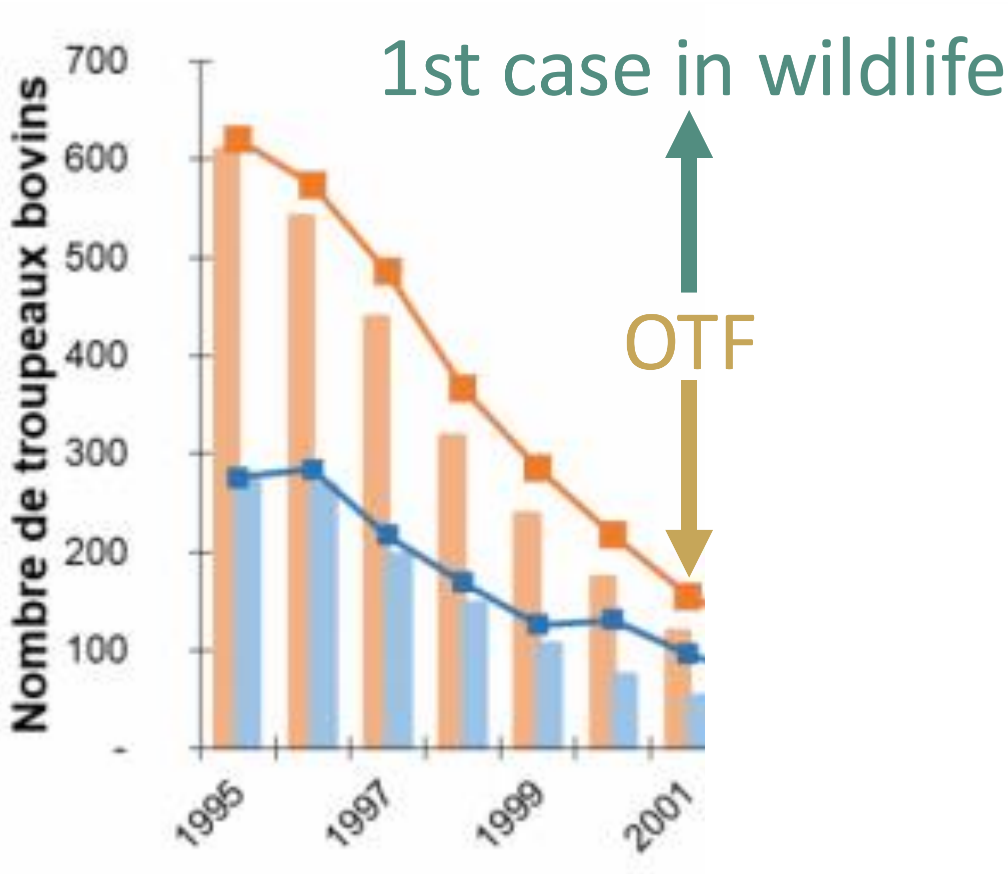
Outbreak reconstruction with a slow evolving multi-host pathogen

A comparative study of three existing methods on *Mycobacterium bovis* outbreaks

Hélène Duault, Benoit Durand, Laetitia Canini
EpiMIM, Laboratoire de Santé Animale, Anses

France's officially bovine tuberculosis-free status under threat

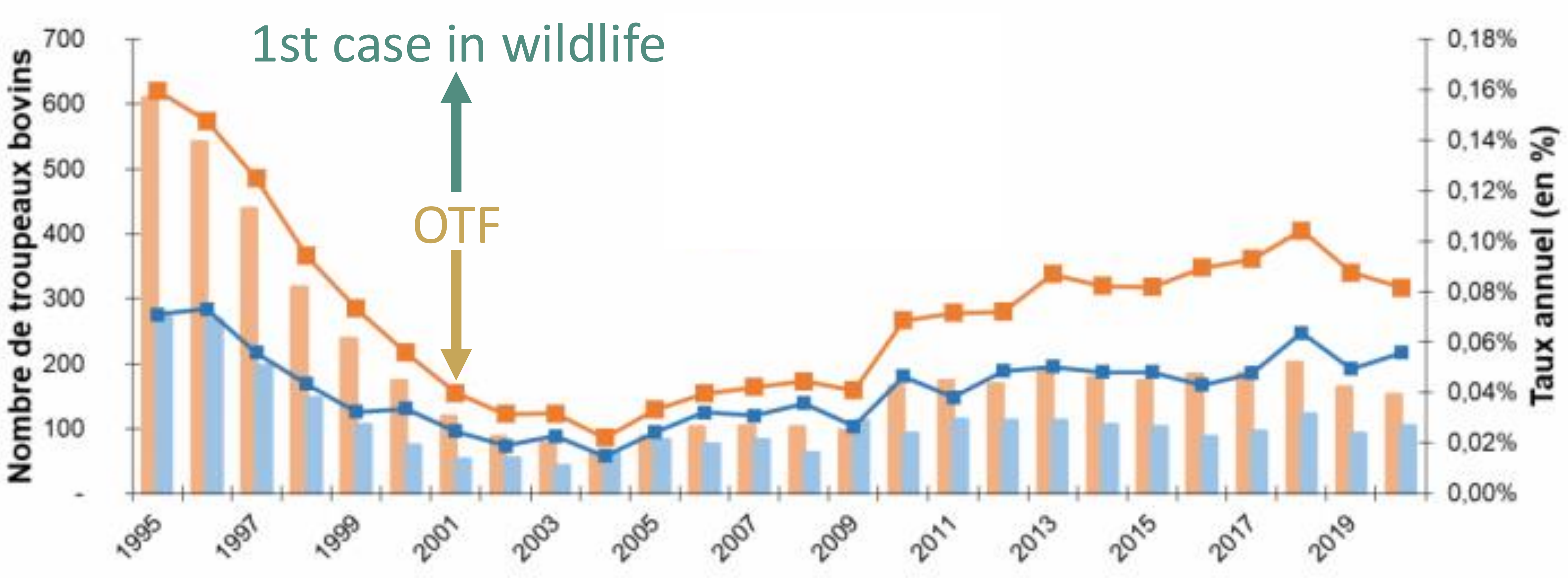


1st case in wildlife

OTF

Trends in the prevalence and incidence of bovine tuberculosis from 1995 to 2020 (© https://www.plateforme-esa.fr)

France's officially bovine tuberculosis-free status under threat



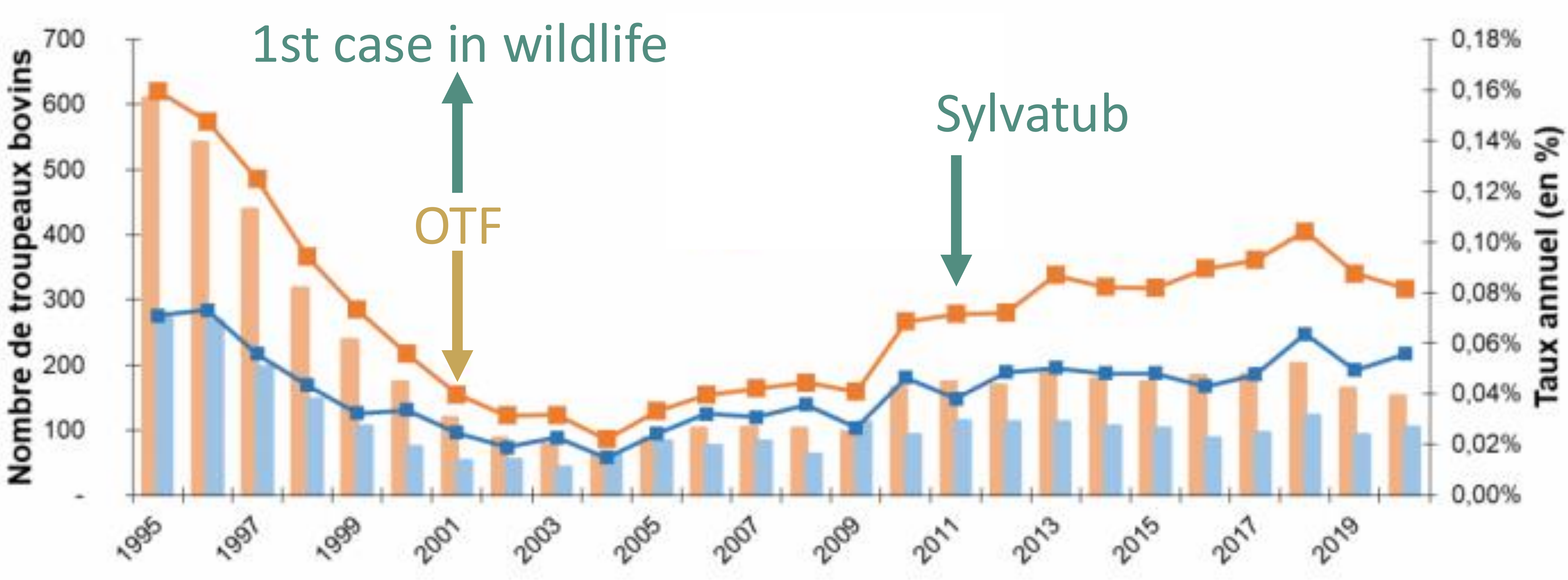
1st case in wildlife

OTF

Trends in the prevalence and incidence of bovine tuberculosis from 1995 to 2020

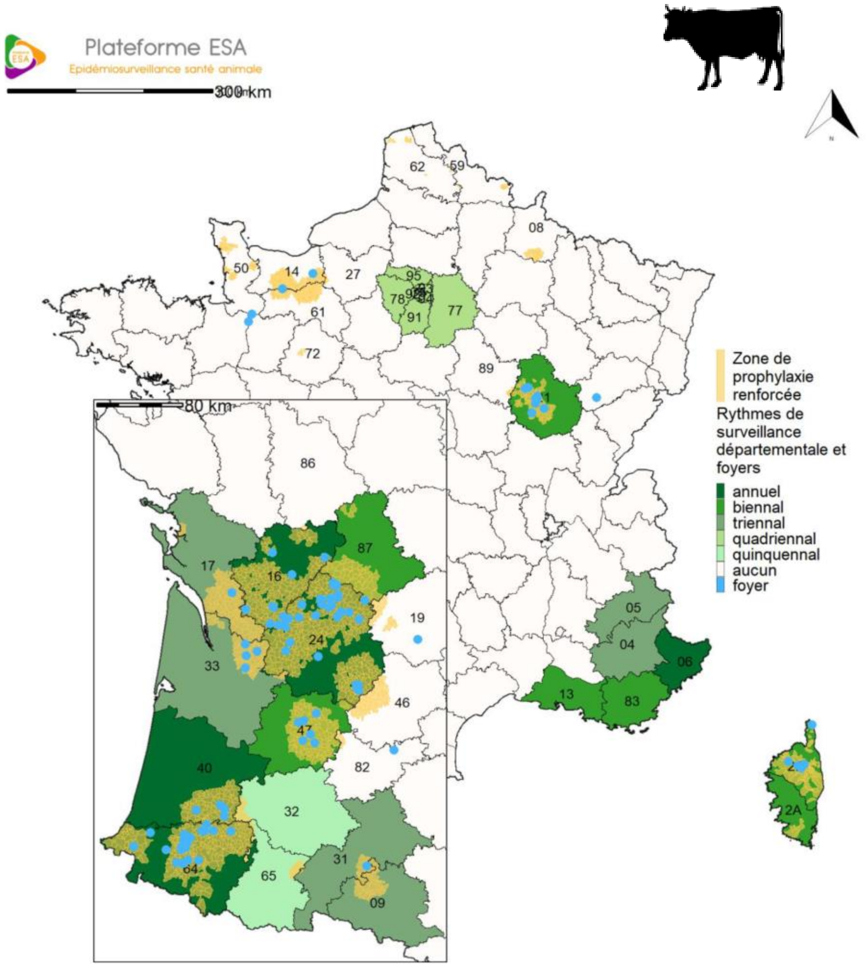
(© https://www.plateforme-esa.fr)

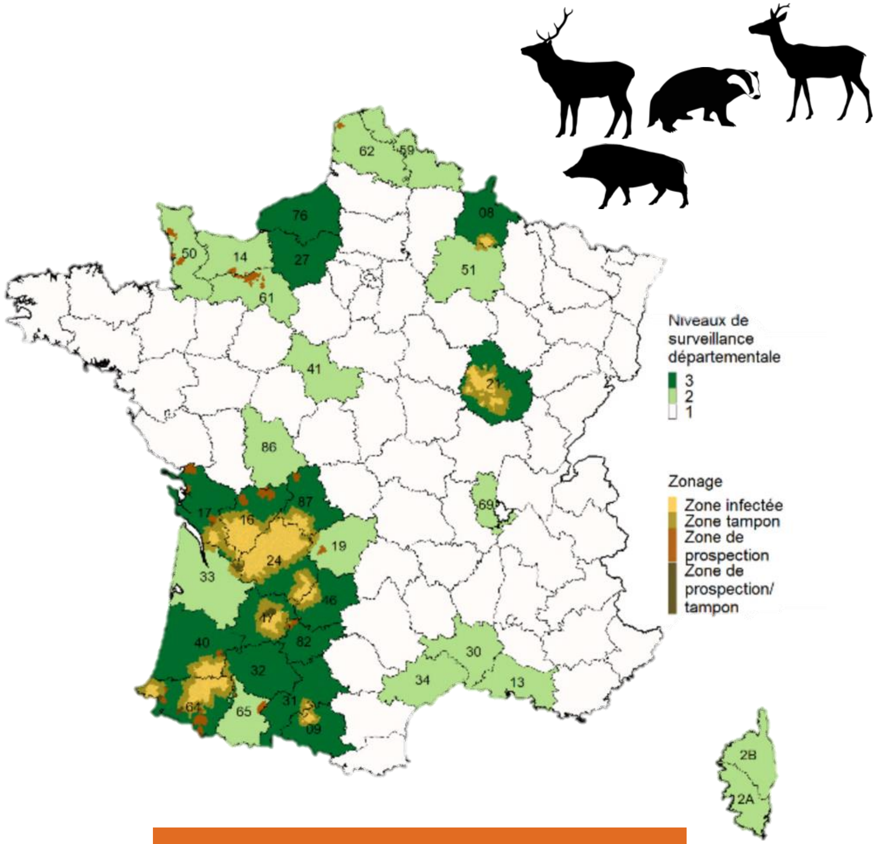
France's officially bovine tuberculosis-free status under threat



