medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Monkeypox virus pangenomics reveals determinants of clade Ib

Gustavo Sganzerla Martinez[1,2], Anuj Kumar[1,2], Eddy Kiganda-Lusamaki[3,4,5], Mansi Dutt[1,2], Tony Wawina Bokalanga[3,4,6], Ali Toloue[1,2], Muyembe Mawete Francisca[3], Jean Claude Makangara-Cigolo[3,4,7], Patricia Kelvin[2], Amuri Aziza Adrienne[3], Christopher D Richardson[1], Emmanuel Lokilo[3], Gradi Luakanda[3], Ahidjo Ayouba[5], Anne W. Rimoin[8], Daniel Mukadi-Bamuleka[3,4,9], Eric Delaporte[5], Genay Pilarowski[10], Jason Kindrachuk[11], Laurens Liesenborghs[6,12], Lisa E. Hensley[13], Lorenzo Subissi[14], Martine Peeters[5], Nicole A. Hoff[8], Olivier Tshiani-Mbaya[15], Sofonias Tessema[16], Jean-Jaques Muyembe Tamfum[3,4], Steve Ahuka-Mundeke[3,4], Alyson A Kelvin[17], John M Archibald[18,19], Placide Mbala-Kingebeni[3,4], Luis Flores-Giron[20], David J Kelvin[1,2]

[1]Department of Microbiology and Immunology, Dalhousie University. Halifax, Nova Scotia, Canada. B3H 4H7

[2]Canadian Center for Vaccinology, Izaak Walton Killam (IWK) Health Center, Department of Pediatrics, Halifax, NS B3H 4H7, Canada

[3]Institut National de Recherche Biomédicale, Kinshasa, Democratic Republic of the Congo

[4]Service de Microbiologie, Département de Biologie Médicale, Cliniques Universitaires de Kinshasa, Université de Kinshasa, Kinshasa, Democratic Republic of the Congo

[5]TransVIHMI (Recherches Translationnelles sur le VIH et les Maladies Infectieuses endémiques et émergentes), Université de Montpellier, French National Research Institute for Sustainable Development (IRD), INSERM, Montpellier, France

[6]Department of Clinical Sciences, Institute of Tropical Medicine, Antwerp, Belgium

[7]Graduate School of Cellular and Biomedical Sciences, University of Bern, Bern, Switzerland

[8]Department of Epidemiology, Jonathan and Karin Fielding School of Public Health, University of California, Los Angeles, CA, USA

[9]Rodolphe Merieux INRB-Goma Laboratory, Goma, Democratic Republic of the Congo

[10]Chan Zuckerberg Biohub, San Francisco, CA 94158, USA

[11]University of Manitoba, Winnipeg, Manitoba, Canada

[12]Department of Microbiology, Immunology and Transplantation, KU Leuven, Leuven, Belgium

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

[13]US Department of Agriculture, Manhattan, KS, USA

[14]World Health Organization, Geneva, Switzerland

[15]Frederick National Laboratory, Leidos Biomedical Research, Clinical Monitoring Research Program Directorate, Frederick, MD, USA

[16]Africa Centers for Disease Control and Prevention (Africa CDC), Addis Ababa, Ethiopia

[17]Vaccine and Infectious Disease Organization (VIDO), University of Saskatchewan, Saskatoon, Canada

[18]Department of Biochemistry and Molecular Biology, Dalhousie University, Halifax, Nova Scotia, Canada

[19]Institute for Comparative Genomics. Dalhousie University, Halifax, Nova Scotia, Canada

[20]Centre de Recherche en Sciences Naturelles de Lwiro, South Kivu, DS Bukavu, Democratic Republic of the Congo

Abstract

Mpox, formerly monkeypox, is a viral zoonotic disease caused by the monkeypox virus (MPXV). MPXV, which is taxonomically divided into clades I and II, was declared a Public Health Emergency of International Concern for the second time in August 2024 due to rapid geographic expansion of clade I viruses including the newly identified clade Ib [1]. With a unique set of genomic mutations and sustained human-to-human transmission, clade Ib has rapidly spread throughout the eastern Democratic Republic of the Congo as well as neighboring non-endemic regions outside the African continent [2-6]. Currently, there is a lack of comparative genomic data with which to address potential zoonotic transmissibility and pathobiology of clade Ib. Here we show MPXV clades I and II share a core genome composed of 68 protein-coding genes which are common in the genomes of other poxviruses such as camelpox, cowpox, and vaccinia. The first documented and all subsequently examined isolates of clade Ib lack the gene pair *OPG032* and *OPG033*, which encode the complement control protein (a vaccinia ortholog associated with virulence), and a Kelch-like protein, respectively. The genomic rearrangement of MPXV suggests a functional evolution that might play an important role in the pathobiology of the novel clade Ib virus. Our results lay the groundwork to exploit the genomic of elements of MPXV as potential targets for therapeutics development/repurposing, vaccine design, and molecular diagnostic expansion, as well as to uncover the viral diversity, and zoonosis of MPXV.

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Keywords: Monkeypox virus; mpox; orthopoxvirus; MPXV; Mpox clade Ib; *OPG032*; pangenome; poxviruses.

INTRODUCTION

Monkeypox virus (MPXV), the causative agent of mpox, is a double-stranded DNA poxvirus endemic to countries in Africa including the Democratic Republic of the Congo (DRC). The first human infection of MPXV was reported in the DRC in 1970 and several outbreaks with human-to-human transmission have since been reported. Taxonomically, MPXV belongs to the genus *Orthopoxvirus* and is divided into two clades: clade I (formerly Congo Basin or Central African Clade) and clade II (formerly West African Clade) [7]. Orthopoxviruses can be taxonomically divided into old world and new world clade [8]. The World Health Organization (WHO) declared a Public Health Emergency of International Concern (PHEIC) in 2022 as MPXV clade IIb viruses rapidly spread to non-endemic countries and territories (n=116) where human infections and deaths were confirmed [9]. In September 2023, a new MPXV clade Ib was first documented in Kamituga, South Kivu, DRC [2]. Clade Ib is currently expanding to Eastern African nations [3,4,6,10] and outside the African continent with confirmed reports of community transmission in the United Kingdom of Great Britain and Northern Ireland. On 13 August 2024, the Africa Centres for Disease Control and Prevention (Africa CDC) declared mpox as a Public Health Emergency of Continental Security [11] just before the WHO's declaration [12] of the outbreak as a second mpox-related PHEIC in less than three years.

The clinical manifestation of both mpox clades is similar [13]. Infected individuals typically develop smallpox-like symptoms including fever and rash that progresses to macules, papules, vesicles, pustules, and scabs, generally beginning on the face and spreading to other body parts [14]. Sexual contact between individuals has been listed as a mode of transmission, but it is not exclusive; transmission by close contact with personal belongings of infected individuals has also been documented [5]. *In-vivo* data suggests the clade Ia virus infection is more virulent with higher morbidity [15]. Limited MPXV clade Ib epidemiological data suggest the novel clade is more transmissible and causes lower case fatality rates when compared to previous clade Ia outbreaks [3,5].

The large genome of MPXV encodes core genes conserved in other orthopoxviruses whose function is involved in DNA replication, transcription and virion assembly. Inhibiting cytokine signaling, blocking apoptosis, and antagonizing innate immune pathways are functions commonly attributed to accessory genes found in MPXV that modulate the host immunity. The gene-rich genome of MPXV allows its adaptation to diverse hosts, highlighting the zoonotic potential and pathogenicity of MPXV [7, 16]

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

We have performed a comparative genomic analysis of all available complete, fully annotated MPXV clade I and clade II genomes, and identified genetic differences predicted to underly zoonosis, transmissibility, and pathogenicity of the novel clade Ib virus. Our findings will support future efforts to (i) track MPXV viral diversity and their pathology; (ii) develop/repurpose therapeutic agents and vaccines; and (iii) create molecular diagnostic models.

