

A stylized DNA double helix structure is the central focus, with one strand colored red and the other blue. The background is dark grey. Various biological illustrations are scattered around the DNA: a black maple leaf in the top left, a green frog in the middle left, a black worm in the middle right, a black insect (possibly a fly or beetle) in the bottom right, a black fish (possibly a salmon) in the bottom right, and several black, elongated, rod-like structures in the bottom left.

THE EVOLUTION OF THE GENOME

Edited by
T. Ryan Gregory

CONTENTS

PART I

The C-value Enigma

1 Genome Size Evolution in Animals

T. Ryan Gregory

Why Should Anyone Care about Genome Size?	4
Genome Size in Animals: A Historical Perspective	5
The Discovery of DNA	5
“A Remarkable Constancy” and the Origin of the “C-value”	6
The C-value Paradox	7
The Modern View: From Paradox to Puzzle	9
The State of Knowledge of Animal Genome Size	10
The <i>Animal Genome Size Database</i>	10
Patterns of Variation	12
Vertebrates (and Nonvertebrate Chordates)	12
Invertebrates	16
Intraspecific Variation in Animals	24
Mechanisms of Genome Size Change	27
“Selfish DNA” and the Spread of Transposable Elements	27
“Junk DNA” and the Accumulation of Pseudogenes	29
Introns	31
Chromosome-Level Events	31
Polyploidy, C-value, and Genome Size	32
Sentinel Sequences and Global Forces	34
Insertion–Deletion Biases	35
Genome Size and G+C Content	39

Assessing the Directionality of Animal Genome Size Evolution	40
Genome Size and Cell Size	41
Explaining the Correlation	48
Genome Size and Organismal Phenotypes	52
Body Size	53
Metabolism	54
Developmental Rate	60
Developmental Complexity	62
Organ Complexity	64
Longevity	66
Measuring Animal Genome Sizes	67
Flow Cytometry	67
Feulgen Microdensitometry	68
Feulgen Image Analysis Densitometry	69
What about Genome Sequencing?	69
Concluding Remarks and Future Prospects	71

2 Genome Size Evolution in Plants

Michael D. Bennett and Ilia J. Leitch

A Brief History of Genome Size Study in Plants	90
The First Estimates of DNA Amounts	90
The Main Areas of Focus of Early Genome Size Studies	92
Impact of the Molecular Revolution on Genome Size Research	93
Genome Size Studies in the Post-Genomic Era	94
The State of Knowledge Regarding Plant Genome Sizes	94
C-values in Cyberspace: Development of the <i>Plant DNA C-values Database</i>	95
Uses and Users of the <i>Plant DNA C-values Database</i>	96
Patterns in Plant Genome Size Evolution	98
The Extent of Variation across Plant Taxa	98
Genome Size in a Phylogenetic Context	99
How Do Plant Genome Sizes Evolve?	103
Sequences Responsible for the Range of Genome Sizes Encountered in Plants	105
What Triggers the Spread of Transposable Elements?	106
Satellite DNA	109
Genome Size Increase by Polyploidy	109
Mechanisms of Genome Size Decrease	112
Key Correlates of Genome Size across Plant Species	114
Early Work on the Phenotypic Consequences of Genome Size Variation in Plants	114
Chromosome Size	116

Nucleus Size	118
Cell Size	118
Cell Division Rate	119
Causation at the Cellular Level: The Nucleotype Concept	121
Pollen and Seeds	122
Minimum Generation Time and Developmental Lifestyle	124
Physiology and Climate Response	128
Ecological and Evolutionary Implications of Genome Size Variation	129
Geographical Distribution and the Large Genome Constraint Hypothesis	130
Genome Size and Plant Response to Human Environmental Change	132
Intraspecific Variation in Genome Size	134
Overview of Intraspecific Variation	134
Genuine Intraspecific Variation in Angiosperms	136
The Special Case of Maize	137
Genuine Intraspecific Variation in Nonangiosperms	138
Intraspecific Variation and Speciation	139
The Mystery of DNA Constancy	140
Methodology for Estimating Genome Size in Plants	141
Chemical Extraction and Reassociation Kinetics	141
Feulgen Microdensitometry	142
Feulgen Image Analysis Densitometry	144
Flow Cytometry	145
Complete Genome Sequencing	147
Some Comments on Plant Genome Size Standards	148
Concluding Remarks and Future Prospects	149
Expansion of the Plant Genome Size Dataset	150
Mechanistic Questions	151
Ecological and Environmental Questions	151
Evolutionary Questions	151