Trends in the prevalence and incidence of bovine tuberculosis from 1995 to 2020 (© https://www.plateforme-esa.fr)

Two surveillance systems with unbalanced sampling schemes

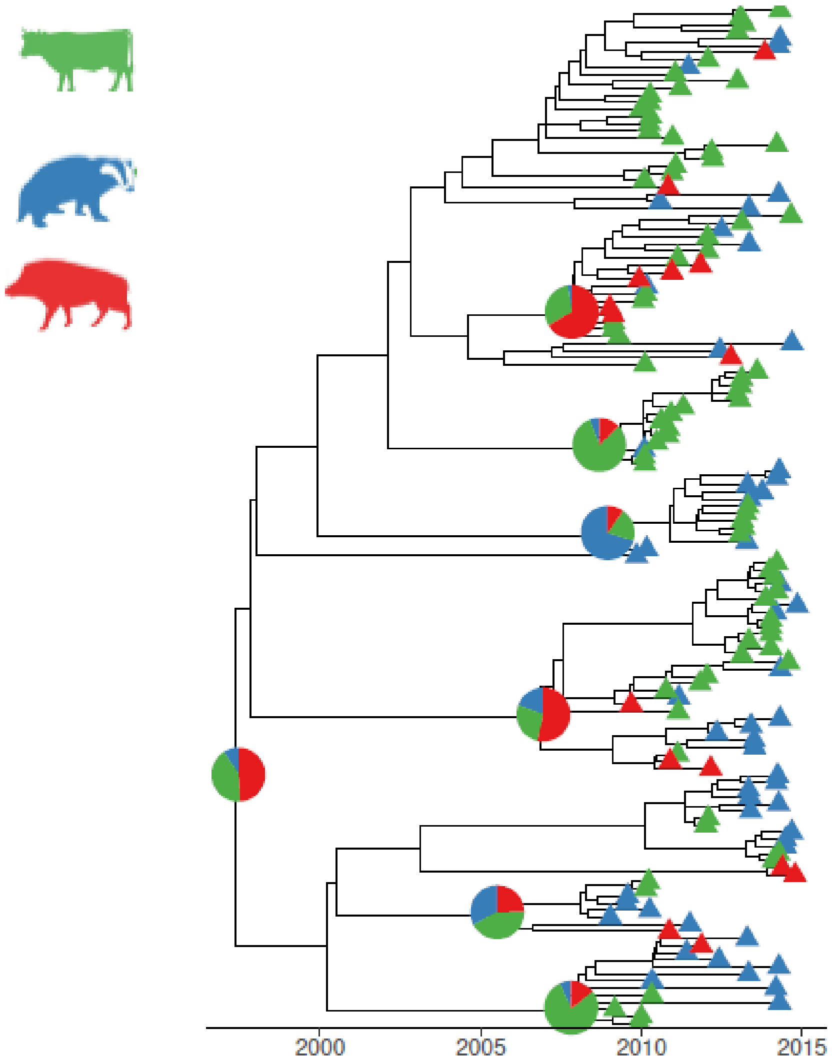


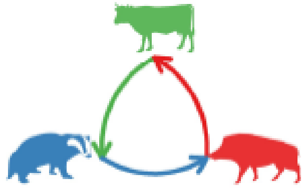
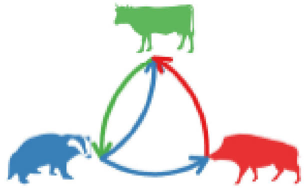
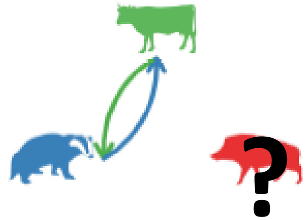


Since 2011

Forfait et al. (2021) Bulletin Epidémiologique

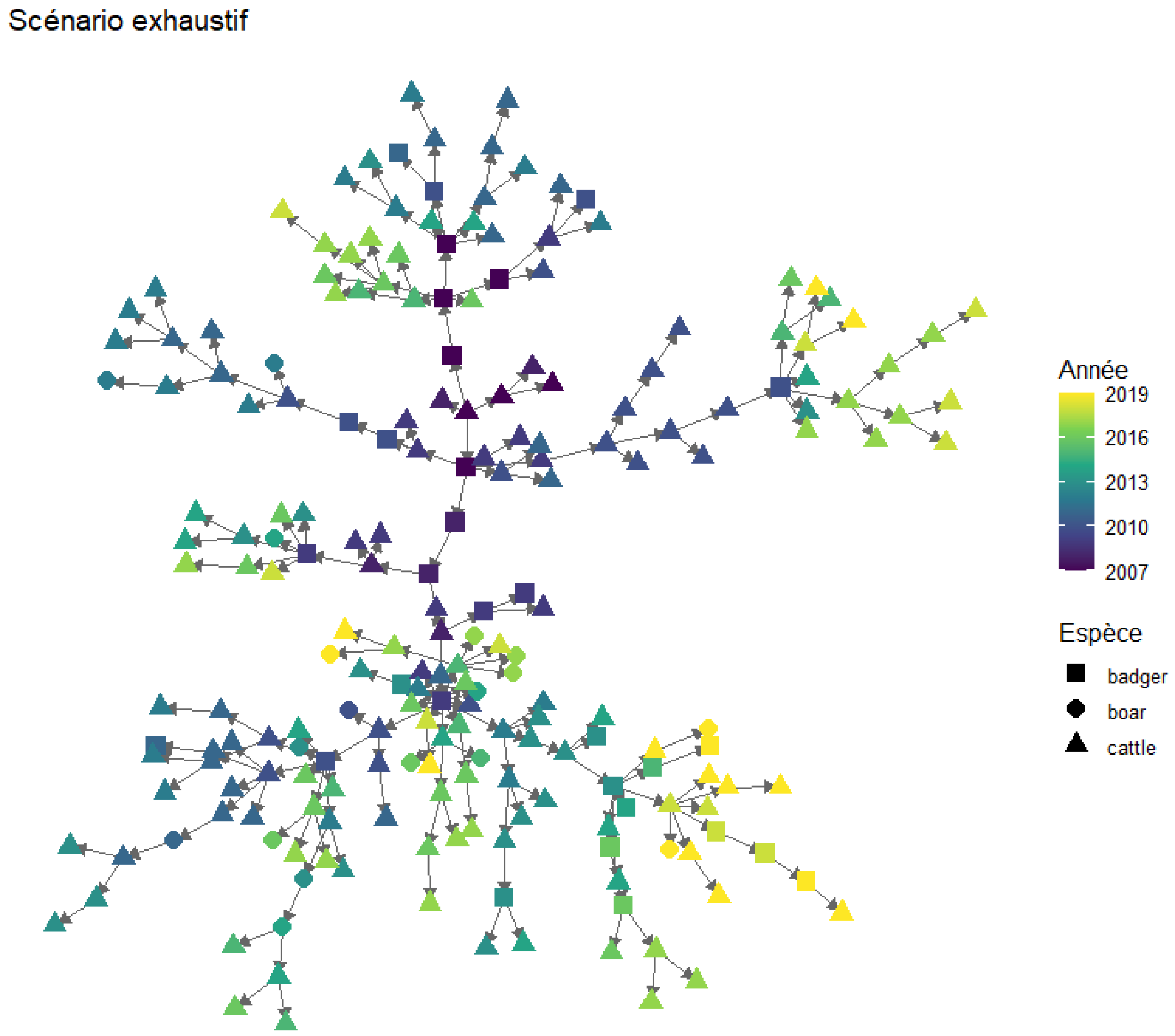
Phylogenetic tree reconstruction was used to study transmission between species



	MRCA	Evolution rate	Inter-species transmission
Côte-d’Or	1997 [1991 – 2002]	0.42 [0.31 – 0.54]	
Dordogne / Haute-Vienne	1991 [1971 – 2000]	0.57 [0.54 – 0.71]	
PA / Landes	1990 [1980 – 1996]	0.41 [0.29 – 0.55]	

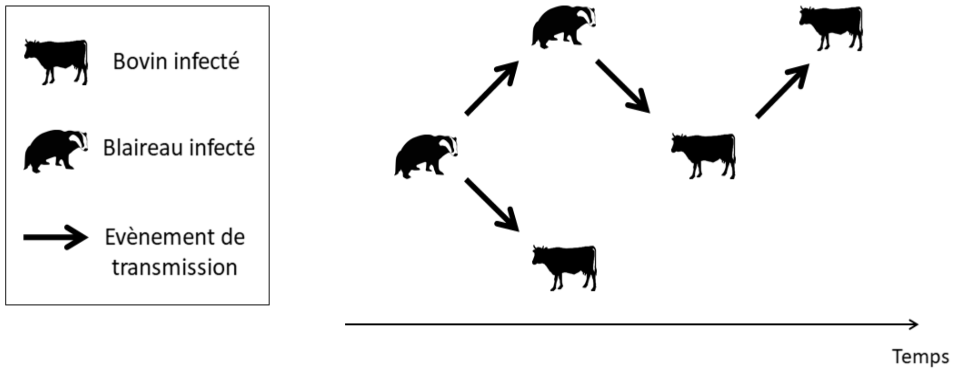
Duault et al. (2022) Veterinary Research
Canini et al. (2023) Microbiology Open

Transmission trees is a better restitution tool



Incidence and genomic data can be combined in transmission trees

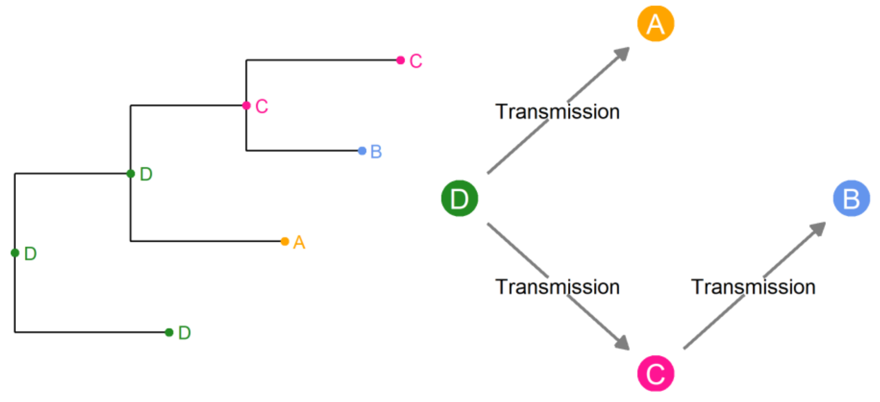
- 22 methods identified...
- None for multi-host systems



