

Taxonomy and morpho-molecular investigations

▣ Summary

I. Basic knowledge in a historical perspective

- a) Linnean systematics
- b) Problems with mite taxonomy
- c) The phylogenetic approach

II. How are phylogenetics and traditional taxonomy working together?

▣ Objectives

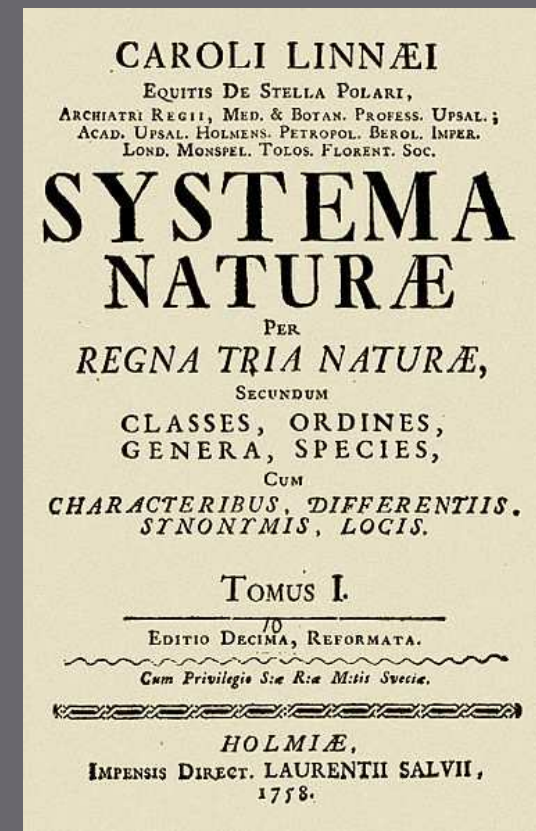
- Clearing up the confusion between legislative part of taxonomy and natural groups / classification consistency
- Identifying the effective contributions of molecular biology
- Understanding how crucial the confrontation of different lines of evidence is (nothing all right, DNA brings complimentary information, never all information)

Linnaeus'

classification of living things

- ▣ Carl von Linné (or Carl Linnaeus) (1707-1778)
 - A swedish naturalist
 - Founder of the modern system for classifying living things
 - Most important work:

Systema Naturæ
(Nature systems)
1758 (Xth edition)



Linnaeus'

classification of living things

- ▣ Linnaeus' binominal nomenclature
 - Objectives = accurately referring to every living species in such a way as to allow worldwide taxonomists understanding each another

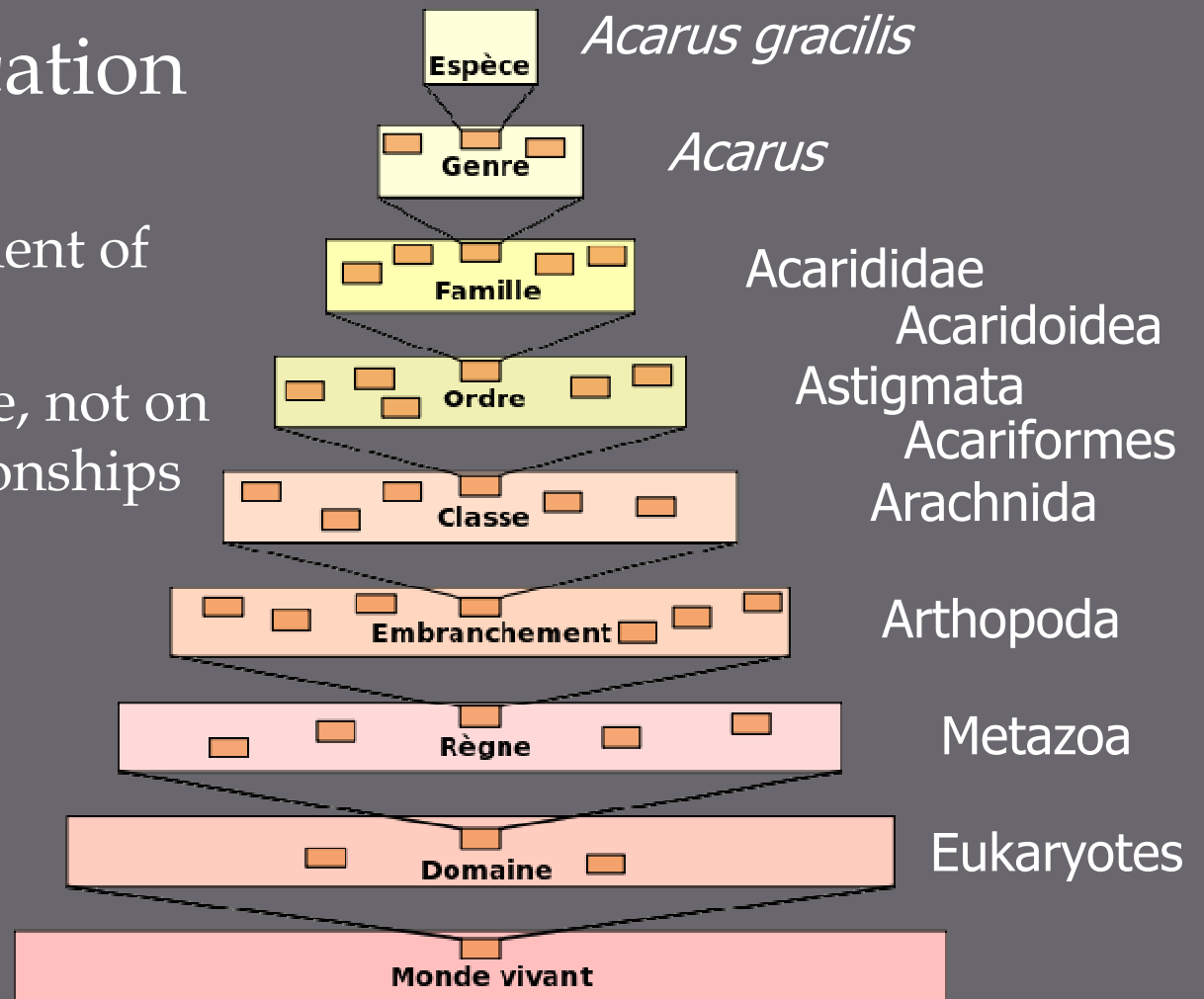


Acarus farinae DeGeer, 1778
Acarus gracilis Hughes, 1957
Acarus calcarabellus (Griffiths, 1965)

Linnaeus' classification of living things

□ Linnaeus' classification

- Hierarchical arrangement of taxa
- Based on ressemblance, not on established interrelationships



Linnaeus' classification of living things

Type-based approach

- ▣ A type = a reference taxon/specimen
 - Representative of a 1-step higher-level taxon
 - Each taxon's description based on characters common to its members
 - Each taxon's description refers to a type taxon

Genus and higher-level taxa:

Type = the first described just below taxon:

E.g. type family for a superfamily, type genus for a family, type species for a genus...

Species and lower-level taxa : 1 principal specimen (holotype) + eventually other specimens having been examined to describe the species (paratypes, allotypes...)

Examples of mite type taxa

- ▣ *Parasitus* Latreille, 1795
= type genus of family Parasitidae Oudemans, 1901
- ▣ *Dermanyssus gallinae* (De Geer, 1778)
= type species of genus *Dermanyssus* Dugès, 1834

Examples of Linnaeus' type-specimens

<http://linnean-online.org/24298/>



© The Linnean Society of London



© The Linnean Society of London

Papilio janira Linnaeus, 1758



© The Linnean Society of London



© The Linnean Society of London

Papilio mneme Linnaeus, 1763

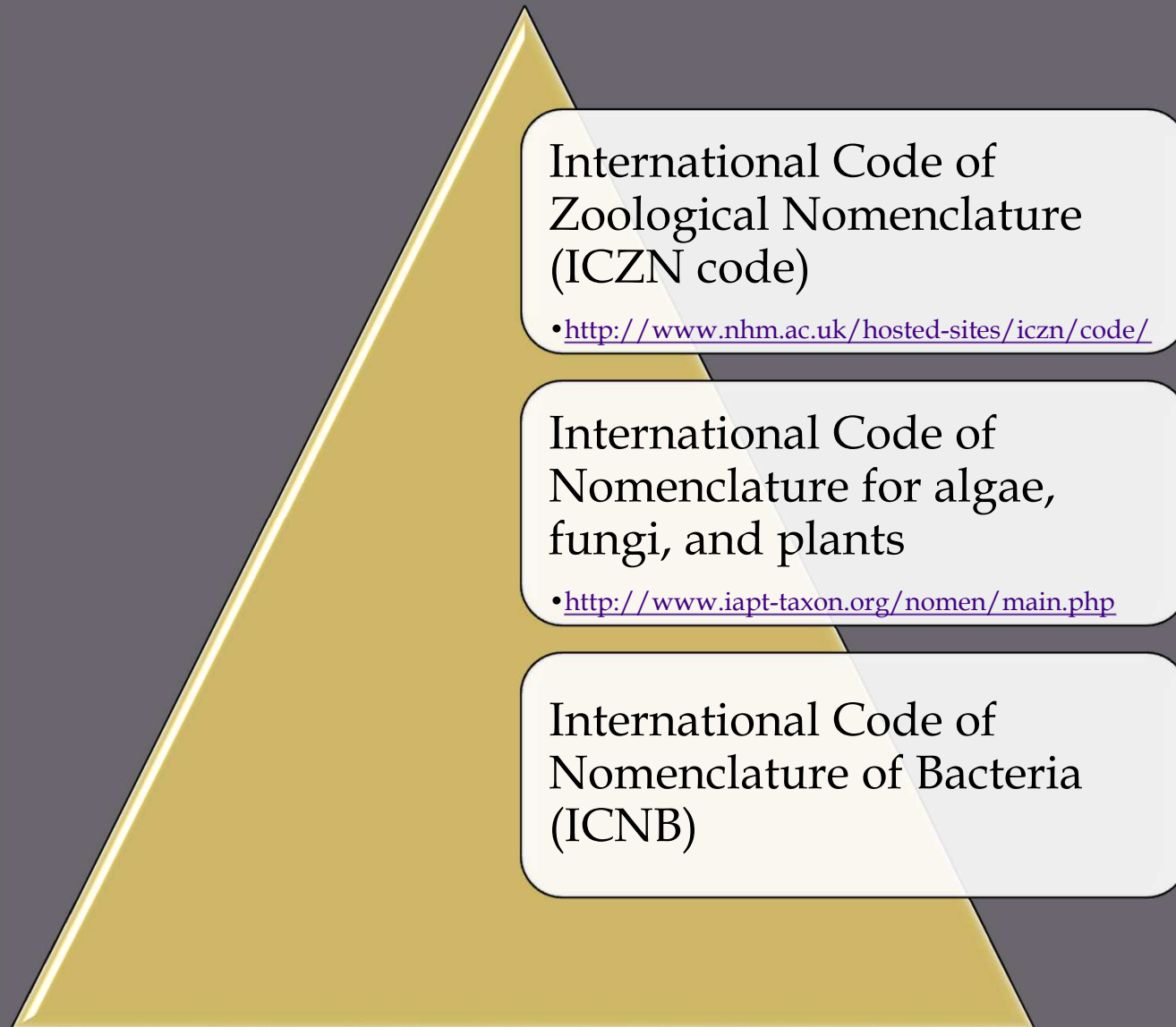
Examples of Linnaeus' type-specimens

- ▣ *Acarus elephantinus*



<http://linnean-online.org/24298/>

Linnaeus' nomenclature = current use



/!\ not to be confused to each another

- ▣ Concept of species
 - Fixist/ evolutionist
- ▣ Criteria for classifying living things
 - Categories' boundaries established on resemblances or relatedness
- ▣ Nomenclature (naming system and rules)

Today:

Concept: evolutionist

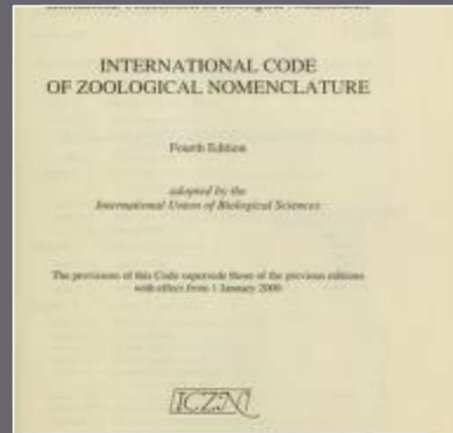
Criteria: diverse (increasing use of relatedness as estimated using phylogenetic tools, simple comparison still in use)

Nomenclature: based on Linnaeus' system

Linnaeus' nomenclature

- ▣ A species description (publication)

= a piece of legislation



must be compliant with a code (ICZN)

refutable only by a
another such piece of
legislation

- → synonymizations

Homonyms & Synonyms

Article 53.2. Homonyms in the genus group. In the genus group, two or more available names established with the same spelling are homonyms.

