

A close-up photograph of a field of tulips. The majority of the flowers are bright yellow, with some in sharp focus and others blurred in the background. A single, vibrant red tulip stands out prominently in the center of the frame, slightly to the left of the middle. The green stems and leaves of the tulips are visible at the bottom and interspersed among the flowers.

WARD C. WHEELER

SYSTEMATICS

A Course of Lectures

 **WILEY-BLACKWELL**

Systematics

Systematics: A Course of Lectures

Ward C. Wheeler

 **WILEY-BLACKWELL**
A John Wiley & Sons, Ltd., Publication

This edition first published 2012 © 2012 by Ward C. Wheeler

Wiley-Blackwell is an imprint of John Wiley & Sons, formed by the merger of Wiley's global Scientific, Technical and Medical business with Blackwell Publishing.

Registered office: John Wiley & Sons, Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK

Editorial offices: 9600 Garsington Road, Oxford, OX4 2DQ, UK
The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, UK
111 River Street, Hoboken, NJ 07030-5774, USA

For details of our global editorial offices, for customer services and for information about how to apply for permission to reuse the copyright material in this book please see our website at www.wiley.com/wiley-blackwell.

The right of the author to be identified as the author of this work has been asserted in accordance with the UK Copyright, Designs and Patents Act 1988.

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, except as permitted by the UK Copyright, Designs and Patents Act 1988, without the prior permission of the publisher.

Designations used by companies to distinguish their products are often claimed as trademarks. All brand names and product names used in this book are trade names, service marks, trademarks or registered trademarks of their respective owners. The publisher is not associated with any product or vendor mentioned in this book. This publication is designed to provide accurate and authoritative information in regard to the subject matter covered. It is sold on the understanding that the publisher is not engaged in rendering professional services. If professional advice or other expert assistance is required, the services of a competent professional should be sought.

Library of Congress Cataloging-in-Publication Data has been applied for
9780470671702 (hardback)
9780470671696 (paperback)

A catalogue record for this book is available from the British Library.

Wiley also publishes its books in a variety of electronic formats. Some content that appears in print may not be available in electronic books.

Set in Computer Modern 10/12pt by Laserwords Private Limited, Chennai, India

For

Kurt Milton Pickett
(1972–2011)
Ave atque vale

Contents

Preface	xv
Using these notes	xv
Acknowledgments	xvi
List of algorithms	xix
I Fundamentals	1
1 History	2
1.1 Aristotle	2
1.2 Theophrastus	3
1.3 Pierre Belon	4
1.4 Carolus Linnaeus	4
1.5 Georges Louis Leclerc, Comte de Buffon	6
1.6 Jean-Baptiste Lamarck	7
1.7 Georges Cuvier	8
1.8 Étienne Geoffroy Saint-Hilaire	8
1.9 Johann Wolfgang von Goethe	8
1.10 Lorenz Oken	9
1.11 Richard Owen	9
1.12 Charles Darwin	9
1.13 Stammbäume	12
1.14 Evolutionary Taxonomy	14
1.15 Phenetics	15
1.16 Phylogenetic Systematics	16
1.16.1 Hennig's Three Questions	16
1.17 Molecules and Morphology	18
1.18 We are all Cladists	18
1.19 Exercises	19
2 Fundamental Concepts	20
2.1 Characters	20
2.1.1 Classes of Characters and Total Evidence	22
2.1.2 Ontogeny, Tokogeny, and Phylogeny	23
2.1.3 Characters and Character States	23
2.2 Taxa	26

2.3	Graphs, Trees, and Networks	28
2.3.1	Graphs and Trees	30
2.3.2	Enumeration	31
2.3.3	Networks	33
2.3.4	Mono-, Para-, and Polyphyly	33
2.3.5	Splits and Convexity	38
2.3.6	Apomorphy, Plesiomorphy, and Homoplasy	39
2.3.7	Gene Trees and Species Trees	41
2.4	Polarity and Rooting	43
2.4.1	Stratigraphy	43
2.4.2	Ontogeny	43
2.4.3	Outgroups	45
2.5	Optimality	49
2.6	Homology	49
2.7	Exercises	50
3	Species Concepts, Definitions, and Issues	53
3.1	Typological or Taxonomic Species Concept	54
3.2	Biological Species Concept	54
3.2.1	Criticisms of the BSC	55
3.3	Phylogenetic Species Concept(s)	56
3.3.1	Autapomorphic/Monophyletic Species Concept	56
3.3.2	Diagnostic/Phylogenetic Species Concept	58
3.4	Lineage Species Concepts	59
3.4.1	Hennigian Species	59
3.4.2	Evolutionary Species	60
3.4.3	Criticisms of Lineage-Based Species	61
3.5	Species as Individuals or Classes	62
3.6	Monoism and Pluralism	63
3.7	Pattern and Process	63
3.8	Species Nominalism	64
3.9	Do Species Concepts Matter?	65
3.10	Exercises	65
4	Hypothesis Testing and the Philosophy of Science	67
4.1	Forms of Scientific Reasoning	67
4.1.1	The Ancients	67
4.1.2	Ockham's Razor	68
4.1.3	Modes of Scientific Inference	69
4.1.4	Induction	69
4.1.5	Deduction	69
4.1.6	Abduction	70
4.1.7	Hypothetico-Deduction	71
4.2	Other Philosophical Issues	75
4.2.1	Minimization, Transformation, and Weighting	75
4.3	Quotidian Importance	76
4.4	Exercises	76

