

Cell, Volume 187

Supplemental information

**Underdetected dispersal and extensive local
transmission drove the 2022 mpox epidemic**

Miguel I. Paredes, Nashwa Ahmed, Marlin Figgins, Vittoria Colizza, Philippe Lemey, John T. McCrone, Nicola Müller, Cécile Tran-Kiem, and Trevor Bedford

Table S1. Geocoding for various country and regional scales used in this study, related to Figures 2 and 3. DTA denotes the samples for the phylogeographic analysis which was subsampled using confirmed case counts as weights. MASCOT-GLM column denotes the sample for the phylodynamic inference which was subsampled by enforcing equal temporal sampling per country per month.

Region	Country	DTA	MASCOT-GLM
North America	Canada	27	45
	United States	533	120
Western Europe	Austria	3	11
	Germany	116	84
	Switzerland	7	7
	United Kingdom	94	78
	France	25	23
	Belgium	15	18
Central Europe	Slovakia	0	11
	Slovenia	0	31
Southern Europe	Italy	6	7
	Portugal	88	75
	Spain	9	11
South America	Colombia	77	57
	Peru	3	10

Table S2. Comparison of time to most recent common ancestor (TMRCA), migration rate (migration events per year), and clock rate (substitutions per site per year) by method, related to Figures 2 and 3. First five rows denote the comparison of key summary statistics from main and alternative models used (with varying tree priors and inclusion of empirical predictors) which include three sequences from March 2022 which were found retrospectively in the UK. The last two rows represent the main analyses but without the three retrospective march 2022 samples.

Analysis	Tree prior	TMRCA		Migration rate		Clock rate	
		Mean	95% HPD	Mean	95% HPD	Mean	95% HPD
With three retrospectively-collected March 2022 sequences from the UK							
MASCOT-GLM	Approximate Structured Coalescent	2021-12-03	(2021-09-21 to 2022-02-01)	1.77	(1.45-2.13)	6.27E-05	(5.62e-5 – 6.96e-5)
MASCOT-Skyline	Approximate Structured Coalescent	2022-01-29	(2021-12-07 to 2022-03-12)	1.83	(1.38-3.00)	5.71E-05	(5.06e-5 – 6.39e-5)
DTA	SkyGrid	2022-03-24	(2022-03-09 to 2022-03-27)	0.72	(0.57-0.89)	8.41E-05	(7.71e-5 – 9.10e-5)
DTA	Skyline	2022-03-18	(2022-02-25 to 2022-03-27)	1.34	(1.03-1.66)	5.64E-05	(5.00e-5 – 6.39e-5)
DTA	Constant	2022-01-23	(2021-12-21 to 2022-02-24)	0.81	(0.68-0.89)	9.44E-05	(8.64e-5 – 1.05e-4)
Without March 2022 sequences from the UK							
MASCOT-GLM	Approximate Structured Coalescent	2021-11-29	(2021-09-19 to 2022-01-28)				
DTA	SkyGrid	2022-03-30	(2022-03-05 to 2022-04-19)				

Table S3. Reproduction numbers and dispersion parameter estimates from the analysis of the size distribution of clusters of identical sequences using a joint-likelihood, related to Figure 7. For each location, we report maximum likelihood estimates (MLE) along 95% likelihood profile confidence intervals. Different assumptions regarding the proportion of infections sequenced are explored. These estimates were obtained by allowing the reproduction numbers to vary between regions but assuming a similar value of the dispersion parameter k across locations.

Assumption regarding the proportion of infections detected	Location	R estimate (maximum likelihood estimate with 95% confidence interval)	k estimate (maximum likelihood estimate with 95% confidence interval)	Proportion of infected individuals with 0 offspring (obtained from R and k MLE)
10%	Portugal	0.87 (0.71-1.03)	0.16 (0.11-0.23)	74%
	United Kingdom	1.10 (0.98-1.22)		72%
	California (USA)	1.14 (1.08-1.20)		72%
	New York (USA)	1.05 (0.94-1.14)		72%
	Washington (USA)	1.06 (0.94-1.19)		72%
50%	Portugal	0.60 (0.44-0.80)	0.30 (0.18-0.54)	72%
	United Kingdom	0.89 (0.73-1.06)		66%
	California (USA)	0.94 (0.85-1.03)		65%
	New York (USA)	0.82 (0.69-0.94)		67%
	Washington (USA)	0.84 (0.69-1.02)		67%
100%	Portugal	0.46 (0.31-0.66)	0.36 (0.21-0.71)	74%
	United Kingdom	0.76 (0.59-0.95)		66%
	California (USA)	0.82 (0.72-0.92)		65%
	New York (USA)	0.68 (0.55-0.82)		68%
	Washington (USA)	0.70 (0.54-0.91)		68%

Table S4. Location-specific reproduction number and dispersion parameter estimates from the analysis of the size distribution of clusters of identical sequences, related to Figure 7. For each location, we report maximum likelihood estimates (MLE) along 95% likelihood profile confidence intervals. Different assumptions regarding the proportion of infections sequenced are explored.

Location	Assumption regarding the proportion of infections detected	<i>R</i> estimate (maximum likelihood estimate with 95% confidence interval)	<i>k</i> estimate (maximum likelihood estimate with 95% confidence interval)	Proportion of infected individuals with 0 offspring (obtained from <i>R</i> and <i>k</i> MLE)
Portugal	10%	0.8 (0.64-0.98)	0.09 (0.02-0.3)	81%
	50%	0.57 (0.40-0.79)	0.19 (0.04-1.8)	77%
	100%	0.45 (0.29-0.67)	0.23 (0.06-6.5)	78%
United Kingdom	10%	0.97 (0.81-1.12)	0.11 (0.03-0.34)	78%
	50%	0.76 (0.58-0.95)	0.23 (0.04-2.6)	72%
	100%	0.63 (0.45-0.84)	0.28 (0.05- >10)	72%
California (USA)	10%	1.03 (0.95-1.11)	0.07 (0.03-0.11)	83%
	50%	0.89 (0.8-0.98)	0.19 (0.08-0.42)	72%
	100%	0.79 (0.68-0.89)	0.25 (0.1-0.75)	70%
New York (USA)	10%	1.21 (1.14-1.28)	1.26 (0.37 - >10)	42%
	50%	0.92 (0.79-1.03)	1.04 (0.24- >10)	52%
	100%	0.75 (0.61-0.88)	0.99 (0.19- >10)	58%
Washington (USA)	10%	1.1 (0.99-1.21)	0.23 (0.07-0.72)	67%
	50%	0.85 (0.71-1.02)	0.38 (0.13-2.3)	64%
	100%	0.71 (0.55-0.9)	0.43 (0.16-2.8)	66%