medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity.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

Authors

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]

Affiliations

[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

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.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 2, 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 .

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

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

[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

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

Abstract

Mpox, formerly monkeypox, is a viral zoonotic disease caused by the monkeypox virus (MPXV). MPXV, which is taxonomically divided into clade 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. 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. Limited epidemiological data suggests clade Ib is less virulent than other clade I and clade II viruses. There currently is a lack of quantitated comparative genomic analyses to address potential zoonotic transmissibility and pathobiology between the two clades. 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 subsequential queried isolates of clade Ib featured the deletion of the *OPG032* gene, which was found present in all queried clade Ia genomes. *OPG032* encodes a vaccinia ortholog which encodes the vaccinia complement protein, an important virulence protein. The schism event in the genomic structure highlighted by an entire gene deletion suggests it plays an important role in the pathobiology of the novel clade Ib. We anticipate our results lay the groundwork for exploiting the genomic of elements of MPXV as potential targets for therapeutics development/repurposing, vaccine design, and molecular diagnostic expansion as well to uncover the viral diversity and function of MPXV.

Keywords: Mpox; MPXV; Mpox clade Ib outbreak; Pangenome; Poxviruses.

INTRODUCTION

Monkeypox virus (MPXV), the causative agent for 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 [7] and since then, several outbreaks with human-to-human transmission have 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) [10]. The World Health Organization (WHO) declared a Public Health Emergency of International Concern (PHEIC) in 2022 as MPXV clade IIb rapidly

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

spread to non-endemic countries and territories (n=116) where human infections were confirmed [11]. Later, in September 2023, a new MPXV clade Ib emerged in Kamituga, South Kivu, DRC [2]. This clade is currently expanding to Eastern Africa nations [3,4,6,12]. Sustained human-to-human transmission and international spread of the novel clade Ib virus to DRC neighboring countries led the Africa Centres for Disease Control and Prevention (Africa CDC), on 13 August 2024, to declare mpox as a Public Health Emergency of Continental Security [13] just before the WHO's declaration of the outbreak as a second mpox-related PHEIC in less than three years.

Mpox viruses from both clades give rise to similar symptoms. Infected individuals typically develop smallpox-like symptoms including fever and rash that progresses to macules, papules, vesicles, pustules, and scabs, which generally begin on the face and spreading to other parts of the body. 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 [7-9]. The genomes of both clade I and II viruses are ~195 kilobase-pairs (kbp) in size and have a nucleotide-level sequence identity of 95%, with the 5% difference being attributed to immunomodulatory proteins [10]. Interestingly, 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].

This study presents a comparative genomic analysis of all available MPXV clade I and II fully annotated genomes, aimed at determining the genetic differences that underly zoonosis, transmissibility, and pathogenicity of MPXV clade Ib viruses. 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 fully annotated complete 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-

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

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 S2) and MPXV clade II (Figure 1-B, the list of protein-coding genes is found in 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.

Functional analysis of the core genome of MPXV (clades I and II)

We predicted the functions of the 68 protein-coding genes that comprise the core genome of MPXV clades I and II. Genes were classified as i) cell entry/attachment; ii) cell-to-cell fusion; iii) immunomodulation; iv) morphogenesis; v) transcription; vi) viral replication/repair; vii) virulence or viii) unknown, following the protein-coding gene classification scheme of [2] and in line with the function predicted in the reference genome NC003310.1. The most abundant functions that were assigned to different MPXV clades I and II core protein-coding genes were found to be morphogenesis (n=19), transcription (n=10), immunomodulation (n=10), and cell entry/attachment (n=9). Viral replication/repair, virulence, cell-to-cell fusion, and unknown functions had a total of 7, 4, 3, and 6 protein-core genes from the MPXV clades I and II core genome, respectively (Figure 2-A). A Gene Ontology (GO) search was then performed upon the shared protein-coding genes. Regarding molecular function (Figure 2-B), the four GO

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

terms associated with most genes were DNA binding (n=13), hydrolase activity (n=11), transferase activity (n=7), and catalytic activity acting on RNA (n=5). Regarding biological processes (Figure 2-C), the GO terms with most genes mapped to were DNA-templated transcription (n=5), programmed cell death (n=4), immunomodulation (n=4), and DNA repair (n=3). Finally, the GO terms with most hits related to the cellular component (Figure 2-D) of the genes were associated with the virion membrane (n=11), host intracellular membrane organelle (n=5), viral envelope (n=5), and host cell membrane (n=5). These results show that the protein-coding genes responsible for viral assembly and viral transcription are well represented in the genomes of MPXV clades I and II.