Article 60.2. Junior homonyms with synonyms. If the rejected junior homonym has one or more available and potentially valid synonyms, the oldest of these becomes the valid name of the taxon with its own authorship and date.

Linnaeus' nomenclature

▣ Synonymization

When one species is shown having already been described

- → another piece of legislation
- = **synonymy by priority**

Cosmolaelaps scimitus Womersley, 1956
Syn. *Hypoaspis antennata* Karg, 1993

When a species is shown better matching another genus

- → another piece of legislation
- = **synonymy by homotypy**

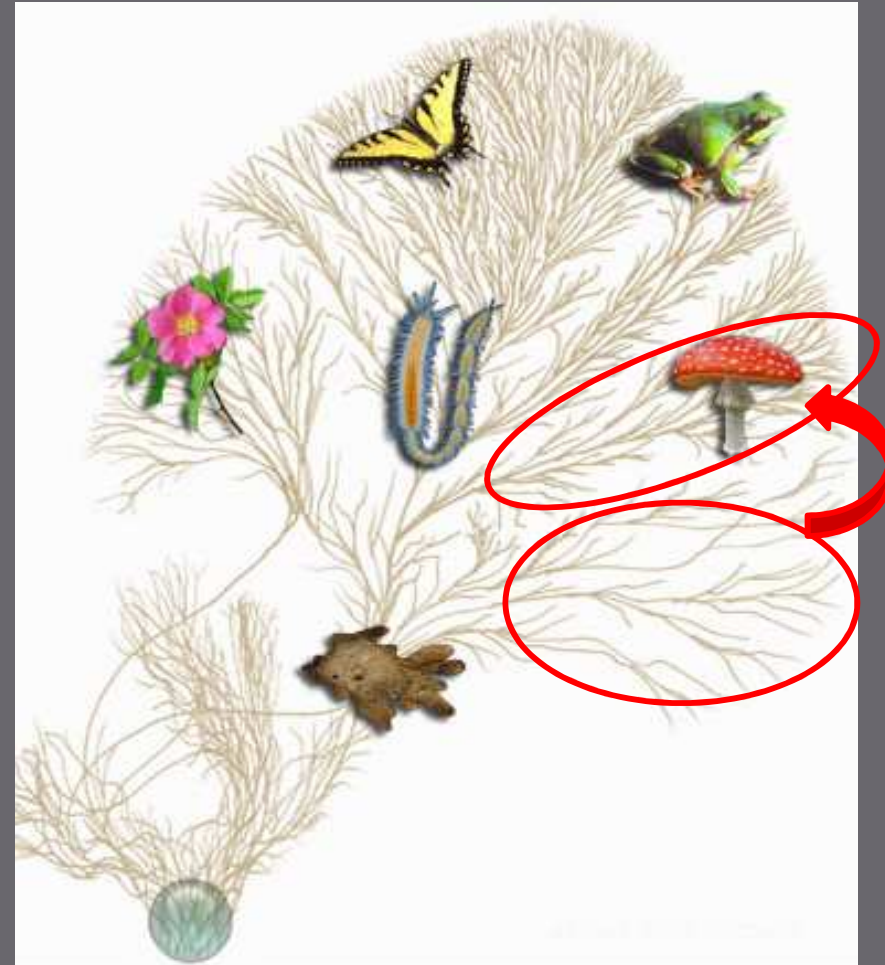
Dermanyssus gallinae (De Geer, 1778)
Syn. *Acarus gallinae* De Geer, 1778

Stratiolaelaps scimitus (Womersley, 1956)
Syn. *Cosmolaelaps scimitus* Womersley, 1956
Syn. *Hypoaspis antennata* Karg, 1993

Linnaeus' nomenclature = today use

▣ Synonymization
→ 'mobility' of specific entities between genera, families... and in some cases much higher-level taxa

e.g. Microsporidia (e.g. *Nosema* spp., honeybee digestive parasites)
long in « protozoan » Apicomplexa
now in fungi



≠ evidence of wrong science!
= just legal statement of science progress (a gradual ongoing process)

Problems in mite taxonomy and molecular tools' utility

Solving practical shortcomings using molecular tools

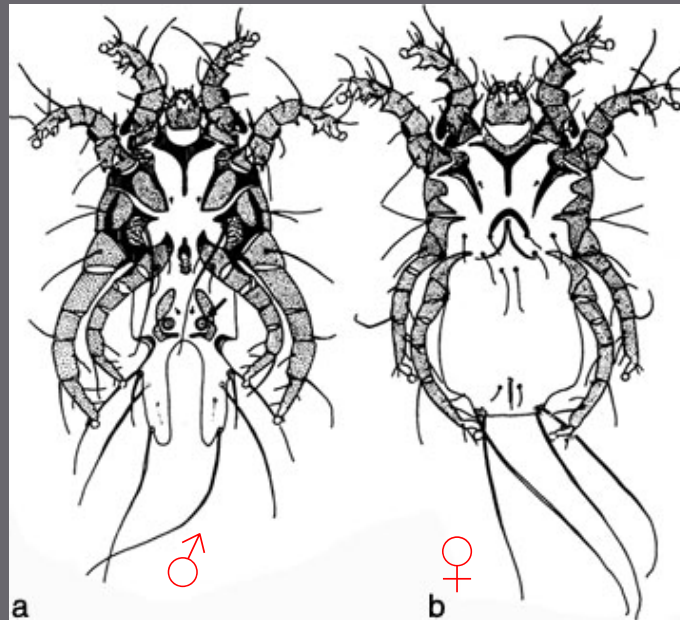
- Sex/stage polymorphism
- Large samples' issue

Contributing to improve the match between Linnean taxa and evolutionary history

- Delineating species boundaries

Assigning specimens to right linnean species

- ▣ Sexual dimorphism, stage polymorphism
 - Source of confusion
 - illusory multispecies entities

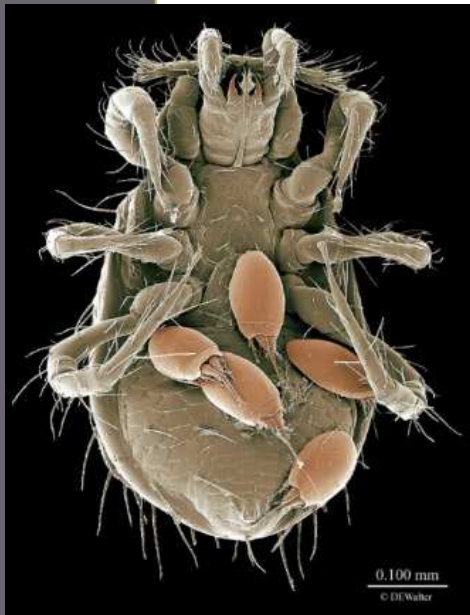


Assigning specimens to right linnean species

- ▣ Stage polymorphism may be long misleading: an example within Astigmata

Dugès 1834...

Description of
genus *Hypopus*



Dujardin 1847-49 ...

Hypopus ≠ a genus
= a young form of
genus *Gamasus*
(Mesostigmata!!)

Claparède 1868...

Hypopus = a male
adult of
Tyroglyphus
(Astigmata OK,
but wrong stage)

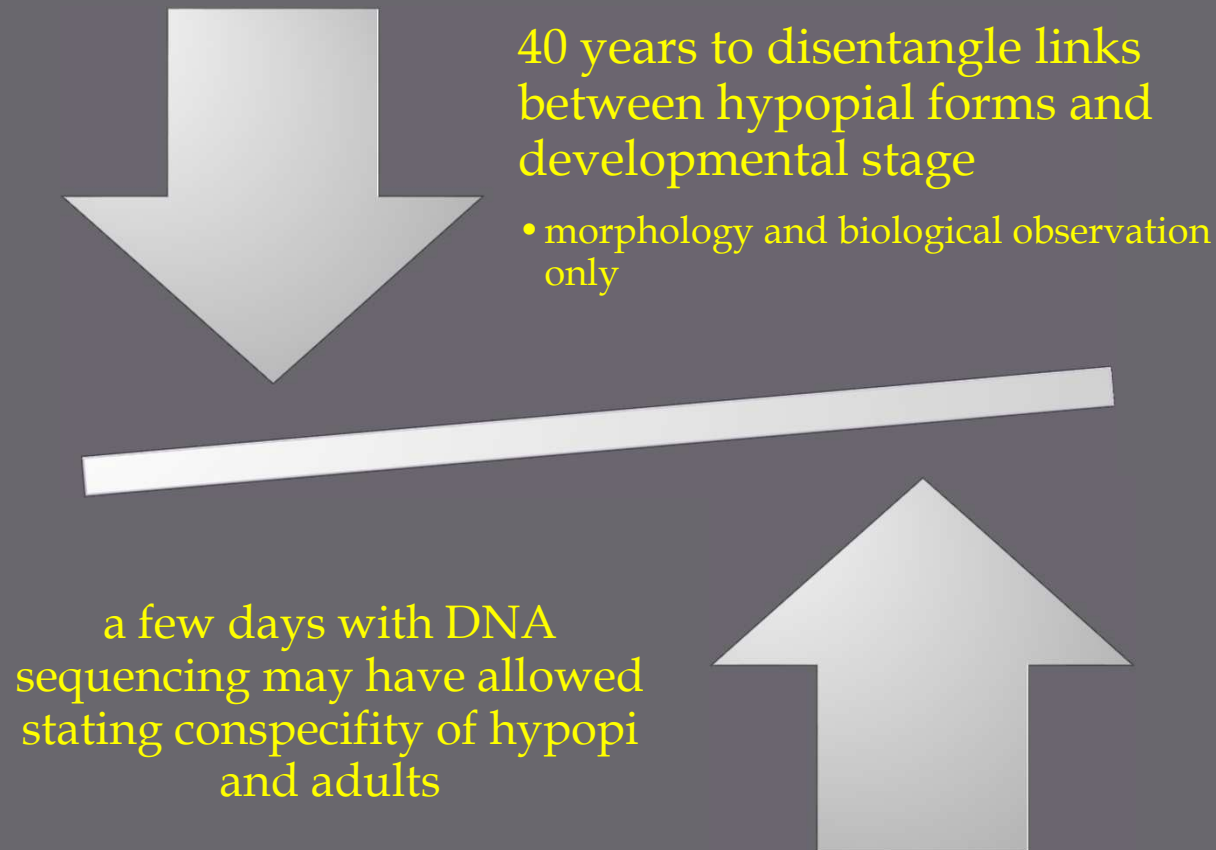
Mégnin 1874

Hypopus = a special
condition of
nymphal stages in
some « Acarina »
to preserve the
species under
adverse
circumstances

From Michael (1884)
The Hypopus Question, or the Life-History of certain Acarina

Assigning specimens to right linnean species

- ▣ Sexual dimorphism, stage polymorphism
 - Discovery of hypopi: DNA sequencing would have speeded up the understanding process ...



