Phylogenetic trees IV
Maximum Likelihood

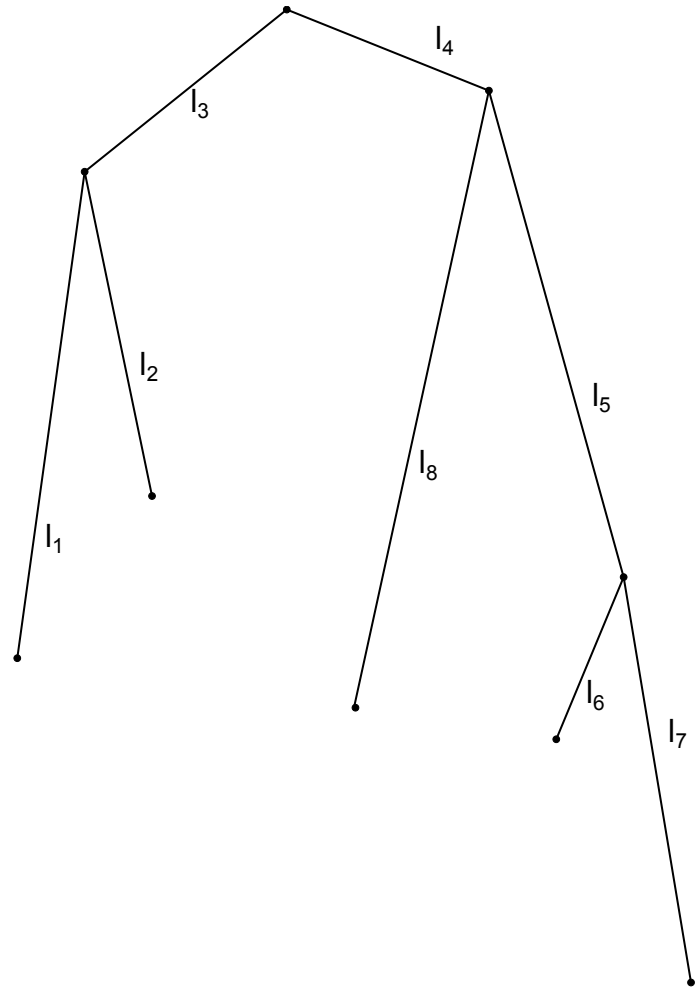
Gerhard Jäger

ESSLLI 2016

Theory

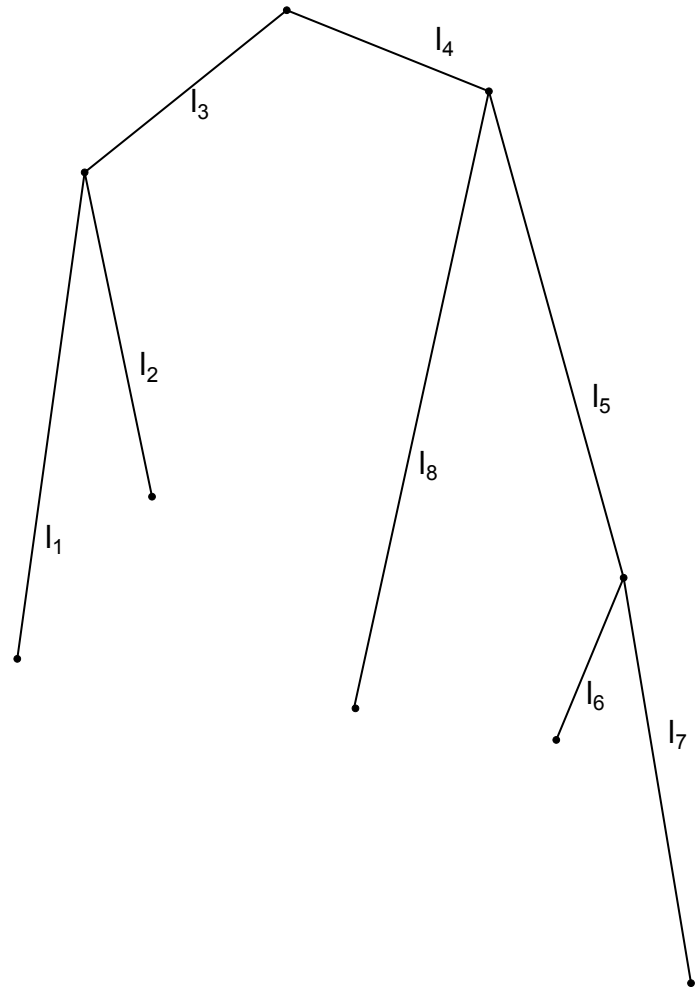
Recap: Continuous time Markov model

$$
\begin{aligned}
P(t) &= \begin{pmatrix} s + re^{-t} & r - re^{-t} \\ s - se^{-t} & r + se^{-t} \end{pmatrix} \\
\pi &= (s, r)
\end{aligned}
$$



Likelihood of a tree

background reading: Ewens and Grant (2005), 15.7



- simplifying assumption: evolution at different branches is independent
- suppose we know probability distributions v_t and v_b over states at top and bottom of branch l_k
- $\mathcal{L}(l_k) = v_t^T P(l_k) v_b$

Likelihood of a tree

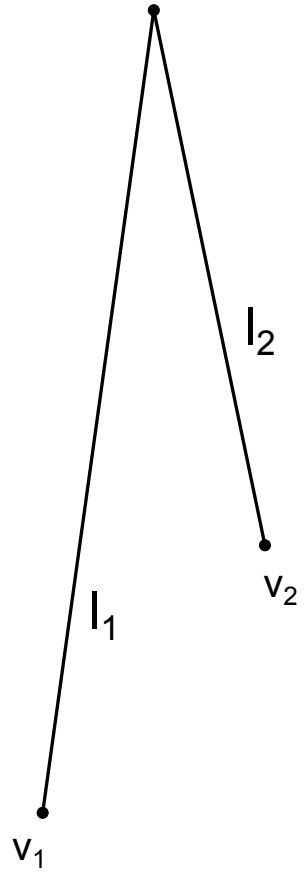
- likelihoods of states $(0, 1)$ at root are

$$v_1^T P(l_1) v_2^T P(l_2)$$

- log-likelihoods

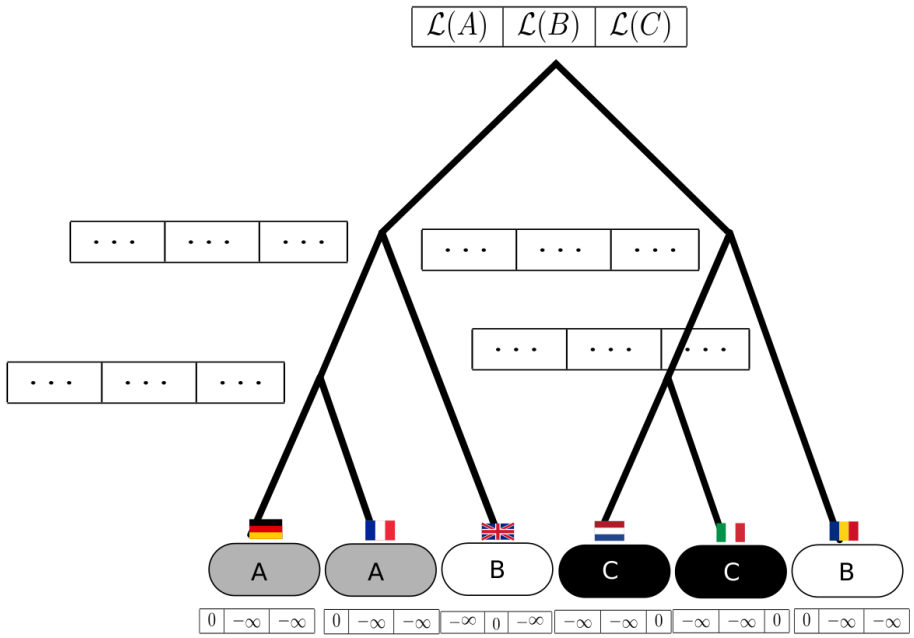
$$\log(v_1^T P(l_1)) + \log(v_2^T P(l_2))$$

- log-likelihood of larger tree: recursively apply this method from tips to root



(Log-)Likelihood of a tree

$$\log \mathcal{L}(\text{tips below}|\text{mother} = s) =$$
$$\sum_{d \in \text{daughters}} \sum_{s' \in \text{states}} \log P(s \rightarrow s'|\text{branchlength}) + \log(\mathcal{L}(\text{tips below } d|d = s'))$$



(Log-)Likelihood of a tree

- this is essentially identical to Sankoff algorithm for parsimony:
  - $\text{weight}(i, j) = \log P(l_k)_{ij}$
  - weight matrix depends on branch length $\rightarrow$ needs to be recomputed for each branch
- overall likelihood for entire tree depends on probability distribution on root
- if we assume that root node is in equilibrium:

$$\mathcal{L}(\text{tree}) = (s, r)^T \mathcal{L}(\text{root})$$

- does not depend on location of the root ($\rightarrow$ time reversibility)
- this is for one character — likelihood for all data is product of likelihoods for each character

(Log-)Likelihood of a tree

- likelihood of tree depends on
  - branch lengths
  - rates for each character
- likelihood for tree *topology:*

$$\mathcal{L}(\text{topology}) = \max_{l_k\text{: } k \text{ is a branch}} \mathcal{L}(\text{tree}|\vec{l_k})$$

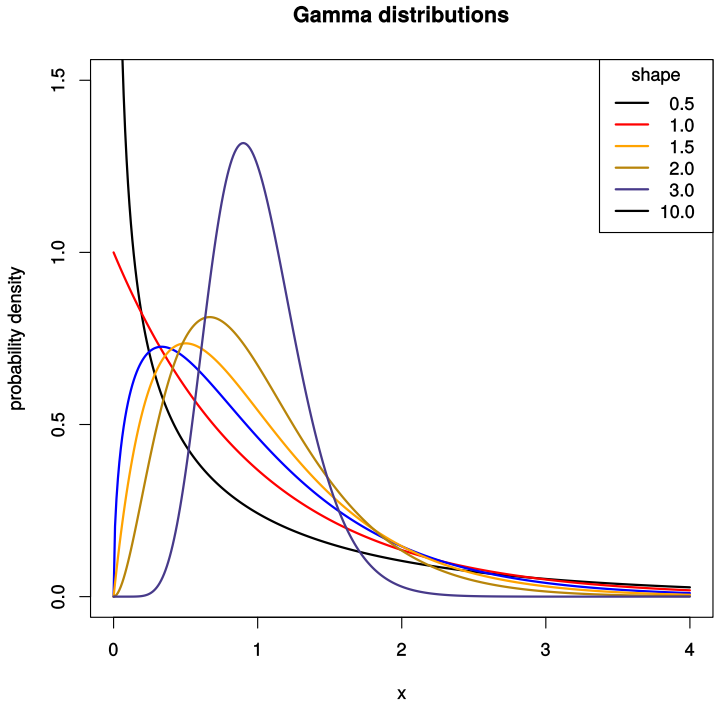
(Log-)Likelihood of a tree

- Where do we get the rates from?
- different options, increasing order of complexity
  1. $s = r = 0.5$ for all characters
  2. $r =$ empirical relative frequency of state 1 in the data (identical for all characters)
  3. a certain proportion p_{inv} (value to be estimated) of characters are *invariant*
  4. rates are *gamma distributed*