[Figure 2]

Figure 2 – Functional analysis of the 68 protein-coding genes in the core genome of MPXV clades I and II. (Figure 2-A) Each protein-coding gene was classified into cell entry/attachment (red), cell-to-cell fusion (blue), immunomodulation (violet), morphogenesis (yellow), transcription (pink), viral replication/repair (light green), virulence (dark green) or, unknown (grey) according to the classes of MPXV genes following their orthopoxvirus gene (OPG) nomenclature specified at [2]. We show the enrichment Gene Ontology (GO)-based quantitative classification of the 68 core protein-coding genes in MPXV clades I and II. Genes were quantified per their GO terms (level two) assignment in Molecular Function (Figure 2-B), Biological Process (Figure 2-C), and Cellular Component (Figure 2-D).

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

We searched for orthologs of all the 68 MPXV clades I and II core genes in 15 distinct chordopoxviruses and show their presence/absence in Figure 3. The most abundant genus, i.e., orthopoxviruses, had proteins of the core MPXV Clade I and II genome mapped to 9 species. All the 68 proteins with predicted functions that compose the core genome of MPXV clade 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,

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

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 6, 5, and 6 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, 4, 2, and 4 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 3]

Figure 3 – 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.

MPXV pangenome map

We created a pangenome map of MPXV (Figure 4) based on the OPG nomenclature associated with the annotation of MPXV clade Ia genome NC003310.1, i.e., the reference genome of MPXV clade I. The MPXV clade I reference genome 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.

[Figure 4]

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 – MPXV pangenome map. Figure 4 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). We also show the length in amino acids (aa) of each protein-coding gene as well as is predicted function (relative to NC003310.1).

OPG032 gene deletion highlighted the introduction of the novel MPXV clade Ib.

The deletion of the *OPG032* gene has been reported as a mark of the current MPXV clade Ib outbreak [2]. This gene encodes a vaccinia homolog, complement control protein, which has been classified as an important virulence factor both in vaccinia [22] and in MPXV [23]. A protein alignment of MPXV clade Ia (NC_003310) *OPG032* and its vaccinia homolog, i.e., vaccinia complement control protein (YP_232907 in the vaccinia genome NC_006998) indicated the two proteins share a sequence identity 95.83% with an e-value of 2.78e-154.

We considered nucleotide content of 195 MPXV clade I (195 clade Ia; 20 clade Ib) and queried the presence of the region that codifies the *OPG032* gene among them. A wider region of the reference clade I genome containing the *OPG032* coordinates (16000:22000) was aligned with DNA segments spanning the same range in 195 MPXV clade I genomes. In Figure 5-A, a low alignment score (<40) is shown for all 20 clade Ib sequences in the locus that encompasses the genomic coordinates of *OPG032* when compared to the reference genome and two random clade Ia sequences (PP601206, and PP601200). In addition, the 175 clade Ia sequences, when aligned to *OPG032*, indicated alignment scores >= 200 for the locus of *OPG032*, indicating the gene presence is common in the clade Ia genomes queried (Supplementary Material S4). Next, we queried all the 195 DNA sequences of clade I in terms of the presence or absence of the genomic region that covers *OPG032* (Figure 5-B) and determined that all the genomes deposited from 1979 to 2023 contained *OPG032*. Up to 2024, *OPG032* was part of the core genome of MPXV clade I, and after a schism event, clade I separated into clade Ia and Ib, at least in part, as a function of the presence or absence of *OPG032*. It will be important to see if MPXV clade Ib becomes a dominant strain of MPXV clade Ia strains when both lineages are cocirculating in the same geographical space as cocirculation has been reported [14] and rectified by our results. We looked for *OPG032* orthologs among the 15 queried chordopoxviruses to find the protein-coding gene is present on the orthopoxviruses camelpox, cowpox, ectromelia, horsepox, taterapox, vaccinia, and variola.

[Figure 5]

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 5 – Presence and absence of *OPG032* across all genomes deposited under the MPXV clade I taxonomy. In Figure 5-A, we show the alignment of the *OPG032* region (query) extracted from the reference clade I genome (NC003310) and the 20 clade Ib sequences (genome IDs from NCBI are included in the y-axis). The first two sequences on top of the panel (PP601206 and PP601200) are two random clade Ia sequences which have alignment scores >= 200 for the *OPG032* locus. Additional alignment containing all the 175 clade I sequences (175 clade Ia) is included in Supplementary Figure S4. In Figure 5-B, we show the distribution over time of MPXV clade I sequences containing and not containing the alignment region matching the locus of *OPG032*.

