The auxiliary replicons of *Butyrivibrio proteoclasticus*

A Thesis presented in fulfilment of the Doctorate of Philosophy degree at Massey University, Palmerston North, New Zealand.

Carl Yeoman
2009

Table of Contents

List of Figures

List of Tables

Abstract

Butyrivibrio proteoclasticus B316^T is the most recently described species of the *Butyrivibrio* / *Pseudobutyrivibrio* assemblage and now the first to have its genome sequenced. The genome of this organism was found to be spread across four replicons: a 3.5 Mb major chromosome and three additional large replicons: 186, 302 and 361 Kb in size. This thesis describes the sequencing, analysis, annotation and initial characterisation of all three *B. proteoclasticus* auxiliary replicons. Most significantly, these analyses revealed that the 302-Kb replicon is a second chromosome. This small chromosome, named BPc2, encodes essential systems for the uptake and/or biosynthesis of biotin and nicotinamide adenine mononucleotide, as well as the enzymes required for utilisation of fumarate as the terminal electron acceptor during anaerobic respiration, none of which are found on the main chromosome. In addition, BPc2 contains two complete rRNA operons, a large number of enzymes involved in the metabolism of carbohydrates, nitrogen and fatty acids. In contrast to BPc2, both megaplasmids appear largely cryptic, collectively encoding 421 genes not previously described in public databases. Nevertheless, only the 186-Kb, but not 361-Kb megaplasmid, could be cured from *Butyrivibrio proteoclasticus* B316^T. The largest megaplasmid has a copy number of 5, while all other replicons are present at a copy number of 1. %GC content and codon usage analyses strongly suggests that all three auxiliary replicons have co-resided with the major chromosome for a significant evolutionary period. Moreover, the replication machineries of these three replicons are conserved. Interestingly, a survey of a number of *Butyrivibrio* / *Pseudobutyrivibrio* species revealed that the megaplasmids are widespread in this assemblage, however these other large plasmids do not show concordance with their 16*S* rRNA phylogeny and appear distinct to those of *B. proteoclasticus* B316^T.

A microarray analysis of gene expression in a co-culture experiment between *B. proteoclasticus* and the important ruminal methanogen, *Methanobrevibacter ruminantium* M1, revealed a potentially mutualistic interspecies interaction. In this relationship *M. ruminantium* appears to provide *B. proteoclasticus* with glutamate, essential to the final step of NAD^+ biosynthesis, while *B. proteoclasticus* appears to provide *M. ruminantium* with formate, hydrogen and carbon dioxide, each important substrate for methanogenesis.

Acknowledgements

It is a privilege and an honour to acknowledge the many great people who have guided me throughout this project in one way or another. But more importantly have helped to shape me into the person that I am today.

To my supervisors Dr Graeme Attwood, Dr William (Bill) Kelly, and Dr Jasna Rakonjac, it is impossible to put into words the sincere gratitude I have for you all. I certainly wouldn't have lasted long without your influence and direction.

To the great people of AgResearch for the practical and financial support, without either this project would not have been possible.

To the kind people of Meat and Wool New Zealand, in particular Alan Frazer, who have funded, essentially, my life throughout not only this Ph.D. but the preceding MSc work.

To the Tertiary Education Commission for their financial support with this Ph.D.

To Dr Keith Joblin, I am eternally greatful to you and Robert Skipp who collectively played a role in bringing me into the AgResearch family and in doing so keeping me in the field of science.

To the Rumen microbiology laboratory, particularly, Dr Sinead Leahy for your help with the assembly, Dr Eric Altermann for your help with all things computational, Zhan-hou Kong and Sam Taylor for your help with the microarrays, Dong Lee for being my go to guy, Diana Pacheco for being my go to gal, Dr Adrian Cookson, Dr Ron Ronimus, Dr Peter Janssen, Graham Naylor, Dr Christina Moon, Carrie Sang, Rechelle Perry, Gemma Henderson, Catherine Tootill, and the ex-pats Dr Lucy Skillman, Dr Nicola Walker, Dr Mathew Nicholson, Dr Karen Olsen, Dr Hassan Husein, Paul Evans, Sam Noel and Nikki Kenters for your friendship.

To my parents Trish and Jim Yeoman thanks for all your support and encouragement.

To my beautiful wife Casey Norris-Yeoman thanks for sticking by me and supporting me while I pursue my dreams it has been a long and tough road but hopefully will be worth it in the long run. Love always!

Dedication

This thesis is dedicated to the most beautiful girls in the world my daughters Summer Ashlee Pamela Yeoman and Sienna Caitlyn Estelle Yeoman.

Whatever you need, whenever you need it, I'll always be there for you both!!

Abbreviations

AWGS	Alan Wilson Centre Genome Sequencing
BAC	Bacterial Artificial chromosome
BER	BLAST-extend-repraze
BSA	Bovine Serum Albumin
bp	Base pair
CDS	Coding sequence
Contigs	Contiguous sequences
DR	Direct repeat
dso	Double-stranded origin
FDR	False discovery rate
g	Gravity
GH	Glycosyl hydrolase
HMM	Hidden Markov-model
IR	Inverted repeat
IVR	Inverse repeat
Kb	Kilobase pair
Mb	Megabase pair
Mpf	Mating pair formation complex
nt	Nucleotide(s)
OD	Optical Density
ORF	Open reading frame
PARP	Poly(ADP-ribose) polymerase
PCB	Polychlorinated biphenyl
PCR	Polymerase chain reaction

PFGE	Pulsed-field gel-electrophoresis
pI	Isoelectric point
polIII	DNA polymerase III
RC	Rolling circle
RE	Restriction endonuclease
rRNA	Ribosomal ribonucleic acid
sso	Single-stranded origin
TA	Toxin-Antitoxin
tRNA	Transfer ribonucleic acid
qPCR	Quantitative real-time PCR