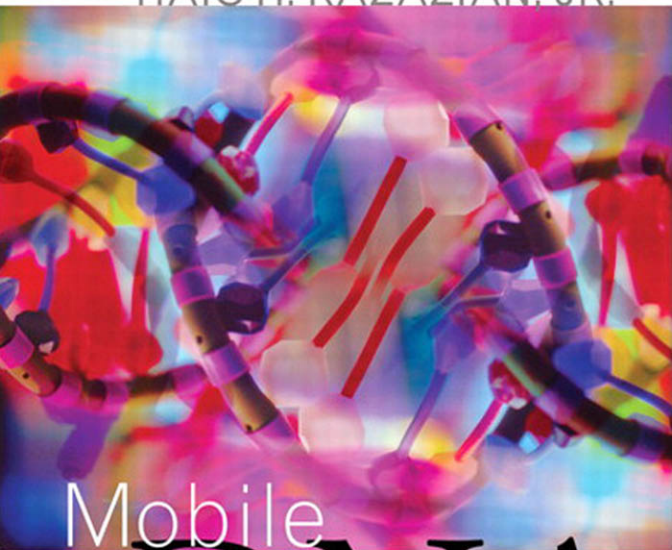


HAIG H. KAZAZIAN, JR.



Mobile

DNA

Finding Treasure in Junk

Mobile DNA

This page intentionally left blank

Mobile DNA

Finding Treasure in Junk

Haig H. Kazazian, Jr.

Vice President, Publisher: Tim Moore
Associate Publisher and Director of Marketing: Amy Neidlinger
Acquisitions Editor: Kirk Jensen
Editorial Assistant: Pamela Boland
Operations Manager: Gina Kanouse
Senior Marketing Manager: Julie Phifer
Publicity Manager: Laura Czaja
Assistant Marketing Manager: Megan Colvin
Cover Designer: Gary Adair
Managing Editor: Kristy Hart
Senior Project Editor: Lori Lyons
Copy Editor: Language Logistics, LLC
Proofreader: Sheri Cain
Indexer: Angela Martin
Compositor: Nonie Ratcliff
Manufacturing Buyer: Dan Uhrig
© 2011 by Pearson Education, Inc.
Publishing as FT Press
Upper Saddle River, New Jersey 07458

FT Press offers excellent discounts on this book when ordered in quantity for bulk purchases or special sales. For more information, please contact U.S. Corporate and Government Sales, 1-800-382-3419, corpsales@pearsontechgroup.com. For sales outside the U.S., please contact International Sales at international@pearson.com.

Company and product names mentioned herein are the trademarks or registered trademarks of their respective owners.

All rights reserved. No part of this book may be reproduced, in any form or by any means, without permission in writing from the publisher.

Printed in the United States of America

First Printing April 2011

ISBN-10: 0-13-707062-4

ISBN-13: 978-0-13-707062-6

Pearson Education LTD.

Pearson Education Australia PTY, Limited.

Pearson Education Singapore, Pte. Ltd.

Pearson Education Asia, Ltd.

Pearson Education Canada, Ltd.

Pearson Educación de México, S.A. de C.V.

Pearson Education—Japan

Pearson Education Malaysia, Pte. Ltd.

Library of Congress Cataloging-in-Publication Data

Kazazian, Haig H. (Haig Hagop), Jr., 1937- author.

Mobile DNA : finding treasure in junk / Haig Kazazian.

p. ; cm.

Includes bibliographical references.

ISBN-13: 978-0-13-707062-6 (hardback : alk. paper)

ISBN-10: 0-13-707062-4 (hardback : alk. paper)

1. Insertion elements, DNA. 2. Transposons. 3. Recombinant, DNA. 4. Genomics. I. Title.

[DNLN: 1. DNA Transposable Elements. 2. DNA, Recombinant—physiology. 3.

Genomics—methods. 4. Sequence Analysis, DNA. QU 470]

QH462.I48K39 2011

572.8'69—dc22

2010050899

*To Lilli, my loving wife of so many years,
who put up with my idiosyncrasies
and encouraged me to write this book.*

This page intentionally left blank

Contents

	Preface: Thoughts on Doing Science	xii
Chapter 1	Introduction to Mobile DNA	1
Chapter 2	Varieties of Mobile DNA	5
Chapter 3	DNA Transposons	19
Chapter 4	Mobile DNA of Model Organisms	29
Chapter 5	Exceptional Scientists Working on Mobile DNA in Lower Organisms	35
Chapter 6	Role of Bioinformatics in Genome Analysis	43
Chapter 7	The Prologue	49
Chapter 8	“Welcome to the Wonderful World of LINES”	59
Chapter 9	An Experimental Breakthrough	73
Chapter 10	Reverse Transcriptase to the Rescue	81
Chapter 11	A Quirk of L1 Elements—A Lousy 3' End Is Important for Genome Evolution	85
Chapter 12	A Tour de Force from Tom Eickbush	89
Chapter 13	“I don’t believe all those colonies represent retrotransposition events.”	93
Chapter 14	L1 Encodes an Endonuclease	101
Chapter 15	The Jocks	105
Chapter 16	The Mayor and the Frenchman	115
Chapter 17	Ostertag’s Coups	121
Chapter 18	The Independent Canadian	133
Chapter 19	The Musician Scientist	141

Chapter 20	Young Ladies in the Back Bay	145
Chapter 21	The Brilliant Young Lady from China	157
Chapter 22	Hiroki's Big Surprises	163
Chapter 23	A Young Man with a Purpose	173
Chapter 24	Other Mobile DNA in Mammalian Genomes	179
Chapter 25	Effects of Retrotransposons on Mammalian Genomes	187
Chapter 26	Host Factors Involved in L1 Retrotransposition	201
Chapter 27	Why Mobile DNA?	207
Chapter 28	The Future of Mobile DNA Research	209
Chapter 29	Predictions for Mobile DNA	221
	References	225
	Glossary	249
	Index	255

Acknowledgments

I thank my early mentors, Lafayette Noda and Lucille Smith, at Dartmouth; Barton Childs at Johns Hopkins; and Harvey Itano at the NIH. I thank colleagues who helped me in entering the mobile DNA field, especially Maxine Singer, Allan Scott, and Jef Boeke. I thank all the lab members, both those mentioned and those unmentioned, for their tireless efforts to contribute to the mobile DNA field. I thank other colleagues, John Moran, Nancy Craig, Vivian Cheung, and Eric Ostertag, for helpful comments and critiques of the book. I also thank Adam Ewing, Dustin Hancks, and Lauren Zeitel for their advice and comments.

