

***Bacillus anthracis* evolutionary history: taking advantage of polytomies and human history to propose dating points**

Gilles Vergnaud gilles.vergnaud@u-psud.fr

Institute for Integrative Biology of the Cell (I2BC), CEA, CNRS, Univ. Paris-Sud, Université Paris-Saclay, 91198, Gif-sur-Yvette cedex, France

Abstract

The document deposited here is a copy of the slides presented in support of my talk during the “Biology of Anthrax” meeting held in Bari, Italy (September 3rd-6th 2019). The talk was a synthesis and global analysis of our recent investigations taking advantage of the existence of rare remarkable features in the phylogeny of *Bacillus anthracis*, polytomies, to propose tentative dating points.

Introduction

In the past few years, whole genome sequence data from hundreds of *B. anthracis* strains collected worldwide have been made public. Although multiple major gaps remain in terms of geographic coverage, this opened the way to relatively large scale phylogenetic analyses using whole genome Single Nucleotide Polymorphism analysis (wgSNP). Currently, a few thousands core-genome SNPs have been identified in the whole *B. anthracis* species (SNPs in non-repeated regions, called in all investigated strains) [1]. This low genetic diversity indicates that the strictly clonal pathogen emerged recently, as shown for other pathogens such as *Mycobacterium tuberculosis*. However, data analysis also illustrated the absence of molecular clock, which could have allowed the dating [1]. This feature is most likely due to the capacity of *B. anthracis* to sporulate and enter a quiescent state for decades. Therefore, the age of the *B. anthracis* species or even of the most recent common ancestor (MRCA) of currently identified lineages is essentially unknown.

In 2014, we described for the first time [2] a major polytomy, subsequently called A.Br.011/009 in keeping with its canSNP [3] location within *B. anthracis*. A polytomy is a star-like pattern in the phylogeny, and six branches, all present in France have been identified so far in A.Br.011/009. One of these branches, called sublineage 2 contained a second more recent three-branches polytomy with two very long branches, constituted by strains isolated in West Africa and North America (WANA) [4, 5]. This remarkable topology prompted us to date the emergence of the WANA polytomy to the 17th century when France was simultaneously colonizing parts of West Africa and North America [4]. In 2018, we updated the analyses with newly available data, showing that some

among the six A.Br.011/009 polytomy branches are also present in Italy, with a pattern suggesting a very early split between French and Italian lineages [5].

In 2016, a third polytomy was described, A.Br.008/011 [1]. The investigation of additional strains assigned to this polytomy, including strains recovered from the Siberian permafrost, allowed us to tentatively link this expansion to the Mongol conquests in the 13th century and thus to propose an additional dating point [5, 6].

The “Régiment de Carignan-Salières” hypothesis

The bottom line of the talk in Bari was that the recent hypothesis linking the emergence of the A.Br.008/011 polytomy to the Mongol conquests is implying that medieval armies were efficient spreaders of *B. anthracis*. The Mongolian armies were driven by animals including horses, so horses or associated equipments would be prime candidates for this role.

In the talk, this hypothesis of *B. anthracis* spreading by military operations was used to further interpret the A.Br.011/009 and WANA polytomies. The intricate Franco-Italian pattern observed in the birth of the A.Br.011/009 polytomy [5] would be the result of the Franco-Italian wars between the 13th and 16th century. The WANA polytomy including the export of *B. anthracis* to North-America in the 17th century CE would be precisely linked to the sending of the Infantry Regiment “Carignan-Salières” to Quebec in 1665 [7-9]. This expedition with more than 1000 soldiers was the most important French military expedition towards Nouvelle-France in the whole 17th century. Interestingly, the Régiment de Carignan-Salières was based in Marsal, North-East France (Lorraine) where all extant French Sublineage 2 strains have been collected [2]. The infantry regiment crossed France from Marsal to La Rochelle in January-February 1665 and left La Rochelle in April and May 1665. As part of the same global operation aimed at strengthening the French colony, 20 horses were transported from La Rochelle in May 1665, 12 of which survived the trip [10, 11]. Even an infantry regiment would include horses for officers and as draft animals so it is tempting to speculate that the 20 horses were selected among the regiment horses. The twelve surviving horses were the first among approximately 80 sent from France in years 1665-1671, which subsequently led to the Canadian horse breed [11].

Conclusion

I concluded the talk by recalling that, although tempting by some aspects, this scenario should be considered as a highly hypothetical working hypothesis. Firstly, whole genome sequence data from many more *B. anthracis* strains will need to be made accessible to challenge the geographic specificity of some lineages. Many countries, including key European countries such as Spain and Portugal,

are not yet represented in public archives. Secondly, the proposed model is predicting that the regiment proved able to carry very efficiently infectious *B. anthracis* from North-East France to Canada after months of travelling across France and the Atlantic Ocean.

One possibility would be that the horses were healthy carriers of an infection, however the existence of such an healthy carrier state is currently not scientifically established. Interestingly, some of the rare circumstantial evidence supporting this possibility was discussed by Robert Gainer in his talk during the Bari conference [12, 13].

A copy of the thirty-nine slides of the talk is appended.

References

1. Sahl, J.W., et al., *A Bacillus anthracis Genome Sequence from the Sverdlovsk 1979 Autopsy Specimens*. MBio, 2016. **7**(5).
2. Girault, G., et al., *High-throughput sequencing of Bacillus anthracis in France: investigating genome diversity and population structure using whole-genome SNP discovery*. BMC Genomics, 2014. **15**(1): p. 288.
3. Van Ert, M.N., et al., *Global genetic population structure of Bacillus anthracis*. PLoS One, 2007. **2**(5): p. e461.
4. Vergnaud, G., et al., *Comparison of French and Worldwide Bacillus anthracis Strains Favors a Recent, Post-Columbian Origin of the Predominant North-American Clade*. PLoS One, 2016. **11**(2): p. e0146216.
5. Timofeev, V., et al., *Insights from Bacillus anthracis strains isolated from permafrost in the tundra zone of Russia*. bioRxiv, 2018.
6. Timofeev, V., et al., *Insights from Bacillus anthracis strains isolated from permafrost in the tundra zone of Russia*. PLoS One, 2019. **14**(5): p. e0209140.
7. Quillivic, B. and J. Nicol. *Le régiment de Carignan*. [cited 2019 20/09/2019]; Available from: <http://www.migrations.fr/>.
8. Fournier, M. and M. Langlois, *Le régiment de Carignan-Salières. Les premières troupes françaises de la Nouvelle-France, 1665-1668*. 2014, Montreal, Canada: Éditions Québec Histoire. 127.
9. Verney, J., *The Good Regiment: The Carignan-Salieres Regiment in Canada, 1665-1668*. 1992.
10. Campeau, C.V. *Navires venus en Nouvelle-France, gens de mer et passagers, des origines à 1699*. 2002 [25/09/2019]; Available from: <https://www.naviresnouvellefrance.net/>.
11. Seguin, R., *Le cheval et ses implications historiques dans l'Amérique française*. Revue d'histoire de l'Amérique française, 1951. **5**(2): p. 227–251.
12. Gainer, R.S., *Spore concentration and modified host resistance as cause of anthrax outbreaks: A practitioner's perspective*. Can Vet J, 2018. **59**(2): p. 185-187.
13. Gainer, R.S., *Commentaries on anthrax in ruminants*. Can Vet J, 1983. **24**(5): p. 163-4.



***Bacillus anthracis* evolutionary history: taking advantage of polytomies and human history to propose dating points**

Gilles VERGNAUD



Bacillus anthracis evolution

- Features which make it easier to follow:
 - Strictly clonal evolution (no genetic exchanges with other bacteria, no intra-species recombination)
- Features which make it more difficult:
 - Erratic evolution rate due to its dormant state (spores)
- Will we be able to pinpoint dates and geographic origins on the phylogeny?

