Manuscript genetics and Perl

or, How I made sense of 101011 manuscripts

Tara L Andrews
Departmental Lecturer
Faculty of History, University of Oxford
tla@mit.edu

Constantinople
Ani
Edessa
© 2009 Tele Atlas
US Dept of State Geographer
© 2009 Europa Technologies
© 2009 Cnes/Spot Image
Google
40°46'40.37" N 34°09'14.32" E elev 530 m
Eye alt 2474.73 km

- How do we reconstruct the past?

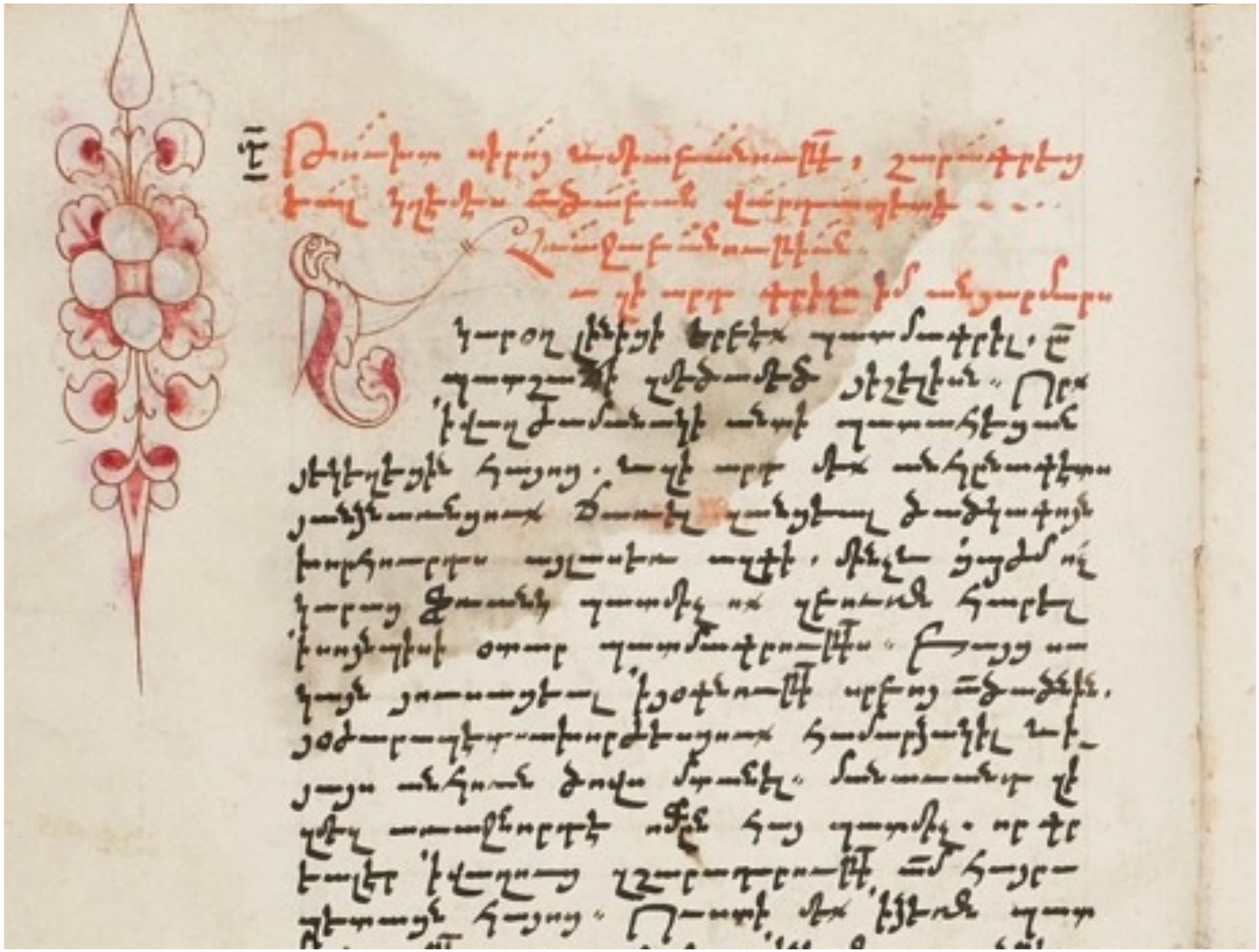
- How do we reconstruct the past?
- Reading what they wrote is a good start.

- How do we reconstruct the past?
- Reading what they wrote is a good start.
- But their originals don’t tend to survive.

- How do we reconstruct the past?
- Reading what they wrote is a good start.
- But their originals don't tend to survive.
- So we have to read what someone else read of what someone else read of what someone else read of what they wrote.

- How do we reconstruct the past?
- Reading what they wrote is a good start.
- But their originals don't tend to survive.
- So we have to read what someone else read of what someone else read of what someone else read of what they wrote.
- So we have to compare the versions.

SCRIPTORIUM MONK AT WORK. (From *Lacroix.*)



This is not a talk about Perl.

This is about a cool use of Perl.

Collation

"The collation of manuscripts requires the infuriating accuracy of a pedant and the obsessive stamina of an idiot. It is therefore an ideal task for a computer."

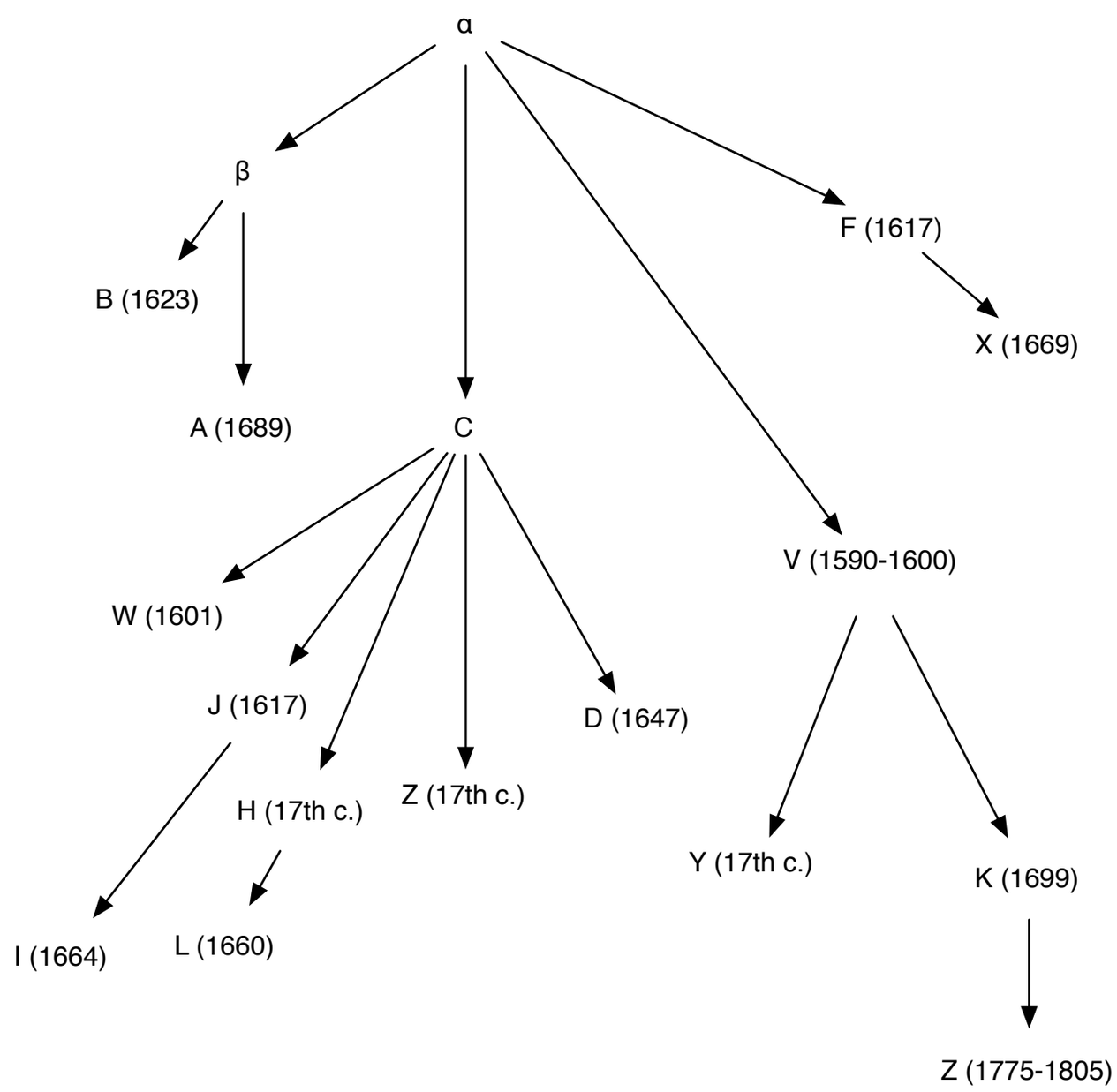
—Peter Robinson, "Collation and Textual Criticism", *LLC* vol. 4 no. 2, 1989

```
Looking at app App452
Context: և մոլորեալ , թողուն զհոգս շինութեան տանց և զյաջողուածս գործոց, և մեջտ և հանապազ ի գինարբուս _____
Following context:
#D                          : դեգերին, վասն սիրոյ և ցանկութեան չար
#W                          : դեգերէին վասն սիրոյ և ցանկութեան չար
#X                          : դէգերեն: վասն սիրոյ և ցանկութեան չար
#Y                          : դեգերին, վասն սիրոց և ցանկութեանց չար՝
#F                          : դէգերին, վասն սիրոց և ցանկութեան չար
#I #Jer #J #K #Z #V #A #B #O: դեգերին, վասն սիրոյ և ցանկութեան չար
-----------

Available readings:
1 : դեգերին (#Jer #K #B #I #A #D #J #O #V #Y #Z )
2 : դէգերեն (#X )
4 : դէգերին (#F )
6 : դեգերէին (#W )

Manuscript details:
3: Witness(es) #X contain a : after դէգերեն
5: Witness(es) #F contain a , after դէգերին
 --> h
Available commands:
accept (#)
accept detail (#)
subst detail (#) (value)
spelling (# alternate) (# canonical)
orth (# alternate) (# canonical)
emend (comment text)
note (note text)
next
help
save
quit
 --> 
```

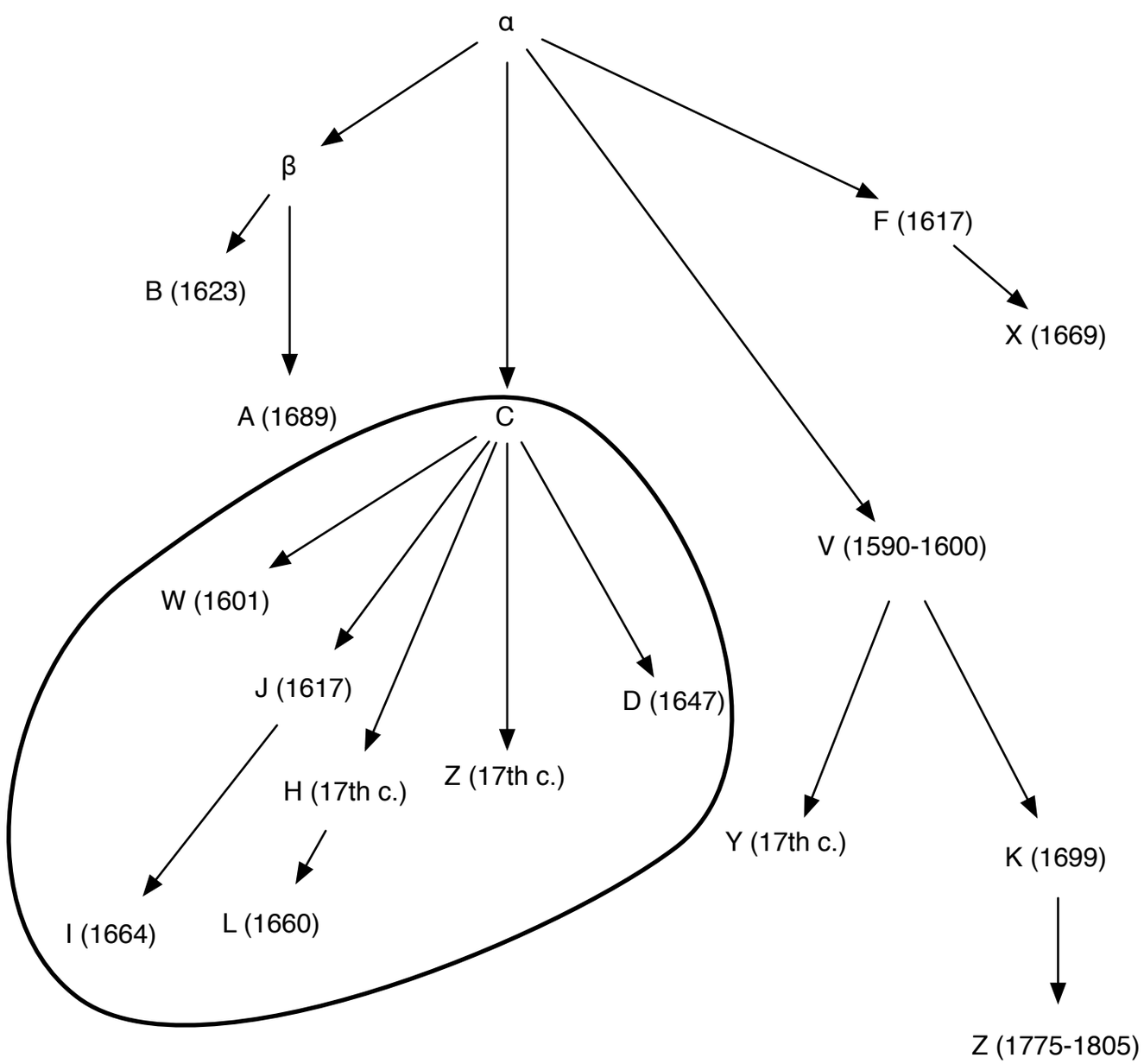
Manuscript genetics

- Not invented here
- But pretty cool anyway
- A way to reconstruct the past, a guide to good editing decisions

A helpful stemma



A helpful stemma



How to make a stemma

How to make a stemma

- Know the answer in advance

How to make a stemma

- Know the answer in advance
- Assume a ‘right’ text that you can compare the existing ‘wrong’ texts to

How to make a stemma

- Know the answer in advance
- Assume a ‘right’ text that you can compare the existing ‘wrong’ texts to
- Assume that the wrong is the child of the right

How to make a stemma

- Know the answer in advance
- Assume a ‘right’ text that you can compare the existing ‘wrong’ texts to
- Assume that the wrong is the child of the right
- Doesn’t really work for texts like mine.

To get the answer, I have to have the answer.

Wasn't this supposed to be labor-saving?

How to make a stemma?

- I have way too many manuscripts to go through them by hand
- I don’t know in advance what the ‘right’ text is
- So I can’t say what is ‘wrong’
- And I don’t believe that all change is error

Pretend the manuscripts are organisms.

յայսմ	այս	յայսմ	այս	այս
ամենայն	ամենայն	ամի	ամենայն	ամենայն
եղելոցն,	եղելոց	եղելոց	եղելոց	եղելոցս
նստուցանեն	նստուցանեն	նստուցանեն	նստուցանեն	նստուցանեն
զաթոռ	զաթոռ	յաթոռ	զաթոռ	զաթոռ
հայրապետութեան	հայրապետութեան	հայրապետութեան	հայրապետութեաննն	հայրապետութեան
ի		ի	ի	
թաւբլուր		թաւաբլուրն:	թաւբլուր	
և		և	և	
կացեալ		կացեալ	կացեալ	
անդ		անդ	անդ	
զամս		զամս	զամս	
գ		գ,	գ	
և		և	և	
ընդ		ընդ	ընդ	
ամենայն		ամենայն	ամենայն	
զ		զ	վեց	
ամ		ամ,	ամ	
կալեալ		կալեալ	կալեալ	
զաթոռ		զաթոռ	զաթոռ	
հայրապետութեաննն		հայրապետութեան	հայրապետութեաննն	
տէր	տէր	տէր	տէր	զտէր
խաչիկ:	խաչիկ:	խաչիկն:	խաչիկ:	խաչիկ:

A	B	A	B	B
A	A	B	A	A
A	B	B	B	C
A	A	A	A	A
A	A	B	A	A
A	A	A	B	A
A	O	A	A	O
A	O	B	A	O
A	O	A	A	O
A	O	A	A	O
A	O	A	A	O
A	O	A	A	O
A	O	A	A	O
A	O	A	A	O
A	O	A	A	O
A	O	A	A	O
A	O	A	B	O
A	O	A	A	O
A	O	A	A	O
A	O	A	A	O
A	O	B	A	O
A	A	A	A	B
A	A	B	A	A

Choosing your data

Choosing your data

- Not all variations are equal
  - Some variations don't come from copying
  - Some variations will get 'corrected by a later scribe
  - Some variations are singularities, so not very useful for judging relationships

Choosing your data

- Not all variations are equal
  - Some variations don't come from copying
  - Some variations will get 'corrected by a later scribe
  - Some variations are singularities, so not very useful for judging relationships
- But how choosy do you need to be?

A little experiment

A little experiment

- Use all the data, corrected for spelling variation.

A little experiment

- Use all the data, corrected for spelling variation.
- Get rid of a common scribal variation (-น at the end of a word.

