Untangling the spread of African Swine Fever virus using whole-genome sequencing
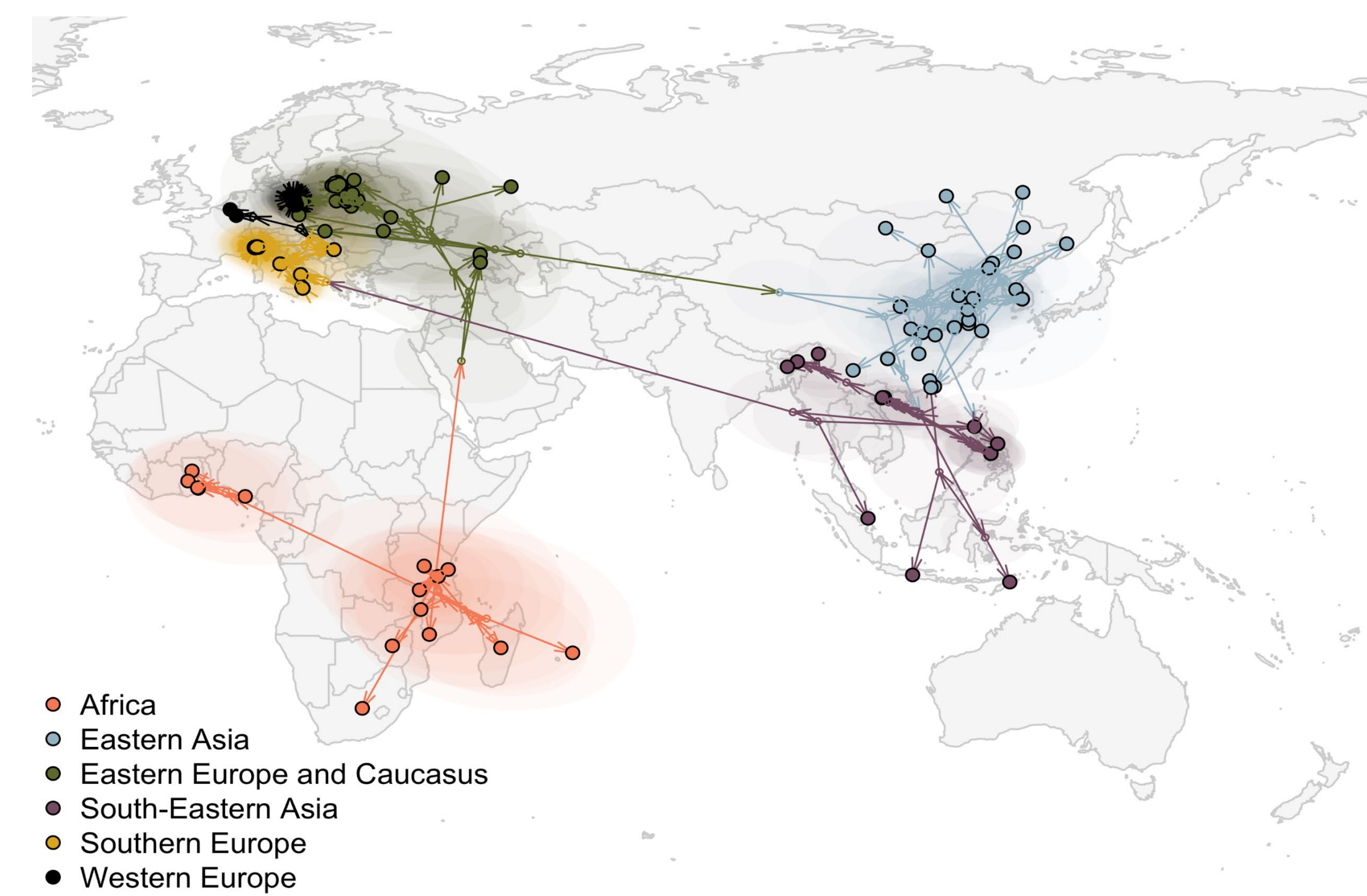
Gianluigi Rossi[1,2*], Jake Graham[1], Roberta Biccheri[3], Cesare Cammà[4], Carmen Iscaro[3], Maurilia Marcacci[4], Claudia Torresi[3], Francesco Feliziani[3], Monica Giammarioli[3], **Samantha J. Lycett**[1,2]