Overview of the presentation

- Polytomies in *Bacillus anthracis* population structure
- The « A.Br.008/011 » polytomy: first tentative dating point
- The « A.Br.011/009 » polytomy: second tentative dating point
- The West Africa-North America (WANA) long branches
- The « Régiment de Carignan-Salières » hypothesis: third dating point
- Conclusions and perspectives

Polytomies in *Bacillus anthracis* population structure

wgSNP analysis of available WGS data

- Currently available whole genome sequence data:
 - 1300 assemblies and read archives
 - 700 non-redundant
- Whole genome SNP calling by mapping using Ames Ancestor as reference genome
 - 7277 core genome SNPs

Global view

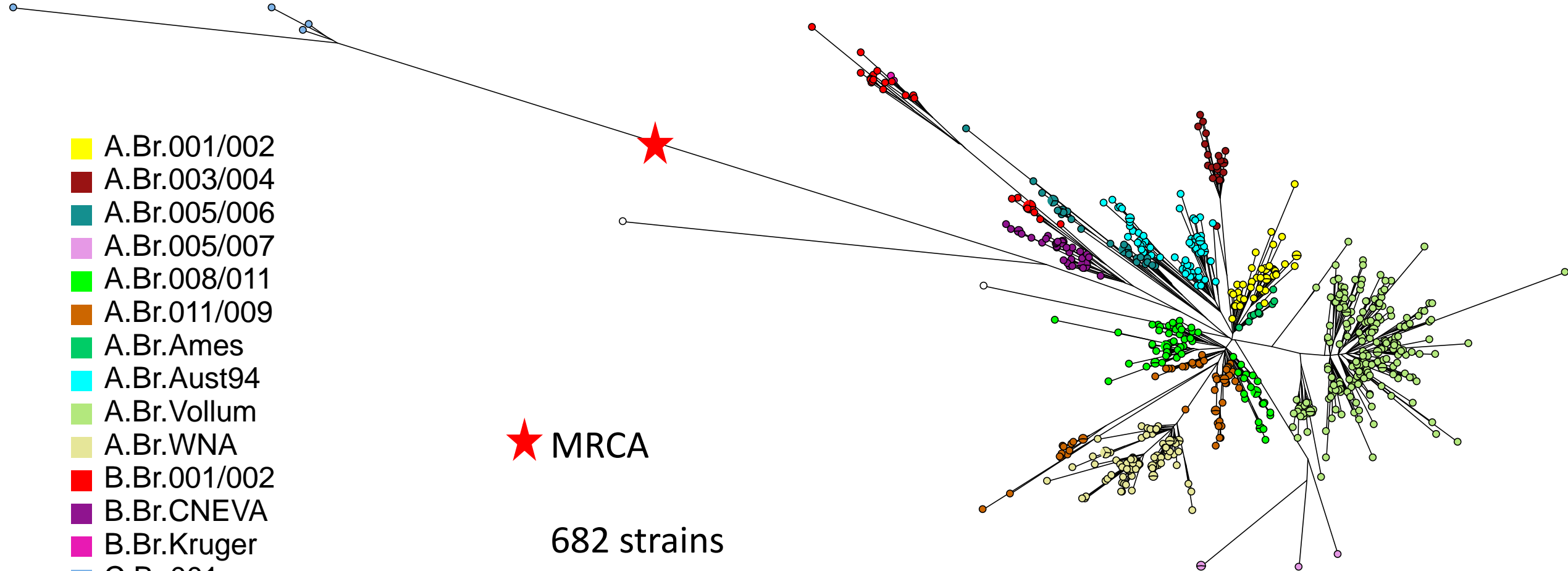
- A.Br.001/002
- A.Br.003/004
- A.Br.005/006
- A.Br.005/007
- A.Br.008/011
- A.Br.011/009
- A.Br.Ames
- A.Br.Aust94
- A.Br.Vollum
- A.Br.WNA
- B.Br.001/002
- B.Br.CNEVA
- B.Br.Kruger
- C.Br.001

