Small RNA control of denitrification and N_2O emissions in bacteria

Sophie Moeller
100157120

A thesis presented for the degree of Doctor of Philosophy
University of East Anglia
School of Biological Sciences
September 2023

Abstract

Anthropogenic influences have led to a shift in the nitrogen cycle resulting in an increase in microbial emissions of the potent greenhouse gas nitrous oxide (N_2O) into the atmosphere. As the genetic, physiological, and environmental factors regulating the microbial processes responsible for the production and consumption of N_2O are not fully understood, this represents a critical knowledge gap in the development of future mitigation strategies. Non-coding small RNAs (sRNAs) regulate a wide range of physiological processes in microorganisms that allow them to rapidly respond to changes in environmental conditions. These have predominantly been studied in a limited number of model organisms but with improvements in the techniques for sRNA discovery it is becoming increasingly clear that sRNAs play a crucial role in environmentally relevant pathways as well. For example, several sRNAs have been shown to control important enzymatic processes within the nitrogen cycle and many more have been identified in model nitrogen cycling organisms such as the model denitrifier *Paracoccus denitrificans*.

The discovery of the sRNA DenR has demonstrated the influence of a single sRNA on the denitrification pathway through interaction with a novel GntR-type regulator NirR, which in turn stalls denitrification at the stage of nitrite reduction. This study aims to further explore the role of this novel regulatory mechanism within the denitrification regulatory network. Furthermore, this thesis presents information on how to further unravel the sRNA regulatory network in the model organism *P. denitrificans* and show indications for the involvement of further sRNAs as well as the involvement of the RNA chaperone Hfq. To address knowledge gaps regarding the transcriptome during denitrification conditions, a global transcription start site analysis is presented which provides a solid foundation for further research into the role of sigma factors and promoter activity during complete and incomplete denitrification. Together these investigations highlight the importance of *P. denitrificans* as a biochemical and physiological model for denitrification to identify potential novel targets for mitigation strategies and combat climate change.

Contents

List of tables

List of figures

5. Structure-based sequence alignment of the C-terminal domains of the HutC-subfamily regulators including NirR
6. NirR test expression
7. Elution profile for NirR protein purification
8. NirR purification using high quantities of cell pellet
9. NirR purification with a reduced cell pellet
10. Bradford assay for the determination of NirR concentration post purification
11. Predicted HutC subfamily regulator recognition motif
12. *nirS* promoter including the overlapping NNR and NirR recognition sites

Chapter 5

1. Bacterial promoters recognized by the primary housekeeping sigma factor contain two conserved regions
2. Categorisation of TSS into primary, internal, and antisense based on their location on the genome
3. TSS loci across the P. denitrificans genome
4. Number of unique TSS present under CuH and CuL conditions
5. Benchmarking the presence if TSS across the *P. denitrificans* genome using the *nos* operon
6. Visualisation of TSS and sRNA locations alongside the *P. denitrificans* genome
7. TSS across the *nar* operon
8. TSS across the *nir* operon
9. TSS across the *nor* operon
10. TSS across the *nos* operon
11. Motif scores for the conserved motifs generated using Improbizer
12. Alignment of the *E. coli* and *P. denitrificans* RpoD genes

Chapter 6

1. Expanded overview of transcriptional regulators in *P. denitrificans*

List of abbreviations

(v/v)	Volume per volume
(w/v)	Weight per volume
ACDS	Decarbonylase/synthase complex
ANR	Anaerobic regulator of nitrite reductase
ATP	Adenosine triphosphate
bp	Base pairs
BSA	Bovine serum albumin
C	Carbon
cDNA	Complementary DNA
$CaCl_2$	Calcium chloride
CDS	Protein coding region
CFU	Colony forming units
CO_2	Carbon dioxide
Cu	Copper
CuH	Copper high
CuL	Copper low
Da	Dalton
DenR	Denitrification repressor
DNA	Deoxyribonucleic acid
FnrP	Fumarate and nitrate reduction protein
Hfq	Host factor for the replication of the bacteriophage Qβ
His	Histidine
HTH	Helix-turn-helix
IF	Inactivating factor

IPTG	Isopropyl β-D-1-thiogalactopyranoside
K	Lysine
Lsm	Sm-like protein
Min	Minute
mL	Millilitre
ML	Machine learning
mM	Millimolar
mm	Millimetre
mRNA	Messenger RNA
MSA	Multiple Sequence Alignment
N	Nitrogen
N_2	Dinitrogen
N_2O	Nitrous oxide
Nap	Periplasmic nitrate reductase
Nar	Membrane-bound nitrate reductase
NarR	Nitrate reductase regulator
Nas	Cytoplasmic assimilatory nitrate reductase
NH_3	Ammonia
NirR	Nitrite reductase repressor
NirS	Nitrite reductase
NNR	Nitrite reductase and nitric oxide reductase regulator
NO	Nitric oxide
NO_2^-	Nitrite
NO_3^-	Nitrate
Nor	Nitric oxide reductase

Nos	Nitrous oxide reductase
nt	nucleotide
O_2	Oxygen
OD	Optical density
ORF	Open reading frame
PCR	Polymerase chain reaction
R	Arginine
RBS	Ribosome binding site
RNA	Ribonucleic acid
RNAP	RNA polymerase
rRNA	ribosomal RNA
RNA-seq	RNA-sequencing
S	Sulfur
SD	Shine dalgarno
SDS	Sodium dodecyl sulphate
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Sec	Second
sRNA	small RNA
tRNA	transport RNA
T	Taurine
TSS	Transcription start site
UTR	Untranslated region
WT	Wild type
µL	Microlitre

Acknowledgements

Firstly, thank you to UEA for funding my studies and for providing the opportunity to present my work at national and international conferences and get involved in outreach events such as the Norwich Science Festival. I am grateful for my supervisory team Dr. Gary Rowley, Prof. David Richardson and Dr. Andrew Gates for providing invaluable expertise, help and advice during my time at UEA, especially through the difficult times caused by COVID-19. I would like to thank Gary specifically for his patience and his guidance. A big thank you to Claire Hews for her support and training when I started in the lab and to all other members of the Rowley lab for making day to day research more sociable.

A special thank you to Dr. Hayley Whitfield for pushing me, when I needed it the most and for being the best support in the lab as well as while writing my thesis. You are the best and I am very lucky to have had you annoy me about my progress more than I sometimes would've liked. I am also truly grateful for Luke Acton. Thank you, for your help and support with my bioinformatics attempts and for being my lab taxi, lockdown buddy and motivator.

To my Mum, Dad and Aaron, thank you for putting up with me during the stressful times and for always being a phone call away when needed. Thank you, mum, for proofreading and giving me advice, even though my science sometimes sounds like a foreign language to you. Also, thank you to Vicki and Mike for being my parents away from home. And finally, a big thanks to my friends in Cologne, who have motivated me to keep going in the final stages of my writing process. Thank you, Vera, for grabbing banana bread with me when the motivation was low and thank you, Jan, for going out for a run or a bike ride with me to clear my head. You are all the best!

1 Introduction

1.1 The biogeochemical nitrogen cycle

Nitrogen is an essential nutrient for sustaining all life forms on Earth. Nitrogen gas (N_2) constitutes 78% of the Earth's atmosphere. However, the majority of living organisms cannot access the gaseous form of nitrogen that is present in the atmosphere. Therefore, the biogeochemical nitrogen cycle is a cornerstone of all living processes and consists of a set of reactions by which nitrogen is converted into different organic and inorganic forms, circulating between atmosphere and biosphere. Nitrogen is able to form covalent bonds with carbon that are integral to the functioning of many organic biomolecules, and it also provides a potent source of electrochemical energy for biological metabolism. As a limiting nutrient it has affected biological evolution and ecology on Earth over time. The partial pressure of nitrogen gas in the atmosphere controls the degree of pressure-broadening of greenhouse gas absorption and thus surface temperature of the planet, making it habitable (Goldblatt et al., 2009). Atmospheric N_2 is only accessible to certain microorganisms, the N_2-fixing Bacteria and Archaea, which are estimated to biologically fix approximately 0.1% of the N_2 pool (Vitousek and Howarth, 1991). Nitrogen gas is a highly stable compound due to the strength of the triple bond between the two nitrogen atoms. Therefore, the fixation process requires eight electrons and at least sixteen ATP molecules (Bernhard, 2010). There is great physiological and phylogenetic diversity among nitrogen-fixing organisms: Some nitrogen-fixing organisms are free-living while others live in close symbiotic relationships with plants, some are aerobic while others are anaerobic. However, they all have a similar enzyme complex called nitrogenase that reduces N_2 to ammonia. The enzyme complex is highly sensitive to oxygen and is deactivated in its presence. The nitrogen fixing microorganisms possessing this key enzyme are called diazotrophs.

Once the largely un-reactive molecular nitrogen has been fixed by microorganisms, the resulting ammonium compounds are transformed into a wide range of amino acids and oxidized compounds and becomes available to plants and animals (Fig. 1.1). Only small quantities of atmospheric nitrogen are fixed by abiotic means including lightning and ultraviolet radiation. Over recent years it has become apparent that the biogeochemical nitrogen cycle has changed markedly over time. With the discovery of the Haber-Bosch process at the beginning of the 20th century it became possible to industrially fix atmospheric

nitrogen converting N_2 into reactive N-forms and shifting the level of available N-forms (Chen et al., 2019). Industrially fixed nitrogen can be utilized to produce N-rich fertilizers that can be applied in agriculture to feed the ever-growing world population. However, this process is powered by fossil fuels which function at temperatures ranging from 400-500°C as well as pressure in the range of 150-300 bar. Therefore, approximately 3-5% of the natural gas produced globally is utilized for this process and nearly 1.2% of the world's energy is used for fertilizer production (IFA, 2009; Wang et al., 2018). It was estimated that in 2002, over half of the world's population consumed food produced with N fixed via the Haber-Bosch process (Smil, 2002). Despite their agricultural importance, the application of these fertilizers can cause huge environmental concerns and major changes to the balance of the biogeochemical nitrogen cycle (Richardson et al., 2009).

Annually, large quantities of reactive nitrogen from fertilisers are lost to the environment due to runoff, or as gaseous products. This can cause soil acidification as well as increased emissions of the greenhouse gas nitrous oxide (N_2O). N_2O has a global warming potential almost 300 times higher than CO_2 (Galloway and Cowling, 2002). Since the beginning of industrialisation, the atmospheric loading of N_2O has increased by over 20% and subsequently it has been listed as one of six gases subjected to restriction in the Kyoto protocol (Richardson et al., 2009; Smith et al., 2012; Prinn et al., 2018). The Intergovernmental Panel on Climate Change (IPCC) estimates that one third of the total global N_2O emissions are a result of anthropogenic activities, with agriculture accounting for the largest fraction (IPCC 2013). The economic costs as a result of nitrogen pollution across Europe are estimated to range from 70 to 320 billion euros a year, mainly due to reduced air and water quality (Sutton et al., 2011).

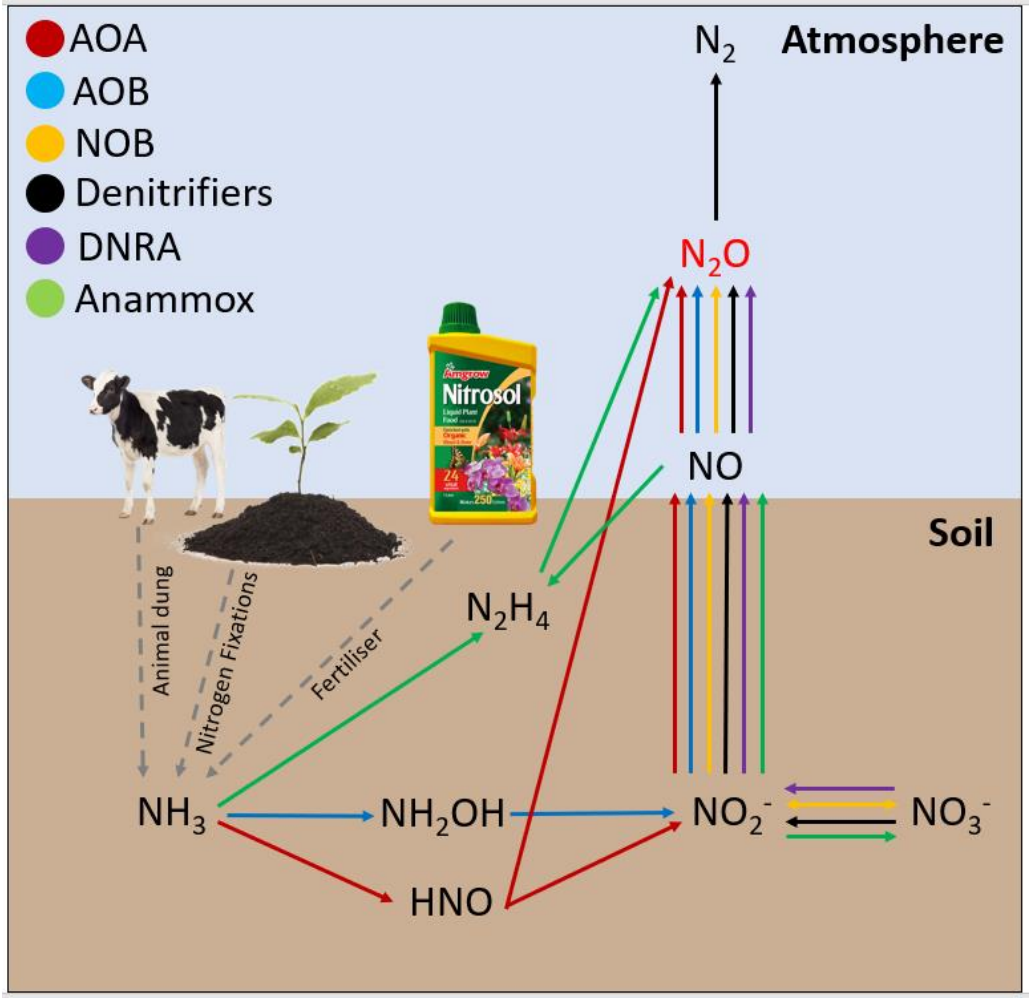


Fig. 1.1: Simplified representation of the major pathways of the nitrogen cycle in the soil (modified from Horta, 2016). Nitrogen fixation as well as the input of animal dung and nitrogen rich fertilisers result in an abundance of NH_3 in the soil, which is microbially converted to atmospheric N_2 via NO_2^-, NO and N_2O. AOA = ammonia oxidizing archaea; AOB = ammonia oxidizing bacteria; NOB = nitrate oxidizing bacteria.

Once nitrogen has been transformed into accessible ammonium compounds, the nitrification pathway is responsible for the conversion of ammonium to nitrite. Nitrification consists of two distinct steps, that are carried out by distinct types of microorganisms. During the first step a group of microbes known as ammonia-oxidizers aerobically convert ammonia to nitrite via the intermediate hydroxylamine. The process requires two enzymes, ammonia monooxygenase and hydroxylamine oxidoreductase (Fig. 1.2). Unlike nitrogen fixation which is carried out by a large range of microbes, ammonia oxidation is less common among prokaryotes. The bacterial ammonia oxidisers belong to the genera *Nitrosomonas*, *Nitrospira* and *Nitrosococcus*. However, in many habitats these are outnumbered by ammonia-oxidizing

Archaea, which are widespread in oceans, soils and salt marshes. The second step in nitrification is the oxidation of nitrite to nitrate carried out by nitrite-oxidizing bacteria belonging to the genera *Nitrospira*, *Nitrobacter*, *Nitrococcus* and *Nitrospina*. Both processes generate very small amounts of energy and thus the growth yields of the microbes are very low.

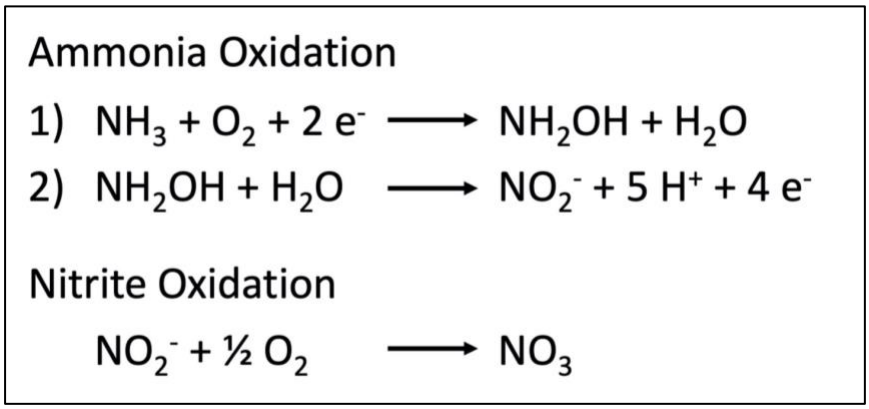


Fig. 1.2: Chemical reactions of Nitrification: During reaction 1 of ammonia oxidation the enzyme ammonia monooxygenase catalyses the conversion of ammonia to hydroxylamine. Reaction 2 converts the intermediate hydroxylamine to nitrite, a process catalysed by hydroxylamine oxidoreductase. The third reaction is the oxidation of nitrite to nitrate.

Nitrates and nitrites are a natural component of plant material, however an increase in nitrite from fertilisers can lead to an accumulation of nitrate in vegetable tissue (Renseigne et al., 2007). High levels of nitrate in food are responsible for methemoglobinemia (blue baby syndrome) in young children (Chan, 2011). Additionally, increased conversion of ammonium can lead to a loss of soil nitrogen through leaching which results in a wastage of fertiliser and water pollution through eutrophication of rivers and lakes (Robertson and Vitousek, 2009; Sutton et al., 2011). In freshwater ecosystems, the levels of nitrite are continuously increasing due to industrial effluents from industries producing metals, dyes, sewage aquaculture and runoff from agricultural soils supplemented with nitrogen fertilisers (Van Maanen et al., 1996; Jensen 2003). As nitrite is rapidly oxidised to nitrate (NO_3^-), NO_3^- is often the predominant N-form found in ground- and surface waters. Elevated levels of nitrite in both sea and fresh-water environments have detrimental effects on aquatic animals by interfering with multiple physiological functions such as ion regulation, respiration, and the cardiovascular system (Jensen et al., 2003). Biological removal of both nitrate and nitrite from aquatic environments

can be achieved aerobically through the processes of nitrate or nitrite assimilation or anaerobically via denitrification. Microorganisms and plants are responsible for the transformation of more than 10^4 megatons of inorganic nitrogen per year via the process of assimilatory nitrate reduction (Guerrero et al., 1981).

Recently, a new type of anoxic ammonia oxidation was discovered in addition to the nitrification process (Strous et al 1999). Anammox (anaerobic ammonia oxidation) is carried out by bacteria belonging to the phylum Planctomycetes which oxidise ammonia by using nitrite as the electron acceptor to produce gaseous nitrogen. They were first discovered in anoxic bioreactors of wastewater treatment plants but have since been found in a range of aquatic systems including low-oxygen zones of the ocean, coastal and estuarine sediments, mangroves, and freshwater lakes. In some parts of the ocean, anammox is thought to be responsible for a significant loss of nitrogen alongside the denitrification process (Kuypers et al., 2005).

The production of N_2O in the soil is primarily attributed to the microbial processes of nitrification and denitrification, although under certain environmental conditions such as nitrate-sufficiency and nitrite accumulation, dissimilatory nitrate and nitrite reduction to ammonium may well significantly contribute to N_2O emissions (Rowley et al., 2012; Stremińska et al., 2012). However, the denitrification process is the only known biological process physiologically capable of the consumption of this greenhouse gas (Bernhard, 2010), disregarding the non-physiological reduction of N_2O by nitrogenase (Hoch et al., 1960). Denitrification is an example of the respiratory flexibility found in prokaryotes with denitrifying microorganisms using nitrogen oxides as alternative electron acceptors in the absence of oxygen. Complete denitrification is a sequential four-step reduction of soluble nitrogen oxides nitrate (NO_3^-) and nitrite (NO_2^-) to the gaseous nitrogen oxides nitric oxide (NO), nitrous oxide (N_2O) and finally to dinitrogen, which takes place mainly in the absence of oxygen (Zumft & Kroneck, 2007). The multi-domain enzymes catalysing these reactions are nitrate reductase, nitrite reductase, nitric oxide reductase and the periplasmic nitrous oxide reductase respectively. As the denitrification process is a modular pathway, some organisms are capable of completely reducing NO_3^- to nitrogen gas whilst others may lack one or several of the enzymes required for the other steps involved in the reduction cascade (Philippot et

al., 2002). Complete denitrification is carried out by bacteria of several genera including *Paracoccus*, while a wide variety of species such as *Ruegeria* are able to perform parts of the process (Rana & Gupta, 2023.

1.2 Nitrogen in the terrestrial environment

The global biogeochemical nitrogen cycle is dominated by microbial processes in soils, sediments and water bodies (Seitzinger et al., 2016). This is a major difference from the global cycling of phosphorus which becomes available to the biosphere mostly through mineral weathering. Depending on the type of the ecosystem, N cycling and storage in soils and vegetation varies considerably. For instance, in agricultural systems, N cycling is dominated by N fertilisation and the removal of crops while in natural and semi-natural systems the cycling of nitrogen depends on climatic and landscape conditions as well as the sum of N inputs via N depositions and biological nitrogen fixation. Nitrogen cycling in terrestrial ecosystems is characterised by a variety of N transformations involving organic and inorganic (ammonium and nitrate) N species converted by microbes and plants. Waste production and decomposition are the major drivers of nitrogen release and nitrogen storage from the pool of organic matter. During the process of decomposition, soil organic matter is cleaved from larger polymers to bio-available monomers that are easily accessible to both plants and microbes.

European soils as well as shrublands, wetlands and forests are often N-limited due to poor soil quality, e.g. high sand content and thus reduced ion exchange capacities and low amounts of organic matter (Butterbach-Bahl & Gundersen, 2011). This has a negative impact on the retention of reactive nitrogen in the system. Anthropogenic influences over centuries have depleted the nutrient reservoirs. Tropical rainforests, however, are often rich in nitrogen. Changes in land-use have a significant effect on N-cycling and therefore a conversion of natural land into arable land is not only characterised by a loss of C stocks, but also by

significant losses of ecosystem N stocks. It is estimated that the conversion of forest soils to agricultural use resulted in an average nitrogen loss of 15% (Murty et al., 2002).

1.3 Nitrogen in the marine environment

Nitrogen occupies a central role in marine biogeochemistry and strongly influences the cycles of the other elements including carbon and phosphorus. It exists in a plethora of chemical forms that are converted by marine organisms as part of their metabolism to obtain nitrogen for the synthesis of structural components or to gain energy for growth (Pajares & Ramos, 2019). The most abundant form of nitrogen in the aquatic environment is dissolved nitrogen gas as the ocean absorbs nitrogen gas from the atmosphere. In the open ocean fixed nitrogen is one of the most important growth-limiting factors for photosynthetic organisms such as algae and marine bacteria (Falkowski, 1997). Denitrification and nitrification in the ocean are both regulated by oxygen concentrations and contribute to the production of nitrous oxide. Anthropogenic nitrogen inputs reach the coastal ocean via river inputs therefore also affecting the balance of the oceanic nitrogen cycle.

1.4 *Paracoccus denitrificans* as a model for denitrification

Paracoccus denitrificans (*P. denitrificans*) is a soil-dwelling member of the pseudomonadota found in a wide range of terrestrial and aquatic environments. Its biochemical apparatus allows the bacterium to utilize a range of electron donors to switch between aerobic and anaerobic respiration. It has become an important model as it is biochemically and genetically tractable and it grows well under denitrifying conditions in the laboratory. The genome of *P. denitrificans* was sequenced in 2006.

1.4.1 The *nar*/*nap*/*nas* gene operons

Excess nitrate is a pollutant, which is removed from the environment via the denitrification process. Nitrate is reduced to nitrite in a two-electron transfer process catalysed by nitrate reductase. *P. denitrificans* possesses three different nitrite reductases, a membrane bound nitrate reductase (Nar), a periplasmic nitrate reductase (Nap) and a cytoplasmic assimilatory nitrate reductase (Nas) (Bertero et al., 2003) (Fig. 1.3). The membrane bound enzyme reduces nitrate to generate metabolic energy whereas the periplasmic enzyme reduces nitrate to counter an over-reduction of the ubiquinone pool and subsequently balance excess redox potential. The Nas enzyme reduces nitrate as a source used for the production of bacterial biomass.

All three nitrate reductases consist of several subunits. Nar comprises of three subunits, a catalytic α subunit encoded by *narG* (Pden_4233), a soluble β subunit encoded by *narH* (Pden_4235) and a membrane bound subunit encoded by *narI* (Pden_4236). Another subunit, NarJ, is required for the assembly of the αβ complex. The *nar* gene operon also contains the genes *narK*, a nitrate-nitrite transporter, as well as *narR*, a nitrate transcriptional regulator which senses the levels of nitrate and oxygen. The location of NarG in the cytoplasm requires nitrate import and subsequent nitrite export. The presence of oxygen inhibits nitrate transport (Wood et al., 2001).

The *nap* operon encoding for the periplasmic enzyme consists of *napEDABC* and is not primarily involved in anaerobic respiration. The catalytic subunit NapA contains a molybdenum (MGD) cofactor as well as a [4Fe-4S] cluster. It receives electrons from the NapB subunit which is membrane bound and contains a tetraheme cytochrome c. The third subunit, NapC is involved in the transfer of electrons from the ubiquinone pool to the NapAB enzyme. The expression of the *nap* operon is unaffected by the availability of ammonium and oxygen in the environment. Instead, Nap activity is induced by nitrate. It has been shown that Nap is required for redox balancing during the oxidative metabolism of certain carbon sources, such as the highly reduced butyrate (Sears et al., 1992).

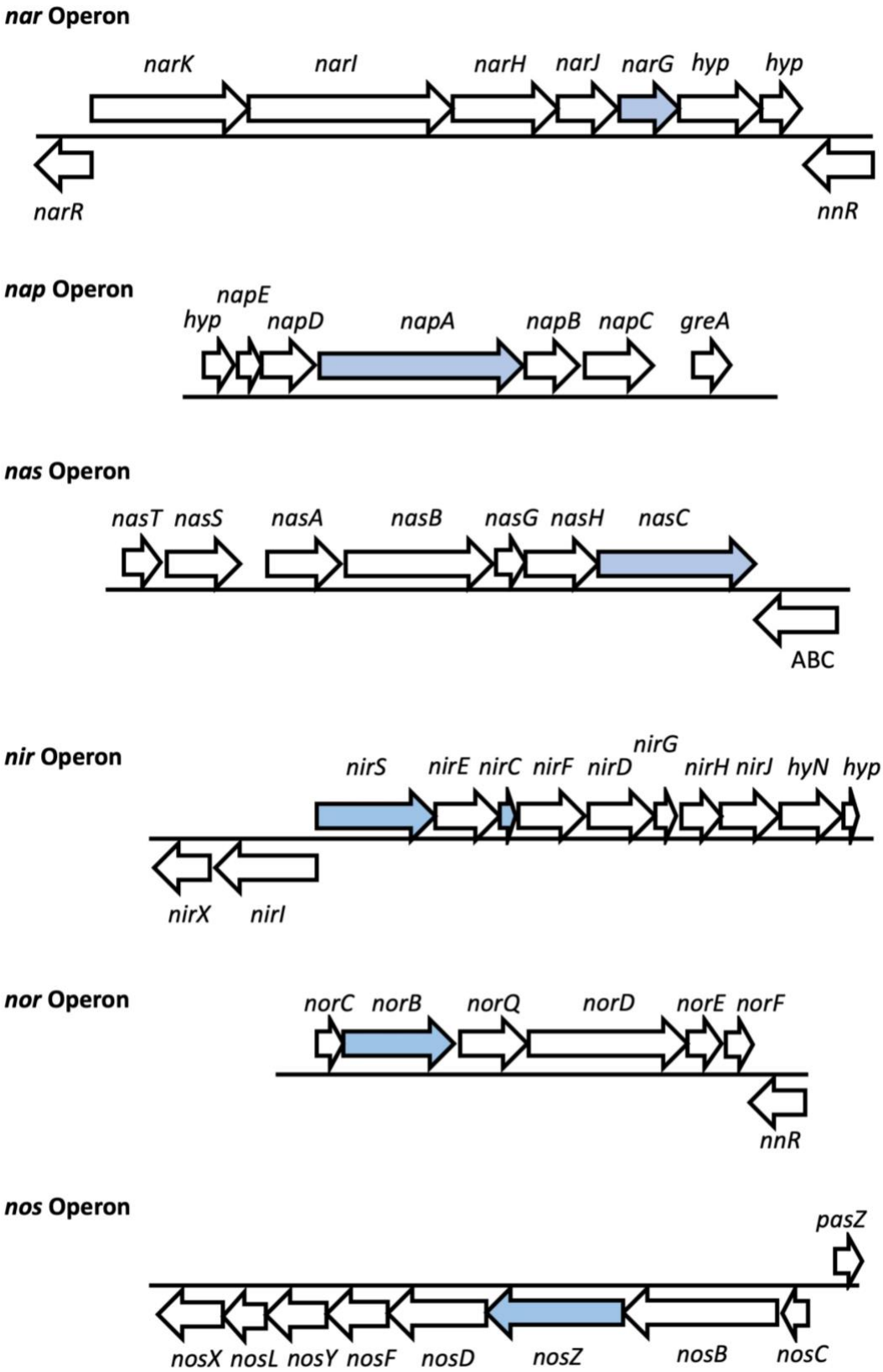


Fig. 1.3: The denitrification operons of *P. denitrificans*; blue arrows indicate the catalytic subunits

If no ammonia is available, the *nasTSABGHC* operon is expressed in *P. denitrificans* to utilise the available nitrate for growth. The catalytic subunit of the assimilatory nitrate reductase, NasC, reduces nitrate to nitrite in a two-electron reduction process. The electrons are transported to the catalytic subunit by the FAD-containing NasC subunit via NasG. NasH and NasA are required for the transport of nitrate and nitrite. The *nas* gene cluster is regulated by the two-component system NasTS (Bertero et al., 2003).

1.4.2 The *nir* gene operon

The next step in denitrification is the conversion of NO_2^- to NO, which is considered the defining step of denitrification as the soluble nitrogen oxide NO_2^- is converted into the gaseous NO (Jones et al., 2013). This is achieved via a one electron transfer and the enzyme responsible for this one-electron reduction reaction is either a copper- or a cytochrome cd_1 nitrite reductase, encoded by *nirK* or *nirS* respectively. Both of these are localised in the periplasm but are varied in structure and catalytic site. The nitrite reductase, NirS, of *P. denitrificans* is a cytochrome cd_1 enzyme. The nitrite reductase consists of two *c*-type hemes as well as two d_1-type hemes that form a functional dimer.

1.4.3 The *nor* gene operon

The reduction of NO to N_2O is carried out by a nitric oxide reductase (Nor). Two moles of nitric oxide are reduced to one mole of nitrous oxide in a two-electron reduction. NO is highly cytotoxic and is often involved in cell signalling as well as host defence. Therefore, many other non-denitrifying microorganisms contain Nor for NO detoxification (Philippot et al., 2013). The NorB subunit is a transmembrane protein that catalyses NO reduction and the NorC subunit receives the necessary electrons from cytochrome c_{550} or from pseudoazurin (Pden_4222).

1.4.4 The *nos* gene operon

The final step in the denitrification pathway, the reduction of nitrous oxide to atmospheric nitrogen gas, is carried out by the enzyme nitrous oxide reductase. The reduction of N_2O to N_2 requires two electrons and two protons and is strongly exergonic. The catalytic subunit of the nitrous oxide reductase enzyme is encoded by the *nosZ* gene. The genes *nosR* and *nosC* are located downstream of the *nosZ* gene and are both involved in *nosZ* expression as well as in Cu transport and assembly of NosZ (Fig. 3). The gene *pasZ* also lies downstream of *nosZ*. It encodes for pseudoazurin, which together with cytochrome c_{550} is responsible for donating electrons to Nos via cytochrome bc_1. The gene product of *nosX* is involved in metal ion transport and has been shown to be a homolog of *nirI* (Saunders et al., 2000). The remaining genes in the *nos* gene cluster, *nosDYFL*, all have an involvement in NosZ maturation. Nitrous oxide reductase forms a functional dimer and each of the monomers is able to bind 6 copper ions. At the N-terminal domain the enzyme possesses a unique catalytic Cu_z center while at the C-terminal domain a Cu_A electron transfer center is located. These copper centers impose a high copper demand on the enzyme and therefore microorganisms may avoid reducing N_2O if the bioenergetic advantage is limited. This can be observed in electron acceptor rich environments, such as nitrate-fertilized fields, resulting in an increased release of N_2O into the atmosphere.

The nitrous oxide reductase (NosZ) protein phylogeny has two distinct groups – clade I, and the recently discovered clade II (Jones et al., 2013; Hallin et al., 2018)). The two clades differ not only in protein phylogeny but also in the *nos* gene cluster organisation, the NosZ translocation pathway as well as the frequency of co-occurrence with other denitrification genes. Clade I organisms are complete denitrifiers which also possess *nirS* or *nirK* genes encoding for nitrite reductase (Jones et al., 2013; Conthe et al., 2018). The majority of the clade II organisms lack complete denitrification machinery and appear to be non-denitrifying N_2O reducers capable of consuming N_2O without contributing to its production, making these organisms of significant interest as they may be potential N_2O sinks in the environment. Despite the pressing need to develop mitigation strategies to combat the ever-increasing N_2O emissions, we still do not fully understand the regulatory network underlying the microbial reaction pathways responsible for the production and consumption of this greenhouse gas.

An enhanced understanding of the ecology of the *nosZ* clade II organisms as well as the conditions under which their activity is favoured is needed (Domeignoz-Horta et al., 2016).

1.5 Gene regulation of denitrification in proteobacteria

Bacteria and Archaea have developed a range of strategies allowing the uptake and utilisation of various nitrogen sources from their environment. These processes are tightly regulated in response to environmental conditions to ensure the correct temporo-spatial control of the pathways and minimise any inappropriate energetic costs as well as maximise the competitive growth advantage. The expression of the denitrification enzymes Nar, Nir, Nor and Nos in *P. denitrificans* is regulated by environmental signals including availability of oxygen, nitrate, nitrite, nitric oxide and copper (Fig. 1.4) (Gaimster et al., 2018). When oxygen levels become limiting, denitrifiers are forced to activate the expression of the denitrification enzymes to avoid entrapment in anoxic conditions without energy. Recent evidence has shown that *P. denitrificans* displays a bet-hedging strategy, a phenomenon that has been observed across a variety of prokaryotes which accept energetic penalties for a fraction of the population to achieve a long-term fitness advantage (Lycus et al., 2018). In the model denitrifier *P. denitrificans* this strategy involves the production of Nos in all cells, while Nir is only synthesised in a small number of cells.

Transcriptional regulation in the model denitrifier *P. denitrificans* is controlled by transcriptional regulators that bind to DNA binding sites (promoters) and subsequently up- or down-regulate the transcription of the gene. Transcriptional regulators can either bind to specific, highly conserved DNA regions, or they can bind non-specifically to control single genes or entire gene networks. In *P. denitrificans*, denitrification is partly controlled by the regulators FnrP (fumarate and nitrate reduction protein), NNR (nitrite reductase and nitric oxide reductase regulator) and NarR (nitrate reductase regulator) (Fig. 1.4).

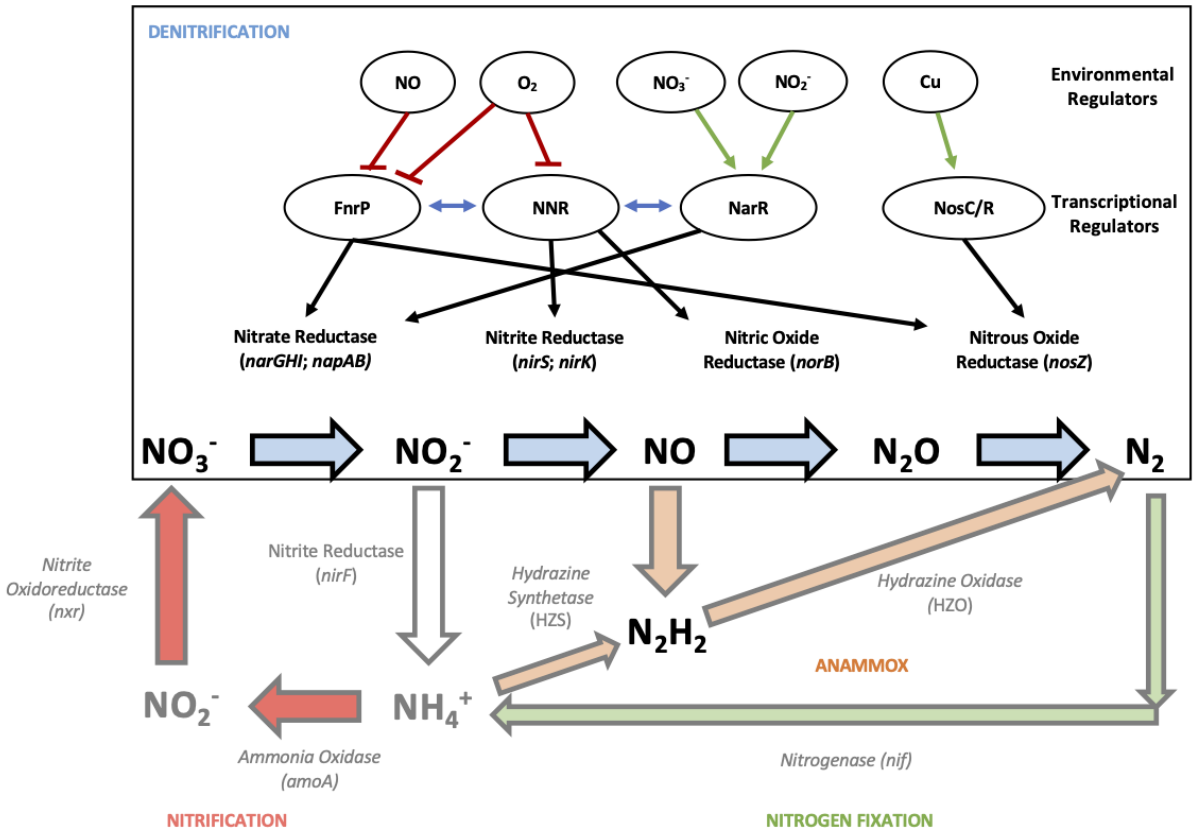


Fig. 1.4: An overview of the known transcriptional and environmental regulators of the denitrification pathway in the model denitrifying bacterium *Paracoccus denitrificans.* The black and red arrows between the upper layer of the environmental regulatory signals and the layer of the regulatory proteins indicate signalling events (red indicates an inhibitory effect whilst green indicates activation), while the arrows between the regulatory proteins and the denitrification enzymes indicate regulation of gene expression. The blue arrows between the transcriptional regulators indicate the cross-talk between the regulators which compete with each other to bind upstream of their targets.

1.5.1 Transcriptional regulators FnrP, NNR and NarR

Both FnrP and NNR are sensitive to oxygen and NO and are therefore involved in the switch between aerobic and anaerobic respiration to achieve maximum energy yields for the given environmental conditions (Van Spanning et al., 1995; Gaimster et al. 2018). To further fine-tune the denitrification network, the three transcriptional regulators FnrP, NnrR and NarR may serve as repressors of each other by competing for the binding upstream of their targets (Fig. 1.4) (Giannopoulos et al., 2017).

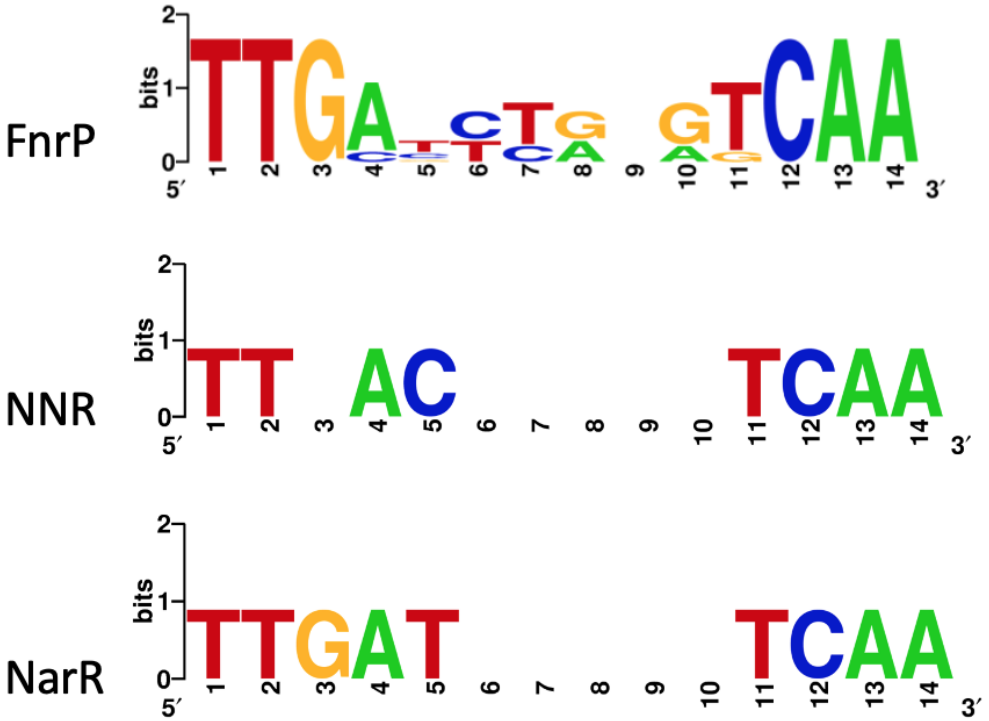


Fig. 1.5: FnrP, NNR and NarR recognition sequences generated using MEME suite showing most likely bases occurring at each position of the consensus sequence (Timothy et al., 2015)

FnrP acts as an activator of the *nar* and *nos* operons and recognises FNR-binding sequences (TTGAGAATTGTCAA and TTGACCTAAGTCAA) in the promoter region of the genes (Fig. 1.5) (Bouchal et al., 2010). Interaction of the FnrP 4Fe-4S cluster with O_2 leads to a separation of the transcriptionally active FnrP dimer into monomers (Crack et al. 2016). Hence, denitrification is switched off in the presence of oxygen as O_2 respiration provides significantly higher ATP yields. Additionally, it has been shown that the FnrP 4Fe-4S cluster interacts with multiple NO molecules leading to a dissociation of the transcriptional regulator into monomers. FnrP is encoded by a single gene, pden_1850, which is located between the high-affinity cbb_3 oxidase operon *cco* and oxygen-independent coproporphyrinogen III oxidase (*hemN*) which is required for heme biosynthesis.

The FNR-family transcriptional regulator NNR is homologous to FnrP and activates expression of the genes encoding nitrite (Nir) and nitric oxide reductases (Nor) (Van Spanning et al., 1995). An NNR mutant shows a two-fold reduction in growth under anaerobic, denitrifying conditions compared to wild-type *P. denitrificans*. The mutant accumulates nitrite due to complete Nir deficiency. NNR is rapidly inactivated when oxygen is introduced, suggesting a direct inactivation of the protein by O_2. Homologs of these transcriptional regulators have been identified in various other bacterial denitrifiers such as *Pseudomonas* species, *P. stutzeri* and *Rhodobacter* (Tosques et al., 1996; Elsen et al., 2004; Schreiber et al., 2007; Torres et al., 2017).

The third known regulator is the nitrite/nitrate-sensitive NarR protein. It is required for maximal expression of Nar and requires the presence of nitrate or nitrite to activate gene expression.

1.5.2 FnrP and NNR homologues in other denitrifiers

In *Pseudomonas* species the regulation of the denitrification machinery is equally dominated by members of the FNR superfamily such as the FnrP equivalent ANR (anaerobic regulator of arginine deaminase and nitrite reductase) as well as DNR (dissimilative nitrate respiration regulator) (Schreiber et al., 2007). This transcription factor ANR activates transcription of genes encoding for a nitrite transporter and a nitrite reductase if oxygen is limited while DNR senses NO concentrations. *P. stutzeri* encodes four FNR-type proteins which lack the cysteine residues required for the formation of a 4Fe-4S cluster (Vollack et al., 1999; Gaimster et al., 2018). ANR induces the expression of the dissimilative nitrate respiration regulator (DNR) which is required for the N-oxide dependent transcriptional activation of genes involved in the denitrification pathway (Giardina et al., 2008). NNR homologs have also been identified in a wide range of denitrifiers including *Rhodobacter sphaeroides* (Tosques et al., 1996). In *R. sphaeroides*, the regulator NnrR activates nitrite reductase and nitric oxide reductase. Outside *P. denitrificans,* other transcriptional regulators have also been shown to be involved

in the regulation of denitrification. These include the RegB/RegA two- component system. First discovered in *Rhodobacter capsulatus*, this system has been shown to regulate a large number of biological processes (Elsen et al., 2004). By controlling the expression of nitrite reductase, the RegB/RegA system in *R spaeroides* acts in concert with the regulator NnrR and therefore plays an important role in the denitrification cascade. In *B. diazoefficiens*, the denitrification machinery is regulated by two interconnected regulatory cascades, FixLJ-$FixK_2$-NnrR and RegSR-NifA that detect low levels of oxygen outside of the cell (Torres et al., 2017).

1.5.3 Environmental factors

The transcriptional regulators, FNR, NNR and NarR underpin the ability of bacteria to sense and respond to oxygen and the denitrification intermediates. However, there are other critical external factors that must be detected and integrated into the regulatory network of the cell. Copper has long been recognised as an important factor in the regulation of NosZ activity (Sullivan et al., 2013). Around 20% of Europe's arable lands are biologically copper deficient and as NosZ requires the unique multi-copper-sulphide centres Cu_Z and Cu_A to bind and activate N_2O, it places a high Cu demand on the bacterium (Sinclair and Edwards, 2008; Pauleta and Moura, 2017). Other bacterial enzymes require Cu for activity, such as haem Cu-oxidases or superoxide dismutases but for all of these enzymes there are non-Cu alternatives that can perform the same function in the absence of Cu (Zumft, 2005). This is however not the case for NosZ. As a result, in Cu deficient conditions, the final reduction step cannot be carried out leading to truncated denitrification and emission of N_2O. Studies carried out in 2012 demonstrated that copper-limited environments indeed lead to a downregulation of *nosZ* expression and an increased net N_2O emission without a significant effect on the biomass of the culture (Sullivan et al., 2013; Felgate et al., 2012). A down-regulation of *nosZ* expression in copper limited medium additionally influences expression of genes controlled by vitamin B_{12} riboswitches as accumulation of N_2O inactivates vitamin B_{12} (Sullivan et al., 2013). This work also showed that the accessory proteins NosC and NosR play an important role in copper-dependent expression of the *nos*-operon. Copper levels can therefore be

manipulated in laboratory studies to create N_2O or N_2 genic conditions and induce global changes in gene expression, a useful tool to further understand the underlying regulatory and biochemical pathways (Felgate et al., 2012).

Other environmental factors such as zinc and pH have also been linked to transcriptional regulation of denitrification enzymes (Bergaust et al., 2010; Gaimster et al., 2018). Zinc depletion has been shown to upregulate the expression of nitric oxide reductase and nitrite reductase as well as *nosC*, which was upregulated nearly 10-fold (Neupane et al., 2017). Low soil pH increases the N_2O:N_2 ratio which has been linked to lowered levels of NosZ protein synthesis and assembly as transcription rates were unaffected by changes in pH (Bergaust et al., 2010). Denitrification in heterotrophs is highly dependent on carbon sources and, therefore, increasing levels of organic carbon in the soil enhance denitrification rates as well as N_2O emissions (Saggar et al., 2013). Both environmental factors and transcriptional regulators strongly influence when denitrification is switched on and, once switched on, affect the denitrification rate. Numerous studies have analysed their influence in both a laboratory environment as well as in an agricultural background. Nevertheless, many variables involved in the switch between N_2O emission and N_2O consumption remain unknown.

1.5.4 Role of sigma factors in the regulation of denitrification

Sigma (σ) factors are required to control transcription initiation of functionally linked genes in bacteria. The σ factor interacts with the RNA polymerase forming the RNA holoenzyme and ensuring recognition of specific sequences upstream of the target gene. Bacteria often contain multiple σ factors. The primary σ factor responsible for essential growth and housekeeping genes is σ^{70}.

Sigma factor σ^{54} (encoded by the *rpoN* gene) was originally implicated in the expression of nitrogen-regulated genes but has since been assigned many other physiological roles. In Ralstonia eutropha σ^{54} is required for anaerobic growth on nitrate, suggesting that this sigma

factor is essential in the regulation of denitrification genes in this microorganism. In *P. stutzeri,* σ^{54} affected nitrite reductase as well as nitric oxide reductase activity. Absence of the sigma factor however did not affect the transcription of their structural genes, suggesting a post-transcriptional role of σ^{54}. Interestingly, *P. denitrificans* encodes a σ^{54} homologue, Pden_4987, as well as a number of other sigma factors.

1.6 Beyond DNA binding proteins – the contribution of small RNAs to the regulation of denitrification

Expression of genes in prokaryotes can also be regulated post-transcriptionally via small non-coding RNAs (sRNAs). Prokaryotic sRNAs are short regulatory RNAs that are heterogenous in length, sequence composition and secondary structure (Wassarman, 2002). They can range from between 50 and 200 nucleotides in length. The majority of sRNAs regulate major biological processes such as stress responses by binding to target regions called seed regions in the mRNA. This can result in either the activation or the repression of gene expression at the posttranscriptional level (Fig. 1.6) (Dutta and Srivastava, 2018). In numerous examples, sRNA binding results in the blocking of the ribosome entry site. Subsequently, translation initiation is inhibited and mRNA cleavage by RNase E as well as Rho-dependent termination of transcription, is induced (Bossi and Figuera-Bossi, 2016; Storz et al., 2004). In *E. coli*, the interaction of RhyB sRNA with the translation initiation region of *sodB* results in inhibition of translation and degradation of the target mRNAs (Masse and Gottesman, 2002). The same interaction of sRNA and target mRNA can lead to promotion of transcription by making 30S ribosomal subunit binding sites available or by blocking mRNA cleavage sites. This has been observed in the post-transcriptional regulation of RpoS, a stationary phase sigma factor, by the sRNA DsrA in both *Salmonella* and *E.* coli (De Lay and Gottesman, 2011). Advances in sRNA research have increased the number of known modes of interaction between sRNAs and their targets. Target interactions generally occur via base-pairing dependent mechanisms that alter translation or mRNA stability (Dutta and Srivastava, 2018).

sRNAs can originate from within a gene of interest or be processed from the 5' or 3' untranslated regions of genes coding for various proteins (Bossi and Figuera-Bossi, 2016). Many are then further processed by RNase E to produce sRNA fragments. This can be observed in the processing of RoxS in *Bacillus subtilis* resulting in an expanded repertoire of target mRNAs (Durand et al., 2015). There are two classes of sRNAs – *cis*-encoded and *trans*-encoded sRNAs. *Cis*-encoded sRNAs are transcribed from a DNA strand complementary to the one from which the target mRNA is transcribed resulting in high levels of complementarity. *Trans*-encoded sRNAs however, are transcribed from regions unrelated to those of their target genes resulting in lesser complementarity (Gottesman, 2005; Bossi and Figuera-Bossi, 2016). Due to a lower level of complementarity, *trans* encoded sRNAs can form base pairing with multiple target mRNAs and result in a global regulation of a physiological response. Interactions of sRNAs and their targets rely on base-pairings between complementary sequences (Georg et al., 2019). Base-pairing with the target is initiated through fast, high affinity binding of a few exposed nucleotides in the stem loop of the sRNA. This initial interaction promotes pairing of additional nucleotides, which frequently results in rearrangement of the RNA secondary structure (Otaka et al., 2011; Dutta and Srivastava, 2018). Structure-driven pairing of sRNAs and their targets in which the sRNA recognises C-rich stretches within accessible loops of the mRNA has also been demonstrated (Storz et al., 2011). Often, *trans*-encoded sRNAs require the presence of an RNA chaperone to facilitate binding to their target mRNA as their sequences are unrelated. In enteric pathogenic bacteria, such as *E. coli* and *Salmonella,* sRNAs have already been particularly well studied and many are associated with pathogenicity (Bossi and Figuera-Bossi, 2016).

1.6.1 The role of RNA chaperones in sRNA-mRNA interactions

The RNA chaperone Hfq is an Sm-like (Lsm) protein in the shape of a homohexameric ring, which can bind both sRNA and mRNA. It was first identified in *E. coli*, in which it acts as a host factor for the replication of the bacteriophage Qβ (Franze de Fernandez et al., 1968). Binding of Hfq acts to protect free sRNA from degradation by the cellular degradosome and increases local mRNA and sRNA concentrations, but Hfq can also recruit the degradosome to induce accelerated decay of the sRNA-mRNA complex (Georg et al., 2019). Most trans-encoded sRNAs contain a 3'-stem loop, which enables anchoring of the sRNA to Hfq via interactions of poly (U) to the inner rim of the Hfq homohexamer. The molecular mechanism of Hfq action has been explained in detail solely for its role in positive regulation of *rpoS* mRNA by the sRNA DsrA in *E. coli* (McCullen et al., 2010). Multilateral interactions between Hfq and the mRNA are formed distorting the mRNA structure to a more compact form, which facilitates binding of the sRNA (McCullen et al., 2010; DeLay and Gottesman, 2011). The binding of the RNA chaperone to a sequence motif in *rpoS* mRNA results in correct positioning of Hfq and is therefore essential for the pairing of this sRNA to its target mRNA. It is not known whether similar mechanisms are used in the Hfq-facilitated regulation of other sRNAs.

In addition to Hfq, recent studies have revealed the existence of a second RNA chaperone, ProQ, that can be found additionally to Hfq in *Salmonella* Typhimurium and *E. coli.* ProQ has been shown to facilitate binding of sRNAs and their targets, the molecular mechanism for this is however unknown (Smirnov et al., 2016; Smirnov et al., 2017; Westermann et al., 2019). In *Salmonella,* a loss of this chaperone results in a loss of virulence, as ProQ controls the expression of genes involved in motility, chemotaxis as well as SPI-1 transcripts (Westermann et al., 2019). The FinO domain of ProQ as well as other chromosomally encoded proteins containing a FinO domain are equally grouped as an additional class of bacterial RNA chaperones (Oleiniczak and Storz, 2017).

Interactions between Lsm proteins and sRNAs have also been observed in archaea. Some archaea, including halophilic archaea, encode a single Lsm protein (Lsm1), others encode two Lsm proteins (Lsm1 and Lsm2) (Fischer et al., 2010). Lsm1 proteins form heptamers capable

of binding DNA. Lsm2 proteins have been shown to associate to hexameric or heptameric complexes in *Archaeoglobus fulgidus.* Crenarchaeota contain a third Lsm3 protein which forms 14-mer complexes. Interestingly, the archaeon *Methanocaldococcus jannaschii* lacks an archaeal Lsm gene and instead contains an Hfq-like protein (Sauter et al., 2003; Nielsen et al., 2007; Vogel and Luisi, 2011). Lsm crystal structures obtained from *Archaeoglobus fulgidus* and *Pyrococcus abyssi* show that they are able to bind U-rich RNA in a similar way to Hfq (Töro et al., 2001; Törö et al., 2002; Thore et al., 2003). However, despite *in vivo* confirmation of the interaction of FLAG-tagged archaeal Lsm protein and sRNAs, the physiological functions remain poorly understood (Fischer et al., 2010; Martens et al., 2015).

1.6.2 Mechanisms of gene repression by sRNA

Regulatory sRNAs can directly or indirectly affect the expression of single or multiple genes. Binding of an sRNA within the physical boundary of the ribosome binding site (RBS) of the target mRNA prevents entry of the 30S ribosomal subunit and therefore blocks translation initiation (Fig. 1.6A) (Udekwu et al., 2005; Morita et al., 2006; Bouvier et al., 2008). Many sRNAs repress their targets by masking the Shine Dalgarno (SD) sequence or the AUG start codon. This mechanism is utilised by the sRNA RhyB found in *E. coli*. RhyB downregulates Fe-storage and non-essential Fe-binding proteins when iron availability is limited (Masse and Gottesman, 2002; Dutta and Srivastava, 2018). Absence of iron increases RhyB expression, which interferes with the binding of the 30S subunit to the RBS of the target mRNAs. An interaction of RhyB with Hfq can also result in the repression of the enzyme methionine sulfoxide reductase by binding to two sites on *msrB* mRNA. Binding to the first site stops ribosome entry at the RBS whereas binding to the second site results in a recruitment of RNase E (Bos et al., 2013; Dutta and Srivastava, 2018). Other sRNAs, such as OxyS, however can bind as far downstream as the 5th codon, without any interaction with the SD or the start codon (Bouvier et al., 2008). When ribosome entry sites are blocked, it is possible for the 30S subunit to bind to 'Standby regions', which are located 100 nucleotides upstream of the translation initiation site (Darfeuille et al., 2007). This mechanism is followed by the cis-acting sRNA Isr-1 and does not require the presence of an Hfq chaperone (Darfeuille et al., 2007).

