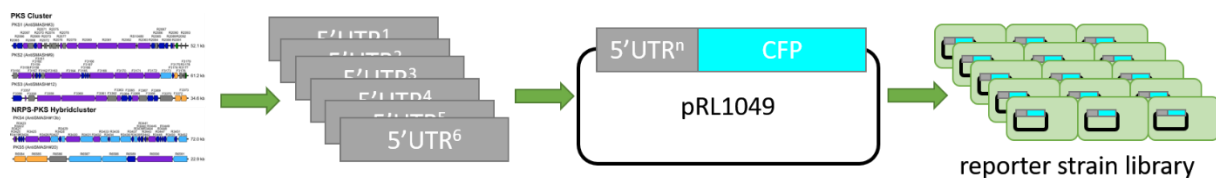


Figure 12 **Schematic workflow for the generation of a *N. punctiforme* reporter strain library for selected secondary metabolite BGCs.**

In cases of NRPS, PKS or hybrid BGCs the 5'UTR of the first gene of a given gene cluster was chosen, whereas for RiPP BGCs the 5'UTR of the first precursor gene was chosen for the CFP fusion construct. However, it was avoided that the gene, the reporter was designed for, was coding for a hypothetical protein and that the intergenic region of the selected gene and the upstream gene was smaller than 100 bp. An overview of the selected target genes that a cyanobacterial reporter strain was designed for and the respective lengths of their promotor regions is shown in Table 4.

Table 4 Selected *N. punctiforme* genes for the construction of transcriptional reporter strains and respective lengths of selected 5'UTRs.

BGC name	Reporter gene	5'UTR length [bp]
Nostopeptolide	NPUN_F2181	674
Anabaenopeptin	NPUN_F2460	957
NRPS1	NPUN_R3027	477
NRPS2	NPUN_CR074	815
PKS1	NPUN_R2091	3194
PKS2	NPUN_F3155	1014
PKS3	NPUN_F3356	706
PKS4	NPUN_R3452	1214
PKS5	NPUN_R6591	1038
RiPP1a	NPUN_R3212	388
RiPP1b	NPUN_F3224	282
RiPP3	NPUN_R3313	516
RiPP4	NPUN_F5045	614
RiPP5	NPUN_AF077	530
RiPP6	NPUN_BF050	451
Mvd	RS38845	488

### 3.1.2 Construction of the cyanobacterial shuttle plasmids and *N. punctiforme* reporter strain mutagenesis

Every *N. punctiforme* reporter strain harbors a cyanobacterial shuttle plasmid coding for a 5'UTR-CFP-fusion construct, a resistance marker and a cyanobacterial origin of replication (*ori*) as well as an *E. coli* ColE1 *ori*. These plasmids are referred to as reporter plasmids and are all designed based on the pRL1049 vector (74, 84). In order to allow for quick assembly of desired reporter plasmids in *E. coli* a precursor plasmid named pDD001 was assembled by insertion of a CFP coding sequence into pRL1049 vector via HiFi-builder mediated homologous recombination. Therefore, a CFP coding sequence insert was amplified via PCR with primers D001/D002 on pECFP-C1 (Takara Bio Europe, France) and thereby attaching 25-bp homology sequence stretches on both ends of the PCR product as well as the missing bases to recover the unique EcoRI recognition site after the assembly. The PCR product was purified via agarose gel electrophoresis followed by gel extraction. The CFP insert was assembled into EcoRI digested pRL1049 by HiFi-builder mediated homologous recombination (see section 2.2.2.8). After transformation and identification of potential positive *E. coli* XL-1Blue mutants by colony PCR with primers D021(S1) and D022(S2), overnight cultures were prepared for plasmid isolation and sequencing was conducted to confirm the correct assembly of the desired plasmid. A vector map of the precursor plasmid pDD001 is depicted in Figure 13.

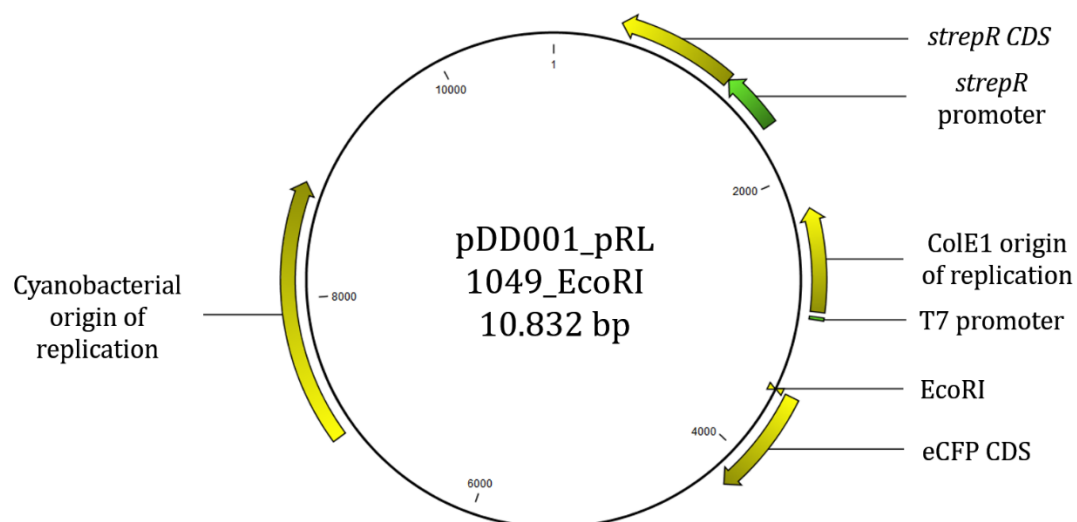


Figure 13 Vector map of final precursor plasmid pDD001.

Except for the reporter plasmids for the nostopeptolide and anabaenopeptin BGC that were constructed as described earlier (71), EcoRI digested pDD001 served as the vector backbone for all reporter plasmids constructed within this study (71, 77). To generate the final reporter plasmids the different 5'UTR sequence inserts of the identified natural product BGCs were amplified via PCR with the respective primers (listed in Table 5) on *N. punctiforme* genomic DNA, thereby extending the PCR products by 25-bp homology sequence stretches on both ends conveying homology towards EcoRI digested pDD001. An overview of the conducted amplifications, target genes and names of the final reporter plasmids are depicted in Table 5.

Table 5 Target 5'UTR amplification primers and corresponding reporter plasmids.

BGC	Target 5'UTR	Primer	Final reporter construct
NRPS1	NPUN_R3027	D005/D006	pDD003
NRPS2	NPUN_CR074	D017/D018	pDD009
PKS1	NPUN_R2091	D110/D111	pDD032
PKS2	NPUN_F3155	D007/D008	pDD004
PKS3	NPUN_F3356	D112/D113	pDD033
PKS4	NPUN_R3452	D013/D014	pDD007
PKS5	NPUN_R6591	D108/D109	pDD031
RiPP1a	NPUN_R3212	D072/D073	pDD025
RiPP1b	NPUN_F3224	F3224-fw/F3224-rv	F3224-CFP
RiPP3	NPUN_R3313	D100/D101	pDD027
RiPP4	NPUN_F5045	D041/D042	pDD011
RiPP5	NPUN_AF077	D154/D155	pDD043
RiPP6	NPUN_BF050	D156/D157	pDD044
Mvd	RS38845	D158/D159	pDD045

After successful PCR amplification of the target 5'UTR inserts, the fragments were purified by agarose gel electrophoresis with subsequent gel extraction. Then the reporter plasmids were assembled by HiFi-builder mediated homologous recombination of the 5'UTRs into EcoRI digested pDD001 backbone (according to section 2.2.2.8). After transformation into *E. coli* XL-1Blue positive mutants were identified via colony PCR with the primers D021(S1) und D023(S3) followed by overnight culture preparation, plasmid isolation and confirmation of the correct assembly by sequencing.

For the generation of *N. punctiforme* transcriptional CFP reporter strains the aforementioned reporter plasmids were transformed into *N. punctiforme* wild-type cells as described in section 2.2.3.2. For the electroporation of the reporter plasmids into *N. punctiforme* large amounts of plasmid DNA were required. Since each mutagenesis approach was run in quadruplicates 40 µg plasmid DNA were required per reporter construct. For that 50 ml LB-medium supplemented with 20 µg/ml streptomycin were inoculated with a given *E. coli* XL-1Blue strain harboring the respective reporter plasmid, incubated over night at 37 °C, 220 rpm and plasmid isolation was conducted the next day (according to section 2.2.2.5) to yield between 50 and 200 µg plasmid DNA. After electroporation of *N. punctiforme* wild-type, initial growth on BG11<sub>0</sub>-agar plates without antibiotics and background die-off and rise of potential mutant colonies after transfer on BG11<sub>0</sub>-agar supplemented with 2 µg/ml streptomycin (Figure 14) was observed, the colonies growing on the HATF membrane were sliced out and placed into BG11<sub>0</sub>-medium supplemented with 2 µg/ml in culture flasks for the transition from solid-phase to liquid growth.

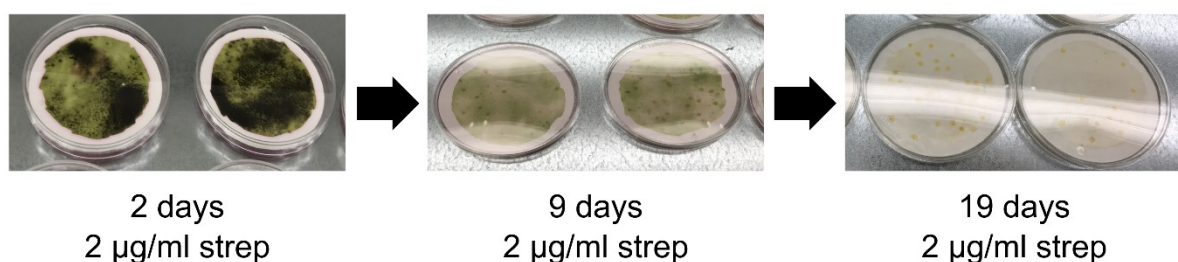


Figure 14 **Time course of an *N. punctiforme* selection process after transfer on BG11<sub>0</sub>-agar plates supplemented with 2 µg/ml streptomycin.**

Those transition cultures were incubated under standard low light conditions and ambient air for up to four weeks or until macroscopic growth was observed. To verify that the potential *N. punctiforme* reporter mutants were harboring the desired reporter

plasmids, plasmid DNA was extracted from samples of these growing transition cultures using the GeneElute Plant Genomic DNA Miniprep Kit (according to section 2.2.3.5). The isolated DNA was used as a colony PCR template to confirm the presence of the desired reporter plasmid. In contrast to *E. coli* colony PCR where mutant cells directly served as a template source here 1 µl of the plasmid DNA isolates were used as a template for the PCR mixture. Using the primers D021(S1) and D022(S2) as universal primers amplifying the 5'UTR-CFP fusion construct on any pDD001-based reporter plasmid together with a positive control (purified reporter plasmid as template) as well as *N. punctiforme* wild-type genomic DNA and no-template negative controls, the presence of a given reporter plasmid in the selected *N. punctiforme* mutant was confirmed.

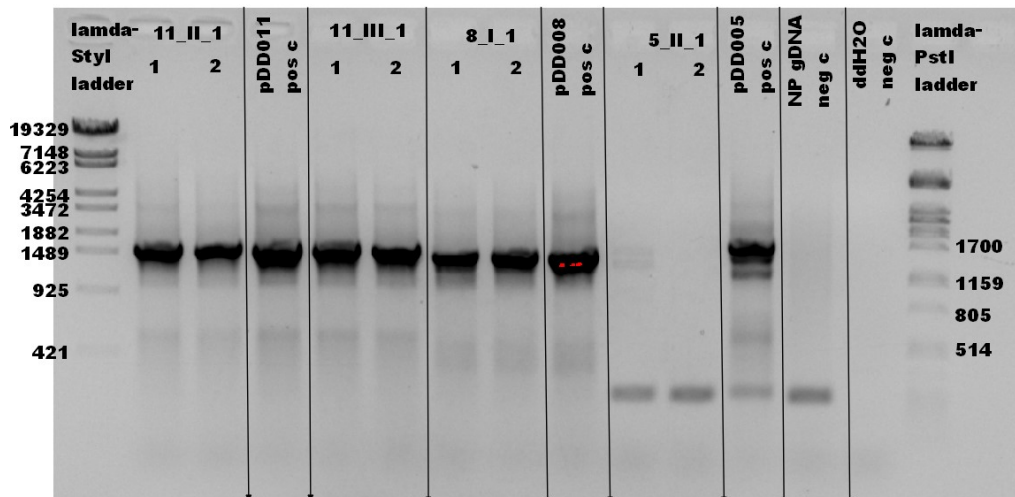


Figure 15 **Result of a *N. punctiforme* reporter strain generation confirmation PCR including all controls.** Primers used were D021(S1) and D022(S2). Expected product sizes: pDD011 - 1571bp; pDD008 - 1485 bp; pDD005 - 1704 bp.

An example of such a confirmation PCR is depicted in Figure 15, where the presence of the reporter plasmids pDD011, pDD008 and pDD005 in selected potential *Nostoc* mutant colony isolates should be confirmed. Each sample PCR was run in duplicates with a single positive control per plasmid and a single *Nostoc* genomic DNA negative control and a single no-template control for all samples. In this example the presence of pDD011 and pDD008 in the selected *Nostoc* mutant colonies could be confirmed, whereas the presence of pDD005 could not be confirmed since the respective PCR was negative (Figure 15). According to the expectation neither the *Nostoc* wild-type genomic DNA isolate nor the no-template control resulted in a PCR product using the primers D021(S1) and D022(S2).

This selection and confirmation procedure was repeated for all the designed *N. punctiforme* reporter strains. Once the reporter strain library was completely

confirmed and transferred to liquid growth it enabled the single cell monitoring of transcriptional responses to different cultivation conditions or external factors for each of the sixteen potential natural product BGCs within *N. punctiforme*.

### 3.1.3 Transcriptome analysis of a growing *N. punctiforme* culture

An alternative way of tracking the transcription levels of selected genes of a given organism is to perform RNA deep sequencing. This method yields values for the average transcription level of every gene of an organism at the timepoint of RNA isolation relative to the transcription levels of selected marker genes. The transcriptome analysis was additionally performed as a time series of a growing *N. punctiforme* wild-type culture. As described in section 2.2.4.8, a *N. punctiforme* culture was spread on HATF membranes and incubated on BG11<sub>0</sub>-agar for up to nine days under standard low light conditions and ambient air. After 0, 1, 3, 5, 7 and 9 days cells were harvested, RNA was isolated and the isolates were then shipped to Dr. Jenke-Kodama's laboratory at the OIST (Okinawa, Japan) for mRNA library preparation, Illumina sequencing and data processing (as described in section 2.2.4.8 and Dehm *et al.* 2019) (80). The time series study was performed in duplicates to reduce potential variations in culture condition or sample processing. A schematic representation of the conducted experimental workflow is depicted in Figure 16.

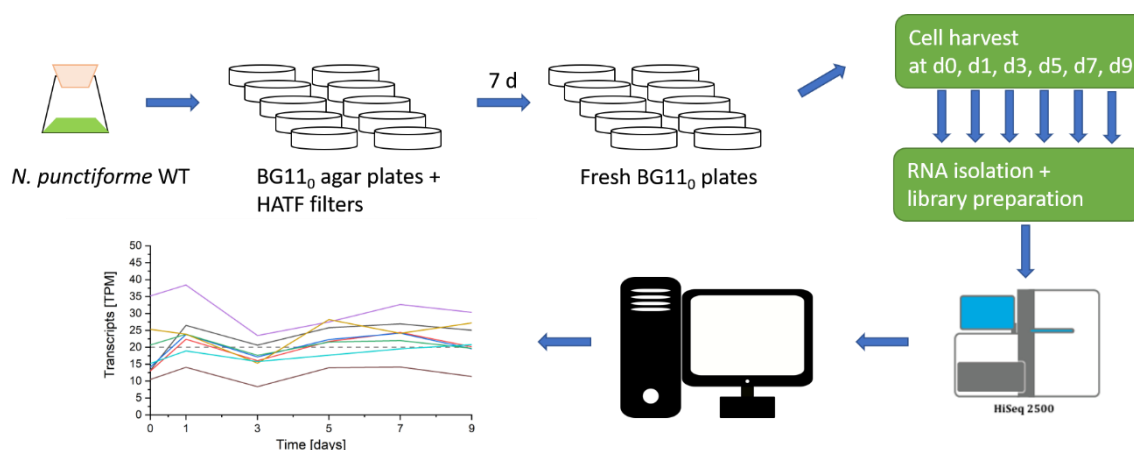


Figure 16 Schematic representation of RNA deep-sequencing experimental workflow.

The RNA sequencing raw data were processed to yield TPM values for all annotated genes of *N. punctiforme*. To check for the transcriptional behavior of the different secondary metabolite BGCs the TPM values of all genes of a given cluster were plotted over the sampling timepoints and according to the study of Amos *et al.* (2017), conducting an interspecies comparison of natural product BGC expression in *Salinispora* strains, an

empirical threshold of 20 TPM for actively expressed genes was assumed and highlighted in the plots (85). The results of the transcriptome analysis of the four NRPS BGCs are depicted in Figure 17.

## NRPS BGCs

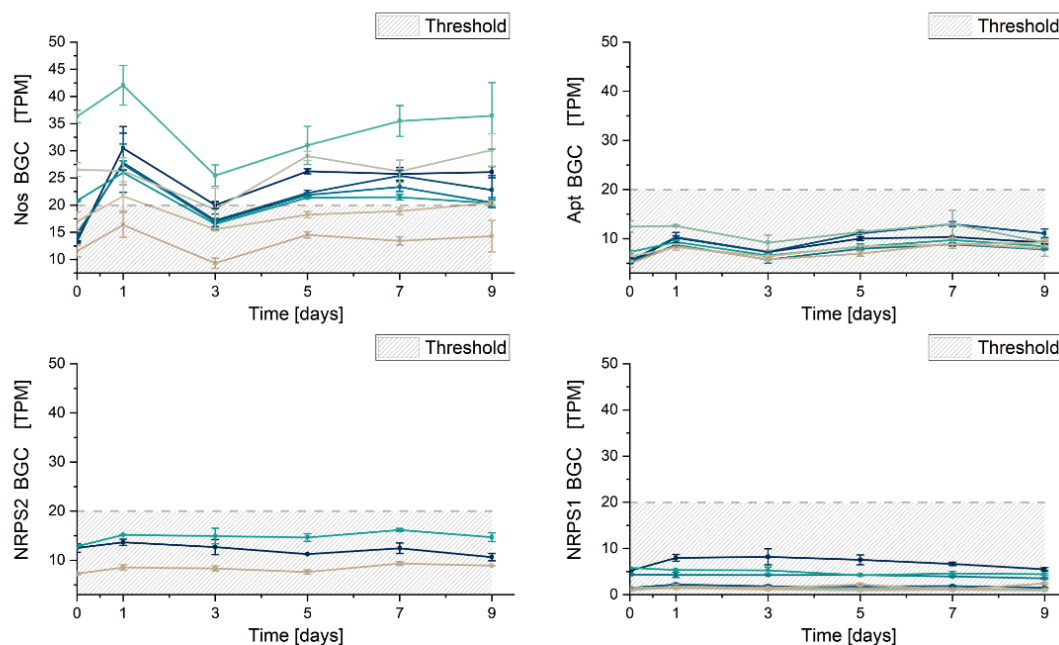


Figure 17 **Full gene cluster transcription levels of *N. punctiforme*'s NRPS BGCs Nos, Apt, NRPS1 and NRPS2.** Dashed grey line indicates empirical threshold for actively expressed genes of 20 TPM. Each curve depicts the transcription level of a single gene of the respective BGC at each timepoint. The color scheme resembles a heatmap from dark blue to light brown, darker genes are located more upstream towards the start of the gene cluster than lighter genes. Detailed plots including gene legends are summarized in appendix X section 0.

The majority of genes of the nostopeptolide NRPS BGC showed transcription levels beyond the empirical threshold of 20 TPM according to expectations since nostopeptolide is one of the three already described natural products of *N. punctiforme* that can be isolated under standard growth conditions. Furthermore, a general increase of transcriptional activity 24h after transfer of the culture carrying membranes to fresh BG11<sub>0</sub>-agar plates could be observed. However, the genes associated with the production of anabaenopeptin all show transcription levels below the active expression threshold. Both the NRPS1 and NRPS2 BGCs also yielded TPM values below the threshold for all sampling timepoints, thereby no differences in transcription levels could be observed whether the BGC is located on the bacterial chromosome (NRPS1) or an external plasmid (NRPS2).

## PKS BGCs

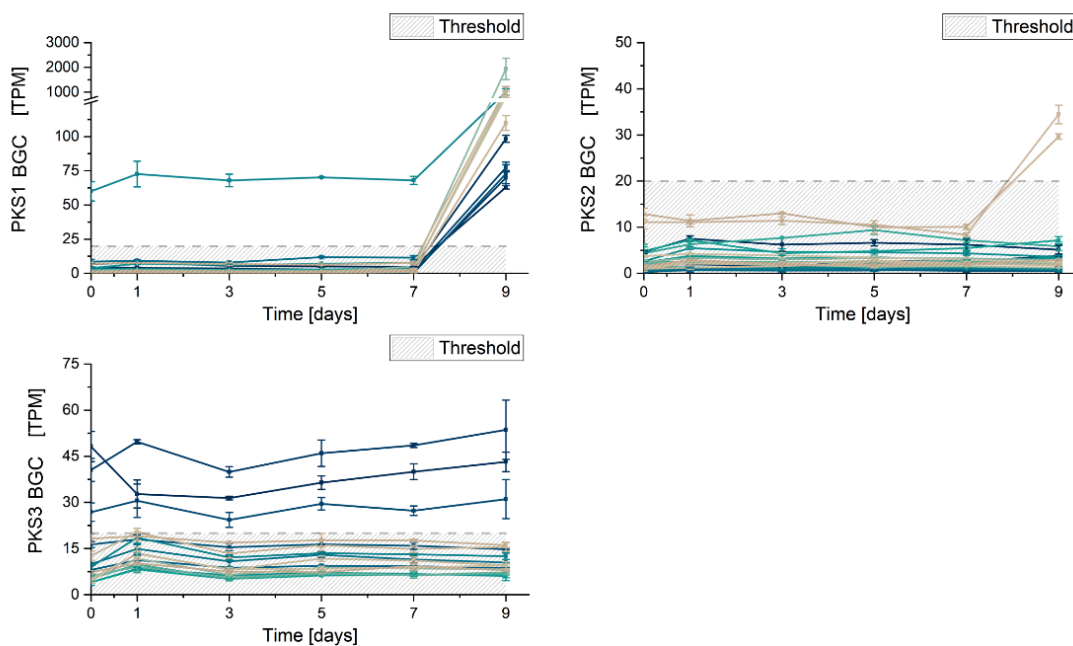


Figure 18 Full gene cluster transcription levels of *N. punctiforme*'s PKS BGCs PKS1, PKS2 and PKS3. Dashed grey line indicates empirical threshold for actively expressed genes of 20 TPM. Each curve depicts the transcription level of a single gene of the respective BGC at each timepoint. The color scheme resembles a heatmap from dark blue to light brown, darker genes are located more upstream towards the start of the gene cluster than lighter genes. Detailed plots including gene legends are summarized in appendix section 0.

Compared to the NRPS BGCs the PKS BGCs of *N. punctiforme* show a different transcriptional behavior both in terms of variation of transcription level of all genes of a gene cluster at a single sampling timepoint as well as transcriptional changes during the course of the nine day time series (Figure 18). Despite the fact that for all three PKS BGCs the majority of genes show transcription levels below the empirical threshold at most of the sampling timepoints unique behaviors of specific genes could be observed. For example, Npun\_R2076 a gene coding for a Hsp90 activator homologue and being part of the PKS1 BGC shows increased transcription levels over the full course of the experiment. Similarly, Npun\_F3356, Npun\_F3357 and Npun\_F3358 of *N. punctiforme*'s PKS3 BGC, coding for an AMP-dependent synthetase and ligase, a phosphopantetheine-binding protein and a conserved cyanobacterial hypothetical protein, respectively, all show elevated transcription levels above the threshold and over the full course of the experiment compared to the remaining genes of the BGC. Additionally, despite one exception all genes of the PKS3 BGC show increased transcription levels in various intensities after 24 h. Another uncommon transcriptional behavior that could be observed was the tremendous increase in transcription levels between day 7 and day 9 of the time course experiment. These elevated TPM values could be observed for all genes of the



PKS1 BGC. However, to a lesser extent a similar behavior could also be observed for two genes of the PKS2 BGC, namely Npun\_R3177 and Npun\_R3178 coding for a polyketide cyclase and a putative mono-oxygenase ydhR, respectively.

## NRPS-PKS hybrid BGCs

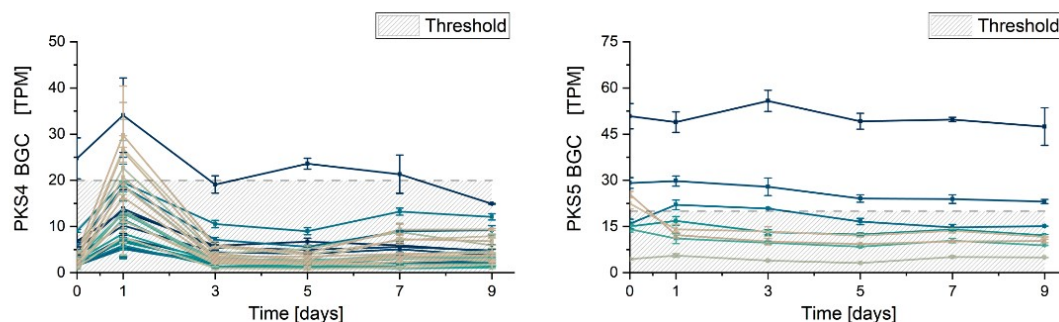


Figure 19 **Full gene cluster transcription levels of *N. punctiforme*'s NRPS-PKS hybrid BGCs PKS4 and PKS5.** Dashed grey line indicates empirical threshold for actively expressed genes of 20 TPM. Each curve depicts the transcription level of a single gene of the respective BGC at each timepoint. The color scheme resembles a heatmap from dark blue to light brown, darker genes are located more upstream towards the start of the gene cluster than lighter genes. Detailed plots including gene legends are summarized in appendix section 0.

The two NRPS-PKS hybrid BGCs show a differential transcription pattern (Figure 19). In the case of the PKS4 BGC again a single gene, namely Npun\_R3420 coding for a MbtH family protein, shows a distinct higher transcription level than the remaining genes of the cluster. Furthermore, a transcriptional peak activity could be observed after 24 h, during that time point five other genes, namely Npun\_R3441 (putative D-alanyl carrier protein), Npun\_R3442 (3-hydroxyacyl-CoA dehydrogenase), Npun\_R3449 (glycosyltransferase family 1 protein), Npun\_R3451 (TauD/TfdA family dioxygenase) and Npun\_R3452 (non-ribosomal peptide synthetase) also surpassed the empirical threshold of 20 TPM. On the other hand, the PKS5 BGC associated genes show no elevated transcription levels after 24h. Interestingly the transport related last three genes of the cluster, namely Npun\_R6584, Npun\_R6585 and Npun\_R6586 coding for a HlyD family efflux transporter periplasmic adaptor subunit, a type I secretion system permease and a FG-GAP repeat-containing protein, respectively, were all transcribed at or above the threshold. In contrast to that the biosynthesis related genes were all transcribed with less than 20 TPM.

## RiPP BGCs

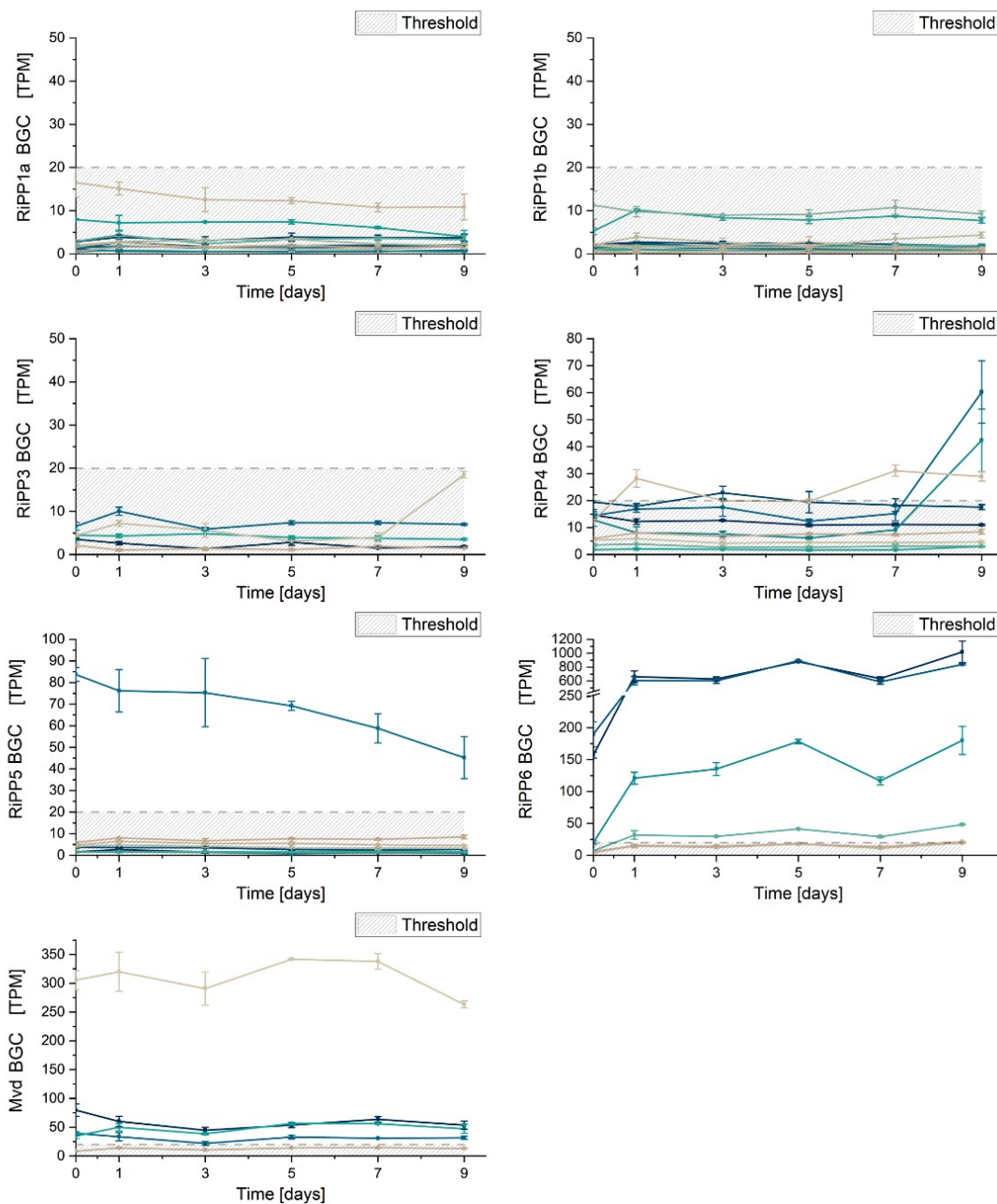


Figure 20 **Full gene cluster transcription levels of *N. punctiforme*'s RiPP BGCs RiPP1a, RiPP1b, RiPP3, RiPP4, RiPP5, RiPP6 and Mvd.** Dashed grey line indicates empirical threshold for actively expressed genes of 20 TPM. Each curve depicts the transcription level of a single gene of the respective BGC at each timepoint. The color scheme resembles a heatmap from dark blue to light brown, darker genes are located more upstream towards the start of the gene cluster than lighter genes. Detailed plots including gene legends are summarized in appendix section 0.

In terms of transcriptional behavior, for the RiPP BGCs of *N. punctiforme* heterogeneous patterns could be observed. Whereas RiPP1a, RiPP1b and RiPP3 were all transcribed at levels clearly below the empirical threshold, certain genes of RiPP4 were transcribed at or slightly above the threshold. These genes include Npun\_F5044 (response regulator), Npun\_F5051 (putative site-specific DNA-methyltransferase) and Npun\_F5045 (Nif11-like

leader peptide family natural product precursor). Interestingly Npun\_F5045 together with another RiPP precursor gene (Npun\_F5046) showed a strong increase in transcriptional activity at day 9. For the two plasmid-located RiPP BGCs RiPP5 and RiPP6 also unique transcriptional patterns could be observed. In case of RiPP5 only a single gene, namely Npun\_AF077, the putative precursor of RiPP5, showed transcription levels above the empirical threshold, whereas all other cluster related genes barely exceeded 10 TPM. On the other hand, in case of RiPP6 all genes of the cluster except two (Npun\_BF054 and Npun\_BF055) did exceed the threshold and moreover the two putative precursor genes showed heavily elevated transcription levels between 600 and 1000 TPM. The chromosomal Mvd BGC showed both overall high transcription levels for the majority of genes and a single gene (Npun\_F2192, a conserved hypothetical protein) with strongly elevated transcription levels at all time points.

In general, the whole transcriptome study confirmed the observed fact that most of *N. punctiforme*'s secondary metabolite genes were transcribed at minimal levels in a culture wide average. However, utilizing the fluorescent transcriptional reporter mutants of *N. punctiforme* further insights into the transcriptional activity of selected secondary metabolite genes on a sub-cultural level could be gathered.