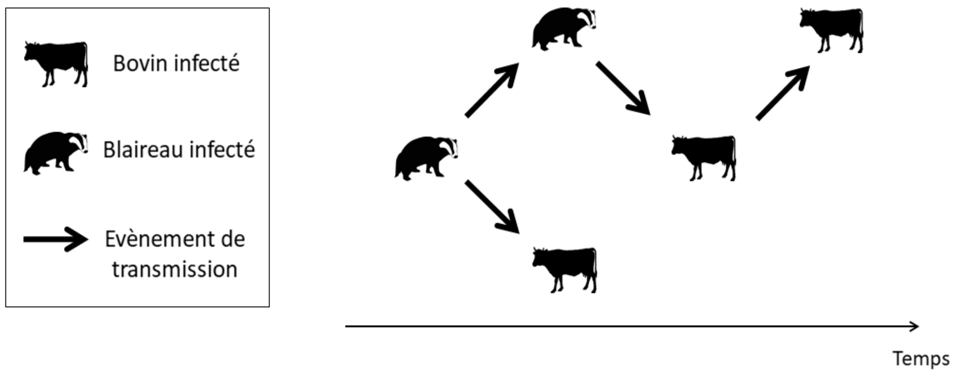
Duault et al. (2022) Pathogens

Incidence and genomic data can be combined in transmission trees

- 22 methods identified...
- None for multi-host systems

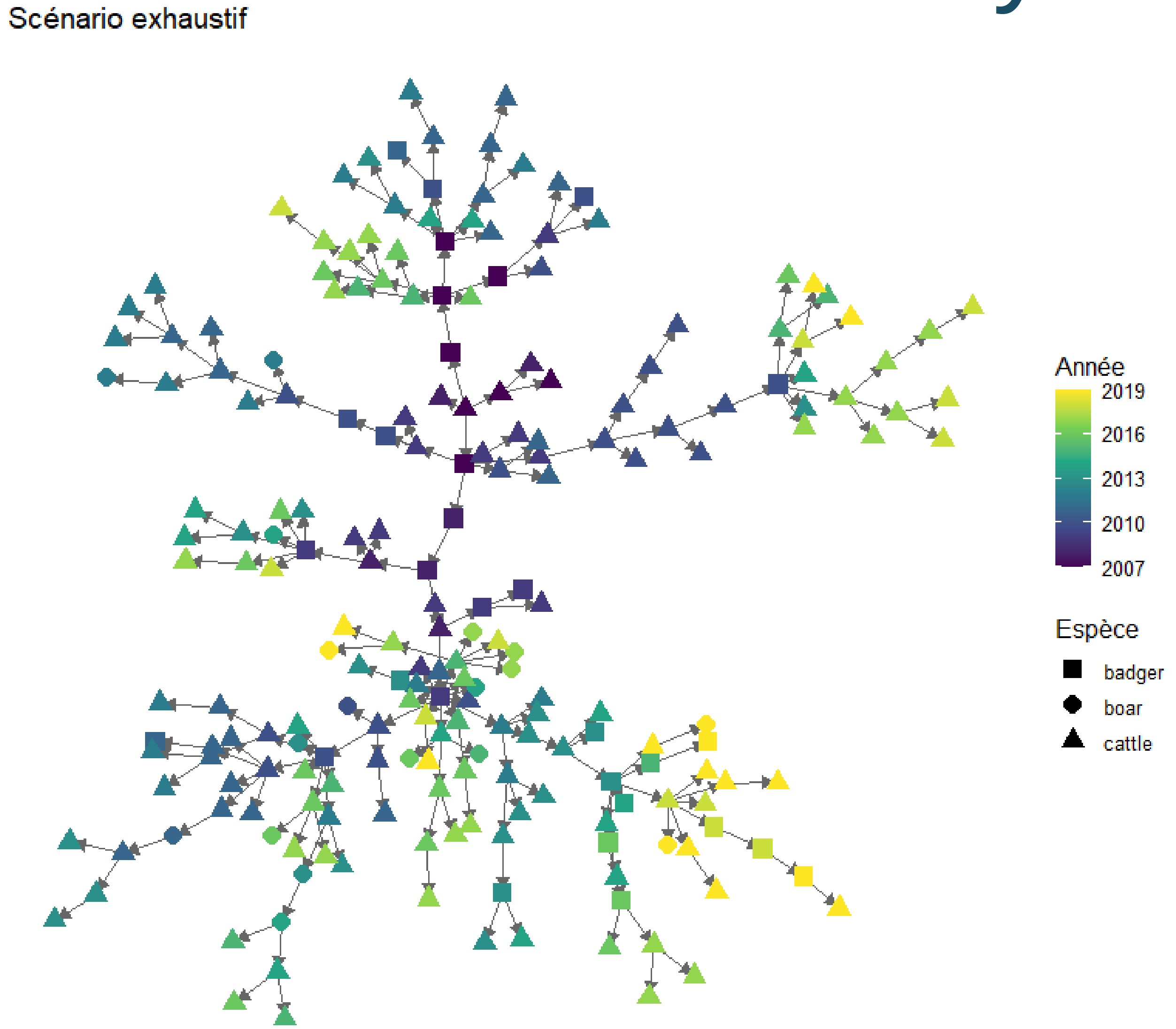
Phylogenetic (PF)		Non-phylogenetic (NPF)
Sequential (SeqPF)	Simultaneous (SimPF)	
N = 5	N = 9	N = 8
Link between phylogenetic trees and transmission trees		Estimation of genetic distances between pairs of sequences
		

	A	B	C	D
A	0	6	4	2
B	6	0	2	4
C	4	2	0	2
D	2	4	2	0

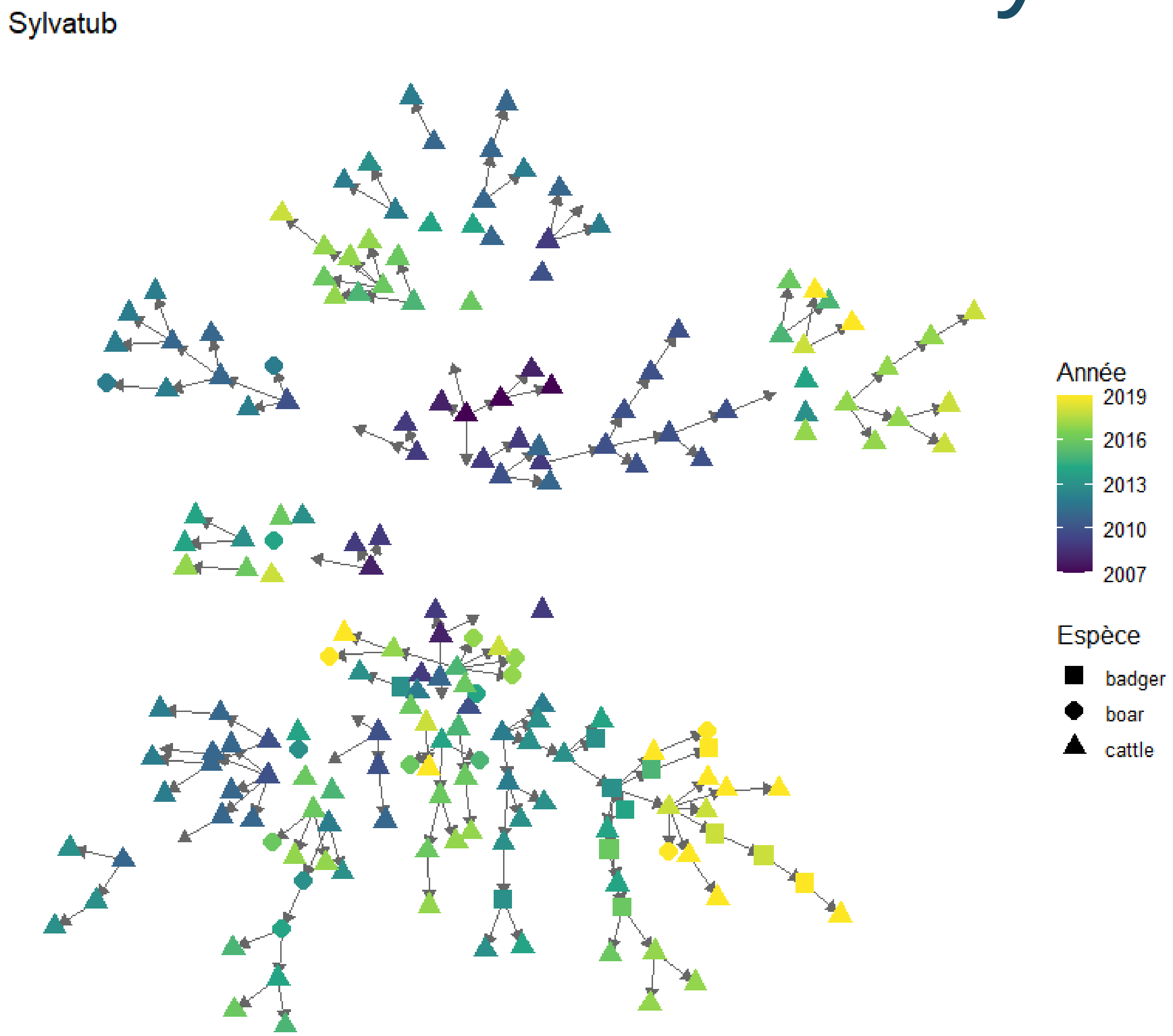


Duault et al. (2022) Pathogens

A better understanding of the impact of sampling bias in a multi-host system

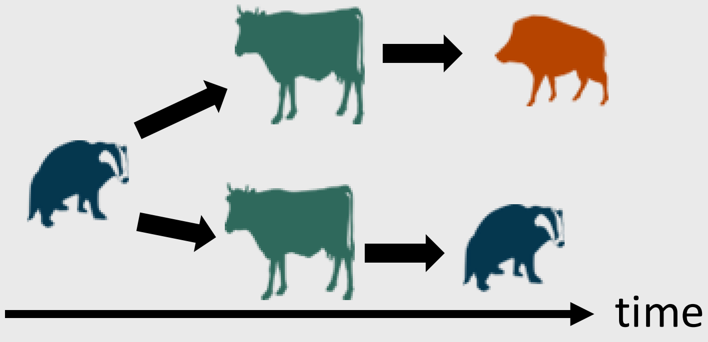


A better understanding of the impact of sampling bias in a multi-host system



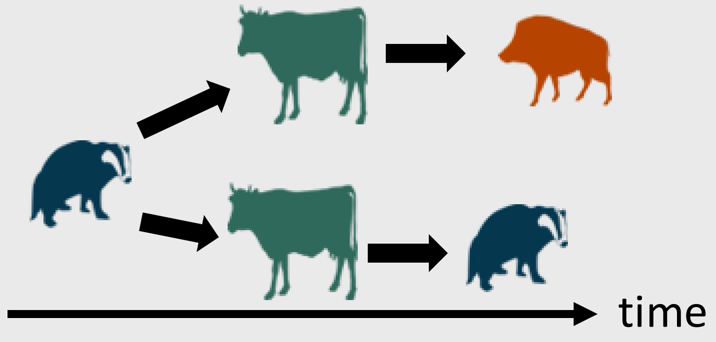
A 5-step process for evaluating methods using different scenarios