Gamma-distributed rates

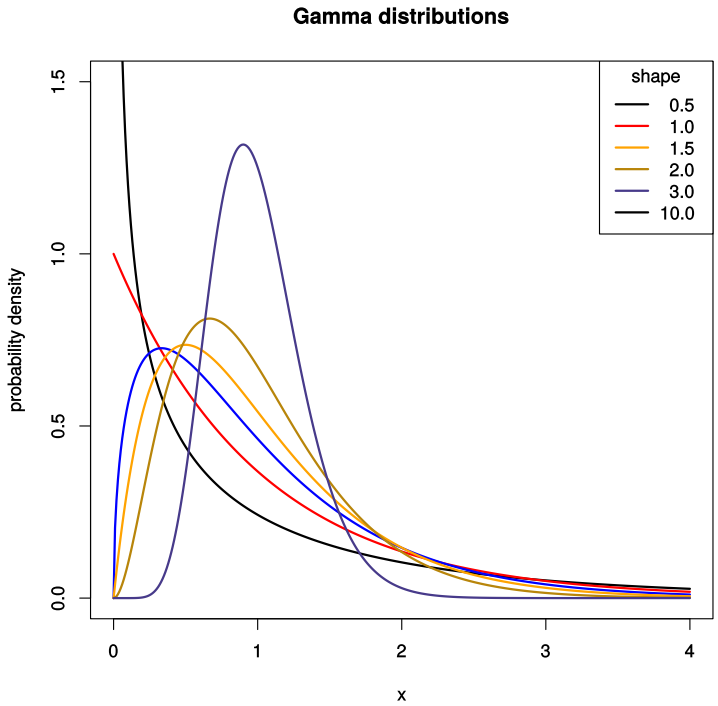
- we want allow rates to vary, but not too much
- common method (no real justification except for mathematical convenience)
  - equilibrium distribution is identical for all characters
  - rate matrix is multiplied with coefficient λ_i for character i
  - λ_i is random variable drawn from a *Gamma distribution*

$$\mathcal{L}(r_i = x) = \frac{\beta^{\beta} x^{(\beta-1)} e^{-\beta x}}{\Gamma(\beta)}$$



Gamma-distributed rates

- overall likelihood of tree topology: integrate over all λ_i, weighted by Gamma likelihood
- computationally impractical
- in practice: split Gamma distribution into n discrete bins (usually $n = 4$) and approximate integration via Hidden Markov Model



Modeling decisions to make

aspect of model	possible choices	number of parameters to estimate
branch lengths	unconstrained	$2n - 3$ (n is number of taxa)
	ultrametric	$n - 1$
equilibrium probabilities	uniform	0
	empirical	1
	ML estimate	1
rate variation	none	0
	Gamma distributed	1
invariant characters	none	0
	p_{inv}	1

This could be continued — you can build in rate variation across branches, you can fit the number of Gamma categories . . .

Model selection

- tradeoff
  - rich models are better at detecting patterns in the data, but are prone to over-fitting
  - parsimoneous models less vulnerable to overfitting but may miss important information
- standard issue in statistical inference
- one possible heuristics: **Akaike Information Criterion** (AIC)

$$\text{AIC} = -2 \times \log \text{likelihood} + 2 \times \text{number of free parameters}$$

- the model minimizing AIC is to be preferred

Example: Model selection for cognacy data/ UPGMA tree

model no.	branch lengths	eq. probs.	rate variation	inv. char.	AIC
1	ultrametric	uniform	none	none	17515.95
2	ultrametric	uniform	none	p_{inv}	17518.39
3	ultrametric	uniform	Gamma	none	17517.89
4	ultrametric	uniform	Gamma	p_{inv}	17519.75
5	ultrametric	empirical	none	none	16114.66
6	ultrametric	empirical	none	p_{inv}	16056.85
7	ultrametric	empirical	Gamma	none	15997.16
8	ultrametric	empirical	Gamma	p_{inv}	16022.21
9	ultrametric	ML	none	none	16034.96
10	ultrametric	ML	none	p_{inv}	16058.83
11	ultrametric	ML	Gamma	none	15981.94
12	ultrametric	ML	Gamma	p_{inv}	16009.90
13	unconstrained	uniform	none	none	17492.73
14	unconstrained	uniform	none	p_{inv}	17494.73
15	unconstrained	uniform	Gamma	none	17494.73
16	unconstrained	uniform	Gamma	p_{inv}	17496.73
17	unconstrained	empirical	none	none	16106.52
18	unconstrained	empirical	none	p_{inv}	16049.28
19	unconstrained	empirical	Gamma	none	16033.21
20	unconstrained	empirical	Gamma	p_{inv}	16011.38
21	unconstrained	ML	none	none	16102.04
22	unconstrained	ML	none	p_{inv}	16051.27
23	unconstrained	ML	Gamma	none	16025.99
24	unconstrained	ML	Gamma	p_{inv}	16001.00

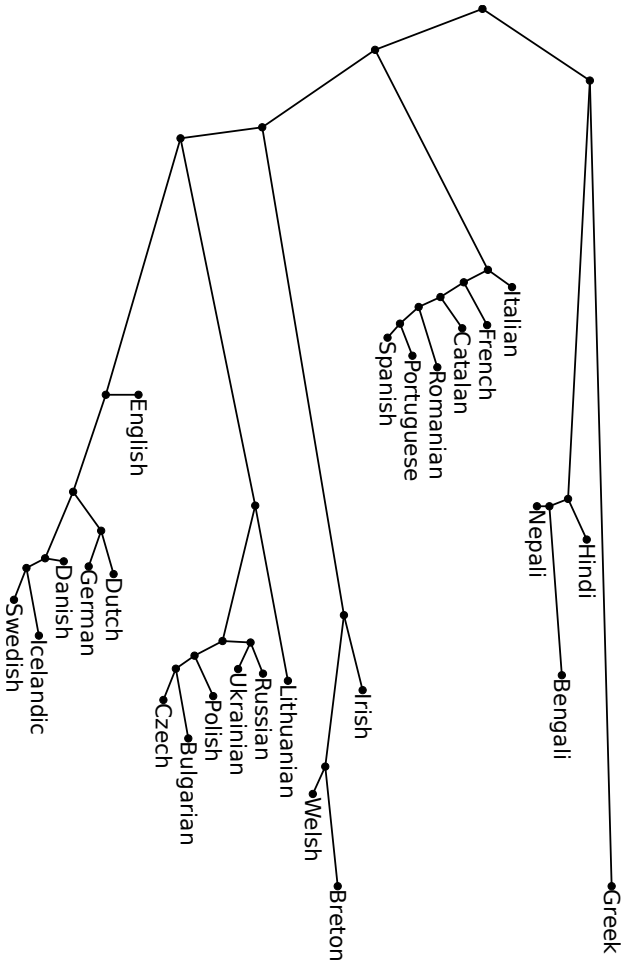
Tree search

- ML computation gives us likelihood of a tree topology, given data and a model
- ML tree:
  - heuristic search to find the topology maximizing likelihood
  - optimize branch lengths to maximize likelihood for that topology
- computationally very demanding!
- *for the 25 taxa in our running example, ML tree search for the full model requires several hours on a single processor; parallelization helps*
- ideally, one would want to do 24 heuristic tree searches, one for each model specification, and pick the tree+model with lowest AIC
- in practice one has to make compromises