About the Author

Haig Kazazian was born and raised in Toledo, Ohio. After attending public schools there, he received his A.B. degree from Dartmouth College in 1959. He then attended Dartmouth Medical School, a two-year school at the time, and finished his M.D. degree at Johns Hopkins University School of Medicine. At Hopkins, he met his wife of nearly 50 years and married during his internship in Pediatrics at the University of Minnesota Hospital. After two years training in Minneapolis, he returned to Johns Hopkins for a two-year fellowship in genetics with Barton Childs, M.D. He then trained for two years in molecular biology in the lab of Harvey Itano, M.D., at the NIH. After a third year of Pediatric training at Johns Hopkins, he joined the faculty there in 1969. He rose through the ranks to become a full professor in 1977, and at that time, he headed the Pediatric Genetics Unit. In 1988, he became Director of the Center for Medical Genetics at Johns Hopkins.

After 25 years on the Hopkins faculty, he was recruited to the University of Pennsylvania School of Medicine as Chair of the Department of Genetics in 1994. At Penn, he recruited 10 young faculty to the department. In 2006, he stepped down as department chair, but remained as the Seymour Gray Professor of Molecular Medicine in Genetics until 2010. In July 2010, he returned to Johns Hopkins as a Professor in the Institute of Genetic Medicine.

Dr. Kazazian is still heavily involved in molecular genetic research, concentrating for the past 20 years on mammalian and human transposable elements, or “jumping genes.” Prior to 1988, he characterized much of the variation in the cluster of genes involved in production of the beta chain of human hemoglobin. With Stuart Orkin at Harvard, his work led to the nearly complete characterization of the mutations causing the β -thalassemias, common anemias in regions of the world endemic for malaria.

Dr. Kazazian is a member of a number of national organizations, including the Institute of Medicine of the National Academy of Sciences and the American Academy of Arts and Sciences. He has received a number of honors for his research, most notably the 2008 William Allan Award, the top honor of the American Society of Human Genetics.

Preface

Thoughts on doing science

Before diving into the subject of mobile DNA and my adventures in the field, I'd like to provide a few personal tips from my experience on working with some success in science for the past 45 years. Doing science is often very difficult and extremely hard work, requiring long hours. In my view, the first thing necessary is very high personal motivation. My original and long-time mentor, Barton Childs, an esteemed Professor of Pediatrics and the "father of pediatric genetics," always used to say, "You've got to burn to do research!" You can't go at it with a half-hearted enthusiasm or self-doubt.

If you do have high personal motivation, you then need to get excellent training, both in didactic class work and in the nuts and bolts of how to do research. You need to find a subject or area that really interests you, no matter what the field. Then find the investigator who is doing excellent research in your field of interest, hopefully someone at the forefront, but also consider that the person's lab is not too large so that he or she will have sufficient time to spend with you. You want someone who will discuss your research with you on a daily basis, perhaps so much so that you feel that he or she is pestering you all the time for new data. That kind of attention means that that individual has great interest in your work. You need that kind of person for both your predoc and your postdoc training. Your trainers also need to be available for discussion of all kinds of problems, both those that you face in the lab and those that are related to other aspects of your life.

Next, you need a dependable mentor. Your mentor could be either your predoc or your postdoc trainer, or it could be a member of your thesis committee or another senior investigator from down the hall. However, you'd like a mentor that you can carry over from your

training days into your first 5–10 years as a faculty member. That mentor can help you with all kinds of problems and questions, giving advice for how to approach various professional and daily life situations. Having an interested, accessible, and experienced mentor is crucial to success in science. Behind every good scientist is an outstanding mentor. I certainly had one in Barton Childs, even if I didn't follow his advice at every turn. He was probably my major mentor for at least 20 years.

I've talked about mentorship from the aspect of the trainee, but what about the importance of being a good mentor? From your first academic job to becoming a long-term lab director, you have the responsibility for mentoring predoc and postdoc trainees. I have usually found it rewarding to give trainees considerable independence, letting them pick their own problems from a wide variety of problems available in the lab. This works well when the trainee is very bright and picks one or more problems that are of real interest to the lab director. When the problem is of little interest to the lab director, there is a good chance that the work will flounder. However, if the problem is important to the mentor, the mentor will add ideas and enthusiasm to the work. I have dealt with both situations over my career, as the reader will discover in this book. A third situation occurs when the student or postdoc needs to finish a period of training and hasn't had much success up to that point. The trick at that time is to find a project that is important for the field (so that the student will take pride in his or her accomplishment), has a clear-cut endpoint, and uses techniques already available in the lab. The design of this project usually requires considerable input from the lab director.

Then there is the question of how one should approach other scientists. Should one be open in discussing new data even with colleagues in the same field, or should one be secretive to avoid being "scooped?" My view has always been that it is better to be open but prudent. It is good to discuss your unpublished work at meetings. If your work is important, your colleagues will respect you for talking

about new data and not rehashing work that has been published and that they've heard previously. Moreover, it is very, very rare that another investigator can start a new experimental tack or line and actually beat you to publication. After all, you've probably been working on that same question for a year or two, so you've got a major head start. It's a rare investigator indeed who would sail off in a new direction hoping to beat you to publication of data that you've just presented. As a general rule, openness in presenting and discussing new data is the best approach.