RESULTS

Identification of the MPXV clade I and clade II core genomes

To identify the protein-coding genes that compose the core genome of MPXV clade I, the proteomes of 181 complete, fully annotated genomes with coverage >= 90% were obtained. The number of protein-coding genes per genome ranged from 170 to 242 (mean number per genome = 184.28, median = 183, standard deviation = ±9.99). Next, we identified a total of 87 protein-coding genes that are present in at least 99% of the 181 genomes considered, composing the core genome of MPXV clade I (Figure 1-A). To determine the genes that compose the core genome of MPXV clade II, the proteomes of 2,390 fully annotated genomes were analyzed. The number of protein-coding genes ranged from 147 to 214 (mean number per genome = 177.84, median = 179, standard deviation = ±3.93). We found a total of 133 protein-coding genes to be present in at least 99% of the 2,390 genomes under investigation, composing the core genome of MPXV clade II (Figure 1-B).

We compared the core genomes of MPXV clades I and II to identify protein-coding genes shared by, and unique to, both clades (Figure 1-C). A total of 68 protein-coding genes (Supplementary Material S1) were found to have orthologs in all the genomes of clades I and II, comprising the core genome of MPXV. We determined that 19 protein-coding genes from clade I had no orthologs in clade II. Finally, 65 protein-coding genes of clade II had no orthologs in clade I. We highlight these proteins as potential targets for molecular diagnostic approaches that could distinguish between different MPXV clades.

[Figure 1]

Figure 1 – Pangenome investigation of MPXV clades I and II genomes. Minimum, maximum, mean, median, and standard deviation of the number of protein-coding genes are shown as well as the protein-coding genes that compose the core genomes of MPXV clade I (Figure 1-A; the list of protein-coding genes is found in Supplementary Material S1) and MPXV clade II (Figure 1-B; the list of protein-coding genes is found in

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Supplementary Material S2). In Figure 1-C, we display the number of protein-coding genes that compose the core genome of MPXV (clades I & II) as well as the protein-coding genes that are unique to clade I genomes (grey) and clade II genomes (green). The intersection between the two sets marks the protein-coding genes that are shared by both clades and consequently comprise the core genome of MPXV clades I and II (Supplementary Material S3).

MPXV clade I pangenome map

We created a pangenome map of MPXV clade I (Figure 2) based on the OPG nomenclature associated with the annotation of MPXV clade Ia reference genome NC003310.1, which has 176 non-redundant protein-coding genes. From these, 68 belong to the core genome of MPXV clades I and II; 19 are accessory proteins found in all MPXV clade I genomes; and the remaining 89 proteins are accessory genes found in some but not all MPXV clade I genomes. Each OPG in Figure 2 is classified according to its function [2]. Morphogenesis (n=19), transcription (n=10), and immunomodulation (n=10) are the prevalent functions in the core genome of MPXV clades I and II.

Detailed quantifications of MPXV core genome protein-coding genes and gene ontology analysis are included in Supplementary Material S4.

[Figure 2]

Figure 2 – MPXV pangenome map[1]. Figure 2 shows the 176 protein-coding genes that compose the pangenome analysis of MPXV with each gene represented by a symbol mapped to its OPG ortholog (derived from genome NC003310.1). Functional classifications are shown on the right. Each gene is labeled as belonging to (i) MPXV clades I and II core genome (blue); (ii) MPXV clade I core genome; (green) or (iii) MPXV clade I accessory (red). Gene product lengths in amino acids (aa) of each protein-coding gene and predicted functions are also shown (relative to NC003310.1).

[1]The *OPG033* gene has been considered as a miscellaneous region in the reference genome of MPXV clade I (NC_003310.1), thus, its product is not present in the reference annotation. *OPG033* ranges from 19,778 to 21,29 nucleotides in NC_003310.1. Initially, we obtained this region from the genome in fasta format and aligned it with the other representative genomes of clade I and clade II. Based on the outcome of the alignment we observed a partial deletion of ~490 nt in the clade Ib, while in the case of clade IIa and b, a large deletion of 1127 nt was noticed. Subsequently, to evaluate the encoding potential of *OPG033* nt sequences, we used the ORFFinder tool (https://www.ncbi.nlm.nih.gov/orffinder/) and predicted as many as 11 ORFs based on the detection of ATG code. Out of 11 predicted ORFs, ORF11 ranging from 1,365-1,042 nt was considered the top-ranked ORF based on the predicted coding sequence and amino acids (aa) length (324nt/107 aa) and 3 reading frames as well. Moreover, the CD-Search program (https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi), predicted the kelch-like protein; provisional functional domain (length 107 aa) in the selected ORF11 of the *OPG033*.

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

The deletion of *OPG032* and *OPG033* genes underpins novel MPXV clade Ib.

The deletion of the *OPG032* gene, an important virulence factor in both vaccinia [17], and MPXV [23], has been reported as a mark of the current MPXV clade Ib outbreak [2]. Interestingly, past evidence of *OPG032* deletions in other poxviruses ocurred in conjunction with its adjacent gene *OPG033* [18]. We considered the nucleotide content of 195 MPXV clade I genome sequences (175 clade Ia; 20 clade Ib) and queried the presence of the region that codifies the gene pair *OPG032* and *OPG033* among them. A wider region of the reference clade I genome corresponding to the *OPG032/OPG033* loci (positions 16000:22000) was aligned with DNA segments spanning the same range in 195 MPXV clade I genomes. Figure 3-A show deletion of the genes *OPG032* and *OPG033* in all 20 clade ib sequences relative to the reference genome and two random clade Ia sequences (PP601206, and PP601200). Additionally, the 175 clade Ia sequences, when aligned to *OPG032*+*OPG033*, have alignment scores >= 200, indicating the gene presence and high sequence conservation across all the queried clade Ia genomes (Supplementary Material S5). *OPG032* and *OPG033* are not found in clade II viruses. Additional genome alignments including MPXV clade II viruses (Supplementary Material S6) also highlight partial deletion of the *OPG033* gene.

Next, we queried all the 195 DNA sequences of clade I in terms of the presence or absence of *OPG032* – *OPG033* genomic region(Figure 3-B) and determined that all the genomes deposited from 1979 to 2023 contained both genes. Up to 2024, *OPG032* and *OPG033* were part of the core genome of MPXV clade I; the subsequential split of clade I into clade Ia and Ib appears directly linked to the presence or absence of *OPG032* and *OPG033*. It remains to be seen if MPXV clade Ib strains will outcompete MPXV clade Ia when both lineages are cocirculating in the same geographical circulation, reported in [14].

[Figure 3]

Figure 3 – Presence and absence of *OPG032* and *OPG033* across all MPXV genomes deposited under the clade I taxonomy. (**A**) Schematic alignment of nucleotide positions 16,000-22,000 (6,000 in total) in all clade I MPXV genomes. In the query, the location of the gene pair *OPG032* and *OPG033*, found in the clade Ia reference genome coordinates (reverse complement) of 19,070-19,710 and 19,776-21,291, respectively. The first two sequences on top of the panel are two random clade Ia sequences which have alignment scores >= 200 for the *OPG032* and *OPG033* locu. An alignment containing all 175 clade Ia sequences is included in Supplementary Material S5. (**B**) Timeline showing MPXV clade I sequences with and without *OPG032* and *OPG033*.

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

In the MPXV clade I reference genome, *OPG033* is listed as a miscellaneous feature without a predicted protein. We searched for orthologs of the protein encoded by MPXV *OPG032* across different chordopoxviruses from the genera orthopox, capripox, leporipox, suipox, and yatapox. Out of nine queried orthopoxviruses (Table 1), the *OPG032* product orthologs were not found in the reference genome of racconpox and volepox viruses (new world clade). We obtained all the complete, fully annotated genomes of the remaining seven orthopoxviruses (old world clade) and queried the presence or absence of *OPG032* product orthologs, which were found in all the queried genomes of camelpox, cowpox, ectromelia, horsepox, taterapox, and variola, composing their core genome. Six out of 91 complete, fully annotated vaccinia virus (VACV) genomes did not have *OPG032* orthologs. All six VACVs were attenuated and modified strains used in vaccine development. Additional information on the individual VACVs missing the *OPG032* ortholog are included in Supplementary Material S7.

