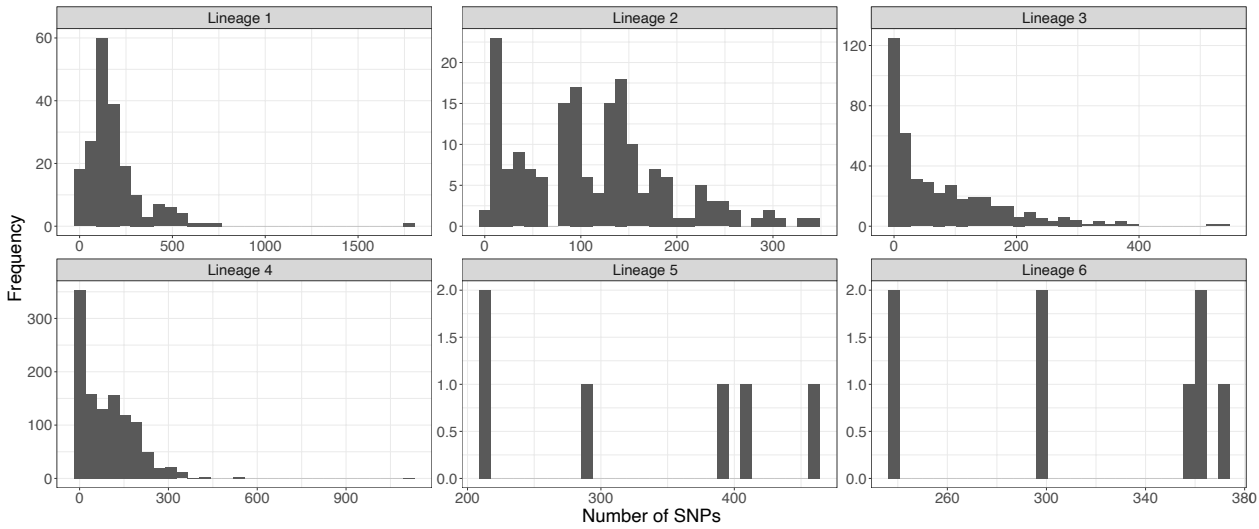
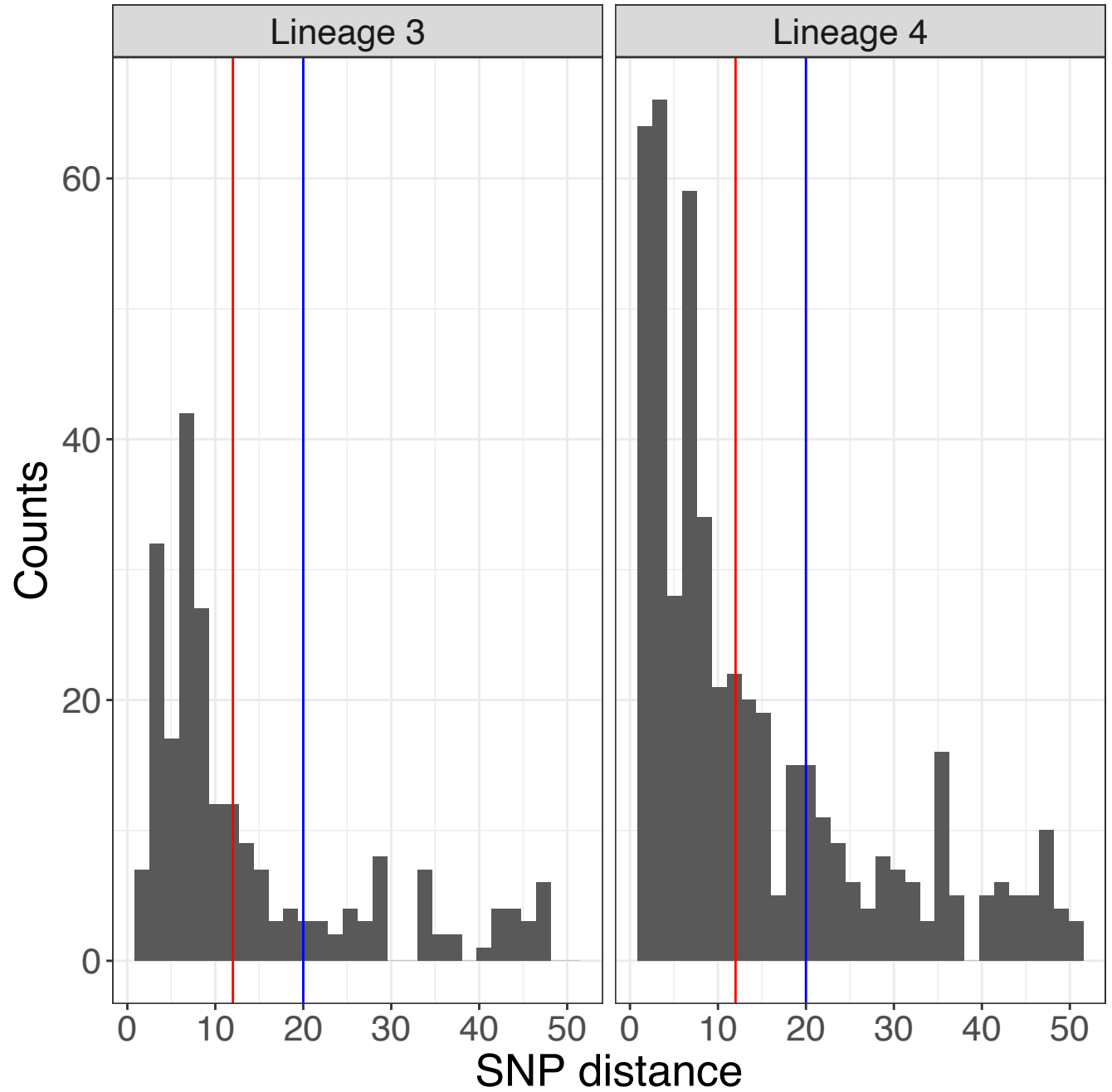


