Learning objectives:

- Explore architecture/structure in economics and biology
- Explore selected structural representations in biology and economics and their possible wider utility.
- Appreciate the differences between dendograms and Cladograms
- Understand the differentiation and similarity of structure arrived at by decomposition and by aggregation and the different constraints that occur in economic and biological systems.
- Examine some recent biological findings from animal development research and how this work links to evolvability
- Explore how the "New Science of Evo Devo" may relate to Engineering Systems.

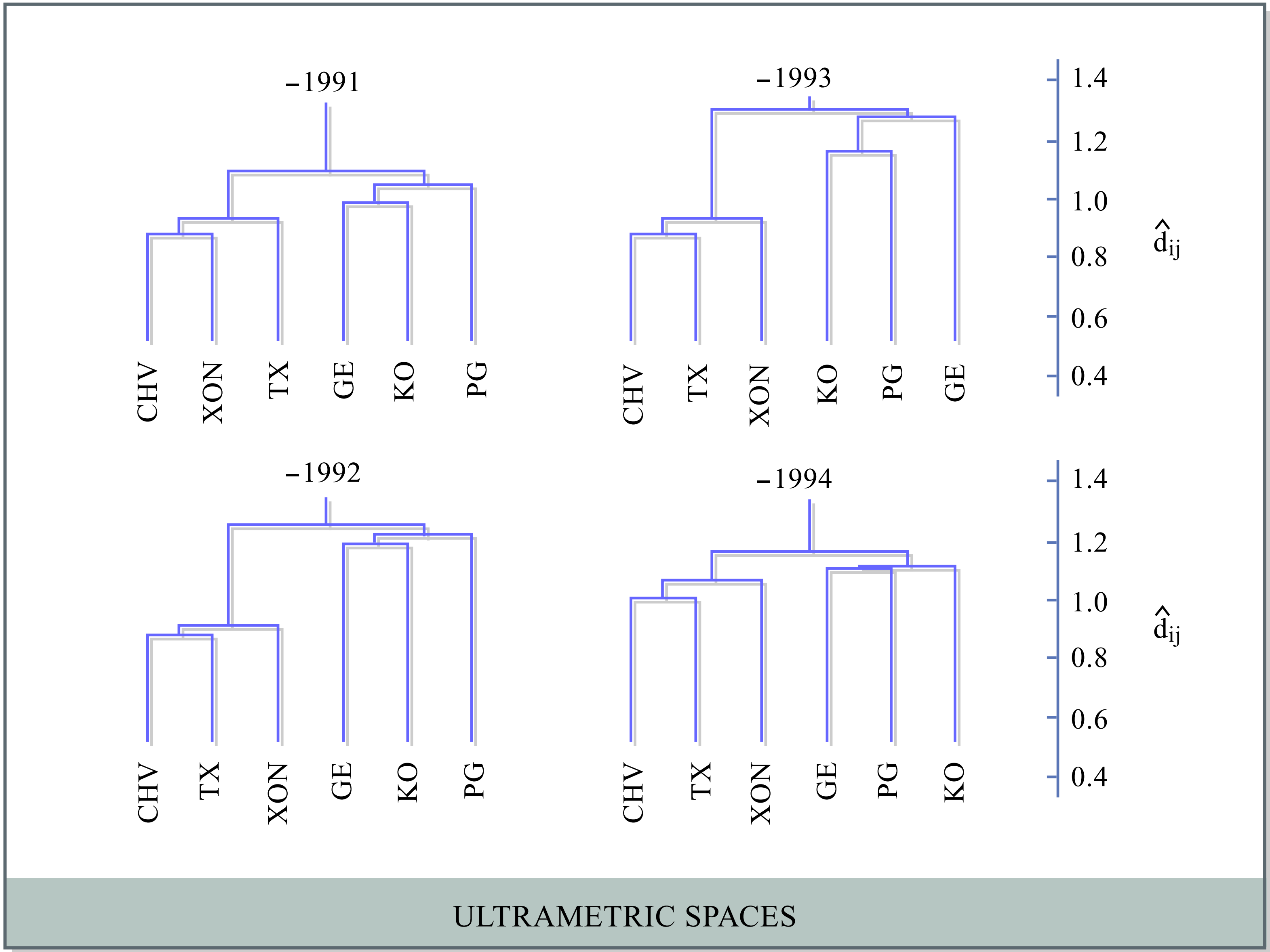
Lecture Outline

- Some structural aspects of economics and relationships to previous structural characterizations
- Abbreviated History of *Systematics* in Biology
- Some current thinking about biological systematics and recent indications of a new synthesis among molecular biology, evolutionary biology and systematics
- Biological Evolvability (and flexibility)
- Innovation and evolutionary thinking

Economics Structure

- **Sub-division of the economy**
  - National boundaries
  - Industry classifications (SIC and now NASIC) are hierarchical with 5 levels
- How are these industry classifications arrived at?
- Partially Input/output relationships: if two plants (locations) use similar inputs (suppliers) and have similar outputs (customers), they are classified together. This is similar to studying community structure in a network with nodes being plants and links being specific product flows
- Broader abstractions that gather similarities at a higher level: for example- agriculture and extractive industries, manufacturing and service
- Market similarities: Chapter from Mantegna and Stanley
  - Companies are classified together according to how correlated their dynamic movement is over time.
  - Metric is $d_{ij} = \sqrt{2(1 - \rho_{ij})}.$