	OPG032 ortholog in ref genome	N. of complete, full annotated genomes	*OPG032* ortholog present in	Orthopoxvirus Clade
Camelpox	YES	11	11	Old world
Cowpox	YES	98	98[2]	Old world
Ectromelia	YES	6	6	Old world
Horsepox	YES	4	4	Old world
Taterapox	YES	2	2	Old world
Vaccinia	YES	91	85[3]	Old world
Variola	YES	76	76	Old world
Raccoonpox	NO	NA	NA	New world
Volepox	NO	NA	NA	New world

Table 1 – *OPG032* orthologs across different orthopoxviruses[1]. Homologs of MPXV *OPG032* were found in different orthopoxvirus genomes. Two orthopoxviruses of the new world clade, i.e., raccoonpox and volepox, did not have obvious *OPG032* orthologs. Six old world orthopoxviruses, i.e., camelpox, cowpox, ectromelia, horsepox, taterapox, and variola, had *OPG032* orthologs in all the queried genomes. An *OPG032* ortholog was found in most but not all the queried genomes of vaccinia (93.45%).

[1]No *OPG032* orthologs were found in other chordopoxviruses than orthopoxvirus. [2]Innitially, the annotation files of the cowpoxviruses under the GenBank IDs MK035757.1, MK035749.1, MK035747.1,

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

MK035746.1, MK035750.1, MK035753.1, MK035759.1, KY549151.1, KY549146.1, and KY549147.1, and MK035748.1 did not include the gene CPXV_034, ortholog to MPXV *OPG032*, We further investigated the nucleotide record of these genomes to find alignments matching the region of the genomic coordinates of CPXV_032, i.e., 35215 to 35916. [3]6 out of 91 vaccinia genomes did not have *OPG032* orthologs. All the genomes lacking the *OPG032* ortholog (MK314710.1, MK314713.1, MK314712.1, AY603355.1, MK336431.1, MT946551.2) are attenuated vaccinia genomes used in the development of smallpox vaccines. Additional information on individual vaccinia genomes lacking *OPG032* is present in Supplementary Material S6.

Comparison of the MPXV Clades I and II core genome with other *chordopoxviruses*

We searched for orthologs of all 68 MPXV clades I and II core genes in 15 distinct chordopoxviruses and show their presence/absence in Figure 4. The most abundant genus, i.e., orthopoxviruses, had proteins of the core MPXV clade I and II genomes mapped to 9 species. All 68 proteins with predicted functions that compose the core genome of MPXV clades I and II were found in the camelpox, cowpox, and vaccinia viruses. In horsepox, we found orthologs of all 68 protein-coding genes that compose the core genome of MPXV clades I and II except for '*MPXV_gp150*, (*OPG173*), totaling 67 orthologs. Taterapox and variola had orthologs for 66 protein-coding genes of the core genome of MPXV. Both lacked the '*Protein F14'* (*OPG058*) and the first also lacked *DNA directed RNA Polymerase subunit'* (*OPG083*), while the second lacked *MPXV_gp150*, (*OPG173*). Ectromelia had 63 orthologs from the core genome of MPXV clades I and II, missing '*Hemagglutinin*' (*OPG185*), '*MPXV_gp150*, (*OPG173*), '*NFkB inhibitor*' (*OPG038*), '*Protein F14*' (*OPG058*), and '*Telomere-binding protein I6*' (*OPG049*). Finally, raccoonpox and volepox viruses did not have orthologs of 17 and 18, respectively, protein-coding genes that compose the core genome of MPXV Clades I and II.

Chordopoxviruses from genera other than orthopoxviruses had fewer orthologs of the core genome of MPXV clades I and II. The three capripoxviruses, goatpox, lumpyskin, and sheeppox, had six, five, and six orthologs, respectively, orthologs from the core genome of MPXV clades I and II with '*DNA-binding phosphoprotein (1)*' (*OPG062*), '*Virion core protein P4a*' (*OPG136*), and '*Zinc finger-like protein (2)*' (*OPG021*) being common among the three viruses. Finally, other genera such as Leporipox (myxoma virus), Suipox (swinepox), and Yatapox (yaba-like-disease virus) had, respectively, four, two, and four orthologs found in the core genome of MPXV clades I and II. In conclusion, the orthologs of protein-coding genes that were most predominant across 15 different chordopoxviruses were '*Virion core protein P4a*' (*OPG136*), '*Zinc finger-like protein (2)*' (*OPG021*), '*DNA-binding phosphoprotein (1)*' (*OPG062*), and '*Glutaredoxin-1*' (*OPG075*), with 14, 12, 12, and 12 hits, respectively.

[Figure 4]

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Figure 4 – Presence/absence of MPXV clades I and II core genome proteins in 15 chordopoxviruses. The figure shows each of the 68 proteins with predicted functions present in the core genome of MPXV clades I and II across 15 different chordopoxviruses. The y-axis (left) shows the name (alphabetically) of each protein per its annotation on MPXV (NC003310.1) for diverse chordopoxviruses (grouped per genus in the x-axis). The right y-axis shows the scale of presence (dark color)/absence (light color). The viruses on the x-axis are colored according to their genus, i.e., orthopoxvirus (red) with the viruses camelpox, cowpox, ectromelia, horsepox, raccoonpox, taterapox, vaccinia, variola, and volepox. Capripoxvirus (purple) with goatpox and sheeppox. Leporipoxvirus (orange) with myxoma. Suipoxvirus (pink) with swinepox. Yatapoxvirus (gray) with yaba-like disease virus.

DISCUSSION

We identified a total of 68 genes that are shared by most of the genomes of both MPXV clades. While clade I was found to have 19 clade-specific genes, clade II had 65; this discrepancy is perhaps due in part to the increased sampling of clade II. Regardless, from the perspective of prevention, vaccine development efforts should consider the conserved elements of MPXV/orthopoxvirus/chordopoxvirus highlighted by our pangenome analysis. These elements could then be complemented with clade-specific attributes in case of insufficient conferred protection. Current vaccines conferring protection against MPXV are (i) Imvamune, a live, non-replicating vaccinia virus vaccine; (ii) ACAM2000, a live, replication-competent vaccinia virus vaccine; and (iii) LC16m8, an attenuated, replication-competent vaccinia virus vaccine. These three vaccines do not have specific MPXV antigens but are instead based on vaccinia virus antigens (with some examples being the genes A27L, B5R, L1R, and A33R), a virus that in our analysis was found to have all 68 protein-coding genes with predicted functions found in the MPXV pangenome. The cross-protection of MPXV and smallpox is found beyond current vaccines in the form of the legacy protection provided by earlier generations of the smallpox vaccine used during the smallpox eradication era. With the reduction of vaccinated individuals in the population due to those born after 1980 not routinely receiving the smallpox vaccine [19], we might expect outbreaks in naive populations with less immunity against orthopoxviruses. Our pangenome data are thus useful for the identification of proteins with potential antigenic effect for the development of vaccines against MPXV such as [20], [21], and [22]. From a treatment perspective, our data can guide therapeutic development of viral protein targets that are conserved across different MPXV isolates. For example, we found that protein-coding genes related to transcriptional activities are a component of the set of genes that belong to the core genome of MPXV clades I and II. The amino acid residues of the protein DNA-dependent RNA polymerase subunit 147 were previously identified as binding targets

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

for inhibition with small molecules in a molecular modelling approach [23]. Taken together, the results of our core genome analysis will be of use in the development of vaccines and therapeutics against MPXV.