DISCUSSION

Here we show that the genomes of MPXV clades I and II share a core genome of 68 protein-coding genes, many of which are also found in other orthopoxviruses and partially in the genome of other chordopoxviruses. The function of the elements that compose the core genome of MPXV is associated with transcription, viral morphogenesis, and immunomodulation. We argue the differences observed in the genomic content of MPXV viruses might, in part, provide insights in the MPXV viral properties, the disease they cause, and their outcomes.

The coding capacity of MPXV clade I and II genomes varies from 147 to 242 protein-coding genes. We identified a total of 68 genes that are shared by most of the genomes of both clades, adding quantitative metrics to the conservation that has been reported among the clades [15]. While clade I was found to have 19 clade-unique genes, clade II had 65, adding more genetic diversity to the second clade. Considering we investigated a larger number of clade II sequences, the large number of proteins unique to clade II may be related to the increased samples of clade II. From the perspective of prevention, we suggest that efforts for developing vaccine components should consider the conserved elements of MPXV/orthopoxvirus/chordopoxvirus highlighted by our pangenome analysis. These could then be complemented with clade-specific attributes in case of insufficient conferred protection. Current vaccines conferring protection against MPXV are Imvamune, a live, non-replicating vaccinia virus vaccine; ACAM2000, a live, replication-competent vaccinia virus vaccine; and 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 the 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

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

vaccinated individuals in the population due to those born after 1980 not routinely receiving the smallpox vaccine [16], we might expect outbreaks in naive populations with less immunity against orthopoxviruses. Thus our pangenome analysis is beneficial for the identification of proteins with potential antigenic effect for the development of vaccines against MPXV such as [17], [18], and [19]. Furthermore, from a treatment perspective, our analysis can benefit therapeutic development for targets that are conserved across different isolates of MPXV. We identified 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 for inhibition with small molecules in a molecular modelling approach [20]. Taken together, the results of our core genome analyses have the potential to be used as foundational information for the development of vaccines and therapeutics against MPXV.

Our results show that the differential coding capacity of MPXV clades Ia and Ib might explain the lower fatality reported for clade Ib infections in humans. A major feature differentiating MPXV clade I genomes is the presence and absence of the *OPG032* gene in clade Ia and Ib, respectively. 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). This gene, in vaccinia, 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 [21]. An MPXV study [22] 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 incorporation of *OPG032* in a clade II virus did not accelerate clinical disease course nor affect disease mortality. Epidemiological data from the current outbreak of MPXV clade Ib in eastern DRC suggests the clade Ib viral infections in humans has 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. Although the genome of clade Ib contains several mutations not found in clade Ia viruses, the complete deletion of the *OPG032* gene is striking. The functional significance of *OPG032* in sustained human-to-human transmission and lowered mortality still needs to be elucidated but our results suggest that the pathobiology of the clade Ib outbreak might be, in part, driven by the

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

deletion of the *OPG032* gene, which was previously part of the core genome of MPXV clade I.

We have 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 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 demonstrate an overall conserved core of genes among *orthopoxviruses*, which might be exploited for developing therapeutics, vaccines, and molecular diagnosis reagents. Finally, by highlighting a functional schism between the MPXV clade Ia and Ib genomes – suggesting the presence or absence of the putative virulence factor *OPG032* – 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.

2,436 fully annotated complete MPXV genomes 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.

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

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 [14]. Genomes are available at https://github.com/inrb-labgenpath/DRC_MPXV_Genomic_Surveillance.

MPXV clade I genome annotation

The non-annotated MPXV clade I genomes (132 clade Ia obtained from the national surveillance program of INRB [14] and 3 clade Ib sequenced in this study) were manually annotated using the Java-based VIGOR4 (Viral Genome ORF Reader) [23] 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.

Orthologs 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) [24]. Additionally, to validate the identified orthologs, we used a reciprocal local implementation of the Basic Local Alignment Search Tool (version 2.15.0) [25] (protein-protein BLASTp). To determine whether a gene belongs 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 the 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

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

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*

We searched GenBank for the complete genome annotation of the 15 *chordopoxviruses* whose taxonomic classification is obtained from [26] and downloaded their protein-coding genes in translated amino acid format (Table 1).