Dendograms From Similarity Metric



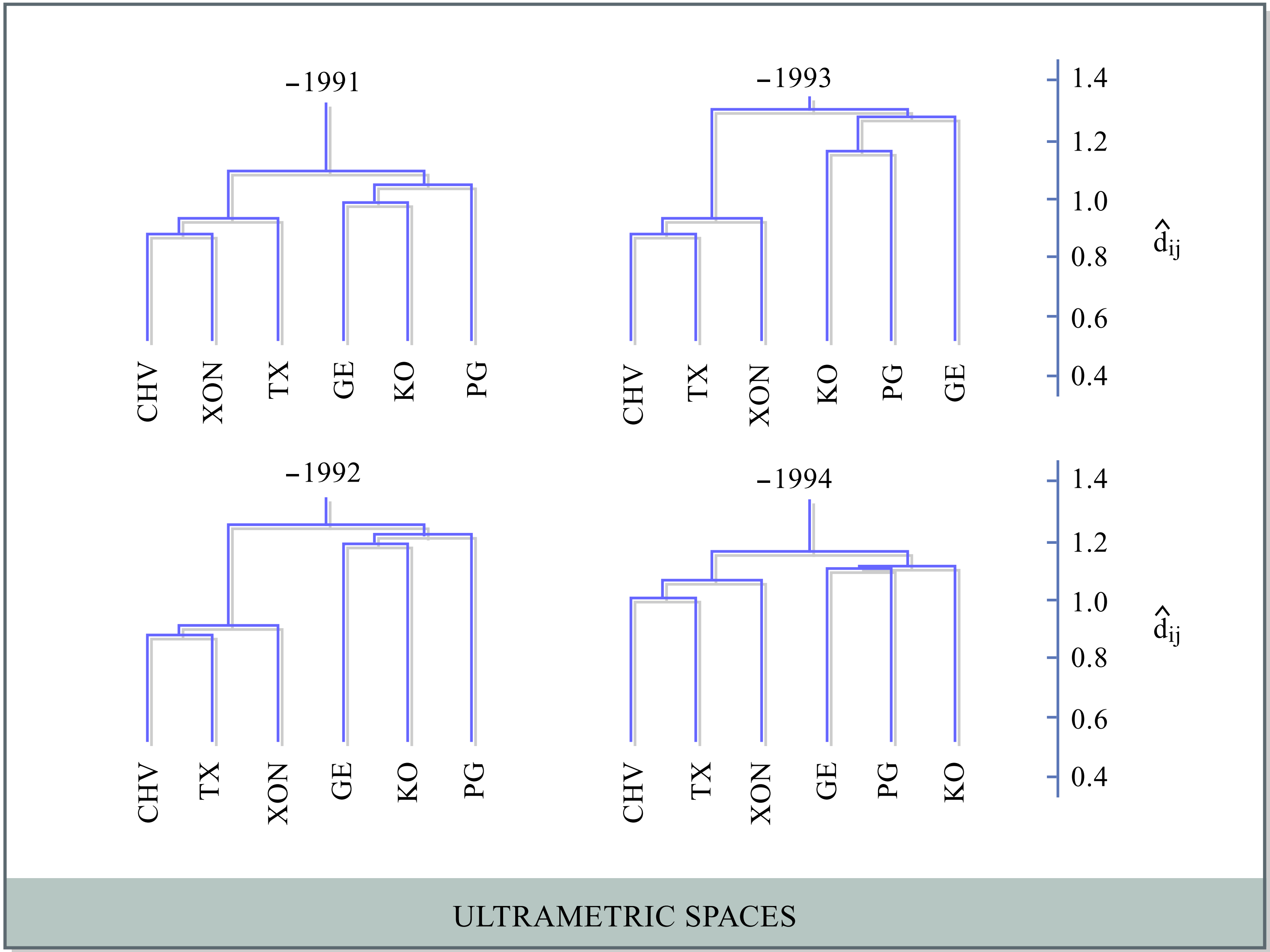
Source: *An Introduction to Econophysics*, R. Mantegna and E. Stanley, page 109, 2000

Figure by MIT OCW.

Dendograms and Network Community Structure

- Dendograms from metrics allow one to see the strength of relationships that determine association of nodes
- Dendograms do not allow the communities to be seen as easily
  - One must make horizontal slices at various levels of similarity and see how many communities arise.

Dendograms From Similarity Metric



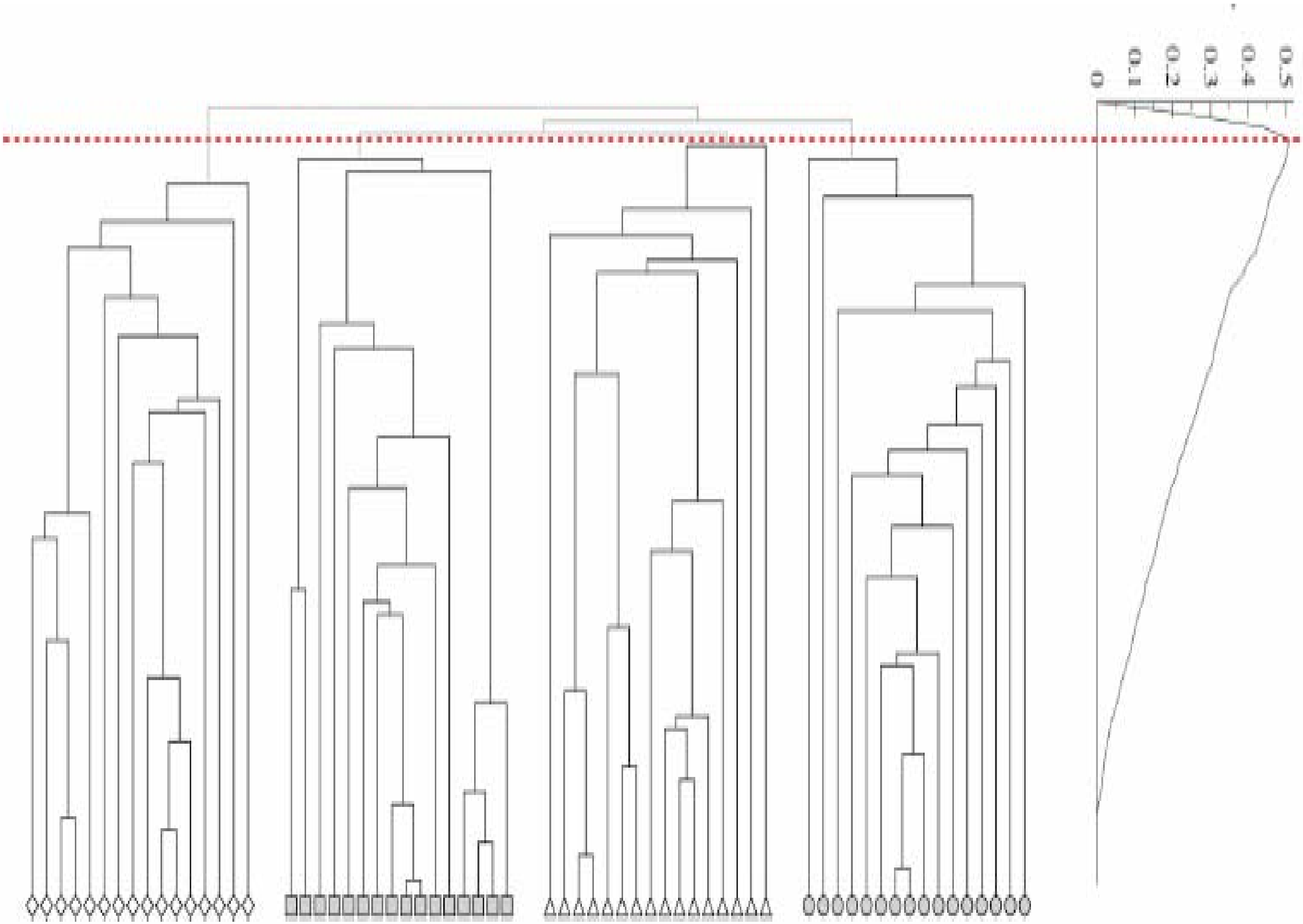
Source: *An Introduction to Econophysics*, R. Mantegna and E. Stanley, page 109, 2000

Figure by MIT OCW.

