Elsevier has created a Monkeypox Information Center in response to the declared public health emergency of international concern, with free information in English on the monkeypox virus. The Monkeypox Information Center is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its monkeypox related research that is available on the Monkeypox Information Center - including this research content - immediately available in publicly funded repositories, with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the Monkeypox Information Center remains active.

CellPress

Med

Letter

Recombination shapes the 2022 monkeypox (mpox) outbreak

Ting-Yu Yeh,[1,*] Zih-Yu Hsieh,[2,7] Michael C. Feehley,[1,7] Patrick J. Feehley,[1,7] Gregory P. Contreras,[1,7] Ying-Chieh Su,[3,4] Shang-Lin Hsieh,[5] and Dylan A. Lewis[6]

Monkeypox (Mpox) is a global health emergency. Yeh et al. analyze tandem repeats and linkage disequilibrium in monkeypox virus (MPXV) sequences from the 2022 pandemic to determine the virus evolution, showing that these are useful tools to monitor and track phylogenetic dynamics and recombination of MPXV.

The 2022 monkeypox outbreak represents the first time this disease has spread widely beyond Central and West Africa. Initially identified in the UK in May 2022, monkeypox case numbers quickly increased in Europe, North and South America, Asia, Africa, and Oceania. On July 23, 2022, the WHO declared the monkeypox outbreak a global health emergency. On August 12, 2022, a total of 35,032 cases were confirmed in nearly 80 countries (Monkeypox data explorer. *Our World in Data*).

Most monkeypox virus (MPXV) sequences from the 2022 outbreak belong to the B.1 clade (except two cases from the US). In this study, we analyze MPXV sequences during 2022 pandemic to investigate whether the virus is adapting for better survival and transmission among the human population.

In a rapidly evolving poxvirus, adaptation is simultaneously driven by two mechanisms: recombination (gene copy number variation, fast) and single nucleotide variants (SNVs, slow) at the same loci.[1] Recombination generates new phenotypes with greatly altered disease potential that are better suited to viral survival. It has been shown that vaccinia viral DNA is swapped back and forth ~18 times per genome in a single round of infection to make recombinant phenotypes. Recombination in poxvirus genomes has been commonly detected by selection or screening in laboratory animals or cell culture for more than 60 years[2,3] Gershon et al. described recombination of four capripoxvirus isolates during natural virus transmission by analyzing physical maps of the viral genome.[4] However, very little is known about poxvirus recombination in nature due to its relatively large genome size and lack of genomic surveillance data, which is now becoming available.

Our efforts focused on discovering the variability of the MPXV genome via recombination to determine the potential risk of new viral strains. Tandem repeats (TRs) were first identified within the inverted terminal repeat of vaccinia virus DNA with a 70-base-pair sequence arranged in two blocks of 13 and 17 copies, respectively.[5] Other poxvirus TRs are small pieces of DNA sequences (3–25 nucleotides) and their sequences and copy numbers vary among different poxvirus family members or isolates.[6] The insertions and deletions of TRs are common events among these poxviruses.[7] Therefore, these TRs exhibit high rates of variation and they represent a target for poxvirus gene truncation and variation.[6,8] Here, we report the first evidence of natural recombination of monkeypox virus by analyzing TRs. We also use linkage disequilibrium, a well-known SNVs-based analysis, to detect new lineages and recombination in MPXV genome of 2022 pandemics.

To determine the genomic diversity in monkeypox genomes in the 2022 outbreak, we first searched TRs in 415 available MPXV sequences (B.1 clade) worldwide from January 1 to July 20, using the Tandem Repeat Finder algorithm. The advantage of this method is that it eliminates the bias caused by sequence alignment error, especially in low complexity sequences like TRs. Based on the criteria of the alignment score (>100) and length (>7 base pair, bp), we identified 6 TRs with variations in their copy numbers (Figures S1A and S1B). TR A/E have identical 16 bp sequences of inverted repeats (5′-TAA CTCTAACTTATGACT-3′ and 5′-AGTC ATAAGTTAGTTA-3′) at both ends of the MPXV genome (Figure S1B). The viral populations were further categorized into six groups based on TR numbers (TRNs) of TRA/E (Figure S1C). Three hundered and seventy-eight cases (90.6%) were associated with TRN = 7.9 versus 14 cases of TRN = 15.9 (3.4%), which were collected in the US (ON959133, ON959134, ON959135, ON959136, ON954773, ON959131, ON959132), Belgium (ONON622712, ON622713, ON880419, ON880420, ON880421, ON880422) and the Czech Republic (ON983168). One

