SAITOU
LABORATORY
NIG SOKENDAI U.TOKYO / MISHIMA

Introduction to Evolutionary Genomics

Handout Part 1

Saitou Naruya

National Institute of Genetics, Mishima

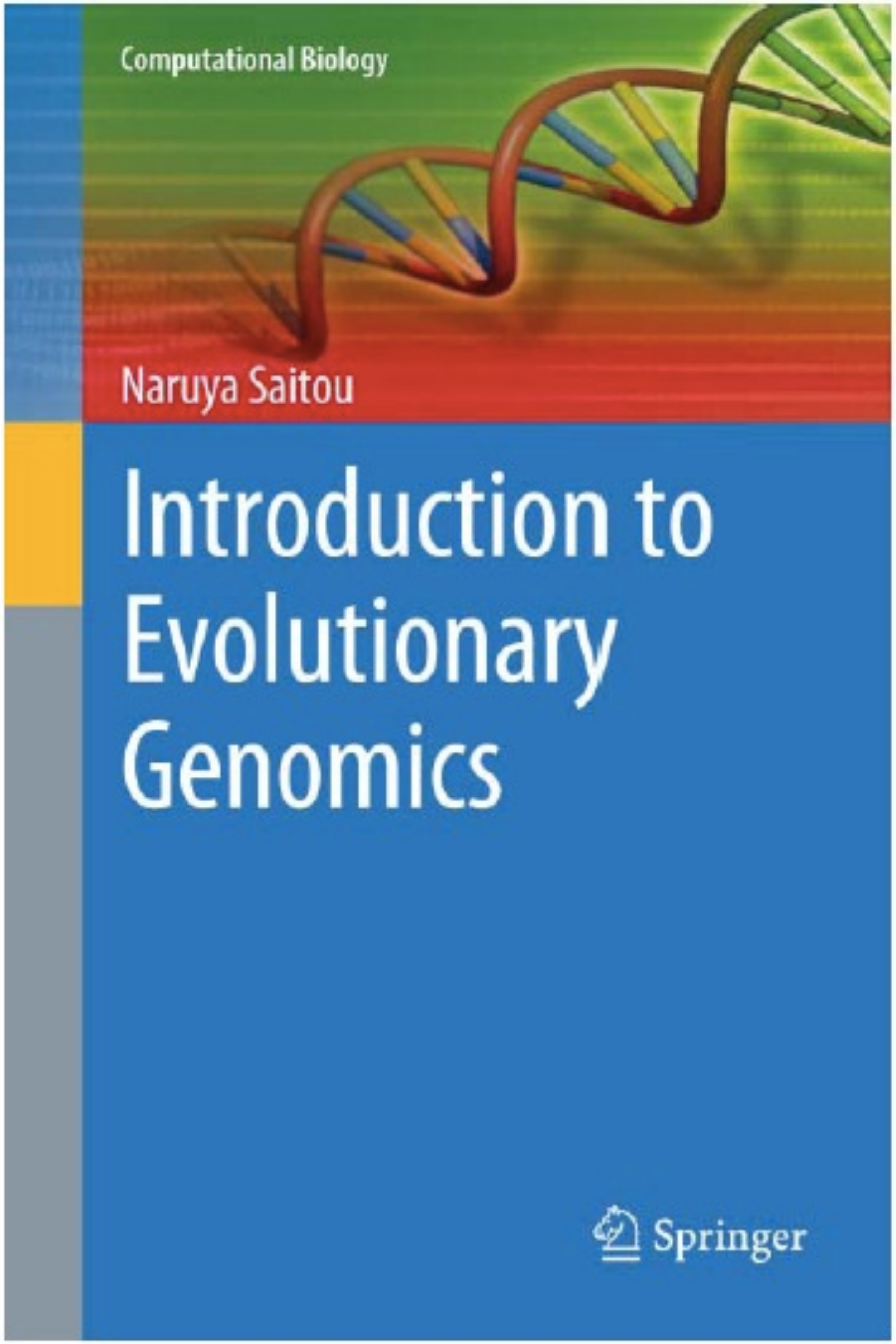
Dept. Genetics, Graduate University for Advanced Studies

Dept. Biological Sciences, University of Tokyo

saitounr@nig.ac.jp

http://www.saitou-naruya-laboratory.org/english

June 18, A.S. 0013
Suzhou, China



Computational Biology

Naruya Saitou

Introduction to Evolutionary Genomics

Springer

To be published from Springer-Verlag (Fall, 2013)

Introduction to Evolutionary Genomics

Part 1 Basic Processes of Genome Evolution

Part 2 Evolving Genomes

Part 3 Methods for Evolutionary Genomics

Part 1 Basic Processes of Genome Evolution

Chapter 1 Replication, Transcription, and Translation

Chapter 2 Mutation

Chapter 3 Phylogeny

Chapter 4 Neutral Evolution

Chapter 5 Natural Selection

Chapter 3 Phylogeny

3.1 DNA replications generate phylogenies
3.2 Genealogy of individuals
3.3 Gene genealogy
3.4 Species phylogeny
3.5 Basic concepts of trees and networks
3.6 Biological nature of trees and networks

Figure 3-2: A schematic representation of the phylogenetic tree of 10bp DNA sequences

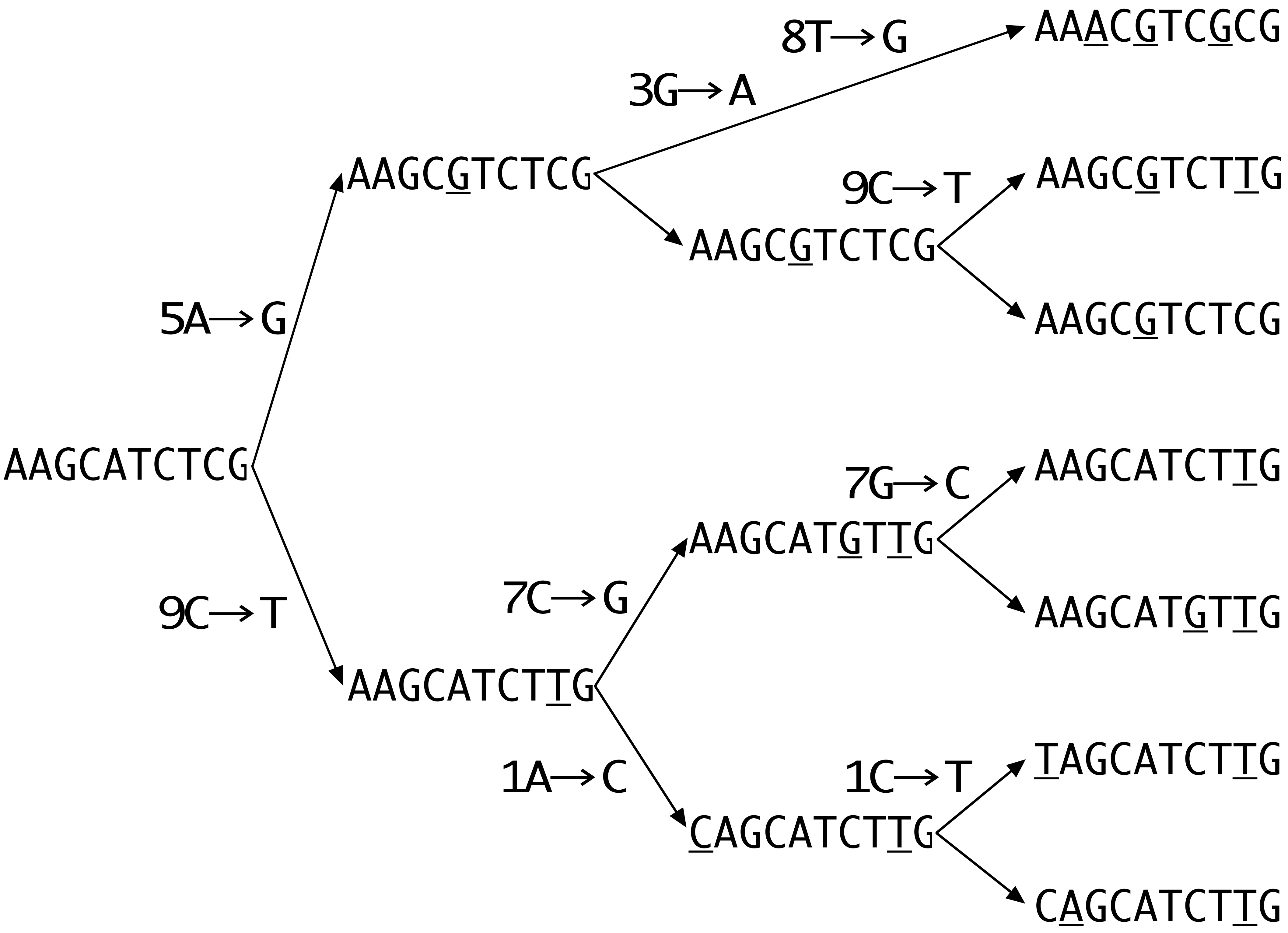
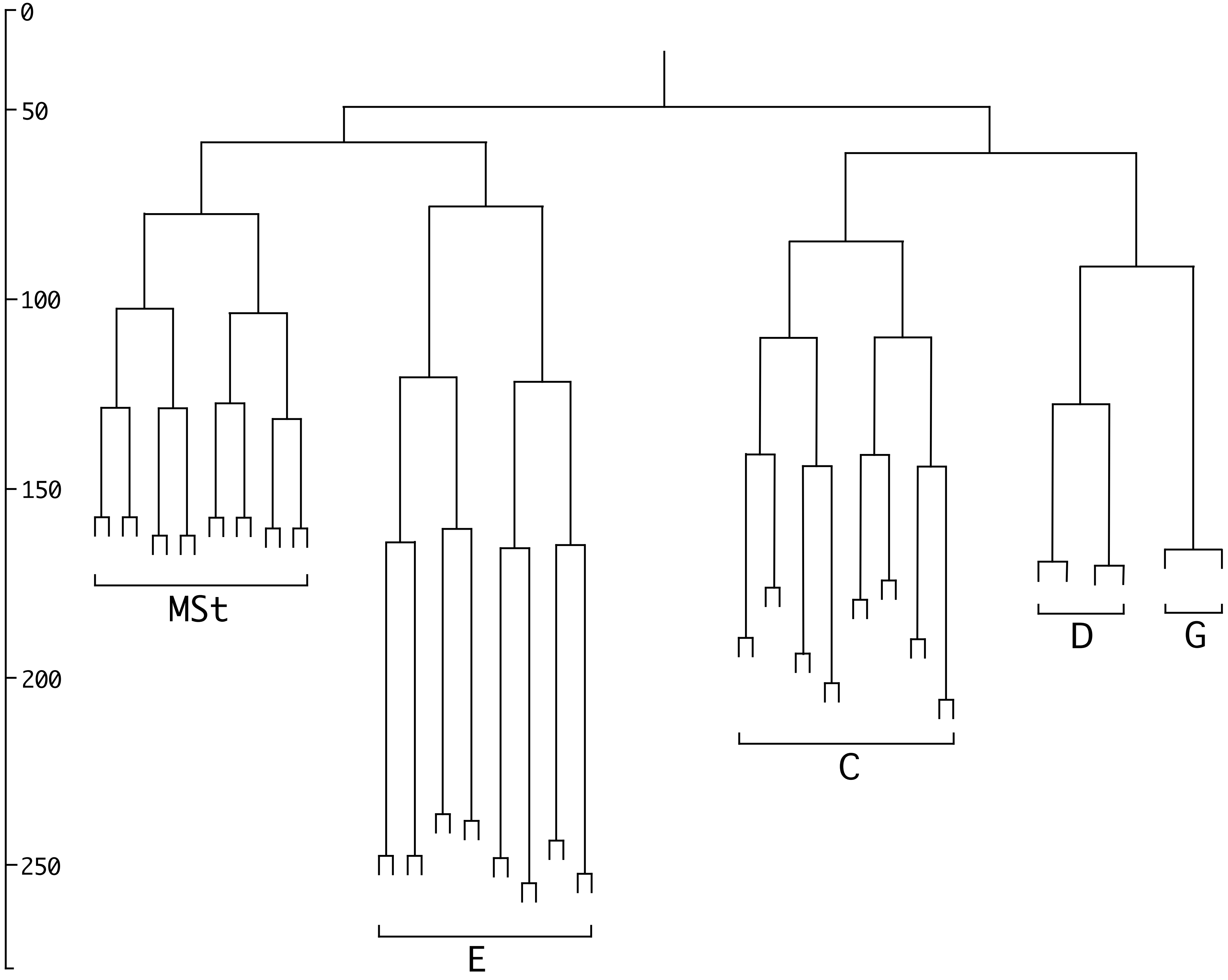
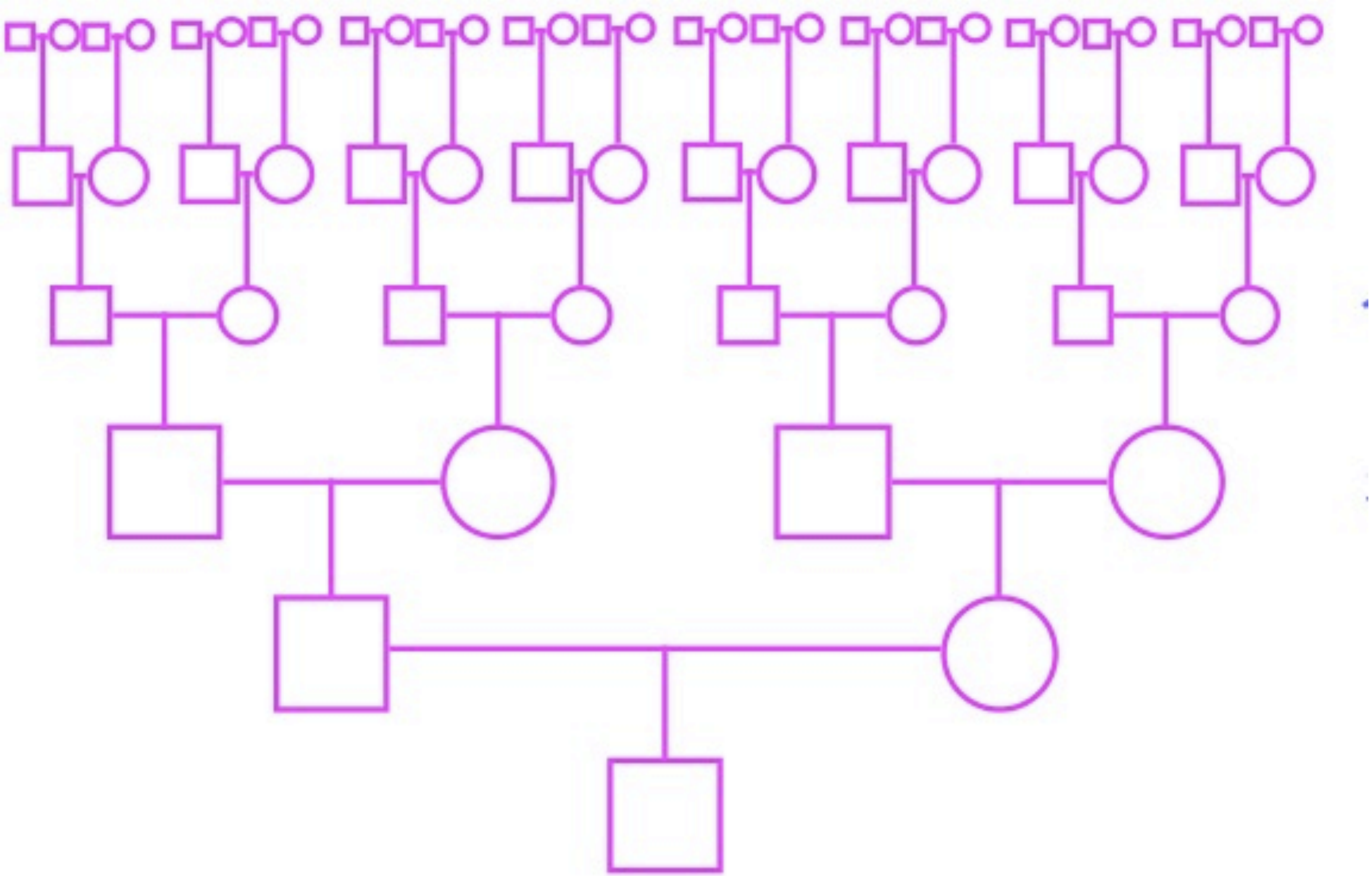


Figure 3-4: Cell genealogy of *C. elegans* (based on ref 3-27)



Individual genealogy for diploid



Individual in question

Figure 3-6: A phylogenetic tree of human individuals (based on ref 3-4)

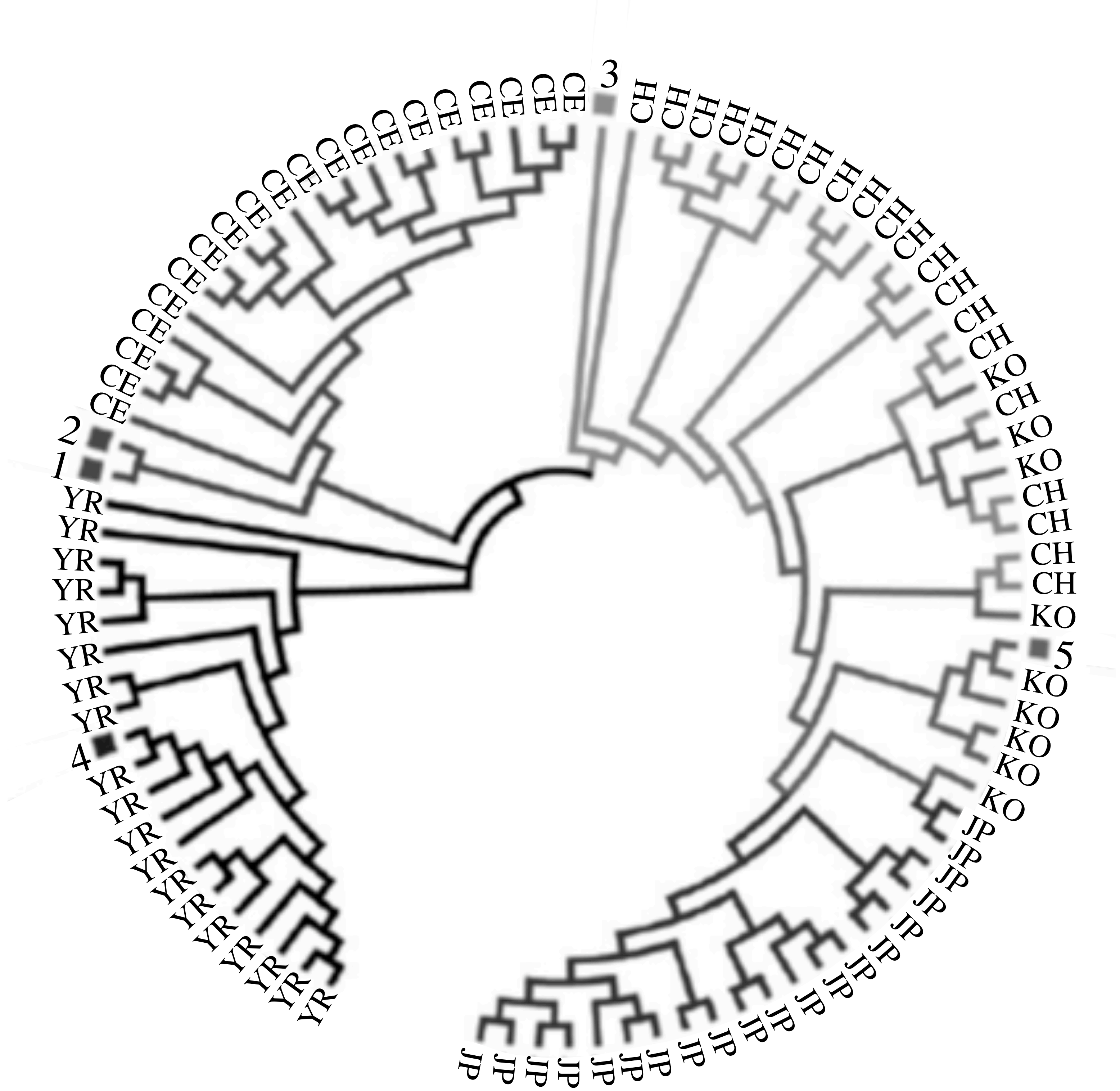


Figure 3-10: Alternative gene genealogies for 4 autosomal genes of two diploid individuals

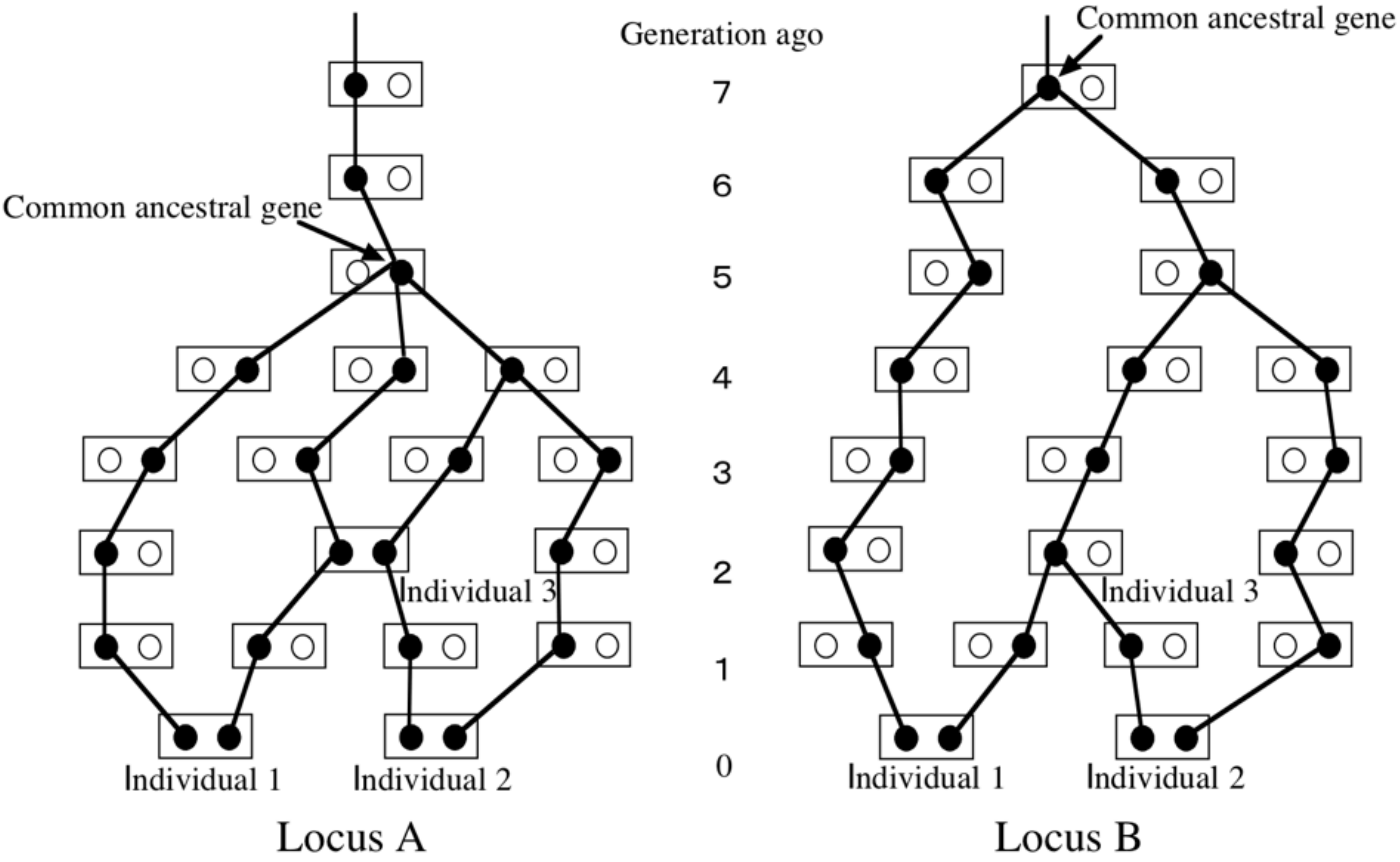


Figure 3-7: Gene genealogy for haploids.
(A) Animal mitochondrial DNA. (B) Mammalian Y chromosomes (C) Avean W chromosomes.

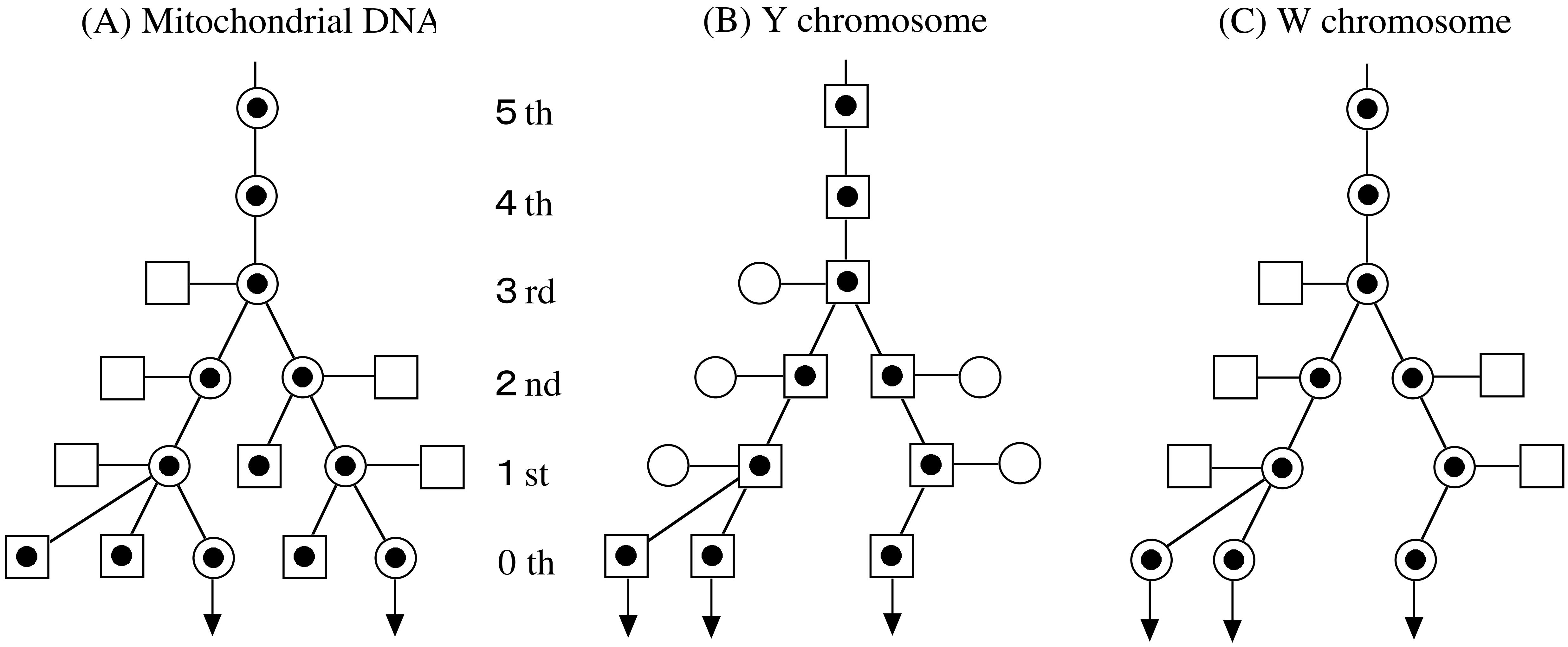


Figure 3-9: Human mitochondrial DNA genome trees for 53 individuals (based on ref 3-8)

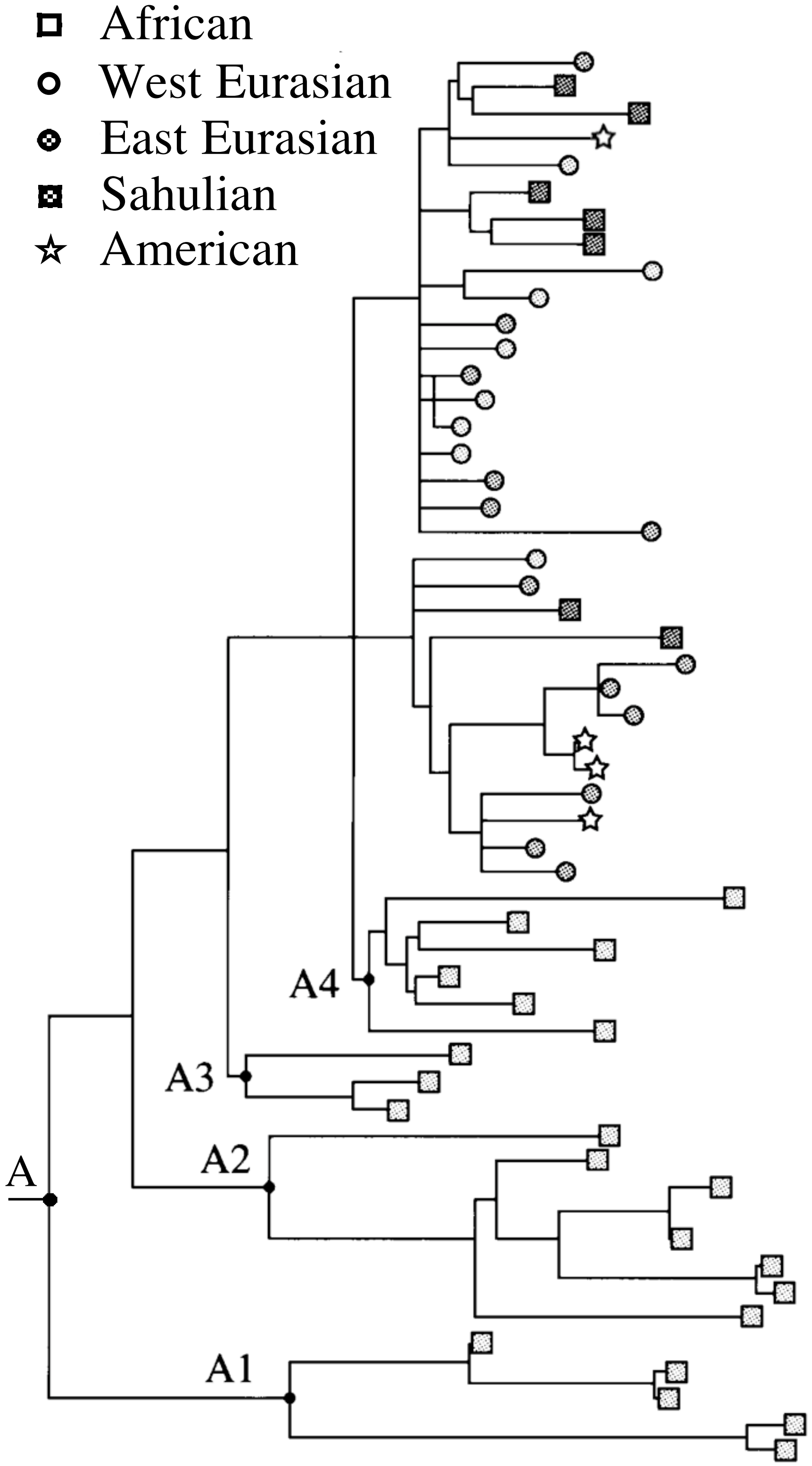


Figure 3-8: Contributions of maternal and paternal lineages of mitochondrial DNA and Y chromosomes