A corollary to openness is to discuss your science with a wide range of other scientists, including those within your immediate field, e.g., mobile DNA, those in the broader field, e.g. human genetics, and those in other fields of biomedical science, e.g., immunology or developmental biology. You never know from where the next good experimental idea will come. The reader will find throughout this book that members of my lab and I personally have gotten ideas from a wide variety of sources who are mentioned at some future time. This plethora of good ideas has come from discussing the work with a large number of other scientists and sources and being as open as possible to new ideas.

I once knew a well-trained, smart young researcher who had a great deal of trouble gaining traction in his field. I always thought that his problem was that he stayed in his lab and did not seek discussion of his science with colleagues. At the other extreme was and still is the Medical Research Council (MRC) laboratory at Cambridge, England, whose investigators have had enormous success over many decades. The Cambridge MRC labs have housed a number of Nobel Laureates, including Francis Crick, Fred Sanger, Sydney Brenner, Aaron Klug, and others. After spending a few months at the MRC early in my career, I felt that a major factor in the success of that lab was the English tradition at that time of a common coffee break in the morning and a common tea break in the afternoon. At 10:30 AM, every investigator, from the trainees to the most senior people, would

gather in the cafeteria for morning coffee and, importantly, discuss science for 30 minutes over coffee. A similar gathering would occur at 3:30 PM over afternoon tea. The number of great new ideas passed from one investigator to another, from past and future Nobel Prize winners to beginning postdocs, and vice versa, was astonishing. Open discussion of science is wonderful for the development of new ideas.

Now I'd like to make a general comment on picking problems in your field on which to work. I've always believed that the problem should be important but potentially solvable with hard effort. All researchers are gamblers. A colleague used to tell me to pick problems with 5 to 1 or 10 to 1 odds of success. Those problems were about right in terms of difficulty. Odds of 50 to 1 or 100 to 1 were too long, and success on those problems was too unlikely. Odds of 2 to 1 meant that the problem was too easy and relatively unimportant, so called "low-hanging fruit." I've also felt that it is best to pick problems that are logical next steps in the project but are important to the field and have those reasonable odds of success, which would be 5–10:1.

My last point is to keep one's mind alert for possible collaboration. Collaborations with other scientists should be welcomed as a way to broaden one's scientific outlook and scope. If two investigators have differing expertise that can be applied to solve a particular problem, this is an ideal situation for collaboration. I once heard it said that collaboration finds its own level, meaning that in order to work best, collaborators should be on the same level of experience and respect in the field. In this way, I've had good collaborations as a postdoc with another postdoc and as a senior scientist with other senior scientists. Many of these collaborations are discussed throughout the book.

This page intentionally left blank

1

Introduction to Mobile DNA

Charles Darwin would be surprised. Indeed, even present day scientists are surprised by the existence of mobile DNA. Consider the skepticism within the scientific community that greeted Barbara McClintock, already a highly-respected scientist, when she announced that she had found what appeared to be mobile DNA in maize plants (McClintock, 1950). DNA was the genetic material, so it must be static, stable, and immobile. The mutation rate had been determined to be $\sim 10^{-8}$ per nucleotide, or building block, of DNA per generation—very low indeed. How and why would some DNA move from place to place in a genome? Scientists are still grappling with these questions. Two hundred years removed from Darwin's birth, and we're still wondering how mobile DNA with all its detrimental effects on organisms could have reached such high proportions in the genomes of mammals and plants. Yet mobile DNA is found in all forms of living things, including plants, animals, bacteria, and archaea. The genome seems to cherish its ability to make rapid changes by rearranging some of its parts as opposed to the slow change afforded by the nucleotide mutation rate.

One theme of this book is that biological scientists have come to expect the unexpected. The study of living things is full of surprises. One of them is the prevalence of mobile DNA in genomes. Another is that most genes are broken up by sections of DNA called introns that need to be removed at the RNA stage in order for the genes to function. A third is that the protein-coding regions of genes make up a very small fraction of mammalian genomes. A fourth surprise is the importance of reverse transcriptase, the enzyme that synthesizes DNA from an RNA template. These are just a few examples of old surprises, or unexpected findings, that have now become hard facts in

all biology textbooks. Many more will be highlighted in the research adventures outlined in this book. These “unexpected observations” provide excitement and anticipation for even the most experienced researchers. What finding will be the next to shatter our present view of the biological world? One can be sure that the future will bring many more surprises to delight the graduate student just beginning his or her studies.

Prior to 1970, scientists thought that the genome was composed mostly of genes lined up like balls on a string with some repetitive DNA in between the balls. Then in the late 1970s, introns were found to break up the regions of genes that encode proteins (Berget et al., 1977; Chow et al., 1977). Protein-coding regions were disrupted by intervening sequences (introns) that required removal from pre-messenger RNA before the intact protein could be synthesized. Soon, we knew that introns were much larger than protein-coding regions, then called exons. The DNA between the genes along with most of the intronic sequences of genes was thought to be functionless, and was called “junk DNA” (Orgel and Crick, 1980). However, now we know that introns make up about 30% of human and mammalian genomes, and exons encode only between 1 and 2% of the human genome (Lander et al., 2001). What a comedown for protein-coding regions! Thus, over 98% of human DNA had been dismissed as “junk.”

