

# Manuscript genetics and Perl

or, How I made sense of 101011 manuscripts

Tara L Andrews  
Departmental Lecturer  
Faculty of History, University of Oxford  
[sla@mit.edu](mailto:sla@mit.edu)



- 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.*)



121

Թուխ սրոյ և մահանալի, շարտրեյ  
եւ ի զիս արժան լարարդեալ . . .  
Հանգստանալիս . . .



Ըրտ ինչի երեւ զարմարել. Ե  
զարշալ շեծած յիշելու. Որ  
ի վարժանիս անտ զարհեցան  
յիշելիցի հոյոյ. Ե զի որ մա անհրմարտ  
յանն անցուս ծանել զանցեալ ծածկարոյն  
խորհուրդս արւսեա արդի, մեզ յոյժ ոչ  
իւրոյ ջաննի զարմիս աս զբարւ հարել  
իսոյնիս օրոր զարմարալիս. Բայց աս  
հոյն յասացել ի յօգնալիս արհոյ արժան.  
յօժարդիս արժանացուս հարձիս ի  
յոյս անհուն ծովս մարնել. Եւ անար զի  
զիս անլնարդի արհ հոյ զարմիս. որ ի  
եւել ի վարժար շարտրեյ, արհ հոյ  
զիս հոյոյ. Որ արհ մա իննու զար

**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 միլ որեւէ , թողուեն զհոգս շինուիթեան առնից և զյաջողուածսս Կորնթոսց , և միջտ և Ըսնաւորազ ի Կիւստանուս \_\_\_\_\_

Following context:

#D : դեպերին , զասն ահրո և Եսնկուիթեան չար  
#W : դեպերեին զասն ահրո և Եսնկուիթեան չար  
#X : դեպերեն : զասն ահրոյ և Եսնկուիթեան չար  
#Y : դեպերին , զասն ահրոց և Եսնկուիթեան չար  
#F : դեպերին , զասն ահրոց և Եսնկուիթեան չար  
#I #Jer #J #K #Z #V #A #B #O : դեպերին , զասն ահրոյ և Եսնկուիթեան չար

Available readings:

1 : դեպերին (#Jer #K #B #I #A #D #J #O #V #Y #Z )  
2 : դեպերեն (#X )  
4 : դեպերին (#F )  
6 : դեպերեին (#W )

Manuscript details:

3: Witness(es) #X contain a : after դեպերեն

5: Witness(es) #F contain a , after դեպերին

--> h

Available commands:

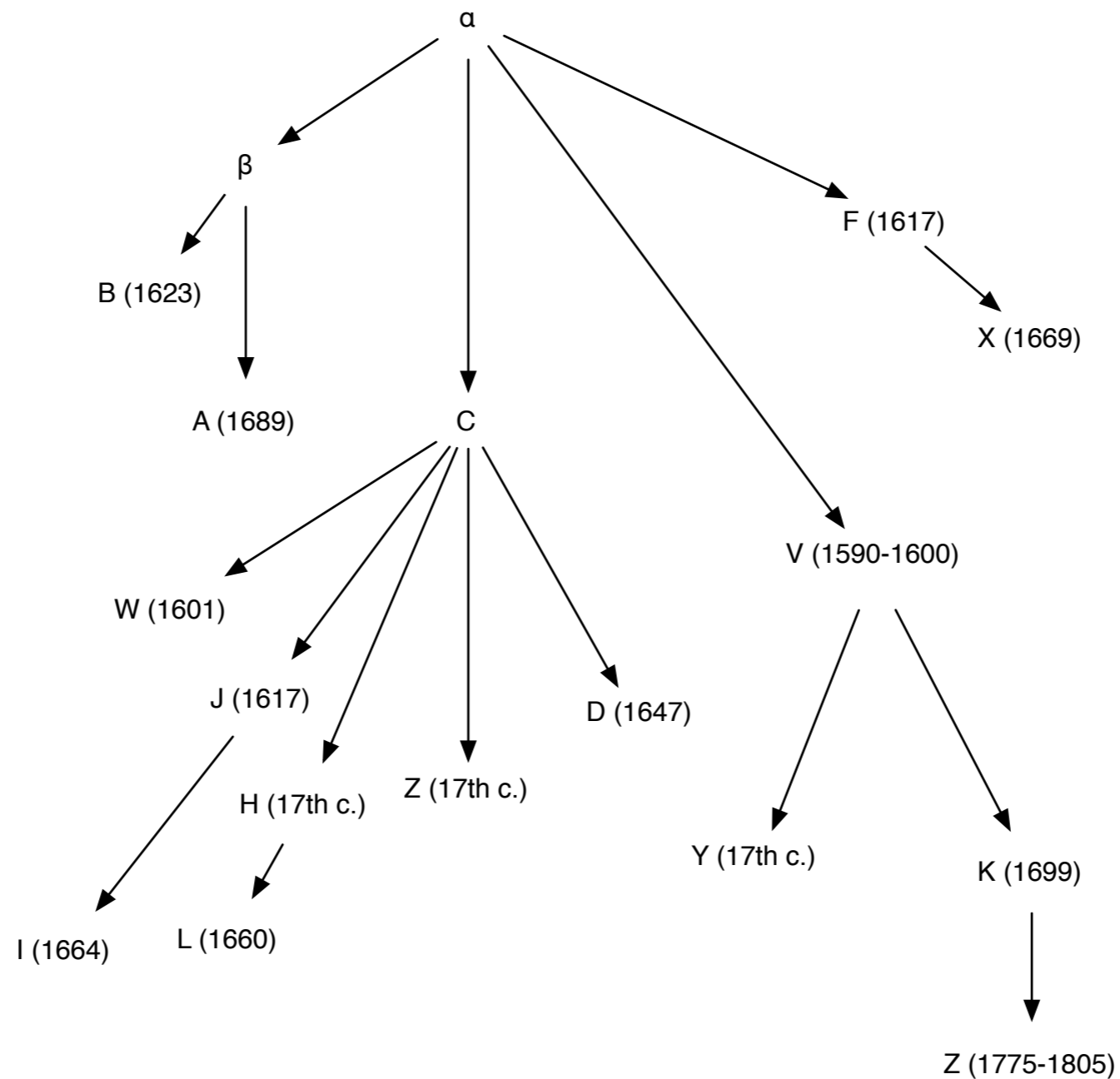
accept (#)  
accept detail (#)  
subst detail (#) (value)  
spelling (# alternate) (# canonical)  
orth (# alternate) (# canonical)  
emend (comment text)  
note (note text)  
next  
help  
save  
quit

-->

# Manuscript genetics

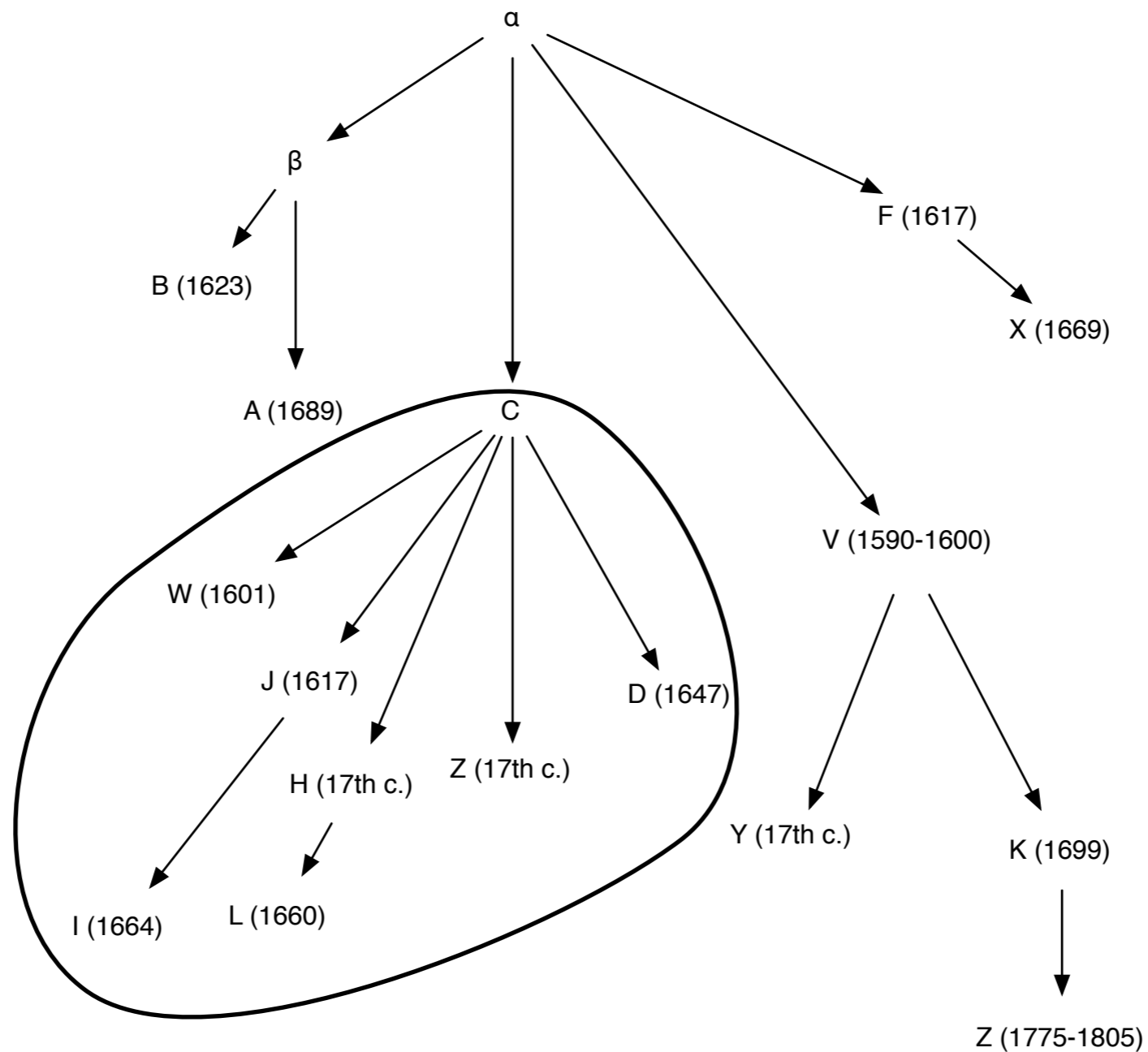
- Not invented here
- But pretty cool anyway
- A way to reconstruct the past, a guide to good editing decisions

# 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  
ամենայն  
եղելոցն,  
նստուցանեն  
զաթոռ  
հայրապետութեան  
ի  
թաւբլուր  
և  
կացեալ  
անդ  
զամս  
գ,  
և  
ընդ  
ամենայն  
զ  
ամ  
կալեալ  
զաթոռ  
հայրապետութեանն  
տէր  
իսաչիկ:

այս  
ամենայն  
եղելոց  
նստուցանեն  
զաթոռ  
հայրապետութեան  
  
  
  
  
  
  
  
  
  
  
  
տէր  
իսաչիկ:

յայսմ  
ամի  
եղելոց  
նստուցանեն  
յաթոռ  
հայրապետութեան  
ի  
թաւբլուրն:  
և  
կացեալ  
անդ  
զամս  
գ,  
և  
ընդ  
ամենայն  
զ  
ամ,  
կալեալ  
զաթոռ  
հայրապետութեան  
տէր  
իսաչիկն:

այս  
ամենայն  
եղելոց  
նստուցանեն  
զաթոռ  
հայրապետութեանն  
ի  
թաւբլուր  
և  
կացեալ  
անդ  
զամս  
գ  
և  
ընդ  
ամենայն  
վեց  
ամ  
կալեալ  
զաթոռ  
հայրապետութեանն  
տէր  
իսաչիկ:

այս  
ամենայն  
եղելոցս  
նստուցանեն  
զաթոռ  
հայրապետութեան  
  
  
  
  
  
  
  
  
  
  
  
զտէր  
իսաչիկ:

A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A

B  
A  
B  
A  
A  
A  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
A  
A

A  
B  
B  
A  
B  
A  
A  
B  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
B  
B

B  
A  
B  
A  
A  
B  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
A  
B  
A  
A  
A  
A  
A

B  
A  
C  
A  
A  
A  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
O  
B  
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 (-h at the end of a word).

