

Evrimsel Biyolojide Deney ve Kuram İlişkisi

Betul Kacar

NASA Astrobiyoloji Enstitüsü

Georgia Tech

betul.kcr@gmail.com

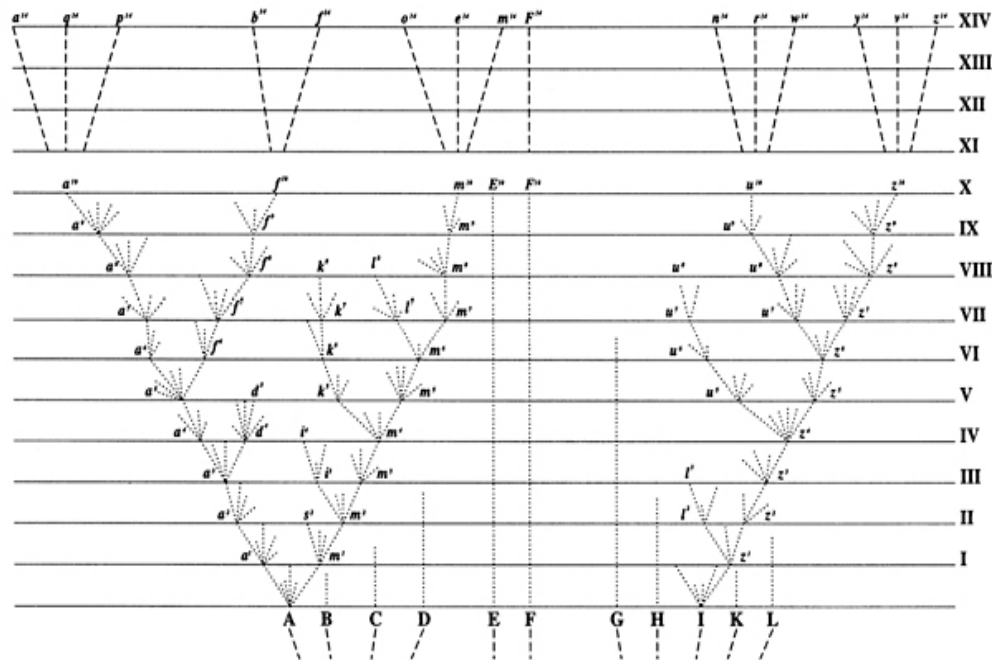
Matematiksel Evrim Çalıştayı

Matematik Köyü, Şirince, İzmir

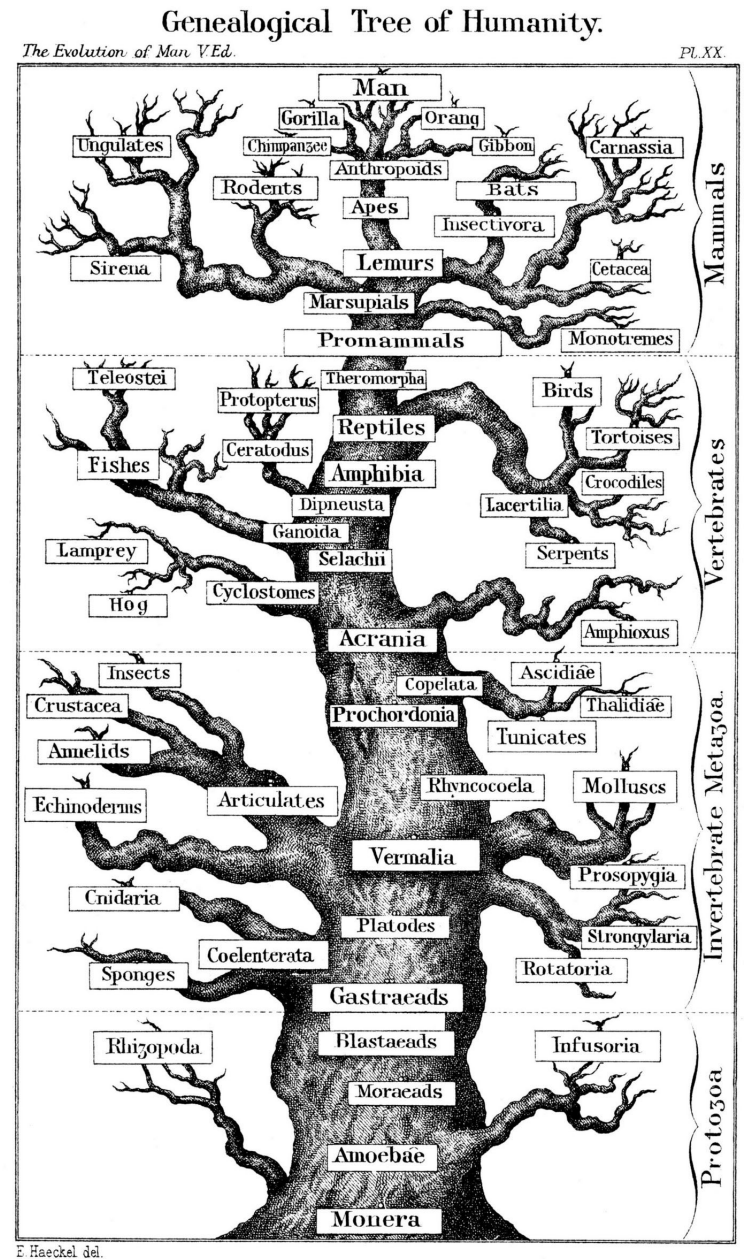
Eylül 2013

Bölüm I: Atasal gen oluşturma

Filogeni Ağaçları

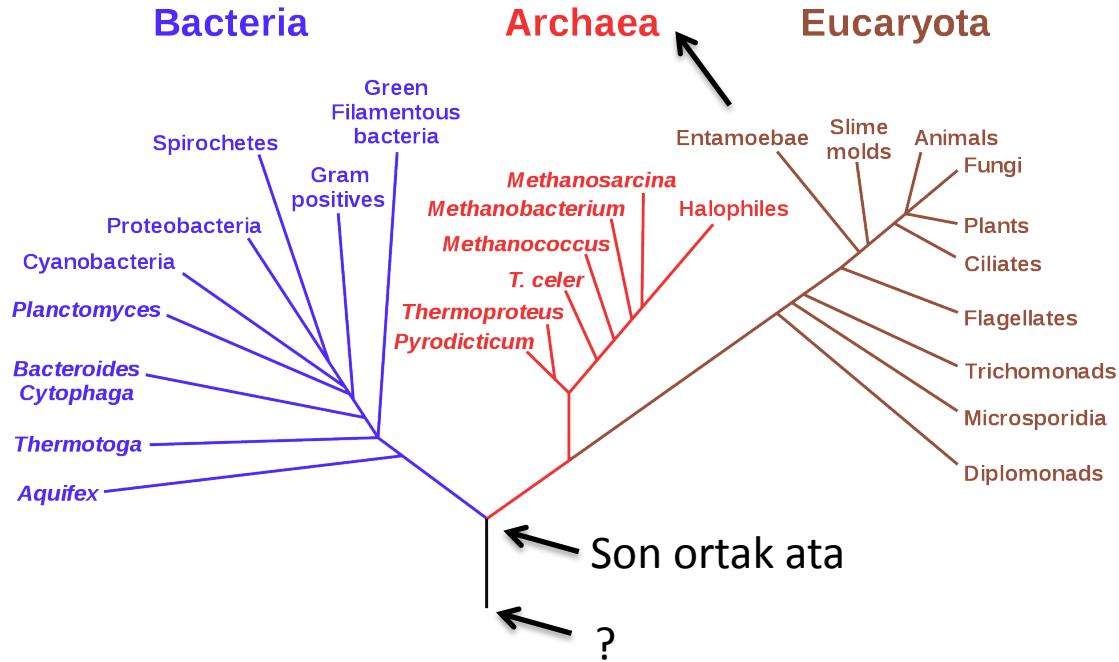


Charles Darwin (1859)
Descent with modification



Ernest Haeckel (1879)
Evolution of man

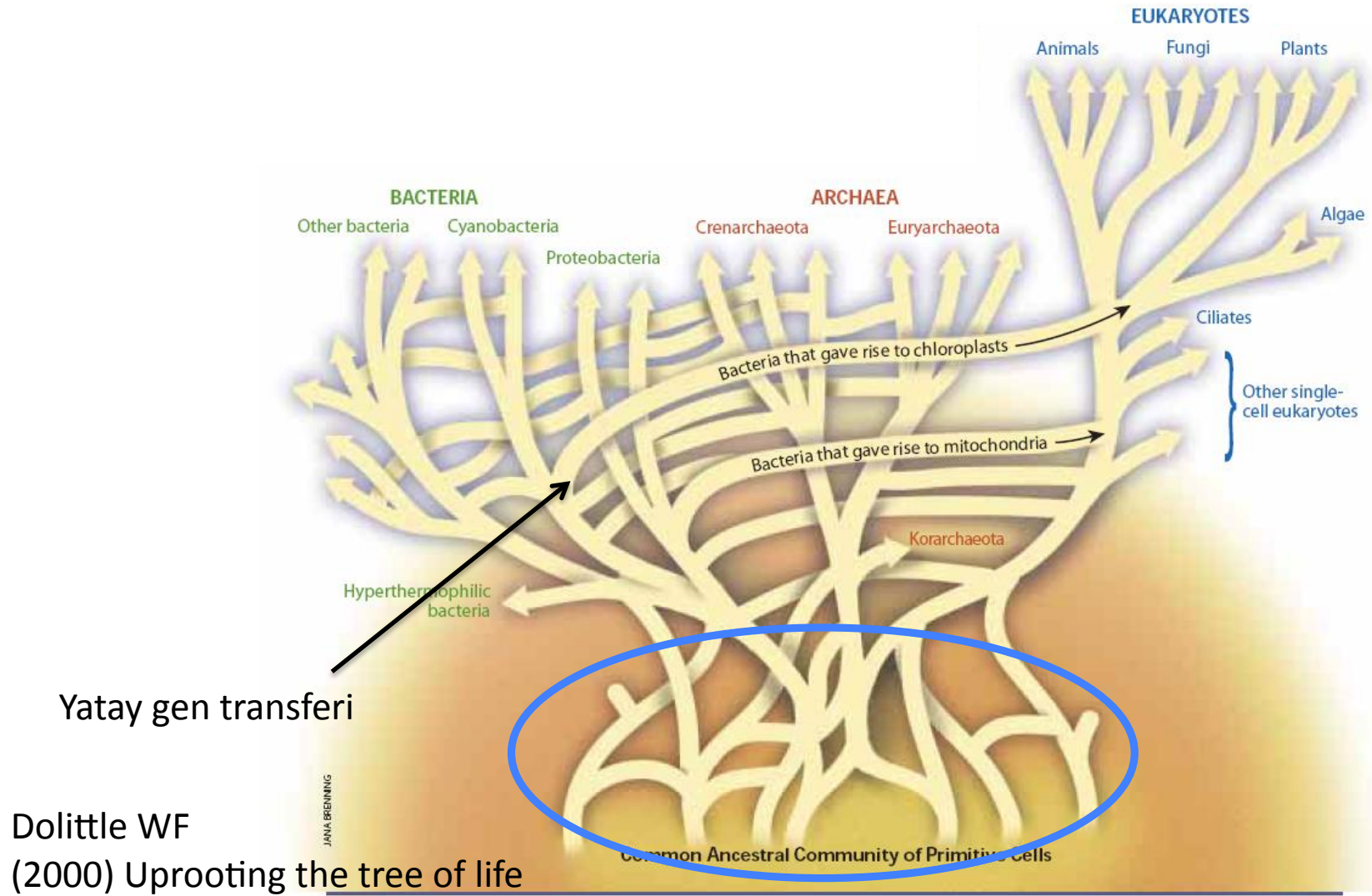
Filogeni Ağaçları



Woese (1990)

Phylogenetic tree of life (uyarlama)

Filogeni Ağaçlarının Evrimi

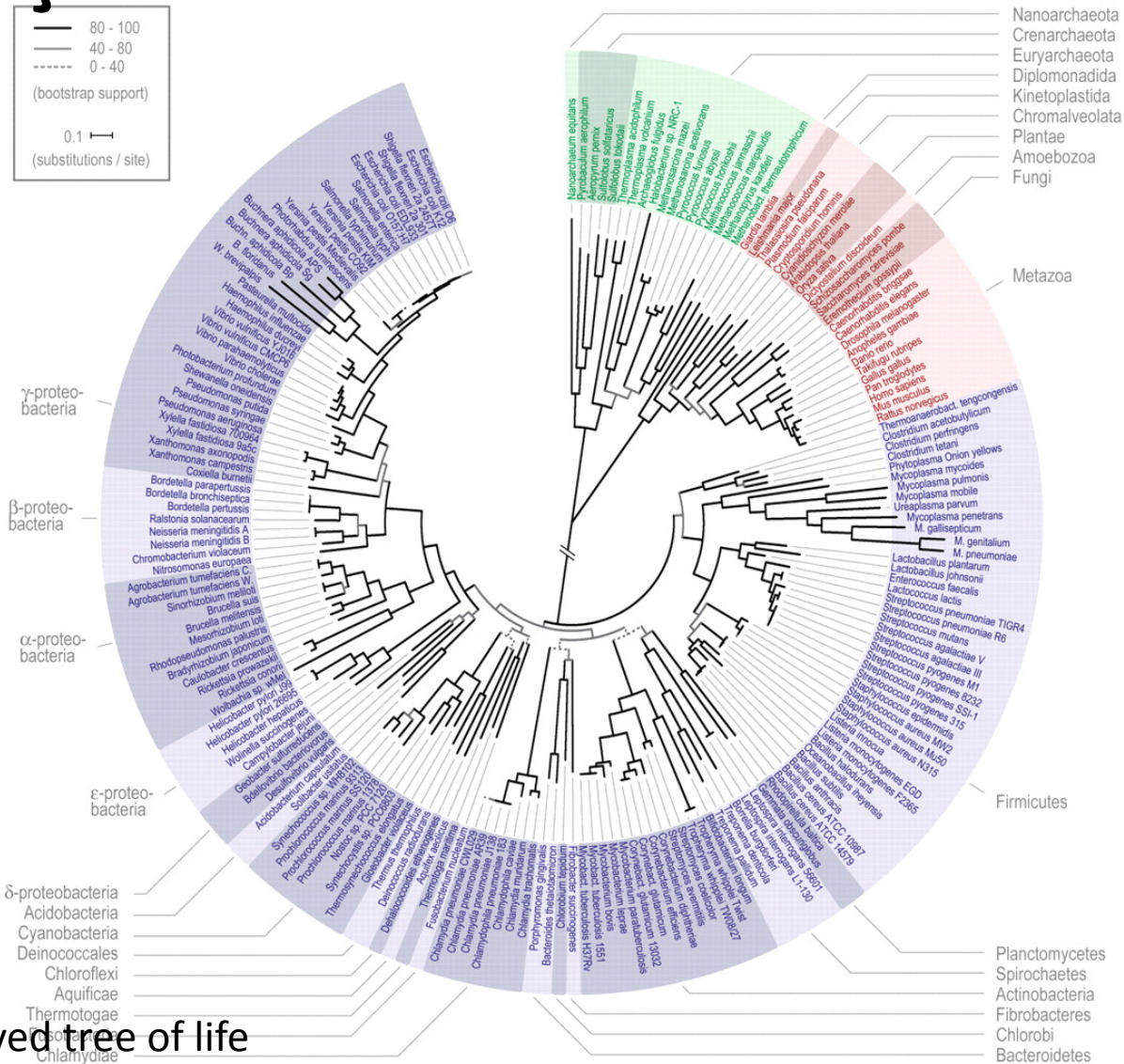


Filogeni Ağaçlarının Evrimi

191 tür,
31 ortolog

Ciccarelli (2006)

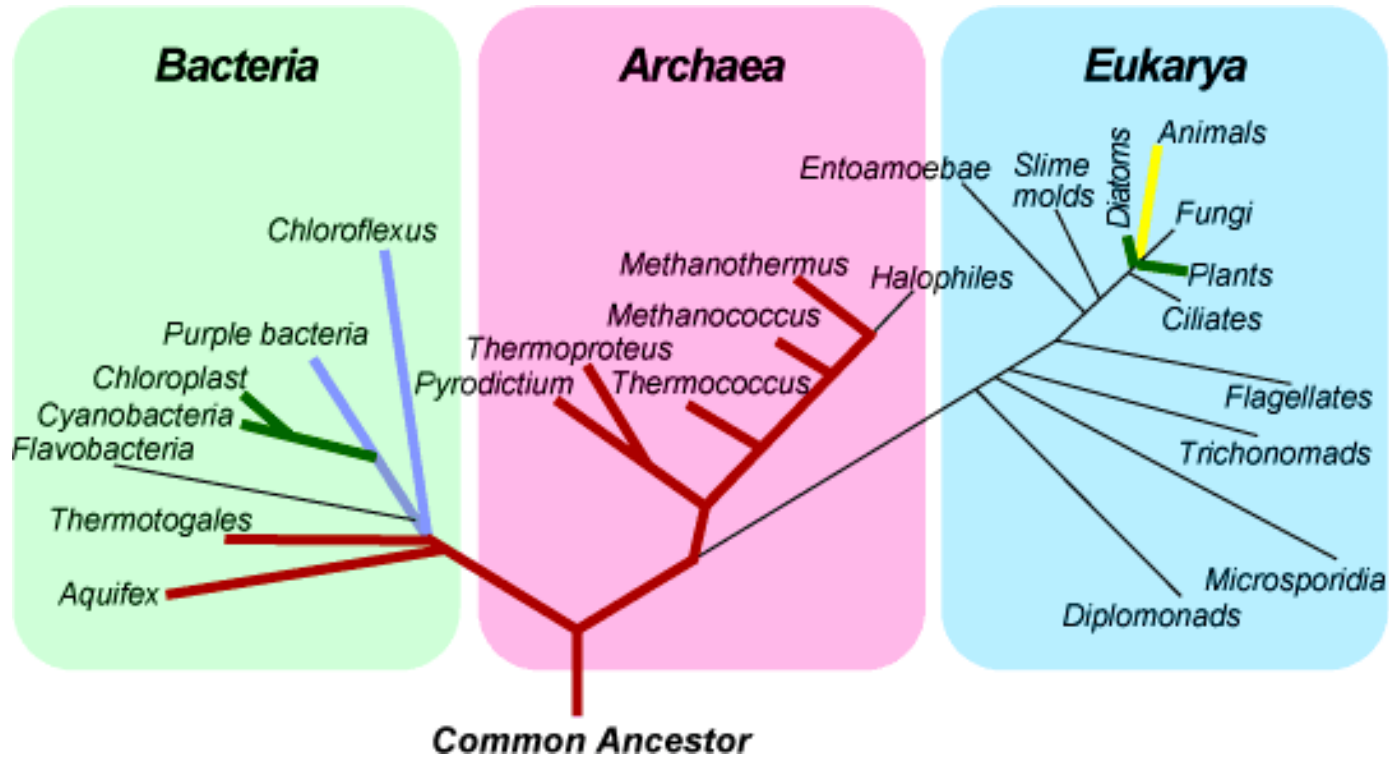
Towards a highly resolved tree of life



Önemli noktalar:

- Yatay gen transferi
- Gen / genom boyutunda azalma / çoğalma
- Her gen farklı bir agac topolojisine sahip olabilir (Koonin NIH / KITP)
- Son ortak ata (ailesi?) (LUCA):
Last Universal Common Ancestor
(Özellikleri?)