PART II

The Evolution of Genomic Parasites

3 Transposable Elements

Margaret G. Kidwell

A Brief History of the Study of Transposable Elements	165
The Discovery of Transposable Elements	166
Early TE Studies in Bacteria	167

Early TE Studies in Fungi	167
Early TE Studies in Plants	168
Early TE Studies in Animals	168
A Recent Explosion of New Information from DNA Sequencing	169
Who Cares about Transposable Elements?	170
How Are TEs Classified?	170
Autonomous and Nonautonomous Elements	170
Classification Based on Mode of Transposition	171
The Relationship Between Class I and II Elements	179
Hallmarks of TE Sequences	180
Dispersed Multigene Families	180
Target Site Duplications	180
Terminal Repeats	180
Coding Regions and Motifs	181
Fixed and Segregating Insertion Sites	182
Methods Used in the Identification and Study of TEs	182
Genetic Analysis of Naturally Occurring Unstable Mutations	182
Methods of Molecular Analysis	183
Data from Genome Sequencing Projects	184
Reconstruction of Ancestral TEs from Incomplete Contemporary Copies	185
Databases for Repetitive DNA Sequences	185
Phylogenetic Analysis	185
Applications of TEs to Other Areas of Biology	186
Transformation Systems Based on Transposable Elements	186
Transposable Element Mutagenesis and Gene Tagging	186
Transposable Elements as Markers in Evolutionary Studies	187
The Use of Mobile Introns for Targeted Gene Manipulation	188
The Prevalence of TEs in Eukaryotic Genomes	188
Ancient Origins	188
Present-Day Prevalence	189
Examples of Common TEs in Familiar Organisms	189
The Distribution of TEs Within Genomes	193
Selection as a Mechanism for Reducing TE Copy Number	193
The Role of Recombination in Determining TE Distributions	194
TE Frequencies in Euchromatin and Heterochromatin	195
Inter- and Intrachromosomal Variation in TE Density	195
TE Target Site Specificities	196
The Dynamics of TE Evolution	197
Long-Term Evolution and TE Life Cycles	197
Mechanisms of Spread and Loss	198

Regulation of TE Activity	201
Element-Mediated Regulation	201
Host-Mediated Regulation of TE Activity	202
Repeat-Induced Gene Silencing	203
Disruption of TE Regulation by Environmental Stresses	204
A Continuum of TE-Host Interactions from Parasitism to Mutualism	205
TEs as Mutagens and Sources of Genomic Variation	205
Coding Sequences and the Evolution of Novel Host Genes	205
Introns	207
Alternative Splicing	207
Gene Regulatory Sequences	207
Telomeres	208
Centromeres	209
Transduction	210
Genome Size	211
Host Genome Structure	212
Concluding Remarks and Future Prospects	213
 4 B Chromosomes	
<i>Juan Pedro M. Camacho</i>	
A Brief History of the Study of B Chromosomes	224
The Name Game	224
What is a "B Chromosome"?	225
B Chromosomes as Genomic Parasites	225
The Frequency of B Chromosome Infection	227
How Widely Distributed Are B Chromosomes?	227
The Likelihood of Infection	230
Variation in the Intensity of Infection	230
The Biology of B Chromosomes	233
Size	233
Structure	234
Composition	234
Meiotic Behavior	240
Mechanisms of Drive	242
Centromeric Drive	245
The Origin(s) of B Chromosomes	246
Derivation from A Chromosomes	247
From Which Species?	248
From Which Population?	249

From Which Chromosome?	249
Accumulation of Transposable Elements	250
Similarities to Sex Chromosomes: Analogy or Homology?	251
Drive to Survive!	252
Interactions with the Host Genome	252
Effects on Gene Expression	253
Recombination in A Chromosomes	253
The Odd–Even Effect	256
Host Resistance	258
Host Tolerance	260
Interactions with the Host Organism	261
Impacts on the Cellular and Organismal Phenotypes	261
B Chromosomes and Host Reproductive Mode	263
Population Dynamics	264
Parasite Prudence	264
The Dynamics of B Chromosome Evolution	265
The B Chromosome Life Cycle	265
The Life Spans of B Chromosomes	268
Can B Chromosomes Become Beneficial?	270
Can B Chromosomes Integrate into the Standard Genome?	271
Concluding Remarks and Future Prospects	273

PART III

Duplications, Duplications...