1. Simulating transmission tree

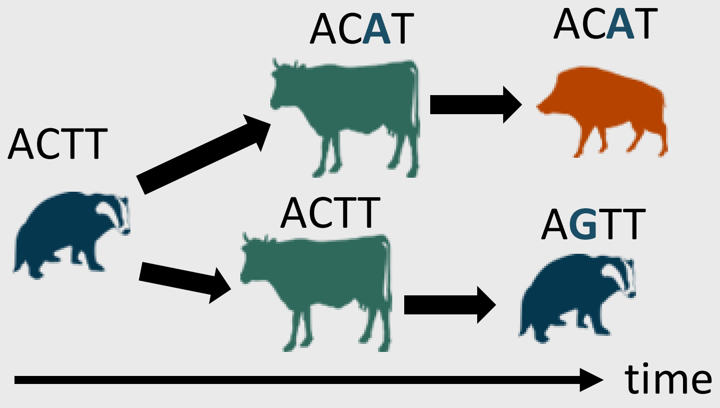


A 5-step process for evaluating methods using different scenarios

1. Simulating transmission tree

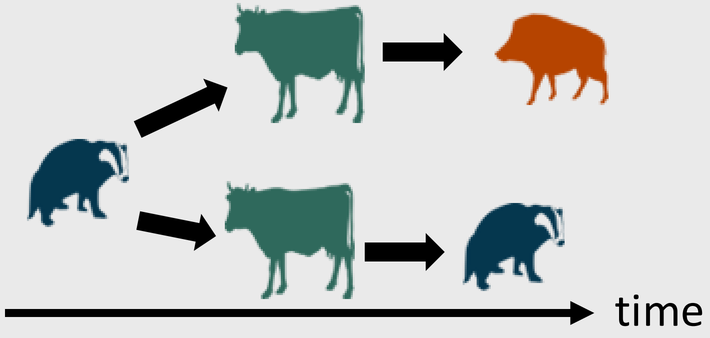


2. Simulating sequences

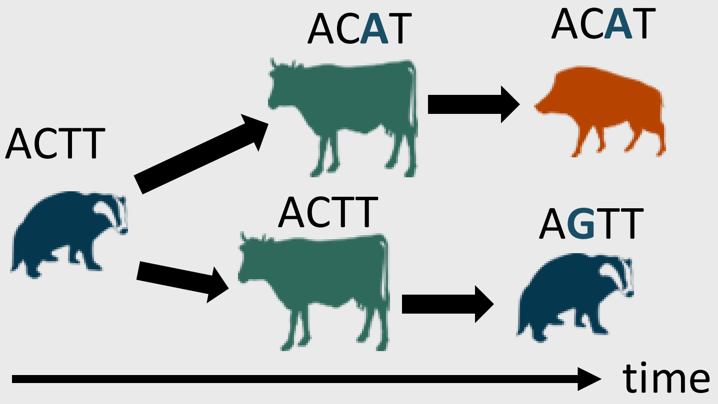


A 5-step process for evaluating methods using different scenarios

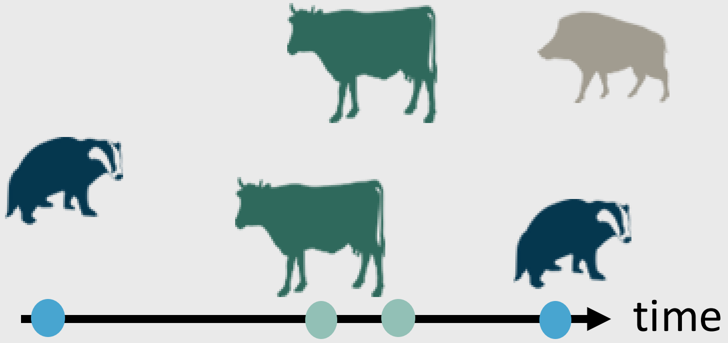
1. Simulating transmission tree



2. Simulating sequences

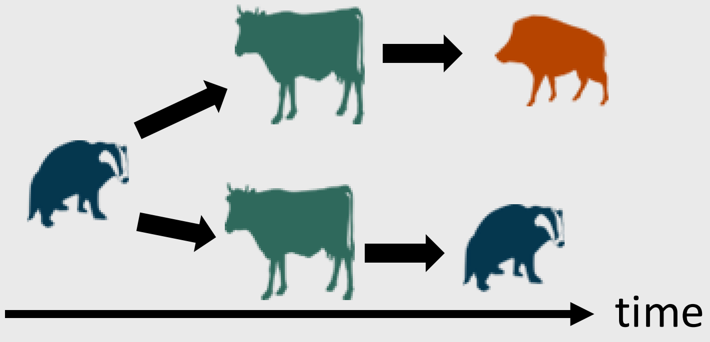


3. Definig sampling bias scenarios

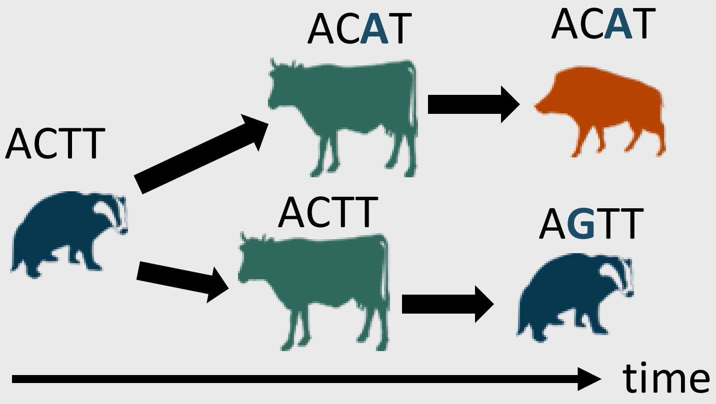


A 5-step process for evaluating methods using different scenarios

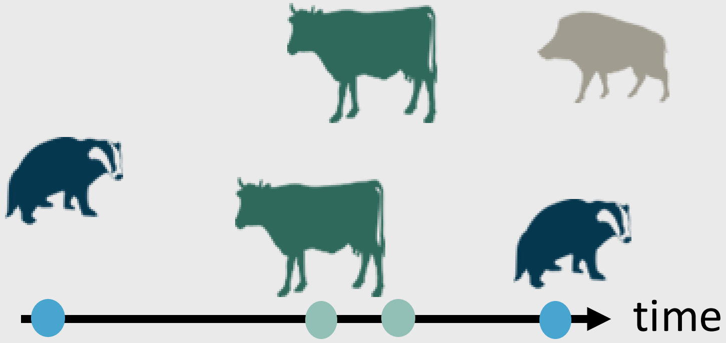
1. Simulating transmission tree



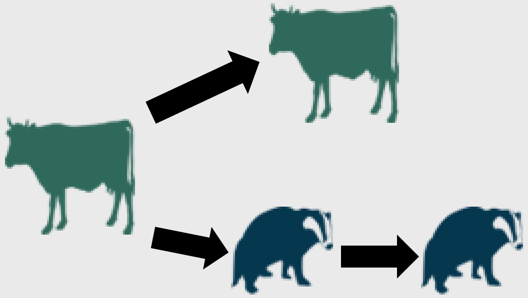
2. Simulating sequences

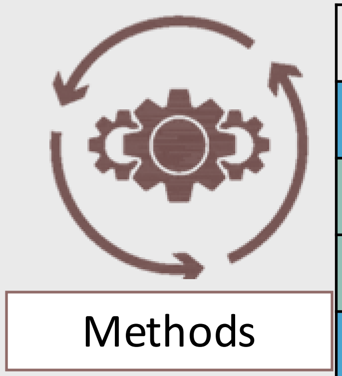


3. Definig sampling bias scenarios



4. Reconstruction

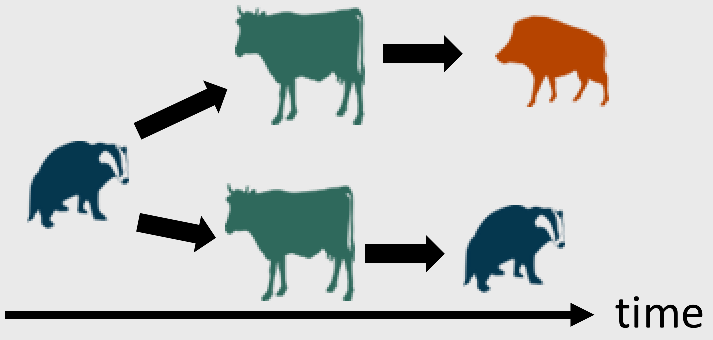




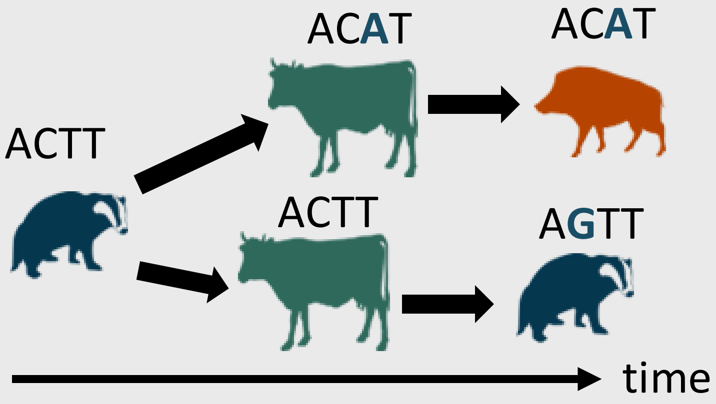
Time	Sequence	Species
T1	ACTT	Badger
T2	ACTT	Cattle
T3	ACAT	Cattle
T4	AGTT	Badger
…		

A 5-step process for evaluating methods using different scenarios

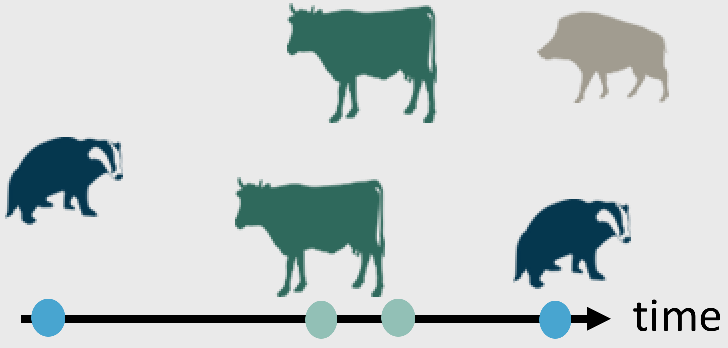
1. Simulating transmission tree



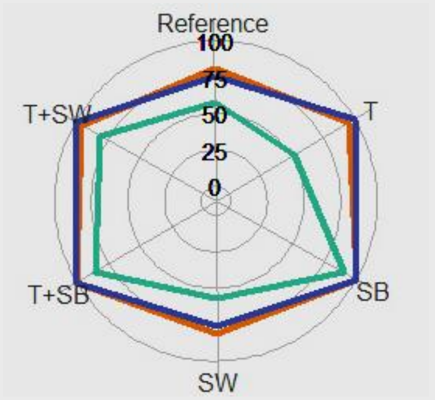
2. Simulating sequences



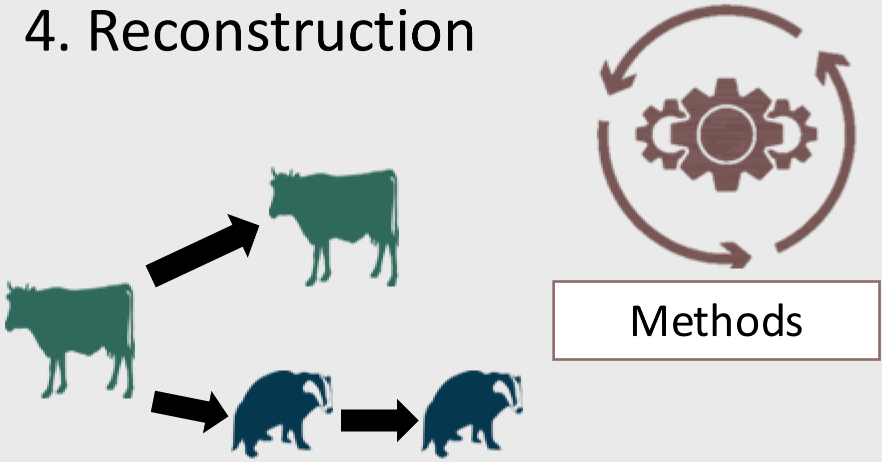
3. Definig sampling bias scenarios



5. Evaluation

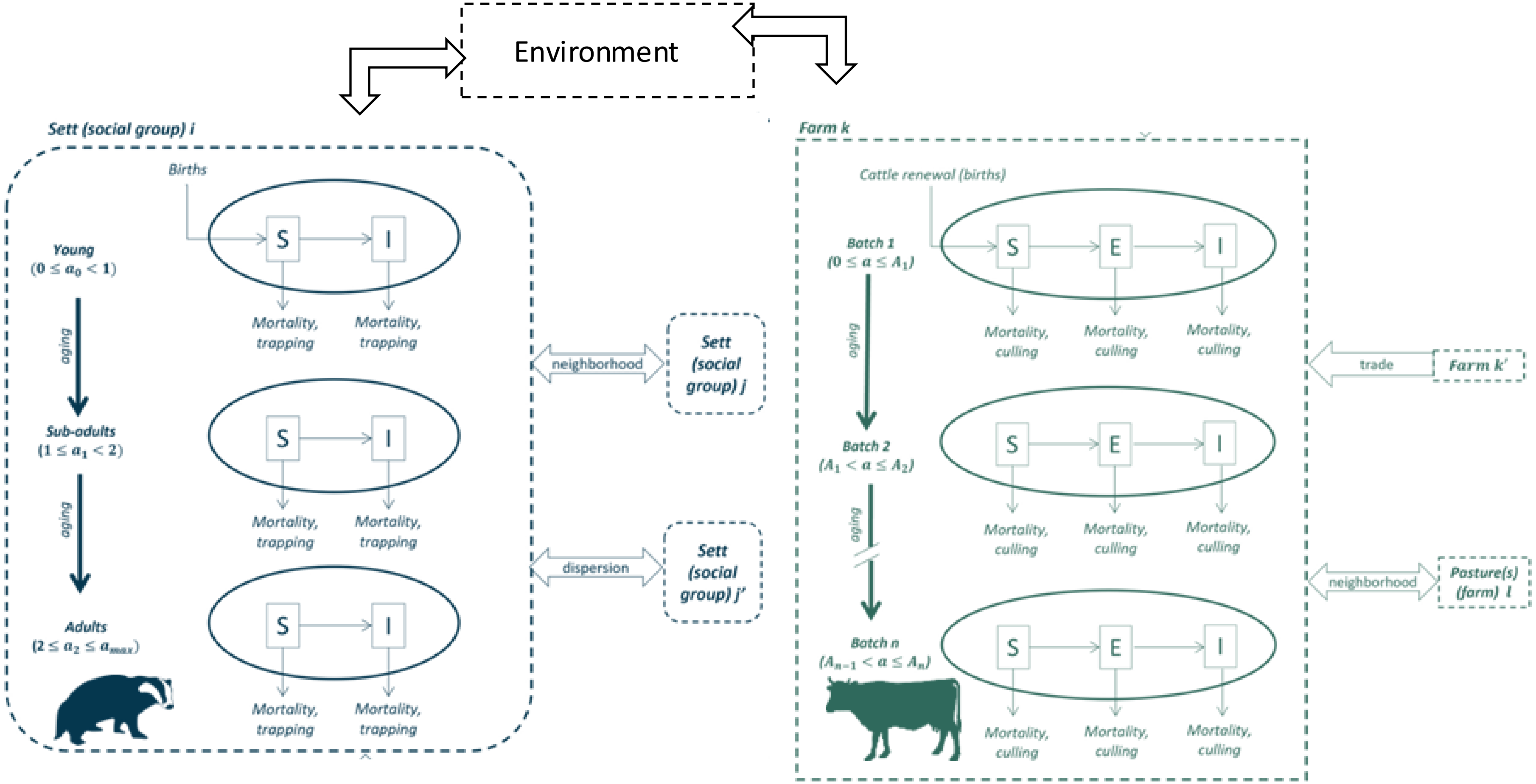


4. Reconstruction



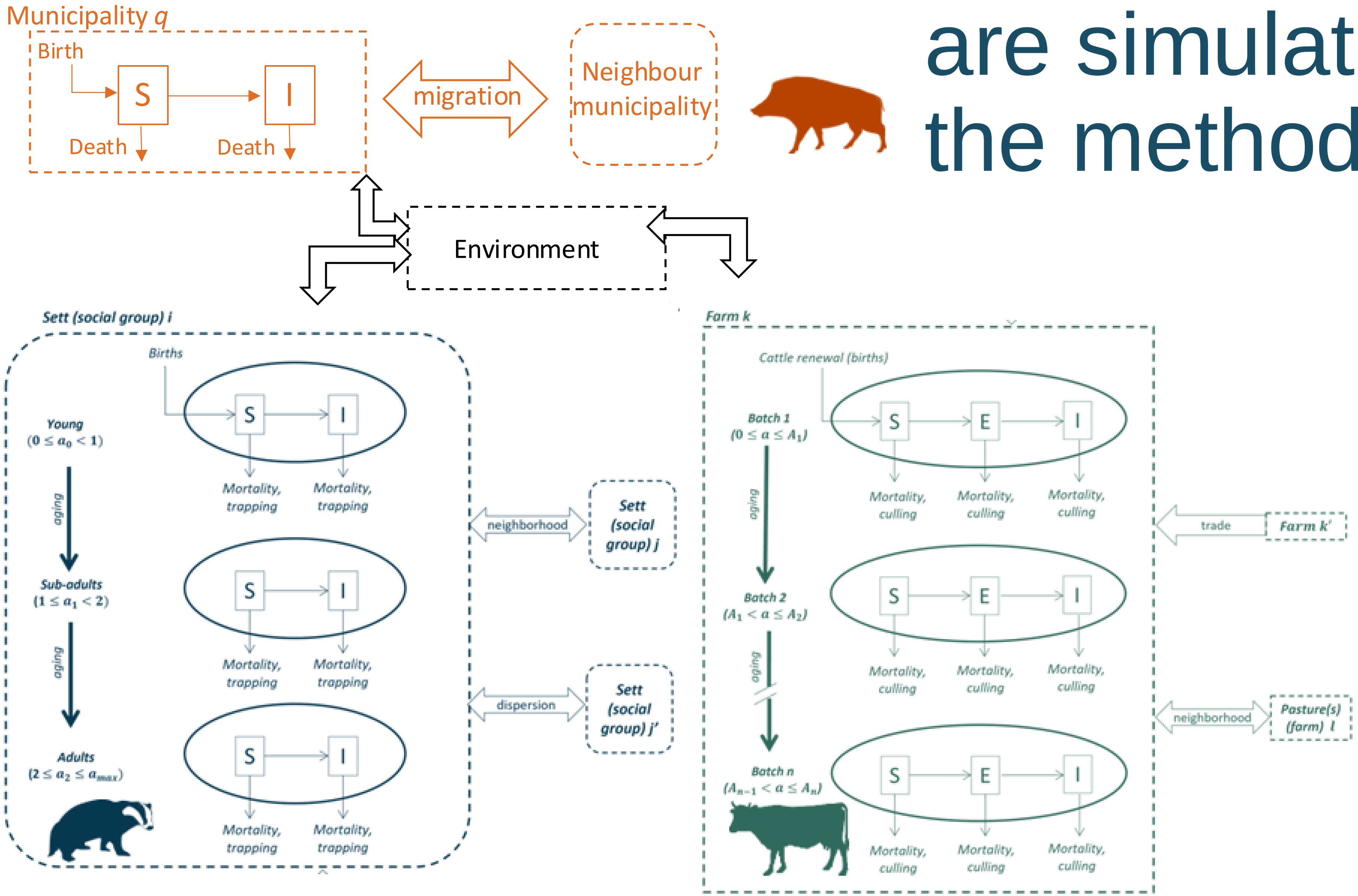
Time	Sequence	Species
T1	ACTT	Badger
T2	ACTT	Cattle
T3	ACAT	Cattle
T4	AGTT	Badger
…		