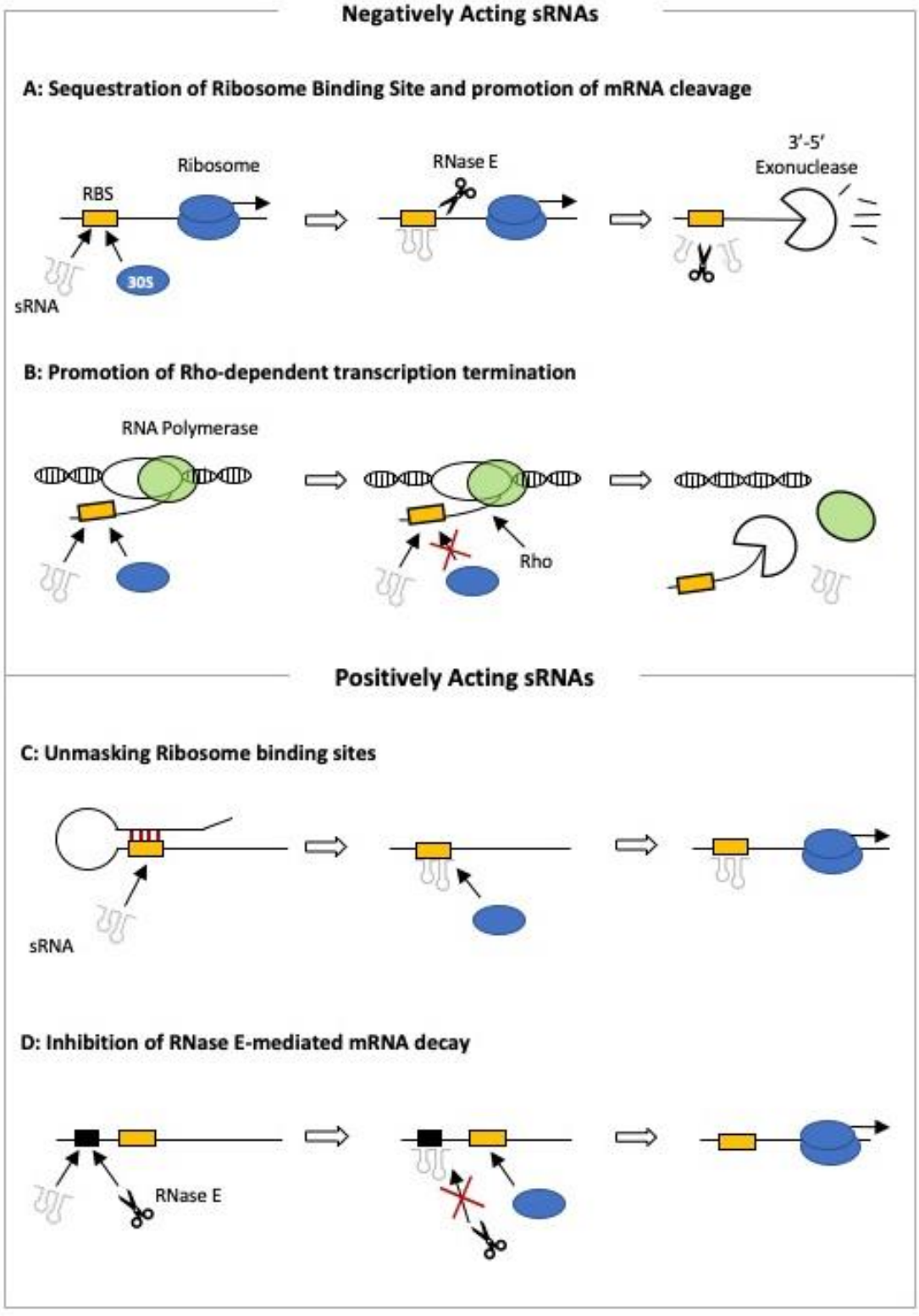


Fig. 1.6: Mechanisms of sRNA induced gene repression and activation. **(A)** When the sRNA target sequence overlaps with the ribosome binding site (RBS) translation initiation is blocked. This leaves the RNA more susceptible to RNase-mediated decay. **(B)** Alternatively, sRNAs can enhance Rho-binding and subsequently cause premature termination of transcription. **(C)**Positively acting sRNAs can bind to hairpin-like structures in their target, causing conformational changes to expose a previously inaccessible RBS and stimulate translation initiation. **(D)** Lastly, sRNAs are able to mask RNase E sites to stabilize their target and activate expression.

In addition to the blocking of ribosome entry sites, base pairing of an sRNA and its target at either the 5'UTR region or at downstream coding sequences can also lead to recruitment of ribonucleases such as RNase E. In prokaryotes, RNase E is a crucial ribonuclease responsible for the turnover of sRNAs and mRNAs (Chao et al., 2017). In some cases, Hfq can act as a protective layer against RNase E degradation by stabilising the sRNA and promoting base-pairing with the target. It has also been shown that Hfq has the capacity to directly bind to an unstructured C-terminal domain within RNase E forming a ribonucleoprotein with the sRNA that induces mRNA decay (Morita et al., 2005). The involvement of RNase E in sRNA induced gene repression has been confirmed for a large number of sRNAs such as RhyB and SgrS in *E. coli*.

Attenuation of transcription is a final mechanism of sRNA-induced gene repression. An example of this is the repression of the virulence gene *icsA* by the sRNA, RnaG, in *Shigella flexneri* (Giangrossi et al., 2010). The promoter of both the RnaG sRNA and the *icsA* virulence gene are convergent and lie less than 120bp apart. Hetero-duplex formation of the sRNA and its target gene results in a conformational change generating an intrinsic terminator that blocks the movement of RNA polymerase and thus attenuating *icsA* transcription.

1.6.3 Mechanisms of sRNA-induced gene activation

sRNAs are also able to mediate activation of genes involved in a wide array of physiological processes (Frohlich and Vogel, 2009; Dutta and Srivastava, 2018). One mechanism of gene activation is the stabilisation of target mRNAs by protecting them from degradation by cellular RNases (Fig. 1.6D). This has been observed for the glucose phosphate stress induced sRNA SgrS found in *E. coli* and *Salmonella* (Vanderpool and Gottesman, 2007). Binding of SgrS to its target mRNA *pldB-yigL* masks an RNase E site within the *pldB* open reading frame and facilitates production of the YigL phosphatase (Papenfort et al., 2013). Often, the secondary structure of mRNAs sequesters the ribosome binding site, which can be liberated for protein synthesis through pairing with an sRNA (Fig. 1.6C). This process is also referred to as 'anti-

antisense' mechanism. This activation of the 5' UTR was first discovered for the sRNA, RNAIII, in *Staphylococcus aureus* (Morfeldt et al., 1995). RNAIII is regulated by cell density through quorum sensing and activates the *hla* gene, which encodes an α-Toxin (Novick and Geisinger, 2008; Papenfort and Vanderpool, 2015). Activation is achieved through an interaction of the 5'-end of the sRNA and the SD-sequence of the target, preventing the formation of a translation-inhibitory structure formation. This 'anti-antisense' mechanism can also be observed in the activation of the sigma factor σ^S in *E. coli*. As the 5' UTR of the σ^S-mRNA (*rpoS*) is unusually long, it forms a complex hairpin structure, making it inaccessible for ribosomal entry (Battesti et al., 2011). Several sRNAs (DsrA, RprA, ArcZ) are able to bind to specific sections within the 5' UTR to rearrange the structure and enhance the rate of σ^S translation (Bossi and Figuera-Bossi, 2016).

In addition to the anti-antisense mechanism observed in *rpoS* activation, a unique transcription anti-termination system has been discovered to play a crucial role in inhibiting Rho-dependent transcription termination in the 5' UTR of *rpoS* (Sedlyarova et al., 2016). Rho is a hexameric helicase protein and together with its cofactor NusG it acts as a global transcription termination factor in prokaryotes (Boudvillain et al., 2013). Rho binds to a stretch of C-rich unstructured RNA that is around 80 nucleotides in length and is located near the transcription terminator and subsequently to mRNA binding, its ATPase activity is stimulated. Under specific circumstances, Rho appears to be active in the 5' UTR, which induces premature termination of transcription. Within *rpoS,* one of these Rho loading sites can be found in the leader sequence. Binding of an sRNA close to this Rho-loading site blocks binding of Rho and enhances transcription and protects from cleavage induced by RNase E (Fig. 1.6D) (McCullen et al., 2010). Hfq further increases the stability of the sRNA-*rpoS* interaction.

In some cases, sRNA can positively regulate expression of an open reading frame (ORF) through interactions with its 5' UTR that can result in a subsequent upregulation of a different cistron of the mRNA which is translationally coupled to the ORF (Dutta and Srivastava, 2018). In *Pseudomonas aruginosa,* the oxygen-responsive sRNA PhrS activates the *ufo-pqsR* operon in the absence of oxygen (Sonnleitner et al., 2011). The transcriptional regulator PqsR controls the expression of several virulence genes in *P. aeruginosa* including the toxic pigment

pyocyanin (PYO) and the quorum sensing and biofilm formation signal PQS. PhrS binds to the 5' region of *ufo* which results in conformational change liberating the RBS. As *ufo* is translationally coupled to *pqsR*, the presence of sRNA PhrS eventually results in enhanced translation of pqsR and increased levels of PYO and PQS (Sonnleitner et al., 2011).

1.6.4 sRNA induced protein sequestration

Certain sRNAs have the ability to directly sequester RNA-binding proteins inhibiting them from carrying out their functions or binding enzymatic proteins to inhibit or modify their enzymatic activity. Therefore, these sRNAs can indirectly regulate the expression of many genes related to this protein. The RNA-binding protein CsrA is a post-transcriptional regulator that has multiple targets, which include several genes involved in carbon flux pathways (Babitzke and Romeo, 2007). The presence of sRNAs such as CsrB results in an inactivation of CsrA activity as CsrB acts as a direct competitor for the CsrA target mRNAs in the cell removing its function and changing the expression level of a large number of genes. Inhibition of a protein's enzymatic activity can be observed for sRNA 6S which binds to RNA polymerase in bacteria interfering with σ^{70}-induced transcription (Wassarman and Storz, 2000). Production of 6S is maximised during the stationary phase and as a result the expression of several genes is reprogrammed to allow the cell to adapt to the given environmental conditions. The RNase BN/Z facilitates 6S RNA decay (Chen et al., 2016).

As sRNAs are significantly smaller than mRNAs and do not require translation into a protein, they have a potential energetic advantage over the production of protein transcription factors (Beisel and Storz, 2010). sRNA copy number within the cell can also be very high but their turnover time however is very short, resulting in a sharp deterioration of sRNA numbers once they have exerted their rapid and effective function in response to an environmental signal. This suggests that sRNA could be crucial in the rapid adaptation to dramatic shocks such as sudden nutrient change that challenge the survival of the microorganism. As more advances are made in sRNA research, more mechanisms are discovered that demonstrate the diverse

mechanisms of action of sRNAs and their association with a large variety of physiological processes.

1.6.5 Physiological Roles of sRNAs

Small RNAs are important regulators that take over diverse roles within the bacterial kingdom. Especially the synthesis of membrane proteins is a predominant target of sRNA regulation: The expression of several membrane proteins including outer membrane β-barrel proteins in *E. coli* are controlled by a number of sRNAs. Additionally, in both *E. coli* and *Salmonella*, numerous sRNAs are responsible for the repression of outer membrane porin production (Vogel and Papenfort, 2006). An example for this is sRNA MicF which is induced by high osmolarity and as a result represses the synthesis of the porin gene *ompF*. Another common target is metabolic remodelling in response to environmental shifts (Gottesman and Storz, 2011). The clearest example for this is RhyB in E. coli which is induced in response to low iron and represses the synthesis of non-essential iron-containing enzymes such as aconitase B and succinate dehydrogenase. As a result, the limited iron in the cell can be used by the critical enzymes (Masse et al., 2007).

It is becoming more and more clear that cell communication during quorum sensing and biofilm formation is regulated by sRNAs. To react to changes in cell density *V. cholerae* possesses two-component membrane-bound sensor kinases. At low cell density, the response regulator LuxO is phosphorylated and activates the expression of five sRNAs that regulate the expression of genes involved in virulence and biofilm formation (Bardill et al., 2011, Michaux et al., 2014). In pathogenesis, sRNAs often modulate expression levels of outer membrane proteins which are targets for the immune system, as well as other responses required for the survival within the host. Members of the CsrB family of sRNAs in *Salmonella, Yersinia, Vibrio* and other pathogenic bacteria have already proven to alter infection by antagonizing global regulators of virulence genes (Waters and Storz, 2009). Other sRNAs are involved in the adaptation to nutrient availability. Switches between nutrient availability and famine trigger major changes in gene expression and require a coordination of regulatory networks. In *E. coli*, the sRNA SgrR modulates the response to an accumulation of glucose 6-

phosphate which is toxic when present at high concentrations (Vanderpool and Gottesman, 2004). Besides biofilm formation and pathogenesis, many of the known sRNAs are involved in stress responses such as oxidative stress, osmotic stress and the switch between aerobic and anaerobic metabolism. This is the case for RprA, a regulatory sRNA found in *E. coli*, which modulates the cellular response to osmotic stress.

1.6.6 sRNAs in the Nitrogen Cycle

Although the focus of sRNA studies has predominantly been on key model bacteria, with a particular focus on stress responses and pathogenesis, it is becoming increasingly clear that sRNAs also play a crucial role in environmentally relevant pathways. With increasing research into the regulatory role of sRNAs it is predicted that large numbers of these influence nitrogen cycle associated metabolism across many microorganisms. Direct involvement of sRNAs in the response to N-fluctuations in the environment or in the regulation of N_2-fixation has not been identified until recently. Indirect participation of sRNAs in N-metabolism however has been reported previously. In cyanobacteria for instance, NsiR1 controls the formation of heterocysts as well as the switch to nitrogen fixation (Ionescu et al., 2010). This trans-encoded sRNA is conserved across many heterocyst-forming cyanobacteria and is dependent on the regulatory protein HetR which is required for cell differentiation in *Anabaena*. Similar to NsiR1, sRNAs NsiR2, NsiR8 and NsiR9 have been shown to be co-expressed with heterocyst-specific genes. However, to date no specific function has been assigned to these three sRNAs. Furthermore, ArrF of *Azotobacter vinelandii* is involved in the regulation of FeSII (DeLay and Gottesman, 2009), which plays a key role in the protection of the nitrogenase (N_2-fixing) enzyme under oxidative conditions (Jung and Kwon, 2008).

sRNAs indirectly involved in nitrogen assimilation include CyaR, GcvB and MmgR. CyaR, present in *E. coli*, inhibits the translation of an ammonium dependent NAD-synthethase responsible for the catalysis of NAD synthesis from either NH_3 or glutamine as well as the nicotinic acid adenine dinucleotide (De Lay and Gottesman, 2009). The sRNA GcvB is one of the most highly conserved Hfq associated sRNAs in Gram-negative bacteria. It inhibits the expression of a number of ABC transporters responsible for transporting amino acids in *E. coli*

and *Salmonella* Typhimurium (Sharma et al., 2007). In *Sinorhizobium meliloti*, hundreds of sRNAs have been identified, with the focus on the sRNA MmgR which shows expression patterns highly dependent on the available nitrogen source (Ceizel Borella et al., 2016). Further work is required to elucidate its exact role.

1.6.7 Small RNAs regulating nitrogen fixation

An sRNA found to be directly involved in N-metabolism, NfiS, was identified in the root associated bacterium *P. stutzeri* A1501. Via a stem loop in the sRNAs secondary structure it is predicted to bind to the 5' region of *nifK* mRNA which encodes a subunit of the nitrogenase enzyme. This interaction increases mRNA half-life and thus increases the translation efficiency of *nifK* optimising N-fixation (Zhan et al., 2016). The stability of NfiS appears to be strongly affected by the presence of Hfq, as the transcript is hardly detectable in an *hfq* deletion strain. A complete knockout of NfiS results in decreased nitrogenase activity, while an overexpression of this sRNA can lead to an increase of up to 150% activity. Although NfiS is highly conserved in *P. stutzeri*, it cannot be found in other bacterial species.

Many mechanistic features of the cellular transcription and translation machinery in archaea are more closely related to the eukaryotic counterparts, however characterisations of archaeal sRNAs have suggested similar mechanisms as observed in bacteria. The best-known examples of sRNA regulation of nitrogen fixation in Archaea are the methanoarchaea *M. mazei* and *M. maripaludis*. Both strains contain the global N-repressor NrpR which is known to transcriptionally regulate a variety of target genes in response to changes in N-levels. An RNA-seq study in *M. mazei* Gö1 under conditions of varied nitrogen availabilities lead to the identification of 242 putative sRNAs (Jäger et al., 2009; Jäger et al., 2012). The discovery of $sRNA_{41}$ in *M. mazei* Gö1 introduced a sRNA in Archaea with a regulatory impact on the metabolic cycles of both carbon and nitrogen (Buddeweg et al., 2017). The sRNA is induced 100-fold in a N-rich environment compared to N-limitation and interacts with the mRNA encoding for an acetyl-coenzyme a decarbonylase/synthase (ACDS) complex (Buddeweg et

al., 2017). In the absence of nitrogen, reduced amounts of $sRNA_{41}$ result in the upregulation of the ACDS complex and a subsequent production of amino acids for the synthesis of nitrogenase. A further sRNA found in *M* mazei, $sRNA_{154}$, was found to be exclusively present under nitrogen-limited conditions (Ehlers et al., 2011). A computational analysis of the transcriptional regulation network in *M. acetivorans* has shown that 5% of genes in this methanoarchaeon are regulated under nitrogen limitation. Two sRNAs, $sRNA_{154}$ and $sRNA_{159}$ were identified which include Nrp binding sites suggesting an involvement in gene regulation under N-limitation (Peterson et al., 2014). The first confirmed directly acting sRNA in *M. mazei*, $sRNA_{154,}$ is under direct control of the global N-repressor NrpR (Weidenbach et al., 2008; Weidenbach et al., 2010). By stabilising the polycistronic mRNA encoding for the nitrogenase enzyme as well as stabilising the transcription of the regulatory protein NrpA it enhances expression of the N-fixing machinery (Prasse et al., 2017). The sequence and structure of this sRNA is highly conserved across members of the *Methanosarcinales.*

Despite nitrification being an important part of the nitrogen cycle, few sRNAs have been shown to be involved in the regulation of this pathway due to a lack of studies around this topic. In the ammonia oxidising archaea *Nitrosopumilus maritimus* six candidates for sRNAs have been identified and it is highly likely that there are many more with a potential involvement in the nitrification process (Walker et al., 2010).

1.6.8 sRNAs controlling Nitrogen assimilation

A differential RNA-seq analysis of the cyanobacteria *Anabaena sp.* PCC7120 in response to N-availability identified over 600 transcriptional start sites indicating an abundance of *cis*- and *trans*-encoded sRNAs involved in the regulation of N-assimilation. Cyanobacteria are of importance in both aquatic and terrestrial ecosystems and are important links between the C- and the N-cycle. A cyanobacterial small RNA directly involved in the regulation of N-assimilation, NsiR4, was first reported by Klähn et al. in 2015. NsiR4 expression in cyanobacteria is stimulated during nitrogen limiting conditions via the transcriptional regulator NtcA which is known to regulate a variety of genes involved in nitrogen metabolism.

It is predicted to interact with the 5' UTR of *gifA* mRNA, encoding for the glutamine synthetase inactivating factor (IF)7. By affecting IF7 expression, the sRNA also alters the activity of glutamine synthetase, a key enzyme in biological nitrogen assimilation (Klähn et al., 2015).

In *Pseudomonas aeruginosa*, the putative sRNA NalA is encoded upstream of the nitrate assimilation operon *nirBD*–PA1779–*cobA*. The transcription of this sRNA is σ^{54} and NtrC-dependent (Romeo et al., 2012). A deletion mutant of NalA was unable to grow in presence of nitrate as the sole nitrogen source, instead it grew similarly to the parental strain in presence of ammonium. The results showed that NalA sRNA and nitrate are required for transcription of the nitrate assimilation operon, being an essential sRNA for the assimilation of nitrate (Romeo *et al.,* 2012). Further studies performed in *P. aeruginosa* allowed the identification of sRNAs related to detoxification of industrial cyanide-containing wastewaters. For this purpose, a differential expression study was carried out by RNA-seq from cells cultured with a cyanide-containing wastewater, sodium cyanide or ammonium chloride as the sole nitrogen source. Among the sRNAs identified, $sRNA_{14}$ (overexpressed in the presence of ammonium) stood out, as its putative target genes include the nitrilase NitC, essential for cyanide assimilation, the FAD-dependent oxidoreductase NitH; and the glutamine synthetase, related to ammonia assimilation. Moreover, $sRNA_{14}$ showed a high conservation among enterobacterial species (Olaya-Abril *et al.*, 2019).

In the archaeon *Haloferax* mediterranei, sRNAs have been studied to elucidate their possible role in the regulation of nitrogen assimilation in Haloarchaea (Payá *et al.*, 2018; Payá *et al.*, 2020). The initial identification of sRNAs in *H. mediterranei* was performed using a library of sRNAs identified in other archaeal species which resulted in the discovery of 295 putative sRNAs genes (hot spots) in the genome of *H. mediterranei*. By way of bioinformatic and RNomic approaches, 88 sRNAs were identified. The differential expression analysis of these 88 sRNAs showed 16 sRNAs with different expression patterns according to the nitrogen source. The expression of their predicted target genes also depends strongly on the nitrogen source. Three regulatory mechanisms mediated by sRNAs were proposed in this study (Fig. 5). The sRNA HM8_S which is overexpressed in presence of nitrate is predicted to target glutamate dehydrogenase, which is repressed in presence of nitrate. Therefore, this sRNA could negatively regulate the expression of glutamate dehydrogenase. Both HM7_S and

HM54_V sRNAs, overexpressed in presence of nitrate, are predicted to target transcriptional regulators belonging to the ArsR family, whose expression depends on the nitrogen source. Finally, the putative target of HM1_A (overexpressed in the presence of ammonium) is an ammonium transporter (expressed in the presence of nitrate or under nitrogen starvation) and therefore this sRNA could be involved in the regulation of ammonium uptake from the extracellular medium. However, more work is needed to confirm these regulatory mechanisms (Payá *et al.*, 2018).

The second step in the identification of sRNAs in *H. mediterranei* results in the identification of the complete sRNAome in presence of ammonium or nitrate as the sole nitrogen source. 460 sRNAs were present in both conditions, 102 of which showed differences in their transcriptional patterns. Specifically, sRNAs with potential target genes related to nitrogen metabolism, such as *nosL, glnK1, gdh, glnA2, nasB, ilvB3, ilvE2, ilvAm*, *rrfh1, tyrA, gst2, gabT, gaD2, argD, gltp, purL, argB, gatD, nadE, fdx, exsB, gcvP1* and *pyrF* also presented differences in their transcriptional expression patterns according to the nitrogen source. From these findings, three potential regulatory mechanisms of nitrogen metabolism pathways mediated by sRNAs were proposed (Fig. 1.4): (1) sRNA228 could be involved in the repression of nitrogen regulatory protein PII (*glnK1*) in the presence of ammonium, potentially through the posttranscriptional degradation of *glnK1* mRNA preventing its transcription and therefore the activation of the GS/GOGAT pathway; (2) the sRNA451 could be involved in the positive regulation of the nitrate/nitrite transporter (nasB) expression in presence of nitrate as nitrogen source, by transcriptional stabilisation of the *nasB* mRNA, increasing nitrate uptake under these conditions; and (3) sRNA238 could be involved in the transcriptional stabilization of the *HFX_RS05100* gene (both overexpressed in presence of nitrate). Although *HFX_RS05100* encodes a signal transduction protein of unknown function, the results of this work suggest that it may be involved in nitrogen metabolism (Payá *et al.*, 2020).

1.6.9 sRNAs controlling denitrification

The importance of sRNA regulation in denitrifiers is a relatively recent discovery. However, 167 putative sRNAs across the *P. denitrificans* genome have now been identified when cultured under denitrifying conditions (Gaimster et al., 2016). Over one third of these sRNAs were differentially expressed between N_2 and N_2O emitting cultures suggesting a role of these sRNAs in production or consumption of the greenhouse gas. Several of these sRNAs showed sequence homology and conservation across other species in the α-proteobacteria. Interestingly, one sRNA, intergenic 28, showed sequence homology to members of the β-proteobacteria, including members of the *Bordetella* genus which include strains of human host-restricted pathogens as well as free-living environmental strains isolated from both aquatic and soil environments. Commonly predicted targets of sRNAs were transcriptional regulators such as Xre, Fis and TetR- family regulators, which may act globally. This is consistent with other studies in which global regulators in other bacterial species have been shown to be subject to regulation by multiple Hfq-dependent sRNAs. *P. denitrificans* is predicted to encode an Hfq protein, Pden_4124, that has 95% sequence identity to Hfq found in *R. sphaeroides* and 54% sequence identity to *P. aeruginosa* Hfq. Many sRNAs found in both these bacteria are Hfq-dependent suggesting the same may be the case in *P. denitrificans* (Gamister et al., 2016). Additional predicted targets for sRNAs in *P. denitrificans* are transport proteins, which could be a conserved role for sRNAs across species as this was also the most commonly predicted sRNA target in the marine denitrifier *R. pomeroyi* (Rivers et al., 2016).

Mechanistic studies carried out by Gaimster et al. then continued to report a novel regulatory pathway controlling denitrification via a single sRNA, sRNA29 (DenR) (Gaimster et al., 2019). DenR is suggested to stabilise the expression of a previously unknown GntR-type transcriptional regulator, NirR, which in turn represses the denitrification rate through repressing NirS, resulting in reduced N_2O emissions. The predicted region of interaction is a 7bp-seed region located within the CDS of *nirR*, and the underlying mechanism is being resolved in this thesis. GntR-type regulators have been identified across many bacterial species in which they play crucial roles in the regulation of intracellular processes. They are named after the gluconate-operon repressor in *Bacillus subtilis* and they consist of a

conserved N-terminal helix-turn-helix DNA-binding domain, which is linked to a C-terminal signalling domain. The overexpression of DenR also results in altered expression levels of 53 other genes that are mostly genes of either unknown function, genes involved in energy metabolism or transport as well as genes involved in carbohydrate metabolism. Interestingly, DenR has been found to be conserved across several denitrifying bacterial species in the *Rhodobacteraceae* genus. This includes the closely related species *Paracoccus aminophilus* but also the more distantly related marine denitrifier *Ruegeria pomeroyi*. All these species encode a transcriptional regulator with homology to NirR, suggesting a similar, conserved mode of action. Although there are limited findings in other denitrifiers, for the opportunistic pathogen *P. aeruginosa*, the anaerobically induced sRNA Pail is known to be required for efficient denitrification by affecting the conversion of nitrite to nitric oxide (Tata et al., 2017).

1.7 Aims

Our understanding of the complex regulatory networks that control the process of denitrification is fundamental to our ability to develop future mitigation strategies for the potent greenhouse gas nitrous oxide. Small RNAs have been proven to play a pivotal role in a plethora of physiological processes and have recently been proven to be involved in switching on and off nitrous oxide production in denitrifying cultures of the model bacteria *Paracoccus denitrificans*. The intention of this thesis is to expand our understanding of the role of 167 previously identified sRNAs in *P. denitrificans* and integrate them into the existing picture of transcriptional regulation while also addressing further questions regarding the transcriptome during denitrification. This thesis specifically aims to:

- Characterize several carefully selected sRNAs in *P. denitrificans* computationally and experimentally to identify the roles of these sRNAs in the regulation of denitrification (Chapter 3)
- Confirm and further explore the suggested mode of action of the previously characterized sRNA DenR via a novel GntR-type regulator NirR (Chapter 4)
- Map transcription start sites across the genome to provide further insights into the use of promoters during denitrification as well as explore the use of alternative sigma factors (Chapter 5)

2 Materials and Methods

2.1 Materials

The chemicals and reagents used are laboratory standard grade or above. To make all media and solutions, dH_2O was used.

2.2 Bacterial Strains and Plasmids

Bacterial strains used in this study are *Paracoccus denitrificans* strain PD1222 as well as *Escherichia coli* K-12 derivatives. Deletion mutants were derived from PD1222 which for ease of description is referred to as 'wild-type' (WT) strain throughout this work. A comprehensive list of strains and plasmids used in this study are described in Table X and Table Y respectively.

2.3 Media and conditions for bacterial growth

Both a complete medium and a minimal medium were used in this thesis: the complete minimum for cell propagation and the defined minimal salt medium for physiological studies. Strains utilised in this study are listed in Table 2.1 and all plasmids utilised are listed in table 2.2. Overexpression plasmid constructs of pLMB509 were generated using the GenScript gene synthesis service. The gene sequences were cloned into the overexpression vector at the first NdeI site and later transformed into P. denitrificans or E. coli by use of the methods outlined in 2.4.3 and 2.4.4.

Table 2.1: Strains used during this study

Strain	Description	Reference/Source
***Paracoccus denitrificans* Strains**		
PD 1222	Wild-type strain, rif^R, $Spec^R$	Sullivan et al., 2013
***Escherichia coli* Strains**		
E. coli JM101	Used as host for small plasmids; supE thi-1 ∆(lac-proAB) (F traD36 proAB laclqZ∆M15)	Invitrogen™
E. coli S17.1		Invitrogen™
E. coli BL21 DE3	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS λ DE3 = λ sBamHIo ΔEcoRI-B int::(lacI::PlacUV5::T7 gene1) i21 Δnin5*	NEB™
E. coli TOP10	F- *mcrA* Δ(*mrr-hsd*RMS-*mcr*BC) Φ80*lac*Z∆M15 Δ *lac*X74 *rec*A1 *ara*D139 Δ(*araleu*)7697 *gal*U *gal*K *rps*L (StrR) *end*A1 *nup*G	Invitrogen™

Table 2.2: Plasmids used during this study

Plasmids	Description	Reference
pLMB509	Pc promoter (taurine inducible); Gm^R	Tett et al., 2012
pET14b	T7 promoter, Amp^R; N-term His + thrombin cleavage site	Novagen™
Overexpression constructs		
pLMB509_DenR	*denR* gene in pLMB509	This study
pLMB509_sRNA5	sRNA 5 gene in pLMB509	This study
pLMB509_sRNA10	sRNA 10 gene in pLMB509	This study
pLMB509_sRNA18	sRNA 18 gene in pLMB509	This study
pLMB509_sRNA36	sRNA 36 gene in pLMB509	This study
pLMB509_sRNA39	sRNA 39 gene in pLMB509	This study
pLMB509_sRNA79	sRNA 79 gene in pLMB509	This study
pLMB509_hfq	hfq gene in pLMB509	This study
His-tagged Protein constructs		
pET14b_GntR	*gntR* gene in pET14b with 7x His	This study

2.3.1 Complete medium

A complete lysogeny broth (LB) medium as described by Luria and Bertani was used to culture *Escherichia coli* and *Paracoccus denitrificans* PD1222. Antibiotics were added as outlined in Table 1. Solid media contained 1.5% (w/v) of agar.

Table 2.3: Concentrations of antibiotics

Antibiotic	**[Stock] (mg mL^{-1})**	**[Final] (μg mL^{-1})**	**Storage**
Gentamicin (gen)	50	20	RT
Rifampicin (rif)	50	50	-20°C
Kanamycin (kan)	50	50	-20°C
Ampicillin (amp)	50	50	-20°C

2.3.2 Minimal Medium

A defined minimal salts media was used for the examination of growth of *P. denitrificans* strains. The medium was prepared at pH 7.5 and contained 29 mmol/liter Na_2HPO_4, 11mmol/liter KH_2PO_4, 10mmol/liter NH_4Cl, 0.4 mmol/liter $MgSO_4$, as well as 20 mmol/liter $NaNO_3$. Succinate and ammonium were used as sole carbon and nitrogen sources, while nitrate served as respiratory electron acceptor when grown anaerobically. The medium was supplemented with 2ml/liter essential trace metal solution (Table 2). To achieve copper (Cu)-low medium, no $CuSO_4$ was added. The pH of the trace metal solution was adjusted to 6.2 using KOH.

Table 2.4: Constituents of Vishniac and Santer trace elements solution for anaerobic growth of *Paracoccus denitrificans*

Compound	M_w	mM	g L^{-1}
EDTA	292.24	342.2	50.00
$ZnSO_4$-$7H_2O$	287.55	15.3	2.20
$MnCl_2$-$4H_2O$	197.91	51.1	5.06
$FeSO_4$-$7H_2O$	278.01	35.9	4.99
$(NH_4)_6Mo_7O_{24}$-$4H_2O$	1235.9	1.8	1.10
$CuSO_4$-$5H_2O$	249.68	12.6	1.57
$CoCl_2$-$6H_2O$	237.93	13.5	1.61
$CaCl_2$-$2H_2O$	147.02	99.8	7.34

2.3.3 Overnight Cultures

Bacterial strains were aseptically streaked from Microbank™ bead stocks or glycerol stocks onto LB agar (1.5% w/v) plates supplemented with the antibiotics when required. For *P. denitrificans,* the plates were then incubated for 48h at 30°C, and for *E. coli*, the plates were incubated overnight (14-18 hours) at 37°C. The plates were stored at 4°C for a maximum of 2 weeks.

Stationary phase overnight cultures were produced by inoculating 10mL of LB broth or minimal media with a single colony picked from a streak plate. Cultures were then incubated at 30°C or 37°C, plus 180-200 rpm agitation for a minimum of 14 hours.

2.3.4 Long-term strain stocks

For long term storage, the strains were stored at -80°C in Microbank™ bead stocks, prepared following manufacturer's instructions, or glycerol stocks made by adding 400uL of overnight culture to 400uL 50% glycerol, inverting to mix and then snap freezing at -80°C.

2.3.5 Aerobic batch culture

Aerobic batch culture was conducted using 50mL of LB media in sterile 250 mL volume conical flasks, supplemented with the required antibiotic and inoculated with 1:100 (v/v) overnight culture. For strains harbouring the pLMB509 overexpression vector, 10mM taurine was added to the media. To ensure maximum aeration, the strains were grown at 30°C or 37°C and agitated at 200rpm.

To measure the cell density, 1mL of culture was removed at regular intervals and the optical density (OD) at 600nm determined. After longer phases of growth, 0.1mL of culture was diluted 1:10 (v/v) in LB before the OD was measured. The rate of growth was calculated by plotting the value of the OD_{600} against time.

2.3.6 Anaerobic batch culture

Anaerobic batch culture was conducted using 250mL Duran bottles filled with 200mL of minimal media and sealed with screw-cap lids and gas-tight silicone septa. The Durans were supplemented with the required antibiotic and inoculated with 1:100 (v/v) overnight culture. For strains harbouring the pLMB509 overexpression vector, 10mM taurine was added to the media. The cultures were sparged with N_2 for 15 minutes and then incubated without agitation at 30°C.

To measure the cell density, 1mL of culture was removed at regular intervals using a 1mL syringe and the OD_{600} was determined.

2.3.6.1 Gas samples from anaerobic batch cultures

To measure headspace gas, 3ml of gas were removed from the Durans at regular intervals using a 5ml gas-tight Hamilton syringe and stored in 3mL pre-evacuated screw cap exetainer vials (LABCO). The vials were stored at 4°C.

2.4 General laboratory techniques

2.4.1 Polymerase Chain Reaction (PCR)

Polymerase chain reaction (PCR) was carried out using Taq DNA Polymerase and Phusion™ high-fidelity DNA Polymerase and their reagents. Purified chromosomal DNA and plasmid DNA were used as templates and stated where appropriate. The annealing temperature for the reactions was determined by calculating the specific primer melting temperatures (T_m). Elongation time was calculated as one minute per one kilobase-pair of product. The PCR programmes were conducted using a DNA engine PTC 300 (BIORAD) PCR machine.

2.4.1.1 Oligonucleotide design

DNA oligonucleotides (primers) were designed using the Artemis Genome Browser (Sanger). An optimal length of 18-22bp and a GC content below 60% was sought where possible. All primers were synthesised by Integrated DNA Technologies (IDT).

2.4.1.2 Colony PCR

In place of purified template DNA, a single bacterial colony was picked from an agar plate, resuspended in 20 µl nuclease-free dH_2O and denatured at 100°C for 10 minutes. The suspension was then centrifuged at high speed for 1 minute and the supernatant used as template for the PCR reaction. The initial denaturing stage of the PCR programme was extended to 5 min to further aid DNA release. The remaining stages were completed as described in 2.4.1.

2.4.1.3 PCR product purification

PCR products were purified using a QIAquick™ PCR Purification Kit by QIAGEN, following the manufacturer's instructions. Products were eluted from the purification columns in 40 µl nuclease-free dH2O and stored at -20°C.

2.4.2 Plasmid extraction and purification

For plasmid extraction and purification, the QIAprep™ Spin Miniprep Kit (QIAGEN) and the QIAprep™ spin Midiprep Kit (QIAGEN) were used, and manufacturer's instructions were followed. Plasmid DNA was eluted from the QIAprep™ spin columns (Miniprep) in 40 µl of nuclease-free dH_2O. The same volume of H_2O was used for resuspension of the air-dried cell pellets produced by the Midiprep kit. Resulting plasmid DNA was analysed for concentration and purity on a NanoDrop 2000c UV-Vis spectrophotometer (Thermo Scientific) with NanoDrop 2000 software (Thermo Scientific) and was then stored at -20°C.

2.4.3 Bacterial transformation

For bacterial transformation, two main methods were adopted during this study: electroporation using electrocompetent cells and heat shock transformation using $CaCl_2$ treated chemically competent cells. For transformation of low yield exogenous DNA electroporation was the preferred methodology.

2.4.3.1 Electrocompetent cells

Overnight cultures of the desired E. coli strains were grown and used to inoculate (1% v/v) 50 mL Lennox broth. The cultures were then incubated at 37°C, 200-250 rpm until an OD_{600} 0.4-

0.6 was achieved. The cells were then harvested by centrifugation at 4000 x g, 4°C. The supernatant was discarded, and the remaining cell pellet was gently resuspended in 15 mL ice-cold 30% (v/v) glycerol and kept on ice. Cell harvesting and resuspension in glycerol was repeated three times. During the final resuspension, a reduced volume of 2 mL was used, and the resuspended cells were aliquoted as 100 µl volumes in pre-chilled 1.5 mL microcentrifuge tubes. The cells were immediately used where possible, otherwise they were snap-frozen and stored at -80°C.

2.4.3.2 Electroporation

MicroPulser™ Electroporation Cuvettes (BIO-RAD) with a 0.2 cm cuvette gap width were pre-chilled at -20°C before use. Exogenous DNA (2-5 mL) was added to defrosted 100 µl electrocompetent cell aliquots. The reaction mix was then transferred to the electroporation cuvette and subjected to a single electrical pulse in the MicroPulser™ Electroporation Apparatus (BIO-RAD) at a voltage of 2.5 kV (programme EC2). Instantly after subjecting the cells to the electrical pulse, 1 mL LB broth was added to the transformation reaction and mixed by gently pipetting. The reaction was then transferred to a clean microcentrifuge tube and incubated at 37°C, 200rpm for 1-2 hours to allow for cell recovery.

The recovered cells were aseptically spread onto agar plates supplemented with the appropriate antibiotic in 100 µl volumes. The remainder of the cells were spun down and resuspended in 100 µl LB broth and plated onto LB agar. Once dried, the plates were incubated at 37°C for 12-16 hours to allow for selection of successful transformants.

2.4.3.3 Chemically competent cells

Overnight cultures of the desired E. coli strains were used to inoculate (1% v/v) 50 mL LB broth and incubated at 37°C, 200-250 rpm until an OD_{600} 0.4-0.6 was achieved. The cells were harvested by centrifugation at 4000 x g, 4°C. The supernatant was discarded, and the cells

were resuspended in 30 mL ice cold 1M $CaCl_2$. Cell harvesting and resuspending was repeated three times. For the final resuspension, the volume was reduced to 2mL and the cells were subsequently aliquoted into 100 µl volumes in 1.5 mL microcentrifuge tubes.

2.4.3.4 Heat Shock Transformation

Exogenous DNA (2-10 µl) was added to the microcentrifuge tubes containing chemically competent cells and the transformation reactions were incubated on ice for one hour. The cells were then subjected to heat shock at 42°C in a water bath (45 seconds) and then returned to ice for 2 min before adding 1 mL, pre-warmed LB broth. The cells were then incubated at 37°C, 200rpm, for a minimum of one hour before being spread onto the appropriate selective media.

2.4.4 In vivo genetic manipulations

Non-self-transmissible plasmids were mobilised from an *E. coli* donor strain to *P. denitrificans* (recipient) by tri-parental mating or biparental conjugation.

2.4.4.1 Tri-parental mating via filter crosses

To transfer plasmids to P. denitrificans, tri-parental mating via filter crosses was used. Plasmids were transformed into an E. coli donor strain and an E. coli helper strain containing the plasmid pRK2013 was utilised to aid transfer of the plasmid from the donor to the recipient *P. denitrificans* strain. A 50 mL overnight culture of the *P. denitrificans* recipient strain as well as two early exponential phase cultures (OD_{600} 0.4-0.6) of the donor and the helper strain were prepared and the cells harvested by centrifugation at 4000 x g, 4°C. The cell pellets were resuspended all together in 1 mL of 50% (v/v) glycerol and pipetted onto a

Whatman filter that was positioned on a solid LB Agar plate. The plate was left to dry and then incubated at 30°C for 48-54 hours. The cells were recovered in 1 mL of 50% (v/v) glycerol and six 100-fold serial dilutions prepared to achieve a dilution of 10^{-6}. A volume of 50 µl of each sample was plated onto LB Agar with the appropriate antibiotics and incubated at 30°C until single colonies could be identified.

2.4.4.2 Bi-parental conjugation

For the selection of rare events such as the generation of mutants, bi-parental conjugation using *E. coli* S17.1 as a donor was utilised. Overnight cultures of the donor strain as well as the *P. denitrificans* recipient strain were prepared. In the same microcentrifuge tube, recipient and donor were mixed in a 3:1 proportion and pelleted via centrifugation at 4000 x g, 4°C. The conjugation mixed pellet was resuspended in 1 mL Minimal Media to wash the cells which were then pelleted again and resuspended in a final volume of 100 µl Minimal Media. The cells were then plated onto Minimal Media plates containing the appropriate antibiotics and incubated at 30°C until single colonies could be identified.

2.4.5 DNA electrophoresis

To separate, visualise and analyse DNA fragments including PCR products, DNA gel electrophoresis was performed using a 1% (w/v) agarose gel prepared using 1x TBE electrophoresis buffer. The gels were stained using SYBR Safe DNA gel stain (Invitrogen). Where required, DNA loading buffer (5x, BIOLINE) was added to the samples prior to loading (1:5) and 1 kb Hyperladder (BIOLINE) was used as a size marker. Electrophoresis was carried out at 110 V for 30-45 minutes using a Sub-Cell GT electrophoresis system (BIOLINE) and DNA was visualised by UV-light using a Molecular Imager Gel Doc System (BIO-RAD).

2.4.6 Extraction of DNA from agarose gels

To extract DNA fragments including enzyme digested vectors or PCR products following agarose gel electrophoresis, gel extraction was conducted using a QIAquick™ Gel Extraction Kit (QIAGEN) following the manufacturer's instructions. The final elution of the extracted DNA from the QIAquick™ Spin Columns was made in 50 µl nuclease free dH_2O and was analysed for concentration and purity on a NanoDrop 2000c UV-Vis Spectrophotometer (Thermo Scientific).

2.4.7 Overexpression from pLMB509

For overexpression of sRNAs and other genes including Hfq, genes were cloned into the overexpression vector. pLMB509, at the first Ndel site using the gene synthesis service provided by GenScript. Plasmids were transformed into *P. denitrificans* triparental mating as described in 2.4.4.1. The cultures of the transformed strains were grown in the defined minimal medium in 250-ml Duran bottles with screw cap lids and gas-tight silicone septa. Cultures were sparged with N_2 for 10 min to impose an anoxic environment and incubated statically at 30°C. To induce sRNA overexpression taurine was added at a final concentration of 10mM.

2.4.8 Measurement of nitrous oxide in cultures

Headspace gas samples (3 ml) were taken using a 5 ml gas-tight Hamilton syringe. The samples were stored in 3 ml preevacuated screw cap Exetainer vials (Labco) and stored at 4°C. The N_2O samples were analysed by gas chromatography through injection of a 50 µL sample into a Clarus 500 gas chromatographer (PerkinElmer) with an electron capture detector and Elite-Plot Q (DVB column, 30 m by 0.53 mm inner diameter [ID]; carrier, N_2; inert portion, 95% [vol/vol] argon -5% [vol/vol] methane). Standards of N_2O (Scientific and Technical Gases) of

5, 100, 1 000, 5 000 and 10 000 ppm) were used to quantify the levels of N_2O. Total amounts of N_2O measured in a sample were calculated by applying Henrys Law constant (K_H) for N_2O at 30°C (K_H^{cc} of 0.5392). The values of N_2O calculated in micromoles were multiplied by 2 to adjust the values to micromoles of N in the form of N_2O; this takes into account the stoichiometry of N in N_2O.

2.4.9 Protein purification

For purification, proteins were tagged with a 6xHis C-terminal epitope and purified from whole cell lysate, post-overexpression, using a 5 mL HiTRap Chelating High Performance column (Cytivia) on an ÄKTAFPLC (GE Healthcare Life Sciences). Genes were cloned into overexpression vectors under the control of an IPTG inducible promoter.

2.4.9.1 Test expression

To identify high yielding conditions for protein overexpression, small-scale test expression assays were conducted. Overnight cultures of host strains harbouring the overexpression construct as well as empty vector controls were used to inoculate five flasks of 10 mL LB broth (1% v/v) supplemented with the appropriate antibiotic. The cultures were incubated at 30°C for 6 hours and a 'zero time point' sample was taken from each culture by collecting 1 mL of sample. The cells were harvested by centrifugation (2min, 14, 000 x g) and the pellets were stored at -20°C. The Inducer IPTG was serial diluted in sterile dH_2O 10-fold four times (1 M – 0.1 mM) and added to the cultures. The cultures were then incubated overnight at 30°C, 200 rpm and 1 mL samples were collected in the same way as the 'zero time point' samples. The cell pellets were then stored at -20°C.

2.4.9.2 Expression sample analysis by SDS-PAGE

To confirm successful overexpression of the recombinant protein prior to large-scale purification, the samples collected during test expression were analysed by SDS-PAGE. The sample pellets were defrosted on ice and resuspended in 50 µl SDS loading buffer before boiling at 100°C, 5 min. Cell debris was pelleted by centrifugation (>12, 000 x g, 30 sec) and 5-10 µl sample was loaded onto SDS-PAGE gels (15% v/v acrylamide) with 3 µl PageRuler^{TM} Prestained Protein Ladder (Thermo Scientific) used as a size marker. The samples were electrophoresed at 180V, one hour, in 1 x TGS running buffer and subsequently stained with SimplyBlue^{TM} SafeStain Coomassie G-250 (Life Technologies) (30-60 min) with gentle agitation. Before imaging, using white light on a Molecular Imager Gel Doc System (BIO-RAD), the gels were transferred to dH_2O and incubated with gentle agitation for 30 minutes.

2.4.9.3 Large-scale cell harvest

Once the ideal overexpression conditions had been established via the test expression assay, the favoured conditions were repeated with a larger culture volume to maximise protein yield. Overnight cultures were used to inoculate 1 L (2% v/v) LB broth in 2 L conical flasks. The cultures were incubated at 30°C, 200 rpm for 6 hours and 1 mL samples were collected as described in 2.4.9.1. Overexpression was induced upon the addition of IPTG at an appropriate concentration. The cultures were then incubated overnight (30°C, 200 rpm) and 1 mL samples were collected and analysed by SDS-PAGE as described in 2.4.9.2 to confirm successful overexpression of the recombinant protein at the larger volume. The remaining culture was transferred to 1L Beckman Coulter centrifugation bottles and the cells were harvested by centrifugation (20 min, 6 000 x g, 4°C) in a Beckman Coulter Avanti J-20 high performance centrifuge using a JLA-8.1000 rotor. Resulting cell pellets were transferred to a 50 mL Falcon tube, snap frozen and stored at -80°C.

2.4.9.4 Cell lysate collection

Frozen cell pellets (2.4.9.3) were defrosted on ice and subsequently resuspended in 10 mL of Buffer A (20mM Na-phosphate, pH 7.0; 150 mM NaCl; 25 mM Imidazole; 10% Glycerol) with one complete, mini, EDTA-free protease inhibitor cocktail tablet, 40 µg DNase I, 40 µg RNase and 50 mg lysozyme. The cells were subjected to three passes through the French Press and the insoluble cell debris was removed from the lysate via centrifugation (42,000 rpm, 40 min, 4°C). His-tagged protein was purified from the supernatant through binding to a HisTrap™ HP column, charged with 0.2M $NiSO_4$, on a ÄKTAFPLC (GE Healthcare Life Sciences), and eluted in 1mL fractions across an imidazole gradient (0-500mM Imidazole). Quality and purity of the eluted protein was ascertained by SDS-PAGE (2.4.10); 10 µL of the fractions were diluted in 50 µL SDS loading buffer and treated as described in 2.4.9.2 prior to imaging.

2.4.10 SDS Polyacrylamide Gel Electrophoresis (PAGE)

SDS-PAGE allows for size-separation of denatured proteins along an acrylamide gel matrix. A standard resolving gel of 15% (w/v) ProtoGel™. Throughout this study, acrylamide/methylene Bis Acrylamide solution (37.5:1 ratio) and a stacking gel (5% w/v) were used and cast utilising Mini-PROTEAN Tetra handcast systems (BIO-RAD) (0.75 mm and 1 mm combs and integrated spacer plates), then left to polymerise for a minimum of 30 min at RT. Gels were electrophoresed at 180 V in 1 x TGS running buffer using Mini-PROTEAN Tetra Cell systems (BIO-RAD) for up to 1 hour and then stained with SimplyBlue™ SafeStain Coomassie G-250 stain (Life Technologies) as per manufacturer's instruction. The resulting gels were imaged using white light on a Molecular Imager Gel Doc System (BIO-RAD).

2.4.11 Bradford assay

A Bradford dye-binding assay allows for quantification of protein at low concentrations by comparing protein samples of unknown concentration to a standard curve generated from bovine serum albumin samples (BSA) of known concentrations. BSA standards and the sample protein of unknown concentration were diluted in Bradford Dye Reagent solution (BIO-RAD) and dH_2O in 1.6 mL cuvettes. The samples were mixed by inverting and absorbance was measured at A_{595}. A standard curve was produced by the plotting BSA standard A_{595} values against their known concentration (mg/mL). This was utilised to deduce the protein concentration of the sample protein based on its A_{595} value.

2.5 Computational tools

2.5.1 Prediction of sRNA folding structures

To predict sRNA folding, the 'mfold' software was used (Zuker, 2003). The core algorithm of the software predicts a minimum free energy, ΔG, as well as minimum free energies for foldings that must contain any particular base pair. sRNA sequence name and RNA sequence were entered into the server input and linear foldings were predicted using the default parameters with a folding temperature of 37°C as well as ionic conditions fixed at $[Na^{+}]$=1M and $[Mg^{++}]$ = 0M. The structure folding plots were extracted as pdf files.

2.5.2 sRNA Target prediction

The key step in downstream sRNA characterisation was the prediction of putative mRNA targets. There are several existing computational methods for sRNA target prediction utilising various aspects of sRNA-mRNA binding properties. Target prediction in this thesis was carried out using two separate approaches: The machine-learning-based approach sRNARFTarget (Naskulwar and Pena-Castillo, 2022) and the web server TargetRNA2 (Kery et al., 2014).

2.5.2.1 Target prediction using the machine learning (ML) tool sRNARFTarget

The ML-based tool sRNARFTarget is generated using a random forest trained on the trinucleotide frequency difference of sRNA-mRNA pairs to predict the probability of interaction. It bases its predictions on sequence alone and can therefore be applied to any sRNA-mRNA pair without the requirement of sequence conservation. The tool was trained using known sRNA-mRNA interactions, identified predominantly in *E. coli, Pasteurella multocida* and *Synechocystis* sp PCC 6803).

All sRNA sequences for *P. denitrificans* as well as all mRNA sequences in FASTA format were inputted into the program to create all possible sRNA-mRNA pairs. The output of the pipeline was a CSV file containing predicted pairing probabilities for each sRNA with an mRNA target in descending order. For further analysis this file was converted to a .xlsx file to sort the pairs according to the sRNA or mRNA of interest. This allowed for identification of the features of the most likely interactions between sRNA and mRNA of interest.

2.5.2.2 Target prediction using TargetRNA2

TargetRNA2 uses four primary features for sRNA target identification: sRNA conservation, sRNA accessibility, mRNA accessibility and energy of hybridization. The performance of the tool was thoroughly tested using a set of verified interaction pairs in *E. coli*. A major benefit of the tool is the short execution time. Input sequences for the prediction were the sRNA nucleotide sequence in FASTA format as well the annotated *P. denitrificans* replicon taken from the NCBI Prokaryotic Genome Annotation Pipeline (NCBI Accession Nos.: NC_008686; NC_008687; NC_008688). For each candidate mRNA target, TargetRNA 2 focuses its search for an sRNA-mRNA interaction in a neighbourhood around the ribosome binding site of the mRNA and outputs a p-value for each target along with information about the target's product and a link to the corresponding gene page from RefSeq (Pruitt et al., 2012).

2.5.3 Extraction of sequence of interest

For the analysis of motifs present within promoter sequences of sRNAs or further analysis of promoter properties, promoter sequences had to be extracted in larger batches. A python script was utilised to extract 60 bp or 100 bp upstream of an sRNA sequence or a TSS. The exact position of each sRNA start codon, or each TSS of interest was inputted as a location position number and the sequence upstream was extracted and outputted in form of a .xlsx file. These sequences could then be utilised for further motif analysis.

2.5.4 Prediction of sequence motifs

A motif is an approximate sequence pattern that occurs repeatedly in a group of related sequences. These recurring motifs are often recognised by regulator proteins to signal regions of regulation or by sigma factors which control expression of regions related to the motif. MEME suite discovers novel, unmapped and recurring motifs within the input sequences and represents these as position-dependent letter-probability matrices that describe the probability of each possible base at each position in the pattern. The input is a group of sequences of interest and the output can contain as many motifs as requested. For the identification of regulator motifs, the promoter sequences of known genes within the regulon were inputted and the recurring motif was plotted as a matrix.

For the identification of putative -10 and -35 boxes, all sequences 60nt or 100nt upstream of all identified primary TSSs were extracted and submitted to MEME as well as Improbizer, a motif-finding algorithm that considers location of sequence patterns within the input sequences and favours motifs that occur at the same location (Ao et al., 2004). The programs were run using default parameters with the number of identifiable motifs set to n=2. The output motifs were ranked based on pattern frequency.

Identified motifs could be searched for within the *P. denitrificans* genome using the Artemis genome browser by using the sequence search bar (Rutherford et al., 2000). This function was utilised to identify the presence of known regulator binding sites within the sRNA promoters

2.5.5 Prediction of protein folding structures

For the prediction and analysis of protein structures and functions the Phyre2 suite was utilized to build 3D models, predict ligand binding sites and analyse domain composition in both secondary and tertiary structures based on homology detection methods (Kelley et al.,

2015). Similar to many other methods for protein structure prediction, Phyre2 relies on the comparison of a sequence of interest with a large database of sequences to construct an evolutionary or statistical profile of the input sequence.

Query amino acid sequences are inputted into the user-friendly interface of the web tool to construct an evolutionary profile based on the protein database. The protein secondary structure is predicted using PSIRED which uses neural networks trained on protein sequence profiles to identify the presence of α-helices, β-strands and coils with an average three-state accuracy of 75–80%. The resulting sequence and structure view displays the predicted secondary and tertiary structure of the input sequence, the confidence in this prediction as well as the amino acid sequence of the modelled regions. The resulting protein structure can be extracted as a PDB file.

2.5.6 Global identification of transcription start sites

It has been shown that copper availability effects the transcription of enzymes involved in N_2O production in *P. denitrificans*, specifically the activity of the final denitrification enzyme NosZ which requires large amounts of copper for its active centre. The transcriptome of *P. denitrificans* under CuH and CuL conditions was determined by Gainster et al., 2016 using differential RNA-seq. The study utilised the Trizol method to extract RNA which was then sequenced on an Illumina NextSeq 500 system using 75bp read length. The resulting raw (fastaq files) and processed date (wig files) are available on the GEO database (Series record number: GSE85362). These datasets were utilised in this study to carry out comprehensive mapping of transcription start sites (TSS) across the genome (Fig. 2.1).

Mapping an alignment of reads was carried out using Galaxy (usegalaxy.org). For quality control of the raw fastaq files, FASTQC was run to provide an overview of the data quality. Mapping of the read was carried out using Bowtie2, a fast and memory-efficient open-source tool with a particular strength in the alignment of sequencing reads of about 50 to 1,000 bases up to large genomes. Both the raw fastaq files and the reference genome were provided as

input parameters and the mapping was run using default parameters. The mapped reads were then extracted as BAM files. The mapped reads were visualised using the Integrated Genome Browser, which allowed for the viewing of sequencing results alongside the *P. denitrificans* reference genome.

TSS were identified using the automatic TSS annotation tool TSSAR. The NCBI accession numbers of the 2 chromosomes (NC_008686; NC_008687) and plasmid (NC_008688) making up of the *P. denitrificans* PD 1222 genome were provided as input alongside the files with the corresponding mapped dRNA-seq data. The default window length of 1000 was kept, as this approximately corresponds to the average gene length in bacteria. Transcription start sites are enriched in a TEX treated library compared to an untreated library, as TEX degrades reads which are not protected by a triphosphate at its 5' end, a characteristic of RNA fragments originating from primary transcripts. The statistical analysis of the TSSAR tool aims to identify significant enriched positions to call a likely TSS while dealing with background noise caused by infallible depletion. TSS were called by comparison of the mapped reads, the –TEX values were subtracted from the +TEX values and the differences between treated and untreated library was considered for each position. Since both libraries follow a Poisson distribution, the difference sample follows a Skellam distribution Enriched signal peaks were called with a cut off value of 10 reads and classed as Primary, Internal and Antisense TSSs based on their position in relation to the coding regions within the genome.