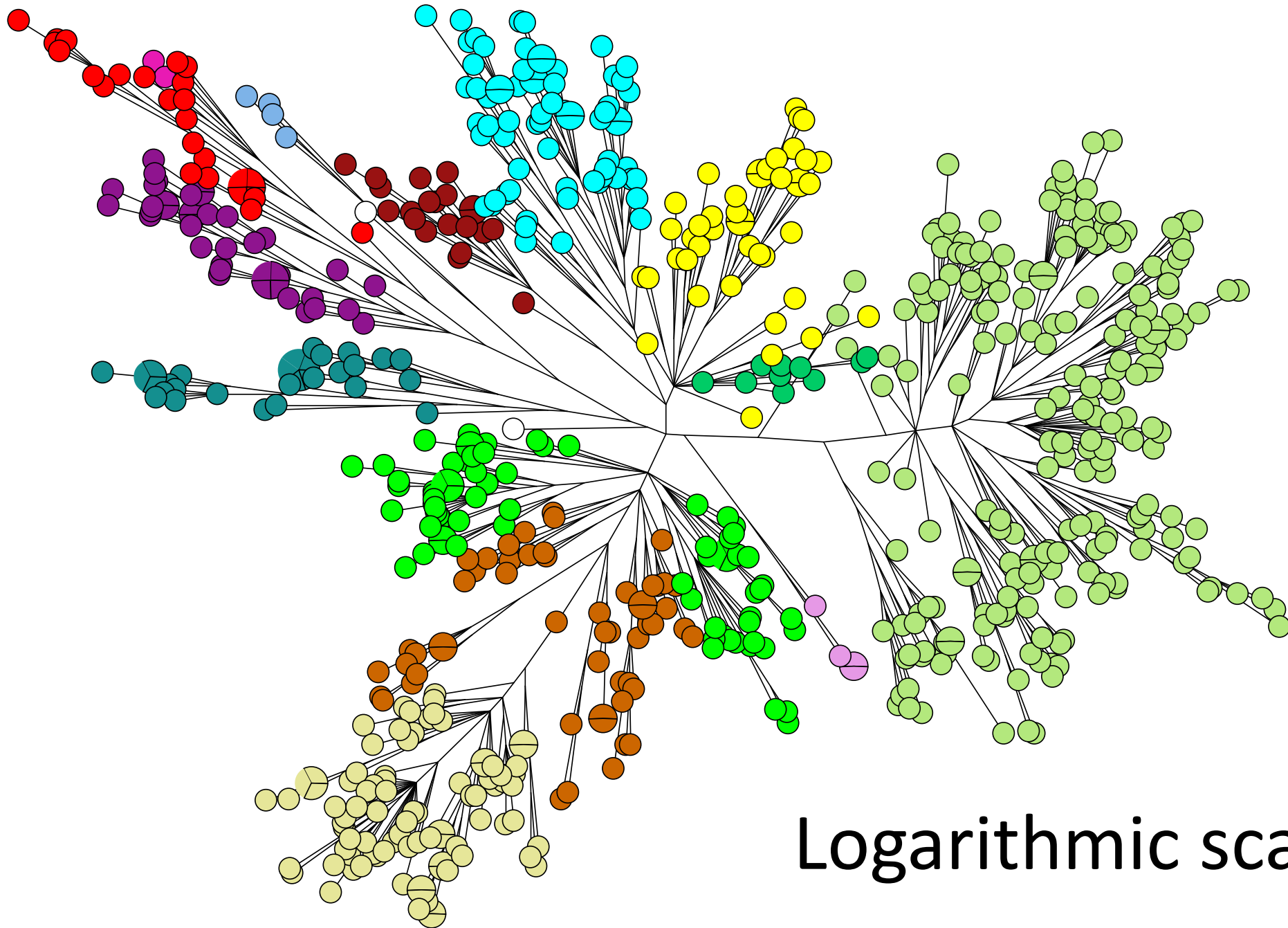
★ MRCA

682 strains

7277 SNPs

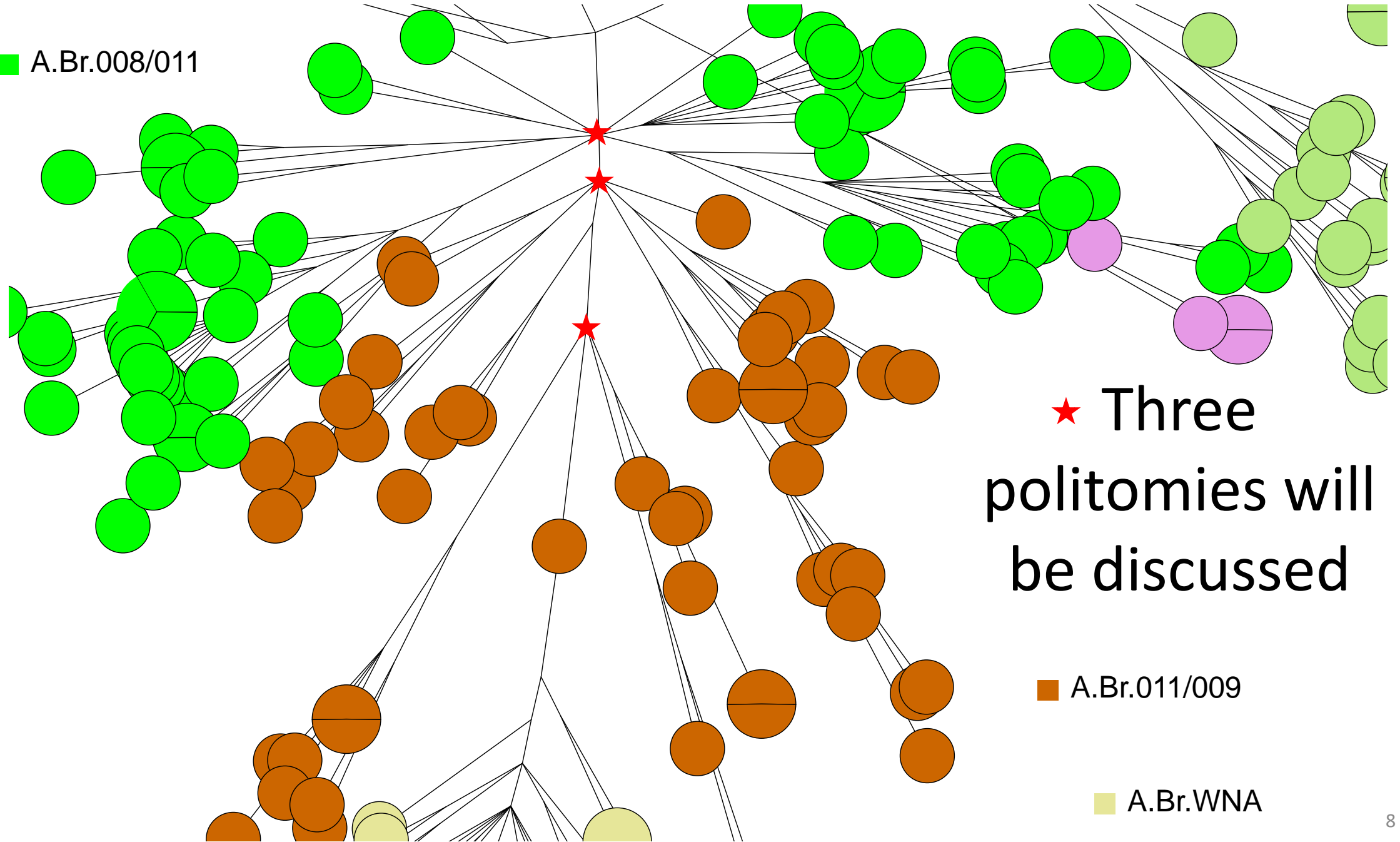
Tree size: 7375



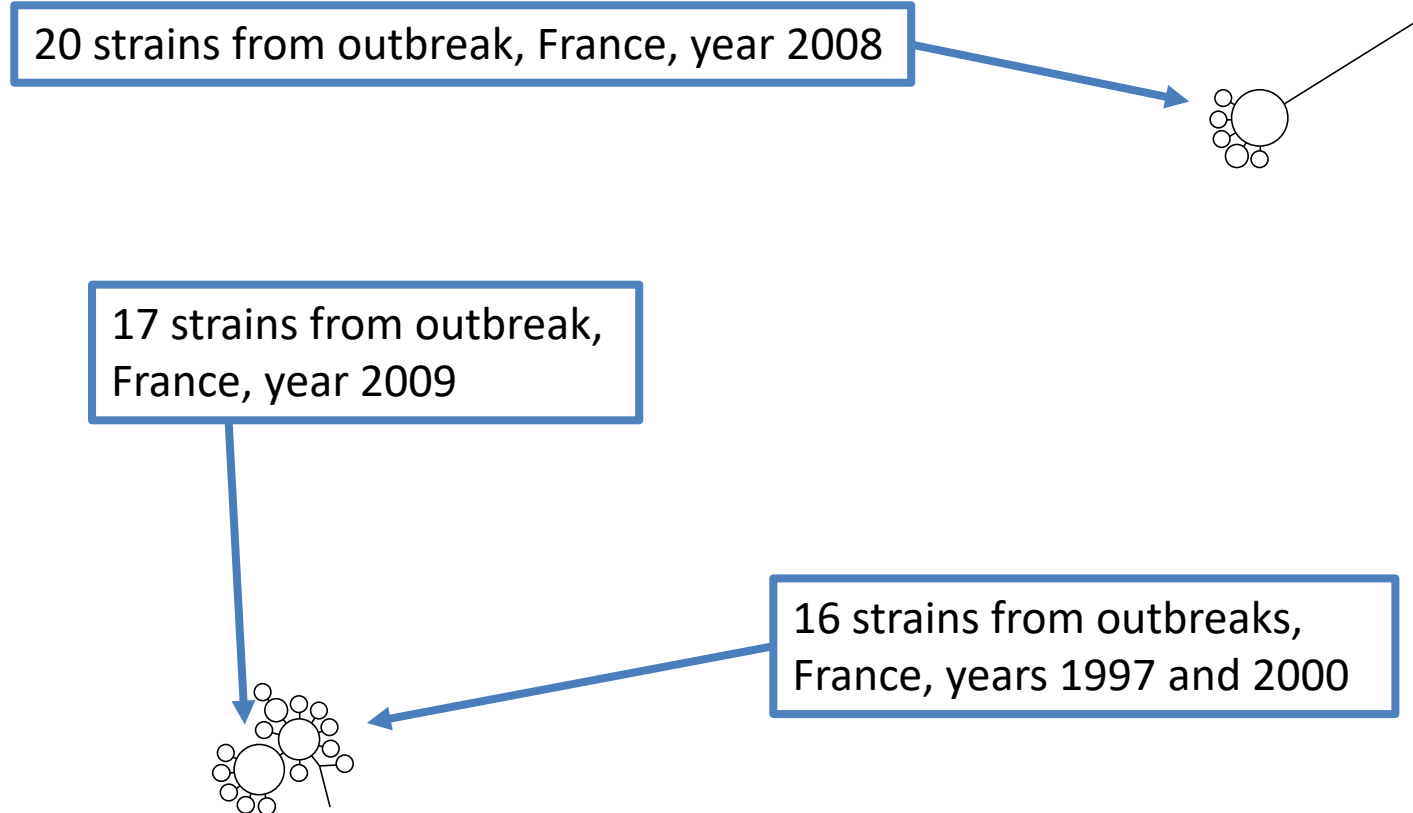


Logarithmic scale

■ A.Br.008/011



The birth of polytomies



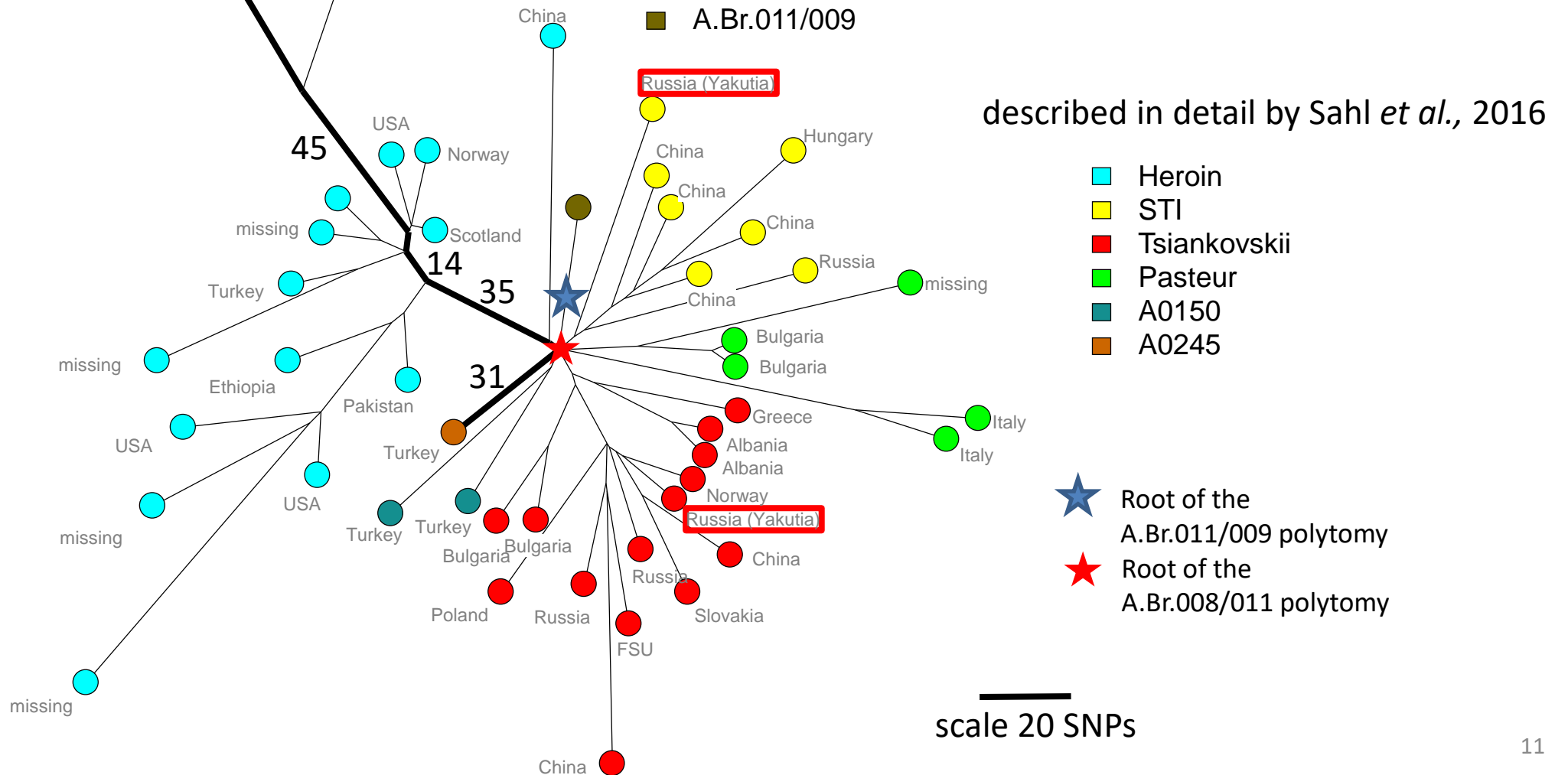
Adapted from Girault *et al.* 2014

Usually, polytomies are not maintained. Long term establishment suggests spreading. Spreading may reflect major historical events.

