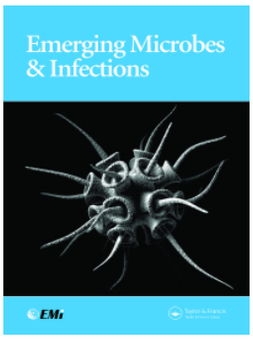

Emerging Microbes & Infections

ISSN: 2222-1751 (Online) Journal homepage: www.tandfonline.com/journals/temi20

Mapping transcriptional patterns of MPXV in human epithelial cells

Ludovica Picarone, Daniele Pietrucci, Davide Mariotti, Marco Milanesi, Cosmina Mija, Licia Bordi, Silvia Meschi, Valentina Mazzotta, Cesare Ernesto Maria Gruber, Carla Mavian, Enrico Girardi, Andrea Antinori, Giulia Matusali, Giovanni Chillemi & Fabrizio Maggi

To cite this article: Ludovica Picarone, Daniele Pietrucci, Davide Mariotti, Marco Milanesi, Cosmina Mija, Licia Bordi, Silvia Meschi, Valentina Mazzotta, Cesare Ernesto Maria Gruber, Carla Mavian, Enrico Girardi, Andrea Antinori, Giulia Matusali, Giovanni Chillemi & Fabrizio Maggi (2026) Mapping transcriptional patterns of MPXV in human epithelial cells, Emerging Microbes & Infections, 15:1, 2627079, DOI: 10.1080/22221751.2026.2627079

To link to this article: https://doi.org/10.1080/22221751.2026.2627079

Emerging Microbes & Infections
2026, VOL. 15, 2627079 (12 pages)
https://doi.org/10.1080/22221751.2026.2627079

Mapping transcriptional patterns of MPXV in human epithelial cells

Ludovica Picarone[a], Daniele Pietrucci[a], Davide Mariotti[b], Marco Milanesi[c], Cosmina Mija[b], Licia Bordi[b], Silvia Meschi[b], Valentina Mazzotta[d], Cesare Ernesto Maria Gruber[b], Carla Mavian[e], Enrico Girardi[f], Andrea Antinori[g], Giulia Matusali[b,h], Giovanni Chillemi[i,j] and Fabrizio Maggi[b]

[a]Department for Innovation in Biological, Agro-Food and Forest Systems (DIBAF), University of Tuscia, Viterbo, Italy; [b]Laboratory of Virology and Laboratories of Biosafety, National Institute for infectious Diseases "Lazzaro Spallanzani" – IRCCS, Rome, Italy; [c]Department of Animal Science, Food and Nutrition – DIANA, Università Cattolica del Sacro Cuore, Piacenza, Italy; [d]Counseling, Test and HIV Prophylaxis and STI Unit, Regional AIDS Reference Center, National Institute for infectious Diseases "Lazzaro Spallanzani" – IRCCS, Rome, Italy; [e]Emerging Pathogens Institute, University of Florida, Gainesville, FL, USA; [f]Scientific Direction, National Institute for Infectious Diseases "Lazzaro Spallanzani" – IRCCS, Rome, Italy; [g]Health Direction, National Institute for infectious Diseases "Lazzaro Spallanzani" – IRCCS, Rome, Italy; [h]Department of Biology and Biotechnology "Charles Darwin", Sapienza University of Rome, Rome, Italy; [i]Bioinformatics Research Unit in Infectious Diseases, National Institute for infectious Diseases "Lazzaro Spallanzani" – IRCCS, Rome, Italy; [j]Department of Experimental Medicine, University of Rome "Tor Vergata", Rome, Italy

ABSTRACT

The 2022 global outbreak of the monkeypox virus (MPXV) clade IIb highlighted changes in transmission dynamics, clinical features, and tissue tropism, yet the impact of host cell origin on viral transcription remains unclear. We performed comparative transcriptomic profiling of MPXV clade IIb in human epithelial cell lines from vaginal, intestinal, ectocervical, and renal tissues. Temporal analysis in vaginal cells revealed two major transcriptional clusters: early genes enriched for host–virus interaction factors in terminal genome regions, and intermediate–late genes encoding transcription, replication, and morphogenesis functions in the core genome. Some genes were expressed at each time point and comprised transcripts encoding signal transduction and inflammation-modulating factors. Single-time point comparisons linked higher viral particle production in vaginal and intestinal cells to intermediate–late gene enrichment, whereas ectocervical and renal cells favoured host–virus interaction transcripts. Several genes, including those that modulate signal transduction pathways, were highly expressed across different cell types. These findings reveal cell-type-dependent modulation of MPXV transcription and identify conserved and variable viral factors that may inform antiviral strategies.

ARTICLE HISTORY Received 10 September 2025; Revised 3 January 2026; Accepted 2 February 2026

KEYWORDS Monkeypox virus; MPXV; viral transcriptome; RNA-seq; hierarchical transcription

Introduction

The monkeypox virus (MPXV) belongs to the Poxviridae family in the Orthopoxvirus genus, comprising other viral pathogens such as variola, vaccinia, and cowpox viruses. Orthopoxviruses' genome is a linear, double-stranded DNA molecule, approximately 197 kilobases in length, containing over 190 open reading frames (ORFs). The genome includes a highly conserved central region, encoding viral transcription, replication, and virion morphogenesis factors, and two peripheral regions containing immune modulating and virulence genes implicated in the determination of the host range and pathogenesis [1,2]. Gene naming was initially based on HindIII restriction enzyme digestion profiles and was later unified under a standardized Orthopoxvirus nomenclature to avoid misinterpretation caused by different reference genomes [3]. Viral replication takes place in the cytoplasm, and a hierarchical transcription of early, intermediate, and late genes has been previously described [4]. The unprecedented 2022 global outbreak caused by MPXV clade IIb lineage B.1, which led to over 100.000 mpox cases worldwide, and the recent upsurge of mpox cases due to clades I and II in the Democratic Republic of Congo (DRC) and neighbouring countries [5], underlined the importance of studying basic biology and viral evolution dynamics of Orthopoxviruses. Importantly, the clade IIb lineage B.1 is characterized by a much higher reproduction number in humans compared to previous strains [6]. Moreover, data from the 2022 outbreak and recent reports from DRC revealed changes in transmission route with the predominant role of sexual transmission in viral spread [7]. The analysis of viral genomes is essential for understanding evolutionary events potentially affecting the interplay of

CONTACT Giulia Matusali giulia.matusali@uniroma1.it; Giovanni Chillemi giovanni.chillemi@inmi.it
Supplemental data for this article can be accessed online at https://doi.org/10.1080/22221751.2026.2627079.

the virus with its host. In the Orthopoxviruses' genome, core genes involved in transcription, replication, and morphogenesis and accessory genes responsible for host-virus interaction are usually described [3]. The analysis of genome expression profile is yet more informative and may be influenced by the host species, genetic background, and the anatomical and cellular environment.

Recent studies have focused on genome evolution and on comparative analyses of viral and host gene expression upon infection by different Orthopoxviruses and MPXV clades [3,6,8–15] to understand the molecular basis of the similarity and divergence in host range, transmissibility, and pathogenesis observed for these viruses and their variants. Genome evolution studies have described gene gains and losses, as well as diversified selection pressure at specific sites in coding sequences for immunomodulatory or core proteins [8–12].

Early comparative analyses of Orthopoxvirus genomes, transcripts, and proteins, dating before 2022, evidenced the impact of some specific factors on the host range, pathogenesis, and virulence [13–15]. One characteristic feature of the MPXV clade II is the absence of the viral complement control protein MOPICE (C3L or OPG032) concurring with the attenuated pathogenicity observed for this clade [16]. More recently, omics approaches have been applied to investigate the genomic basis of the increased person-to-person transmission observed in clade IIb, and the differences in the pathogenesis among clades [17,18].