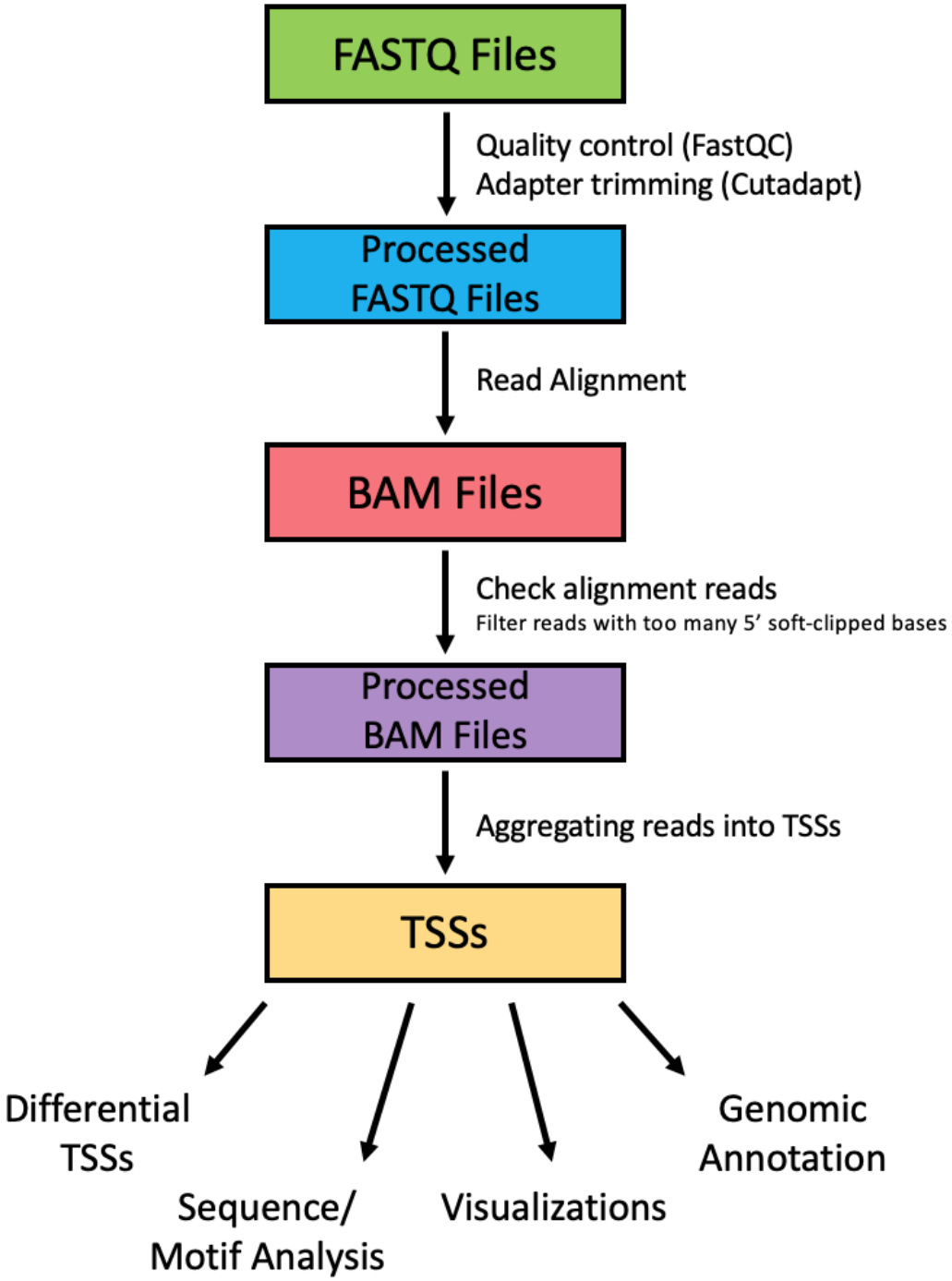


Fig. 2.1: Workflow for the computational processing of TSS mapping data into different formats with software used for each step. More information on each piece of software listed can be found at the following URLS: FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc); Cutadapt (https://cutadapt.readthedocs.io/en/stable)

2.5.7 Generating a phylogenetic tree

For the generation of a phylogenetic tree, a multiple sequence alignment (MSA) of each sigma factor amino acid sequence was generated using MAFFT (v.7.470) by aligning sequences sourced from the NCBI protein database (REF -MAFFT). THe resulting 940 amino acid alignment was used to generate a maximum likelihood phylogenetic tree using IQTree 2 (v2.1.4-beta) with model selection carried out using ModelFinder (Minh et al., 2020; Kalyaanamoorthy et al., 2017). The selected substitution model was Q.pfam+F+I+G4 with an ultrafast bootstrap value of 1,000 (Hoang et al., 2018).

3. Further characterisation of small RNAs in *Paracoccus denitrificans*

3.1 Introduction

Advances in both computational and experimental approaches to identify sRNAs have accelerated the pace of sRNA discovery across all types of bacteria. More and more sRNAs are characterised and their functions elucidated, giving further insights into the complex regulatory networks found across all types of environments. Since their initial characterisation in the 1980s, approximately 550 distinct sRNA families have been identified (Boutet et al., 2022). Bacterial sRNAs have been shown to be involved in the regulation of many important biological processes such as stress responses and infection and their abundance in microbial communities suggests they may play an even greater role in fine-tuning bacterial responses than previously known.

Initial screens for sRNA in bacteria relied primarily on computational screens of intergenic regions to identify conserved sequences or orphan promoter and terminator sequences. The discovery was enhanced through advances in whole genome expression profiling. No known bacteria have had their exact number of sRNAs, and the individual regulatory roles fully characterised. However, it seems likely that bacteria will have a few hundred regulatory sRNAs rather than thousands of sRNAs (Gottesman and Storz, 2011, Diallo et al., 2020). The number of sRNAs identified in a single bacterium is heavily dependent on the size of the genome, with larger genomes generally coding for a larger number of regulatory sRNAs. Small RNAs are prominent in highly researched model organisms with *Enterobacteriaceae* encoding 145 distinct sRNAs, while other bacterial families may only have an average of 7-10 sRNAs (Boutet et al., 2022). Therefore, sRNAs often outnumber the sigma factors encoded in a bacterium. For instance, *E. coli* encodes 7 sigma factors and 98 sRNAs. Many of the putative sRNAs that have been detected still have no known function whereas other sRNAs with important functions may be missed during sRNA detection due to their expression exclusively under very specific conditions (Sittka et al., 2008; Liu et al., 2011).

A large number of bacterial sRNA are transcribed from regions opposite of annotated genes. These so called cis-encoded sRNAs share extensive complementarity with the corresponding transcripts (Fig. 3.1A). The most prevalent roles for antisense sRNAs in bacteria have been the

repression of genes encoding toxic gene products (Gerdes und Wagner, 2007; Fozo et al., 2008) or the directed cleavage of the mRNA encoded on the opposite strand. The most extensively studied sRNAs however, are those with limited target complementarity that are encoded at a different genomic location (Fig. 3.1A). These are often referred to as trans-encoded sRNAs and are generally expressed under specific growth conditions ranging from cell envelope stress to glucose starvation (Hör et al., 2020). Pairing between a trans-encoded sRNA and its target generally involves a conserved seed region of 6-8 contiguous base-pairs (Bandyra et al., 2012). This region can often interact with multiple targets and can result in a number of regulatory outcomes. Some sRNAs bind to the ribosome binding site (RBS) thus blocking translation by preventing the entry of the ribosome. Translation can also be prevented when the region of interaction is 50 or more nucleotides upstream of the RBS (Sharma et al., 2007). Other sRNAs can pair with regions downstream of the Shine-Dalgarno Sequence without affecting ribosome entry. This has been observed for Salmonella MicC which causes an acceleration of *ompD* mRNA RNAse E-dependent mRNA decay instead without blocking ribosome entry (Pfeiffer et al., 2009).

sRNAs can also activate translation which generally occurs through preventing or overcoming the formation of an inhibitory secondary structure (Prevost et al., 2007). The sRNA-mRNA interaction can result in the remodelling of the mRNA structure making the ribosome entry site accessible and allowing translation. Binding can occur distant from the start codon making computational target prediction more difficult.

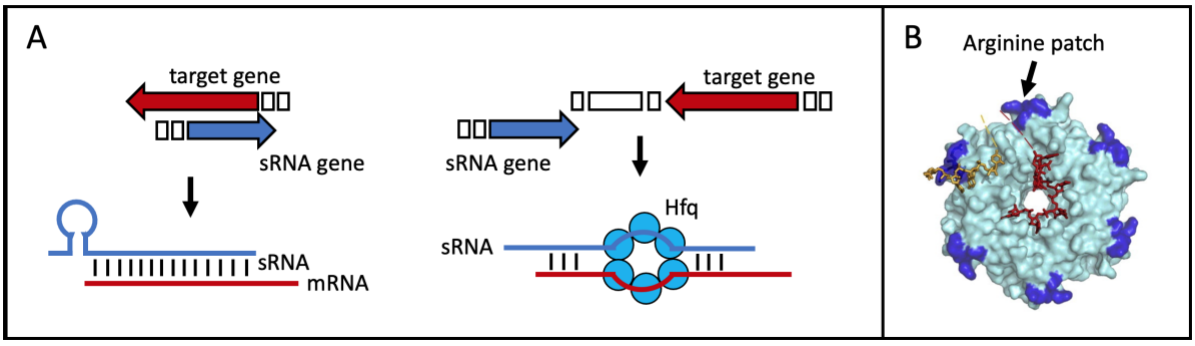


Fig. 3.1 A cis encoded sRNAs are transcribed from regions related to their targets and therefore show high levels of sequence complementarity, while trans-encoded sRNAs often require the presence of the RNA chaperone Hfq due to low levels of complementarity. **B** The tertiary structure of the *E. coli* Hfq hexamer in complex with RydC sRNA (red sticks and yellow sticks). Arginine patches are shown in blue (adapted from Zheng et al., 2017).

Many sRNAs that act via limited base pairing have been found to require the RNA chaperone Hfq which binds both sRNAs and mRNAs to stimulate their pairing (Fig. 3.1A). Hfq-dependent sRNAs contain an mRNA base-pairing region, also known as the seed region, an Hfq-binding site, and a Rho-independent transcription terminator. The seed region allows the formation of an RNA-RNA hybrid with the mRNA target to regulate both its translation and its stability. The Rho-independent terminator is a GC-rich palindrome sequence followed by a run of U residues and is responsible for transcription termination and the formation of distinct sRNA molecules (Ishikawa et al., 2012). Regions of Hfq binding have thus far only been identified for a limited number of transcripts are frequently AU-rich single stranded regions. When bound to Hfq, the RNA secondary structure is altered and local concentration of both mRNA and sRNA is increased (Brennan and Link, 2007). Hfq is also able to interact with ribosomes and with RNAse E. The impact of these interactions on the functions of Hfq are however still unclear.

The involvement of sRNAs in nitrogen cycle associated metabolism is a more recent discovery and with the identification of 167 sRNAs across the *P. denitrificans* genome via a combination of RNA-seq and bioinformatics approaches, these short regulatory RNAs were first shown to play a role in the regulation of denitrification, as introduced in **1.6.8** (Gaimster et al., 2016). The presence and size of a selected number of candidate sRNA identified in this study via RNA-seq were confirmed using RT-PCR and verified by Sanger sequencing. However, their role within the regulatory network of the model denitrifier is still unclear.

The known regulatory network controlling the denitrification apparatus in *P. denitrificans* consists of the key transcriptional regulators FnrP, NNR and NarR which respond to environmental signals including nitrate, nitrite, NO, oxygen and copper and regulate the expression of the Nar, Nir, Nor and Nos enzymes (introduced in **1.5**). Although several important links in the regulatory network of the model denitrifier have been unravelled, much remains to be discovered before a full understanding of the phenotypic response can be obtained. For instance, phenotypic responses of a variety of *P. denitrificans* mutants suggest that transcription of *nar* requires a dual control via *FnrP* and *NarR*, while the transcription of *nosZ* is equally effective with only *FnrP* or *NNR*.

3.1.1 sRNAs – the missing link in denitrification regulatory network?

The sRNAs in *P. denitrificans* are distributed across the two chromosomes and the plasmid, that make up the *P. denitrificans* genome: The majority, 110, of the sRNAs are located on chromosome 1 while 39 can be found encoded on chromosome 2 and the remaining 18 on the plasmid (Fig. 3.2). Their expression levels vary across high and low N_2O anaerobic conditions. For instance, sRNA intergenic 100 is expressed threefold higher under low N_2O emitting conditions compared to the expression level during high N_2O emitting conditions. In total, 35% of the sRNAs were differentially expressed by twofold higher or lower. Interestingly, seven sRNAs showed a greater than 7-fold change in expression and a further 3 showed a change in expression greater than 10-fold (Gaimster et al., 2016).

In order to gain a further insight into the newly identified sRNAs in *P. denitrificans* various online tools were used in the study (Gaimster et a., 2016). The secondary structures were predicted using Mfold. All sRNAs were predicted to form highly structured molecules with more than one hairpin loop suggesting that they have the potential to form complex conformations comparable to other directly acting RNA transcripts. Putative gene targets were predicted using TargetRNA and the most frequently predicted targets were transcriptional regulators such as the Xre, Fis and TetR families as well as transporter proteins including metal and ABC transporters.

Over one third of the sRNAs showed differential expression between N_2 and N_2O emitting culture conditions suggesting an involvement of the sRNAs in the switch between these two conditions. Putative gene targets for many of the sRNAs included genes encoding protein products involved in transcriptional regulation as well as transport proteins. Conservation of sRNA sequences across other species in α-proteobacteria classes such as the *Rhodobacteracea* and *Rhizobiales* was also shown for half of the confirmed sRNAs.

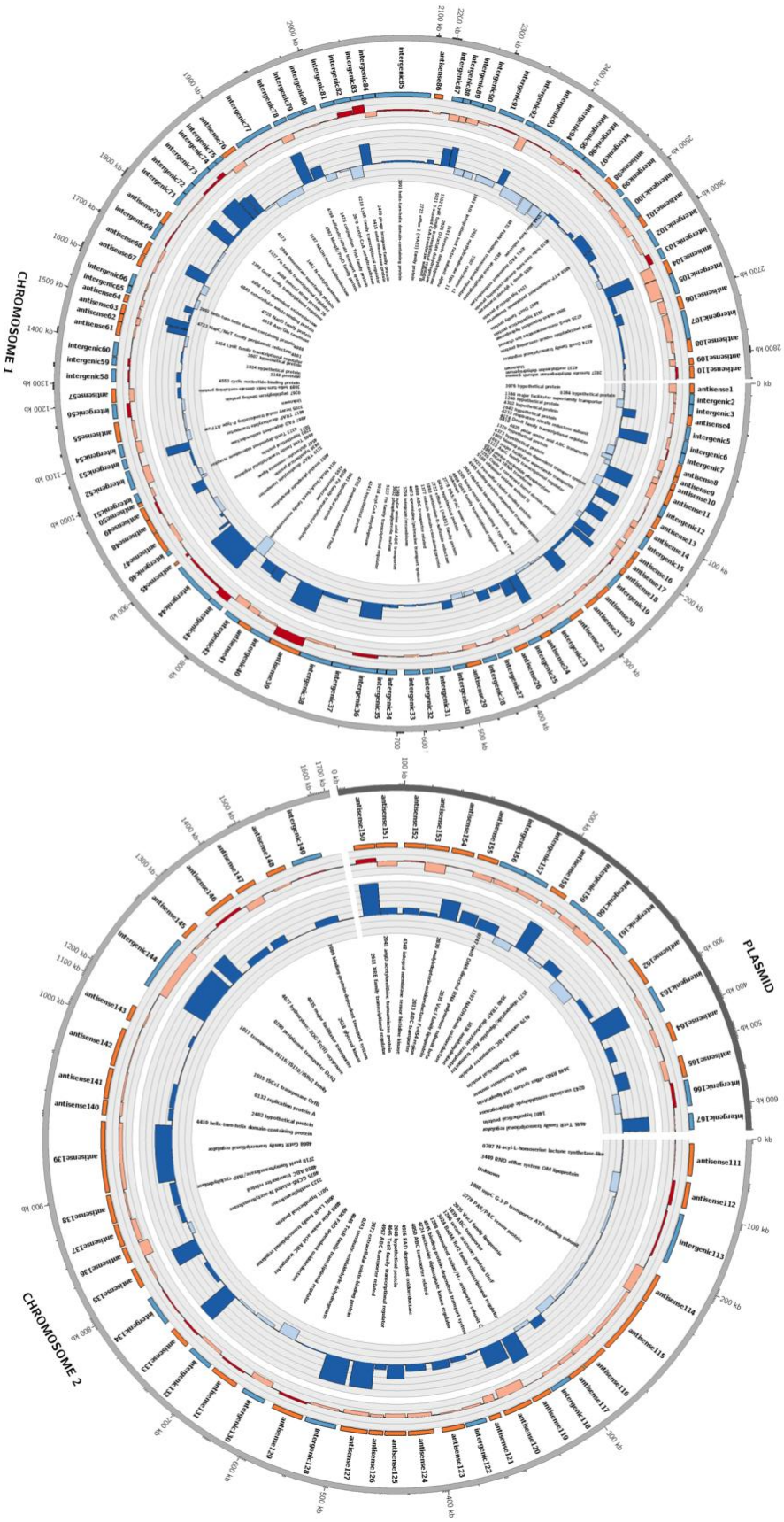


Fig. 3.2 Summary of sRNAs identified in the *P. denitrificans* transcriptome in high N_2O (anaerobic), low N_2O (anaerobic) and zero N_2O emitting (aerobic) conditions. Outer-to-inner rings: position in the *P. denitrificans*

chromosome 1, chromosome 2 or plasmid; sRNA name; sRNA relative size and location, color-coded according to intergenic (blue) or antisense to ORF (orange) positions: sRNA expression level, color coded as increased expression in high N_2O anaerobic compared to low N_2O anaerobic (dark red), or lower expression in high N_2O anaerobic compared to low N_2O anaerobic (pale red), increased expression in low N_2O anaerobic compared to zero N_2O aerobic (dark blue) or lower expression in low N_2O anaerobic compared to zero N_2O aerobic (pale blue), with each ring representing increments of 2 $\log_2$-fold units of differential expression; predicted target for sRNA, Gene identifier (pden) number is included along with gene name when known. Note: for spacing purposes the gene names for predicted targets for 5 sRNAs on chromosome 1 could not be included; 4173 TonB-dependent receptor, 4861 ABC transporter related, 4986 ATP-NAD/AcoX, kinase 0810 solute-binding protein, 5071 hypothetical protein. (Gaimster et al., 2016)

A closer characterisation of one sRNA, sRNA-29 demonstrated, that this sRNA is an important regulator of denitrification (Gaimster et al., 2019). Modulation of its expression levels impacts on nitrite reduction, therefore affecting NO and N_2O emissions from bacterial cultures. The results of this study strongly suggest, that sRNAs are crucial and yet widely uncharacterised nodes in regulation of denitrification and could be key targets for controlling cellular production and emissions of the greenhouse gas N_2O.

3.2 Aims

The aim of this chapter was to further enhance the insight on the sRNA regulatory network in *P. denitrificans* using two separate approaches for the selection of sRNAs for experimental characterisation. The first approach uses putative targets and expression patterns of sRNAs identified through computational approaches as well as RNA-seq. The second approach is based on potential regulation of sRNAs through known regulators of denitrification in *P. denitrificans*. This approach relies on the hypothesis that sRNAs regulated by one of the important regulators of denitrification are likely to be of important links in the regulatory network. Therefore, this chapter aims to answer the following questions: Which sRNAs are crucial for the switch between N_2O production and N_2O consumption? How are they involved in the denitrification pathway and when are they active and available?

Chapter hypotheses:

- There are further sRNAs present in P. denitrificans that have an impact on N_2O emissions
- Some sRNAs are likely regulated by known denitrification regulators
- The RNA chaperone Hfq has an important function in sRNA action in *P. denitrificans*

3.3 Results

The small RNAs characterised in this thesis were identified by Gaimster et al. (2016) via a combination of RNA-seq and bioinformatics approaches. For sRNA characterisation, the sRNA sequences (Table 3.1) were cloned into the overexpression vector, pLMB509 (Tett et al., 2012), at the first NdeI site using the gene synthesis service from GenScript and subsequently transformed into *P. denitrificans* by the use of triparental mating.

Table 3.1: Name, position, size and sequence of the sRNAs characterised in this thesis

sRNA name	start site	end site	size (bp)	sequence
sRNA 5	273173	273337	164	CUUUGGCGAAACGGUCCUCGUGCAGG CGGAUCGCGAACUUGAUAAUGCCAUUA UUCAGGAACUUUACGCACUCGGUCCCGA CGAACUCGCGAUCCGUGACCAGAAGCUG GAGCGUCGAGACCGGGAAAUGCGCGAGA CAGCGCGAUCCGGGUCGCGGUCGUGCU
sRNA 10	319497	319579	82	UCAUCCGAGCGGGAUCUCCGGGCGGUGCC AUGGAAGACCCCAUUGUGGGCACCGGCCC AAAGAACGACGCUCUGGACCCGAG
sRNA 18	229605	229644	39	GGAGGGAAUCACGGCGUCAUUGA CGUUGACGACCGCGAUA
sRNA 36	726444	726596	152	GCGAGUGUGUGAUUCGUCAUACGC UCGUGAUCUUCUGUCCGAGACGGA GGGCGGGCCGCAAGGUCGCUAAGU CCUUCCUGAGAUGGGGAAUCAGAC GAGAAUUCAGAUCAACGGCGCUAGC GCGAUCCAUUGGUUGGAAACAUCUG GACCUU
sRNA 39	744622	744742	120	CCGGGCCUGCUCGUGCAUGUAGC CGCAUCGCUGCCGAUCUUUCCAGA AUCGGUUGCCUUGUGCCGGGCCGC CAUAUCAUGGGGCGGCCUGCUCUU GUUUAUACGGUUCGACGCUGGCGGU
sRNA 79	1965193	1965271	78	CCACGAUGCUGUGGUAGCUCAGUG GUAGAGCACUCCCUUGGUAAGGG AGAGGUCGAGAGUUCAAUCCUCU CUCACAGC

3.3.1 Folding Structure of sRNAs

In order to gain more insight into the sRNAs in *P. denitrificans* and select putative candidates with an involvement in the denitrification regulatory network various online tools were used to provide further information. The secondary structure of sRNAs was predicted using the online tool Mfold using the default parameters (Zuker, 2003). The core algorithm of the software predicts a minimum free energy, ΔG, as well as minimum free energies for foldings that must contain any particular base pair.

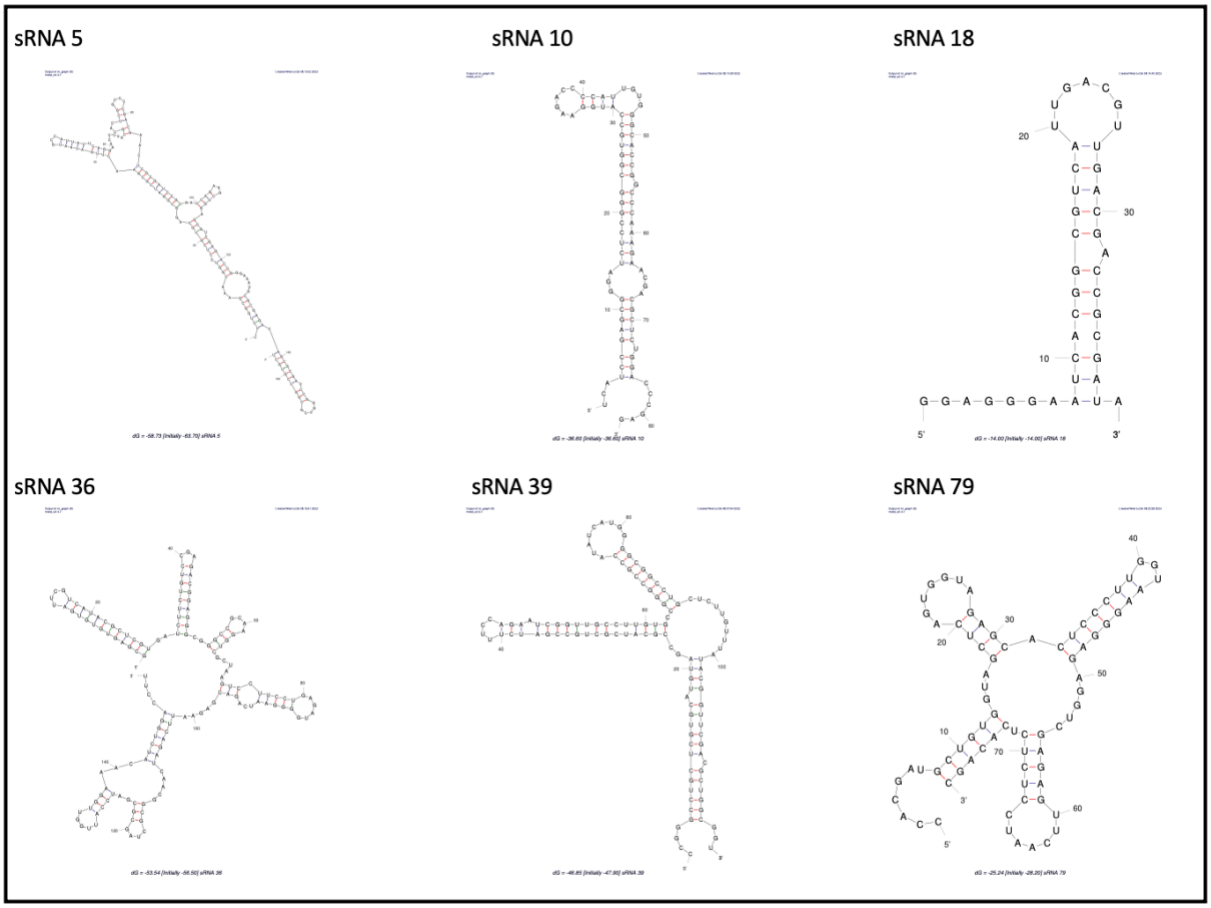


Fig. 3.3: The secondary structure for six sRNAs predicted using Mfold (Zuker, 2003). All six sRNAs are shown to fold into complex structures including one or more stem loops

All of the sRNAs, including the six sRNAs shown in Fig. 3.3 were shown to have significant predicted secondary structures. Most of the sRNAs were predicted to form highly structured molecules with more than one hairpin loop which suggests that they are likely to be capable

of forming complex conformations similar to those observed in other directly acting bacterial sRNAs. The six sRNAs shown in Fig. 3.3 were selected for further investigation due to further computational predictions made in the following sections.

3.3.2 Target prediction

Putative gene targets for all sRNAs in *P. denitrificans* were predicted using two different tools. The first approach was made using the machine-learning tool sRNARfTarget, which determined the number and probability of putative mRNA targets for each sRNA out of all known sRNAs and mRNAs across the genome (Naskulwar & Pena-Castillo, 2021). The transcriptome-wide target prediction programme was trained on the trinucleotide frequency difference of sRNA-mRNA pairs using a dataset consisting of 745 confirmed interacting sRNA-mRNA pairs, basing its predictions on sequence alone. Therefore, it does not require sequence conservation and could be applied to all sRNA identified in *P. denitrificans* (Appendix 1). This analysis resulted in three distinct sRNA groups that varied in target numbers and scores. Some sRNAs were classified as broad range sRNAs (Fig. 3.4.1), with a high number of putative mRNA targets of high target probability. For these sRNA the prediction probability score ranged between 0.3 and 0.5, with a higher score attributing to a higher interaction probability. Others only had few hits with high target probability and were therefore classified as more specific (Fig. 3.4.2). For the specific hits the prediction probability could reach a score of up to 0.55. These results suggest that the roles of sRNAs within *P. denitrificans* may vary with some of them regulating a larger number of targets and potentially taking over a global regulatory role while others only have a small number of targets, suggesting that this class of sRNAs may have a more specific regulatory role. The third group of sRNAs (Fig. 3.4.3) showed a large number of specific targets scoring around or above 0.5, whilst only having few targets with lower prediction scores.

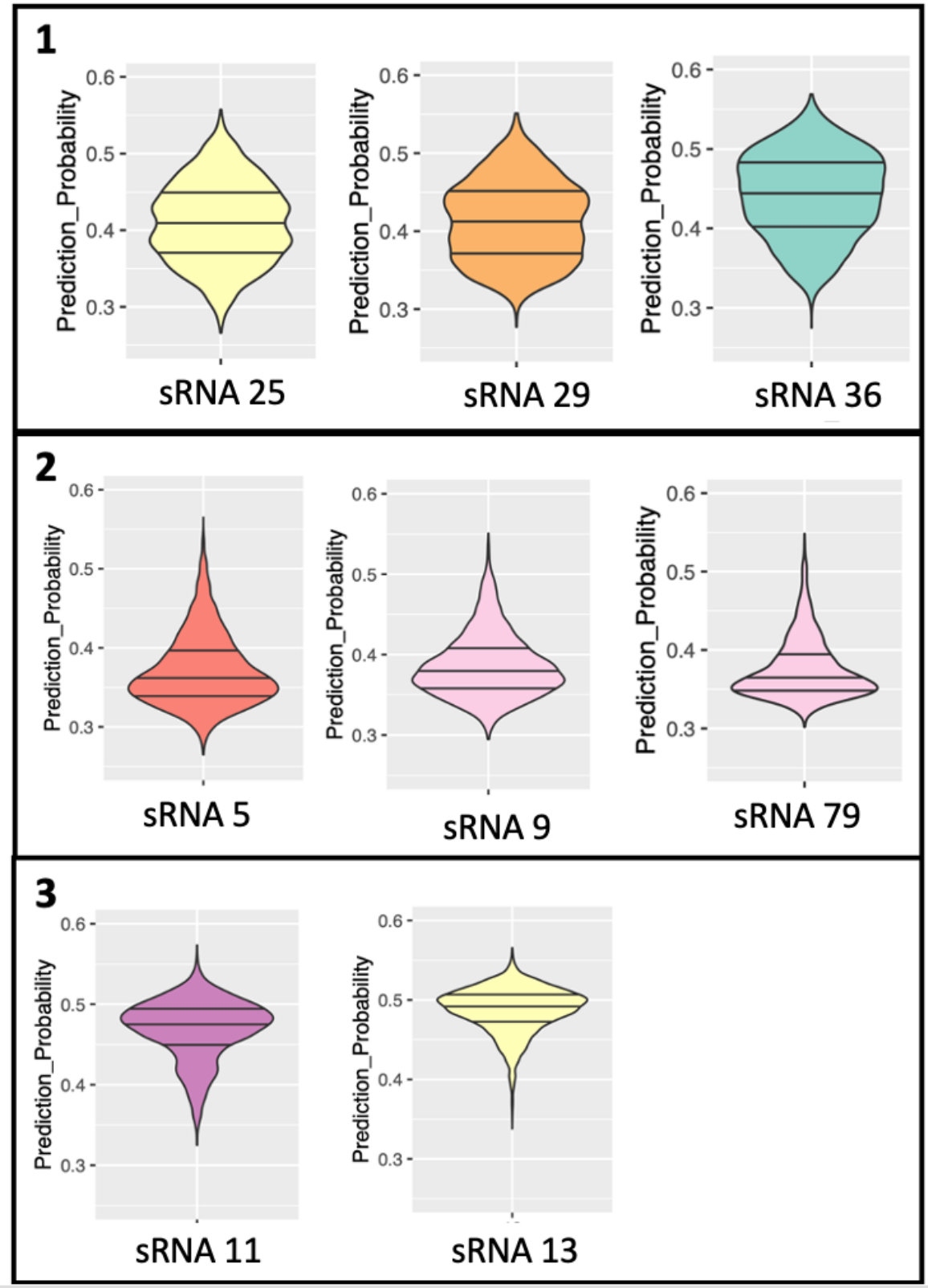


Fig. 3.4 Violin plots for the prediction probabilities for the interaction of sRNAs found in P. denitrificans with all mRNA transcripts present generated using sRNARFTarget (Naskulwar & Pena-Castillo, 2021) **1:** Globally acting sRNAs with a broad target range showing large number of putative targets with average specificity. **2:** sRNAs with high specificity for a small number of targets. **3:** sRNAs with high specificity for a larger number of targets.

Table. 3.2 Prediction probabilities for the top 5 mRNA Targets predicted with sRNARFTarget (Naskulwar & Pena-Castillo, 2021)

sRNA name	mRNA ID	prediction probability
5	CP000489.1_cds_ABL68319.1_202_[locus_tag=Pden_0205] protein=hypothetical_protein	0.54566
	CP000489.1_cds_ABL68279.1_162_[locus_tag=Pden_0162 protein=MltAinteracting_MipA_family_protein	0.54374
	CP000489.1_cds_ABL70504.1_2387_[locus_tag=Pden_2416]_ protein=conserved_hypothetical_protein	0.54221
	CP000489.1_cds_ABL69619.1_1502_[locus_tag=Pden_1519] protein=hypothetical_protein	0.54164
	CP000489.1_cds_ABL69793.1_1676_[locus_tag=Pden_1696] protein=Phosphoribulokinase	0.54135
10	CP000489.1_cds_ABL68215.1_98_[locus_tag=Pden_0098] protein=Conjugal_transfer_TraD_family_protein	0.5582
	CP000489.1_cds_ABL69793.1_1676_[locus_tag=Pden_1696] protein=Phosphoribulokinase	0.54797
	CP000489.1_cds_ABL70004.1_1887_[locus_tag=Pden_1909] protein=protein_of_unknown_function	0.54471
	CP000489.1_cds_ABL70504.1_2387_[locus_tag=Pden_2416] protein=conserved_hypothetical_protein	0.54445
	CP000489.1_cds_ABL69171.1_1054_[locus_tag=Pden_1062] protein=flavin_reductase_domain_protein__FMN-binding_protein	0.53992
18	CP000489.1_cds_ABL69041.1_924_[locus_tag=Pden_0930] protein=hypothetical_protein	0.56034
	CP000489.1_cds_ABL68540.1_423_[locus_tag=Pden_0426] protein=transcription_elongation_factor_GreA	0.55042
	CP000489.1_cds_ABL70504.1_2387_[locus_tag=Pden_2416] protein=conserved_hypothetical_protein	0.54566
	CP000489.1_cds_ABL70308.1_2191_[locus_tag=Pden_2216] protein=RNA_polymerase__sigma_32_subunit__RpoH]	0.53329
	CP000489.1_cds_ABL70179.1_2062_[locus_tag=Pden_2087] protein=conserved_hypothetical_protein	0.53312

36	CP000489.1_cds_ABL70157.1_2040_[locus_tag=Pden_2065] protein=cold-shock_DNA-binding_protein_family	0.56051
	CP000489.1_cds_ABL69087.1_970_[locus_tag=Pden_0976] protein=hypothetical_protein	0.55714
	CP000489.1_cds_ABL69720.1_1603_[locus_tag=Pden_1621] protein=hypothetical_protein	0.55608
	CP000489.1_cds_ABL69128.1_1011_[locus_tag=Pden_1017] protein=transposase_IS116/IS110/IS902_family_protein]	0.55284
	lCP000489.1_cds_ABL69571.1_1454_[locus_tag=Pden_1471] protein=transcriptional_regulator__LysR_family]	0.55255
39	CP000489.1_cds_ABL69780.1_1663_[locus_tag=Pden_1683] protein=monosaccharide_ABC_transporter_ATP-binding_protein CUT2_family	0.56055
	CP000489.1_cds_ABL69793.1_1676_[locus_tag=Pden_1696] protein=Phosphoribulokinase	0.55535
	CP000489.1_cds_ABL69065.1_948_[locus_tag=Pden_0954] protein=cold-shock_DNA-binding_protein_family	0.55486
	CP000489.1_cds_ABL69078.1_961_[locus_tag=Pden_0967] protein=hypothetical_protein	0.55398
	CP000489.1_cds_ABL69994.1_1877_[locus_tag=Pden_1899] protein=Substrate binding region of ABC-type glycine betaine transport system	0.55389
79	CP000489.1_cds_ABL70120.1_2003_[locus_tag=Pden_2028] protein=protein_of_unknown_function_DUF465	0.54812
	CP000489.1_cds_ABL69164.1_1047_[locus_tag=Pden_1055] protein=hypothetical_protein	0.53758
	CP000489.1_cds_ABL68227.1_110_[locus_tag=Pden_0110] protein=hypothetical_protein	0.53473
	CP000489.1_cds_ABL69471.1_1354_[locus_tag=Pden_1370] protein=hypothetical_protein	0.5329
	CP000489.1_cds_ABL69332.1_1215_[locus_tag=Pden_1227] protein=hypothetical_protein]	0.52304

Across all 167 sRNAs, as well as for the six sRNAs shown in Table 3.2, over one third of sRNARFTarget-predicted targets were hypothetical proteins or proteins of unknown function. Although their function is yet to be elucidated, the study of sRNAs may be helpful in eventually assigning a function to these proteins. Another large group of commonly predicted targets were transcriptional regulators such as the Xre, Fis and LysR families. An example for this is sRNA 36, which has a prediction score greater than 0.5 for a LysR family regulator protein. Other commonly predicted targets were proteins involved in DNA and RNA metabolism, such as the RNA polymerase sigma 32 subunit RpoH which is predicted to be targeted by sRNA 18. Transporters such as metal and ABC transporters are also commonly predicted targets. Gaining further insights into these targets and assigning functions to unknown proteins targeted by sRNAs in *P. denitrificans* will aid in further deciphering the complex denitrification regulatory network.

In order to see if any sRNAs were likely to target any of the catalytic subunits of denitrification reductases, the sRNARFTarget output was searched for the highest scores of sRNA-mRNA interactions of the denitrification-associated mRNAs (Fig. 3.5). For all four reductases NirS, NarG, NorB and NosZ there were several sRNAs predicted to be capable of interacting with the mRNA transcript. For both NirS and NosZ, five of the sRNAs showed predicted binding scores of above 0.5 suggesting a higher likelihood of interaction with one or more of the predicted sRNAs. The predicted sRNAs to interact with *nosZ* mRNA included sRNA 39, sRNA 47 and sRNA 55. Two sRNAs, sRNA 55 and sRNA 85 were predicted to target both *nirS* and *nosZ* mRNAs. Uncovering direct sRNA-induced regulation of the denitrification enzymes would confirm the importance of these regulatory molecules in the denitrification regulatory network and shed further light onto the roles of the individual sRNAs in *P. denitrificans*.

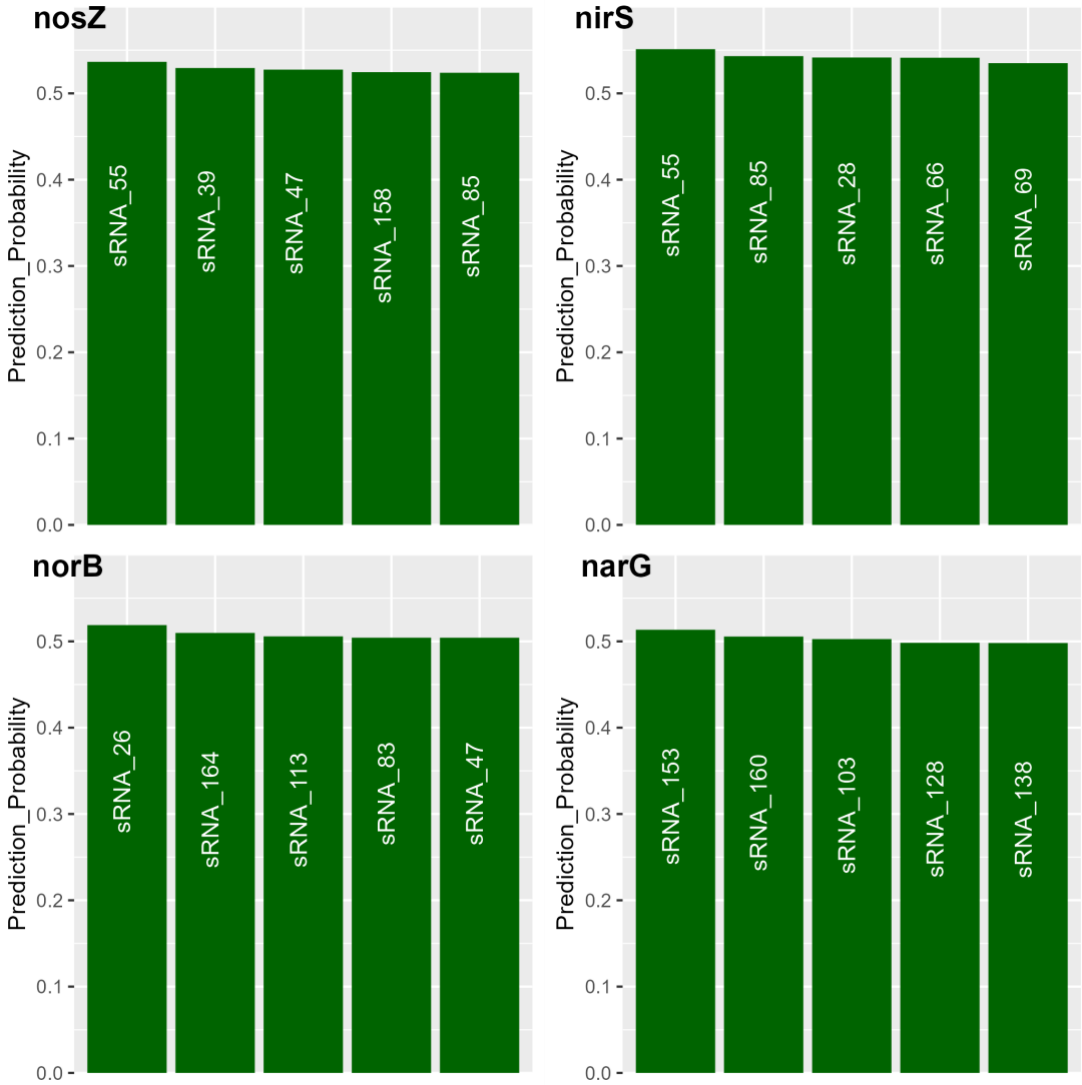


Fig. 3.5 Prediction scores for the 5 top scoring sRNAs for interaction with the catalytic subunits of the four key denitrification reductases NosZ, NirS, NorB and NarG predicted using sRNARFTarget

The second approach for target prediction was via TargetRNA 2 (Kery et al., 2014). Target predictions were generated using default parameters. The top three most energetically favourable targets for each of the six sRNAs also included hypothetical proteins as well as transcriptional regulators such as the GntR ad Fis families (Table 3.3). Transcriptional regulators were predicted as targets in 118/167 sRNAs. For instance, both sRNA 36 and sRNA 79 were predicted to target the Fis-family transcriptional regulator Pden_5127.

Table 3.3: Putative gene targets for sRNAs 5, 10, 18, 36, 39 and 79 predicted using Target RNA 2

sRNA	Putative gene targets (Target RNA)
sRNA 5_490	**Chromosome 1** 1. Pden_2778 PAS/PAC sensor protein 2. Pden_2510 lytic murein transglycosylase 3. hisI phosphoribosyl-AMP cyclohydrolase **Chromosome 2** 1. Pden_4084 response regulator receiver protein Plasmid - 1. Pden_5130 FecR 2. Pden_5125 monooxygenase 3. Pden_4985 pyruvate dehydrogenase (acetyl-transferring)
sRNA 10_490	**Chromosome 1** 1. Pden_1288 monovalent cation/H+ antiporter subunit C 2. Pden_2218 sporulation domain-containing protein 3. Pden_1166 major facilitator superfamily transporter **Plasmid** 1. Pden_4573 hypothetical protein 2. Pden_5044 aldo/keto reductase 3. Pden_4986 ATP-NAD/AcoX kinase
sRNA 18_489	**Chromosome 2** 1.Pden_3255 Cupin 2, conserved barrel domain protein 2. Pden_4274 GntR family transcriptional regulator; K00375 GntR family transcriptional regulator / MocR family aminotransferase 3.pyrE orotate phosphoribosyltransferase **Plasmid** 1.Pden_5016 Serine O-acetyltransferase; K00640 2.Pden_4998 ABC transporter related; amino acid/amide ABC transporter ATP-binding protein 1, HAAT family; K01995 branched-chain amino acid transport system ATP-binding protein 3.Pden_4748 glycine betaine/L-proline ABC transporter, ATPase subunit; K02000 glycine betaine/proline transport system ATP-binding protein
sRNA 36_489	**Chromosome 2** 1. Pden_3492 4-carboxymuconolactone decarboxylase 2. Pden_2983 pseudoazurin 3. Pden_3561 transposase IS116/IS110/IS902 family protein **Plasmid** 1. Pden_5127 Fis family transcriptional regulator 2. Pden_5002 rhodanese domain-containing protein 3. Pden_4919 methionine aminopeptidase
sRNA 39_489	**Chromosome 2** 1.Pden_3981 hypothetical protein; protein of unknown function DUF983 2.Pden_2948 antifreeze protein, type I; 3.Pden_3752 hypothetical protein; conserved **Plasmid** 1.Pden_5103 poly-beta-hydroxybutyrate polymerase domain-containing protein; K03821 polyhydroxyalkanoate synthase 2.Pden_4545 TRAP C4-dicarboxylate transport system permease DctM subunit; 3.Pden_4822 short-chain dehydrogenase/reductase SDR
sRNA 79_489	**Chromosome 1** 1. ihfA integration host factor subunit alpha 2. Pden_2502 hypothetical protein

3. Pden_1041 aspartate dehydrogenase
Chromosome 2
1. Pden_4169 sulfonate/nitrate transport system substrate-binding protein
2. Pden_4147 SoxS
3. Pden_4422 alkylhydroperoxidase
Plasmid
1. Pden_5127 Fis family transcriptional regulator
2. Pden_4724 nucleoside diphosphate kinase regulator
3. Pden_4798 phenylacetic acid degradation protein paaN

Another class of commonly predicted targets were again transporters such as metal and ABC transporters. These were predicted in 100/167 including the transporter Pden_4998, an ABC transporter related amino acid/amide ABC transporter ATP-binding protein predicted to be targeted by sRNA 18. ATP-binding cassette (ABC) transporters are integral ATP-powered membrane proteins that are required for the translocation of many substrates across membranes (Rees et al., 2009). Some may be involved in the translocation of substrates required in the denitrification regulatory cascade.

3.3.3 Experimental sRNA screening

It was previously identified that copper as well as oxygen availability has an impact on the regulation of the denitrification apparatus and subsequently affects emissions of N_2O from a culture (Felgate et al., 2012; Sullivan et al., 2013). A manipulation of these well-characterised parameters made it possible to identify the sRNA complements transcribed by *P. denitrificans* under N_2O-producing (Cu-restricted conditions) and N_2O-consuming (Cu repletion) conditions (Gaimster et al., 2016). Using this information, a number of sRNA were picked for further screening, based on their expression patterns during complete and incomplete denitrification conditions. Batch cultures of *P. denitrificans* were grown under anaerobic batch-denitrifying conditions at 30°C with nitrate as the electron acceptor with or without the addition of copper. The sRNAs were then overexpressed in *trans* from a taurine-inducible promoter to determine their impact on N_2O emissions.

3.3.3.1 Characterisation of sRNAs 5 and 10

The sRNAs were overexpressed from a taurine-inducible promoter in the vector pLMB509, where an addition of 10 mM taurine is known to induce an increase in expression of up to 15-20-fold. The OD_{600} of the overexpression cultures for sRNA 5 and sRNA 10 as well as an empty vector control were collected at regular intervals (Fig. 3.6A).

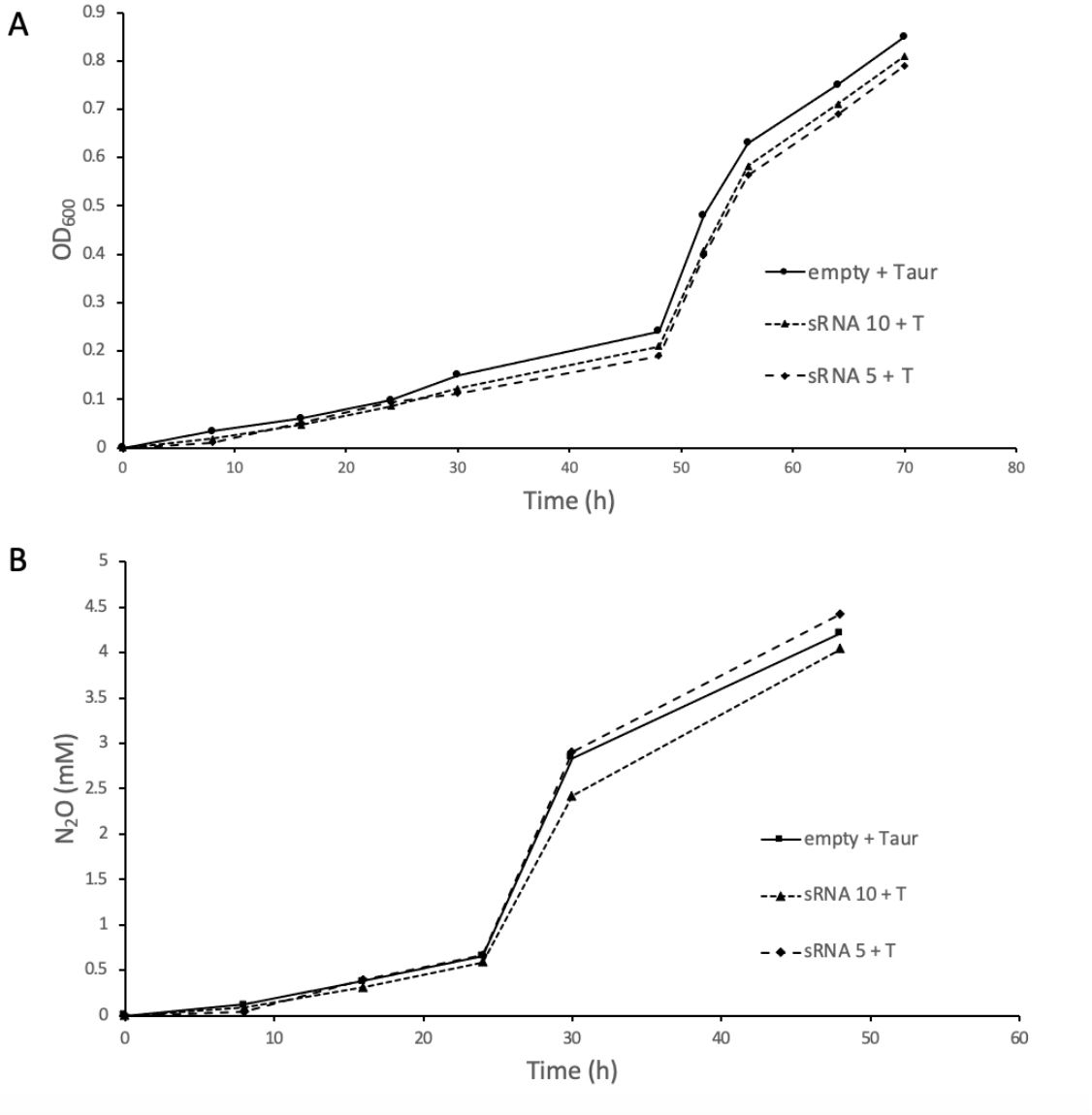


Fig. 3.6 Overexpression of sRNAs 5 and 10 does not affect the growth rate or the amount of N_2O produced by P*. denitrificans*. (A) Cultures of *P. denitrificans* plus empty pLMB509 vector (squares) as well as cultures containing pLMB509 plus sRNA 5 (diamonds) or sRNA 10 (triangles) were grown under denitrification conditions with 10mM added taurine. The OD_{600} was measured at regular intervals (B) N_2O levels were measured

For both sRNA there was no change in growth when compared to the empty vector control, therefore suggesting that these two sRNAs did not promote or block any important processes involved in cell growth and replication. To assess the impact of sRNA activity on N_2O emissions, gas samples were taken at regular intervals to measure any potential sRNA-induced changes in emissions. For sRNA 5 and 10 no changes in N_2O emissions could be observed suggesting that their role in *P. denitrificans* may not be directly related to the switch between N_2O production and consumption and further examinations will be required to uncover their role within the regulatory network (Fig. 3.6B). sRNAs 5 and 10 were not taken forward for further study in this thesis.

3.3.4 sRNA 39 overexpression increases N_2O emissions from denitrifying cultures

One sRNA, sRNA 39, however was selected for further investigation through the previously mentioned screening process, as after overexpression of this sRNA, a significant increase in the amount of N_2O emitted from the culture could be observed compared to the amount measured in the empty vector control cultures (Fig. 3.7B). While the OD_{600} remained the same despite overexpression of sRNA 39 (Fig. 3.7A), the rate of N_2O emissions increased more rapidly after 20 h of incubation time. Approximately 4.2 mM N_2O were produced by the empty vector control culture after 48 h compared to 5.1 mM measured in the sRNA 39 overexpression culture. sRNA-39 is a 120 bp long intergenic sRNA that is located within Pden_757. It is predicted to form a highly complex structure with stem loops (Fig. 3.3).

A BLASTn comparison of the sRNA-39 sequence showed that there is a conservation within the order of Rhodobacteraceae as homology was observed within the strain *Paracoccus aminophilus JCM 7686.* sRNA 39 is most highly expressed under low N_2O emitting anaerobic conditions (Fig. 3.7C). Lowest expression levels can be observed under anaerobic, high N_2O emitting conditions. It has a high number of mRNA targets that it is likely to interact with (Fig. 3.7D). These include ABC-transporters, hypothetical proteins, as well as the catalytic subunit

of the denitrification enzyme NosZ, suggesting a direct involvement of this sRNA in denitrification and the switch between N_2O emission and N_2O reduction.

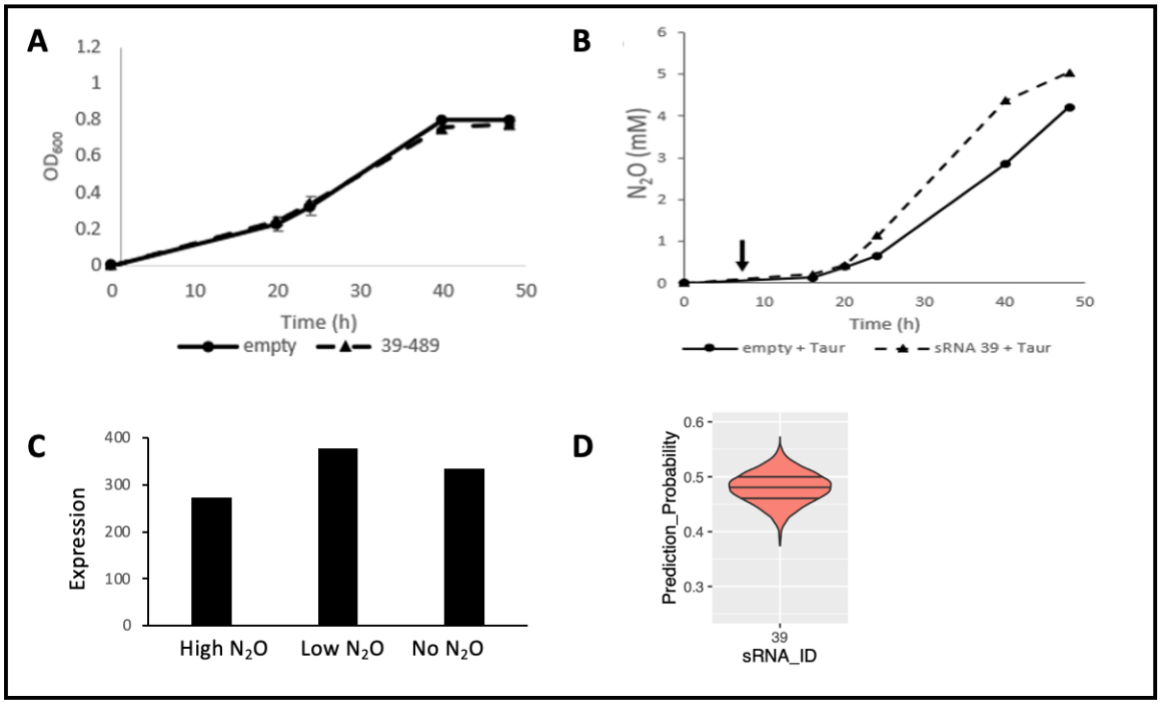


Fig. 3.7 sRNA 39 causes an increase in the amount of N_2O produced by *P. denitrificans*. Overexpression cultures of *P. denitrificans* plus empty pLMB509 vector as well as *P. denitrificans* plus pLMB509 plus sRNA 39 were grown under denitrifying conditions. (A) The OD600 was measured and (B) the levels of nitrous oxide emitted from the cultures were measured at regular intervals. (C) Expression levels shown as reported previously (Gaimster et al., 2016). (D) Distribution of predicted mRNA targets for sRNA 39.

3.3.4 sRNAs 18, 36 and 79 possess a denitrification regulator motif in their promoter region

A second approach to screen sRNAs for further investigation was to search for sRNAs with a putative denitrification regulator binding site upstream of their promoter. The denitrification regulators in *P. denitrificans* are known to recognise specific sequences upstream of their target genes (Fig. 3.8). In its active form FnrP forms a dimer and contains a helix-turn-helix domain that binds to the motif TTGATnnnnnTCAA. NNR recognizes the sequence TTnACnnnnnTCAA and NarR binds to the sequence TTGATnnnnnTCAA. sRNAs with a binding motif within 100bp of the sRNA sequence were considered likely to be regulated by one of

the denitrification regulators. These were identified using the Artemis genome browser by searching for the presence of recognition motifs within 100bp of the sRNA promoter (Carver et al., 2011). Through the initial screening process, eight sRNAs with a putative denitrification regulator binding site were identified (Fig. 3.9). Five sRNAs, sRNA 12, 18, 27, 36 and 79 had a putative FnrP site upstream of their promoter. Four sRNAs, sRNA 1, 12, 54 and 96 were identified to have a putative NNR binding site upstream of their promoter and a single sRNA, sRNA 27 was identified to have a putative NarR motif.

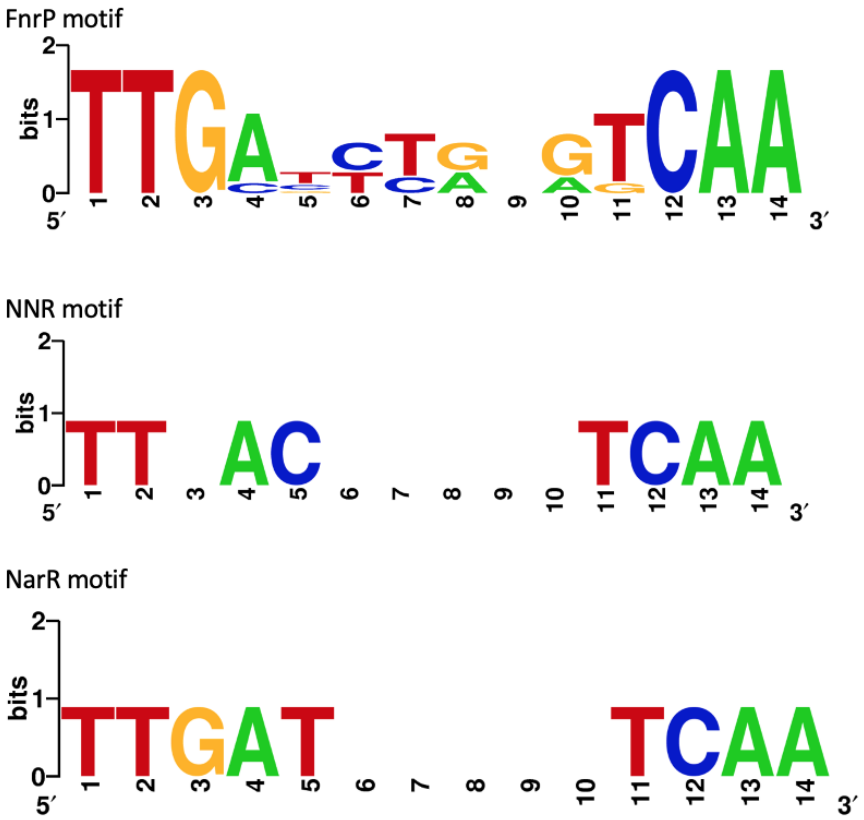


Fig. 3.8 FnrP, NNR and NarR motifs generated using MEME Suite based on known promoter sequences, bits show the confidence levels for each base at every position (Timothy et al., 2015)

The expression levels (Gaimster et al., 2016) of the identified sRNAs were highly variable with some, such as sRNA 18 and 36, showing highest expression levels during high N_2O emitting conditions and lowest expression levels during aerobic conditions. Others, including sRNA 12

and sRNA 54 showed highest expression under aerobic, zero N_2O emitting conditions and lowest expression in denitrifying, high N_2O-emission cultures. One sRNA, sRNA 96 showed little to no expression under anaerobic culture conditions, but high levels of expression under aerobic conditions. These observations suggest that expression of these sRNA is switched on and off by different cues resulting in a regulation of different responses depending on when they are activated or inactivated. Three of these sRNA, sRNA 18, 36 and 79 were selected for further experimental investigation to elucidate their role in the regulatory network controlling denitrification in *P. denitrificans*.

