A giant virus genome is densely packaged by stable nucleosomes within virions

Graphical abstract

Highlights
- Marseillevirus nucleosomes tightly wrap 121 bp of DNA within virions
- Nucleosomes lack linkers and resist nuclease and chemical cleavage
- Marseillevirus genes lack nucleosome positioning or 5' phasing
- Reconstitution on closely spaced multi-copy arrays confirms tight packing

Authors
Terri D. Bryson, Pablo De Ioannes, Marco Igor Valencia-Sánchez, ..., Bernard La Scola, Karim-Jean Armache, Steven Henikoff

Correspondence
karim-jean.armache@nyulangone.org (K.-J.A.), steveh@fredhutch.org (S.H.)

In brief
The authors show that the doublet histones of Marseillevirus tightly wrap 121 base pairs of DNA within virions and lack linkers. This distinct form of higher-order chromatin suggests adaptation for protection within virions. Doublet histones may resemble an early stage of histone differentiation leading to the eukaryotic octameric nucleosome.
A giant virus genome is densely packaged by stable nucleosomes within virions

Terri D. Bryson, Pablo De Ioannes, Marco Igor Valencia-Sánchez, Jorja G. Henikoff, Paul B. Talbert, Rachel Lee, Bernard La Scola, Karim-Jean Armache, and Steven Henikoff

SUMMARY

The two doublet histones of Marseillevirus are distantly related to the four eukaryotic core histones and wrap 121 base pairs of DNA to form remarkably similar nucleosomes. By permeabilizing Marseillevirus virions and performing genome-wide nuclease digestion, chemical cleavage, and mass spectrometry assays, we find that the higher-order organization of Marseillevirus chromatin fundamentally differs from that of eukaryotes. Marseillevirus nucleosomes fully protect DNA within virions as closely abutted 121-bp DNA-wrapped cores without linker DNA or phasing along genes. Likewise, we observed that nucleosomes reconstituted onto multi-copy tandem repeats of a nucleosome-positioning sequence are tightly packed. Dense promiscuous packing of fully wrapped nucleosomes rather than “beads on a string” with genic punctuation represents a distinct mode of DNA packaging by histones. We suggest that doublet histones have evolved for viral genome protection and may resemble an early stage of histone differentiation leading to the eukaryotic octameric nucleosome.

INTRODUCTION

The association of histones with DNA in the eukaryotic nucleus was known by the late 19th century, but it was the revolutionary discovery of nucleosomes in the early 1970s that established the fundamental subunit structure of eukaryotic chromatin. Other nucleosome configurations were described for homotetrameric archaeal histones, which were found to wrap 60 base pairs of DNA and form higher order “slinkies” (Mattioli et al., 2017). Some archaeal histones are doublets with two histone fold domains, a differentiated form hypothesized to predate the evolution of eukaryotic nucleosomes (Malik and Henikoff, 2003). Although the DNA wrap of archaeal nucleosomes resembles that of eukaryotic nucleosomes, the histones are too dissimilar in sequence from the four core eukaryotic histones to identify correspondences with specific domains of histone doublets. However, the sequencing of the giant Marseillevirus discovered in 2007 led to the realization that its genome encodes two histone doublets that are paired homologs of the four eukaryote core histones (Boyer et al., 2009). In Marseillevirus, H3-H5-H7-H8 is homologous to eukaryotic H2B and H2A, and H6-H7 is homologous to H4 and H3. Related viruses of the family Marseilleviridae have since been discovered infecting Acanthamoeba species worldwide (Sahmi-Bounsiar et al., 2021). Phylogenetic analysis places H3, H6, H7, and H8 as sister groups respectively to their H2A, H2B, H3, and H4 eukaryotic counterparts, consistent with divergence from the last eukaryotic common ancestor for all four histone folds (Erives, 2017).

We and others have previously used biochemical reconstitution and cryo-EM imaging to solve the high-resolution structure of Marseillevirus nucleosomes, which show a striking resemblance to their eukaryotic counterparts (Valencia-Sánchez et al., 2021; Liu et al., 2021). Unlike octameric eukaryotic nucleosomes, which wrap 147 bp of DNA, tetrameric Marseillevirus nucleosomes wrap only 121 bp of well-ordered DNA, despite being reconstituted on the Widom 601 artificial positioning sequence, selected for 147-bp wrapping of nucleosome cores. Because Marseillevirus doublets assemble into nucleosomes that are not fully wrapped by DNA, it was proposed that they are inherently unstable, perhaps to facilitate expression during the early stages of infection or for gene regulation (Liu et al., 2021). However, these studies were performed using a nucleosome-positioning sequence and cross-linking, and without isolation of viral chromatin in its native state.
form, the extent to which reconstituted Marseillevirus nucleosomes are representative of their conformation in virions is debatable (Vannini and Marazzi, 2021). By mapping the chromatin landscape of the viral genome, we aimed to understand the functional and evolutionary basis for viral nucleosomal packaging.

Here, we show that chromatin from Marseillevirus can be released from viral particles by breaching the capsid and permeabilizing the lipid membrane followed by either enzymatic or chemical cleavage. Using mass spectrometry, MNase-seq (Micrococcal Nuclease sequencing), and MPE-seq (Methidiumpropyl-EDTA sequencing), we confirmed that the 121-bp wrap observed by cryo-EM for reconstituted nucleosomes is also observed for endogenous histone doublets in the virion. However, in contrast to eukaryotic nucleosomes, which are 147-bp particles separated by 50-bp linkers, 121-bp Marseillevirus nucleosomes are tightly packed within the virion with 125-bp average spacing. Unlike eukaryotic nucleosomes, which are depleted from transcription start sites and phased downstream, we observed no depletion or phasing at genes. To determine whether tight packaging seen in the virion is inherent to Marseillevirus histone doublets, we reconstituted histones on multicopy arrays with a strong 147-bp positioning sequence. We found that most of the arrays were fully occupied by histone doublets and were resistant to chemical cleavage, suggesting that neighboring nucleosomes stabilize wrapping. Our findings indicate that Marseillevirus histone doublets have evolved for tight packing of 121-bp particles without sequence or genic preference, consistent with a viral packaging function, thus representing a previously unknown mode of genome packaging by histones.

