Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.
Pectin degradation and metabolism in
Monoglobus pectinilyticus 14\textsuperscript{T} from human faeces

By Caroline Chae-hyun Kim

A thesis presented in partial fulfilment of the
requirements for the degree of

Doctor of Philosophy
in
Microbiology

at Massey University, Manawatu, New Zealand

2017

Copyright © 2017 by Caroline Kim
Supervisory Committee

Dr Douglas I. Rosendale (The New Zealand Institute for Plant and Food Research)
Supervisor

Dr Mark L. Patchett (Institute of Fundamental Sciences, Massey University)
Supervisor

Dr Zoe Jordens (Institute of Fundamental Sciences, Massey University)
Supervisor

Dr William J. Kelly (Donvis Limited)
Supervisor

Professor Gerald W. Tannock (Department of Microbiology and Immunology, University of Otago)
Supervisor
Abstract

Pectin is a conspicuous plant polysaccharide, comprising one third of the dry weight of dietary fibre in common vegetables and fruit. Although pectin is almost completely digested by the human gut microbiota, few bacterial species are known to possess a comprehensive glycobiome to challenge the structurally complex pectin. The current understanding of the colonic degradation of pectin is incomplete, as the knowledge has almost exclusively derived from studying the sequestration system of *Bacteroides* spp. Here I report the isolation and characterization of *Monoglobus pectinilyticus*, and the sequencing of its genome which so far encodes the most pectin-specialized repertoire of carbohydrate active enzymes (CAZymes) found from the human gut. *M. pectinilyticus* also possesses an extracellular pectin degradation system consisting of novel protein constituents which did not find significant sequence homology and functional matches using the most up-to-date nucleotide and protein sequence databases. Proteome analysis of *M. pectinilyticus* using iTRAQ quantification revealed that pectin-degrading CAZymes and the potential constituents of the novel pectin degradation system were differentially up-regulated in response to the availability of pectin. Finally, using quantitative PCR, a positive correlation was observed between the prevalence of *M. pectinilyticus* and the consumption of fibre, vegetables, and pectin in individuals living in NZ. The discovery of *M. pectinilyticus* may add a new layer of complexity onto our interpretation of the colonic pectin degradation by presenting a system highly relevant to the pectin-rich diet of humans, and by suggesting a possibility outside the established paradigms of microbial polysaccharide degradation. The presence of *M. pectinilyticus* and the related uncultured bacteria in the gastrointestinal systems of humans and animals indicated that the organisms of this lineage are frequent terrestrial gut commensals, prompting an investigation into the genomic and molecular properties underlying their carbohydrate degradation potentials.
Acknowledgement

The support and help of many individuals are gratefully acknowledged. I would like to thank all my supervisors for providing me with the opportunity to do this research. I am deeply grateful for their mentorship, support, guidance, and optimism which were unreservedly given to me throughout the duration of my studies. I would also like to extend my sincere appreciations to all members of Food, Nutrition & Health group at Plant & Food Research, for their friendship and contributions to my works. A special thanks to Dr Ian Sims and Dr Tracey Bell for their help with respect to carbohydrate analysis. I am also grateful for the administrative assistance provided by many staffs at Plant & Food Research and Massey University. I am endlessly thankful to my parents and sisters for their encouragement, patience, and understanding throughout my studies. Finally, this work would not have been possible without the financial support of Ministry of Business, Innovation and Employment of New Zealand (‘Foods for Health at Different Life Stages’ C11X1312).
Table of Contents

