The biodiversity of phytate cycling in soils

Gregory David Rix

A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy at the University of East Anglia, School of Biological Sciences

September 2021

Abstract

Phytic acid, *myo*-inositol hexakisphosphate, $InsP_6$, is the major storage form of phosphate in seeds and grains that constitute a major part of the diets of monogastric animals such as swine and poultry. Monogastrics lack enough enzymes in the right part of the gut to digest dietary phytate. Consequently, phytases, a group of enzymes capable of releasing inorganic phosphate from phytate, are added to commercial poultry and swine diets. These adjunct phytases are a major sector of the global enzyme market with an estimated value of $5 billion in 2015. There is continued commercial interest in the discovery and development of more effective and cost-effective enzymes.

The soil environment is microbially diverse and therefore offers significant potential for the isolation of novel phytases. In this thesis, I have developed new methods for the culture-dependent isolation of phytases from different soil environments by first analysing phytase activity of the soil microbiome using HPLC. The isolation of a multiple inositol polyphosphate phosphatase, MINPP, from *Acinetobacter* sp. represents one of the first phytases of its kind to be isolated from the soil environment. This study provides a robust characterisation of the protein, identifying an outstanding long-term stability at room temperature and activity from 37.6-101.3% over 755 days. The expression of the phytase was examined using β-galactosidase and qPCR assays which showed that while expression was enhanced in the presence of impure phytate by β-galactosidase, it was significantly repressed in the qPCR experiment. Additionally, a long-term phytase isolation experiment was performed using well-characterised Rothamsted soils. Of the sixty-six isolates that were re-streaked from mixed plates, seventeen showed a diverse array of phytase degradation profiles, highlighting the diversity of phytase activity in the soil environment. I have also undertaken metagenomic analysis to examine the diversity of phytases in environmental and enteric environments. This highlighted the dominance of MINPP genes in enteric environments above all other known classes of phytases. In soil and aquatic metagenomes, the relative abundances were significantly less than in enteric environments and here the Multiple Inositol Polyphosphate Phosphatase (MINPP) gene was not the overwhelming majority, instead the Beta-Propeller Phytase (BPPhy), Histidine Acid Phytase (HAPhy) and Protein Tyrosine Phosphatase-like Phytase (PTPhy) were equal to or even higher in abundances. Additionally, in this analysis I also examined the prospect of horizontal gene transfer in the MINPP dataset using the program T-rex. Here I identified multiple HGT events occurring between both enteric and environmental bacteria, with one transfer occurring between environments.

Acknowledgements

As this project finally draws to a close there are many people I would like to give my heartfelt thanks to.

I would first like to thank my supervisor Professor Charles Brearley for his never-ending depth of knowledge, advice, and guidance throughout this PhD process, it can't have been easy! I would also like to thank my co-supervisor Professor Jonathan Todd for his advice pertaining to all things genomic, I have learnt a lot, and I express my gratitude to you and the members of the Todd lab. As well as Professor Andy Neal from Rothamsted Research for his invaluable expertise in metagenome analysis and AB Vista for providing additional funding.

I would like to thank all the members of the Brearley lab for keeping me somewhat sane for the last four years. To Colleen, I couldn't have asked for a better lab mate, I had an absolute blast in the lab, and it was never boring, though I'm sure Charles was concerned for us on multiple occasions. And thanks to Dr Hayley Whitfield for your cool headedness in the lab and your patience.

I would like to thank each member of the Pamela Salter office, especially Tom McLean for being an absolute legend, you were an excellent teacher and a great friend. I would also like to thank everyone in the weekly football sessions especially Dr Darrell Green for organising it each week.

I should also thank current and former members of Speedwell Way, I couldn't have asked for better housemates over the last 4 years, you've all been amazing, in particular Sam and Bridie who somehow latest the full duration.

I would also be remiss if I did not include all my friends in the 'Sorrow' discord channel for the hours of entertainment, games, and D&D sessions, thank you Sam for introducing me to this group.

Last but certainly not least, I would like to thank my girlfriend Libby for putting up with me during this stressful final year and write-up period and finally to my parents and nan for all the overwhelming love and support for all of my time at university, thank you so much.

Publications arising from the work in this thesis

Rix GD, Todd JD, Neal AL, Brearley CA. (2020) Improved sensitivity, accuracy and prediction provided by a high-performance liquid chromatography screen for the isolation of phytase-harbouring organisms from environmental samples. *Microbial Biotechnology*. 14:4, p.1409-1421.

Rix GD, Sprigg C, Hemmings AM, Todd, JD, Brearley CA. (2021) Characterisation of a Soil MINPP Phytase with Remarkable Long-Term Stability and Activity from *Acinetobacter* sp. *In review.*

Abbreviations

A_{280} – Absorbance measured at 280 nm

Amp – Ampicillin

APS – Ammonium persulfate

ATP – Adenosine triphosphate

***Bl*MINPP** - *Bifidobacterium longum* susp. *Infantis ATCC 15698 MINPP*

BPPhy – Beta-propeller phytase

BSA – Bovine serum albumin

***Bt*MINPP** – *Bacteroides thetaiotaomicron* MINPP

Cam – Chloramphenicol

cDNA – complementary DNA

Ct – Cycle threshold

DNA – Deoxyribonucleic acid

dsDNA – double stranded DNA

EBI – European Bioinformatics Institute

EC – Enzyme Commission

EDTA – Ethylenediaminetetraacetic acid

ENA – European Nucleotide Archive

Gent – Gentamycin

GF – Gel filtration

HAPhy – Histidine Acid Phytase

HCl – Hydrochloric acid

HGT – Horizontal gene transfer

HPLC – High performance liquid chromatography

IMAC – Immobilised metal affinity chromatography

IMC – International Minerals and Chemicals Corporation

IMP – *myo*-inositol monophosphatase

$InsP_1$ – Inositol monophosphate

$InsP_2$ – Inositol bisphosphate

$InsP_3$ – Inositol trisphosphate

$InsP_4$ – Inositol tetrakisphosphate

$InsP_5$ – Inositol pentakisphosphate

$InsP_6$ – Inositol hexakisphosphate

IS_6 – Inositol hexasulfate

IP_3 – Inositol 1,4,5-trisphosphate

IPK – Inositol polyphosphate multikinase

IPTK – Inositol trisphosphate kinase

iTOL – Interactive tree of life

Kan – Kanamycin

K_{cat} – Enzymatic turnover number

kDA – kilodaltons

KEGG – Kyoto Encyclopaedia of Genes and Genomes

K_i – Inhibitor constant

K_m – Michaelis constant

KO – KEGG Orthology

LB – Lysogeny broth

MAFFT – Multiple alignment using FAST Fourier Transform

MINPP – Multiple inositol polyphosphate phosphatase

MPE – Metallophosphoesterase superfamily

MSA – Multiple sequence alignment

MUSCLE – Multiple Sequence Comparison by Log-Expectation algorithm

MW – Molecular weight

n/a – Not applicable

NaCl – Sodium chloride

NADP - Nicotinamide adenine dinucleotide phosphate

NADPH - Nicotinamide adenine dinucleotide phosphate hydrogen

NaOH – Sodium hydroxide

NASA – National Aeronautics and Space Administration

NCBI – National Center for Biotechnology Information

NMR – Nuclear magnetic resonance

O/N – Overnight

OD_{600} – Optical density measured at 600 nm

ONPG – *ortho*-nitrophenyl-β-galactopyranoside

PAGE – Polyacrylamide gel electrophoresis

PAPhy – Purple acid phytase

PCR – Polymerase chain reaction

PD – Proton donor

PDB – Protein databank

PEP – Phosphoenolpyruvic acid

P_i – Inorganic phosphate

pNPP – *para*-nitrophenylphosphate

PQQ – Pyrroloquinoline quinone

PSM – Phytase Specific Medium

PTFE – Polytetrafluoroethylene

PTPhy – Protein tyrosine phosphatase-like phytase

qPCR – Quantitative PCR

Rif – Rifampicin

RNA – Ribonucleic acid

RPM – Revolutions per minute

rRNA – ribosomal RNA

SDS – Sodium dodecyl sulfate

Spec – Spectinomycin

TEMED – Tetramethylethylenediamine

T_m – Melting temperature

T-rex – Tree and Reticulogram Reconstruction

Tris – 2-Amino-2-hydroxymethyl-propane-1,3-diol

U – Units

USGS – United States Geological Survey

VGT – Vertical gene transfer

V_{max} – Maximum rate of catalysis

WT – Wild-type

Contents

Contents 7

Chapter 1. Introduction 12

1.1. Phosphorus and Phosphate 12

1.1.1. A Brief History of Phosphorus 12

1.1.2. Phosphate Cycling in the Environment 13

1.1.3. Phosphate Rock: Uses, Future Issues, and Environmental Consequences 15

1.1.4. Overcoming phosphorus pollution 17

1.2. The Inositol Phosphates 17

1.2.1. *Myo*-Inositol Hexakisphosphate 17

1.2.2. Synthesis of Inositol hexakisphosphate 19

1.2.3. The Other Inositol Phosphates 21

1.2.4. The Issues Concerning *myo*-InsP_6 21

1.3. Overview of Phytases 23

1.3.1. The History of Phytases 23

1.3.2. Histidine Acid Phytases 24

1.3.3. β-propeller Phytases 27

1.3.4. Protein-Tyrosine Phytases 30

1.3.5. Purple Acid Phytases 32

1.4. Phytases in the Animal Feed Industry 34

1.4.1. The History of Phytases in Industry 35

1.4.2. The Economic and Environmental Benefit of Phytases 36

1.4.3. The Future of Phytases 37

1.5. Project Aims 38

Chapter 2. A Critical Analysis of the Phytase Isolation Methods in the Literature and Development of New Methods 39

2.1. Materials and Methods 40

2.1.1. HPLC analysis of inositol phosphates 40

2.1.2. Media preparations 40

2.1.3. Acid extraction of inositol phosphates from phytase-specific media plates 41

2.1.4. Preparations of Soil Cultures 41

2.1.5. Preparation of Soil Cultures for HPLC Analysis 41

2.1.6. The Microbial Communities' response to phytate degradation with and without a mixed carbon source 42

2.1.7. Culturing Isolates 42

2.1.8. Phytase activity from Church Farm Isolates without trace metals 42

2.1.9. 16S rRNA amplification 42
2.1.10. Gel Electrophoresis 43
2.1.11. Degenerate primer design 43
2.1.12. Genome Sequencing of strains 44
2.1.13. The growth of *Acinetobacter* sp. on different Carbon Sources 44
2.1.14. Nomenclature of the Inositol Phosphates 44
2.1.15. Measuring the phytate degradation potential of Broadbalk, bare fallow and arable soils from Rothamsted 45
2.1.16. Long-term phytase isolation experiment 45
2.1.17. Measuring the effect of soil and calcium on the adsorption of phytate 45
2.1.18. Calculating the absorption capacity of Broadbalk, Highfield bare fallow and Highfield ley arable soils through the addition of 10 mM phytate 46
2.1.19. Examining phytate degradation of liquid, aerobic and anaerobic layers of a salt marsh 46
2.2. Results and Discussion 47
2.2.1. Analysing the Pitfalls with regards to the Isolation of Phytases from the Environment 47
2.2.2. Review of the State of the Literature with regards to the pitfalls 51
2.2.3. Lack of standardisation on Phytase Reporting 53
2.2.4. Acid extraction of phytate from PSM plates 55
2.2.5. Assay of phytate degradation by mixed population soil cultures 57
2.2.6. Variation of $InsP_6$ Degradation by Microbial Cohorts When $InsP_6$ is the Sole Source of Carbon and Phosphate. 61
2.2.7. Isolation of phytase-producing microorganisms from the environment 63
2.2.8. Identification of phytases present in *Acinetobacter sp.* and *Buttiauxella* sp. 67
2.2.9. Evolutionary differences between *Acinetobacter* sp. AC1-2 and *Buttiauxella* sp. CH 10^{-6}-4 Ac. 72
2.2.10. *Acinetobacter* sp. AC1-2 – Bacterial Pathways 73
2.2.11. *Acinetobacter* sp. AC1-2 growth on different carbon sources 76
2.3. Using culture-dependent techniques to investigate diversity of phytases amongst soil and salt marsh environments 78
2.3.1. Measuring phytate degradation in Rothamsted soils 78
2.3.2. Isolating phytases from Rothamsted soils 80
2.3.3. Identifying Inositol Phosphatase Activity 86
2.3.4. Isolate sequencing 88
2.4. Examining the adsorption of phytate onto soil 90
2.5. Analysing the phytate degradation in aerobic, anaerobic and water sediments from Blakeney Iron and Sulphur cores. 93

2.6. Conclusion 102
Chapter 3. The Purification of a Multiple Inositol Polyphosphate Phosphatase from *Acinetobacter* sp. AC1-2 105
3.1. Materials and Methods 105
3.1.1. Phenol-chloroform method to extract genomic DNA. 105
3.1.2. Plasmids and Bacterial Strains 106
3.1.3. Gateway cloning of the *Acinetobacter* sp. AC1-2 *Minpp* gene 108
3.1.4. Generation of the Entry Clone 109
3.1.5. Integration of the Entry Clone into the Entry Vector (BP Reaction) 110
3.1.6. Transformation of the Entry Vector into DH5α cells 110
3.1.7. Confirming correct Entry Vector Inserts 111
3.1.8. Assessment of Sequencing Quality 111
3.1.9. Transformation of the Entry Clone into the Destination Vector 111
3.1.10. Transformation of the Expression Clone into the Expression Host 111
3.1.11. Small Scale Expression and Protein Purification 112
3.1.12. SDS-PAGE Gel Electrophoresis 112
3.1.13. Bradford Assay 113
3.1.14. Measuring Phytase Activity using the Molybdenum Blue Method for Phosphate Release 113
3.1.15. Large Scale Protein Purification Prep 114
3.1.16. Optimization of the Solubility and Purity of Minpp AC1-2 115
3.1.17. Cloning into sHuffle® T7 express and ArcticExpress (DE3) 115
3.1.18. Small Scale Expression Trials of AC1-2 no signal construct 116
3.1.19. Transformation and Expression Trials in *Pichia pastoris* 117
3.2. Results and Discussion 119
3.3. Conclusion 135
Chapter 4. The Characterisation of a Multiple Inositol Polyphosphate Phosphatase from *Acinetobacter* sp. AC1-2 137
4.1. Materials and Methods 138
4.1.1. Phytase Activity Assays 138
4.1.2. pH Profile 138
4.1.3. Time course 138
4.1.4. Inhibition 138
4.1.5. Substrate specificity towards phosphate esters 139
4.1.6. Thermostability 139
4.1.7. Long-term stability assay 139

4.1.8. Effect of storage buffer on long-term stability 139
4.1.9. Michaelis-Menten parameters of AC1-2 MINPP 139
4.1.10. Structural Biology 140
4.2. Results and Discussion 140
4.2.1. Characterisation of AC1-2 MINPP 140
4.2.2. Modelling of AC1-2 MINPP 155
4.3. Conclusion 161
Chapter 5. Using culture independent techniques to investigate the diversity of phytases amongst soil and other environments. 162
5.1. Materials and Methods 165
5.1.1. Preparing a diverse MINPP dataset for metagenomic studies 165
5.1.2. Metagenome Selection 168
5.1.3. Metagenomic Pipeline 170
5.1.4. Refinement of the dataset to examine horizontal gene transfer (HGT). 170
5.2. Results and Discussion 173
5.2.1. Examining the distribution of MINPP genes in the environment 173
5.2.2. The placement of MINPP genes on the phylogenetic tree 177
5.3. Examining the diversity of phytase genes in the environment 184
5.4. Examining the MINPP dataset for evidence of Horizontal Gene Transfer 189
5.4.1. HGT events 195
5.5. Conclusion 209
Chapter 6. Control of phytase expression in *Acinetobacter* sp. 210
6.1. Materials and Methods 211
6.1.1. Generation of β-galactosidase Promoter-Reporter constructs for phytase expression 211
6.1.2. Ligation of promoter regions with the pS4 vector 212
6.1.3. Preparation of competent *E. coli* 803 cells 213
6.1.4. Triparental Mating – Patch Cross 214
6.1.5. β-Galactosidase Assay 215
6.1.6. Generation of an *Acinetobacter* MINPP gene knockout using a double crossover technique 216
6.1.7. Ligation of flanking regions 1 and 2 into the pK18mobsacB plasmid. 218
6.1.8. Triparental mating – Filter Cross 220
6.1.9. Confirming gene knockout 220
6.1.10. Rif-mutant and MINPP Knockout soil survivability experiment 221
6.1.11. Measurement of phytase expression by quantitative PCR 221

6.1.12. RNA extraction and quantification 222
6.1.13. Dnase Treatment 222
6.1.14. cDNA synthesis 222
6.1.15. qPCR primer design 223
6.1.16. Quantitative PCR 223
6.1.17. Data analysis 224
6.2. Results and Discussion 224
6.2.1. Promoter-reporter constructs 224
6.2.2. Induction of phytase expression in the presence of phytate and inorganic phosphate 226
6.2.3. Regulation of AC1-2 *MINPP* expression, assessed by reverse transcriptase quantitative PCR (RT-qPCR) 228
6.2.3.1. Calculating qPCR primer efficiencies. 229
6.2.3.2. qPCR analysis of *Acinetobacter* sp. *MINPP* expression in LB and LB + $InsP_6$ liquid cultures 231
6.2.3.3. Calculation of Fold Change between the LB and LB + $InsP_6$ treatments 233
6.2.3.4. qPCR analysis of *Acinetobacter* sp. *MINPP* expression in Minimal media (MM), MM + $InsP_6$ and MM + P_i treatments 234
6.2.3.5. Calculation of Fold Change between the MM, MM + $InsP_6$ and MM + P_i treatments. 235
6.2.4. Generation of an *Acinetobacter* ΔMINPP knockout 237
6.2.5. Competition experiment to compare performance of Rif-resistant *Acinetobacter* sp. and *Acinetobacter* sp. Δ*MINPP* in soil. 240
6.3. Conclusion 243
6.4. Supplementary Information 245
Chapter 7. Final Discussion 255
References 259

Chapter 1. Introduction

1.1. Phosphorus and Phosphate

1.1.1. A Brief History of Phosphorus

The element phosphorus is required by all life on Earth (Ahemad *et al*., 2009). It is however unusual when compared with the other fundamental elements for life such as carbon, nitrogen and oxygen in that it is not found as a free element by itself (Desmidt *et al*., 2015). Instead, it is commonly found in the form of phosphate.

Phosphorus was first discovered by the German alchemist Hennig Brandt who managed to isolate it from urine while searching for the legendary "Philosopher's Stone" (Sharpley *et al*., 2018), becoming the first element discovered using modern scientific techniques, partly because it is so reactive in elemental form that it never occurs freely in nature. Following this, in the 17th and 18th century, phosphorous was unsuccessfully used for a variety of questionable medicinal purposes (Ashley *et al*., 2011). In 1771, a more suitable method for obtaining phosphorous was discovered. Scheele and Gahn, two Swedish scientists found that bones contained a more abundant source of phosphorus in the form calcium phosphate which is still used today in the form of bone meal for organic fertilisers (Pierrou, 1976).

It was not until the mid-19th century when significant evidence towards the roles of phosphate in the environment was discovered. Liebig, recognised by many as the father of agricultural chemistry, alongside the work of his colleague Sprengel, identified the critical roles that nitrogen, phosphorus and potassium play in plant growth, dubbed "The Theory of Mineral Nutrition" (van der Ploeg *et al*., 1999). This replaced the "Vitalistic Humus Theory' which had previously been suggested, where plants feed on decaying plants and animals as part of an organic cycle (Nath, 1940).

Soon after, Lawes, an English entrepreneur, experimented on the effects of various manures on the growth of plants, and in 1842 he patented an artificial fertiliser formed by treating crushed bones, rich in calcium phosphate, with sulfuric acid. In the succeeding years he enlisted the services of Gilbert, who worked under Liebig, and together founded Rothamsted Research, to undertake long-term trails of the effectiveness of mineral and organic fertilisers on crop yields (Ivell, 2012). In 1906, researchers Harden and William John Young discovered that salts of orthophosphoric acid stimulate alcoholic fermentation of sugar in yeast (Föllmi, 1996). They were able to develop an apparatus for collecting and measuring gases evolved during fermentation leading to the discovery of fructose-1,6-diphosphate. The first chemical

intermediate discovered in fermentation (Harden *et al*., 1911; Harden and Young, 1913), further demonstrating the important role phosphorus plays in the lifecycle of plants and eukaryotes.

The boom of phosphate, however, did not occur until post-World War II, where the use of mineral phosphate sources grew exponentially. Rich and accessible sources of phosphate were found in phosphate rock, used to refer to apatite minerals, which can contain as much as 10-20% phosphate on the basis of weight (Cooper *et al*., 2011). Phosphate rock mining was significantly increased to keep up with the nitrogen fertiliser demand caused by the invention of the Harber-Bosch Process. This led to a sextuple increase in fertiliser use between 1950 and 2000 (Ashley *et al*., 2011; Smil, 2000). This 'Green Revolution' saw the feeding of billions globally (Cullather, 2004). Since then, phosphate has been recognised as one of the key substrates in energy metabolism, biosynthesis of nucleic acids and membranes (Komaba and Fukagawa, 2016). It also plays a very important role in cell signalling, respiration and photosynthesis (Berridge and Irvine, 1989).

The importance of phosphorus cannot be understated. As was once quoted by Issac Asimov.

"Life can multiply until all the phosphorus is gone, and then there is an inexorable halt which nothing can prevent. We may be able to substitute nuclear power for coal, and plastics for wood, and yeast for meat, and friendliness for isolation – but for phosphorus there is neither substitute nor replacement".

1.1.2. Phosphate Cycling in the Environment

Despite being the 11th most abundant element on the planet, phosphorus is considered to be one of the least abundant plant nutrients in the soil (Maciá, 2005; Pereira and Castro, 2014), with approximately 5.7 billion hectares of land worldwide deficient for optimal plant growth (Raghothama and Karthikeyan, 2005). This issue is particularly common in Africa, where the production of crops without fertiliser results in continuous removal of nutrients in the soil without replacement (Sanchez, 2002).

One of the reasons for this is, despite the presence of phosphate in these soils, it occurs in organic and inorganic forms which are unavailable for root uptake (Sharma *et al*., 2013); with as little as 0.1% of the total phosphate available to plants for uptake (Molla *et al*., 1984; Zou *et al*., 1992). Inorganic phosphate is adsorbed by a range of iron and aluminium containing compounds such as goethite/boehmite, fixing the phosphate in the solid soil phase (McLaughlin *et al*., 1981). It can also bind organic compounds such as soil colloids which can

be made up of humus and clay (Jiao *et al*., 2007). In addition to soil environments, phosphate plays a key role in marine environments. In nearly all modern aquatic systems, growth and the primary production of organic matter is thought to be limited by either phosphorus or nitrogen (Howarth, 1988; Planavsky *et al*., 2010). Whilst nitrogen is obtained through nitrogen fixation from the abundant gases in the environment, phosphorus does not have a gaseous cycle and therefore its global cycle is remarkably different. In fact unlike the other fundamental elements for life on earth, hydrogen, carbon, nitrogen, oxygen and sulfur, phosphorus cannot be found in any significant gaseous state (Jahnke, 1992).

Without a gaseous phase, the global phosphorus cycle occurs through the land and water. The main source of phosphorus is from the weathering of continental materials (Föllmi, 1996; Planavsky *et al*., 2010). The cycle can be described in 4 major points. (i) Tectonic uplift, causing the exposure of phosphate-containing rocks to weathering, (ii) physical and chemical weathering of rocks, producing soils and dissolved phosphorus into waterways, (iii) transport of phosphorus along rivers and into lakes and oceans and (iv) the sedimentation of phosphorus associated with minerals and organic matter. The cycle begins again with the uplift of sediments to weathering (Ruttenberg, 2003; Smil, 2000). This cycle completion takes place over long, geologic timescales, 10^7-10^9 years (Ruttenberg, 2003), therefore the phosphorus cycle is often regarded as the slowest biogeochemical cycle. (Daneshgar *et al*., 2018).

However, there are many smaller, more localised cycles of phosphate in both marine and terrestrial environments. This involves the life cycle of living organisms such as bacteria, phytoplankton, and plants. A large share of the nutrients that have been taken up are returned back to their environments through the death and decay of these organisms (Vadstein, 2000). This, however, does cause issues with farming as the phosphorus taken up by the growing crops are not returned to their local environment, causing a break in the phosphorus cycle. Interestingly, the deposition of phosphate-containing dust has been investigated as another method of phosphorus cycling. For many years now, the presence of African dust in the Amazon has been observed (Baars *et al*., 2011). Recently, with the help of NASA and satellite surveillance, they managed to track the trans-Atlantic journey of phosphate-containing dust travelling from ancient lake beds in Africa to the Amazon rainforest (H. Yu *et al*., 2015).

Altogether, there are many issues involving the global cycling of phosphorus, as stated previously soils worldwide are deficient in phosphate (Raghothama and Karthikeyan, 2005)

and therefore for farming purposes have been intensively supplemented with fertilisers over the last hundred years.

1.1.3. Phosphate Rock: Uses, Future Issues, and Environmental Consequences

Originally, organic sources of phosphorus were used such as crushed bones, guano and animal manure however, the necessary amounts required to provided sufficient phosphorus meant that their application was economically unfeasible (Chien *et al*., 2011). Therefore, phosphate rock is commonly used in favour of organic alternatives.

Phosphate rock is an apatite mineral that contains as much as 10-20% phosphate by weight. Commonly found in sedimentary and igneous deposits (Walan *et al*., 2014), phosphate rock is a non-renewable resources that has taken tens of millions of years to form (Cordell and White, 2014). Therefore, it is in a similar situation to oil whose "death" is frequently heralded in the media (Becken, 2014; Curtis, 2009). Phosphorus depletion however doesn't receive as much public attention. Like oil, the distribution of these phosphorus reserves is only localised to a handful of locations known as phosphorite giants. The leader in which is the country of Morocco which contains the bulk of the reserves with nearly 77% of global supply (Cooper *et al*., 2011), Figure 1.1.

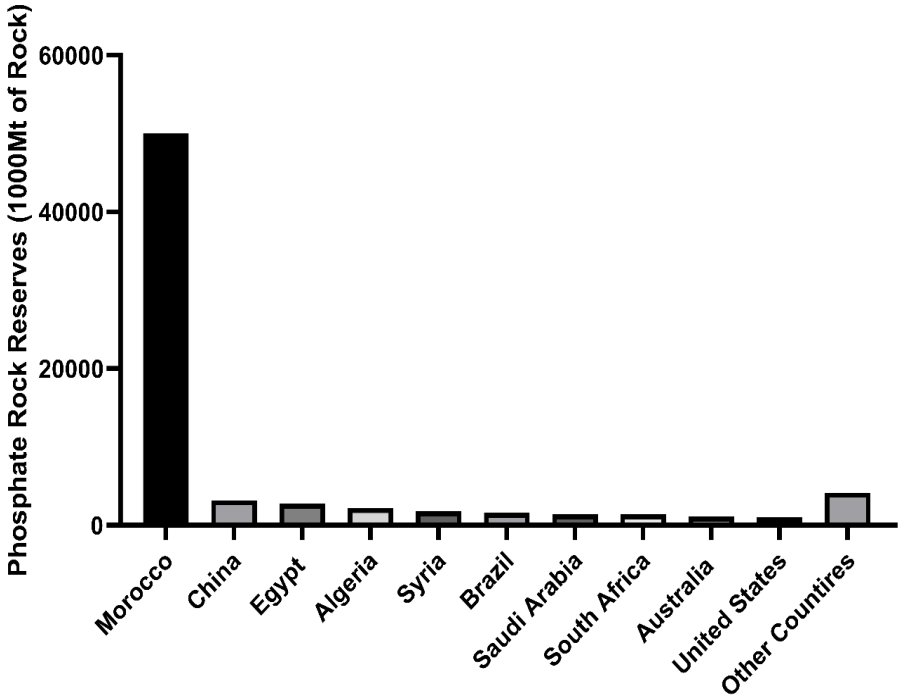


Figure 1.1 – The distribution of world phosphate rock reserves. Drawn from data published by the United State Geological Survey (Stephen Jasinski, 2021).

According to the United States Geological Survey (USGS), as of 2021, the current global phosphate rock reserves stand at 71,000 Mt with mined phosphate rock for the year standing at 223 Mt. The potential depletion of phosphate rock has been a hotly debated topic amongst the scientific community with an array of different predictions of peak phosphate, the hypothetical point in time in which the global phosphate production reaches its maximum rate, after which production will gradually decline, predicted anywhere from 2030-2136 (Cordell and White, 2011; Walan *et al*., 2014). It should be noted that for papers before 2010, predictions will be considerably lower as Moroccan estimated reserves increased from 5700 Mt to 51,000 Mt. The International Fertilizer Development Centre predicts that full depletion of phosphate rock could be between 3-400 years (Cordell and White, 2014). While there is much debate to be had on the lifetime of phosphate rock reserves, there are several clear issues. It has been accepted that phosphate rock reserves are decreasing in grade (%P_2O_5) (Cordell *et al*., 2011; Schröder *et al*., 2011; Van Vuuren *et al*., 2010), higher in contaminants such as Pb and U (Ma and Rao, 1997; Sabiha-Javied *et al*., 2009) and more difficult to access and extract (Cordell *et al*., 2009), making the continued future of phosphate rock unsustainable.

This is however, ignoring the issues of today.

Of the phosphate rock mined, only one fifth of it makes it from mine to human consumption, with significant loss of phosphorus at each step, from mining, extraction, application and harvest, livestock rearing and finally onto the plate (Cordell *et al*., 2012). As described in Section 1.1.2, most of the inorganic phosphate in the soil is bound in inaccessible forms for plant uptake. This is also the case when phosphorus is added in the forms of fertilisers (Bhattacharyya *et al*., 2015). The added phosphorus is readily adsorbed, with as much as 80% of the applied phosphorus becoming unavailable for uptake shortly after application (Zhu *et al*., 2018). In order to provide enough phosphorus for optimal crop growth, this had led to the over supplementation of phosphate fertilisers (Li *et al*., 2011; Woo *et al*., 2012). This, however, has led to serious environmental ramifications. The excessive application of nutrients causes surface run-off and leaching that can contaminate ground or surface waters (Paik, 2001) flowing into freshwater lakes and rivers. This stimulates the rapid growth of algae and water plants; which have disastrous effects to the local environment. Many algae blooms can produce bioactive compounds, including potent toxins that can cause livestock and wildlife death (Trevino-Garrison *et al*., 2015). Additionally, non-toxic algae blooms, although not producing any dangerous toxins can still have a significant effect. These become so densely concentrated that they create anoxic conditions resulting in the death of any fish

or underwater creatures (Hallegraeff, 1992). The death and decay of these blooms impacts water quality, restricting use for fisheries, recreation, industry and drinking (Khan and Mohammad, 2014). These are not however, isolated to freshwater. Eventually, the phosphorus reaches the coastal marine ecosystem. With approximately 40% of the World's population within 100 km of the coastline and providing more than $10 trillion in annual resources these blooms are contributing to significant hypoxic zones across the globe causing economic consequences (Wallace *et al*., 2014).

1.1.4. Overcoming phosphorus pollution

Ninety percent of total phosphorus consumption is used in both crop farming and animal husbandry (Neset and Cordell, 2012; Scholz *et al*., 2013). Most of it, however, is not absorbed by either. This has led to the over supplementation of phosphorus in both fertilisers and animal feeds (Bhattacharyya *et al*., 2015; Kebreab *et al*., 2008) causing significant environmental damage. Attaining sustainable supply/reuse of phosphorus has never been more critical. The "wake-up" call came in 2008, where rock phosphate prices spiked dramatically (Cordell and White, 2015, 2013), with the price going from $262 to $1218 per tonne. This price rise was driven by a combination of factors, an imbalance between supply and rapidly increasing demand, particularly in China and India. Increased biofuels production in the U.S. and Europe as well as increased livestock production, thereby requiring more grain and thus more fertiliser. Economically, the price was also affected by the imposition of high tariffs by China on fertiliser exports. This emphasised the fragility of the world's food supply to even minor disruptions (Cordell and White, 2015, 2014) and the importance of developing sustainable approaches that can minimise the use of rock phosphate

One such approach, involves the abundant organic phosphate molecule phytate (Inositol hexakisphosphate, $InsP_6$, IP_6).

1.2. The Inositol Phosphates

1.2.1. *Myo*-Inositol Hexakisphosphate

The inositol phosphates are a class of organic phosphate compounds that are widely found in nature, including in animals. These are derived from an inositol ring where the alcohol groups are replaced with phosphate. These range from inositol monophosphate to inositol hexakisphosphate.

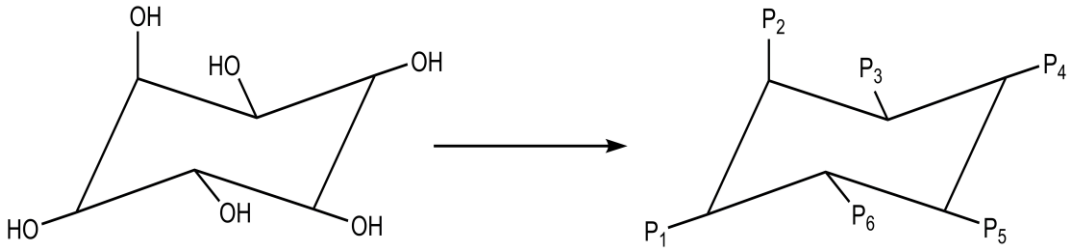


Figure 1.2 – The structures of *myo*-inositol and *myo*-inositol hexakisphosphate. Where each P represents an ester bonded orthophosphate, -OPO_3^{2-}. Figures were drawn in Chemdraw.

There are nine different stereoisomeric forms (*myo-, neo-, scyllo-, d-chiro-, L-chiro-, epi, muco, allo* or cis-$InsP_6$) that differ in their axial or equatorial conformation of phosphates. Of these, only four has been confirmed in nature, these are *myo-, neo-, scyllo-* and *d-chiro-*, Figure 1.3 (Giles *et al*., 2011).

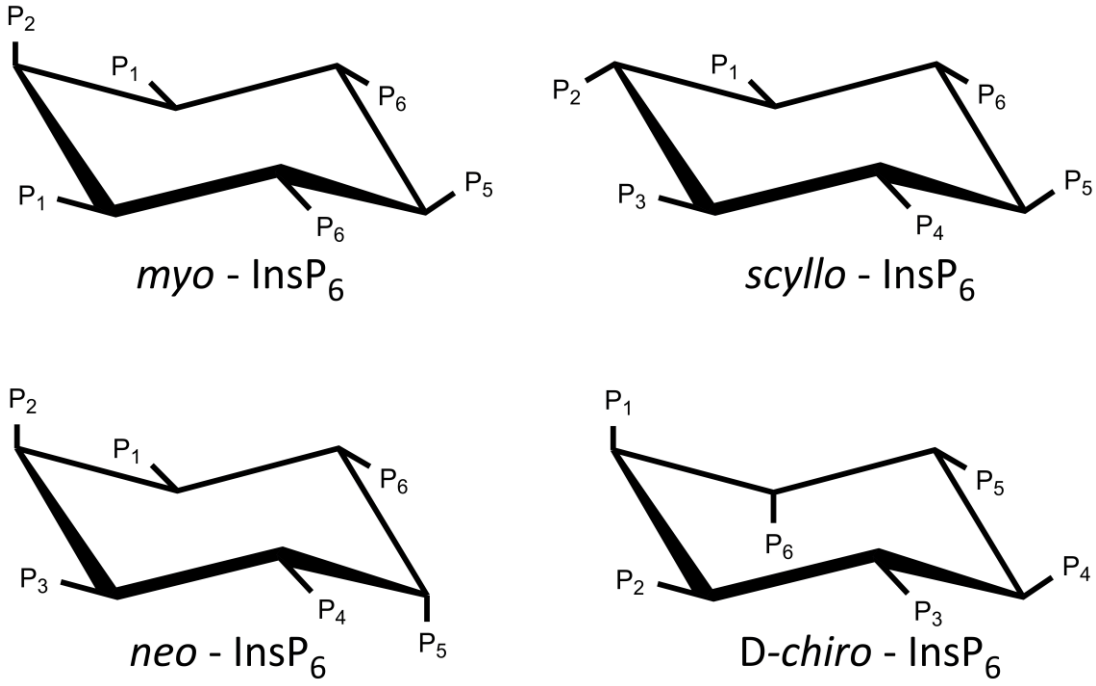


Figure 1.3 – The four naturally occurring inositol hexakisphosphates. *myo*-inositol hexakisphosphate contains five equatorial phosphates in the 1-, 3-, 4-, 5-, 6- positions with one phosphate in the axial 2-position, *scyllo*-$InsP_6$ the phosphates are all in the equatorial position, *neo*-$InsP_6$ has four equatorial phosphate groups in the enantiomeric positions, 1/3 and 4/6 with two axial groups in the 2-, 5- position. D-*chiro* also has four equatorial phosphates at the 2-, 3-, 4- and 5-positions and two axial groups at the 1- and 6-positions.

Of the four naturally occurring inositol hexakisphosphates, *myo*-$InsP_6$ garners the most attention as it is the primary storage form of phosphate in plants, seeds and grains (Kumar *et al*., 2010; Raboy and Dickinson, 1993), accounting for as much as 1-5% of dry weight and

up to 60-82% total phosphate (Ravindran *et al*., 1994; Reddy and Sathe, 2001). In developing seeds, more phosphorus is accumulated than required for cellular function. In higher plants this excess phosphorus is incorporated into $InsP_6$ and deposited as mixed salts referred to as phytin (Raboy, 2009).

1.2.2. Synthesis of Inositol hexakisphosphate

There are two biosynthetic pathways for the synthesis of inositol hexakisphosphate (Gillaspy, 2013). The first pathway is called the lipid-dependent pathway that is present in all eukaryotic organisms. In this pathway, phosphatidylinositols (PtdIns), a family of lipids with an inositol head group (Freed *et al*., 2020) specifically phosphatidylinositol 4,5-bisphosphate, PIP_2, is acted upon by the enzyme phospholipase C (PLC) producing $Ins(1,4,5)P_3$. This is then acted upon by specific inositol kinases, which catalyse the stepwise phosphorylation of free hydroxyls on the inositol ring. For example, in *Arabidopsis*, Inositol polyphosphate multikinase, IPK2, catalyses the conversion $Ins(1,4,5)P_3$ to $Ins(1,3,4,5,6)P_5$, after which Inositol Pentakisphosphate-2-kinase, IPK1, generates the $InsP_6$ molecule (Cridland and Gillaspy, 2020). Their importance was confirmed by knocking out both IPK1 and IPK2. A single IPK1 knockout saw a 83% reduction of seed phytate levels which increased to >95% when also knocking out IPK2 (Stevenson-Paulik *et al*., 2005). Similar processes have been observed in flies, human and rat homologs (Monserrate and York, 2010). Additionally, there is another family of enzymes that are involved in the phosphorylation of the InsPs, namely Inositol trisphosphate Kinases, ITPK1-4), which can catalyse the conversion of $InsP_3$ to $InsP_4$ and $InsP_4$ to $InsP_5$ (Monserrate and York, 2010).

The alternative pathway is the lipid-independent pathway. This pathway can initially proceed via two different enzymes. In the first path, inositol is first acted upon by the enzyme *myo*-inositol kinase (MIK) which converts the inositol to inositol (1,4,5)-phosphate. The next pathway converts glucose 6-phosphate to inositol (3)-phosphate using the enzyme *myo*-Inositol (3)-phosphate-synthase (MIPS) (Shi *et al*., 2005), which can be converted to *myo*-inositol by *myo*-inositol monophosphatase (IMP) or phosphorylated to *myo*-inositol bisphosphate. Mutants in both the maize *MIK* and maize and barley *MIPS* genes exhibited a 50% and 30-50% reduction in seed $InsP_6$ levels respectively (Dorsch *et al*., 2003; Shi *et al*., 2005). The conversion of $InsP_1$ to $InsP_2$ and $InsP_3$ has not been completely deciphered yet. The ITPKs have been shown to have a high catalytic flexibility and therefore may be involved

in this process (Laha *et al*., 2019), however this is not confirmed and still debated. A general schematic of both pathways is displayed below in Figure 1.4.

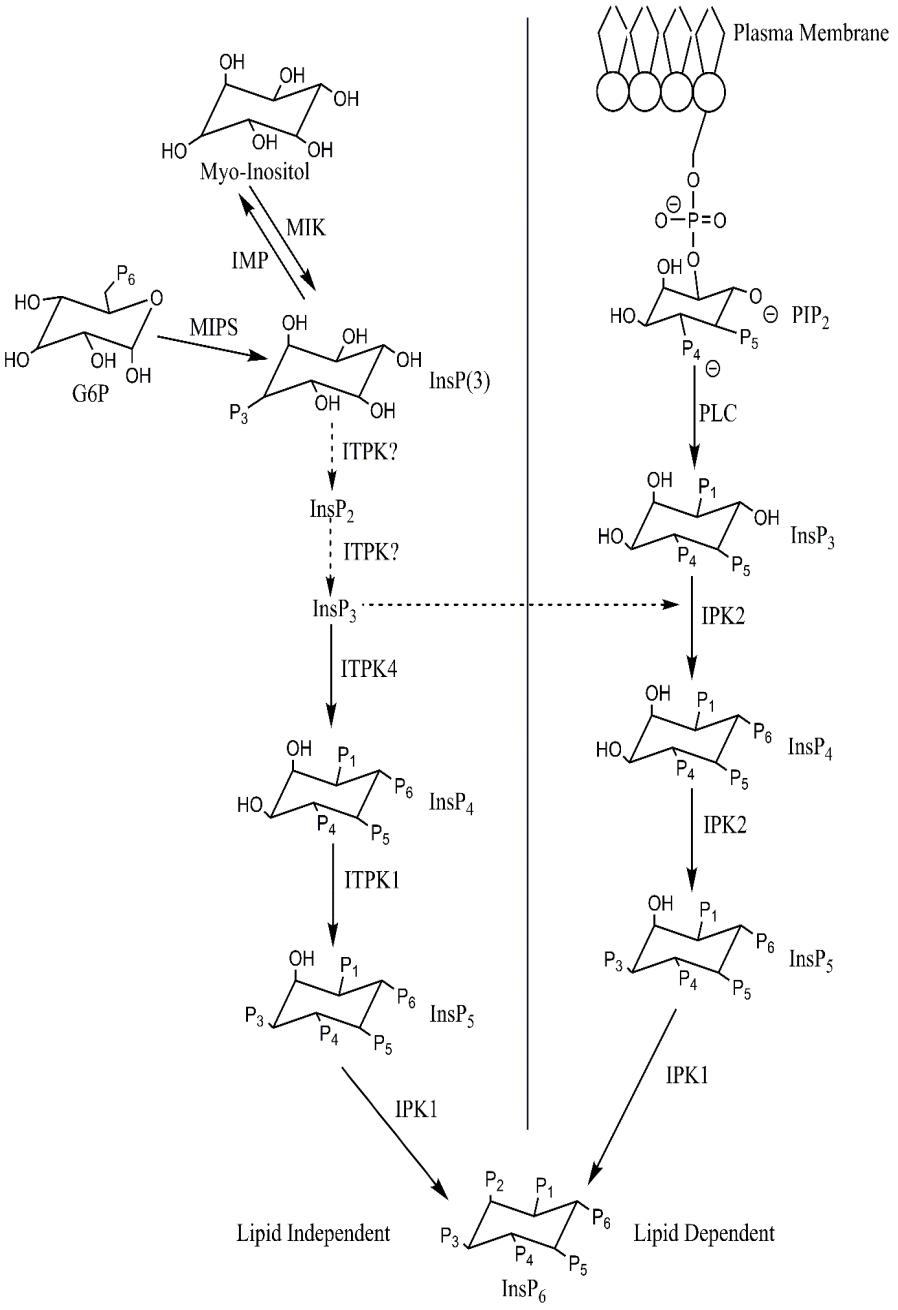


Figure 1.4 – A schematic detailing the lipid dependent and lipid independent pathways for $InsP_6$ synthesis. The lipid-dependent pathway proceeds through breakdown of phosphatidylinositol (PIP_2) by phospholipase C (PLC). It is then acted upon by the Inositol phosphate kinases (IPKs) which sequentially phosphorylate to $InsP_6$. The lipid-independent pathway can occur through the phosphorylation of *myo*-inositol by *myo*-inositol kinase (MIK) to Ins3P or by conversion of G6P to Ins3P. The next steps have yet to be fully elucidated but may occur through Inositol trisphosphate kinases (ITPKs). These can phosphorylate the InsP molecule similarity to the IPKs.

1.2.3. The Other Inositol Phosphates

The $InsP_6$ pathway is fundamentally important to a variety of biological processes. The structure of the six-membered cyclohexane ring with the potential of varying levels of phosphorylation, 64 unique phosphorylation molecules (York, 2006), allows a diverse range of compounds to be produced and used in a variety of different biological processes, the diversity can be further increased by the attachment of lipid groups forming phosphoinositides or the addition of phosphates to existing phosphate-monoesters forming inositol pyrophosphates, $InsP_7$ and $InsP_8$ (Livermore *et al*., 2016).

The best characterised inositol phosphate other then $InsP_6$ is the calcium signalling molecule $Ins(1,4,5)P_3$ ($InsP_3$). $InsP_3$ binds to IP_3-dependent calcium channels causing the release of calcium from intracellular stores. This signalling molecule plays an important role in many biological processes, including cell proliferation, apoptosis and muscle contraction (Berridge, 2008; Hanson *et al*., 2004; Taylor and Thorn, 2001). Concentrated at the cytosolic surface of membranes, phosphoinositides have key roles in signal transduction, regulation of membrane traffic and the permeability and transport functions of membranes (Di Paolo and De Camilli, 2006). Inositol pyrophosphates are small energy-rich molecules that have been ubiquitously found in all eukaryotes. These are involved in a wide range of cellular functions, including DNA repair, osmoregulation, immune signally and ribosome synthesis (Thota and Bhandari, 2015; Wilson *et al*., 2013).

1.2.4. The Issues Concerning *myo*-$InsP_6$

As shown in Figure 1.4, the $InsP_6$ pathway and the family of inositol phosphates are involved in a variety of biologically important processes. $InsP_6$ is abundantly found in animal cells, at concentrations ranging from 10-60 μM. Plants, seeds, and grains are the major storage form of phosphate. Likewise, inositol phosphates are abundant in the soil environment, forming a large part of the organic soil phosphate pool, being anywhere from 10-80% of the total organic phosphate (Giles *et al*., 2011).

$InsP_6$ is an immense reservoir of phosphate. In the year 2000 it was estimated that as much as 51 million tonnes of phytate can be found every year in commercially produced fruits and crop seeds (Lott *et al.*, 2000; Gerke, 2015), which according to the USGS is equivalent to as much as 37% of the phosphate rock mined that year. Therefore, the importance of $InsP_6$ in the global P cycle cannot be understated.

Thus, we have an organic phosphate molecule that is found throughout nature. It can be found in the cells of eukaryotes as part of an important family of signalling molecules, it is highly abundant in growing plants, seeds, and grains as the major storage form of phosphate, and it can form a significant part of the organic soil phosphates. So, what is the problem?

$InsP_6$ is a very highly charged molecule with a negative charge that can range from 0-12 depending on the pH. At a neutral pH this charge can be as high as 6-8 (Turner *et al*., 2002). This high charge causes it to form very stable complexes with mineral ions, as well clays and calcite (Crea *et al*., 2008; Mamedov *et al*., 2016) which renders it inaccessible to be absorbed as nutrients. Additionally, with a stronger adsorption capacity then phosphate, phytate binds more tightly to soil particles. Therefore, even further reducing the accessibility of phosphate in the soil. This binding does not just affect $InsP_6$ in the soil, but it can also negatively affect the abundance of mineral ions in animals. Monogastric organisms such as humans, swine and poultry lack the sufficient activity of $InsP_6$-degrading enzymes in their digestive tract to properly break down the $InsP_6$ (Gupta *et al*., 2015). It has been reported that $InsP_6$ inhibits the absorption of calcium, iron, magnesium, manganese, and zinc ions, reducing their bioavailability (Nissar *et al.*, 2017), although the extent to which absorption is hindered is debated. $InsP_6$ can also form complexes with amino acids and proteins, having an adverse effect on protein utilisation and digestibility (Kies *et al*., 2006; Selle *et al*., 2000). This has significant effects in the swine and poultry industry.

In 2014 there were 21 billion chickens and 798 million pigs worldwide. These are all fed a diet that is largely cereal- and/or grain-based meaning that the dominant form of phosphate in these feeds are $InsP_6$ at approximately 10 g kg^{-1} feed (Selle *et al*., 2009). The animal then does not absorb sufficient phosphate to help it grow or develop, in addition to the anti-nutritional properties described above. These insoluble nutrients and proteins are unable to be absorbed and are subsequently excreted as waste (Phillippy, 2006). Additional phosphorus can be supplemented in the form of phosphate rock, but this can further exacerbate the overuse of phosphorus and environmental issues. Similar to phosphate, phytate can be adsorbed onto the solid soil phase where it can accumulate in the soil. Adsorbed phytate can be subjected to weathering, or enzymatic degradation which further increases the amount of phosphate run-off and leaching.

Therefore, from an environmental and economic outlook, there must be a way in which $InsP_6$ can be degraded, releasing the inorganic phosphate and other mineral ions for absorption by the animal, as well as diminishing the amount of exogenous phosphate that needs to be

supplemented into the animal's diet. One method that has been used since the early 1990's is the supplementation of animals feeds with the biological enzyme, phytase.

1.3. Overview of Phytases

Phytases are enzymes that catalyse the sequential dephosphorylation of $InsP_6$ into the lower inositol phosphates with the release of inorganic phosphate at each hydrolysis step. The Enzyme Nomenclature Committee currently recognises three types of phytases based on the position at which the first phosphate is cleaved from the *myo*-inositol ring. These are the D-3-phytase (EC 3.1.3.8), which can also be referred to as a L-1-phytase, D-4-Phytase (EC 3.1.3.26), which can also be referred to as a L-6-phytase, and 5-phytase (EC 3.1.3.72). There is also the addition of a multiple inositol polyposphatase phosphatase (EC 3.1.3.62) which can remove the first phosphate from multiple positions. These enzymes have been found in plants, bacteria, fungi and animals. Phytases have been detected in calves, birds and reptiles as well as rice and wheat plants, alongside many other species (Haefner *et al*., 2005). Phytases have been found to be extracellular and secreted into the surroundings or remain intracellularly (Azeke *et al*., 2011; Olstorpe *et al*., 2009; Ullah and Gibson, 1987).

1.3.1. The History of Phytases

The first mention of phytase in scientific literature was in the early 20th century (Suzuki and Yoshimura, 1907; Dox and Golden, 1911). Suzuki *et al*., examined the cleavage of "phytin" into inosite and phosphoric acid by suspending rice bran in water for several days. Upon returning, they discovered that phosphoric acid was formed at the expense of the organic phosphorus compound to such an extent, that an enzyme might be responsible for this reaction. This was confirmed by precipitating the enzyme in alcohol and testing in vitro with a phytin solution. Later on, Dox and Golden examined the production of phytase in fungi, particularly *Aspergillus niger*.

Over the next 50 years, most of the research into phytases involved the discovery of phytases in a variety of different species, such as plants, vertebrates such as frogs, snakes and certain species of birds and fish (Rapoport, Leva, and Guest, 1941), as well as mammals (Pileggi, 1959). But as noted by Peers in the early 1950's they had not been as intensively studied as many of the other phosphatases (Peers, 1953). It wasn't until the 1960s however, that interest in isolating, characterising and commercialising phytate degrading organisms and their enzymes started to grow (Hill and Richardson, 2007). With the International Minerals and Chemicals Corporation (IMC), now known as IMC Global, in 1962 making efforts to

develop a commercial enzyme, foreseeing a potential market in a phytase that would break down $InsP_6$ in plant meals. Despite being discontinued in 1968, it did provide a valuable isolate for the future from *Aspergillus* (Lei *et al*., 2013). It was the last couple of decades where phytase research truly grew exponentially. Phytases were identified with differing catalytic mechanisms for $InsP_6$ degradation (Mullaney and Ullah, 2007; Vincent *et al*., 1992).

Phytases are commonly split into four broad classes: the most common and widely studied is the Histidine Acid Phytases (HAPhys), the other three are the β-propeller phytase (BPPhys, also known as an alkaline phytase), the Purple Acid Phytases (PAPhys; metalloenzymes) and the Protein Tyrosine Phosphatase-like phytase (PTPhys).

1.3.2. Histidine Acid Phytases

The histidine acid phytases (HAPhy) are the most widely studied of the four classes of phytases, with most bacterial, fungal and plant phytases belonging to this class (Kumar *et al*., 2012). This class of enzymes belongs in the superfamily of histidine acid phosphatases, a diverse group of functional proteins who shares a conserved catalytic core centred on a histidine residue which is phosphorylated during the reaction (Rigden, 2008). This superfamily is comprised of two branches. Branch 1 consists of proteins which function in metabolic regulation and developmental processes whilst the smaller branch 2 consists of phosphatases and phytases (Coker *et al*., 2013). Phytases from this branch have a low sequence homology however each member possesses a conserved α/β and a more variable α-domain (Rigden, 2008), Figure 1.5.

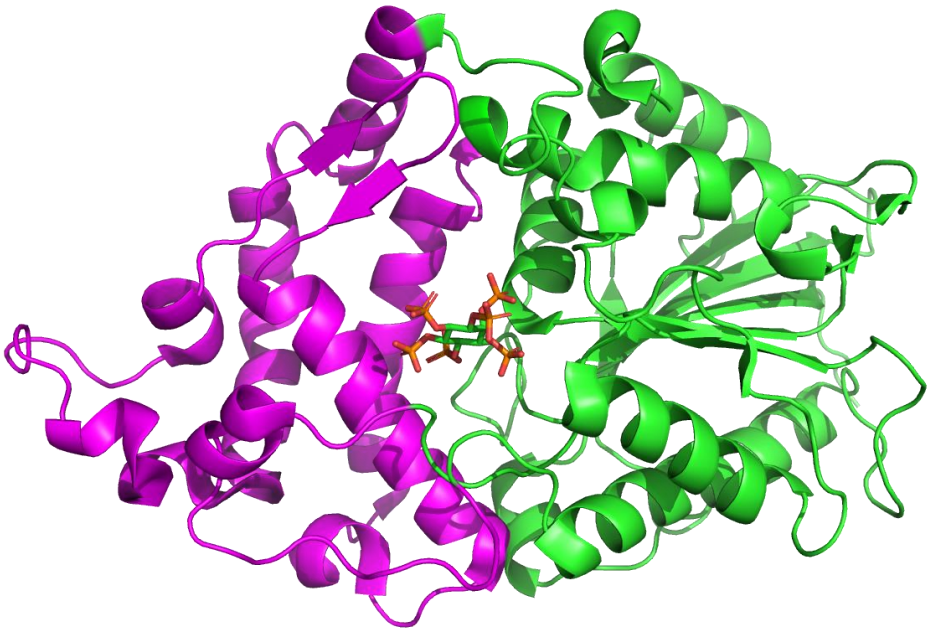


Figure 1.5 – A cartoon representation of the *Escherichia coli* phytase, PDB 1DKP (Lim *et al*., 2000), that has been rendered in PyMOL. All histidine phytases possess a variable α-domain, pink, and a conserved α/β domain, green. The interface between the two provides the catalytic site, in which, in this structure, the $InsS_6$ analogue of $InsP_6$ is bound.

In-between these two domains are the two important catalytic motifs, that are found in all HAPhys. The first is the active site heptapeptide sequence motif, RHGxRxP, with the RHG (aRginine, Histidine and Glycine) being very highly conserved while the xRxP is more variable. The second is the proton donor motif, HD (Histidine and Aspartic Acid, Figure 1.6.

```
Gilliamella.apicola       QYQLDKVVILSRHGIRTPL …------------LLVGHDSTIAALLGAL
Pseudomonas.syringae      GLRLDKVVLVMRHGIRPAT …---------KWLAYVGHDSNIAQLRTLL
Lonsdalea.quercina        PFTLENVVTLSRHGVRPQT …RFQA-----PLVMFVGHDTNIAQIQTML
Pantoea.deleyi            TYQLEKVVELSRHGVRPPT …---------RWFILVAHDTNIAMVRTLM
Zymomonas.mobilis         RYMLEKVVELSRHGVRPPT …QAIAGPPDVAYLLYVAHDTNIAFIRRLL
Acidobacteriaceae         GADLQMVVMLSRHGVRSPI …IGNPGD---RLVLLVGHDTNIVAVAGAL
Sphingomonas.sp           PLQVDRVVLMMRHGIRPPT …---------KVALIAGHDTNVANLAGVL
Citrobacter.amalon.       DMKLERVVIVSRHGVRAPT …KLPV-----SLLFIAGHDTNLANLSGVF
Edwardsiella.tarda        GYRLDKVVIVSRHGVRAPT …AN-------RLLLLVGHDTNLANLSGLL
                             :: ** : ***:*.                 ..**:.:. :   :
```

Figure 1.6 – Multisequence alignment of nine diverse bacterial HAPhys using MUSCLE, Multiple Sequence Comparison by Log-Expectation, from EMBL-EBI. The heptapeptide sequence motif and proton donor motif are highlighted in red and blue respectively.

The acid-base catalysis of $InsP_6$ involves a two-step reaction in which phosphate is cleaved from the $InsP_6$. As $InsP_6$ is a highly negatively-charged molecule, the binding pocket must reflect this, the two arginine, R, residues are positively charged which helps the binding (Lei *et al*., 2007). A schematic of the reaction for the fungal 3-phytase from *Aspergillus niger phyA*

is shown in Figure 1.7. In this reaction, $InsP_6$ enters the active site of the phytase, helped by the attraction of the two positively charged arginine residues Arg-58 and Arg-62, as well as additional arginine residues that are not part of the active site. Here it is acted upon by the nucleophilic lone pair of electrons on the catalytic histidine residue, His-59, which attacks one of the phosphate groups. This leads to the formation of a phosphohistidine intermediate. Aspartic acid, Asp-339, donates a hydrogen ion to form an alcohol group which replaces the lost phosphate, forming the molecule $InsP_5$. In the next step of the reaction, the negatively charged Asp-339 residue attacks a water molecule which in turn attacks the phosphohistidine intermediate, generating the free, inorganic phosphate and resetting the catalytic system (Mullaney and Ullah, 2007; Oakley, 2010).

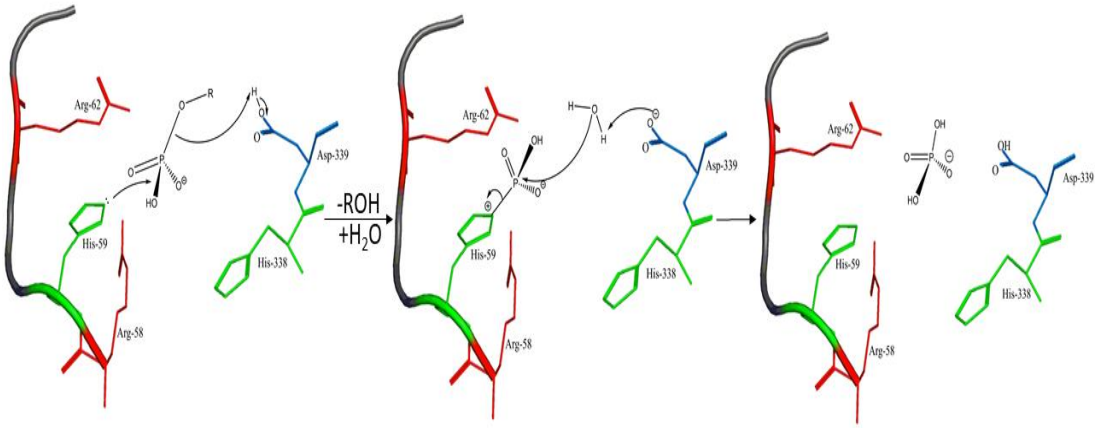


Figure 1.7 – The catalytic mechanism of the Histidine Acid Phytase from *Aspergillus niger*, PDB 3K4Q (Oakley, 2010)**.** Residues are numbered as in *Aspergillus niger*. 1. Nucleophilic attack and phosphorylation of the catalytic histidine residues, and release of $InsP_5$. 2. Protein donation by the HD motif, and breakdown of the phosphohistidine intermediate. 3. Release of inorganic phosphate. This mechanism was produced using PyMOL and Chemdraw.

The $InsP_5$ can re-enter the active site to be degraded further into the lower inositol phosphates. In many cases, the full conversion of inositol hexakisphosphate to *myo*-inositol however, does not occur. In fact, only two known phytase that are found in fungi, *Debaryomyces castellii* and *Schwanniomyces occidentalis* HP, have shown to be capable of degrading the axial 2-inositol monophosphate (Segueilha *et al.*, 1992; Boze *et al*., 2011).

Nearly all of the HAPhy's show strong stereospecificity towards the position of attack on the $InsP_6$ molecule (Balaban *et al*., 2018). For example, the HAPhy from *Aspergillus niger* shows predominant dephosphorylation from the D3-position, whereas in *Escherichia coli*, degradation initially begins at the D6-position (Greiner *et al*., 2009, 2000). A sub-branch of the histidine acid phytases however, shows less selectivity and these are the Multiple Inositol

Polyphosphate Phosphatases (MINPP). These phytases were first described in 1995 and characterised in 1997 (Craxton, Ali and Shears, 1995; Craxton *et al.*, 1997). The first was isolated from the liver of a rat. It was immediately noted that this phytase catalyses hydrolysis of $InsP_6$ without specificity towards any particular phosphate group, the 2-position being the exception. The MINPPs were found throughout mammals such as humans and birds before being isolated in plants (Chi *et al*., 1999; Cho *et al*., 2006; Dionisio *et al*., 2007). The profile of intermediates generated from $InsP_6$ by different MINPPs are variable (Acquistapace *et al*., 2020).

The MINPPs were previously considered to only function inside animal cells and plants as a means of regulating $InsP_6$ and $InsP_5$ levels (Chi *et al*., 2000; Romano *et al*., 1998), however, it wasn't until 2014 where the first bacterial MINPP was reported from *Bacteroides thetaiotaomicron* (Stentz *et al*., 2014).

Despite these enzymes sharing a low sequence homology, and functional diversity to the other members of the histidine acid phytases, they still possess the conserved α/β and more variable α-domain. They also contain the heptapeptide sequence motif, RHGxRxP, but they differ in their proton donor motif. Instead of the usual bipeptide motif, HD, these enzymes contain a tripeptide motif such as HAE (Stentz *et al*., 2014).

1.3.3. β-propeller Phytases

The β-propeller phytases (BPPhy) are widely distributed in nature playing a key role in the cycling of $InsP_6$ in both soil and aquatic microbial communities, BPPhy being the only class of phytase that is commonly found in aquatic systems and the most common in soil (Huang *et al*., 2009b; Lim *et al*., 2007a; Neal *et al*., 2017). These phytases, unlike the others, are structurally unique, exhibiting no homology to any known phosphatase and therefore showing no dephosphorylation activity for any phosphate-containing compound other then $InsP_6$ (Shin *et al*., 2001). They are aptly named, containing a ring of six β-sheets that is reminiscent of a six-bladed propeller. The activity of BPPhys are metal ion-dependent as they require calcium for activity and stability (S. Fu *et al*., 2008) which can be examined in the crystal structures, Figure 1.8. Most of the BPPhy described come from the genus *Bacillus*.

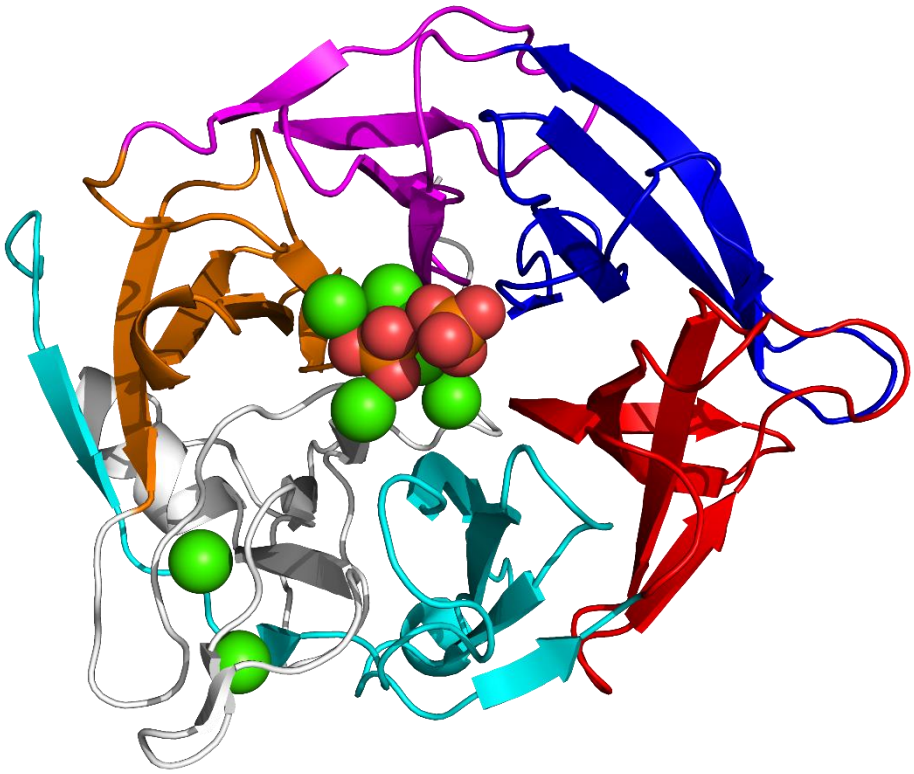


Figure 1.8 – A cartoon representation of the β-propeller phytase from *Bacillus amyloliquefaciens*, PDB 1H6L (Shin *et al*., 2001) that has been rendered in PyMOL. The structure forms a "six-bladed propeller" made up of β-sheets, highlighted in red, blue, pink, orange, grey and cyan. The crystal structure of BPPhy contains multiple calcium binding sites, green, and two phosphate binding groups, red/orange.

These enzymes have shown to have a high thermostability with an optimal pH of around 7-8, therefore they are commonly known as an alkaline phytase. This is different from the HAPhys which typically show optimal activity at a more acidic pH (Tang *et al*., 2006). Analysis of crystal structures and the amino acid sequences revealed a negatively-charged active site with multiple calcium binding sites allowing the favourable binding of Ca^{2+}-phytate (Kumar *et al*., 2017). Oh *et al* (2001) reports that in the absence of calcium ions, phytase activity ceases, indicating the important role that calcium plays (Oh *et al*., 2001). All of the structures of the BPPhy have shown six calcium binding sites, a high-affinity triad important for enzyme stability and a low-affinity triad that forms the catalytic centre. In Figure 1.8 however, a seventh calcium was located.

These positively charged calcium ions facilitate the binding of the substrate by creating a favourable environment for the negatively charged phosphate molecules on $InsP_6$. The positively charged substrate, Ca-$InsP_6$, favours the negatively charged electrostatic environment of the active site (Ha *et al*., 2000). This is present in the form of a "cage" of six aspartate and glutamate residues (Kumar *et al*., 2017) .

These calcium binding sites lay on "top" of the propeller, Figure 1.8. The mechanism of $InsP_6$ dephosphorylation in *Bacillus amyloliquefaciens* is believed to occur through the formation of a bidentate ligand between two oxyanions of two adjacent phosphate groups of the phytate and calcium (P-Ca^{2+}-P). These phosphates bind in two pockets, the cleavage and affinity site. This initiates dephosphorylation beginning at the 3-position with the 4-position in the affinity site. The calcium ions in the active site form a bridging activated water molecule that cleaves the phosphate bond, the role of the general acid ($B:H^+$) can be served by adjacent lysine residues. This is followed by degradation of the 1:2 position and then 5:4. This continues until inositol triphosphate where degradation is considerably slowed/stopped as there is no longer an adjacent phosphate (Kim *et al.*, 2010; Oh *et al.*, 2006; Shin *et al.*, 2001). A proposed schematic is shown in Figure 1.9.

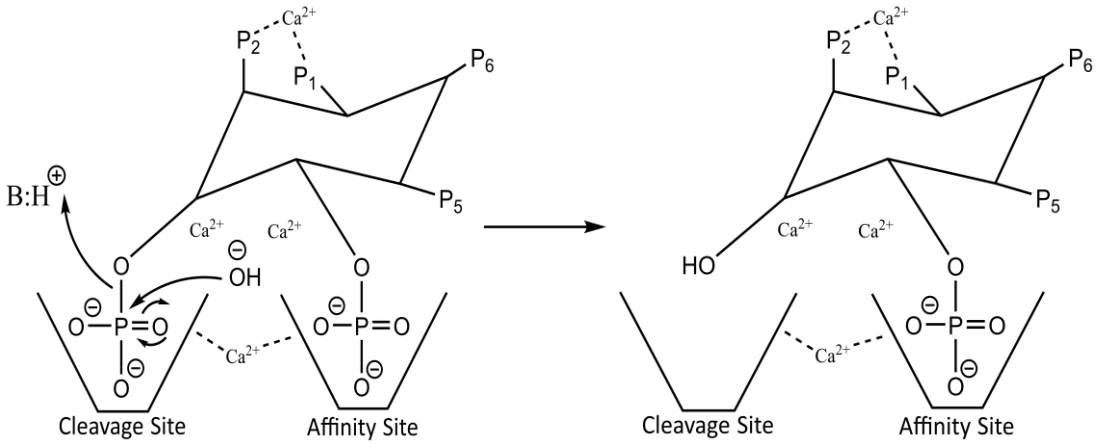


Figure 1.9 – A proposed mechanism for the degradation of $InsP_6$ by β-propeller phytases. 1. Binding of two adjacent phosphates of $InsP_6$ to the cleavage and affinity pockets within the protein. 2. Nucleophilic attack by a bridging water molecule to cleave the phosphate bond. 3. Phosphate release.

Sequence alignments of the BPPhys have identified two common motifs that are ubiquitous amongst the BPPhys, these are the DAADPAIW and NNVD motifs, Figure 1.10 (Kumar *et al.*, 2017).

```
Thioploca.ingrica              VVATVETKPVSVDGDAADDSAIWVHPQDPA.....HMNNVDLRYHFPL
Marine.bacterium.AO1-C         VLADMETEPVNSADDAADDPAVWVHPKKVG.....RINNVDVRYGFLL
Paenibacillus.jamilae          VLPSVETEAVEDGEDAADDPAIWLNPVDPG.....KMNNVDLRYGFTL
Bacillus.amyloliquefaciens     VNAAAETEPVDTAGDAADDPAIWLDPKTPQ.....KLNNVDIRYDFPL
Bacillus.licheniformis         VTADAETEPVDTPDDAADDPAIWVHPKQPE.....KLNNVDLRYNFPL
Caulobacter.mirabilis          VTATVETQPVEGGGDAADDPAIWVHPTDPS.....RMNNVDLRDGFKL
Sphingopyxis.flava             ARPTVETVPVETGGDAADDPAIWINPADPA.....RMNNVDLRSGFRL
                               . .  ** .*.   *****.*:*: *      .:****:*  * *
```

Figure 1.10 - The Multisequence alignment of seven diverse BPPhys using MUSCLE, Multiple Sequence Comparison by Log-Expectation, from EMBL-EBI.

1.3.4. Protein-Tyrosine Phytases

Ruminants, unlike monogastric organisms, have a particularly metabolically active gut microbiota that is capable of degrading dietary phytate. Consequently, ruminants are able to access the phosphate content of phytate for nutrition. (Yanke *et al*., 1999). Among the cohort of enzymes possessed by the ruminant gut microflora are a group of phytases, the protein-tyrosine phytases, otherwise known as the cysteine phytases, that were first identified in the anaerobic ruminant bacterial species *Selenomonas ruminantium*. This enzyme has no homology to the other classes of phytase, instead it bears the characteristic protein tyrosine phosphatase (PTP) active site sequence motif, HCxxGxGRT. The PTP superfamily is commonly associated with controlling the levels of cellular protein tyrosine phosphorylation and is known to play a significant role in initiating, sustaining and terminating cellular signalling (Andersen *et al*., 2001). They can be split into three classes, (i) the classical PTPs, (ii) the dual-specificity PTPs and (iii) low-molecular-weight phosphatases. These enzymes act through the formation of a cysteinyl-phosphate intermediate. Their structure, Figure 1.11, contains two domains, a large domain, consisting of a 4-stranded β-sheet which is sandwiched by α-sheets on both sides (shown in green), the smaller domain consists of a 5-stranded β-sheet (shown in magenta). At the interface of these two domains resides the phytate-binding pocket. Two loop structures lie at the bottom of the pocket. The P-loop which contains active-site sequence motif HCxxGxGRT, red, and the WPD-loop, orange, which contains the donor Asp-212, blue, and Cys-241, yellow, serves as the nucleophile and general acid respectively to catalyse the hydrolysis of the phosphate bond of phytate, green, Figure 1.11.

Figure 1.11 – The cartoon representation of the protein tyrosine phosphatase-like phytase (PTPhy) from *Selenomonas ruminantium,* PDB 1U26 (Chu *et al*., 2004), rendered using PyMOL. The structure is made up of two domains, a 4-stranded β-sheet which is sandwiched by α-sheets on both sides, green, the smaller domain consists of a 5-stranded β-sheet, magenta. The P-loop contains the active-site sequence motif HCxxGxGRT, red, and the WPD-loop, orange, which contains the donor Asp-212, blue, and Cys-241, yellow, serves as the nucleophile and general acid respectively to catalyse the hydrolysis of the phosphate bond of phytate, green.

The mechanism of action is believed to occur firstly through the binding of phytate to the active site pocket which is facilitated by interactions between the basic charges in the pocket and the negatively charged phytate molecule. This is then orientated through hydrogen bonds, such as those from Arg-247 which help stabilise the molecule. The thiolate anion of Cys-241 then attacks the phytate, initially at the 5-position, creating a phosphocysteine intermediate. Asp-212 acts as a general acid by donating a hydrogen ion and breaking the phosphate bond. For the second step, Asp-212 acts as a general base by taking a proton from a bound water molecule which in turn attacks the phospho-cysteine intermediate, removing the phosphate and resetting the catalytic system (Chu *et al*., 2004). The schematic diagram is detailed in Figure 1.12.

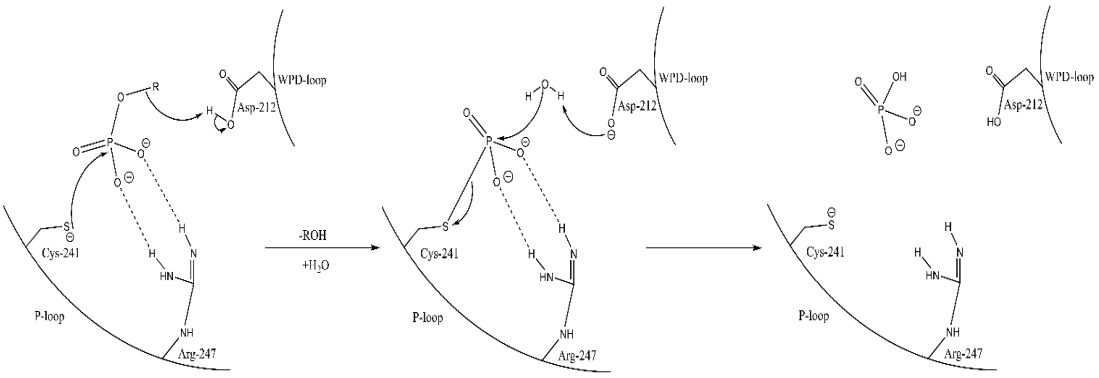


Figure 1.12 – A proposed mechanism for the dephosphorylation of phytate by a protein tyrosine phosphatase like phytase (PTPLP) by *Selenomonas ruminantium*. 1. Nucleophilic attack by a catalytic cysteine residue and release of $InsP_5$. 2. Proton donation by a catalytic aspartic acid residue and release of the phosphocysteine intermediate. 3. Release of phosphate.

1.3.5. Purple Acid Phytases

The purple acid phytases (PAPhy) are part of the metallophosphoesterase superfamily (MPE). This superfamily catalyses phosphate ester hydrolysis with the aid of a binuclear metallic centre. In animals this centre is typically two irons, whilst in plants the second iron can be replaced by either a zinc or manganese ion (Lei *et al*., 2007). The first reported discovery of a PAP that was capable of degrading $InsP_6$ was in 2001 and isolated from germinating soyabeans (Hegeman and Grabau, 2001). Their names are given to them because they exhibit a purple colour in their inactive oxidised form, and a pink colour in their active reduced form. This colour change is due to the oxidation of the heterovalent binuclear iron core which occurs through a charge transfer transition with a highly conserved tyrosine residue (Matange *et al*., 2015; Mitić *et al*., 2006).

The purple acid phytases are commonly referred to as plant phytase, as they had only been found in plants such as wheat, barley, maize and rice (Dionisio *et al*., 2011). In 2018 however, a potential PAPhy was isolated from an earthworm cast bacterium (Nasrabadi *et al.*, 2018). Furthermore, metagenomic studies identified phytase activity from purple acid phosphatases from nonvegetal origins (Villamizar, Nacke, Boehning, *et al.*, 2019).

Members of the MPE superfamily have a diverse array of functions such as nucleases, phosphoprotein phosphatases, pyrophosphatases and phospholipases, however, they still share the core MPE fold and conserved active sites and motifs (Matange *et al*., 2015). In general, the PAPs from both plant and animals appear to occur in at least two forms which can be distinguished by their molecular weights, a smaller 35 kDa and larger 55-60 kDa form. The smaller form is typically a monomer, while the large forms a homodimer, however, this

can range up to a hexameric state (Matange *et al*., 2015; Schenk *et al*., 2013). The fold is described as a βαβαβ architecture, the heart of which harbours a β-sandwich which is surrounded by α-helices. The active site of the MPE's is located at the top of the β-sandwich and consists of two metal ions. These ions are coordinated by seven conserved amino acids which are contained in five highly conserved sequence motifs. These are G**D**xx**Y**, red, G**N**HE, magenta, G**H**x**H**, cyan, G**D**xG, blue, and Vxx**H**, yellow, Figure 1.13.

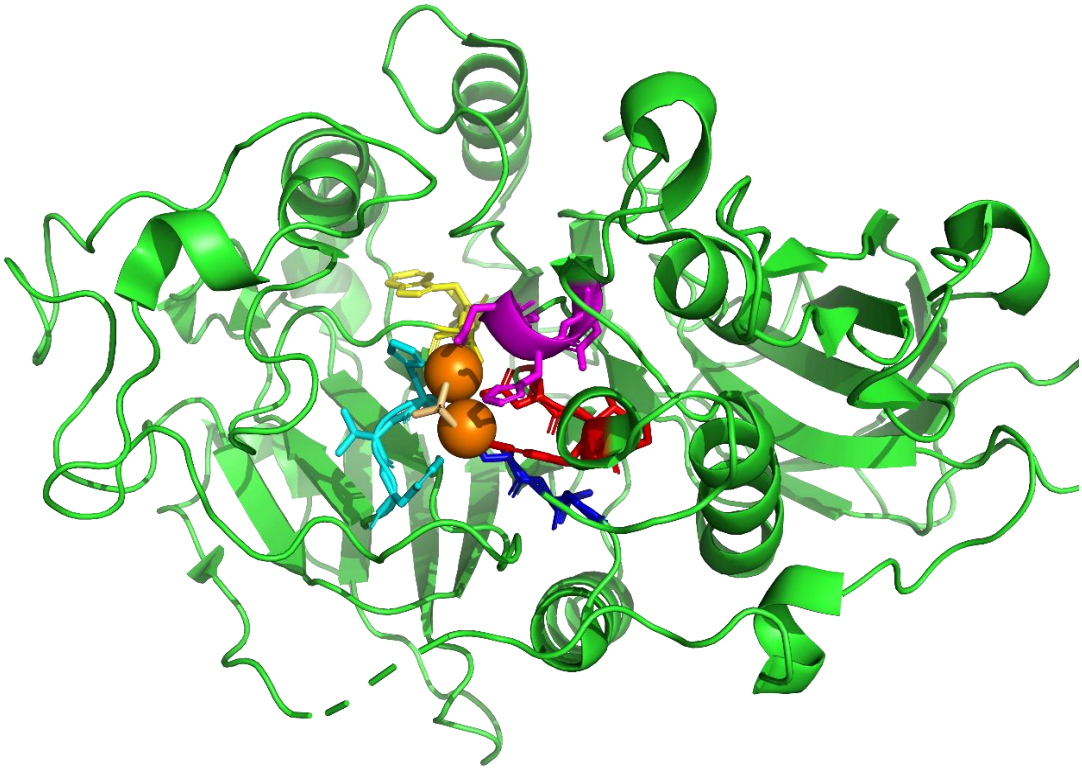


Figure 1.13 – The cartoon representation of the Purple Acid Phytase from *Triticum aestivum* Isoform B2, PDB 6GJA (Unpublished), rendered using PyMOL. Each coloured segments represent one of the five highly conserved sequence motifs, GDxxy (red), GNHE (magenta), GHxH (cyan), GDxG (blue) and VxxH (yellow).

In the mechanism of action, residues displayed are from *Triticum aestivum*, the $InsP_6$ binds to the divalent metal ions facilitated by the exchange of the bound water ligand. This step is stabilised by the coordinated amino acid residues. The coordination of the substrate is followed by nucleophilic attack of a metal-bound hydroxide which in turn causes the formation of an -OH group on the $InsP_5$ via protonation by an active site amino acid and release of the $InsP_5$ molecule leaving behind the bound phosphate. In the next steps, the regeneration of the enzyme and reset of the catalytic system, are the least understood of the mechanism. It is believed that the bound phosphate is displaced and at least two water

molecules are used in this process. The proposed mechanism is shown in Figure 1.14 (Mitić *et al*., 2006; Schenk *et al*., 2013, 2008).

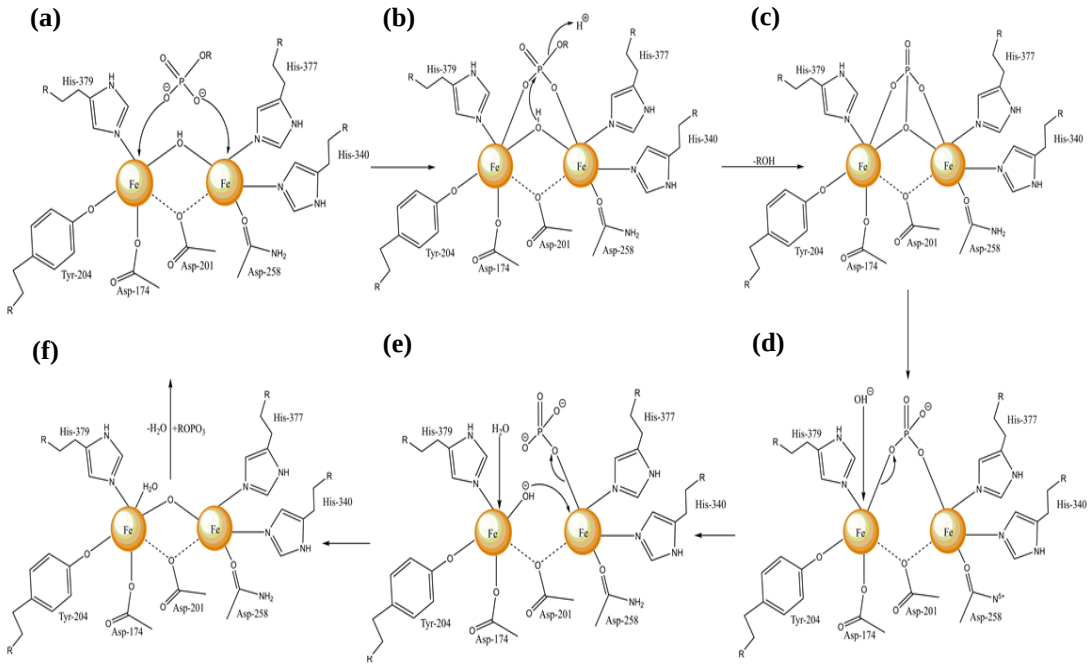


Figure 1.14 – The proposed mechanism for the dephosphorylation of $InsP_6$ by the Purple Acid Phytase from *Triticum aestivum*. 1. $InsP_6$ binds to the catalytic diiron centre. 2. Nucleophilic attack by bridging water molecule and release of $InsP_5$. 3. Further nucleophilic attack by two water molecules, reforming the bridging water molecule and releasing phosphate.

1.4. Phytases in the Animal Feed Industry

Phytases are found throughout nature and have been shown to have catalytically diverse ways of degrading $InsP_6$ in a wide range of environments, from Himalayan, volcanic soils to the guts of fish and animals and even the oceans of the Antarctic (Huang *et al*., 2009b; Jorquera *et al*., 2008; Kumar *et al*., 2013; Nakashima *et al*., 2007; P. Yu *et al*., 2015).

Phytases are a taxonomically widespread group of enzymes that adopt diverse fold, catalytic mechanism and display great variation in specificity of attack on $InsP_6$, relative activity on alternative substrates, pH optima, protein stability and catalytic activity (turnover). Earlier in this thesis, there was a discussion for the need for $InsP_6$ to be degraded in the guts of monogastric organisms such as swine and poultry to improve the nutrition of the animal, as well as contribute to lessening the need for the over-supplementation of fertilisers and phosphate rock. Since 1991, phytases have been engineered and supplemented into animal feeds. Since then, phytases have become a major sector of a global enzyme market of

estimated value $5 billion in 2015, with annual growth estimated at 6-8% from 2016 to 2020 (Guerrand, 2018).

1.4.1. The History of Phytases in Industry

Despite the isolation efforts of the International Minerals and Chemicals (IMC) in 1962 failing to find a suitable phytase for commercial use, one of the isolates from the study, from the fungal species *Aspergillus niger* (ATCC 66876) provided the basis for the first commercial phytase launched in 1991 under the name Natuphos® (BASF), a HAPhy. Originally, the screening of novel phytases focused on the identification of microorganisms that exhibit high phytase activity. However, with the development of recombinant DNA technology, the gene from *Aspergillus niger* was cloned, characterised and overexpressed (Mullaney *et al*., 1991; van Hartingsveldt *et al*., 1993). The overexpression of the phytase was significantly important as this allowed the necessary production for commercial use. This subsequently led to a shift in the search for novel phytases towards enzymatic properties. Phytases should exhibit a high catalytic efficiency, a wide pH profile that is capable of degrading phytase along the animals gastrointestinal track, pH 2.5-7 (Mabelebele *et al*., 2013; Merchant *et al*., 2011a), a high thermostability, that is necessary to remain active during the feed pelleting process (Slominski *et al*., 2007) and resistance to proteolytic digestion (Dersjant-Li *et al*., 2015). The first generation of phytases were based on fungal HAPhys, with many developed from the phytase encoded by the gene originating from *A. niger*. It wasn't until the early 2000's when a bacterial candidate was discovered. The *Escherichia coli* genes *appA* and *appA2* were identified to code for more effective phytases, showing superior feed efficacy (Lei *et al*., 2013; Rodriguez *et al*., 1999a, 1999b).

The phytase market developed slowly however, since the introduction of Natuphos® in 1991, because there were no cost benefits to using phytase instead of supplementing with mined phosphate rock. But this changed in 2007-2008. A significant increase in the price of inorganic phosphate, in addition to new environmental legislation made the use of phytase much more cost effective and led to a new drive in phytase researching resulting in the second generation of phytases based on the *E. coli appA* phytase (Lei *et al*., 2013), and in 2007 the European Commission authorised the use of Danisco's Phyzyme® XP for use as a feed additive (Gabriele Aquilina, *et al.*, 2012).

There are now many different phytases on the market that are derived from a number of sources. Currently, all of the phytases shown in Table 1.1 are from the class of histidine acid phytases.

Table 1.1 –Commercial Phytases in use today

Trademark	Supplier	Organism	Expression Host
Natuphos	BASF	*Aspergillus niger*	*Aspergillus niger*
Quantum Blue	AB Vista	*Escherichia coli*	*Pichia pastoris*
Phyzyme XP	Danisco	*Escherichia coli*	*Saccharomyces pombe*
Axtra PHY	Danisco	*Buttiaxuella* sp.	*Trichoderma reesei*
Ronozyme HiPhos	DSM/Novozymes	*Citrobacter braakii*	*Aspergillus oryzae*
Ronozyme NP	DSM/Novozymes	*Peniophora lycii*	*Aspergillus oryzae*
OptiPhos	Huvepharma	*Escherichia coli*	*Pichia pastoris*

These phytases are typically expressed in eukaryotic host systems rather than their original host, with these systems providing numerous benefits that the original does not. Eukaryotic systems can provide post-translational modifications such as glycosylation, which have been shown to improve thermostability (Yao *et al*., 2013) as well as generating higher protein yields (Ma *et al*., 2020).

1.4.2. The Economic and Environmental Benefit of Phytases

The advent and development of phytases for use in industries is regarded as one of the top ten landmark discoveries in swine nutrition in the past century (Cromwell, 2009) alongside their environmental benefits. The benefits of addition of phytase into the diets of monogastric animals is multifaceted. As phytase breaks down the phytate that has been consumed in the form of animal feed, the released inorganic phosphate is available for nutrition of the animal. This obviates the addition of inorganic phosphate to the diet (in place of phytate), it ameliorates excretion of phosphate, as inositol phosphates, in animal manures and obviates the antinutrient effects of phyate. These antinutrient effects extend to interference with protein digestibility and mineral provision. In the reports by Selle and Ravindran (Selle and Ravindran, 2007; Selle and Ravindran, 2008) they reviewed experiments where phytases were supplementation in poultry and swine diets. The results showed that a large amount of the phytate phosphate was made available to the animal, leading to an increased weight gain. In addition to this, supplementation also improves the bioavailability

of previously bound minerals such as calcium, zinc and iron (Stahl *et al*., 1999). In conjunction to the nutritional effect the supplementation of phytases has, there is also a significant environmental impact. With less unutilised dietary phosphorus entering the environment, there will also be a reduction in environmental pollution as there is less phosphorus run-off into lakes and rivers.

1.4.3. The Future of Phytases

Despite the new advances in the second generation of phytases, there are still some issues in their activity. Many of the HAPhys are strictly positional stereospecific, meaning that the D6- and D3-phytases that are used commercially will have a preference for these positions, with only small quantities of the other $InsP_5$'s forming. The $InsP_5$'s are rapidly degraded to $InsP_4$ and $InsP_3$ but these lower inositol phosphates begin to accumulate and the rate of hydrolysis significantly slows (Balaban *et al*., 2018).

Therefore, the search and development for the 'ideal' phytase continues. This can diverge into two directions: Identifying novel phytases in nature or engineering desired characteristics in known phytases, or a combination of both.

Strategies for identifying novel phytases in the environment have included screening on media where phytate is the sole source of phosphorus (Howson and Davis, 1983; Kerovuo *et al*., 1998; Kumar *et al*., 2013) and more recently through prospecting environmental metagenomes and the creation of genomic libraries (Tan *et al.*, 2014; Farias, Almeida and Meneses, 2018; Villamizar, Funkner, *et al.*, 2019; Villamizar, Nacke, Boehning, *et al.*, 2019).

Phytases have also been developed using protein engineering. There are three broad strategies, directed evolution, rational design and a combination of the two, as a semi-rational design. Directed evolution is an engineering method that mimics natural selection in order to alter the structural and functional properties of an enzyme. These may employ techniques such as random mutagenesis which introduces random point mutations into the gene of interest (Sen *et al*., 2007). Another means of inducing variability is DNA shuffling, which involves the recombination of homologous genes which are randomly fragmented by DNase 1 and amplified using PCR (Joern, 2003). These techniques are useful because they can be performed without prior knowledge about the structure and mechanistic pathway of the target enzyme. It does however require the use of a high-throughput assay to analyse the mutant proteins generated. Both of these techniques were performed in the study by Tang *et al*., (2018) generating mutants of *Aspergillus niger* N25 with higher thermostability

and half-life (Tang *et al*., 2018). Rational design is the opposite of directed evolution, requiring detailed knowledge of the sequence and structure of the enzyme. Fakhravar and Hesampour demonstrated this using the crystal structure of the *E. coli* (K12) *appA* phytase, bioinformatics and docking binding energy measurement. They produced a S392F (Serine to Phenylalanine) mutant with a 25.6% improvement in catalytic efficiency compared to the WT (Fakhravar and Hesampour, 2018). Semi-rational design combines the rational element of rational design with the randomness of direction evolution. In Lehmann *et al* (2020) a consensus phytase designed using a sequence alignment of 13 fungal phytases showed a 15-22°C increase in thermostability.

The expression system also plays an important role in the production of new phytases. For commercial use, the cost of production is a dominant factor in the adoption of the new enzyme. As shown in Table 1.1, there are numerous expression hosts that have been used to produce commercial phytases, typically eukaryotic systems, each host provides the necessary machinery to properly fold and post-translationally modify such as disulfide bond formation and glycosylation (Leis *et al*., 2013). Suleimanova *et al.,* (2021) examined the effect of expressing a histidine acid phytase from *Pantoea* sp. 3.5.1 in three different expression systems, *Escherichia coli*, *Pichia pastoris*, and *Yarrowia lipolytica* (Suleimanova *et al*., 2021). Despite all producing the identical protein, there were differences in the biochemical characteristics, such as pH activity and stability, as well as thermostability. Therefore, it is important to develop an expression host that can produce large quantities of the protein at a lower cost, as well as providing the best conditions to enhance its stability and activity.

1.5. Project Aims

The aims of this PhD project were to explore the diversity of phytases in soils using both culture-dependent and culture-independent methods. The main objectives of this project can be outlined in five points: (1) method development for phytase isolation, (2) the investigation of phytase activity and isolation of phytases from different environments, (3) detailed characterisation of a novel soil MINPP, (4) environmental metagenomics to examine the diversity of phytases in diverse environments, (5) examining the controls of phytase expression.

Chapter 2. A Critical Analysis of the Phytase Isolation Methods in the Literature and Development of New Methods

The methods that are commonly used to isolate phytases, typically from the soil environment, involves incubating soil suspensions onto a "phytase-specific media" where the phytate has precipitated forming a cloudy surface (Baranes-Bachar *et al*., 2018; Howson and Davis, 1983). The believe is that the presence of phytase-producing bacteria will lead to the formation of clearing zones around the bacterial colony, as the phytate is degradation from the plate.

This chapter describes the many pitfalls associated with this method, alongside a literature review to investigate how many manuscripts, out of a sample of twenty-five, have considered these potential errors. This led to the further development of the "phytase-specific media" to better identify phytase degraders whilst avoiding the many pitfalls associated with this method. This was followed by the development of a new phytase isolation method using High-performance liquid chromatography (HPLC) to track phytate degradation in a soil slurry. This led to the isolation of two distinct phytases from environmental soil samples that was published in Microbial Biotechnology (Rix *et al*., 2020).

By taking the lessons learnt in the first part of this chapter, a long-term isolation study was performed using soil from Rothamsted to isolate and investigate a diverse range of phytate-degraders using HPLC to accurately determine whether the isolate is indeed a phytate degrader, as well as identify the degradation pathway.

2.1 Materials and Methods

2.1.1. HPLC analysis of inositol phosphates

Inositol phosphates were separated by anion exchange HPLC on a 3 mm × 250 mm CarboPac PA200 column (Dionex, Sunnyvale, CA) fitted with a 3 mm × 50 mm guard column of the same material, eluted with a gradient of methanesulfonic acid (Whitfield *et al*., 2018). The injection volume was 20 µL. Peak areas were integrated with ChromNav (Jasco) software and compared to that of standards of $Na_{12}InsP_6$ (Merck Millipore – Calbiochem Cat: 407125-25 MG Lot: 2663470). The Jasco LC-2000 HPLC system comprised: an AS-2055i autosampler, PU-2085i quaternary gradient pump, a PU-2085 pump for delivery of ferric nitrate and a UV-2075 UV-visible detector. The column and reaction coil were held in a thermostatted oven at 35 °C. Chromatography data was exported as x,y data and redrawn in GraphPad Prims v8.0.

2.1.2. Media preparations

Phytase Specific Media (PSM) – Comprised: 15 g agar, 10 g glucose, 2 g $CaCl_2$, 5 g NH_4NO_3, 0.5 g KCl, 0.5 g $MgSO_4.7H_2O$, 0.01 g $FeSO_4.7H_2O$, 0.01 $MnSO_4.H_2O$ L^{-1}, autoclaved at 121 °C for 20 minutes. After, 4 mL of filter-sterilised (0.22 µM Pore PTFE syringe filter, Kinesis, UK) 100 g L^{-1} phytate (Test sample provided by our industrial partner, AB Vista), pH 7 to a final concentration of 4 g L^{-1}.

Phytase Isolation Minimal Media, L^{-1} – The base media modified from Neal *et al* (2017), comprised: 18.7 mM NH_4Cl, 8.6 mM NaCl, 1 mM $MgSO_4$, 0.1 mM $CaCl_2$. This was supplemented with a filter-sterilised mixed carbon source: 50 µL per 10 mL of medium, pH 7, 200 mM succinic acid (Sigma), 200 mM glucose (Formedium), 200 mM sucrose (Sigma), 200 mM pyruvic acid (Sigma), 200 mM glycerol. Ten µL of vitamin solution (containing 10 mg pyridoxine-HCl, 5 mg thiamine-HCl, 5 mg riboflavin, 5 mg para-amino benzoic acid, 5 mg nicotinic acid, 2 mg vitamin B12, 2 mg folic acid, L^{-1}) and with micronutrients: 10 µL (2 g nitriloacetic acid, 1 g $MnSO_4.6H_2O$, 0.8 g $Fe(NH_4)_2(SO_4)_2$, 0.2 g $CoCl_2.6H_2O$, 0.2 g $ZnSO_4.7H_2O$, 20 mg $CuCl_2.2H_2O$, 20 mg $NiCl_2.6H_2O$, $NaMoO_4.2H_2O$ L^{-1}). For environmental samples, the medium included 0.1-0.2 mg mL^{-1} cycloheximide to inhibit fungal growth (Neal *et al*., 2017). This was supplemented with 1 mM phytate, pH 7 (AB Vista).

Lysogeny Broth and Agar – Comprised: 10 g tryptone (Formedium), 5 g yeast extract (Formedium), 10 g NaCl (Sigma) and 15 g agar (Sigma), L^{-1}.

All Culture Medium – All in one mixture comprised of: 20 g proteose peptone, 3 g beef extraction, 3 g yeast extract, 3 g malt extract, 5 g dextrose, 0.2 g ascorbic acid and 10 g agar, L^{-1}.

BG11 (Blue-Green Medium), pH 7.1 – Stocks of each individual solution were prepared (Per 500 mL): 75 g $NaNO_3$, 2 g K_2HPO_4, 3.75 g $MgSO_4.7H_2O$, 1.8 g $CaCl_2.2H_2O$, 0.3 g Citric acid, 0.3 g Ammonium ferric citrate green, 0.05 g $EDTANa_2$, 1 g Na_2CO_3, 10 mL of each stock was combined and made up to 1 L deionised water, 1 mL of trace element solution (containing 2.86 g H_3BO_3, 1.81 g $MnCl_2.4H_2O$, 0.22 g $ZnSO_4.7H_2O$, 0.39 g $Na_2MoO_4.2H_2O$, 0.08 g $CuSO_4.5H_2O$, 0.05 g $Co(NO_3)_2.6H_2O$ L^{-1}) was separately added. This was supplemented with 1 mM phytate, pH 7 (AB Vista).

To 250 mL of media, the following components were added – 250 µL 100 mg/mL cycloheximide (prepared in 100% ethanol) and 1.5 mL mixed carbon source.

2.1.3. Acid extraction of inositol phosphates from phytase-specific media plates

The control strains were streaked onto PSM and grown for three days at 30 °C. After sufficient clearing zones had been produced, bacterial cells were washed off the plate using dH_2O, and 100 mg samples of both cleared and cloudy agar zones were extracted with 400 µl 0.8 M HCl with vortexing after disruption of the agar with a plastic stirrer. Samples were extracted for 15 min at room temperature and centrifuged at 16060 g for one minute. The supernatant was removed with a HPLC needle and syringe and filtered through a 13 mm diameter 0.45 µm pore PTFE syringe filter (Kinesis, UK) into a borosilicate glass HPLC vial (Chromacol 03-FISV(A)).

2.1.4. Preparations of Soil Cultures

Agricultural Soil (0.5 g), Fakenham, Norfolk, UK, was added to 10 mL of minimal media, pH 7, in a 30 mL universal bottle. This was supplemented with a mixed carbon source, vitamin solution, micronutrients solution and cycloheximide to inhibit fungal growth, as described in Section 2.1.1. The soil suspensions were incubated under shaking at 180 RPM and 30 °C for six days, unless stated otherwise, taking samples for HPLC each day.

2.1.5. Preparation of Soil Cultures for HPLC Analysis

Five hundred µL of a well-mixed soil culture in minimal media was centrifuged at 16060 g for 5 minutes. The supernatant was filtered through a 13 mm diameter 0.45 µM pore PTFE

syringe filter (Kinesis, UK) and centrifuged again, and an aliquot (200 μL) was dispensed into an HPLC vial.

2.1.6. The Microbial Communities' response to phytate degradation with and without a mixed carbon source

To analyse the effect of carbon sources, and the ability of the soil microbial communities to utilise phytase as the sole carbon source, 10 mL of minimal media was supplemented with 2 mM phytate, with and without the addition of the mixed carbon source. One hundred μL of the phytase-active soil culture was used to inoculate the fresh minimal media, these soil suspensions were incubated under shaking at 180 RPM and 30 °C for six days, taking samples for HPLC each day.

2.1.7. Culturing Isolates

One hundred μL of the active soil culture, Agricultural, Fakenham, and Church Farm, a John Innes field experimental station, was serially diluted, 10^{-4}-10^{-6} onto LB and All culture agar media. These were incubated at 30 °C overnight. Single colonies were re-streaked onto fresh plates and reinoculated into liquid media to confirm phytase degradation. Reinoculated isolates were incubated under shaking at 180 RPM and 30 °C, taking samples for HPLC each day, until phytate degradation was complete.

2.1.8. Phytase activity from Church Farm Isolates without trace metals

Isolates were grown in 10 mL LB broth overnight under shaking at 180 RPM and 30 °C. One mL of cell growth was spun down at 16060 g and washed three times with 0.9% saline. These were then resuspended in three different solutions, 0.2 M Na-acetate pH 3.5, 0.2 M Na-acetate pH 5.5- and 20-mM Tris-HCl pH 7, each containing 1 mM $InsP_6$ and 0.1% NaCl. These were incubated under shaking at 180 RPM and 30 °C. Samples were taken on Day 0 and Day 2 and examined for phytate degradation on the HPLC.

2.1.9. 16S rRNA amplification

Single bacterial colonies were purified, and their 16S rRNA gene was amplified using the primers (IDT) 28F (5'-GAGTTTGATCNTGGCTCAG-3') and 519R (5'-GWNTTACNGCGGCKGCTG-3') from genomic DNA using colony PCR. GoTaq® G2 flexi DNA polymerase (Promega) using the manufacturer's protocol, 2.5 mM $MgCl_2$ and 0.2 μM primers and 5 μL DNA template, was used to generate a single band that was resolved on a 1 % agarose gel. This was purified using

a QIAquick Gel Extraction Kit (QIAGEN). The PCR conditions were: initial denaturation, 95 °C, 3:00 (min:sec), denaturation, 95 °C, 0:30, annealing, 55 °C, 0:45, elongation 72 °C, 0:45, repeated 30 times, and final elongation 72 °C, 4:00.

Sequencing of these PCR products at Eurofins (MWG, Ebersberg, Germany) identified the two isolates further examined in this chapter as strains of *Acinetobacter* sp, and *Buttiauxella* sp. The amplified 16S rRNA gene of strain AC1-2 was identical to that of *Acinetobacter calcoaceticus* strain IIPRDSCP-11, *Acinetobacter* sp. strain YAZ49 and *Acinetobacter calcoaceticus* strain EB11. The 16S rRNA gene of strain CH 10^{-6}-4 Ac was identical to that of *Buttiauxella agrestis* strain EB112, *Buttiauxella* sp. CL_136_AN_40 and *Buttiauxella* sp. SA_136_AN_45. These were deposited under the accession number MT450216 and MT450213-MT450215 respectively.

2.1.10. Gel Electrophoresis

One percent agarose was prepared by dissolving 0.5 g agarose (Melford) in 50 mL TAE buffer. At point of use, this was heated to melt the agarose, allowed to cool, and supplemented with 1.5 µL ethidium bromide. Once set, 5 µL of DNA sample and 1 µL of the 6X loading dye were loaded onto the gel, and the gel run at 100 Volts for 45 minutes.

2.1.11. Degenerate primer design

To confirm that the isolated *Buttiauxella* sp. CH-10^{-6}-4 contained a histidine acid phytase, degenerate primers were designed to the *appA* gene of sequenced *Buttiauxella spp*. genomes obtained from NCBI blast using the *appA* gene from *Citrobacter amalonaticus* (Accession number: ABI98040.1) as the query sequence. (*Buttiauxella* sp. JUb87, *Buttiauxella* sp. A111*, Buttiauxella agrestis, Buttiauxella ferragutiae, Buttiauxella brennerae, Buttiauxella gaviniae, Buttiauxella noackiae, Buttiauxella* sp, BIGb0552, *Buttiauxella* sp. *3AFRM03*. Accession numbers – WP_133522247.1, WP_183270673.1, WP_115627260.1, WP_083963259.1, WP_083967908.1, WP_083962918.1, WP_034459427.1, WP_134184667.1 and WP_121814872.1 respectively*).* Two primer sets were used: Set 1 – Forward 5'-GTGGTTTRACTATTCAYCACC-3', Reverse 5'-GGCATGGATTGCVCTAATC-3' and Set 2 - Forward 5'-GCGAGAARTTTCAACARCAGG-3', Reverse 5'-GTGYCCGGCAAKAAACAGG-3' to amplify a 259 and 725-bp product from the *Buttiauxella* sp. isolate respectively. The PCR conditions were: initial denaturation, 95 °C, 0:30 (min:sec), denaturation, 95 °C, 0:30, annealing, 55 °C, 0:30, elongation 72 °C, 1:00, repeated 35 times, and final elongation 72 °C 5:00. The 725-bp PCR products were sequenced by Eurofins, and their identity to ratified

Buttiauxella spp. appA genes was established by BLAST analysis, the sequence was deposited in GenBank under the accession MT680195.

2.1.12. Genome Sequencing of strains

The *Acinetobacter* sp. strain AC1-2 genome was sequenced by MicrobesNG (University of Birmingham, UK) using illumina technology. This Whole Genome Shotgun Project has been deposited at DDBJ/ENA/Gen-Bank under the accession JABFFO010000000. Genomic completeness was analysed using BUSCO v3 (Simão *et al*., 2015), an open-source software that provides quantitative measures for genomic completeness based on evolutionary informed expectations of gene content from near-universal single-copy orthologs selected. The *Acinetobacter* sp. strain AC1-2 completeness was measured at 98 and 98.9% from both BUSCO's bacterial and *Gammaproteobacteria* databases, respectively.

2.1.13. The growth of *Acinetobacter* sp. on different Carbon Sources

Acinetobacter sp. strain AC1-2 were grown on minimal media plates and liquid cultures containing $InsP_6$ as the sole phosphate source and supplemented with either 1 mM glucose, sucrose, succinate, citrate, *myo*-inositol, glycerol, D-mannitol or pyruvate. The plates were incubated at 30 °C, and the liquid cultures grown under shaking at 180 RPM and 30 °C.

2.1.14. Nomenclature of the Inositol Phosphates

The term 'Ins' with the prefix 1D-*chiro*-, *myo*-, *neo*-, or s*cyllo*- is commonly used as an abbreviation of particular inositols and their phosphate substituted derivatives. Therefore, *myo*-Inositol 1,2,3,4,5,6-hexakisphosphate can be abbreviated to *myo*-Ins(1,2,3,4,5,6)P_6, or, as this molecule is unique, *myo*-$InsP_6$. Additionally, *myo*-$InsP_6$ is the only stereoisomer of inositol hexakisphosphate used in this project and therefore the term $InsP_n$ is used to signify a *myo*-inositol phosphate with n phosphates. According to the relaxation of rules for the numbering of carbons in *myo*-inositol (IUBMB, Nomenclature recommendations, 1989), I have used the '1D-'numbering convention, with or without the '1D-' prefix, but for $InsP_5$s, e.g., Ins(2,3,4,5,6)P_5, I also use the shorthand form $InsP_5$(1-OH). Where chromatography does not allow the resolution of enantiomers e.g., Ins(2,3,4,5,6)P_5 and Ins(1,2,4,5,6)P_5, I use the term $InsP_5$(1/3-OH). For co-eluting enantiomers of $InsP_4$ e.g., Ins(1,4,5,6)P_4 and Ins(3,4,5,6)P_4, I use the term D/L-Ins(1456)P_4.For $InsP_4$s where there is only one stereoisomer, I use the term Ins(2,4,5,6)P_4.

2.1.15. Measuring the phytate degradation potential of Broadbalk, bare fallow and arable soils from Rothamsted

Soil samples from Rothamsted Research were collected in 50 mL Falcon tubes and immediately brought back for experimentation. Samples were taken from the Highfield Ley Arable, Highfield Bare Fallow and Broadbalk experimental plots (September 2019).

Each of the three soils were inoculated into phytase isolation minimal media (0.5 g in 10 mL) supplemented with 1 mM phytate (AB Vista) and analysed each day for phytate degradation by HPLC. The Broadbalk soil removed all of the phytate from the liquid media such that neither phytate nor lower inositol phosphates were recovered at day 0. As such, the culture was supplemented with an additional 1 mM phytate.

2.1.16. Long-term phytase isolation experiment

Each of the soil samples were inoculated into Phytase Isolation Minimal Media (0.5 g per 10 mL) at 30 °C, shaken for three days before 100 µL was serially diluted 10^{-4} – 10^{-7} onto Minimal media, PSM, 1.10 and 1.100 dilutions of LB agar plates. These plates were left at ambient temperature on the laboratory bench.

After a month of growth, 80 isolates were streaked onto PSM plates, an additional 49 isolates were streaked onto MM, LB 1/10 and 1.100 and blue green plates. Unfortunately, many of the isolates were lost due to fungal growth, however, 66 of the isolates were grown and tested for phytase activity using HPLC.

2.1.17. Measuring the effect of soil and calcium on the adsorption of phytate

To 0.5 g of Highfield ley arable soil in 10 mL phytase isolation minimal media, additions of 1 mM $InsP_6$ and 1 mM $InsP_6$ + 1 mM $CaCl_2$ were added separately. These cultures were shaken for 30 minutes, before samples were analysed by HPLC.

Measurements for phytate absorption was calculated by peak integrations. For example, a 20 µL injection of a 1 mM solution, 20 nmol, gave a peak area of 2,876,000. Therefore, a peak area of 9,483,421, obtained with the Highfield ley Arable sample, is equivalent to (9,483,421/2,876,486)*20 nmole = 66.4 nmol, and a solution concentration of 3.3 mM.