Relationship between hierarchical clustering and community structure analysis

Dendograms and Network Community Structure

- Dendograms from metrics allow one to see the strength of relationships that determine association of nodes
- Dendograms do not allow the communities to be seen as easily
  - One must make horizontal slices at various levels of similarity and see how many communities arise.
- Community structure analysis arrives at the structure by *decomposition* (removing links from a full network) whereas hierarchical clustering (dendograms) is done by adding links in the order of most similar to least similar (*aggregation*)


0
0.1
0.2
0.3
0.4
0.5

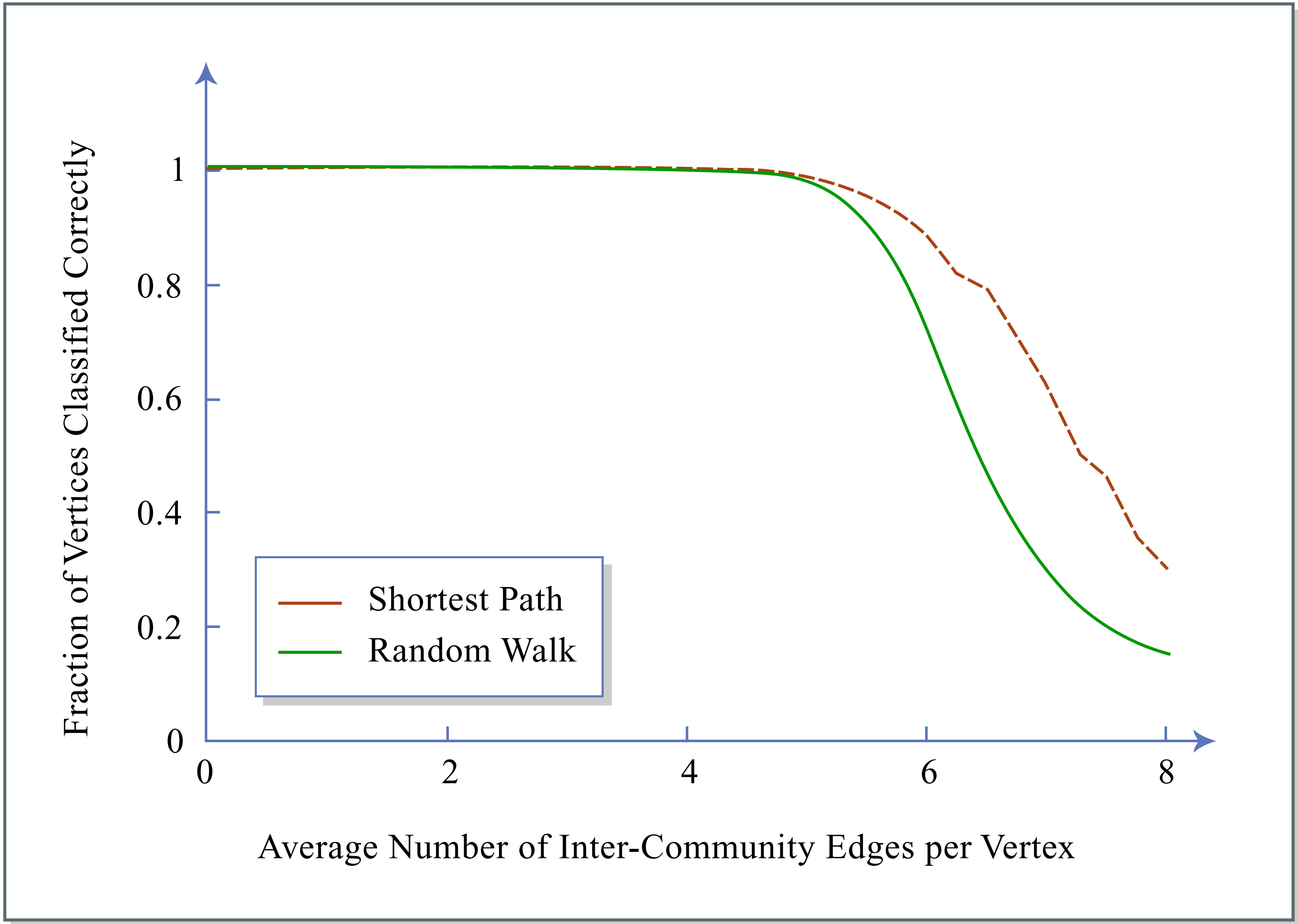
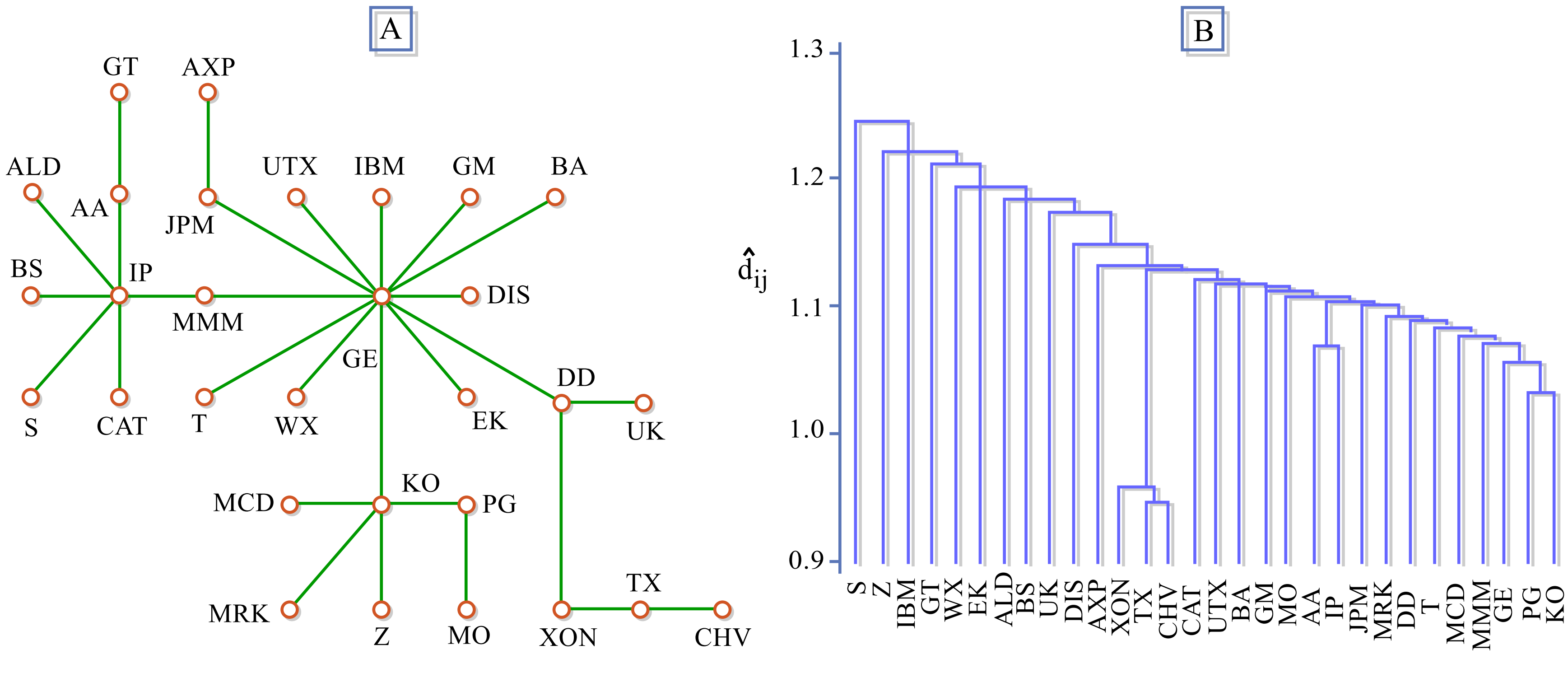


Figure by MIT OCW.