Table 1 – Reference proteomes from isolates of the *Chordopoxvirinae* subfamily.

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

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

Swinepox	Suipoxvirus	146	GCF_000839965.1	300880
Myxoma	Leporipoxvirus	158	GCF_000843685.1	31530

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 in providing input in 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

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).

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

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-

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

Kumbana, F., Mutimbwa-Mambo, L., Belesi-Siangoli, F., Mujula, Y., 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, N., Makangara-Cigolo, J.-C., Pratt, C., Vakaniaki, E. H., Hoff, N. A., Luakanda, G., Akil-Bandali, P., Nundu, S. S., Mulopo-Mukanya, N., Ngimba, M., Modadra-Madakpa, B., Diavita, R., Paku, 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. *MedRxiv*, 2024.08.13.24311951. https://doi.org/10.1101/2024.08.13.24311951
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. 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
8. 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
9. Hutson, C. L., Olson, V. A., Carroll, D. D., Abel, J. A., Hughes, C. M., Braden, Z. H., Weiss, S., Self, J., Osorio, J. E., Hudson, P. N., Dillon, M., Karem, K. L., Damon, I. K., & Regnery, R. L. (2009). A prairie dog animal model of systemic orthopoxvirus disease using west African and Congo Basin strains of Monkeypox virus. *Journal of General Virology*, *90*(2). https://doi.org/10.1099/vir.0.005108-0
10. 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
11. https://doi.org/10.2807/1560-7917.ES.2024.29.42.2400666

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

12. 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).
13. https://www.who.int/emergencies/situations/monkeypox-oubreak-2022
14. https://africacdc.org/news-item/africa-cdc-declares-mpox-a-public-health-emergency-of-continental-security-mobilizing-resources-across-the-continent/
15. 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
16. Americo, J. L., Earl, P. L., & Moss, B. (2023). Virulence differences of mpox (monkeypox) virus clades I, IIa, and IIb.1 in a small animal model. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(8). https://doi.org/10.1073/pnas.2220415120
17. 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
18. 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).
19. 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. *BioRxiv*, 2023.08.04.551871. https://doi.org/10.1101/2023.08.04.551871
20. Zhang, R. R., Wang, Z. J., Zhu, Y. L., Tang, W., Zhou, C., Zhao, S. Q., Wu, M., Ming, T., Deng, Y. Q., Chen, Q., Jin, N. Y., Ye, Q., Li, X., & Qin, C. F. (2023). Rational development of multicomponent mRNA vaccine candidates against mpox. *Emerging Microbes and Infections*, *12*(1). https://doi.org/10.1080/22221751.2023.2192815
21. 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 molecules as potential inhibitors against DNA-dependent RNA polymerase using molecular modeling approach. *Journal of Cellular Biochemistry*, *124*(5). https://doi.org/10.1002/jcb.30397
22. 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

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

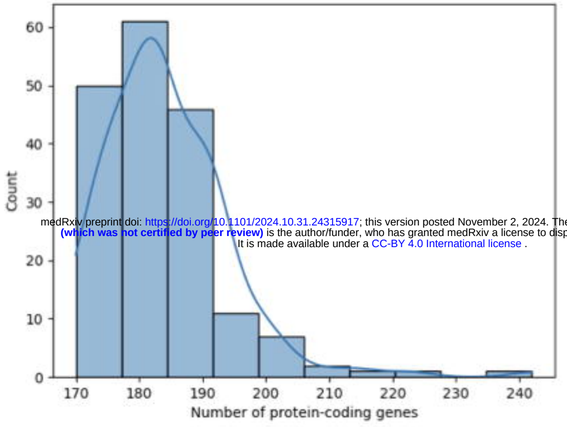
Responses during Infection. *Journal of Virology*, *85*(6). https://doi.org/10.1128/jvi.01474-10

23. 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
24. 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
25. 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
26. 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
27. 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

MPVX 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 2, 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

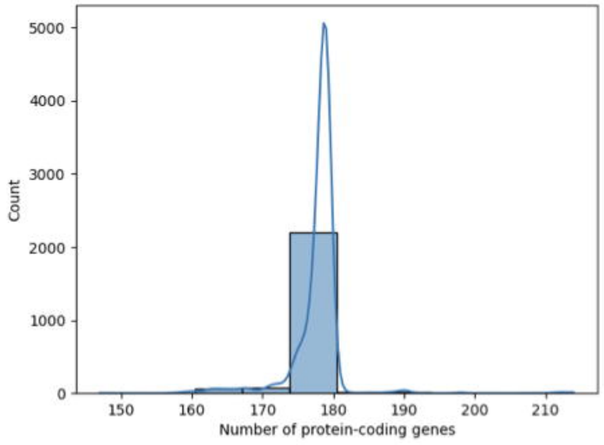
MPVX Clade II

2390 genomes

Statistics

(n. of protein-coding genes)