2.1.18. Calculating the absorption capacity of Broadbalk, Highfield bare fallow and Highfield ley arable soils through the addition of 10 mM phytate

Soil (0.5g) was added to an empty Poly-Prep® Chromatography Columns, after which 1 mL of 10 mM $InsP_6$ was added to the column and allowed to settle for 10 minutes to allow the phytate to adsorb to the soil before being allowed to drain. The elutate from the column was centrifuged 21460 g for 5 mins and filtered using a 13 mm diameter 0.45 µM pore PTFE syringe filter (Kinesis, UK) to remove microparticulates. The sample was then analysed by HPLC, Section 2.11. The amount of phytate that had been adsorbed was estimated by integrating the $InsP_6$ peak and comparison with a control sample of a known concentration, taking account of dilutions, to see how much was left in the liquid phase. Additionally, one month-old samples of each soil type were subjected to the same treatment to examine how weathering of the soil sample may affect its adsorption capacity.

Following sampling for HPLC, the soil suspension was extracted with 0.25 M NaOH, 0.05 M EDTA (Doolette *et al*., 2010; Giles *et al*., 2011), in a 5:1 solution:soil ratio at 30 °C by overnight shaking. Samples of supernatant of the centrifuged 9965 g, 5 mins were aliquoted into an Eppendorf and spun further, 21460 g, 5 mins to remove all soil particulates and the soil suspension analysed by HPLC to measure $InsP_6$.

2.1.19. Examining phytate degradation of liquid, aerobic and anaerobic layers of a salt marsh

Alongside examining soils environments that are predominantly inland, sampling was conducted in the Blakeney salt marsh on the north Norfolk coast (52.977°N, 0.9778°E), that is part of a larger salt marsh complex that stretches for >200 km along the North Sea Coast (van de Velde *et al*., 2020). Push core liners made of polyvinyl chloride were used to extract vertical columns of sediments from two pools at low tide, 24/11/2020, these pools are regularly covered with the high tide, 2-2.1 m on the day. One pool was ferruginous, with ferrous iron concentrations up to 3 mM and the other was sulfidic, with sulfide concentrations up to 8 mM (Wilkening *et al*., 2019). A layer of pond water was left at the top of the core, which was sealed at top and bottom to prevent ingress of air.

A sample of 0.5 g of either sediment or water layer was inoculated into 10 mL phytase-isolation minimal media supplemented with 1 mM phytate (Sigma P8810) under shaking at

30 °C with samples taken each day for HPLC analysis. After the time-course was run, the iron-core samples were incubated with 0.25 M NaOH and 0.05 EDTA as described in Section 6.1.6.

2.2. Results and Discussion

2.2.1. Analysing the Pitfalls with regards to the Isolation of Phytases from the Environment

Since the isolation of the first phytase in the early 20th century (Suzuki, and Yoshimura., 1907), phytases have been isolated and characterised from a variety of different environments, animals and microorganisms. Each new addition adding to the growing list of differing characteristics, be it activity, substrate degradation pattern, thermostability or kinetics. The small fraction of environmental organisms amenable to culture however has the consequence that the biodiversity of phytase producers is grossly underestimated. Consequently, metagenomic and metaproteomic approaches have supplanted culture-based approaches for the study of the relationship of microbiological diversity and soil phosphorus (Chen *et al*., 2019; Neal *et al*., 2017; Yao *et al*., 2018). Alternatively, others have employed amplicon sequences of functional phosphatases using *phoD* alkaline phosphatase specific primers (Ragot *et al*., 2015). When allied with heterologous expression, metagenomic methods have revealed novel catalytic diversity among phytate degraders extending classification beyond the four canonical classes (Villamizar *et al.*, 2019a; Villamizar *et al.*, 2019b) as have more conventional functional genomic methods (Sarikhani *et al*., 2019). Irrespective of the methods of identification of candidate phytases, whether as commercial product leads or contributors to environmental processes, both culture-dependent and culture-independent approaches rely on informative enzyme assays for characterisation of the reactions catalysed (Rix *et al*., 2020).

The opportunity for mischaracterisation of isolates or 'purified' enzymes was noted as early as the 1970's when Cosgrove (Cosgrove *et al*., 1970) identified isolates active on impurities in the phytate substrate. Indeed, Cosgrove, Tate and co-workers, following the method of Ballou (Hendrickson and Ballou., 1964), pioneered methods for stereochemical and enantiomeric identification of inositol phosphates in soils (Cosgrove and Irving., 1980) and as products of enzyme activity on phytate (Irving and Cosgrove, 1971). The work of Ballou, Cosgrove, Tate, and co-workers truly set the standards of inositol phosphate analysis that are rarely matched today.

In the report by Cosgrove *et al*., (Cosgrove *et al*., 1970), they identified that numerous papers report on the presence of low levels of phytase activity in microorganisms, however, these should be looked at with caution due to the possibility that such activity is due to the action of phosphatases on the lower inositol phosphates that are often present in commercially available phytate. This is still the case in the present day, with many of the commercial phytates that had sufficient purity being discontinued (Madsen *et al*., 2019), with many labs facing difficulties in finding a suitable replacement. Currently, 25 mg of substrate advertised as ≥99% pure from Selleckchem costs approximately £80. Analysis by HPLC reveals that approximately 25% of the inositol phosphate content of this material is a mixture of inositol pentakisphosphates, Figure 2.1A. Lower purity phytate from Sigma-Aldrich (P8810 – sodium phytate from rice) is as much as £55-102 for 10 grams, Figure 2.1B. For multiple phytase isolation experiments in a variety of conditions, or rigorous characterisation of phytases, it much more economically feasible to buy lower purity phytate. This should, however, be noted in the literature which is often lacking.

Importantly, the purity of phytate can be detected through techniques such as high-performance liquid chromatography (HPLC), Figure 2.1 shows the HPLC trace after injection of 20 µL of a 1 mM phytic acid solution from Selleckchem (top) and that from Sigma-Aldrich (bottom).

The largest peak on the HPLC chromatogram running at 36 minutes is the $InsP_6$ peak. However, there are four other peaks representing the four different forms of $InsP_5$ that can be resolved by the HPLC, these are $InsP_5$(2-OH), $InsP_5$(1/3-OH), $InsP_5$(4/6-OH), $InsP_5$(5-OH). In addition to this, small traces of $InsP_4$ peaks are also detectable. The HPLC chromatograms tell us that the two products are less than >90% pure, and the advertised 99% pure sample from Selleckchem is false. This indicates that both samples are susceptible to the action of phosphatases on the lower inositol phosphates, that would be undetectable without the use of HPLC.

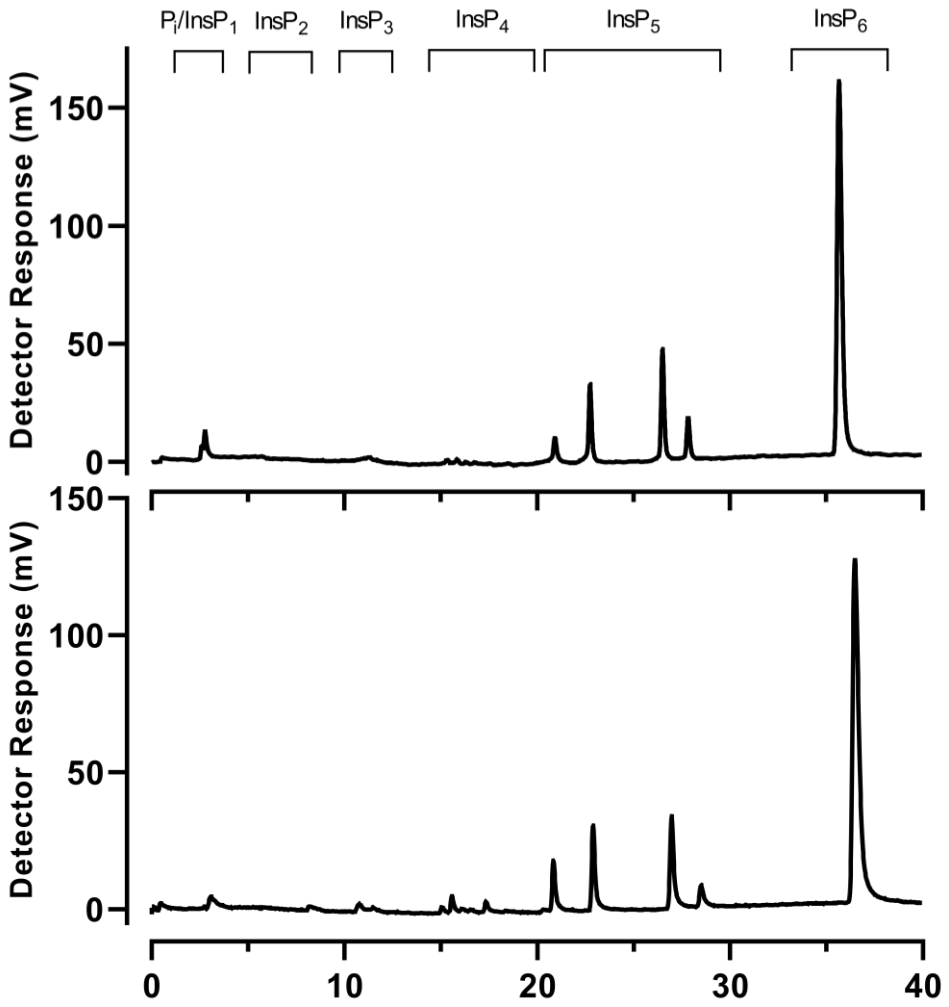


Figure 2.1 – HPLC trace of 1 mM Phytic acid sodium salt hydrate from Selleckhem (top) and phytate sodium salt hydrate from Sigma-Aldrich (bottom). Significant impurities of the lower inositol phosphates as well as inorganic phosphate are present in this solution. Regions of the chromatogram in which different inositol phosphates bearing different numbers of phosphate substituents are indicated.

In the same report, Cosgrove *et al*., (Cosgrove *et al*., 1970) isolated three soil bacteria that were able to utilise *myo*-inositol phosphates. However, only one, a *Pseudomonas* sp. was capable of degrading phytate. The others had no phytase activity but were able to degrade $InsP_5$ and the lower inositol phosphates. Therefore, the potential for false positives degrading the lower inositol phosphates is high, especially regarding the high levels of impurities shown in Figure 2.1 from commercial phytate.

Another issue that is commonly perpetuated was from the development of a solid medium agar plate called Phytase Specific Media (PSM) (Howson and Davis, 1983; Kerovuo *et al*.,

1998) which have been cited 269 and 520 times in the literature (Google Scholar). In this technique, the phytate precipitates upon addition to the agar media, causing the formation of an opaque surface when set. Bacterial and environmental isolates are then streaked onto the plate and over time clearing zones are formed around some bacterial colonies. This has been taken as a sign of phytate degradation as the precipitated phytate has disappeared. There are several key issues with this however, the first is that in the original text by Howson and Davis, after the addition of phytate to the medium, it was sterilised through autoclaving at 121 °C for 20 minutes. This has been shown to significantly degrade the phytate into the lower inositol phosphates and cause the release of inorganic phosphate (Fredrikson *et al*., 2002). Using relatively pure $InsP_6$ has shown that as much as 60% of the phytate is degraded after autoclaving, and there is an increase in the levels of the lower inositol phosphates and inorganic phosphate, Figure 2.2.

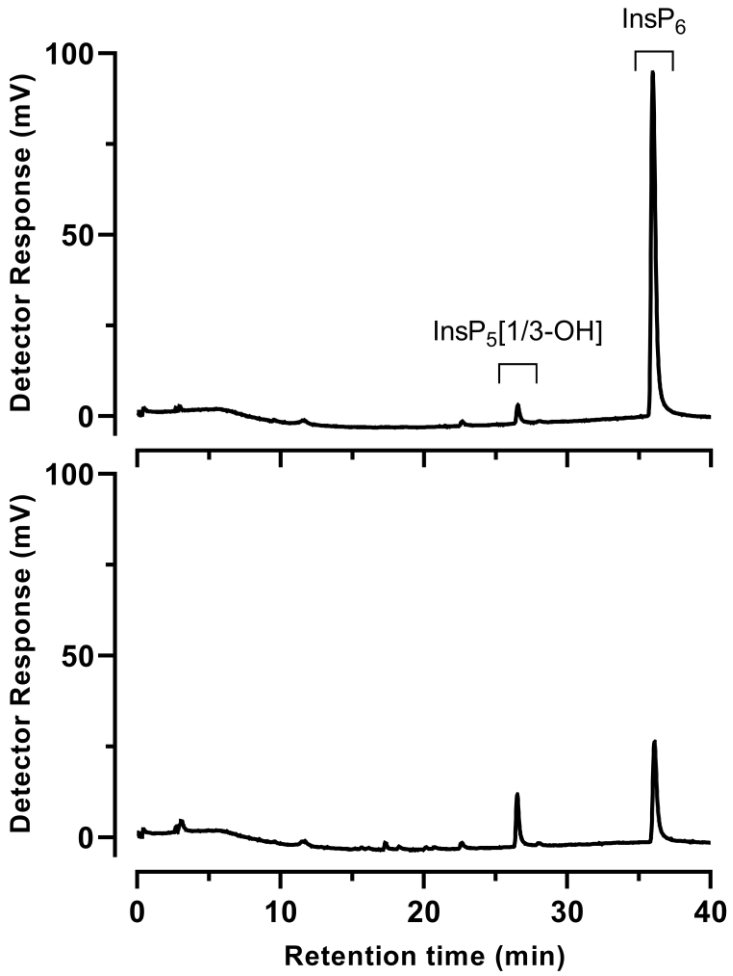


Figure 2.2 – The degradation of 1 mM $InsP_6$ after autoclaving at 121 °C for 20 minutes. Aliquots of 20 nM $InsP_6$ (equivalent to 20 µL of a 1 mM solution) were injected onto the HPLC column and analysed before and after autoclaving.

With significant degradation of the $InsP_6$ in the agar, it is also possible that the microorganisms grown on the plate will be using the inorganic phosphate that have been released or the lower inositol phosphates to grow. Additionally this approach to isolation of phytase producers also suffers from 'false positives' arising from the bacterial secretion of low molecular weight organic acids capable of solubilising the precipitated phytate causing the formation of clearing zones (Iyer *et al*., 2017). This in itself highlights another issue with the approach – that it is not suitable for screening at low pH – a condition for which many commercial enzymes have been optimised – and hence a condition which might be used in screening. There are also associated pH changes to the media with the addition of phytate. Without buffering, a solution of phytate can have pH, 2-4, which will change the pH of the media significantly. Therefore, the phytate should be buffered and upon addition to the media it should all be re-buffered as precipitation of the phytate causes additional pH changes. While solubilisation may be overcome by a two-step counterstaining test to re-precipitate acid-solubilised phytate (Bae *et al*., 1999), re-precipitation does not indicate to what extent the available phytate has been degraded, since other 'higher' inositol phosphates can also be re-precipitated.

Overall, clearing zones may not exclusively indicate enzymatic hydrolysis of phytate in the plate, while pH limitations of the method will necessarily be selective of the organisms cultured.

2.2.2. Review of the State of the Literature with regards to the pitfalls

The "phytase-specific" media (PSM) has frequently been used throughout the literature despite the many issues that have been associated with it. The two main papers associated have been cited a combined 789 times (Howson and Davis, 1983; Kerovuo *et al*., 1998). Of this list, twenty-five manuscripts were selected without any prior knowledge on the paper, Table 2.1. Four questions were posed of the manuscripts, 1: Is the purity of $InsP_6$ described? As discussed in the reports by Cosgrove and Madsen, there are considerable issues with obtaining pure $InsP_6$ for experiments. This could lead to false positives on the plate as microorganisms capable of degrading the lower inositol phosphates can obtain enough phosphate to grow (Cosgrove *et al*., 1970; Madsen *et al*., 2019). 2. Is the $InsP_6$ filter-sterilised and added separately to the other media components? This is due to $InsP_6$ degrading after autoclaving releasing inorganic phosphate and lower inositol phosphates. This again can lead to false positives. 3. Are these isolates selected due to the production of clearing zones

surround the colony, do they use the counterstaining technique? Microorganisms can secrete low molecular weight acids that solubilise the phytate causing the formation of a clearing zone, another false positive. 4. Do they use a technique to quantify $InsP_6$ and the lower inositol phosphates? Techniques such as HPLC, Figure 2.1, can show the composition of the inositol phosphates in the medium and how it changes over time.

Table 2.1 – A literature review of 25 randomly selected manuscripts which use Phytase Specific Media (PSM) to screen for phytase-producing microorganisms. * Indicate whether the counterstaining technique was used.

	Is $InsP_6$ Purity Described?	*Is $InsP_6$ filter-sterilised?*	*Screen Isolates on Clearing Zones*	*Do they show $InsP_6$ degradation?*	*Reference*
1	N	N	Y	N	(Tseng *et al*., 2000)
2	N	N	Y	N	(Gulati *et al*., 2006)
3	N	N	Y	N	(Aziz *et al*., 2015)
4	N	N	Y*	N	(Ocampo Betancur *et al*., 2012)
5	N	N	Y*	Y	(Sajidan *et al*., 2004)
6	N	Y	Y	N	(Hill *et al*., 2007)
7	N	N	Y	N	(Xiong *et al*., 2004)
8	N	N	Y	Y	(Suleimanova *et al*., 2015)
9	N	N	Y	N	(Kumar *et al*., 2013)
10	N	N	Y	N	(Imelda, Joseph, 2007)
11	N	N	Y	N	(Quan *et al*., 2001)
12	N	N	Y	N	(El-Toukhy *et al*., 2016)
13	N	Y	Y	N	(Escobin-Mopera *et al*., 2012)
14	N	N	Y	N	(Wang, Xueying *et al*., 2004)
15	N	N	Y	N	(Kim, Young-Hoon *et al*., 2002)
16	N	N	Y	N	(Milko A. Jorquera *et al*., 2011)
17	N	N	Y	N	(Roy, Moushree *et al*., 2012)
18	N	N	Y	N	(Muslim *et al*., 2018)
19	N	N	Y	N	(Mehmood *et al*., 2019)
20	N	N	Y	N	(Hosseinkhani, Baharak, 2009)
21	N	N	Y	N	(Alias *et al*., 2018)
22	N	N	Y*	N	(Hariprasad and Niranjana, 2009)
23	N	Y	Y	N	(Yoon *et al*., 1996)
24	N	N	Y	N	(Huang *et al*., 2009a)
25	N	N	Y*	N	(Chanderman *et al*., 2016)

The information from Table 2.1 indicates that many of the issues described above have not been considered when isolating phytases from the environment, 0/25 of manuscripts referenced the potential of impurities in the $InsP_6$ and the potential effect this may have on

the results. Only 3/25 filter-sterilised the $InsP_6$ and added to the medium after it had been autoclaved to avoid $InsP_6$ degradation. In all of these reports the formation of clearing zones were used to select isolates for phytate degradation, however, only four of these mentioned the counterstaining technique used by (Bae *et al*., 1999) and only two of these identified the possibilities of false positives. Finally, only 2/25 manuscripts used advanced techniques such as HPLC to properly examine the quantities of $InsP_6$ and the lower inositol phosphates and the changes to them during degradation.

Overall, these considerations are typically not taken into account and therefore the chance for false positives being reported as low phytase activity isolates is considerably increased (Cosgrove *et al*., 1970).

2.2.3. Lack of standardisation on Phytase Reporting

Another potential issue in phytase reporting is the lack of standardisation between methods for measurement of phytase activity. Phytase activity is typically measured using the molybdenum blue method for phosphate release. However, this is often regarded as a 'black box' in the analytical chemistry literature with the underlying chemistry only addressed superficially (Nagul *et al*., 2015). In brief, the release of inorganic phosphate can be determined after a set time with the additional of an ammonium molybdate developing solution (containing other components such as acetone, sulfuric acid, ammonium vanadate, citric acid). This causes the formation of a deep blue colour depending on the amount of inorganic phosphate released which can then be analysed using a spectrophotometer and then compared with a phosphate calibration curve. However, this method is constantly being modified and reoptimized. In the report by Qvirist *et al.,* (Qvirist *et al*., 2015) they assessed phytase activity in five different colourimetric methods at 15- and 30-minute time points. The calculated activities varied substantially, at 15 minutes ranging from 275-586 mU/mL in comparison with $InsP_6$ analysis by HPLC at 152 mU/mL. What is more, these techniques only analyse total phosphate release, and therefore do not take into account degradation of the lower inositol phosphates, which we know can occur. Therefore, while these methods are useful in detecting phosphatase activity, they may cause some false positives with regards to phytase activity. Once more, phytase activity is best addressed by directly analysing $InsP_6$ concentrations through techniques such as HPLC.

Of the 25 manuscripts analysed above, the authors' methods for phytase activity determination are shown in Table 2.2.

Table 2.2 – A literature review of the different of methods of measuring phytase activity in 25 randomly selected manuscripts.

	Measured at (nm)	*Assay Components*	*Developing Solution*	*Phytase Assay Method*	*Reference*
1	415	Sodium phytate, sodium acetate (pH 5.5)	Ammonium molybdate, Ammonium vanadate	1	(Tseng *et al*., 2000)
2	820	Sodium phytate, sodium acetate (pH 5.5)	Ammonium molybdate, sulphuric acid, ascorbic acid	2	(Gulati *et al*., 2006)
3	660	Sodium phytate, sodium acetate (pH 5.5)	Taussky-Schorr reagent	3	(Aziz *et al*., 2015)
4	N/A	N/A	N/A		(Ocampo Betancur *et al*., 2012)
5	355	Sodium phytate, sodium acetate (pH 5.4)	Ammonium molybdate, sulfuric acid, acetone, citric acid	4	(Sajidan *et al*., 2004)
6	N/A	N/A	N/A		(Hill *et al*., 2007)
7	380	Sodium phytate, acetate buffer (pH 5.5)	Ammonium molybdate, sulfuric acid, acetone, citric acid	4	(Xiong *et al*., 2004)
8	355	Sodium phytate, sodium acetate (pH 4.5)	Ammonium molybdate, sulfuric acid, acetone, citric acid	4	(Suleimanova *et al*., 2015)
9	415	Sodium phytate. citrate buffer (pH 5.5)	Ammonium molybdate solution	5	(Kumar *et al*., 2013)
10	NA	N/A	N/A		(Imelda, Joseph, 2007)
11	660	Sodium phytate, acetate buffer (pH 5.5)	Taussky-Schorr reagent	3	(Quan *et al*., 2001)
12	355	Sodium phytate, sodium acetate (pH 5)	Ammonium molybdate, sulfuric acid, acetone, citric acid	4	(El-Toukhy *et al*., 2016)
13	380	Sodium phytate, acetate buffer (pH 5.5)	Ammonium molybdate, sulfuric acid, acetone, citric acid	4	(Escobin-Mopera *et al*., 2012)
14	415	Sodium phytate solution (pH 5.5)	Ammonium molybdate, ammonium vanadate, nitric acid	6	(Wang, Xueying *et al.,* 2004)
15	820	Sodium phytate, Tris-HCl buffer (pH 7.0)	Ammonium molybdate, sulphur acid ascorbic acid, water	7	(Kim, Young-Hoon *et al*., 2002)
16	355	Sodium phytate, sodim acetate (pH 4.5)	Ammonium molybdate, sulfuric acid, acetone, citric acid	4	(Milko A. Jorquera *et al*., 2011)
17	610	Sodium phytate, acetate buffer (pH 5.5)	Ammonium molybdate, sulphur acid ascorbic acid	2	(Roy, Moushree *et al*., 2012)
18	355	Sodium phytate, acetate buffer (pH 5.5)	Ammonium molybdate, sulfuric acid, acetone, citric acid	4	(Muslim *et al*., 2018)
19	415	Sodium phytate solution (pH 5.5)	Ammonium molybdate, ammonium vanadate, nitric acid	6	(Mehmood *et al*., 2019)
20	355	Sodium phytate, acetate buffer (pH 5.5)	Ammonium molybdate, sulfuric acid, acetone	8	(Hosseinkhani, Baharak, 2009)
21	820	Sodium phytate, sodium acetate (pH 5.5)	Ammonium molybdate, sulphuric acid, ascorbic acid	2	(Alias *et al*., 2018)
22	700	Sodium phytate, sodium acetate (pH 5)	Ammonium molybdate, sulphuric acid, ferrous sulphate solution	9	(Hariprasad and Niranjana, 2009)
23	410	Sodium phytate, sodium acetate (pH 5)	Ammonium molybdate, sulphuric acid, acetone	10	(Yoon *et al*., 1996)
24	700	Sodium phytate, Tris-HCl buffer (pH 7.0)	Ammonium molybdate, sulphuric acid, ferrous sulphate solution	9	(Huang *et al*., 2009a)
25	390-420	Sodium phytate, Tris-HCl buffer (pH 8.0)	Ammonium molybdate, sulfuric acid, acetone, citric acid	4	(Chanderman *et al*., 2016)

The information from Table 2.2 demonstrates that there is a significant diversity in the measurements of phytase activity. There were eight different spectrophotometric measurements, at 355, 380, 410, 415, 610, 660, 700 and 820 nm, as well as slight differences in assays components, typically sodium phytate and sodium acetate were used, although at higher pH Tris-HCl is often used. All of the developing solutions used ammonium molybdate,

which is complexed with a variety of different solutions such as acetone, ammonium vanadate, ferrous sulphate.

Overall while the molybdenum blue method is a simple and easy to perform method, it does not truly define phytase activity as impurities may lead to lower inositol phosphatase activity, which will affect the results. Therefore, it is very important to determine change in phytate concentration through advanced techniques such as HPLC to categorically confirm phytase activity.

2.2.4. Acid extraction of phytate from PSM plates

The PSM plate approach is one of the most commonly used methods for the isolations of phytase-positive microorganisms from environmental samples, but it is not without the substantial drawbacks discussed above. The formation of clearing zones is accepted to be a sign of phytate degradation on the plate, and if there is no phytase activity, then it is the production secretion of small molecular weight acids. However, none of these papers have measured the extent to which phytate is degraded, or if it is even degraded at all.

Therefore, I performed such an analysis by extracting the inositol phosphates from the PSM plates after being incubated with control phytate degraders. The PSM plates were prepared as stated by (Kerovuo *et al*., 1998), with the exception that the $InsP_6$ was filter-sterilised, adjusted to pH 7 and added after autoclaving. The purity of the $InsP_6$ used was analysed by HPLC and considered 'clean', with less than 5% contaminating lower inositol phosphate. This $InsP_6$ was provided by AB Vista and is not available commercially. The precipitation of $InsP_6$ upon addition to solution causes a change in the pH and therefore was supplemented with sterile NaOH to bring it back to a pH of 7 (Kerovuo *et al*., 1998). Three control strains were tested. A *Escherichia coli*-pDEST-17-*BtMinpp* harbouring a plasmid-borne MINPP from *Bacteroides thetaiotaomicron* (Stentz *et al*., 2014), *Bacillus subtilis* strain ESKAPE (predicted to contain a BPPhy) and *Pseudomonas putida* J450 (predicted to contain a BPPhy) were each streaked onto PSM plates and allowed to grow over three days at 30 °C. All isolates generated clearing zones around their biomass on these PSM plates. Cores of agar from 'cleared' and 'non-cleared, cloudy' zones were extracted with HCl, and the inositol phosphate profile thereof examined by HPLC, Figure 2.3. While there can be slight differences in the efficiency of extraction between the cleared and cloudy zones, comparison of individual peaks within the respective profiles makes evident the different extents and pathways of phytate degradation by the strains (Rix *et al*., 2020).

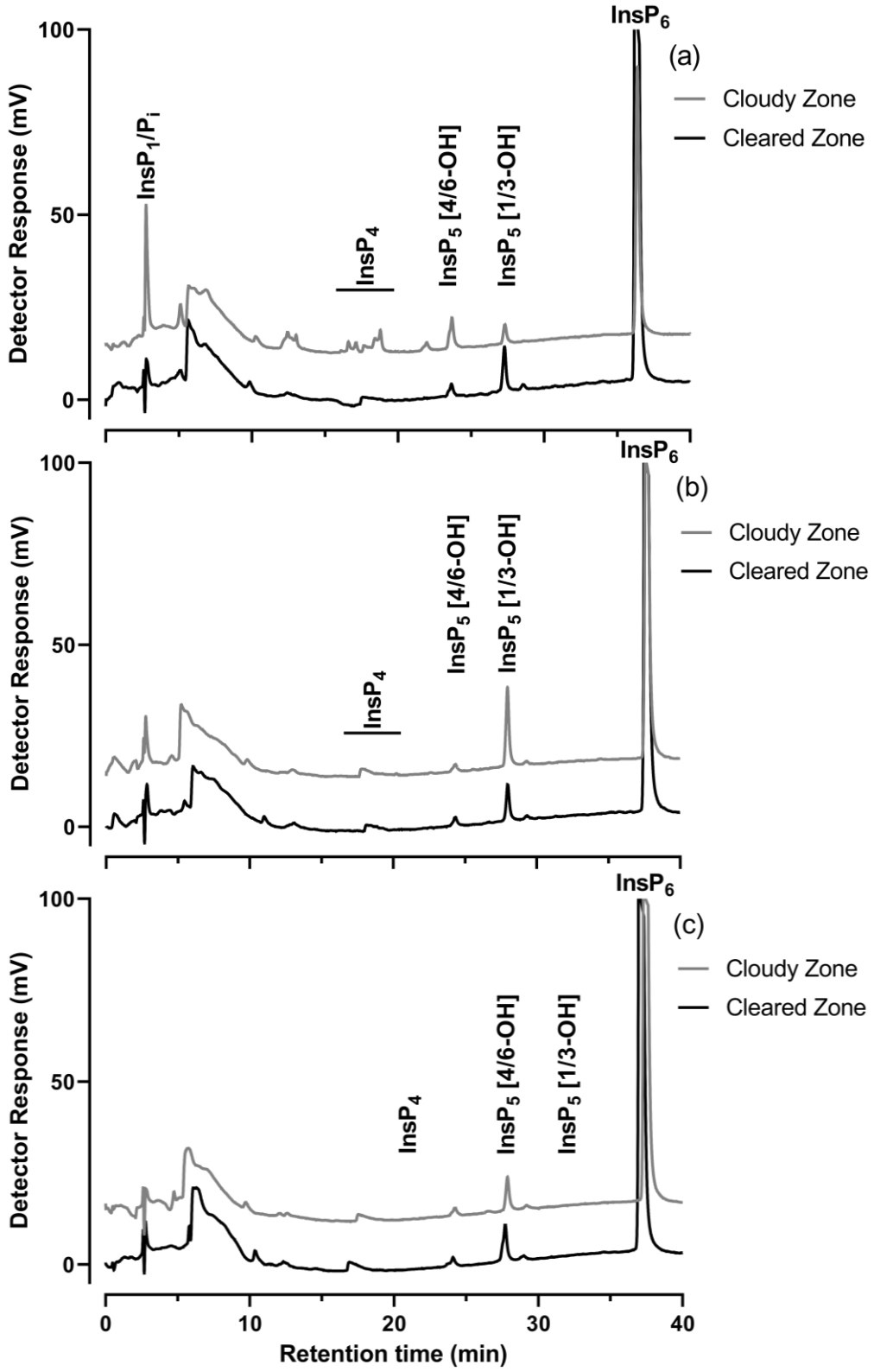


Figure 2.3 – Acid extraction of cleared and cloudy zones from the PSM plates to extract $InsP_6$. Samples were run on the HPLC. A, *Escherichia coli*-pDEST17-*BtMinpp, B, Bacillus subtilis* ESKAPE strain and C, *Pseudomonas putita* J450. Non-cleared agar, black lines; cleared agar, grey lines.

All profiles from the 'non-cleared' zones show the predominant peak of $InsP_6$ with a retention time of *c.* 37 min and a smaller peak of $InsP_5$ (1/3-OH) contamination with a retention time of *c.* 28 min, representing approximately 5% of total inositol phosphate in this 'clean' $InsP_6$ substrate. Inorganic phosphate (P_i) elutes with the solvent front at *c*. 2.8 min.

In respect of the 'cleared' zones, the *E. coli*-pDEST17-*BtMinpp* strain, Figure 2.3a, showed considerable activity, producing multiple peaks of $InsP_5$, $InsP_4$ and $InsP_3$ intermediates, characteristic of a MINPP phytase (Haros *et al*., 2009; Stentz *et al*., 2014; Tamayo-Ramos *et al*., 2012). There is also a larger P_i peak. In the *B. subtilis* strain ESKAPE, Figure 2.3b, there was only a small amount of $InsP_6$ degradation with a concomitant increase in $InsP_5$ (1/3-OH). The increase in the $InsP_5$ (1/3-OH) peak is the expected product of the known $InsP_6$ D-3-phosphatase activity of the BPPhy (Kerovuo *et al*., 1998), however, in this experiment, much of the $InsP_6$ remained. Finally, the *P. putida* J450 strain, Figure 2.3c, showed little difference in the profile of 'cleared' vs. 'non-cleared' agar despite the known $InsP_6$ D-3-phosphatase activity of other *Pseudomonas* sp. (Cosgrove *et al*., 1970; Irving and Cosgrove, 1971).

Collectively, these comparisons demonstrate that zone clearing without careful normalisation is a poor assay for phytate degradation even of well-characterised organisms. In all cases $InsP_6$ remained the dominant form of inositol phosphate. This experiment does illustrate, however, that HPLC can be combined, with media-based culture and extraction of agar for testing of phytate degradation to provide high sensitivity and diagnostic analysis of the likely enzyme activity, by the simple expedient of observation of the occurrence of InsP not present in 'non-cleared' regions of agar plates (Rix *et al*., 2020).

2.2.5. Assay of phytate degradation by mixed population soil cultures

Phytase degradation can be demonstrated with mixed cultures that might ordinarily be subjected to standard dilution and culture techniques for discrimination of individual isolates. For this, a minimal media containing $InsP_6$ as the sole source of phosphate was developed, derived from (Neal *et al*., 2017). The first soil used to test this technique was untilled (for the season) agricultural soil from Fakenham, Norfolk, UK. In this experiment, 0.5 g of agricultural soil was mixed in 10 mL of the minimal media and incubated with shaking at 30 °**C**. Samples were taken each day and ran on the HPLC. Figure 2.4 shows the HPLC traces of Day 0, 3 and 5 which has been focused in to examine the appearance and accumulation of the lower inositol phosphates.

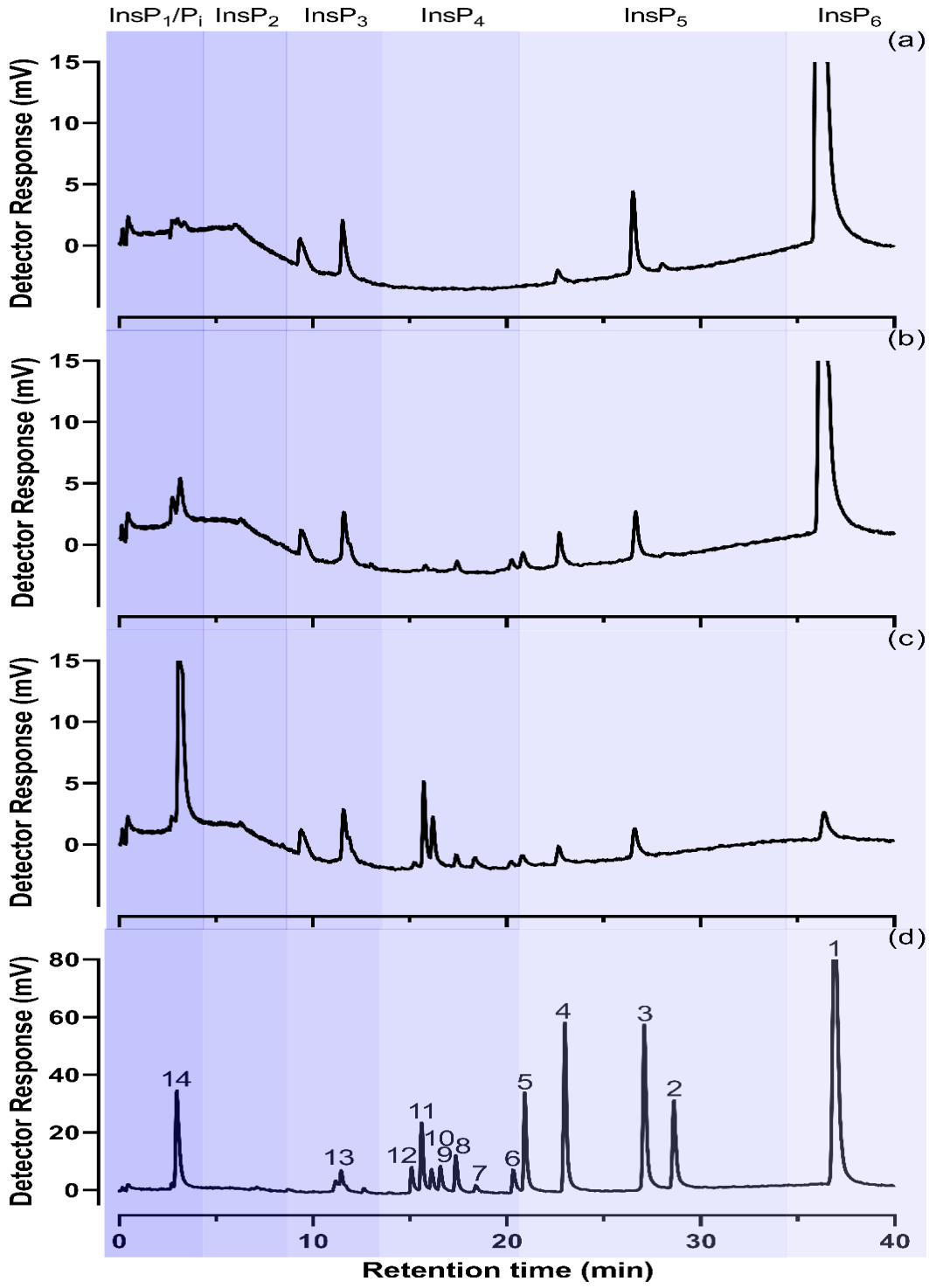


Figure 2.4 - HPLC analysis of phytate degradation in minimal media of a mixed culture isolated from farmland soil, Fakenham. a) day 0, b) day 3, c) day 5, and d) A set of $InsP_n$ standards prepared by acid reflux of phytate: the peaks identified are 1: $InsP_6$, 2: $InsP_5$ (2-OH), 3: $InsP_5$ (1/3-OH), 4: $InsP_5$ (4/6-OH), 5: $InsP_5$ (5-OH), 6: D/L-Ins(1456)P_4, 7: Ins(2,4,5,6)P_4, 8: D/L-Ins(1256)P_4, 9: D/L-Ins(1345)P_4, 10: D/L-Ins(1245)P_4, 11: D/L-Ins(1234)P_4, 12: Ins(1,2,4,6)P_4, 13: $InsP_3$, and 14: $InsP_1$/Pi.

Degradation of $InsP_6$ was initially observed on day 3, Figure 2.4b, and by Day 5 there was a significant increase in the $InsP_1/P_i$ and less than <5% of the starting $InsP_6$ remaining. On Day 3, there was a reduction in the $InsP_5$ (1/3-OH) and (2-OH) peaks that were present on day 0, Figure 2.4a. Alongside this, there was an increase in $InsP_5$ (4/6-OH) and the formation of the $InsP_5$ (5-OH) peak. There was also the formation of three new $InsP_4$ peaks. D/L-Ins(1456)P_4 after the removal of phosphate from the 2 and 3-positions or the 1 and 2-positions. D/L-Ins(1256)P_4 forms after the removal from the 3 and 4-positions or the 1 and 6-positions and D/L-In(1234)P_4 forms after the removal from either the 5 and 6-postion or 4 and 5-positions, these positions agreeing with the changes occurring to the $InsP_5$'s. There is also an increase in the P_i peak, which coelutes with $InsP_1$ on this column-gradient method, which occurs due to the removal of phosphate, leading to the formation of all inositol phosphate peaks. On day 5, Figure 2.4c, there was further degradation of the inositol phosphates with three new $InsP_4$ peaks forming. These were Ins(2,4,5,6)P_4 after the removal of the 1 and 3-positions, D/L-Ins(1245)P_4 after the removal of either 3 and 6-positions or the 1 and 4-positions. Finally, Ins(1,2,4,6)P_4 which is formed from the removal of 3 and 5-positions. Figure 2.4d is a set of InsP standards that have been prepared by acid reflux of $InsP_6$. On Day 6 (data not shown), all the inositol phosphates had been degraded with only a large $InsP_1/P_i$ peak remaining. There was, however, no indication of $InsP_2$ present in this time course and the only indication of $InsP_3$ is the slight broadening of the minimal media component at *c.* 11:46 min where the $InsP_n$ standards predict that the $InsP_3$'s are supposed to run.

The formation of multiple inositol phosphate peaks at all stages of degradation probably arises as a consequence of the action of several phytase enzymes, since the classification of phytases reflects predominant attack in discrete sequences and predominant accumulation of single $InsP_5$ and $InsP_4$ species.

This experiment was repeated on a soil sampled from Church Farm, the field study site of the John Innes Centre in Bawburgh, Norwich UK (TG 15138 08433). Samples were taken on day 0 and day 2, however, degradation was very rapid, with no $InsP_6$, $InsP_5$ and only small amounts of $InsP_4$ present, Figure 2.5.

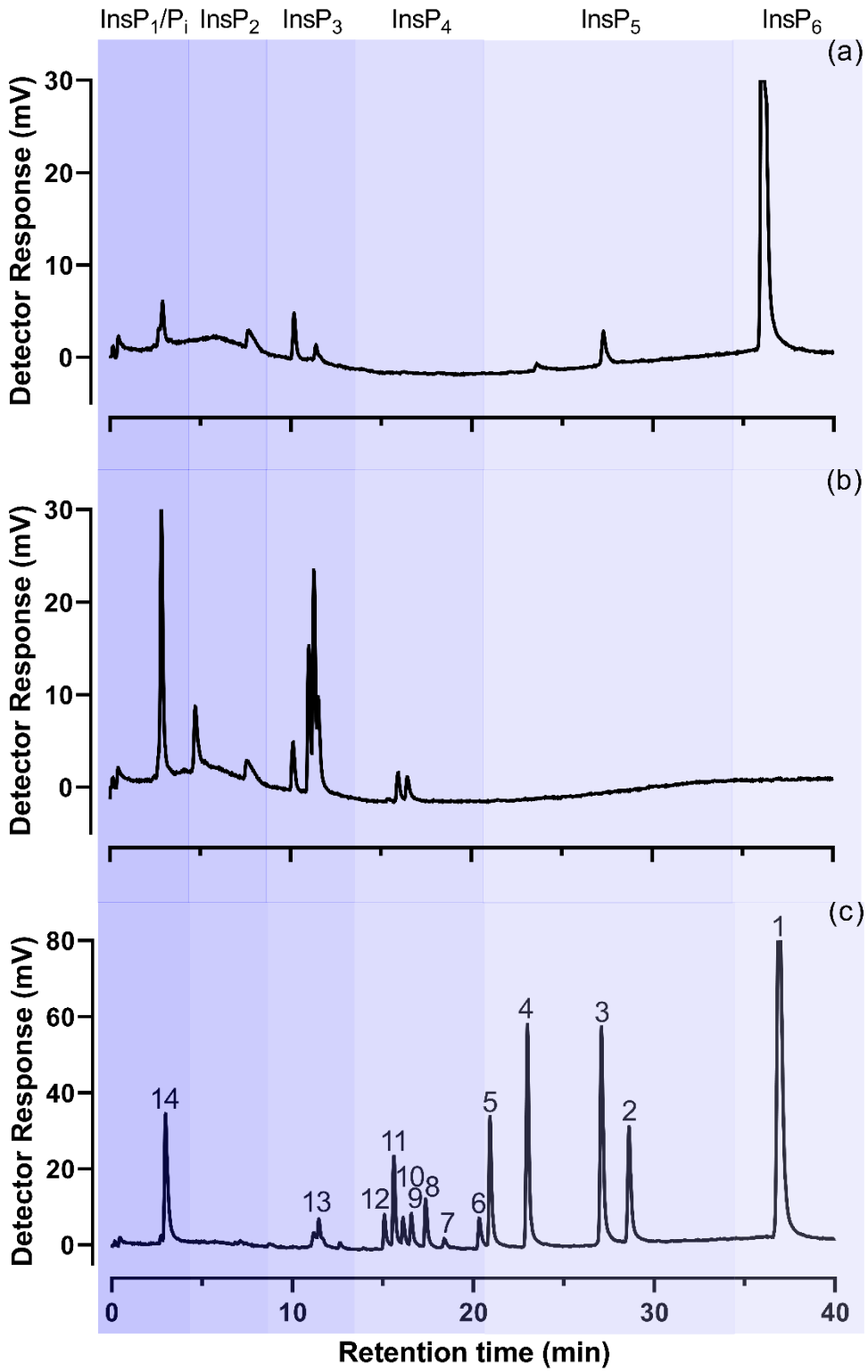


Figure 2.5 - HPLC analysis of degradation of a 1 mM phytate sample in minimal media of a mixed soil culture using soil from Church Farm, John Innes. (a) day 0, (b) day 2 and (c) $InsP_6$ hydrolysate for peak identification.

The soil culture from Figure 2.5 was considerably more active than the Fakenham farmland soil, Figure 2.4. Whilst the farmland soil showed detectable degradation from $InsP_6$ into $InsP_5$ and $InsP_4$, there was no presence of $InsP_6$ and $InsP_5$ and only small amounts of $InsP_4$, one of

which is barely discernible by Day 2 of the Church farm soil, D/L-Ins(1345)P_4, D/L-Ins(1245)P_4, D/L-Ins(1234)P_4. There was also the formation of a broad $InsP_3$ peak that split into three peaks at its apex, a $InsP_2$ peak and lastly a large $InsP_1$/P_i peak.

This technique identified that both of these soils were phytase active with different degradation profiles, as well as different overall phytase activity. Unlike many other experiments that have been described in the literature, this technique makes use of HPLC which definitively confirms phytase activity, without the chance of mischaracterisation due to lower inositol phosphate degradation.

2.2.6. Variation of $InsP_6$ Degradation by Microbial Cohorts When $InsP_6$ is the Sole Source of Carbon and Phosphate.

Phytase enzymes are used as a supplement in animal feeds to improve phosphate bioavailability. These are typically 'superdosed' to ensure the complete degradation of phytate and the other lower inositol phosphates which may also pose an anti-nutritional effect (Broch *et al*., 2018; Lee *et al*., 2017). The complete degradation of phytate leads to the production of *myo*-inositol which can be subsequently absorbed and utilised in a number of biological functions in animals (Lee and Bedford, 2016). Therefore, an important industrial attribute of a bacterial phytase and microbial community is the ability to degrade to, and utilise *myo*-inositol as a carbon source.

In the following experiment, the phytase-active Fakenham farmland soil was sub-cultured into fresh minimal media with and without a mixed carbon source, Section 2.1.1. These were incubated under shaking at 30 °C for one week, with samples taken for HPLC each day. The results are displayed in Figure 2.6.

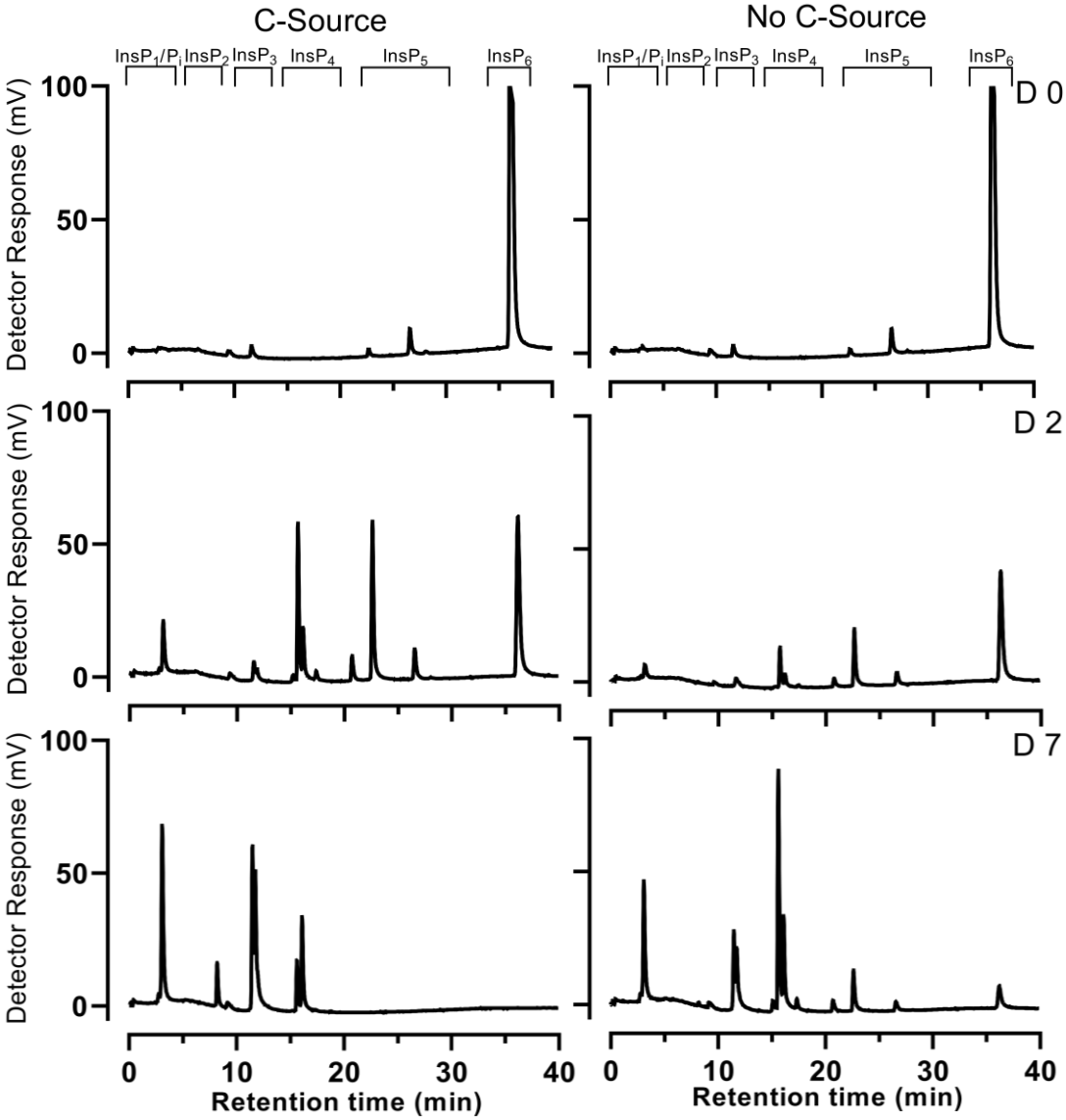


Figure 2.6 – Analysis of the Fakenham Soil's microbial community's response to provision of $InsP_6$ as sole phosphate and carbon source over a 7-day period. Farmland soil from Fakenham was cultured into minimal media with and without a carbon source, and the speed of phytate degradation measured by HPLC on day 0, 2 and 7.

Phytate degradation occurs much more quickly when the soil community has access to an additional carbon source, showing faster degradation of $InsP_6$ into the lower inositol phosphates and concurrent increase in the $InsP_1/P_i$ peak. By day 7, no $InsP_6$ or $InsP_5$'s remained whereas when the soil bacterial community has to utilise phytate as the sole carbon source as well, this degradation was considerably slowed with $InsP_6$ and $InsP_5$'s remaining by day 7 albeit at low concentrations.

Therefore, whilst phytate degradation in the microbial community is slowed by the lack of an additional carbon source, phytate is still processed down to $InsP_2$, $InsP_1$ and P_i with the potential use of *myo*-inositol as the sole carbon source.

2.2.7. Isolation of phytase-producing microorganisms from the environment.

From the actively growing farmland soil culture grown in minimal media, 100 μL, was serially diluted, 10^{-4}-10^{-6}, onto LB and All Culture agar media. These were grown overnight, and single colonies were reinoculated into fresh minimal media and re-streaked onto either LB or All Culture agar to obtain a pure isolate. Of the fourteen isolates that were tested, only eight showed continued phytase activity. Their 16S rRNA was amplified using colony PCR and the results were examined using NCBI nucleotide blast using the Nucleotide collection (nr/nt) database, to determine the bacterial species, Table 2.3.

Table 2.3 – 16S rRNA sequencing of the eight bacterial isolates from Fakenham farmland soil that showed phytase activity. Four isolates were grown on LB and the other four isolates came from All Culture agar.

Isolate 1 – Enterobacterales	Isolate 5 – *Acinetobacter calcoaceticus*
Isolate 2 – *Lelliottia amnigena*	Isolate 6 – *Acinetobacter calcoaceticus*
Isolate 3 – *Serratia* sp.	Isolate 7 – *Acinetobacter calcoaceticus*
Isolate 4 - Enterobacteriaceae	Isolate 8 – *Acinetobacter calcoaceticus*

Isolates 1-4 were isolated from LB plates. Identifying the potential species for Isolate 1 is very difficult due to a 100% percentage identity and identical Max-score for the first ninety-four results. The top bacterial hits were from *Enterobacter* and *Leclercia sp.* Therefore, the isolate has been regarded as an isolate from the Order of Enterobacterales. For Isolate 2 the highest scoring result was for *Lelliottia amnigena* with a Query Cover and Percentage identity of 100%. For Isolate 3, the first 26 results had a 100% sequence identity, however, all of them were from *Serratia* species. The highest scoring results for Isolate 4 were from *Klebsiella*, *Lelliottia*, and *Buttiauxella* species, all with a 100% sequence identity and Max-score of 800. Therefore, this was regarded as an isolate from the family of Enterobacteriaceae.

It is very difficult to determine the correct species of an isolate solely by 16S rRNA sequencing, especially when the primers used covered less than 500 bp which is not ideal when comparing 16S sequences in a database. These isolates all come from the order of Entereobacterales, and they all have very similar sequences when aligned together. However, it has been noted that the resolution at the genus and/or species level with 16S rRNA data for many bacterial groups is common, this includes the family Enterobacteriaceae (Janda and Abbott, 2007, 2002). Therefore, the species determination for Isolate 1-4 in Table 2.3 is not conclusive.

Isolates 5-8 were more easily determines. These were all isolated from All culture plates, all of the top scoring 16S rRNA sequences from *Acinetobacter* sp., indicating that they may be the same isolate.

To investigate the distribution of different phytases between the bacterial species highlighted in Figure 2.3 and the 16s rRNA sequencing information. BLAST searches were conducted using ratified examples of each of the five phytase families. For the Purple Acid Phytases (PAPhy), where originally most of the known sequences originate from plants, we followed Nasrabadi *et al.,* (2018) using *Lupinus luteus* (AJ505579) as query to blast *Sphingobium spp.,* genomes, returning hits with percentage identity 23-26% with E value of 5^{-12} to 9^{-10} (Ghorbani Nasrabadi *et al*., 2018). Subsequently, the full gene from *Sphingobium yanoikuyae* (CP060122) was used as query. The reference sequences used were as follow: multiple inositol-polyphosphate phosphatase (Minpp), *Bacteroides thetaiotaomicron* (WP_009040027); histidine acid phytase (HAPhy), *Citrobacter amalonaticus* (DQ975370.1; beta-propeller phytase (BPPhy) *Bacillus amyloliquefaciens* (WP_013352583); protein tyrosine phytase (PTPhy), *Selenomonas lacticifex* (ABC69367) and Purple Acid Phytase (PAPhy), *Sphingobium yanoikuyae* (CP060122), Table 2.4.

Table 2.4 – The number of different phytase families identified in genera related to Isolates 1 to 8 obtained from Farmland Soil. These were determined using NCBI blast that was limited to search for each individual family of canonical phytases. The max target sequences were set to 1000, with an E-value cutoff at $1e^{-10}$.

Genus	*HAPhy*	*BPPhy*	*PAPhy*	*PTPhy*	*MINPP*
Enterobacter sp.	528	0	0	0	0
Leclercia sp.	20	0	0	0	0
Pantoea sp.	616	4	3	0	0
Lelliottia sp.	16	0	0	0	0
Serratia sp.	298	1	0	0	0
Buttiauxella sp.	24	0	0	0	0
Acinetobacter sp.	9	80	1	0	445

From analysis of the deposited data from NCBI, isolates 1-4 most likely contains a histidine acid phytase. Isolates 5-8, which from now will be called *Acinetobacter* sp. AC1-2, taking a colony from Isolate 5, however, most likely contains a MINPP, a subclade of the histidine acid phytases. This isolate was subjected to further HPLC analysis to determine the profile of inositol phosphates generated from phytate over two days, Figure 2.7.

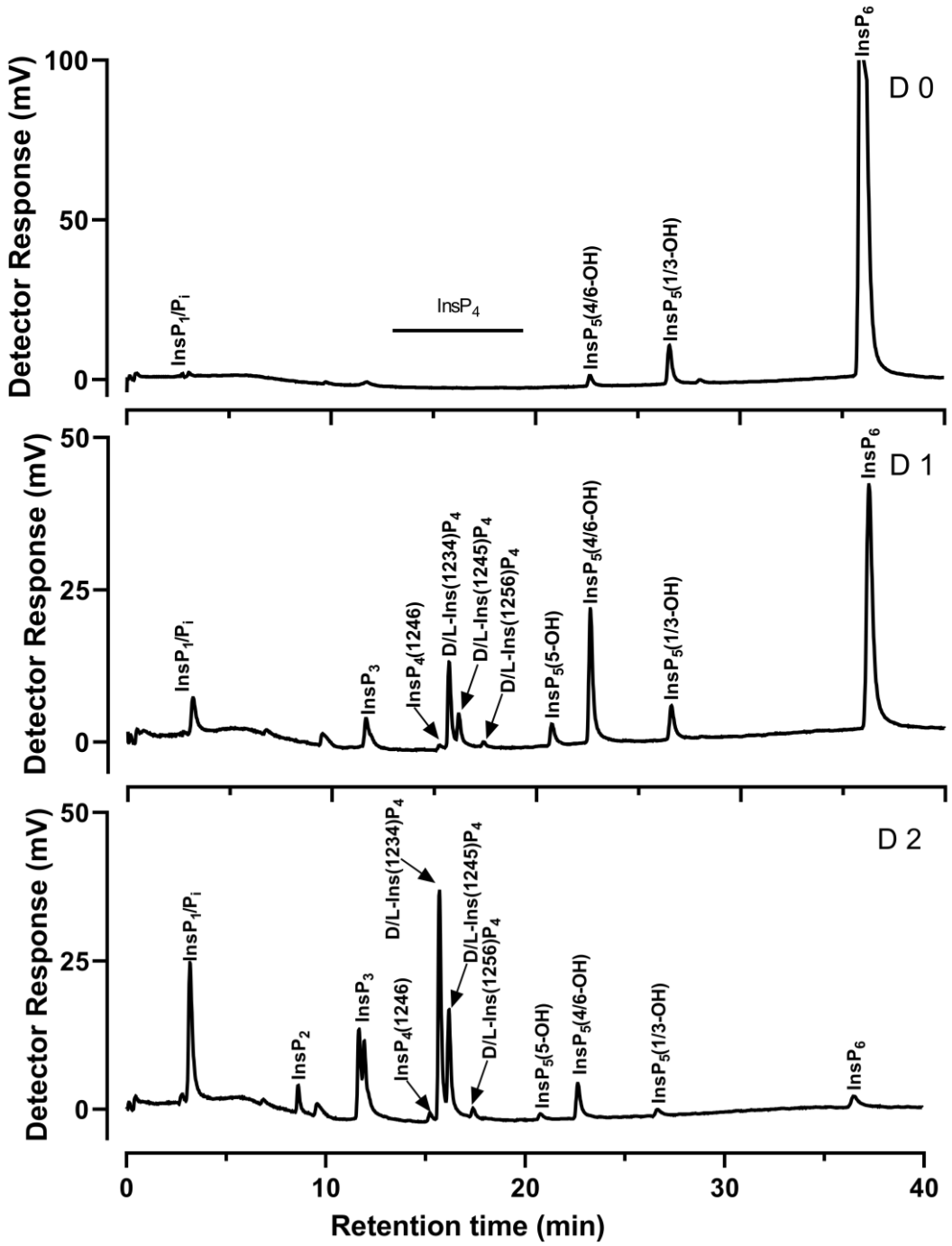


Figure 2.7 – HPLC chromatogram of the degradation profile of 1 mM Phytate by *Acinetobacter* sp. AC1-2 over two days. Samples were taken on Day 0, 1 and 2. The y-axis on Day 1 and 2 is shortened to allow for visualisation of the smaller peaks.

HPLC analysis of the degradation profile of *Acinetobacter* sp. AC1-2 demonstrates the phytase associated with AC1-2 is promiscuous in their site of initial attack on the phytate substrate, yielding among $InsP_5$ isomers a dominant 4/6-OH peak, a smaller 5-OH peak and little to no detectable degradation at the 1/3-position. There was also the generation of four $InsP_4$ intermediates, D/L-Ins(1256)P_4, D/L-Ins(1245)P_4, D/L-Ins(1234)P_4 and Ins(1,2,4,6)P_4, with the promiscuous behaviour, the ability to produce multiple phytate degradation products, strengthening AC1-2's credentials as a Minpp, the genome was sent for sequencing.

This technique was repeated on phytase-active Churchfarm soil, Figure 2.5, with one exception. The minimal media that had previously been used in these experiments contains trace metals of Fe, Zn, Cu and Mn, which have been shown to act as phytase inhibitors (Maenz *et al*., 1999; Santos *et al*., 2015). Therefore, in this instance, potential phytase producers were analysed in either 0.2 M sodium acetate buffer pH 3.5, 5.5-, or 20-mM Tris-HCl pH 7 each containing 1 mM $InsP_6$ and 0.1% NaCl. Here another phytase-producing isolate was identified that had not been growing nor producing phytases in the minimal media. This isolate was designated Ch 10^{-6}-4 Ac and sent for 16S rRNA sequencing, which identified the isolate's 16S rRNA gene as identical to that of *Buttiauxella agrestis* strain EB112, *Buttiauxella* sp. CL_136_AN_40 and *Buttiauxella* sp. SA_136_AN_45. The phytate degradation pattern of the strain is shown in Figure 2.8.

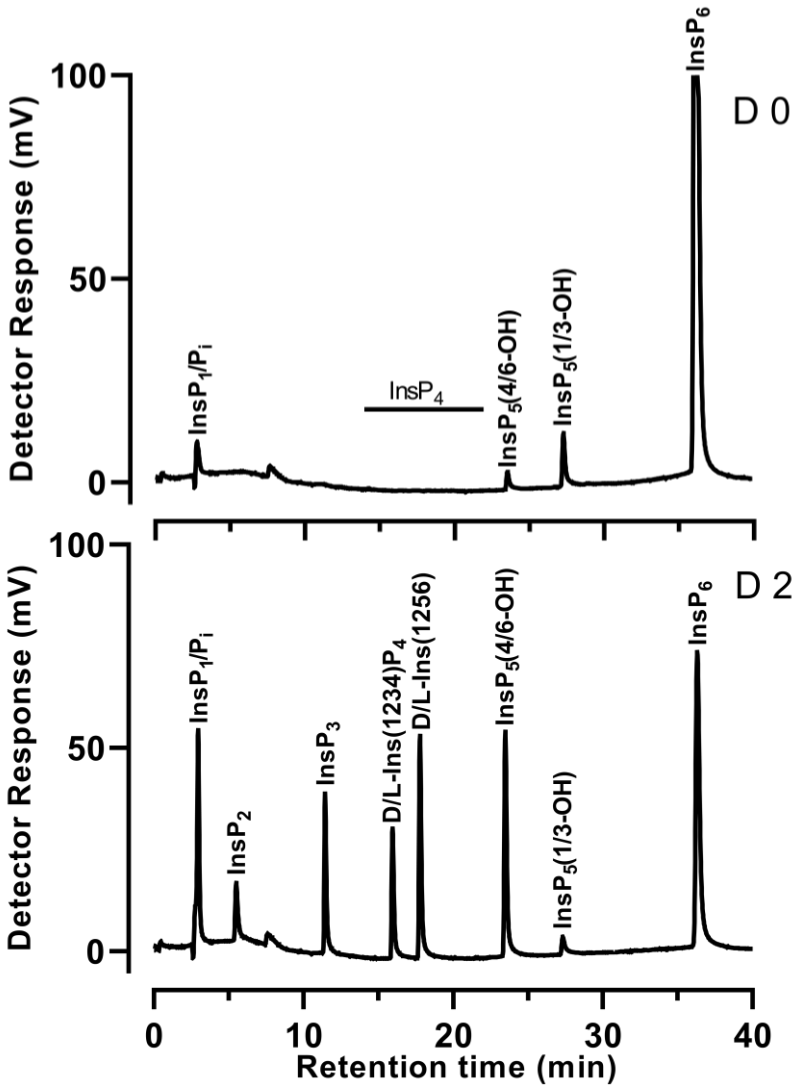


Figure 2.8 – HPLC chromatogram of the degradation profile of 1 mM Phytate by *Buttiauxella* sp. CH 10^{-6}-4 Ac over two days. Samples were taken on Day 0 and Day 2.

The *Buttiauxella* sp. CH-10^{-6}-4 Ac showed a high specificity towards the initial position of attack on phytate, generating $InsP_5$ (4/6-OH) predominantly among $InsP_5$ products, consistent with the published properties of *Buttiauxella* phytase (Cervin *et al.*, 2008) and its industrial use (Herrmann *et al*., 2019; Ushasree *et al*., 2017).

2.2.8. Identification of phytases present in *Acinetobacter sp.* and *Buttiauxella* sp.

While both *Acinetobacter* and *Buttiauxella* strains showed preferential 1D-4/6 selectively of attack on phytate, they differ in terms of the resulting $InsP_4$ intermediates: the *Acinetobacter* strain produced four $InsP_4$ intermediates, while the *Buttiauxella* strain produced two, a predominant peak with the chromatographic properties of D/L-Ins(2,3,4,5)P_4 and a minor peak of D/L-Ins(1,2,3,4)P_4. Again, HPLC can be shown to distinguish between classes of

phytase without assistance of 16S rRNA gene, Table 2.3. The phytate degradation profile of the *Buttiauxella* isolate is characteristic of 1D-6-directed histidine acid phytase, that of the *Acinetobacter* strain was indicative of the MINPP subclass of the HAPhys (Stentz *et al*., 2014; Tamayo-Ramos *et al*., 2012).

To confirm that the isolated *Buttiauxella* sp. CH-10^{-6}-4 AC contained a histidine acid phytase, degenerate primers were designed to the *appA* gene using 12 sequenced *Buttiauxella spp*. genomes obtained from NCBI blast using *Citrobacter amalonaticus* (DQ975370.1) as the query sequence. Alignments are shown in Figure 2.9.

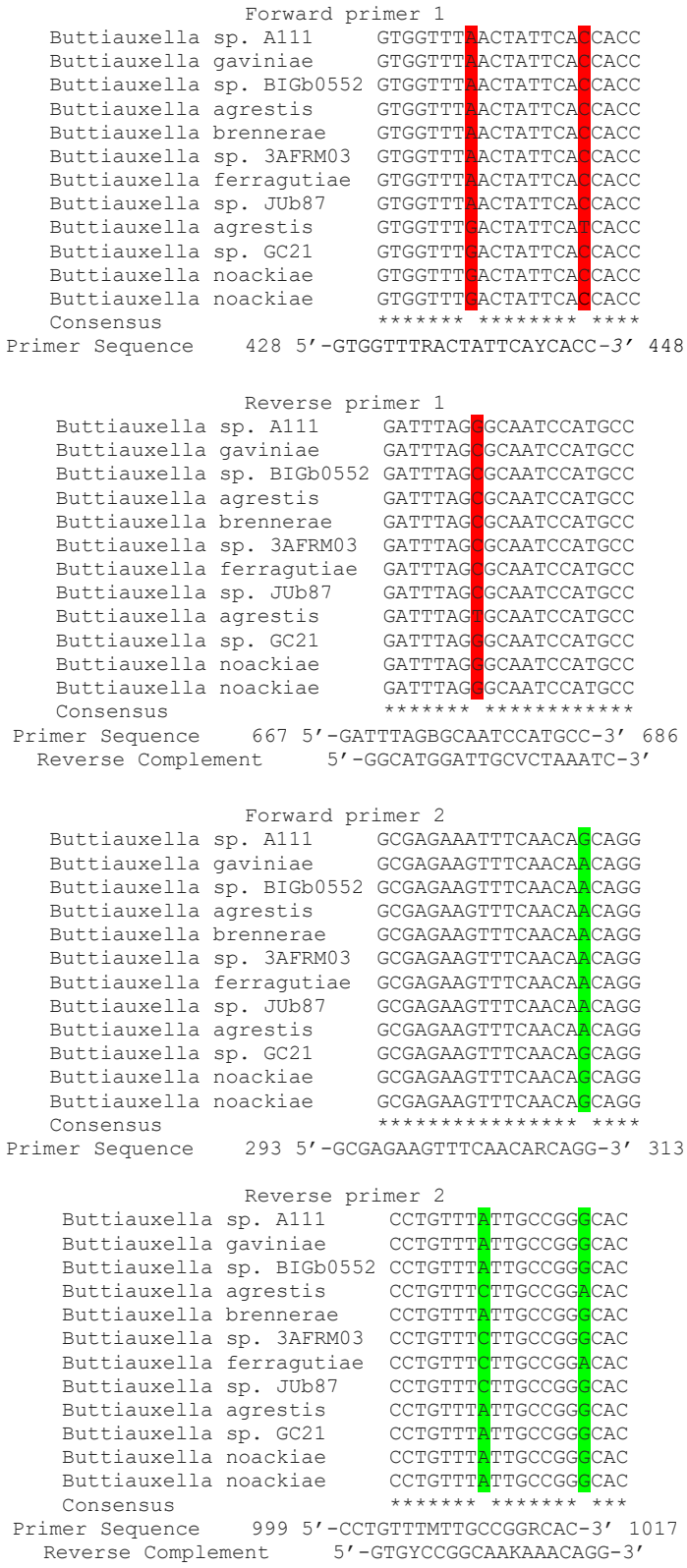


Figure 2.9 – The nucleotide alignment of twelve *appA* phytases from *Buttiauxella* species using MUSCLE. Degeneracy in the sequences are highlighted in red and green.

These sequences were aligned in MUSCLE and conserved regions chosen for each primer set, producing 259 and 725 bp products. These primers were used on the *Buttiauxella* isolate in triplicate by colony PCR, the gel image is displayed in Figure 2.10.

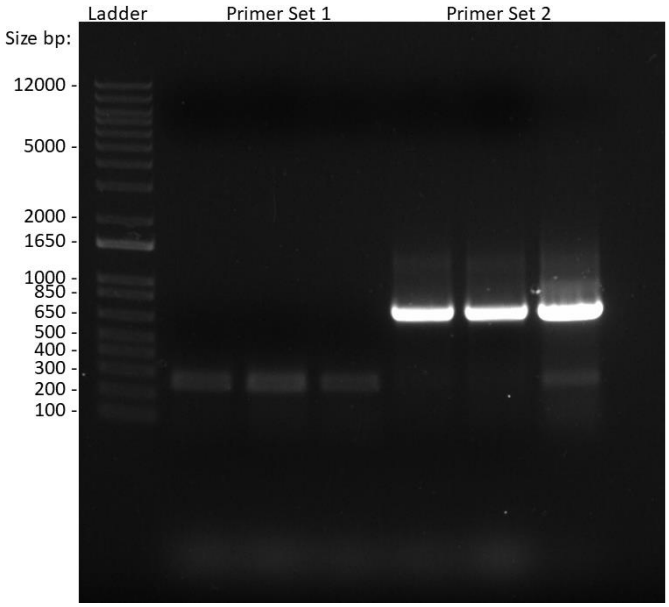


Figure 2.10 – 1% agarose gel of the two degenerate *appA Buttiauxella* sp. primers. Lanes 1-3, primer set 1, lanes 4-6, primer set 2. Expected bands for Primer set 1 and 2 were 259 and 625 respectively.

The amplified DNA was sent for sequencing, the results showed that *Buttiauxella* sp. CH 10^{-6}-4 Ac contained a histidine acid phytase that was 100% identical at the amino acid level to that in *Buttiauxella ferragutiae* (Accession WP_083963259).

The *Acinetobacter* sp. AC1-2 isolate was sent for genome sequencing (microbesng), and the assembled sequence genome (JABFFO000000000) was found to harbour a single histidine acid phosphatase of the MINPP class, rather than a canonical HAPhy as predicted by its HPLC activity. The gene contained the heptapeptide sequence motif RHGSRGL, a catalytic motif that is typical in histidine acid phytases, RHGxRxP (Lei and Porres, 2003). But it contained a tripeptide proton donor motif, HAE, instead of the dipeptide motif that are common in HAPhys, this tripeptide motif is a useful identifier in MINPPs (Stentz *et al*., 2014). The *Acinetobacter* AC1-2 phytase contained 523 amino acids and a molecular weight of 58.6 kDa, calculated by ProtParam (Gasteiger *et al*., 2005). This phytase is considerably larger than many of its histidine acid phytase counterparts, as well as other MINPPs, Table 2.5, with the exception of *Bifidobacterium longum* with which the phytase shares its closest similarity. The sequence of the phytase is displayed in Table 2.5.

Table 2.5 – The size analysis of various Histidine acid phytases (HAPhy) and Multiple inositol polyphosphate phosphatase (MINPP) with comparison to the *Acinetobacter* sp. AC1-2 isolated in this study. The sequence of the *Acinetobacter* phytase gene with the catalytic motif and proton donor motif highlighted.

Phytase Type	*Bacterial Species*	*Size (kDa)*	*Reference*
HAPhy	*Escherichia coli*	42	(Greiner *et al*., 1993)
HAPhy	*Citrobacter amalonaticus*	46.3	(Li *et al*., 2015)
HAPhy	*Hafnia alvei*	46.3	(Ariza *et al*., 2013)
HAPhy	*Yersinia mollaretti*	47	(Körfer *et al*., 2018)
HAPhy	*Buttiauxella* sp.	45.4	(Yu *et al*., 2014)
MINPP	*Bacteroides thetaiotaomicron*	49.2	(Stentz *et al*., 2014)
MINPP	*Bifidobacterium longum*	67	(Acquistapace *et al*., 2020)
MINPP	*Aeromicrobium* sp.	45.2	(Rigden, 2008)
MINPP	*Streptomyces* sp.	49.8	NCBI – Blast WP_123466415
MINPP	*Amycolatopsis jejuensis*	49.1	NCBI – Blast WP_033295522
MINPP	*Acinetobacter calcoaceticus*	58.4	NCBI – Blast WP_017389539
MINPP	*Acinetobacter* sp. AC1-2	59.6	This Study

Acinetobacter sp. AC1-2
MNILFKTTMLATSLFLVACNNNDDQDDQPTTSPTTQSKYYQTKTPYQPQQDLKSYEQAPN
GFQPVFTELVARHGSRGLSSLKYDLALYNLWKQAKAENALTPLGEQLGADLEAMMKANIL
LGYGVEGIRQYGYGNETMTGILEHRGIADRLLQRLPNLLNPQAGILVQSSGVDRAVDSAK
FFTAELIKQQPQLKDKIVPVSYTNLSSESVPSVIDGGVDRFKLYFHSLNADEDLTQPLSA
SQQKIYDASQAYQDFEENNNDLAQKLDELSKNTQAEKTAQTVLNPIFKADFIKKLGTAGY
SFSNTGSFTVTSPKGEQITEKGKGKNTIASAVDAAAYVYELYSISGGMKDELKGIDFDKY
MPIEAAKFYAEFNDANDFYEKGPSFTESNLVTSEIAQGLKQDMFQQVDAVVNKAQPYKAV
LRFAHAEIIIPLATSLDLHNMMQPLPLRQTYNYSTSTWRGEVVSPMAANVQWDIYQNGQG
NTLVKMLYNEKETLFKSACNYARYSPTSFYYDYIKLKQCYQIQ*

The genome of *Acinetobacter* sp. AC1-2 encoded a MINPP 98.28% identical at the amino acid level to that in *Acinetobacter calcoaceticus* (Accession – WP_017389539). To ensure the genome was of good quality, genomic completeness was analysed using BUSCO v3 (Simão *et al*., 2015). BUSCO is an open-source software that provides quantitative measures for genomic completeness based on evolutionarily informed expectations of gene content from near-universal single-copy orthologs. The *Acinetobacter* sp. strain AC1-2 was analysed using

both BUSCO's bacterial and Gammaproteobacterial databases with completeness measured at 98 and 98.9% respectively.

2.2.9. Evolutionary differences between *Acinetobacter* sp. AC1-2 and *Buttiauxella* sp. CH 10^{-6}-4 Ac.

To interrogate the evolutionary differences between the *Acinetobacter* sp. AC1-2 and *Buttiauxella* sp. CH 10^{-6}-4 Ac phytases, an alignment of thirty-one Histidine Acid Phytases and twenty-seven MINPPs was made using the online multisequence alignment tool (Katoh *et al*., 2019), with the output reported as an Interactive Tree of Life (iTOL) (Letunic and Bork, 2019), Figure 2.11. The results of this analysis split MINPP sequences into two clades, those whose origins are from animals and plants (Cho *et al*., 2006; Dionisio *et al*., 2007), and those from bacteria (Haros *et al*., 2009; Stentz *et al*., 2014; Tamayo-Ramos *et al*., 2012). Both are distinct from bacterial Histidine Acid Phytases, with bacterial MINPPs more closely related to eukaryotic MINPPs than bacterial HAPhys. Of the bacterial MINPPs, the *Acinetobacter* enzyme was more deeply rooted than the MINPPs of previously characterised gut commensals *Bifidobacterium* and *Bacteroides* spp.

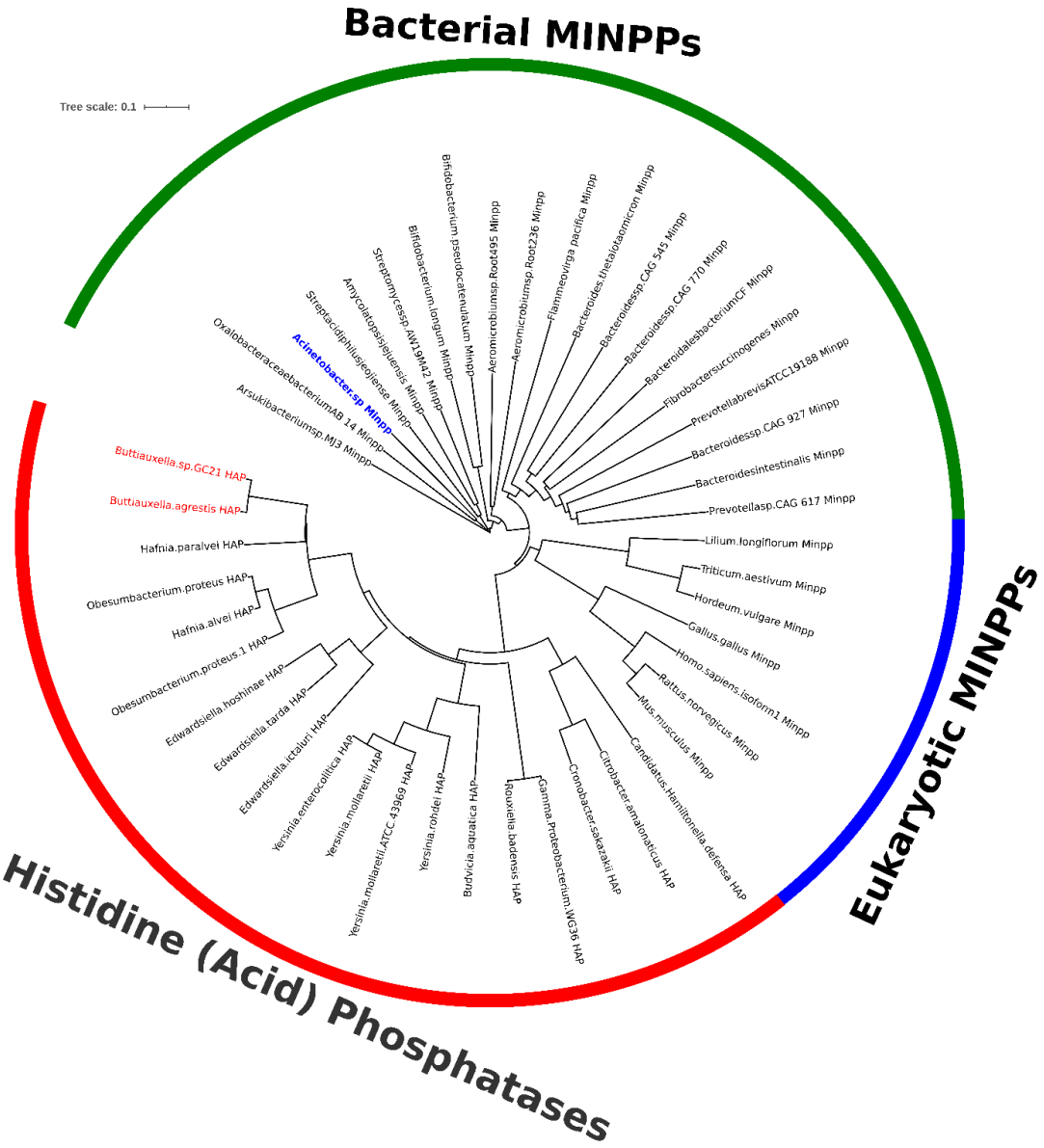


Figure 2.11 – Phylogram of thirty-one histidine acid phytases (HAPhy) and twenty-seven multiple inositol-polyphosphate phosphatases (Minpp) showing the evolutionary differences between the two sets of genes. The *Acinetobacter* sp. AC1-2 gene sequenced (JABFFO000000000) is highlighted in blue. The *Buttiauxella* strains highlighted in red is a species similar to that identified by 16S rRNA sequencing of CH 10^{-6}-4 Ac (accession MT680195).

2.2.10. *Acinetobacter* sp. AC1-2 – Bacterial Pathways

Bacteria from the *Acinetobacter* genus are strictly aerobic, gram-negative, oxidase-negative organisms that tend to be paired non-motile cocci when viewed under the microscope. One of the characterisitic features of this genus is their inability to utilise glucose as the sole source of carbon and energy (Beardmore-Gray and Anthony, 1986). Many members are able to oxidise glucose to gluconic acid by the direct route, as the only pathway for glucose oxidiation. This reaction is catalysed by a quinone glucose-dehydrogenase which requires

pyrroloquinoline quinone (PQQ) as a coenzyme, both of which are present in the *Acinetobacter* sp. AC1-2 genome, PQQ in the form of its synthase. An analysis of the glycolysis/gluconeogenesis pathway using the KEGG (Kyoto Encyclopedia of Genes and Genomes) database and their BlastKOALA tool was performed. Molecular functions are stored in the KO (KEGG Orthology) database containing orthologs of experimentally characterised genes/proteins. BlastKOALA then assigns these KO identifiers to individual genes in the genome. These are then used to reconstruct the KEGG pathways enabling interpretation of high-level functions (Kanehisa, 2019, 2000). The glycolysis pathway is shown in Figure 2.12, the genes present in the *Acinetobacter* sp. genome are highlighted in green.

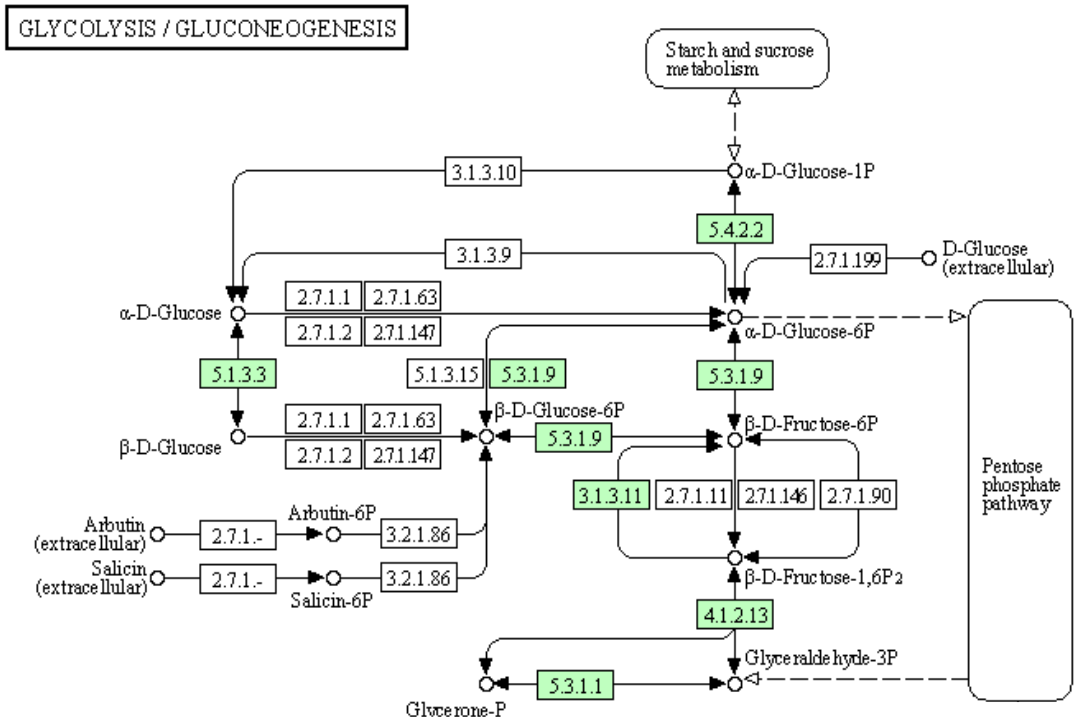


Figure 2.12 – The KEGG pathway analysis of the glycolysis/gluconeogenesis of *Acinetobacter* sp. AC1-2. *Acinetobacter* sp. are unable to utilise glucose as the sole source of carbon and energy as they are lacking the fundamental D-glucose phosphotransferase (2.7.1.199) which catalyses the formation of glucose-6-phosphate from glucose.

This model indicates that *Acinetobacter* sp. AC1-2 is missing EC 2.7.1.199 which encodes a D-glucose phosphotransferase which transfers a phosphate group from ATP to D-glucose forming D-glucose-6-phosphate. Later on in the glycolysis cycle, it is also missing EC 2.7.1.11 and 2.7.1.146 which encodes a fructose-6-kinase and ADP-6-phosphofructokinase respectively which catalyses the formation of D-fructose-1-6-bisphosphate from fructose-6-phosphate.

Therefore, if *Acinetobacter* sp. AC1-2 cannot get energy from glucose, there must be an alternative pathway or use for it (Van Schie *et al*., 1987). One such use, is the Entner-Doudoroff pathway which can be used as an alternative to glycolysis. Unique enzymes such as phosphogluconate dehydratase and 2-keto-deoxy-6-phosphogluconate (KDPG) aldolase are used, both present in AC1-2's genome, along with other common metabolic enzymes to catalyse the formation of pyruvate (Conway, 1992). This process yields 1 ATP and 1 NADP and 1 NADPH molecules in comparison to glycolysis which forms 2 ATP and 2 NADP per glycose. However, *Acinetobacter* sp. AC1-2 lacks pyruvate kinase, EC 2.7.1.40, which may indicate why it doesn't grow on glucose and cannot use this pathway. This has been identified in other *Acinetobacter* species such as *Acinetobacter baylyi* (Young *et al*., 2005).

Even if the gluconic acid cannot be utilised via the Entner-Doudoroff pathway it is not without any use. As a small-molecular weight acid, gluconic acid can be used to acidify the cell surroundings, AC1-2 contains a high-affinity gluconate transporter. This mechanism can be extremely useful in phosphate solubilisation in the environment. As has already been described, phosphate is bound in many inaccessible forms in the environment that renders it unavailable for microorganism and plant uptake (Prabhu *et al*., 2019). The acidification of phosphate through the secretion of organic acids, however, has been demonstrated frequently throughout the literature. This acidification leads to the release of phosphate ions from their inaccessible forms due to H^+ substitution for the cation bound to phosphate (Goldstein, 1994; Sharma *et al*., 2013). Evidence from an abiotic study using HCl and gluconic acid to solubilise phosphate indicated that chelation of Al^{3+} by gluconic acid may have been a factor in the solubilisation of colloidal aluminium phosphate (Whitelaw *et al*., 1999). Gluconic acid was the major organic acid produced during phosphate solubilisation as reported for *Acinetobacter calcoaceticus, Azospirillum lipoferum, Burkholderia cepacian,* and *Serratia marcescens* (Gulati *et al*., 2010). The use of *Acinetobacter* species for phosphate solubilisation and plant growth promotions have also been examined. *Acinetobacter* strains were shown to be able to colonise the roots of *Oryza sativa,* after inoculation of chickpea and barley there was an increase of 15-20% and 56-66% of dry matter weight with respect to the uninoculated control, phosphate content per plant was also significantly higher (Islam *et al*., 2007; Ogut *et al*., 2010; Peix *et al*., 2009).

2.2.11. *Acinetobacter* sp. AC1-2 growth on different carbon sources

If *Acinetobacter* sp. AC1-2 is unable to utilise glucose as the sole carbon source, then what carbon sources can it grow on?

To answer this question, *Acinetobacter* sp. AC1-2 was grown on different carbon sources in solid and liquid cultures to identify what it can grow on as the sole carbon source. The carbon sources used in this experiment were 1 mM citrate, pyruvate, succinate, glucose, D-mannitol, D-galactose, myo-inositol, sucrose compared with a no-carbon source control. In this experiment, colonies grew on citrate, pyruvate, and succinate, and did not grow on the other carbon sources, Figure 2.13. The same results occurred in the liquid cultures. Interestingly with glucose as the sole carbon-source, inoculations into the culture showed no growth, however, a significant decrease in pH from 7 to below 5 occurred. With the decrease most likely due to the production of gluconic acid that is insufficient to sustain growth.

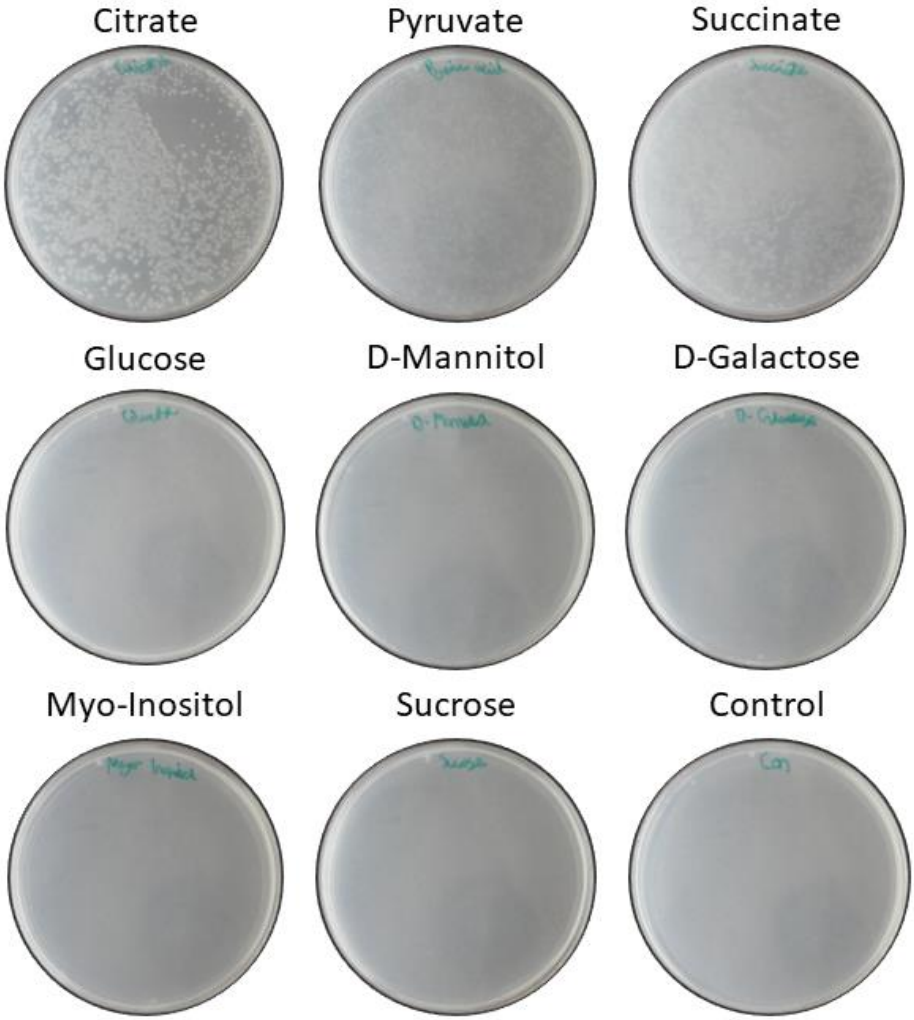


Figure 2.13 – *Acinetobacter* sp. AC1-2 grown onto solid agar media with a 1 mM sole carbon source.

Following from this, the KEGG pathway analysis for the Citric Acid Cycle was examined to examine the completeness of this cycle and investigate the genes involved. Despite missing several key enzymes for glycolysis, it contains all the key enzymes necessary for this cycle, Figure 2.14. Also note that succinate, citrate and pyruvate, the only sole sources of carbon that AC1-2 would grow on, are key substrates in this cycle.

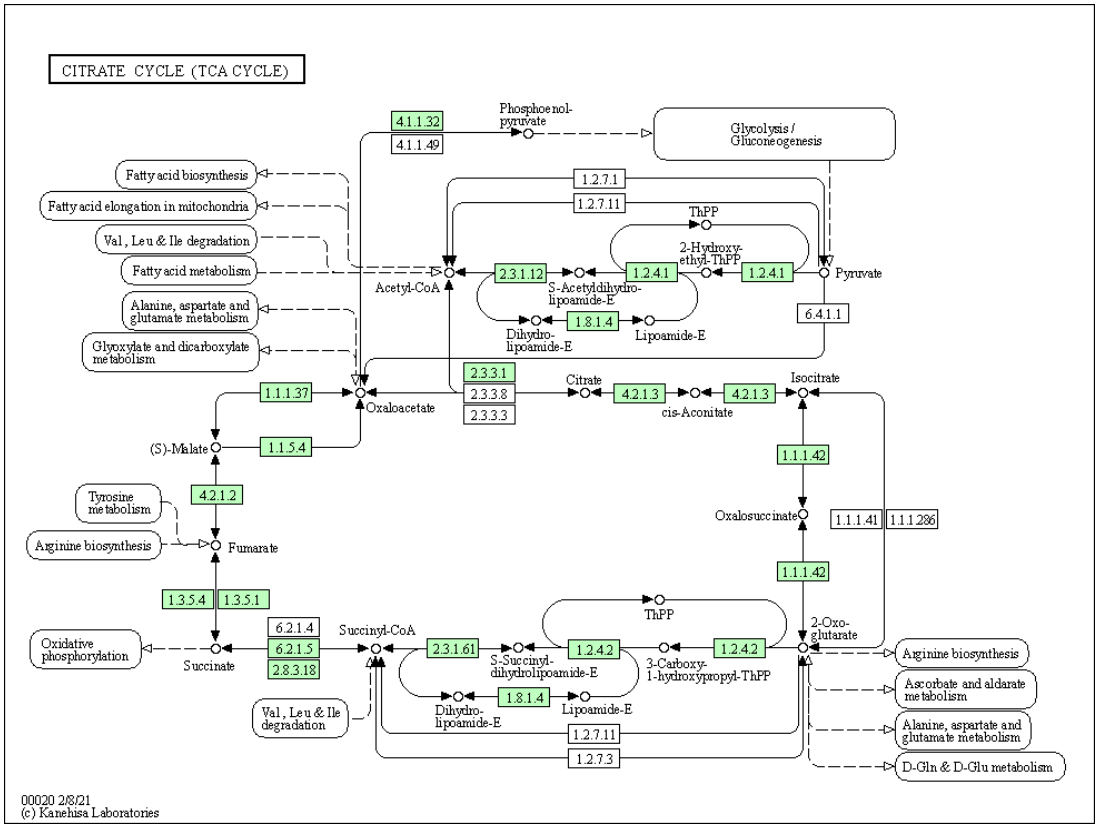


Figure 2.14 – The KEGG pathway analysis of The Citric Acid Cycle of *Acinetobacter* sp. AC1-2. The main energy yielding pathway of this organism is directed to respiration by the citric acid cycle. Demonstrating why AC1-2 can grow on carbon sources such as succinate, citric and pyruvic acid.

2.3. Using culture-dependent techniques to investigate diversity of phytases amongst soil and salt marsh environments

2.3.1. Measuring phytate degradation in Rothamsted soils

In this Chapter, new techniques have been developed to accurately and effectively isolate phytases from the soil environment, whilst avoiding the potential pitfalls that are frequently underreported in the literature.

Building on the techniques developed in this Chapter, a long-term phytase isolation experiment was performed on three well-characterised soil environments obtained from Rothamsted Research to examine the diversity of phytate degradation.

When using culture-dependent techniques to isolate phytase-degrading microorganisms, it is first important to determine the environment is biologically active as a whole. As shown in Figure 2.4, it is possible to detect phytase activity in the soil environment using HPLC. This time, the experiment was repeated on three well-characterised soil environments obtained from Rothamsted Research. The first sample was obtained from continuous arable plots growing winter wheat (*Triticum aestivum* L.) of the Highfield Ley-Arable experiment, Figure 2.15a. Also from this site, soil was sampled from permanent bare fallow plots that had been maintained crop- and weed-free by regular tilling for over 50 years, Figure 2.15b. Lastly, soil was collected from a plot of the Broadbalk Winter Wheat experiment which had been running since 1843, Figure 2.15c

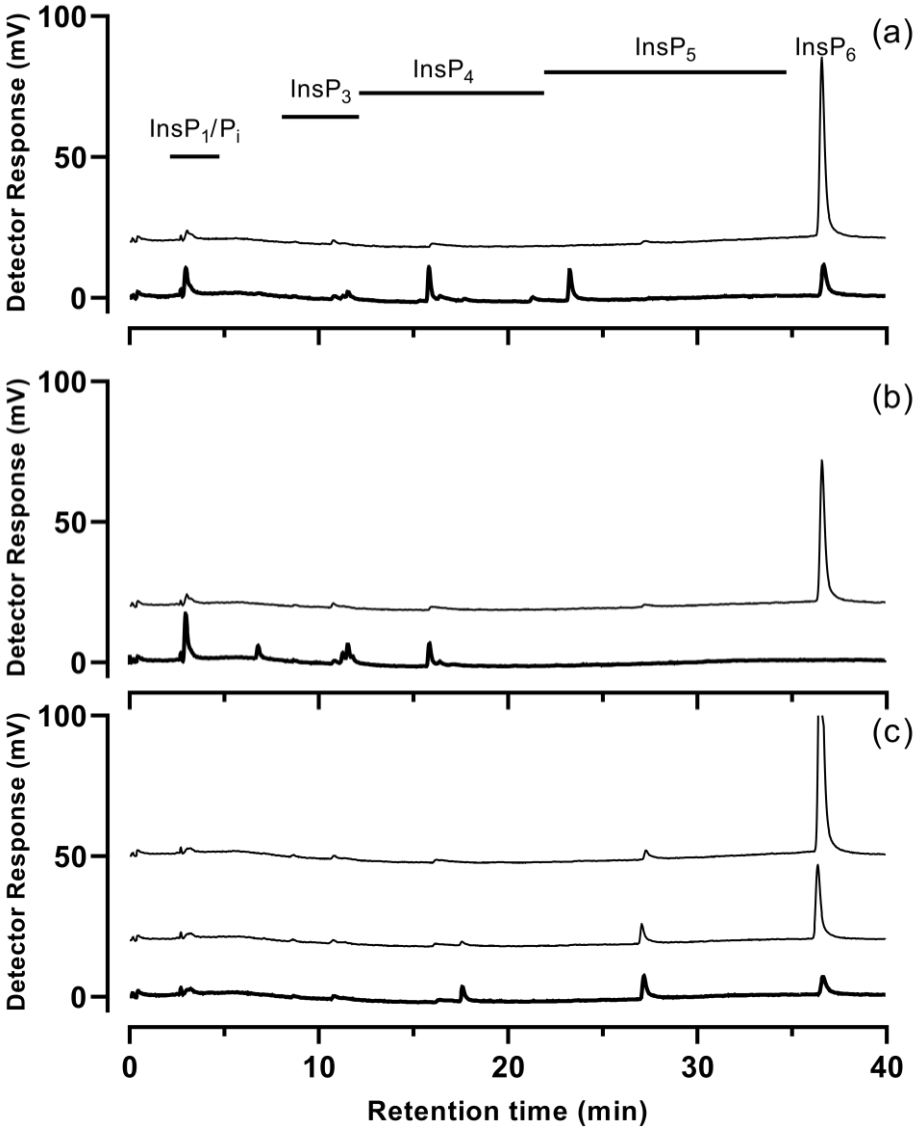


Figure 2.15 – HPLC chromatogram of the degradation profile of 1 mM Phytate by three different soil environments, Arable (a), Barefallow (b) and Broadbalk (c) were obtained from Rothamsted Research. The broadbalk soil (c) was spiked with an additional 1 mM phytate as all of the phytate had been adsorbed and couldn't be detected. The traces were offset on the Y-scale. Light grey lines, day 3 (c), black lines, day 1 (a, b) and day 7 (c), dark black line, day 2 (a, b) and day 8 (c).

All of the soil types when grown in liquid culture showed distinct profiles of inositol phosphates concurrent with disappearance of the peak of $InsP_6$. Lower inositol phosphates were generated at different rates in the different soils. The Highfield ley arable soil, Figure 2.15a, had degraded most of the $InsP_6$, with peaks of $InsP_5$s, $InsP_4$s and $InsP_3$s appearing on the trace. Whereas in the bare fallow soil, Figure 2.15b, there was no presence of $InsP_6$ nor $InsP_5$. Instead, $InsP_4$s, $InsP_3$s and $InsP_2$s were present with the largest peak that of $InsP_1$ and P_i. For the Broadbalk soil, Figure 2.15c, the experiment was extended because the Broadbalk soil removed phytate from the liquid media such that neither phytate nor lower inositol

phosphates were recovered. This may be attributable to adsorption of phytate to soil particles as this soil has clay contents up to 39% (Watts *et al*., 2006). Consequently, this condition was supplemented with additional 1 mM phytate. The soil subsequently degraded phytate, monitored over eight days. Additionally, while the Highfield arable and bare fallow soils showed an accumulation of inorganic phosphate in the media, this was not the case for the Broadbalk soils. Interestingly, there were differences in the dominant $InsP_5$s detectable in the arable and Broadbalk soils, in the arable soils the $InsP_5$(4/6-OH) was the dominant peak, whereas the Broadbalk yielded predominantly $InsP_5$(1/3-OH), highlighting how treatment of the soil can affect the way in which phytate is degraded, with different phytases at work in differing environments.

2.3.2. Isolating phytases from Rothamsted soils

Following this, each of these soil types were grown in liquid phytase isolation media for three days, before being serially diluted onto a variety of agar plates which had been supplemented with phytate. These plates were incubated at room temperature for 1 month to allow the development of slow-growing soil bacteria. It is very common in isolation efforts, including phytase-isolation, to only incubate cultures for less than a week (Jorquera *et al.*, 2011; Kumar *et al.*, 2013), and then isolate and examine single colonies. However, many bacteria that require longer incubation times are often overlooked, instead the faster growing, more common bacteria are isolated (Davis *et al*., 2005; Pham and Kim, 2012). After 1 month of growth, 80 single colonies from the plates were selected and streaked onto phytase-specific medium, additionally 49 isolates were streaked onto minimal media, LB 1/10 and 1.100 and blue green plates, however, many isolates were lost due to fungal contaminants, Figure 2.16.

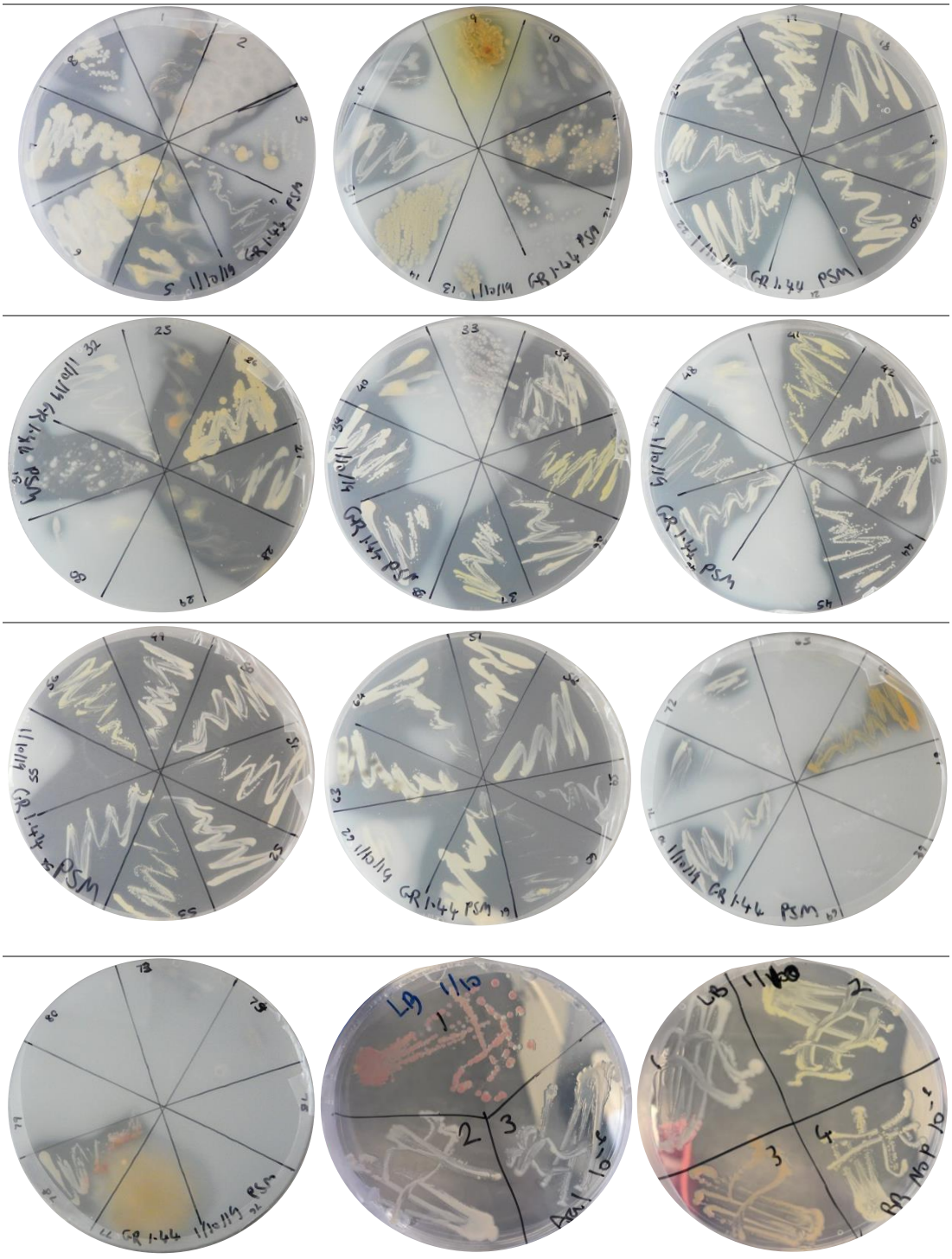
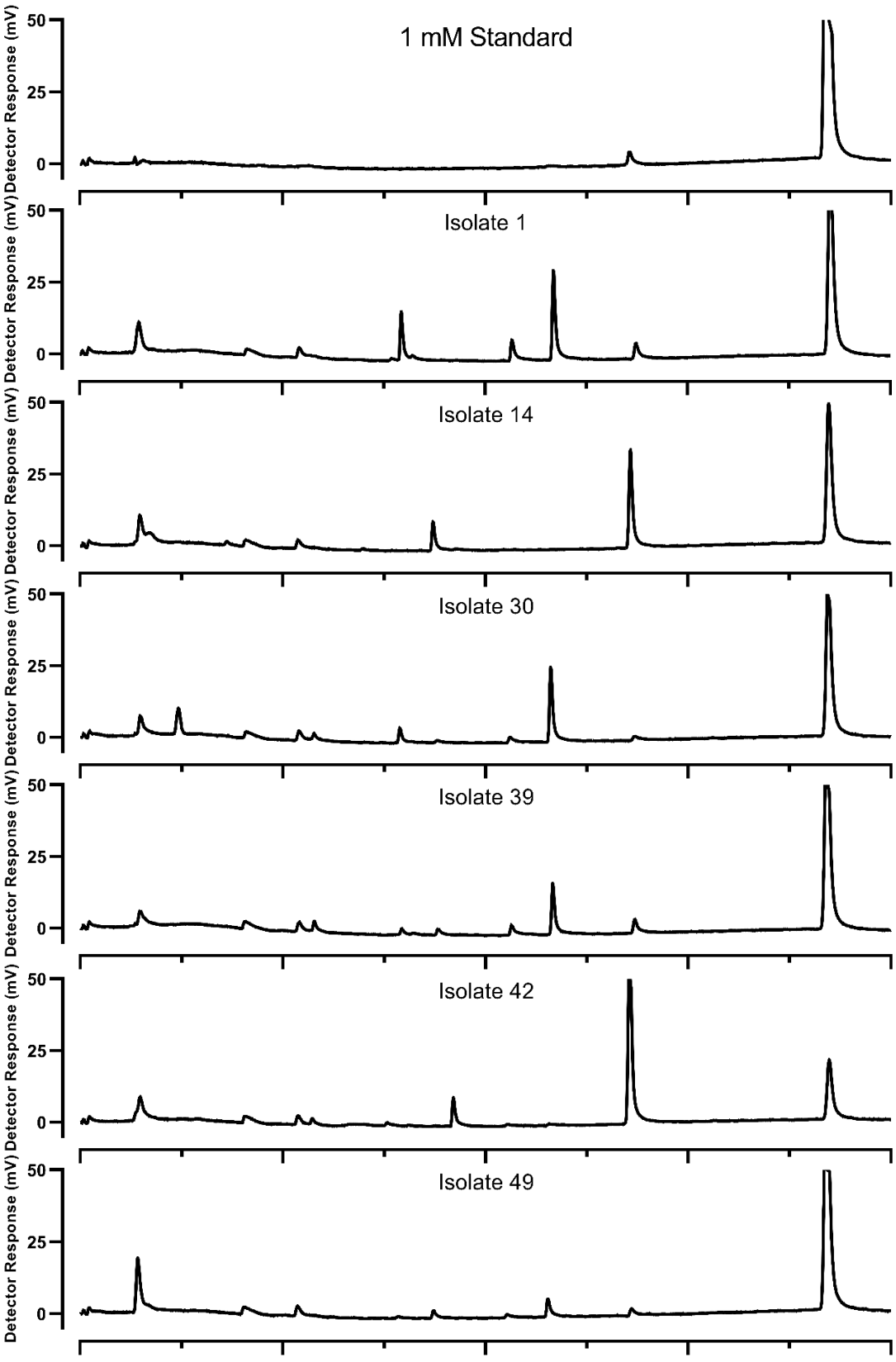


Figure 2.16 – Plate images of 87 isolates grown after one month of incubation in the presence of phytate.