[1]Auxergen Inc., Columbus Center, Baltimore, MD 21202, USA

[2]Lake Washington High School, Kirkland, WA 98033, USA

[3]Department of Thoracic Surgery, Chi-Mei Medical Center, Tainan City, 710, Taiwan

[4]Department of Biotechnology and Bioindustry Sciences, National Cheng Kung University, Tainan City, 710, Taiwan

[5]MacKay Memorial Hospital, Taipei City, 104, Taiwan

[6]Monte Vista High School, Danville, CA 94526, USA

[7]These authors contributed equally

*Correspondence: yehty@auxergen.com

https://doi.org/10.1016/j.medj.2022.11.003

case of TRN = 5.6, TRN = 3.6, and TRN = 2.6 was found in the UK (ON 619837, ON619835, ON022171). There were 21 cases with different TR numbers between TR A and E ("mismatch" in Figure S1C). This result shows that genome diversity can be grouped by TR polymorphism among MPXV populations in the 2022 pandemic.

TR B, C, D, and F are direct repeats and located at either the intergenic regions (TR B, C, D) or 3′ inverted terminal repeat (TR F) (Figure S1C, detail information of TR B, C, D in Mendeley data: 10.17632/txgdw36vxc.1). Each TR C and F contain one and three copies of 9 bp sequence (5′-TAT GATGGA-3′), respectively. While the majority of viral sequences contain TR F (TRN = 3.5, 97.1%), 50.8 and 48.2% of total samples either have TR C with TRN = 9.7 and TRN = 7.7, respectively (Figure S1D). Based on TR C/F pattern, the viral populations can be classified into 4 lineages (M, 210 cases; U, 193 cases; I, 7 cases; and one uncategorized, Figure S1D). Further, in combination with TRNs of TR C/F and TR A/E, we are able to categorize viral populations into 11 subgroups. TR D is a 9 base-pair sequence (5′-ATATCATT-3′) with more various copy numbers, ranging from TRN = 2 to 54.6 (Figure S1E). Interestingly, the sequences with high TRNs of TR D (TRN>30) also contain higher TRNs of TR A/E (TRN>15.9, 20 cases) in the group U, indicating that lineage U has more TR diversity than the others.

Taken together, our data demonstrate that TRs diverged frequently during natural transmission within the B.1 clade and that the virus is evolving as its population expands (Zeng's $E = -1.65$, Achaz's $Y = -2.52$, $p < 0.001$).

Poxvirus recombination between two co-infecting parental viruses generates genetic diversity.[1,2] MPXV recombination in natural infection has not been reported to date. Using TR polymorphism, we identified eight genomes with recombinant crossovers (Figure S1F). Case FVG-ITA-01 (ON755039) in Italy may be generated from parental sequences from the group I and M. Case VIDRL01 (ON631963) in Australia comes from parental sequence of the group M and U (Figure S1D), as well as six cases in Slovenia (ON838178, ON631241, ON60 9725, ON754985, ON754986, ON754987). This is the first report of recombination of MPXV in natural transmission to our knowledge. Our results also suggests that six Slovenian cases may have evolved into a new lineage.

We then employed single nucleotide polymorphism (SNPs) analysis using the DNASP v6 algorithm to detect the occurrence of recombination (Rozas's $Za = 0.0005$, $p < 0.05$). We also used Haploview algorithm to visualize the patterns of linkage disequilibrium (LD) between variants with minor alleles in at least two MPXV isolates and to detect the possible recombination (Figure S1G). In the absence of evolutionary forces or natural selection, D', the normalized coefficient of LD, converges to zero along the time axis at a rate depending on the magnitude of the recombination rate between the two loci. Since most SNPs were at very low frequencies, many SNP pairs had low values of squared coefficient of correlation (r^2) and the log of the odds (LOD) (Figures S1J and S1K).