RESULTS

Release of Marseillevirus chromatin from viral particles by nuclease digestion

The capsids of giant viruses that infect amoebae are resistant to treatments that disrupt other viruses. Previous attempts to release chromatin intact from Marseilleviridae capsids were reported to have been unsuccessful (Liu et al., 2021); however, by dialyzing in low pH conditions, Schrad and co-workers (Schrad et al., 2020) had shown that giant Mimivirus capsids can be breached, although without releasing their contents. We followed their protocol to open the Marseillevirus capsid and then dialyzed into a neutral buffer for micrococcal nuclease (MNase) digestion. Capillary gel electrophoresis revealed a striking ladder of protected particles with mono- and di-nucleosomes dominating and discernable tri- and tetra-nucleosome peaks (Figure 1A). Under the same permeabilization and digestion conditions, Drosophila mono-nucleosomes dominate, with only a minor di-nucleosomal peak (Figure 1B).

High yields of digested chromatin were obtained from samples treated with 0.5% NP-40 over a time course ranging from 1 min to 18 h (Figures 1A and S1A). By contrast, 0.1% digitonin had no effect (Figures 1C, S1B, and S1C). As NP-40 permeabilizes membranes by sequestering lipids, whereas digitonin permeabilizes cell membranes by displacing membrane sterols, our results suggest that the chromatin of Marseilleviridae, like that of giant Mimiviridae (Kuznetsov et al., 2010; Xiao et al., 2009; Zauberman et al., 2008; Suzan-Monti et al., 2007), is enclosed within an impermeable lipid membrane that would have protected it from low pH conditions, and subsequent permeabilization allowed for MNase to gain access to chromatin. Under our permeabilization and digestion conditions, 85%–90% of the DNA in virions was recovered as intact MNase-protected particles (Figure S1).
To ascertain the protein composition of Marseillevirus chromatin, we centrifuged viral suspensions following NP40 permeabilization and MNase treatment, then extracted total protein from the pellet, supernatant, and washes. We performed SDS-PAGE and excised bands from silver-stained gels as indicated (Figure 2A), including the \( \gamma C_24 \) 25 kDa bands corresponding to the predicted molecular weight (MW) of both Hb-Ha (25,949) and Hd-Hg (25,218) and neighboring bands. We also excised bands from the total protein extracted from reconstituted chromatin using histones produced in Escherichia coli and from a blank gel lane. Protein samples were digested with trypsin and subjected to mass spectrometry followed by a comparison of peptide MWs to those predicted for trypsinized Marseillevirus ORFs and likely contaminants. Only Hb-Ha and Hd-Hg peptides were found at consistently high levels in both viral chromatin and reconstituted chromatin samples (Figure 2B). A distantly related variant of Hb-Ha, Hc-He was (MW = 19,033), was detected in the chromatin fraction at a much lower level in the expected size range, consistent with previous mass spectrometry of viral extracts (Boyer et al., 2009). Taken together with the high DNA yield obtained from NP40-permeabilized and MNase-digested chromatin, our mass spectrometry (MS) data confirm that the Hb-Ha and Hd-Hg composition of viral chromatin matches that of the chromatin used for cryo-EM (Valencia-Sánchez et al., 2021).

We were intrigued by the presence in viral chromatin of the Hc-He variant, which is only 26% identical to Hb-Ha and lacks the long lysine-rich C-terminal tail, and wondered if the variant could be incorporated together with Hd-Hg into nucleosomal particles. To address this possibility, we used standard histone assembly conditions with equal molar amounts of Hz-He and Hz-Hy relative to X.l. and Xenopus laevis (X.l.) assemblies are shown as controls.
array and MNase digestion conditions, which suggests lower stability of DNA wrapping by the variant nucleosome.

Global analysis of the Marseillevirus genomic landscape

We performed paired-end DNA sequencing of Marseillevirus fragments and used Bowtie2 to align them to the GenBank T19 (later named Marseillevirus marseillevirus) genome assembly. This mapping revealed the over-representation of specific regions detected using raw reads from the primary Marseillevirus marseillevirus (T19) assembly (Boyer et al., 2009), sequencing without MNase treatment (no MNase), an MNase-seq digestion series (where 1 × = 10 U/million cells, 1/2 × = 5 U/million cells, and so on), and an MPE-seq digestion series (1, 10, 100, and 1,000 μM H₂O₂).

To confirm the validity of the GenBank assembly of the T19 circular genome, we performed de novo genome assembly for a representative T19 sample. We ran the SPAdes program on the 50-bp end reads, which provided >2,000× genomic coverage per sample. We obtained >99.9% coverage by the 8 largest contigs, where regions of over-representation are spanned in our assembly (Table S1). This demonstrates that

Figure 3. Regional over-representation of DNA during culturing of Marseillevirus

(A) IGV screenshots reveal an over-representation of specific regions detected using raw reads from the primary Marseillevirus marseillevirus (T19) assembly (Boyer et al., 2009), sequencing without MNase treatment (no MNase), an MNase-seq digestion series (where 1 × = 10 U/million cells, 1/2 × = 5 U/million cells, and so on), and an MPE-seq digestion series (1, 10, 100, and 1,000 μM H₂O₂).

(B) Same as (A) for G648 using the published assembly (Sahmi-Bounsari et al., 2021).

(C) Cumulative plot showing the rank of each over-represented region among random samples of the span from regions of the rest of the genome, where Rank 1 is the most over-represented sample. We excluded 1 kb on either side of the regions and at either end of the linearized genome assembly of the circular genome.

(D) To identify novel junctions at the borders of over-represented regions, we compared maps of total fragment occupancy with maps of 25-mers, where only a single 25-mer fragment end was mapped (unpaired). The y axis is autoscaled within groups. A novel junction will be recognized because any fragment spanning the junction without overlap will not be mapped, as the 25-mer ends will be oriented 3′-to-5′ relative to one another rather than 5′-to-3′.
the over-represented regions seen in T19 are inherent to the genome used for profiling chromatin and do not reflect a misassembly artifact.