1. University of Edinburgh, Scotland (*g.rossi@ed.ac.uk);
2. EPIC Scotland; 3. CEREP, IZSUM, Italy; 4. GENPAT, IZS Teramo, Italy

BACKGROUND – African Swine Fever (ASF) is a hemorrhagic disease of suids that is currently threatening the pig industry worldwide. The current panzootic of ASF virus (ASFV) genotype II was first detected in Georgia in 2007 and has now spread to several countries across Eastern Asia, Europe and the Caribbean. The disease was detected in the European Union (2nd largest pigmeat producer, 1st exporter) in 2014, and since then it spread to many countries including major pig producers like Poland, Germany and Italy.

OBJECTIVE – **Understand the spatio-temporal dynamic spread of African Swine Fever globally, with particular focus on the four outbreaks in Italy, using the virus' whole-genome sequences (WGS) and their metadata.**

Region	# seqs	Countries	Oldest sequence	Youngest sequence
Africa (AFR)	15	Ghana (4), Madagascar, Malawi (2), Mauritius, Mozambique, Nigeria, South Africa, Tanzania (2), Zimbabwe	1998	2022
Eastern Asia (EAS)	42	China (30), Mongolia, South Korea (16), Eastern Russia (4)	2018	2022
Eastern Europe and Caucasus EEUR	31	Czech Republic, Georgia (2), Hungary, Lithuania, Moldova, Poland (12), Russia (12), Ukraine	2007	2019
South-Eastern Asia (SEAS)	23	India (4), Indonesia, Philippines (13), Singapore, Timor-Leste, Vietnam (3)	2019	2023
Southern Europe SEUR	85	Italy (80), Serbia (5)	2019	2023
Western Europe (WEUR)	25	Belgium (2), Germany (23)	2018	2021

Africa
Eastern Asia
Eastern Europe and Caucasus
South-Eastern Asia
Southern Europe
Western Europe

Figure 1: the spatial distribution of all 228 ASFV WGSs in the original ASFV alignment.
The region is based on United Nations geo-scheme, we aggregated regions when they had less than 5 sequences.

Table 1: Summary of the 221 sequences used in the study, after .