The « A.Br.008/011 » polytomy: first tentative dating point

The A.Br.008/011 polytomy

(approx. 70 strains)



Observation 1

- The A.Br.008/011 shows deep branching lineages present accross Eurasia (China, North-Eastern Siberia, Turkey, Bulgaria, Italy, Russia, Pakistan ...)

Interpretation

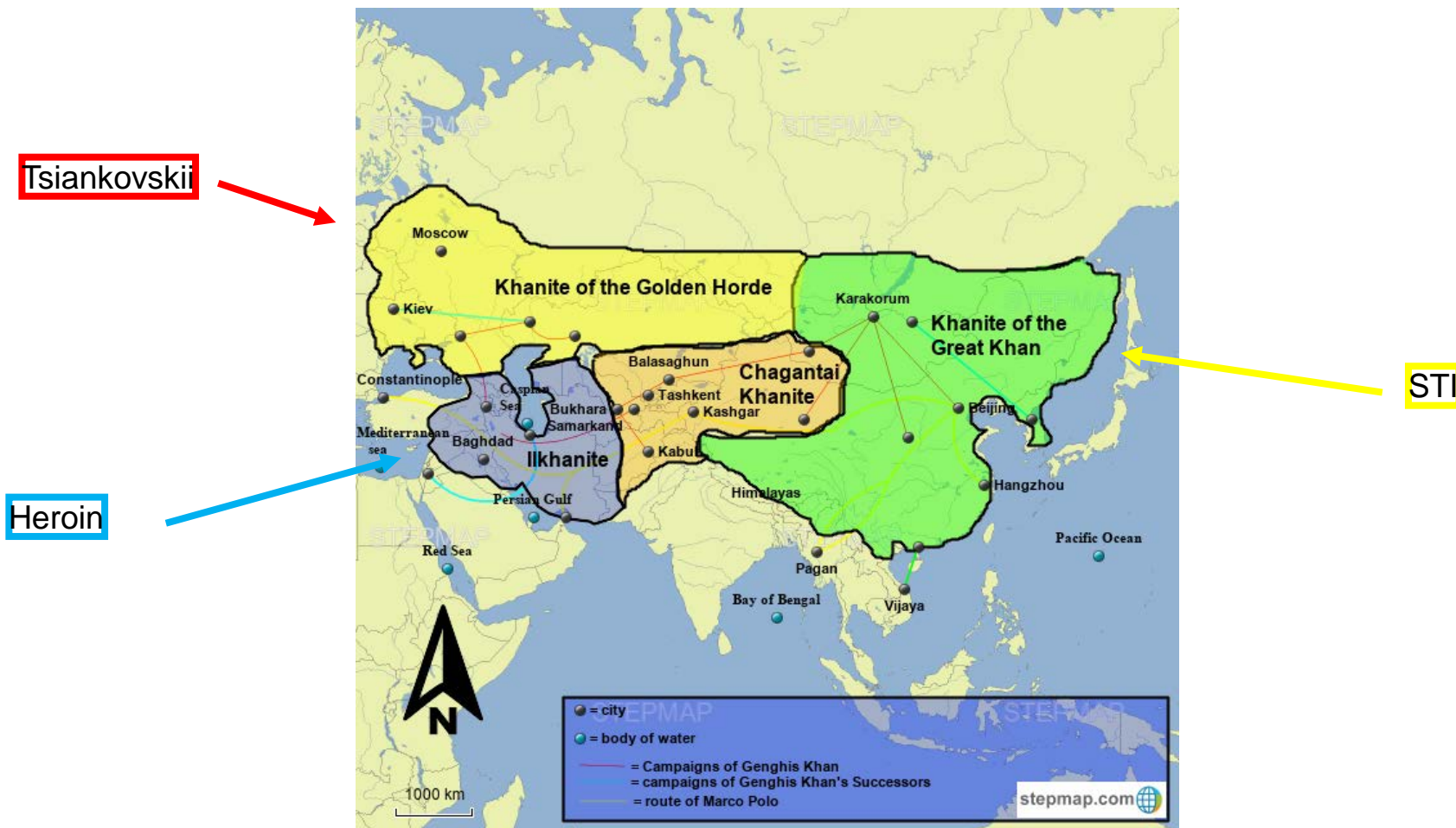
- We proposed to link the spread of this polytomy with the Mongol empire (Timofeev et al., 2019)
- Eastern Europe (as far as Hungaria and Poland) was invaded circa 1237-1241 and 1280



What might be so special about the Mongol invasions?

- Gathering of the Mongolian armies in Mongolia occurred twice during the 13th century for the election of a new Khan
 - In 1227 after the death of Genghis Khan
 - By that time, the Mongol empire had reached the Caspian sea
 - In 1242 after the death of Ögedei Khan his successor
- During the second half of the 13th century, growing independence of four khanates

The four khanates

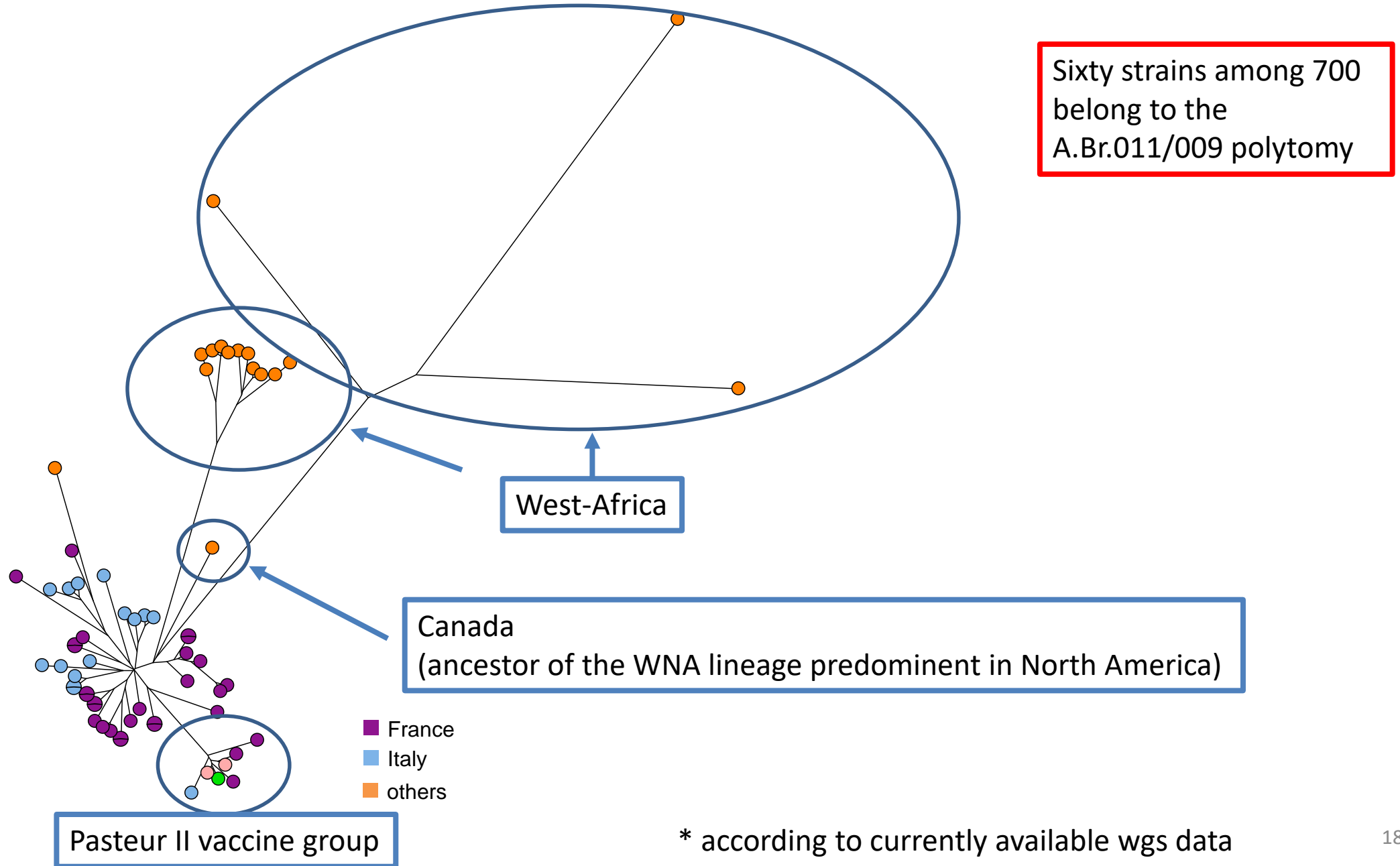


« Military operations involving animals » as large-scale spreaders of *B. anthracis*?