TAXONOMY OF A STOCK PORTFOLIO

Figure by MIT OCW.
After Mantegna and Stanley.

Is the relatively low degree of correlation among DJI stocks "explainable"?

Source: *An Introduction to Econophysics*, R. Mantegna and E. Stanley, page 110, 2000

Economics Structure

- **Sub-division of the economy**
  - National boundaries
  - Industry classifications (SIC and now NASIC) are hierarchical with 5 levels
- How are these industry classifications arrived at?
- Partially Input/output relationships
- Broader abstractions that gather similarities at a higher level: for example- agriculture and extractive industries, manufacturing and service
- Market similarities: Chapter from Mantegna and Stanley
- **Sub-division of the field of study (e.g. publications)**
  - Suggestions…

Journal of Economic Literature Classification System (19 categories)

- General Economics and Teaching
- Schools of Economic Thought and Methodology
- Mathematical and Quantitative Methods
- Microeconomics
- Macroeconomics and Monetary Economics
- International Economics
- Financial Economics
- Public Economics
- Health, Education and Welfare
- Labor and Demographic Economics
- Law and Economics
- Industrial Organization
- Business Administration and Business Economics; Marketing; Accounting
- Economic History
- Economic Development, Technological Change, and Growth
- Economic Systems
- Agricultural and Natural Resource Economics: Environmental and Ecological Economics
- Urban, Rural and Regional Economics
- Other Special Topics

Journal of Economic Literature Classification System (19 top categories)

- General Economics and Teaching
- Schools of Economic Thought and Methodology
- Mathematical and Quantitative Methods
- Microeconomics
- Macroeconomics and Monetary Economics
- International Economics
- Financial Economics
- Public Economics
- Health, Education and Welfare
- Labor and Demographic Economics
- Law and Economics
- **Industrial Organization**
- Business Administration and Business Economics; Marketing; Accounting
- Economic History
- Economic Development, Technological Change, and Growth
- Economic Systems
- Agricultural and Natural Resource Economics: Environmental and Ecological Economics
- Urban, Rural and Regional Economics
- Other Special Topics

Journal of Economic Literature Classification System: Typical Final Substructure in which books and papers are the next level of detail

- L. Industrial Organization (one of 19 highest level categories)
  - L9 – Industry Studies: Transportation and Utilities
    - L90- General
    - L91- Transportation: General
    - L92- Railroads and Other Surface Transportation: Autos, Buses, Trucks and Water Carriers; Ports
    - L93- Air Transportation
    - L94- Electric Utilities
    - L95- Gas Utilities; Pipelines; Water Utilities
    - L96- Telecommunications
    - L97-Utilities: General
    - L98- Government Policy
    - L99- Other

Economics Structure

- **Sub-division of the economy**
  - National boundaries
  - Industry classifications (SIC and now NASIC) are hierarchical with 5 levels
- How are these industry classifications arrived at?
- Partially Input/output relationships
- Broader abstractions that gather similarities at a higher level: for example- agriculture and extractive industries, manufacturing and service
- Market similarities: Chapter from Mantegna and Stanley
- **Sub-division of the field of study (e.g. publications)**
  - *JEL* Classification System

Decomposition Approaches to Determining Structure

Economics Structure

- **Sub-division of the economy**
  - National boundaries
  - Industry classifications (SIC and now NASIC) are hierarchical with 5 levels
- How are these industry classifications arrived at?
- Partially Input/output relationships
- Broader abstractions that gather similarities at a higher level: for example- agriculture and extractive industries, manufacturing and service
- Market similarities: Chapter from Mantegna and Stanley
- **Sub-division of the field of study (e.g. publications)**
  - *JEL* Classification System

Aggregation Approaches to Determining Structure

Structure in Biology

- **Sub-division of the field of study (e.g. publications)**
  - No parallel found to *JEL* Classification System

Categorical Topics from on-line Biology Text (19 given!)

- Animals
- Behavior
- Biochemistry
- Cancer
- Cell Biology
- Chemistry
- Diversity of Life
- DNA and RNA
- Ecology
- Development
- Evolution
- General Science
- Genetics
- Immunology
- Microbiology
- Molecules
- Physiology
- Plant Biology
- Viruses

Structure in Biology

- **Sub-division of the field of study (e.g. publications)**
  - No parallel found to *JEL* Classification System
  - Textbook topics are at best approximately homologous with the activities in the field. Areas such as bioinformatics, mathematical modeling and observational techniques receive relatively more listings than for economics so would probably appear on an equivalent to the *JEL* Classification. Also, it is clear that systematics is an old and still very active field not much covered in modern undergraduate textbooks. Paleontology is not part of biology but is essential to Biology Systematics.
- **Sub-division (and aggregation) of the objects studied**
  - Extensive, historically dominant in field and still very active

Biological Classification History

- Aristotle (had lots of graduate students?) identified hundreds of differing kinds of animals and plants based upon morphology and function.
- John Ray: 1628-1705, [Cambridge UK] restarted biological classification based on morphology and structural similarity.
- Linnaeus 1707-1778, [Swedish] is the most recognized biological classifier. He introduced the idea of a unified hierarchical tree (7 levels) and the "binomial" nomenclature for species and both of these still stand. However, the characteristics he used for classification (sexual reproduction modes and organs) are now not as important as the ones used by Ray.
- Aggasiz, Paley, Sedgwick, Buckland (19th century): paleontology, embryology, ecology, and biogeography all became important in classification through their work.
- Robert Whitaker in 1969 proposed 5 kingdoms whereas Linnaeus only had 2 (plants and animals).

Lecture Outline

- Some structural aspects of economics and relationships to previous structural characterizations
- Abbreviated History of *Systematics* in Biology
- Some current thinking about biological systematics and recent indications of a new synthesis among molecular biology, evolutionary biology and systematics
- Biological Evolvability (and flexibility)
- Innovation and evolutionary thinking

Biological Classification

- 5 Kingdoms: Plants, Animals, Fungi, Prokaryotes, Protists
- The 7 layer hierarchy continues:
  - Phylum (for animals) and Divisions (for plants and Fungi)
  - Class
  - Order
  - Family
  - Genus
  - Specie (named by genus + Latinized specific)
- The bottom 2 layers and the top are fairly well specified but the middle is a real muddle and extremely hard to make sensible in all 5 kingdoms.
- Darwin changed the way of thinking about the hierarchical tree and thus an evolutionist framework with alternatives and testability has emerged. The systematists (also cladists) are now biology's "high volume" classifiers with the kinds of problems discussed in the Gould reading for today