30 transmission trees are simulated to test the methods



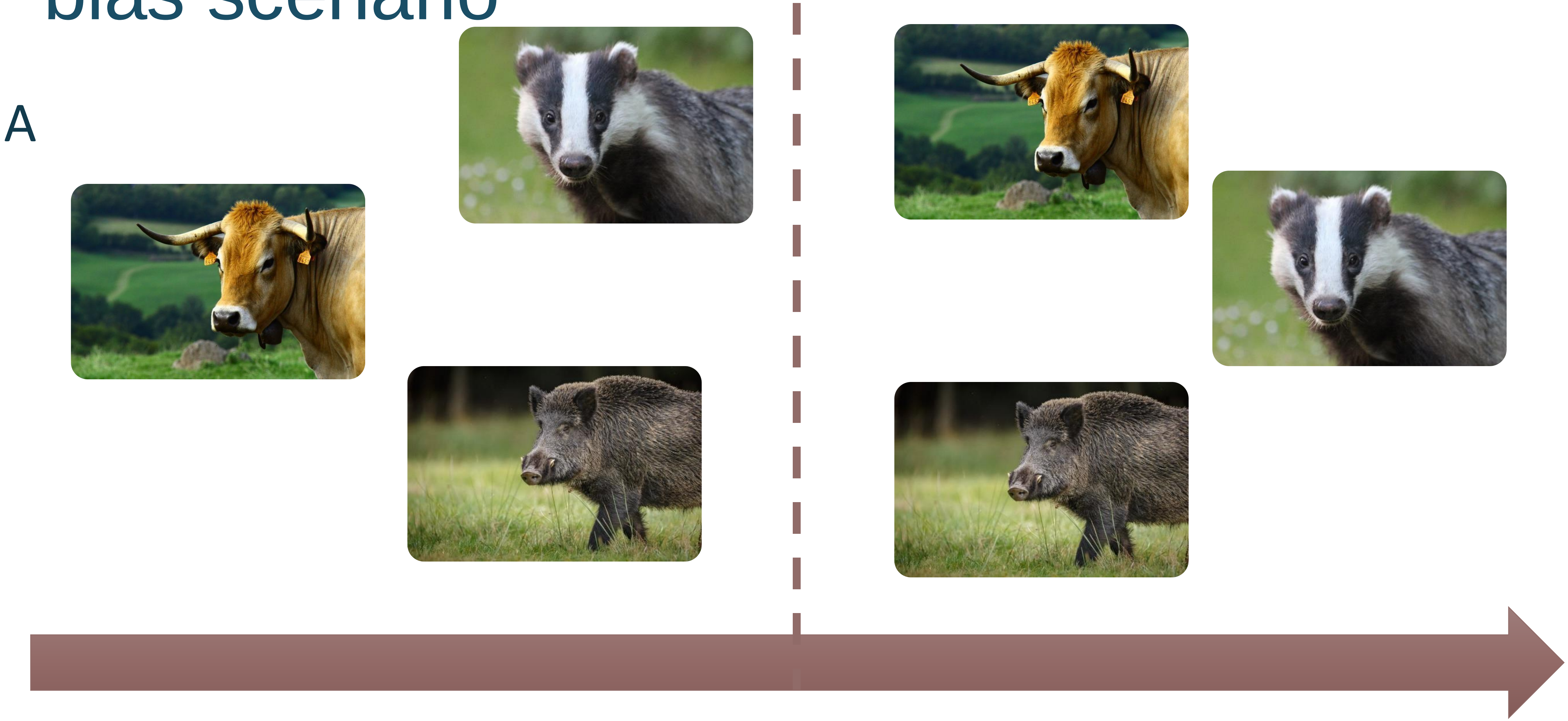
D'après Bouchez-Zacria *et al.*, 2023

30 transmission trees are simulated to test the methods



D'après Bouchez-Zacria *et al.*, 2023

30 trees are simulated for each sampling bias scenario



30 trees are simulated for each sampling bias scenario

T

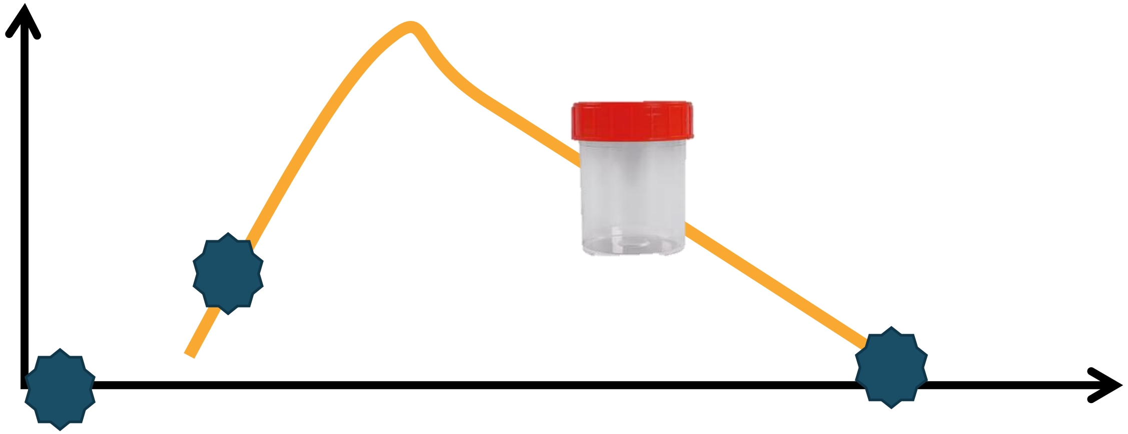
30 trees are simulated for each sampling bias scenario

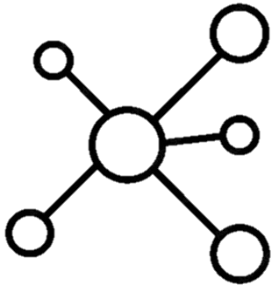
S_B

30 trees are simulated for each sampling bias scenario

T + S_B

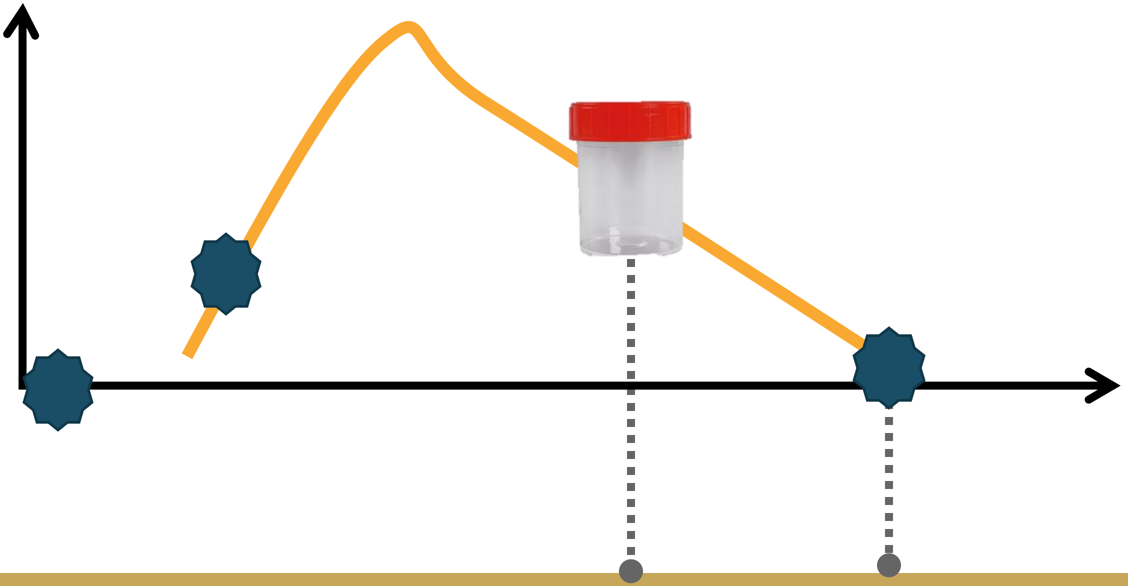
For bovine tuberculosis, according to the available observations…





Duault et al. (2022) Pathogens

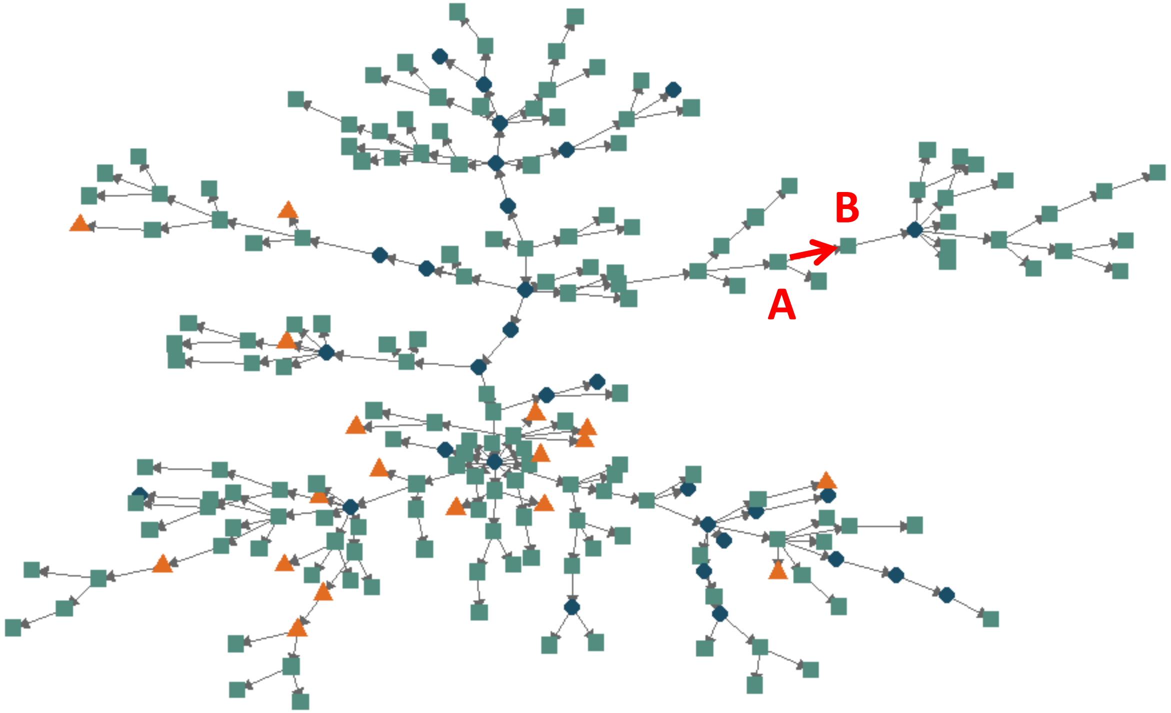
… 4/22 methods can be used



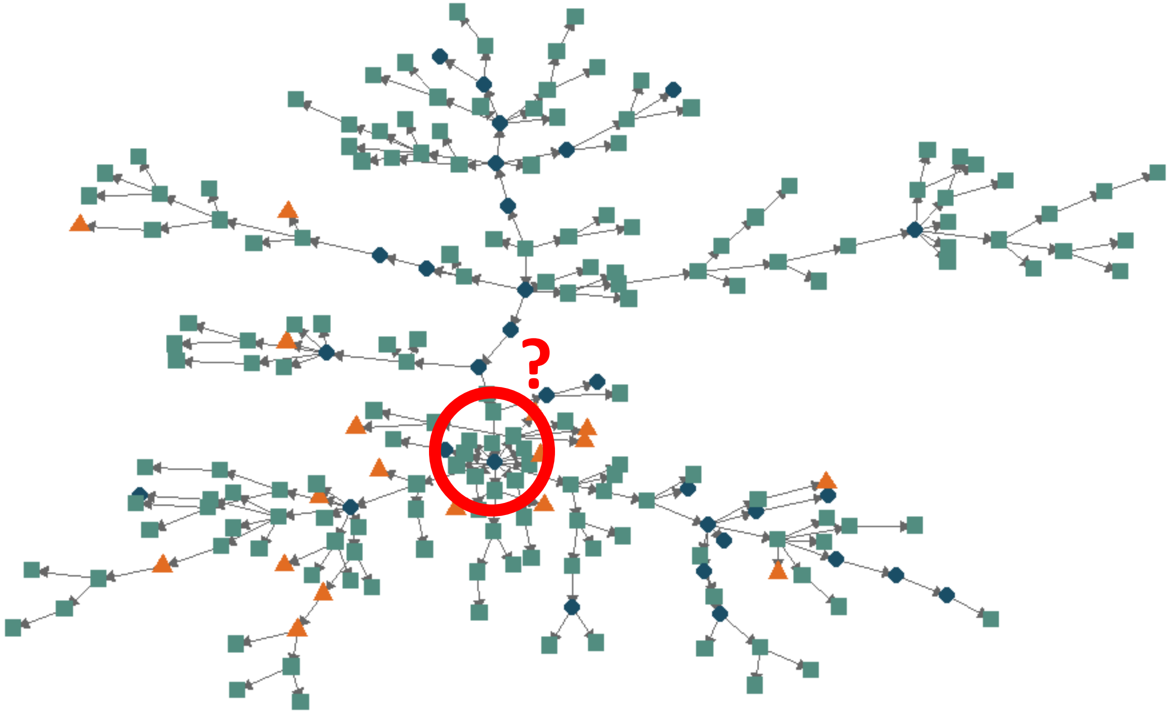
seqTrack
outbreaker2
TransPhylo
phybreak