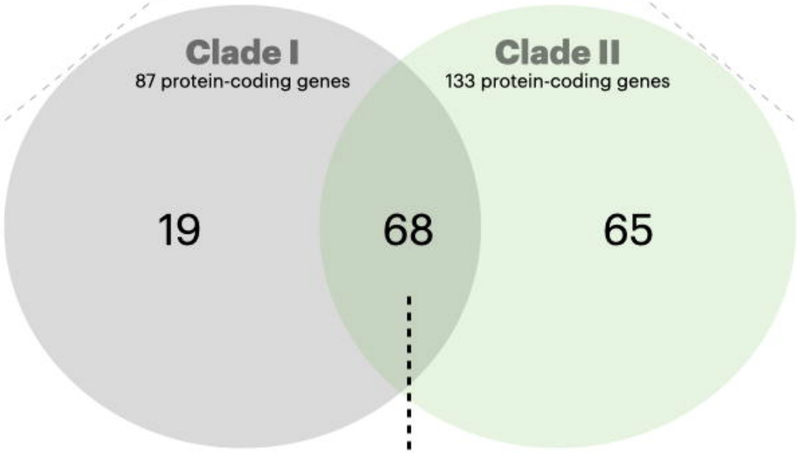
min = 147
max = 214
mean = 177.84
median = 179
standard deviation = 3.93



Clade II Core Genome

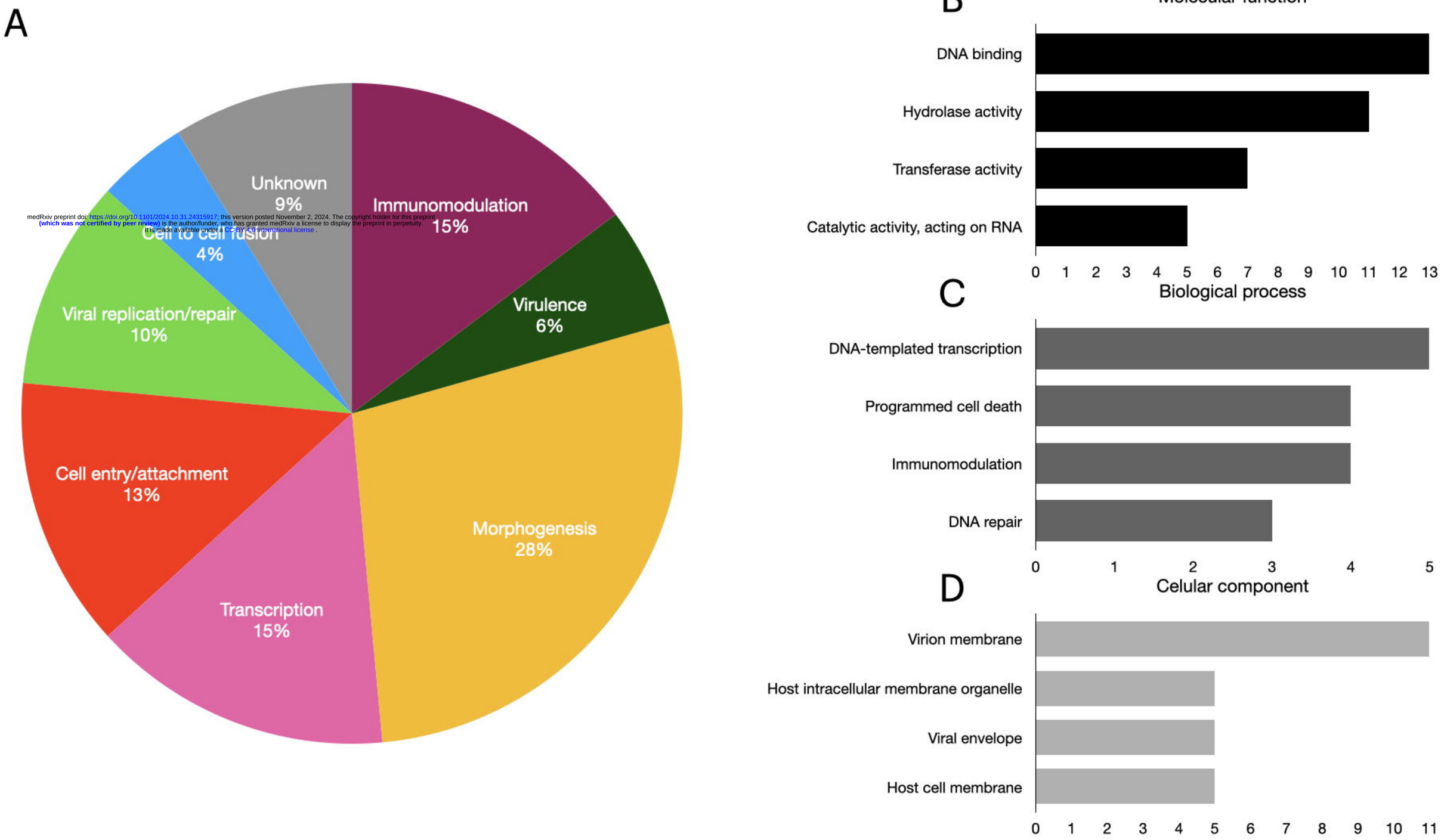
133 protein-coding genes are shared by at least 99% of the 2390 genomes