METHODS – We used an alignment (188,125 bp) which included **228 ASFV WGSs**. This included:
82 Italian WGSs complete with metadata (*sample date*, *coordinates* and *host species*);
146 WGSs from **GenBank,** including **Africa, Asia** and **Europe** (in case of incomplete metadata, they were approximated to the best of our knowledge).
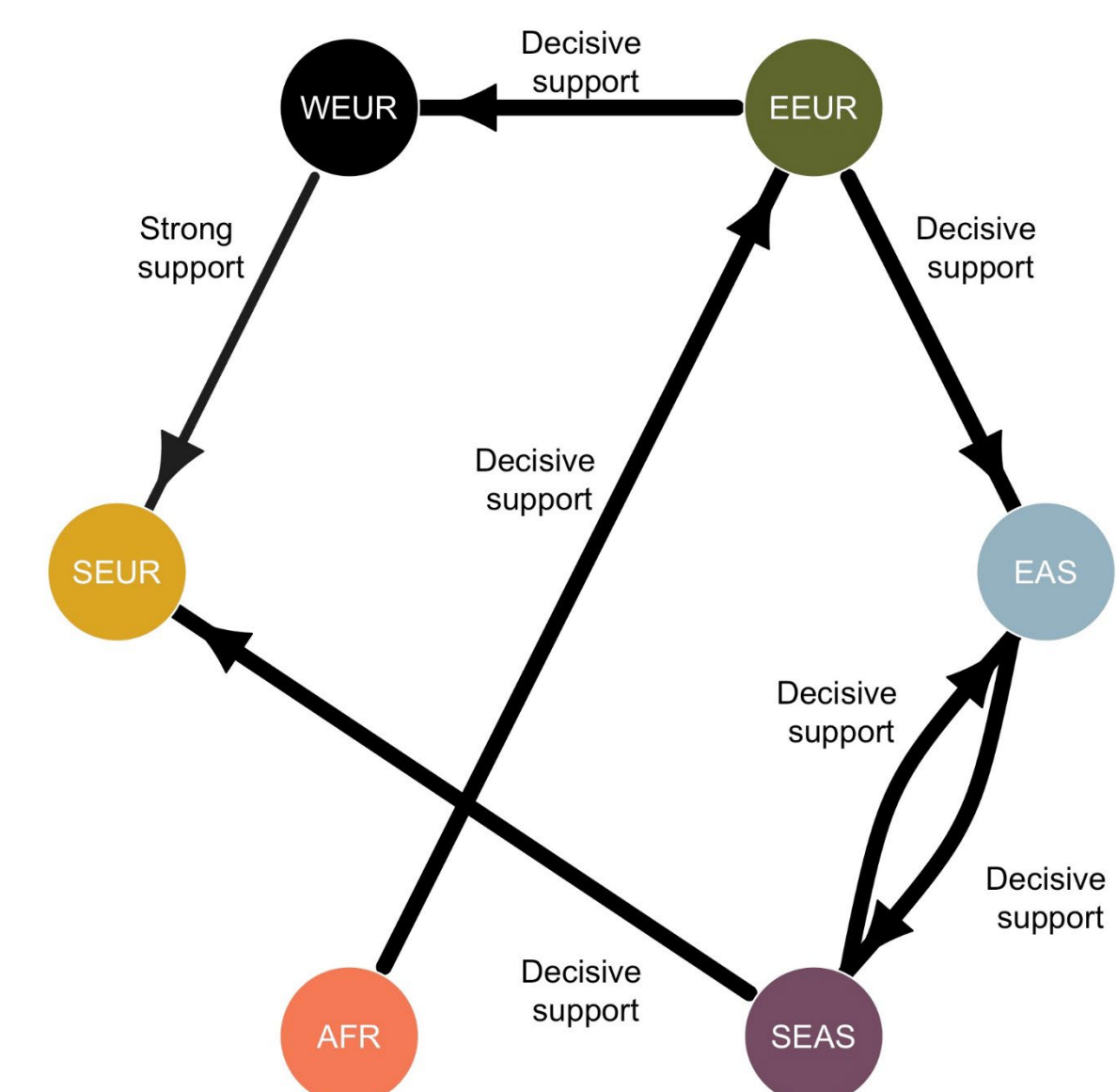We ran the alignment in **iqtree2**[1] (v2.3.6) to obtain a maximum likelihood