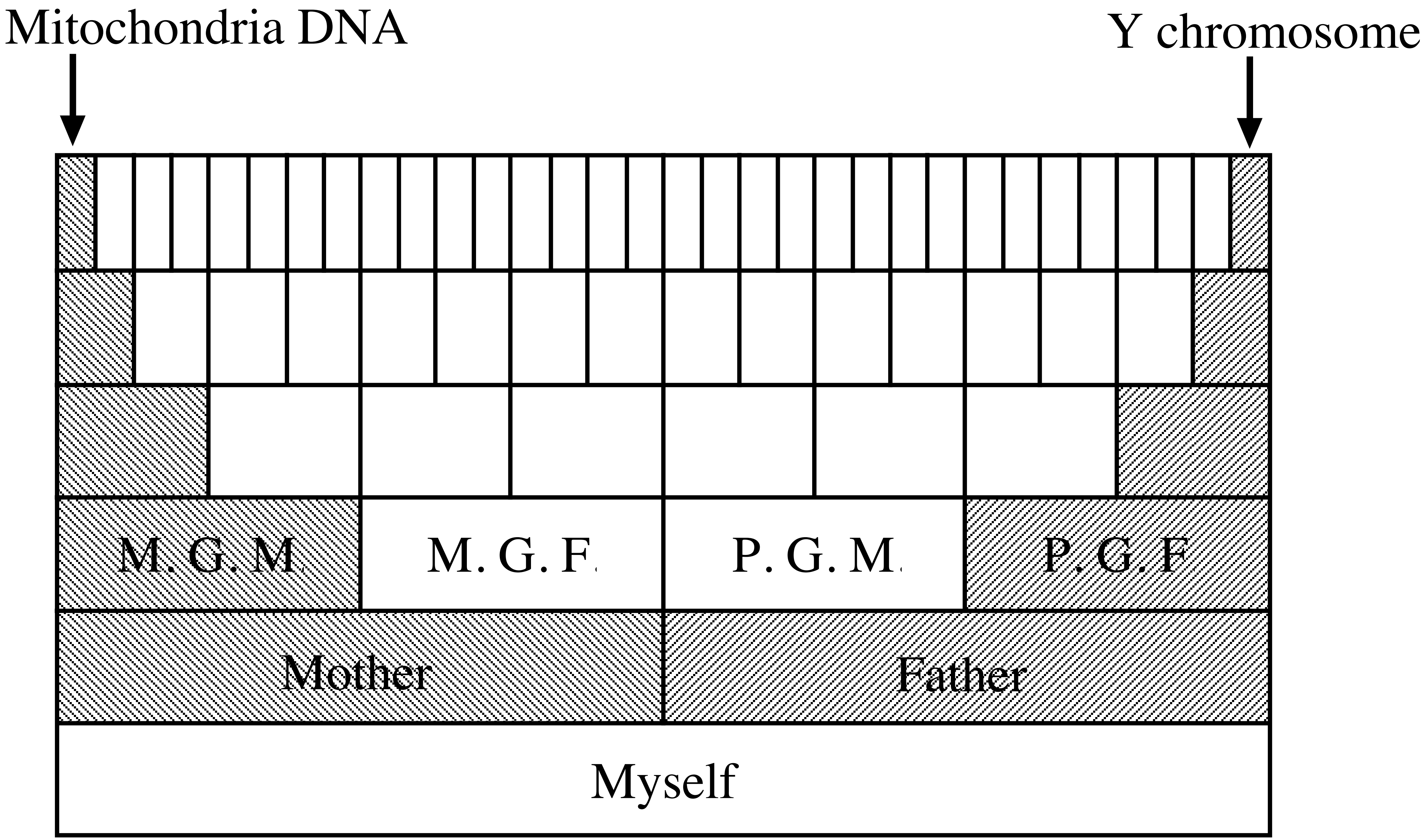


Figure 3-11: Gene phylogeny with recombinations for gibbon ABO blood group genes (from ref 3-9)

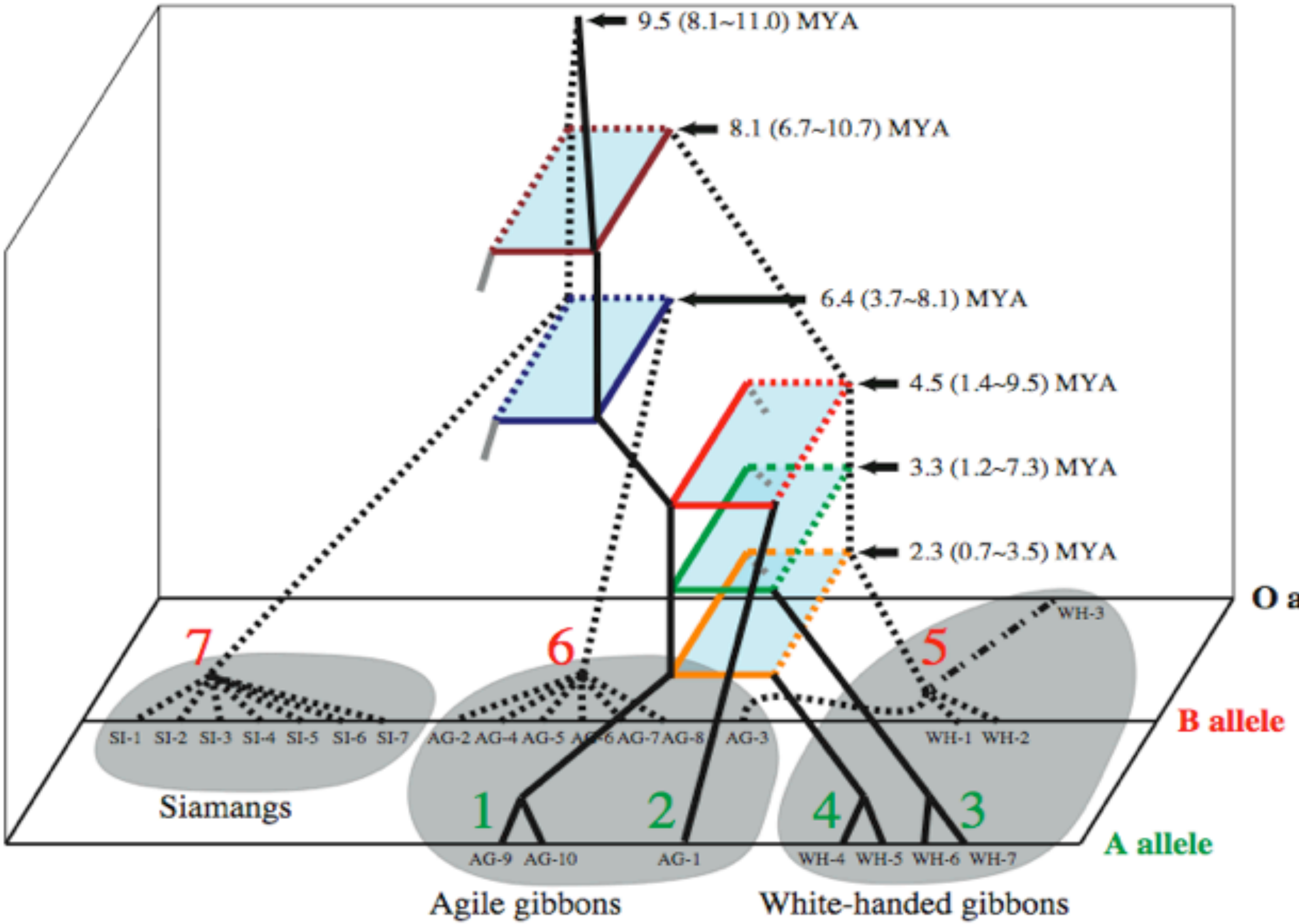


Fig. 3 (Kitano et al.)

From Kitano et al. (2010)

Figure 3-12: Three kinds of gene phylogenies. (A) Temporal and mutational gene phylogeny. (B) Mutational gene phylogeny. (C) Estimated gene phylogeny (from ref 3-13).

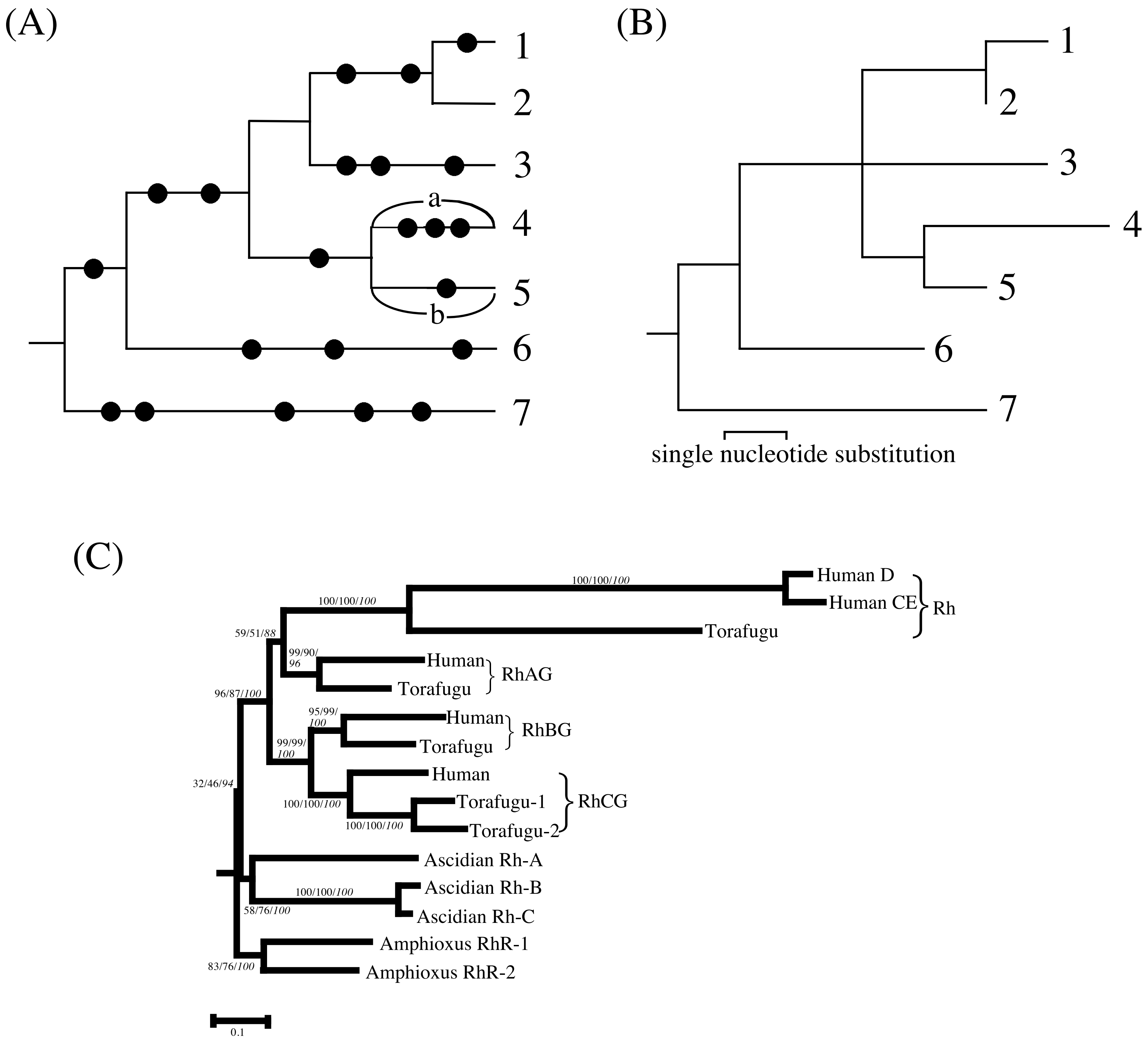


Figure 3-13: A linialized gene tree of Figure 3-12C (from ref 3-13)

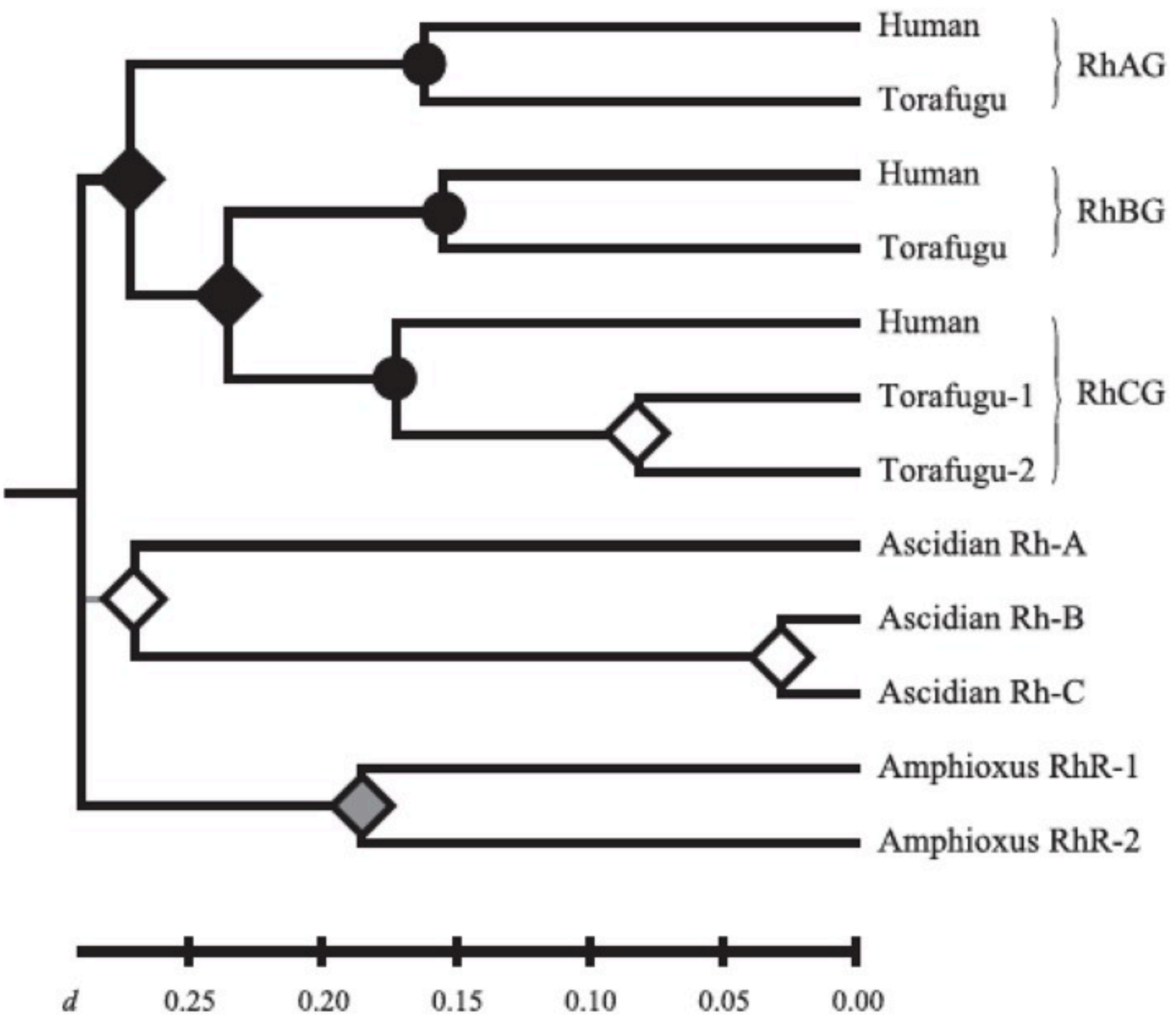


Figure 3-14: Two possible trees regarding gene duplications
(A) Gene duplication was misinterpreted as speciation. (B) Two gene duplication occurred after speciation

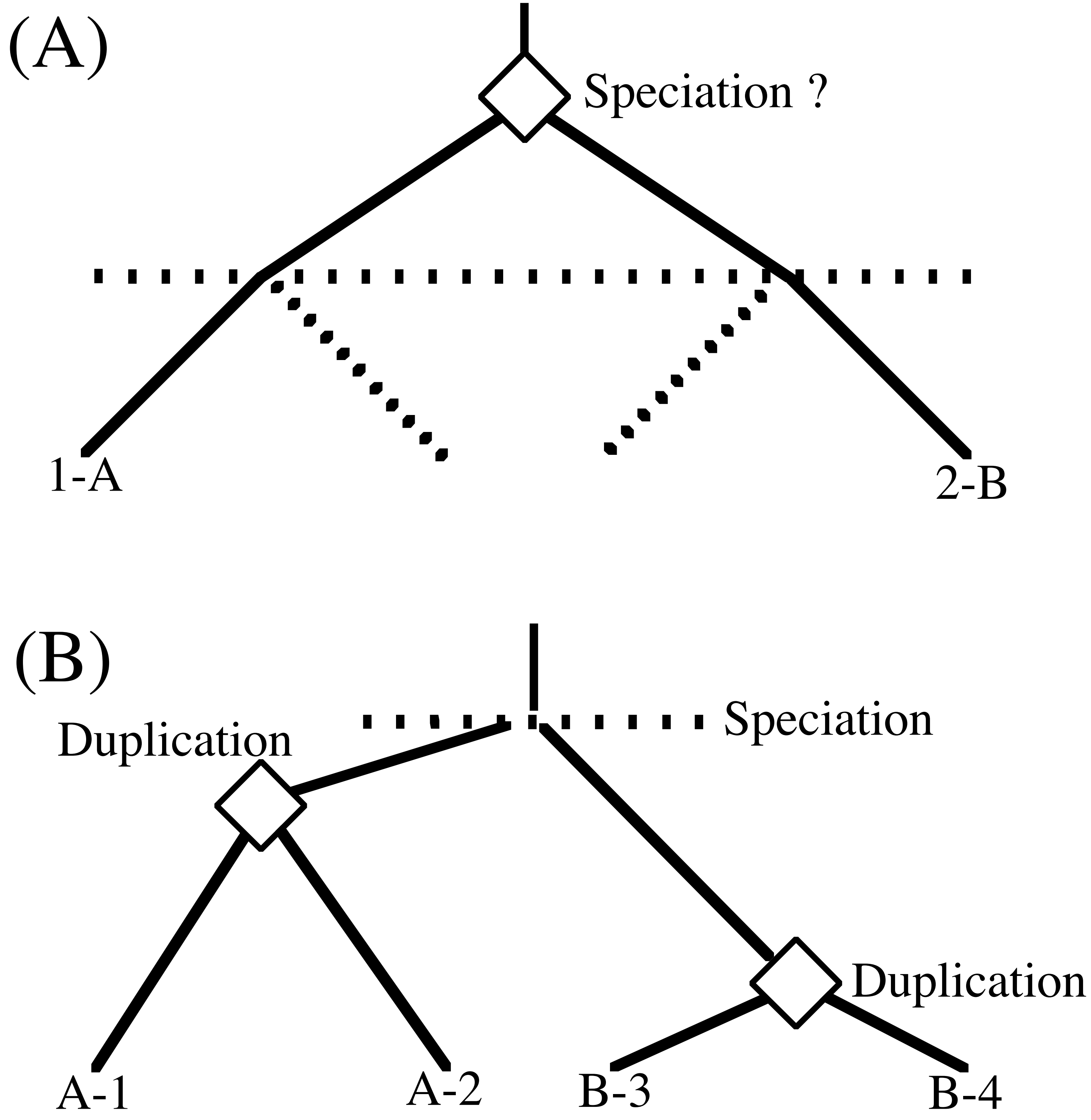


Figure 3-15: Duplogs transformed to be 'alleles'

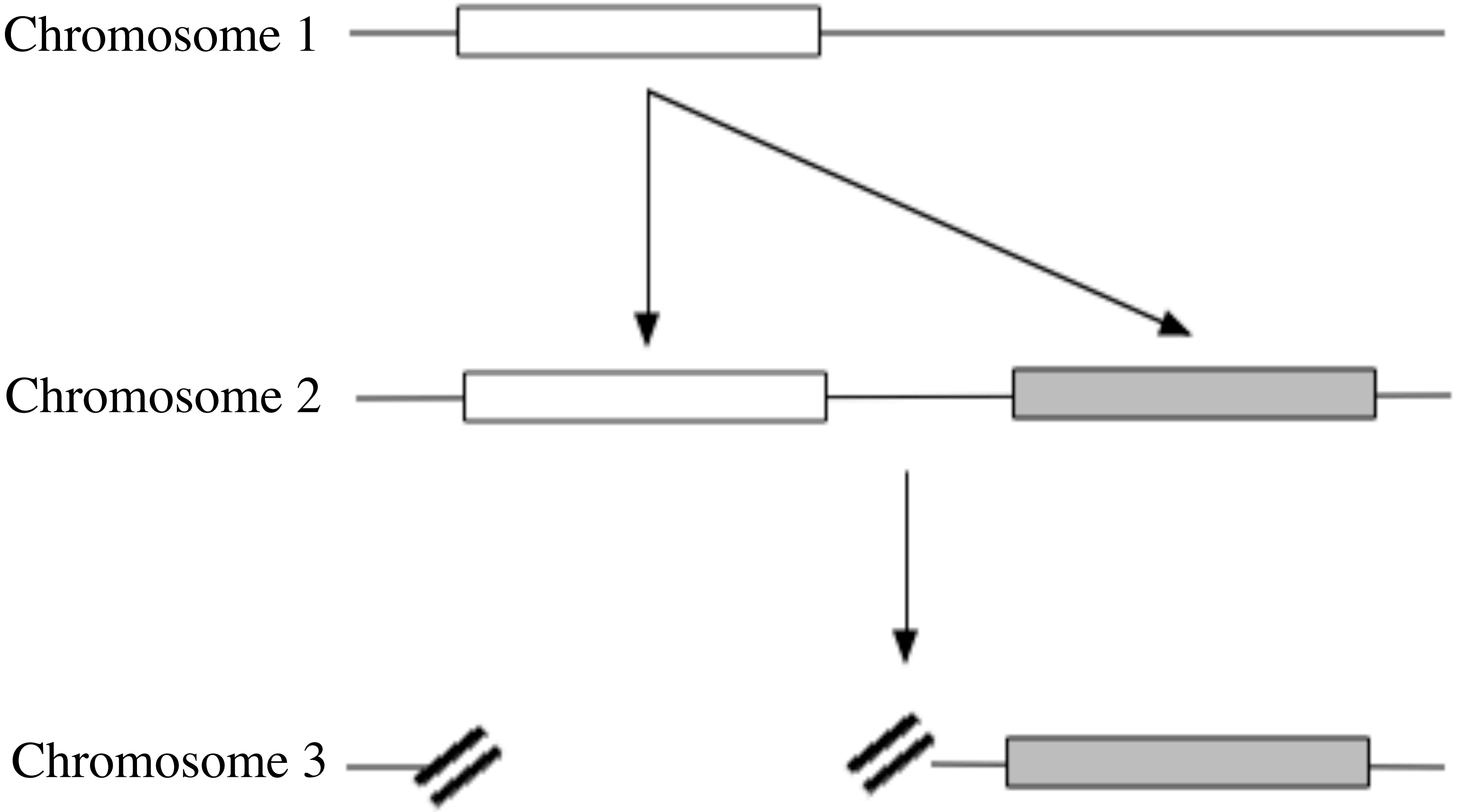


Figure 3-16: Phylogenetic relationship of hominoid immunoglobulin A gene region (from ref 3-20)

(A)Spurious tree when nucleotide sequence data were simply used for tree construction

(B)Proposed tree where one gene duplication occurred before speciation of human and great apes

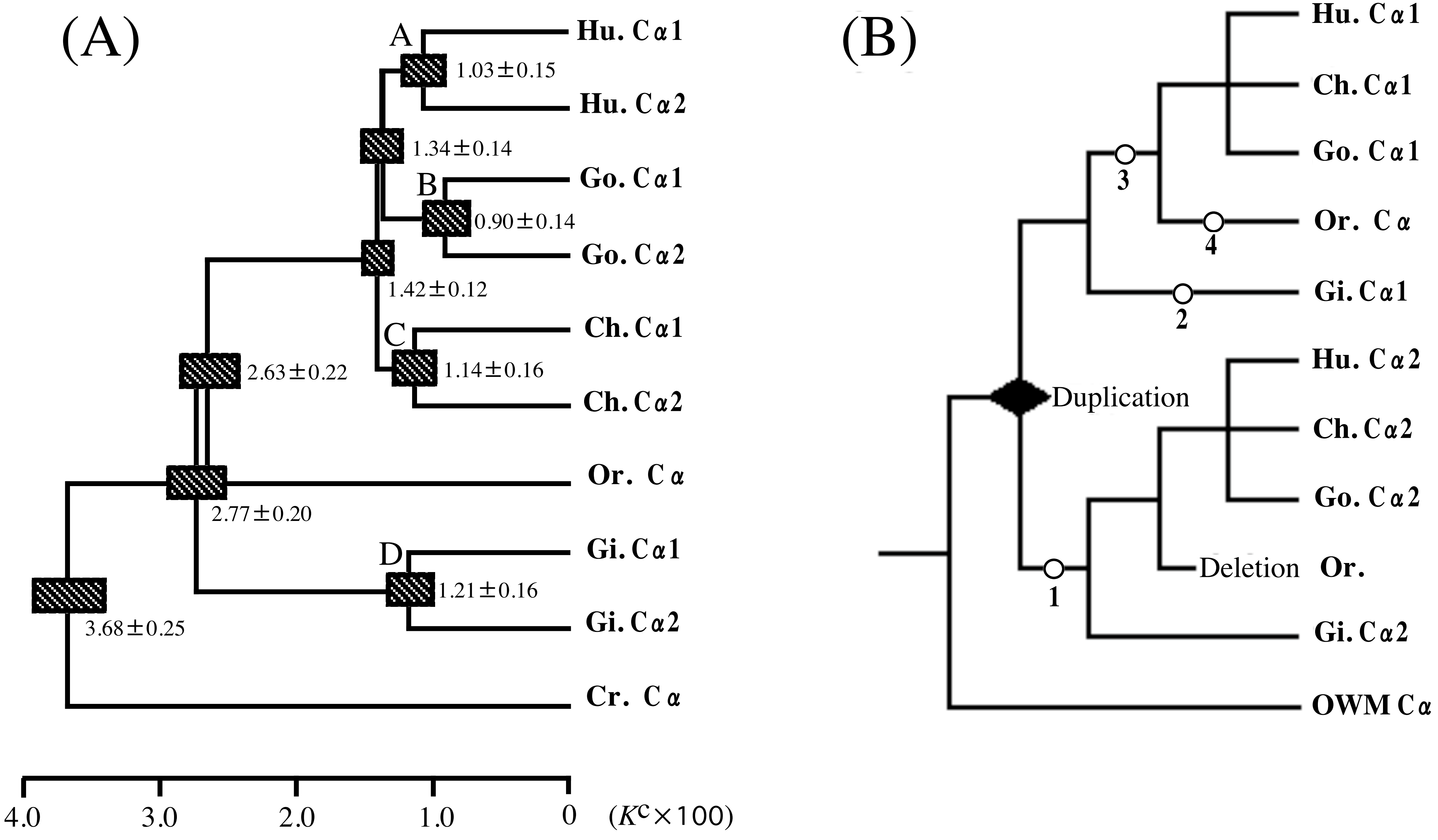


Figure 3-18: Two layers of species tree (based on Saitou 2007)
(A) Expected or true species tree. (B) Estimated species tree.