(Phylogenetic) Cladistics

- Cladistics is now accepted as the best method available for phylogenetic analysis, for it provides an explicit and testable hypothesis of organismal relationships.
- Three assumptions of cladistics
  - Any group of organisms are *related by descent* from a common ancestor.
  - There is a *bifurcating pattern* of cladogenesis.
  - *Change* in characteristics occurs in lineages over time.
- A Clade is a monophyletic taxon; a group of organisms which includes the most recent common ancestor of all of its members and all of the descendants of that most recent common ancestor. From the Greek word "klados", meaning branch or twig.
- The basis for similarity is shared derived characteristics among the members of the clade

An Example Cladogram

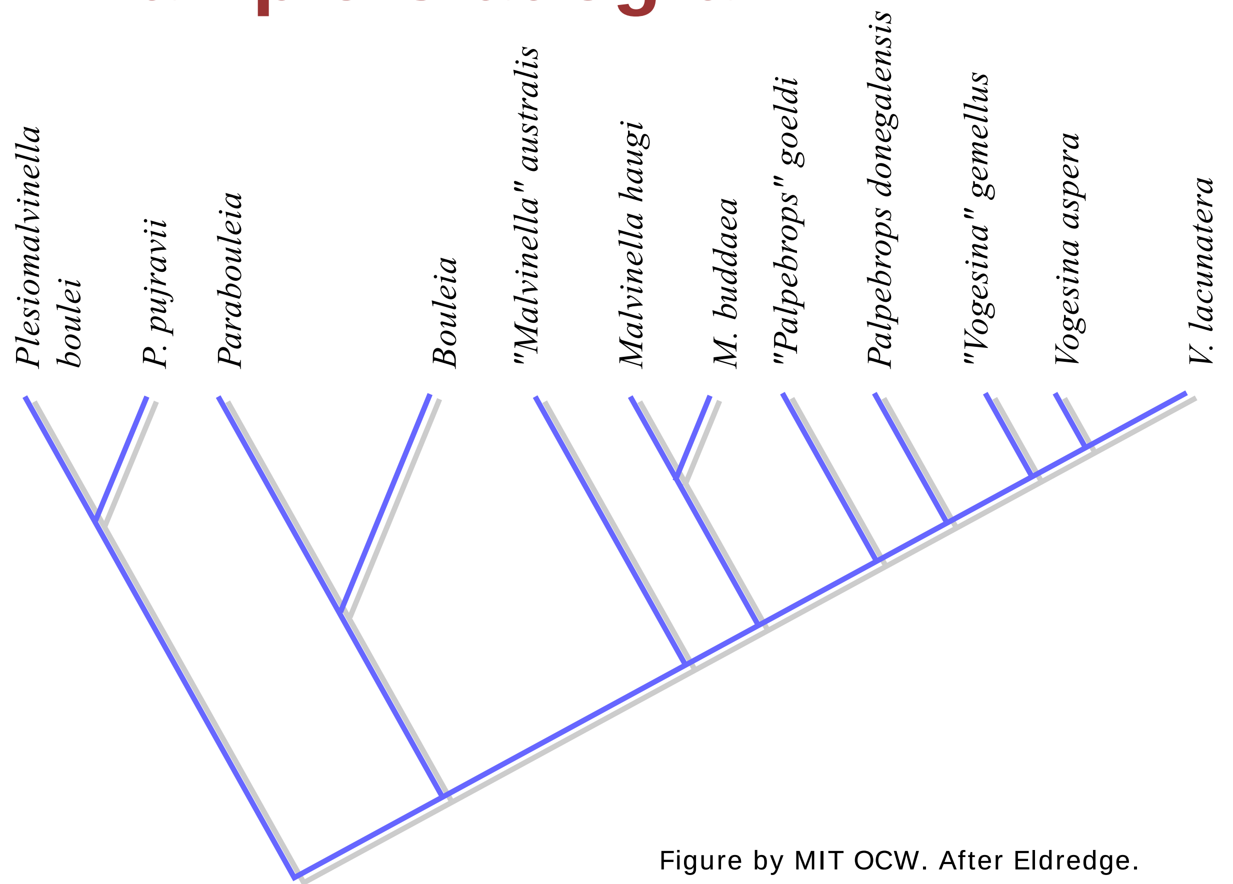


Figure by MIT OCW. After Eldredge.

An evolutionary tree or "cladogram," showing relationships among a group of Devonian trilobites from the southern hemisphere. This diagram is typical of the thousands of cladograms published every year by systematists studying the evolutionary relationships of fossil and recent organisms. In the majority of cases, it is possible to specify with confidence which species are most closely related.