Of the 129 isolates that were streaked onto the plates, only 66 isolates grew, and were tested for phytase activity. Only 17 of the 64 isolates showed signs of phytase activity, 4 of which were very minimal, constrained to small peaks of the $InsP_5$. Very pure >98% phytate was used

for each of these isolates, containing only small amounts of $InsP_5$(1/3-OH) impurities, Figure 2.17.



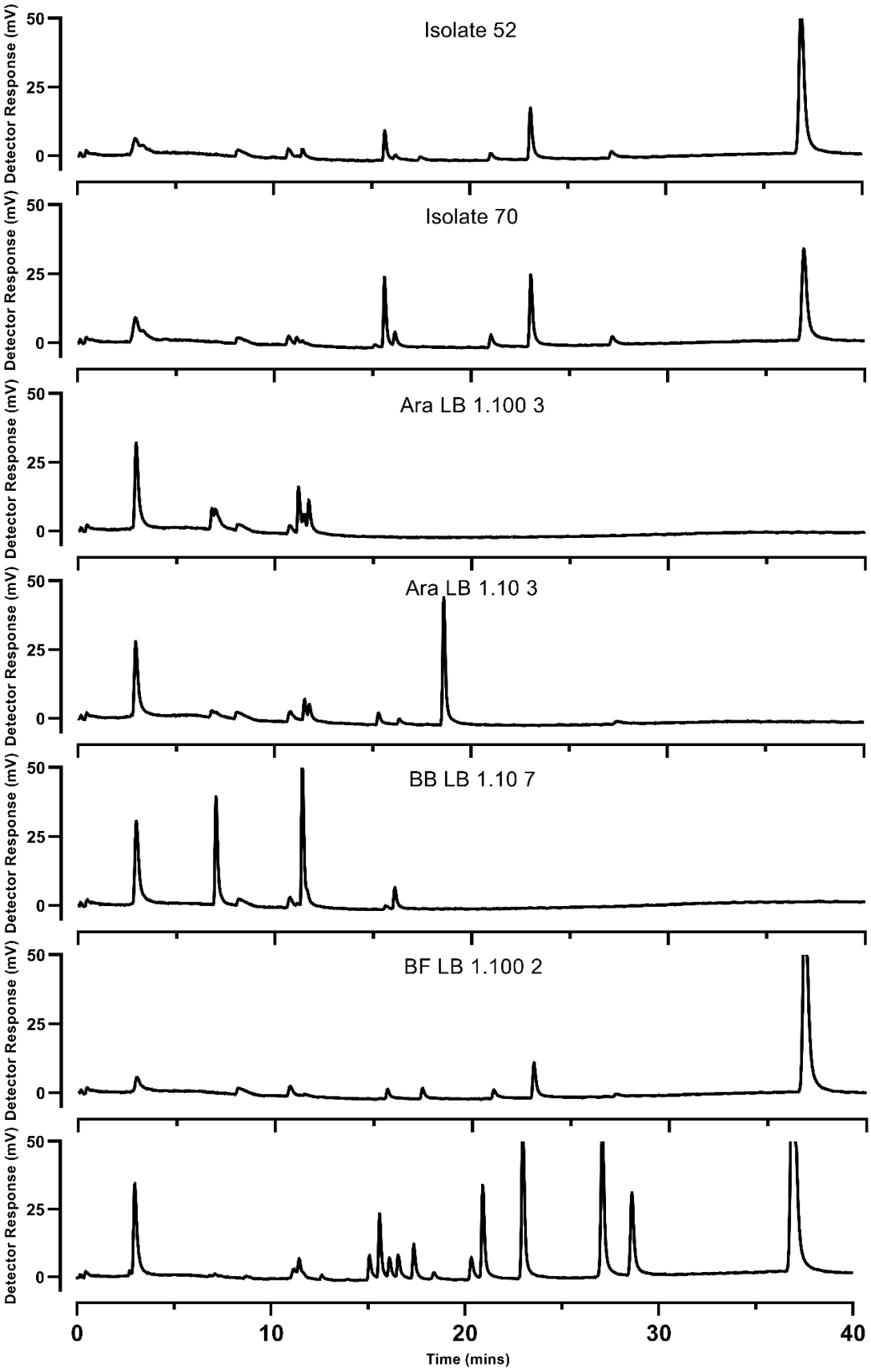


Figure 2.17 – HPLC chromatogram of the degradation profile of 1 mM phytate by thirteen soil isolates that showed phytase activity. These isolates were incubated in minimal media for three days at 30 °C. The chromatograms were compared with a phytate hydrolysate, bottom panel.

The thirteen isolates showed a variety of degradation profiles demonstrating that in the soil environment, there are many ways in which phytate can be degraded. Isolate 1 showed predominant phytase activity at the $InsP_5$(4/6-OH) position with minor $InsP_5$(5-OH) activity, three $InsP_4$ peaks were produced, a larger D/L-Ins(1234)P_4 and two very small peaks of Ins(1,2,4,6)P_4 and D/L-Ins(1245)P_4. Isolate 14 on the other hand showed (1/3-OH) activity with a single D/L-Ins(1256)P_4 peak. Isolate 30 was (4/6-OH) active with a small amount of (5-OH) produced, two $InsP_4$ peaks were present, mostly D/L-Ins(1234)P_4 with a smaller D/L-Ins(1256)P_4 peak with a single, small $InsP_3$ peak. Isolate 39 produced similar peaks to Isolate 30, however the abundances of the two $InsP_4$ peaks were the same, it also produced a small amount of $InsP_3$. Isolate 42 was a (1/3-OH) phytase, however unlike Isolate 14 it instead produced a single Ins(2,4,5,6)P_4 peak, as well as trace amounts of $InsP_3$. Isolate 49 was unusual in that it low phytase activity but produced a larger $InsP_1/P_i$ than the previously isolates. It showed (4/6-OH) activity with a slight amount of (5-OH) activity, it produced two $InsP_4$ peaks, D/L-Ins(1256)P_4 and D/L-Ins(1234)P_4, the larger of which was D/L-Ins(1256)P_4. Isolate 52 was another (4/6-OH) phytase, producing three $InsP_4$ intermediates, an D/L-Ins(1234)P_4 peak and two smaller D/L-Ins(1245)P_4 and D/L-Ins(1256)P_4 peaks, and trace $InsP_3$. Isolate 70 produced a similar degradation pattern to isolate 1 with the exception that it produced a second smaller $InsP_3$ peak. Isolate Ara LB 1.100 3 had degraded the phytate down to three $InsP_3$ and two $InsP_2$ peaks, the initial specificity was undetermined. For isolate Ara LB 1.10 3, there was only a minute amount of $InsP_5$(4/6-OH), there were three $InsP_4$ peaks, a very large Ins(2,4,5,6)P_4, and two significantly smaller Ins(1,2,4,6)P_4 and Ins(1245)P_4, additionally two smaller $InsP_3$ and one $InsP_2$ peaks were visible. Isolate BB LB 1.10 7 had also degraded phytate past the $InsP_5$s, instead two $InsP_4$ peaks were present, a larger Ins(1245)P_4 and smaller D/L-Ins(1234)P_4, with considerably larger, single $InsP_3$ and $InsP_2$ peaks. Finally, isolate BF LB 1.100 2 showed (4/6-OH) activity with trace amounts of (5-OH), it contained two $InsP_4$ peaks with similar abundances, D/L-Ins(1234)P_4 and D/L-Ins(1256)P_4 of these phytase isolates, six came from the Highfield arable soil, six came from the Highfield bare fallow soil and five came from Broadbalk.

Through the long-term isolation of different types of soil bacteria, a diverse set of phytases all with different degradation profiles can be isolated and examined, and this is just scratching the surface of what the soil community is capable of, with such a significant amount of bacteria unamenable to isolation, this will be discussed in Chapter 4, and diverse environments to examine (Nunes da Rocha *et al*., 2009).

The editorial piece by Kumar *et al* 2016 (Kumar *et al.*, 2016) provides an excellent basis for culture dependent work that has been performed by various research groups in soil, water, plants and animal environments and allows for suitable comparison to the work discussed in this Chapter. In the reports by Jorquera *et al.,* and Kumar *et al.,* (Jorquera *et al*., 2008; Kumar *et al*., 2013) phytases were isolated from Himalayan and volcanic soils and in both cases "phytase-specific" media, PSM, was used to identify phytase isolates, the drawbacks of which have already been discussed. Both papers used the appearance of clearing zones as an indicator of phytase activity and only the latter used the counterstaining technique to re-solubilise phytate on the plate. Additionally, the method used to measure phytase activity in the latter paper was based on molybdovanadate as the colouring reagent by Engelen *et al* (Engelen *et al*., 1994) using sodium phytate from rice, which has been demonstrated to contain $InsP_5$ and lower inositol phosphate impurities, Figure 2.1. While this method is suitable for quick analysis of multiple isolates, it is by no means a definitive identification of phytase activity (Madsen *et al*., 2019; Qvirist *et al*., 2015), this requires techniques such as HPLC to directly examine the degradation of phytate neither of which have been performed in the two papers. Many other papers describe an array of culture-dependent techniques isolating and investigating hundreds of bacterial isolates for investigation and finding many of them to be phytase active (Aziz *et al*., 2015; Hill *et al*., 2007; Unno *et al*., 2005). However, for these reports there is no further interrogation of gene identity or examination of the degradation profile, nor is there any comparison between the isolates, whereas in Figure 2.17, HPLC has allowed us to probably visualise the diversity of phytase degradation that can be obtained from a culture-dependent soil isolation experiment.

There is another side of examining phytase diversity that does not rely on culture-dependent techniques, and this involves the use of in silico methods. For example, in the study by Lim *et al* (Lim *et al*., 2007a), representative genes from the three classes of phytases, HAPhy, BPPhy, PTPhy, with known phytase activity were used as probes to BLAST microbial genome databases from NCBI and the Moore Marine Microbial Genome Sequencing Project and positive hits were retrieved and used to construct a phylogenetic tree, which were then used to BLAST the environmental sequence databases from NCBI and CAMERA. From this study, multiple bacterial hits for each of the classes were identified in a variety of different environments. Interestingly, the report identified the lack of HAPhy or PTPhy sequences in the genomes of aquatic bacteria. It also identified the closely associated relationship of the BPPhy with the TonB-dependent receptor-like gene in operons, suggesting the two genes may be functionally linked.

In a more recent report from Neal *et al* (Neal *et al*., 2017), environmental metagenomics was used to investigate the abundance of the HAPhy, BPPhy, PTPhy in addition to phosphatase genes in Highfield bare fallow, arable and grass soils. The HAPhy and PTPhy gene were more abundant in the grass environment whereas the BPPhy was more abundant in the bare fallow environment.

Despite identifying a plethora of phytase sequences, these papers lack the proper investigation into these diverse sequences, how active are they? What is their degradation profile? What are they inhibited by? How stable are they? Culture-dependent experiments that would provoke a well-rounded analysis. Therefore, a combination of both the culture-dependent and culture-independent techniques would provide the best analysis for phytase diversity in the environment.

2.3.3. Identifying Inositol Phosphatase Activity

Following from the phytase activity and profile analysis above, it was noticed on the HPLC chromatograms of the isolates that phytase degradation was not the only feature that was observed. As first referenced by Cosgrove *et al* (Cosgrove *et al*., 1970), isolates may have activity towards the lower inositol phosphates and not phytate. Twenty-one of the isolates displayed either inositol pentakisphosphate activity or unusual activity, Figure 2.18.

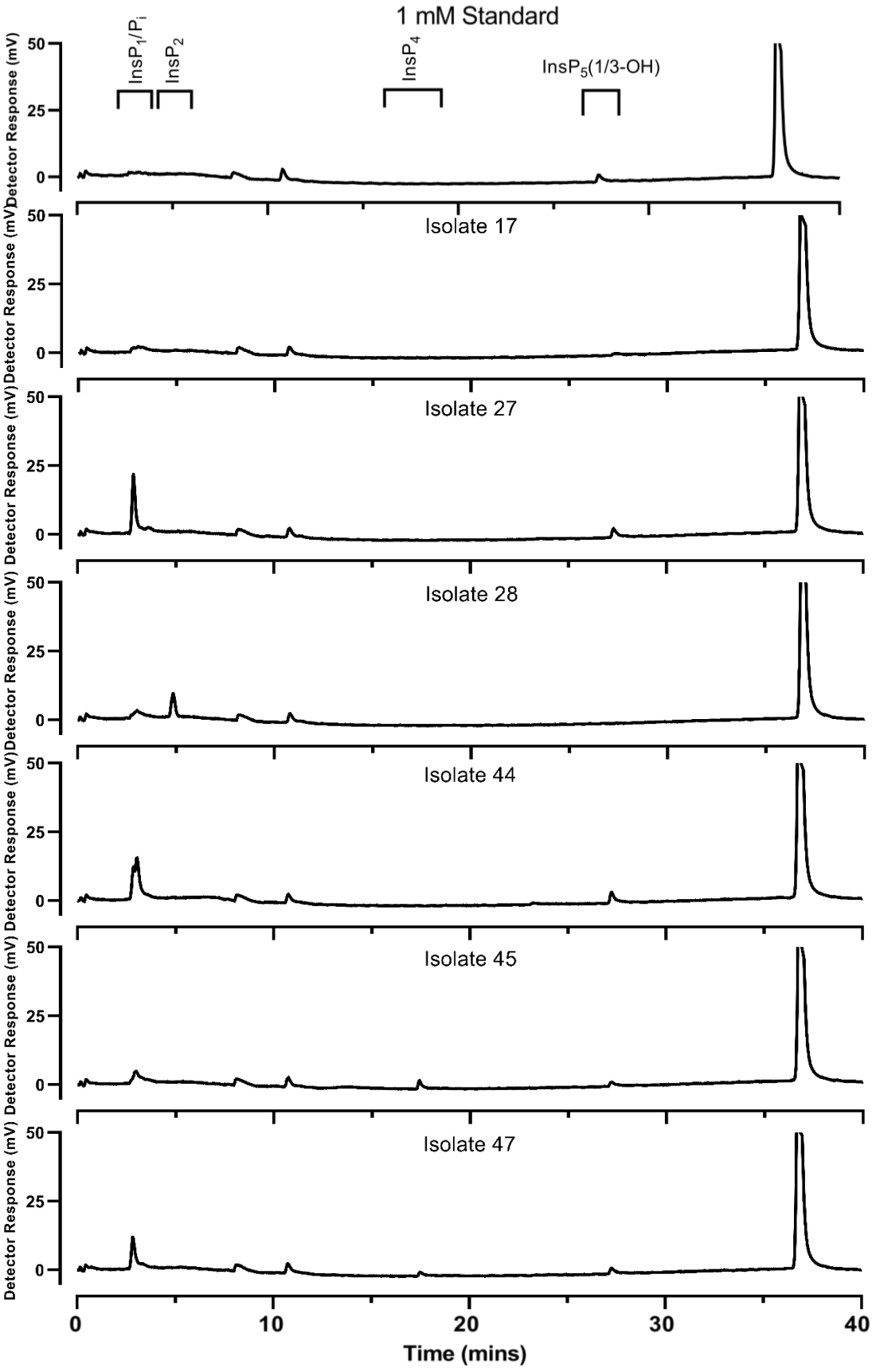


Figure 2.18 – HPLC traces of six of the 21 $InsP_5$/unusual degraders that were isolated from Rothamsted soils.

These isolates were compared with a 1 mM Phytate standard with areas of change highlighted.

Isolate 17, whose activity profile was shared by Isolates 18, 20, 23, 32, 35, 37 and 38 (not shown), showed degradation of the small amount of D/L-Ins(1,2,4,5,6)P_5 impurities found in the 1 mM phytate sample, the $InsP_6$, however, was not touched and there was no real change in the $InsP_1/P_i$ peak. Isolate 27 and similarly Isolate 44 did not show any $InsP_6$ nor $InsP_5$ degradation, however, a large $InsP_1/P_i$ peak appeared in the chromatogram. Unusually, this peak in Isolate 44 which is commonly a sharp, single peak was broader and split into two at the tip of the peak. Isolate 28 was interesting in that there was $InsP_5$ degradation and the subsequent appearance of an $InsP_2$ peak with no presence of $InsP_4$ or $InsP_3$ products. There was also no significant increase in the $InsP_1/P_i$ peak either. Isolate 44 showed $InsP_5$ degradation and an increase in D/L-Ins(1256)P_4, with a small increase in the $InsP_1/P_i$ peak. Similarly Isolate 47 showed the same activity, but a much larger increase in the $InsP_1/P_i$ peak.

Similar to the results shown by Cosgrove *et al* (Cosgrove *et al*., 1970), the isolation of lower inositol phosphate degraders has occurred as well as other isolates showing interesting phosphate build-up. In the case of all of these isolates, without the use of HPLC it would be difficult to determine if the activity and accumulation of phosphate was due to low activity phytases or lower inositol phosphate degraders.

2.3.4. Isolate sequencing

All these isolates that were involved in phytate or $InsP_5$ degradation were subjected to 16S rRNA sequencing, Table 2.6. Unfortunately, many of the 16S rRNA genes of these isolates did not amplify, nor did the bacteria grow when they were restreaked onto fresh plates. Therefore only 6 phytate degraders, 6 $InsP_5$ degraders, and two isolates displaying unusual inorganic phosphate activity were sequenced.

Table 2.6 – 16S rRNA sequencing of the Rothamsted Isolates using bacterial primers. The highest scoring genus is displayed below.

Isolate Number	Bacteria	Percentage Identity (%)	Type	Accession Number
11	*Curtobacterium* sp.	100	Phytase	CP041259
17	*Arthrobacter/ Pseudoarthrobacter* sp.	99.71	$InsP_5$	KY753220
23	*Arthrobacter* sp.	99.01	$InsP_5$	OM060450
24	*Serratia* sp.	100	Unusual P_i	KY90700
27	*Serratia* sp.	100	Unusual P_i	KY90700
37	*Pseudoarthrobacter* sp.	100	$InsP_5$	KU647202
38	*Arthrobacter* sp.	99.06	$InsP_5$	KU350608
41	*Serratia* sp.	99.90	Phytase	MN511730
42	*Bacillus* sp.	100	Phytase	LC667823
70	*Serratia* sp.	99.90	Phytase	MN422008
BB LB 1.10 7	*Acinetobacter* sp.	100	Phytase	EU118781

Blast analysis was performed on each of the strains as per Section 2.26 to investigate the type of phytase genes each of corresponding genus contains, Table 2.7.

Table 2.7 – The number of different phytase families identified in genera related to the isolates obtained from Rothamsted soils. These were determined using NCBI blast that was limited to search for each individual family of canonical phytases. The max target sequences were set to 1000, with an E-value cutoff at $1e^{-10}$.

Bacteria	HAPhy	MINPP	BPPhy	PTPhy	PAPhy
***Curtobacterium* sp.**	0	58	0	0	101
***Serratia* sp.**	303	0	1	1	2
***Bacillus* sp.**	3	0	371	1	64
***Acinetobacter* sp.**	9	445	80	0	1

Curtobacterium sp. provided the most interesting results, showing 58 MINPP hits and 101 PAPhy hits. Both phytase types that up until recently were not considered to be present in bacteria in the environment (Ghorbani Nasrabadi *et al*., 2018; Rix *et al*., 2020), with MINPPs being predominantly isolated from enteric sources (Cho *et al*., 2006; Stentz *et al*., 2014) and PAPhys being from plants (Dionisio *et al*., 2011; Lung *et al*., 2008). The *Curtobacterium* isolate produced low activity in the 4/6 position which has been identified in both HAPhys and

PAPhys (Faba Rodriguez, 2018; Rix *et al*., 2020). The HAPhys produced the most hits in *Serratia* sp., with only 1-2 hits for the other phytase types. The activity of Isolate 41 was in the 4/6-position which is in agreement with HAPhys from *Serratia* sp. whose genomes have been mined to clone and characterise novel phytases, and have also been identified in enteric and environmental samples (Salaet *et al*., 2021; Shedova *et al*., 2008; Zhang *et al*., 2011a). Isolate 42 had the characteristic activity of a BPPhy with activity in the 1/3-position (Kim *et al*., 2010; Shin *et al*., 2001). The BPPhy is extremely common in *Bacillus* species and has been regularly identified in the literature, particularly pertaining to aquatic environments (Cheng and Lim, 2006; Huang *et al*., 2009b). There was also 64 hits from the PAPhy which haven't been isolated from *Bacillus* before. Finally, phytase activity in the *Acinetobacter* genus has already been spoken about in Section 2.2.7. The activity was similar to that of *Acinetobacter* sp. AC1-2 and therefore the phytase from this isolate is most likely that of a MINPP.

Despite the unfortunate loss of many of the isolates, this experiment has categorically identified ways in which a diverse set of phytase degraders can be isolated from the soil environment using a long-term incubation period.

2.4. Examining the adsorption of phytate onto soil

As identified in the Broadbalk soil in Section 2.3.1, when working with soil-liquid cultures as a means of examining phytate degradation, some of the phytate may be adsorbed onto the soil particles, removing the accessibility of phytate, depending on the contents of the soil. Soils that are high in iron and aluminium oxides such as goethite and boehmite (Gerke, 2015; Johnson *et al*., 2012) are able to complex the phytate molecule, and in alkaline soils the same occurs with clays and calcite (Wan *et al*., 2016), of which Broadbalk can be up to 39% clay (Watts *et al*., 2006).

The binding of $InsP_6$ to Highfield arable soil with and without calcium was examined in Figure 2.19. Calcium, Ca^{2+}, was used as it is a very abundant and important metal ion in soil, the cation capable of bridging soil particles into aggregates, having a stabilising effect on the structural stability of soil (Wuddivira and Camps-Roach, 2007). When in the form of calcite ($CaCO_3$) in calcareous soils, phosphate ions can be sorped onto the surface of the calcite, occurring via the replacement of CO_3^{2-}. As such, calcium may have a dual role in this experiment, it first acts to stabilise the soil and secondly binds and precipitates phytate from the liquid phase.

The results were compared to a 1 mM $InsP_6$ culture in water, Figure 2.19a. When a similar concentration of $InsP_6$ is added to the Highfield ley arable soil-liquid culture, Figure 2.19b, almost 62% of the phytate is removed from the liquid phase and is adsorbed onto soil particles, and this increases further to 93% when 1 mM calcium is added, Figure 2.19c. This may be due to sorption of the phytate onto soil calcium complexes, or the formation of insoluble precipitates.

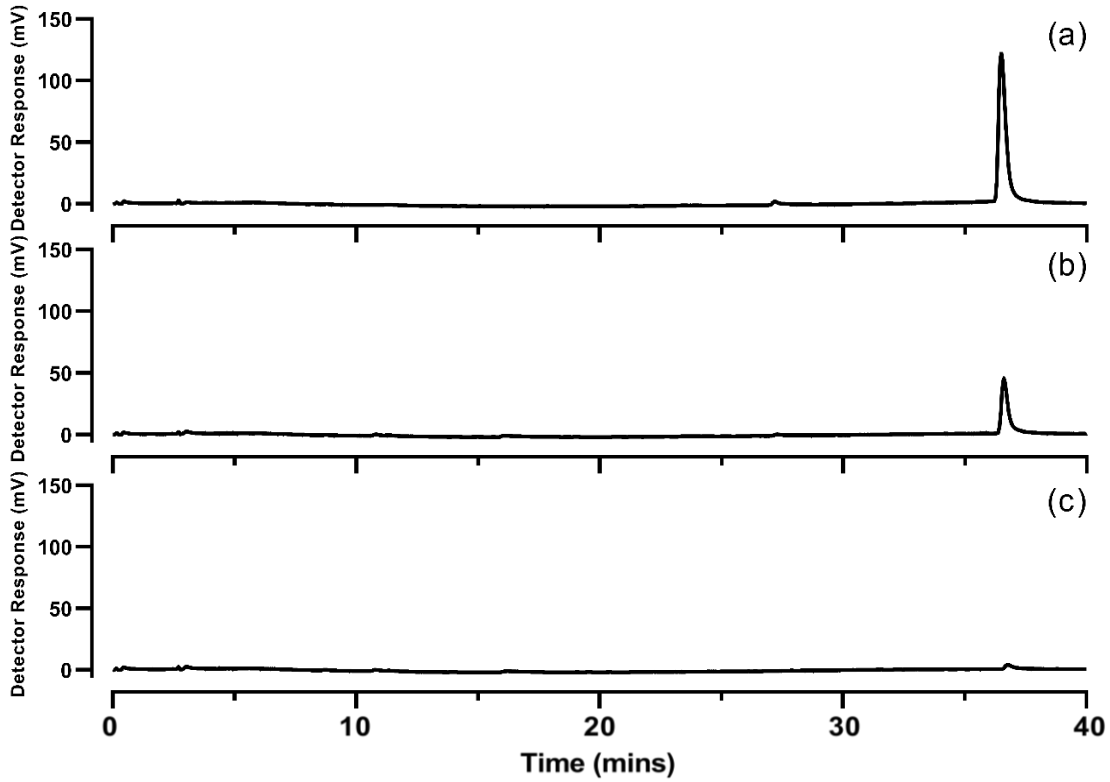


Figure 2.19 - HPLC chromatogram of the degradation profile of 1 mM phytate by Highfield ley arable soil. (a) Minimal media control with no added soil, (b) Soil and Minimal Media (c) Soil and Minimal media spiked with 1 mM calcium ions. Samples were incubated for 30 minutes before sampling.

The extent of phytate adsorption in each of the soil types, Broadbalk, Highfield ley arable and Highfield bare fallow, was tested by the addition of 1 mL of 10 mM $InsP_6$ solution to each 0.5 g soil sample. The concentration of $InsP_6$ that remained in solution was calculated by peak integration using the HPLC software (Jasco ChromNav v.1) and comparison with the peak area of a 1 mM $InsP_6$ solution in dH_2O. The HPLC traces and peak integrations are displayed in Table 2.8. The adsorption calculations for this experiment can be seen in Section 2.1.16

Table 2.8 – Peak integrations of the $InsP_6$ peaks to determine the adsorption of 1 mL of 10 mM $InsP_6$ after 30-minute incubations at 30 °C.

Environment	IP_6 (nmol)	Concentration of IP_6 in solution (mM)	% Adsorped
Control	20	1	0
Highfield Ley Arable	66.4	3.3	66.82
Highfield Bare Fallow	57.5	2.9	71.23
Broadbak	1.2	0.06	99.40

The adsorption capacity of each of the soils was substantial, with 66.82, 71.23 and 99.40% of the added phytate adsorbed onto it for the Highfield arable, bare fallow and Broadbalk soils respectively. Following adsorption, phytate was extracted from the soil suspension with NaOH:EDTA, a method which is commonly used to extract $InsP_6$ for ^{31}P NMR studies (Doolette *et al*., 2010; Giles *et al*., 2011), Table 2.9.

Table 2.9– The recovery of $InsP_6$ that had been adsorbed onto soil using NaOH:EDTA extractions of the Highfield Ley Arable, Highfield Barefallow and Broadbalk soil saples

Environment	$InsP_6$ Recovered (mM)	Recovered Percentage (%)
Highfield Ley Arable	**0.66**	**10.31**
Highfield Bare Fallow	**0.30**	**4.86**
Broadbalk	**1.42**	**14.98**

Following this, the stocks of each soil types were left for a month, during which they became desiccated and finely grained. As such, the experiment was repeated to see how much of an effect this has on the adsorption capacity of phytate, Table 2.10.

Table 2.10 – The adsorption capacity of three desiccated and finely grained soil, Highfield Ley Arable, Highfield Bare Fallow and Broadbalk, after one month incubation.

Soil Type	Moles of IP6 (nmoles)	Concentration of IP6 in solution (mM)	% Adsorped
Control	20	1	0
Highfield Ley Arable	112.1	5.6	44
Highfield Bare Fallow	109.6	5.5	45
Broadbalk	50.5	2.5	75

It may be that the desiccation of the soil has had a negative impact on the ability of the soil to adsorb phytate, with only 44, 45 and 75% of the phytate adsorbed by the Highfield arable, Bare fallow and Broadbalk soils respectively, in comparison to 66.8, 71.2 and 99.4% previously. Indeed, it may be that due to changes in the structure of soil and soil minerals which had initially stabilised phytate may not have the capacity as it had previously (Doolette *et al*., 2010). While these results provide an interesting view into the ability of phytate to be adsorbed onto three different soil types, and the consequence of the time-dependent drying and desiccation of soils. There are limitations to this experiment. The moisture content of the soil was not analysed for the fresh and one month old samples, nor were there replication to allow for statistical analysis.

2.5. Analysing the phytate degradation in aerobic, anaerobic and water sediments from Blakeney Iron and Sulphur cores.

All of the soils examined for phytase degradation and isolation in this study have currently been from inland soils, however phytases and phytase activity have not just been identified in soil environments but also from aquatic environment and sediments (Cheng and Lim, 2006; P. Yu *et al*., 2015). Therefore, this section examines phytase degradation from a local salt marsh. For these experiments, commercial phytate (Sigma P8810) was used due to a lack of pure phytate.

Two cores were taken from a salt marsh located in Norfolk near the town of Blakeney, these cores were from either iron-rich or sulfide rich sediments (Hutchings *et al*., 2019). Aerobic samples were taken within a 1 cm depth and anaerobic samples at 24 cm, water samples were taken from the top of the core. Salt marshes provide an interesting environment to study phytate degradation (Antler *et al*., 2019), because these wetlands are highly productive, regularly covered by tidal movements, and due to high nutrient loading from terrestrial run-off, productivity exceeds that of many other environments (Giblin and Wieder, 1992). As such, six conditions/environments were analysed for phytate degradation, the first condition was the water sample of the core of sulfidic sediment, Figure 2.20.

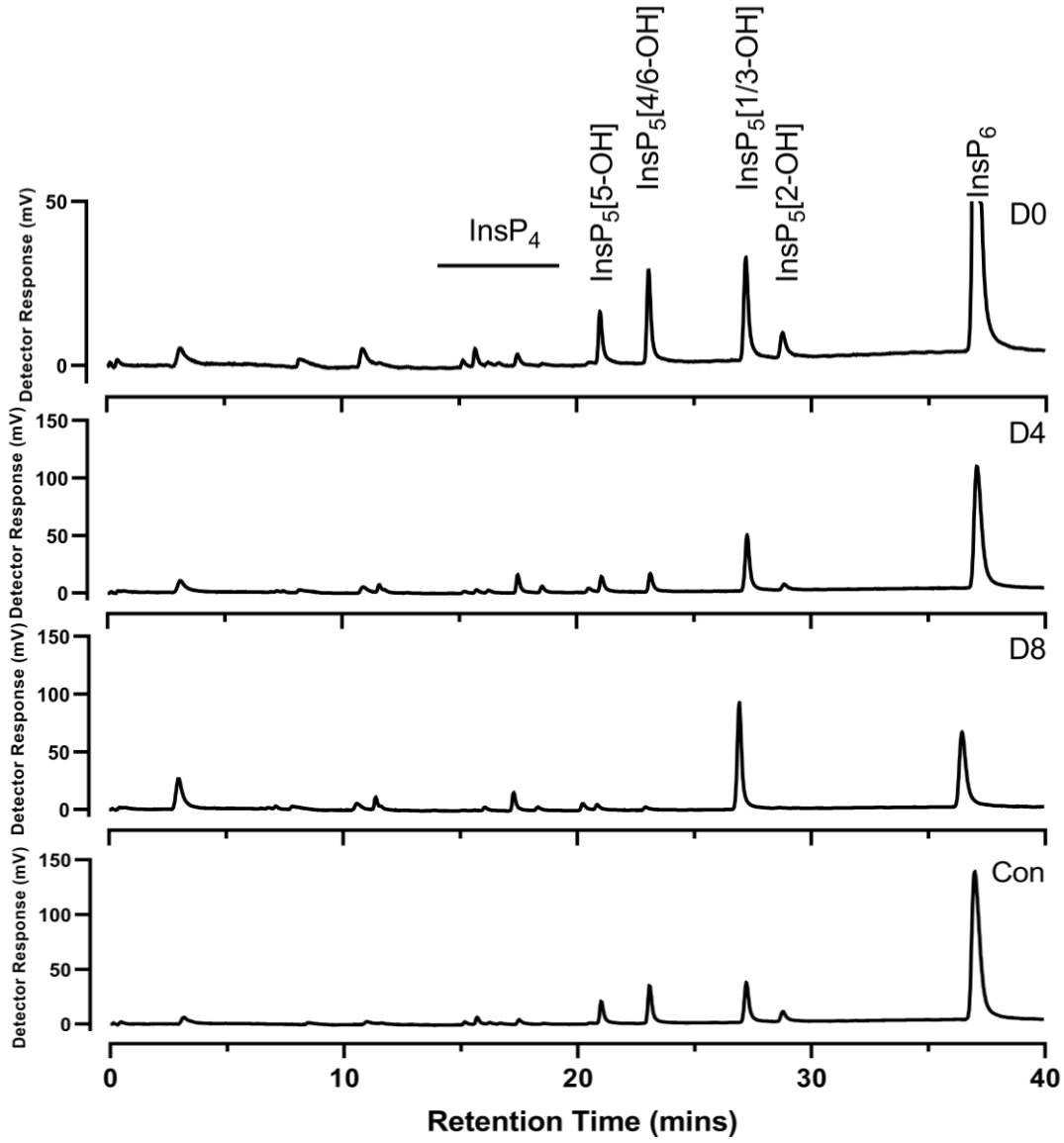


Figure 2.20 – HPLC chromatogram of the degradation profile of 1 mM phytate by the sulphur core water sample in minimal media. Samples were taken each day to examine the progression of phytate degradation.

In this sample, the primary degradation pathway occurs in the 1/3-OH position, at the same time, the other $InsP_5$ impurities are removed from solution after eight days. For the $InsP_4$s,

there is a slight increase in D/L-Ins(1456)P_4, Ins(2,4,5,6), D/L-Ins(1245)P_4 and a larger increase in D/L-Ins(1256)P_4, there is also a small increase in both $InsP_3$ and $InsP_2$.

The next condition was the aerobic sample of the sulphur core, Figure 2.21

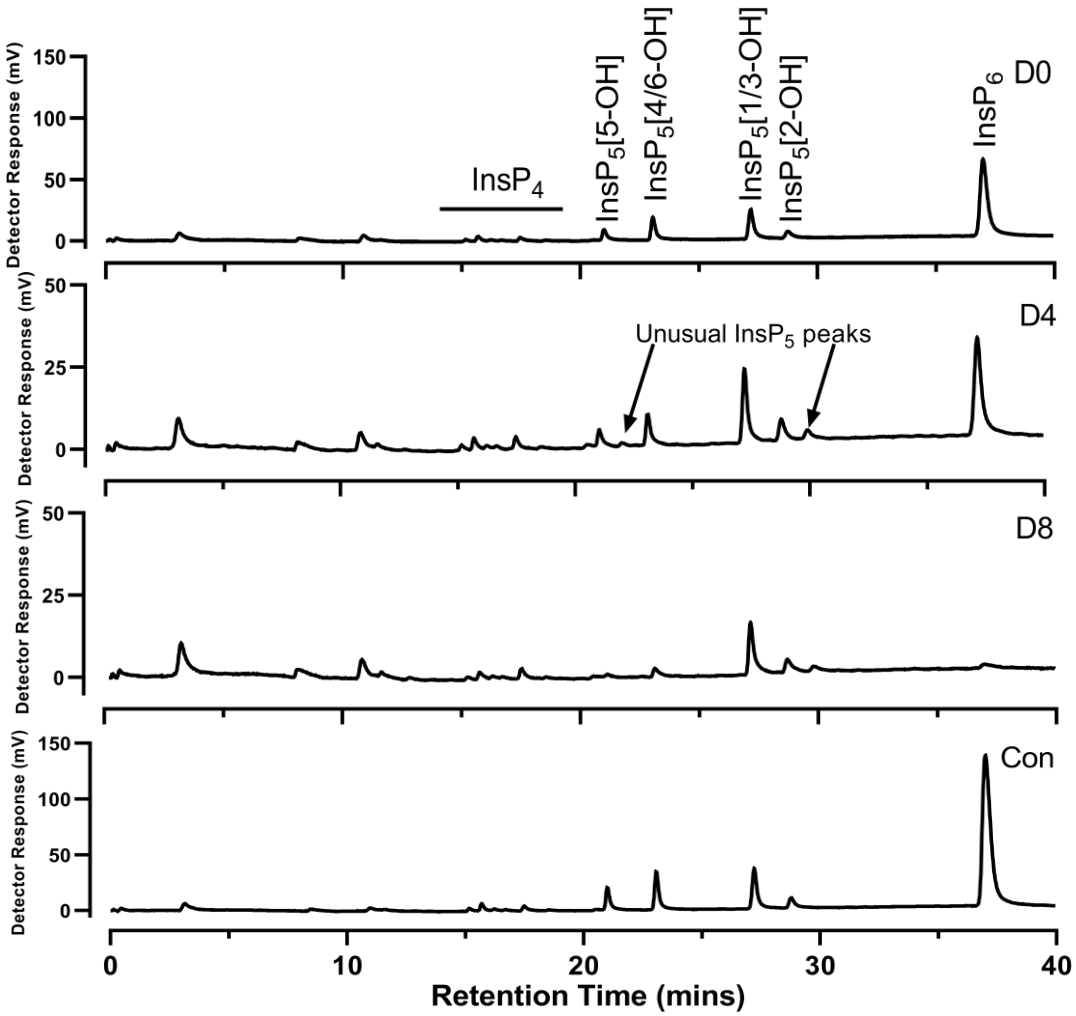


Figure 2.21 – HPLC chromatogram of the degradation profile of 1 mM phytate by the sulphur core aerobic sample in minimal media. Samples were taken each day to examine the progression of phytate degradation.

The first difference between the aerobic sample and the water sample is that almost immediately as much as 60% of the phytate has been adsorbed into the soil. Also interestingly, there are two new peaks that are not observed for $InsP_6$ hydrolysis that appear just before the $InsP_5$(2-OH) peak and also $InsP_5$(5-OH). The report from Whitfield *et al*, (Whitfield *et al*., 2018) has shown that a *scyllo*-$InsP_5$ peak runs just after $InsP_5$(2-OH) which may be what is present in the chromatogram. The second peak running in-between $InsP_5$(5-OH) and $InsP_5$(4/6-OH) may be its degradation product, a *scyllo*-$InsP_4$, or it may also be a L-*chiro* or D-*chiro*-$InsP_5$ which elutes between these two peaks.

Additionally, there is also a slight change in the abundances of the $InsP_4$s, for the $InsP_5$'s, both $InsP_5$(4/6-OH) and $InsP_5$(5-OH) impurities are degraded, however despite the $InsP_6$ peak decreasing, there does not appear to be a concurrent increase in any of the $InsP_5$ peaks.

The next condition was the anaerobic sample of the sulphur core, Figure 2.22.

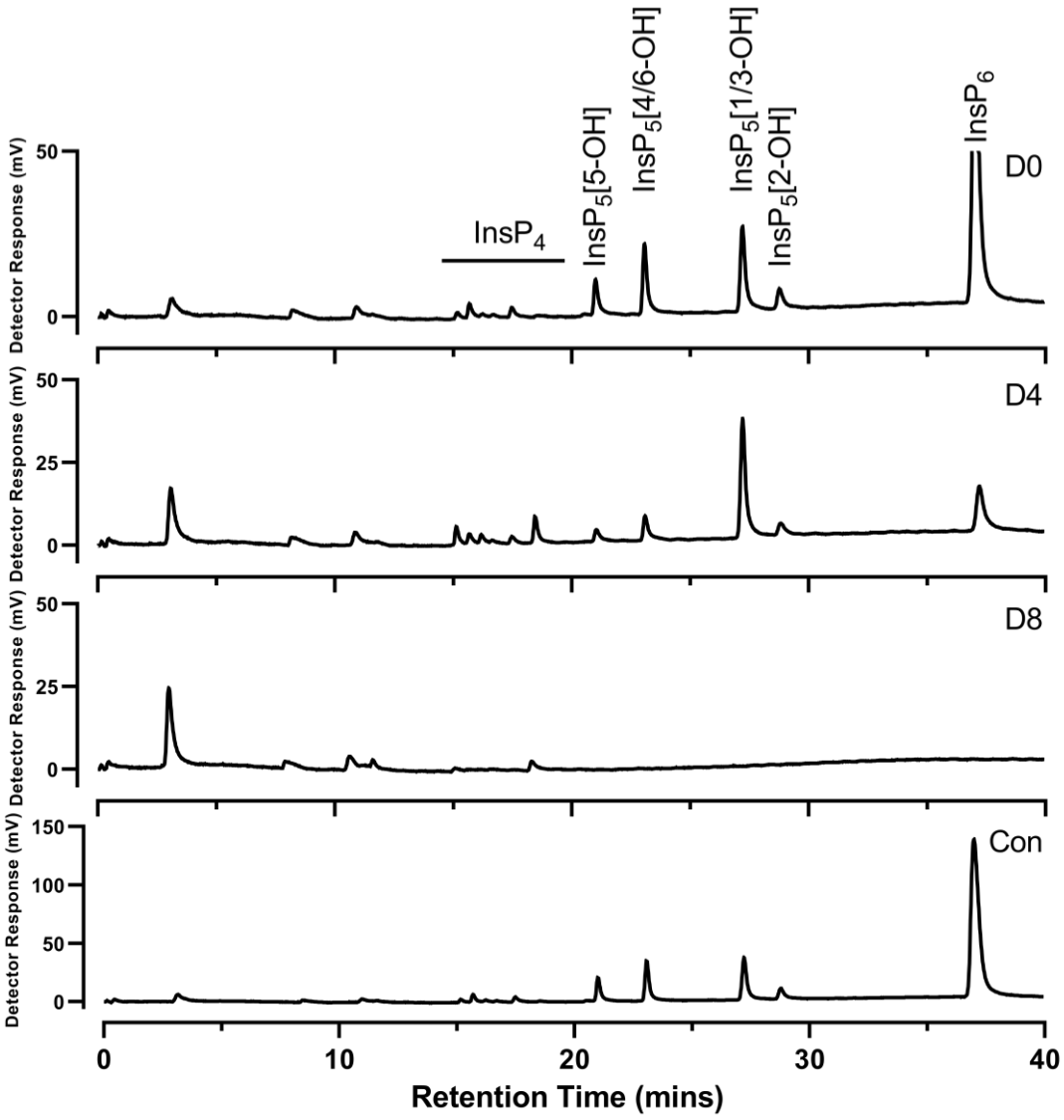


Figure 2.22 - HPLC chromatogram of the degradation profile of 1 mM phytate by the sulphur core anaerobic sample in minimal media. Samples were taken each day to examine the progression of phytate degradation.

Once again, the addition of phytate to the anaerobic sediment has caused nearly 50% of the phytate to be adsorbed. There is a modest increase over time in $InsP_5$(1/3-OH) with no increase in the other $InsP_5$s. For the $InsP_4$s there are slight increases in D/L-Ins(1245)P_4 and Ins(1,2,4,6)P_4 and a larger increase in Ins(2,4,5,6)P_4 with no increase in $InsP_3$. At the same time, the $InsP_1$/P_i peak continues to grow in intensity from Day 0 – Day 8.

Each depth of the sulphur core sample produced a different HPLC chromatogram, highlighting an interesting diversity in inositol phosphates in water, aerobic and anaerobic conditions. A comparative figure of Day 4 of each sample is shown in Figure 2.23.

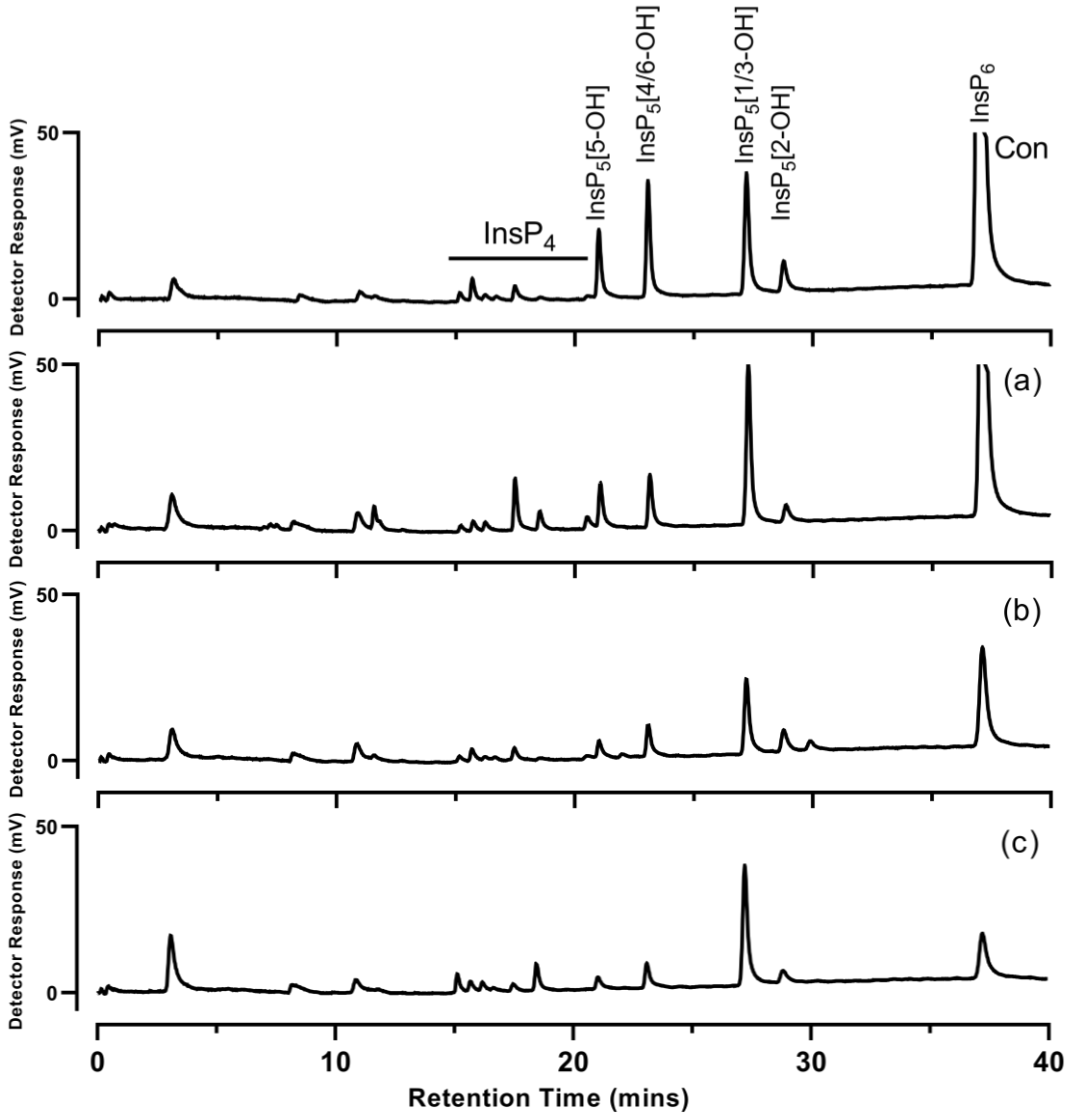


Figure 2.23 – HPLC trace of the phytate degradation profiles of the three sulphur core conditions after 4 days. Water (a), aerobic (b) and anaerobic (c) samples from the sulphur core, compared with a control sample.

In panels a-c we can see that the presence of phytate has induced a different profile in the water, aerobic and anaerobic samples.

The next three conditions were from the iron core, firstly the iron water sample was examined, Figure 2.24.

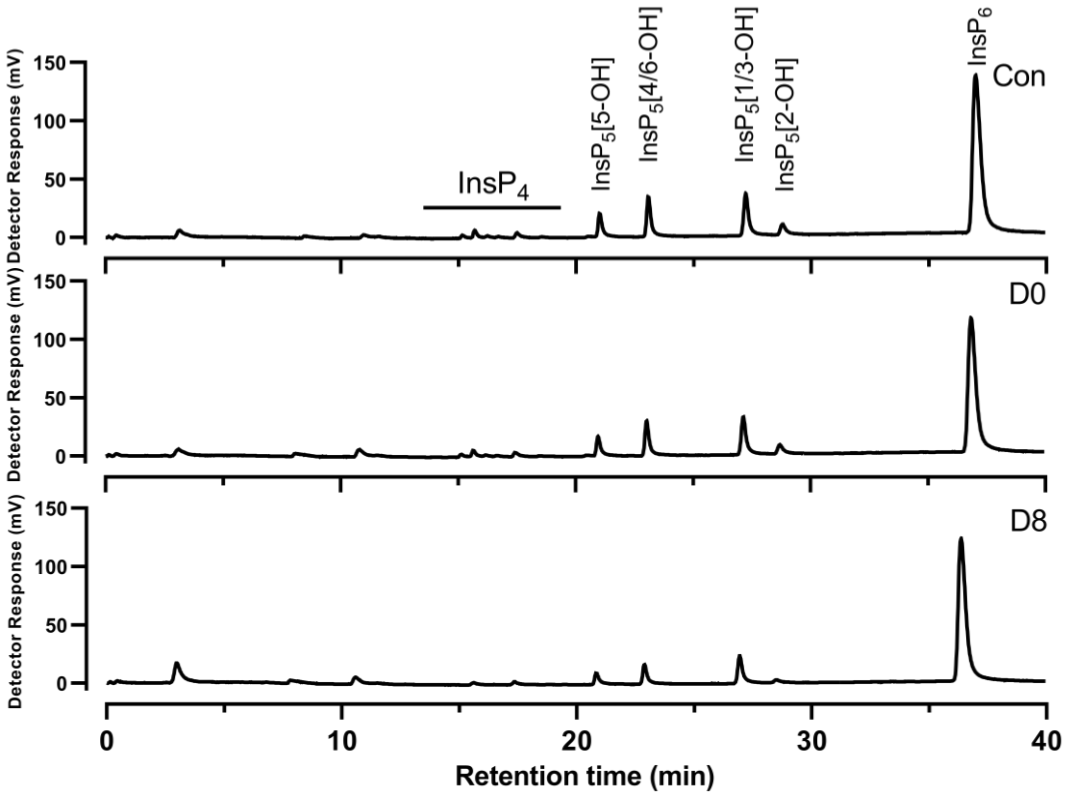


Figure 2.24 – HPLC chromatogram of the degradation profile of 1 mM phytate by the iron core water sample in minimal media. Samples were taken each day to examine the progression of phytate degradation.

In this sample there does not appear to be any substantial indication of phytase degradation. The only change is in the abundances of the $InsP_4$ impurities which decreases and a slight increase in the $InsP_1/P_i$ peak. This may be due to a time-dependent desorption of phosphate by the added $InsP_6$ and lower inositol phosphates as small amounts of sediment were present in the samples.

The next sample was the aerobic sample of the iron core, Figure 2.25.

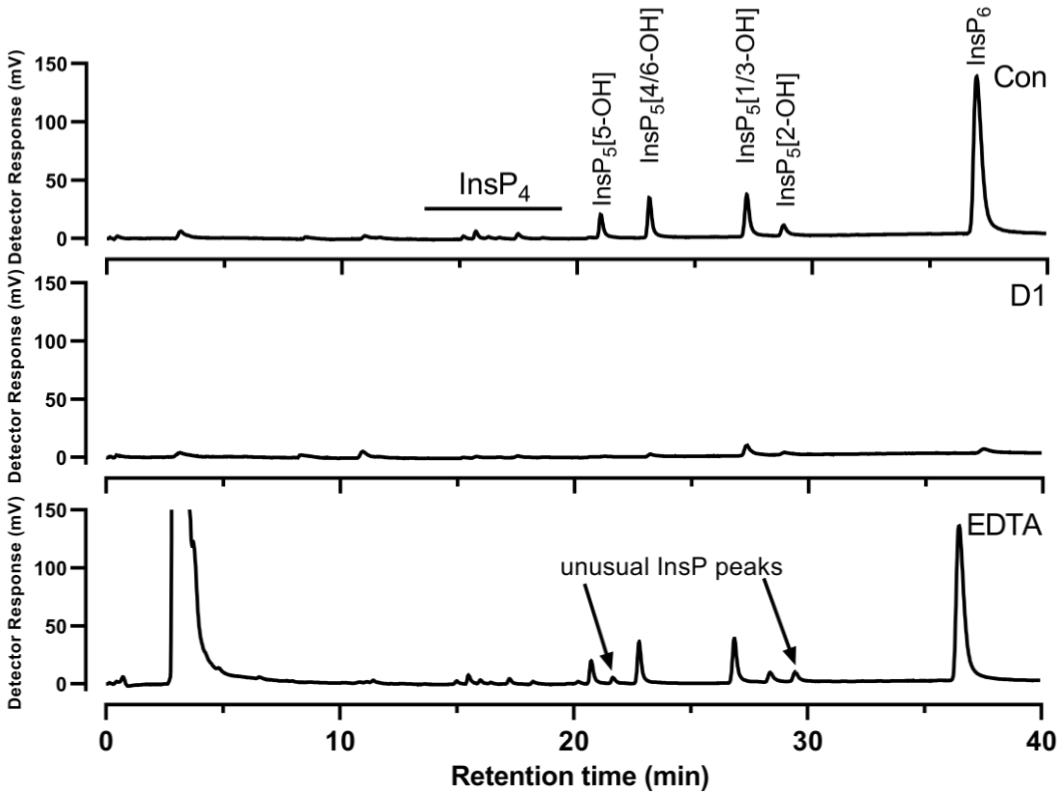


Figure 2.25 – HPLC trace of phytate degradation of the iron core aerobic sample in minimal media using 1 mM $InsP_6$. Samples were taken each day to examine the progression of phytate degradation.

Interestingly, by Day 1 all of the phytate had been adsorbed onto the sediment phase and therefore it wasn't possible to detect any phytase degradation during the time-course. After eight days, the samples were subjected to 0.25 M NaOH, 0.05 M EDTA extractions to recover any adsorbed $InsP_6$ and lower inositol phosphates. In comparison to the control sample, the EDTA extraction appeared to recover nearly all of the bound $InsP_6$, and while there may not be conclusive signs of phytate degradation, there is the appearance of the two unusual peaks that were present in the sulphur core aerobic sample.

The final sample was the anaerobic sample of the iron core, Figure 2.26. This profile was nearly identical to the iron-core aerobic sample, with the phytate immediately adsorbed onto the soil, and when the sample was extracted using 0.25 M NaOH, 0.05 M EDTA, the two unusual Inositol phosphate peaks were present once more.

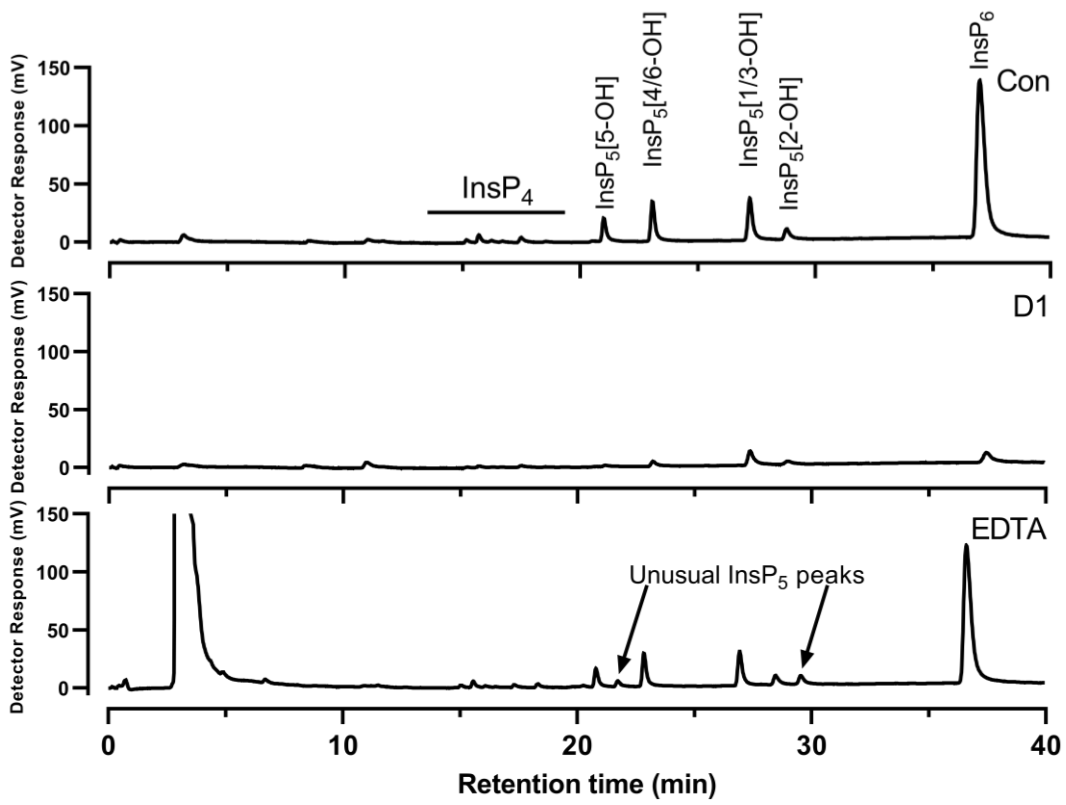


Figure 2.26 – HPLC chromatogram of the degradation profile of 1 mM phytate by the iron core anaerobic sample in minimal media. Samples were taken each day to examine the progression of phytate degradation.

In this experiment two environments have been examined for phytate degradation at different depths. Samples from the iron/sulphur cores were taken from the top water layer, aerobic and anaerobic layers and mixed with phytate. Samples were taken daily and examined for their degradation profile. For the sulphur cores, phytase degradation is most visible with the water sample, which shows a large peak of $InsP_5$(1/3-OH) which is commonly associated with beta propeller phytases which are ubiquitous in water environments (Cheng and Lim, 2006). In addition to this the $InsP_5$ impurities which are present in commercial phytate (Madsen *et al*., 2019) are also broken down from Day 0 to Day 8, there is a slight build-up of $InsP_4$s. In the sulphur core aerobic sample, there is no real increase in the $InsP_5$s as the $InsP_6$ peak decreases. It may be that in time, more and more of the phytate is becoming adsorbed out of the liquid culture, there is an increase in the $InsP_1$/P_i peak, indicating the release of inorganic phosphate through some sort of phosphatase activity, but this may be due to the breakdown of the $InsP_5$ impurities and not phytase activity. At the same time, two new peaks are present on the HPLC chromatogram which are not associated with phytate degradation and may have arose due to other processes. In the sulphur anaerobic sample, the HPLC chromatogram changes once more. There is an increase in the abundance of $InsP_5$(1/3-OH) and an increase in different $InsP_4$ products in comparison with the sulphur

water environment. Therefore, in each of these environments, phytate appears to be acted upon in different ways. The differences in the rate of phytate degradation observed in the sulphur core water and sediment is similar to the results identified by Stout *et al* which say 16S rRNA-normalised BPPhy copies were four orders of magnitude higher in water than in sediments (Stout *et al*., 2016).

Furthermore, the difference between the phytate degradation profiles of the anaerobic and aerobic samples of phytate saw low activity in the anaerobic sample and no detectable sign in the aerobic sample. Visually, while it looks like there was degradation from both the aerobic and anaerobic samples with the $InsP_6$ peak decreasing. There was no concurrent increase in any $InsP_5$ peaks for the aerobic sample, while there was for the anaerobic sample. Therefore, disappearance of the $InsP_6$ may be a case of adsorption rather than phytase activity.

A similar occurrence can be examined in the report by Suzumura and Kamatani (Suzumura and Kamatani, 1995) where in riverine sediments under simulated marine environments, $InsP_6$ in anaerobic conditions were decomposed almost completely within 40 days, whereas about 50% of $InsP_6$ remained under aerobic conditions after 60 days. It may be that the experiment did not run long enough for the aerobic bacterial community to begin phytate degradation. Returning to the report by Stout *et al* (Stout *et al*., 2016) another feature of their analysis of BPPhy genes identified that there was a greater abundance closer to the agricultural farmland where both inorganic phosphate and phytate concentrations were high, and the phytate mineralised by bacteria in the water, excess phytate may accumulate in sediments or transported downstream but enzyme activity was limited. It may be that due to the proximity of the sample pools to the ocean there is little to no phytate exposure to the microbial community and therefore there is less phytase activity in these sediments.

Sampling and analysis of the iron cores showed a different problem, it is well known that iron acts as an inhibitor to phytase activity (Santos *et al*., 2015), in the iron core water sample, there are no signs of phytate degradation despite a slight increase in the $InsP_1/P_i$ peak which may have occurred through degradation of the $InsP_4$ impurities. For both the aerobic and anaerobic samples, by the first day nearly all of the phytate and its impurities have been adsorbed onto the solid soil phase. In an investigation of Camargue sediments by De Groot and Golterman (De Groot and Golterman, 1993) they noted that while phytate was found to precipitate with all polyvalent cations tested, it was Fe(OOH) that phytate most strongly adsorbed onto, which may explain its immediate absorption and stability in the Blakeney

Iron-core sediments. This was examined by extracting the sample using EDTA-NaOH (Doolette *et al*., 2010), no phytate degradation occurred in these samples, however, the appearance of the same two peaks previously shown for the sulphur core aerobic sample suggest something is occurring in this soil.

To conclude this experiment, not just changing the environments, but also the depths at which samples are taken could change the activity and degradation profiles of phytases. There may also be soil environments where phytase activity is low due to the presence of inhibitors or strong phytate adsorbers.

2.6. Conclusion

The first part of this chapter provided a critical review into many of the pitfalls with regards to isolating phytases from the environment, many issues of which are not considered in phytase isolation studies in the literature, as shown in Table 2.1. As such, this review concluded that unless phytate degradation is directly measured, through techniques such as HPLC, then low phytase activities may actually derive from lower inositol phosphates degraders.

In the next part of the chapter, the frequently use culture medium "Phytase-specific media" was examined and developed to ensure that all of the potential pitfalls have been taken into account. Control strains of bacteria were plated onto the medium and examined as large halos formed around of each of the colonies. These halos were then extracted using acid to resolubilise the phytate and examined for degradation using HPLC. This confirmed it was possible to examine phytate degradation this way, however, the dominant peak of phytate still remained and one of the strains showed no sign of phytate degradation despite containing a phytase gene in its genome. Therefore this method is a poor indicator of phytase activity, even when using known phytase degraders, this is before the issues regarding acid secretion which also majorly affects the agar plates (Bae *et al*., 1999), a feature which is very common for bacteria (Archana *et al*., 2012). A study of 36 bacterial isolates identified eight different kinds of organic acids secreted as a means of solubilising phosphate (Chen *et al*., 2006), therefore in the next section a new technique was developed which avoided the pitfalls and did not rely of "PSM" media.

This technique used a mixed soil culture spiked with phytate with samples taken each day for HPLC analysis to examine for phytate degradation. This led to the isolation of two phytase-producing isolates *Acinetobacter* sp. AC1-2 and *Buttiauxella* sp. CH 10^{-6}-4, genome

sequencing and degenerate primers confirmed the presence of a Multiple Inositol Polyphosphate Phosphatase gene and Histidine Acid Phytase respectively. The *Acinetobacter* phytase is further described in this thesis as it represents one of the first MINPP phytases that have been isolated from the soil environment.

Furthermore, from the knowledge gained from the previous isolation experiment an additional phytase isolation experiment was performed on the well-known Rothamsted soils. This time the bacterial isolates were allowed to grow for over a month to allow for diverse, slow-growing bacteria to develop on the plates.

Of the 66 isolates that were restreaked from the mixed bacterial plates, 17 of the isolates showed a diverse array of phytase activities highlighting the diversity of phytase reactions that can occur in the soil environment. Four gram-positive bacteria were isolated, *Curtobacterium* sp., *Pseudarthrobacter* sp., *Arthobacter* sp., *Bacillus subtilis* and two-gram negative bacterium, *Acinetobacter pittii* and *Serratia* sp., Unfortunately however, not all of the isolates could be sequenced due to sample contamination. In addition to phytate degraders, there were also 21 identified $InsP_5$ degraders who showed no signs of inositol phosphate degradation but produced an unusual HPLC chromatogram, further highlighting how important it is to definitively detect phytate degradation.

Finally, the water, aerobic and anaerobic samples from iron/sulphur cores from the Blakeney salt marsh, Norfolk, were investigated for phytate degradation. The previous "inland" soils that had been examined in this chapter, primarily accumulated $InsP_5$ in the 4/6-position, however, when moving to an aquatic environment instead the primary peak was the 1/3-position indicating a higher abundance of β-propeller phytases. The strongest signs of phytate degradation were present in the water sample of the sulphur core, which is in agreement with previous literature analysis which saw that BPPhy's were four orders of magnitude higher in water than in sediment (Stout *et al*., 2016). The iron core samples showed no signs of phytase activity, as iron is a strong inhibitor of activity however, this is not an unexpected result. Interestingly however, two additional peaks appeared on the HPLC chromatogram from the sulphur/iron aerobic samples and iron anaerobic samples which may indicate the presence of *scyllo*-$InsP_5$ and a *scyllo*-$InsP_4$ or L-*chiro* or D-*chiro*-$InsP_5$.

Currently, this thesis has successfully used culture-dependent techniques to examine the diversity of phytase degrading microorganisms in a variety of environments. However, it has not looked into the characteristics of any of these isolates nor has it examined culture-independent techniques to examine the unculturable phytase degraders. Chapters 3 and 4

describe the purification and characterisation of the soil MINPP from *Acinetobacter* sp. AC1-2, Chapter 5 describes a metagenomic study into the diversity of MINPP, HAPhy, BPPhy and PTPhys genes in a variety of environments. Finally, Chapter 6 examines the controls of phytase expression in *Acinetobacter* sp. AC1-2.

Chapter 3. The Purification of a Multiple Inositol Polyphosphate Phosphatase from *Acinetobacter* sp. AC1-2

The activity of the phytase isolated from *Acinetobacter* sp. AC1-2 was similar to that of a Multiple Inositol Polyphosphate Phosphatase (MINPP). This was confirmed by genomic sequencing and inspecting the assembled genome. The gene contained the heptapeptide sequence motif that is typical in the histidine acid phytases (Lei and Porres, 2003) but instead of the dipeptide motif, HD, it contained a tripeptide donor motif, HAE, that is a useful identifier in MINPPs (Stentz *et al*., 2014). This phytase type was first associated with eukaryotic organisms, being identified in plants and animals (Cho *et al*., 2006; Mehta *et al*., 2006) and was later discovered in enteric bacteria (Acquistapace *et al*., 2020; Stentz *et al*., 2014).

The discovery of a soil MINPPs from *Acinetobacter* sp. AC1-2 represents one of the first to be isolated from a soil bacterium. In this Chapter, expression trials were performed to identify the best expression system to produce a pure, active protein.

3.1. Materials and Methods

3.1.1. Phenol-chloroform method to extract genomic DNA.

Acinetobacter sp. AC1-2 cells were grown overnight in two 10 mL LB cultures at 180 RPM and 37 °C. These were pelleted for 10 minutes at 8422 g and resuspended in 500 µL Buffer P1 (QIAGEN resuspension buffer). Lysozyme, 5 mg/mL stock concentration, was added and incubated for an hour mixing regularly. SDS was added to 1% (≈60 µL of 10% SDS) and an equal volume of phenol chloroform added. These were vortexed thoroughly and centrifuged at 16060 g for 5 minutes. The top layer was removed to a fresh tube, and the steps from the addition of phenol chloroform repeated 3 times and finally with chloroform once. Next, three volumes of 100% ethanol were added, and the sample incubated at -20 °C for 30 minutes. The DNA was pelleted at 16060 g for 20 minutes and the supernatant removed. The DNA was washed with 500 µL of 80% ethanol, the ethanol decanted, and the sample pelleted at 16060 g for 1 minute. The samples were air-dried in the flow hood for up to 5 minutes, resuspended in 100 µL dH_2O and allowed to rehydrate overnight in the fridge. The concentration of the DNA was measured using a NanoDrop™ (ThermoScientific), yielding concentrations of 244 ng/µL and 158 ng/µL.

3.1.2. Plasmids and Bacterial Strains

Plasmids

The *Minpp* phytase gene was cloned using Gateway Cloning which employed three plasmids. pDONR™ 207, a gentamycin resistance plasmid that is used first to transfer the gene of interest to the expression vector. pDEST-17, chloramphenicol resistance plasmid that encodes an IPTG-inducible T7 expression system and a 6x His-tag. Likewise, pH9-GW, kanamycin, is a pET28a(+)-derived plasmid that has been rendered Gateway compatible. It contains a T7 expression system and a 9x His-tag. The plasmid maps are shown in Figure 3.1-3.3. Note, pH9-GW does not have a plasmid map and so pET28a(+) has been used instead.

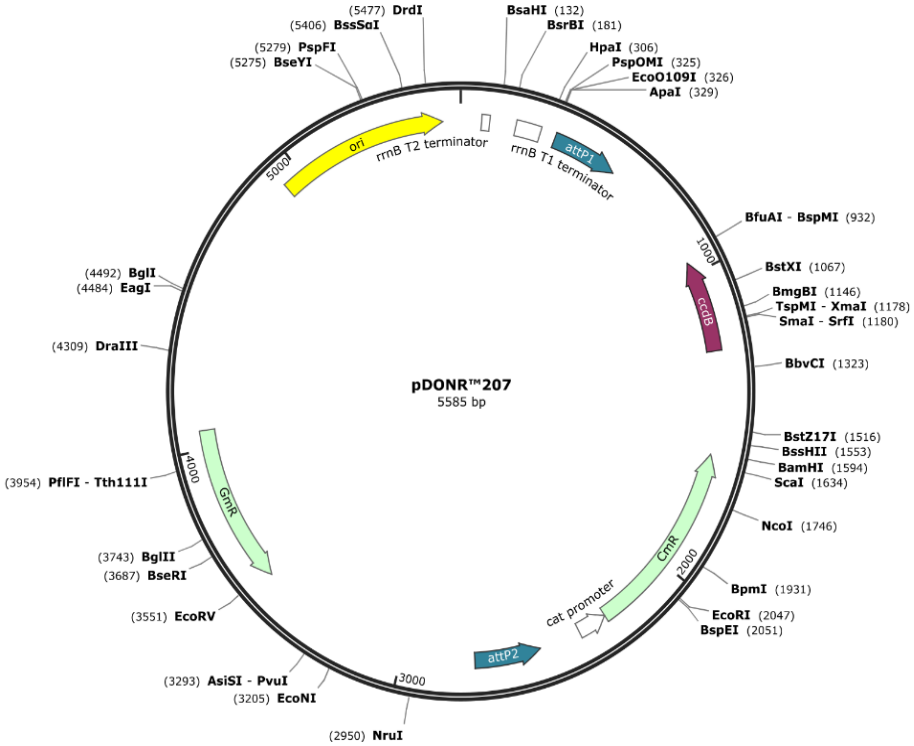


Figure 3.1 – Plasmid map of pDONR™207. pDONR™ 207 is the Entry Vector for Gateway cloning containing the characterisitic *att*P1 and *att*P2 sites that are described in Section 3.13. Vector maps were made using Snapgene (https://www.snapgene.com).

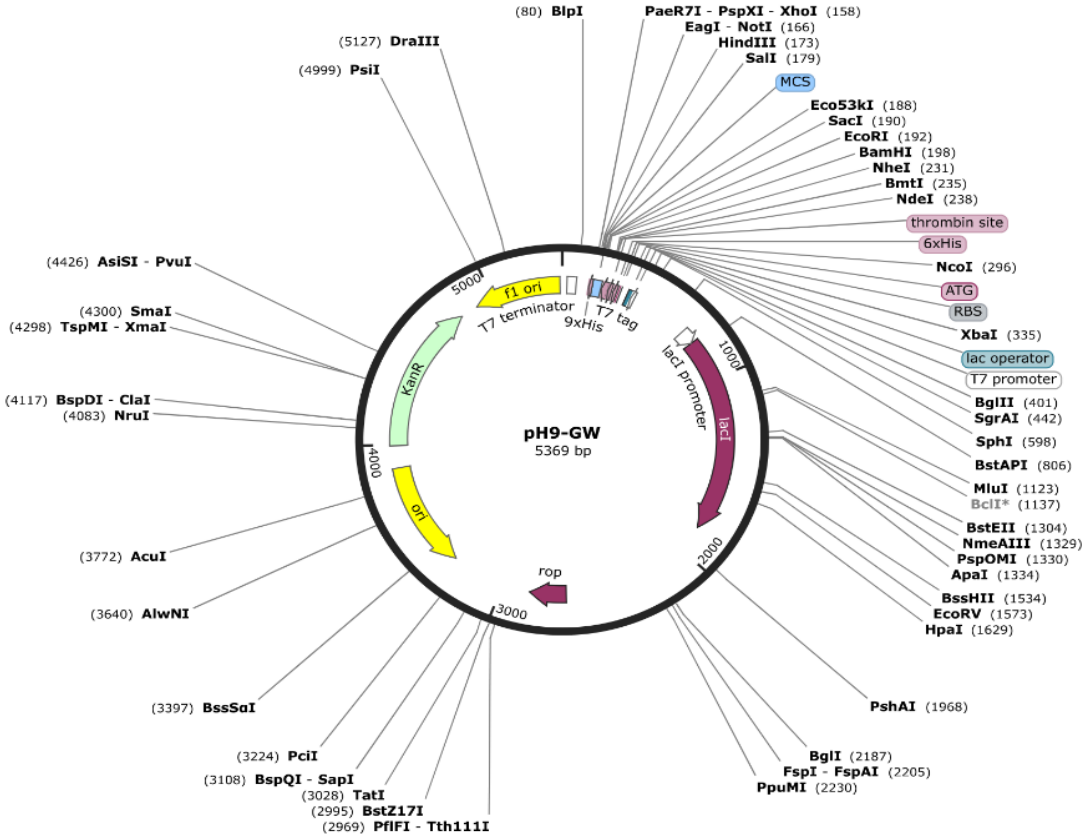


Figure 3.2 - Plasmid map of pH9-GW. pH9-GW is a pET28(+)-derived plasmid that has been made Gateway compatible. Its expression system is control by a T7 promoter and it contains a 9x His-tag that is before the *att*R1 sites. Vector maps were made using Snapgene (https://www.snapgene.com).

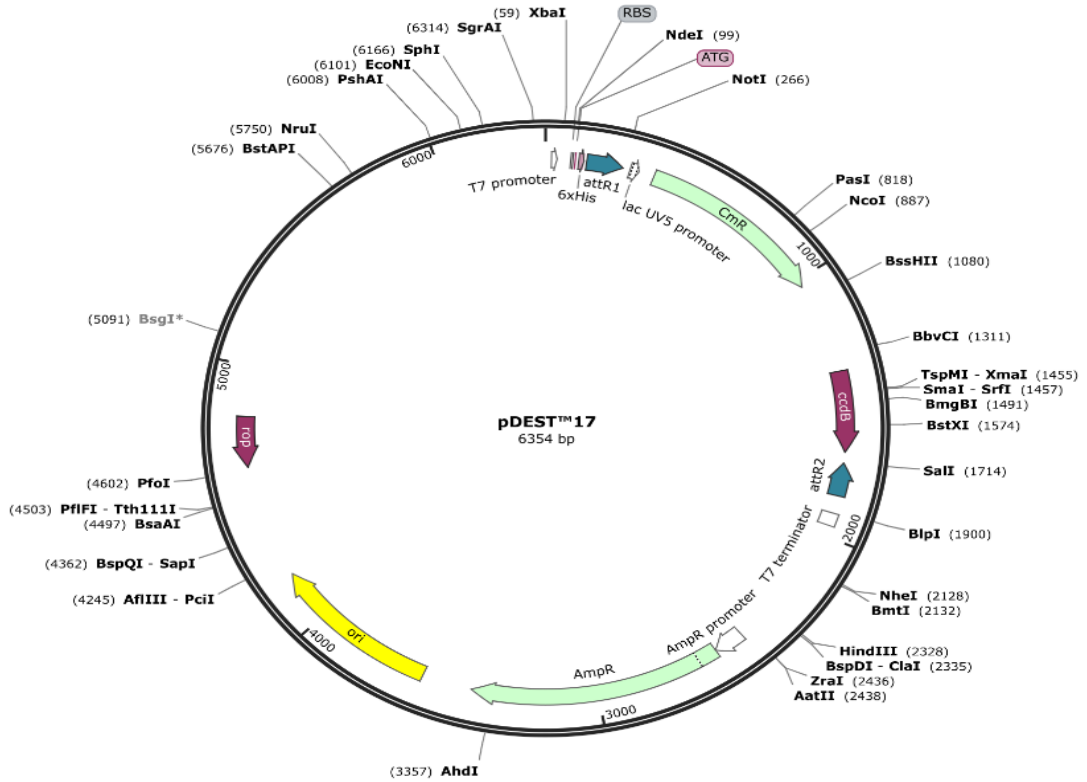


Figure 3.3 - Plasmid map of pDEST™ 17. pDEST™ 17 contains a 6x His-tag and is controlled by a T7 promoter. Vector maps were made using Snapgene (https://www.snapgene.com/).

Bacterial Strains

The strains used in this chapter are shown in Table 3.1.

Table 3.1 – Bacterial strains used in plasmid amplification and protein expression.

Strain	Resistance	Origin	Characteristics
DH5α	None	B line	Chemically competent cells with high transformation efficiency. Useful for routine subcloning of genes into plasmid vectors
BL21	None	B line	Widely used non-T7 expression. Suitable for transformation and protein expression
BL21 pLysS	Chloramphenicol	BL21 derivative	pLysS encodes T7 lysozyme which suppresses the basal expression of uninduced genes.
Rosetta 2™ (pLysS)	Chloramphenicol	BL21 derivative	Enhances the expression of eukaryotic proteins that contain codons rarely used in *E. coli.* pLysS encodes T7 lysozyme which suppresses the basal expression of uninduced genes.
Shuffle T7 Express	Spectinomycin and Streptomycin	BL21 derivative	Enhanced capacity to correctly fold proteins with multiple disulfide bridges.
Arctic Express (DE3)	Gentamycin	BL21 derivative	Expression of proteins at low temperatures for improved protein solubility

3.1.3. Gateway cloning of the *Acinetobacter* sp. AC1-2 *Minpp* gene

Gateway cloning technology is based on the site-specific recombination reactions of bacteriophage λ in *E. coli*. Bacteriophage λ can integrate site-specifically into the genome of *E. coli* by a process called lysogenisation. The overall process is mediated by proteins encoded by lambda and *E. coli.* The process can be represented as follows:

*att***B** x *att***P** ⇔ *att***L** x *att***R** (where x signifies recombination)

The four *att* sites, in Donor, Entry Clone, Expression Clone and Destination Vector, contain binding sites for the proteins that mediate the recombination reactions. By combination of

the BP and LR reactions, genes can be moved between Entry Clones and Expression Clones. A basic schematic can be seen in Figure 3.4.

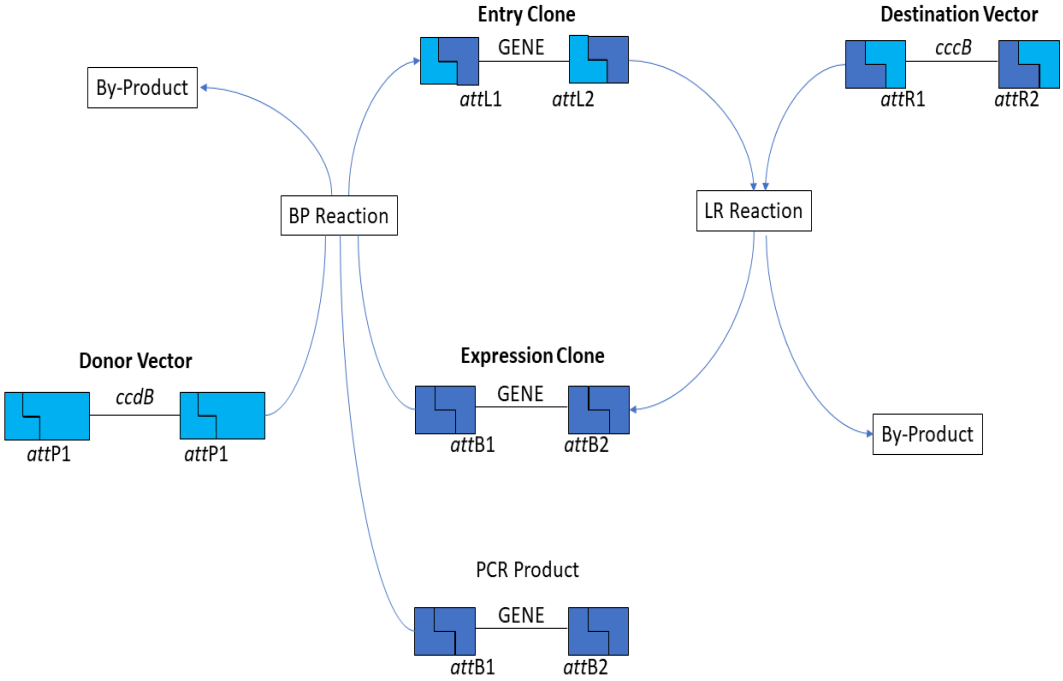


Figure 3.4 – Gateway Cloning process for the transformation of a desired gene into an expression clone. The Gene of interested is cloned with *att*B1 sites and can be transferred into the Donor Vector (Via the BP reaction) to form the Entry Clone. This can then be transferred into the Destination Vector (via the LR reaction) to generation the Expression Clone which can then be used for protein expression. This image was remade from the Gateway Cloning Technology manual.

3.1.4. Generation of the Entry Clone

Primers were designed for Gateway Cloning of the *MINPP* gene in accordance with the Gateway™ Cloning Technology manual from Life Technologies, Table 3.2.

Table 3.2: Primer sets used for Gateway cloning. The highlighted red sequence is the 3C protease cleavage site. The highlighted blue sequences are the *att*B1 and *att*B2 recombination sites.

Primer Set 1:	**FOR: 5′ -CTGGAAGTTCTGTTTCAGGGCCCG…ATGAATATTTTATTTAAAACGACGAT-3′** **REV: 5′ -CAAGAAAGCTGGGTTTAA…TTGTATTTGATAGCACTGTTTC-3′**
Primer Set 2:	FOR: 5′ -GGGGACAAGTTTGTACAAAAAAGCAGGCTTCCTGGAAGTTCTGTT-3′ REV: 5′ -GGGGACCACTTTGTACAAGAAAGCTGGGTT-3′
Primer Set 3:	FOR: 5′ -CTGGAAGTTCTGTTTCAGGGCCCG…AACAATAATGACGATCAAGATG-3′ REV: 5′ -CAAGAAAGCTGGGTTTAA…TTGTATTTGATAGCACTGTTTC-3′

These were used to amplify a PCR product of the *MINPP* gene flanked by *attB* recombination sites, bearing a 5' localized 3C-protease site. Because the primers were too large to clone both the *attB* and 3C-protease sites, a two-step PCR was employed with the products from the first reaction used in the second. Primer Sets 1 and 2 and Primer Sets 2 and 3 were used to generate the entry clone.

The gene was amplified using a high-fidelity polymerase (Phusion® High-Fidelity DNA Polymerase, New England Biolabs) using template genomic DNA from 1.25 - 20 ng/µL. Each PCR mixture contained: 12.2 µL H_2O, 4 µL 5X HF Buffer, 1 µL Primers (10 µM), 1 µL dNTPs (10 µM), 0.6 µL 100% DMSO, 0.2 µL Phusion Polymerase and 1 µL DNA.

PCR protocol: initial denaturation, 95 °C, 3:00 (min:sec); denaturation, 95 ℃, 0:20; annealing, 50 ℃, 0:30; elongation 72 ℃, 1:00; with 30 cycles and final elongation at 72 ℃ for 10:00. The PCR products were purified using the QIAquick PCR purification Kit (Qiagen)

3.1.5. Integration of the Entry Clone into the Entry Vector (BP Reaction)

The PCR product was transferred into the Donor Vector, pDONR207 (gentamycin resistance), using the BP reaction. A typical BP reaction contained 2 µL TE buffer, 1 µL BP Clonase, 1 µL template DNA and1 µL pDONR 207 plasmid. Reaction mixtures were vortexed briefly and incubated for 3hr30 min at 25 °C, before inactivation of the reaction by the addition of 0.5 µL Proteinase K for 10 min at 37 ℃.

3.1.6. Transformation of the Entry Vector into DH5α cells

After the inactivation of the BP reaction, the donor vector was transformed into DH5α cells. 1 µL of the BP reaction was added to 20 µL DH5α cells. These were mixed gently, left on ice for 30 min, heat-shocked at 42 ℃ for 35 seconds before being returned to ice for 2 minutes. 180 µL SOC (Super Optimal Broth with catabolite repression – ThermoScientific) was added and the reaction incubated at 37 °*C* for 1 hour with shaking at 180 rpm. 30 µL was spread onto gentamycin, or other appropriate antibiotic plates, 20 µg/mL. The remainder was pelleted at 4738 g for 2 minutes, 90 µL of the supernatant was removed, and the remaining media used to resuspend the pellet and the whole streaked again onto gentamycin. Plates were incubated overnight at 37 ℃.

3.1.7. Confirming correct Entry Vector Inserts

Twenty-three colonies that had grown overnight were screened for the insertion of the entry vector by colony PCR using the same PCR protocol as 2.1.8. Internal vector primers that included the *att*L1 and *att*L2 were used: **F: 5'-**GCAGTTCCCTACTCTCGC**-3'** and **R: 5'-**CATCAGAGATTTTGAGACAC**-3'.** The amplified DNA was resolved using a 1% w/v agarose gel, Section 2.1.9. The remaining DNA were purified and sent for sequencing at Eurofins (MWG, Ebersberg, Germany). The sequencing results however, struggled to show the whole gene, particularly in the middle where it calls 3 G's instead of 2. Therefore, an internal primer was designed, 5' TGA TTT AGA AGC AAT GAT G 3,' 300 bases downstream of this area. Sequencing confirmed correct integration of the *MINPP* gene.

3.1.8. Assessment of Sequencing Quality

Eurofins (MWG) provides quality assessment for each sequencing sample. The quality is based on a 0-9, 10-19, 20-29 and 30 =< quality score. The beginning and ends of the sequences were trimmed to ensure the best quality data. These were then compared with the expected sequences to highlight any errors/mismatches.

3.1.9. Transformation of the Entry Clone into the Destination Vector

Correct colonies containing the entry vector were grown overnight in 10 mL LB cultures at 37 °*C* with shaking at 180 RPM and the vector DNA isolated using the QIAprep Spin Miniprep Kit (Qiagen). The PCR product was transferred from the donor vector to the destination vector using the LR clonase (LR reaction). A typical LR reaction contains 2 µL TE buffer, 1 µL LR Clonase, 1 µL pDONR™207 DNA and 1 µL pDEST17 (Ampicillin) or pH9GW (Kanamycin) plasmid. The reactions were vortexed briefly and incubated for 3hr30 min at 25 °*C,* before inactivation of the reaction by the addition of 0.5 µL Proteinase K for 10 min at 37 ℃. The resulting reaction was cloned into DH5α, and correct inserts identified as described in Sections 3.1.6 and 3.1.7.

3.1.10. Transformation of the Expression Clone into the Expression Host

Colonies containing correctly inserted expression clone were grown overnight in 10 mL LB cultures at 37 °*C* with shaking at 180 RPM. Vector DNA was isolated using the QIAprep Spin Miniprep Kit (Qiagen). The pDEST17 and pH9GW expression clones were subsequently

transformed into *E. coli* BL21 (New England Biolabs), *E. coli* BL21 pLysS (chloramphenicol) (Invitrogen) and Rosetta™ 2 (pLysS) (chloramphenicol) (Novagen) as described in Section 3.1.6 and 3.1.7.

3.1.11. Small Scale Expression and Protein Purification

Single colonies from Rosetta 2 – pDEST17 MINPP were grown overnight in 10 mL LB cultures at 37 ℃ with appropriate antibiotic and with shaking at 180 RPM. The cultures were used to inoculate fresh 10 mL LB which was grown at 37 ℃ with shaking at 180 RPM. At an OD_{600} of 0.4-0.6, the temperature was lowered to 20 ℃, IPTG added to a final concentration of 0.2 mM, and the samples incubated overnight.

The cells were harvested by centrifugation for 10 min at 8422 g in 50 mL falcon tubes, the pellet was stored at -80 ℃ for 30 min and then thawed. The pellet was resuspended in 1 mL lysozyme buffer containing 50 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% (w/v) glycerol, 20 mM Imidazole, 1% (w/v) Triton X-100, cOmplete Protease Inhibitor (Half a tablet - Roche), 1 mM EDTA and lysozyme (1 mg/mL) was added. This was incubated at 30 ℃ for 15 min with shaking at 230 RPM. After which, a 10 µL of Benzonase solution (Sigma-Aldrich) was added and the samples incubated at 30 ℃ for 15 min with shaking at 230 RPM. The samples were transferred to a 1.5 mL microfuge tube and centrifuged at 21460 g for 15 min. The supernatant (clarified cell lysate) was added to 100 µL of HisPur™ Ni-NTA agarose resin (ThermoScientific) that had been pre-equilibrated in a 1.5 mL microfuge tube by washing twice with 200 µL lysis buffer containing 50 mM Tris-HCl pH 8, 500 mM NaCl, 20 mM imidazole, 10% (w/v) glycerol and 1% (w/v) Triton X-100. The cell lysates were incubated at 4 °C for 1 hour with gentle shaking.

The samples were centrifuged for 1 min at 2106 g and the supernatant discarded. The resin was washed with 400 µL lysis buffer three times, with the supernatant discarded each time. Subsequently the resin was washed with 400 µL of 50 mM Tris-HCl pH 8, 500 mM NaCl, 20 mM imidazole, 10% (w/v) glycerol, three times, before elution of protein from the resin with 100 µL elution buffer; 50 mM Tris-HCl pH 8, 500 mM NaCl, 600 mM imidazole. The resin-lysates were incubated at 4 °C for 30 min with gentle shaking and centrifuged for 3 min at 2106 g. The supernatant was collected, and aliquots thereof analysed by SDS-PAGE.

3.1.12. SDS-PAGE Gel Electrophoresis

Two 12% (w/v) SDS Polyacrylamide resolving gels were prepared using 3.96 mL 30% acrylamide, 3.33 mL dH_2O, 2.5 mL 1.5 M Tris pH 8.8, 0.1 mL 10% SDS, 0.1 mL 10% APS and 4

µL TEMED. 2X 5% stacking gel was prepared using 0.83 mL 30% acrylamide, 3.4 mL dH_2O, 0.6 mL 1 M Tris pH 6.8, 50 µL 10% SDS, 50 µL 10% APS and 5 µL TEMED. The gel was run at 120 V until the dye had reached the bottom.

3.1.13. Bradford Assay

A Bradford assay was performed to measure the concentration of protein from the small-scale purifications. These were prepared in a 96-well plate (Thermoscientific, Nunc™, polystyrene) A calibration curve was obtained by serial dilution of BSA (0.5, 0.25, 0.125, 0.0625, 0.03125, 0.015625 and 0 mg/mL) in elution buffer. For resin-purified proteins, 10 µL were added to 10 µL of elution buffer in wells. Colour was measured at 595 nm in a HidexSense (Turkuu, Finland) plate reader after the addition of 200 µL Bradford solution (Quick Start™ Bradford protein assay, BIO-RAD) to each of the wells. The absorbance was then subtracted from the y-intercept of the calibration curve divided by the gradient to give the concentration of protein in mg/mL.

3.1.14. Measuring Phytase Activity using the Molybdenum Blue Method for Phosphate Release

Phytase specific activity (U/mg) was determined using the molybdenum blue method for phosphate release. One phytase unit was defined as the amount of the enzyme that released 1 µM inorganic phosphate per minute under the assay conditions. All samples were assayed in triplicate (Dokuzparmak *et al*., 2017).

Briefly, 8.75 µL enzyme solution (concentration stated in the results section) was diluted in 148.75 µL of 0.2 M Na-acetate buffer pH 5.5. Forty-five µL of the enzyme buffer solution was mixed with 5 µL 50 mM $InsP_6$ in triplicate on ice before being heated for 15 minutes at 37 °C in a PCR machine. After which, 50 µL of a 4:1 ratio of ammonium molybdate sulphuric acid solution (400 mL, 6 g $NH_4Mo_7O_{24}.4H_2O$ and 22 mL 98 % H_2SO_4) and ferrous sulphate solution (20 mL, 2.16 g iron (II) sulphate heptahydrate, 2 drops 98% H_2SO_4) was added causing the solution to go blue from phosphate release. Absorbance at 700 nm was measured in a Hidex sense plate reader after 15 min and compared with a phosphate calibration curve. Initial, pilot calibrations were performed with AC1-2 MINPP solutions of known concentrations to determine appropriate concentration of protein for routine assays.

3.1.15. Large Scale Protein Purification Prep

A pre-culture of single colonies of Rosetta™ 2 (pLyss) – pDEST17-Minpp, and Rosetta™ 2 (pLyss) – pH9GW-MINPP were each grown in 100 mL LB with the appropriate antibiotics overnight at 180 RPM and 37 °C. Four 500 mL LB cultures were inoculated with 12.5 mL of the pre-culture each, two for each plasmid type, with the appropriate antibiotics, chloramphenicol was not added, however. These were grown until an OD between 0.5-0.7 and IPTG was added to a final concentration of 0.2 mM to induce protein expression. The flasks were incubated overnight at 180 RPM and 20 °C. Flasks of the same plasmid type were combined and spun down at 3243 g at 20 °C (JLA 8.1 rotor in a J-20 centrifuge) for 20 minutes. The supernatant was poured off and the cell mass resuspended in 25 mL Binding buffer: 50 mM Tris-HCl pH 7.5, 300 mM NaCl and 20 mM imidazole. To the resuspended cells, half a cOmplete Protease Inhibitor tablet and lysozyme (1 mg/mL) was added, and cell lysis was carried using a French press (3-4 cycles, 1000 PSI, Internal pressure of 18000 PSI). The soluble and insoluble fractions were separated by centrifugation at 29596 g at 4 °C (JA 25.50 rotor in a J-20 centrifuge) for 45 minutes.

The supernatant was then passed through a 1 mL Histrap HP column using a ÄKTA pure protein purification system for Immobilised Metal Affinity Chromatography (IMAC) and eluted with elution buffer: 50 mM Tris-HCl pH 7.5, 300 mM NaCl and 500 mM imidazole. Fractions emitting UV absorbance were pooled together and then concentrated down to 1.5 mL using a Amicon® Ultra-15 centrifugal filter unit with a 30 kDa cutoff point. The concentrated protein was then injected into a 1.5 mL loop and then further purified using a HiLoad 16/600 Superdex 75 pg column for size-exclusion chromatography. Gel Filtration buffer: 50 mM Tris-HCl pH 7.5 and 300 mM NaCl.

Three peaks were present on the chromatograph; therefore, all the fractions were tested for activity as per Section 3.1.14. Only the fractions from the first peak showed phytase activity, and so these were pooled together and concentrated using a Amicon® Ultra-15 centrifugal filter unit with a 10 kDa cut-off and 12 mL was concentrated down to around 1.5 mL. A small sample was taken to examine the protein concentration in solution. This was done using a nanodrop, measuring absorbance at 280 nm – 1 mg/mL. The average protein concentration was 0.682 ± 0.029 mg/mL which is approximately 11.37 µM. Stock concentrations of 4 µM protein aliquots were prepared with 25% trehalose.

The activity of the Minpp was determined using the molybdenum blue method for phosphate release and the purity examined using SDS-PAGE. The activity was very low, and

the SDS-PAGE showed significant protein presence in the pellet indicating that the protein isn't very soluble, the concentrated gel filtration samples were also analysed showing multiple bands present in the sample.

3.1.16. Optimization of the Solubility and Purity of Minpp AC1-2

Removal of the Signal Peptide

One method of improving the solubility of the MINPP phytase was removal of the signal peptide. The signal peptide of AC1-2 MINPP was determined using Signal P 5.0, a web-based program (Almagro Armenteros *et al*., 2019), Figure 3.5. This predicted the signal peptide to be a lipoprotein signal peptide (Sec/SPII) with the cleavage site between positions 18 and 19 based on a probability score, highlighted in green, LFLVACNNND. Therefore, an additional Primer Set was designed (Section 3.1.4 Primer Set 3).

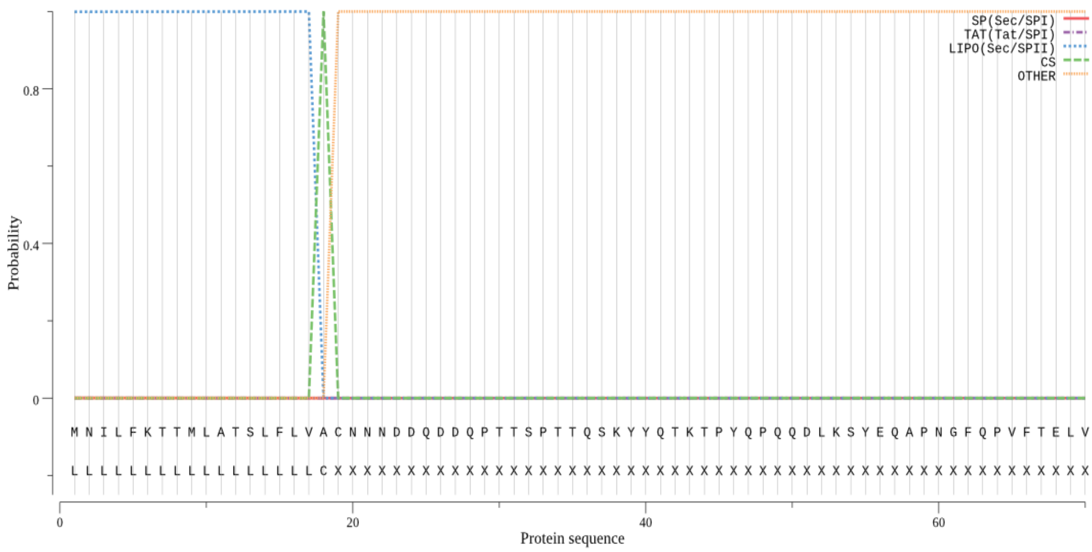


Figure 3.5 – The prediction of the signal peptide cleavage site using SignalP-5.0 based on the protein-coding sequence of AC1-2 MINPP.

Primer Set 3 was used to clone the *Minpp* gene without the signal peptide via Gateway Cloning as described, Section 3.1.3-3.1.13.

3.1.17. Cloning into sHuffle® T7 express and ArcticExpress (DE3)

Additionally, the expression clone lacking signal peptide was also transformed into expression strains sHuffle® T7 Express and ArcticExpress (DE3).

sHuffle® T7 Express was developed to tackle one of the problematic areas of cloning into *E. coli*: the lack of post-translational modification of disulphide bonds. These can be formed

inter- and intramolecularly and can be critical for proper protein folding, stability, and or activity (Lobstein *et al*., 2012). ArcticExpress (DE3) was developed to address the issues that can occur during the overexpression of heterologous proteins in *E. coli*. This strain is suited to low-temperature cultivation, it expresses the cold-adapted chaperonins Cpn10 and Cpn60, which, within the strain, aid protein refolding at 4-12 °C (Belval *et al*., 2015).

3.1.18. Small Scale Expression Trials of AC1-2 no signal construct

As described in Section 3.1.3-3.1.13, the AC1-2 *MINPP* no signal construct was expressed and MINPP purified on a small-scale using Rosetta™ 2 (pLysS), sHuffle® T7 Express and ArcticExpress (DE3). IPTG at concentrations 0, 0.01 mM, 0.05 mM, 0.1 mM and 0.5 mM were used to identify whether over-inducing the sample causes insolubility. The proteins were all tested for phytase activity and analysed by SDS-PAGE to examine purity. A protein band obtained from Rosetta™ 2 (pLysS) lysates was sent for sequencing using Protein Mass Fingerprinting at the John Innes Centre to confirm the expression of the AC1-2 MINPP protein.

The analysis reported a protein score of 51, with 14% cover of the protein sequence: protein scores greater than 51 are significant ($p<0.05$), Figure 3.6.

The matched peptides are shown in bold red in the sequence below.

1 HHHHHHLEST SLYKKAGFLE VLFQGPNNND DQDDQPTTSP TTQSKYYQTK

51 TPYQPQQDLK **SYEQAPNGFQ PVFTELVAR**H GSRGLSSLKY DLALYNLWKQ

101 AKAENALTPL GEQLGADLEA MMK**ANILLGY GVEGIR**QYGY GNETMTGILE

151 HRGIADR**LLQ RLPNLLNPQA GILVQSSGVD R**AVDSAKFFT AELIKQQPQL

201 KDKIVPVSYT NLSSESVPSV IDGGVDRFKL YFHSLNADED LTQPLSASQQ

251 KIYDASQAYQ DFEENNNDLA QKLDELSKNT QAEKTAQTVL NPIFKADFIK

301 KLGTAGYSFS NTGSFTVTSP KGEQITEKGK GKNTIASAVD AAAYVYELYS

351 ISGGMKDELK GIDFDKYMPI EAAKFYAEFN DANDFYEKGP SFTESNLVTS

401 EIAQGLKQDM FQQVDAVVNK **AQPYKAVLR**F AHAEIIIPLA TSLDLHNMMQ

451 PLPLRQTYNY STSTWRGEVV SPMAANVQWD IYQNGQGNTL VKMLYNEKET

501 LFKSACNYAR **YSPTSFYYDY IK**LKQCYQIQ

Figure 3.6 – The protein sequence of His-tagged AC1-2 MINPP phytase, with peptides that were identified by mass-spectrometry highlighted in red.

The peptides identified in Figure 3.6 were blasted against the *E. coli* genome to ensure that they did not correspond to *E. coli* gene products. The E-values were significantly greater than that of AC1-2 MINPP, confirming the identity of the purified protein as AC1-2 MINPP phytase.

Subsequently, large-scale protein preparation was repeated, from a 2 L culture, to confirm the isolation of a purified and active phytase.

3.1.19. Transformation and Expression Trials in *Pichia pastoris*

Attempts were also made to improve the solubility of the AC1-2 MINPP phytase by transformation and expression in *Pichia pastoris* X-33 using the Easy Select™ *Pichia* Expression kit (Invitrogen).

Use of eukaryotic organisms, yeasts, such as *Pichia pastoris,* as expression system provides several benefits that are not available with bacterial expression hosts. The proteins are directly secreted into the media, the yeast cells have the advantage of higher eukaryotic expression systems such as protein processing, protein folding and post-translational modification. Of the latter, one important post-translational modification is glycosylation, the attachment of carbohydrate chains, glycans, to the protein. N-glycosylation has been shown to increase the stability, thermostability and proteolytic resistance in phytases (Niu *et al*., 2016). In fact, many of the commercial phytases are also glycosylated, despite having bacterial origins (Yao *et al*., 2013).

The program NetNGlyc 1.0 was used to predict N-glycosylation sites based on the sequence of AC1-2 MINPP (Gupta and Brunak, 2002), Figure 3.7. This program uses artificial neural networks that examine the sequence context of Asn-Zaa-Ser/Thr sequons, which are a sequence of consecutive amino acids that can serve as an attachment site to a polysaccharide. This identified three potential N-glycosylation sites whose glycosylation potential was higher than the threshold value, Figure 3.7.

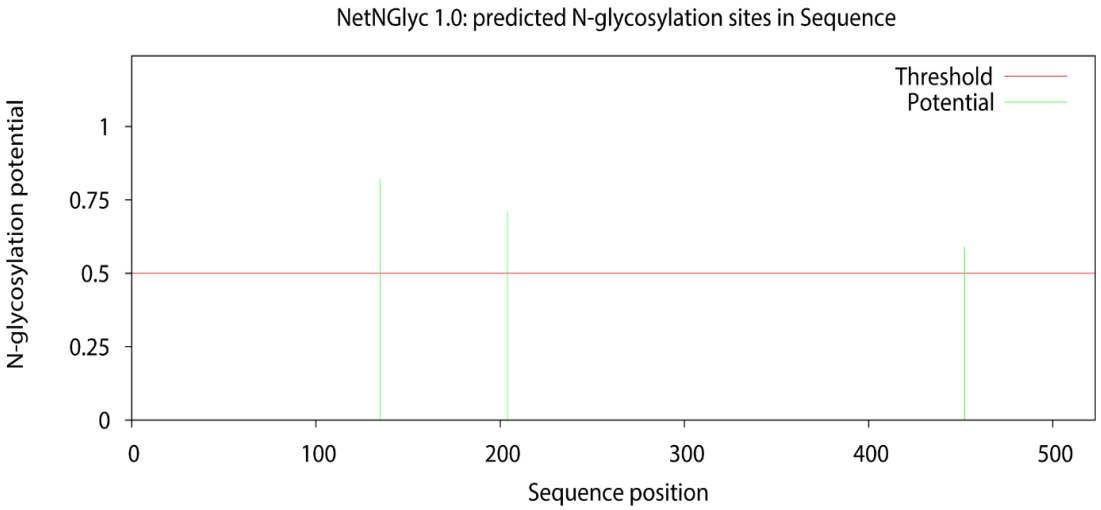


Figure 3.7 – The predicted of N-terminal glycosylation sites of AC1-2 MINPP using the online-based program NetNGlyc 1.0 which predicts glycosylation sites based on Asn-Zaa-Ser/Thr sequons, which are a sequence of consecutive amino acids that can serve as an attachment site to a polysaccharide.

Subsequently, a primer set was designed in order to clone the gene into the yeast vector pPICZα (Zeocin resistance) using Gateway Cloning as discussed in section 3.13. Preliminary experiments were unsuccessful, as there was no protein expression from the yeast expression experiment. No further detail is provided here, but because the yeast cloning was ran concurrently with the cloning to remove the bacterial signal peptide and as such several gels contain bands from both experiments.

3.2. Results and Discussion

The genomic DNA was extracted from *Acinetobacter* sp. AC1-2 cells using the phenol chloroform method from two bacterial preps, Section 3.1.1, Primers designed for Gateway Cloning of the *MINPP* gene were used to amplify the gene and clone the *att*B1 and *att*B2 sites and 3C protease recognition site for cloning into the donor vector. This required a two-step PCR as described in Section 3.1.4. Dilutions of the two genomic samples were used as template DNA. All PCR products were run on an agarose gel to confirm amplification of an appropriately sized band. These were run against a control strain of 3-phytase from *Aspergillus niger, phyA*, Figure 3.8.

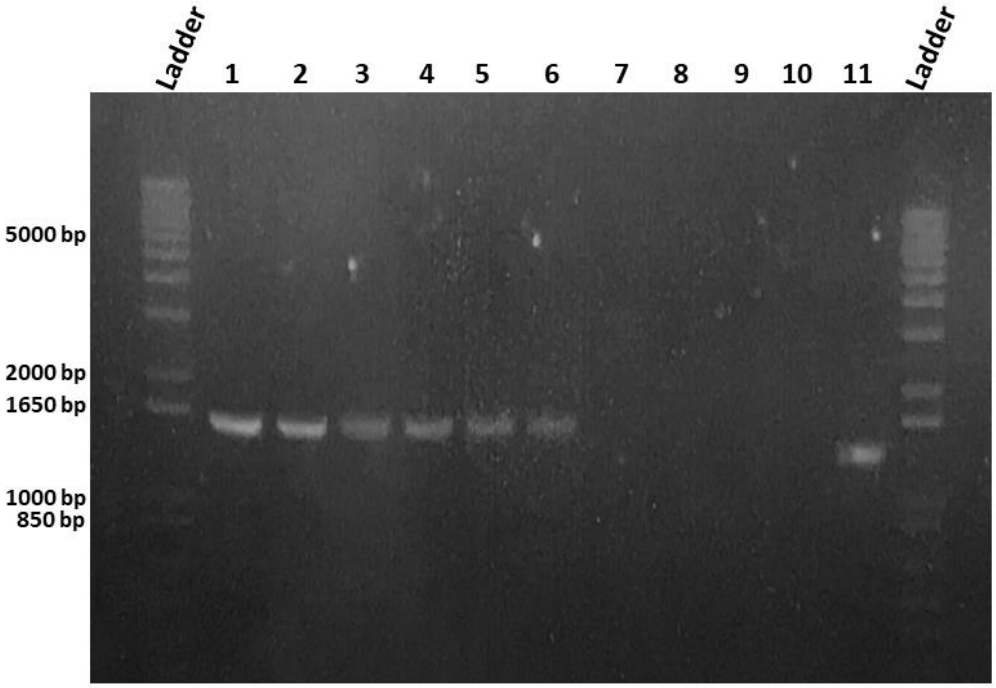


Figure 3.8 – 1% agarose gel image of the first PCR reaction required to amplify the *MINPP* gene, 3C protease site, and part of the *att*B2 adapter. Dilutions of the two DNA preps were used as a template, Lane 1-5 contained DNA concentrations from Sample 1 of 244, 20, 10, 5, 2.5 ng/µL respectively. Lane 6-10 contained DNA concentrations from Sample 2 of 158, 15, 7.5, 3.75 ng/µL respectively. Lane 11 contained a AnPhyA positive control.

The desired band size was 1614 bp, and the band of the AnphyA control at 1401 bp. Weak amplification was observed for Sample 1 and no amplification was seen in Sample 2. However, the bands for Sample 1 were the correct size and was in agreement with the positive AnPhyA control. Afterwards, Lanes 1 and 2 were diluted to 1/25, 1/50 and 1/100 and the second step of the reaction was performed to fully complete the cloning of the *att*B sites, Figure 3.9.