Assigning specimens to right linnean species

▣ Analyzing large samples in community ecology

A community = a group of populations of different species, different genera, different families...,
living together in an ecosystem

- ▣ Characterizing communities to compare different habitats commonly requires mite inventories from a couple hundred samples
 - usually each containing >1000 mites

Assigning specimens to right linnean species

- ▣ Practical shortcomings
 - 2. Constraints in community ecology

E.g. Bioptipou project

Mean no of mites per sample: 1,469.44 ($\pm$ 1,349.83)



Slide-mounting +
microscope identification

100 samples x 5 sampling
campaigns → ca. 735,000 mites
to be identified

Too much time-consuming

Joint morpho-molecular
study

DNA sequences from
environmental soup + linkage
mite cuticules - DNA sequences
on a subsample of mites

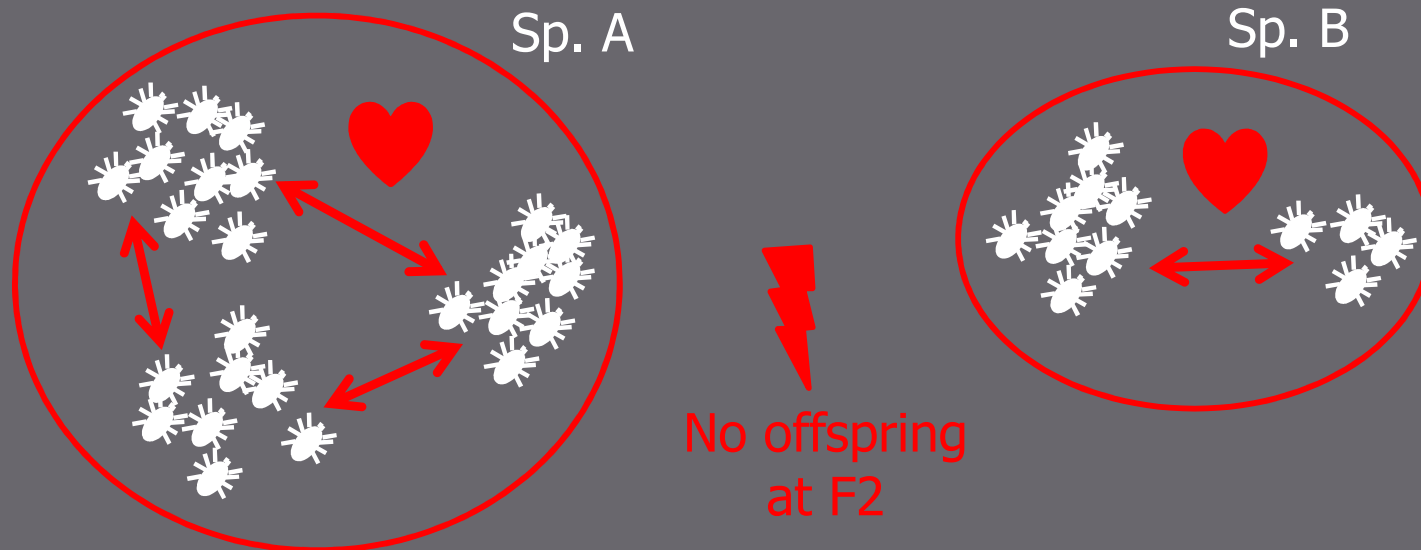
Feasible within
a 1-year time-frame

Improving match between Linnean taxa and evolutionary history

- Species = the only taxonomic level with a biological standard basis

« Species are groups of natural populations that can potentially interbreed to produce fertile offsprings; every species is isolated in terms of a reproduction from all other species. »

Ernst Mayr (1942)



Improving match between Linnean taxa and evolutionary history

▣ Main problems with species delineation in mites



Mites and the reproductive isolation evidence	• Highly difficult (if not impossible) to biologically test for reproductive incompatibility in most cases
Intraspecific diversity insufficiently represented by type materials	• Highly (morphologically) variable species commonly described as several different species
Cryptic species common in mites	• Morphological differentiation not yet visible, whilst reproductive incompatibility reached • Species difficult to be described

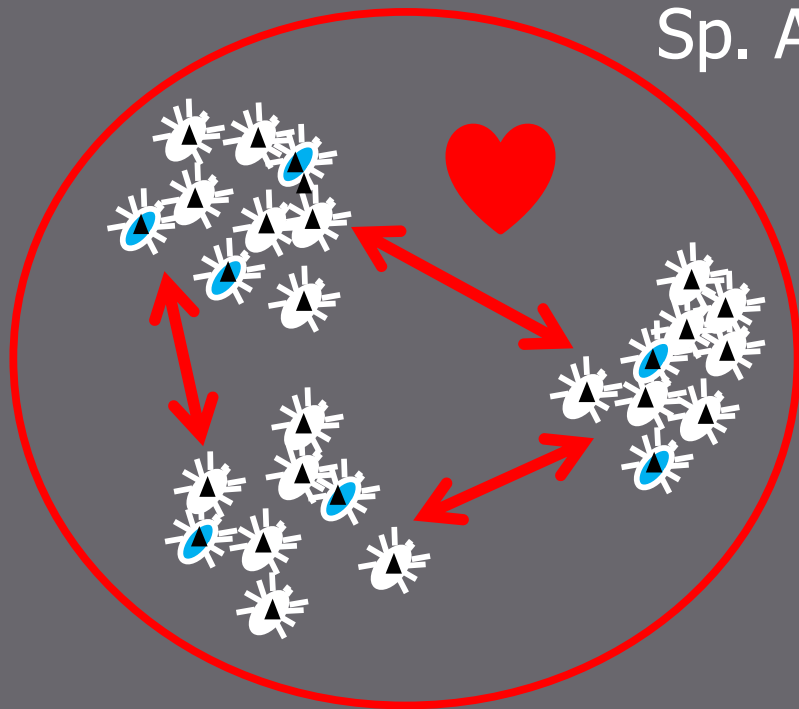
Improving match between Linnean taxa and evolutionary history

- ▣ Intraspecific diversity and type materials

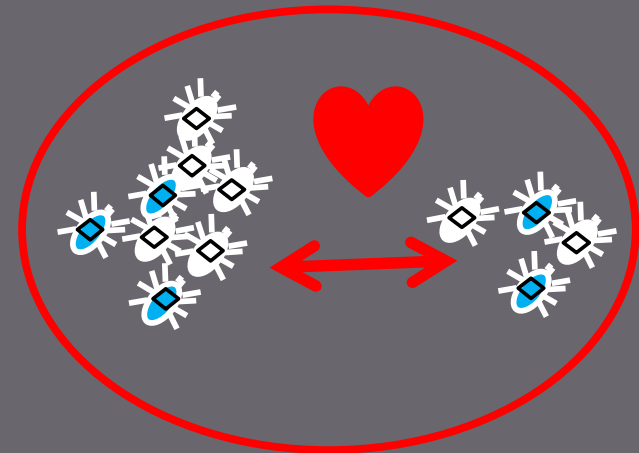


Blue dorsum: not a good diagnostic character!

Sp. A



Sp. B



No offspring
at F2

Cryptic species

- Some apparently generalist species in fact are a complex of specialists with no or almost no morphological differences



Contents lists available at SciVerse ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev



Cryptic species of mites (Uropodoidea: *Urobovella* spp.) associated with burying beetles (Silphidae: *Nicrophorus*): The collapse of a host generalist revealed by molecular and morphological analyses

Wayne Knee^{a,b,*}, Frédéric Beaulieu^b, Jeffrey H. Skevington^b, Scott Kelso^b, Mark R. Forbes^a

^a Carleton University, 1125 Colonel By Drive, Department of Biology, 209 Nesbitt Building, Ottawa, Ont., Canada K1S 5B6

^b Canadian National Collection of Insects, Arachnids and Nematodes, Agriculture and Agri-Food Canada, 960 Carling Avenue, Ottawa, Ont., Canada K1A 0C6

ARTICLE INFO

Article history:

Received 17 February 2012

Revised 12 June 2012

Accepted 17 June 2012

Available online 22 June 2012

Keywords:

Urobovella nova

ABSTRACT

Urobovella (Mesostigmata: Uropodoidea: Urodinychidae) species are among the most common mites associated with carrion-feeding *Nicrophorus* (Silphidae) beetles. Previous taxonomic understanding suggests that a single host generalist, *U. nova*, disperses and lives with *Nicrophorus* species worldwide (reported from at least seven host species). Using morphometrics and morphological characteristics, as well as partial cytochrome oxidase I (COI) and the entire internal transcribed spacer 2 (ITS2) markers, we tested whether this apparent generalist is truly a generalist or rather a complex of cryptic species with narrower host ranges. Based on deutonymph mites collected from 14 host species across six countries and 17 provinces or states, we show that *U. nova* represents at least five morphologically similar species

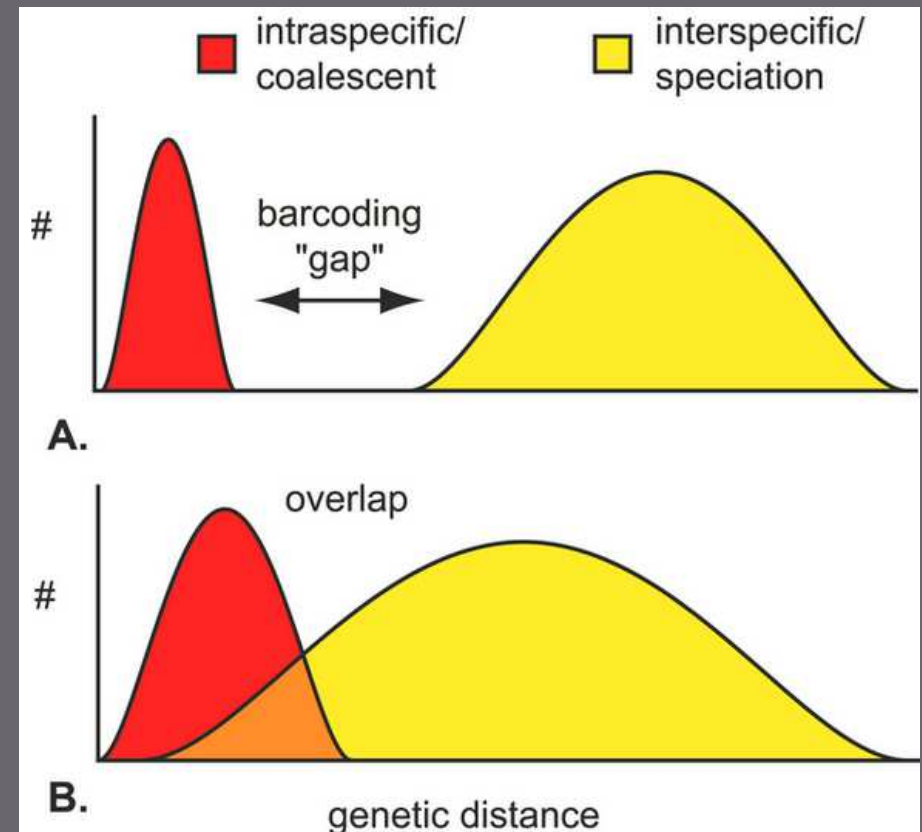
DNA sequence analyses utility

- 1) Checking conspecificity between different morphs
- 2) Fast recording different MOTUs from a soup or from individuals
- 3) Testing the level of reproductive isolation between populations
- 4) Checking how informative a morphological character is

DNA sequences useful, though ≠ a simple species tattoo ID provided by nature

- ▣ In any case, preliminary work required:
Link between a DNA sequence and a given linnean
taxon is not straight

Preliminary molecular analyses of dozens/hundred morphologically identified individuals from the different species under test are required to decide which DNA sequence(s) are to be assigned to which taxon



DNA sequence analyses utility and constraints

- 1) Checking conspecificity between different morphs
- 2) Fast recording different MOTUs from a soup or from individuals

Two very easy actions provided that:

- (1) the gene region under test is variable enough to discriminate between species,
- (2) one knows how variable it is within the species under consideration (not a single DNA sequence per species, different levels of variation among species)



(2) Another possibility is to forgive getting linnean id (identifying molecular operational taxonomic units), **but if so, important biological information is lost!!**

DNA sequence analyses utility and constraints

3) Testing the level of reproductive isolation between populations

Possible by confronting different phylogenetic trees from different gene regions instead of bioassays

One tree per gene region → tree-like or network-like interrelationships?