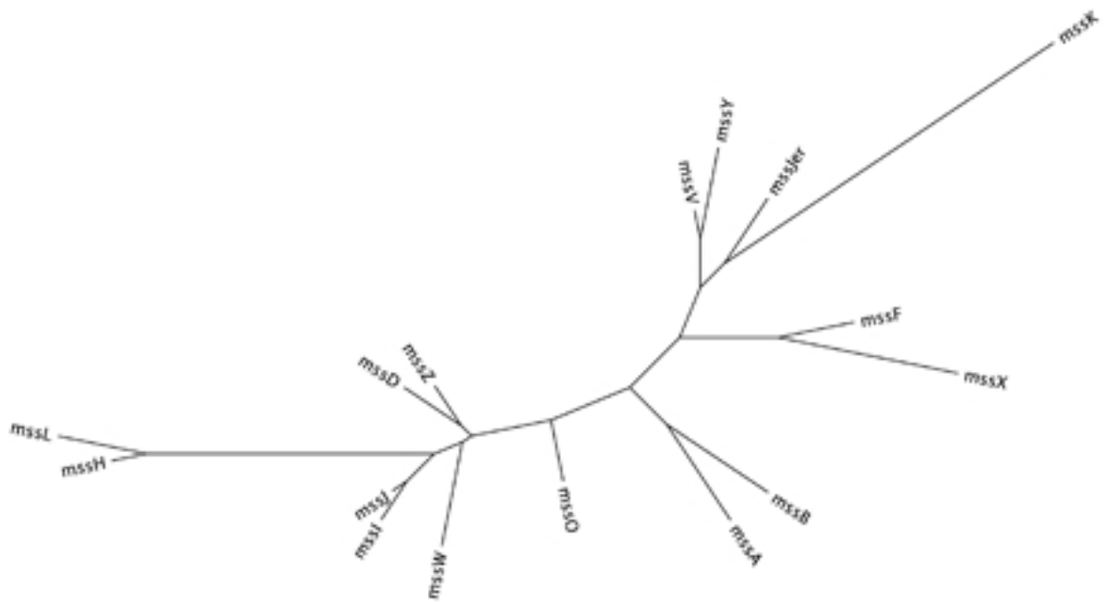
A little experiment

- Use all the data, corrected for spelling variation.
- Get rid of a common scribal variation (-น at the end of a word.
- Only use data where there are at least two witnesses for each variant.

A little experiment

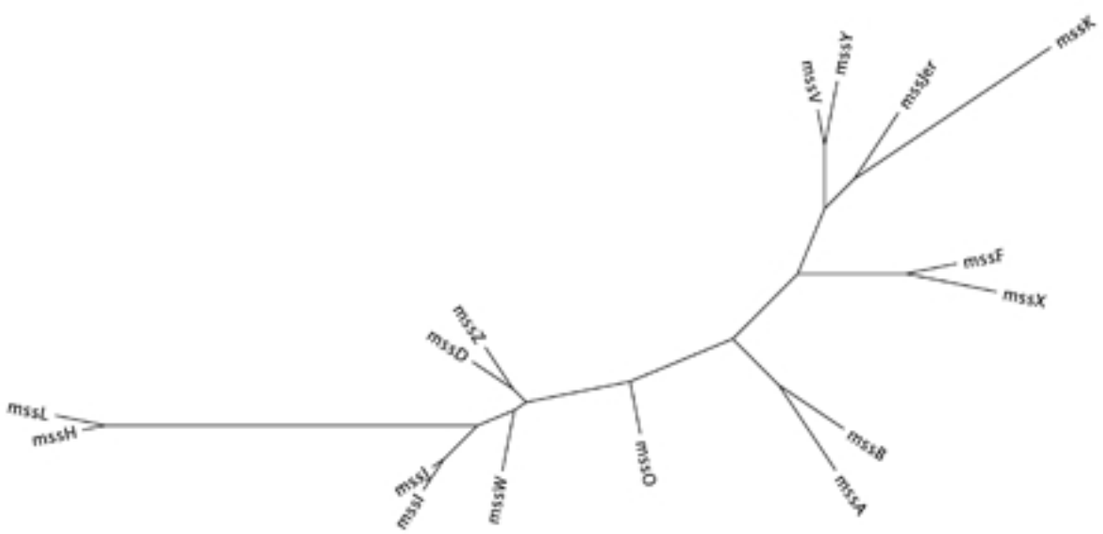
- Use all the data, corrected for spelling variation.
- Get rid of a common scribal variation (-น at the end of a word.
- Only use data where there are at least two witnesses for each variant.
- See if there is any difference.

1. All readings

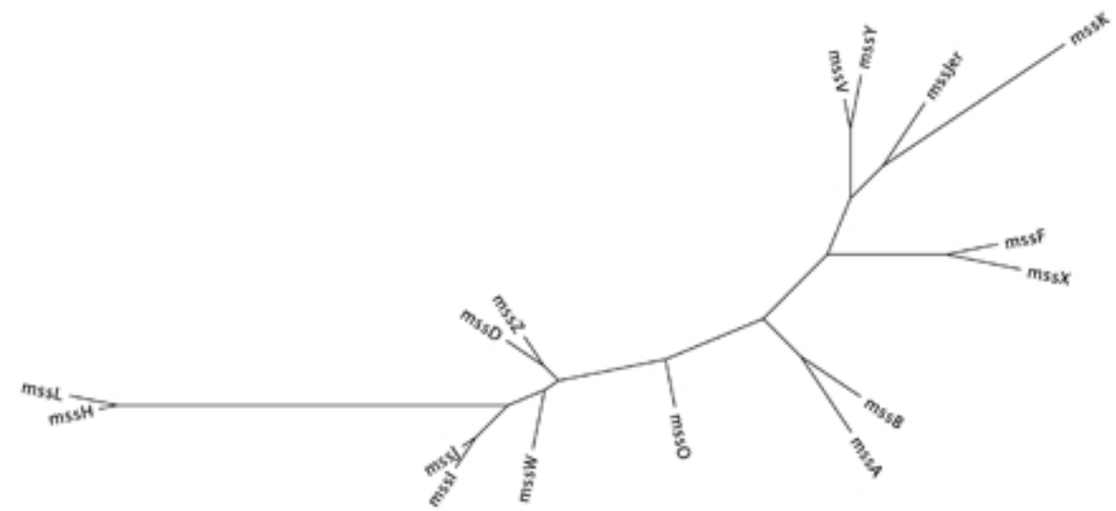


3. Selected readings

2. All readings, strict pairs only



4. Selected readings, strict pairs only



Sampling techniques

Sampling techniques

- Three samples, one from each book of the Chronicle

Sampling techniques

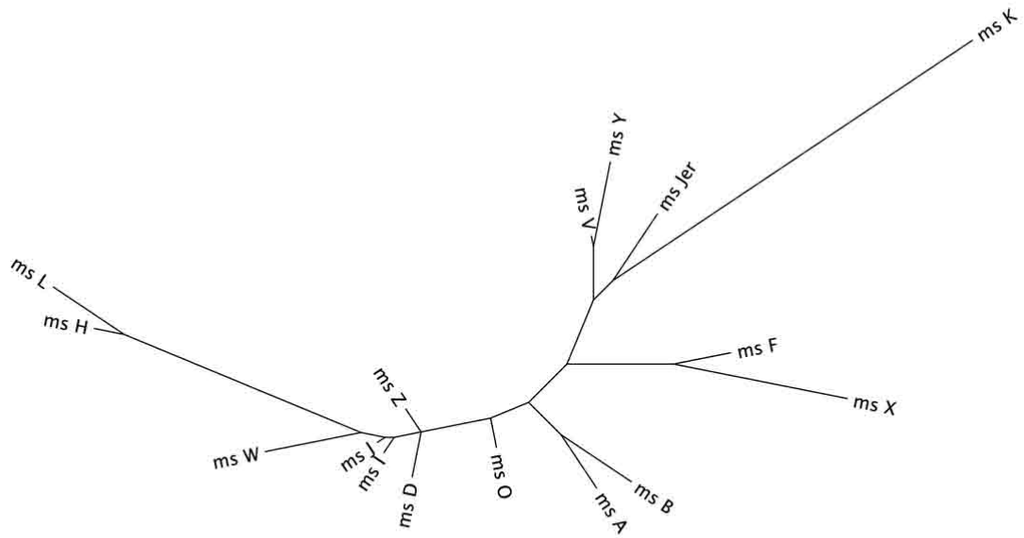
- Three samples, one from each book of the Chronicle
- 5,000 words in total for an 80,000-word text

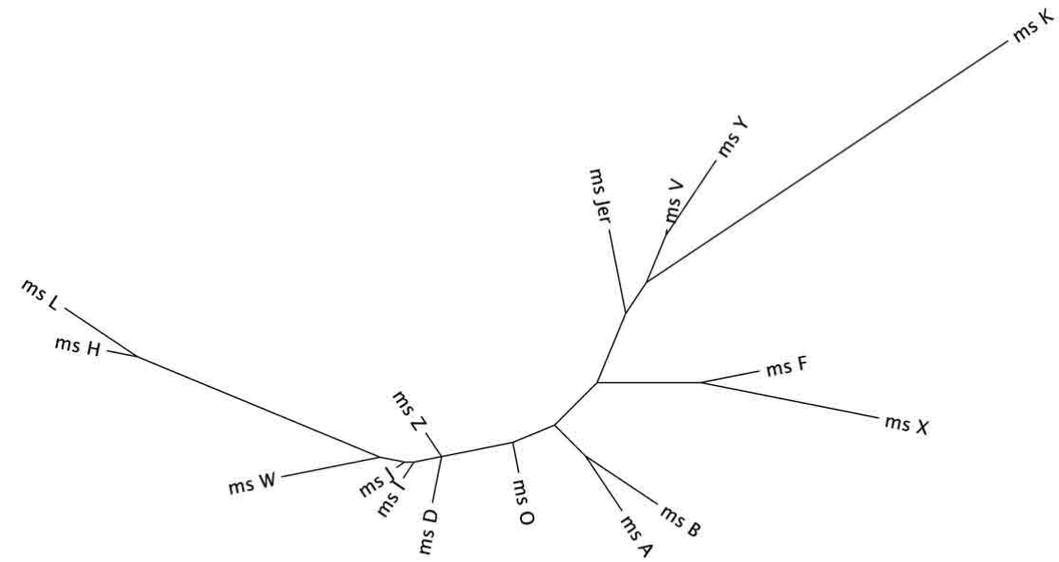
Sampling techniques

- Three samples, one from each book of the Chronicle
- 5,000 words in total for an 80,000-word text
- Needed to give a single answer

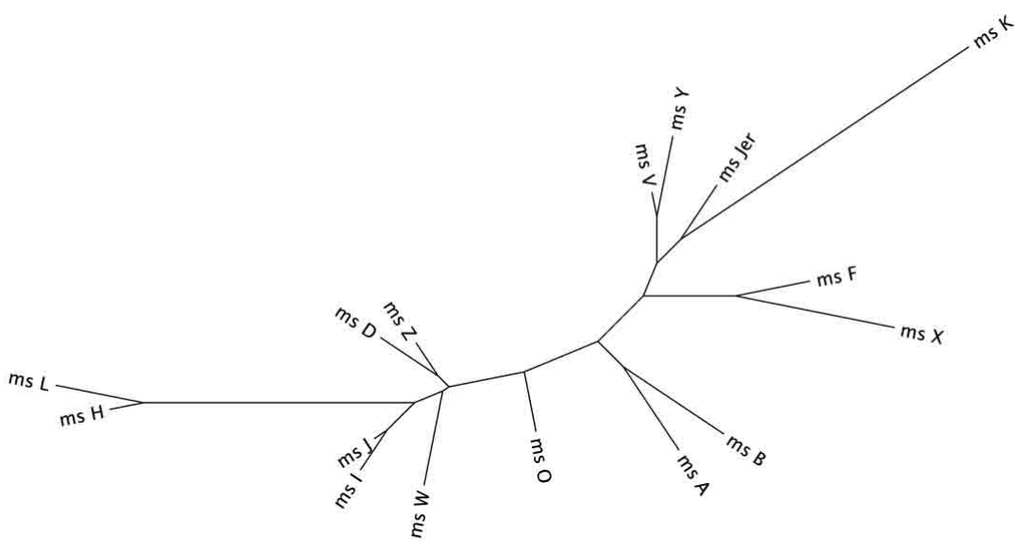
Sampling techniques

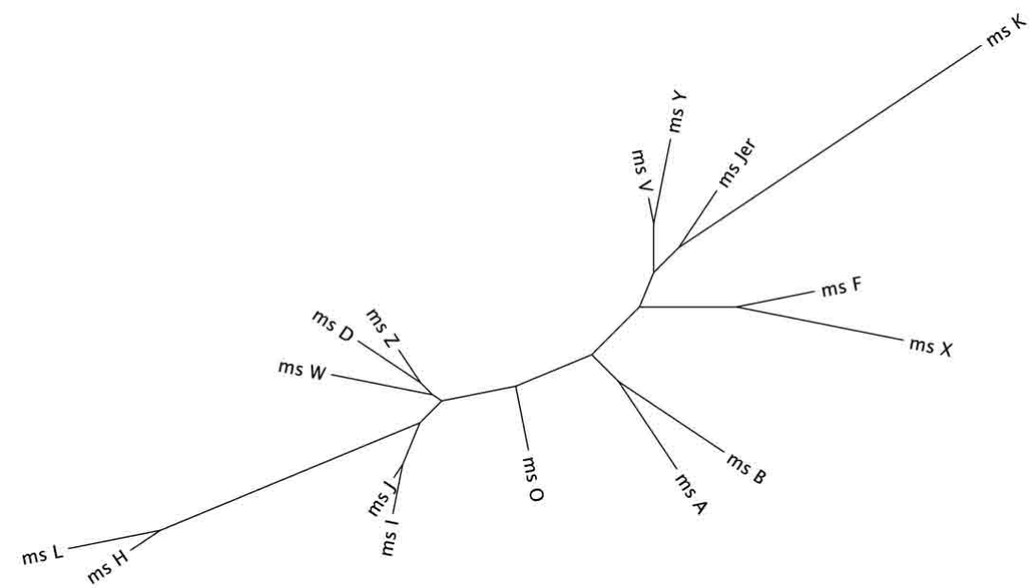
- Three samples, one from each book of the Chronicle
- 5,000 words in total for an 80,000-word text
- Needed to give a single answer
- Needed to give consistent answers across samples



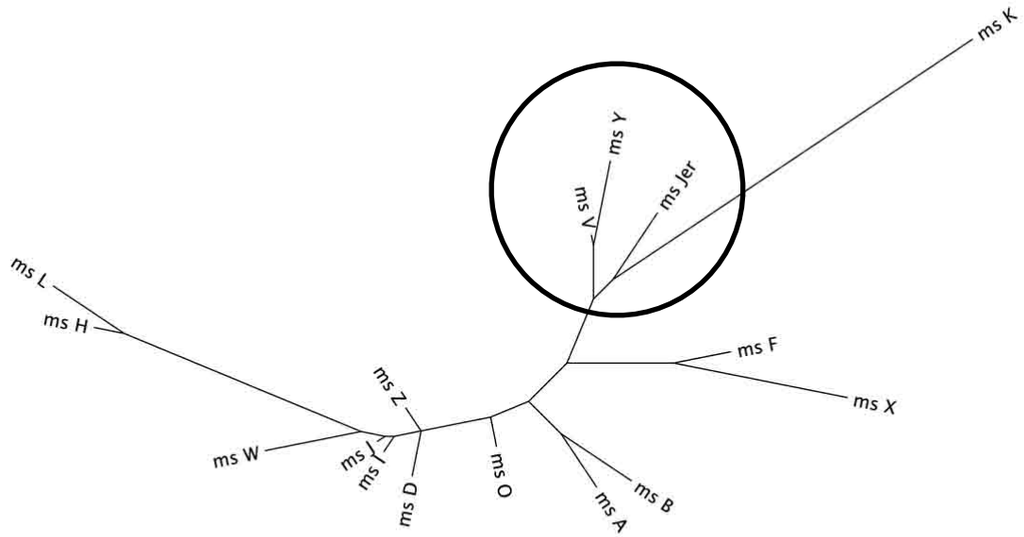


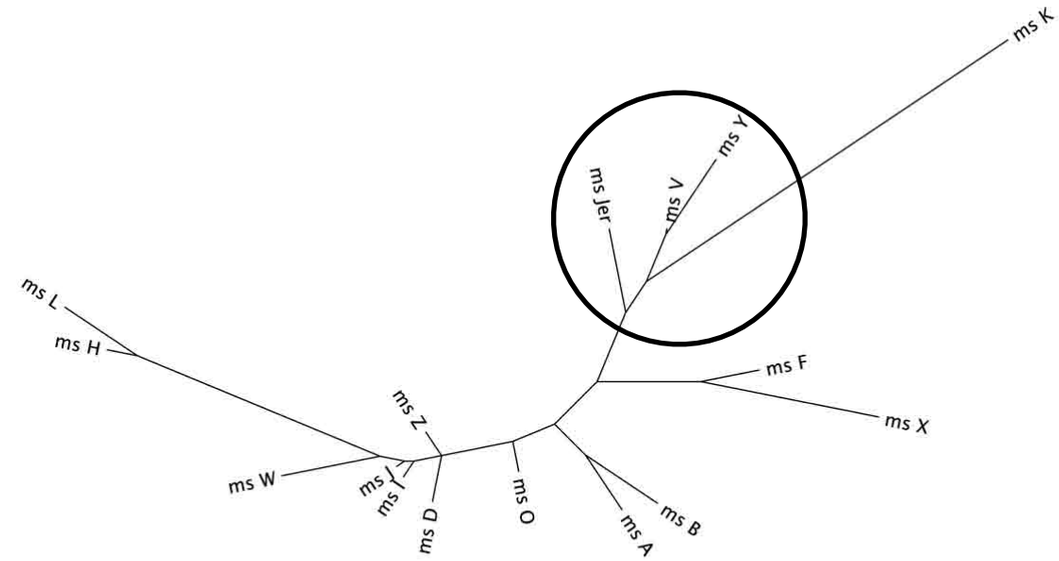
Possible trees from the text of Book 2



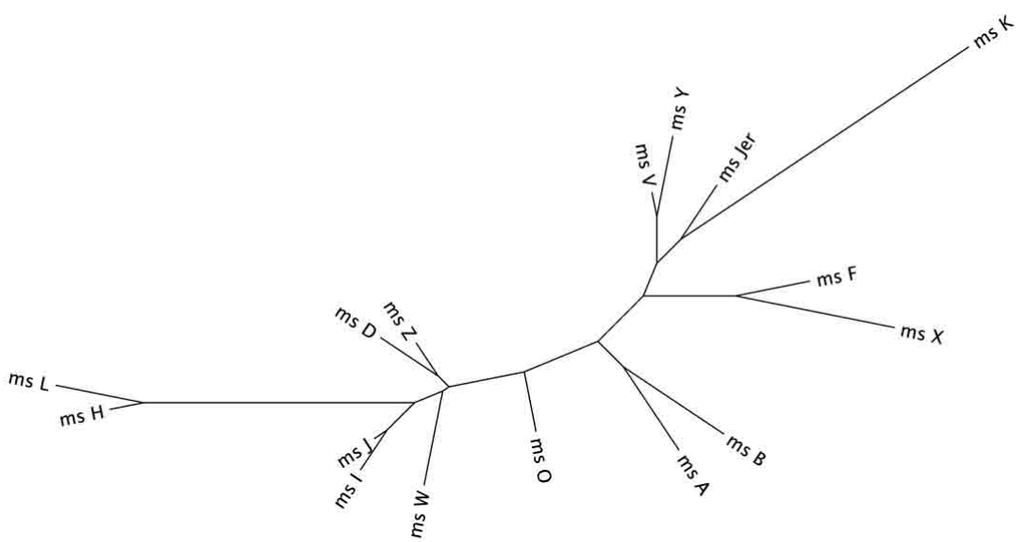


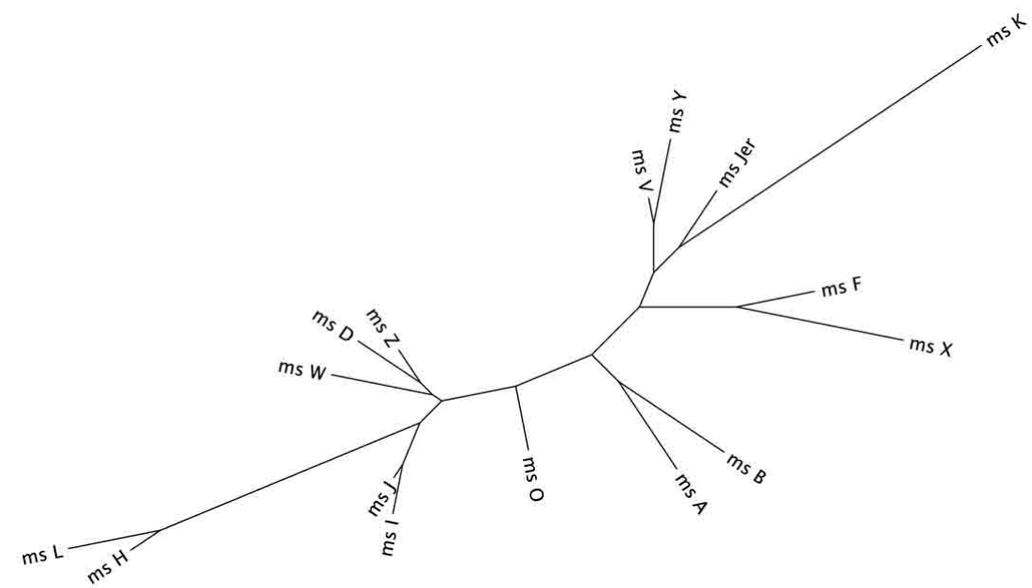
Possible trees from the text of the prophecy