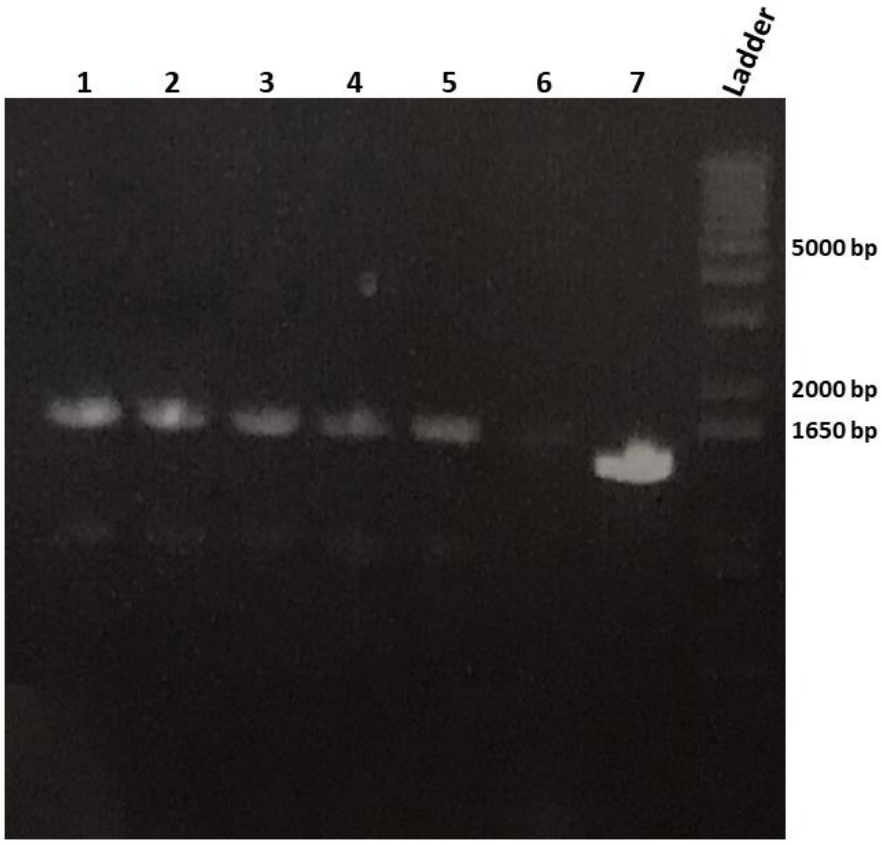


Figure 3.9 – 1% agarose gel image of the second PCR reaction to finish the amplification of the *att*B1 and *att*B2 sites using the amplified DNA from Lanes 1 and 2 in Figure 3.8. Lanes 1-3 contained dilutions of Lane 1 of 1/25, 1/50 and 1/100. Lane 4-6 contained dilutions of Lane 2 of 1/25, 1/50 and 1/100. Lane 7 was AnphyA control from the previous gel.

The desired band was 1655 bp which is in agreement with the ladder and the AnPhyA control. With the *att*B adapters cloned either site of the *MINPP* gene, the BP reaction was performed to clone the PCR product into the donor vector, pDONR207, to form the entry clone, Figure 3.2, Section 3.1.5.

These were then transformed into DH5α *E. coli* cells and grown overnight at 37 °C with gentamycin selection, Section 3.1.6. Colony PCR was performed on 23 colonies from the plate to confirm the insertion of the pDONR207 vector, Section 3.1.7 using internal primers, *att*L1 and *att*L2. This was compared with an AnPhyA pDONR207 control, Figure 3.10.

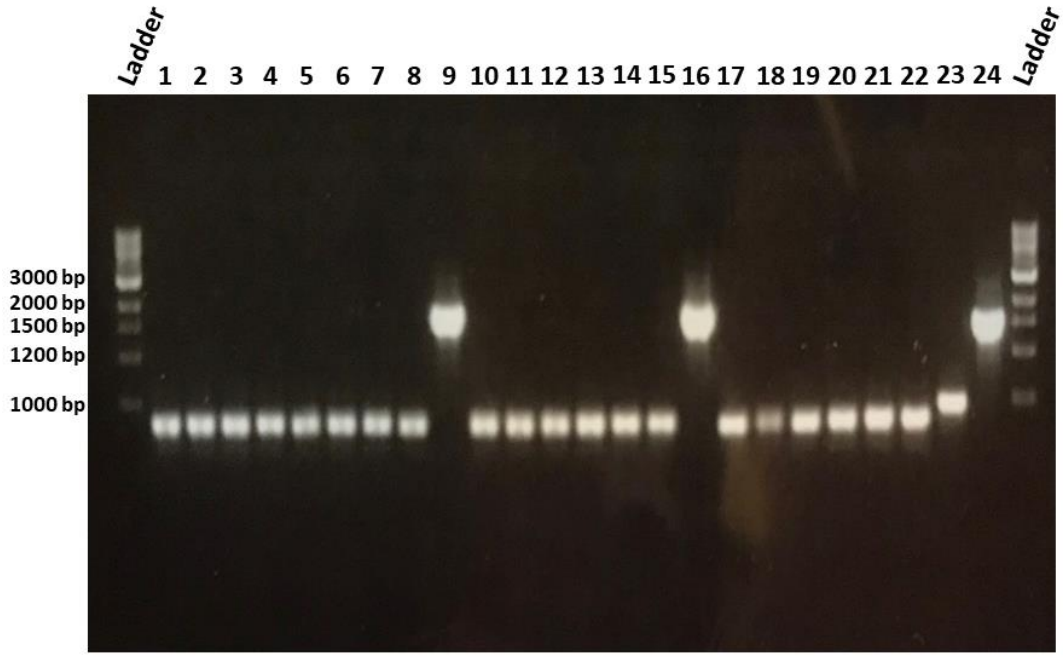


Figure 3.10 – Confirmatory agarose gel image of the BP reaction. Lanes 1-23 were the amplified single colonies; Lane 24 was the AnPhyA pDONR 207 positive control.

Only two of the twenty-three colonies analysed contained a band at the appropriate size. This was compared with the AnPhyA positive control which is predicted to be lower. The pDONR207 vector DNA for Samples 9 and 16 were mini-prepped and sent for sequencing. The *att*L1 and *att*L2 primers were used alongside a gene internal primer. These confirmed that the sequencing within the vector contained a 100% sequence identity with the *MINPP* gene.

Following this, the LR reaction was performed to transfer the entry clone into the destination vector, pDEST17 and pHGW. These were transformed into DH5α *E. coli* cells, section 3.1.6, with ampicillin and kanamycin respectively. One isolate from the pH9GW plate and six from the pDEST17 plate were used to examine insertion of the vector using colony PCR. This was compared with an AnPhyA pDEST17 control, Figure 3.11.

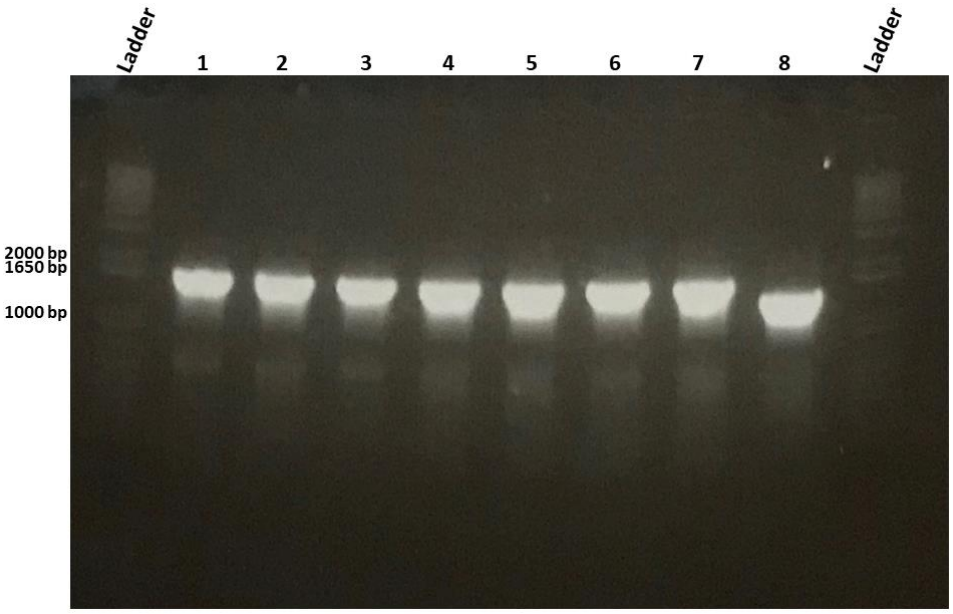


Figure 3.11 – Agarose gel image of the confirmatory colony PCR of the LR reaction. Lane 1 – pH9GW colony, Lane 2-7 - pDEST17 colonies, Lane 8 – AnPhyA pDEST17 positive control.

Both the pH9GW and PDEST17 colonies contained bands at the appropriate size. These were transformed into the expression host, BL21, BL21-pLysS and Rosetta™ 2 (pLysS).

A small-scale expression and purification experiment using Rosetta™ 2 (pLysS) pDEST17 colonies, Section 3.1.11, was performed to see how well the protein was expressed. This used Ni-NTA agarose resin that can be used to purify recombinant proteins containing a poly histidine tag. Samples were taken at each step of the purification and examined by SDS-PAGE. The samples were taken from induced cells (lanes 1-3), clarified lysate (lanes 4,6), pellet (lanes 7-9) and eluted protein (lanes 10-16), Figure 3.12.

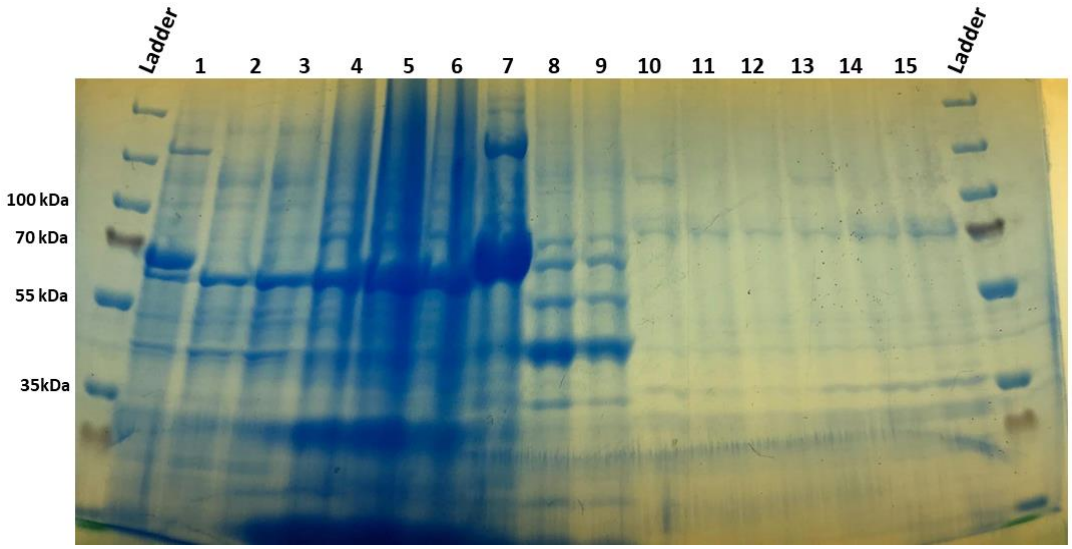


Figure 3.12 – SDS-PAGE image of the small-scale protein purification experiment using Rosetta™ 2 (pLysS)-*MINPP*-pDEST17 cells. The samples taken were induced cells, Lanes 1-3, clarified lysate, Lanes 4,6, pellet, Lanes 7-9, and eluted protein, Lanes 10-16.

With an expected size of 60 kDa, the Minpp protein was not present in the eluted protein fraction, it was present however, in the induced cells and clarified lysate samples, as expected. However, the protein was also present in the pellet sample which indicated solubility issues. In the eluted protein fraction, the MINPP band was absent, and instead a faint band around 70 kDa was visible.

Following this, a larger scale protein preparation was performed using 2 L cultures for the Rosetta™ 2 (pLysS) pDEST17 and pH9GW samples, section 3.16. After overnight expression using IPTG, the cultures were pelleted and resuspended in binding buffer. The cells were then lysed at 1000 PSI using a French Press. The sample was centrifuged to obtain a clarified lysate. The clarified lysate was then applied to a 1 mL Histrap HP column using a ÄKTA pure protein purification system for Immobilised Metal Affinity Chromatography (IMAC), Figure 3.13. There are three stages to this process: Firstly, sample application, the clarified lysate is passed through the column, the N-terminal 6/9X His-tag on the Minpp protein binds strongly to the nickel column, whilst other proteins do not bind at all or only weakly bind to the column. In the second, wash, stage, weakly bound proteins can be removed through washing the column, in Figure 3.13 we can see this stage by the large detector response signalling the elution of proteins off the column. This stage continues until the absorbance plateaus. Finally, in the third stage, the elution step, an increasing concentration of imidazole, symbolised by the green line, is added to the column. Imidazole competes with the His-tag for binding to the column, eventually, eluting the His-tagged protein.

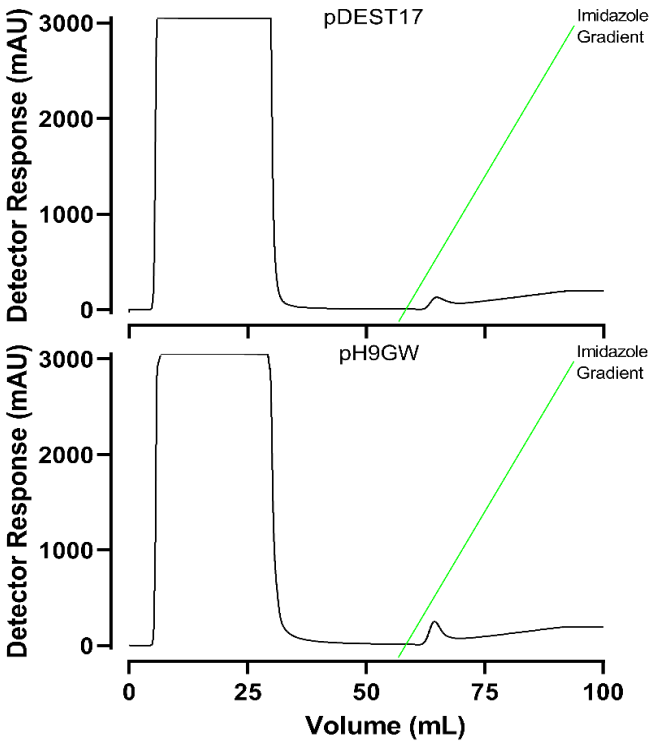


Figure 3.13 – The immobilised metal affinity chromatography (IMAC) trace of the AC1-2 MINPP protein preparations obtained from pDEST17 and pH9GW expression vectors.

Fractions showing UV absorbance above baseline were pooled from the pDEST17 and pH9GW expressed samples and the pooled sample was concentrated with a Amicon® Ultra-15 centrifugal filter unit (10 kDa cut-off). The resulting sample was applied to a HiLoad 16/600 Superdex 75 PG column for size-exclusion chromatography - as an additional purification step, Figure 3.14.

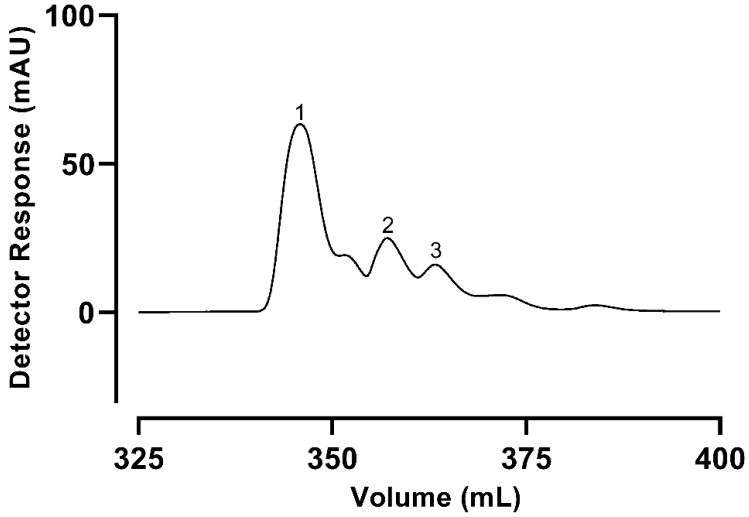
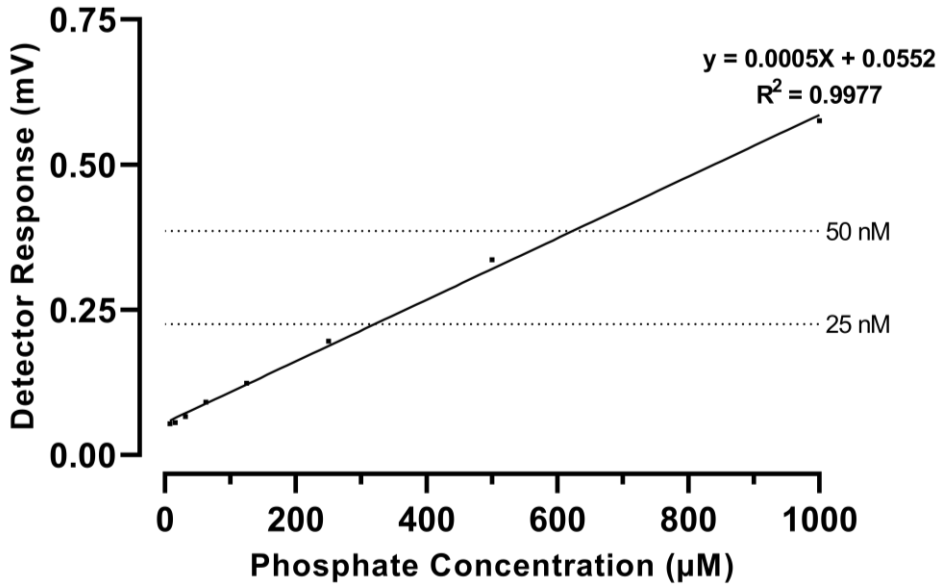


Figure 3.14 – The size exclusion chromatography (Gel Filtration) of a combined pDEST17 and pH9GW MINPP sample.

Three UV-absorbing peaks were present on the gel filtration chromatogram at elution volume 340-360 mL. Each 1 ml fraction was tested for phytase activity, and only fractions showing activity were processed further. A final pooled protein sample (12 mL) was concentrated to 1.5 mL and protein measured, yielding 0.682 ± 0.029 mg/mL (11.37 µM). The protein was combined with 50% (w/v) trehalose (50:50 ratio of gel filtration buffer and 50% (w/v) trehalose) which can act as a stabilisation agent and cryoprotectant (Leibly *et al*., 2012).

A serial dilution of enzyme from 200 to 1.5625 nM was performed to identify the appropriate concentration of AC1-2 MINPP for determination of kinetic parameters, Section 3.15. A calibration curve for phosphate release is shown, Figure 3.15.



AC1-2 Concentration (nm)	Mean Absorbance	Standard Deviaton
200	1.19	0.082
100	0.63	0.027
50	0.367	0.021
25	0.22	0.008
12.5	0.14	0.0064
6.25	0.11	0.0055
3.125	0.089	0.0031
1.5625	0.08	0.0045

Figure 3.15 – Determining the appropriate concentration of AC1-2 MINPP for phytase activity assays. Enzyme activity of a serial dilution of known protein concentration was measured using the molybdenum blue method. This was measured at 700 nm and compared with a phosphate calibration curve as absorbance is linear to phosphate concentration to a certain extent. Highlighted are the absorbance reading of 25 nM and 50 nM AC1-2 MINPP within an appropriate range.

Based on the results, the optimal concentration for phosphate release (phytase assay) is 50 nM as it yields an absorbance in the middle of the calibration curve. These absorbance values gave AC1-2 MINPP an activity value (µmol/min/mg of protein or U/mg) of 12.94 ± 0.89. This activity however, appeared to be very low for a MINPP phytase. Samples of the purified MINPP were ran on an SDS-gel with (lanes 1-2) and without (lanes 3-4) DTT, a reducing agent commonly used to disrupt protein disulphide bonds, Figure 3.16

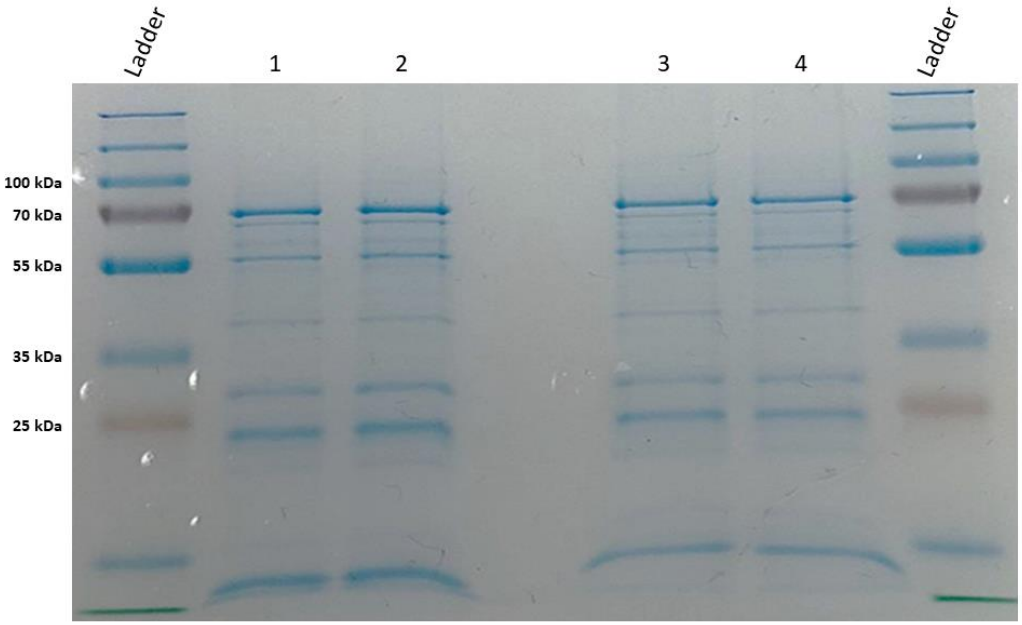


Figure 3.16 – SDS-PAGE of a large-scale protein purification from combined Rosetta™ 2 (pLysS) pDEST17 and Rosetta™ 2 (pLysS) pH9GW expressions. Lanes 1,2 purified MINPP, lanes 3,4 purified MINPP + DTT.

SDS-PAGE confirmed the low expression of the MINPP construct. The expected band of AC1-2 MINPP is 60 kDa, however there are multiple bands present on the gel, in addition to a faint band just above the 55 kDa ladder has an appropriate mobility on PAGE to be AC1-2 MINPP. Overall, the low purity and low-level expression offers an explanation of the low activity measured, Figure 3.12.

Subsequently, attempts were made to try improving the solubility and purity of the phytase, Section 3.1.16. Firstly, primer sets were designed to remove the signal peptide from the start of the sequence. Using mini-prepped DNA from the AC1-2 *MINPP* pDEST17 construct, Gateway cloning was performed as described in Sections 3.13 for transformation into *E. coli* and *Pichia pastoris*. The gel images confirming successful cloning are shown, Figure 3.17.

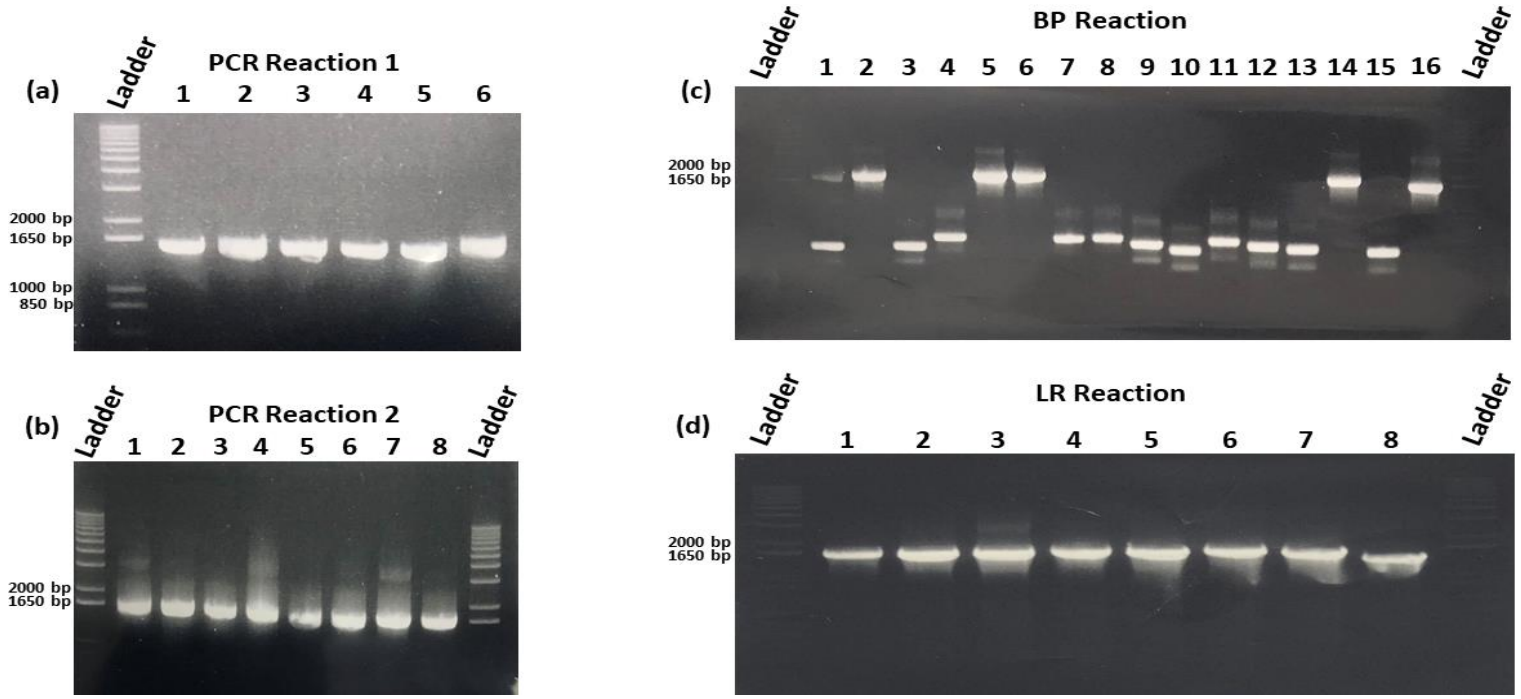


Figure 3.17 – Gateway cloning of AC1-2 *MINPP* without the signal peptide. Using the original AC1-2 *MINPP* pDEST-17 vector, dilutions were used as the template DNA. The *att*B1 adapters were cloned using primer set 3, PCR amplification products were resolved by agarose gel electrophoresis, Figure a, Lane 1 – 1/30 bacterial miniprep, Lane 2 – 1/50 bacterial miniprep, Lane 3-4 – 1/30 Yeast miniprep, Lane 5 – 1/50 Yeast Miniprep, Lane 6 – AC1-2 MINPP pDEST17 control and primer set 2, Figure b, Lane 1 - PCR-1-Sample 1 undiluted, Lane 2 – PCR-1-Sample 2 undiluted, Lane 3 – PCR-1-Sample 1 1/50 dilution, Lane 4 – PCR-1-Sample 2 1/50 dilution, Lane 5 – PCR-1-Sample 3 undiluted, Lane 6 – PCR-1-Sample 3 1/50 dilution, Lane 7 – PCR-1-Sample 4, Lane 8 – 1/50 dilution AC1-2 MINPP pDEST17 control. The PCR fragment was cloned into the pDONR207 vector via the BP reaction and transformed into DH5α. Colonies were examined by colony PCR to confirm successful transformation, Figure c, Lane 1-8 DH5α from *E. coli* plates, 9-15 DH5α colonies from Yeast plates, Lane 16 – AnPhyA control. Using the LR reaction, successful colonies with the *MINPP* gene were transformed into the pDEST17 vector and examined by colony PCR, Figure d, Lane 1-2 – BP reaction Sample 2, Lane 3-4 – BP reaction Sample 6, Lane 5-7 BP reaction Sample 14, Lane 15 – AnPhyA pDEST17 control. Successful transformants were subsequently cloned into expression hosts for protein expression trials.

The pDEST17 vector containing the no-signal peptide AC1-2 *MINPP* was cloned into Rosetta™ 2 (pLysS). Additionally, as described in Section 3.1.17. it was also cloned into sHuffle® T7 Express and ArcticExpress (DE3) in a further attempt to improve the solubility. A small-scale purification was performed, Section 3.1.11, using the optimal conditions suggested by the manufacturer. ArcticExpress (DE3) was grown at 30 °C without antibiotics with expression induced at 13 °C. sHuffle® T7 Express was also grown at 30 °C with antibiotic selection before being induced at 16 °C. These were all induced at varying IPTG concentrations, 0 mM, 0.01 mM, 0.05 mM, 0.1 mM and 0.5 mM. Higher concentrations of IPTG have been shown to have a negative effect on cell growth and soluble protein yield and productivity. Therefore it is important to optimise the induction process (Larentis *et al*., 2014). Figure 3.18 shows the clarified lysate and concentrated protein from the three bacterial strains under differing concentrations of IPTG.

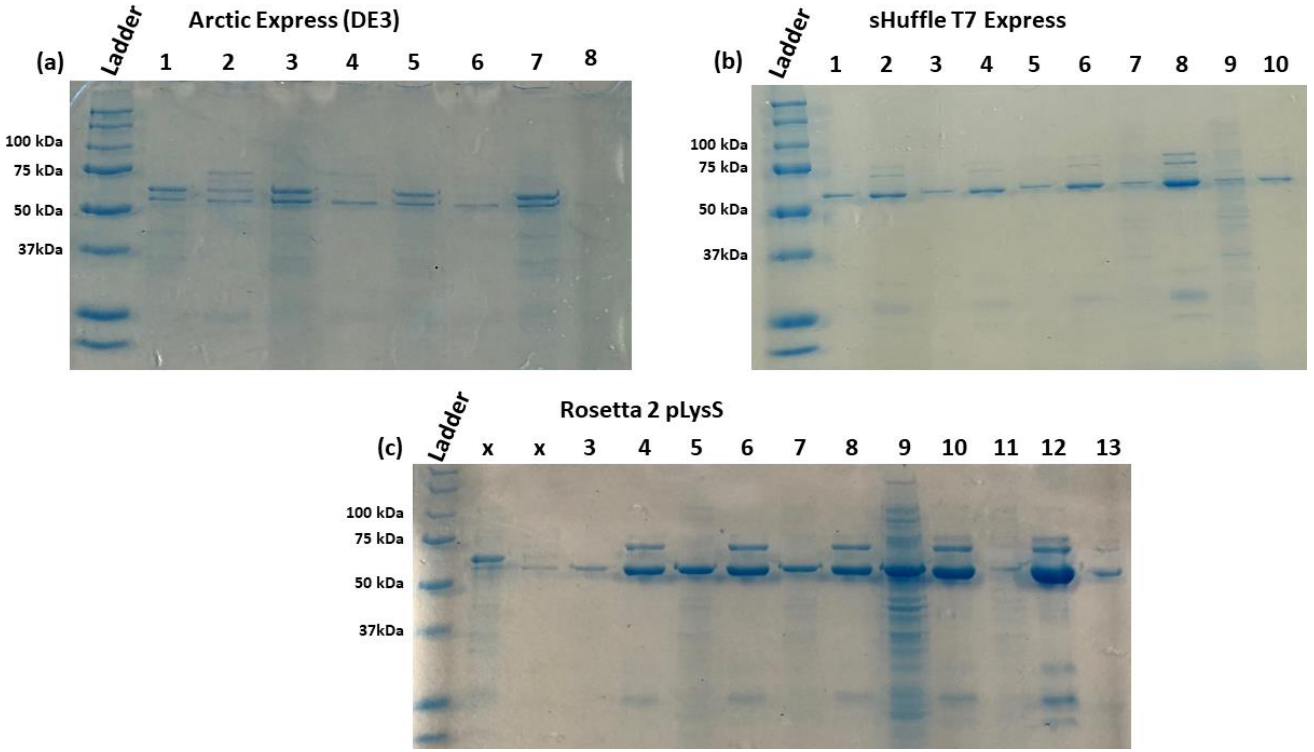
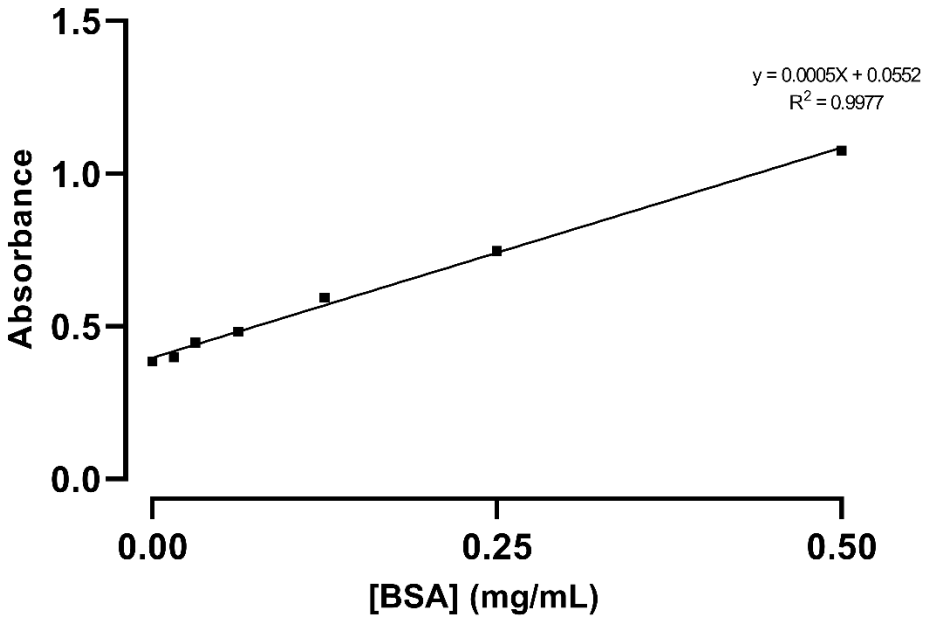


Figure 3.18 – The SDS-PAGE small-scale protein purification using ArcticExpress (DE3) (a), sHuffle T7 express (b) and Rosetta 2 pLysS (c). Figure a, Lane 1 – 0.5 mM IPTG clarified lysate, Lane 2 – 0.5 mM concentrated protein, Lane 3 – 0.1 mM IPTG clarified lysate, Lane 4 – 0.1 mM concentrated protein, Lane 5 – 0.05 mM IPTG clarified lysate, Lane 6 – 0.05 mM concentrated protein, Lane 7 – 0.01 mM IPTG clarified lysate, Lane 8 – 0.01 mM concentrated protein. sHuffle® T7 Express, Figure b; Lane 1 – 0.5 mM IPTG clarified lysate, Lane 2 – 0.5 mM concentrated protein, Lane 3 – 0.1 mM IPTG clarified lysate, Lane 4 – 0.1 mM concentrated protein, Lane 5 – 0.05 mM IPTG clarified lysate, Lane 6 – 0.05 mM concentrated protein, Lane 7 – 0.01 mM IPTG clarified lysate, Lane 8 – 0.01 mM concentrated protein, and Rosetta™ 2 (pLysS), Figure c; Lane 3 - 0.5 mM IPTG clarified lysate, Lane 4 – 0.5 mM concentrated protein, Lane 5 – 0.1 mM IPTG clarified lysate, Lane 6 – 0.1 mM concentrated protein, Lane 7 – 0.05 mM IPTG clarified lysate, Lane 8 – 0.05 mM concentrated protein, Lane 9 – 0.01 mM IPTG clarified lysate, Lane 10 – 0.01 mM concentrated protein, Lane 11 - 0 mM IPTG clarified lysate, Lane 12 – 0 mM concentrated protein, Lane 13 – N/A

In contrast to experiments with signal peptide-encoded constructs, Figure 3.12, there is clear indication in Figure 3.18 that AC1-2 *MINPP* has been expressed and can be purified using Ni-NTA agarose resin. Following this, a Bradford assay was performed to estimate the protein concentration. This was compared with a BSA standard calibration curve. The results are displayed in Figure 3.19.



	A595	mg/ml	µg/ml	µM
ArcticExpress (DE3) 0.5 mM IPTG	0.651	0.16	160	2.66
ArcticExpress (DE3) 0.1 mM IPTG	0.534	0.09	90	1.50
ArcticExpress (DE3) 0.05 mM IPTG	0.47	0.05	52	0.86
ArcticExpress (DE3) 0.01 mM IPTG	0.412	0.02	17	0.29
sHuffle® T7 Express 0.5 mM IPTG	0.66	0.17	165	2.75
sHuffle® T7 Express 0.1 mM IPTG	0.613	0.14	137	2.28
sHuffle® T7 Express 0.05 mM IPTG	0.648	0.16	158	2.63
sHuffle® T7 Express 0.01 mM IPTG	0.869	0.29	290	4.83
Rosetta™ 2 (pLysS) 0.5 mM	0.796	0.25	246	4.10
Rosetta™ 2 (pLysS) 0.1 mM	0.923	0.32	322	5.36
Rosetta™ 2 (pLysS) 0.05 mM	0.905	0.31	311	5.18
Rosetta™ 2 (pLysS) 0.01 mM	1.052	0.40	399	6.64
Rosetta™ 2 (pLysS) 0 mM	1.526	0.68	681	11.35*

Figure 3.19 – Bradford assay results for the expression of AC1-2 *MINPP* in ArcticExpress (DE3), sHuffle® T7 Express and Rosetta™ 2 (pLysS) under different IPTG concentrations. *Highlights an anomalously high result for the condition.

For the ArcticExpress cells, Figure 3.18a, the higher the IPTG concentration, the higher the concentration of expressed protein. There were two dominant bands, one, around 60 kDa which is the expected size of AC1-2 MINPP. The other, just below this band is most likely the Cpn60 protein (~57 kDa), which is constitutively expressed from the pACYCbased plasmid, and which is typically detected in protein gel analysis of ArcticExpress cell lysates. The highest yield of protein, 2.66 µM, was obtained with 0.5 mM IPTG. In the sHuffle® T7 Express cells, Figure 3.18b, the highest yield of protein was 4.83 µM, obtained with 0.01 mM IPTG. The other IPTG concentrations only had slight variability amongst themselves. The gel image indicated a higher purity of the sHuffle-expressed protein in comparison with the ArcticExpress cells. The Rosetta™ 2 (pLysS) cells, Figure 3.18c, showed considerably higher expression at each IPTG concentrations. The opposite of the ArcticExpress cells, Rosetta™ 2 (pLysS) showed higher expression at lower IPTG concentrations. The 0 mM datapoint has been omitted as an error as the pLysS should prevent any leaky expression. Therefore, the highest concentration of protein was expressed using 0.01 mM IPTG at 6.64 µM.

While the Rosetta™ 2 (plysS) expressed protein was not as pure as the sHuffle-expressed protein, there was higher expression of protein, and the secondary band was sufficiently far enough away, 3-5 kDa, that it offered opportunity for further purification of AC1-2 MINPP by size exclusion chromatography. Before any further purification, the potential phytase band was sent for sequencing using Protein Mass Fingerprinting at the John Innes Centre to confirm the expression of the AC1-2 MINPP protein.

Once confirmed, the construct in Rosetta™ 2 (pLysS) was expressed on a large scale with induction at 0.1 mM IPTG. The immobilised metal affinity chromatography (IMAC) and gel filtration (GF) chromatograms are displayed in Figure 3.20.

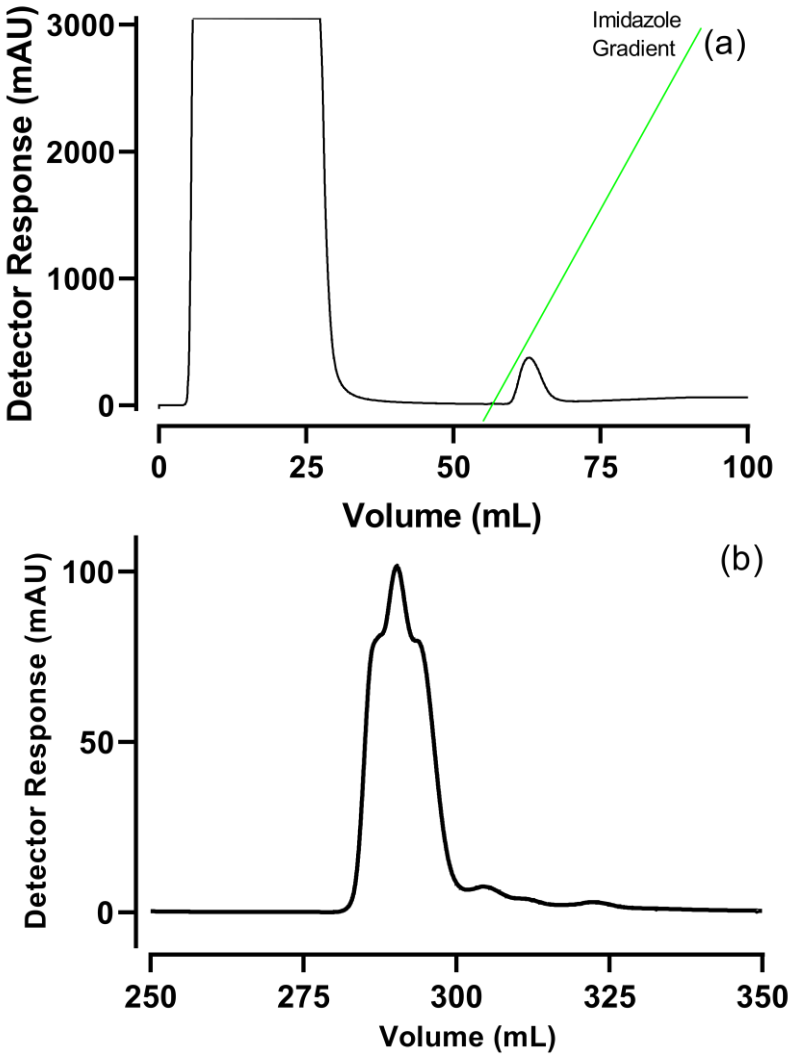


Figure 3.20 – Chromatograms of the Immobilised metal affinity chromatography (IMAC) (a) and Gel Filtration (b) of the signal peptide-lacking AC1-2 MINPP expressed using Rosetta™ 2 (pLysS) cells.

The removal of the signal peptide from AC1-2 MINPP significantly improved protein expression and subsequent purification (*cf*. Figures 3.20 and 3.13/3.14). The IMAC chromatogram, Figure 3.20, shows a significantly larger and broader peak, and the gel filtration chromatogram is also a considerably larger single peak with an absorbance nearly twice that of Figure 3.11 despite the reduced culture volume (2 litres vs 4 litres). To confirm the purity of product, Fractions 10-20 were run on an SDS-PAGE and the purity of each fraction examined, Figure 3.21

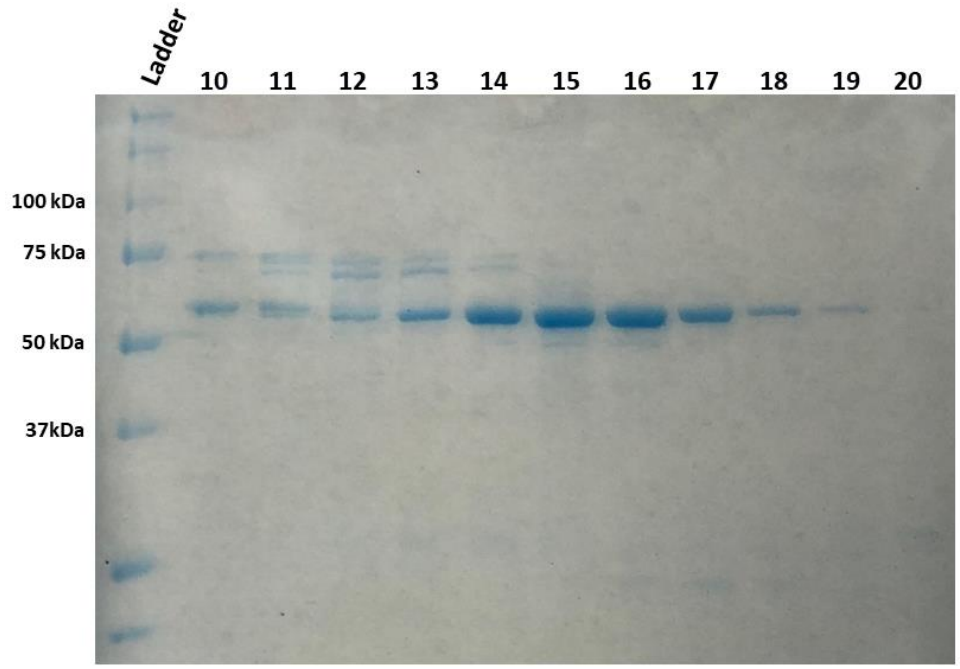


Figure 3.21 – SDS-Page gel electrophoresis to examine the purity of gel filtration fractions. Lanes 10-20 – Fractions 10-20.

SDS-PAGE showed that the large band above AC1-2 Minpp, Figure 3.15, had been significantly reduced using gel filtration. To further improve purity, Fractions 16-19 were pooled and concentrated. Samples from each stage of the purification process were taken for SDS-PAGE gel electrophoresis, Figure 3.22.

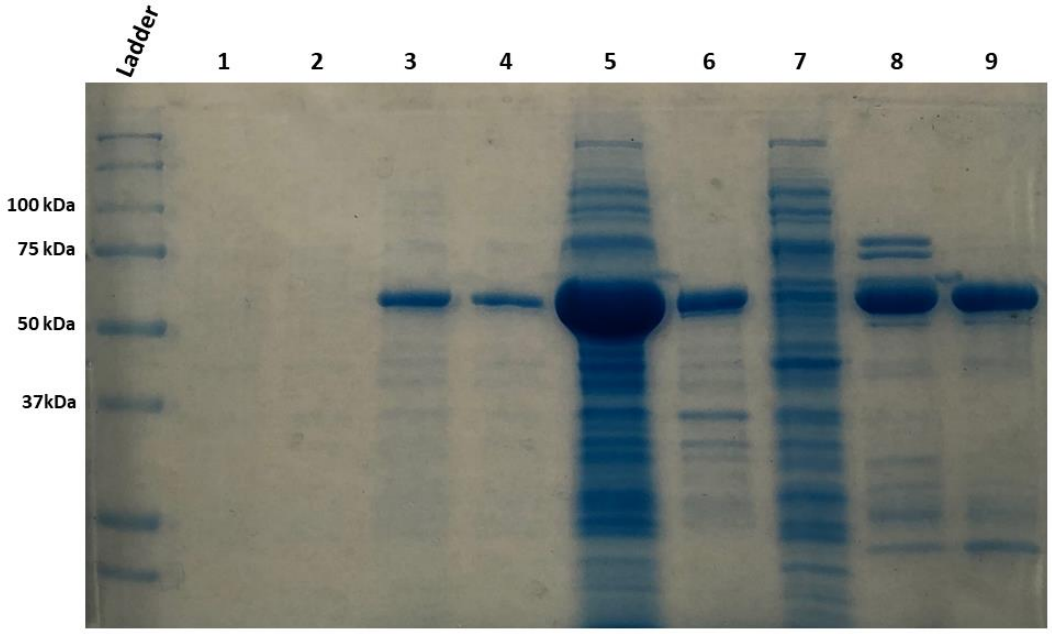
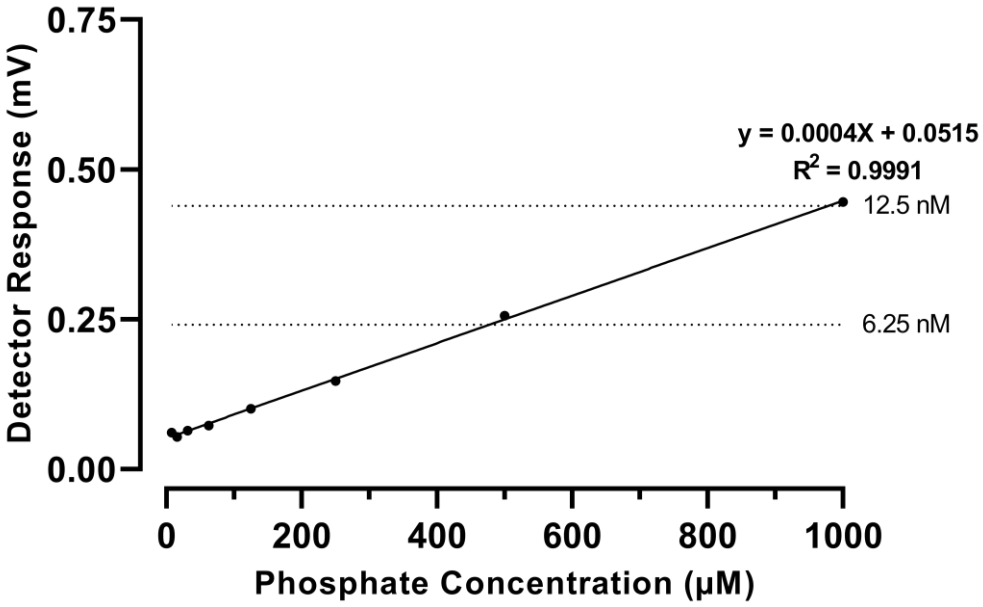


Figure 3.22 – Summary SDS-PAGE of the purification of AC1-2 MINPP without its signal peptide. Lanes 1-2 – Uninduced sample, Lanes 3-4 – Induced sample, Lane 5 – lysate, Lane 6 – pellet, Lane 7 – clarified lysate, Lane 8 – concentrated IMAC, Lane 9 – concentrated gel filtration.

Examining the concentrated gel filtration sample from Figure 3.22 and comparing with the gel filtration sample of the signal peptide MINPP protein, Figure 3.16, shows that the AC1-2 MINPP protein has a considerably higher expression rate, and has been concentrated into a pure sample with minor impurities, although some of the protein remains insoluble in the pellet fraction.

Similar to Figure 3.15, a serial dilution was performed to identify the optimal concentration of phytase to use in degradation assays and the initial activity, Figure 3.23.



AC1-2 Concentration	*Absorbance*		
100	1.896	1.827	1.871
50	1.881	1.991	1.794
25	0.91	1.162	0.981
12.5	0.439	0.583	0.488
6.25	0.241	0.298	0.261
3.125	0.143	0.165	0.161
1.5625	0.114	0.115	0.11
0.78125	0.09	0.083	0.091

Figure 3.23 – Determining the appropriate concentration of the no-signal peptide AC1-2 MINPP for phytase activity assays. A serial dilution of phytase activity was measured using the molybdenum blue method for phosphate release. This was measured at 700 nm and compared with a phosphate calibration curve.

The optimal concentration of AC1-2 MINPP (without signal peptide, hereafter called AC1-2 MINPP) for use in activity assays was determined to be 12.5 nM and the activity of the protein was determined to be 120.94 ± 5.1 U/mg. This is a 9.3-fold increase in specific activity from the previous purification of the AC1-2 Minpp with the signal peptide.

3.3. Conclusion

The foregoing experiments show how optimization of protein expression is an empirical process that needs to be determined for individual proteins and their constructs. The reasons for sub-optimal expression are varied, there may be too much inducing agent added, causing too much protein formation and development of inclusion bodies. Therefore, it may be necessary to identify the optimal concentration of inducing agent to use. Additionally, many proteins require different machinery for folding and expression that may not be available in particular expression hosts. In this work, three different expression hosts were used, Rosetta™ 2 (pLysS), sHuffle® T7 Express and ArcticExpress (DE3), each of which adding a different feature to aid in protein expression. Nevertheless, in this case, it was necessary to return to and modify the original design of the construct (Duong-Ly and Gabelli, 2014), here by removal of the signal peptide.

One of the challenges with protein purification is finding the optimal conditions that can overexpress the protein as well as producing a largely pure product. The first steps should be identifying the presence of the signal peptide, which has been shown in this chapter, can significantly interfere with the solubility of the protein. Similarly in the expression of the phytase from *Yersinia kristeensenii* the coding region of the phytase gene was cloned without the signal peptide (Fu *et al.*, 2008).

Secondly, there are many expression vectors that can be used to overexpress the phytase. For example, the β-propeller phytase from *Janthionbacterium* sp. TN115 was expressed in pET-22b(+) (Novagen) which contains a His-tag for immobilised metal chromatography, the purification method the same as in this chapter (Zhang *et al*., 2011b). The gene in this case, was cloned into the vector using restriction enzymes whereas in this chapter it was cloned using gateway cloning. It is therefore very important to select your cloning method and vector appropriately. Choosing your expression strain is also important and a variety have been used in the literature. Commonly, expression strains derived from *Escherichia coli* are used for bacterial systems, such as BL21 (DE3), Top10 and Rosetta (Acquistapace *et al*., 2020; Gu *et al*., 2009; Zhang *et al*., 2011b). Phytases are also frequently expressed in eukaryotic

systems such as *Pichia pastoris* using vectors suitable for cloning into eukaryotes such as pPIC9 (Xiong *et al*., 2005).

To summarize this chapter, the MINPP phytase from *Acinetobacter* sp. AC1-2 was successfully cloned, expressed and purified, with expression trials performed to produce the highest purity protein.

Chapter 4. The Characterisation of a Multiple Inositol Polyphosphate Phosphatase from *Acinetobacter* sp. AC1-2

The phytase from *Acinetobacter* sp. AC1-2 was successfully expressed in Rosetta™ 2 (pLyss) cells using the gateway vector pDEST17, producing a pure protein that was active towards phytate. Phytase specific activity is typically expressed as U/mg in the literature to allow for a direct comparison between different phytases. This is defined as the amount of enzyme required to release 1 µM per minute under the assay conditions (Kim *et al*., 2003; Wu *et al*., 2014).

Due to the widespread nature of natural phytases, their properties vary considerable in activity, stability, temperature and pH optima. However, there are currently no single phytase that is ideal for commercial applications. An ideal phytase is one that is catalytically efficient, resistant to proteolysis by the animal's digestive enzymes, has a high thermostability as the pelleting process requires the enzyme to withstand temperatures up to 80°C and not lose significant activity, as well as being stable in storage. The phytase should also be easy and cheap to manufacture (Yao *et al*., 2012).

The aim of this Chapter was to perform a detailed characterisation of the phytase, investigating the features that would be of industrial interest.

4.1. Materials and Methods

4.1.1. Phytase Activity Assays

Phytase specific activity (U/mg) was determined using the molybdenum blue method for phosphate release (Ariza *et al*., 2013; Crouch and Malmstadt, 1967). One phytase unit was defined as the amount of the enzyme that released 1 µM inorganic phosphate per minute under the assay conditions. All samples were assayed in triplicate (Dokuzparmak *et al*., 2017).

These reactions were performed from a working stock of 4 µM AC1-2 MINPP in 25 % w/v trehalose and gel filtration buffer unless stated otherwise. In brief, 8.75 µL of 250 nM AC1-2 MINPP was diluted in 148.75 µL of 0.2 M Na-Acetate pH 5 buffer leaving a final concentration of 12.5 nM. 45 µL of the enzyme-buffer solution was mixed with 5 µL of 50 mM $InsP_6$ in triplicate on ice before being heated at 37 °C for 15 minutes in a PCR machine. The reaction was stopped by addition of 50 µL of a 4:1 ratio of ammonium molybdate sulphuric acid solution, prepared by mixing solutions of ammonium molybdate (6 g $NH_4Mo_7O_{24}.4H_2O$ and 22 mL 98 % H_2SO_4 in 400 mL) with ferrous sulphate solution (2.16 g iron (II) sulphate heptahydrate, 2 drops 98% H_2SO_4 in 20mL). Absorbance was measured at 700 nm after 15 minutes. Activity was determined by comparing the absorbance values against a calibration curve starting from 1 mM NaH_2PO_4. AC1-2 MINPP was used at a final concentration of 12.5 nM unless stated otherwise.

4.1.2. pH Profile

The pH profile of AC1-2 MINPP was measured in 0.2M buffer: glycine HCl, pH 2-3.5; sodium acetate, pH 4-5.5; Bis-Tris, pH 6-7; Tris HCl, pH 8-8.5. For these, 8.75 µL of 250 nM AC1-2 was diluted in 148.75 µL of the appropriate buffer.

4.1.3. Time course

For analysis of the inositol phosphate products of AC1-2 MINPP action on $InsP_6$, assays were performed in 0.2 M Na-Acetate pH 5 working from a 4 µM AC1-2 MINPP in 25 % trehalose + 0.5 mg/mL BSA stock. $InsP_6$ was incubated with 5 mM $InsP_6$ for 30 min, 2 h, 4 h and 8 h, at which point, the reaction was heated to 95 °C for 5 mins to inactivate the enzyme.

4.1.4. Inhibition

The effect of metal ions on enzyme activity was investigated with or without the addition of 1 mM K^+, Mn^{2+}, Cu^{2+}, Co^{2+}, Mg^{2+}, Ca^{2+}, Zn^{2+}, $Fe^{2+/3+}$ or IS_6, the substrate analogue of $InsP_6$,

directly to the reaction mixture. For the reaction, 8.75 µL of 250 nM AC1-2 was diluted in 145.25µL 0.2 M Na-acetate pH 5.

4.1.5. Substrate specificity towards phosphate esters

The substrate specificity of the enzyme was investigated using various phosphate-containing compounds at 5 mM concentrations. ATP, glycerol 3-phosphate, glucose 6-phosphate, pyrophosphate, para-nitrophenyl phosphate (pNPP), phosphoenolpyruvic acid (PEP) and creatine phosphate were compared beside $InsP_6$. The assays were performed as described in Section 4.1.1.

4.1.6. Thermostability

The short-term thermostability of AC1-2 MINPP was measured by firstly incubating the enzyme-buffer mixture at 4, 37, 50, 60 and 70 °C for 10 minutes. Following this, assays were performed as described above at 37 °C.

4.1.7. Long-term stability assay

Two solutions of an impure AC1-2 MINPP protein, with signal peptide, were stored at either 4°C or at room temperature at 4 µM concentration in in 25% (w/v) trehalose. The samples were assayed for activity at intervals thereafter at a final concentration of 50 nM.

4.1.8. Effect of storage buffer on long-term stability

The stability of AC1-2 MINPP was measured over a long period of time. The protein was stored in 4 µM aliquots in different cryoprotectants, 50% (w/v) trehalose, 50% (w/v) trehalose and 1 mg/mL BSA, 50% (w/v) sucrose, 50% (w/v) sucrose and 1 mg/mL BSA, 50% (w/v) glycerol, 50% (w/v) glycerol and 1 mg/mL BSA, 1 mg/mL BSA and gel filtration buffer at room temperature. The microfuge tubes in which the samples were stored were wrapped with parafilm to prevent evaporation. On occasions thereafter, aliquots were tested for activity.

4.1.9. Michaelis-Menten parameters of AC1-2 MINPP

AC1-2 MINPP was assayed over 12 substrate concentrations, 3750, 2500, 1750, 1250, 625, 350, 175, 125, 62.5, 35, 17.5, and 12.5 µM. The progress of reaction curve was fitted to a non-linear regression model for substrate inhibition using GraphPad Prism 8.0.1.

4.1.10. Structural Biology

Homology modelling of the structure of AC1-2 MINPP was performed using SWISS-MODEL (Waterhouse *et al*., 2018). Two models of the enzyme, with the signal peptide removed as predicted by SignalP 5.0 (Almagro Armenteros *et al*., 2019) were produced based on X-ray crystallography structures of the highest scoring sequence homologue, the MINPP from *Bifidobacterium longum* (*Bl*Minpp) with which AC1-2 MINPP shares 34% sequence identity. The first model was based on the structure of apo-*Bl*Minpp (PDB entry 6RXD) (Acquistapace *et al*., 2020) and the other structure of the enzyme complexed with the non-hydrolysable substrate analogue inhibitor, *myo*-inositol hexasulfate (IS_6) (PDB entry 6RXE). The Global Model Quality Estimate scores of the two models were 0.60 and 0.61, respectively. Residues forming the specificity pocket of AC1-2 MINPP were inferred from the predicted structure and its complex with IS_6. The corresponding residues in specificity pockets of the closest AC1-2 MINPP homologues for which crystal structures are available were identified through analysis of MINPPs from *Bifidobacterium longum* (sequence identity 34%; PDB 6RXE) and *Bacteroides thetaiotaomicron* (*Bt*MINPP) (sequence identity 21%; PDB 4FDU) (Stentz *et al*., 2014), and from the extracellular histidine acid phytase from *Aspergillus fumigatus* (sequence identity 15%; PDB 1SK8) (Liu *et al*., 2004).

4.2. Results and Discussion

4.2.1. Characterisation of AC1-2 MINPP

As described in Chapter 3, a useful working concentration of AC1-2 MINPP phytase was determined to be 12.5 nM and a specific activity 120.94 ± 5.1 U/mg was determined. The reaction used a stock solution of 0.2 M Na-Acetate pH 5.5 for this measurement. However, phytases can act over a wide range of differing pH's with pH optima varying from acidic to alkaline therefore it is important to identify the optimal pH for characterisation assays. The activity of AC1-2 MINPP in the range pH 2-8.5 was determined, Figure 4.1.

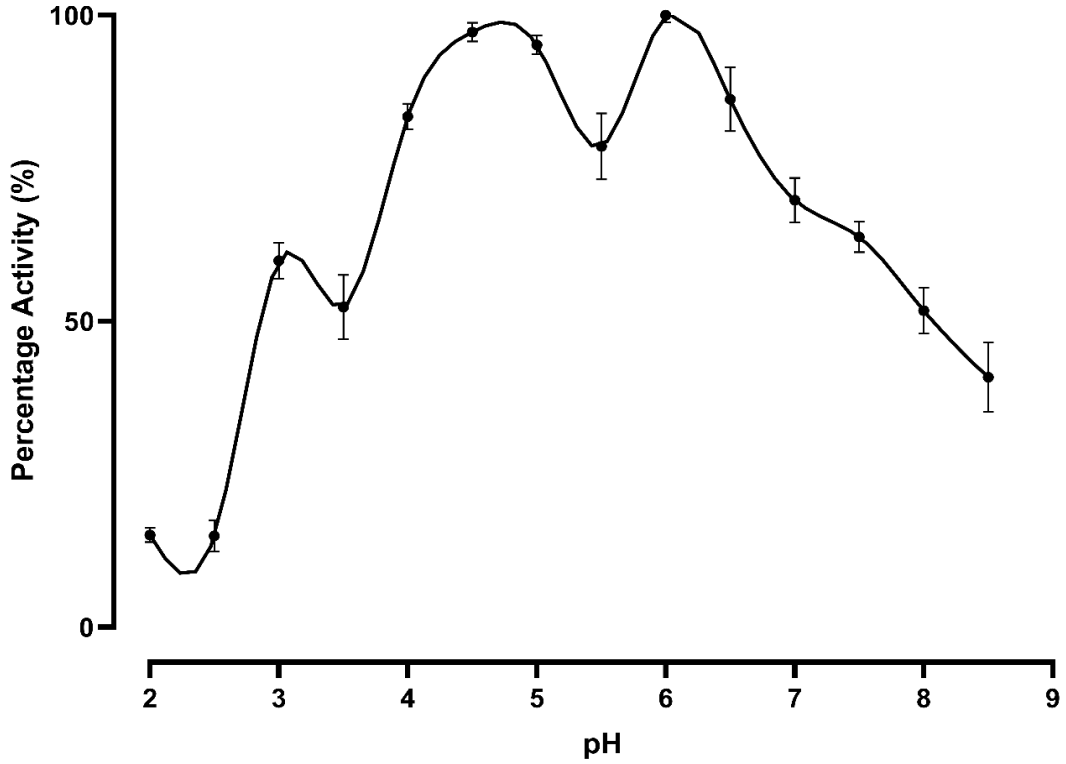


Figure 4.1 – The pH profile, 2-8.5, of phytase activity for the AC1-2 MINPP protein. N=3.

AC1-2 MINPP showed activity in the range pH 2-8.5 tested with three peaks at pH 3, 4.5-5 and 6, with at least 70% of maximum within the pH range 4-7. Here AC1-2 MINPP differs from many phytases which typically exhibit one or two maxima, Table 4.1 (Konietzny and Greiner, 2002; Pandey *et al*., 2001). The exception is the MINPP phytase from *Bacteroides thetaiotaomicron*. The small majority of current MINPP studies describe a phytase that is commonly found in eukaryotic organisms (Cho *et al*., 2006; Dionisio *et al*., 2007) and has also been described in gut commensal bacteria (Haros *et al*., 2009; Stentz *et al*., 2014). It may be the unique properties of these enteric environments which allow the soil-derived MINPP to function over a wide range of pH, similar to the pH range in the digestive tract (Merchant *et al*., 2011b). This experiment indicated that use of 0.2 M Na-Acetate pH 5.5 underestimates the maximal activity of the phytase and buffers at a pH of 4.5-5 or 6 would be more appropriate.

Table 4.1 – The pH optima of 20 different naturally occurring phytases from Bacteria, Fungal and Plants.

Species	pH Optima	pH Range	Phytase Type	Reference
Escherichia coli	4.5	1.5-7.5	HAPhy	(Greiner *et al*., 1993)
Citrobacter braakii	4	2-8.5	HAPhy	(Kim *et al*., 2003)
***Buttiauxella* sp. GC21**	4.5	2-7	HAPhy	(Shi *et al*., 2008)
***Shingella* sp. CD2**	5.5	2.5-8.5	HAPhy	(Pal Roy *et al*., 2016)
Aspergillus niger	2.5, 5.5	2-7	HAPhy	(Nagashima *et al*., 1999)
Aspergillus fumigatus	4, 6.0-6.5	2.5-8	HAPhy	(Pasamontes *et al*., 1997)
***Penicillium* sp.**	5.5	3-8	HAPhy	(Zhao *et al*., 2010)
Peniophora lycii	5	4-8	HAPhy	(Ullah and Sethumadhavan, 2003)
Hafnia alvei	4.5	2-7.5	HAPhy	(Ariza *et al*., 2013)
Oryza sativa L.	3.5	2.5-6.5	HAPhy	(Li, Ruijuan *et al*., 2011)
Bacillus subtilis	7.0-7.5	3.5-8.5	BPPhy	(Oh *et al*., 2004)
Bacillus amloliquefaciens	7.0-7.5	4-10	BPPhy	(Oh *et al*., 2004)
***Janthinobacterium* sp. TN115**	8.5	5-10	BPPhy	(Zhang *et al*., 2011b)
***Streptomyces* sp. US42**	7.0	5.5-10	BPPhy	(Boukhris *et al*., 2016b)
Bacteroides thetaiotaomicron	2.5, 4.0, 7.5	1.5-8	Minpp	(Stentz *et al*., 2014)
Triticum aestivum L.	4.5	2-7.5	Minpp	(Dionisio *et al*., 2007)
Hordeum vulgare L.	4.5	2-7.5	Minpp	(Dionisio *et al*., 2007)
Arabidopsis thaliana	4.5, 5.5	3-7.5	PAPhy	(Kuang *et al*., 2009)
Nicotiana tabacum	5-5.5	3.5-7	PAPhy	(Lung *et al*., 2008)
Stylosanthes guianensis	5	3-8	PAPhy	(Liu *et al*., 2018)
Bdellovibrio bacteriovorus	4	3-7	PTPhy	(Gruninger *et al*., 2014)
Selenomonas ruminantium	5	3.5-7	PTPhy	(Puhl *et al*., 2007)
Megasphaera elsdenii	5	2.5-7	PTPhy	(Puhl *et al*., 2009)

Following this, an experiment was performed to determine the profile of inositol phosphate products generated from $InsP_6$, with the additional purpose of confirming that the phytate degradation profile was the same as that of the bacterial phytase shown in Figure 2.7. To this end, HPLC analysis was undertaken on products of reaction terminated at 30 min, 2 h, 4 h, or 8, Figure 4.2.

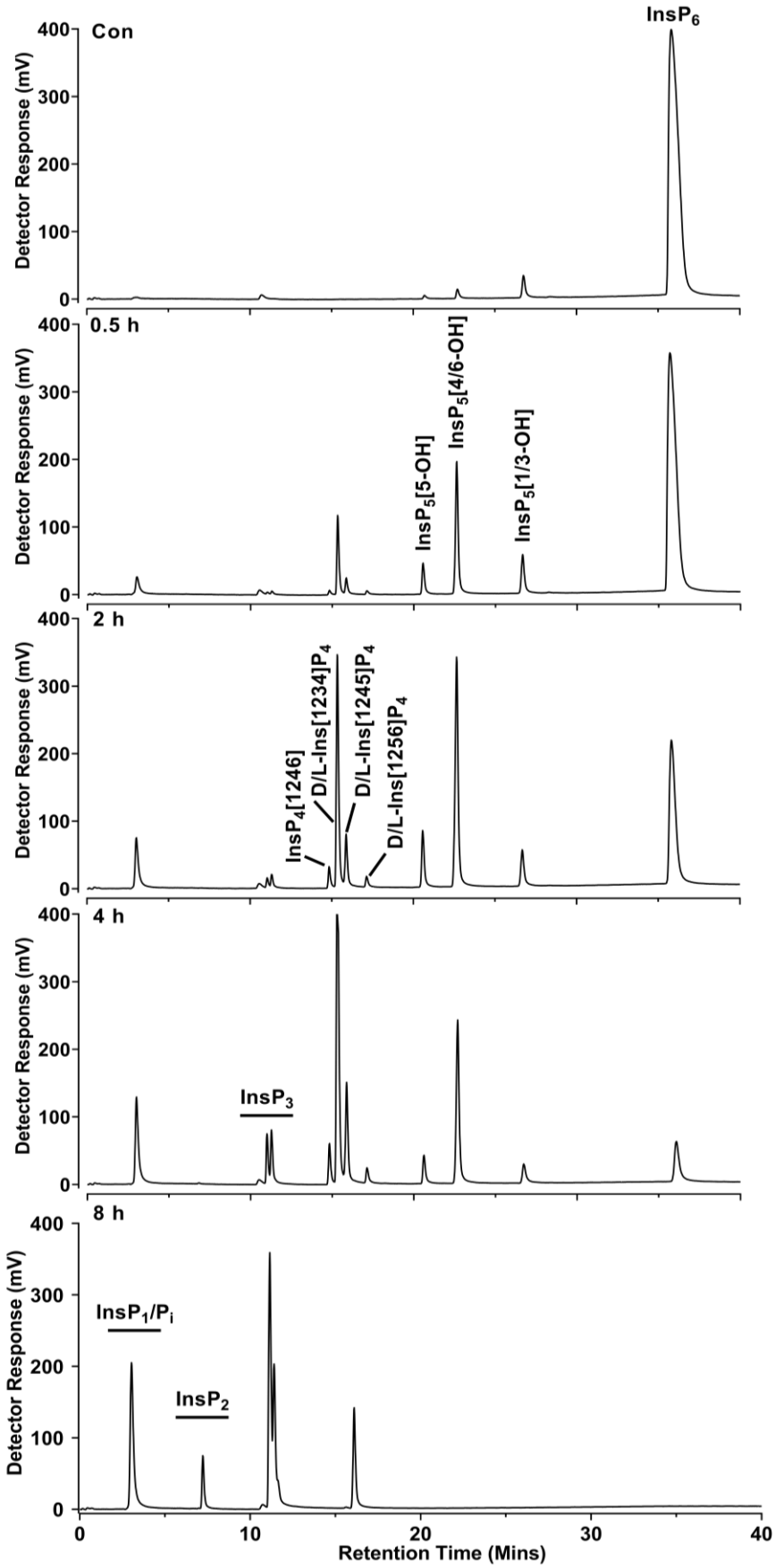


Figure 4.2 – The time-course degradation profile of 1 mM phytate by the pure AC1-2 MINPP enzyme. Degradation was measured at 0.5-, 2-, 4- and 8-hour time points

HPLC analysis of the phytate degradation profile of the AC1-2 MINPP protein shows that the pattern is identical to that of *Acinetobacter* sp. AC1-2, Figure 2.7. Following from this, the peak areas of the data shown in Figure 4.2 were integrated and displayed to show the change from one inositol phosphate to the other, Figure 4.3.

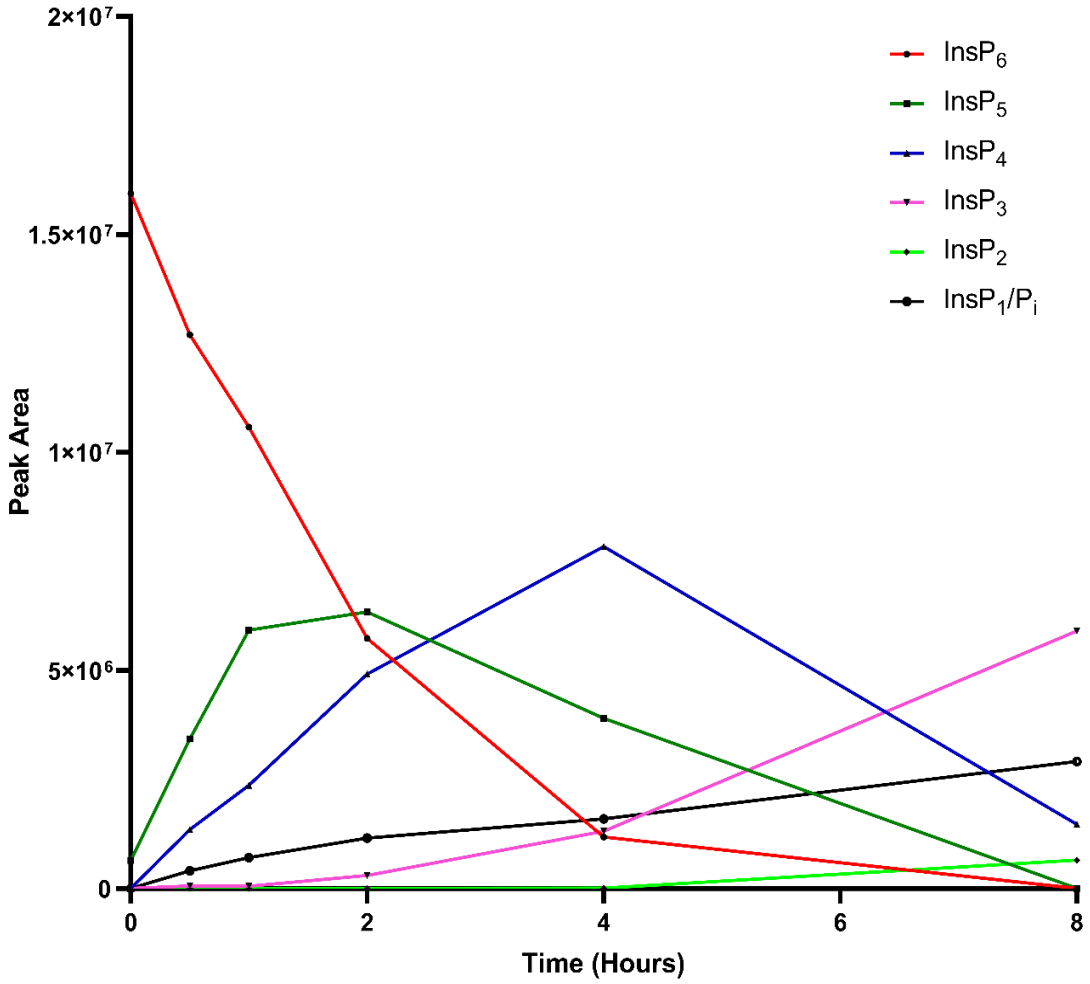


Figure 4.3 – The depletion and accumulation of the inositol phosphates over time by recombinant AC1-2 MINPP.

Here we analyse the products of phytate degradation by recombinant AC1-2 MINPP. Figure 4.2-4.3 shows sequential degradation of phytate over a period of 8h with 12.5 nM protein assayed at pH 5. At early stages of degradation three peaks of $InsP_5$ were detected with a predominance of $InsP_5$(4/6-OH) and near equal amounts of $InsP_5$(1/3-OH), most of which were initially present as impurities but there was a detectable increase from the control to 30-minute sample, and $InsP_5$(5-OH). The absence of $InsP_5$(2-OH) among products indicates that the enzyme does not attack the single axial-orientated phosphate on the 2-position. The generation of multiple $InsP_5$s is typical of the commensal bacterial MINPPs characterized to date (Acquistapace *et al*., 2020; Haros *et al*., 2009; Stentz *et al*., 2014; Tamayo-Ramos *et al*.,

2012). The co-production of $InsP_4$ with $InsP_5$ at 30 min, before the peak of accumulation of $InsP_5$, suggests that, even in the presence of excess $InsP_6$, $InsP_5$s are better substrates than $InsP_6$. Multiple peaks of $InsP_4$ co-exist in the degradation products with $InsP_5$ and $InsP_3$ products, until the $InsP_5$s are wholly consumed at which point a single peak of Ins(1,2,4,6)P_4 was observed. $InsP_4$s are particularly well resolved on the CarboPac PA200 column (Madsen *et al*., 2019) allowing identification of the major and minor routes of initial and subsequent dephosphorylations. There were 4 $InsP_4$ products, the most dominant of which was D/L-Ins(1234)P_4 followed by Ins(1,2,4,5)P_4, D/L-Ins(1246)P_4 and D/L-Ins(1256)P_4.

$InsP_3$, of which there are 20 possible stereoisomers is less well resolved on this column, but it is evident that as $InsP_4$ degradation proceeds to exhaustion – one of the two resolvable peaks (which could contain multiple isomers) predominates, concurrent with the appearance of a single peak of $InsP_2$. While there are fifteen possible stereoisomers of $InsP_2$, we may reasonably assume that the peak is comprised of species retaining the 2-phosphate. Similarly, we may assume that the monophosphate product is Ins2P, which along with other monophosphates co-elutes with Pi at the solvent front of this HPLC system – accounting for the progressive accumulation of Pi over the course of the assay.

As discussed earlier in this thesis, phytate is often regarded as an antinutritional factor for humans and animals as it may chelate nutritionally important cations, reducing their absorption. The histidine acid phosphatase class typically hydrolyses metal-free phytate at an acidic/neutral pH range, therefore the conjugation of the metal ion and substrate can cause reduced activity and inhibition of the phytase (Oh *et al*., 2004), although conversely, phytases have also been reported to increase the bioavailability of some metal ions such as zinc and copper (Revy *et al*., 2004).

In the report by Santos *et al*., (Santos *et al*., 2015) they examined the inhibition of three HAPhys from *Aspergillus niger*, *Escherichia coli* and *P. lycii* using Fe, Zn, Cu and Mn ions. What became apparent was that each of the tested HAPhys displayed different responses to the same inhibitor despite being from the same class. Fe was universally a strong inhibitor of phytase activity whilst Zn severely repressed *E. coli* and *P. lycii* and whilst it also inhibited *A. niger* the inhibition was considerably less. Cu was a potent inhibitor of *A. niger* and *E. coli*, and only moderately inhibited *Peniophora. lycii*. Mn had the least effect, only slightly inhibiting *A. niger*. Despite any differences between the effects each inhibitor has on the HAPhys, the three commonly potent inhibitory molecules throughout the literature are Zn, Fe and Cu (Greiner *et al*., 1997; Huang *et al*., 2006; Wyss *et al*., 1999; Yoon *et al*., 1996).

The next characterisation assay performed was an inhibition experiment using an array of metal ions and the substrate analogue IS_6. The results are displayed in Figure 4.4.

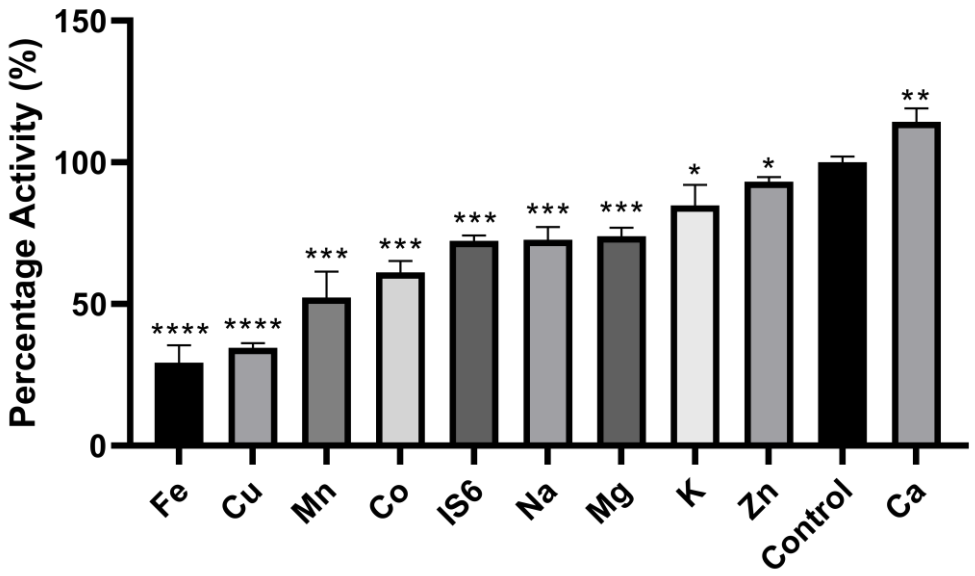


Figure 4.4 – The inhibition of AC1-2 Minpp using 1 mM concentrations of potential inhibitors. Phytase activity is described as a percentage relative to the activity of the control sample. N=3.

Every agent tested in this experiment had a statistically significant effect on the activity of AC1-2 MINPP. Both $Fe^{2+/3+}$ and Cu^{2+} showed the largest inhibitory effect, reducing activity to below 50%, which is in agreement with the literature discussed above. One of the more interesting features was the lack of strong inhibition from the Zn^{2+} ion. Typically, zinc is referenced as one of the most potent phytase inhibitors (Wang, Xueying *et al*., 2004; Dionisio, Holm and Brinch-Pedersen, 2007), however in this instance, activity was only reduced to 93.2%. Furthermore the substrate analogue IS_6, also regarded as a potent inhibitor only reduced activity to 72.4% (Ullah *et al*., 2000). Another interesting result was the enhancement of activity by the addition of calcium, Ca^{2+}, increasing activity to 114.3% which is commonly only associated with Beta-propeller phytases (Cheng and Lim, 2006). However, calcium activation of HAPhy is not unprecedented. The MINPP from *Bifidobacterium longum* which shares 34% sequence identity with AC1-2 MINPP saw calcium activation at low concentrations, 2 mM (Tamayo-Ramos *et al*., 2012).

Within the HAPhy classification there are two subgroups, those that have a broad substrate specificity but a low specific activity and conversely, those that have a narrow substrate specificity and a high specific activity towards phytate (Böhm *et al*., 2010). For example, the *Aspergillus niger* phytase (PhyB) shows over 1000% activities for *para*-nitrophenyl phosphate (pNPP), glucose-6-phosphate and fructose 1,6-biphosphate and over 100% activity for many

other phosphate-esters, when compared with sodium phytate as 100% activity, whereas *Escherichia coli* showed strict substrate specificity towards Na-InsP_6 activity with the next highest activity with pNPP at 12.3% (Oh *et al*., 2004).

In this experiment an array of different phosphate esters was tested for phosphatase activity and compared with phytate as 100% activity. The results are displayed in Figure 4.5.

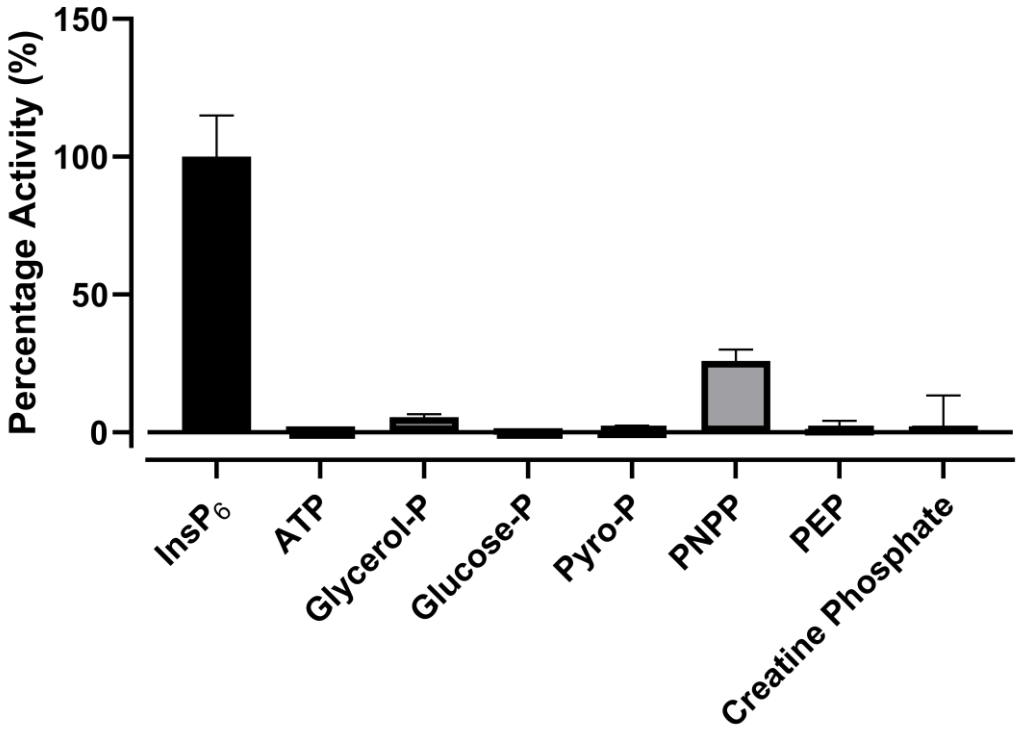


Figure 4.5 – The specificity of AC1-2 MINPP towards different P-containing compounds. InsP_6 is defined as 100% of activity. N=3.

Inspection of the results show that AC1-2 MINPP is in the latter group, exhibiting a narrow substrate specificity with low activity towards other phosphorylated substrates such as p-NPP, 25.8% and glycerol 3-phosphate, 5.4%, as compared to InsP_6. This narrow specificity is in agreement with previous analyses of wheat, barley and avian MINPPs (Cho *et al*., 2006; Dionisio *et al*., 2007).

One of the most sought-after features of a phytase is the ability to withstand high temperatures and remain active. Contamination is a significant hazard during the production of animal feeds, harmful bacteria such as *Salmonella spp.* which can cause animal contamination (Jones, 2011), as such many countries require, by law, the control of *Salmonella* using heat treatments (Jones and Richardson, 2004). Additionally, feeds are typically pelleted to reduce contamination, increase digestibility, and make it easier to

consume by the livestock. This process involves steaming the pellets which could reach temperatures up to 90°C (Abdollahi *et al*., 2013). Many microbial phytases however, are unable to withstand such temperatures and lose significant activity, which reduces the effectiveness of the phytase in the animal feed (Mrudula Vasudevan *et al*., 2019). This can be bypassed by the application of liquid phytase formulations after pelleting to minimise activity loss from heating, however, this process is time consuming and costly (Mrudula Vasudevan *et al*., 2019).

Therefore, many of the current commercial phytases have undergone significant engineering as discussed in Section 1.4. The native *E. coli* phytase which is the basis for many of the second generation phytase (Lei *et al*., 2013), only retains 24.4% activity after undergoing pelleting at 80 °C (Simon and Igbasan, 2002). Even the more thermostable fungal phytases *Aspergillus fumigatus* and *Aspergillus niger* only recovered 51 and 30% activity respectively after heating at 85°C (Wyss *et al*., 1998), highlighting the importance of the protein engineering and the development of new techniques. There has also been considerable attention drawn to identifying intrinsically thermostable phytases from thermophiles as a new source of commercial-based phytases.

In the next experiment, AC1-2 MINPP was incubated at increasing temperatures before the activity was assayed at 37 °C. The results are displayed in Figure 4.6.

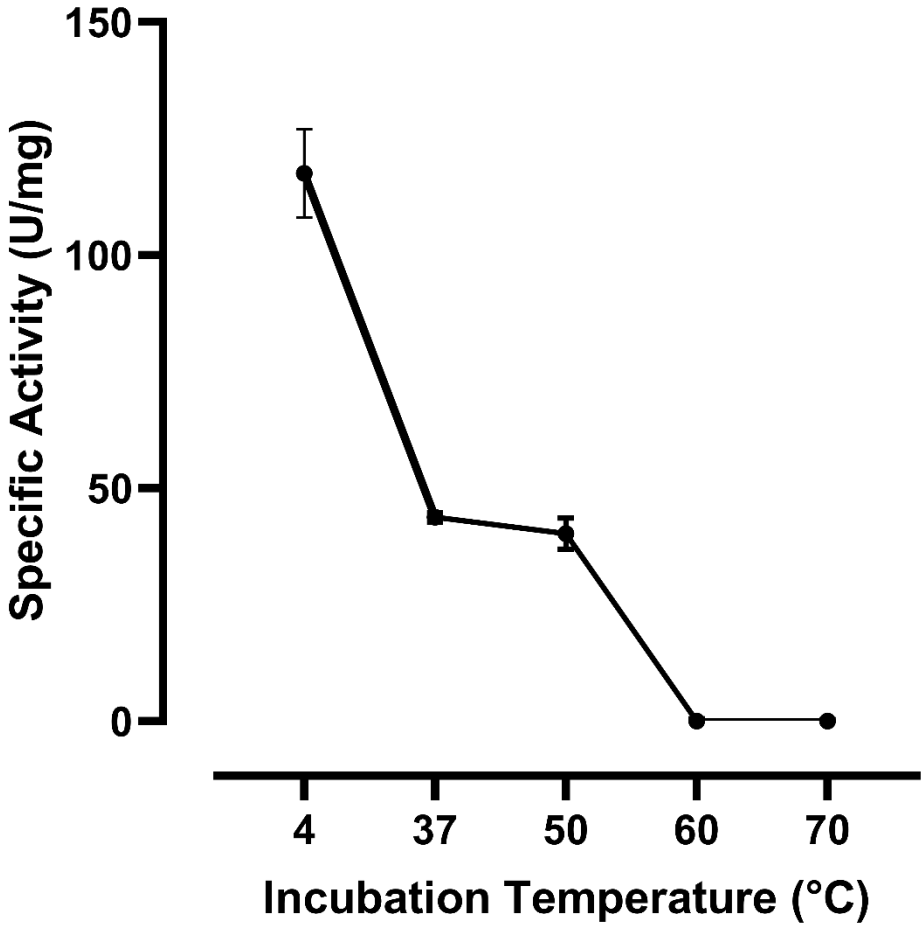


Figure 4.6 – The thermostability of AC1-2 MINPP after incubation at 4, 37, 50, 60 and 70°C for 10 minutes. N=3.

The results from Figure 4.6 identify AC1-2 MINPP as a poorly thermotolerant phytase with activity all but diminished after heating at 60 °C. It also shows a significant reduction in activity after heating at 37 and 50 °C. While we may assume that instability at higher temperatures reflects protein unfolding and denaturation, this is not tested formally here. Methods such as differential scanning calorimetry as employed (Lehmann *et al*., 2000; Schoene *et al*., 2016) are most informative.

Another important feature of a phytase is stability. Ideally, the phytase should have a long 'shelf-life' retaining its activity during storage and application, this is particularly pertaining to liquid enzyme formulations which are added post-pelleting (Haefner *et al*., 2005).

Therefore, in this next experiment, a long-term stability experiment was performed. Firstly, using a sample of the impure, signal peptide-containing protein which was stored at either room temperature or at 4 °C. This experiment had a duration of 903 days with samples taken at random timepoints for analysis of enzyme activity, Figure 4.7.

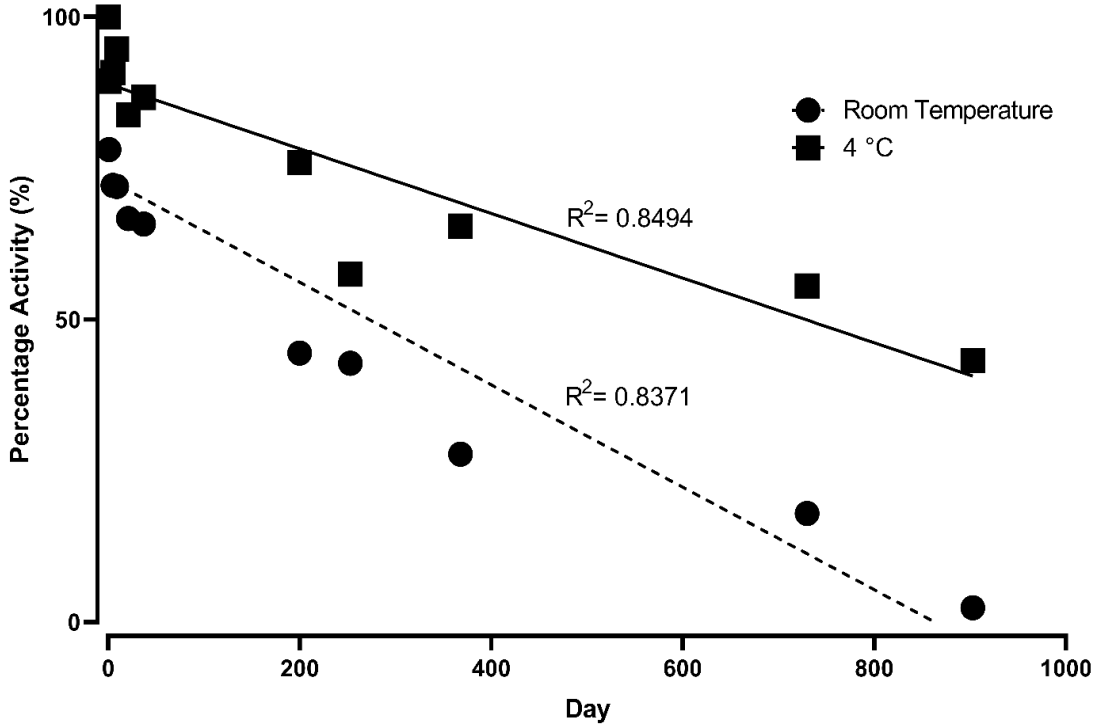


Figure 4.7 – The long-term stability of the impure AC1-2 MINPP after storage at room temperature and at 4 °C

Following this experiment, another long-term stability experiment was performed using the signal peptide-cleaved protein (AC1-2 MINPP) in a range of different storage buffers. These storage buffers contained the osmolytes trehalose, sucrose and glycerol, sugars and polyols, as well as bovine serum albumin (BSA) as a stabilising agent. Osmolytes are small organic compounds that have been shown to be used by bacterial cells in response to osmotic stress from the external environment. These can include resistance against salinity, temperature increases and freeze-thaw treatments (Dandapath *et al*., 2017). They have also been shown to protect intracellular proteins against denaturing environmental stress. While the mechanisms are not fully known, Bolen and Baskakov (Bolen and Baskakov, 2001) proposed a mechanism called the osmophobic effect, in which an unfavourable interaction between the peptide backbone and the osmolyte shifts the equilibrium in favour of the native state, the interacting forces causing the protein to fold up more compact, reducing exposure of the peptide backbone (Yancey, 2005). BSA is an extremely common laboratory protein, with numerous biochemical applications, the most common of which is used to determine the concentration of other proteins via the Bradford assay. In this experiment, BSA acts as a crowding agent. Proteins have evolved to function inside the crowded conditions of a cell, with proteins taking up as much as 20-30% of the total volume (Wang *et al*., 2012). This 'crowding' helps to stabilise proteins in their folded state in what is often explained in terms of excluded volume, the volume inaccessible to proteins due to their interaction with protein

crowders (Minton, 2000). This makes the folded state of the protein more energetically favourable than the denatured, more entropic state (Kim and Stites, 2008).

In the following experiment, 8 µM stocks of AC1-2 MINPP were mixed in equal volumes with 50% (w/v) trehalose, sucrose, and glycerol with and without 1 mg/mL BSA. These were left at room temperature and the activity measured over the course of 755 days. In the experimental data there was a sharp increase in activity on Day 24 in comparison to Day 0 and Day 3, therefore the two data points were removed from Figure 4.8.

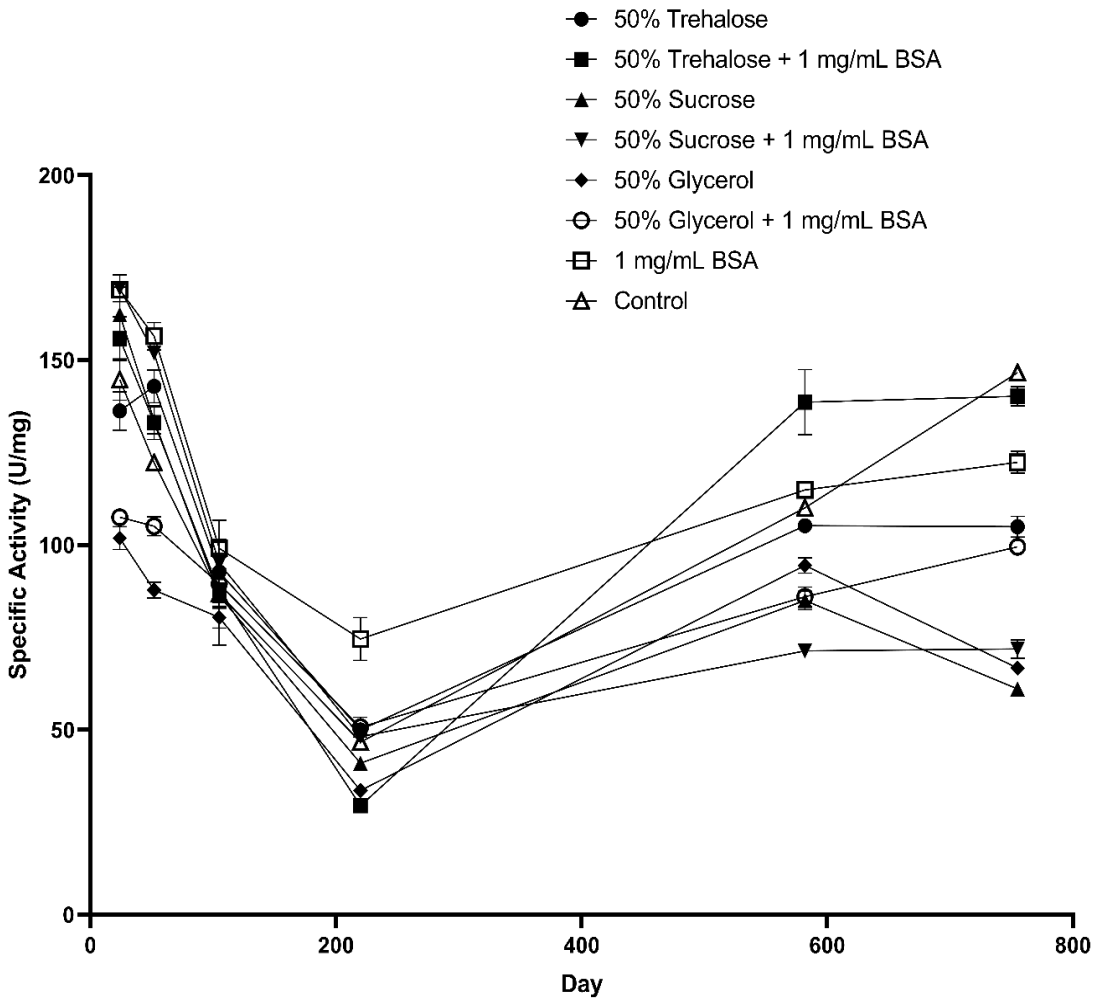


Figure 4.8 - The long-term stability of the pure AC1-2 MINPP in different storage buffers after storage at room temperature. These were measured at day 24, 52, 105, 220, 582 and 755, Day 0 and Day 3 datapoints were omitted. N=3.

Over the first 220 days of testing, a significant decrease in phytase activity was observed, with a loss of activity between 81-52.8%. However, between day 220 and 582 we see that activity had been regained. The trehalose + BSA sample saw a return from 18.9% of initial activity back up to 89%. On Day 755, activity had only slightly diminished for sucrose and glycerol and remained fairly consistent for sucrose + BSA, trehalose and trehalose + BSA and

had increased for the control sample, AC1-2 Minpp without additives, and the glycerol + BSA samples. A comparison between the activities at Day 24 and Day 755 are shown in Table 4.2.

Table 4.2 – The change in AC1-2 MINPP's activity after being stored at room temperature for 755 days under different storage conditions.

Condition	Day 24 Specific Activity (U/mg)	Day 755 Specific Activity (U/mg)	Percentage Change (%)
50% trehalose	136.3	105	77.0
50% trehalose + 1 mg/mL BSA	155.8	140.2	90.0
50% sucrose	162.4	61.1	37.6
50% sucrose + 1 mg/mL BSA	169.4	71.9	45.4
50% glycerol	101.8	66.7	65.5
50% glycerol + 1 mg/mL BSA	107.5	99.5	92.6
1 mg/mL BSA	169.1	122.4	72.4
Control	144.8	146.6	101.3

In each case, the addition of BSA as a stabilising agent saw an increase in stability in comparison with just the addition of the osmolyte. Interestingly, the control sample which saw no stabilising agent and was in a solution of 50 mM Tris-HCl pH 7.5 and 300 mM NaCl retained all of its activity and even increased by 1.3% after more than 731 days sitting at room temperature. Interesting, during the course of the stability measurements, this included the summer months where temperatures in the lab reached 30-35 °C for several days, which somewhat contrasts with the low thermostability indicated in Figure 4.6 which shows a significant decrease in activity after heating at 37 °C. This may indicate that while AC1-2 MINPP is denatured after abrupt temperature changes, activity can be regained. One potential issue of this experiment may be that the protein has concentrated in the microfuge tube due to evaporation of liquid after over 2 years on the bench, however, each microfuge tube was wrapped in parafilm and spun down before each sample was taken. Regardless, there is still strong activity after two years.

Undoubtably, this experiment represents one of the longest time-based phytase stability experiment in the literature. Sulabo *et al.,* (Sulabo *et al*., 2011) analysed the stability of three commercial enzymes Ronozyme P (DSM Nutritional Products), OptiPhos (Phytex LLC) and

Phyzyme (Danisco Animal Nutrition) over a period of 180 days at room temperature (23 °C). When the enzymes were stored at room temperature as pure products, phytase activity were more than 85% of the initial activity after 180 days. When stored at 37 °C phytases stored in its pure form had only retained activities ranging from 1 to 53%. As such, AC1-2 MINPP can be regarded as in good company, with its impressive storage stability.

The final element of characterisation of AC1-2 MINPP undertaken in this study was an analysis of kinetic parameters. Here AC1-2 MINPP activity was determined in the range of substrate concentrations from 12.5 – 3750 µM $InsP_6$, that is, after preliminary experiments to determine the enzyme concentration and duration of assay limiting substrate depletion to give first-order reaction conditions with respect to substrate. The enzyme activity in units of phytase activity per mg of enzyme, U/mg, was calculated and a non-linear curve fit using the substrate inhibition model provided by GraphPad Prism 8.0.1 was used to determine the V_{max}, K_m and K_i, Figure 4.9.

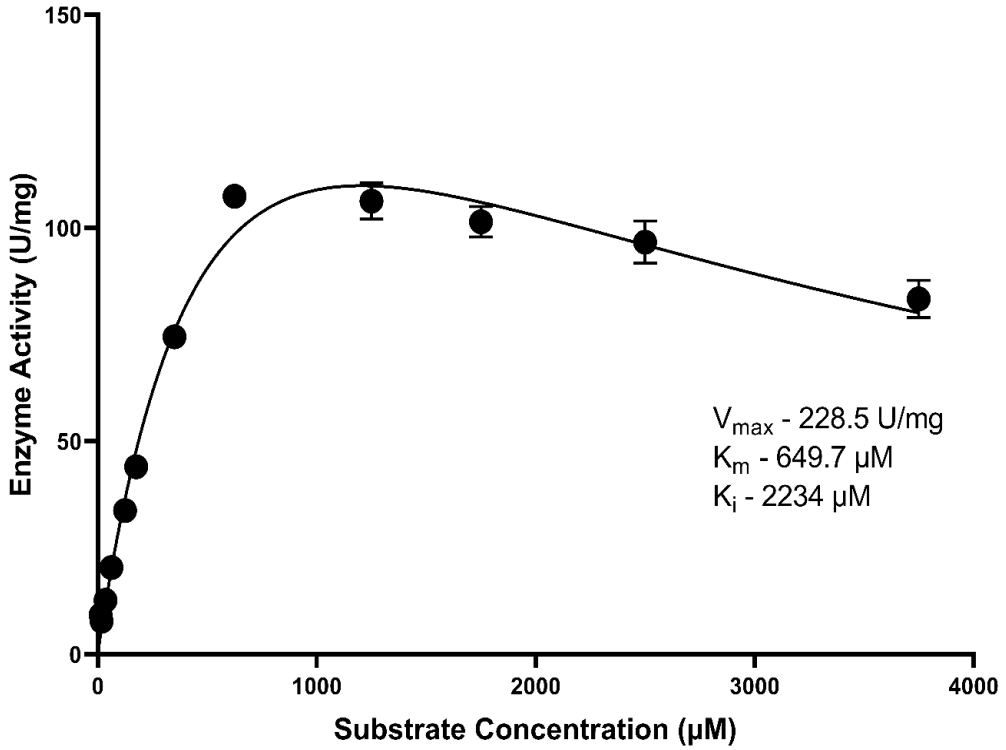


Figure 4.9 –Kinetic parameters of AC1-2 MINPP phytase. Curve fitting was performed by Graphpad Prism 8.0.1 using their substrate inhibition model for V_{max}, K_m and K_i calculations. Error bars for measurements with standard deviation smaller than the symbol are subsumed in the symbol. N=3.

The substrate inhibition model had an R^2 value of 0.9903 which indicates a good fit to the dataset. The best fit values for V_{max}, K_m and K_i were determined to be 228.5 U/mg, 649.7 µM and 2234 µM respectively. This model predicts the maximum velocity if the substrate didn't

also inhibit enzyme activity. Using the known concentration of enzyme, the values for K_{cat} and K_{cat}/K_m were determined to be 226.8 sec^{-1} and 3.49 x 10^5 $M^{-1}sec^{-1}$. The kinetics of AC1-2 MINPP was compared with other HAPhys in the literature, and combined into a table, Table 4.3.

Table 4.3 – The kinetic parameters of four MINPPs and 11 different Histidine Acid Phytases

Species	V_{max}	K_m mM	K_i mM	K_{cat} Sec^{-1}	K_{cat}/K_m	Reference
***Acinetobacter* sp. AC1-2**	228 U/mg	0.65	2.2	226.8	$3.49x10^6$ ($M^{-1}S^{-1}$)	This Study
***Bacteroides thetaiotaomicron* (MINPP)**	178±32 U/mg	0.18±4.1	n/a	146	n/a	(Stentz *et al*., 2014)
***Triticum aestivum* (MINPP)**	1.4 µmol $min^{-1}ml^{-1}$	0.25±0.038	n/a	n/a	n/a	(Dionisio *et al*., 2007)
***Hordeum vulgare* (MINPP)**	1.5 µmol $min^{-1}ml^{-1}$	0.33±0.031	n/a	n/a	n/a	(Dionisio *et al*., 2007)
***Gallus gallus* (MINPP)**	0.715±0.031 U/mg	0.14±0.012	n/a	n/a	n/a	(Cho *et al*., 2006)
***Shigella* sp. CD2**	149.1 µmol min^{-1}	0.18	n/a	2230	$1.23x10^7$ ($M^{-1}S^{-1}$)	(Pal Roy *et al*., 2016)
Yersinia kristensenii	2960 U/mg	0.078	n/a	n/a	n/a	(D. Fu *et al*., 2008)
Obesumbacterium proteus	435 µmol min^{-1}	0.34	n/a	n/a	n/a	(Lassen *et al*., 2001)
***Klebsiella* sp.**	224 (U mg)	0.28	n/a	n/a	$2.357x10^5$ ($M^{-1}S^{-1}$)	(Sajidan *et al*., 2004)
Eupenicillium parvum	88.24±1.06 U/mg	0.58±0.01	n/a	102.94±2.14	$1.78x10^5$ ($M^{-1}S^{-1}$)	(Shimizu, 1993) (Fugthong *et al*., 2010)
Aspergillus niger	135 U/mg	0.148	n/a	168.75±2.1	$1.14x10^6$ ($M^{-1}S^{-1}$)	(Hesampour *et al*., 2015)
Erwinia carotovora	792 U/mg	0.252	n/a	950	n/a	(Huang, 2009)
Klebsiella pneumoniae	99 U/mg	0.280	n/a	66	n/a	(Huang, 2009)
Escherichia coli	3191 U/mg	0.550	n/a	2950	n/a	(Huang, 2009)
Klebsiella terrigena	205 U/mg	0.3	>2	180	n/a	(Greiner *et al*., 1997)

The K_m (0.65 mM) of AC1-2 MINPP is consistent with values of other HAPhys described in the table (0.078-0.65 mM) and those reported anonymously (0.15-1.37 mM) for a range of commercial phytases (Salaet *et al*., 2021). The V_{max} (228 U/mg) was higher than the bacterial MINPPs of *Bacteroides thetaiotaomicron* and considerably faster than those from *Triticum*

aestvum, *Hordeum vulgare* and *Gallus gallus*. For the HAPhys, AC1-2 MINPP was higher than the bacterial strains of *Shigella* sp. CD2 and the two *Klebsiella* strains whilst lower than that of *E. coli* (3191 U/mg) and *Yersinia kristensenii* (2960 U/mg). It was also greater than that (23 – 196 U/mg) of a range of fungal phytases (Tomschy *et al*., 2002; Wyss *et al*., 1999) with which AC1-2 MINPP and other bacterial MINPPs (Acquistapace *et al*., 2020) show greater structural similarity but less than that (≈ 2000 U/mg) of *Peniophora lycii.*

Nearly all of the papers studied in Table 4.3 did not examine the potential for substrate inhibition. Despite this, substrate inhibition by phytate has been well documented (Konietzny and Greiner, 2002). In Figure 4.9, substrate inhibition begins with $InsP_6$ concentrations above 1 mM, with a K_i value of 2.2 mM. This is similar to the inhibition of the bacteria *Klebsiella terrigena*, *Citrobacter braakii* YH-15 and the fungi *Aspergillus ficcum*, which showed inhibition at phytate concentrations above 2 mM, 1.5 mM and 1.2 mM respectively (Greiner *et al*., 1997; Kim *et al*., 2003).

4.2.2. Modelling of AC1-2 MINPP

Another useful experiment in addition to the biochemical characterisation of the AC1-2 MINPP is the computational modelling of the protein structure. The online program SWISS-MODEL is a fully automated protein structure homology-modelling server (Waterhouse *et al*., 2018) which performs a comparative analysis based on published X-ray crystallography structures of the highest scoring sequence homologue to predict the structure of the similar AC1-2 MINPP. These models are important as they can provide a useful aid to the rational design of experiments such as site directed mutagenesis, or in understanding protein stability and function (Forster, 2002).

In this study, two structural models of AC1-2 MINPP were produced using the program SWISS-MODEL, based on published X-ray crystallography structures of the highest scoring sequence homologue, the MINPP from *Bifidobacterium longum* (*Bl*MINPP) with which AC1-2 MINPP shares 34% sequence identity. Homology modelling was used to generate models for the structure of AC1-2 MINPP in the apo-state and bound to the non-hydrolysable substrate analogue inhibitor, *myo*-inositol hexasulfate (IS_6). The overall structure of AC1-2

MINPP is therefore predicted to resemble *Bl*MINPP having α/β- and α-domains with an active site arranged between the two domains, Figure 4.10.