Universal (Dünyanın) Hayat Ağacı



Mesofilik=20-40 °C Thermofilik=45-75 °C Hypertermofilik≥80 °C

Hipotez Last Universal Common Ancestor (LUCA) hidrotermal bacada yasayan bir hipertermofili (>80 °C) idi

Termofilik son ortak ata? (okumalar)

- Brochier & Philippe 2002 Nature
- Forterre 1996 Cell 85, 789-792
- Galtier N, Tourasse N, Gouy M 1999 Science, 283
- Di Giulio 2000 Gene, 281, 11-17
- Xu & Glansdorff 2002 Comp Biochem Phys A, 677-688

Atasal Genleri Canlandırma

- Molekuler filogeni (gecmis → bugun)
Protein üzerindeki evrimsel kısıtlamalar?
- ASR (bugun → gecmis)

Atasal Genleri Canlandırma

Uygulamalı Filogeni

Neden?

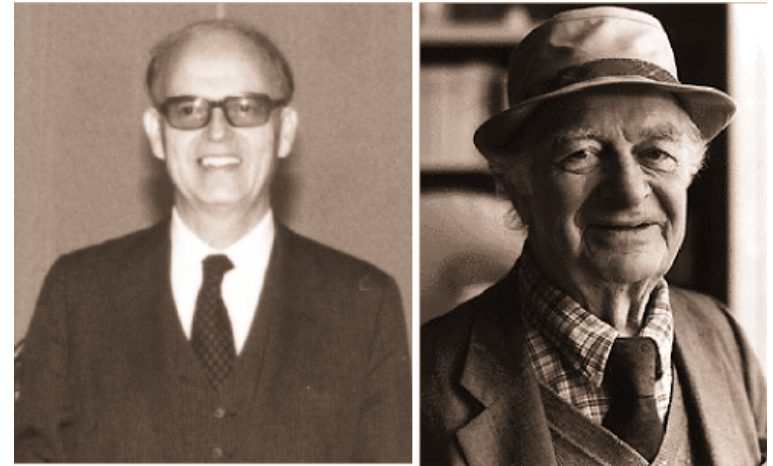
- Protein ve nükleotid diziliminin evrimi
 - DNA, amino asit dizilimlerin proteinin yapısına (ve işlevine etkileri)
 - Protein işlevinin zaman içindeki değişimlerini gözlemleme imkanı
- Antik çevre şartları hakkında bilgi
 - Antik proteinlerin fiziko/biyokimyasal özellikleri
 - Paleo-kuşullar hakkında hipotezler
 - Hayatin başlangıcı

Atasal Genleri Canlandırma

Tarihçe

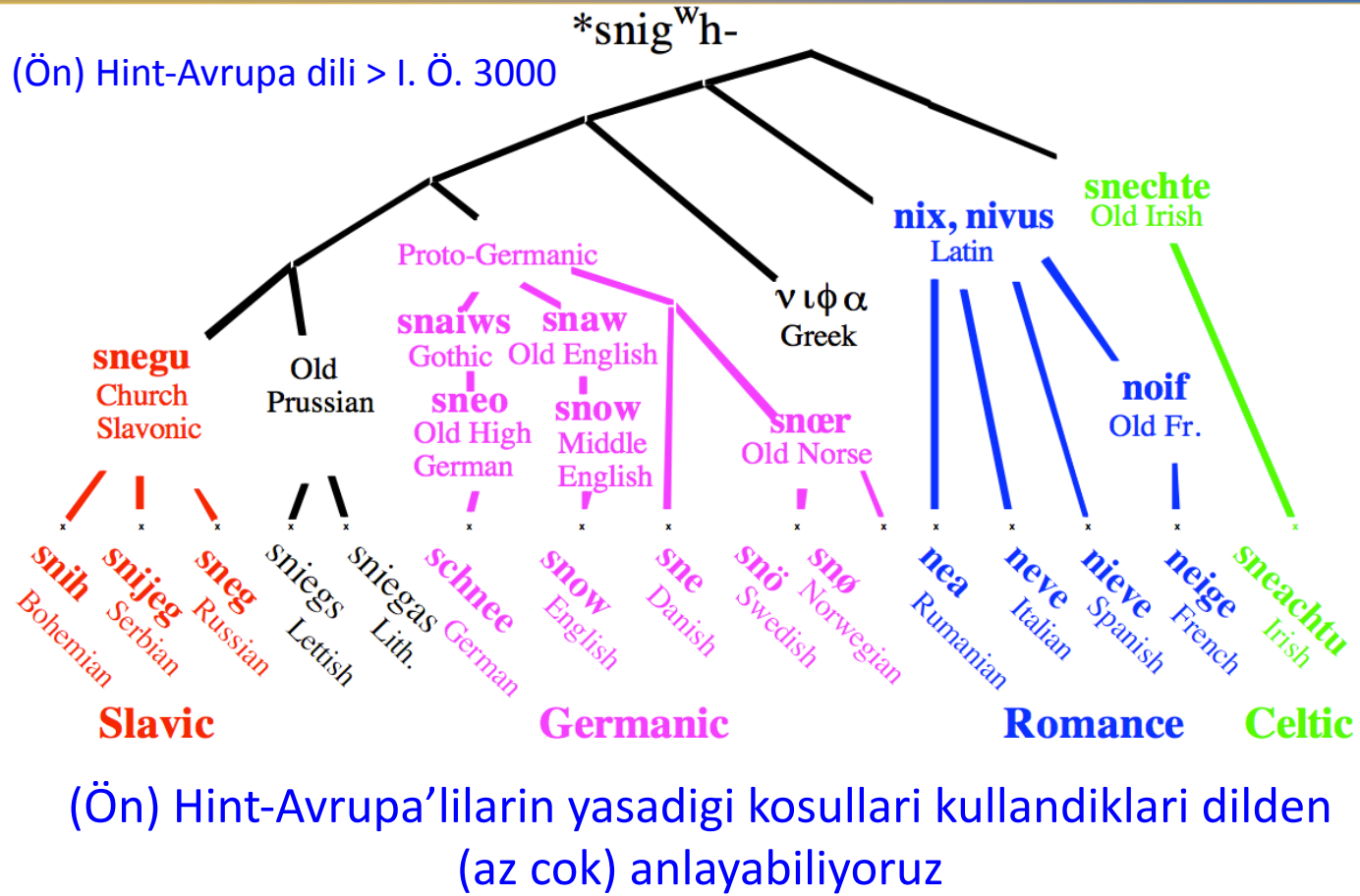
- Pauling, L. and Zuckerkandl, E. (1963) Chemical Paleogenetics 'molecular restoration' studies of extinct forms of life.

Acta Chem. Scand. 17, S9–S16

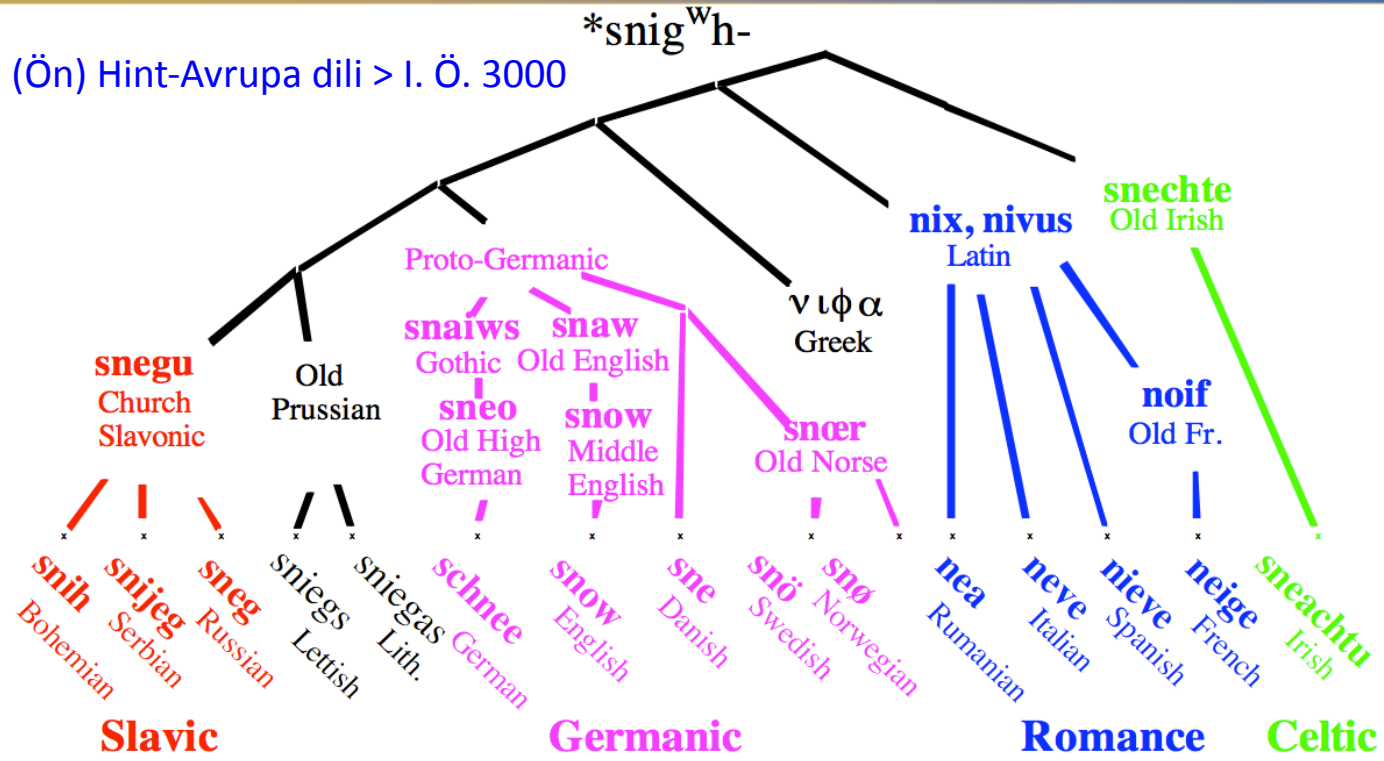


- Paleogenetik & DNA bazlı moleküler dil bilimciliği?