5 Small-Scale Gene Duplications

John S. Taylor and Jeroen Raes

The Long Pedigree of Gene Duplication Research	290
Early Chromosomal Studies	290
Studies at the Protein Level: Evidence for Gene Duplication and Divergence from Isozymes	294
The Advent of PCR and DNA Sequencing	294
Gene Duplications in the Post-Genomic Era	296
Mechanisms of Gene Duplication	298
Aneuploidy	298
Duplicative Transposition	298
Local Tandem Duplication	300
The Life and Death of Gene Duplicates in the Genome	301
The Birth and Death of Gene Duplicates	301

The Evolution of Gene Families	302
The Contribution of Gene Duplication to Genome Structure	305
What Happens to Duplicated Genes?	306
Nonfunctionalization	306
Those That Beat the Odds	306
Sequence Divergence	308
Changes in Expression Patterns	310
Neofunctionalization	312
Subfunctionalization	314
Reversion	316
Hox Gene Duplication and the Evolution of Animal Development	316
The General Evolutionary Importance of Gene Duplications	319
Concluding Remarks and Future Prospects	320
 6 Large-Scale Gene and Ancient Genome Duplications	
<i>Yves Van de Peer and Axel Meyer</i>	
Historical Perspectives on the Importance of Large-Scale Duplications	330
Mechanisms of Large-Scale Duplication	331
Autopolyploidy	331
Allopolyploidy	332
Aneuploidy	333
Block Duplications	334
Tandem Duplications	334
How Large-Scale Gene Duplications Are Studied	334
Identification of Block Duplications	334
The Map-Based Approach	335
Hidden Duplications, Ghost Duplications, and Multiplicons	340
Genomic Profiles: An Extension to the Map-Based Approach	343
Dating Duplication Events	344
Absolute Dating Based on Synonymous Substitutions	345
Protein-Based Distances	345
Dating by Phylogenetic Means	346
Putting Theory into Practice: Evidence for Large-Scale Gene Duplication Events	347
1R/2R: Genome Duplications in Vertebrates	348
3R: An Additional Round of Genome Duplication in Teleost Fishes	350
Ancient Genome Duplications in Plants	354
Large-Scale Duplications in the Evolutionary Process	356
The Maintenance of Duplicated Genes	356

Which Genes Are Maintained, and Why?	358
The Maintenance of Duplicated Genomes	359
Speciation and Divergent Resolution	361
Concluding Remarks and Future Prospects	363

PART IV

...And More Duplications

7 Polyploidy in Plants

Jennifer A. Tate, Douglas E. Soltis, and Pamela S. Soltis

History of the Study of Polyploidy in Plants	372
Types of Polyploids	373
Frequency of Polyploids	379
How Common Is Polyploidy in Plants?	379
The Frequency of Allopolyploidy versus Autopolyploidy	383
Polyploid Formation and Establishment	384
Mechanisms and Chances of Formation	384
Likelihood of Establishment	388
Multiple Origins of Polyploid Species	390
The Rule, Not the Exception	390
Genotypic and Phenotypic Consequences of Multiple Origins	394
A Case of Particular Interest: The Arctic Flora	395
Impacts of Polyploidization at the Cellular and Organismal Levels	396
Cell Size	397
Reproductive Biology	398
Physiology and Development	400
Geographic Distribution	401
Plant–Animal Interactions	402
Impacts of Polyploidization at the Genome Level	404
Genomic Rearrangements	405
Genomic Downsizing and “Diploidization”	407
The Fates of Plant Genes Duplicated by Polyploidy	408
Transposable Elements	413
Nuclear–Cytoplasmic Interactions	413
Concluding Remarks and Future Prospects	414

8 Polyploidy in Animals

T. Ryan Gregory and Barbara K. Mable

The Origins and Classification of Polyploid Animals	428
Autopolyploidy and Allopolyploidy	428