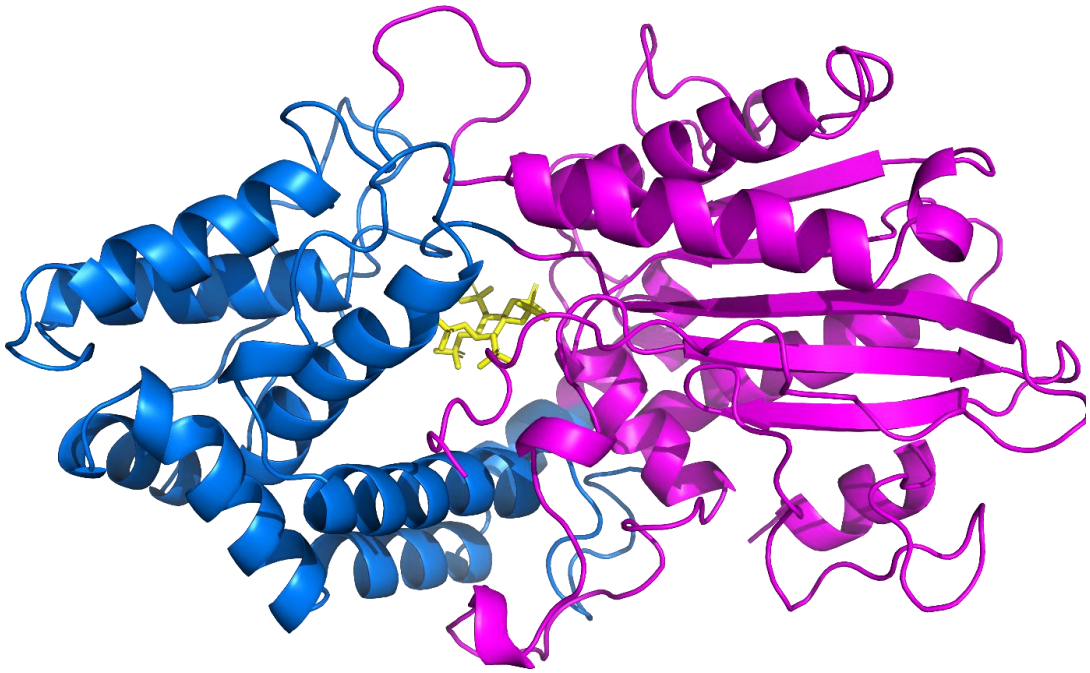


Figure 4.10 – Homology modelling of the structure of AC1-2 MINPP phytase using the program SWISS-MODEL. The active site (yellow) is sandwiched between domains. An α/β- (pink) and α-domains (blue).

Phytases can be grouped according to the position of the phosphate group on the phytate molecule at which hydrolysis first occurs. These can be either 3-phytases (EC 3.1.3.8), 6-phytases (EC. 3.1.3.26) or 5-phytases (EC 3.1.3.72). Homology modelling and structural alignment was employed in an effort to investigate the residues affecting positional specificity of AC1-2 MINPP. Alignment of the structures of sequence homologues to that of AC1-2 MINPP revealed residues forming the specificity pockets in related clade 2 histidine phytases (HP2P) (Rigden, 2008). The sequence homologues were chosen on the basis for which high resolution crystal structures of their complexes with IS_6, were available. Each alignment shows spatially equivalent residues in the specificity pockets of each enzyme which lie within 5Å of the phosphorus of the corresponding phosphate group on the substrate.

The sequence homologues include the MINPPs from *Bifidobacterium longum*, a predominant 4/6-phytase (Acquistapace *et al*., 2020), and from *Bacteroides thetaiotaomicron*, a predominant 5-phytase (Stentz *et al*., 2014). The stereospecific histidine phytase from *Aspergillus fumigatus* is a 3-phytase (EC:3.1.3.8) (Liu *et al*., 2004). In the following, the

specificity pocket nomenclature used was adopted by Acquistapace *et al* (Acquistapace *et al*., 2020) in their study of the cell-surface anchored MINPP from *Bifidobacterium longum*.

Consideration of the spatially conserved residues in the different specificity pockets gives possible insights into the molecular basis for the predominant 4/6-phytase activity of AC1-2 MINPP. Pocket A is occupied by the scissile phosphate group of the substrate. The phosphate here is intimately bound and enjoys six polar contacts with specificity pocket residues. The role of the pocket is to engage with the scissile phosphate and to position and orient it for hydrolysis and, as a consequence, pocket A residues are highly conserved between the enzymes considered. On the other hand, residues in pockets B, C and F have fewer contacts with their corresponding phosphate groups but those residues involved in direct polar contacts with the substrate analogue (and therefore by inference with the substrate), while predominantly conserved with *Bl*MINPP, vary in the *B.thetaiotaomicron* and *A.fumigatus* enzymes. A representation of the specificity pockets is shown, Figure 4.11. Amino acid residues are numbered according to the AC1-2 MINPP sequence.

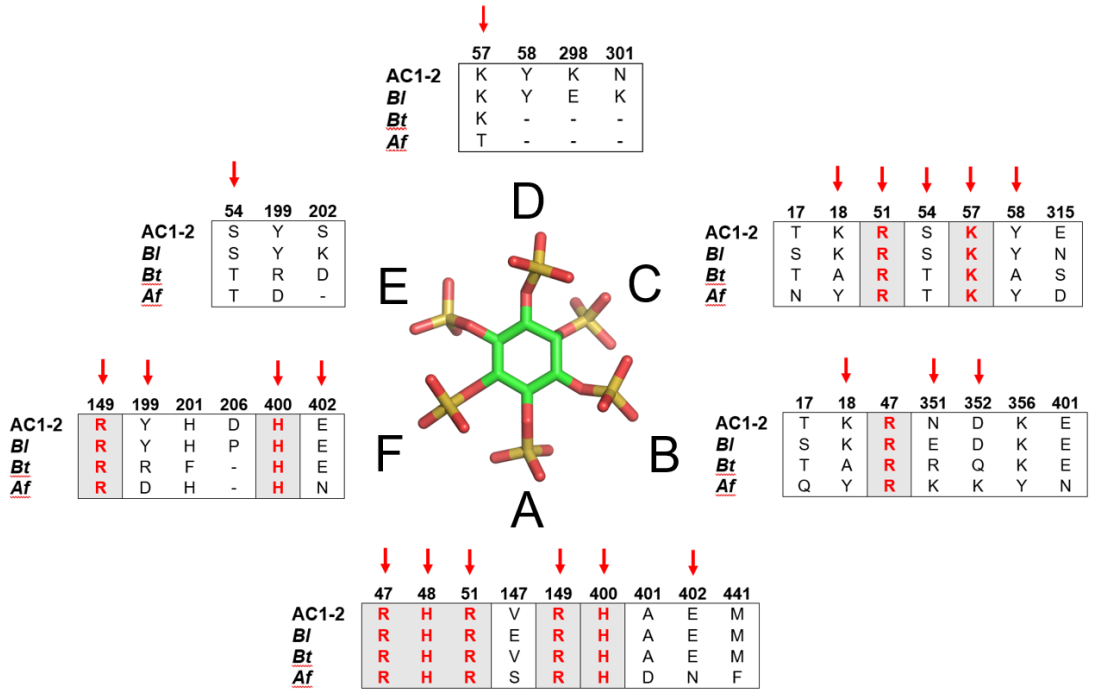


Figure 4.11 – Residues in the specificity pockets of AC1-2 MINPP and selected histidine phytases. This figure was produced with the aid of Professor Andrew Hemmings. Residues that are completely conserved are highlighted in bold with red text. Red arrows indicate residues that have the closest interactions with the substrate analogue inhibitor. AC1-2 – *Acinetobacter* sp. AC1-2, Bl – *Bifidobacterium longum*, Bt – *Bacteroides thetaiotaomicron*, Af – *Aspergillus fumigatus*.

Following this, residues within 5 Å of the phosphorus atoms of the substrate inhibitor IS_6 that were non-catalytic residues, RHGxRxP or HAE active site and proton donor motif respectively, were rendered in PyMOL (DeLano, 2002) to garner a tentative analysis of the interactions between the inhibitor and the enzyme. Eleven residues were identified, Figure 4.12.

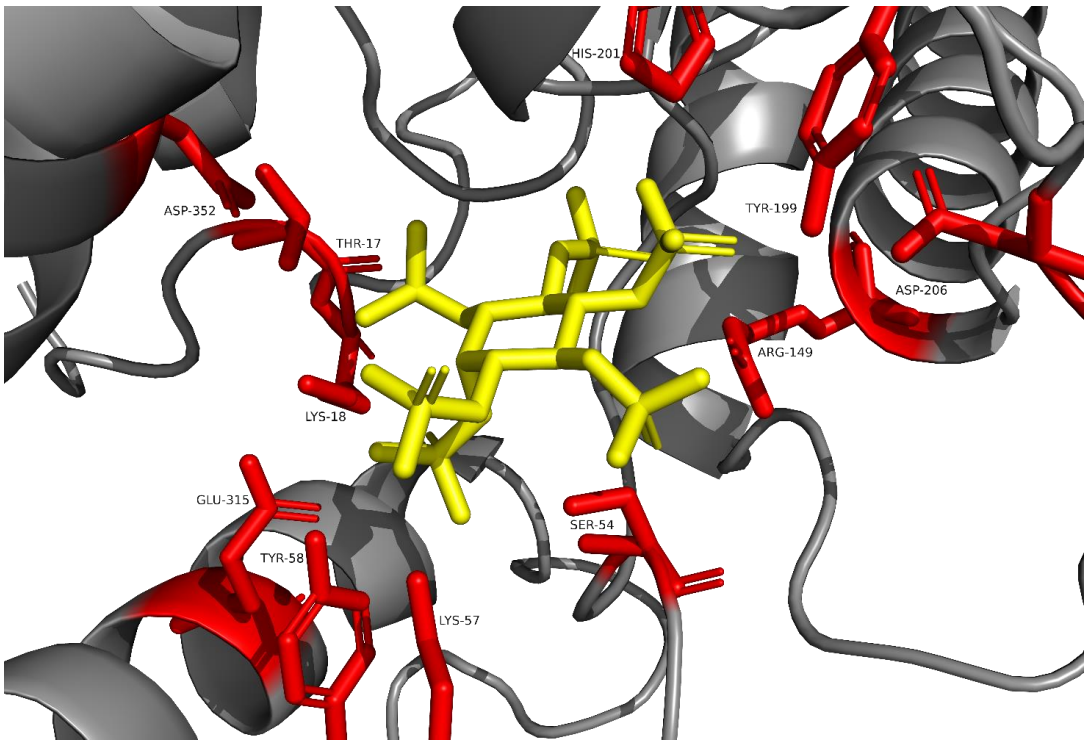


Figure 4.12 – Non-catalytic residues within 5 Å of the phosphorus atom on the phytate molecule of the AC1-2 MINPP model. This was visualised using PyMOL.

These eleven residues are all within the substrate specificity pockets highlighted in Figure 4.11 and therefore likely play a key role in the substrate and positional specificity. The amino acids, Lys-18, Lys-57 and His-201 are all basic amino acids, Asp-206, Glu-315 and Asp-352 are all acidic amino acids, Thr-17, Ser-54, Tyr-58, Tyr-199 are all uncharged amino acids, however they all contain -OH groups which are capable of participation in hydrogen bonding interactions.

In future work, it may be interesting to target these residues for random mutagenesis trials to see the impact this has on binding, activity, and specificity. For example the mutation of R183D in *Bt*Minpp (equivalent to residue 199 in AC1-2 MINPP) converts *Bt*MINPP, a predominant 5-phytase, to an *A. fumigatus*-like 1/3-phytase. It may be that the interactions made by the substrate phosphate groups in these specificity pockets provides the determination of the positional specificity of histidine phytases as a whole. Following this

idea, the alignments of the substrate specificty pockets for AC1-2 MINPP a 4/6-phytase and *Bl*MINPP a 4/6 phytase were compared against those for *Bt*MINPP, a 5-phytase and *Aspergillus fumigatus* a 3-phytase. Residues which were the same between the two 4/6-phytases and different in the 3- and 5-phytase may be of interest for mutagenesis studies, potential targets include K18, Y199 and D352.

Another feature that is present in the structures of the MINPPs is the presence of a large α-domain polypeptide insert termed the U-loop, Figure 4.13. Phylogenetic analysis has revealed three groups of polypeptide inserts in MINPP sequences which have been given the identifiers A, B, or C depending on insert length. This spans the active site and has an influence of substrate specificity pockets. Mutagenesis of U-loop residues has been shown to have an impact on the thermostability and kinetics of the enzyme, C278A and C291A saw a reduction in melting temperature by 10 °C in *Bl*MINPP. In AC1-2 MINPP, residues K298 and N301 which form specificity pocket D are found on the U-loop, Figure 4.11. In the report by Acquistapace *et al* (Acquistapace *et al*., 2020) they captured a large ligand-driven α-domain movement to allow for substrate access to the active site using four crystal structures of *Bifidobacterium longum* obtained by x-ray crystallography. This movement allowing the recruitment of residues to participate in specificity pockets. Figure 4.13B shows the molecular surface representation of the apo- and IS_6- bound AC1-2 MINPP which has been predicted by homology modelling.

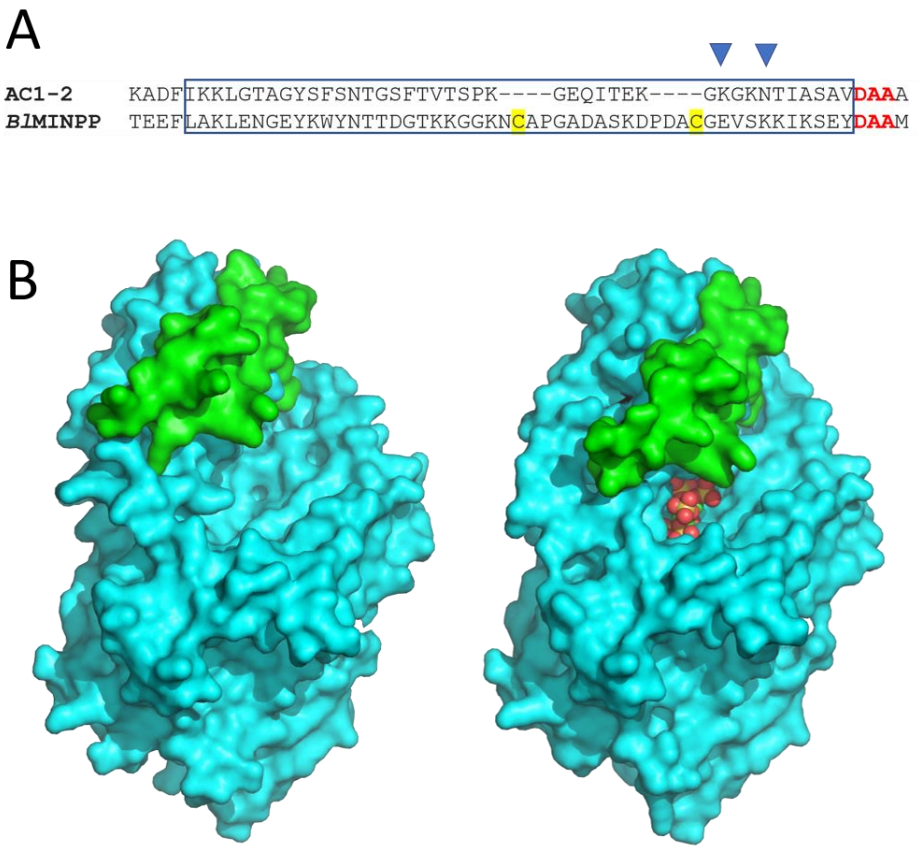


Figure 4.13 – Alignment of AC1-2 MINPP and *Bl*MINPP in the region of the U-loop., A. A blue box delimits U-loop residues. Cysteine residues forming a disulphide bridge in the crystal structure of *Bl*MINPP (PDB entry 6RXD (Acquistapace *et al*., 2020) are highlighted in yellow. Positions of residues contributing to specificity pockets are indicated by inverted blue triangles. B. Molecular surface representations of the structures of apo- (left) and IS_6-bound (right) AC1-2 MINPP predicted by homology modelling. The U-loop residues are coloured green with the remainder of the molecule in cyan. Atoms of the substrate analogue IS_6 are shown as spheres and coloured red (oxygen), cyan (carbon) and orange (sulfur). This image was produced with the aid of Professor Andrew Hemmings.

AC1-2 MINPP has a 41-residue polypeptide insertion in the α-domain which maps to the type A U-loop found in the MINPP from *Bifidobacterium longum* (*Bl*MINPP) (Figure 4.13A). The insertion is followed immediately by a characteristic tetrapeptide motif (DAAM in BlMINPP and DAAA in AC1-2) which is absent in sequences that do not contain a U-loop. The AC1-2 MINPP insertion is, however, shorter than the type A U-loop in *Bl*MINPP by eight residues and lacks the two cysteine residues which form a stabilizing disulphide bridge in the latter. The U-loop residues in *Bl*MINPP span the active site and close down onto a bound IS_6 inhibitor (and by inference onto a bound substrate molecule) through a rigid body motion involving a major part of the α-domain (Acquistapace *et al*., 2020). The prediction of a large type A U-loop in AC1-2 MINPP strongly suggests the presence of similar rigid body domain

motions, presumably to allow the imposition of additional contacts with phytate in the complex, particularly in specificity pocket D (Figure 4.13B).

Interestingly, the lack of the two stabilising cysteine residues in the U-loop that are seen in *Bl*MINPP may infer as to why AC1-2 MINPP exhibits such a low thermostability, Figure 4.6.

4.3. Conclusion

In this chapter, the recombinant version of AC1-2 MINPP that has been produced in *E. coli* was characterised through a variety of different experiments to learn the attributes and kinetics of this enzyme. Despite arising from a soil bacterium, AC1-2 MINPP was more closely related to the enteric MINPP genes from *Bacteroides thetaiotaomicron* and *Bifidobacterium longum* (Acquistapace *et al*., 2020; Stentz *et al*., 2014) than its fellow soil phytases and whilst its kinetic attributes indicate that it is not the fastest of phytases in comparison with the literature, it is one of the most significantly stable phytases over a long period of time, retaining up to 37.6-101.3% activity over 755 days, comparable to that of commercial phytases (Sulabo *et al.*, 2011).

This chapter also highlights the usefulness of structural biology in regards to homology modelling. Despite not having the crystallographic model of AC1-2 MINPP, the structure of AC1-2 MINPP has been inferred using the crystal structure of its closest MINPP homologue *Bl*MINPP. This allowed the identification of eleven non-catalytic residues which may influence the positional specificity and substrate specificity of the enzyme. Additionally, comparisons of the specificity pockets of AC1-2 MINPP with the MINPP phytases from *Bl*MINPP, *Bt*MINPP and the HAPhys from *Aspergillus niger* which have 4/6-, 5-, and 3-positional specificity respectively have highlighted the residues 18, 199 and 352 for candidates for residue determinants of positional specificity.

So far in this thesis, the diversity of phytase-producing microorganisms in the soil environment has been examined through culture dependent studies, showing an array of differing phytase activities in soils and isolating candidate phytases. One phytase from *Acinetobacter* sp. AC1-2 was taken through to completion, from the bacterial strain to the characterised pure enzyme. This approach however, belies the true extent of phytase diversity from the soil environment with more than 99% of bacterial species in soil unculturable by traditional techniques (Pham and Kim, 2012). Therefore, in this next chapter, I examined phytase diversity using a culture-independent process, environmental metagenomics.

Chapter 5. Using culture independent techniques to investigate the diversity of phytases amongst soil and other environments.

This chapter was performed in collaboration with Dr Andy Neal from Rothamsted Research. Using environmental metagenomics, the diversity of the phytase genes, MINPP, HAPhy, BPPhy and PTPhy were examined amongst environmental metagenomes.

One of the main issues of culture-dependent approaches, via traditional plating techniques, for the isolation of microorganisms from the environment is that it does not reflect the overall microbial biodiversity (Austin, 2017). The soil environment is a rich "treasure-trove" of untapped potential, 10^9 individual cells, $>10^6$ distinct taxa in each gram of soil (Curtis, 2005; Gans, 2005), however, more than 99% of these bacterial species cannot be cultured by traditional techniques (Bodor *et al*., 2020; Pham and Kim, 2012). Many bacteria require highly specific growth conditions and therefore arbitrary choice of medium, incubation temperature, aerobic/anaerobic conditions will all impact the types of bacteria that can be isolated. Additionally, the incubation duration is also very important. Short incubation periods of 24-48 hours are typical in the laboratory, however this will preclude slower growing organisms (Mallory *et al*., 1977). While culture will still play a significant role in the future of microbiology, important environmentally relevant bacteria, in this case phytase producers, may be overlooked if only culture-dependent approaches are taken (Bodor *et al*., 2020). To avoid this, experiments should be combined with culture-independent approaches such as isotopic labelling (Kreuzer-Martin, 2007), quantitative real-time PCR (RT-qPCR) (Sharma *et al*., 2007), microarrays (McGrath *et al*., 2010) and environmental metagenomics (Hirsch *et al*., 2010). Metagenomics is described as the direct genetic analysis of genomes contained within an environmental sample (Thomas *et al*., 2012). The first large-scale environmental sequence project was carried out by the J. Craig Venter Institute (JCVI) in 2004 (Sleator *et al*., 2008), in which fragments of DNA from the entire microbial population of the nutrient-limited Sargasso sea were sequenced. A shotgun sequencing approach yielded over 1.6 billion base pairs of DNA and led to the discovery of 1.2 million new genes.

Recently, environmental metagenomics has been used to reveal novel phytases that have little to no homology to the current phytases (Villamizar, Funkner, *et al.*, 2019; Villamizar, Nacke, Boehning, *et al.*, 2019). Additionally, function-based screening from a forest soil metagenome identified the first protein-tyrosine phosphatase with phytase activity

(Villamizar, Nacke, Griese, *et al.*, 2019). The combination of culture-independent techniques to identified new and novel phytases with culture-dependent techniques to directly characterise phytase activity will allow for a more comprehensive understanding of phytase diversity. The first part of this Chapter examines the diversity of MINPP genes in enteric and environmental metagenomes to garner a better understanding of the distribution of MINPPs in the environment in comparison to the β-propeller phytases, Histidine acid phytases and protein tyrosine phytases.

In the second part of this Chapter, the MINPP gene was investigated for Horizontal Gene Transfer (HGT) events that has occurred within the dataset to examine whether this gene is amenable to transfer between the same and differential bacterial species.

Horizontal gene transfer between bacteria is described as the exchange of genetic material from donor to recipient cells, and persistence of the genetic information in the new host via integration into the hosts DNA or autonomous replication (via plasmids) (Van Elsas *et al*., 2003), this can occur between bacteria of the same or different species.

There are three classical mechanisms of DNA transfer in nature: conjugation, transformation, and transduction, the basic mechanisms are highlighted in Figure 5.1.

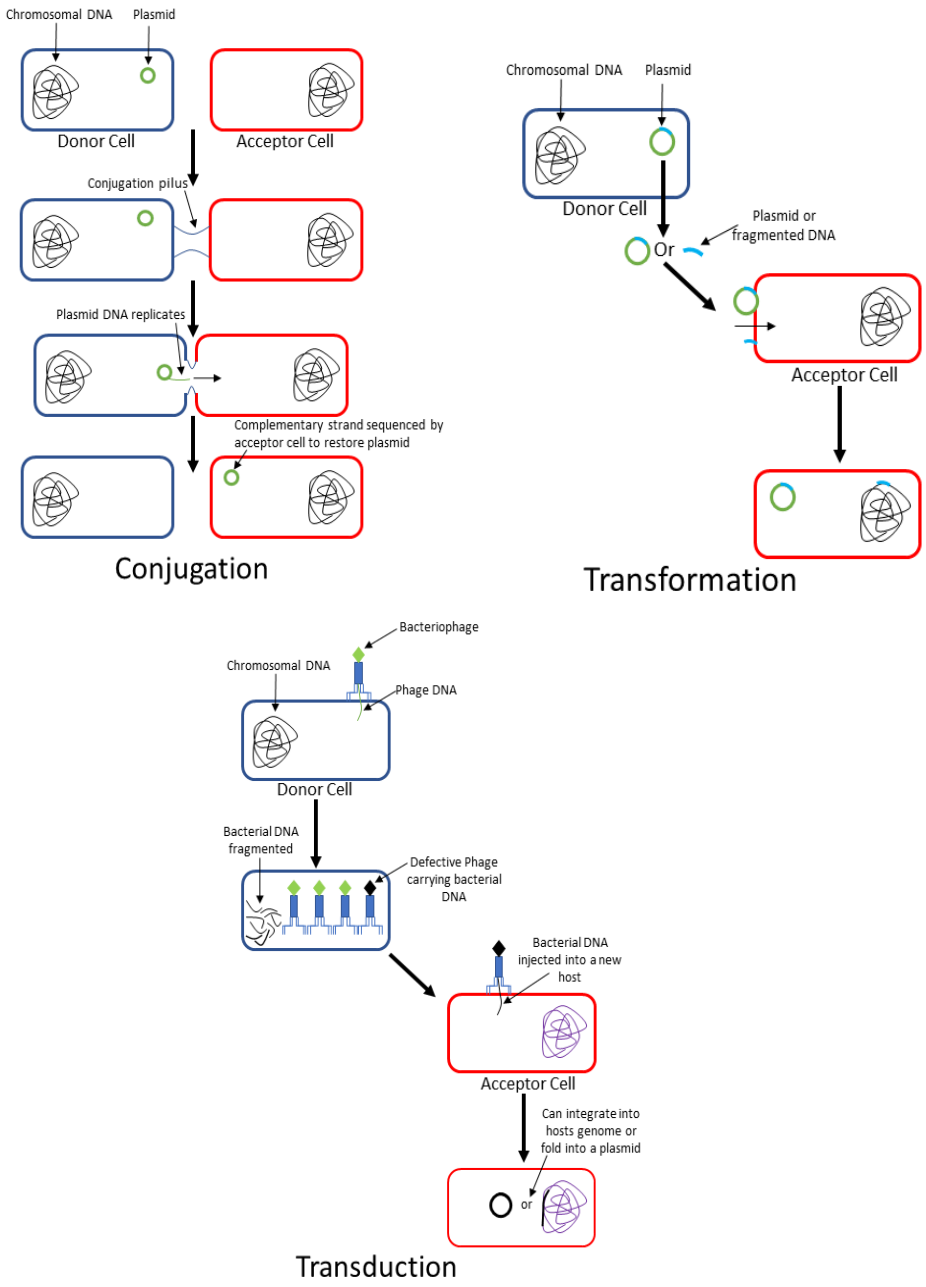


Figure 5.1 – The three classical mechanisms of horizontal gene transfer. These are conjugation, transformation, and transduction. This figure was designed in Microsoft PowerPoint.

In brief, conjugation is the transfer of genetic information between cells via direct cell-to-cell contact. An extension from the donor cell known as a conjugation pilus forms contact with the acceptor cell and draws the two cells in until they are touching. One strand of the plasmid to be transferred is replicated and transferred to the acceptor cell, the recipient then synthesises a complementary strand restoring the complete plasmid. In transformation, there is no direct cell-to-cell contact instead, transformation is controlled entirely by the acceptor cell. Exogenous DNA fragments or plasmids may be released by the death of a

bacterial cell, this can then be taken up by another bacterial cell provided it goes into a special physiological state known as competency. This change causes the cell membrane to become more permeable which allows the uptake of extracellular DNA. Finally, transduction is process in which foreign DNA is introduced into a cell via a virus or viral vector. The bacteriophage infects a bacterial cell with its own DNA and takes advantage of the bacteria reproductive mechanisms to replicate its own DNA and create new phages, at the same time the bacterial DNA is broken down. In some instances, however, bacterial DNA is loaded onto the phage instead of viral DNA which is then injected into a new bacterial host. This injected DNA can be recycled by the bacteria, if the injected DNA was originally a plasmid it may re-circularised and become active again, or if the DNA is homologous to a region within the bacterial genome, recombination between the two sets of DNA may occur (Sun, 2018; von Wintersdorff *et al*., 2016). One worrying aspect of HGT is it is one the main drivers of antibiotic resistance, which has arisen through overuse of antibiotics and development of multidrug resistance bacteria, "superbugs" (Porco *et al*., 2012; Sun *et al*., 2019). Left unchecked, deaths attributable to antimicrobial resistance have been projected to reach 10 million per annum, overtaking the number of deaths caused by cancer (de Kraker *et al*., 2016; Shallcross *et al*., 2015).

This Chapter provides a culture-independent view of the distribution and abundances of phytases in ruminant, monogastric and environmental metagenomes.

5.1. Materials and Methods

5.1.1. Preparing a diverse MINPP dataset for metagenomic studies

A set of 17 bacterial MINPPs were used as starting points to generate a set of reference proteins. Sixteen of the bacterial MINPPs were kindly provided by Dr Acquistapace. The seventeenth was from *Acinetobacter* sp. AC1-2 from the previous chapters.

Of the sixteen, one came from deep-sea sediments in the Pacific Ocean, *Flammeovirga pacifica*; one came from a cold alkaline environment in Ikka Fjord, Greenland, *Arsukibacterium* sp. MJ3; two came from ruminant environments, *Prevotella brevis* ATCC 19188 and *Fibrobacter succinogenes;* four came from a human gut metagenome, *Bacteroides* sp. *CAG:927*, *Bacteroides* sp. CAG:545, *Bacteroides* sp. CAG:770, *Prevotella* sp. CAG:617; one from human faeces, *Bacteroides intestinalis*; one from a chloroform degrading mixed culture, *Bacteroidales bacterium* CF; two from the roots of *Arabidopsis thaliana*, *Aeromicrobium* sp.

Root495, *Aeromicrobium* sp. Root 236; one from a low pH uranium mine sediment, *Oxalobacteraceae bacterium* AB; one from soil from a cave located in South Korea, *Streptacidiphilus jeojiense*; one from dried bat dung, *Amycolatopsis jejuensis*, and finally one from a sea squirt microbiota, *Streptomyces* sp. AW19M42. These were supplemented with the sequence of *Acinetobacter* sp. AC1-2 from untilled farmland soil. These sequences were first aligned into a multisequence alignment (MSA) using MAFFT, the online multisequence alignment tool (Katoh *et al*., 2019), employing the E-INS-I iterative refinement method, BLOSUM62 and a gap opening penalty of 1.53. JackHMMER, part of the HMMER v3.3 web server (Potter *et al*., 2018) was then used to generate a collection of homologous protein sequences from the UniprotKB database using a bitscore cut-off of 300, and the search restricted to Eubacteria. JackHMMER works by generating a profile hidden Markov model (Eddy, 2011). The BLOSUM62 substitution scoring matrix was used together with gap opening and extension penalties of 0.02 and 0.4, respectively. Iterations were continued until no new sequences were included into the dataset. The initial dataset included 1133 sequences.

Due to similarities of the Histidine Acid Phytases and the Multiple Inositol Polyphosphate Phosphatases, a similar JackHMMER was performed using 21 diverse bacterial HAPhys. JackHMMER provides a model for the most likely amino acids at each position. This allowed the manual removal of any HAPhys based on conserved motifs within the MINPPs. There were 4 criteria that were examined to confirm whether the sequence was from a HAPhy or a MINPP. Figures 5.2 and 5.3 highlight conserved features of the HAPhys and MINPPs around the heptapeptide sequence motif and proton donor motif.

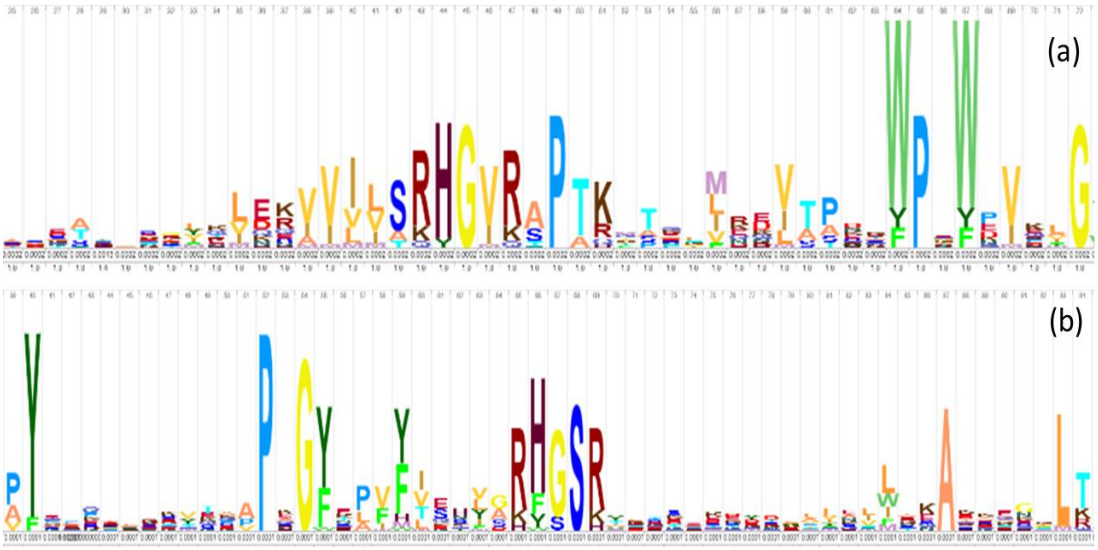


Figure 5.2 – A conserved amino acid model of the HAPhys (a) and MINPPs (b) at the heptapeptide sequence motif, produced using JackHMMER (Potter *et al*., 2018)

The first strong signal when determining the differences between the HAPhys and MINPPs is the Wx(P)xW motif that follows the HAPhys heptapeptide sequence motif, which is not conserved in the MINPP sequences.

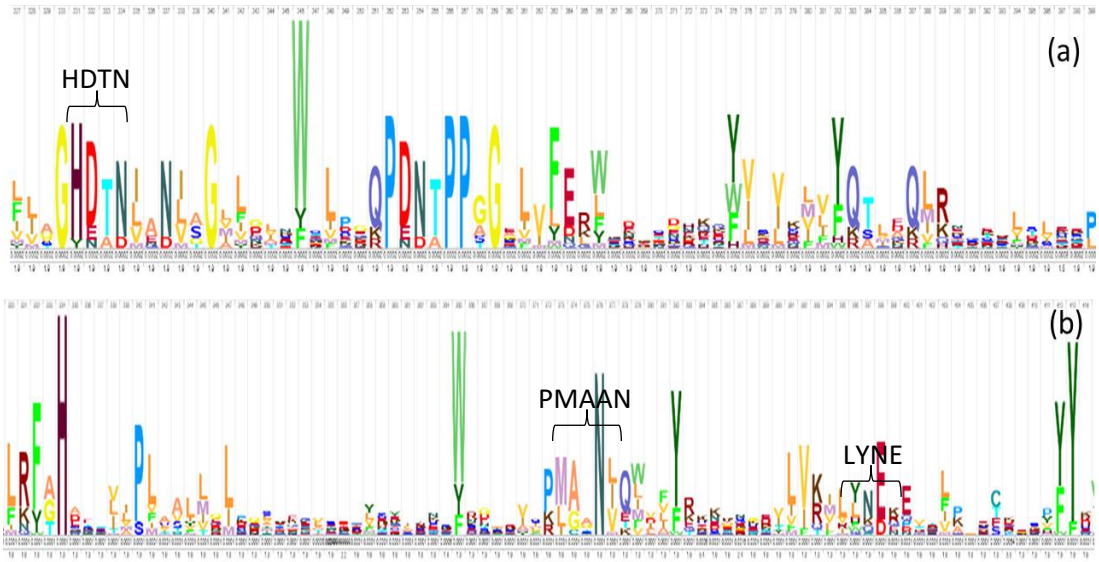


Figure 5.3 – A conserved amino acid model of the HAPhys (a) and MINPPs (b) at the proton donor motif, produced using JackHMMER (Potter *et al*., 2018)

Following this, the tetrapeptide HDTN is another common motif present in the HAPhys though this is not as definitive: HAPhys have other donor motifs, while MINPPs also have the HD motif. Four motifs WRG, PMAAN, LYNE and YY are very common in MINPPs and not found in the HAPhys. The initial datasets were aligned using MAFFT and each sequence examined for the motifs, or similar, as described. Thirty-four sequences were removed leaving a dataset of 1099 protein sequences. The DNA sequences of the complementary protein sequences were downloaded, and redundancy of the dataset was reduced to 99.9% removing a further 384 sequences leaving 715 sequences. Redundancy was removed using CD-HIT (Fu *et al*., 2012; Huang *et al*., 2010), a widely used program for clustering protein or nucleotide sequences to reduce sequence redundancy.

This final nucleotide dataset was again aligned through MAFFT. Maximum-Likelihood phylogenetic trees were generated using RAxML version 7.2.8 (Stamatakis, 2006), employing the PROTGAMMA amino acid substitution evolutionary model and Dayhoff matrix and bootstrapping. The number of bootstrapping replicates were determined using autoMRE convergence checking. The best scoring maximum-likelihood trees were visualised using iTOL version 6 (Letunic and Bork, 2019).

Additional datasets from the HAPhy, BPPhy and PTPhy classes were kindly provided by Dr Andy Neal for metagenomic analysis.

5.1.2. Metagenome Selection

Publicly available metagenomic datasets were used to determine the global distribution of MINPPs. Shotgun metagenomes sequences generated using Illumina® sequencing containing a minimum of 30 million reads were downloaded in FASTQ format from the European Bioinformatics Institute (EBI), National Center for Biotechnology Information (NCBI) and the European Nucleotide Archive (ENA). Datasets were download at Rothamsted Research, Nov 2018.

The specific datasets were chosen because they examine globally diverse environments. Three ruminant datasets were chosen from cattle and sheep. Five monogastric datasets were chosen, two from humans and one each from pig, mouse, and cat. Twelve environmental datasets were chosen, eight were from aquatic environments and four from soil environments. These are described in Table 5.1.

Each metagenome was uploaded to the Rothamsted Galaxy Server. Here they underwent quality control. Sequences were limited to a minimum quality score of 25 using a sliding window of 4 bases, and a minimum read length of 70 bases using Trimmomatic (Bolger *et al.*, 2014).

Table 5.1 – List of the metagenomes used to identify the abundance of MINPPs, HAPhys, BPPhys and PTPhys in diverse environments.

Metagenome	Reads	Sample Number	
Cattle rumen 2	151,613,082	PRJEB21624	**Ruminant**
Sheep Rumen	421,718,257	SRS429584	
Cattle rumen	84,106,366	ERS812543	
Pig Gut	61,427,996	ERS970331	**Monogastric**
Mouse Intestine	206,302,133	ERS823135	
Human Faecal	73,571,387	ERS433544	
Human Dysbiosis	64,177,051	ERS537243	
Feline faecal	39,981,617	ERS723639	
Angelo Meadow Soil	86,572,024	SRR2546430	**Environmental**
Manikaran Hot Springs	44,907,232	ERS1262240	
Activated Sludge	57,049,807	ERS1107856	
Australian Soil	80,189,228	ERR671923	
Hypersaline Lake	31,236,897	PRJEB18068	
Swedish Lake	97,650,020	ERS433967	
Noosa River	247,892,796	ERR688352	
Alaskan Tundra Soil	220,605,353	ERR1035438	
Gulf of Mexico	63,987,136	SRR4027974	
Columbia River	162,059,859	ERR864075	
Fricke Cemetery Soil	33,382,004	ERR346662	
Marcell Forest Peatbog	97,892,632	SRR1157608	

5.1.3. Metagenomic Pipeline

The MINPP MSA nucleotide dataset was used to generate a pHMM (Profile Hidden Markov Model) (Eddy, 1998) using HMMbuild (HMMER version 3.1b1), adopting an assembly-free, gene-centric approach of Neal *et al* (Neal *et al*., 2018; Neal and Glendining, 2019). The pHMMs were then used to search the unassembled metagenomic datasets using HMMsearch employing an $E < 1 \times 10^{-5}$ along the full sequence length. The probability thresholds for the multiple segment Viterbi, Viterbi and Forward filters were 0.02, 0.001 and 1×10^{-5} respectively. pplacer version 1.1 (Matsen *et al*., 2010) was used to phylogenetically place the recovered metagenomic hits upon reference Maximum-Likelihood phylogenetic trees, such as Figure 5.3.

To allow for meaningful comparison between metagenomic datasets, gene abundance was expressed as a proportion of the total estimated number of genomes in each dataset. This was assessed by estimating the number of ubiquitous, single-copy genes, *gyrB*, *recA* and *atpD* (Gaunt *et al*., 2001; Santos and Ochman, 2004). Metagenomic-derived homologues for each single-copy gene were size normalised to the length of the shortest gene, *recA* accounting for differences in length between genes, this is because the likelihood of getting a hit increases with gene length. To do this, the modal length of *recA* (1164 nucleotides) was divided by the modal length of the other single-copy genes (1392 for *atpD* and 2618 for *gyrB*), and this value was multiplied by each single-copy gene count. The size normalised abundance of the MINPP was then calculated for each environment as [Target gene hits] / [mean normalised single copy genes] (Neal *et al*., 2018). To get 1 in n, each value was divided from 1. pHMMs of each of the single-copy genes were kindly provided by Dr Andy Neal from Rothamsted.

The distribution of the MINPP genes in each metagenome were visualised using iTOL. The abundance of MINPP, HAPhy, BPPhy and PTPhy genes were calculated against the mean gene count of *gyrB*, *recA* and *atpD*.

5.1.4. Refinement of the dataset to examine horizontal gene transfer (HGT).

The 715 MINPP MSA was further refined to examine horizontal gene transfer. The redundancy of the MSA was set to 90% using CD-HIT resulting in the removal of 337 sequences, leaving 378 sequences. While considerable redundancy was removed, sequences from *Prevotella* sp. (105) and *Bacteroides* sp. (53) were abundant. In order to remove bias

but retain the natural diversity of the model, the number of *Prevotella* species was reduced to 54, whilst *Bacteroides* was reduced to 51 leaving a MINPP dataset of 325 sequences.

Complementary 16S rRNA sequences were downloaded for these genes using the SILVA database (Quast *et al*., 2012), which provides comprehensive, regularly updated datasets for 16s/18S ribosomal RNA. Unfortunately, many 16S rRNA sequences were not recovered, reducing the dataset to 257 sequences.

Another issue that was encountered was that ten of the bacterial sequences contained two or more MINPP genes, Table 5.2.

Table 5.2 – The ten bacterial species containing two or more MINPP genes.

Bacterial Species	Number of MINPP genes
Bacteroides faecichinchillae	2
Bacteroides oleiciplenus	3
Dysgonomonas gadei	2
Dysgonomonas macrotermitis	2
Prevotellaceae bacterium	3
Prevotella ruminicola	2
Prevotella ruminicola 2	3
Prevotella ruminicola 3	3
Bacteroides sp. *770*	2
Bacteroidales bacterium WCE2004	3

For the horizontal gene analysis to work, identical naming convention is needed for the MINPP and 16S rRNA datasets, without replication. Thus, MINPPs from organisms with multiple copies were manually curated to determine which one to remove, while maintaining overall MINPP diversity. An example of this process is shown in Figure 5.4.

All three of the *Bacteroides oleiciplenus* genes are surrounded by similar *Bacteroides* species so the loss of two of the three should not significantly impact the overall diversity. Therefore, *Bacteroides oleiciplenus_1* was chosen to remain as the other two had closer relatives.

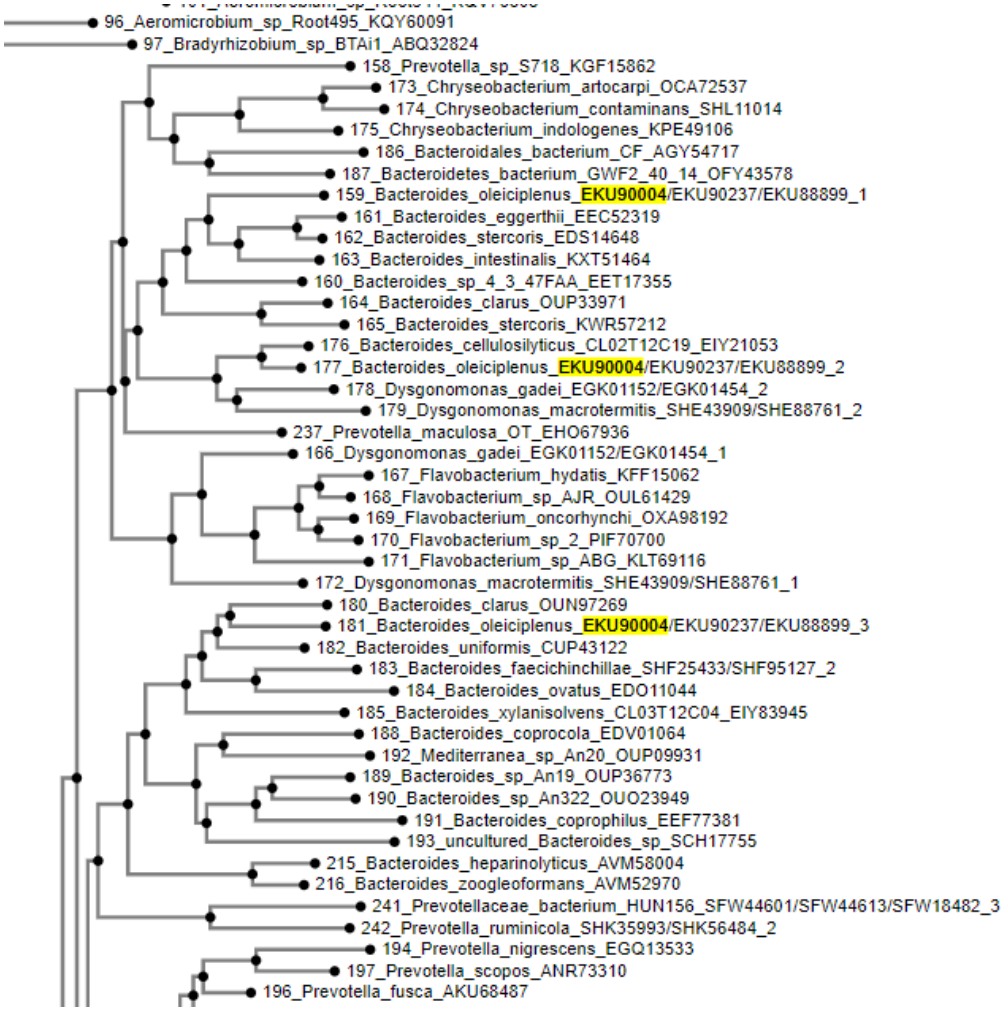


Figure 5.4 – Manually curating the MINPP dataset to remove organisms with multiple copies. Highlighted (yellow) are the three *Bacteroides oleiciplenus* MINPP genes.

With a remaining 241 sequences, both the 16S rRNA and MINPP datasets were aligned using MAFFT. For the MINPP dataset maximum-likelihood trees were generated using RaxML as described above, with 100 bootstraps performed. A single best-tree of the 16S rRNA dataset was also generated using RaxML. The online webserver T-REX (Boc *et al*., 2012; Li *et al*., 2005) was used to infer horizontal gene transfer for the given pair of species and gene trees. In total 134 HGT events were calculated with a bootstrap value ranging from 0-100%.

5.2. Results and Discussion

5.2.1. Examining the distribution of MINPP genes in the environment

The MINPP datasets described in Section 5.1.3 was used to interrogate selected metagenomes for MINPP genes. The number of "hits" for each metagenome are shown in Table 5.3.

Table 5.3 – The number of MINPP hits in each metagenome analysed.

Metagenome	*MINPP* Hits	
Cattle rumen 2	7204	**Ruminant**
Sheep Rumen	16656	
Cattle rumen	1357	
Pig Gut	460	**Monogastric**
Mouse Intestine	1316	
Human Faecal	423	
Human Dysbiosis	699	
Feline faecal	485	
Angelo Meadow Soil	100	**Environmental**
Manikaran Hot Springs	34	
Activated Sludge	18	
Australian Soil	13	
Hypersaline Lake	3	
Swedish Lake	1	
Noosa River	3	
Alaskan Tundra Soil	0	
Gulf of Mexico	0	
Columbia River	0	
Fricke Cemetery Soil	0	
Marcell Forest Peatbog	0	

While this initial data tells us the number of MINPP hits in each metagenome, it does not tell us the relative abundance of each gene, nor does it allow a true comparison between the different metagenomes. This is because they have all been sequenced to different depths. For example, the Sheep Rumen had over 421 million reads whereas the Hypersaline Lake only had just over 31 million reads. Therefore, a higher amount of MINPP hits is expected in the former. Consequently, gene abundance was normalized: expressed in relation to the number of single-copy genes, *gyrB*, *recA* and *atpD*. These calculations are shown in Table 5.4.

Table 5.4 – Calculation of the relative abundance of MINPP genes in 20 different metagenomes using three single-copy genes, *gyrB*, *recA*, and *atpD*.

Metagenome	*gyrB*	*recA*	*atpD*	mean gene count	*MINPP*	norm *MINPP*	Ratio 1 in n	
Cattle Rumen 2	55958	18022	19974	19868.0	7204	0.363	2.76	**Ruminant**
Sheep Rumen	128261	46548	50444	48585.4	16656	0.343	2.92	
Cattle Rumen	19495	8588	9838	8494.1	1357	0.160	6.26	
Pig Gut	14139	6698	6927	6258.9	460	0.073	13.61	**Monogastric**
Mouse Intestine	44135	21737	18237	18870.0	1316	0.070	14.34	
Human Faecal	17970	7963	9930	8085.4	423	0.052	19.11	
Human Dysbiosis	21280	9125	11807	9486.5	699	0.074	13.57	
Feline faecal	15639	5452	9613	6814.6	485	0.071	14.05	
Angelo Meadow Soil	19652	8436	11964	9059.3	100	0.011	90.59	**Environmental**
Manikaran Hot Springs	7928	3540	4766	3683.4	34	0.009	108.3	
Activated Sludge	10767	4414	5841	4695.1	18	0.004	260.8	
Australian Soil	14932	6961	9267	7116.4	13	0.002	547.4	
Hypersaline Lake	8956	2805	5224	3718.4	3	0.001	1239	
Swedish Lake	11418	5438	6886	5424.2	1	0.0002	5424	
Noosa River	47786	19450	30608	22097.0	3	0.0001	7365	
Alaskan Tundra Soil	32120	16081	25132	17125.9	0	0.000	0.00	
Gulf of Mexico	19583	8500	18101	10781.0	0	0.000	0.00	
Columbia River	2429	1098	1786	1223.8	0	0.000	0.00	
Fricke Cemetary	4674	2304	3224	2359.4	0	0.000	0.00	
Marcell Forest Peatbog	12225	8679	10409	7606.2	0	0.000	0.00	
pHMM length	2618	1164	1392					

This information gives us significant insight into the abundance and location of MINPP genes. Previously, MINPPs had only been considered to function inside animal cells, and it was in 2009 when one of the first bacterial MINPPs was characterised from the prevalent gut bacteria *Bifidobacterium pseudocatenulatum*, followed by *Bacteroides thetaiotaomicron* in

2014 (Haros *et al*., 2009; Stentz *et al*., 2014). More recently MINPPs have been isolated from the soil bacterium *Acinetobacter* sp. (Rix *et al*., 2020).

Examination of the metagenomes began with the ruminant environment which had the highest MINPP abundance. Ruminants, made up of large hoofed herbivorous mammals such as cattle, sheep and deer, are characterised by a four-compartmented stomach (Owensby *et al*., 1996), the largest of which is the rumen. This is the primary site of digestion, where plant-based foods are fermented in the stomach to release nutrients, principally through microbial action. This is therefore the most obvious place to examine for microorganisms capable of degrading phytate.

Ruminants were assumed to efficiently utilise phytate phosphate in their plant-based diet through the activity of bacteria found in the rumen, this was based on earlier experiments where the inherent phytase activity of rumen microbes digested nearly all of the $InsP_6$ into inorganic phosphate (Raun *et al.,* 1956), and phytate-phosphate was believed to be fully available to the animal when fed diets to meet their phosphate requirements (Morse *et al*., 1992) and in 1999, the first PTPhy from *Selenomonas ruminantium* was isolated (Yanke *et al*., 1999)

There are, however, still some indications that the use of exogenous phytases may provide better nutrition to the animal. In high-producing ruminants such as dairy cows, faster passage of digesta and suboptimal fermentation conditions may limit $InsP_6$ degradation because of the shorter duration exposure to microbial phytase (Jarrett *et al*., 2014). In an experiment by Park *et al.,* (Park *et al*., 1999), ruminal phytate digestibility decreased from 84% to 69% and then 57% when passage rate increased from 0.02/h to 0.05/h and 0.08/h respectively. Additionally, the use of high-grain diets have also been associated with a reduction in the secretion of saliva, thus possibly reducing the amount of salivary P available for absorption in the small intestine (Humer and Zebeli, 2015).

Regardless, higher levels of phytate degradation in ruminants is supported by the metagenomics data with the MINPP gene highly abundant in these environments. In the digestive tracks of ruminants, approximately 36.2, 34.2 and 16% of bacteria from the environments Cattle Rumen 2, Sheep Rumen and Cattle Rumen contained a MINPP gene, respectively.

Following the ruminants is monogastrics organisms, animals with a simple single-chambered stomach, that have often been touted for their inability to properly utilise phytate

phosphorus. While it is widely reported that monogastrics are limited in their ability to process dietary phytate, hence the presence of a billion dollar per annum adjunct phytase sector, Section 1.4, it is clear that monogastrics have their own endogenous MINPP enzymes (Cho *et al*., 2006) but not enough activity in the right part of the digestive tract (Beeson *et al*., 2017).

Another common misconception is the assumed abundance of phytases in monogastric environments (Brinch-Pedersen *et al*., 2002; Dao, 2005; George *et al*., 2005; Leytem and Maguire, 2005; Singh and Satyanarayana, 2008). As can be seen from Table 5.4, the abundance of MINPPs in these environments ranged from 1 in 13.57 to 1 in 19.11 (7.4-5.2%). It is often written that phosphorus from dietary phytic acid is poorly available in monogastrics due to a lack of endogenous phytases in the gut, and indeed similar has been remarked upon in the introduction of this report (Gupta *et al*., 2015; Lorenz *et al*., 2008). However, several studies have shown effective endogenous phytase activity exists in the gastrointestinal tract of swine and poultry (Cowieson *et al*., 2016; Maenz and Classen, 1998; Tamim *et al*., 2004). In the reports by Tamin *et al*., (Tamim *et al*., 2004) and Wilkinson *et al.,* (Wilkinson *et al*., 2014) they both investigated the influence of dietary calcium on phytate phosphorus hydrolysis and nutrient digestibility. Here Tamin *et al.,* showed that the addition of calcium (0-0.5%) resulted in a reduction of ileal phytate phosphorus degradation from 69.2% to 25.4%, concluding that the ability of poultry to utilise phytate-phosphorus could be significantly enhanced through the use of a form of calcium that is unreactive with the phytate molecule. Inclusion of a 3-phytase improved ileal phytate phosphorus disappearance from 25.4 – 58.9% indicating that an adjunct phytase does improve the efficiency of phytate degradation, however addition of a 6-phytase produced a weaker improvement. Another calcium-based study performed by Wilkinson *et al.,* (Wilkinson *et al*., 2014) encouraged consumption of a separate calcium, $CaCO_3$, source while reducing the dietary calcium from feedstuffs in poultry. This led to a significant increase in ileal digestibility of phosphorus, nitrogen, and other nutrients.

Overall, we see a larger presence of the MINPP gene in both monogastrics and ruminants with the abundance of MINPPs higher in ruminants when compared with monogastrics. These results are predicted, perhaps, by early isolation/characterization of MINPPs from the gastrointestinal tract or the microbiota (Chi *et al*., 1999; Cho *et al*., 2006; Haros *et al*., 2009; Stentz *et al*., 2014; Tamayo-Ramos *et al*., 2012).

Following the two enteric environments, the abundance of MINPPs in environmental metagenomes were examined. Prior to the work described elsewhere in this thesis (Chapter 1, and published (Rix *et al*., 2020)), descriptions of MINPPs have been restricted to plants such as wheat and barley (Dionisio, Holm and Brinch-Pedersen, 2007), to animals (Chi *et al*., 1999; Cho *et al*., 2006) and to gut commensal bacteria (Haros *et al*., 2009; Stentz *et al*., 2014; Tamayo-Ramos *et al*., 2012). Consequently, this metagenomic undertaking is perhaps the first to be performed for MINPPs. As can be seen from the data in Figure 5.4, the abundance of MINPPs in the environment is significantly less than that in enterics, and quite varied depending on the environment. Only one of the twelve environmental conditions had over 1% of bacterial genomes containing MINPP genes, this was from Angelo Meadow. Angelo Meadow is part of a 7,400-acre reserve located in California which has been protected from major human disturbances since the late 1930s. These meadows contain vibrant populations of grasses and forbs which hosts complex food webs. It may be that animal faeces play a role in the accumulation of MINPP genes in this environment. This is followed closely by the sediment of the Manikaran hot springs, India, where only 0.92% of the bacterial genomes contain the MINPP gene. This represents a high temperature environment where temperatures of the sediment were around 57 °C. This is an interesting result with 34 hits from the metagenome. As discussed previously, the ability of a phytase to withstand high temperatures such as those experienced during the pelleting process is very valuable to the phytase industry. The next two environments activated sludge communities from Switzerland, and Australian soil, saw a further decrease in the abundance of MINPPs at 0.38 and 0.18% respectively. The three aquatic environments, the Iranian hypersaline lake, Swedish Lake, and the Noosa River, Australia were even less enriched for MINPP, with between 1-3 hits, 0.08, 0.018 and 0.014% respectively. Finally, the last remaining five metagenomes did not return any MINPP hits.

The abundance of MINPPs genes across this diverse set of environmental metagenomes show a gene that is considerably less abundant in the environment than in enteric niches.

5.2.2. The placement of MINPP genes on the phylogenetic tree

Each of the MINPP gene hits from each metagenome were placed onto the phylogenetic tree, Figure 5.3, and visualised using iTOL. This data is displayed in Figure 5.5, 5.6 and 5.7 for the ruminant, monogastric and environmental metagenomes, respectively. The size of the circles represents how many hits have been assigned by pplacer to the same position on the phylogenetic tree. Unfortunately, this is not directly comparable between the different

metagenomes, as those from ruminants are considerably more abundant than that of the metagenomes, and, consequently, similar scaling would make the circles too small for lowly populated environments. In the representations shown in Figures 5.5-5.7, the phylogenetic clusters have been highlighted and numbered from 1-18. The range of bacteria present in each cluster is highlighted in Table 5.5.

Table 5.5 – The bacterial distribution of the MINPP phylogenetic tree which have been placed into 17 clusters. The number of bacterial species in each cluster is highlighted within brackets.

Cluster	Bacterial gene sequences
1	*Prevotella* sp (45)., *Prevotellaceae* sp (1)., *Bacteroidales* bacterium (1).
2	*Prevotella* sp (32).
3	*Prevotella* sp (15)., *Prevotellaceae* bacterium (4).
4	*Bacteroides* sp (37)., *Mediterranea* sp (1)., *Prevotella* sp (29)., Alloprevotella sp (1)., *Muribaculum* sp (1)., *Fibrobacter* sp (22).
5	*Gilliamella* sp (17)., *Frischella* sp (1).
6	*Chitinophaga* sp (5)., *Filimonas* sp (1)., *Flexibacter* sp (1)., *Arachidicoccus* sp (1)., *Mucilaginibacter* sp (5).
7	*Paenibacillus* sp (10)., *Halomonas* sp (2)., *Sediminispirochaeta* sp (1)., *Cetobacterium* sp (2)., *Fusobacterium* sp (1).
8	*Bacteroides* sp (15)., *Odoribacter* sp (2)., *Parabacteroides* sp (1)., *Barnesiella* sp (2).
9	*Chryseobacterium* sp (6)., *Flavobacterium* sp (6)., *Dysgonomonas* sp (3).
10	*Bacteroides* sp (13).
11	*Bacteroides* sp (6)., *Dysgonomonas* sp (3).
12	*Bacteroides* sp (13)., *Prevotella* sp (20)., *Alistipes* sp (13)., bacterium sp (1)., *Bacteroidales* sp (8)., *Prevotellaceae* sp (3).
13	*Streptomyces* sp (51)., *Actinobacteria* bacterium (1)., Luteimicrobium sp (1)., *Kitasatospora* sp (4)., *Aeromicrobium* sp (5)., *Bradyrhizobium* sp (1).
14	*Bifidobacterium* sp (20)., *Actinomyces* sp (2)., *Curtobacterium* sp (4)., *Microbacterium* sp (1)., *Actinomycetales* sp (1)., *Agromyces* sp (1).
15	*Clavibacter* sp (14)., *Agreia* sp (2)., *Cryobacterium* sp (1)., *Rhodococcus* sp (8)., *Glutamicibacter* sp (1)., *Arthrobacter* sp (2)., *Microbacterium* sp (1)., *Aeromonas* sp (22).
16	*Vibrio* sp (5)., *Marinomonas* sp (5)., *Photobacterium* sp (9)., Oleibacter sp (1)., *Oceanospirillaceae* sp (2)., *Legionella* sp (1).
17	*Acinetobacter* sp (80).
18	*Moraxella* sp (1)., *Duganella* sp (3)., *Janthinobacterium* sp (27)., *Herbaspirillum* sp (3)., *Variovorax* sp (4)., Andreprevotia sp (1)., *Methlomangum* sp (1)., *Burkholderia* sp (62).

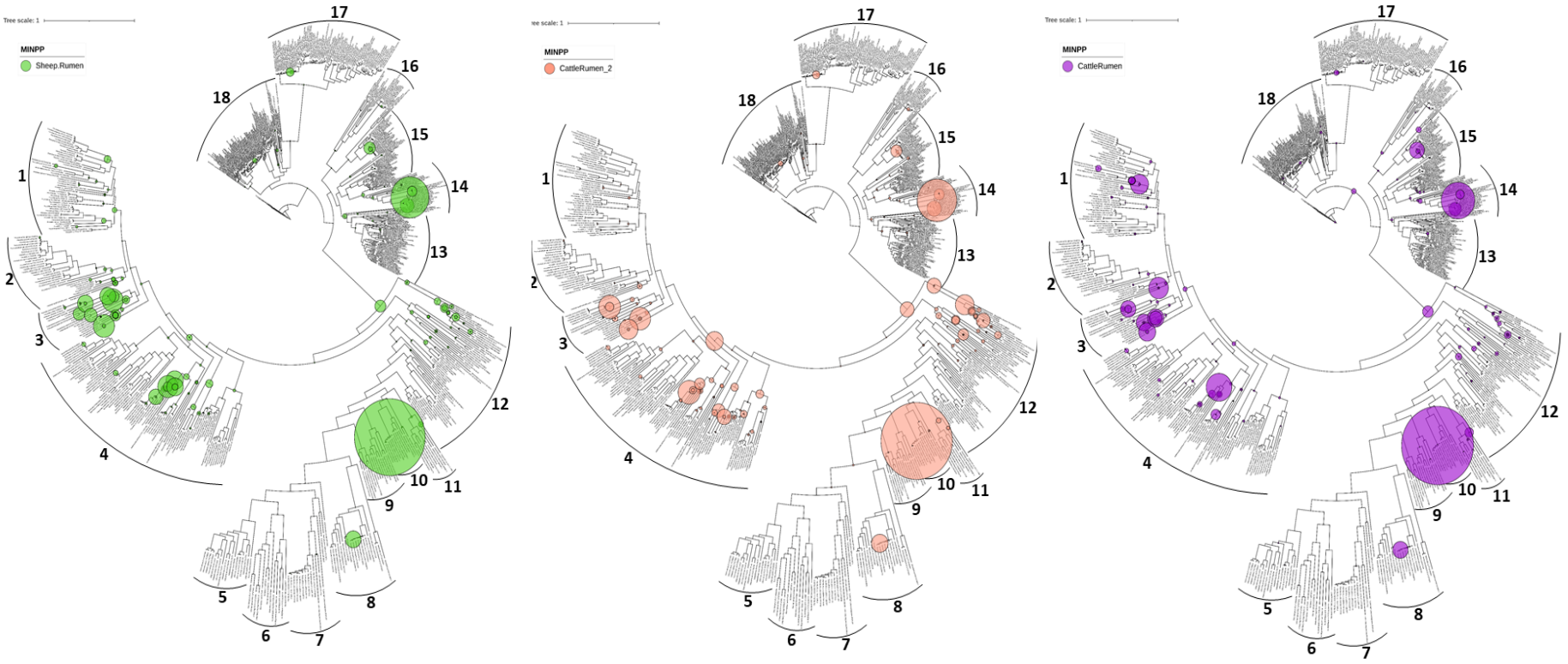


Figure 5.5 – The phylogenetic placement of metagenomic sequences from ruminant metagenomes, Sheep (green), Cattle (purple) and Cattle 2 (orange). The size of the circles is proportional to the normalised relative abundance within the individual metagenome. Maximum hits for each metagenome were 1243, 135 and 579 for Sheep, Cattle and Cattle 2 respectively.

Figure 5.6 - The phylogenetic placement of metagenomic sequences from monogastric metagenomes, feline faecal (green), human dybiosis (purple), human faecal (blue), pig gut (red) and mouse gut (orange). The size of the circles is proportional to the normalised relative abundance within the individual metagenome. Maximum hits for each metagenome were 64, 118, 53, 58 and 151 respectively.

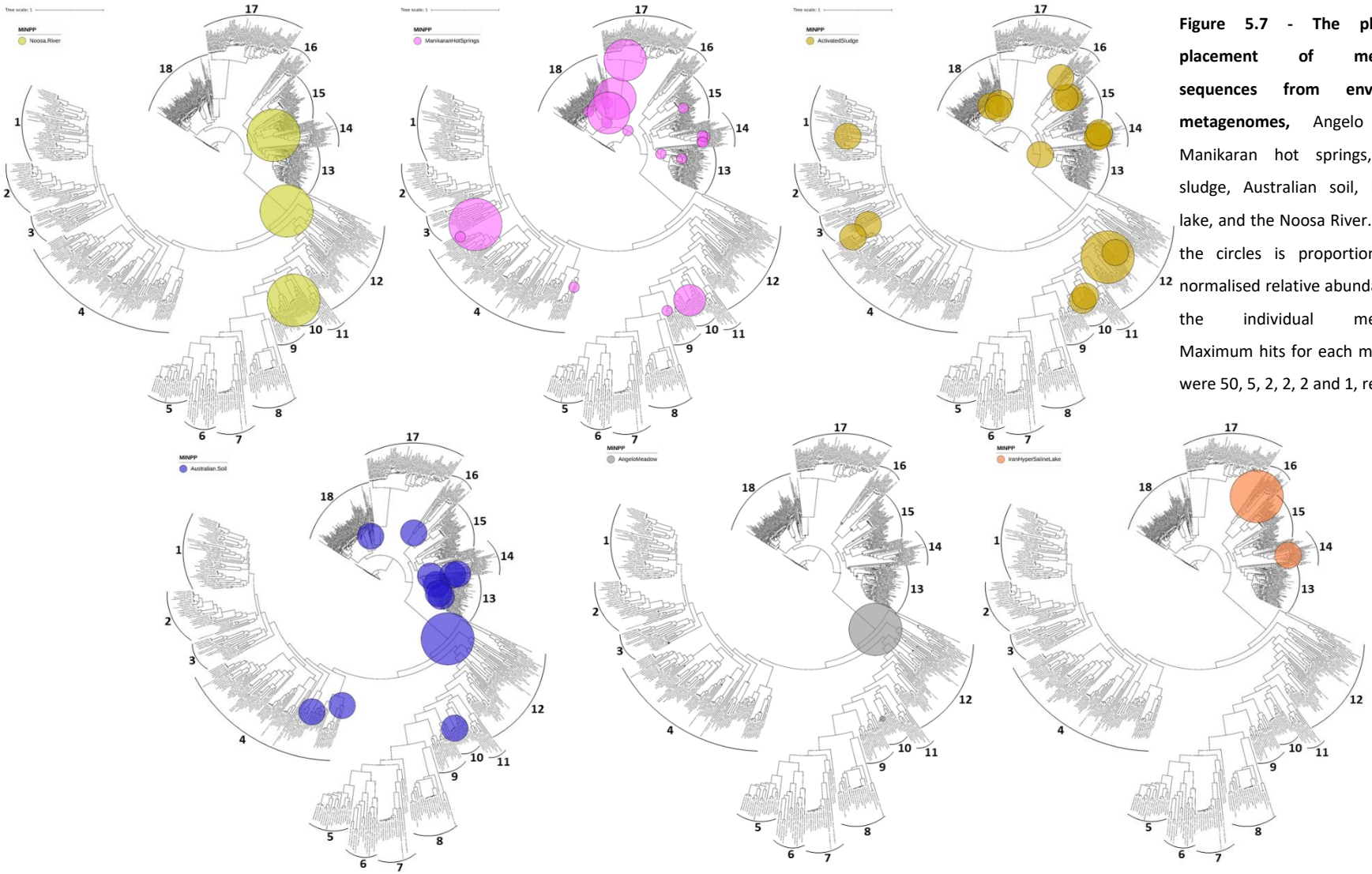


Figure 5.7 - The phylogenetic placement of metagenomic sequences from environmental metagenomes, Angelo Meadow, Manikaran hot springs, activated sludge, Australian soil, hypersaline lake, and the Noosa River. The size of the circles is proportional to the normalised relative abundance within the individual metagenome. Maximum hits for each metagenome were 50, 5, 2, 2, 2 and 1, respectively.

All three of the ruminant environments, Figure 5.5, showed a large number of hits at cluster 10, exclusively homologous MINPPs from *Bacteroides* sp. This is unsurprising as *Bacteroides* species are gram negative, obligate anaerobic bacteria that make up a substantial part of the mammalian gastrointestinal tract (Salyers, 1984; Wexler, 2007). The next largest hits are at cluster 14 which is centred on the *Bifidobacterium* species, these gram-positive bacteria are also very common inhabitants of the gastrointestinal tract (O'Callaghan and van Sinderen, 2016). Next, we see smaller clusters of hits at cluster 3 and 4 which are predominantly *Prevotella*, *Bacteroides* and *Fibrobacter* species. Once again, both *Prevotella* and particularly *Fibrobacter* are common rumen and gut bacteria (Shinkai, Ueki and Kobayashi, 2010; Henderson *et al.*, 2015; Ley, 2016). Undoubtably though, homologous MINPP genes present in the rumen metagenomes are much more diverse and spread throughout the tree, even in the environmental cluster, only hindered through the sheer abundance of hits in the aforementioned clusters.

The monogastric environments are less abundant but fairly similar to their ruminant counterparts. In Figure 5.6, clusters 10 and 14 are some of the most populous clusters. Clusters 3 and 4 however, are less abundant in the feline, human dybiosis and pig gut metagenomes. Interestingly, cluster 1 has a higher abundance of hits in feline and pig gut metagenomes, this cluster is predominantly *Prevotella* whose MINPP diversity is spread throughout clusters 1-4. The mouse gut also has a high proportion of hits in cluster 12, in the *Bacteroides/Prevotella* region.

Finally, the environmental metagenomes were substantially less abundant and diverse then its enteric counterparts. The dominant hit of Angelo Meadow making up half of the total hits was deep within the tree between the two enteric bacteria clusters. The Manikaran hot springs contained more MINPP hits in the environmental cluster of the metagenomic tree while still containing hits from *Prevotella* in cluster 3. The activated sludge and Australian soil were similar, with a lot of hits clustering in the environmental region, but still showing homology hits with more enteric bacteria. The Iranian hypersaline lake and Noosa River only had three hits apiece and the Swedish lake (not shown) only had one hit.

This experiment represents one of the first studies of MINPP gene diversity across environments. In the rumen metagenomes there is a considerably higher number of MINPP hits that are spread throughout the clusters. The extent of the diversity is hidden in Figure 5.5 due to the sheer abundance of MINPP hits in clusters 10 and 14. Removing the two largest

number of hits in Cattle Rumen, Figure 5.8, shows more homologous hits in the environmental cluster and throughout the phylogenetic tree.

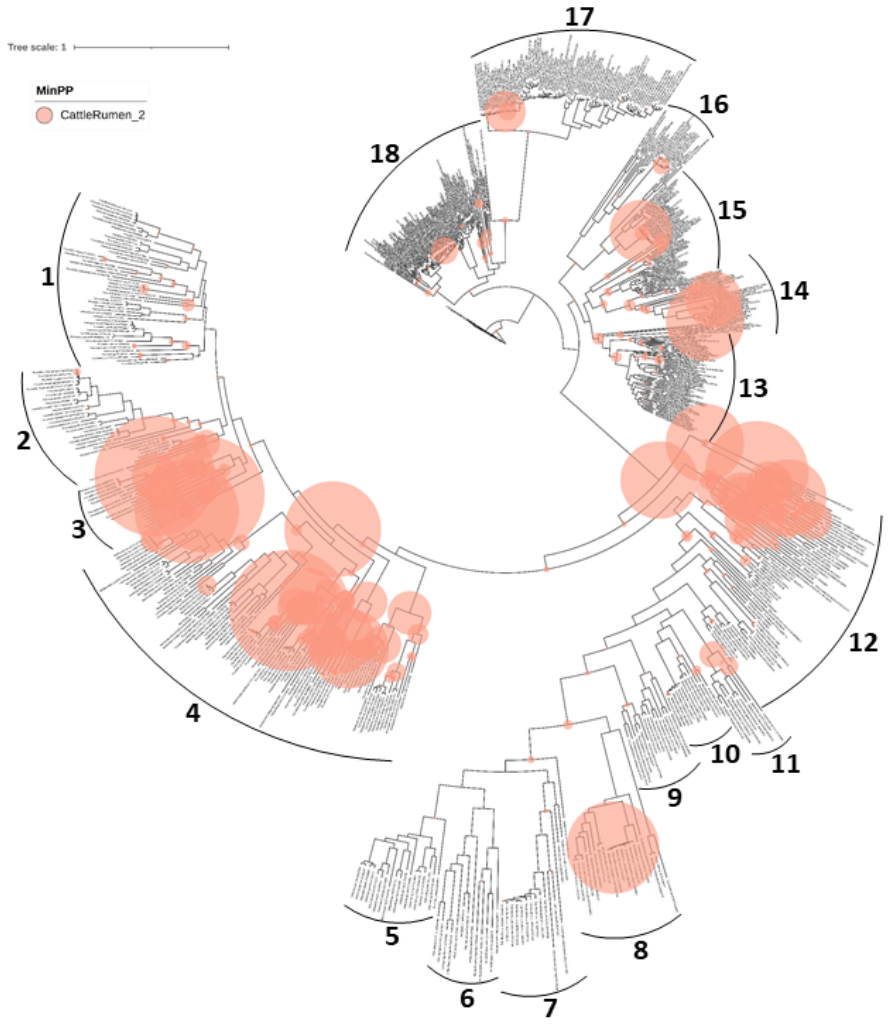


Figure 5.8 – The phylogenetic placement of metagenomic sequences from Cattle Rumen 2 after the removal of the two largest hits.

The homologous MINPPs in monogastric metagenomes are similar to the ruminant in that they are dominant by the abundant gut bacteria *Bacteroides*, *Bifidobacterium* and *Prevotella*, there is, however, less abundance and diversity. The environmental metagenomes were more varied in their larger hits then the ruminant and monogastric environments which all followed the cluster 10 and 14 as the dominant hits. Due to the lower abundance of homologous environmental MINPPs, smaller number of hits were more apparent. Generally, these hits were clustered around the environmental cluster with the exception of Angelo Meadow with a large number of hits deep in the tree between the two enteric clusters. Removal of this cluster however shows many small clusters of hits in the environmental cluster, Figure 5.9.

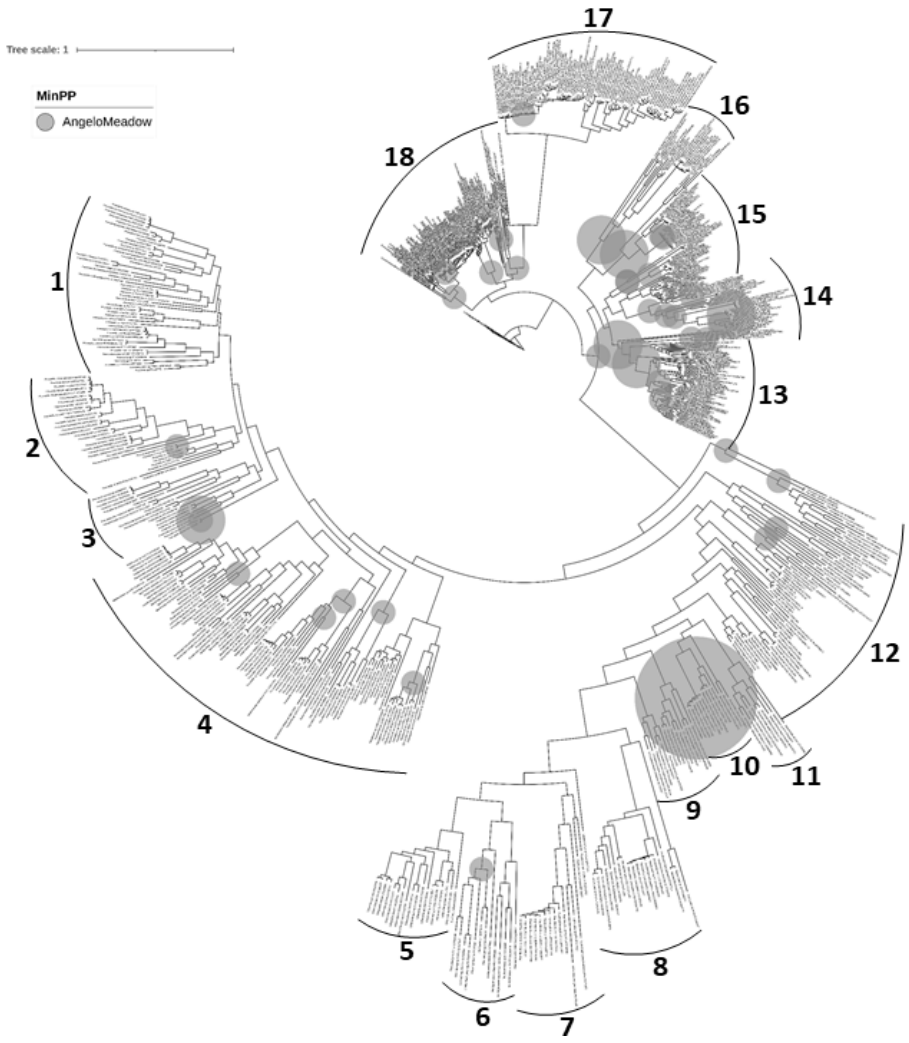


Figure 5.9 – The phylogenetic placement of metagenomic sequences from Angelo Meadow after the removal of the largest hits.

Overall, we can see that the *MINPP* genes in ruminant and monogastric environments are predominantly between clusters 1-12 containing enteric bacteria such as *Prevotella* and *Bacteroides* species, as well as cluster 14 containing *Bifidobacterium* in a region of more environmentally present bacteria. *MINPP* genes from environmental metagenomes such as soils and rivers are widespread throughout the clusters, but they are generally found in the environmental cluster.

5.3. Examining the diversity of phytase genes in the environment

More broadly, the curation of data from the aforementioned metagenomes allows analysis of the *MINPP* gene abundance relative to other histidine acid phytases, protein-tyrosine-like phytases and β-propeller phytases. The purple acid phytases were omitted in this study as they have been considered to be restricted to plants (Dionisio *et al*., 2011), however, recent

publications has revealed this may not be the case (Castillo Villamizar *et al*., 2019b; Ghorbani Nasrabadi *et al*., 2018) and that purple acid phytases can be found in bacteria. The abundance of *HAPhy, PTPhy* and *BPPhy* genes were examined in each of the 20 metagenomes used for the *MINPPs*, the relative abundance is shown in Table 5.6.

Table 5.6 - The relative abundance of four phytase genes, *MINPP, HAPhy, PTPhy* and *BPPhy* in 20 metagenomes.

	Percentage (%)				
	MINPP	***HAPhy***	***PTPhy***	***BPPhy***	
Cattle Rumen 2	36.23	4.0	3.81	0.010	**Ruminant**
Sheep Rumen	34.25	2.0	1.37	0	
Cattle Rumen	15.97	0.48	1.88	0	
Pig Gut	7.35	0.048	0.67	0	**Monogastric**
Mouse Intestine	6.97	0.18	0	0	
Human Faecal	5.23	0.43	2.12	0	
Human Dysbiosis	7.37	0.15	0	0	
Feline faecal	7.12	0	0.059	0	
Angelo Meadow Soil	1.10	0.23	0.95	0.75	**Environmental**
Manikaran Hot Springs	0.92	0.35	0	0.87	
Activated Sludge	0.38	0.021	0.38	1.92	
Australian Soil	0.18	0.23	0.34	4.82	
Hypersaline Lake	0.08	0	0	11.48	
Swedish Lake	0.018	0.018	0.33	0.074	
Noosa River	0.014	0.0091	0.027	1.40	
Alaskan Tundra Soil	0	0.82	0.12	0.041	
Gulf of Mexico	0	0	0	0.074	
Columbia River	0	0	0	0.25	
Fricke Cemetary	0	0	2.33	1.14	
Marcell Forest Peatbog	0	0.36	0.12	0.11	
Mean gene count calculated using *gyrB*, *recA*, *atpD*					

The most obvious result to arise from this comparison is the preponderant abundance of *MINPP* genes in the rumen of ruminants and monogastric environments, with an average percentage of 28.8% compared to 2.16, 2.35 and 0.0034% for the *HAPhy*, *PTPhy* and *BPPhy* respectively in the ruminant metagenomes and 6.8% compared to 0.16, 0.57 and 0% in the monogastric environments.

The PTPhys or cysteine phytases were first identified in the anaerobic bacteria *Selenomonas ruminantium* (Yanke *et al*., 1999, 1998), and moreover, most of these phytases have been isolated from anaerobic culture (Nakashima *et al*., 2007). Therefore, its presence in the rumen metagenomes should not be surprising, though at 3.8, 1.4 and 1.9% for Cattle Rumen 2, Sheep Rumen and Cattle Rumen respectively, its abundance is considerably lower than that of the MINPPs (36.2, 34.2 and 16%). The HAPhys, of which MINPPs are a subclass, are also present in the rumen at similar abundances to the *PTPhys* 4, 2, 0.48%, though literature reports of HAPhys in this environment are lacking. Both of these phytases are classed as acid phosphatases (pH optima 4-5), however, the pH of the rumen ranges from 6-6.5 (Duffield *et al*., 2004; Huang *et al*., 2011), therefore both the HAPhys and PTPhys may not be as active nor efficient for phytate hydrolysis. The MINPPs on the other hand tolerate a breadth of pH that encompasses this range. The MINPP from *Bifidobacterium longum* (Tamayo-Ramos *et al*., 2012) shows activity above >75% at pH 6 and AC1-2 MINPP, Figure 4.1, is optimally active, 100%, at this pH as well. This may account for the preference and considerable abundance of *MINPP* genes in the rumen.

There is a significant difference in the abundance of HAPhy and PTPhy in the monogastric metagenomes. The HAPhys were on average more than 10-fold less abundant in the monogastric environment, there were also no hits with regards to the Feline metagenome, this may however, be an issue with the lack of depth, with just under 40 million reads. The PTPhys also contained less hits in all but the Human Faecal metagenome, additionally, the Mouse Intestine and Human Dysbiosis had no hits. Regardless, the MINPPs are once again the dominant phytase in this environment, 7.3, 7.0, 5.2, 7.4 and 7.1%, once again the ability of the phytase to act over a wide range of pH's suits the pH changes along the gastrointestinal tract, the pH activity of the MINPP from *Bacteroides thetaiotaomicron* overlaps with the human GI tract suiting phytate hydrolysis across its length (Fallingborg, 1999; Stentz *et al*., 2014).

Finally, in these enteric environments the diversity of the beta propeller phytases were examined. As expected, all of the environments with the exception of Cattle Rumen 2 (2 hits) had 0 hits. These phytases are referred to as alkaline phytases with activity from the neutral to alkaline range (Kumar *et al*., 2017), therefore they would be unsuitable in the rumen and gastrointestinal tract of monogastrics. An exception to this is aquatic animals such as monogastric and agastric fish. They have a digestive tract with a neutral environment which is more suited to BPPhys and as such have been isolated from the intestinal contents of grass carp (Huang *et al*., 2009a, 2009b).

The environmental metagenomes were more varied and were less populous than the enteric environments, additionally we now see the emergence and preference of the BPPhys. In Angelo Meadow, the MINPPs no longer dominated in the metagenome with 1.1, 0.232, 0.95 and 0.75% for the MINPP, HAPhy, PTPhy and BPPhys respectively. Surprisingly the PTPhy were the second highest abundance in this environment, it was only in 2019, where the first protein tyrosine phosphatase with phytase activity were characterised from a soil metagenome (Castillo Villamizar *et al*., 2019c). The presence of PTPhy genes had been identified earlier in environmental metagenomes but had not been characterised (Neal *et al*., 2017), this may be an indication that the PTPhys are not solely an enteric phytase in the rumen, but have a wider importance in the cycling of phytate. The higher abundance of the BPPhys is also a feature that should be expected, a study by Lim *et al* in 2007 (Lim *et al*., 2007b) highlighted BPPhy as the dominant class of phytase in aquatic environments, as well as its distribution in the soil environment, which was later analysed by Neal *et al* (Neal *et al*., 2017), being present in bare fallow, arable and grassland soils.

The Manikaran hot springs metagenome saw both MINPP and BPPhy being the dominant phytases, followed by HAPhy, surprisingly there was no hits for PTPhy, 0.923, 0.35, 0 and 0.87% for the MINPP, HAPhy, PTPhy and BPPhys respectively. The activated sludge from a wastewater treatment plant had lower abundances of MINPP, HAPhy and PTPhy but an increased abundance of the BPPhy, 0.383, 0.021, 0.38 and 1.92%. The Australian soil was similar with the BPPhy once again being the dominant phytase, BPPhy, MINPP, PTPhy, 0.183, 0.22, 0.34, 4.8%. Interestingly the hypersaline lake from Lake Meyghan, Iran where salinity ranged from 5, 20 to over 30% total salt and a alkaline pH (7.7-8.8) (Naghoni *et al*., 2017) had the highest abundance of BPPhy hits, a very small number of MINPP hits and no HAPhy and PTPhy, 0.08, 0, 0, and 11.48%. Surprisingly, the BPPhy are not the most abundant in the Swedish lake metagenome and the MINPP, HAPhy and BPPhy abundances are extremely low, and it is the PTPhy that has the highest abundance, 0.018, 0.018, 0.33 and 0.074%. This is an interesting result as the presence of PTPhys in aquatic systems are unknown and not described in the literature. The Noosa River is similar except the BPPhys are the most abundant, and the others are extremely rare, 0.0091, 0.0091, 0.027 and 1.4%. In the Alaskan tundra soil, for the first the HAPhys are the most abundant in this environment, 0, 0.82. 0.12 and 0.04%. Also, interestingly in the Gulf of Mexico and Columbia River metagenomes, metagenomic hits were only confined to the BPPhys, at 0.074 and 0.25% respectively. In the Fricke Cemetery, a tallgrass prairie ecosystem, there was no presence of MINPP and HAPhy genes and instead there was a high abundance of the PTPhys and BPPhys, 2.33 and 1.14%.

Finally, the Marcell Forest peatbog had a higher abundance of HAPhys, and lower abundances of PTPhy and BPPhy, 0, 0.35, 0.12 and 0.11%, the low abundance of the BPPhy in this environment is expected as peatbogs environment are typically acidic (Lin *et al*., 2014).

In summary, this study has allowed inspection of the diversity of phytase genes in diverse gut niches and a range of environments from around the globe. The main message from this research is the sheer abundance of MINPP genes in the enteric environments of ruminants and monogastrics that are many times higher than the commonly described ruminant phytase, PTPhy, and the familial HAPhys. There is also a higher abundance of the three phytase genes in the ruminant metagenomes than the monogastric environment, which appears to agree with the literature on the abundances of phytase genes and ease of phytate degradation. In the external environment, of rivers, lakes, seas and soil, there is a sharp decrease in the abundances of these three genes and an increase in the abundances of the BPPhy which are barely seen in enteric environments. The distribution and abundances of the phytase genes are now even more varied, but typically the BPPhy or the PTPhys are the most abundant in these environments.

The advances in next-generation sequencing and metagenomic techniques have provided a great boon to the discovery of novel enzymes from metagenomes (Daniel, 2005; Nacke and Daniel, 2014). The report by Berini *et al* provides a comprehensive list of 332 industrially relevant enzymes that had been discovered from metagenomes within the last three years of publishing (Berini *et al*., 2017), these include lipases, esterase, proteases, hydrogenases and many others. These were discovered using a variety of different methods, for example a novel cellulase and xylanase was identified through soil enrichments and the DNA was extracted and cloned into metagenomic libraries (Mori *et al*., 2014). A total of 23,000-40,000 clones were screened for cellulase and xylanase activity and positive clones were identified with no significant homology to known cellulase genes. This study used a plasmid library and positive colonies were identified based on a function-based screening approach. These libraries can also be based on cosmids, fosmids and bacterial artificial chromosome (BAC) which allow for greater cloning capacity (Nierman and Feldblyum, 2001). In addition to the type of library, the type of screening approach is also varied. These can include, phenotypical detection, heterologous complementation, induced gene expression, genetic screening, PCR amplification of metagenomic library DNA and PCR, *in silico* analysis of metagenomic library DNA and *in silico* analysis of shotgun eDNA (Berini *et al*., 2017).

These methods have provided a better understanding of the diversity of organisms outside what was possible to culture in the laboratory (Howe and Chain, 2015). The investigation of the diversity and potential of environmental phytases remains largely unexplored so far, with currently, almost all of the functionally characterised phytases derived from cultured organisms (Castillo Villamizar *et al*., 2019b). In the metagenomic studies by Villamizar *et al*, Farias *et al* and Tan *et al* (Castillo Villamizar *et al*., 2019b; Farias *et al*., 2018; Tan *et al*., 2014), novel phytases with unique characteristics were isolated, indicating that the classical concept of phytase classes may need to be modified. Villamizar *et al* identified a previously unreported phytase activity of the alkaline phosphatase and sulfatase superfamily of the purple acid phosphatases from nonvegetal origins, one of the first reports identifying a purple acid phytase from a bacterial source. While not phytase related, they also identified gene products carrying domains, SNARE-associated domain DedA, that have never been associated with phosphatase activity before. In the report by Farias *et al* (2018) they isolated a BPPhy with closest similarity to an unculture bacterium and was set apart from other known phytases in the context of a phylogenetic tree. Furthermore, in the report by Tan *et al* (2014) two novel phytases were identified, one was an unusual HAPhy containing an additional amino acid residue in the conserved heptapeptide catalytic motif of RHGxRxP, instead being RHGLRYYP, with the addition of a tyrosine, Y, residue in the sixth position. The second phytase contained within one of 14,400 clones possessed no conserved motifs of any known phytase suggesting that this clone possesses an unknown type of phytase.

My metagenomic study has further expanded the knowledge of phytase diversity in the environment and can provide a basis for future investigations and screening approaches of environmental and enteric metagenomes.

5.4. Examining the MINPP dataset for evidence of Horizontal Gene Transfer

One interesting feature of the MINPP phylogeny tree was the presence of *Bifidobacterium* species in a cluster of predominantly environmental bacteria, Figure 5.10. *Streptomyces*, *Curtobacterium, Clavibacter* are all aerobic, gram-positive bacteria that are part of the Actinobacteria phylum. These are typically found in aquatic and terrestrial habitats throughout the world (Servin *et al*., 2007). *Bifidobacterium* is also from the Actinobacteria phylum; however, it is an anaerobic and ubiquitous inhabitant of the gastrointestinal tract.

This poses the queston how homologous genes could be spread between two different bacteria from two widely different environments. Most likely, it is due to both of these bacteria's evolving from a common ancestor, however, an interesting mechanism for the transfer of genes between the same and different species may also play a role in the phylogenetic history of the *MINPP* gene. This process known as Horizontal gene transfer (HGT).

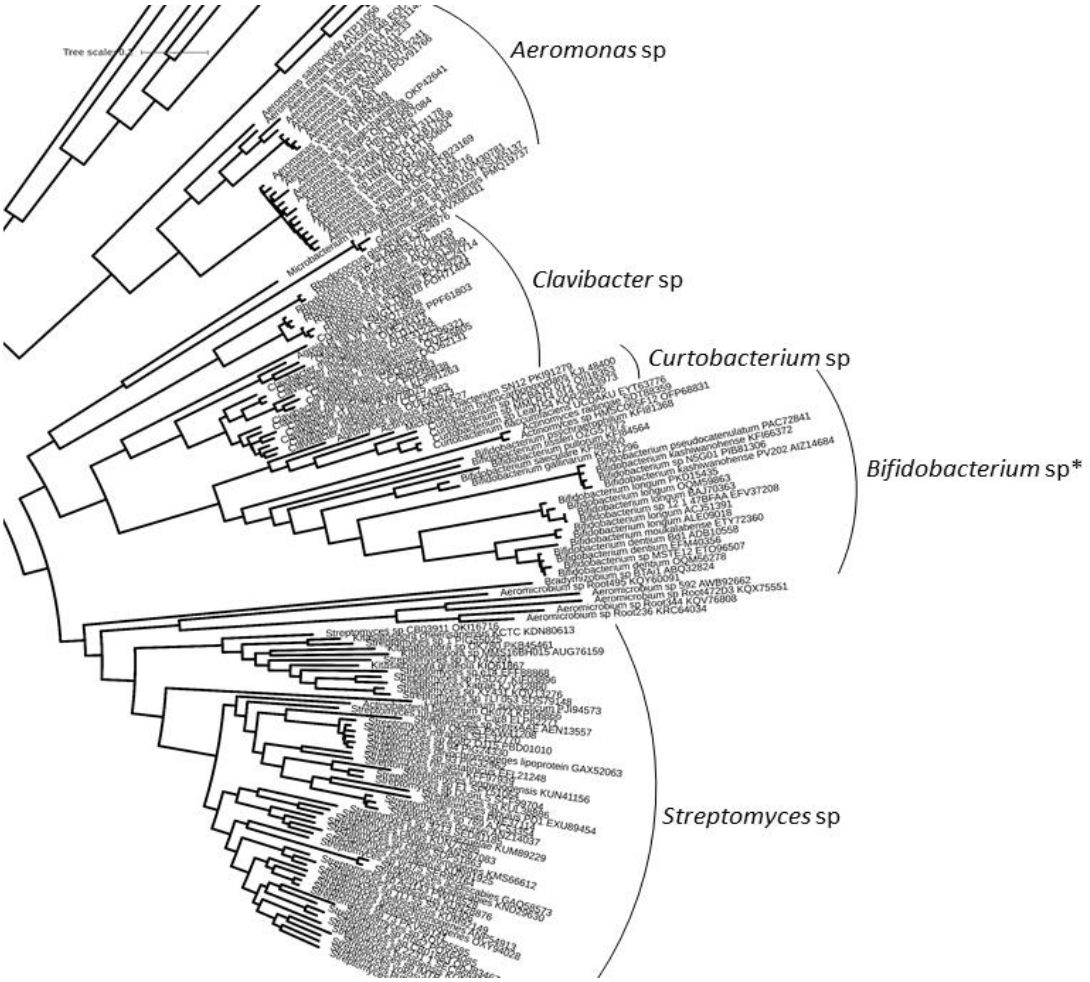


Figure 5.10 – An exploded view of the MINPP phylogenetic tree highlighting the cluster of bacteria that contain MINPP genes that are similarly related to those of *Bifidobacterium* sp.

The *MINPP* gene is an excellent candidate for HGT, it does not appear to be an essential gene nor is it ubiquitously present in all organisms. Essential genes that are ubiquitously expressed like those encoding transcription, translation and replication operons are less likely to undergo transfer as bacterial genomes naïve to their functions are rare (Dutta and Pan, 2002; Lawrence, 1999). It is therefore, the genes whose mechanism/products catalyse steps in a

single pathway and provide for weakly selected function that are preferable for HGT (Lawrence and Roth, 1996).

Therefore, to analyse the potential HGT events of *MINPP*s further analysis was made of the phylogenetic tree and metageonomic datasets. The dataset was refined once more, as too many leaves demand excessive computing requirements (time and power). It was first refined using CD-Hit to reduce redundancy and then manually curated to reduce the abundancy of the *Prevotella* and *Bacteroides* genes, as they may skew the results. At the same time, the 16S rRNA data corresponding to each of the MINPPs were downloaded, further reducing the dataset to a final number of 241 leaves.

In order to examine HGT the program T-Rex (Tree and Reticulogram Reconstruction) was used, which infers HGT events based on the MINPP and 16S rRNA species phylogenetic trees (Boc *et al*., 2012; Li *et al*., 2005; Makarenkov, 2001). Simply, a phylogenetic tree is a structure used to represent the process of evolution between different species, and to analyse these differences statistically (Joseph Felsenstein, 2004) . The leaves at the ends of the trees represents the species under study (the dataset), the nodes within the tree represents a "common ancestor" and the vertical branches are the evolutionary events. There are, however, several complex evolutionary mechanisms that cannot be adequately described by a phylogenetic model, such as recombination and HGT, these can be defined as reticulate evolution which cannot be represented by the phylogenetic tree model (Legendre and Makarenkov, 2002). A reticulogram reveals relationships amongst differing organisms that may have one or more path connecting one to another. Combining the linked 16S rRNA and gene datasets, the program T-Rex calculates the gene and species matrices and uses these to detect potential HGTs based on incongruence between the two matrices, plotting this information onto a HGT reticulogram, Figure 5.12.

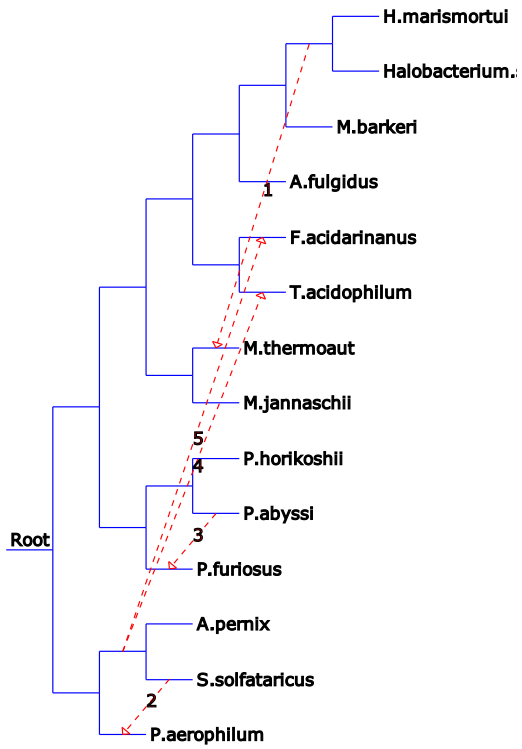


Figure 5.12 – A horizontal gene transfer reticulogram based on the gene *rpl12e*, ribosomal protein, provided by the T-rex server and originally considered by (Matte-Tailliez *et al*., 2002).

Furthermore, this data analysis only represents a single 16S rRNA species tree compared with a single gene tree. By performing a bootstrap analysis, 100 gene trees, the robustness of the obtained HGTs can be analysed, providing a "score" for each HGT based on frequency of occurrence among gene trees in which the HGT was observed, Figure 5.13.

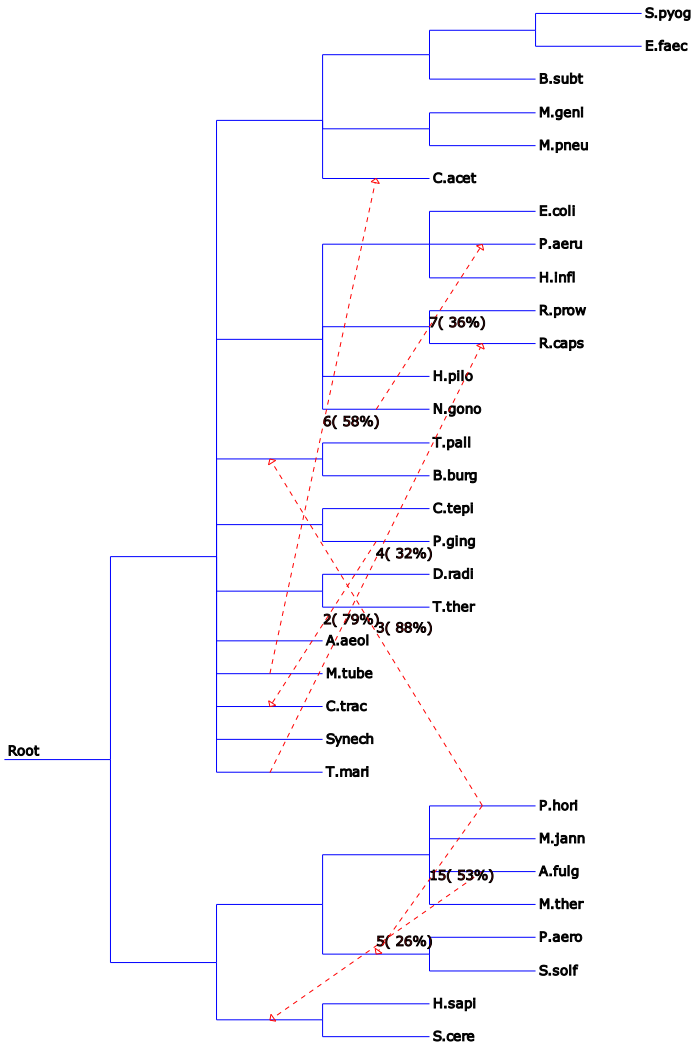


Figure 5.13 – A horizontal gene transfer reticulogram showing horizontal transfers of the gene *phe*RS, provided by the T-rex server and originally considered by (Woese *et al.*, 2000). One hundred bootstrapped gene trees were provided, and the scores are shown as percentages.

In this experiment, the best fit 16S rRNA species tree was generated using RAxML alongside 100 bootstrapped MINPP trees. An overview of the whole HGT reticulogram is shown below in Figure 5.14. A total of 134 HGT events were identified, symbolised by red arrows, indexed with number and bootstrap value, the higher the bootstrap value, the greater the confidence in assignment.

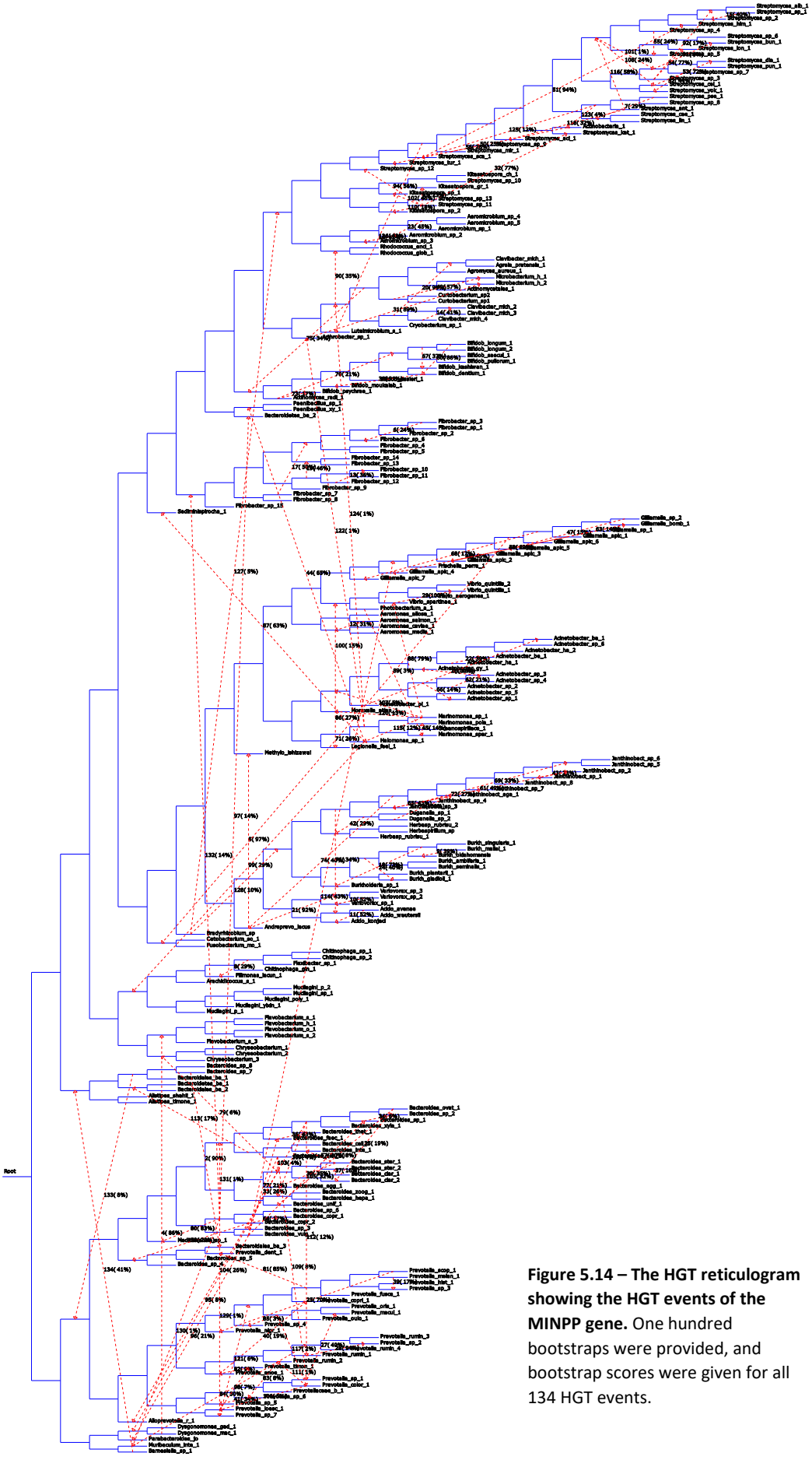


Figure 5.14 – The HGT reticulogram showing the HGT events of the MINPP gene. One hundred bootstraps were provided, and bootstrap scores were given for all 134 HGT events.

5.4.1. HGT events

Overall, there were 33 HGT events with a 50% bootstrap confidence score or higher, or with an inverse bootstrap score of >50%. Twelve of the HGT events identified were between different species, Table 5.7.

Table 5.7 - The HGT events with a >50% bootstrap confidence score.

HGT Event		Bootstrap Value Forward and Inverse (%)	HGT between different species?
1	Acinetobacter_pi_1 to subtree Acinetobacter_ba_1	100.0, 1.0	N
2	Acinetobacter_sp_3 to subtree Acinetobacter_gy_1	55.0, 7.0	N
3	Bifidob_dentium_1 to subtree Bifidob_moukalab_1	91.0, 9.0	N
4	Bifidob_longum_1 to subtree Bifidob_dentium_1, Bifidob_moukalab_1	86, 14.0	N
5	Prevotella_scop_1 to subtree Prevotella_nigr_1	70.0, 1.0	N
6	Prevotella_rumin_1 to subtree Prevotella_sp_2	94.0, 1.0	N
7	Prevotella_fusca_1, Prevotella_hist_1, Prevotella_melan_1, Prevotella_nigr_1, Prevotella_scop_1, Prevotella_sp_3 to subtree Prevotella_dent_1	85.0, 14.0	N
8	Kitasatospora_gr_1, Streptomyces_kat_1, Streptomyces_sp_11 to subtree Streptomyces_sp_12	56.0, 1.0	N
9	Streptomyces_sp_11 to subtree Streptomyces_kat_1	77.0, 1.0	N
10	Streptomyces_sca_1 to subtree Streptomyces_him_1	94.0, 7.0	N
11	Streptomyces_cel_1 to subtree Streptomyces_dia_1	72.0, 1.0	N
12	Streptomyces_sp_3 to subtree Streptomyces_lon_1	77.0, 21.0	N
13	Streptomyces_ant_1 to subtree Streptomyces_bun_1, Streptomyces_cel_1, Streptomyces_dia_1, Streptomyces_pun_1, Streptomyces_sp_4, Streptomyces_sp_5, Streptomyces_sp_6, Streptomyces_sp_7, Streptomyces_yok_1	58.0, 31.0	N
14	Actinobacteria_1, Streptomyces_aci_1, Streptomyces_alb_1, Streptomyces_ant_1, Streptomyces_bun_1, Streptomyces_cae_1, Streptomyces_cel_1, Streptomyces_dia_1, Streptomyces_him_1, Streptomyces_lin_1, Streptomyces_lon_1, Streptomyces_mir_1, Streptomyces_pse_1, Streptomyces_pun_1, Streptomyces_sca_1,	62.0, 1.0	Y

	Streptomyces_sp_1, Streptomyces_sp_2, Streptomyces_sp_3, Streptomyces_sp_4, Streptomyces_sp_5, Streptomyces_sp_6, Streptomyces_sp_7, Streptomyces_sp_8, Streptomyces_sp_9, Streptomyces_tur_1, Streptomyces_yok_1 to subtree Luteimicrobium_s_1		
15	Variovorax_sp_2 to subtree Variovorax_sp_1	52.0, 49.0	N
16	Bacteroides_sp_7 to subtree Bacteroidales_ba_3	90.0, 10.0	N
17	Acido_avenae to subtree Acido_konjaci	52.0, 49.0	N
18	Fibrobacter_sp_7 to subtree (Fibrobacter_sp_1, Fibrobacter_sp_2, Fibrobacter_sp_3, Fibrobacter_sp_4, Fibrobacter_sp_5, Fibrobacter_sp_6	50, 2.0	N
19	Gilliamella_bomb_1 to subtree Gilliamella_apic_4	62.0, 38.0	N
20	Burkh_mallei_1, Burkh_oklahomensis, Burkh_singularis_1 to subtree Burkholderia_sp_1	73.0, 1.0	N
21	Clavibacter_mich_2, Clavibacter_mich_3, Clavibacter_mich_4 to subtree Agreia_pratensis_1, Clavibacter_mich_1	99.0, 1.0	N
22	Vibrio_spartinae_1 to subtree Vibrio_quintilis_1	100.0, 1.0	N
23	Duganella_sp_2 to subtree Janthinobact_aga_1	100, 1.0	Y
24	Dysgonomonas_mac_1 to subtree Flavobacterium_h_1, Flavobacterium_o_1, Flavobacterium_s_1, Flavobacterium_s_2, Flavobacterium_s_3	86.0, 2.0	Y
25	Andreprevo_lacus to subtree Methylo_ishizawai	97.0, 1.0	Y
26	Andreprevo_lacus, Methylo_ishizawai to subtree Variovorax_sp_1, Variovorax_sp_2, Variovorax_sp_3	92.0, 6.0	Y
27	Arthrobacter_sp_1 to subtree Microbacterium_h_2	89.0, 1.0	Y
28	Halomonas_sp_1 to subtree Paenibacillus_sp_1, Paenibacillus_xy_1	65.0, 33.0	Y
29	Curtobacterium_sp1, Curtobacterium_sp2 to subtree Microbacterium_h_1	57.0, 9.0	Y
30	Bacteroides_copr_1, Bacteroides_copr_2, Bacteroides_sp_4, Bacteroides_sp_5, Bacteroides_sp_6 to subtree Mediterranea_sp_1	83.0, 6.0	Y
31	Muribaculum_inte_1 to subtree Prevotella_rumin_3, Prevotella_rumin_4	50.0, 14.0	Y
32	Halomonas_sp_1, Paenibacillus_sp_1, Paenibacillus_xy_1 to subtree Sediminispirocha_1	63.0, 3.0	Y

33	Marinomonas_sp_1 to subtree Vibrio_quintilis_1, Vibrio_spartinae_1	79.0, 22.0	Y
34	Aeromicrobium_sp_2 to subtree Aeromicrobium_sp_3	42.0, 57.0	N
35	Aeromicrobium_sp_4, Aeromicrobium_sp_5 to subtree Aeromicrobium_sp_2, Aeromicrobium_sp_3	45.0, 53.0	N
36	Bacteroides_ovat_1 to subtree Bacteroides_clar_2	19.0, 58.0	N
37	Bacteroides_faec_1 to subtree Bacteroides_thet_1	43.0, 50.0	N
38	Bacteroides_ster_1 to subtree Bacteroides_egg_1	35.0, 62.0	N
39	Prevotella_melan_1 to subtree Prevotella_sp_3	17.0, 66.0	N
40	Streptomyces_sp_4 to subtree Streptomyces_sp_5	24.0, 61.0	N
41	Bacteroides_egg_1, Bacteroides_ster_1 to subtree Bacteroides_inte_1	30.0, 71.0	N

Of the HGT events between the same bacterial genus, it was *Streptomyces* that had the most transfer events (7). Interestingly, these ubiquitous soil bacteria are one of nature's most commonly exploited source of antibiotics, from 1955-1962 around 80% of the antibiotics originated from actinomycetes amongst which *Streptomyces* was a major contributor, and nearly half of the antibiotics in clinical use are derived from the natural products from these bacteria and close relatives. While there was a substantial decline in development in the late 1980s and 1990s (Watve *et al*., 2001) this is not due to exhaustion of the resource, in fact a large majority of the biosynthetic gene clusters coding for natural products are not expressed in the laboratory settings and thus, there is still significantly diversity left to be explored (McLean *et al*., 2019). What makes this bacteria so amenable to HGT is its large, linear chromosome (ranging from 8.7-10.1 Mb), with its essential genes located in an inner core region. This means that the ends of which (often defined as contingency/nonessential regions) are more susceptible to genetic instability (Bentley *et al*., 2002). In the review by Choulet *et al*, they examined the coding DNA sequences at these ends between three *Streptomyces* species, coming to the conclusion that the low levels of conservation between these homologous suggests the introduction of large amounts of foreign genes by HGT (Choulet *et al*., 2006). Furthermore, these hotspots of genetic activity have been shown to contain biosynthetic gene clusters whose products may confer an inhibitory capacity (Tidjani *et al*., 2019).