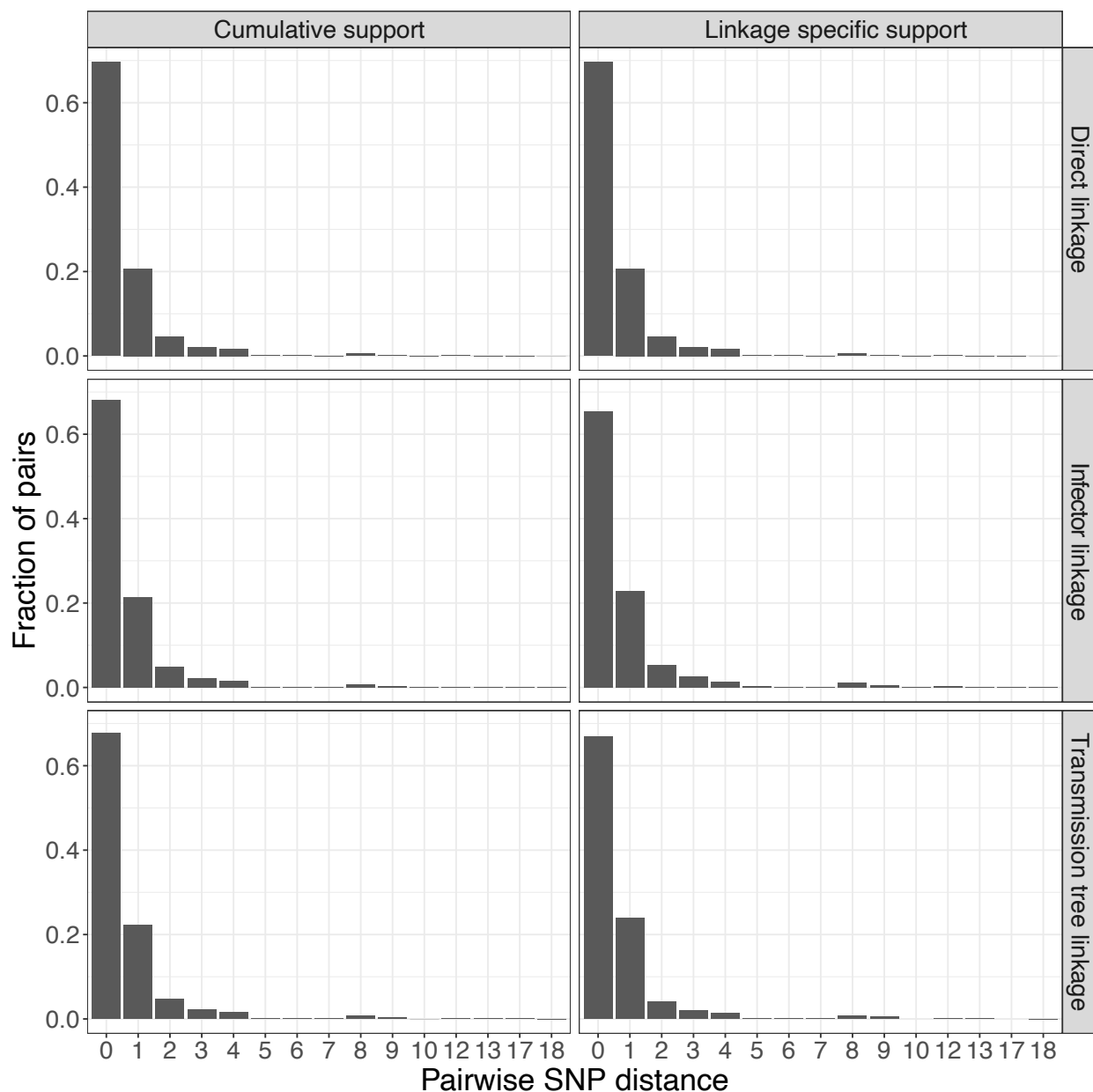
Supplementary Figure 1: Frequency of pairwise SNP distances between isolates in a main lineage. The minimum distance of an isolate to all other isolates in the same lineage is shown.



