The Dynamic Eukaryote Genome:
Evolution, Mobile DNA, and
the TE-Thrust Hypothesis

This Thesis is presented for the award of Doctor of Philosophy
Murdoch University, Perth, Australia

BY

Keith Robert Oliver

BA (Philosophy) Edith Cowan University
BSc (Multidisciplinary Science) Curtin University
BSc (Hons 1) (Biological Science) Murdoch University

DATE: May 23 2012
Declaration that this Thesis is My Own Work

This Thesis is entirely my own original work, and contains no previously published material, or material written by any other person, except where due reference has been clearly indicated in the text.

I have clearly stated any contribution of others to my thesis as a whole, including any research work or hypotheses not entirely attributable to me. The content of this Thesis is the result of work I have carried out since the commencement of my Research Higher Degree candidature, and does not include any significant parts of work that I have submitted to qualify for the award of any other Degree or Diploma in any University or other Tertiary Institution.

Signed: Keith Robert Oliver…………………………………….

Date………………………..
The Dynamic Eukaryote Genome: Evolution, Mobile DNA, and the TE-Thrust Hypothesis

Abstract

The discovery of transposable elements (TEs) by Barbara McClintock in the 1940s, triggered a new dawning in the development of evolutionary theory. However, similar to Gregor Mendel’s development of the laws of heredity in the nineteenth century, it was a long time before the full significance of this discovery was appreciated. Nevertheless, by the beginning of the 21st century, the study and recognition of TEs as significant factors in evolution was well underway. However, many evolutionary biologists still choose to ignore them, to highlight the loss of fitness in some individuals caused by TEs, or concentrate on the supposed parasitic nature of TEs, and the diseases they cause.

The major concept and theme of this thesis is that the ubiquitous and extremely ancient transposable elements are not merely “junk DNA” or “selfish parasites” but are instead ‘powerful facilitators of evolution’. They can create genomic dynamism, and cause genetic changes of great magnitude and variety in the genotypes and phenotypes of eukaryotic lineages.

A large variety of data are presented supporting the theme of TEs as very significant forces in evolution. This concept is formalised into a hypothesis, the TE-Thrust hypothesis, which explicitly
Abstract

presents detail of how TEs can facilitate evolution. This hypothesis opens the way to explaining otherwise inexplicable aspects of evolution, such as the mismatch between the phyletic gradualism theory, and the punctuated equilibrium concept, which is based on the fossil record.

Data from the studies of many metazoans are analysed, with a focus on the well studied mammals, especially the primates. Data from the seed plants are also included, with a strong focus on Darwin’s ‘abominable mystery’, the rapid origin, and the extraordinary success of the flowering plants.

TEs are ubiquitous and many of them are extremely ancient, probably dating back to the origin of the eukaryotes, and some are also found in prokaryotes. TEs can build, sculpt and reformat genomes by both active and passive means. Active TE-Thrust is due to transpositions by members of the TE consortium, or their retrotransposition of retrocopy genes, or by new acquisitions of TEs, or by the endogenisation of retroviruses, and other similar phenomena. Major results of this are that the promoters carried by TEs can result in very significant alterations in gene expression, and that sequences from the TEs themselves can become exapted or domesticated as novel genes. TEs can also cause exon shuffling, possibly building novel genes. Passive TE-Thrust is due to large homogenous consortia of inactive TEs that can act passively by causing ectopic recombination, resulting in genomic deletions, duplications, and possibly karyotypic changes. TE-Thrust often works together with other facilitators of evolution, such as point mutations, which can occur in duplicated, or
Abstract

retrocopy genes, sometimes resulting in new functions for such genes.

A major concept in the TE-Thrust hypothesis is that although TEs are sometimes harmful to individuals, and can lower the fitness of a population, they endow the lineage of that population with adaptive potential and evolutionary potential. These are extremes of a continuum of intra-genomic potential, and are not separate entities. This adaptive/evolutionary potential due to the presence and activities of the TE consortium of the genomes in a lineage, greatly enhance the future survival prospects of the lineage, and its ability to undergo evolutionary transitions, and/or to radiate into a clade of multiple divergent lineages. Lineages may acquire a TE consortium by new infiltrations of TEs, either by horizontal transposon transfer, *de novo* synthesis, or endogenisation of retroviruses. Lineages lacking an effective TE consortium are likely to lack adaptive/evolutionary potential and could fail to diversify, become “living fossils”, or even become extinct, as many lineages ultimately do.