A major feature differentiating MPXV clade I genomes is the presence and absence of the gene pair *OPG032* and *OPG033* in clades Ia and Ib, respectively. The joint loss of *OPG032* and *OPG033* has previously been documented in the evolutionary history of orthopoxviruses [18] where *OPG033* has been linked to inflammation inhibition and reduction of immunopathology [24]. Evidence of a more transmissible MPXV clade I virus was first observed by our team in Kamituga, South Kivu, DRC, in September 2023 [2,3,5]. The rapid spread of this virus in 2024 led the WHO to declare a PHEIC in August 2024. The MPXV gene *OPG032* has an ortholog in vaccinia, i.e., Vaccinia Complement Control Protein (VCP). In vaccinia, this gene is described as a virulence factor and it modulates the complement system activation and inhibits early steps of the complement cascade by dissociating the C3 and C5 convertase enzymes that start and maintain the complement cascade. VCP is described as a virulence factor. Mice infected with a VCP-knockout vaccinia virus developed significantly smaller smallpox lesions than those infected with the wildtype virus [25]. An MPXV study [26] incorporated and removed the *OPG032* from clade II and I MPXV viruses, respectively. Upon infecting prairie dogs with the mutated viruses, it was identified that the removal of *OPG032* resulted in reduced mpox disease morbidity and mortality, indicating the gene could be a significant virulence factor. However, the addition of *OPG032* in a clade II virus did not accelerate clinical disease course nor affect disease mortality. Limited and context-dependent epidemiological data from the current outbreak of MPXV clade Ib in eastern DRC suggests the clade Ib viral infections in humans have lower case fatality rates than clade Ia infections in humans; however, other factors such as access to health care and the presence of coinfections may play a role as well.

We found that MPXV *OPG032* orthologs are core genes of old world orthopoxviruses including cowpox, vaccinia, and horsepox. Poxviruses in general have a lower rate of point mutations [27], and genomic rearrangements, gene losses, duplications, and gains by either recombination or horizontal gene transfer have been documented [16, 27]. *OPG032* and *OPG033* are not found in MPXV clade II genomes, responsible for the multi-country outbreak of 2022. Interestingly, the two viruses that led the WHO to declare mpox as a PHEIC lacked *OPG032* and *OPG033*. While functional characterization and validation of the roles of these genes in sustained human-to-human transmission is needed, our results suggest that the deletion of *OPG032* and *OPG033* play a role in driving MPXV clade Ib zoonosis, transmissibility, and pathogenicity.

In conclusion, we identified 68 protein-coding genes that are present in the genomes of MPXV clades I and II viruses. Most of these proteins are involved in housekeeping

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

activities of the virus, such as morphogenesis and transcription. The presence of the entire core genome of MPXV clades I and II in camelpox, cowpox, and vaccinia viruses demonstrates an overall conserved core of genes among *orthopoxviruses*, which might be exploited for developing therapeutics, vaccines, and molecular diagnosis reagents. By highlighting a functional schism between the MPXV clade Ia and Ib genomes – defined by the the presence or absence of the gene pair *OPG032* and *OPG033* – we lay the groundwork for questioning the impact of the complement control system in the virulence and sustained human-to-human transmission of MPXV.

MATERIALS & METHODS

Fully annotated MPXV complete genomes from public data sources

Fully annotated complete MPXV genomes (n=2,436) were obtained from the National Center for Biotechnology Information (NCBI) on 2024-06-01. A total of 46 MPXV clade I (43 clade Ia and 3 clade Ib) and 2,390 clade II genomes were considered for the identification of protein-coding genes unique to clade I, unique to clade II, and shared by both clades.

Non-public MPXV clade I genomes retrieval

To increase the number of MPXV clade I genomes, we obtained 429 additional sequences collected and sequenced from the DRC during the period of 2018-2024 [4]. Due to the number of ambiguous nucleotides in the sequencing runs, we filtered each genome to not have its entire genome content equal to or greater than 10% of N bases. By applying the filter, we were left with 132 MPXV clade I genomes to be manually annotated. All the additional 132 MPXV genomes are clade Ia.

DNA sequencing of three MPXV clade Ib samples

Due to the lower number of fully sequenced and annotated MPXV clade Ib genomes available in public repositories (n=3), we obtained three additional MPXV clade Ib isolates. DNA extraction of three MPXV clade Ib samples part of the Institut National de Recherche Biomédicale (INRB) surveillance program was performed using the QIAamp® DNA Mini kit (Qiagen, Hilden, Germany). To sequence the full-length MPXV genome, libraries were made using a hybridization probes enrichment protocol with the Twist kit and the Comprehensive Viral Research Panel (Twist Biosciences). Obtained libraries were loaded onto the GridION sequencer. Consensus generation was performed as described previously [28]. Genomes are available at https://github.com/inrb-labgenpath/DRC_MPXV_Genomic_Surveillance.

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

MPXV clade I genome annotation

The non-annotated MPXV clade I genomes (132 clade Ia obtained from the national surveillance program of INRB [28] and 3 clade Ib sequenced in this study) were manually annotated using the Java-based VIGOR4 (Viral Genome ORF Reader) [29] pipeline mapping each genome to the MPXV references database embedded in VIGOR4. The resulting output files included predicted CDSs, proteins, alignments, GFF3 files, and GenBank tbl.

Ortholog identification

We searched for orthologs of the protein-coding genes of 2,570 MPXV viruses (181 clade I [175 clade Ia; 6 clade Ib], and 2,390 clade II) using OrthoFinder (version 2.5.5) [30]. Additionally, to validate the identified orthologs, we used a reciprocal local implementation of the Basic Local Alignment Search Tool (version 2.15.0) [31] (protein-protein BLASTp). To determine whether a gene belonged to the core genome of clade I, clade II, or both, we considered a threshold of at least 80% sequence identity and an e-value equal to or less than $1x10^{-10}$. Moreover, for a gene to be considered part of the core genome of clade I or clade II, it had to be present in 99% of all the genomes of the clade.

Partial genomes of MPXV clade Ib for querying the *OPG032* gene

In addition to the 6 MPXV clade Ib genomes (3 complete and annotated sequences from GenBank and 3 obtained in this study), we obtained 14 MPXV genomes deposited to GenBank under the taxonomy of clade Ib that matched the inclusion criteria of having its genomic content equal to or higher than 90% of unambiguous nucleotides. As the extra 14 clade Ib genomes do not fit the inclusion criteria of being sequenced completely and being fully annotated, they were not considered as part of the pangenome analysis. In the annotation of the reference genome of MPXV clade Ia (NC003310), *OPG032* consists of 651 nucleotides and is found in the complement sequence within the coordinates 19060 – 19710. A flanking region spanning from 16000 to 22000 was considered. All clade I genomes were aligned using the multiple sequence alignment tool at https://blast.ncbi.nlm.nih.gov/ with the megablast algorithm using default parameters.

Proteome of *chordopoxviruses*

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

We searched GenBank for the complete genome annotation of the 15 *chordopoxviruses* whose taxonomic classification were obtained from [32] and downloaded their protein-coding genes in translated amino acid format (Table 2).

Virus	Genus	Protein-coding genes	Genome Accession	Taxonomy ID
Camelpox	Orthopoxvirus	261	GCF_000839105.1	203172
Cowpox		233	GCA_025738115.1	10243
Ectromelia		180	GCA_006458505.1	1452750
Horsepox		228	GCA_006451015.1	397342
Raccoonpox		207	GCF_001029045.1	10256
Vaccinia		214	GCA_000860085.1	10245
Variola		211	GCA_000859885.1	10255
Volepox		204	GCA_001744115.1	28874
Taterapox		220	GCA_000869985.1	28871
Goatpox	Capripoxvirus	149	GCA_006465485.1	1416740
Lumpyskin		156	GCA_000839805.1	59509
Sheeppox		147	GCA_000840205.1	10266
Yaba-like disease	Yatapoxvirus	155	GCA_000847185.1	99000
Swinepox	Suipoxvirus	146	GCF_000839965.1	300880
Myxoma	Leporipoxvirus	158	GCF_000843685.1	31530

Table 2 – Reference proteomes from isolates of the *Chordopoxvirinae* subfamily.