5	Computational Concepts	77
5.1	Problems, Algorithms, and Complexity	77
5.1.1	Computer Science Basics	77
5.1.2	Algorithms	79
5.1.3	Asymptotic Notation	79
5.1.4	Complexity	80
5.1.5	Non-Deterministic Complexity	82
5.1.6	Complexity Classes: P and NP	82
5.2	An Example: The Traveling Salesman Problem	84
5.3	Heuristic Solutions	85
5.4	Metricity, and Untrametricity	86
5.5	NP-Complete Problems in Systematics	87
5.6	Exercises	88
6	Statistical and Mathematical Basics	89
6.1	Theory of Statistics	89
6.1.1	Probability	89
6.1.2	Conditional Probability	91
6.1.3	Distributions	92
6.1.4	Statistical Inference	98
6.1.5	Prior and Posterior Distributions	99
6.1.6	Bayes Estimators	100
6.1.7	Maximum Likelihood Estimators	101
6.1.8	Properties of Estimators	101
6.2	Matrix Algebra, Differential Equations, and Markov Models	102
6.2.1	Basics	102
6.2.2	Gaussian Elimination	102
6.2.3	Differential Equations	104
6.2.4	Determining Eigenvalues	105
6.2.5	Markov Matrices	106
6.3	Exercises	107
II	Homology	109
7	Homology	110
7.1	Pre-Evolutionary Concepts	110
7.1.1	Aristotle	110
7.1.2	Pierre Belon	110
7.1.3	Étienne Geoffroy Saint-Hilaire	111
7.1.4	Richard Owen	112
7.2	Charles Darwin	113
7.3	E. Ray Lankester	114
7.4	Adolf Remane	114
7.5	Four Types of Homology	115
7.5.1	Classical View	115
7.5.2	Evolutionary Taxonomy	115

7.5.3	Phenetic Homology	116
7.5.4	Cladistic Homology	116
7.5.5	Types of Homology	117
7.6	Dynamic and Static Homology	118
7.7	Exercises	120
8	Sequence Alignment	121
8.1	Background	121
8.2	“Informal” Alignment	121
8.3	Sequences	121
8.3.1	Alphabets	122
8.3.2	Transformations	123
8.3.3	Distances	123
8.4	Pairwise String Matching	123
8.4.1	An Example	127
8.4.2	Reducing Complexity	129
8.4.3	Other Indel Weights	130
8.5	Multiple Sequence Alignment	131
8.5.1	The Tree Alignment Problem	133
8.5.2	Trees and Alignment	133
8.5.3	Exact Solutions	134
8.5.4	Polynomial Time Approximate Schemes	134
8.5.5	Heuristic Multiple Sequence Alignment	134
8.5.6	Implementations	135
8.5.7	Structural Alignment	139
8.6	Exercises	145
III	Optimality Criteria	147
9	Optimality Criteria—Distance	148
9.1	Why Distance?	148
9.1.1	Benefits	149
9.1.2	Drawbacks	149
9.2	Distance Functions	150
9.2.1	Metricity	150
9.3	Ultrametric Trees	150
9.4	Additive Trees	152
9.4.1	Farris Transform	153
9.4.2	Buneman Trees	154
9.5	General Distances	156
9.5.1	Phenetic Clustering	157
9.5.2	Percent Standard Deviation	160
9.5.3	Minimizing Length	163
9.6	Comparisons	170
9.7	Exercises	171

10 Optimality Criteria–Parsimony	173
10.1 Perfect Phylogeny	174
10.2 Static Homology Characters	174
10.2.1 Additive Characters	175
10.2.2 Non-Additive Characters	179
10.2.3 Matrix Characters	182
10.3 Missing Data	184
10.4 Edge Transformation Assignments	187
10.5 Collapsing Branches	188
10.6 Dynamic Homology	188
10.7 Dynamic and Static Homology	189
10.8 Sequences as Characters	190
10.9 The Tree Alignment Problem on Trees	191
10.9.1 Exact Solutions	191
10.9.2 Heuristic Solutions	191
10.9.3 Lifted Alignments, Fixed-States, and Search-Based Heuristics	193
10.9.4 Iterative Improvement	197
10.10 Performance of Heuristic Solutions	198
10.11 Parameter Sensitivity	198
10.11.1 Sensitivity Analysis	199
10.12 Implied Alignment	199
10.13 Rearrangement	204
10.13.1 Sequence Characters with Moves	204
10.13.2 Gene Order Rearrangement	205
10.13.3 Median Evaluation	207
10.13.4 Combination of Methods	207
10.14 Horizontal Gene Transfer, Hybridization, and Phylogenetic Networks	209
10.15 Exercises	210
11 Optimality Criteria–Likelihood	213
11.1 Motivation	213
11.1.1 Felsenstein’s Example	213
11.2 Maximum Likelihood and Trees	216
11.2.1 Nuisance Parameters	216
11.3 Types of Likelihood	217
11.3.1 Flavors of Maximum Relative Likelihood	217
11.4 Static-Homology Characters	218
11.4.1 Models	218
11.4.2 Rate Variation	219
11.4.3 Calculating $p(D T, \theta)$	221
11.4.4 Links Between Likelihood and Parsimony	222
11.4.5 A Note on Missing Data	224
11.5 Dynamic-Homology Characters	224
11.5.1 Sequence Characters	225