Source: *Darwin, Discovering the Tree of Life*, Niles Eldredge, page 235, 2005

The cladistic pattern of great apes and humans

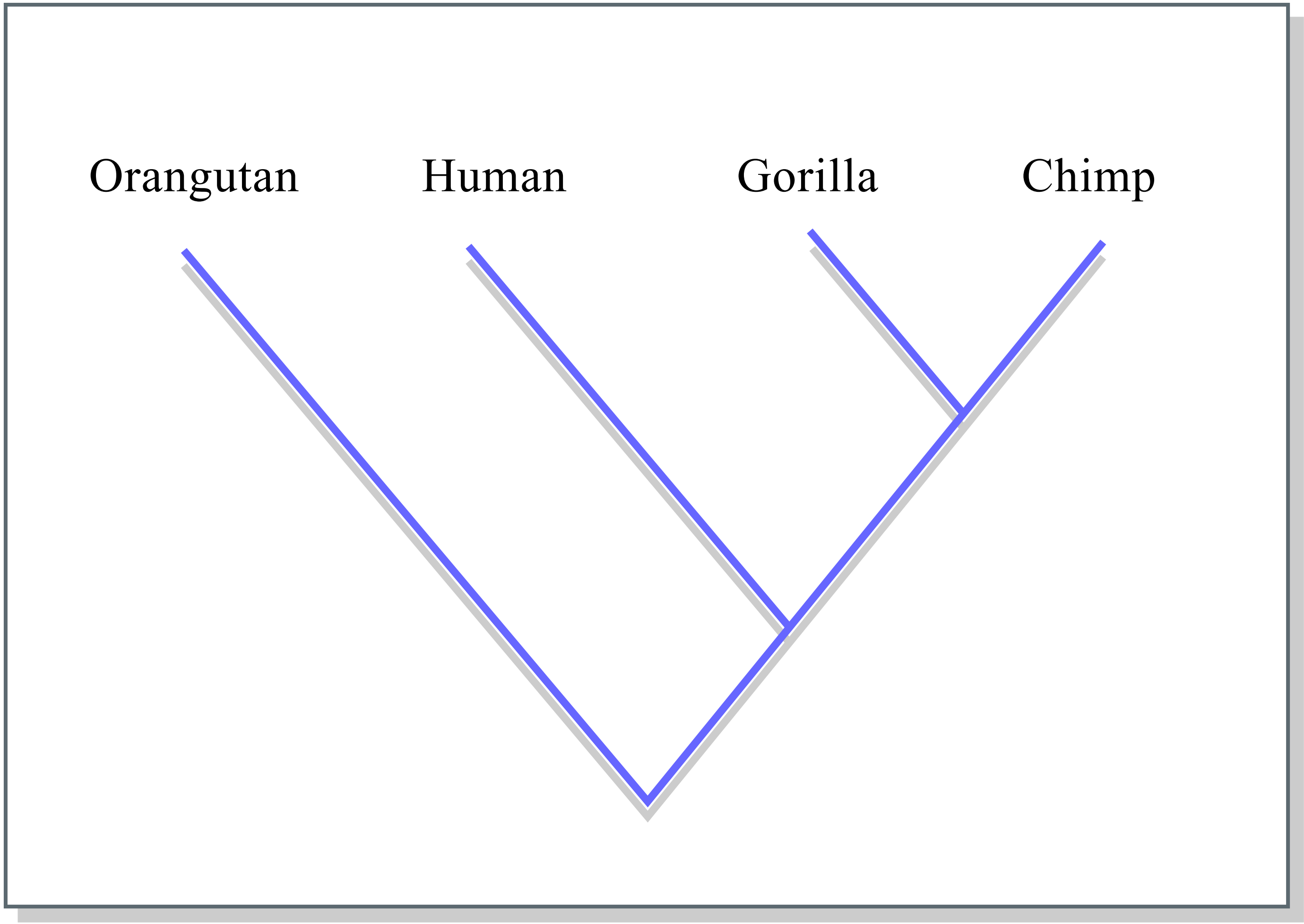


Figure by MIT OCW. After Gould and Monnier.

Source: *Hen's Teeth and Horse's Toes*, Stephen Jay Gould, page 357, 1983 (reprinted from Natural History, drawing by Joe Le Monnier)

Structure in Biology

- **Sub-division of the field of study (e.g. publications)**
  - No parallel found to *JEL* Classification System
  - Textbook Topics list OK but biology would be much more difficult to structure usefully than is economics as it is much larger and not as insular.
- **Sub-division (and aggregation) of the objects studied**
  - Body plans, morphology, sexual reproduction modes and organs
  - Paleontology, embryology, ecology, and biogeography
  - Evolutionary relationships and Molecular Biology (particularly in regard to switches and regulatory genes) are now the main drivers for ongoing new systematics findings.

Summary observations on Structure in Biology and Economics

- Uniformity of closeness by separate criteria is not the usual case (Gould- nature is more interesting than this- and so is technology or humanity's creations)
- Pure decomposability of all dimensions is rare (works for pure chemical elements but not for mixtures –alloys)
- Categorization/abstraction is highly useful but cannot be expected to be perfect
- Statics are not enough, deeper understanding arises through understanding changes over time.
- Deeper biological understanding is now particularly coming about by combination of embryology (development) and long-term change (evolution) – "EVO-DEVO"

Developmental Biology (Classical Embryology with Molecular Biology)

- One of the most rapidly developing areas of understanding for biology
- May yield important insights relative to the "Evolvability" of complex, large-scale engineering systems
- Some Observations from simple animals:
  - Embryos possess "site maps" very early-cells learn where they are and what type of cell they will be in the full form.
  - "Tool-kit genes" establish longitude, latitude and "mark the spot" of future appendages in the earliest embryo
  - Genes that encode proteins are ~1.5% of DNA but regulatory material make up another ~3%
  - This regulation is best thought of as a "switch"

The Role of Gene Regulatory Switches in Development

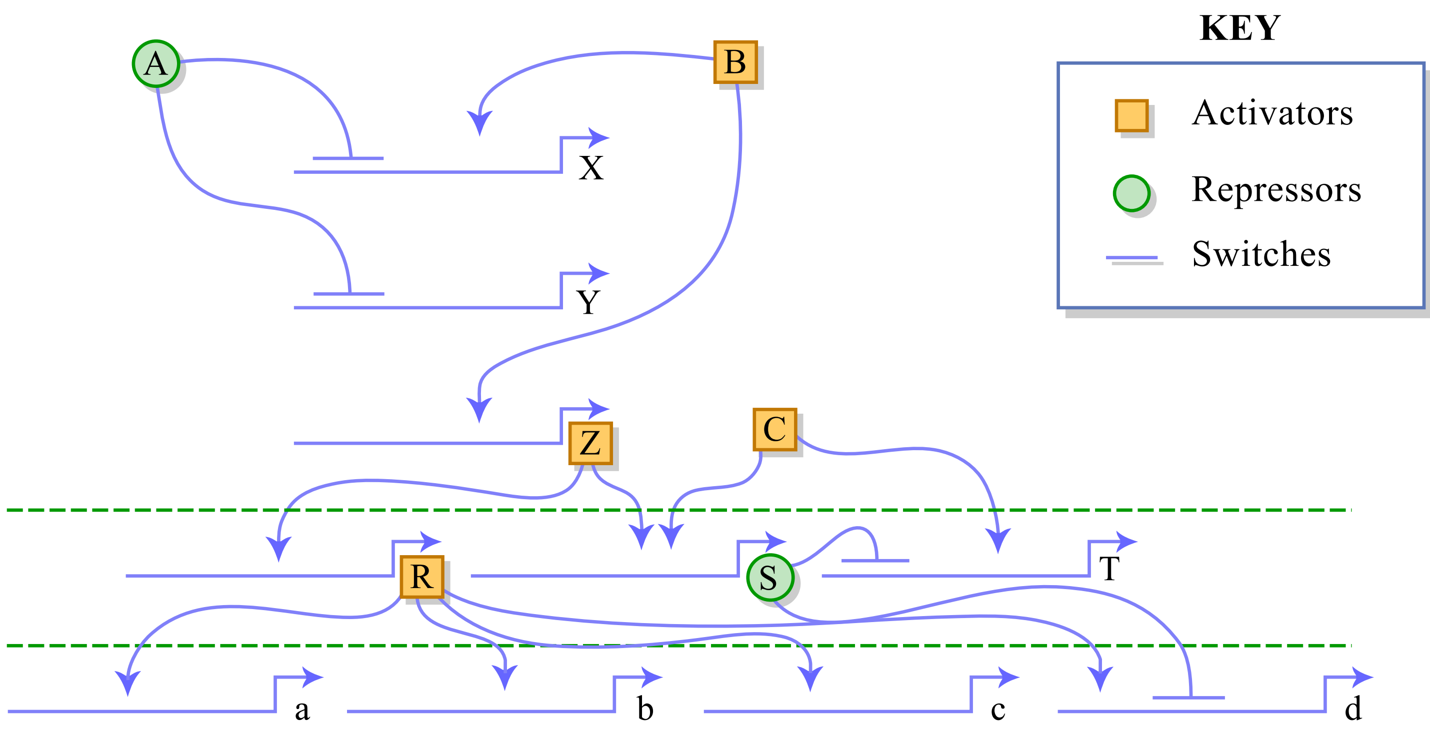
- Individual switches are long sequences of DNA bound by a large number and variety of proteins which can activate or suppress gene transcription. Switches "compute" the input of multiple proteins and convert complex sets of inputs into simpler 3D on/off patterns of gene expression.
- A single gene can be controlled by multiple switches so that a single gene is used differently in many places (and times) during development.
- The makeup of every switch is different but they are composed of "reused" signature sequences of DNA that bind various proteins in different ways.
- Simply varying the position of a few signature sequences opens up a combinatorially large (cornucopia) number of switch varieties
- There has been good experimental confirmation of the action of these switches (in flies) as "GPS integrators".