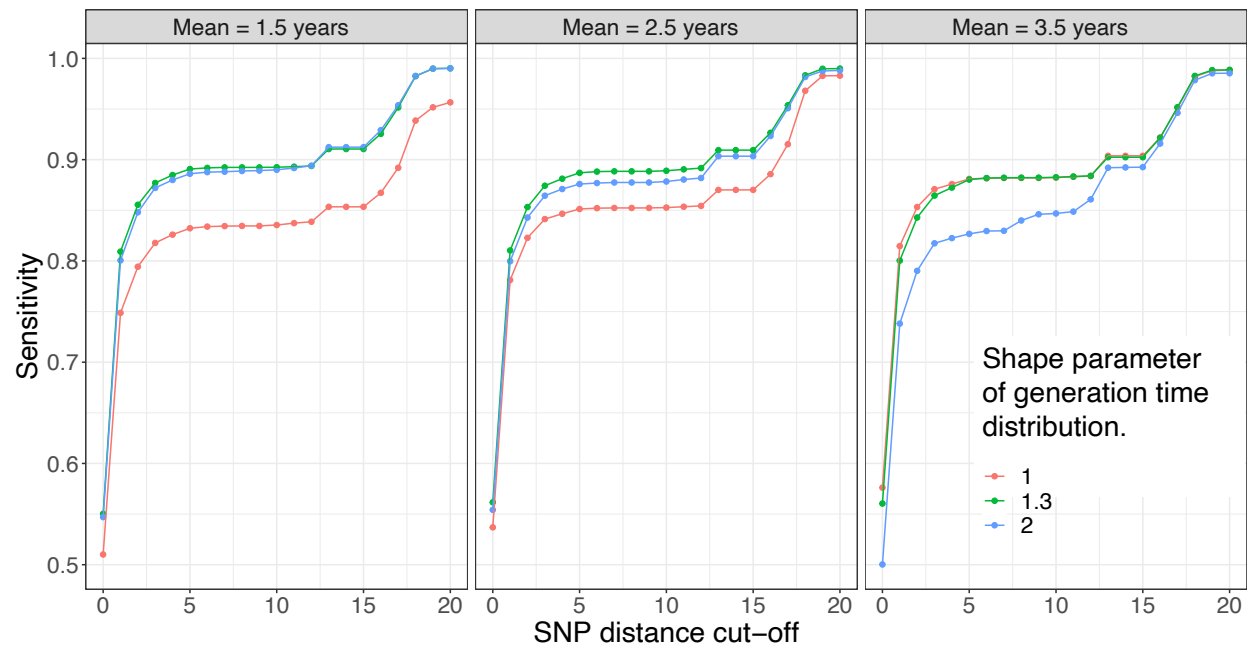
Supplementary Figure 2: Distribution of minimal SNP distances up to 50 SNPs for isolates in lineages 3 and 4 (45% of the isolates in these lineages). The red line indicates the 12 SNPs cut-off, found as a maximal cut-off in literature, and the blue line indicates the 20 SNPs cut-off.