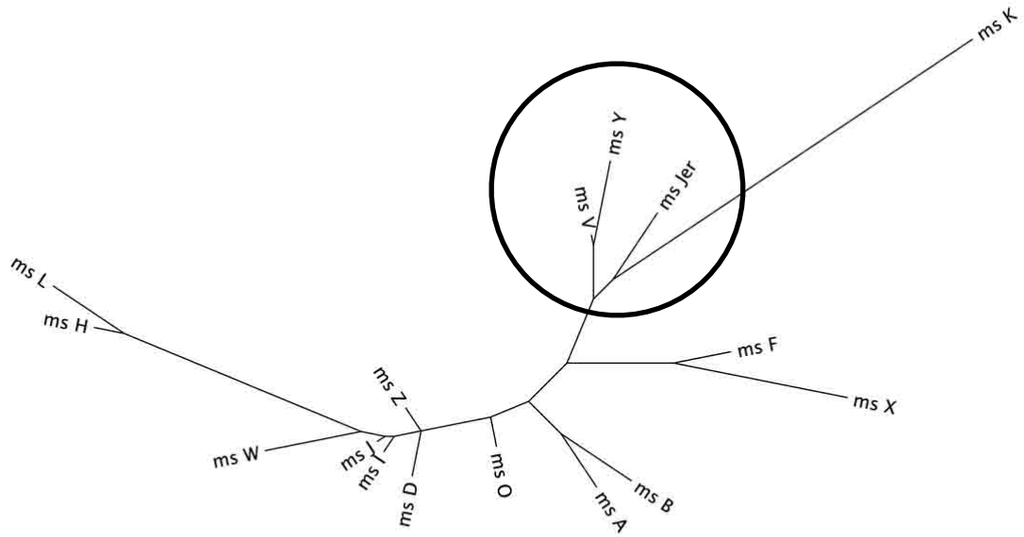


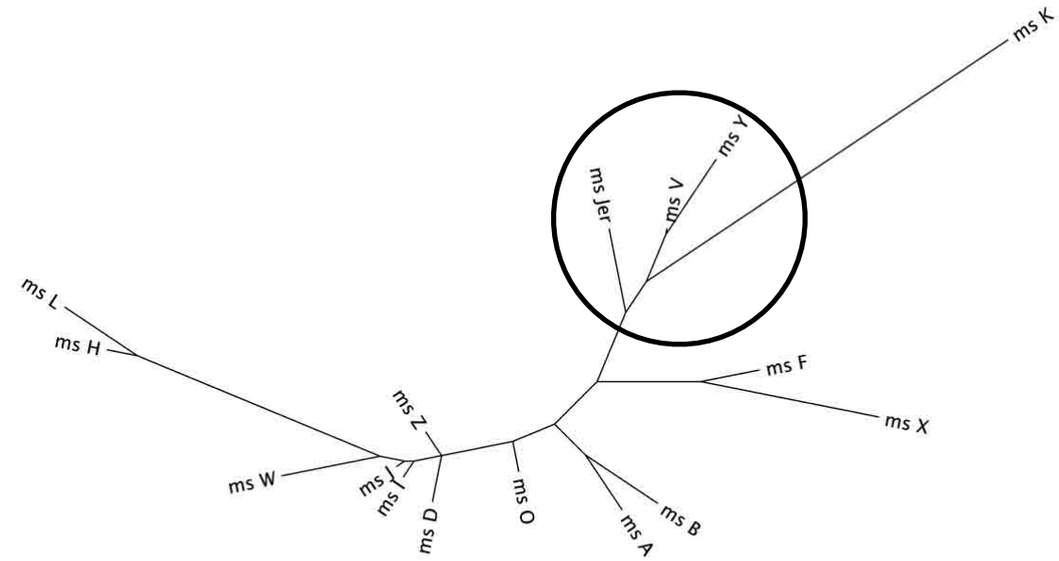
Possible trees from the text of Book 2



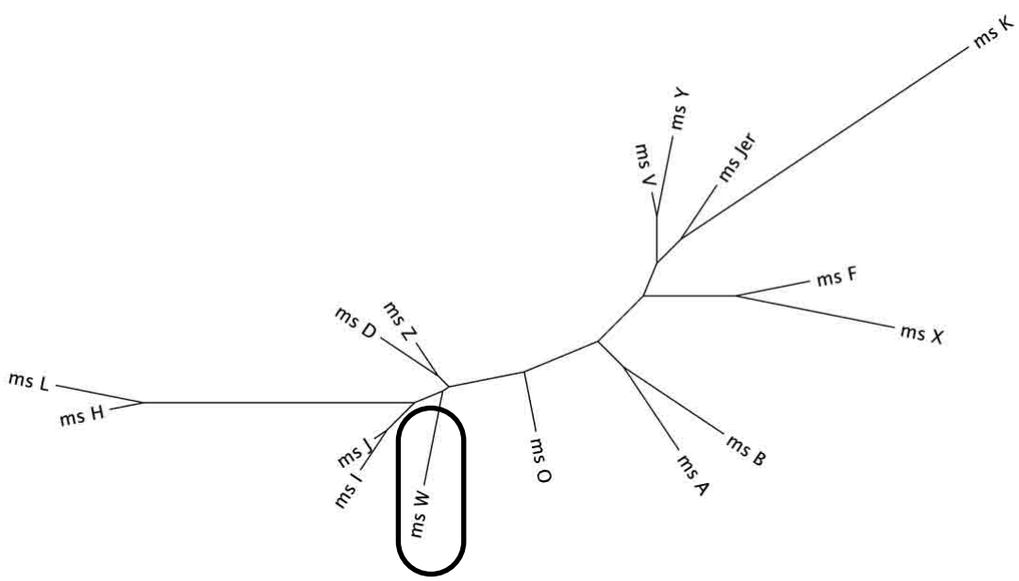


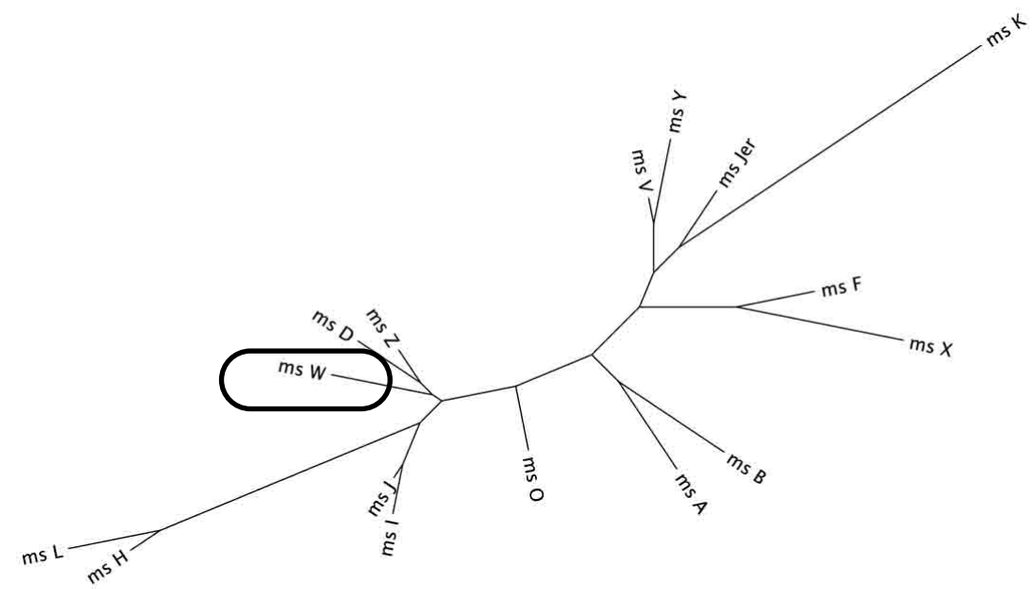
Possible trees from the text of the prophecy