The opposite of extinction is the fecund radiation of lineages, and it is shown here that fecund species-rich lineages such as rodents (Order Rodentia) and bats (Order Chiroptera) and the angiosperms, are all well endowed with many viable active TEs. The Simian Primates which have undergone major evolutionary transitions are also well endowed with viable and periodically active TEs, and/or large homogenous populations of TEs. Data on the “living fossils” such as the coelacanth and the tuatara are very
limited, but indicate a lack of new acquisitions of TEs, and/or the mutational decay of ancient TE families in their genomes.

Lineages are often in stasis, but a new acquisition of TEs, or other factors such as stress, hybridisation, or whole genome duplications (especially in angiosperms) may trigger a major burst of activity in the TE consortium, resulting in an evolutionary punctuation event. The TE-Thrust hypothesis thus offers an explanation for the punctuated equilibrium, frequently observed in the fossil record.

There are many other known facilitators of evolution, such as point mutations, whole genome duplications, changes in allele frequency, epigenetic changes, symbiosis, hybridisation, simple sequence repeats, karyotypic changes, drift in small populations, allopatric and sympatric reproductive isolation, co-evolution, environmental and ecological changes, and so on. In addition, there may be some as yet unknown facilitators of evolution. However, TEs usually make up between 20 to 80 percent of the genomes of eukaryotes, as against one or two percent of coding genes, and are known to be able to make genomic modifications (“mutations”) that cannot be made by other facilitators of evolution. TEs also come in many superfamilies, and in thousands of families, which make up the mobile DNA of the earth’s biota. It is apparent then that their influence on, and facilitation of, eukaryotic evolution has been very significant indeed. In this thesis data are presented, which indicate that these ubiquitous and extremely ancient TEs are powerful facilitators of change, essential to the evolution of the earth’s biota.
Abstract

The TE-Thrust hypothesis, when fully explored, developed, and tested, if confirmed, must result in an extension to the Modern Synthesis, or even become a part of a new paradigm of evolutionary theory.
# Table of Contents

Acknowledgements...............................................................................................-1-

List of Published Manuscripts.............................................................................-3-

Statement of Contributions to Jointly Authored Works

   In this Thesis........................................................................................................-5-

Abbreviations used in this Thesis..........................................................................-7-

**Chapter One**: General Introduction..............................................................1

**Chapter Two**: Transposable Elements: Powerful Facilitators
   of Evolution..........................................................................................................17

**Appendix to Chapter Two**: Biological Evidence Supporting
   Punctuated Equilibrium Type Evolution...........................................................51

**Chapter Three**: Mobile DNA and the TE-Thrust Hypothesis:
   Supporting Evidence from the Primates............................................................62

**Chapter Four**: The TE-Thrust Hypothesis and Plants: Darwin’s
   “Abominable Mystery” and Other Puzzles.........................................................113

**Appendix to Chapter Four**: TE-Thrust and Evolution,
   Devolution, and Background Extinction in the Metazoans..151

**Chapter Five**: Transposable Elements and Viruses as Factors
   in Adaptation and Evolution: an Expansion and
   Strengthening of the TE-Thrust Hypothesis....................................................165
Acknowledgements

I am very grateful for the continuing support, guidance, and encouragement I have received from my supervisors, Associate Professor Wayne Greene, Emeritus Professor Jen McComb, and the philosopher, Dr Alan Tapper, and also to Professor Stuart Bradley early in my studies, before other duties precluded him from continuing in a supervisory role.

I am especially grateful to Jen McComb who has been an encouraging friend, as well as a supervisor, and to Wayne Greene who shared his vast store of knowledge of molecular biology with me, in addition to his supervisory role. In addition I thank Wayne Greene for the PowerPoint diagrams he did for me.

I thank the talented Margot Wiburd very much for her persistence and dedication in sorting out some computer and setting out problems, which I encountered along the way.

The contributions to my rapidly growing knowledge in the subject matter of this Thesis, by overseas email contacts, have also been outstanding, and I am especially grateful to Frank Ryan and Luis Villarreal, but also to David King, and George Parris, as well as Norihiro Okada, Cedric Feschotte, and several others.
Acknowledgements

Fay Uhe, together with some of my family, unfortunately, endured a measure of neglect, while I was busy with this study. I thank Fay, and my family, for their tolerance and understanding.