GEÇMİŞİ GÜNÜMÜZDEN YOLA ÇIKARAK TEKRAR OLUSTURMAK

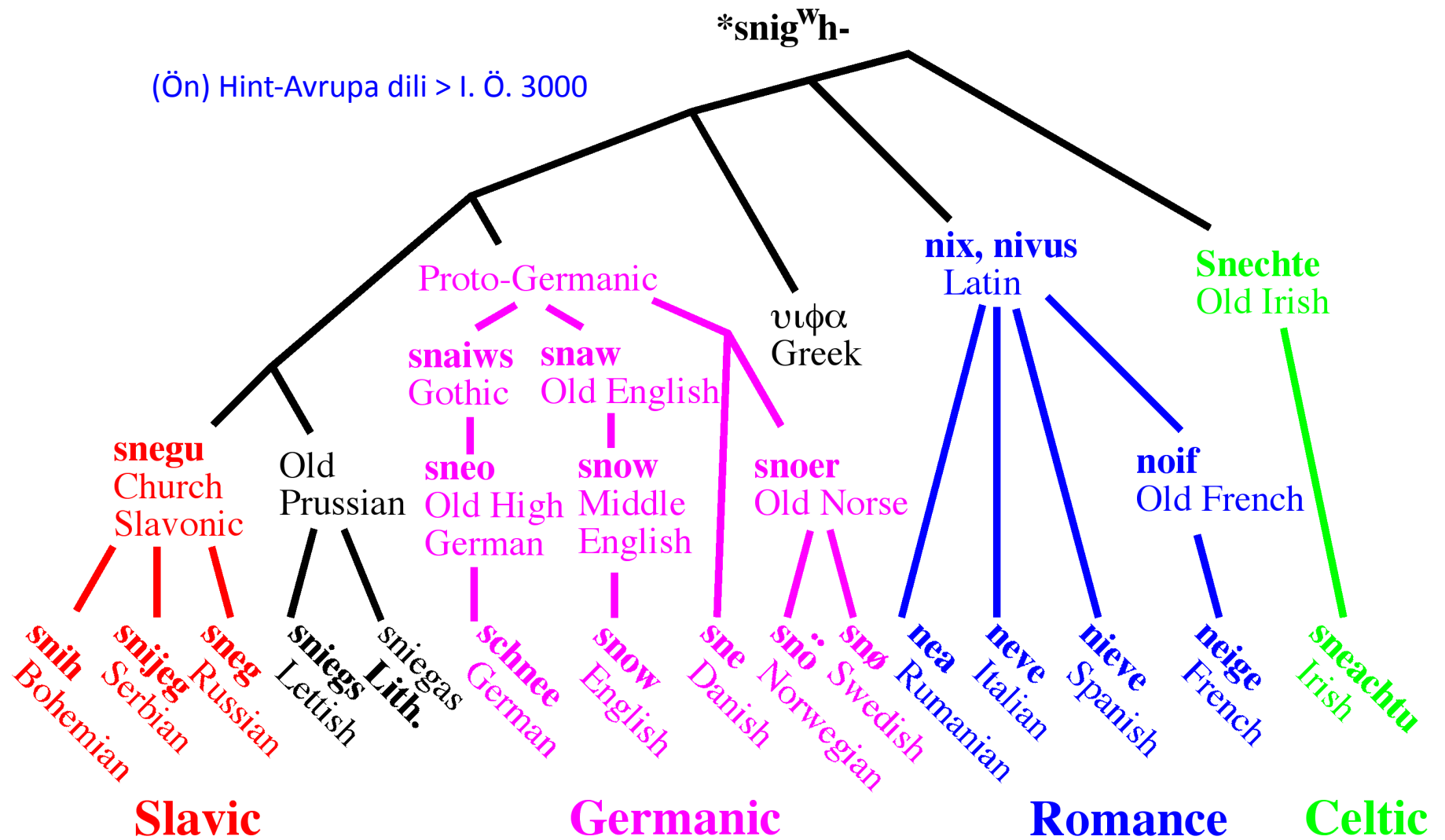


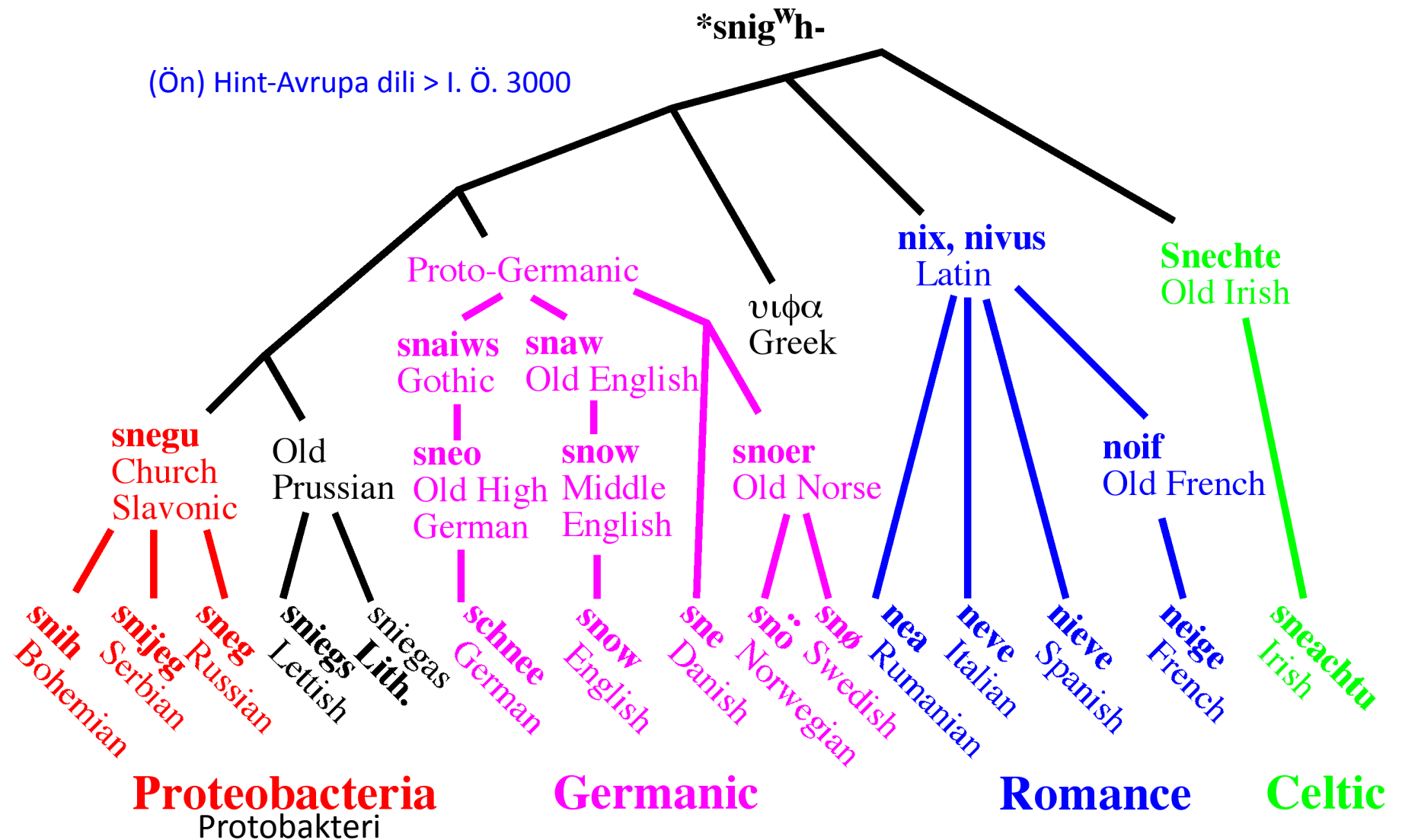
GEÇMİŞİ GÜNÜMÜZDEN YOLA ÇIKARAK TEKRAR OLUSTURMAK

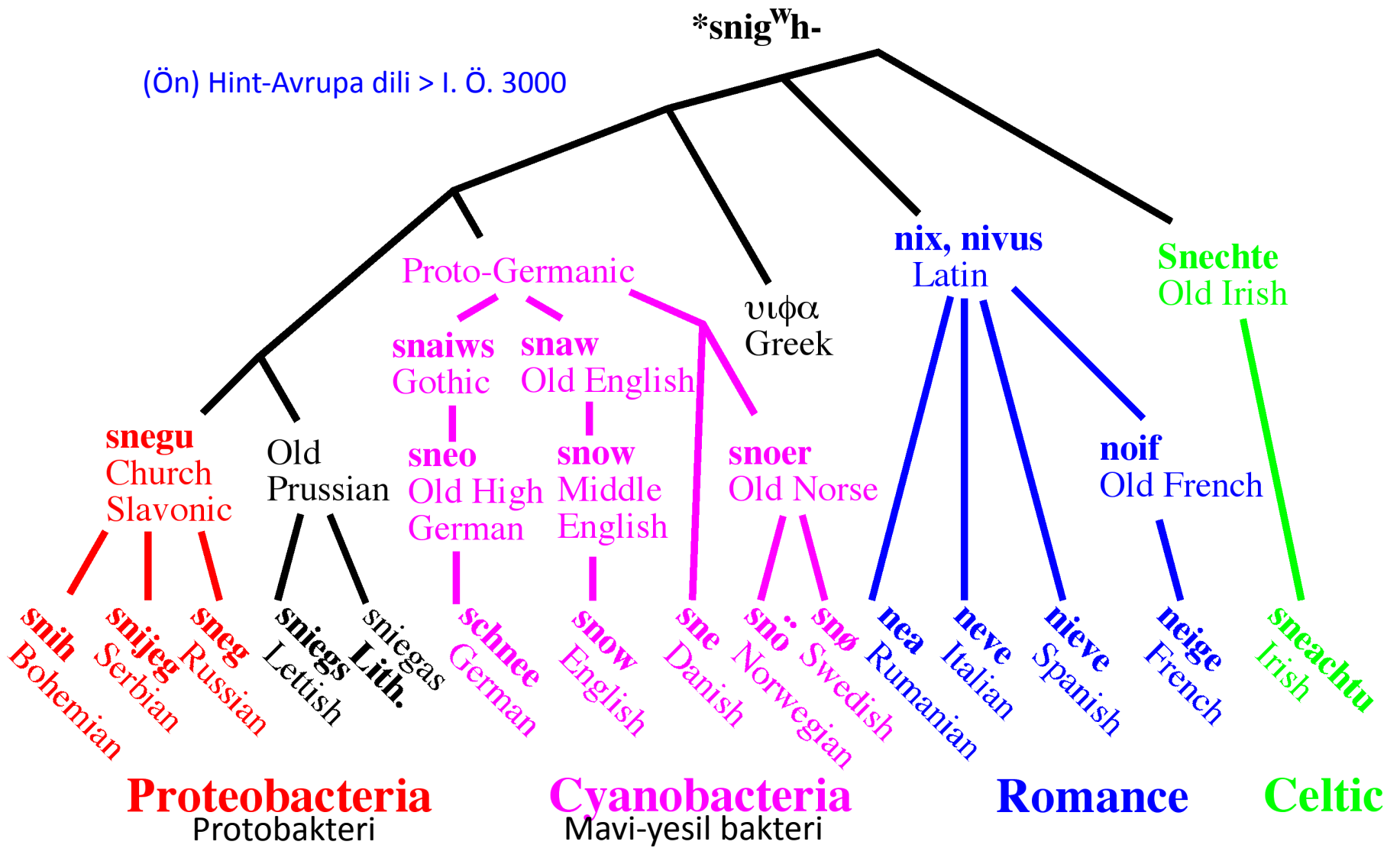


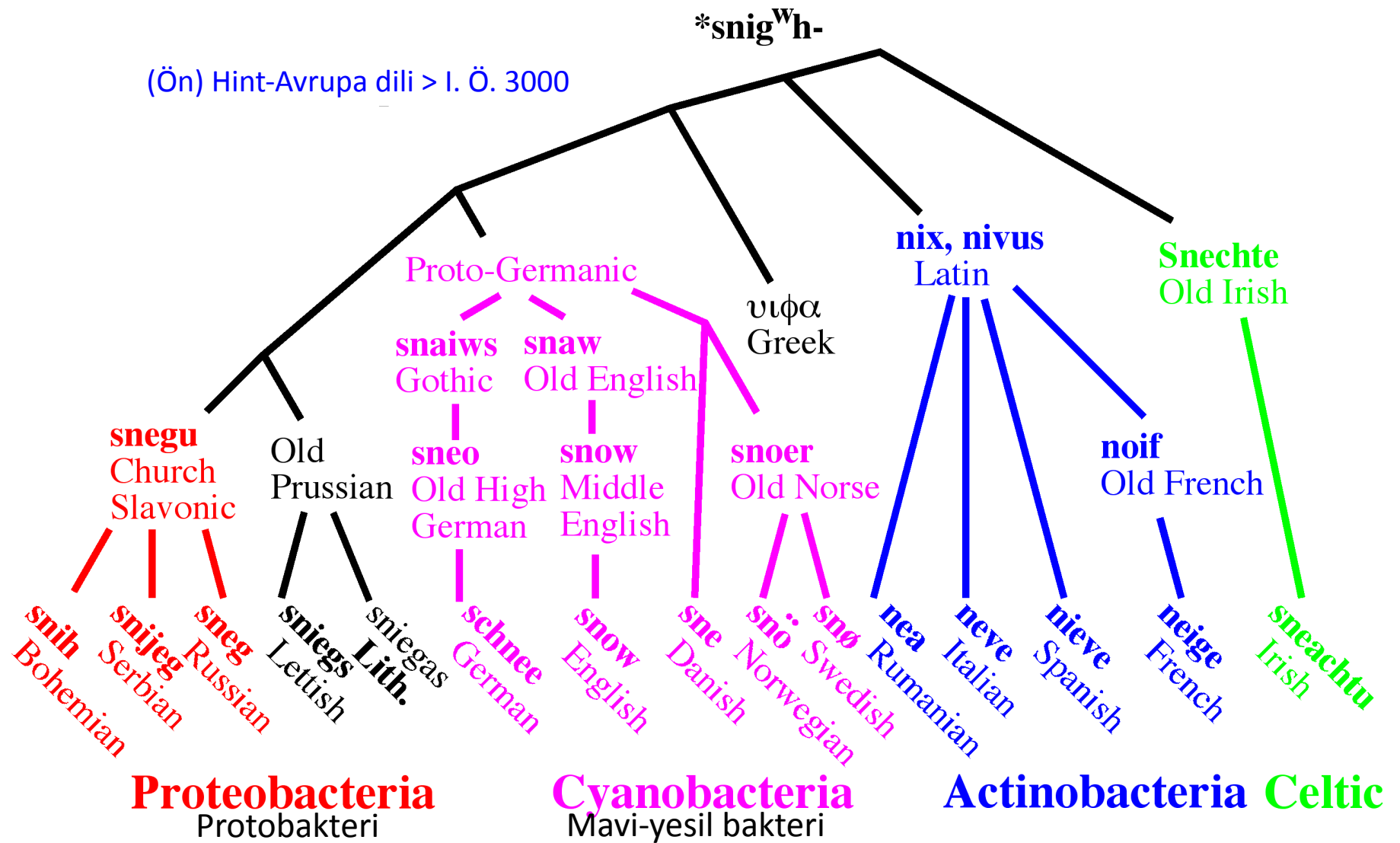
(Ön) Hint-Avrupa'lilarin yasadigi kosullari kullandiklari dilden
(az cok) anlayabiliyoruz

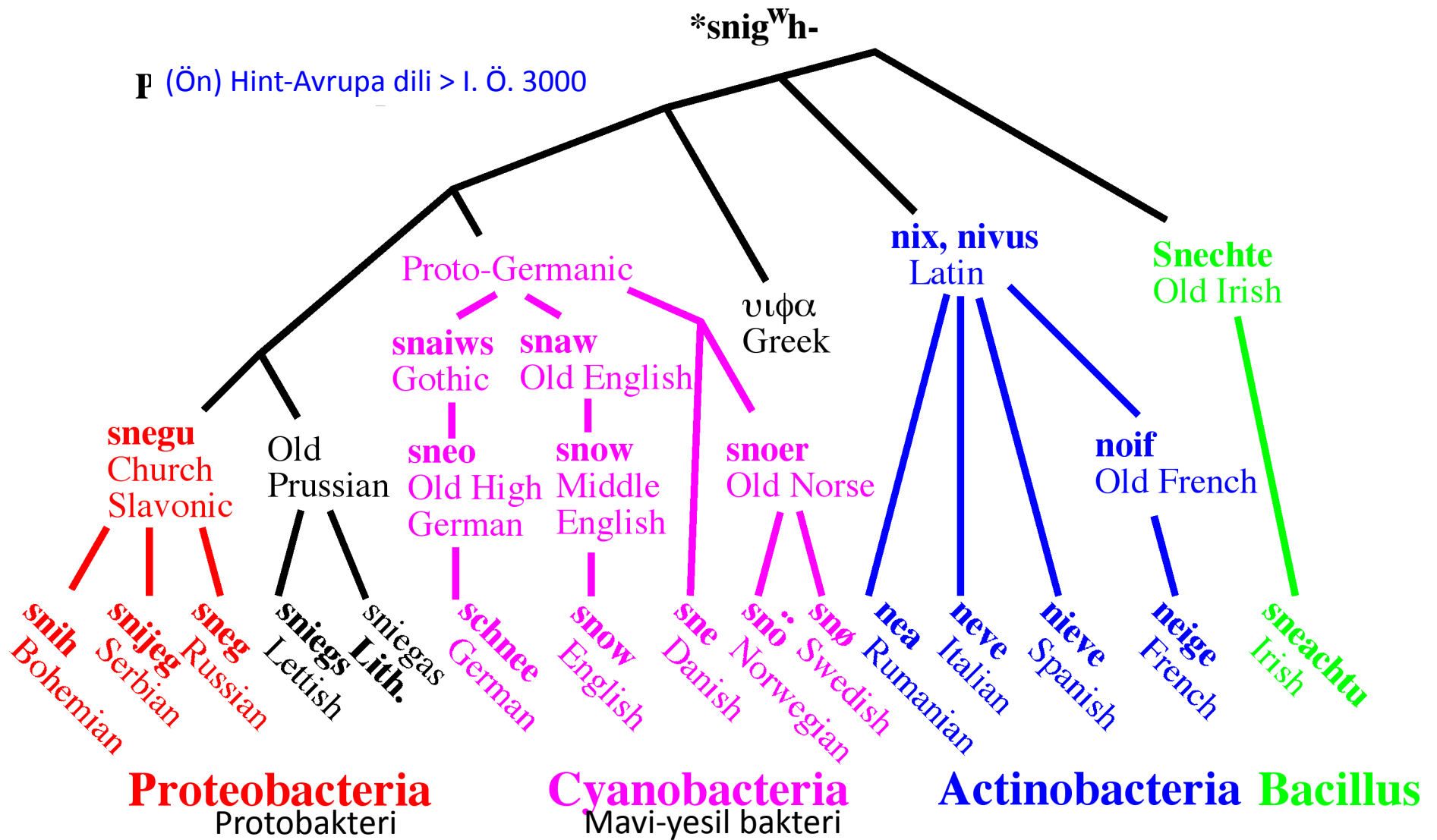
Karli yerlerde yasamislar..