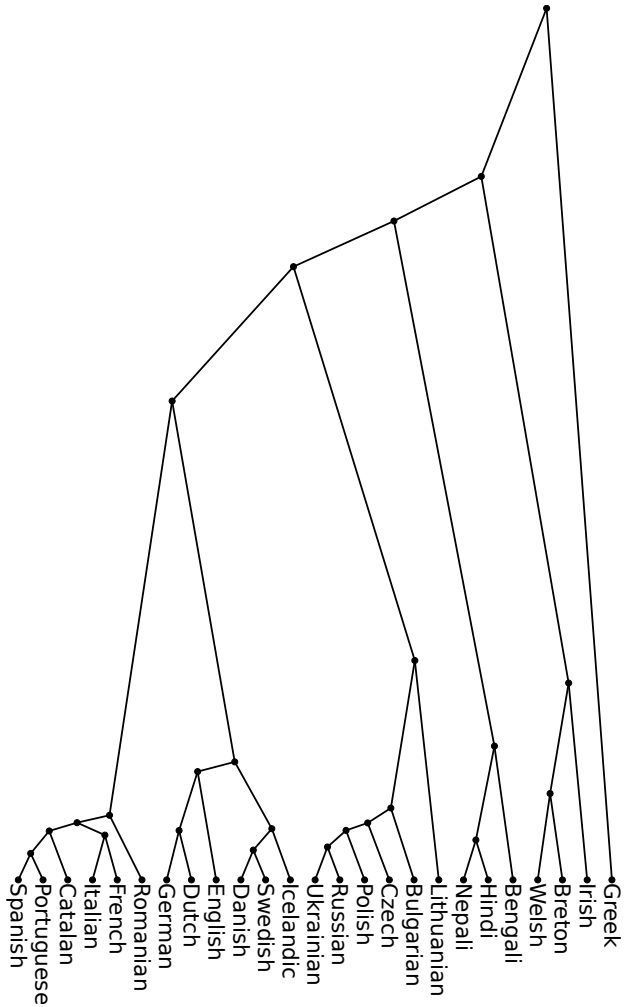
Running example

Running example: cognacy data

unconstrained branch lengths:
AIC = 7929

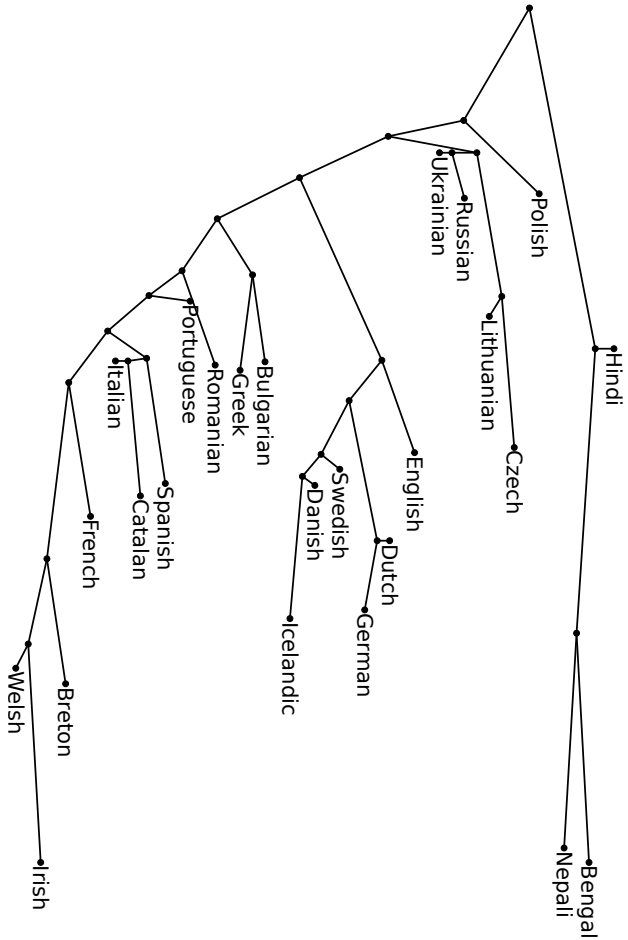


ultrametric:
AIC = 7972

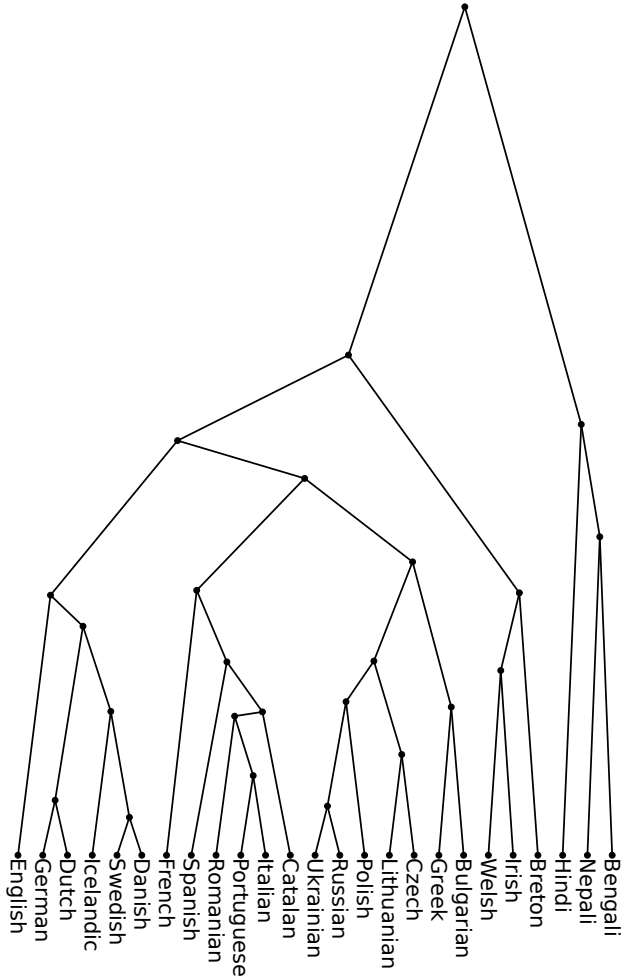


Running example: WALS data

unconstrained branch lengths:
AIC = 2752

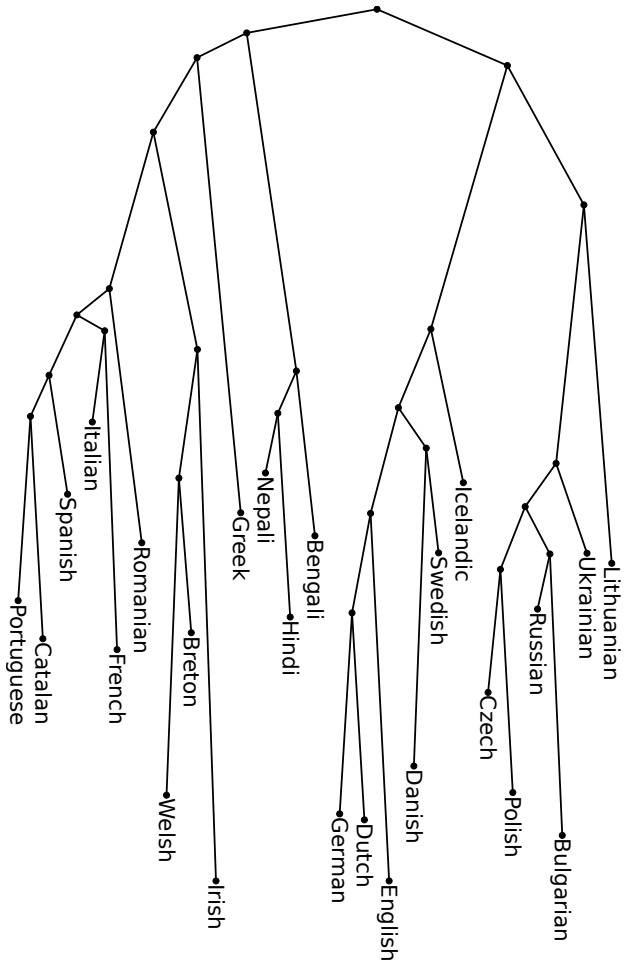


ultrametric:
AIC = 2828

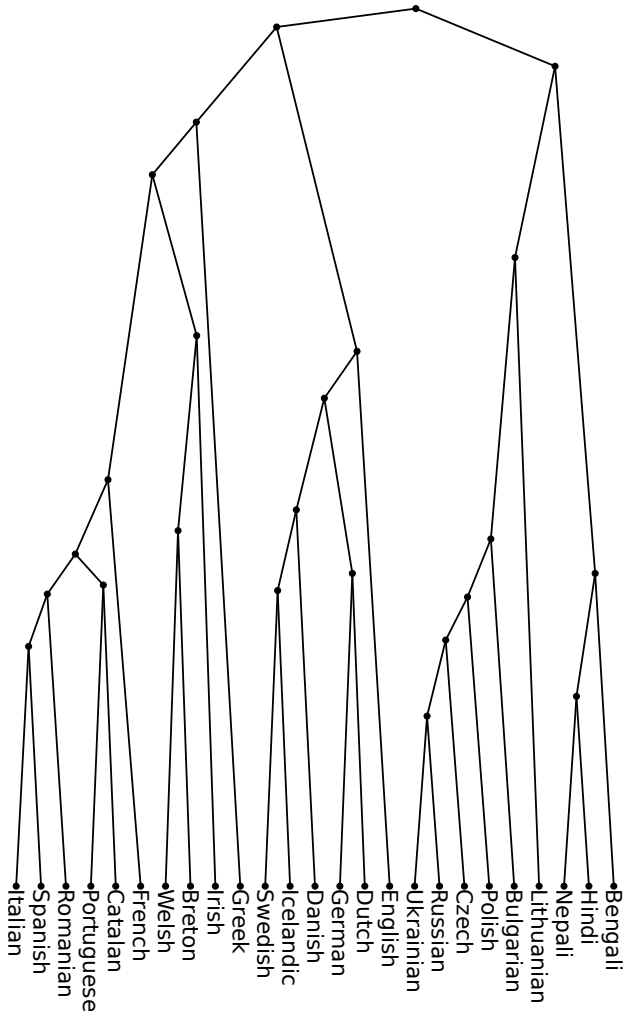


Running example: phonetic data

unconstrained branch lengths:
AIC = 89871



ultrametric:
AIC = 90575



Wrapping up

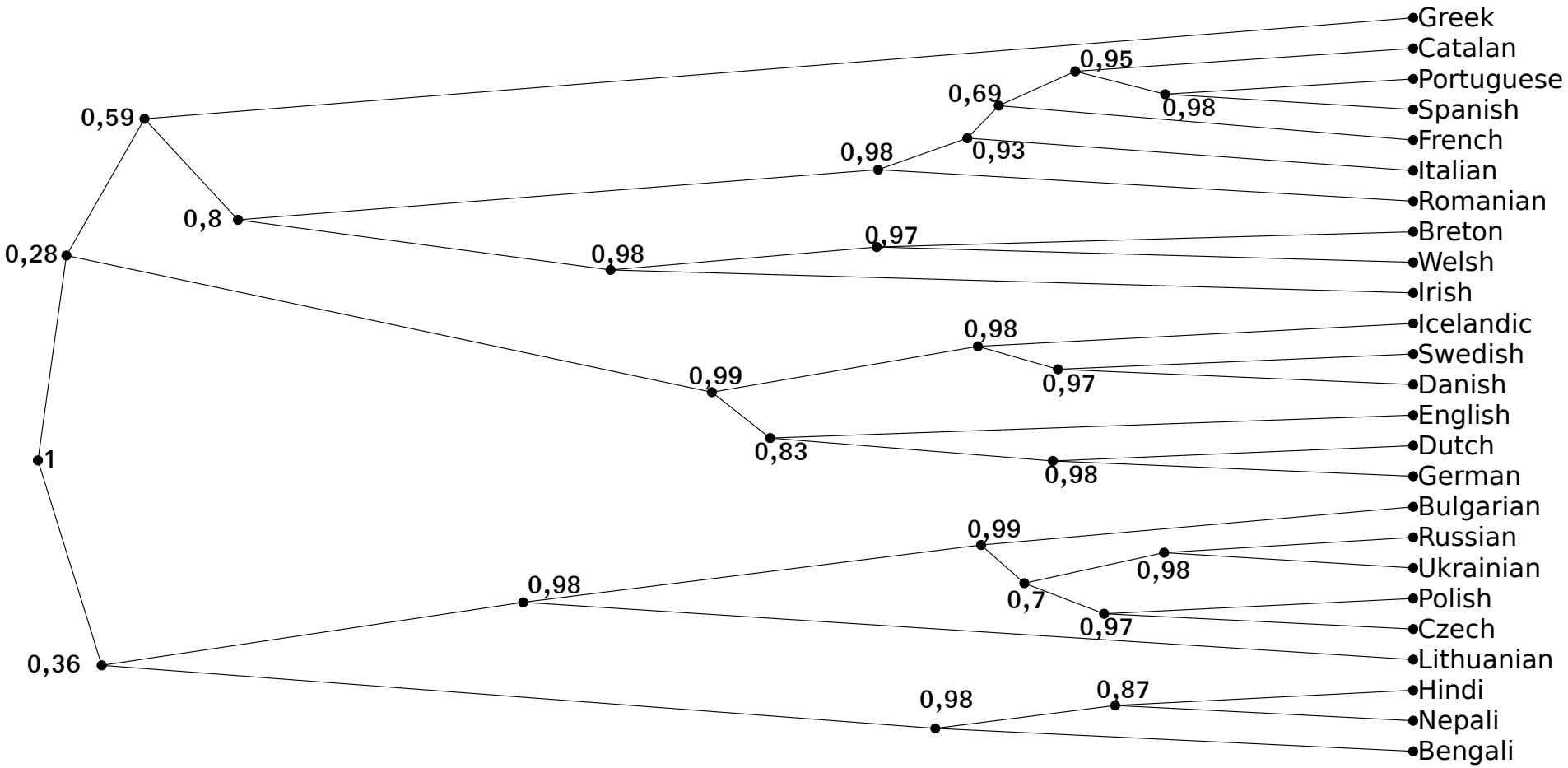
- ML is conceptually superior to MP (let alone distance methods)
  - different mutation rates for different characters are inferred from the data
  - possibility of multiple mutations are taken into account — depending on branch lengths
  - side effect of likelihood computation: probability distribution over character states at each internal node can be read off
- disadvantages:
  - computationally demanding
  - many parameter settings makes model selection difficult
    (note that the ultrametric trees in our example are sometimes better even though they have higher AIC)
  - ultrametric constraint makes branch lengths optimization computationally more expensive $\Rightarrow$ not feasible for larger data sets (more than 100–200 taxa)

Cleaning up from yesterday

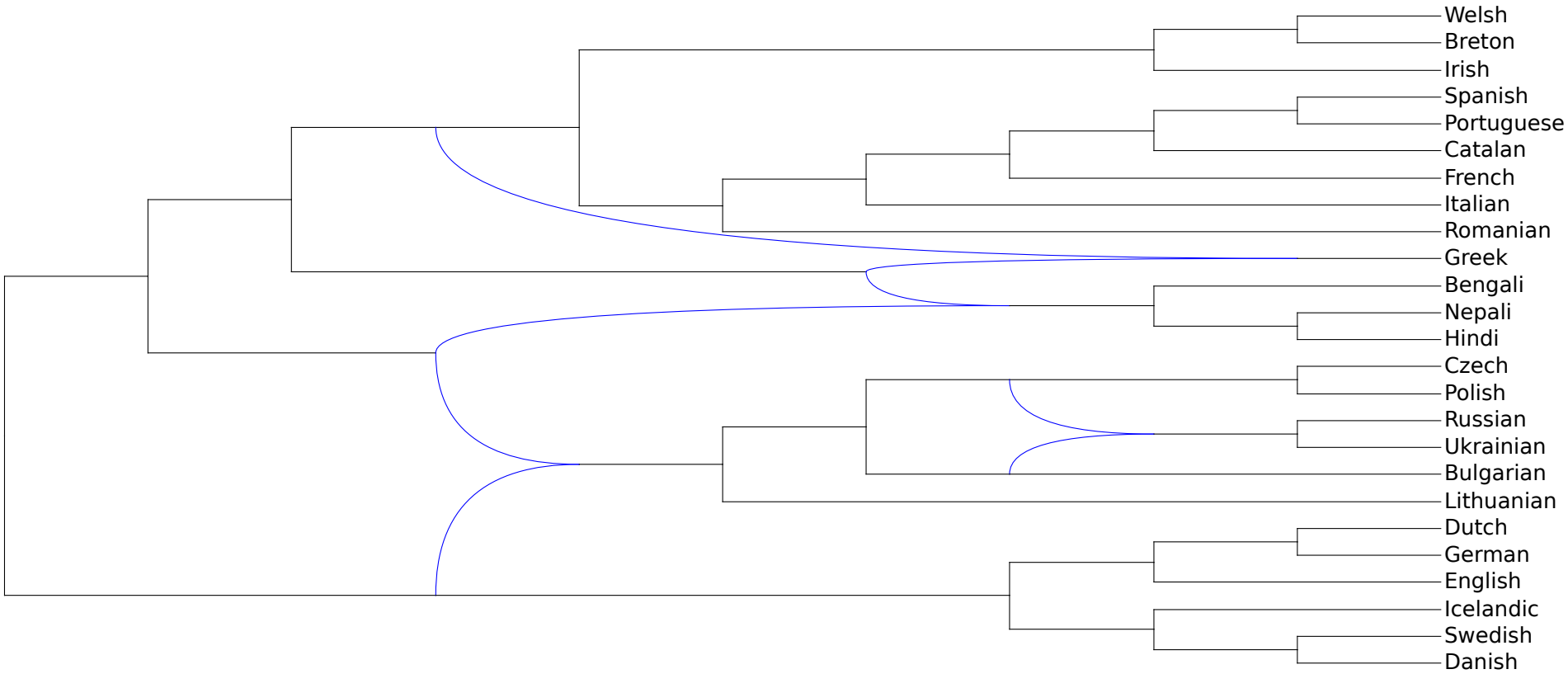
Using all data and the most sophisticated model...