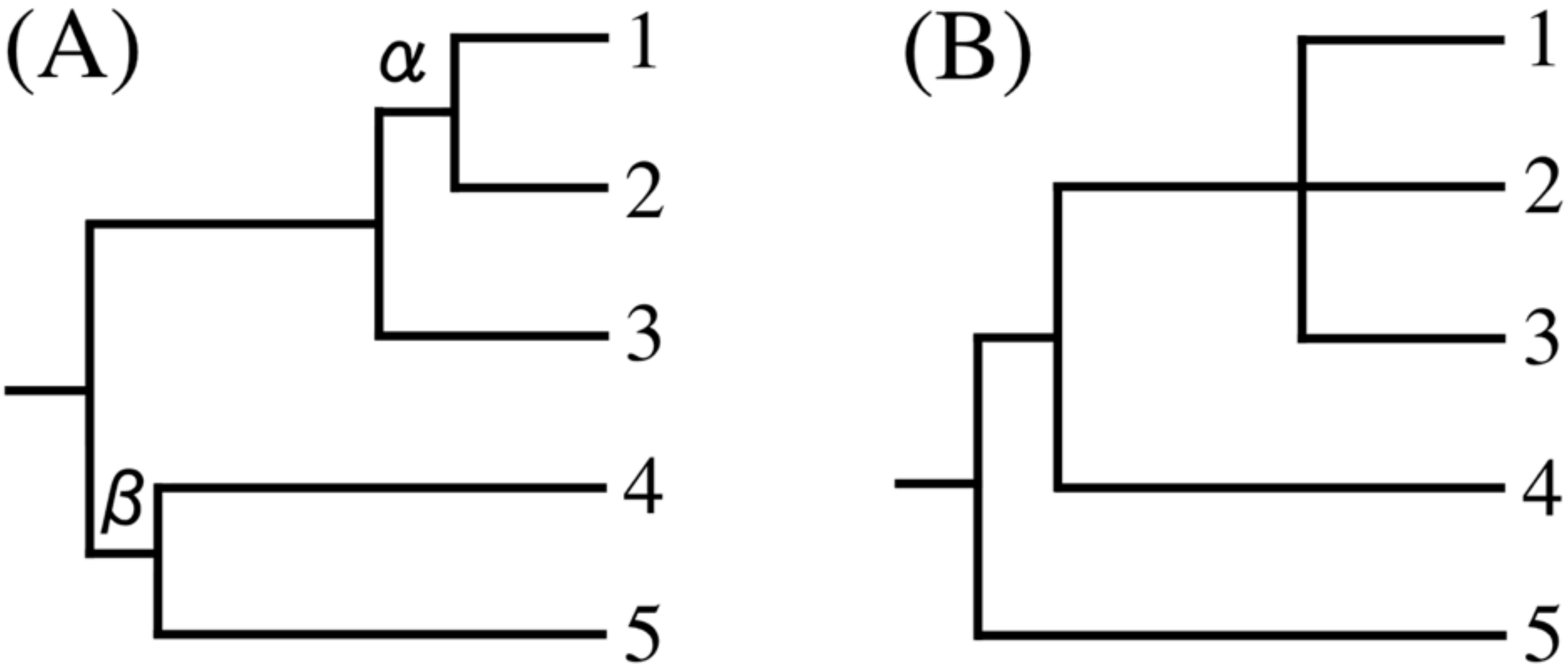


Figure 3-19: A mitochondrial DNA genome tree as approximation of a species tree. Human, chimpanzee, and bonobo, 5 individuals each were used.

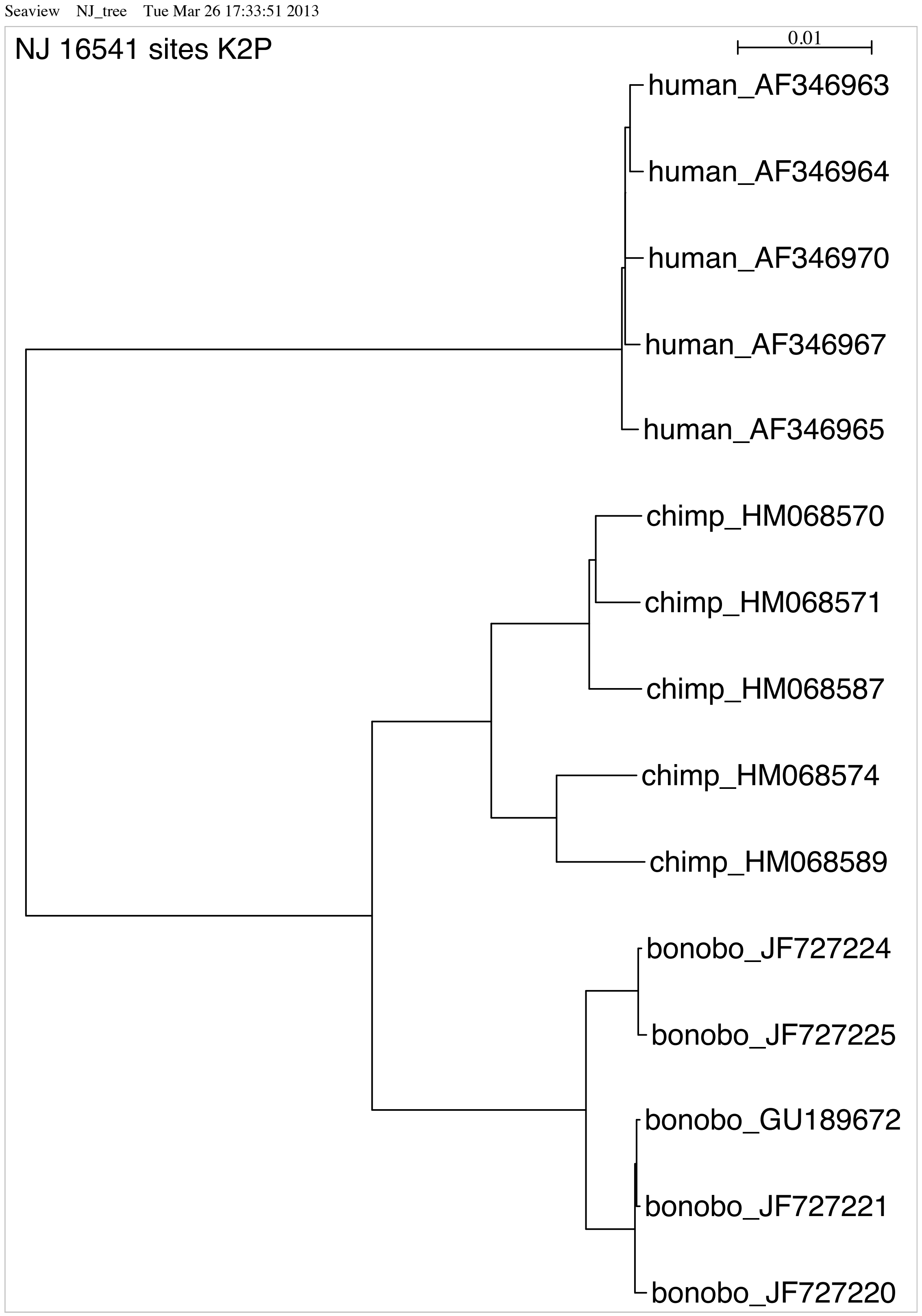


Figure 3-21: Trees for 6 OTUs. (A) A rooted tree. (B) An unrooted tree.

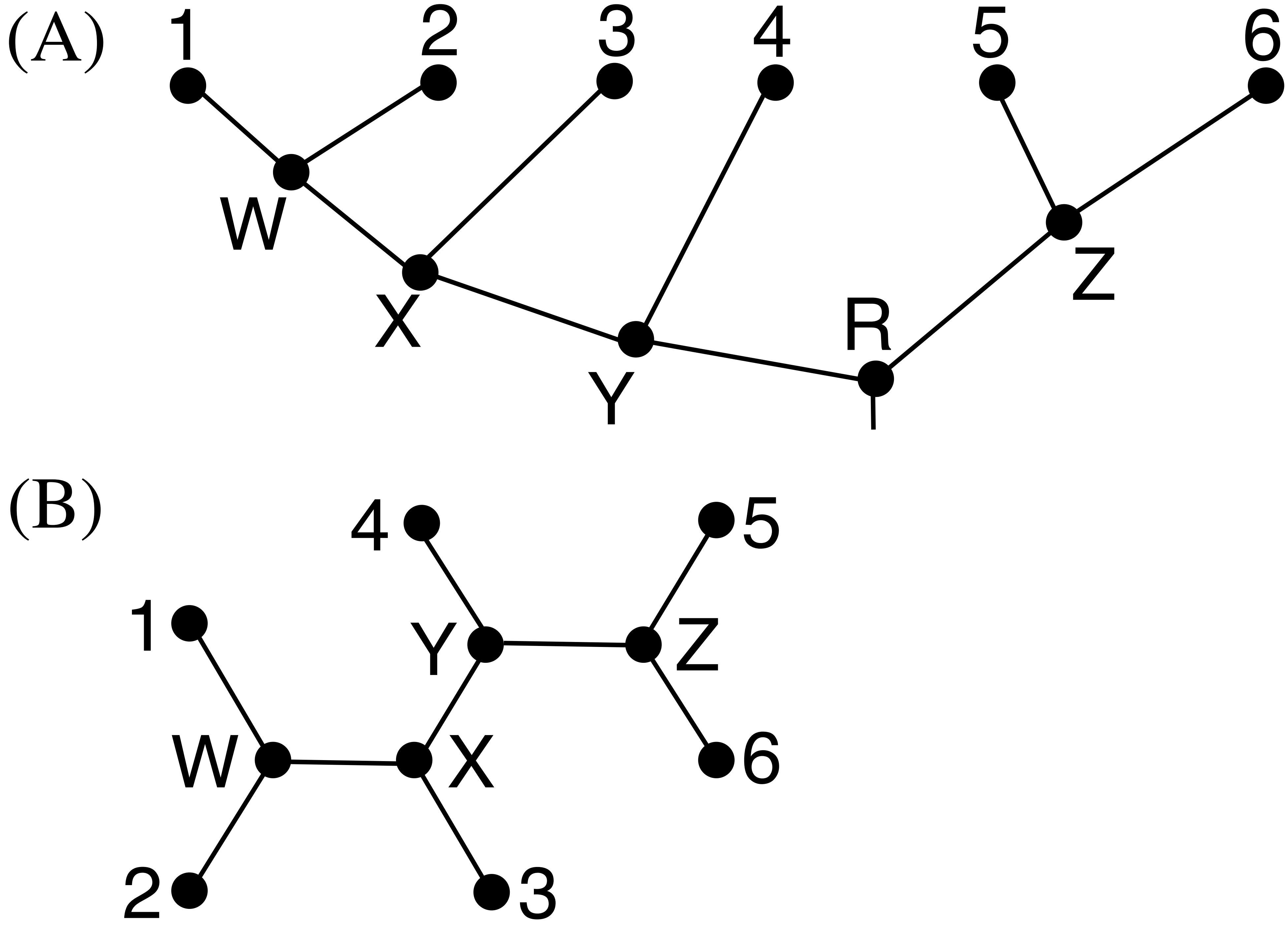
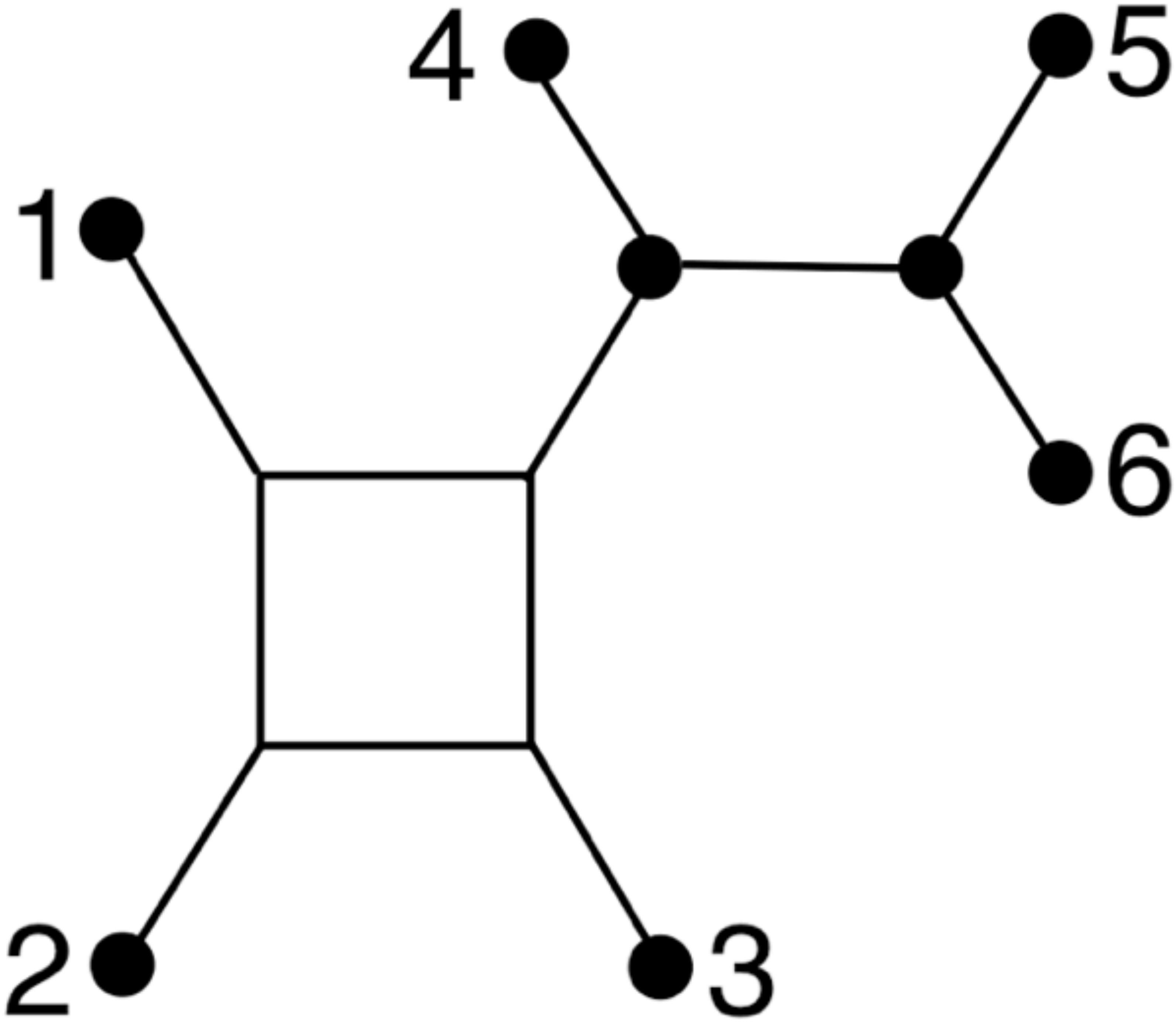
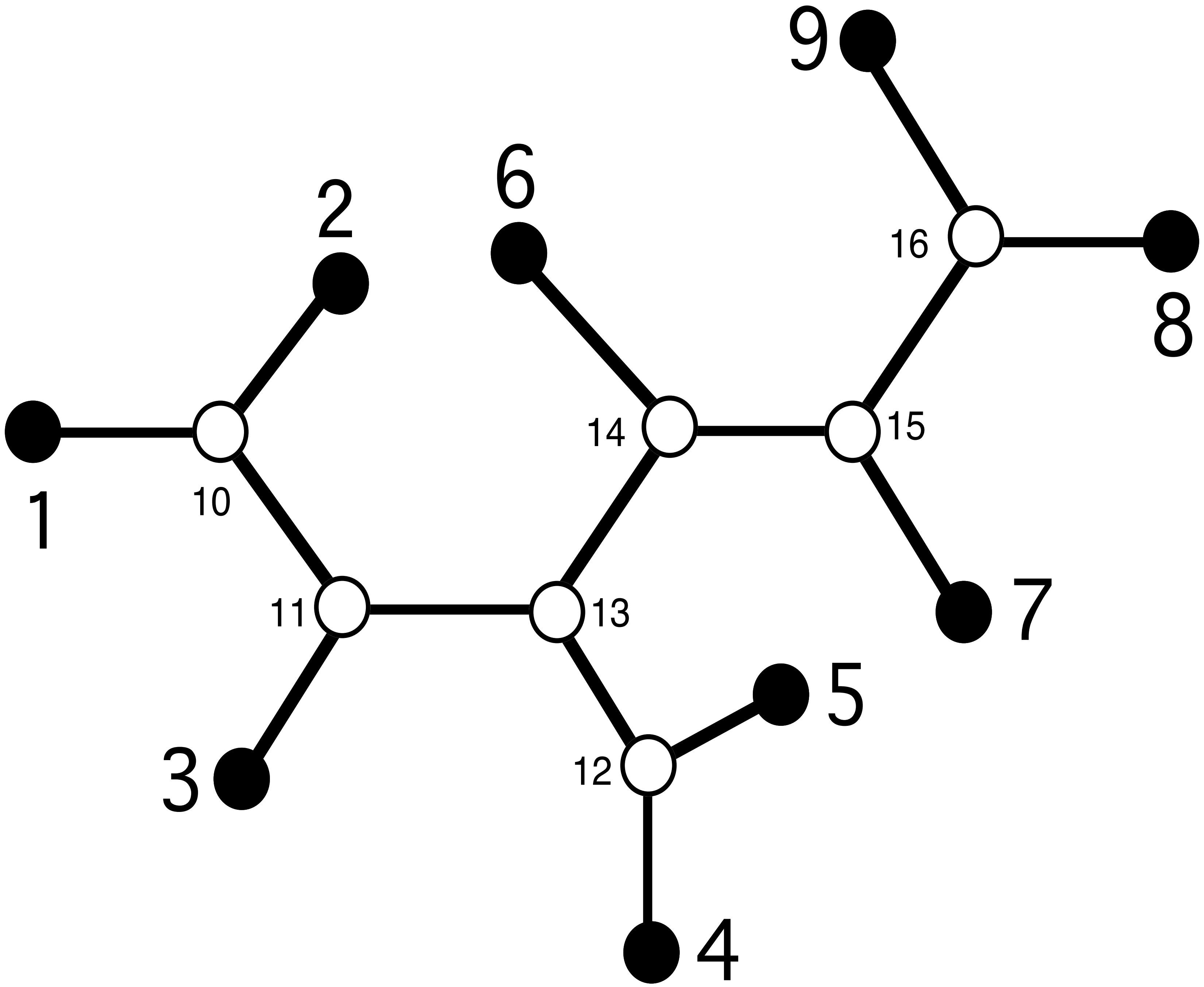
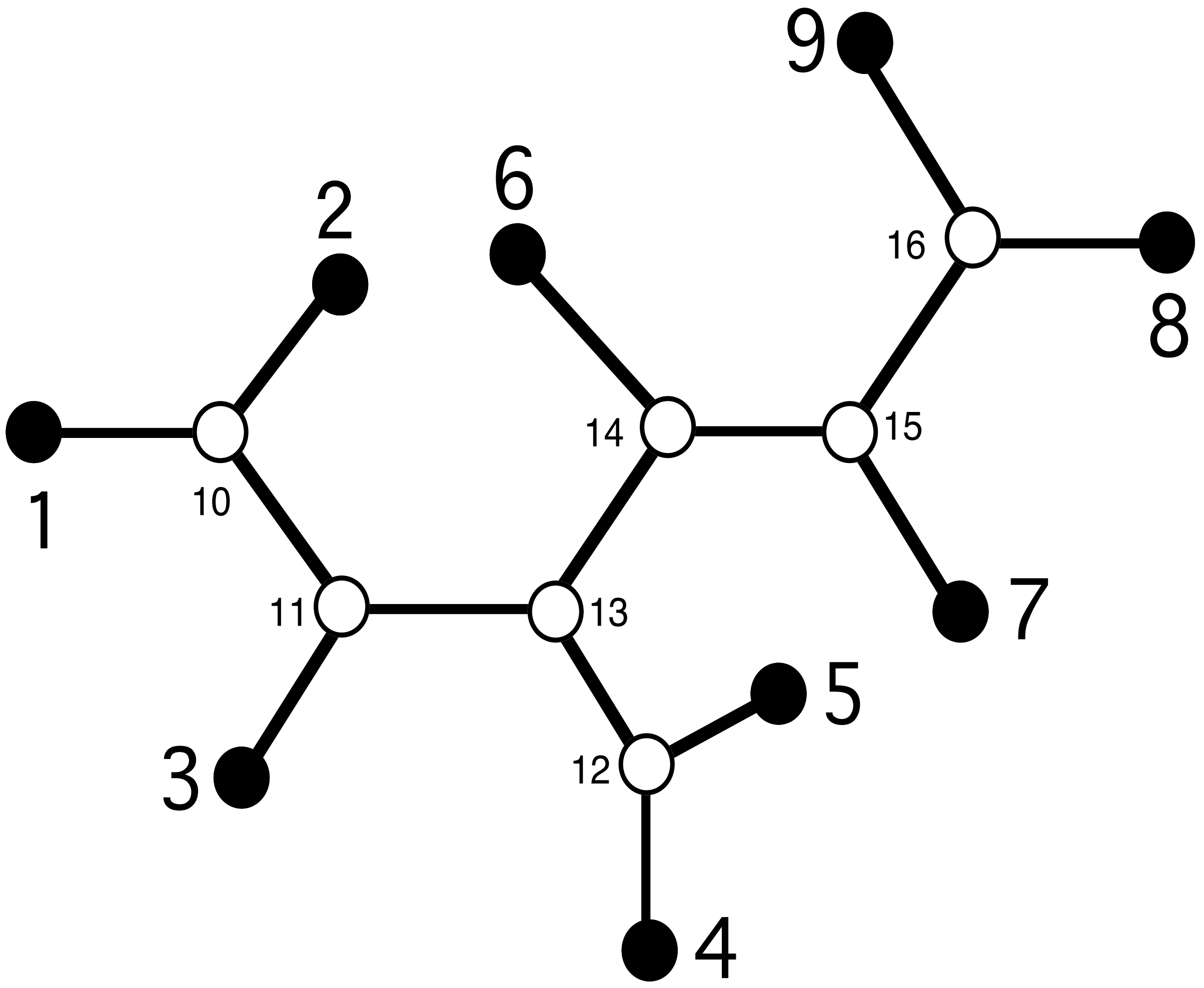


Figure 3-22: Example of nontree networks for 6 OTUs




9
6
2
16
8
14
15
1
10
7
11
13
5
3
12
4

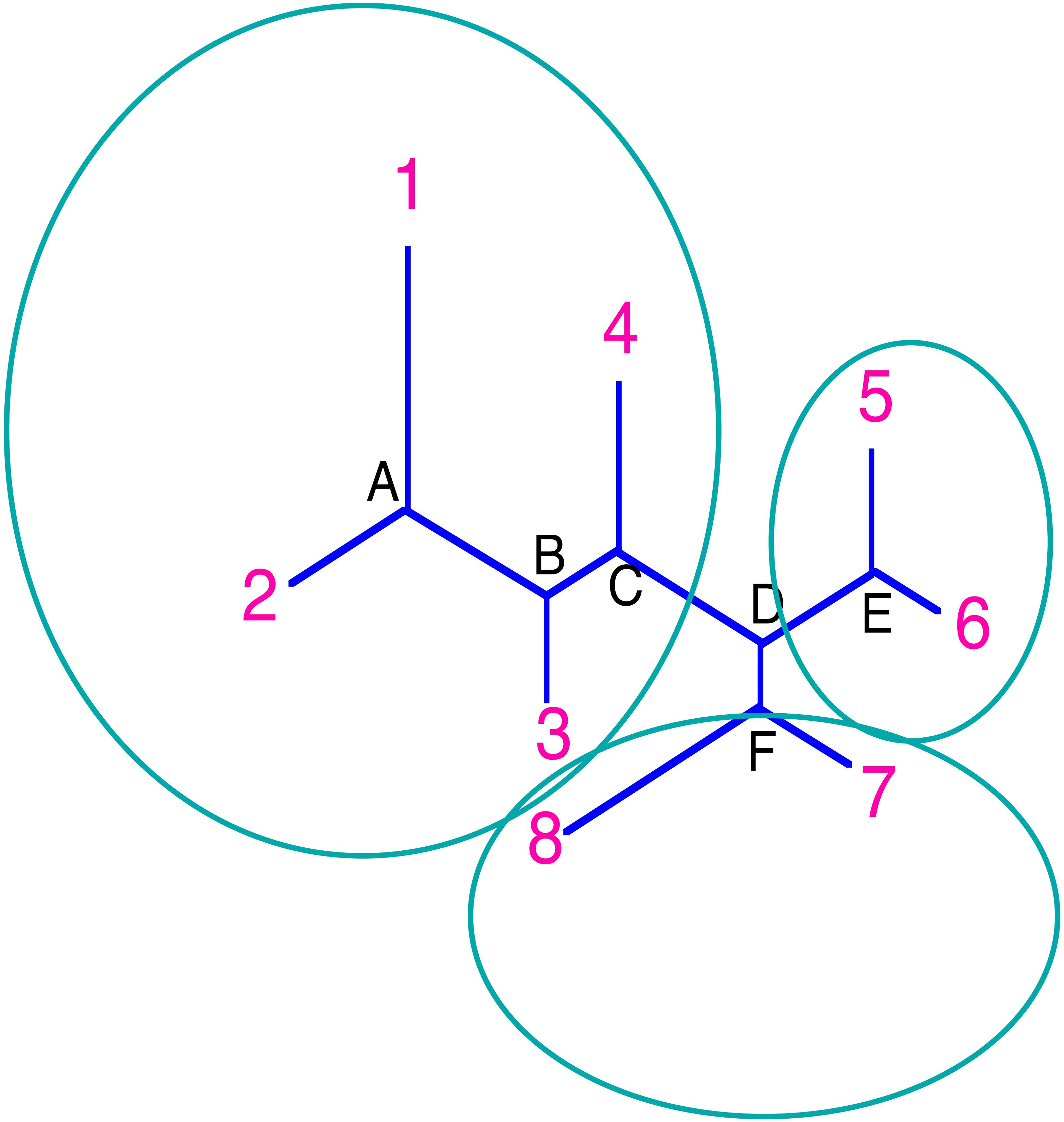
How to describe one tree?



List of branches

[1, 10]
[2, 10]
[3, 11]
[4, 12]
[5, 12]
[6, 14]
[7, 15]
[8, 16]
[9, 16]
[10, 11]
[11, 13]
[12, 13]
[13, 14]
[14, 15]
[15, 16]

How to describe one unrooted tree?



Newick (New Hampshire) Format

- (I, II, III);
- I = 1,2,3,4   II = 5,6   III = 7,8
- I = (1,2)
- I = ((1,2),3)
- I = (((1,2),3),4)
- ((((1,2),3),4),(5,6),(7,8));
- = (((((1,2),3),4),(5,6)),7,8);

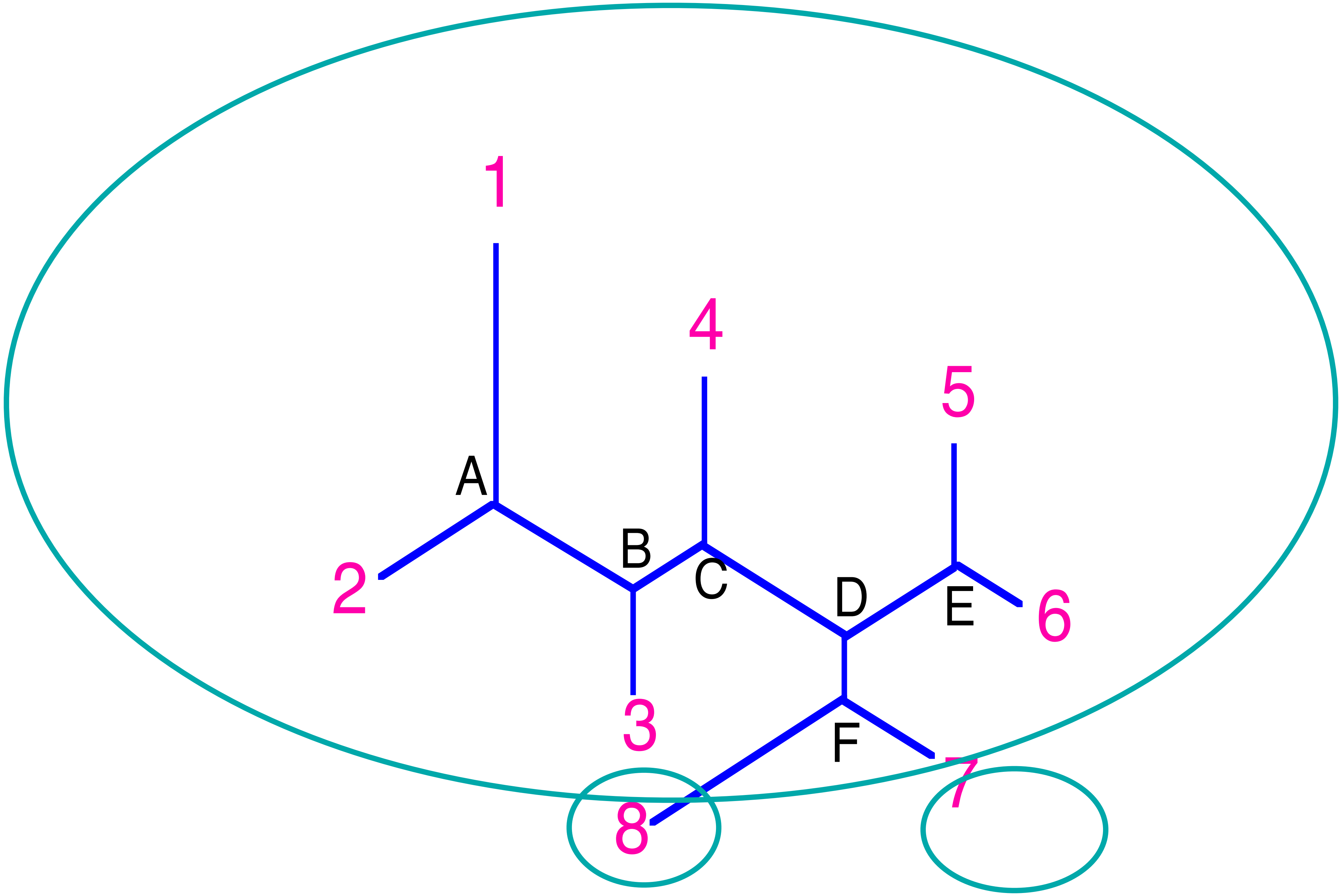
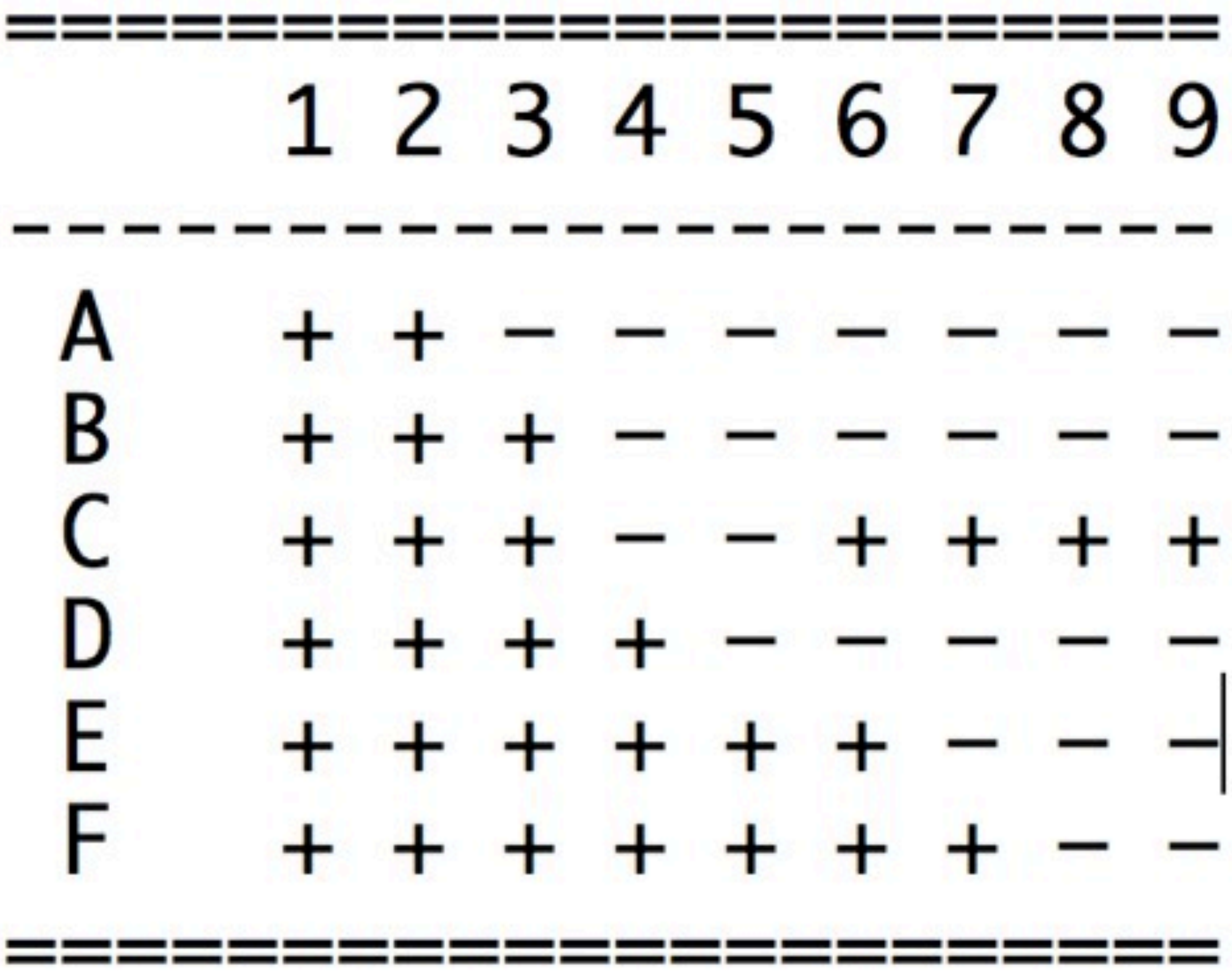