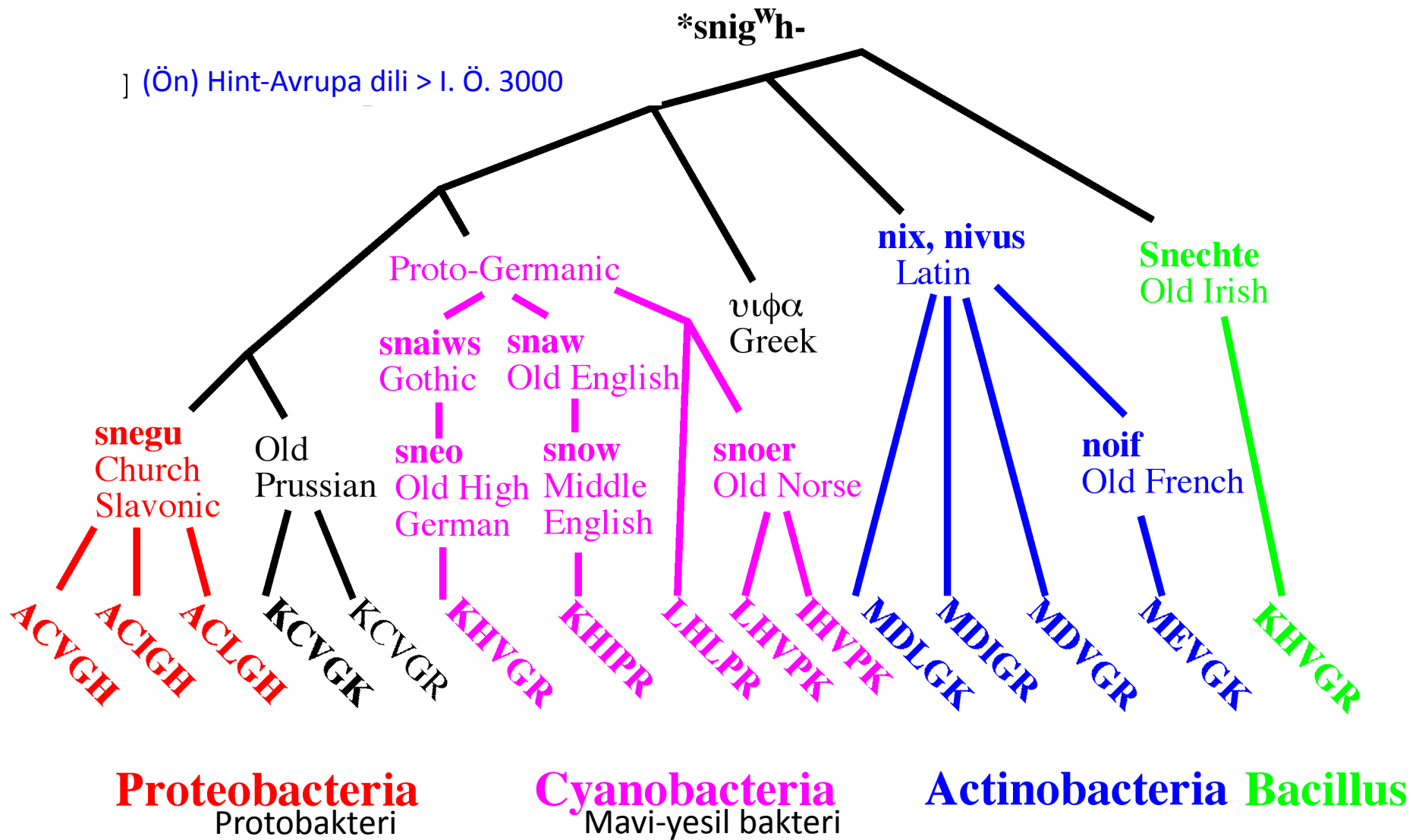


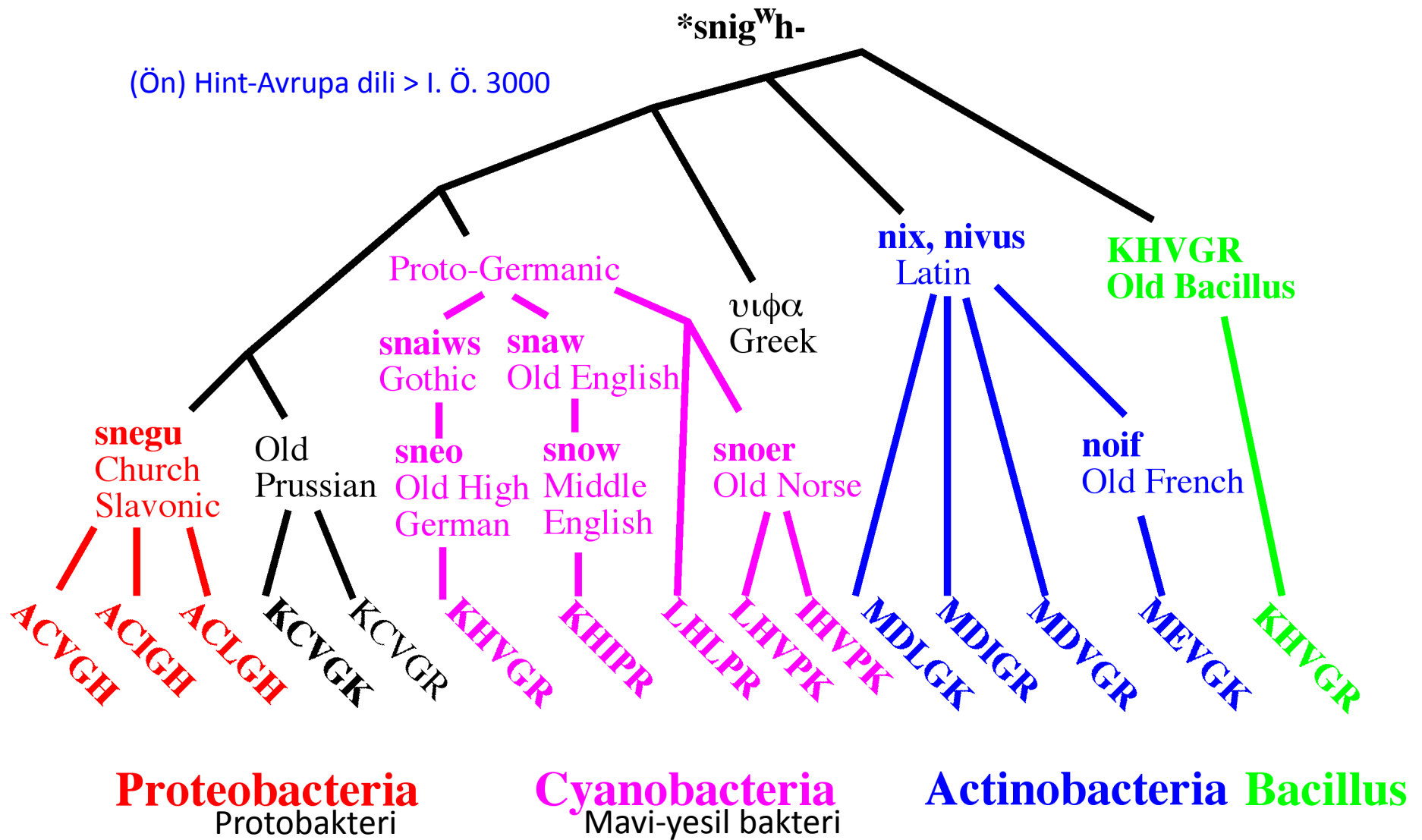


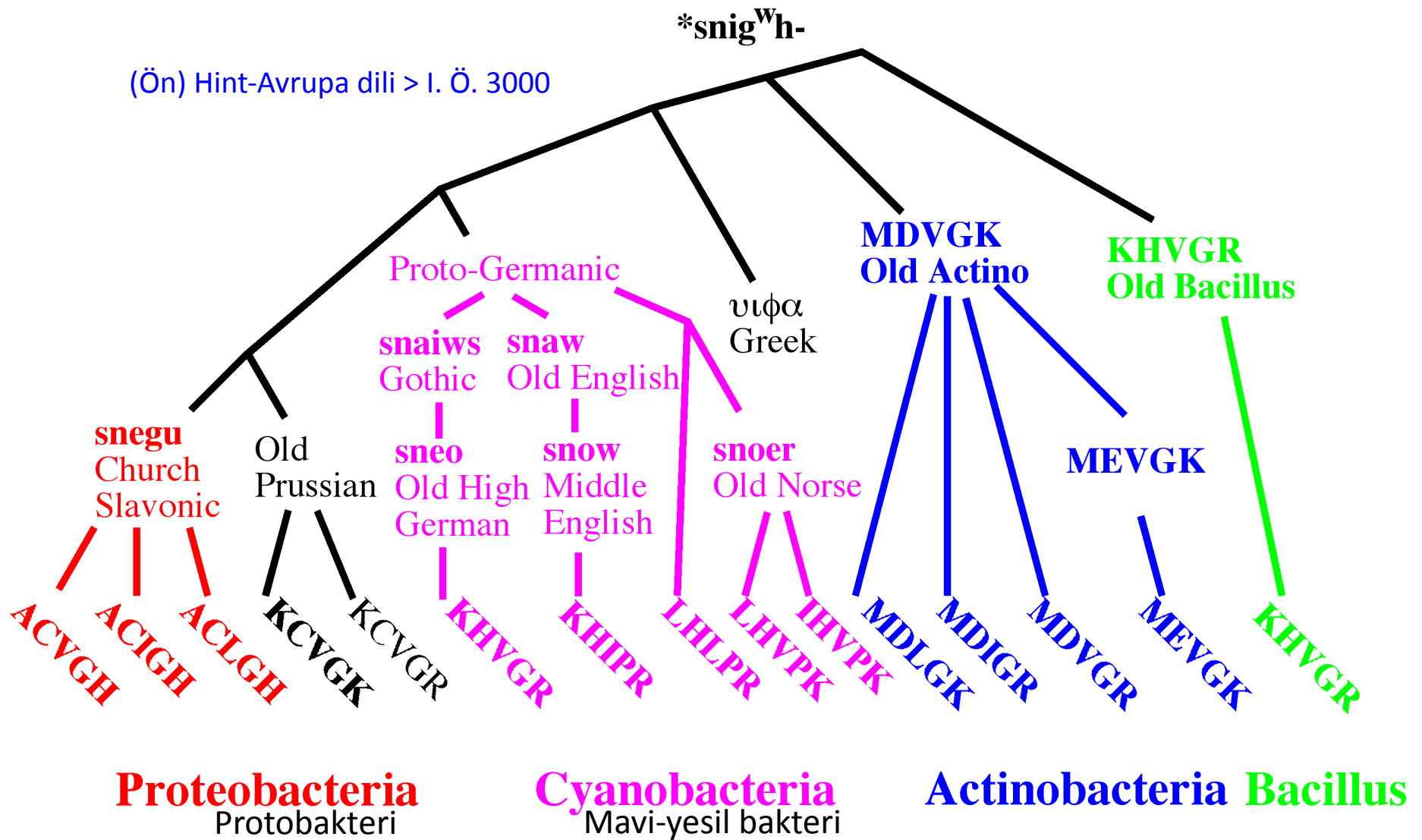


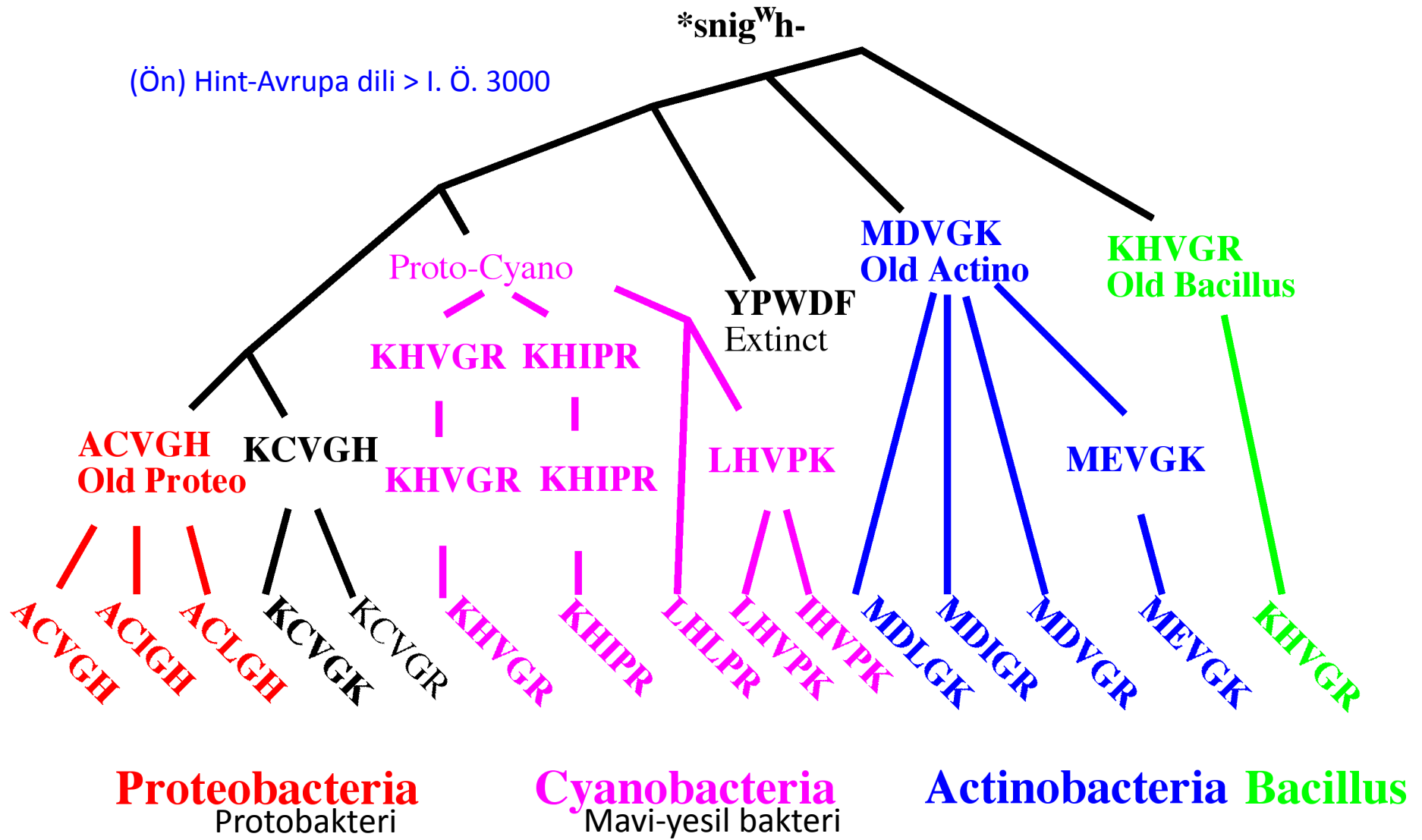












Atasal Genleri Canlandırma

Tarihçe

- Kuram & Modelleme & Sentetik biyoloji

Atasal gen canlandırma üzerine ilk örnek:

Stackhouse, J., et al. **(1990)**,
'The ribonuclease from an extinct bovid
ruminant', *FEBS Lett*, 262 (1), 104-6.

(laboratuarda sentezlenen ilk atasal protein)



Steve Benner

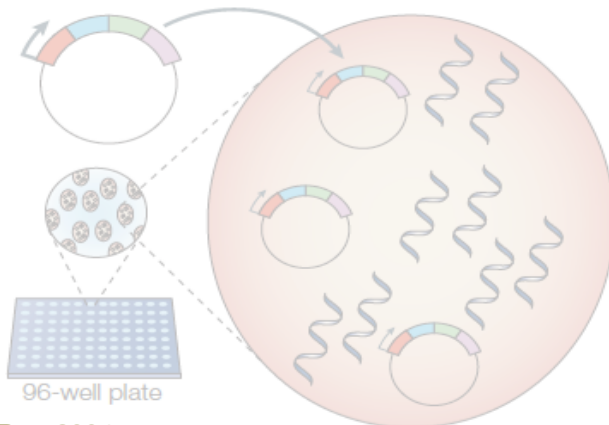
Nasıl?



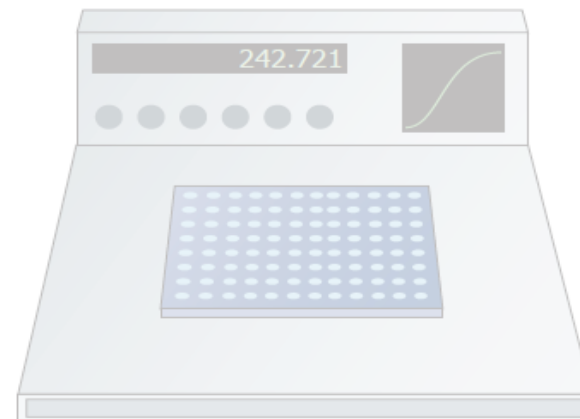
1) Her gen/protein için bir filogeni ağacı oluşturma

- Hem DNA hem amino asit kullanılabilir
- (Filogeni ağaçları yapımı için bkz. Bilgin ders notları)

d Subclone assembled gene into vector, transform cultured cells and express ancestral protein



e Purify ancestral protein (if necessary) and characterize function using *trans*-activation, binding or other assay



a Infer phylogenetic tree from aligned sequences and determine best-fitting evolutionary model



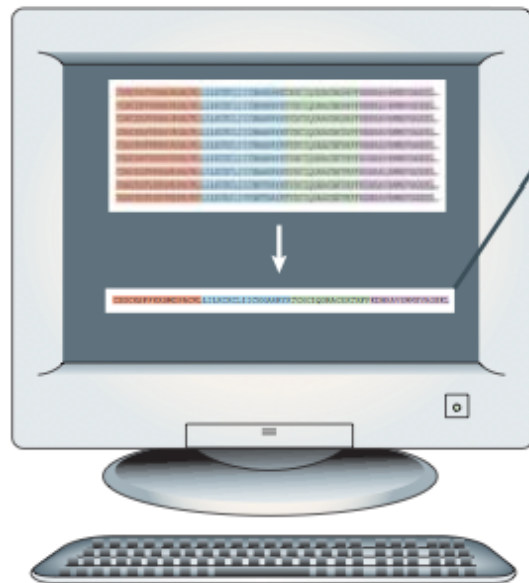
b Reconstruct protein sequence at ancestral node by maximum likelihood

Ancestral sequence

CEGCKSFFKKSMDPACMLLILKCKCLII CHHARRYKTCHCIQGRACEKTKFFKDNRAVEMMEVAGEKL

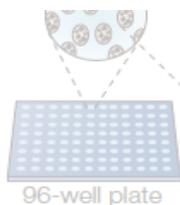
Ancestral sequence

CEGCKSFFKKSMDPACMLLILKCKCLII CHHARRYKTCHCIQGRACEKTKFFKDNRAVEMMEVAGEKL

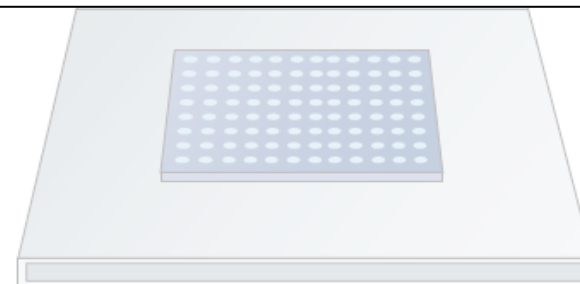


2) Evrimsel modeller / İstatistik yöntemleri kullanarak ortak ataya ulasma

- İki başlıca yöntem: Maximum Likelihood & Parsimony



96-well plate



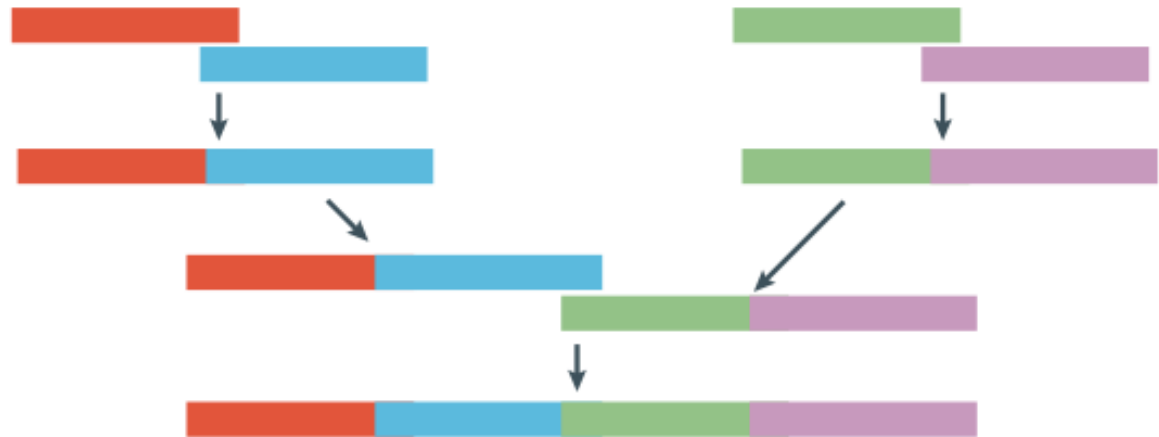
a Infer phylogenetic tree from aligned sequences and determine best-fitting evolutionary model



b Reconstruct protein seq at ancestral node by max likelihood



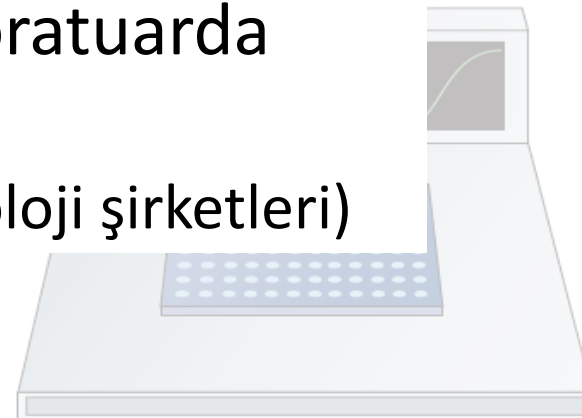
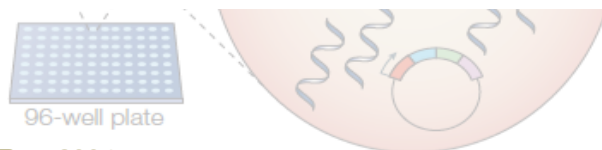
c



d Subclone assembled ge cultured cells and expe

3) Tahmin edilen dizilerin laboratuvarda sentezlenmesi

- PCR (yada bazı biyoteknoloji şirketleri)



a Infer phylogenetic tree from aligned sequences and determine best-fitting evolutionary model



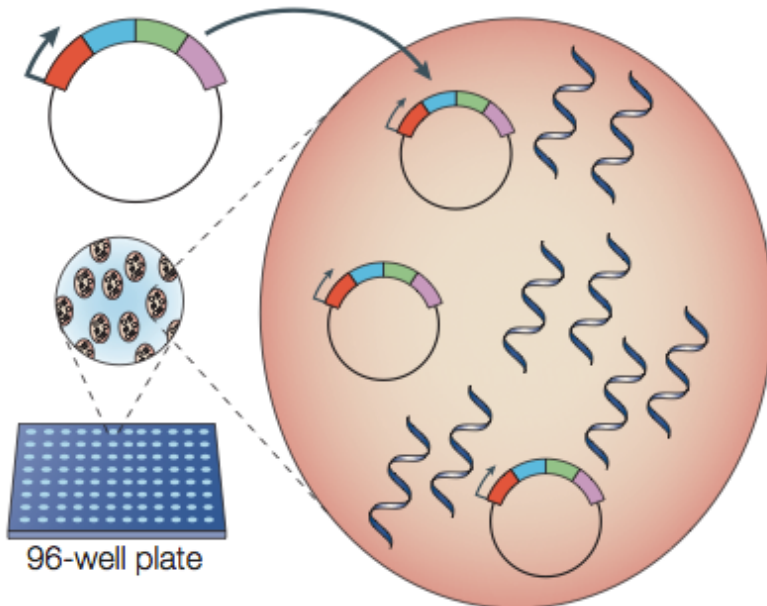
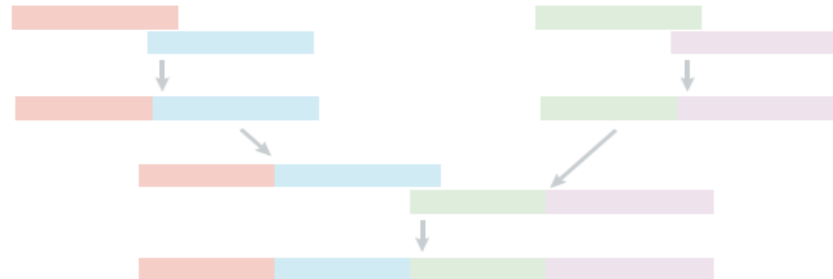
b Reconstruct protein sequence at ancestral node by maximum likelihood



Ancestral sequence

CEGCKSFFKKSMDPACMLLILKCKCLIICHHARRYKTCHCIQGRACEKTKFFKDNRAVEMMEVAGEKL...

c Synthesize oligonucleotides and assemble gene for ancestral protein by stepwise PCR



4) Antik proteinin klonlanması & saflastırılması gibi moleküler biyoloji deneyleri

a Infer phylogenetic tree from aligned sequences and determine best-fitting evolutionary model



b Reconstruct protein sequence at ancestral node by maximum likelihood

Ancestral sequence

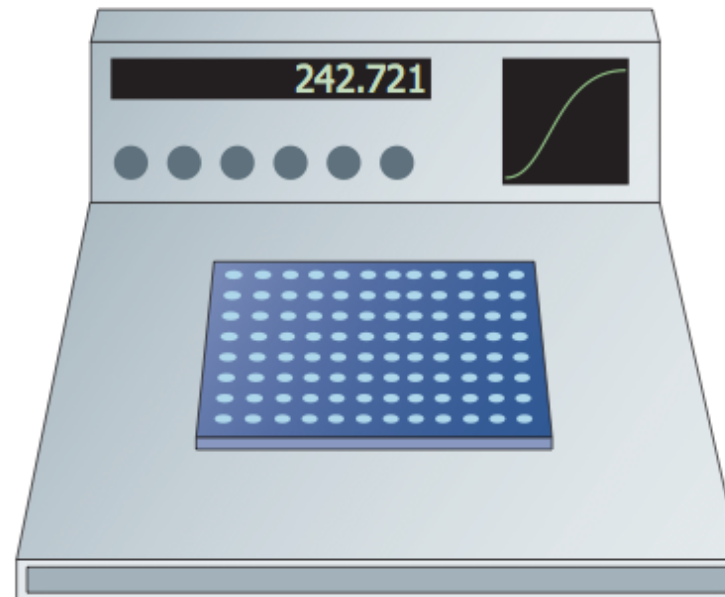
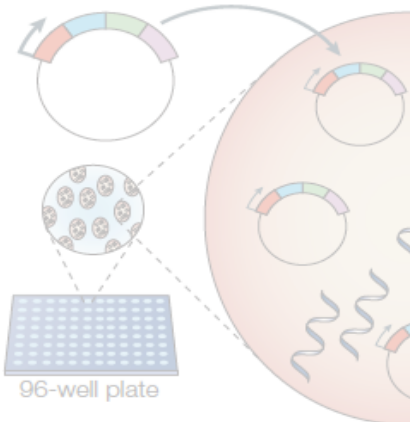
CEGCKSFFKKSMDPACMLLILKCKCLII CHHARRYKTCHCIQGRACEKTKFFKDNRAVEMMEVAGEKL



5) Antik proteinin biyokimyasal/ fiziksel analizi

- *In-vitro*
- *In-vivo* (Kacar & Gaucher)

d Subclone assembled gene into v cultured cells and express ances

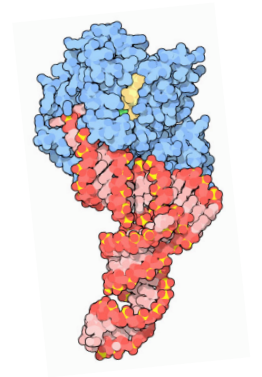
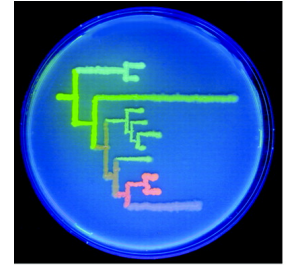
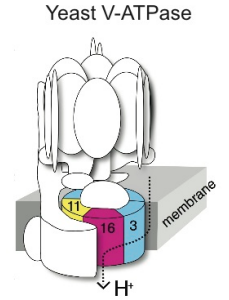


Neler öğrendik?