tree, which we later used in **TempEst**[2] *(v1.5.3)* to check the genetic signal in the data. We removed 7 sequences:
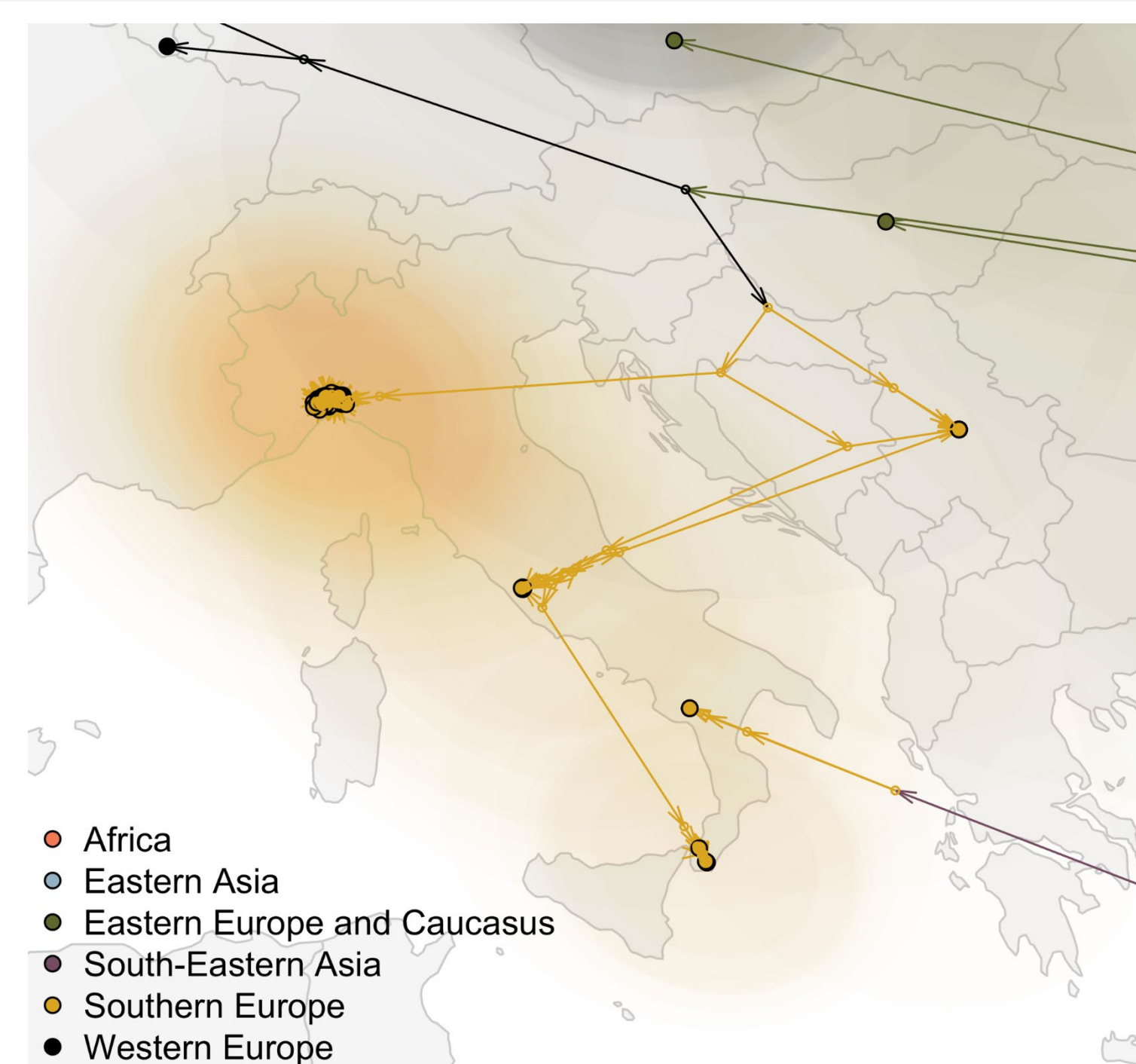- 4 were potential duplicates;
- LR881473 (Armenia), LS478113 (Estonia), and OM105587 (China) were removed because of low support and/or potential recombination.
We run the **221 sequences alignment** in **BEAST X**[3] (v1.10.5), to obtain the Maximum Clade Credibility tree. After trial