- using both cognacy characters and phonetic characters
- Bayesian phylogenetic inference (related to Maximum Likelihood, but quite a bit more complex)
- 10 Gamma categories
- **relaxed** molecular clock $\Rightarrow$ rates are allowed to vary between branches, but only to a limited degree

Using all data and the most sophisticated model...



Using all data and the most sophisticated model...



Application: Ancestral State Reconstruction

joint work with Johann-Mattis List

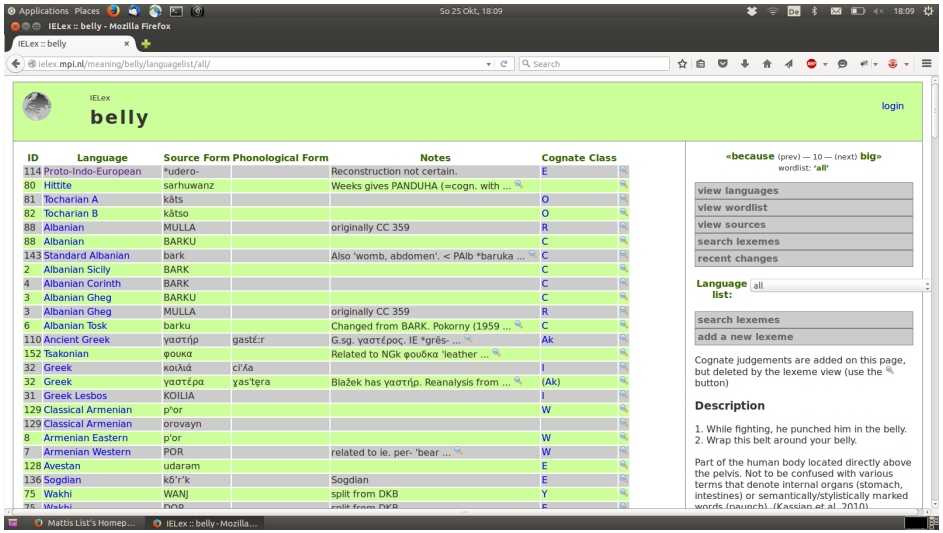
What is Ancestral State Reconstruction?

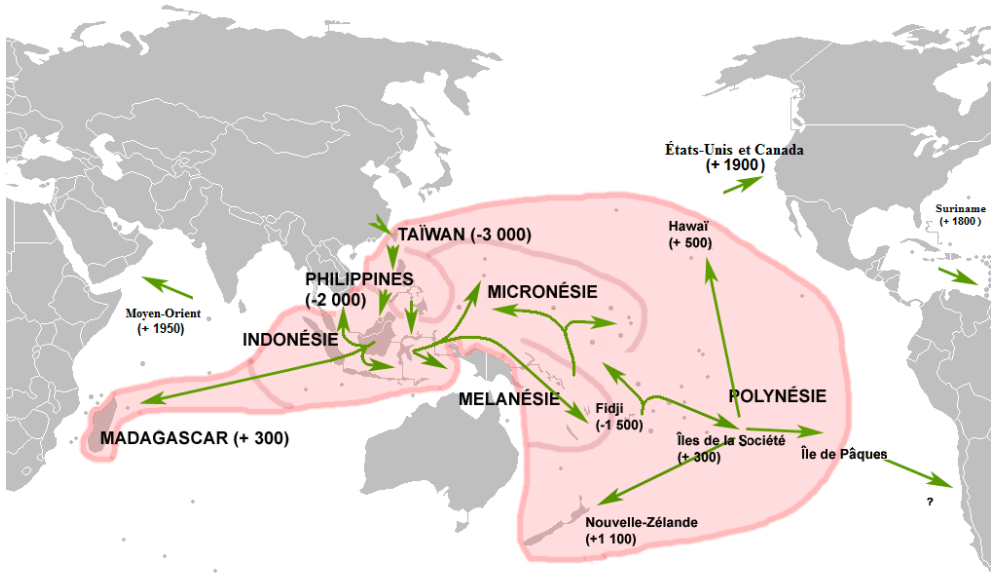
- While tree-building methods seek to find branching diagrams which explain how a language family has evolved, ASR methods use the branching diagrams in order to explain what has evolved concretely.
- Ancestral state reconstruction is very common in evolutionary biology but only spuriously practiced in computational historical linguistics (Bouchard-Côté et al., 2013)
- In classical historical linguistics, on the other hand, linguistic reconstruction of proto-forms and proto-meanings is very common and one of the main goals of the classical *comparative method* (Fox 1995).

ASR of Lexical Replacement Patterns

- If we look for words corresponding to one meaning in a wordlist and know which of the words are cognate or not, we may ask which of the word forms was the most likely candidate to be used in the proto-language of all descendant languages.
- This question resembles the task of "semantic reconstruction", but in contrast to classical semantic reconstruction, we are only operating within one concept slot here, disregarding all words with a different meaning which may also be cognate with the words in our sample.
- As a result of this restriction, it is quite likely that we cannot recover the original form from our data.
- It is, however, very interesting to see to which degree we *can* propose a good candidate word form (cognate set) for the proto-language.

Data


IELex :: belly - Mozilla Firefox
IELex :: belly
ielex.mpi.nl/meaning/belly/languagelist/all/
IELex
belly
login
ID Language Source Form Phonological Form Notes Cognate Class
114 Proto-Indo-European *udero- Reconstruction not certain. E
80 Hittite sarhuwanz Weeks gives PANDUHA (=cogn. with ...)
81 Tocharian A kāts O
82 Tocharian B kātso O
88 Albanian MULLA originally CC 359 R
88 Albanian BARKU C
143 Standard Albanian bark Also 'womb, abdomen'. < PAlb *baruka ... C
2 Albanian Sicily BARK C
4 Albanian Corinth BARK C
3 Albanian Gheg BARKU C
3 Albanian Gheg MULLA originally CC 359 R
6 Albanian Tosk barku Changed from BARK. Pokorny (1959 ... C
110 Ancient Greek γαστήρ gastɛ́:r G.sg. γαστέρος, IE *gʰēs- ... Ak
152 Tsakonian φουκα Related to NGk φουδκα 'leather ... C
32 Greek κοιλιά cí'ʎa I
32 Greek γαστέρα ɣas'tɛ́ra Blažek has γαστήρ. Reanalysis from ... (Ak)
31 Greek Lesbos KOILIA I
129 Classical Armenian pʰor W
129 Classical Armenian orovayn
8 Armenian Eastern pʻor W
7 Armenian Western POR related to ie. per- 'bear ... W
128 Avestan udarəm E
136 Sogdian kδ'r'k Sogdian E
75 Wakhi WANJ split from DKB Y
«because (prev) — 10 — (next) big»
wordlist: 'all'
view languages
view wordlist
view sources
search lexemes
recent changes
Language list: all
search lexemes
add a new lexeme
Cognate judgements are added on this page, but deleted by the lexeme view (use the button)
Description
1. While fighting, he punched him in the belly.
2. Wrap this belt around your belly.
Part of the human body located directly above the pelvis. Not to be confused with various terms that denote internal organs (stomach, intestines) or semantically/stylistically marked


États-Unis et Canada
(+ 1900)
Suriname
(+ 1800)
Hawaï
(+ 500)
TAÏWAN (-3 000)
PHILIPPINES
(-2 000)
MICRONÉSIE
Moyen-Orient
(+ 1950)
INDONÉSIE
MELANÉSIE
Fidji
(-1 500)
POLYNÉSIE
MADAGASCAR (+ 300)
Îles de la Société
(+ 300)
Île de Pâques
?
Nouvelle-Zélande
(+1 100)

Data

IELex

- 153 Indo-European doculects
- 207 concepts
- entries for Proto-Indo-European for 135 concepts → used as gold standard
- arbitrarily split into training set and test set:
  - training set: 67 concepts, 1127 cognate classes (83 occur in PIE)
  - test set: 68 concepts, 957 cognate classes (79 from PIE)

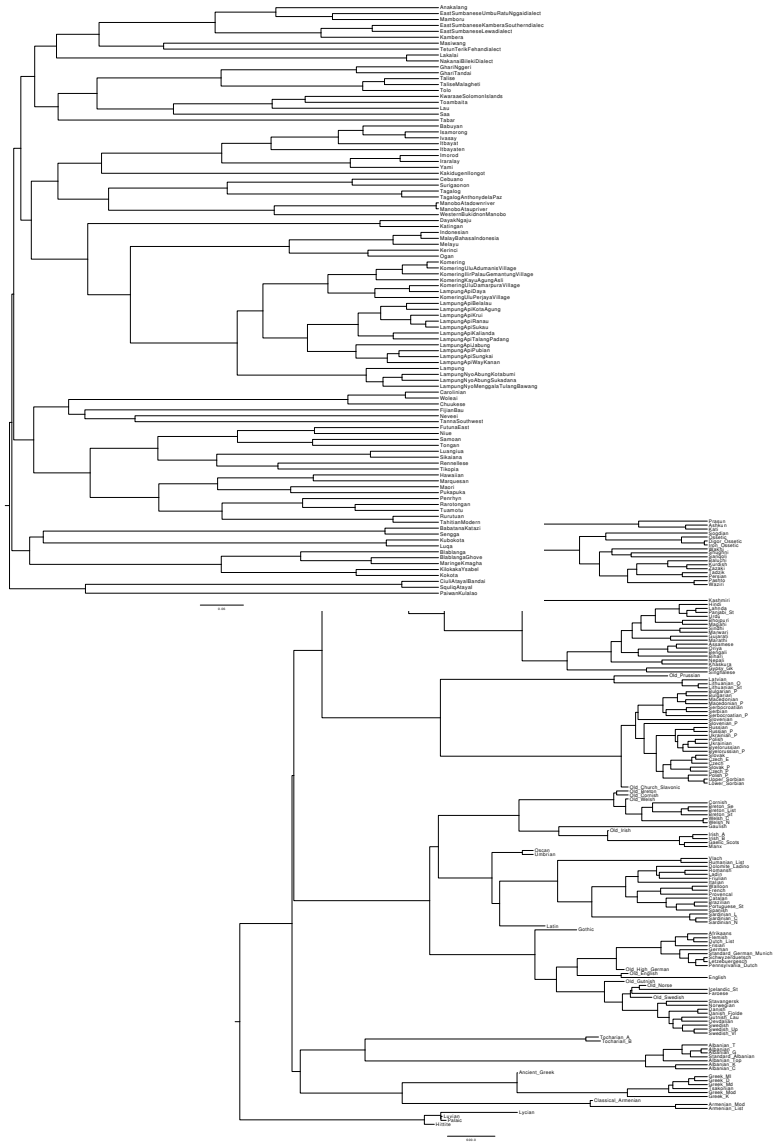
ABVD

- 743 Austronesian doculects → 100 were selected at random
- 210 concepts; for 154 of them entries for Proto-Austronesian
- split into training set and test set:
  - training set: 81 concepts, 1695 cognate classes (88 occur in PAn)
  - test set: 74 concepts, 1584 cognate classes (79 occur in PAn)