To quantify over-representation, we randomly sampled fragments of similar lengths to each over-represented feature from the rest of the genome 1,000 times and plotted the abundance of each sample on a log-log cumulative plot for the raw reads from the original 2007 virus culture and the No-MNase paired-end reads from the 2020 virus culture used in this study (Figure 3C, left panel). This revealed that the 20-kb region over position 27,000 was ~2-fold over-represented in the original 2007 viral culture and ~2.7-fold over-represented in the 2020 culture relative to their respective genome-wide median abundances.

Likewise, the 2-kb region over position 318,400 was ~1.1-fold over-represented in the original 2007 viral culture and ~10-fold over-represented in the 2020 culture. Although the magnitude of over-representation of the 2-kb region in the original 2007 viral culture is small, it is ranked between #1 and #2 of the 1,000 randomly sampled regions, so it likely represents the incipient over-representation of this region in some virions in that culture. It is evident that over-representation of the 20-kb region continued to be maintained at 2–3-fold over-abundance during successive passages in Acanthamoeba polyphaga culture at limiting conditions (Chereji et al., 2019) used here. For G648, we observed a different region of over-representation in which an ~50-kb segment approximately centered over position 319,000 is present in all ten samples in the MNase concentration series (Figure 3B). Although the primary reads from the fastq files used for the original assembly of G648 were relatively sparse, quantitative sampling analysis showed that this ~50-kb region is also strongly over-represented, with a 1.9-fold excess, compared with a 2.6-fold excess for the no MNase sample (Figure 3C, right panel). As for T19, de novo assembly yielded >99.9% coverage by the 8 largest contigs, confirming the validity of the original assembly (Table S1).

Mapping of unpaired fragments from paired-end sequencing libraries shows them to be in large excess immediately outside of the over-represented regions as expected for the presence of novel tandem repeat junctions (Figure 3D). Our observations of different over-represented regions in two Marseillevirus isolates and increases in the abundance of over-represented regions in T19 during passage indicate that regional over-representation occurs during culturing of Marseillevirus in A. polyphaga. Analysis of long repeat-spanning fragments indicated that regional over-representation is accounted for by intrachromosomal tandem repeats rather than extrachromosomal circles (Figure S2).

The marseillevirus genome is densely packaged by 121-bp nucleosomes without linkers

Marseillevirus mono-nucleosomal DNA fragments released by MNase are smaller than those of Drosophila and show much less spacing between nucleosomes (Figure 1D). For example, the average size difference between both Marseillevirus di- and tri-nucleosomes on gels is 390 – 270 = 120 bp, which compares with 540 – 330 = 210 bp for Drosophila. Given that a eukaryotic nucleosome wraps 147 bp, our MNase digestion results suggest that there are no linkers separating 121-bp Marseillevirus nucleosomes but rather that they are abutted against one another. Such tight packing of Marseillevirus nucleosomes might explain why they remain insoluble even after MNase digestion, in contrast to Drosophila nucleosomes, which are quantitatively released by MNase into solution (Figures 1C and S1), as if tight packing within the permeabilized virion prevents the release of mono-nucleosomes.

At the highest MNase digestion levels, control Drosophila nucleosomes showed a dominant 147-bp peak but also a smear of sub-nucleosome-sized MNase digestion products punctuated by 70-, 90-, 110-, and 130-bp peaks (Figure 4A). We might attribute the presence of these intranucleosomal cleavages seen for the adapter-ligated and PCR-amplified sequenced fragments but not by gel analysis (Figure 1B) to single-strand nicking that would result in termination of Taq polymerase extension during PCR. Using the highest digestion level for both Marseillevirus T19 and G648, we observed a dominant 121-bp MNase-protected mono-nucleosome peak, a smaller ~250-bp di-nucleosome peak, and a small, variable, ~70-bp peak (Figures 4B and S3), confirming both the ~125-bp peak-to-peak spacing seen by gel analysis (Figure 1D) and the cryo-EM observations of a well-ordered 121-bp DNA wrap of reconstituted nucleosome cores (Valencia-Sánchez et al., 2021; Liu et al., 2021). These observations also indicate that Marseillevirus nucleosomes are less sensitive to intranucleosomal cleavages than are eukaryotic nucleosomes.

MPE-seq confirms the dense packing of marseillevirus nucleosomes

MNase is an endo/exo-nuclease that is known to preferentially digest AT-rich regions (Chung et al., 2010; McGhee and Felsenfeld, 1983) and to “nibble” on ends of intact nucleosomes under the limiting conditions (Chereji et al., 2019) used here (Figures 4A–4C and S4). Consistent with these observations, we found that ~90% of the cleavage sites occur between A/T base pairs (Figure S5A) and long AT-rich regions are preferentially digested (Figure S5B). To avoid this artifact, we used a small-molecule DNA cleavage reagent, methidiumpropyl-EDTA-Fe(II) [MPE(FeII)], which intercalates into DNA and hydrolyzes phosphodiester bonds, where H2O2 provides reactive oxygen for MPE-Fe(II)-catalyzed DNA cleavage (Cartwright et al., 1983). MPE-seq is performed similarly to MNase-seq but without exonuclease activity and without sequence bias (Ishii et al., 2015). When we treated Drosophila nuclei and permeabilized Marseillevirus particles with MPE-Fe(II), we observed mostly mono-nucleosome-sized particles for Drosophila, with an ~150-bp peak and periodic 10-bp internal cleavages (Figure 4E). MPE-seq of Marseillevirus chromatin from both T19 and G648 also revealed an ~150-bp peak with close concordance between samples and no 70-bp peak (Figures 4F and 4G). Whereas MNase digestion resulted in a 26-bp fragment size difference between Marseillevirus and Drosophila nucleosomes (Figure 4D),
MPE cleavage resulted in fragments of nearly the same size (Figure 4H). The difference between the sharp 121-bp MNase-seq peak and the broad MPE-seq distribution centered at ~150 bp may be accounted for by the exonucleolytic activity of MNase in which the endonuclease cleavage of the exposed DNA in linker regions is followed by end-nibbling (Ramachandran et al., 2017). By contrast, MPE(FeII) intercalates into DNA and efficiently cleaves in the presence of H₂O₂, releasing fragments with a minimum size of 4 bp (Van Dyke et al., 1982), but has no end-nibbling activity. As we recently reported for the cryo-EM structure assembled onto 147-bp Widom 601 DNA, only the central 121 bp are well ordered at high resolution, and the 13 bp at the DNA entrance and exit of the viral nucleosome are less ordered (Valencia-Sánchez et al., 2021). This looser wrap of the less-ordered 13-bp ends is expected to make them susceptible to MNase end-nibbling but evidently sufficient to block MPE(FeII) intercalation and cleavage (Figure S4).