The temporal expression definition of Orthopoxvirus genomes has been challenging due to the high number of viral genes, limiting the efficacy of classical molecular biology approaches, the very close timing of expression of intermediate and late genes, and extensive readthroughs of many transcripts [4]. Moreover, information on viral expression kinetics comes mostly from Vaccinia virus (VACV) studies and has been subsequently applied to MPXV with the assumption that orthologs in MPXV and VACV are expressed in the same temporal classes [4]. Studies employing microarray technologies [14] provided only partial insight, focusing on a subset of viral genes, while a more recent long-read sequencing provided high-resolution maps of viral transcript structure in an MPXV infected African green monkey cell line [19]. One study was carried out analyzing viral transcripts upon infection of human cells with MPXV clade I and VACV, and only one report described the viral protein dynamics in MPXV IIb infected human fibroblasts, mapping phosphosites and suggesting a role for MAPK pathways in the regulation of the viral life cycle [20].

To profile the temporal expression of viral transcripts upon MPXV clade IIb infection in epithelial cells (the first for MPXV infection and viral replication) of human genital origin, we applied a transcriptomic approach and analyzed temporal clusters and gene expression trajectories coupled to up-to-date viral transcript functional annotations.

Finally, the analysis of viral dynamics in MPXV-infected patients from the 2022 outbreak showed the high prevalence of virus-induced lesions and marked viral replication in the anogenital area as well as differential patterns of viral shedding from different bodily fluids [21,22]. Animal studies underlined the role of the infection site in the modulation of viral dissemination, showing that inoculation through rectal and vaginal mucosae led to increased shedding and replication [23].

To the best of our knowledge, none of the prior investigations addressed how tissue-specific cellular contexts – such as epithelial cells from different anatomical sites – might influence viral transcriptional timing, abundance, or gene expression kinetics. To address this gap, here we compared the viral transcription profile in four human epithelial cell lines from different anatomical districts; namely vaginal, intestinal, ectocervical, and renal tissues; and showed differences in patterns of viral transcript expression and susceptibility to infection.

Materials and methods

Cells and viral stock preparation

Vero E6 (immortalized kidney epithelial cells from the African green monkey) and Caco-2 (colorectal adenocarcinoma cells, ATCC) were maintained in Eagle's Modified Essential Medium (E-MEM; Corning), containing glutamine (2 mM), 1% penicillin/streptomycin, and supplemented with 10% and 20% Fetal Bovine Serum (FBS; Euroclone), respectively. VK2/E6E7 (human epithelial vaginal cells, ATCC) and Ect1/E6E7 (ectocervical epithelial cells, ATCC) were maintained in Keratinocyte-Serum Free medium (K-SFM) with 0.1 ng/ml human recombinant EGF, 0.05 mg/ml bovine pituitary extract (BPE), and additional calcium chloride (0.4 mM; Gibco). HK-2 (epithelial kidney cells, ATCC) were cultured in K-SFM with 5 ng/ml EGF and 0.05 mg/ml BPE. All cell lines were grown at 37°C in a humidified 5% CO_2 incubator.

Preparation of viral stock was performed on Vero E6 cells, infected with the isolate of the MPXV virus, strain hMpxV/Italy/INMI-Pt2/2022, clade/lineage IIb B.1 (GISAID: EPI_ISL_13251120, Gen-Bank: ON745215.1), obtained from a skin lesion of an MPXV-infected patient. After three freezing-thawing cycles, cell lysates were cleared, aliquoted, and stored at −80°C. Virus titration was performed on Vero E6 cells with a limiting dilution assay, and the results were reported as a 50% tissue culture infectious dose ($TCID_{50}$/ml). All procedures described below involving infectious MPXV were performed in a Biosafety Level 3 (BSL3) facility at the National Institute for Infectious Diseases "Lazzaro Spallanzani" (INMI) in Rome.

MPXV infection

VK2/E6E7, Ect1/E6E7, HK-2, and Caco-2 cells were plated in 6-well plates, cultured for 24 h at 37°C and 5% CO_2. Cells were then infected with hMpxV/Italy/un-INMI-Pt2/2022 Clade IIb isolate at a multiplicity of infection (MOI) of 0.1. After 1 h and 30 min of incubation, the virus was discarded, cells were rinsed twice with phosphate-buffered saline (DPBS) and cultured in either complete KSFM or E-MEM supplemented with 2% FBS, P/S and Gln. Supernatants were used to measure viral infectivity by back-titration, while total RNA was extracted from cells for transcriptomic analysis.

Viral infectivity

Viral infectivity was established in supernatants from VK2/E6E7, Ect1/E6E7, HK-2, and Caco-2 that were exposed to MPXV for up to 48 h, using a limiting dilution assay. More in detail, supernatants were serially diluted in E-MEM 2% FBS, added to Vero E6 cells in a 96-well plate, and incubated at 37°C and 5% CO_2 for 6 days. The cytopathic effect (CPE) occurrence was observed by light microscope, and viral titre was calculated according to the Reed and Muench method, and expressed as tissue culture infectious dose per ml ($TCID_{50}$/ml).

RNA extraction and sequencing

Total RNA was extracted from MPXV-infected cells at 2, 24, and 48 h post-infection (hpi) from the VK2/E6E7 and at 48 hpi from the other cell lines, using Quick-RNA MicroPrep kit (Zymo Research Corporation, Irvine, CA, USA) and eluted in 15 µl of Elution buffer. A DNase I treatment step was included in the extraction protocol. RNA concentration was measured by NanoDrop™ 2000/2000c Spectrophotometers (Thermo Scientific, Italy). RNA samples were processed to remove ribosomal RNA. Stranded Total RNA libraries were then prepared and sequenced on the Illumina NovaSeq 6000 platform using paired-end 2 × 100 cycles. The average number of reads per sample was 14,980,800, ranging from a minimum of 9,235,664 to a maximum of 18,246,167.

Aligning reads to the MPXV and host genome

Short sequencing reads were quality checked using FastQC v0.11.9 [24], and then Trimmomatic v0.39 [25] was used to filter and trim the sequences. Filtered reads were aligned using STAR v2.7.9a [26] to the human (GRCh38/hg38) and the MPXV reference genomes (GCF_014621545.1) using the RefSeq annotation file (NC_063383.1). Gene quantification was performed using FeatureCounts v2.0.8 for both human and virus genes [27] using the aforementioned annotation file.

Count normalization and cluster analysis

RNA-seq data analysis was conducted using R v4.2.2. For MPXV expression analysis, raw counts for each time point in the VK2/E6E7 cell line were normalized using the DESeq2 package v1.38.3 and the method "median of ratios" [28]. To assess the robustness of normalization, viral gene expression values obtained using DESeq2 median-ratio normalization were compared with length-normalized TPM values. DESeq2 rlog-transformed counts and log2(TPM + 1) values were used for this comparison to stabilize variance and ensure comparability between normalization approaches on comparable log-like scales. This analysis was performed both for the VK2/E6E7 time-course dataset and for viral gene expression profiles across multiple cell lines at 48 hpi (Supplementary Figure 1) and showed comparable trends, supporting the robustness of the normalization strategy. Subsequently, a regularized log transformation was applied to the normalized counts, followed by Z-score standardization. To enhance comparability for visualization purposes, a min–max normalization was applied, scaling the Z-scores between 0 and 1. After normalization, gene expression profiles were clustered using both hierarchical and K-means clustering approaches to identify global transcriptional patterns.