The three selected sRNA were further characterised using the computational approaches outlined in 3.3.1 and 3.3.2. All three sRNA were predicted to fold into complex structures (Fig. 3.9 B, D, F) and putative targets were transcriptional regulators such as a Fis-family transcriptional regulator predicted to be targeted by sRNA 36 as well as transport proteins including a sulfonate/nitrate transport related protein targeted by sRNA 79 (Table 3.4). Sequence conservation was investigated using BLASTn, and all hits with nucleotide identity high than 60% combined with a coverage between query and subject sequence higher than 80% were considered to be conserved. All three sRNAs showed sequence conservation across other Gram-negative bacteria such as *Neorhizobium* or closely related members of the *Paracoccus* clade including *Paracoccus aminovorans* and *Paracoccus suum*. It seems likely, that these sRNAs are specific to closely related bacteria.

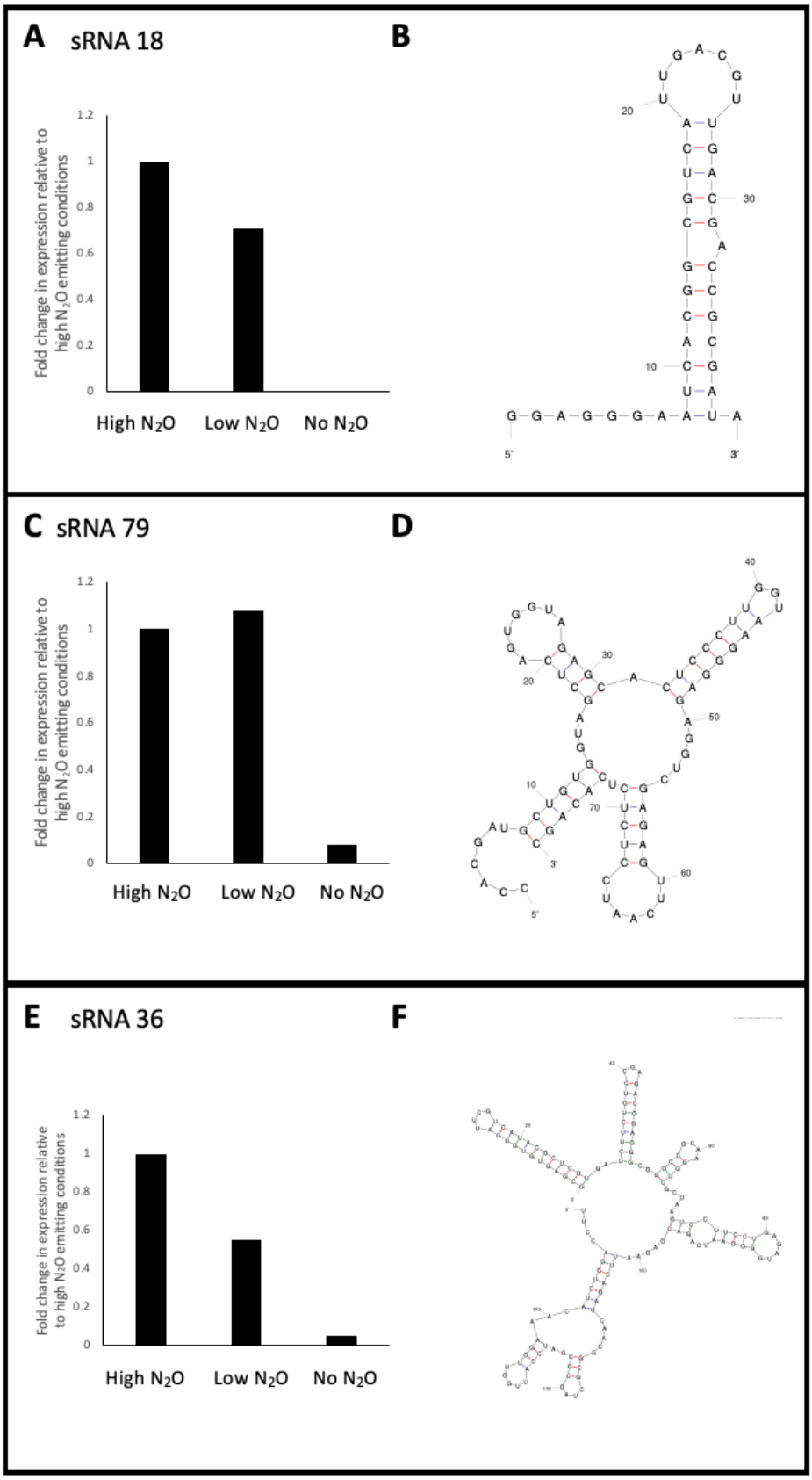


Fig. 3.9 Fold change in sRNA expression relative to expression during high N_2O emissions on the y axis for sRNAs 18 (A), sRNA 79 (C) and sRNA 36 (E) shown as reported previously (Gaimster et al., 2016) as well as their corresponding folding structures predicted using mfold (B, D, F)

Table 3.4: Summary of sRNA 18, 36 and 79 denitrification regulator motifs identified upstream of the sRNA sequence, putative targets identified using TargetRNA and conservation of the sequences across different orders, classes, and species of proteobacteria

sRNA Promoter	motif	putative regulator	putative sRNA target	Conservation (blastn)
18_489	**TTG**TCGATGAA**CCA**	FnrP, NarR	Pden_3255 Cupin 2, conserved barrel domain protein	N*eorhizobium* sp. NCHU2750, *Rhizobium indicum* strain JKLM 13E; *Mesorhizobium* sp. WSM1497; Mesorhizobium amorphae CCNWGS0123; *Mesorhizobium cicero* strain CC1192; *Paracoccus suum* strain SC2-6; *Aureimonas* sp. AU20
36_489	**TTG**GCTCTTGG**CAA**	FnrP	Pden_5127 Fis family transcriptional regulator	*Octadecabacter sp*. SW4; *Xuhuaishuia manganoxidans* strain DY6-4; *Celeribacter manganoxidans* strain DY25; *Paracoccus sanguinis* strain OM2164; *Rhodobacterace*a QY30
79_489	**TTG**GATTTTTT**CCA**	FnrP	Pden_4169 sulfonate/nitrate transport system substrate-binding protein	*Paracoccus zhejangensis* strain J6; *Paracoccus aminovorans* JCM7685; *Paradevosia shaoguanensis* strain J5-3; Devosia sp. D6-9; *Paracoccus yeei* strain FDAARGOS_643; *Haematobacter massiliensis* strain OT1, *Paracoccus mutanolyticus* strain RSP-02

3.3.5 sRNA 18 overexpression causes reduction in N_2O emissions from denitrifying cultures

Using the same overexpression as mentioned in 3.3.3, sRNA 18, 36 and 79 were overexpressed from a taurine-inducible promoter in the vector pLMB509. An addition of 10 mM taurine was added to induce expression up to 15-20-fold. Subsequently the OD_{600} of the overexpression cultures as well as an empty vector control were collected at regular intervals. All three sRNA overexpression cultures showed similar growth levels as observed in the empty vector control culture, with a maximum OD_{600} of around 0.8 reached after 60-65 hours post inoculation (Fig. 3.10 A). N_2O emissions were also monitored at regular intervals. While sRNA

79 showed similar levels of N_2O emissions to the empty vector control, the overexpression of sRNA 18 and sRNA 36 resulted in changes to the levels of N_2O measured in the cultures.

The overexpression of sRNA 36 resulted in an enhanced rate of N_2O production in the first 32 hours of the experiment (Fig. 3.10B). While the N_2O production rate during the first 32 hours of the experiment was 0.11 mM N_2O/h for sRNA 36, the rate measured for the empty vector control was 0.076 mM N_2O/h. After 32 hours the rate of N_2O production decreased and peaked at comparable levels measured to those in the empty vector control culture. This suggests that increased levels of sRNA 36 activate the denitrification machinery required for the production of N_2O.

The overexpression of sRNA 18 however, resulted in significantly lower levels of N_2O emitted from the culture, with the highest amount of 0.2 mM measured at 56 hours post inoculation. This clear change in the level of N_2O flux indicated, that the sequential reduction of NO_3^- to N_2 was some way perturbed by an overexpression of sRNA 18.

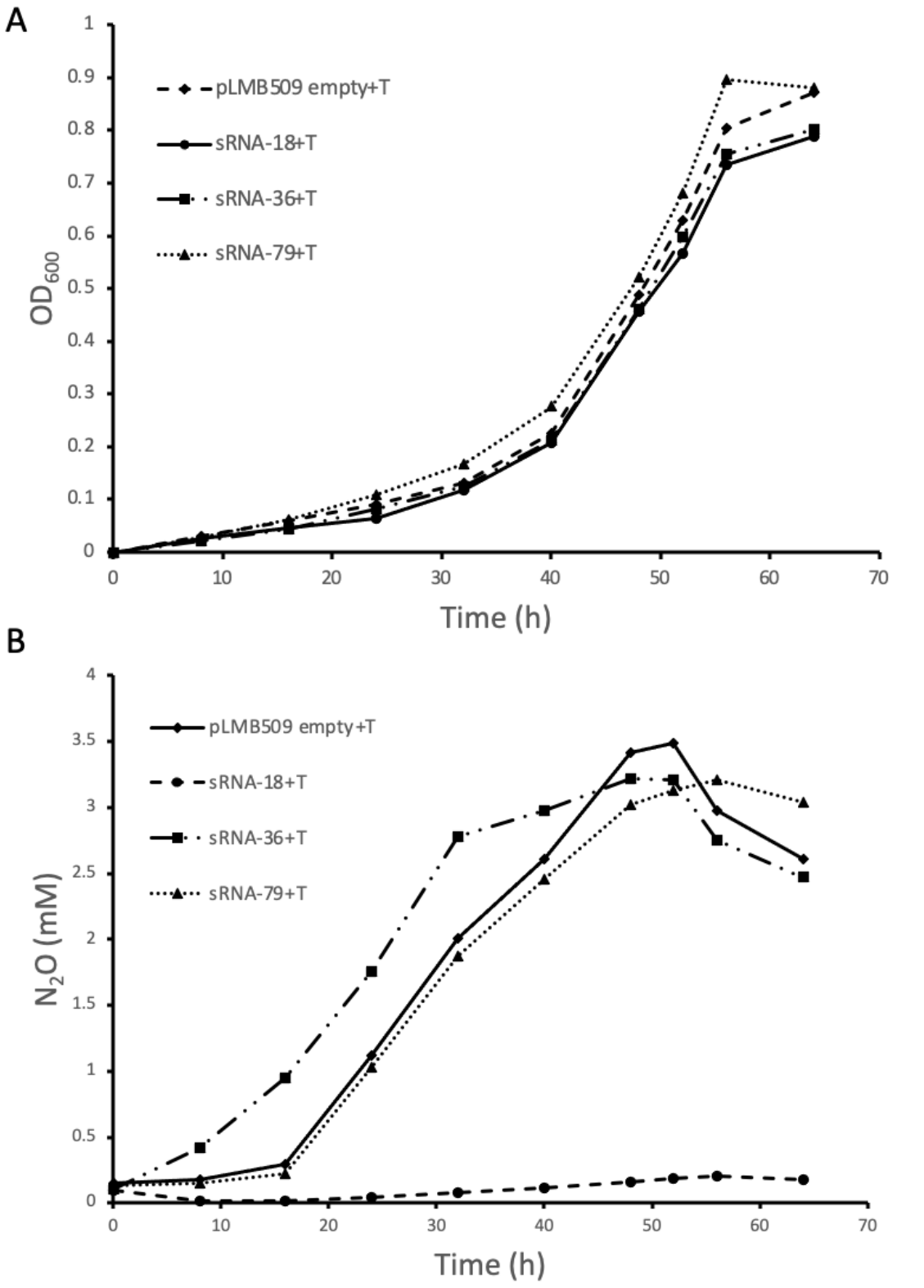


Fig. 3.10 Representative dataset for the overexpression of sRNAs 18, 36 and 79. For all three the overexpression does not affect the growth rate. The amount of N_2O produced by *P. denitrificans* is however affected by sRNA 18 and sRNA 36 **(A)** Cultures of *P. denitrificans* plus empty pLMB509 vector (squares) as well as cultures containing pLMB509 plus sRNAs were grown under denitrification conditions with 10mM added taurine. The OD_{600} was measured at regular intervals **(B)** N_2O levels were measured at regular intervals

3.3.6 Arginine patch on the Hfq rim predicts RNA annealing activity

Hfq facilitates interactions between sRNA molecules and their mRNA targets. In many bacteria an *hfq* deletion results in stress intolerance or reduced virulence. Based on the RNA annealing activity of Hfq from *E. coli*, *Pseudomonas aeruginosa*, *Listeria monocytogenes*, *Bacillus subtilis* and *Staphylococcus aureus* it was established, that RNA annealing activity increases with the number of arginines in a semi-conserved patch on the rim of the Hfq hexamer (Zheng et al., 2016). Thus, the amino acid sequence of the arginine patch can predict the chaperone function of Hfq in sRNA regulation in a bacterial species. As the role of Hfq (Pden_4124) in *P. denitrificans* is yet to be elucidated, the amino acid sequence was aligned to those of the other bacterial species listed in the study by Zheng et al., 2016, the arginine patch was identified and subsequently compared to the sequence of the other listed bacterial species.

The amino acid sequence of the arginine patch in *P. denitrificans* Hfq consists of one arginine followed by a lysine residue, a glycine residue and finally another lysine residue. The lower arginine content suggests that *P. denitrificans* has a lower sRNA annealing activity than bacteria with strong arginine patches such as *E. coli* and *P. aeruginosa*, however, based on the sequence it is likely to have similar impact on sRNA regulation as observed in L. monocytogenes (Fig. 3.11; Table 3.5). Therefore, Hfq is likely to be of importance for sRNA function in *P. denitrificans*, and further characterisation of the RNA chaperone is necessary.

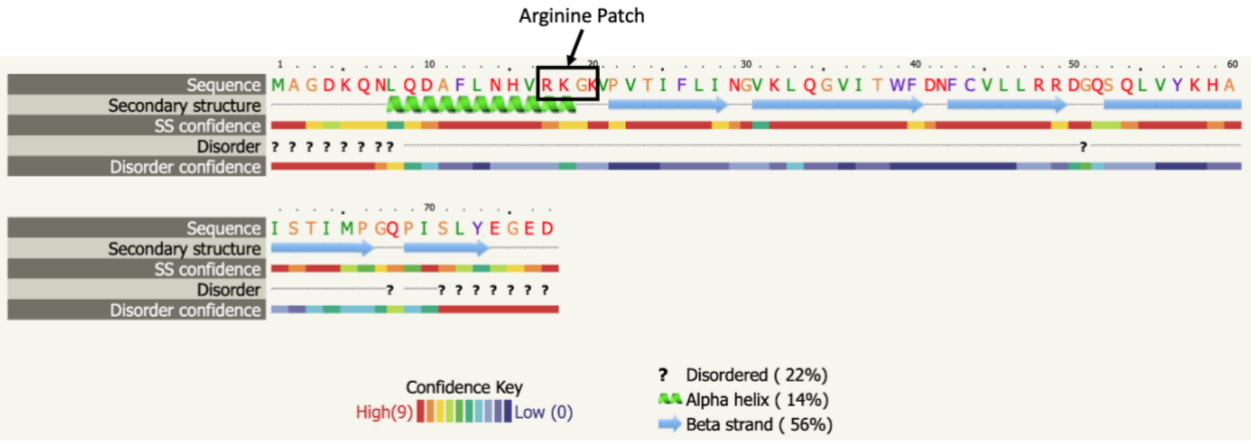


Fig. 3.11 Predicted secondary structure of the *P. denitrificans* Hfq monomer with marked location of the arginine patch, predicted using PHYRE2 (Kelley, 2015). Alpha helices are marked in green; beta sheets are marked in blue.

Table 3.5: Amino acid sequences of the conserved patches on the Hfq rim across several bacterial species (adapted from Zheng et al., 2016). Stronger sequences with a higher arginine content resulted in higher sRNA binding ability and therefore a stronger involvement in sRNA regulation

Rim	Hfq	Annealing Rate	sRNA binding	sRNA regulation
RRER	*E. coli*	100X	3nM	++
RKER	*P. aeruginosa*	10X	33nM	++
RKEK	*L. monocytogenes*	3X	18nM	+
RKEN	*B. subtilis*	1X	ND	-
KANQ	*S. aureus*	1X	110nM	-
RKGK	*P. denitrificans*	?	?	?

To further study the role of Hfq in the model denitrifier *P. denitrificans*, Hfq was overexpressed using the same overexpression system as mentioned in 3.3.3. The hfq sequence was cloned into the overexpression vector pLMB509 at the first NdeI site and transformed into *P. denitrificans* by the use of triparental mating. The RNA chaperone was overexpressed from the taurine-inducible promoter in pLMB509 and an addition of 10 mM taurine was added to induce expression up to 15-20-fold in both aerobic and anaerobic batch cultures of *P. denitrificans*. The OD_{600} of the overexpression cultures as well as an empty vector control were collected at regular intervals. In aerobic conditions, an overexpression of Hfq through the addition of taurine resulted in a small reduction in growth compared to the empty vector control culture as well as the culture with the Hfq-containing vector pLMB509

without an addition of taurine. A maximum OD_{600} of ~1 was reached compared to ~1.2 reached by the empty vector control culture (Fig. 3.12A).

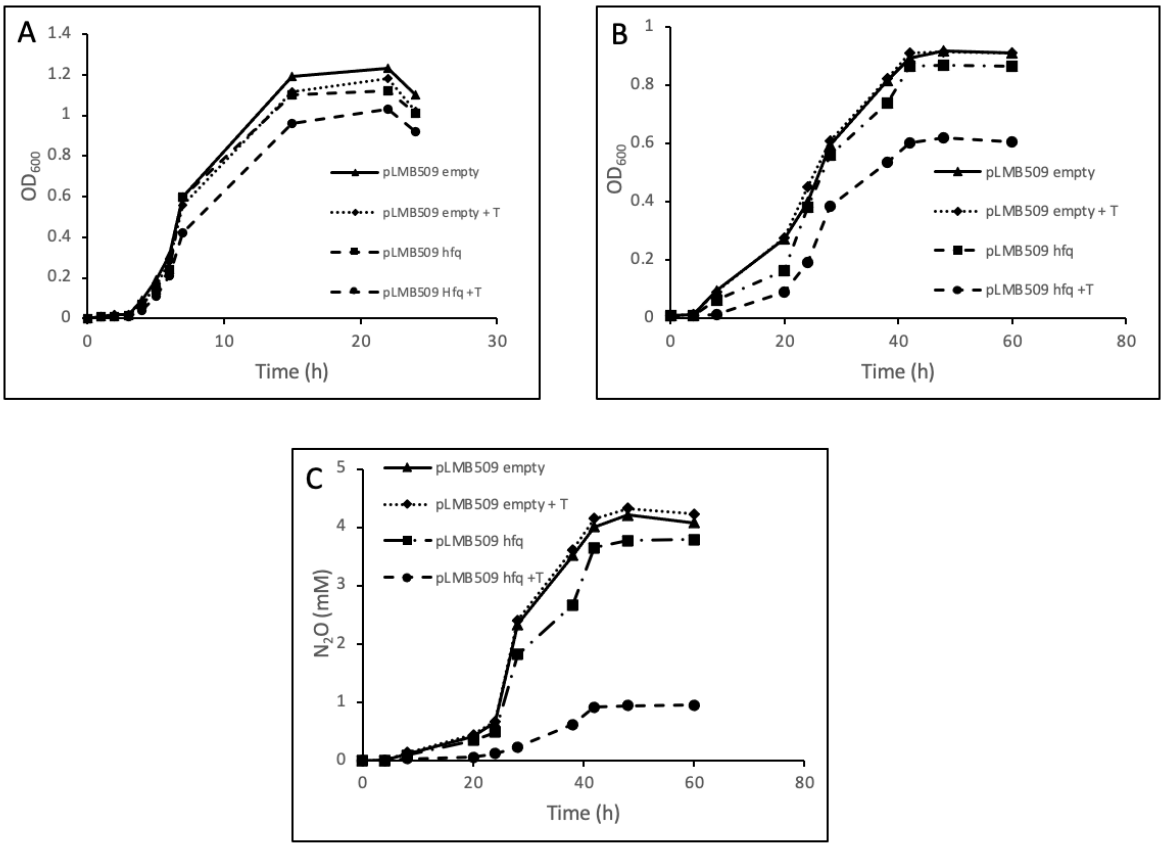


Fig. 3.12 Overexpression of Hfq from pLMB509 affects growth of *P. denitrificans* cultures in both aerobic, non-denitrifying (A) and anaerobic, denitrifying (B) conditions. **(A)** Cultures of *P. denitrificans* plus empty pLMB509 vector (squares) as well as cultures containing pLMB509 plus sRNAs were grown under aerobic, non- denitrifying conditions and **(B)** denitrification conditions with 10mM added taurine. The OD_{600} was measured at regular intervals **(C)** N_2O levels were measured from the anaerobic, denitrifying cultures at regular intervals

In aerobic culture conditions there was an even clearer impact of Hfq overexpression on growth. The clear reduction in growth observed in the Hfq overexpression culture resulted in a maximum OD_{600} of ~0.6 compared to ~0.9 observed in the empty vector control and the culture with the Hfq-containing vector pLMB509 without an addition of taurine (Fig. 3.12 B). When monitoring the N_2O levels emitted from the denitrification cultures, it was clear that an overexpression of Hfq did not only impact on culture growth but also perturbed the

sequential reduction of NO_3^- to N_2 resulting in a decrease in N_2O emissions measured in the Hfq-overexpression culture. In the presence of 10mM taurine, approximately 1mM N_2O was produced by the Hfq overexpression culture compared to ~4.2 mM measured in the empty vector control culture. Where no taurine was added to the empty vector control or the Hfq-containing vector, no significant differences in N_2O emissions were observed.

3.4 Discussion

The ubiquity and importance of sRNAs across bacterial genomes is now widely accepted and what remains is the challenge of identifying and deciphering the roles of these regulatory molecules within the context of bacterial regulatory circuits. The large number of sRNAs in *P. denitrificans* makes it a difficult and time-consuming process to experimentally characterise the exact role of each sRNA in the denitrification regulatory network. Denitrification is an important factor in the microbially driven flux of N_2O to the atmosphere and to date a comprehensive understanding of the molecular mechanisms and environmental cues that underpin this process is still lacking. Uncovering the role of sRNAs within the important biogeochemical cycle comes with a number of challenges that have to be tackled to complete the picture of the complex regulatory circuits to enhance the understanding of their impact on the production of the harmful greenhouse gas N_2O.

Hypothesis-driven research is the bedrock of modern science and although often error-prone, computational approaches are a useful tool to generate hypotheses as a direction for experimental design. Therefore, several computational tools were used to initially characterise structure and targets of select sRNA to have an enhanced understanding of their potential function before further analysing them experimentally. Several sRNAs were selected based on two approaches – the first approach utilising expression patterns and predicted targets to select sRNA for further investigation and a second approach based on putative denitrification regulator sites present upstream of select sRNA sequences.

3.4.1 sRNA Target prediction gives insights into putative sRNA functions

To better understand the function and regulatory networks of the sRNAs in *P. denitrificans*, it is important to identify their targets. There are several bioinformatics methods for sRNA target prediction such as CopraRNA (Patrick et al., 2013) and TargetRNA2 (Kery et al., 2014). The most accurate method, CopraRNA requires sequence conservation of sRNA and mRNA in at least four bacterial species and must be run one sRNA at a time. Therefore, this method is

highly time-consuming and is unsuitable for species-specific sRNA, making this method impossible to use for the 167 sRNAs identified in *P. denitrificans*.

The novel bioinformatics tool named sRNARFTarget is the first machine learning-based method that is capable of predicting the probability of interaction between an sRNA-mRNA pair (Naskulwur and Pena-Castillo, 2021). Machine learning (ML) is a field of study that makes it possible for computers to learn without being explicitly programmed (Samuel, 1959). Methods based on ML use data to build models, discover statistically significant patterns and relationships, and consequently make predictions on novel data (Bishop, 2006). The tool is generated using a random forest trained on the trinucleotide frequency difference of sRNA-mRNA pairs (Breimann, 2001). Therefore, it does not require sequence conservation of either sRNA or mRNA and can be applied to any sRNA-mRNA pair making it a suitable method for transcriptome-wide sRNA target prediction in *P. denitrificans*. Another advantage of sRNARFTarget is its speed. The tool is 100 times faster than other non-comparative genomics programs.

The results of the ML-based tool clearly grouped the sRNAs into specific sRNAs with a small number of highly predicted targets and non-specific sRNAs with a large number of highly predicted targets. The presence of both specific and broad target range sRNAs is common in bacteria with well-characterised sRNA-landscapes including *E. coli* and *Salmonella*. The broad-range sRNA, RNAIII present in S. aureus is able to modulate a global response by regulating gene expression directly via RNA-RNA interactions as well as indirectly controlling transcript levels through other regulators, therefore exhibiting a major impact on virulence (Jakobsen et al., 2017). Common targets of the sRNAs in *P. denitrificans* included transcriptional regulators such as GntR, Fis and TetR family regulators as well as transport proteins and a large number of hypothetical proteins or proteins of unknown functions and based on the presence of many broad-range sRNAs some may also be able to act globally via several modes of interaction. As shown in the 2016 study by Gaimster et al., transcriptional regulator targets are promising targets to expand the knowledge around the denitrification regulatory network as they provide a further step towards linking sRNA action with downstream regulatory processes affecting the biochemical pathways in a microbe making them important players in larger regulatory networks. Indeed, sRNAs such as *E. coli* Spot 42 RNA can be parts of

multioutput loops directly and indirectly acting on the expression of important genes (Baekkedal et al., 2015).

The second approach, TargetRNA 2 uses a variety of features to identify targets of sRNA action (Kery et a., 2014). These include sRNA sequence conservation, sRNA accessibility, mRNA accessibility and the energy of sRNA-mRNA hybridization. This second approach gave an additional insight into the most likely targets for each sRNA and similar to the results obtained through sRNARFTarget, many of these also included transcriptional regulators, transport proteins or proteins involved in DNA and RNA metabolism, such as the RNA polymerase sigma 32 subunit RpoH predicted to be targeted by sRNA 18. One of the first base-pairing sRNA targets to be characterized was the *rpoS* gene encoding for the stationary phase sigma factor sigma S in *E. coli* (Gottesman and Storz, 2010). The gene is positively regulated by sRNAs DsrA and RprA and negatively regulated by OxyS which shows the importance of sRNA regulation also regarding sigma factor expression. Sigma factors are of major importance in the initiation of transcription during different environmental conditions as they enable RNA polymerase binding to gene promoters. *P. denitrificans* possesses nine sigma factor homologues. Their exact roles are yet to be characterised and identifying sRNAs with a direct involvement in sigma factor regulation may give further insights into the sigma factors active during denitrification conditions, making the predicted targets of sRNA 18 an interesting finding.

Transporters such as metal and ABC transporters are also commonly predicted targets. The expression of membrane proteins such as transporters and porins have commonly been shown to be controlled by sRNAs in *E. coli* and *Salmonella* (Gottesman and Storz, 2010). Why these are such predominant targets is still not completely clear but due to the abundance of identified sRNA examples, there must be a regulatory advantage to this mode of regulation. Metal transporters are of importance for the denitrification process, as metal ions are used as electron donors, cofactors and components of the catalytic subunits of important enzymes (Tavares et al., 2006). Regulation of metal transport and cellular metal availability therefore greatly affects the rates of denitrification, making sRNAs with metal transporter targets of high importance. The predicted targets are an indication of what role an sRNA may be taking over and where in the denitrification regulatory network they may act, however, confirming

these predicted targets requires an experimental approach, to validate the computational findings. Despite many advances in computational platforms the need for experimental validation remains.

3.4.2 sRNA 39 overexpression causes an increased rate of N_2O emissions

Based on target predictions, structural folding predictions and existing knowledge of sRNA expression patterns during denitrifying and non-denitrifying conditions, three sRNAs were selected for experimental screening. While two sRNAs, sRNAs 5 and 10 were shown to have no impact on growth or N_2O emissions, an overexpression of sRNA 39 resulted in an enhanced rate of denitrification compared to the control cultures. The increased N_2O flux observed in the overexpression cultures combined with the results of the target prediction tool sRNARFTarget suggest that sRNA 39 may be capable of directly negatively regulating the expression of *nosZ* resulting in a decreased activity of the final denitrification reductase and therefore stalling denitrification at the level of N_2O reduction which would make it the first sRNA known to directly interact with a denitrification enzyme (Fig. 3.5 + 3.7). To confirm an interaction of sRNA 39 with *nosZ* mRNA several further experimental steps will have to be carried out in the future, including qPCR looking at levels of NosZ expression in the context of sRNA 39 overexpression to confirm sRNA 39 induced changes in expression as well as DNA-RNA binding studies to confirm interaction of the sRNA-mRNA pair and to give insights into the involvement of Hfq.

3.4.3 FNR-regulated sRNAs may be important links in the denitrification regulatory network

Across most bacterial species sRNAs have few, if any, transcriptional regulators identified. There are few exceptions such as *E. coli* sRNAs MicF and GadF which have 8 and 10 reported regulators respectively, but for most other sRNAs their regulators remain unknown (Keseler

et al., 2011). Therefore, a second approach to identifying sRNAs with an important function within the regulation of the denitrification network was to search for regulator motifs of known denitrification regulators upstream of all sRNA sequences. Roughly 20% of documented interactions regulating sRNAs involve sigma factors that respond to a multitude of stresses, but due to the lack of knowledge around sigma factors in *P. denitrificans*, the consensus search was limited to the known sites of FnrR, NarR or NNR (Bossi, 2016). Indeed, there were eight sRNAs that contained putative regulator binding sites upstream to their start codons suggesting these may be regulated be one or several of the known denitrification regulators and therefore may be taking over important roles in regulating the denitrification process.

Out of the three sRNAs selected for further experimental characterisation, sRNA 18 as well as sRNA 36 impacted on N_2O emissions when overexpressed in denitrification cultures. While sRNA 36 caused an enhanced rate of N_2O emissions similar to the rate observed in sRNA 39-overexpression cultures (Fig. 3.7), sRNA 18 resulted in a significant reduction of N_2O emissions. One of the putative targets predicted for sRNA 36 is a LysR transcriptional regulator (Fig. 5). LysR regulators are abundant regulators in prokaryotic organisms, that control single or operonic gene expression as activators and repressors (Minezaki et al., 2005). The functions of LysR type regulators in *P. denitrificans* remain unclear, and a confirmation of the involvement in the denitrification pathway would give further insight into the role of this important regulator group in the model denitrifier. Many previous sRNAs have shown to cause global regulatory responses through the regulation of a transcriptional regulator, as has been suggested for the sRNA DenR (Gaimster et al., 2016). Here it would be of interest to further examine the expression levels of this regulator protein to compare expression levels with those of the sRNA during high and low N_2O emitting conditions to identify an overlap and a potential mode of sRNA action. Furthermore, the identification of a seed region of interaction would provide insights into an interaction between sRNA and mRNA.

The second sRNA that caused a change in N_2O emissions, sRNA 18, is predicted to target a TraD family protein, the RNA polymerase sigma 32 subunit RpoH as well as a number of hypothetical proteins and proteins of unknown functions. 3.4.4 Proteins are classed as hypothetical, if they are predicted to be expressed from an open reading frame, but there is

no experimental evidence of their translation. Across many genomes a large number of genes remain whose function within the organism is still functionally unknown. Despite the lack of functional characterisation, these proteins are of great importance as many of them might be associated with important biochemical and physiological processes and uncovering their function may give a broader understanding of the whole functional landscape of an organism. Especially important biochemical processes such as the denitrification process would benefit from uncovering the roles of proteins with unknown functions to fill in the knowledge gaps and aid in better understanding the switches between N_2O emission and N_2O reduction to N_2. Therefore, a further investigation of the potential targets of sRNA 18 is required to confirm where in the denitrification cascade sRNA 18 influences levels of N_2O emissions. Measurements of the denitrification intermediates accumulating within the cultures would give further insights into whether the denitrification cascade is stalled at previous stages before N_2O production, or whether N_2 production is enhanced.

As both sRNA 18 and sRNA 36 showed significant changes in the levels of N_2O emitted from the cultures, future work should focus on elucidating their targets and their roles. As seen during the overexpression of DenR, the novel sRNA 18 reduces the levels of N_2O emitted from the cultures when overexpressed, making it an interesting target for the knock-down of N_2O emissions. An interesting starting point could be the putative sigma factor gene target, as sigma factors have been established as common sRNA targets with important regulatory function in model organisms such as *E. coli* and *Salmonella* (Gottesman and Storz, 2010). This may also be a starting point for the characterisation of the nine sigma factor homologues present in *P. denitrificans*, to gain further insights into the activity of sigma factors during denitrification. Furthermore, sRNA 36 is the first sRNA in *P. denitrificans* to have been shown to increase the rate of denitrification. Identifying the mode of action and potentially uncovering a further denitrification regulator positively regulating N_2O emissions would strengthen the knowledge around the slowly growing regulatory picture. This work should also continue via qPCR experiments to identify which of the denitrification enzymes are specifically affected by changing expression levels of the sRNAs, continuing with measurements of denitrification intermediates and an RNA-seq experiment to identify, which genes are differentially expressed in response to sRNA overexpression. This would provide insights into sRNA targets and aid in the further integration into the regulatory network.

3.4.4 Hfq may have a crucial role in sRNA action in *P. denitrificans*

Despite the widely accepted role of Hfq as an 'RNA chaperone' in several bacterial species, the detailed mechanisms by which it promotes sRNA-mRNA interactions remains ambiguous. While the focus has long been on E. coli Hfq, studies of the RNA chaperone in a range of bacterial species will help uncover its full role in posttranscriptional regulation. High resolution structures of Hfq bound to RNA oligomers combined with in vitro binding assays with Hfq mutants have revealed the presence of several RNA binding sites on both the proximal and distal faces of the hexameric Hfq ring. Nearly all known Hfq sequences possess a conserved patch on the outer rim of the hexamer which contains an arginine at position 16. Position 17 is usually R or K, followed by a neutral polar or acidic amino acid. This is also the case for the predicted Hfq sequence, Pden_4124, in *P. denitrificans* (Table 3.3) which is comparable to the sequence of Hfq in the gram-positive food-borne pathogen *L. monocytogenes* which causes serious infections in humans (Christiansen et al., 2004). Mutants of *L. monocytogenes* without Hfq were more sensitive to environmental stresses including salt and ethanol stress as well as during entry into the stationary phase suggesting that it contributes to stress tolerance and pathogenesis through interactions with sRNA molecules (Christiansen et al., 2004).

Based on structure alone, the predicted Hfq molecule in *P. denitrificans* seems likely to have an involvement in the sRNA regulatory network and may well be a requirement for correct sRNA function. Indeed, an overexpression of Hfq in *P. denitrificans* cultures impacted on growth in both aerobic and anaerobic culture conditions as well as nitrous oxide emissions from denitrifying cultures (Fig. 3.12). There have previously been reports that Hfq is closely related to cell growth, due to its association with stress resistance and cell survival under nutrient limitation (Chao & Vogel, 2010). An increased level of Hfq expression may promote cell growth by enhancing stress-resistance in some organisms such as in *Pseudomonas fluorescens*, while in others it may disrupt cellular physiology by alteration of protein expression resulting in retardation of cell growth (Vo et al., 2021; Wu et al., 2021). In *P.*

aeruginosa, hfq overexpression led to a decrease in cell motility as well as a dysregulation of the small RNA *phrS*, which is known to interact with Hfq (Fernandez et al., 2015). In the case of *P. denitrificans* an overexpression of Hfq from a taurine-inducible promoter resulted in attenuated growth under both aerobic and anaerobic conditions suggesting a significant alteration of sRNA-regulated protein expression impacting on cell growth and therefore supporting the previously unconfirmed role of Hfq in RNA metabolism in *P. denitrificans*.

Additional to the impact on growth, an overexpression of Hfq clearly affected levels of N_2O emitted from the denitrifying cultures. Levels of N_2O were approximately 4-fold lower in the Hfq overexpression cultures which could also be attributed to the changes in sRNA-regulated protein expression caused by increased availability of the RNA chaperone. As it has been shown that several sRNAs are capable of affecting N_2O emissions, it is plausible, that some of these sRNAs require the presence of Hfq to bind to their target as has been the case for many characterised sRNAs in model organisms such as *E. coli* and *Salmonella*. Increased availability of Hfq may therefore result in a dysregulation of action of certain sRNAs and therefore result in an alteration of the levels of N_2O emissions from the cultures. To gain further insights into the importance of Hfq, the generation of an Hfq knockout mutant would be beneficial, to compare the phenotypic effects of overexpression and deletion and more importantly, to assess the impact of sRNA overexpression in the absence of Hfq. Additionally, binding assays of Hfq and select sRNAs could give the confirmation of interaction required to validate the importance of Hfq in sRNA-induced regulation in *P. denitrificans*.

3.4.5 Conclusion

The identification of further sRNAs with an involvement in the denitrification pathway demonstrates the significant gaps in the understanding of this process. These gaps need to be filled to help inform future development of N_2O mitigation. We show that advances in sRNA-target predictions tools allow for initial sRNA screening and aid in the identification of putative targets for experimental screening. Furthermore, the presence of known regulator

consensus sites upstream of an sRNA can give insights into sRNAs potentially involved in a regulatory response. Using these two approaches three sRNAs were identified that alter N_2O emissions. Once their exact roles and functions within the complex denitrification network are uncovered, these and other sRNAs could possibly be utilised to specifically target and knock down N_2O emissions from *P. denitrificans*. In this study the RNA chaperone is shown to be of importance in *P. denitrificans* and its exact role in sRNA action will have to be elucidated to give an overview which sRNAs require its presence, and to determine how it fits into the existing regulatory network.

Clearly, this has been an understudied area to date and this, as well as further research, will be required to fill in the existing knowledge gaps. Future work should further combine studies within model organisms such as *P. denitrificans* with meta-analyses of sRNA in the environment.

4. NirR - a novel regulator of denitrification

4.1 Introduction

Small RNAs (sRNAs) are now considered important components of bacterial regulatory networks and recent work in our laboratory has confirmed the involvement of one sRNA, DenR (**Den**itrification **R**epressor - sRNA 29), in the regulation of nitrous oxide levels from *P. denitrificans* denitrification cultures. This was, to our knowledge, the first example of an sRNA directly regulating denitrification (Gaimster et al., 2016). DenR is expressed most highly under aerobic conditions and lowest when complete denitrification is occurring. It is encoded on the positive strand of *P. denitrificans* chromosome 1 and is transcribed in the opposite direction of a nucleoside ABC transporter membrane protein (Pden_0526). An overexpression of DenR resulted in a significant reduction in the amount of N_2O emitted from the cultures, compared to the empty vector control. Mutation of either a 3-bp region or a 6-bp region between bases 72 and 80 of the sRNA resulted in disruption of DenR secondary structure and removed its impact on denitrification confirming the direct effect of the sRNA on denitrification.

Measurements of the denitrification intermediates indicated that denitrification was stalled at nitrite reduction (Gaimster et al., 2016). The overexpression of DenR resulted in a change in expression of a total of 53 genes. Expression levels of nitrite reductase gene *nirS* (Pden_2487), the *norB* (Pden_2483) and *norC* (Pden_2484) genes encoding catalytic subunits of nitric oxide reductase, as well as *nosZ* (Pden_4219) encoding nitrous oxide reductase were downregulated as a result of DenR overexpression. Among the other gene products regulated by DenR there were proteins involved in energy metabolism, transport proteins as well as proteins involved in carbohydrate metabolism. Most interestingly, the expression levels of four transcriptional regulators were altered in response to DenR. A TetR family regulator (Pden_0668) and a TraR/DksA family regulator (Pden_0916) were downregulated while a GntR-type regulator (Pden_2475) and a LysR family regulator (Pden_4369) were upregulated. All four of these regulators were previously uncharacterised in *P. denitrificans*, or with regard to denitrification.

One-component systems are the most common prokaryotic signal transduction mechanism. They consist of a single polypeptide that contains both a sensory domain and a DNA-binding

domain. Among these, the TetR family of regulators are known to regulate numerous aspects of bacterial physiology and are capable of interacting with a vast array of ligands (Cuthbertson & Nodwell, 2013). The members of the TetR family consist of an N-terminal DNA-binding domain and a larger C-terminal domain which can interact with one or more small-molecule or protein ligands. The enormous diversity within this family has shown that the members can function as both activators and repressors while serving as either local or global regulators. TraR and DksA are transcription factors encoded in *E. coli* (Gopalkrishnan et al., 2017). They are highly conserved across major bacteriophage and bacterial groups suggesting that their regulatory functions are strongly selected for in evolution. In *E. coli* TraR acts as a global regulator inhibiting certain promoters while activating others.

Of the two types of regulators upregulated by DenR overexpression, LysR reglators are a well-characterised group that are highly conserved and ubiquitous amongst prokaryotes (Maddocks & Oyston, 2008). They have a conserved structure with an N-terminal DNA-binding helix-turn-helix (HTH) motif and a C-terminal co-inducer binding domain and regulate a diverse set of genes including those involved in virulence, metabolism, quorum sensing and motility. The GntR family of transcriptional regulators was first described in 1991 and named after the gluconate-operon repressor in *Bacillus subtilis* (Suvorova et al., 2015). Members of the family share a highly similar N-terminal HTH DNA-binding domain but differ in the C-terminal effector-binding and oligomerization domains. The C-terminal domain does not bind to DNA, but it can impose steric constraints on the DNA-binding domain affecting the HTH motif.

According to the type of the C-terminal domain, the members of the GntR family are divided into four main (FadR, HutC, MocR and YtrA) and two minor (AraR and PlmA) subfamilies. The FadR subfamily is the largest subfamily and comprises about 40% of the known GntR regulators. Members of this subfamily bind effectors, small organic ligands, such as carboxylic acids and subsequently undergo conformational changes that affect DNA binding (Zheng et al., 2009). The C-terminal domain of the second subfamily, HutC, has a length of around 170 amino acids and contains a combination of α-helical and β-sheet structures (Rigali et al., 2002). Its folding structure suggests that it may bind smaller effector molecules, such as histidine (HutC), fatty acids (FarR), sugars (TreR), and alkylphosphonates (PhnF) (Suvorova et

al., 2015). The third subfamily, MocR, differs from the other groups due to a large C-terminal domain with an average length of 350 amino acids while members of the final subfamily, YtrA, have a reduced C-terminal domain with only two α-helices and an average length of about 50 amino acids.

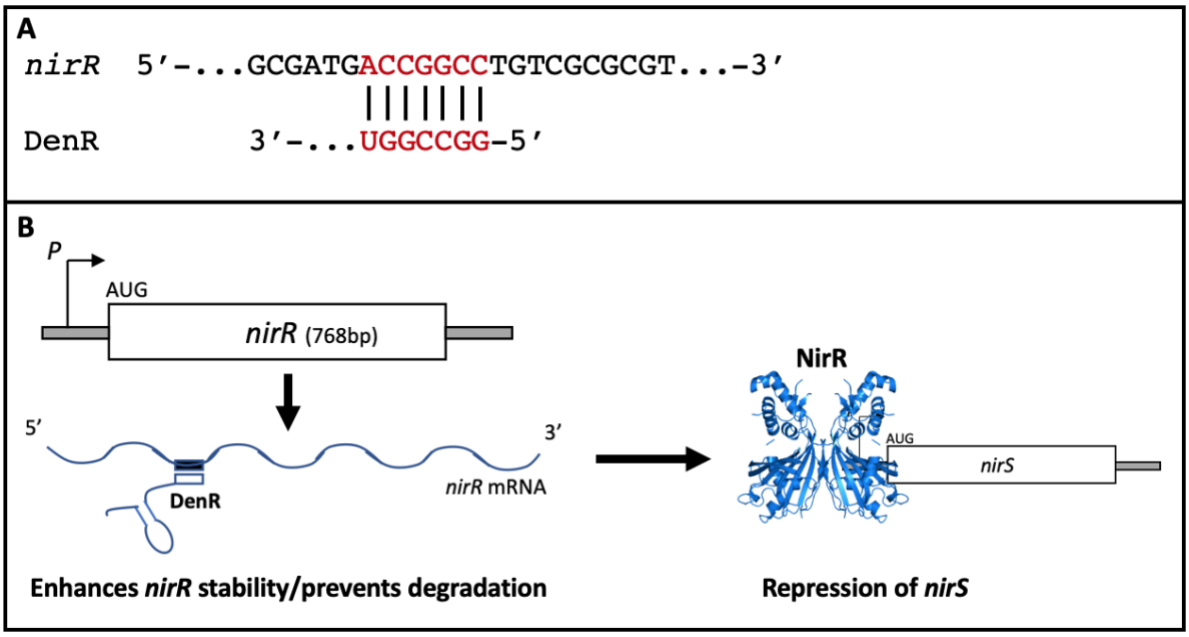


Fig. 4.1 Proposed mechanism of DenR-induced repression of *nirS* by the novel GntR-type regulator NirR. **A:** The 7-bp seed region of sRNA-mRNA interaction. **B:** Proposed mechanism of repression by which DenR interact with the seed region within *nirR* mRNA, enhancing *nirR* stability to facilitate repression of *nirS*

The GntR- type regulator in *P. denitrificans* displayed conserved expression patterns identical to those seen with DenR (Gaimster et al., 2016). Overexpression of the regulator phenocopied DenR overexpression and resulted in nitrite accumulation and a reduction in N_2O emissions. Furthermore, the regulator shares a 7-bp region of sequence homology with DenR which is located within the CDS of *gntR* (Fig 4.1A). The study proposed a mechanism by which sRNA-29 (DenR) stabilises *gntR* mRNA, possibly by blocking the activity of RNase-mediated decay. In turn, the GntR-type regulator represses the expression of nitrite reductase resulting in the stalling of denitrification at the stage of nitrite reduction observed in the study (Gaimster et al., 2016). Therefore, the GntR-type regulator was given the name NirR (nitrite reductase repressor).

4.2 Aims

This study aims to gain further insights into the genes targeted by DenR. Furthermore, it aims to evaluate the interaction of DenR with *gntR* mRNA in more detail through the purification and experimental characterisation of NirR and explores the detailed molecular interactions of the proposed regulatory model.

Chapter Hypotheses:

- DenR and NirR interact via a 7bp seed region
- NirR is a novel regulator of denitrification
- NirR can be purified for downstream binding assays between DenR and NirR

4.3 Results

4.3.1 Eight of the 53 DenR differentially regulated genes contain the 7bp seed recognition site

In order to gain further insights into the 53 genes differentially regulated as a result of DenR overexpression, sequences of all 53 genes were screened for the presence of the 7-bp seed region predicted as region of interaction between NirR and DenR. Consequentially, seven further genes were confirmed to possess the sequence (Fig. 4.2). Of the seven genes, three genes (Pden_1996, Pden_3414 and Pden_4058) encoded for hypothetical proteins or proteins of unknown function. Two of the remaining genes encoded for transport proteins; a sulfate ABC transporter (Pden_1798) and a cation transporter (Pden_4058). Pden_2495 encodes for a cytochrome d1 and Pden_2487 encodes for nitrite reductase, suggesting that DenR may also be capable of directly regulating levels of NirS expression.

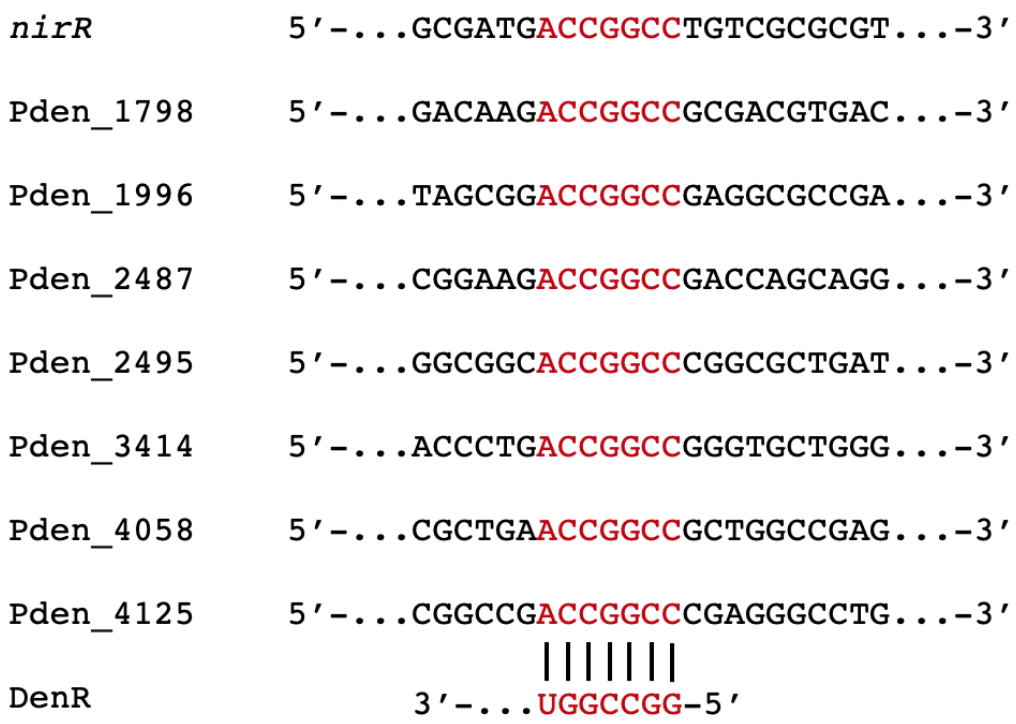


Fig. 4.2: Computationally predicted genes possessing the 7-bp seed region of interaction identified for DenR and NirR in their promoter regions

4.3.2 Based on the structure of the C-terminal domain, NirR is classified as HutC subfamily regulator

To date, NirR remains structurally and functionally uncharacterised in *P. denitrificans* and therefore several computational tools were used to gain further insights into folding structure and ligand binding capabilities. The NirR protein is 255 amino acids long and has a mass of 27,464 Da. Phyre2 is a web-based tool for prediction and analysis of protein structure and function (Kelley et al., 2015). The *nirR* sequence was submitted to Phyre2 to generate a prediction for both secondary and tertiary structures of the regulator protein (Fig. 4.3 and 4.4). The secondary structure of a protein is determined by the pattern of hydrogen bonding and NirR contains both α-helical and β-sheet structures which were predicted with high confidence levels (Fig. 4.3). The C-Terminal domain contains a combination of both α-helices and β-sheets while the N-terminal domain is dominated by α-helices.

The entirety of the NirR prediction contains regions of disorder, which are often regions of functional importance that make prediction difficult. The predictions scores for secondary structure as well as disorder prediction is lower than average which may be due to a lack of sequence homologues detected in the Phyre2 database.

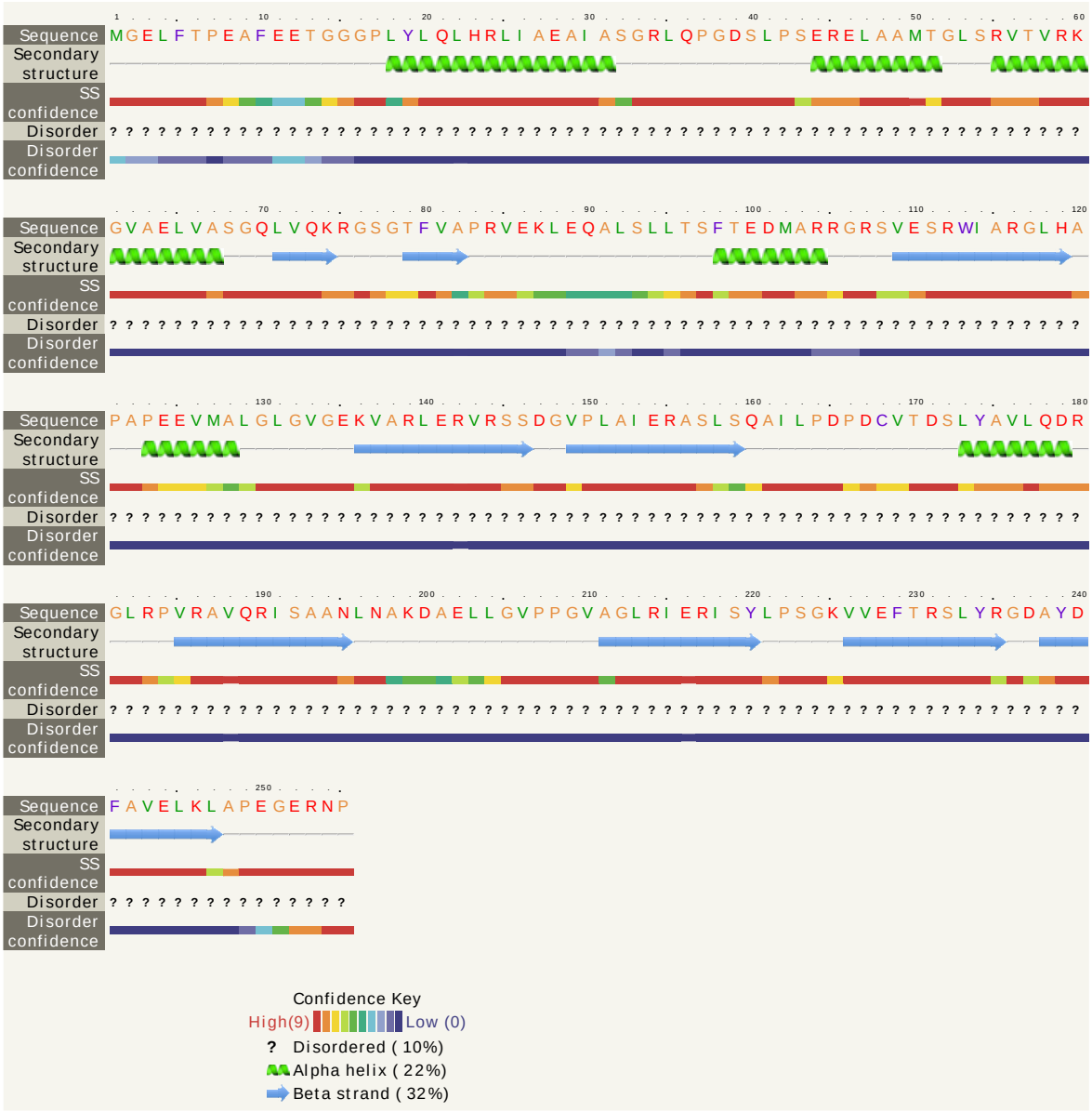


Fig. 4.3 Secondary structure of NirR predicted using Phyre2 (Kelley et al., 2015). Predicted alpha helices are shown in green, while beta sheets are marked as blue arrows. The SS confidence line indicates the confidence in the prediction, with red being high confidence and blue low confidence. The Disorder line contains the prediction of disordered regions in the NirR protein. These are often regions of high functional importance.

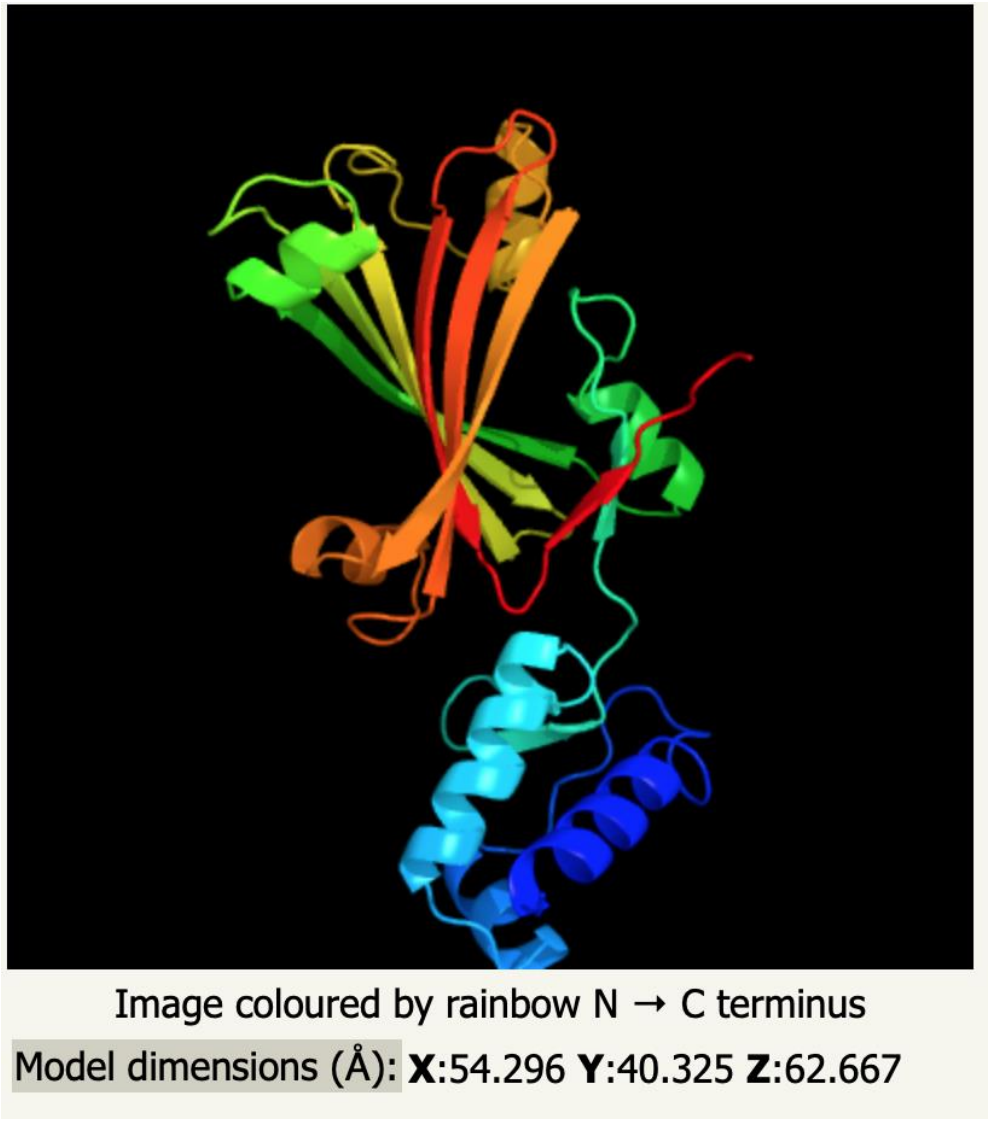


Fig. 4.4 Tertiary structure of NirR predicted using Phyre2 and based on the hth-type transcriptional repressor dasr, which was selected as highest scoring template. Of the NirR sequence 225 residues (89% of the sequence) have been modelled with 100% confidence.