MPE cleavage resulted in fragments of nearly the same size (Figure 4H). The difference between the sharp 121-bp MNase-seq peak and the broad MPE-seq distribution centered at ~150 bp may be accounted for by the exonucleolytic activity of MNase in which the endonuclease cleavage of the exposed DNA in linker regions is followed by end-nibbling (Ramachandran et al., 2017). By contrast, MPE(FeII) intercalates into DNA and efficiently cleaves in the presence of H₂O₂, releasing fragments with a minimum size of 4 bp (Van Dyke et al., 1982), but has no end-nibbling activity. As we recently reported for the cryo-EM structure assembled onto 147-bp Widom 601 DNA, only the central 121 bp are well ordered at high resolution, and the 13 bp at the DNA entrance and exit of the viral nucleosome are less ordered (Valencia-Sánchez et al., 2021). This looser wrap of the less-ordered 13-bp ends is expected to make them susceptible to MNase end-nibbling but evidently sufficient to block MPE(FeII) intercalation and cleavage (Figure S4).

Although the genomes of both Marseillevirus isolates have the same ~44% GC content as Drosophila, they have characteristic A/T homopolymeric runs in promoter regions (Oliveira et al., 2017) that might contribute to low-level internal MNase cleavages seen for Marseillevirus nucleosomes (Figure S5A). The lack of periodic 10-bp internal cleavages with MPE digestion (Figures 4F–4H) is unlikely to have resulted from non-histone protein binding, in that there are no plausible candidates detected by MS (Figure 2B), but rather demonstrates that Marseillevirus chromatin is inherently highly refractory to intranucleosomal cleavages by both MNase and MPE (Figure 4).

To examine the genome-wide organization of Marseillevirus nucleosomes, we separated fragments by size in 100-bp intervals and displayed a representative 10-kb region of the T19 and G648 genomes for MNase- and MPE-generated replicate samples. Typical Drosophila chromatin profiles showed patches of positioned nucleosomes for 101–200-bp fragments but little if any consistent positioning between replicates for subnucleosomal 1–100 bp or supranucleosomal 201–300 bp fragments for both MNase-seq (Figure 5A) and MPE-seq (Figure 5B). Although we observed patches of positioning for Marseillevirus T19 and G648, this was seen for all three size classes, but only for MNase-seq, with little if any distinctness in the chromatin landscape for MPE-seq. We also examined these representative 10-kb regions for evidence of nucleosome phasing by performing autocorrelation analysis (Figure 5C). This revealed consistent periodicities for both the MNase-seq and MPE-seq Drosophila

Figure 4. MNase- and MPE-digested Marseillevirus chromatin yields tightly packed nucleosome-sized DNA fragments

MNase-seq and MPE-seq were performed on Drosophila melanogaster S2 cells and on viral particles after pH 2 treatment, neutralization, and NP40 treatment, followed by Illumina DNA library preparation and PE25 paired-end sequencing. Densities (normalized counts per base pair) at each base pair are plotted.

(A–D) MNase-seq (A) Drosophila (3 samples), (B) Marseillevirus T19 (17 samples from two experiments), (C) Marseillevirus G648 (8 samples), and (D) superimposition of average for samples shown in (A)–(C). For the best resolution consistent with nucleosome integrity, a 5-min digestion time at 20 U/mL was used for all samples, based on a 512-fold dilution series (Figure S3).

(E–H) MPE-cleaved samples: (E) Drosophila (4 samples), (F) Marseillevirus T19 (6 samples), (G) Marseillevirus G648 (4 samples), and (H) superimposition of average for samples shown in (E)–(G). Curves for individual samples processed in parallel are superimposed to illustrate the degree of variation between replicate samples digested under different conditions used in this study.
data, most prominently for the 101–200 bp subset, as expected for nucleosome phasing. By contrast, the slight periodicities we observed for Marseillevirus T19 and G648 were mostly inconsistent between MNase-seq and MPE-seq. This lack of consistent phasing in Marseillevirus T19 and G648 nucleosomes suggests that they are densely packed without intervening linkers characteristic of eukaryotic nucleosomes.

The marseillevirus chromatin landscape lacks genic differentiation

We wondered whether the minor regularities seen in the MNase-digested Marseillevirus nucleosome landscapes and by autocorrelation analysis (Figure S6) correspond to genic regions, which are annotated as open reading frames (ORFs) and are separated by very short spans that are usually AT-rich (Oliveira et al., 2017). To investigate this possibility, we aligned all 191 ORFs that are R600 bp at their 5’ ends and averaged each nucleotide position over gene bodies. For comparison, we chose the first 191 ORFs ≥600 bp from the chronological list of Drosophila ORFs. The alignment of MNase-generated 101–200 bp fragments to the 191 Drosophila ORFs revealed the characteristic translational phasing pattern, with a prominent +1 nucleosomal peak just downstream of the ORF 5’ end and phased +2 and +3 nucleosomes with reduced occupancy for all digestion levels over a 128-fold range (Figure 6A). MPE-generated fragments showed the same 5’-aligned translational phasing patterns for 101–200 bp fragments, confirming that maps produced using MPE and MNase are approximately concordant for control eukaryotic nucleosomes. By contrast, MNase-generated Marseillevirus T19 and G648 101–200 bp fragments displayed little if any 5’ phasing when plotted on the same scale (Figure 6A), although when the scale was expanded a positioned +1 nucleosome was observed. However, no nucleosome positioning was seen for MPE-generated T19 and G648 5’-aligned ORFs (Figures 6A and 6C), indicating that the minor +1 nucleosome positioning that was seen with MNase is likely attributable to its aggressive endo/exonuclease activity.
Figure 6. Phasing observed by MNase-seq for Marseillevirus is not observed by MPE-seq
(A and B) All 191 Marseillevirus ORFs ≥600 bp were aligned (A) at their 5' ends and (B) at the stop codon at their 3' ends. Normalized counts are plotted at a 1-bp resolution over the 600-bp span. Left to right: Drosophila S2 cells (8 MNase and 4 MPE samples), Marseillevirus T19 (8 MNase and 4 MPE samples), and G648 (10 MNase and 8 MPE samples) plotted on the same scale, and T19 and G648 plotted on a vertically expanded scale to display the differences between MPE and MNase digestions. Curves with bluish colors are from MNase-seq for different samples digested in parallel, and curves with reddish colors are for MPE-seq samples. The dotted line is for a no MNase control.