Figure 5.15 shows the HGTs detected within the *Streptomyces* genus, there were many HGT events, but only six of them had a bootstrap value of >50% and one had an inverse score of >50%. An inverse result is where the potential transfer on the gene from one subtree to another, is actually more likely in the reverse scenario.

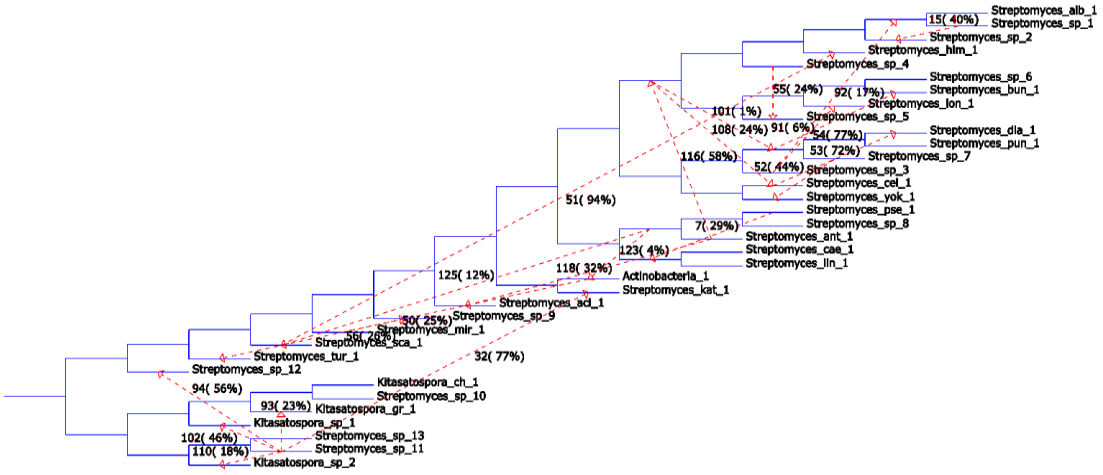


Figure 5.13 – The HGT movement of the *MINPP* gene in the *Streptomyces* cluster.

Following *Streptomyces*, the enteric bacteria had the next highest frequency of HGT events, *Bacteroides* (5), *Prevotella* (4) and *Bifidobacterium* (2). In the gastrointestinal tract HGT is extremely common amongst bacteria, particularly in the case of the transfer of genes encoding antibiotic resistance (Lerminiaux and Cameron, 2019). From this, it seems like the *MINPP* gene is frequently transferred between bacteria of the same genus and likely provides a useful benefit to the host, as the gene persists in the genome.

In this next section, the HGT events between different bacterial genus was examined. The first HGT event with a bootstrap value of 97% was from subtree *Andreprevotia lacus* to subtree *Methylomagnum ishizawai*, Figure 5.16.

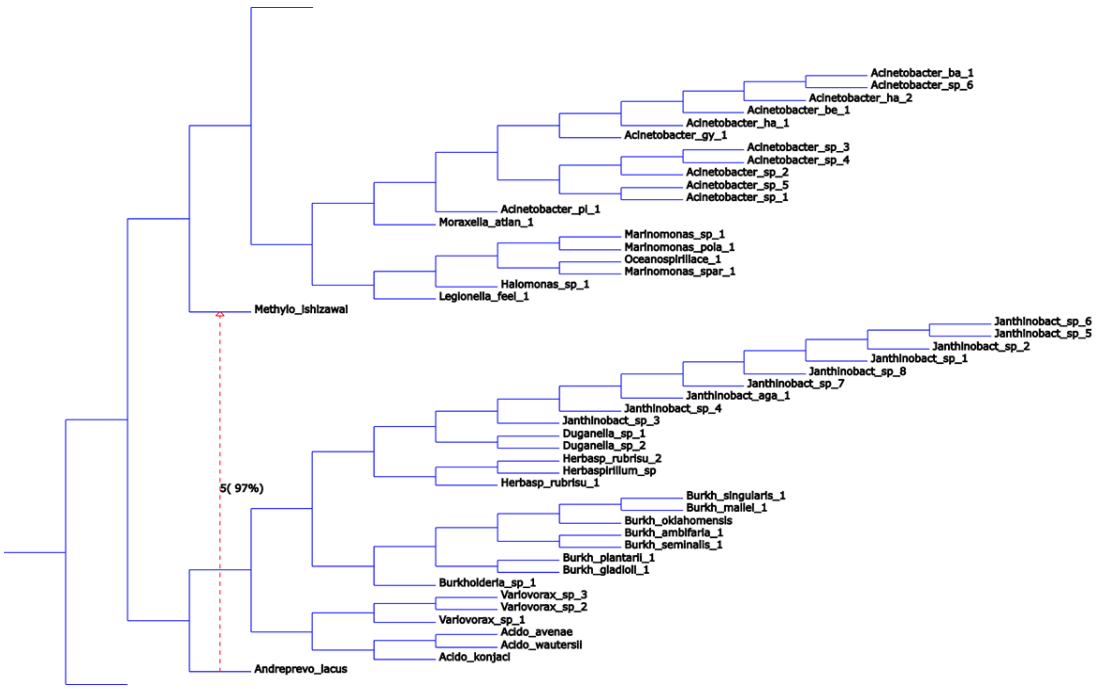


Figure 5.16 – The HGT of the *MINPP* gene from subtree *Andreprevotia lacus* to substree *Methylomagnum ishizawai*, with a bootstrap score of 97%.

Andreprevotia lacus was first isolated from a fish-culture pond in Taiwan becoming the second member of the family *Neisseriaceae*, which had only encompassed one species, *Andreprevotia chitinilytica* which was isolated from Halla Mountain Forest soil (Sheu *et al*., 2009). *Methylomagnum ishizawai*, a methane-oxidising bacteria, on the other hand was isolated from the rhizosphere of rice in a Japanese rice paddy field (Khalifa *et al*., 2015), this can also be an aquatic environment as rice fields are frequently flooded as a means of killing off weeds whilst rice are tolerant of anoxic environments, interestingly, many farmers also put small fish in their fields which help by eating some of the insects and pests as well as providing fertiliser in the form of excrements (Fernando, 1993). It is, therefore, not a surprising result that bacteria inhabiting similar environments may pass the *MINPP* between themselves, even if they are relatively unrelated bacteria on the 16S rRNA tree. *Andreprevotia lacus* is also involved in another HGT event, with a bootstrap value of 92% we can see the potential transfer from subtree *Andreprevotia lacus* to subtree *Variovorax* sp 1-3, Figure 5.17.

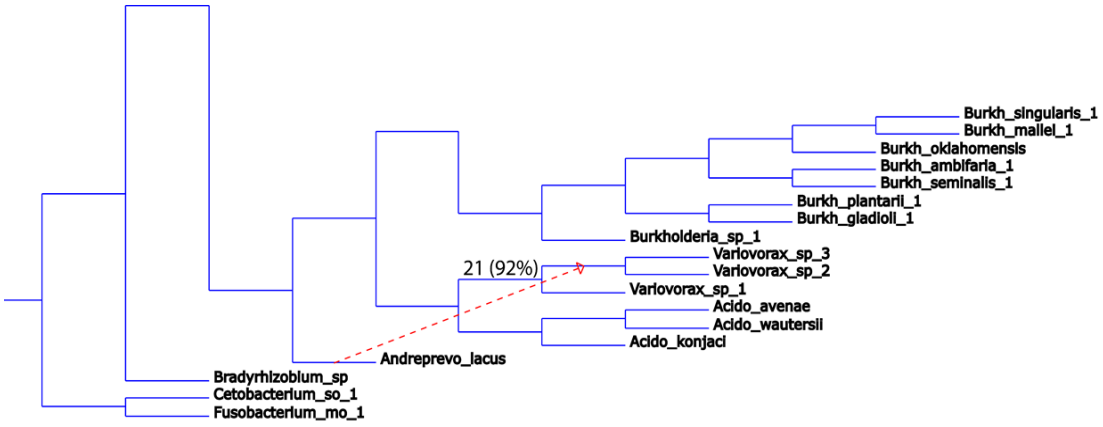


Figure 5.15 – The HGT of the *MINPP* gene from subtree *Andreprevotia lacus* to subtree *Variovorax* sp 1-3, with a bootstrap score of 97%.

Variovorax sp. is another bacterium that are common inhabitants of soil and water environments (Ghio *et al*., 2012; Robertson *et al*., 2018; Satola *et al*., 2013). Curiously, *Variovorax* sp. is associate with a large variety of diverse and uncommon metabolic features such as isoprene oxidation, linuron degradation, aromatic catabolism and polymer degradation, reports have detailed that many of these processes may have been acquired through horizontal gene transfer (Dawson *et al*., 2020; Öztürk *et al*., 2020; Satola *et al*., 2013), which suggests high genetic tractability, therefore, it is not unusual that *Variovorax* sp. has taken up the *MINPP* gene.

The following HGT event occurs from subtree *Arthrobacter* sp_1 to subtree *Microbacterium hydrocarbonoxydans*_2, with a bootstrap value of 89%, Figure 5.18.

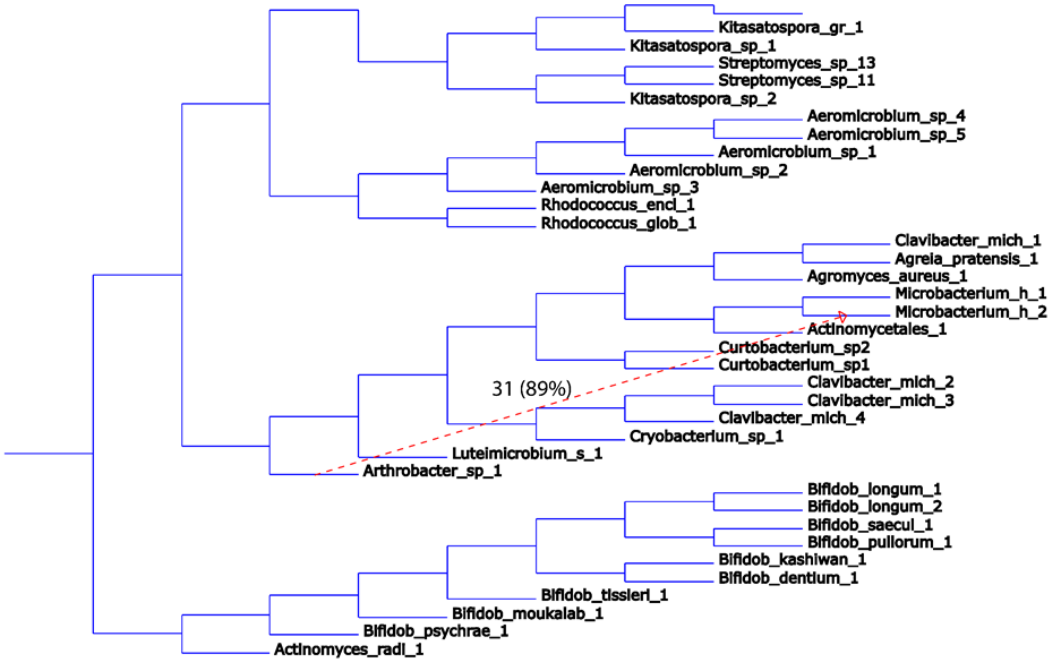


Figure 5.18 – The HGT of the *MINPP* gene from subtree *Arthrobacter* sp_1 to subtree *Microbacterium hydrocarbonoxydans*_2, with a bootstrap value of 89%.

Arthrobacter sp. is an aerobic bacteria from the actinobacteria phylum, that is commonly found in soils, the surface of plants, and wastewater sediments (Gobbetti and Rizzello, 2014). They possess a broad range of interesting physiological functions such as catabolism of organic pollutants and pesticides, metal resistance and extreme resistance to drying (Jerke *et al*., 2008), isolates have also been found in oil brines at depths of 200-700 m (Gobbetti and Rizzello, 2014). *Microbacterium hydrocarbonoxydans* is described as a novel crude oil degrading bacterium (Schippers *et al*., 2005), members of the *Microbacterium* genus have been shown to survive in heavy metal contaminated environments suggesting these strains can survive in heavily contaminated environments. The report by Corretto *et al.,* (Corretto *et al*., 2020) highlighted that the pan-genome of *Microbacterium* is open, a feature that is usually associated with organisms in a community where there is a high rate of horizontal gene transfer. Both of these bacteria can live in contaminated environments and therefore transfer of the *MINPP* gene is unsurprising.

The next HGT event occurs from subtree *Bacteroides coprocola 1-2* and *Bacteroides* sp. 4-6 to subtree *Mediterranea* sp*.,* with a bootstrap value of 83%, Figure 5.19.

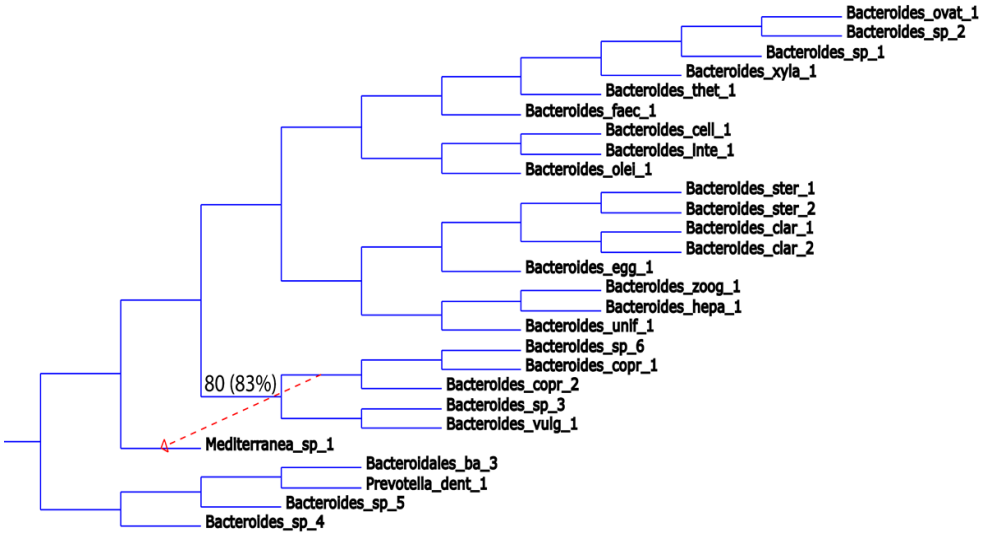


Figure 5.19 – The HGT of the *MINPP* gene from subtree *Bacteroides coprocola 1-2* and *Bacteroides* sp. 4-6 to subtree *Mediterranea* sp., with a bootstrap value of 83%.

This HGT event occurs to be between two members of the Bacteroidaceae family, *Bacteroides coprocola* and *Mediterranea* sp. *Mediterranea* sp. was isolated after whole genome sequencing of gut anaerobes isolated from chicken caecum (Medvecky *et al*., 2018) and has also been identified in the human colon (Mailhe *et al*., 2016). The enteric environment is one in which HGT is common, and in some cases deadly, as the human gastrointestinal tract provides an ideal environment for antibiotic resistance genes to arise and spread throughout the bacterial population (Huddleston, 2014). As seen in Figure 5.5 and 5.6, *Bacteroides* sp. are dominant MINPP carriers and as such, the transfer of the gene to *Mediterranea*, a fellow enteric bacterium is likely.

Another HGT event occurs between subtree *Curtobacterium* sp. 1-2 to subtree *Microbacterium hydrocarbonoxydans_2* with a bootstrap value of 57%, Figure 5.20.

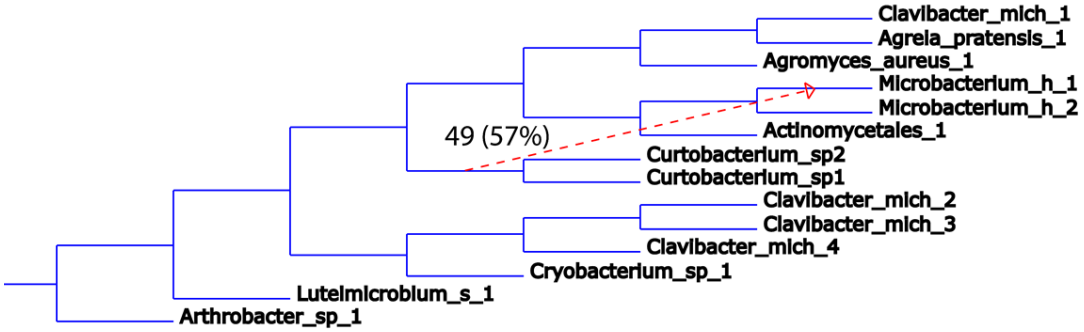


Figure 5.20 – The HGT of the *MINPP* gene from subtree *Curtobacterium* sp. 1-2 to subtree *Microbacterium hydrocarbonoxydans_2* with a bootstrap value of 57%.

Curtobacterium sp. is another bacterium from the Acintobacteria phylum, with a predominant presence in soil ecosystems globally (Chase *et al*., 2016), often focused on for its role as a plant pathogen. *Microbacterium hydrocarbonoxydans* has been described above as a strain that can survive in heavy metal and other contaminated environments, similarly, *Curtobacterium* sp. have also been described to be resistant to heavy metals (Egamberdieva, 2018) and was demonstrated to be capable of removing both Ni(II) and Pb(II) from aqueous environments (Masoumi *et al*., 2016). Therefore, this is another fairly typical HGT event.

The next HGT event occurred between subtree *Duganella* sp. to *Janthinobacterium* sp_4 with a bootstrap value of 100%, Figure 5.21.

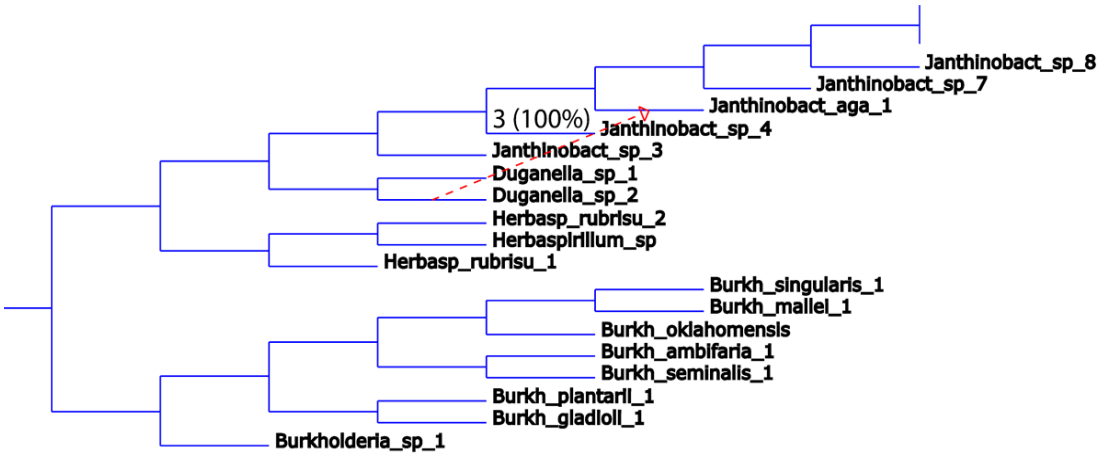


Figure 5.21 – The HGT of the *MINPP* gene from subtree *Duganella* sp. to subtree *Janthinobacterium* sp_4, with a bootstrap value of 100%.

Duganella, was first isolated from sewage and polluted water and has since been isolated from the rhizosphere (Aranda *et al*., 2011; Tanner *et al*., 1998). Its main interest is in the production of Violacein, a natural purple pigment that has a variety of biological activities including, anti-bacterial, anti-viral, anti-tumor and anti-inflammatory properties (Venegas *et al*., 2019). *Janthinobacterium* is closely related to *Duganella* as shown by Figure 5.21, and is also associated with violacein production (Pantanella *et al*., 2006). The HGT event between these two species is highly likely with a 100% bootstrap score.

A further HGT event occurred between subtree *Dysgonomonas macrotermitis_1* to subtree *Flavobacterium hydatis*, *Flavobacterium oncorhynchi*, *Flavobacterium* sp., 1-3 with a bootstrap value of 86%, Figure 5.22.

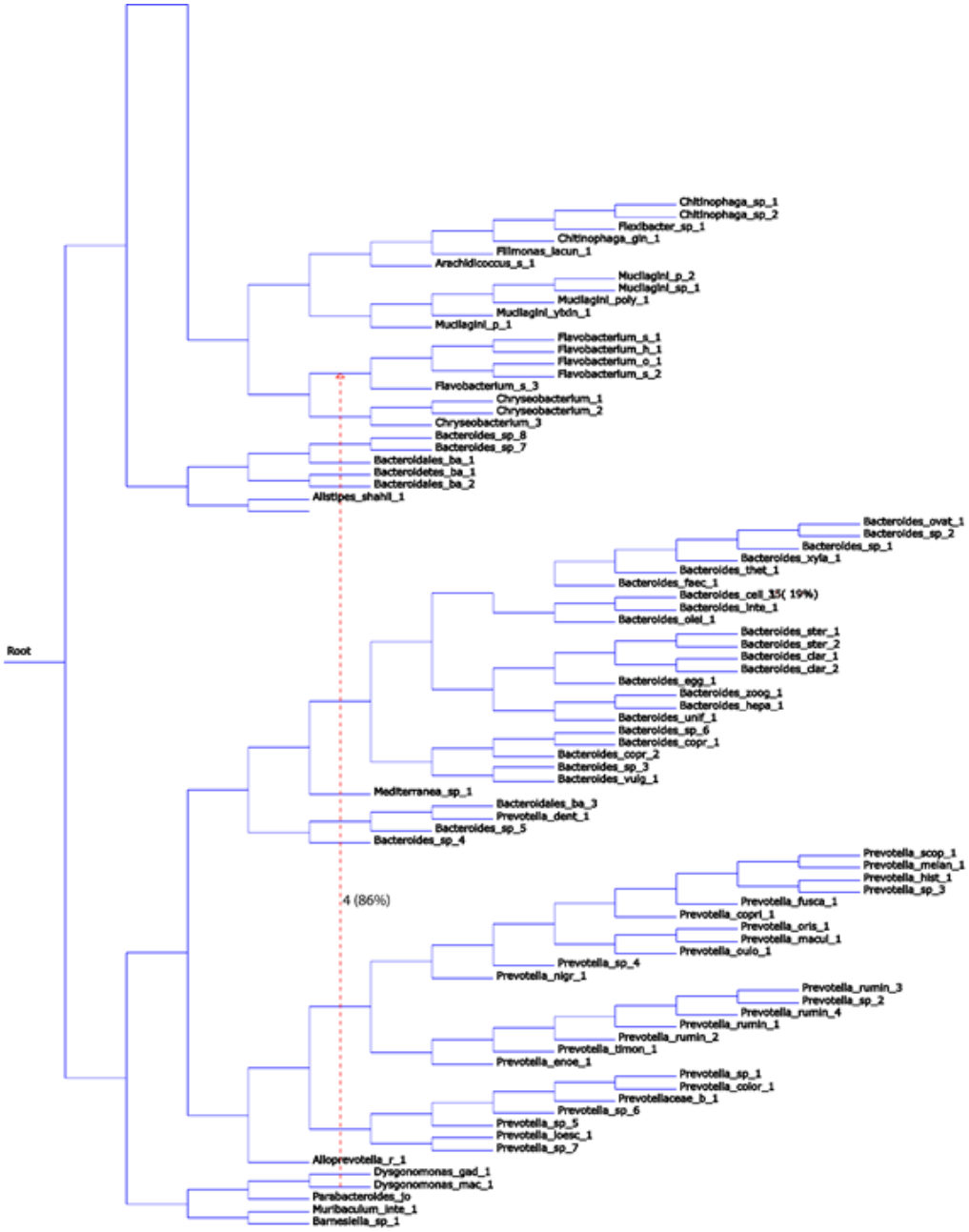


Figure 5.22 - The HGT of the *MINPP* gene from subtree *Dysgonomonas macrotermitis_1* to subtree *Flavobacterium hydatis*, *Flavobacterium oncorhynchi*, *Flavobacterium* sp 1-3 with a bootstrap value of 86%.

Dysgonomonas macrotermitis is a gram-negative, anaerobic bacteria that was isolated from the hindgut of a fungus growing termite, *Macrotermes barneyi* (Yang *et al*., 2014). *Flavobacterium* strains meanwhile are widely distributed in nature, and have been isolated in various habitats such as both fresh and saltwater sediments, soil, microbial mats and in

diseased fish (Waśkiewicz and Irzykowska, 2014), unlike with *Dysgonomonas* they are strictly aerobic bacteria. This HGT event may be the first instance of a *MINPP* being transferred from an enteric bacterium to one with a more environmental niche. This is supported by a review of many of the recognised species in the genus *Flavobacterium* (Bernardet and Bowman, 2006), many of which are present in soil and waterways as well as in fish, additionally, none of the enteric or monogastric metagenomes examined in Figure 5.5 and 5.6 contained hits relating to this genus of bacteria. It can therefore be tentatively inferred that the *MINPP* was passed from *Dygonomonas macrotermitis* from the hindgut of a termite to *Flavobacterium* in the soil environment, indeed termites uses soil together with saliva and faeces to construct their nests.

An additional HGT event occurred between subtree *Halomonas* sp_1 to subtree *Paenibacillus* sp_1 and *Paenibacillus xylanexedens* with a bootstrap value of 65%, Figure 5.23.

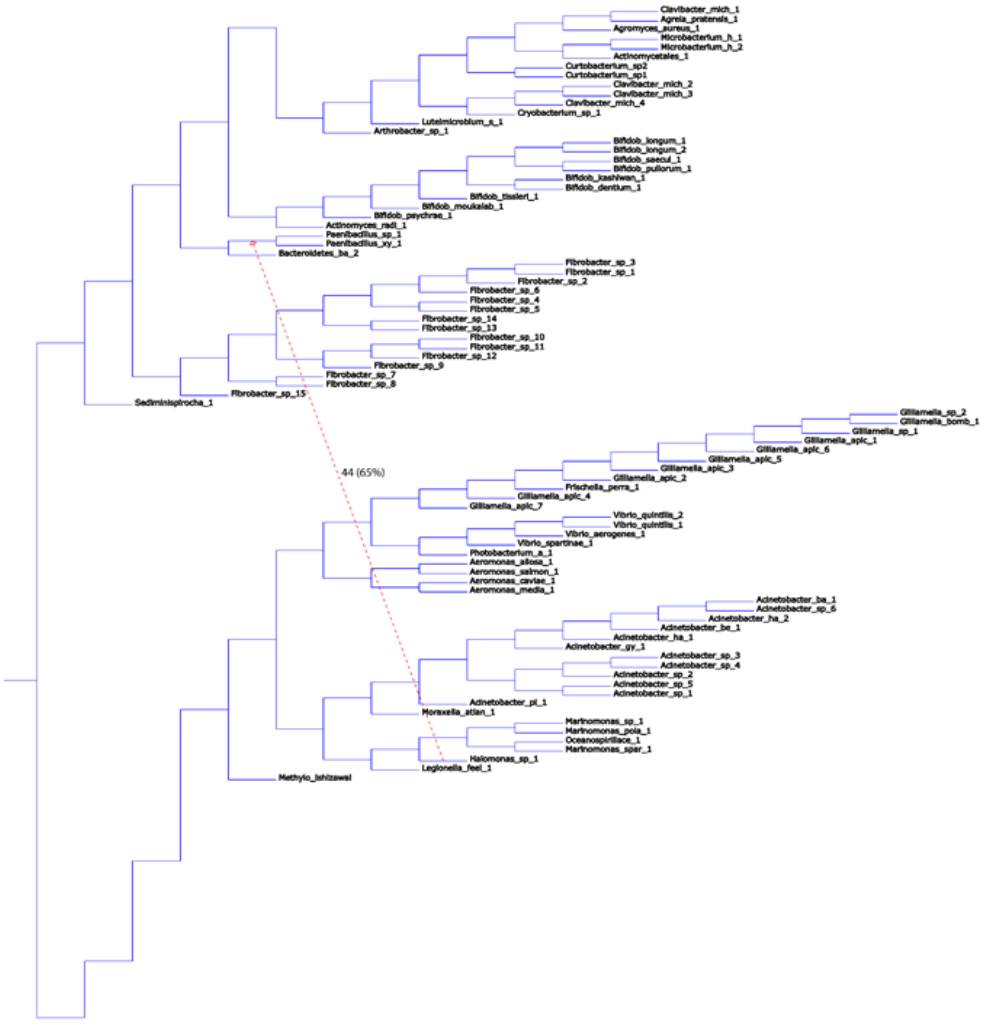


Figure 5.23 - The HGT of the *MINPP* gene from subtree *Halomonas* sp_1 to subtree *Paenibacillus* sp_1 and *Paenibacillus xylanexedens* with a bootstrap value of 65%.

Bacteria from the *Halomonas* genus are halophilic (salt tolerant) and have been found in a broad variety of saline environments, such as oceans, hypersaline lakes and estuaries (Ventosa *et al*., 2021). *Paenibacillus* (paene – almost) have been isolated from plants and the environment, as well as humans and animals, predominantly however, they are found in the soil (Grady *et al*., 2016). As *Halomonas* species have also been isolated from the soil environment, this HGT event is not significant (Gan *et al*., 2018; Quillaguamán *et al*., 2004).

The next HGT event occurs between subtree *Halomonas* sp_1 to subtree *Sediminispirochaeta smaragdinae* with a bootstrap value of 63%, Figure 5.24.

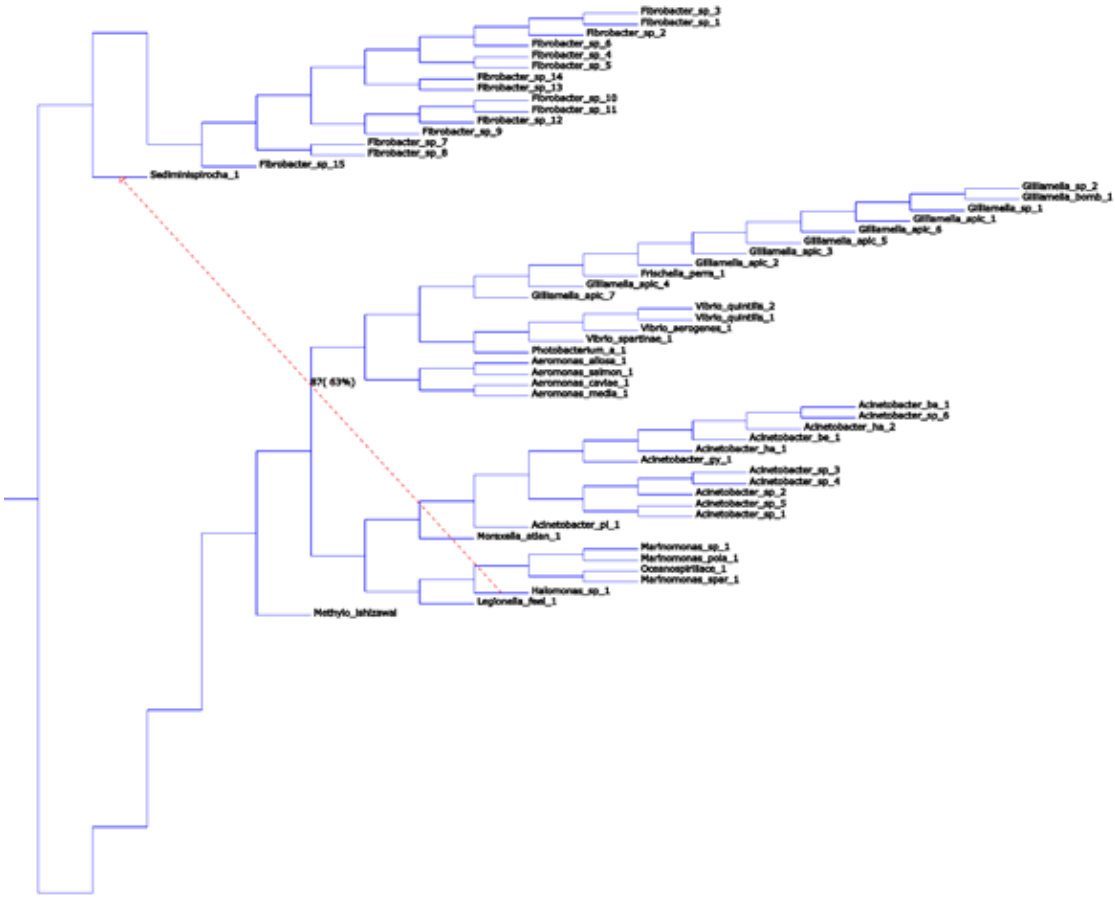


Figure 5.24 - The HGT of the *MINPP* gene from subtree *Halomonas* sp_1 to subtree *Sediminispirochaeta smaragdinae* with a bootstrap value of 63%.

Originally known as *Spirochaeta smaragdinae*, this bacteria has since be reclassified to a new genus and is now known as *Sediminispirochaeta smaragdinae* (Shivani *et al*., 2016). This strain was first isolated from an oil-injection production water sample off the coast of a Congo oilfield (Magot *et al*., 2006), it is characterised by being a gram-negative, strict anaerobe that can survive halophilic environments (1-10% NaCl) (Mavromatis *et al*., 2010). It is therefore found in similar environments to *Halomonas* sp_1 and as such the HGT event

is not significant. Interestingly this is the second HGT event that *Halomonas* sp_1 which may suggest a readiness of the genus to both donate and receive genetic information. Examining the literature suggests that genes encoding for carbohydrate transport and metabolism, nitrogen metabolism and heavy metal resistance may have been taken up by the bacteria from the genus supporting this theory (Balderrama-Subieta and Quillaguamán, 2013; Guzmán *et al*., 2012; Huo *et al*., 2014).

The next HGT event was between subtree *Marinomonas* sp_1 to subtree *Vibrio quintillis* and *Vibrio spartinae* with a bootstrap value of 79%, Figure 5.25.

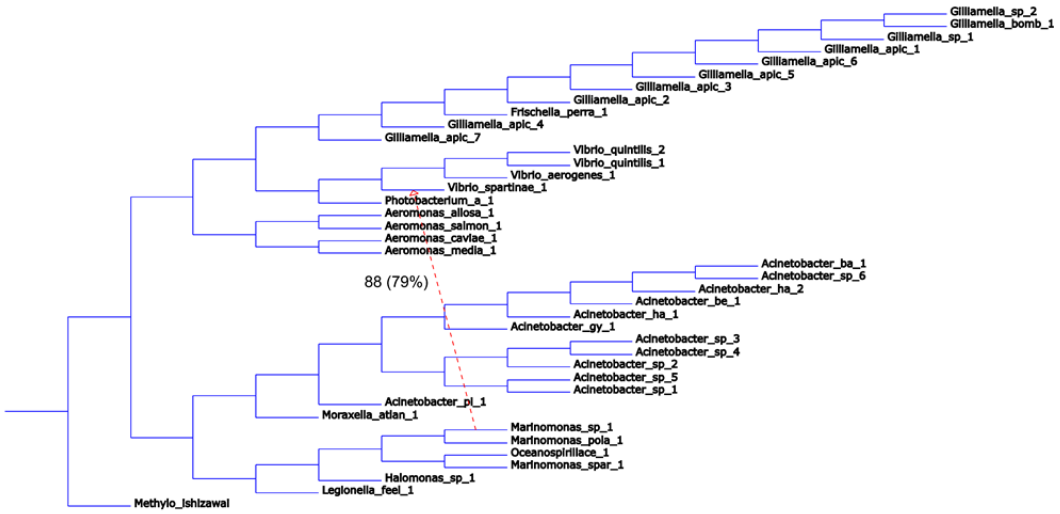


Figure 5.25 - The HGT of the *MINPP* gene from subtree *Marinomonas* sp_1 to subtree *Vibrio quintillis* and *Vibrio spartinae* with a bootstrap value of 79%.

Marinomonas sp. are aerobic bacteria from the family of Oceanospirillaceae that are characterised by their halophilic nature, living in predominantly marine environments. (Ivanova, 2005; Kang *et al*., 2012). Bacteria of the *Vibrio* genus are an opportunistic human pathogen that are typically found in a wide variety of aquatic and marine habitats (Baker-Austin *et al*., 2018). Infections are due to exposure to contaminated water or the consumption of undercooked contaminated seafood, one of the main infection causing bacteria is *Vibrio cholerae* which causes the disease cholera (Faruque *et al*., 1998). Interestingly bacteria from the *Vibrio* genus are reported to possess two circular chromosomes (Trucksis *et al*., 1998), a larger chromosome which is relatively constant amongst them and a smaller chromosome which is more varied (Okada *et al*., 2005). Le Roux and Blokesch (Le Roux and Blokesch, 2018), have reviewed the extent of HGT in the *Vibrio* genus, and its role in pathogen emergence. Overall, they highlighted features that enhance HGT events such as superintegron islands, gene capturing platforms that can incorporate exogenous DNA and chitin, a polysaccharide that induces competency, as well as identifying

horizontally acquired regions. Therefore, it is likely that bacteria from the *Vibrio* sp. would take up the *MINPP* gene from *Marinomonas* sp. as both bacteria live in similar environments and *Vibrio* sp. have the necessary mechanisms for accepting and incorporating the gene.

The next HGT event was between subtree *Muribaculum intestinale* to subtree *Prevotella rumincola_3-4* with a bootstrap value of 50%, Figure 5.26.

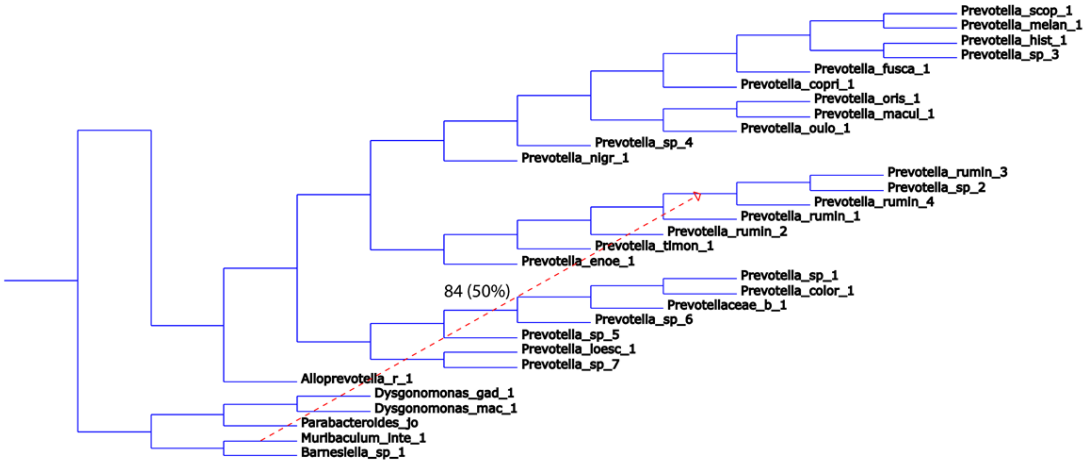


Figure 5.26 - The HGT of the *MINPP* gene from subtree *Muribaculum intestinale* to subtree *Prevotella rumincola_3-4* with a bootstrap value of 50%.

Muribaculum intestinale is a strictly anaerobic bacterium that is part of the Muribaculaceae family which is known to be a dominant bacterial group in the mouse gut (Lagkouvardos *et al*., 2016) and is also known to be present in the rumen microbiota (Islam *et al*., 2021). The *Prevotella* genus constitutes one of the most dominant bacterial groups in the rumen (Kim *et al*., 2017), *Prevotella rumincola* plays a significant role in the metabolism of proteins and peptides. Both of these bacteria inhabit similar environments and therefore the crossover event is not significant.

Finally, the next HGT event was between a bacterium deeply rooted within the *Streptomyces* subtree to subtree *Luteimicrobium subarcticum* with a bootstrap value of 62%, Figure 5.27.

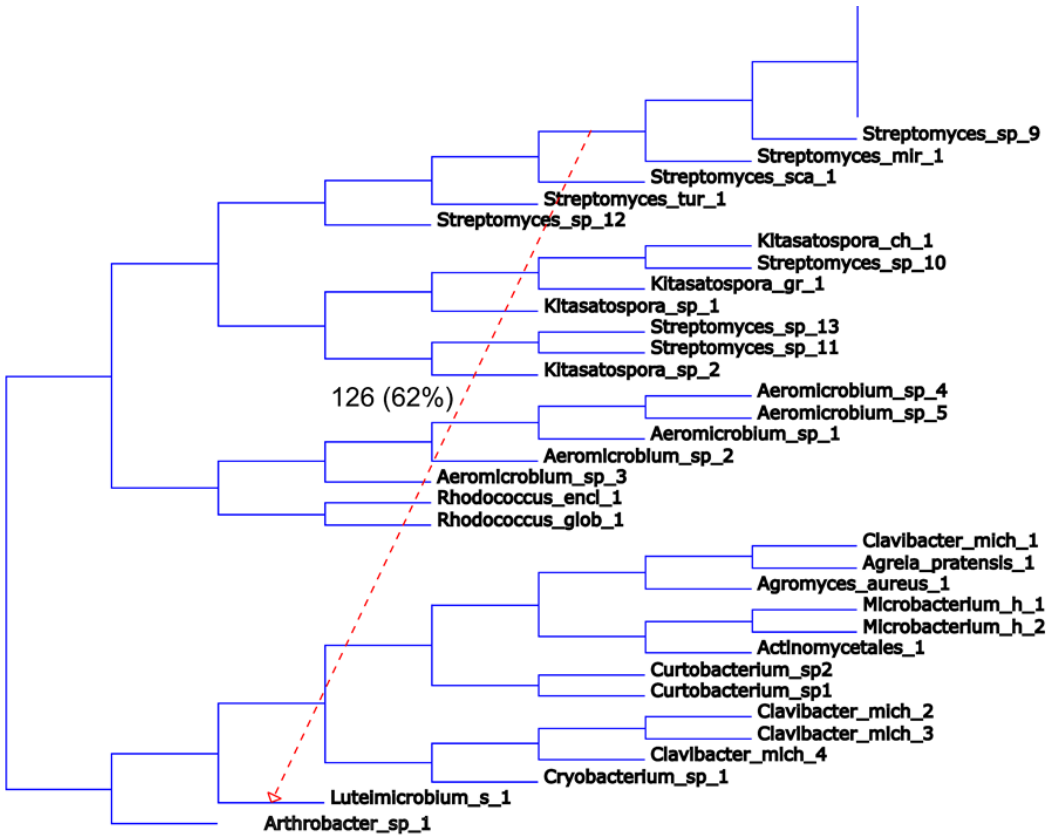


Figure 5.27 - The HGT of the *MINPP* gene from deep within the *Streptomyces* subtree to subtree *Luteimicrobium subarcticum* with a bootstrap value of 62%.

Similar to *Streptomyces* which have already been discussed in this chapter, *Luteimicrobium subarcticum* is a gram-positive bacteria of the Actinobacteria that was first isolated from a soil sample from Rishiri Island, Japan (Hamada *et al*., 2010), as both strains are similarly related and present in the same environment this is not an uncommon result.

Overall, nearly all of the HGT events between different bacteria occur in the same environment between bacteria that are typically closely related to one another. The most interesting HGT event examined was from *Dysgonomonas macrotermitis* to *Flavobacterium* strains, an enteric environment to a more natural one. Based on the metagenomic data highlighted in Table 5.4, the abundance of *MINPPs* is substantially higher in enteric and ruminant environments, therefore animals with a close connection to the soil environment, through the spread of faeces may also cause the transfer of the *MINPP* gene to the environment.

5.5. Conclusion

This study represents one of the first instances that the *MINPP* gene has been examined throughout a wide range of metagenomes. Here, the diversity and abundance of the gene was established through metagenomic hits placed onto a phylogenetic tree that had been built through generation of a consensus sequence alignment. The MINPP is especially abundant in enteric environments with as much as 16-36.2% of all bacterial species in the rumen containing a MINPP gene, highlighting the importance of this gene in phytate degradation of the rumen. Despite this abundance, discussions of the MINPP genes in the rumen and guts of monogastric animals is extremely rare with less than 10 identifiable literature reports (Per GoogleScholar) at least mentioning MINPP genes in enteric environments. Therefore, there is much to come on the study of the MINPP phytase.

Following this, the genetic history of the MINPP gene was examined using the program T-rex which predicts potential horizontal gene transfer events based on the co-evolution of the 16S rRNA gene tree in comparison with the *MINPP* gene. Forty-one HGT events were predicted by the software with a bootstrapping score over 50%, predominantly, such transfers typically occurred between bacteria of the same genus. However, twelve of these transfer events occurred between different bacterial genera, each event of which has been highlighted from Figures 5.14-5.25, however these events were also between similar bacteria from similar environments. Only one HGT event, Figure 5.20 provided evidence of a HGT event between the enteric environment of a hindgut of a termite, and the soil. Due to the sheer abundance of MINPP genes in enteric environments and relatively low abundance in environmental metagenomes there may be a pathway in which MINPP genes flow from the enteric environment into the natural environment.

Chapter 6. Control of phytase expression in *Acinetobacter* sp.

The expression and controls of phytase production is currently poorly understood and does not appear to be consistent across bacteria (Jain *et al*., 2016; Konietzny and Greiner, 2004). In this chapter, the controls of phytase expression in *Acinetobacter* sp. AC1-2 were examined through promoter-reporter assays and qPCR.

Promoter-reporter assays provide opportunity to identify factors or conditions under which transcriptional control is influenced. They allow identification of sequences of DNA that encode promoters and/or repressors, even without precise identification of the response element. They also allow identification of conditions under which particular genes or whole operons are transcriptionally up- or down-regulated. This is a consequence of the juxtaposition of response element and gene. In one context, the cloning of a putative promoter region, from genomic DNA upstream of a gene of interest, in frame with a heterologous reporter gene, allows testing of transcription and translation combined. In this Chapter, *Acinetobacter* sp. AC1-2 was transformed with a plasmid encoding the *lacZ* gene encoding β-galactosidase, downstream of a region of genomic DNA cloned from *Acinetobacter*. In this assay, the transformed *Acinetobacter* cells were grown under the different concentrations of phytate and/or phosphate, after which, the cells are lysed and enzyme activity as proxy of gene transcription is measured against *ortho*-nitrophenyl-β-galactopyranoside (ONPG) substrate. The induction or repression of the promoter region determines the amount of β-galactosidase transcribed and translated. The enzyme assay is based on the production of nitrophenolate ion from ONPG, the product has yellow colour at an alkaline pH. The intensity of colour correlates to the amount of β-galactosidase produced and thus quantifies promoter strength (Smale, 2010). Additionally, qPCR was performed to examine the expression of the MINPP phytase in *Acinetobacter* sp. AC1-2 in rich and minimal media in the presence of phytate and phytate/phosphate.

6.1. Materials and Methods

6.1.1. Generation of β-galactosidase Promoter-Reporter constructs for phytase expression

To produce the promoter-reporter constructs, the program Artemis (Carver *et al*., 2012) was used as a genome browser and annotation tool to allow the visualisation of sequence features. In this instance, it was used to identify the location of the AC1-2 *MINPP* gene in the chromosome of *Acinetobacter* sp. AC1-2 and to see whether the gene lies within any cluster of genes, that might be recognizable in other MINPP-bearing bacteria. In this instance, AC1-2 *MINPP* appeared to be part of a cluster and as the location of the promoter was unknown, three promoters-reporter constructs were prepared, Figure 6.1.

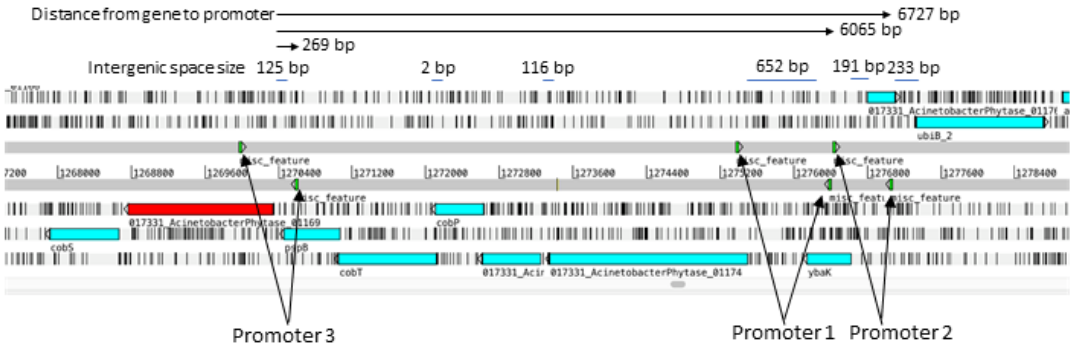


Figure 6.1 – Generation of the promoter region primers through visualisation of the *Acinetobacter* sp. AC1-2 genome using the program Artemis, highlighting the location of the *MINPP* gene (red) amidst the other genes in the genome (blue). Three potential promoter regions have been identified (green arrows), Promoters 1 and 2 were located after the gene cluster, Promoter 3 just before the *MINPP* gene.

Primer sets were designed for each of these promoter regions, Table 6.1.

Table 6.1 – Primer sets used to clone potential promoter regions of the *MINPP* gene. A spacer of GATC, underlined, was added before each restriction site, highlighted yellow. The restriction site GAATTC was used for each forward primer, encoding a EcoR1 cut site, the restriction site CTGCAG was used for each reverse primer, encoding a Pst1 cut site.

Primer Set 1	F- 5'-GATCGAATTCGTAGCTACGGCGTTTGGATG-3' R- 5'-GATCCTGCAGCCTGTCCAATGAGTTCTTGAG-3' Size – 1029 bp
Primer Set 2	F- 5'-GATCGAATTCCAATGATGATGCTAGGAAGG-3' R- 5'-GATCCTGCAGCCGGCAAGACAGCAACG-3' Size – 636 bp
Primer Set 3	F- 5'- GATCGAATTCCCATGGCGGTGTCATTAAGC-3' R- 5'- GATCCTGCAGCCTAACAGAATATTGGCTT-3' Size – 634 bp

A β–galactosidase construct was made encoding Promoter 3 and transformed into *Acinetobacter* sp. AC1-2 the construct however, showed low/no β -galactosidase (β-gal) activity, and therefore constructs were designed with Promoters 1 and 2 and transformed into *Acinetobacter* sp. AC1-2.

To do this, *Acinetobacter* sp. AC1-2 cells from a glycerol stock were streaked onto LB and grown overnight at 37 °C. Single colonies were mixed with 100 µL sterile dH_2O and heated for 5 minutes at 95 °C to lyse the cells. Using this as a template, promoter regions were amplified using the high-fidelity polymerase Phusion® as described in Section 3.1.4. The PCR products were purified using the QIAquick PCR purification Kit (QIAGEN)

6.1.2. Ligation of promoter regions with the pS4 vector

Insert Digestion

The purified amplicon with the promoter region insert was digested using the EcoR1 and Pst1 restriction enzymes using the following protocol: 16 µL purified PCR product, 2 µL restriction enzyme buffer (Roche buffer H), 1 µL restriction enzyme *Pst1* (Roche), 1 µL restriction enzyme *EcoR1* (Roche). The microfuge tube was lightly flicked to mix the components and briefly spun down, before being incubated at 37 °C for two hours.

Vector Digestion

The purified vector pS4 which does not contain a presentable plasmid map, derived from pmp220 (Spaink *et al*., 1987) was also digested using the EcoR1 and Pst1 restriction enzyme using the following protocol: 1 µL pS4 vector (approximately 50-200 ng/µL), 2 µL restriction enzyme buffer (Roche buffer H), 1 µL restriction enzyme Pst1 (Roche), 1 µL restriction enzyme EcoR1 (Roche), 15 µL dH_2O. The microfuge tube was lightly flicked to mix the components and briefly spun down, before being incubated at 37 °C for 2-4 hours.

After 2 hours, all of the digested PCR product and 1 µL of the vector digest (the remaining digest was left to continue digesting) were run on a 1% agarose gel. The PCR digests were then purified from the gel using the QIAquick Gel Extraction Kit (QIAGEN) as per the manufacturer's protocol.

Vector Dephosphorylation

The vector digest was examined to ensure that the vector had been digested, afterwards the remaining vector digestion was heated at 80 °C for 10 minutes to inactivate the enzymes. The vector was dephosphorylated to prevent re-circularisation during the ligation reaction:

19 µL (heat-killed PCR digest), 2.5 µL phosphatase buffer, 2.5 µL dH_2O and 1 µL phosphatase (Thermofisher Scientific) enzyme were incubated at 37 °C for 30 minutes.

Ligation

Both the digested promoter region and digested vector were combined in a ligation reaction to produce the pS4 plasmid containing the promoter region with the heterologous reporter gene. This reaction used the following protocol: 2 µL dephosphorylated vector, 15 µL gel extracted promoter digest (15 µL dH_2O for religated controls), 2 µL T4 ligase buffer, 1 µL T4 ligase (Promega), incubated on ice overnight, with the ice allowed to melt increasing the temperature of the reaction.

Transformation

The ligation reactions were transformed into *E. coli* 803 cells, as per Section 3.1.6, and streaked onto a LB plate with the antibiotic spectinomycin, 200 µg/mL final concentration, and grown overnight at 37 °C. Alongside this, control plates of the uncut vector, digested vector, re-ligated vector, and cells only helped to confirm the efficacy of the transformation.

Confirmation

Successful transformants were determined using colony PCR and sequencing to ensure the promoter insert was successful in the pS4 vector. Promoters 1 and 2 were examined by colony PCR using their appropriate primer sets. Promoter 3 was examined by digestion of the pS4 vector with two bands corresponding to the vector and Promoter insert.

6.1.3. Preparation of competent *E. coli* 803 cells

LB (5 mL) was inoculated with *E. coli* 803, using a single colony from a fresh plate, and grown overnight at 37 °C at 180 RPM. From this, 1 mL of culture was used to inoculate 100 mL LB and grown at 37 °C shaking until the OD_{600} reached approximately 0.3-0.4. To two sterile universals, 40 mL of culture was aseptically transferred, the cells were pelleted at 3750 g for 10 minutes and kept on ice as the supernatant was removed carefully. The cells were resuspended in 10 mL ice-cold 0.1 M $CaCl_2$ and left on ice for 30 minutes, pelleted again at 3750 g, the supernatant removed, and the cells resuspended in 2 mL ice-cold 0.1 M $CaCl_2$. Cells were stored overnight at 4 °C overnight and used for transformation.

6.1.4. Triparental Mating – Patch Cross

Promoter: reporters are, perhaps, most useful when expressed in the native host. To transfer plasmid from *E. coli* to host, *Acinetobacter* sp., triparental mating is commonly employed. This uses three bacterial strains. Briefly, a plasmid present in the *E. coli* strain (pRK2013 – kanamycin resistance), known as the helper strain, carries a conjugative plasmid that encodes for genes required for conjugation and DNA transfer. This plasmid is transferred to the donor strain, *E. coli* 803 containing the promoter-pS4 plasmid, increasing the transmissibility. Finally, there is the recipient strain, *Acinetobacter* sp. AC1-2, which accepts the promoter-pS4 plasmid.

For the following protocol, resistance to the antibiotic rifampicin was induced in *Acinetobacter* sp. AC1-2 as a means to select between the *E. coli* 803 – pS4 and the *Acinetobacter* sp. AC1-2 that has taken up the pS4 vector.

Generation of a spontaneous Rif resistant *Acinetobacter* strain

A single colony of *Acinetobacter* sp. AC1-2 was inoculated into 10 mL LB and grown overnight at 37 °C with shaking at 180 RPM. Cells were pelleted at 9965 g, 9 mL of the supernatant removed, and the pellet was resuspended. The remaining culture was streaked onto rifampicin LB plates, 20 µg/mL and incubated at 37 °C for three days until growth appeared on the plates. The colonies were then restreaked onto rifampicin LB plates and grown overnight at 37 °C.

Patch Cross method

Ten mL cultures of *E. coli* 803 – pS4, *E. coli* PRK2013 and *Acinetobacter* sp. AC1-2 with the appropriate antibiotics, spectinomycin 200 µg/mL, kanamycin 20 µg/mL and rifampicin 20 µg/mL respectively, were grown overnight at 37 °C. One mL of *Acinetobacter* sp. AC1-2 cells were spun down at 16060 g and the supernatant removed. The pellet was washed in 500 µL 0.9% NaCl three times to remove residual antibiotics with the final spin down resuspending the pellet in 100 µL LB media. Then 0.5 mL of both the *E. coli* 803 and *E. coli* PRK2013 cells were spun down and washed with the same method.

The three strains were plated onto a LB plate with no antibiotics that had been split into three sections, these were grown overnight at 37 °C. The next day, the three strains were plated onto a fresh LB plate and mixed together in the centre of the plate, these were grown overnight at 37 °C. The next day, growth was streaked onto LB plates with spectinomycin and rifampicin antibiotics and grown at 37 °C until single colonies appeared. *Acinetobacter* sp.

AC1-2 colonies that had taken up the plasmid were confirmed by sequencing using the internal vector primers M13F and M13R, Supplementary Information.

6.1.5. β-Galactosidase Assay

Initial testing of each promoter via β-galactosidase assays showed that Promoter 3, the region closest to the *MINPP* gene showed very low β-galactosidase activity, indicating that this was not the location of the *MINPP* promoter region and that instead the gene is likely activated as part of the cluster. Comparisons of the β-galactosidase activity between Promoters 1 and 2 showed that Promoter 1 was the most active, with Promoter 2 less than half as active. Therefore Promoters 1 and 3 were taken forwards, as Promoter 3 which shows basal β-galactosidase activity can be used as a negative control.

Acinetobacter sp. AC1-2:Promoter 1 and *Acinetobacter* sp. AC1-2:Promoter 2, were grown in 10 mL LB overnight at 37 °C, after which their OD_{600} was measured. These were inoculated to an OD 0.025 into fresh 10 mL minimal media supplemented with 1 mM, 500 µM, 20 µM and 0 µM (Impure, Sigma P8810) $InsP_6$ and 1 mM P_i and grown for 8 hours at which cell growth is in the exponential/late exponential phase.

The grown cells were immediately used for the β-Galactosidase assays. A 0.5 mL aliquot of culture was removed and placed in a 2 mL microfuge tue and made up to 1 mL by adding fresh Z-buffer (1 mL of 3 M $Na_2HPO_4.7H_2O$, 0.5 mL 4 M $NaH_2PO_4.2H_2O$, 0.5 mL 1 M KCl, 0.5 mL 0.1 M $MgSO_4.7H_2O$ and 175 µL β-mercaptoethanol, made up to 50 mL with dH_2O) using three replicates for each condition. The cells were lysed by adding 2 drops of chloroform and 1 drop of 0.1% SDS to each microfuge tube and vortexed for 10 seconds and incubated at 28 °C for 5 minutes. Then 0.2 mL ONPG (4 mg/mL in dH_2O, freshly made) was added to each tube, inverted 4-6 times and incubated for 5 minutes at 28 °C to develop a sufficient yellow colour, then 0.5 mL of 1 M Na_2CO_3 was added to stop the reaction. The microfuge tube was centrifuged at 16060 g to remove cell pellet debris and transferred to a cuvette, and the OD_{420} measured. Additionally, 1 mL of cell culture was placed in a separate cuvette and the OD_{600} measured. The units of β-gal activity (Miller Units) were calculated by:

$$Miller\ units = \frac{1000 \times OD_{420}}{t \times v \times OD_{600}}\ (Where\ t = time\ of\ incubation\ and\ v = volume\ of\ cell\ culture\ used\ (mL)$$

6.1.6. Generation of an *Acinetobacter* MINPP gene knockout using a double crossover technique

A single gene knockout of the *MINPP* gene was created in *Acinetobacter* sp. AC1-2 to abolish phytase activity. Gene knockout via homologous recombination is a frequently used technique to generate specific mutants, allowing functional analysis of the gene (Suzuki and Kurosawa, 2017). Homologous recombination is the exchange of genetic material between two strands of DNA that contain similar base sequences (Kowalczykowski *et al*., 1994), using this fundamental process, a truncated version of the *MINPP* was cloned into a pk18mobsacB (kanamycin resistance), Figure 6.2, suicide vector that contained the *sacB* gene which is lethal in the presence of 10% sucrose.

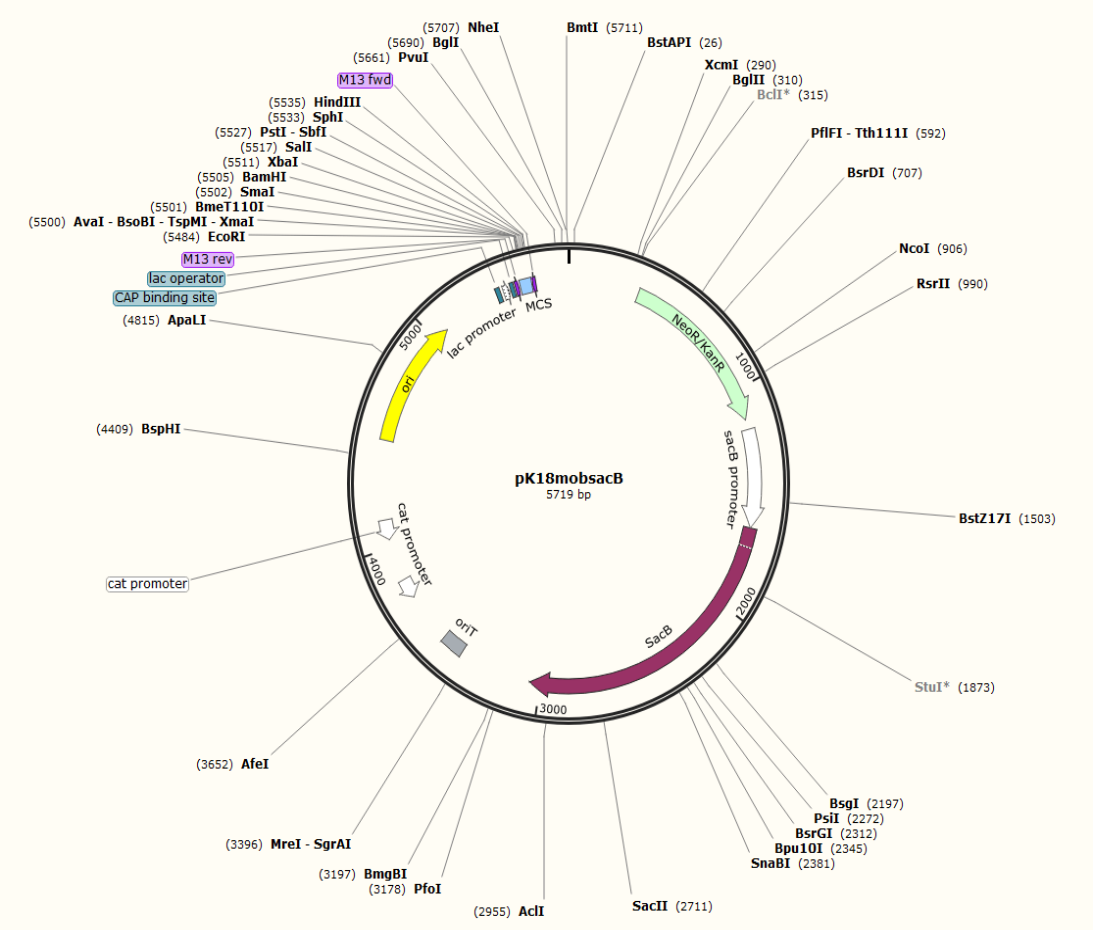


Figure 6.2 - **Plasmid map of pK18mobsacB.** pK18mobsacB is a cloning vector that can integrate genetic information into the host chromosome via homologous recombination. Excision of the plasmid can be facilitrated by selection on medium containing 10% sucrose. Vector maps were made using Snapgene (https://www.snapgene.com).

Therefore in bacteria where homologous recombination has occurred, *SacB* can be used as a counter selectable marker (Pelicic *et al*., 1996). The process is shown in Figure 6.3.

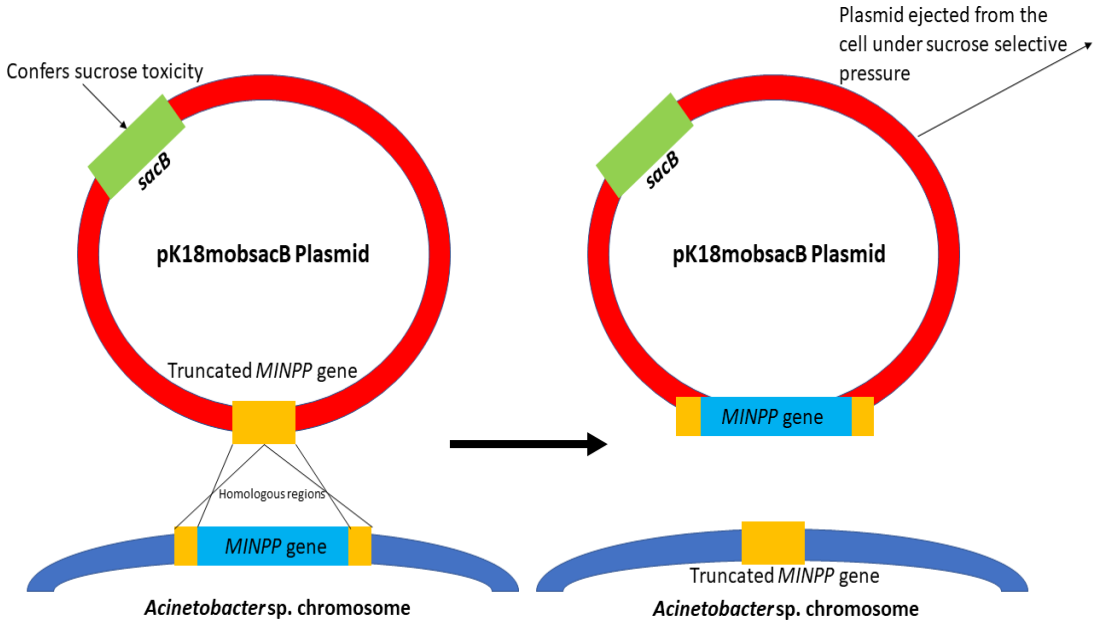


Figure 6.3 – Schematic showing the homologous recombination of a truncated *MINPP* gene replacing the full length *MINPP* gene in the *Acinetobacter* sp. chromosome.

To produce the knockout constructs the program Artemis was used again (Carver *et al*., 2012). Two primer sets were developed to produce a truncated version of the *MINPP* as shown by Figure 6.4, encoding for regions outside of the gene and the ends of it, Primer sets are displayed in Table 6.2.

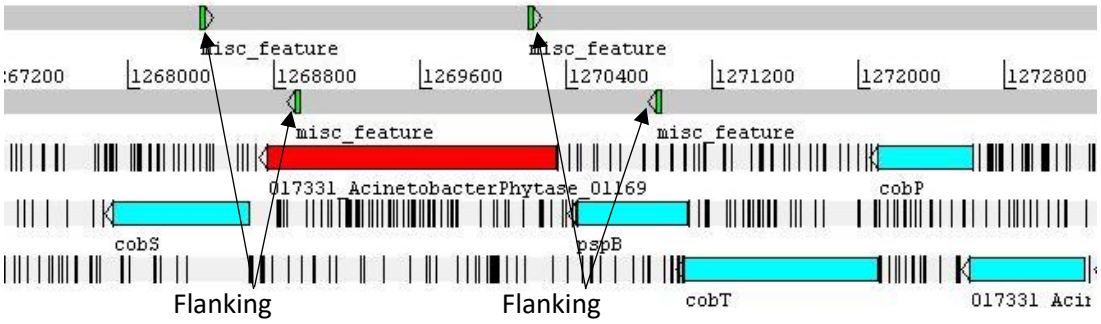


Figure 6.4 – Generation of the knockout primers through visualisation of the *Acinetobacter* genome using the program Artemis, highlighting the location of the *MINPP* gene (red) amidst the other genes in the genome. Two flanking regions made up of the non-coding regions, *cobS* and *pspB* genes, and the ends of the *MINPP* genes were produced to knockout the *MINPP* gene.

Table 6.2 – Primer sets used to clone the flanking regions of the *MINPP* gene for gene knockout. A spacer of GATC, underlined, was added before each restriction site, highlighted yellow. For Flanking region 1, the restriction site GAATTC was used for the forward primer, encoding a EcoR1 cut site, and the restriction site GGATCC was used for the reverse primer, encoding a BamHI cut site. For Flanking region 2, the restriction site GGATCC was used for the forward primer, encoding a BamHI cut site, and CTGCAG for the reverse primer encoding a PstI cut site.

Flanking Region 1 (F1)	F- 5'- GATCGAATTCGGGATGTCATTGTAAGTTC-3' R- 5'- GATCGGATCCGTAAGGCGTTTTGGTTTG -3' Size – 719 bp
Flanking Region 2 (F2)	F- 5'- GATCGGATCCGCTGCCAATGTACAGTGGG -3' R- 5'- GATCCTGCAGCCCACGCATCTGCCGTATC -3' Size – 542 bp

6.1.7. Ligation of flanking regions 1 and 2 into the pK18mobsacB plasmid.

Acinetobacter sp. AC1-2 DNA was prepared as per Section 6.1.1. The two flanking regions were amplified using the high-fidelity polymerase Phusion® as described in Section 3.1.4, and the PCR products were purified using the QIAquick PCR purification Kit (QIAGEN). Restriction digests of the flanking regions and vectors were performed in accordance with Section 6.1.1.1, using the following reactions. pK18mobsacB – EcoR1 and BamHI, Pst1 and BamHI, EcoR1 and Pst1, F1 – EcoR1 and BamHI, Bam and F2 – BamHI and PstI and Bam. These were run on a 1% agarose gel as described and the flanking regions gel extracted, following these digests, the following ligations were performed:

1. F1 + EcoR1:BamHI pK18mobsacB,
2. H_2O + EcoR1:BamHI pK18mobsacB (religated control),
3. F2 + PstI:BamHI pk18mobsacB
4. H_2O + PstI:BamHI pK18mobsacB (religated control).

After which, the ligations were transformed into *E. coli* 803 as described, Section 6.1.1.1. Four single colonies from the F1 + pK18mobsacB plates and eight from the F2 + pK18mobsacB plates were examined using colony PCR to confirm the presence of the flanking regions. To further confirm the ligation of the two flanking regions into the vector, three colonies from each plate were used to inoculate 10 µL of cell suspension in 10 mL LB (20 µg/mL kanamycin) and grown overnight at 37 °C, with shaking at 180 RPM. The plasmid pK18mobsacB DNA was isolated (QIAprep Spin Miniprep Kit – Qiagen) and subjected to restriction enzyme digestion

and agarose gel electrophoresis to confirm for the presence of the vector and flanking region bands. The digestion was allowed to continue and the F1 and F2 regions were gel extracted.

To introduce the other flanking region into the vector, the F1 pK18mobsacB vector was digested with BamHI + PstI, and the F2 pk18mobsacB vector with BamHI + EcoRI, an overview is shown in Figure 6.5.

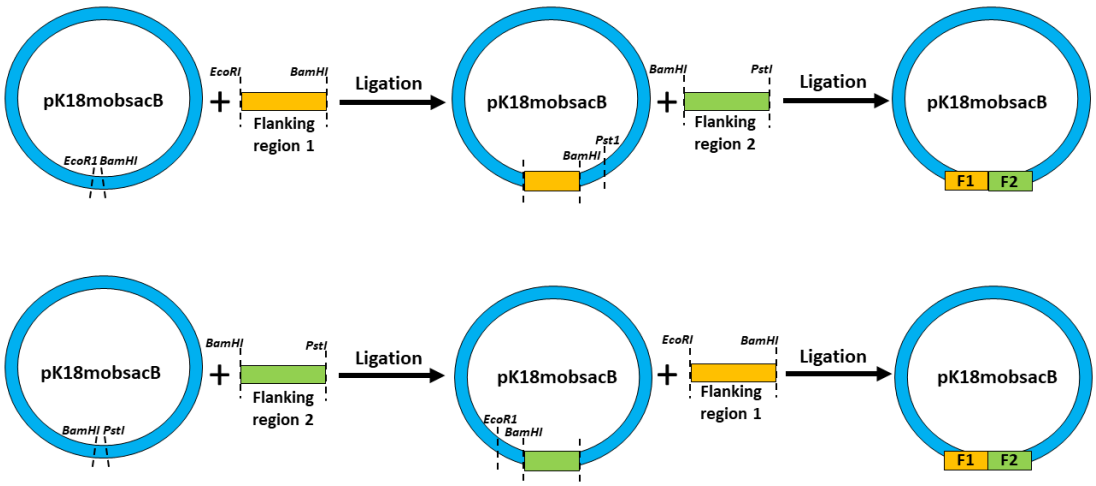


Figure 6.5 – A diagram of how both flanking regions were cloned into the pK18mobsacB vector using restriction enzymes.

The following ligations were performed as described in Section 6.1.1.1 (where S1 refers to Sample 1 from the previous gel):

1. F1S1 Insert (BamHI-EcoRI) – F2S1 vector (BamHI-EcoRI)
2. F1S2 Insert (BamHI-EcoRI) – F2S3 vector (BamHI-EcoRI)
3. F1S3 Insert (BamHI-EcoRI) – F2S6 vector (BamHI-EcoRI)
4. F2S1 Insert (BamHI-PstI) – F1S1 vector (BamHI-PstI)
5. F2S3 Insert (BamHI-PstI) – F1S2 vector (BamHI-PstI)
6. F2S6 Insert (BamHI-PstI) – F1S3 vector (BamHI-PstI)
7. dH_2O + F1S1 vector (BamHI-PstI)
8. dH_2O + F2S1 vector (BamHI-EcoR1)

After overnight ligation, all 8 samples were transformed into *E. coli* 803 with the additional controls stated in the previous section, (cut vector, uncut vector, cells only). These were grown overnight at 37 °C. Colony PCR was performed on 24 single colonies taken from the F1S1 insert – F2S1 vector, F2S3 insert-F1S2 vector and F2S6 insert-F1S3 vector plates using

the F1 forward primer and F2 reverse primer. Of the 24 single colonies examined, 17 showed a band of the correct size c. 1260 bp.

Colonies 3, 12 and 24 were grown in 10 mL LB (kanamycin) overnight and plasmid DNA extracted. Restriction digestion was used to confirm that the two flanking regions had been inserted into the vector using EcoRI, PstI and BamHI. Colonies S12 and S24 showed appropriately sized bands for both flanking regions.

6.1.8. Triparental mating – Filter Cross

The filter-cross method is similar to the patch cross method shown in Section 6.1.1.3, with the exception that 100 µL of the resuspended washed, donor (*E. coli* 803), helper (*E. coli* PRK2013) and host (*Acinetobacter* sp. AC1-2) strains were inoculated onto a sterile filter paper disc (Whatman®) on a LB plate and mixed using a sterile loop. These were left in a 28 °C incubator to dry and then turned over and incubated overnight. The next day, using sterilised forceps, the filter with cell growth on, was removed and placed in a sterile universal. Cells were washed off the filter using 2 mL of 50% glycerol, of which 100 µL was plated onto a LB (rifampicin + kanamycin) plate and grown overnight at 37 °C. Colony PCR was performed on 16 colonies using the F1-F and F2-R primers and on the wild-type bacteria. Nine colonies contained a band around 1260 bp indicating that the plasmid had successful been transformed into the *Acinetobacter* cells, while the control gave a band of c. 2500 bp.

6.1.9. Confirming gene knockout

In order to remove the plasmid containing the WT gene from the bacteria, 5 colonies from the previous PCR (2, 6, 8, 9 and 11) were grown in 10 mL LB (rifampicin) overnight and then 100 µL was plated onto LB rifampicin + 10% sucrose plates and incubated overnight at 37 °C. Colony PCR was performed on 15 single colonies using AC1-2 MINPP specific primers from Section 3.1.4. Successful homologous recombination between the WT gene and the truncated gene should yield, a band of c. 300 bp, while the WT gene should yield a band of 1572 bp. Of the 15 colonies, only 3 gave appropriately sized bands, these colonies were subsequently grown in minimal media with phytate and examined for phytase activity by HPLC. Sample 14 showed no activity towards phytate, indicating a successful knockout and as such was sent for sequencing.

6.1.10. Rif-mutant and MINPP Knockout soil survivability experiment

Following successful transformation an experiment was undertaken to compare the performance in soil of WT-Rif resistant *Acinetobacter* sp AC1-2. and *Acinetobacter* sp. AC1-2 Δ*MINPP*. The experiment was performed in thrice-autoclaved soil. For this, single colonies from WT-Rif resistant *Acinetobacter* sp AC1-2. and *Acinetobacter* sp. AC1-2 Δ*MINPP* were grown in 10 mL LB (Rif) overnight at 37 °C and adjusted to the same OD600 (1.692-1.736). Subsequently, three samples were prepared in Eppendorfs and wrapped in parafilm to examine the survivability and competition of the two strains in autoclaved soil to which phytate, phytate and inorganic phosphate or water were added:

1. 200 µL Rif-*Acinetobacter* sp. AC1-2 + 200 µL *Acinetobacter* sp. AC1-2 Δ*MINPP* + 100 µL dH_2O to 1 g of soil.
2. 200 µL Rif-*Acinetobacter* sp. AC1-2 + 200 µL *Acinetobacter* sp. AC1-2 Δ*MINPP* + 10 mM $InsP_6$ (100 µL) to 1 g of soil.
3. 200 µL Rif-*Acinetobacter* sp. AC1-2 + 200 µL *Acinetobacter* sp. AC1-2 Δ*MINPP* + 10 mM $InsP_6$ (50 µL) + 10 mM P_i (50 µL) to 1 g of soil).

These samples were incubated at room temperature for 7 days, before being serially diluted and streaked onto LB plates, 10^{-1} – 10^{-7} dilutions and incubated at 37 °C.

Single colonies were examined by colony PCR using the Flanking region 1 forward and Flanking region 2 reverse primers for the presence of either the wild-type *MINPP* or truncated *MINPP* gene. In total, 97, 93 and 101 single colonies were used for samples 1, 2 and 3 respectively.

6.1.11. Measurement of phytase expression by quantitative PCR.

While promoter: reporters can be used in destructive, endpoint; and non-destructive context, e.g., real-time monitoring of gene expression with promoter: GFP constructs, there are other approaches to monitor gene expression. Quantitative PCR (qPCR) is a robust approach that is used widely to monitor gene expression. Here, PCR is used to amplify DNA that has been generated by reverse transcription from RNA, with careful calibration of dose-response of product formation.

6.1.12. RNA extraction and quantification

Ten mL of *Acinetobacter* sp. cells growing in either LB or Minimal Media spiked with 1 mM $InsP_6$ was harvested after reaching exponential/late exponential phase. The cells were pelleted at 3750 g for 5 min. To the cells, 1 mL of TRIzol/TRI reagent (Invitrogen) was added, and the tubes were vortexed thoroughly and incubated at room temperature for 5 min. To this, 0.2 mL of chloroform (Sigma) was added and the tubes vortexed briefly again until the contents were homogenously cloudy, then incubated at room temperature for 2 min. The tubes were centrifuged at 16060 g for 10 min at 4 °C. Of the three separated layers, the top, aqueous layer was taken and transferred to a fresh tube. 500 µL of isopropanol, was added, the tube inverted 5 times to mix and incubated for 10 min at room temperature to precipitate RNA. Following centrifugation at 16060 g for 10 min at 4 °C, the supernatant was removed, and the tube tapped to dry on tissue. Ethanol, 1 mL of 75% ethanol, made up with DEPC-treated or RNase-free water (Qiagen), was added and the tube vortexed briefly. RNA was pelleted at 16060 g for 2 min at 4 °C and then the majority of the ethanol was carefully removed, followed by a second centrifugation before the tube and its RNA was air dried for 5-10 minutes. Finally, 30 µL of DEPC/RNase-free water was added, and the RNA resuspended.

6.1.13. Dnase Treatment

Dnase treatment was performed using the Promega RQ1 RNase-free Dnase according to the protocol. Briefly, 1 µg of RNA (maximum 8 µL per reaction), 1 µL Dnase (Promega RQ1 RNase-free Dnase), 1 µL 10X Reaction buffer were made up to 10 µL using sterile RNase-free water. This was incubated at 37 °C for 30 min, before addition of 1 µL of RQ1 stop solution and incubation at 65 °C for 10 min. Dnase treated RNA was then used to produce complementary DNA, cDNA.

6.1.14. cDNA synthesis

cDNA synthesis was performed according to the First Strand cDNA Synthesis Standard Protocol from New England BioLabs using Protoscript II Reverse Transcriptase.

Briefly: To 1 µg of DNase-treated RNA, 2 µL Random Hexamers (NEB Random Primer Mix) and 1 µL dNTP mix (10 mM New England Biolabs) were added to a total volume of 10 µL. After heating at 65 °C for 5 min, 4 µL 5X Protoscript II Buffer (NEB), 2 µL 0.1 M DTT, 1 µL Protoscript II Reverse Transcriptase (NEB) and 3 µL RNase-free H_2O were added. This was incubated at 25 °C for 5 min, heated at 42 °C for 60 min, with a final incubation at 65 °C for

20 min to inactivate enzyme. The cDNA was quantified in a NanoDrop™ (ThermoScientific) and diluted to a concentration of 250 ng/µL.

6.1.15. qPCR primer design

qPCR Primer sets were made using the Primer Express Software 3.0.1, Table 6.3, with the AC1-2 *MINPP* gene followed by Primer Blast against the *Acinetobacter* sp. AC1-2 genome to ensure no secondary products should be made.

Table 6.3: Primer sets used for qPCR

	Position	Sequence (5' -> 3')	Primer Length (bp)
Forward Primer-1	82	CAACCAACGACATCACCTACAAC	23
Reverse Primer-1	221	CCATGACGCGCCACTAACT	19
Length	140 bp		
Forward Primer-2	1079	TATGCCTATTGAAGCCGCAAA	21
Reverse Primer-2	1278	CAGCATGGGCAAAACGTAAA	20
Length	200 bp		

Primer efficiencies were calculated using a two-step qPCR protocol described below from a dilution series of cDNA, 1:1, 1:5, 1:10, 1:25, 1:100. The primer efficiency for primer set 1 and 2 was calculated to be 105.44 and 97.49% respectively. Primer Set 1 was used for the LB and LB + $InsP_6$ samples whereas primer set 2 was used for MM and MM + $InsP_6$. This was because Primer Set 1 caused multiple bands in the minimal media datasets.

6.1.16. Quantitative PCR

qPCRs were performed for all individual samples in biological and technical triplicates in 20 µL reaction volume, using SensiFast SYBR Hi-Rox Kit (Bioline), with a final primer concentration of 400 nM. These were run on a StepOne Plus Real-Time PCR System (Applied Biosystems). The housekeeping gene *recA* was used alongside these experiments.

qPCR was performed using a two-step protocol: initial denaturation 95 °C 3:00, denaturation 95 °C 0:03, annealing 60 °C 0:30 with 40 cycles. In the second step, a melt curve was obtained: 95 °C 0:15, annealing 60 °C 1:00, with step and hold in +0.3 °C increments to 95 °C 0:15.

6.1.17. Data analysis

All experiments were carried out in technical and biological triplicates. The collected data was subjected to statistical analysis, performed using Prism 8.0.1. The ΔCt values were first investigated for outliers using the ROUT method, with any outliers being removed from the dataset. The dataset was then analysed for normality and lognormality using the Anderson-Darling test, which indicated the normalised distribution of the datasets. Therefore, significance between the two datasets were analysed using either an unpaired, parametric T-test, or an unpaired nonparametric Mann-Whitney test.

The statistical significance of the fold-change values was determined using a one-way ANOVA with the means of each column compared with the mean of every other column.

6.2. Results and Discussion

6.2.1. Promoter-reporter constructs

Single colonies of *Acinetobacter* sp. were used as template DNA for the amplification of Promoter regions 1, 2 and 3, Figure 6.1, by colony PCR which was run on a 1% agarose gel, Figure 6.6a and b. These inserts, alongside the pS4 vector were digested with the appropriate restriction enzymes and the inserts ran on a 1% agarose gel and gel extracted to obtain the purified digested inserts. A ligation reaction was performed overnight, and the reactions transformed into *E. coli* 803 cells and streaked onto LB plates with the appropriate antibiotics. Colonies for Promoters 1 and 2 were examined by colony PCR, Figure 6.6c, whereas for Promoter 3, the vector was digested by the EcoRI and PstI restriction enzymes separating the vector and the insert, Figure 6.6d. The expected band sizes for Promoters 1, 2 and 3 were 1029, 636 and 634 bp respectively.

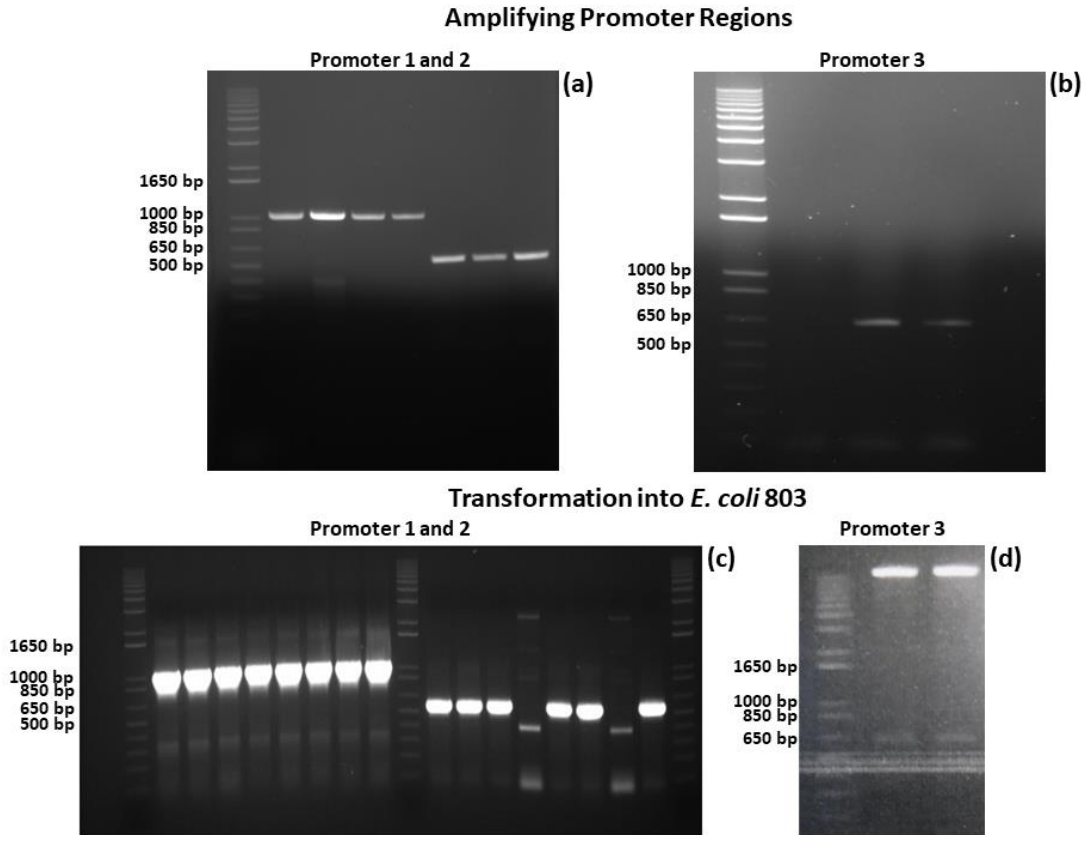


Figure 6.6 – Preparation of Promoters 1, 2 and 3 for transformation into *Acinetobacter* sp. The promoter regions were amplified using *Acinetobacter* sp. template DNA, (a) Lane 1-4 Promoter 1, Lane 5-8 Promoter 2, b) Lane 1-3 Promoter 3. (c) ligation of promoters 1 and 2 into the pS4 vector and transformation into *E. coli* 803, Lane 1-8 Promoter 1, Lane 10-17 Promoter 2. (c) Restriction digest of Promoter 3, Lane 1-2.

Triparental mating was performed by streaking together *Acinetobacter* sp. (host), *E. coli* 803-pS4 (donor) and *E. coli* PRK2013 (helper) cells which were grown together overnight, after which, the cell mass was plated onto antibiotic selection. Single colonies were examined for the successful transformation of the pS4 vector, Figure 6.7. As such, each *E. coli* pS4:Promoter strain underwent sequencing to confirm the correct sequence of each promoter region.

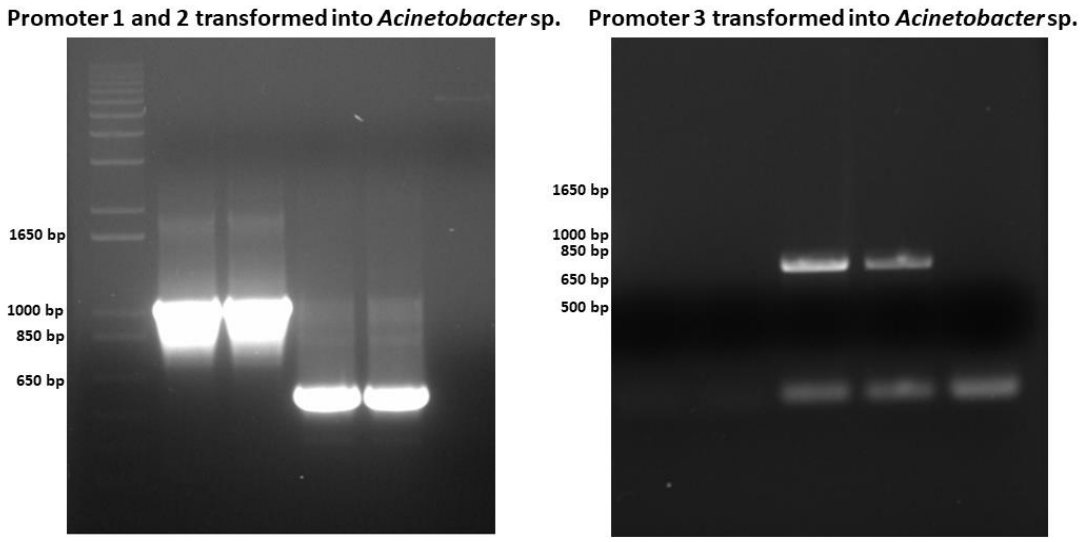


Figure 6.7 – Confirmation of the presence of Promoter 1, 2 and 3 in the pS4 vector into *Acinetobacter* sp., through conjugation. Note* the Promoter 3 gel does not have a ladder as the gel was run with samples from another experiment.

Bands with appropriate size for Promoters 1, 2 and 3, 1030, 636 and 634 bp, respectively, were observed, Figure 6.7.

6.2.2. Induction of phytase expression in the presence of phytate and inorganic phosphate

As discussed in the methods section, only Promoters 1 and 3 were used in the following experiment, as Promoter 1 is located at the beginning of the cluster and gives a larger β-galactosidase response it was preferred over Promoter 2. Additionally, Promoter 3, which is located before the *MINPP* gene, gave only basal β-galactosidase activity and was used as a negative control. In this experiment, *Acinetobacter* sp. cells were grown to exponential/late exponential in the presence of different concentrations of (Impure) $InsP_6$ and P_i. The results are displayed in Figure 6.8.

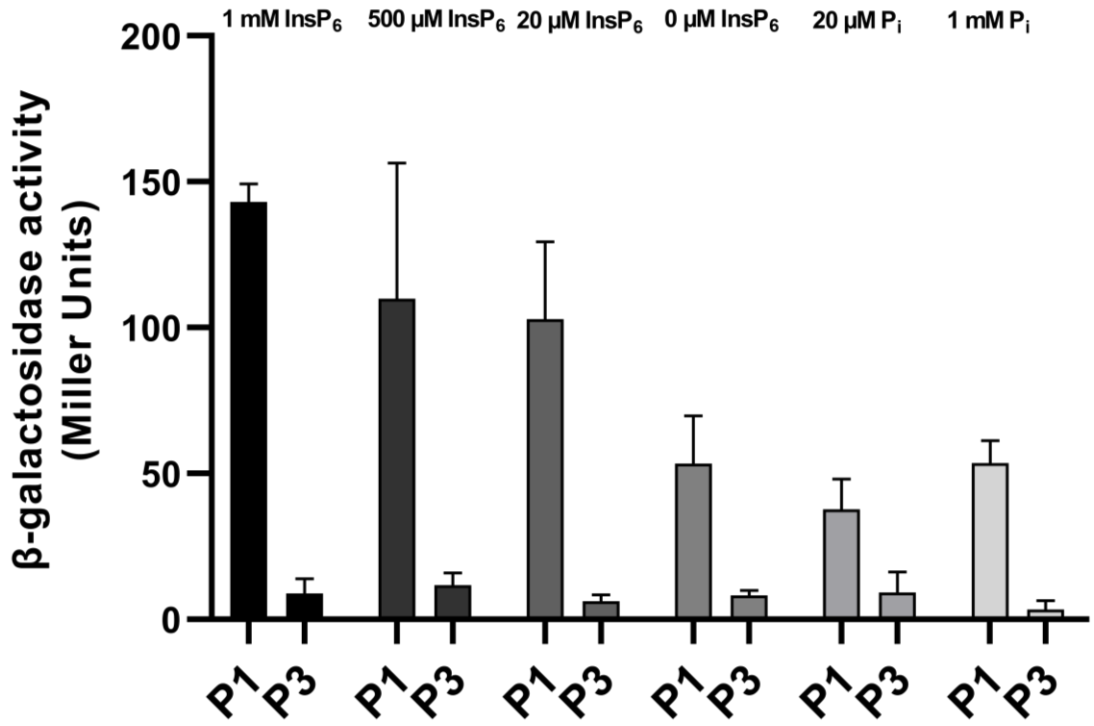


Figure 6.8 – The β-galactosidase activity of Promoter regions 1 and 3 in the presence of different concentrations of impure $InsP_6$ (Sigma P8810) (1 mM, 500 µM, 20 µM) and Inorganic phosphate, P_i (20 µM and 1 mM).

The results from the β-galactosidase assay indicate weak induction of the promoter in the presence of $InsP_6$ that is not scaled based on the concentration of $InsP_6$ used. Additionally, there is a much weaker induction by inorganic phosphate. In the literature bacterial phytases have been shown to be repressed in the presence of inorganic phosphate (Konietzny and Greiner, 2004). This is not the only feature of phytase repression however, and in fact, phytase expression is controlled by a complex regulatory system that is not universal. For example, expression of phytase in *Klebsiella terrigena* was halted in the exponential phase of growth, only resuming in the stationary phase in the presence of phytate where carbon starvation was a limiting factor (Konietzny and Greiner, 2004). In *E. coli*, phytase production was shown to be stimulated by a limitation in inorganic phosphate, interestingly phytate had no influence of the synthesis of the *E. coli* phytase (Greiner *et al*., 1997; Jain *et al*., 2016). Therefore, in the case of *Acinetobacter* sp. AC1-2 we see stimulation of promoter activity in the presence of increasing concentrations of phytate, which is similar to that of *Klebsiella terrigena* and unlike *E. coli*. Therefore, a further experiment should be conducted to examine the stimulation of the promoter region in regards to changing carbon concentrations.

Furthermore, in order to examine whether these results are corroborated using another method, RT-qPCR was performed.

6.2.3. Regulation of AC1-2 *MINPP* expression, assessed by reverse transcriptase quantitative PCR (RT-qPCR)

The experiment and results from Section 6.2.2 describe for liquid culture of *Acinetobacter* sp. AC1-2, alone, the increased expression of the gene cluster containing AC1-2 *MINPP* in the presence of $InsP_6$ with lesser induction when P_i is the sole source of phosphate.

In this experiment, AC1-2 *MINPP* expression in rich and minimal media was also examined directly by RT-qPCR. The RNA was extracted from *Acinetobacter* sp. cells after growing in either LB, or phytase isolation minimal media with or without 1 mM (impure, Sigma P8810) $InsP_6$ until exponential/late exponential phase. These were visualised on a non-denaturing agarose gel and examined for purity of the RNA and signs of degradation, Figure 6.15.

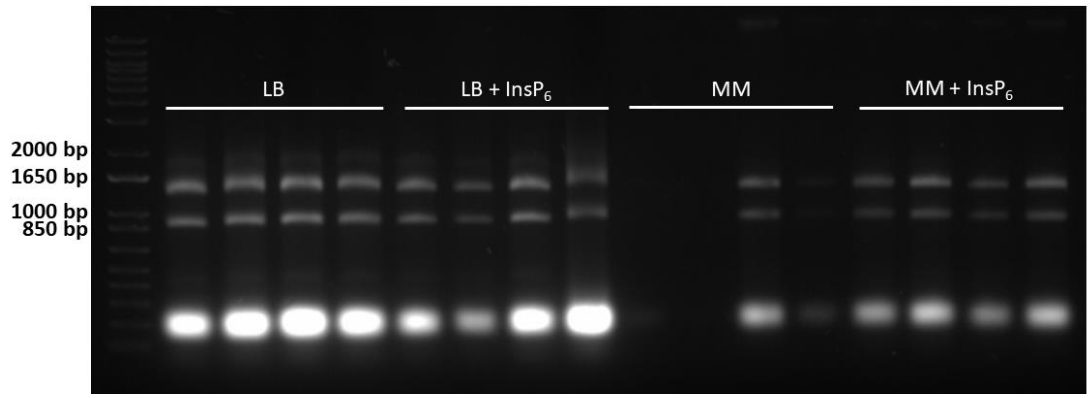


Figure 6.15 - A non-denaturing agarose gel analysis of RNA from extracted samples, Samples 1-4, *Acinetobacter* sp. grown in LB; Samples 5-8, *Acinetobacter* sp. grown in LB + 1 mM $InsP_6$; Samples 9-11, *Acinetobacter* sp. grown in MM and Samples 12-14 *Acinetobacter* sp. grown in MM + 1 mM $InsP_6$.

When RNA samples are run on an agarose gel, there should be two sharp rRNA bands (23s and 16S for bacterial samples), with the top band (23S) approximately twice as intense as the lower band (16S). For partially or completely degraded RNA there will be a smeared appearance on the gel which will lack the sharp rRNA bands (As described by the ThermoFisher Scientific RNA protocols). Examination of the gel determined that RNA of good quality was extracted, and confirmed using a NanoDrop™, where RNA concentration ranged from 58.4-937 ng/µL and Absorbance 260/280 from 1.78-2.07, Supplementary Table S1. Using Reverse Transcriptase, 1 µg of RNA was converted into cDNA.

6.2.3.1. Calculating qPCR primer efficiencies.

Two qPCR primers sets were designed to examine the expression of the *MINPP* gene. First, it was necessary to determine the primer efficiency of the two primer sets. This is because the data analysis method assumes that there is near 100% amplification efficacy of the target genes (Mallona *et al*., 2011). Therefore, a dilution series 1:1, 1:5, 1:10, 1:25 and 1:100 of known cDNA concentrations was performed, Figure 6.16 – Primer Set 1. Here, RN is the normalised reporter value, the fluorescent signal of the SYBR green normalised to the signal of the passive reference dye and ΔRN is the Rn value of an experimental reaction minus the Rn value of the baseline signal generated by the instrument. When the ΔRN value reaches 10, it means that the amplified DNA fluorescence is detected as statistically significant above the background, this is called the threshold cycle or Ct value. The Figures for calculating efficiency for Primer Set 2 can be found in the supplementary information.

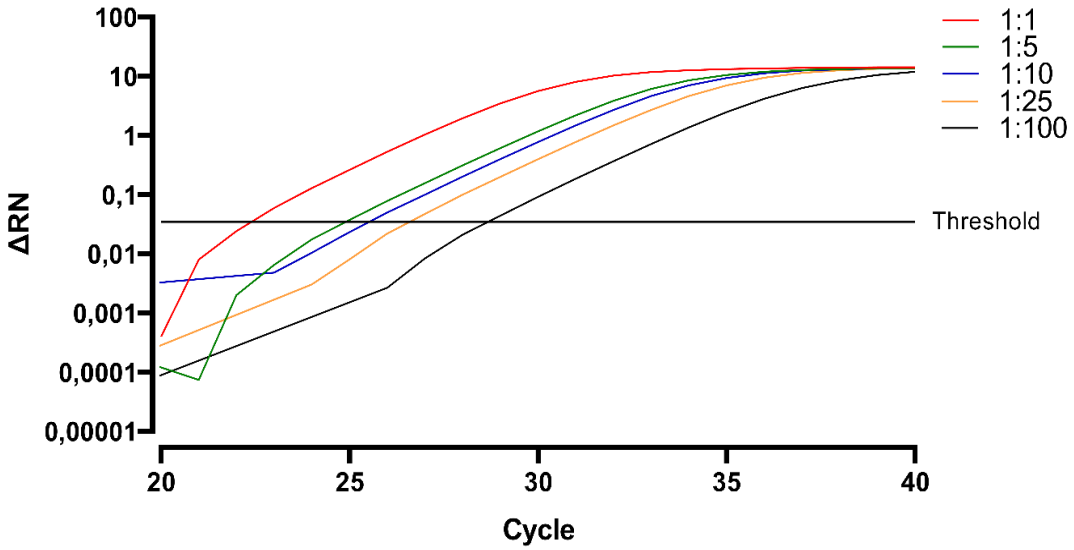


Figure 6.16 - The average ΔRN amplification data for Primer Set 1 for each set of the serial dilutions in triplicate was plotted against the qPCR cycle number.

As a method of quality control, after the end of each qPCR run, a melt-curve analysis was performed (Taylor *et al*., 2010). These are often used to assess whether the qPCR reaction has produced a single, specific product. Briefly, the intercalating dyes used in qPCR fluoresce only when they are bound to double-stranded DNA (dsDNA). Starting at 60 °C, the thermocycler measures the fluorescence and then incrementally (+0.3 °C) raises the temperature and measures fluorescence. As the temperature increases, dsDNA denatures into single-stranded DNA (ssDNA) and the dye dissociates causing a decrease in fluorescence.

The change in slope of this curve is then plotted as a function of temperature to obtain the melt curve, Figure 6.17.

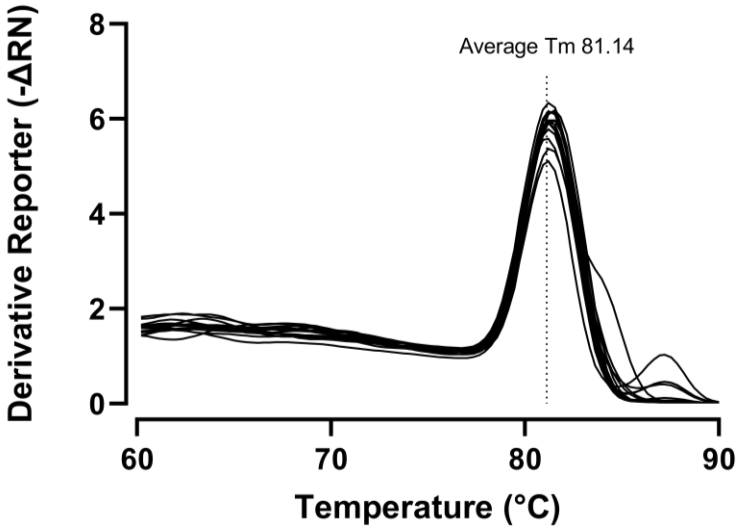


Figure 6.17 – Melt Curve Analysis of Primer Set 1 to asses the quality of the qPCR. Each dilution, 1:1, 1:5, 1:10, 1:25 and 1:100, in triplicate were analysed.

Melt curve analysis is particularly useful for identification of improper reactions, or artifacts. Thus, in Figure 6.17 there was a broadening of one side of the peak. All of the qPCR reactions were subseqeuntly analysed on an agarose gel to examine whether the broadening of the melt-curve was due to secondary products, Figure 6.18.

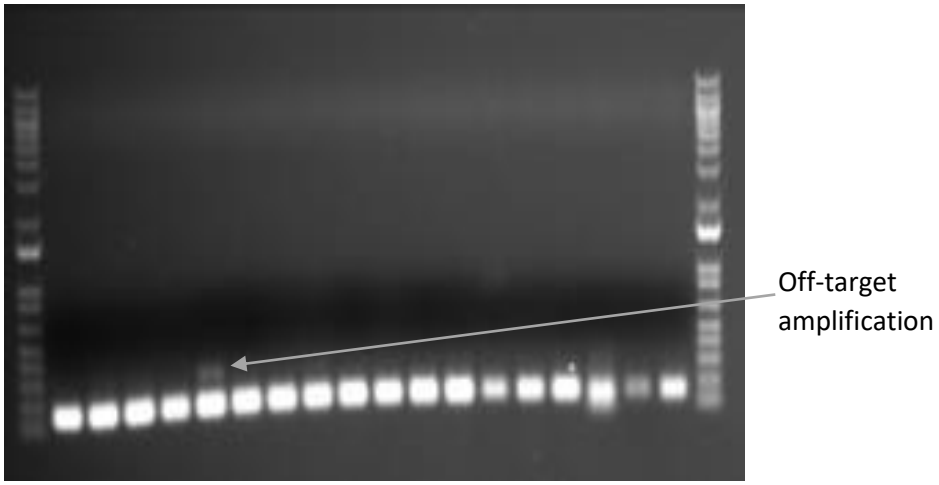


Figure 6.18 – The agarose gel-electrophoresis of the qPCR reactions for the Primer Set 1 dilutions.

As one of the samples contained an impurity, this sample was discarded for the primer efficiency calculations. The threshold, Ct values obtained from melt curve analysis were used to calculate primer efficiency against the $\log_{10}$ of the cDNA concentration, Table S2, linear regression yielded the slope of the line and R^2 value, Figure 6.19.

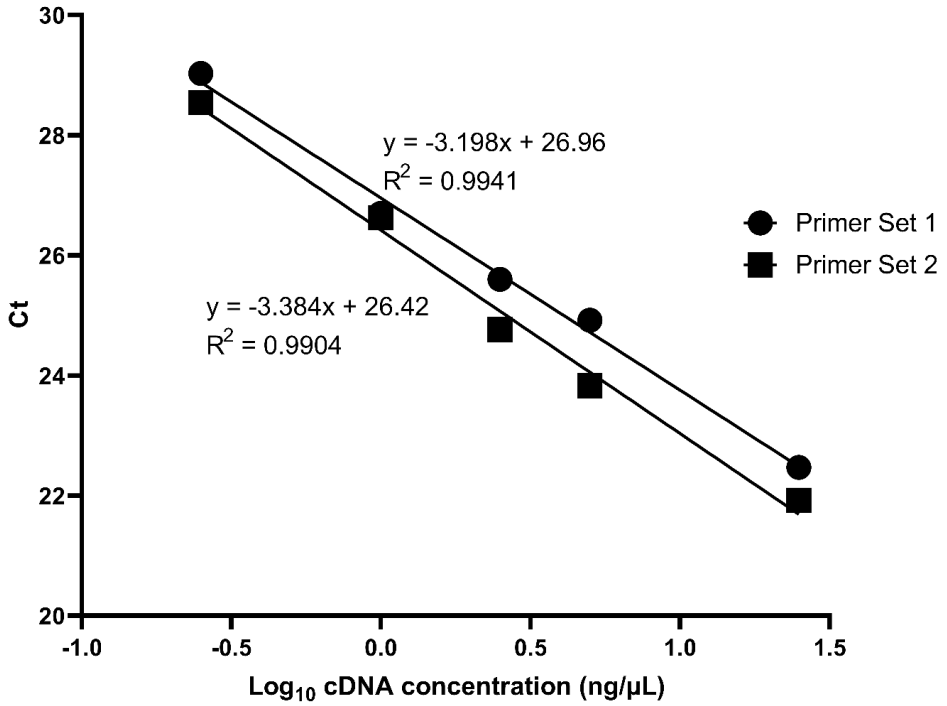


Figure 6.19 – Linear regression of the qPCR primer sets to calculate efficiency.

Efficiency was calculated using the following equation:

$$Efficiency\ (\%) = \ 10^{\left(\frac{-1}{Slope}\right)} - 1 \times 100$$

Therefore, the Primer Efficiencies were calculated to be 105 and 97% for Primer Sets 1 and 2. Primer Set 1 was used for the LB and LB + $InsP_6$ Samples and Primer Set 2 used for the MM, MM + $InsP_6$ and MM + P_i Samples.

6.2.3.2. qPCR analysis of *Acinetobacter* sp. *MINPP* expression in LB and LB + $InsP_6$ liquid cultures

qPCR was performed using biological and technical replicates. Firstly, the LB sample with and without $InsP_6$ was performed alongside the housekeeping gene *recA*. The housekeeping genes are traditionally used to control for error between samples and different treatments/environments. Genes such as *recA* are commonly used as it is expected that their expression remains constant between different experimental conditions (Dheda *et al*., 2004). Similarly, to Section 6.2.3.1, the qPCR reaction was visualised in real time on the StepOne

Plus PCR System, Figure 6.16, and the melt curve, Figure 6.17, and agarose gels, Figure 6.18, were used as a quality control to ensure the reaction did not have any secondary products. As such, the LB sample qPCR information is shown below, Figure 6.20-6.22 and the LB + $InsP_6$ data in the supplementary information, Figure S4-S6.