The LD analysis reveals five SNP pairs located at C22736T/G74357A, G34305A/G148421A, G34305A/G189246A, G148421A/G189246A, and G186153A/C188379T (8 cases) with the high log of odds (>10) and strong evidence of LD ($\chi 2$ test, $p < 0.0001$), in which the upper 95% confidence bound of D' is above 0.98 and the lower bound is above 0.7 (Figures S1H and S1I). C22736T/G74357A SNPs are present in 28 cases, including 25 in Germany, one in Austria, one in the UK, and one in Portugal. G186153A/C188379T SNP pairs has eight German cases. There are 14 Canadian cases containing G34305A/G148421A/G189246A SNPs (Mendeley data: https://doi.org/10.17632/txgdw36vxc.1). Our results suggest that virus has evolved into at least three new lineages.

Moreover, the upper 95% confidence bound of D' for SNP pairs C25641T/C70777T and G5592A/G78031A was 0.34 and 0.87, respectively, showing strong evidence of recombination. This result suggests that two Germany cases (ON959149 and ON637939) and one Spain case (ON720849) already gained their mutations via recombination (Figure S1I).

In this study, we show the first report of natural recombination of MPXV genome based on TR (SNP-independent analysis) and LD (SNP-dependent analysis). Kugelman et al. have investigated MPXV genome diversity from 60 human samples collected in the Democratic Republic of the Congo from 2005 through 2007 based on four regions with TRs.[6] MPXV populations in 2022 pandemics can be categorized into 4 lineages with 11 different subgroups in clade B.1. Like Kugelman et al.'s studies, we found that none of the TR of MPXV strains were located at the protein coding region. However, all of their 60 samples (2005–2007) had identical right and left TRs. In contrast, we have detected 21 genomes (5.1%) with mismatch TR A/E during the 2022 pandemic to date (Figure S1C).

LD analysis also detected three new lineages (G22736T/G74357A, G186153A/G188379T, G34305A/G148421A/G189246A), suggesting that MPXV has diverged during the 2022 pandemic. Based on a neutrality test, directional selection appears to not yet be significant (normalized Wu and Fay's $DH = -0.74$, $p > 0.05$), consistent with the idea that SNVs' mutation rates are generally slower than recombination in poxvirus evolution.[7]

It has been shown TRs with the diverse length (54, 70, 125 bp) near the ends of vaccinia virus genome can provide a novel

marker to detect unequal crossing over and compare the relatedness of poxviruses.[9] However, the TR diversity has been underappreciated in poxvirus study. To prevent the power of TR annotation being compromised by sequencing-assembly error, we also validated TR A to F using multiple TR detection TRAL algorithms combined with evolutionary and statistical analyses (Mendeley data: 10.17632/txgdw36vxc.1). We have not detected ambiguous sequences at the boundary of TRs in MPXV recombinants; therefore, it is very unlikely that their sequences resulted from mixed infection of different MPXV isolates. Taken together, these data confirmed that recombination did occur in 2022 monkeypox outbreak. Our results indicate that TR analysis is well-suited for detecting poxvirus recombination events. These data also demonstrated that TR and LD analysis can detect different recombination events (Figures S1F versus S1I). In combination with genomic surveillance, TR and LD analysis are both useful tools to monitor and track phylogenetic dynamics and genetic epidemiology of monkeypox transmission.

We speculate that the MPXV genomes in the 2022 outbreak emerged most likely from a single origin, gained mutations or TRs, and then evolved into different lineages and subgroups. Then, co-infection of viruses from two parental lineages occurred, followed by homologous recombination via multiple possible mechanisms.[2] Therefore, the progeny MPXV recombinants have mosaic pattern of TRs or mutations. So far, we have not detected any defective MPXV virus arising from a single infection.

TR insertions in the promoter and 3′-untranslated regions of MPXV may also have influence on gene expression and regulation (Mendeley data: 10.17632/txgdw36vxc.1). It has been reported that the 3′-to-5′ exonuclease activity of viral DNA polymerase plays an essential role in promoting extraordinarily high levels of genetic recombination in vaccinia infection.[2] Recombination is involved in vaccinia virus adaptation to counteract the interferon-induced antiviral host cell response mediated by the double-stranded RNA-dependent protein kinase (PKR). One of the weak PKR inhibitors, vaccinia virus K3L, undergoes recurrent gene amplification with a beneficial SNV (His47Arg) via active recombination driven by selection.[1,10] This event is mediated by vaccinia RNA polymerase and leads to rapid homogenization of K3L gene arrays.[10] It is worth noting that natural MPXV C3L (vaccinia K3L homologue) is truncated at amino acid 43 by the stop codon and loss of binding site to PKRs and eIF2B (Mendeley data: 10.17632/txgdw36vxc.1). It is unclear whether MPXV virulence is changed due to the loss of the anti-interferon activity.

