Mining the Genetic Diversity of the Pathogenic Protozoan, *Neospora caninum*

By

Larissa Calarco

January
2020

School of Life Sciences
Faculty of Science
University of Technology Sydney

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy (PhD) at the University of Technology Sydney (UTS)

Certificate of Original Authorship

I, Larissa Calarco, declare that this thesis is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Life Sciences at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise referenced or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

This document has not been submitted for qualifications at any other academic institution.

This research is supported by the Australian Government Research Training Program.

Signature:

Production Note:
Signature removed prior to publication.

Date: January 2020

Dedication

This thesis is dedicated to my wonderful parents, Dominic and Kathy Calarco. I would not be who or where I am today without their love, sacrifices, generosity, and constant support, not only throughout my studies, but everything leading up to this point. I am eternally grateful for the ample opportunities they have strived to provide me with, always ensuring that I have the means to reach my full potential.

Acknowledgements

I would like to extend my sincerest thanks and gratitude to my supervisors, Professor John Ellis and Dr Joel Barratt, who I am indebted to. Completion of this thesis and research degree would not have been impossible without their constant and unfaltering support, succour, encouragement, patience, and guidance. It has truly been an honour and privilege being part of the Ellis research group for the last four years, and being supervised by such eminent, renowned professionals in their field has been an extremely rewarding and humbling experience.

I would like to thank Alexa Kaufer, Rory Gough, and the other members of the Ellis research group for their support, comradery, solidarity, and assistance. It's been a pleasure working alongside all of you.

I would also like to thank the University of Technology Sydney for the generous PhD scholarship, which has supported me throughout the course of my degree.

Last by not least, I would like to thank the ongoing support, love, and encouragement of my parents, grandparents, immediate and extended families, and friends.

Table of Contents

Publications arising from thesis

Calarco, L., Barratt, J., Ellis, J., 2018. *Genome wide identification of mutational hotspots in the apicomplexan parasite Neospora caninum and the implications for virulence*. **Genome Biol Evol**, 10 (9), 2417-2431.

Calarco, L. and Ellis, J., 2019. *Annotating the 'hypothetical' in hypothetical proteins: in-silico analysis of uncharacterised proteins for the Apicomplexan parasite, Neospora caninum*. **Vet Parasitology**, 265, 29-37.

Calarco, L., Barratt, J., Ellis, J., 2020. *Detecting sequence variants in clinically important protozoan parasites*. **Int J Parasitol,** 50 (1), 1-18.

Calarco, L. and Ellis, J., 2020. *Contribution of introns to the species diversity associated with the apicomplexan parasite, Neospora caninum*. **Parasitol Res**, 1-15.

Conference presentations arising from thesis

Calarco, L., Bush, S., Barratt, J. & Ellis, J. *Validation of a variant analysis pipeline for the identification of vaccine candidates for eukaryotic pathogenic organisms*. International Congress for Tropical Medicine and Malaria; Brisbane, Australia (2016).

Calarco, L., Barratt, J. & Ellis, J. *Validation of a variant analysis pipeline for the identification of vaccine candidates for eukaryotic pathogenic organisms*. Australian Society for Parasitology Annual Conference; Leura, Australia (2017).

Calarco, L., Barratt, J. & Ellis, J. *Genome wide identification of mutational hotspots in the apicomplexan parasite Neospora caninum, and the implications for virulence*. 14th International Congress of Parasitology; Daegu, Korea (2018).

Calarco, L., Barratt, J. & Ellis, J. *Genome wide identification of mutational hotspots in the apicomplexan parasite Neospora caninum, and the implications for virulence.* Australian Society for Parasitology Annual Conference; St Kilda, Australia (2018).

Calarco, L. & Ellis, J. *Contribution of introns to the species diversity associated with the apicomplexan parasite, Neospora caninum*. Australian Society for Parasitology Annual Conference; Adelaide, Australia (2019).

List of Figures

List of Tables

Supplementary Information

The thesis chapters 2-4 make reference to supplementary files, online resources, tables, and spreadsheets, containing additional supporting information related to respective methods and results. This information is contained within files accessible either at the URL specified at the end of published chapters, or on the Supplementary Information disk provided.

Thesis Abstract

Neospora caninum is a cyst-forming apicomplexan parasite, responsible for economic and reproductive losses to cattle industries worldwide, and represents a serious neurological disease in canines. Although discovered over three decades ago, progress towards treatment and control strategies against neosporosis, remains stagnant. Currently, common practices to combat the disease include passivity, or expensive culling of seropositive dams. However, vaccination represents a cost-effective and efficacious option, especially using live, attenuated isolates.

Members of the Apicomplexa consist of populations that vary enormously in their disease-causing potential, where *in vivo* experiments have demonstrated pathogenic variability between *N. caninum* isolates. The underlying question therefore, is what is the genetic basis of virulence within the species, and consequently, how can such information be exploited in vaccine development? Thus far, conventional techniques have been employed to study the intraspecies genetic diversity associated with *N. caninum*, generally involving PCR-based approaches targeting repetitive elements. However, a direct causal relationship between such diversity and important parasite phenotypes such as virulence, is yet to be established.

Alternatively, burgeoning next generation sequencing (NGS) technologies and *in-silico* tools have provided new opportunities to perform genome-wide scans in such organisms. Hence the objective of this body of work was to compare the genomes and transcriptomes of two distinct *N. caninum* isolates, using NGS data and bioinformatics workflows, to identify sequence variants in coding and non-coding DNA. Annotation of variable regions would reveal potential virulence markers distinguishing isolates of this species. Challenges accompanying such research include the lack of optimisation and

standardisation of NGS analysis tools for non-model organisms such as pathogenic Protozoa. This is compounded by the dubious accuracy of the *N. caninum* reference genome, as well as the disturbingly large number of proteins described as 'hypothetical' or 'uncharacterised'.

This body of research represents a thesis by compilation, consisting of four publications, and one chapter under review. Each chapter represents an independent study, which collectively address the research objective and current gaps in the literature. The results present polymorphic "hotspots" in concentrated windows of the *N. caninum* genome, where there is a correlation between hypervariable regions within protein-coding genes, and non-coding regions. Furthermore, an *in-silico* pipeline is developed to annotate uncharacterised proteins, subsequently identifying a subset of proteins potentially implicated in crucial parasite mechanisms, conducive to *N. caninum*'s success. It is trusted that this thesis contributes vital knowledge pertaining to *N. caninum* intraspecies diversity, aiding in the quest to develop a vaccine against neosporosis.