The optimal number of clusters was determined using independent quantitative criteria. Specifically, the Elbow method based on the Within-Cluster Sum of Squares (WCSS) identified a clear inflection point at k = 2 (Supplementary Figure 2A). Consistently, Silhouette method analysis showed the highest average silhouette width at k = 2 (Supplementary Figure 2B), supporting an optimal separation of genes into two groups.

For hierarchical clustering and data visualization, the pheatmap v1.0.12 package was used [29]. K-means clustering was performed using the kmeans function from the base stats package in R, with the number of clusters set to 2 (k = 2).

Viral gene expression counts for every cell line at 48 hpi were clustered using the same approach and the same number of clusters (Supplementary Figure 3).

Functional classification

We carried out a functional characterization of the MPXV genes, considering that the Poxviridae family belongs to the group of nucleocytoplasmic large DNA viruses (NCLDV), viral families that share a common evolutionary origin. The study of Yutin et al. [30] classified the genes of NCLDV into clusters

of orthologous groups (NCVOG), many of which could be linked to potential functional categories [30]. The functional annotation of the MPXV genes was first obtained using the InterProScan Tool [31,32]. We then classified them into functional groups based on the NCVOG categories [30]. Twenty genes (11.17%) were uncharacterized, while the remaining genes were divided into seven categories (Figure 1B), resulting in a total of eight categories. The most abundant is the "Host-virus interaction," comprising 57 genes (31.84%), followed by "Virion structure and morphogenesis," which included 48 genes (26.82%), and "Transcription and RNA processing," with 29 genes (16.20%). Other categories included "DNA replication, recombination and repair" (12 genes, 6.70%), "Nucleotide metabolism" (5 genes, 2.79%), "Other metabolic functions" (2 genes, 1.12%), and "Miscellaneous" (6 genes, 3.37%).

Temporal regulation of MPXV gene expression identified by likelihood ratio testing (LRT)

To statistically validate the temporal dynamics of MPXV gene expression, we applied a likelihood ratio test (LRT) using the DESeq2 [28] framework to viral gene expression counts across the three time points (2-, 24-, and 48-h post-infection). The test compared a full model (~ replicate + time) with a reduced model (~ replicate), where the factor replicates account for replicate variability, and time represents the progression of infection. This approach identifies genes whose expression is significantly regulated over the course of infection. To further characterize the expression dynamics, we used the degPatterns function, part of the DEGreport v1.34.0 R package [33] – a pattern-based clustering method that groups genes based on the shape of their expression profiles across time points.

Statistical analysis

For Figure 3A and B, data are presented as mean ± standard deviation (SD) of three independent biological replicates per condition (n = 3). To compare the TCID50/ml between cell lines, the Kruskal–Wallis with post-Dunn test was performed.

Results

Temporal MPXV gene expression in infected epithelial cells

To profile the temporal expression of MPXV genes, we performed the whole transcriptomic profiling of MPXV-infected VK2/E6E7 vaginal epithelial cells over the course of infection from an early time point (2 h post-infection; hpi) to 24 and 48 hpi when the cytopathic effect became evident. Figure 1A shows the percentages of mapped reads on the human genome and on the MPXV genome as a function of the infection time. The statistics show that the three replicates have good reproducibility, with similar values in both genomes. As expected, the percentages of reads aligned on the MPXV genome increased over time, and the values for the MPXV mapped reads ranged from an average of 0.04% at 2 hpi, to 9.13% at 24 hpi, and 29.02% at 48 hpi (Figure 1A).

To identify temporal patterns in MPXV gene expression in infected VK2/E6E7 cells, a hierarchical cluster analysis was carried out on DESeq2-normalized counts for every time point. The expression of individual MPXV genes also shows excellent reproducibility in the three replicates for each time point (Supplementary Table 1). The hierarchical cluster analysis, carried out with two independent methods, highlighted two major temporal groups (Supplementary Figure 4): (i) genes highly expressed at the beginning of the viral infection, as early as 2 hpi; (ii) genes highly expressed between 24 and 48 hpi. To further validate the gene expression profiles of these two temporal groups and assign the MPXV genes more accurately, we performed a second clustering analysis with the *k*-means clustering method, imposing k = 2. This approach allowed us to quantitatively partition the genes into two distinct expression patterns over time as suggested by the hierarchical clustering. According to the NC_063383.1 annotation file, the MPXV genome contains 179 annotated protein-coding genes, all of which were detected and included in the downstream expression analyses. Each of these genes was uniquely assigned to one of the two clusters identified by the *k*-means analysis. Specifically, 63 genes were assigned to cluster 1 (35.2%, shown in blue in Figure 1B), and 116 genes to cluster 2 (64.8%, shown in green in Figure 1B) (see Supplementary Table 2).

Subsequently, we characterized the functions of all the genes for each cluster using the MPXV gene annotations (see the Experimental Procedure Section and Figure 1C) and highlighted the main viral genomic functions for each cluster (Figure 1D). In cluster 1, MPXV gene functions are mostly associated with host-virus interaction mechanisms, such as viral entry into the host cell, transport to viral replication sites, and immune evasion mechanisms. The other genes are associated with DNA replication and transcription. Genes highly expressed from 24 to 48 hpi (cluster 2) are mainly associated with intermediate and late viral life cycle processes, such as DNA replication, transcription, and virion structure and morphogenesis. Genes in cluster 1 are mainly located on the terminal regions of the genome (Figure 1E), while genes in cluster 2 are in the central region of the genome (Figure 1F).