Hayatın geçmişteki ve günümüzdeki evrimsel koşulları

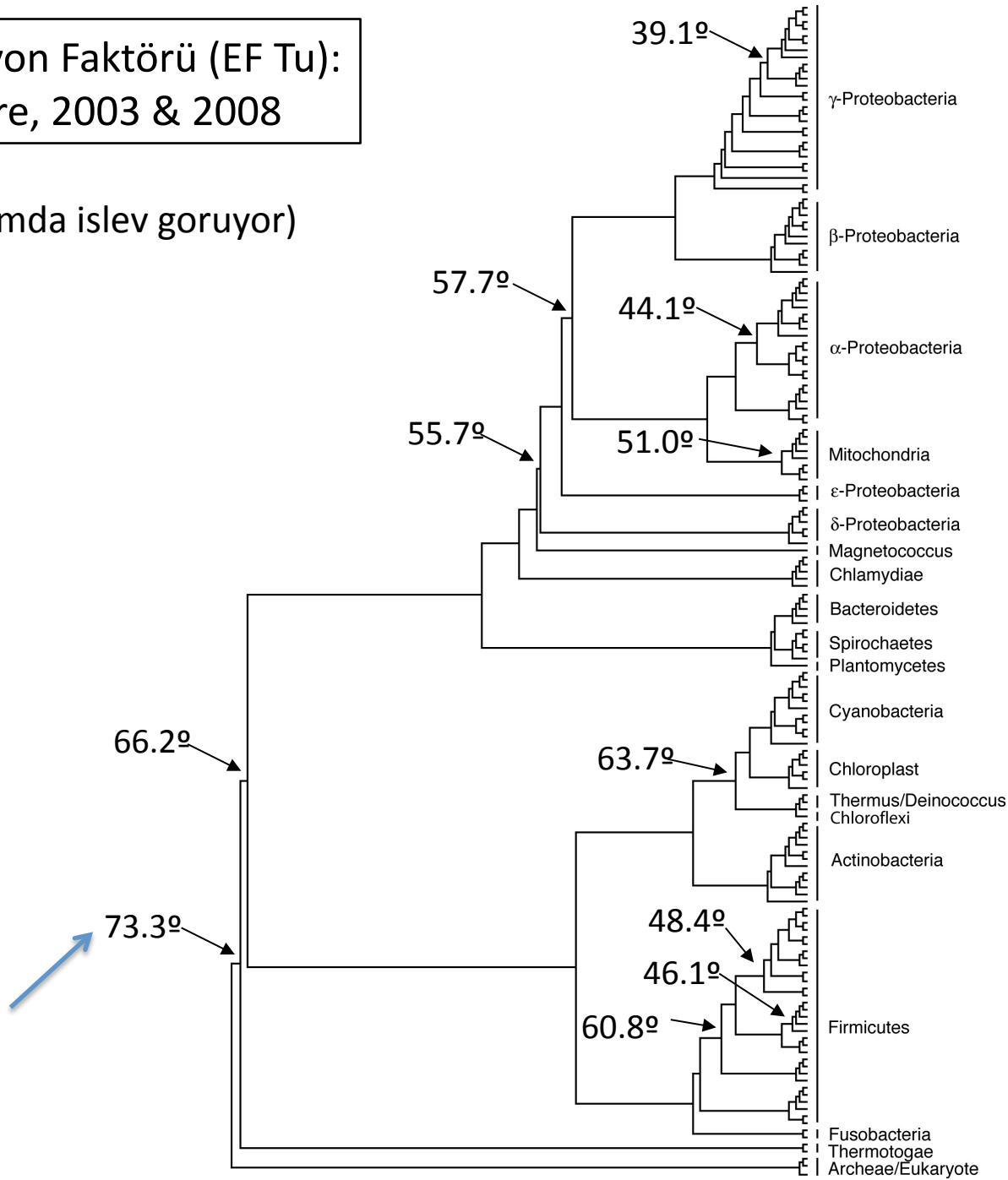
Antik termofili hayat üzerine örnekler

- Kinaz enzimleri (Akanuma 2013, PNAS)
- Redüktif enzimler (Perez-Jimenez, 2012, NSB) (pH ve Sıcaklık)
- Elongasyon faktörleri (Gaucher 2003/8, Nature)
- Kompleks hayatın evrimi (Finnigan 2012, Nature)
- In-vitro moleküler evrim (Ugande 2004, Science)
- Alkol dehidrojenaz enziminin geçmişi (Thomson 2005, Nature)
- Antik metabolizmalar (Voordeckers 2012, PLoS)



Elongasyon Faktörü (EF Tu):
Nature, 2003 & 2008

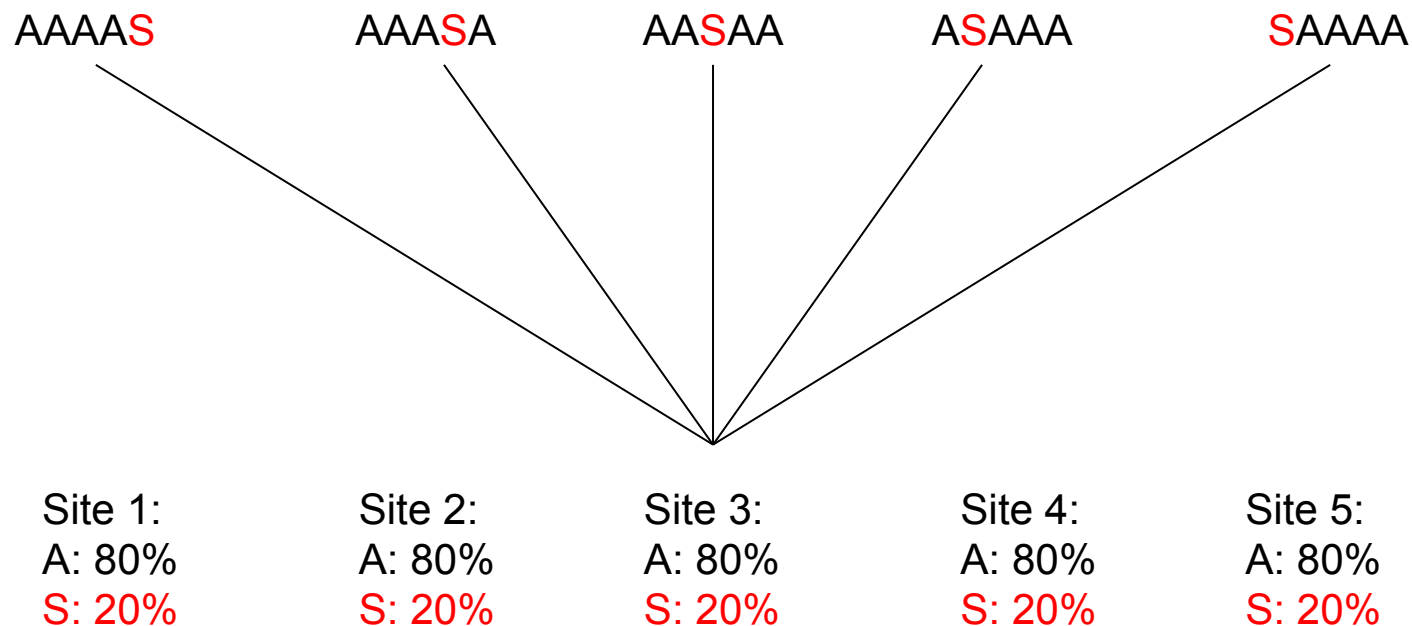
(Ribozomda islev goruyor)



16 ata
Yeniden oluřturuldu

Muhtemel Problemler

Atasal genotip ne kadar doğru? (Atasal genotipi bulmak mümkün mü?)



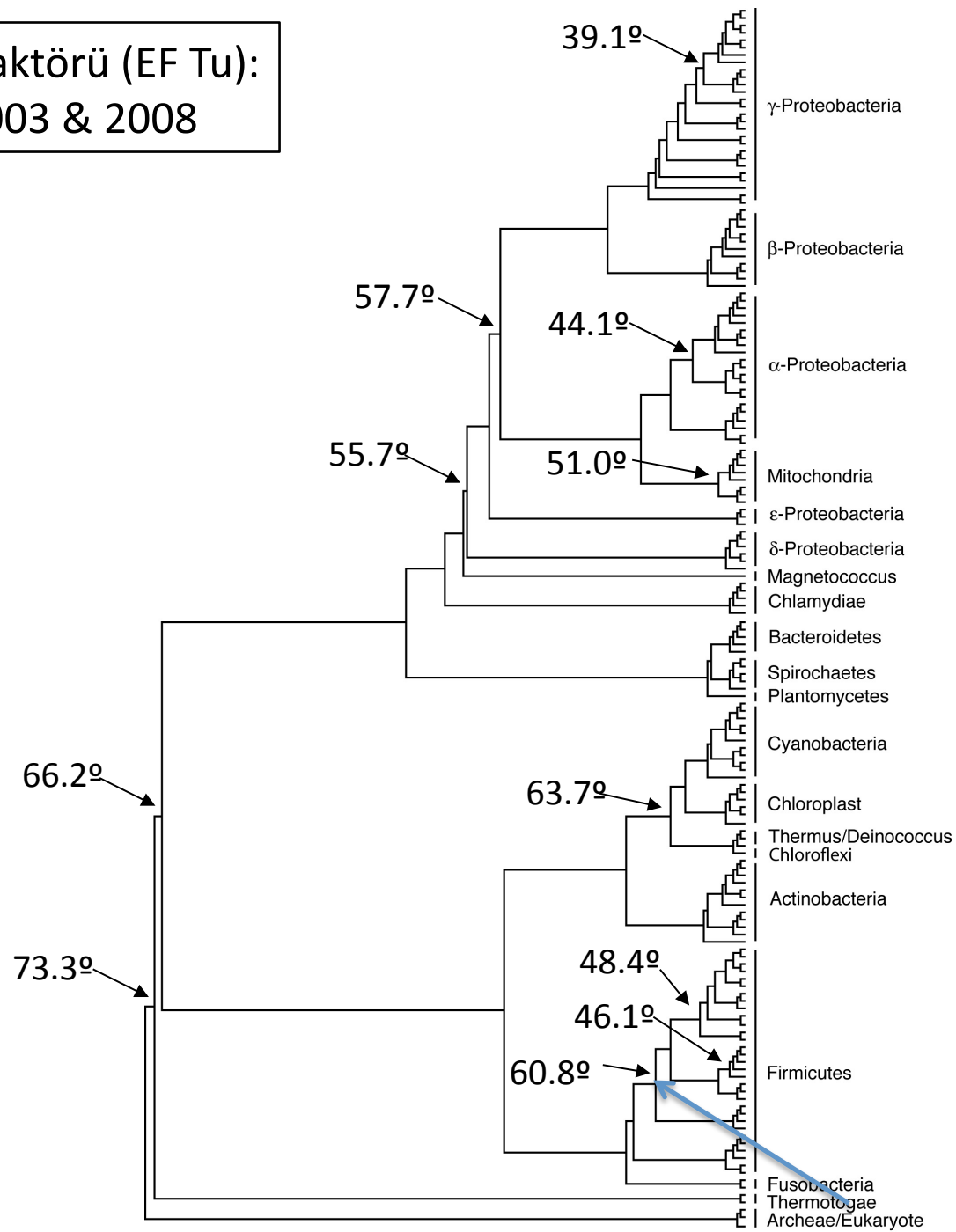
Most-Probabilistic Ancestral Sequence:

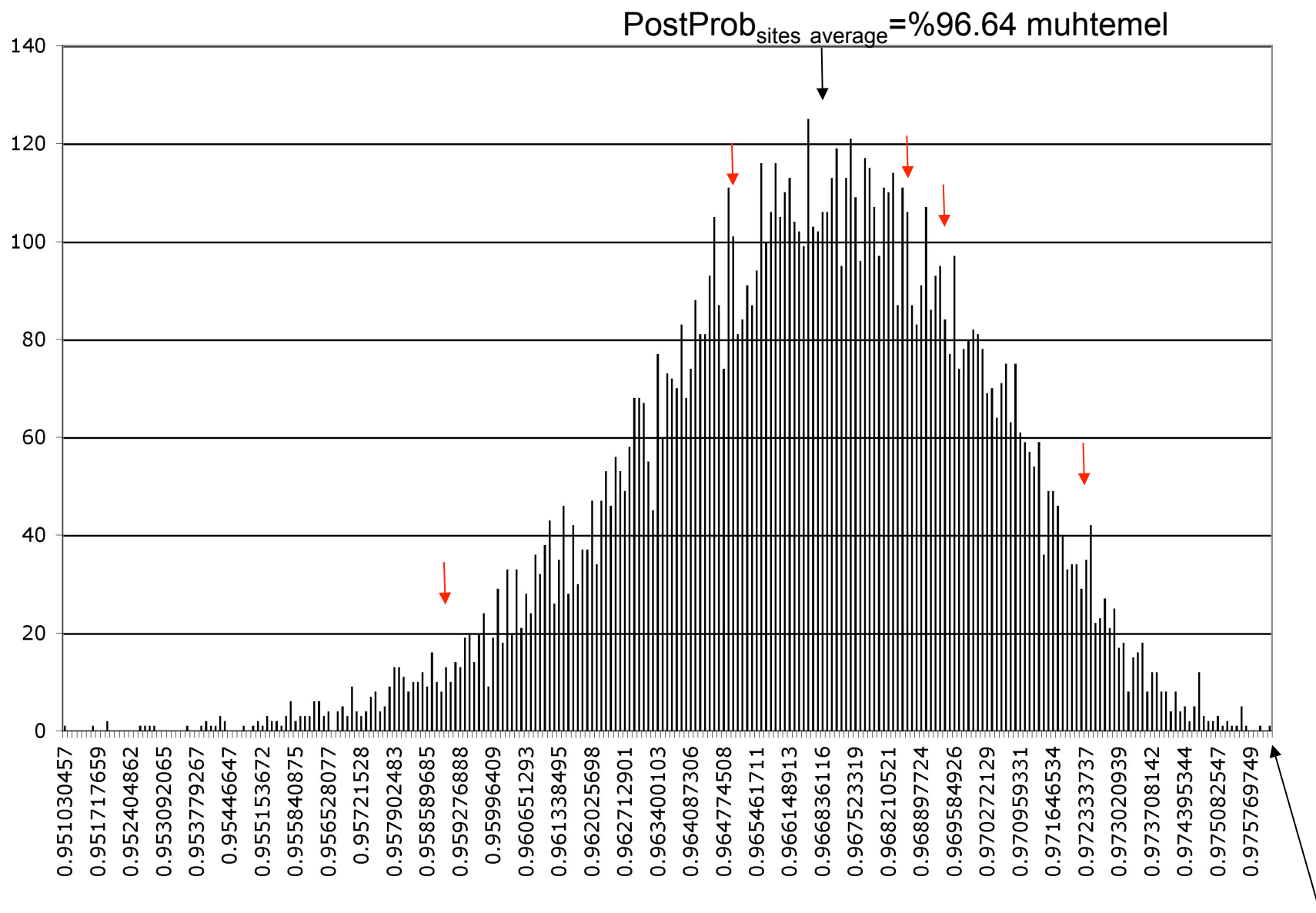
AAAAA

Weighted random-sample Posterior Distribution:

2 (AA A S A & A S A A A)

Elongasyon Faktörü (EF Tu):
Nature, 2003 & 2008





Most-probabilistic ancestral sequence (MPAS)
PostProb_{sites average}=%97.62 en muhtemel

MPAS vs. WRS

- Richard Goldstein and David Pollock:
MPAS ancestral sequence can be biased towards thermostability but that weighted-random samples from the posterior distribution can circumvent such bias)

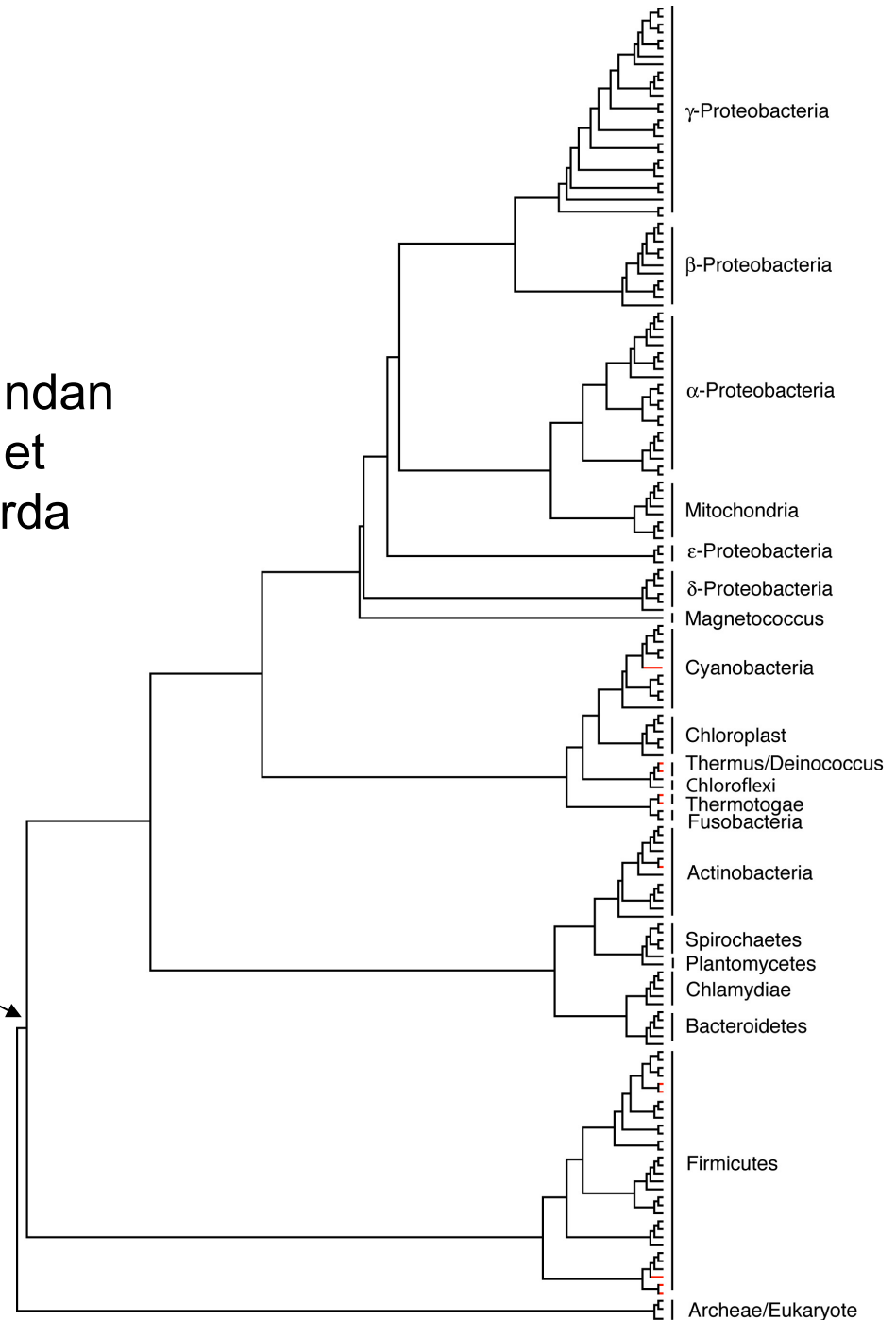
Elongation Factor – Tu *Nature*, 2003 & 2008

- 10,000 rastgele dizilim arasından Posterior dağılımını sample et
5 tane dizilim laboratuarda canlandırılıyor

- Analizler
 - 1) 66.3°
 - 2) 64.6°
 - 3) 62.2°
 - 4) 60.5°
 - 5) 60.5°

[7-18 AA farkı]

64.8°



Muhtemel Problemler / Kısıtlamalar

Atasal genotip ne kadar doğru? (Atasal genotipi bulmak mümkün mü?)

- En yaygın kullanılan yöntem Maximum Likelihood (PALM Modelleme)
- Zhao makalesi Maximum Likelihood ve Parsimony yöntemleri güvenilirliğini kıyaslıyor

Filogeni analizlerine ne kadar güvenebiliriz?