11.5.2	Calculating ML Pairwise Alignment	227
11.5.3	ML Multiple Alignment	230
11.5.4	Maximum Likelihood Tree Alignment Problem	230
11.5.5	Genomic Rearrangement	232
11.5.6	Phylogenetic Networks	234
11.6	Hypothesis Testing	234
11.6.1	Likelihood Ratios	234
11.6.2	Parameters and Fit	236
11.7	Exercises	238
12	Optimality Criteria—Posterior Probability	240
12.1	Bayes in Systematics	240
12.2	Priors	241
12.2.1	Trees	241
12.2.2	Nuisance Parameters	242
12.3	Techniques	246
12.3.1	Markov Chain Monte Carlo	246
12.3.2	Metropolis–Hastings Algorithm	246
12.3.3	Single Component	248
12.3.4	Gibbs Sampler	249
12.3.5	Bayesian MC ³	249
12.3.6	Summary of Posterior	250
12.4	Topologies and Clades	252
12.5	Optimality versus Support	254
12.6	Dynamic Homology	254
12.6.1	Hidden Markov Models	255
12.6.2	An Example	256
12.6.3	Three Questions—Three Algorithms	258
12.6.4	HMM Alignment	262
12.6.5	Bayesian Tree Alignment	264
12.6.6	Implementations	264
12.7	Rearrangement	266
12.8	Criticisms of Bayesian Methods	267
12.9	Exercises	267
13	Comparison of Optimality Criteria	269
13.1	Distance and Character Methods	269
13.2	Epistemology	270
13.2.1	Ockham’s Razor and Popperian Argumentation	271
13.2.2	Parsimony and the Evolutionary Process	272
13.2.3	Induction and Statistical Estimation	272
13.2.4	Hypothesis Testing and Optimality Criteria	272
13.3	Statistical Behavior	273
13.3.1	Probability	273
13.3.2	Consistency	274
13.3.3	Efficiency	281
13.3.4	Robustness	282

13.4	Performance	282
13.4.1	Long-Branch Attraction	283
13.4.2	Congruence	285
13.5	Convergence	285
13.6	Can We Argue Optimality Criteria?	286
13.7	Exercises	287
IV	Trees	289
14	Tree Searching	290
14.1	Exact Solutions	290
14.1.1	Explicit Enumeration	290
14.1.2	Implicit Enumeration—Branch-and-Bound	292
14.2	Heuristic Solutions	294
14.2.1	Local versus Global Optima	294
14.3	Trajectory Search	296
14.3.1	Wagner Algorithm	296
14.3.2	Branch-Swapping Refinement	298
14.3.3	Swapping as Distance	301
14.3.4	Depth-First versus Breadth-First Searching	302
14.4	Randomization	304
14.5	Perturbation	305
14.6	Sectorial Searches and Disc-Covering Methods	309
14.6.1	Sectorial Searches	309
14.6.2	Disc-Covering Methods	310
14.7	Simulated Annealing	312
14.8	Genetic Algorithm	316
14.9	Synthesis and Stopping	318
14.10	Empirical Examples	319
14.11	Exercises	323
15	Support	324
15.1	Resampling Measures	324
15.1.1	Bootstrap	325
15.1.2	Criticisms of the Bootstrap	326
15.1.3	Jackknife	328
15.1.4	Resampling and Dynamic Homology Characters	329
15.2	Optimality-Based Measures	329
15.2.1	Parsimony	330
15.2.2	Likelihood	332
15.2.3	Bayesian Posterior Probability	334
15.2.4	Strengths of Optimality-Based Support	335
15.3	Parameter-Based Measures	336
15.4	Comparison of Support Measures—Optimal and Average	336
15.5	Which to Choose?	339
15.6	Exercises	339

16 Consensus, Congruence, and Supertrees	341
16.1 Consensus Tree Methods	341
16.1.1 Motivations	341
16.1.2 Adams I and II	341
16.1.3 Gareth Nelson	344
16.1.4 Majority Rule	347
16.1.5 Strict	347
16.1.6 Semi-Strict/Combinable Components	348
16.1.7 Minimally Pruned	348
16.1.8 When to Use What?	350
16.2 Supertrees	350
16.2.1 Overview	350
16.2.2 The Impossibility of the Reasonable	350
16.2.3 Graph-Based Methods	353
16.2.4 Strict Consensus Supertree	355
16.2.5 MR-Based	355
16.2.6 Distance-Based Method	358
16.2.7 Supertrees or Supermatrices?	360
16.3 Exercises	361
 V Applications	 363
17 Clocks and Rates	364
17.1 The Molecular Clock	364
17.2 Dating	365
17.3 Testing Clocks	365
17.3.1 Langley–Fitch	365
17.3.2 Farris	366
17.3.3 Felsenstein	367
17.4 Relaxed Clock Models	368
17.4.1 Local Clocks	368
17.4.2 Rate Smoothing	368
17.4.3 Bayesian Clock	369
17.5 Implementations	369
17.5.1 r8s	369
17.5.2 MULTIDIVTIME	370
17.5.3 BEAST	370
17.6 Criticisms	370
17.7 Molecular Dates?	373
17.8 Exercises	373
 A Mathematical Notation	 374
Bibliography	376
Index	415