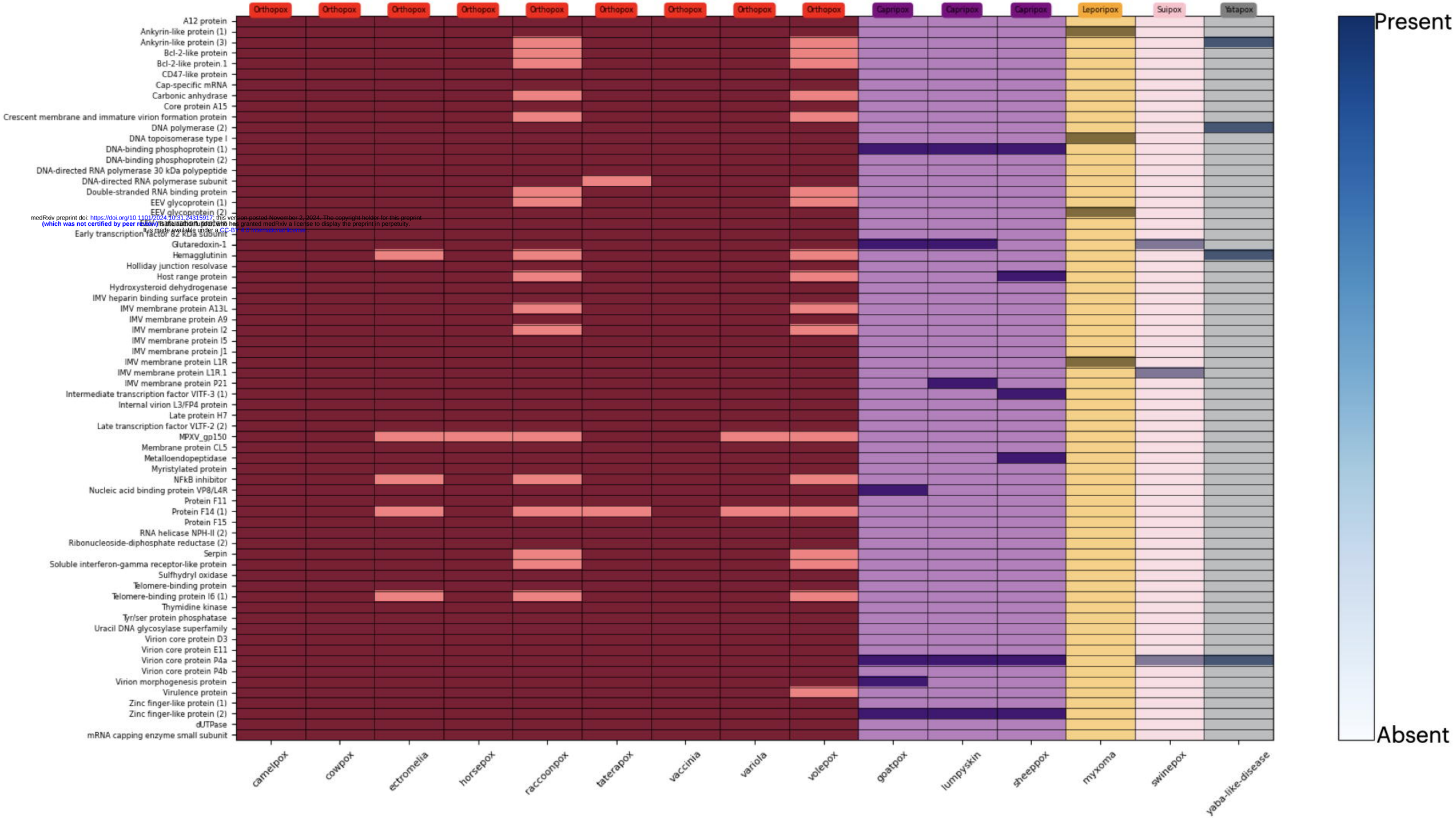
C



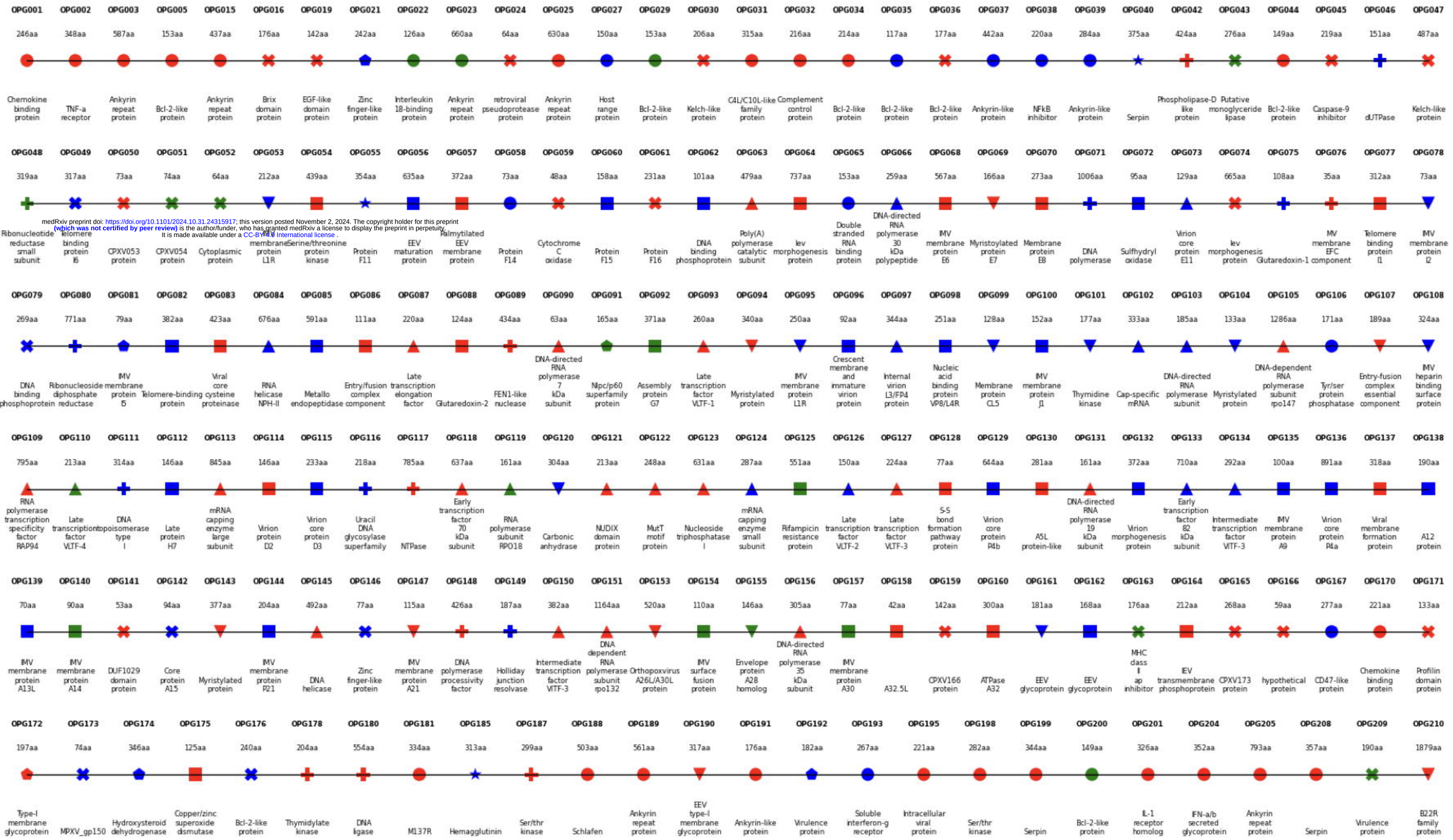
MPVX Clades I & II Core Genome

medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .

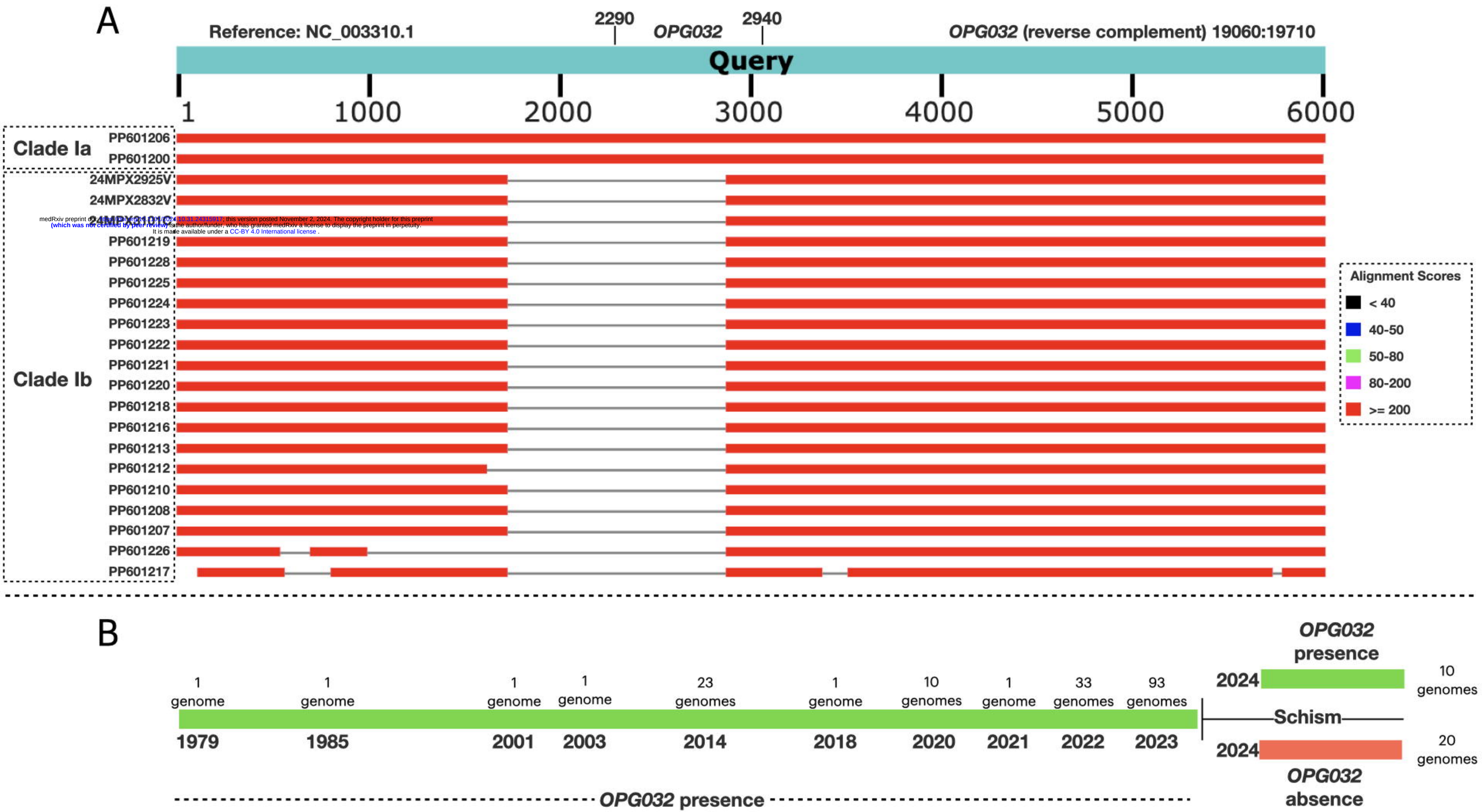




medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 2024. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license.



medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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-NC-ND 4.0 International license .



medRxiv preprint doi: https://doi.org/10.1101/2024.10.31.24315917; this version posted November 2, 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 .