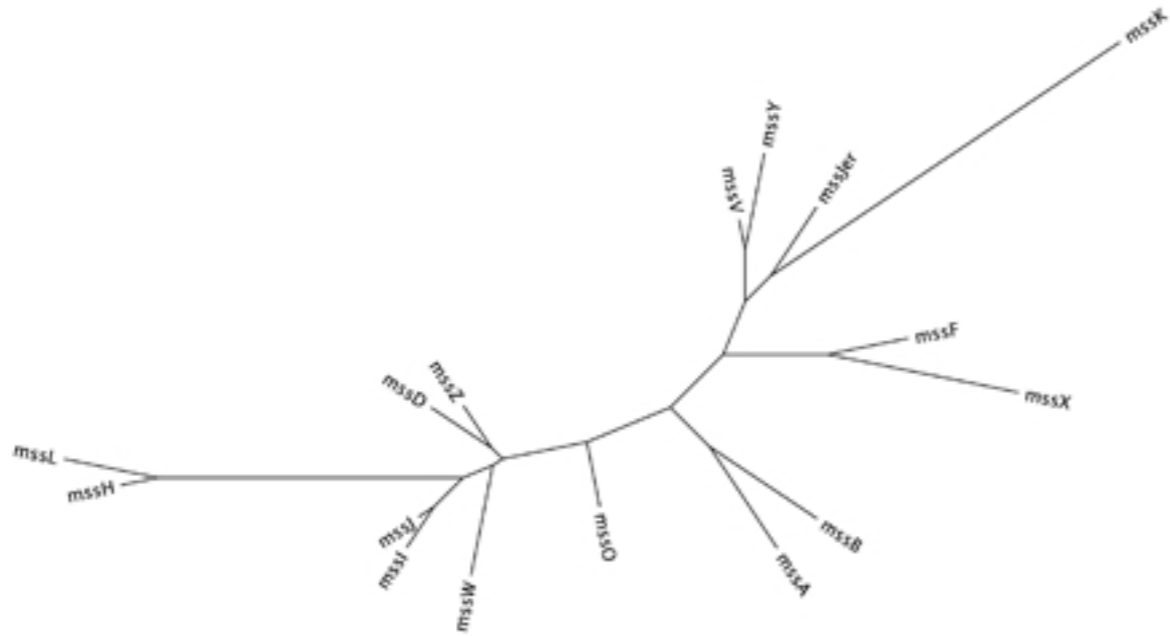
# A little experiment

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

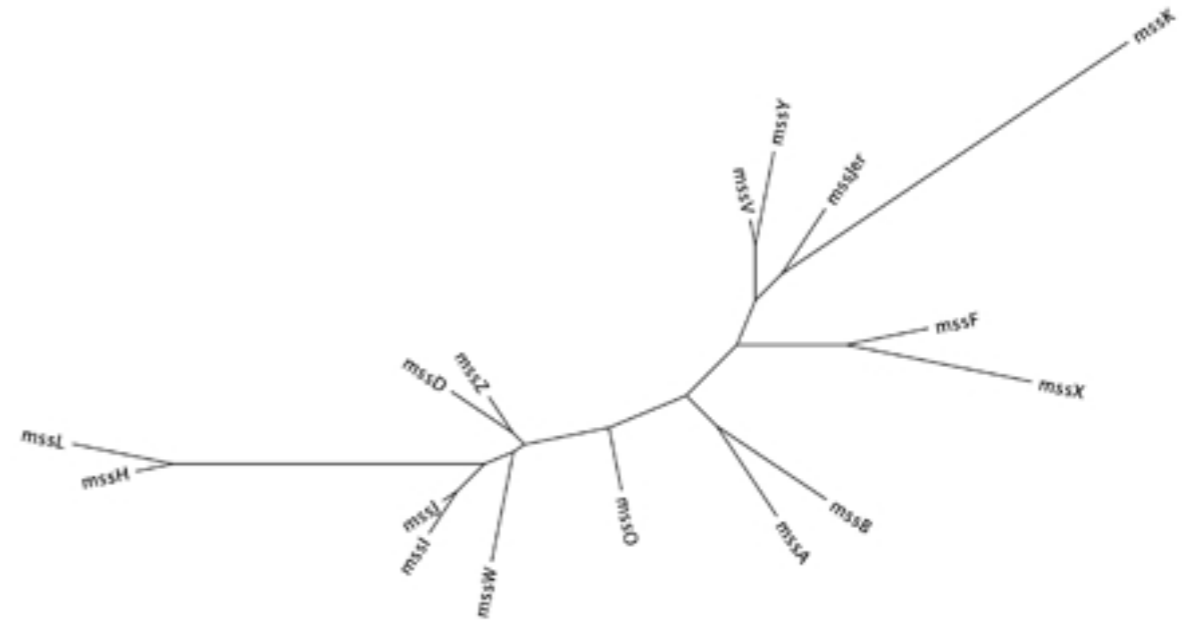
# A little experiment

- Use all the data, corrected for spelling variation.
- Get rid of a common scribal variation (-ŭ 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.

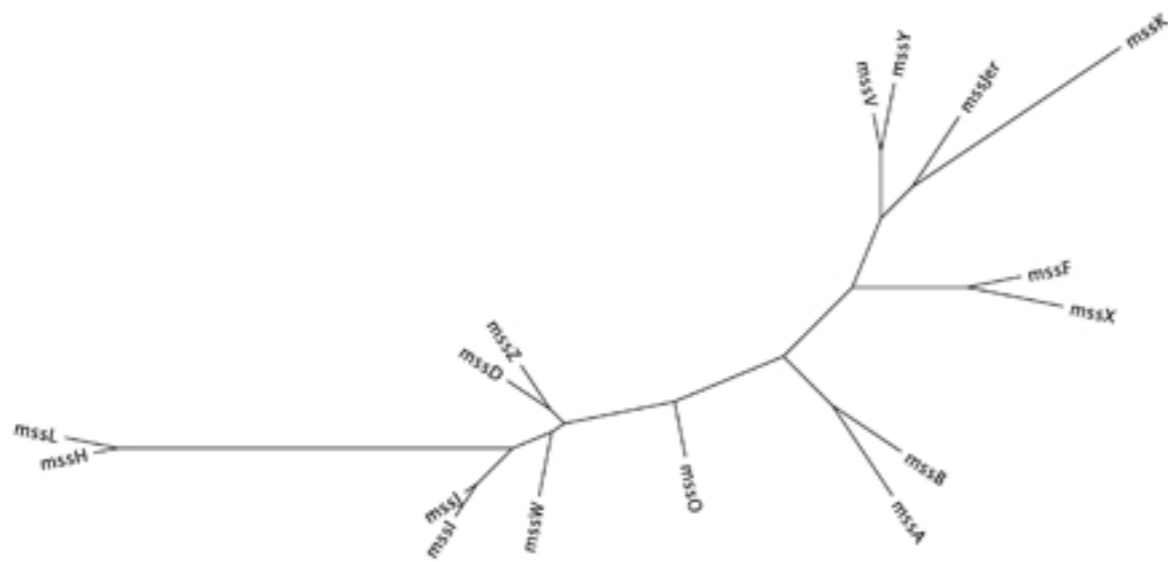
1. All readings



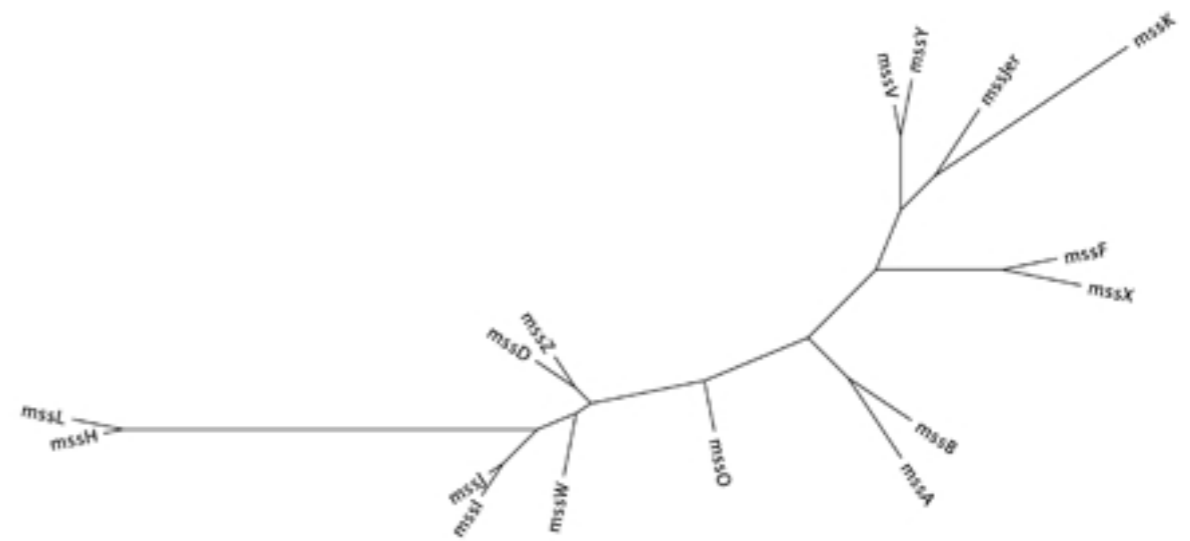
3. Selected readings



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

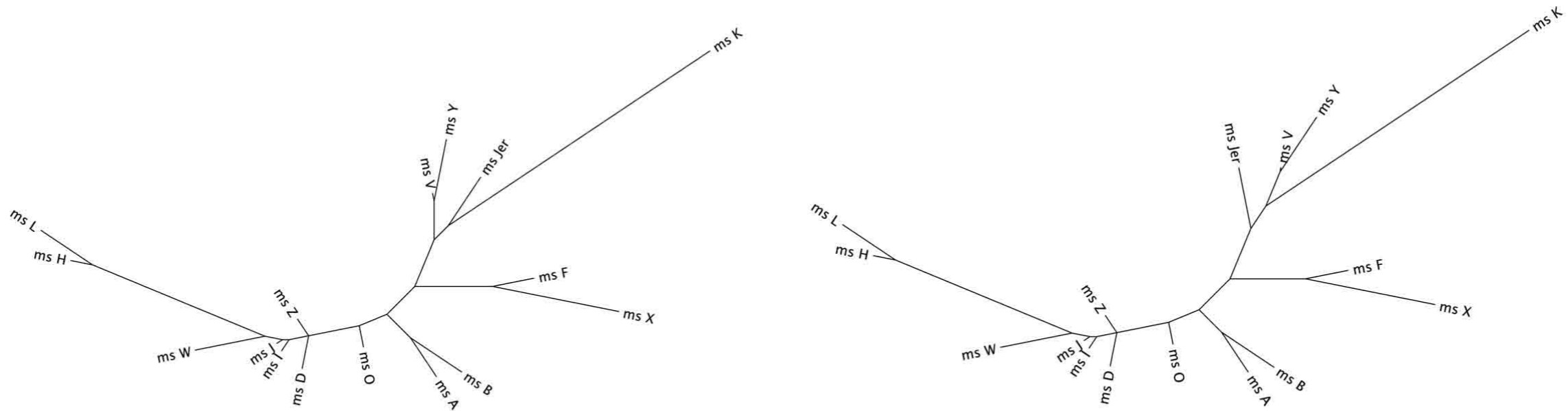


# Sampling techniques

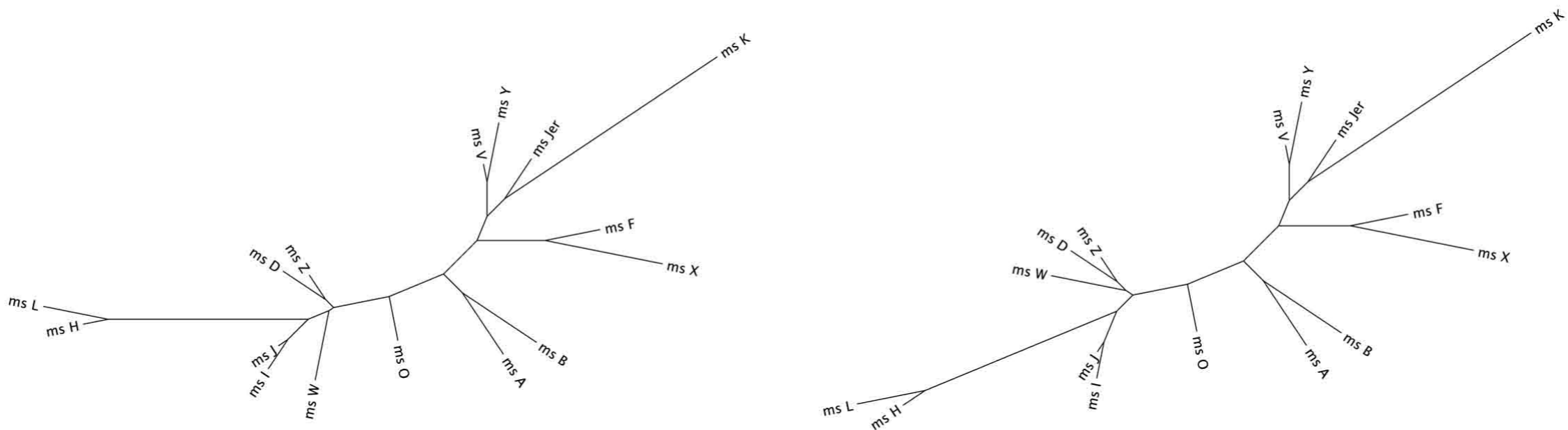
- Three samples, one from each book of the Chronicle
- 5,000 words in total for an 80,000-word text
- Needed to give a single answer

# Sampling techniques

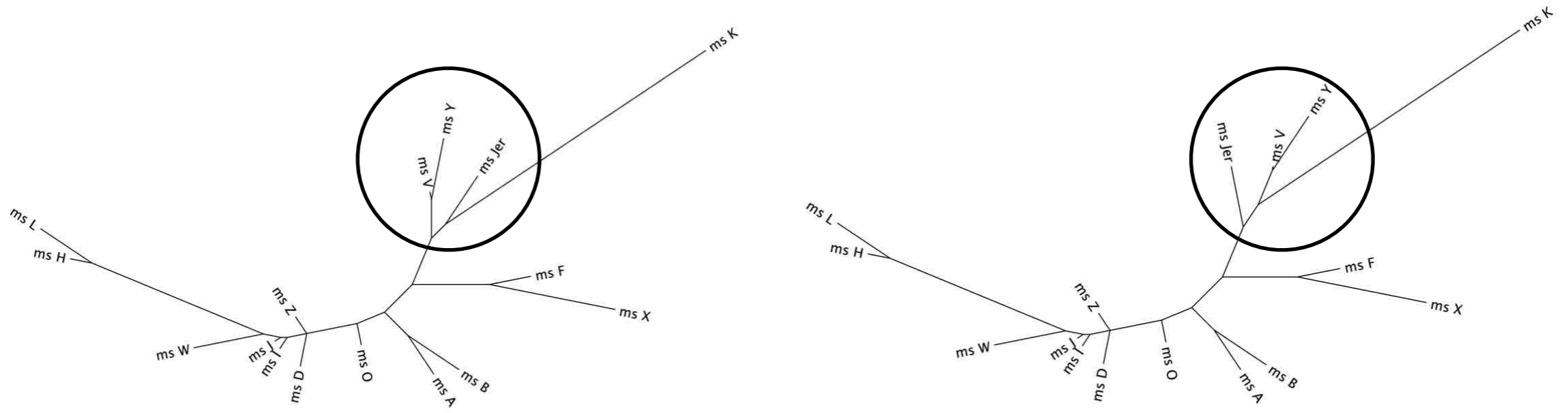
- Three samples, one from each book of the Chronicle
- 5,000 words in total for an 80,000-word text
- Needed to give a single answer
- Needed to give consistent answers across samples