Color plate section between pp. 76 and 77

Preface

These notes are intended for use in an advanced undergraduate or introductory level graduate course in systematics. As such, the goal of the materials is to encourage knowledge of core systematic literature (*e.g.* works of Aristotle, Linné, Mayr, Hennig, Sokal, Farris, Kluge, Felsenstein) and concepts (*e.g.* Classification, Optimality, Optimization, Trees, Diagnosis, Medians, Computational Hardness). A component of this goal is specific understanding of methodologies and theory (*e.g.* Cluster Analysis, Parsimony, Likelihood, String Match, Tree Search). Exercises are provided to enhance familiarity with concepts and common analytical tools. These notes are focused on the study of pattern in biodiversity; notions of process receive limited attention and are better discussed elsewhere.

Each chapter covers a topic that could easily be the subject of an entire book-length treatment and many have. As a result, the coverage of large literatures is confined to what I think could be covered in a lecture or two, but may seem brief, idiosyncratic, but hopefully not too superficial. These notes are not meant to be the last word in systematics, but the first.

Students should have basic knowledge of biology and diversity including anatomy and molecular genetics. Some knowledge of computation, statistics, and linear algebra would be nice but not required. Relevant highlights of these fields are covered where necessary.

Using these notes

This is not a fugue. In most cases, sections can be rearranged, or separated entirely without loss of intelligibility. Several sections do build on others (*e.g.* sections on tree searching and support), while others can be deleted entirely if students have the background (*e.g.* sections on computational and statistical basics). The book was developed for a single semester course and, in general, each chapter is designed to be covered in a single 90 minute class period. The chapters on Parsimony, Likelihood, Posterior Probability, and Tree Searching are exceptions, spanning two such classes.

Exercises are of three types: those that can be worked by hand, those that require computational aids, and lastly those that are more suited to larger projects or group work. Hopefully, they are useful.

Acknowledgments

I would like to thank the support of the National Science Foundation (NSF) as well as the National Aeronautics and Space Administration (NASA) and the Defense Advanced Research Projects Agency (DARPA) for supporting the research that has gone into many sections of these notes.

John Wenzel and Adam Kashuba were supportive and persistent in urging me to publish these notes.

All errors, polemics, and disturbing asides are of course my own.

I thank the following people for offering expert advice in reviewing text, suggesting improvements, and identifying errors:

- Benjamin de Bivort, Ph.D.
Rowland Institute at Harvard
- James M. Carpenter, Ph.D.
Division of Invertebrate Zoology
American Museum of Natural History
- Megan Cevasco, Ph.D.
Coastal Carolina University
- Ronald M. Clouse, Ph.D.
Division of Invertebrate Zoology
American Museum of Natural History
- Louise M. Crowley, Ph.D.
Division of Invertebrate Zoology
American Museum of Natural History
- John Denton, M.A.
Division of Vertebrate Zoology
American Museum of Natural History
- James S. Farris, Ph.D.
Molekylärsystematiska laboratoriet
Naturhistoriska riksmuseet
- John V. Freudenstein, Ph.D.
Director of the Herbarium and Museum of Zoology
Herbarium, 1350 Museum of Biological Diversity
- Darrel Frost, Ph.D.
Division of Vertebrate Zoology
American Museum of Natural History
- Taran Grant, Ph.D.
Departamento de Zoologia
Instituto de Biociências
Universidade de São Paulo

- Gonzalo Giribet, Ph.D.
Department of Organismic and Evolutionary Biology
Museum of Comparative Zoology
Harvard University
- Pablo Goloboff, Ph.D.
Instituto Superior de Entomologia
Miguel Lillo
- Lin Hong, M.S.
Division of Invertebrate Zoology
American Museum of Natural History
- Jaakko Hyvönen, Ph.D.
Plant Biology (Biocenter 3)
University of Helsinki
- Daniel Janies, Ph.D.
Biomedical Informatics
The Ohio State University
- Isabella Kappner, Ph.D.
Division of Invertebrate Zoology
American Museum of Natural History
- Lavanya Kannan, Ph.D.
Division of Invertebrate Zoology
American Museum of Natural History
- Arnold G. Kluge, Ph.D.
Museum of Zoology
The University of Michigan
- Nicolas Lucaroni, B.S.
Division of Invertebrate Zoology
American Museum of Natural History
- Brent D. Mishler, Ph.D.
Department of Integrative Biology
and Jepson Herbaria
University of California, Berkeley
- Jyrki Muona, Ph.D.
Division of Entomology
Zoological Museum
University of Helsinki
- Paola Pedraza, Ph.D.
Institute of Systematic Botany
New York Botanical Garden