Having predicted the secondary and tertiary folding structures, it was possible to determine the subfamily of GntR type regulator NirR belongs to as the subfamilies are distinguished based on the structure of their C-terminal domains. The length of the C-terminal domain, as well as the combination of α-helices and β-sheets suggested, that NirR was part of the HutC-subfamily. An alignment of the C-Terminal domain of NirR with the sequences of known HutC-subfamily regulators collected by Rigali et al, 2011, confirmed the presence of six conserved alpha helices and seven beta sheet structures which therefore confirmed the classification of GntR as a HutC-subfamily regulator (Fig. 4.5). This potentially gives further insight into its

mechanism of action and can aid in the further characterisation of this novel regulatory protein.

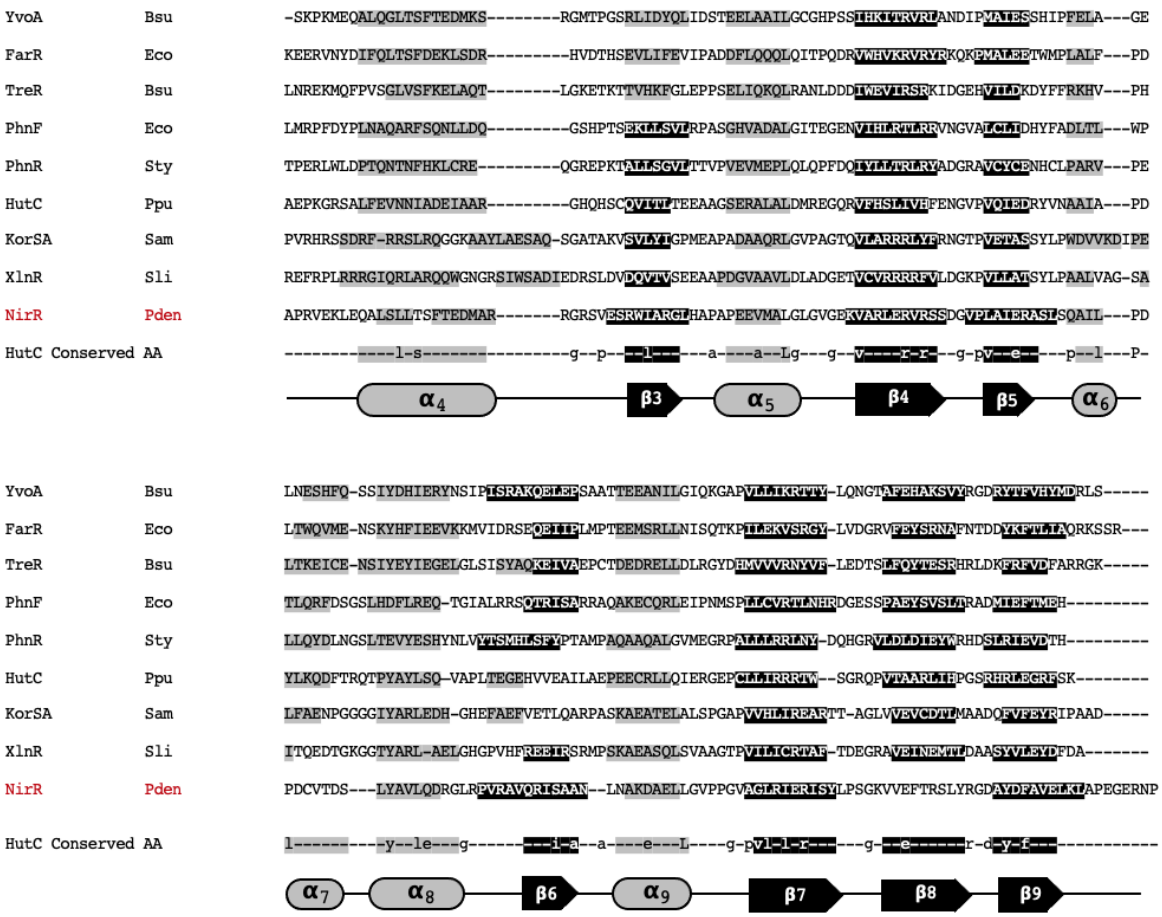


Fig. 4.5 Structure-based sequence alignment of the C-terminal domains of proteins of the Hut-C family of GntR-type regulators (adapted from Rigali et al., 2001). Alpha helices are marked in grey and beta sheets are showm in black. Bacterial abbreviations as followed: Bsu = *Bacillus subtilis*, Eco = *Escherichia coli*, Sty = *Salmonella* Typhimurium, Ppu = *Pseudomonas putida*, Sam = *Streptomyces ambofaciens*, Sli = *Streptomyces lividans*, Pden = *Paracoccus denitrificans*.

4.3.3 Purification of *P. denitrificans* NirR for use in DNA binding assays

Test expression of NirR was carried out under a number of conditions with varying concentrations of the transcription-triggering reagent isopropyl β-D-1-thiogalactopyranoside (IPTG) from 0.001mM to 1mM to induce protein expression from the lac operon (Fig. 4.6). NirR protein was produced at increased levels when induced with 0.01mM, 0.1mM and 1mM. The highest levels of NirR protein expression (molecular weight: 27 kDa) however, were observed at induction with 0.1mM IPTG overnight at 30°C.

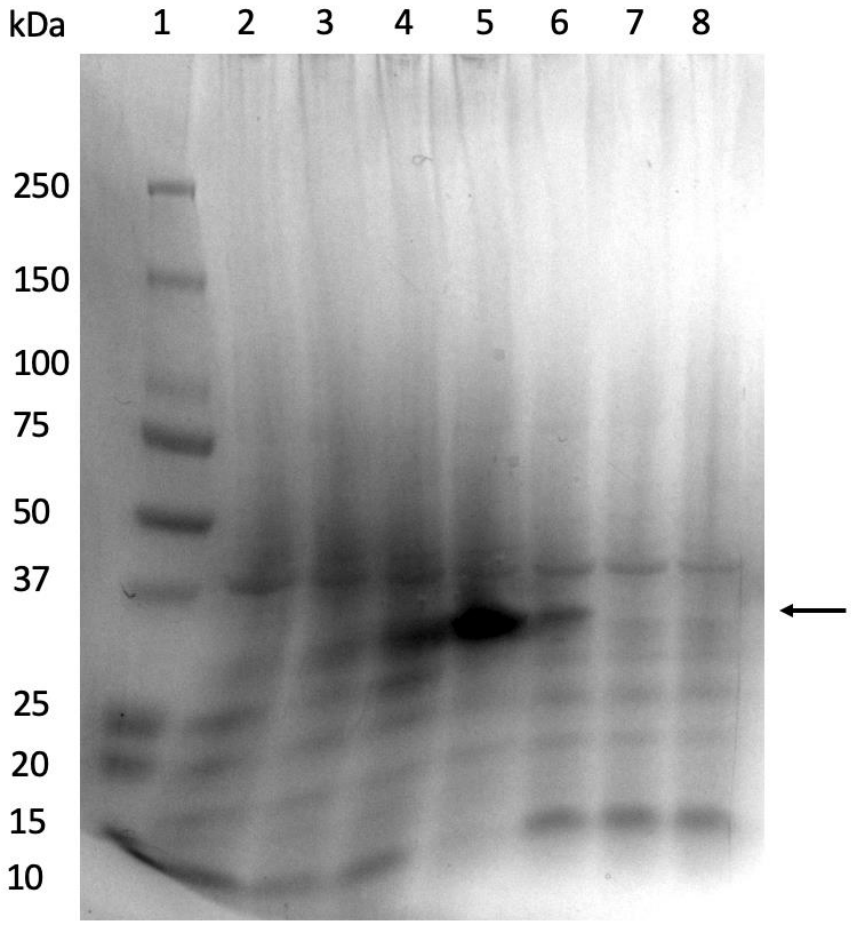


Fig. 4.6: Test Expression Lane 1: Precision Plus Protein Standards Ladder, Lane 2: BL21 pET14b GntR (pre induction), Lane 3: BL21 pET14b GntR (0.001mM IPTG), Lane 4: BL21 pET14b GntR (0.01 mM IPTG), Lane 5: BL21 pET14b GntR (0.1 mM IPTG), Lane 6: BL21 pET14b GntR (1 mM IPTG), Lane 7: BL21 Empty pET14b Control (0.1mM IPTG), Lane 8: BL21 Empty pET14b Control (1mM IPTG)

Having determined the ideal expression conditions, His-tagged NirR was purified following overexpression in *E. coli* BL21 (DE3). Following successful overexpression 10His-NirR was purified as described in 2.4.9. The elution profile showed the release of NirR protein in two stages from fractions 2.B.8 to 2.C.6 (Fig. 4.7). When using the cell pellet of 2L overexpression culture, there were two clear bands (Fig. 4.8). The first band corresponding to the molecular mass of 27kDA of 10His-NirR was clearly visible. It was accompanied by a second band at around 50kDa suggesting the presence of a NirR dimer at high concentrations of protein in the fractions. The fractions in lanes 9-17 were collected and buffer exchanged into a storage buffer.

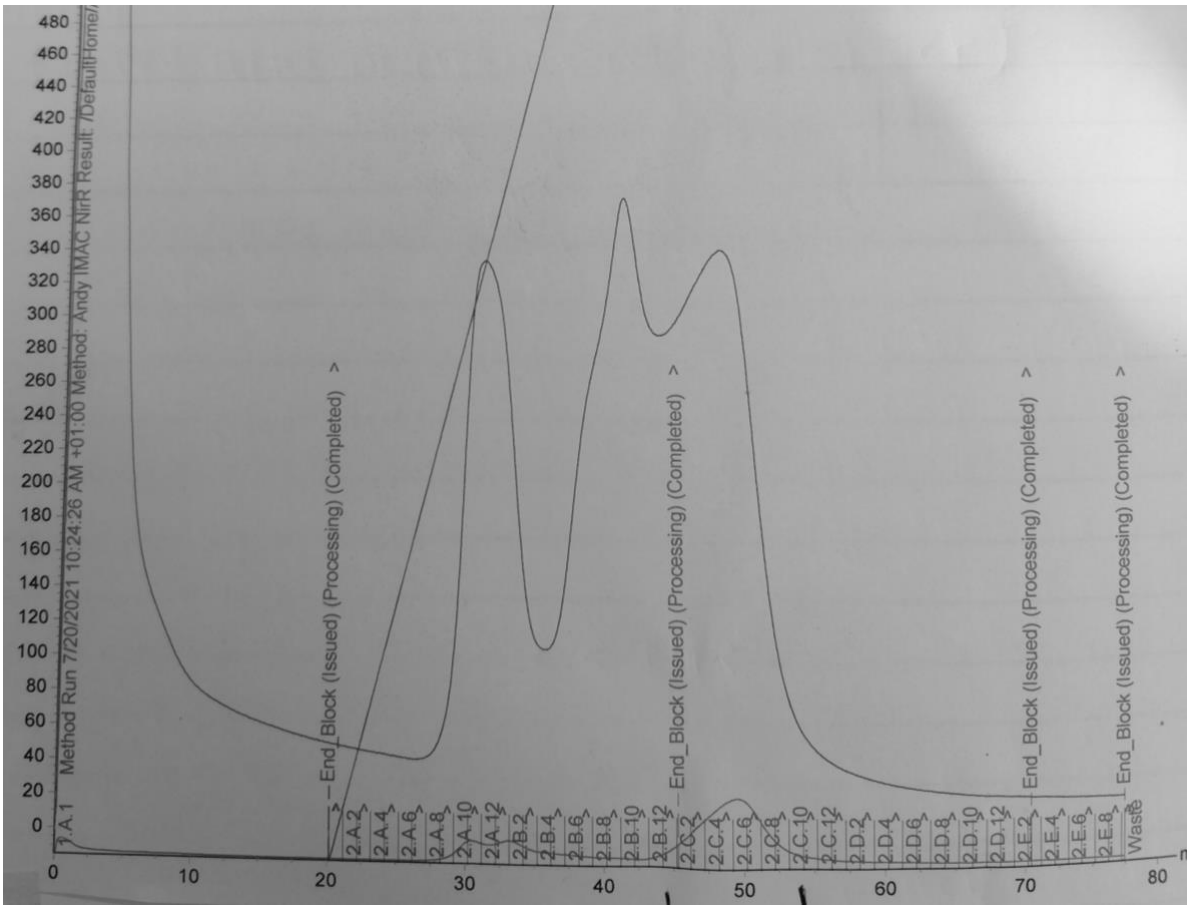


Fig. 4.7 The elution profile for NirR from the ÄKTA™ Chromatography System. Elution of NirR was measured at 280nm; the first peak shows the release of unbound material followed by the elution of bound protein with an increase of buffer B. Fractions 2B8 to 2C6 were collected.

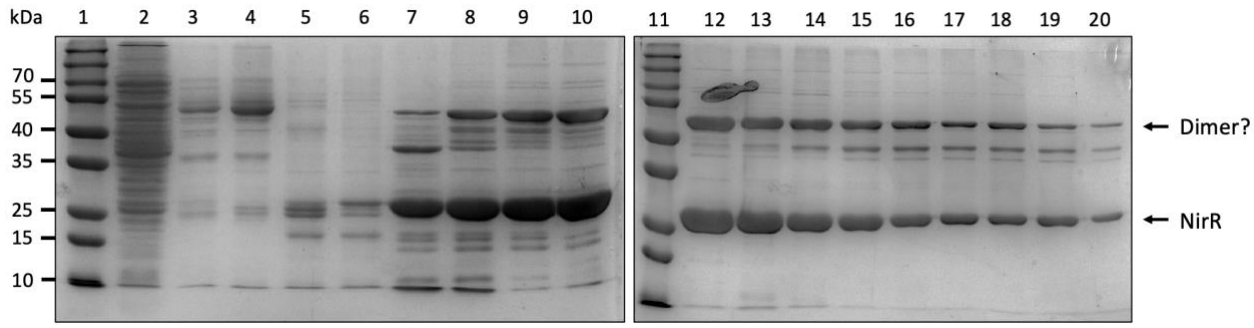


Fig. 4.8: Samples from 17 fractions collected (lanes 3-10 and 12-20) during the elution stage of the purification procedure were resolved by SDS-PAGE. Proteins were stained with InstantBlue™. Lane 1: PrecisionPlus Protein Dual Color Standard. Lane 2: Column Flowthrough

When reducing the amount of cell pellet used during protein purification, the protein appeared to be present mainly as a monomer, forming a clear band at 27kDa corresponding to the molecular weight of 10-His-NirR (Fig. 4.9).

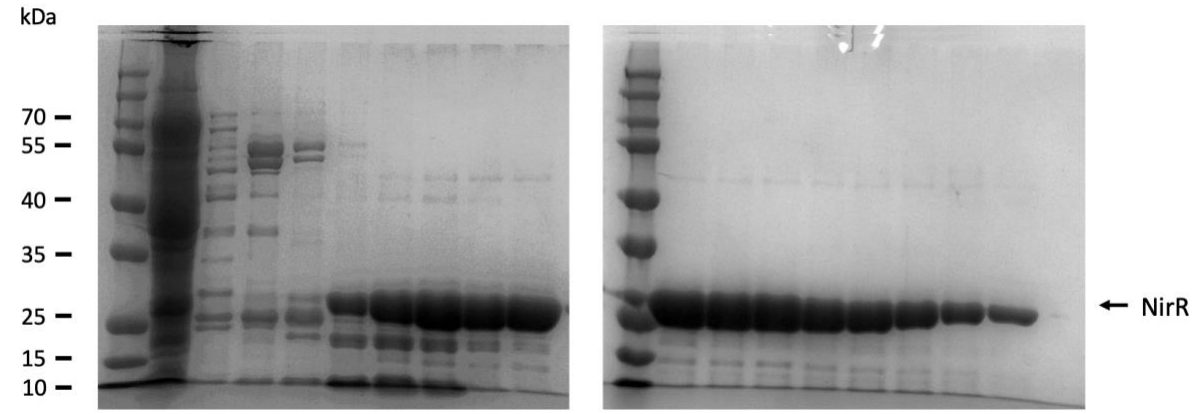


Fig. 4.9 Samples from 17 fractions collected (lanes 3-10 and 12-20) during the elution stage of the purification procedure were resolved by SDS-PAGE. Proteins were stained with InstantBlue™. Lane 1: PrecisionPlus Protein Dual Color Standard. Lane 2: Column Flowthrough

The protein concentration of the final elution was determined using a Bradford assay, a colorimetric protein assay based on an absorbance shift of the dye Coomassie brilliant blue G-250 (Bradford, 1976). Binding of protein molecules to the dye results in a colour change from brown to blue. The difference between the dyes is greatest at 595nm, making this the optimal wavelength to measure the colour of the Coomassie dye-protein complex. Based on

the standard curve generated using bovine serum albumin (BSA), the protein concentration in the elution fractions was determined to be 310 µg (Fig. 4.10).

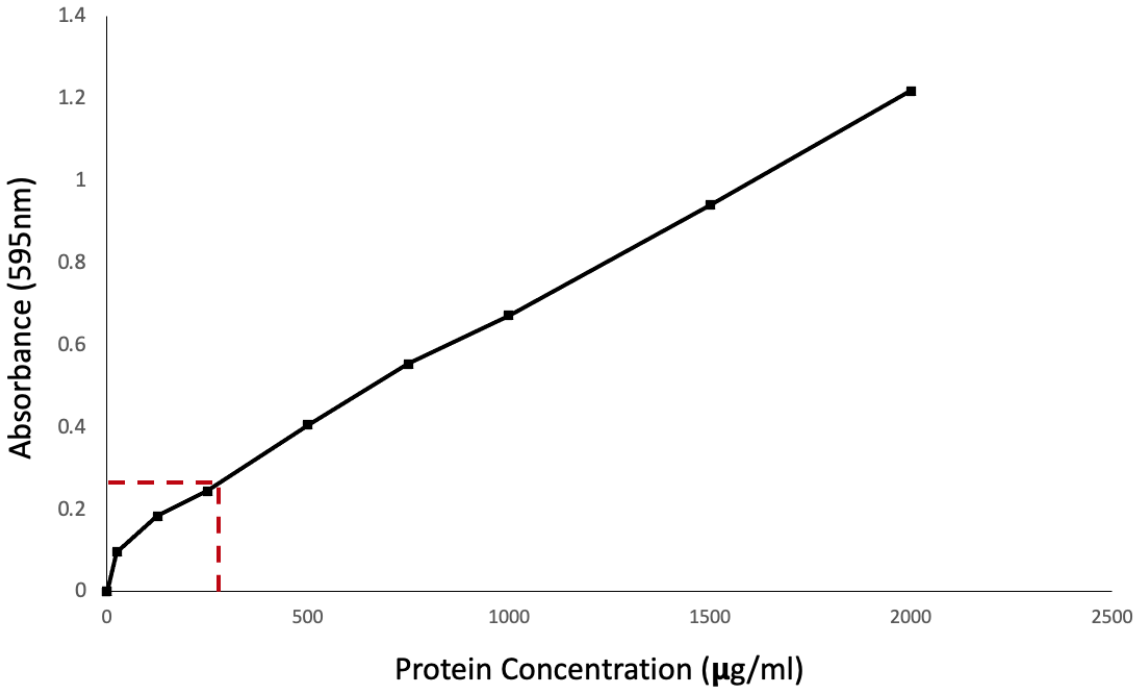


Fig. 4.10 Bradford assay for the determination of NirR protein concentration in the purification elution fractions. Absorbance was measured at 595 nm and NirR was present at a concentration of 310 µg.

4.3.4 Competitive binding of the denitrification regulators NNR and NirR

The consensus sequence of all known motifs of HutC-subfamily transcription factors resembles the motif identified for members of the FadR subfamily. To generate a putative HutC regulator binding motif, sequences from 1300+ known HutC recognition sequences collected by Suvorova et al., were submitted to MEME Suite, a motif-based sequence analysis tool. The putative motif contained a highly conserved G and T at positions 2 and 3 of the motif as well as a highly conserved A/G and C at positions 8 and 9 of the motif.

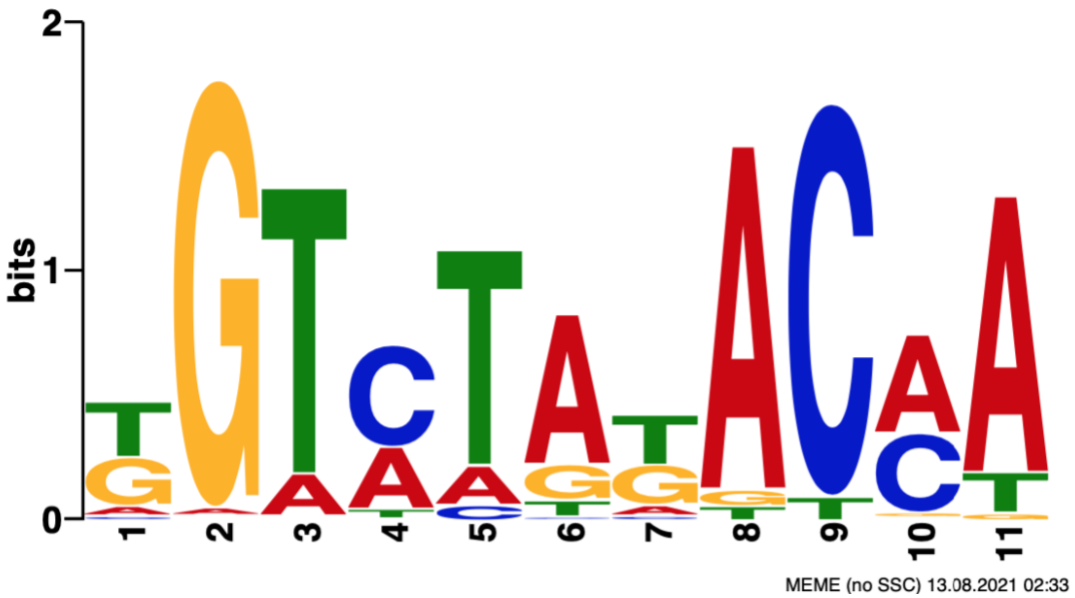


Fig. 4.11 HutC subfamily regulator recognition Motif Generated using MEME, with sequences from Suvorova et al. (1300+ sequences of recognition sequences of HutC subfamily regulators)

To identify the location of the NirR consensus sequence upstream of *nirS,* the *nirS* promoter sequence (up to 100 bp upstream of the start codon) was submitted to the web tool Motif Matcher. The computationally predicted putative NirR motif is located 30 nt upstream of the TSS and contains the conserved GT and GC that were identified using MEME Suite (Fig. 4.11). The predicted site overlapped with the known FNR-box recognised by the denitrification regulator NNR which is known to bind upstream of *nirS* to upregulate expression in response to decreasing oxygen levels (Fig. 4.12). The FNR-box is located at a position -41.5 relative to the transcription start sites of both the *nirI* and the *nirS* genes. An overlap in the binding sites as well as the counter-function of the two regulators suggests a potential competition for the binding upstream of *nirS*. Competitive binding affects expression levels of a gene and therefore a further investigation of the relation of both NirR and NNR could give further insights into the fine-tuning of *nirS* expression.

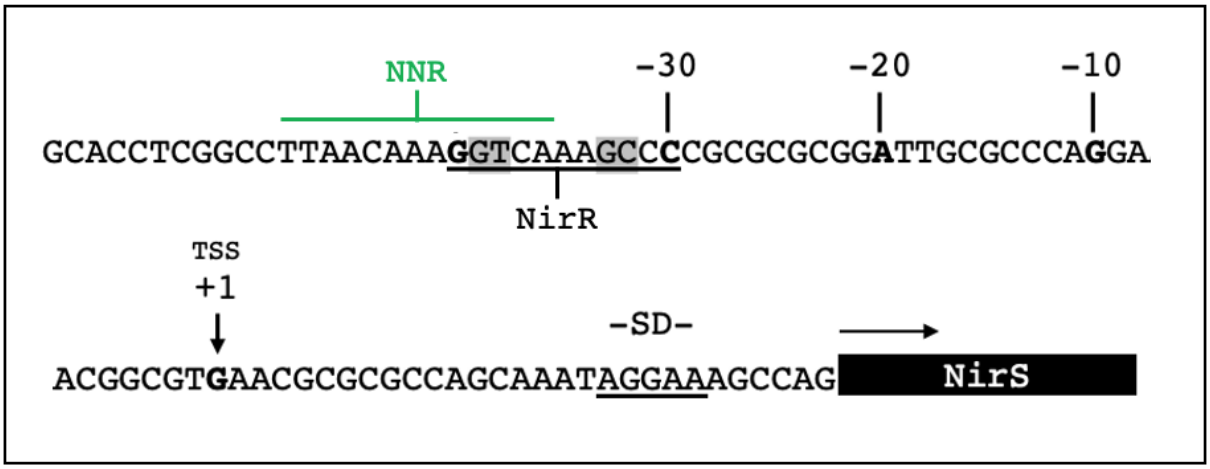


Fig. 4.12 DNA sequence of the region upstream of the *nirS* gene. Shown is the transcriptional start site (TSS) for the nirS gene, the suggested NNR and NirR binding sites

4.3.5 DNA Binding Assays

Surface Plasmon Resonance (SPR) is an optical technique that can be used without radio- or other labelling to measure molecular interactions in real time making it a useful tool for initial screening of NirR-DNA binding. The ligand-analyte pair is represented in response units (RZ) and R_{max} is the maximal feasible SPR signal generated by an interaction. Five DNA probes of the region upstream of *nirS* were utilised and purified NirR was added at a range of 1nM to 10uM. The reactions were carried out in triplicates. A low-level response of up to around 45% R_{max} was observed at the higher concentrations of 5uM and 10uM (Fig. 4.13). This low-level response, however, was likely not sequence specific. Although an R_{max} of 100% is rarely achieved, as this would require 100% of the probes to be occupied by a ligand, a specific response would be expected to have an R_{max} of above 60%, which was not the case for the NirR experiment. For the lower protein concentrations, no interaction of the protein and the DNA probes on the chip were detected. Further studies will be required to provide concrete evidence of NirR binding to the promoter sequence upstream of *nirS.*

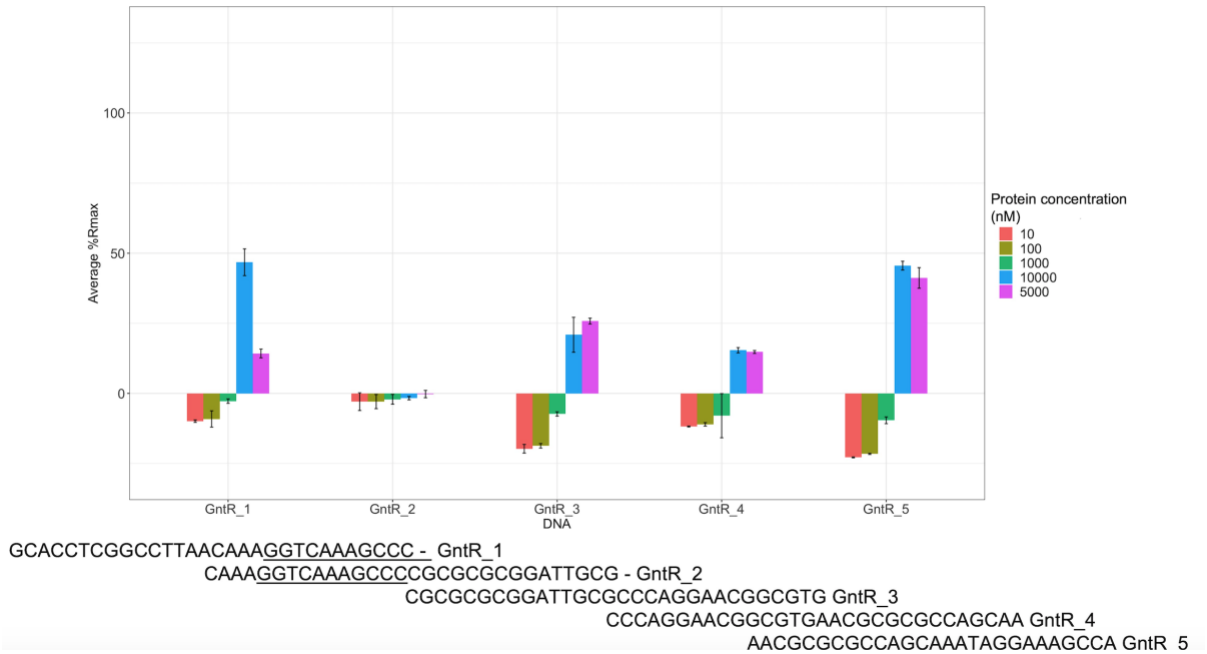


Fig. 4.13 Surface plasmon resonance to assay NirR binding to the *nirS* promoter region. Purified NirR was added at a range of 1nM to 10uM with each reaction carried out in triplicate. For all 5 promoter probes GntR_1 to GntR_5 low-level response units (R_{max}) were observed at the higher concentrations of 5uM and 10uM. Lower protein concentrations did not show a specific binding response.

4.4 Discussion

4.4.1 The DenR regulon

DenR is known to differentially regulate the expression of 53 genes including a novel GntR type regulator, named NirR due to its proposed function of repressing the expression of NirS (Gaimster et al., 2016). Comparable to other sRNA-target pairs, NirR and DenR share a 7-bp seed region of complementarity. Some bacteria have one region of interaction while others, such as Spot42 or FnrS, have multiple seed regions that pair with different mRNA targets (Durand & Storz, 2010; Beisel & Storz, 2011). The seed region was identified to be present in 7 further genes within *P. denitrificans* suggesting that these genes may also be directly regulated by DenR if the seed region is used for more than one target, potentially resulting in a global regulatory response. Surprisingly, *nirS* was among the identified genes suggesting that DenR may be able to directly target *nirS* expression, while also regulating its expression via the novel transcriptional regulator NirR, an interaction that had been overlooked in previous studies (Gaimster et al., 2016). This could further explain the significant impact of DenR overexpression on denitrification. Many well-characterised sRNAs across the bacterial kingdom are known to have more than one mRNA target and therefore it is likely that DenR is indeed capable of interacting with further transcripts. The region of interaction and the targeting of other genes will have to be explored further to confirm the interaction.

4.4.2 Classification of NirR

The GntR-family is known to regulate transcription through allostery on binding to metabolites (Haydon & Guest, 1991). Allostery is the process wherein binding of a ligand or effector molecule, such as a denitrification intermediate alters the activity of a protein at a distant site. This can alter the affinity of a regulator protein to DNA, resulting in it becoming a molecular switch. Ligand binding occurs at the C-terminal domain affecting oligomerisation of the transcription modulator. Many well characterised GntR regulators require the presence

of metal cofactors such as iron to efficiently bind DNA targets (Suvorova et al., 2015). In light of metals being of high importance during denitrification, it is not unlikely, that *P. denitrificans* NirR also requires the presence of a metal cofactor such as iron or copper. Having identified the importance of the previously uncharacterised GntR regulator NirR in regulating the denitrification machinery, this chapter demonstrates that NirR is part of the HutC regulator subfamily. The two-domain proteins with a smaller N-terminal HTH-domain and a larger C-terminal domain have been implicated in processes like antibiotic production and sensing of the nutritional status in the immediate environment as well as other important physiological responses (Rigali et al., 2002).

4.4.3 Regulation of *nirS* by NirR

Regulatory networks for activation and repression of gene expression lie at the heart of cellular computation. Extracellular signals such as changes in nutrient availabilities are integrated by transcription factors which in turn switch on and off regions of DNA. Regulatory regions are often highly complex and consist of both repetitive and overlapping transcription factor binding sites resulting in a cooperation as well as competition of transcriptional regulators with one another. Competition for regulation is still poorly understood as few studies have directly addressed the competitive interactions between transcription factors. The FNR-box upstream of *nirS* recognised by the NNR transcriptional activator dimer allows for an activation of nirS gene expression. An induced sequence change of the motif resulted in a complete inactivation of the promoter confirming the importance of this site for NNR action (Saunders et al., 2002).

The NirR recognition motif was predicted utilising known sequences of HutC family regulator binding sites by inputting these sequences into the motif finding algorithm Motif Matcher. The consensus sequence of all analysed binding sites of the HutC-subfamily is an A/T-rich sequence. The common distance between the conserved GT and AC is 4nt (Suvorova et al., 2015). A protein-DNA correlation analysis showed that positions significant for binding specificity of HutC transcription factors resemble the ones identified for FadR from *E. coli* and

the FadR subfamily in general. Amino acids in position 28 of the HTH domain correlate with nucleotides 8/17 (Suvorova et al., 2015). The most frequent amino acid in this position is Arg, as is the case for NirR. According to electrochemical characteristics it strongly prefers the interaction with the G/C pair within the motif, possibly interacting with the G.

The overlap in predicted binding sites for NirR and NNR upstream of *nirS* suggests a potential competition for binding to the promoter region and regulating the expression of the crucial denitrification enzyme. While NirR acts as a transcriptional repressor, decreasing the levels of *nirS* expression, NNR acts as a transcriptional activator, increasing the levels of *nirS* expression in response to lowered oxygen levels. Further insights into this competitive interaction would be given by overexpressing NirR in the presence as well as the absence of NNR and comparing expression of *nirS*.

4.4.4 Role of effector molecules in the binding of NirR upstream to *nirS*

GntR regulators are known to require effector molecules that interact with the effector binding domain at the C-terminus of the protein. This usually results in a conformational change and the dimerization of the regulator to allow for subsequent binding of a DNA operator sequence. Allostery is an important phenomenon that allows bacterial regulatory proteins to function as molecular switches turning on and off the expression of certain genes in the presence or absence of certain molecules. At high concentrations post purification, two distinct bands suggested the presence of a NirR dimer, a conformation that is seen in many other GntR regulators in their active form (Jain, 2015). Many known GntR regulators require the presence of effector molecules such as iron or molybdenum that are products or substrates for the metabolic pathway they regulate.

As protein-DNA binding assays of NirR and the promoter sequence upstream of *nirS* remained unspecific, it seems likely that NirR also requires the presence of metal ions, a specific effector molecule or that the protein was not present in its active form post purification. Especially in the context of denitrification, metal cofactors are a likely requirement for a denitrification-

specific regulator protein. Equally, other transcriptional regulators require the presence of denitrification intermediate to become transcriptionally active. For example, the RegSR proteins in *B. diazoefficiens* respond to the presence of nitrate and oxygen, which alter protein conformation and hence transcriptional activity (Elsen et al., 2004). Further DNA-protein interaction studies such as an expanded SPR experiment or an EMSA with an addition of denitrification intermediates or additional metal ions may provide further insights into the binding capability of NirR upstream to *nirS.*

4.4.5 Conclusion

To our knowledge, this was the first purification of NirR protein. The further characterisation of DenR and NirR has provided more insights into the interaction and has opened paths for future studies to determine the exact roles of the two regulators within the denitrification pathway. DenR is proposed to interact with NirR via a 7-bp seed region, that is present in 7 other genes potentially regulated by the sRNA. In turn, NirR stalls denitrification by repressing expression of the denitrification enzyme NirS which is responsible of nitrite reduction. Binding of the regulator upstream of the *nirS* operon may be dependent on the presence of effector molecules such as the denitrification intermediates. This should be the subject of future studies regarding the novel denitrification regulator. A full understanding of DenR and NirR would further expand the picture of the denitrification regulatory network and strengthen the use of *P. denitrificans* as a regulatory model of denitrification.

5. Genome-wide mapping of transcriptional start sites

5.1 Introduction

The prokaryotic transcriptome is highly dynamic and encompasses the complete set of RNA. It changes in response to varying environmental conditions, growth stages or developmental stages. Initially it was thought that transcriptomes were nearly entirely composed of rRNAs, tRNAs, several housekeeping RNAs and protein-coding mRNAs. Comparative transcriptome studies have shown that transcriptional maps may differ even among closely related species and in recent years it has become clear that both eukaryotes and prokaryotes contain large numbers of additional non-coding regulatory RNA molecules which have been termed sRNAs.

Transcription of DNA into RNA is one of the key processes in molecular biology and is executed in all cellular organisms (Crick, 1970). The enzyme responsible for transcription, the DNA-dependent RNA polymerase (RNAP) is composed of multiple subunits that vary substantially across the three domains of life (Werner & Grohmann, 2011). However, a core set of subunits are conserved in general architecture reflecting the universal functions that the enzyme performs. RNA polymerases locate promoters which are sequences that direct the enzyme to the gene. For the majority of housekeeping genes these promoters are centred -35bp and -10bp upstream from the transcriptional start site (Fig. 5.1). The first base of a gene to be transcribed by RNA polymerase which corresponds to the 5'-most base of the resulting transcript is referred to as the transcription start site (TSS). The RNAP subsequently opens and unwinds the double-stranded DNA and catalyses templated *de-novo* polymerisation of ribonucleotides. After initiation of transcription, the enzyme transitions to an elongation complex until transcription of the gene is completed and the process terminated.

In bacteria, a single RNAP performs all transcription. The enzyme consists of an evolutionary conserved catalytic core which requires the presence of an initiation factor, sigma (σ), for promoter-specific DNA binding and unwinding (Burgess et al., 1969). Sigma factors are responsible for determining promoter specificity and control how efficiently transcription is initiated (Burgess, 2001). An interaction of the σ factor with the core RNAP results in the formation of the active RNAP holoenzyme. All bacteria possess a primary housekeeping σ factor, σ^{70}, which controls transcription of essential genes during growth. Promoters that are

recognized by the primary factor are generally comprised of two conserved sequence motifs upstream of the transcription start site: the -35 element (TTGACA) and the -10 element (TATAAT). Bacteria can also express different sigma factors in response to different environmental conditions allowing adaptation to specific niches. For instance, in *E. coli* under nitrogen starvation conditions the alternative sigma factor RpoN (σ^{54}) is responsible for activating transcription. These sigma factors often also recognise different promoter elements. For RpoN this may be the conserved -24 and -12 consensus sequences (Thöny and Hennecke, 1989).

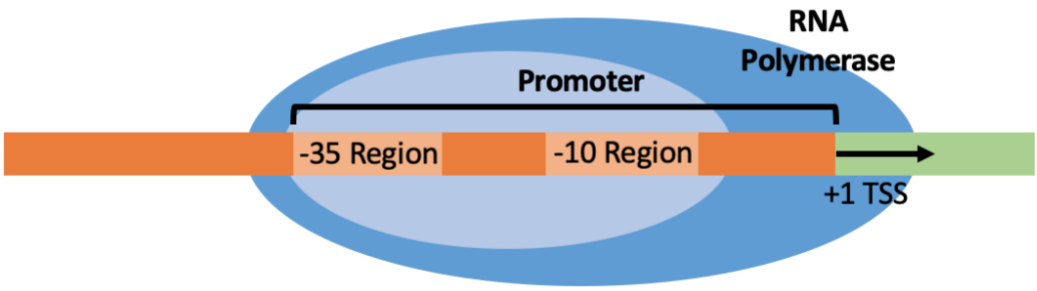


Fig. 5.1: Bacterial promoters recognised by the primary sigma factor are comprised of two conserved regions: the -35 element and the -10 element

To date little is known about the involvement of sigma factors in denitrification. In *Ralstonia eutropha* (formerly known as *Alcaligenes eutrophus*) sigma factor σ^{54} has been implicated in anaerobic growth on nitrate suggesting an importance in the regulation of denitrification-related genes in this organism (Romermann et al., 1989). The organism is known to utilize the σ^{54}-dependent regulator NorR to activate transcription of the *nor* operon (Pohlmann et al., 2000). In contrast, a *Pseudomonas aeruginosa rpoN* mutant is able to anaerobically grow on nitrate. In this organism, σ^{54} was shown to control a diverse set of genes such as those encoding glutamine synthetase, urease and flagellin (Totten et al., 1990). In *P. stutzeri*, expression levels of Nar, Nir, Nor and Nos proteins in *rpoN*-null mutants showed that σ^{54} influenced both nitrite reductase and nitric oxide reductase activity without affecting transcription of their structural genes (Härtig & Zumft 1998). This suggests a role of the sigma factor in post-translational processes rather than at the level of gene expression or potentially

an involvement via an indirect effect. Additionally, the RpoN regulon controls genes involved in the response to nitrogen limitation, nitric oxide stress, availability of alternative carbon sources and nucleic acid damage in *Salmonella* Typhimurium (Hartman et al., 2016). In the model organisms *E.* coli, RpoN affects both nitrogen and carbon metabolism as well as fermentation, cell envelope biogenesis and stress resistance (Reitzer and Schneider, 2001; Riordan and Mitra, 2017). P*. denitrificans* encodes a homologue to σ^{54} as well as homologues to other sigma factors and further insights into the role of these factors in regulation of the denitrification apparatus is required to uncover the transcriptional switches in the model denitrifier (Table 5.1).

Table 5.1: The seven sigma factors of *Escherichia coli* RNA polymerase

Factor	Gene	Size (kDa)	Consensus binding site	Genes regulated
σ^{70} (σ^{D})	*rpoD*	70	TTGACA-N17-TATAAT	Housekeeping
σ^{54} (σ^{N})	*rpoN*	54	CTGGCAC-N5-TTGCA	Nitrogen metabolism
σ^{38} (σ^{S})	*rpoS*	38	TTGACA-N12-TGTGCTATACT	Stationary phase
σ^{32} (σ^{H})	*rpoH*	32	CTTGAA-N14-CCCC	Heat shock
σ^{28} (σ^{F})	*fliA*	28	TAAA-N15-GCCGATAA	Flagellar proteins
σ^{E}	*rpoE*	24	GAACTT-N16-TCTGA	Extreme heat shock
σ^{fecI}	*fecI*	19	GGAAAT-N17-TC	Iron transport

Pinpointing the TSS of an RNA permits the identification of potential regulator binding sites or sigma-factor recognition sites that specify a promoter region. Identifying the regulatory signals required for transcriptional regulation is crucial to understanding the nature of the factors involved in a transcriptional response. A major challenge in transcriptomic studies is the construction of reliable transcriptome maps and the identification of transcription start

sites across the genome. Due to the importance of TSS identification in numerous biological contexts, a large number of methods for global TSS Profiling have been discovered and developed. The first published transcriptome maps were based on a manual curation of putative TSSs, a process that is time-consuming and not easily scalable to the entirety of a genome or to a larger number of organisms and conditions. Advances in sequencing and the reduction of sequencing costs have allowed for the enhanced rate of identification of TSSs.

5.2 Aims

This study aims to explore the differences in the transcriptome of the model denitrifier *P. denitrificans* under denitrifying and non-denitrifying conditions. The overarching aim of this study is to further develop *P. denitrificans* as a regulatory as well as biochemical model of bacterial denitrification. Therefore, it aims to identify and annotate the transcription start sites to provide further insights into promoter activity and explore the role of sigma factors in the regulation of denitrification.

Hypotheses:

- Identifying the transcription start sites will provide insights into the promoter usage in *P. denitrificans*
- There is a sigma factor controlling transcription during denitrification

5.3 Results

5.3.1 Categorisation of TSS

It has been shown that copper availability affects the transcription of enzymes involved in N_2O production in *P. denitrificans*, specifically the activity of the final denitrification enzyme NosZ. To allow further insight into the transcriptional landscape during denitrifying and non-denitrifying conditions, the transcriptome of *P. denitrificans* under CuH and CuL conditions was determined using differential RNA-seq (Gaimster et al., 2016). These datasets generated by Gainster et al., 2016 were utilised to carry out comprehensive mapping of transcription start site across the genome. Active promoters were determined by identifying transcription start sites (TSS) under the two culture conditions. TSS were identified using dRNA-seq data and the automatic TSS annotation tool TSSAR by comparison of the mapped reads, the –TAP values were subtracted from the +TAP values. Then, peaks were called with a cut off value of 10 reads and classed as Primary, Internal and Antisense TSSs (Fig. 5.2A).

For chromosome 1, 6091 individual TSS were annotated under CuH conditions (p-Value cut off 1e-05; noise threshold 2; max range to merge 5). 4844 remained after consecutive TSS were merged. Under CuL conditions, 5074 individual TSS were annotated (p-Value cut off 1e-05; noise threshold 2; max range to merge 5). 3967 TSS remained after consecutive TSS were merged. For chromosome 2, 3507 individual TSS were annotated under CuH conditions (p-Value cut off 1e-05; noise threshold 2; max range to merge 5). 2854 remained after consecutive TSS were merged. Under CuL conditions, 2868 individual TSS were annotated (p-Value cut off 1e-05; noise threshold 2; max range to merge 5). 2313 TSS remained after consecutive TSS were merged. For the plasmid, 1023 individual TSS were annotated under CuH conditions (p-Value cut off 1e-05; noise threshold 2; max range to merge 5). 862 remained after consecutive TSS were merged. Under CuL conditions, 760 individual TSS were annotated (p-value cut off 1e-05; noise threshold 2, max range to merge 5). 645 TSS remained after consecutive TSS were merged.

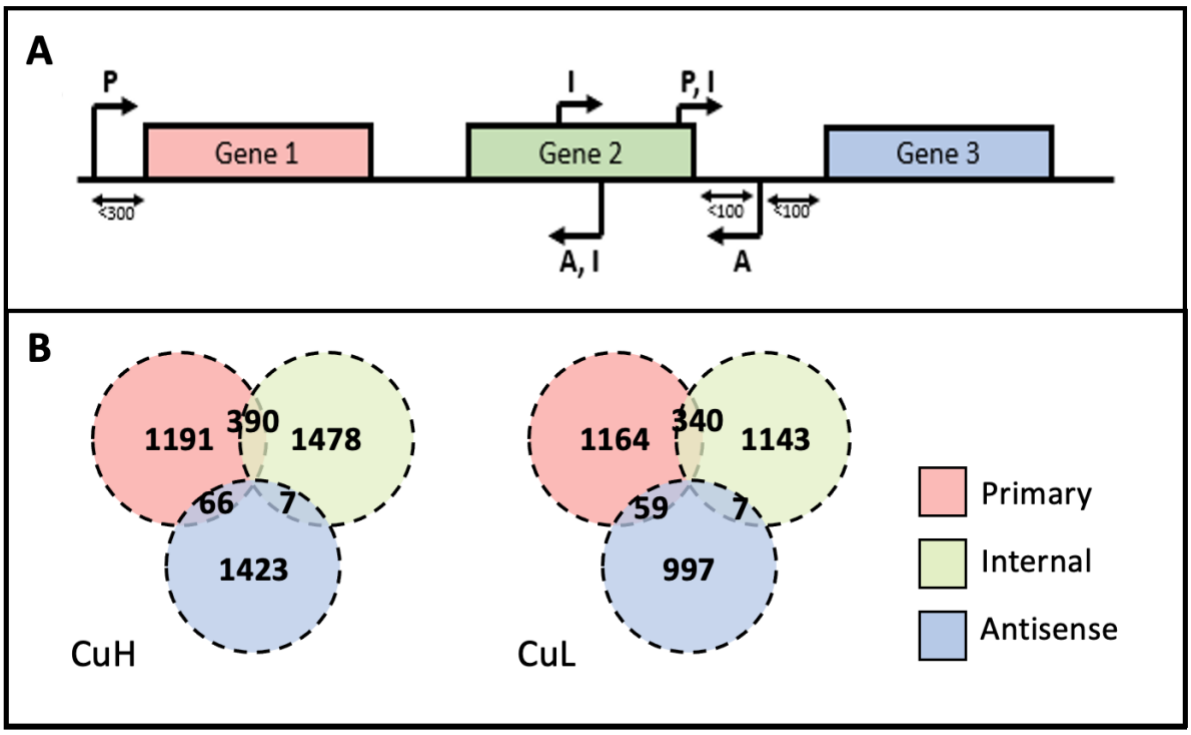


Fig. 5.2 A: Categorization of TSS into different classes; TSS positioned upstream of annotated gene (primary), TSS located internally to a gene (internal) and antisense TSS (antisense) **B:** The three classes of TSS identified across *P. denitrificans* Chromosome 1 under CuH and CuL conditions

The location of TSS were mapped onto the genome using the web-based tool Proksee, an expert system for genome assembly, annotation, and visualisation (Fig. 5.3) (Grant et al., 2023). The Proksee output showed the presence of TSS (dark green) alongside the coding sequences (dark blue) present across the two chromosomes and the plasmid under copper high and copper low conditions highlighting that TSSs were evenly dispersed across the genome relative to the number of genes present on each chromosome. There were significantly more unique TSS present under Cu-high conditions (Fig. 5.4A) than under Cu-low conditions (Fig. 5.4B). This could be attributed to the ability to carry out complete denitrification, with several genes involved in the pathway switched on rather than stunted denitrification resulting in the release of N_2O, as observed in the absence of copper.

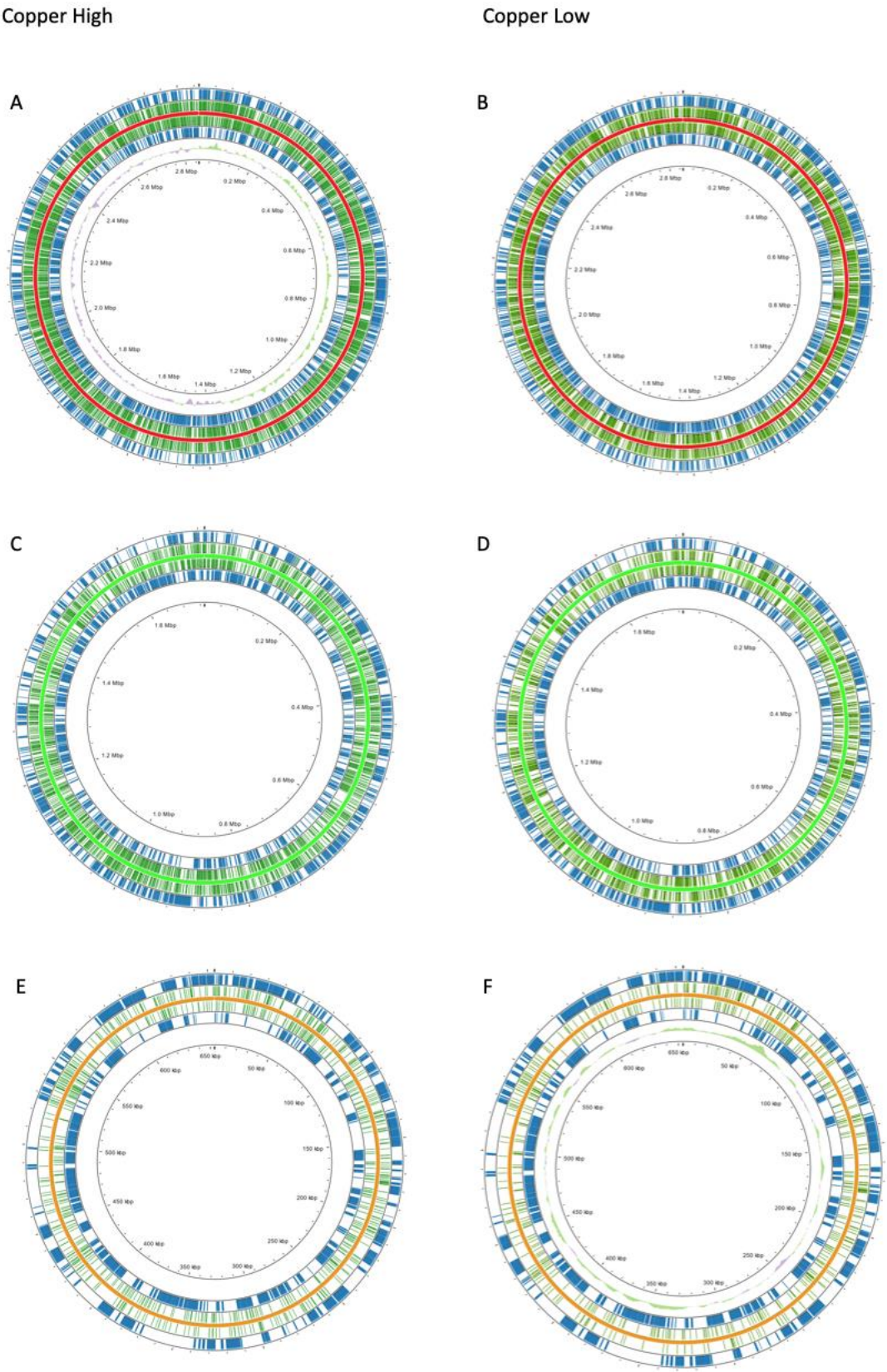


Fig. 5.3 TSS loci across the two *P. denitrificans* chromosomes and the plasmid generated using Proksee (Grant et al., 2023). **A:** chromosome 1 (red), Cu-high. **B:** Chromosome 1 (red), Cu-low. **C:** chromosome 2 (light green), Cu-high. **D:** chromosome 2 (light green), Cu-low. **E**: plasmid (orange), Cu-high. **F:** plasmid (orange), Cu-low. Coding Sequences are shown in dark blue and location of TSS is marked in dark green

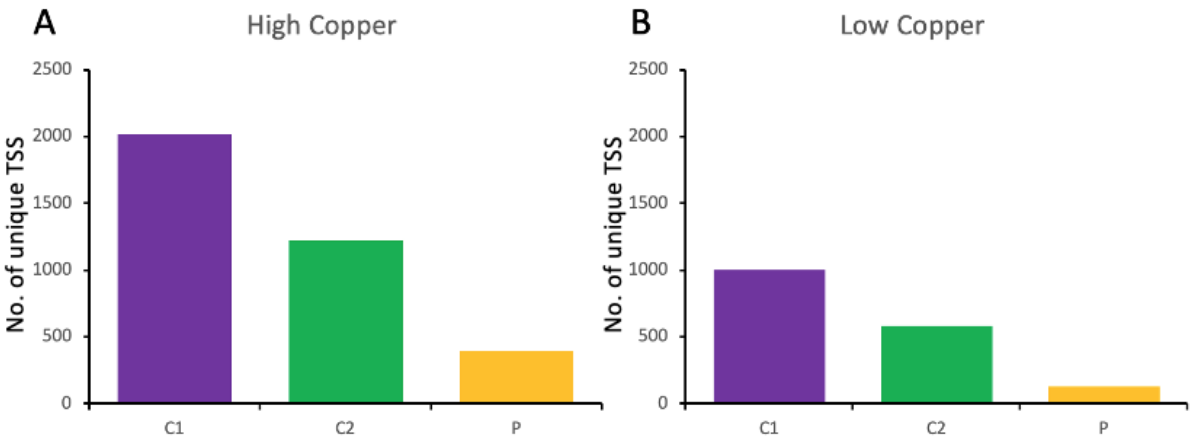


Fig. 5.4 Number of unique TSS under CuH and CuL conditions on chromosome 1 (C1), chromosome 2 (C2) and the plasmid (P). Highest numbers of unique TSS are present on chromosome 1 corresponding to chromosome size. **A** During CuH conditions more unique TSS were Identified.

5.3.2 Benchmarking the differences in transcriptional landscape using the *nos* operon

To benchmark the differences in the transcriptional landscape during denitrification (Cu-high) conditions and non-denitrification conditions (Cu-low), transcription of the *nos* operon was analysed in close detail. The structural gene for the enzyme nitrous oxide reductase (*nosZ*) is located within a gene cluster consisting of genes encoding for products required for the maturation of the denitrification enzyme. The transcriptional landscape of the operon for each condition was visualized using the Integrated Genome Browser to allow comparison and to manually annotate the transcription start sites based upon the results obtained from the transcription start site analysis. As expected, expression levels of *nosZ* and *nosR* appeared to be upregulated under CuH conditions with several primary and internal TSSs detected across both genes (Fig. 5.5). Under CuL conditions however, the expression levels are low and fewer TSSs were identified. Only one internal TSS within the *nosZ* gene and one primary TSS upstream of *nosR* remained, in comparison to the TSSs detected under CuH conditions.

The *nosC* gene however is more highly expressed under Cu limitation and has an additional primary TSS 227bp upstream of the *nosC* gene (Fig. 5.5). There are two further TSSs that are

only present under CuL conditions, which are located antisense to the *nosF* and the *nosD* gene.

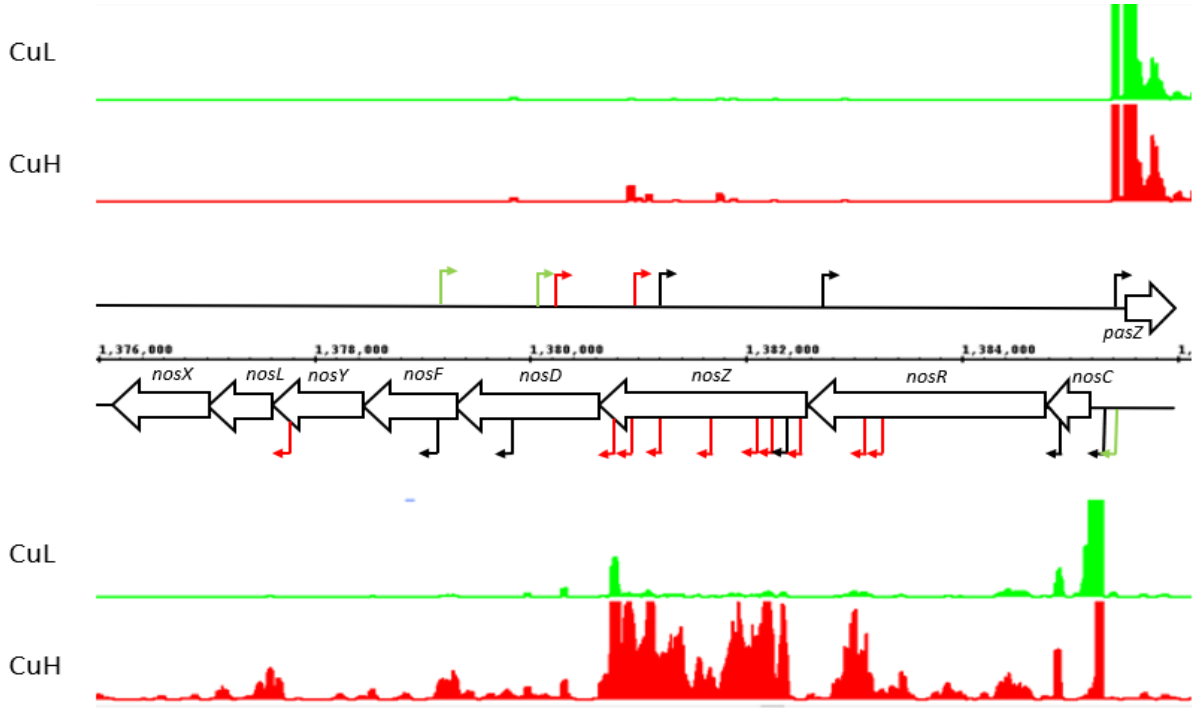


Fig. 5 Visualization of the mapped sequence reads of the nos operon in *Paracoccus denitrificans* under CuH (red) and CuL (green) conditions (Integrated Genome Browser). The names of coding genes are labelled in black and the TSS are indicated by cornered arrows with the TSS pointing towards the right indicating TSS on the positive strand and TSS oriented towards the left indicating the negative strand. TSS only present under CuH conditions are indicated in red, whilst TSS present only under CuL conditions are visualised in green. TSS present under both conditions are shown in black.