#### **3.1.4 Evaluation of reporter strain CFP signal patterns in comparison to the RNA deep-seq data**

To further investigate the transcriptional activity of all *N. punctiforme* secondary metabolite BGCs the above described CFP reporter strain library was generated. Utilizing the different reporter mutants, the transcriptional activity of selected genes could be monitored as CFP signals on a single cell level by fluorescence microscopy. First a proof of concept was generated by transformation of the unmodified pDD001 plasmid (carrying the CFP CDS without any promoter sequence) as a negative control into *N. punctiforme*. Compared to the constitutively expressed PKS3 BGC Npun\_F3356 reporter strain it could be shown that the negative control shows no unspecific CFP signal nor a phenotype (Figure 21).

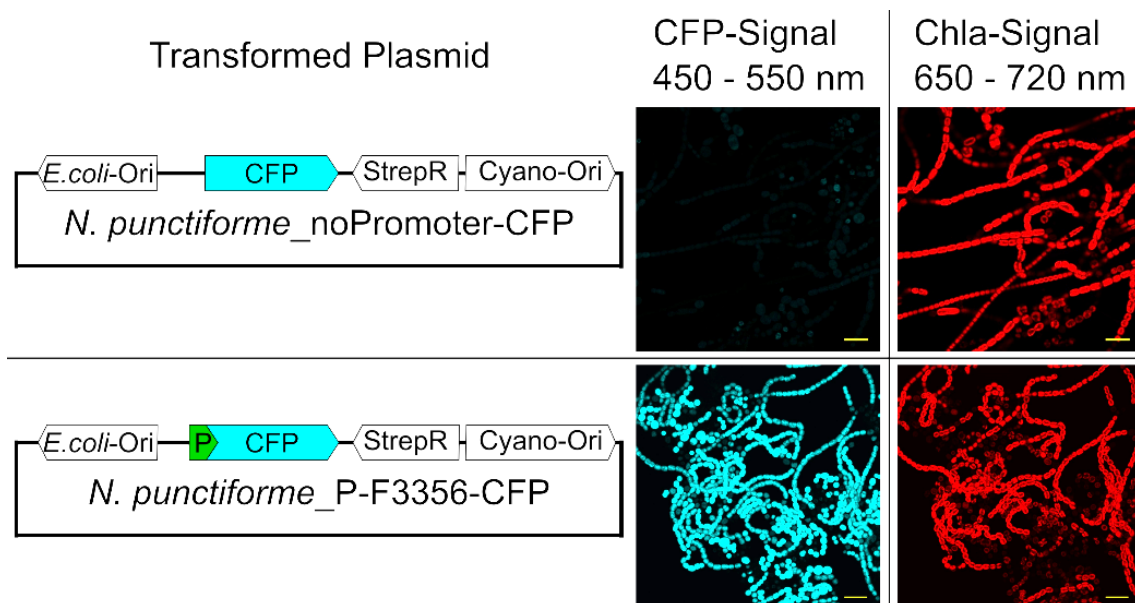


Figure 21 **Proof of concept for *N. punctiforme* transcriptional reporter strains.** Lefthand side shows schemes of the transformed reporter plasmid variants, righthand side shows confocal fluorescence micrographs of the corresponding *N. punctiforme* reporter mutants under standard low light conditions and ambient air. Upper row shows the *N. punctiforme* no-promoter negative control mutant strain and lower row shows the *N. punctiforme* PKS3 Npun\_F3356 reporter mutant strain.

Next, the transcriptional activity of all generated reporter strains under standard low light conditions was investigated and in parallel compared to the results of the time course RNA deep-sequencing study.

Based on the RNA deep-sequencing results, the majority of the *N. punctiforme* reporter mutants were expected to show no or very faint CFP signals indicating the low expression rates observed for most of the secondary metabolite BGCs. The transcriptional reporter strains for PKS1 (Npun\_R2091) and PKS2 (Npun\_F3155) confirmed the expected behavior as depicted in Figure 22. The chlorophyll  $\alpha$  mediated auto-fluorescence (Figure 22, B, right image) proves that the filaments were vital but no CFP signal was observed (Figure 22, B, left image). Accordingly, the respective genes Npun\_R2091 and Npun\_F3155 were transcribed at minimum levels in the culture-wide RNA deep-sequencing experiment (Figure 22, A).

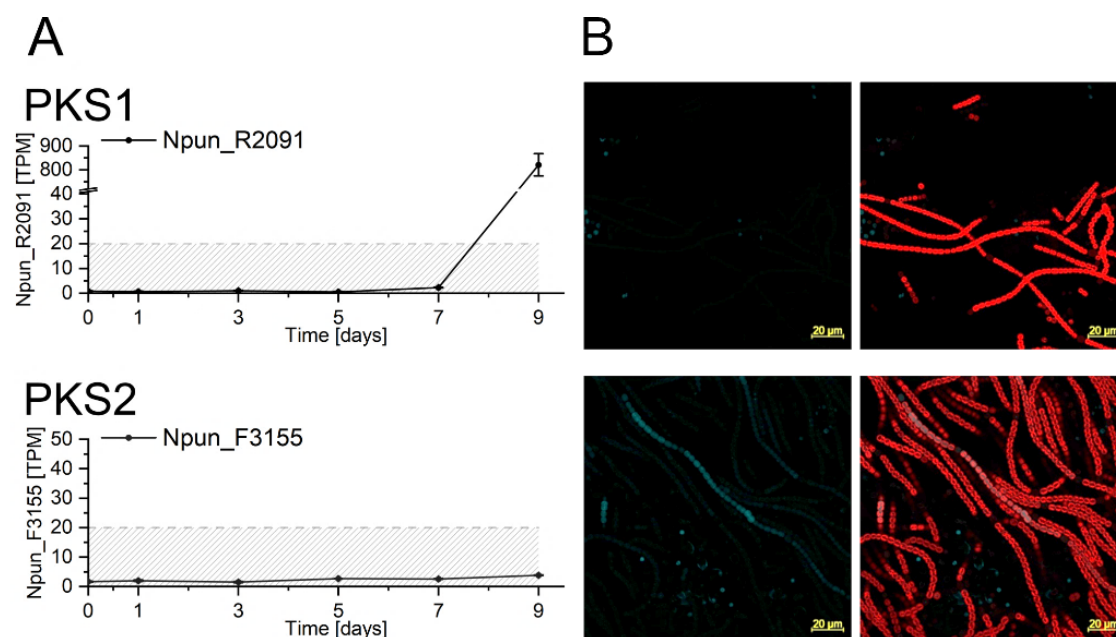


Figure 22 *N. punctiforme* secondary metabolite BGCs showing no or faint expression under standard lowlight conditions. A) Average TPM value of the gene that was selected for the construction of the reporter plasmid in the RNA deep-sequencing time course experiment. Dashed grey line indicates the empirical threshold of 20 TPM for actively expressed genes; B) confocal fluorescence microscopy picture of the respective *N. punctiforme* reporter strain under standard low light cultivation and ambient air. Left image shows CFP channel only and right image CFP/Chl $\alpha$  channels merged. Scale bar indicates 20  $\mu$ m.

Three reporter strains, namely nostopeptolide (Npun\_F2181), PKS3 (Npun\_F3356) and RiPP6 (Npun\_BF050), showed a prominent CFP signal under standard low light conditions and ambient air (Figure 23, B). The nostopeptolide reporter strain showed a homogeneous CFP signal among all cells of a filament under standard low light conditions. Although the overall CFP intensity was weakest among the three reporter strains this observation was consistent with the lower TPM values of Npun\_2181 just slightly above the empirical threshold. Since nostopeptolide is one of the few described natural products of *N. punctiforme* and is produced in quantitative amounts under regular growth conditions, this result confirmed the expected behavior. For the RiPP6 transcriptional reporter strain also a homogeneous CFP signal was detected, but also a strong phenotype with shorter filaments and more roundish cells with partially increased cell size could be observed. However, although the Npun\_F3356 and Npun\_BF050 reporter strains showed intense CFP signals throughout the culture together with elevated TPM values for the respective genes in the RNA deep-sequencing experiment, no compound could be associated with the PKS3 or RiPP6 BGCs yet.

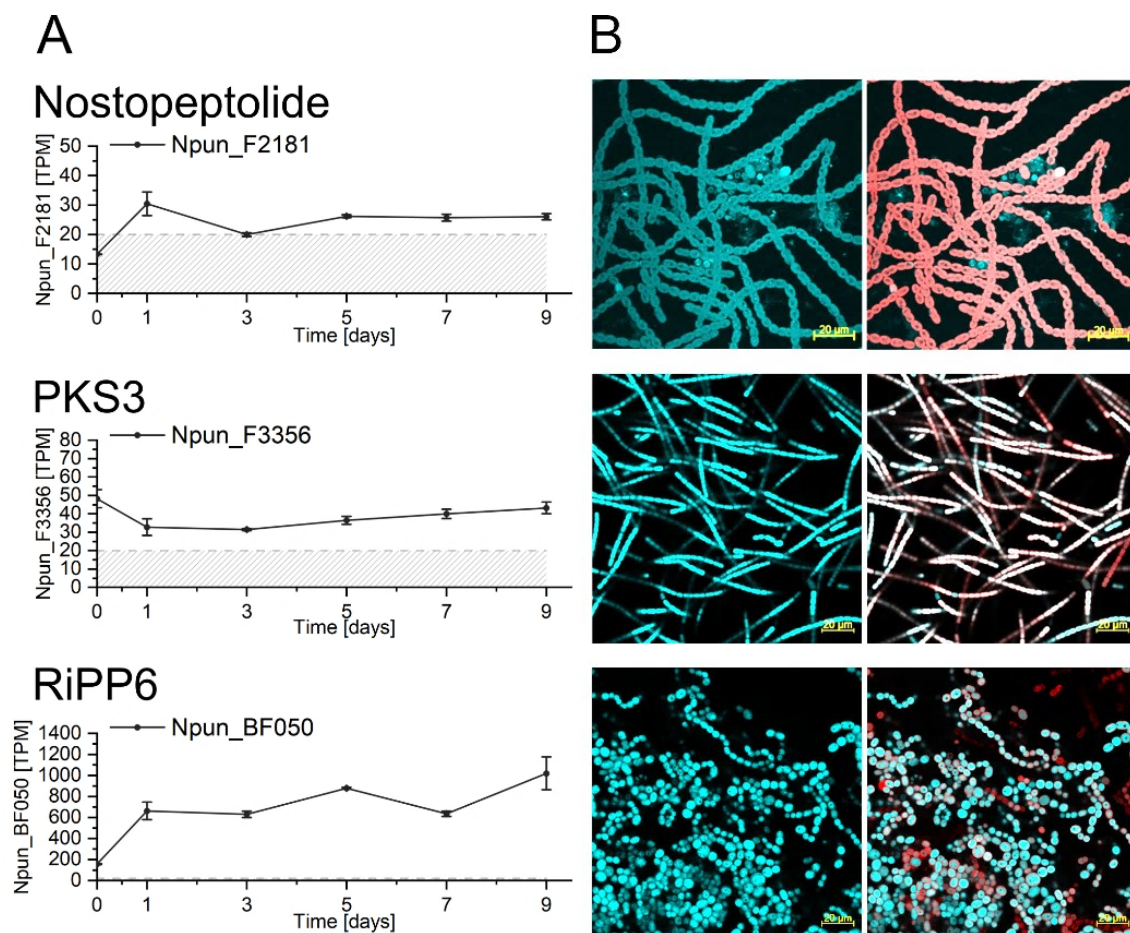


Figure 23 *N. punctiforme* secondary metabolite BGCs showing a constitutive expression under standard lowlight conditions. A) Average TPM value of the gene that was selected for the construction of the reporter plasmid in the RNA deep-sequencing time course experiment. Dashed grey line indicates the empirical threshold of 20 TPM for actively expressed genes; B) confocal fluorescence microscopy picture of the respective *N. punctiforme* reporter strain under standard low light cultivation and ambient air. Left image shows CFP channel only and right image CFP/ Chl $\alpha$  channels merged. Scale bar indicates 20  $\mu$ m.

The above observed transcriptional behaviors can in general be categorized as “constitutively ON” and “OFF” under standard low light conditions and ambient air. However, the eleven remaining reporter strains could neither be associated to the first nor the latter category. Those reporter strains showed recurring, very specific and clearly distinguishable CFP signals under standard low light conditions and ambient air. Thus, these reporter strains were categorized into an independent group that will be referred to as “spatially restricted” expression. However, this group can be subdivided into three expression patterns shared among certain reporter strains.

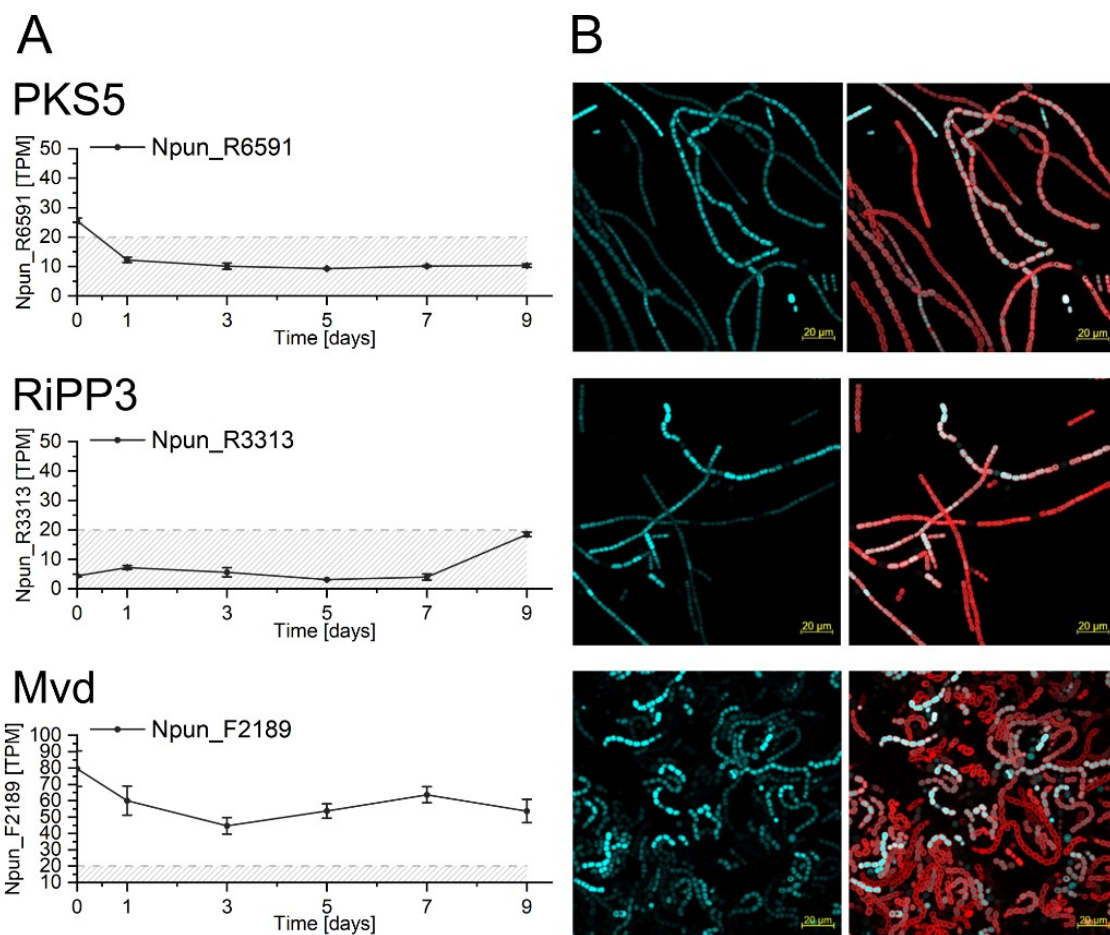


Figure 24 *N. punctiforme* secondary metabolite BGCs showing a spatially restricted expression pattern in long single filaments under standard lowlight conditions. A) Average TPM value of the gene that was selected for the construction of the reporter plasmid in the RNA deep-sequencing time course experiment. Dashed grey line indicates the empirical threshold of 20 TPM for actively expressed genes; B) confocal fluorescence microscopy picture of the respective *N. punctiforme* reporter strain under standard low light cultivation and ambient air. Left image shows CFP channel only and right image CFP/ Chl $\alpha$  channels merged. Scale bar indicates 20  $\mu$ m.

The first subgroup consists of the PKS5, RiPP3 and Mvd BGCs. The unique signal pattern observed among reporter strains of this group is restricted to single long vegetative filaments within a culture (Figure 24, B). Compared to the surrounding cells the CFP signal was very potent in every cell of these isolated long filaments. However, there was still a faint signal detectable in those non-potent filaments. This transcriptional activity could not be confirmed by the average activity observed in the RNA deep-sequencing experiment for the PKS5 and RiPP3 BGCs fitting to the observation that the fraction of active potent filaments is comparatively small (Figure 24, A). In case of the Mvd reporter the potent filaments were shorter compared to the other two reporter strains, but the number of active filaments was also elevated. However, the general transcriptional activity of the microviridin-like BGC is also higher than the activity of the PKS5 and RiPP3 BGCs (Figure 19, Figure 20) and thus fitting to the increased number of active filaments.

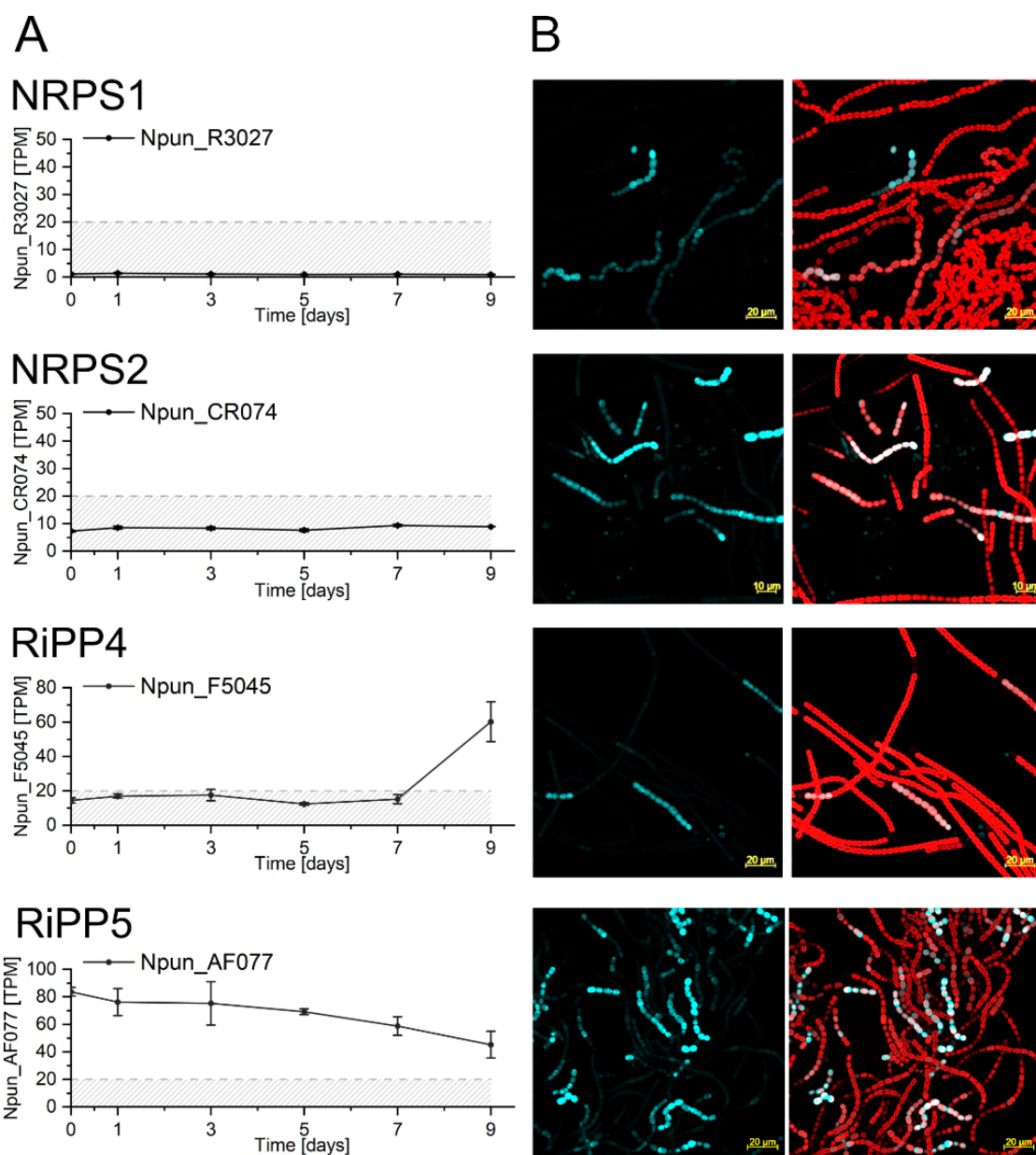


Figure 25 *N. punctiforme* secondary metabolite BGCs showing a spatially restricted expression pattern in short isolated filaments under standard lowlight conditions. A) Average TPM value of the gene that was selected for the construction of the reporter plasmid in the RNA deep-sequencing time course experiment. Dashed grey line indicates the empirical threshold of 20 TPM for actively expressed genes; B) confocal fluorescence microscopy picture of the respective *N. punctiforme* reporter strain under standard low light cultivation and ambient air. Left image shows CFP channel only and right image CFP/ Chl $\alpha$  channels merged. Scale bar indicates 10 or 20  $\mu$ m as specified above.

Second, there was a CFP signal pattern observed that could be described as expression in short isolated filaments with low background activity in neighboring filaments. To this group the NRPS1, NRPS2, RiPP4 and RiPP5 reporter strains could be associated (Figure 25, B). Comparable to the former group for three of the four strains the general transcriptional activity monitored within the RNA deep-sequencing experiment was below the empirical threshold and merely the plasmid based RiPP5 BGC gene Npun\_AF077 showed elevated TPM values.



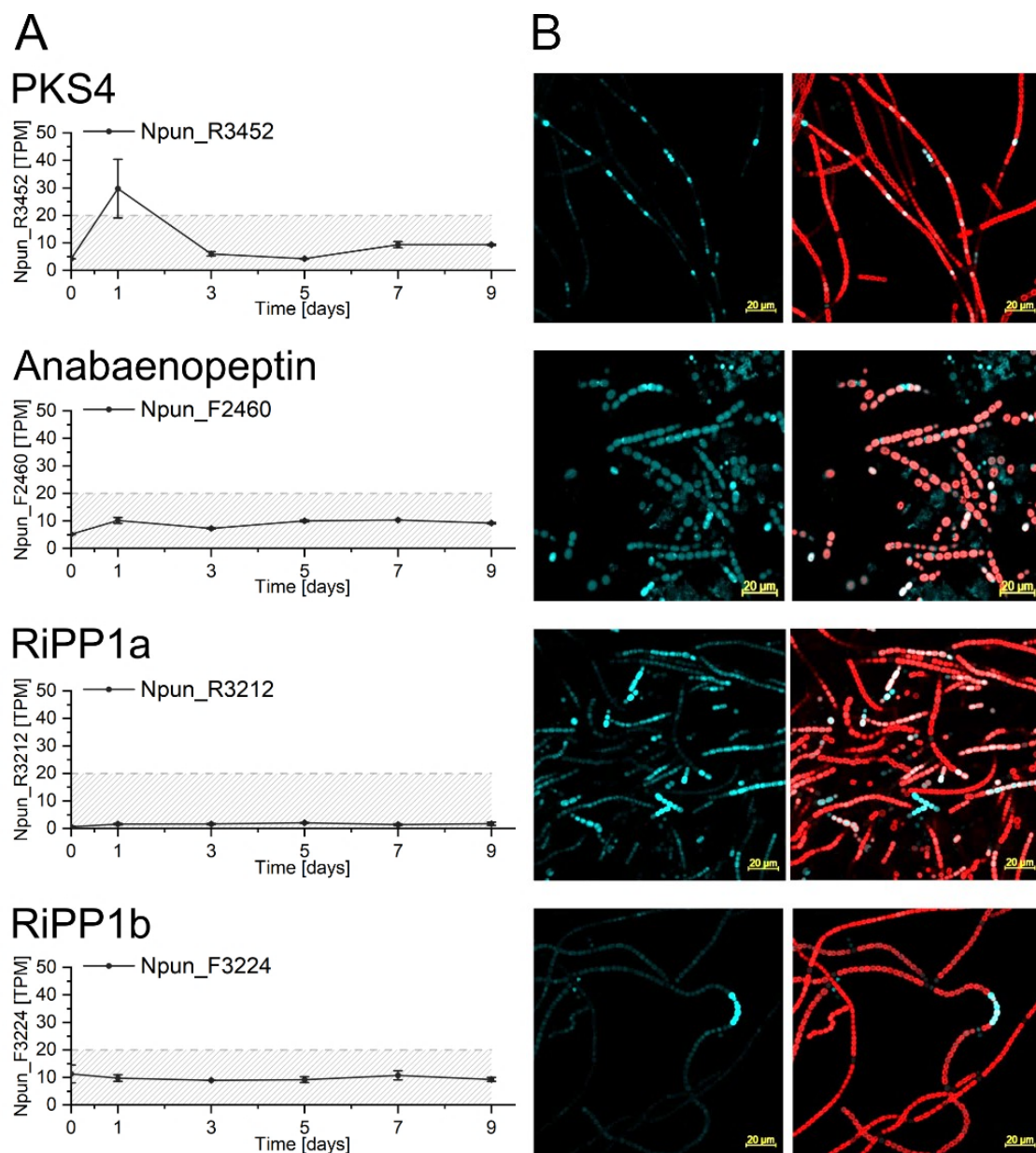


Figure 26 *N. punctiforme* secondary metabolite BGCs showing a spatially restricted expression pattern in short cell stretches or single cells within a filament under standard lowlight conditions. A) Average TPM value of the gene that was selected for the construction of the reporter plasmid in the RNA deep-sequencing time course experiment. Dashed grey line indicates the empirical threshold of 20 TPM for actively expressed genes; B) confocal fluorescence microscopy picture of the respective *N. punctiforme* reporter strain under standard low light cultivation and ambient air. Left image shows CFP channel only and right image CFP/ Chl $\alpha$  channels merged. Scale bar indicates 20  $\mu$ m.

The last category that could be described consists of the PKS4, anabaenopeptin, RiPP1a and RiPP1b transcriptional reporter. For these reporters the unique transcription pattern could be best described as CFP signal in single cells or small cell stretches within a filament (Figure 26). Noteworthy, the PKS4 reporter strain is the only mutant that showed a very regular signal pattern whereas the others again showed more heterogenous signal patterns.

### **3.1.5 Cell type specific expression of CFP in *N. punctiforme* reporter mutant strains**

The *N. punctiforme* reporter mutant strain library could also be used to examine in detail the presence or absence of a cell type specific expression of CFP. Indeed, in addition to the above-mentioned categories it could be observed that there were differences in the frequency at which certain cell types were found to give a CFP signal among the different reporter mutant strains. Strikingly, there was not a single case observed where a CFP signal in heterocysts could be detected and only a single reporter strain, namely the PKS5 reporter (Figure 27, M), showed hormogonial cells expressing CFP. An interesting tendency within the RiPP5 and RiPP6 reporter mutant strains (Figure 27, O and P) was observed as that especially stressed or amorphous vegetative cells showed a prominent CFP signal. For the NRPS1 and the Apt reporters (Figure 27, A and C) mainly pre-akinetes and akinetes were found to express CFP. However, besides these special cases the majority of reporter strains showed CFP signals only within the vegetative cell type (Figure 27).

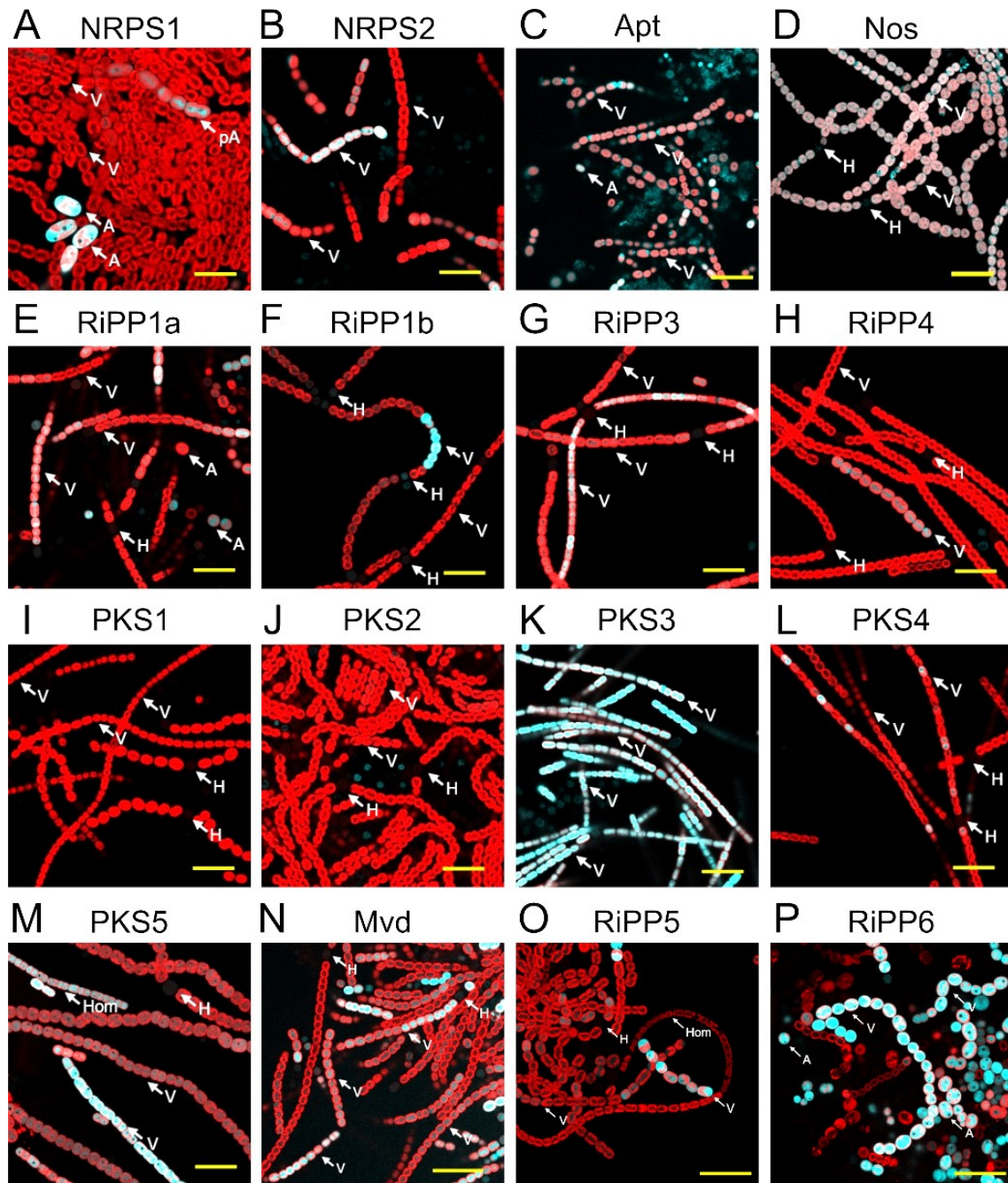


Figure 27 *N. punctiforme* CFP reporter strain fluorescence micrograph close-ups. Labeled are distinct cell types to highlight differences in cell type restricted CFP expression. Depicted are CFP-channel/Chl $\alpha$ -channel merged images. V: vegetative cells; H: heterocysts; Hom: hormogonia; pA: pre-akinetes; A: akinetes. Yellow scale bar indicates 20  $\mu$ m.

Taking the multitudes of specific signal patterns of this last category into account it seems very likely that *N. punctiforme* deploys a strong regulatory network that tightly controls the secondary metabolite production.

## **3.2 Influencing *N. punctiforme*'s secondary metabolite production**

It is already described in the literature that environmental conditions can have a tremendous impact on the secondary metabolome of *N. punctiforme* (71, 77). These factors can be divided into two sections: abiotic factors like light intensity or local carbon dioxide concentration and biotic factors like culture density or proximity of a potential host organism. To further investigate the underlying principles of the regulation of *N. punctiforme*'s secondary metabolome these biotic and abiotic factors were altered independently and the induced changes in transcriptional activity were monitored utilizing the reporter strain library as well as semi quantitative RT qPCR.

### **3.2.1 High density cultivation of *N. punctiforme***

As described by Guljamow and Kreische et al. 2017, the cultivation of *N. punctiforme* to very high cell densities has a strong impact on the produced secondary metabolites (77). In a cooperative study with Julia Krumbholz this observation should be assessed in more detail utilizing the reporter strain library and the CellDeg high-density cultivator. Figure 28 depicts the HPLC chromatograms of two aqueous *N. punctiforme* cell extractions, one cultivated under standard low light conditions and ambient air and the second one cultivated to high cell densities using the CellDeg HD-cultivator as described in section 2.2.1.2.