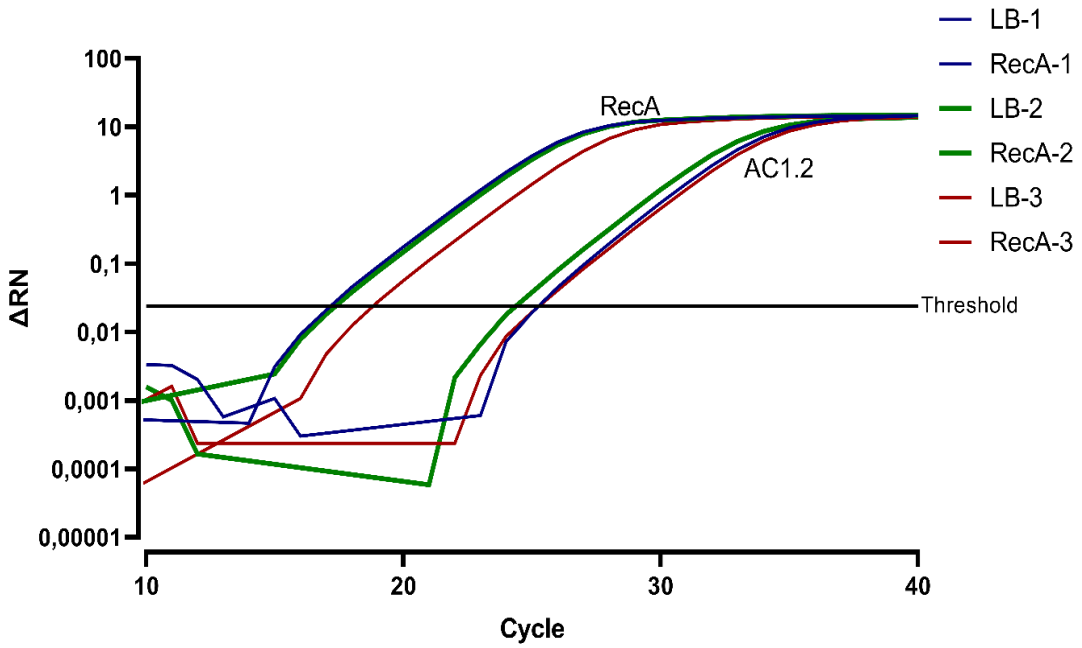


Figure 6.20 – qPCR plot displaying average ΔRN amplification versus cycle number for each LB sample comparing the expression of the *MINPP* gene relative to housekeeping gene *RecA* when grown to exponential/late exponential phase.

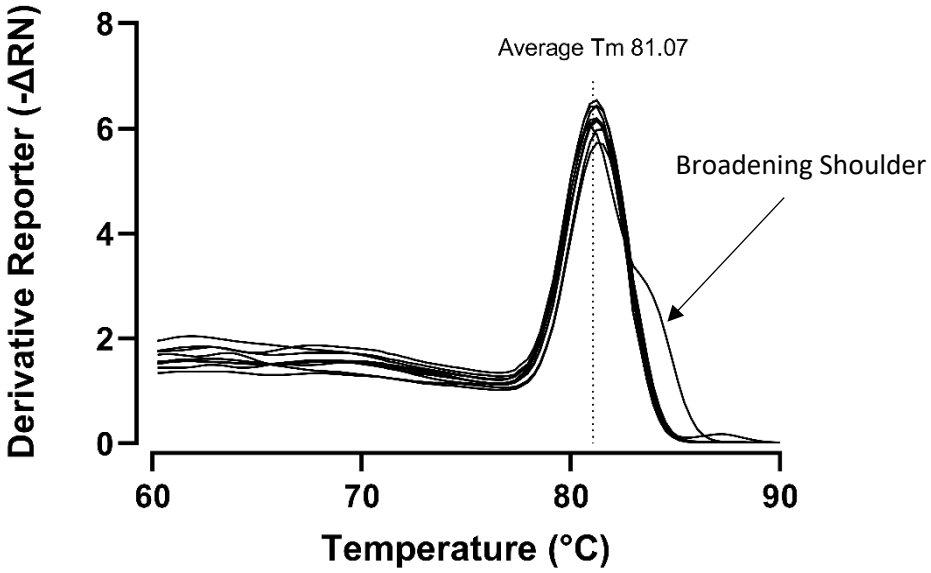


Figure 6.21 - Melt Curve Analysis of each LB Sample dataset to assess the quality of the qPCR. A broadening shoulder indicates potential impurities.

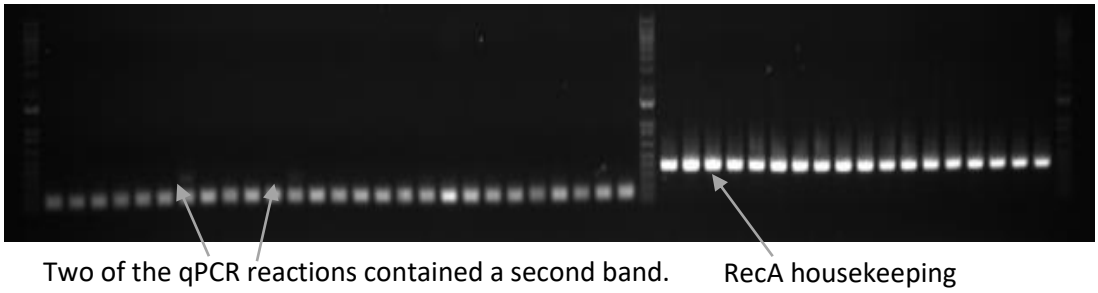


Figure 6.22 - The agarose gel-electrophoresis of the qPCR reactions for the LB Samples. Lanes 1-9, LB-1; Lanes 10-18, LB-2; Lanes 19-27, LB-3; Lanes 29-37, RecA-1 and Lanes 38-46, RecA-2.

Lanes containing a second band were omitted from the dataset.

6.2.3.3. Calculation of Fold Change between the LB and LB + $InsP_6$ treatments

As discussed in Section 6.2.3.1, the Ct values are the number of cycles required for the amplified DNA to reach the threshold limit. However, due to the technical variability between experiments, it is necessary for the Ct values to be normalised. Consequently, each Ct value was normalised to the average Ct value of its corresponding housekeeping gene, ΔCt.

$$\Delta Ct = Ct(MINPP) - Average\ Ct(RecA)$$

Once normalised, the ΔCt of the control sample (LB) was subtracted from the experimental sample (LB + $InsP_6$), giving ΔΔCt which is the differences between the between the control and experimental samples.

$$\Delta\Delta Ct = \Delta Ct(Experimental) - Average\ \Delta Ct(Control)$$

Finally, Fold change was calculated using the following equation:

$$Fold\ Change = 2^{-\Delta\Delta Ct}$$

This is necessary because the Ct values are in the logarithmic scale, as DNA is approximately doubled in every qPCR cycle (Yuan *et al*., 2006). The average fold change of *MINPP* expression the LB + $InsP_6$ treatment in comparison to the LB Sample was 1.45±0.105, 1.00±0.211 and 0.8725±0.220 for Samples 1, 2, 3 respectively with an overall average of 1.088±0.312.

The Ct value is inversely proportional to the quantity of DNA, therefore high ΔCt values represents low expression, while low ΔCt values represent high expression. A fold change of 1 means that there is no change in expression between the control and experimental

conditions (equivalent to 100%). A fold-change value above 1 shows up-regulation of the gene of interest relative to the control, for example, a fold change of 2 is equivalent to 200% gene expression relative to the control. A fold-change below 1 shows gene down-regulation relative to the control, for example, a fold change of 0.5 is equivalent to 50% gene expression relative to the control.

Therefore, to conclude, in the rich LB culture medium, there is no change in AC1-2 *MINPP* expression when the medium is spiked with $InsP_6$.

6.2.3.4. qPCR analysis of *Acinetobacter* sp. *MINPP* expression in Minimal media (MM), MM + $InsP_6$ and MM + P_i treatments.

Similar analysis was performed for the MM and MM + $InsP_6$ Samples using Primer Set 2. Once again, all the Samples were plotted on a logarithmic graph versus the cycle number, and quality control was performed by analysing the melt curves and agarose gel, Figure 6.23-6.25. The MM + $InsP_6$ Figures can be found in the supplementary information, Figures S7-S9.

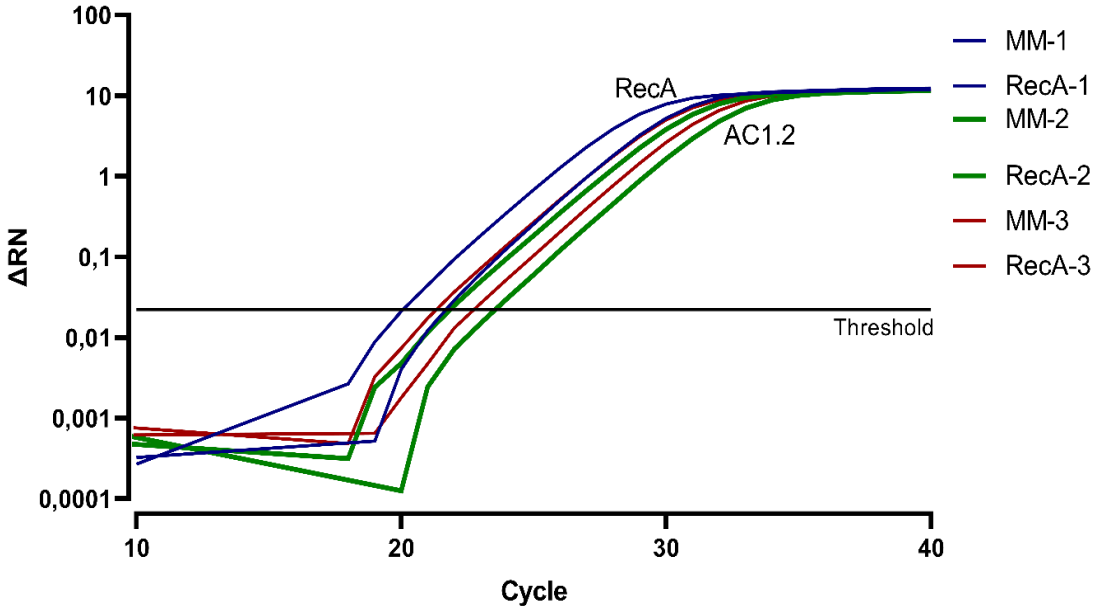


Figure 6.23 - qPCR plot displaying average ΔRN amplification versus cycle number for each MM treatment, comparing the expression of the *MINPP* gene to the housekeeping gene *RecA*.

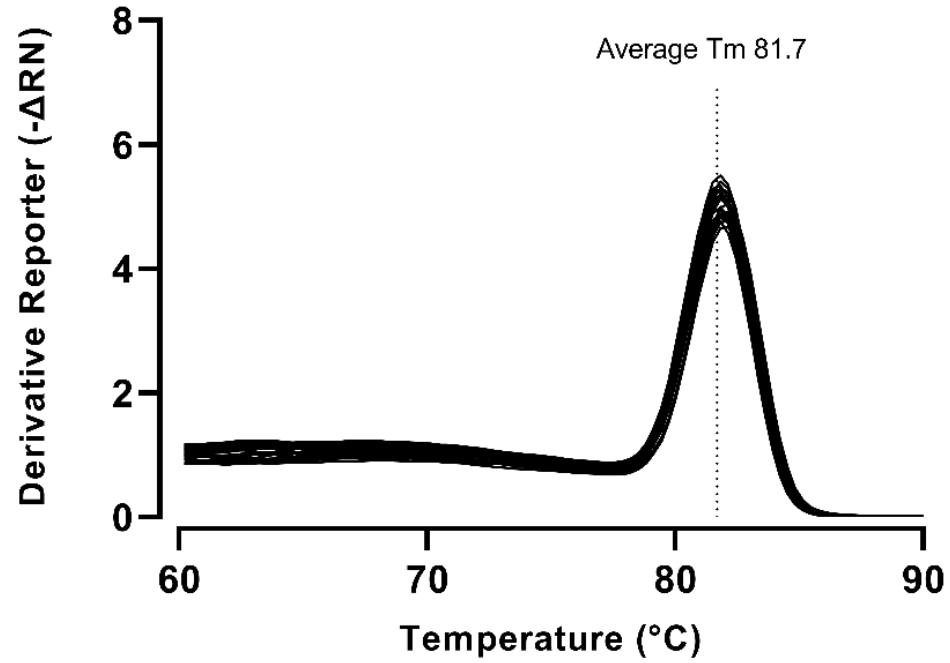


Figure 6.24 - Melt Curve Analysis of each MM Sample dataset to assess the quality of the qPCR.

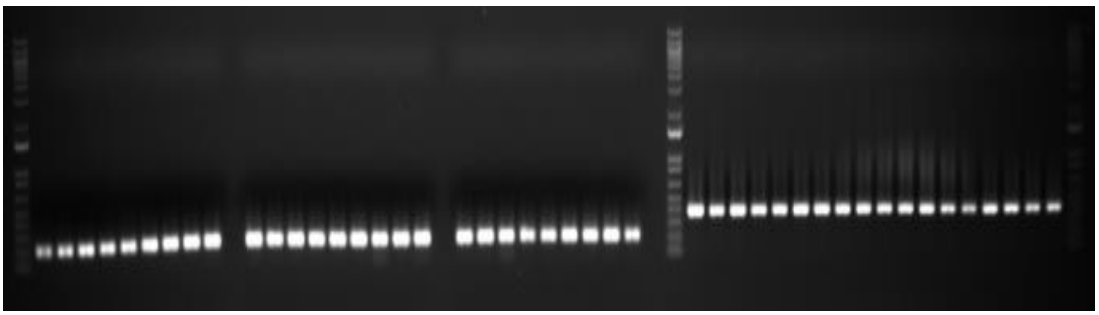


Figure 6.25 - The agarose gel-electrophoresis of the qPCR reactions for the MM treatment. Lanes 1-9, MM-1; Lanes 11-19, MM-2; Lanes 21-29, MM-3; Lanes 32-40, RecA-1 and Lanes 42-50, RecA-2.

6.2.3.5. Calculation of Fold Change between the MM, MM + $InsP_6$ and MM + P_i treatments

Fold change was calculated as described per Section 6.2.3.3 for the MM + $InsP_6$ condition using the Ct values, Supplementary Table S4. The average fold change of the MM + $InsP_6$ Sample in comparison to the MM Sample was 0.114±0.029, 0.167±0.028 and 0.123±0.0118 for Samples 1,2 and 3 respectively with an overall average of 0.135±0.029.

These results indicate that in a minimal media where phytate is the sole phosphate source, AC1-2 *MINPP* expression is repressed by 7.5-fold in comparison with the minimal media with no phosphate additives. This was initially a surprising result, however, the impurities of the $InsP_6$ sample had not been considered. The concentration of inorganic phosphate therein

was measured, and equivalent concentration supplemented into minimal media. These figures can be found in the Supplementary Information, Figures S10-S11.

The average fold change of the MM + P_i treatment in comparison to the MM treatment was 0.309±0.028, 0.589±0.126 and 0.273±0.0344 for treatments 1, 2 and 3 respectively with an overall average of 0.390±0.161. Therefore, there is a reduction in gene expression by 2.6-fold in the presence of inorganic phosphate in liquid culture, Figure 6.25.

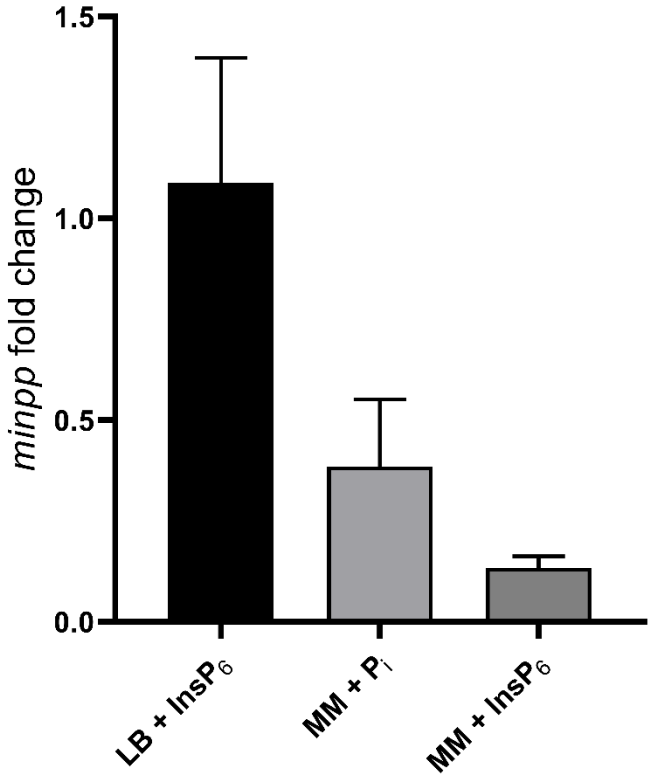


Figure 6.26 –Fold-change of AC1-2 *MINPP* in LB + $InsP_6$, MM + P_i and MM + $InsP_6$ treatments.

Interestingly, these results appear to differ from the β-galactosidase assays in Section 6.2.2, which show an induction of the gene cluster in the presence of the same $InsP_6$ sample, with the same contaminants. However, this does not mean either result are wrong, β-galactosidase assays only examines activation of the promoter region and is not a direct *MINPP* assay whereas qPCR is directly examining AC1-2 *MINPP* expression from the mRNA, thus the β-galactosidase assays could be considered less accurate. To explain differences between the two, there may be other overlapping regulatory elements that have not been encoded within the promoter region that could create transcriptional interference (Smanski *et al*., 2014).

Regardless, this study has shown that AC1-2 *MINPP* expression was unaltered by $InsP_6$, and associated phosphate, in the rich LB medium. In minimal media however, the inclusion of inorganic phosphate or $InsP_6$ (containing inorganic phosphate impurities) reduced gene expression ($p<0.05$) by 2.6-fold and 7.5-fold respectively. A general feature commonly observed in microbial phytase producers is the regulatory inhibition of phytase production by inorganic phosphate levels (Konietzny and Greiner, 2004) which may influence the RT-qPCR results displayed in this section. Nevertheless, this shows a dual aspect to regulation of AC1-2 MINPP phytase activity, by inorganic phosphate and phytate. Returning to the *Klebsiella terrigena* phytase discussed in Section 6.2.2, phytase expression was halted in the exponential phase of growth, with resumption of expression upon entering the stationary phase. As this sample was taken in the exponential/late exponential phase of growth, this may also be impacting *MINPP* expression.

6.2.4. Generation of an *Acinetobacter* ΔMINPP knockout

The *Acinetobacter* MINPP gene was knocked out using the double crossover technique, Figure 6.3. Two flanking regions covering a small portion of either side of the *MINPP* gene, Figure 6.4, were cloned into the suicide vector pK18mobsacB, Figure 6.9a and Figure 6.9b. Using restriction enzyme digestions, these flanking regions were combined in the vector and transformed into *E. coli* 803, this was initially confirmed by colony PCR Figure 6.9c and Figure 6.9d and restriction enzyme digests, Figure 6.9e to show the vector and two flanking regions. The expected size of Flanking regions 1 and 2 were 719 and 542 bp respectively, with a combined size of 1261 bp.

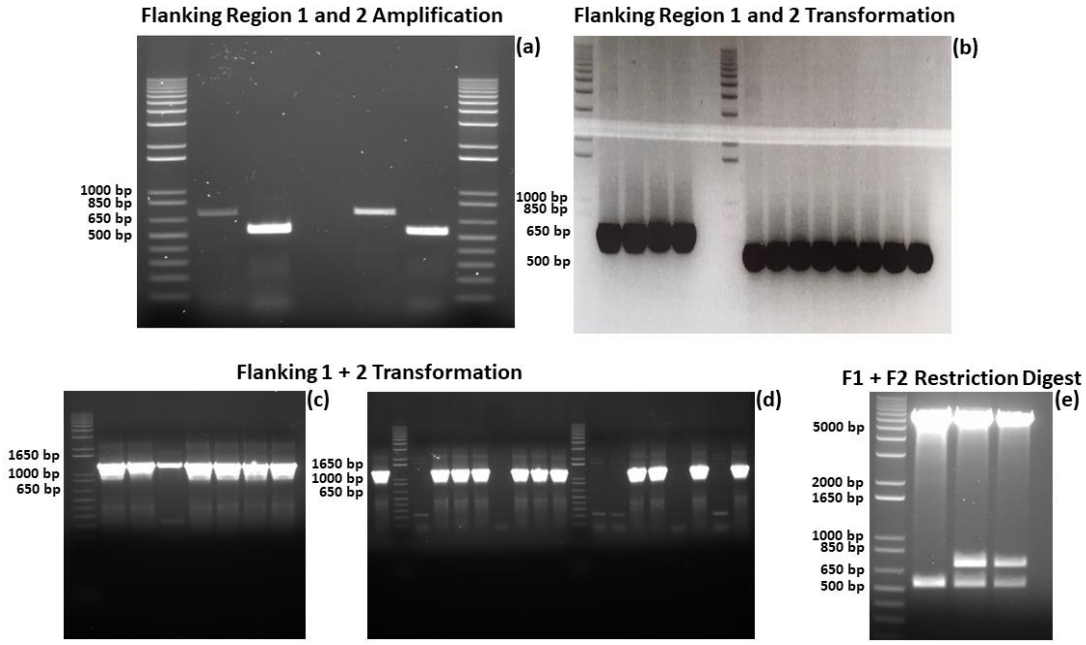


Figure 6.9 – Preparation of the *MINPP* gene knockout. Flanking regions 1 and 2 were amplified by colony PCR (a), Lane 1, 4 Flanking region 1, Lane 2,5 Flanking region 2, (b) Flanking regions 1 and 2 ligated into the pK18mobsacB vector, Lane 1-4 Flanking region 1, Lane 6-13 Flanking region 2. (c + d) DNA was isolated from both vectors and digested with restriction enzymes to yield the flanking regions which were ligated into the vector containing the other flanking region. Flanking region 1 (c) Lane 1-7, Flanking region 2 (d) Lane 1-17, with further confirmation that the correct ligation and transformation had occurred by restriction digests (e) Lane 1-3.

With both flanking regions inserted in the pK18mobsacB vector, triparental mating was performed to transform the vector from *E. coli* into *Acinetobacter* sp. Colony PCR was performed initially to confirm integration of the truncated gene into the *Acinetobacter* sp. gene: a 2500 bp band is indicative of the WT enzyme (the size of the gene + flanking regions), whilst a band at approximately 1250 bp indicates integration into the host genome, Figure 6.10a. There is still however, the problem of the pK18mobsacB vector which would still reside in the *Acinetobacter* sp. cell. Therefore, isolates which appeared to have worked were first propagated without kanamycin and then plated onto LB media with 10% sucrose. Here, the presence of the *sacB* cassette in the vector leads to cell death, selectively pressuring the cells to remove the vector. Once again, the isolates were examined specifically for the *MINPP* gene: yielding a 300 bp product for a successful knockout, and 1572 bp for an unsuccessful one, Figure 6.10b.

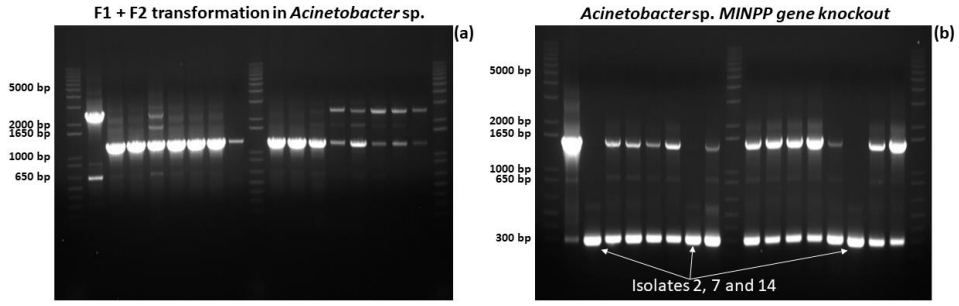


Figure 6.10 – The transformation of the pK18mobsacB vector into *Acinetobacter* sp. and replacement of the *MINPP* gene with a truncated version (a) Lane 1 – WT, Lane 2-16 - Isolates, followed by removal of the vector by 10% sucrose and confirmation that the gene has been removed (b) Lane 1 – WT, Lane 2-16 – Isolates.

Isolates 2, 7 and 14 from the 10% sucrose plates showed a truncated *MINPP* gene when examined by colony PCR, Figure 6.10, and these were grown to test for phytase degradation by HPLC, Figure 6.11.

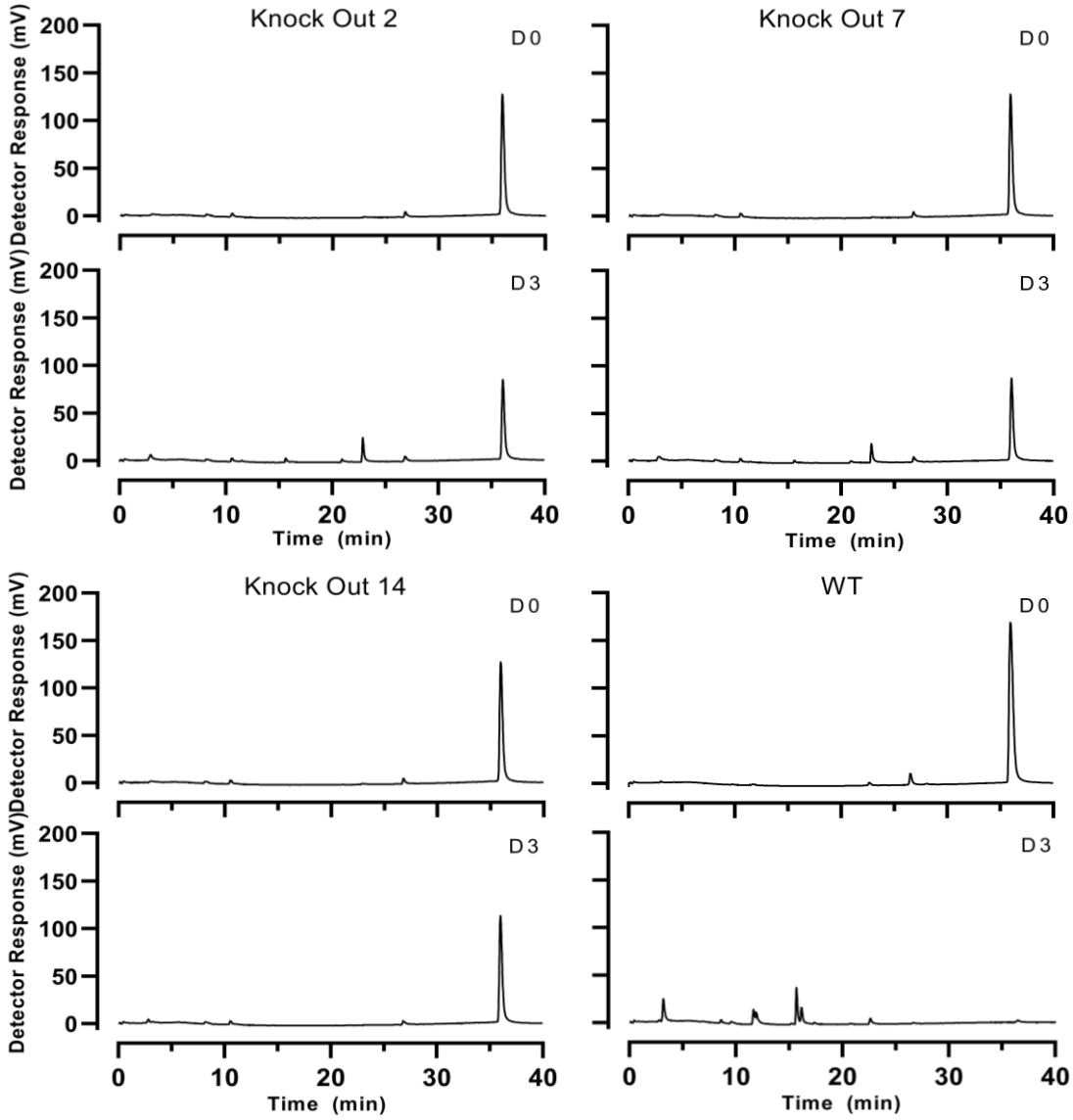


Figure 6.11 – HPLC assay of phytase activity of *Acinetobacter* sp. *ΔMINPP* isolates in comparison to the WT bacteria.

Both isolates 2 and 7 showed low levels of phytase activity, evidenced by the generation of peaks of $InsP_5$(4/6-OH), which is a known AC1-2 MINPP product. Despite this however, phytase activity has been significantly slowed in these isolates in comparison to the WT bacteria. Isolate 14 showed no signs of phytate degradation (the small peak of $InsP_5$ (1/3-OH) is a minor contaminant), confirming phytase activity has successfully been eliminated from the bacteria.

6.2.5. Competition experiment to compare performance of Rif-resistant *Acinetobacter* sp. and *Acinetobacter* sp. Δ*MINPP* in soil.

A competition experiment between the Rif-resistant *Acinetobacter* sp. and *Acinetobacter* sp. Δ*MINPP* was performed in the soil environment to investigate whether presence of the phytase gene confers advantage. Wild-type *Acinetobacter* sp. AC1-2 was not used as rifampicin resistance was induced into the knockout as a selective marker, this mutation occurs in the *rpoB* gene which encodes the β-subunit of RNA polymerase (Goldstein, 2014), which has an effect on the growth rate of the bacteria (Cai *et al*., 2017). Therefore, the WT bacteria would have a growth advantage on the Δ*MINPP* knockout and therefore the Rif-resistance bacteria with the WT *MINPP* gene was used.

Three different conditions were investigated: a thrice-autoclaved soil with no additives, the same spiked with 10 mM $InsP_6$ and, again, spiked with 10 mM $InsP_6$ and P_i. To these soils, equal amounts of Rif-resistant *Acinetobacter* sp. and *Acinetobacter* sp. Δ*MINPP* were added and the samples incubated at room temperature for one week. Each sample was then serially diluted onto LB plates, 10^{-1}-10^{-7} and incubated at 37 °C. Single colonies were taken at random from the plate and examined by colony PCR for the presence of either the WT gene or the truncated one. Ninety-seven colonies from Sample 1 were examined, Figure 6.12, 93 colonies from Sample 2 were examined, Figure 6.13 and 101 colonies were examined from Sample 3, Figure 6.14.

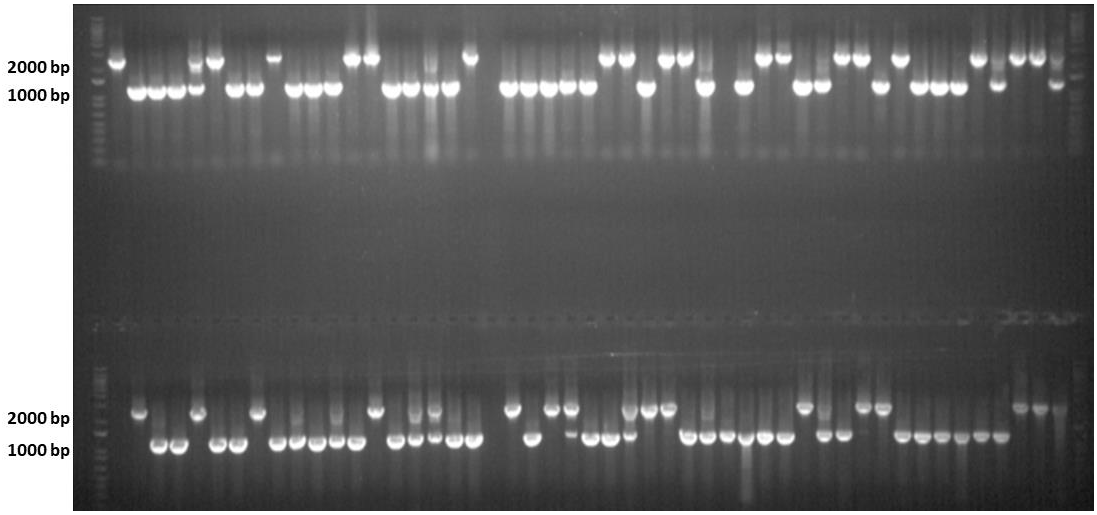


Figure 6.12 – PCR verification of growth of *Acinetobacter* sp. and *Acinetobacter* sp. Δ*MINPP* in autoclaved, nil phosphate soil. The WT gene has an expected band at around 2500 bp, while the truncated gene (KO) has an expected band at 1350 bp.

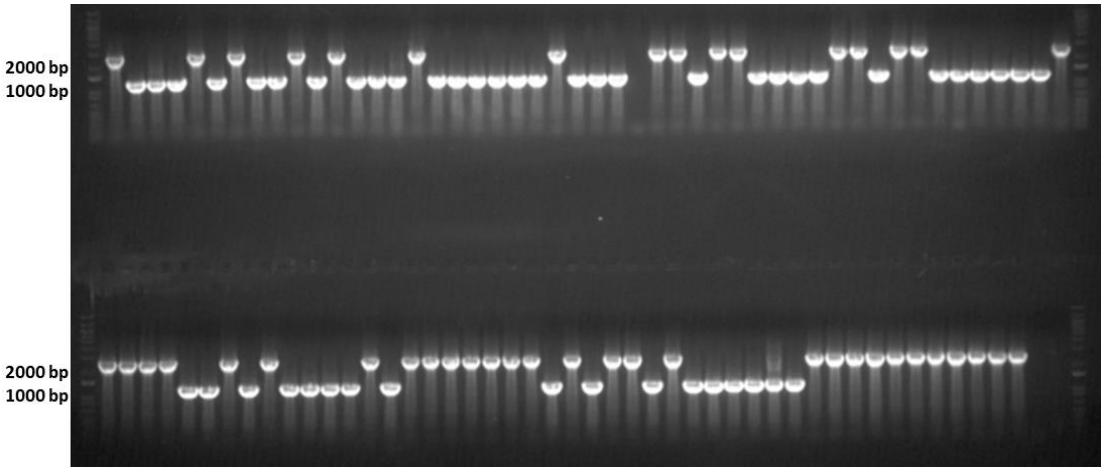


Figure 6.13 – PCR verification of growth of *Acinetobacter* sp. and *Acinetobacter* sp. Δ*MINPP* in soil spiked with $InsP_6$.

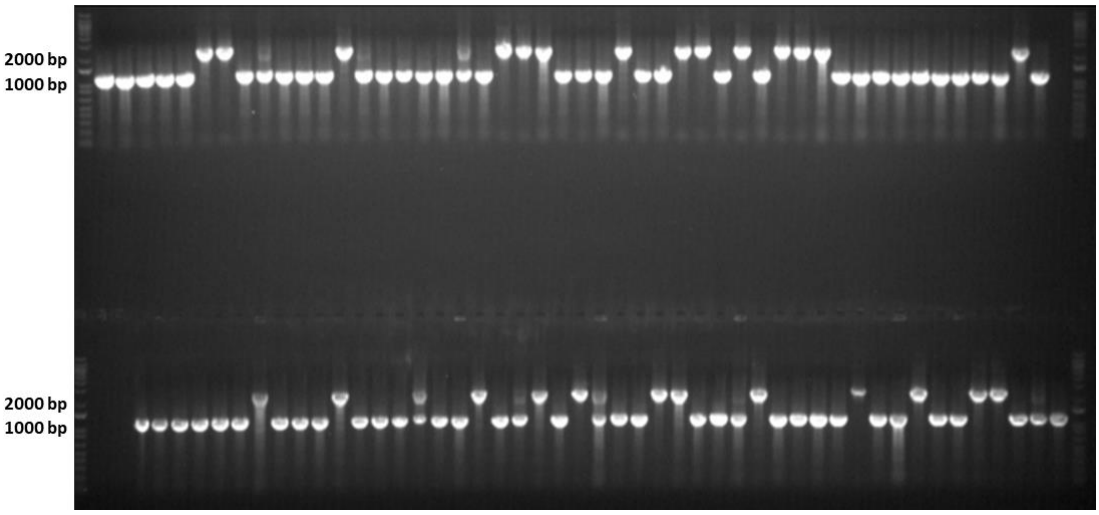


Figure 6.14 – PCR verification of growth of *Acinetobacter* sp. and *Acinetobacter* sp. Δ*MINPP* in soil spiked with $InsP_6$ and Pi.

The ratio of the occurrence of WT and truncated *MINPP* genes for Samples 1, 2 and 3 were 35:62, 45:48, 27:74 respectively. In Sample 1, in the presence of autoclaved soil with nil phosphate, there appeared to be a higher survivability of the *Acinetobacter* sp. Δ*MINPP* to grow over the Rif-resistant *Acinetobacter* sp., whereas in Sample 2, in the presence of added $InsP_6$, there were near equal amounts of the WT and truncated *MINPP* gene, the WT bacteria appearing to better respond to the presence of $InsP_6$. There is a stark difference in Sample 3, in soil spiked with both $InsP_6$ and inorganic phosphate, with the *Acinetobacter* sp. Δ*MINPP* growing significantly better than the Rif-resistant *Acinetobacter* sp.

While this initial study was only completed once for each condition and therefore results may be subject to change, a number of questions arise. For sample 1, the indication is that in the presence of very scarce environments, with little to no $InsP_6$ concentrations (In this case due to autoclaving the soil) it was not favourable for the Rif-resistance *Acinetobacter* sp. to grow over the Δ*MINPP* bacteria. It may be that the activation of the *MINPP* gene cluster in response to low inorganic phosphate concentrations is not energetically favourable for the bacteria and results in a lower survivability. In the presence of $InsP_6$ however, we still do not see the dominance of the Rif-resistant *Acinetobacter* sp. with the *MINPP* gene. It may be due to the nature of phytase expression, many of which are secreted into the soil environment (Richardson *et al*., 2001; Singh and Satyanarayana, 2011) that leads to this phenomena. The *Acinetobacter* sp. secretes the MINPP phytase into its surroundings which contains a mixture of the Δ*MINPP* and wild-type variants, with the phytases releasing inorganic phosphate from the phytate into the soil. The inorganic phosphate in the soil can therefore be taken up by any bacteria in its surroundings regardless of whether they secreted the phytase to release it or not, an example of 'social cheaters' which is excellently described in the report by Hibbing *et al* (Hibbing *et al*., 2010). This concept can be described as part of the complex interactions that can occur both within and between species, originally, the bacteria cooperate to obtain a group specific advantage, however 'cheaters', in this instance caused by the gene knockout, unfairly procure an excessive share of the group-generated resources, while only making small contributions themselves (Hibbing *et al*., 2010; Velicer, 2003). For example, when *P. aeruginosa* is grown under conditions requiring quorum sensing-regulated extracellular proteases, social cheaters with mutations in *lasR*, the quorum sensing regulator, accumulate within 100 generations, the cheaters benefiting from the protease activity of the enzymes secreted by their neighbours without expending energy themselves (Hibbing *et al*., 2010). This may be occurring here in regards to the *MINPP* gene knockout.

Finally, Sample 3, which contrasts to Sample 1, with high $InsP_6$ and inorganic phosphate concentrations provided a nutrient rich environment where the Δ*MINPP* was able to prosper. Again, it may be that the WT bacteria is responding to the presence of $InsP_6$ as shown by Figure 6.6 while the Δ*MINPP* is taking both inorganic phosphate and the inorganic phosphate released from the secreted phytase.

6.3. Conclusion

There is still much to be understood about the induction and repression of phytase genes, the expression of which does not appear to be uniformly controlled amongst bacteria (Jain *et al*., 2016; Konietzny and Greiner, 2004). In this chapter, phytase expression was first analysed using β-galactosidase assays in which the induction of the promoter region of AC1-2 MINPP, which lies at the beginning of a gene cluster, was measured in the exponential/late exponential with varying phytate and inorganic phosphate concentrations in minimal media, here there was a greater response in the presence of phytate than with inorganic phosphate.

qPCR was then conducted to directly measure AC1-2 expression in rich, LB medium and minimal media spiked with phytate. Surprisingly, despite the abundance of phosphate in the LB media, phytate expression remained unaltered whereas in the impure $InsP_6$ sample, which contained phosphate impurities (365 µM) phytase expression was heavily repressed by 7.5-fold, likewise the minimal media spiked with inorganic phosphate also saw phytase expression repressed by 2.6-fold. In the report by Fredrikson *et al* (Fredrikson *et al*., 2002) they noticed a similar occurrence, in the richer pea flour media that they used, access to inorganic phosphate was not sufficient to repress phytase synthesis and subsequent addition of inorganic phosphate still did not induce repression. Therefore, in this instance there may be a component in the LB media that is promoting the synthesis of the phytase. Whereas in the minimal media, phytase expression is repressed by the inorganic phosphate. In the paper by Boukhris *et al* working with *Streptomyces coelicolor* (Boukhris *et al*., 2016a) they identified that while expression of the phytase gene had been shown to be induced by phosphate limitation by the response regulator PhoP, the presence of a Direct Repeat putative operator site was the site of strong negative regulation by an unknown repressor.

Therefore, from the β-galactosidase assays which show an induction of the promoter regions from samples with impure phytate, there may be an additional mechanism that is causing repression of the phytase gene in the presence of phytate and inorganic phosphate which

was detected by qPCR. One such mechanism that has been described in the literature is how phytase synthesis has been repressed in the exponential phase of growth, only resuming once it has reach stationary/lag phase (Greppi *et al*., 2015; Konietzny and Greiner, 2004).

Additionally, a soil survivability experiment was performed to examine the competition of a mixed soil culture of the rif-resistant *Acinetobacter* sp. AC1-2 with the wild-type *MINPP* gene and the rif-resistant *Acinetobacter* sp. AC1-2 Δ*MINPP*. Three samples were examined, without added phosphate, with phytate, and with phytate and phosphate, these were left to grow for a week before being streaked onto LB plates and random colonies selected for PCR to identify if they contained the WT or truncated *MINPP* gene. The *Acinetobacter* AC1-2 *ΔMINPP* grew better in the conditions without phosphate, and with phytate and phosphate. While near even amounts of each bacterium were present in the phytate sample. This indicates that there may be an advantage to not secreting a phytase to increase the availability of inorganic phosphate as the *Acinetobacter* sp. AC1-2 Δ*MINPP* may be acting as 'social cheaters'. This experiment however, does need additional repeats to provide statistical significance.

6.4. Supplementary Information

Highlighted yellow areas were confirmed by sequencing results.

Promoter Region 1

>filtered DNA sequence consisting of 1030 bases.

gtagctacggcgtttggatgtaaaaagttacaatggctaatcttaaagatgctgaacg
tttaacgggatatttggtaggagggattagccctttaggccaaaaaaaacgtttaaaaac
tgtcattgatgcatcggctcaaaactttagcaaaatttatgtcagtggcggaaaaagagg
gctggatattggtttaaatcctaaagatttagcacagcttttaaatgctcaatttgttga
tgtgcttgatcagtaagaaaagtttagaggaaatactctaagaattaatcattatttta
tctacattgagaaaatactttgttgagaggtatagctcattgcggatccttctttatttt
tcaaaaaagagataagcacaaataattgcaaaaaagcttgaatcattaaatgaaaatgat
tatcattattgttaatggcgatgtagtcctagtcccgattaggggaggagattgattgt
aatataaagcccatattctaagtttatgctcattttctaatgaaatttagagttatacca
gaacctcaaacaagaagaacagcgaaagacgagttgatttctggtatgtgacacctgcaa
gacattataaaaatatgcgggttcaagcagcggtaaatggaactaagaaaaataaatact
atcagcgcaataaaaagaataaatcaagagtagcaagcgaggatggcttgctactcttgt
atttaaagttagtttcatacttatttatattgttcaacctttttagagttttagctctgc
tcagtaacagattacttatatttttgattaaaaaatgatcactagtgtttcatttattaa
catcctgccttttagccaatttttttgtataccctgaataaaacatgatcaaaaaagagg
taaaagcgtgcttatttttcatggaaaacctgttcatggagcaatttttgatatggatgg
cacaatgtttgatactgagcggttgcggtttcaaacattgcaacaagcatctcaagaact
cattggacagg

Promoter 2

>filtered DNA sequence consisting of 636 bases.

gcaggcgattcgtctgaatatgttgcaataacatctgtttgactgaatcatgcataaaa
ccaagtgcatgagcttttacaatgatgatgctaggaaggtgtaaaagaagtacttaaca
tcatcaagagttgcgttattaaaaataccttgttcaattaacgtctgctgcaaaaattgc
ttttcatttaatacagaatctgaagattcatcccaaaaacggtgacgcatcgccatcaga
ttgttgtgtttataactcataaaattgtcagctcagacttcaaatctaaataaaatatat
ggagtcgagcaacgaatatcaagttgagattatttttttcttcagttgctgtttgaatttt
tatgtaactcttttaaatagcctaagcaaatttaatttttttttaaaaaaaactataatc
gaaacaatttcatctataacactacttatcatgactcctgcatgtagattgttaaaaagt

aataaaatcgaattttctattcatgaatatgaacatgattcaaatgccaaaagttttgga

ttagaagcagctgaaaaattagatttagatgtaaacgaagtatttaaaactctaatggta

agtgatgataaaaattatttcgttgctgtcttgccgg

Promoter Region 3

>filtered DNA sequence consisting of 634 bases. reverse complement

ccatggcggtgtcattaagctattggcttgtttagccaagcaacaacctttagatgattt

acttaaaatgccagcagaacttgggaaattatattctctcgaattttctgaaatggatgg

tcaattaacttttatgctgagatagcaaatttatatttttatacacttgtaagttattat

ttttataaagatatttgaagtcatgtaatcgtactgtcattgctaaaaattatgctatgg

gcatttcttattaatcatttatcagacttatgaatattttatttaaaacgacgatgcttg

caacatcattatttcttgtggcctgcaacaataatgacgatcaagatgatcaaccaacga

catcacctacaactcaatcaaagtattatcaaaccaaaacgccttaccaaccacaacaag

atttaaaaagctatgaacaggcaccaaatggattccagccagtttttacagagttagtgg

cgcgtcatggttcaagaggtttatcaagtctcaaatatgatttagcactttataatttat

ggaagcaggcaaaagcagaaaatgccttaacgccgttaggtgagcaattaggtgctgatt

tagaagcaatgatgaaagccaatattctgttagg

Table S1 – RNA-extraction concentration and quality

Condition	RNA Conc. (ng/uL)	260/280	260/230
LB1	607.6	1.98	1.52
LB2	785.8	1.97	1.56
LB3	704.1	1.93	1.52
LB4	513.1	1.97	1.58
LB+IP61	873	1.95	0.76
LB+IP62	937.5	1.94	0.78
LB+IP63	549.3	1.87	0.84
LB+IP64	726.9	1.91	0.97
MM1	176.7	1.75	0.35
MM2	468.1	2	0.66
MM3	66	1.6	0.27
MM4	58.4	1.78	0.23
MM+IP61	500	2.01	1.84
MM+IP62	254.7	2.07	1.54
MM+IP63	584.2	2.02	1.04
MM+IP64	481.6	2.02	0.85

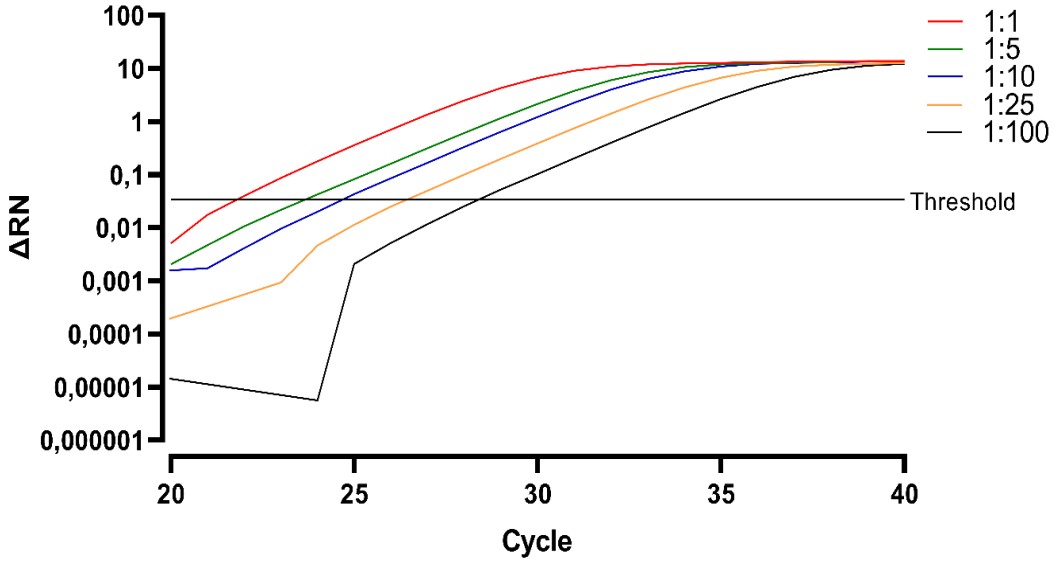


Figure S1 - The average ΔRN amplification data for Primer set 2 for each set of the serial dilutions in triplicate was plotted against the qPCR cycle number.

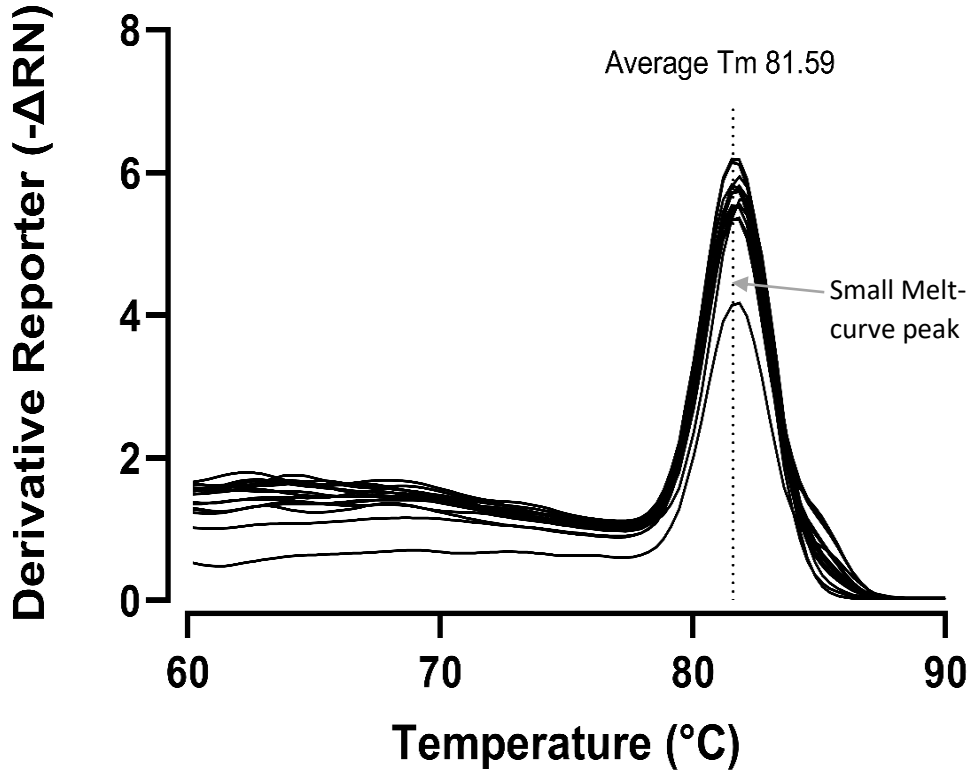


Figure S2 – Melt Curve Analysis of Primer Set 2 to asses the quality of the qPCR. Each dilution, 1:1, 1:5, 1:10, 1:25 and 1:100, in triplicate were analysed.

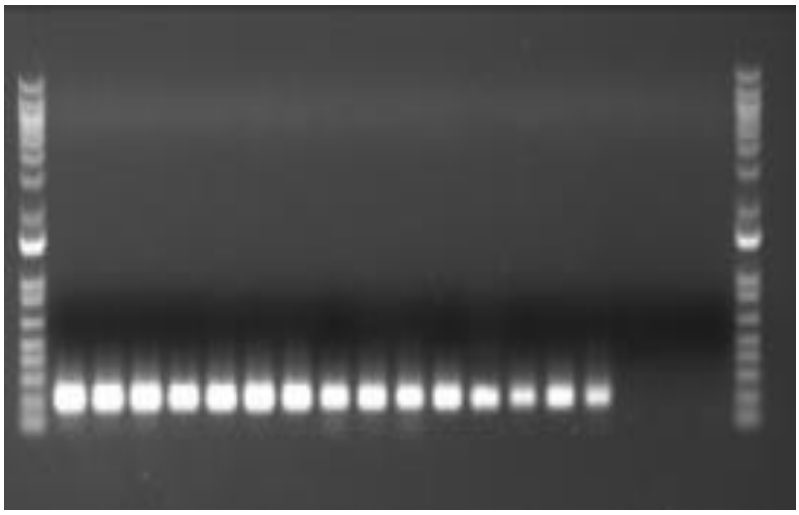


Figure S3 – The agarose gel-electrophoresis of the qPCR reactions for the Primer Set 2 dilutions.

Table S2 – The Ct values used to calculate Primer Efficiency

Standard conc. (cDNA ng/µl)	*Primer set 1*			*Primer set 2*		
	Tech rep 1	Tech rep 2	Tech rep 3	Tech rep 1	Tech rep 2	Tech rep 3
25	22.56	22.33	22.52	Removed	21.70	22.14
5	24.98	Removed	24.86	24.35	23.73	23.42
2.5	25.75	25.49	25.57	24.77	24.76	24.76
1	26.80	26.97	26.32	27.13	26.38	26.37
0.25	30.32	28.45	28.32	28.78	27.93	28.91

Standard conc. (cDNA ng/µL)	***Primer Set 1 (Average Ct)***	***Primer Set 2 (Average Ct)***
0.25	29.03	28.54
1	26.70	26.63
2.5	25.60	24.76
5	24.92	23.83
25	22.47	21.92

Log10 data	***Primer Set 2 (Average Ct)***	***Primer Set 3 Average Ct***
-0.60206	29.03	28.54
0	26.70	26.63
0.39794001	25.60	24.76
0.69897	24.92	23.83
1.39794001	22.47	21.92

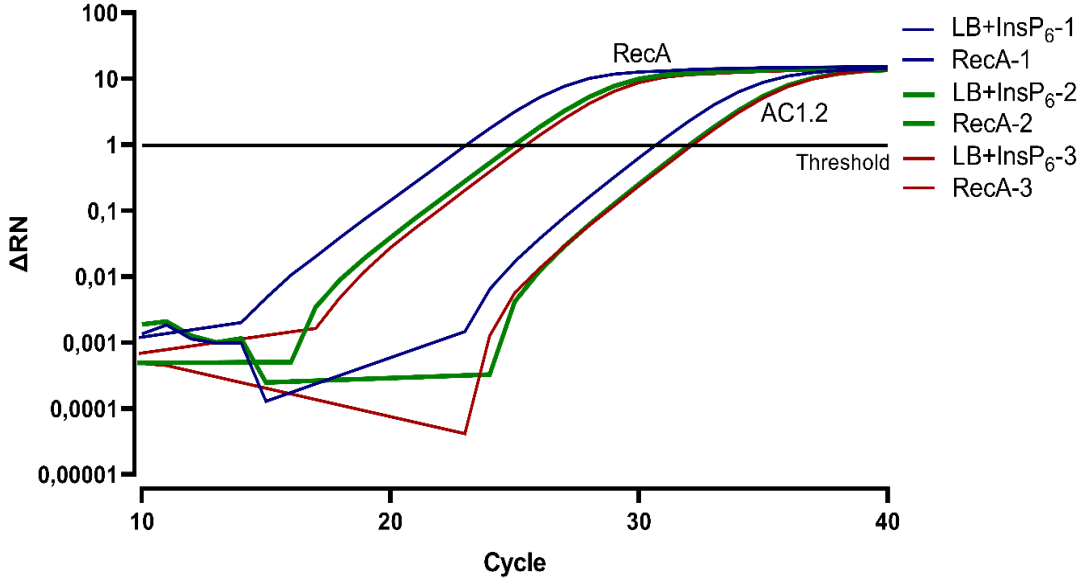


Figure S4 – qPCR plot displaying average ΔRN amplification versus cycle number for each LB + $InsP_6$ sample comparing the expression of the *MINPP* gene to the housekeeping gene *RecA*.

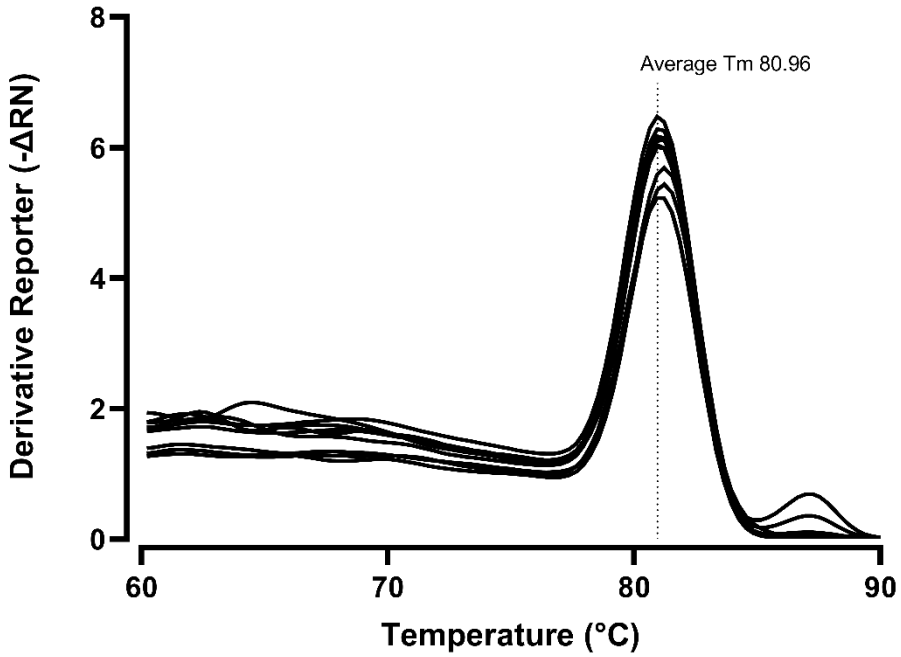


Figure S5 - Melt Curve Analysis of each LB + $InsP_6$ Sample dataset to assess the quality of the qPCR.

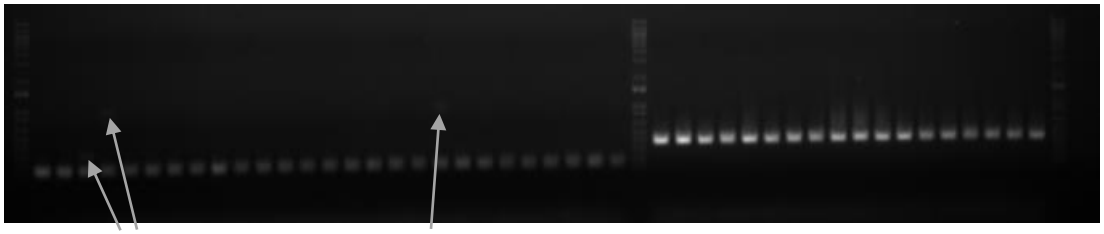


Figure S6 - The agarose gel-electrophoresis of the qPCR reactions for the LB + $InsP_6$ Samples. Lanes 1-9 LB-1, Lanes 10-18 LB-2, Lanes 19-27 LB-3, Lanes 29-37 RecA-1, Lanes 38-46 RecA-2.

Table S3 – Ct values for the LB and LB + $InsP_6$ Samples.

LB Sample Ct values

AC1.2 1	AC1.2 2	AC1.2 3	RecA 1	RecA 2	RecA 3
25.50	24.51	24.90	17.39	17.53	18.86
25.84	Removed	25.14	17.47	17.49	19.21
25.31	24.55	24.97	17.61	17.40	18.80
Removed	24.27	25.67	17.00	17.37	18.94
25.50	24.88	25.75	17.42	17.25	18.65
25.03	24.31	25.33	16.97	17.60	18.85
25.42	24.50	25.92	17.30	17.45	19.27
24.96	24.43	25.52	17.17	17.64	18.53
25.39	24.47	25.31	16.79	17.41	18.92

LB + $InsP_6$ Sample Ct Values

AC1.2 1	AC1.2 2	AC1.2 3	RecA 1	RecA 2	RecA 3
30.47	32.59	Removed	23.00	25.22	25.69
30.78	31.96	32.42	23.24	24.95	25.32
Removed	32.37	32.61	23.74	24.96	25.43
Removed	31.93	31.75	23.06	25.23	25.44
30.66	31.91	31.99	22.55	24.87	25.43
30.73	31.66	31.68	22.85	25.09	25.68
30.73	31.64	32.44	22.79	24.87	25.42
30.68	32.27	32.49	23.17	24.85	25.41
30.75	32.15	32.11	23.30	24.81	25.31

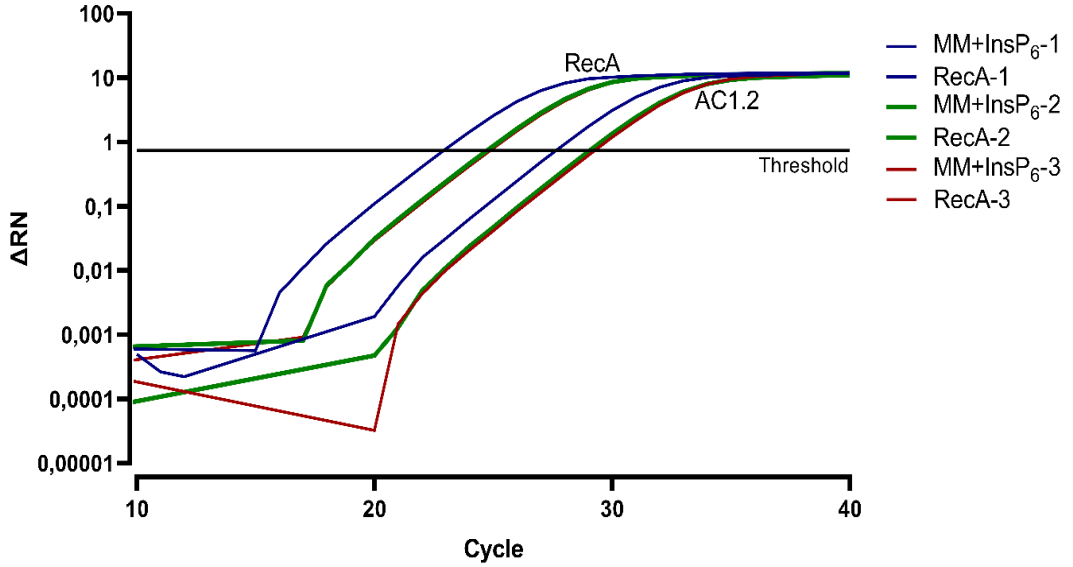


Figure S7 - qPCR plot displaying average ΔRN amplification versus cycle number for each MM + $InsP_6$ Sample comparing the expression of the *MINPP* gene to the housekeeping gene *RecA*.

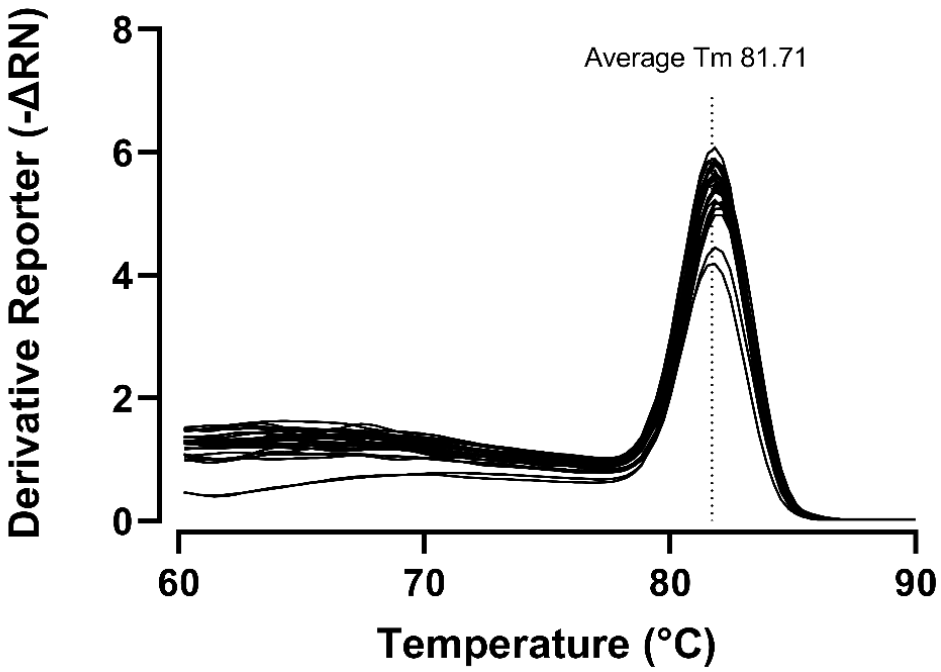


Figure S8 - Melt Curve Analysis of each MM + $InsP_6$ Sample dataset to assess the quality of the qPCR.

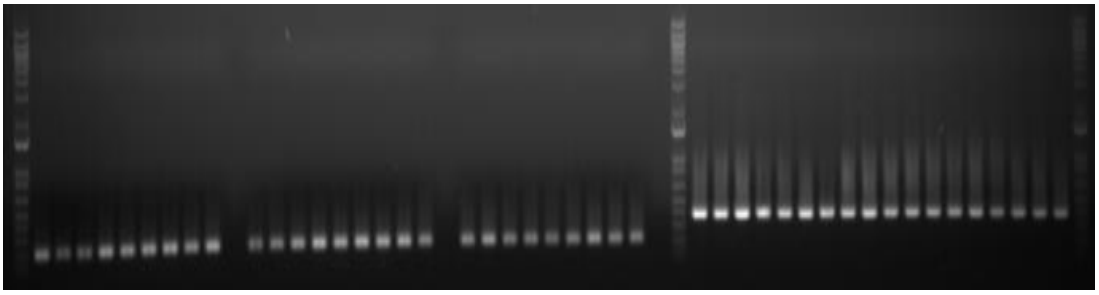


Figure S9 - The agarose gel-electrophoresis of the qPCR reactions for the MM + $InsP_6$ Samples. Lanes 1-9 MM+$InsP_6$-1, Lanes 11-19 MM+$InsP_6$-2, Lanes 21-29 MM+$InsP_6$-3, Lanes 32-40 RecA-1, Lanes 42-50 RecA-2.

Table S4 – Ct values for the MM and MM + $InsP_6$ Samples

MM Sample Ct Values					
AC1.2 1	**AC1.2 2**	**AC1.2 3**	**RecA 1**	**RecA 2**	**RecA 3**
26.92	25.26	27.47	24.44	23.26	26.49
26.70	25.53	27.01	24.49	22.93	25.19
26.60	25.62	26.93	24.41	23.17	25.32
26.98	25.41	28.00	24.32	22.92	25.56
26.99	25.70	27.64	24.35	22.98	24.98
26.78	25.47	27.42	24.27	22.86	25.14
26.55	25.90	26.77	24.37	22.84	25.56
26.70	25.74	27.23	24.39	23.16	24.84
26.46	25.55	28.76	24.43	22.78	25.18

MM+IP6 Sample Ct Values					
AC1.2 1	**AC1.2 2**	**AC1.2 3**	**RecA 1**	**RecA 2**	**RecA 3**
25.22	27.47	27.28	20.20	21.53	22.22
25.18	26.92	27.36	20.51	21.49	22.20
25.63	26.45	27.45	20.33	21.54	22.18
25.59	26.52	26.75	20.20	21.44	22.14
25.37	26.32	27.43	20.28	21.56	22.18
25.97	Removed	26.81	20.23	21.57	22.21
25.38	26.58	27.32	20.38	21.51	22.15
Removed	26.44	27.20	20.37	21.74	22.01
25.35	26.75	27.95	20.18	21.57	21.72

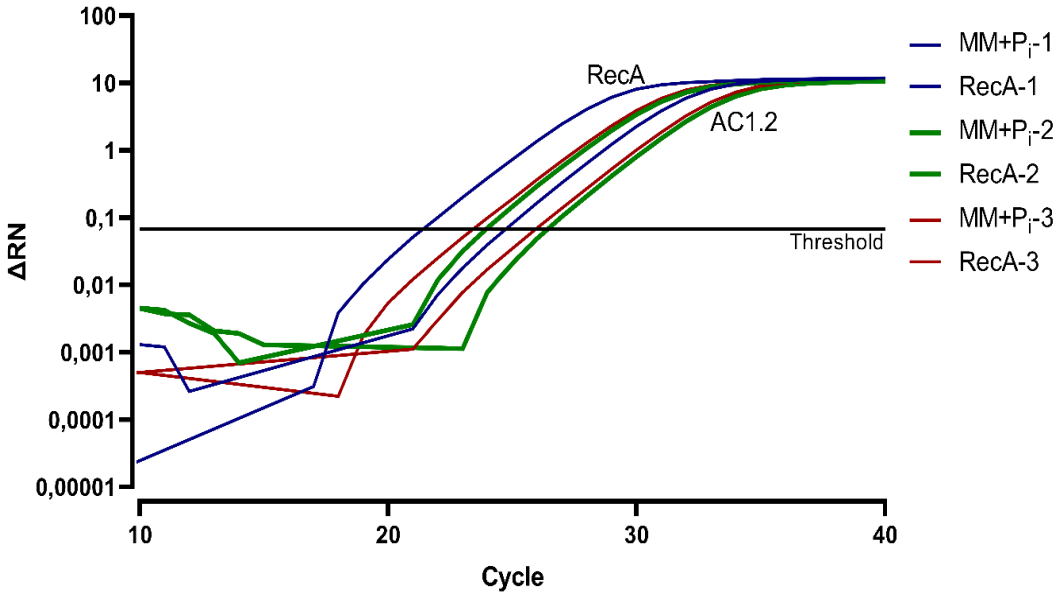


Figure S10 - qPCR plot displaying average ΔRN amplification versus cycle number for each MM + P_i Sample comparing the expression of the *MINPP* gene to the housekeeping gene *RecA*.

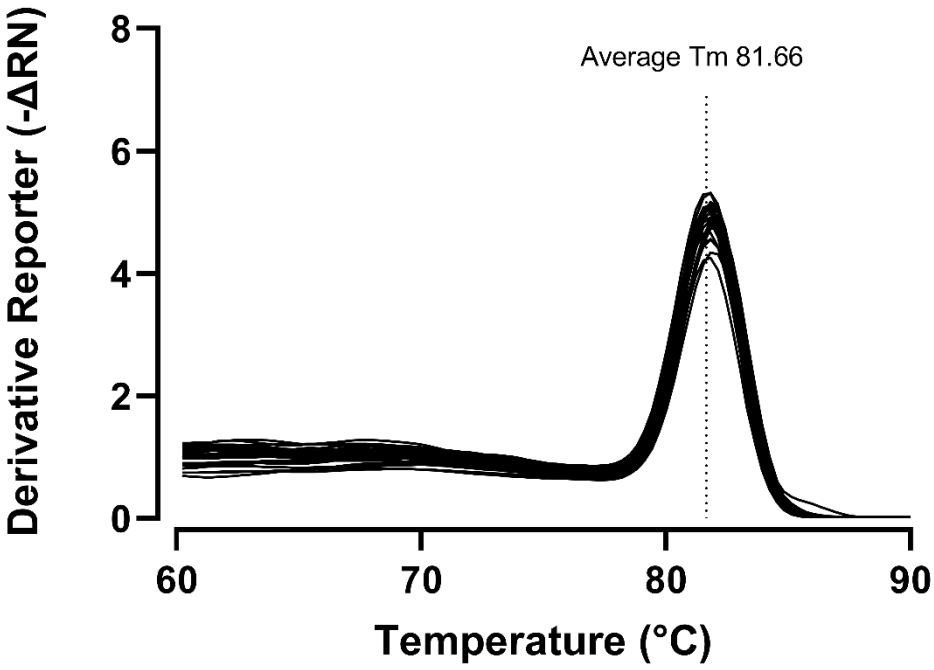


Figure S11 - Melt Curve Analysis of each MM + P_i Sample dataset to assess the quality of the qPCR.

Chapter 7. Final Discussion

This thesis investigated the diversity of phytases using both culture-dependent and culture-independent experiments in the soil and other environments.

Chapter 2 examined the current state of phytase isolation in the literature and noted many of the pitfalls of a frequently used isolation method "Phytase Specific Media," and developed a method in which phytate can be extracted from the agar plate and examined using HPLC. This method is typically based on the formation of clearing zones around the colony which is taken as signs of phytate degradation on the plate. However, HPLC analysis of three phytase positive controls still showed a dominant phytate peak, regardless of whether the phytate was had been degrading, suggesting the formation of clearing zones colony are not due to phytate degradation, and are most likely due to acid secretion from the bacteria (Bae *et al*., 1999; Fredrikson *et al*., 2002).

Following this, a new phytase isolation method was developed through examination of a mixed soil culture in minimal media spiked with phytate to induce a community response to synthesise phytase. As phytate degradation was observed from this mixed culture using HPLC, the culture was streaked onto agar plates and single colonies were isolated. In the first experiment, two soils were examined for phytate degradation and phytase-producing bacteria were isolated. From the agricultural soil, Fakenham, the bacteria *Acinetobacter* sp. was isolated, displaying a lack of positional specificity which is indicative of the MINPP phytases, a subgroup of Histidine Acid Phytases which are the primary class of phytase used commercially. This finding represented one of the first MINPPs to be isolated from the soil environment, as MINPPs were thought of as predominantly enteric in nature. The bacteria and gene were examined further in Chapter 3, 4 and 6.

A further long-term isolation experiment was performed in this chapter using three well-established Rothamsted soils. A total of 66 isolates were examined for phytate degradation, of which 17 were phytase active. These isolates showed a diverse array of phytase degradation profiles indicating that in the soil environment there are many different routes in which phytase degradation can occur and that they can be isolated and measured through culture-dependent means. Finally, the degradation profile of two core samples from Blakeney salt marsh were examined, these came from two pools of water, one with high concentrations of iron and one with high concentrations of sulphur. This experiment allowed the investigation of phytase degradation in the water, aerobic and anaerobic sediments of each core. In the iron environment, there was no sign of phytate degradation, most likely

due to the strength of adsorption of phytate onto iron particles that sequester phytate from being degraded. Despite this, there was the appearance of two unusual chromatogram peaks that are typically not associated with *myo*-$InsP_6$ degradation, but may be associated with a *scyllo*-$InsP_5$, the second peak running earlier may be a L/D-*chiro*-$InsP_4$ or a *scyllo*-$InsP_4$ (Whitfield *et al*., 2018). The sulphur environments exhibited phytase activity particularly the water sample which had strong $InsP_5$[1/3-OH] activity which is typically derived from the BPPhys. However, activity was still significantly slower in comparison with the previously tested soil samples that were further inland. This may be a feature of where the samples were taken, phytate that has leached into the waterways are acted upon by microbial phytases that there is little remaining activity as they reach the ocean (Stout *et al*., 2016), therefore the tidally washed pools may have limited exposure to phytate. It could also be due to the minimal media developed being less suitable to the saline environments of the salt marsh.

Chapter 3 details the expression trials and purification of the MINPP phytase from *Acinetobacter* sp. The *MINPP* was cloned into the expression host using gateway cloning, however initial effects only yielded an impure protein with minimal activity. The gene was then cloned without its signal peptide and expression trials were performed to establish the best conditions for protein expression. The final conditions used Rosetta™ 2 (pLysS) cells which were induced with 0.1 mM IPTG, producing a pure, active protein which was characterised in Chapter 3.

In Chapter 4, a thorough characterisation of AC1-2 MINPP was performed, many of the enzyme characteristics such as pH profile, substrate specificity, metal ion inhibitors, stability, thermostability and kinetics were examined using enzymatic assays and HPLC. The most interesting feature was its long-term stability which saw the enzyme retain as much as 100% of activity after 755 days on incubation at room temperature, while not the most active of phytases, this stability is an impressive feature of industrial concern. Additionally, the structure of AC1-2 MINPP was modelled using its closest sequence homologue, *Bl*MINPP. Through this, the specificity pockets involved in phytate binding were investigated and compared with other selected phytases with different positional specificities, three residues that were conserved for *Bl*MINPP, the other 4/6-phytase but not for *Bt*MINPP, a 5-phytase and *Af*HAPhy, a 3-phytase were identified as potential targets for mutagenic studies.

Chapter 5 provided a culture-independent, metagenomic examination of the diversity of the MINPPs, HAPhys, BPPhys and PTPhys in both environmental and enteric environments.

While currently not prominent in phytase research it showed the importance of the MINPPs in enteric environments, with the relative abundance as much as 16-36.2%, which is many times higher than that of the other phytases, 0.0101-4%. It also showed the abundance of the BPPhy in environmental metagenomes and in some cases, Gulf of Mexico, and Columbia River, being the only phytase type present in the environment. In addition to examining the diversity of phytases, I also examined the potential of horizontal gene transfer with the abundant MINPPs, particularly pertaining to transfers between different bacterial species. Due to the MINPPs dominance in enteric environments I was interested if we could see transfer of the gene from the enteric environment into the environmental. While 41 HGT events were identified with a bootstrap value greater than 50%, only 1 occurred between environments occurring between *Dygonomonas macrotermitis* from the hindgut of a termite to *Flavobacterium* in the soil environment.

Finally, in Chapter 6, both promoter:reporter and qPCR assays were performed to examine the expression of phytase in *Acinetobacter* sp. Three promoter:reporter constructs were prepared and analysed for activity by β-galactosidase assays, of which one produced sufficient activity. In the presence of increasing concentrations of impure phytate, there was an increased response in β-galactosidase activity and a diminished response in the presence of inorganic phosphate. In the qPCR analysis however, significant repression of the phytase gene occurred in both impure phytate, and inorganic phosphate conditions. Therefore, there may be an additional mechanism that is controlling the expression of the phytase.

To summarise, this thesis has made several important advances with regards to phytase understanding. One of the main advances was the development of new methods of phytase isolation in the soil environment and shown the breadth of phytase activity and diversity that can be found. The strength of this project lies in its use of HPLC to qualitatively determine phytase activity a feature which is commonly lacking in the literature. Frequently, there are many examples throughout the literature that have shown it to have an overreliance on the use of phytase isolation media without understanding of its limitations, which has been demonstrated on and improved in this thesis. Furthermore, the use of long-term incubations to allow for the development of diverse bacteria complemented with HPLC to show the diversity of phytase degradation that can be found, is a significant advance in culture-dependent studies. The next steps would be to repeat these experiments in a variety of different environments in combination with HPLC to further isolate diverse phytase producers with an array of different degradation patterns.

Another important advance in phytase understanding was the isolation of the MINPP phytase from the soil bacterium *Acinetobacter* sp. AC1-2, the first to be isolated from this environment. This phytase showed incredible stability over time that would be an important attribute for a candidate industrial phytase. There is, however, a lot of future work needed to be done for this enzyme to be suitable. One of the first experiments required would be to engineer improved thermostability so that it could withstand the high temperatures associated with the animal feed pelleting process. Experiments could include engineering new disulphide bonds, or by expressing the protein in a eukaryotic system to encourage glycosylation to take place, as was unsuccessfully attempted in this thesis (Dombkowski *et al.*, 2014; Niu *et al.*, 2016).

Combining this new discovery with the recent discoveries of a bacterial purple acid phytase, and unearthing new catalytic domains associated with phytase activity tells us that the current canonical phytases that are commonly considered, HAPhy, BPPhy, PAPhy and PTPhy, needs to be extended (Farias, Almeida and Meneses, 2018; Nasrabadi *et al.*, 2018; Villamizar, Funkner, *et al.*, 2019). The MINPPs also need to be added to the conversation of a commonly considered phytase as well. This was demonstrated in the environmental metagenomics shown in this thesis. The MINPPs were significantly more abundant in the metagenomes of enterics and monogastrics then all other phytase classes that they simply must be considered a common phytase. It is therefore very important that more research is conducted into this severely underrepresented class of phytase. This includes future isolation studies and characterisation of new MINPPs.

Additionally in this thesis, the controls of phytase expression for the MINPP from *Acinetobacter* sp. AC1-2 was examined using β-galactosidase assays and qPCR. This produced contradictory findings with the β-galactosidase assays showing an enhancement of expression in the presence of phytate, whereas repression was shown in the presence of phytate in the qPCR experiments. This was explained in Section 6.2.3.5 but showed an experimental shortcoming, as phytase expression had been shown to be repressed during the exponential phase of growth, which was when the RNA samples were taken. A future experiment would be to examine phytase expression during the stationary phase, to see if phytase expression is induced.

Overall, this thesis has provided the groundwork for future isolation studies and has demonstrated the importance and abundance of the MINPP class of phytase.

References

Abdollahi, M.R., Ravindran, V., Svihus, B., 2013. Pelleting of broiler diets: An overview with emphasis on pellet quality and nutritional value. Anim. Feed Sci. Technol. 179, 1–23. https://doi.org/10.1016/j.anifeedsci.2012.10.011

Acquistapace, I.M., Zi˛etek, M.A., Li, A.W.H., Salmon, M., Kühn, I., Bedford, M.R., Brearley, C.A., Hemmings, A.M., 2020. Snapshots during the catalytic cycle of a histidine acid phytase reveal an induced-fit structural mechanism. J. Biol. Chem. 295, 17724–17737. https://doi.org/10.1074/jbc.RA120.015925

Ahemad, M., Zaidi, A., Khan, M.S., Oves, M., 2009. Biological importance of phosphorus and phosphate solubilizing microbes - An overview, in: Phosphate Solubilizing Microbes for Crop Improvement, 1-14.

Alias, N., Shunmugam, S., Ong, P.Y., 2018. Isolation and molecular characterization of phytase producing bacteria from Malaysia hot springs. J. Fundam. Appl. Sci. 9, 852. https://doi.org/10.4314/jfas.v9i2s.53

Almagro Armenteros, J.J., Tsirigos, K.D., Sønderby, C.K., Petersen, T.N., Winther, O., Brunak, S., von Heijne, G., Nielsen, H., 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotechnol. 37, 420–423. https://doi.org/10.1038/s41587-019-0036-z

Andersen, J.N., Mortensen, O.H., Peters, G.H., Drake, P.G., Iversen, L.F., Olsen, O.H., Jansen, P.G., Andersen, H.S., Tonks, N.K., Møller, N.P.H., 2001. Structural and Evolutionary Relationships among Protein Tyrosine Phosphatase Domains. Mol. Cell. Biol. 21, 7117–7136. https://doi.org/10.1128/MCB.21.21.7117-7136.2001

Antler, G., Mills, J. V., Hutchings, A.M., Redeker, K.R., Turchyn, A. V., 2019. The Sedimentary Carbon-Sulfur-Iron Interplay – A Lesson From East Anglian Salt Marsh Sediments. Front. Earth Sci. 7, 140. https://doi.org/10.3389/feart.2019.00140

Aranda, S., Montes-Borrego, M., Landa, B.B., 2011. Purple-Pigmented Violacein-Producing Duganella spp. Inhabit the Rhizosphere of Wild and Cultivated Olives in Southern Spain. Microb. Ecol. 62, 446–459. https://doi.org/10.1007/s00248-011-9840-9

Archana, G., Buch, A., Kumar, G.N., 2012. Pivotal Role of Organic Acid Secretion by Rhizobacteria in Plant Growth Promotion, in: Microorganisms in Sustainable Agriculture and Biotechnology. Springer Netherlands, Dordrecht, pp. 35–53. https://doi.org/10.1007/978-94-007-2214-9_3

Ariza, A., Moroz, O. V., Blagova, E. V., Turkenburg, J.P., Waterman, J., Roberts, S.M., Vind, J., Sjøholm, C., Lassen, S.F., De Maria, L., Glitsoe, V., Skov, L.K., Wilson, K.S., 2013. Degradation of Phytate by the 6-Phytase from *Hafnia alvei*: A Combined Structural and Solution Study. PLoS One, 8, e65062. https://doi.org/10.1371/journal.pone.0065062

Ashley, K., Cordell, D., Mavinic, D., 2011. A brief history of phosphorus: From the philosopher's stone to nutrient recovery and reuse. Chemosphere, 84, 737–746. https://doi.org/10.1016/j.chemosphere.2011.03.001

Austin, B., 2017. The value of cultures to modern microbiology. Antonie Van Leeuwenhoek, 110(10), 1247–1256. https://doi.org/10.1007/s10482-017-0840-8

Azeke, M.A., Greiner, R., Jany, K.-D., 2011. Purification and Characterization of two intraceullular phytases from the tempeh fungus *Rhizopus oligosporus*. J. Food Biochem. 35, 213–227. https://doi.org/10.1111/j.1745-4514.2010.00377.x

Aziz, G., Nawaz, M., Anjum, A.A., Yaqub, T., Ahmed, M.U.D., Nazir, J., Khan, S.U., Aziz, K., 2015. Isolation and characterization of phytase producing bacterial isolates from soil. J. Anim. Plant Sci. 25, 771–776.

Baars, H., Ansmann, A., Althausen, D., Engelmann, R., Artaxo, P., Pauliquevis, T., Souza, R., 2011. Further evidence for significant smoke transport from Africa to Amazonia. Geophys. Res. Lett. 38(20). https://doi.org/10.1029/2011GL049200

Bae, H.D., Yanke, L.J., Cheng, K.J., Selinger, L.B., 1999. A novel staining method for detecting phytase activity. J. Microbiol. Methods, 39, 17–22. https://doi.org/10.1016/S0167-7012(99)00096-2

Baker-Austin, C., Oliver, J.D., Alam, M., Ali, A., Waldor, M.K., Qadri, F., Martinez-Urtaza, J., 2018. Vibrio spp. infections. Nat. Rev. Dis. Prim. 4, 1–19. https://doi.org/10.1038/s41572-018-0005-8

Balaban, N.P., Suleimanova, A.D., Shakirov, E. V., Sharipova, M.R., 2018. Histidine Acid Phytases of Microbial Origin. Microbiology, 87, 745–756. https://doi.org/10.1134/S0026261718060024

Balderrama-Subieta, A., Quillaguamán, J., 2013. Genomic studies on nitrogen metabolism in *Halomonas boliviensis*: Metabolic pathway, biochemistry and evolution. Comput. Biol. Chem. 47, 96–104. https://doi.org/10.1016/j.compbiolchem.2013.08.002

Baranes-Bachar, K., Levy-Barda, A., Oehler, J., Reid, D.A., Soria-Bretones, I., Voss, T.C., Chung, D., Park, Y., Liu, C.,

Yoon, J.B., Li, W., Dellaire, G., Misteli, T., Huertas, P., Rothenberg, E., Ramadan, K., Ziv, Y., Shiloh, Y., 2018. The Ubiquitin E3/E4 Ligase UBE4A Adjusts Protein Ubiquitylation and Accumulation at Sites of DNA Damage, Facilitating Double-Strand Break Repair. Mol Cell, 69, 866-878 e7. https://doi.org/10.1016/j.molcel.2018.02.002

Beardmore-Gray, M., Anthony, C., 1986. The Oxidation of Glucose by *Acinetobacter calcoaceticus*: Interaction of the Quinoprotein Glucose Dehydrogenase with the Electron Transport Chain. Microbiology, 132, 1257–1268. https://doi.org/10.1099/00221287-132-5-1257

Becken, S., 2014. Oil depletion or a market problem? A framing analysis of peak oil in The Economist news magazine. Energy Res. Soc. Sci. 2, 125–134. https://doi.org/10.1016/j.erss.2014.03.018

Beeson, L.A., Walk, C.L., Bedford, M.R., Olukosi, O.A., 2017. Hydrolysis of phytate to its lower esters can influence the growth performance and nutrient utilization of broilers with regular or super doses of phytase. Poultry Science, 96, 2243–2253. https://doi.org/10.3382/ps/pex012

Belval, L., Marquette, A., Mestre, P., Piron, M.-C., Demangeat, G., Merdinoglu, D., Chich, J.-F., 2015. A fast and simple method to eliminate Cpn60 from functional recombinant proteins produced by *E. coli* Arctic Express. Protein Expr. Purif. 109, 29–34. https://doi.org/10.1016/j.pep.2015.01.009

Bentley, S.D., Chater, K.F., Cerdeño-Tárraga, A.-M., Challis, G.L., Thomson, N.R., James, K.D., Harris, D.E., Quail, M.A., Kieser, H., Harper, D., Bateman, A., Brown, S., Chandra, G., Chen, C.W., Collins, M., Cronin, A., Fraser, A., Goble, A., Hidalgo, J., Hornsby, T., Howarth, S., Huang, C.-H., Kieser, T., Larke, L., Murphy, L., Oliver, K., O'Neil, S., Rabbinowitsch, E., Rajandream, M.-A., Rutherford, K., Rutter, S., Seeger, K., Saunders, D., Sharp, S., Squares, R., Squares, S., Taylor, K., Warren, T., Wietzorrek, A., Woodward, J., Barrell, B.G., Parkhill, J., Hopwood, D.A., 2002. Complete genome sequence of the model actinomycete *Streptomyces coelicolor* A3(2). Nature, 417, 141–147. https://doi.org/10.1038/417141a

Berini, F., Casciello, C., Marcone, G.L., Marinelli, F., 2017. Metagenomics: novel enzymes from non-culturable microbes. FEMS Microbiol. Lett. 364(21), fnx211. https://doi.org/10.1093/femsle/fnx211

Bernardet, J.-F., Bowman, J.P., 2006. The Genus *Flavobacterium*, in: The Prokaryotes. Springer New York, New York, NY, 481–531. https://doi.org/10.1007/0-387-30747-8_17

Berridge, M.J., 2008. Inositol trisphosphate and calcium signalling mechanisms. BBA - Mol. Cell Res. 1793, 933–940. https://doi.org/10.1016/j.bbamcr.2008.10.005

Berridge, M.J., Irvine, R.F., 1989. Inositol phosphates and cell signalling. Nature, 341, 197–205. https://doi.org/10.1038/341197a0

Bhattacharyya, P., Nayak, A.K., Shahid, M., Tripathi, R., Mohanty, S., Kumar, A., Raja, R., Panda, B.B., Lal, B., Gautam, P., Swain, C.K., Roy, K.S., Dash, P.K., 2015. Effects of 42-year long-term fertilizer management on soil phosphorus availability, fractionation, adsorption–desorption isotherm and plant uptake in flooded tropical rice. Crop J. 3, 387–395. https://doi.org/10.1016/j.cj.2015.03.009

Boc, A., Diallo, A.B., Makarenkov, V., 2012. T-REX: a web server for inferring, validating and visualizing phylogenetic trees and networks. Nucleic Acids Res. 40, W573–W579. https://doi.org/10.1093/nar/gks485

Bodor, A., Bounedjoum, N., Vincze, G.E., Erdeiné Kis, Á., Laczi, K., Bende, G., Szilágyi, Á., Kovács, T., Perei, K., Rákhely, G., 2020. Challenges of unculturable bacteria: environmental perspectives. Rev. Environ. Sci. Bio/Technology, 19(1), 1–22. https://doi.org/10.1007/s11157-020-09522-4

Böhm, K., Herter, T., Müller, J.J., Borriss, R., Heinemann, U., 2010. Crystal structure of Klebsiella sp. ASR1 phytase suggests substrate binding to a preformed active site that meets the requirements of a plant rhizosphere enzyme. FEBS J. 277, 1284–1296. https://doi.org/10.1111/j.1742-4658.2010.07559.x

Bolen, D., Baskakov, I. V, 2001. The osmophobic effect: natural selection of a thermodynamic force in protein folding. J. Mol. Biol. 310, 955–963. https://doi.org/10.1006/jmbi.2001.4819

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics, 30, 2114–2120. https://doi.org/10.1093/bioinformatics/btu170

Boukhris, I., Dulermo, T., Chouayekh, H., Virolle, M.-J., 2016a. Evidence for the negative regulation of phytase gene expression in *Streptomyces lividans* and *Streptomyces coelicolor*. J. Basic Microbiol. 56, 59–66. https://doi.org/10.1002/jobm.201500417

Boukhris, I., Farhat-Khemakhem, A., Bouchaala, K., Virolle, M.-J., Chouayekh, H., 2016b. Cloning and characterization of the first actinomycete β-propeller phytase from *Streptomyces* sp. US42. J. Basic Microbiol. 56, 1080–1089. https://doi.org/10.1002/jobm.201500760

Brinch-Pedersen, H., Sørensen, L.D., Holm, P.B., 2002. Engineering crop plants: Getting a handle on phosphate. Trends Plant Sci, 7(3), 118-125. https://doi.org/10.1016/S1360-1385(01)02222-1

Broch, J., Nunes, R. V., Eyng, C., Pesti, G.M., de Souza, C., Sangalli, G.G., Fascina, V., Teixeira, L., 2018. High levels of dietary phytase improves broiler performance. Animal Feed Science and Technology. 244, 56–65. https://doi.org/10.1016/j.anifeedsci.2018.06.001

Cai, X.-C., Xi, H., Liang, L., Liu, J.-D., Liu, C.-H., Xue, Y.-R., Yu, X.-Y., 2017. Rifampicin-Resistance mutations in the rpoB Gene in *Bacillus velezensis* CC09 have Pleiotropic Effects. Front. Microbiol. 8, 178. https://doi.org/10.3389/fmicb.2017.00178

Carver, T., Harris, S.R., Berriman, M., Parkhill, J., McQuillan, J.A., 2012. Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics, 28, 464–469. https://doi.org/10.1093/bioinformatics/btr703

Castillo Villamizar, G.A., Funkner, K., Nacke, H., Foerster, K., Daniel, R., 2019a. Functional Metagenomics Reveals a New Catalytic Domain, the Metallo-β-Lactamase Superfamily Domain, Associated with Phytase Activity. mSphere 4, e00167-19. https://doi.org/10.1128/mSphere.00167-19

Castillo Villamizar, G.A., Nacke, H., Boehning, M., Herz, K., Daniel, R., 2019b. Functional Metagenomics Reveals an Overlooked Diversity and Novel Features of Soil-Derived Bacterial Phosphatases and Phytases. MBio 10(1), e01966-18. https://doi.org/10.1128/mBio.01966-18

Castillo Villamizar, G.A., Nacke, H., Griese, L., Tabernero, L., Funkner, K., Daniel, R., 2019c. Characteristics of the first Protein Tyrosine Phosphatase with Phytase Activity from a Soil Metagenome. Genes, 10(2), 101. https://doi.org/10.3390/genes10020101

Cervin, M. A., Kensch, O., Kettling, U., Leuthner, B., Mias-nikov, A., and Pellengahr, K., 2008. Variant *Buttiauxella* sp. phytases having altered properties. Danisco US INC Genencor DIV, US8143036B2

Chanderman, A., Puri, A.K., Permaul, K., Singh, S., 2016. Production, characteristics and applications of phytase from a rhizosphere isolated *Enterobacter* sp. ACSS. Bioprocess Biosyst. Eng. 39, 1577–1587. https://doi.org/10.1007/s00449-016-1632-7

Chase, A.B., Arevalo, P., Polz, M.F., Berlemont, R., Martiny, J.B.H., 2016. Evidence for Ecological Flexibility in the Cosmopolitan Genus *Curtobacterium*. Front. Microbiol. 7, 1874. https://doi.org/10.3389/fmicb.2016.01874

Chen, X., Jiang, N., Condron, L.M., Dunfield, K.E., Chen, Z., Wang, J., Chen, L., 2019. Impact of long-term phosphorus fertilizer inputs on bacterial phoD gene community in a maize field, Northeast China. Sci Total Env. 669, 1011–1018. https://doi.org/10.1016/j.scitotenv.2019.03.172

Chen, Y.P., Rekha, P.D., Arun, A.B., Shen, F.T., Lai, W.-A., Young, C.C., 2006. Phosphate solubilizing bacteria from subtropical soil and their tricalcium phosphate solubilizing abilities. Appl. Soil Ecol. 34, 33–41. https://doi.org/10.1016/j.apsoil.2005.12.002

Cheng, C., Lim, B.L., 2006. Beta-propeller phytases in the aquatic environment. Arch. Microbiol. 185, 1–13. https://doi.org/10.1007/s00203-005-0080-6

Chi, H., Tiller, G.E., Dasouki, M.J., Romano, P.R., Wang, J., O'Keefe, R.J., Puzas, J.E., Rosier, R.N., Reynolds, P.R., 1999. Multiple Inositol Polyphosphate Phosphatase: Evolution as a Distinct Group within the Histidine Phosphatase Family and Chromosomal Localization of the Human and Mouse Genes to Chromosomes 10q23 and 19. Genomics, 56(3), 324–336. https://doi.org/10.1006/geno.1998.5736

Chi, H., Yang, X., Kingsley, P.D., O'Keefe, R.J., Puzas, J.E., Rosier, R.N., Shears, S.B., Reynolds, P.R., 2000. Targeted deletion of Minpp1 provides new insight into the activity of multiple inositol polyphosphate phosphatase in vivo. Mol Cell Biol, 20, 6496–6507. https://doi.org/10.1128/mcb.20.17.6496-6507.2000

Chien, S.H., Prochnow, L.I., Tu, S., Snyder, C.S., 2011. Agronomic and environmental aspects of phosphate fertilizers varying in source and solubility: an update review. Nutr. Cycl. Agroecosystems, 89, 229–255. https://doi.org/10.1007/s10705-010-9390-4

Cho, J., Choi, K., Darden, T., Reynolds, P.R., Petitte, J.N., Shears, S.B., 2006. Avian multiple inositol polyphosphate phosphatase is an active phytase that can be engineered to help ameliorate the planet's "phosphate crisis". J. Biotechnol. 126, 248–259. https://doi.org/10.1016/j.jbiotec.2006.04.028

Choulet, F., Aigle, B., Gallois, A., Mangenot, S., Gerbaud, C., Truong, C., Francou, F.-X., Fourrier, C., Guérineau, M., Decaris, B., Barbe, V., Pernodet, J.-L., Leblond, P., 2006. Evolution of the Terminal Regions of the Streptomyces Linear Chromosome. Mol. Biol. Evol. 23, 2361–2369. https://doi.org/10.1093/molbev/msl108

Chu, H.M., Guo, R.T., Lin, T.W., Chou, C.C., Shr, H.L., Lai, H.L., Tang, T.Y., Cheng, K.J., Selinger, B.L., Wang, A.H.J., 2004. Structures of *Selenomonas ruminantium* phytase in complex with persulfated phytate: DSP phytase fold and mechanism for sequential substrate hydrolysis. Structure, 12(11), 2015–2024. https://doi.org/10.1016/j.str.2004.08.010

Coker, O.O., Warit, S., Rukseree, K., Summpunn, P., Prammananan, T., Palittapongarnpim, P., 2013. Functional characterization of two members of histidine phosphatase superfamily in *Mycobacterium tuberculosis*. BMC Microbiol. 13(1). 1-12. https://doi.org/10.1186/1471-2180-13-292

Conway, T., 1992. The Entner-Doudoroff pathway: history, physiology and molecular biology. FEMS Microbiol. Lett. 103, 1–28. https://doi.org/10.1111/j.1574-6968.1992.tb05822.x

Cooper, J., Lombardi, R., Boardman, D., Carliell-Marquet, C., 2011. The future distribution and production of global phosphate rock reserves. Resour. Conserv. Recycl. 57, 78–86. https://doi.org/10.1016/j.resconrec.2011.09.009

Cordell, D., Drangert, J.-O., White, S., 2009. The story of phosphorus: Global food security and food for thought. Glob. Environ. Chang. 19, 292–305. https://doi.org/10.1016/j.gloenvcha.2008.10.009

Cordell, D., Neset, T.-S.S., Prior, T., 2012. The phosphorus mass balance: identifying 'hotspots' in the food system as a roadmap to phosphorus security. Curr. Opin. Biotechnol. 23, 839–845. https://doi.org/10.1016/j.copbio.2012.03.010

Cordell, D., Rosemarin, A., Schröder, J.J., Smit, A.L., 2011. Towards global phosphorus security: A systems framework for phosphorus recovery and reuse options. Chemosphere, 84, 747–758. https://doi.org/10.1016/j.chemosphere.2011.02.032

Cordell, D., White, S., 2015. Tracking phosphorus security: indicators of phosphorus vulnerability in the global food system. Food Secur. 7(2), 337-350. https://doi.org/10.1007/s12571-015-0442-0

Cordell, D., White, S., 2014. Life's Bottleneck: Sustaining the World's Phosphorus for a Food Secure Future. Annu. Rev. Environ. Resour. 39, 161–188. https://doi.org/10.1146/annurev-environ-010213-113300

Cordell, D., White, S., 2013. Sustainable Phosphorus Measures: Strategies and Technologies for Achieving Phosphorus Security. Agronomy, 3(1), 86-116. https://doi.org/10.3390/agronomy3010086

Cordell, D., White, S., 2011. Peak Phosphorus: Clarifying the Key Issues of a Vigorous Debate about Long-Term Phosphorus Security. Sustainability, 3, 2027–2049. https://doi.org/10.3390/su3102027

Corretto, E., Antonielli, L., Sessitsch, A., Höfer, C., Puschenreiter, M., Widhalm, S., Swarnalakshmi, K., Brader, G., 2020. Comparative Genomics of Microbacterium Species to Reveal Diversity, Potential for Secondary Metabolites and Heavy Metal Resistance. Front. Microbiol. 11, 1869. https://doi.org/10.3389/fmicb.2020.01869

Cosgrove, D and Irving, C., 1980. Inositol phosphates: their chemistry, biochemistry and physiology, Elsevier Scientific Pub. Co.

Cosgrove, D.J., Irving, G.C.J., Bromfield, S.M., 1970. Inositol phosphate phosphatases of microbiological origin. The isolation of soil bacteria having inositol phosphate phosphatase activity. Aust. J. Biol. Sci. 23, 339–344. https://doi.org/10.1071/BI9700339

Cowieson, A.J., Ruckebusch, J.P., Knap, I., Guggenbuhl, P., Fru-Nji, F., 2016. Phytate-free nutrition: A new paradigm in monogastric animal production. Anim. Feed Sci. Technol. 222, 180–189. https://doi.org/10.1016/j.anifeedsci.2016.10.016

Craxton, A., Ali, N., Shears, S.B., 1995. Comparison of the activities of a multiple inositol polyphosphate phosphatase obtained from several sources: a search for heterogeneity in this enzyme. Biochem. J. 305, 491–498. https://doi.org/10.1042/bj3050491

Craxton, A., Caffrey, J.J., Burkhart, W., Safrany, T.S., Shears, B.S., 1997. Molecular cloning and expression of a rat hepatic multiple inositol polyphosphate phosphatase. Biochem. J. 328, 75–81. https://doi.org/10.1042/bj3280075

Crea, F., De Stefano, C., Milea, D., Sammartano, S., 2008. Formation and stability of phytate complexes in solution. Coord. Chem. Rev. 252, 1108–1120. https://doi.org/10.1016/j.ccr.2007.09.008

Cridland, C., Gillaspy, G., 2020. Inositol Pyrophosphate Pathways and Mechanisms: What Can We Learn from Plants? Molecules, 25, 2789. https://doi.org/10.3390/molecules25122789

Cromwell, G.L., 2009. ASAS Centennial Paper: Landmark discoveries in swine nutrition in the past century. J. Anim. Sci. 87, 778–792. https://doi.org/10.2527/jas.2008-1463

Crouch, S.R., Malmstadt, H. V., 1967. A Mechanistic Investigation of Molybdenum Blue Method for Determination of Phosphate. Anal. Chem. 39, 1084–1089. https://doi.org/10.1021/ac60254a027

Cullather, N., 2004. Miracles of Modernization: the Green Revolution and the apotheosis of technology. Dipl. Hist. 28, 227–254. https://doi.org/10.1111/j.1467-7709.2004.00407.x

Curtis, F., 2009. Peak globalization: Climate change, oil depletion and global trade. Ecol. Econ. 69, 427–434.

https://doi.org/10.1016/j.ecolecon.2009.08.020

Curtis, T.P., 2005. Microbiology: Exploring Microbial Diversity--A Vast Below. Science, 309(5739), 1331–1333. https://doi.org/10.1126/science.1118176

Dandapath, I., Chatterjee, M., Sarkar, D., Gupta, A., Rabbani, G., Minakshi, R., 2017. Bacterial osmolyte system and its physiological roles, in: Cellular Osmolytes, 229–249. https://doi.org/10.1007/978-981-10-3707-8_10

Daneshgar, S., Callegari, A., Capodaglio, A., Vaccari, D., 2018. The Potential Phosphorus Crisis: Resource Conservation and Possible Escape Technologies: A Review. Resources, 7, 37. https://doi.org/10.3390/resources7020037

Daniel, R., 2005. The metagenomics of soil. Nat. Rev. Microbiol, 3(6), 470–478. https://doi.org/10.1038/nrmicro1160

Dao, T.H., 2005. Ligand effects on inositol phosphate solubility and bioavailability in animal manures., in: Inositol Phosphates: Linking Agriculture and the Environment, Inositol phosphates, 169–185. https://doi.org/10.1079/9781845931520.0169

Davis, K.E.R., Joseph, S.J., Janssen, P.H., 2005. Effects of Growth Medium, Inoculum Size, and Incubation Time on Culturability and Isolation of Soil Bacteria. Appl. Environ. Microbiol. 71(2), 826–834. https://doi.org/10.1128/AEM.71.2.826-834.2005

Dawson, R.A., Larke-Mejía, N.L., Crombie, A.T., Ul Haque, M.F., Murrell, J.C., 2020. Isoprene Oxidation by the Gram-Negative Model bacterium *Variovorax* sp. WS11. Microorganisms, 8(3), 349. https://doi.org/10.3390/microorganisms8030349

De Groot, C.J., Golterman, H.L., 1993. On the presence of organic phosphate in some Camargue sediments: evidence for the importance of phytate. Hydrobiologia, 252, 117–126.

de Kraker, M.E.A., Stewardson, A.J., Harbarth, S., 2016. Will 10 Million People Die a Year due to Antimicrobial Resistance by 2050? PLOS Med. 13, e1002184. https://doi.org/10.1371/journal.pmed.1002184

DeLano, W., 2002. Pymol: An open-source molecular graphics tool. CCP4 Newsl. Protein Crystallogr, 40(1), 82-92.

Dersjant-Li, Y., Awati, A., Schulze, H., Partridge, G., 2015. Phytase in non-ruminant animal nutrition: A critical review on phytase activities in the gastrointestinal tract and influencing factors. J. Sci. Food Agric, 95(5), 878-896. https://doi.org/10.1002/jsfa.6998

Desmidt, E., Ghyselbrecht, K., Zhang, Y., Pinoy, L., Van der Bruggen, B., Verstraete, W., Rabaey, K., Meesschaert, B., 2015. Global Phosphorus Scarcity and Full-Scale P-Recovery Techniques: A Review. Crit. Rev. Environ. Sci. Technol. 45, 336–384. https://doi.org/10.1080/10643389.2013.866531

Dheda, K., Huggett, J.F., Bustin, S.A., Johnson, M.A., Rook, G., Zumla, A., 2004. Validation of housekeeping genes for normalizing RNA expression in real-time PCR. Biotechniques, 37, 112–119. https://doi.org/10.2144/04371RR03

Di Paolo, G., De Camilli, P., 2006. Phosphoinositides in cell regulation and membrane dynamics. Nature, 443(7112), 651-657. https://doi.org/10.1038/nature05185

Dilfuza Egamberdieva, P.A., 2018. Plant Microbiome: Stress Response, Microorganisms for Sustainability, Springer Singapore. https://doi.org/10.1007/978-981-10-5514-0

Dionisio, G., Holm, P.B., Brinch-Pedersen, H., 2007. Wheat (*Triticum aestivum* L.) and barley (*Hordeum vulgare* L.) multiple inositol polyphosphate phosphatases (MINPPs) are phytases expressed during grain filling and germination. Plant Biotechnol. J. 5, 325–338. https://doi.org/10.1111/j.1467-7652.2007.00244.x

Dionisio, G., Madsen, C.K., Holm, P.B., Welinder, K.G., Jørgensen, M., Stoger, E., Arcalis, E., Brinch-Pedersen, H., 2011. Cloning and Characterization of Purple Acid Phosphatase Phytases from Wheat, Barley, Maize, and Rice. Plant Physiol. 156, 1087–1100. https://doi.org/10.1104/pp.110.164756

Dokuzparmak, E., Sirin, Y., Cakmak, U., Saglam Ertunga, N., 2017. Purification and characterization of a novel thermostable phytase from the thermophilic *Geobacillus* sp. TF16. Int. J. Food Prop. 20, 1104–1116. https://doi.org/10.1080/10942912.2016.1203930

Dombkowski, A.A., Sultana, K.Z., Craig, D.B., 2014. Protein disulfide engineering. FEBS Lett. 588, 206–212. https://doi.org/10.1016/j.febslet.2013.11.024

Doolette, A.L., Smernik, R.J., Dougherty, W.J., 2010. Rapid decomposition of phytate applied to a calcareous soil demonstrated by a solution 31P NMR study. Eur. J. Soil Sci. 61, 563–575. https://doi.org/10.1111/j.1365-2389.2010.01259.x

Dorsch, J.A., Cook, A., Young, K.A., Anderson, J.M., Bauman, A.T., Volkmann, C.J., Murthy, P.P., Raboy, V., 2003. Seed phosphorus and inositol phosphate phenotype of barley low phytic acid genotypes. Phytochemistry, 62, 691–706.

Dox, A.W., Golden, R., 1911. Phytase in lower fungi. J. Biol. Chem. 10, 183–186. https://doi.org/10.1016/S0021-9258(18)88798-1

Duffield, T., Plaizier, J.C., Fairfield, A., Bagg, R., Vessie, G., Dick, P., Wilson, J., Aramini, J., McBride, B., 2004. Comparison of Techniques for Measurement of Rumen pH in Lactating Dairy Cows. J. Dairy Sci. 87, 59–66. https://doi.org/10.3168/jds.S0022-0302(04)73142-2

Duong-Ly, K.C., Gabelli, S.B., 2014. Explanatory Chapter: Troubleshooting Protein Expression: what to do when the protein in not soluble. 541, 231–247. https://doi.org/10.1016/B978-0-12-420119-4.00018-5

Dutta, C., Pan, A., 2002. Horizontal gene transfer and bacterial diversity. J. Biosci. 27, 27–33. https://doi.org/10.1007/BF02703681

Eddy, S.R., 2011. Accelerated Profile HMM Searches. PLoS Comput. Biol. 7, e1002195. https://doi.org/10.1371/journal.pcbi.1002195

Eddy, S.R., 1998. Profile hidden Markov models. Bioinformatics 14, 755–763. https://doi.org/10.1093/bioinformatics/14.9.755

El-Toukhy, N.M., Youssef, A.S., Mikhail, M.G., 2016. Isolation, purification and characterization of phytase from *Bacillus subtilis* MJA. African J. Biotechnol, 12(20).