Supervisory Committee ........................................................................................................ ii
Abstract ................................................................................................................................ iii
Acknowledgement ................................................................................................................ iv
Table of Contents .................................................................................................................. v
List of Figures ........................................................................................................................ ix
List of Tables .......................................................................................................................... x
List of Appendices ................................................................................................................ xi
List of copyrighted material for which permission was obtained ........................................ xii
Abbreviations ....................................................................................................................... xiii
Chapter 1 .................................................................................................................................. 1
Literature review ..................................................................................................................... 1
1.1 Plant cell wall degradation by colonic bacteria ................................................................. 1
1.2 Plant cell wall polysaccharides ......................................................................................... 3
   1.2.1 Pectin ..................................................................................................................... 3
   1.2.2 Cellulose ............................................................................................................... 7
   1.2.3 Hemicellulose ....................................................................................................... 8
1.3 Fermentability of plant cell wall polysaccharides in the human large intestine ................ 10
1.4 Microbial strategies of complex polysaccharide degradation ........................................... 12
   1.4.1 The cellulosome system ....................................................................................... 12
   1.4.2 The polysaccharide utilization locus (PULs) ......................................................... 16
1.5 CAZy database and CAZymes ....................................................................................... 19
1.6 Pectin-degrading enzymes ............................................................................................. 21
   1.6.1 Homogalacturonan (HG)-degrading enzymes ...................................................... 22
   1.6.2 Rhamnogalacturonan-I (RG-I)-degrading enzymes ............................................. 29
   1.6.3 Rhamnogalacturonan-II (RG-II)-degrading enzymes ........................................... 34
   1.6.4 Pectin-associated carbohydrate binding modules (CBMs) ................................... 35
1.7 Pectin metabolism ............................................................................................................ 38
1.8 Pectin degradation in the human large intestine ............................................................... 39
1.9 Rationale of the research ............................................................................................... 41
1.10 Aims ............................................................................................................................... 42
Chapter 2  Materials and Methods ....................................................................................... 44
2.1 General methods .............................................................................................................. 44
   2.1.1 Media and additives .............................................................................................. 44
   2.1.2 Basal nutrient (BN) medium ............................................................................... 44
   2.1.3 Basal nutrient (BN) agar medium prepared in roll-tubes ..................................... 45
   2.1.4 Mineral medium ................................................................................................... 45
Chapter 3   Isolation and characterization

3.1 Introduction .......................................................... 83
3.2 Constituent sugar analysis of pectin extracts from kiwifruit ............... 84
  3.2.1 Soluble and insoluble kiwifruit fractions .................................. 84
  3.2.2 Fractionation of branched pectin from insoluble kiwifruit materials ....... 85
3.3 Isolation and selection of strain 14\textsuperscript{T} ........................................ 86
3.4 Morphological descriptions of strain 14\textsuperscript{T} .............................. 87
  3.4.1 TEM examination ..................................................... 87
  3.4.2 Endospore formation .................................................. 89
Chapter 4   Genomic overview of Monoglobus pectinilyticus strain 14T ................................. 101
4.1 Introduction .................................................................................................................... 101
4.2 de novo genome assembly and primer walking ............................................................... 101
4.3 Assessing the genome completeness ............................................................................. 104
4.4 Genome-based phylogenetic analysis ........................................................................... 105
4.5 Genome organization and general features ................................................................. 107
4.6 Extraneous DNA ........................................................................................................... 111
        4.6.1 Prophage genomes ............................................................................................... 111
        4.6.2 Horizontal gene transfer (HGT) ........................................................................... 113
4.7 CRISPR ......................................................................................................................... 114
4.8 The glycobiome of M. pectinilyticus ............................................................................ 115
4.9 S-layer homology domain-containing proteins ............................................................. 122
4.10 Metabolic capacity of M. pectinilyticus ..................................................................... 126
4.11 Organization of pectin gene clusters .......................................................................... 129
4.12 Endospore formation, cofactor synthesis and motility ................................................... 131
Chapter 5   Proteomic profiling of Monoglobus pectinilyticus using iTRAQ .......................... 133
5.1 Introduction .................................................................................................................... 133
5.2 iTRAQ protein identification ......................................................................................... 135
5.3 Differential expression of pellet proteins identified using iTRAQ ................................ 136
5.4 Differential expression of supernatant proteins identified using iTRAQ ...................... 138
Chapter 6   Ecological abundance of Monoglobus pectinilyticus in the human gut .............. 141
6.1 Introduction .................................................................................................................... 141
6.2 Quantification of M. pectinilyticus by qPCR ................................................................. 141
        6.2.1 Primer design and specificity testing .................................................................... 141
        6.2.2 Construction of qPCR standard curves ............................................................... 142
        6.2.3 Participant recruitment and dietary composition analysis .................................... 143
        6.2.4 The relative abundance of M. pectinilyticus in stool DNA samples .................. 143
6.3 The presence of M. pectinilyticus in the metagenomic datasets .................................... 147
Chapter 7 Discussions .......................................................................................................... 149
7.1 Isolation of a novel pectin-degrading bacterium .......................................................... 149
7.2 Genomic characterization suggests *M. pectinilyticus* is a pectin-degrading specialist ..........153
7.3 Proteomics supports putative extracellular pectin degradation ........................................161
7.4 Ecological presence of *M. pectinilyticus* in human colonic microbiota .............................167
7.5 A proposed model for pectin degradation by *M. pectinilyticus* ..................................169
7.6 Conclusions ........................................................................................................................171
7.7 Future directions ..................................................................................................................171
References ..................................................................................................................................174
Appendices ..................................................................................................................................200
List of Figures