Identifying Polyploids	429
Chromosome Number and Nuclear DNA Content	429
Cell and Nucleus Size	430
Meiotic Chromosome Behavior	430
Protein Electrophoresis	432
Why Is Polyploidy Less Common in Animals than in Plants?	433
Disruption of Sex Determination	434
Degenerate Sex Chromosomes and Dosage Compensation	436
Impediments to Meiotic Disjunction	437
Interploidy Crosses and Triploid Sterility	438
Disruption of Development	438
Not Enough Hybridization . . . or Maybe Too Much?	439
Nucleotypic Constraints	440
Polyploidy and Unisexuality	440
Definitions	441
Why Are Polyploidy and Unisexuality Linked?	442
Polyploidy in Vertebrates	444
Jawless Fishes	444
Cartilaginous Fishes	445
Lungfishes	445
Chondrosteans	446
Teleosts	447
An Entire Family of Polyploids: Salmonidae	448
Another Entirely Polyploid Family: Catostomidae	451
Several Species of Polyploids: Cyprinidae	452
Special Cases: Poeciliidae	456
Miscellaneous Polyploid Fishes	459
Amphibians	461
Diploid–Polyploid Species Pairs in Frogs	462
Polyploidy in Salamanders: Ambystomatidae	469
A Polyploid Family of Salamanders (?): Sirenidae	471
Rare Triploidy in Newts	472
Reptiles	472
The Genus <i>Cnemidophorus</i>	473
Other Polyploid Reptiles	475
Mammals and Birds	475
Vertebrate Polyploids: A Summary	477
Polyploidy in Invertebrates	478
Crustaceans	478
Water Fleas (Order Cladocera)	479
Brine Shrimp (Order Anostraca)	480

Insects	480
Molluscs	486
Bivalves	487
Gastropods	488
Annelids	490
Flatworms	493
Miscellaneous Invertebrates	494
Nematodes	494
Rotifers	495
Tardigrades	495
Arachnids	496
Polyploidy and Geographic Distribution in Invertebrates	496
Unisexuality	497
Genetic Factors	498
Physiology, Development, and Ecology	499
History	499
The Evolutionary Fate of Polyploids	500
Concluding Remarks and Future Prospects	501

PART V

Sequence and Structure

9 Comparative Genomics in Eukaryotes

Alan Filipinski and Sudhir Kumar

The Early History of Comparative Eukaryotic Genomics	522
The Basics of Eukaryotic Chromosome Structure	522
Karyotyping: The Beginning of Comparative Genomics	524
Genome Architecture	531
Working with Eukaryotic Genomes	533
Mapping: Genetic and Physical	534
Sequencing: The Holy Grail of Comparative Genomics	534
Annotation: Making Biological Sense of the Letters	535
The Genesis of Large-Scale Sequencing Projects for Eukaryotes	536
Sequencing the Human Genome: The Most Ambitious Idea	536
Private versus Public Efforts	537
Genome Sequencing in Fungi	541
<i>Saccharomyces cerevisiae</i> : The First Eukaryote to Be Sequenced	541
Other Fungal Sequencing Projects	542

<i>Caenorhabditis elegans</i> and <i>Drosophila melanogaster</i> : The First Animal Genomes to Be Sequenced	544
The Worm Project	544
The Fruit Fly Project	546
The Human Genome Project	546
Genome Variation in Human Populations	552
Pufferfish Synergy	554
The Mouse and Rat Genomes: The Rise of Modern Mammalian Comparative Genomics	555
Genome Sequencing in Plants and Their Pathogens	559
Comparative Genomics of <i>Arabidopsis</i>	559
The Rice Genome	560
The Rice Blast Fungus: <i>Magnaporthe grisea</i>	562
Other Invertebrate Animal Genomes	562
The Mosquito Genome	562
The Sea Squirt: A Primitive Chordate	563
Genomewide Duplications in Vertebrates?	564
Protist Genomes	564
<i>Encephalitozoon cuniculi</i> : A Parasitic Eukaryote with a Tiny Genome	564
<i>Plasmodium</i> : The Malaria Pathogen	565
<i>Dictyostelium</i> : The "Slime Mold"	566
Comparative Genomics and Phylogenetics in Eukaryotes	566
Concluding Remarks and Future Prospects	569
Complete Genome Sequencing	569
Partial-Genome Comparisons	572
The Tree of Life	573
The Charter of Genomics	574

10 Comparative Genomics in Prokaryotes

T. Ryan Gregory and Rob DeSalle

What Is a Prokaryote?	586
Classifying Prokaryotes the Old-Fashioned Way	586
The Deepest Split of All?	589
The Rise of Complete Prokaryotic Genome Sequencing	590
From Viruses to Venter	590
The Prokaryote Genome Sequencing Explosion	592
General Insights about Prokaryote Genomes	593
Genome Organization: Assumptions and Exceptions	594
Structure of Prokaryotic Chromosomal DNA	596
DNA Replication in Prokaryotes	597