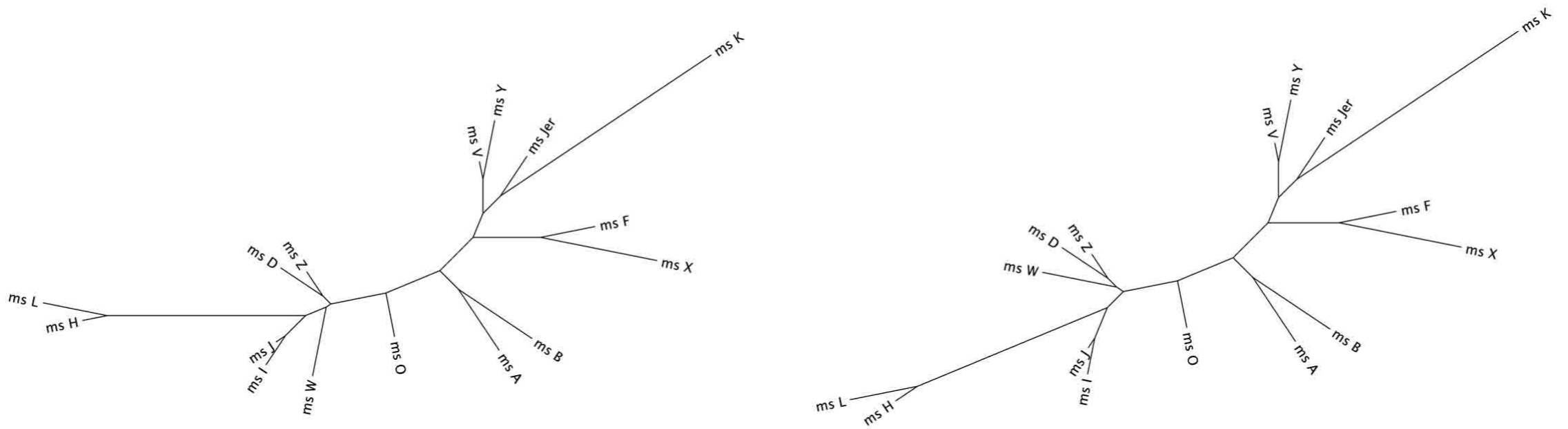
Possible trees from the text of Book 2



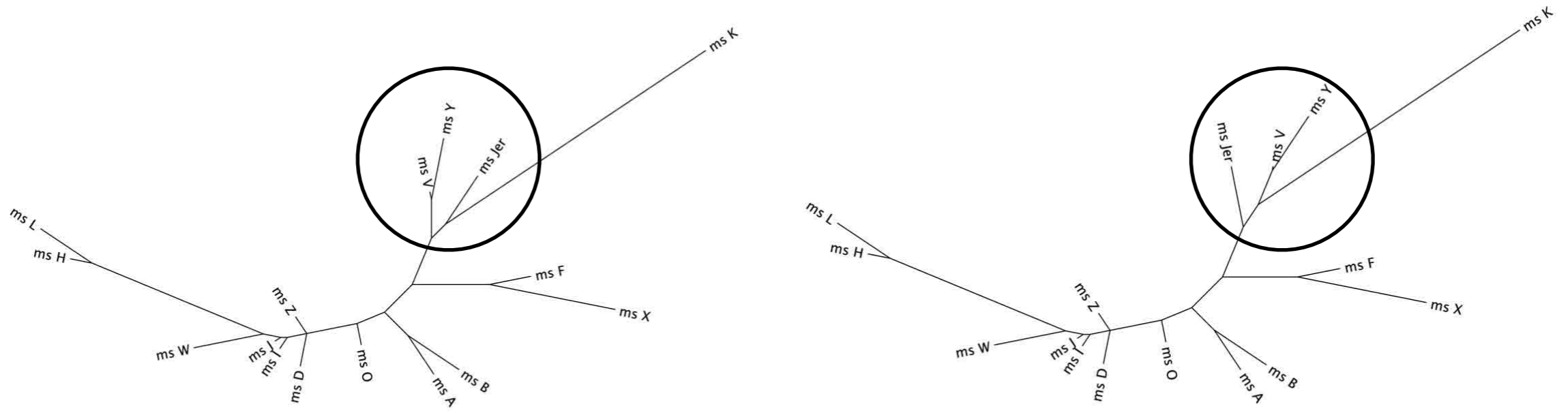
Possible trees from the text of the prophecy



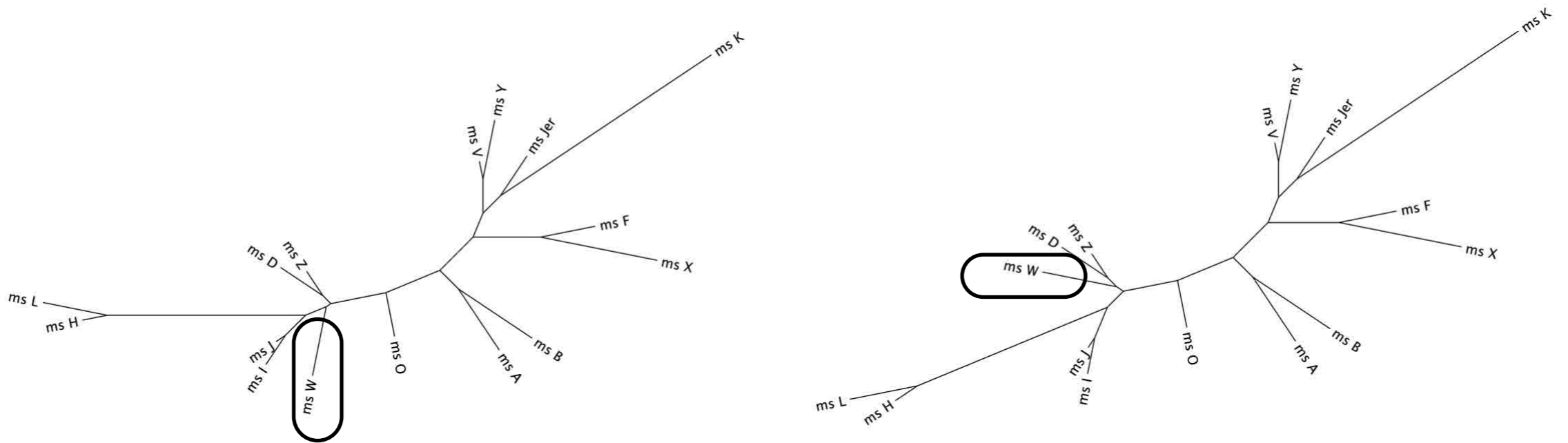
Possible trees from the text of Book 2



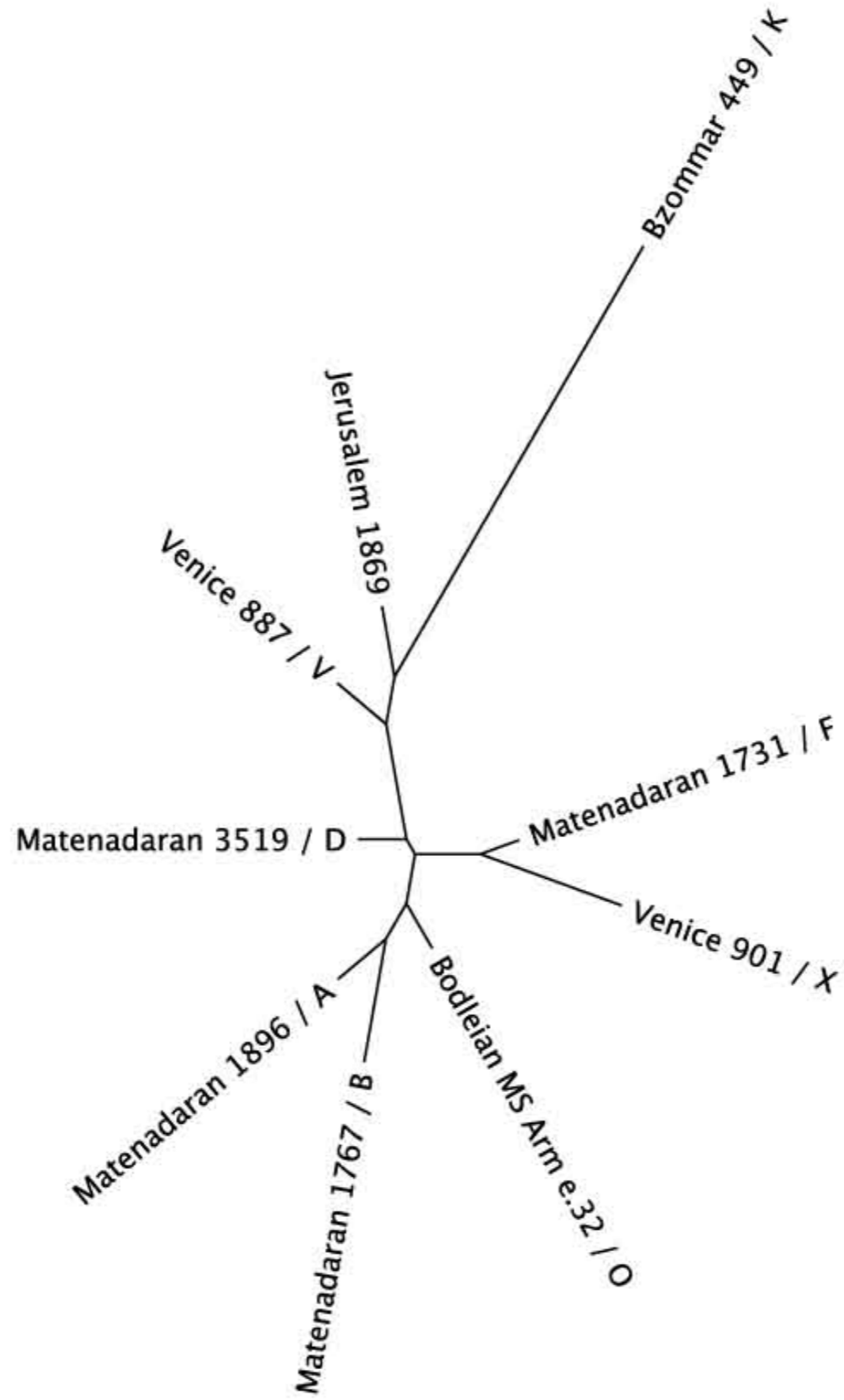
Possible trees from the text of the prophecy

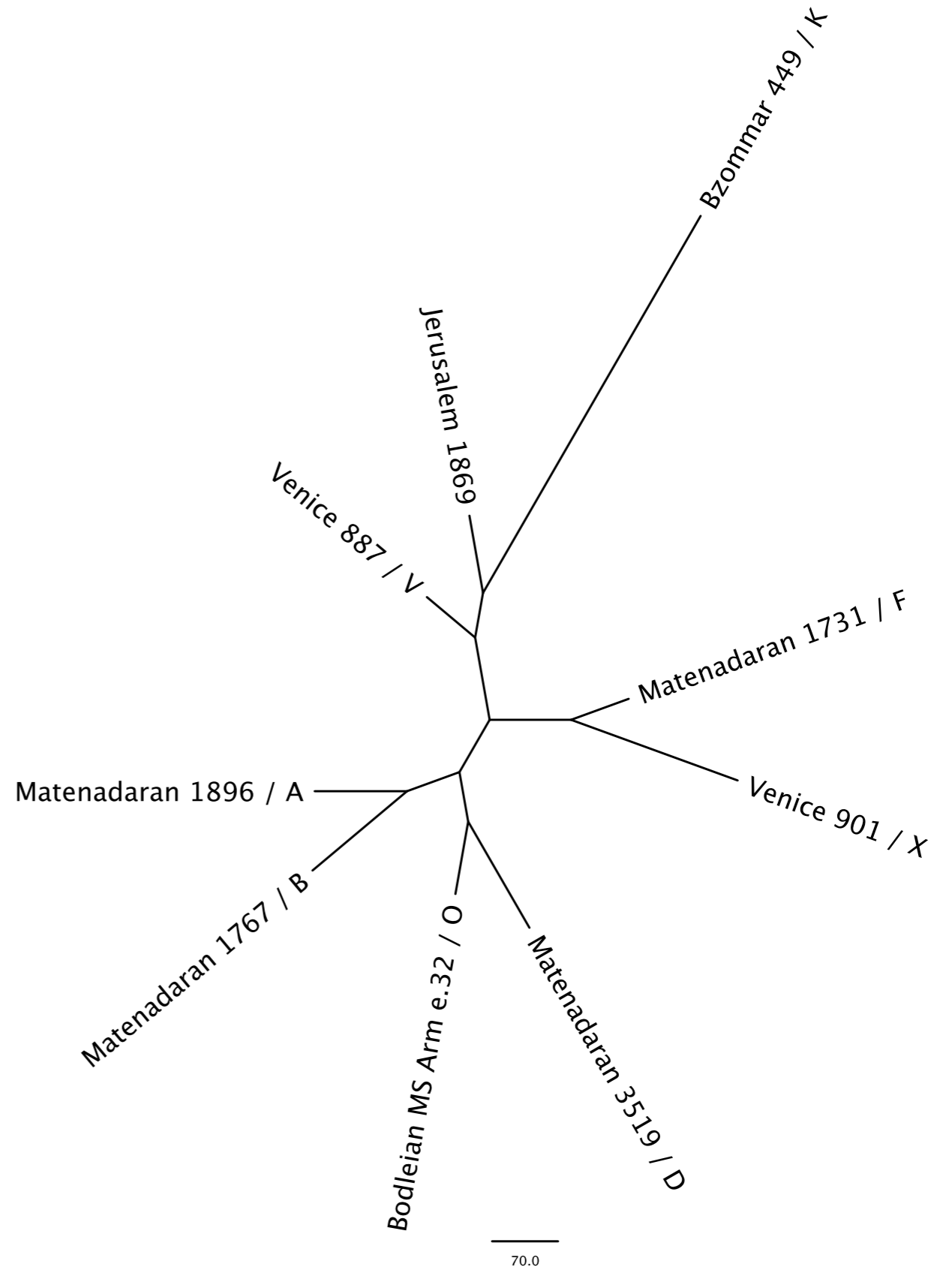
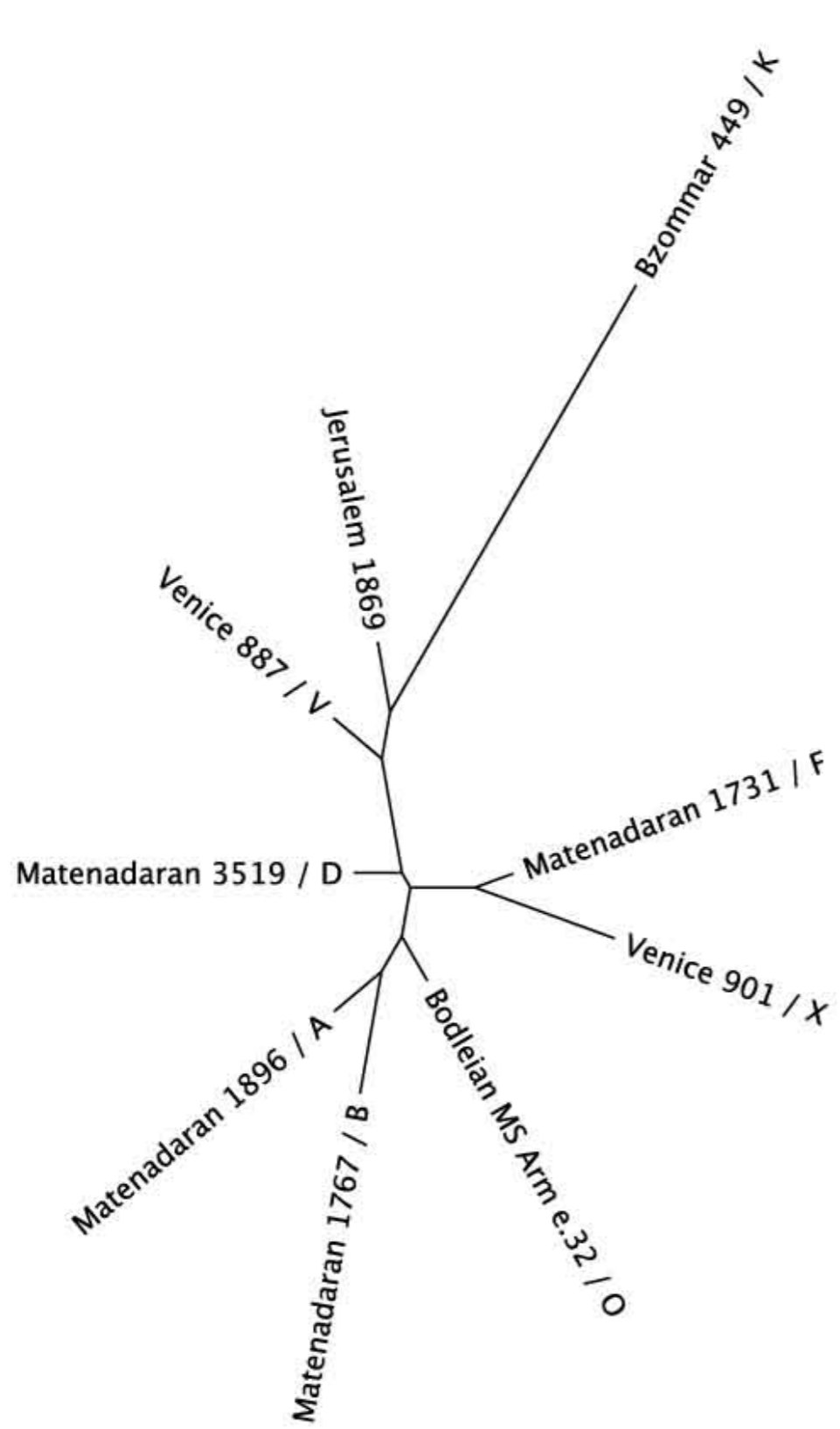