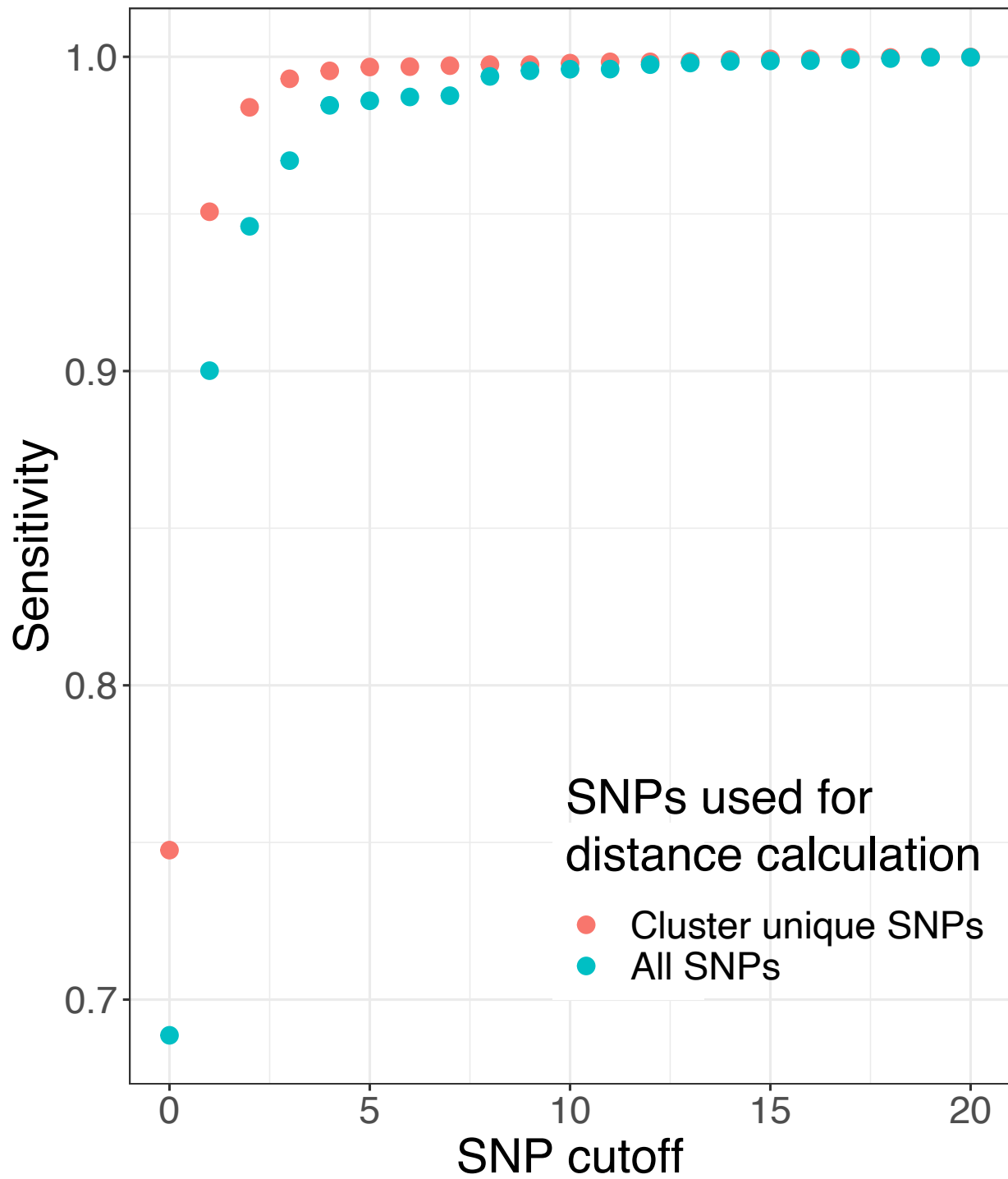
Supplementary Figure 3: Histograms of pair-wise SNP distances for the three sets of transmission links:
direct linkage, infector linkage, and transmission tree linkage. Cumulative: SNP distance distribution is calculated with cumulative support for the pairs. Infector linkage also includes direct linkage pairs, whereas transmission tree linkage includes all linkage types. Separate: SNP distance distribution is calculated with support for pairs to have a certain link. Only the support of pairs with the specific linkage is used, e.g. SNP distance distribution of infector linkage pairs includes only the support of pairs for infector linkage and not the support for direct linkage. Almost all pairs have support for all three types of linkage, and thus no difference in SNP distance distribution is found between cumulative and linkage-specific support.



