

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.

Non-protein-coding-RNA processing in
the deep-branching protozoan parasite
Giardia intestinalis

A thesis presented in partial fulfilment of the requirements
for the degree of PhD in Molecular Genetics

at Massey University, Palmerston North
New Zealand

Xiaowei (Sylvia) Chen

2008

Acknowledgement

It has been a challenging but rewarding passage towards the completion of this thesis. I am grateful to all the people who helped during the work.

At first I would like to express my sincere gratitude to my supervisor Prof. David Penny, co-director of Allan Wilson Centre, Massey University. His wide knowledge and inspiring ideas have been great value for all his students, including me. His excellent guidance provided a good basis for this work.

I am grateful to my co-supervisor Dr. Lesley J. Collins, Allan Wilson Centre, Massey University for her continuous support and detailed comments throughout this work. Her understanding and encouragement are valuable for me, and also it has been very enjoyable to share the same office with her.

I wish to express my warm thanks to my co-supervisor Prof. Jürgen Brosius, ZMBE, University of Münster, Germany, who kindly provided me the opportunity to work with them in Münster, and to learn the valuable skills required for completion of my work.

I also would like to express my gratitude to all the lab people who helped me during different periods of my work. In particular, I am grateful to our lab manager Trish McLenachan. Her warm personality and enthusiasm have been a great help to my smooth progress. Also it has been a great time to work with the people in Münster. I wish to thank Associate Prof. Jürgen Schmitz for his friendly support, Dr. Timofey S. Rozhdestvensky and Anja Zemann for their great help of constructing the RNA library.

For the excellent help on computational side of the work, I would like to express my special thanks to my good friend Tim White. He has provided high quality technical assistance as well as great entertainment during this work.

Also I would like to thank Dr. George Ionas and Errol Kuan in MicroAquaTech, Massey University for kindly providing *Giardia* cell culture. My sincere thanks are due to Dr. Kathryn Stowell for her kind instruction on radioisotope techniques, and to Dr. Gill Norris for her instruction on protein analysis.

For the critical and useful suggestions which gave me more thoughts about this work during the thesis examination, I would like to thank my examiners, Dr. Austen Ganley, Dr. Peter K. Dearden and Dr. Paul P. Gardner.

I owe my most sincere gratitude to my Mum and Dad for their love and support throughout my life. It is always a pure joy to see them every year.

Finally, thanks to all the people in Allan Wilson Centre. Life has been great here.

This work is supported by the New Zealand Marsden fund; the Allan Wilson Centre for Molecular Ecology and Evolution; European Union (EU: LSHG-CT-2003-503022 and the Nationales Genomforschungsnetz (NGFN: 0313358A).

Content

	Page
Introduction to the thesis	1
 Chapter 1: Evolution of non-protein-coding RNAs	
1.1 The origin of non-protein-coding RNAs	13
1.2 Relics from the RNA-World	
1.2.1 The history of universal house-keeping RNAs	18
1.2.2 The antiquity of catalytic RNAs	
1.2.2.1 Small catalytic RNAs	23
1.2.2.2 Self-splicing introns	25
1.2.2.3 RNase P and MRP	28
1.2.3 Ancient RNA regulators – Riboswitches	31
1.3 Evolution of ncRNAs in eukaryotes	
1.3.1 Introns and their roles in eukaryotic evolution	33
1.3.2 Divergence of regulatory RNAs	
1.3.2.1 RNAs in the nucleus – small nuclear RNAs	35
1.3.2.2 RNA editing and small nucleolar RNAs (snoRNAs)	39
1.3.2.3 Enrichment of ncRNAs in eukaryotes and systematic gene regulation	42
1.4 The present and future of ncRNA evolution	44
 Chapter 2: Identification of novel non-protein-coding RNAs from <i>Giardia intestinalis</i>	
Combined experimental and computational approach to identify non-protein-RNAs in the deep-branching eukaryote <i>Giardia intestinalis</i> <i>Nucleic Acids Research</i> 35 (14): 4619-4628.....	48

Chapter 3: Analysis of the ncRNA library of *Giardia intestinalis*

3.1 Background: Molecular biology of *Giardia intestinalis* and techniques in the studies of ncRNA

3.1.1 <i>Giardia</i> – a deep branching unicellular eukaryote	59
3.1.2 The genome of <i>Giardia</i>	62
3.1.3 Techniques in the studies of ncRNAs.....	64

3.2 Analysis of the novel ncRNA candidates

3.2.1 Novel ncRNA candidates from the cDNA library of <i>Giardia</i>	69
3.2.2 Analysis of potential promoter elements of characterized ncRNAs from <i>Giardia</i>	71
3.2.3 Analysis of the upstream sequence elements and internal sequence elements of novel <i>Giardia</i> ncRNAs	83
3.2.4 Polymerase III transcription factors of <i>Giardia</i>	88
3.2.5 Structural analysis of uncharacterized novel ncRNA candidates	92

3.3 Conclusion

97

Chapter 4: Studies of the major spliceosomal snRNAs in *Giardia*

4.1 Introduction: Does *Giardia* have a functional spliceosome?

99

4.2 Searching for U-snRNAs in *Giardia*

4.2.1 Prediction of <i>Giardia</i> U1-snRNA candidate	102
4.2.2 Prediction of <i>Giardia</i> U2-snRNA candidate	106
4.2.3 Prediction of <i>Giardia</i> U6 and U4 snRNA candidates	110

4.3 *Giardia* homologue of Prp8 protein – the central protein component of the spliceosome

118

4.3.1 Bioinformatical analysis of the <i>Giardia</i> homologue of Prp8 protein	119
---	-----

4.3.2 Analysis of the potential RNA-recognition motif of *Giardia* Prp8 protein

4.3.2.1 Computational analysis of the likelihood of Gp8d1	
---	--

being a protein domain	122
4.3.2.2 Cloning and <i>in vitro</i> recombinant protein expression	126
4.3.2.3 RNA-protein binding assays of Gp8d1 <i>versus Giardia</i> ncRNA candidates	127
4.4 Conclusion and overview of the major spliceosomal components in <i>Giardia</i>	131
4.5 Experimental materials and methods	
4.5.1 PCR amplification and cloning	132
4.5.2 Plasmid preparation	133
4.5.3 <i>In vitro</i> recombinant protein expression	133
4.5.4 <i>In vitro</i> RNA transcription	134
4.5.5 Affinity capturing of proteins with ability to bind U5- and U1-snRNA Candidates	134
4.5.6 Gel shift assay with radio-isotope labelled RNAs	135
4.5.7 Reverse transcription (RT) PCR	135
 Chapter 5: Unusual ncRNAs in <i>Giardia</i> and the putative RNAi pathway	
5.1 Introduction: The mechanism of dsRNA-induced gene silencing and RNAi in unicellular eukaryotes	137
5.2 The unusual ncRNA repeats in <i>Giardia</i>	144
5.3 Protein components of the putative RNAi pathway in <i>Giardia</i>	152
5.4 The possible existence of a truncated <i>Giardia</i> Dicer protein	160
5.5 Conclusion and overview of the unusual RNAs found in this study....	167
5.6 Experimental material and methods	
5.6.1 Primers used in generating <i>in vitro</i> transcription template for Girep RNAs and Girep1RNA and primers used for recombinant truncated Dicer peptide	168
5.6.2 Recombinant truncated Dicer peptide expression and purification	168
5.6.3 Nuclease activity assay	169

Final Words	171
Appendix-1	175
Appendix-2	215
Appendix-3	237
Appendix-4	257
Bibliography	263