Due to the limitation of available demographic information, our study does not imply that MPXV recombination occurred in any specific populations of sex, gender, age, ethnicity, or socioeconomic status.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.medj.2022.11.003.

ACKNOWLEDGMENTS

We sincerely thank scientists worldwide for providing sequence information during 2022 monkeypox outbreak. This paper is made possible by their collective work on MPXV genomic surveillance. T.-Y. Yeh designed the overall experiments. T.-Y Yeh and G. P. Contreras had unrestricted access to all data. T.-Y Yeh, Z.-Y. Hsieh, G. P. Contreras performed the TR analysis, T.-Y Yeh, M. C. Feehley and P.J. Feehley performed phylogenetic analysis and summarized the sequence information. T.-Y Yeh, Y.-C. Su, S.-L. Hsieh, and D. A. Lewis performed the LD analysis. T.-Y Yeh, and Z.-Y. Hsieh performed statistical analyses. T.-Y Yeh and G. P. Contreras wrote the article. All authors read and approved the final article and take responsibility for its content.

DECLARATION OF INTERESTS

T.Y.Y. and G.P.C. are founders of Auxergen, Inc. Y.C.S. and S.L.H. are stockholders of Auxergen, Inc. All authors and Auxergen, Inc., declare no competing or financial interests.

REFERENCES

1. Sasani, T.A., Cone, K.R., Quinlan, A.R., and Elde, N.C. (2018). Long read sequencing reveals poxvirus evolution through rapid homogenization of gene arrays. Elife *7*, e35453. https://doi.org/10.7554/eLife.35453.

2. Evans, D.H. (2022). Poxvirus recombination. Pathogens *11*, 896. https://doi.org/10.3390/pathogens11080896.

3. Fenner, F., and Comben, B.M. (1958). Genetic studies with mammalian poxviruses. I. Demonstration of recombination between two strains of vaccinia virus. Virology *5*, 530–548. https://doi.org/10.1016/0042-6822(58)90043-6.

4. Gershon, P.D., Paul Kitching, R., Hammond, J.M., and Black, D.N. (1989). Poxvirus genetic recombination during natural virus transmission. J. Gen. Virol. *70*, 485–489. https://doi.org/10.1099/0022-1317-70-2-485.

5. Wittek, R., and Moss, B. (1980). Tandem repeats within the inverted terminal repetition of vaccinia virus DNA. Cell *21*, 277–284. https://doi.org/10.1016/0092-8674(80)90135-X.

6. Kugelman, J.R., Johnston, S.C., Mulembakani, P.M., Kisalu, N., Lee, M.S., Koroleva, G., McCarthy, S.E., Gestole, M.C., Wolfe, N.D., Fair, J.N., et al. (2014). Genomic variability of monkeypox virus among humans, Democratic Republic of the Congo. Emerg. Infect. Dis. *20*, 232–239. https://doi.org/10.3201/eid2002.130118.

7. Coulson, D., and Upton, C. (2011). Characterization of indels in poxvirus genomes. Virus Gene. *42*, 171–177. https://doi.org/10.1007/s11262-010-0560-x.

8. Hatcher, E., Wang, C., and Lefkowitz, E. (2015). Genome variability and gene content in Chordopoxviruses: dependence on microsatellites. Viruses *7*, 2126–2146. https://doi.org/10.3390/v7042126.

9. Baroudy, B.M., and Moss, B. (1982). Sequence homologies of diverse length tandem repetitions near ends of vaccinia virus genome suggest unequal crossing over. Nucleic Acids Res. *10*, 5673–5679. https://doi.org/10.1093/nar/10.18.5673.

10. Elde, N., Child, S., Eickbush, M., Kitzman, J., Rogers, K., Shendure, J., Geballe, A., and Malik, H. (2012). Poxviruses deploy genomic accordions to adapt rapidly against host antiviral defenses. Cell *150*, 831–841. https://doi.org/10.1016/j.cell.2012.05.049.