Comparative genomics of rumen methanogens

A thesis presented in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biochemistry

At Massey University, Palmerston North, New Zealand

Yang Li

2016

Abstract

Methane (CH_4) emissions from agriculture represent around 9% of global anthropogenic greenhouse gas emissions. The single largest source of this CH_4 is animal enteric fermentation, predominantly from ruminant livestock, where it is produced mainly in their fermentative forestomach (or reticulo-rumen) by a group of archaea known as methanogens.

In order to reduce CH_4 emissions from ruminants, it is necessary to understand the role of methanogenic archaea in the rumen, and to identify their distinguishing characteristics that can be used to develop CH_4 mitigation technologies. To gain insights into the role of methanogens in the rumen environment, two methanogens have been isolated from ovine rumen and their genomes were sequenced: methanogenic archaeon ISO4-H5 of the order Methanomassiliicoccales and *Methanobrevibacter* sp. D5 of *Methanobrevibacter gottschalkii* clade.

Genomic analysis suggests ISO4-H5 is an obligate hydrogen-dependent methylotrophic methanogen, able to use methanol and methylamines as substrates for methanogenesis. Like other organisms within this order, ISO4-H5 does not possess genes required for the first six steps of hydrogenotrophic methanogenesis. Comparison between the genomes of different members of the order Methanomassiliicoccales revealed strong conservation in energy metabolism, particularly in genes of the methylotrophic methanogenesis pathway, as well as in the biosynthesis and use of pyrrolysine. Unlike members of Methanomassiliicoccales from human sources, ISO4-H5 does not contain the genes required for production of coenzyme M (CoM), and requires external supply of CoM to survive.

Methanobrevibacter sp. D5 is a hydrogenotrophic methanogen predicted to utilise $CO_2 + H_2$ and formate as substrates. Comparisons between the available *Methanobrevibacter* genomes has revealed a high conservation in energy metabolism and characteristics specific to each clade. The coexistence of different *Methanobrevibacter* species in the rumen may be partly due to the physical association *Methanobrevibacter* species with different microorganisms and host surface, which allow unique niches to be established.

Acknowledgements

It was my privilege to work under the guidance of Dr. Graeme Attwood, Dr. Sinead Leahy and Associate Professor Jasna Rakonjac, I am grateful for their contant guidance and unflinching support, they are a source of inspiration and motivation. This thesis would not be possible without them.

I am grateful to the New Zealand Agricultural Greenhouse Gas Research Centre and Massey University for the financial support and the opportunity they gave me to conduct a truly meaningful research project contributing towards CH_4 mitigation. I would also like to thank AgResearch Ltd. for providing the excellent facility for me to conduct my research at the Grasslands Research Centre. My gratitude also to the Pastoral Greenhouse Gas Research Consortium (PGgRc) for kindly providing the methanogen cultures.

The experience and skills I have learned in my time here are invaluable, but the best thing that happened during my PhD is finding my better, wiser half, Jana Muller, who is also a PhD student here in AgResearch. She and my family provided me the strength to finish this long journey, which often seems never ending.

My project would not be possible without the kind assistance and continued support offered by all the nice people of our rumen microbiology team, my fellow PhD students and the experts of Massey University. I would like to thank Dr. Bill Kelly for teaching me Pulsed Field Gel Electrophoresis and the scientific insights he offered, Dr. Eric Altermann for his expertise in bioinformatics and insights into prophage, Dr Peter Janssen for his knowledge in methanogen biology and insights into bioenergetics, Dr Gemma Henderson for kindly providing the source culture H6, and her scientific insights. I would like to extend my thanks to my predecessor, Dr Jeyamalar Jeyanathan who I never met, for her excellent work on Rumen Cluster C enrichment and kindly providing the Rumen Cluster C cultures, Dr. Suzanne Lambie for teaching me how to extract genomic DNA from methanogens, Dong Li for his scientific insights and assistance in trouble shooting molecular biology problems, Dr. Janine Kamke and Priya Soni for working together on the transcriptome project, Caroline Kim for her experience in *Methanobrevibacter* sp. D5 isolation, Dr Siva Ganesh for his expertise on statistical analysis and Rstudio, Debjit Dey for teaching me techniques on anaerobic cultivation, Dr Linda Samuelson and Dr. Patrick Edwards for their assistance and expertise on ^{1}H- nuclear magnetic resonance (NMR) and metabolomics, Douglas Hopcroft and Jordan Taylor of Manawatu Microsocopy and Imaging