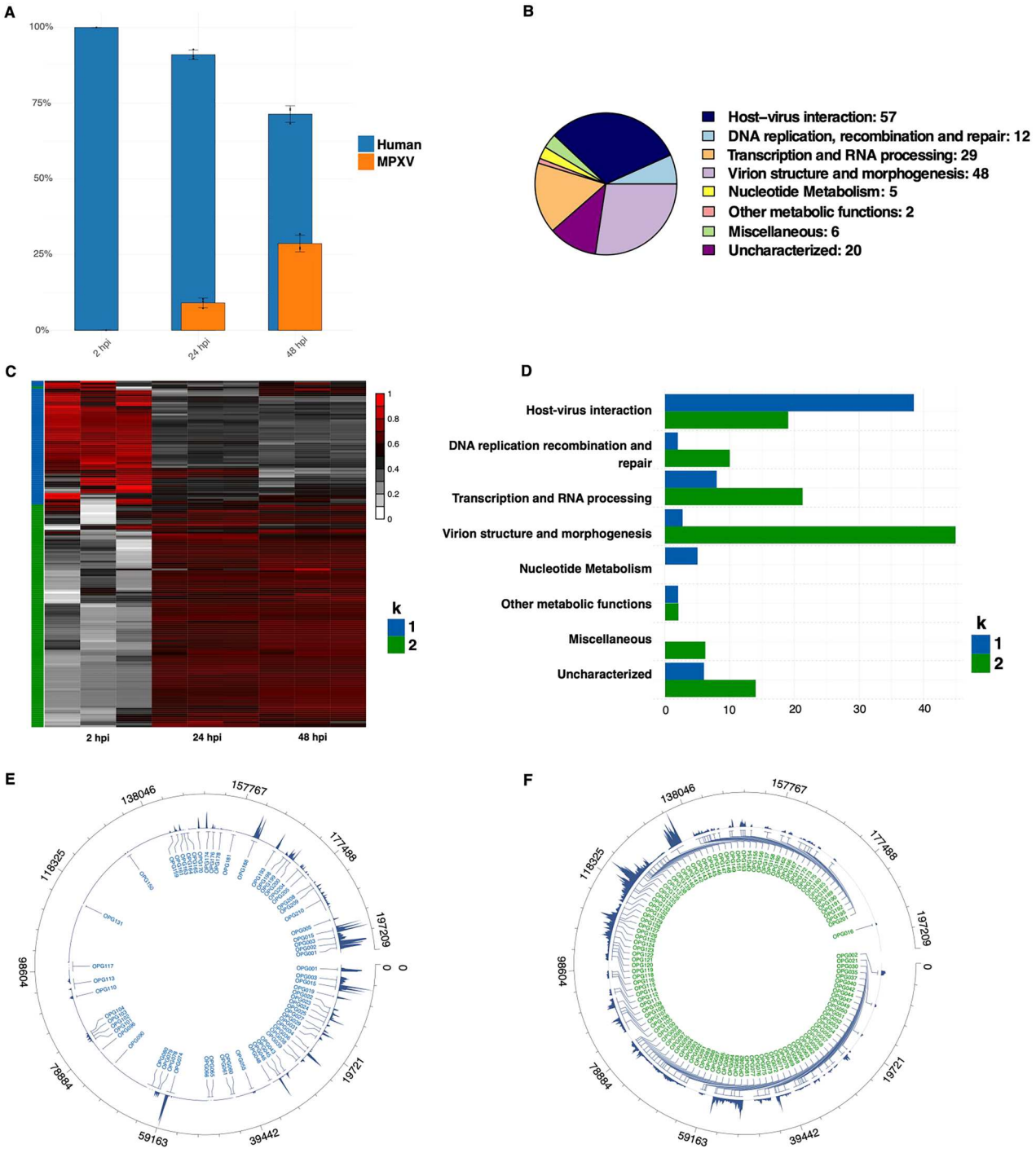

Figure 1. Characterization of MPXV transcriptome during the infection of the VK2/E6E7 cell line at different infection time points. (A) Percentages of aligned reads to Human reference and MPXV reference across different hpi. (B) Functional annotation of MPXV viral genes (see the Experimental Procedure Section). (C) Heat map of MPXV gene expression profiles at different hpi divided into two clusters (k = 2) by k-means clustering method. (D) Functional annotation of viral clusters. (E) Coverage of viral genes associated with cluster 1 calculated on peak expression times (2 hpi). (F) Coverage of viral genes associated with cluster 2 calculated on peak expression times (24–48 hpi).

Although all MPXV genes were uniquely assigned to one of the two temporal clusters identified by the *k*-means analysis, we further investigated the top 20 most highly expressed genes across all time points and identified the most transcriptionally active components of the viral genome (Table 1). The analysis confirmed differential expression dynamics across the infection timeline: genes assigned to cluster 1 showed the highest read counts at 2 hpi, while those in cluster 2 exhibited increased expression at later time points. The two most highly expressed genes, OPG136 and OPG071, belong to cluster 2 and encode the viral core protein P4a and the DNA polymerase, respectively. The most represented functional category among the top 20 was "Virion structure and morphogenesis" which included six genes (OPG070, OPG069, OPG153, OPG129, OPG136, and OPG138), followed by "Host-virus interaction," represented by six genes

Table 1. Top 20 highly expressed genes in VK2/E6E7 cell line over the entire course of infection. For each gene, the following information is reported: RefSeq annotation (NC_063383.1); the value of k obtained using the k-means clustering method (column "K", Figure 1B); protein name; functional classification ("Functional category" column. Figure 1C); number of normalized reads values across different time points in thousands, computed using the "median of ratios" of the DESeq2 package.

MPXV genes	K	Protein	Functional category	2 hpi	24 hpi	48 hpi
OPG055	1	Protein F11	Host-virus interaction	14.32	1.28	0.87
OPG188	1	Schlafen (1)	Host-virus interaction	12.21	1.41	1.44
OPG110	1	Late transcription factor VLTF-4 (1)	Transcription and RNA processing	11.34	3.97	2.95
OPG074	1	lev morphogenesis protein	Host-virus interaction	7.59	7.80	5.70
OPG080	1	Ribonucleoside-diphosphate reductase (2)	Nucleotide Metabolism	5.73	2.12	1.94
OPG039	1	Ankyrin-like protein (3)	Host-virus interaction	5.62	3.07	2.20
OPG210	1	B22R family protein	Host-virus interaction	5.58	2.81	2.79
OPG174	1	Hydroxysteroid dehydrogenase	Host-virus interaction	5.01	2.14	3.49
OPG113	1	mRNA-capping enzyme large subunit	Transcription and RNA processing	4.07	3.95	2.60
OPG071	2	DNA polymerase (2)	DNA replication, recombination, and repair	3.73	10.76	7.51
OPG105	2	DNA-dependent RNA polymerase subunit rpo147	Transcription and RNA processing	3.74	3.93	4.19
OPG151	2	DNA-dependent RNA polymerase subunit rpo132	Transcription and RNA processing	1.72	3.96	5.27
OPG129	2	Virion core protein P4b	Virion structure and morphogenesis	0.59	5.99	8.86
OPG077	2	Telomere-binding protein I1	DNA replication, recombination and repair	0.19	6.39	4.39
OPG136	2	Virion core protein P4a	Virion structure and morphogenesis	0.17	9.32	15.18
OPG125	2	Rifampicin resistance protein	Transcription and RNA processing	0.14	5.11	5.59
OPG138	2	A12 protein	Virion structure and morphogenesis	0.08	4.81	8.21
OPG153	2	Orthopoxvirus A26L/A30L protein	Virion structure and morphogenesis	0.05	3.62	6.91
OPG070	2	Membrane protein E8	Virion structure and morphogenesis	0.01	8.81	6.07
OPG069	2	Myristoylated protein E7	Virion structure and morphogenesis	0.00	6.36	4.67

(OPG039, OPG188, OPG074, OPG055, OPG174, and OPG210). Genes in cluster 1 included several transcripts encoding proteins involved in the modulation of IFN response (OPG024, 025, 065, 188, 193, 204), NF-kB, and antiapoptotic (OPG038, 039, 176, 200, 005, 199) pathways. Six genes were abundantly expressed at all time points: four genes were implicated in DNA transcription and replication (OPG071, OPG105, OPG110, and OPG113) and included genes encoding the DNA and the RNA polymerase; one gene (OPG074) is supposed to be involved in the modulation of the ERK1/2 signalling pathway, and another one is involved in the modulation of host inflammation (OPG174).

We also analyzed gene expression on a time-point-specific basis by identifying the top 20 most highly expressed genes at each time of infection (Supplementary Table 3). This time-resolved analysis confirmed that four genes – OPG071, OPG110, OPG105, and OPG074 – were consistently among the top 20 most expressed genes at every time point. Interestingly, hierarchical clustering assigned these genes to different temporal clusters, with two falling into cluster 1 and two into cluster 2. This highlights how genes with consistently high expression levels can still differ in their temporal expression profiles.

To statistically validate the temporal dynamics of MPXV gene expression, we applied a likelihood ratio test (LRT) to viral gene expression counts across the three time points. The analysis identified 151 MPXV genes as significantly regulated over time. To characterize the expression dynamics, we used a pattern-based clustering method that groups genes based on the shape of their expression profiles across time points. Four distinct expression trajectories, or groups, were identified (Figure 2). Group 1 identified 46 genes with high expression at the onset of infection (2 hpi), followed by progressive downregulation; Group 2 identified 102 genes with increasing expression over time and peaking at 24–48 hpi. Groups 3 ($n = 2$ genes) and 4 ($n = 1$ genes) displayed unique expression profiles distinct from the two main clusters. Group 1 was dominated by host-virus interaction factors, and no structural genes appear in this group. In contrast, Group 2 was strongly enriched in virion structure and morphogenesis and transcription-related genes, reflecting the shift towards viral replication and assembly (Supplementary Table 4). The presence of uncharacterized and miscellaneous genes in this cluster may point to late-stage processes not yet fully understood. Group 3 included only OPG031 and OPG074, both modulating cell signal transduction pathways (NFkB and ERK, respectively). The single gene in Group 4, OPG159, is a putative ATPase. This statistical modelling complemented unsupervised clustering methods, offering robust validation of the transcriptional progression during infection. Finally, 28 of the 179 detected MPXV genes did not show significant temporal regulation and included genes in the functional categories: host-virus interaction, transcription and RNA processing, DNA replication, recombination and repair, and virion structure and morphogenesis. Their stable expression across time points suggests functions required throughout the infection cycle (Supplementary Table 5).