runs, we set parameters as follows: **HKY** substitution model, **relaxed exponential** clock model, **SkyGrid**[4] tree model (27 years max, 54 bins).
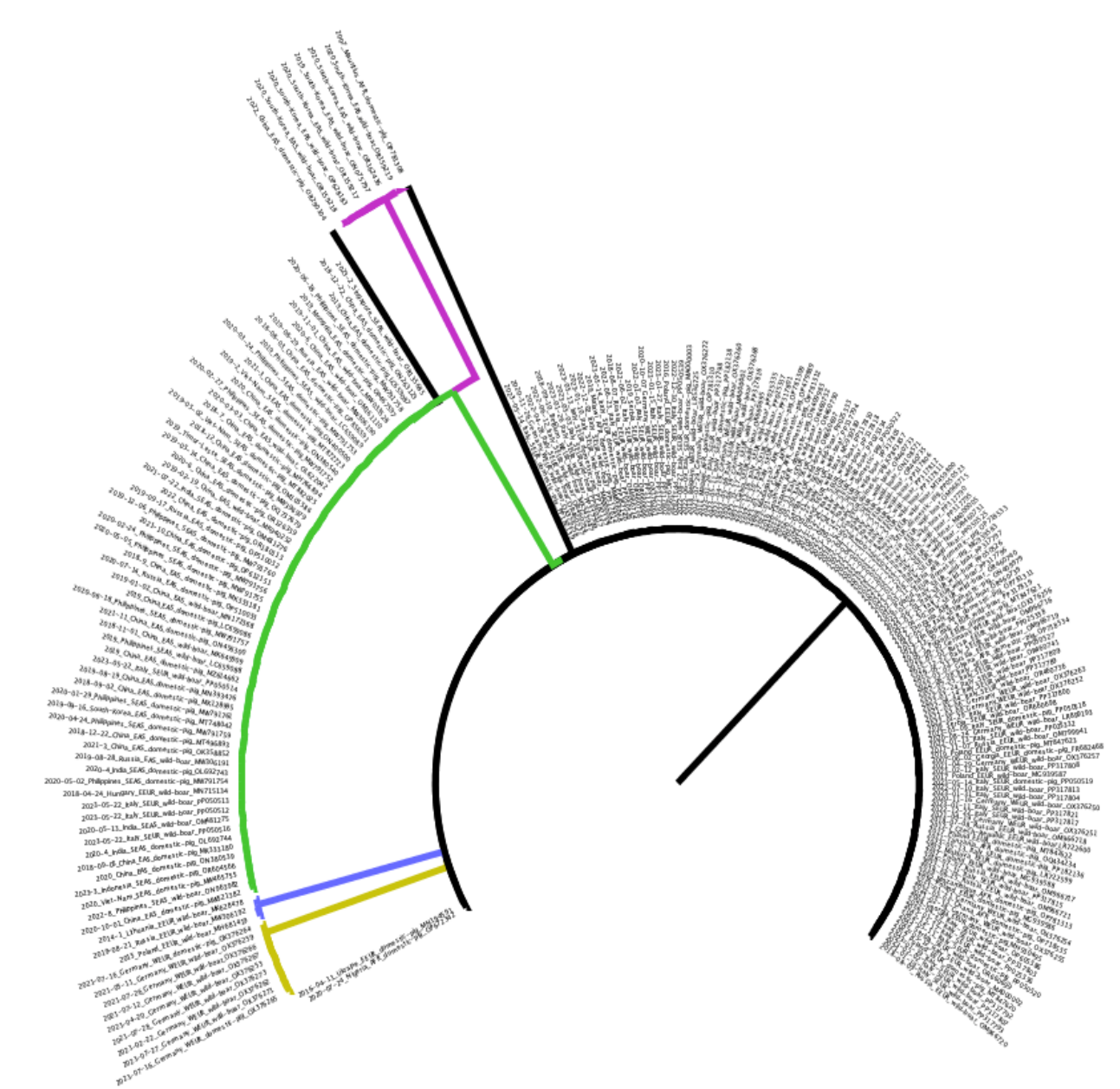We used two discrete traits in the main BEAST computation, *host species* (wild boar, n = 134 or domestic pig, n = 87) and *region*. We re-run BEAST with a sample of empirical trees, to assign the coordinates to the nodes, and calculate phylogeographic variables.

