Genomics Working Group

Shameless plug: DPH coding working group

- https://massgov-my.sharepoint.com/:x:/g/personal/mary_godec_mass_gov/EWs-bcWyYn1JkZI0FlJkhkQBW0jLQ_WOO5M8TpCeidQrtg?CID=698d6b0a-170b-a877-e9de-e372df9e5d3c

- Short-term plans: compile user list of everyone who does code/wants to code more at DPH
- Medium-term: establish good coding practices/standards for DPH
- Long-term plans – future seminars, workshops, one-on-one help, code review, development opportunities, etc

- Underdetected dispersal and extensive local transmission drove the 2022 mpox epidemic
  Paredes, Miguel I. et al.
  Cell, Volume 187, Issue 6, 1374 - 1386.e13
- https://doi.org/10.1016/j.cell.2024.02.003

Cell

CellPress

Volume 187, Issue 6, 14 March 2024, Pages 1374-1386.e13

Article

Underdetected dispersal and extensive local transmission drove the 2022 mpox epidemic

Miguel I. Paredes [1 2 8], Nashwa Ahmed [2 3], Marlin Figgins [2 4], Vittoria Colizza [5], Philippe Lemey [6], John T. McCrone [2], Nicola Müller [2], Cécile Tran-Kiem [2], Trevor Bedford [1 2 7]

Show more

Add to Mendeley Share Cite

https://doi.org/10.1016/j.cell.2024.02.003

Get rights and content

Highlights

- Phylodynamic models reveal swift early mpox spread between five global regions
- Extensive, underdetected dissemination promoted rapid local transmission
- Later mpox introductions played a negligible role in prolonging regional epidemics
- N. America epidemic declined before 10% of high-risk group had vaccine-induced immunity

Overview: Spatially resolved analyses of interventions/spread

- Despite the heightened focus on public health surveillance of emerging infections since the start of the SARS-CoV-2 pandemic, MPXV sparked regional epidemics around the world, contributing to a high degree of morbidity among those affected.[24,27,28]
- In this study, we present both a global and regional view of mpox detection, expansion, and containment by jointly analyzing genomic, mobility, and epidemiological data.
- We find evidence of rapid spread following initial regional viral seeding events, community transmission prior to detection by local public health surveillance, differential changes in case detection throughout the epidemic, a limited role of viral introductions in prolonging regional epidemics, a large degree of transmission heterogeneity, and limited impact of vaccination campaigns during the early phases of the North American epidemic.

Mpox Clade IIb

- Despite double-stranded DNA viruses typically exhibiting a slower evolutionary rate than RNA viruses,[29] clade IIb of MPXV has been found to have a significantly faster evolutionary rate since transitioning to sustained human-to-human transmission driven by APOBEC3 editing.[30] Although the evolutionary rate of the variola virus (a closely related poxvirus to MPXV) has been previously estimated to be about9×10−6 substitutions per site per year,[31] we infer the evolutionary rate of the B.1 lineage of MPXV to be 8.41×10−5 (95% HPD 7.71×10−5 to 9.10×10−5) substitutions per site per year or approximately 16.6 substitutions per genome per year (compared with the 1–2 substitutions per genome per year for variola virus). This increased evolutionary rate approaches the rate of many RNA viruses[32] and allows for a strong phylogenetic signal (Figures S2 and S5) to analyze epidemic spread and dynamics.

Evaluated interventions to see what affected spread

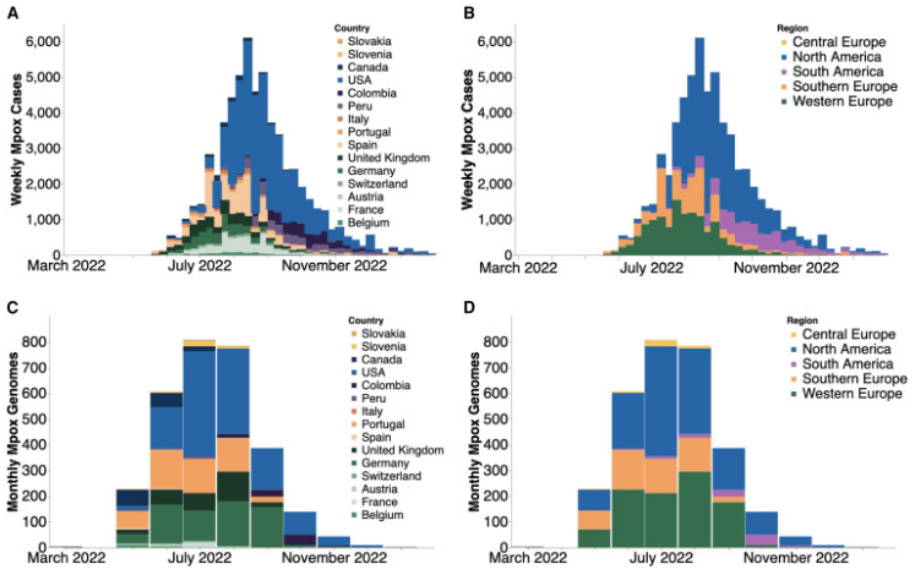
- “An outstanding question raised during the beginning of the mpox epidemic that remains unclear is the potential impact of interventions in preventing and controlling spread.[40]”
- Vaccine campaign/introduction
- Travel recommendations/warnings from CDC
- Behavioral modifications by MSM community