1
4
5
A
2
B
C
D
E
6
3
F
7
8

Table 3-2: A splits matrix that describes the tree shown in Figure 3-28



	1	2	3	4	5	6	7	8	9
A	+	+	−	−	−	−	−	−	−
B	+	+	+	−	−	−	−	−	−
C	+	+	+	−	−	+	+	+	+
D	+	+	+	+	−	−	−	−	−
E	+	+	+	+	+	+	−	−	−
F	+	+	+	+	+	+	+	−	−

Completely Bifurcating Tree

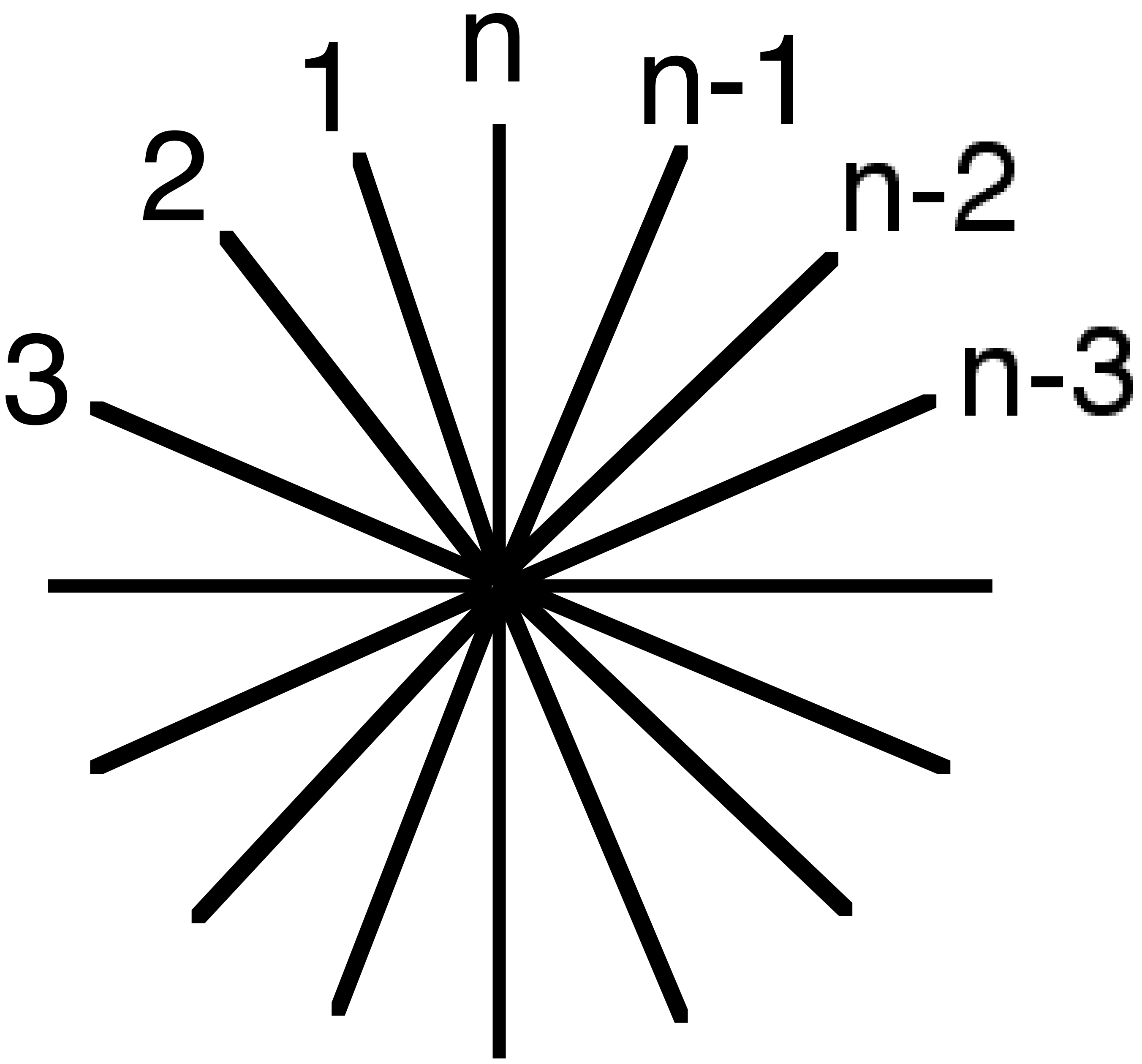


Figure 3-23: Three possible unrooted trees for 4 OTUs and their corresponding rooted trees

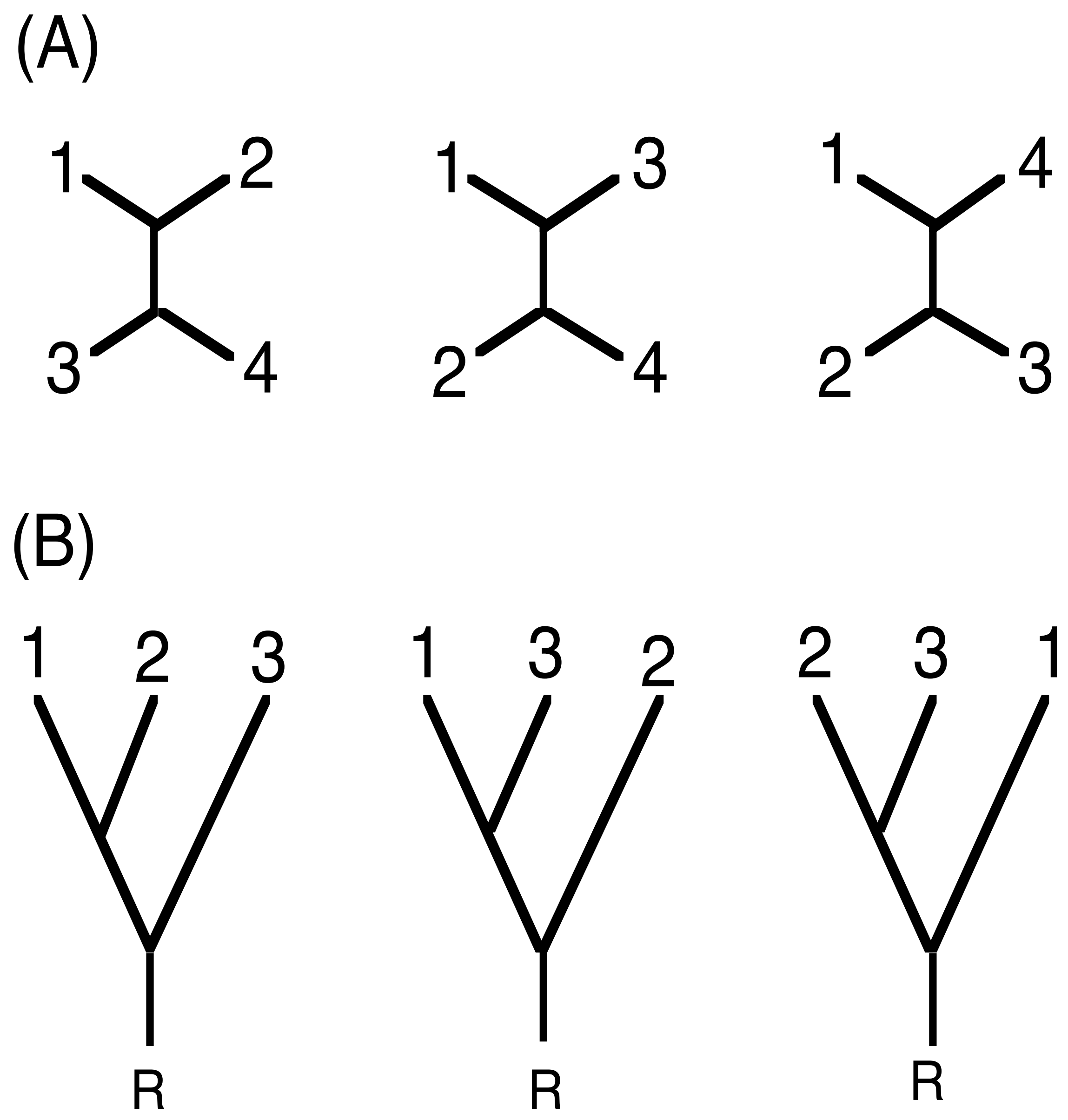
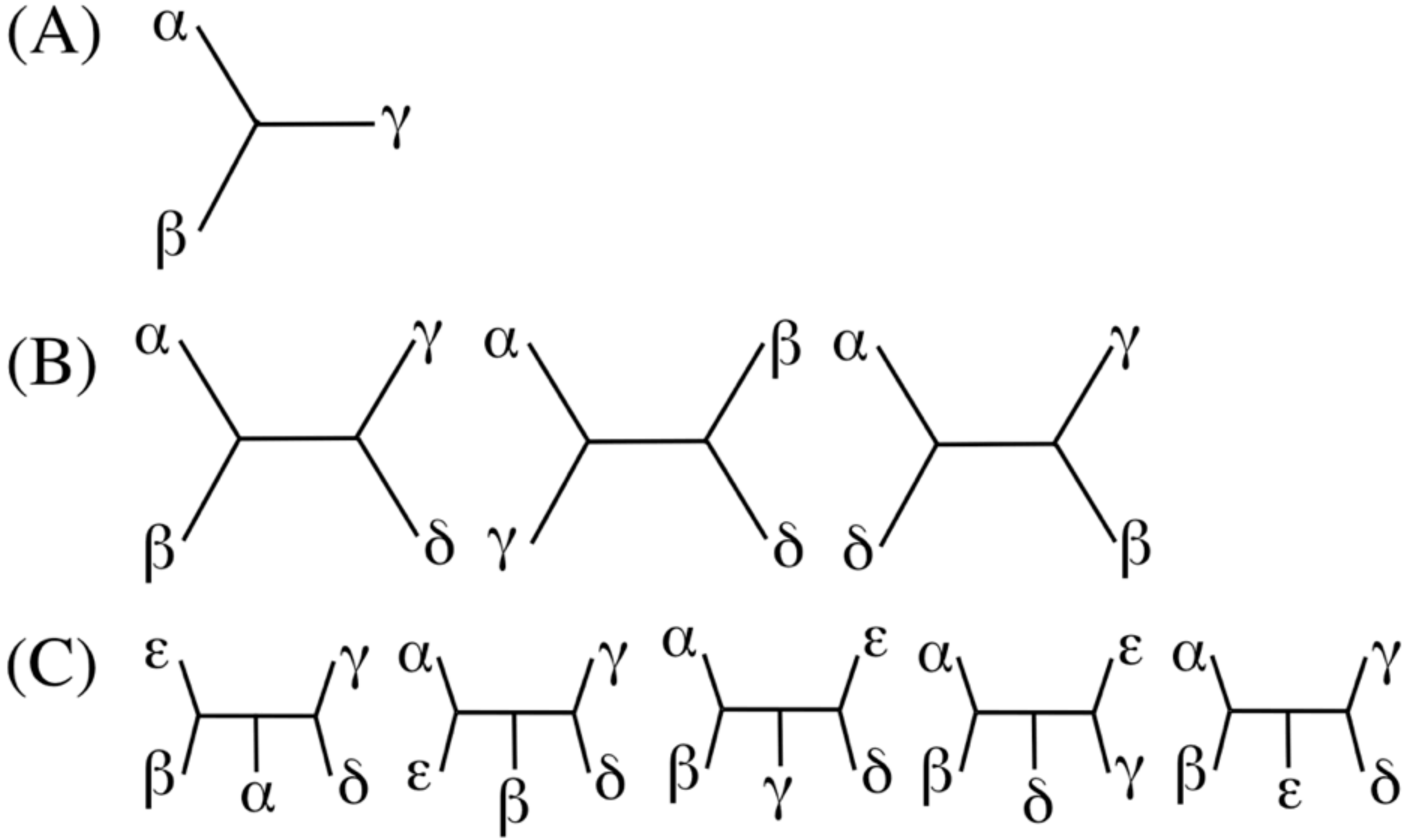


Figure 3-24: Sequential addition of new OTUs



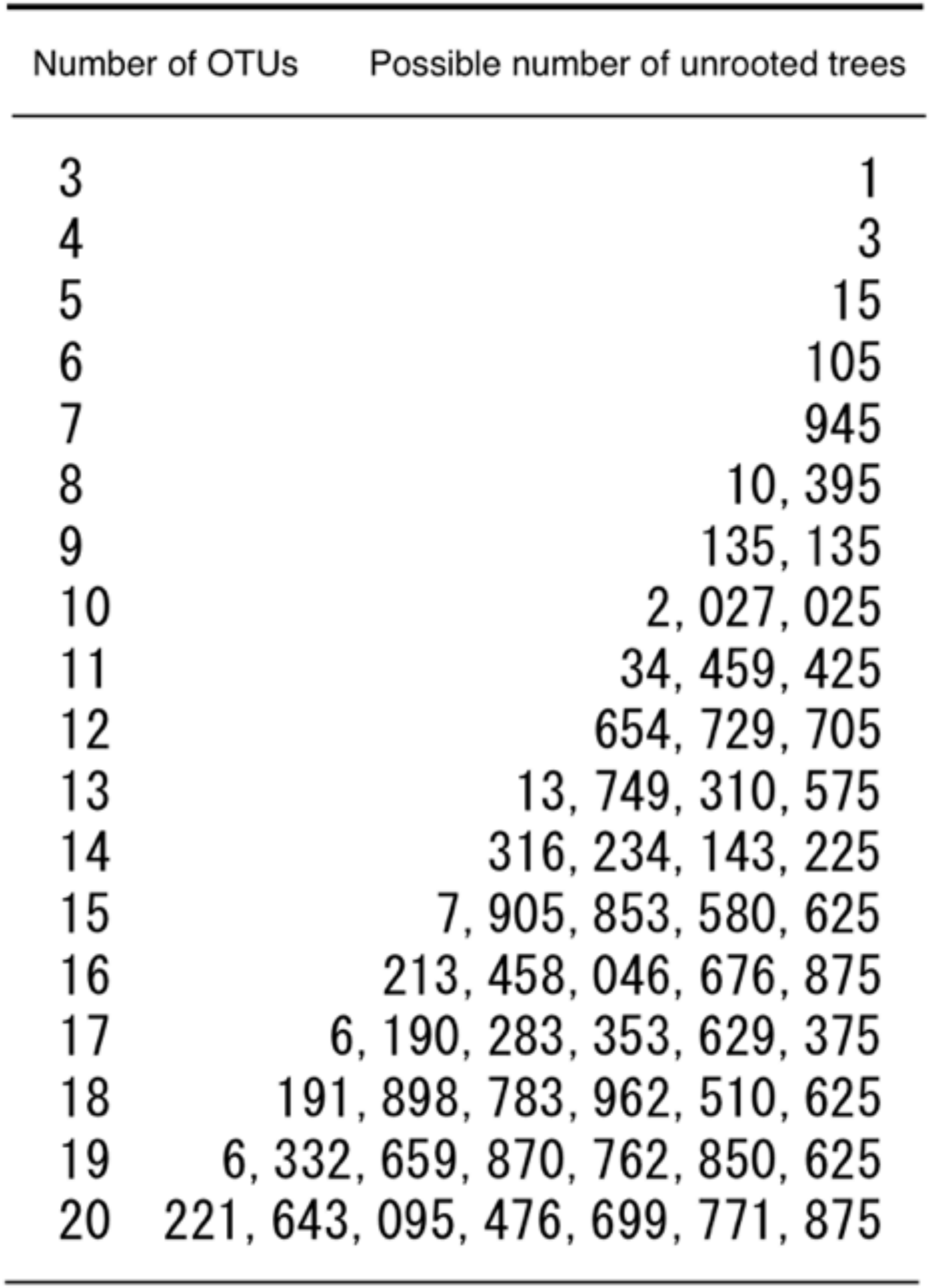
General Formula

$$Nr(n) = 1\times3\times5\times \ldots \times(2n-3)$$

$$Nu(n) = 1\times3\times5\times \ldots \times(2n-5)$$

$$= (2n - 5)! / [2^{n-3} (n - 3)!]$$

Table 3-1: Number of possible unrooted tree topologies for up to 20 OTUs



Number of OTUs	Possible number of unrooted trees
3	1
4	3
5	15
6	105
7	945
8	10, 395
9	135, 135
10	2, 027, 025
11	34, 459, 425
12	654, 729, 705
13	13, 749, 310, 575
14	316, 234, 143, 225
15	7, 905, 853, 580, 625
16	213, 458, 046, 676, 875
17	6, 190, 283, 353, 629, 375
18	191, 898, 783, 962, 510, 625
19	6, 332, 659, 870, 762, 850, 625
20	221, 643, 095, 476, 699, 771, 875

Figure 3-29: The hypersphere of tree topologies: from the star tree to completely bifurcating tree when N = 7

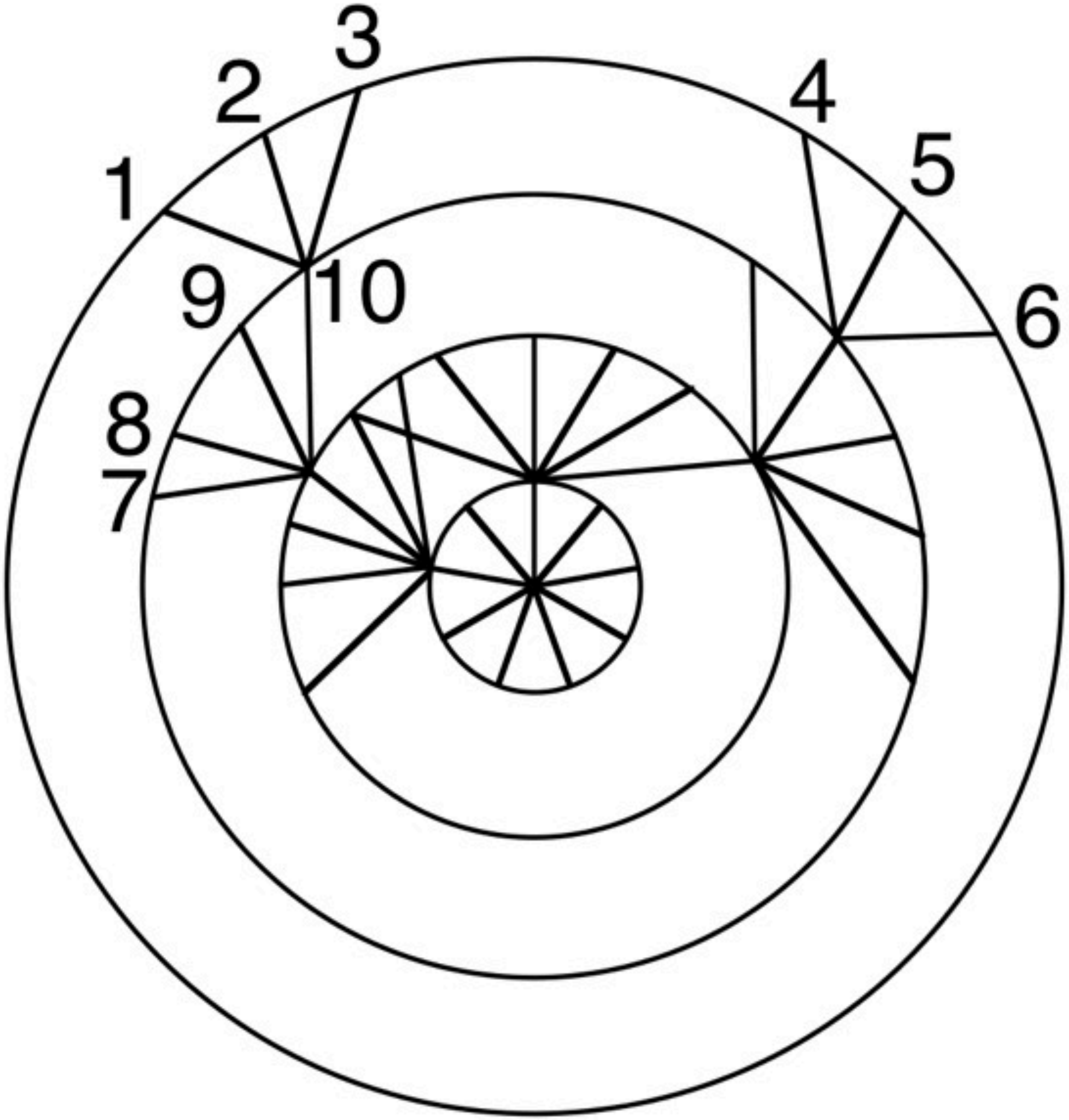
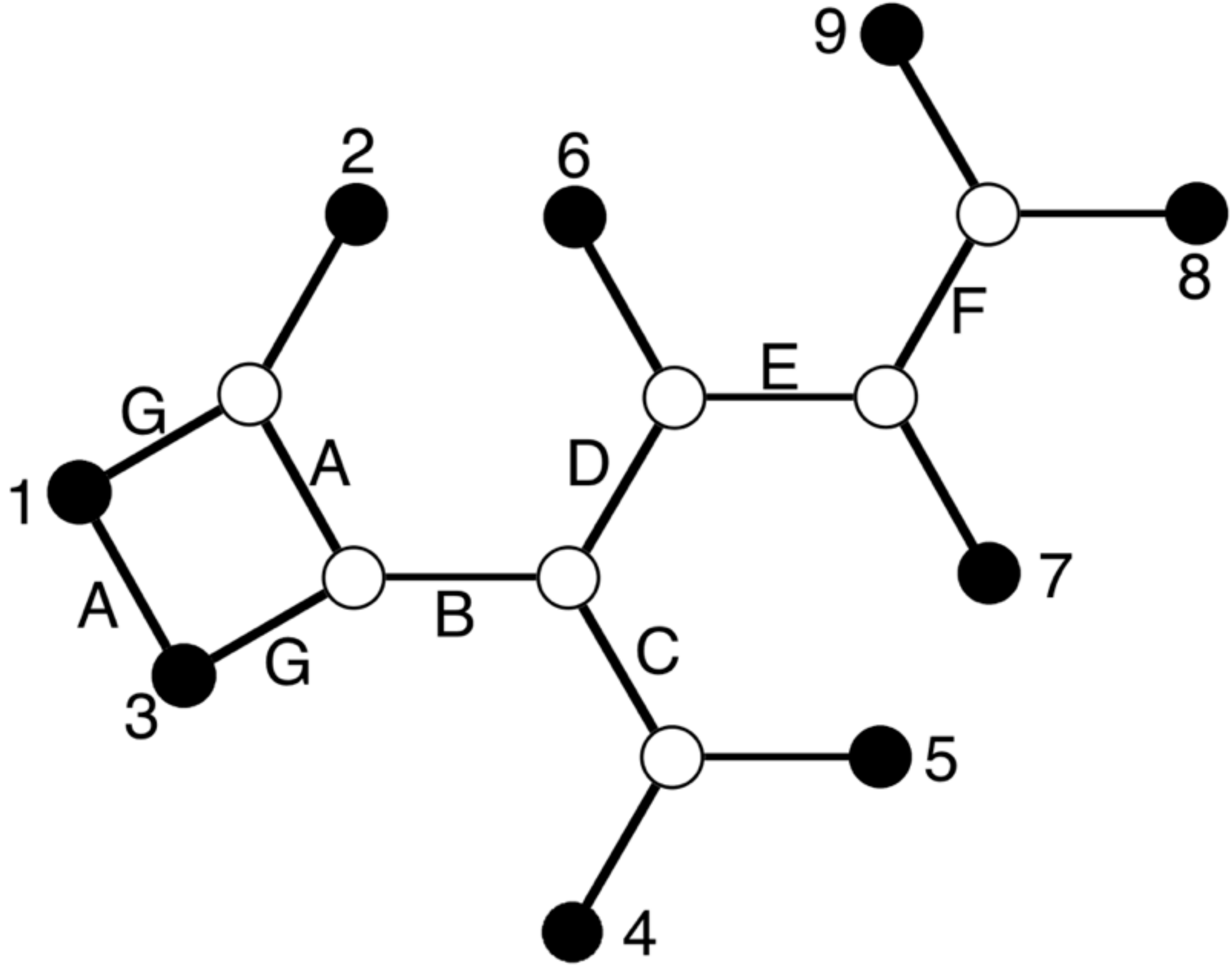


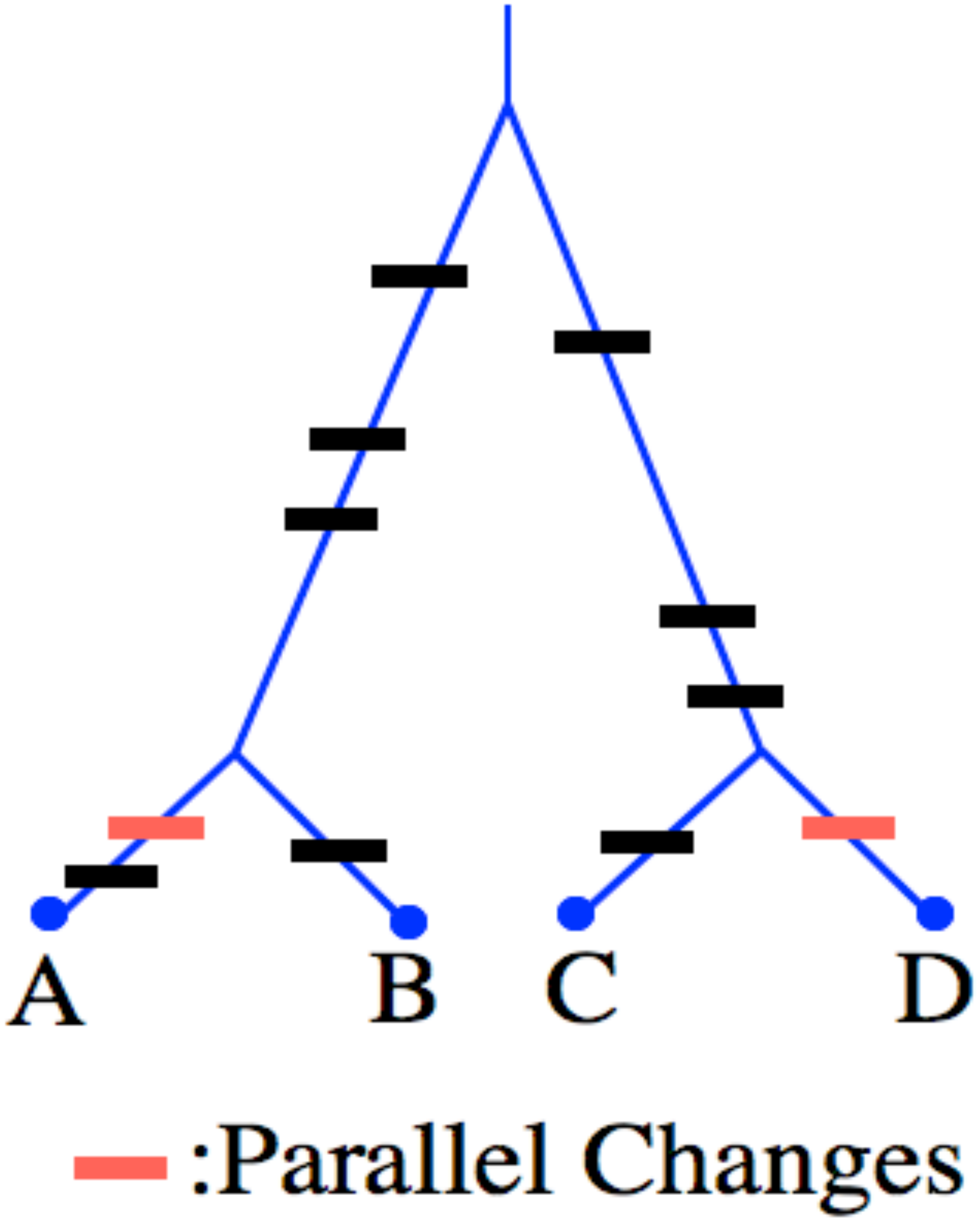
Table 3-3: A splits matrix for 9 OTUs that are not mutually compatible or nested

	1	2	3	4	5	6	7	8	9
A	+	+	−	−	−	−	−	−	−
B	+	+	+	−	−	−	−	−	−
C	+	+	+	−	−	+	+	+	+
D	+	+	+	+	−	−	−	−	−
E	+	+	+	+	+	+	−	−	−
F	+	+	+	+	+	+	+	−	−
G	+	−	+	−	−	−	−	−	−

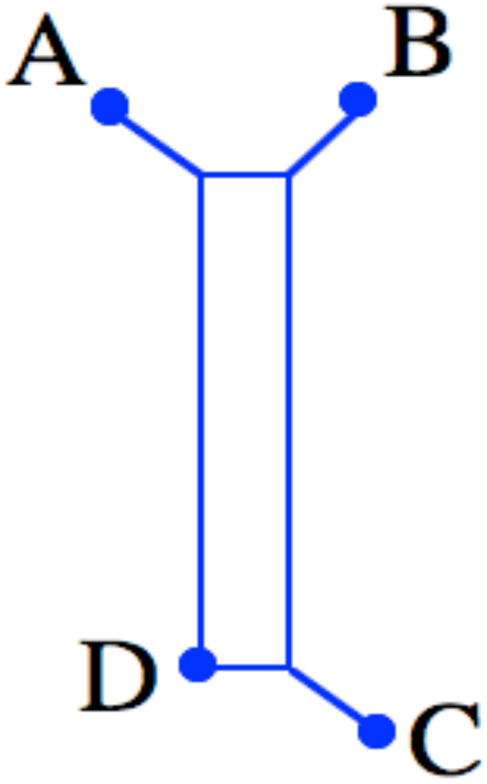
Figure 3-26: A nontree network corresponding to the split matrix shown in Table 3-3.