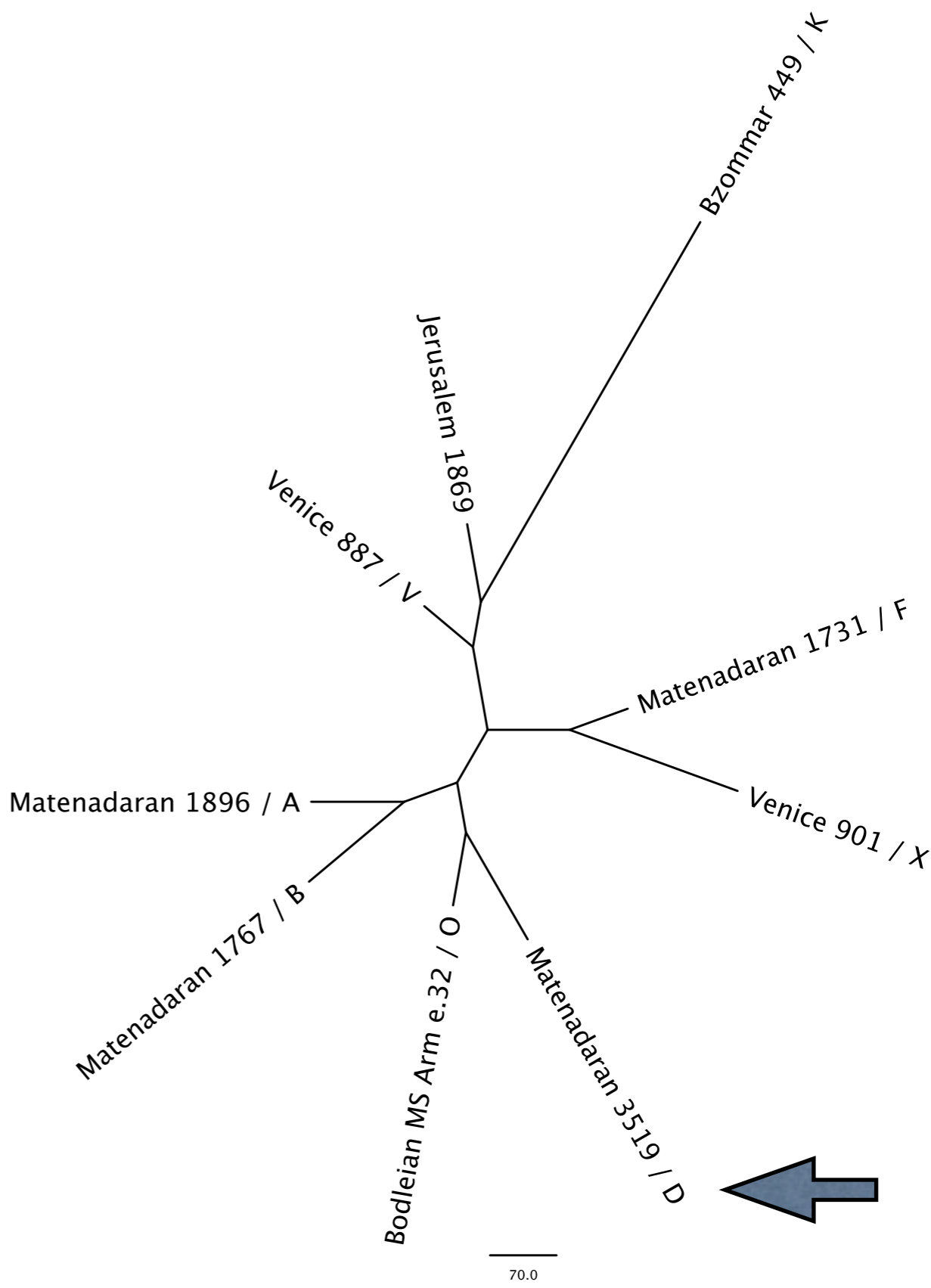
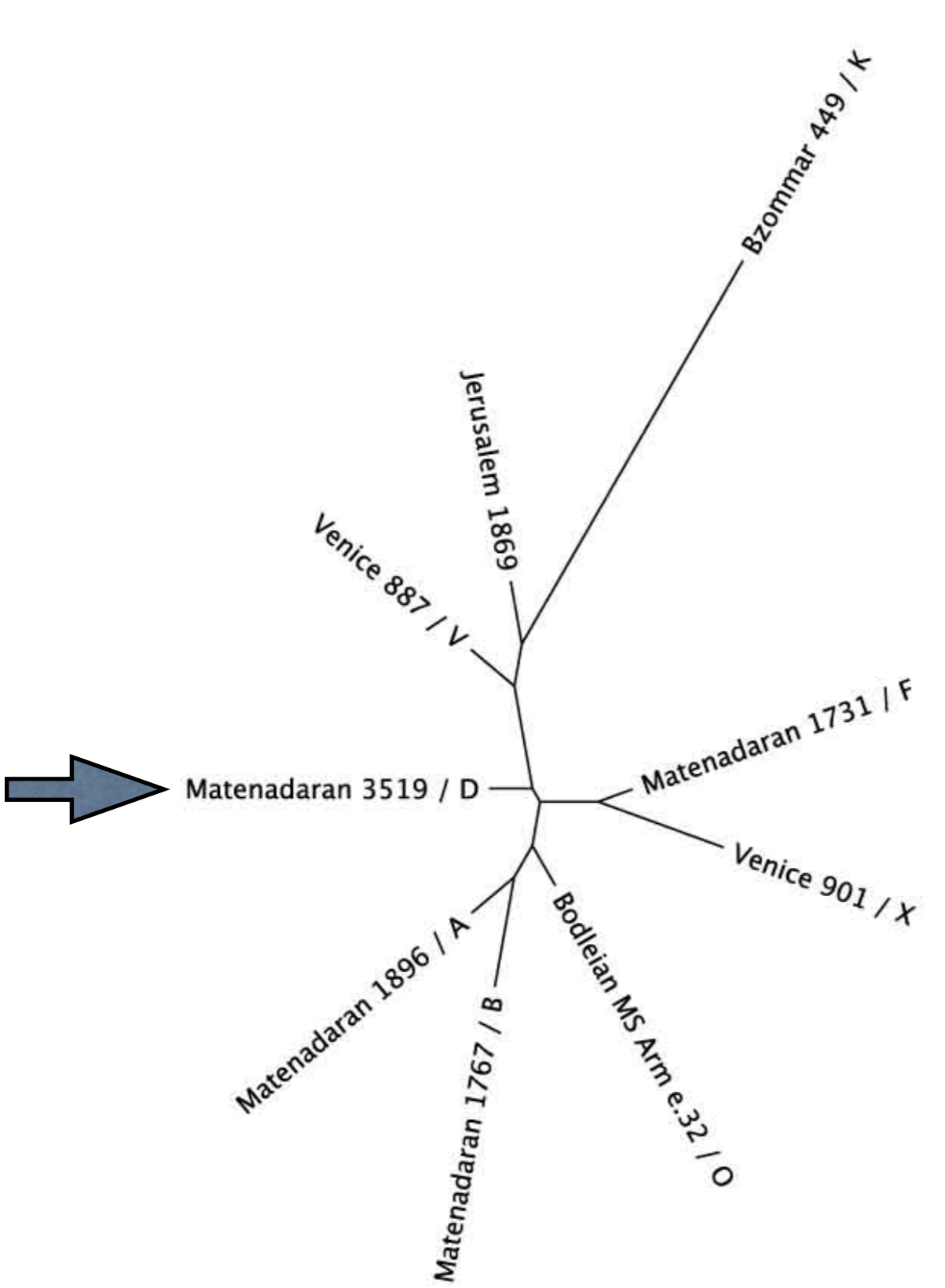
Possible trees from the text of Book 2



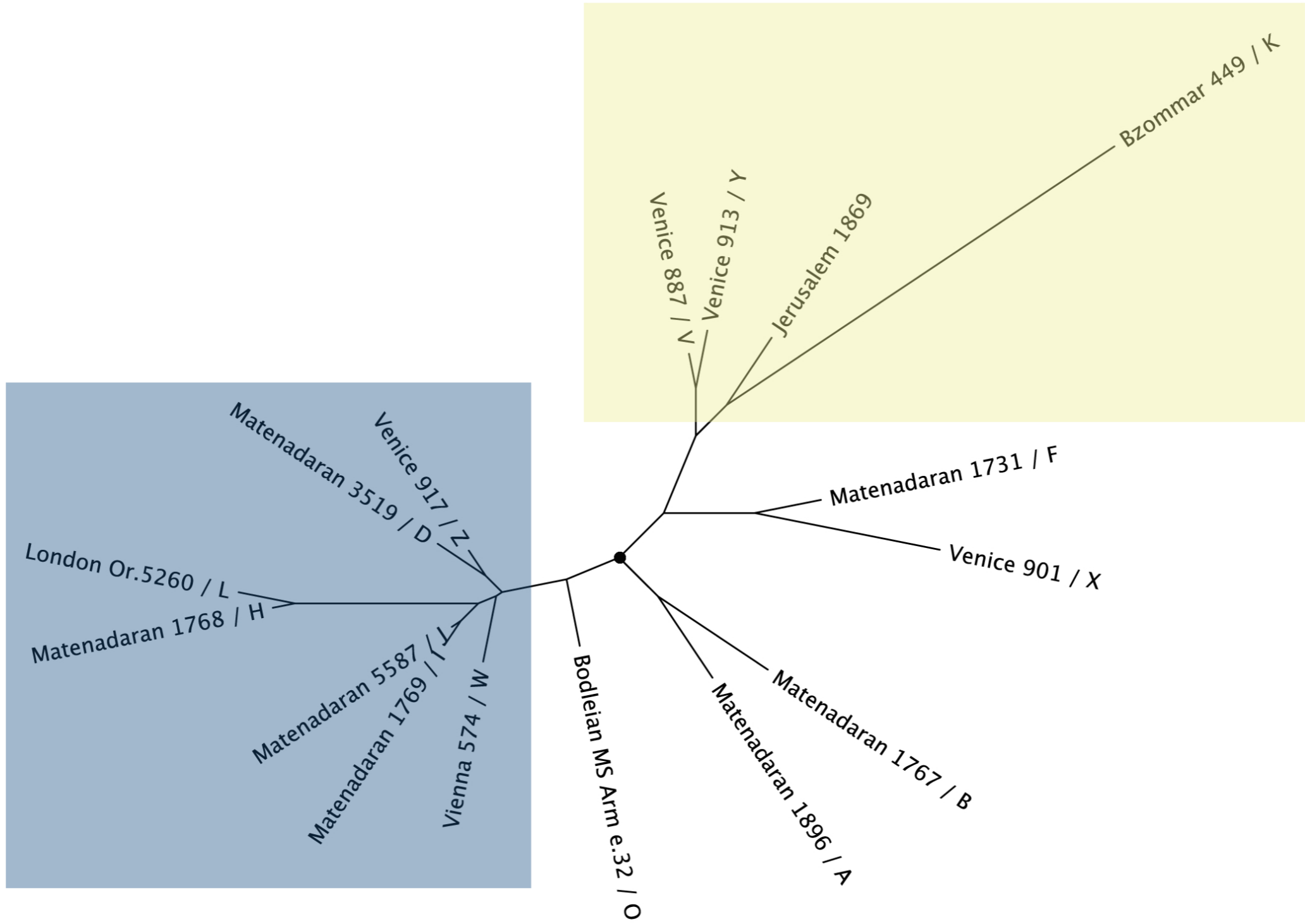
Possible trees from the text of the prophecy