Distal clades congruent interspecifically, incongruent intraspecifically

Requirement = fundamentals of the phylogenetic approach

DNA sequence analyses utility and constraints

4) Checking how informative a given morphological character is

Possible by comparing different populations and mapping morphological characters onto phylogenetic trees

- need to examine dozens individuals from different populations of close species
- some intraspecifically variable characters may be detected

Requirement = fundamentals of the phylogenetic approach

Phylogenetics and molecular biology: no link in first

1858: elaboration of evolutionary theory, Darwin and Wallace

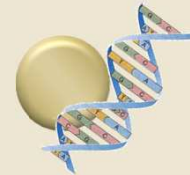
1950's:
Hennig,
Phylogenetic
classification

1970's:
Molecular
phylogeny,
Fitch

1859: *Origin of Species*
by Darwin

1866: Ernst Haeckel,
first phylogeny-
based evolutionary
tree, precursor
concept

1953: Discovery of
the structure of
DNA double helix,
Watson & Crick



Phylogeny

- ▣ Phylogenetic tree
 - A key construction in biology
 - Illustrates in 2 dimensions parental relationships between species or groups of species: time x diversity
- ▣ Tree of life and evolution of species

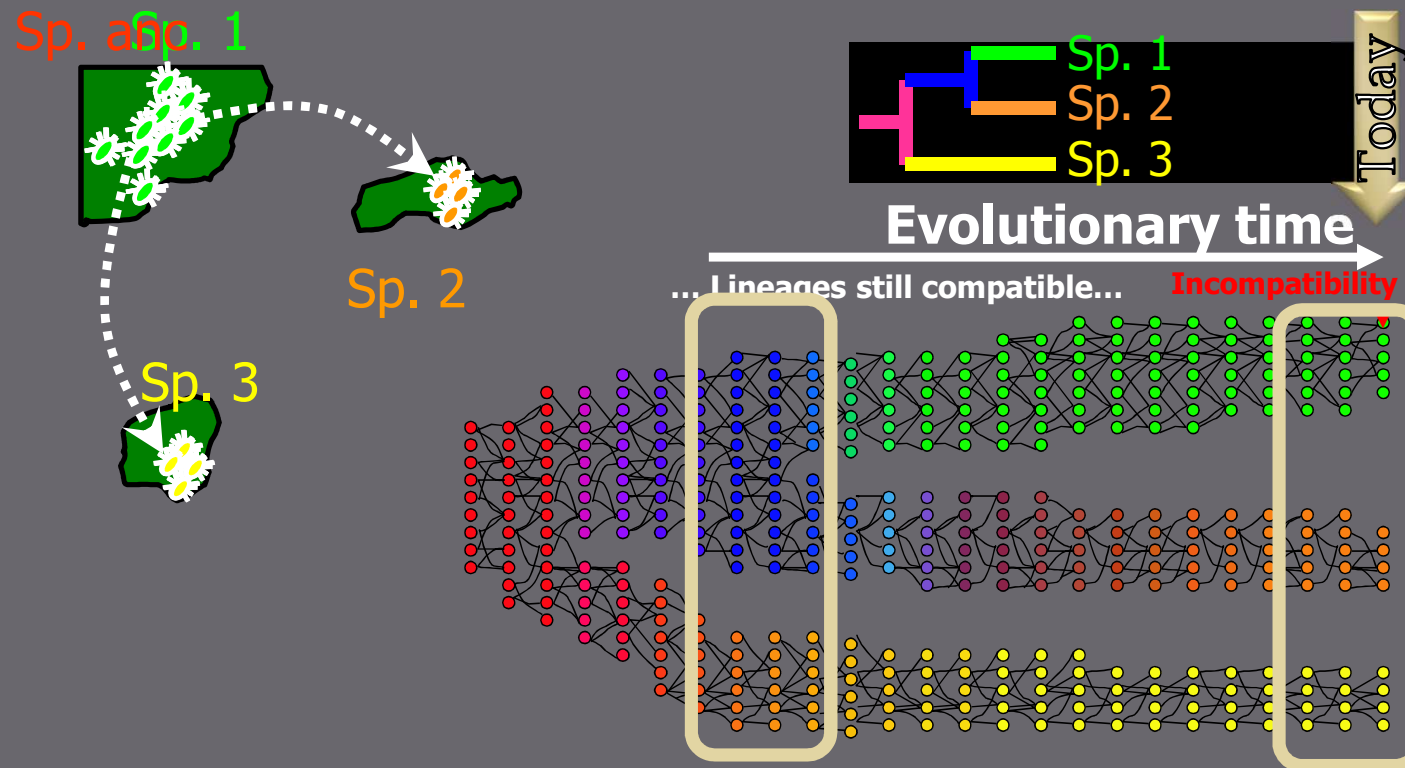
Etymology and definition of phylogeny

Ancient greek *phûlon*, tribe, race → genus, species... and *-gen-*, same root as in generate, genesis...

Following Darlu and Tassy (1993)

- **History of progenies in living organisms**

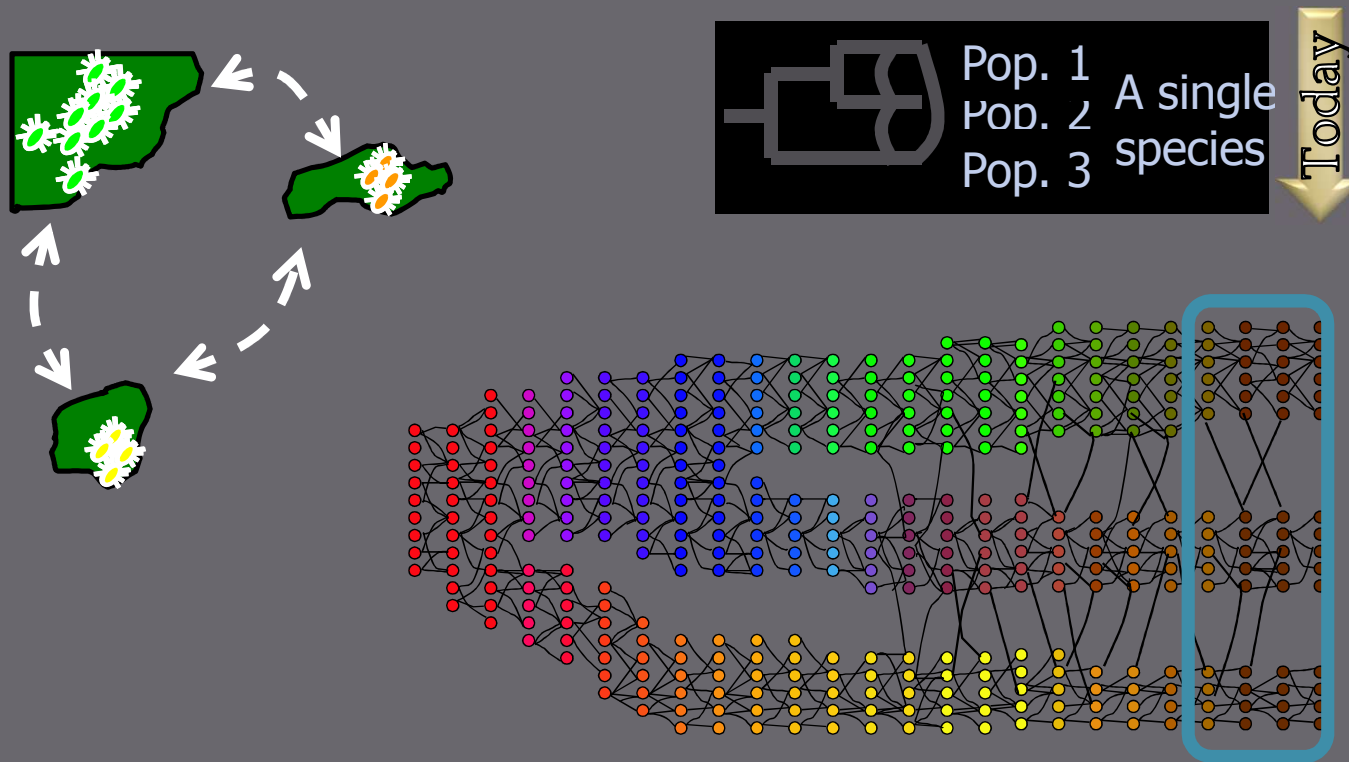
Evolutionary framework: species and speciation



Speciation = a continuous process (though not regular)

Perpetual differentiation of populations at any level (DNA sequences, morphology, biochemical activity ...)