- Norman Platnick, Ph.D.
Division of Invertebrate Zoology
American Museum of Natural History
- Christopher P. Randle, Ph.D.
Department of Biological Sciences
Sam Houston State University
- Randall T. Schuh, Ph.D.
Division of Invertebrate Zoology
American Museum of Natural History
- William Leo Smith, Ph.D.
Department of Zoology
The Field Museum of Natural History
- Katherine St. John, Ph.D.
Dept. of Mathematics & Computer Science
Lehman College, City University of New York
- Alexandros Stamatakis, Ph.D.
Heidelberg Institute for Theoretical Studies (HITS GmbH)
- Andrés Varón, Ph.D.
Jane Street Capital
- Peter Whiteley, Ph.D.
Division of Anthropology
American Museum of Natural History

List of Algorithms

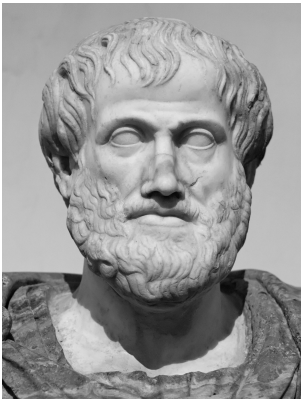
2.1	Farris1974GroupDetermination.	37
5.1	SingleLoop	81
5.2	NestedLoops	81
8.1	PairwiseSequenceAlignmentCost	125
8.2	PairwiseSequenceAlignmentTraceback	126
8.3	PairwiseSequenceAlignmentUkkonen.	130
9.1	UPGMA	158
9.2	DistanceWagner	165
9.3	NeighborJoining	168
10.1	AdditiveDownPass.	176
10.2	AdditiveUpPass	177
10.3	NonAdditiveDownPass	180
10.4	NonAdditiveUpPass	181
10.5	MatrixDownPass	183
10.6	DirectOptimizationFirstPass	194
10.7	DirectOptimizationTraceback.	195
12.1	Viterbi Algorithm	259
12.2	Forward-Backward Algorithm	261
14.1	ExplicitEnumeration.	291
14.2	RecurseAllTrees	291
14.3	Branch-and-Bound Tree Search	293
14.4	BoundRecurse	293
14.5	WagnerBuild	296
14.6	NearestNeighborInterchange	301
14.7	SubTreePruningandRegrafting	302
14.8	TreeBisectionandRegrafting	304
14.9	RatchetRefinement	308
14.10	SectorialSearch.	311
14.11	Rec-I-DCM3	313
14.12	TreeDrift	314
14.13	GeneticAlgorithmTreeSearch	317

Part I

Fundamentals

Chapter 1

History



Roman bust of Aristotle
(384–322 BCE)



Ibn Rushd (Averroes)
(1126–1198)

Systematics has its origins in two threads of biological science: classification and evolution. The organization of natural variation into sets, groups, and hierarchies traces its roots to Aristotle and evolution to Darwin. Put simply, systematization of nature can and has progressed in absence of causative theories relying on ideas of “plan of nature,” divine or otherwise. Evolutionists (Darwin, Wallace, and others) proposed a rationale for these patterns. This mixture is the foundation of modern systematics.

Originally, systematics was natural history. Today we think of systematics as being a more inclusive term, encompassing field collection, empirical comparative biology, and theory. To begin with, however, taxonomy, now known as the process of naming species and higher taxa in a coherent, hypothesis-based, and regular way, and systematics were equivalent.

1.1 Aristotle

Systematics as classification (or taxonomy) draws its Western origins from Aristotle¹. A student of Plato at the Academy and reputed teacher of Alexander the Great, Aristotle founded the Lyceum in Athens, writing on a broad variety of topics including what we now call biology. To Aristotle, living things (*species*) came from nature as did other physical classes (*e.g.* gold or lead). Today, we refer to his classification of living things (Aristotle, 350 BCE) that show similarities with the sorts of classifications we create now. In short, there are three features of his methodology that we recognize immediately: it was functional, binary, and empirical.

Aristotle’s classification divided animals (his work on plants is lost) using functional features as opposed to those of habitat or anatomical differences: “Of land animals some are furnished with wings, such as birds and bees.” Although he recognized these features as different in aspect, they are identical in use.

¹Largely through translation and commentary by Ibn Rushd (Averroes).

Features were also described in binary terms: “Some are nocturnal, as the owl and the bat; others live in the daylight.” These included egg- or live-bearing, blooded or non-blooded, and wet or dry respiration.

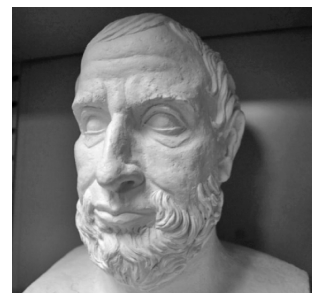
An additional feature of Aristotle’s work was its empirical content. Aspects of creatures were based on observation rather than ideal forms. In this, he recognized that some creatures did not fit into his binary classification scheme: “The above-mentioned organs, then, are the most indispensable parts of animals; and with some of them all animals without exception, and with others animals for the most part, must needs be provided.” Sober (1980) argued that these departures from Aristotle’s expectations (Natural State Model) were brought about (in Aristotle’s mind) by errors due to some perturbations (hybridization, developmental trauma) resulting in “terata” or monsters. These forms could be novel and helped to explain natural variation within his scheme.