(C) ORFs showing high variability based on autocorrelation (Figure S5) downstream of the 5' end for the 101–200 bp size class were displayed within 2-kb regions for four technical replicates each. For MNase-seq, samples were digested in (top to bottom) 50, 25, 12.5, and 6.25 mU MNase/reaction. For MPE-seq, samples were digested using (top to bottom) 1, 10, 100, and 1,000 mM H₂O₂. In contrast to Drosophila nucleosomes, which show phasing of 101–200 bp fragments resulting from either MNase or MPE digestion, Marseillevirus nucleosomes are not detectably phased when assayed by MPE-seq, and MNase-seq shows phasing irrespective of fragment size, indicating internal cleavages. Tracks are group autoscaled within sets of four.
(Figure S3A). The lack of 5’ phasing, a universal characteristic of eukaryotic genes, indicates that the packaged Marseillevirus genome is likely to be transcriptionally inactive upon release from the capsid during infection.

Alignment of the 191 ORFs at the stop codon of their 3’ ends showed extreme sensitivity of Drosophila nucleosomes to MNase levels not seen for MPE-generated fragments (Figure 6B), which is consistent with the partial unwrapping of AT-rich ORF 3’-end DNA from nucleosome cores and sensitivity to MNase exonucleolytic activity. A similar MNase sensitivity and MPE insensitivity were seen for T19 and G648 nucleosomes at the very 3’ ends of ORFs (Figure 6B), which is likewise attributable to AT-rich regions at Marseillevirus 3’ ends. However, unlike chromatin at the 3’ ends of Drosophila ORFs, both T19 and G648 chromatin displayed an average peak of MNase resistance just upstream of the 3’ end. As this peak was absent from MPE-digested average profiles, we attribute it to internal MNase cleavage of a subset of nucleosomes that are relatively excluded from neighboring AT-rich regions rather than to 3’ nucleosome phasing. The lack of any chromatin accessibility features punctuating genic regions implies that Marseillevirus nucleosomes have evolved exclusively for packaging within the virion.

**Dense packing of reconstituted marseillevirus chromatin on widom 601 arrays**

We next asked whether previous reconstitutions of Marseillevirus nucleosomes on single 601 sequences (Valencia-Sánchez et al., 2021; Liu et al., 2021) failed to show fuller wrapping because of the lack of neighboring nucleosomes, which we find are closely abutted with ~125-bp spacing in virions. Accordingly, we performed MPE and MNase digestion on native and cross-linked reconstituted Marseillevirus chromatin that had been assembled at 60 or 300 μg/mL concentrations onto a three-copy 601 array and onto a 12-copy 601 array. Following DNA extraction, we prepared sequencing libraries and performed paired-end sequencing, aligning fragments from native or cross-linked 3-copy and 12-copy chromatin digests to the 3-copy 601 array. For the 3-copy native chromatin, MPE-seq revealed similar chromatin landscapes for the 60 and 300 μg/mL samples, with higher occupancy over the three 147 bp 601 positioning sequences than over the intervening 40-bp linkers, although not as high as for Xenopus control core histones assembled on the same 3-copy array (Figure 7A, left). The transitions between the 601 and linker sequence were sharply defined, indicating the precise positioning of 147-bp 601 particles over the 601 sequence. A much smaller fraction of 147-bp 601 particles was observed for MPE digestion of cross-linked 3-copy chromatin (Figure 7A, right).

To quantify the relative abundance of cleavages precisely at 147-bp 601 particle ends, we plotted the length distributions for each 3-copy sample (Figure S7A). In each case, MPE digestion of 3-copy array chromatin resulted in 1-bp wide fragment length peaks at 147, 334, and 521 bp (Table S2). As the entire 3-copy array is 521 bp, and 334 bp is exactly the size expected for a 601-linker-601 spanning fragment, it is evident that MPE endonucleolytically digests to completion within linker regions without detectable encroachment into the nucleosome-wrapped particles. Nineteen percent of the nucleosomes in the 60 μg/mL sample were 147 bp and precisely phased over the 601 positioning sequence, and 147, 334, and 521 bp fragment ends accounted for 39% of the total. By comparison, control Xenopus nucleosomes assembled on the same 3-copy 601 array and subjected to MPE-seq yielded 38% 147-bp particles, and 147, 334, and 521 bp fragment ends accounted for 73% of the total. This illustrates the inherent tendency of adjacent Marseillevirus nucleosomes to interact and partially overcome DNA sequence-directed positioning. High histone occupancy over the linkers between 601 positioning sequences contrasts with the 121-bp wrap seen for Marseillevirus nucleosomes reconstituted onto single 601 sequences (Valencia-Sánchez et al., 2021; Liu et al., 2021) and suggests that closely abutted Marseillevirus nucleosomes stabilize one another and prevent unwrapping (Figure S3A).