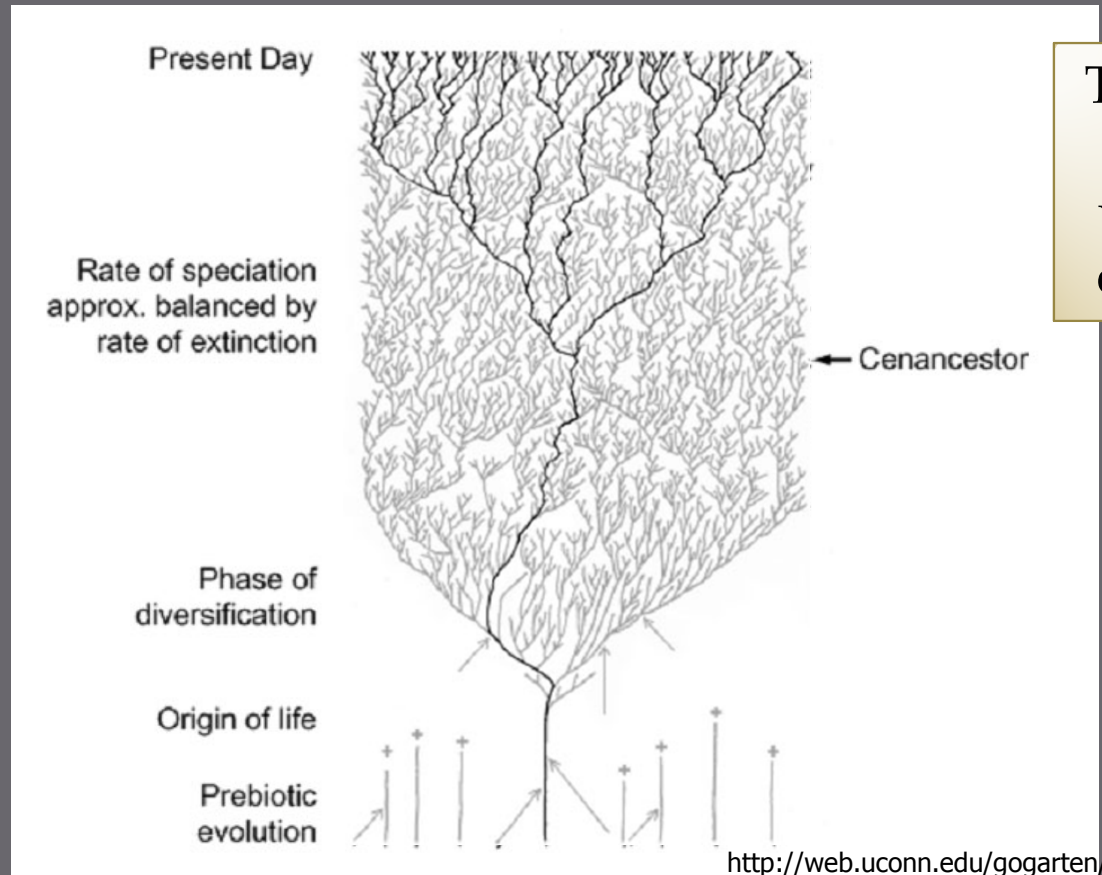
Evolutionary framework: species and speciation



Speciation = a **REVERSIBLE** process, if connexion between populations is restored before incompatibility is reached

Evolutionary framework: species and speciation

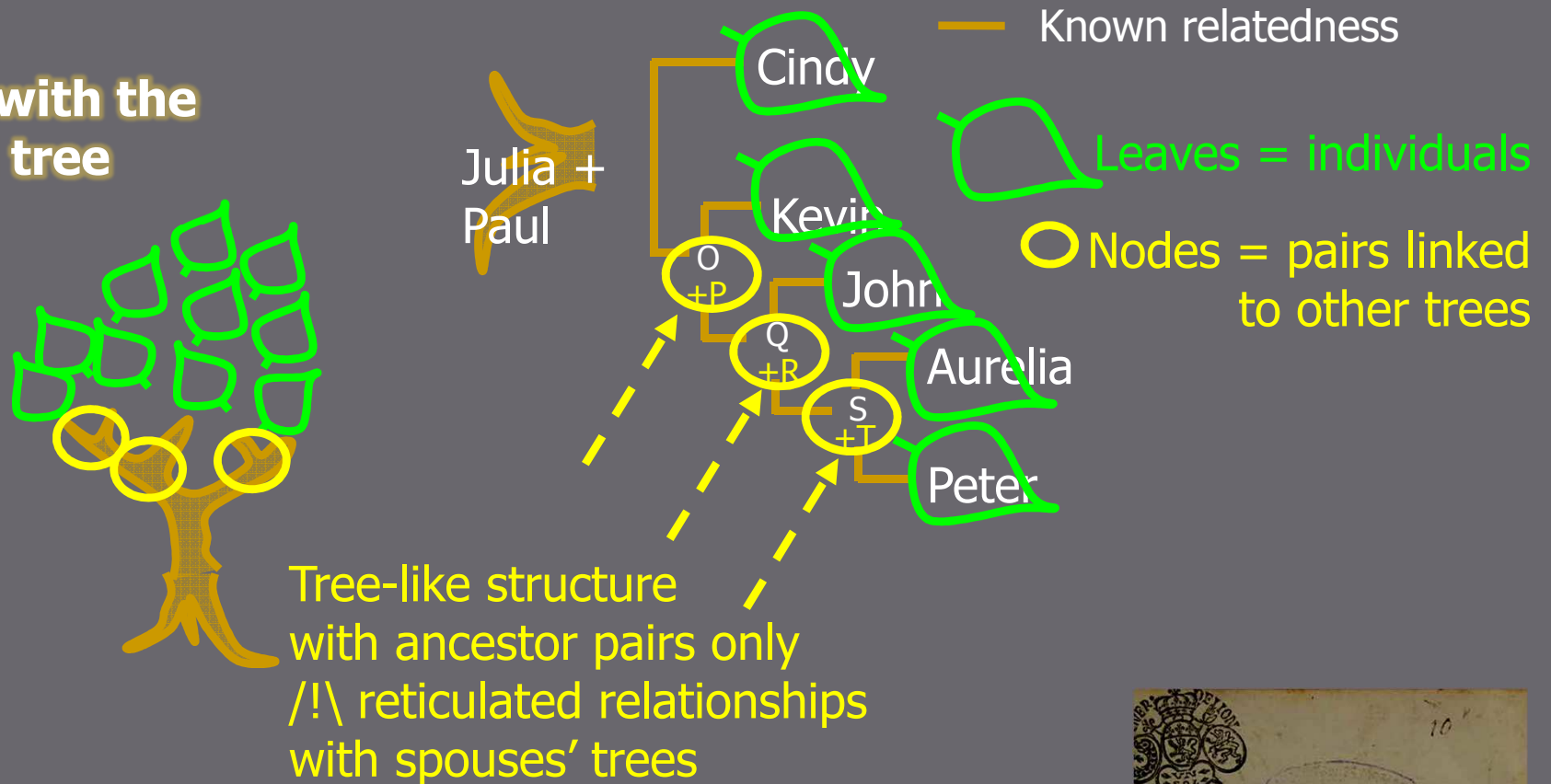
Current global biodiversity = a smallest part of life history



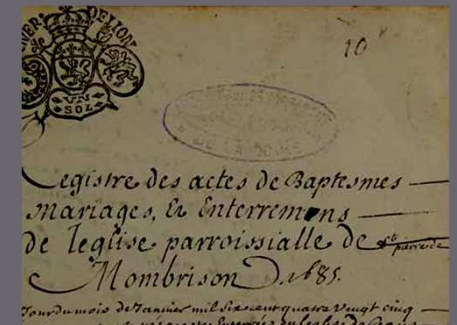
Tree of life, with
(grey) and
without (black)
extinct lineages

Phylogenetic tree

Comparison with the genealogical tree

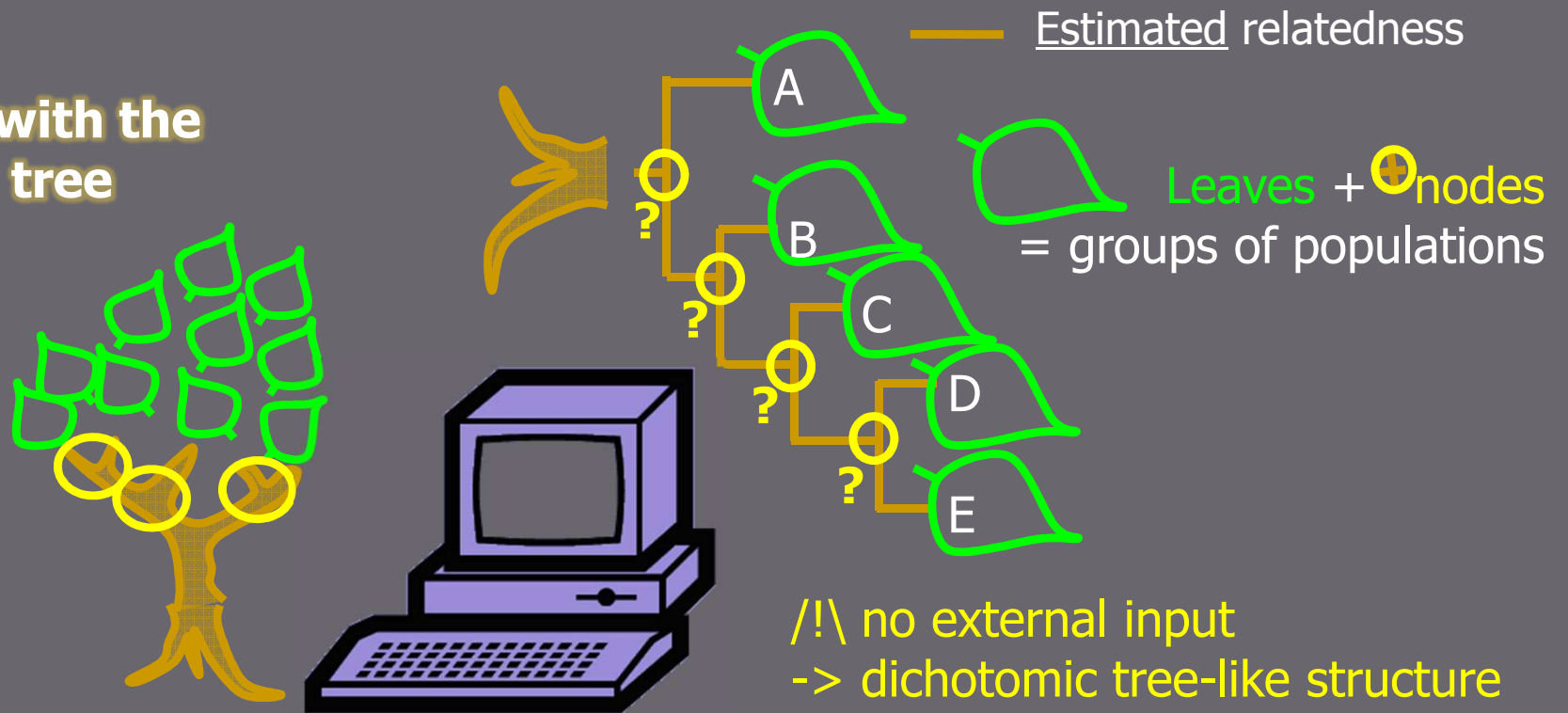


Genealogical tree : Roots+nodes+leaves at least partly known
Part of the network (within a branch)