Characterization of the MPXV transcriptional profile across different epithelial cell lines

To understand how the expression of the viral genome varies in epithelial cells derived from different anatomical districts that may differ in their susceptibility

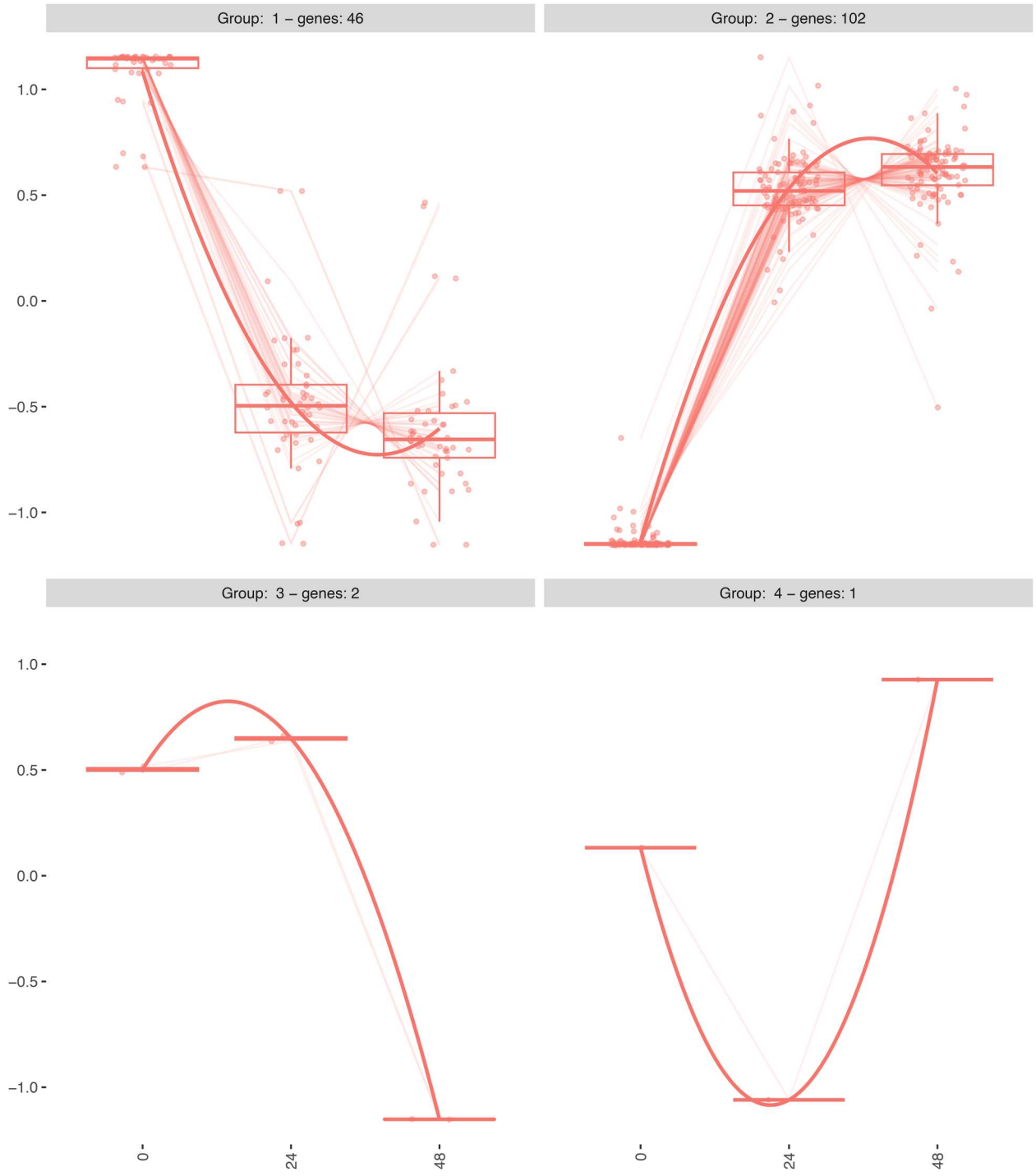

Figure 2. Temporal regulation of MPXV gene expression identified by likelihood ratio testing (LRT). A total of 151 MPXV genes showing significant temporal regulation (likelihood ratio test, LRT) were clustered based on their expression patterns across three time points (2-, 24-, and 48-hpi). Each panel represents one of four distinct expression groups, with individual gene trajectories (thin lines and dots) and the group mean profile (bold line).

to infection, we compared the viral transcriptional profiles in MPXV-infected epithelial cell lines from the vagina (VK2/E6E7), ectocervix (Ect1/E6E7), colon (Caco-2), and kidney (HK-2). The proportion of viral reads relative to human reads differed markedly among the cell lines (Figure 3). VK2/E6E7 and Caco-2 cells exhibited a higher percentage of viral reads, whereas Ect1/E6E7 and HK-2 cells showed considerably lower values (Figure 3A). Notably, HK-2 cells displayed the lowest percentage of viral reads across all cell lines, accounting for only 3.55%. Individual biological replicates are shown, highlighting the relatively low variability observed within each cell line.

The different cell lines also demonstrated differential levels of productive infection, measured by titration of the virus in the supernatant of infected cells (Figure 3B), reflecting the abundance of viral reads. Indeed, Caco-2 and VK-2/E6E7 showed the highest release of infectious particles in the supernatant,