Figure 2: the spatial spread of ASFv Genotype II.
Arrows and tips colours represent different world regions.
Internal nodes do not reflect the actual position of the ancestor, as long-range transmissions were likely mediated by human activity (through movement of fomites or contaminated material).

Africa
Eastern Asia
Eastern Europe and Caucasus
South-Eastern Asia
Southern Europe
Western Europe

RESULTS:

MCC TREE – Fig. 3 reports the Maximum Clade Credibility (MCC) tree obtained. The median **clock rate estimate is 1.29x10^{-5}** (95th high-posterior density 1.15-1.75x10^{-5}), similar to the existing literature (1.10-1.47x10^{-5})[5,6]. The most recent common ancestor dates back to **November 1996** (HPD February '93 – December '98).

GLOBAL SPREAD – The tree is monophyletic, hinting to only one incursion of the genotype II outside Africa, where the genotype was endemic. After the early circulation in Eastern Europe, three macro-clades could be identified in the tree: two in Europe (Eastern Europe/Germany, and Southern Europe/Belgium) and one that from Eastern Europe branched off to China and the rest of Northeast and Southeast Asia. The transitions between regions are reported in Fig. 4, showing ASF spread from Africa to Caucasus and Eastern Europe, and from there to Western Europe and Eastern Asia. Evidence was less strong for the transition from Western to Southern Europe (Serbia and Italy). This could be due to a lack of sequences from Eastern European countries, which might be the source of introduction in some parts of Western and Southern Europe.

ASFV INTRODUCTIONS IN ITALY – The four Italian outbreaks are in the North-West, Lazio (Rome municipality), Campania, and Calabria (Fig. 5). Our results suggest there were three different introductions: the first in the North-West estimated in May 2021, from either Balkans or Eastern Europe (both not definitively supported). The second introduction was in Rome in September 2021, likely from the Balkans area. It further spread to Calabria in January 2023. The third and final introduction was from South-Eastern Asia to Campania, in February 2023 (Fig.3, 4, 5). This result was confirmed by the analysis of the **MGF 360-10L ge**ne. Fig. 6 shows a differentiation between the African-European lineage (black line) and the Asiatic lineages (green line), which include sequences from Campania (Italy, 4) and Hungary (1).

LIMITATIONS – The lack of sequences from some countries, especially in Eastern Europe, could have biased the rooting of the internal nodes. The MCC tree was sensible to the selected parameters in BEAST, which might indicate poor assembly in some of the GenBank sequences, or recombination.

References
1. Minh, et al. (2020) 'IQ-TREE 2: New models and efficient methods for phylogenetic inference in the genomic era'. Mol. Biol. Evol., 37:1530-1534.
2. Rambaut, et al. (2016) 'Exploring the temporal structure of heterochronous sequences using TempEst'. Virus Evol. 2, vew007.
3. Suchard, et al. (2018) 'Bayesian phylogenetic and phylodynamic data integration using BEAST 1.10'. Virus Evol. 4, vey016.
4. Gill et al. (2013) 'Improving Bayesian Population Dynamics Inference: A Coalescent-Based Model for Multiple Loci'. Mol. Biol. Evol. 30:713–724
5. Shen, et al. (2022) 'Bayesian Phylodynamic Analysis Reveals the Dispersal Patterns of African Swine Fever Virus'. Viruses 14, 5: 889.
6. Zhang, et al. (2023) 'Tracing the Origin of Genotype II African Swine Fever Virus in China by Genomic Epidemiology Analysis'. Transb. and Emer. Dis., 4820809.

Figure 3: the MCC time-scaled tree obtained with BEAST of 221 whole-genome sequences of ASFV. Branch colours identify different regions, while vertical lines represent different clades (solid) and subclades (dashed).

Figure 4: ASFV transition graph between regions. Decisive support: Bayes Factor > 100, Strong support: BF 10<31.6. Not strongly supported transitions not shown.

Africa
Eastern Asia
Eastern Europe and Caucasus
South-Eastern Asia
Southern Europe
Western Europe

Figure 5: Three separate introductions of ASFV in Italy: two from South or Eastern Europe and one from South-Eastern Asia.

Figure 6: Phylogenetic tree of the ASFV MGF360-10L gene.

TAKE HOME MESSAGE – Our results showed many potentially human mediated ASFV transmissions between geographic areas, as they are beyond the movement range of wild boars or other wild carriers of the disease. Italy was likely subject to three separate ASFV introductions: the successful establishment of the disease might be due to the presence of wild boars in peri-urban areas, where they can easily get in contact with potentially contaminated human waste. While the risk for Scotland and the UK might be lower due to small and highly localised feral pig population, authorities needs to be vigilant for potential introduction from contaminated pork products.