Transposable elements were then found in human DNA, and this active mobile DNA along with the remnants of many transposition events over hundreds of millions of years is now known to account for at least 50% of human genomic DNA. This transposable element DNA, both those relatively few sequences that are presently mobile, and the many remnants of old events are now demonstrating function. However, this function is evident only in the many ways mobile DNA can modify the genome over evolutionary time. It can be co-opted for useful purposes but has not yet been definitively shown to have a useful function in the individual organism. Moreover, DNA encoding small RNAs of different types and functions has been discovered amidst the “junk.” Enhancer sequences at great distances from the genes upon which they act are being found continually. Segmental duplications of hundreds to many thousands of nucleotide pairs of DNA are strewn around the genome and are further grist in

the mill of genome plasticity and malleability. The bottom line is that “junk” DNA is gradually being eroded away as function is found for a greater and greater fraction of genomic DNA. In this book, I concentrate on the “junk” DNA that is mobile or has been over the millennia. This is the DNA that those of us in the mobile DNA field have come to treasure.

In the next several chapters, I provide details on important topics in the mobile DNA field as well as discuss a number of top scientists who have been pioneers in many areas involving mobile DNA. I then discuss the state of the human mobile DNA field prior to my involvement in it, what led to my fascination with mobile DNA, and why I jumped at the chance to work on it when the opportunity presented itself. Later, I discuss many of the people who worked in my lab up to the present time, their most important work, and the relationship of that work to what is known about L1 biology today. This is followed by important findings of other labs working on mammalian mobile DNA, ending with some thoughts about the future of the field. Yes, DNA as genetic material would have surprised Charles Darwin, but mobile DNA would have really made his head spin!

This page intentionally left blank

Index

NUMBERS

1000 Genomes Project, 47
14kb (HLA-A gene), 174

A

A (deoxyadenosine monophosphate), 9
A3G (APOBEC3G), 203
Ac (activator) control element, 19
adenomatous polyposis coli (APC) gene, 211
African origin of modern humans, 47
alleles, 78
alternative splicing, SVA elements, 174-176
Alu elements
 in trans retrotransposition by L1s, 189-192
 insertions, 187
 mammalian genomes, 179-180
Alu restriction endonuclease site, 179
AluYa5 subfamily (Alu elements), 179
American Society of Human Genetics meeting (1983), 61

AmnSINE1 ultraconserved SINE, 186
AMV (avian myeloblastosis virus), 50
antisense promoter effects (retrotransposition), 196
Antonarakis, Stylianos, 60
APC (adenomatous polyposis coli) gene, 211
APOBEC3 (apoprotein B-editing catalytic polypeptide 3) proteins, 202-203
APOBEC3G (A3G), 203
Arabidopsis thaliana
 model organism, 33
 role of small RNAs, 41
attTn7 insertion site, 25
autonomous retrotransposons, 6
avian myeloblastosis virus (AMV), 50

B

β -thalassemia mutations, 61
B1 elements, 179, 185
B2 elements, 186
Babushok, Dasha
 genomic distribution of *de novo* insertions, 149, 152
 PIPSL gene, 153-156

Badge, Richard, 209
 Baker, Tania, 35-38
 Baltimore, David, 50
 Batzer, Mark, 177
 Belfort, Marlene, 35, 38
 Benetzen, Jeff, 35
 beta-actin (CAG) promoter, 163
 biochemical characterization of
 retrotransposons, 218-220
 biochemistry of transposition, 37
 bioinformatics, 43-48
Biology of Homo Sapiens, 52
 biology of L1, 73-79
 Bishop, J. Michael, 35
 Boeke, Jef, 35, 39, 56, 81, 101
 Boissinot, Stephanie, 70
 Britten, Roy, 50
 Brouha, Brook, 105, 107-112
 Brown, Pat, 35, 38
 Bucheton, Alain, 145
 Buzdin, Anton, 136

C

C (deoxycytidine
 monophosphate), 9
C. albicans (zorros), 29-31, 202
C. elegans
 DNA transposons, 41
 model organism, 32
 CAG (beta-actin) promoter, 163
 Capecchi, Mario, 117
 Carstens, Russ, 176
 cell culture assay,
 retrotransposition, 86, 93-99
 cell transfections, 97
 cellular stress, influence on
 retrotransposition, 220
 Central Dogma of Biology, 9
 Chaconas, George, 37
 Chandler, Mick, 35-38

Cheung, Vivian, 209-210
 Childs, Barton, 59
 chimeras, template
 switching, 197
 chimpanzee genome, 44
 choroideremia, L1 insertion in
 embryogenesis, 172
 chromatin modification, HDACs
 (histone deacetylases), 206
cis preference, 76, 189
 classes of mobile DNA
 DNA transposons, 6-8
 effect on genome evolution,
 14-18
 retrotransposons, 8-14
 cold spots, DNA
 polymorphisms, 61
 Cold Spring Harbor
 meeting (1986), 52
 composition
 DNA, 9
 DNA transposons, 7
 RNA, 9
 consensus sequence (L1s), 54
 control elements, DNA
 transposons, 19
 Cooke, Bob, 60
 copy and paste mobility
 mechanism, 6
 Coufal, Nicole, 172
 Craig, Nancy, 35-37
 CRE-1 transposable element, 81
 Curcio, Joan, 35, 39
 cut and paste mobility
 mechanism, 6-7