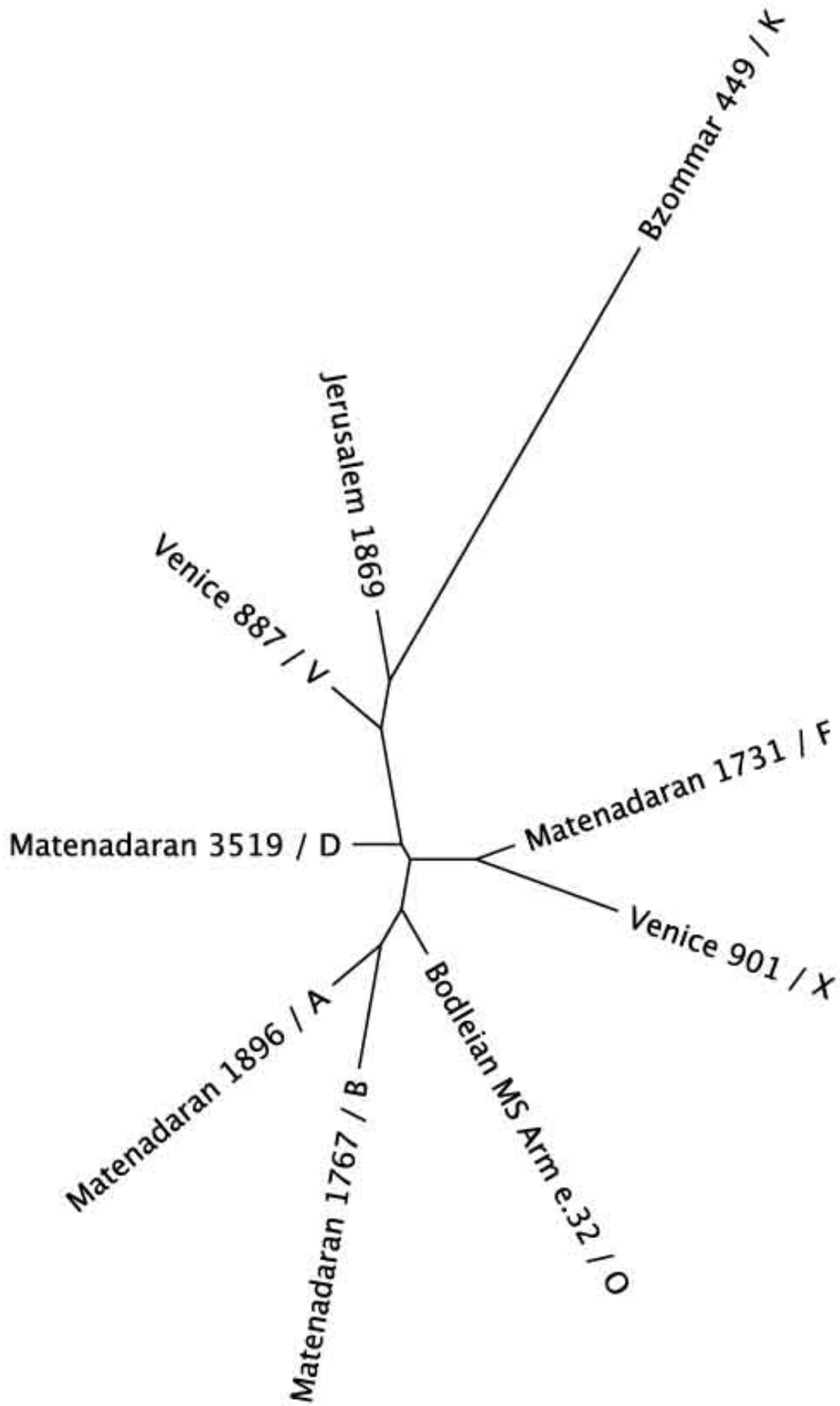


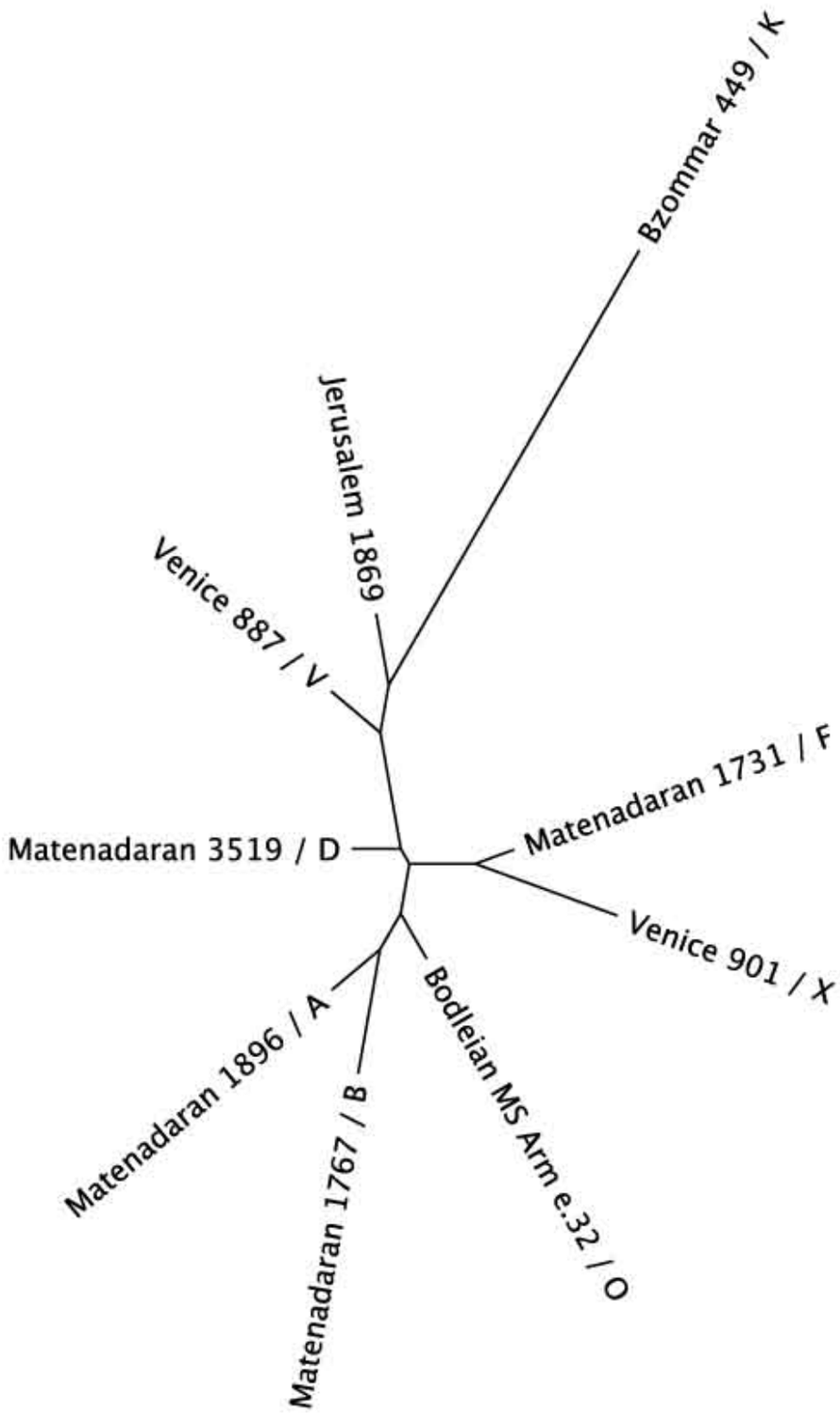
Possible trees from the text of Book 2

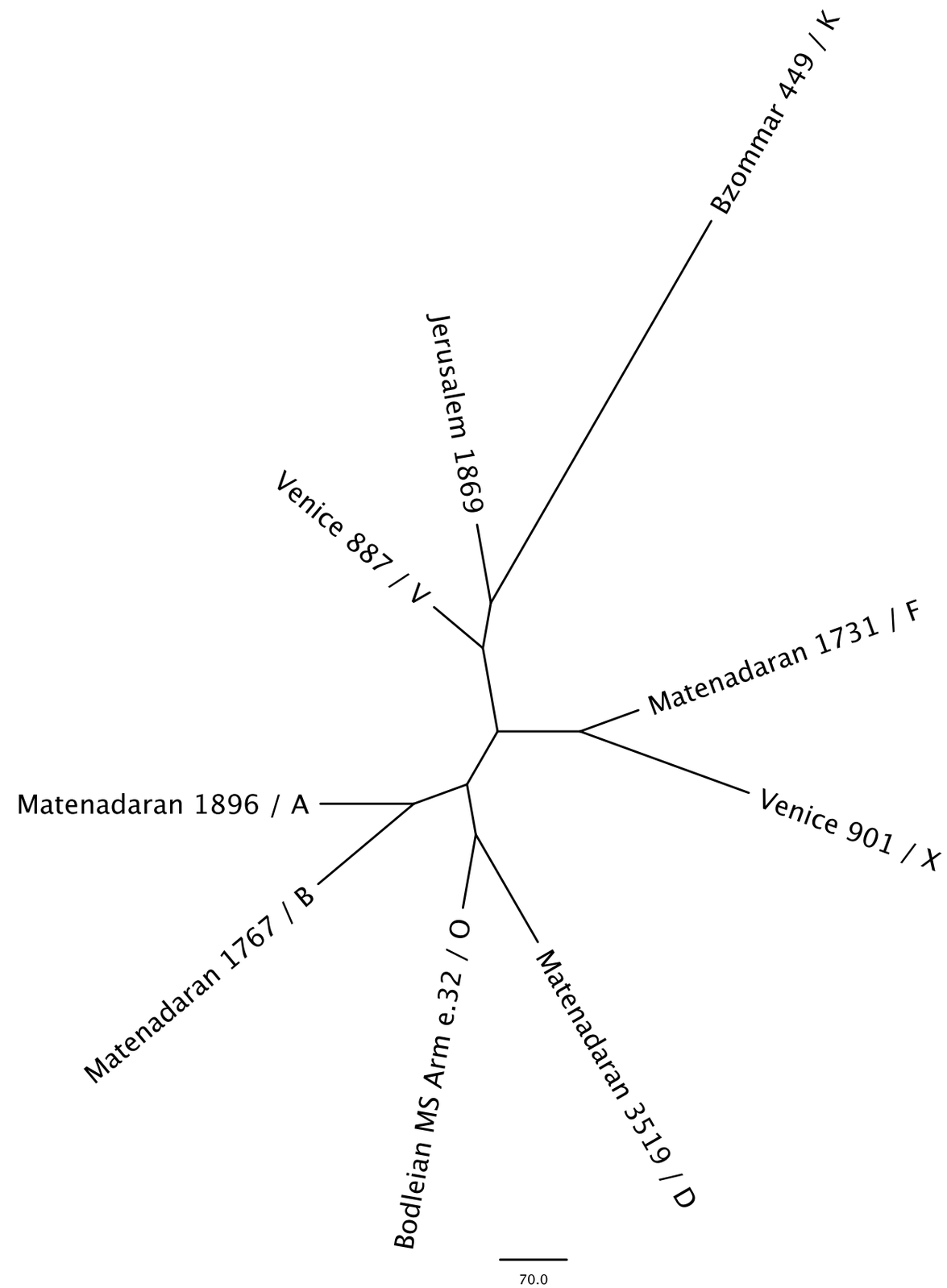


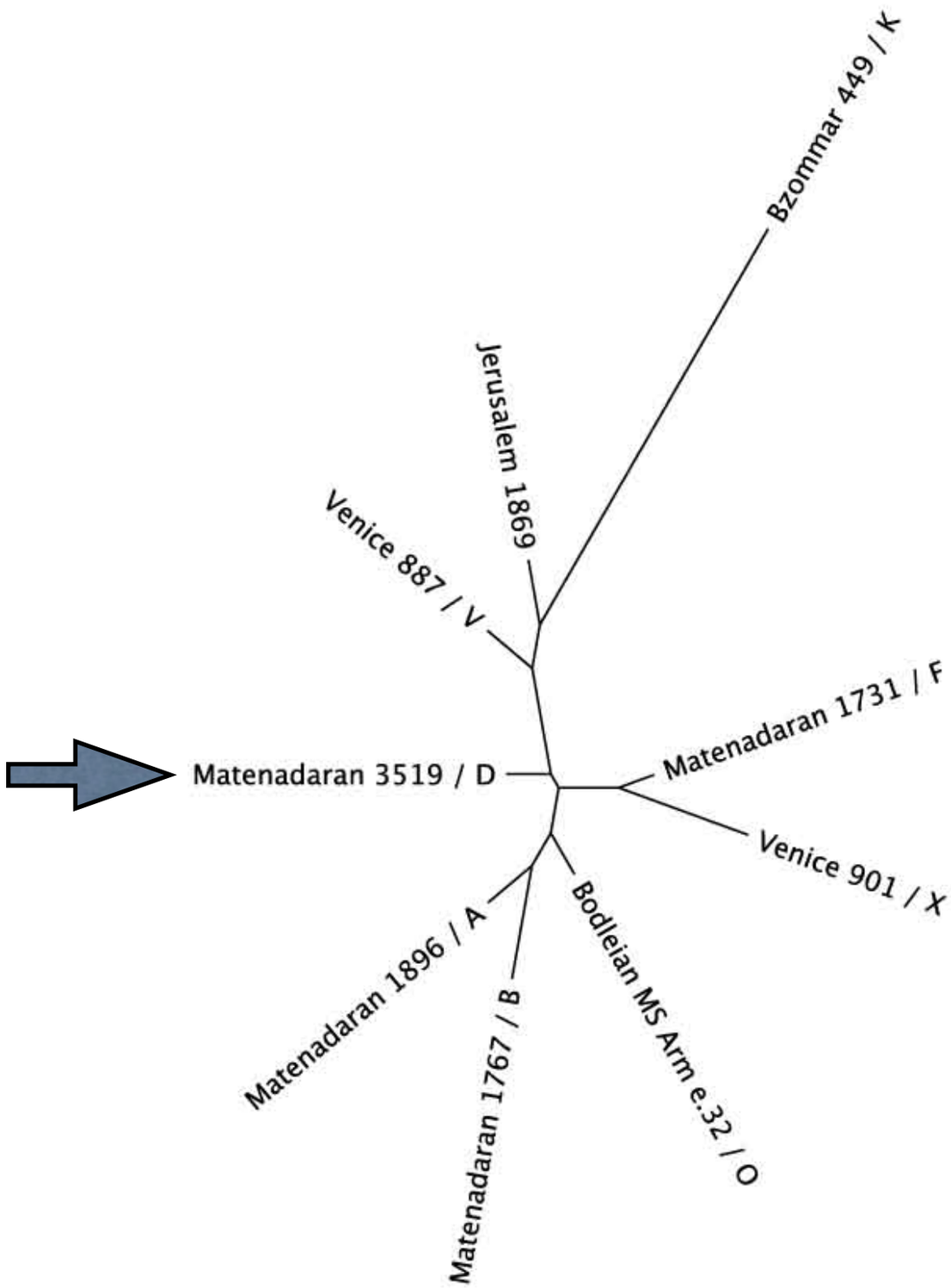


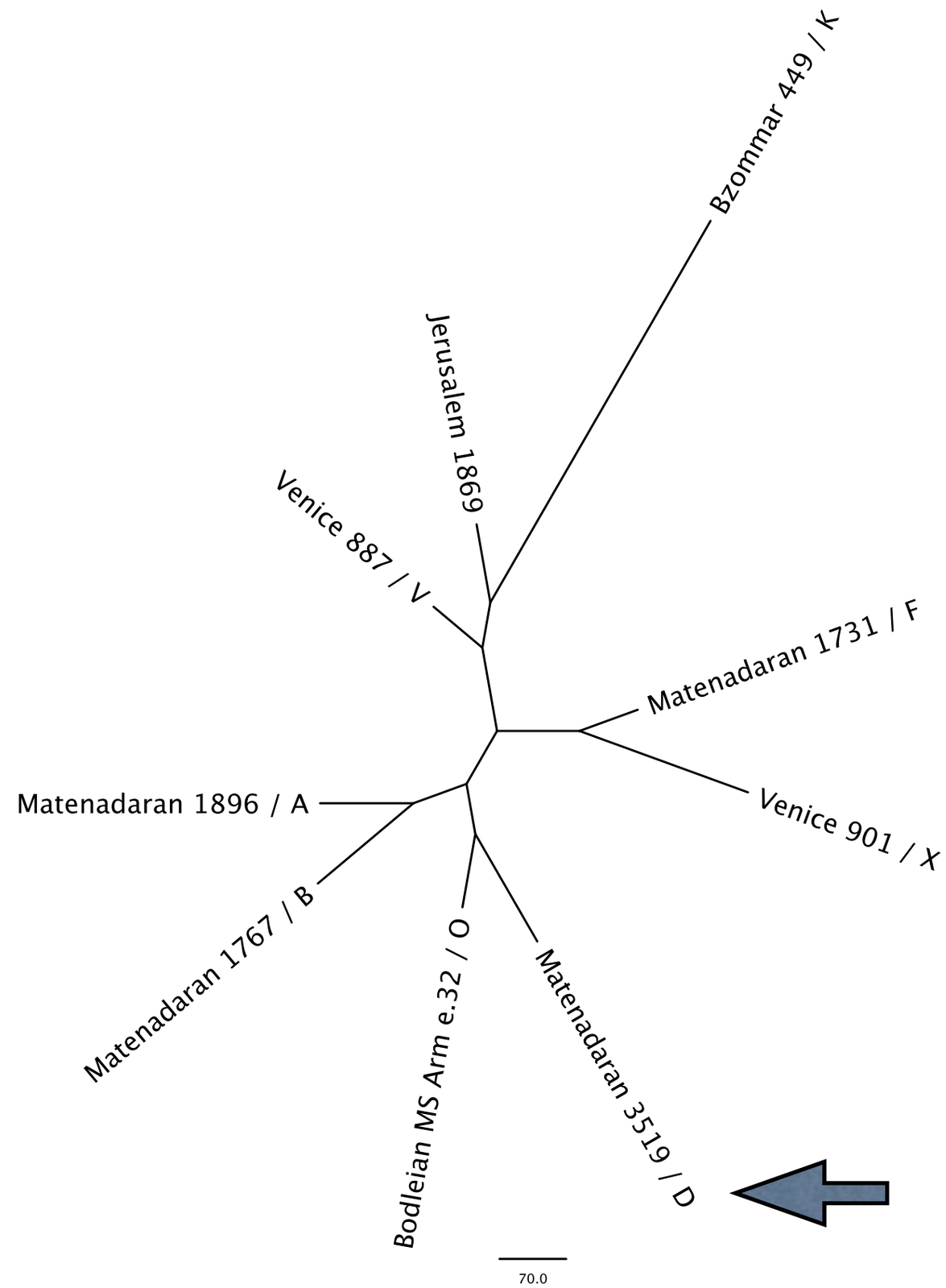
Possible trees from the text of the prophecy

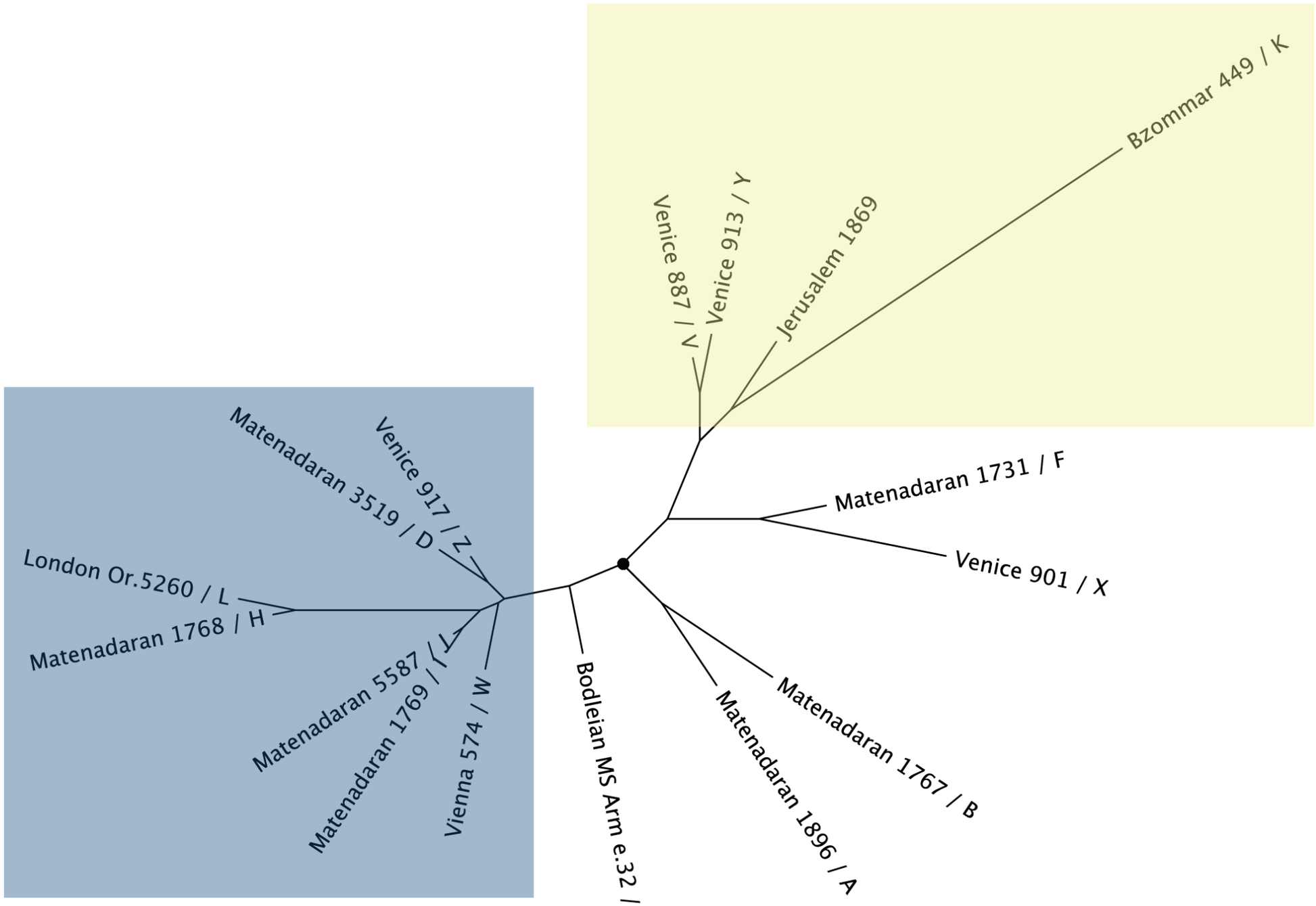

Bzommar 449 / K
Jerusalem 1869
Venice 887 / V
Matenadaran 1731 / F
Matenadaran 3519 / D
Venice 901 / X
Bodleian MS Arm e.32 / O
Matenadaran 1896 / A
Matenadaran 1767 / B

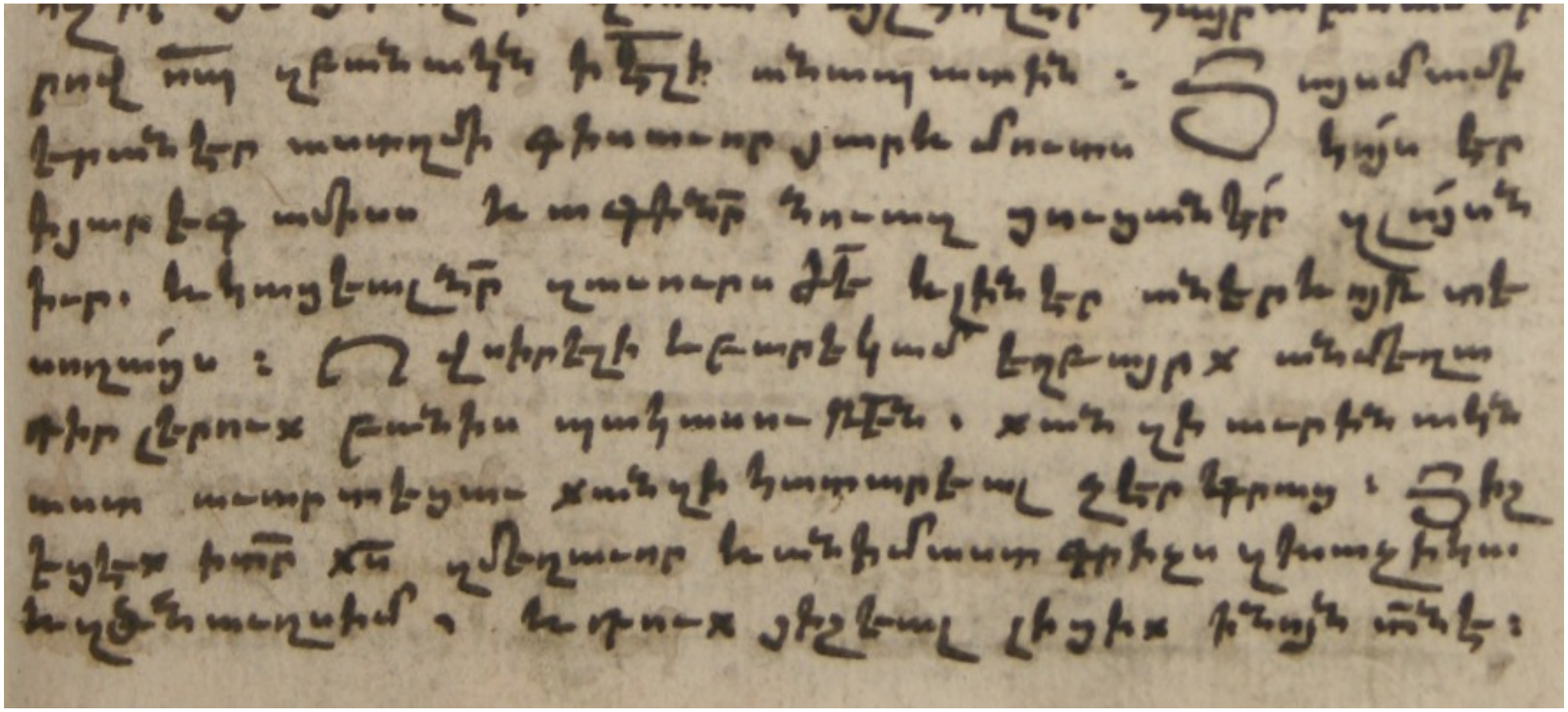

Bzommar 449 / K
Jerusalem 1869
Venice 887 / V
Matenadaran 1731 / F
Matenadaran 3519 / D
Venice 901 / X
Matenadaran 1896 / A
Matenadaran 1767 / B
Bodleian MS Arm e.32 / O