Duault et al. (2022) Pathogens

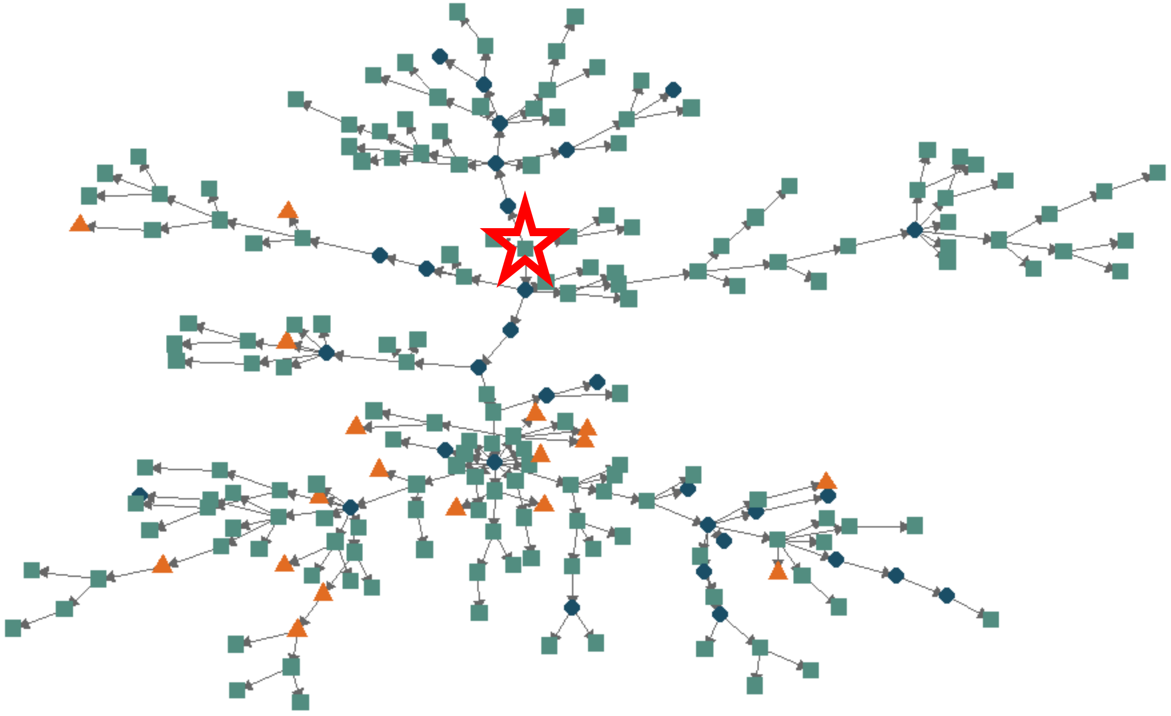
Indicators are defined to evaluate methods

Indicator	Definition	Method
Accuracy		seqTrack outbreaker2 Transphylo
Super-spreaders		
Index case host species		
Outbreak size		outbreaker2 Transphylo
Host species contribution		

Indicators are defined to evaluate methods

Indicator	Definition	Method
Accuracy		seqTrack outbreaker2 Transphylo
Super-spreaders	 < 10% of infected responsible for > 80% of transmission events	
Index case host species		
Outbreak size		outbreaker2 Transphylo
Host species contribution		

Indicators are defined to evaluate methods

Indicator	Definition	Method
Accuracy		seqTrack outbreaker2 Transphylo
Super-spreaders		
Index case host species		
Outbreak size		outbreaker2 Transphylo
Host species contribution		

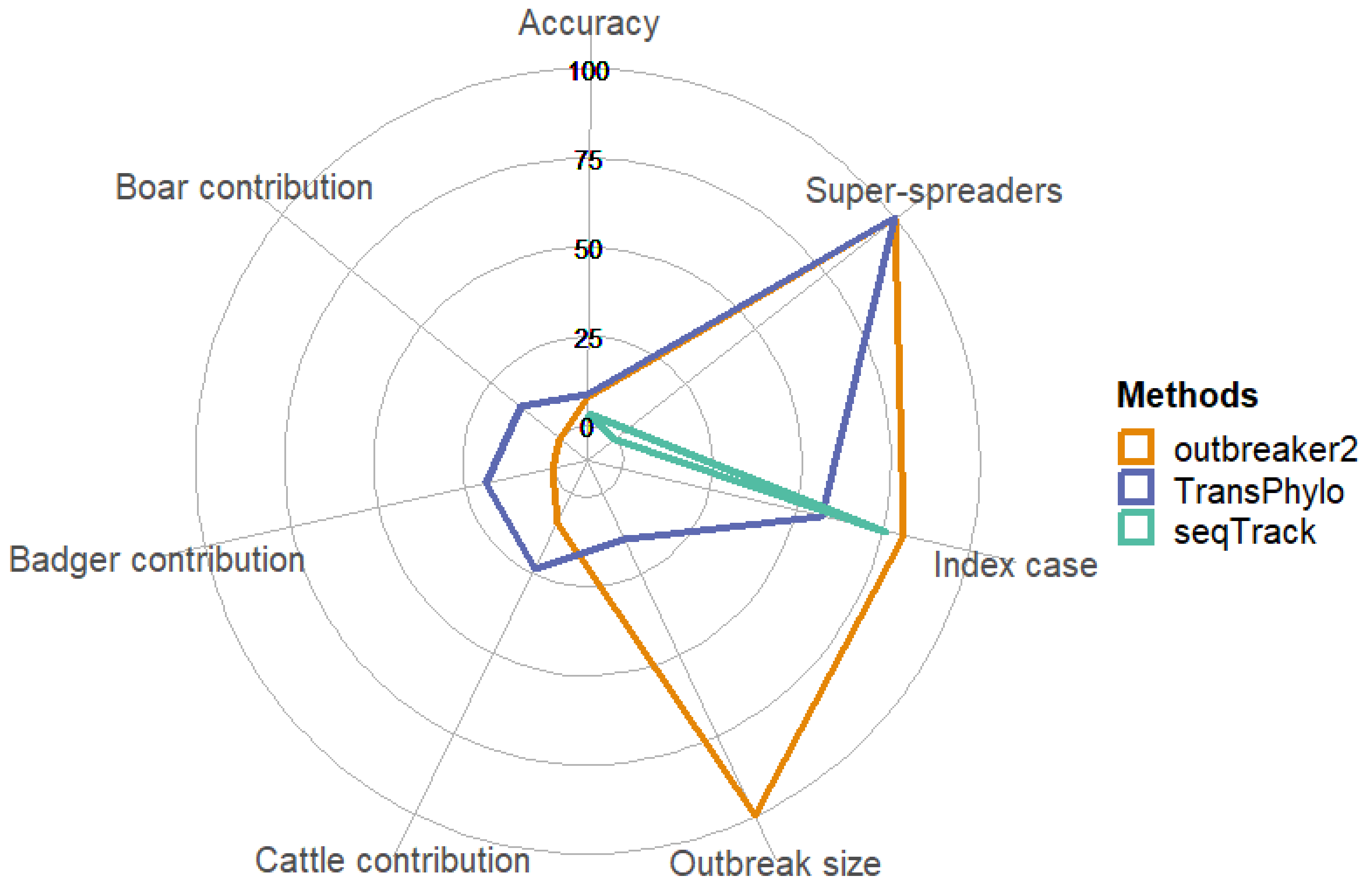
Indicators are defined to evaluate methods

Indicator	Definition	Method
Accuracy		seqTrack outbreaker2 Transphylo
Super-spreaders		
Index case host species		
Outbreak size	In reconstructed trees : $\frac{no\ of\ sampled\ hosts}{median\ sampling\ proportion}$	outbreaker2 Transphylo
Host species contribution		

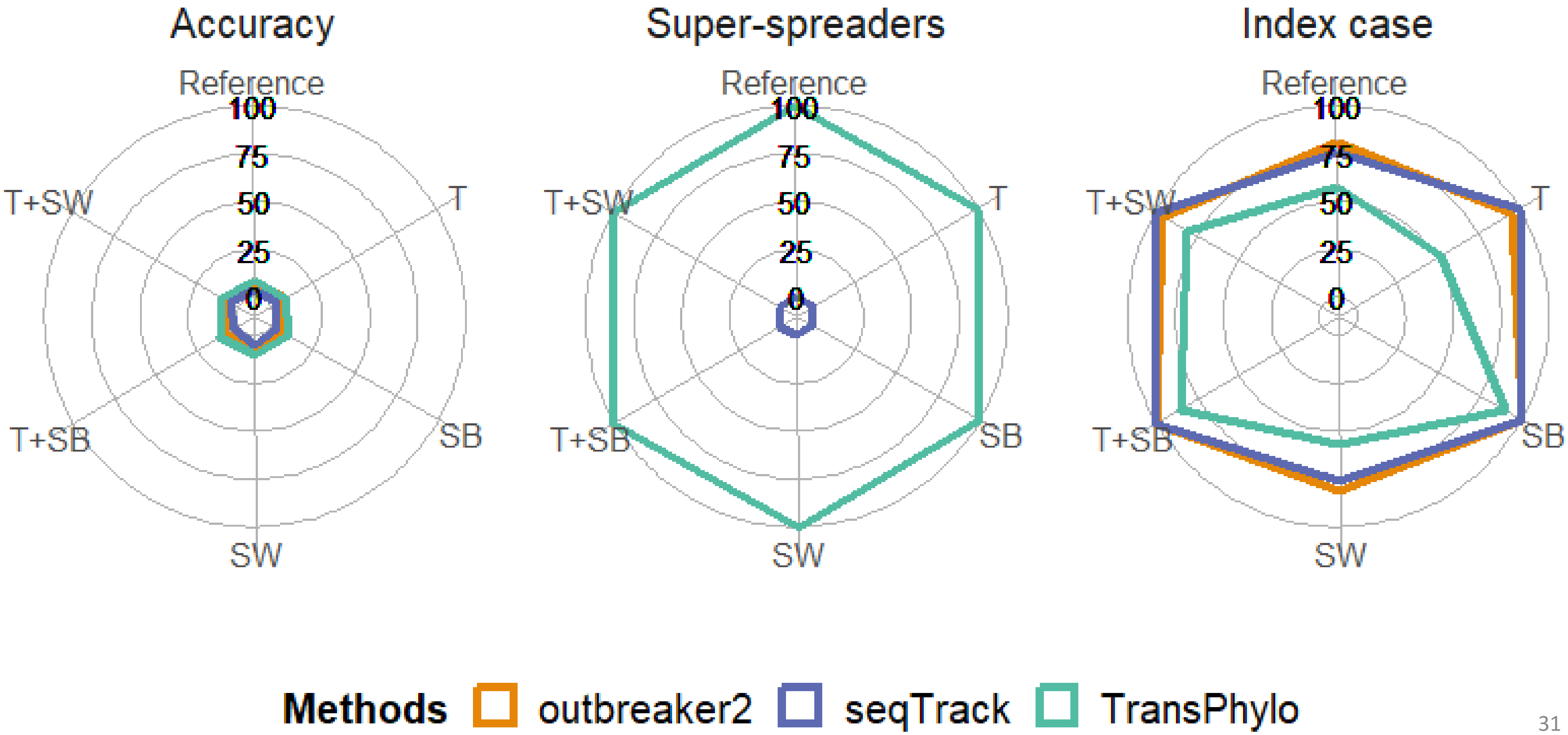
Indicators are defined to evaluate methods

Indicator	Definition	Method
Accuracy		seqTrack outbreaker2 Transphylo
Super-spreaders		
Index case host species		
Outbreak size		outbreaker2 Transphylo
Host species contribution	In reconstructed trees : $\left(\frac{no\ of\ transmission\ events\ between\ sampled\ hosts}{median\ of\ sampling\ proportion}\right)_{species}$	