D

DDE superfamily of
 recombinases, 21
 DDM1 gene, 33
de novo L1 insertions, 211

DeBerardinis, Ralph

in vivo retrotransposition,
117-119

mouse L1 elements, 115-117

T_F insertions in mouse, 117-118

deletions, L1/L1-mediated

insertions, 194

deoxyadenosine

monophosphate (A), 9

deoxycytidine

monophosphate (C), 9

deoxyguanine

monophosphate (G), 9

dissociator (Ds) control

element, 19

distribution, retrotransposons,

187-189

DNA

composition, 9

contrast with RNA, 9

hybridization technique, 50

hypomethylation, 205

methylation, 205

polymorphisms, 61

renaturation, 50

transposons. *See* transposons

Dombroski, Beth, 73-79**donor sites, 19****double helix structure, 9*****Drosophila melanogaster***

model organism, 31

P-element

horizontal transmission, 26

hybrid dysgenesis, 21-22

small RNAs, 203

Ds (dissociator) control

element, 19

dsRNA-binding protein,

Loquacious (Loqs), 204

Duchenne muscular dystrophy,

86-88

Duvernell, David, 173

dystrophin gene 48, 86-88

E

early transposon (Etn), 184

Edgell, Marshall, 51

EGFP (enhanced green

fluorescent protein), 102

substitution for neo gene, 121

tracking an embryonic

retrotransposition event

(Prak), 141-143

Eickbush, Tom, 89-91

embryonic development

L1 transcripts in various

developmental stages (Kano),

168, 172

retrotransposition in embryos

lacking L1 transgene

(Kano), 168

embryonic retrotransposition

event (Prak), 141-143

endo-siRNAs, 204**endogenous retroviruses, 32,**

181-183

endonuclease

endonuclease-independent

insertions, 199

L1 biology, 101, 104

enhanced green fluorescent

protein (EGFP), 102

substitution for neo gene, 121

tracking an embryonic

retrotransposition event

(Prak), 141-143

enhancer sequences, 2**envelope (env) gene, 11****epigenetic effects, reverse**

transcription of L1, 205-206

ES cells (human)
retrotransposition support of a
transfected active L1, 172

Escherichia coli, Tn7
transposon, 24

Etn (early transposon), 184
evolution of genomes, 14-18
Ewing, Adam, 47, 174, 209-210
exceptional scientists, 35-41
exonization, SVA elements
(Hancks), 176-177
exons, 2
expression of genes, 195-196

F

FACS (fluorescence activated
cell sorting), 109
factor VIII, 63
characterization of
mutations, 63
Southern blotting, 65-72
falciparum malaria, 62
Farley, Alex, 143
Federoff, Nina, 35, 56
Felsenfeld, Gary, 163
Feng, Qinghua, 102
Fink, Gerry, 56
fluorescence activated cell
sorting (FACS), 109
forensic applications, mobile
DNA as molecular marker, 47
formation of inversions
(Ostertag), 125-127
Furano, Tony, 70
future
predictions, 221-223
research
biochemical
characterization of
retrotransposons, 218-220

*genome-wide analysis of
retrotransposition events,*
209-217
*role of retrotransposition in
disease, 217-218*

G

G (deoxyguanine
monophosphate), 9
Gabriel, Abram, 81
Gage line, 164
Garfinkel, David, 35, 39, 56
Gellert, Marty, 35, 40
gene-trapping, SVA elements
(Hancks), 176-177
genes
APC (adenomatous polyposis
coli), 211
expression, 195-196
HLA-A (14kb), 174
ISL1, 186
MAST2, 174
neo, EGFP substitution for, 121
Nkg2d, 186
PIPSL, 153-156
syncytin, 183
Xist, 198
genome-wide association studies
(GWAS), 212
genomes
analysis, 43-48
evolution, 14-18, 85-88
mammalian
Alu elements, 179-180
effects of retrotransposons,
187-200
HERVs (human
endogenous retroviruses),
181-183
LINE elements, 180-181

mice, LTR-retrotransposons,
184-186
protein-coding regions, 1-2
sequencing
 human genomes, 43
 Tn7 random insertion, 25
unexpected findings, 1
G_F Subfamily (L1), 135
Giemsa chromosomal bands, 188
Gilbert, Nicolas, 138
Goodier, John, 88
 analysis of 3' transductions,
 133-134
 G_F subfamily of L1s, 135
 location of L1 proteins in
 human cells, 137-139
granulomatous disease, disease-
causing L1 insertion, 112
Grindley, Nigel, 35-37
Group I introns, 39
Group II introns, 39, 93
GWAS (genome-wide association
studies), 212
gypsy, 32

H

Hackett, Perry, 22
Haldane, J.B.S., 63
Hancks, Dustin, 173
 SVA alternative splicing,
 174-176
 SVA gene-trapping and
 exonization, 176-177
Harshey, Rasika, 37
hAT superfamily transposases, 20
HDACs (histone
deacetylases), 206
Heidmann, Thierry, 94, 189
HeLa cells, 97

hemoglobin genes
 β -thalassemia mutations, 61
 falciparum malaria, 62
hemophilia A, characterization of
mutations, 63-72
Hermes transposon, 20
HERV-K (Human Endogenous
RetroVirus-K), 32
HERVs (human endogenous
retroviruses), 181-183
heterochromatin, 30
heterologous promoters, human
L1 transgenic project (Kano),
163-165
 L1 transcripts in various
 developmental stages, 168-172
 retrotransposition in embryos
 lacking L1 transgene, 168
HCWD (human genome working
draft), 108
histone deacetylases
(HDACs), 206
HLA-A gene (14kb), 174
Holmes, Susan, 85
 insertion in dystrophin gene,
 86-88
 ORF1 protein, 85
 retrotransposition in cell
 cultures, 86
horizontal transmission, DNA
transposons, 26-27
host organisms, effects on L1
retrotransposition, 201
 APOBEC3 proteins, 202-203
 epigenetic effects, 205-206
 inhibition of non-LTR
 retrotransposons by small
 RNAs, 203-205
 Poulter and Han discoveries,
 201-202

hot L1s, 145-147, 214
 hot spots, DNA
 polymorphisms, 61
 human DNA, transposable
 elements, 2
 Human Endogenous
 RetroVirus-K (HERV-K), 32
 human endogenous retroviruses
 (HERVs), 181-183
 human ES cells, 172
 human genome working draft
 (HGWD), 108
 human genomes
 chimp comparison, 44
 L1 families, 43
 predictions for mobile DNA,
 221-223
 sequenced, 43
 human L1RP, 164
 human origins, analysis of
 retrotransposon insertion, 47
 Hutchison, Clyde, 51
 hybrid dysgenesis, 21-22
 hybridization, 50