Gene Content	597
Gene Order: Plasticity and Stability	601
Base Pair Composition	603
Repeat Content	606
Horizontal Transfer of Genetic Material	606
Identifying and Characterizing Horizontal Transfers	607
Transfer from Viruses to Bacteria: Prophages in Bacterial Genomes	608
Transfer among Bacteria	610
Transfer across Domains	612
Implications for Prokaryote Evolution and the Study Thereof	615
Highlights from Specific Prokaryote Genome Sequencing Projects	616
<i>Haemophilus influenzae</i> (1995)	617
<i>Methanocaldococcus jannaschii</i> (1996)	618
<i>Escherichia coli</i> (1997)	619
<i>Mycobacterium tuberculosis</i> (1998)	621
<i>Deinococcus radiodurans</i> (1999)	622
<i>Vibrio cholerae</i> (2000)	623
<i>Streptococcus</i> spp. (2001)	624
<i>Streptomyces coelicolor</i> (2002)	627
<i>Bacillus anthracis</i> (2003)	628
<i>Wolbachia pipiens</i> (2004)	630
The Evolution of Genome Size in Prokaryotes	631
Factors that Limit Prokaryote Genome Size	634
Mechanisms of Genome Size Increase	636
Genome Reductions in Obligate Parasites and Endosymbionts	637
Bacteria, Organelles, or Something in Between?	640
The Minimal Genome Concept	641
Breaking it Down	642
Building it Up	644
Genomic Insights into Prokaryotic Abundance and Diversity	645
What Is a Prokaryotic Species?	645
Genetic Delineation of Species	646
Genomic Perspectives on Prokaryote Diversity	647
Shotgun First, Ask Questions Later	649
Applications of Prokaryote Genomics	651
Medicine	651
Industry and the Environment	654
Agriculture and Food Production	655
Evolutionary Biology	655
Concluding Remarks and Future Prospects	656
Which Prokaryotes Can (or Cannot) Be Sequenced?	657
Policy Issues: From Sequence Completeness to Bioterrorism	658
Prokaryote Genomics: The End of the Beginning	660

PART VI

The Genome in Evolution

11 Macroevolution and the Genome

T. Ryan Gregory

Part One—Macroevolutionary Theory and Genome Evolution	680
A Brief History of Evolutionary Theory	680
From Darwin to Neo-Darwinism	680
Genomes, Fossils, and Theoretical Inertia	682
Is a Theory of Macroevolution Necessary?	684
Microevolution, Macroevolution, and Extrapolationism	684
Critiques of Strict Extrapolationism	686
Reductionism in the Post-Genomic Era	692
The Structure of Macroevolutionary Theory	693
Group Selection	694
Species Selection: Concepts and Challenges	694
Aggregate versus Emergent Characters	695
Selection versus Sorting	696
Bottom-up Processes: The Effect Hypothesis versus Emergent Fitness	696
Top-down Processes: Context-dependent Sorting	697
Macroevolutionary Theory: A Summary	698
A Macroevolutionary Look at the Genome	698
Did the Genome Originate by Group Selection?	698
Genomic Parasites Require a Hierarchical Interpretation	700
Genome Size, Emergent Fitness, and an Upside-down Acid Test	704
Context-dependent Sorting of Genes and Nongenes	706
Genomes in the Evolutionary Hierarchy	707
Part Two—"Nonstandard" Genomic Processes and Major Evolutionary Transitions	708
The Origin of Genomes and Cells	709
The Evolution of Sex	709
The Origin of Eukaryotes	710
Linear Chromosomes	710
Increased Genetic Complexity	711
The Origin of Multicellularity and the Emergence of Complex Metazoa	712
Transposable Elements and Gene Regulation	712
Spliceosomal Introns and Exon Shuffling	712
Gene Duplication and Developmental Complexity	713
The Evolution of Immunity	714

The Origin of Vertebrates	715
Gen(om)e Duplication	715
Silencing and/or Splicing	716
Human Uniqueness	716
Diversity in Gene Expression	717
The Role of Alu Elements	717
Gene Duplication and Back Again	719
Nonstandard Genomic Processes: A Summary	719
Concluding Remarks and Future Prospects	720
From Reductionism to Integrationism	720
A Post-Genomic Evolutionary Synthesis	721
Genomes and the Future of Biology	722

Index	731
-------	-----