Filogeni simulasyonları:

|
ATCG

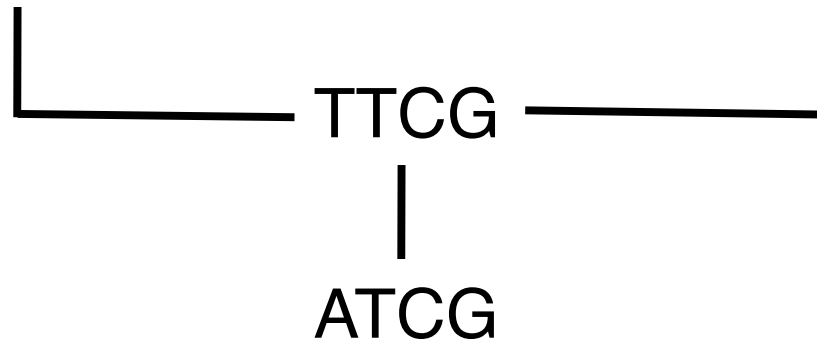
Filogeni analizlerine ne kadar güvenebiliriz?

Filogeni simulasyonları:

TTCG
|
ATCG

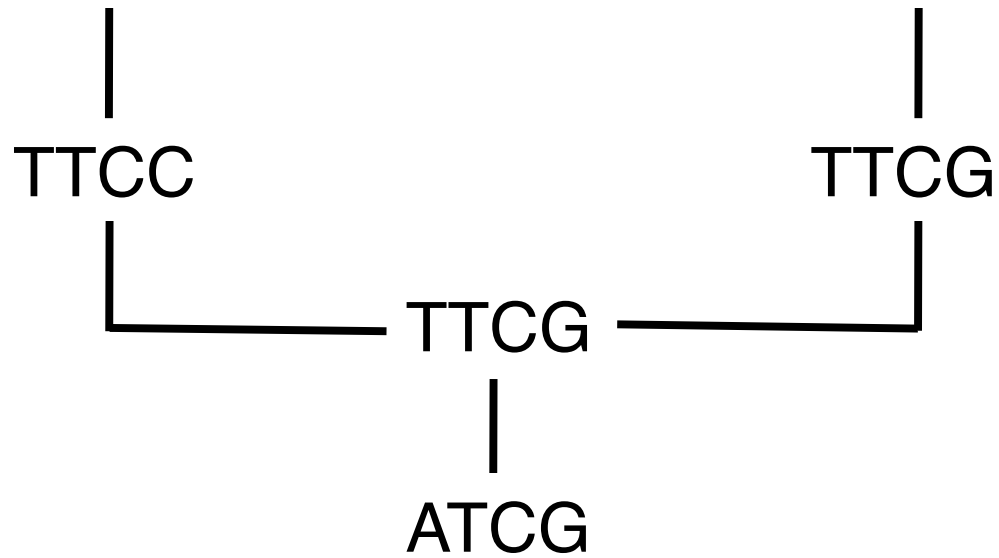
Filogeni analizlerine ne kadar güvenebiliriz?

Filogeni simülasyonları:



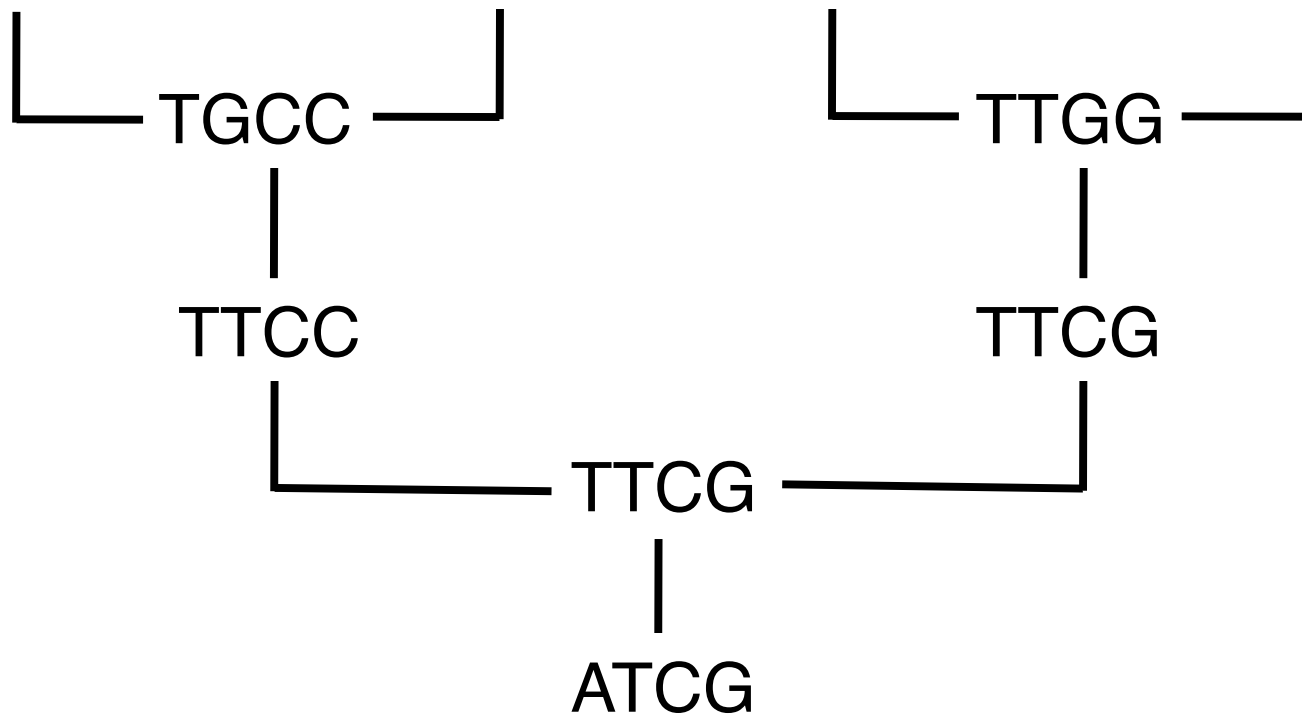
Filogeni analizlerine ne kadar güvenebiliriz?

Filogeni simulasyonları:



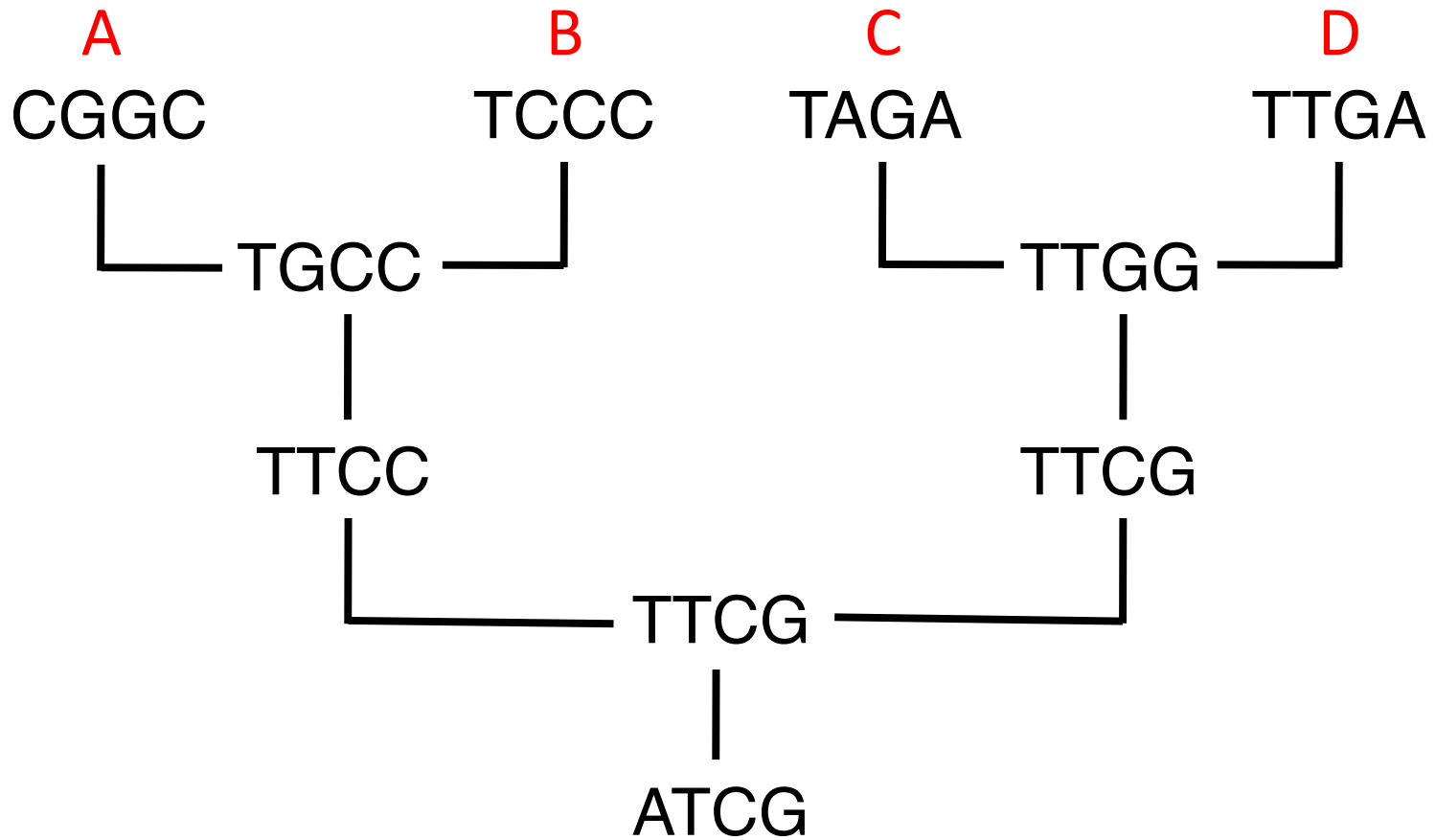
Filogeni analizlerine ne kadar güvenebiliriz?

Filogeni simulasyonları:



Filogeni analizlerine ne kadar güvenebiliriz?

Filogeni simulasyonları:

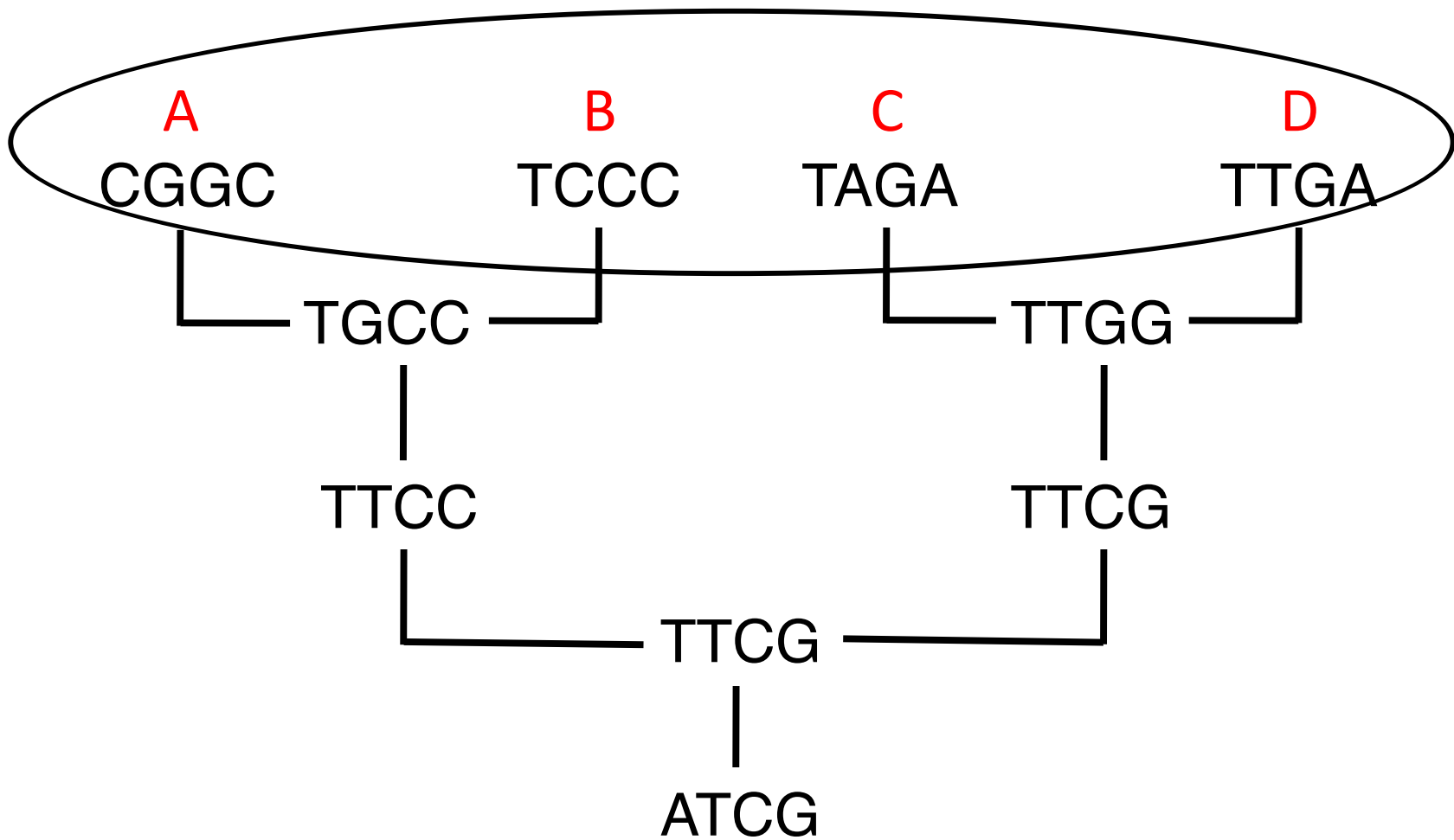


Filogeni analizlerine ne kadar güvenebiliriz?

Filogeni simülasyonları:

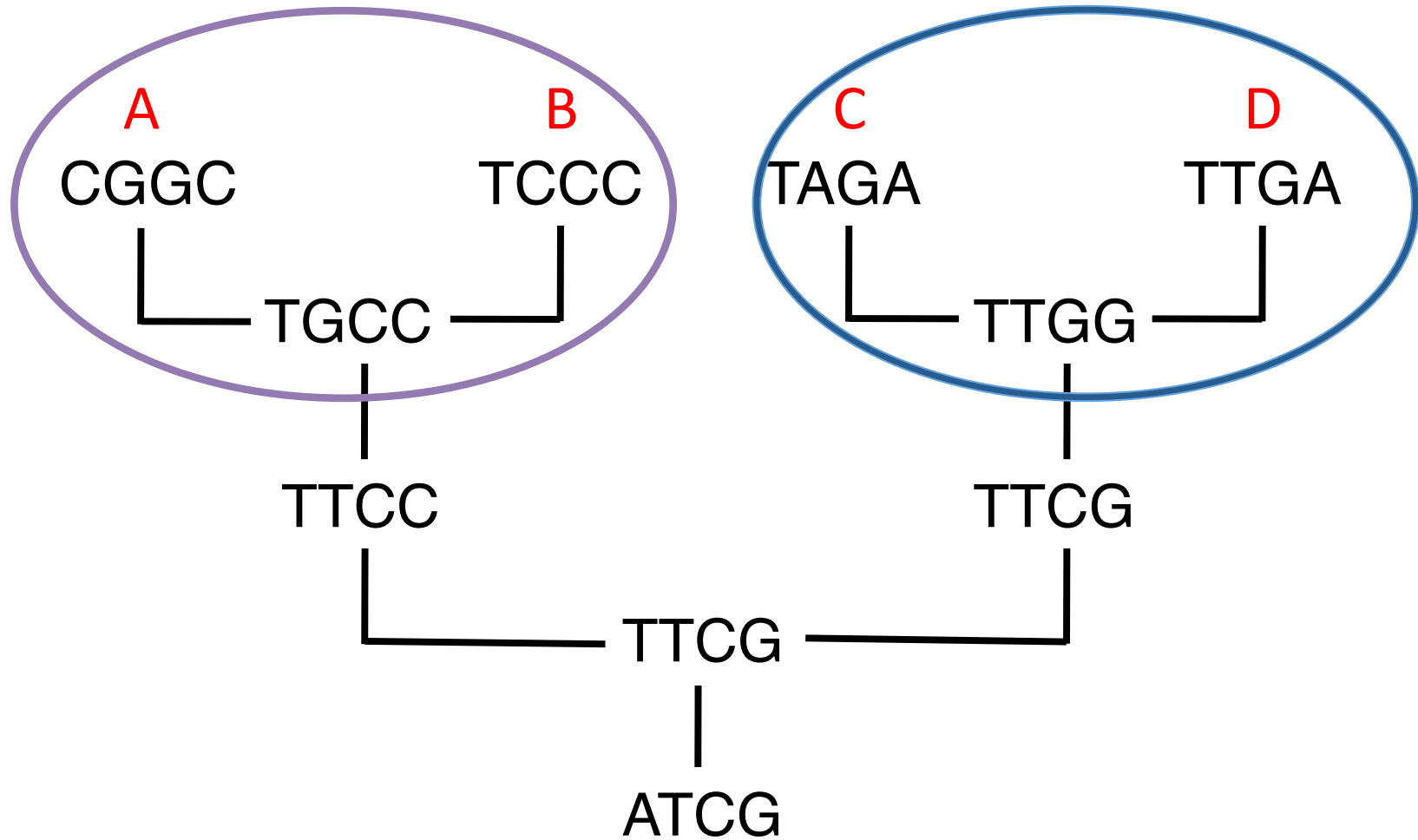
Tip sekansları

(simülasyon sonucu inferred ağaç tipi ile uyumlu mu?)



Filogeni analizlerine ne kadar güvenebiliriz?

Filogeni simulasyonları:



Filogeni analizlerine ne kadar güvenebiliriz?

Filogeni simülasyonları?

Deneysel çözüm:

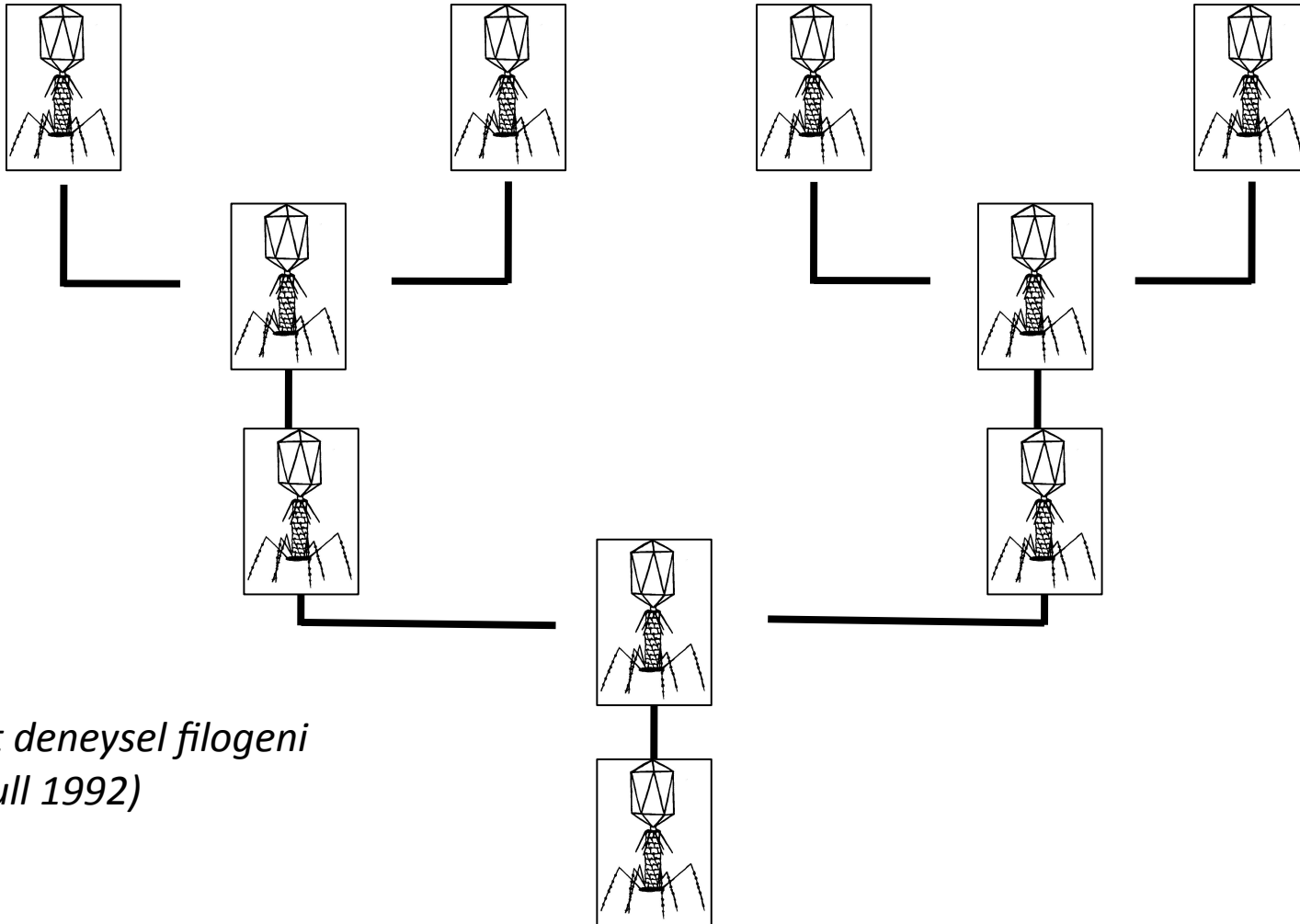
Laboratuarda deneysel olarak filogeni oluşturup simülasyon ile karşılaştırma

Filogeni analizlerine ne kadar güvenebiliriz?