I

I factor, 31
 IAPs (intracisternal
 A-particles), 184
 identity testing, mobile DNA as
 molecular marker, 47
in trans mobility, 180
in trans retrotransposition,
 189-192
in vitro cell culture assay, 53
in vitro system of retroviral
 integration, 38
in vitro transposition systems, 36

in vivo retrotransposition,
 117-119
 inactivation of X chromosome,
 effects of retrotransposition, 198
 inhibition, 203-205
 insertional mutagenesis, 20
 insertions
 Alu elements, 187
 de novo genomic distribution,
 149-152
 endonuclease-independent, 199
 factor VIII genes, 66
 known disease-producing
 insertions, 68
 L1/L1-mediated, 194
 LINE, 45
 Mendelian disease-causing
 insertions, 216
 non-repetitive sequence
 (Ostertag), 127, 131
 retrotransposons, 44
 SINE, 45
 somatic, 199-200
 insulator line, 164
 intracisternal A-particles
 (IAPs), 184
 introns, 1-2
 inversions (Ostertag), 125-127
 inverted repeat sequences, DNA
 transposons, 7
 ISL1 neuro-developmental
 gene, 186
 isolation of active human
 transposable elements, 81-83
 Itano, Harvey, 60

J-K

J subfamily (Alu elements), 179
 junk DNA, 2

- Kano, Hiroki, transgenic project**
human L1 without a
heterologous promoter,
163-165
L1 transcripts in various
developmental stages, 168-172
retrotransposition in embryos
lacking L1 transgene, 168
- Kennett, Roger, 97**
- Kidwell, Margaret, 21**
- Kleckner, Nancy, 35-36**
- Kunkel, Lou, 86**

L

- L1 (LINE1) biology, 49**
chimeras, template
switching, 197
cloning of full-length L1s of
Ta subset
retrotransposition assays,
106-107
reverse transcriptase
activity, 105-106
de novo insertions in
tumors, 211
DeBerardinis and Naas paper,
115-117
disease-causing insertion,
granulomatous disease, 112
distribution of
retrotransposition activity,
107-112
dystrophin gene, insertion into
exon 48, 86-88
endonuclease, 101-104
G_F subfamily, 135
hot L1s, 214
in vitro cell culture assay, 53
in vivo retrotransposition,
117-119
- insertions
as an insertional
mutation, 164
L1-mediated insertions, 194
resulting deletions, 194
transfected HeLa cells, 97
- isolation
active human transposable
elements, 81-83
precursor to insertion,
71-79
- L1 families, 43
life cycle, 136
locating L1 proteins in human
cells, 137-139
non-LTR retrotransposons, 91
ORF1 protein, 85
retrotransposition
Babushok model, 151
host factors, 201-206
mouse model, 122
- RNAs, 52
T_F subfamily, 117-118
transgenic experiment without a
heterologous promoter (Kano),
163-165
L1 transcripts in various
developmental stages,
168-172
retrotransposition in
embryos lacking L1
transgene, 168
- LIENp, nicking activities,**
101-104
- L2 elements, 180**
- L3 elements, 180**
- Lambowitz, Alan, 35-38**
- Levin, Henry, 35, 40**
- life cycle of L1, 136
- LINE elements, 180-181**

LINE insertions, 45
 LINE-SINE pairs, 180
 LINEs (long interspersed elements), 49
 living organisms
 mobile DNA proportions, 5
 unexpected findings, 1
 local hopping, 7, 22
 locating L1 proteins in human cells (Goodier), 137-139
 long interspersed elements (LINEs), 49
 long terminal repeat (LTR) retrotransposons. *See* LTR-retrotransposons
 Loqs (Loquacious) dsRNA-binding protein, 204
 LTR (long terminal repeat) retrotransposons, 6, 44
 HERVs (human endogenous retroviruses), 181-183
 mice genomes
 Etn (early transposon), 184
 IAPs (intracisternal *A*-particles), 184
 MaLR (mammalian apparent LTR-retrotransposons), 185
 SINEs, 185
 ultraconserved SINEs, 186
 LTR-transposons, 29
 Luan, Dongmei, 89
 lymphocyte DNA, somatic insertions, 213

M

Mager, Dixie, 95
 MaLR (mammalian apparent LTR-retrotransposons), 185
 mammalian apparent LTR-retrotransposons (MaLR), 185
 mammalian genomes
 Alu elements, 179-180
 effect of retrotransposons, 187
 3' and 5' transductions, 194-195
 antisense promoter effects, 196
 deletions resulting from L1/L1-mediated insertions, 194
 endonuclease-independent insertions, 199
 gene expression, 195-196
 in trans retrotransposition of Alu, SVA, and mRNA, 189-192
 L1 chimeras and template switching, 197
 non-allelic homologous recombination, 192-193
 purifying selection on retrotransposon distribution, 187-189
 somatic insertions, 199-200
 X chromosome inactivation, 198
 HERVs (human endogenous retroviruses), 181-183
 LINE elements, 180-181
 predictions for mobile DNA, 221-223
 mammalian mobile DNA, 53
 Mandal, Prabhat, 160
 Martienssen, Rob, 35, 41
 Martin, Sandy, 137
 MAST2 gene, 174
 Mathias, Steve, 82
 Mauriceville plasmid, 11
 Mayer, Jens, 182
 McClintock, Barbara, 19, 35