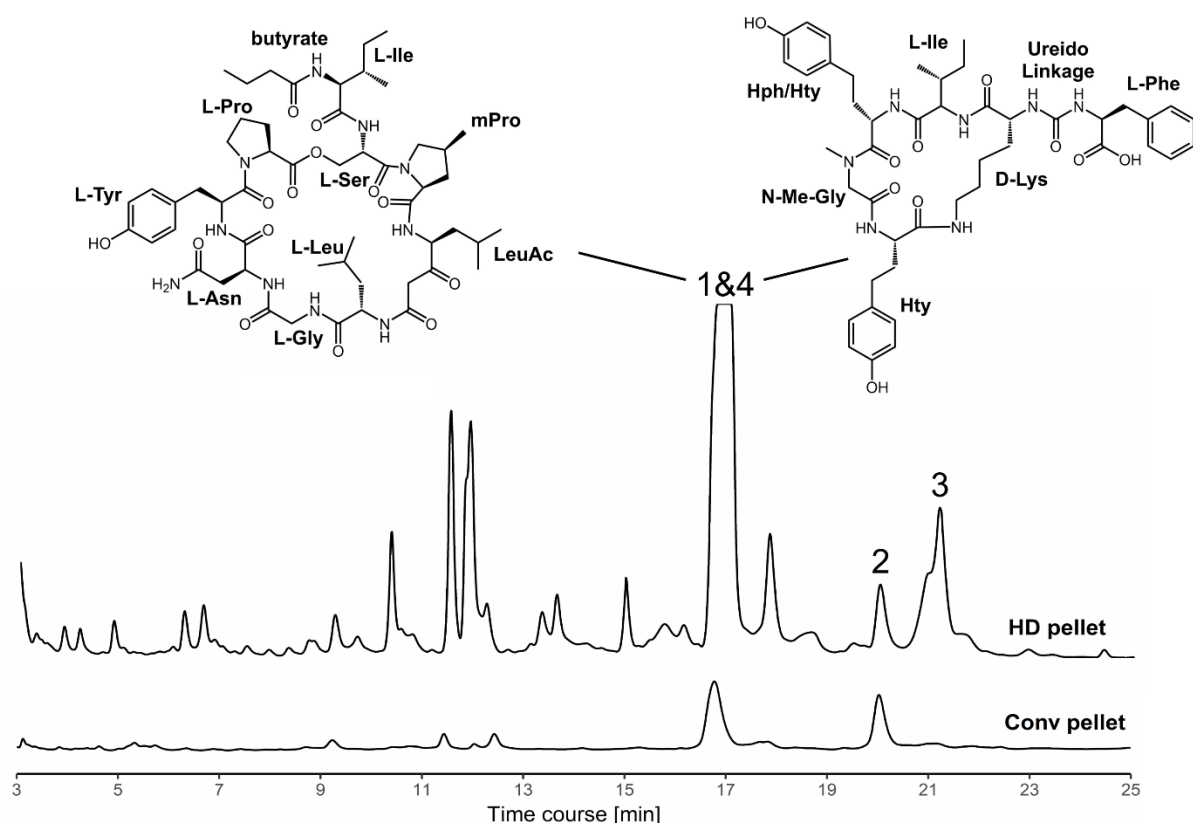
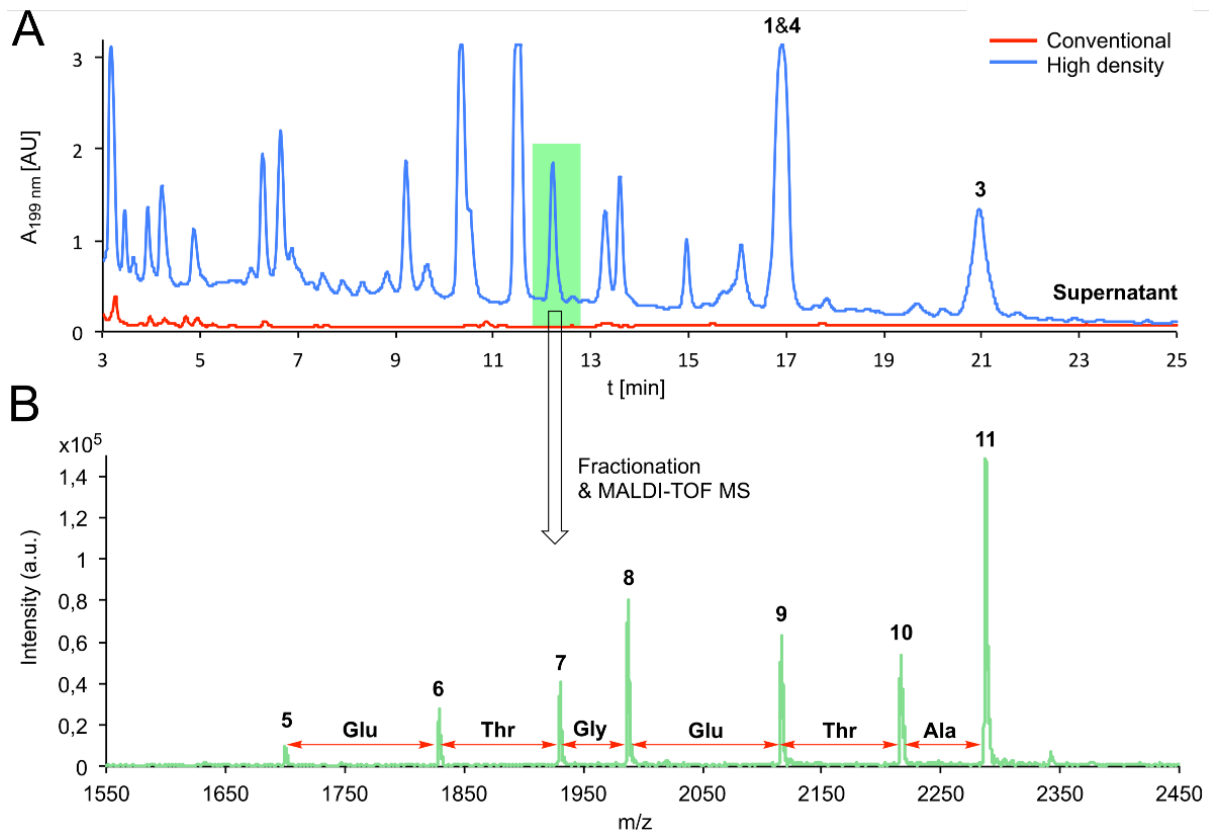


Figure 28 **Comparison of HPLC chromatograms of equal amounts of aqueous cell extracts generated from either conventionally cultivated or high density cultivated *N. punctiforme*.** Upper trace: cell extract of a 30 days high density cultivation of *N. punctiforme*; lower trace: cell extract of a 30 days conventional cultivation. 1: nostopeptolide1052, 2: nostopeptolideA, 3: nostamideA, 4: anabaenopeptin NZ857. Sample processing and HPLC analysis conducted as described in methods sections 2.2.5.5 and 2.2.5.6. Data provided by Julia Krumbholz.

Already described compounds including nostopeptolides and anabaenopeptins were produced in very high amounts when compared to conventional low light cultivation but more important a multitude of unknown peaks appeared pointing towards an overall increased metabolomic activity under high cell density conditions.

This unexpected high impact on *N. punctiforme* metabolic activity resembled a unique chance for the search of novel compounds. In a follow up cooperative study lead by Martin Baunach and Vincent Wiebach, high density cultivated *N. punctiforme* extracts were analyzed by different mass spectrometric techniques. Thereby, a cluster of masses was found in the high-density culture supernatant that was categorized as potential new RiPPs due to their respective masses of 1700 – 2287 Da. Since heavy post-translational modifications are unique traits of RiPPs the mass differences between single peaks of this prominent mass cluster were closer investigated. The analysis revealed that the difference of one peak to its neighbor corresponded to a specific proteinogenic amino acid (Figure 29).



**Figure 29 HPLC analysis of supernatant extracts of conventional and HD cultivated *N. punctiforme* cultures and consecutive MALDI-TOF mass spectrometry.** A: Comparative HPLC chromatograms of high density cultivated *N. punctiforme* and a control culture cultivated under standard low light conditions and ambient air. Increased production of known products nostopeptolide1052 (1), nostamideA (3) and anabaenopeptin NZ857 (4) and various undescribed compounds. B: MALDI-TOF mass spectrometric analysis of an HPLC fraction containing the undescribed peak highlighted in green. 5-11: potential RiPP products and theoretical mass differences correlated to distinct proteinogenic amino acids. Data provided by Julia Krumbholz, Martin Baunach and Vincent Wiebach.

However, the identified amino acid sequence ETGETA (or ATEGTE) could not be related to any of the putative RiPP core sequences predicted in the *N. punctiforme* genome. Surprisingly, after expanding the search frame to also include the leader peptide sequences, the sequence ATEGTE was found to be part of the predicted leader peptide sequence of the microviridin precursor peptide located right upstream of the N-terminus of the core peptide. Taking the typical intramolecular cyclizations into account the various compound masses could theoretically be fit to a set of unusual 15- to 20-membered novel microviridins. These compounds could finally be confirmed by combination of MALDI-TOF/TOF analysis and chemical transformation with the missing acetylation being in good agreement with the predicted microviridin cluster constitution also lacking an acetyltransferase. To emphasize the unique elongated peptide structure the novel microviridins were named microviridin N3-N9 corresponding to the increased number of N-terminal amino acids of up to nine (Figure 30).

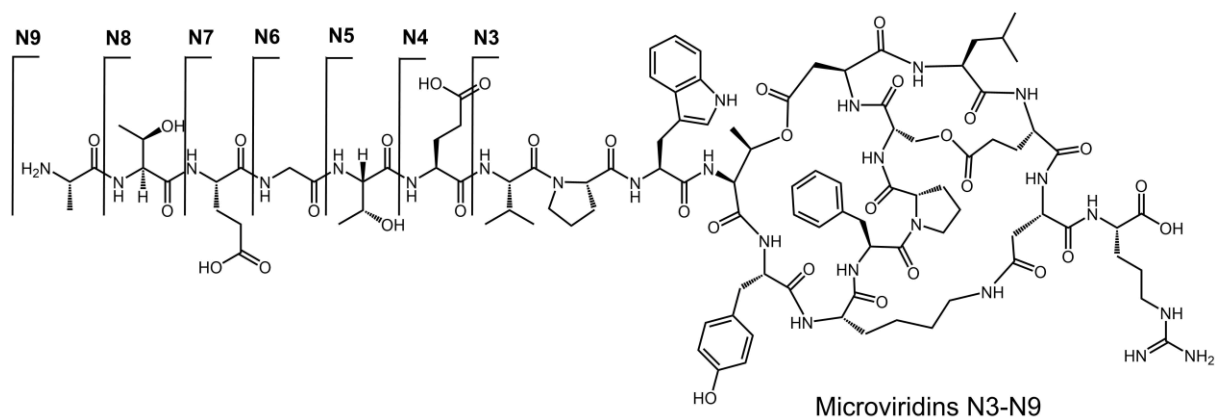


Figure 30 **Structure of novel microviridins N3-N9.** The 11-membered core peptide sequence is N-terminally extended by up to 7 amino acids (WPV-E-T-G-E-T-A) leading to the final sequence of microviridins N3 to N9. Structure confirmation achieved by various means of mass spectrometric analyses in combination with chemical transformation. Analyses conducted by Vincent Wiebach and Matrin Baunach. Detailed information summarized in Dehm et al 2019 (80).

To assess whether high cell-density cultivation further impacts the activity of other predicted secondary metabolite BGCs, an RT qPCR study was conducted comparing the relative transcription levels of selected secondary metabolite BGC genes of a high density cultivated culture and a conventional culture of *N. punctiforme*.

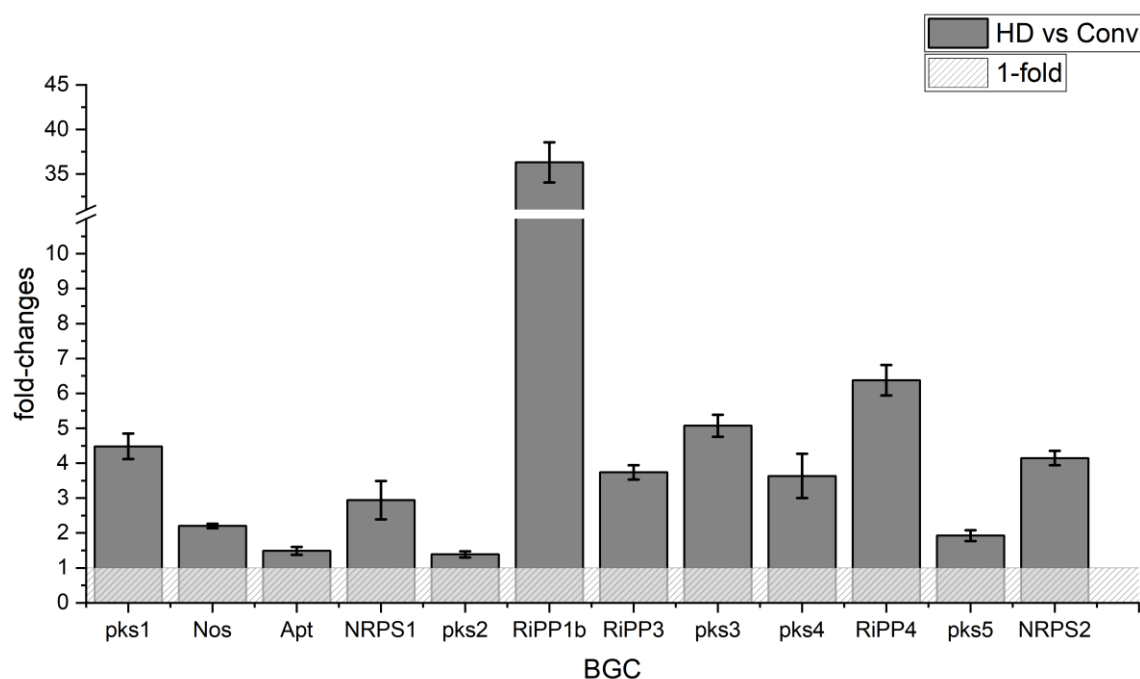
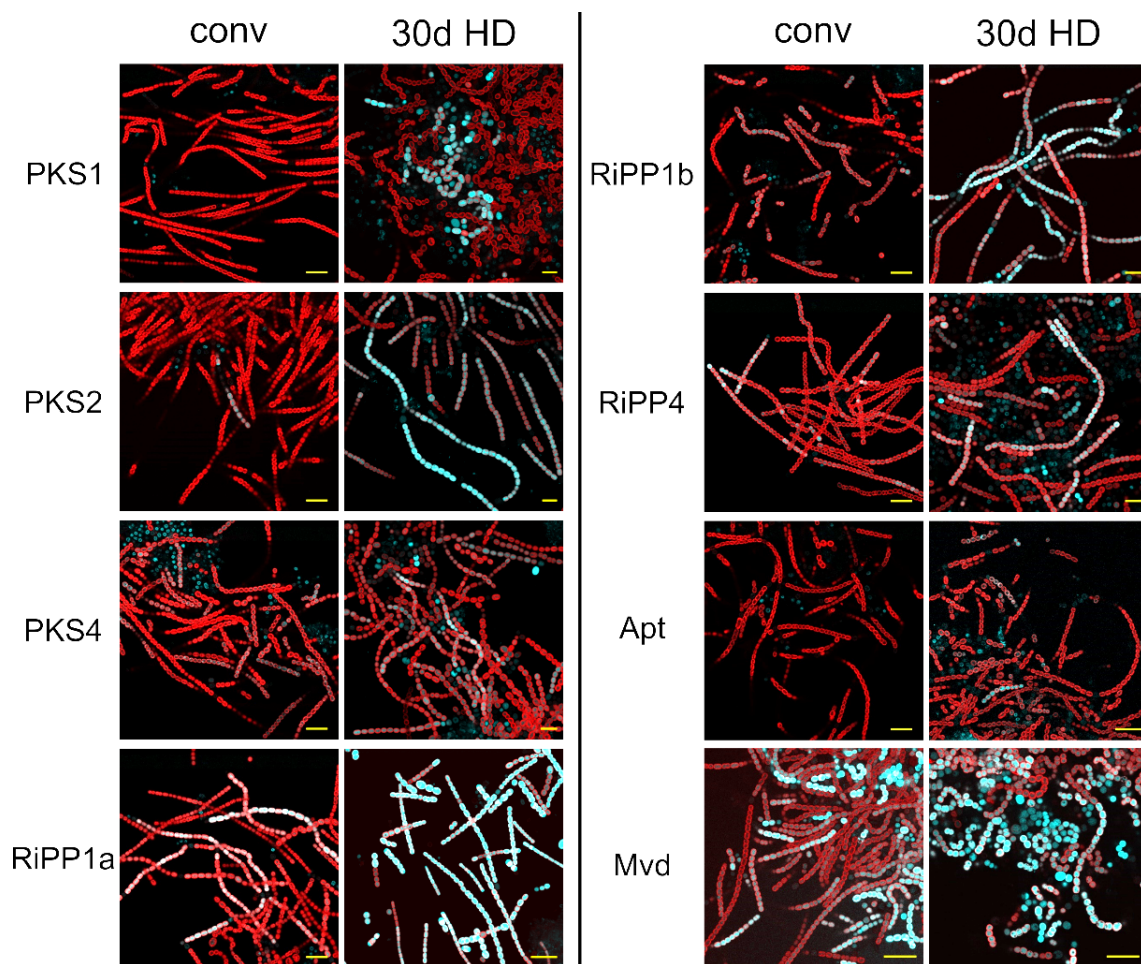


Figure 31 **Relative transcription levels of selected *N. punctiforme* secondary metabolite BGC genes of a high density cultivated culture in comparison to a conventionally cultivated culture.** Grey area highlights the fold-change range of 1 – 0, indicating a downregulation of the respective gene in comparison to the control. RNA extraction, qPCR sample preparation, RT primers and calculation of fold-changes with the  $\Delta\Delta C_t$ -method are described in section 2.2.4.7. High density culture sample provided by Julia Krumbholz.

Interestingly, all the representative genes of the secondary metabolite BGCs showed positive fold-changes between 1.5 and 7 with one prominent exception showing a fold-change of 36 (Figure 31). However, stronger upregulation with fold-change values  $> 4$  could be observed for the BGCs PKS1, PKS3, RiPP4, NRPS2 and RiPP1b. These results indicated that in general expression of secondary metabolite BGCs got upregulated when the culture was cultivated to very high cell densities. To confirm this observation with an independent approach also the *N. punctiforme* reporter strains were cultivated in CellDeg culture vessels to very high cell densities using the CellDeg HD-cultivator. After 30 days samples were examined by confocal fluorescence microscopy to assess differences in CFP expression compared to respective standard low light cultivated control samples.



**Figure 32 Confocal fluorescence micrographs of high density cultivated and conventionally cultivated *N. punctiforme* reporter cultures.** Depicted are merged CFP-channel/Chl $\alpha$ -channel images. Both culture types were cultivated for 30 days and samples were examined using confocal fluorescence microscopy as described in section 2.2.6.1. Experiment conducted in collaboration with Julia Krumbholz. Yellow scalebar indicates 20  $\mu$ m.

For eight secondary metabolite BGCs a pronounced increase in CFP signal after 30 days of high-density cultivation could be observed (Figure 32). Unexpectedly, even PKS1 and



PKS2 could be induced by this cultivation method, two BGCs that were found silent under standard low light conditions and ambient air. The strongest differential CFP signal was observed for Mvd, RiPP1a and RiPP1b. Furthermore, the induction of PKS1 and Mvd BGC leads to atypical cell shapes that might hint towards increased cell stress.

To tackle the question whether a secreted factor might induce the upregulation of these secondary metabolite BGCs during high cell density cultivation, conventional cultures of those reporter strains that showed a positive response towards high density cultivation were treated with cell free supernatant of a 30 days old high density *N. punctiforme* wild-type culture for 7 days. After the incubation time samples were taken and CFP levels were estimated using confocal fluorescence microscopy.

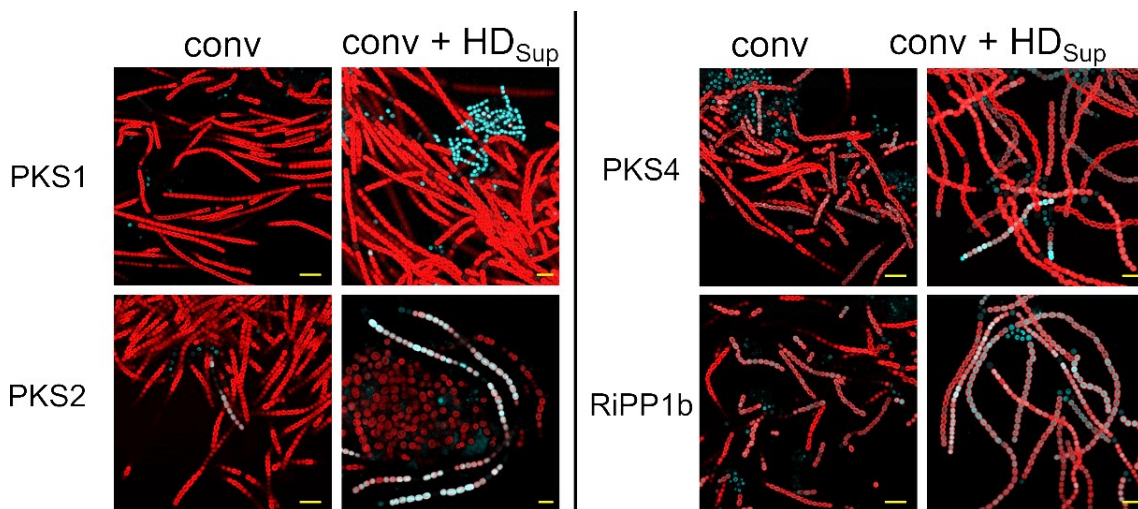


Figure 33 **Confocal fluorescence micrographs of high density supernatant treated and conventionally cultivated *N. punctiforme* reporter cultures.** Depicted are merged CFP-channel/Chl $\alpha$ -channel images. The treated cultures were cultivated for 7 days and samples were examined using confocal fluorescence microscopy as described in in section 2.2.6.1. Experiment conducted in collaboration with Julia Krumbholz. Yellow scalebar indicates 20  $\mu$ m.

From the eight secondary metabolite BGCs that showed a positive response in high density cultivation four could also be induced by treatment with cell free high density culture supernatant (experimental design and execution by Julia Krumbholz). Despite the response being less prominent no other induction pattern compared to the high density cultivated reporter strains could be observed.

Cultivation of the various *N. punctiforme* reporter strains under high density conditions brought further information about which specific secondary metabolite BGCs contribute to the overall increase in metabolic activity observed. This additional information should be utilized in a targeted approach discover these specific products. Therefore, all differential peaks in HD cultivated culture extracts were analyzed by various means of

mass spectrometry (experimental design and execution by Julia Krumbholz). Unfortunately, only the newly discovered microviridins could be detected and no other mass correlated with any of the predicted products of the upregulated secondary metabolite BGCs. This might point towards issues connected with the complexity of the extracts or unanticipated inaccuracies in the prediction of potential products.

Since high density cultivation includes increased CO<sub>2</sub> availability and higher light intensities for the photosynthetic organism during cultivation, these abiotic factors should also be investigated. To test for the impact of a higher CO<sub>2</sub> partial pressure aliquots of the high density responsive *N. punctiforme* reporter strains were cultivated with dialysis bags filled KHCO<sub>3</sub>/K<sub>2</sub>CO<sub>3</sub> buffer that produces a 32 mbar CO<sub>2</sub> partial pressure over the solution (as described in section 2.2.1.1) for 3 days. This experimental setting mimicked the CO<sub>2</sub> condition during high-density cultivation without changing any other factor of standard low light condition. Interestingly the mere increase of CO<sub>2</sub> levels did not induce elevated expression levels in seven out of the eight high density responsive reporters. However, for the RiPP4 reporter strain a pronounced increase in CFP signal could be observed (Figure 34).

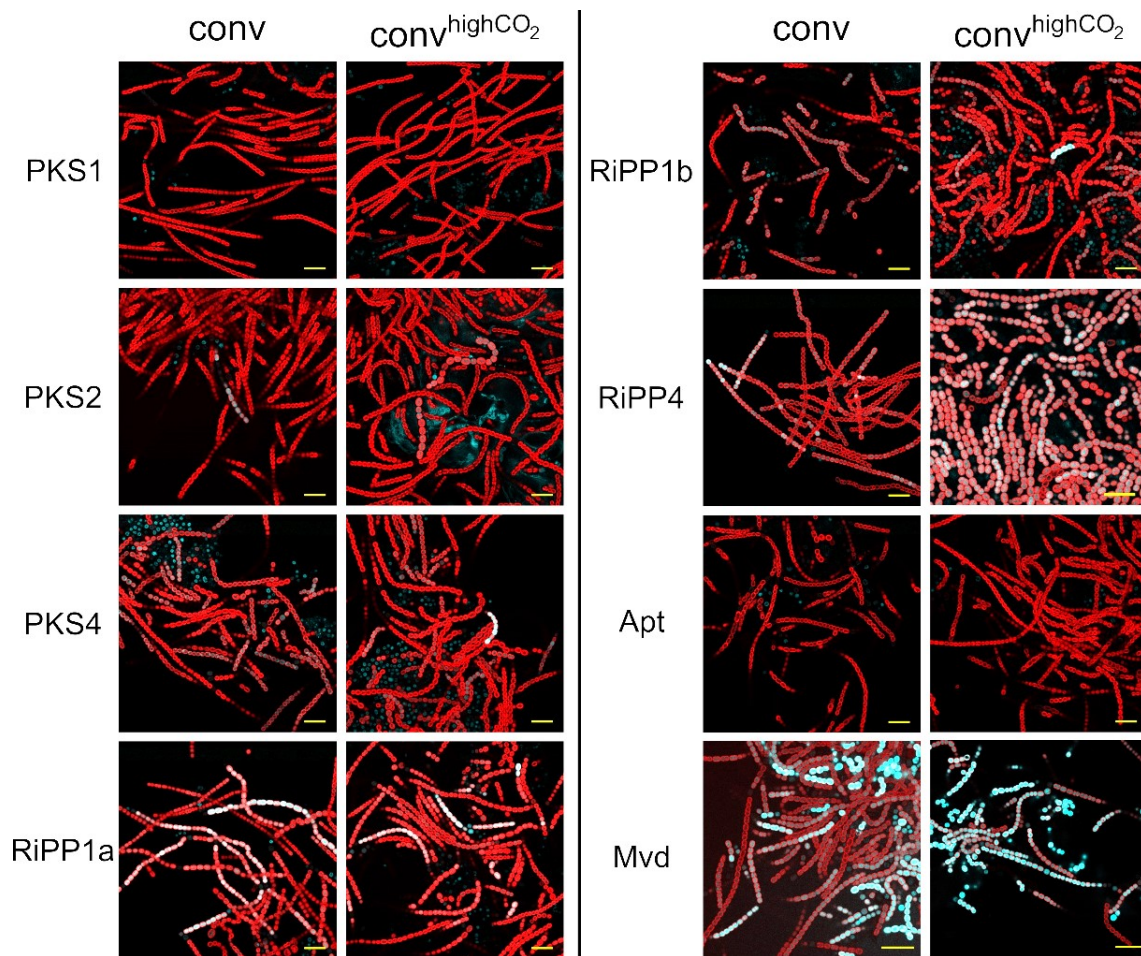


Figure 34 **Confocal fluorescence micrographs of high density responsive *N. punctiforme* reporter strains cultivated with  $\text{KHCO}_3/\text{K}_2\text{CO}_3$  buffer filled dialysis bags to induce high  $\text{CO}_2$  availability.** Depicted are merged CFP-channel/ $\text{Chl}\alpha$ -channel images. The treated cultures were cultivated for 3 days and samples were examined using confocal fluorescence microscopy as described in section 2.2.6.1. Yellow scalebar indicates 20  $\mu\text{m}$ .

The second abiotic factor that was altered independently was the illumination light intensity during cultivation. To mimic high density cultivation light conditions aliquots of the high density responsive *N. punctiforme* reporter strains were incubated for 24 h at 107  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  and afterwards CFP signal levels were determined using confocal fluorescence microscopy. Increased illumination light intensities had an inducing effect on the PKS2, RiPP1a and Mvd reporter strains but additionally a stress response could be observed for the RiPP1a and Mvd reporter cultures, since an increased number of shortened and amorphous filaments could be detected (Figure 35).

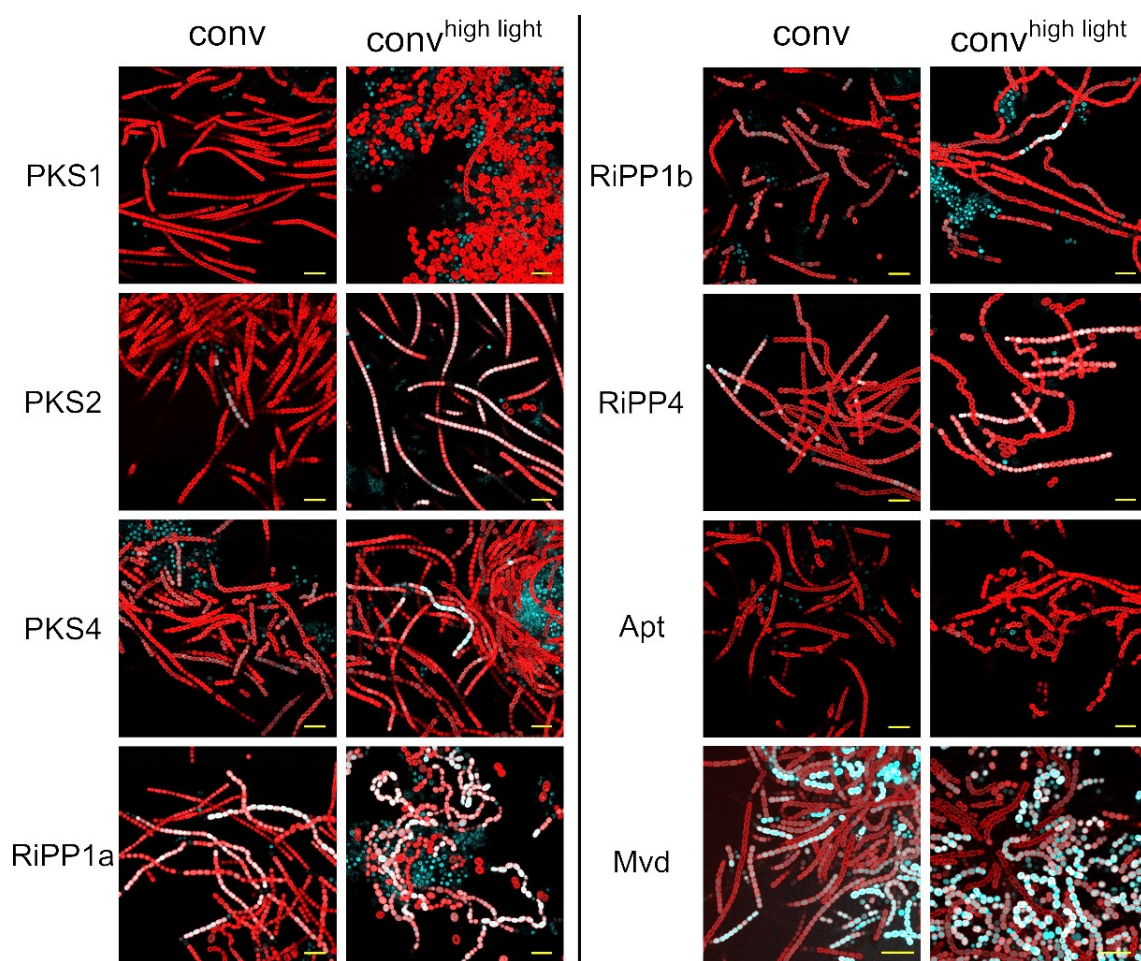


Figure 35 Confocal fluorescence micrographs of high density responsive *N. punctiforme* reporter strains cultivated under high illumination light intensity of  $107 \mu\text{mol photons m}^{-2} \text{s}^{-1}$ . Depicted are merged CFP-channel/Chl $\alpha$ -channel images. The treated cultures were cultivated for 24 h and samples were examined using confocal fluorescence microscopy as described in in section 2.2.6.1. Yellow scalebar indicates 20  $\mu\text{m}$ .

### 3.2.2 Cyanobiont – host plant interaction studies of *N. punctiforme* and *Blasia pusilla*

Besides the observation that very high cell densities have a strong impact on the secondary metabolome of *N. punctiforme* it was already shown that also the interaction with a potential host plant induces a metabolic reprogramming inside the symbiotic *N. punctiforme* (71). In order to investigate the genetic response of *N. punctiforme* towards the presence of a potential host plant a series of experiments were conducted and *N. punctiforme* was either exposed to cell free *B. pusilla* media exudates, resembling a mere chemical interaction, or to *B. pusilla* cultures, resembling a true physical interaction.

#### 3.2.2.1 *N. punctiforme* – *B. pusilla* chemical interaction

Since the initial contact between the cyanobiont and its host is loose and without physical contact, first the chemical interaction of *N. punctiforme* with *B. pusilla* was investigated. Therefore, a cell free conditioned *B. pusilla* exudate was produced as described in section

2.2.6.2. Preliminary studies confirmed that a total of 50% culture volume *B. pusilla* exudate is sufficient to trigger a genetic response within *N. punctiforme* (71). To diminish medium exchange effects the conditioned N-deprived host plant exudate was concentrated to reduce the total volume that had to be added to any given sample culture. A schematic representation of the conducted experimental workflow is depicted in Figure 36.