After cross-linking, only ~1% of 3-copy Marseillevirus chromatin arrays corresponded to 147-bp particles, and 147, 334, and 521 bp fragment ends accounted for only ~2% of the total, superimposed over a broad distribution of fragment lengths (Figure S7B). For 12-copy cross-linked reconstructed arrays, MPE-seq resulted in the total absence of 1-bp peaks. By comparison, control Xenopus nucleosomes reconstituted on 3-copy arrays yielded 14% 147-bp particles, and 147, 334, and 521 bp fragment ends accounted for 26% of the total (Figure S7B). This suggests that a larger fraction of the MPE cleavages are uniformly distributed on these arrays using Marseillevirus histones than Xenopus histones.

In contrast to the results with MPE, MNase digestion of the same reconstituted chromatin showed a distribution of fragment lengths ~20-25 bp smaller than the three discrete fragment lengths produced by MPE digestion of Marseillevirus chromatin. MNase digestion of native Marseillevirus chromatin produced fewer than 1% precisely positioned cleavages and resulted in a rough profile, whereas for Xenopus nucleosomes assembled on the same 3-copy Widom 601 array, 70% of the cleavages were precisely positioned and resulted in a clean sawtooth pattern (Figure 7A). Reduced nucleosome positioning on 601 arrays measured by both MPE-seq and MNase-seq distinguishes Marseillevirus doublet histones from eukaryotic core histones and recapitulates the situation in viri. One interpretation of the better positioning of Xenopus than Marseillevirus nucleosomes is that the Widom 601 sequence was evolved by SELEX (systematic evolution of ligands by exponential enrichment) selection from chemically synthesized random DNA sequences using eukaryotic histone cores (Lowary and Widom, 1998) and so might have a stronger inherent preference for Xenopus over Marseillevirus nucleosomes. If so, then we expect that by removing most of the 40-bp linker DNA, differences in relative positioning between Marseillevirus and Xenopus core particles would diminish with the reduction in alternative positions along the array. Accordingly, we reconstituted Marseillevirus and Xenopus nucleosomes on plasmid DNA containing a 7-copy Widom 601 sequence with 12-bp linkers. In this assembly reaction, the 601 × 7 insert and the pUC19 vector were separated by digestion with BamHI and KpnI restriction endonucleases prior to assembly so that the vector control would serve as a negative control. For assembly, we used six different histone:DNA ratios ranging from 0.25 to 2.5. Interestingly, the
Marseillevirus assemblies were successful for all 6 samples, whereas the Xenopus assemblies completely precipitated out of solution when the histone:DNA ratio exceeded 1.5 (Figure 7B). Unlike the artificial Widom 601 sequence, the pUC vector is eukaryotic, which implies a remarkable adaptation of Marseillevirus doublet histones for diverse sequences that they have never encountered. We observed nearly identical protections for the assembled samples by both MPE-seq and MNase-seq (Figures 7C, S7C, and S7D). Thus, when Marseillevirus nucleosomes are forced to be packed tightly together, they are at least as stable as eukaryotic nucleosomes.

The overall occupancy of the 601 × 7 array was 3-fold higher than for the pUC vector using MPE-seq and 12-fold higher using MNase-seq with only minor differences between samples and between fully assembled Marseillevirus and Xenopus nucleosomes (Figure 7C). This implies that Marseillevirus tetramers and eukaryotic octamers follow the same DNA sequence energetic preferences. With MPE cleavage, the very weak 147 + 12 = 159-bp periodicity for the 601 array with 12-bp linkers compared with the strong 147 + 40 = 187-bp periodicities for the 3- and 12-copy 601 arrays is attributable in part to steric hindrance, and indeed, the 4 bp “bump” rather than a “dip”
between nucleosomes from both Marseillevirus and Xenopus is very likely attributable to protection by intercalated MPE(Fell) itself (Van Dyke et al., 1982). The very similar protections of both the 601 7-copy array and the pUC19 vector using both MPE-seq and MNase-seq emphasize the overall structural similarities between Marseillevirus and eukaryotic core particles when they are closely packed together on a strong nucleosome-positioning sequence.

Finally, we assembled Marseillevirus and Xenopus histones at histone:DNA ratios of 0.25–2.5 on a 3-copy 147-bp DNA array that lacked linkers (441 bp). After MNase digestion, DNA purification, and capillary gel electrophoresis, we observed fragments showing complete protection of the DNA, with some sub-nucleosomal and nucleosome-sized particles appearing with increasing histone:DNA ratios for both Marseillevirus and Xenopus (Figure 7D). It would appear that adjacent positioning sequences with zero base pairs in between result in interference between particles during assembly regardless of whether the histone cores comprise eukaryotic octamers or Marseillevirus doublet tetramers.

**DISCUSSION**

We have shown that Marseillevirus nucleosomes can be recovered intact from virions and used to elucidate nucleosome organizational features using mass spectrometry, MNase-seq, and MPE-seq. These methods reveal particles that differ from eukaryotic nucleosomes in being refractory to internal cleavages and tightly packed into a landscape without linker DNA or phasing around genes. Taken together, our findings reveal a dense chromatin landscape that may have evolved to maximize the protection of viral DNA for survival during infection in the amoeba cytoplasm. This mode of chromatin organization differs drastically from that of eukaryotes, where nucleosomes not only protect DNA but also have evolved for gene regulation by limiting access to regulatory elements (Kornberg and Lorch, 2020). Given the close structural superimposition of the Marseillevirus nucleosome with the eukaryotic nucleosome (Valencia-Sánchez et al., 2021; Liu et al., 2021), our finding that Marseillevirus chromatin lacks linkers or genic punctuation is especially remarkable. Marseillevirus chromatin is also unlike that of well-studied archaeal chromatin in which 60-bp single-wrapped units are thought to form long slinkies that dynamically open and close (Mattiroli et al., 2017; Bowerman et al., 2021). Rather, tight packing without linkers implies an inherently stiff fiber. Therefore, Marseillevirus nucleosomes represent a previously unknown mode of chromatin packaging. While our manuscript was under review, a cryo-EM structure of human telomeric chromatin was published showing a tightly packed columnar structure with 132-bp spacing (Soman et al., 2022), similar to our model for published showing a tightly packed columnar structure with review, a cryo-EM structure of human telomeric chromatin was molecular packaging. While our manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manuscript was under manifold.