- The split of the Mongolian empire quickly allowed the subsequent fixation of a number of lineages
- No prediction on the geographic origin of the contamination.
- Mongolian armies were driven by horses
 - Estimates of >200.000 horses involved in Eastern Europe invasion 1235-1242

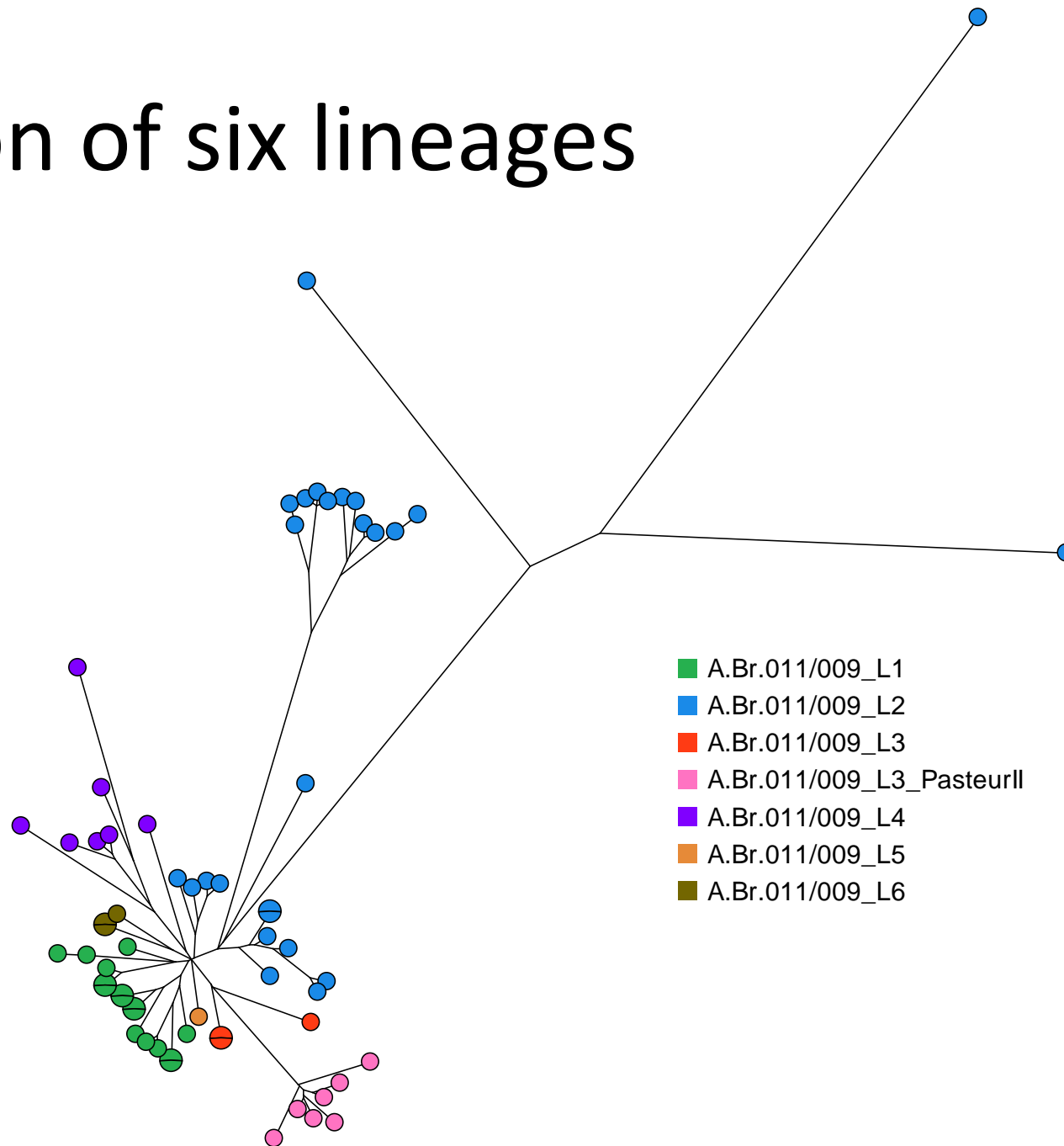
The « A.Br.011/009 » polytomy: second tentative
dating point

The A.Br.011/009 polytomy in the world*



* according to currently available wgs data

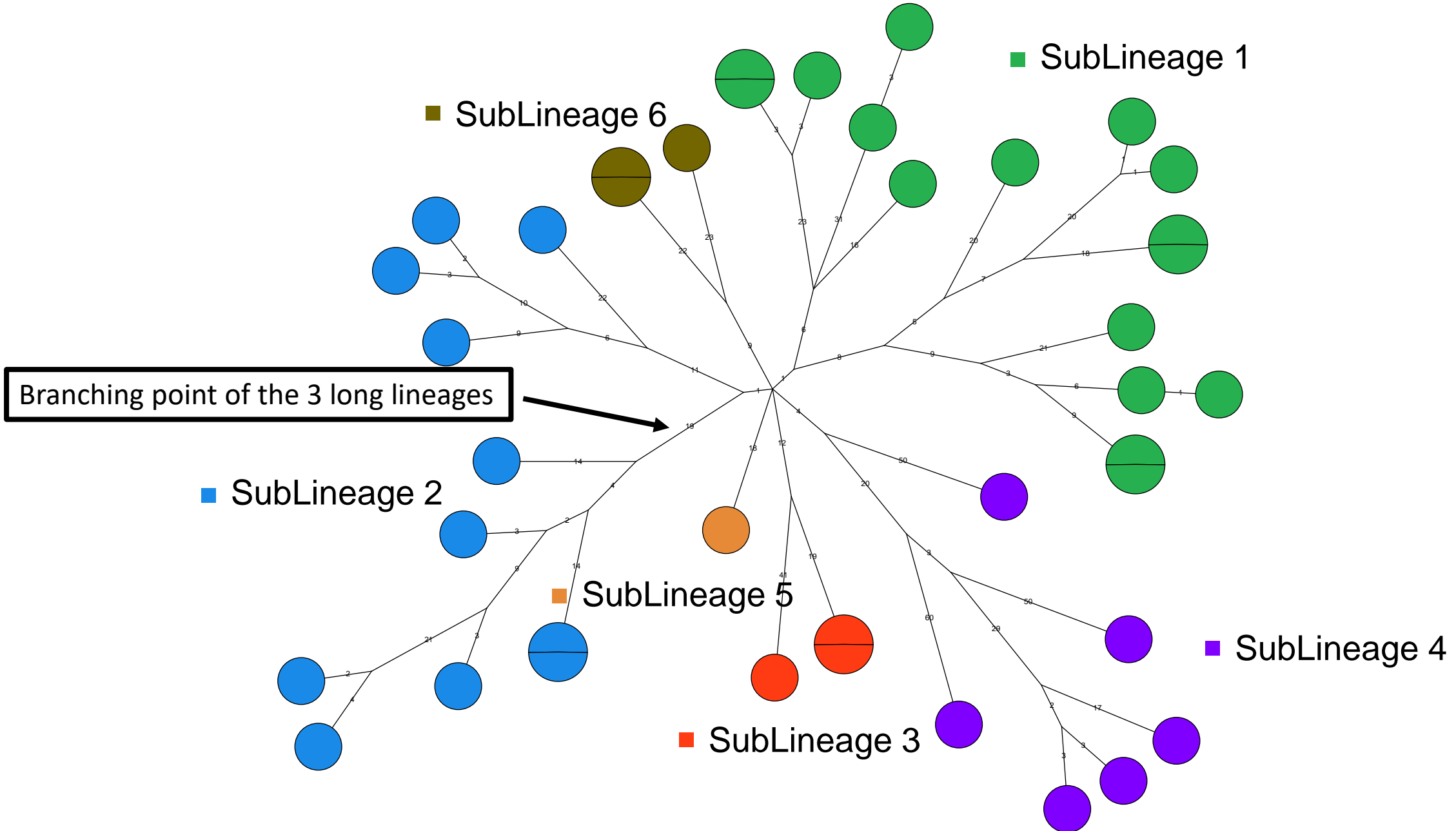
Definition of six lineages



Observation 2

- Three remarkable lineages share an identical MRCA within A.Br.011/009 sublineage 2
- Two are present in West Africa, the third is associated with the most widespread *B. anthracis* lineage in North America (WNA)
- Expansion rates along these lineages are much faster (10x)