Figure 1.1 Plant cell wall structure .................................................................................. 1
Figure 1.2 Pectin structure .............................................................................................. 4
Figure 1.3 Cellulose structure .......................................................................................... 7
Figure 1.4 Hemicelluloses structure .................................................................................. 9
Figure 1.5 Schematic diagram of cellulosome system in Ruminococcus flavefaciens .... 13
Figure 1.6 Schematic diagram of starch utilization system of Bacteroides thetaiotaomicron ... 17
Figure 1.7 Activities of pectin-degrading enzymes on HG, RG-I, and RG-II ................. 24
Figure 1.8 Pectin degradation pathway characterized in Dickeya dadantii ..................... 38
Figure 3.1 Micrographs of strain 147 ............................................................................. 88
Figure 3.2 Neighbor-joining phylogenetic tree of strain 147 ......................................... 90
Figure 3.3 Phylogenetic tree of strain 147 and the closest 16S rRNA sequence relatives ... 93
Figure 3.4 Growth curves of strain 147 growing at different temperatures .................... 94
Figure 3.5 Growth curves of strain 147 over an acidity range of pH 5.0 – pH 9.0 .......... 95
Figure 3.6 Organic acid production from pectins ......................................................... 96
Figure 3.7 Organic acid production from monomeric sugars ........................................ 97
Figure 3.8 Size exclusion chromatography .................................................................... 99
Figure 4.1 Genome sequencing gaps ............................................................................. 102
Figure 4.2 Primer walking procedures for closing the genome sequencing gaps .......... 103
Figure 4.3 Matrix diagram comparing bacterial genome and 16S rRNA gene sequence similarities ........................................................................................................ 106
Figure 4.4 A genome map of M. pectinilyticus chromosome ......................................... 109
Figure 4.5 Prophage-associated gene clusters present within the M. pectinilyticus genome .... 112
Figure 4.6 CRISPR-Cas systems within the M. pectinilyticus genome ............................ 114
Figure 4.7 Comparison of PLs, GHs, and CEs with pectin-specific activities between M. pectinilyticus and other pectin-degrading strains ......................................................... 117
Figure 4.8 Comparison of the number of pectin-specific CAZymes in relative to the total number of CAZymes ..................................................................................................... 118
Figure 4.9 PL1 and CE8 domain sequences of M. pectinilyticus forming discrete species-specific clusters ........................................................................................................ 121
Figure 4.10 The 42 SLH module-containing proteins of M. pectinilyticus ....................... 124
Figure 4.11 Percentage identity matrix of SLH proteins of M. pectinilyticus .................. 125
Figure 4.12 Reconstruction of sugar metabolism by M. pectinilyticus ............................ 127
Figure 4.13 Gene clusters concerning the pectin degradation and utilization .................. 130
Figure 5.1 Workflow diagram of iTRAQ quantitative proteomics .................................. 134
Figure 5.2 Protein ratios between M. pectinilyticus grown on fructose control and pectins ... 137
Figure 6.1 Quantitative PCR comparison of log_{10} concentration of M. pectinilyticus and the consumption of different food categories ..................................................................... 146
Figure 7.1 Simplified diagram of sugar utilization and organic acid/gas production .......... 160
Figure 7.2 A proposed model for pectin degradation by M. pectinilyticus ....................... 170
List of Tables