Prerequisites: Trees

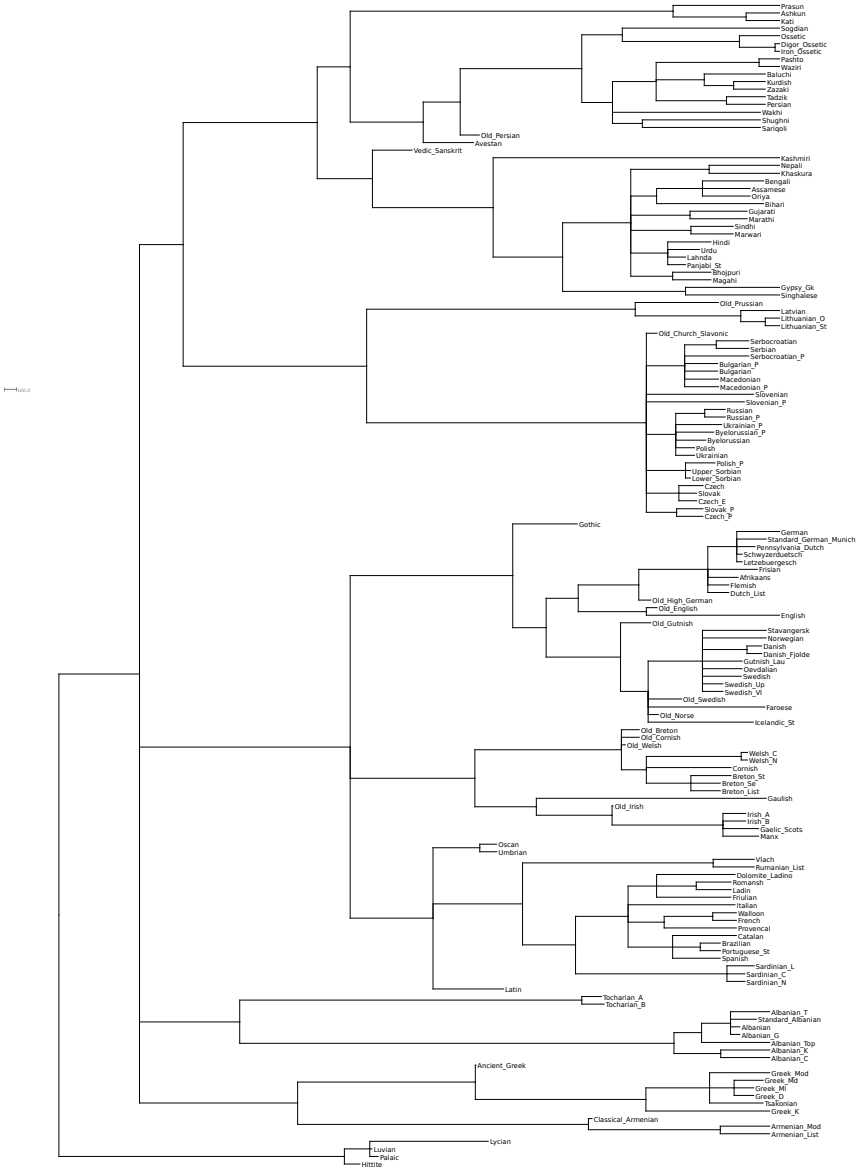
Trees

- trees were inferred with full data set (training + test data) via Bayesian inference
  - IELex outgroup: Anatolian
  - ABVD outgroup: Malayo-Polynesian
- random samples of 1000 trees from posterior distributions
- maximum clade credibility trees



Phylogenetic uncertainty

- proper way to deal with it: work with posterior sample rather than with a single tree
- poor man's method:
  - remove all short branches (shorter than some threshold)
  - do ASR with resulting multifurcating tree



Summary on Indo-European ASR

Error Type	GS	ASR	Number
Missing forms	A	∅	7
Different forms	A	B	9
Additional forms in ASR	A	A, B	5
Missing root in ASR	A, B	A	4
Summary			25

Evaluating the Differences

We evaluate the differences qualitatively by checking

- the reflection of the proposed root in the branches, especially with semantically shifted word forms,
- the likelihood of semantic shift of the given root with help of the Database of Cross-Linguistic Colexifications (CLICS, List et al. 2013 and 2014),
- thoroughly whether the cognate sets in the data are really reflexes of the proposed PIE root.

Based on this check, we distinguish four grades of root quality:

erroneous	problematic	possible	good

Indo-European ASR: Missing forms

Concept	Form	Meaning in Reflexes	Comment
SEE	*derḱ-	to see	Only reflected in Indo-Iranian, cognates also problematic.
SEE	*weid-	to see or to know	Safe root for Indo-European.
SING	*kan-	to sing or the rooster	Root is proposed for PIE on the basis of Germanic reflexes meaning "rooster" which is a highly unlikely semantic change
SMELL	*$h_3$ed-	to smell	Potential root for PIE, but only reflected in Greek and Romance
SMALL	*mei-	small	Wrong cognate judgments in the database, since neither Russian *malenkij* nor English *small* go back to this root
THINK	*teng-	to think or to feel	Root only reflected in Germanic languages with spurious reflexes in semantically shifted form in other branches. A better candidate for PIE would be *men- "the mind or to think".
WASH	*leh_2w-	to wash or to pour	Wrong cognate assignment in the source since Romance and Albanian reflexes are not annotated.
WASH	*$neig^w$-	to wash or water monster	Very unlikely cognate assignment, due to the extreme shift from "to wash" to "water monster" (cf. English *nix*) in the Germanic languages.
WET	*wed-	water or wet	Semantic change from "water" to "wet" is likely according to CLICS, but it is not clear why this should have already happened in PIE times.

erroneous | problematic | possible | good

Indo-European ASR: Missing forms

Concept	Form	Meaning in Reflexes	Comment
SEE	*derḱ-	to see	Only reflected in Indo-Iranian, cognates also problematic.
SEE	*weid-	to see or to know	Safe root for Indo-European.
SING	*kan-	to sing or the rooster	Root is proposed for PIE on the basis of Germanic reflexes meaning "rooster" which is a highly unlikely semantic change
SMELL	*$h_3$ed-	to smell	Potential root for PIE, but only reflected in Greek and Romance
SMALL	*mei-	small	Wrong cognate judgments in the database, since neither Russian *malenkij* nor English *small* go back to this root
THINK	*teng-	to think or to feel	Root only reflected in Germanic languages with spurious reflexes in semantically shifted form in other branches. A better candidate for PIE would be *men- "the mind or to think".
WASH	*leh_2w-	to wash or to pour	Wrong cognate assignment in the source since Romance and Albanian reflexes are not annotated.
WASH	*$neig^w$-	to wash or water monster	Very unlikely PIE root, due to the extreme shift from "to wash" to "water monster" (cf. English *nix*) in the Germanic languages.
WET	*wed-	water or wet	Semantic change from "water" to "wet" is likely according to CLICS, but it is not clear why this should have already happened in PIE times.

erroneous | problematic | possible | good

Indo-European ASR: Different Forms

Concept	GS	ASR	Comment
RIVER	*$h_2ek^weh_2$	*h_2ep-	Form in GS meant "water" in PIE. Although a shift from "water" to "river" is likely according to CLICS, this meaning is an innovation in Germanic. The ASR form is reflected across multiple branches and a much better candidate.
RUB	*$melh_1$-	*$terh_1$-	Form in GS is not reflected in the standard literature (LIV and LIN), form in ASR is reflected in the meaning "to rub, to bore".
SCRATCH	*$gerb^h$-	*kes-	Form in GS is only reflected in few Germanic languages, probably with a wrong cognate assignment. Following Derksen (2008), assuming the GSR form is a much better candidate for the PIE word for "scratch".
SKIN	*pel	*(s)kewH-	Form in GS is a good PIE root, but not necessarily with the meaning "skin", as the meaning of the reflexes differs greatly. The GSR form derives from a PIE verb meaning "to cover", but the cognate should not contain Slavic words (Derksen 2008).
WALK	*$ǵ^heh_1$	*h_1ei-	The GS form is only reflected in Germanic. The ASR form is a clear PIE root, but the meaning may also have been "to go".
WATER	*$h_2ek^weh_2$	*wódr̥	The ASR form is a much better candidate for "water" in PIE, due to its high number of reflexes in all branches.
WHITE	*$h_2elb^hós$	*$h_2erǵó$-	The GS form is only reflected in Romance in this meaning and as meaning "cloud" in Hittite. The ASR form is a much better candidate, with a much more plausible connection between reflexes meaning "shine" and "white", as also confirmed by CLICS.
WORM	*wr̥mi-	*$k^wr̥mis$	The ASR form is reflected in more different branches of PIE, while the GS form is only reflected in Germanic and Romance.

erroneous | problematic | possible | good

Indo-European ASR: Different Forms

Concept	GS	ASR	Comment
RIVER	*$h_2ek^weh_2$	*h_2ep-	Form in GS meant "water" in PIE. Although a shift from "water" to "river" is likely according to CLICS, this meaning is an innovation in Germanic. The ASR form is reflected across multiple branches and a much better candidate.
RUB	*$melh_1$-	*$terh_1$-	Form in GS is not reflected in the standard literature (LIV and LIN), form in ASR is reflected in the meaning "to rub, to bore".
SCRATCH	*$gerb^h$-	*kes-	Form in GS is only reflected in few Germanic languages, probably with a wrong cognate assignment. Following Derksen (2008), assuming the GSR form is a much better candidate for the PIE word for "scratch".
SKIN	*pel	*(s)kewH-	Form in GS is a good PIE root, but not necessarily with the meaning "skin", as the meaning of the reflexes differs greatly. The GSR form derives from a PIE verb meaning "to cover", but the cognate should not contain Slavic words (Derksen 2008).
WALK	*$ǵ^heh_1$	*h_1ei-	The GS form is only reflected in Germanic. The ASR form is a clear PIE root, but the meaning may also have been "to go".
WATER	*$h_2ek^weh_2$	*wódr̥	The ASR form is a much better candidate for "water" in PIE, due to its high number of reflexes in all branches.
WHITE	*$h_2elb^hós$	*$h_2erǵó$-	The GS form is only reflected in Romance in this meaning and as meaning "cloud" in Hittite. The ASR form is a much better candidate, with a much more plausible connection between reflexes meaning "shine" and "white", as also confirmed by CLICS.
WORM	*wr̥mi-	*$k^wr̥mis$	The ASR form is reflected in more different branches of PIE, while the GS form is only reflected in Germanic and Romance.

erroneous | problematic | possible | good

Indo-European ASR: Additional Forms

Concept	Form in ASR	Comment
MOON	*lewk-s-nh$_2$	This form would go back to a PIE root meaning "to shine" and is often said to have independently turned to mean "moon" in Romance and Slavic and other branches. The shift from "shine" to "moon" is however not very likely (no evidence in CLICS), so it is also possible that the word meant already "moon" in PIE as an epithet (Vaan 2008).
SNOW	*ǵ^{h}éi-mn̥-	The form has probably independently shifted from the original meaning "frost, cold", which is a very likely shift according to CLICS.
SUCK	*suḱ-	The root is present in this meaning in many subbranches and a good candidate for PIE in this meaning.
THIS	*so / *to	The root is a clear PIE demonstrative (Meier-Brgger 2010), but the reflexes in the daughter languages vary greatly, due to analogical levelling.
WITH	*sm̥	A very good candidate for the meaning with reflexes in Greek, Indo-Iranian and Slavic.