The six A.Br.011/009 sublineages



The six lineages according to geography

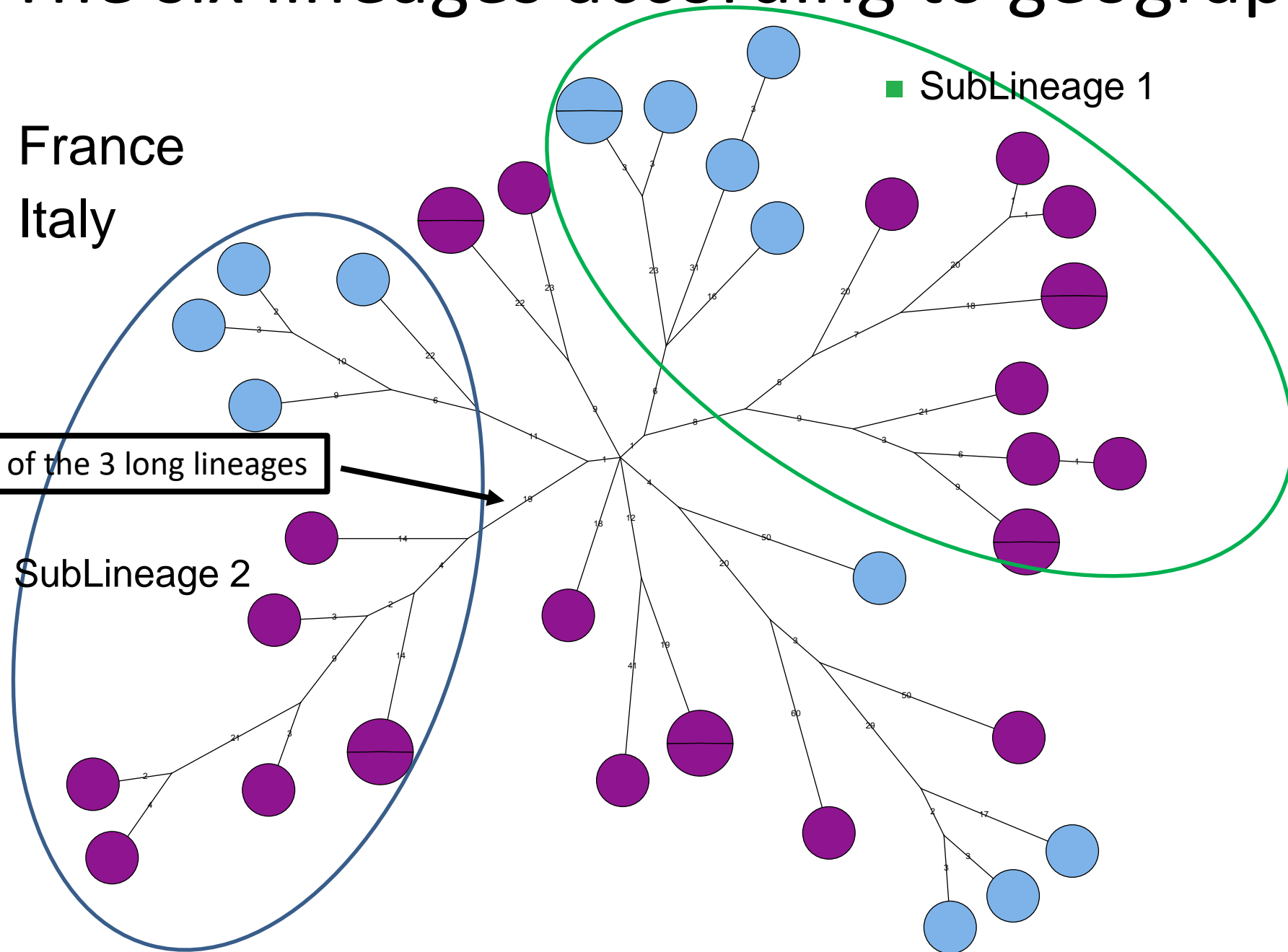
■ France

■ Italy

■ SubLineage 1

■ SubLineage 2

Branching point of the 3 long lineages



Observation 3

- Splits occurred very early on during the A.Br.011/009 polytomy expansion, leading to the definition of separate « French » and « Italian » lineages
- As if an outbreak had been ongoing simultaneously in France and Italy

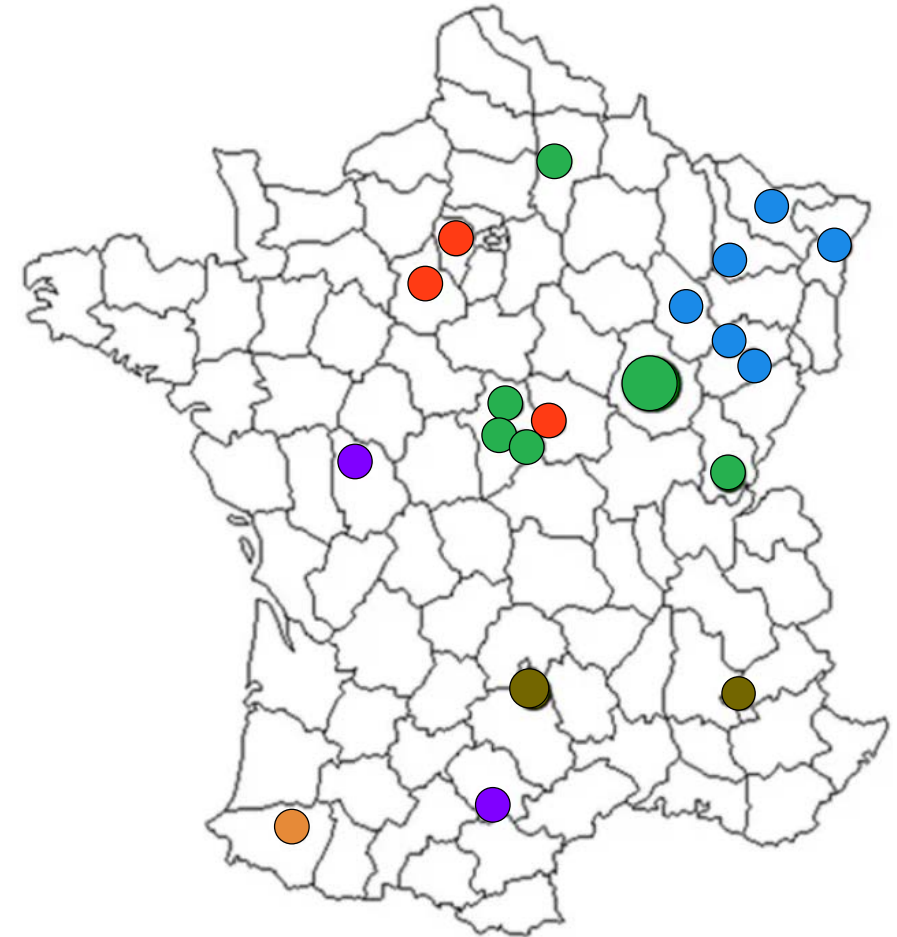
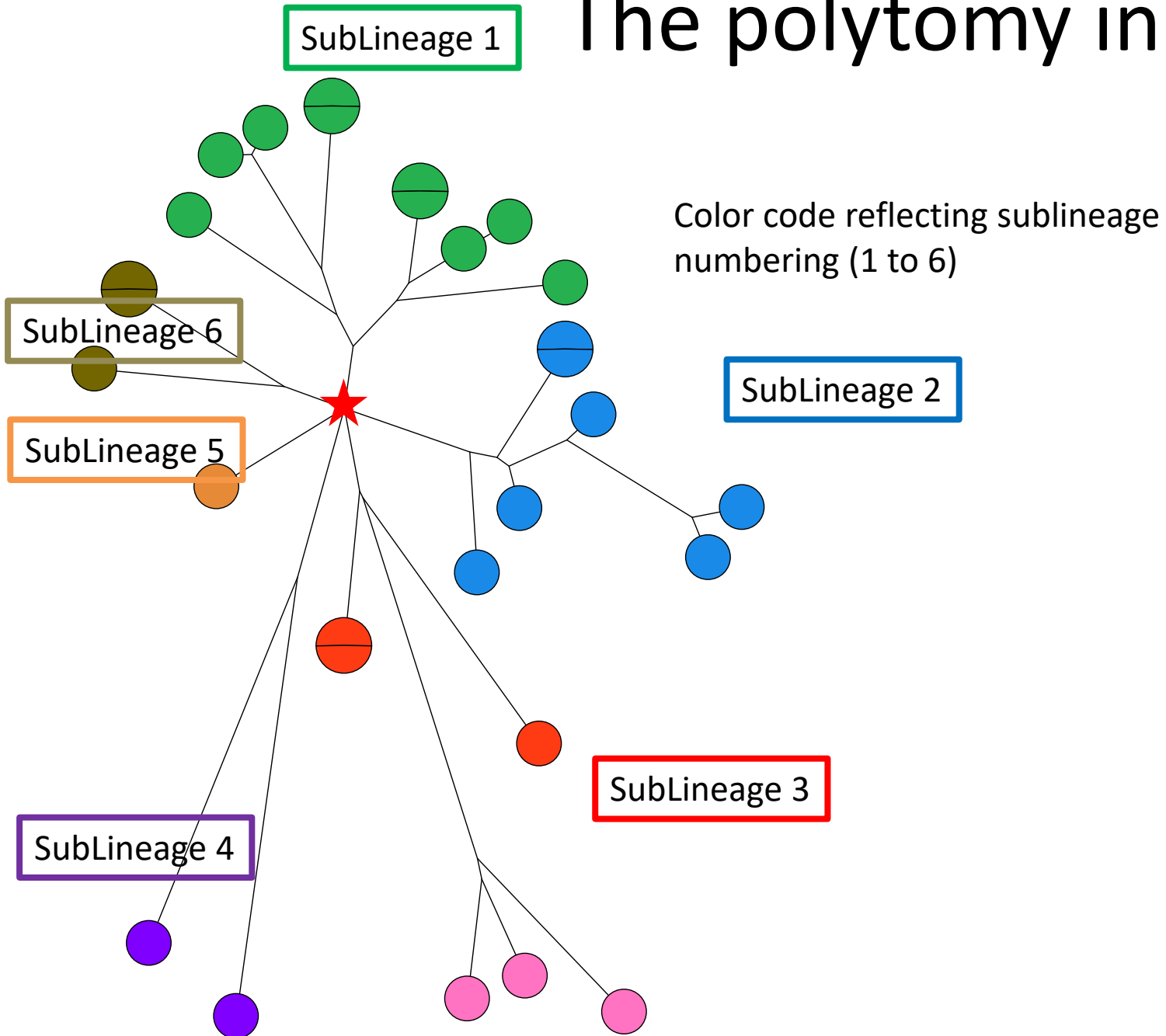
Applying the « military operations » hypothesis:

- The intricate France-Italy star-like pattern would be the by-product of conflicts between France and Italy, causing the gathering of animals (war horses including cavalry horses), contamination, and subsequent spread back to France
- The most intense among such conflicts after the 13th century occurred in 1260-1300 and 1498-1559

Franco-Italian wars

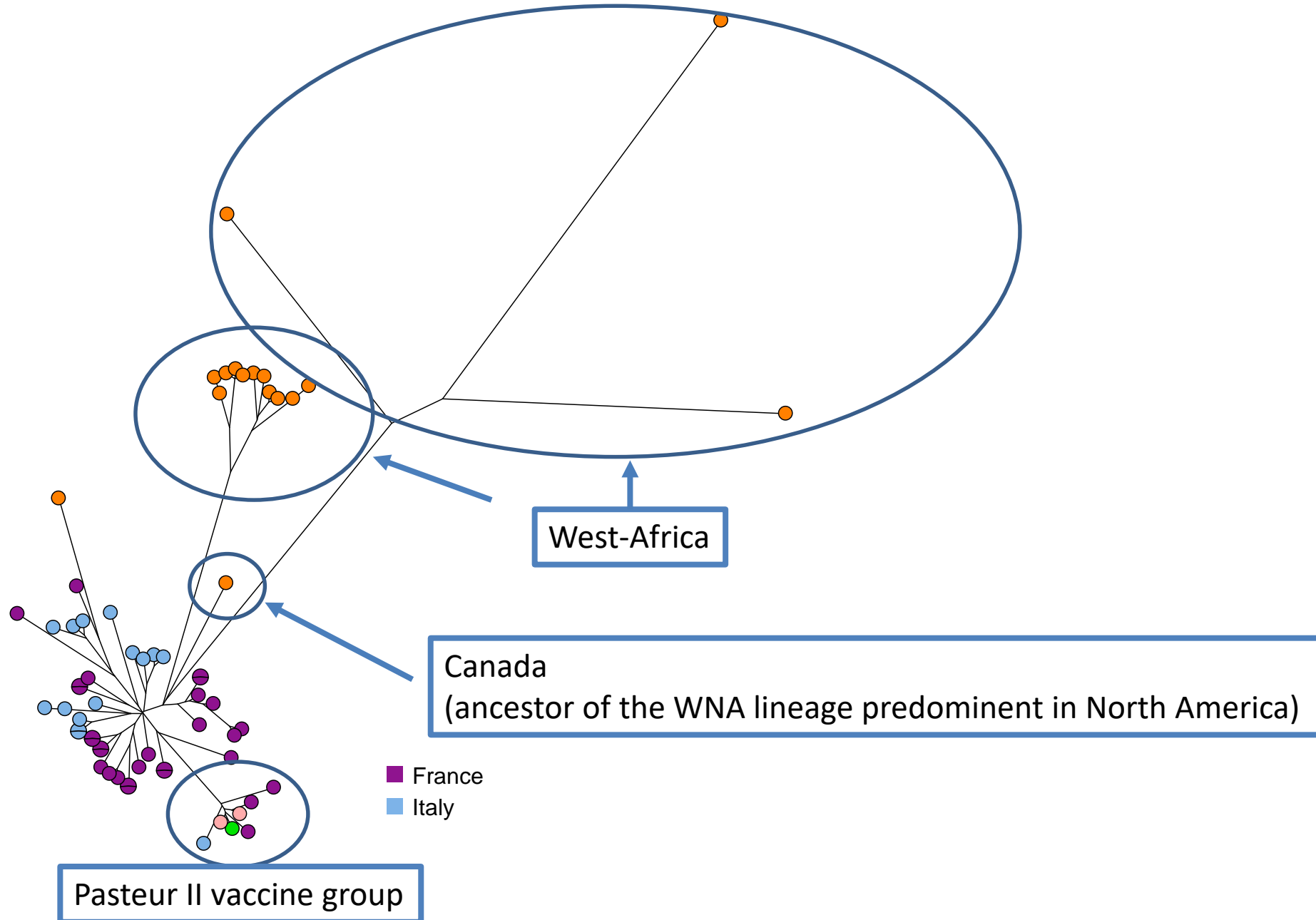
- The French claims on the Royaume de Naples were one main reason for the Franco-Italian wars, 1498-1559
- Sublineage 2 is currently present in North-East France (Lorraine) and the South part of Italy (including the Pollino National Park)