Molecular Cell 82, 4458–4470, December 1, 2022
to heterotypic nucleosomes that later evolved into the four eukaryotic core histones (Malik and Henikoff, 2003). Our evidence that Marseillevirus doublet histones are well-suited for viral packaging is consistent with the possibility that eukaryotic histones have evolved from virus-encoded histone doublits that infected a host proto-eukaryote just as present-day Marseilleviridae infect Acetamonadea around the globe. There has been considerable recent interest in the hypothesis that the eukaryotic nucleus evolved from a viral factory (Talbert et al., 2022; Liu and Krupovic, 2022), and if so, the as-yet-unknown mechanism whereby a tightly packed viral particle transitions to a fully functional viral factory may shed light on the earliest stages of eukaryotic evolution.

Limitations of the study
Although previous work has shown that Marseillevirus histones are abundantly present in viral factories (Liu et al., 2021), we expect that the viral chromatin landscape will be extensively re-modeled to accommodate replication and transcription. The machinery that accomplishes this transition is also highly uncertain, insofar as most Marseillevirus-encoded proteins are unique to the Marseilleviridae (Talbert et al., 2022) with no known eukaryotic homologies to suggest function, and it is possible that the transition is entirely governed by host-encoded factors. Elucidating the machinery and dynamics of this transition, and testing for the role of the H2A-H2B variant in the transition, remains an exciting future challenge that has the potential of providing insights into the evolution of eukaryotes and the origin of the nucleus.

STAR*METHODS
Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **RESOURCE AVAILABILITY**
  - Lead contact
  - Materials availability
  - Data and code availability
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
- **METHOD DETAILS**
  - Viruses
  - Drosophila cells
  - Viral capsid opening and permeabilization
  - MNase-seq
  - MPE-seq
  - DNA extraction
  - Proteomics
  - Library preparation
  - Expression and purification of marseillevirus histone tetramers
  - Purification of widom 601 DNA arrays
  - Reconstitution of nucleosomes on widom 601 arrays
  - Gradient fixation (GraFix) of nucleosome arrays
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
  - De novo genome assembly
- **DATA AND SOFTWARE AVAILABILITY**

SUPPLEMENTAL INFORMATION
Supplemental information can be found online at https://doi.org/10.1016/j.molcel.2022.10.020.

ACKNOWLEDGMENTS
We thank Jim Kadonaga for MPE, Fred Hutch Genomics Shared Resource for sequencing services, Fred Hutch Proteomics & Metabolomics Shared Resource for mass spectrometry services, Pamela Dyer for information on the 601-147-12 copy pUC19 plasmid, and Kami Dyer for information on the manuscript. T.D.B., P.B.T., and S.H. are supported by the Howard Hughes Medical Institute. P.D.I., M.I.V.-S., and K.-J.A. were supported by a grant from the David and Lucile Packard Foundation.

AUTHOR CONTRIBUTIONS
Conceptualization, K.-J.A. and S.H.; investment, T.D.B., P.D.I., M.I.V.-S., J.G.H., P.B.T., B.L.S., K.-J.A., and S.H.; writing – original draft, S.H.; writing – review & editing, T.D.B., P.D.I., M.I.V.-S., J.G.H., P.B.T., K.-J.A., and S.H.

DECLARATION OF INTERESTS
The authors declare no competing interests.

Received: April 6, 2022
Revised: September 9, 2022
Accepted: October 19, 2022
Published: November 11, 2022

REFERENCES
Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Pyshelski, A.D., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. 19, 455–477.
Bhardwaj, S.K., Mall, S.G., Olufemi, L., Bramha, S., Kundu, S., Hota, S.K., Persinger, J., and Bartholomew, B. (2020). Minicell-specific specificity and allosteric switch of the ISW1a ATP-dependent chromatin remodeler in transcription regulation. Nat. Commun. 11, 5913.
Bowerman, S., Wereszczyński, J., and Luger, K. (2021). Archaeal chromatin ‘slinkies’ are inherently dynamic complexes with defected DNA wrapping pathways. eLife 10.
Boyer, M., Yutin, N., Pagnier, I., Barrassi, L., Fournous, G., Espinosa, L., Robert, C., Azza, S., Sun, S., Rossmann, M.G., et al. (2009). Giant Marseillevirus highlights the role of amoebae as a melting pot in emergence of chimeric microorganisms. Proc. Natl. Acad. Sci. USA 106, 21848–21853.
Cartwright, I.L., Hertzberg, R.P., Dervin, P.B., and Elgin, S.C. (1983). Cleavage of chromatin with methidiumpropyl-EDTA: iron(II). Proc. Natl. Acad. Sci. USA 80, 3213–3217.
Chereji, R.V., Bryson, T.D., and Henikoff, S. (2019). Quantitative MNase-seq accurately maps nucleosome occupancy levels. Genome Biol. 20, 198.
Chua, E.Y., Vasudevan, D., Davey, G.E., Wu, B., and Davey, C.A. (2012). The mechanics behind DNA sequence-dependent properties of the nucleosome. Nucleic Acids Res. 40, 6338–6352.
Chung, H.R., Dunkel, I., Heise, F., Linke, C., Krobitsch, S., Ehrenhofer-Murray, A.E., Sperling, S.R., and Vingron, M. (2010). The effect of micrococcal nuclease digestion on nucleosome positioning data. PLoS One 5, e15754.
Dyer, P.N., Edayathumangalam, R.S., White, C.L., Bao, Y., Chakravarthy, S., Muthurajan, U.M., and Luger, K. (2004). Reconstitution of nucleosome core particles from recombinant histones and DNA. Methods Enzymol. 375, 23–44.
Erives, A.J. (2017). Phylogenetic analysis of the core histone doublet and DNA top II genes of Marseilleviridae: evidence of proto-eukaryotic provenance. Epigenetics Chromatin 10, 55.
Furuya, T., Codomo, C.A., and Henikoff, S. (2013). Reconstitution of histones on budding yeast centromeric DNA. Nucleic Acids Res. 41, 5769–5783.

Grau, D., Zhang, Y., Lee, C.H., Valencia-Sánchez, M., Zhang, J., Wang, M., Holder, M., Svetlov, V., Tan, D., Nuider, E., et al. (2021). Structures of monomer and dimeric PRC2:EZH1 reveal flexible modules involved in chromatin compaction. Nat. Commun. 12, 714.