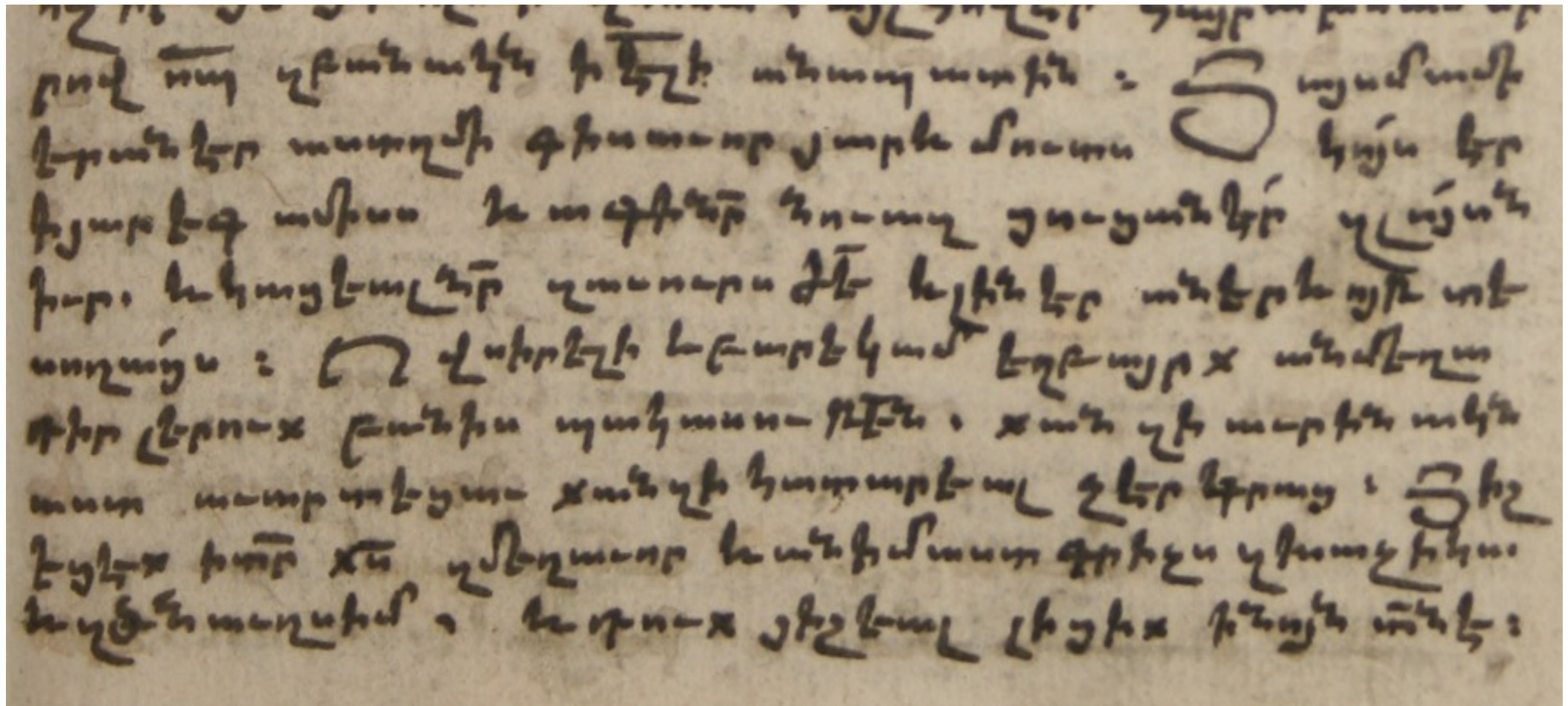




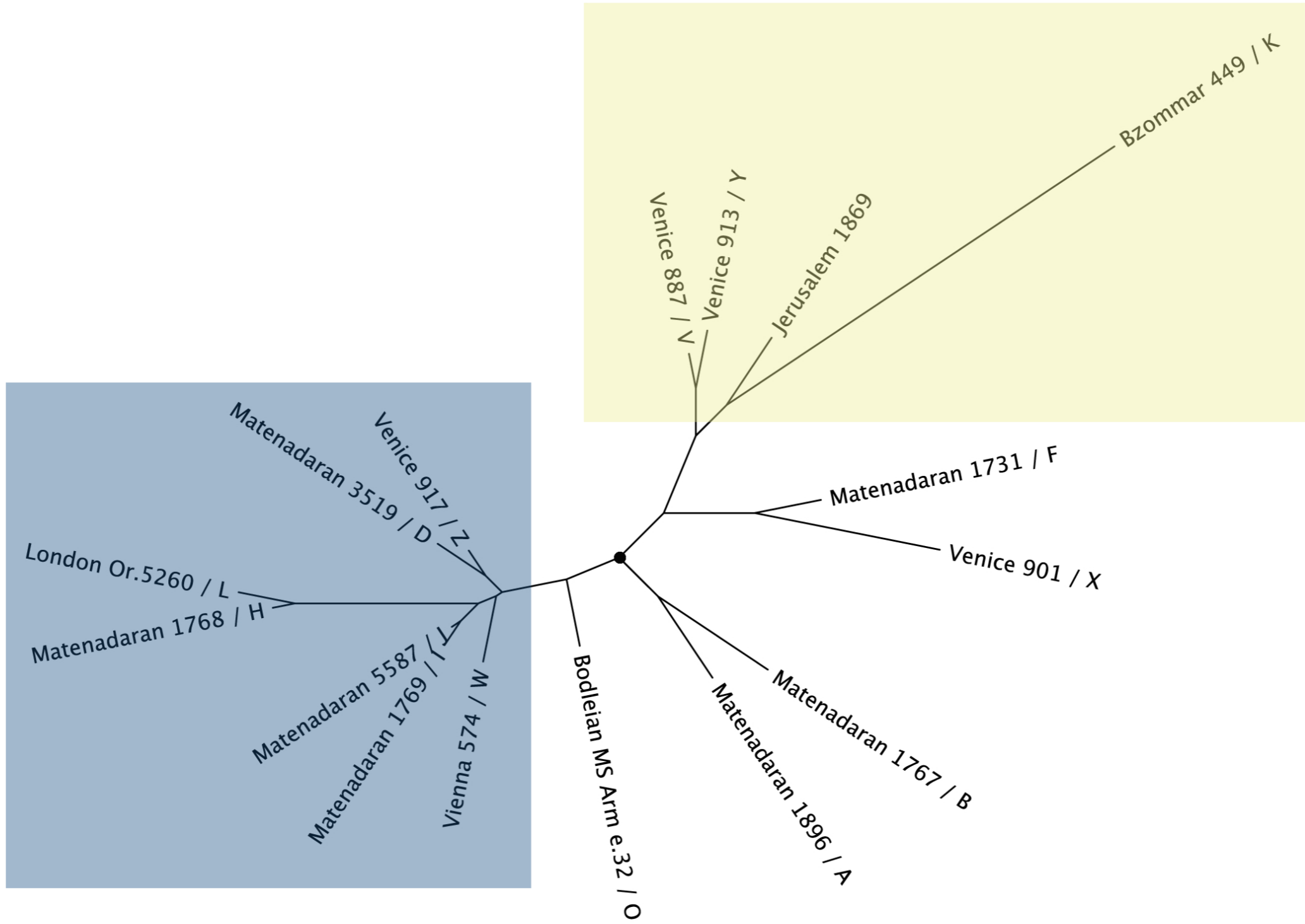


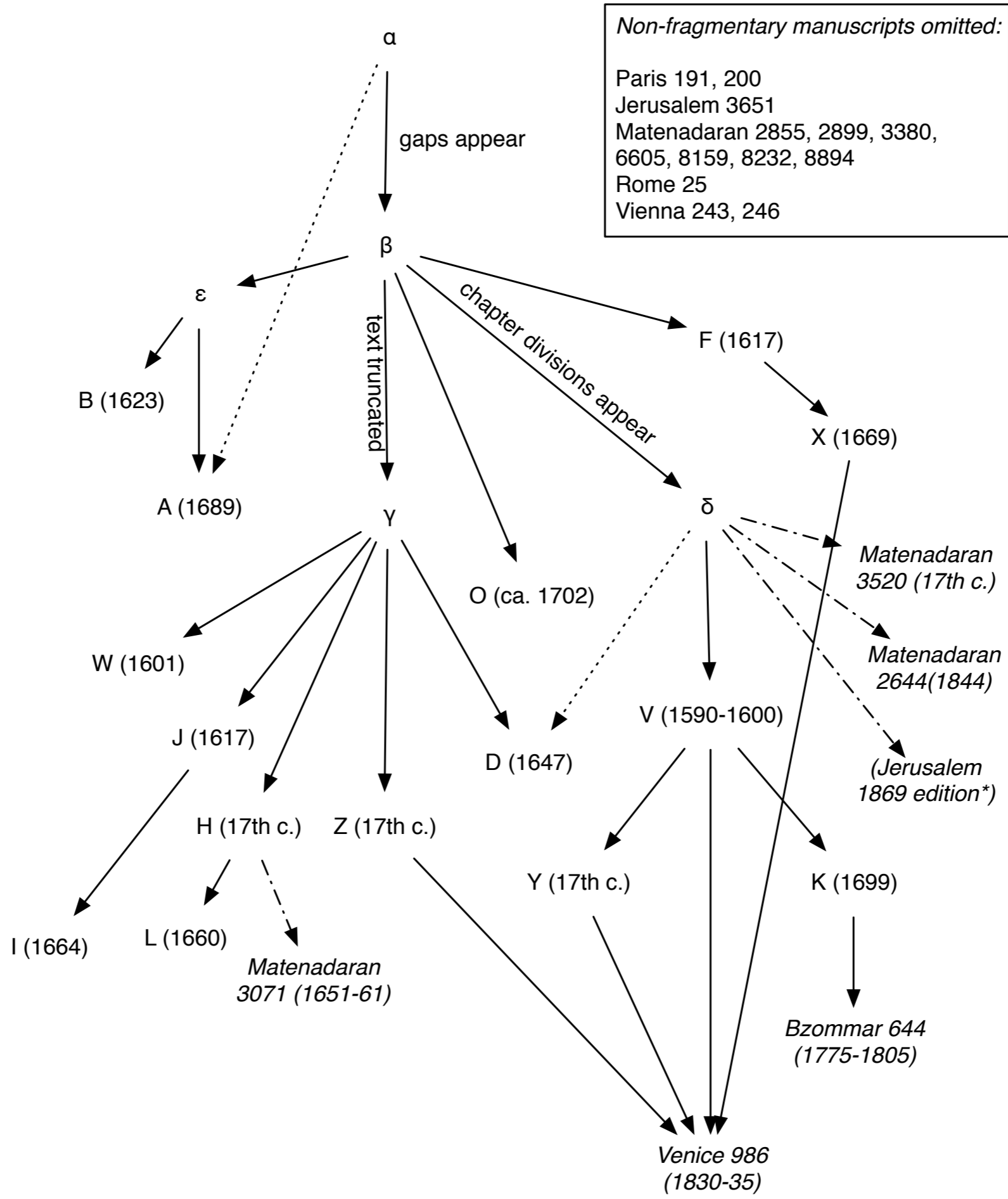


Handwritten text in Armenian script, likely a manuscript page. The text is dense and appears to be a historical or religious document. The script is cursive and somewhat difficult to read due to the handwriting style and the age of the document. The text is written in dark ink on aged, yellowish paper. The lines of text are closely spaced and run horizontally across the page. The overall appearance is that of an old, well-preserved manuscript.



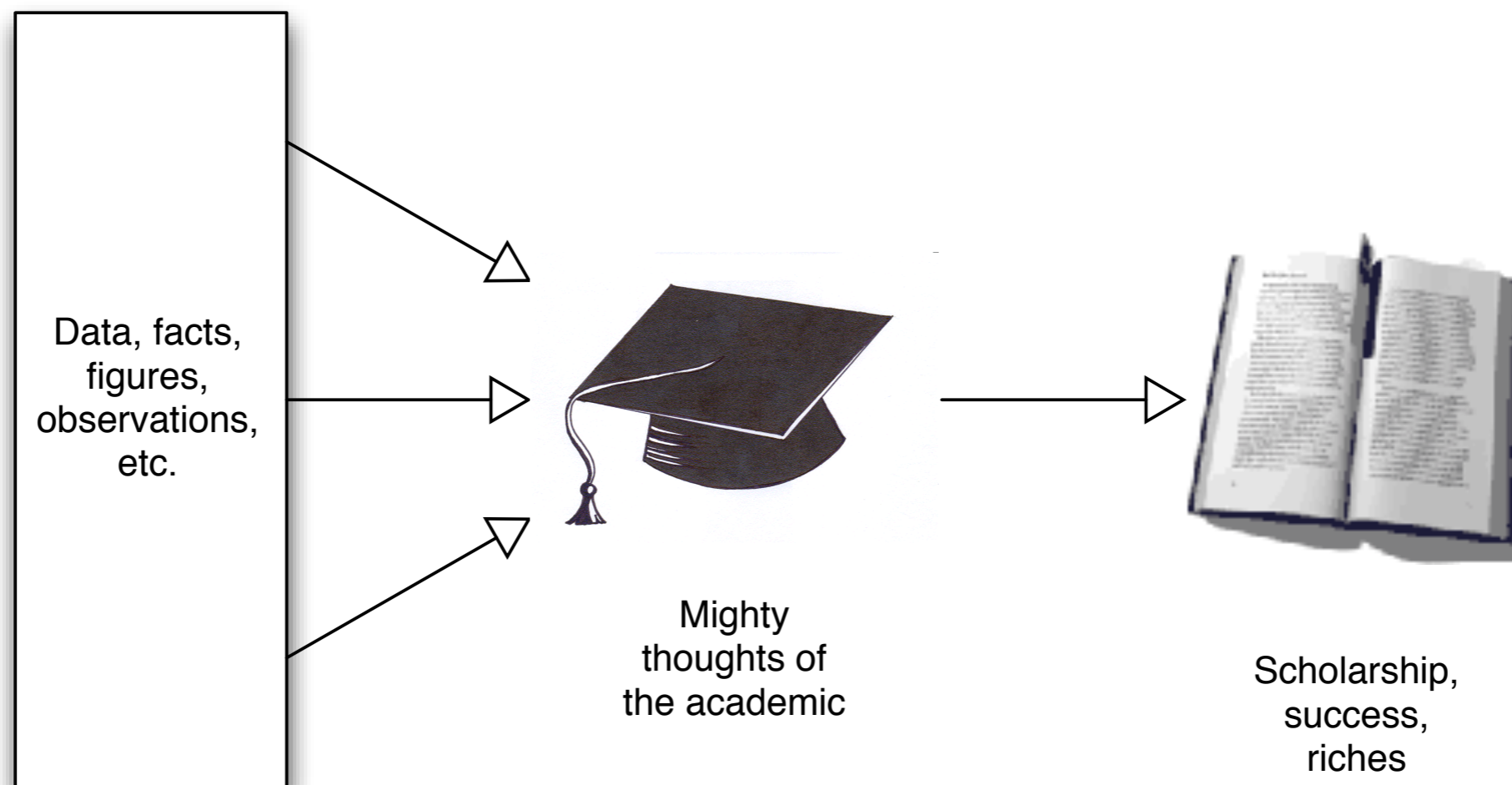
“In this year there appeared a hairy star [i.e. comet] moving toward the west in the month of Areg and its tail gave off a little of its light; and it was there for 15 days and then became invisible to viewers. O kind and friendly brothers, pardon the lack of words, because the exemplar ended here because 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.