OPG032 orthologs in orthopoxviruses

We ran an ortholog search querying the *OPG032* gene product (NC003310 as reference) in the protein-coding genes present in Table 2. We considered an ortholog to have equal or higher than 80% of sequence identity and an e-value equal to or less than 1e-10 in protein alignments. Taxonomically, we divided orthopoxviruses into two clades: old world and new world [33]. *OPG032* orthologs were found only in old world orthopoxviruses. We obtained from GenBank (November 4 2024) all the complete, fully annotated genomes of camelpox (n=11), cowpox (n=98), ectromelia (n=2), horsepox (n=4), taterapox (n=2), vaccinia (n=91), and variola (n=76) and queried the presence or absence of the *OPG032* orthologs. We identified 11 cowpox virus genomes (additional

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

information on each genome is included in Supplementary Material S7) without orthologs of the *OPG032* mpox clade Ia gene being present in the annotation files. Further investigation of the DNA sequence of the eleven genomes showed matching alignments to the nucleotide content that spans *OPG032* region.

Functional analysis

For the genes found to be unique to MPXV clades I and II, a gene ontology (GO) enrichment analysis was performed using the BLAST2GO program embedded in the OmicsBox (version 3.3) platform. BLASTX-fast was selected by specifying the non-redundant protein sequences (nr v5) and *Orthopox* (genus) as a taxonomic filter. E-value cutoff was set at $1x10^{-5}$, while other parameters were set at default. We also manually classified the genes that compose the core of MPXV (clades I and II) according to the function of the genes previously specified at [2] into immunomodulation, virulence, cell-to-cell fusion, viral replication/repair, cell entry/attachment, transcription, morphogenesis, and unknown.

MPXV clade I pangenome map

We built a pangenome map of MPXV using the reference genome NC003310.1 as a template due to its universal OPG nomenclature. Protein-coding genes identified in the core genome of MPXV clades I and II, in the MPXV clade I core genome, or in the MPXV clade I accessory set were matched to their OPG orthologs. Their functions were classified as being related to immunomodulation, virulence, cell-to-cell fusion, viral replication/repair, cell entry/attachment, transcription, morphogenesis, or ‘unknown’ [2]; their lengths in amino acid were inferred, and their predicted descriptions were linearly plotted using Python’s (version 3.12) matplotlib (version 3.8.3) library.

Acknowledgements

We appreciate the assistance provided by Dr. Nikki Kelvin for input regarding poxviruses. We also acknowledge participation of the Digital Research Alliance of Canada through its Atlantic Canadian provider ACENET for providing high-performance computing infrastructure to partially run the analyses depicted in this work.

Funding

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

This work was supported by awards from the Canadian Institutes of Health Research, the Mpox Rapid Research Funding initiative (CIHR MZ1 187236), Li-Ka Shing Foundation (DJK), Research Nova Scotia (DJK).

Research in the Archibald Lab is supported by the Natural Sciences and Engineering Research Council of Canada (RGPIN- 2019-05058) and the Gordon and Betty Moore Foundation (GBMF5782).

The genomes were generated through DRC national genomic surveillance activities which are supported by many initiatives such as the Africa CDC Pathogen Genomics Initiative (Africa PGI) (grants BMGF-INV-018278, INV-033857, Saving Lives and Livelihoods program, and NU2HGH000077); AFROSCREEN project (grant agreement CZZ3209, coordinated by ANRS-MIE Maladies infectieuses emergentes in partnership with Institut de Recherche pour le Developpement [IRD] and Pasteur Institute) funded by Agence Francaise de Developpement; PANAFPOX project funded by ANRS-MIE; Belgian Directorate-General Development Cooperation and Humanitarian Aid and the Research Foundation, Flanders (FWO, grant number G096222 N to L.L.); Department of Defense, Defense Threat Reduction Agency, and Monkeypox Threat Reduction Network; USDA Non-Assistance Cooperative Agreement #20230048; International mpox Research Consortium (IMReC), through funding from the Canadian Institutes of Health Research and International Development Research Centre (grant no. MRR-184813); and E.K.-L. received a PhD grant from the French Foreign Office.

Conflict of Interest

The authors G.S.M, A.K., M.D., P.K, and D.J.K. are shareholders of the company BioForge Canada Limited. The authors declare the interests of the company had no impact in the study.

References

1. https://www.who.int/news/item/14-08-2024-who-director-general-declares-mpox-outbreak-a-public-health-emergency-of-international-concern
2. Masirika, L. M., Kumar, A., Dutt, M., Ostadgavahi, A. T., Hewins, B., Nadine, M. B., Steeven, B. K., Mweshi, F. K., Mambo, L. M., Mbiribindi, J. B., Siangoli, F. B., Kelvin, A. A., Udahemuka, J. C., Kelvin, P., Flores, L., Kelvin, D. J., & Martinez, G. S. (2024). Complete Genome Sequencing, Annotation, and Mutational Profiling of the Novel Clade I Human Mpox Virus, Kamituga Strain. *Journal of Infection in Developing Countries*, *18*(4), 600–608. https://doi.org/10.3855/JIDC.20136
3. Vakaniaki, E. H., Kacita, C., Kinganda-Lusamaki, E., O'Toole, Á., Wawina-Bokalanga, T., Mukadi-Bamuleka, D., Amuri-Aziza, A., Malyamungu-Bubala, N., Mweshi-Kumbana, F., Mutimbwa-Mambo, L., Belesi-Siangoli, F., Mujula, Y.,

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

Parker, E., Muswamba-Kayembe, P. C., Nundu, S. S., Lushima, R. S., Makangara-Cigolo, J. C., Mulopo-Mukanya, N., Pukuta-Simbu, E., … Mbala-Kingebeni, P. (2024). Sustained human outbreak of a new MPXV clade I lineage in eastern Democratic Republic of the Congo. *Nature Medicine 2024*, 1–5. https://doi.org/10.1038/s41591-024-03130-3

4. Kinganda-Lusamaki, E., Amuri-Aziza, A., Fernandez-Nuñez, N., Makangara-Cigolo, J.-C., Pratt, C., Vakaniaki, E. H., Hoff, N. A., Luakanda-Ndelemo, G., Akil-Bandali, P., Nundu, S. S., Mulopo-Mukanya, N., Ngimba, M., Modadra-Madakpa, B., Diavita, R., Paku-Tshambu, P., Pukuta-Simbu, E., Merritt, S., O'Toole, Á., Low, N., … Ahuka-Mundeke, S. (2024). Clade I mpox virus genomic diversity in the Democratic Republic of the Congo, 2018–2024: Predominance of zoonotic transmission. *Cell*, *0*(0). https://doi.org/10.1016/J.CELL.2024.10.017
5. Murhula Masirika, L., Udahemuka, J. C., Ndishimye, P., Sganzerla Martinez, G., Kelvin, P., Bubala Nadine, M., Kitwanda Steeven, B., Kumbana Mweshi, F., Mutimbwa Mambo, L., Oude Munnink, B. B., Bengehya Mbiribindi, J., Belesi Siangoli, F., Lang, T., Malekani, J. M., Aarestrup, F., Koopmans, M., Schuele, L., Musabvimana, J. P., Umutoni, B., … Flores Girona, L. (2024). Epidemiology, clinical characteristics, and transmission patterns of a novel Mpox (Monkeypox) outbreak in eastern Democratic Republic of the Congo (DRC): an observational, cross-sectional cohort study. *MedXriv*.
6. Nzoyikorera, N., Nduwimana, C., Schuele, L., Nieuwenhuijse, D. F., Koopmans, M., Otani, S., Aarestrup, F. M., Ihorimbere, T., Niyomwungere, D., Ndihokubwayo, A., Diawara, I., Niyomwungere, A., Nizigiyimana, D., Uwineza, M. N., Oude Munnink, B. B., & Nyandwi, J. (2024). Monkeypox Clade Ib virus introduction into Burundi: first findings, July to mid-August 2024. *Euro Surveillance□: Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin*, *29*(42). https://doi.org/10.2807/1560-7917.ES.2024.29.42.2400666
7. Likos, A. M., Sammons, S. A., Olson, V. A., Frace, A. M., Li, Y., Olsen-Rasmussen, M., Davidson, W., Galloway, R., Khristova, M. L., Reynolds, M. G., Zhao, H., Carroll, D. S., Curns, A., Formenty, P., Esposito, J. J., Regnery, R. L., & Damon, I. K. (2005). A tale of two clades: Monkeypox viruses. *Journal of General Virology*, *86*(10). https://doi.org/10.1099/vir.0.81215-0
8. Molteni, C., Forni, D., Cagliani, R., Mozzi, A., Clerici, M., & Sironi, M. (2023). Evolution of the orthopoxvirus core genome. *Virus Research*, *323*. https://doi.org/10.1016/j.virusres.2022.198975
9. Laurenson-Schafer, H., Sklenovská, N., Hoxha, A., Kerr, S. M., Ndumbi, P., Fitzner, J., Almiron, M., de Sousa, L. A., Briand, S., Cenciarelli, O., Colombe, S., Doherty, M., Fall, I. S., García-Calavaro, C., Haussig, J. M., Kato, M., Mahamud, A. R., Morgan, O. W., Nabeth, P., … le Polain de Waroux, O. (2023). Description of the first global outbreak of mpox: an analysis of global surveillance data. *The Lancet Global Health*, *11*(7). https://doi.org/10.1016/S2214-109X(23)00198-5
10. Nzoyikorera, N., Nduwimana, C., Schuele, L., Nieuwenhuijse, D. F., Koopmans, M., Otani, S., Aarestrup, F. M., Ihorimbere, T., Niyomwungere, D., Ndihokubwayo, A., Diawara, I., Niyomwungere, A., Nizigiyimana, D., Uwineza, M. N., Oude Munnink, B. B., & Nyandwi, J. (2024). Monkeypox Clade□Ib virus introduction into Burundi: first findings,