The Role of Gene Regulatory Switches in Development II

- Experimental work on flies has also shown that each stripe in a fly is controlled by its own unique switch. Thus, Turing's "mathematically beautiful theory" of the organization of periodic patterns across whole structures was not found but instead that the periodic patterns are the composite of numerous individual elements.
- The tool kit genes are used over and over in different contexts differently due to differing switches with spectacular diversity and exquisite geographic specificity.
- Hierarchies of genes and switches operate in development. For example, one switch at the level of the *Hox* gene which acts in each body segment and other switches that are recognized by *Hox* Proteins that control how other genes are expressed in specific modules.
- The forms of animals and body parts are not due to specific switches but

  "large sets of interconnected switches and proteins form 'local circuits' that govern the development of complex structures. *Animal architecture is a product of genetic regulatory network architecture."* (from S. Carroll P 129 italics added)

An example fragment of the regulatory network



Activators and repressors act on switches. Arrows are activation events, lines ending in flat denote repression. Note the multiple tiers or hierarchy that enter into this regulation.

GENETIC WIRING DIAGRAM OF REGULATORY LOGIC

Figure by MIT OCW. After Carroll.

Source: *Endless Forms Most Beautiful*, Sean B. Carroll, page 129, 2005

The connection between devo and evo

- A paradox arose when it was discovered that the same or very similar tool kit genes were found in very disparate animals- how can the same genes be used to build such different (endless and beautiful) forms?
- A clue to resolving this came with the discovery of arrays of switches that enable individual tool kit genes to be used again and again in one animal in slightly or dramatically different ways in serially repeated structures.
- *"Switches enable the same tool kit genes to be used differently in different animals."* (S. Carroll, *EVO-DEVO*, p131)
  - The major evolutionary step was the Cambrian explosion beginning about 540 million years ago –bilateralism and a cornucopia of increasingly complex body types evolved over the next 10-20 million years

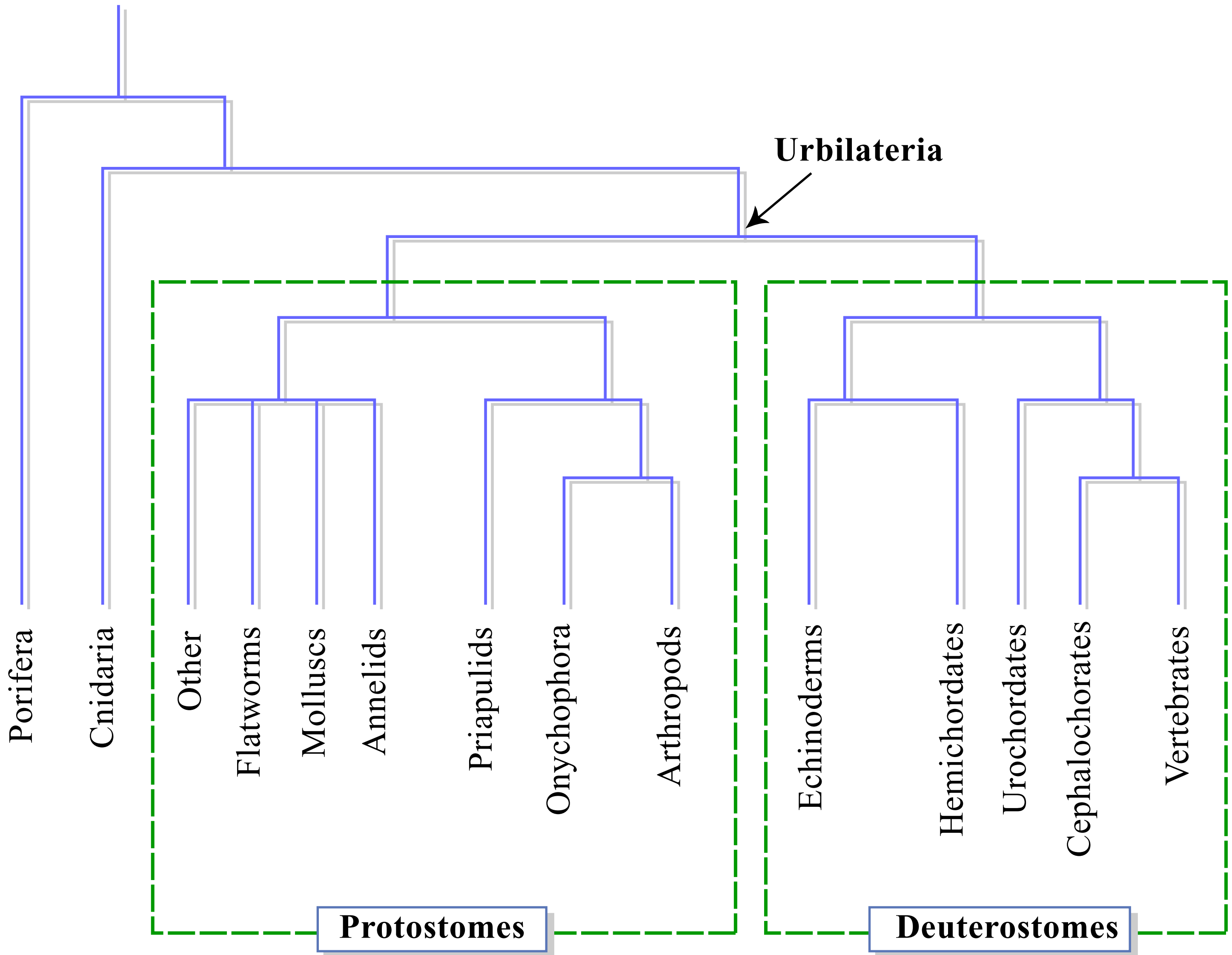


Figure by MIT OCW. After Carroll.
Source: *Endless Forms Most Beautiful,* Sean B.Carroll, page 142, 2005

The connection between devo and evo

- A paradox arose when it was discovered that the same or very similar tool kit genes were found in very disparate animals- how can the same genes be used to build such different (endless and beautiful) forms?
- A clue to resolving this came with the discovery of arrays of switches that enable individual tool kit genes to be used again and again in one animal in slightly or dramatically different ways in serially repeated structures.
- *"Switches enable the same tool kit genes to be used differently in different animals."* (S. Carroll, *EVO-DEVO*, p131)
  - The major evolutionary step was the Cambrian explosion beginning about 540 million years ago –bilateralism and a cornucopia of increasingly complex body types evolved over the next 10-20 million years
  - "The main story in the Cambrian is that of evolving different kinds and numbers of repeated body parts." p. 159
  - EVO-DEVO message: all genes for this explosion present for 50 million or more years earlier. P. 159

