THE PATTERN AND PROCESSES OF GENOME CHANGE IN ENDOSYMBIONTS OLD AND NEW

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

Doctor of Philosophy

in Evolutionary Biology

Submitted by

Barbara Inge Karoline Schönfeld

2012

Institute of Molecular BioSciences

Massey University

New Zealand

Abstract

Bacterial endosymbionts are an important part of eukaryote evolution as they allow their hosts to exploit bacterial abilities. Plastids, the organelles that enable plants and eukaryotic algae to photosynthesise are ancient cyanobacterial endosymbionts. Since the initial symbiosis ~1.5 billion years ago the majority of their genes has been lost or transferred to their host's nucleus. This process has carried on independently in the different lineages following the diversification of the lineage.

I have compiled a comprehensive data set of fully sequenced plastid genomes to systematically study the frequency of gene transfers from the plastid to the nucleus across the different lineages. Following the reconstruction of the Plantae phylogenetic tree from plastid encoded proteins, gene loss events were reconstructed along its branches. My calculations show that gene losses have occurred at a relative high frequency and in a lineage specific way. This challenges the original idea that gene transfers from the organelle to the nucleus are rare and chance driven events.

Bacteria and eukaryotes continue to form endosymbioses and the study of these relationships produces valuable insights into the early stages of organelle evolution, bacterial metabolic pathways and metabolic regulation. They also allow us a glimpse into the ancient history of eukaryote evolution. For this reason, diatoms that have acquired cyanobacterial endosymbionts with the capability to fix molecular nitrogen were chosen to explore the potential and limitations of high-throughput sequencing technologies for investigating this type of relationship when DNA sequences are obtained from environmental samples and in the presence of bacterial contaminants. The results of this work confirmed the suitability of this relatively new technology to sequence mixed samples but also highlighted i) difficulties in sample preparation which can bias the composition of metagenomic samples obtained, and also ii) the varying suitability of different types of samples used in high-throughput sequencing.

In Gedenken an meinen Vater Günther Schönfeld

Acknowledgements

I would like to thank my supervisors Pete Lockhart and Lesley Collins for their support and guidance. They truly care and I will be forever grateful.

Many others have helped me along the way and should know that it is deeply appreciated:

Trish McLenachan and Phil Novis volunteered their time and expertise to help me in the lab and in the field, and have taught me much. Mike Steel and Tim White worked magic with numbers and scripts to make my results 14% more likely. Bill Martin provided inspiration, support, enthusiasm and helpful criticism when it was needed most. Uwe Maier's keen interest shaped the course of this project. David Penny and the assortment of brilliant people he has gathered around him over the years who have allowed me to share in the passion, the wonder, and the drinking of malt-based beverages, that is science.

The Institute of Molecular BioSciences and the Allan Wilson Centre for Molecular Ecology and Evolution gave the financial and practical support that made my work possible.

My warmest thanks have to go to all the friends and new family who have made my time in New Zealand the best of my life:

The Piripis and Ellicots who made me feel like family. The van Hoves and McComishs for being family. The flatmates, the crafters, the gamers, the foragers, the time travellers and the stargazers, whose friendship means the world to me.

Thank you, Bennet, for being there.

Finally and most importantly I would like to acknowledge my parents, my family and my good friends back home, on whose support I can always count and whose love is always with me.

Table of Contents

List of Figures

List of Tables