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

July to mid-August 2024. *Euro Surveillance□: Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin*, *29*(42), 2400666. https://doi.org/10.2807/1560-7917.ES.2024.29.42.2400666/CITE/REFWORKS

11. https://africacdc.org/download/mpox-continental-preparedness-and-response-plan-for-africa/
12. https://www.who.int/news/item/14-08-2024-who-director-general-declares-mpox-outbreak-a-public-health-emergency-of-international-concern
13. Hutson, C. L., Abel, J. A., Carroll, D. S., Olson, V. A., Braden, Z. H., Hughes, C. M., Dillon, M., Hopkins, C., Karem, K. L., Damon, I. K., & Osorio, J. E. (2010). Comparison of West African and Congo Basin monkeypox viruses in BALB/c and C57BL/6 mice. *PLoS ONE*, *5*(1). https://doi.org/10.1371/journal.pone.0008912
14. Gupta, A. K., Talukder, M., Rosen, T., & Piguet, V. (2023). Differential Diagnosis, Prevention, and Treatment of mpox (Monkeypox): A Review for Dermatologists. In *American Journal of Clinical Dermatology* (Vol. 24, Issue 4). https://doi.org/10.1007/s40257-023-00778-4
15. Hutson, C. L., Carroll, D. S., Self, J., Weiss, S., Hughes, C. M., Braden, Z., Olson, V. A., Smith, S. K., Karem, K. L., Regnery, R. L., & Damon, I. K. (2010). Dosage comparison of Congo Basin and West African strains of monkeypox virus using a prairie dog animal model of systemic orthopoxvirus disease. *Virology*, *402*(1). https://doi.org/10.1016/j.virol.2010.03.012
16. Hendrickson, R. C., Wang, C., Hatcher, E. L., & Lefkowitz, E. J. (2010). Orthopoxvirus genome evolution: The role of gene loss. *Viruses*, *2*(9). https://doi.org/10.3390/v2091933
17. Girgis, N. M., DeHaven, B. C., Xiao, Y., Alexander, E., Viner, K. M., & Isaacs, S. N. (2011). The Vaccinia Virus Complement Control Protein Modulates Adaptive Immune Responses during Infection. *Journal of Virology*, *85*(6). https://doi.org/10.1128/jvi.01474-10
18. Senkevich, T. G., Yutin, N., Wolf, Y. I., Koonin, E. v., & Moss, B. (2021). Ancient gene capture and recent gene loss shape the evolution of orthopoxvirus-host interaction genes. *MBio*, *12*(4). https://doi.org/10.1128/mBio.01495-21
19. Henderson, D. A. (2011). The eradication of smallpox - An overview of the past, present, and future. *Vaccine*, *29*(SUPPL. 4). https://doi.org/10.1016/j.vaccine.2011.06.080
20. Li, E., Guo, X., Hong, D., Gong, Q., Xie, W., Li, T., Wang, J., Chuai, X., & Chiu, S. (2023). Duration of humoral immunity from smallpox vaccination and its cross-reaction with Mpox virus. *Signal Transduction and Targeted Therapy*, *8*(1). https://doi.org/10.1038/s41392-023-01574-6
21. Kumar, A., Dutt, M., Dehury, B., Sganzerla Martinez, G., Singh, K. P., Kelvin, D. J. (2024). Formulation of next-generation polyvalent vaccine candidates against three important poxviruses by targeting DNA-dependent RNA polymerase using an integrated immunoinformatics and molecular modeling approach. *Journal of Infection and Public Health*, *17*(7).
22. Martinez, G. S., Dutt, M., Kumar, A., & Kelvin, D. (2023). PoxiPred: An artificial intelligence-based method for the prediction of potential antigens and epitopes to accelerate vaccine development efforts against poxviruses. *Biology*, 13(2).
23. Dutt, M., Kumar, A., Rout, M., Dehury, B., Martinez, G., Ndishimye, P., Kelvin, A. A., & Kelvin, D. J. (2023). Drug repurposing for Mpox: Discovery of small mole-

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint **(which was not certified by peer review)** is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.

It is made available under a CC-BY 4.0 International license .

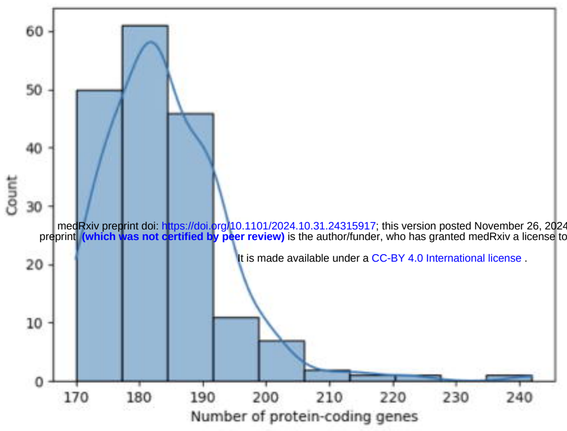
cules as potential inhibitors against DNA-dependent RNA polymerase using molecular modeling approach. *Journal of Cellular Biochemistry*, *124*(5).