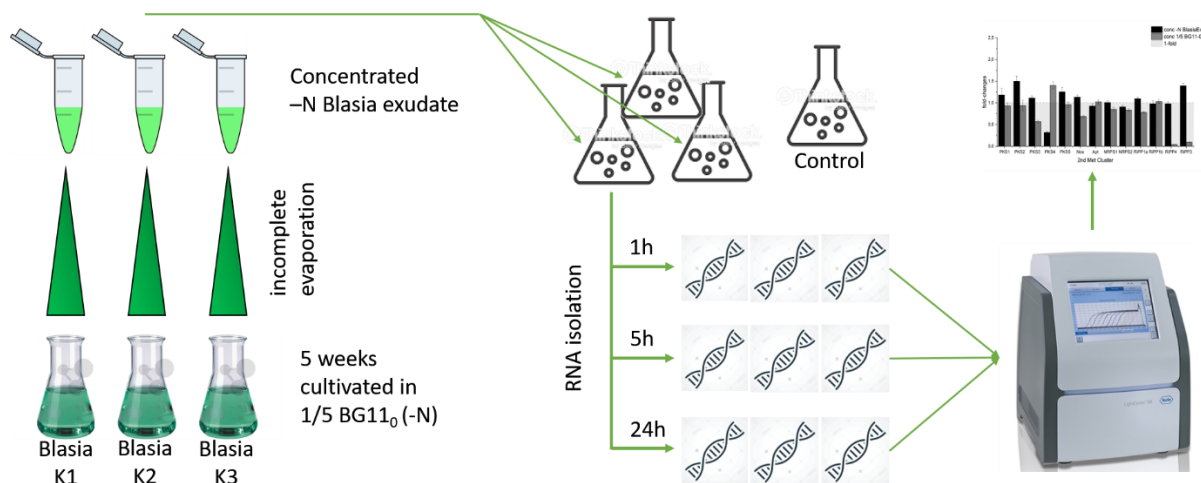


Figure 36 **Schematic overview of the *N. punctiforme* - *Blasia pusilla* chemical interaction experimental workflow.** For a detailed description see section 2.2.6.3.

To monitor the transcriptomic response of *N. punctiforme* to effective concentrations of conditioned *B. pusilla* exudate a time course experiment was conducted including the determination of relative transcription levels via RT qPCR. For that purpose, aliquots of *N. punctiforme* wild type cultures were supplemented with concentrated conditioned *B. pusilla* exudate (final concentration of 50% (v/v) total culture volume) and incubated for 24 h under standard low light conditions and ambient air. After 1 h, 5 h and 24 h samples were taken and RNA was isolated and prepared for RT qPCR as described in section 2.2.4.2. To investigate the influence on secondary metabolite BGC expression, the relative transcription levels of representative genes of given BGCs in comparison to an untreated control culture of *N. punctiforme* were determined and fold-changes were calculated (as described in section 2.2.4.7). The time course experiment was run in biological triplicates to rule out any variances during culture preparation or sample processing.

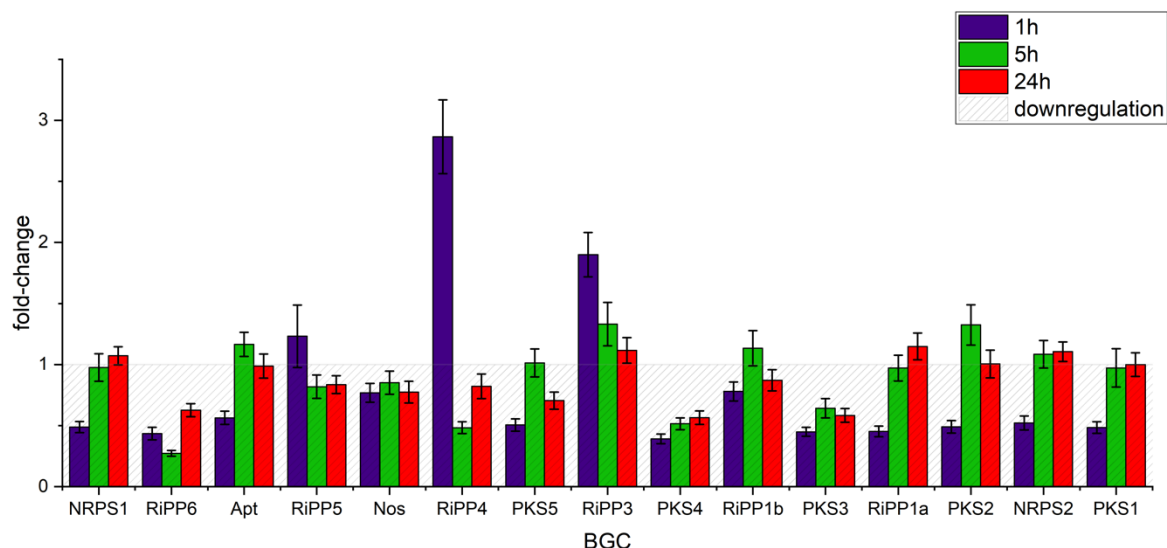


Figure 37 Average fold-changes of three biological replicates of *N. punctiforme* treated with conditioned *B. pusilla* exudate for 1 h, 5h and 24 h relative to an untreated control culture. Purple bars: average fold-change of three biological replicates after 1 h of incubation with *B. pusilla* exudate; green: average fold-change of three biological replicates after 5 h of incubation with *B. pusilla* exudate; red: average fold-change of three biological replicates after 24 h of incubation with *B. pusilla* exudate. Grey area highlights the fold-change range of 1 – 0, indicating a downregulation of transcription compared to the control sample. RNA isolation and RT qPCR preparation conducted as described in sections 2.2.4.2 to 0.

The impact on transcriptional activity was most intense at the 1 h time point. All secondary metabolite BGCs were downregulated to varying degrees except RiPP3, RiPP4 and RiPP5, with RiPP4 showing a comparatively strong upregulation of  $2.87 \pm 0.3$  (Figure 37). After 5 h transcription levels were less differential and nine BGCs were close to a 1-fold transcription compared to the untreated control. However, PKS3, PKS4 and RiPP6 still were strongly downregulated with RiPP6 showing a minimum of  $0.27 \pm 0.02$  (Figure 37). Interestingly, RiPP5 and RiPP3 settled close below and above unchanged transcription, respectively, RiPP4 showed a fully inversed transcriptional behavior and was downregulated at  $0.48 \pm 0.05$  at the 5 h time point. After 24 h, all BGCs were back in the range of  $1 \pm 0.2$ -fold compared to the untreated control except PKS3, PKS4, PKS5 and RiPP6 that remained downregulated to approx. 0.6-fold (Figure 37).

Taking all the results of the different BGCs and sampling time points together it could be shown that the mere presence of a potential host plant exudate is sufficient to trigger a genomic response within *N. punctiforme*. Although in many cases the initial response was a downregulation of the expression of a given BGC the response of RiPP4 points towards a dynamic regulation process that can be triggered by the host plant. Thus, in the next step the physical interaction of *N. punctiforme* with *B. pusilla* and its impact on the secondary metabolome should be investigated.

### 3.2.2.2 *N. punctiforme* – *B. pusilla* physical interaction

Later stages of the establishment of a cyanobacterial – plant symbiosis involve physical contact of the cyanobiont with its host. Since a chemical interaction was already shown to have a considerable impact on the transcription of different secondary metabolite BGCs in the course of this study the effects of a truly physical interaction should be examined. Thus, a similar experimental set-up as for the chemical interaction was designed with the difference that instead of the addition of a conditioned host plant exudate *N. punctiforme* was directly co-cultivated with *B. pusilla* (as described in section 2.2.6.4).

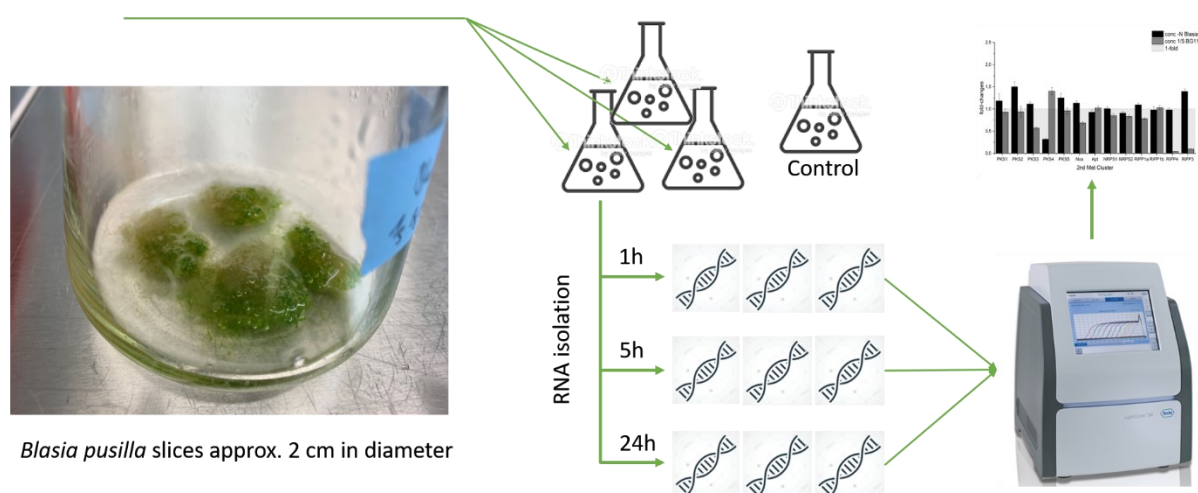


Figure 38 **Schematic overview of the *N. punctiforme* – *Blasia pusilla* physical interaction experimental workflow.** For a detailed description see section 2.2.6.4.

Again, sampling was conducted in fixed periods of 1 h, 5 h and 24 h and afterwards RNA was isolated and prepared for RT qPCR mediated determination of transcription levels of representative genes of *N. punctiforme*'s secondary metabolite BGCs. Similar to the first study the full experiment was run in three biological replicates (Figure 38).

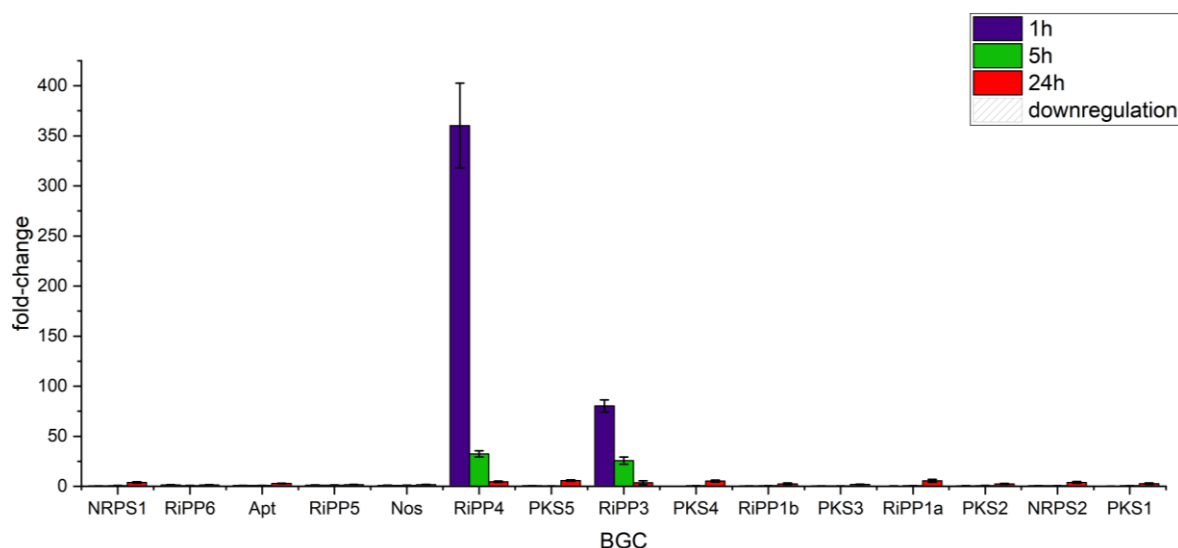


Figure 39 Average fold-changes of three biological replicates of *N. punctiforme* co-cultivated with *B. pusilla* for 1 h, 5h and 24 h relative to a control culture. Purple bars: average fold-change of three biological replicates after 1 h of incubation with *B. pusilla* exudate; green: average fold-change of three biological replicates after 5 h of incubation with *B. pusilla* exudate; red: average fold-change of three biological replicates after 24 h of incubation with *B. pusilla* exudate. RNA isolation and RT qPCR preparation conducted as described in sections 2.2.4.2 to 0.

In the co-cultivation study again the BGCs RiPP4 and RiPP3 showed a positive transcriptional response after 1 h. Strikingly, this upregulation after 1 h of RiPP4 ( $360.42 \pm 42.36$ ) and RiPP3 ( $80.27 \pm 6.00$ ) triggered by co-cultivation with *B. pusilla* showed dimensions of induction 125 times and 40 times higher than those observed in the chemical interaction study. Although the trend of declining transcriptional activity after 5 h could also be observed in the co-cultivation study, the transcriptional induction was still 68 times higher for RiPP4 ( $32.47 \pm 3.06$ ) and 19 times higher for RiPP3 ( $25.68 \pm 3.52$ ) after 5 h (Figure 37, Figure 39).



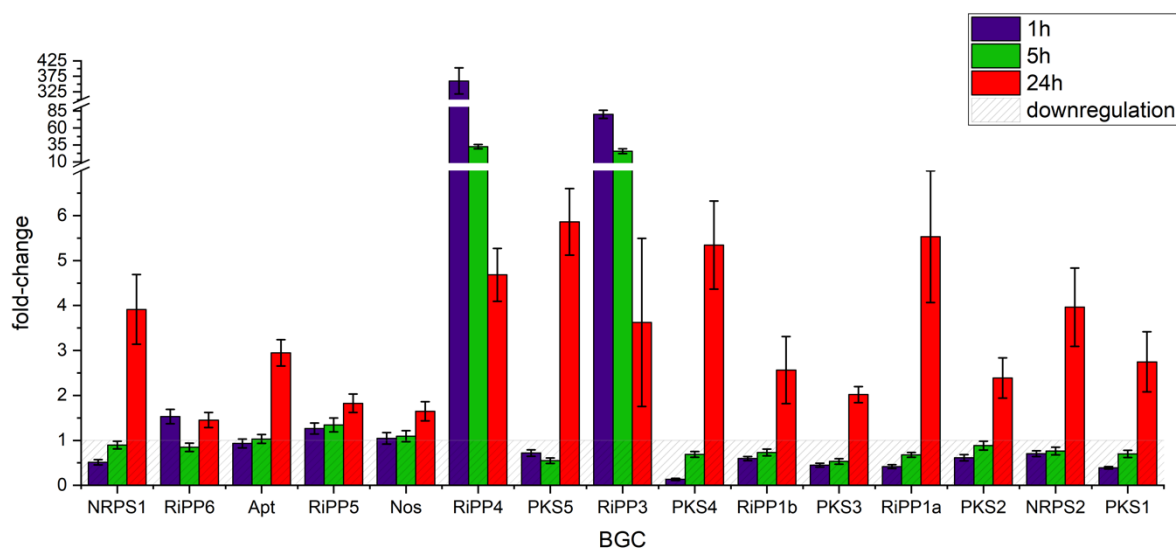


Figure 40 **Average fold-changes of three biological replicates of *N. punctiforme* co-cultivated with *B. pusilla* for 1 h, 5h and 24 h relative to a control culture.** Purple bars: average fold-change of three biological replicates after 1 h of incubation with *B. pusilla* exudate; green: average fold-change of three biological replicates after 5 h of incubation with *B. pusilla* exudate; red: average fold-change of three biological replicates after 24 h of incubation with *B. pusilla* exudate. Grey area highlights the fold-change range of 1 – 0, indicating a downregulation of transcription compared to the control sample. RNA isolation and RT qPCR preparation conducted as described in sections 2.2.4.2 to 0.

For the other BGCs investigated a differential behavior could be observed (Figure 40). Whereas after 1 h Nos, Apt and RiPP5 showed an unchanged transcription compared to the control culture NRPS1, PKS5, PKS4, RiPP1b, PKS3, RiPP1a, PKS2, NRPS2 and PKS1 were downregulated. Interestingly, compared to the 1 h time point in the chemical interaction study the PKS4 BGC was even stronger downregulated at  $0.13 \pm 0.02$ . After 5 h of co-cultivation no differential behavior compared to the 1 h sampling time point could be observed except for the PKS5 BGC that showed a further decrease in transcriptional activity. However, after 5 h in general the transcription levels were higher for all BGCs downregulated after 1 h (Figure 40). After 24 h of co-cultivation a general upregulation of transcriptional activity could be observed and merely the BGCs RiPP6, RiPP5 and Nos showed a fold-change of less than 2.0. For those BGCs initially downregulated PKS5, PKS4 and RiPP1a showed the strongest differential transcriptional activity after 24 h of co-cultivation. For the two super-responsive BGCs RiPP4 and RiPP3 fold-changes of only  $4.68 \pm 0.60$  and  $3.62 \pm 1.87$ , respectively, compared to the untreated control could be observed (Figure 40).

Since both chemical and physical interaction with *B. pusilla* induced the transcription of RiPP3 and RiPP4 in a follow-up study these observed effects should be confirmed utilizing the respective CFP reporter mutant strains.

### 3.2.2.3 *N. punctiforme* CFP reporter mutant strains in chemical and physical interaction with *B. pusilla*

As shown in the previous studies, exposure of *N. punctiforme* to either conditioned *B. pusilla* exudate or *B. pusilla* had tremendous effects on the transcriptional behavior of certain secondary metabolite BGCs. The greatest positive effects could be observed for the RiPP3 and RiPP4 BGCs. Thus, there was a follow-up study designed to confirm the results of the RT qPCR study utilizing the respective *N. punctiforme* reporter mutant strains. Therefore, the reporter mutants for the BGCs RiPP3, RiPP4 and PKS5 (non-responsive control) were co-cultivated with either conditioned *B. pusilla* exudate or *B. pusilla* under standard low light conditions and ambient air for up to 24 h. After 1.5 h, 6 h and 24 h samples were taken and examined via confocal fluorescence microscopy to track the transcriptional activity of the respective BGCs by determination of the CFP expression levels within the cultures (as described in section 2.2.6.1).

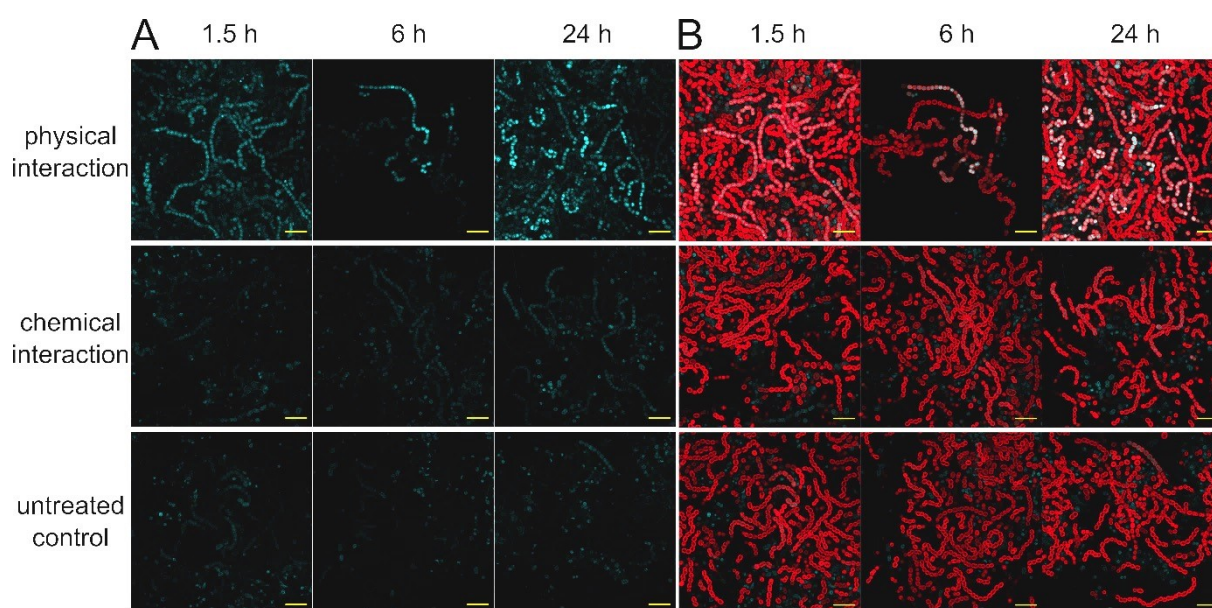


Figure 41 *N. punctiforme* RiPP4 CFP reporter mutant strain co-cultivated with *B. pusilla* or *B. pusilla* conditioned exudate for up to 24 h in comparison to an untreated control. Physical interaction: co-cultivation with *B. pusilla* under standard low light conditions and ambient air; chemical interaction: co-cultivation with concentrated conditioned *B. pusilla* exudate (final concentration of 50% total culture volume); untreated control: *N. punctiforme* RiPP4 CFP reporter mutant cultivated under standard low light conditions and ambient air. A: confocal fluorescence micrographs depicting the signal captured in CFP channel after 1.5 h, 6 h and 24 h of cultivation; B: confocal fluorescence micrographs depicting the merged signals captured in CFP and Chl *a* channel after 1.5 h, 6 h and 24 h of cultivation. Yellow scalebar indicates 20  $\mu\text{m}$ .

The *N. punctiforme* reporter mutant for RiPP4 showed a CFP signal increase after 1.5 h of physical interaction with *B. pusilla* (Figure 41, A) and thus confirmed the RT qPCR results (Figure 39) of the previous study. Furthermore, the CFP signal increased over the course of the experiment and peaked at 24 h. Although this inducing effect could also be observed

while in chemical interaction in the RT qPCR study (Figure 37) the RiPP4 reporter mutant cultivated with conditioned *B. pusilla* exudate showed no increased CFP signal after 1.5h. However, after 6 h a very weak signal increase could be observed that was slightly more prominent after 24h (Figure 41). This observation resembles the weaker transcriptional induction of the RiPP4 BGC observed during chemical interaction compared to physical interaction. Overall, the cultures were healthy and showed no differential phenotype compared to the untreated control while maintaining the spatially restricted CFP patterns observed during standard low light cultivation (Figure 41, B).

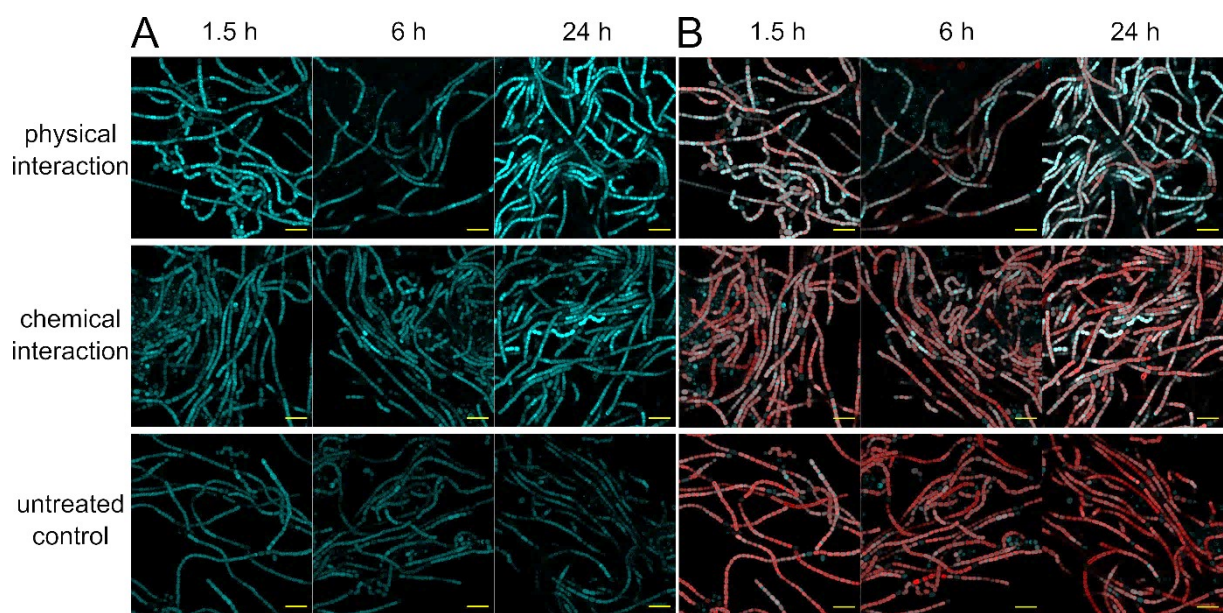


Figure 42 *N. punctiforme* RiPP3 CFP reporter mutant strain co-cultivated with *B. pusilla* or *B. pusilla* conditioned exudate for up to 24 h in comparison to an untreated control. Physical interaction: co-cultivation with *B. pusilla* under standard low light conditions and ambient air; chemical interaction: co-cultivation with concentrated conditioned *B. pusilla* exudate (final concentration of 50% total culture volume); untreated control: *N. punctiforme* RiPP4 CFP reporter mutant cultivated under standard low light conditions and ambient air. A: confocal fluorescence micrographs depicting the signal captured in CFP channel after 1.5 h, 6 h and 24 h of cultivation; B: confocal fluorescence micrographs depicting the merged signals captured in CFP and Chl $\alpha$  channel after 1.5 h, 6 h and 24 h of cultivation. Yellow scalebar indicates 20  $\mu$ m.

Interestingly, the degree of transcriptional induction observed in the RT qPCR study was overall lower for the RiPP3 BGC in both chemical and physical interaction (Figure 37, Figure 40) but in the reporter mutant experiment on the other hand the differential CFP signals detected were much higher. The RiPP3 CFP reporter showed intense CFP signal increase already after 1.5h of either physical or chemical interaction (Figure 42, A), furthermore the intensity of the detected signals increase over the course of 24 h in both interaction types and was strongest after 24 h. However, there could again be a difference observed between chemical and physical interaction with *B. pusilla* as the overall CFP signal intensity was stronger in the physical interaction culture at all sampling time points

(Figure 42, A). The auto-fluorescence (Figure 42, B) proved that the cultures were in a healthy state and no differential morphology could be observed compared to the untreated control.

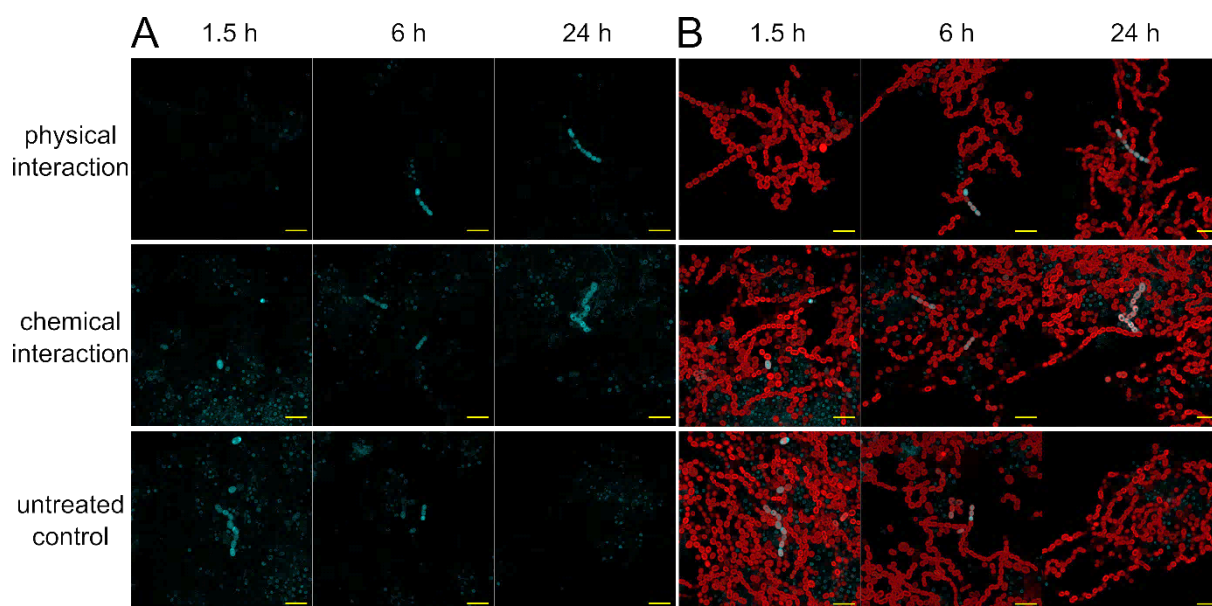


Figure 43 *N. punctiforme* PKS5 CFP reporter mutant strain co-cultivated with *B. pusilla* or *B. pusilla* conditioned exudate for up to 24 h in comparison to an untreated control. Physical interaction: co-cultivation with *B. pusilla* under standard low light conditions and ambient air; chemical interaction: co-cultivation with concentrated conditioned *B. pusilla* exudate (final concentration of 50% total culture volume); untreated control: *N. punctiforme* RiPP4 CFP reporter mutant cultivated under standard low light conditions and ambient air. A: confocal fluorescence micrographs depicting the signal captured in CFP channel after 1.5 h, 6 h and 24 h of cultivation; B: confocal fluorescence micrographs depicting the merged signals captured in CFP and Chl $\alpha$  channel after 1.5 h, 6 h and 24 h of cultivation. Yellow scalebar indicates 20  $\mu$ m.

As a non-responsive control the PKS5 CFP reporter strain was also cultivated in chemical or physical contact with *B. pusilla*. According to the expectation neither of the two co-cultivation approaches showed an impact on the CFP levels expressed inside the cultures (Figure 43, A). This confirmed the negative/neutral transcriptional response of the PKS5 BGC observed during the RT qPCR time course experiments (Figure 37, Figure 40). However, the auto-fluorescence depicted in the Chl $\alpha$ -channel (Figure 43, B) again showed that the mode of cultivation in both cases had no negative impact on the culture morphology observed.

#### 3.2.2.4 HPLC and mass spectrometric analysis of *N. punctiforme* wild-type cultures in physical contact with *B. pusilla*

Based on the promising results of both the RT qPCR study as well as the CFP reporter study of *N. punctiforme* co-cultivated with *B. pusilla* as a follow-up approach an experiment was designed to assign compounds to both the RiPP3 and RiPP4 secondary

metabolite BGCs. Therefore *N. punctiforme* wild-type cultures were co-cultivated with *B. pusilla* and cells were harvested after 2, 6 and 24 h and both culture supernatant and cell extracts were analyzed via HPLC to monitor the peptide spectrum produced within the sample. Differential peptide peaks not occurring in the respective *N. punctiforme* control culture without *B. pusilla* were then fractionated and analyzed via MALDI-TOF mass spectrometry to identify the main mass occurring in the isolated peak fraction.

In the first step of the experiment the *N. punctiforme* co-cultivation samples and control samples were harvested at the respective time points. Culture supernatant and cells extracts were prepared for HPLC and analyzed as described in sections 2.2.5.5 and 2.2.5.6.

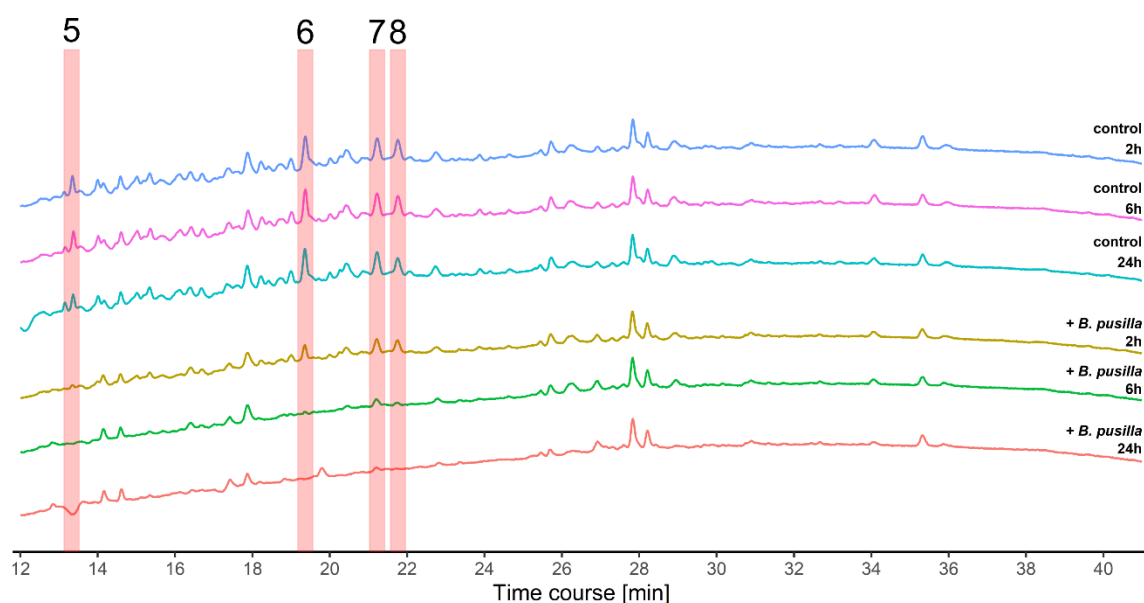


Figure 44 **HPLC traces of culture supernatant samples of *N. punctiforme* with and without *B. pusilla* in co-cultivation.** Samples were harvested after 2, 6 and 24 h and prepared for HPLC analysis as described in section 2.2.5.5. HPLC program including the acetonitrile gradient applied as described in section 2.2.5.6.