14 **Supplementary Figure 4: Sensitivity analysis of generation time distribution.** Sensitivity scores are shown
 15 for all SNP cut-offs with inferred direct transmission events as gold standard. The mean and shape of the Gamma
 16 distribution are varied, with a mean of 2.5 and shape of 1.3 used in the main analysis.

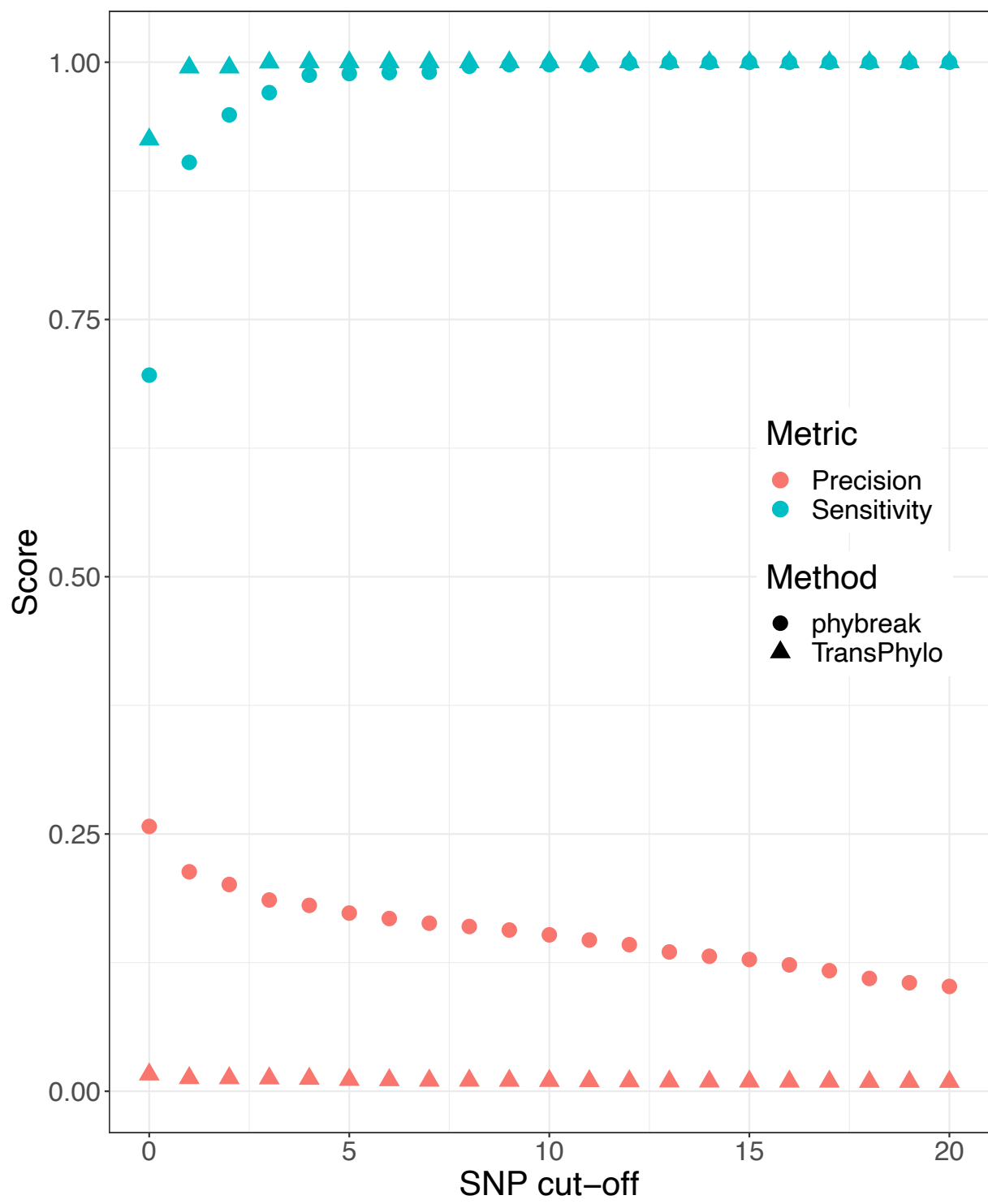


Supplementary Figure 5: Sensitivity analysis of cluster-unique SNPs. Sensitivity scores are shown for all SNP cut-offs with inferred direct transmission events as gold standard. Pair-wise distances between isolates are computed with only cluster-unique SNPs (in red) or all SNPs (in blue).



20 **Supplementary Figure 6: Comparison of metrics with different inference methods.** Precision (red) and
 21 sensitivity (blue) are computed for all SNP cut-offs using direct transmission links from phybreak (dots) or

TransPhylo (triangles) as gold standard.



23 **Supplementary Table 1: Characteristics of all patients with sequenced isolates.**

	All Cases (n = 2008)	Cases in lineages 3 and 4 (n = 1645)	Cases in clusters (n = 478)
Age (mean $\pm$ sd)	41 ($\pm$ 28)	40 ($\pm$ 19)	34 ($\pm$ 17)
Sex (Male)	62.7%	64.5%	70.3%
Geographical distribution^a			
Amsterdam	12.1%	12.0%	12.6%
Rotterdam	8.7%	8.3%	11.5%
The Hague	6.2%	6.1%	4.4%
Utrecht (city)	2.7%	2.6%	2.3%
North Netherlands	26.0%	26.9%	29.7%
Middle Netherlands	24.5%	23.8%	20.9%