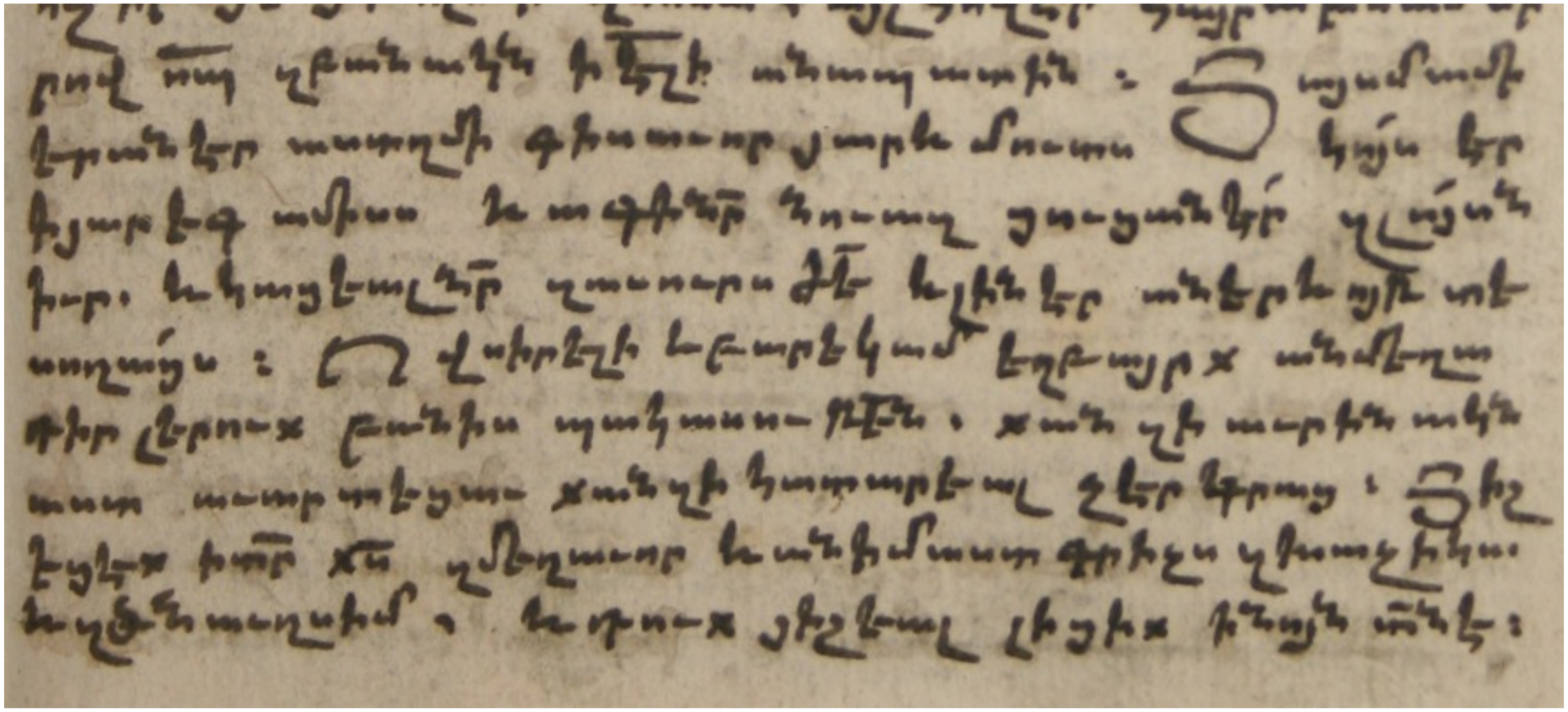

Bzommar 449 / K
Jerusalem 1869
Venice 887 / V
Matenadaran 1731 / F
Venice 901 / X
Matenadaran 1896 / A
Matenadaran 1767 / B
Bodleian MS Arm e.32 / O
Matenadaran 3519 / D
70.0


Bzommar 449 / K
Jerusalem 1869
Venice 887 / V
Matenadaran 3519 / D
Matenadaran 1731 / F
Venice 901 / X
Matenadaran 1896 / A
Matenadaran 1767 / B
Bodleian MS Arm e.32 / O

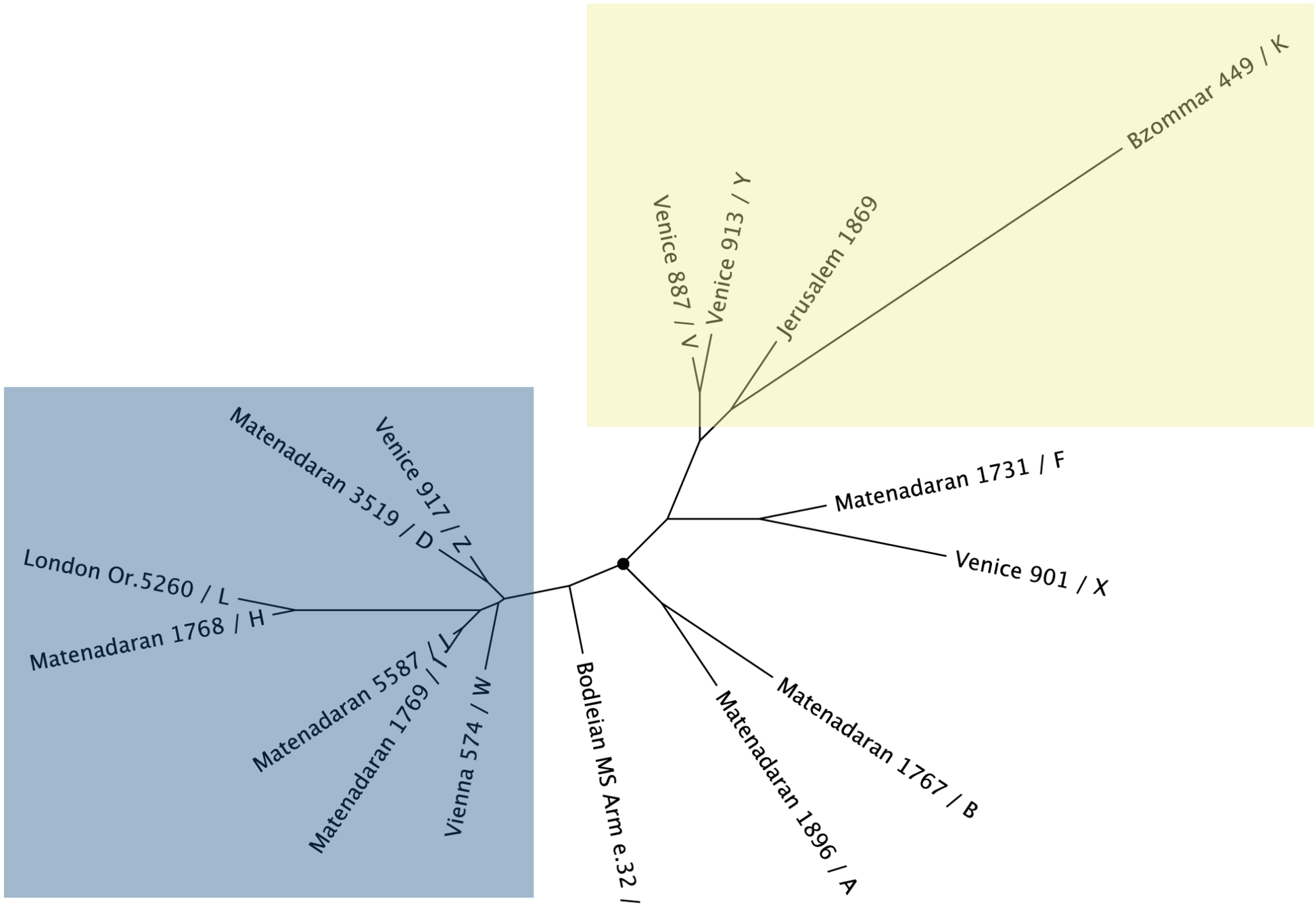

Bzommar 449 / K
Jerusalem 1869
Venice 887 / V
Matenadaran 1731 / F
Venice 901 / X
Matenadaran 1896 / A
Matenadaran 1767 / B
Bodleian MS Arm e.32 / O
Matenadaran 3519 / D
70.0

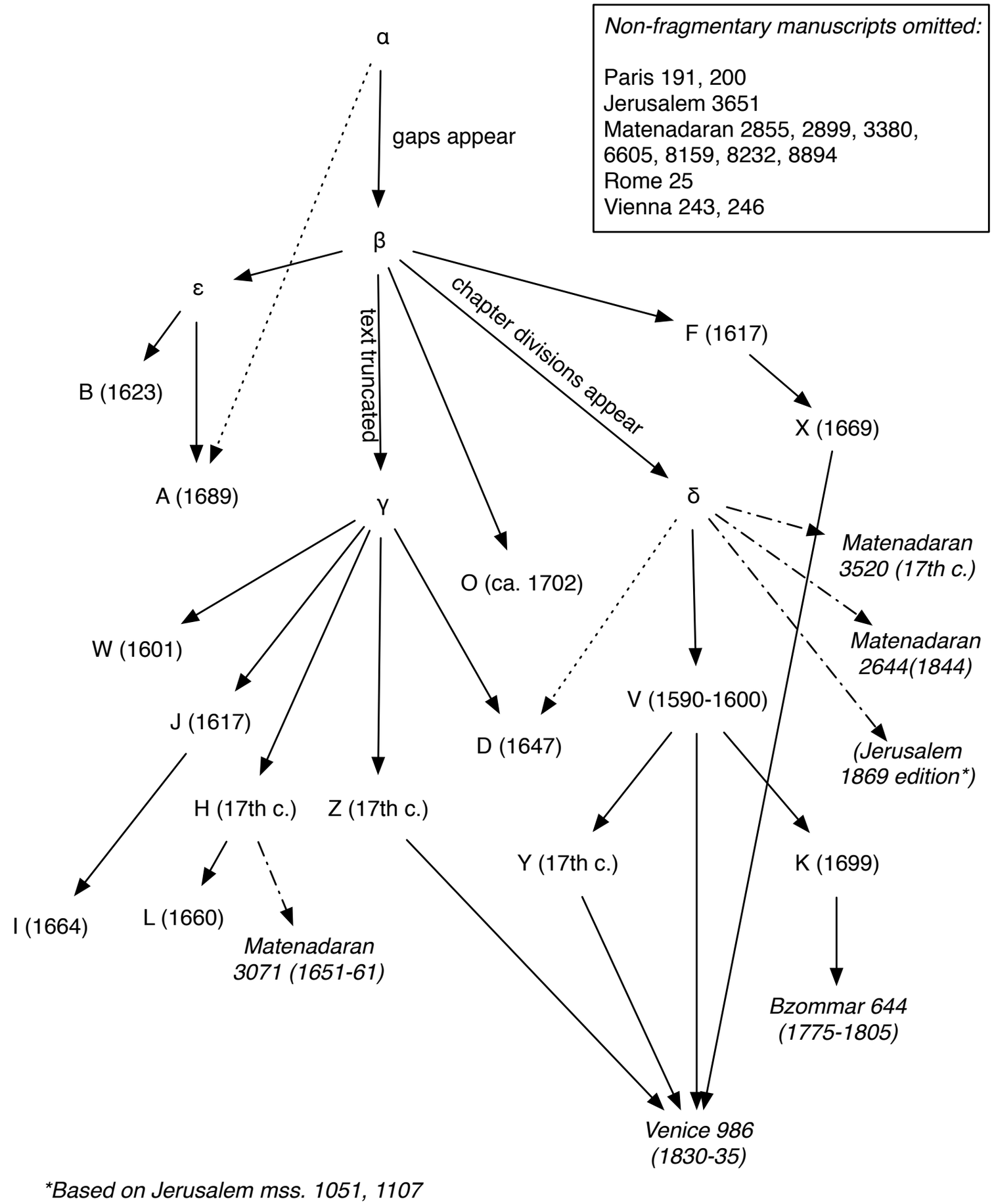

Bzommar 449 / K
Jerusalem 1869
Venice 913 / Y
Venice 887 / V
Matenadaran 1731 / F
Venice 901 / X
Matenadaran 1767 / B
Matenadaran 1896 / A
Bodleian MS Arm e.32 / O
Vienna 574 / W
Matenadaran 1769 / J
Matenadaran 5587 / T
Matenadaran 1768 / H
London Or.5260 / L
Matenadaran 3519 / D
Venice 917 / Z



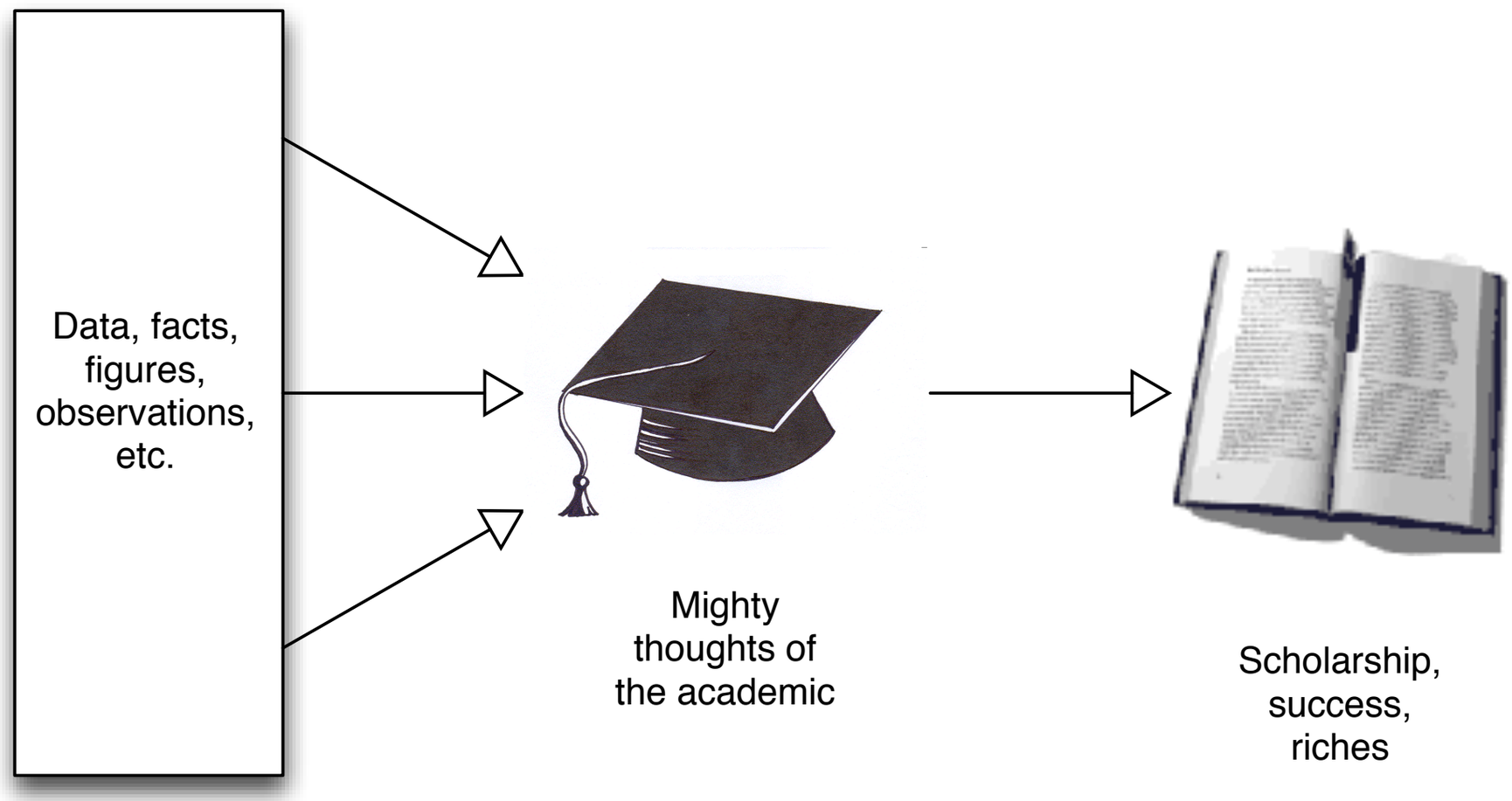


“In this year there appeared a hairy star [i.e. comet] moving toward the west in the month of Areg and its tail gave off a little of its light; and it was there for 15 days and then became invisible to viewers. O kind and friendly brothers, pardon the lack of words, because the exemplar ended here because the text was not finished. Remember to Jesus Christ the sinful and foolish scribe Khachik, and my parents, and you shall also be remembered by the same Lord.