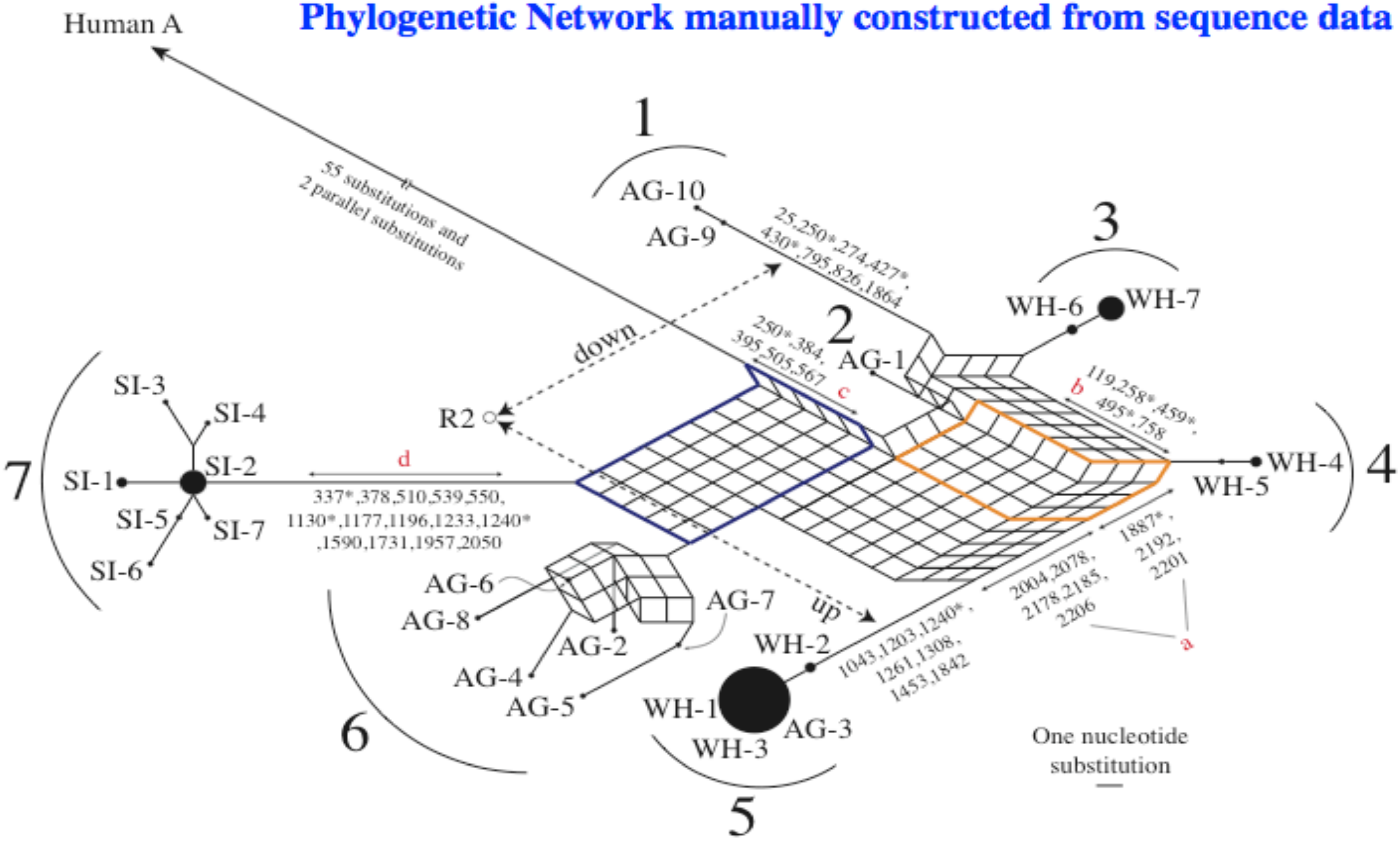
Gene Tree



Phylogenetic Network



Phylogenetic Network manually constructed from sequence data



From Kitano et al. (2010)

Part 2 Evolving Genomes

Chapter 6 Brief History of Life
Chapter 7 Prokaryote Genomes
Chapter 8 Eukaryote Genomes
Chapter 9 Vertebrate Genomes
Chapter 10 Human Genomes

SAITOU LABORATORY
NIG SOKENDAI U.TOKYO / MISHIMA

Introduction to Evolutionary Genomics

Handout Part 2

Saitou Naruya

National Institute of Genetics, Mishima

Dept. Genetics, Graduate University for Advanced Studies

Dept. Biological Sciences, University of Tokyo

saitounr@nig.ac.jp

http://www.saitou-naruya-laboratory.org/english

June 18, A.S. 0013
Suzhou, China

Part 2 Evolving Genomes

Chapter 6 Brief History of Life
Chapter 7 Prokaryote Genomes
Chapter 8 Eukaryote Genomes
Chapter 9 Vertebrate Genomes
Chapter 10 Human Genomes

Part 3 Methods for Evolutionary Genomics

Chapter 11 Genome Sequencing

Chapter 12 Omic Data Collection

Chapter 13 Databases

Chapter 14 Sequence Homology Handling

Chapter 15 Evolutionary Distances

Chapter 16 Tree and Network Building

Chapter 17 Population Genomics

Table 15-1: Example of a distance matrix (data from Table 3 of Ishida et al. [1995; ref 15-54])

	1	2	3	4	5	6	7
1	0	9	11	6	42	38	35
2	9	0	6	5	45	41	38
3	11	6	0	7	47	43	40
4	6	5	7	0	42	38	35
5	42	45	47	5	0	46	43
6	38	41	43	5	46	0	29
7	35	38	40	5	43	29	0

1: Thoroughbred horse (*Equus caballus*), 2: Przewalskii's wild horse (*E. caballus*), 3: Mongolian native horse (*E. caballus*), 4: Japanese native horse (*E. caballus*), 5: mountain zebra (*E. zebra*), 6: donkey (*E. asinus*), and 7: Grevy's zebra (*E. grevyi*).

Table 15-2: Example of a distance matrix showing only lower-triangle values (data from Table 3 of Ishida et al. [1995; ref 15-54])

	1	2	3	4	5	6
2	8.0					
3	11.7	5.4				
4	6.5	5.6	5.4			
5	41.8	49.0	46.5	43.7		
6	35.1	41.3	45.8	35.5	47.9	
7	36.6	34.8	34.8	38.2	41.5	29.1

1-7: Same as those of Table 15-1

Figure 15-1: An example of nucleotide sequence evolution only through substitutions

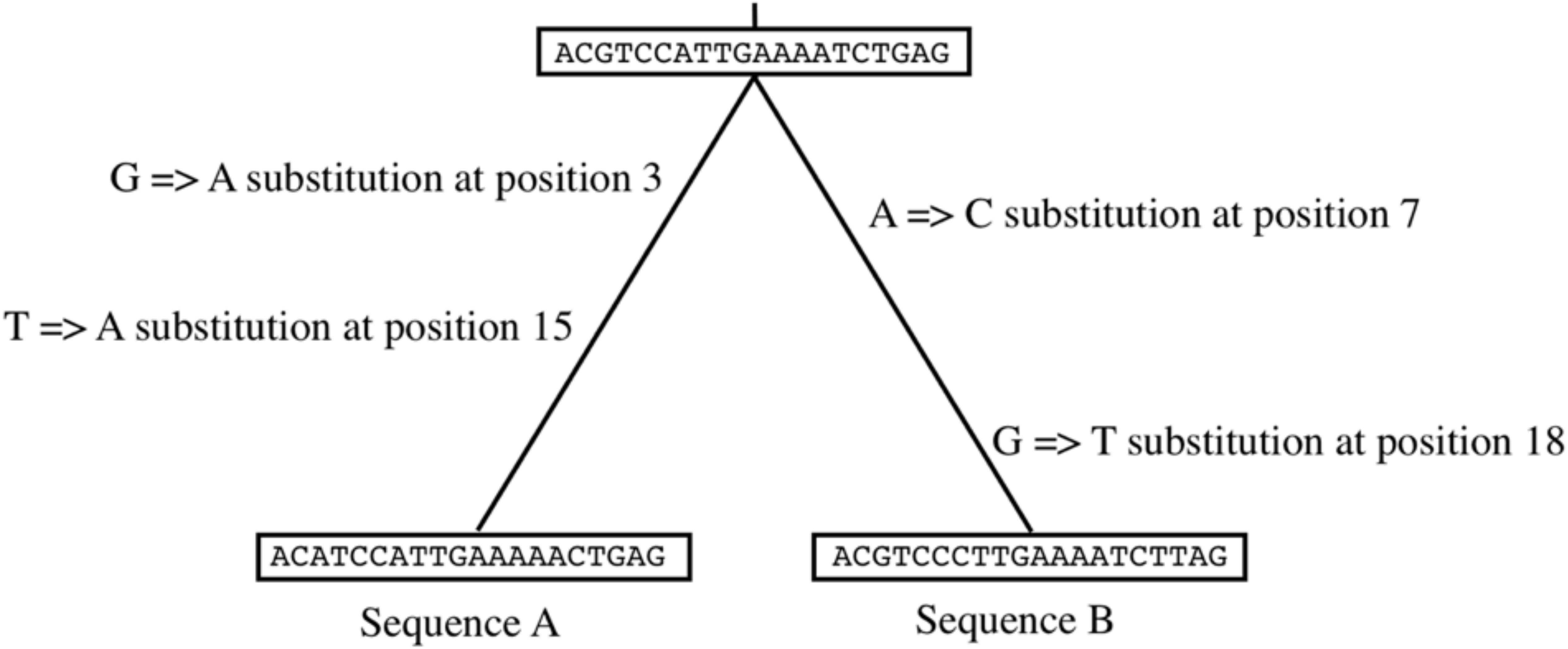


Figure 15-2: Another example of nucleotide sequence evolution only through substitutions

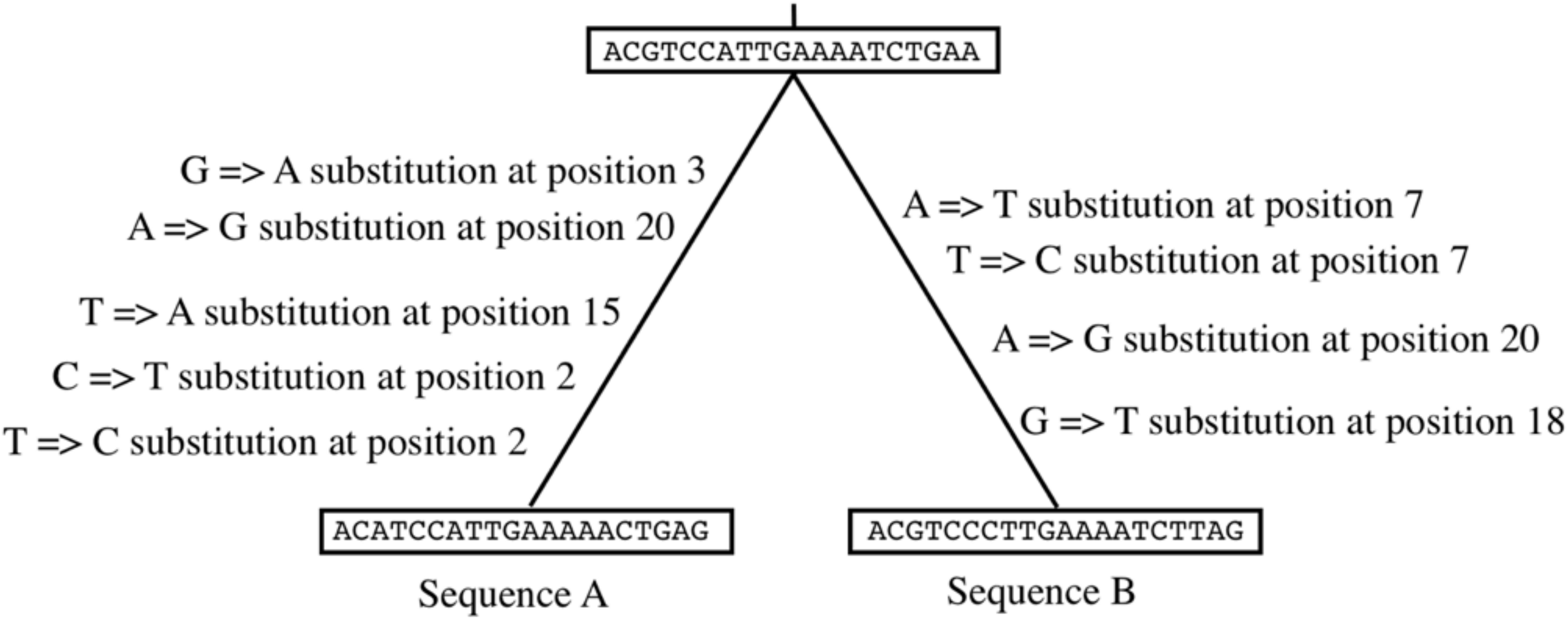


Table 15-3: One-parameter and two-parameter models of nucleotide substitution matrix

(A) One-parameter model

		NEW			
		A	C	T	G
O	A	1–3α	α	α	α
L	C	α	1–3α	α	α
D	T	α	α	1–3α	α
	G	α	α	α	1–3α

(B) Two-parameter model

		NEW			
		A	C	T	G
O	A	1–α–2β	β	β	α
L	C	β	1–α–2β	α	β
D	T	β	α	1–α–2β	β
	G	α	β	β	1–α–2β

Figure 15-3: An evolutionary scheme between two presentday sequences

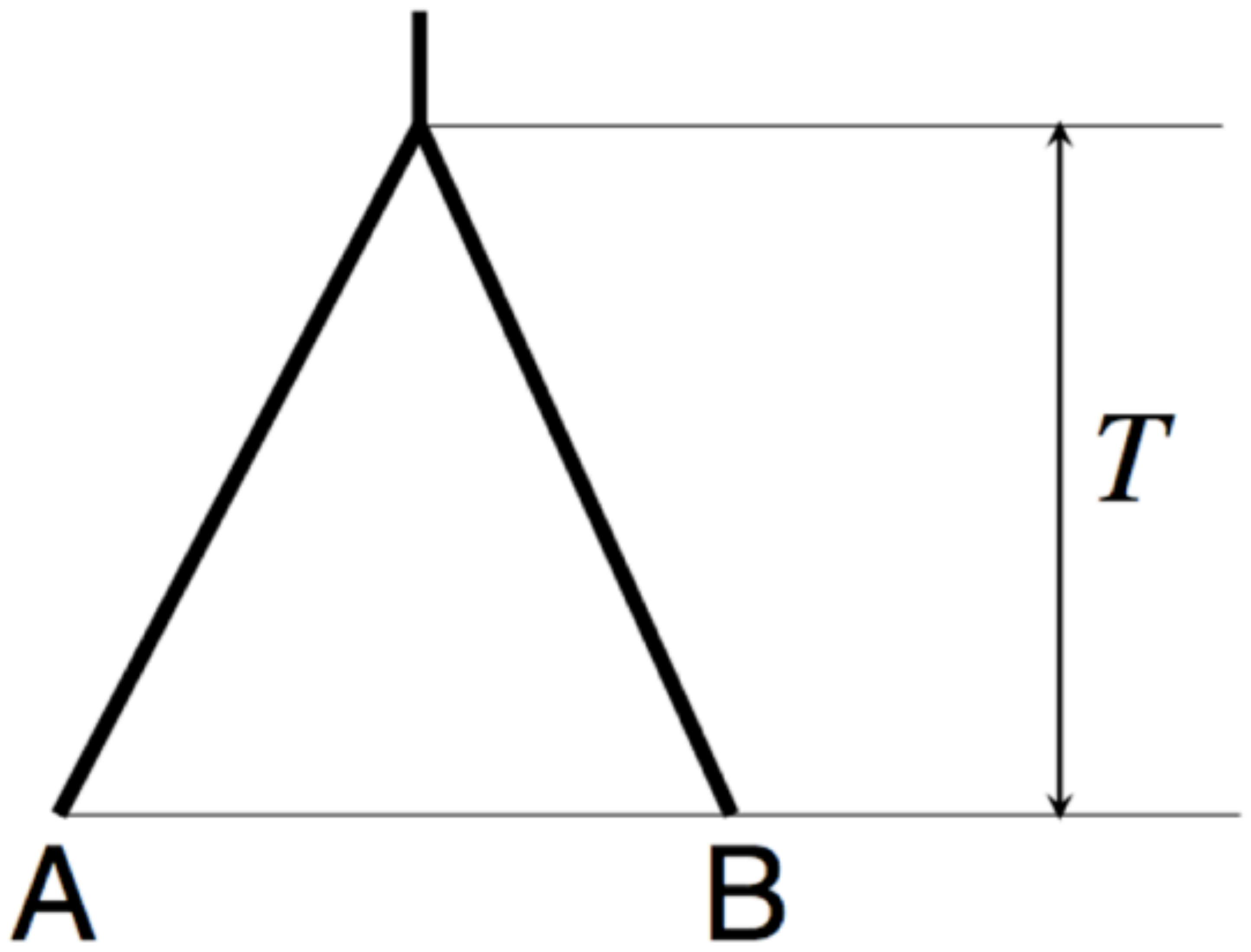


Figure 15-4: Relationship between *p* and *d* under the one-parameter method

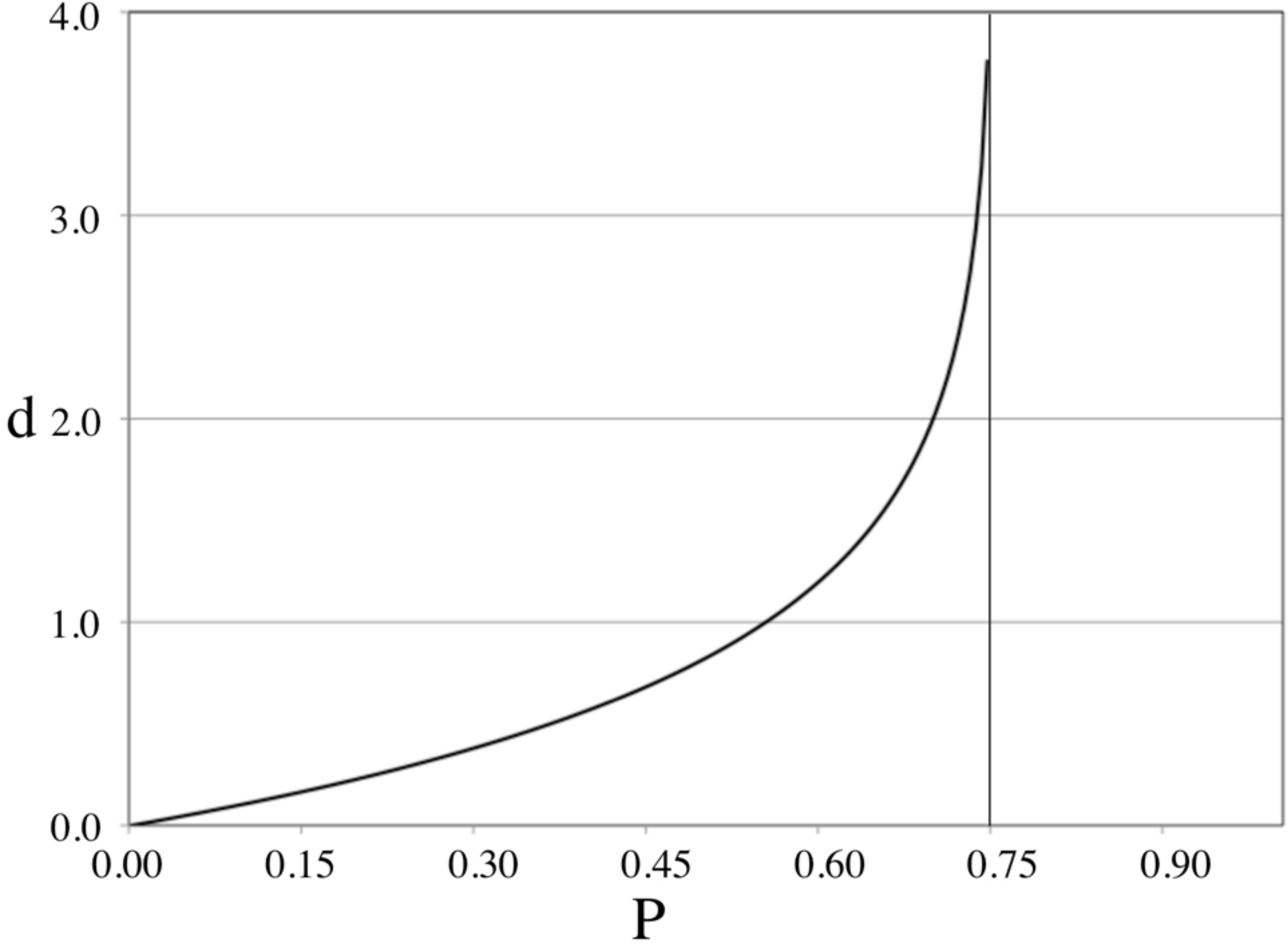


Figure 15-5: Temporal changes of *P* and *Q* under the two-parameter model with α=10β

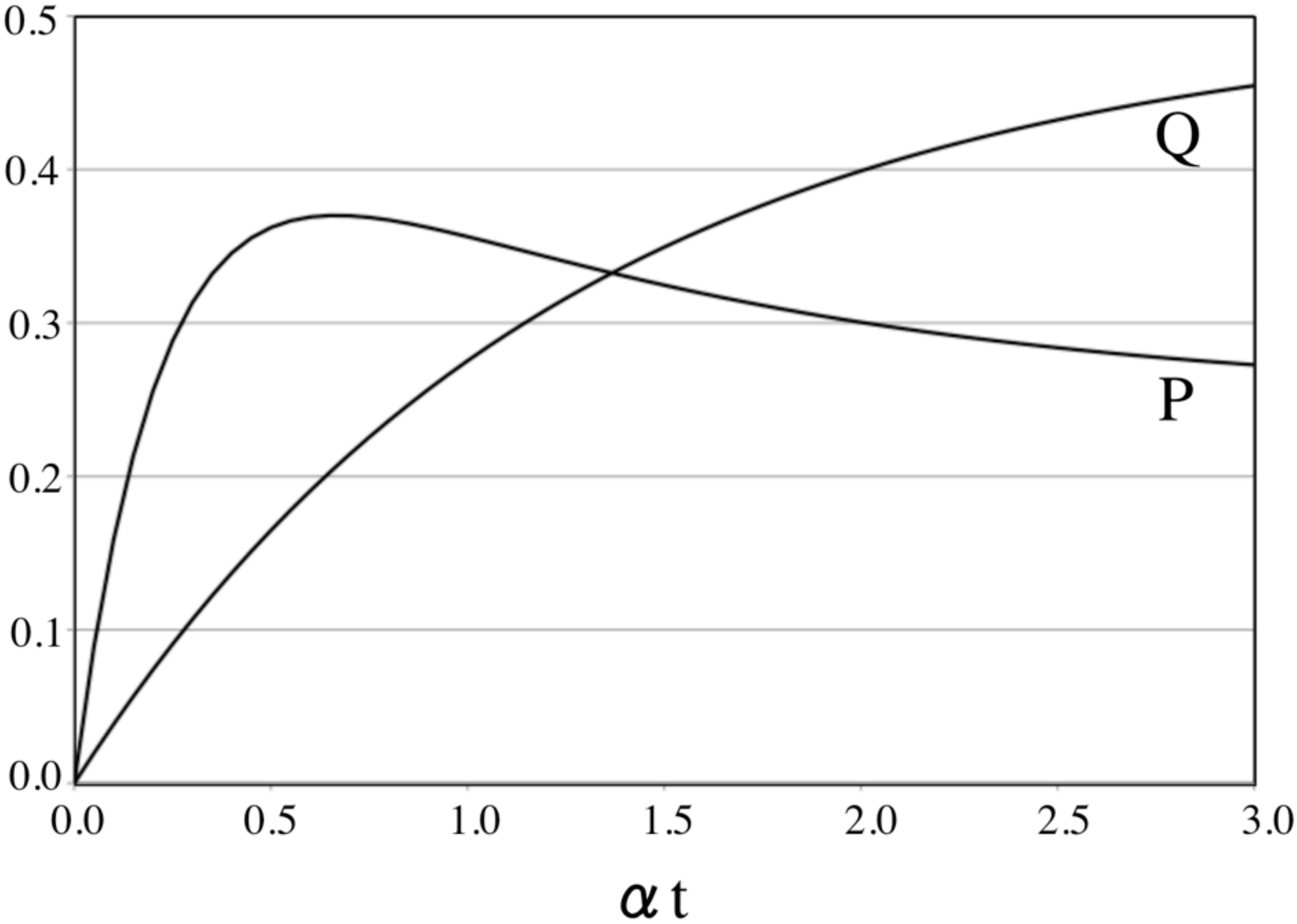


Table 15-5: Pattern of nucleotide substitutions for the human mitochondrial genome (from Kawai, Kikuchi, and Saitou, unpublished)

		N E W			
		A	C	T	G
O	A	-	0.0031	0.0030	0.0901
L	C	0.0070	-	0.0593	0.0028
D	T	0.0042	0.2566	-	0.0048
	G	0.5574	0.0080	0.0037	-

Table 15-6: Models of nucleotide substitution matrices incorporating nucleotide frequencies

(A) Equal-input model

		NEW			
		A	C	T	G
O	A	$1-\Sigma\lambda_{A\bullet}$	$\pi_C\alpha$	$\pi_T\alpha$	$\pi_G\alpha$
L	C	$\pi_A\alpha$	$1-\Sigma\lambda_{C\bullet}$	$\pi_T\alpha$	$\pi_G\alpha$
D	T	$\pi_A\alpha$	$\pi_C\alpha$	$1-\Sigma\lambda_{T\bullet}$	$\pi_G\alpha$
	G	$\pi_A\alpha$	$\pi_C\alpha$	$\pi_T\alpha$	$1-\Sigma\lambda_{G\bullet}$

(B) Equal-output model

		NEW			
		A	C	T	G
O	A	$1-\Sigma\lambda_{A\bullet}$	$\pi_A\alpha$	$\pi_A\alpha$	$\pi_A\alpha$
L	C	$\pi_C\alpha$	$1-\Sigma\lambda_{C\bullet}$	$\pi_C\alpha$	$\pi_C\alpha$
D	T	$\pi_T\alpha$	$\pi_T\alpha$	$1-\Sigma\lambda_{T\bullet}$	$\pi_T\alpha$
	G	$\pi_G\alpha$	$\pi_G\alpha$	$\pi_G\alpha$	$1-\Sigma\lambda_{G\bullet}$

(C) Hasegawa-Kishino-Yano model

		NEW			
		A	C	T	G
O	A	$1-\Sigma\lambda_{A\bullet}$	$\pi_C\beta$	$\pi_T\beta$	$\pi_G\alpha$
L	C	$\pi_A\beta$	$1-\Sigma\lambda_{C\bullet}$	$\pi_T\alpha$	$\pi_G\beta$
D	T	$\pi_A\beta$	$\pi_C\alpha$	$1-\Sigma\lambda_{T\bullet}$	$\pi_G\beta$
	G	$\pi_A\alpha$	$\pi_C\beta$	$\pi_T\beta$	$1-\Sigma\lambda_{G\bullet}$