- Blooded Animals
 - Live-bearing animals
 - humans
 - other mammals
 - Egg-laying animals
 - birds
 - fish
- Non-Blooded Animals
 - Hard-shelled sea animals: Testacea
 - Soft-shelled sea animals: Crustacea
 - Non-shelled sea animals: Cephalopods
 - Insects
 - Bees
- Dualizing species (potential “terata,” errors in nature)
 - Whales, seals and porpoises—in water, but bear live young
 - Bats—have wings and can walk
 - Sponges—like plants and like animals.

Aristotle clearly had notions of biological progression (*scala naturae*) from lower (plant) to higher (animals through humans) forms that others later seized upon as being evolutionary and we reject today. Aristotle’s classification of animals was neither comprehensive nor entirely consistent, but was hierarchical, predictive (in some sense), and formed the beginning of modern classification.

1.2 Theophrastus

Theophrastus succeeded Aristotle and is best known in biology for his *Enquiry into Plants* and *On the Causes of Plants*. As a study of classification, his work



Theophrastus
(c.371–c.287 BCE)

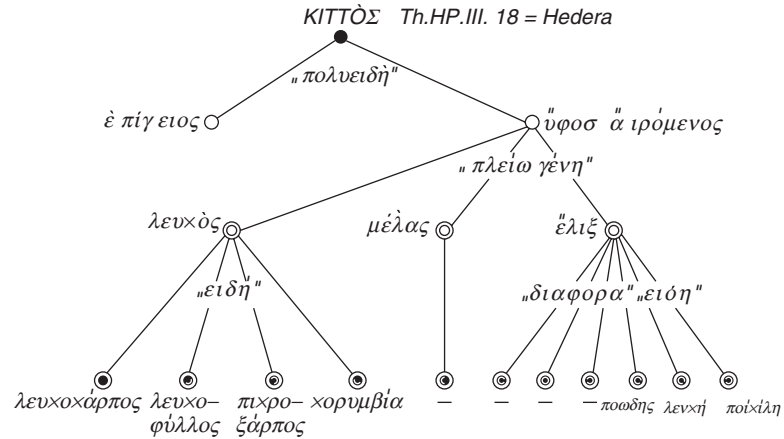


Figure 1.1: Branching diagram after Theophrastus (Vácsy, 1971).



Pierre Belon
(1517–1564)

on ivy (κίττος) discussed extensively by Nelson and Platnick (1981), has been held to be a foundational work in taxonomy based (in part at least) on dichotomous distinctions (*e.g.* growing on ground versus upright) of a few essential features.

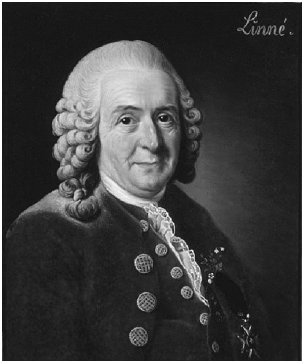
Theophrastus distinguished ivies based on growth form and color of leaves and fruit. Although he never presented a branching diagram, later workers (including Nelson and Platnick) have summarized these observations in a variety of branching diagrams (Vácsy, 1971) (Fig. 1.1).

1.3 Pierre Belon

Trained as a physician, Pierre Belon, studied botany and traveled widely in southern Europe and the Middle East. He published a number of works based on these travels and is best known for his comparative anatomical representation of the skeletons of humans and birds (Belon, 1555) (Fig. 1.2).

1.4 Carolus Linnaeus

Carolus Linnaeus (Carl von Linné) built on Aristotle and created a classification system that has been the basis for biological nomenclature and communication for over 250 years. Through its descendants, the current codes of zoological, botanical, and other nomenclature, his influence is still felt today. Linnaeus was interested in both classification and identification (animal, plant, and mineral species), hence his system included descriptions and diagnoses for the creatures he included. He formalized the custom of binomial nomenclature, genus and species we use today.



Carl von Linné
(1707–1778)

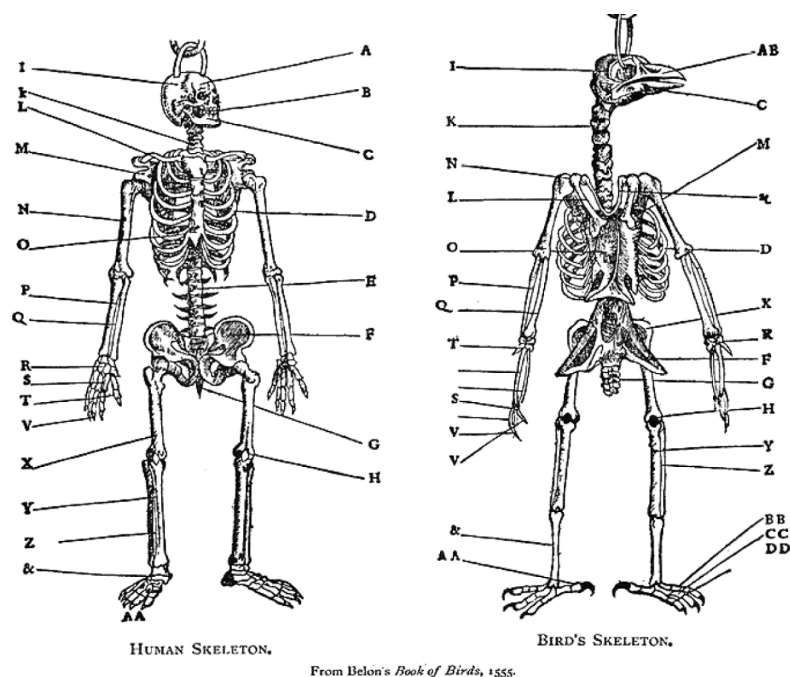


Figure 1.2: Belon's funky chicken (Belon, 1555).