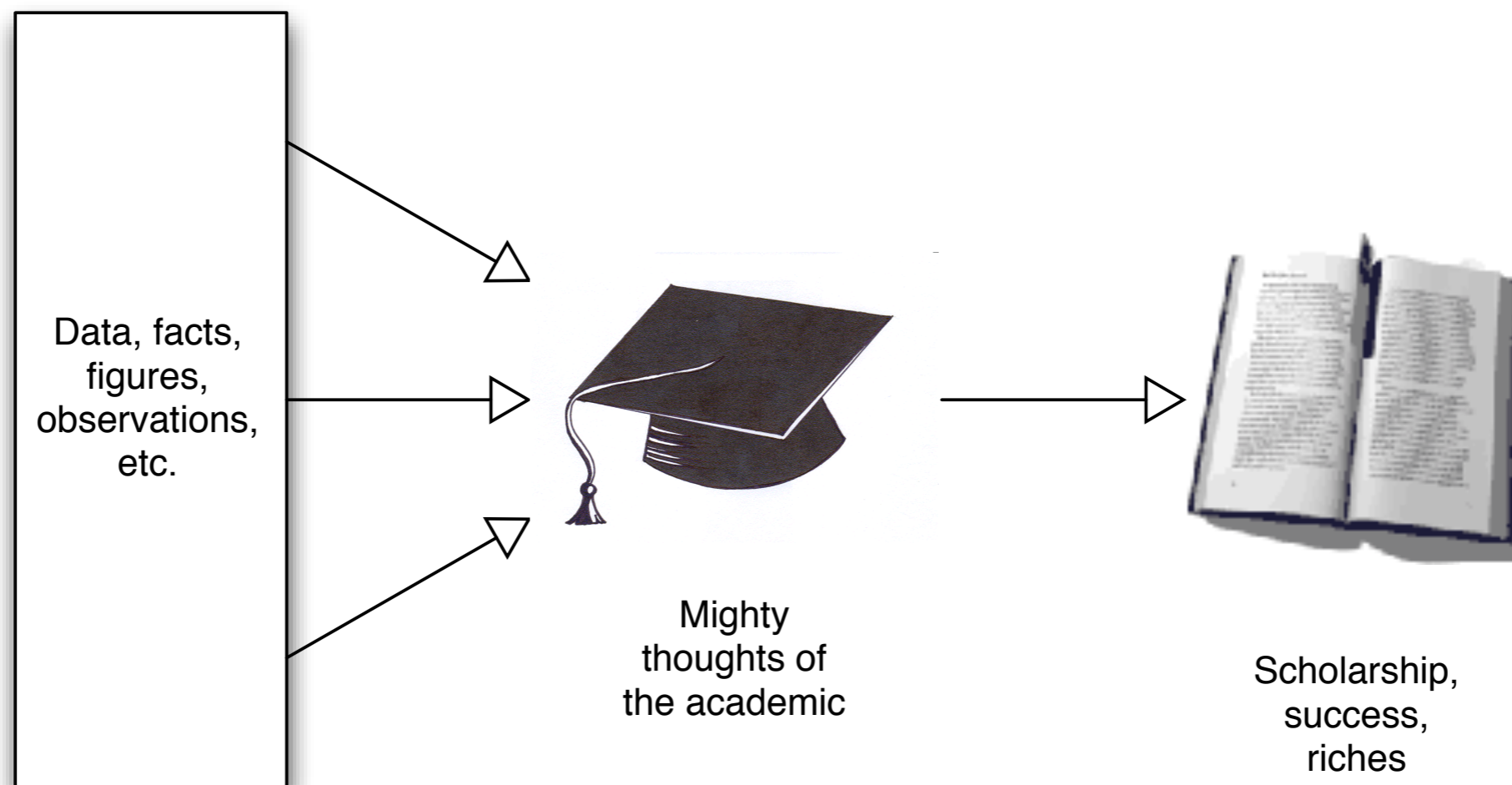


\*Based on Jerusalem mss. 1051, 1107

# How academics see their jobs

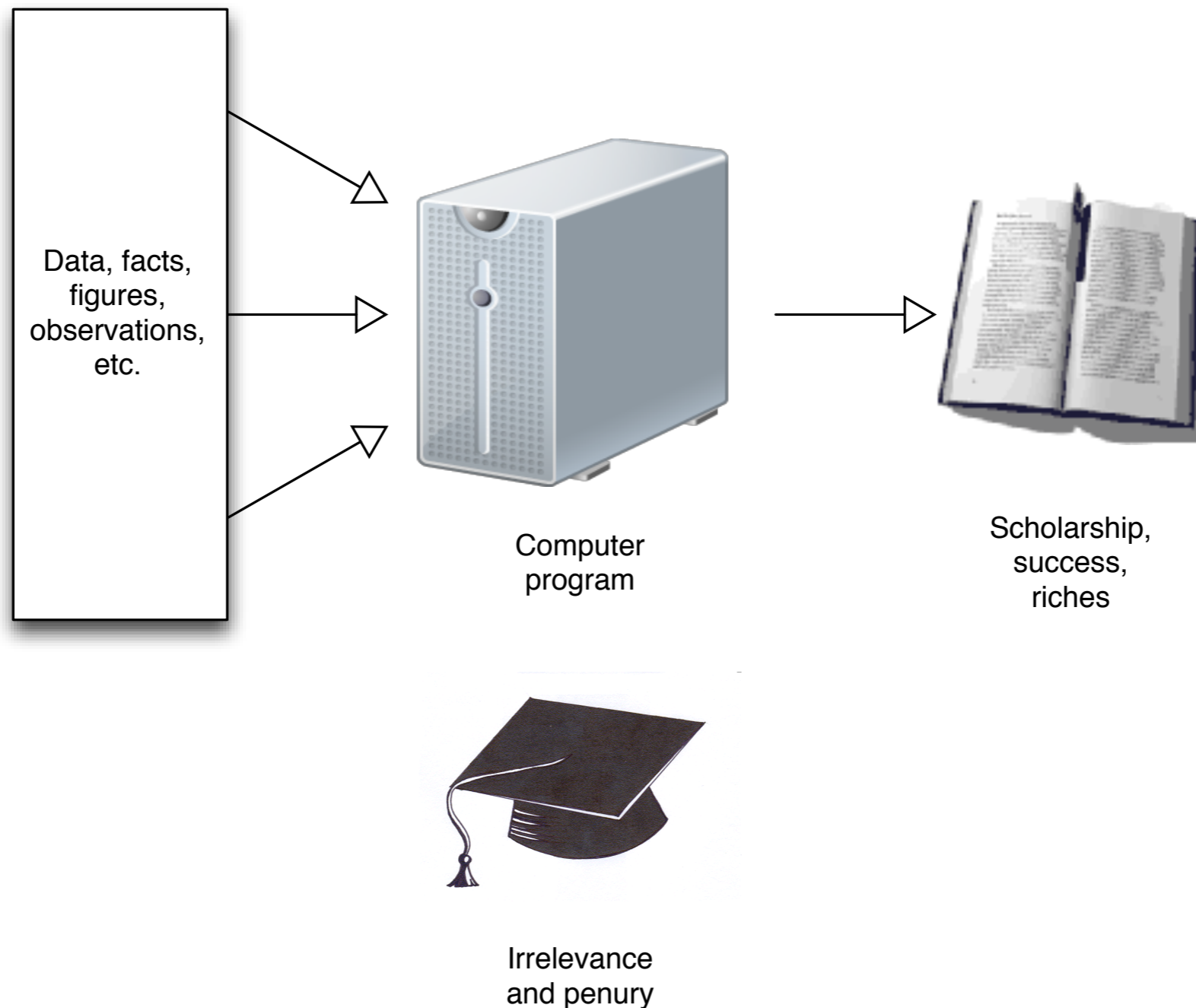


# How academics see their jobs



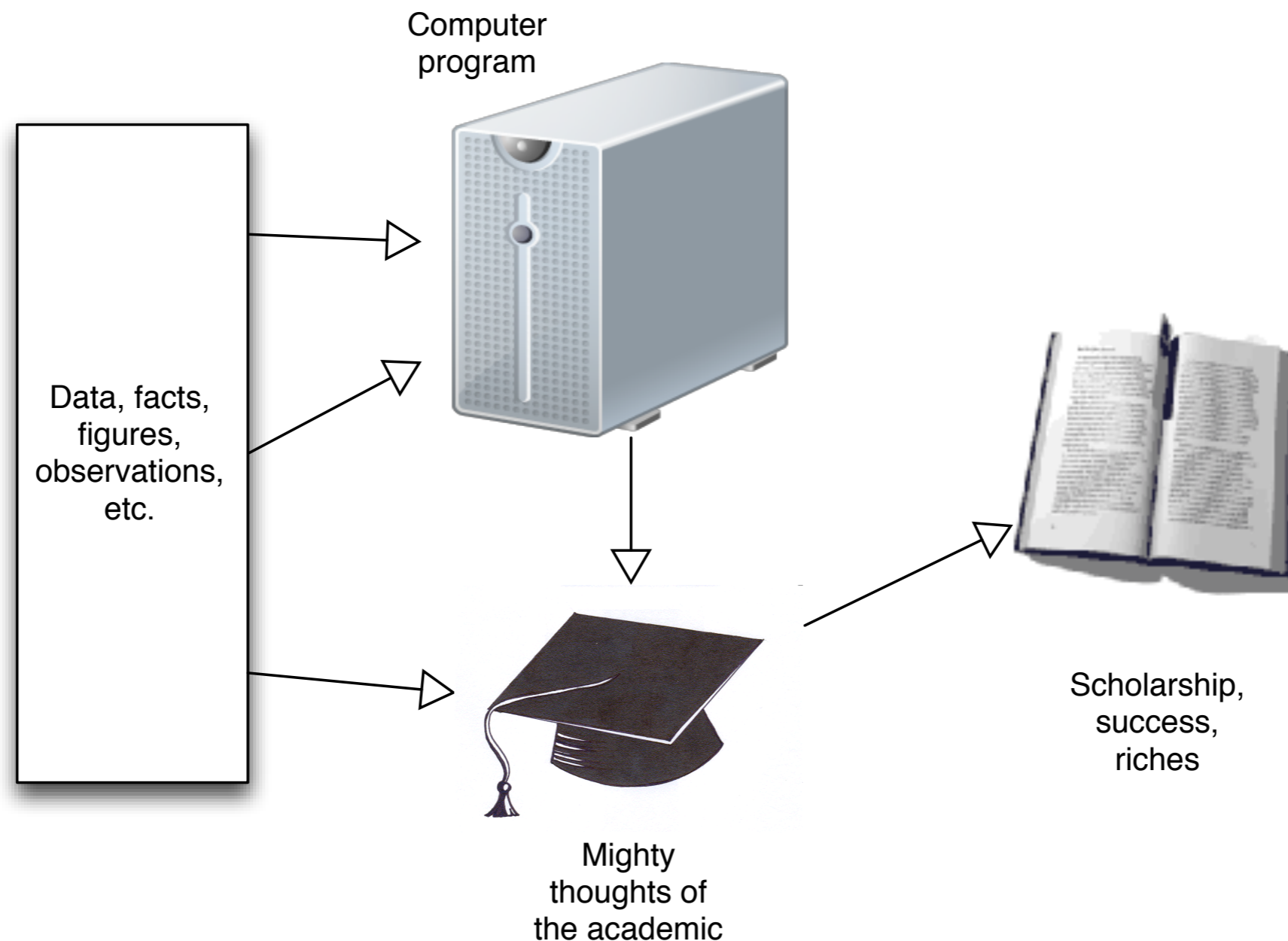
\*warning: gross oversimplification and/or wishful thinking

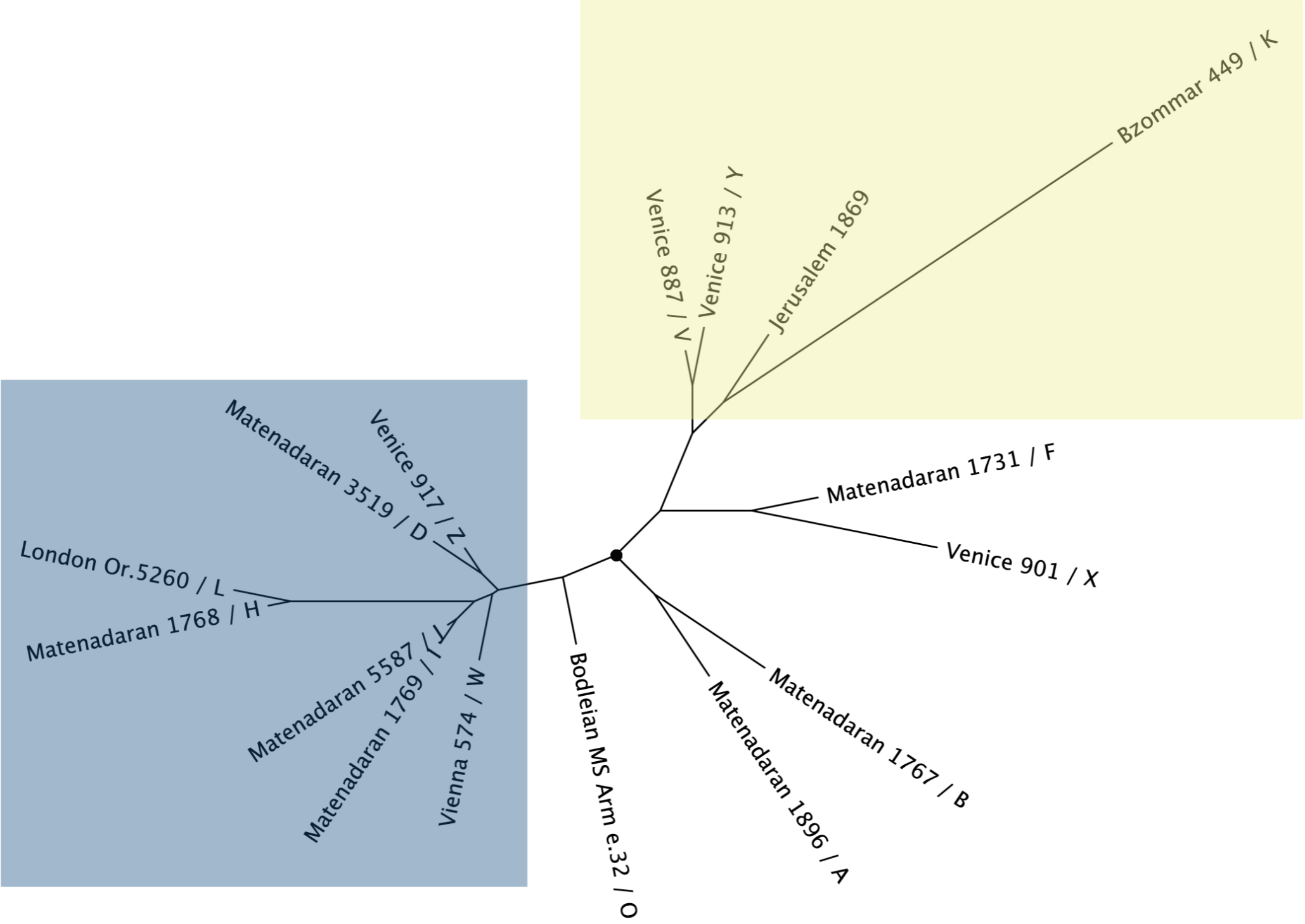
# How academics see computers in their jobs