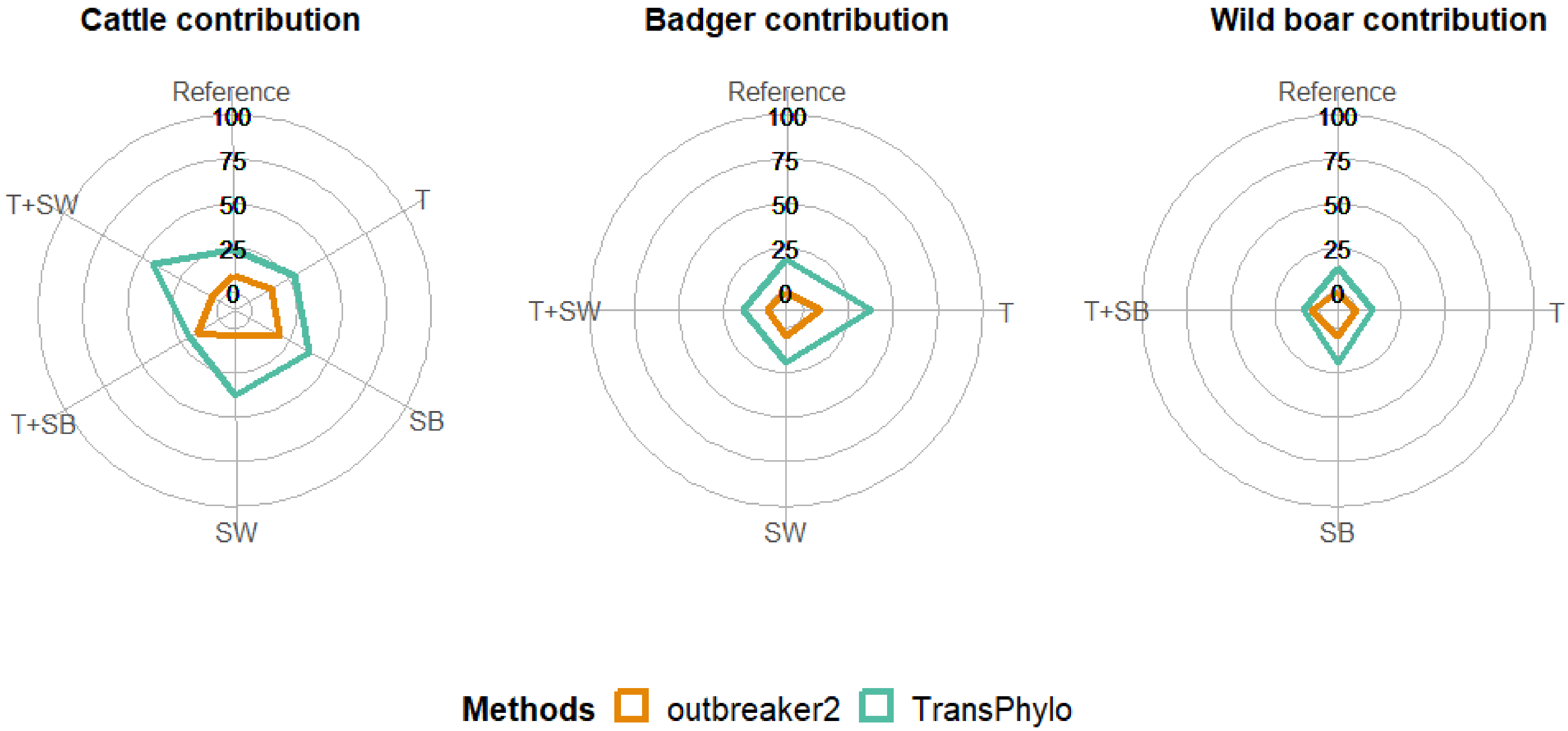
Overall, the performance of the three methods is mediocre with the reference scenario



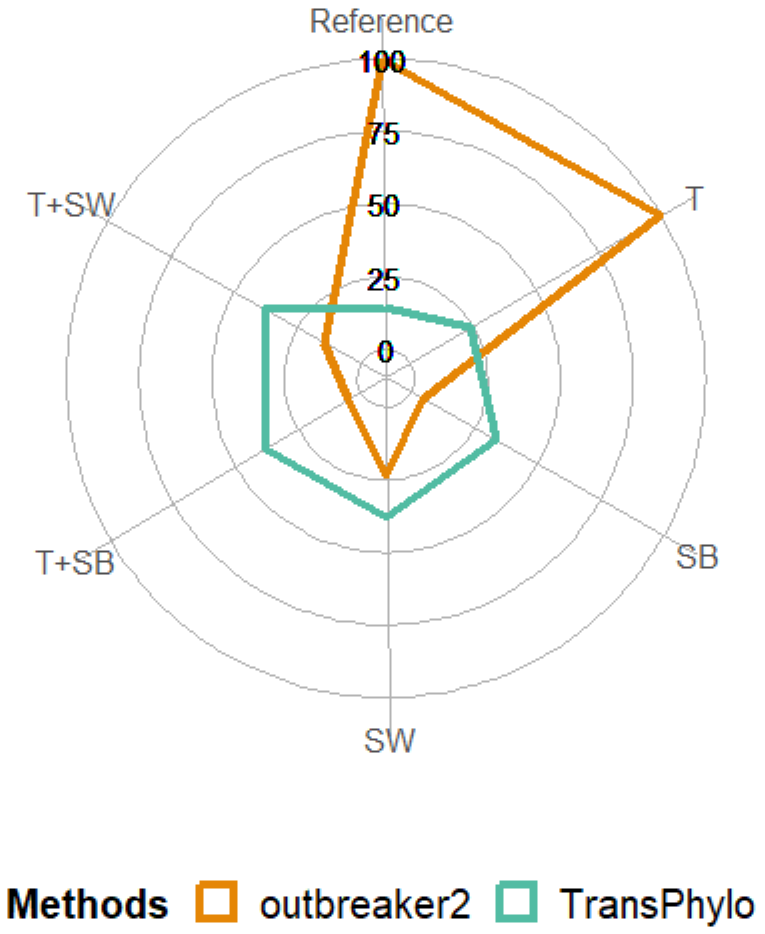
Sampling biases have little effect on accuracy or the index case…



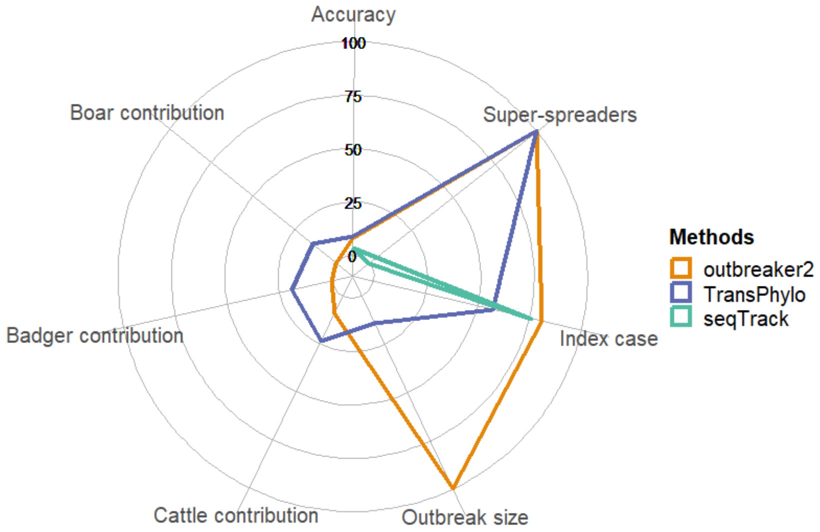
… and host species contribution



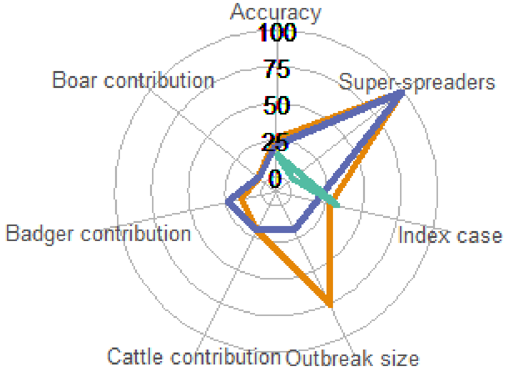
outbreaker2's good performance for outbreak size deteriorates rapidly with species bias



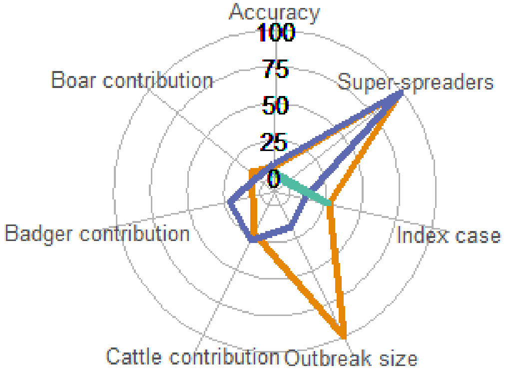
In other epidemiological contexts, the results are no better



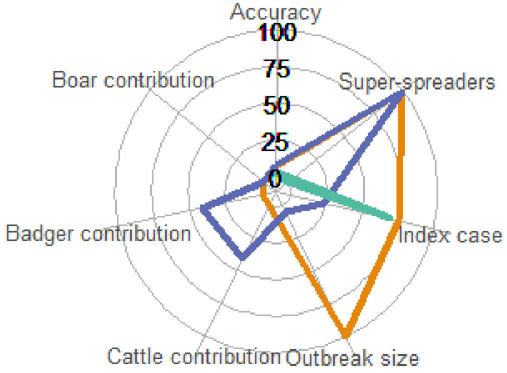
Mutation rate x10



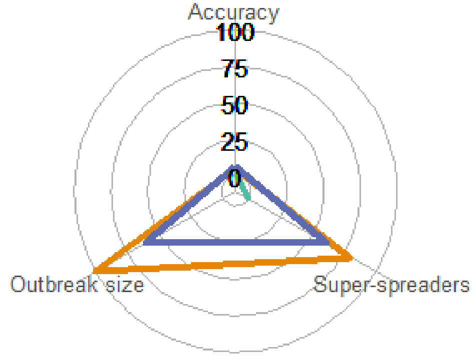
Badger as Index case



Boars as dead-end hosts



Single-host system



What question do we expect to answer ?

Indicator	
Accuracy	Who-infected-who ?
Super-spreaders	
Index case host species	Can we have a correct picture of the outbreak ?
Outbreak size	
Host species contribution	

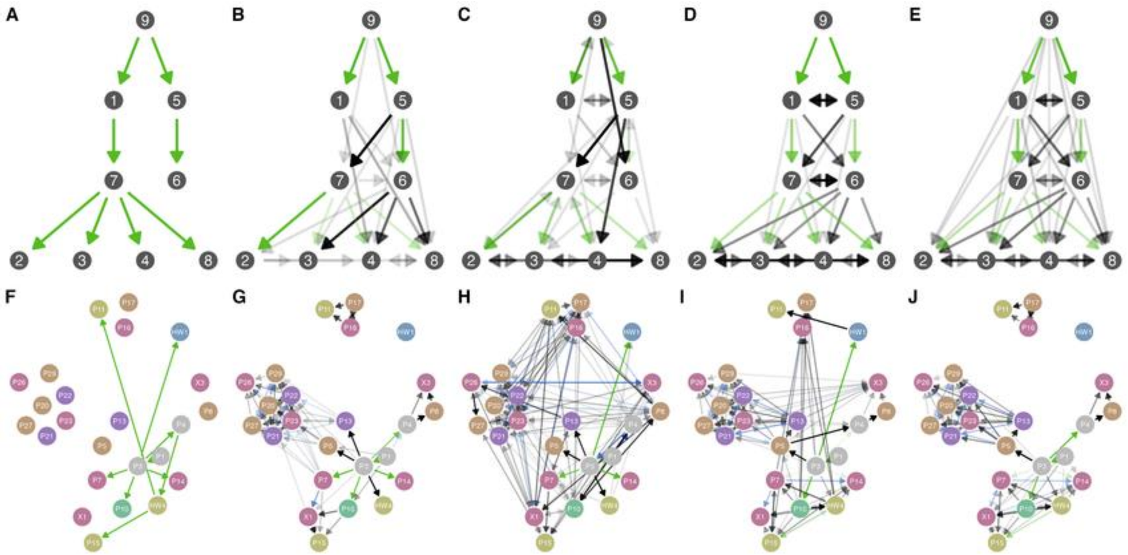
How can we improve the results ?

- Test simultaneous phylogenetic method
  - ➢ phybreak ?

MolEvolEpid/
biophybreak

- Improve convergence for transPhylo

- Increase the level of information in the data
  - ➢ Contact data ?