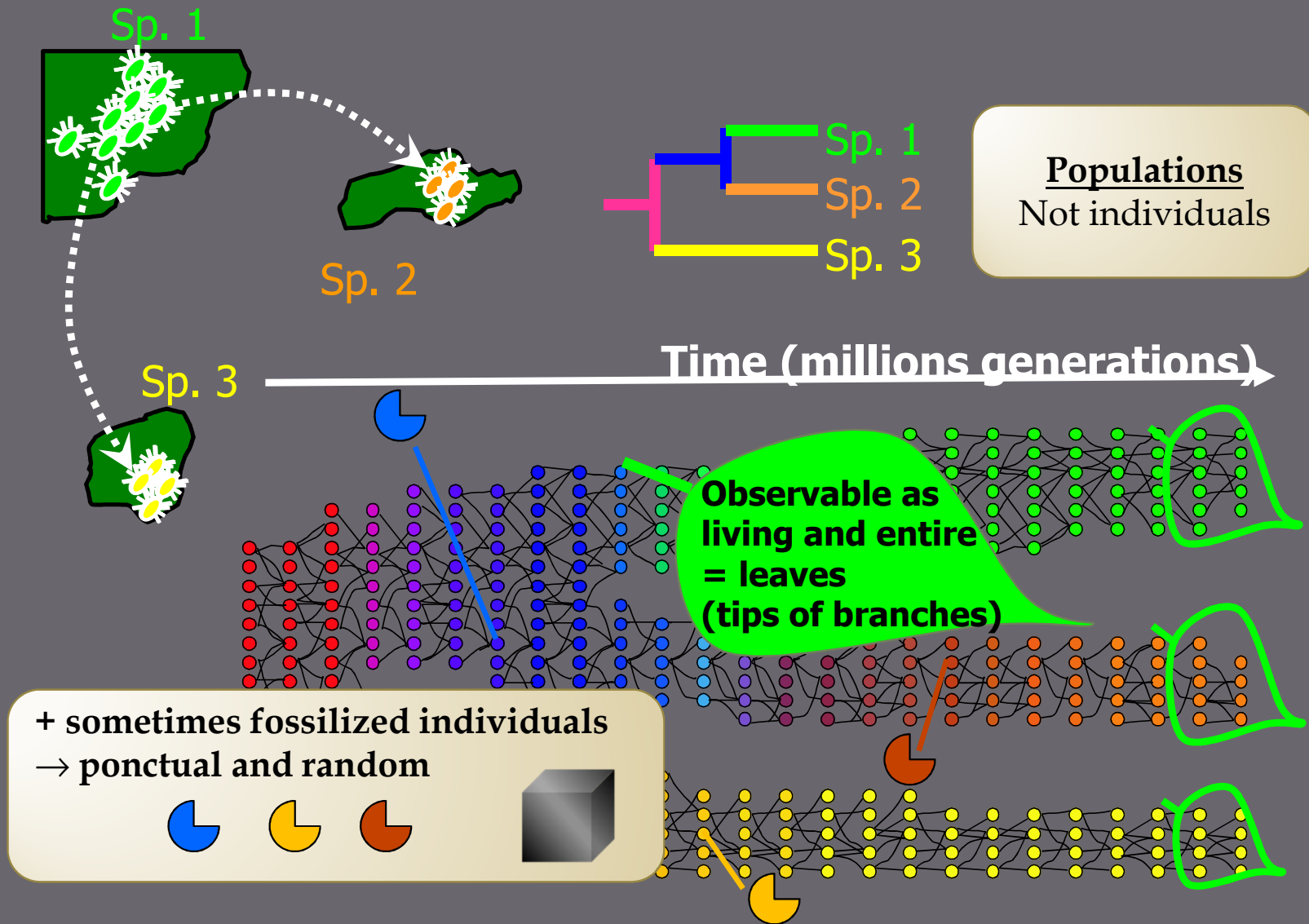
Phylogenetic tree

Comparison with the genealogical tree



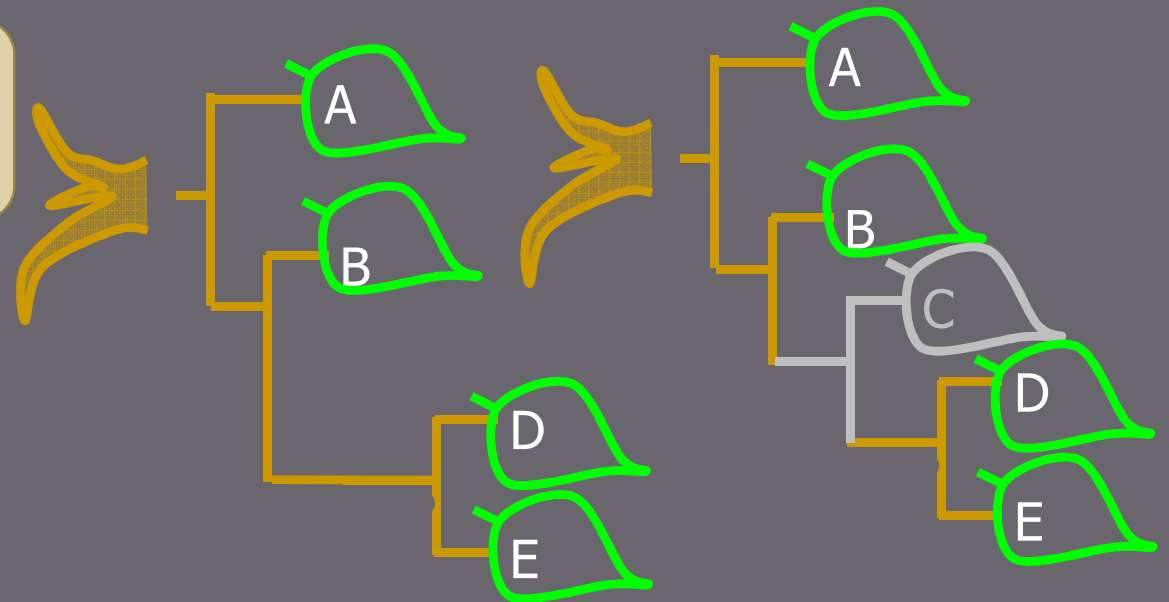
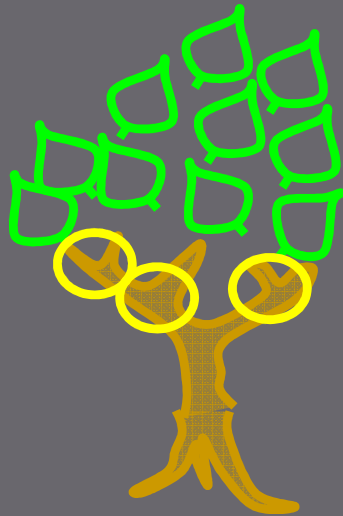
The **phylogenetic** tree : **Reconstruction of progenies' history by calculation**
from population leaves' (ie observable) characters
Tree-like representation OK, at least basally

Phylogenetic tree



Phylogenetic tree

/!\ Keep in mind that
leaves' sample is also
sparse and +/- random



Often not all leaves are available in a given study:

- Extinct species and characters mostly unavailable as fossils
- Some extant species rather unfrequent and/or difficult to be collected

→ some lacking steps

How to read a phylogenetic tree?

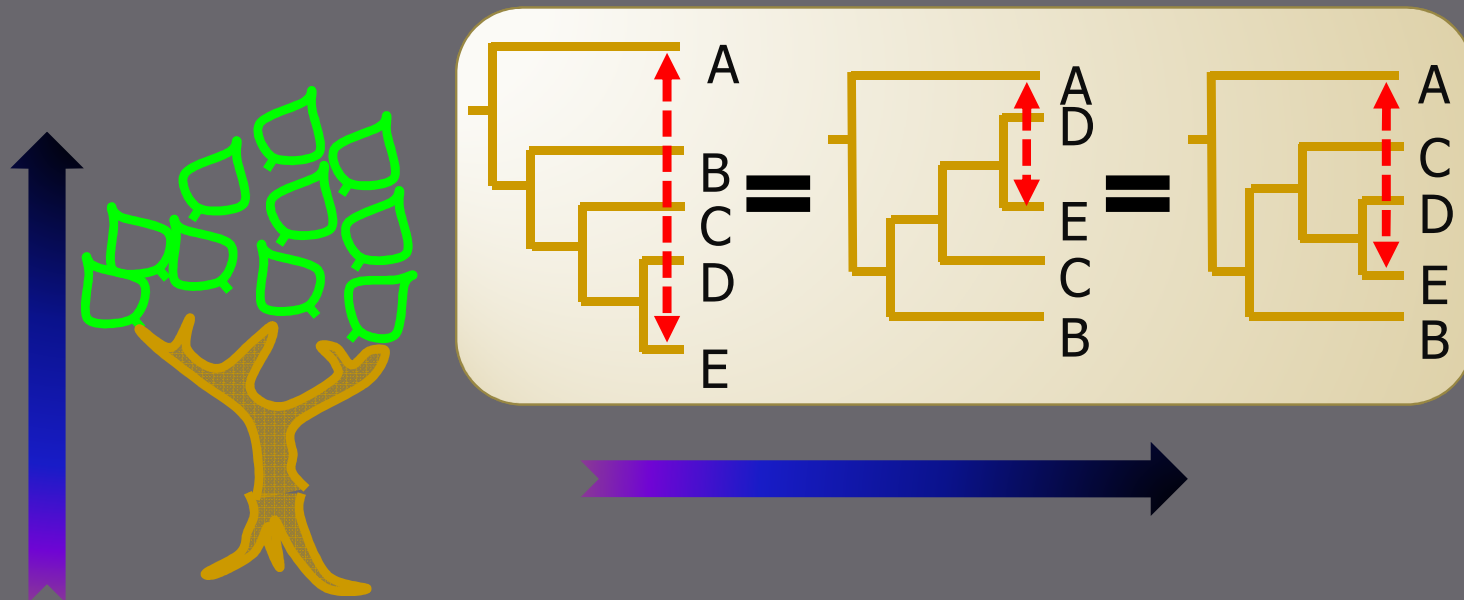
👉 Transverse distances
between branches without any signification
(no more than a drawing effect)



Reading sense: from roots to leaves

How to read a phylogenetic tree?

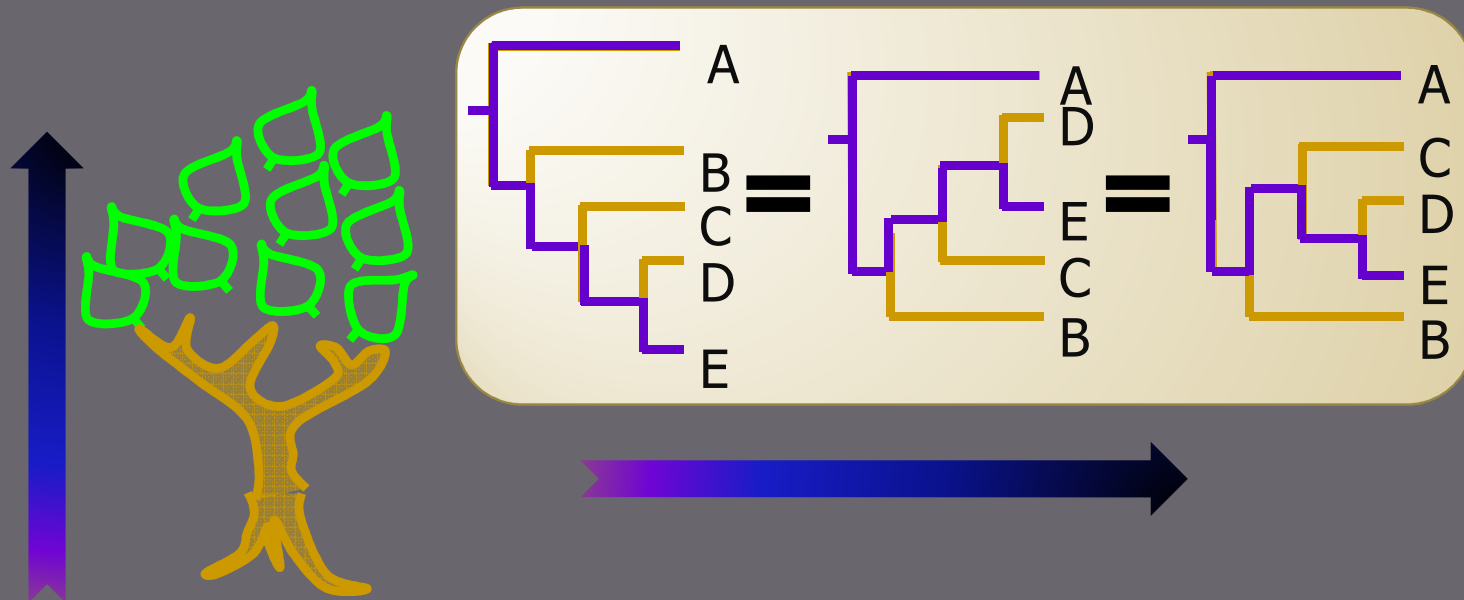
✋ Transverse distances
between branches without any signification
(no more than a drawing effect)



Reading sense: from roots to leaves

How to read a phylogenetic tree?