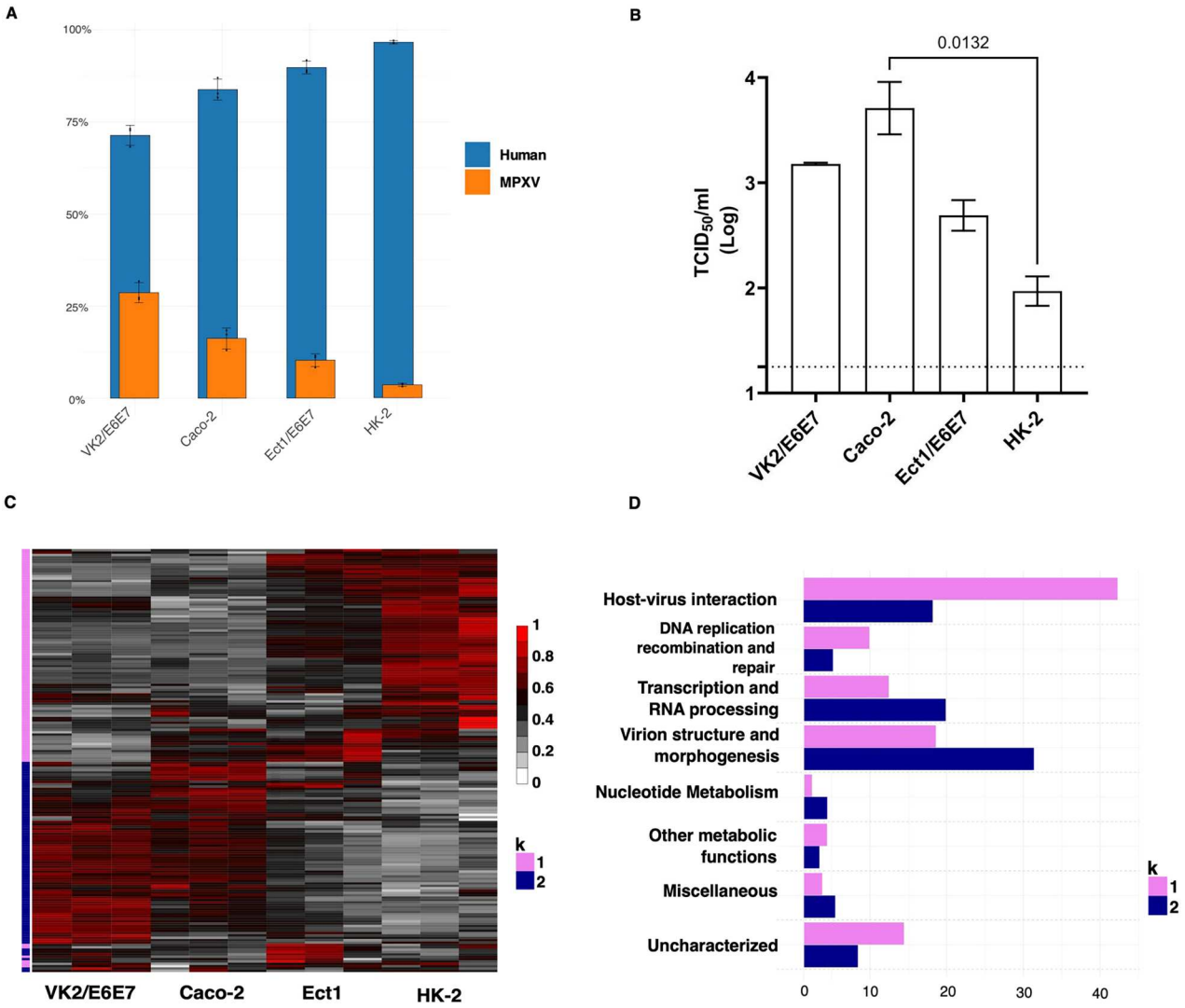

Figure 3. Characterization of MPXV transcriptome during infection of epithelial cell lines at 48 hpi. (A) Percentages of reads aligned to the Human reference and MPXV reference (B) Infectious MPXV titres in infected cell supernatants expressed as Log TCID50/ml (n = 3), dotted line represents the limit of quantification of the assay (C) Heat map of MPXV gene expression profiles divided into two clusters (k = 2) by k-means clustering method. (D) Functional annotation of viral clusters.

followed by Ect1/E6E7 and HK-2 (Figure 3B). To investigate whether differences in viral read abundance influence the overall viral transcriptome, we clustered and annotated MPXV transcripts across all four cell lines. Expression patterns were highly reproducible among the three biological replicates for each cell line (Supplementary Table 6). This analysis revealed two distinct gene expression profiles: (i) 83 (46.9%) highly expressed genes in Ect1/E6E7 and HK-2 cells; (ii) 96 (53.7%) highly expressed genes in VK2/E6E7 and Caco-2 cells (Figure 3B, Supplementary Figure 5, Supplementary Table 7). Functional characterization of the genes associated with the two profiles shows that the first one is characterized by genes of all Functional categories, with a high abundance in "Host-virus interaction"; cluster 2 is strongly associated with "Transcription and RNA processing" and "Virion structure and morphogenesis" (Figure 3C and D).

We selected the 20 most highly expressed MPXV genes in each cell line (Supplementary Table 7) and identified a group of eleven genes that were consistently expressed across all four cell types (Table 2). Notably, nearly half of these shared genes are associated with the functional category "Virion structure and morphogenesis," reflecting the ability of all the cell lines to produce transcripts for structural components of the virus. In addition, seven of these eleven genes are assigned to cluster 2, which includes genes with higher expression in VK2/E6E7 and Caco-2 cells.

In contrast, five genes were in the top 20 list only in the renal cell line, two of them implicated in the inhibition of antiviral response (OPG065, OPG188), one modulating the NFkB pathway (OPG200), the OPG040 coding for the Serpin, a superinfection exclusion protein, and OPG161 encoding the EEV Glycoprotein. The OPG039, producing a factor involved in host range, NFkB inhibition, and antagonism of

Table 2. Genes expressed in all the cell lines at 48 hpi. For each gene, the following information is reported: RefSeq annotation (NC_063383.1); the value of k obtained using the k-means clustering method (column "K", Figure 1B); protein name; functional classification ("Functional category" column. Figure 1C); number of normalized reads values across different time points in thousands, computed using the "median of ratios" of the DESeq2 package.

MPXV genes	K	Protein	Functional category	VK2/E6E7	Caco-2	Ect1/E6E7	HK-2
OPG069	1	Myristoylated protein E7	Virion structure and morphogenesis	26.16	23.61	28.55	45.80
OPG070	1	Membrane protein E8	Virion structure and morphogenesis	34.06	33.76	35.03	53.05
OPG071	1	DNA polymerase(2)	DNA replication, recombination and repair	42.12	60.46	55.61	78.62
OPG074	2	lev morphogenesis protein	Host-virus interaction	31.79	34.17	32.86	36.73
OPG077	1	Telomere-binding protein I1	DNA replication, recombination and repair	24.62	20.77	26.27	44.48
OPG105	2	DNA-dependent RNA polymerase subunit rpo147	Transcription and RNA processing	23.53	27.40	21.14	20.19
OPG125	2	Rifampicin resistance protein	Transcription and RNA processing	31.32	29.93	28.38	27.19
OPG129	2	Virion core protein P4b	Virion structure and morphogenesis	49.75	44.77	32.31	28.66
OPG136	2	Virion core protein P4a	Virion structure and morphogenesis	85.12	99.83	59.36	44.78
OPG151	2	DNA-dependent RNA polymerase subunit rpo132	Transcription and RNA processing	29.53	31.01	28.24	22.28
OPG138	2	A12 protein	Virion structure and morphogenesis	45.99	41.53	27.26	22.38

antiviral response, was highly expressed by both the HK-2 and Ect-1/E6E7 cells. Of note, HK-2 at 48 h maintained high expression of transcripts observed in VK2/E6E7 at an early time point (2 hpi). These transcripts include immune evasion-encoding products such as OPG038 (an early protein involved in T cell response evasion), OPG039 (antiviral response antagonist), OPG065 (inhibiting the PKR pathway), and OPG188 (evading cGAS-mediated innate immunity) [4]. In addition to viral gene expression profiling, a global modulation of the host transcriptome was observed, as indicated by principal component analysis of the human gene expression data (Supplementary Figure 6).

Discussion

The unprecedented global spread of MPXV, which started in 2022 due to clade IIb, shed light on a previously neglected pathogen and fuelled the interest of scientists towards the understanding of the biology and evolution of this virus.

Here, we focused our attention on the temporal and differential expression of MPXV clade IIb genes in cells derived from a range of anatomical districts to analyze in depth the hierarchical regulation of viral genes and identify MPXV gene expression in compartments that can be differentially susceptible to viral infection.

Temporal profiling was performed in VK2/E6E7 vaginal epithelial cells at three time points post-infection (2, 24 and 48 h), enabling the reconstruction of viral gene expression kinetics throughout the course of infection. We also profiled viral gene expressions in three other epithelial cell lines at a single time point post-infection. This design enabled a direct comparison of viral transcriptional profiles across all four cell types, highlighting both temporal regulation and tissue-specific differences in transcript abundance. Through clustering, statistical modelling, and functional annotation of viral transcripts, we aimed to characterize the temporal dynamics and tissue-specific features of the MPXV transcriptional landscape *in vitro*.