5.3.3 Visualisation of TSS

To allow for fast and convenient visualisation of the TSSs across the complete *P. denitrificans* genome, the TSS information files for each chromosome were converted to GFF files which can be easily viewed alongside the chromosome using the Artemis Genome Browser. As a result, the browser displays location and name and number of the TSS on the forward or the reverse strand of the genome. Furthermore, an sRNA information file was manually created to allow the display of sRNA locations on each chromosome (Fig. 5.6). The sRNAs upregulated under high N_2O emitting conditions were marked in pink, while the sRNAs downregulated

under high N_2O emitting conditions were marked in turquoise. The sRNAs upregulated under low N_2O emitting conditions were marked in green and finally, the sRNAs downregulated under low N_2O emitting conditions were marked in orange. A visualisation of TSSs as well as sRNAs will allow for further ease of transcriptome analysis and will aid in addressing further questions surrounding the *P. denitrificans* transcriptome.

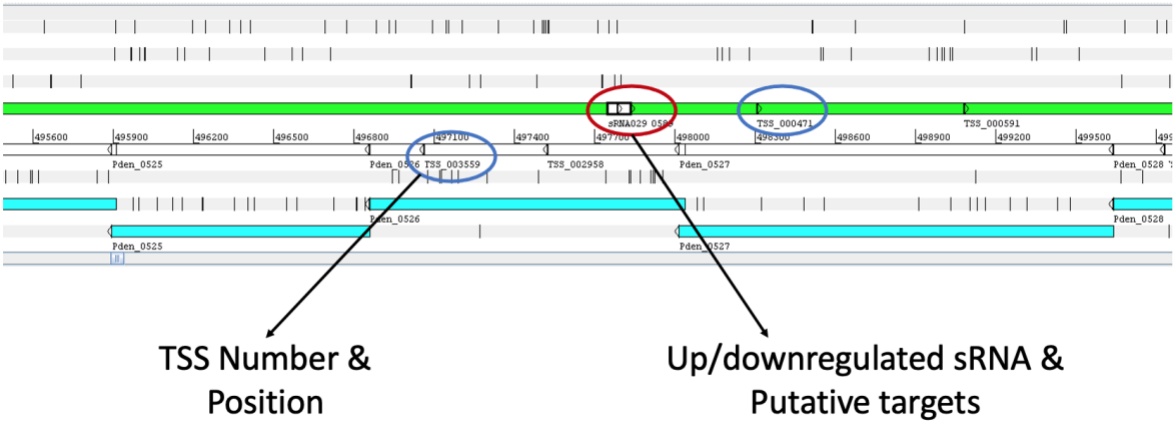


Fig. 5.6 Visualisation of TSSs and sRNAs alongside the *P. denitrificans* genome using the Artemis Genome Browser. TSSs are shown as arrows, as indicted by the blue circle while sRNAs are labelled with their name as part of the genome annotation file

5.3.4 Identification of TSS across the denitrification enzyme genes

Having identified the location of TSSs across the *P. denitrificans* genome, further gene operons important for the switch between complete denitrification and N_2O emitting incomplete denitrification conditions could be analysed to compare the transcriptional activity during denitrifying and non-denitrifying conditions. The first operons to be analysed were the four denitrification enzyme operons: *nar*, *nir*, *nor* and *nos* encoding for nitrate reductase, nitrite reductase, nitric oxide reductase and nitrous oxide reductase. For all four operons differences between the two conditions could be observed, with some TSSs only present under CuH conditions and others only present under CuL conditions.

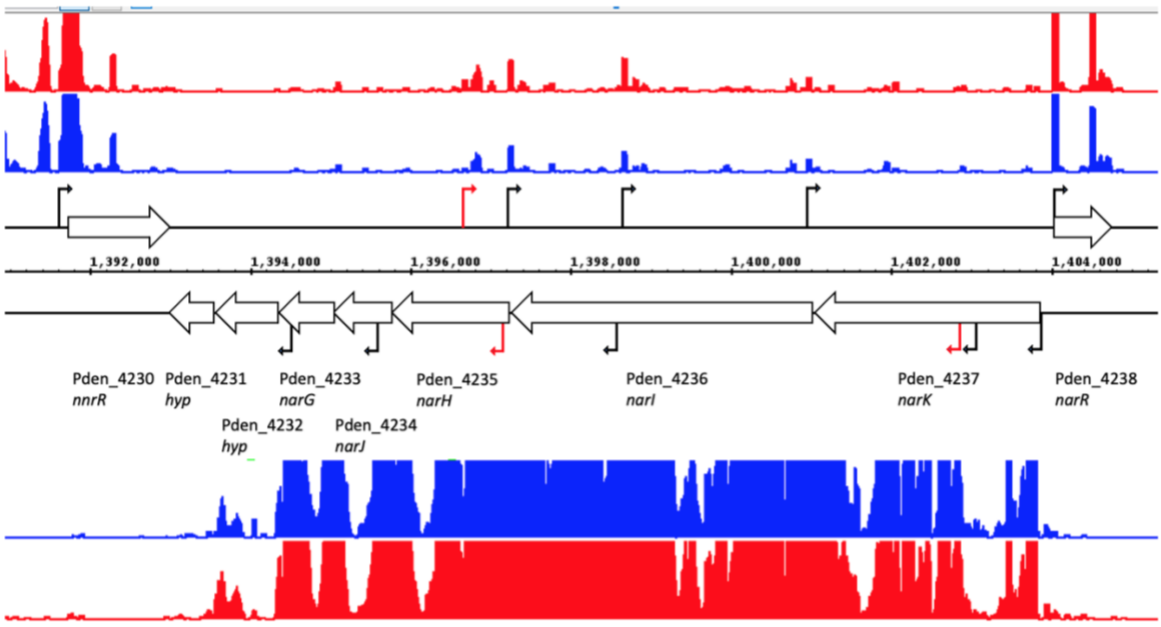


Fig. 5.7 Visualization of the mapped sequence reads of the *nar* operon in *Paracoccus denitrificans* under CuH (red) and CuL (blue) conditions (Integrated Genome Browser). The names of coding genes are labelled in black and the TSS are indicated by cornered arrows with the TSS pointing towards the right indicating TSS on the positive strand and TSS oriented towards the left indicating the negative strand. TSS only present under CuH conditions are indicated in red. TSS present under both conditions are shown in black

For the nitrate reductase operon (*nar*) there were three additional TSSs identified under CuH conditions, with two present on the negative strand and one on the positive strand (Fig. 5.7). The two TSSs present on the negative strand lie within the nar*H* and the nar*K* genes and are classified as internal TSSs. The additional TSS on the positive strand is an antisense TSS to *narH.* The expression patterns however only differed at very low levels, confirming that *nar* expression is not significantly affected by the lack of copper from the culture media.

For the nitrite reductase operon (nir) there were also three additional TSSs annotated under CuH conditions (Fig. 5.8). All three of them are located on the positive strand and are internal TSSs to the *nirS*, the *nirH* and the *nirN* genes. Similar to the expression changes observed in

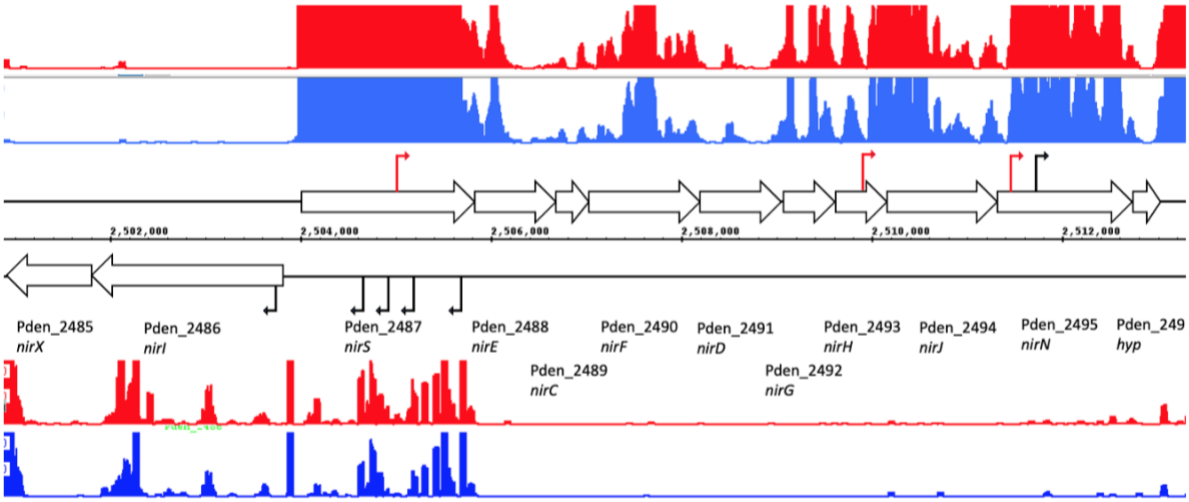


Fig. 5.8 Visualization of the mapped sequence reads of the *nir* operon in *Paracoccus denitrificans* under CuH (red) and CuL (blue) conditions (Integrated Genome Browser). The names of coding genes are labelled in black and the TSS are indicated by cornered arrows with the TSS pointing towards the right indicating TSS on the positive strand and TSS oriented towards the left indicating the negative strand. TSS only present under CuH conditions are indicated in red. TSS present under both conditions are shown in black.

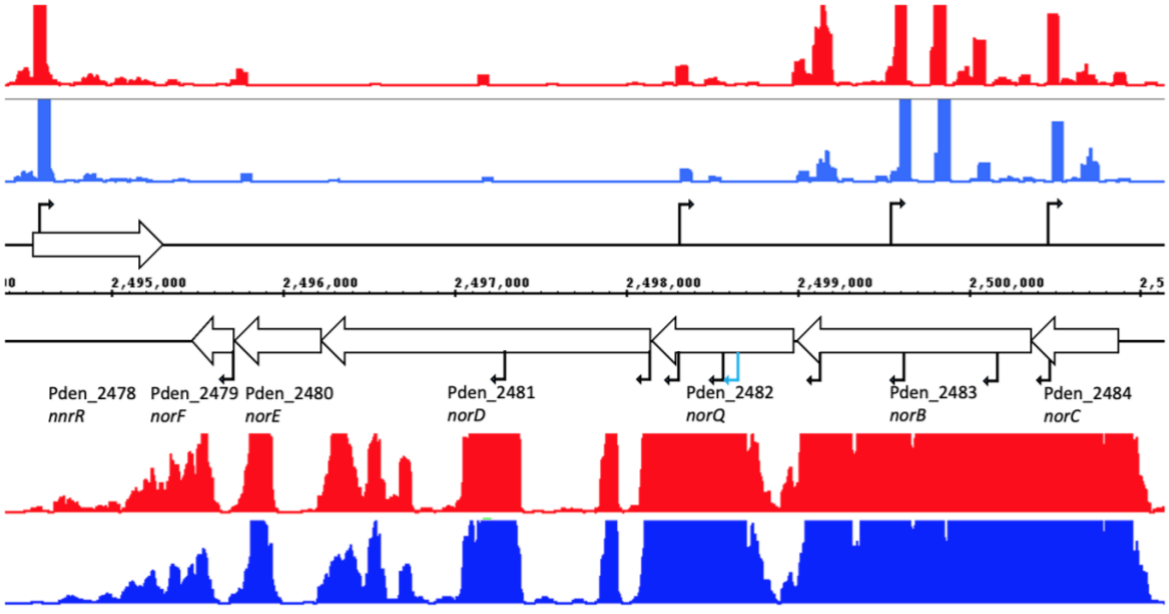


Fig. 5.9 Visualization of the mapped sequence reads of the *nor* operon in *Paracoccus denitrificans* under CuH (red) and CuL (blue) conditions (Integrated Genome Browser). The names of coding genes are labelled in black and the TSS are indicated by cornered arrows with the TSS pointing towards the right indicating TSS on the positive strand and TSS oriented towards the left indicating the negative strand. TSS present only under CuL conditions are visualised in blue. TSS present under both conditions are shown in black

the nar operon, the expression of the nir operon is also not significantly affected by the lack of copper from the culture media.

The TSSs of the *nar* operon under CuH and CuL conditions only differ at one single TSS which is present only in the absence of copper from the culture media (Fig. 5.9). This TSS is an internal TSS located on the negative strand within the narQ gene. All other TSSs remained the same under both conditions and the expression patterns of the operon did not change significantly between the two conditions.

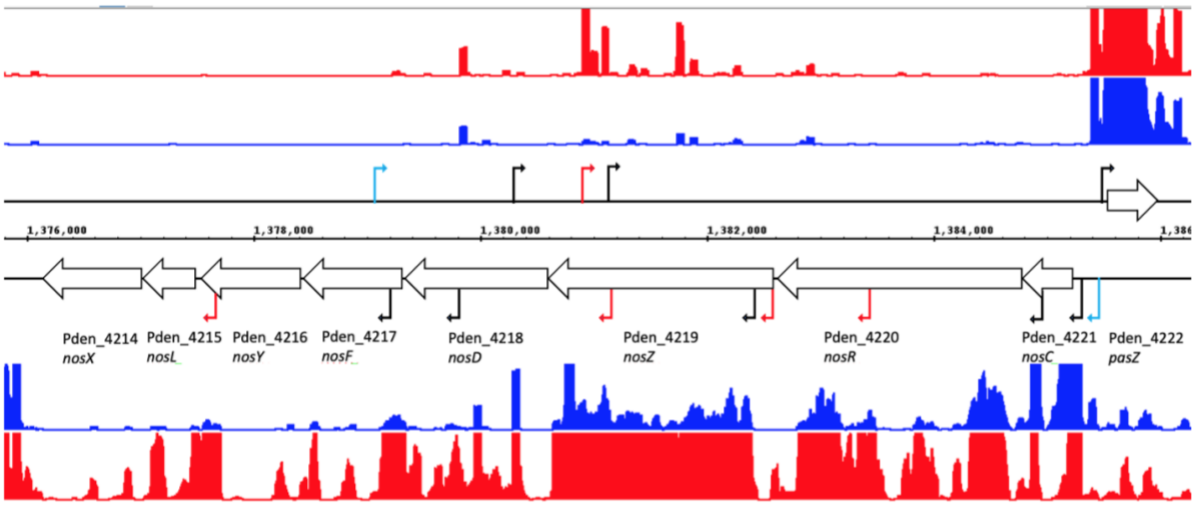


Fig. 5.10 Visualization of the mapped sequence reads of the *nos* operon in *Paracoccus denitrificans* under CuH (red) and CuL (blue) conditions (Integrated Genome Browser). The names of coding genes are labelled in black and the TSS are indicated by cornered arrows with the TSS pointing towards the right indicating TSS on the positive strand and TSS oriented towards the left indicating the negative strand. TSS only present under CuH conditions are indicated in red, whilst TSS present only under CuL conditions are visualised in blue. TSS present under both conditions are shown in black

The final denitrification operon, the *nos* operon was most significantly affected by the presence or absence of copper from the culture media as seen in 5.3.2. After merging consecutive TSSs, four additional TSSs were identified purely under CuH conditions and two TSSs were present only under CuL conditions (Fig. 5.10). Three of the four TSSs present only in the presence of copper were located on the negative strand, with two present internally of the *nosZ* and the *nosR* gene and one located upstream of *nosZ*, therefore classified as primary TSS. The fourth TSS present only under CuH denitrifying culture conditions is located on the

positive strand antisense to *nosZ*. Of the two TSSs present only under CuL conditions, one TSS is located antisense to *nosF* while the other is a primary TSS to *nosC*. The expression patterns of the operon are changed significantly in the presence or absence of copper and therefore, the nos operon is the denitrification enzyme operon that is most heavily impacted by a change in copper concentrations.

5.3.5 Identification of conserved promoter motifs

As little is known about conserved promoter motifs across the *P. denitrificans* genome, the initial search focussed on an identification of the core promoter regions with its -10 and -35 elements. Sequences up to 60nt upstream of all identified primary TSSs were extracted and submitted to Improbizer, a motif-finding algorithm that considers location of sequence patterns within the input sequences and favours motifs that occur at the same location (Ao et al., 2004). A majority of sequences possess a motif centred at position -35/-36 (CTTGCC) as well as a second motif at position -11 to -16 (GCATT) (Fig. 5.11). The second motif is A and T rich, which is also the case for the highly conserved -10 hexamer of *E. coli* with the consensus sequence TATAAT. Despite the presence of As and Ts, the sequence differs from the *E. coli* consensus sequence and therefore neither of the two identified motifs are closely related to the -10 and -35 boxes known to be recognised by the housekeeping sigma factor in *E. coli*.

.

CuH489

Profile Motif 1:

motif score: 4.2549	motif position: 22.74 (sd=1.17)			consensus=CTTGCC	
a	0.2	0.135	0.099	0.003	0.337	0.249
c	0.435	0.127	0.119	0.003	0.541	0.402
g	0.23	0.142	0.271	0.81	0.004	0.175
t	0.135	0.596	0.511	0.183	0.118	0.175

Profile Motif 2:

motif score: 3.7905	motif position: 44.17 (sd=3.06)			consensus=GCATT	
a	0.099	0.003	0.58	0.441	0.281
c	0.3	0.601	0.003	0.004	0.154
g	0.404	0.392	0.003	0.003	0.278
t	0.197	0.003	0.413	0.552	0.288

Fig. 5.11 Base scores and motif position for two motifs in the *P. denitrificans* promoter sequences identified using Improbizer. Scores for each base at its corresponding position are shown and overall motif score was calculated from the distribution of bases across the motif position.

5.3.6 Sigma factor utilisation in *P. denitrificans*

A bioinformatics survey of the *P. denitrificans* genome has revealed the presence of genes predicted to encode the major sigma factors that are known to be present in related bacteria. These include RpoD (Pden_4072), RpoN (Pden_4987), RpoH (Pden_2216) and RpoE (Pden_2820) (Table 5.2). An alignment of the predicted housekeeping sigma factor RpoD of *P. denitrificans* with RpoD of *E. coli* showed a similar length of the two gene transcripts as well as high levels of sequence similarity (Fig. 5.12). Additionally, a number of extracellular sigma factors, ECF 1-4 are present. Interestingly, a second gene encoding for a sigma factor of similar length to RpoD was identified. Its C-terminal shares significant sequence similarity to the flagellum-specific sigma factor RpoF (24% identity, 46% similarity to the full-length sequence of FliA from *E. coli*).

Table 5.2: Predicted *P. denitrificans* sigma factors and their % similarity to the *E. coli* homologs. * indicates % similarity to *E. coli* RpoE.

σ Factor	Gene	*P. denitrificans* homolog	% Similarity
70	***rpoD***	**Pden_4072**	**64**
54	***rpoN***	**Pden_4987**	**44**
38	***rpoS***		
32	***rpoH***	**Pden_2216**	**63**
28	***rpoF***	**Pden_2086**	**46**
24	***rpoE***	**Pden_2820**	**48**
ECF1		**Pden_4533**	**46***
ECF2		**Pden_1137**	**43***
ECF3		**Pden_4558**	**46***
ECF4		**Pden_1629**	**44***

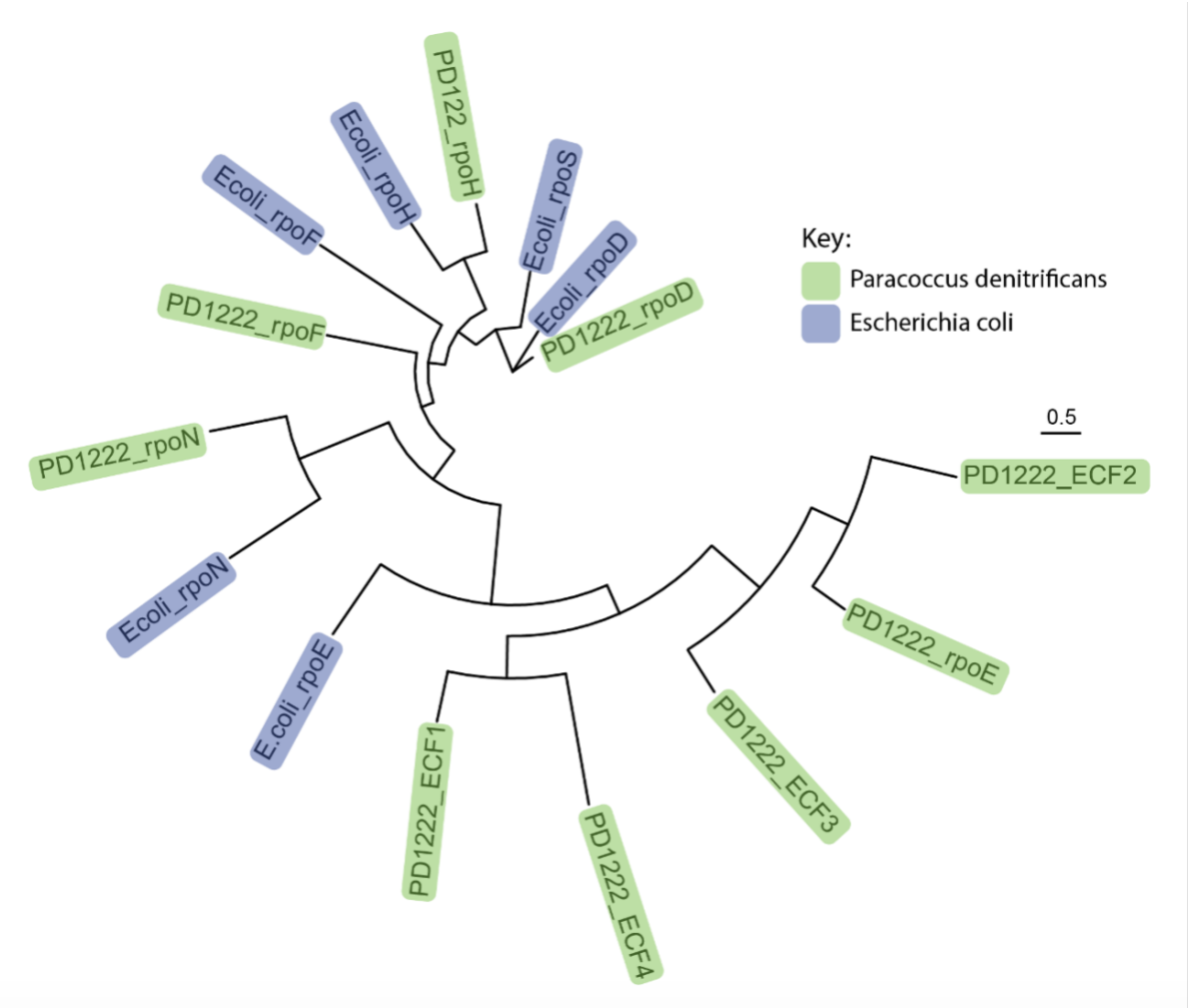


Fig. 5.12: Maximum likelihood phylogenetic tree based on amino acid variation from the multiple sequence alignment of the ten sigma factor homolog genes present in *P. denitrificans* (green) and the seven sigma factor genes in the model bacterium *E. coli* (blue).

Homologs in *P. denitrificans* and *E. coli* cluster together which suggests sequence homology (Fig. 5.12)**.** This applies to all sigma factors, except for RpoS, of which a homolog is not present in P. denitrificans as well as for *P. denitrificans* rpoE, which clusters together with the extracellular sigma factors ECF1, ECF2, ECF3 and ECF 4.

5.4 Discussion

This study is the first genome wide analysis of TSS in *P. denitrificans*. Through differential RNA-Sequencing combined with bioinformatics methods this chapter annotated primary, internal, and antisense TSS for genes expressed under complete denitrification (Cu-high) conditions and incomplete denitrification (Cu-low) conditions. As the majority of Europe's arable soils are copper deficient, it is important to gain an insight into the transcriptional landscape during incomplete denitrification conditions in the absence of copper to identify targets for future mitigation strategies (Sinclair and Edwards, 2008). The presence of antisense and internal TSS alongside primary TSS suggests that the *P. denitrificans* transcriptome is highly complex. The number of TSS identified across the *P. denitrificans* genome is comparable to the number of TSS identified in other genome-wide transcription start site profiling studies, such as in *Burkholderia cenocepacia* J2315, in which 3908 TSS were annotated (Sass et al., 2015). An identification of TSS allows for an addition of relevant information to the Pden_1222 annotation which will provide the starting point for future evaluation of regulation of denitrification in the model denitrifier *P. denitrificans*.

Significantly more unique TSS were present, when the growth media was rich in copper, which can be explained by the active expression of several genes involved in the final denitrification step, the conversion of the greenhouse gas nitrous oxide to the harmless atmospheric dinitrogen. The levels of *nosZ* and *nosR* expression displayed as a result of the differential RNA-seq confirms previously published data by Sullivan et al., 2013 with *nosZ* being expressed at low levels in the absence of copper due to the high Cu-demand of the enzyme. The products of *nosR* and *nosC* are involved in copper transport and *nosZ* assembly and have therefore been shown to influence the indirect transcriptional regulation of *nosZ*. The gene *nosX* is homologous to *nirI* (Pden_2486) and NosX is likely to be involved in metal ion transport and the assembly of *nosZ* (Giannopoulos, 2014). The remaining genes *nosDYFL* show sequence similarities with the equivalent genes found in *Pseudomonas,* in which they have been assigned a role in NosZ maturation.

NosR has been shown to contain residues for binding Cu, which could explain its involvement in Cu-dependency of N_2O reduction. In anaerobic denitrification conditions, the oxygen

sensitive transcriptional regulator FNR binds upstream of *nosZ* where two FNR-binding sequences are present. The first one is located at position -21. The second one is located at position -126 and is not thought to be involved in transcriptional binding as it is located far away from the start codon. FNR also binds upstream of *pasZ,* which encodes for pseudoazurin, which together with cytochrome c_{550} is responsible for donating electrons to Nos via cytochrome bc_1 (Giannopoulos, 2014). Indeed, in a proteomic study *pasZ* had been shown to have been induced by FNR (Bouchal et al., 2010). The *nosC* gene however is more highly expressed under Cu limitation which could be a result of an additional primary TSS 227bp upstream of the *nosC* gene (Fig. 5.3). NosC could therefore act as a repressor to *nosR*, explaining the lower levels of the NosR protein in the absence of copper. Further analysis of the differential transcription start sites under high and low N_2O emitting conditions across the genes known to be involved in denitrifications, as well as novel genes that show differential expression across the two conditions will allow a more in depth overview of all the promoters turned on or off depending on the availability of copper in the environment.

One of the most surprising knowledge gaps in terms of integrating our understanding of transcriptional regulation of denitrification is the absence of data on which sigma factors associate with the core RNA polymerase during anaerobic denitrification. *P. denitrificans* possesses homologs to multiple sigma factors known to target diverse sets of promoters in *E. coli* forming complex networks. The presence of sigma factor homologs suggests that transcription in *P. denitrificans* is also dependent on the presence of specific sigma factors that recognize distinct promoter sequences and activate gene expression in response to changes in environmental signals. Despite the association of sigma 54 (RpoN) with nitrogen-regulated genes, there has been no direct association of RpoN with denitrification.

The layout of a sigma factor network directly impacts global transcriptional regulation and therefore dictates the phenotype of the bacteria. Bacterial promoters usually contain specific sequences for binding of RNA polymerase-associated sigma factors, for example the -35 and the -10 boxes of the housekeeping sigma factor RpoD in *E. coli* (Paget, 2015). These motifs however can vary strongly between alternative sigma factors with different recognition motifs, extended motifs and structural features of the sigma proteins that influence DNA melting capacity (Feklistov et al., 2014). Furthermore, sigma factors can be present as singly

copy gene or as multi copy genes across the different bacterial families. The heat shock sigma factor, RpoH, is present as a single copy gene in gammaproteobacterial and as a multi copy gene in many alphaproteobacterial genomes (Green and Donohue, 2006, Martinez-Salazar et al., 2009; Lopez-Leal et al., 2016). These copies deal with different environmental responses, with *rpoH*$_1$ responsible for heat shock and *rpoH*$_2$ responsible for osmotic-shock responses with both recognising diverse consensus sequences (Green and Donohue, 2006). The denitrifier *P. denitrificans* is predicted to have only one copy of each sigma factors, however, the presence of homologs to the ECF sigma factors provides further putative regulatory depth. Some sigma factors, such as the ECF group, have a concerted, focused response to very specific environmental conditions targeting only small regulons which can lead to a lack in global consensus motifs.

Initial motif searches carried out in the -10 and -35 regions upstream of the TSS revealed the presence of two conserved motifs in *P. denitrificans*. Despite the significant abundance of these two motifs across the active promoters during denitrification, these two motifs did not have close resemblance to the consensus sequences of the housekeeping sigma factor RpoD or the nitrogen-associated transcription factor RpoN in *E. coli* (Paget, 2015). In the facultative phototrophic Alphaproteobacterium *Rhodobacter sphaeroides*, sigma factors and their recognition sites also vary significantly from those found in *E. coli*. It has however been shown, that *R. sphaeroides* possesses a very intricate network with sigma factors of overlapping functions, allowing the organisms to carefully adapt to a variety of environmental conditions (Zhang, 2023). This could also be the case for sigma factor regulation in *P. denitrificans*, making it harder to elucidate the exact regulatory functions and recognition motifs of each sigma factor and to identify consensus sequences.

It is likely, that the identified consensus sequences in this study are of importance to the genes activated under denitrification conditions and may be sequences recognized by one or more sigma factors responsible for controlling this physiological process, however, which sigma factor is responsible remains to be answered and the search will have to be expanded to identify smaller groups of motifs within the active promoters during denitrification. As past research has shown, the higher the specificity of a sigma factor, the smaller the regulon, which can range from as few as two to three genes to as many as several hundred (Feklistov et al.,

2014). To avoid an oversight of important promoter groups, the motif search will have to be fine-tuned, to answer the questions around active sigma factors during denitrification. Furthermore, chromatin immunoprecipitation sequencing could be used to screen the P. denitrificans genome for sites bound by the sigma factors to identify exactly which regions are captured by which sigma factors and establish their regulons.

The consensus view remains that transcription from prokaryotic promoters is unidirectional, a concept that has long been challenged in eukaryotes in which it has been shown that many RNAP II promoters simultaneously stimulate antisense transcription while driving the production of a canonical sense RNA (Core et al., 2008). In *E. coli*, it has been shown, that bidirectional promoters have inherent symmetry allowing RNAP to bind the same section of duplex DNA in either orientation (Warman et al., 2021). This can allow for coregulation of divergent operons through binding of transcriptional repressors at binding sites that overlap a bidirectional promoter sequence adding a further level of complexity. Having established the importance of a fine-tuned transcriptional response in the switch between N_2O production and N_2O reduction, it is not unlikely, that *P. denitrificans* possesses bidirectional promoters in a similar manner to those characterised by Warman et al., 2021, especially as these bidirectional promoters are speculated to be capable of giving rise to antisense RNAs impacting adjacent genes. Studying promoter motifs in *P. denitrificans* more extensively, may therefore give further insights not only into sigma factor utilisation but also help in the integration of further sRNAs into the complex regulatory networks.

This thesis has provided a starting point for the future analysis of promoter usage and sigma factor utilisation and has opened up avenues for further in-depth analysis of the transcriptional complexity in *P. denitrificans*, to strengthen its use as a regulatory and biochemical model for denitrification. As a next step, it would be beneficial, to expand TSS analysis in the denitrifier to those present under aerobic conditions, to compare the TSS present under non-denitrifying conditions to those established in this thesis, to gain a better understanding of the changes occurring during the switch between N_2O emission and N_2O consumption. Furthermore, the 5’ untranslated regions should be analysed by identifying sequences between the TSS and the start codon of a gene or gene cluster. This will contribute to the understanding of *P. denitrificans* genes not only involved in the final step of the

denitrification but also in the global switch between aerobic and anaerobic metabolism as well as the transcriptional regulators involved in switching these genes on and off and will yield new targets for future research and the development of novel mitigation strategies.

5.4.1 Conclusion

The data presented in this study provides the starting point for evaluation of the denitrification regulatory switches and has opened paths for the identification of further insights into the role of sigma factors in these regulatory processes. TSS and sRNA locations have been mapped onto the genome and have been made available to view using the Artemis Genome Browser. This tool has the potential to be developed further to include transcriptional regulator binding sites as well as changing transcription levels with changing environmental conditions to give a global overview over the transcriptional landscape by extending the research into sigma factor usage and enhanced promoter complexity.

6. Discussion and Concluding remarks

As the individual results of each chapter have been discussed in the previous chapters, this general discussion chapter will highlight the major outcomes of this study, the wider context, and the potential impact.

6.1 Context

Nitrous oxide is a harmful greenhouse gas that has nearly 300 times the global warming potential of carbon dioxide. Moreover, nitrous oxide remains in the atmosphere for an average of 114 years, where it may be converted to nitrogen oxides capable of depleting the stratospheric ozone layer (Goldblatt et al., 2009). While about 60% of global nitrous oxide emissions occur naturally, the remaining emissions are caused by anthropogenic influences such as agriculture (Vitousek and Howarth, 1991). The global rise in nitrous oxide emissions from 2020 to 2021 was higher than the average annual growth rate over the past 10 years, highlighting the pressing need for mitigation strategies to reduce emissions of this harmful gas while also sustaining the agricultural processes required to feed the ever-growing world population (Richardson et al., 2009). To identify novel approaches for the development of mitigation strategies it is crucial to enhance our understanding of the microbial processes underlying the production of nitrous oxide in the natural environment, as these are the major contributors to the production of nitrous oxide. The model denitrifier *Paracoccus denitrificans* has thus far been used as a model for denitrification and it has become increasingly clear that there are many gaps regarding our knowledge around the regulation of denitrification and the emissions of nitrous oxide from this model.

The discovery of small RNAs as regulatory molecules has shown that our knowledge around bacterial regulatory networks in the most diverse environments is far from complete (Wassarman, 2002; Möller et al., 2021). Small RNAs act as regulatory switches in many of the

major physiological processes and have been proven to be of importance for the adaptation to environmental stress and changes in nutrient availability. Small RNAs may act in different ways by binding to mRNA or protein targets to in turn repress or activate the expression of genes to change a transcriptional response. With more sRNAs being studied, more mechanisms of sRNA-induced regulation are pinpointed. The discovery of 167 of these short regulatory molecules has suggested an importance of sRNAs in denitrification, and indeed, one single sRNA was capable of altering nitrous oxide emission levels of denitrifying cultures when overexpressed (Gaimster et al., 2016).

6.2. Why study sRNAs in *P. denitrificans*

Although sRNAs regulate a wide range of important biological processes, our current understanding of their role is far from complete. A manipulation of only a handful of these sRNAs in the lab, including DenR, can lead to drastic changes in the response of the lab organism to an experimental condition. This suggests that there are much larger networks of unknown sRNAs required in an entire environmental response. Uncovering these sRNA regulatory networks, understanding the environmental stimuli involved and integrating these into the current picture of transcriptional regulator networks and other regulatory molecules known to be of importance will give us a broader insight into how all major biochemical processes including denitrification are regulated (Moeller et al., 2021). The use of high-throughput RNA seq has sped up the identification of sRNAs across bacterial genomes, however only a few selected sRNAs have been characterised in detail. This study has further characterised several sRNA in *P. denitrificans* capable of both enhancing and slowing the rates of N_2O emissions from a culture and has provided further evidence that they are crucial links in the development of novel mitigation strategies that have to be integrated into the existing regulatory networks.

Besides the nitrogen cycle, bacterial small RNAs have also been shown to play crucial roles in other biogeochemical cycles highlighting the importance of uncovering their roles not only in

the context of denitrification, but also in the context of other biogeochemical cycles. Studies focussing on marine cyanobacteria have identified several sRNAs involved in the regulation of the photosystem, which is a key player in photosynthesis. Furthermore, the discovery of 99 putative sRNAs under carbon and nitrogen-limiting conditions in the model marine bacterium *Ruegeria pomeroyi* is of interest, as this microbe is suggested to scavenge for alternate sources of organic carbon, affecting the ratios of major biomolecules in carbon-limited conditions (McCarren et al., 2010; Rivers et al., 2016). Interestingly, one of these sRNAs showed homology to 6S RNA, an important sRNA in *E. coli* known to downregulate multiple genes during environmental stress such as nutrient limitation (Cavanagh and Wassarman, 2014). In *R. pomeroyi*, 6S was upregulated under C- and N-limitation and showed altered expression patterns in the context of sulfur metabolism making this sRNA a player in not one but several biogeochemical cycles. This finding highlights the importance, of examining crucial sRNA players in *P. denitrificans* with homologues in other bacteria, to identify conserved roles across bacterial species and potentially across biogeochemical cycles.

Carbon and sulfur cycling in the biosphere are tightly interwoven through bacterial processes carried out by marine microorganisms. Dimethylsulfonioproprionate (DMSP) represents a major carbon and reduced sulfur source in the aquatic environment. The genes involved in DMS-production pathways have been extensively studied, but knowledge of the regulatory networks controlling the pathways is still lacking (Williams & Todd, 2019). The discovery of 182 putative sRNAs in *R. pomeroyi* when grown on DMSP with predicted targets involved in regulation, transport and signalling suggested a potential involvement of sRNAs in posttranscriptional regulation of both DMSP metabolic pathways (Burns et al., 2016). Further sRNAs, for instance sRNA42 and sRNA53 have been implicated in the regulation of the DMSP lyase *dddQ.* Uncovering the roles of these and other sRNAs found in marine bacteria will enhance the understanding of the cycling of carbon and sulfur as well as other elements in the ocean which is an important reservoir for all crucial elements. The confirmed existence of a plethora of sRNA with crucial regulatory roles in biogeochemical pathways further underlines the importance of integrating their functions across cycles, studying their existence and their roles across species.

Recent work has demonstrated, that bacterial small RNAs could be applied in the development of novel diagnostic tools (Tarallo et al., 2019; Mjelle et al., 2020; Vogel et al., 2020). Colorectal cancer patients display differences in their stool sRNA profile compared to healthy patient samples which is a result of host-microbiome dysbiosis (Tarallo et al., 2019). Furthermore, pathogenic bacteria such as *Mycobacterium tuberculosis* secrete sRNAs which can subsequently be detected in the patient plasma. These sRNAs therefore act as biomarkers for diseases. A pathogen commonly associated with colorectal cancer is *Fusarium nucleatum* (Brennan and Garrett, 2019). A depletion of this pathogen from the colon by administration of a short antisense sRNA targeting an essential gene of the bacterium could provide further insights into the role of this pathogen in the development of the disease (Vogel, 2020). Despite major remaining knowledge gaps regarding the transcriptome structure and cell envelope of these potential target organisms, programmable RNA 'antibiotics' are a promising approach in the targeting of antibiotic resistances in the future.

This concept can also be applied in non-host microbiomes. Gaining an overview of the entirety of sRNA present in an environmental sample or in an environmental response such as denitrification could provide valuable insights into the specific responses and the regulatory switches present at the current state of the sample. Similar to the use of sRNA biomarkers in infection, sRNA, such as DenR, could act as 'Eco-markers' to determine whether a microbial soil community is contributing to nitrous oxide production, or whether it is actively removing the greenhouse gas (Moeller et al., 2021). To achieve this, it would be crucial to gain a detailed overview of which *P. denitrificans* sRNAs are expressed under a certain condition in an environmental response, not exclusively in an isolated setting. Furthermore, sRNA could be engineered to target key microbes or key enzymes, such as NosZ, within the denitrification cascade and added to fertilizers to allow a control of emissions from the agriculture industry. Therefore, a further understanding of the sRNA regulatory circuits within the model organism *P. denitrificans* is crucial to continue along this path.

6.3. Novel regulators of denitrification

Transcription regulation is a fundamental biological process, and much effort has been made to identify the important players in this process across all environments to understand how they work together to regulate levels of gene expression. The different niches through which bacteria move during the cycle of their life require rapid and well-adapted responses to changing environmental cues. For denitrifying bacteria such as *P. denitrificans* these factors include the presence of a plethora of environmental factors such as metal ions, denitrification intermediates and oxygen (Moeller et al., 2021). The regulatory network of *P. denitrificans* is therefore composed of a variety of regulatory molecules that work together as well as opposingly to each other to allow for an ideal response to a change in conditions. The small RNAs characterised in this thesis have shown, that there are both small RNAs that enhance N_2O emission rates, such as sRNA 36 as well as small RNAs repressing N_2O emissions, for example DenR and sRNA 18 which strongly suggests that the network of small RNAs is far vaster and more interwoven than initially suspected (Fig. 6.1). Within the putative targets of these small RNAs there is a plethora of novel regulator proteins, transporters as well as sigma factor genes, which are all additional players in regulating a physiological response. Uncovering their exact roles and understanding more of the complexity of regulation that results from promoter architecture is a crucial next step for future research. Having identified the importance of sRNAs in the complex network of interactions involved in an environmental response it has become clear that our picture of regulation is far from complete. This is also the case for denitrification specifically.

The characterisation of a single sRNA, DenR, has introduced a novel transcriptional regulator to the denitrification regulatory network, that is still to be fully characterised to determine its impact on N_2O emissions (Gaimster et al., 2019). This thesis has provided further insights into the interactions of DenR and NirR which are suggested to interact to affect expression levels of the core denitrification enzyme NirS and consequently altering levels of N_2O emissions. This has highlighted the importance of identifying and characterising the targets of other regulatory small RNAs such as sRNA 18, 36 and 39 which have all been shown to play a role in the modulation of nitrous oxide emissions from the model denitrifier *P. denitrificans* by

positively or negatively altering levels of N_2O emitted from an overexpression culture. As their putative targets contain a plethora of uncharacterised proteins of unknown functions as well as functionally uncharacterised transcriptional regulators, there is a strong potential for the identification of further key players in the regulators of denitrification besides the known regulators FnrP, NNR and NarR. Bioinformatics provides powerful tools for the identification of sRNAs targeting genes associated with a regulatory function or genes directly involved in the denitrification response. This can form the basis for further experimental work and can serve as a starting point for the identification of potential regulatory molecules that could be the key for the development of future mitigation strategies.

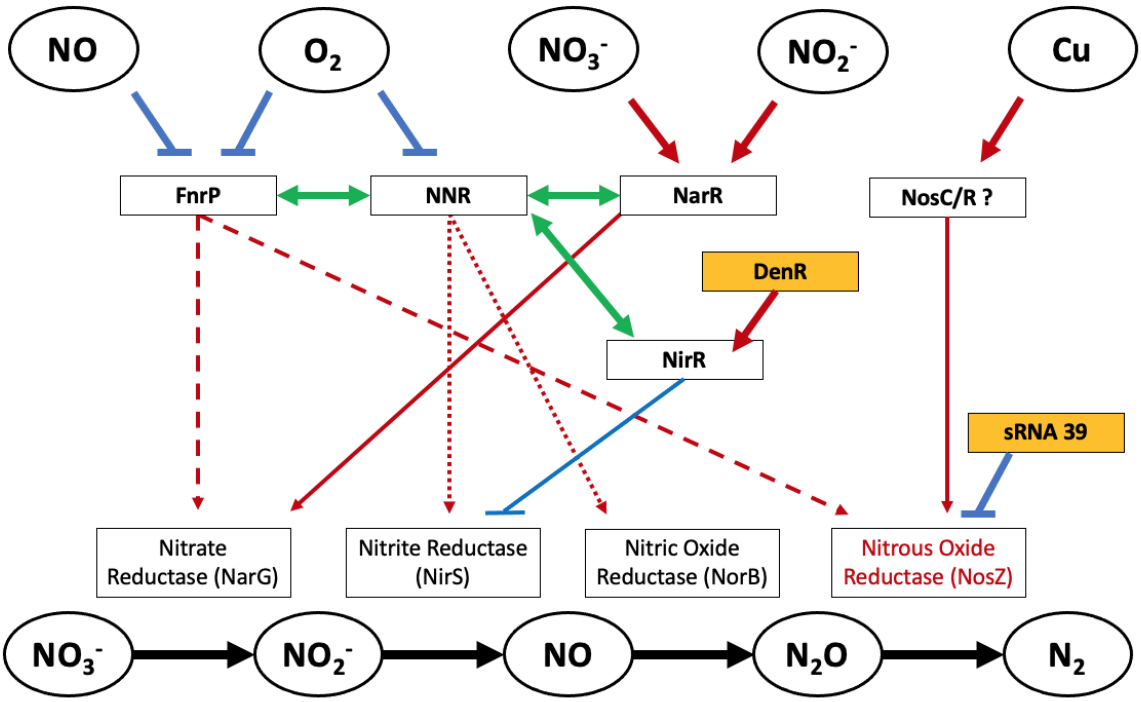


Fig. 6.1: An overview of the known transcriptional and environmental regulators of the denitrification pathway in the model denitrifying bacterium *Paracoccus denitrificans* with the addition of two suggested sRNAs studied in this thesis. The blue and red arrows between the upper layer of the environmental regulatory signals and the layer of the regulatory proteins indicate signalling events (blue indicates an inhibitory effect whilst red indicates activation), while the arrows between the regulatory proteins and the denitrification enzymes indicate regulation of gene expression, blue indicates transcriptional repression. The green arrows between the transcriptional regulators indicate the cross-talk between the regulators which compete with each other to bind upstream of their targets. The yellow boxes indicate sRNAs involved in the regulatory network.

6.4 Transcriptional profiling

Knowledge of the exact position of a transcriptional start site provides evidence for the presence of an active promoter recognised by RNA polymerase with the help of sigma factors. Not all promoters are active at any given time, therefore a TSS map across variable conditions can provide insights into which genes are active and which are inactive. It can also be exploited to predict specific promoter motifs and provide insights into consensus sequences present.

TSS maps have already proven useful in a large number of bacterial species with some TSS maps proving helpful in the understanding of bacterial pathogenicity or survival in stressful environments. For instance, in *Sinorhizobium meliloti* the identification of TSS and promoters on a global scale provided a solid foundation for a range of studies focussing on gene organisation, sigma factors as well as transcription factors and regulatory RNAs, which all remain to be interconnected in *P. denitrificans* (Schlüter et al., 2013). In the natural environment, the availability of copper strongly affects the rate of nitrous oxide emissions, with copper deficient soils contributing to the production of the harmful greenhouse gas, due to a lack of activity of the final denitrification enzyme NosZ. As a result of a dRNA-seq study it has been proven, that gene expression patterns differ between Cu-rich and Cu-limited culture conditions. However, it remains unclear which sigma factor is responsible for switching on transcription under denitrifying conditions.

To further establish *P. denitrificans* as a regulatory model for the complex denitrification processes carried out by bacteria in soil and aquatic environments a global transcription start site map has been generated and has been made accessible as a tool for future research to facilitate further studies on gene regulation and sigma factor usage in *P. denitrificans*. This map is a valuable resource for both future global studies of the *P. denitrificans* transcriptome as well as for in-depth analyses of specific genes and their regulation. Screening for promoter motifs did not reveal the presence of conserved motifs comparable to those known in other well-studied bacteria such as *E. coli*. Consequently, further studies focussed on the sigma factor homologs identified in *P. denitrificans* will be needed.

6.5 Advancing the field of sRNA research outside of model organisms

Although sRNAs have proven to be key players in the major biogeochemical cycles including the nitrogen, carbon and sulfur cycles, our current understanding of their role is far from complete, particularly with respect to the microbial ecology of diverse environments. Manipulating a handful of sRNAs can lead to drastic changes in the response of a lab organism to an experimental condition. This is the case for the DenR, which is capable of decreasing levels of N_2O emissions when overexpressed in *P. denitrificans* (Gaimster et al., 2019) but also for sRNAs 18, 36 and 39 presented in this thesis. This suggests the existence of much larger sRNA networks with unknown sRNAs that regulate the entirety of an environmental response. Uncovering these networks will contribute greatly to our understanding of the environmental stimuli that cause the switch between complete and incomplete denitrification not only in the model organism *P. denitrificans* but also in other environmentally relevant organisms.

The use of high-throughput RNA-seq has led to significant advances in the number of sRNAs discovered. However, most of available information on sRNAs associated with biogeochemical cycles is restricted to a limited number of model organisms. Even within these models limited numbers of sRNAs and their targets have been fully characterised. Traditionally, coding genes were annotated during genome annotation using automated pipelines. Meanwhile, non-coding regions were overlooked (Moeller et al., 2021). The diverse characteristics of sRNAs such as their variable length and secondary structure as well as a lack of sequence conservation across distantly related genomes have made a computational discovery of sRNAs an extremely difficult task (Bossi, 2016). Advances in computational biology such as the development of comparative genomics approaches, RNA-structure analysis and thermodynamic stability-based approaches have contributed to an identification of a vast number of sRNAs (Sridhar & Gunasekaran, 2013) Though the majority have been identified using comparative genomics, transcriptional signal-based approaches are promising in the discovery of novel intergenic sRNAs (Rajendran et al., 2020). The existing breadth of fully sequenced bacterial and archaeal genomes and pre-existing whole transcriptome studies may be a source of future sRNA identification. In the pathogen *Salmonella* Typhimurium SL1344 novel and refined identification methods for sRNA discovery

have uncovered a large number of previously unknown sRNA, suggesting that these short regulatory molecules may even outnumber protein coding genes (Houserova et al., 2021).

Post-identification the detection of sRNA targets is a critical bottleneck in functional sRNA characterisation. Target identification enables the integration of an sRNA into existing regulatory network models responsible for fine-tuning all microbial responses. Experimental target confirmation via genetic screens, knockouts and overexpression often followed by proteomics and microarrays or qRT-PCR are not only time-consuming but also laborious (Backofen & Hess, 2010; Georg et al., 2019). Therefore, the development of reliable computational tools such as sRNARFTarget is highly desirable and has proven in this thesis, that a fast and global target prediction approach can enable the generation of hypothesis for experimental work. Furthermore, it is crucial to extend the study of sRNAs in model organism to a metagenomic scale. In the natural environment, microorganisms do not live in isolation, instead they are found in communities in which they interact to form an environmental response. It is estimated that laboratory culturing techniques provide information on less than 1% of the bacterial diversity in a given environmental sample (Torsvik et al., 1990). Integrative -omics approaches can give insights into genes, RNA molecules, proteins and metabolites present in an entire microbial community that take over a particular function in the environment. Novel sequencing methods based on sRNAs present in a variety of human samples provide insights into the bacterial diversity of patient samples. This is a promising tool for the analysis of the entirety of the sRNA content as well as the microbial sample profile (Mjelle et al., 2020).

Metatranscriptomic datasets from the ocean have revealed the abundance of sRNA molecules present in a microbial community and have highlighted the importance of bacterial sRNAs in processes such as carbon metabolism and nutrient acquisition (Shi et al., 2009). These sRNAs were identified as part of a pyrosequencing study of total RNA extracted from the Hawaiian ocean. A large fraction of sequences identified in this study shared no homology with known proteins and many were identified to be novel sRNA sequences. A second study identified an abundance of sRNAs from an extremophilic microbial community in the Atacama Desert (Gelsinger et al., 2020). Predicted targets for these sRNAs were genes involved in osmotic adjustments to major rain events in the desert as well as nutrient acquisition which

further underpins the importance of sRNAs in the community stress response in the natural environment. Community studies like these two are important contributors to uncovering the vast number of sRNAs *in situ* that are of environmental importance. Therefore, metatranscriptomic studies of denitrifying communities in soil and aquatic environments around model organisms such as *P. denitrificans* will be required to gain an insight into sRNA activity in the environment and identify targets for future mitigation strategies.

6.7 Concluding remarks

The work presented in this thesis has expanded our understanding of the importance of sRNAs in regulating denitrification in the model denitrifier *P. denitrificans.* In particular, the results presented for sRNAs 18, 36 and 39 have shown that the previously characterised sRNA DenR is not the only sRNA in *P. denitrificans* to have an impact on N_2O emissions. Their targets include a range of uncharacterised proteins which may well be important players in the denitrification regulatory network. Furthermore, many potential novel regulators such as NirR remain to be integrated into the far from complete picture of this regulatory network.

There are many avenues of future work resulting from this research and specific experiments are described in the 'Discussion' sections of each results chapter. Further sRNAs and their targets remain to be characterised with the computational characterisation carried out in this thesis forming the foundation of experimental work and the study of the P*. denitrificans* transcriptome should be taken further to provide insights into sigma factor usage during denitrification. Furthermore, the study of sRNAs implicated in the biogeochemical nitrogen cycle and the denitrification response specifically should be taken to a metagenomic level to elucidate the key players in this response. This knowledge could provide targets for future mitigation strategies to combat excessive nitrous oxide emissions from agricultural soils. Given the increasing temperatures on Earth and the pressing need for fertilisation to provide sufficient crops to sustain the growing world population, there is a desperate need for novel approaches. There is still a lot more to be discovered and understood before we can fully

understand the regulatory switches underpinning denitrification, but first steps have been made to facilitate future projects.

7 References

Ao W, Gaudet J, Kent WJ, Muttumu S, Mango SE. Environmentally induced foregut remodeling by PHA-4/FoxA and DAF-12/NHR. Science. 2004;305(5691):1743–6

Babitzke, P., Romeo, T. (2007). CsrB sRNA family: sequestration of RNA binding regulatory proteins. Curr. Opin. Microbiol. 10: 156-163. doi: 10.1016/j.mib.2007.03.007.

Backofen, R., Hess, W. (2010) Computational prediction of sRNAs and their targets in bacteria. RNA Biol. 7, 1: 33-42. doi: 10.4161/rna.7.1.10655

Bandyra KJ, Said N, Pfeiffer V, Górna MW, Vogel J, Luisi BF. The seed region of a small RNA drives the controlled destruction of the target mRNA by the endoribonuclease RNase E. Mol Cell. 2012 Sep 28;47(6):943-53. doi: 10.1016/j.molcel.2012.07.015. Epub 2012 Aug 16. PMID: 22902561; PMCID: PMC3469820.

Bardill, J., Zhao, X., Hammer, B. (2011) the Vibrio cholerae quorum sensing response is mediated by Hfq-dependent sRNA/mRNA base pairing interactions. Mol. Microbiol. 80: 1381-1394. doi: 10.1111/j.1365-2958.2011.07655

Battesti, A., Majdalani, N., Gottesman, S. (2011). The RpoS-mediated general stress response in *Escherichia coli.* FEMS Microbiol. Rev. 34: 866-882. doi: 10.1146/annurev-micro-090110-102946.

Beisel, C., Storz, G. (2010) Base pairing small RNAs and their role in global regulatory networks. FEMS Microbiol. Rev. 34, 5: 866-882. doi: 10.111/j.1574-6976.2010.00241.x

Bergaust, L., Mao, Y., Bakken, L.R., Frostegård, Å. (2010). Denitrification response patterns during the transition to anoxic respiration and posttranscriptional effect of suboptimal pH on nitrogen oxide reductase in *Paracoccus denitrificans*. Appl. Environ. Microbiol. 76: 6837-6896. doi: 10.1128/AEM.00608-10.

Bergaust L., vanSpanning R., Frostegård Å., Bakken L. R. In press Expression of nitrous oxide reductase in *Paracoccus denitrificans* is regulated by oxygen and nitric oxide through FnrP and NNR. *Microbiology* (doi:10.1099/mic.0.054148-0)

Bernhard, A. (2010). The Nitrogen Cycle: Processes, Players and Human Impact. Nature Education, Knowledge Project. 3, 10: 25.

Bertero M, Rothery A, Palak M, Hou C, Lim D, Blasco F, Weiner J, Strynadka N (2003) *Insights into the respiratory electron transfer pathway from the structure of nitrate reductase A;* Nat Struct Mol Bio, Vol. 10 (9), 681-687

Bishop, C.M.; Nasrabadi, N.M. *Bishop, Pattern Recognition and Machine Learning (Information Science and Statistics)*; Springer: Berlin/Heidelberg, Germany, 2006

Bos, J., Duverger, Y., Thouvenot, B., Chiaruttini, C., Branlant, C., Springer, M., et al. (2013). The sRNA RhyB regulates the synthesis of the *Escherichia* coli methionine sulfoxide reductase MrsB but not MrsA. PLoS One. 8, 5: e63647. doi: 10.1371/journal.pone.0063647.

Bose, A., Pritchett, M., Metcalf, W. (2008). Genetic analysis of the methanol- and methylamine-specific methyltransferase 2 genes of *Methanosarcina acetivorans* C2A. J. Bacteriol. 190: 4017-4026. doi: 10.1128/JB.00117-08.

Bossi, L., and Figuera-Bossi, N. (2016). Competing endogenous RNAs: a target-centric view of small RNA regulation in bacteria. Nat. Rev. Microbiol. 14, 12: 775-784. doi: 10.1038/nrmicro.2016.129.

Bouchal, P., Struharova, I., Budinska, E., Sedo, O., Vhylidalova, T., Zdrahal, Z., et al. (2010). Unraveling an RNR based regulatory circuit in *Paracoccus denitrificans* using a proteomics-based approach. Biochim. Biophys. Acta. 1804, 6: 1350-1368. doi: 10.1016/j.bbapap.2010.01.016.

Boudes, M., Lazar, N., Graille, M., Durand, D., Gaidenko, T., Steward, V., Van Tilbeurgh, H. (2012) The structure of the NasR transcription anititerminator reveals a one-component system with a NIT nitrate receptor coupled to an ANTAR RNA-binding effector. Mol. Microbiol. 85: 431-444. doi: 10.1111/j.1365-2958.2012.08111.x

Boudvillain, M., Figueroa-Bossi, N., Bossi, L. (2013). Terminator still moving forward: expanding roles for rho factor. Curr. Opin. Microbiol. 16: 118-124. doi: 10.1016/j.mib.2012.12.003.

Boutet E, Djerroud S, Perreault J. Small RNAs beyond Model Organisms: Have We Only Scratched the Surface? Int J Mol Sci. 2022 Apr 18;23(8):4448. doi: 10.3390/ijms23084448. PMID: 35457265; PMCID: PMC9029176

Bouvier, M., Sharma, C., Mika, F., Nierhaus, K., Vogel, J. (2008). Small RNA binding to 5‘ mRNA coding region inhibits translational initiation. Mol. Cell. 32, 6: 827-837. doi: 10.1016/j.molcel.2008.10.027.

Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem. 1976 May 7;72:248-54. doi: 10.1006/abio.1976.9999. PMID: 942051.

Breiman Leo . Random Forests. Machine Learning, 45(1), 2001.

Brennan RG, Link TM 2007. Hfq structure, function and ligand binding. *Curr Opin Microbiol*10: 125–133

Brennan, C. and Garrett, W. (2019) Fusobacterium nucleatum – Symbiont, opportunist and oncobacterium. Nat. Rev. Microbiol. 17: 156-166. doi: 10.1038/s41579-018-0129-6

Brusseau, M. (2019). Ecosystems and Ecosystem Services. Environmental and Pollution Science. Acad Press. Third Edition: 89-102.