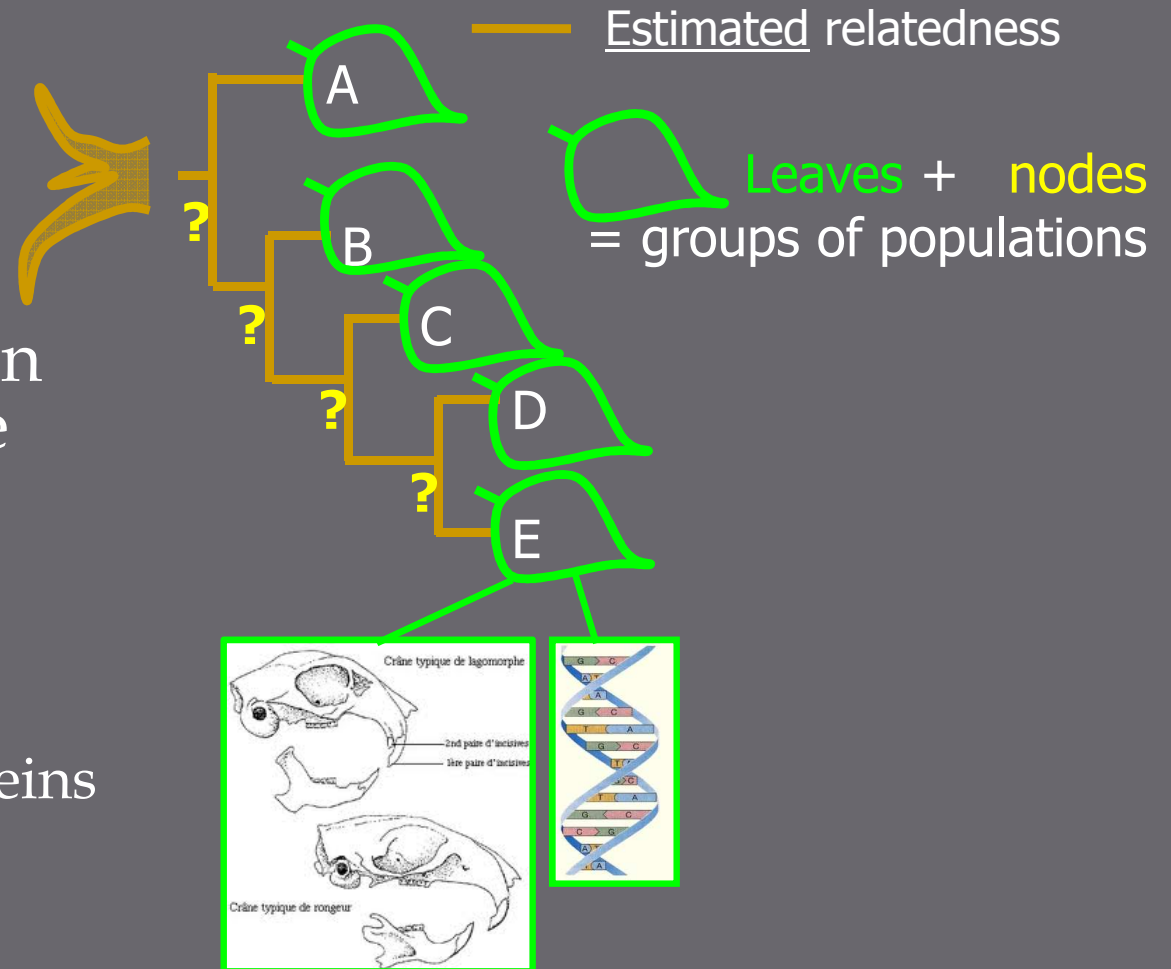
👉 Transverse distances
between branches without any signification
(no more than a drawing effect)



Reading sense: from roots to leaves

Phylogenetic reconstruction

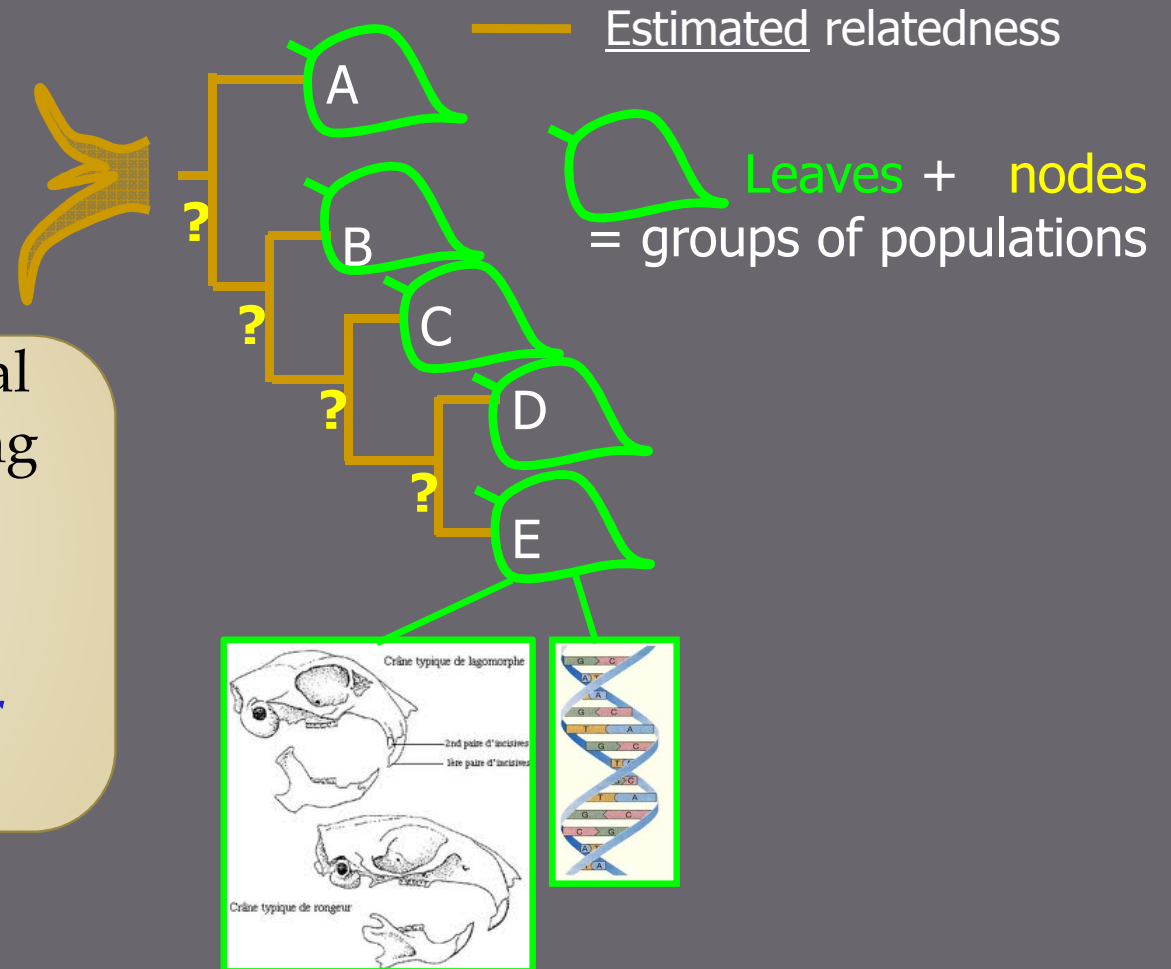
- ▣ Types of characters in leaves which may be used to process calculations:
 - External morphology
 - Internal anatomy
 - DNA sequences, proteins
 - Life traits ...



Phylogenetic reconstruction

Basic concept = historical path of progeny implying modification

→ *Similarity is to be understood in terms of common ancestry*

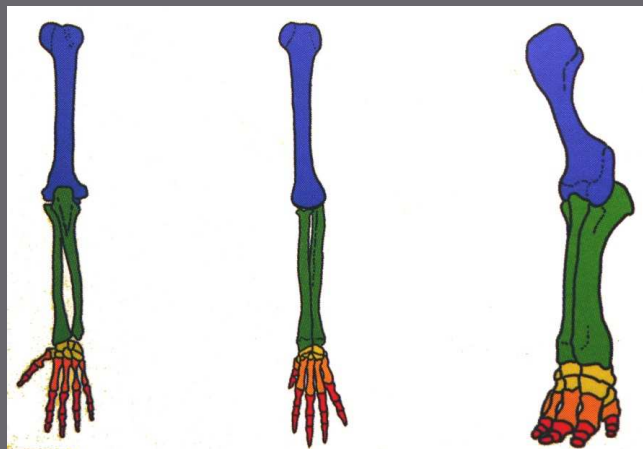


Phylogenetic reconstruction

▣ The concept of homology

Homology = a link between two characters that are present in two (or more) different species because they both inherited it from a common ancestor

Ex. structure of bones in forelegs of mammals



Hervé & Poinso (2013)

→ identical organization **because** inherited from a common ancestor
→ humeri (green) are homologous to each another

Phylogenetic reconstruction

▣ The concept of homology

Some of the similarities are not the result of heritage from a common ancestor – more resulting from similar selective pressure

- ▣ **Convergence and reversion (homoplasies)**
= misleading similarities

E.g. convergence :
mole cricket (insect)
and mole (mammal)

Two burrowing species → similar
selective pressure though no close
common ancestor



Mole cricket (*Gryllotalpa gryllotalpa*)



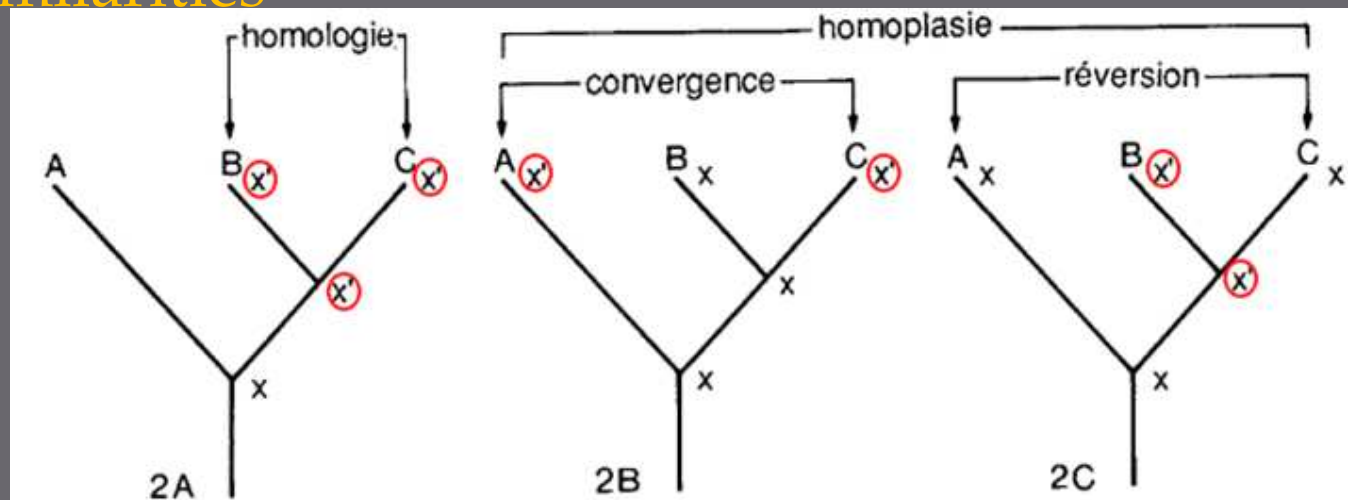
Mole (*Talpa europaea*)

Phylogenetic reconstruction

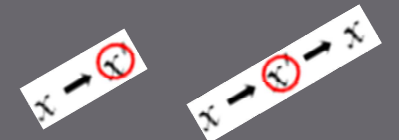
▣ The concept of homology

Some of the similarities are not the result of heritage from a common ancestor – more resulting from similar selective pressure

- ▣ Convergence and reversion (homoplasies) = misleading similarities



Evolution of character X (2 states: x et x') along the phylogenetic history of a hypothetical group of species named A, B et C resp.