24. Young, B., Seifert, S. N., Lawson, C., & Koehler, H. (2024). Exploring the genomic basis of Mpox virus-host transmission and pathogenesis. *MSphere*. https://doi.org/10.1128/MSPHERE.00576-24
25. Girgis, N. M., DeHaven, B. C., Xiao, Y., Alexander, E., Viner, K. M., & Isaacs, S. N. (2011). The Vaccinia Virus Complement Control Protein Modulates Adaptive Immune Responses during Infection. *Journal of Virology*, *85*(6). https://doi.org/10.1128/jvi.01474-10
26. Hudson, P. N., Self, J., Weiss, S., Braden, Z., Xiao, Y., Girgis, N. M., Emerson, G., Hughes, C., Sammons, S. A., Isaacs, S. N., Damon, I. K., & Olson, V. A. (2012). Elucidating the role of the complement control protein in monkeypox pathogenicity. *PLoS ONE*, *7*(4). https://doi.org/10.1371/journal.pone.0035086
27. Brennan, G., Stoian, A. M. M., Yu, H., Rahman, M. J., Banerjee, S., Stroup, J. N., Park, C., Tazi, L., & Rothenburg, S. (2023). Molecular Mechanisms of Poxvirus Evolution. In *mBio* (Vol. 14, Issue 1). https://doi.org/10.1128/mbio.01526-22
28. Wawina-Bokalanga, T., Akil-Bandali, P., Kinganda-Lusamaki, E., Lokilo, E., Jansen, D., Amuri-Aziza, A., Makangara-Cigolo, J. C., Pukuta-Simbu, E., Ola-Mpumbe, R., Muyembe, M., Kacita, C., Paku-Tshambu, P., Dantas, P. H., Tshiani-Mbaya, O., Luakanda, G., Nkuba-Ndaye, A., Matondo, M., Vakaniaki, E. H., Tessema, S., … Mbala-Kingebeni, P. (2024). Co-circulation of monkeypox virus subclades Ia and Ib in Kinshasa Province, Democratic Republic of the Congo, July to August 2024. Euro Surveillance : Bulletin Europeen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin, 29(38), 2400592. https://doi.org/10.2807/1560-7917.ES.2024.29.38.2400592/CITE/REFWORKS
29. Wang, S., Sundaram, J. P., & Spiro, D. (2010). VIGOR, an annotation program for small viral genomes. BMC Bioinformatics, 11, 451. https://doi.org/10.1186/1471-2105-11-451
30. Emms, D. M., & Kelly, S. (2019). OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biology*, *20*(1). https://doi.org/10.1186/s13059-019-1832-y
31. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3). https://doi.org/10.1016/S0022-2836(05)80360-2
32. Buller, R. M. L., & Palumbo, G. J. (1991). Poxvirus pathogenesis. *Microbiological Reviews*, *55*(1). https://doi.org/10.1128/mr.55.1.80-122.1991
33. Molteni, C., Forni, D., Cagliani, R., Mozzi, A., Clerici, M., & Sironi, M. (2023). Evolution of the orthopoxvirus core genome. *Virus Research*, *323*. https://doi.org/10.1016/j.virusres.2022.198975

MPXV Clade I

181 genomes

A



Statistics
(n. of protein-coding genes)

min = 170
max = 242
mean = 184.28
median = 183
standard deviation = 9.99

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license .

Clade I Core Genome

87 protein-coding genes are shared by at least 99% of the 181 genomes

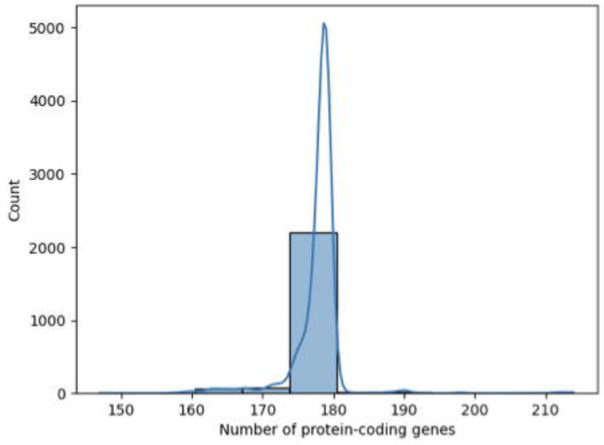
B

MPXV Clade II

2390 genomes

Statistics
(n. of protein-coding genes)

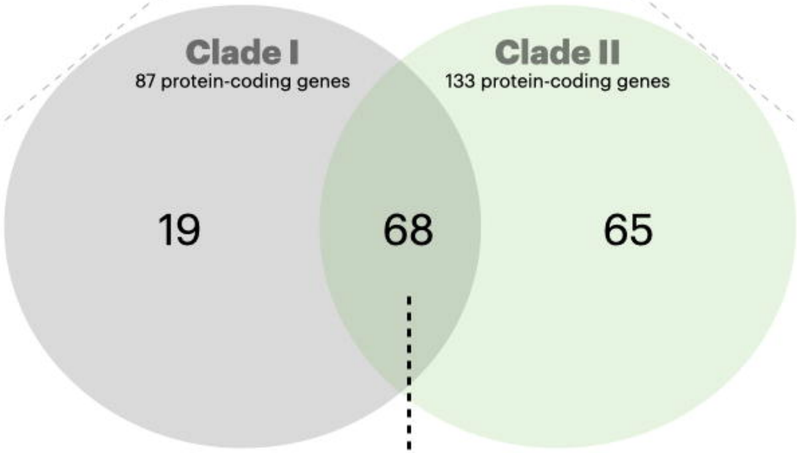
min = 147
max = 214
mean = 177.84
median = 179
standard deviation = 3.93