Buddeweg, A., Sharma, K., Urlaub, H., Schmitz, R. (2018). sRNA41 affects ribosome binding sites within polycistronic mRNAs in *Methanosarcina mazei* Gö1. Mol. Microbiol. 107: 595-609. doi: 10.1111/mmi.13900.

Bui Quang Minh, Heiko A. Schmidt, Olga Chernomor, Dominik Schrempf, Michael D. Woodhams, Arndt von Haeseler, and Robert Lanfear (2020) IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era. Mol. Biol. Evol., in press. https://doi.org/10.1093/molbev/msaa015

Burgess RR, Travers AA, Dunn JJ & Bautz EKF Factor Stimulating Transcription by RNA Polymerase. *Nature* 221, 43–46 (1969).

Burgess RR, Sigma Factors, Editor(s): Sydney Brenner, Jefferey H. Miller, Encyclopedia of Genetics, Academic Press, 2001, Pages 1831-1834, ISBN 9780122270802, https://doi.org/10.1006/rwgn.2001.1192.

Burns, S., Bullock, H., Smith, C., Huang, Q., Whitman, W., Moran, M. (2016). Small RNAs expressed during dimethylsulfoniopropionate degradation by a model marine bacterium. Environ. Microbiol. Rep. 8, 5: 763-773. doi: 10.1111/1758-2229.12437.

Carver T, Harris SR, Berriman M, Parkhill J and McQuillan JA. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. *Bioinformatics (Oxford, England)* 2011;28;4;464-9

Cavanagh, A., and Wassarman, K. (2014). 6S RNA, a global regulator of transcription in *Escherichia coli, Bacillus subtilis* and beyond. Annu. Rev. Microbiol. 68: 45-60. doi: 10.1146/annurev-micro-092611-150135.

Ceizel Borella, G., Lagares, A., Valverde, C. (2016). Expression of the *Sinorhizobium meliloti* small RNA gene mmgR is controlled by the nitrogen source. FEMS. Microbiol. Lett. 363, 9: fnw069. doi: 10.1093/femsle/fnw069

Chai, W. and Stuart, V. (1999) RNA sequence requirements for NasR-mediated, nitrate-responsive transcription antitermination of the Klebsiella oxytoca M5al nasF operon leader. J. Mol. Biol. 292: 203-216. doi: 10.1006/jmbi.1999.3084

Chan, T. (2011) Vegetable-borne nitrate and nitrite and the risk of methaemoglobinaemia. Toxicol. Lett. 200: 107-108. doi: 10.1016/j.toxlet.2010.11.002.

Chao, Y., Li, L., Girodat, D., Forstner, K., Said, N., Corcoran, C., et al. (2017). In vivo cleavage map illuminates the central role of RNase E in coding and non-coding RNA pathway. Mol. Cell. 65: 39-51. doi: 10.1016/j.molcel.2016.11.002.

Chao Y, Vogel J. The role of Hfq in bacterial pathogens. *Curr Opin Microbiol.* 2010;13(1):24–33. doi: 10.1016/j.mib.2010.01.001.

Chen, H., Dutta, T., Deutscher, M. (2016). Growth-phase dependent variation of RNase BN/Z affects sRNAs: regulation of 6S RNA. J. Biol. Chem. 291: 26435-26442. doi: 10.1074/jbc.M116.757450.

Chen, S., Perathoner, S., Ampelli, C., Centi, G. (2019). "Electrochemical Dinitrogen Activation: To Find a Sustainable Way to Produce Ammonia" Studies in Surface Science and Catalysis, Vol. 178, Elsevier: 31-46.

Christiansen JK, Larsen MH, Ingmer H, Søgaard-Andersen L, Kallipolitis BH. The RNA-binding protein Hfq of Listeria monocytogenes: role in stress tolerance and virulence. J Bacteriol. 2004 Jun;186(11):3355-62. doi: 10.1128/JB.186.11.3355-3362.2004. PMID: 15150220; PMCID: PMC415768.

Conthe, M., Wittorf, L., Gjis Kuenen, J., Kleerebezem, R., Hallin, S., Van Loosdrecht, M. (2018) Growth yield and selection of *nosZ* clade II types in a continuous enrichment culture of N_2O respiring bacteria. Env. Microbiol. Rep. 10, 3: 239-244. doi: 10.1111/1758-2229.12630.

Core LJ, Waterfall JJ, Lis JT. Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. *Science.* 2008;322:1845–1848

Crack, J.C., Hutchings, M.I., Thomson, A.J., Le Brun, N.E. (2016). Biochemical properties of *Paracoccus denitrificans* FnrP: reactions with molecular oxygen and nitric oxide J. Biol. Inorg. Chem. 21, 1: 71-82. doi: 10.1007/s00775-015-1326-7

Crick F. Central Dogma of Molecular Biology. *Nature* 227, 561–563 (1970).

Curson, A., Williams, B., Pinchbeck, B., Sims, L., Martinez, A., Rivera, P., et al. (2018) DSYB catalyses the key step of dimethylsulfoniopropionate biosynthesis in many phytoplankton. Nat. Microbiol. 3: 430-439. doi: 10.1038/s41564-018-0119-5

Cuthbertson L, Nodwell JR. The TetR family of regulators. Microbiol Mol Biol Rev. 2013 Sep;77(3):440-75. doi: 10.1128/MMBR.00018-13. PMID: 24006471; PMCID: PMC3811609.

Darfeuille, F., Unoson, C., Vogel, J., Wagner, E. (2007). An antisense RNA inhibits translation by competing with standby ribosome. Mol. Cell. 26: 381-392. doi: 10.1016/j.molcel.2007.04.003.

De Lay, N., and Gottesman, S. (2009). The Crp-activated small noncoding regulatory RNA CyaR (RyeE) links nutritional status to group behavior. J. Bacteriol. 191: 461-476. doi: 10.1128/JB.01157-08

De Lay, N., and Gottesman, S. (2011). Role of polynucleotide phosphorylase in sRNA function in *Escherichia coli*. RNA. 17, 6: 1172-1189. doi: 10.1261/rna.2531211.

Diallo I, Provost P. RNA-Sequencing Analyses of Small Bacterial RNAs and their Emergence as Virulence Factors in Host-Pathogen Interactions. Int J Mol Sci. 2020 Feb 27;21(5):1627. doi: 10.3390/ijms21051627. PMID: 32120885; PMCID: PMC7084465.

Diep Thi Hoang, Olga Chernomor, Arndt von Haeseler, Bui Quang Minh, and Le Sy Vinh (2018) UFBoot2: Improving the ultrafast bootstrap approximation. Mol. Biol. Evol., 35:518–522. https://doi.org/10.1093/molbev/msx281

Dixon, R., Buck, M., Drummond, M., Hawkes, T., Khan, H., MacFarlane, S., et al. (1986) Regulation of the nitrogen fixation genes in *Klebsiella pneumoniae*: Implications for genetic manipulation. Plant and Soil. 90: 225-233. Doi: 10.1007/BF02277399

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

Domeignoz-Horta, L., Putz, M., Spor, A., Bru, D., Breuil, M., Hallin, S., Philippoz, L. (2016) Non-denitrifying nitrous oxide-reducing bacteria – An effective N_2O sink in soil. Soil Biol Biochem. 103: 376-379. Doi: 10.1016/j.soilbio.2016.09.010

Durand, S., Tomasini, A., Braun, F., Condon C, Romby P (2015) sRNA and mRNA turnover in Gram-positive bacteria*.* FEMS. Microbiol. Rev. 39, 3: 316-330. doi: 10.1093/femsre/fuv007.

Dutta, T., and Srivastava, S. (2018). Small RNA-mediated regulation in bacteria: A growing palette of diverse mechanisms*.* Gene. 20, 656: 60-72. doi: 10.1016/j.gene.2018.02.068.

Ehlers, C., Jäger, D., Schmitz, R. (2011) Establishing a Markless Genetic Exchange System for *Methanosarcina mazei* Strain Gö1 for Constructing Chromosomal Mutants of Small RNA Genes. Archaea. 2011: 439608. Dooi: 10.1155/2011/439608

Elsen, S., Swem, L., Swem, D., Bauer, C (2004) RegB/RegA, a Highly Conserved Redox-Responding Global Two-Component Regulatory System. Microbiol Mol Biol Rev. 68, 2: 263-279. doi: 10.1128/MMBR.68.2.263-279.2004

Falkowski, P. G. (1997). Evolution of the nitrogen cycle and its influence on the biological sequestration of CO_2 in the ocean. *Nature* 387, 272–275. doi: 10.1038/387272a0

Feklistov A, Sharon B, Darst S, Gross C (2014) Bacterial Sigma Factors: A Historical, Structural, and Genomic Perspective; Annu. Rev. Microbio. Vol. 68: 357-376; doi: 10.1146

Felgate, H., Giannopoulos, G., Sullivan, M., Gates, A., Clarke, T., Baggs, E., et al. (2012). The impact of copper, nitrate and carbon status on the emission of nitrous oxide by two species of bacteria with biochemically distinct denitrification pathways*.* Environ. Microbiol. 14, 7: 1788-1800. doi: 10.1111/j.1462-2920.2012.02789.x.

Fernández L, Breidenstein EBM, Taylor PK, Bains M, de la Fuente-Núñez C, Fang Y, Foster LJ, Hancock REW. Interconnection of post-transcriptional regulation: The RNA-binding protein Hfq is a novel target of the Lon protease in Pseudomonas aeruginosa. Sci Rep. 2016 May 27;6:26811. doi: 10.1038/srep26811. PMID: 27229357; PMCID: PMC4882532.

Fischer, S., Benz, J., Späth, B., Maier, L.K., Straub, J., Granzow, M., et al. (2010). The archaeal Lsm protein binds to small RNAs. J. Biol. Chem. 285, 45: 34429-34438. doi:10.1074/jbc.M110.118950

Fozo EM, Hemm MR, Storz G 2008. Small toxic proteins and the antisense RNAs that repress them. *Microbiol Mol Biol Rev* 72: 579–589

Franze de Fernandez, M.T., Eoyang, L., August, J.T. (1968). Factor fraction required for the synthesis of bacteriophage Qbeta-RNA. Nature. 219, 5154: 588-590. doi: 10.1038/219588a0.

Frohlich, K., and Vogel, J. (2009). Activation of Gene expression by small RNA. Curr. Opin. Microbiol. 12: 674-682. doi: 10.1016/j.mib.2009.09.009.

Gaimster, H., Alston, M., Richardson, D., Gates, A., Rowley, G. (2018). Transcriptional and Environmental control of Bacterial Denitrification and N2O Emissions. FEMS. Microbiol. Lett., 365, 5. doi: 10.1093/femsle/fnx277.

Gaimster, H., Chalklen, L., Alston, M., Munnoch, J., Richardson, D., Gates, A., Rowley, G. (2016). Genome-wide discovery of putative sRNAs in *Paracoccus denitrificans* expressed under nitrous oxide emitting conditions. Front. Microbiol., 14, 7: 1806. doi: 10.3389/fmicb.2016.01806.

Gaimster, H., Hews, C., Griffiths, R., Soriano-Laguna, M., Alston, M., Richardson, D., Gates, A., Rowley, G. (2019). A central small RNA regulatory circuit controlling Bacterial Denitrification and N_2O emissions. mBio. 10, 4: e01165-19. doi: 10.1128/mBio.01165-19.

Galloway, J., and Cowling, E. (2002). Reactive nitrogen and the world: 200 years of change. Ambio. 31, 2: 64-71. doi: 10.1579/0044-7447-31.2.64.

Gelsinger, D., Uritskiy, G., Reddy, R., Munn, A., Farney, K., DiRuggiero, J. (2020) Regultory Noncoding Small RNAs are diverse and abundant in an extremophilic microbial community. ASM. 5. 1. doi: 10.1128/mSystems.00584-19

Georg, J., Lalaouna, D., Hou, S., Lott, S., Caldelari, I., Marzi, S., Hess, W., Romby, P. (2019). The power of cooperation: Experimental and computational approaches in the functional characterisation of bacterial sRNAs. Mol. Microbiol. 113, 3: 603-612. doi: 10.1111/mmi.14420.

Gerdes K, Wagner EG 2007. RNA antitoxins. *Curr Opin Microbiol* 10: 117–124

Giangrossi, M., Prosseda, G., Tran, C., Brandi, A., Colonna, B., Falconi, M. (2010). A novel antisense RNA regulates at transcriptional level the virulence gene icsA of *Shigella flexneri.* Nucleic. Acids. Res. 38, 10: 3362-3375. doi: 10.1093/nar/gkq025.

Giannopoulos, G., Sullivan, M., Hartop, K., Rowley, G., Gates, J., Watmough, N., et al. (2017) Tuning the modular Paracoccus denitrificans respirome to adapt from aerobic respiration to anaerobic denitrification. Environ. Microbiol. 19. 12: 4953-4964. doi: 10.1111/1462-2920.13974

Giannopoulos G (2014) *The regulation of denitrification in P. denitrificans;* Thesis for the degree of Doctor of Philosophy; School of Biological Sciences, University of East Anglia, UK

Giardina G, Rinaldo S, Johnson K, Di Matteo A, Brunori M, Cutruzzolà F; NO sensing in Pseudomonas aeruginosa: Structure of the Transcriptional Regulator DNR, Journal of Molecular Biology, Vol. 378, 5, 2008, pp. 1002-1015, ISSN 0022-2836, https://doi.org/10.1016/j.jmb.2008.03.013.

Gierga, G., Voss, B., Hess, W. (2012). Non-coding RNAs in marine *Synechococcus* and their regulation under environmentally relevant stress conditions. ISME. J. 6, 8: 1544-1557. doi: 10.1038/ismej.2011.215.

Grant JR, Enns E, Marinier E, Mandal A, Herman EK, Chen C, Graham M, Van Domselaar G, and Stothard P. Proksee: in-depth characterization and visualization of bacterial genomes Nucleic Acids Research, 2023, gkad326, https://doi.org/10.1093/nar/gkad326

Goldblatt, C., Claire, M., Lenton, T. *et al.* Nitrogen-enhanced greenhouse warming on early Earth. *Nature Geosci* 2, 891–896 (2009). https://doi.org/10.1038/ngeo692

Gopalkrishnan, S.; Ross, W.; Chen, A.; Gourse, R. TraR directly regulates transcription initiation by mimicking the combined effects of the global regulators DksA and ppGpp (2017) 114 (28) E5539-E5548. https://doi.org/10.1073/pnas.1704105114

Gottesman, G. (2005). Micros for microbes: noncoding regulatory RNAs in bacteria.Trends Genet. 21, 7: 399-404. doi: 10.1016/j.tig.2005.05.008.

Green H. A., Donohue T. J. (2006). Activity of *Rhodobacter sphaeroides* RpoHII, a second member of the heat shock sigma factor family. *J. Bacteriol.* 188, 5712–5721. 10.1128/JB.00405-06

Guerrero, M.G., Vega, J.M., Losada, M. (1981). The assimilatory nitrate-reducing system and its regulation. Annu. Rev. Plant. Physiol. 32: 169-204. doi: 10.1146/annurev.pp.32.060181.001125

Halbleib, C., Ludden, P (2000) Regulation of Biological Nitrogen Fixation. The Journal of Nutrition 130, 5: 1081-1084. doi: 10.1093/jn/130.5.1081

Hallin, S., Philippot, L., Löffler, F., Sanford, R., Jones, C. (2018) Genomics and Ecology of Novel N_2O-reducing Microorganisms. Trends Microbiol. 16. 1: 43-55. doi: 10.1016/j.tim.2017.07.003.

Hao, Y., Updegrove, T.B., Livingston, N.N., Storz, G. (2016). Protection against deleterious nitrogen compounds: role of sigma(S)-dependent small RNAs encoded adjacent to sdiA. Nucleic. Acids. Res. 44, 14: 6935-6948. doi: 10.1093/nar/gkw404.

Härtig E, Zumft WG. The requirement of RpoN (sigma factor ς54) in denitrification by *Pseudomonas stutzeri* is indirect and restricted to the reduction of nitrite and nitric oxide. *Appl Environ Microb* 1998; 64:3092–5.

Hartman CE, Samuels DJ, Karls AC. Modulating *Salmonella Typhimurium's* response to a changing environment through bacterial enhancer-binding proteins and the RpoN regulon. Front Mol Biosci. 2016;3:41.

Haydon, D. J. and Guest, J. R. (1991) A new family of bacterial regulatory proteins. *FEMS Microbiol. Lett*. 63, 291–295.

Heilmann, B., Hakkila, K., Georg. J., Tyystjärvi, T., Hess, W.R., Axmann, I.M., Dienst, D. (2017) 6S RNA plays a role in recovery from nitrogen depletion in *Synechocystis* sp. PCC 6803. BMC Microbiol. 17, 229. doi: hhtps://doi.ord/10.1186/s12866-017-1137-9

Herrero, A., Muro-pastor, A., Flores, E. (2001) Nitrogen control in cyanobacteria. J Bacteriol. 183, 2: 411-425. doi: 10.1128/jb.183.2.411-425.2001.

Hoch, G., Schneider, K., Burris, R. (1960) Hydrogen evolution and exchange, and conversion of N2O to N2 by soybean root nodules. Biochim. Biophys. Acta. 37. 2: 273-279. doi: 10.1016/0006-3002(60)90234-1

Hör J, Matera G, Vogel J, Gottesman S, Storz G. Trans-Acting Small RNAs and Their Effects on Gene Expression in *Escherichia coli* and *Salmonella enterica*. EcoSal Plus. 2020

Houserova, D., Dahmer, D., Amin, S., Crucello, A., King, M., Barnhill, E., Huang, Y., Zambrano, M., Dean, M., Aria, K., Spector, M., Borchert, G. (2021) Characterisation of 475 novel, putative small RNAs (sRNAs) in Carbon-starved *Salmonella enterica* serovar Typhimurium. bioRxiv. preprint. doi: 10.1101/2021.01.11.426214

Hu, J., Wang, Q. (2018). Regulatory sRNAs in Cyanobacteria. Front. Microbiol. 9, 2399. doi: 10.3389/fmicb.2018.02399.

IFA (2009). Energy efficiency and CO2 emissions in ammonia production. Available at: https://www.fertilizer.org/images/Library_Downloads/2009_IFA_energy_efficiency.pdf (Accessed January 10, 2024).

Ionescu, D., Voss, B., Oren, A., Hess, W., Muro-Pastor, M. (2010). Heterocyst-Specific Transcription of NsiR1, a Non-coding RNA encoded in a tandem array of direct repeats in cyanobacteria. J. Mol. Biol. 398, 2: 177-188. doi: 10.1016/j.jmb.2010.03.010.

Ishikawa H, Otaka H, Maki K, Morita T, Aiba H. The functional Hfq-binding module of bacterial sRNAs consists of a double or single hairpin preceded by a U-rich sequence and followed by a 3' poly(U) tail. RNA. 2012 May;18(5):1062-74. doi: 10.1261/rna.031575.111. Epub 2012 Mar 27. PMID: 22454537; PMCID: PMC3334693

Stocker, T., Qin, D., Plattner, G., Tignor, M., Allen, S., Boschung, J., et al. IPCC (2013) Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York. NY. USA. 1535 pp.

Jäger, D., Pernitzsch, S., Richter, A., Backofen, R., Sharma, C., Schmitz, R. (2012). An archaeal sRNA targeting *cis*- and *trans*- encoded mRNAs via two distinct domains. Nucleic. Acids. Res. 40, 21: 10964-10979. doi: 10.1093/nar/gks847.

Jäger, D., Sharma, C., Thomsen, J., Ehlers, C., Vogel, J., Schmitz. (2009) Deep Sequencing analysis of the *Methanosarcina mazei* Gö1 transcriptome in response to nitrogen availability. PNAS. 106, 51: 21878-21882. Doi: 10.1073/pnas.0909051106

Jain, D. (2015), Allosteric control of transcription in GntR family of transcription regulators: A structural overview. IUBMB Life, 67: 556-563. https://doi.org/10.1002/iub.1401

Jensen, F.B. (2003). Nitrite disrupts multiple physiological functions in aquatic animals. Comp. Biochem. Physiol. 135A: 9-24. doi: 10.1016/s1095-6433(02)00323-9.

Jones, C., Graf, D., Bru, D., Philippot, L., Hallin, S. (2013) The unaccounted yet abundant nitrous oxide-reducing microbial community: a potential nitrous oxide sink. ISME Journal 7: 417-426. Doi: 10.1038/ismej.2012.125

Jung, Y.S., and Kwon, Y.M. (2008). Small RNA ArrF regulates the expression of *sodB* and *feSII* genes in *Azotobacter vinelandii*. Curr. Microbiol. 57: 593-597. doi: 10.1007/s00284-008-9248-z

Kalyaanamoorthy S., Bui Quang Minh, Thomas KF Wong, Arndt von Haeseler, and Lars S Jermiin (2017) ModelFinder: Fast model selection for accurate phylogenetic estimates. Nature Methods, 14:587–589.https://doi.org/10.1038/nmeth.4285

Kapoor, V., Li, X., Elk, M., Chandran, K., Impellitteri, C., Santo Domingo, J. (2015) Impact of heavy metals on transcriptional and physiological activity of nitrifying bacteria. Environ. Sci. Technol. 49: 13454-13462. doi: 10.1021/asc.est.5b02748

Kelley LA *et al.* The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Protocols* 10, 845-858 (2015)

Kery, Mary Beth ; Feldman, Monica; Tjaden, Brian; Livny, Jonathan. TargetRNA2: identifying targets of small regulatory RNAs in bacteria. Nucleic Acids Research, 42(W1):W124–W129, 04 2014.

Keseler I. M., Collado-Vides J., Santos-Zavaleta A., Peralta-Gil M., Gama-Castro S., Muniz-Rascado L., et al.. (2011). Ecocyc: A Comprehensive Database of Escherichia Coli Biology. *Nucleic Acids Res.* 39, D583–D590. 10.1093/nar/gkq1143

Kettles, N., Kopriva, S., Malin, G. (2014) Insights intot he Regulation of DMSP Synthesis in the Diatom Thalassiosira pseudonana through APR Activity, Proteomics and Gene Expression Analyses on Cells Acclimating to Changes in Salinity, Light and Nitrogen. PLoS One. 9. 4. doi: 10.1371/journal.pone.0094795

Klähn, S., Schaal, C., Georg, J., Baumgartner, D., Knippen, G., Hagemann, M., et al. (2015). The sRNA NsiR4 is involved in nitrogen assimilation control in cyanobacteria by targeting

glutamine synthetase inactivating factor IF7. Proc. Natl. Acad. Sci. USA. 112, 45: E6243-52. doi: 10.1073/pnas.1508412112.

Kopf, M., and Hess,W. (2015). Regulatory RNAs in photosyntheric cyanobacteria. FEMS. Microbiol. Rev. 39, 3: 301-315. doi: 10.1093/femsre/fuv017.

Krätzer, C., Carini, P., Deppenmaier, U., Kratzer, C. (2009). Transcriptional profiling of methyltransferase genes during growth of *Methanosarcina mazei* on trimethylamine. J.Bacteriol. 191: 5108-5115. doi: 10.1128/JB.00420-09.

Kumar, K., Chakraborty, A., Chakrabarti, S. (2020) PresRAT: A server for identification of bacterial small-RNA sequences and their targets with probable binding region. RNA Biol. doi: 10.1080/15476286.2020.1836455

Liu JM, Camilli A: Discovery of bacterial sRNAs by high-throughput sequencing. Methods in molecular biology (Clifton, N.J.). 2011, 733: 63-79. 10.1007/978-1-61779-089-8_5.

Livny J, Waldor MK 2007. Identification of small RNAs in diverse bacterial species. *Curr Opin Microbiol* 10: 96–101

López-Leal G, Cevallos MA, Castillo-Ramírez S. Evolution of a Sigma Factor: An All-In-One of Gene Duplication, Horizontal Gene Transfer, Purifying Selection, and Promoter Differentiation. Front Microbiol. 2016 Apr 25;7:581. doi: 10.3389/fmicb.2016.00581. PMID: 27199915; PMCID: PMC4843759.

Luque-Almagro, V., Gates, A., Moreno-Vivian, C., Ferguson, S., Richardson, D., Roldan, M. (2011). Bacterial nitrate assimilation: gene distribution and regulation. Biochem. Soc. Trans. 39, 6: 1838-1843. doi: 10.1042/BST20110688.

Luque-Almagro, V., Manso, I., Sullivan, M., Rowley, G., Ferguson, S., Moreno-Vivian, C., et al (2017). Transcriptional and translational adaptation to aerobic nitrate anabolism in the denitrifier *Paracoccus denitrificans*. Biochem. J. 474, 11: 1769-1787. doi: 10.1042/BCJ20170115

Lycus, P., Soriano-Laguna, M., Kjos, M., Richardson, D., Gates, A., Milligan, D., et al. (2018). A bet-hedging strategy for denitrifying bacteria curtails their release of N_2O. Proc. Natl. Acad. Sci. USA. 115, 46: 11820-11825. doi: 10.1073/pnas.1805000115

Maddocks SE, Oyston PCF. Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. Microbiology (Reading). 2008 Dec;154(Pt 12):3609-3623. doi: 10.1099/mic.0.2008/022772-0. PMID: 19047729.

Madsen, E. (2011). Microorganisms and their roles in fundamental biogeochemical cycles. Curr. Opin. Biotechnol. 22, 3: 456-464. doi: 10.1016/j.copbio.2011.01.008.

Mar; 9(1):10.1128/ecosalplus.ESP-0030-2019. doi: 10.1128/ecosalplus.ESP-0030-2019. PMID: 32213244; PMCID: PMC7112153.

Märtens, B., Bezerra, G.A., Kreuter, M.J., Grishkovskaya, I., Manica, A., Arkhipova, V., et al. (2015). The Heptameric SmAP1 and SmAP2 proteins of the crenarchaeon *Sulfolobus solfataricus* bind to common and distinct RNA targets. Life. 5: 1264-1281. doi: 10.3390/life5021264

Märtens, B., Sharma, K., Urlaub, H., Bläsi, U. (2017) The SmAP2 RNA binding motif in the 3'UTR affects mRNA stability in the crenarchaeum *Sulfolobus solfataricus*. Nucleid Acids Res. 45, 15: 8957-8967. doi: 10.1093/nar/gkx581

Martinez-Argudo, I., Little, R., Shearer, N., Johnson, P., Dixon, R. (2004). The NifL-NifA system: a multidomain transcriptional regulatory complex that integrates environmental signals. J. Bacteriol. 186: 601-610. doi: 10.1128/jb.186.3.601-610.2004.

Martinez-Salazar J. M., Salazar E., Encarnacion S., Ramirez-Romero M. A., Rivera J. (2009b). Role of the extracytoplasmic function sigma factor RpoE4 in oxidative and osmotic stress responses in *Rhizobium etli*. *J. Bacteriol*. 191, 4122–4132. 10.1128/JB.01626-08

Martinez-Salazar J. M., Sandoval-Calderon M., Guo X., Castillo-Ramirez S., Reyes A., Loza M. G., et al.. (2009a). The *Rhizobium etli* RpoH1 and RpoH2 sigma factors are involved in different stress responses. *Microbiology* 155, 386–397. 10.1099/mic.0.021428-0

Masse, E., and Gottesman, S. (2002). A small RNA regulates the expression of genes involved in iron metabolism in *Escherichia coli.* Proc. Natl. Acad. Sci. USA. 99, 7: 4620-4625. doi: 10.1073/pnas.032066599.

McCarren, J., Becker, J., Repeta, D., Shi, Y., Young, C., Malmstrom, R., et al. (2010). Microbial community transcriptomes reveal microbes and metabolic pathways associated with dissolved organic matter turnover in the sea. Proc. Natl. Acad. Sci. USA. 107, 38: 16420-16427. doi: 10.1073/pnas.1010732107.

McCullen, C., Benhammou, J., Majdalani, N., Gottesman, S. (2010). Mechanism of positive regulation by DsrA and RprA small noncoding RNAs: Pairing Increases Translation and Protects rpoS mRNA from Degradation. J. Bacteriol. 192, 21: 5559-5571. doi: 10.1128/JB.00464-10.

Michaux, C., Vernuil, N., Hartke, A., Giard, J. (2014) Physiological roles of small RNA molecules. Microbiol. 160, 6: 1007-1019. doi: 10.1099/mic.0.076208-0

Mjelle, R., Aass, K., Sjursen, W., Hofsli, E., Saetrom, P. (2020) sMETAseq: Combined Profiling of Microbiota and Host Small RNAs. iScience. 23, 5. doi: 10.1016/j.sci.2020.101131

Minezaki, Y., Homma, K., Nishikawa, K. (2005). Genome-wide survey of transcription factors in prokaryotes reveals many bacteria-specific families not found in archaea. *DNA Res.* 12, 269–280. doi: 10.1093/dnares/dsi016

Moeller S, Payá G, Bonete MJ, Gates AJ, Richardson DJ, Esclapez J, Rowley G (2021) Microbial small RNAs – the missing link in the nitrogen cycle? *Frontiers in Environmental Science* 9:660055. https://doi.org/10.3389/fenvs.2021.660055

Moller, T., Franch, T., Udesen, C., Gerdes, K., Valentin-Hansen, P. (2002). Spot 42 RNA mediates discoordinate expression of the *E. coli* galactose operon. Genes Dev. 16, 13: 1696-1706. doi: 10.1101/gad.231702.

Morfeldt, E., Taylor, D., von Gabain, A., Arvidson, S. (1995). Activation of alpha-toxin translation in *Staphylococcus aureus* by the trans-encoded antisense RNA, RNAII. EMBO. J. 14: 4569-4577.

Morita, T., Maki, K., Aiba, H. (2005) RNase E-based ribonucleoprotein complexes: mechanical basis of mRNA destabilization mediated by bacterial noncoding RNAs. Genes Dev. 34: 21-22. Doi: 10.1101/gad.1330405

Morita, T., Mochizuki, Y., Aiba, H. (2006). *Translational repression is sufficient for gene silencing by bacterial small noncoding RNAs in the absence of mRNA destruction.* Proc. Natl. Acad. Sci. USA. 103, 13: 4858-4863. doi: 10.1073/pnas.0509638103.

Muro-Pastor, M., Reyes, J., Florencio, F. (2005). Ammonium assimilation in cyanobacteria. Photosynth. Res. 83: 135-150. doi: 10.1007/s1120-004-2082-7

Murty, D., Kirschbaum, M. U. F., McMurtrie, R. E. and McGilvray, A. (2002). Does conversion of forest to agricultural land change soil carbon and nitrogen? A review of the literature. *Global Change Biology*, 8, 105–123.

Mustachio, L., Aksit, S., Mistry, R., Scheffler, R., Yamada, A., Liu, J. (2012). The *Vibrio cholerae* Mannitol Transporter is regulated posttransciptionally by the MtlS Small Regulatory RNA. J. Bacteriol. 194, 3: 598-606. doi: 10.1128/JB.06153-11

Naskulwar,K.; Peña-Castillo, L.; sRNARFTarget: A fast machine-learning-based approach for transcriptome-wide sRNA Target Prediction. bioRxiv 2021.03.05.433963; doi: https://doi.org/10.1101/2021.03.05.433963

Neha Rana, Piyush Gupta; Integrative Strategies for Bioremediation of Environmental Contaminants, Volume Two, 2023

Neupane, D.P, Jacquez, B., Sundararajan, A., Ramaraj, T., Schilkey, F.D., Yukl, E.T. (2017). Zinc-dependent transcriptional regulation in *Paracoccus denitrificans*. Front. Microbiol. 8: 569. doi: 10.3389/fmicb.2017.00569.

Nielsen, J.S., Bøggild, A., Andersen, C. B., Nielsen, G., Boysen, A., Brodersen, D.E., Valentin-Hansen, P. (2007). An Hfq-like protein in archaea: crystal structure and functional characterization of the Sm protein from *Methanococcus jannaschii*. RNA. 13: 2213-2223. doi: 10.1261/rna.689007

Norton, J., and Stark, J. (2011). Regulation and Measurement of Nitrification in terrestrial systems. Methods. Enzymol. 486: 343-368. doi: 10.1016/B978-0-12-381294-0.00015-8

Norton, J., and Ouyang, Y. (2019) Controls and Adaptive Management of Nitrification in Agricultural Soils. Front. Microbiol. 10. doi: 10.3389/fmicb.2019.01931

Novick, R., and Geisinger, E. (2008). Quorum sensing in *Staphylococci*. Annu. Rev. Genet. 42: 541-564. doi: 10.1146/annurev.genet.42.110807.091640.

Olaya-Abril, A., Luque-Almagro, V.M., Pérez, M.D., López, C.M., Amil, F., Cabello, P., et al. (2019). Putative small RNAs controlling detoxification of industrial cyanide-containing wastewaters by *Pseudomonas pseudoalcaligenes* CECT5344. PLoS. One. 14, 2: e0212032. doi: 10.1371/journal.pone.0212032.

Oleijniczak, M., and Storz, G. (2017). ProQ/FinO-domain protein: another ubiquitous family of RNA matchmakers? Mol. Microbiol. 104, 6: 905-915. doi: 10.1111/mmi.13679.

Otaka, H., Ishikawa, H., Morita, T., Aiba, H. (2011). PolyU Tail of rho-independent terminator of bacterial small RNA is essential for Hfq action. Proc. Natl. Acad. Sci. USA. 108: 13059-13064. doi: 10.1073/pnas.110705010

Pajares Silvia, Ramos Ramiro (2019) Processes and Microorganisms Involved in the Marine Nitrogen Cycle: Knowledge and Gaps; Frontiers in Marine Science; Vol. 6; doi: 10.3389

Papenfort, K., and Vanderpool, C. (2015). Target activation of regulatory RNAs in bacteria. FEMS Microbiol. Rev. 39: 362-378. doi: 10.1093/femsre/fuv016.

Papenfort, K., Sun, Y., Miyakoshi, M., Vanderpool, C., Vogel, J. (2013). Small RNA-mediated activation of sugar phosphate mRNA regulates glucose homeostasis. Cell. 153: 426-437. doi: 10.1016/j.cell.2013.03.003

Paget MS. Bacterial Sigma Factors and Anti-Sigma Factors: Structure, Function and Distribution. Biomolecules. 2015 Jun 26;5(3):1245-65. doi: 10.3390/biom5031245. PMID: 26131973; PMCID: PMC4598750.

Patrick R. Wright, Andreas S. Richter, Kai Papenfort, Martin Mann, Jörg Vogel, Wolf-gang R. Hess, Rolf Backofen, and Jens Georg. Comparative genomics boosts target prediction for bacterial small RNA. Proceedings of the National Academy of Sciences, 110(37):E3487– E3496, 2013

Pauleta, S., Moura, I. (2017) Assembly of CuZ and CuA in Nitrous Oxide Reductase. Encycl Inorg Bioinorg Chem.: 1-11 Doi: 10.1002/9781119951438.eibc2477

Pawlowsky, K., Klosse, U., the Brujin, F. (1991) Characterization of a novel *Azorhizobium caulinodans* ORS571 two-component regulatory system, NtrY/NtrX, involved in nitrogen fixation metabolism. Mol. Gen. Genet. 231, 1: 124-138. doi: 10.1007/BF00293830.

Payá, G., Bautista, V., Camacho, M., Bonete, M.J., Esclapez, J. (2020). New proposal of nitrogen metabolism regulation by small RNAs in the extreme halophilic archaeon *Haloferax mediterranei.* Mol. Genet. Genomics. 295, 3: 775-785. doi: 10.1007/s00438-020-01659-9.

Payá, G., Bautista, V., Camacho, M., Castejón-Fernández, N., Alcaraz, L.A., Bonete, M.J., Esclapez, J. (2018). Small RNAs of *Haloferax mediterranei:* Identification and Potential Involvement in Nitrogen Metabolism. Genes (Basel). 9, 2: 83. doi: 10.3390/genes9020083

Pei, G., Sun, T., Chen, S., Chen, L., Zhang, W. (2017). Systematic and functional identification of small non-coding RNAs associated with exogenous biofuel stress in cyanobacterium *Synechocystis* sp. PCC 6803. Biotechnol. Biofuels. 10: 57. doi: 10.1186/s13068-017-0743-y.

Peterson, J., Labhsetwae, P., Ellermeier, J., Kohler, P., Jain, A., Ha, T., Metcalf, W., Luthey-Schulten, Z. (2014). Towards a Computational Model of a Methane Producing Archaeum. Archaea. 2014: 898453. doi: 10.1155/2014/898453.

Pfeiffer V, Papenfort K, Lucchini S, Hinton JCD, Vogel J 2009. Coding sequence targeting by MicC RNA reveals bacterial mRNA silencing downstream of translational initiation. *Nat Struct Mol Biol* 16: 840–846

Philippot, L., Piutti, S., Martin-Laurent, F., Hallet, S., Germon, J. (2002). Molecular analysis of the nitrate-reducing community from unplanted and maize-planted soils. Appl. Environ. Microbiol. 68, 12: 6121-6128. doi: 10.1128/aem.68.12.6121-6128.2002.

Placella, S. and Firestone, M. (2013) transcriptional Response of Nitrifying Communities to Wetting of Dry Soil. Appl. Environ. Microbiol. 79, 10: 3294-3302. doi: 10.1128/AEM.00404-13

Pohlmann A, Cramm R, Schmelz K et al. A novel NO-responding regulator controls the reduction of nitric oxide in *Ralstonia eutropha*. *Mol Microbiol* 2000; 38:626–38.

Prasse, D., and Schmitz, R. (2018). Small RNAs Involved in Regulation of Nitrogen Metabolism. Microbiol Spectr. 6, 4. doi: 10.1128/microbiolspec.RWR-0018-2018.

Prasse, D., Förstner, K., Jäger, D., Backofen, R., Schmitz, R. (2017). sRNA 154 a newly identified regulator of nitrogen fixation in *Methanosarcina mazei* Strain Gö1. RNA. Biol. 14, 11: 1544-1558. doi: 10.1080/15476286.2017.1306170

Prévost K, Salvail H, Desnoyers G, Jacques JF, Phaneuf E, Massé E 2007. The small RNA RyhB activates the translation of *shiA* mRNA encoding a permease of shikimate, a compound involved in siderophore synthesis *Mol Microbiol* 64: 1260–1273

Prinn, R.G., Weiss, R.F., Arduini, J., Arnold, T., DeWitt, H.L., Fraser, P.J., Ganesan, A.L., et al. (2018). History of chemically and radiatively important atmospheric gases from the Advanced Global Atmospheric Gases Experiment (AGAGE). Earth. Syst. Sci. 10: 985-1018. doi: 10.5194/essd-10-985-2018

Prosser, J. (2007) The ecology of nitrifying nacteria. Biology of the nitrogen cycle: 223-243. Doi: 10.1016/B978-044452857

Pruitt K.D., Tatusova T., Brown G.R., Maglott D.R. NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy. *Nucleic Acids Res.* 2012;**40**:D130–D135.

Rees, D., Johnson, E. & Lewinson, O. ABC transporters: the power to change. *Nat Rev Mol Cell Biol* 10, 218–227 (2009). https://doi.org/10.1038/nrm2646

Reisch, C., Crabb, W., Gifford, S., Teng, Q., Stoudemayer, M., Moran, M., Whitman, W. (2013). Metabolism of dimethylsulphoniopropionate by *Ruegeria pomeroyi* DSS-3. Mol. Microbiol. 89: 774-791. doi: 10.1111/mmi.12314.

Reitzer L, Schneider BL. Metabolic context and possible physiological themes of sigma(54)-dependent genes in *Escherichia coli*. *Microbiol Mol Biol Rev.* 2001;**65**(3):422–44. doi: 10.1128/MMBR.65.3.422-444.2001

Renseigne, N., Umar, S., Iqbal, M. (2007) Nitrate accumulation in plants, factors affecting the process, and human health implication. A review. Agron Sustain Dev. Springer/EDP Sciences/INRA. 27, 1: 45-57. Doi: 10.1051/agro:2006021

Richardson, D., Felgate, H., Watmough, N., Thomson, A., Baggs, E. (2009). Mitigating release of the potent greenhouse gas N2O from the nitrogen cycle – could enzymatic regulation hold the key*?* Trends. Biotechnol. 27, 7: 388-397. doi: 10.1016/j.tibtech.2009.03.009.

Richardson, D., Edwards, M., White, G., Baiden, N., Hartshorne, R., Fredrickson, J., et al. (2012) Exploring the biochemistry at the extracellular redox frontier of bacterial mineral Fe(III) respiration. Biochem. Soc. Trans. 40.3: 493-500. doi: 10.1042/BST20120018

Rigali S, Derouaux A, Giannotta F, Dusart J. Subdivision of the helix-turn-helix GntR family of bacterial regulators in the FadR, HutC, MocR, and YtrA subfamilies. J Biol Chem. 2002 Apr 12; 277(15):12507– 15. PMID: 11756427

Riordan JT, Mitra A. Regulation of *Escherichia coli* Pathogenesis by Alternative Sigma Factor N. EcoSal Plus. 2017;7(2).

Rivers, A., Burns, A., Chan, L., Moran, M. (2016). Experimental Identification of Small Non-coding RNAs in the Model Marine Bacterium *Ruegeria Pomeroyi* DSS-3. Front. Microbiol. 29, 7: 380. doi: 10.3389/fmicb.2016.00380.

Robertson, P., Vitousek, P. (2009) Nitrogen in Agriculture: Balancing the Cost of an Essential Resource. Ann Rev Environ Res. 34, 1: 97-125. Doi: 10.1146/annurev.environ.032108.105046

Romermann D, Warrelmann J, Bender RA et al. An rpoN-like gene of *Alcaligeneseutr*ophus and *Pseudomonas facilis* controls expression of diverse metabolic pathways, including hydrogen oxidation. *J Bacteriol* 1989; 171:1093–9.

Romeo, A., Sonnleitner, E., Sorger-Domenigg, T., Nakano, M., Eisenhaber, B., Bläsi, U. (2012). Transcriptional regulation of nitrate assimilation in *Pseudomonas aeruginosa* via transcriptional antitermination within the nirBD-PA1779-cobA operon. Microbiol. 158: 1543-1552. doi: 10.1099/mic.0.053850-0.

Rowley, G., Hensen, D., Felgate, H., Arkenberg, A., Appia-Ayme, C., Prior, K., et al. (2012) Resolving the contributions of the membrane-bound and periplasmic nitrate reductase systems to nitric oxide and nitrous oxide production in Salmonella enterica serovar Typhimurium. Biochem. J. 441. 2: 755-762. doi: 10.1042/BJ20110971

Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA and Barrell B **Artemis: sequence visualization and annotation.** Bioinformatics (Oxford, England) 2000;16;10;944-5 PUBMED: 11120685

Saggar, S., Jha, N., Deslippe, J., Bolan, N., Luo, J., Giltrap, D., Kim, D., Zaman, M., Tillman, R. (2013). Denitrification and N_2O:N_2 production in temperate grasslands: Processes, measurements, modelling and mitigating negative impacts.*;* Sci. Total. Environ. 465: 173-95. doi: 10.1016/j.scitotenv.2012.11.050.

Samuel, A.L. Some Studies in Machine Learning Using the Game of Checkers. *IBM J. Res. Dev.* 1959, *3*, 210–229.

Sánchez, C., Itakura, M., Okubo, T., Matsumoto, T., Yoshikawa, H., Gotoh, A., et al. (2014) The nitrate-sensing NasTS system regulates nitrous oxide reductase and periplasmic nitrate reductase in Bradyrhizobium japonicum. Environm. Microbiol. 16, 3263-3274. doi: 10.1111/1462-2920.12546

Sanders, D., Gillece-Castro, B., Burlingame, A., Hoshland, D. (1992) Phosphorylation site of NtrC, a protein phosphatase whose covalent intermediate activates transcription. J Bacteriol. 174: 5117-5122. doi: 10.1128/jb.174.15.5117-5122.1992.

Saunders, N.F.W., Houben, E.N.G., Koefoed, S., De Weert, S., Reijnders, W.N.M., Westerhoff, H.V., De Boer, A.P.N. and Van Spanning, R.J.M. (1999), Transcription regulation of the *nir* gene cluster encoding nitrite reductase of *Paracoccus denitrificans* involves NNR and NirI, a novel type of membrane protein. Molecular Microbiology, 34: 24-36. https://doi.org/10.1046/j.1365-2958.1999.01563.x

Sauter, C., Basquin, J., Suck, D. (2003). Sm-like proteins in Eubacteria: the crystal structure of the Hfq protein from *Escherichia coli.* Nucleic. Acids. Res. 31: 4091-4098. doi: 10.1093/nar/gkg480

Sass, A.M., Van Acker, H., Förstner, K.U. *et al.* Genome-wide transcription start site profiling in biofilm-grown *Burkholderia cenocepacia* J2315. *BMC Genomics*16, 775 (2015). https://doi.org/10.1186/s12864-015-1993-3

Schlüter et al., 2013 Global mapping of transcription start sites and promoter motifs in the symbiotic a-proteobacterium SInorhizobium meliloti1021

Schreiber, K., Krieger, R., Benkert, B. Eschbach, M., Arai, H., Schobert, M., Jahn, D. (2007). The anaerobic regulatory network required for *Pseudomonas aeruginosa* nitrate respiration. J Bacteriol. 189, 11: 4310-4314. doi: 10.1128/JB.00240-07

Schröder G, Krause S, Zechner EL, Traxler B, Yeo HJ, Lurz R, Waksman G, Lanka E. TraG-like proteins of DNA transfer systems and of the Helicobacter pylori type IV secretion system: inner membrane gate for exported substrates? J Bacteriol. 2002 May;184(10):2767-79. doi: 10.1128/JB.184.10.2767-2779.2002. PMID: 11976307; PMCID: PMC135038.

Sedlyarova, N., Shamovskym, I., Bharati, B., Epshtein, V., Chen, J., Gottesman, S., et al. (2016). sRNA-mediated control of transcription termination in *E. coli.* Cell. 167: 111-121. doi: 10.1016/j.cell.2016.09.004.

Seitzinger, S., Harrison, J. A., Böhlke, J. K. *et al*. (2006). Denitrification across landscapes and waterscapes: a synthesis. *Ecological Applications*, 16, 2064–2090.

Sharma, C., Darfeuille, F., Plamtinga, T., Vogel, J. (2007). A small RNA regulates multiple ABC transporter mRNAs by targeting C/A- rich elements inside and upstream of ribososme-binding sites. Genes Dev. 21, 21: 2804-2817. doi: 10.1101/gad.447207.

Sharma, C., Hoffman, S., Darfeuille, F., Reignier, J., Findeiss, S., Sittka, A., et al. (2010) The primary transcriptome of the major human pathogen *Heliobacter pylori*. Nature. 464: 250-255. doi: 10.1038/nature08756

Shi, Y., Tyson, G., DeLong, E. (2009). Metatranscriptomics reveals unique microbial small RNAs in the ocean's water column. Nature. 459, 7244: 266-269. doi: 10.1038/nature08055

Sinclair, A., Edwards, A. (2008) Micronutrient deficiency problems in agricultural crops in Europe. Micronutrient deficiencies in global crop production. Springer, Dordrecht: 225-244

Sittka A, Lucchini S, Papenfort K, Sharma C, Rolle K, Binnewies T, Hinton J, Vogel J: Deep sequencing analysis of small noncoding RNA and mRNA targets of the global post-transcriptional regulator, Hfq. PLoS Genet. 2008, 4 (8)

Smil, V. (2002). Nitrogen and food production: Proteins for human diets. Ambio. 31: 126-132. doi: 10.1579/0044-7447-31.2.126.

Smirnov, A., Forstner, K., Holmqvist, E., Otto, A., Gunster, R., Becher, D., et al. (2016). Grad-seq guides discovery of ProQ as a major small RNA-binding protein. Proc. Natl. Acad. Sci. USA. 113: 11591-11596. doi: 10.1073/pnas.1609981113.

Smirnov, A., Wang, C., Drewry, L., Vogel, J. (2017). Molecular Mechanism of mRNA repression in trans by a ProQ dependent small RNA. EMBO. J. 36: 1029-1045. doi: 10.15252/embj.201696127.

Smith K., Mosier, A., Crutzen P. (2012). The Role of N_2O derived from crop-based biofuels, and from agriculture in general, in Earth's climate. Philos. Trans. R. Soc. Lond. B. Biol. Sci. 367, 1593: 1169-1174. doi: 10.1098/rstb.2011.0313.

Sonnleitner, E., Gonzalez, N., Sorger-Domenigg, T., Heeb, S., Richter, A., Backofen, R., et al. (2011). The small RNA PhrS stimulates synthesis of the *Pseudomonas aeruginosa* quinolone signal. Mol. Micobiol. 80: 868-885. doi: 10.1111/j.1365-2958.2011.07620.x.

Sridhar, J., Gunasekaran, P. (2013) Computational Small RNA Prediction in Bacteria. Bioinform. Biol. Insights. 7: 83-95. doi: 10.4137/BBI.S11213

Storz, G., Opdyke, J., Zhang, A. (2004). Controlling mRNA stability and translation with small, Noncoding RNAs. Curr. Opin. Microbiol. 7, 2: 140-144. doi: 10.1016/j.mib.2004.02.015.

Storz, G., Vogel, J., Wassarman, K. (2011). Regulation by small RNAs in bacteria: Expanding Frontiers. Mol. Cell. 43, 6: 880-891. doi: 10.1016/j.molcel.2011.08.022.

Stremińska, M., Felgate, H., Rowley, G., Richardson, D., Baggs, E. (2012) Nitrous oxide production in soil isolates of nitrate-ammonifying bacteria. Environ. Microbiol. Rep. 4. 1: 66-71. doi: 10.1111/j.1758-2229.2011.00302.x.

Sullivan, M., Gates, A., Appia-Ayme, C., Rowley, G., Richardson, D. (2013). Copper control of bacterial nitrous oxide emission and its impact on vitamin B12-dependent metabolism. Proc. Natl. Acad. Sci. USA. 110, 49: 19926-19931. doi: 10.1073/pnas.1314529110.

Sutton, M., Howard, C., Erisman, J., Billen, G., Bleeker, A., van Grinsven, H., et al. (2011) The European Nitrogen Assessment: Sources, Effects and Policy Perspectives. Cambridge: Cambridge University Press. ISBN 9781107006126

Suvorova IA, Korostelev YD, Gelfand MS (2015) GntR Family of Bacterial Transcription Factors and Their DNA Binding Motifs: Structure, Positioning and Co-Evolution. PLoS ONE 10(7): e0132618. https://doi.org/10.1371/journal.pone.0132618

Tata, M., Amman, F., Pawar, V., Wolfinger, M.T., Weiss, S., Häussler, S., Bläsi, U. (2017). The anaerobically induced sRNA PaiI affects denitrification in *Pseudomonas aeruginosa* PA14. Front. Microbiol. 8: 2312. doi: 10.3389/fmicb.2017.02312.

Teikari, J., Österholm, J., Kopf, M., Battchikova, N., Wahlsten, M., Aro, E., et al. (2015). Transcriptomic and Proteomic Profiling of *Anabaena sp.* Strain 90 under Inorganic Phosphorus Stress. Appl. Environ. Microbiol. 81, 15: 5212-5222. doi: 10.1128/AEM.01062-15

Thöny B, Hennecke H. The -24/-12 promoter comes of age. FEMS Microbiol Rev. 1989 Dec;5(4):341-57. doi: 10.1016/0168-6445(89)90028-4. PMID: 2517036

Thore, S., Mayer, C., Sauter, C., Weeks, S., Suck, D. (2003). Crystal structures of the *Pyrococcus abyssi* Sm core and its complex with RNA: common features of RNA binding in Archaea and Eukarya. J. Biol. Chem. 278: 1239-1247. doi: 10.1074/jbc.M207685200

Timothy L. Bailey, James Johnson, Charles E. Grant, William S. Noble, "The MEME Suite", *Nucleic Acids Research*, 43(W1):W39-W49, 2015.

Tett AJ, Rudder SJ, Bourdès A, Karunakaran R, Poole PS. 2012. Regulatable vectors for environmental gene expression in Alphaproteobacteria. *Appl Environ Microbiol* 78:7137–7140. doi: 10.1128/AEM.01188-12.

Törö, I., Basquin, J., Teo-Dreher, H., Suck, D. (2002). Archaeal Sm proteins form heptameric and hexameric complexes: crystal structures of the Sm1 and Sm2 proteins from the hyperthermophile *Archaeoglobus fulgidus.* J. Mol. Biol. 320: 129-142. doi: 10.1016/S0022-2836(02)00406-0

Törö, I., Thore, S., Mayer, C., Basquin, J., Séraphin, B., Suck, D. (2001). RNA binding in an Sm core domain: X-ray structure and functional analysis of an archaeal Sm protein complex. EMBO. J. 20: 2293-2303. doi: 10.1093/emboj/20.9.2293

Torres, M., Bueno, E., Jiminez-Leiva, A., Cabrera, J., Bedmar, E., Mesa, S., Delgado, M. (2017). FixK2 is the main transcriptional activator of *Bradyrhizobium diazoefficiens nosRZDYFLX* genes in response to low oxygen. Front. Microbiol. 8: 1621. doi: 10.3389/fmicb.2017.01621.

Torsvik, V., Goksoyr, J., and Daae, F. L. (1990). High diversity in DNA of soil bacteria. *Appl. Environ. Microbiol*. 56, 782–787.

Tosques, I., Shi, J., Shapleigh, J. (1996) Cloning and characterisation of nnR, whose product is required for the expression of proteins involved in nitric oxide metabolism. J Bacteriol. 178: 4958-4965. doi: 10.1128/jb.178.16.4958-4964.1996.

Totten PA, Lara JC, Lory S. The rpoN gene product of *Pseudomonas aeruginosa* is required for expression of diverse genes, including the flagellin gene. *J Bacteriol* 1990; 172:389–96.

Udekwu, K., Darfeuille, F., Vogel, J., Reimegard, J., Holmqvist, E., Wagner, E. (2005). Hfq-dependent regulation of OmpA synthesis is mediated by an antisense RNA. Genes. Dev. 19, 19: 2355-2366. doi: 10.1101/gad.354405.

Van Maanen, J.M.S., Welle, I.J., Hageman, G., Dallinga, J.W., Mertens, P.L.J.M., Kleinjans, J.C.S. (1996). Nitrate contamination of drinking water: relationship with HPTR variant frequency in lymphocyte DNA and urinary excretion of N-nitrosamines. Environ. Health. Perspect. 104: 522-528. doi: 10.1289/ehp.96104522.

Van Spanning, R., De Boer, A., Reijnders, W., Spiro, S., Westerhoff, H.V., Stouthamer, A.H., Van der Oost, J. (1995). Nitrite and nitric oxide reduction in *Paracoccus denitrificans* is under the control of NNR, a regulatory protein that belongs to the FNR family of transcriptional activators. FEBS. Lett. 360, 2: 151-154. doi: 10.1016/0014-5793(95)00091-m.

Vanderpool, C., and Gottesman, S. (2007). A dual function for a bacterial small RNA: SgrS performs base pairing-dependent regulation and encodes a functional polypeptide. PNAS. 104, 51: 20454-20459. Doi: 10.1073/pnas.0708102104

Vanderpool, C., and Gottesmann, S. (2004). Involvement of a novel transcriptional activator and small RNA in post-transcriptional regulation of the glucose phosphoenolpyruvate phosphotransferase system. Mol.r Microbiol. 54, 4: 1076-1089. doi: 10.1111/j.1365-2958.2004.04348.x.

Vanderpool, C., Gottesman, S. (2007) The Novel Transcription Factor SgrR coordinates the response to glucose-phosphate stress. J Bacteriol. 189, 6: 2238-2248. Doi: 10.1128/JB.01689-06

Veit, K., Ehlers, C., Schmitz, R. (2005). Effect of nitrogen and carbon sources on transcription of soluble methyltransferases in *Methanosarcina mazei* strain Gö1. J. Bacteriol. 187: 6147-6154. doi: 10.1128/JB.187.17.6147-6154.2005.

Vitousek, P., and Howarth, R.W (1991). Nitrogen limitation on land and in the sea: How can it occur? Biogeochemistry. 13: 87-115. doi: 10.1007/BF00002772.

Vo PNL, Lee HM, Ren J, Na D. Optimized expression of Hfq protein increases Escherichia coli growth. J Biol Eng. 2021 Feb 18;15(1):7. doi: 10.1186/s13036-021-00260-x. PMID: 33602295; PMCID: PMC7890833

Vogel, J., and Luisi, B.F. (2011). Hfq and its constellation of RNA. Nat. Rev. Microbiol. 9: 578e589. doi: 10.1038/nrmicro2615

Vogel, J. (2020) An RNA biology perspective on species-specific programmable RNA antibiotics. Mol. Microbiol. 113, 3: 550-559. doi: 10.1111/mmi.14476

Vollack, K., Härtig, E., Körner, H., Zumft, W.G. (1999). Multiple transcription factors of the FNR family in denitrifying *Pseudomonas stutzeri*: characterisation of four fnr-like genes, regulatory responses and cognare metabolic processes. Mol. Microbiol. 31, 6: 1681-1694. doi: 10.1046/j.1365-2958.1999.01302.x.

Wagner, E. (2013) Cycling of RNAs on Hfq. RNA Biol. 10: 619-626. doi: 10.4161/rna.24044

Walker, C., Torre, J., Klotz, M., Urakawa, H., Pinel, N., Arp, D., et al. (2010) *Nitrosopumilus maritimus* genome reveals unique mechanisms for nitrification and autotrophy in globally distributed marine crenarchaea. PNAS. 107, 19: 8818-1123. doi: 10.1073/pnas.0913533107

Wang, L., Xia, M., Wang, H., Huang, K., Qian, C., and Maravelias, C. T. (2018). Greening ammonia toward the solar ammonia refinery. *Joule* 2 (6), 1–20. doi:10.1016/j.joule.2018.04.017

Warman EA, Forrest D, Guest T, Haycocks JJRJ, Wade JT, Grainger DC. Widespread divergent transcription from bacterial and archaeal promoters is a consequence of DNA-sequence symmetry. Nat Microbiol. 2021 Jun;6(6):746-756. doi: 10.1038/s41564-021-00898-9. Epub 2021 May 6. PMID: 33958766; PMCID: PMC7612053.

Wassarman, K. (2002) Small RNAs in bacteria: diverse regulators of gene expression in response to environmental changes. Cell. 109, 2: 141-144. Doi: 10.1016/S0092-8674(02)00717-1

Wassarman, K., and Storz, G. (2000). 6S RNA regulates *E. coli* RNA polymerase activity. Cell. 101: 613-623. doi: 10.1016/s0092-8674(00)80873-9.

Waters, L., Storz, G. (2009) Regulatory RNAs in Bacteria. Cell. 136, 4: 615-628. doi: 10.1016/j.cell.2009.01.043

Weidenbach, K., Ehlers, C., Kock, J., Ehrenreich, A., Schmitz, R.A. (2008) Insights into the NrpR regulon in *Methanosarcina mazei* Go1. Arch. Microbiol. 190: 319-332. doi: 10.1007/s00203-008-0369-3.