Filogeni simülasyonları?

Deneysel çözüm:

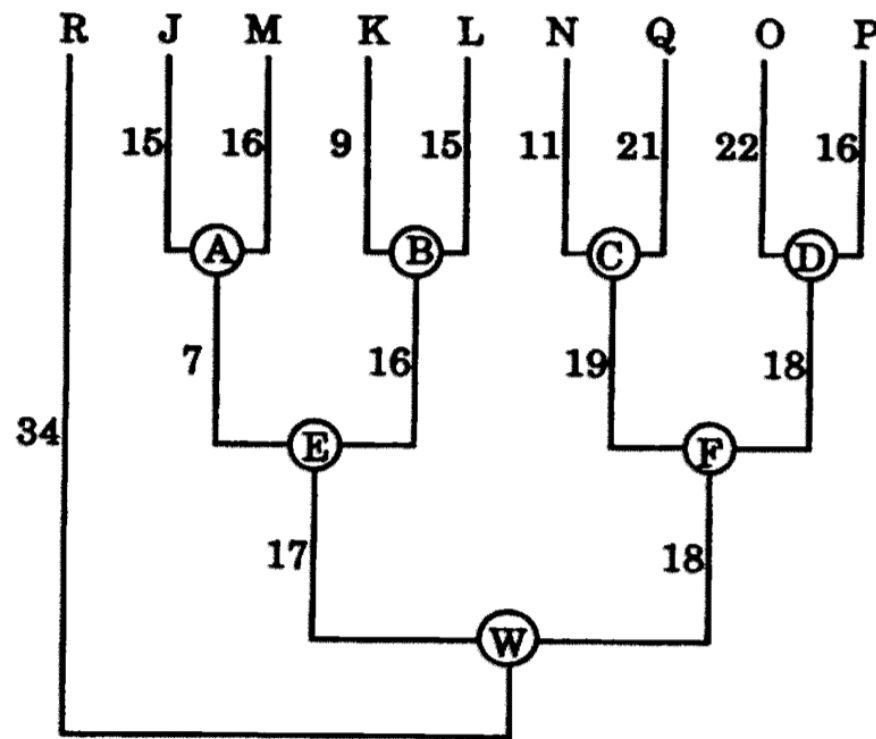
Laboratuarda deneysel olarak filogeni oluşturup simülasyon ile karşılaştırma



*T7 faj'a ait deneysel filogeni
(Hillis & Bull 1992)*

Experimental Phylogenetics: Generation of a Known Phylogeny

DAVID M. HILLIS, JAMES J. BULL, MARY E. WHITE,
MARTY R. BADGETT, IAN J. MOLINEUX

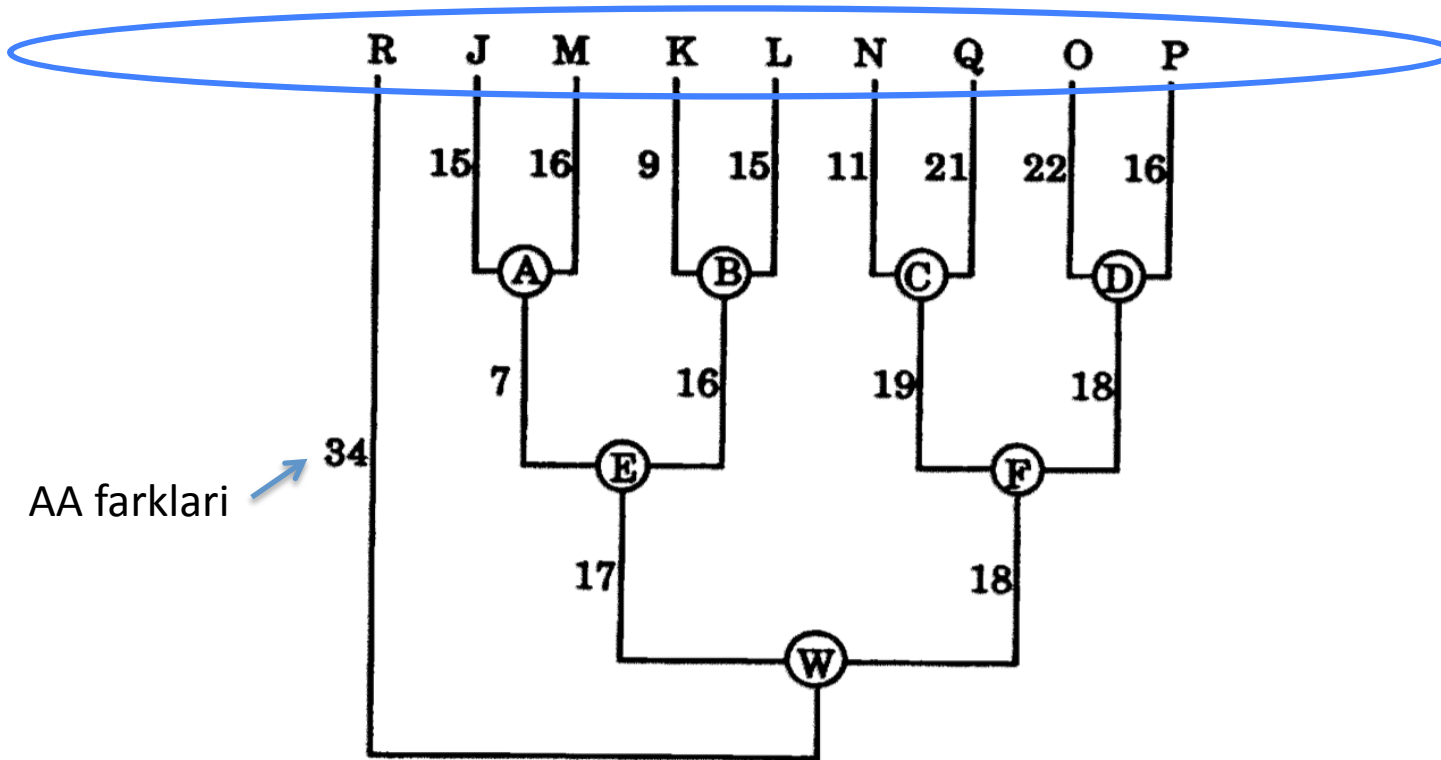


Experimental Phylogenetics: Generation of a Known Phylogeny

Science

AAAS

DAVID M. HILLIS, JAMES J. BULL, MARY E. WHITE,
MARTY R. BADGETT, IAN J. MOLINEUX



Modern gen	Canlandırılan Ata	Atanın yaklaşık yaşı	ML yada MP?
Ribonükleaz	Bufaloya ait ortolog	5-10 milyon	MP
Retrotranspozon	Fareye ait paralog	Bir kac milyon (tam yaş belirtilmemiş)	Konsensüs
Proteaz	Memelilerin atası	80 milyon	MP
RNAaz	Primatların atası	31 milyon	MP
Hormon reseptörleri	Protostom atası	1000 milyon	ML
Elongasyon Faktörleri	Öbakteri atası	3600 milyon	ML
Rodopsin		240 milyon	ML

ML Maximum Likelihood
MP Maximum Parsimony

Muhtemel Problemler / Kısıtlamalar

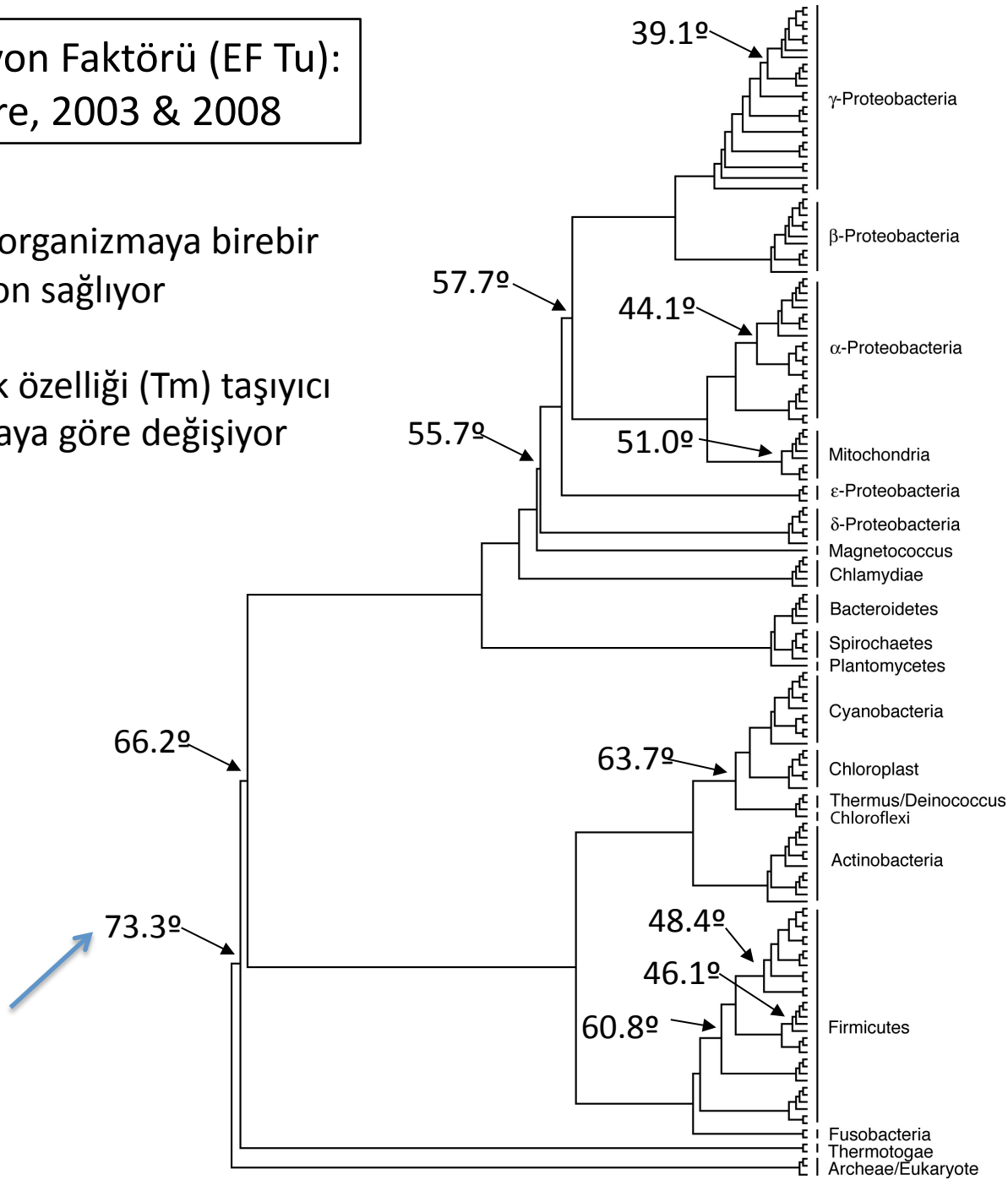
Atasal genotip ne kadar doğru? (Atasal genotipi bulmak mümkün mü?)

Diğer sorunlar:

- ‘Molekuler clock’ problemi: Hedges topolojisi
- Evrimsel substitution rate constant kabul edilmesi & sadece simulasyonlar ile dizilimi test etmek ne kadar gerçekçi
- Deneysel yöntemlerle bu simulasyonlar ve evrimsel modeller test edilmeli
- Her protein canlandırılabilir mi? (functional divergence – adaptive protein olması önemli)

Elongasyon Faktörü (EF Tu): Nature, 2003 & 2008

- Yaşadığı organizmaya birebir adaptasyon sağlıyor
- Fenotipik özelliği (T_m) taşıyıcı organizmaya göre değişiyor