I am grateful to the Council of the Royal Society of Western Australia for enabling me to organise, together with Vic Semeniuk, a Symposium on Evolutionary Biology, in 2009, which marked the 150th Anniversary of the publication of the ‘Origin of Species’, the 200th Anniversary of the publication of Lamarck’s ‘Philosophie Zoologique’, and the 200th Anniversary of the birth of Darwin. A variety of aspects of, and opinions on, evolutionary biology were presented at this very successful and well attended event.

I have long had a passionate interest in evolutionary theory, and I am grateful to all the people who have discussed this subject with me, both the orthodox, where the late Julian Ford is remembered, and the unorthodox, where Ted Steele comes to mind. I also thank the more neutral microbiologist Graham O’Hara for some very interesting discussions.

Finally I am grateful for my long association with Murdoch University, and the School of Biological Sciences and Biotechnology, with whom I have been a mature-age undergraduate student, an honorary research associate, an adjunct lecturer, and finally a post-graduate student. This has been a fruitful association indeed.

Keith Oliver, 2012
List of Published Manuscripts


Oliver K. R., Greene W. K. (2012a) Transposable Elements and Viruses as Factors in Adaptation and Evolution: an Expansion and Strengthening of the TE-Thrust Hypothesis. This has been submitted for publication.

List of Published Manuscripts

Oliver K R, McComb J A and Greene W K (2012) The TE-Thrust Hypothesis and Plants: Darwin’s “Abominable Mystery” and Other Puzzles. This is being reformatted and submitted for publication.
The manuscripts included in this Thesis (Chapters 2 to 5), and the Appendices 1, 2 and 3 were joint contributions. Specific contributions to the work, analysis or synthesis, and/or writing and editing, of these manuscripts were as follows:


Keith Oliver was responsible for conception and theoretical design, analysis and interpretation of data, manuscript writing and editing. Wayne Greene, as supervisor, contributed to theoretical design, analysis and interpretation of data, manuscript writing and editing. This also applies to Chapters 3 (published) and 5 (submitted for publication), and Appendices 1, 2, and 3 (all published).

(Chapter 4) Oliver K R, McComb J. A. and Greene W K. The TE-Thrust Hypothesis and Plants: Darwin’s “Abominable Mystery” and Other Puzzles. (This is ready for reformatting for publication).

Keith Oliver was responsible for conception and theoretical design, analysis and interpretation of data, manuscript writing and editing. Jen McComb and Wayne Greene as supervisors, contributed to theoretical design, analysis and interpretation of data, manuscript writing and editing.


Abbreviations used in this Thesis

Alu: a SINE specific to the primates.
ARMD: Alu-Recombination-Mediated Deletion.
DNA-TE: DNA Transposable Element, or Transposon (Class II TE).
EBN: Endosperm Balance Number.
ECR-LTR: Envelope-Class Retrovirus-like LTR retro-TE.
ERV: Endogenous RetroVirus.
ERVs/sLTRs: Endogenous Retroviruses and/or solo Long Terminal Repeats.
EVE: Endogenous Viral Element.
HTT: Horizontal Transposon Transfer.
ITR: Inverted Terminal Repeat.
LINE: Long INterspersed Element.
L1 or LINE-1: (autonomous) Long INterspersed Element 1.
LTR: Long Terminal Repeat.
LTR retro-TE: Long Terminal Repeat retro-TE.
MIR (SINE): ancient Mammalian-wide Interspersed Repeat.
MITE: (non autonomous) Miniature Inverted-repeat Transposable Element.
Mya: Million years ago.
Myr: Million years.
retro-TE: Retrotransposable Element or Retroposon (Class I TE).
RT: Reverse Transcriptase.
ORF: Open Reading Frame.
RIP: Repeat-Induced Point mutation.
SINE: Short INterspersed Element.
Abbreviations used in this Thesis

sLTR: solo Long Terminal Repeat.
sLTR/ERV: solo LTR and/or an ERV.
SVA: SINE-VNTR-Alu chimaeric retro-TE.
TE: Transposable Element, transposon, or retroposon.
TEd-alleles: alleles deactivated or destroyed, by TE insertions.
TEm-alleles: alleles modified in either function or regulation, or duplicated, by TE insertions.
TSD: Target Site Duplication.
UTR: UnTranslated Region.
VIMS: Variation Inducing Mobile Sequences.
VNTR: Variable Number Tandem Repeat.