The polytomy in France



The West Africa-North America long branches

The long branches expand from SubLineage 2



Interpretation: export towards North America and West Africa

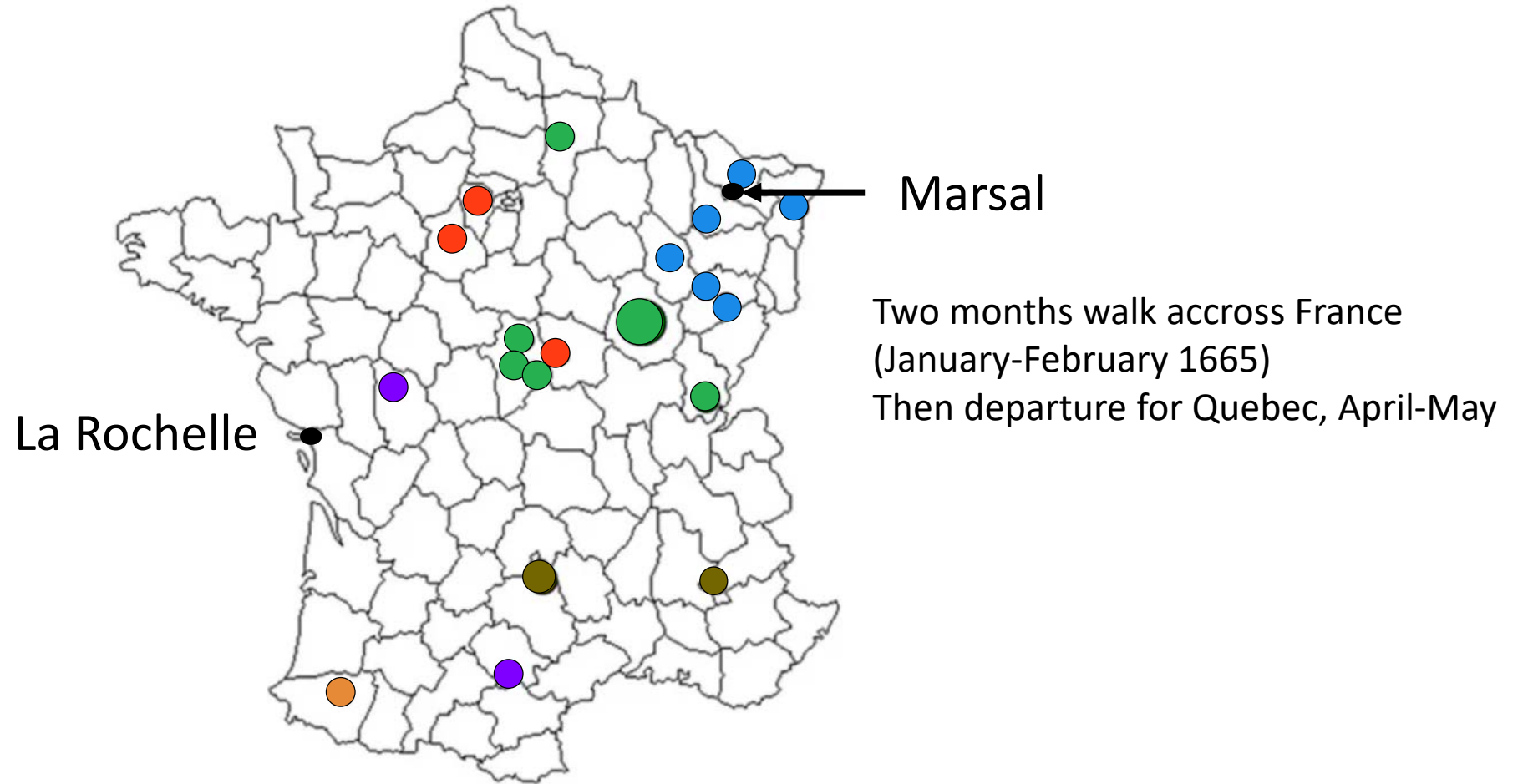
- At the « same » time and from the « same » place
- Current favorite working hypothesis: France, seventeenth century
- France was settling in both North America (Canada) and West Africa (Senegal) via the royal « Compagnie des Indes Occidentales »

The « Régiment de Carignan-Salières » hypothesis:
third dating point

Search for significant military operations between France and Canada in the 17th century

- A unique such event, the « Régiment de Carignan-Salières » expedition in 1665
- Considered an historical event in Canada

Infantry regiment sent from Marsal, North-East France to Canada, 1665



- At the same time, horses were sent from France for breeding

Horses from the Carignan-Salières regiment

- 10 companies left from Marsal
- Each company presumably included 5 to 10 horses
- What happened to the horses (not mentioned on the boats transporting soldiers)?

Horses in « Nouvelle France »

- The Canadian horse breed is derived from 80 horses sent from France in years 1665-1672
- The first horses arrived on July 16th 1665 (12 among 20 horses survived transport)
- The ship departed from La Rochelle on May 10th 1665
- The Regiment companies departed from La Rochelle on April 19th, and from May 13th to 24th
- Hypothesis: the first shipment of horses comprised horses from the Régiment de Carignan-Salières

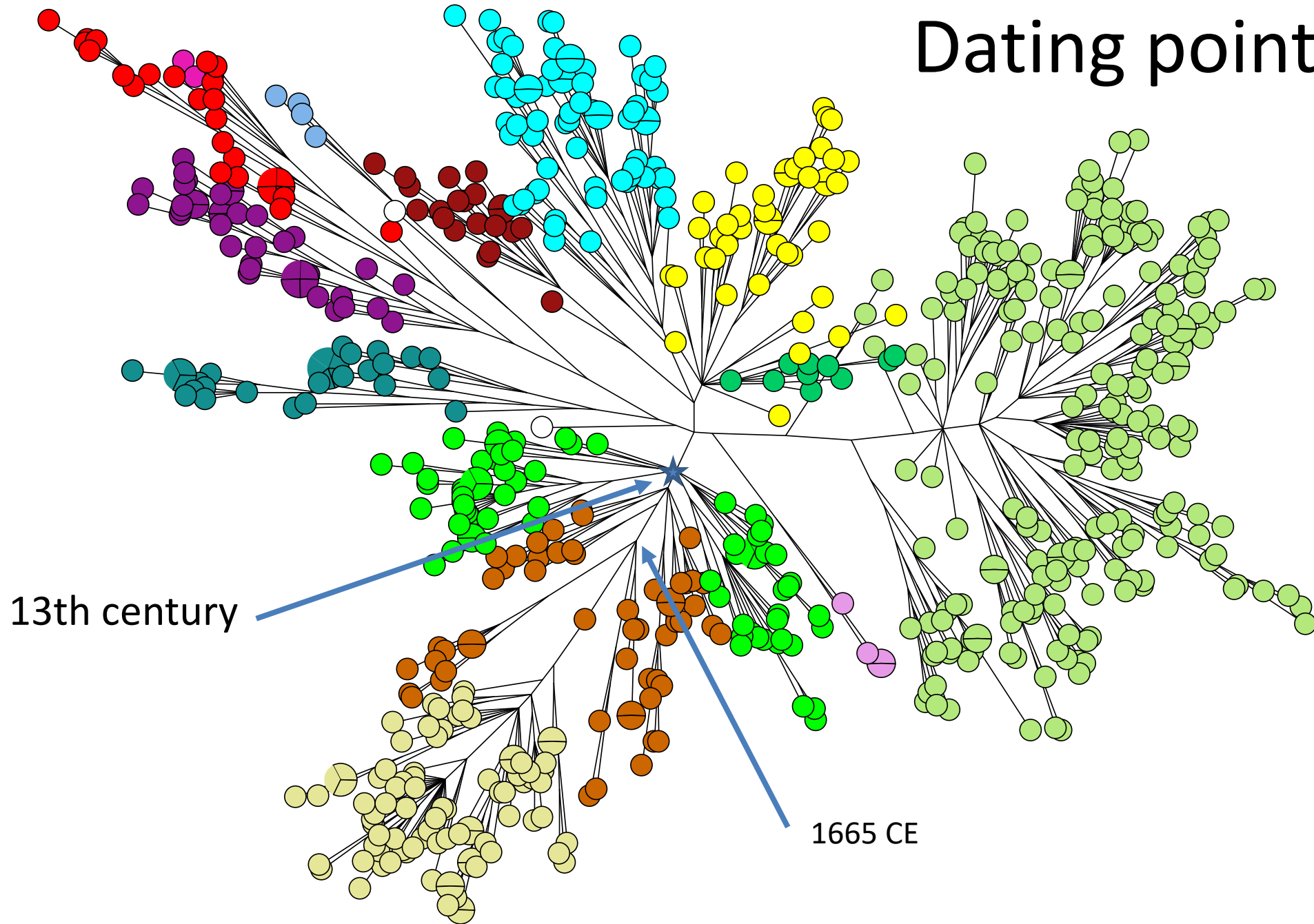
Scenario for the West Africa export

- Not all horses (only 20) were sent to Quebec
- During the 17th century, La Rochelle was a major player in the slave trade
- At that time, horses were traded for slaves
- Prediction: some of the Regiment horses were sent from La Rochelle to West Africa in 1665

Current scenario

- 13th century, Mongol invasions to Europe allowed the spreading and fixation of 7 independant lineages, constituting the A.Br008/011 polytomy
- Some of these lineages subsequently reached Italy
- 14th-17th century: conflicts between France and Italy allowed the spreading and fixation of 6 independant lineages, constituting the A.Br.011/009 polytomy present almost exclusively in France and Italy
- Regiment Carignan-Salières brings SubLineage 2 from Marsal to « Nouvelle France » in 1665, via La Rochelle
- Some of the horses are exchanged for slaves in Senegambia and/or Guinea (along the « slave coast »)

Dating points



13th century

1665 CE

Conclusion

- Animals used in military operations appear to provide interesting leads
- More data needed
- More polytomies to be explored
- Specific historical research needed
- The proposed scenario implies that horses have been healthy carriers of *B. anthracis* for months (major role for latent infections in maintenance and long-distance spreading?)

Thanks

- To Nora Madani, in charge of the French national reference laboratory for *B. anthracis* in animals
- To Christine Pourcel, for our many discussions on these topics
- To the microbiologist community for the deposition of assemblies and sequence reads archives
- To EBI-ENA and NCBI for storing and making data easily accessible
- To the developers of great software for whole genome sequence analysis
- To our local Bioinformatics infrastructures support
- To the French defense procurement agency, DGA, for financing the sequencing of hundreds of strains from dangerous pathogens

