THE ORIGINS OF GENOME ARCHITECTURE

Michael Lynch

Indiana University



Contents

Preface xiii

1 The Origin of Eukaryotes 1

Entry into the DNA World 4

A viral origin of DNA? 6 Membranes early or late? 7

The Three Domains of Cellular Life 9

Tree, ring, or web of life? 13 Eukaryotic gene acquisition from endosymbionts 14 The unity of the eukaryotes 15

The Stem Eukaryote 16

The Eukaryotic Radiation 18

The unikonts and bikonts 21 A eukaryotic big bang? 22

Genome Repatterning and the Eukaryotic Radiation 23

A Synopsis of the First 2 Billion Years of Biology 26

69

2 Genome Size and Organismal Complexity 29

Genome Size and Complexity 30 The Selfish-DNA and Bulk-DNA Hypotheses 32 The Metabolic Cost of DNA 34 Directional Mutation Pressures on Genome Size 35 Population Size and the Mutational Hazard of Excess DNA 39

3 The Human Genome 43

Gene Number 43 Introns and Exons 49 **Regulatory DNA** 51 **Mobile Genetic Elements** 56 LINEs and SINEs 57 LTR elements 59 60 Transposons 60 **P**seudogenes The Unique Trajectory of Human Evolution 63 **4** Why Population Size Matters Random Genetic Drift at a Neutral Locus 70

 The Effects of Selection
 74

 The Importance of Linkage
 76

 The theory of genetic draft
 78

 Empirical considerations
 81

 The Three Genomic Perils of Evolving Large Body Size
 83

 Smaller population size
 84

 Higher mutation rates
 84

 Reduced recombination in large genomes
 86

The Global Effective Population Sizes of Species 90

Sources of bias 92 Resolution of the "paradox of variation" 95

Mutation as a Weak Selective Force 97

5 Three Keys to Chromosomal Integrity 101

Origins of Replication 102

ORI specification 104 Evolutionary consequences 105

Centromeres 107

The centromeric-drive hypothesis 109 Centromere positioning 112 Centromere definition 113

Telomeres 116

6 The Nucleotide Composition Landscape 121

The Molecular Spectrum of Mutational Effects 122

Biased Gene Conversion 127

Evolutionary Consequences of Replication 133

Isochores 136

The Nonadaptive Basis of Genome-Wide Nucleotide Composition Variation 139

Codon Usage Bias 141

Prokaryotes 144 Eukaryotes 146 Unsolved problems 149

7 Mobile Genetic Elements 151

Non-LTR Retrotransposons 153

The IR system of hybrid dysgenesis 157

LTR Retrotransposons 159 The gypsy element 163
Transposons 164 The P element 166
Rates of Insertion and Fitness Effects 167
Regulation of Mobile Element Activity169The opportunity for self-regulation170Host-encoded mechanisms for mobile element containment172
The Population Biology of Mobile Elements174Conditions for establishment174Conditions for element number stabilization179Insights from population surveys181Mobile elements and species extinction185
Positive Effects of Mobile Element Insertions 189

8 Genomic Expansion by Gene Duplication 193

The Evolutionary Demography of Duplicate Genes 194 Origins of segmental duplications 199 Whole genome duplication 202 Mechanisms for the Preservation of Duplicate Genes 209 Neofunctionalization 209 The masking effect of duplicate genes 213 Subfunctionalization 215 The Fates of Duplicated Protein Sequences 219 Patterns of molecular evolution 250 The Case for Subfunctionalization 225 Speciation via the Divergent Resolution of Duplicate Genes 228

9 Genes in Pieces 237

The Spliceosome(s) 238 The Introns Early–Introns Late Debate 242 A Group II Origin? 244 Mechanisms of Origin and Loss 247 The Intrinsic Cost of Introns 251 253 Introns and Population Size **Rates of Intron Gain and Loss** 256 262 Adaptive Exploitation of Introns Modifiers of the recombination rate 262 Alternative splicing 263 Messenger RNA surveillance 265

10 Transcript Production 271

Transcription and Translation Initiation272The Premature Translation Initiation Problem276Transcription Termination282The Deployment of transSplicing287287Evolution of Modular Gene Organization292The passive emergence of modularity294The demise of operons297

11 Expansion and Contraction of Organelle Genomes 303

Establishment and Stabilization of Organelles 305 The early stages of mitochondrial evolution 305 Intergenomic transfer 308 The origins of plastids 310 Genome Content and Organization 312 The Population Genetic Environment 316 Mutation 317 Transmission and recombination 321 Effective population size 323

Mutation Pressure and the Diversification of Organelle Genomes 326 The proliferation of noncoding DNA 327 Modifications of the genetic code 330 Messenger RNA editing 333 Muller's Ratchet and the Mitochondrion 336 Insights from More Recent Endosymbioses 342 **17** Sex Chromosomes 347 The Origins of Sex Chromosomes 349 The Population Genetic Environment 351 Degeneration of the Y Chromosome 354 **Retailoring the X Chromosome** 358 **13** Genomfart 363 The Origins of Modern Evolutionary Theory 364

Evolution as a Science in the Post-Darwinian World 366 Nothing in Evolution Makes Sense Except in the Light of Population Genetics 370 The Passive Emergence of Genomic Complexity by Nonadaptive Processes 372 Do the Roots of Organismal Complexity Also Reside in Nonadaptive Processes? 377 An Entrée to Understanding the Origins of Genomic Determinants of Organismal Complexity 380 Evolvability 386 **Closing Comments** 389 Glossary 391 Literature Cited 401 471 Author Index

Subject Index 483