South Netherlands	19.9%	20.2%	18.6%
Country of birth^b			
The Netherlands	17.1%	17.2%	16.5%
Europe ^c	11.3%	11.3%	7.95%
Africa	23.1%	23.3%	42.9%
Americas	4.6%	4.7%	6.3%
Eastern Mediterranean	23.5%	23.1%	23.2%
South-East Asia	11.4%	11.9%	1.5%
Western Pacific	6.1%	5.7%	0.4%
Ethnicity			
First generation migrant	77.6%	76.5%	75.3%
Second generation migrant	6.1%	6.8%	9.8%
Native Dutch	14.0%	14.4%	11.5%
Unknown	2.4%	2.3%	3.4%
Risk factors			

HIV (positive)	3.2%	2.7%	1.9%
Diabetes mellitus	5.7%	6.1%	3.1%
Homelessness	3.2%	3.3%	5.0%
Alcohol abuse	2.1%	2.4%	4.2%
Drug abuse	2.3%	2.4%	3.3%
Arrival in the Netherlands	(n = 1558)	(n = 1258)	(n = 360)
1947 - 2000	15.4%	15.6%	12.8%
2001 - 2005	3.5%	2.8%	1.5%
2006 - 2010	4.2%	3.5%	1.1%
2011 - 2013	5.2%	4.7%	3.6%
2014	5.4%	6.3%	8.8%
2015	9.2%	10.0%	12.8%
2016	6.3%	6.0%	7.7%
2017	6.0%	5.9%	7.5%
2018	5.1%	5.1%	7.1%
2019	1.8%	1.7%	1.1%
Unknown	38.0%	38.4%	36.2%
Years between arrival in the Netherlands and TB sample			
0	10.1%	9.5%	11.3%
1	9.6%	10.2%	13.2%
2	7.5%	8.0%	9.6%
3	5.1%	5.3%	7.7%
4	3.0%	3.2%	3.6%
5	1.8%	1.6%	1.9%
6 - 10	5.4%	4.6%	2.1%
More than 10	36.7%	32.3%	31.0%
Unknown	20.8%	25.1%	19.7%

^a Provinces excluding the mentioned cities. North Netherlands: Drenthe, Flevoland, Friesland, Groningen, Noord-Holland, Overijssel; Middle Netherlands: Gelderland Utrecht, Zuid-Holland; South Netherlands: Limburg, Noord-Brabant, Zeeland. ^b

WHO regions. ^c Excluding the Netherlands