Our analyses evidenced two main clusters: early viral gene expression (2 hpi) and intermediate–late (24–48 hpi). The early cluster included genes involved in the host–virus interaction functional category and mapping to the terminus region; the intermediate–late cluster was enriched in transcripts of genes involved in viral transcription, replication, and morphogenesis and located in the core genome. We also identified distinct expression trajectories with the vast majority of viral genes showing either an early or an intermediate–late expression, consistent with the identification of two main clusters and being enriched in virus–host interaction or transcription-replication-morphogenesis encoding factors, respectively. Nevertheless, for a minor subset of genes, three other possible dynamics were observed: an early-intermediate, an early-late, and a continuous expression. Genes implicated in the evasion of the IFN-I mediated response (e.g. OPG024, OPG025, OPG026, OPG027, OPG029) belonged to the early cluster and first trajectory (highly expressed at 2 hpi and less expressed at 24–48 hpi). This is in line with the significant delay in the expression of IFN-I observed in cells infected either with MPXV clade I [14] and IIb [34].

The viral genes continuously expressed at each time point tested comprised some transcripts encoding signal transduction and inflammation-modulating factors, as well as the viral polymerases and their cofactors. OPG001, 002, and 003 have been implicated in chemokine binding, TNF-alpha pathway modulation, and antibody-dependent complement enhanced neutralization of infectivity, respectively. NF-kB inhibitions have been described for OPG047 and OPG058 [4], which are also expressed at each time point in our model. TNF-alpha is an inflammatory cytokine produced in response to viral infections following the activation of NF-kB. The importance of

modulating the TNF pathway is evidenced by the adoption during the evolution of poxviruses of different molecular mechanisms to contrast its action, including soluble TNF receptors such as the crmB encoded by OPG002 [35].

The hierarchical expression of Orthopoxviruses' genes was previously described, together with the non-random expression of genes in the core and terminal regions, but our contribution is of importance since most published data comes from VACV or less recent MPXV strain studies [4], and for the application of bioinformatic tools, which allowed us to better dissect MPXV expression patterns.

A proteomic analysis comparing the MPXV clade II from the 2003 US outbreak and the IIb from the 2022 global epidemics [6] underlined the higher production by the recent strain of the proteins encoded by OPG039 and OPG074, which are implicated in antiviral response modulation and extracellular signal-regulated kinase 1/2 (ERK1/2) activation, respectively. In our analysis, OPG039 is an early factor highly expressed at 2 hpi, while OPG074 showed a unique pattern of expression, being highly expressed at the early time point and at 24 hpi. The multiomics analysis by Huang Y and colleagues [20] on MPXV clade IIb-infected human foreskin fibroblasts (MOI = 3) underlined the importance of phosphorylation in the regulation of the viral life cycle, suggesting an interplay between viral and host kinases and phosphatases during the course of infection, with OPG074 showing phosphomotifs for over 20 MAPKs. The biological function of OPG074 in enhancing ERK1/2 activation and promoting viral replication is described in Vaccinia virus infection, but the sequence conservation and phosphorylation pattern suggest a conserved activity among these Orthopoxviruses [20]. Targeted approaches could help in the comprehension of the role of human and MPXV encoded kinases in the regulation of viral infectivity or pathogenesis.

Our MPXV clade IIb transcriptome comparative study, carried out at a later time point from infection (i.e. 48 hpi), showed a higher production of infectious viral particles from vaginal and intestinal cells than from ectocervical and renal epithelial cell lines, which was accompanied by a clustering of genes highly expressed in the most productive and in the less productive cells. The vaginal and intestinal cells at this time mainly expressed genes involved in the intermediate–late phase of replication; the ectocervical and renal cells were enriched in transcripts associated with the host-virus interaction functional category. The differential transcriptome investigation across cell lines also revealed genes that were expressed in all the analyzed cell lines. The "common genes" included transcripts for both accessory (e.g. OPG069, OPG074) and core genes. OPG074 is among the genes highly expressed in all the cell lines, again indicating the importance of ERK 1/2 activation for the viral life cycle.

The renal cells, which produced the lower number of infectious particles, also showed high expression of numerous genes involved in antiapoptotic mechanisms, signal transduction, and immune evasion pathways. Whether the differences in viral release and viral-gene expression observed here reflect the in vivo replication capacity of the MPXV clade IIb remains to be established. Studies analyzing viral presence in bodily fluids from 2022 MPXV-infected patients [21,22,36,37] described high viral load in skin lesions from the anogenital area, and a high percentage of viral DNA detection in rectal and faecal samples. Lower viral load and DNA detection rate were observed in urine, with the hypothesis of the virus originating from the urethra rather than from the upper urinary tract [21,22,37].

Compared to previous epidemics, the 2022 outbreak only partially involved female patients, but sexual transmission represented the primary route of transmission, and genital lesions were observed in both male and female patients [7,38,39]. Moreover, animal studies recognized vaginal and rectal exposure as a preferential route for systemic dissemination of MPXV IIb [23].

One limitation of our study is the lack of temporal analysis in all the cell lines tested; nevertheless, our findings suggest viral dynamics modulation according to the cellular target.

In line with this, the principal component analysis of the human gene expression data shows that variance is primarily driven by the cellular background, with clear separation among the four lines, and the vaginal and ectocervical lines closer than the others. The global gene expression profiles, therefore, mask the effect of the infection that – however – is still visible. Nonetheless, a detailed analysis of host response to infection is beyond the scope of this manuscript. These findings further underline the importance of studying viral dynamics across different cellular compartments.

Further studies are needed to better understand the role of the MPXV products in the human host, ideally through the coupling of in vivo and in vitro analyses with the use of advanced biological models and bioinformatic tools. A fine comparison of clades Ib and IIb infections will also be essential for the understanding of common and diverse host-virus interactions to be targeted for therapeutic purposes. Moreover, the use of the unified nomenclature for Orthopoxvirus genes should be preferred to avoid the misinterpretation of published data. The evolution and emergence of MPXV clades need to be monitored and investigated through omics and targeted approaches to fill the gaps in our knowledge of the biology of this pathogen.

Acknowledgments

We acknowledge the CINECA award under the ISCRA initiative, for the availability of high performance computing resources and support. We acknowledge ELIXIR-IIB (elixir-italy.org), the Italian Node of the European ELIXIR infrastructure (elixireurope.org).

Authors contributions

conceptualization, G.C., F.M., G.M.; bioinformatic analyses, L.P., D.P., M.M., C.E.M.G.; investigation, D.M., C.Mi., S.M., G.M., L.B., V.M.; writing – original draft, G.M., L.P., G.C.; writing – review & editing, G.M., L.P., G.C., F.M., C.M., A.A.; funding acquisition, F.M.; supervision, G.C., G.M., F.M.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Ministero della Salute.

Data and code availability

RNA-seq data have been submitted to the NCBI Sequence Read Archive (SRA) under BioProject accession PRJNA1395380 and will be made publicly available upon publication. All scripts used for data processing and analysis are publicly available at https://github.com/LPicarone/MPXV_RNAseq_analysis.