Table 2.1 List of bioinformatics tools used in this study. .................................................................69
Table 2.2 List of long-range PCR primers used to amplify genome gaps...............................70
Table 2.3 List of microbial genomes used to construct user-specified Prokka databases.........72
Table 2.4 Sequences of universal and M. pectinilyticus-specific PCR primers used in this study 79
Table 3.1 Sugar composition of insoluble and soluble fractions from kiwifruit materials ..........84
Table 3.2 Sugar compositions of pectic oligosaccharides extracted from digested kiwifruit ......86
Table 3.3 Differential phenotypic characteristics between strain 14T and related species.........98
Table 4.1 Key genomic properties of M. pectinilyticus.................................................................108
Table 6.1 qPCR comparison of M. pectinilyticus and the total bacterial populations............144
Table 7.1 Bacteria that produce CAZyme domain-containing proteins with SLH modules ......157
List of Appendices

Appendix 1 CAZymes involved in the degradation of HG, RG-I, and RG-II ............................................. 201
Appendix 2 CBM domains associated with binding the pectic polysaccharides or the monomeric components of pectin .................................................................................................................. 205
Appendix 3 Pectin-degrading bacteria reported from the human gut ................................................................. 207
Appendix 4 Uncultured bacterial 16S rRNA gene sequences from GenBank with ≥ 92 % sequence similarity to strain 14T .................................................................................................................... 208
Appendix 5 M. pectinilyticus 14T CDSs encoding putative enzymes and CBMs involved in polysaccharide degradation/modification .............................................................................................................. 210
Appendix 6 Identifiable protein domain architectures in plant carbohydrate-associated PLs, GHs, and CEs of M. pectinilyticus ............................................................................................................................ 214
Appendix 7 Extended phylogenetic trees of PL1 and CE8 catalytic domain sequences ........................................ 223
Appendix 8. EC numbers and enzyme names associated with Figure 4.12 ........................................................... 224
Appendix 9 Sporulation-related genes of M. pectinilyticus ...................................................................................... 225
Appendix 10 Geometric distribution of z-scores for individual protein expression ratios ...................................... 227
Appendix 11 iTRAQ proteomics expression ratios between apple pectin-grown M. pectinilyticus cells and fructose-grown cells ............................................................................................................................................. 228
Appendix 12 iTRAQ proteomics expression ratios between citrus pectin-grown M. pectinilyticus cells and fructose-grown cells ............................................................................................................................................. 234
Appendix 13 iTRAQ proteomics expression ratios between kiwifruit pectin-grown M. pectinilyticus cells and fructose-grown cells ............................................................................................................................................. 240
Appendix 14 iTRAQ proteomics expression ratios across nine replicates from all pectins combined ........................................................ ............................................................................................................................................. 246
Appendix 15 iTRAQ expression ratios in M. pectinilyticus extraproteomes .................................................................. 252
Appendix 16 Testing the target specificity of M. pectinilyticus-specific qPCR primers ............................................... 255
Appendix 17 Enumeration of cell numbers of M. pectinilyticus (V6-V8) and E. coli (V3-V4) by viable cell count and qPCR ............................................................................................................................................. 256
Appendix 18 Quantitative PCR standard curves constructed by plotting log_{10} cell numbers of M. pectinilyticus and E. coli against the crossing points (CP) ..................................................................................... 257
Appendix 19 Participant information and the daily food group intake calculated from four sets of 3-day dietary records ............................................................................................................................................. 258
Appendix 20 Determining the split point for separating M. pectinilyticus-positive and M. pectinilyticus-negative groups ............................................................................................................................................. 260
Appendix 21 SLH protein sequences of M. pectinilyticus identified in the NIH human stool metagenome databases ............................................................................................................................................. 261
Appendix 22 US nationwide dietary intake data for the years 2003 - 2006 .................................................................... 267
List of copyrighted material for which permission was obtained