Engelen, A.J., Van Der Heeft, F.C., Randsdorp, P.H.G., Smtt, E.L.C., 1994. Simple and Rapid Determination of Phytase Activity. J. AOAC Int. 77, 760–764. https://doi.org/10.1093/jaoac/77.3.760

Escobin-Mopera, L., Ohtani, M., Sekiguchi, S., Sone, T., Abe, A., Tanaka, M., Meevootisom, V., Asano, K., 2012. Purification and characterization of phytase from *Klebsiella pneumoniae* 9-3B. J. Biosci. Bioeng. 113, 562–567. https://doi.org/10.1016/j.jbiosc.2011.12.010

Faba Rodriguez, R., 2018. Structure-function studies of a purple acid phytase, PhD thesis, University of East Anglia, England.

Fakhravar, A., Hesampour, A., 2018. Rational design-based engineering of a thermostable phytase by site-directed mutagenesis. Mol. Biol. Rep. 45, 2053–2061. https://doi.org/10.1007/s11033-018-4362-x

Fallingborg, J., 1999. Intraluminal pH of the human gastrointestinal tract. Dan. Med. Bull. 46, 183–96.

Farias, N., Almeida, I., Meneses, C., 2018. New Bacterial Phytase through Metagenomic Prospection. Molecules, 23, 448. https://doi.org/10.3390/molecules23020448

Faruque, S.M., Albert, M.J., Mekalanos, J.J., 1998. Epidemiology, Genetics, and Ecology of Toxigenic *Vibrio cholerae*. Microbiol. Mol. Biol. Rev. 62, 1301–1314. https://doi.org/10.1128/MMBR.62.4.1301-1314.1998

Fernando, C.H., 1993. Rice field ecology and fish culture — an overview. Hydrobiologia, 259, 91–113. https://doi.org/10.1007/BF00008375

Föllmi, K., 1996. The phosphorus cycle, phosphogenesis and marine phosphate-rich deposits. Earth-Science Rev. 40, 55–124. https://doi.org/10.1016/0012-8252(95)00049-6

Forster, M.J., 2002. Molecular modelling in structural biology. Micron 33, 365–384. https://doi.org/10.1016/S0968-4328(01)00035-X

Fredrikson, M., Andlid, T., Haikara, A., Sandberg, A.S., 2002. Phytate degradation by micro-organisms in synthetic media and pea flour. J. Appl. Microbiol. 93, 197–204. https://doi.org/10.1046/j.1365-2672.2002.01676.x

Freed, C., Adepoju, O., Gillaspy, G., 2020. Can Inositol Pyrophosphates Inform Strategies for Developing Low Phytate Crops? Plants, 9(1), 115. https://doi.org/10.3390/plants9010115

Fu, D., Huang, H., Luo, H., Wang, Y., Yang, P., Meng, K., Bai, Y., Wu, N., Yao, B., 2008. A highly pH-stable phytase from Yersinia kristeensenii: Cloning, expression, and characterization. Enzyme Microb. Technol. 42, 499–505. https://doi.org/10.1016/j.enzmictec.2008.01.014

Fu, L., Niu, B., Zhu, Z., Wu, S., Li, W., 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics, 28, 3150–3152. https://doi.org/10.1093/bioinformatics/bts565

Fu, S., Sun, J., Qian, L., Li, Z., 2008. Bacillus Phytases: Present Scenario and Future Perspectives. Appl. Biochem. Biotechnol. 151, 1–8. https://doi.org/10.1007/s12010-008-8158-7

Fugthong, A., Boonyapakron, K., Sornlek, W., Tanapongpipat, S., Eurwilaichitr, L., Pootanakit, K., 2010. Biochemical characterization and in vitro digestibility assay of *Eupenicillium parvum* (BCC17694) phytase expressed in *Pichia pastoris*. Protein Expr. Purif. 70, 60–67. https://doi.org/10.1016/j.pep.2009.10.001

Gabriele Aquilina, Georges Bories, Andrew Chesson, Pier Sandro Cocconcelli, Joop de Knecht, N., Albert Dierick, Mikolaj Antoni Gralak, Jürgen Gropp, Ingrid Halle, Christer Hogstrand, Reinhard Kroker, L.L., Secundino López Puente, Anne-Katrine Lundebye Haldorsen, Alberto Mantovani, Giovanna Martelli, M.M., Derek Renshaw, Maria Saarela, K.S. and J.W., 2012. Scientific Opinion on the safety and efficacy of Phyzyme XP (6-phytase) as a feed additive for minor poultry species. EFSA J. 10(3), 2619. https://doi.org/10.2903/j.efsa.2012.2619

Gan, L., Long, X., Zhang, H., Hou, Y., Tian, J., Zhang, Y., Tian, Y., 2018. *Halomonas saliphila* sp. nov., a moderately halophilic bacterium isolated from a saline soil. Int. J. Syst. Evol. Microbiol. 68, 1153–1159. https://doi.org/10.1099/ijsem.0.002644

Gans, J., 2005. Computational Improvements Reveal Great Bacterial Diversity and High Metal Toxicity in Soil. Science, 309(5739), 1387–1390. https://doi.org/10.1126/science.1112665

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M.R., Appel, R.D., Bairoch, A., 2005. Protein Identification and Analysis Tools on the ExPASy Server, in: The Proteomics Protocols Handbook. Humana Press, Totowa, NJ, 571–607. https://doi.org/10.1385/1-59259-890-0:571

Gaunt, M.W., Turner, S.L., Rigottier-Gois, L., Lloyd-Macgilp, S.A., Young, J.P., 2001. Phylogenies of atpD and recA support the small subunit rRNA-based classification of rhizobia. Int. J. Syst. Evol. Microbiol. 51, 2037–2048. https://doi.org/10.1099/00207713-51-6-2037

George, T.S., Quiquampoix, H., Simpson, R.J., Richardson, A.E., 2005. Interactions between phytase and soil constituents: implications for hydrolysis of inositol phosphates. Inositol Phosphates Soil–Plant–Animal Syst. Link. Agric. Environ. 47–48. https://doi.org/10.1079/9781845931520.0221

Gerke, J., 2015. Phytate (Inositol Hexakisphosphate) in Soil and Phosphate Acquisition from Inositol Phosphates by Higher Plants. A Review. Plants, 4(2), 253–266. https://doi.org/10.3390/plants4020253

Ghio, S., Lorenzo, G.S. Di, Lia, V., Talia, P., Cataldi, A., Grasso, D., Campos, E., 2012. Isolation of Paenibacillus sp. and Variovorax sp. strains from decaying woods and characterization of their potential for cellulose deconstruction. Int. J. Biochem. Mol. Biol. 3, 352–64.

Ghorbani Nasrabadi, R., Greiner, R., Yamchi, A., Nourzadeh Roshan, E., 2018. A novel purple acid phytase from an earthworm cast bacterium. J. Sci. Food Agric. 98, 3667–3674. https://doi.org/10.1002/jsfa.8845

Giblin, A.E., Wieder, R.K., 1992. Sulphur cycling in marine and freshwater wetlands. Sulphur Cycl. Cont. Wiley, New York 85–117.

Giles, C., Cade-Menun, B., Hill, J., 2011. The inositol phosphates in soils and manures: Abundance, cycling, and measurement. Can. J. Soil Sci. 91, 397–416. https://doi.org/10.4141/cjss09090

Gillaspy, G.E., 2013. The role of phosphoinositides and inositol phosphates in plant cell signaling. Adv Exp Med Biol, 991, 141–157. https://doi.org/10.1007/978-94-007-6331-9_8

Gobbetti, M., Rizzello, C.G., 2014. *Arthrobacter*, in: Encyclopedia of Food Microbiology. Elsevier, 69–76. https://doi.org/10.1016/B978-0-12-384730-0.00009-4

Goldstein, A.H., 1994. Involvement of the Quinoprotein Glucose Dehydrogenase in the Solubilization of Exogenous Phosphates by Gram-Negative Bacteria, ASM Press, Washington DC, 197–203.

Goldstein, B.P., 2014. Resistance to rifampicin: a review. J. Antibiot. 67(9), 625–630. https://doi.org/10.1038/ja.2014.107

Grady, E.N., MacDonald, J., Liu, L., Richman, A., Yuan, Z.-C., 2016. Current knowledge and perspectives of *Paenibacillus*: a review. Microb. Cell Fact. 15(1), 1-18. https://doi.org/10.1186/s12934-016-0603-7

Greiner, R., Carlsson, N.-G., Alminger, M.L., 2000. Stereospecificity of myo-inositol hexakisphosphate dephosphorylation by a phytate-degrading enzyme of *Escherichia coli*. J. Biotechnol. 84, 53–62. https://doi.org/10.1016/S0168-1656(00)00331-X

Greiner, R., Haller, E., Konietzny, U., Jany, K.-D., 1997. Purification and Characterization of a Phytase from *Klebsiella terrigena*. Arch. Biochem. Biophys. 341, 201–206. https://doi.org/10.1006/abbi.1997.9942

Greiner, R., Konietzny, U., Jany, K.D., 1993. Purification and Characterization of Two Phytases from *Escherichia coli.* Arch. Biochem. Biophys. 303, 107–113. https://doi.org/10.1006/abbi.1993.1261

Greiner, R., Silva, L.G. da, Couri, S., 2009. Purification and characterisation of an extracellular phytase from *Aspergillus niger* 11T53A9. Brazilian J. Microbiol. 40, 795–807. https://doi.org/10.1590/S1517-

83822009000400010

Greppi, A., Krych, Ł., Costantini, A., Rantsiou, K., Hounhouigan, D.J., Arneborg, N., Cocolin, L., Jespersen, L., 2015. Phytase-producing capacity of yeasts isolated from traditional African fermented food products and PHYPk gene expression of *Pichia kudriavzevii* strains. Int. J. Food Microbiol. 205, 81–89. https://doi.org/10.1016/j.ijfoodmicro.2015.04.011

Gruninger, R.J., Thibault, J., Capeness, M.J., Till, R., Mosimann, S.C., Sockett, R.E., Selinger, B.L., Lovering, A.L., 2014. Structural and biochemical analysis of a unique phosphatase from *Bdellovibrio bacteriovorus* reveals its structural and functional relationship with the protein tyrosine phosphatase class of phytase. PLoS One 9, e94403. https://doi.org/10.1371/journal.pone.0094403

Gu, W., Huang, H., Meng, K., Yang, P., Fu, D., Luo, H., Wang, Y., Yao, B., Zhan, Z., 2009. Gene Cloning, Expression, and Characterization of a Novel Phytase from Dickeya paradisiaca. Appl. Biochem. Biotechnol. 157, 113–123. https://doi.org/10.1007/s12010-008-8329-6

Guerrand, D., 2018. Economics of food and feed enzymes: Status and prospectives. Status and prospectives., in: Enzymes in Human and Animal Nutrition: Principles and Perspectives. Elsevier Inc., 487–514. https://doi.org/10.1016/B978-0-12-805419-2.00026-5

Gulati, A., Sharma, N., Vyas, P., Sood, S., Rahi, P., Pathania, V., Prasad, R., 2010. Organic acid production and plant growth promotion as a function of phosphate solubilization by *Acinetobacter* rhizosphaerae strain BIHB 723 isolated from the cold deserts of the trans-Himalayas. Arch. Microbiol. 192, 975–983. https://doi.org/10.1007/s00203-010-0615-3

Gulati, H.K., Chadha, B.S., Saini, H.S., 2006. Production and characterization of thermostable alkaline phytase from *Bacillus laevolacticus* isolated from rhizosphere soil. J. Ind. Microbiol. Biotechnol. 34, 91–98. https://doi.org/10.1007/s10295-006-0171-7

Gupta, R., Brunak, S., 2002. Prediction of glycosylation across the human proteome and the correlation to protein function. Pacific Symp. Biocomput. 310–22.

Gupta, R.K., Gangoliya, S.S., Singh, N.K., 2015. Reduction of phytic acid and enhancement of bioavailable micronutrients in food grains. J. Food Sci. Technol. 52, 676–684. https://doi.org/10.1007/s13197-013-0978-y

Guzmán, D., Balderrama-Subieta, A., Cardona-Ortuño, C., Guevara-Martínez, M., Callisaya-Quispe, N., Quillaguamán, J., 2012. Evolutionary patterns of carbohydrate transport and metabolism in *Halomonas boliviensis* as derived from its genome sequence: influences on polyester production. Aquat. Biosyst. 8(1), 1-12. https://doi.org/10.1186/2046-9063-8-9

Haefner, S., Knietsch, A., Scholten, E., Braun, J., Lohscheidt, M., Zelder, O., 2005. Biotechnological production and applications of phytases. Appl. Microbiol. Biotechnol, 68(5), 588-597. https://doi.org/10.1007/s00253-005-0005-y

Hallegraeff, G.M., 1992. Harmful algal blooms in the Australian region. Mar. Pollut. Bull. 25, 186–190. https://doi.org/10.1016/0025-326X(92)90223-S

Hamada, M., Otoguro, M., Yamamura, H., Tamura, T., Suzuki, K., Hayakawa, M., 2010. Luteimicrobium subarcticum gen. nov., sp. nov., a new member of the suborder Micrococcineae. Int. J. Syst. Evol. Microbiol. 60, 796–800. https://doi.org/10.1099/ijs.0.014597-0

Hanson, C.J., Bootman, M.D., Roderick, H.L., 2004. Cell Signalling: IP3 Receptors Channel Calcium into Cell Death. Curr. Biol. 14, R933–R935. https://doi.org/10.1016/j.cub.2004.10.019

Harden, A., Thompson, J., Young, W.J., 1911. Apparatus for Collecting and Measuring the Gases evolved during Fermentation. Biochem. J., 5, 230–235. https://doi.org/10.1042/bj0050230

Harden, A., Young, W.J., 1913. The Enzymatic Formation of Polysaccharides by Yeast Preparations. Biochem. J. 7, 630–636. https://doi.org/10.1042/bj0070630

Hariprasad, P., Niranjana, S.R., 2009. Isolation and characterization of phosphate solubilizing rhizobacteria to improve plant health of tomato. Plant Soil, 316, 13–24. https://doi.org/10.1007/s11104-008-9754-6

Haros, M., Carlsson, N.-G., Almgren, A., Larsson-Alminger, M., Sandberg, A.-S., Andlid, T., 2009. Phytate degradation by human gut isolated *Bifidobacterium pseudocatenulatum* ATCC27919 and its probiotic potential. Int. J. Food Microbiol. 135, 7–14. https://doi.org/10.1016/j.ijfoodmicro.2009.07.015

Hegeman, C.E., Grabau, E.A., 2001. A Novel Phytase with Sequence Similarity to Purple Acid Phosphatases Is Expressed in Cotyledons of Germinating Soybean Seedlings. Plant Physiol. 126, 1598–1608. https://doi.org/10.1104/pp.126.4.1598

Hélène Boze, André Aumelas, G.M., 2011. *Debaryomyces castellii* phytase. US8003360B2.

Henderson, G., Cox, F., Ganesh, S., Jonker, A., Young, W., Janssen, P.H., 2015. Rumen microbial community composition varies with diet and host, but a core microbiome is found across a wide geographical range. Sci. Rep. 5(1), 1-15. https://doi.org/10.1038/srep14567

Hendrickson, H. Stewart, and C.E.B., 1964. Ion exchange chromatography of intact brain phosphoinositides on diethylaminoethyl cellulose by gradient salt elution in a mixed solvent system. J. biol. Chem, 239, 1369–1373.

Herrmann, K.R., Ruff, A.J., Infanzón, B., Schwaneberg, U., 2019. Engineered phytases for emerging biotechnological applications beyond animal feeding. Appl. Microbiol. Biotechnol., 103(16), 6435-6448. https://doi.org/10.1007/s00253-019-09962-1

Hesampour, A., Siadat, S.E.R., Malboobi, M.A., Mohandesi, N., Arab, S.S., Ghahremanpour, M.M., 2015. Enhancement of Thermostability and Kinetic Efficiency of Aspergillus niger PhyA Phytase by Site-Directed Mutagenesis. Appl. Biochem. Biotechnol. 175, 2528–2541. https://doi.org/10.1007/s12010-014-1440-y

Hibbing, M.E., Fuqua, C., Parsek, M.R., Peterson, S.B., 2010. Bacterial competition: surviving and thriving in the microbial jungle. Nat. Rev. Microbiol. 8, 15–25. https://doi.org/10.1038/nrmicro2259

Hill, J.E., Kysela, D., Elimelech, M., 2007. Isolation and assessment of phytate-hydrolysing bacteria from the DelMarVa Peninsula. Environ. Microbiol. 9, 3100–3107. https://doi.org/10.1111/j.1462-2920.2007.01420.x

Hirsch, P.R., Mauchline, T.H., Clark, I.M., 2010. Culture-independent molecular techniques for soil microbial ecology. Soil Biol. Biochem. 42, 878–887. https://doi.org/10.1016/j.soilbio.2010.02.019

Hosseinkhani, Baharak, and G.H., 2009. Analysis of phytase producing bacteria (*Pseudomonas* sp.) from poultry faeces and optimization of this enzyme production. African J. Biotechnol. 8(17).

Howarth, R.W., 1988. Nutrient Limitation of Net Primary Production in Marine Ecosystems. Annu. Rev. Ecol. Syst. 19, 89–110. https://doi.org/10.1146/annurev.es.19.110188.000513

Howe, A., Chain, P.S.G., 2015. Challenges and opportunities in understanding microbial communities with metagenome assembly (accompanied by IPython Notebook tutorial). Front. Microbiol., 6, 678. https://doi.org/10.3389/fmicb.2015.00678

Howson, S.J., Davis, R.P., 1983. Production of phytate-hydrolysing enzyme by some fungi. Enzyme Microb. Technol. 5, 377–382. https://doi.org/10.1016/0141-0229(83)90012-1

Huang, H., 2009. Novel Low-Temperature-Active Phytase from *Erwinia carotovora* var. *carotovota* ACCC 10276. J. Microbiol. Biotechnol. 19(10), 1085-1091. https://doi.org/10.4014/jmb.0901.039

Huang, H., Luo, H., Yang, P., Meng, K., Wang, Y., Yuan, T., Bai, Y., Yao, B., 2006. A novel phytase with preferable characteristics from *Yersinia intermedia*. Biochem. Biophys. Res. Commun. 350, 884–889. https://doi.org/10.1016/j.bbrc.2006.09.118

Huang, H., Shao, N., Wang, Y., Luo, H., Yang, P., Zhou, Z., Zhan, Z., Yao, B., 2009a. A novel beta-propeller phytase from P*edobacter nyackensis* MJ11 CGMCC 2503 with potential as an aquatic feed additive. Appl. Microbiol. Biotechnol. 83, 249–259. https://doi.org/10.1007/s00253-008-1835-1

Huang, H., Shi, P., Wang, Y., Luo, H., Shao, N., Wang, G., Yang, P., Yao, B., 2009b. Diversity of Beta-Propeller Phytase Genes in the Intestinal Contents of Grass Carp Provides Insight into the Release of Major Phosphorus from Phytate in Nature. Appl. Environ. Microbiol. 75, 1508–1516. https://doi.org/10.1128/AEM.02188-08

Huang, H., Zhang, R., Fu, D., Luo, J., Li, Z., Luo, H., Shi, P., Yang, P., Diao, Q., Yao, B., 2011. Diversity, abundance and characterization of ruminal cysteine phytases suggest their important role in phytate degradation. Environ. Microbiol. 13, 747–757. https://doi.org/10.1111/j.1462-2920.2010.02379.x

Huang, Y., Niu, B., Gao, Y., Fu, L., Li, W., 2010. CD-HIT Suite: a web server for clustering and comparing biological sequences. Bioinformatics 26, 680–682. https://doi.org/10.1093/bioinformatics/btq003

Huddleston, J.R., 2014. Horizontal gene transfer in the human gastrointestinal tract: potential spread of antibiotic resistance genes. Infect. Drug Resist. 7, 167. https://doi.org/10.2147/IDR.S48820

Humer, E., Zebeli, Q., 2015. Phytate in feed ingredients and potentials for improving the utilization of phosphorus in ruminant nutrition. Anim. Feed Sci. Technol. 209, 1–15. https://doi.org/10.1016/j.anifeedsci.2015.07.028

Huo, Y.-Y., Li, Z.-Y., Cheng, H., Wang, C.-S., Xu, X.-W., 2014. High quality draft genome sequence of the heavy metal resistant bacterium Halomonas zincidurans type strain B6T. Stand. Genomic Sci. 9, 30. https://doi.org/10.1186/1944-3277-9-30

Hutchings, A.M., Antler, G., Wilkening, J. V., Basu, A., Bradbury, H.J., Clegg, J.A., Gorka, M., Lin, C.Y., Mills, J. V., Pellerin, A., Redeker, K.R., Sun, X., Turchyn, A. V., 2019. Creek Dynamics Determine Pond Subsurface

Geochemical Heterogeneity in East Anglian (UK) Salt Marshes. Front. Earth Sci., 7, 41. https://doi.org/10.3389/feart.2019.00041

Imelda, Joseph, and R.P., 2007. Isolation and characterization of phytase producing *Bacillus* strains from mangrove ecosystem. Journal Mar. Biol. Assoc. India, 49(2), 177–182.

Irving, G.C.J., Cosgrove, D.J., 1971. Inositol phosphate phosphatases of microbiological origin. some properties of a partially purified bacterial (*Pseudomonas* sp.) phytase. Aust. J. Biol. Sci. https://doi.org/10.1071/BI9710547

Islam, M., Kim, S.-H., Son, A.-R., Ramos, S.C., Jeong, C.-D., Yu, Z., Kang, S.H., Cho, Y.-I., Lee, Sung-Sill, Cho, K.-K., Lee, Sang-Suk, 2021. Seasonal Influence on Rumen Microbiota, Rumen Fermentation, and Enteric Methane Emissions of Holstein and Jersey Steers under the Same Total Mixed Ration. Animals, 11(4), 1184. https://doi.org/10.3390/ani11041184

Islam, M.T., Deora, A., Hashidoko, Y., Rahman, A., Ito, T., Tahara, S., 2007. Isolation and Identification of Potential Phosphate Solubilizing Bacteria from the Rhizoplane of *Oryza sativa* L. cv. BR29 of Bangladesh. Zeitschrift für Naturforsch. C 62, 103–110. https://doi.org/10.1515/znc-2007-1-218

Ivanova, E.P., 2005. *Marinomonas pontica* sp. nov., isolated from the Black Sea. Int. J. Syst. Evol. Microbiol. 55, 275–279. https://doi.org/10.1099/ijs.0.63326-0

Ivell, D.M., 2012. Phosphate Fertilizer Production – From the 1830's to 2011 and Beyond. Procedia Eng. 46, 166–171. https://doi.org/10.1016/j.proeng.2012.09.461

Iyer, B., Rajput, M.S., Rajkumar, S., 2017. Effect of succinate on phosphate solubilization in nitrogen fixing bacteria harbouring chick pea and their effect on plant growth. Microbiol. Res. 202, 43–50. https://doi.org/10.1016/j.micres.2017.05.005

Jahnke, R.A., 1992. The Phosphorus Cycle. Int. Geophys., 50, 301-315. https://doi.org/10.1016/S0074-6142(08)62697-2

Jain, J., Sapna, Singh, B., 2016. Characteristics and biotechnological applications of bacterial phytases. Process Biochem. 51, 159–169. https://doi.org/10.1016/j.procbio.2015.12.004

Janda, J.M., Abbott, S.L., 2007. 16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls. J. Clin. Microbiol. 45, 2761–2764. https://doi.org/10.1128/JCM.01228-07

Janda, J.M., Abbott, S.L., 2002. Bacterial Identification for Publication: When Is Enough Enough? J. Clin. Microbiol., 40, 1887–1891. https://doi.org/10.1128/JCM.40.6.1887-1891.2002

Jane E. Hill and Alan E. Richardson, 2007. Inositol phosphates: linking agriculture and the environment. CABI, Wallingford. https://doi.org/10.1079/9781845931520.0000

Jarrett, J.P., Wilson, J.W., Ray, P.P., Knowlton, K.F., 2014. The effects of forage particle length and exogenous phytase inclusion on phosphorus digestion and absorption in lactating cows. J. Dairy Sci. 97, 411–418. https://doi.org/10.3168/jds.2013-7124

Jerke, K., Nakatsu, C.H., Beasley, F., Konopka, A., 2008. Comparative analysis of eight *Arthrobacter* plasmids. Plasmid 59, 73–85. https://doi.org/10.1016/j.plasmid.2007.12.003

Jiao, Y., Whalen, J.K., Hendershot, W.H., 2007. Phosphate Sorption and Release in a Sandy-Loam Soil as Influenced by Fertilizer Sources. Soil Sci. Soc. Am. J. 71, 118–124. https://doi.org/10.2136/sssaj2006.0028

Joern, J.M., 2003. DNA Shuffling, in: Directed Evolution Library Creation. Humana Press, New Jersey, 85–90. https://doi.org/10.1385/1-59259-395-X:85

Johnson, B.B., Quill, E., Angove, M.J., 2012. An investigation of the mode of sorption of inositol hexaphosphate to goethite. J. Colloid Interface Sci. 367, 436–442. https://doi.org/10.1016/j.jcis.2011.09.066

Jones, F.T., 2011. A review of practical Salmonella control measures in animal feed. J. Appl. Poult. Res. 20, 102–113. https://doi.org/10.3382/japr.2010-00281

Jones, F.T., Richardson, K.E., 2004. Salmonella in Commercially Manufactured Feeds. Poult. Sci. 83, 384–391. https://doi.org/10.1093/ps/83.3.384

Jorquera, M A, Crowley, D.E., Marschner, P., Greiner, R., Fernandez, M.T., Romero, D., Menezes-Blackburn, D., De La Luz Mora, M., 2011. Identification of beta-propeller phytase-encoding genes in culturable *Paenibacillus* and *Bacillus* spp. from the rhizosphere of pasture plants on volcanic soils. FEMS Microbiol Ecol 75, 163–172. https://doi.org/10.1111/j.1574-6941.2010.00995.x

Jorquera, M.A., Hernández, M.T., Rengel, Z., Marschner, P., De La Luz Mora, M., 2008. Isolation of culturable phosphobacteria with both phytate-mineralization and phosphate-solubilization activity from the rhizosphere of plants grown in a volcanic soil. Biol. Fertil. Soils 44, 1025–1034. https://doi.org/10.1007/s00374-008-0288-0

Joseph Felsenstein, 2004. Inferring phylogenies, Sinauer Associates, Sunderland, MA., 2, 664.

Kanehisa, M., 2019. Toward understanding the origin and evolution of cellular organisms. Protein Sci. 28, 1947–1951. https://doi.org/10.1002/pro.3715

Kanehisa, M., 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res. 28, 27–30. https://doi.org/10.1093/nar/28.1.27

Kang, I., Jang, H., Oh, H.-M., Cho, J.-C., 2012. Complete Genome Sequence of *Marinomonas* Bacteriophage P12026. J. Virol. 86, 8909–8910. https://doi.org/10.1128/JVI.01328-12

Katoh, K., Rozewicki, J., Yamada, K.D., 2019. MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. Brief. Bioinform. 20, 1160–1166. https://doi.org/10.1093/bib/bbx108

Kebreab, E., Odongo, N.E., McBride, B.W., Hanigan, M.D., France, J., 2008. Phosphorus Utilization and Environmental and Economic Implications of Reducing Phosphorus Pollution from Ontario Dairy Cows. J. Dairy Sci. 91, 241–246. https://doi.org/10.3168/jds.2007-0432

Kerovuo, J., Lauraeus, M., Nurminen, P., Kalkkinen, N., Apajalahti, J., 1998. Isolation, characterization, molecular gene cloning, and sequencing of a novel phytase from *Bacillus subtilis*. Appl. Environ. Microbiol. 64, 2079–2085.

Khalifa, A., Lee, C.G., Ogiso, T., Ueno, C., Dianou, D., Demachi, T., Katayama, A., Asakawa, S., 2015. *Methylomagnum ishizawai* gen. nov., sp. nov., a mesophilic type I methanotroph isolated from rice rhizosphere. Int. J. Syst. Evol. Microbiol. 65, 3527–3534. https://doi.org/10.1099/ijsem.0.000451

Khan, M.N., Mohammad, F., 2014. Eutrophication: Challenges and Solutions, in: Eutrophication: Causes, Consequences and Control. Springer Netherlands, Dordrecht, 1–15. https://doi.org/10.1007/978-94-007-7814-6_1

Kies, A.K., De Jonge, L.H., Kemme, P.A., Jongbloed, A.W., 2006. Interaction between Protein, Phytate, and Microbial Phytase. In Vitro Studies. J. Agric. Food Chem. 54, 1753–1758. https://doi.org/10.1021/jf0518554

Kim, Young-Hoon, Moon-Nam Gwon, Si-Yong Yang, Tae-Kyu Park, Chan-Gil Kim, Chang-Won Kim, and M.-D.S., 2002. Isolation of Phytase-Producing *Pseudomonas* sp. and Optimization of its Phytase Production. J. Microbiol. Biotechnol. 12, 279–285.

Kim, H.-W., Kim, Y.-O., Lee, J.-H., Kim, K.-K., Kim, Y.-J., 2003. Isolation and characterization of a phytase with improved properties from *Citrobacter braakii*. Biotechnol. Lett. 25, 1231–1234. https://doi.org/https://doi.org/10.1023/A:1025020309596

Kim, J.N., Méndez–García, C., Geier, R.R., Iakiviak, M., Chang, J., Cann, I., Mackie, R.I., 2017. Metabolic networks for nitrogen utilization in *Prevotella ruminicola*, 23. Sci. Rep. 7, 7851. https://doi.org/10.1038/s41598-017-08463-3

Kim, O.H., Kim, Y.O., Shim, J.H., Jung, Y.S., Jung, W.J., Choi, W.C., Lee, H., Lee, S.J., Kim, K.K., Auh, J.H., Kim, H., Kim, J.W., Oh, T.K., Oh, B.C., 2010. beta-propeller phytase hydrolyzes insoluble Ca(2+)-phytate salts and completely abrogates the ability of phytate to chelate metal ions. Biochemistry, 49, 10216–10227. https://doi.org/10.1021/bi1010249

Kim, Y.H., Stites, W.E., 2008. Effects of Excluded Volume upon Protein Stability in Covalently Cross-Linked Proteins with Variable Linker Lengths. Biochemistry, 47, 8804–8814. https://doi.org/10.1021/bi800297j

Komaba, H., Fukagawa, M., 2016. Phosphate—a poison for humans? Kidney Int. 90, 753–763. https://doi.org/10.1016/j.kint.2016.03.039

Konietzny, U., Greiner, R., 2004. Bacterial phytase: potential application, in vivo function and regulation of its synthesis. Brazilian J. Microbiol. 35, 11–18.

Konietzny, U., Greiner, R., 2002. Molecular and catalytic properties of phytate-degrading enzymes (phytases). Int. J. Food Sci. Technol. 37(7), 791-812. https://doi.org/10.1046/j.1365-2621.2002.00617.x

Körfer, G., Novoa, C., Kern, J., Balla, E., Grütering, C., Davari, M.D., Martinez, R., Vojcic, L., Schwaneberg, U., 2018. Directed evolution of an acid *Yersinia mollaretii* phytase for broadened activity at neutral pH. Appl. Microbiol. Biotechnol. 102, 9607–9620. https://doi.org/10.1007/s00253-018-9308-7

Kowalczykowski, S.C., Dixon, D.A., Eggleston, A.K., Lauder, S.D., Rehrauer, W.M., 1994. Biochemistry of homologous recombination in *Escherichia coli*. Microbiol. Rev. 58, 401–465. https://doi.org/10.1128/mr.58.3.401-465.1994

Kreuzer-Martin, H.W., 2007. Stable Isotope Probing: Linking Functional Activity to Specific Members of Microbial Communities. Soil Sci. Soc. Am. J. 71, 611–619. https://doi.org/10.2136/sssaj2006.0093

Kuang, R., Chan, K.-H., Yeung, E., Lim, B.L., 2009. Molecular and Biochemical Characterization of AtPAP15, a

Purple Acid Phosphatase with Phytase Activity, in *Arabidopsis*. Plant Physiol. 151, 199–209. https://doi.org/10.1104/pp.109.143180

Kumar, V., Singh, G., Verma, A.K., Agrawal, S., 2012. In Silico Characterization of Histidine Acid Phytase Sequences. Enzyme Res. 2012, 1–8. https://doi.org/10.1155/2012/845465

Kumar, V., Singh, P., Jorquera, M.A., Sangwan, P., Kumar, P., Verma, A.K., Agrawal, S., 2013. Isolation of phytase-producing bacteria from Himalayan soils and their effect on growth and phosphorus uptake of Indian mustard (*Brassica juncea*). World J. Microbiol. Biotechnol. 29, 1361–1369. https://doi.org/10.1007/s11274-013-1299-z

Kumar, V., Sinha, A.K., Makkar, H.P.S., Becker, K., 2010. Dietary roles of phytate and phytase in human nutrition: A review. Food Chem. 120(4), 945-959. https://doi.org/10.1016/j.foodchem.2009.11.052

Kumar, V., Yadav, A.N., Verma, P., Sangwan, P., Saxena, A., Kumar, K., Singh, B., 2017. β-Propeller phytases: Diversity, catalytic attributes, current developments and potential biotechnological applications. Int. J. Biol. Macromol. 98, 595-609. https://doi.org/10.1016/j.ijbiomac.2017.01.134

Lagkouvardos, I., Pukall, R., Abt, B., Foesel, B.U., Meier-Kolthoff, J.P., Kumar, N., Bresciani, A., Martínez, I., Just, S., Ziegler, C., Brugiroux, S., Garzetti, D., Wenning, M., Bui, T.P.N., Wang, J., Hugenholtz, F., Plugge, C.M., Peterson, D.A., Hornef, M.W., Baines, J.F., Smidt, H., Walter, J., Kristiansen, K., Nielsen, H.B., Haller, D., Overmann, J., Stecher, B., Clavel, T., 2016. The Mouse Intestinal Bacterial Collection (miBC) provides host-specific insight into cultured diversity and functional potential of the gut microbiota. Nat. Microbiol. 1, 16131. https://doi.org/10.1038/nmicrobiol.2016.131

Laha, D., Parvin, N., Hofer, A., Giehl, R.F.H., Fernandez-Rebollo, N., von Wiren, N., Saiardi, A., Jessen, H.J., Schaaf, G., 2019. *Arabidopsis* ITPK1 and ITPK2 Have an Evolutionarily Conserved Phytic Acid Kinase Activity. ACS Chem Biol, 14, 2127–2133. https://doi.org/10.1021/acschembio.9b00423

Larentis, A.L., Nicolau, J.F.M.Q., Esteves, G. dos S., Vareschini, D.T., de Almeida, F.V.R., dos Reis, M.G., Galler, R., Medeiros, M.A., 2014. Evaluation of pre-induction temperature, cell growth at induction and IPTG concentration on the expression of a leptospiral protein in *E. coli* using shaking flasks and microbioreactor. BMC Res. Notes, 7, 671. https://doi.org/10.1186/1756-0500-7-671

Lassen, S.F., Breinholt, J., Østergaard, P.R., Brugger, R., Bischoff, A., Wyss, M., Fuglsang, C.C., 2001. Expression, Gene Cloning, and Characterization of Five Novel Phytases from Four Basidiomycete Fungi: *Peniophora lycii*, *Agrocybe pediades* , a *Ceriporia* sp., and *Trametes pubescens*. Appl. Environ. Microbiol. 67, 4701–4707. https://doi.org/10.1128/AEM.67.10.4701-4707.2001

Lawrence, J.G., 1999. Gene transfer, speciation, and the evolution of bacterial genomes. Curr. Opin. Microbiol. 2, 519–523. https://doi.org/10.1016/S1369-5274(99)00010-7

Lawrence, J.G., Roth, J.R., 1996. Selfish Operons: Horizontal Transfer May Drive the Evolution of Gene Clusters. Genetics 143(4), 1843 - 1860.

Le Roux, F., Blokesch, M., 2018. Eco-evolutionary Dynamics Linked to Horizontal Gene Transfer in *Vibrios*. Annu. Rev. Microbiol. 72, 89–110. https://doi.org/10.1146/annurev-micro-090817-062148

Lee, S.A., Bedford, M.R., 2016. Inositol - An effective growth promotor? Worlds. Poult. Sci. J. 72, 743–760. https://doi.org/10.1017/S0043933916000660

Lee, S.A., Nagalakshmi, D., Raju, M., Rama Rao, S. V, Bedford, M.R., 2017. Effect of phytase superdosing, myo-inositol and available phosphorus concentrations on performance and bone mineralisation in broilers. Anim Nutr. 3, 247–251. https://doi.org/10.1016/j.aninu.2017.07.002

Legendre, P., Makarenkov, V., 2002. Reconstruction of Biogeographic and Evolutionary Networks Using Reticulograms. Syst. Biol. 51, 199–216. https://doi.org/10.1080/10635150252899725

Lehmann, M., Kostrewa, D., Wyss, M., Brugger, R., D'Arcy, A., Pasamontes, L., van Loon, A.P.G.M., 2000. From DNA sequence to improved functionality: using protein sequence comparisons to rapidly design a thermostable consensus phytase. Protein Eng. Des. Sel. 13, 49–57. https://doi.org/10.1093/protein/13.1.49

Lei, X.G., Porres, J.M., 2003. Phytase enzymology, applications, and biotechnology. Biotechnol. Lett, 25(21), 1787-1794. https://doi.org/10.1023/A:1026224101580

Lei, X.G., Porres, J.M., Mullaney, E.J., Brinch-Pedersen, H., 2007. Phytase: Source, Structure and Application, in: Industrial Enzymes. Springer Netherlands, Dordrecht, 505–529. https://doi.org/10.1007/1-4020-5377-0_29

Lei, X.G., Weaver, J.D., Mullaney, E., Ullah, A.H., Azain, M.J., 2013. Phytase, a New Life for an "Old" Enzyme. Annu. Rev. Anim. Biosci. 1, 283–309. https://doi.org/10.1146/annurev-animal-031412-103717

Leibly, D.J., Nguyen, T.N., Kao, L.T., Hewitt, S.N., Barrett, L.K., Van Voorhis, W.C., 2012. Stabilizing Additives

Added during Cell Lysis Aid in the Solubilization of Recombinant Proteins. PLoS One, 7, e52482. https://doi.org/10.1371/journal.pone.0052482

Leis, B., Angelov, A., Liebl, W., 2013. Screening and Expression of Genes from Metagenomes. Advances in applied microbiology, 83, 1–68. https://doi.org/10.1016/B978-0-12-407678-5.00001-5

Lerminiaux, N.A., Cameron, A.D.S., 2019. Horizontal transfer of antibiotic resistance genes in clinical environments. Can. J. Microbiol. 65, 34–44. https://doi.org/10.1139/cjm-2018-0275

Letunic, I., Bork, P., 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res. 47, W256–W259. https://doi.org/10.1093/nar/gkz239

Ley, R.E., 2016. Prevotella in the gut: choose carefully. Nat. Rev. Gastroenterol. Hepatol. 13, 69–70. https://doi.org/10.1038/nrgastro.2016.4

Leytem, A.B., Maguire, R.O., 2005. Environmental implications of inositol phosphates in animal manures., in: Inositol Phosphates: Linking Agriculture and the Environment. CABI, Wallingford, 150–168. https://doi.org/10.1079/9781845931520.0150

Li, Ruijuan, Wenjing Lu, Juntao Gu, Xiaojuan Li, Chengjin Guo, and K.X., 2011. Molecular characterization and functional analysis of OsPHY2, a phytase gene classified in histidine acid phosphatase type in rice (*Oryza sativa* L.). African J. Biotechnol. 10, 11110–11123.

Li, C., Lin, Y., Zheng, X., Pang, N., Liao, X., Liu, X., Huang, Y., Liang, S., 2015. Combined strategies for improving expression of Citrobacter amalonaticus phytase in *Pichia pastoris*. BMC Biotechnol. 15, 88. https://doi.org/10.1186/s12896-015-0204-2

Li, H., Huang, G., Meng, Q., Ma, L., Yuan, L., Wang, F., Zhang, W., Cui, Z., Shen, J., Chen, X., Jiang, R., Zhang, F., 2011. Integrated soil and plant phosphorus management for crop and environment in China. A review. Plant Soil, 349, 157–167. https://doi.org/10.1007/s11104-011-0909-5

Li, Z., Wang, L., Zhong, Y., 2005. Detecting horizontal gene transfer with T-REX and RHOM programs. Brief. Bioinform. 6, 394–401. https://doi.org/10.1093/bib/6.4.394

Lim, B.L., Yeung, P., Cheng, C., Hill, J.E., 2007a. Distribution and diversity of phytate-mineralizing bacteria. ISME J. 1, 321–330. https://doi.org/10.1038/ismej.2007.40

Lim, D., Golovan, S., Forsberg, C.W., Jia, Z., 2000. Crystal structures of *Escherichia coli* phytase and its complex with phytate. Nat. Struct. Biol. 7, 108–113. https://doi.org/10.1038/72371

Lin, X., Tfaily, M.M., Steinweg, J.M., Chanton, P., Esson, K., Yang, Z.K., Chanton, J.P., Cooper, W., Schadt, C.W., Kostka, J.E., 2014. Microbial Community Stratification Linked to Utilization of Carbohydrates and Phosphorus Limitation in a Boreal Peatland at Marcell Experimental Forest, Minnesota, USA. Appl. Environ. Microbiol. 80, 3518–3530. https://doi.org/10.1128/AEM.00205-14

Liu, P., Cai, Z., Chen, Z., Mo, X., Ding, X., Liang, C., Liu, G., Tian, J., 2018. A root-associated purple acid phosphatase, SgPAP23, mediates extracellular phytate-P utilization in *Stylosanthes guianensis*. Plant. Cell Environ. 41, 2821–2834. https://doi.org/10.1111/pce.13412

Liu, Q., Huang, Q., Lei, X.G., Hao, Q., 2004. Crystallographic Snapshots of *Aspergillus fumigatus* Phytase, Revealing Its Enzymatic Dynamics. Structure 12, 1575–1583. https://doi.org/10.1016/j.str.2004.06.015

Livermore, T.M., Azevedo, C., Kolozsvari, B., Wilson, M.S.C., Saiardi, A., 2016. Phosphate, inositol and polyphosphates. Biochem. Soc. Trans. 44, 253–259. https://doi.org/10.1042/BST20150215

Lobstein, J., Emrich, C.A., Jeans, C., Faulkner, M., Riggs, P., Berkmen, M., 2012. SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microb. Cell Fact. 11, 753. https://doi.org/10.1186/1475-2859-11-56

Lorenz, A.J., Scott, M.P., Lamkey, K.R., 2008. Genetic Variation and Breeding Potential of Phytate and Inorganic Phosphorus in a Maize Population. Crop Sci. 48, 79–84. https://doi.org/10.2135/cropsci2007.03.0136

Lott, J.N.A., Ockenden, I., Raboy, V., Batten, G.D., 2000. Phytic acid and phosphorus in crop seeds and fruits: a global estimate. Seed Sci. Res. 10, 11–33. https://doi.org/10.1017/S0960258500000039

Lung, S.-C., Leung, A., Kuang, R., Wang, Y., Leung, P., Lim, B.-L., 2008. Phytase activity in tobacco (*Nicotiana tabacum*) root exudates is exhibited by a purple acid phosphatase. Phytochemistry, 69, 365–373. https://doi.org/10.1016/j.phytochem.2007.06.036

Ma, L.Q., Rao, G.N., 1997. Effects of Phosphate Rock on Sequential Chemical Extraction of Lead in Contaminated Soils. J. Environ. Qual. 26, 788–794. https://doi.org/10.2134/jeq1997.00472425002600030028x

Ma, Y., Lee, C.-J., Park, J.-S., 2020. Strategies for Optimizing the Production of Proteins and Peptides with

Multiple Disulfide Bonds. Antibiotics, 9, 541. https://doi.org/10.3390/antibiotics9090541

Mabelebele, M., Alabi, O.J., Ng`ambi, J.W., Norris, D., Ginindza, M.M., 2013. Comparison of Gastrointestinal Tracts and pH Values of Digestive Organs of Ross 308 Broiler and Indigenous Venda Chickens Fed the Same Diet. Asian J. Anim. Vet. Adv. 9, 71–76. https://doi.org/10.3923/ajava.2014.71.76

Maciá, E., 2005. The role of phosphorus in chemical evolution. Chem. Soc. Rev., 34(8), 691-701. https://doi.org/10.1039/b416855k

Madsen, C.K., Brearley, C.A., Brinch-Pedersen, H., 2019. Lab-scale preparation and QC of phytase assay substrate from rice bran. Anal. Biochem. 578, 7–12. https://doi.org/10.1016/j.ab.2019.04.021

Maenz, D., Classen, H., 1998. Phytase activity in the small intestinal brush border membrane of the chicken. Poult. Sci. 77, 557–563. https://doi.org/10.1093/ps/77.4.557

Maenz, D.D., Engele-Schaan, C.M., Newkirk, R.W., Classen, H.L., 1999. The effect of minerals and mineral chelators on the formation of phytase-resistant and phytase-susceptible forms of phytic acid in solution and in a slurry of canola meal. Anim. Feed Sci. Technol. 81, 177–192. https://doi.org/10.1016/S0377-8401(99)00085-1

Magot, M., Fardeau, M.-L., Arnauld, O., Lanau, C., Ollivier, B., Thomas, P., Patel, B.K.C., 2006. *Spirochaeta smaragdinae* sp. nov., a new mesophilic strictly anaerobic spirochete from an oil field. FEMS Microbiol. Lett. 155, 185–191. https://doi.org/10.1111/j.1574-6968.1997.tb13876.x

Mailhe, M., Ricaboni, D., Benezech, A., Khelaifia, S., Fournier, P.-E., Raoult, D., 2016. '*Mediterranea massiliensis* ' gen. nov., sp. nov., a new human-associated bacterium isolated from the right and left colon lavage of a 58-year-old patient. New Microbes New Infect. 13, 54–55. https://doi.org/10.1016/j.nmni.2016.06.002

Makarenkov, V., 2001. T-REX: reconstructing and visualizing phylogenetic trees and reticulation networks. Bioinformatics, 17, 664–668. https://doi.org/10.1093/bioinformatics/17.7.664

Mallona, I., Weiss, J., Egea-Cortines, M., 2011. pcrEfficiency: a Web tool for PCR amplification efficiency prediction. BMC Bioinformatics, 12, 404. https://doi.org/10.1186/1471-2105-12-404

Mallory, L.M., Austin, B., Colwell, R.R., 1977. Numerical taxonomy and ecology of oligotrophic bacteria isolated from the estuarine environment. Can. J. Microbiol. 23, 733–750. https://doi.org/10.1139/m77-110

Mamedov, A.I., Bar-Yosef, B., Levkovich, I., Rosenberg, R., Silber, A., Fine, P., Levy, G.J., 2016. Amending Soil with Sludge, Manure, Humic Acid, Orthophosphate and Phytic Acid: Effects on Infiltration, Runoff and Sediment Loss. L. Degrad. Dev. 27, 1629–1639. https://doi.org/10.1002/ldr.2474

Masoumi, F., Khadivinia, E., Alidoust, L., Mansourinejad, Z., Shahryari, S., Safaei, M., Mousavi, A., Salmanian, A.-H., Zahiri, H.S., Vali, H., Noghabi, K.A., 2016. Nickel and lead biosorption by *Curtobacterium* sp. FM01, an indigenous bacterium isolated from farmland soils of northeast Iran. J. Environ. Chem. Eng. 4, 950–957. https://doi.org/10.1016/j.jece.2015.12.025

Matange, N., Podobnik, M., Visweswariah, S.S., 2015. Metallophosphoesterases: structural fidelity with functional promiscuity. Biochem. J. 467, 201–216. https://doi.org/10.1042/BJ20150028

Matsen, F.A., Kodner, R.B., Armbrust, E.V., 2010. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. BMC Bioinformatics, 11, 538. https://doi.org/10.1186/1471-2105-11-538

Matte-Tailliez, O., Brochier, C., Forterre, P., Philippe, H., 2002. Archaeal Phylogeny Based on Ribosomal Proteins. Mol. Biol. Evol. 19, 631–639. https://doi.org/10.1093/oxfordjournals.molbev.a004122

Mavromatis, K., Yasawong, M., Chertkov, O., Lapidus, A., Lucas, S., Nolan, M., Del Rio, T.G., Tice, H., Cheng, J.-F., Pitluck, S., Liolios, K., Ivanova, N., Tapia, R., Han, C., Bruce, D., Goodwin, L., Pati, A., Chen, A., Palaniappan, K., Land, M., Hauser, L., Chang, Y.-J., Jeffries, C.D., Detter, J.C., Rohde, M., Brambilla, E., Spring, S., Göker, M., Sikorski, J., Woyke, T., Bristow, J., Eisen, J.A., Markowitz, V., Hugenholtz, P., Klenk, H.-P., Kyrpides, N.C., 2010. Complete genome sequence of *Spirochaeta smaragdinae* type strain (SEBR 4228T). Stand. Genomic Sci. 3, 1–9. https://doi.org/10.4056/sigs.1143106

McGrath, K.C., Mondav, R., Sintrajaya, R., Slattery, B., Schmidt, S., Schenk, P.M., 2010. Development of an Environmental Functional Gene Microarray for Soil Microbial Communities. Appl. Environ. Microbiol. 76, 7161–7170. https://doi.org/10.1128/AEM.03108-09

McLaughlin, J.R., Ryden, J.C., Syers, J.K., 1981. Sorption of inorganic phosphate by Iron- and Aluminium-containing components. J. Soil Sci. 32, 365–378. https://doi.org/10.1111/j.1365-2389.1981.tb01712.x

McLean, T.C., Wilkinson, B., Hutchings, M.I., Devine, R., 2019. Dissolution of the Disparate: Co-ordinate Regulation in Antibiotic Biosynthesis. Antibiotics, 8, 83. https://doi.org/10.3390/antibiotics8020083

Medvecky, M., Cejkova, D., Polansky, O., Karasova, D., Kubasova, T., Cizek, A., Rychlik, I., 2018. Whole genome sequencing and function prediction of 133 gut anaerobes isolated from chicken caecum in pure cultures. BMC Genomics, 19, 561. https://doi.org/10.1186/s12864-018-4959-4

Mehmood, A., Baneen, U., Zaheer, A., Wasim Sajid, M., Hussain, A., Saleem, S., Yousafi, Q., Rashid, H., Riaz, H., Ihsan, A., Jamil, F., Sajjad, Y., Zahid, N., Shahzad Anjam, M., Arshad, M., Mirza, Z., Karim, S., Rasool, M., 2019. Physical and chemical mutagens improved *Sporotrichum* thermophile, strain ST20 for enhanced Phytase activity. Saudi J. Biol. Sci. 26, 1485–1491. https://doi.org/10.1016/j.sjbs.2019.07.006

Mehta, B.D., Jog, S.P., Johnson, S.C., Murthy, P.P., 2006. Lily pollen alkaline phytase is a histidine phosphatase similar to mammalian multiple inositol polyphosphate phosphatase (MINPP). Phytochemistry, 67, 1874–1886. https://doi.org/10.1016/j.phytochem.2006.06.008

Merchant, H.A., McConnell, E.L., Liu, F., Ramaswamy, C., Kulkarni, R.P., Basit, A.W., Murdan, S., 2011a. Assessment of gastrointestinal pH, fluid and lymphoid tissue in the guinea pig, rabbit and pig, and implications for their use in drug development. Eur. J. Pharm. Sci. 42, 3–10. https://doi.org/10.1016/j.ejps.2010.09.019

Minton, A.P., 2000. Effect of a Concentrated “Inert” Macromolecular Cosolute on the Stability of a Globular Protein with Respect to Denaturation by Heat and by Chaotropes: A Statistical-Thermodynamic Model. Biophys. J. 78, 101–109. https://doi.org/10.1016/S0006-3495(00)76576-3

Mitić, N., Smith, S.J., Neves, A., Guddat, L.W., Gahan, L.R., Schenk, G., 2006. The Catalytic Mechanisms of Binuclear Metallohydrolases. Chem. Rev. 106, 3338–3363. https://doi.org/10.1021/cr050318f

Molla, M.A.Z., Chowdhury, A.A., Islam, A., Hoque, S., 1984. Microbial mineralization of organic phosphate in soil. Plant Soil, 78, 393–399. https://doi.org/10.1007/BF02450372

Monserrate, J.P., York, J.D., 2010. Inositol phosphate synthesis and the nuclear processes they affect. Curr. Opin. Cell Biol. 22, 365–373. https://doi.org/10.1016/j.ceb.2010.03.006

Mori, T., Kamei, I., Hirai, H., Kondo, R., 2014. Identification of novel glycosyl hydrolases with cellulolytic activity against crystalline cellulose from metagenomic libraries constructed from bacterial enrichment cultures. Springerplus, 3, 365. https://doi.org/10.1186/2193-1801-3-365

Morse, D., Head, H.H., Wilcox, C.J., 1992. Disappearance of Phosphorus in Phytate from Concentrates In Vitro and from Rations Fed to Lactating Dairy Cows. J. Dairy Sci. 75, 1979–1986. https://doi.org/10.3168/jds.S0022-0302(92)77957-0

Mrudula Vasudevan, U., Jaiswal, A.K., Krishna, S., Pandey, A., 2019. Thermostable phytase in feed and fuel industries. Bioresour. Technol. 278, 400–407. https://doi.org/10.1016/j.biortech.2019.01.065

Mullaney, E.J., Gibson, D.M., Ullah, A.H.J., 1991. Positive identification of a lambda gt11 clone containing a region of fungal phytase gene by immunoprobe and sequence verification. Appl. Microbiol. Biotechnol. 35(5), 611-614. https://doi.org/10.1007/BF00169625

Mullaney, E.J., Ullah, A.H.J., 2007. Phytases: attributes, catalytic mechanisms and applications., in: Inositol Phosphates: Linking Agriculture and the Environment. CABI, Wallingford, 97–110. https://doi.org/10.1079/9781845931520.0097

Muslim, Sahira Nsayef, Mohammed Ali, A.N., AL-Kadmy, I.M.S., Khazaal, S.S., Ibrahim, S.A., Al-Saryi, N.A., Al-saadi, L.G., Muslim, Sraa Nsayef, Salman, B.K., Aziz, S.N., 2018. Screening, nutritional optimization and purification for phytase produced by Enterobacter aerogenes and its role in enhancement of hydrocarbons degradation and biofilm inhibition. Microb. Pathog. 115, 159–167. https://doi.org/10.1016/j.micpath.2017.12.047

Nacke, H., Daniel, R., 2014. Approaches in Metagenome Research: Progress and Challenges, in: Encyclopedia of Metagenomics. Springer New York, New York, NY, 1–7. https://doi.org/10.1007/978-1-4614-6418-1_790-3

Nagashima, T., Tange, T., Anazawa, H., 1999. Dephosphorylation of Phytate by Using the Aspergillus niger Phytase with a High Affinity for Phytate. Appl. Environ. Microbiol. 65, 4682–4684. https://doi.org/10.1128/AEM.65.10.4682-4684.1999

Naghoni, A., Emtiazi, G., Amoozegar, M.A., Cretoiu, M.S., Stal, L.J., Etemadifar, Z., Shahzadeh Fazeli, S.A., Bolhuis, H., 2017. Microbial diversity in the hypersaline Lake Meyghan, Iran. Sci. Rep. 7, 11522. https://doi.org/10.1038/s41598-017-11585-3

Nagul, E.A., McKelvie, I.D., Worsfold, P., Kolev, S.D., 2015. The molybdenum blue reaction for the determination of orthophosphate revisited: Opening the black box. Anal. Chim. Acta. https://doi.org/10.1016/j.aca.2015.07.030

Nakashima, B.A., McAllister, T.A., Sharma, R., Selinger, L.B., 2007. Diversity of phytases in the rumen. Microb.

Ecol. 53, 82–88. https://doi.org/10.1007/s00248-006-9147-4

Nam-Chul Ha, Byung-Chul Oh, Sejeong Shin, Hyun-Ju Kim, Tae-Kwang Oh, Young-Ok Kim, K.Y.C.& B.-H.O., 2000. Crystal structures of a novel, thermostable phytase in partially and fully calcium-loaded states. Nat. Struct. Biol. 7, 147–153.

Nath, B.V., 1940. A century of Liebig's theory of mineral nutrition of plants and of soil fertility. Curr. Sci. 9, 529–532.

Neal, A.L., Blackwell, M., Akkari, E., Guyomar, C., Clark, I., Hirsch, P.R., 2018. Phylogenetic distribution, biogeography and the effects of land management upon bacterial non-specific Acid phosphatase Gene diversity and abundance. Plant Soil, 427, 175–189. https://doi.org/10.1007/s11104-017-3301-2

Neal, A.L., Glendining, M.J., 2019. Calcium exerts a strong influence upon phosphohydrolase gene abundance and phylogenetic diversity in soil. Soil Biol. Biochem. 139, 107613. https://doi.org/10.1016/j.soilbio.2019.107613

Neal, A.L., Rossmann, M., Brearley, C., Akkari, E., Guyomar, C., Clark, I.M., Allen, E., Hirsch, P.R., 2017. Land-use influences phosphatase gene microdiversity in soils. Environ. Microbiol. 19, 2740–2753. https://doi.org/10.1111/1462-2920.13778

Neset, T.S.S., Cordell, D., 2012. Global phosphorus scarcity: Identifying synergies for a sustainable future. J. Sci. Food Agric. 92, 2–6. https://doi.org/10.1002/jsfa.4650

Nierman, W.C., Feldblyum, T.V., 2001. Genomic Library, in: Encyclopedia of Genetics. Elsevier, 865–872. https://doi.org/10.1006/rwgn.2001.0559

Nissar, J., Ahad, T., Naik, H.R. and Hussain, S.Z., 2017. A review phytic acid: As antinutrient or nutraceutical. J. Pharmacogn. Phytochem. 6, 1554–1560.

Niu, C., Luo, H., Shi, P., Huang, H., Wang, Y., Yang, P., Yao, B., 2016. N -Glycosylation Improves the Pepsin Resistance of Histidine Acid Phosphatase Phytases by Enhancing Their Stability at Acidic pHs and Reducing Pepsin's Accessibility to Its Cleavage Sites. Appl. Environ. Microbiol. 82, 1004–1014. https://doi.org/10.1128/AEM.02881-15

Nunes da Rocha, U., Van Overbeek, L., Van Elsas, J.D., 2009. Exploration of hitherto-uncultured bacteria from the rhizosphere. FEMS Microbiol. Ecol. 69, 313–328. https://doi.org/10.1111/j.1574-6941.2009.00702.x

O'Callaghan, A., van Sinderen, D., 2016. *Bifidobacteria* and Their Role as Members of the Human Gut Microbiota. Front. Microbiol. 7, 925. https://doi.org/10.3389/fmicb.2016.00925

Oakley, A.J., 2010. The structure of *Aspergillus niger* phytase PhyA in complex with a phytate mimetic. Biochem. Biophys. Res. Commun. 397, 745–749. https://doi.org/10.1016/j.bbrc.2010.06.024

Ocampo Betancur, Maritza, Patiño Cervantes, Luisa Fernanda, Marín Montoya, Mauricio, Salazar Yepes, Mauricio, & Gutiérrez Sánchez, P.A., 2012. Isolation and Characterization of Potential Phytase-Producing Fungi from Environmental Samples of Antioquia (Colombia). Rev. Fac. Nac. Agron. Medellín, 65, 6291–6303.

Ogut, M., Er, F., Kandemir, N., 2010. Phosphate solubilization potentials of soil *Acinetobacter* strains. Biol. Fertil. Soils, 46, 707–715. https://doi.org/10.1007/s00374-010-0475-7

Oh, B.-C., Chang, B.S., Park, K.-H., Ha, N.-C., Kim, H.-K., Oh, B.-H., Oh, T.-K., 2001. Calcium-Dependent Catalytic Activity of a Novel Phytase from *Bacillus amyloliquefaciens* DS11. Biochemistry, 40, 9669–9676. https://doi.org/10.1021/bi010589u

Oh, B.-C., Choi, W.-C., Park, S., Kim, Y. -o., Oh, T.-K., 2004. Biochemical properties and substrate specificities of alkaline and histidine acid phytases. Appl. Microbiol. Biotechnol. 63, 362–372. https://doi.org/10.1007/s00253-003-1345-0

Oh, B.-C., Kim, M.H., Yun, B.-S., Choi, W.-C., Park, S.-C., Bae, S.-C., Oh, T.-K., 2006. Ca^{2+}-Inositol Phosphate Chelation Mediates the Substrate Specificity of β-Propeller Phytase. Biochemistry, 45, 9531–9539. https://doi.org/10.1021/bi0603118

Okada, K., Iida, T., Kita-Tsukamoto, K., Honda, T., 2005. Vibrios Commonly Possess Two Chromosomes. J. Bacteriol. 187, 752–757. https://doi.org/10.1128/JB.187.2.752-757.2005

Olstorpe, M., Schnürer, J., Passoth, V., 2009. Screening of yeast strains for phytase activity. FEMS Yeast Res. 9, 478–488. https://doi.org/10.1111/j.1567-1364.2009.00493.x

Owensby, C.E., Cochran, R.C., Auen, L.M., 1996. Effects of Elevated Carbon Dioxide on Forage Quality for Ruminants, in: Carbon Dioxide, Populations, and Communities. Elsevier, 363–371. https://doi.org/10.1016/B978-012420870-4/50056-6

Öztürk, B., Werner, J., Meier-Kolthoff, J.P., Bunk, B., Spröer, C., Springael, D., 2020. Comparative Genomics Suggests Mechanisms of Genetic Adaptation toward the Catabolism of the Phenylurea Herbicide Linuron in *Variovorax*. Genome Biol. Evol. 12, 827–841. https://doi.org/10.1093/gbe/evaa085

Paik, I.K., 2001. Management of Excretion of Phosphorus, Nitrogen and Pharmacological Level Minerals to Reduce Environmental Pollution from Animal Production - Review -. Asian-Australasian J. Anim. Sci. 14, 384–394. https://doi.org/10.5713/ajas.2001.384

Pal Roy, M., Mazumdar, D., Dutta, S., Saha, S.P., Ghosh, S., 2016. Cloning and Expression of Phytase appA Gene from Shigella sp. CD2 in Pichia pastoris and Comparison of Properties with Recombinant Enzyme Expressed in *E. coli*. PLoS One 11, e0145745. https://doi.org/10.1371/journal.pone.0145745

Pandey, A., Szakacs, G., Soccol, C.R., Rodriguez-Leon, J.A., Soccol, V.T., 2001. Production, purification and properties of microbial phytases. Bioresour. Technol. 77, 203–214. https://doi.org/10.1016/S0960-8524(00)00139-5

Pantanella, F., Berlutti, F., Passariello, C., Sarli, S., Morea, C., Schippa, S., 2006. Violacein and biofilm production in *Janthinobacterium lividum*. J. Appl. Microbiol., 102(4), 992-999. https://doi.org/10.1111/j.1365-2672.2006.03155.x

Park, W.-Y., Matsui, T., Konishi, C., Kim, S.-W., Yano, F., Yano, H., 1999. Formaldehyde treatment suppresses ruminal degradation of phytate in soyabean meal and rapeseed meal. Br. J. Nutr. 81, 467–471. https://doi.org/10.1017/S0007114599000823

Pasamontes, L., Haiker, M., Wyss, M., Tessier, M., van Loon, A.P., 1997. Gene cloning, purification, and characterization of a heat-stable phytase from the fungus Aspergillus fumigatus. Appl. Environ. Microbiol. 63, 1696–1700. https://doi.org/10.1128/aem.63.5.1696-1700.1997

Peers, F.G., 1953. The phytase of wheat. Biochem. J. 53, 102–110. https://doi.org/10.1042/bj0530102

Peix, A., Lang, E., Verbarg, S., Spröer, C., Rivas, R., Santa-Regina, I., Mateos, P.F., Martínez-Molina, E., Rodríguez-Barrueco, C., Velázquez, E., 2009. *Acinetobacter* strains IH9 and OCI1, two rhizospheric phosphate solubilizing isolates able to promote plant growth, constitute a new genomovar of *Acinetobacter calcoaceticus*. Syst. Appl. Microbiol. 32, 334–341. https://doi.org/10.1016/j.syapm.2009.03.004

Pelicic, V., Reyrat, J.M., Gicquel, B., 1996. Expression of the *Bacillus subtilis* sacB gene confers sucrose sensitivity on mycobacteria. J. Bacteriol. 178, 1197–1199. https://doi.org/10.1128/jb.178.4.1197-1199.1996

Pereira, S.I.A., Castro, P.M.L., 2014. Phosphate-solubilizing rhizobacteria enhance *Zea mays* growth in agricultural P-deficient soils. Ecol. Eng. 73, 526–535. https://doi.org/10.1016/j.ecoleng.2014.09.060

Pham, V.H.T., Kim, J., 2012. Cultivation of unculturable soil bacteria. Trends Biotechnol. 30, 475–484. https://doi.org/10.1016/j.tibtech.2012.05.007

Phillippy, B.Q., 2006. Transport of calcium across caco-2 cells in the presence of inositol hexakisphosphate. Nutr. Res. 26, 146–149. https://doi.org/10.1016/j.nutres.2006.02.008

Pierrou, U., 1976. The Global Phosphorus Cycle. Ecol. Bull. No. 22, 75–88.

Pileggi, V.J., 1959. Distribution of phytase in the rat. Arch. Biochem. Biophys. 80, 1–8. https://doi.org/10.1016/0003-9861(59)90334-0

Planavsky, N.J., Rouxel, O.J., Bekker, A., Lalonde, S. V, Konhauser, K.O., Reinhard, C.T., Lyons, T.W., 2010. The evolution of the marine phosphate reservoir. Nature, 467. https://doi.org/10.1038/nature09485

Porco, T.C., Gao, D., Scott, J.C., Shim, E., Enanoria, W.T., Galvani, A.P., Lietman, T.M., 2012. When Does Overuse of Antibiotics Become a Tragedy of the Commons? PLoS One 7, e46505. https://doi.org/10.1371/journal.pone.0046505

Potter, S.C., Luciani, A., Eddy, S.R., Park, Y., Lopez, R., Finn, R.D., 2018. HMMER web server: 2018 update. Nucleic Acids Res. 46, W200–W204. https://doi.org/10.1093/nar/gky448

Prabhu, N., Borkar, S., Garg, S., 2019. Phosphate solubilization by microorganisms, in: Advances in Biological Science Research. Elsevier, 161–176. https://doi.org/10.1016/B978-0-12-817497-5.00011-2

Puhl, A.A., Greiner, R., Selinger, L.B., 2009. Stereospecificity of myo-inositol hexakisphosphate hydrolysis by a protein tyrosine phosphatase-like inositol polyphosphatase from *Megasphaera elsdenii*. Appl. Microbiol. Biotechnol. 82, 95–103. https://doi.org/10.1007/s00253-008-1734-5

Puhl, A.A., Gruninger, R.J., Greiner, R., Janzen, T.W., Mosimann, S.C., Selinger, L.B., 2007. Kinetic and structural analysis of a bacterial protein tyrosine phosphatase-like myo -inositol polyphosphatase. Protein Sci. 16, 1368–1378. https://doi.org/10.1110/ps.062738307

Quan, C., Zhang, L., Wang, Y., Ohta, Y., 2001. Production of phytase in a low phosphate medium by a novel yeast Candida krusei. J. Biosci. Bioeng. 92, 154–160. https://doi.org/10.1016/S1389-1723(01)80217-6

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., Glöckner, F.O., 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 41, D590–D596. https://doi.org/10.1093/nar/gks1219

Quillaguamán, J., Hatti-Kaul, R., Mattiasson, B., Alvarez, M.T., Delgado, O., 2004. *Halomonas boliviensis* sp. nov., an alkalitolerant, moderate halophile isolated from soil around a Bolivian hypersaline lake. Int. J. Syst. Evol. Microbiol. 54, 721–725. https://doi.org/10.1099/ijs.0.02800-0

Qvirist, L., Carlsson, N.-G., Andlid, T., 2015. Assessing phytase activity-methods, definitions and pitfalls. J. Biol. Methods, 2, e16. https://doi.org/10.14440/jbm.2015.58

Raboy, V., 2009. Approaches and challenges to engineering seed phytate and total phosphorus. Plant Sci. 177, 281–296. https://doi.org/10.1016/j.plantsci.2009.06.012

Raboy, V., Dickinson, D.B., 1993. Phytic acid levels in seeds of Glycine max and *G.soja* as influenced by phosphorus status. Crop Sci. 33, 1300–1305. https://doi.org/10.2135/cropsci1993.0011183X003300060036x

Raghothama, K.G., Karthikeyan, A.S., 2005. Phosphate Acquisition. Plant Soil 274, 37–49. https://doi.org/10.1007/s11104-004-2005-6

Ragot, S.A., Kertesz, M.A., Bunemann, E.K., 2015. phoD Alkaline Phosphatase Gene Diversity in Soil. Appl Env. Microbiol 81, 7281–7289. https://doi.org/10.1128/AEM.01823-15

Rapoport, S., Leva, E., and Guest, G.M., 1941. Phytase in plasma and erythrocytes of various species of vertebrates. J. Biol. Chem. 139, 621–632.

Raun, Arthur, Edmund Cheng, and W.B., 1956. Ruminant nutrition, phytate phosphorus hydrolysis and availability to rumen microorganisms. J. Agric. Food Chem. 4, 869–871.

Ravindran, V., Ravindran, G., Sivalogan, S., 1994. Total and phytate phosphorus contents of various foods and feedstuffs of plant origin. Food Chem. 50, 133–136. https://doi.org/10.1016/0308-8146(94)90109-0

Reddy, N.R., Sathe, S.K., 2001. Food phytates, CRC Press. https://doi.org/10.1201/9781420014419

Revy, P.., Jondreville, C., Dourmad, J.., Nys, Y., 2004. Effect of zinc supplemented as either an organic or an inorganic source and of microbial phytase on zinc and other minerals utilisation by weanling pigs. Anim. Feed Sci. Technol. 116, 93–112. https://doi.org/10.1016/j.anifeedsci.2004.04.003

Richardson, A.E., Hadobas, P.A., Hayes, J.E., 2001. Extracellular secretion of Aspergillus phytase from Arabidopsis roots enables plants to obtain phosphorus from phytate. Plant J. 25, 641–649. https://doi.org/10.1046/j.1365-313x.2001.00998.x

Rigden, D.J., 2008. The histidine phosphatase superfamily: structure and function. Biochem. J. 409, 333–348. https://doi.org/10.1042/BJ20071097

Rix, G.D., Todd, J.D., Neal, A.L., Brearley, C.A., 2020. Improved sensitivity, accuracy and prediction provided by a high-performance liquid chromatography screen for the isolation of phytase-harbouring organisms from environmental samples. Microb. Biotechnol. 14(4), 1409-1421. https://doi.org/10.1111/1751-7915.13733

Robertson, A.W., McCarville, N.G., MacIntyre, L.W., Correa, H., Haltli, B., Marchbank, D.H., Kerr, R.G., 2018. Isolation of Imaqobactin, an Amphiphilic Siderophore from the Arctic Marine Bacterium *Variovorax* Species RKJM285. J. Nat. Prod. 81, 858–865. https://doi.org/10.1021/acs.jnatprod.7b00943

Rodriguez, E., Han, Y., Lei, X.G., 1999a. Cloning, Sequencing, and Expression of an *Escherichia coli* Acid Phosphatase/Phytase Gene (appA2) Isolated from Pig Colon. Biochem. Biophys. Res. Commun. 257, 117–123. https://doi.org/10.1006/bbrc.1999.0361

Rodriguez, E., Porres, J.M., Han, Y., Lei, X.G., 1999b. Different Sensitivity of Recombinant *Aspergillus niger* Phytase (r-PhyA) and *Escherichia* coli pH 2.5 Acid Phosphatase (r-AppA) to Trypsin and Pepsinin Vitro. Arch. Biochem. Biophys. 365, 262–267. https://doi.org/10.1006/abbi.1999.1184

Romano, P.R., Wang, J., O'Keefe, R.J., Puzas, J.E., Rosier, R.N., Reynolds, P.R., 1998. HiPER1, a phosphatase of the endoplasmic reticulum with a role in chondrocyte maturation. J. Cell Sci. 111, 803–813. https://doi.org/10.1242/jcs.111.6.803

Roy, Moushree Pal, Madhumita Poddar, Kamal Krishna Singh, and S.G., 2012. Purification, characterization and properties of phytase from *Shigella* sp. CD2. Indian J. Biochem. Biophys. 49, 266–271.

Ruttenberg, K.C., 2003. The Global Phosphorus Cycle, in: Treatise on Geochemistry. Elsevier, 585–643. https://doi.org/10.1016/B0-08-043751-6/08153-6

Sabiha-Javied, Mehmood, T., Chaudhry, M.M., Tufail, M., Irfan, N., 2009. Heavy metal pollution from phosphate rock used for the production of fertilizer in Pakistan. Microchem. J. 91, 94–99. https://doi.org/10.1016/j.microc.2008.08.009

Sajidan, A., Farouk, A., Greiner, R., Jungblut, P., Muller, E.C., Borriss, R., 2004. Molecular and physiological characterisation of a 3-phytase from soil bacterium Klebsiella sp. ASR1. Appl Microbiol Biotechnol, 65, 110–118. https://doi.org/10.1007/s00253-003-1530-1

Salaet, I., Marques, R., Yance-Chávez, T., Macías-Vidal, J., Gimémez-Zaragoza, D., Aligue, R., 2021. Novel Long-Term Phytase from Serratia odorifera : Cloning, Expression, and Characterization. ACS Food Sci. Technol. 1, 689–697. https://doi.org/10.1021/acsfoodscitech.0c00074

Salyers, A.A., 1984. Bacteroides of the Human Lower Intestinal Tract. Annu. Rev. Microbiol. 38, 293–313. https://doi.org/10.1146/annurev.mi.38.100184.001453

Sanchez, P.A., 2002. Soil Fertility and Hunger in Africa. Science, 295(5562), 2019–2020. https://doi.org/10.1126/science.1065256

Santos, S.R., Ochman, H., 2004. Identification and phylogenetic sorting of bacterial lineages with universally conserved genes and proteins. Environ. Microbiol. 6, 754–759. https://doi.org/10.1111/j.1462-2920.2004.00617.x

Santos, T., Connolly, C., Murphy, R., 2015. Trace Element Inhibition of Phytase Activity. Biol. Trace Elem. Res. 163, 255–265. https://doi.org/10.1007/s12011-014-0161-y

Sarikhani, M.R., Malboobi, M.A., Aliasgharzad, N., Greiner, R., 2019. Identification of two novel bacterial phosphatase-encoding genes in Pseudomonas putida strain P13. J Appl Microbiol 127, 1113–1124. https://doi.org/10.1111/jam.14376

Satola, B., Wübbeler, J.H., Steinbüchel, A., 2013. Metabolic characteristics of the species *Variovorax paradoxus*. Appl. Microbiol. Biotechnol. 97, 541–560. https://doi.org/10.1007/s00253-012-4585-z

Schenk, G., Elliott, T.W., Leung, E., Carrington, L.E., Mitić, N., Gahan, L.R., Guddat, L.W., 2008. Crystal structures of a purple acid phosphatase, representing different steps of this enzyme's catalytic cycle. BMC Struct. Biol. 8(1), 1-13. https://doi.org/10.1186/1472-6807-8-6

Schenk, G., Mitić, N., Hanson, G.R., Comba, P., 2013. Purple acid phosphatase: A journey into the function and mechanism of a colorful enzyme. Coord. Chem. Rev. 257, 473–482. https://doi.org/10.1016/j.ccr.2012.03.020

Schippers, A., Bosecker, K., Spröer, C., Schumann, P., 2005. *Microbacterium oleivorans* sp. nov. and *Microbacterium hydrocarbonoxydans* sp. nov., novel crude-oil-degrading Gram-positive bacteria. Int. J. Syst. Evol. Microbiol. 55, 655–660. https://doi.org/10.1099/ijs.0.63305-0

Schoene, C., Bennett, S.P., Howarth, M., 2016. SpyRing interrogation: analyzing how enzyme resilience can be achieved with phytase and distinct cyclization chemistries. Sci. Rep., 6(1), 1-12. https://doi.org/10.1038/srep21151

Scholz, R.W., Ulrich, A.E., Eilittä, M., Roy, A., 2013. Sustainable use of phosphorus: A finite resource. Sci. Total Environ., 461, 799-803. https://doi.org/10.1016/j.scitotenv.2013.05.043

Schröder, J.J., Smit, A.L., Cordell, D., Rosemarin, A., 2011. Improved phosphorus use efficiency in agriculture: A key requirement for its sustainable use. Chemosphere 84, 822–831. https://doi.org/10.1016/j.chemosphere.2011.01.065

Segueilha, L., Lambrechts, C., Boze, H., Moulin, G., Galzy, P., 1992. Purification and properties of the phytase from *Schwanniomyces castellii*. J. Ferment. Bioeng. 74, 7–11. https://doi.org/10.1016/0922-338X(92)90259-W

Selle Ravindran, V., P.H., 2007. Microbial phytase in poultry nutrition. Anim. Feed Sci. Technol. 135(1-2), 1-41. https://doi.org/10.1016/j.anifeedsci.2006.06.010

Selle, P.H., Cowieson, A.J., Ravindran, V., 2009. Consequences of calcium interactions with phytate and phytase for poultry and pigs. Livest. Sci. 124(1-3), 126–141. https://doi.org/10.1016/j.livsci.2009.01.006

Selle, P.H., Ravindran, V., 2008. Phytate-degrading enzymes in pig nutrition. Livest. Sci. 113, 99–122. https://doi.org/10.1016/j.livsci.2007.05.014

Selle, P.H., Ravindran, V., Caldwell, A., Bryden, W.L., 2000. Phytate and phytase: consequences for protein utilisation. Nutr. Res. Rev. 13(2), 255-278. https://doi.org/10.1079/095442200108729098

Sen, S., Venkata Dasu, V., Mandal, B., 2007. Developments in Directed Evolution for Improving Enzyme Functions. Appl. Biochem. Biotechnol. 143, 212–223. https://doi.org/10.1007/s12010-007-8003-4

Servin, J.A., Herbold, C.W., Skophammer, R.G., Lake, J.A., 2007. Evidence Excluding the Root of the Tree of Life

from the Actinobacteria. Mol. Biol. Evol. 25, 1–4. https://doi.org/10.1093/molbev/msm249

Shallcross, L.J., Howard, S.J., Fowler, T., Davies, S.C., 2015. Tackling the threat of antimicrobial resistance: from policy to sustainable action. Philos. Trans. R. Soc. B Biol. Sci. 370(1670), 20140082. https://doi.org/10.1098/rstb.2014.0082

Sharma, S., Radl, V., Hai, B., Kloos, K., Mrkonjic Fuka, M., Engel, M., Schauss, K., Schloter, M., 2007. Quantification of functional genes from procaryotes in soil by PCR. J. Microbiol. Methods 68, 445–452. https://doi.org/10.1016/j.mimet.2006.10.001

Sharma, S.B., Sayyed, R.Z., Trivedi, M.H., Gobi, T.A., 2013. Phosphate solubilizing microbes: sustainable approach for managing phosphorus deficiency in agricultural soils. Springerplus 2(1), 1-14. https://doi.org/10.1186/2193-1801-2-587

Sharpley, A., Jarvie, H., Flaten, D., Kleinman, P., 2018. Celebrating the 350th Anniversary of Phosphorus Discovery: A Conundrum of Deficiency and Excess. J. Environ. Qual. 47, 774–777. https://doi.org/10.2134/jeq2018.05.0170

Shedova, E., Lipasova, V., Velikodvorskaya, G., Ovadis, M., Chernin, L., Khmel, I., 2008. Phytase activity and its regulation in a rhizospheric strain of Serratia plymuthica. Folia Microbiol. (Praha). 53, 110–114. https://doi.org/10.1007/s12223-008-0016-z

Sheu, S.-Y., Chiu, T.F., Chou, J.-H., Sheu, D.-S., Arun, A.B., Young, C.-C., Chen, C.A., Wang, J.-T., Chen, W.-M., 2009. Andreprevotia lacus sp. nov., isolated from a fish-culture pond. Int. J. Syst. Evol. Microbiol. 59, 2482–2485. https://doi.org/10.1099/ijs.0.009233-0

Shi, J., Wang, H., Hazebroek, J., Ertl, D.S., Harp, T., 2005. The maize low-phytic acid 3 encodes a myo-inositol kinase that plays a role in phytic acid biosynthesis in developing seeds. Plant J, 42, 708–719. https://doi.org/10.1111/j.1365-313X.2005.02412.x

Shi, P., Huang, H., Wang, Y., Luo, H., Wu, B., Meng, K., Yang, P., Yao, B., 2008. A novel phytase gene appA from *Buttiauxella* sp. GC21 isolated from grass carp intestine. Aquaculture, 275, 70–75. https://doi.org/10.1016/j.aquaculture.2008.01.021

Shimizu, M., 1993. Purification and Characterization of Phytase and Acid Phosphatase Produced by *Aspergillus oryzae* K1. Biosci. Biotechnol. Biochem. 57, 1364–1365. https://doi.org/10.1271/bbb.57.1364

Shin, S., Ha, N.C., Oh, B.C., Oh, T.K., Oh, B.H., 2001. Enzyme mechanism and catalytic property of beta propeller phytase. Structure., 9(9), 851-858.

SHINKAI, T., UEKI, T., KOBAYASHI, Y., 2010. Detection and identification of rumen bacteria constituting a fibrolytic consortium dominated by Fibrobacter succinogenes. Anim. Sci. J. 81, 72–79. https://doi.org/10.1111/j.1740-0929.2009.00698.x

Shivani, Y., Subhash, Y., Sasikala, C., Ramana, C. V., 2016. Description of '*Candidatus Marispirochaeta* associata' and reclassification of *Spirochaeta bajacaliforniensis*, *Spirochaeta smaragdinae* and *Spirochaeta sinaica* to a new genus *Sediminispirochaeta* gen. nov. as *Sediminispirochaeta bajacaliforniensis* comb. nov. Int. J. Syst. Evol. Microbiol. 66, 5485–5492. https://doi.org/10.1099/ijsem.0.001545

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E. V., Zdobnov, E.M., 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics., 31(19), 3210-3212. https://doi.org/10.1093/bioinformatics/btv351

Simon, O., Igbasan, F., 2002. In vitro properties of phytases from various microbial origins. Int. J. Food Sci. Technol., 37(7), 813-822. https://doi.org/10.1046/j.1365-2621.2002.00621.x

Singh, B., Satyanarayana, T., 2011. Microbial phytases in phosphorus acquisition and plant growth promotion. Physiol. Mol. Biol. Plants, 17, 93–103. https://doi.org/10.1007/s12298-011-0062-x

Singh, B., Satyanarayana, T., 2008. Phytase production by a thermophilic mould Sporotrichum thermophile in solid state fermentation and its potential applications. Bioresour. Technol. 99, 2824–2830. https://doi.org/10.1016/j.biortech.2007.06.010

Sleator, R.D., Shortall, C., Hill, C., 2008. Metagenomics. Lett. Appl. Microbiol. 47, 361–366. https://doi.org/10.1111/j.1472-765X.2008.02444.x

Slominski, B.A., Davie, T., Nyachoti, M.C., Jones, O., 2007. Heat stability of endogenous and microbial phytase during feed pelleting. Livest. Sci. 109, 244–246. https://doi.org/10.1016/j.livsci.2007.01.124

Smale, S.T., 2010. Beta-Galactosidase Assay. Cold Spring Harb. Protoc., 2010(5), pdb-prot5423.

Smanski, M.J., Bhatia, S., Zhao, D., Park, Y., B A Woodruff, L., Giannoukos, G., Ciulla, D., Busby, M., Calderon, J.,

Nicol, R., Gordon, D.B., Densmore, D., Voigt, C.A., 2014. Functional optimization of gene clusters by combinatorial design and assembly. Nat. Biotechnol. 32, 1241–1249. https://doi.org/10.1038/nbt.3063

Smil, V., 2000. PHOSPHORUS IN THE ENVIRONMENT: Natural Flows and Human Interferences. Annu. Rev. Energy Environ. 25, 53–88. https://doi.org/10.1146/annurev.energy.25.1.53

Spaink, H.P., Okker, R.J.H., Wijffelman, C.A., Pees, E., Lugtenberg, B.J.J., 1987. Promoters in the nodulation region of the Rhizobium leguminosarum Sym plasmid pRL1JI. Plant Mol. Biol. 9, 27–39. https://doi.org/10.1007/BF00017984

Stahl, C.H., Han, Y.M., Roneker, K.R., House, W.A., Lei, X.G., 1999. Phytase Improves Iron Bioavailability for Hemoglobin Synthesis in Young Pigs. J. Anim. Sci. 77, 2135. https://doi.org/10.2527/1999.7782135x

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics, 22, 2688–2690. https://doi.org/10.1093/bioinformatics/btl446

Stentz, R., Osborne, S., Horn, N., Li, A.W.H., Hautefort, I., Bongaerts, R., Rouyer, M., Bailey, P., Shears, S.B., Hemmings, A.M., Brearley, C.A., Carding, S.R., 2014. A Bacterial Homolog of a Eukaryotic Inositol Phosphate Signaling Enzyme Mediates Cross-kingdom Dialog in the Mammalian Gut. Cell Rep. 6, 646–656. https://doi.org/10.1016/j.celrep.2014.01.021

Stephen Jasinski, 2021. Mineral Commodity Summaries. https://doi.org/https://doi.org/10.3133/mcs2021

Stevenson-Paulik, J., Bastidas, R.J., Chiou, S.T., Frye, R.A., York, J.D., 2005. Generation of phytate-free seeds in Arabidopsis through disruption of inositol polyphosphate kinases. Proc Natl Acad Sci. 102(35), 12612–12617. https://doi.org/10.1073/pnas.0504172102

Stout, L.M., Nguyen, T.T., Jaisi, D.P., 2016. Relationship of Phytate, Phytate-Mineralizing Bacteria, and Beta-Propeller Phytase Genes along a Coastal Tributary to the Chesapeake Bay. Soil Sci. Soc. Am. J. 80, 84–96. https://doi.org/10.2136/sssaj2015.04.0146

Sulabo, R.C., Jones, C.K., Tokach, M.D., Goodband, R.D., Dritz, S.S., Campbell, D.R., Ratliff, B.W., DeRouchey, J.M., Nelssen, J.L., 2011. Factors affecting storage stability of various commercial phytase sources. J. Anim. Sci. 89, 4262–4271. https://doi.org/10.2527/jas.2011-3948

Suleimanova, A.D., Beinhauer, A., Valeeva, L.R., Chastukhina, I.B., Balaban, N.P., Shakirov, E. V., Greiner, R., Sharipova, M.R., 2015. Novel glucose-1-phosphatase with high phytase activity and unusual metal ion activation from soil bacterium *Pantoea* sp. strain 3.5.1. Appl. Environ. Microbiol. 81, 6790–6799. https://doi.org/10.1128/AEM.01384-15

Suleimanova, A.D., Bulmakova, D.S., Itkina, D.L., Sharipova, M.R., 2021. Expression of *Pantoea* sp. 3.5.1 AgpP Phytase in Three Expression Systems. Bionanoscience 11, 648–652. https://doi.org/10.1007/s12668-021-00855-0

Sun, D., 2018. Pull in and Push Out: Mechanisms of Horizontal Gene Transfer in Bacteria. Front. Microbiol. 9, 2154. https://doi.org/10.3389/fmicb.2018.02154

Sun, D., Jeannot, K., Xiao, Y., Knapp, C.W., 2019. Editorial: Horizontal Gene Transfer Mediated Bacterial Antibiotic Resistance. Front. Microbiol., 10, 1933. https://doi.org/10.3389/fmicb.2019.01933

Suzuki, U., K. Yoshimura, and M.T., 1907. Ueber ein Enzym “Phytase” das “Anhydro-oxy-methylen diphosphorsaure” Spaltet. Tokyo Imper Univ Coll Agric Bull 7, 503–12.

Suzuki, S., Kurosawa, N., 2017. Development of the Multiple Gene Knockout System with One-Step PCR in Thermoacidophilic Crenarchaeon Sulfolobus acidocaldarius. Archaea, 2017, 1–12. https://doi.org/10.1155/2017/7459310

Suzumura, M., Kamatani, A., 1995. Mineralization of inositol hexaphosphate in aerobic and anaerobic marine sediments: Implications for the phosphorus cycle. Geochim. Cosmochim. Acta, 59, 1021–1026. https://doi.org/10.1016/0016-7037(95)00006-2

Tamayo-Ramos, J.A., Sanz-Penella, J.M., Yebra, M.J., Monedero, V., Haros, M., 2012. Novel Phytases from *Bifidobacterium pseudocatenulatum* ATCC 27919 and *Bifidobacterium longum* subsp. infantis ATCC 15697. Appl. Environ. Microbiol. 78, 5013–5015. https://doi.org/10.1128/AEM.00782-12

Tamim, N.M., Angel, R., Christman, M., 2004. Influence of dietary calcium and phytase on phytate phosphorus hydrolysis in broiler chickens. Poult. Sci. 83, 1358–1367. https://doi.org/10.1093/ps/83.8.1358

Tan, H., Mooij, M.J., Barret, M., Hegarty, P.M., Harington, C., Dobson, A.D.W., O'Gara, F., 2014. Identification of Novel Phytase Genes from an Agricultural Soil-Derived Metagenome. J. Microbiol. Biotechnol. 24(1), 113–118. https://doi.org/10.4014/jmb.1307.07007

Tang, J., Leung, A., Leung, C., Lim, B.L., 2006. Hydrolysis of precipitated phytate by three distinct families of phytases. Soil Biol. Biochem. 38, 1316–1324. https://doi.org/10.1016/j.soilbio.2005.08.021

Tang, Z., Jin, W., Sun, R., Liao, Y., Zhen, T., Chen, H., Wu, Q., Gou, L., Li, C., 2018. Improved thermostability and enzyme activity of a recombinant phyA mutant phytase from *Aspergillus niger* N25 by directed evolution and site-directed mutagenesis. Enzyme Microb. Technol. 108, 74–81. https://doi.org/10.1016/j.enzmictec.2017.09.010

Tanner, M.A., Goebel, B.M., Dojka, M.A., Pace, N.R., 1998. Specific Ribosomal DNA Sequences from Diverse Environmental Settings Correlate with Experimental Contaminants. Appl. Environ. Microbiol. 64, 3110–3113. https://doi.org/10.1128/AEM.64.8.3110-3113.1998

Taylor, C.W., Thorn, P., 2001. Calcium signalling: IP 3 rises again… and again. Curr. Biol. 11, 352–355.

Taylor, S., Wakem, M., Dijkman, G., Alsarraj, M., Nguyen, M., 2010. A practical approach to RT-qPCR—Publishing data that conform to the MIQE guidelines. Methods 50, S1–S5. https://doi.org/10.1016/j.ymeth.2010.01.005

Thomas, T., Gilbert, J., Meyer, F., 2012. Metagenomics - a guide from sampling to data analysis. Microb. Inform. Exp. 2, 3. https://doi.org/10.1186/2042-5783-2-3

Thota, S.G., Bhandari, R., 2015. The emerging roles of inositol pyrophosphates in eukaryotic cell physiology. J. Biosci. 40, 593–605. https://doi.org/10.1007/s12038-015-9549-x

Tidjani, A.-R., Lorenzi, J.-N., Toussaint, M., van Dijk, E., Naquin, D., Lespinet, O., Bontemps, C., Leblond, P., 2019. Massive Gene Flux Drives Genome Diversity between Sympatric *Streptomyces* Conspecifics. MBio 10(5), e01533-19. https://doi.org/10.1128/mBio.01533-19

Tomschy, A., Brugger, R., Lehmann, M., Svendsen, A., Vogel, K., Kostrewa, D., Lassen, S.F., Burger, D., Kronenberger, A., van Loon, A.P.G.M., Pasamontes, L., Wyss, M., 2002. Engineering of Phytase for Improved Activity at Low pH. Appl. Environ. Microbiol. 68, 1907–1913. https://doi.org/10.1128/AEM.68.4.1907-1913.2002

Trevino-Garrison, I., DeMent, J., Ahmed, F., Haines-Lieber, P., Langer, T., Ménager, H., Neff, J., van der Merwe, D., Carney, E., 2015. Human Illnesses and Animal Deaths Associated with Freshwater Harmful Algal Blooms—Kansas. Toxins, 7, 353–366. https://doi.org/10.3390/toxins7020353

Trucksis, M., Michalski, J., Deng, Y.K., Kaper, J.B., 1998. The Vibrio cholerae genome contains two unique circular chromosomes. Proc. Natl. Acad. Sci. 95, 14464–14469. https://doi.org/10.1073/pnas.95.24.14464

Tseng, Y.H., Fang, T.J., Tseng, S.M., 2000. Isolation and characterization of a novel phytase from *Penicillium simplicissimum*. Folia Microbiol. 45(2), 121-127. https://doi.org/10.1007/BF02817409

Turner, B.L., Papházy, M.J., Haygarth, P.M., Mckelvie, I.D., 2002. Inositol phosphates in the environment. Philosophi. transac., 357(1420), 449-469 https://doi.org/10.1098/rstb.2001.0837

Ullah, A.H.., Sethumadhavan, K., 2003. PhyA gene product of *Aspergillus ficuum* and *Peniophora lycii* produces dissimilar phytases. Biochem. Biophys. Res. Commun. 303, 463–468. https://doi.org/10.1016/S0006-291X(03)00374-7

Ullah, A.H.J., Gibson, D.M., 1987. Extracellular Phytase (E. C. 3.1.3.8) from *Aspergillus Ficuum* NRRL 3135: Purification and Characterization. Prep. Biochem. 17, 63–91. https://doi.org/10.1080/00327488708062477

Ullah, A.H.J., Sethumadhavan, K., Lei, X.G., Mullaney, E.J., 2000. Biochemical Characterization of Cloned *Aspergillus fumigatus* Phytase (phyA). Biochem. Biophys. Res. Commun. 275, 279–285. https://doi.org/10.1006/bbrc.2000.3271

Unno, Y., Okubo, K., Wasaki, J., Shinano, T., Osaki, M., 2005. Plant growth promotion abilities and microscale bacterial dynamics in the rhizosphere of Lupin analysed by phytate utilization ability. Environ. Microbiol. 7, 396–404. https://doi.org/10.1111/j.1462-2920.2004.00701.x

Ushasree, M.V., Shyam, K., Vidya, J., Pandey, A., 2017. Microbial phytase: Impact of advances in genetic engineering in revolutionizing its properties and applications. Bioresour. Technol. 245, 1790-1799. https://doi.org/10.1016/j.biortech.2017.05.060

Vadstein, O., 2000. Heterotrophic, Planktonic Bacteria and Cycling of Phosphorus., Advances in microbial ecology, 115–167. https://doi.org/10.1007/978-1-4615-4187-5_4

van de Velde, S.J., Hidalgo-Martinez, S., Callebaut, I., Antler, G., James, R.K., Leermakers, M., Meysman, F.J.R., 2020. Burrowing fauna mediate alternative stable states in the redox cycling of salt marsh sediments. Geochim. Cosmochim. Acta 276, 31–49. https://doi.org/10.1016/j.gca.2020.02.021

van der Ploeg, R.R., Böhm, W., Kirkham, M.B., 1999. On the Origin of the Theory of Mineral Nutrition of Plants and the Law of the Minimum. Soil Sci. Soc. Am. J. 63, 1055–1062. https://doi.org/10.2136/sssaj1999.6351055x

Van Elsas, J.D., Turner, S., Bailey, M.J., 2003. Horizontal gene transfer in the phytosphere. New Phytol. 157, 525–537. https://doi.org/10.1046/j.1469-8137.2003.00697.x

van Hartingsveldt, W., van Zeijl, C.M.J., Harteveld, G.M., Gouka, R.J., Suykerbuyk, M.E.G., Luiten, R.G.M., van Paridon, P.A., Selten, G.C.M., Veenstra, A.E., van Gorcom, R.F.M., van den Hondel, C.A.M.J.., 1993. Cloning, characterization and overexpression of the phytase-encoding gene (phyA) of *Aspergillus niger*. Gene 127, 87–94. https://doi.org/10.1016/0378-1119(93)90620-I

Van Schie, B.J., De Mooy, O.H., Linton, J.D., Van Dijken, J.P., Kuenen, J.G., 1987. PQQ-Dependent Production of Gluconic Acid by *Acinetobacter, Agrobacterium* and *Rhizobium* Species. Microbiology 133, 867–875. https://doi.org/10.1099/00221287-133-4-867

Van Vuuren, D.P., Bouwman, A.F., Beusen, A.H.W., 2010. Phosphorus demand for the 1970-2100 period: A scenario analysis of resource depletion. Glob. Environ. Chang. 20, 428–439. https://doi.org/10.1016/j.gloenvcha.2010.04.004

Velicer, G.J., 2003. Social strife in the microbial world. Trends Microbiol., 11, 330–337. https://doi.org/10.1016/S0966-842X(03)00152-5

Venegas, F.A., Köllisch, G., Mark, K., Diederich, W.E., Kaufmann, A., Bauer, S., Chavarría, M., Araya, J.J., García-Piñeres, A.J., 2019. The Bacterial Product Violacein Exerts an Immunostimulatory Effect Via TLR8. Sci. Rep. 9(1), 1-17. https://doi.org/10.1038/s41598-019-50038-x

Ventosa, A., Haba, R.R., Arahal, D.R., Sánchez-Porro, C., 2021. Halomonas, in: Oren, A. (Ed.), Bergey's Manual of Systematics of Archaea and Bacteria. Wiley, 1–111. https://doi.org/10.1002/9781118960608.gbm01190.pub2

Vincent, J.B., Crowder, M.W., Averill, B.A., 1992. Hydrolysis of phosphate monoesters: a biological problem with multiple chemical solutions. Trends Biochem. Sci. 17, 105–110. https://doi.org/10.1016/0968-0004(92)90246-6

Vinod Kumar, Ajar Nath Yadav, Abhishake Saxena, P.S. and, Dhaliwal, H.S., 2016. Unravelling Rhizospheric Diversity and Potential of Phytase Producing Microbes. SM J. Biol. 2(1), 1009.

von Wintersdorff, C.J.H., Penders, J., van Niekerk, J.M., Mills, N.D., Majumder, S., van Alphen, L.B., Savelkoul, P.H.M., Wolffs, P.F.G., 2016. Dissemination of Antimicrobial Resistance in Microbial Ecosystems through Horizontal Gene Transfer. Front. Microbiol. 173. https://doi.org/10.3389/fmicb.2016.00173

Walan, P., Davidsson, S., Johansson, S., Höök, M., 2014. Phosphate rock production and depletion: Regional disaggregated modeling and global implications. Resour. Conserv. Recycl. 93, 178–187. https://doi.org/10.1016/j.resconrec.2014.10.011

Wallace, R.B., Baumann, H., Grear, J.S., Aller, R.C., Gobler, C.J., 2014. Coastal ocean acidification: The other eutrophication problem. Estuar. Coast. Shelf Sci. 148, 1–13. https://doi.org/10.1016/j.ecss.2014.05.027

Wan, B., Yan, Y., Liu, F., Tan, W., Chen, X., Feng, X., 2016. Surface adsorption and precipitation of inositol hexakisphosphate on calcite: A comparison with orthophosphate. Chem. Geol. 421, 103–111. https://doi.org/10.1016/j.chemgeo.2015.12.004

Wang, Xueying, Suchart Upatham, Watanalai Panbangred, Duangnate Isarangkul, Pijug Summpunn, Suthep Wiyakrutta, and V.M., 2004. Purification, characterization, gene cloning and sequence analysis of a phytase from *Klebsiella pneumoniae* subsp. *pneumoniae* XY-5. Sci. Asia 30, 383–390.

Wang, Y., Sarkar, M., Smith, A.E., Krois, A.S., Pielak, G.J., 2012. Macromolecular Crowding and Protein Stability. J. Am. Chem. Soc. 134, 16614–16618. https://doi.org/10.1021/ja305300m

Waśkiewicz, A., Irzykowska, L., 2014. Flavobacterium spp. – Characteristics, Occurrence, and Toxicity, in: Encyclopedia of Food Microbiology. Elsevier, 938–942. https://doi.org/10.1016/B978-0-12-384730-0.00126-9

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F.T., de Beer, T.A.P., Rempfer, C., Bordoli, L., Lepore, R., Schwede, T., 2018. SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res. 46, W296–W303. https://doi.org/10.1093/nar/gky427

Watts, C.W., Clark, L.J., Poulton, P.R., Powlson, D.S., Whitmore, A.P., 2006. The role of clay, organic carbon and long-term management on mouldboard plough draught measured on the Broadbalk wheat experiment at Rothamsted. Soil Use Manag. 22, 334–341. https://doi.org/10.1111/j.1475-2743.2006.00054.x

Watve, M., Tickoo, R., Jog, M., Bhole, B., 2001. How many antibiotics are produced by the genus *Streptomyces*? Arch. Microbiol., 176, 386–390. https://doi.org/10.1007/s002030100345

Wexler, H.M., 2007. Bacteroides : the Good, the Bad, and the Nitty-Gritty. Clin. Microbiol. Rev., 20, 593–621.

https://doi.org/10.1128/CMR.00008-07

Whitelaw, M.A., Harden, T.J., Helyar, K.R., 1999. Phosphate solubilisation in solution culture by the soil fungus *Penicillium radicum*. Soil Biol. Biochem., 31, 655–665. https://doi.org/10.1016/S0038-0717(98)00130-8

Whitfield, H., Riley, A.M., Diogenous, S., Godage, H.Y., Potter, B.V.L., Brearley, C.A., 2018. Simple synthesis of (32)P-labelled inositol hexakisphosphates for study of phosphate transformations. Plant Soil, 427, 149–161. https://doi.org/10.1007/s11104-017-3315-9

Wilkening, J. V., Turchyn, A. V., Redeker, K.R., Mills, J. V., Antler, G., Carrión, O., Todd, J.D., 2019. The Production and Fate of Volatile Organosulfur Compounds in Sulfidic and Ferruginous Sediment. J. Geophys. Res. Biogeosciences, 124, 3390–3402. https://doi.org/10.1029/2019JG005248

Wilkinson, S.J., Selle, P.H., Bedford, M.R., Cowieson, A.J., 2014. Separate feeding of calcium improves performance and ileal nutrient digestibility in broiler chicks. Anim. Prod. Sci., 54, 172. https://doi.org/10.1071/AN12432

Wilson, M.S.C., Livermore, T.M., Saiardi, A., 2013. Inositol pyrophosphates: between signalling and metabolism. Biochem. J., 452, 369–379. https://doi.org/10.1042/BJ20130118

Woese, C.R., Olsen, G.J., Ibba, M., Söll, D., 2000. Aminoacyl-tRNA Synthetases, the Genetic Code, and the Evolutionary Process. Microbiol. Mol. Biol. Rev., 64, 202–236. https://doi.org/10.1128/MMBR.64.1.202-236.2000

Woo, J., MacPherson, C., Liu, J., Wang, H., Kiba, T., Hannah, M.A., Wang, X.-J., Bajic, V.B., Chua, N.-H., 2012. The response and recovery of the *Arabidopsis thaliana* transcriptome to phosphate starvation. BMC Plant Biol. 12, 62. https://doi.org/10.1186/1471-2229-12-62

Wu, T.-H., Chen, C.-C., Cheng, Y.-S., Ko, T.-P., Lin, C.-Y., Lai, H.-L., Huang, T.-Y., Liu, J.-R., Guo, R.-T., 2014. Improving specific activity and thermostability of *Escherichia coli* phytase by structure-based rational design. J. Biotechnol. 175, 1–6. https://doi.org/10.1016/j.jbiotec.2014.01.034

Wuddivira, M.N., Camps-Roach, G., 2007. Effects of organic matter and calcium on soil structural stability. Eur. J. Soil Sci. 58, 722–727. https://doi.org/10.1111/j.1365-2389.2006.00861.x

Wyss, M., Brugger, R., Kronenberger, A., Rémy, R., Fimbel, R., Oesterhelt, G., Lehmann, M., van Loon, A.P.G.M., 1999. Biochemical Characterization of Fungal Phytases (myo -Inositol Hexakisphosphate Phosphohydrolases): Catalytic Properties. Appl. Environ. Microbiol. 65, 367–373. https://doi.org/10.1128/AEM.65.2.367-373.1999

Wyss, M., Pasamontes, L., Rémy, R., Kohler, J., Kusznir, E., Gadient, M., Müller, F., van Loon, A.P.G.M., 1998. Comparison of the Thermostability Properties of Three Acid Phosphatases from Molds: *Aspergillus fumigatus* Phytase, *A. niger* Phytase, and *A. niger* pH 2.5 Acid Phosphatase. Appl. Environ. Microbiol., 64, 4446–4451. https://doi.org/10.1128/AEM.64.11.4446-4451.1998

Xiong, A.-S., Yao, Q.-H., Peng, R.-H., Han, P.-L., Cheng, Z.-M., Li, Y., 2005. High level expression of a recombinant acid phytase gene in Pichia pastoris. J. Appl. Microbiol., 98, 418–428. https://doi.org/10.1111/j.1365-2672.2004.02476.x

Xiong, A.S., Yao, Q.-H., Peng, R.-H., Li, X., Fan, H.-Q., Guo, M.-J., Zhang, S.-L., 2004. Isolation, Characterization, and Molecular Cloning of the cDNA Encoding a Novel Phytase from *Aspergillus niger* 113 and High Expression in *Pichia pastoris*. BMB Rep. 37, 282–291. https://doi.org/10.5483/BMBRep.2004.37.3.282

Yancey, P.H., 2005. Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. J. Exp. Biol. 208, 2819–2830. https://doi.org/10.1242/jeb.01730

Yang, Y., Zhang, N., Ji, S., Lan, X., Zhang, K., Shen, Y., Li, F., Ni, J., 2014. *Dysgonomonas macrotermitis* sp. nov., isolated from the hindgut of a fungus-growing termite. Int. J. Syst. Evol. Microbiol. 64, 2956–2961. https://doi.org/10.1099/ijs.0.061739-0

Yanke, L.J., Bae, H.D., Selinger, L.B., Cheng, K.J., 1998. Phytase activity of anaerobic ruminal bacteria. Microbiology 144, 1565–1573. https://doi.org/10.1099/00221287-144-6-1565

Yanke, L.J., Selinger, L.B., Cheng, K. -J., 1999. Phytase activity of *Selenomonas ruminantium* : a preliminary characterization. Lett. Appl. Microbiol. 29, 20–25. https://doi.org/10.1046/j.1365-2672.1999.00568.x

Yao, M.-Z., Wang, X., Wang, W., Fu, Y.-J., Liang, A.-H., 2013. Improving the thermostability of *Escherichia coli* phytase, appA, by enhancement of glycosylation. Biotechnol. Lett. 35, 1669–1676. https://doi.org/10.1007/s10529-013-1255-x

Yao, M.Z., Zhang, Y.H., Lu, W.L., Hu, M.Q., Wang, W., Liang, A.H., 2012. Phytases: Crystal structures, protein engineering and potential biotechnological applications. J. Appl. Microbiol. 112, 1–14. https://doi.org/10.1111/j.1365-2672.2011.05181.x

Yao, Q., Li, Z., Song, Y., Wright, S.J., Guo, X., Tringe, S.G., Tfaily, M.M., Paša-Tolić, L., Hazen, T.C., Turner, B.L., Mayes, M.A., Pan, C., 2018. Community proteogenomics reveals the systemic impact of phosphorus availability on microbial functions in tropical soil. Nat. Ecol. Evol. 2, 499–509. https://doi.org/10.1038/s41559-017-0463-5

Yoon, S.J., Choi, Y.J., Min, H.K., Cho, K.K., Kim, J.W., Lee, S.C., Jung, Y.H., 1996. Isolation and identification of phytase-producing bacterium, *Enterobacter* sp. 4, and enzymatic properties of phytase enzyme. Enzyme Microb. Technol. 18, 449–454. https://doi.org/10.1016/0141-0229(95)00131-X

York, J.D., 2006. Regulation of nuclear processes by inositol polyphosphates. Biochim. Biophys. Acta. Mol. Cell Biol. Lipids., 1761(5-6), 552-559. https://doi.org/10.1016/j.bbalip.2006.04.014

Young, D.M., Parke, D., Ornston, L.N., 2005. Opportunities for genetic investigation afforded by *Acinetobacter baylyi,* a nutritionally versatile bacterial species that is highly competent for natural transformation. Annu. Rev. Microbiol. 59, 519–551. https://doi.org/10.1146/annurev.micro.59.051905.105823

Yu, H., Chin, M., Yuan, T., Bian, H., Remer, L.A., Prospero, J.M., Omar, A., Winker, D., Yang, Y., Zhang, Y., Zhang, Z., Zhao, C., 2015. The fertilizing role of African dust in the Amazon rainforest: A first multiyear assessment based on data from Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations. Geophys. Res. Lett. 42, 1984–1991. https://doi.org/10.1002/2015GL063040

Yu, P., Wang, X.-T., Liu, J.-W., 2015. Purification and characterization of a novel cold-adapted phytase from *Rhodotorula mucilaginosa* strain JMUY14 isolated from Antarctic. J. Basic Microbiol. 55, 1029–1039. https://doi.org/10.1002/jobm.201400865

Yu, S., Kvidtgaard, M.F., Isaksen, M.F., Dalsgaard, S., 2014. Characterization of a Mutant *Buttiauxella* Phytase using Phytic Acid and Phytic Acid-Protein Complex as Substrates. Anim. Sci. Lett., 1(1), 18-32.

Yuan, J.S., Reed, A., Chen, F., Stewart, C.N., 2006. Statistical analysis of real-time PCR data. BMC Bioinformatics 7(1), 85. https://doi.org/10.1186/1471-2105-7-85

Zhang, R., Yang, P., Huang, H., Shi, P., Yuan, T., Yao, B., 2011a. Two types of phytases (Histidine acid phytase and b-propeller phytase) in *serratia* sp. TN49 from the gut of *Batocera horsfieldi* (Coleoptera) Larvae. Curr. Microbiol. 63, 408–415. https://doi.org/10.1007/s00284-011-9995-0

Zhang, R., Yang, P., Huang, H., Yuan, T., Shi, P., Meng, K., Yao, B., 2011b. Molecular and biochemical characterization of a new alkaline β-propeller phytase from the insect symbiotic bacterium *Janthinobacterium* sp. TN115. Appl. Microbiol. Biotechnol. 92, 317–325. https://doi.org/10.1007/s00253-011-3309-0

Zhao, Q., Liu, H., Zhang, Ying, Zhang, Yuzhen, 2010. Engineering of protease-resistant phytase from *Penicillium* sp.: High thermal stability, low optimal temperature and pH. J. Biosci. Bioeng., 110, 638–645. https://doi.org/10.1016/j.jbiosc.2010.08.003

Zhu, J., Li, M., Whelan, M., 2018. Phosphorus activators contribute to legacy phosphorus availability in agricultural soils: A review. Sci. Total Environ., 612, 522–537. https://doi.org/10.1016/j.scitotenv.2017.08.095

Zou, X., Binkley, D., Doxtader, K.G., 1992. A new method for estimating gross phosphorus mineralization and immobilization rates in soils. Plant Soil, 147, 243–250. https://doi.org/10.1007/BF00029076