Antik termofilik hayat üzerine örnekler

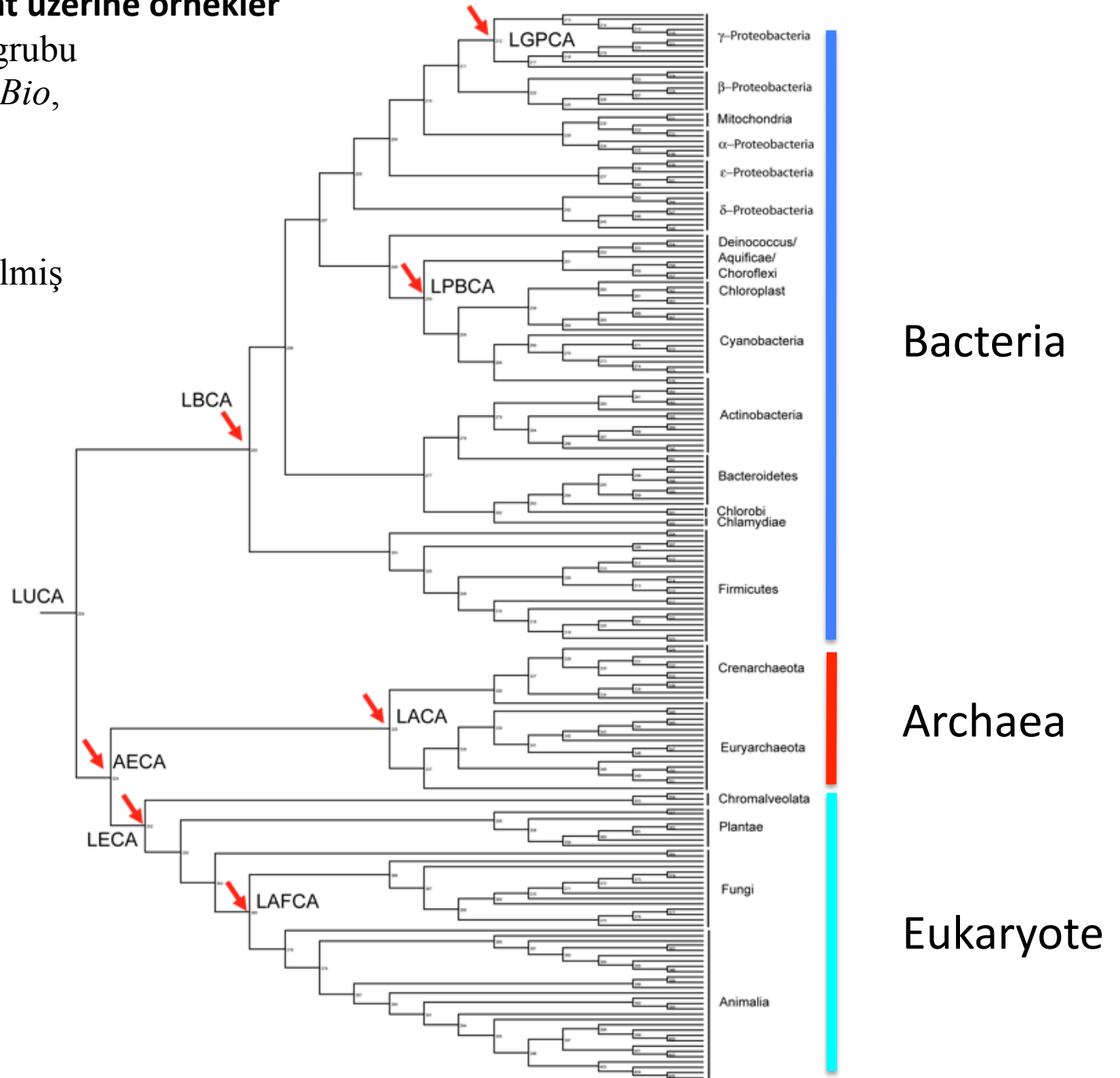
Thioreduktaz enzim grubu

Nature Struct & Mol Bio,

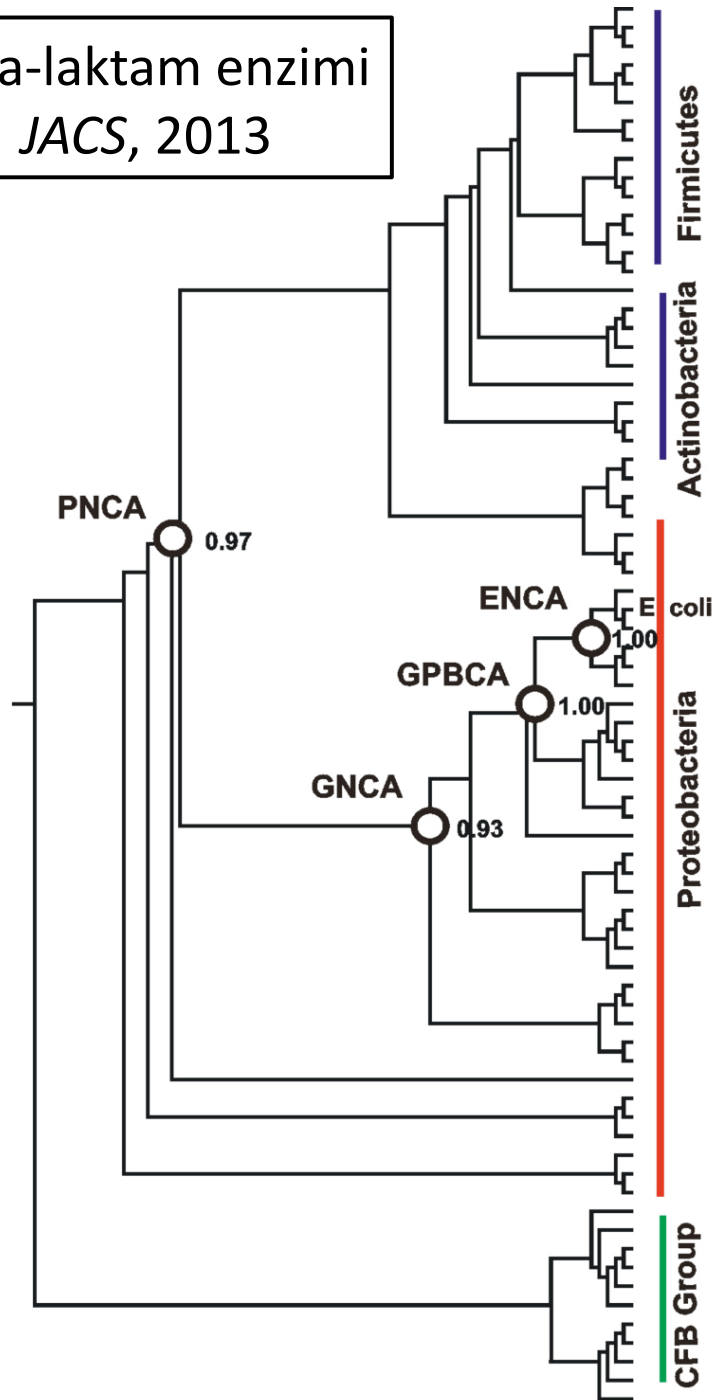
2009 & 2011

>200 dizilim

>3 ana grup dahil edilmiş



Beta-laktam enzimi
JACS, 2013



Antik termofilik hayat üzerine örnekler

75 sekans

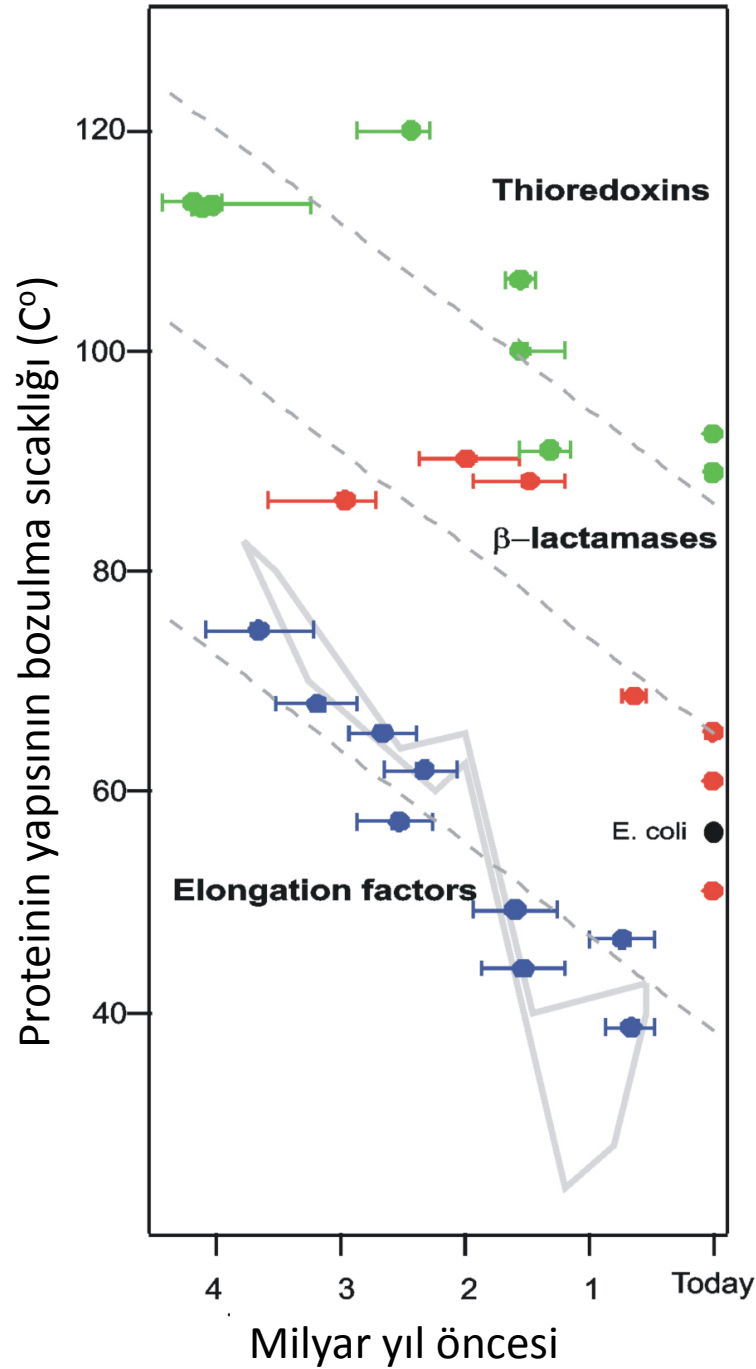
ENCA=Enterobakteri

GPBCA=Gamma Proteobakteri

GNCA=Gram Negatif

PNCA=Gram +/-

..bakterilerinin son ortak atası
canlandırılmış



3 farklı ve birbirine fonksiyonel olarak uzak protein grubu benzer bir davranış gösteriyor.

Her protein aynı davranışı gösterebilir mi? Hayır.



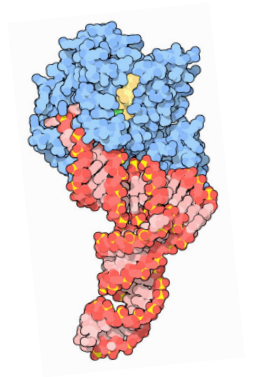
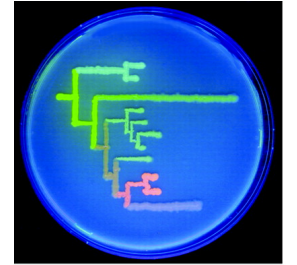
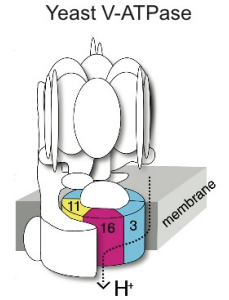
Antik hava şartları:
Sıcak ve Asidik

Neler öğrendik?

Hayatın geçmişteki ve günümüzdeki evrimsel koşulları

Antik termofili hayat üzerine örnekler

- Kinaz enzimleri (Akanuma 2013, PNAS)
- Redüktif enzimler (Perez-Jimenez, 2012, NSB) (pH ve Sıcaklık)
- Elongasyon faktörleri (Gaucher 2003/8, Nature)
- Kompleks hayatın evrimi (Finnigan 2012, Nature)
- In-vitro moleküler evrim (Ugande 2004, Science)
- Alkol dehidrojenaz enziminin geçmişi (Thomson 2005, Nature)
- Antik metabolizmalar (Voordeckers 2012, PLoS)



Neler öğrendik?

Hayatın geçmişteki ve günümüzdeki evrimsel koşulları

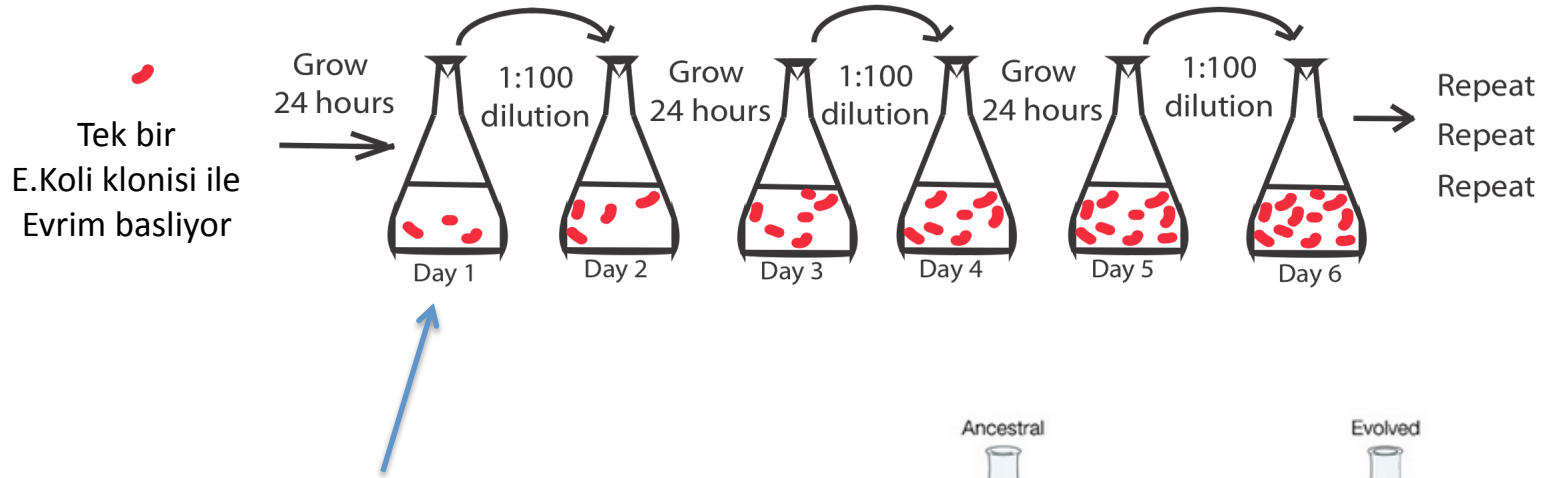
- Antik hormon reseptörleri
 - Aralarında 200 milyon yıl yaş farkı olan iki hormon reseptörü canlandırılıyor
 - Amino asit farkları belirleniyor
- Amino asitleri değiştirerek fonksiyonel özelliklere yön verebilir miyiz?
 - Genç olan protein eskisine dönüşemiyor (epistatik etkileşim)
- Evrimsel yönelimler tek yönlü mü?
 - (one way ticket to the functional specificity?)

Bölüm II: Deneysel Evrim

Deneysel Evrim (Başlıca örnekler)

1. Lenski uzun-vadeli evrim deneyi (tarihçe ve amaçlar)
2. Bakteri büyüme ortamı ve fitness (~~seçim değeri~~ yeti) hesaplamaları
3. Faydalı mutasyonları belirleme
4. Mutasyon seçim değerleri (değişkenler)
5. Contingency

Deneysel Evrim (Seri transfer) - (Rekabet)

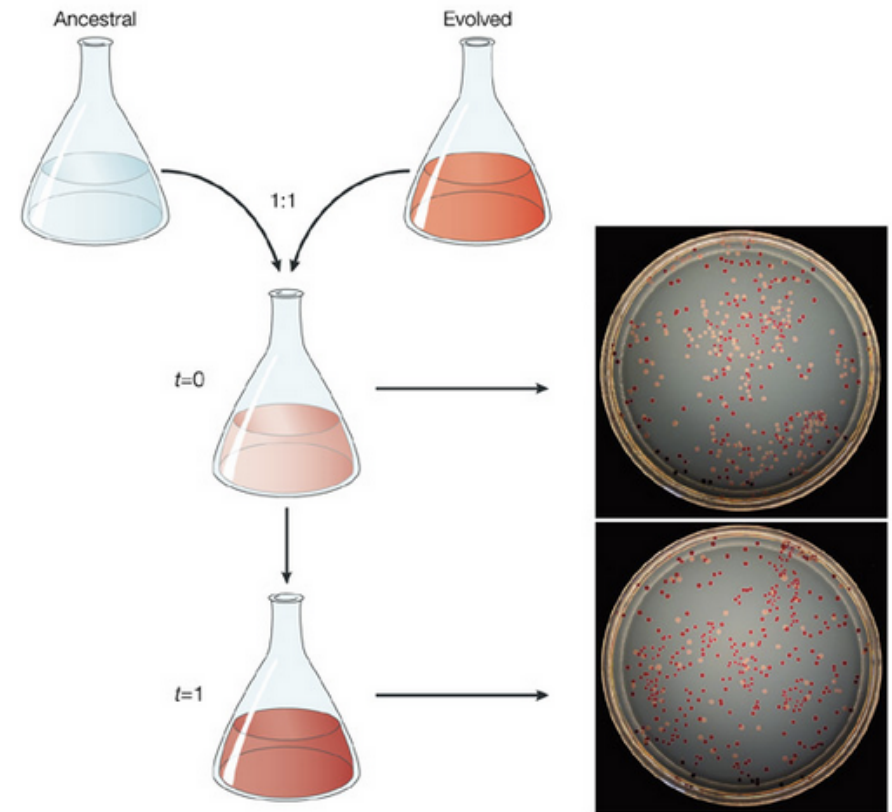


Degiskenler:

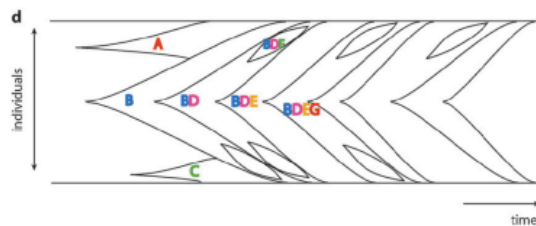
Buyume ortami (minimal, zengin vs.)

Sicaklik & basinc

Oksijen miktarı



Adaptation



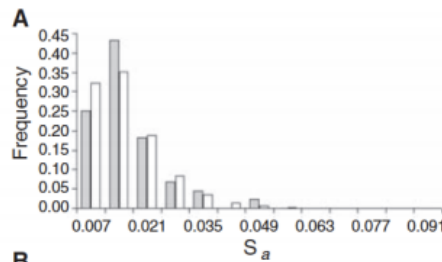
Dynamics of adaptation:

Population structure

Clonal interference

Multiple mutations

Mutation and recombination rates?

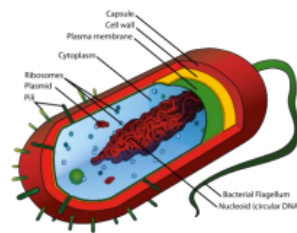


Fitness landscape:

~~Mutation rates~~

Distribution of fitness effects

Epistasis



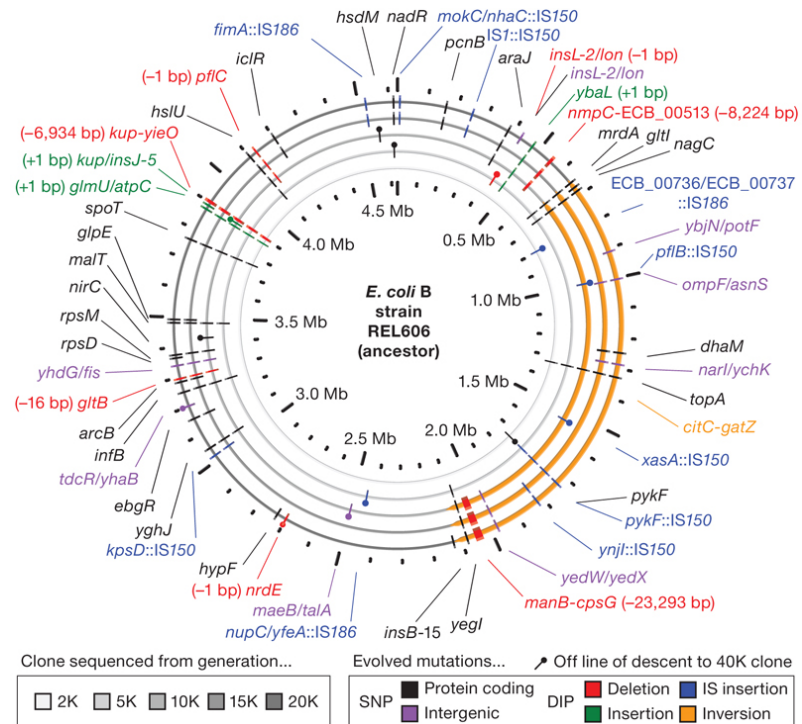
Environment + cellular

architecture:

Regulatory networks

Proteins

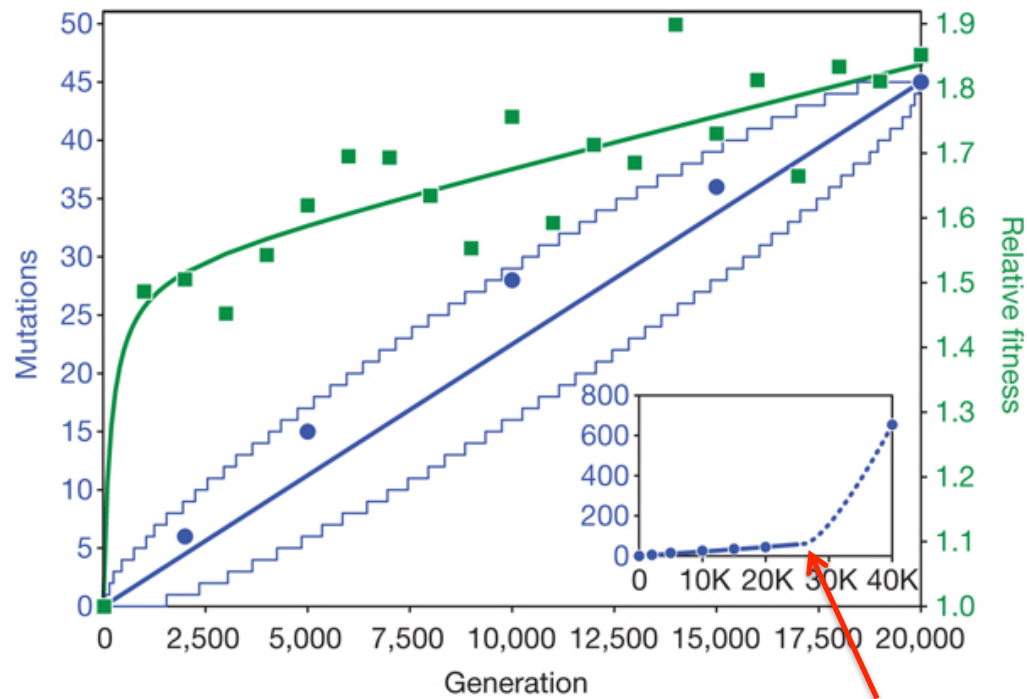
Mutations found by sequencing genomes sampled between 2,000 and 20,000 generations from an evolution experiment with *E. coli*.



JE Barrick *et al.* *Nature* **000**, 1-5 (2009) doi:10.1038/nature08480

nature

Rates of genomic evolution and fitness improvement.



JE Barrick *et al.* *Nature* **000**, 1-5 (2009) doi:10.1038/nature08480

nature

20,000. Nesil koloni genomu

Gene or region	Fitness effect (%)	Significance	Source
<i>topA</i>	13.3	***	Ref. 32
<i>pykF</i> [*]	11.1	***	D.S., R.E.L., unpublished
<i>spoT</i>	9.4	***	Ref. 31
<i>nadR</i> [†]	8.1	***	D.S., R.E.L., unpublished
<i>glmU</i> promoter	4.9	***	M. Stanek, T. Cooper, R.E.L., unpublished
<i>fis</i>	2.9	***	Ref. 32
<i>rbs</i> operon [†]	2.1	***	Ref. 43
<i>malT</i>	0.4	**	Ref. 44
<i>ompF</i> [‡]	-9.7	**	D.S., R.E.L., unpublished

20,000. Nesil koloni genomu

Gene or region	Function	Parallel mutations (%)	Source
<i>nadR</i>	Transcriptional regulator	100	Ref. 42
<i>pykF</i>	Pyruvate kinase	100	Ref. 42
<i>rbs</i> operon	Ribose catabolism	100	Ref. 43
<i>malT</i>	Transcriptional regulator	64	Ref. 44
<i>spoT</i>	Stringent response regulator	64	Ref. 31
<i>mrda</i>	Cell-wall biosynthesis	45	Ref. 42
<i>infB</i>	Translation initiation factor 2	45*	This study
<i>fis</i>	Nucleoid-associated protein	27	E. Crozat, D.S., unpublished
<i>topA</i>	DNA topoisomerase I	27	E. Crozat, D.S., unpublished
<i>pcnB</i>	Poly(A) polymerase	27	This study
<i>ompF</i>	Outer-membrane porin	18*	This study
<i>rpsD</i>	30S ribosomal protein	18*	This study
<i>rpsM</i>	30S ribosomal protein	0	This study
<i>glmU</i> promoter	Cell-wall biosynthesis	0	M. Stanek, R.E.L., unpublished
*In addition to populations with substitutions, one or more others were polymorphic.			