# How academics should see computers

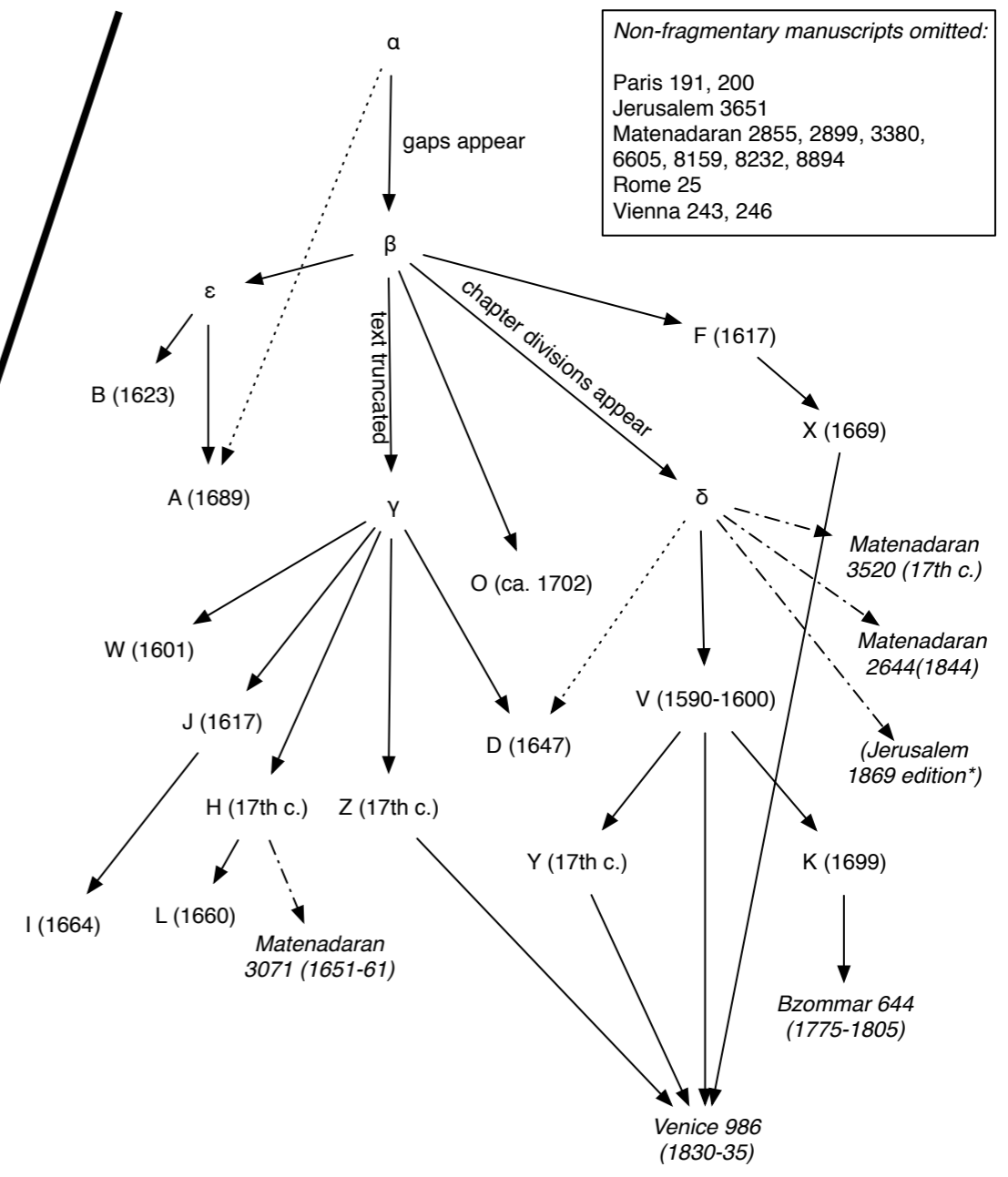


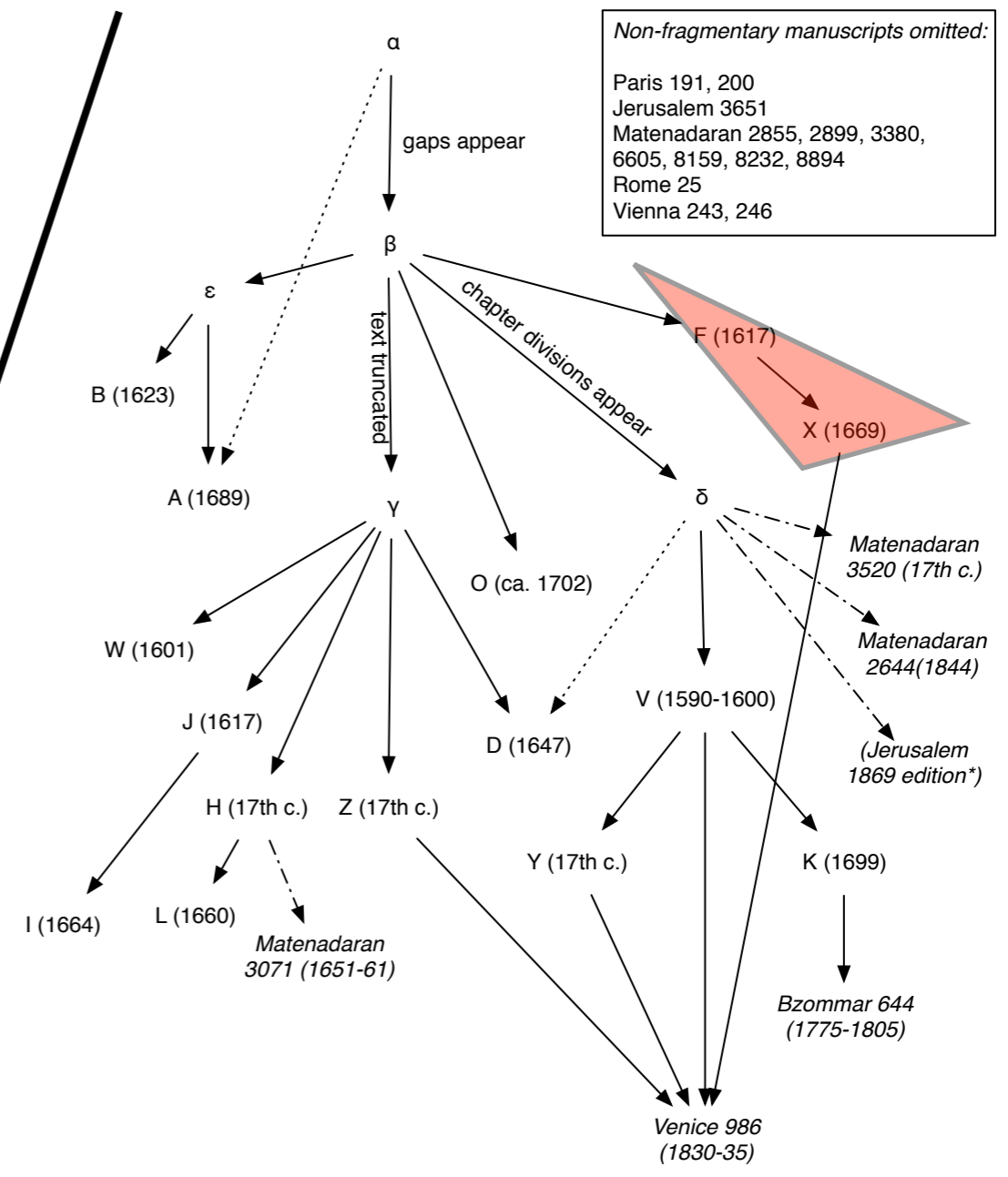
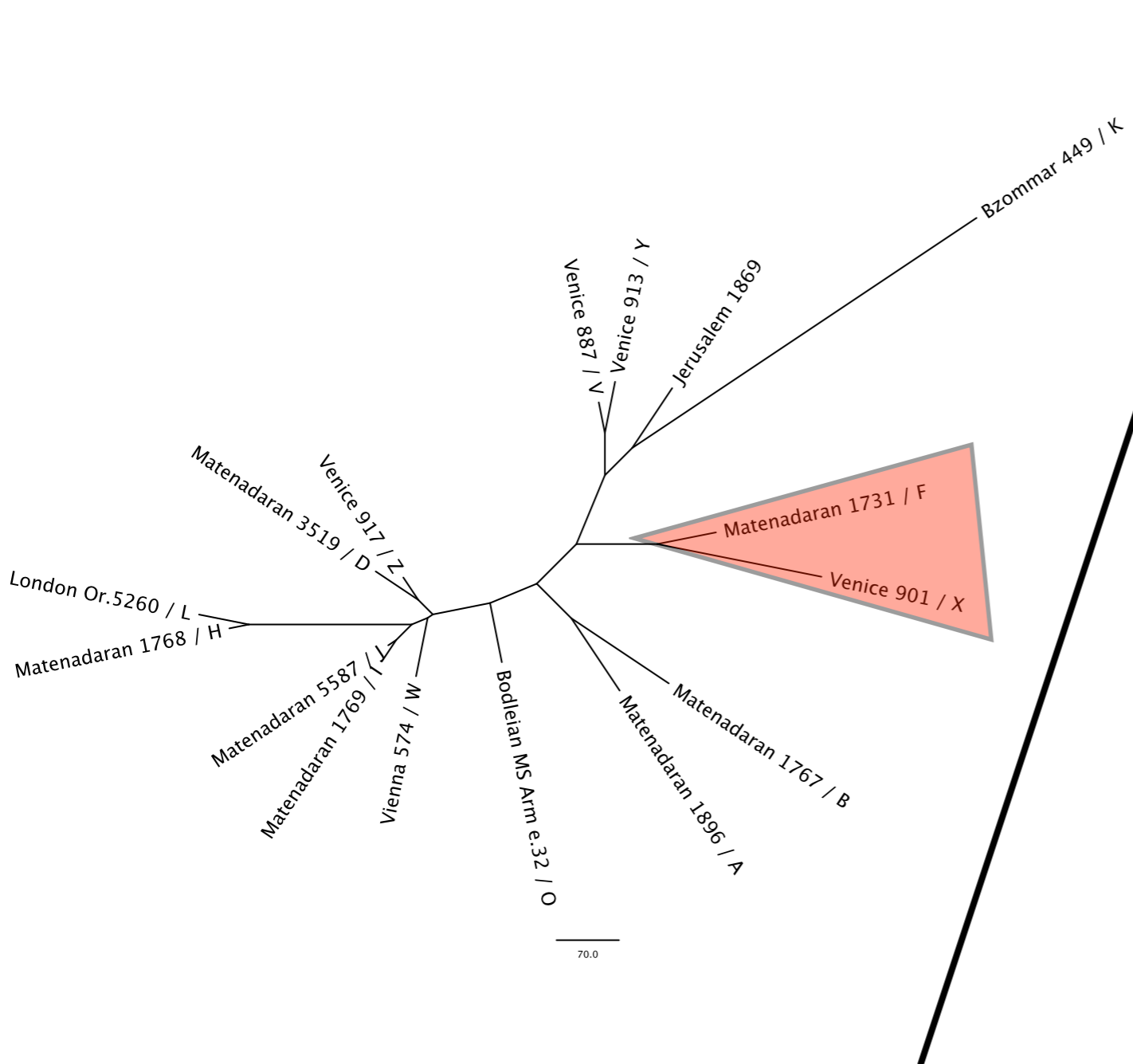


**Յ**այսմ ամի ժողով ձեռն արար  
Բարե արժեքն զամ Են յառեաց  
Բէն 'իսահմանս համոյ. և զամ  
կողմն յարեւելոյ Տիե. և ասիս  
զի Բազմապի Իսաղացեալ  
գայր հասաներ 'ի վրայ յայսր հիե  
աւանդաց. և ի վրայ Բանակ հար  
կաներ 'ի վրայ անուանի քնքին որ  
կոչի Տիբեճ. Ձերմ 'ի հաւաքա  
զարթին. որ էր հաղակաւոր քնք  
աւանդաց. և ասանի պիտի Եղև  
և ի վանս գ Էտրի զԲազմաքն ձեռ  
Բ. Եւ 'ի յեւանս Իսահմանս յայ  
գայն և անսիս զօրացն. և զիս  
այսքան և Եղև ի նոյն հրամայ  
եց Բարե կանգնել Բաբանս

ասա Բարե արժեքն. որ ասանս  
ժողովարար 'ի կողմնս Բարե  
պիտի. և Եղև Եւ զայս Բարե  
զնաց ժողովարար:

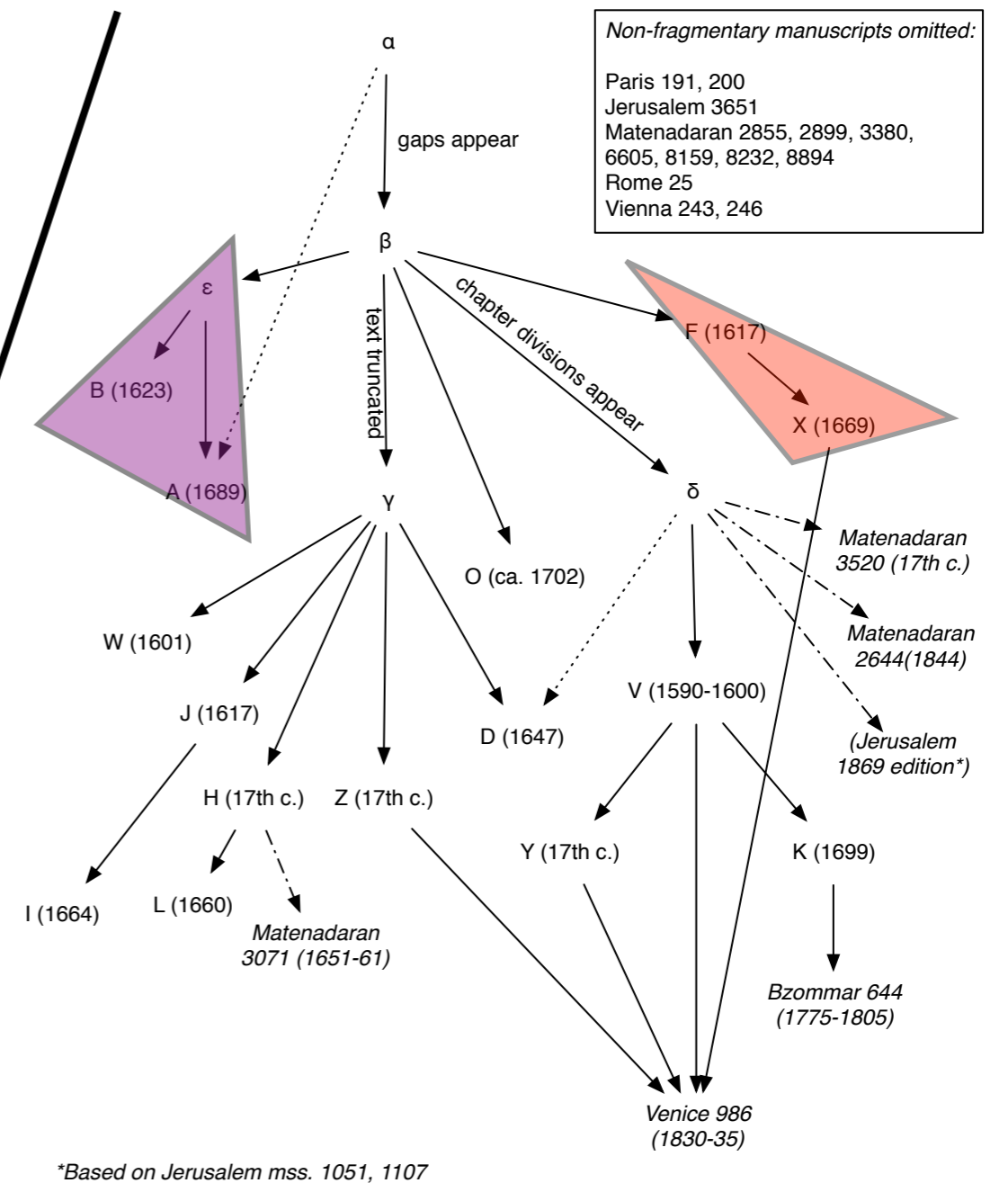
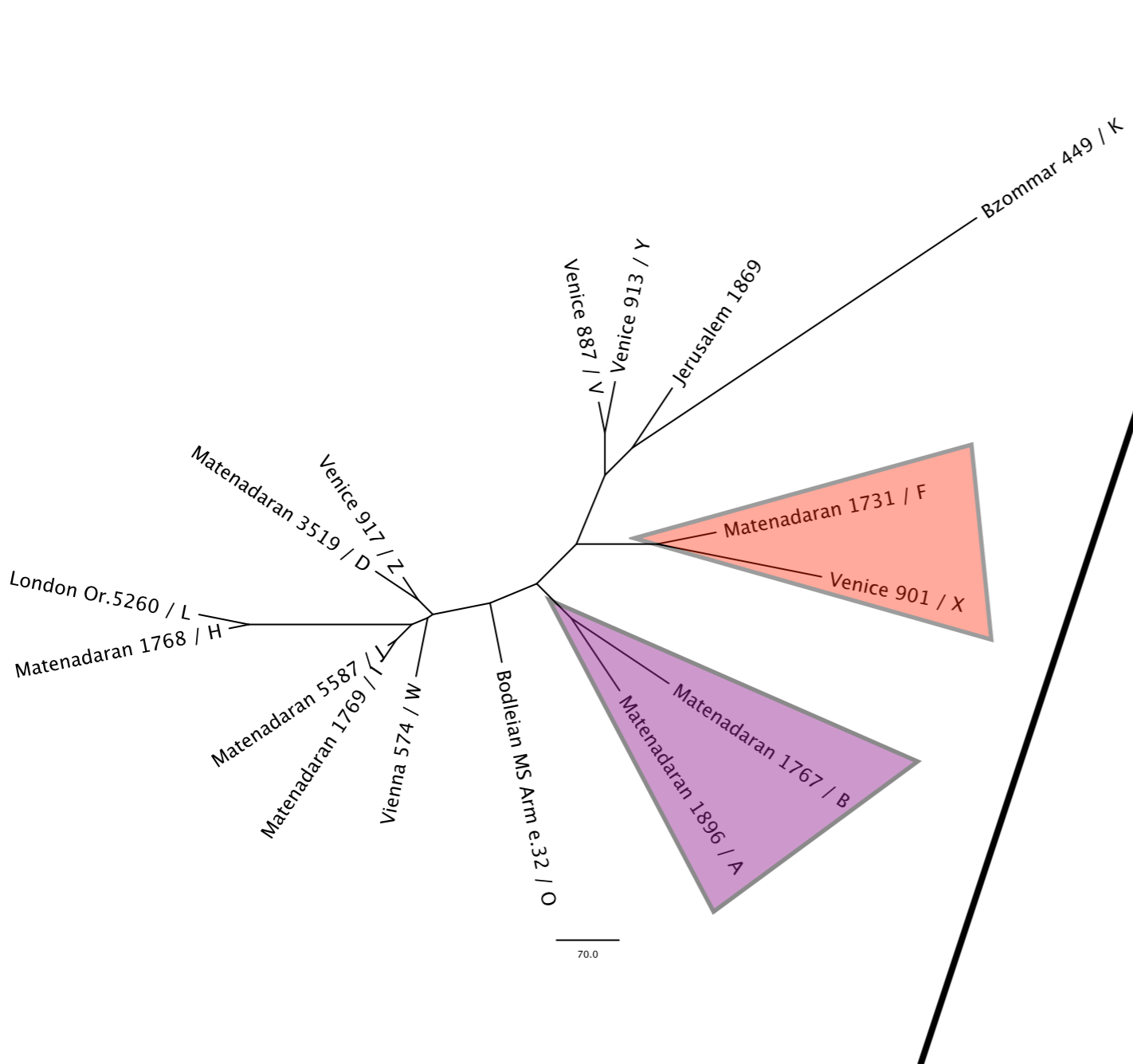
**Յ**այսմ ամի Եւ Եւ Եւ  
Զայ ամիրայ Տի կար ի ձ ան  
ուան. և կաներ գաղտնաւ  
տամ Բիւլի պարսից. և ԿԷ  
Եւ Եւ 'ի կողմնս Եւ Եւ Եւ  
Սիս. աս յառեաց Բազմ  
արն. և էր սն յապգի  
աբասլան առ կարան ին.  
և Բազմ զօրոյ հասա  
նէր 'ի սեփասար իս. և

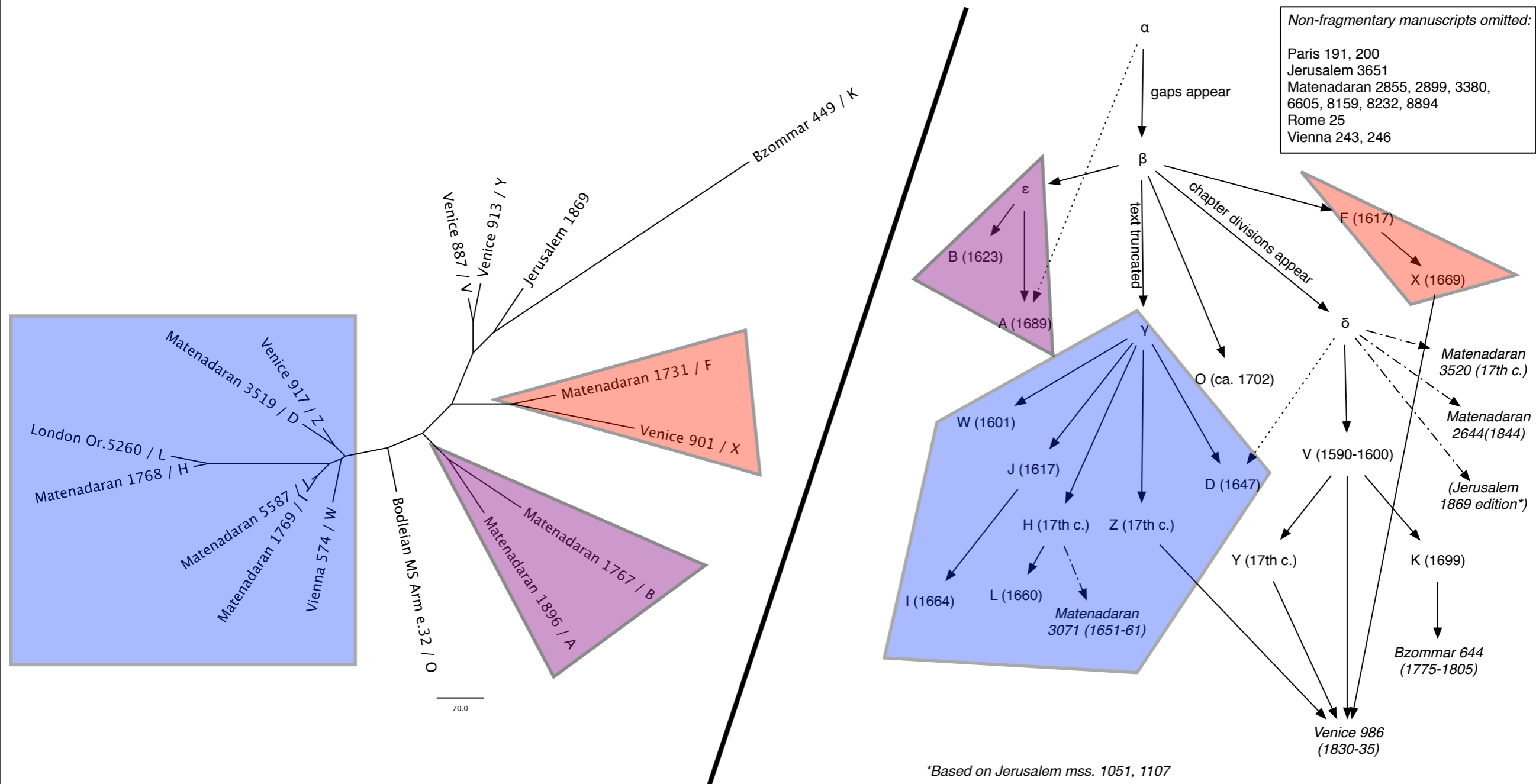




Non-fragmentary manuscripts omitted:  
 Paris 191, 200  
 Jerusalem 3651  
 Matenadaran 2855, 2899, 3380,  
 6605, 8159, 8232, 8894  
 Rome 25  
 Vienna 243, 246

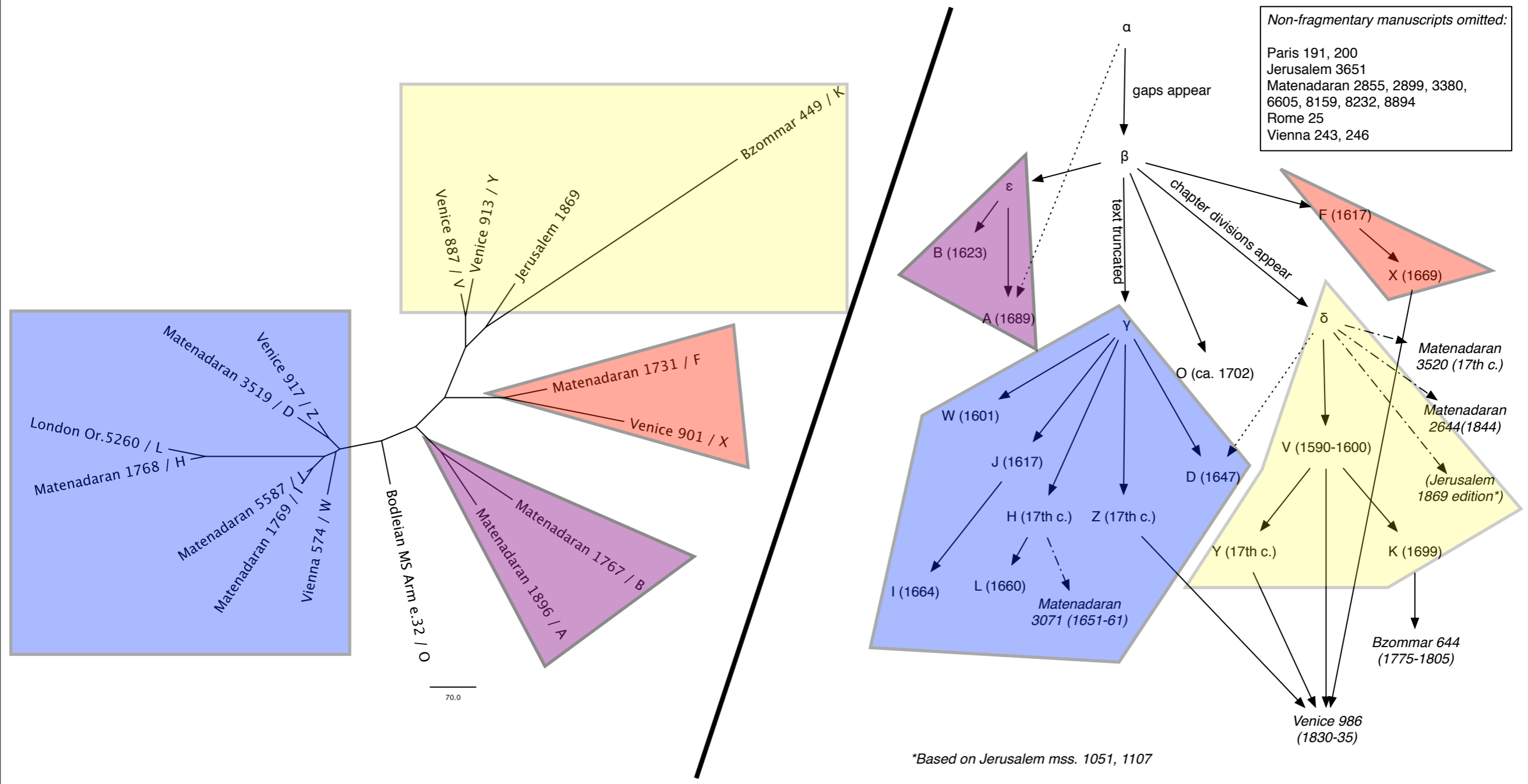
\*Based on Jerusalem mss. 1051, 1107





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





Thank you

*Obrigada*

Շնորհակալություն

*[shnorhakalutiun]*