Weidenbach, K., Ehlers, C., Kock, J., Schmitz, R.A. (2010). NrpRII mediates contacts between NrpRI and general transcription factors in the archaeon *Methanosarcina mazei* Gö1. FEBS. J. 277: 4398-4411. doi: 10.1111/j.1742-4658.2010.07821.x

Wenner, N., Maes, A., Cotado-Sampayo, M., Lapouge, K. (2014). NrsZ: a novel, processed, nitrogen-dependent, small non-coding RNA that regulates *Pseudomonas aeruginosa* PAO1 virulence. Environmental Microbiology. 16, 4: 1053-1068. doi: https://doi.org/10.1111/1462-2920.12272

Werner F & Grohmann D Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol* 9, 85–98 (2011).

Westermann, A., Venturini, E., Sellin, M., Förstner, K., Hardt, W., Vogel, J. (2019). The Major RNA-binding protein ProQ impacts virulence Gene expression in *Salmonella enterica* Serovar Typhimurium. mBIO. 10, 1: e02504-18. doi: 10.1128/mBio.02504-18.

Williams, B., and Todd, J. (2019). A day in the life of marine sulfonates. Nat. Microbiol. 4, 10:1610-1611. doi: 10.1038/s41564-019-0576-5.

Wu P, Wang Z, Zhu Q, Xie Z, Mei Y, Liang Y, Chen Z (2021) Stress preadaptation and overexpression of rpoS and hfq genes increase stress resistance of Pseudomonas fluorescens ATCC13525, Microbiological Research, Vol. 250, 126804, ISSN 0944-5013,

Zhan, Y., Yan, Y., Deng, Z., Chen, M., Lu, W., Shang, L., et al. (2016). The Novel regulatory ncRNA, NfiS optimises nitrogen fixation via base pairing with the nitrogenase gene nifK mRNA in *Pseudomonas stutzeri* A1501. Proc. Natl. Acad. Sci. USA. 113, 30: E4348-56. doi: 10.1073/pnas.1604514113.

Zhang J, Zheng M, Tang Z, Zhong S, Bu T, Li Q. The Regulatory Functions of the Multiple Alternative Sigma Factors RpoE, RpoHI, and RpoHII Depend on the Growth Phase in *Rhodobacter sphaeroides*. *Microorganisms*. 2023; 11(11):2678. https://doi.org/10.3390/microorganisms11112678

Zheng M, Cooper DR, Grossoehme NE, Yu M, Hung LW, Cieslik M, et al. Structure of Thermotoga mar- itima TM0439: implications for the mechanism of bacterial GntR transcription regulators with Zn2 +-binding FCD domains. Acta Crystallogr D Biol Crystallogr. 2009 Apr; 65(Pt 4):356–65. doi: 10.1107/ S0907444909004727 PMID: 19307717

Zheng, A.; Panja, S.; Woodson, S.A. Arginine Patch Predicts the RNA Annealing Activity of Hfq from Gram-Negative and Gram-Positive Bacteria, Journal of Molecular Biology, Volume 428, Issue 11, 2016, Pages 2259-2264, ISSN 0022-2836, https://doi.org/10.1016/j.jmb.2016.03.027.

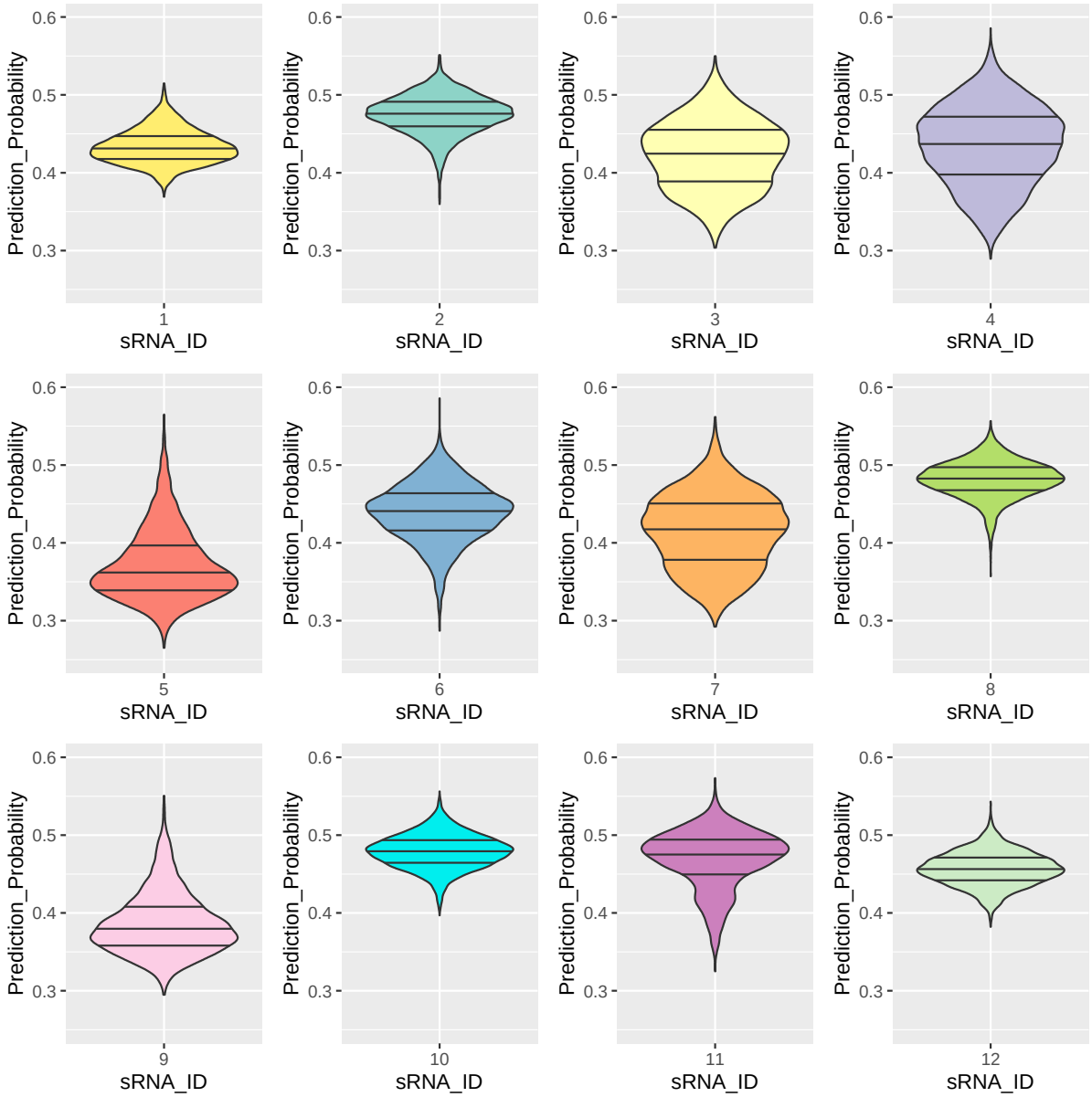
Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res. 2003 Jul 1;31(13):3406-15. doi: 10.1093/nar/gkg595. PMID: 12824337; PMCID: PMC169194.

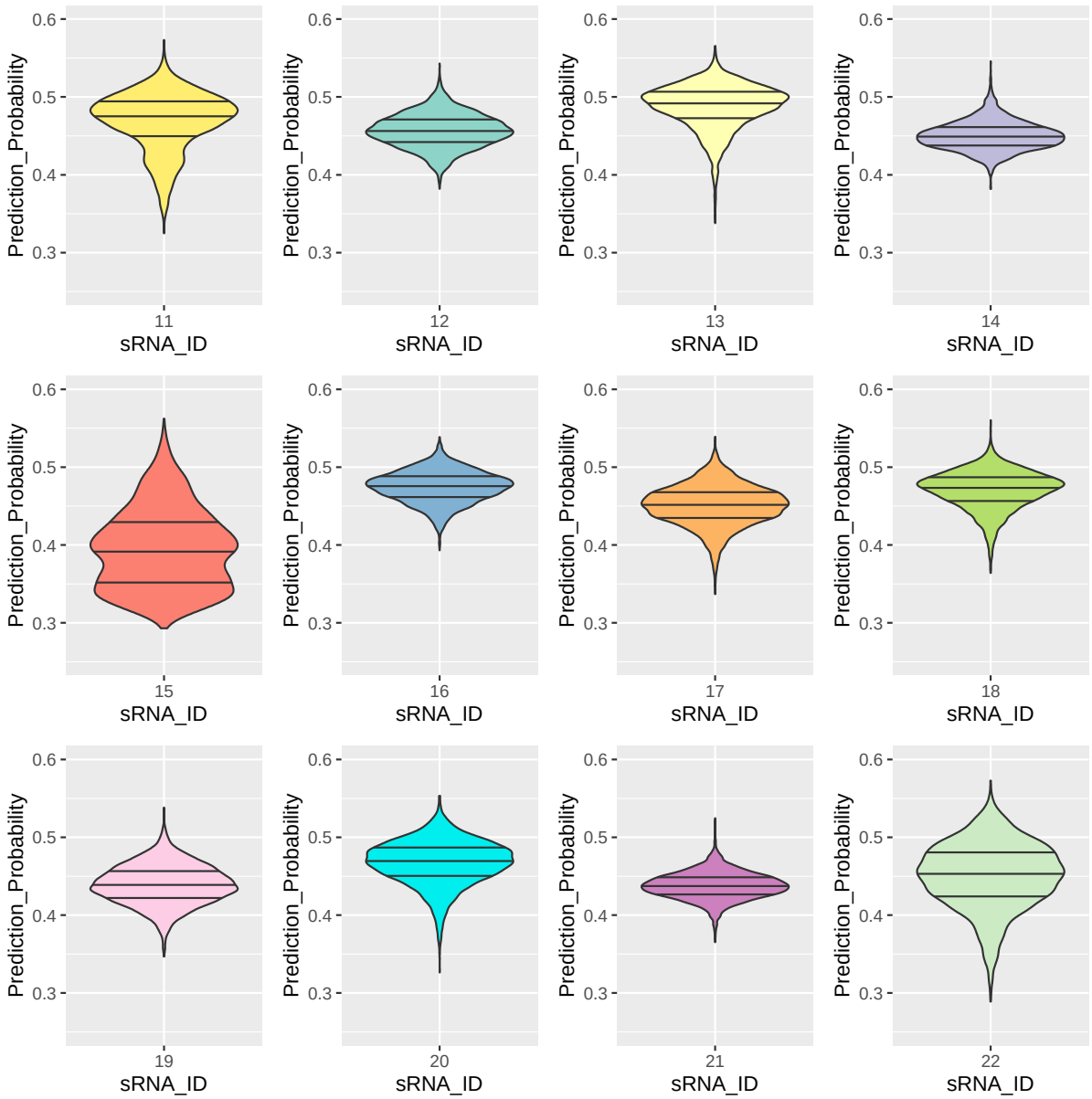
Zumft, W. (2005). Nitric oxide reductases of prokaryotes with emphasis on the respiratory, heme-copper oxidase type. J. Inorg. Biochem. 99, 1: 194-215. doi: 10.1016/j.jinorgbio.2004.09.024.

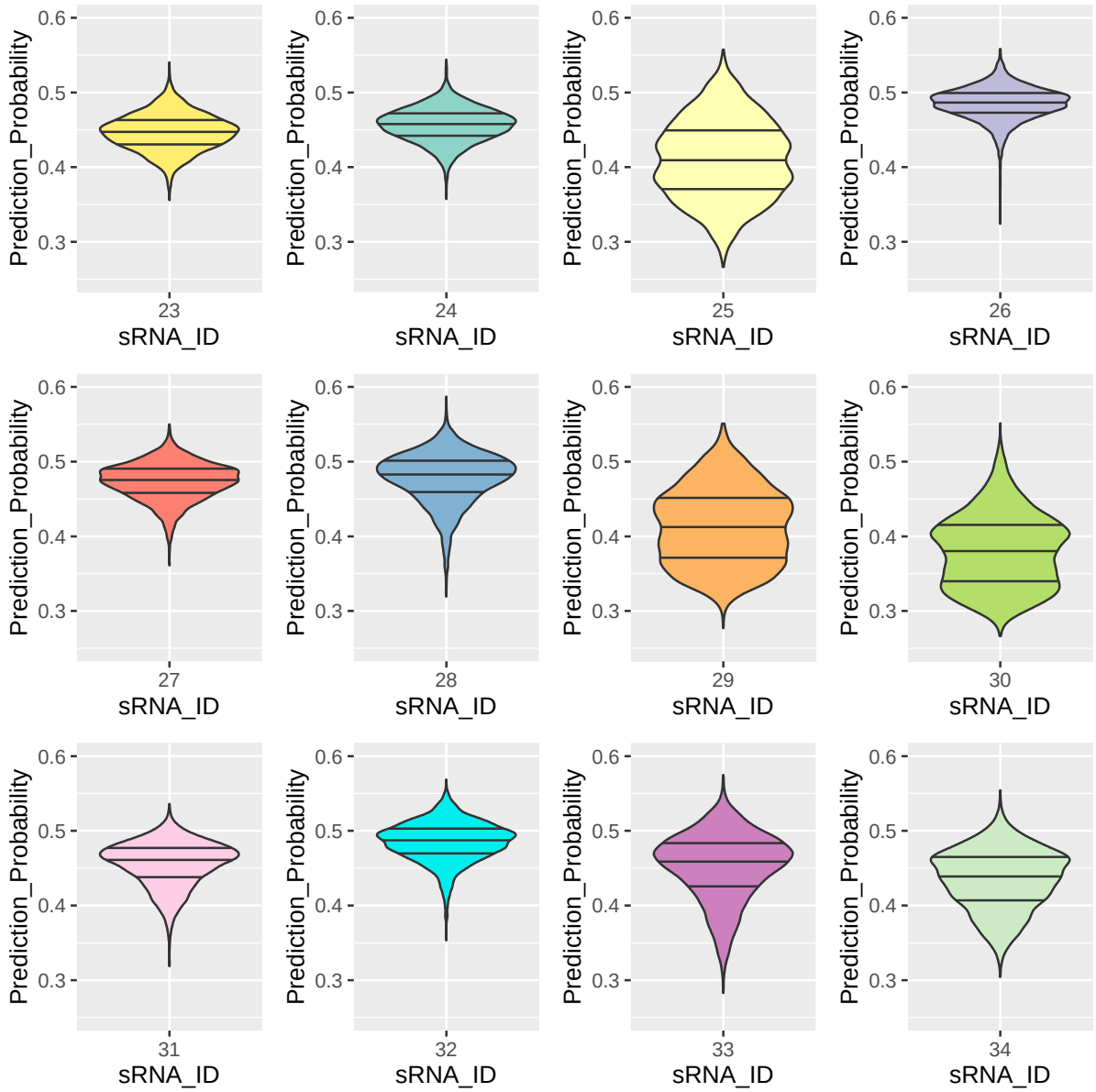
Zumft, W., and Kroneck, P. (2007). Respiratory transformation of nitrous oxide to dinitrogen by Bacteria and Archaea. Adv. Microb. Physiol. 52: 107-227. doi: 10.1016/S0065-2911(06)52003-X.

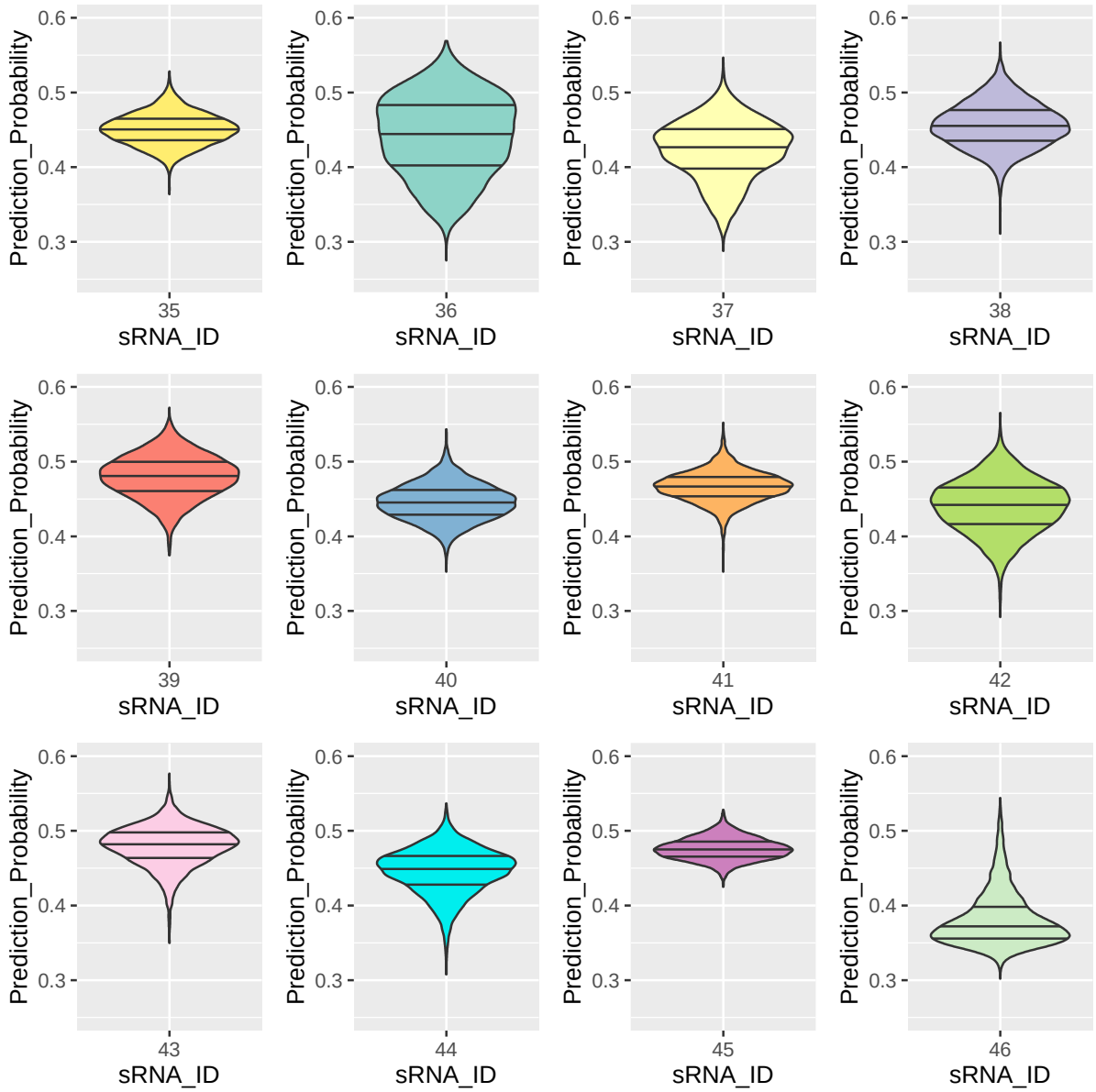
9 Appendices

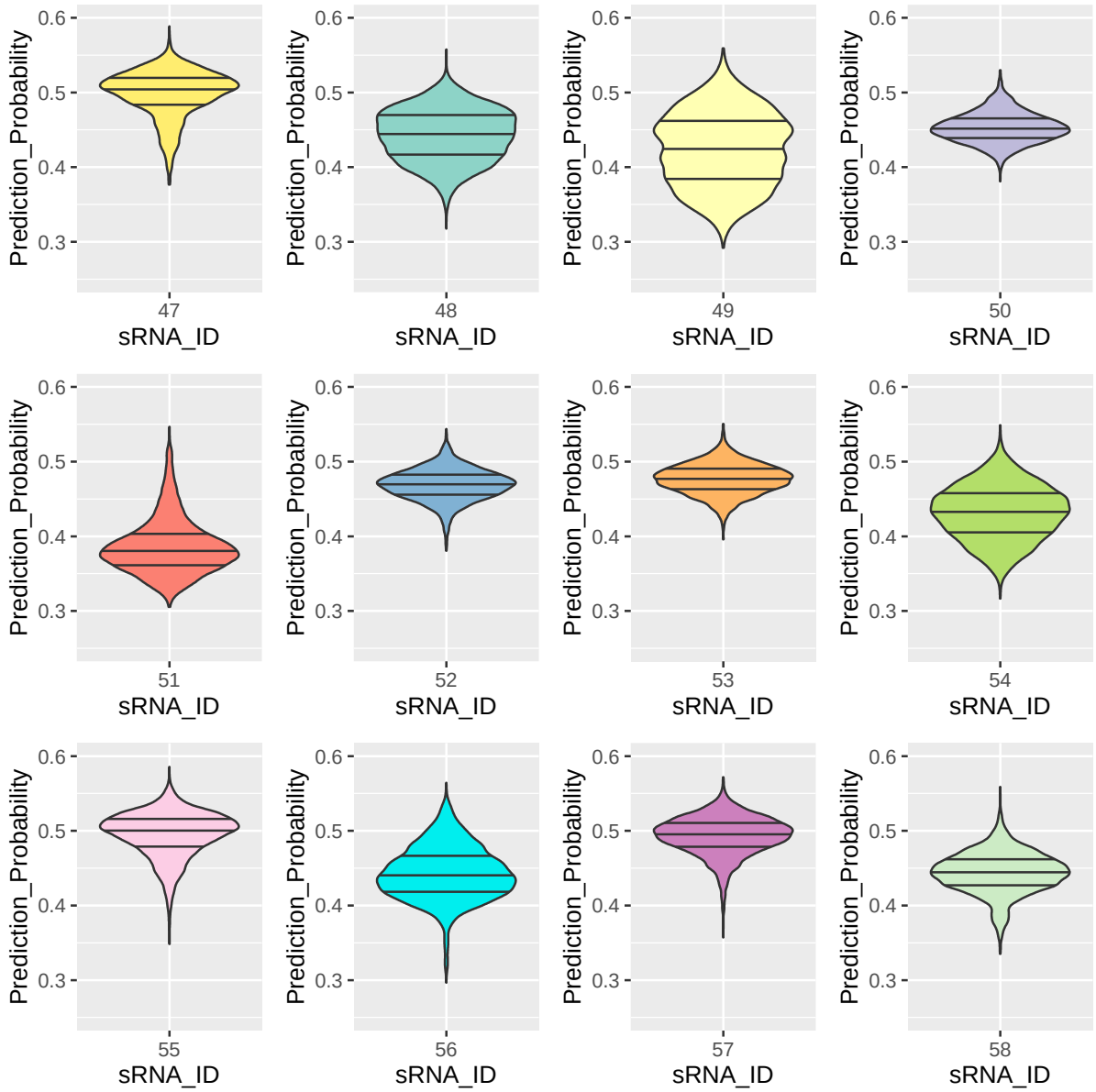
Appendix A: sRNARFTarget results for all sRNAs identified across the *P. denitrificans* genome

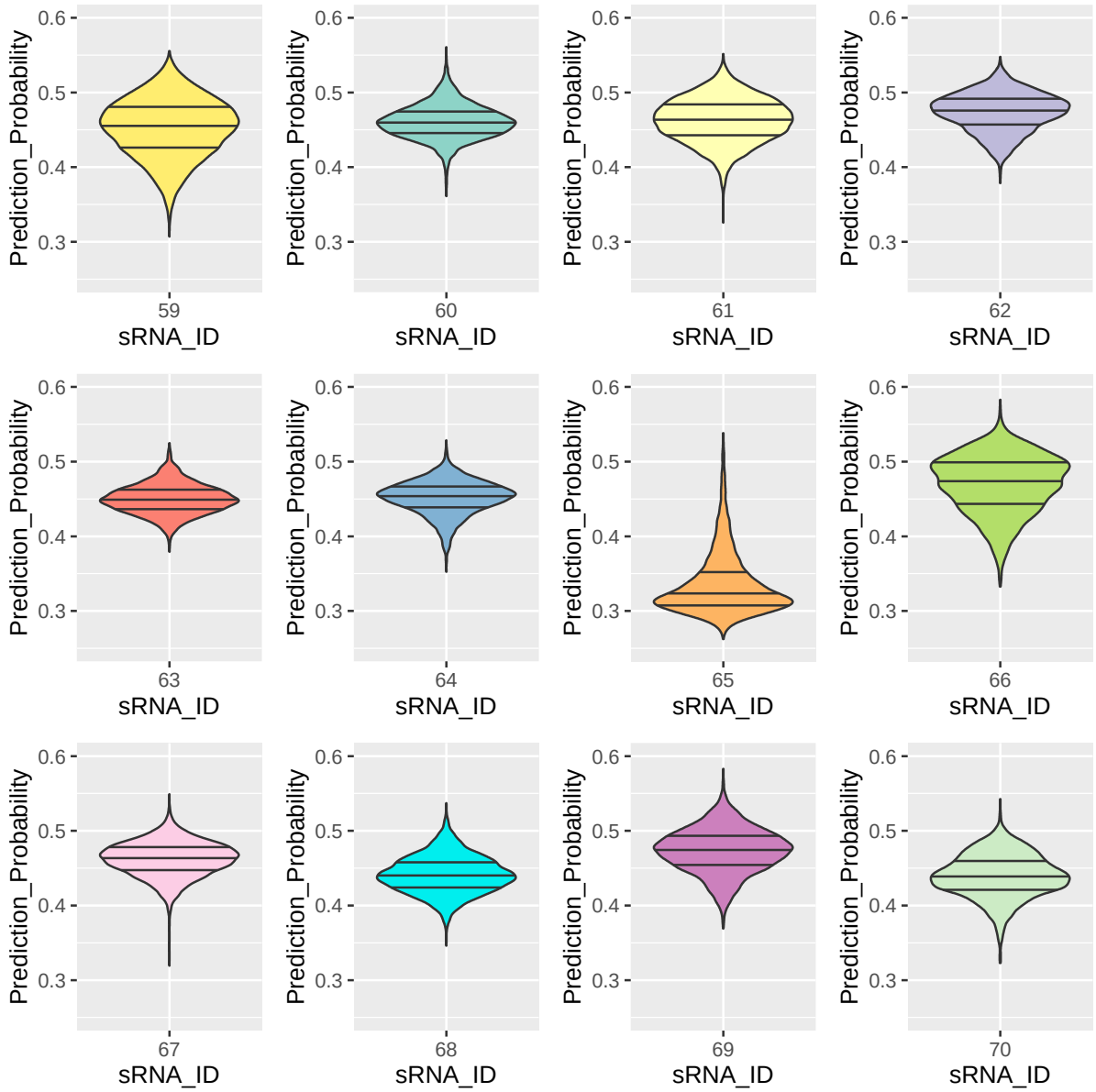


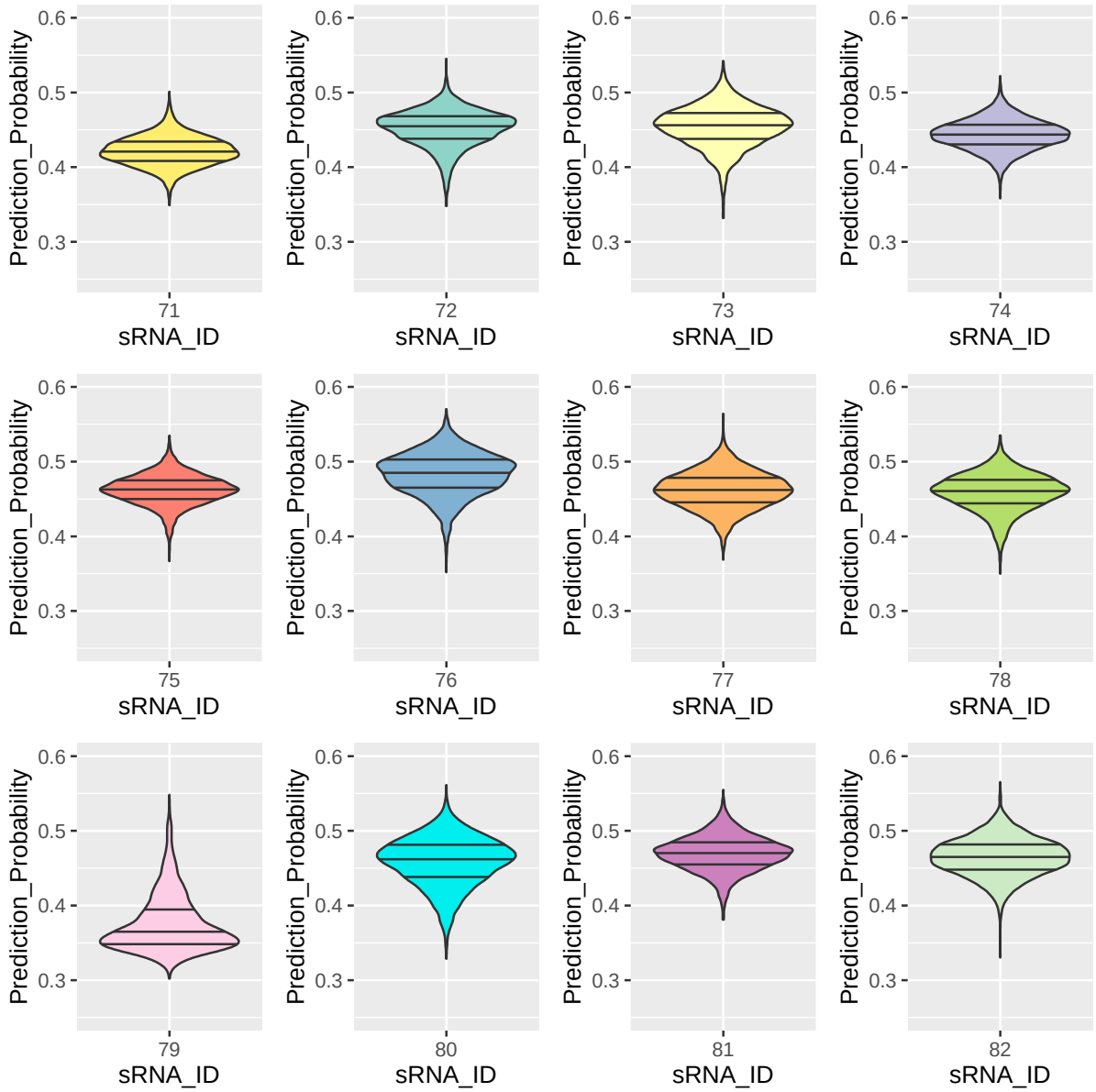

Prediction_Probability
0.6
0.5
0.4
0.3
11
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
12
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
13
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
14
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
15
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
16
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
17
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
18
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
19
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
20
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
21
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
22
sRNA_ID

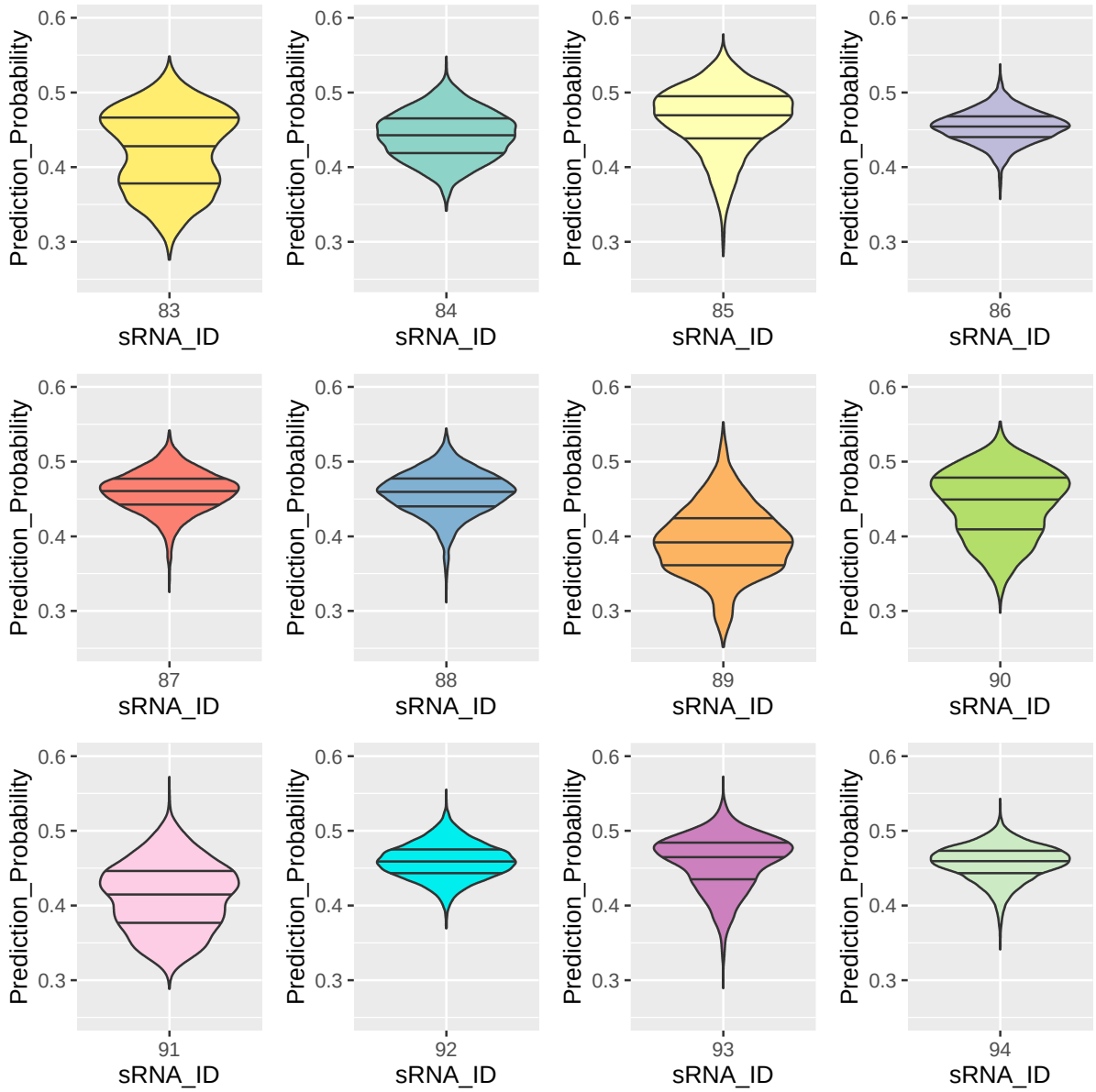

Prediction_Probability
0.6
0.5
0.4
0.3
23
24
25
26
27
28
29
30
31
32
33
34
sRNA_ID

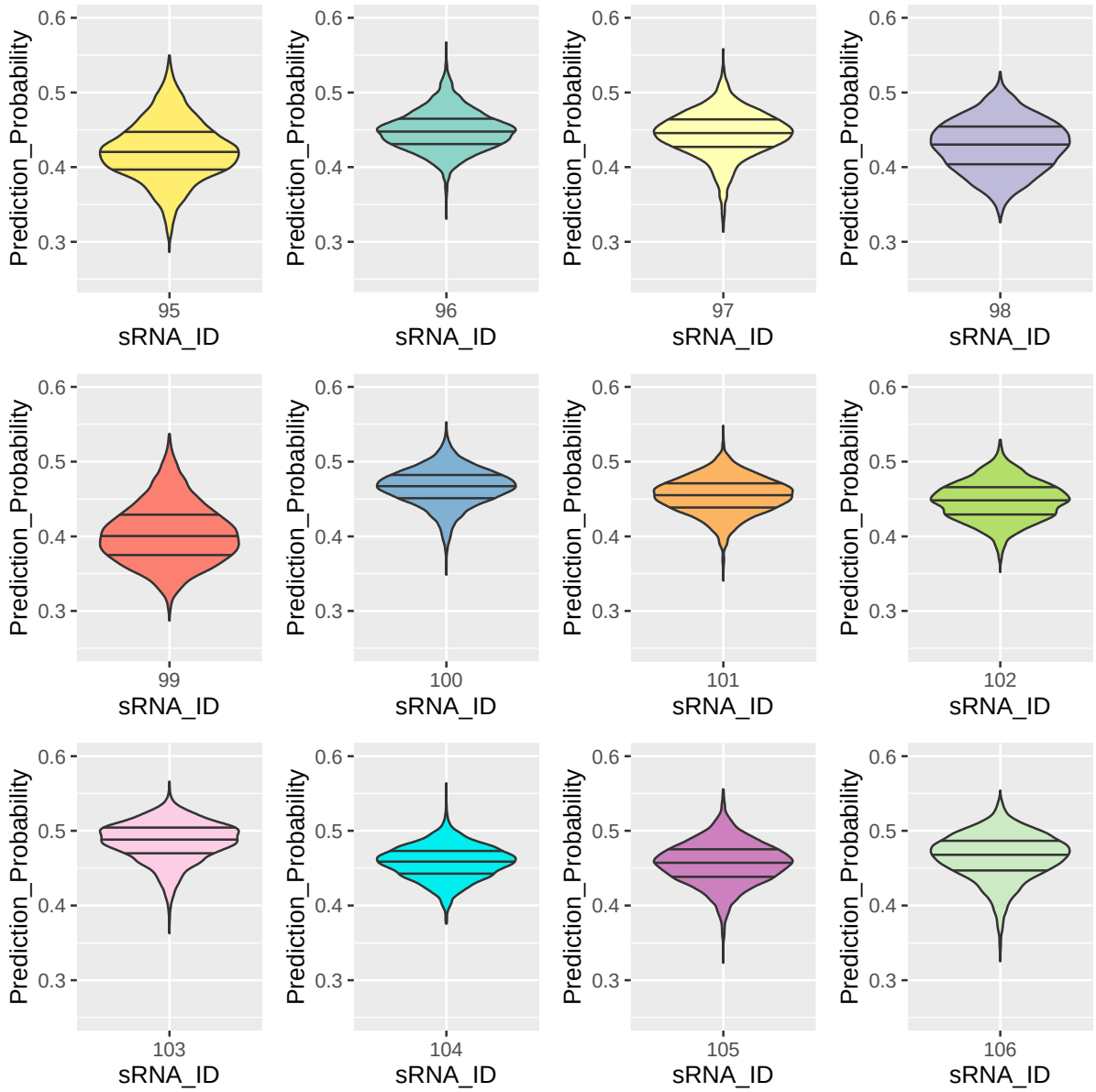

Prediction_Probability
0.6
0.5
0.4
0.3
35
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
36
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
37
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
38
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
39
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
40
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
41
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
42
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
43
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
44
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
45
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
46
sRNA_ID

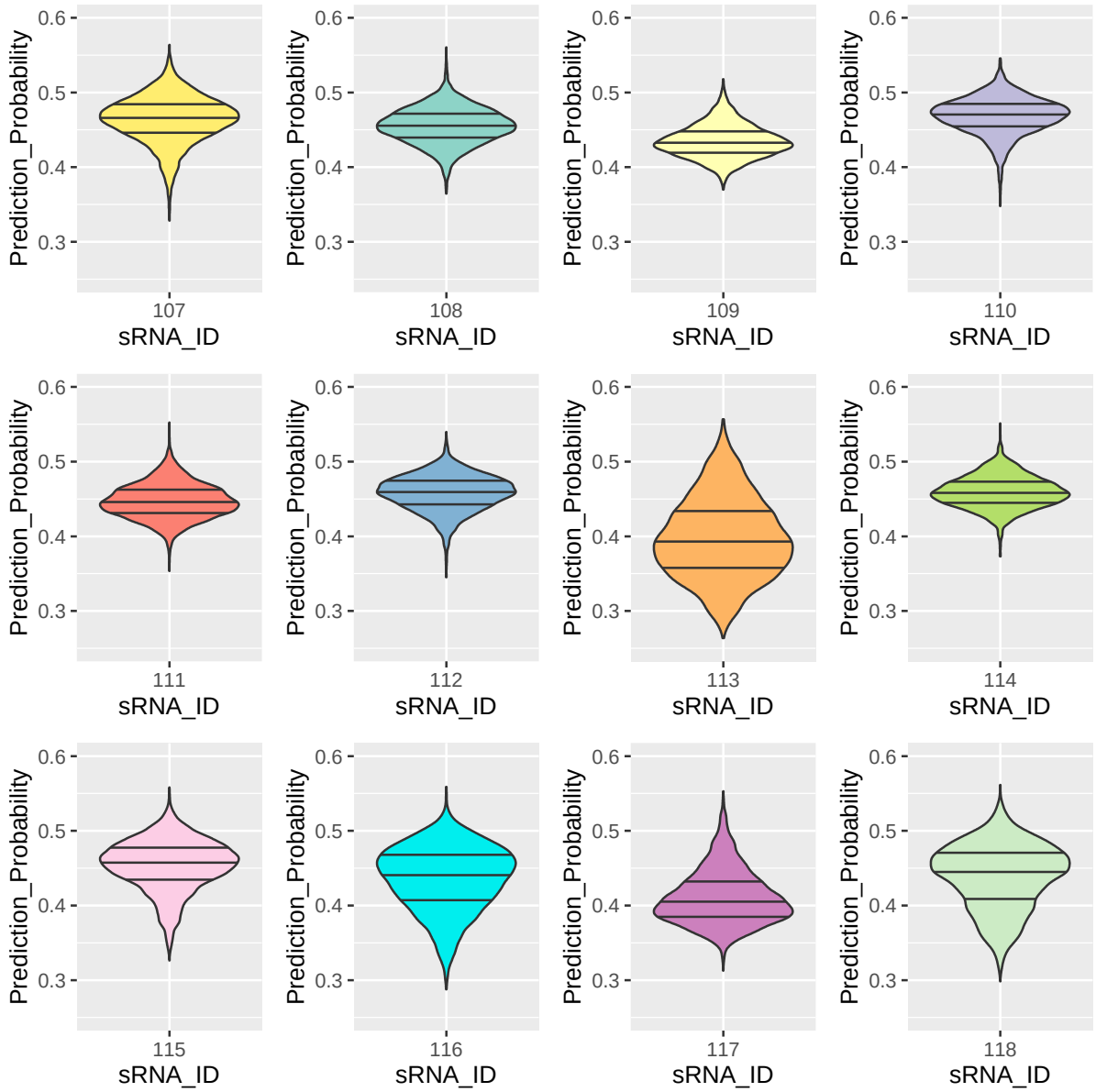

Prediction_Probability
0.6
0.5
0.4
0.3
47
sRNA_ID
48
49
50
51
52
53
54
55
56
57
58

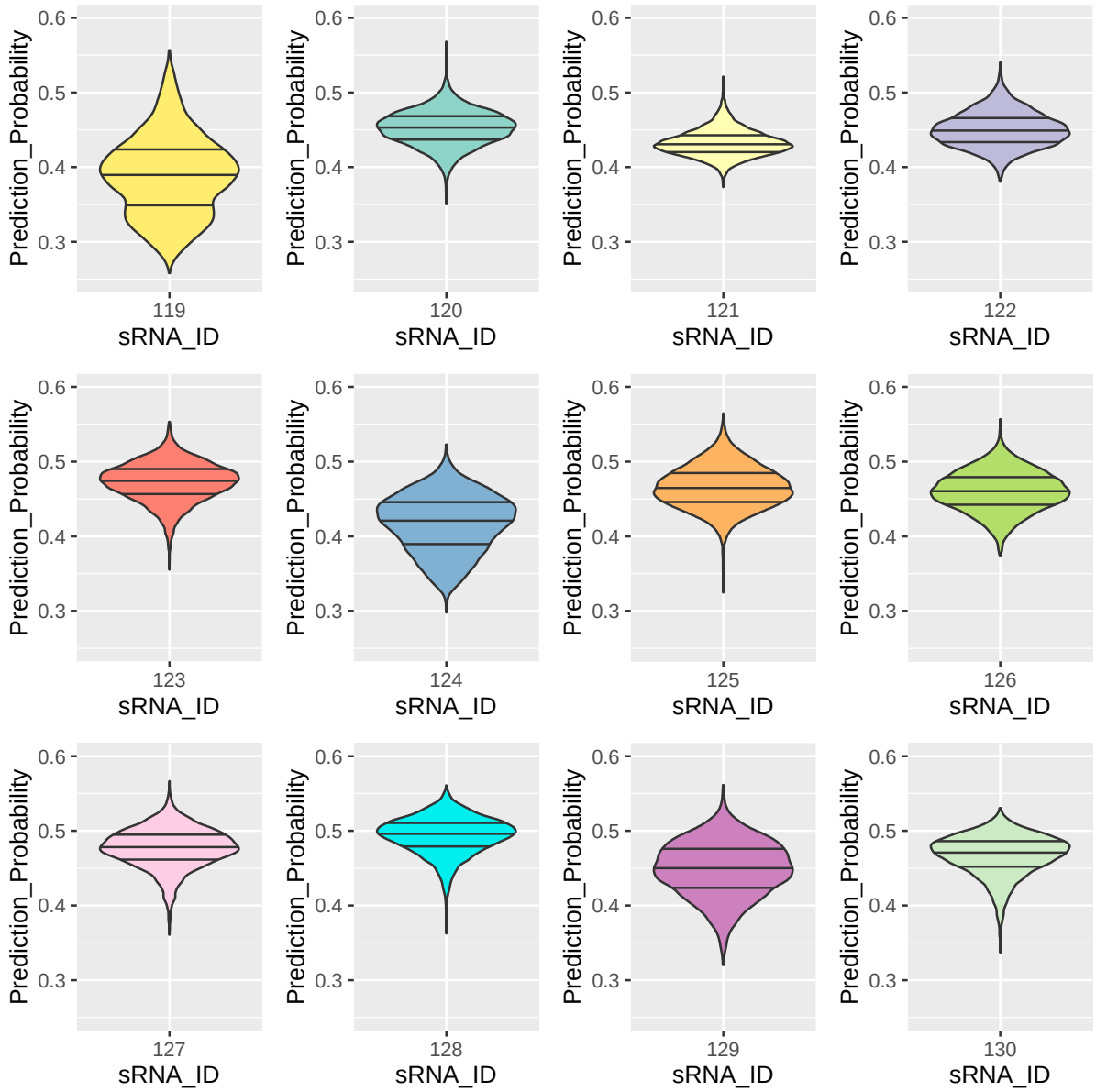

Prediction_Probability
0.6
0.5
0.4
0.3
59
60
61
62
63
64
65
66
67
68
69
70
sRNA_ID

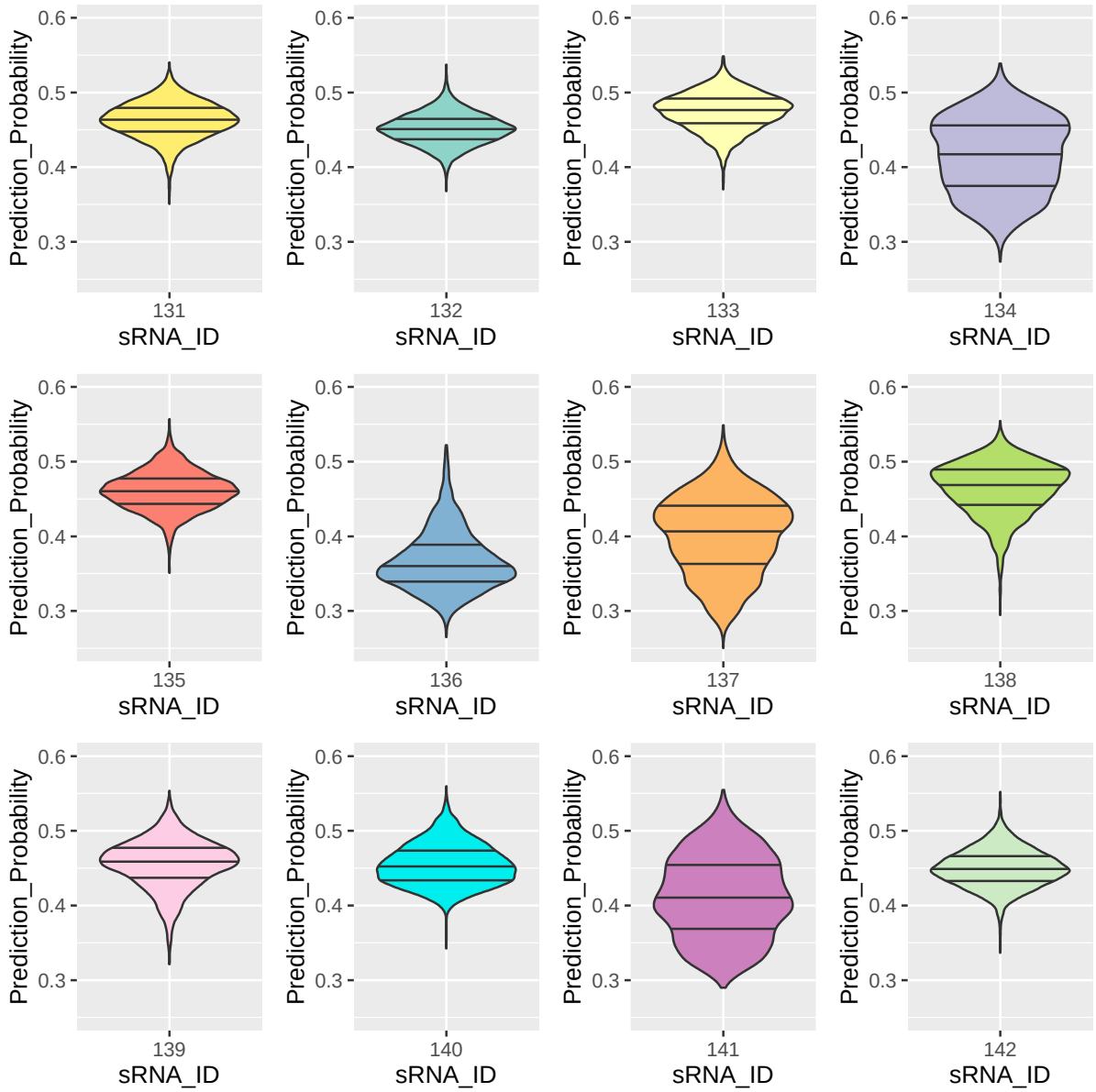

Prediction_Probability
0.6
0.5
0.4
0.3
71
72
73
74
75
76
77
78
79
80
81
82
sRNA_ID

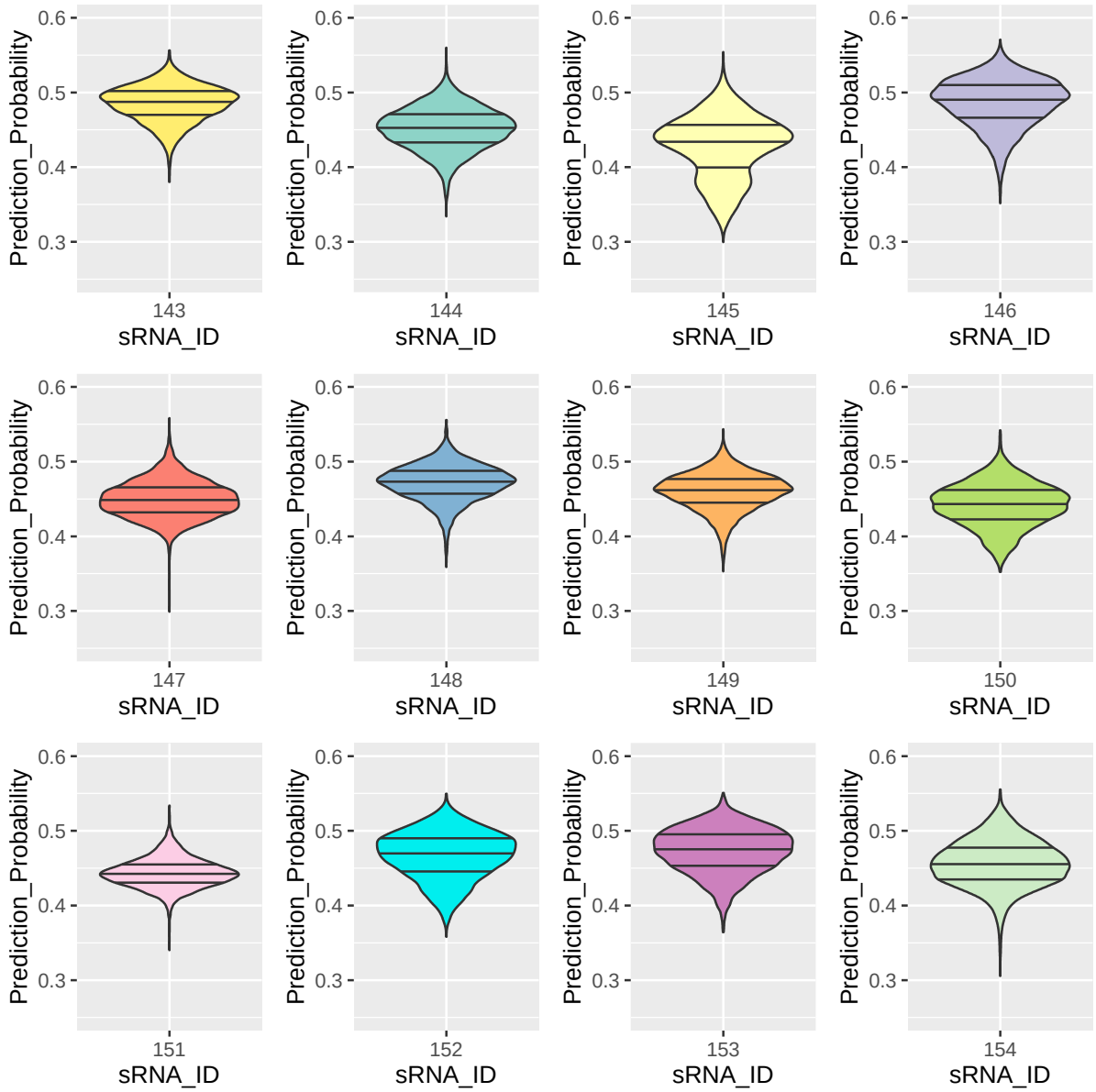

Prediction_Probability
0.6
0.5
0.4
0.3
83
84
85
86
87
88
89
90
91
92
93
94
sRNA_ID

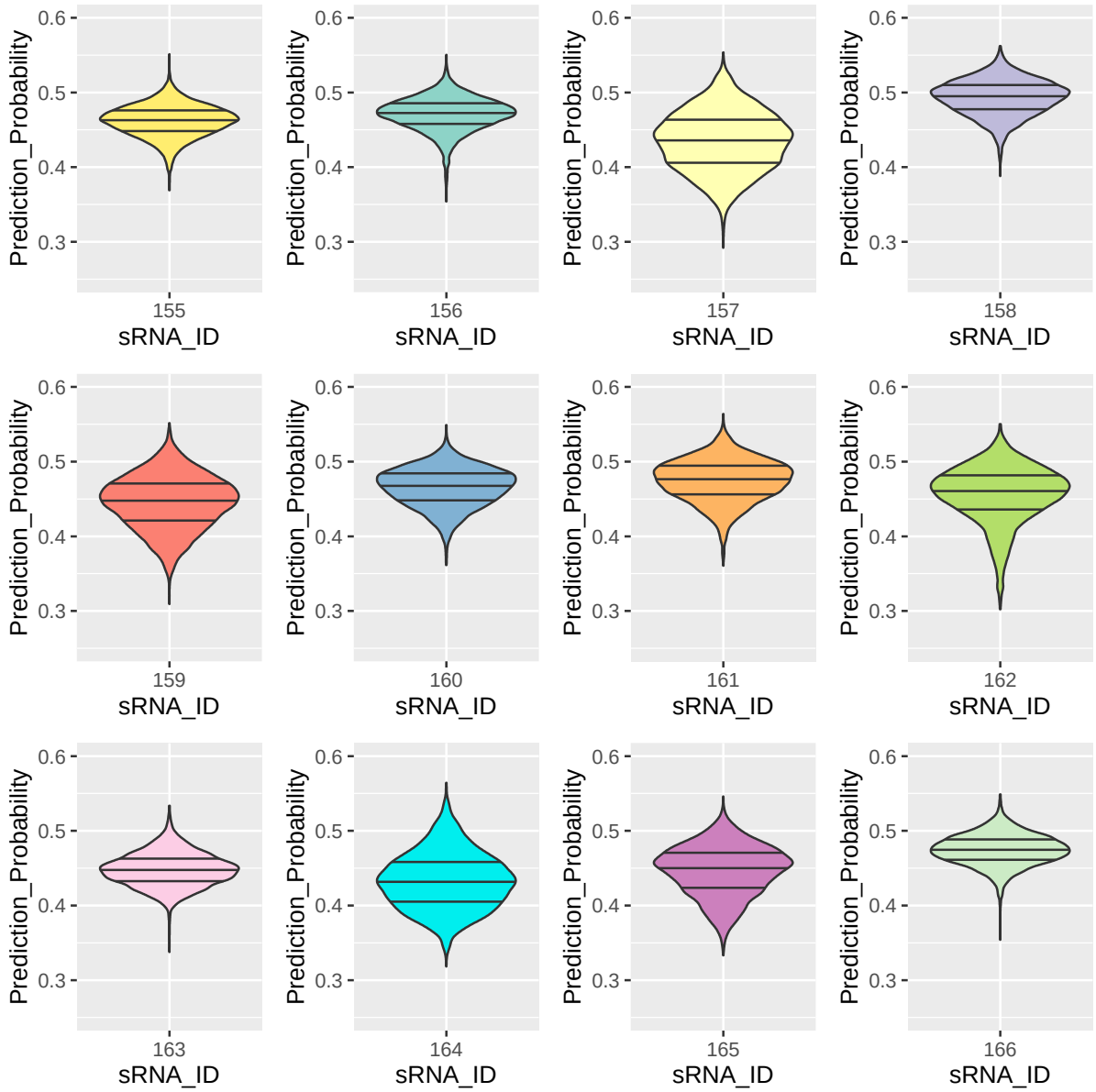

Prediction_Probability
0.6
0.5
0.4
0.3
95
96
97
98
99
100
101
102
103
104
105
106
sRNA_ID


Prediction_Probability
0.6
0.5
0.4
0.3
107
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
108
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
109
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
110
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
111
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
112
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
113
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
114
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
115
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
116
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
117
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
118
sRNA_ID


Prediction_Probability
0.6
0.5
0.4
0.3
119
120
121
122
123
124
125
126
127
128
129
130
sRNA_ID


Prediction_Probability
0.6
0.5
0.4
0.3
131
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
132
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
133
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
134
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
135
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
136
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
137
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
138
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
139
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
140
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
141
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
142
sRNA_ID


Prediction_Probability
0.6
0.5
0.4
0.3
143
sRNA_ID
144
145
146
147
148
149
150
151
152
153
154


Prediction_Probability
0.6
0.5
0.4
0.3
155
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
156
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
157
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
158
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
159
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
160
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
161
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
162
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
163
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
164
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
165
sRNA_ID
Prediction_Probability
0.6
0.5
0.4
0.3
166
sRNA_ID