(D) Tamura-Nei model

		NEW			
		A	C	T	G
O	A	$1-\Sigma\lambda_{A\bullet}$	$\pi_C\beta$	$\pi_T\beta$	$\pi_G\alpha_1$
L	C	$\pi_A\beta$	$1-\Sigma\lambda_{C\bullet}$	$\pi_T\alpha_2$	$\pi_G\beta$
D	T	$\pi_A\beta$	$\pi_C\alpha_2$	$1-\Sigma\lambda_{T\bullet}$	$\pi_G\beta$
	G	$\pi_A\alpha_1$	$\pi_C\beta$	$\pi_T\beta$	$1-\Sigma\lambda_{G\bullet}$

Table 15-7: Various models of nucleotide substitutions

(A) Kimura 3P (ref 15-17)

		NEW			
		A	C	T	G
O	A	$1-\Sigma\lambda_{A\bullet}$	γ	β	α
L	C	γ	$1-\Sigma\lambda_{C\bullet}$	α	β
D	T	β	α	$1-\Sigma\lambda_{T\bullet}$	γ
	G	α	β	γ	$1-\Sigma\lambda_{G\bullet}$

(B) Takahata-Kimura 4P (ref 15-18)

		NEW			
		A	C	T	G
O	A	$1-\Sigma\lambda_{A\bullet}$	$\theta\alpha$	β	α
L	C	$\theta\gamma$	$1-\Sigma\lambda_{C\bullet}$	γ	β
D	T	β	α	$1-\Sigma\lambda_{T\bullet}$	$\theta\alpha$
	G	γ	β	$\theta\gamma$	$1-\Sigma\lambda_{G\bullet}$

(C) Takahata-Kimura 5P (ref 15-18)

		NEW			
		A	C	T	G
O	A	$1-\Sigma\lambda_{A\bullet}$	δ	β	α
L	C	ε	$1-\Sigma\lambda_{C\bullet}$	γ	β
D	T	β	α	$1-\Sigma\lambda_{T\bullet}$	δ
	G	γ	β	ε	$1-\Sigma\lambda_{G\bullet}$

(D) Gojobori-Ishii-Nei 6P (ref 15-19)

		NEW			
		A	C	T	G
O	A	$1-\Sigma\lambda_{A\bullet}$	α	α_1	α
L	C	β	$1-\Sigma\lambda_{C\bullet}$	β	α_2
D	T	β_1	α	$1-\Sigma\lambda_{T\bullet}$	α
	G	β	β_2	β	$1-\Sigma\lambda_{G\bullet}$

Figure 15-6: Relationship of various nucleotide substitution models

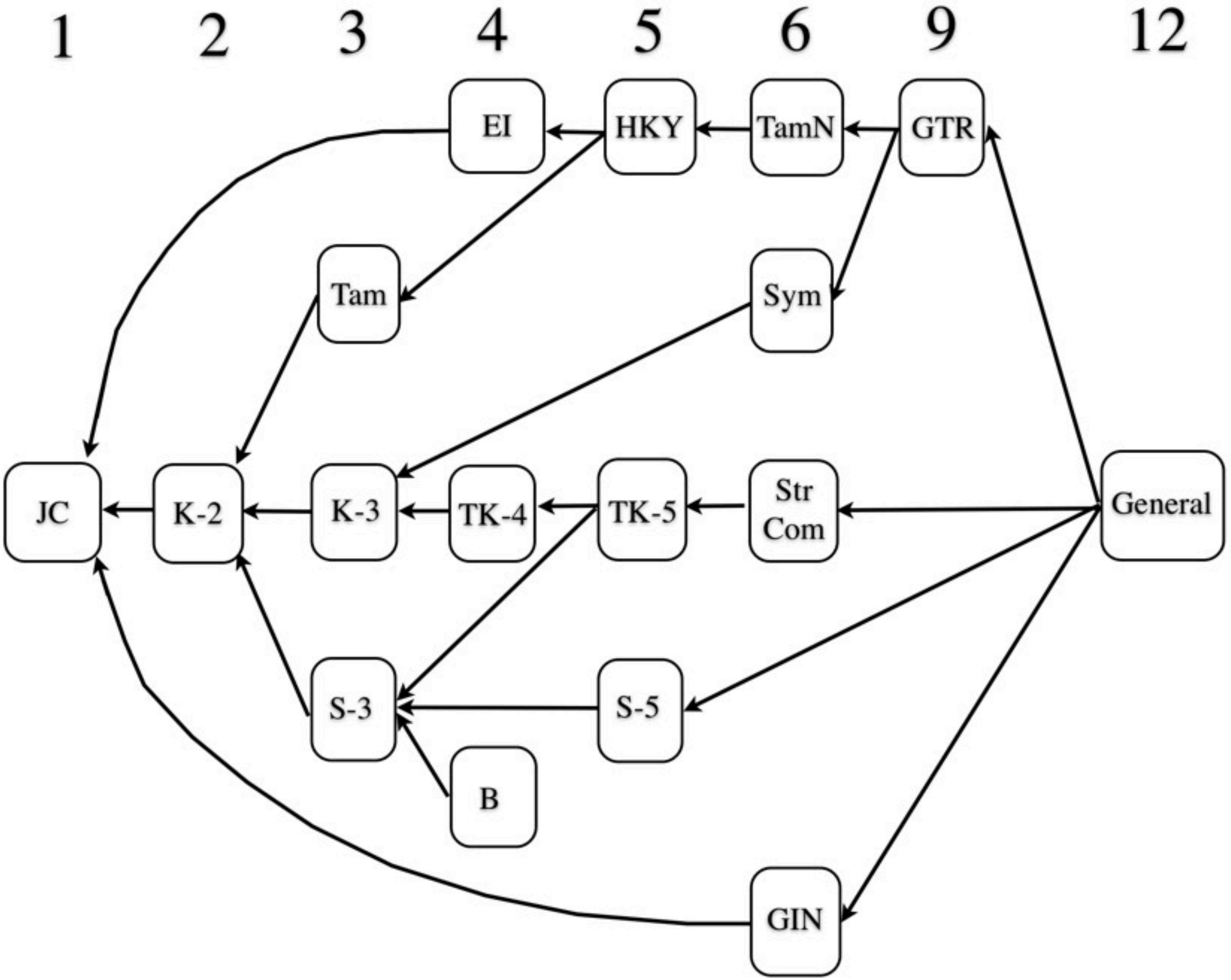


Table 15-9: Amino acid substitution matrix for vertebrate mitochondrial DNA coded proteins

	Ala	Arg	Asn	Asp	Cys	Gln	Glu	Gly	His	Ile	Leu	Lys	Met	Phe	Pro	Ser	Thr	Trp	Tyr	Val
Ala	***	2	9	2	3	1	2	64	4	84	35	0	63	5	26	243	380	0	2	67
Arg	7	***	4	0	7	51	0	13	38	0	23	29	0	3	14	5	2	6	0	3
Asn	17	2	***	134	2	31	17	25	114	22	30	102	17	4	39	327	161	3	44	4
Asp	7	0	279	***	0	14	135	32	34	6	1	3	0	4	4	42	24	2	3	0
Cys	35	20	14	0	***	9	0	18	41	42	60	0	0	40	8	200	140	102	75	0
Gln	2	38	49	10	2	***	73	4	148	10	57	98	25	16	68	43	78	0	11	5
Glu	7	0	28	105	0	75	***	11	12	0	0	59	0	0	4	39	12	0	6	9
Gly	82	4	18	11	2	2	5	***	0	6	2	3	1	0	0	81	7	2	0	1
His	10	25	159	23	9	132	10	10	***	13	17	15	0	21	21	40	44	2	198	0
Ile	70	0	10	1	3	3	0	4	4	***	465	2	216	40	6	23	285	0	9	461
Leu	15	3	7	0	2	9	0	1	3	241	***	2	244	118	20	49	90	8	12	33
Lys	0	24	173	2	0	105	61	8	18	8	12	***	34	5	22	60	102	8	15	1
Met	86	0	13	0	0	12	0	1	0	354	772	15	***	42	8	73	408	6	9	155
Phe	6	1	3	1	4	7	0	0	10	57	328	2	37	***	8	42	23	2	124	2
Pro	35	5	28	2	1	31	2	0	11	10	62	10	7	9	***	100	96	1	4	3
Ser	246	1	179	11	18	15	13	64	16	28	116	20	54	36	76	***	452	8	16	0
Thr	315	1	72	5	10	22	3	5	14	283	172	27	247	16	60	369	***	4	8	80
Trp	0	4	4	2	2	0	0	4	2	0	48	6	12	5	3	21	12	***	7	2
Tyr	4	0	53	2	14	8	4	0	169	23	62	10	14	227	7	34	22	6	***	2
Val	112	1	4	0	0	3	5	2	0	918	128	0	118	3	4	0	160	2	2	***

Table 16-1: Examples of distance matices

(A) When the evolutionary rate is constant

	A	B	C	D	E	F
A	0	2	6	4	16	16
B	2	0	6	4	16	16
C	6	6	0	6	16	16
D	4	4	6	0	16	16
E	16	16	16	16	0	10
F	16	16	16	16	10	0

(B) When all distances are purely additive (from ref 16-1)

	1	2	3	4	5	6	7	8
1	0	7	8	11	13	16	13	17
2	7	0	5	8	10	13	10	14
3	8	5	0	5	7	10	7	11
4	11	8	5	0	8	11	8	12
5	13	10	7	8	0	6	6	10
6	16	13	10	11	5	0	9	13
7	13	10	7	8	6	9	0	8
8	17	14	11	12	10	13	8	0

(C) Distance matrix estimated from real nulceotide sequence data (from ref 16-2)

```
------------------------------------------------------------------------------------
            1         2         3         4         5         6         7         8         9
------------------------------------------------------------------------------------
   2    0.0516
   3    0.0550    0.0031
   4    0.0483    0.0221    0.0253
   5    0.0582    0.0651    0.0685    0.0549
   6    0.0094    0.0416    0.0450    0.0384    0.0549
   7    0.0125    0.0584    0.0619    0.0551    0.0651    0.0157
   8    0.0284    0.0687    0.0722    0.0654    0.0754    0.0317    0.0285
   9    0.0925    0.1221    0.1259    0.1185    0.1370    0.0820    0.0786    0.0927
  10    0.1921    0.2183    0.2228    0.2054    0.2309    0.1798    0.1795    0.1833    0.1860
------------------------------------------------------------------------------------
```

OTU ID; 1 = *M. m. domesticus* functional gene, 2 = *M. m. domesticus* pseudogene, 3 = *M. m. castaneus* pseudogene, 4 = M. spicilegus pseudogene, 5 = *M. leggada* pseudogene, 6 = *M. m. domesticus* functional gene, 7 = *M. leggada* functional gene, 8 = *M. platythrix* functional gene, 9 = *Rattus norvegicus* functional gene, 10 = *Homo sapiens* functional gene.

Table 16-2: Operation of UPGMA for distance matrix of Table 16-1A

(A) After OTUs A and B are clustered

	AB	C	D	E	F
AB	0	6	4	16	16
C	6	0	6	16	16
D	4	6	0	16	16
E	16	16	16	0	10
F	16	16	16	10	0

(B) After OTUs AB and D are clustered

	ABD	C	E	F
ABD	0	6	16	16
C	6	0	16	16
E	16	10	0	10
F	16	16	10	0

(C) After OTUs ABD and C are clustered

	ABDC	E	F
ABDC	0	16	16
E	16	0	10
F	16	10	0

(D) After OTUs E and F are clustered

	ABDC	EF
ABDC	0	16
EF	16	0

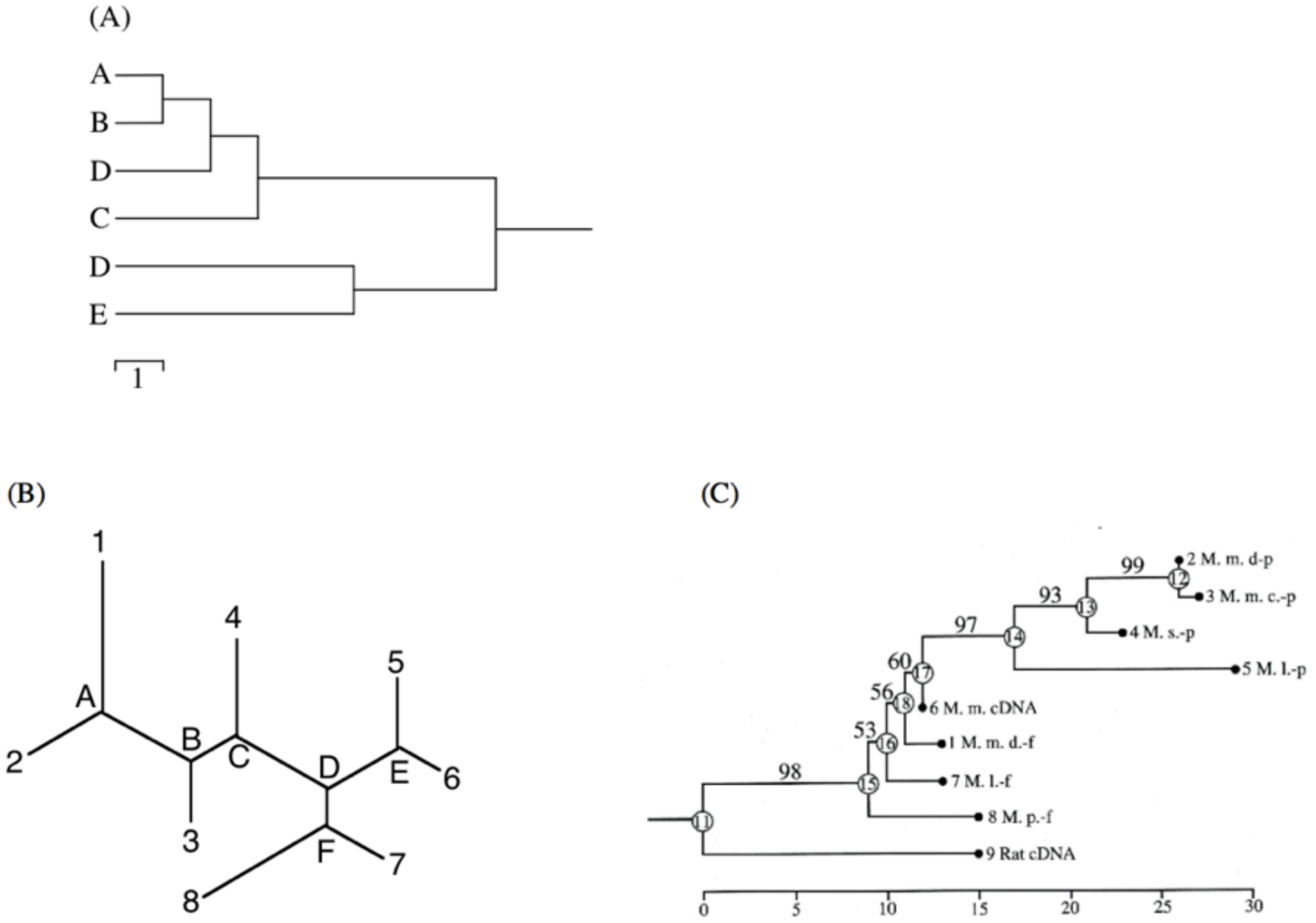


Figure 16-1: Phylogenetic trees corresponding to distance matices shown in Table 16-1

Figure 16-2: A phylogenetic tree constructed by using UPGMA from distance matix of Table 16-1B

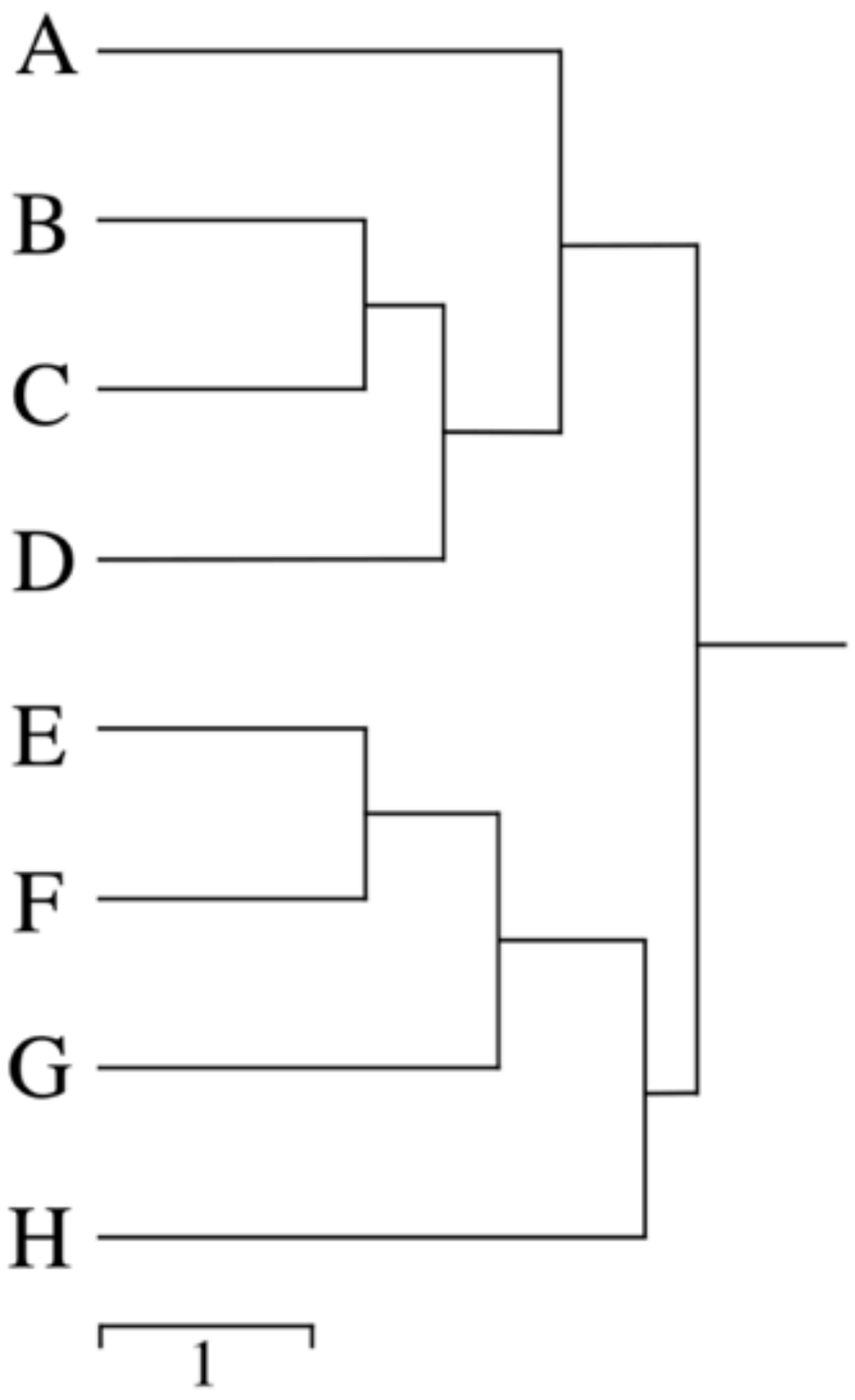


Figure 16-6: Two types of multifurcating trees
(A) Star phylogeny with no interior branch. (B) Multifurcating tree with only one interior branch.

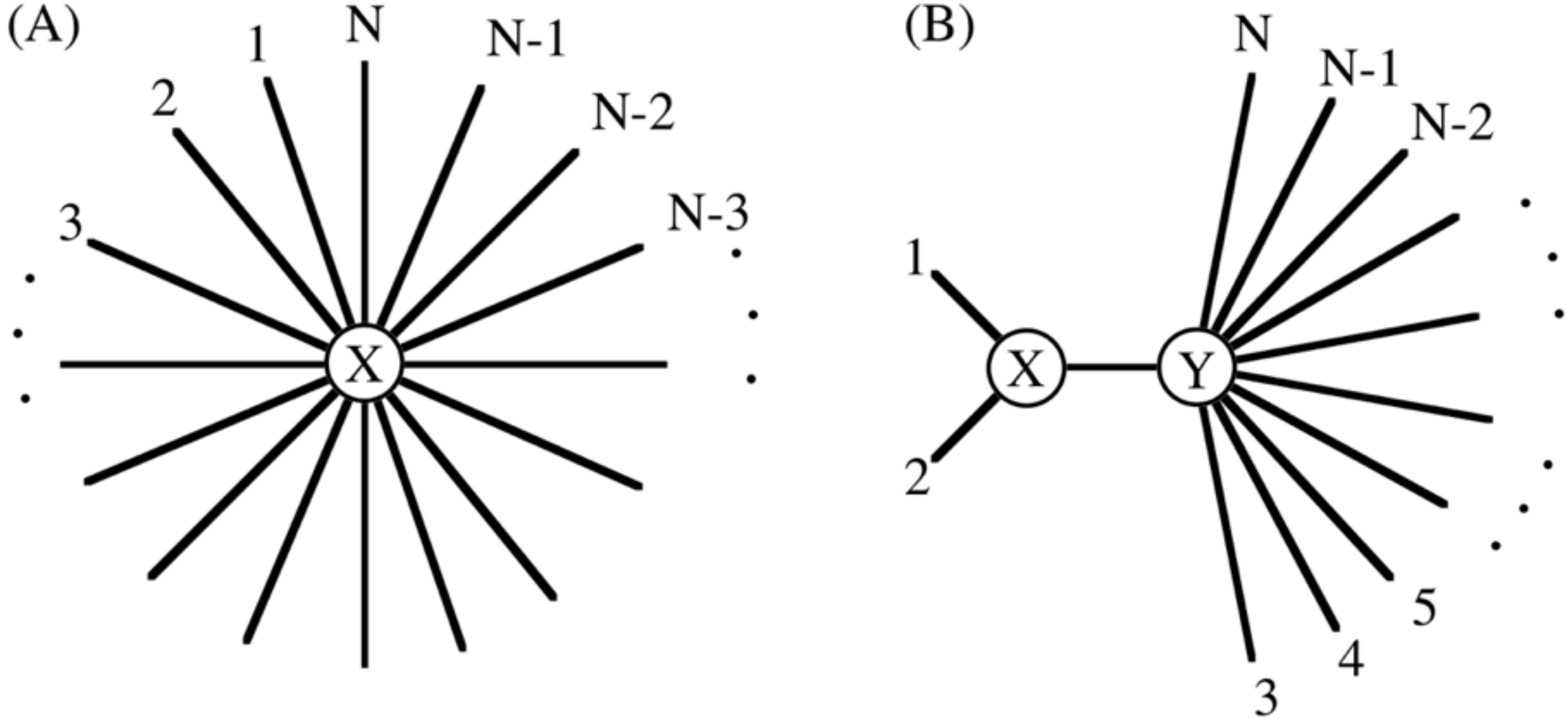


Figure 16-7: Construction of the neighbor-joining tree for 8 OTUs (from ref 16-1)

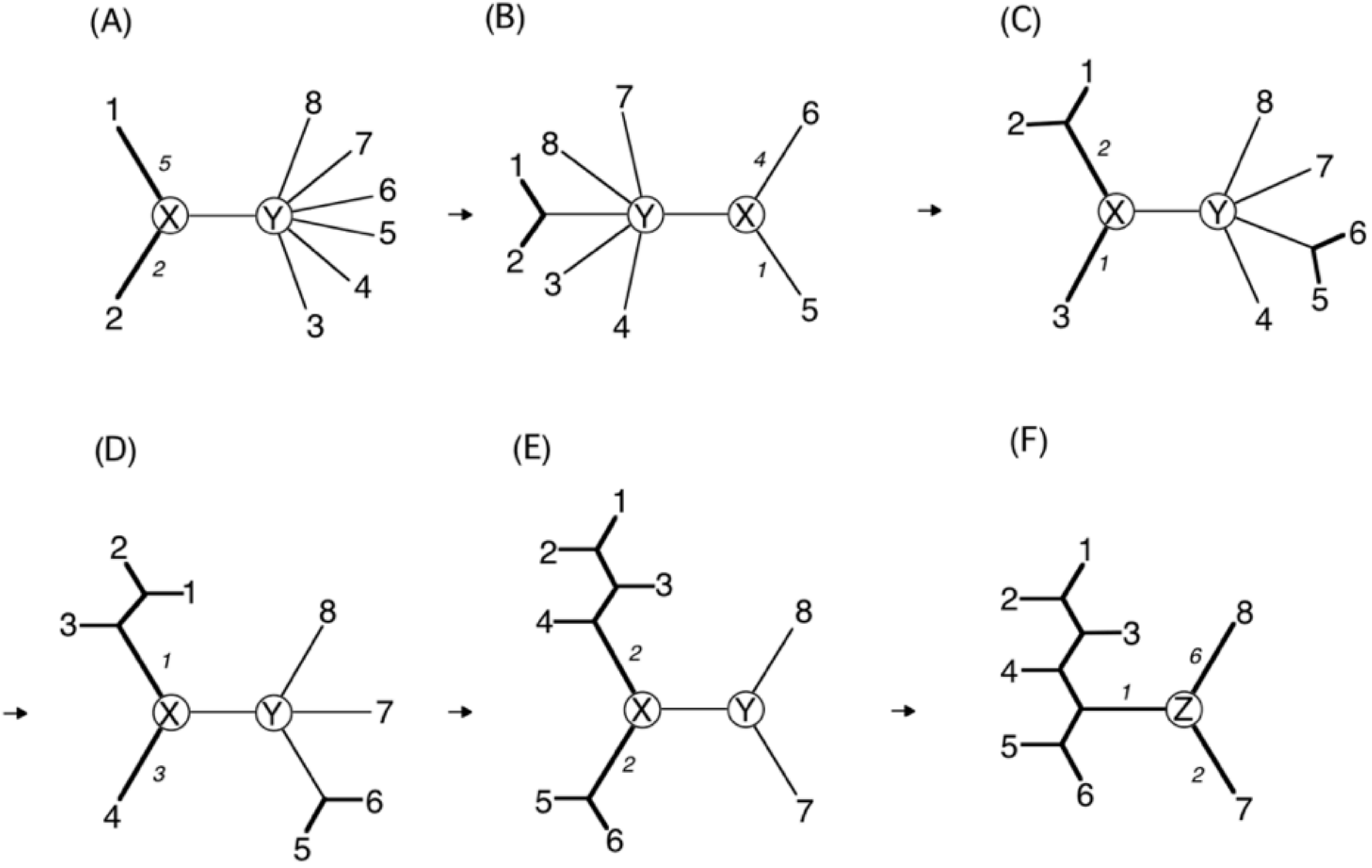


Figure 16-8: Expected tree for 10 sequences which produced the distance matrix of Table 16-1C