Phylogenetic reconstruction

▣ 2 different approaches

Phenetic approach

- Sokal et Sneath (1963, 1973)

Principle = global analysis of similarities (including homologies + convergences + reversions)
Method = calculating 1°) a distance matrix, 2°) similarity indices

« Birds of a feather flock together »

Evolutionary approach

- Simpson (1961)
- Mayr (1969)

Principle = reconstruction of progeny history based on homologous similarities
Method = observing how the different character states are mapped on tree(s)

Tracing the most likely historical path of progenies

Phylogenetic reconstruction

A single
historical
path of
progenies



A
single
true
tree

A huge number of possible
trees!

E.g. 15 evolutionary histories
possibles for 4 leaves (a, b, c, d)



Vekemans

Phylogenetic reconstruction



Number T of possible trees according to the number k of leaves under test:

$$T'_n = \prod_{k=2}^n (2k-3)$$

- With 10 leaves ($k = 10$) : 34,459,425 possible trees
- With 20 leaves ($k = 20$) : 8×10^{21} possible trees

Phylogenetic reconstruction

- ▣ Evolutionist approach:
 - Exploration of different alternative possibilities; heavy analyses (long calculations)
 - Much more efficient (more resolved trees), since the methods reduced noise brought by convergence and reversions (called « homoplasies »)

Important = order of relatedness links over time
Not similarity as such

Keep in mind that phylogenetics predicts past

Phylogenetic reconstruction

- ▣ Evolutionist approach:
 - Hennig (1950, 1966) : taking into account whether a character state is ancestral *vs* derived

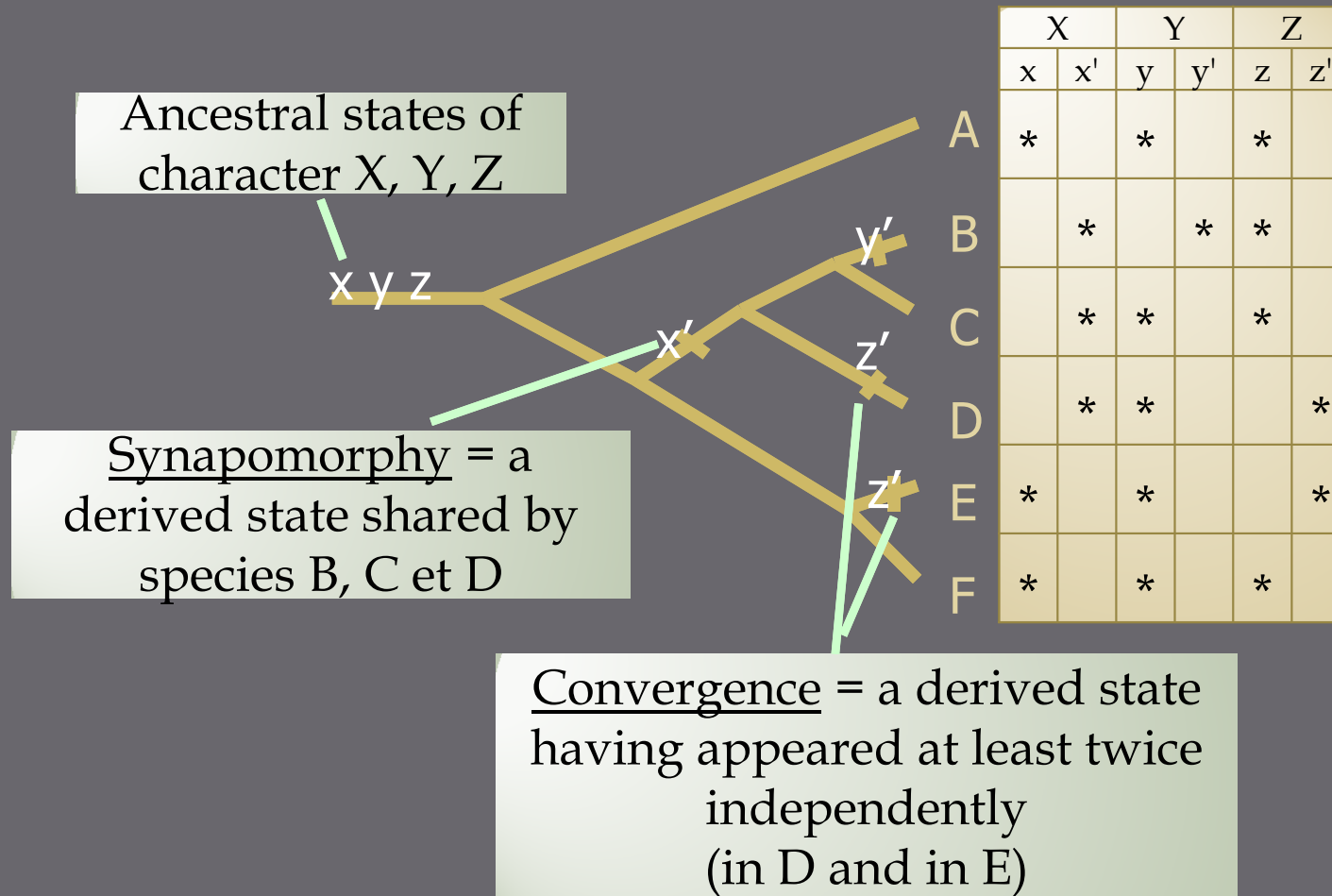
Informative and non informative homology

Characters which keep having the ancestral state along a branch do not allow stating about relatedness

Focus on derived characters (**apomorphic**),
that (seem to) have been **transmitted as such**
(**modified**) to the different members of a branch
= **synapomorphies**
(e.g. feathers in birds amongst Amniota)

Phylogenetic reconstruction

- ▣ Identifying derived and ancestral states: from root to leaves



Phylogenetic reconstruction

- ▣ Calculations to be processed from dozens characters
→ choosing the tree with the least misleading similarities

Different algorithms in different methods

Maximum parsimony

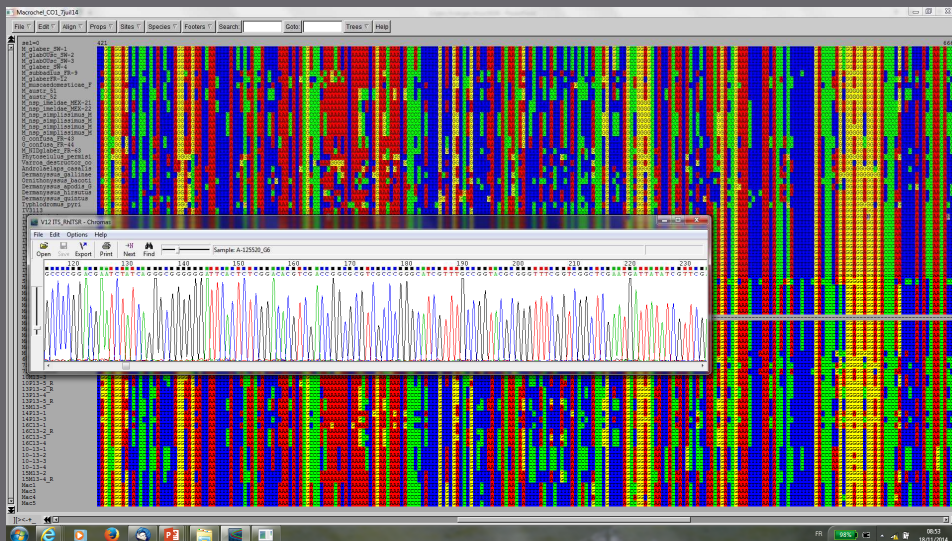
Minimizing the number of steps (changes of character states) over the whole tree (not suitable with DNA)

Maximum likelihood, Bayesian

Using likelihood or Bayesian methods, to select the best tree following some models of evolution

Biotechnological progress and improvement of phylogenetics

- From the discovery of DNA molecular structure (Watson et Crick, 1953), new biotechnologies very quickly developed and opened the way of in-depth investigation of biological diversity



1956-1988 : PCR
(Polymerase Chain
Reaction)

1975-1980 : Sanger
sequencing

2005-... : high-
throughput
sequencing
methods (NGS:
Next Generation
Sequencing)

Progresses thanks to molecular biology



- ▣ More characters, much easier to be encoded for phylogenetic reconstruction

[illegible]

Example of a matrix of morphological characters in *Dermanyssus*

[illegible]

A matrix of DNA characters
= sequence alignment

Progresses thanks to molecular biology

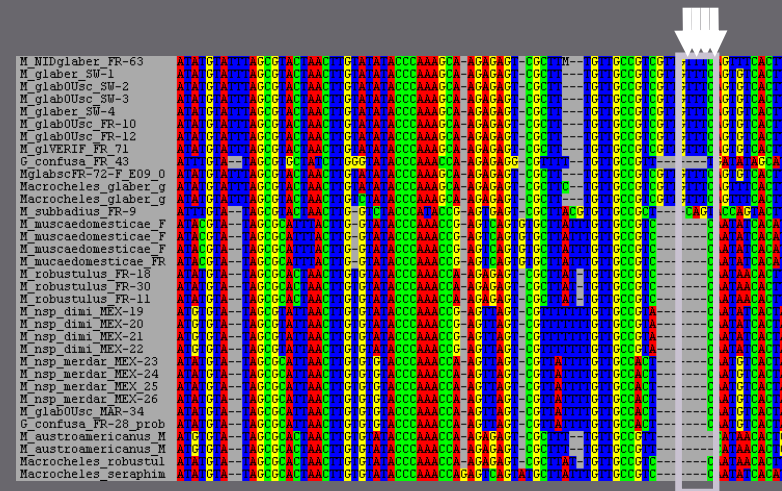
- Homology of characters states (columns of matrix) required both for morphology and DNA sequences
- How to assign homologous status to nucleotides?
 - A DNA sequence covers a series of loci: sequence alignments makes homologous nucleotides aligned as columns

Confronting homologous character states

- E.g. **Man's arm and bat's wing YES**
- ~~forelegs in ??? And in ???~~ NO



⇒ When dealing with DNA sequences,
the quality of alignment
is a crucial issue



True or false?

- ~~▫ “Morphological characters are not interesting as they are under the selective pressure, as opposed to DNA sequences, so mostly homoplastic (eg convergence), as opposed to DNA”~~

FALSE – See Insecticide resistance: genes coding targets of insecticides undergo same mutations in different families of insect and mite
E.g. *kdr* mutation in the voltage-gated sodium channel found in aphids and flies (pyrethroid resistance)

- “Morphological characters are much more difficult and long to be observed and encoded than DNA sequences”

TRUE – morphological skill transfer requires years or dozens years training while molecular lab practices may be acquired rather quickly

Progresses thanks to molecular biology



- ▣ More characters, much easier to be encoded for phylogenetic reconstruction
 - DNA sequences = 'words' written using a 4-letter alphabet
 - A standard and straight definition of the 4 states in sequence alignments, as opposed to much more tricky definition of the n states of the so much polymorphic morphological characters
- ▣ New models for reconstructing the history of progenies
 - General rules of DNA evolution easier to be evidenced than general rules of morphological evolution
- ▣ Makes possible studies on the speciation process, with questions at the interface between inter- and intraspecific levels