Mpox in the US and Canada spread predominantly among high-risk MSM populations,[50] suggesting that the majority of the North American sequences in our study were derived from a similar (but not identical) population as used to estimate vaccine coverage.
Our conclusions are concordant with those from the CDC which also found that *Rt* fell below one in August 2022 when only about 1.3% of the high-risk population in the US had any vaccine-induced immunity.[51]
Similarly, modeling of mpox in Washington, D.C. suggests that behavioral modifications within the MSM community were the main contributing factor to slowing initial mpox spread, but that vaccination campaigns were ultimately needed to definitively curb the local epidemic and prevent future outbreaks.[52,53] A UK-based modeling study focusing on MSM found that vaccination could not explain the drop in mpox incidence in the region but rather attribute the declining incidence to changes in behavior within the same community.[54]
Together, these findings highlight the significant effect of behavioral change among MSM in curbing the epidemic as well as emphasize the need for prompt public health response in order to maximize the population-level effectiveness of vaccination campaigns.

- Early undetected spread drove the outbreak

Early mpox spread in Western Europe sparks prolonged outbreaks in Southern Europe, North America, and South America

Following initial detection in the UK on May 7, 2022, the number of mpox cases reported worldwide grew rapidly (Figure 1). In early May, reported cases were found mainly in Western and Southern, and then Central, Europe where the epidemic peaked around mid-July (Figures 1A and 1B). Beginning in mid-May, cases began to be reported in North America, which ultimately led to the largest number of reported cases of any global region studied, peaking at the beginning of August. Around the same time as the North American peak, cases were detected and started rising in South America, which substantially contributed to the later tail of the 2022 epidemic. Similarly, the number of sequences collected increased as more cases were detected, with heterogeneity between regions and with North America (primarily the US) submitting the largest number of sequences to GenBank. (Figures 1C and 1D).



Download: Download high-res image (641KB)

Download: Download full-size image

Figure 1. Case counts and publicly available sequences by geographic region

(A and B) Confirmed positive weekly mpox cases by country (A)and global region (B)smoothed using a 7-day rolling average on daily data and then aggregating into weekly counts. Only countries with greater than 5 sequences on GenBank were included.

(C and D) Monthly count of publicly available MPXV genomes found on GenBank by country (C)and global region (D).

METHODS

- In order to analyze within-region transmission dynamics and enhance inference via the joint integration of genomic and epidemiological metadata, we then employed an approximate structured coalescent model (MASCOT) with a generalized linear model (GLM) approach with estimated prevalence and air passenger data as empirical predictors on 587 sequences to infer the effective population size and migration rates within and between each region, respectively (Figure S4A). We also included a predictor for each month to account for potential changes in case detection over time. The included sequences were subsampled with equal temporal weighting to increase representation of undersampled regions such as Central Europe (Figures S1A and S1C; see STAR Methods for more information).
- Using a minimum of 5∗107Markov chain Monte Carlo (MCMC) steps to promote convergence, these runtimes translate to 25.5 days of computational demand for DTA and 34.3 days for MASCOT-GLM. As such, we reduced the number of sequences to 587 for MASCOT-GLM to allow for inference within actionable timescales (Figure S1C).
- The MASCOT-GLM subsampling scheme is different from that of DTA as the structured coalescent is more robust to differences in sampling across regions and is subsequently informed by regional prevalence.[17,19]
- We used a GLM approach to draw inferential power from relevant predictors and reduce uncertainty relative to inferences using the coalescent alone.