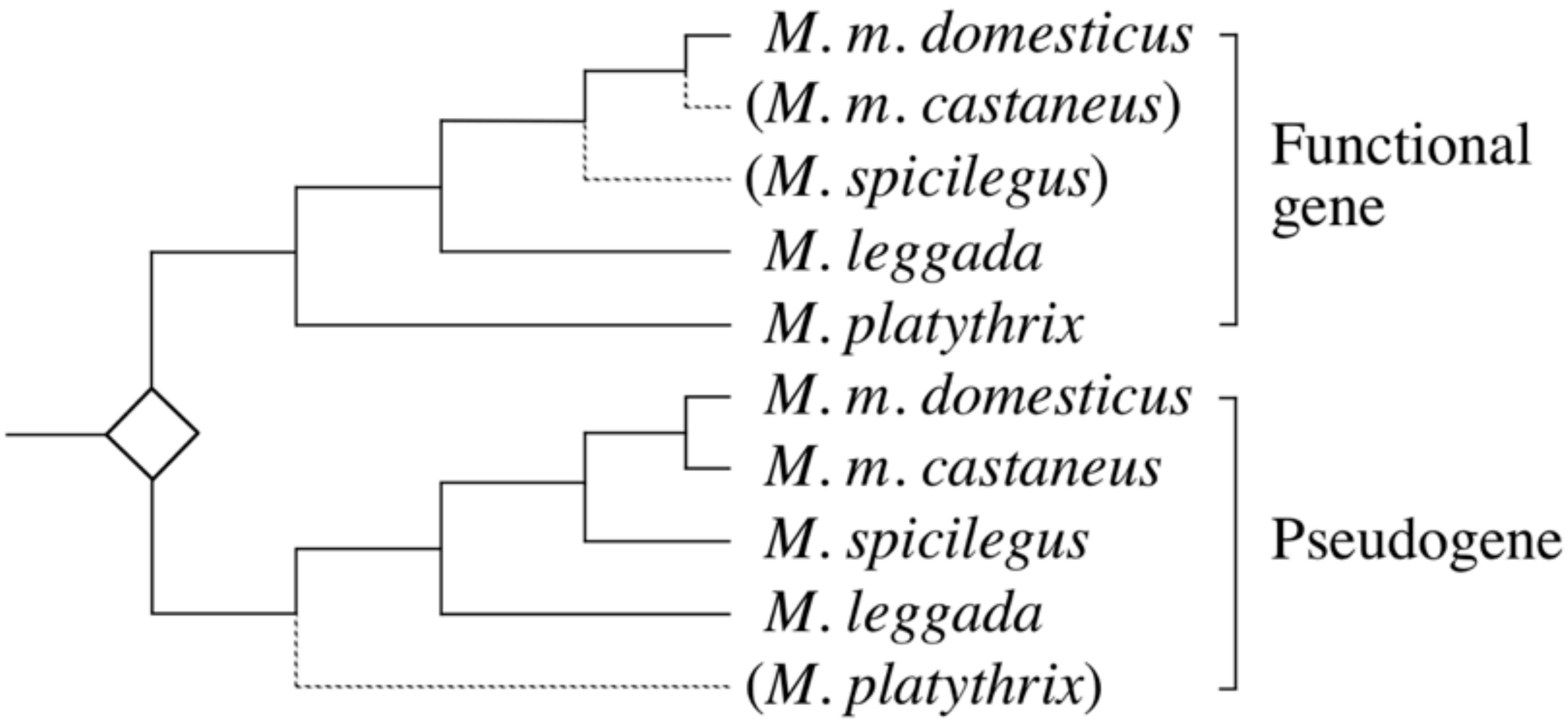


Figure 16-9: A phylogenetic network with two splits for four OTUs

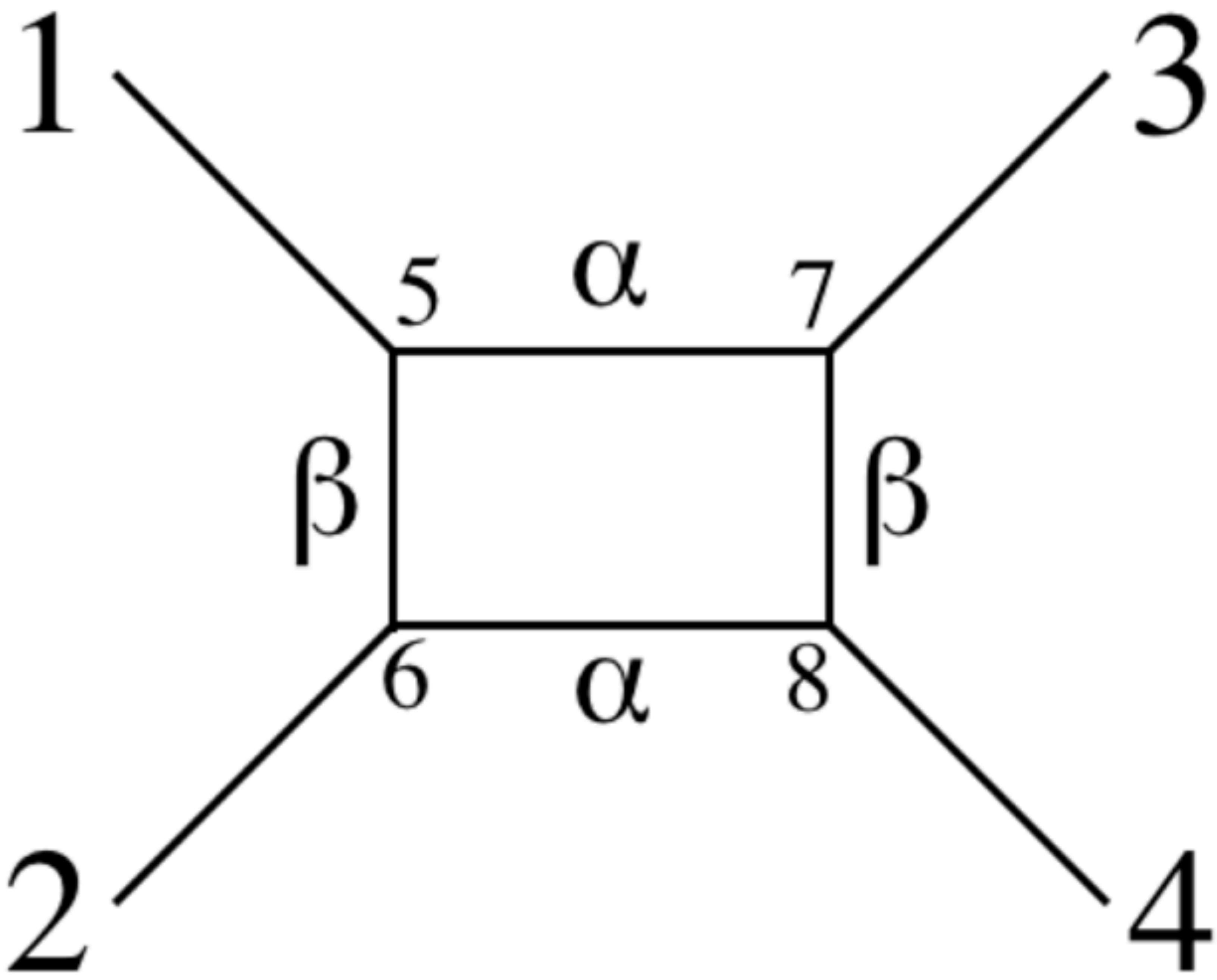


Figure 16-10: Two phylogenetic networks constructed from a distance matrix data of 2? gibbon ABO blood group gene partial sequences (from ref 16-43)
(A) When the Split Decomposition method (ref 16-42) was used. (B) When the Neighbor-Net method (ref 16-44; ref 16-45) was used

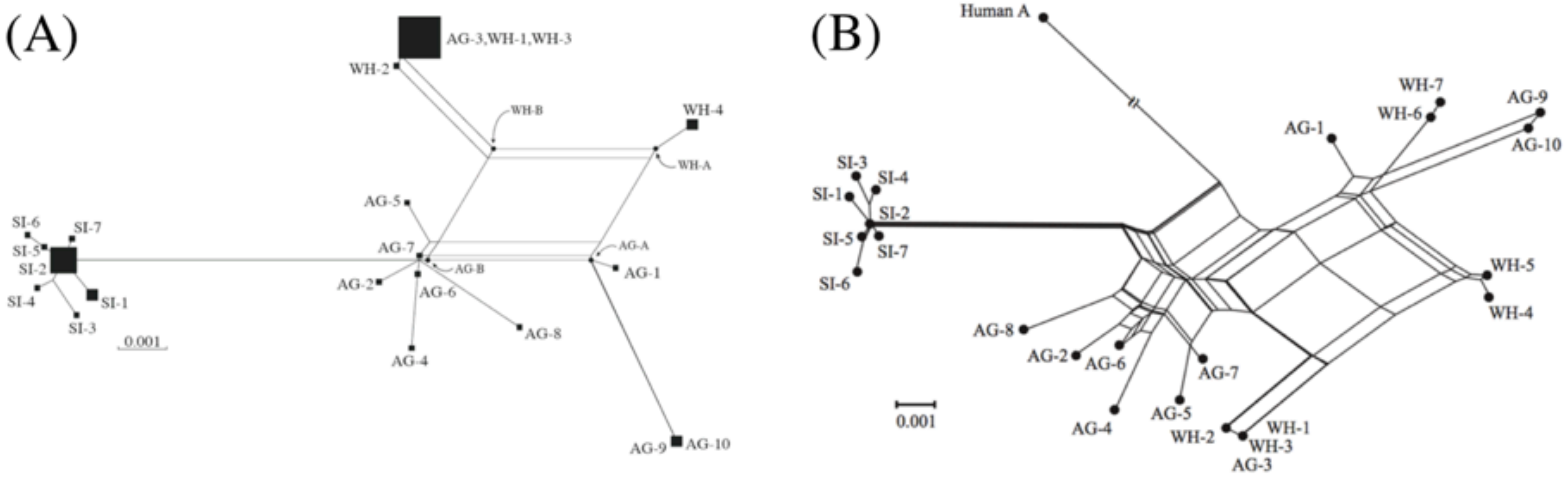
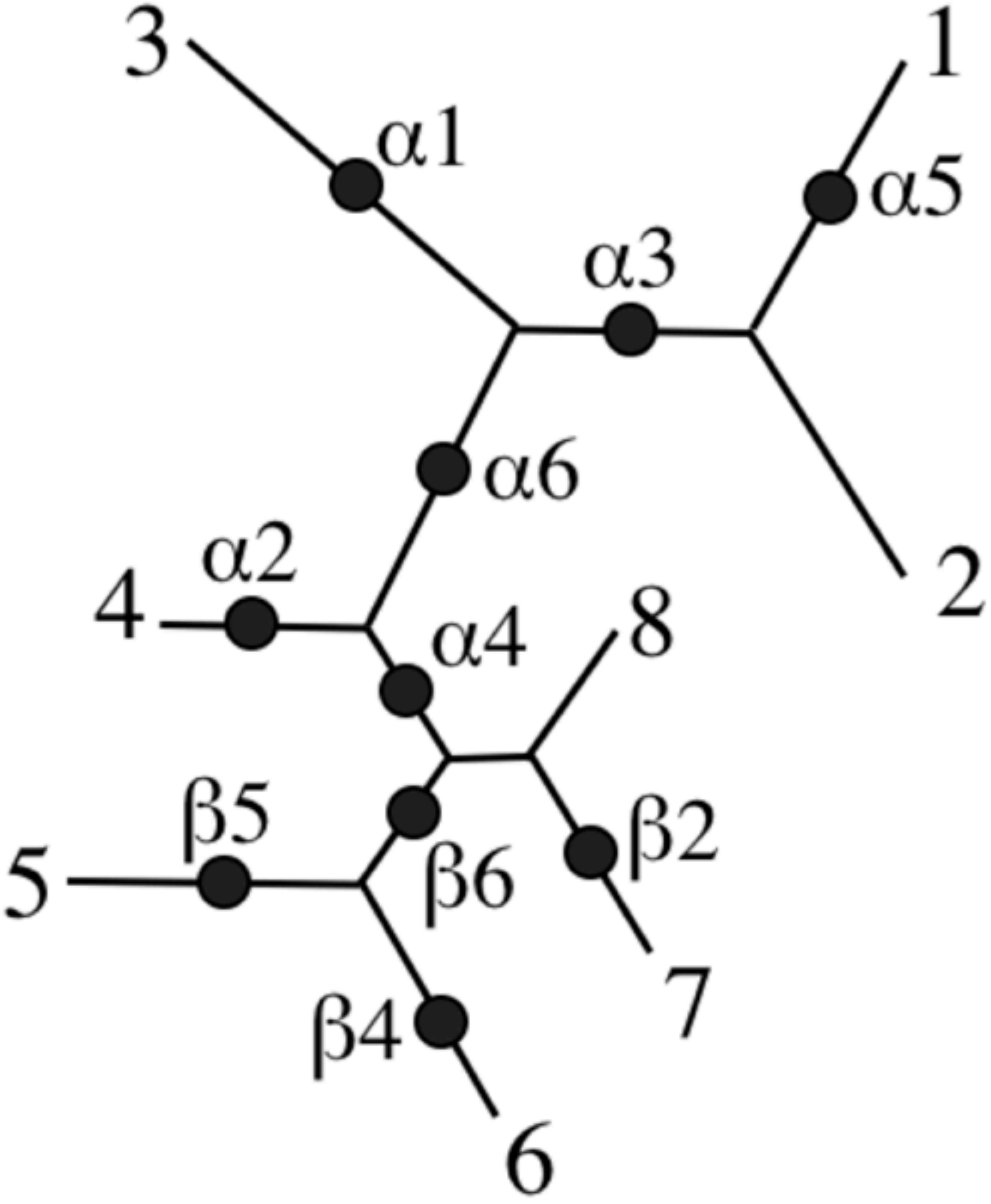


Figure 16-11: The maximum parsimony tree (left) for the sequence data (right)



```
==========
S 12345678
----------
1 AAGAAAAA
2 TTTCTTGT
3 CCGGGGGG
4 GGGGTCTT
5 ACCCACCC
6 CCCGTTGG
==========
```

Table 16-7: Ten informative configurations for five nucleotide sequences

	A	B	C	D	E
1	X	X	*	*	*
2	X	*	X	*	*
3	X	*	*	X	*
4	X	*	*	*	X
5	*	X	X	*	*
6	*	X	*	X	*
7	*	X	*	*	X
8	*	*	X	X	*
9	*	*	X	*	X
10	*	*	*	X	X

Note: * and X are diffent nucleotides with each ohter.

Table 16-8: Application of the SSJ algorithm to an ideal sequence data

Step 0: Initial sequences after elimination of invarianct sites and sequence identity check

```
=========================
  00000000011111111112222
  12345678901234567890123
-------------------------
A aacgtttcattgagatacgtgca
B ..........gc...........
C c.........g.g..........
D cta.......g..tt........
E ctacga....g....cg......
F ctacgagt..g......a.....
G ctacgagt..g.......ca...
H ctacgag.c.g.........t..
I ctacgag.cgg.........g.
J ctacgag.cgg..........c
=========================
```

Step 1: After elimination of non-informative sites

```
============
  0000000001
  1234567890
------------
A aacgtttcat
B ..........
C c.........
D cta.......
E ctacga....
F ctacgagt..
G ctacgagt..
H ctacgag.c.
I ctacgag.cg
J ctacgag.cg
============
```

Step 2: After joining identical sequences

```
=============
   0000000001
   1234567890
-------------
AB aacgtttcat
C  c.........
D  cta.......
E  ctacga....
FG ctacgagt..
H  ctacgag.c.
IJ ctacgag.cg
=============
```

Step 3: After elimination of non-informative sites

```
==========
   0000000
   2345679
----------
AB acgttta
C  .......
D  ta.....
E  tacga..
FG tacgag.
H  tacgagc
IJ tacgagc
==========
```

Step 4:

```
===========
    2345679
-----------
ABC acgttta
D   ta.....
E   tacga..
FG  tacgag.
HIJ tacgagc
===========
```

Step 5:

```
========
    4567
--------
ABC gttt
D   ....
E   cga.
FG  cgag
HIJ cgag
========
```

Step 6:

```
==========
      4567
----------
ABCD  gttt
E     cga.
FGHIJ cgag
==========
```

Figure 16-12: The maximum parsimony tree for sequence data shown in Table 16-8

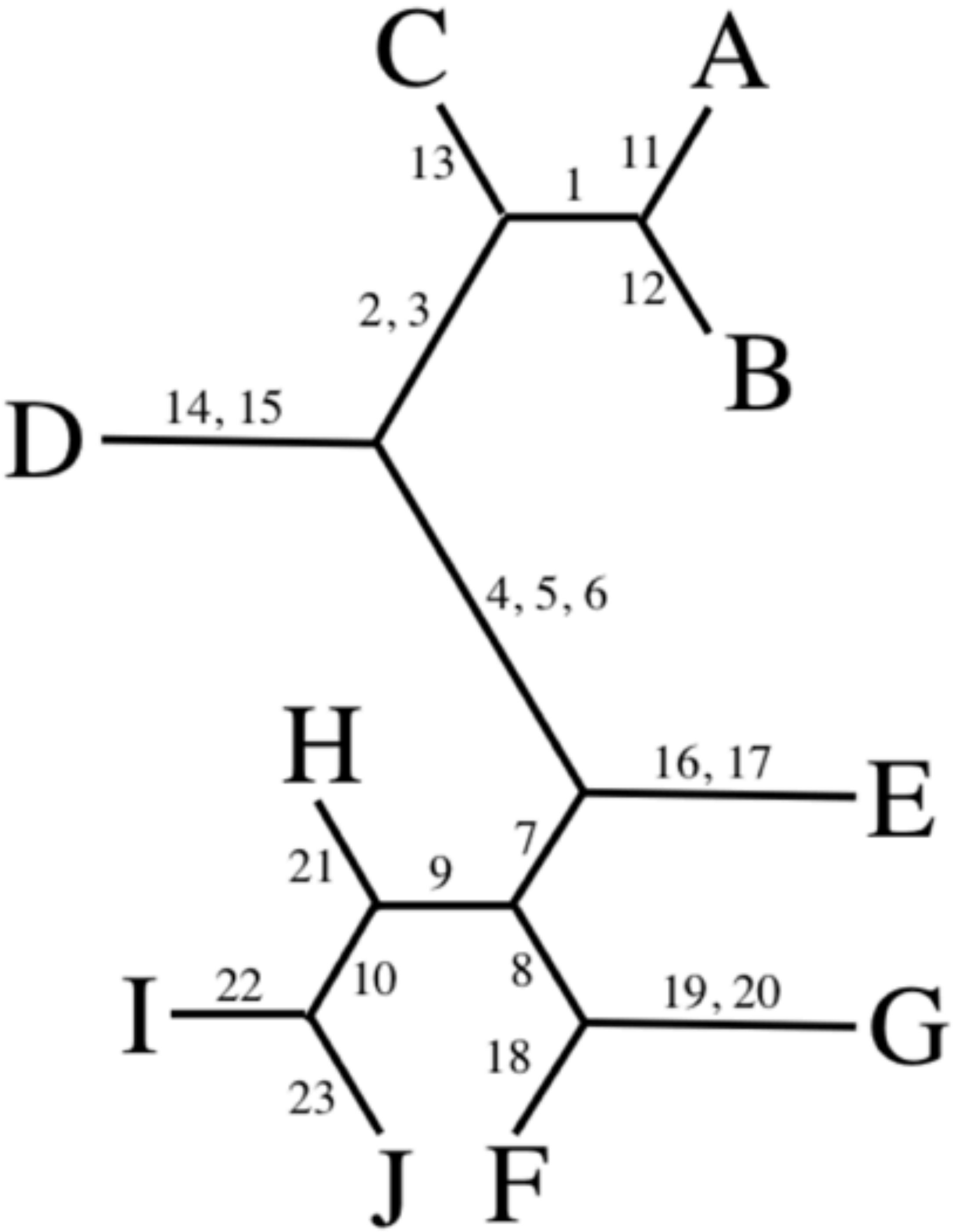


Figure 16-14: Likelihood values for three possible trees with four sequences (from ref 16-71)

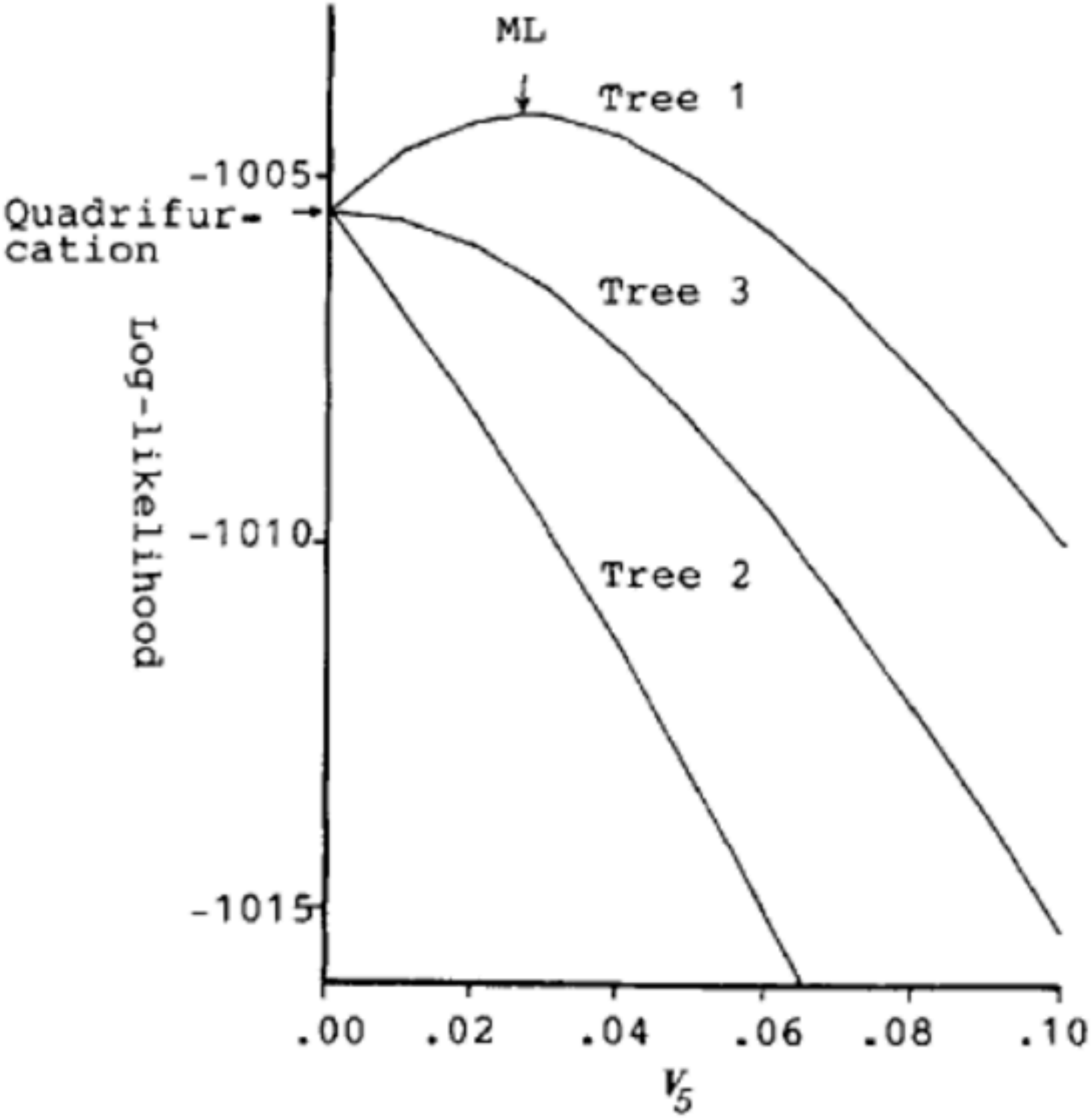


Figure 16-15: Application of Saitou's NJ-like stepwise clustering search using ML method (from ref 16-74)

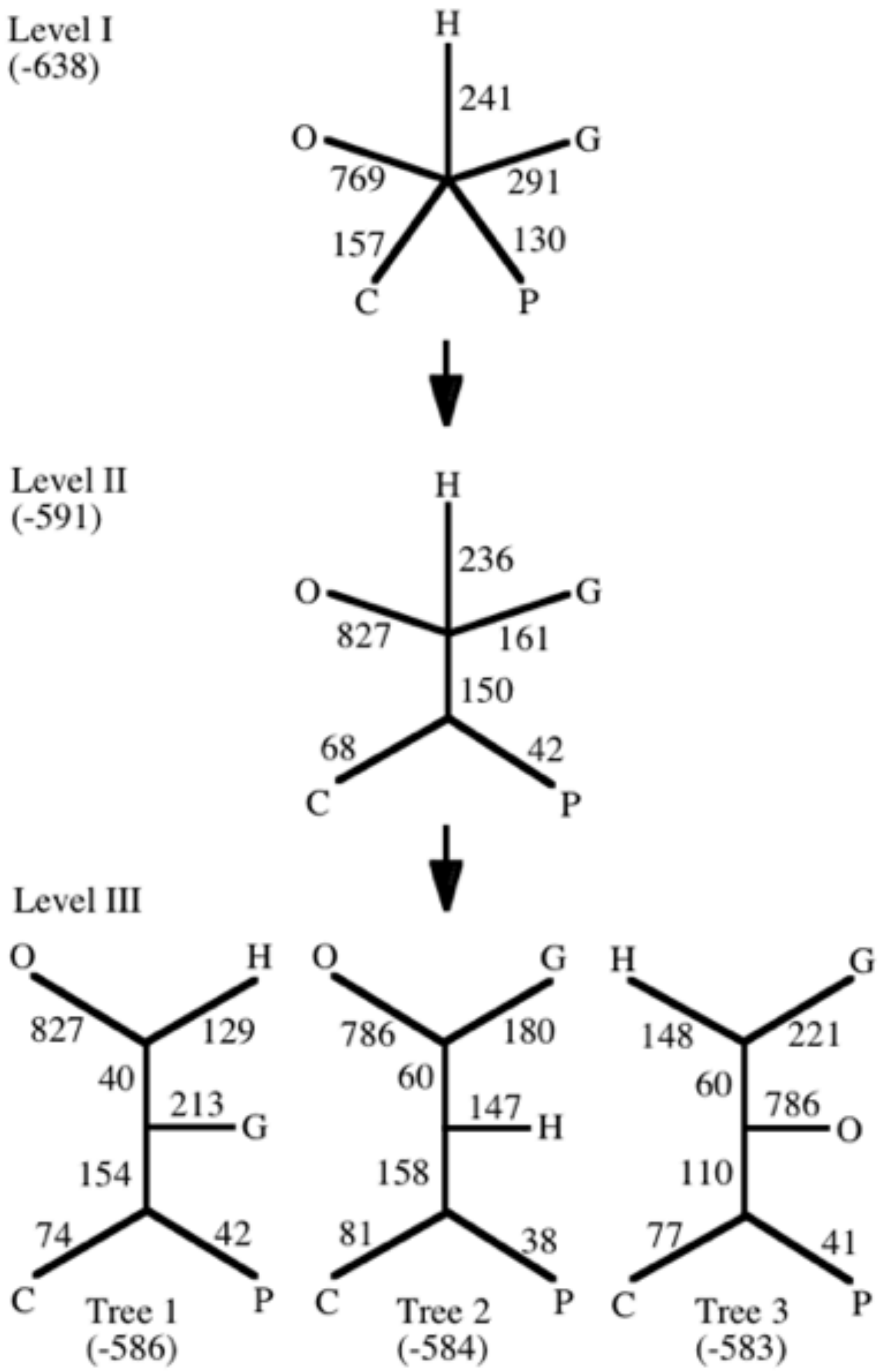
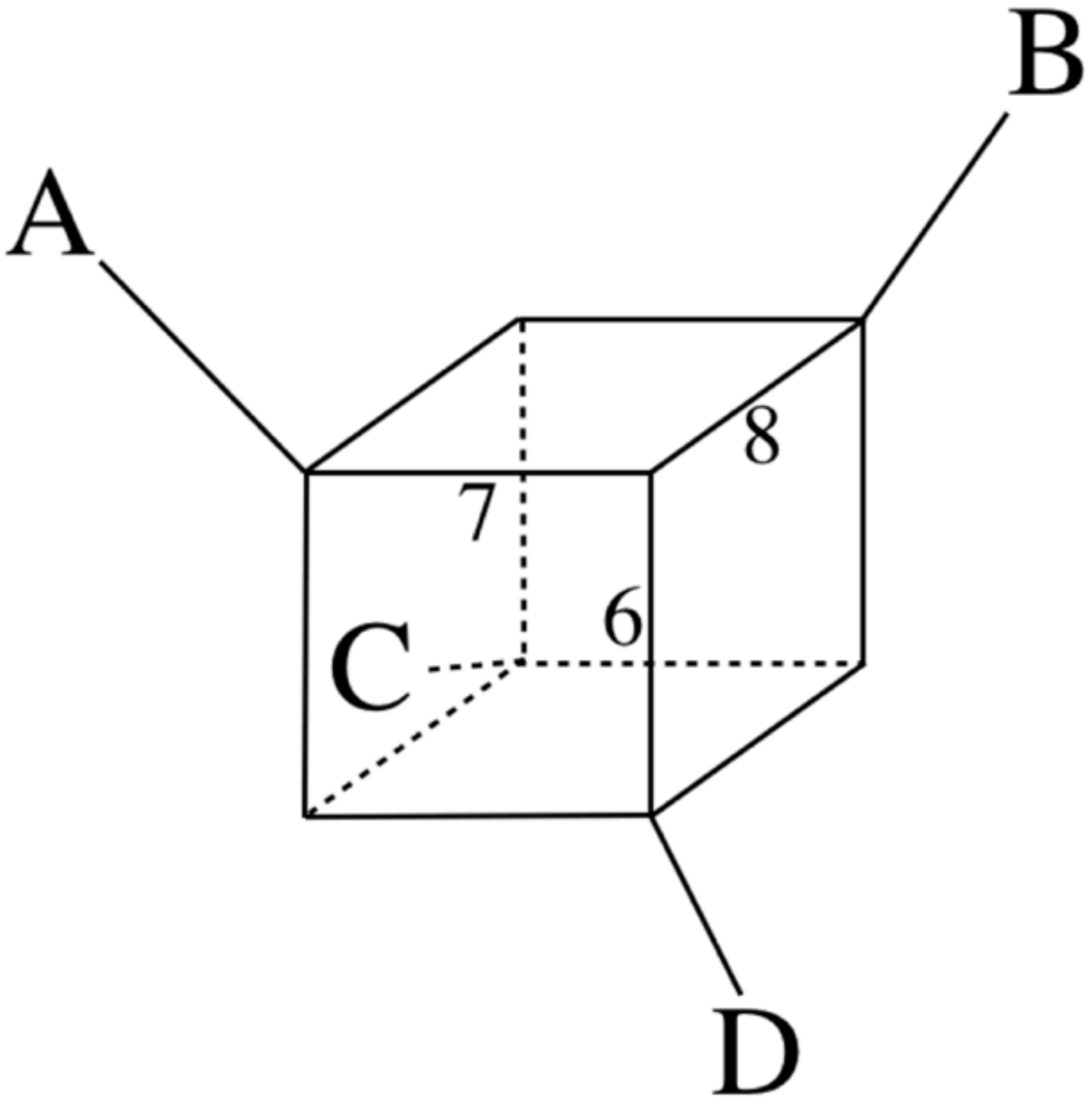