erroneous	problematic	possible	good

Indo-European ASR: Additional Forms

Concept	Form in ASR	Comment
MOON	*lewk-s-nh$_2$	This form would go back to a PIE root meaning "to shine" and is often said to have independently turned to mean "moon" in Romance and Slavic and other branches. The shift from "shine" to "moon" is however not very likely (no evidence in CLICS), so it is also possible that the word meant already "moon" in PIE as an epithet (Vaan 2008).
SNOW	*ǵ^{h}éi-mn̥-	The form has probably independently shifted from the original meaning "frost, cold", which is a very likely shift according to CLICS.
SUCK	*sukʷ-	The root is present in this meaning in many subbranches and a good candidate for PIE in this meaning.
THIS	*so / *to	The root is a clear PIE demonstrative (Meier-Brügger 2010), but the reflexes in the daughter languages vary greatly, due to analogical levelling.
WITH	*sm̥	A very good candidate for the meaning with reflexes in Greek, Indo-Iranian and Slavic.

erroneous	problematic	possible	good

Indo-European ASR: Missing Forms in ASR

Concept	Form in GS	Comment
NOT	*meh_1	This form is reflected in Old Greek as a prohibitive negation and also reconstructed as such. Whether it was the normal negation in PIE is less clear.
SLEEP	*drem	This form is mainly reflected in Latin and spuriously in Indian and Greek. It is much more likely that it meant something else in PIE and then shifted into this meaning.
VOMIT	*$h_1$rewg-	No need to reconstruct this form back to PIE, since it is only reflected in two languages of Romance.
YEAR	*ieHr-	This form has only reflexes in Germanic languages. Generally, the meaning "year" is difficult to reconstruct, due to the high potential for shift from "summer", "winter", "time", etc. as shown in CLICS.

erroneous	problematic	possible	good

Indo-European ASR: Missing Forms in ASR

Concept	Form in GS	Comment
NOT	*meh_1	This form is reflected in Old Greek as a prohibitive negation and also reconstructed as such. Whether it was the normal negation in PIE is less clear.
SLEEP	*drem	This form is mainly reflected in Latin and spuriously in Indian and Greek. It is much more likely that it meant something else in PIE and then shifted into this meaning.
VOMIT	*$h_1$rewg-	No need to reconstruct this form back to PIE, since it is only reflected in two languages of Romance.
YEAR	*ieHr-	This form has only reflexes in Germanic languages. Generally, the meaning "year" is difficult to reconstruct, due to the high potential for shift from "summer", "winter", "time", etc. as shown in CLICS.

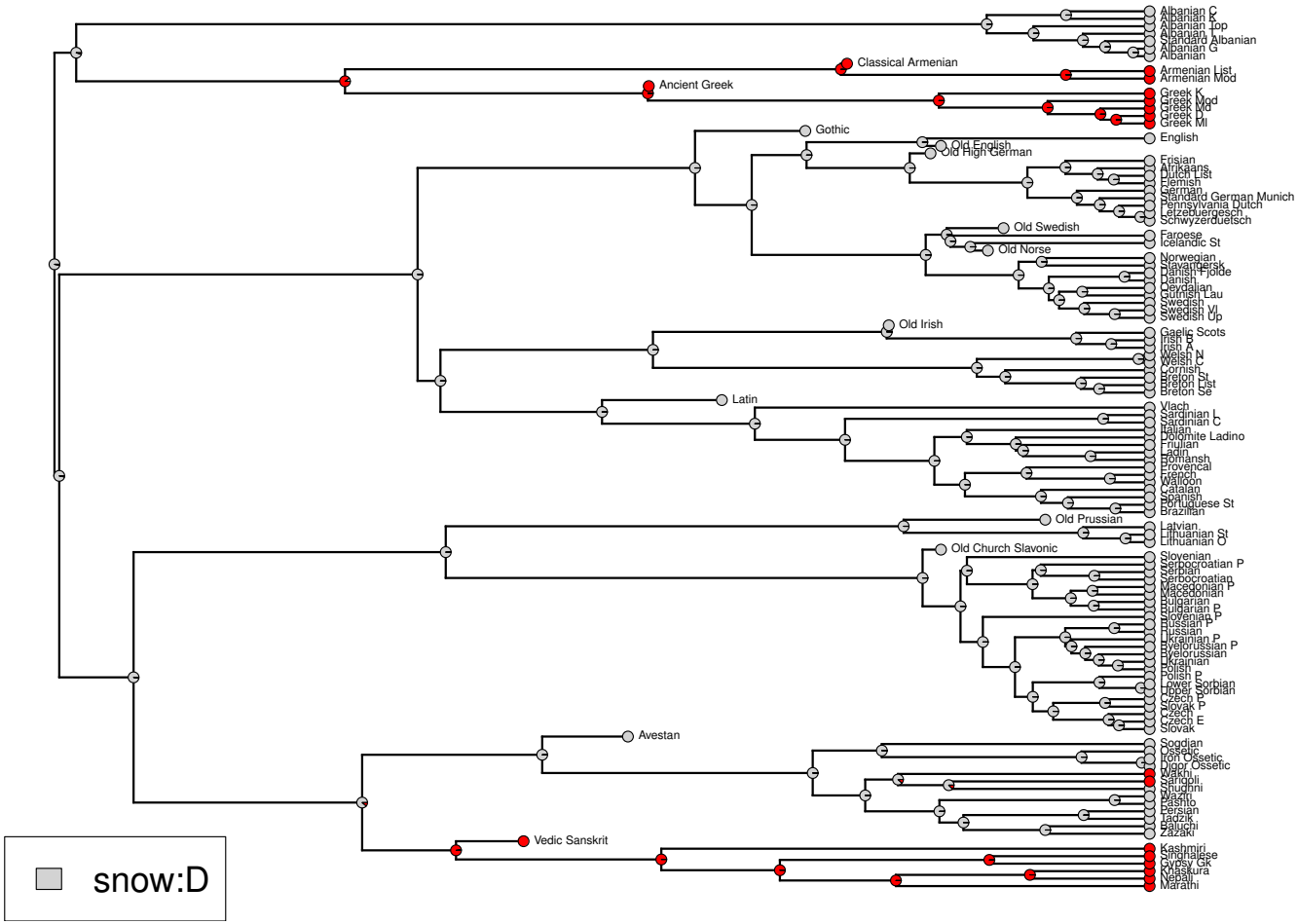
erroneous	problematic	possible	good

Evaluation against our manually created gold standard

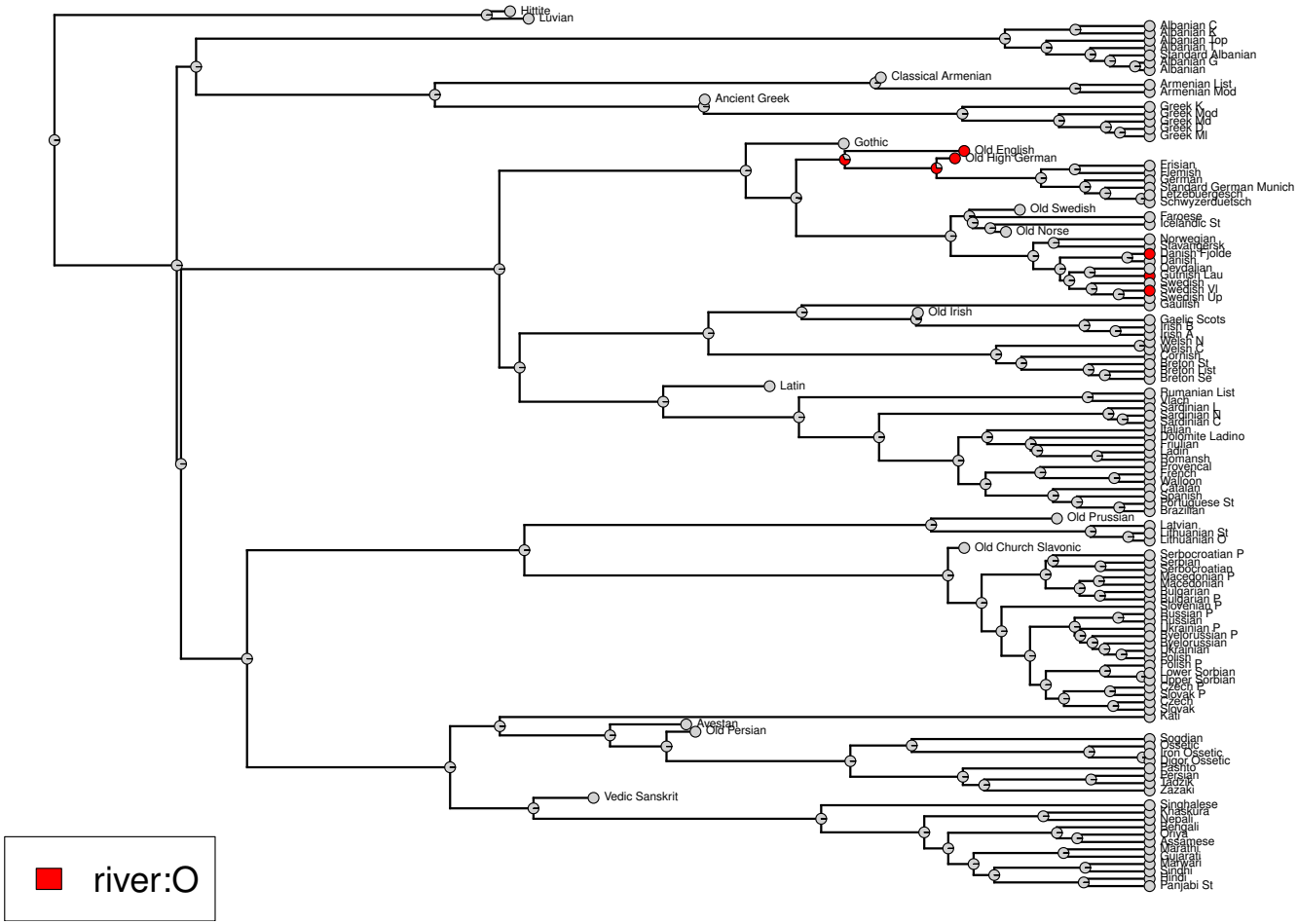
- precision: 0.986 (1 false positive)
- recall: 0.895 (8 false negatives)
- F-score: 0.938[1]

[1] The IELex PIE entries have an F-score of 0.854.