The chordate evolutionary tree and the expansion of *Hox* clusters in vertebrate evolution

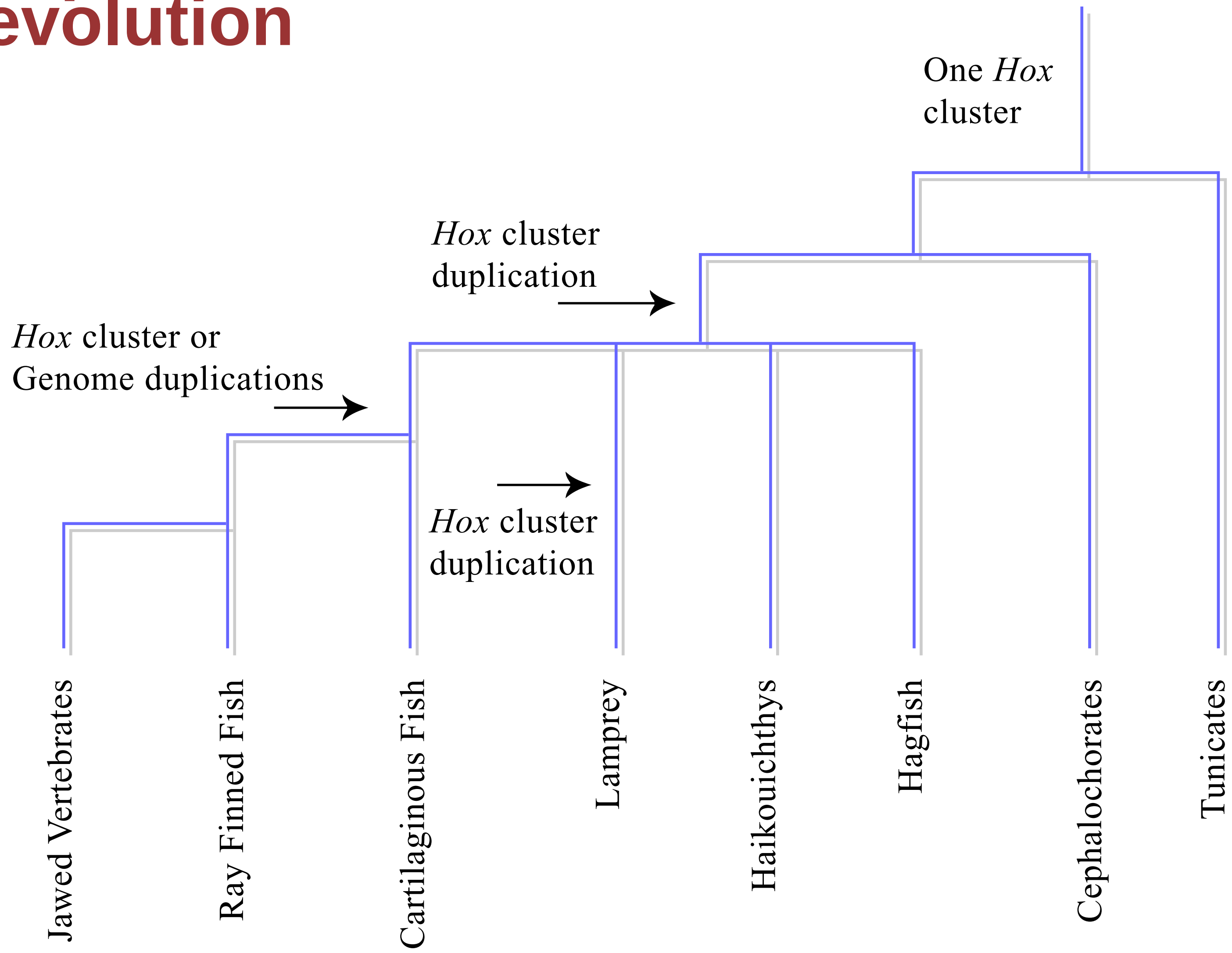


Figure by MIT OCW.
Source: *Endless Forms Most Beautiful,* Sean B.Carroll, page 155, 2005

Some macro understandings before molecular biology came along

- The Title of Gould’s book is *Hen’s Teeth and Horses Toes*. His article on this written in the late 70’s points the way towards regulatory (switches) importance in evolution.
- “Is the zebra white with black stripes or black with white stripes?” is the title of Chapter 29 in *Hen’s Teeth and Horses Toes*
- Melanism in various species (flies, mice, birds, etc.) is now well understood at the genetic, molecular level and the important protein is common (MC1R). The regulation of this protein can be used to make spots, stripes and any complex pattern one wants by switch arrays.
- Bard in 1977 (before any knowledge of MC1R or switches was available), “explained” the different stripe patterns of different zebras by a scaling argument (D’Arcy-Thompson principle applied)

Bard's model for stripe initiation at constant spacing in embryo and variable timing in development

Stripes in embryo at 21 days

Image removed for copyright reasons.
Image of Equus burchelli (Burchelli's zebra).

Stripes in embryo at ~ 28 days

Image removed for copyright reasons.
Image of Equus zebra (mountain zebra).

Stripes in embryo at ~35 days

Image removed for copyright reasons.
Image of Equus grevyi (Grevy's zebra).

Is the zebra white with black stripes or black with white stripes?

- Melanism in various species (flies, mice, birds, etc.) is now well understood at the genetic, molecular level and the important protein is common (MC1R). The regulation of this protein can be used to make spots, stripes and any complex pattern one wants by switch arrays.
- Bard in 1977 (before any knowledge of MC1R or switches was available), "explained" the different stripe patterns of different zebras by a scaling argument (D'Arcy-Thompson principle applied)
- The model is consistent with switches and the multiple evolution of striping and with other observations but the proof (embryology studies) is not (yet) available.
- This example shows the value of macro-thinking combined with new molecular approaches as does the entire exciting area of EVO-DEVO.

Figure removed for copyright reasons.
Source: Eldredge, Niles. *Darwin: Discovering the Tree of Life*. W. W. Norton & Company, Incorporated. New York, NY: 2005. p. 226.

Evolutionary relationship of all major groups of organisms alive today. Contributions of many different scientists are compiled based on a combination of anatomical, **developmental**, and genetic information.

Figure removed for copyright reasons.
Source: Eldredge, Niles. *Darwin: Discovering the Tree of Life*. W. W. Norton & Company, Incorporated. New York, NY: 2005. p. 226.

References

- *Endless Forms Most Beautiful: The New Science of EVO Devo* by Sean B. Carroll, (2005)
- *Darwin: Discovering the Tree of Life*, by Niles Eldredge (2005)
- *Animal Evolution: Interrelationships of the Living Phyla* by Claus Nielsen (2nd Edition, 2001)
- *Hen's Teeth and Horse's Toes: Further Reflections on Natural History*, by Stephen J. Gould (1983) [Chapter 28 was read for the class)
- *An Introduction to Econophysics: Correlations and Complexity in Finance*, by R. N. Mantegna and H. E. Stanley, (2000) [Chapter 13 was read for class]