Figure 16-17: A phylogenetic network with three splits for four sequences



Chapter 17 Population Genomics

17.1 Evolutionary distances between populations

17.1.1 Distance between populations based on gene genealogy

17.1.2 Distance between species based on gene genealogy

17.1.3 Distance between populations based on allele frequency differences

17.1.4 Evolutionary distance between genomes

17.2 Mitochondrial DNA population genomics

17.2.1 Inference of gene genealogy

17.2.2 Population size fluctuation

17.2.3 Estimation of nucleotide substitution patterns

17.3 Population genomics of prokaryotes

17.4 Population genomics of nuclear genomes

17.4.1 Relationship of individuals and populations

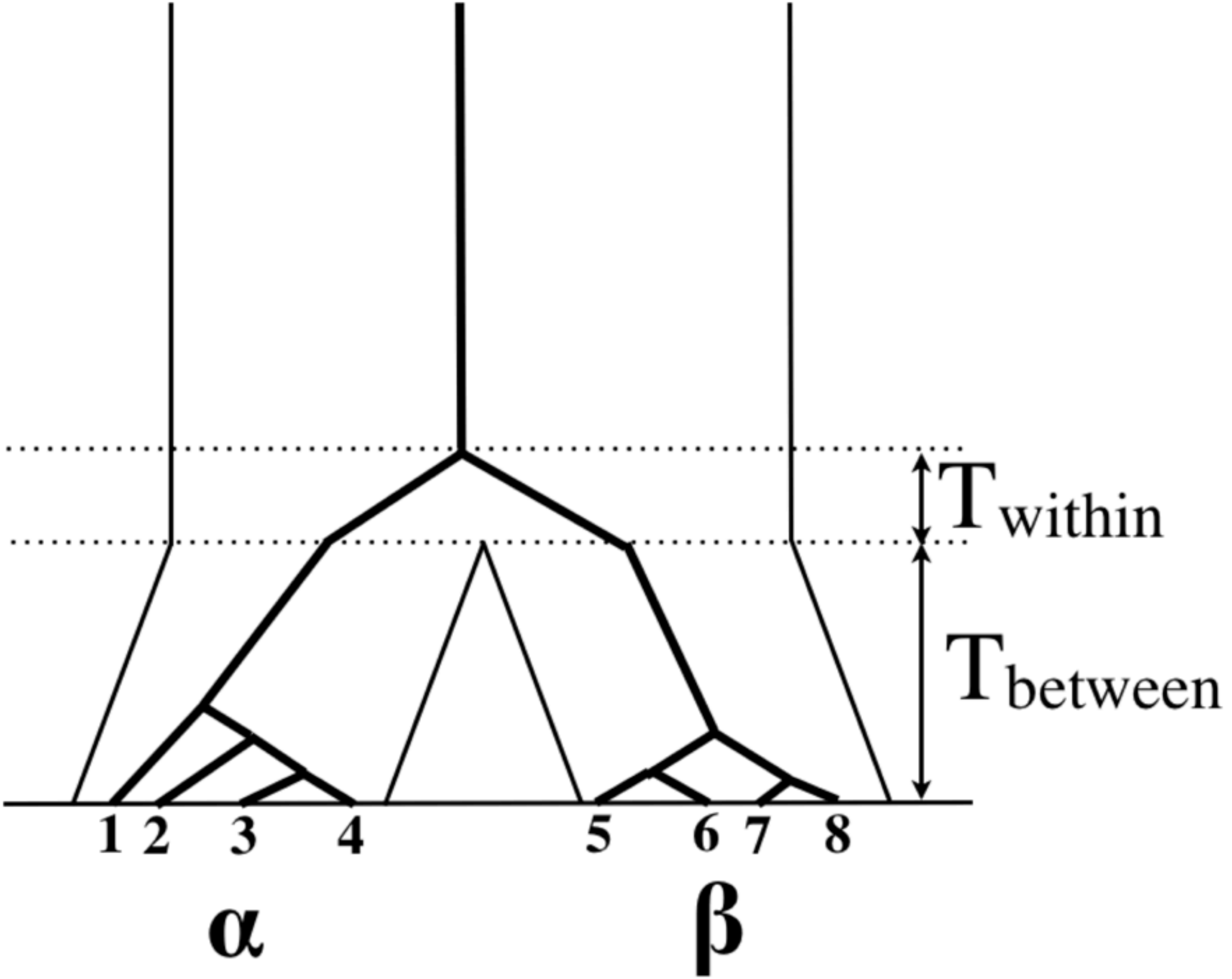
17.4.2 Admixture

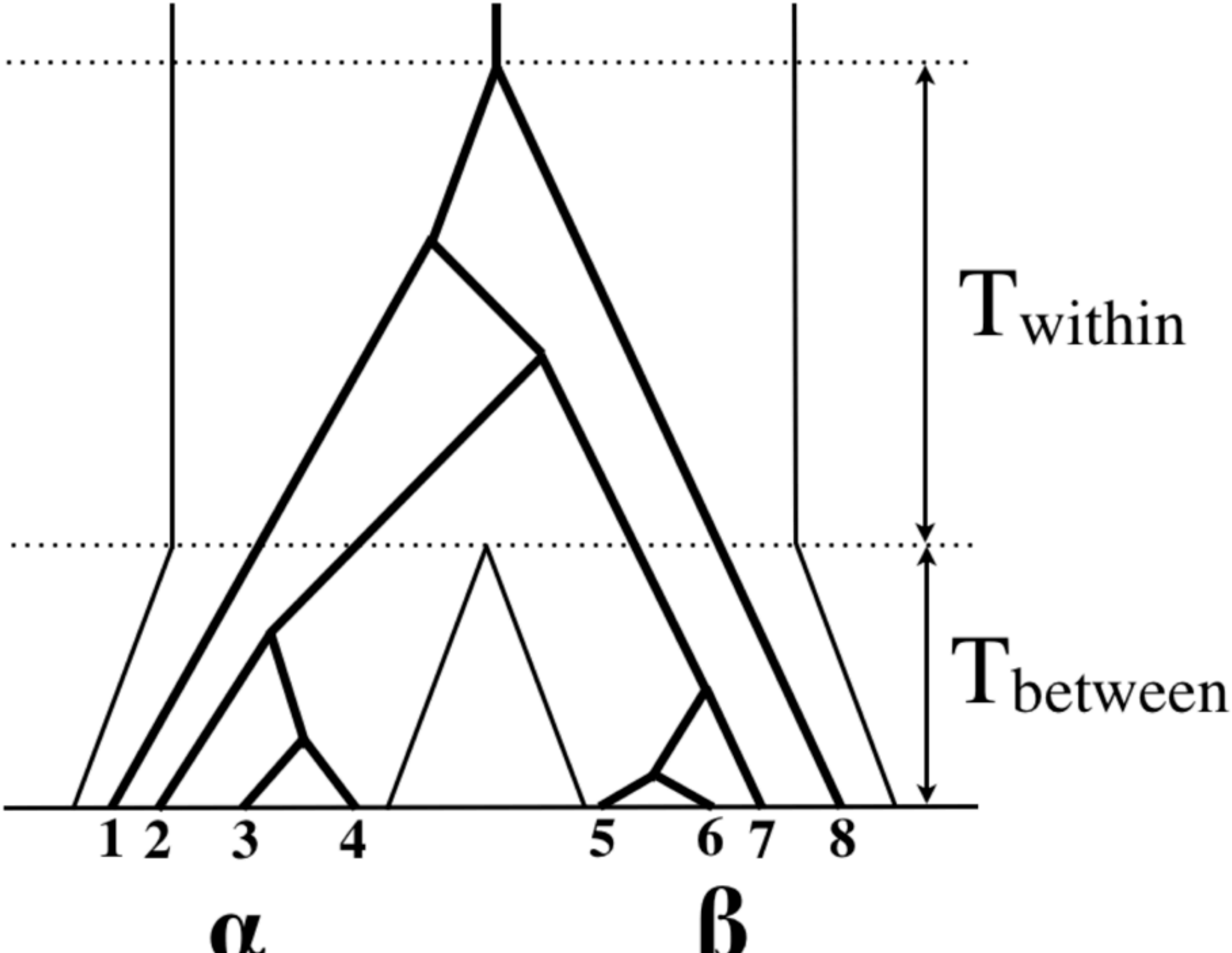
17.4.3 Introgression

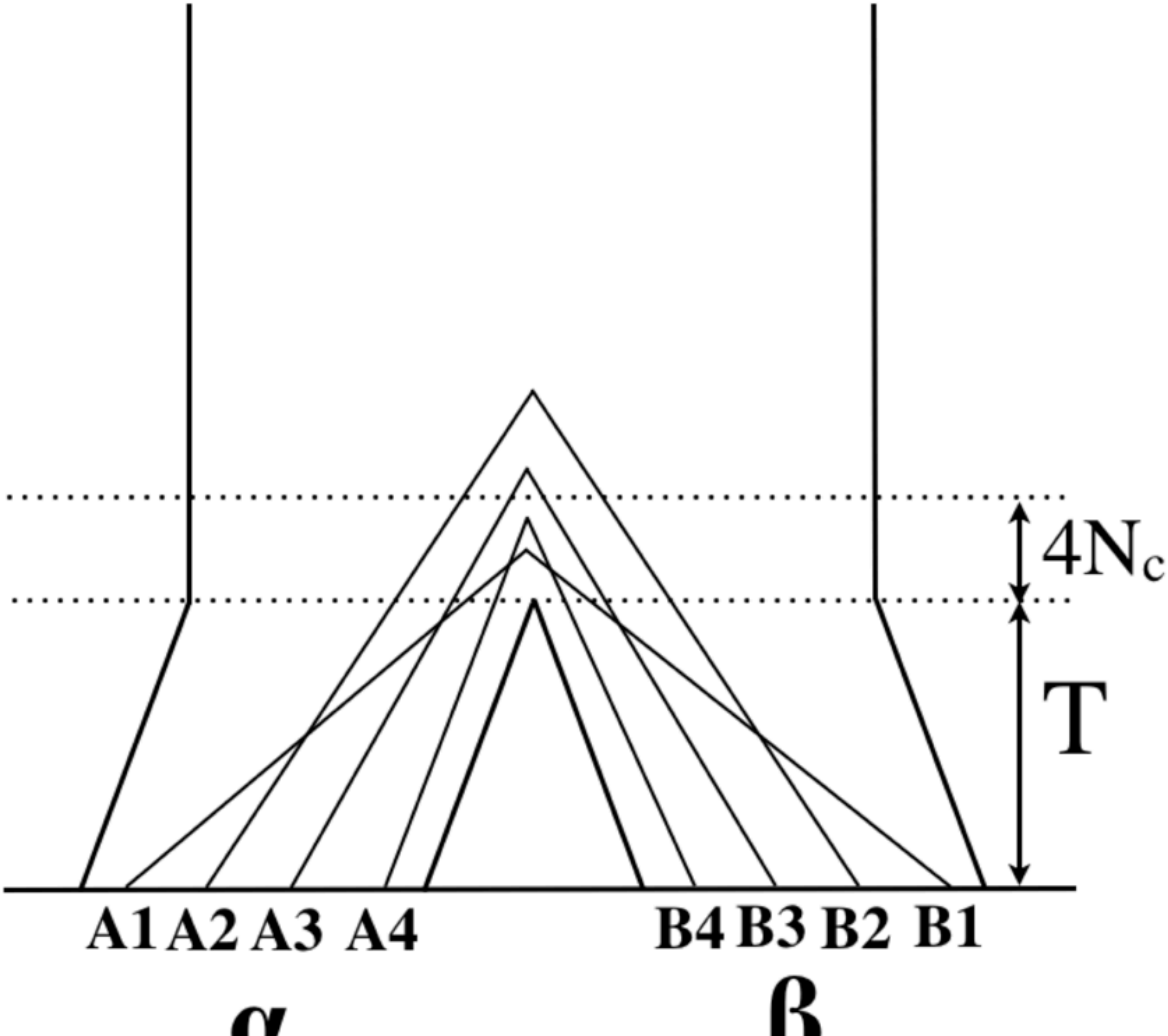
17.4.4 Population size fluctuation

17.4.5 Genome wide association study

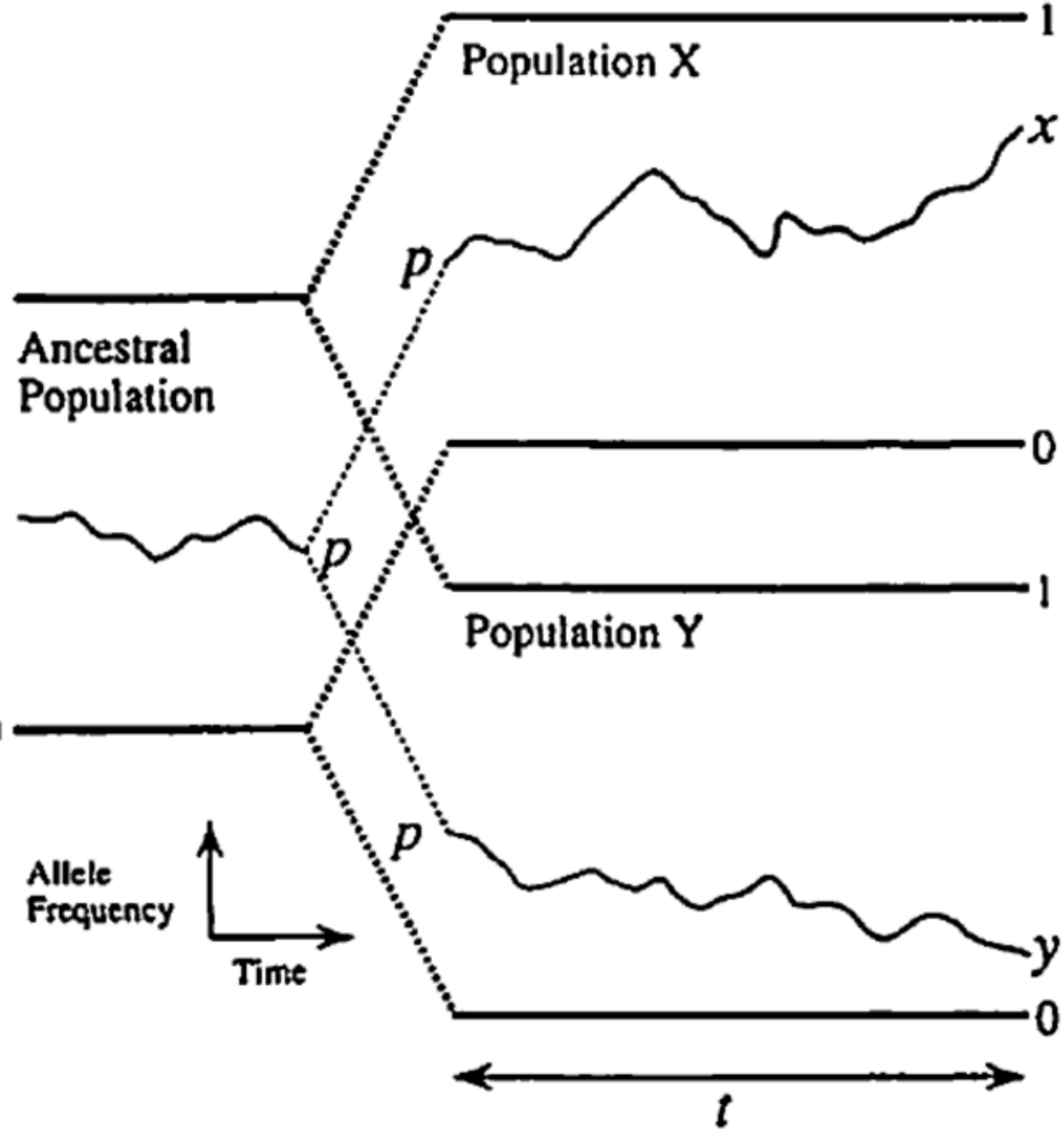
Figure 17-1: A schematic gene genealogy of two popualtions which differenciated long time ago



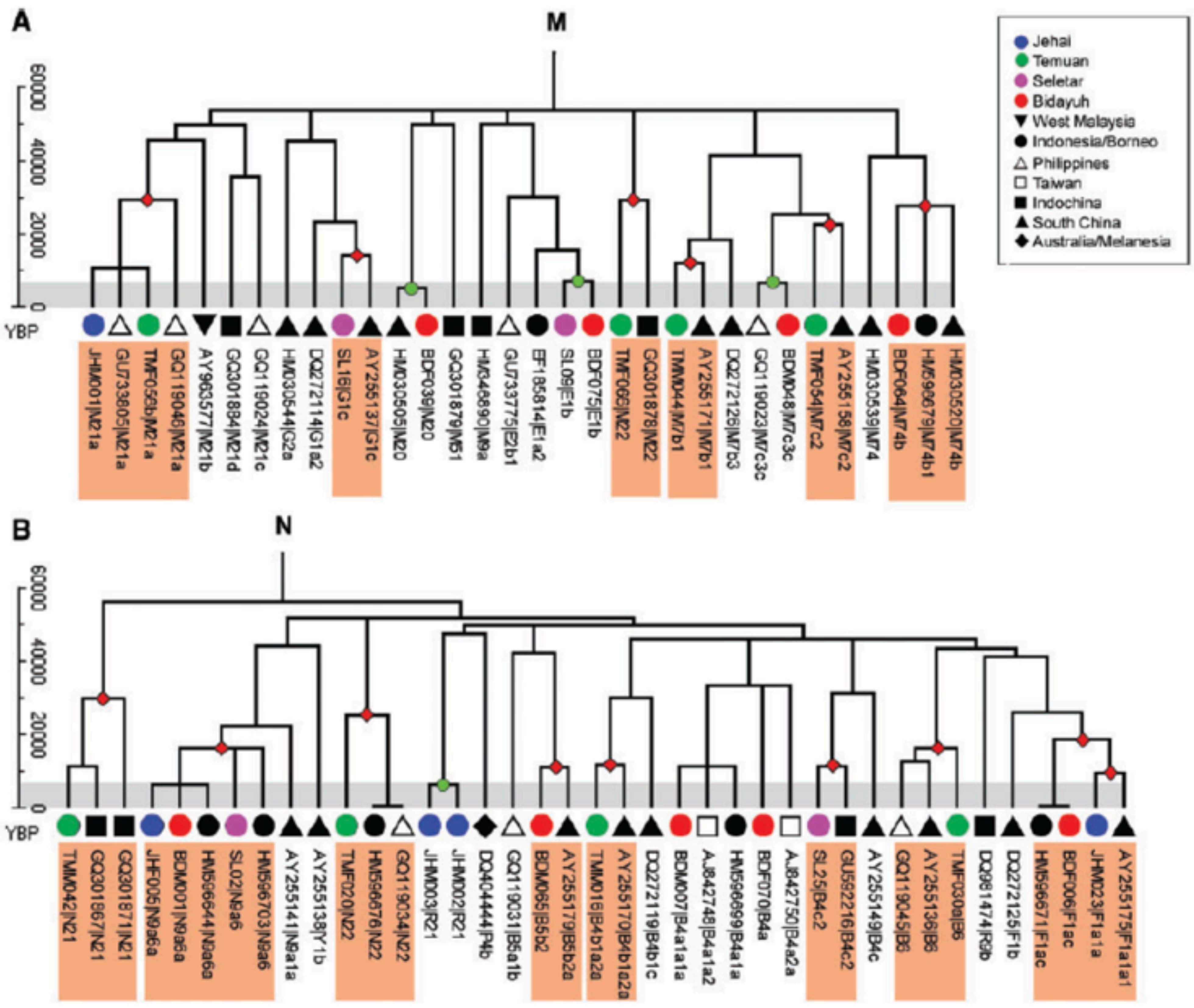

Twithin
Tbetween
1
2
3
4
5
6
7
8
α
β

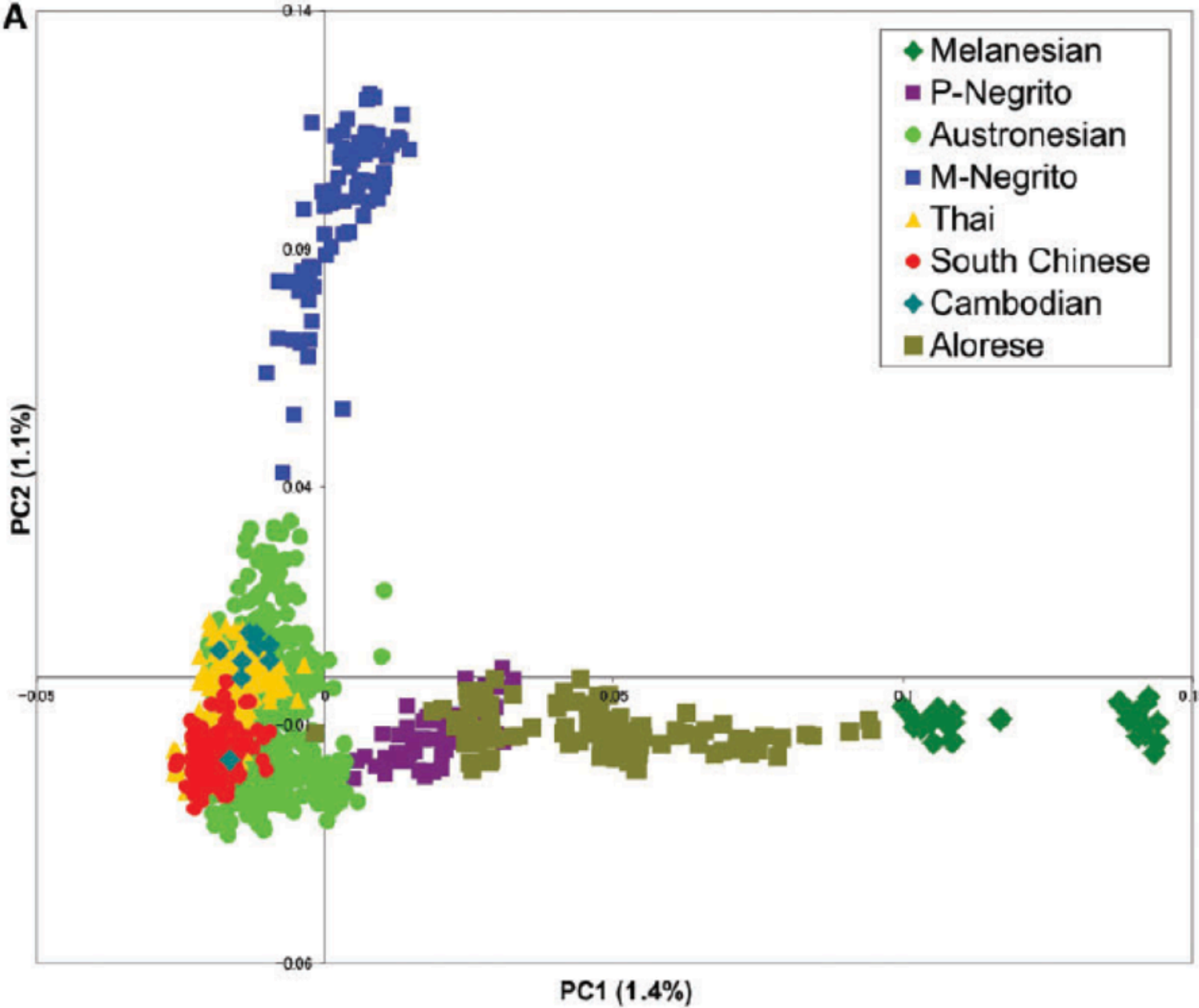

4N_c
T
A1
A2
A3
A4
B4
B3
B2
B1
α
β

ure 17-4: Dynamics of allele frequency changes during the population differenciation (from ref 17



Phylogenetic trees of mtDNA sequences of Malaysians (from Jinam et al. 2012)




A
Melanesian
P-Negrito
Austronesian
M-Negrito
Thai
South Chinese
Cambodian
Alorese
0.14
0.09
0.04
-0.01
-0.06
-0.05
0
0.05
0.1
0.15
PC2 (1.1%)
PC1 (1.4%)

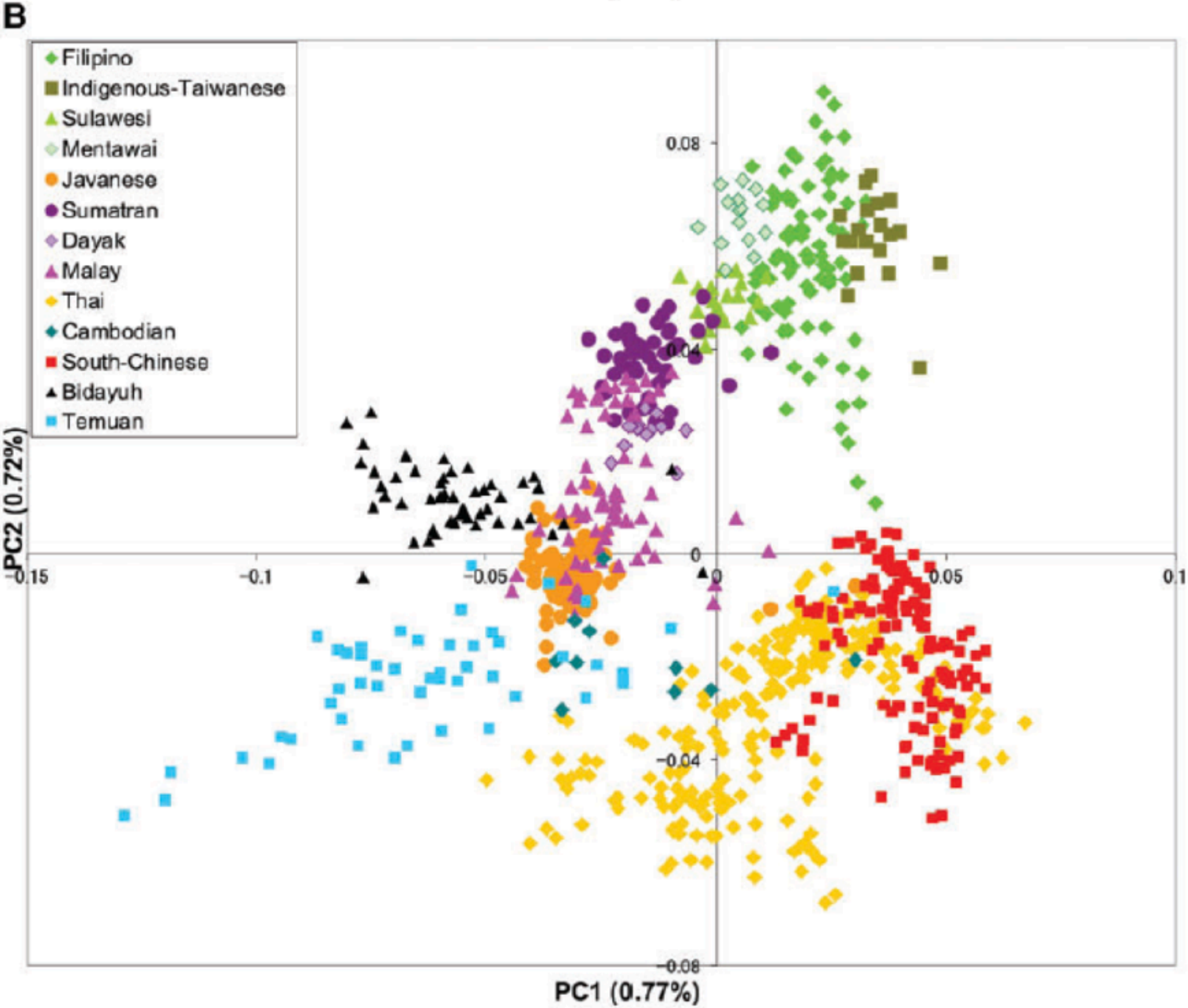

B
Filipino
Indigenous-Taiwanese
Sulawesi
Mentawai
Javanese
Sumatran
Dayak
Malay
Thai
Cambodian
South-Chinese
Bidayuh
Temuan
PC2 (0.72%)
PC1 (0.77%)
0.08
0.04
0
-0.04
-0.08
-0.15
-0.1
-0.05
0
0.05
0.1

Figure 17-7: distributions of sequence differences between all possible pairs of individuals for non-African populations (from ref 17-19)

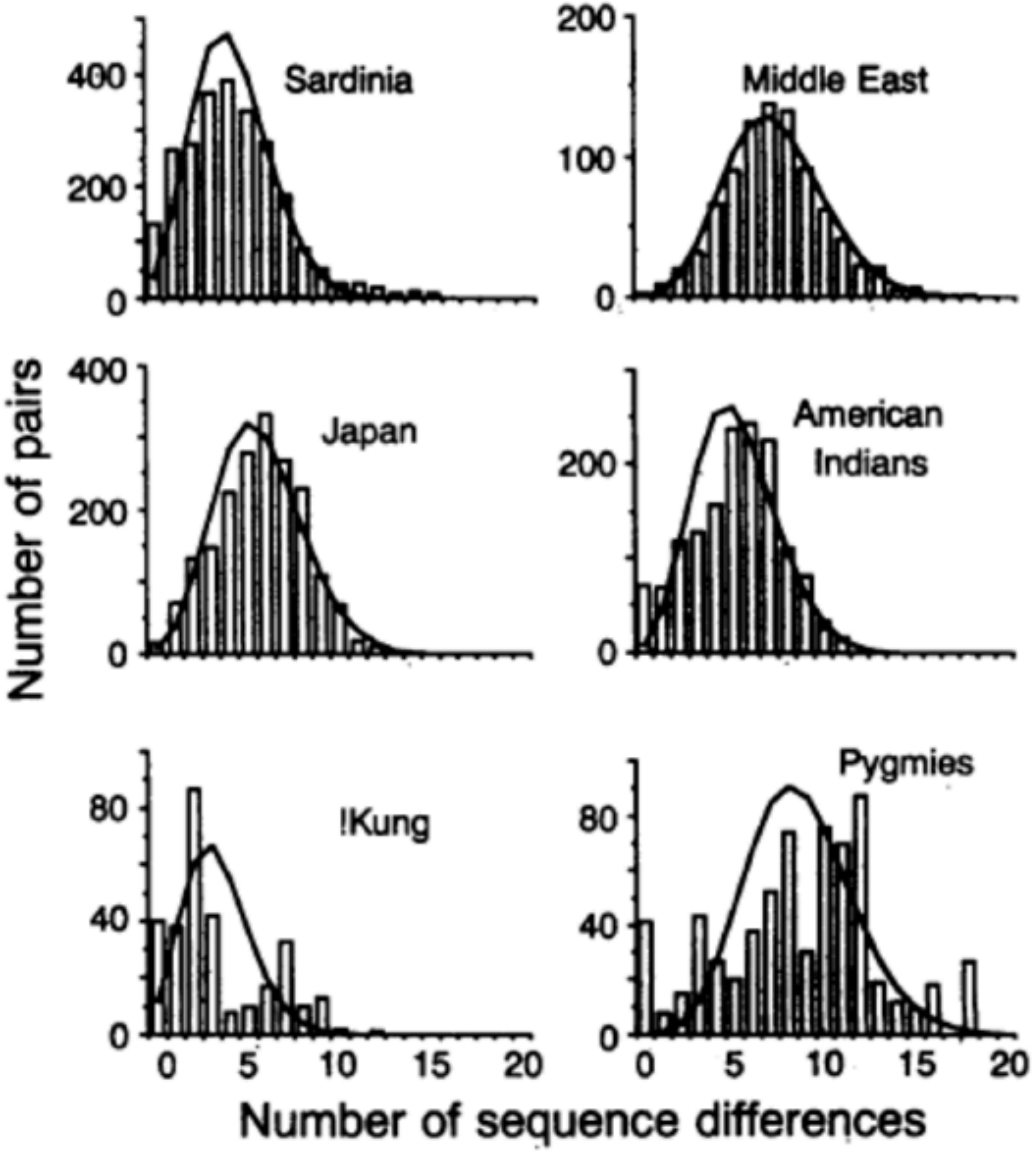


Figure 17-8: An example of the Bayesian Skyline Plot (from ref 17-16)