- This suggests that transmission heterogeneity alone (without a reproduction number greater than 1) is unlikely to explain the size of the large polytomy observed at the beginning of the epidemic. **Overall, the large first polytomy is highly consistent with a reproduction number greater than 1 at the beginning of the mpox outbreak**. This aligns with reproduction number estimates obtained from our phylodynamic analysis (Figure 5), which is indicative of mpox spread within the community.

High degree of transmission heterogeneity observed in the declining phase of the mpox epidemic

Upon separating out each introduction and its inferred descendants from the MCC tree (Figure 3A), we noticed that a small number of introductions resulted in a sustained expansion of local transmission, whereas the remaining majority produced few downstream infections. The extent to which some individuals tend to contribute disproportionately to infection events is measured by the dispersion parameter *k*, which quantifies transmission heterogeneity.[21] Lower values of the dispersion parameter correspond to a higher degree of heterogeneity in transmission. When transmission heterogeneity is high, interventions targeting the most infectious individuals can have a considerable impact on epidemic burden. Quantifying transmission heterogeneity is hence important to guide control efforts. We thus sought to quantify mpox transmission heterogeneity using a method relying on the analysis of the size distribution of clusters of identical sequences (46).

PHYLODYNAMIC INFERENCE CONCLUSIONS

- After separating out each introduction and its inferred descendants from the maximum clade credibility (MCC) tree and comparing them to confirmed case counts, **we see strong evidence of viral circulation before initial detection in each global region (Figure 3A).**
- Additionally, we revealed that **the largest downstream outbreak clusters arise from introductions prior to detection from public health surveillance**, whereas **introductions after detection are more likely to be a single case and extinguish quickly** (Figure S4B).

Limitations of the study

Our study has noteworthy limitations. Our genomic data from GenBank only cover a small selection of countries and regions, suggesting that we are missing transmission events that involve unsampled countries, especially from regions such as Asia, Oceania, and Africa, although mpox cases in these areas in summer 2022 were limited and unlikely to significantly impact our results. The changing availability of genomic sequencing, as well as unequal sampling across the regions study affect the probability that a case shows up as a sequence in our dataset. If viruses migrate frequently between our study countries and countries that lack genomic sampling, the lack of samples that might interdigitate with samples from the study country may affect our ability to distinguish separate introductions. Despite this potential bias, the 2022 mpox epidemic mainly affected Europe and the Americas, which are regions that are well represented in our study, limiting the effect of this bias. Additionally, we attempted to account for this variation by weighting the subsampling for DTA according to confirmed case counts and by oversampling undersampled regions (and downsampling overrepresented regions) in our MASCOT-GLM analysis (Figure S1) as well as by adding in estimated prevalence as a predictor in the model to account for this variation.

Bayesian coalescent models assume random sampling of infected individuals, meaning that targeted sampling of superspreader events, or via contact tracing, could bias our phylodynamic estimations. We attempt to quantify the extent of transmission heterogeneity via our estimates of overdispersion (Figure 7). In the analysis of transmission heterogeneity, we explicitly accounted for the fraction of cases sequenced and explored several assumptions regarding the proportion of infections detected by the surveillance system. This was done assuming that all infections had the same probability of being detected as cases and sequenced. Active surveillance targeting larger clusters could lead to underestimating the extent of transmission heterogeneity.[23,58]

We see a discrepancy in the time to MRCA (TMRCA) between various models (Table S2) and find our estimates to be highly dependent on the tree prior and thus should be interpreted with caution. Inference of TMRCA is dependent on the estimate of effective population size in early 2022. Different tree priors assume different parametric forms of effective population size and so differ in TMRCA estimates. The rapid exponential growth observed in early 2022 suggests that effective population size should be low in January–March 2022. This information is used by the DTA skyline and skygrid models, as well as the MASCOT-skyline model, resulting in TMRCA estimates close to the earliest March sequences. Consistently, the MASCOT-GLM model estimates the coefficient of the monthly predictor for April 2022 and earlier at −1.09 (95% HPD: −1.89–0.00, Figure 3D), again supporting a small effective population size in this time period. We suggest a conservative interpretation of these results, supporting a TMRCA between September 2021 and March 2022.

community resulted in a sharp decline in *Rt* in North America ahead of vaccination rollout in the US. Our findings are relevant for policymakers in promoting broader routine specimen screening as a core tenant of pandemic preparedness. Recent emerging disease outbreaks—Zika, Ebola, SARS-CoV-2, and now mpox—have been characterized by