References

[1] Bratke KA, McLysaght A, Rothenburg S. A survey of host range genes in poxvirus genomes. Infect Genet Evol. 2013;14:406–425. doi:10.1016/j.meegid.2012.12.002

[2] Lansiaux E, Jain N, Laivacuma S, et al. The virology of human monkeypox virus (hMPXV): a brief overview. Virus Res. 2022;322:198932. doi:10.1016/j.virusres.2022.198932

[3] Senkevich TG, Yutin N, Wolf YI, et al. Ancient gene capture and recent gene loss shape the evolution of Orthopoxvirus-host interaction genes. mBio. 2021;12:e0149521. doi:10.1128/mbio.01495-21

[4] Deng Y, Navarro-Forero S, Yang Z. Temporal expression classes and functions of vaccinia virus and mpox (monkeypox) virus genes. mBio. 2025;16:e0380924. doi:10.1128/mbio.03809-24

[5] Mathieu E, Spooner F, Dattani S, et al. Mpox. *Our World in Data*; 2022.

[6] McGrail JP, Mondolfi AP, Ramírez JD, et al. Comparative analysis of 2022 outbreak MPXV and previous clade II MPXV. J Med Virol; 2024;96:e70023. doi:10.1002/jmv.70023

[7] Vakaniaki EH, Kacita C, Kinganda-Lusamaki E, et al. Sustained human outbreak of a new MPXV clade I lineage in eastern Democratic Republic of the Congo. Nat Med. 2024;30:2791–2795. doi:10.1038/s41591-024-03130-3

[8] Giorgi FM, Pozzobon D, Di Meglio A, et al. Genomic and transcriptomic analysis of the recent Mpox outbreak. Vaccine. 2024;42:1841–1849. doi:10.1016/j.vaccine.2023.12.086

[9] Monzón S, Varona S, Negredo A, et al. Monkeypox virus genomic accordion strategies. Nat Commun. 2024;15:3059. doi:10.1038/s41467-024-46949-7

[10] Mavian C, López-Bueno A, Martín R, et al. Comparative pathogenesis, genomics and phylogeography of mousepox. Viruses. 2021;13:1146. doi:10.3390/v13061146

[11] Ferrareze PAG, Pereira e Costa RA, Thompson CE. Genomic characterization and molecular evolution of human monkeypox viruses. Arch Virol. 2023;168:278. doi:10.1007/s00705-023-05904-5

[12] Molteni C, Forni D, Cagliani R, et al. Evolution of the orthopoxvirus core genome. Virus Res. 2023;323:198975. doi:10.1016/j.virusres.2022.198975

[13] Rubins KH, Hensley LE, Bell GW, et al. Comparative analysis of viral gene expression programs during poxvirus infection: a transcriptional map of the vaccinia and monkeypox genomes. PLoS One. 2008;3:e2628. doi:10.1371/journal.pone.0002628

[14] Rubins KH, Hensley LE, Relman DA, et al. Stunned silence: gene expression programs in human cells infected with monkeypox or vaccinia virus. PLoS One. 2011;6:e15615. doi:10.1371/journal.pone.0015615

[15] Shchelkunov SN. Orthopoxvirus genes that mediate disease virulence and host tropism. Adv Virol. 2012;2012:1–17. doi:10.1155/2012/524743

[16] Hudson PN, Self J, Weiss S, et al. Elucidating the role of the complement control protein in monkeypox pathogenicity. PLoS One. 2012;7:e35086. doi:10.1371/journal.pone.0035086

[17] Thornhill JP, Gandhi M, Orkin C. Mpox: the reemergence of an old disease and inequities. Annu Rev Med. 2024;75:159–175. doi:10.1146/annurev-med-080122-030714

[18] Young B, Seifert SN, Lawson C, et al. Exploring the genomic basis of Mpox virus-host transmission and pathogenesis. mSphere. 2024;9:e0057624. doi:10.1128/msphere.00576-24

[19] Kakuk B, Dörmő Á, Csabai Z, et al. In-depth temporal transcriptome profiling of monkeypox and host cells using nanopore sequencing. Sci Data. 2023;10:262. doi:10.1038/s41597-023-02149-4

[20] Huang Y, Bergant V, Grass V, et al. Multi-omics characterization of the monkeypox virus infection. Nat Commun. 2024;15:6778. doi:10.1038/s41467-024-51074-6

[21] Meschi S, Colavita F, Carletti F, et al. MPXV DNA kinetics in bloodstream and other body fluids samples. Sci Rep. 2024;14:13487. doi:10.1038/s41598-024-63044-5

[22] Kim H, Kwon R, Lee H, et al. Viral load dynamics and shedding kinetics of mpox infection: a systematic review and meta-analysis. J Travel Med. 2023;30:taad111. doi:10.1093/jtm/taad111

[23] Port JR, Riopelle JC, Smith SG, et al. Infection with mpox virus via the genital mucosae increases shedding and transmission in the multimammate rat (Mastomys natalensis). Nat Microbiol. 2024;9:1231–1243. doi:10.1038/s41564-024-01666-1

[24] Andrews S, Krueger F, Segonds-Pichon A, et al. FastQC; 2012.

[25] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–2120. doi:10.1093/bioinformatics/btu170

[26] Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29:15–21. doi:10.1093/bioinformatics/bts635
[27] Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30:923–930. doi:10.1093/bioinformatics/btt656
[28] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550. doi:10.1186/s13059-014-0550-8
[29] Kolde R. pheatmap: Pretty Heatmaps; 2025. https://github.com/raivokolde/pheatmap.
[30] Yutin N, Wolf YI, Raoult D, et al. Eukaryotic large nucleo-cytoplasmic DNA viruses: clusters of orthologous genes and reconstruction of viral genome evolution. Virol J. 2009;6:223. doi:10.1186/1743-422X-6-223
[31] de Souza FG, Abrahão JS, Rodrigues RAL. Comparative analysis of transcriptional regulation patterns: understanding the gene expression profile in Nucleocytoviricota. Pathogens. 2021;10(8):935. doi:10.3390/pathogens10080935
[32] Jones P, Binns D, Chang H-Y, et al. InterProScan 5: genome-scale protein function classification. Bioinformatics. 2014;30:1236–1240. doi:10.1093/bioinformatics/btu031
[33] Pantano L, et al. DEGreport: Report of DEG analysis. New Jersey: R package version; 2019; 1: 10–18129.
[34] Bordi L, D'Auria A, Frasca F, et al. MPXV infection impairs IFN response but is partially sensitive to IFN-γ antiviral effect. Med Microbiol Immunol. 2024;213:25. doi:10.1007/s00430-024-00808-w
[35] Alvarez-de Miranda FJ, Alonso-Sánchez I, Alcamí A, Hernaez B. TNF decoy receptors encoded by poxviruses. Pathogens 2021;10:1065. doi:10.3390/pathogens10081065
[36] Suñer C, Ubals M, Tarín-Vicente EJ, et al. Viral dynamics in patients with monkeypox infection: a prospective cohort study in Spain. Lancet Infect Dis. 2023;23:445–453. doi:10.1016/S1473-3099(22)00794-0
[37] Piralla A, Mileto D, Rizzo A, et al. Dynamics of viral DNA shedding and culture viral DNA positivity in different clinical samples collected during the 2022 mpox outbreak in Lombardy, Italy. Travel Med Infect Dis. 2024;59:102698. doi:10.1016/j.tmaid.2024.102698
[38] Bragazzi NL, Woldegerima WA, Wu J, et al. Epidemiological and clinical characteristics of Mpox in cisgender and transgender women and non-binary individuals assigned to female sex at birth: a comprehensive, critical global perspective. Viruses. 2024;16:325. doi:10.20944/preprints202401.2099.v1
[39] Portela-Dias J, Sereno S, Falcão-Reis I, et al. Monkeypox infection with localized genital lesions in women. Am J Obstet Gynecol. 2022;227:906. doi:10.1016/j.ajog.2022.08.046