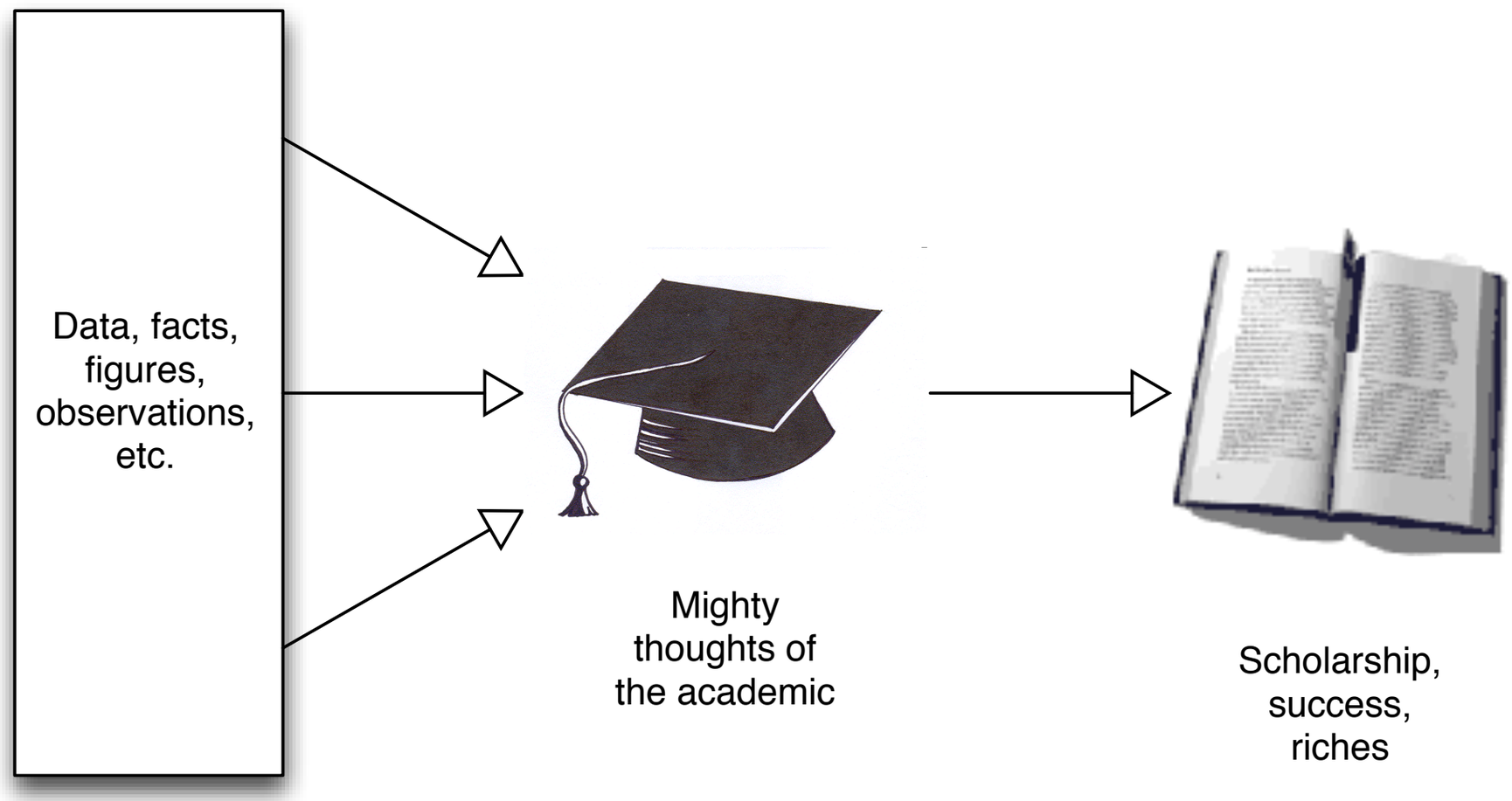

Bzommar 449 / K
Jerusalem 1869
Venice 913 / Y
Venice 887 / V
Matenadaran 1731 / F
Venice 901 / X
Matenadaran 1767 / B
Matenadaran 1896 / A
Bodleian MS Arm e.32 / O
Vienna 574 / W
Matenadaran 1769 / J
Matenadaran 5587 / T
Matenadaran 1768 / H
London Or.5260 / L
Matenadaran 3519 / D
Venice 917 / Z


Non-fragmentary manuscripts omitted:
Paris 191, 200
Jerusalem 3651
Matenadaran 2855, 2899, 3380, 6605, 8159, 8232, 8894
Rome 25
Vienna 243, 246
α
gaps appear
β
ε
B (1623)
A (1689)
text truncated
γ
chapter divisions appear
F (1617)
X (1669)
δ
Matenadaran 3520 (17th c.)
Matenadaran 2644(1844)
(Jerusalem 1869 edition*)
O (ca. 1702)
W (1601)
J (1617)
D (1647)
V (1590-1600)
H (17th c.)
Z (17th c.)
Y (17th c.)
K (1699)
I (1664)
L (1660)
Matenadaran 3071 (1651-61)
Bzommar 644 (1775-1805)
Venice 986 (1830-35)
*Based on Jerusalem mss. 1051, 1107

How academics see their jobs

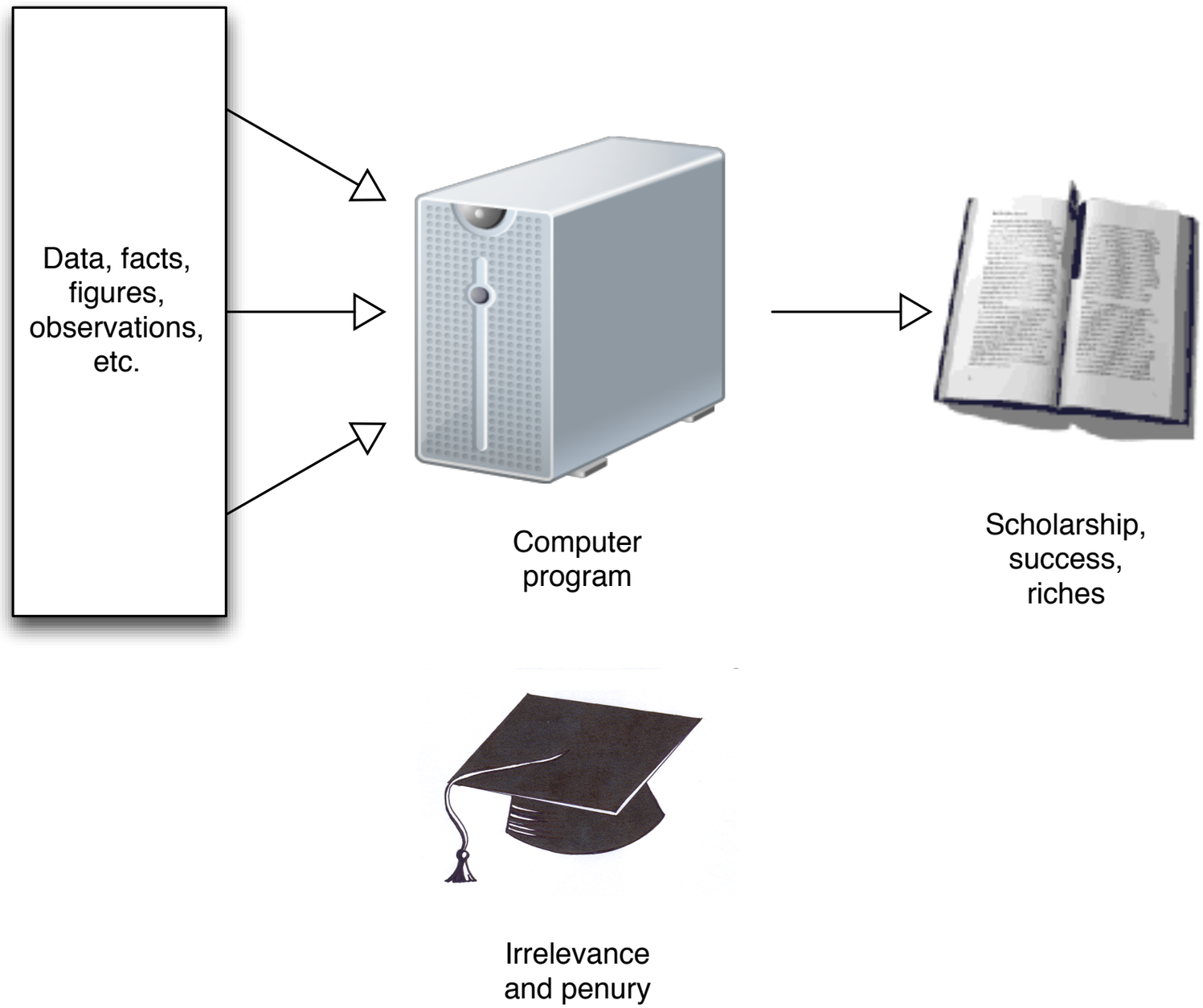


How academics see their jobs

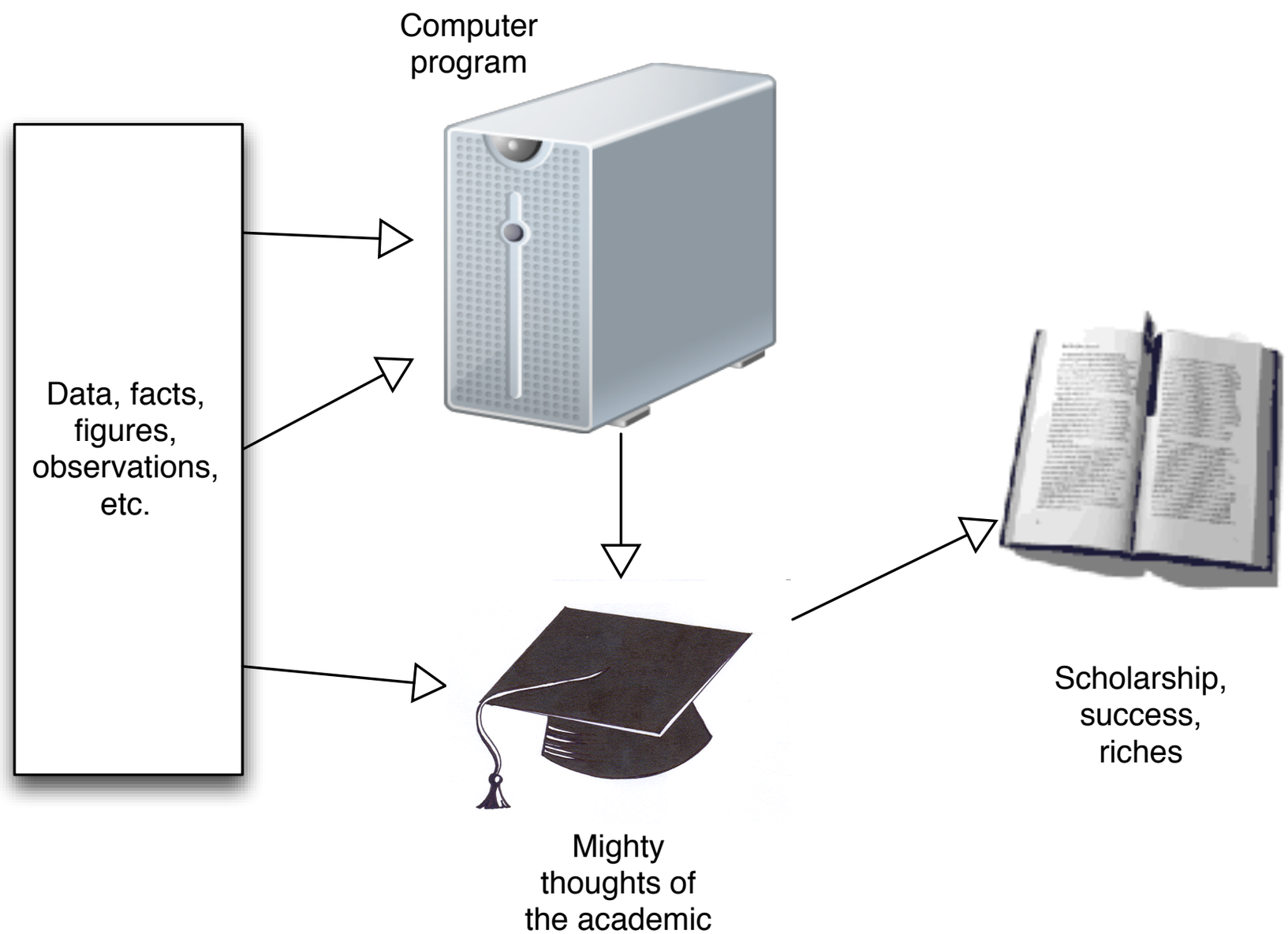


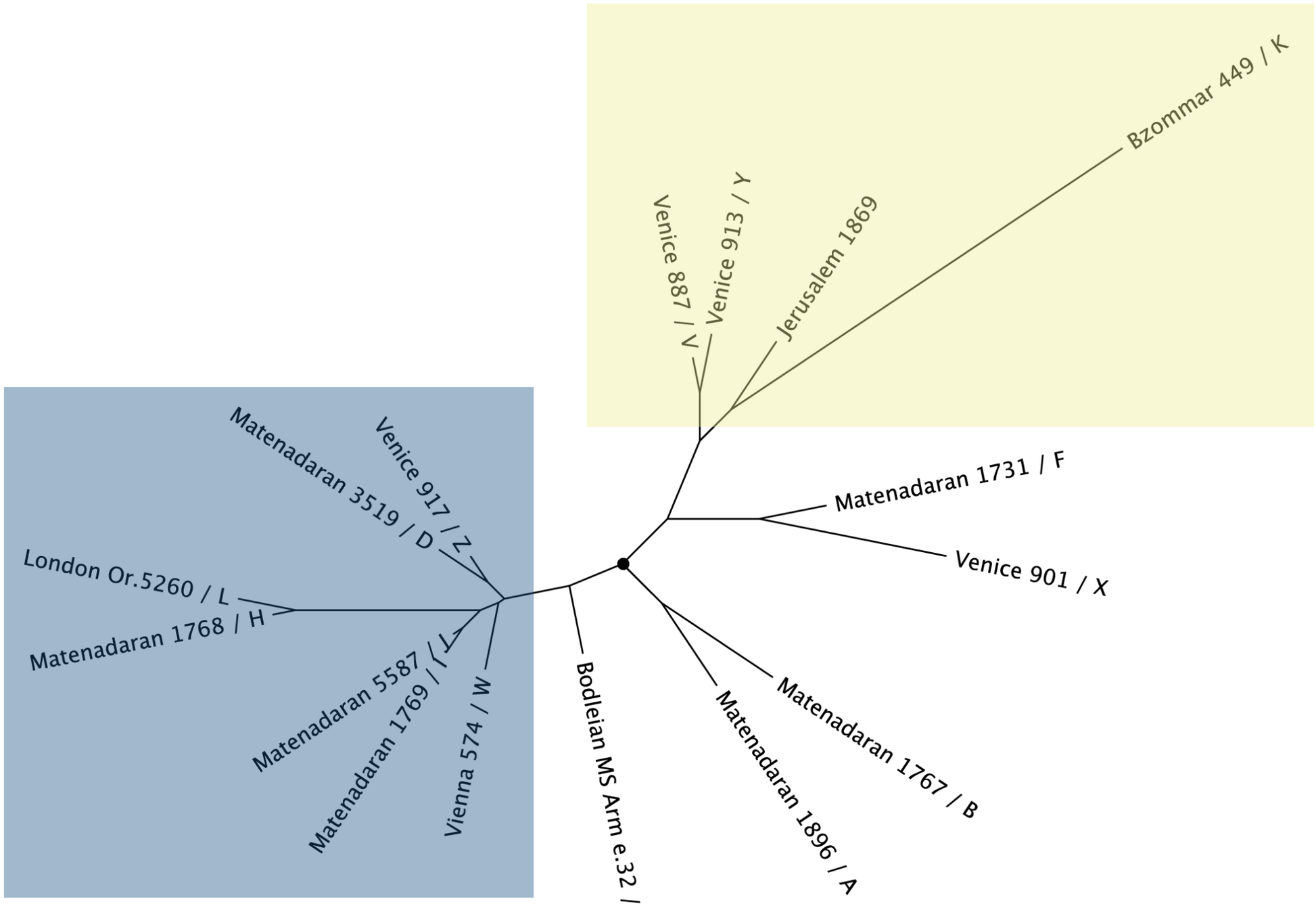
*warning: gross oversimplification and/or wishful thinking

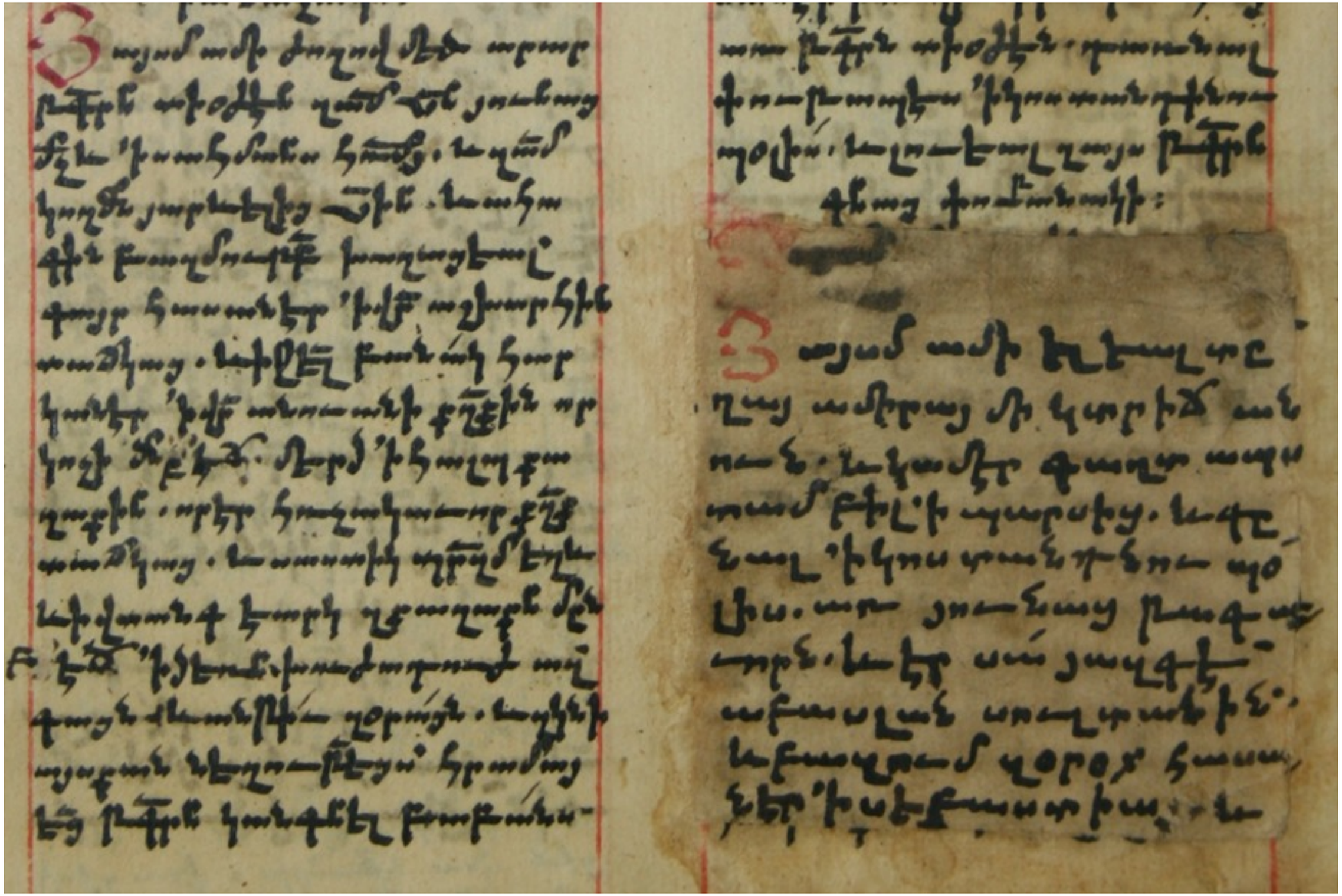
How academics see computers in their jobs