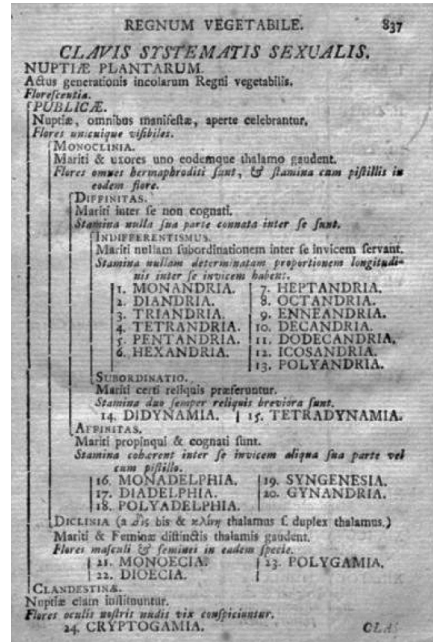
Linnaeus was known, somewhat scandalously in his day, for his sexual system of classification (Fig. 1.3). This was most extensively applied to plants, but was also employed in the classification of minerals and fossils. Flowers were described using such terms as visible (public marriage) or clandestine, and single or multiple husbands or wives (stamens and pistils). Floral parts were even analogized to the foreskin and labia.

Nomenclature for many fungal, plant, and other eukaryote groups² is founded on the *Species Plantarum* (Linnaeus, 1753), and that for animals the 10th Edition of *Systema Naturae* (Linnaeus, 1758). The system is hierarchical with seven levels reflecting order in nature (as opposed to the views of Georges Louis Leclerc, 1778 [Buffon], who believed the construct arbitrary and natural variation a result of the combinatorics of components).

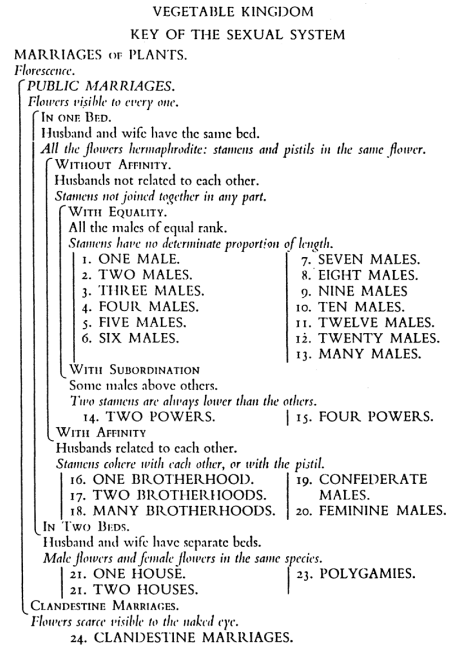
- Imperium (Empire)—everything
- Regnum (Kingdom)—animal, vegetable, or mineral
- Classis (Class)—in the animal kingdom there were six (mammals, birds, amphibians, fish, insects, and worms)
- Ordo (Order)—subdivisions of Class
- Genus—subdivisions of Order

²For the current code of botanical nomenclature see <http://ibot.sav.sk/icbn/main.htm>.

- Species—subdivisions of Genus
- Varietas (Variety)—species varieties or “sub-species.”



(a) Sexual system for plants (Linnaeus, 1758).



(b) English translation.

Figure 1.3: Linnaeus' sexual system for classification (a) with English translation (b) (Linnaeus, 1758).



Georges Louis Leclerc, Comte de Buffon
(1707–1788)

The contemporary standard hierarchy includes seven levels: Kingdom, Phylum, Class, Order, Family, Genus, and Species, although other levels are often created as needed to describe diversity conveniently (*e.g.* McKenna and Bell, 1997).

1.5 Georges Louis Leclerc, Comte de Buffon

Georges Louis Leclerc, Comte de Buffon, began his scientific career in mathematics and probability theory³. He was appointed director of the *Jardin du Roi* (later *Jardin des Plantes*), making it into a research center.

Buffon is best known for the encyclopedic and massive *Histoire naturelle, générale et particulière* (1749–1788). He was an ardent anti-Linnean, believing taxa arbitrary, hence there could be no preferred classification. He later thought, however, that species were real (due to the *moule intérieur*—a concept at the

³Buffon's Needle: Given a needle of length l dropped on a plane with a series of parallel lines d apart, what is the probability that the needle will cross a line? The solution, $\frac{2l}{d\pi}$ can be used to estimate π .

foundation of comparative biology). Furthermore, Buffon believed that species could “improve” or “degenerate” into others, (*e.g.* humans to apes) changing in response to their environment. Some (*e.g.* Mayr, 1982) have argued that Buffon was among the first evolutionary thinkers with mutable species. His observation that the mammalian species of tropical old and new world, though living in similar environments, share not one taxon, went completely against then-current thought and is seen as the foundation of biogeography as a discipline (Nelson and Platnick, 1981).