The culture supernatant samples showed a diverse spectrum of peaks with comparatively low overall intensities (Figure 44). Within the control and the sample group no prominent signal accumulation of a given peak could be observed over the course of 24 h. However, comparing the co-cultivation samples with the *N. punctiforme* control culture some minor differences in the peak spectrum could be observed. Certain peaks that were present in the 2 h co-cultivation sample disappeared over the course of 24 h (6, 7, 8), some peaks could not be observed in the sample culture at all (5) but the most of the observed peaks were also found in the control (Figure 44).

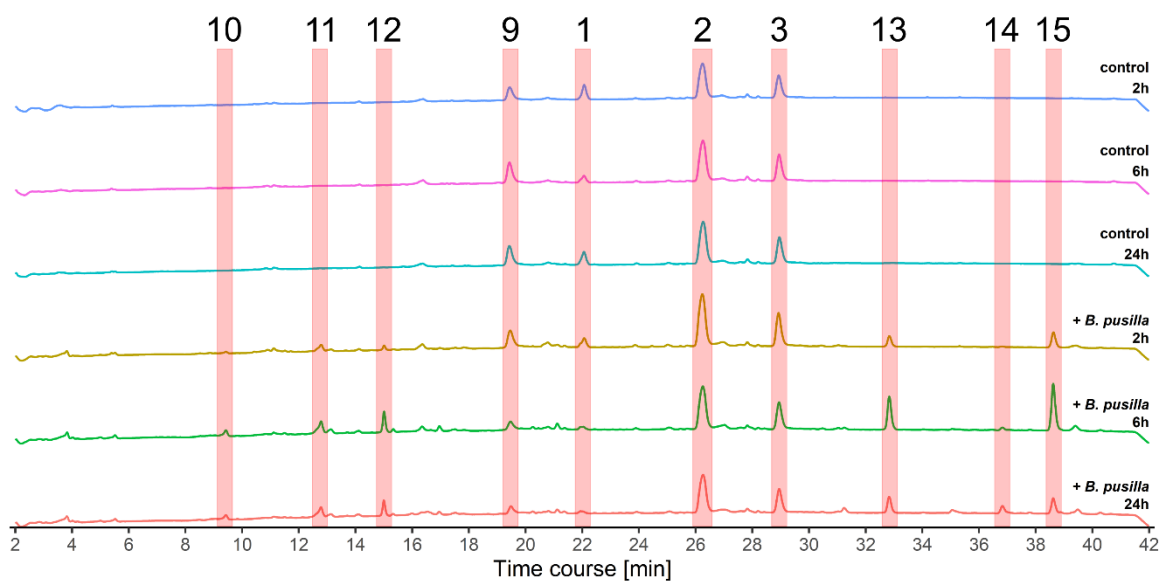


Figure 45 **HPLC traces of cell extracts of *N. punctiforme* with and without *B. pusilla* in co-cultivation.** Samples were harvested after 2, 6 and 24 h and prepared for HPLC analysis as described in section 2.2.5.5. 1: Microviridin, 2: Nostopeptolide1052, 3: NostopeptolideA. HPLC program including the acetonitrile gradient applied as described in section 2.2.5.6.

As expected, the cell extract samples were less heterogenous with smaller peak counts but higher intensities. Within the control samples besides the known compounds, microviridin (1), the two nostopeptolide variants (2, 3) and a conserved unknown peak (9), there were only a few peaks with neglectable intensities observed (Figure 45). However, within the co-cultivation sample a set of unique peaks could be identified. Some of these new peaks accumulated over the course of the experiment (12, 13, 14) while others remained constant (10, 11) or showed a differential behavior (15) (Figure 45). The most prominent differential peaks eluted late between 30 and 40 min corresponding to 45 to 60 % acetonitrile and thus pointing towards compounds with higher hydrophobicity (Figure 45).

Unfortunately, in a direct comparison of the 6 h co-cultivation culture with a *B. pusilla* cell extract and *N. punctiforme* control culture extract the majority of differential peaks could be associated to the plant host (Figure 46).

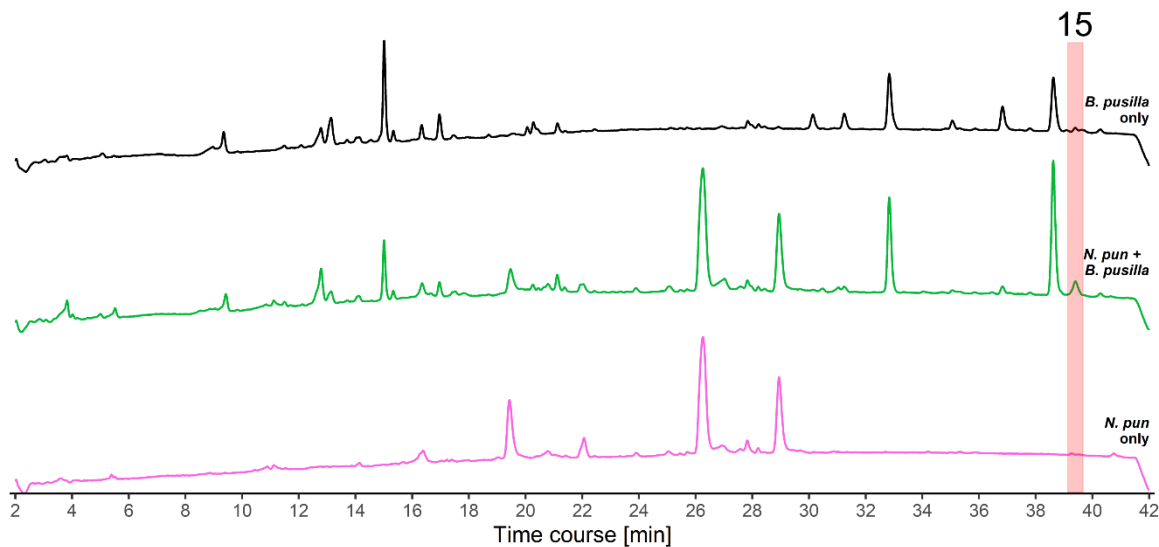


Figure 46 HPLC traces of cell extracts of *B. pusilla* and *N. punctiforme* in single culture in comparison to the 6 h co-cultivation cell extract of *N. punctiforme* with *B. pusilla*. Sample processing conducted as described in section 2.2.5.5. HPLC program including applied acetonitrile gradient as described in section 2.2.5.6.

However, peak **15** could not neither be associated to *B. pusilla* nor to *N. punctiforme* and resembled a promising candidate for a compound specifically induced by bacterial-host interaction (Figure 46).

Because of the higher abundance in the 6 h co-cultivation sample, this extract was used to isolate peak **15** by fractionation (as described in section 2.2.5.6). The fractions were pooled according to the respective peaks and prepared for mass spectrometry via MALDI-TOF (as described in sections 2.2.5.7 and 2.2.5.8).

The isolated peak showed a heterogenous MALDI-TOF mass spectrum with three prominent masses of 556.26 Da 597.56 Da and 790.52 Da. The core peptides of RiPP3 and RiPP4 precursor genes Npun\_R3313, Npun\_F5045 and Npun\_F5046 have theoretical masses of 2204.875 Da, 1974.769 Da and 1946.763 Da, respectively (Figure 47).

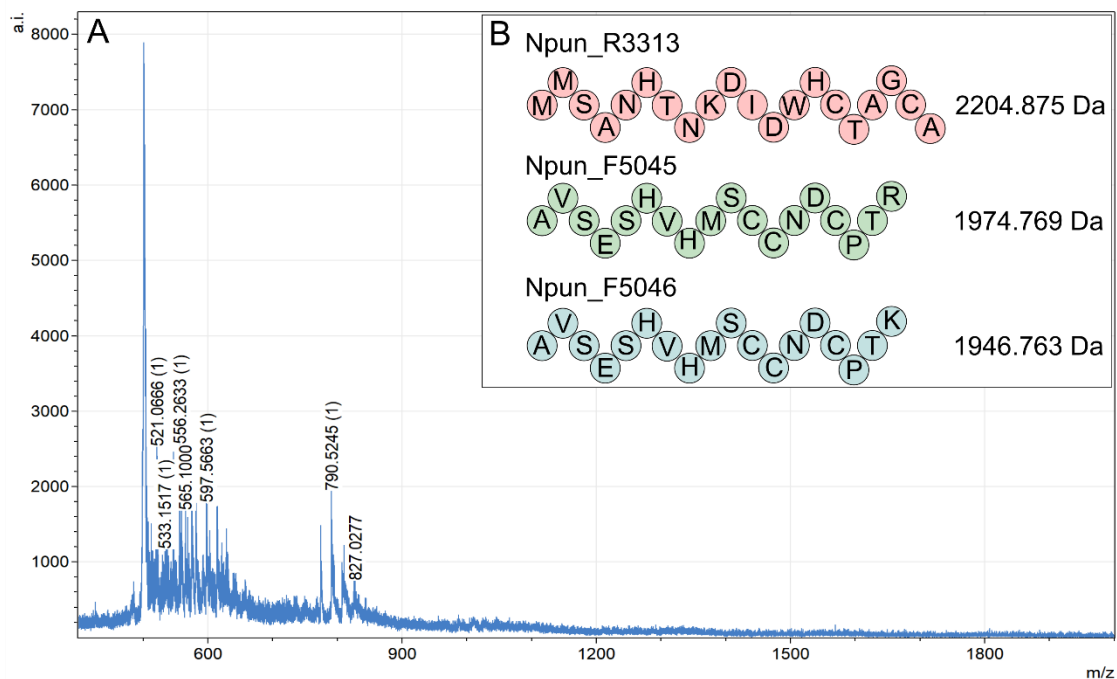


Figure 47 **MALDI-TOF mass spectrum of isolated peak 15 including RiPP3 and RiPP4 core peptides.** A: Mass spectrum of peak 15 cropped from 400 – 2000 Da. MALDI-TOF sample preparation and measurement conducted as described in sections 2.2.5.7 and 2.2.5.8. B: Schematic representation of the RiPP3 and RiPP4 core peptides and associated theoretical masses, annotations refer to the respective precursor genes.

Unfortunately, no mass detected within peak 15 exceeded 827 Da and it is thus unlikely that this isolate contains a potential RiPP3 or RiPP4 BGC product.

The outcome of the analysis of the broad increase in metabolic activity under high cell density conditions and the analysis of the symbiotic-interaction mediated distinct induction of specific secondary metabolite BGCs are strikingly similar as in the latter study no compound could be assigned to isolated induced BGCs. This again points towards the presence of a general issue in analytics or product prediction accuracy.



### 3.3 Mutagenesis of *N. punctiforme*'s NRPS-PKS hybrid BGC PKS4

Although it became clear in the course of this study that there are potent ways to specifically induce the expression of certain *N. punctiforme* secondary metabolite BGCs, those BGCs could not be assigned to a product. An alternative method to approach the multitude of substances detected in the HPLC analyses is to interrupt or alter a BGC of interest and backtrack the induced impact on the array of substances that can be detected. A common way to achieve this is to knock-out selected genes of the BGC or exchange its promoter. One of the largest secondary metabolite BGC detected within *N. punctiforme*'s genome is the NRPS-PKS hybrid gene cluster referred to as PKS4 (Figure 48). Since this specific BGC showed a very interesting transcriptional behavior during the RNA deep-sequencing time course study (Figure 19) and the respective CFP reporter mutant strain also showed a unique spatially restricted CFP expression pattern even under standard low light conditions (Figure 26) this BGC was chosen for a mutagenesis study.

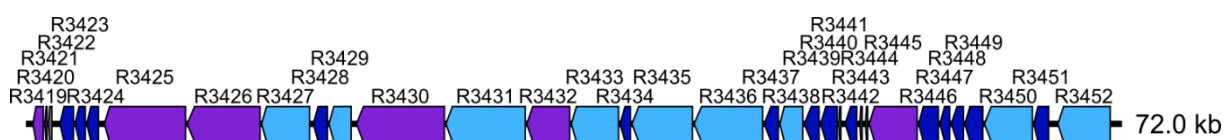


Figure 48 **Map of the *N. punctiforme* NRPS-PKS hybrid BGC PKS4.** Genes colored in light blue highlight NRPS genes, purple highlights PKS genes, dark blue highlights other biosynthetic genes and grey indicates putative genes of unknown function.

Utilizing the tri-parental mating technique (as described in section 2.2.3.3) in four independent approaches different parts of the BGC should be altered. The first two approaches aimed at knocking-out either only the first gene of the cluster (Npun\_R3452) or the leading two genes of the cluster (Npun\_R3452 and Npun\_R3451). The third approach aimed at the deletion of the whole PKS4 BGC while the final approach aimed at the substitution of the BGC's first promoter region with the weak constitutive nostopeptolide BGC promoter (promoter region of Npun\_F2181). Therefore, based on the pRL271 vector different integrative suicide plasmids were constructed in *E. coli* by HiFi-Builder mediated homologous recombination.

Table 6 Overview of the integration plasmids constructed for *N. punctiforme*'s PKS4 BGC mutagenesis.

Construct	Integration result	Final size [bp]
pDD014	Npun_R3452::StrepR	9141
pDD016	Npun_R3451-3452::StrepR	9141
pDD018	Npun_R3419-3452::StrepR	9141
pDD023	3'UTR(Npun_R3452)::P <sub>Npun_2181</sub> StrepR	10 063

In the first step the common base plasmid pDD012 was built. Therefore, pRL271 was linearized with *SacI* and a streptomycin resistance cassette was introduced via homologous recombination. Thereby there were also two unique restriction enzyme recognition sites upstream and downstream of the newly inserted sequence introduced, namely *AfeI* (5') and *EheI* (3'). Utilizing these two linearization spots in two independent HiFi-Builder mediated cloning steps each of the final plasmids pDD014, pDD016 and pDD018 was built by introduction of two 1000 bp homology sequences targeting the 5' region and 3' region of the deletion sequence, respectively, for site directed integration of the final mutagenesis sequence. For the construction of the promoter substitution construct pDD023 a second cloning step at the 5' end of the streptomycin cassette was necessary for the introduction of the NPun\_2181 promoter sequence.

After successful assembly of the desired integration plasmids the *E. coli* XL1-blue cells carrying the respective constructs were directly used for tri-parental conjugation into *N. punctiforme* as described in section 2.2.3.3.

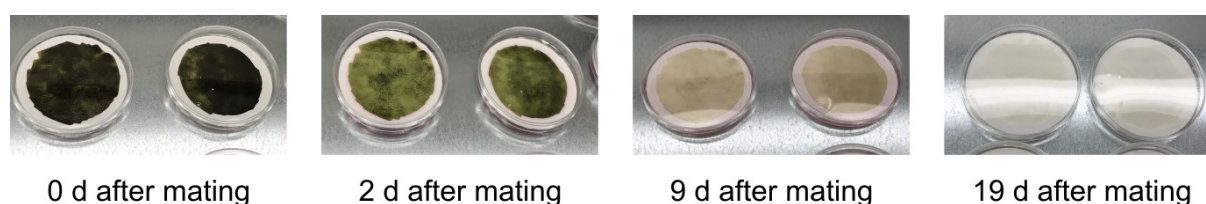


Figure 49 Typical course of PKS4 mutagenesis approach via tri-parental mating.

Unfortunately, several attempts to genetically modify *N. punctiforme*'s PKS4 BGC using the afore described suicide plasmids were unsuccessful. Even after extended incubation of the transformation membranes on BG11<sub>0</sub> agar plates supplemented with streptomycin no potential mutant colonies could be identified (Figure 49).

A specific feature of the pRL271 based constructs is that these plasmids do not carry a cyanobacterial origin of replication and are thus referred to as suicide plasmids that cannot be propagated inside *N. punctiforme*. To increase the theoretical retention time of such an integration plasmid inside the cell, in a second approach the pRL1049 cyanobacterial origin of replication was cloned into the plasmids pDD014, pDD016, pDD018 and pDD023 linearized with XhoI by HiFi-Builder mediated homologous recombination yielding the constructs pDD034, pDD035, pDD036 and pDD037. However, after successful assembly of these plasmids again no potential mutant colonies could be found after tri-parental conjugation.

### 3.4 Heterologous expression of *N. punctiforme*'s PKS5 BGC in *E. coli*

Another alternative way to decipher the complex mixture of small compounds produced by an organism is to transfer a gene cluster of interest into a new, better described genetic framework. This heterologous expression then allows to specifically detect changes in the peptide spectrum produced by the host organisms and thus streamlines the assignment of a product to the transferred BGC. Accordingly, the aim of this side project was to heterologously express the *N. punctiforme* secondary metabolite BGC PKS5. Although cloning more than 20 kb of target sequence is not trivial the PKS5 BGC was a promising candidate since its core genes are densely located and consist of only eight genes (Figure 50).

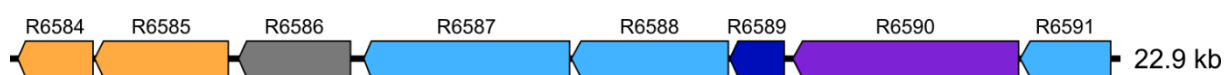


Figure 50 **Map of the *N. punctiforme* NRPS-PKS hybrid BGC PKS5.** Genes colored in light blue highlight NRPS genes, purple highlights PKS genes, dark blue highlights other biosynthetic genes, yellow highlights transport related genes and grey indicates putative genes of unknown function.

PKS5 contains three NRPS units (Npun\_R6587, Npun\_R6588 and Npun\_R6591), a single PKS unit (Npun\_6590), an acetyl transferase (Npun\_R6589) an FG-GAP repeat containing protein of unknown function (Npun\_R6586) and two transporter genes (Npun\_R6584 and Npun\_R6585). Based on the calculated activation domain specificities AntiSMASH predicted a small putative product consisting of two valines, an unknown amino acid and a malonate. Taking the missing acetylation into account the predicted product mass ranges around 417 Da.

The assembly of the heterologous expression construct was designed as a three step *in vivo* assembly utilizing only the fullRecET LLHR technique as described by Wang et al. 2016 and in section 2.2.2.9 (72). Therefore, the PKS5 BGC sequence was separated into three parts of equal size that should consecutively be assembled into the fullRecET expression vector pBR322\_tetR\_tetO\_hygg\_ccdB as described by the authors. However, due to technical issues it was impossible to PCR amplify the second and third sequence section of the PKS5 BGC. To still accomplish the synthesis of the heterologous expression plasmid the cloning strategy was altered to enable the utilization of those PCR products that could reliably be generated. An overview of the five-step cloning approach that ultimately lead to the correct assembly of 27.384 kb pBR322\_PKS5\_Frag1\_Frag2\_Frag3 plasmid is summarized in Figure 51.

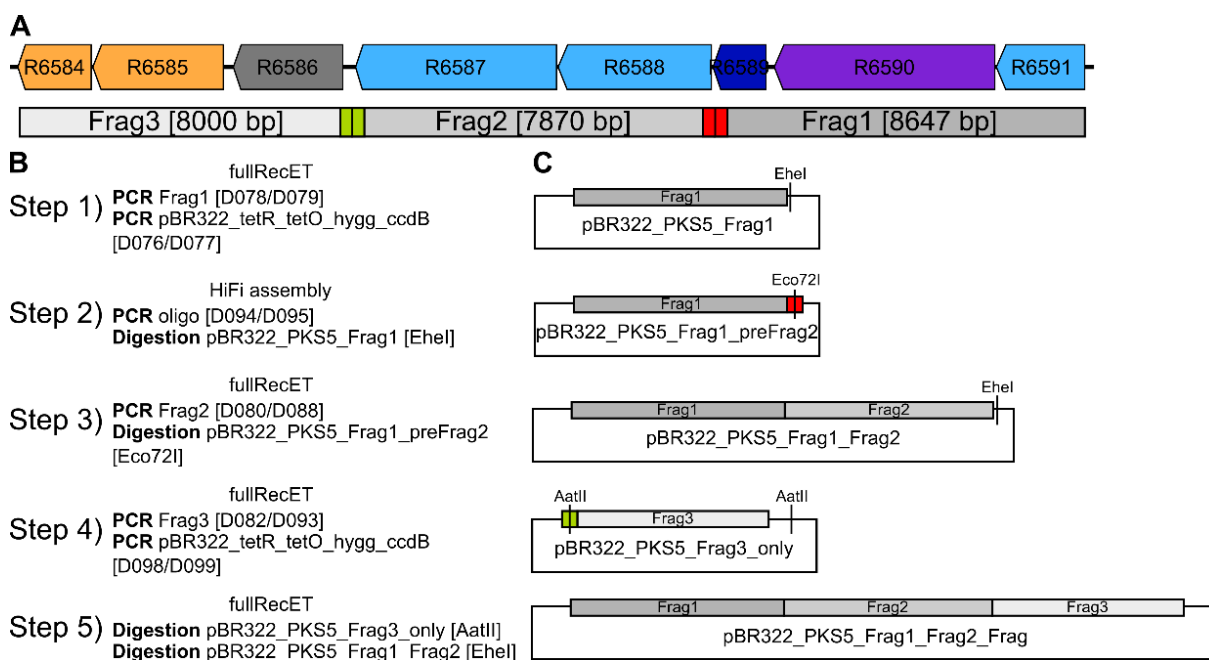


Figure 51 **Stepwise overview of the conducted cloning steps to assemble the *N. punctiforme* PKS5 heterologous expression plasmid in *E. coli*.** A: PKS5 BGC map highlighting the separation into three equally sized fragments for the assembly approach; B: Overview of the different cloning steps including the required parts and methods used for the assembly of the desired construct of the given step; C: Schematic overview of the resulting construct of each cloning step.

The primers used within this project are listed in appendix section 6.1. Each intermediate construct as well as the final expression plasmid were confirmed via sequencing prior to any expression attempt.

For the heterologous expression of the PKS5 construct the inducible vector pBR322\_PKS5\_Frag1\_Frag2\_Frag3 was retransformed into the expression strain *E. coli* GB05-MtaA, an engineered *E. coli* strain that had the *Bacillus subtilis* phosphopantetheinyl transferase MtaA, crucial for the activity of PKS and NRPS enzymes, integrated into its genome.

In order to express the PKS5 construct three times 100 ml LB medium supplemented with 100 µg/ml ampicillin (LB-amp) were inoculated with an overnight culture of *E. coli* GB05-MtaA\_pBR322\_PKS5\_Frag1\_Frag2\_Frag3 to an OD<sub>600 nm</sub> of 0.1. Additionally, 100 ml LB-amp were inoculated with *E. coli* GB05-MtaA as a no-plasmid control. The expression cultures were incubated at 30 °C and 220 RPM until an OD<sub>600 nm</sub> of 0.7 was reached. Then two *E. coli* GB05-MtaA\_pBR322\_PKS5\_Frag1\_Frag2\_Frag3 culture was induced with 0.5 µg/ml tetracycline, one was kept as a non-induced control and all cultures were incubated at RT and 220 RPM for 5 days. Afterwards the cells were harvested, extracted and prepared for HPLC analysis as described in sections 2.2.5.5 and 2.2.5.6.

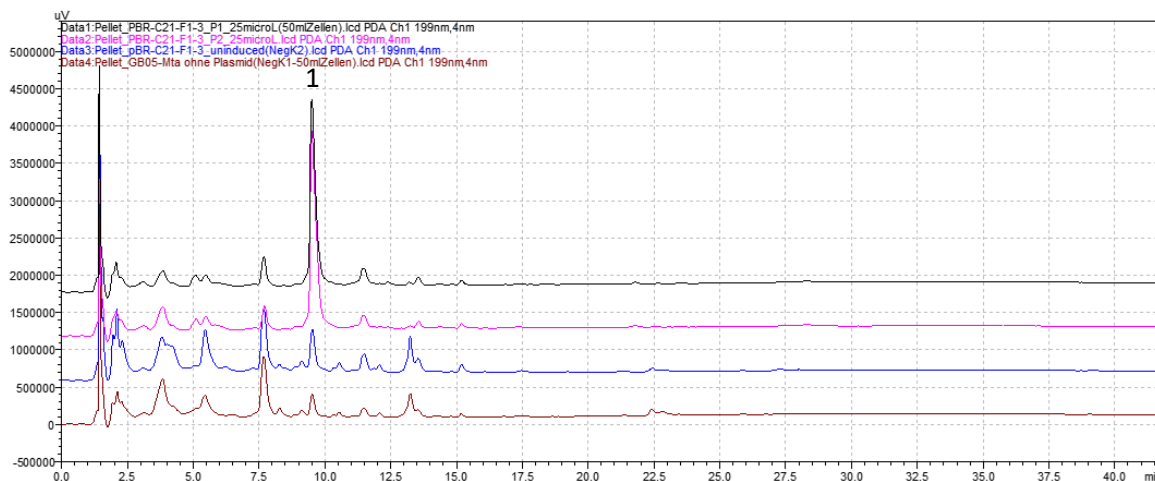


Figure 52 HPLC traces of cell extracts of *E. coli* GB05-MtaA harboring the PKS5 expression plasmid. Black: induced PKS5 expression culture 1; Purple: induced PKS5 expression culture 2; Blue: non-induced control culture; Brown: no-plasmid control culture. HPLC program used as described in section 2.2.5.6.

After induction and expression of the PKS5 construct in *E. coli* GB05-MtaA there could a reproduceable peak at 9.5 min be detected (**1**) that was neither present in the non-induced control nor in the no-plasmid control (Figure 52). To check whether the compound could resemble a putative PKS5 product, the peak was fractionated and analyzed via MALDI-TOF mass spectrometry.

Although there were additional masses detected within the peak (**1**) fraction, there was a signal at 439.95 Da. This mass fell in the range of the predicted putative PKS5 product. For structural elucidation of the compound harvested expression cultures as well as protein extracts were sent to Dr. Keishi Ishida at the HKI in Jena (Germany). Unfortunately, LC-MS/MS analysis revealed that the main compound in the isolated PKS5 expression fraction was L-tryptophan. This hints towards an unexpected side effect and needs to be further examined and the heterologous expression attempt was thus not successful under afore described conditions. Next steps could include the co-expression of *Anabaena sp.* PCC7120 HetI phosphopantetheinyl transferase in *E. coli* as this coenzyme is more closely related to *N. punctiforme* or to switch to a cyanobacterial heterologous expression host like *Synechocystis sp.* PCC6803.

## 4 Discussion

### 4.1 *N. punctiforme* PCC73102: a suitable target for cyanobacterial natural product research

One major motivation of this study was to assess the suitability of the terrestrial filamentous cyanobacterium *N. punctiforme* PCC73102 as a model strain for genomic mining of cyanobacterial natural products. Establishing such a model system was important, since (I) there is still an imbalance in available data between terrestrial and aquatic strains due to the higher interest in potentially toxic bloom-forming cyanobacteria and (II) there is no strategy available enabling broad scale analysis of secondary metabolism in filamentous cyanobacteria. Being a terrestrial cyanobacterium and having a largely untapped secondary metabolome, *N. punctiforme* satisfies both needs perfectly. Moreover, *N. punctiforme* is amenable to genetic modification so that genetics-based new strategies for the systematic analysis of the organism's secondary metabolite machinery were deemed feasible. For this purpose, a fluorescent reporter gene approach promised to be suitable. This concept was successfully realized by the generation of a *N. punctiforme* CFP reporter strain library consisting of sixteen mutant strains each harboring a single transcriptional reporter plasmid for an individual secondary metabolite BGC. This library has not only led to new and unexpected insights into *N. punctiforme*'s secondary metabolome, but it also set the ground for the creation of a model system to obtain a better understanding of the tight regulatory machinery that underlies cyanobacterial natural product biosynthesis.

Further employing the reporter strain library, the second major question regarding the impact of symbiotic interactions on the expression of distinct secondary metabolites was tackled in this work. While earlier studies already pointed towards a genetic reprogramming of *N. punctiforme*'s secondary metabolome in host plants (71), the utilization of the reporter strain library enabled a systematic assignment of differentially expressed BGCs and furthermore the differentiation of physical and chemical interaction effects. Using both the reporter mutant system and RT qPCR approaches it could be shown that, as expected, the exposition to a potential host plant caused differential effects in terms of transcriptional activity of different natural product BGCs. Particularly two distinct secondary metabolite BGCs, RiPP3 and RiPP4, were strongly upregulated in these experimental set ups. The results of the interaction experiments conducted in this study

further emphasized that the complex lifestyle of *N. punctiforme* is closely connected to the regulation of its secondary metabolome.

Taking the major insights of the current study together it could be proven that *N. punctiforme* is a suitable model organism for the investigation of the secondary metabolism within terrestrial, filamentous cyanobacteria, that the reporter gene based mining approach offers huge benefits over classical genomic mining strategies and that the investigation of the symbiotic lifestyle of *N. punctiforme* with respect to natural products might be a new way of discovering so far undescribed, tightly regulated compounds.

## **4.2 Genomic mining is a potent tool for natural product research**

There is rising evidence that the overall abundance of natural product BGCs encoded in genomes of diverse microbial phyla is substantially higher than the number of identified products synthesized by these organisms under standard laboratory conditions. Thus, new screening techniques are needed to uncover those hidden treasures (86). Driven by recent advances in high-throughput genome sequencing and accompanying price drops, genomic mining became a potent new way of discovering so far undescribed high value compounds by combining powerful computational approaches with the increasing amount of sequencing data becoming available (35). In contrast to more classical screening methods that are either bioactivity-based, or, later on, target-based, genomic mining strategies are more knowledge-driven and based on structural or functional homologies of an undescribed organism or sequence compared to a described system (35–37).

Initially the genomic mining strategy was nothing more than a reverse genetics approach, where a given genome was browsed for the presence of one or more sequences coding for already known key biosynthetic enzymes, but conducted *in silico*. However, the methodology evolved rapidly and several unique strategies arose to actually identify new secondary metabolites by identifying the underlying BGCs which were then further examined by a variety of genetic approaches to reveal the associated compound (43). Common strategies involve (I) the exchange of the endogenous promoter with a constitutive promoter, as successfully conducted for streptocollin, a RiPP produced by *Streptomyces collinus* (87), (II) overexpression of cluster-associated regulator genes, as conducted during the discovery of stambomycin a PKS produced by *Streptomyces*



*ambofaciens* (88) or (III) heterologous expression of the full BGC in a suitable expression host (87). Other approaches involve knock-out strategies or optimization of culture conditions in order to track down a gene clusters product (89, 90).

A rather new strategy was deployed in a large comparative genomic mining study, involving transcriptomics data to assess global BGCs expression (85). Within this study the authors mined a total of 119 *Salinispora* genomes and selected four strains that encoded for a total of 49 different BGCs of which 36 were orphan. One topic the authors wanted to highlight was the relationship between orphan BGCs and transcription levels. Utilizing clusters with described products they defined an empirical threshold for active BGC expression, yet even though an average of 50 % of the orphan BGCs among the investigated strains were transcribed above this threshold, a product could not be associated. This observation is also reflected in the transcriptomics data gathered in this study, as some of *N. punctiforme*'s orphan BGCs show clear transcriptional activities above the defined threshold and yet no product could be associated. Another outcome of the *Salinispora* study was that, on average, the investigated strains merely used 50 % of their biosynthetic potential (85). Taking only the transcriptomics data into account, the situation in *N. punctiforme* is even worse, since only for five of the sixteen BGCs more than one gene of a given cluster showed expression levels above the described threshold, making up for only 31 % of *N. punctiforme*'s biosynthetic potential (Figure 17 - Figure 20).

Because of these findings, within this study another layer of information was gathered by not only combining genomic mining data with transcriptomics data but also by a fluorescence reporter gene approach to track secondary metabolite cluster activity in *N. punctiforme*. Therefore, a mutant strain library was produced with a single promoter region-CFP fusion reporter strain for each of the sixteen secondary metabolite BGCs investigated in this study. Undeniably, this approach does have certain limitations, because the organism of interest has to be amenable to genetic modification such as the introduction and maintenance of an extrachromosomal plasmid. On top of that, such genetic alterations may cause unwanted side effects due to the employment of a selection marker or the overall increase of the metabolic burden caused by the expression of the additional modified reporter sequence. However, there are a number of strong benefits supporting a reporter gene based approach for investigation of secondary metabolite BGC expression in a target organism, since (I) gene cluster activity can be easily monitored *in vivo* by fluorescence microscopy, (II) by the same method, the detection can take place

on a single-cell level in contrast to the culture average information gathered during RTqPCR and (III) experimental set-up and sample preparation require minimal effort and time so that even short-term effects can be easily investigated. A similar approach was also successfully conducted for investigation and induction of silent secondary metabolite BGCs in two *Streptomyces* strains, however instead of focusing on fluorescence signals the authors employed a colorimetric signal approach monitoring the enzymatic conversion of catechol (91, 92). The main reason why silent secondary metabolite BGCs are especially challenging is that the system has three variables since an unknown elicitor may activate an undescribed gene cluster inducing the production of a new product. However, bioinformatics and metabolomics can often deduce the chemical structure of a new metabolite and thus backtrack the associated set of biosynthetic genes once a given silent BGCs is induced so that the main obstacle that remains is a robust screening system allowing for an easy detection of favorable expression conditions (93, 94). For *N. punctiforme* such a system was realized in terms of the reporter mutant strain library and successfully used to screen different elicitor factors like high density cultivation, conditioned medium exudates and biotic interactions but also to gain basic information about the state of *N. punctiforme*'s secondary metabolome under standard laboratory conditions. However, there are still many questions that will be addressed by utilization of the reporter-based expression monitoring system, e.g. the effects of other biotic and abiotic interactions or to which extent a quorum sensing mechanism regulates the expression of certain secondary metabolite BGCs in *N. punctiforme*.