mechanism of inversion
 formation (Ostertag), 125-127
MeCP2 protein, 205
medaka fish, Tol2 transposon, 21
Mendelian disease-causing
 insertions, 216
messenger RNA (mRNA), 10
mice genomes, LTR-
 retrotransposons
 Etn (early transposon), 184
 IAPs (intracisternal
 A-particles), 184
 MaLR (mammalian apparent
 LTR-retrotransposons), 185
 SINEs, 185
 ultraconserved SINEs, 186
MicroRNAs (miRNAs), 203
migrations of humans, analysis of
 retrotransposon insertion, 47
MILI (Piwi protein
 homologue), 204
miniature inverted repeat
 transposable elements
 (MITEs), 41
miRNAs (MicroRNAs), 203
MITEs (miniature inverted
 repeat transposable
 elements), 41
MIWI2 (Piwi protein
 homologue), 204
Mizuuchi, Koichi, 35-36
mobile elements, model
 organisms, 29-33
mobility mechanisms
 copy and paste, 6
 cut and paste, 6-7
model organisms, 29-33
Moran, John, 93-99
mouse model of L1
 retrotransposition
 (Ostertag), 122

mRNA (messenger RNA), 10,
 189-192
Mu transposition, 37
multiple retrotransposon
 insertions, 44
Muotri, Alysson, 164
MusD endogenous
 proviruses, 185
mutant proteins, nicking
 activities, 101-104

N

Naas, Thierry, mouse L1
 elements, 115-117
neo gene
 EGFP substitution for, 121
 neomycin resistance, 95-97
neuro-developmental genes, 186
Neurospora crassa, 11
NHEJ (non-homologous end
 joining), 127
nicking activities, L1 ENp and
 mutant proteins, 101-104
Nienhuis, Art, 50
Nkg2d gene, 186
Noda, Lafayette, 59
non-allelic homologous
 recombination, 192-193
non-homologous end joining
 (NHEJ), 127
non-LTR retrotransposons, 6
 inhibition by small RNAs,
 203-205
 TPRT, 91
non-repetitive sequence
 insertions (Ostertag), 127-131
nonautonomous
 retrotransposons, 6
nucleotide combinations, 9
nucleotide pairs, 2

O

- Oettinger, Marjorie, 35, 40
- Ohshima, Koichi, 153
- Okada, Nori, 153
- Oligonucleotides, Southern blotting, 73-79
- open reading frame (ORF), 55
- ORF (open reading frame), 55
- ORF1 protein, 85
- ORF1p protein, 55, 137
- ORF2p protein, 55, 136
- organisms, model, 29-33
- Orkin, Stuart, 61
- Ostertag, Eric, 88, 121
 - mouse mechanism of inversion formation, 125-127
 - mouse model of L1 retrotransposition, 122
 - non-repetitive sequence insertions, 127-131
 - retrotransposition cassette, 121

P

- P-element (*Drosophila melanogaster*)
 - horizontal transmission, 26
 - hybrid dysgenesis, 21-22
- Pardue, Mary Lou, 35
- Perlman, Phil, 35-38, 93
- phage clones, 87
- phage library, 77
- Phusion polymerase, 146
- Pickeral, Oksana, 88
- piggyBac* transposon, 20
- PIPSL gene, 153-156
- piRNAs (Piwi-interacting RNAs), 204-205
- plant transposons, 41

- Plasmodium falciparum*, hemoglobin genes, 62
- Plasterk, Ron, 35, 41
- poly A tails (non-LTR retrotransposons), 6
- polymorphisms, 61
- Poulter, Russell, 201
- pPolII (RNA polymerase II) promoter, 141, 165
- Prak, Nina Luning, 141-143
- predictions for the future, 221-223
- processed pseudogene formation, 190
- promoters, CAG (beta-actin), 163
- proportions, transposable elements, 8
- protein-coding regions, 1-2
- proteins
 - APOBEC3, 202-203
 - L1
 - locating in human cells, 137-139
 - ORF1p, 137
 - ORF2p, 136
 - MeCP2, 205
- purifying selection, 187-189

Q-R

- R2Bm retrotransposition, 89-91
- random insertion, Tn7
 - transposon, 25
- recombination
 - DNA polymorphisms, 61
 - non-allelic homologous, 192-193
- renaturation of DNA, 50
- repetitive DNA, 43, 49

research (future of)

- biochemical characterization of retrotransposons, 218-220
- genome-wide analysis of retrotransposition events, 209-217
- role of retrotransposition in disease, 217-218

retrotransposition

- Alu elements, 180
- assays of full-length L1s of Ta subset, 106-107
- cassettes, EGFP substituted for neo gene, 121
- cell cultures, 86, 93-99
- distribution of activity in L1s, 107-112
- embryonic, 141-143
- future research
 - biochemical characterization of retrotransposons*, 218-220
 - genome-wide analysis of events*, 209-217
 - role in disease*, 217-218
- HERVs (human endogenous retroviruses), 182
- host factors, 201
 - APOBEC3 proteins*, 202-203
 - epigenetic effects*, 205-206
 - inhibition of non-LTR retrotransposons by small RNAs*, 203-205
 - Poulter and Han discoveries*, 201-202
- influence of cellular stress, 220
- L1, 151
- Ostertag's mouse model, 122

R2Bm, 89-91

SVA elements

- sequence of events*, 127-130
- SVA alternative splicing*, 174-176

retrotransposons, 5-11

- biochemical characterization, 218-220
- contrast to DNA transposons, 11-14
- effect on genome evolution, 14-18
- effect on mammalian genomes, 187
 - 3' and 5' transductions*, 194-195
 - antisense promoter effects*, 196
 - deletions resulting from L1/L1-mediated insertions*, 194
 - endonuclease-independent insertions*, 199
 - gene expression*, 195-196
 - in trans retrotransposition of Alu, SVA, and mRNA*, 189-192
 - L1 chimeras and template switching*, 197
 - non-allelic homologous recombination*, 192-193
 - purifying selection on retrotransposon distribution*, 187-189
 - somatic insertions*, 199-200
 - X chromosome inactivation*, 198
- LTR, 184-186
- multiple insertion points, 44
- retroviral integration, 38

retroviruses, 11, 32
 endogenous, 181-183
 reverse transcriptase activity, 50
 reverse transcriptase, 1, 9
 activity in full-length L1s of Ta subset, 105-106
 activity in retroviruses, 50
 isolation of active human transposable elements, 81-83
 oldest known, 11
 Reznikoff, Bill, 35-37
 ribonucleoprotein particles (RNPs), 53
 RNA
 composition, 9
 contrast with DNA, 9
 inhibition of non-LTR retrotransposons, 203-205
 L1s, 52
 RNA polymerase II (pPolII)
 promoter, 141, 165
 RNPs (ribonucleoprotein particles), 53
 Rubin, Gerry, 21, 35
 Rykowski, in situ hybridization of Alu elements, 188