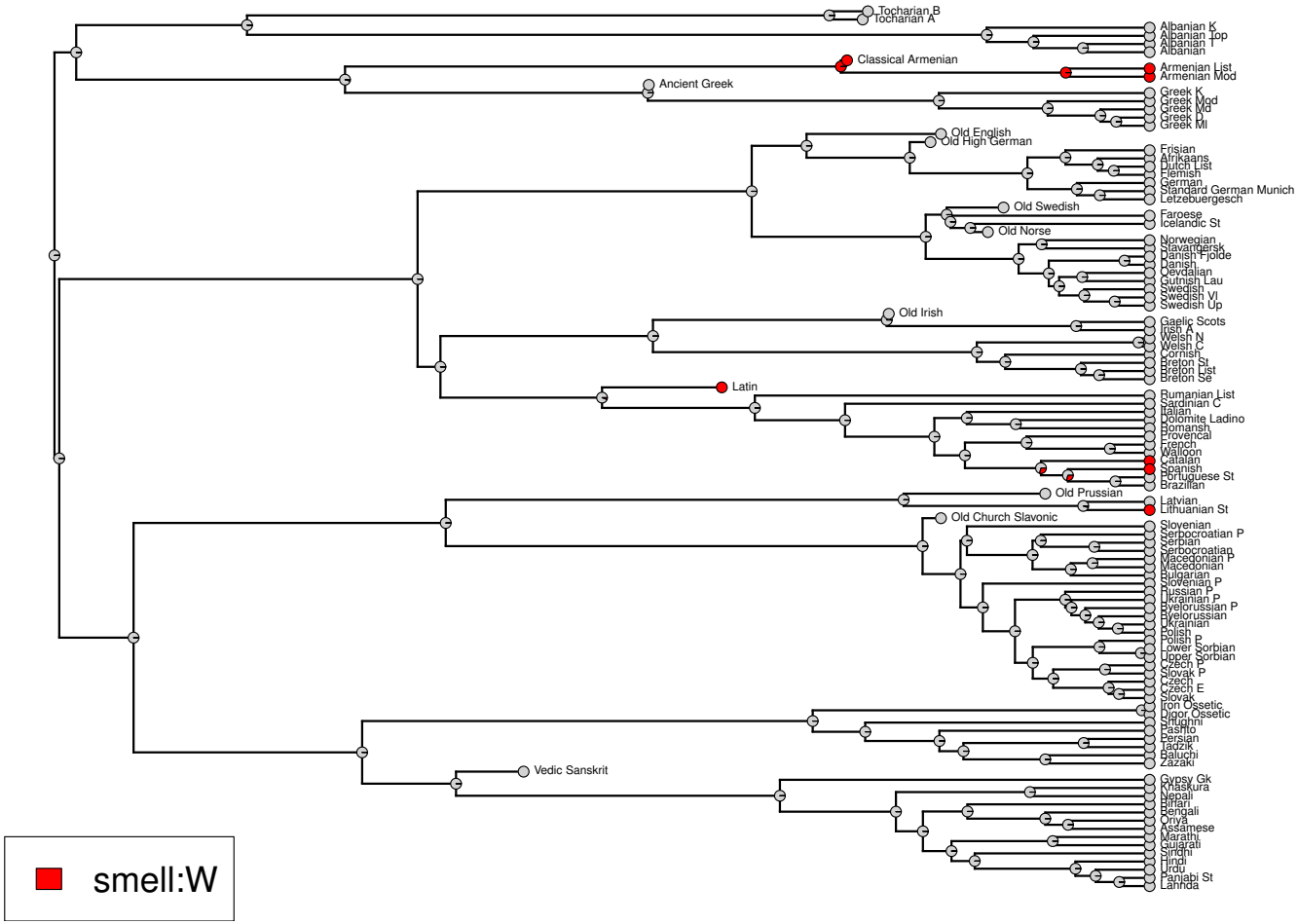
False positive



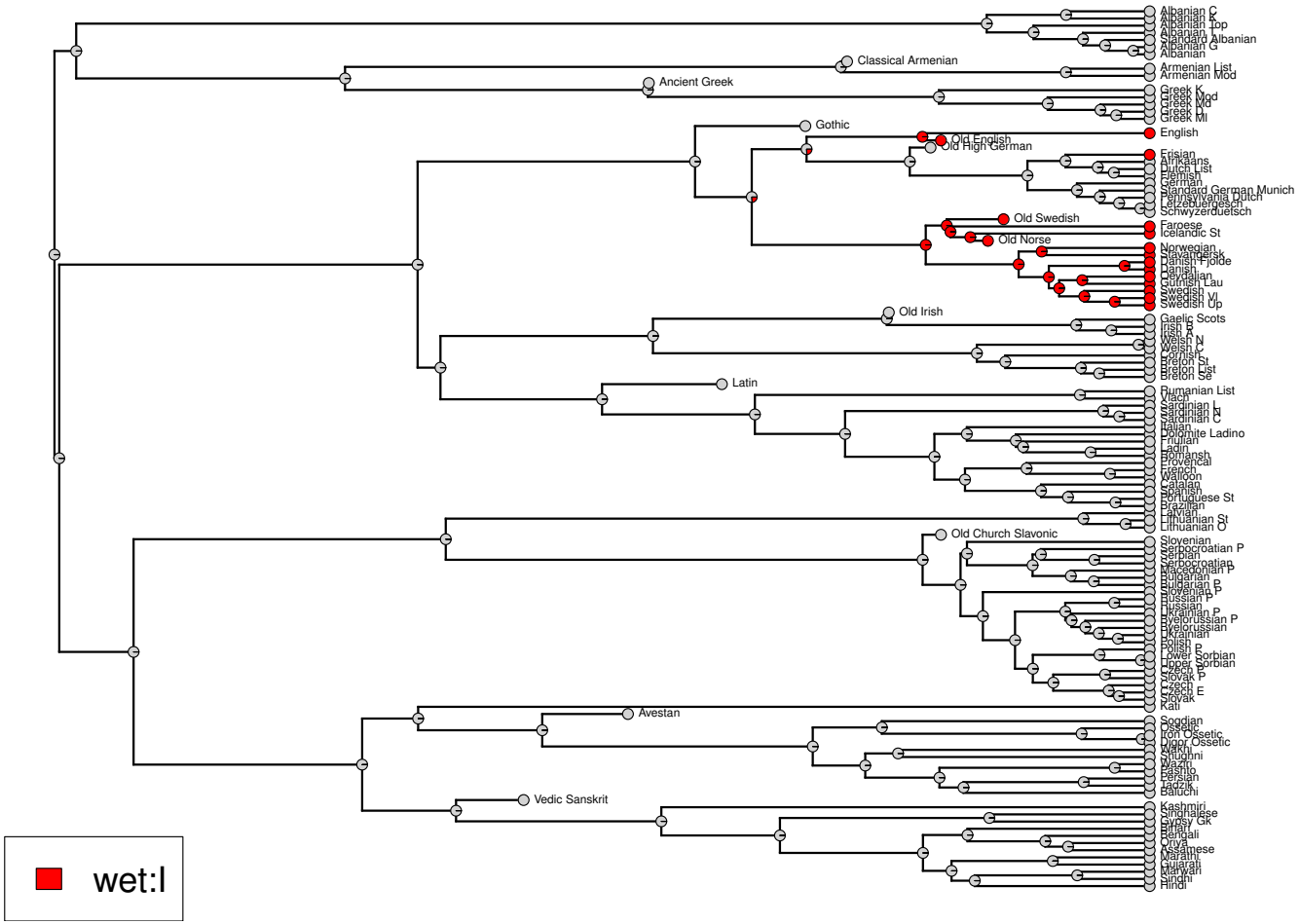
False negatives



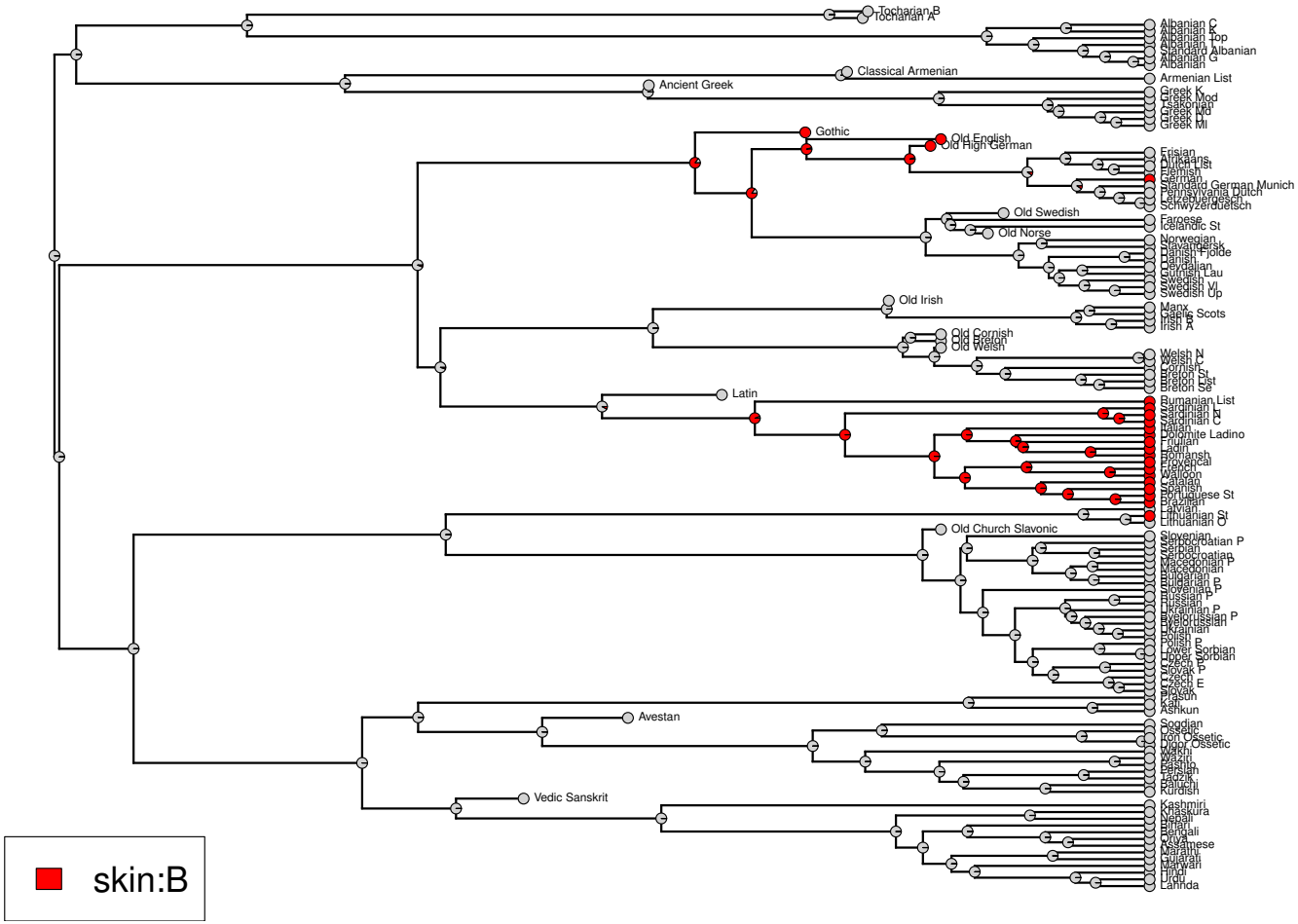
False negatives



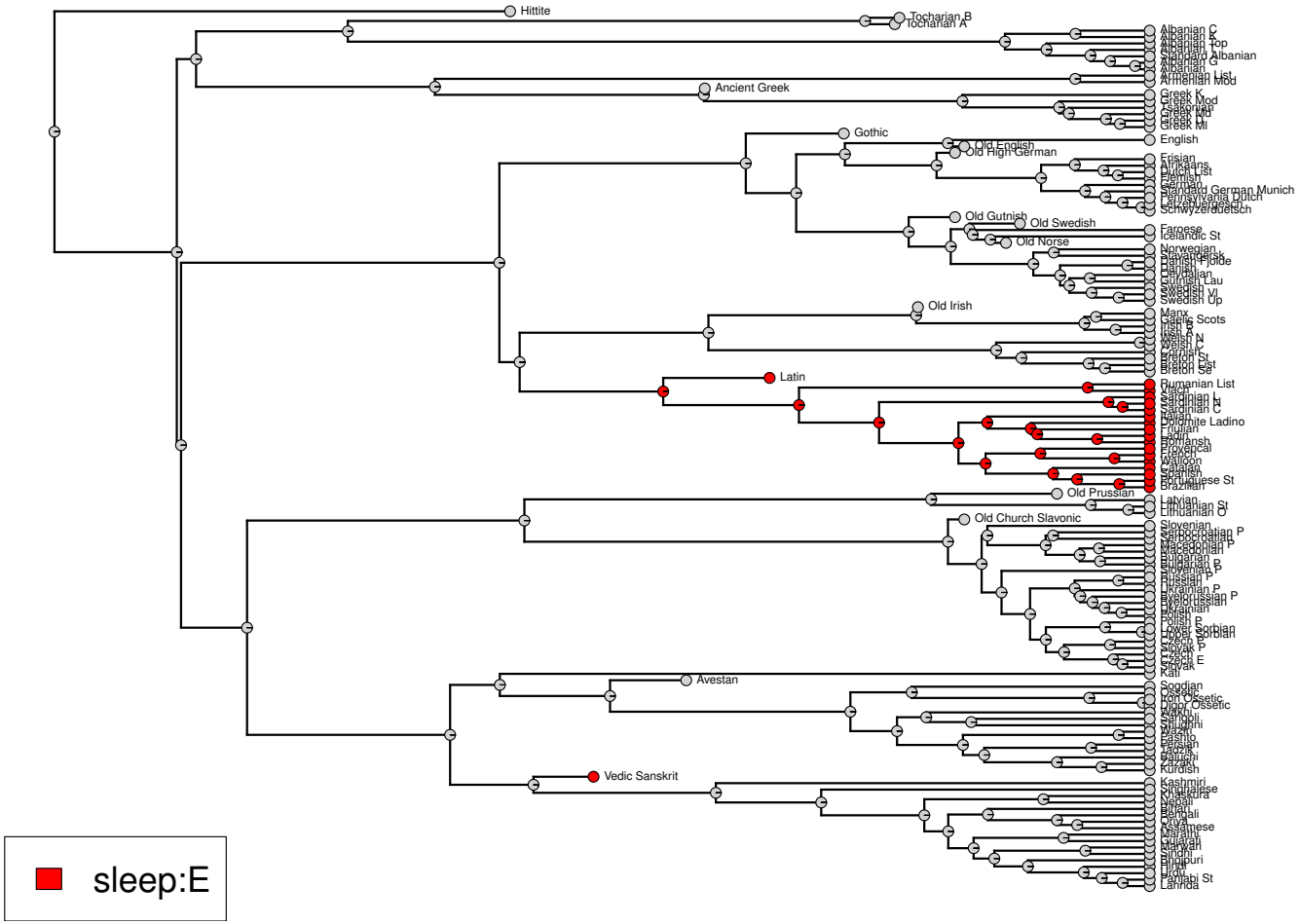
False negatives



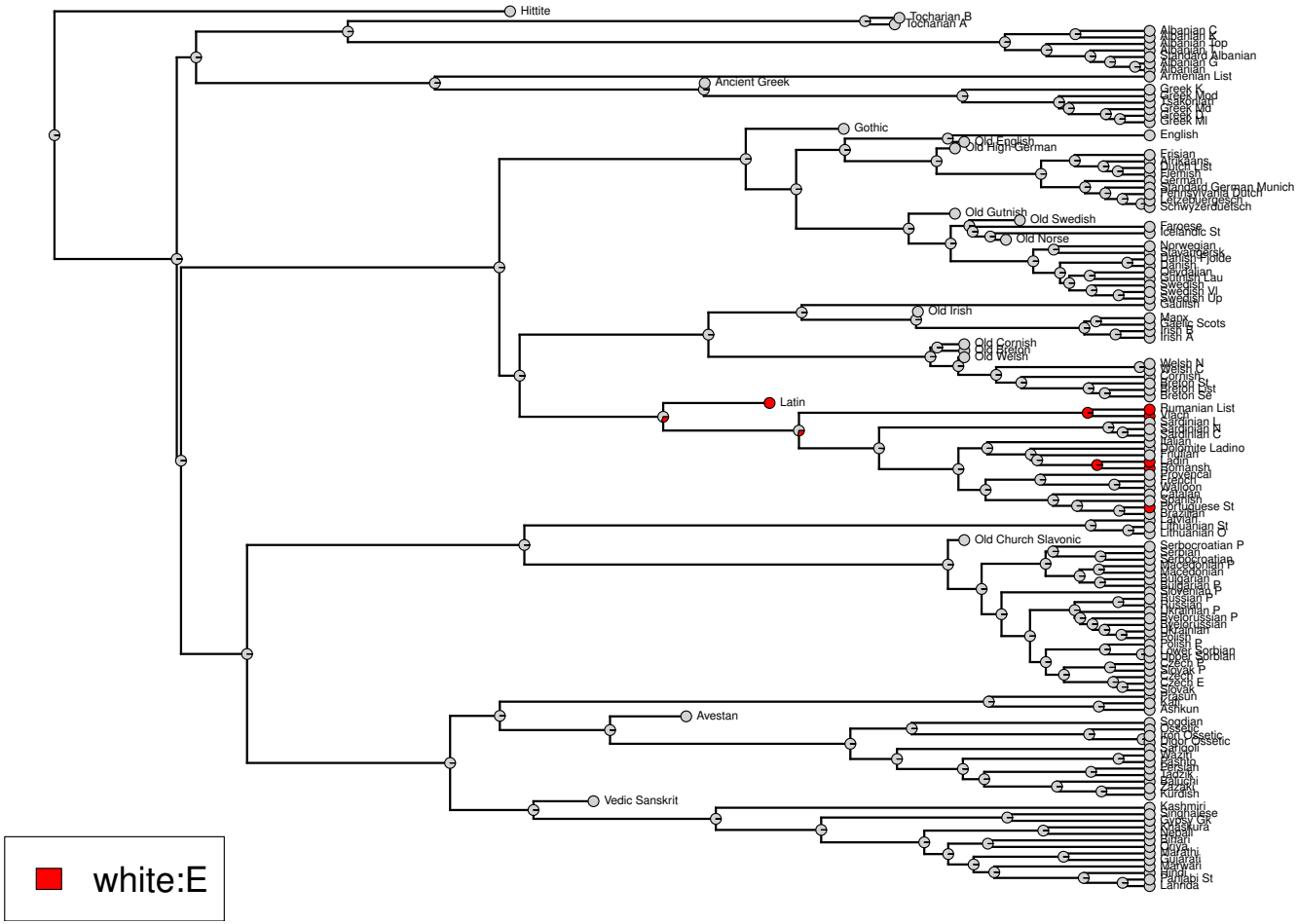
False negatives



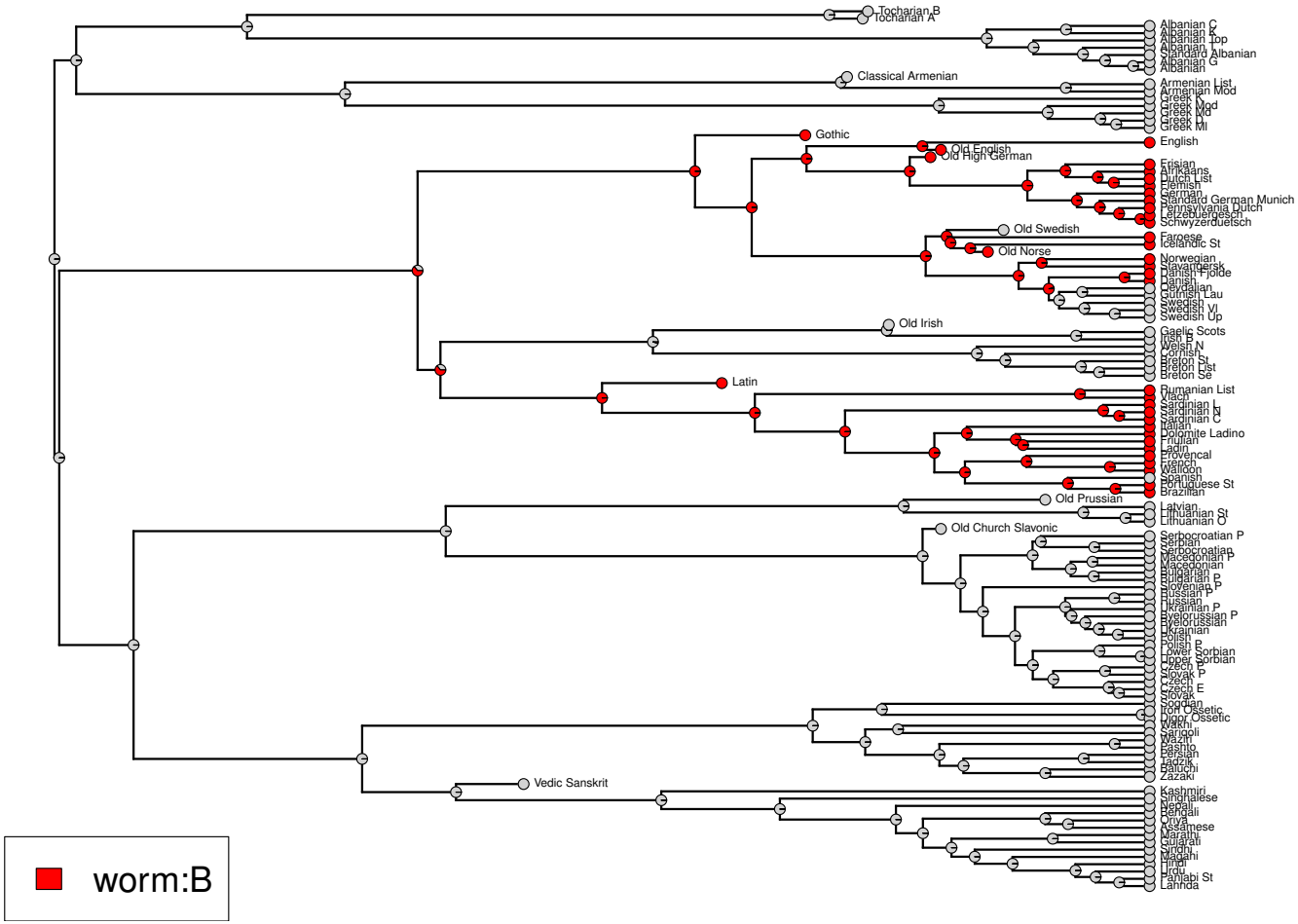
False negatives



False negatives



False negatives



Summary on Indo-European

As the qualitative evaluation shows, the proto-forms proposed to be reconstructed back to PIE by our best ASR method are mostly equally good if not even better candidates than those which we found in the gold standard. Given the general and well-known uncertainties in semantic reconstruction in classical historical linguistics, it seems that ASR methods could provide actual help in semantic reconstruction by providing objective evolutionary scenarios for word evolution along a given tree which follow a specific evolutionary model.

Hands-on

How to run Maximum-Likelihood tree estimation in Paup*

- Load your nexus file in to Paup*
  ```
  >paup4 soundConcept.bin.nex
  ```
- set optimality criterion to likelihood
  ```
  paup> set criterion=likelihood
  ```
- choose model:
  - optimized rate parameter
    ```
    paup> lset basefreq=estimate
    ```
  - ultrametric tree:
    ```
    paup> lset clock=yes
    ```
  - gamma-distributed rates
    ```
    paup> lset rates=gamma shape = estimate
    ```
  - assume invariant sites
    ```
    paup> lset pinvar=estimate
    ```

Hands-on

How to run Maximum-Likelihood tree estimation in Paup* (cont.)

- perform heuristic search
  `paup> hsearch`
- display tree
  `paup> describetree /plot=phylo`
- show log-likelihood and AIC
  `lscores /aic=yes`

Bouchard-Côté, A., D. Hall, T. L. Griffiths, and D. Klein (2013). Automated reconstruction of ancient languages using probabilistic models of sound change. *Proceedings of the National Academy of Sciences*, **36**(2):141–150.

Ewens, W. and G. Grant (2005). *Statistical Methods in Bioinformatics: An Introduction*. Springer, New York.