1.6 Jean-Baptiste Lamarck

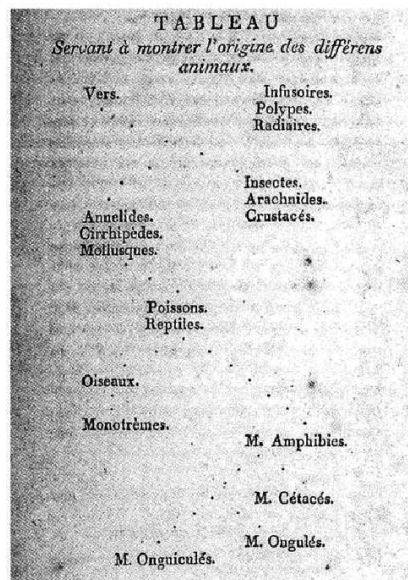
Jean-Baptiste Lamarck (who coined the word “Biologie” in 1802) believed that classifications were entirely artificial, but still useful (especially if dichotomous). His notion of classification is closer to our modern keys (Nelson and Platnick, 1981). An example of this comes from his *Philosophie zoologique* (Lamarck, 1809), with the division of animal life into vertebrates and invertebrates on the presence or absence of “blood” (Fig. 1.4(a)).



Jean-Baptiste Lamarck
(1744–1829)

1. Les Mammifères.	} Animaux vertébrés.
2. Les Oiseaux.	
3. Les Reptiles.	
4. Les Poissons.	
5. Les Mollusques.	} Animaux invertébrés.
6. Les Cirrhipèdes.	
7. Les Annelides.	
8. Les Crustacés.	
9. Les Arachnides.	
10. Les Insectes.	
11. Les Vers.	
12. Les Radiaires.	
13. Les Polypes.	
14. Les Infusoires.	

(a) Lamarck's classification of animals.



(b) Lamarck's transmutational tree.

Figure 1.4: Lamarck's division of animal life (a) and transmutational tree (b) (Lamarck, 1809).

Lamarck is best known for his theory of Transmutation (Fig. 1.4(b))—where species are immutable, but creatures may move through one species to another based on a motivating force to perfection and complexity, as well as the familiar “use and dis-use.” Not only are new species created in this manner, but species can “re-evolve” in different places or times as environment and innate drive allow.



Georges Cuvier
(1769–1832)

1.7 Georges Cuvier

The hugely influential Léopold Chrétien Frédéric Dagobert “Georges” Cuvier divided animal life not into the *Scala Naturae* of Aristotle, or two-class Vertebrate/Invertebrate divide of Lamarck, but into four “embranchements”: Vertebrata, Articulata, Mollusca, and Radiata (Cuvier, 1812). These branches were representative of basic body plans or “archetypes” derived (in Cuvier’s view) from functional requirements as opposed to common genealogical origin of structure. Based on his comparative anatomical work with living and fossil taxa, Cuvier believed that species were immutable but could go extinct, (“catastrophism”) leaving an unfillable hole. New species, then, only appeared to be new, and were really migrants not seen before. Cuvier established the process of extinction as fact, a revolutionary idea in its day.



Étienne Geoffroy Saint-Hilaire
(1772–1844)

1.8 Étienne Geoffroy Saint-Hilaire

Although (like Lamarck), the comparative anatomist Étienne Geoffroy Saint-Hilaire is remembered for his later evolutionary views⁴, Geoffroy believed that there were ideal types in nature and that species might transform among these immutable forms. Unlike Lamarck, who believed that the actions of creatures motivated transmutation, Geoffroy believed environmental conditions motivated change. This environmental effect was mediated during the development of the organism. He also believed in a fundamental unity of form for all animals (both living and extinct), with homologous structures performing similar tasks. In this, he disagreed sharply with Cuvier and his four archetypes (embranchements), not with the existence of archetypes, but with their number.



Johann Wolfgang von Goethe
(1749–1832)

1.9 Johann Wolfgang von Goethe

With Oken and Owen, Goethe was one of the foremost “ideal morphologists” of the 19th century in that he saw universal patterns underlying the forms of organisms. He coined the term “Morphology” to signify the entirety of an organism’s form through development to adult as opposed to “gestalt” (or type—which was inadequate in his view). This is similar to Hennig’s concept of the “semaphoront” to represent the totality of characters expressed by an organism over its entire life cycle.

Goethe applied these ideas to the comparative morphology and development of plants (von Goethe, 1790)⁵ as Geoffroy did to animals, creating morphological ideals to which all plants ascribed. He claimed, based on observation, that

⁴ “The external world is all-powerful in alteration of the form of organized bodies... these are inherited, and they influence all the rest of the organization of the animal, because if these modifications lead to injurious effects, the animals which exhibit them perish and are replaced by others of a somewhat different form, a form changed so as to be adapted to the new environment” (Saint-Hilaire, 1833).

⁵ In his spare time, he wrote a book called *Faust*.