<table>
<thead>
<tr>
<th>Material</th>
<th>Source</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 1.1</td>
<td>Flint <em>et al.</em>, 2012. Gut Microbes. 3, 289-306.</td>
<td>No permission required</td>
</tr>
<tr>
<td>Figure 1.2</td>
<td>Mohnen, 2008. Current Opinion in Plant Biology. 11, 266-277.</td>
<td>Permission granted</td>
</tr>
<tr>
<td>Figure 1.3</td>
<td>Simoneit <em>et al.</em>, 1999. Atmospheric Environment. 33, 173-182.</td>
<td>Permission granted</td>
</tr>
<tr>
<td>Figure 1.4</td>
<td>Scheller and Ulvskov, 2010. Annual Review of Plant Biology. 61, 263-89</td>
<td>No permission required</td>
</tr>
<tr>
<td>Figure 1.5</td>
<td>Rincon <em>et al.</em>, 2007. Journal of Bacteriology. 189, 4774-4783.</td>
<td>No permission required</td>
</tr>
<tr>
<td>Figure 1.6</td>
<td>Martens and Koropatkin, 2009. Journal of Biological Chemistry. 284, 24673-24677.</td>
<td>No permission required</td>
</tr>
<tr>
<td>Figure 1.7</td>
<td>Ndeh <em>et al.</em>, 2017. Nature. 544, 65-70.</td>
<td>Permission granted</td>
</tr>
<tr>
<td>Figure 1.8</td>
<td>Condemine and Robert-Baudouy, 1987. Journal of Bacteriology. 169, 1972-1978.</td>
<td>No permission required</td>
</tr>
<tr>
<td>Appendix 21</td>
<td>National Health and Nutrition Examination Survey (NHANES), 2003-2006.</td>
<td>No permission required</td>
</tr>
</tbody>
</table>
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>2-O-Me-Fuc</td>
<td>2-O-methyl L-fucose</td>
</tr>
<tr>
<td>2-O-Me-Xyl</td>
<td>2-O-methyl-xylose</td>
</tr>
<tr>
<td>AceA</td>
<td>L-aceric acid</td>
</tr>
<tr>
<td>Api</td>
<td>D-apiose</td>
</tr>
<tr>
<td>AraF</td>
<td>arabinofuranose</td>
</tr>
<tr>
<td>AraP</td>
<td>arabinopyranose</td>
</tr>
<tr>
<td>BN</td>
<td>basal nutrient medium</td>
</tr>
<tr>
<td>CAZyme</td>
<td>carbohydrate active enzyme</td>
</tr>
<tr>
<td>CBM</td>
<td>carbohydrate-binding module</td>
</tr>
<tr>
<td>CDS</td>
<td>coding sequence</td>
</tr>
<tr>
<td>CE</td>
<td>carbohydrate esterase</td>
</tr>
<tr>
<td>COG</td>
<td>Cluster of orthologous groups</td>
</tr>
<tr>
<td>CRISPR</td>
<td>Clustered regularly interspaced short palindromic repeats</td>
</tr>
<tr>
<td>DDH</td>
<td>DNA-DNA hybridization</td>
</tr>
<tr>
<td>DE</td>
<td>degree of esterification</td>
</tr>
<tr>
<td>DhaA</td>
<td>3-deoxy-D-lyxo-heptulosaric acid</td>
</tr>
<tr>
<td>DSMZ</td>