### **4.3 The majority of *N. punctiforme* secondary metabolite BGCs is not silent under standard laboratory conditions**

One of the most unexpected findings of this study was the observation that secondary metabolite gene expression is not switched on or off globally as in the case of the PKS3 or the PKS1 cluster, respectively (Figure 53 A, B). Instead, nine out of sixteen natural product BGCs fell in one of three unique categories of spatially restricted expression (Figure 53 C, D, E).

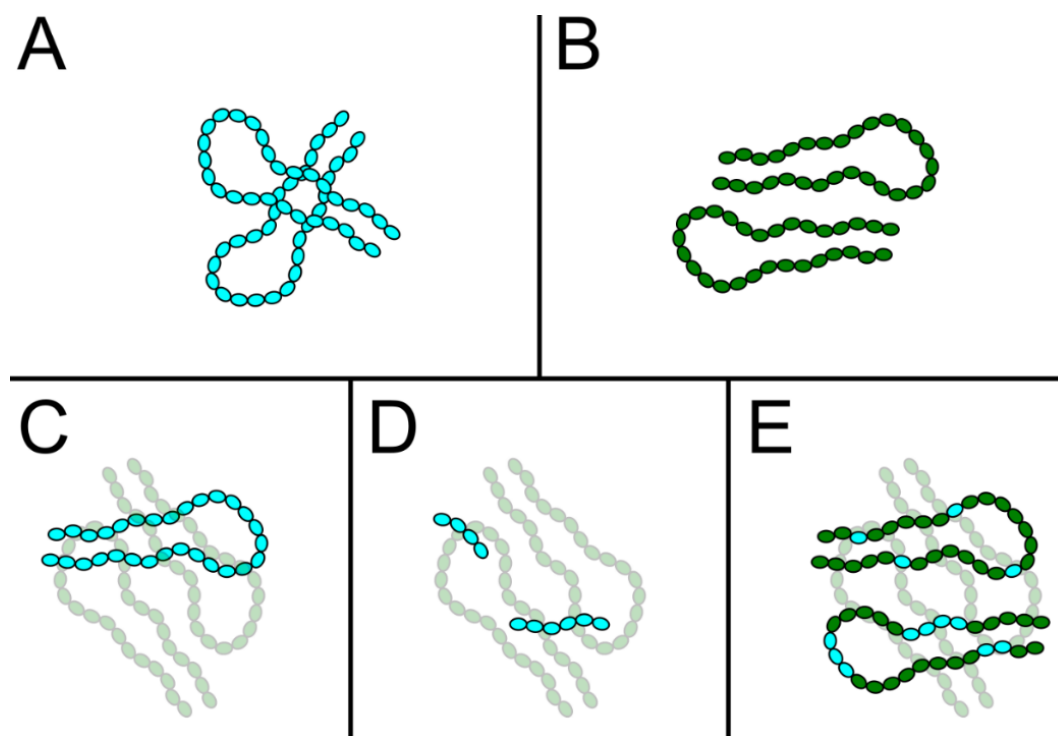


Figure 53 **Schematic expression patterns of *N. punctiforme* secondary metabolite BGCs observed under standard laboratory conditions.** A: constitutive expression; B: no expression (silent); C: expression in long isolated filaments; D: expression in short isolated cell stretches; E: expression in single cells or short cell stretches within a filament.

A similarly defined expression pattern was observed for the heterocyst cell differentiation factor PatS in the filamentous cyanobacterium *Anabaena* sp. PCC7120. Heterocysts differentiate from apparently equivalent vegetative cells but in a highly structured manner with approximately ten vegetative cells between two heterocysts in a single filament. However, it turned out that besides environmental factors the expression level of the *patS* gene product, a short peptide, is essential for the maintenance of this structured pattern. Overexpression mutants of PatS were unable to differentiate heterocysts while deletion mutants developed multiple heterocysts without a detectable pattern. Furthermore, the authors highlighted that the *patS* expression additionally followed a temporal pattern important for fixation of cell fate (95). Indications that a similar mechanic could be true for *N. punctiforme*'s secondary metabolite BGCs were found for the PKS4 cluster (Figure 19). For this BGC not only a time dependency could be observed, as the transcription peaked after 24 h, but also a very regular patterning of actively expressing cells was detected, pointing towards a potential differentiation regulating role of the product. To maintain these highly structured patterns a tight regulation has to be employed by *N. punctiforme*. In the case of PatS mediated heterocyst differentiation a concentration dependent feedback regulation seems to be the primary

mechanism. However, the underlying regulation strategy of PKS4 expression has yet to be figured out.

Despite PKS4, for the majority of BGCs such unique temporal and spatial behavior could not be monitored, and the observed patterns were rather stable and not linked to a specific cell type. This might hint towards products, which carry out functions in other biological processes such as signal transduction or defense against competitors. If the patterns are not related to cellular differentiation another explanation of the spatially restricted expression might be an early type of division of labor. West and Cooper define three major requirements for true division of labor that must be met, (I) phenotypic variation, i.e. different tasks are performed by different individuals within a population, (II) cooperation between these differently acting individuals and (III) adaption, i.e. the behavior maximizes the overall fitness of the population (96). An obvious example that filamentous cyanobacteria are readily conducting division of labor is the fact that species like *Anabaena* and *Nostoc* differentiate heterocysts, a specialized cell type for the mere purpose of nitrogen fixation. Under the aspect of division of labor heterocystous cyanobacteria meet all three requirements, because (I) the heterocysts are clearly phenotypically distinguishable from vegetative cells, (II) both cell types cooperate as heterocysts provide combined nitrogen for the population and vegetative cells supply heterocysts with carbon sources from photosynthesis and (III) although individual heterocysts lack the ability of cell division, the whole colony would starve without their presence and thus heterocysts do in fact increase the overall fitness of the population. Another prominent example of microbial division of labor is sporulation, as seen in myxobacteria where a large fraction of cells undergo programmed cell death upon formation of the fruiting bodies (97). There are also division of labor types discussed in apparently unicellular organisms like *Saccharomyces cerevisiae* where formation of large biofilms or apoptotic behavior of a subset of cells is considered division of labor under the aspects of cooperation and adaption (98). Separation and specialization also make sense in terms of the production of secondary metabolites like antibiotics or secreted enzymes, since growth inhibition is a frequently observed trade-off for the costly synthesis of these products. However, it is sufficient that a subfraction of cells produce these compounds to enable an overall advantage in the habitat for both producers and non-producers while still maintaining efficient propagation rates in the non-producing subfraction (97). Although this topic is rarely investigated in prokaryotes there is clear evidence available

for fungi that represent a filamentous system resembling that of *N. punctiforme*. For example in *Aspergillus* heterogeneity in gene expression and translational activity lead to specialized subclasses of hyphae that were found to fulfil unique tasks (99). The observed spatially restricted expression patterns for the various secondary metabolite BGCs of *N. punctiforme* might be explained with a similar development of specialized subpopulations within a colony, either because the production is costly and the connected growth trade-off needs to be restricted or the product itself is harmful to the producer and a subfraction of cells is sacrificed for the increase of total populations fitness. An example supporting the latter hypothesis is part of the study published by Guljamow and co-workers in 2017 (77). Therein new anabaenopeptin variants were identified from *Nostoc* KVJ2, a strain distantly related to *N. punctiforme*, that were only detectable under inductive cultivation conditions. Interestingly the authors were able to backtrack the already described allelopathic trait of *Nostoc* KVJ2 against a second *Nostoc* strain, isolated from the very same habitat, to one of these newly identified anabaenopeptin variants. Certain BGCs of *N. punctiforme* might produce small molecules with similar anti-cyanobacterial effects. This hypothesis is supported by the observation that for some of the reporter strains especially the highly active cells showed amorphous enlarged phenotypes that point towards elevated stress levels (Figure 54)

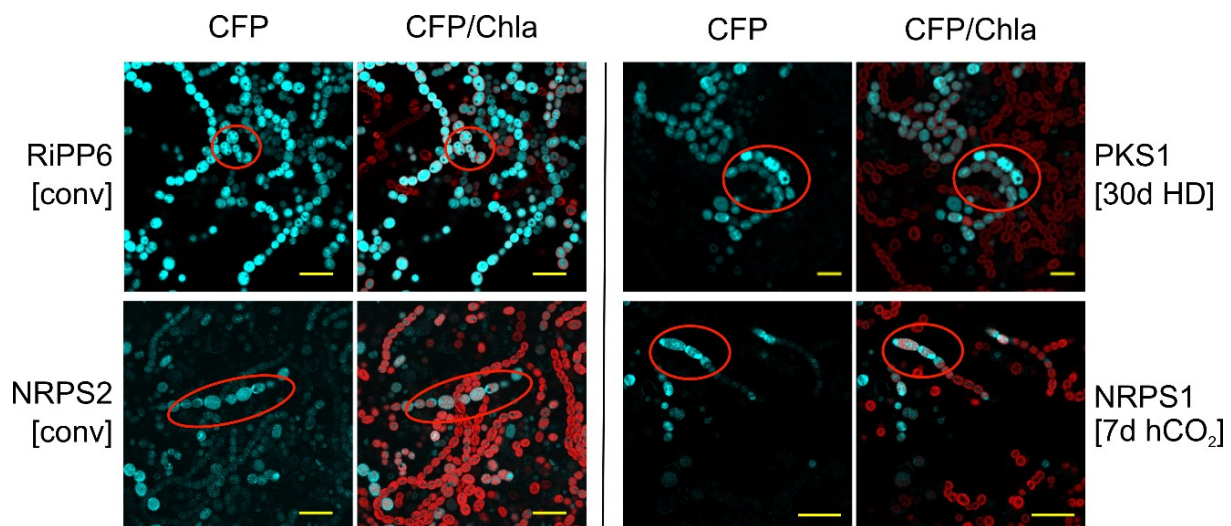


Figure 54 **Stressed phenotypes of *N. punctiforme* reporter mutants under various cultivation conditions.** Stressed cells are highlighted in red. RiPP6 and NRPS2 reporter mutants were cultivated under standard low light condition. PKS1 reporter was cultivated for 30 d in a high density cultivator (experimental procedure and data collection conducted by Julia Krumbholz). NRPS1 reporter was cultivated for 7 d with a buffer-filled dialysis bag producing a 32 mbar partial pressure of CO<sub>2</sub> over the culture. Yellow scale bar indicates 20 μm.

Large scale genomic mining studies show that especially late branching filamentous cyanobacteria tend to show a high density of secondary metabolite BGCs (29). This

observation is in good agreement with the findings of the current study, as the type IV cyanobacterium *N. punctiforme* also dedicates a substantial fraction of its genome to secondary metabolism. In the light of natural product biosynthesis, spatially restricted expression and the aspect of division of labor may resemble factettes of multicellularity. A mode of living that can be interpreted as an evolutionary strategy to provide the framework to build up a complex system of regulation, that is needed by a producing organism to utilize such high impact products while still maintaining effective growth.

#### **4.4 Alternative genomic mining strategies**

Apart from these rather new combinatorial approaches connecting genomic mining data with transcriptomics in order to decipher cryptic secondary metabolite BGCs in its native host, it is still common to isolate a given natural product BGC and transfer it to a well-known heterologous expression system such as *E. coli* or *Streptomyces coelicolor* to identify its product (43). Based on the advances in molecular biology the efficiency of the assembly of heterologous expression plasmids has drastically increased in recent years (100). One of these sophisticated cloning techniques was also used in this study to attempt the heterologous expression of *N. punctiforme*'s NRPS-PKS hybrid BGC PKS5. Unfortunately, even after successful assembly of the 27 kb inducible expression plasmid in *E. coli*, extensive analytics revealed that the plasmid caused an unexpected production of L-tryptophane and no actual product could be identified. This example is one among many showcasing that especially heterologous production of complex gene clusters is not trivial. Often the chosen production host is not suitable because of missing crucial co-factors like a PPTase enzyme. But there are also other obstructions like enzyme ratios, input material availability and proper enzyme folding conditions that have to be taken into account for a heterologous expression attempt (101).

Nevertheless, there are promising examples of successful heterologous expression of secondary metabolite BGCs available, like the lyngbyatoxin or microcystin production in *E. coli* published by the Neilan lab in 2013 and 2017, respectively (102, 103). Interestingly, until now heterologous expression attempts were only successful for natural products with described chemical structures (104). This method seems not suitable for the characterization of orphan BGCs, since (I) expression efficiency is hard to monitor, (II) protein purification is challenging due to the limited information available for an unknown product and (III) the complex constitution of crude extracts renders product identification very tough. Those aspects in conjunction with the results of this study point

out that there are specific drawbacks connected to the heterologous expression of natural products. As discussed above, many of *N. punctiforme*'s orphan or cryptic secondary metabolite BGCs seem to be tightly controlled or their expression might be even separated spatially from proliferating cells. Taking a given cluster out of its genetic context interrupts such control mechanisms and, depending on the biological function of these unknown compounds, overproduction likely leads to lethal effects. Nevertheless, such attempts are reasonable when the native producer is unculturable under laboratory conditions or it is not amenable to genetic modifications.

If the native producer can be cultivated and modified, approaches to induce or alter the expression of a given cluster in its natural context are more likely to be successful. Common knock-out or knock-in strategies involve the deletion of core biosynthetic enzymes in order to identify accumulating intermediate products or to the exchange of native promoters with constitutive ones in order to unlink the expression from environmental conditions (105). A similar approach was conducted within this study to gather further information about *N. punctiforme*'s BCG PKS4 which showed interesting temporal transcription patterns during the RNA deep-sequencing experiment. Utilizing tri-parental conjugation, it was attempted to generate a number of deletion mutants of the PKS4 cluster, ranging from single gene knock-outs to a full cluster deletion variant. In addition to that, the endogenous promoter was planned to be substituted with the weak constitutive promoter of *N. punctiforme*'s nostopeptolide BGC. However, in several attempts not a single PKS4 mutant variant could be produced while other *N. punctiforme* mutants could be generated consistently. This experience might again emphasize the importance of the tight regulation of these secondary metabolite BGCs under spatially restricted expression, since alteration of their leading genes or 5'UTRs again would interrupt these regulatory elements. This in turn might be the reason why no vital mutants could be generated. On the other hand, the complete deletion of the cluster was also unsuccessful pointing towards the cluster product being of very high importance for *N. punctiforme*.

The results of the heterologous expression approach of the PKS5 gene cluster and the mutagenesis approach of the PKS4 gene cluster both highlight the importance of finding alternative ways of switching on an organism's conditional secondary metabolome. Thereby, the development of robust ways of screening for favorable conditions or elicitor

factors that trigger metabolite biosynthesis is crucial for the identification of hidden natural products.

#### **4.5 Symbiotic interactions trigger major genetic reprogramming in *N. punctiforme***

Previous investigations revealed that the exposure of *N. punctiforme* to a nitrogen-deprived host plant supernatant is sufficient to trigger a specific genetic response. A transcriptional reporter for the nostopeptolide BGC indicated a strong downregulation of the gene cluster upon exposure to conditioned *B. pusilla* exudate (71). This observation together with the physiological changes that *N. punctiforme* undergoes when in vicinity to a potential host plant already give an idea of the extent of the genetic reprogramming that takes place in the organism during establishment of a symbiosis (106). For a broad scale analysis of the impact of symbiotic interactions with *B. pusilla* specifically on *N. punctiforme*'s secondary metabolite BGCs, both chemical and physical interaction was investigated. Therefore, transcription levels of representative biosynthesis genes of different natural product BGCs during such interactions were monitored in this study. Using similar techniques in a large comparative transcriptomics study, Warshan and co-workers investigated the genetic responses of *N. punctiforme*, the model symbiotic strain, a feathermoss isolated *Nostoc* strain and a non-symbiotic *Nostoc* strain in chemical and physical contact with the feathermoss *Pleurozium schreberi* (107). Although the symbiotic association in the case of *P. schreberi* is truly epiphytic and not endophytic as in the case of *B. pusilla*, the pre-symbiotic cyanobiont-host communication and associated genetic changes before and during symbiosis are likely very similar. The authors identified 32 gene families unique to the symbiotic *Nostoc* strains that were upregulated during chemical contact with the feathermoss, with their main functions including motility, chemotaxis, signaling and aliphatic sulfonate-transport (107). Most interestingly, a protein, containing a conserved ligand-binding motif known from  $\alpha$ -integrins called FG-GAP repeat, (Npun\_R6586) was found highly upregulated in chemical contact but downregulated in physical contact and proposed to play a role in establishment of symbiosis since it was also abundant among the secreted proteins. This gene is part of the NRPS-PKS hybrid BGC PKS5 of *N. punctiforme* (107). While Warshan et al. observed an upregulation of *pks5* after 24h, a varying response of *N. punctiforme*'s BGC PKS5 could be observed in the present investigation. In chemical contact, (Figure 37) the gene cluster was downregulated and in physical contact it was upregulated after 24h (Figure 40). This



24h response pattern is especially interesting in connection with the observation that some PKS5 reporter strain cultures showed a distinct CFP signal among hormogonia (Figure 27) as cell differentiation to hormogonia peaks after 24 h of a synchronized fresh *N. punctiforme* culture. To further investigate the temporal behavior of PKS5 more experiments with synchronized fresh reporter cultures need to be deployed e.g. using the time laps video system of the confocal laser scanning microscope in conjunction with higher sampling frequency for HPLC and expression analysis. Comparing the 5h timepoint of Warshan et al. with the current study the response of pks5 was similar in the *P. schreberi* and *B. pusilla* interactions studies, since here the transcription level was highest among all time points detected for PKS5 during chemical contact (Figure 37) and lowest among all time points during physical contact (Figure 40).

For the other secondary metabolite BGCs there was a general trend of either downregulated or unchanged transcription during chemical contact with *B. pusilla* at all three sampling time points. In terms of physical contact with *B. pusilla* after 1 h and 5 h, a similar behavior could be observed but after 24 h an overall increase of transcriptional activity could be observed compared to the untreated control culture. For both chemical and physical contact there were two exceptions as the RiPP3 and RiPP4 BGCs were also induced. However, for both clusters the degree of transcriptional activation was substantially stronger during physical contact with *B. pusilla*. Utilizing the corresponding CFP reporter strains, the induction of these two secondary metabolite BGCs could be confirmed by an independent technique (Figure 41 and Figure 42). Unfortunately, with the analytics available, no compound could be associated to any of the two orphan BGCs. The behavior of the other gene clusters in comparison to RiPP3 and RiPP4 impressively demonstrates the specificity of this induction and that the treatment with a potential host plant is a promising approach to expose *N. punctiforme* towards a single, rather complex, eliciting factor that can ultimately lead to a better understanding of the regulatory mechanisms deployed by the organism. The combination of genomic mining data with transcriptomics and reporter-guided expression monitoring is a sophisticated approach to gather detailed information about an organism's secondary metabolome that can yield new insights into yet undescribed regulatory networks and biological functions of small molecule natural products.

#### **4.6 Inaccurate predictions and analytical limits are the main bottlenecks of natural product discovery**

Several unique approaches to alter and monitor secondary metabolite production in *N. punctiforme* - high cell-density cultivation, exposure to a symbiotic host organism, RTqPCR-mediated expression monitoring and a BGC specific expression reporter mutant strain library - were combined in this study. This not only led to the striking insight that in general the secondary metabolite production is rarely plain silent but instead shows very specific and so far undescribed patterns. It also allowed drawing more specific conclusions about the source of the observed overall increase in metabolic activity under high cell density cultivation conditions. Approximately 50% of *N. punctiforme*'s predicted secondary metabolite BGCs were shown to be induced under those conditions. Despite this very specific information just a single new product group, namely the miroviridins N3-N9, could be identified. The in-depth analyses of the other newly detected peaks remained inconclusive. In contrast to the general activation effect of high cell density cultivation the co-cultivation of *N. punctiforme* with *B. pusilla* led to a very specific induction of two distinct RiPP BGCs, namely RiPP3 and RiPP4. Unfortunately, even this very specific information did not lead to a correlation of a new compound to either of these orphan RiPP BGCs.

Combining the insights generated by the novel combination of genomic mining approaches with potent monitoring techniques it becomes apparent that the translational status of secondary metabolite BGCs might not be the issue. Instead the results point out that some of the main challenges include analytical limitations frequently caused by our limited understanding of the spectrum of modifications that natural products can go through upon maturation. This is especially true for ribosomal products as several studies highlight that, despite the genome mining tools reliably detecting undescribed gene clusters and our growing understanding of natural product biosynthesis, the information is not sufficient to detect new compounds straight in their natural producers. Instead going back to part-wise heterologous expression of single enzymes is still an efficient approach for deciphering pathways (108). On the other hand there is an ongoing process of detecting unanticipated mechanisms of posttranslational modifications like new ways of incorporation of D-amino acids or  $\beta$ -amino acids in proteins (108, 109). The resulting inaccuracies in structure prediction might as well contribute to the general bias in the ratio of detected to predicted compounds.

Since active translation of many natural product BGCs seems not an issue, one could exploit this to focus on bypassing the analytical limitations. Approaches to selectively enrich target products in complex extracts may include his-tagging of leader peptides or core biosynthetic enzymes, incorporation of non-canonical amino acids to deploy click-chemistry or the generation of fusion products with other affinity purification extensions (108, 110, 111).

## 5 References

1. Editorial, All Natural. *Nature chemical biology* **3**, 351 (2007).
2. David E. Cane, Christopher T. Walsh, Chaitan Khosla, Harnessing the Biosynthetic Code: Combinations, Permutations, and Mutations. *Science* **282**, 63–68 (1998).
3. H. S. Yoon, J. W. Golden, PatS and products of nitrogen fixation control heterocyst pattern. *Journal of bacteriology* **183**, 2605–2613 (2001).
4. R. Finking, M. A. Marahiel, Biosynthesis of nonribosomal peptides<sup>1</sup>. *Annual review of microbiology* **58**, 453–488 (2004).
5. K. J. Weissman, P. F. Leadlay, Combinatorial biosynthesis of reduced polyketides. *Nature reviews. Microbiology* **3**, 925–936 (2005).
6. D. J. Newman, G. M. Cragg, Natural Products as Sources of New Drugs from 1981 to 2014. *Journal of natural products* **79**, 629–661 (2016).
7. E. Dittmann, M. Gugger, K. Sivonen, D. P. Fewer, Natural Product Biosynthetic Diversity and Comparative Genomics of the Cyanobacteria. *Trends in microbiology* **23**, 642–652 (2015).
8. J. Gershenzon, N. Dudareva, The function of terpene natural products in the natural world. *Nature chemical biology* **3**, 408–414 (2007).
9. P. G. Arnison, M. J. Bibb, G. Bierbaum, A. A. Bowers, T. S. Bugni, G. Bulaj, J. A. Camarero, D. J. Campopiano, G. L. Challis, J. Clardy, P. D. Cotter, D. J. Craik, M. Dawson, E. Dittmann, S. Donadio, P. C. Dorrestein, K.-D. Entian, M. A. Fischbach, J. S. Garavelli, U. Göransson, C. W. Gruber, D. H. Haft, T. K. Hemscheidt, C. Hertweck, C. Hill, A. R. Horswill, M. Jaspars, W. L. Kelly, J. P. Klinman, O. P. Kuipers, A. J. Link, W. Liu, M. A. Marahiel, D. A. Mitchell, G. N. Moll, B. S. Moore, R. Müller, S. K. Nair, I. F. Nes, G. E. Norris, B. M. Olivera, H. Onaka, M. L. Patchett, J. Piel, M. J. T. Reaney, S. Rebuffat, R. P. Ross, H.-G. Sahl, E. W. Schmidt, M. E. Selsted, K. Severinov, B. Shen, K. Sivonen, L. Smith, T. Stein, R. D. Süßmuth, J. R. Tagg, G.-L. Tang, A. W. Truman, J. C. Vederas, C. T. Walsh, J. D. Walton, S. C. Wenzel, J. M. Willey, van der Donk, Wilfred A., Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Natural product reports* **30**, 108–160 (2013).
10. H. Wang, D. P. Fewer, L. Holm, L. Rouhiainen, K. Sivonen, Atlas of nonribosomal peptide and polyketide biosynthetic pathways reveals common occurrence of nonmodular enzymes. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 9259–9264 (2014).
11. J. Beld, E. C. Sonnenschein, C. R. Vickery, J. P. Noel, M. D. Burkart, The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life. *Natural product reports* **31**, 61–108 (2014).
12. B. Shen, Polyketide biosynthesis beyond the type I, II and III polyketide synthase paradigms. *Current Opinion in Chemical Biology* **7**, 285–295 (2003).
13. Shiina I., Katoh T., Nagai S., Hashizume M., Evaluation of the Efficiency of the Macrolactonization Using MNBA in the Synthesis of Erythromycin A Aglycon. *Chem Rec.* **6**, 305–320 (2009).
14. C. T. Walsh, The chemical versatility of natural-product assembly lines. *Accounts of chemical research* **41**, 4–10 (2008).
15. J. Vestola, T. K. Shishido, J. Jokela, D. P. Fewer, O. Aitio, P. Permi, M. Wahlsten, H. Wang, L. Rouhiainen, K. Sivonen, Hassallidins, antifungal glycolipopeptides, are widespread among cyanobacteria and are the end-product of a nonribosomal pathway. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E1909-17 (2014).

16. E. Dittmann, D. P. Fewer, B. A. Neilan, Cyanobacterial toxins: biosynthetic routes and evolutionary roots. *FEMS microbiology reviews* **37**, 23–43 (2013).
17. L. M. Miller, M. T. Mazur, S. M. McLoughlin, N. L. Kelleher, Parallel interrogation of covalent intermediates in the biosynthesis of gramicidin S using high-resolution mass spectrometry. *Protein science : a publication of the Protein Society* **14**, 2702–2712 (2005).
18. Du, L and Shen, B, Identification and characterization of a type II peptidyl carrier protein from the bleomycin producer *Streptomyces verticillus* ATCC 15003. *Chemistry & biology* **6**, 507–517 (1999).
19. J. H. Crosa, C. T. Walsh, Genetics and Assembly Line Enzymology of Siderophore Biosynthesis in Bacteria. *Microbiology and Molecular Biology Reviews* **66**, 223–249 (2002).
20. J. A. McIntosh, M. S. Donia, E. W. Schmidt, Ribosomal peptide natural products: bridging the ribosomal and nonribosomal worlds. *Nat. Prod. Rep.* **26**, 537 (2009).
21. T. J. Oman, van der Donk, Wilfred A., Follow the leader: the use of leader peptides to guide natural product biosynthesis. *Nature chemical biology* **6**, 9–18 (2010).
22. M. Trabi, J. S. Mylne, L. Sando, D. J. Craik, Circular proteins from *Melicocyctus* (Violaceae) refine the conserved protein and gene architecture of cyclotides. *Organic & biomolecular chemistry* **7**, 2378–2388 (2009).
23. Heather E. Hallen, Hong Luo, John S. Scott-Craig, and Jonathan D. Walton, Gene family encoding the major toxins of lethal *Amanita* mushrooms. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 19097–19101 (2007).
24. M. F. Freeman, C. Gurgui, M. J. Helf, B. I. Morinaka, A. R. Uria, N. J. Oldham, H. G. Sahl, S. Matsunaga, J. Piel, Metagenome Mining Reveals Polytheonamides as Posttranslationally Modified Ribosomal Peptides. *Science (New York, N.Y.)* **338**, 384–387 (2012).
25. D. H. Haft, M. K. Basu, D. A. Mitchell, Expansion of ribosomally produced natural products: a nitrile hydratase- and Nif11-related precursor family. *BMC biology* **8**, 70 (2010).
26. H. Wang, K. Sivonen, D. P. Fewer, Genomic insights into the distribution, genetic diversity and evolution of polyketide synthases and nonribosomal peptide synthetases. *Current opinion in genetics & development* **35**, 79–85 (2015).
27. M. Fischbach, C. T. Walsh, Assembly-line enzymology for polyketide and nonribosomal Peptide antibiotics: logic, machinery, and mechanisms. *Chemical reviews* **106**, 3468–3496 (2006).
28. J. Clardy, C. T. Walsh, Lessons from natural molecules. *Nature* **432**, 829–837 (2004).
29. P. M. Shih, D. Wu, A. Latifi, S. D. Axen, D. P. Fewer, E. Talla, A. Calteau, F. Cai, N. Tandeau de Marsac, R. Rippka, M. Herdman, K. Sivonen, T. Coursin, T. Laurent, L. Goodwin, M. Nolan, K. W. Davenport, C. S. Han, E. M. Rubin, J. A. Eisen, T. Woyke, M. Gugger, C. A. Kerfeld, Improving the coverage of the cyanobacterial phylum using diversity-driven genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 1053–1058 (2013).
30. R. I. Aminov, A brief history of the antibiotic era: lessons learned and challenges for the future. *Frontiers in microbiology* **1**, 1–7 (2010).
31. A. L. Demain, Importance of microbial natural products and the need to revitalize their discovery. *Journal of industrial microbiology & biotechnology* **41**, 185–201 (2014).
32. Ventola CL, The Antibiotic Resistance Crisis. *P&T* **40**, 277–283 (2015).
33. M. C. White, D. M. Holman, J. E. Boehm, L. A. Peipins, M. Grossman, S. J. Henley, Age and cancer risk: a potentially modifiable relationship. *American journal of preventive medicine* **46**, S7-15 (2014).

34. Challis GL, Hopwood DA, Synergy and contingency as driving forces for the evolution of multiple secondary metabolite production by *Streptomyces* species. *Proc Natl Acad Sci USA* **100**, 14555-14561 (2003).
35. L. Katz, R. H. Baltz, Natural product discovery: past, present, and future. *Journal of industrial microbiology & biotechnology* **43**, 155–176 (2016).
36. W. Wohlleben, Y. Mast, E. Stegmann, N. Ziemert, Antibiotic drug discovery. *Microbial biotechnology* **9**, 541–548 (2016).
37. M. L. T. Ang, P. Murima, K. Pethe, Next-generation antimicrobials: from chemical biology to first-in-class drugs. *Archives of pharmaceutical research* **38**, 1702–1717 (2015).
38. S. N. Cohen, A.C.Y. Chang, H. W. Boyer, R. B. Helling, Construction of Biologically Functional Bacterial Plasmids In Vitro. *Proc Natl Acad Sci USA* **70**, 3240–3244 (1973).
39. D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison, H. O. Smith, Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat Meth* **6**, 343–345 (2009).
40. N. Kouprina, V. Larionov, Transformation-associated recombination (TAR) cloning for genomics studies and synthetic biology. *Chromosoma*. 10.1007/s00412-016-0588-3 (2016).
41. H. Wang, Z. Li, R. Jia, J. Yin, A. Li, L. Xia, Y. Yin, R. Müller, J. Fu, A. F. Stewart, Y. Zhang, ExoCET: Exonuclease in vitro assembly combined with RecET recombination for highly efficient direct DNA cloning from complex genomes. *Nucleic Acids Research*. 10.1093/nar/gkx1249 (2017).
42. M. Xu, G. D. Wright, Heterologous expression-facilitated natural products' discovery in actinomycetes. *Journal of industrial microbiology & biotechnology* **46**, 415–431 (2019).
43. N. Ziemert, M. Alanjary, T. Weber, The evolution of genome mining in microbes - a review. *Natural product reports* **33**, 988–1005 (2016).
44. Stephen F. Altschu, Warren Gish, Webb Miller, Eugene W. Myers and David J. Lipman, Basic Local Alignment Search Tool. *Journal of molecular biology* **215**, 403–410 (1990).
45. R. D. Finn, J. Clements, S. R. Eddy, HMMER web server: interactive sequence similarity searching. *Nucleic Acids Research* **39**, W29-37 (2011).
46. T. Weber, K. Blin, S. Duddela, D. Krug, H. U. Kim, R. Brucoleri, S. Y. Lee, M. A. Fischbach, R. Müller, W. Wohlleben, R. Breitling, E. Takano, M. H. Medema, antiSMASH 3.0—a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res* **43**, W237-W243 (2015).
47. P. Cimermancic, M. H. Medema, J. Claesen, K. Kurita, L. C. Wieland Brown, K. Mavrommatis, A. Pati, P. A. Godfrey, M. Koehrsen, J. Clardy, B. W. Birren, E. Takano, A. Sali, R. G. Lington, M. A. Fischbach, Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. *Cell* **158**, 412–421 (2014).
48. J. Zucko, P. F. Long, D. Hranueli, J. Cullum, Horizontal gene transfer and gene conversion drive evolution of modular polyketide synthases. *Journal of industrial microbiology & biotechnology* **39**, 1541–1547 (2012).
49. N. Sélem-Mojica, C. Aguilar, K. Gutiérrez-García, C. E. Martínez-Guerrero, F. Barona-Gómez, EvoMining reveals the origin and fate of natural product biosynthetic enzymes. *Microbial genomics*. 10.1099/mgen.0.000260 (2019).
50. M. N. Thaker, W. Wang, P. Spanogiannopoulos, N. Waglechner, A. M. King, R. Medina, G. D. Wright, Identifying producers of antibacterial compounds by screening for antibiotic resistance. *Nature Biotechnology* **31**, 922–927 (2013).
51. E. Michta, K. Schad, K. Blin, R. Ort-Winklbaue, M. Röttig, O. Kohlbacher, W. Wohlleben, E. Schinko, Y. Mast, The bifunctional role of aconitase in *Streptomyces viridochromogenes* Tü494. *Environmental microbiology* **14**, 3203–3219 (2012).