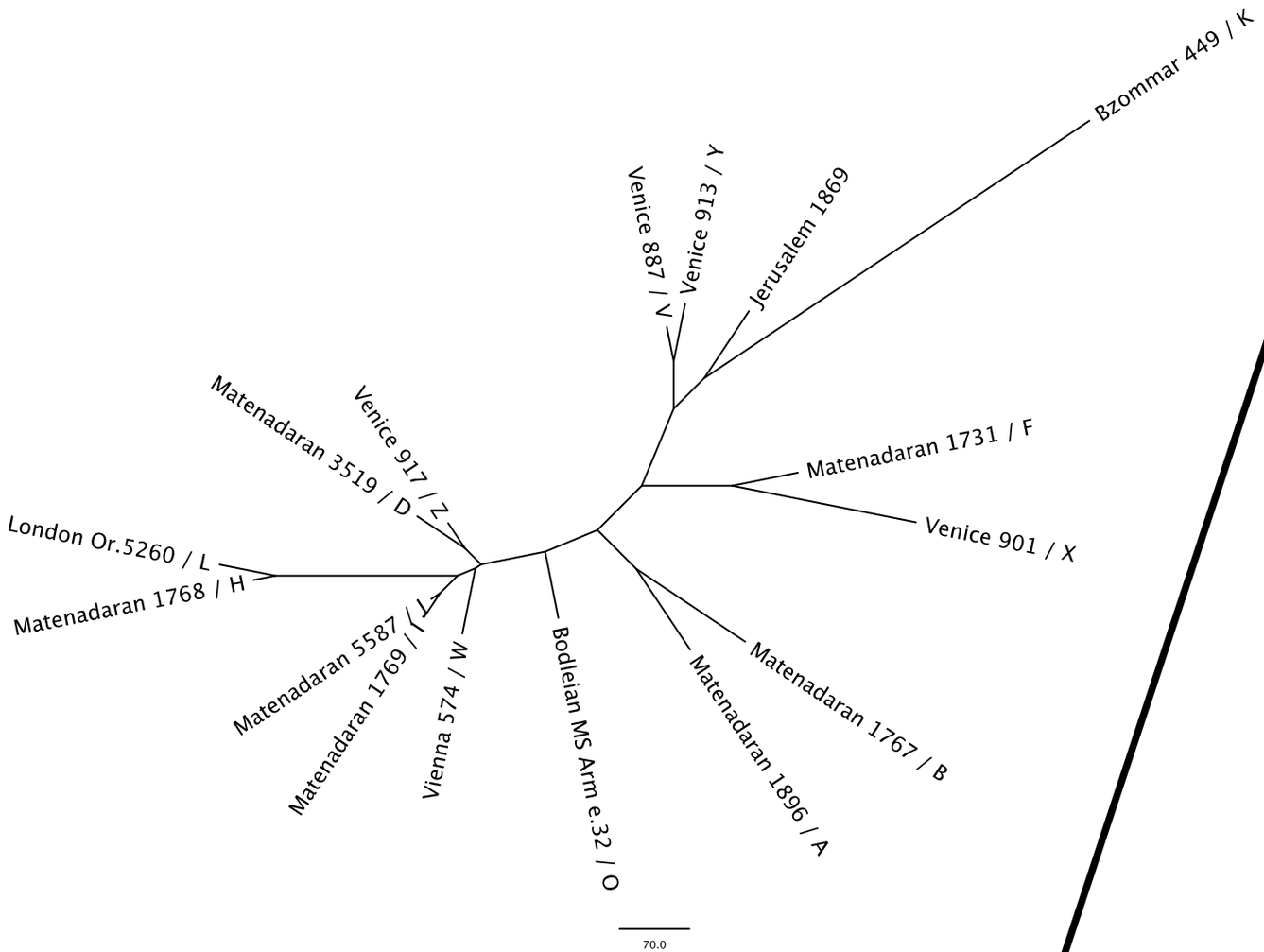


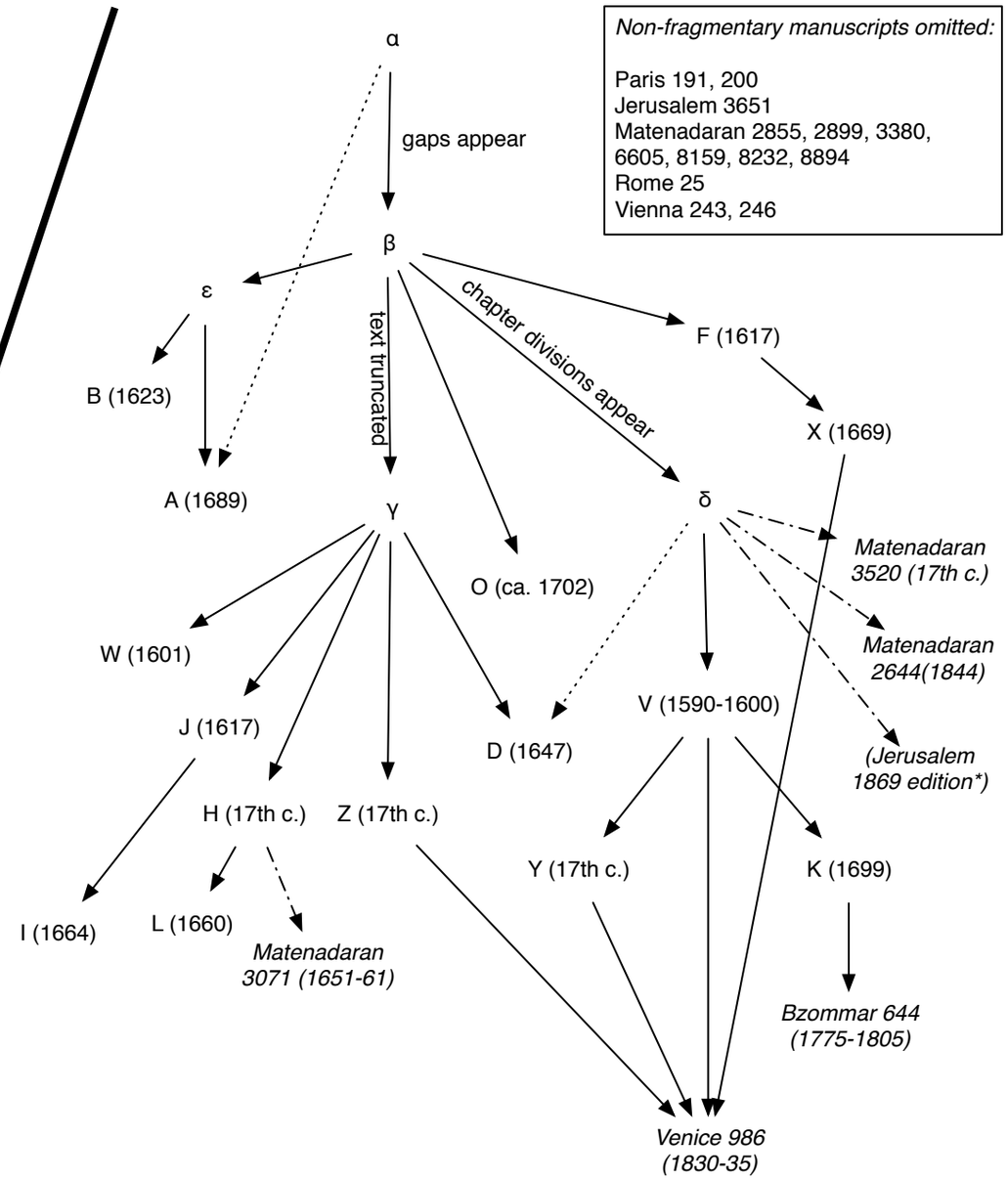
How academics should see computers

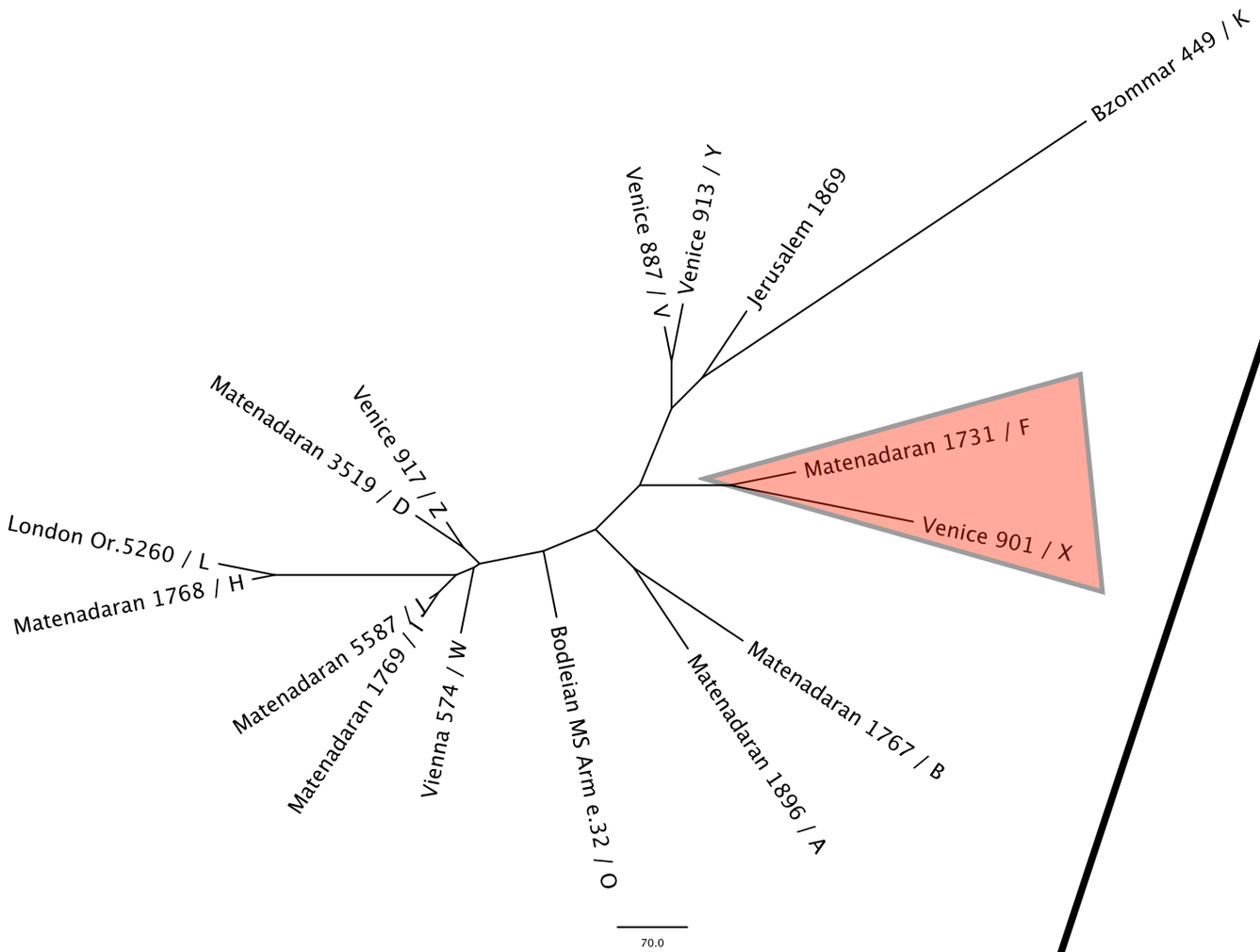


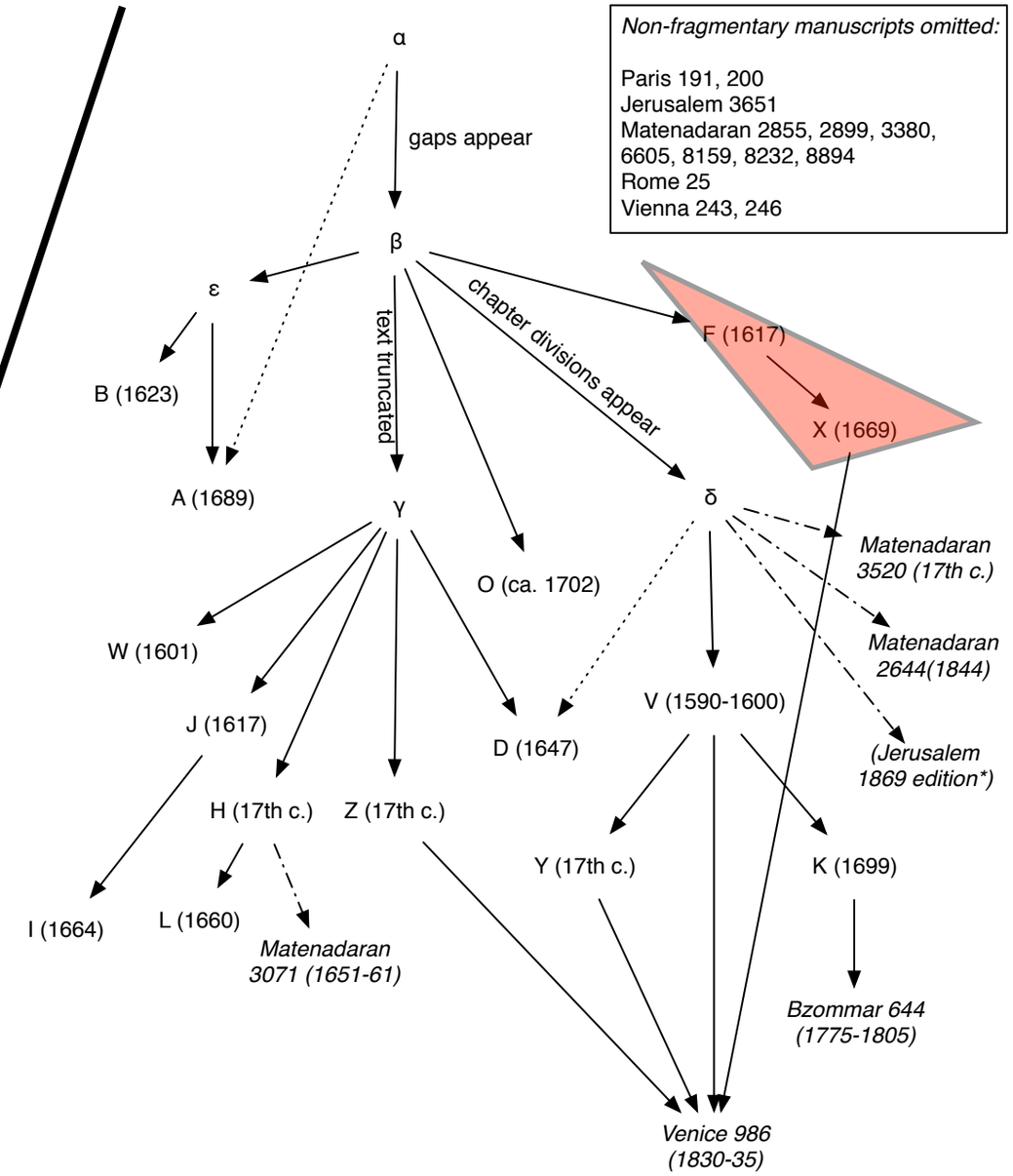

Bzommar 449 / K
Jerusalem 1869
Venice 913 / Y
Venice 887 / V
Matenadaran 1731 / F
Venice 901 / X
Matenadaran 1767 / B
Matenadaran 1896 / A
Bodleian MS Arm e.32 / O
Vienna 574 / W
Matenadaran 1769 / J
Matenadaran 5587 / T
Matenadaran 1768 / H
London Or.5260 / L
Matenadaran 3519 / D
Venice 917 / Z

