The origin and characterisation of new nuclear genes originating from a cytoplasmic organellar genome

Andrew Henry Mark Lloyd

A thesis submitted for the degree of Doctor of Philosophy

Discipline of Genetics

School of Molecular and Biomedical Science

The University of Adelaide

December 2010

Table of Contents

Abstract

Endosymbiotic transfer of DNA and functional genes from the cytoplasmic organelles (mitochondria and chloroplasts) to the nucleus has been a major factor driving the origin of new nuclear genes, a process central to eukaryote evolution. Recent developments have allowed the experimental reconstruction of DNA transfer and functional gene transfer, enabling investigation of the molecular mechanisms involved in these important evolutionary processes.

To simulate the process of functional endosymbiotic gene transfer, a chloroplast reporter gene *aadA,* which had been transferred from the chloroplast to the nucleus, was monitored for nuclear activation. In total 16 plant lines were screened, each line representing an independent nuclear insertion of the *aadA* gene. For each line ~50 million cells were screened resulting in three plants being recovered in which *aadA* showed strong nuclear activation. Activation occurred by acquisition of the CaMV 35S nuclear promoter or by nuclear activation of the native chloroplast promoter. Two fortuitous sites resident within the 3' UTR of *aadA* mRNA both promoted poly-adenylation without any sequence change. In addition, cryptic nuclear activity of the chloroplast promoter was revealed which became conspicuous when present in multiple nuclear copies.

To determine the method of chloroplast DNA insertion into the nucleus the insertion site was sequenced in line kr2.2. Complete characterisation of the nuclear sequence before and after gene transfer demonstrated simultaneous insertion of multiple chloroplast DNA fragments *via* synthesis dependent non-homologous end joining, probably at a site of double strand break (DSB) repair.

To further investigate the role of DSB repair in the nuclear insertion of organelle DNA, DSBs were induced at a specific nuclear location using the rare-cutting endonuclease I-SceI and the resulting repair events were observed. Analysis of ~300 DSB repair events indicated that most involved the loss of nucleotides from one or both ends being joined. Insertions were observed in five repair events. None of the inserted sequences were of organelle origin. Notably, the amount of nuclear sequence deleted was significantly larger in repair events involving insertion than in those without insertion, indicating that the two types of repair may be mediated by distinct pathways.

Declaration

This work contains no material which has been accepted for the award of any other degree or diploma in any university or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text.

I give consent to this copy of my thesis, when deposited in the University Library, being made available for loan and photocopying, subject to the provisions of the Copyright Act 1968.

I also give permission for the digital version of my thesis to be made available on the web, *via* the University's digital research repository, the Library catalogue, the Australasian Digital Theses Program (ADTP) and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

Signed ……*Date*……

Acknowledgements

A PhD is quite a journey and has taken me to many places that I would have otherwise never known, both literally and figuratively. I have learnt a great many things writing and working on this thesis, things about science, about people, about failure and success and about myself. It is hard to believe that it is now coming to an end.

First and foremost I would like to thank my supervisor Jeremy Timmis for all of his support, guidance, encouragement and supervision. Jeremy has a wonderful approach to science, and great warmth, patience and knowledge. From the time spent in his lab as an undergraduate student he has been a wonderful mentor and has instilled in me a great enthusiasm and respect for science.

I would like to thank all members of the Timmis lab, past and present, for making it such an enjoyable, engaging time. In particular I would like to thank Anna Sheppard for all of her help, especially when starting out on this journey and also for many wonderfully productive conversations that helped this work take shape. Thanks must also go to Yuan Li for all of her good cheer and technical assistance along the way. Much of this would have been impossible without her. Thanks also to Matt, Dong, Rory and Sven who as well as being a lot of fun, have been great people to work with and bounce ideas off. I would also like to thank everyone in the Genetics discipline, it has been a wonderful, supportive and enjoyable place to work.

To my family and friends your support and encouragement have been unending, you are truly wonderful people. In particular I would like to thank all past and present members of 'Palmerston'. It has been an amazing, enriching and thoroughly enjoyable time and a welcome escape from the lab.

Finally, I would like to Mairead. For all of your love, support, understanding and encouragement throughout the last three years I am truly grateful.

List of abbreviations

4-MU	4-Methylumbelliferone
A	adenine
ATP	adenosine triphosphate
BAC	bacterial artificial chromosome
bp, kb	base pairs, kilobase pairs
C	cytosine
CaMV	cauliflower mosaic virus
˚C	degrees Celsius
cDNA	complementary deoxyribonucleic acid
cv	cultivar
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleoside triphosphate
DSB	double strand break
EDTA	ethylenediaminetetraacetic acid disodium salt
FISH	fluorescence *in situ* hybridisation
g	force of gravity
g, mg, µg, ng, pg	gram, milligram, microgram, nanogram, picogram
G	guanine
hr	hour
iPCR	inverse PCR
kr	kanamycin resistant
L, mL, µL	litre, millilitre, microlitre
M, µM	moles per litre, micromoles per litre
min	minute
mRNA	messenger RNA
cm, mm	centimetre, millimetre
MUG	4-methylumbelliferyl-beta-D-glucuronide
NHEJ	non-homologous end joining
numt	nuclear integrant of mitochondrial DNA
nupt	nuclear integrant of plastid DNA
norg	nuclear integrant of organellar DNA
nt	nucleotide
PCR	polymerase chain reaction
pmol	picomole
PVPP	polyvinylpolypyrrolidone
RACE	rapid amplification of cDNA ends
RLM-RACE	RNA ligase-mediated

RNA	ribonucleic acid
RNase A	ribonuclease A
rRNA	ribosomal RNA
RT-PCR	reverse transcription PCR
RT-QPCR	real-time quantitative PCR
sec	second
smPCR	single molecule PCR
sr	spectinomycin/streptomycin resistant
T	thymine
T-DNA	transfer DNA
TAE	Tris-acetate-EDTA
TAIL-PCR	thermal asymmetric interlaced PCR
Taq	*Thermus aquaticus*
Tris	tris(hydroxymethyl)aminomethane
tRNA	transfer RNA
U	uracil
U	Unit(s) of enzyme
UTR	untranscribed region
v/v	volume per volume
w/ v	weight per volume