Thank you

EpiMIM

LNR Tuberculose animale

Hélène DUAULT

Benoit DURAND

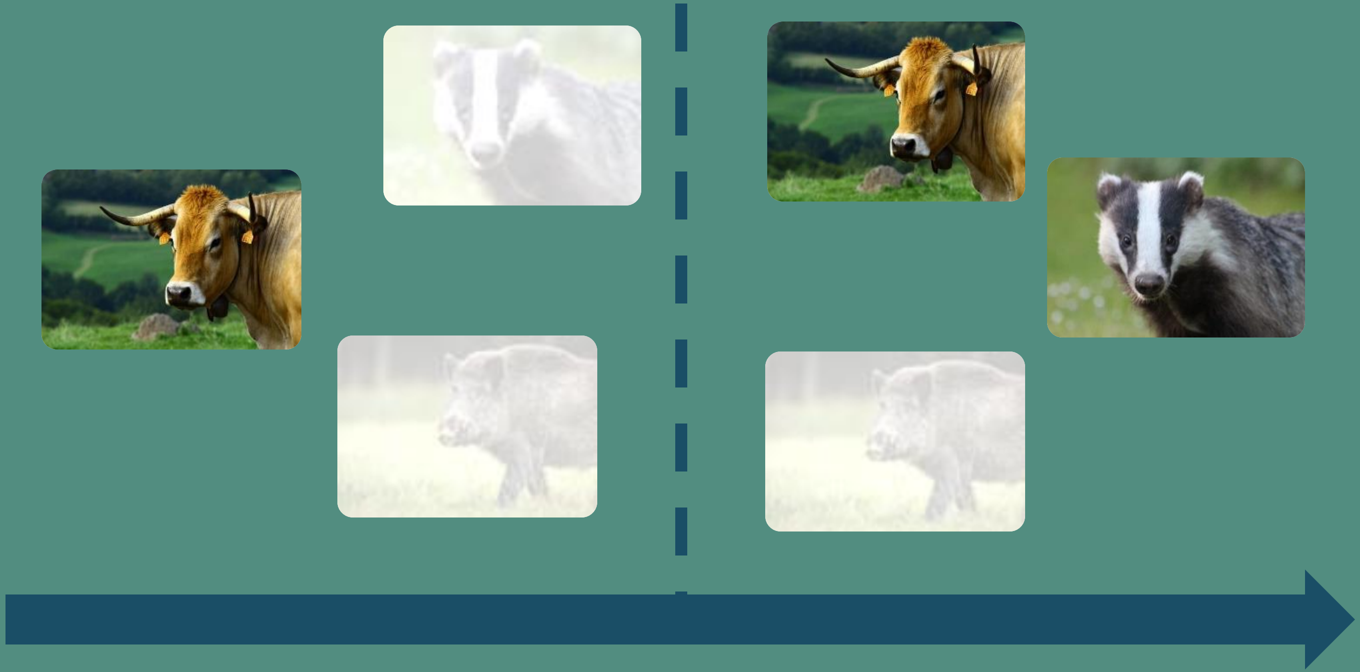
Gabriela MODENESI

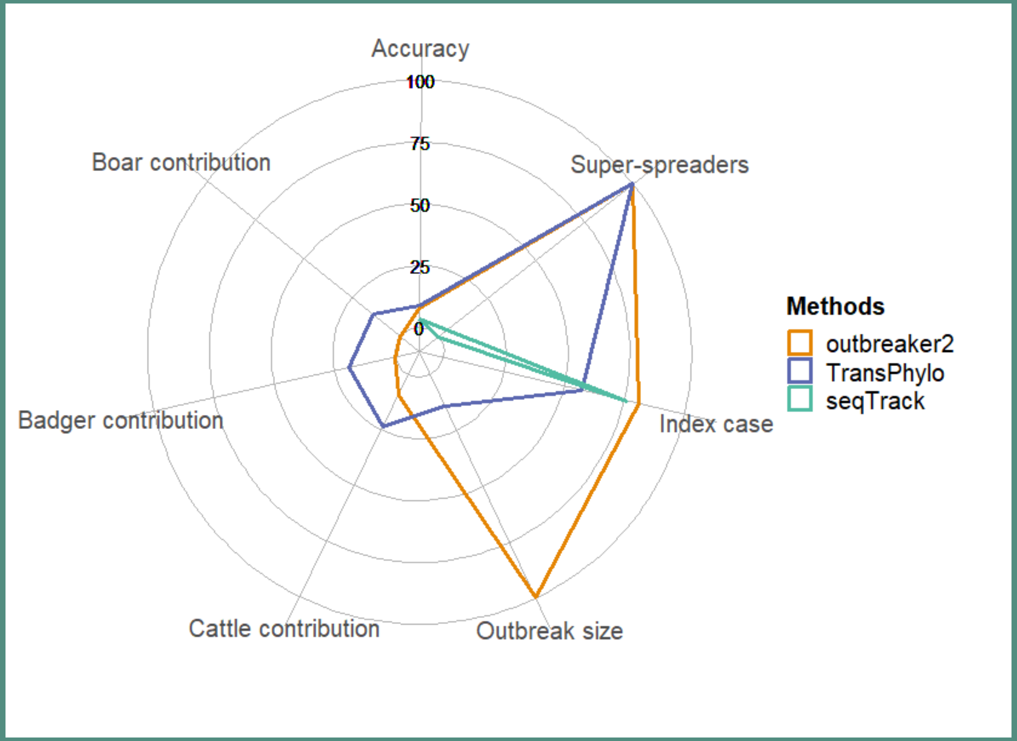
Aurélie COURCOUL

Maria-Laura BOSCHIROLI

Lorraine MICHELET

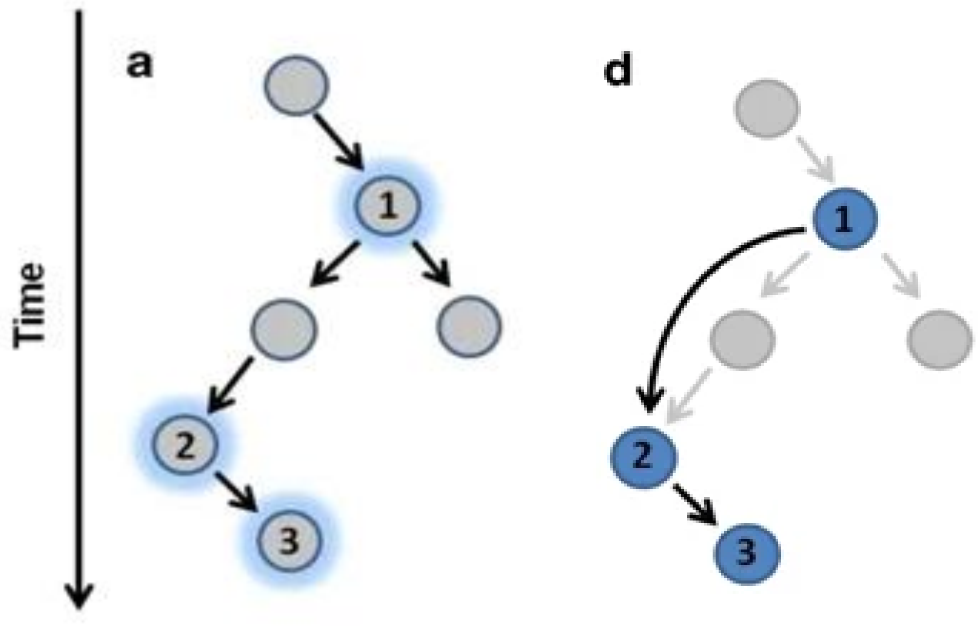
We need methods adapted to multi-host systems with sampling bias





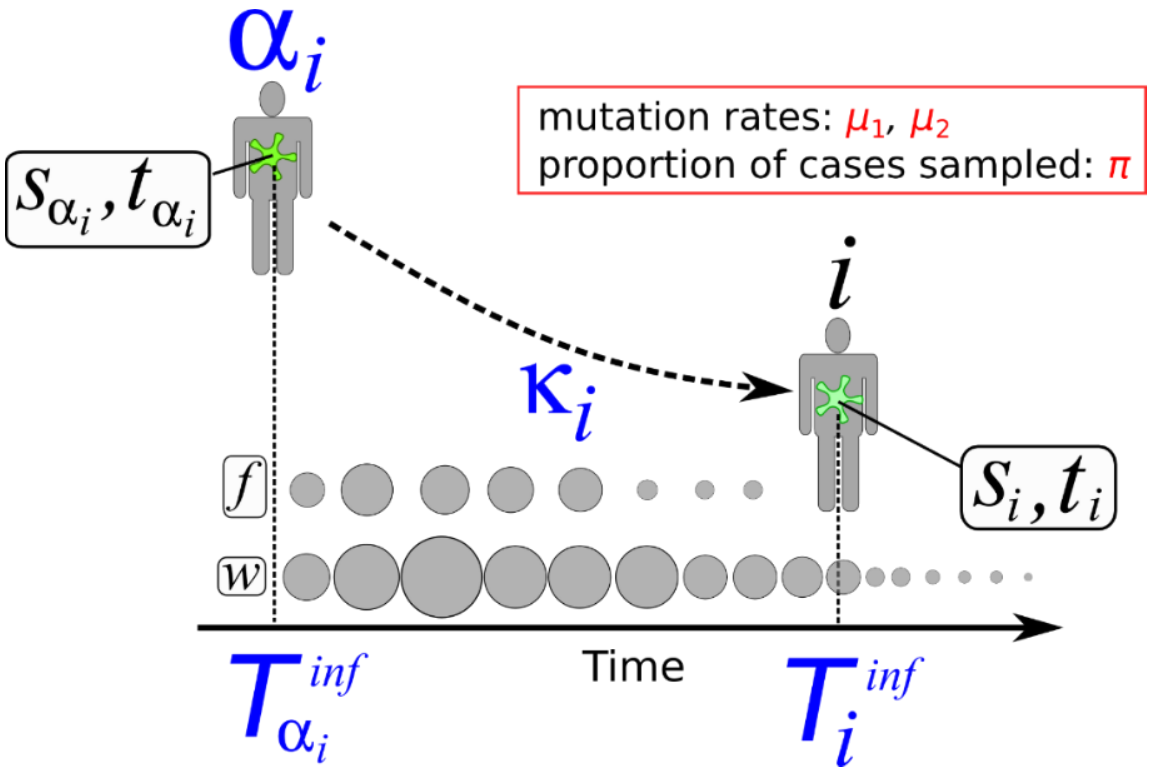
Non-phylogenetic methods for studying bTB

SeqTrack



Jombart et al. 2011, Heredity

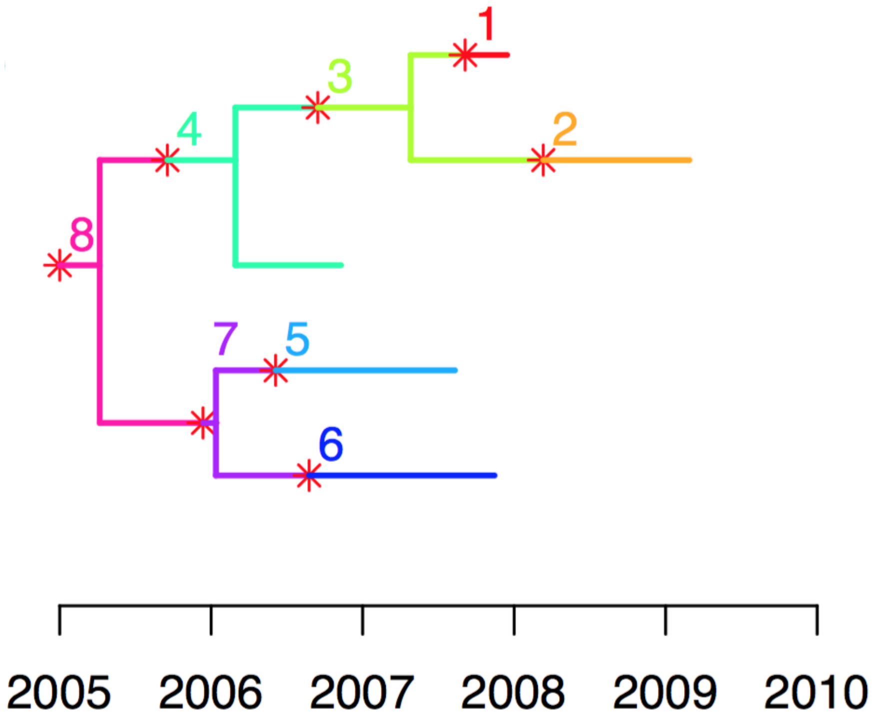
Outbreaker2



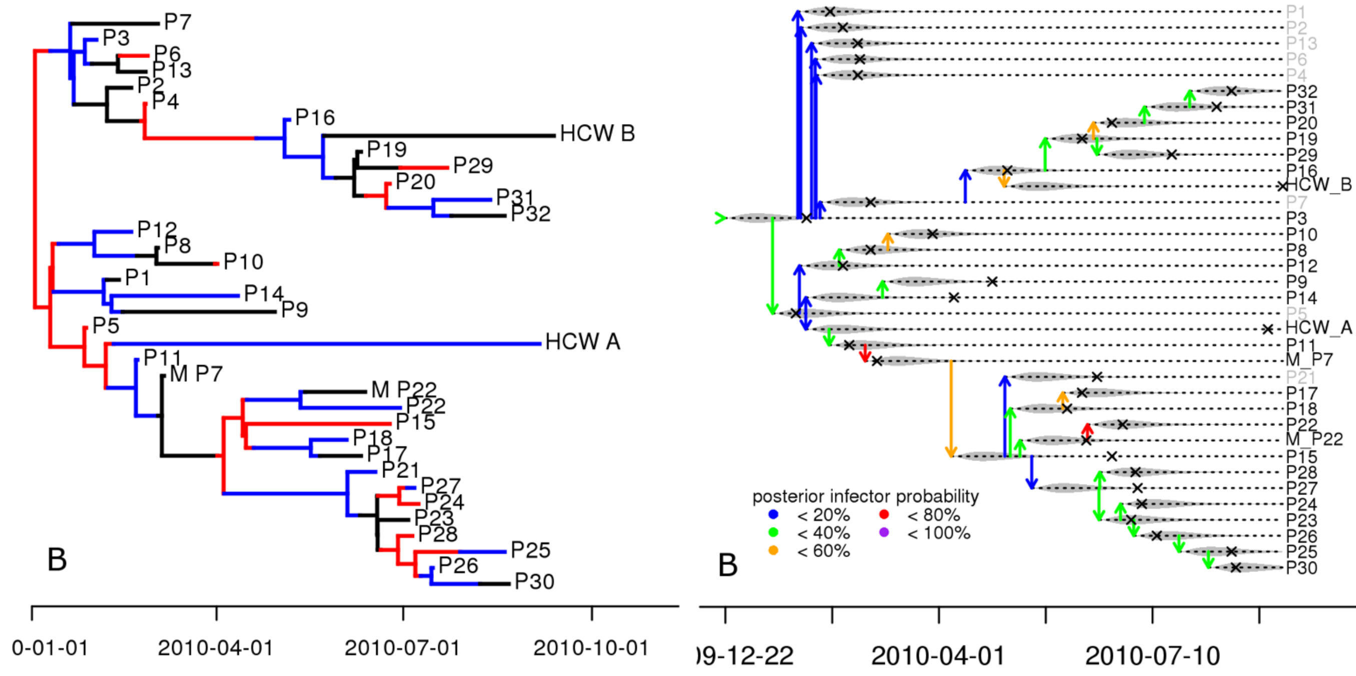
Campbell et al. 2019, BMC Bioinformatics

Phylogenetic methods for studying bTB

Sequential : TransPhylo



Simultaneous : phybreak



Didelot et al. 2017, Molecular Biology and Evolution

Klinkenberg et al. 2017, PLoS CB