Centre (MMIC) for assistance in transmission electron microscopy, Dr. Linley Schofield, Dr. Dragana Gagic, Dr Sandra Kittleman, Dr. Henning Seedorf and Dr. Ron Ronimus for their scientific insights, Kerri Reilly and Savannah Devente for the singing and humming in the lab, Carrie Sang for keeping lab consumables supplied, Dr. Kenneth Teh for the his insights on the taxonomic relationship of PhD students and zombies, Dr. Sonal Shewaramani for sharing her experience in RNA extraction and genome assembly, and special thanks to Faith Cox for her friendship, teaching me techniques on anaerobic cultivation, purification of *Succinivibrio dextrinisolvens* H5 and working together on Rumen Cluster C. Special thanks to Dr. Nick Palevich and Dr. Preeti Raju for their friendship, being great office mates and supporting each other throughout our PhD. I would also like to thank my fellow PhD students for their friendship and support, and Dr. Sunny Sun and his family for their friendship and support.

Table of contents

List of Tables

Page

List of Figures

Page

Abbreviations

Non-standard abbreviations:

aa	Amino acids (length of peptide chain or sequence identity)
ANOVA	Analysis of variance
ATP	Adenosine triphosphate
BES	2-bromoethanesulfonic acid
BLAST	Basic Local Alignment Search Tool
BLOSUM	BLOcks SUbstitution Matrix
BRIG	BLAST Ring Image Generator
CAI	Codon adaptation index
CDS	Coding DNA sequence
CH_4	Methane
CO	Carbon monoxide
CO_2	Carbon dioxide
CoA	Coenzyme A
CoB	Coenzyme B
COG	Clusters of Orthologous Groups
CoM	Coenzyme M
CRISPR	Clustered regularly interspaced short palindromic repeat
D_2O	Deuterated water
DEPC	Diethylpyrocarbonate
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
$F_{390/420/430}$	Cofactor $F_{390/420/430}$
FDR	False discovery rate
Fdx	Ferredoxin
FGD	Functional genome distribution
GHGs	Greenhouse gas
GIT	Gastrointestinal
H_4MPT	Tetrahydromethanopterin
HMM	Hidden Markov model
HSQC	Heteronuclear Single Quantum Coherence Spectroscopy
IS	Insertion sequence

IVOM	Interpolated variable ordered motif
KEGG	Kyoto Encyclopedia of Genes and Genomes
KW	Kruskal-Wallis rank sum test
M3MSP	Methyl-3-methylthiopropionate
M3SP	Methylmercaptopropionate
MCL	Maximum Composite Likelihood
MDS	Multidimensional scaling
MF	Methanofuran
MMIC	Manawatu Microscopy and Imaging Centre
mRNA	Messenger RNA
N_2	Nitrogen (gas or liquid)
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate
NCBI	National Center for Biotechnology Information
NMR	Nuclear magnetic resonance
NRPS	Non-ribosomal peptide synthase
NZ	New Zealand
O_2	Oxygen
ORB	Origin recognition box
ORF	Open reading frame
PCoA	Principal coordinate analysis
PCR	Polymerase chain reaction
PFGE	Pulsed field gel electrophoresis
RCC	Rumen Cluster C
RNA	Ribonucleic acid
snRNA	Small nuclear RNA
snoRNA	Small nucleolar RNA
SSPGMS	*Succinivibrio* spent pectin growth media supernatant
TAE	Tris acetate EDTA
TBE	Tris borate EDTA
TE	Tris EDTA
TEM	Transmission electron micrograph
TMH	Transmembrane helix
TOCSY	Total Correlation Spectroscopy

tRNA	Transfer RNA
UPGMA	Unweighted pair group method with arithmetic mean
UV	Ultra violet
VFA	Volatile fatty acid

Measurement Units:

^{o}C	Degree Celsius
µg	Microgram
µL	Microlitre
µm	Micrometer
µM	Micromolar
bp	Base pair
h	Hour
kcal	kilocalorie
kb	kilobase
kDa	kilodaltons
kpa	kilopascal
kV	kilovolts
L	Litre
M	Molar
Mb	Million base pairs
mg	Milligram
MHz	Mega hertz
min	Minutes
mL	Millilitre
mm	Millimeter
mM	Millimolar
mV	Millivolts
ng	Nanogram
nm	Nanometer
ppm	Parts per million
s	Seconds
v/v	Volume/volume
w/v	Weight/volume