52. J. W. Schopf, B.M. Packer, Early Archean (3.3-billion to 3.5-billion-year-old) microfossils from Warrawoona Group, Australia. *Science* **237**, 70–73 (1987).
53. J. C. Meeks, J. Elhai, T. Thiel, M. Potts, F. Larimer, J. Lamerdin, Predki, P and Atlas, R, An overview of the genome of *Nostoc punctiforme*, a multicellular, symbiotic cyanobacterium. *Photosynthesis Research*, 85–106 (2001).
54. S. Mazard, A. Penesyan, M. Ostrowski, I. T. Paulsen, S. Egan, Tiny Microbes with a Big Impact: The Role of Cyanobacteria and Their Metabolites in Shaping Our Future. *Marine drugs* **14** (2016).
55. R. W. Castenholz, G. M. Garrity, D. R. Boone, Phylum BX. Cyanobacteria. *Bergey's manual of systematic bacteriology 2nd edn, vol.1*, 473–487 (2001).
56. J. Demay, C. Bernard, A. Reinhardt, B. Marie, Natural Products from Cyanobacteria: Focus on Beneficial Activities. *Marine drugs* **17** (2019).
57. R. B. Dixit, M. R. Suseela, Cyanobacteria: potential candidates for drug discovery. *Antonie van Leeuwenhoek* **103**, 947–961 (2013).
58. J. C. Meeks, E. L. Campbell, M. L. Summers, F. C. Wong, Cellular differentiation in the cyanobacterium *Nostoc punctiforme*. *Archives of microbiology* **178**, 395–403 (2002).
59. R. W. Castenholz, J. B. Waterbury, Oxygenic photosynthetic bacteria. Group I. Cyanobacteria. *Bergey's manual of systematic bacteriology vol. 3*, 1710–1789 (1989).
60. D. G. Adams, P. S. Duggan, Heterocyst and akinete differentiation in cyanobacteria. *New Phytol.* **144**, 3–33 (1999).
61. M. Ekman, S. Picossi, E. L. Campbell, J. C. Meeks, E. Flores, A *Nostoc punctiforme* sugar transporter necessary to establish a Cyanobacterium-plant symbiosis. *Plant physiology* **161**, 1984–1992 (2013).
62. A. Liaimer, H. Jenke-Kodama, K. Ishida, K. Hinrichs, J. Stangeland, C. Hertweck, E. Dittmann, A polyketide interferes with cellular differentiation in the symbiotic cyanobacterium *Nostoc punctiforme*. *Environmental Microbiology Reports* **3**, 550–558 (2011).
63. S. W. Hunsucker, K. Klage, S. M. Slaughter, M. Potts, R. F. Helm, A preliminary investigation of the *Nostoc punctiforme* proteome. *Biochemical and biophysical research communications* **317**, 1121–1127 (2004).
64. L. Rouhiainen, J. Jokela, D. P. Fewer, M. Urmann, K. Sivonen, Two alternative starter modules for the non-ribosomal biosynthesis of specific anabaenopeptin variants in *Anabaena* (Cyanobacteria). *Chemistry & biology* **17**, 265–273 (2010).
65. A. N. Rai, E. Söderbäck, B. Bergman, Cyanobacterium-plant symbioses. *New Phytol.*, 449–481 (2000).
66. A. Liaimer, J. B. Jensen, E. Dittmann, A Genetic and Chemical Perspective on Symbiotic Recruitment of Cyanobacteria of the Genus *Nostoc* into the Host Plant *Blasia pusilla* L. *Frontiers in microbiology* **7**, 1693 (2016).
67. A. N. Eily, K. M. Pryer, F.-W. Li, A first glimpse at genes important to the *Azolla*–*Nostoc* symbiosis. *Symbiosis* **78**, 149–162 (2019).
68. D. Warshan, A. Liaimer, E. Pederson, S.-Y. Kim, N. Shapiro, T. Woyke, B. Altermark, K. Pawlowski, P. D. Weyman, C. L. Dupont, U. Rasmussen, Genomic Changes Associated with the Evolutionary Transitions of *Nostoc* to a Plant Symbiont. *Molecular biology and evolution* **35**, 1160–1175 (2018).
69. K. M. Usher, B. Bergman, J. A. Raven, Exploring Cyanobacterial Mutualisms. *Annu. Rev. Ecol. Evol. Syst.* **38**, 255–273 (2007).

- 70.J. Rikkinen, "Cyanobacteria in Terrestrial Symbiotic Systems" in *Modern Topics in the Phototrophic Prokaryotes: Environmental and Applied Aspects*, P. C. Hallenbeck, Ed. (Springer International Publishing, 2017), pp. 243–294.
- 71.A. Liaimer, E. J. N. Helfrich, K. Hinrichs, A. Guljamow, K. Ishida, C. Hertweck, E. Dittmann, Nostopeptolide plays a governing role during cellular differentiation of the symbiotic cyanobacterium *Nostoc punctiforme*. *Proceedings of the National Academy of Sciences of the United States of America* **112**, 1862–1867 (2015).
- 72.H. Wang, Z. Li, R. Jia, Y. Hou, J. Yin, X. Bian, A. Li, R. Muller, A. F. Stewart, J. Fu, Y. Zhang, RecET direct cloning and Redalpha recombineering of biosynthetic gene clusters, large operons or single genes for heterologous expression. *Nature protocols* **11**, 1175–1190 (2016).
- 73.R. A. Levin, S. K. Farrand, M. P. Gordon, E. W. Nester, Conjugation in *Agrobacterium tumefaciens* in the Absence of Plant Tissue. *Journal of bacteriology* **127**, 1331–1336 (1976).
- 74.C. P. Wolk, A. Vonshak, P. Kehoe, J. Elhai, Construction of shuttle vectors capable of conjugative transfer from *Escherichia coli* to nitrogen-fixing filamentous cyanobacteria. *PNAS*, 1561–1565 (1984).
- 75.R. Rippka, J. Deruelles, J. B. Waterbury, Herdman, Michael and Stanier, Roger Y., Generic Assignments, Strain Histories and Properties of Pure Cultures of Cyanobacteria. *Journal of General Microbiology* **111**, 1–61 (1979).
- 76.Y. Pörs, A. Wüstenberg, R. Ehwald, A Batch Culture Method for Microalgae and Cyanobacteria with CO<sub>2</sub> Supply Through Polyethylene Membranes. *Journal of phycology* **46**, 825–830 (2010).
- 77.A. Guljamow, M. Kreische, K. Ishida, A. Liaimer, B. Altermark, L. Bähr, C. Hertweck, R. Ehwald, E. Dittmann, High-density cultivation of terrestrial *Nostoc* strains leads to reprogramming of secondary metabolome. *Applied and environmental microbiology*. 10.1128/AEM.01510-17 (2017).
- 78.J. Sambrook, D. W. Russell, *Molecular Cloning - A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 2001).
- 79.Y. Cai, C. P. Wolk, Use of a Conditionally Lethal Gene in *Anabaena* sp. Strain PCC 7120 To Select for Double Recombinants and To Entrap Insertion Sequences. *Journal of bacteriology* **6**, 3138–3145 (1990).
- 80.D. Dehm, J. Krumbholz, M. Baunach, V. Wiebach, K. Hinrichs, A. Guljamow, T. Tabuchi, H. Jenke-Kodama, R. D. Süßmuth, E. Dittmann, Unlocking the Spatial Control of Secondary Metabolism Uncovers Hidden Natural Product Diversity in *Nostoc punctiforme*. *ACS chemical biology* **14**, 1271–1279 (2019).
- 81.U. K. Laemmli, Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680–685 (1970).
- 82.M. H. Medema, K. Blin, P. Cimermancic, V. de Jager, P. Zakrzewski, M. A. Fischbach, T. Weber, E. Takano, R. Breitling, antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Research* **39**, 46 (2011).
- 83.K. Blin, T. Wolf, M. G. Chevrette, X. Lu, C. J. Schwalen, S. A. Kautsar, H. G. Suarez Duran, E. L. C. de Los Santos, H. U. Kim, M. Nave, J. S. Dickschat, D. A. Mitchell, E. Shelest, R. Breitling, E. Takano, S. Y. Lee, T. Weber, M. H. Medema, antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Research* **45**, W36-W41 (2017).
- 84.T. A. Black, C. P. Wolk, Analysis of a Het- mutation in *Anabaena* sp. strain PCC 7120 implicates a secondary metabolite in the regulation of heterocyst spacing. *J. Bacteriol.* **176**, 2282–2292 (1994).
- 85.G. C. A. Amos, T. Awakawa, R. N. Tuttle, A.-C. Letzel, M. C. Kim, Y. Kudo, W. Fenical, B. S Moore, P. R. Jensen, Comparative transcriptomics as a guide to natural product discovery and biosynthetic



- gene cluster functionality. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E11121–E11130 (2017).
86. S. D. Bentley, K. F. Chater, A.-M. Cerdeño-Tárraga, G. L. Challis, N. R. Thomson, K. D. James, D. E. Harris, M. A. Quail, H. Kieser, D. Harper, A. Bateman, S. Brown, G. Chandra, C. W. Chen, M. Collins, A. Cronin, A. Fraser, A. Goble, J. Hidalgo, T. Hornsby, S. Howarth, C.-H. Huang, T. Kieser, L. Larke, L. Murphy, K. Oliver, S. O'Neil, E. Rabbinowitsch, M.-A. Rajandream, K. Rutherford, S. Rutter, K. Seeger, D. Saunders, S. Sharp, R. Squares, S. Squares, K. Taylor, T. Warren, A. Wietzorrek, J. Woodward, B. G. Barrell, J. Parkhill, D. A. Hopwood, Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). *Nature* **417**, 141–147 (2002).
87. J. P. Gomez-Escribano, M. J. Bibb, *Streptomyces coelicolor* as an expression host for heterologous gene clusters. *Methods in enzymology* **517**, 279–300 (2012).
88. L. Laureti, L. Song, S. Huang, C. Corre, P. Leblond, G. L. Challis, B. Aigle, Identification of a bioactive 51-membered macrolide complex by activation of a silent polyketide synthase in *Streptomyces ambofaciens*. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 6258–6263 (2011).
89. T. A. Knappe, U. Linne, S. Zirah, S. Rebuffat, X. Xie, M. A. Marahiel, Isolation and structural characterization of capistruin, a lasso peptide predicted from the genome sequence of *Burkholderia thailandensis* E264. *Journal of the American Chemical Society* **130**, 11446–11454 (2008).
90. J. Tian, H. Chen, Z. Guo, N. Liu, J. Li, Y. Huang, W. Xiang, Y. Chen, Discovery of pentangular polyphenols hexaricins A-C from marine *Streptosporangium* sp. CGMCC 4.7309 by genome mining. *Applied microbiology and biotechnology* **100**, 4189–4199 (2016).
91. F. Guo, S. Xiang, L. Li, B. Wang, J. Rajasärkkä, K. Gröndahl-Yli-Hannuksela, G. Ai, M. Metsä-Ketelä, K. Yang, Targeted activation of silent natural product biosynthesis pathways by reporter-guided mutant selection. *Metabolic engineering* **28**, 134–142 (2015).
92. S.-H. Xiang, J. Li, H. Yin, J.-T. Zheng, X. Yang, H.-B. Wang, J.-L. Luo, H. Bai, K.-Q. Yang, Application of a double-reporter-guided mutant selection method to improve clavulanic acid production in *Streptomyces clavuligerus*. *Metabolic engineering* **11**, 310–318 (2009).
93. F. Xu, Y. Wu, C. Zhang, K. M. Davis, K. Moon, L. B. Bushin, M. R. Seyedsayamdost, A genetics-free method for high-throughput discovery of cryptic microbial metabolites. *Nature chemical biology* **15**, 161–168 (2019).
94. M. R. Seyedsayamdost, High-throughput platform for the discovery of elicitors of silent bacterial gene clusters. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 7266–7271 (2014).
95. Ho-Sung Yoon and James W Golden, Heterocyst Pattern Formation Controlled by a Diffusible Peptide. *Science*.
96. S. A. West, G. A. Cooper, Division of labour in microorganisms: an evolutionary perspective. *Nature reviews. Microbiology* **14**, 716–723 (2016).
97. Z. Zhang, D. Claessen, D. E. Rozen, Understanding Microbial Divisions of Labor. *Frontiers in microbiology* **7**, 2070 (2016).
98. D. M. Wloch-Salamon, R. M. Fisher, B. Regenbreg, Division of labour in the yeast: *Saccharomyces cerevisiae*. *Yeast (Chichester, England)* **34**, 399–406 (2017).
99. A. Vinck, M. Terlouw, W. R. Pestman, E. P. Martens, A. F. Ram, C. A. M. J. J. van den Hondel, H. A. B. Wösten, Hyphal differentiation in the exploring mycelium of *Aspergillus niger*. *Molecular microbiology* **58**, 693–699 (2005).

100. L. Huo, J. J. Hug, C. Fu, X. Bian, Y. Zhang, R. Müller, Heterologous expression of bacterial natural product biosynthetic pathways. *Natural product reports*. 10.1039/c8np00091c (2019).
101. W. Li, X. Zhou, P. Lu, Bottlenecks in the expression and secretion of heterologous proteins in *Bacillus subtilis*. *Research in microbiology* **155**, 605–610 (2004).
102. S. E. Ongley, X. Bian, Y. Zhang, R. Chau, W. H. Gerwick, R. Müller, B. A. Neilan, High-titer heterologous production in *E. coli* of lyngbyatoxin, a protein kinase C activator from an uncultured marine cyanobacterium. *ACS chemical biology* **8**, 1888–1893 (2013).
103. T. Liu, R. Mazmouz, S. E. Ongley, R. Chau, R. Pickford, J. N. Woodhouse, B. A. Neilan, Directing the Heterologous Production of Specific Cyanobacterial Toxin Variants. *ACS chemical biology* **12**, 2021–2029 (2017).
104. A. Cullen, L. A. Pearson, R. Mazmouz, T. Liu, A. H. Soeriyadi, S. E. Ongley, B. A. Neilan, Heterologous expression and biochemical characterisation of cyanotoxin biosynthesis pathways. *Natural product reports* **36**, 1117–1136 (2019).
105. D. Ferreira, F. Garcia-Pichel, Mutational Studies of Putative Biosynthetic Genes for the Cyanobacterial Sunscreen Scytonemin in *Nostoc punctiforme* ATCC 29133. *Frontiers in microbiology* **7**, 735 (2016).
106. D. G. Adams, P. S. Duggan, Cyanobacteria-bryophyte symbioses. *Journal of Experimental Botany* **59**, 1047–1058 (2008).
107. D. Warshan, J. L. Espinoza, R. K. Stuart, R. A. Richter, S.-Y. Kim, N. Shapiro, T. Woyke, N. C. Kyrpides, K. Barry, V. Singan, E. Lindquist, C. Ansong, S. O. Purvine, H. M. Brewer, P. D. Weyman, C. L. Dupont, U. Rasmussen, Feathermoss and epiphytic *Nostoc* cooperate differently: Expanding the spectrum of plant-cyanobacteria symbiosis. *The ISME journal*. 10.1038/ismej.2017.134 (2017).
108. X. Yang, van der Donk, Wilfred A., Post-translational Introduction of D-Alanine into Ribosomally Synthesized Peptides by the Dehydroalanine Reductase NpnJ. *Journal of the American Chemical Society* **137**, 12426–12429 (2015).
109. B. I. Morinaka, E. Lakis, M. Verest, M. J. Helf, T. Scalvenzi, A. L. Vagstad, J. Sims, S. Sunagawa, M. Gugger, J. Piel, Natural noncanonical protein splicing yields products with diverse  $\beta$ -amino acid residues. *Science* **359**, 779 (2018).
110. Q. Zhang, X. Yang, H. Wang, van der Donk, Wilfred A., High divergence of the precursor peptides in combinatorial lanthipeptide biosynthesis. *ACS chemical biology* **9**, 2686–2694 (2014).
111. Y. Shi, X. Yang, N. Garg, van der Donk, Wilfred A., Production of lantipeptides in *Escherichia coli*. *Journal of the American Chemical Society* **133**, 2338–2341 (2011).

## 6 Appendix

### 6.1 List of Figures

- Figure 1 **General organization of a PKS assembly line.** A: Schematic display of a PKS assembly line. AT: Acyltransferase; ACP: Acyl carrier protein; KS: Ketosynthase; KR: Ketoreductase; DH: Dehydratase; ER: Enoyl reductase; TE: Thioesterase. Adopted from Kehr, et al., 2011. *Beilstein J Org Chem.* 7:1622–1635. B: Examples of chemical structures of PKS derived compounds (13)....2
- Figure 2 **General organization of a NRPS assembly line.** A: Schematic display of a NRPS assembly line. C: Condensation domain; A: Adenylation domain; PCP: Peptidyl carrier protein; KR: Ketoreductase; MT: Methyltransferase; E: Epimerase; TE: Thioesterase; PPT: Phosphopantetheinyl Transferase. Adopted from Kehr, et al., 2011. *Beilstein J Org Chem.* 7:1622–1635. B: Examples of chemical structures of NRPS derived compounds (16, 17).....3
- Figure 3 **Schematic representation of a RiPP gene cluster and precursor peptide features. A: Process of RiPP maturation.** Adopted from Kehr, et al., 2011. *Beilstein J Org Chem.* 7:1622–1635. B: Simplified presentation of two mature RiPP products (9).....5
- Figure 4 Fluorescence-microscopic picture of *N. punctiforme* PCC73102. .... 11
- Figure 5 **Schematic representation of the various cell types *N. punctiforme* can differentiate including differentiation triggering conditions.** HIF: Hormogonia inducing factors; HRF: Hormogonia reducing factors. Adopted from Liaimer, et al., 2015. *PNAS.* 112:1862–1867. .... 12
- Figure 6 **Chemical structures of already described natural products produced by *N. punctiforme*...** 13
- Figure 7 **Overview of various potential symbiotic hosts of *N. punctiforme* including the different modes of symbiotic interaction established with the respective host.** A1: *Blasia pusilla* (source: photo courtesy of Anton Liaimer), A2: Fluorescence micrograph of *Blasia pusilla* colonized by *N. punctiforme* (source: Warshan et al. 2018, *Mol. Biol. Evol.* 35(5):1160–1175); B1: *Macrozamia communis* (source: Photo courtesy of Missouri Botanical Garden), B2: Coralloid root section with highlighted symbiotic tissue (source: Rai et al., 2000, *New Phytol.* (2000), 147, 449–481); C1: *Gunnera manicata* (source: Photo courtesy of Missouri Botanical Garden), C2: *Gunnera manicata* stem gland colonized by *N. punctiforme* (source: Warshan et al. 2018, *Mol. Biol. Evol.* 35(5):1160–1175); D1: Micrograph of *Geosiphon pyriformis* showing mycelial structures and bladder bodies (source: Wolf, E., 2002, Dissertation), D2: Schematic of a bladder body with intracellular cyanobacterial filaments (source: Courtesy of Schüßler, A, [http://www.geosiphon.de/geosiphon\\_home.html](http://www.geosiphon.de/geosiphon_home.html), opened 01.09.2018)..... 15
- Figure 8 **Potential NRPS BGCs identified in the *N. punctiforme* genome using AntiSMASH.** Genes colored in light blue highlight NRPS genes, purple highlights PKS genes, yellow highlights transport related genes and grey indicates putative genes of unknown function. Gene annotation resemble the “NPun”-type annotations conducted and published by the authors of the original genome sequencing data in 2008, later detected and added genes follow the “RS”-type annotation nomenclature..... 45
- Figure 9 **Potential PKS BGCs identified in the *N. punctiforme* genome using AntiSMASH.** Genes colored in light blue highlight NRPS genes, purple highlights PKS genes, dark blue highlights other biosynthetic genes, yellow highlights transport related genes, green highlights regulatory genes and grey indicates putative genes of unknown function. Gene annotation resemble the “NPun”-type annotations conducted and published by the authors of the original genome sequencing data in 2008, later detected and added genes follow the “RS”-type annotation nomenclature..... 45
- Figure 10 **Potential NRPS-PKS hybrid BGCs identified in the *N. punctiforme* genome using AntiSMASH.** Genes colored in light blue highlight NRPS genes, purple highlights PKS genes, dark blue highlights other biosynthetic genes, yellow highlights transport related genes and grey indicates putative genes of unknown function. Gene annotation resemble the “NPun”-type annotations conducted and published by the authors of the original genome sequencing data in 2008, later detected and added genes follow the “RS”-type annotation nomenclature..... 46
- Figure 11 **Potential RiPP BGCs identified in the *N. punctiforme* genome using AntiSMASH.** Genes colored in light red highlight RiPP precursor genes, dark red highlights RiPP tailoring genes, yellow highlights transport related genes, green highlights regulatory genes and grey indicates putative genes of unknown function. Gene annotation resemble the “NPun”-type annotations

	conducted and published by the authors of the original genome sequencing data in 2008, later detected and added genes follow the “RS”-type annotation nomenclature.....	46
Figure 12	<b>Schematic workflow for the generation of a <i>N. punctiforme</i> reporter strain library for selected secondary metabolite BGCs.</b> .....	47
Figure 13	<b>Vector map of final precursor plasmid pDD001.</b> .....	49
Figure 14	<b>Time course of an <i>N. punctiforme</i> selection process after transfer on BG11<sub>0</sub>-agar plates supplemented with 2 µg/ml streptomycin.</b> .....	50
Figure 15	<b>Result of a <i>N. punctiforme</i> reporter strain generation confirmation PCR including all controls.</b> Primers used were D021(S1) and D022(S2). Expected product sizes: pDD011 - 1571bp; pDD008 - 1485 bp; pDD005 - 1704 bp.....	51
Figure 16	<b>Schematic representation of RNA deep-sequencing experimental workflow.</b> .....	52
Figure 17	<b>Full gene cluster transcription levels of <i>N. punctiforme</i>'s NRPS BGCs Nos, Apt, NRPS1 and NRPS2.</b> Dashed grey line indicates empirical threshold for actively expressed genes of 20 TPM. Each curve depicts the transcription level of a single gene of the respective BGC at each timepoint. The color scheme resembles a heatmap from dark blue to light brown, darker genes are located more upstream towards the start of the gene cluster than lighter genes. Detailed plots including gene legends are summarized in appendix X section 0.....	53
Figure 18	<b>Full gene cluster transcription levels of <i>N. punctiforme</i>'s PKS BGCs PKS1, PKS2 and PKS3.</b> Dashed grey line indicates empirical threshold for actively expressed genes of 20 TPM. Each curve depicts the transcription level of a single gene of the respective BGC at each timepoint. The color scheme resembles a heatmap from dark blue to light brown, darker genes are located more upstream towards the start of the gene cluster than lighter genes. Detailed plots including gene legends are summarized in appendix section 0.....	54
Figure 19	<b>Full gene cluster transcription levels of <i>N. punctiforme</i>'s NRPS-PKS hybrid BGCs PKS4 and PKS5.</b> Dashed grey line indicates empirical threshold for actively expressed genes of 20 TPM. Each curve depicts the transcription level of a single gene of the respective BGC at each timepoint. The color scheme resembles a heatmap from dark blue to light brown, darker genes are located more upstream towards the start of the gene cluster than lighter genes. Detailed plots including gene legends are summarized in appendix section 0. ....	55
Figure 20	<b>Full gene cluster transcription levels of <i>N. punctiforme</i>'s RiPP BGCs RiPP1a, RiPP1b, RiPP3, RiPP4, RiPP5, RiPP6 and Mvd.</b> Dashed grey line indicates empirical threshold for actively expressed genes of 20 TPM. Each curve depicts the transcription level of a single gene of the respective BGC at each timepoint. The color scheme resembles a heatmap from dark blue to light brown, darker genes are located more upstream towards the start of the gene cluster than lighter genes. Detailed plots including gene legends are summarized in appendix section 0.....	56
Figure 21	<b>Proof of concept for <i>N. punctiforme</i> transcriptional reporter strains.</b> Lefthand side shows schemes of the transformed reporter plasmid variants, righthand side shows confocal fluorescence micrographs of the corresponding <i>N. punctiforme</i> reporter mutants under standard low light conditions and ambient air. Upper row shows the <i>N. punctiforme</i> no-promoter negative control mutant strain and lower row shows the <i>N. punctiforme</i> PKS3 Npun_F3356 reporter mutant strain.....	58
Figure 22	<b><i>N. punctiforme</i> secondary metabolite BGCs showing no or faint expression under standard lowlight conditions.</b> A) Average TPM value of the gene that was selected for the construction of the reporter plasmid in the RNA deep-sequencing time course experiment. Dashed grey line indicates the empirical threshold of 20 TPM for actively expressed genes; B) confocal fluorescence microscopy picture of the respective <i>N. punctiforme</i> reporter strain under standard low light cultivation and ambient air. Left image shows CPF channel only and right image CFP/ Chl $\alpha$ channels merged. Scale bar indicates 20 µm. ....	59
Figure 23	<b><i>N. punctiforme</i> secondary metabolite BGCs showing a constitutive expression under standard lowlight conditions.</b> A) Average TPM value of the gene that was selected for the construction of the reporter plasmid in the RNA deep-sequencing time course experiment. Dashed grey line indicates the empirical threshold of 20 TPM for actively expressed genes; B) confocal fluorescence microscopy picture of the respective <i>N. punctiforme</i> reporter strain under standard low light cultivation and ambient air. Left image shows CPF channel only and right image CFP/ Chl $\alpha$ channels merged. Scale bar indicates 20 µm. ....	60

- Figure 24 ***N. punctiforme* secondary metabolite BGCs showing a spatially restricted expression pattern in long single filaments under standard lowlight conditions.** A) Average TPM value of the gene that was selected for the construction of the reporter plasmid in the RNA deep-sequencing time course experiment. Dashed grey line indicates the empirical threshold of 20 TPM for actively expressed genes; B) confocal fluorescence microscopy picture of the respective *N. punctiforme* reporter strain under standard low light cultivation and ambient air. Left image shows CPF channel only and right image CFP/ Chl $\alpha$  channels merged. Scale bar indicates 20  $\mu$ m. .... 61
- Figure 25 ***N. punctiforme* secondary metabolite BGCs showing a spatially restricted expression pattern in short isolated filaments under standard lowlight conditions.** A) Average TPM value of the gene that was selected for the construction of the reporter plasmid in the RNA deep-sequencing time course experiment. Dashed grey line indicates the empirical threshold of 20 TPM for actively expressed genes; B) confocal fluorescence microscopy picture of the respective *N. punctiforme* reporter strain under standard low light cultivation and ambient air. Left image shows CPF channel only and right image CFP/ Chl $\alpha$  channels merged. Scale bar indicates 10 or 20  $\mu$ m as specified above. .... 62
- Figure 26 ***N. punctiforme* secondary metabolite BGCs showing a spatially restricted expression pattern in short cell stretches or single cells within a filament under standard lowlight conditions.** A) Average TPM value of the gene that was selected for the construction of the reporter plasmid in the RNA deep-sequencing time course experiment. Dashed grey line indicates the empirical threshold of 20 TPM for actively expressed genes; B) confocal fluorescence microscopy picture of the respective *N. punctiforme* reporter strain under standard low light cultivation and ambient air. Left image shows CPF channel only and right image CFP/ Chl $\alpha$  channels merged. Scale bar indicates 20  $\mu$ m. .... 63
- Figure 27 ***N. punctiforme* CFP reporter strain fluorescence micrograph close-ups.** Labeled are distinct cell types to highlight differences in cell type restricted CFP expression. Depicted are CFP-channel/Chl $\alpha$ -channel merged images. V: vegetative cells; H: heterocysts; Hom: hormogonia; pA: pre-akinetes; A: akinetes. Yellow scale bar indicates 20  $\mu$ m. .... 65
- Figure 28 **Comparison of HPLC chromatograms of equal amounts of aqueous cell extracts generated from either conventionally cultivated or high density cultivated *N. punctiforme*.** Upper trace: cell extract of a 30 days high density cultivation of *N. punctiforme*; lower trace: cell extract of a 30 days conventional cultivation. 1: nostopeptolide1052, 2: nostopeptolideA, 3: nostamideA, 4: anabaenopeptin NZ857. Sample processing and HPLC analysis conducted as described in methods sections 2.2.5.5 and 2.2.5.6. Data provided by Julia Krumbholz. .... 67
- Figure 29 **HPLC analysis of supernatant extracts of conventional and HD cultivated *N. punctiforme* cultures and consecutive MALDI-TOF mass spectrometry.** A: Comparative HPLC chromatograms of high density cultivated *N. punctiforme* and a control culture cultivated under standard low light conditions and ambient air. Increased production of known products nostopeptolide1052 (1), nostamideA (3) and anabaenopeptin NZ857 (4) and various undescribed compounds. B: MALDI-TOF mass spectrometric analysis of an HPLC fraction containing the undescribed peak highlighted in green. 5-11: potential RiPP products and theoretical mass differences correlated to distinct proteinogenic amino acids. Data provided by Julia Krumbholz, Martin Baunach and Vincent Wiebach. .... 68
- Figure 30 **Structure of novel microviridins N3-N3.** The 11-membered core peptide sequence is N-terminally extended by up to 7 amino acids (WPV-E-T-G-E-T-A) leading to the final sequence of microviridins N3 to N9. Structure confirmation achieved by various means of mass spectrometric analyses in combination with chemical transformation. Analyses conducted by Vincent Wiebach and Matrín Baunach. Detailed information summarized in Dehm et al 2019 (80). .... 69
- Figure 31 **Relative transcription levels of selected *N. punctiforme* secondary metabolite BGC genes of a high density cultivated culture in comparison to a conventionally cultivated culture.** Grey area highlights the fold-change range of 1 – 0, indicating a downregulation of the respective gene in comparison to the control. RNA extraction, qPCR sample preparation, RT primers and calculation of fold-changes with the  $\Delta\Delta$ Ct-method are described in section 2.2.4.7. High density culture sample provided by Julia Krumbholz. .... 69
- Figure 32 **Confocal fluorescence micrographs of high density cultivated and conventionally cultivated *N. punctiforme* reporter cultures.** Depicted are merged CFP-channel/Chl $\alpha$ -channel

- images. Both culture types were cultivated for 30 days and samples were examined using confocal fluorescence microscopy as described in section 2.2.6.1. Experiment conducted in collaboration with Julia Krumbholz. Yellow scalebar indicates 20  $\mu\text{m}$ ..... 70
- Figure 33 **Confocal fluorescence micrographs of high density supernatant treated and conventionally cultivated *N. punctiforme* reporter cultures.** Depicted are merged CFP-channel/Chl $\alpha$ -channel images. The treated cultures were cultivated for 7 days and samples were examined using confocal fluorescence microscopy as described in section 2.2.6.1. Experiment conducted in collaboration with Julia Krumbholz. Yellow scalebar indicates 20  $\mu\text{m}$ ..... 71
- Figure 34 **Confocal fluorescence micrographs of high density responsive *N. punctiforme* reporter strains cultivated with  $\text{KHCO}_3/\text{K}_2\text{CO}_3$  buffer filled dialysis bags to induce high  $\text{CO}_2$  availability.** Depicted are merged CFP-channel/Chl $\alpha$ -channel images. The treated cultures were cultivated for 3 days and samples were examined using confocal fluorescence microscopy as described in section 2.2.6.1. Yellow scalebar indicates 20  $\mu\text{m}$ . ..... 73
- Figure 35 **Confocal fluorescence micrographs of high density responsive *N. punctiforme* reporter strains cultivated under high illumination light intensity of 107  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$ .** Depicted are merged CFP-channel/Chl $\alpha$ -channel images. The treated cultures were cultivated for 24 h and samples were examined using confocal fluorescence microscopy as described in section 2.2.6.1. Yellow scalebar indicates 20  $\mu\text{m}$ ..... 74
- Figure 36 **Schematic overview of the *N. punctiforme* - *Blasia pusilla* chemical interaction experimental workflow.** For a detailed description see section 2.2.6.3. .... 75
- Figure 37 **Average fold-changes of three biological replicates of *N. punctiforme* treated with conditioned *B. pusilla* exudate for 1 h, 5h and 24 h relative to an untreated control culture.** Purple bars: average fold-change of three biological replicates after 1 h of incubation with *B. pusilla* exudate; green: average fold-change of three biological replicates after 5 h of incubation with *B. pusilla* exudate; red: average fold-change of three biological replicates after 24 h of incubation with *B. pusilla* exudate. Grey area highlights the fold-change range of 1 - 0, indicating a downregulation of transcription compared to the control sample. RNA isolation and RT qPCR preparation conducted as described in sections 2.2.4.2 to 0. .... 76
- Figure 38 **Schematic overview of the *N. punctiforme* - *Blasia pusilla* physical interaction experimental workflow.** For a detailed description see section 2.2.6.4. .... 77
- Figure 39 **Average fold-changes of three biological replicates of *N. punctiforme* co-cultivated with *B. pusilla* for 1 h, 5h and 24 h relative to a control culture.** Purple bars: average fold-change of three biological replicates after 1 h of incubation with *B. pusilla* exudate; green: average fold-change of three biological replicates after 5 h of incubation with *B. pusilla* exudate; red: average fold-change of three biological replicates after 24 h of incubation with *B. pusilla* exudate. RNA isolation and RT qPCR preparation conducted as described in sections 2.2.4.2 to 0. .... 78
- Figure 40 **Average fold-changes of three biological replicates of *N. punctiforme* co-cultivated with *B. pusilla* for 1 h, 5h and 24 h relative to a control culture.** Purple bars: average fold-change of three biological replicates after 1 h of incubation with *B. pusilla* exudate; green: average fold-change of three biological replicates after 5 h of incubation with *B. pusilla* exudate; red: average fold-change of three biological replicates after 24 h of incubation with *B. pusilla* exudate. Grey area highlights the fold-change range of 1 - 0, indicating a downregulation of transcription compared to the control sample. RNA isolation and RT qPCR preparation conducted as described in sections 2.2.4.2 to 0. .... 79
- Figure 41 ***N. punctiforme* RiPP4 CFP reporter mutant strain co-cultivated with *B. pusilla* or *B. pusilla* conditioned exudate for up to 24 h in comparison to an untreated control.** Physical interaction: co-cultivation with *B. pusilla* under standard low light conditions and ambient air; chemical interaction: co-cultivation with concentrated conditioned *B. pusilla* exudate (final concentration of 50% total culture volume); untreated control: *N. punctiforme* RiPP4 CFP reporter mutant cultivated under standard low light conditions and ambient air. A: confocal fluorescence micrographs depicting the signal captured in CFP channel after 1.5 h, 6 h and 24 h of cultivation; B: confocal fluorescence micrographs depicting the merged signals captured in CFP and Chl $\alpha$  channel after 1.5 h, 6 h and 24 h of cultivation. Yellow scalebar indicates 20  $\mu\text{m}$ ..... 80
- Figure 42 ***N. punctiforme* RiPP3 CFP reporter mutant strain co-cultivated with *B. pusilla* or *B. pusilla* conditioned exudate for up to 24 h in comparison to an untreated control.** Physical interaction: co-cultivation with *B. pusilla* under standard low light conditions and ambient air;