<td>German Collection of Microorganisms and Cell Cultures</td>
</tr>
<tr>
<td>Fru</td>
<td>fructose</td>
</tr>
<tr>
<td>Fucp</td>
<td>fucose pyranose</td>
</tr>
<tr>
<td>Galp</td>
<td>galactopyranose</td>
</tr>
<tr>
<td>GalpA</td>
<td>galacturonic acid</td>
</tr>
<tr>
<td>GH</td>
<td>glycoside hydrolase</td>
</tr>
<tr>
<td>GI</td>
<td>gastrointestinal tract</td>
</tr>
<tr>
<td>GlcA</td>
<td>glucuronic acid</td>
</tr>
<tr>
<td>Glu</td>
<td>glucose</td>
</tr>
<tr>
<td>GT</td>
<td>glycosyltransferase</td>
</tr>
<tr>
<td>HG</td>
<td>homogalacturonan</td>
</tr>
<tr>
<td>HGT</td>
<td>horizontal gene transfer</td>
</tr>
<tr>
<td>HMM</td>
<td>Hidden Markov model</td>
</tr>
<tr>
<td>HMP</td>
<td>Human Microbiome Project</td>
</tr>
<tr>
<td>HPAEC-PAD</td>
<td>High-Performance Anion-Exchange Chromatography Coupled with Pulsed Amperometric Detection</td>
</tr>
<tr>
<td>HSP</td>
<td>high-scoring segment pairs</td>
</tr>
<tr>
<td>IJSEM</td>
<td>International Journal of Systematic and Evolutionary Microbiology</td>
</tr>
<tr>
<td>JCM</td>
<td>Japan Culture Collection</td>
</tr>
<tr>
<td>kdoA</td>
<td>2-keto-3-deoxy-D-manno-octulosonic acid</td>
</tr>
<tr>
<td>KEGG</td>
<td>Kyoto Encyclopedia of Genes and Genomes</td>
</tr>
<tr>
<td>LB</td>
<td>Luria-Bertani medium</td>
</tr>
<tr>
<td>LPS</td>
<td>lipoprotein signal peptides</td>
</tr>
<tr>
<td>Man</td>
<td>mannose</td>
</tr>
<tr>
<td>ML</td>
<td>middle lamella</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>ORF</td>
<td>open reading frame</td>
</tr>
<tr>
<td>Ori</td>
<td>origin of replication</td>
</tr>
<tr>
<td>PBS</td>
<td>phosphate buffered saline</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>PCW</td>
<td>plant cell wall</td>
</tr>
<tr>
<td>PL</td>
<td>polysaccharide lyase</td>
</tr>
<tr>
<td>POCP</td>
<td>the percentage of conserved proteins</td>
</tr>
<tr>
<td>PUL</td>
<td>polysaccharide utilization locus</td>
</tr>
<tr>
<td>RC</td>
<td>reinforced clostridial medium</td>
</tr>
<tr>
<td>RG</td>
<td>rhamnogalacturonan</td>
</tr>
<tr>
<td>Rhap</td>
<td>rhamnopyranose</td>
</tr>
<tr>
<td>SLH</td>
<td>S-layer homology</td>
</tr>
<tr>
<td>Sus</td>
<td>starch utilization system or sequestration system</td>
</tr>
<tr>
<td>TEM</td>
<td>transmission electron microscope</td>
</tr>
<tr>
<td>XGA</td>
<td>xylogalacturonan</td>
</tr>
<tr>
<td>Xylp</td>
<td>xylopyranose</td>
</tr>
</tbody>
</table>