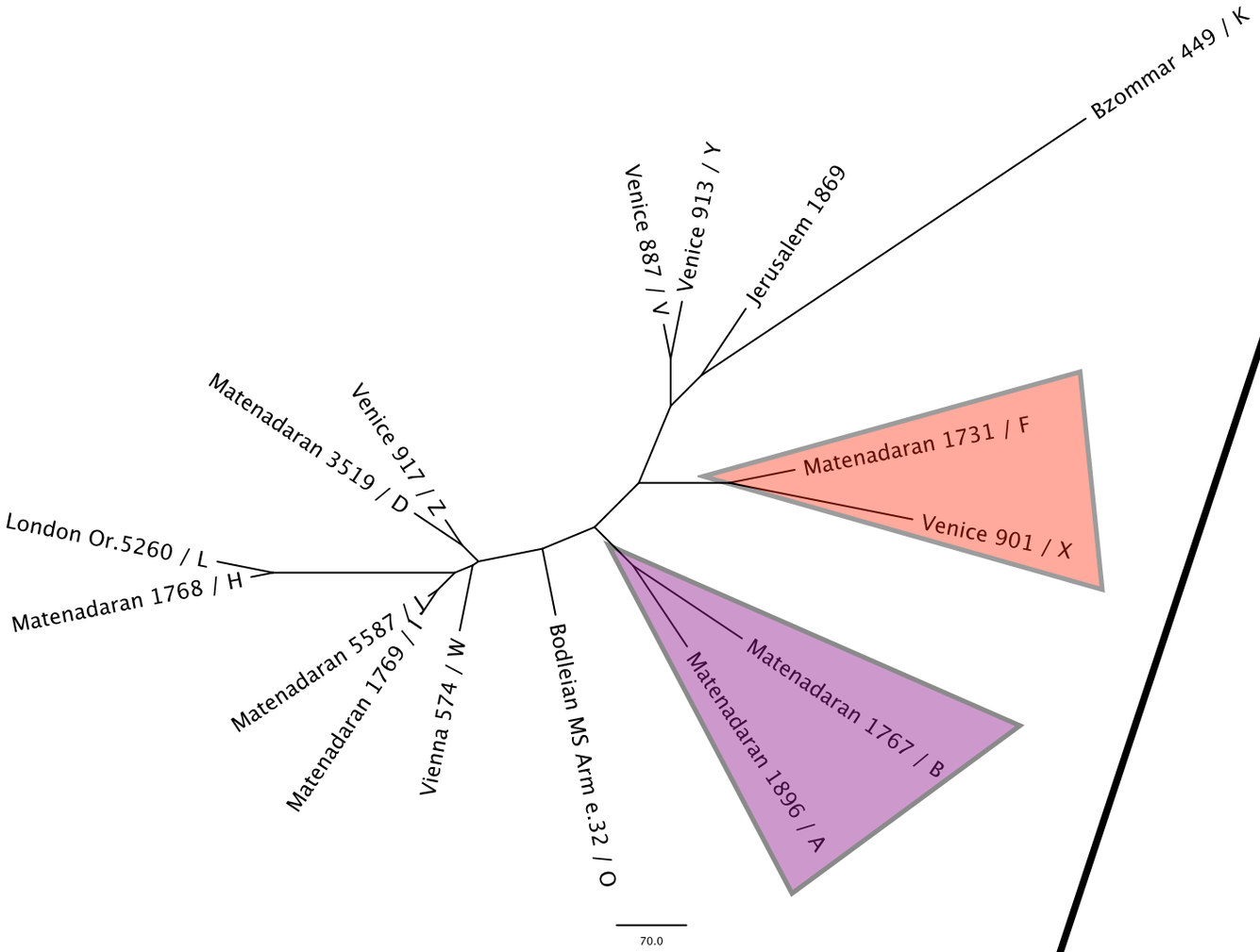


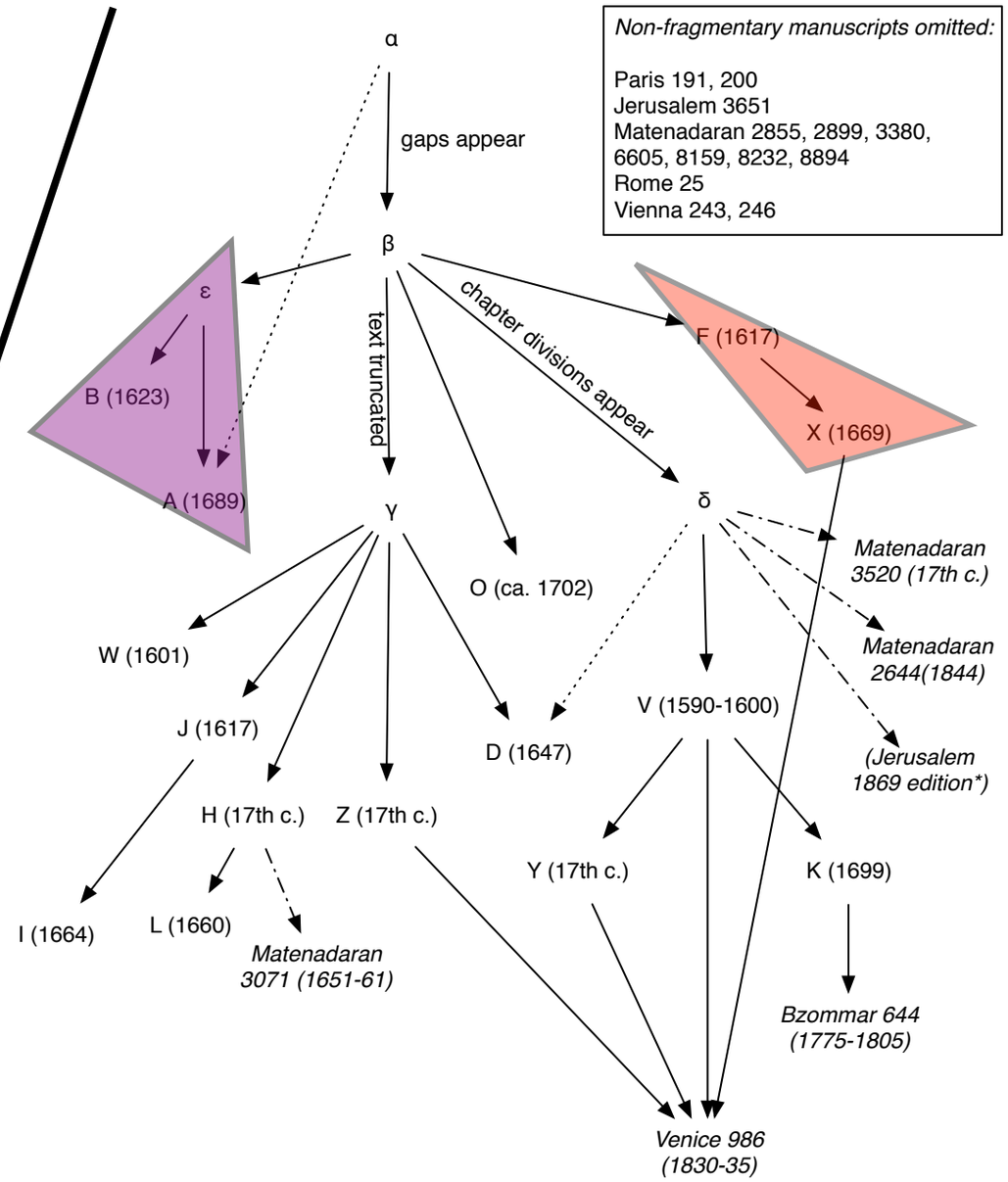

Bzommar 449 / K
Venice 913 / Y
Venice 887 / V
Jerusalem 1869
Matenadaran 1731 / F
Venice 901 / X
Venice 917 / Z
Matenadaran 3519 / D
London Or.5260 / L
Matenadaran 1768 / H
Matenadaran 5587 / J
Matenadaran 1769 / I
Vienna 574 / W
Bodleian MS Arm e.32 / O
Matenadaran 1896 / A
Matenadaran 1767 / B
70.0

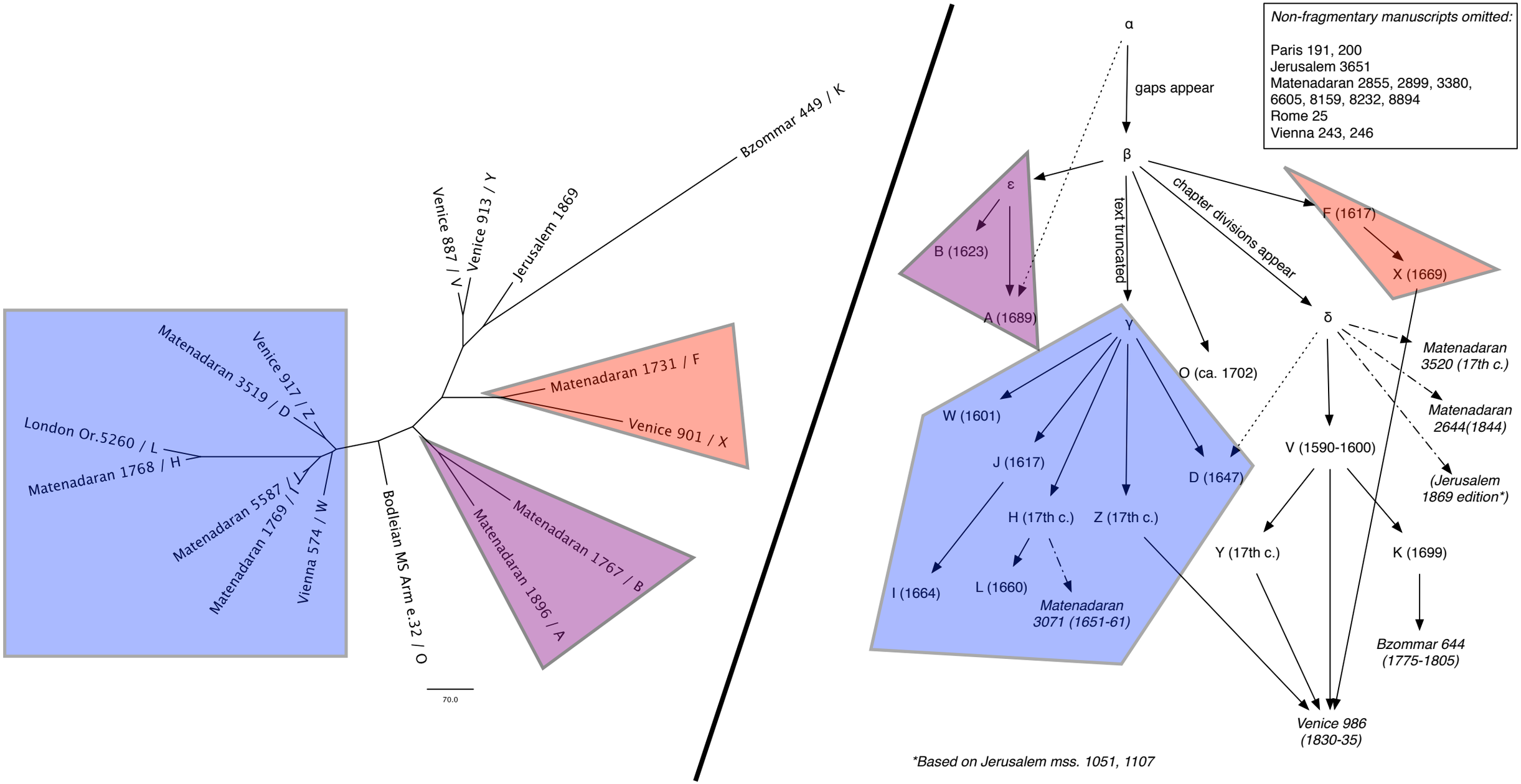

Non-fragmentary manuscripts omitted:
Paris 191, 200
Jerusalem 3651
Matenadaran 2855, 2899, 3380, 6605, 8159, 8232, 8894
Rome 25
Vienna 243, 246
α
gaps appear
β
ε
text truncated
chapter divisions appear
F (1617)
B (1623)
X (1669)
A (1689)
γ
δ
Matenadaran 3520 (17th c.)
O (ca. 1702)
Matenadaran 2644(1844)
W (1601)
V (1590-1600)
J (1617)
D (1647)
(Jerusalem 1869 edition*)
H (17th c.)
Z (17th c.)
Y (17th c.)
K (1699)
I (1664)
L (1660)
Matenadaran 3071 (1651-61)
Bzommar 644 (1775-1805)
Venice 986 (1830-35)
*Based on Jerusalem mss. 1051, 1107

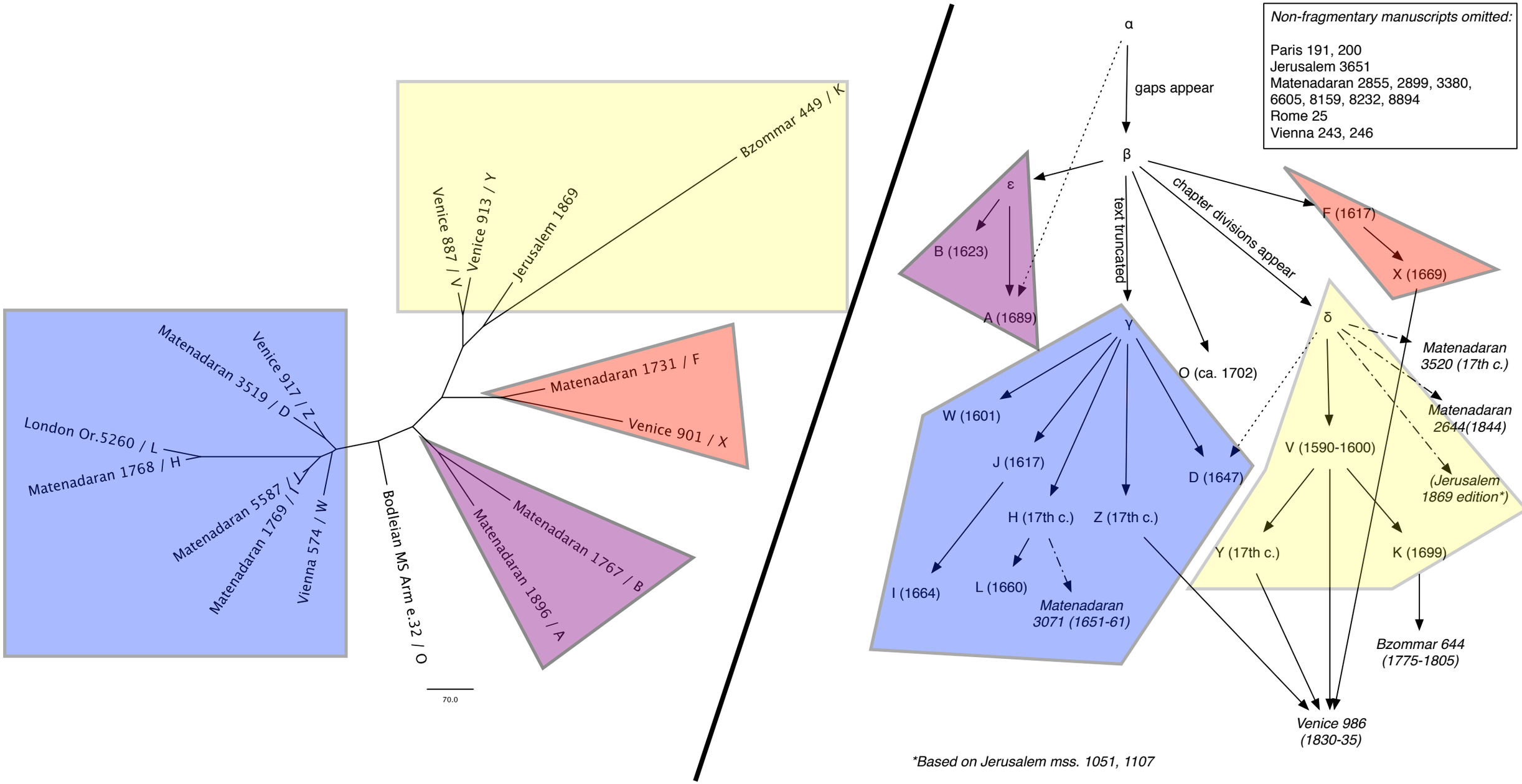

Bzommar 449 / K
Jerusalem 1869
Venice 913 / Y
Venice 887 / V
Matenadaran 1731 / F
Venice 901 / X
Venice 917 / Z
Matenadaran 3519 / D
London Or.5260 / L
Matenadaran 1768 / H
Matenadaran 5587 / L
Matenadaran 1769 / I
Vienna 574 / W
Bodleian MS Arm e.32 / O
Matenadaran 1896 / A
Matenadaran 1767 / B
70.0


Non-fragmentary manuscripts omitted:
Paris 191, 200
Jerusalem 3651
Matenadaran 2855, 2899, 3380, 6605, 8159, 8232, 8894
Rome 25
Vienna 243, 246
α
gaps appear
β
ε
B (1623)
A (1689)
text truncated
chapter divisions appear
F (1617)
X (1669)
γ
δ
O (ca. 1702)
Matenadaran 3520 (17th c.)
Matenadaran 2644(1844)
W (1601)
J (1617)
V (1590-1600)
D (1647)
(Jerusalem 1869 edition*)
H (17th c.)
Z (17th c.)
Y (17th c.)
K (1699)
I (1664)
L (1660)
Matenadaran 3071 (1651-61)
Bzommar 644 (1775-1805)
Venice 986 (1830-35)
*Based on Jerusalem mss. 1051, 1107


Bzommar 449 / K
Jerusalem 1869
Venice 913 / Y
Venice 887 / V
Matenadaran 1731 / F
Venice 901 / X
Venice 917 / Z
Matenadaran 3519 / D
London Or.5260 / L
Matenadaran 1768 / H
Matenadaran 5587 / L
Matenadaran 1769 / I
Vienna 574 / W
Bodleian MS Arm e.32 / O
Matenadaran 1767 / B
Matenadaran 1896 / A
70.0


Non-fragmentary manuscripts omitted:
Paris 191, 200
Jerusalem 3651
Matenadaran 2855, 2899, 3380, 6605, 8159, 8232, 8894
Rome 25
Vienna 243, 246
α
gaps appear
β
ε
B (1623)
A (1689)
text truncated
chapter divisions appear
F (1617)
X (1669)
γ
δ
O (ca. 1702)
Matenadaran 3520 (17th c.)
Matenadaran 2644(1844)
W (1601)
J (1617)
V (1590-1600)
D (1647)
(Jerusalem 1869 edition*)
H (17th c.)
Z (17th c.)
Y (17th c.)
K (1699)
I (1664)
L (1660)
Matenadaran 3071 (1651-61)
Bzommar 644 (1775-1805)
Venice 986 (1830-35)
*Based on Jerusalem mss. 1051, 1107


Bzommar 449 / K
Venice 887 / V
Venice 913 / Y
Jerusalem 1869
Matenadaran 1731 / F
Venice 901 / X
Venice 917 / Z
Matenadaran 3519 / D
London Or.5260 / L
Matenadaran 1768 / H
Matenadaran 5587 / J
Matenadaran 1769 / I
Vienna 574 / W
Bodleian MS Arm e.32 / O
Matenadaran 1767 / B
Matenadaran 1896 / A
70.0
α
gaps appear
β
text truncated
chapter divisions appear
ε
B (1623)
A (1689)
F (1617)
X (1669)
γ
O (ca. 1702)
δ
Matenadaran 3520 (17th c.)
Matenadaran 2644(1844)
(Jerusalem 1869 edition*)
W (1601)
J (1617)
D (1647)
V (1590-1600)
H (17th c.)
Z (17th c.)
Y (17th c.)
K (1699)
I (1664)
L (1660)
Matenadaran 3071 (1651-61)
Bzommar 644 (1775-1805)
Venice 986 (1830-35)
Non-fragmentary manuscripts omitted:
Paris 191, 200
Jerusalem 3651
Matenadaran 2855, 2899, 3380, 6605, 8159, 8232, 8894
Rome 25
Vienna 243, 246
*Based on Jerusalem mss. 1051, 1107


Bzommar 449 / K
Venice 887 / V
Venice 913 / Y
Jerusalem 1869
Matenadaran 1731 / F
Venice 901 / X
Venice 917 / Z
Matenadaran 3519 / D
London Or.5260 / L
Matenadaran 1768 / H
Matenadaran 5587 / I
Matenadaran 1769 / J
Vienna 574 / W
Bodleian MS Arm e.32 / O
Matenadaran 1767 / B
Matenadaran 1896 / A
70.0
Non-fragmentary manuscripts omitted:
Paris 191, 200
Jerusalem 3651
Matenadaran 2855, 2899, 3380, 6605, 8159, 8232, 8894
Rome 25
Vienna 243, 246
α
gaps appear
β
text truncated
chapter divisions appear
ε
B (1623)
A (1689)
F (1617)
X (1669)
γ
O (ca. 1702)
δ
Matenadaran 3520 (17th c.)
Matenadaran 2644(1844)
(Jerusalem 1869 edition*)
W (1601)
J (1617)
D (1647)
V (1590-1600)
H (17th c.)
Z (17th c.)
Y (17th c.)
K (1699)
I (1664)
L (1660)
Matenadaran 3071 (1651-61)
Bzommar 644 (1775-1805)
Venice 986 (1830-35)
*Based on Jerusalem mss. 1051, 1107

Thank you

Obrigada

Շնորհակալություն

[shnorhakalutiun]