- chemical interaction: co-cultivation with concentrated conditioned *B. pusilla* exudate (final concentration of 50% total culture volume); untreated control: *N. punctiforme* RiPP4 CFP reporter mutant cultivated under standard low light conditions and ambient air. A: confocal fluorescence micrographs depicting the signal captured in CFP channel after 1.5 h, 6 h and 24 h of cultivation; B: confocal fluorescence micrographs depicting the merged signals captured in CFP and Chl $\alpha$  channel after 1.5 h, 6 h and 24 h of cultivation. Yellow scalebar indicates 20  $\mu$ m..... 81
- Figure 43 ***N. punctiforme* PKS5 CFP reporter mutant strain co-cultivated with *B. pusilla* or *B. pusilla* conditioned exudate for up to 24 h in comparison to an untreated control.** Physical interaction: co-cultivation with *B. pusilla* under standard low light conditions and ambient air; chemical interaction: co-cultivation with concentrated conditioned *B. pusilla* exudate (final concentration of 50% total culture volume); untreated control: *N. punctiforme* RiPP4 CFP reporter mutant cultivated under standard low light conditions and ambient air. A: confocal fluorescence micrographs depicting the signal captured in CFP channel after 1.5 h, 6 h and 24 h of cultivation; B: confocal fluorescence micrographs depicting the merged signals captured in CFP and Chl $\alpha$  channel after 1.5 h, 6 h and 24 h of cultivation. Yellow scalebar indicates 20  $\mu$ m..... 82
- Figure 44 **HPLC traces of culture supernatant samples of *N. punctiforme* with and without *B. pusilla* in co-cultivation.** Samples were harvested after 2, 6 and 24 h and prepared for HPLC analysis as described in section 2.2.5.5. HPLC program including the acetonitrile gradient applied as described in section 2.2.5.6..... 83
- Figure 45 **HPLC traces of cell extracts of *N. punctiforme* with and without *B. pusilla* in co-cultivation.** Samples were harvested after 2, 6 and 24 h and prepared for HPLC analysis as described in section 2.2.5.5. 1: Microviridin, 2: Nostopeptolide1052, 3: NostopeptolideA. HPLC program including the acetonitrile gradient applied as described in section 2.2.5.6..... 84
- Figure 46 **HPLC traces of cell extracts of *B. pusilla* and *N. punctiforme* in single culture in comparison to the 6 h co-cultivation cell extract of *N. punctiforme* with *B. pusilla*.** Sample processing conducted as described in section 2.2.5.5. HPLC program including applied acetonitrile gradient as described in section 2.2.5.6..... 85
- Figure 47 **MALDI-TOF mass spectrum of isolated peak 15 including RiPP3 and RiPP4 core peptides.** A: Mass spectrum of peak 15 cropped from 400 – 2000 Da. MALDI-TOF sample preparation and measurement conducted as described in sections 2.2.5.7 and 2.2.5.8. B: Schematic representation of the RiPP3 and RiPP4 core peptides and associated theoretical masses, annotations refer to the respective precursor genes..... 86
- Figure 48 **Map of the *N. punctiforme* NRPS-PKS hybrid BGC PKS4.** Genes colored in light blue highlight NRPS genes, purple highlights PKS genes, dark blue highlights other biosynthetic genes and grey indicates putative genes of unknown function. .... 87
- Figure 49 **Typical course of PKS4 mutagenesis approach via tri-parental mating.**..... 88
- Figure 50 **Map of the *N. punctiforme* NRPS-PKS hybrid BGC PKS5.** Genes colored in light blue highlight NRPS genes, purple highlights PKS genes, dark blue highlights other biosynthetic genes, yellow highlights transport related genes and grey indicates putative genes of unknown function. .... 90
- Figure 51 **Stepwise overview of the conducted cloning steps to assemble the *N. punctiforme* PKS5 heterologous expression plasmid in *E. coli*.** A: PKS5 BGC map highlighting the separation into three equally sized fragments for the assembly approach; B: Overview of the different cloning steps including the required parts and methods used for the assembly of the desired construct of the given step; C: Schematic overview of the resulting construct of each cloning step. .... 91
- Figure 52 **HPLC traces of cell extracts of *E. coli* GB05-MtaA harboring the PKS5 expression plasmid.** Black: induced PKS5 expression culture 1; Purple: induced PKS5 expression culture 2; Blue: non-induced control culture; Brown: no-plasmid control culture. HPLC program used as described in section 2.2.5.6..... 92
- Figure 53 **Schematic expression patterns of *N. punctiforme* secondary metabolite BGCs observed under standard laboratory conditions.** A: constitutive expression; B: no expression (silent); C: expression in long isolated filaments; D: expression in short isolated cell stretches; E: expression in single cells or short cell stretches within a filament. .... 97
- Figure 54 **Stressed phenotypes of *N. punctiforme* reporter mutants under various cultivation conditions.** Stressed cells are highlighted in red. RiPP6 and NRPS2 reporter mutants were cultivated under standard low light condition. PKS1 reporter was cultivated for 30 d in a high density cultivator (experimental procedure and data collection conducted by Julia Krumbholz).

NRPS1 reporter was cultivated for 7 d with a buffer-filled dialysis bag producing a 32 mbar partial pressure of CO<sub>2</sub> over the culture. Yellow scale bar indicates 20 μm..... 99

Figure 55 NRPS BGCs transcriptomics data with detailed legends..... 121

Figure 56 PKS BGCs transcriptomics data with detailed legends..... 121

Figure 57 RiPP BGCs transcriptomics data with detailed legends..... 122

## 6.2 List of Tables

Table 1 Genomic elements of *N. punctiforme* PCC73102 ..... 13

Table 2 List of organisms used in the current study. .... 17

Table 3 Accession numbers of all genetic elements of *N. punctiforme* PCC73102 used for this study. .... 18

Table 4 Selected *N. punctiforme* genes for the construction of transcriptional reporter strains and respective lengths of selected 5'UTRs. .... 48

Table 5 Target 5'UTR amplification primers and corresponding reporter plasmids. .... 49

Table 6 Overview of the integration plasmids constructed for *N. punctiforme*'s PKS4 BGC mutagenesis. .... 88

Table 7 List of all DNA oligos used in this study. .... 118

## 6.3 Table of DNA Oligos

Table 7 List of all DNA oligos used in this study.

Application	Name	Sequence 5'-3'
pDD001	D001	tcgatgataagctgtcaaacatgagaattcATGGTGAGCAAGGGCGAG
	D002	cgcaagaggcccttctctcaaaTTACTGTACAGCTCGTCCATG
pDD003	D005	tcgatgataagctgtcaaacatgagaattcTGCCAAGTTTTCCATAAAAC
	D006	acagctcctgcccttctcaccatTAGTATAGCAATCCTATTTAAATTGTAAAATTAG
pDD004	D007	tcgatgataagctgtcaaacatgagaattcATTTTTGGTTATGTATTTCTGAG
	D008	acagctcctgcccttctcaccatAAGTCTTTAATACACTAAAAGCAAC
pDD007	D013	tcgatgataagctgtcaaacatgagaattcGGTAGAGACTACTTGAGCAAAAACAAC
	D014	acagctcctgcccttctcaccatTTCCGCCTCAAGCAATCAG
pDD009	D017	tcgatgataagctgtcaaacatgagaattcATCTCAATAAATTGCAGCAC
	D018	acagctcctgcccttctcaccatAAATTAGTCCTTTTGGCGTTAC
Sequencing	D021(S1)	AGGATGACGATGAGCGCATT
Sequencing	D022(S2)	CGGCACTCGACAGAATTGGC
Sequencing	D023(S3)	GCTGAACCTTGTGGCCGTTTA
Sequencing	D036(S4)_pBR-Spez	TATGGAAAAACGCCAGCAAC
Sequencing	D037(S5)	TCTAAAGGGCAAAAAGTGAGT
Sequencing	D038(S6)_P15A-Spez	GGAGCCTATGGAAAAACGG
pDD011	D041	tcgatgataagctgtcaaacatgagaattcTTTTAGATAGGACTGACGC
	D042	acagctcctgcccttctcaccatCCGTCTATCCTCCGTTTT
pDD012	D043	acaactactggtgaggaagttaaagagcgctTTATTTGCCACTACTTGGTGATCTC
	D044	ttgctgcaggcatgcaagctttcgcggcgccTGAGGGAAGCGGTGATCGCC
pDD013	D045	acaactactggtgaggaagttaaagagcgctAGAAATCACTGTCTTTCCATTCC
	D046	gatcaccaaggttagtcggcaaataaATTTGAAAGGTTGAGGTTAATG
pDD014	D047	acttcggcgatcacccgcttcctcaTTCCGCCTCAAGCAATCAG
	D048	ctgcaggcatgcaagctttcgcggcgccGTCTGCTCTTTGGTAATCAGGC
pDD015	D049	acaactactggtgaggaagttaaagagcgctTCGCTAATGTCCAGTTTTCC
	D050	gatcaccaaggttagtcggcaaataaGGCAAAAATACAGGCTAGAAATATG
pDD016	D051	acttcggcgatcacccgcttcctcaTTCCGCCTCAAGCAATCAG
	D052	ctgcaggcatgcaagctttcgcggcgccGTCTGCTCTTTGGTAATCAGGC
pDD017	D053	acaactactggtgaggaagttaaagagcgctCATCAAAACCACGAAGTAGATTG
	D054	gatcaccaaggttagtcggcaaataaAGTCAATACATTGGCAAATAAG
pDD018	D055	acttcggcgatcacccgcttcctcaTTCCGCCTCAAGCAATCAG
	D056	ctgcaggcatgcaagctttcgcggcgccGTCTGCTCTTTGGTAATCAGGC
pDD019	D057	acttcggcgatcacccgcttcctcaTTTTCTGTTCATTCCAGTC
	D058	ctgcaggcatgcaagctttcgcggcgccTAAACCCCTTAAAATCCTG
Sequencing	D064(S7)	AGTAACCGGCAAAATCGC
Sequencing	D065(S8)	ATTGAAGATGAGGAGCGAT
pDD022	D066	ATTGTCAGGATTTAAAGGGTTTTAATGTCTTTAATTCTCAATATATTC



	D067	TTGCTGCAGGCATGCAAGCTTTCGCGGCGCCTACCATAACAAGGATAGAACCC
pDD023	D068	ACAACACTGGTGAGGAAAGTTAAAGAGCGCTTGGCGGTAGTAATACCAGATAC
	D069	GATCACCAGGTAGTCGGCAAATAAATGAAGTCTAAACAGTTTGTTTC
pDD025	D072	tcgatgataagctgtcaaacatgagaattcCTTGCATATATACTTTGTTTCAT
	D073	acagctcctgcacctgtcaccatATTACCTTACCTTGAATTACGG
PKS5 het. ex. assembly	D076	TCTGATGATAATGATGTTTGTAGACTCAGTTACCAGAAATTGATTTTTATTTTGTAGCTAAT GAATCAGTTATTACTTTTCATAGATCCTTCTCCTTTAGATC
	D077	TAGATTTGTAGAGGTTCCAATGCGTAATTTTAAAGTCTACAATCTTGGCTAAACAATG AATTTAAATGCCAATACTTTGGGCGCAGATCCGAAAACCCCAAGTTACG
	D078	ATGAAAAGTAATAACTGATTC
	D079	AAAGTATTGGCATTAAATTC
	D080	TAGATTTGTAGAGGTTCCAATGCGTAATTTTAAAGTCTACAATCTTGGCTAAACAATGAATTTAA TGCCAATACTTTGCTAAGAAAAAATTAATTTTAGTT
	D081	GCCCAGTCTTTCGACTGAGCCTTTCGTTTATTTGATGCTGGAGATCCTTAAGATCCGTAACCTGGGG TTTTCGGATCTGGCGCCAGGCTTTATTAGCGATCGCA
	D082	GATGAGGTTTTATTGCGGTTAATAACTTCAACTATATATCAACTATATATGTGTAGATGC GATCGCTAATAAAGCCTTAGATAAGTCTAAACTGGAT
D083	CAGTCTTTCGACTGAGCCTTTCGTTTTATTGATGCTGGAGATCCTTAAGATCCGTAACCTGGGGT TTTCGGATCTGGCGCTCAACGTACCTCTTAAAGAT	
PKS5 het. ex. Sequencing	D084_Check-5'Step1C21	CTATAGGTGAGATGCTGATT
	D085_Check-3'Step1C21	GGTTCATTGGTTTGTCAAG
	D086_Check-5'Step2C21	CAAGACCAACCATCACAAT
	D087_Check-3'Step2C21	TGCTCTACAAGTTTACTCAG
	D088_Check-5'Step3C21	GACTGTATAGTTGGTGGGAG
	D089_Check-3'Step3C21	AGATAATGAACAGCGTTTGG
PKS5 het. ex. assembly	D090_Frag2-only-fw	GCTAAGAAAAAATTAATTTTAGTT
	D091_Frag2-only-rv	AGGCTTTATTAGCGATCGCA
	D092_Frag3-only-fw	TTAGATAAGTCTAAACTGGAT
	D093-Frag3-only-rv	TCAACGTACCTCTTAAAGAT
	D094-OligoFW	CAATGAATTTAAATGCCAATACTTTCACGTGGATGAGGTTTTATTGCGGTTAATAACTTCAACTATATA TCAACTATAT
	D095-OligoRV	GTAACCTGGGGTTTTTCGGATCTGGCGCCAGGCTTTATTAGCGATCGCATCTACACATATATATAGTTGA TATATAGTTGA
	D096-OligoOUT-fw	CAATGAATTTAAATGCCAATAC
	D097-OligoOUT-rv	GTAACCTGGGGTTTTTCGG
	D098	AGGCTTTATTAGCGATCGCATCTACACATATATATAGTTGATATATAGTTG AAGTTATAACCGCAAATAAACCTCATGACGTCAGATCCTTCTCCTTTAGATC
D099	TACGCGAAGAACGGACTGTGCTTAGTTTATTCAGTGTGTTTACCAACAATTGA TACTCTTAAAGAGGTACGTTGAAGATCCGAAAACCCCAAGTTACG	
pDD027	D100	tcgatgataagctgtcaaacatgagaattcGTTGCTCTGAGCAACCGGT
	D101	acagctcctgcacctgtcaccatAAAATTTCTCCGTTCTTGTG
pDD031	D108	tcgatgataagctgtcaaacatgagaattcCCTAAAGATGATTACGATTAC
	D109	acagctcctgcacctgtcaccatATTTTATTCCGCAAAAAATAATC
pDD032	D110	tcgatgataagctgtcaaacatgagaattcATTTTGTCCAAATATTGATGTG
	D111	acagctcctgcacctgtcaccatCATTTCTCTGCTAATGATTAC
pDD033	D112	tcgatgataagctgtcaaacatgagaattcCACAACAGAAGTGAATTGTAG
	D113	acagctcctgcacctgtcaccatTTTATAGACACCTTATTTAAATAAC
pDD034-37	D114	ggatgaatggcagaaattcgatctagatACTACGCATGATCTGCATAC
	D115	atgtccctggcgaggggaagtatccagAGCTTGAAAACCTGGCAAGGG
Sequencing Sequencing	D116(S9)	GCTGGCGATTCAAGTTTCAT
	D117(S10)	CATCAGTGAAATCCAGCTGG
	D118_R6591_fw d	AGAATCCGAAAAGACCTAC
	D119_R6589_fw d	ATTTATTGTGCGTTTTGCC
	D120_R6589_re v	GAATGCTTGGAGAAGGAAAA

	D121_R6584_fw d	GACGGTAACAGCAAACACTAGA
	D122_R6584_re v	TCAAAGTCTTCCACAACAAC
Sequencing	D129(S11)	GTAACACAGTATTGCAAGG
Sequencing	D130(S12)	CCCCCACTCTATTGTAAC
Sequencing	D131(S13)	TTCTTCGCCCCGTTTTCA
RT qPCR PKS2	D132_RT_F3155	GCGGACTAGCTCATCAGACC
	D133_RT_F3155	GTACCATCCCCACAACCCAT
RT qPCR RiPP1b	D134_RT_F3224	CGAAAGAAGCAGTCATCAAA
	D135_RT_F3224	TTGCTTTTCTTGATCACTG
RT qPCR PKS3	D136_RT_F3356	AGAACGGGCGCTACTCTTTT
	D137_RT_F3356	TTTCTTGGTGGATAGGCGGG
RT qPCR PKS4	D138_RT_R345 2	ACTCTCAGGCAATGTTCCA
	D139_RT_R345 2	CCTGAAATTGACGCGCAGAT
RT qPCR RiPP4	D140_RT_F5045	GCGATCTCTCAAATGCTGGC
	D141_RT_F5045	CGTGTGATTGAGAGCGGCT
RT qPCR RiPP6	D148_RT- BF050-fw	AACCCTAAACAGGCAAGTG
	D149_RT- BF050-rev	CACCAGCAATCTGCTCAA
RT qPCR RiPP5	D150_RT- AF077-fw	GCTAGGAGCTACTGAGAACA
	D151_RT- AF077-rev	AGCATCTACCTCCTGAATCG
pDD043	D154	tcgatgataagctgtcaaacatgagaattcGGCTGTACAAAAGTAACAAAT
	D155	acagctcctcgccctgtcaccatGACTATTTTCTCTATTGATGAA
pDD044	D156	tcgatgataagctgtcaaacatgagaattcTTATAGCATTATATTGTGATATC
	D157	acagctcctcgccctgtcaccatAATTTCTCCTAACTGAAATTTAAAG
pDD045	D158	tcgatgataagctgtcaaacatgagaattcCAGTTAGACAATCTGTTCACT
	D159	acagctcctcgccctgtcaccatAGATTTTCTCTGAAATTAATTTGA
RT qPCR rnpB	rnpB_RT3_FW	GCGGTTGCAGATCAGTCATA
	rnpB_RT3_RV	TCTGTGGCACTATCCTCACG
F3224-CFP	F3224-fw	CTAGACGCGTGAATTCAGGATAGAATTA
	F3224-rv	CATGCATATGGTTGCTTTTCAGCGACTTC
PKS1	C3_fw	CGTGGTTGACTGGAGATGCT
	C3_rv	AAGCCTCTCGTGCCGTTTTA
PKS5	C21	GGGGAATGGAAAGCATGGGA
	C21	ATTAACGCCCTTCCCTGTG
Nos	NosA_RT4_fw	GTTTGCCTCTCTGCTGAAC
	NosA_RT4_rv	GCGGTAAGCAGGGTATCAA
Apt	apt_RT_fw	GAAATTGAGCGCTTTTGAG
	apt_RT_rv	GGCTAGTGACGCTCACATCA
NRPS1	C8_fw	AGCGGCAACATATTCCCAA
	C8_rv	ACGCCAACCTGCTCTATTTC
NRPS2	Cp_fw	CTCATGTGGGGTGCAGCTTA
	Cp_rv	CCTCAATCCAAGTCAGGCGT
RiPP1a	C10Li_fw	GGCAGAATTGGGAGGACGAA
	C10Li_rv	TCCCAAACCATCATTGAGCA
RiPP3	C11a_fw	AGCAGACATCATAGCTCCACT
	C11a_rv	GGGTGCAGAAAAGGGCTACA

## 6.4 Transcriptomics Diagrams including Gene Annotations

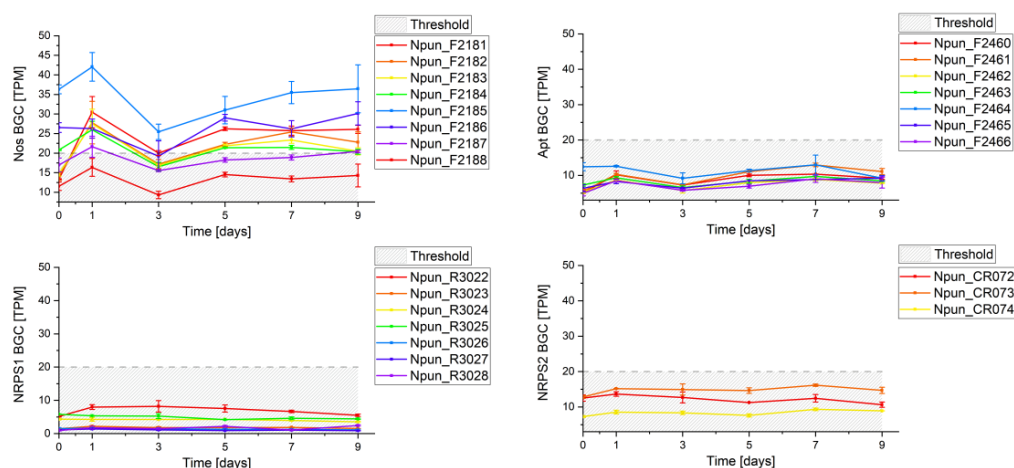


Figure 55 NRPS BGCs transcriptomics data with detailed legends.

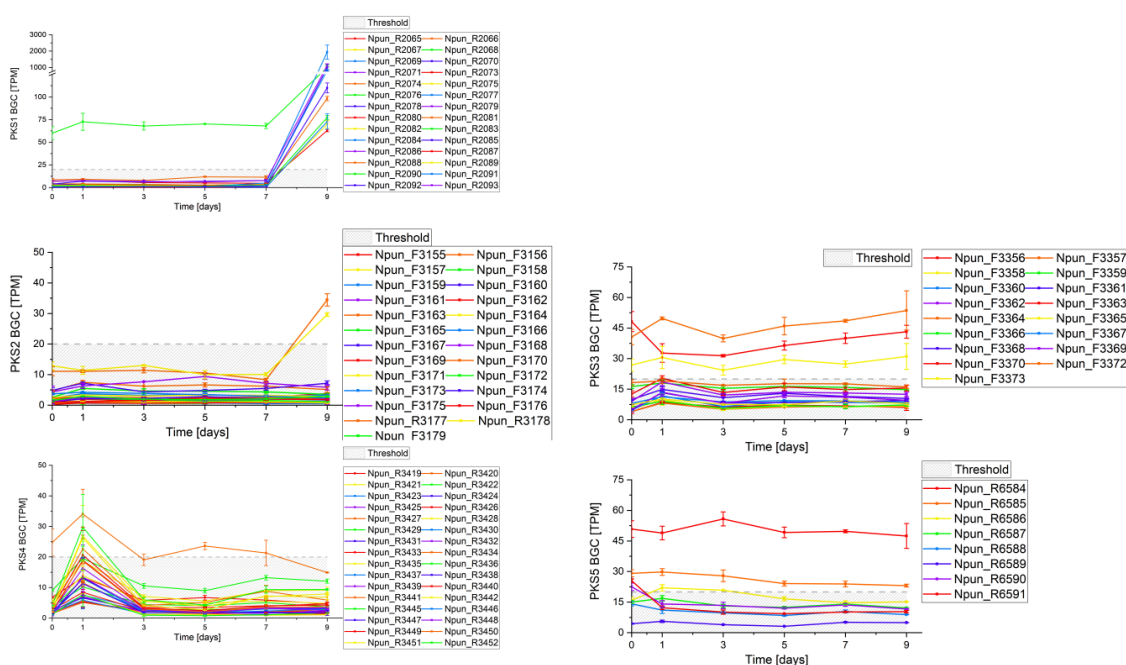


Figure 56 PKS BGCs transcriptomics data with detailed legends.

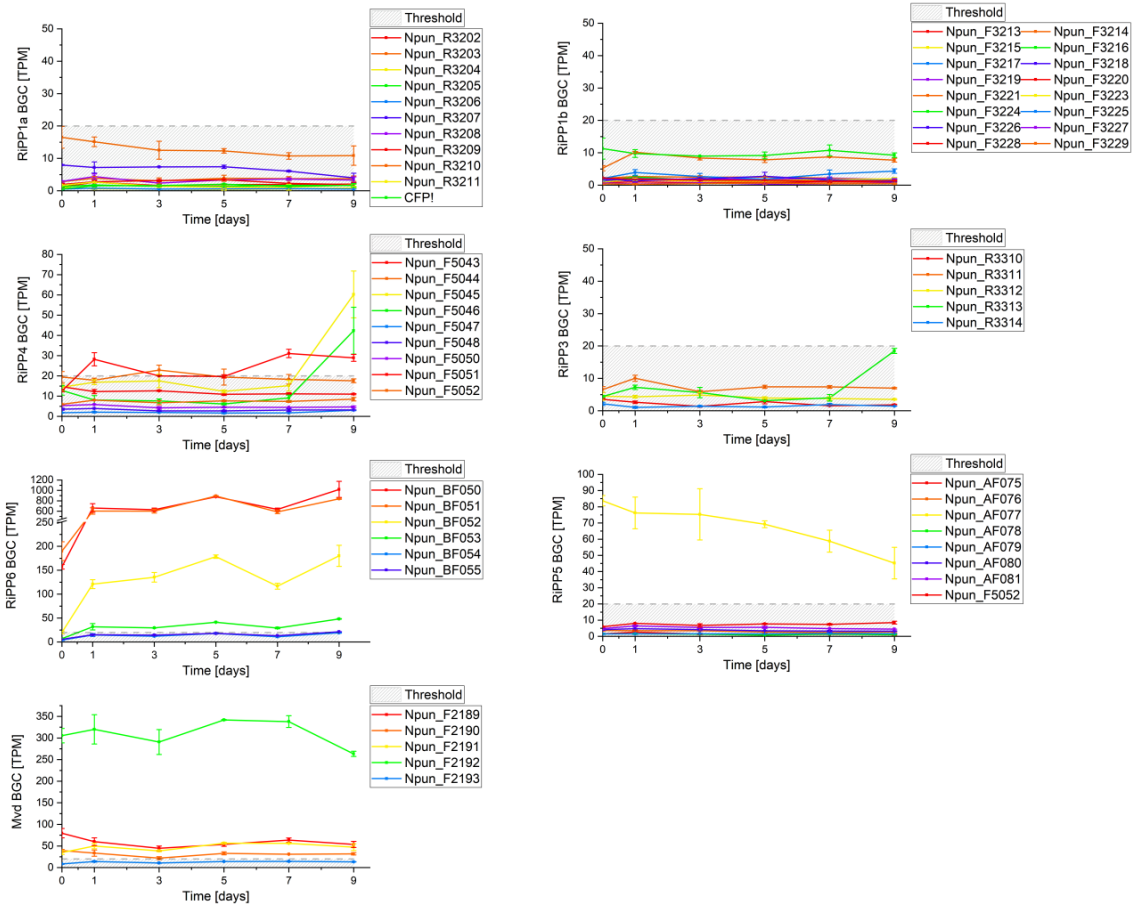


Figure 57 RiPP BGCs transcriptomics data with detailed legends.

## Publications and Conference Contributions

### Publications

D. Dehm, J. Krumbholz, M. Baunach, V. Wiebach, K. Hinrichs, A. Guljamow, T. Tabuchi, H. Jenke-Kodama, R. D. Süssmuth, E. Dittmann, Unlocking the Spatial Control of Secondary Metabolism Uncovers Hidden Natural Product Diversity in *Nostoc punctiforme*. *ACS chemical biology* **14**, 1271–1279 (2019).

J. Teikari, D. Dehm et al., Induction of cyanobacterial symbiosis (in preparation)

### Conference Contributions

#### September 2018

Genomic Mining of Plant-induced Secondary Metabolites in Symbiotic Cyanobacteria. *Poster presentation* at “Cyano 2018 Young Researcher Symposium”, Freiburg im Breisgau, Germany

#### August 2018

Genomic Mining of Plant-induced Secondary Metabolites in Symbiotic Cyanobacteria. *Oral presentation* at “Annual Conference of the Association for General and Applied Microbiology”, Frankfurt am Main, Germany.

#### September 2017

Genomic Mining of Plant-induced Secondary Metabolites in Symbiotic Cyanobacteria. *Poster presentation* at “Annual Conference of the Association for General and Applied Microbiology”, Tübingen, Germany.

#### August 2017

Genomic Mining of Plant-induced Secondary Metabolites in Symbiotic Cyanobacteria. *Poster presentation* at 10<sup>th</sup> European Conference on Molecular Biology of Cyanobacteria, Cluj-Napoca, Romania

### September 2016

Genomic Mining of Plant-induced Secondary Metabolites in Symbiotic Cyanobacteria. *Poster presentation* at “Annual Conference of the Association for General and Applied Microbiology”, Freiburg im Breisgau, Germany.

### June 2016

Biochemical Dissection of the Natural Diversification of Microcystin Provides Lessons for Synthetic Biology of NRPS. *Poster presentation* at 1<sup>st</sup> International Workshop on Cyanobacterial Natural Products, Center for Nuclear Energy in Agriculture, Piracicaba, SP, Brazil

## Acknowledgements

Although the successful completion of this study is the result of fruitful collaborations with many great people at the very first, I would like to express my gratefulness to Professor Elke Dittmann, not only for providing with me the opportunity to work on this exciting project in the first place, but also for the exhaustive scientific supervision, invaluable support, countless ideas and endless positivism until the very end.

Next, I would like to thank Professor Helge Bode and Dr. Dennis Nürnberg for reviewing this work.

My sincere appreciation also for all the external collaborators - Holger Jenke-Kodama, Anton Liaimer, Keishi Ishida and Vicent Wiebach, only to name a few - that shaped the outcome of this project with their valuable contributions. Special appreciation also for my collaborating colleagues Julia Krumbholz and Martin Baunach, thank you for the great and seamless teamwork at all times!

A big shout out to the whole MiBi crew! First of all, Katrin who gave me all the tools necessary when working with *Nostoc* (including the development of a very much needed 'gut-feeling') but also supporting me with open ears for all the work-related and not-so-much-work-related issues. Next, Arthur – thank you for all the helpful discussions and supportive questions, dealing with my written raw material and last but not least your enthusiasm for good music and volleyball! Big thanks also to Julia not only for crucial collaborations as a fellow *Nostoc*-buddy but also for the good time we had in the lab and on joint conference trips (and showing us the full spectrum of tea flavors)! Thank you, Markus for taking over Katrins spot in every imaginable perspective, keeping the engine running and on top of that providing our lab with many new technical upgrades. Annika, Jan-Christoph, Martin, Jonna, Tino, Sophie, Sabine, Emmanuel, Stella – you all contributed on various scientific and social levels to what finally made the last four years a great experience and an unforgettable time, thank you!!

Finally, but certainly not least, I want to express my deepest gratitude to my mum – if it wasn't for your unconditional support and full backing, I would have never been able to tread this path and become what I am today. This one's for you.

## Eigenständigkeitserklärung

Gemäß §9 sowie §12 der Promotionsordnung der Mathematisch-Naturwissenschaftlichen Fakultät I der Universität Potsdam in der Fassung vom 18. September 2013,

erkläre ich,

dass ich die vorliegende Arbeit selbständig angefertigt und ohne fremde Hilfe verfasst habe, keine außer den angegebenen Hilfsmitteln und Quellen dazu verwendet habe, die den benutzten Werken wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe und die Arbeit in der vorliegenden Fassung bisher an keiner anderen Hochschule eingereicht wurde.

---

Daniel Dehm, Potsdam, 19.4.2020