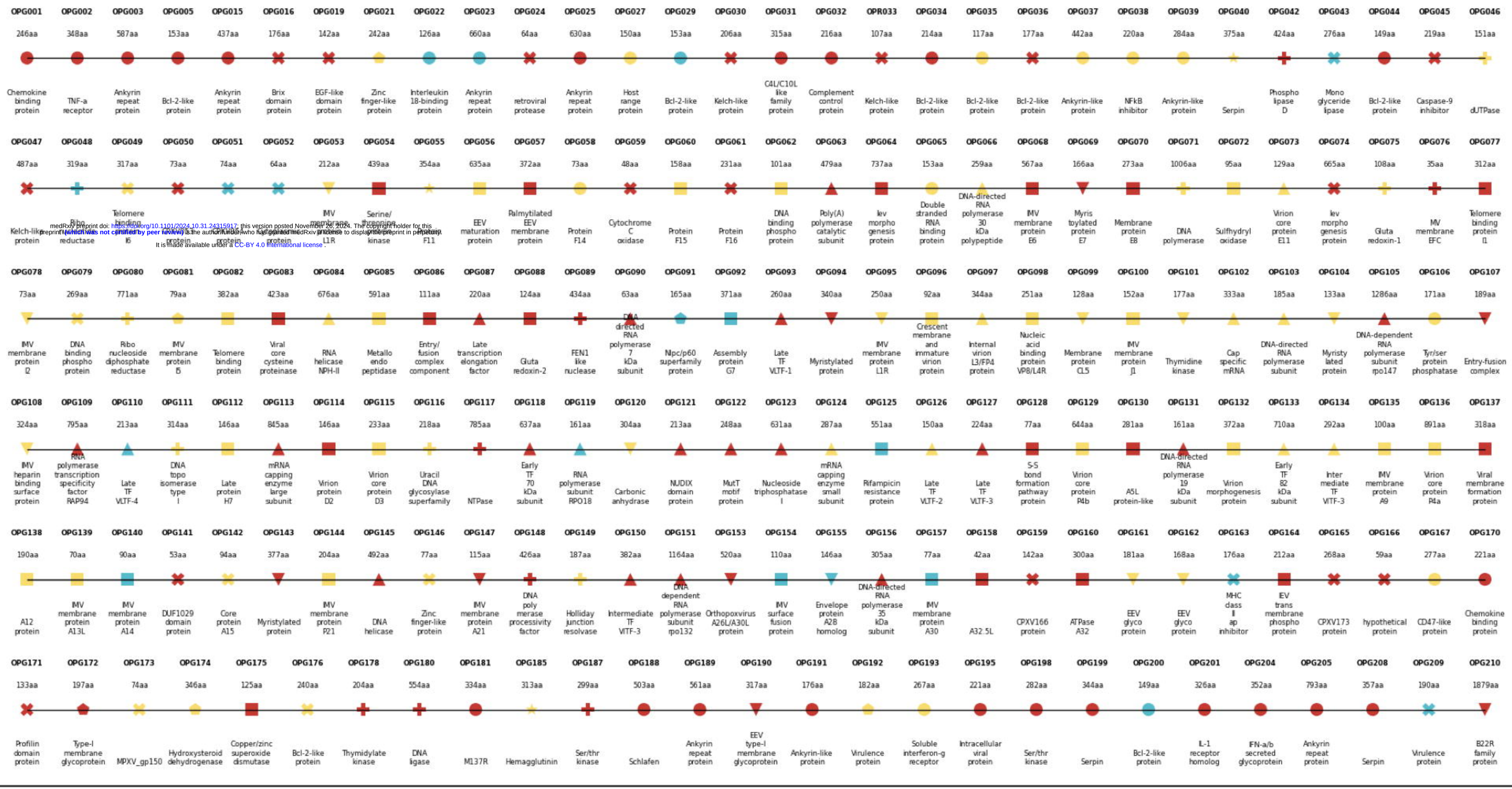
Clade II Core Genome

133 protein-coding genes are shared by at least 99% of the 2390 genomes

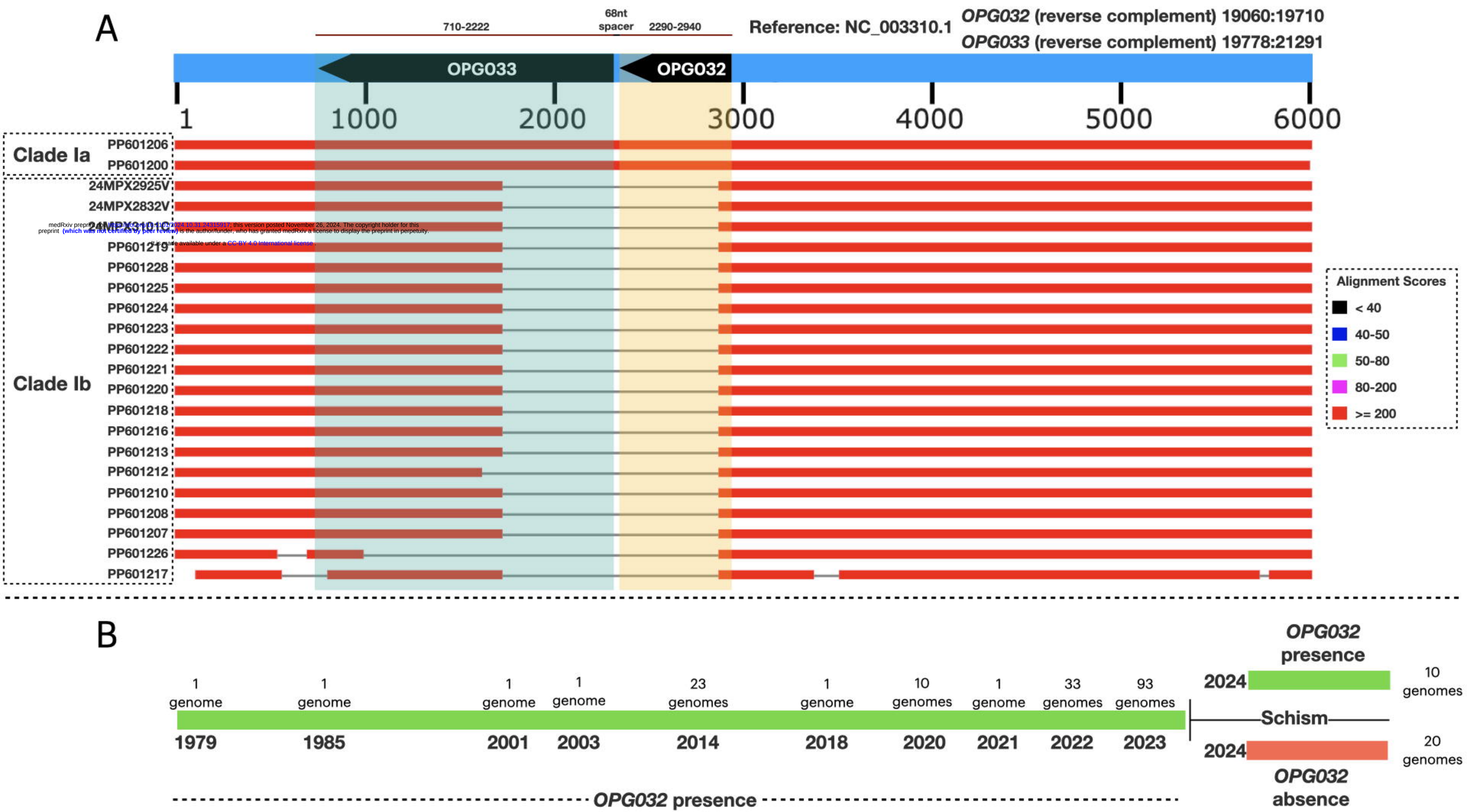
C



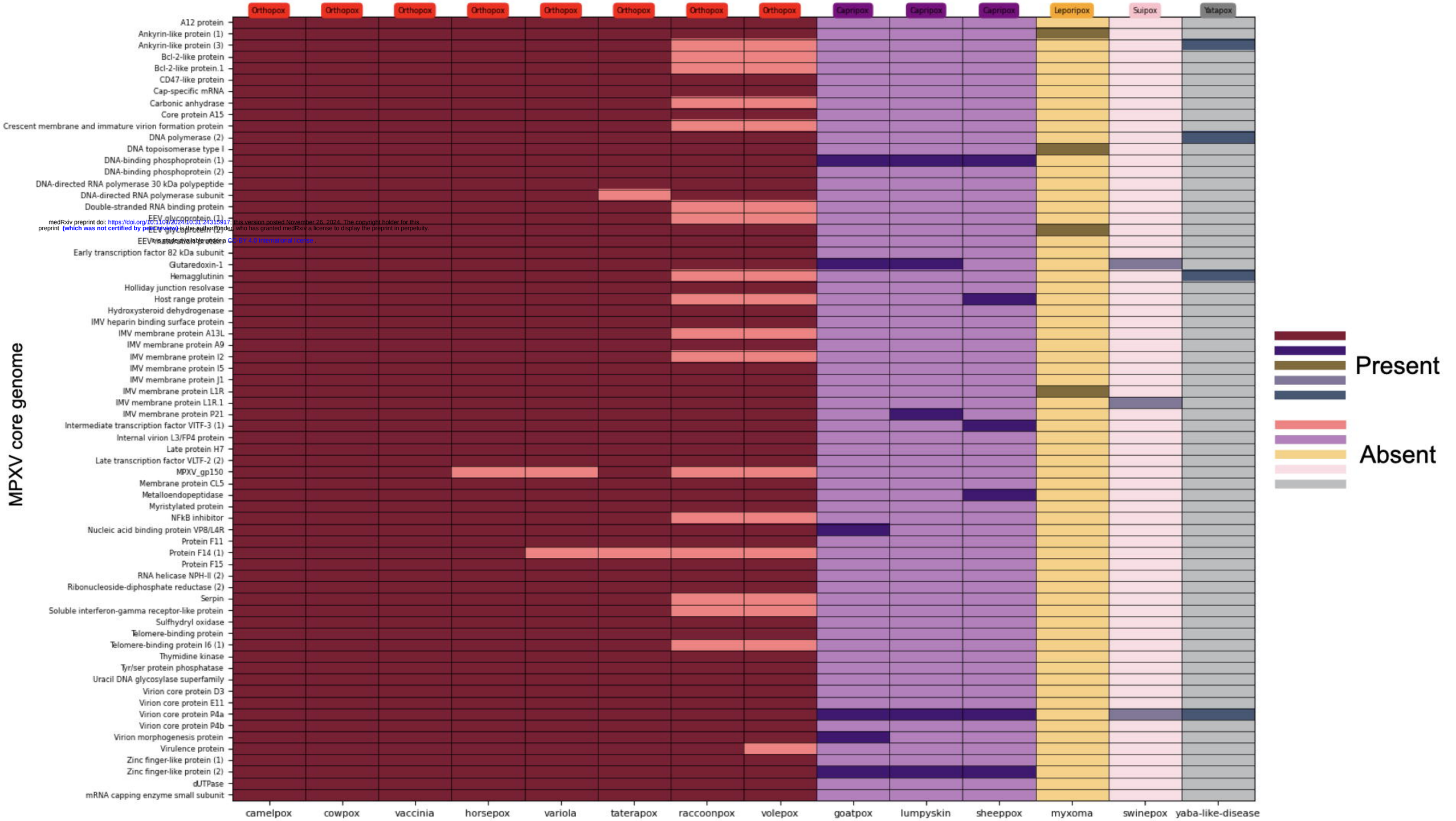
MPXV Clades I & II Core Genome



medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.



medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.
It is made available under a CC-BY 4.0 International license.



medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 26, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY 4.0 International license.