S

S subfamily (Alu elements), 179
S. cerevisiae, 25
 model organism, 29
 Ty elements, 39
S. pombe (*Schizosaccharomyces pombe*)
 model organism, 29
 Tf mobile elements, 30
 Tf1 retrotransposon, 40
 Sakaki, Yoshi, 54
 salmon transposons, 22
 Sandmeyer, Suzanne, 35-39
 Sanger, Fred, 68

Sassaman, Donna, cloning of full-length L1s of Ta subset
 retrotransposition assays, 106-107
 reverse transcriptase activity, 105-106
 Schatz, David, 35, 40
Schizosaccharomyces pombe (*S. pombe*)
 model organism, 29
 Tf mobile elements, 30
 Tf1 retrotransposon, 40
 scientists, 35-41
 Scott, Alan, 54
 selection, transposable elements, 207-208
 Seleme, Maria del Carmen, 145-147
 SETMAR, 46
 Shapiro, Larry, 71
 short interspersed elements (SINEs), 49
 Shustak, Josh, 109
 SINEs (short interspersed elements), 49
 insertions, 45
 ultraconserved, 185-186
 Singer, Maxine, 49
 site-specific recombination, 40
 Skowronski, Jacob, 52
 Sleeping Beauty DNA transposon, 22-24
 small RNAs, 203-205
 Smit, Arian, 183
 Smith, Lucille, 59
 Smithies, Oliver, 117
 solo LTRs, 29
 somatic insertions
 effects of retrotransposition, 199-200
 lymphocyte DNA, 213

Southern blotting, 65-66, 73-79
Southern, Ed, 65
species diversification,
 transposable elements, 46
Spradling, Allan, 21
Stamatoyannopoulos, George, 60
Styles, C.A., 56
subfamilies, Alu elements, 179
superfamilies, 33
survival, transposable elements,
 207-208
SVA elements
 alternative splicing (Hancks),
 174-176
 exonization (Hancks), 176-177
 gene-trapping (Hancks),
 176-177
 in trans retrotransposition by
 L1s, 189-192
 insertions, 207
 sequence of retrotransposition
 events, 127-130
Swergold, Gary, 157, 209
syncytin genes, 183

T

T (thymine monophosphate), 9
Ta (L1 RNAs), 52
TAIL-PCR (Thermal Asymmetric
 Interlaced PCR) technique, 150
target primed reverse
 transcription (TPRT) reaction,
 31, 91, 126
target sites, 19
taxonomic classification, 45
Temin, Howard, 50
template switching,
 L1 chimeras, 197
T_F subfamily (L1s), 117
 insertions in mouse, 117-118
 mobile elements, 30

Tf1 element, 30
Tf1 retrotransposon, 40
Thermal Asymmetric
 Interlaced PCR (TAIL-PCR)
 technique, 150
thymidine kinase (TK) poly A
 signal, 121
thymine monophosphate (T), 9
TK (thymidine kinase) poly A
 signal, 121
Tn series transposons, 24
Tn10 transposition, 36
Tn5 transposition, 37
Tn7
 insertion site, 37
 transposon, 24
TnsD transposon, 25
Tol2 transposon, 21
TPRT (target primed reverse
 transcription) reaction, 31,
 91, 126
trans mobilization,
 Alu elements, 180
trans-preference, 76
transposable elements, 2
 selection and function, 207-208
 species diversification, 46
Transposagen, 131
transposition mechanisms
 biochemistry, 37
 DNA transposons, 20-26
transposons, 5-6
 C. elegans, 41
 classification by transposition
 mechanism, 20-26
 composition, 7
 contrast to retrotransposons,
 11-14
 control elements, 19
 effect on genome evolution,
 14-18

- horizontal transmission, 26-27
- mammalian genomes, 183
- mobility mechanism, 7
- proportions in various organisms, 8
- tumors, *de novo* L1
 - insertions, 211
- twin-priming hypothesis (Ostertag), 125
- Ty elements
 - LTR-transposons, 29
 - S. cerevisiae*, 39
 - Ty5 element, 30
- types of mobile DNA
 - DNA transposons. *See* transposons
 - effect on genome evolution, 14-18
 - retrotransposons, 5-14

U-V

- U (uridine monophosphate), 9
- U6-L1 chimeras, 197
- ultraconserved SINEs, 186
- uridine monophosphate (U), 9
- V(D)J recombination, 40
- Variable Number Tandem Repeat (VNTR) region, 174

- varieties of mobile DNA. *See* types of mobile DNA
- Varmus, Harold, 35
- Venter, Craig, 43
- VNTR (Variable Number Tandem Repeat) region, 174
- Voytas, Dan, 35, 39

W

- Wallace, Bruce, 73
- Watson, James, 43
- Weichenrieder, Oliver, 219
- Wessler, Sue, 35, 41
- Wong, Corinne, 68-69
- wooly mammoth, analysis of transposable elements, 46
- Worton, Ron, 86

X-Y-Z

- X chromosome inactivation, 198
- Xist gene, 198
- Y subfamily (Alu elements), 179
- Yang, Nuo, 157
- Young, Bill, 59
- Youssoufian, Hagop, 63-66
- zorros (*C. albicans*), 31, 202