Hewish, D.R., and Burgoyne, L.A. (1973). Chromatin sub-structure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. Biochem. Biophys. Res. Commun. 52, 504–510.

Ishii, H., Kadonaga, J.T., and Ren, B. (2015). MPE-seq, a new method for the compaction. Nat. Commun. 6, 12746.

Liu, Y., Bisio, H., Toner, C.M., Jeudy, S., Philippe, N., Zhou, K., Bowerman, S., Liu, Y., and Krupovic, M. (2022). Genome chromatinization in giant double-stranded DNA viruses. Trends Biochem. Sci. 47, 3–5.

Lowary, P.T., and Widom, J. (1998). New DNA sequence rules for high affinity binding to histone octamer and sequence-directed nucleosome positioning. J. Mol. Biol. 276, 19–42.

Luk, E., Ranjan, A., Fitzgerald, P.C., Mizuguchi, G., Huang, Y., Wei, D., and Wu, C. (2010). Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosomes. Cell 143, 725–736.

Malik, H.S., and Henikoff, S. (2003). Phylogenomics of the nucleosome. Nat. Struct. Biol. 10, 882–891.

Martin, M. (2010). Cutadapt removes adapter sequences From high-throughput sequencing reads. EMBnet. j. 17, 1138–1143. https://doi.org/10.14806/ej.17.1.200.

Mattrilli, F., Bhattacharyya, S., Dyer, P.N., White, A.E., Sandman, K., Burkhart, B.W., Byrne, K.R., Lee, T., Ahn, N.G., Santangelo, T.J., et al. (2017). Structure of histone-based chromatin in Archaea. Science 357, 609–612.

McGhee, J.D., and Felsenfeld, G. (1983). Another potential artifact in the study of nucleosome phasing by chromatin digestion with micrococcal nuclease. Cell 32, 1205–1215.

Noll, M. (1974). Subunit structure of chromatin. Nature 251, 249–251.

Oliveira, G.P., Lima, M.T., Arantes, T.S., Assis, F.L., Rodrigues, R.A.L., da Fonseca, F.G., Bonjardim, C.A., Kroon, E.G., Colson, P., La Scola, B., and Abrahão, J.S. (2017). The investigation of promoter sequences of Marseillieviruses highlights a remarkable abundance of the AAATATT motif in intergenic regions. J. Virol. 91 e01088–17.

Pagnier, I., Reteno, D.G., Saadi, H., Boughalmi, M., Gaia, M., Slimani, M., N'gounga, T., Bekoltz, M., Colson, P., Raoult, D., and Scola, B.L. (2013). A decade of improvements in Mimiviridae and Marseillieviridae isolation from amoeba. Intervirology 56, 354–363.

Pereira, S.L., Grayling, R.A., Lurz, R., and Reeve, J.N. (1997). Archaeal nucleosomes. Proc. Natl. Acad. Sci. USA 94, 12633–12637.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842.

Ramachandran, S., Ahmad, K., and Henikoff, S. (2017). Transcription and remodeling produce asymmetrically unwrapped nucleosomal intermediates. Mol. Cell 68, 1038–1053.e4.

Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 44, W160–W165.

Sahmi-Bounsiar, I., Rolland, C., Aherfi, S., Boudjemaa, H., Levasseur, A., La Scola, B., and Colson, P. (2021). Marseillieviruses: an update in 2021. Front. Microbiol. 12, 648731.

Schad, J.R., Abrahão, J.S., Cortines, J.R., and Parent, K.N. (2020). Structural and proteomic characterization of the initiation of giant virus infection. Cell 181, 1046–1061.e6.

Shchevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996). Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal. Chem. 68, 850–858.

Soman, A., Wong, S.Y., Korolev, N., Surya, W., Lattmann, S., Vogirala, V.K., Chen, Q., Berezhnoy, N.V., Noort, J.V., Rhodes, D., and Nordenskiod, L. (2022). Columnar structure of human telomeric chromatin. Nature 609, 1048–1055. https://doi.org/10.1038/s41586-022-05236-5.

Stark, H. (2010). GraFix: stabilization of fragile macromolecular complexes for single particle cryo-EM. Methods Enzymol. 481, 109–126.

Suzan-Monti, M., La Scola, B., Barrassi, L., Espinosa, L., and Raoult, D. (2007). Ultrastructural characterization of the giant volcano-like virus factory of Acanthamoeba polyphaga Mimivirus. PLoS One 2, e328.

Talbert, P.B., Armache, K.J., and Henikoff, S. (2022). Viral histones: pick-pocket’s prize or primordial progenitor? Epigenetics Chromatin 15, 21.

Valencia-Sánchez, M.I., Abini-Agbomson, S., Wang, M., Lee, R., Vasilyev, N., Zhang, J., De Ioannes, P., La Scola, B., Talbert, P., Henikoff, S., et al. (2021). The structure of a virus-encoded nucleosome. Nat. Struct. Mol. Biol. 28, 413–417.

Van Dyke, M.W., Hertzberg, R.P., and Dervan, P.B. (1982). Map of distamycin, netropsin, and actinomycin binding sites on heterogeneous DNA: DNA cleavage-inhibition patterns with methidiumpropyl-EDTA-Fe(II). Proc. Natl. Acad. Sci. USA 79, 5470–5474.

Vannini, A., and Marazzi, I. (2021). A small nucleosome from a weird virus with a fat genome. Mol. Cell 81, 3447–3448.

Xiao, C., Kuznetsov, Y.G., Sun, S., Hafenstein, S.L., Kostyuchenko, V.A., Chipman, P.R., Suzan-Monti, M., Raoult, D., McPherson, A., and Rossmann, M.G. (2009). Structural studies of the giant Mimivirus. PLoS Biol. 7, e52.

Zauberman, N., Mutsafi, Y., Halevy, D.B., Shimoni, E., Xiao, C., Sun, S., and Minsky, A. (2008). Distinct DNA exit and packaging portals in the virus Acanthamoeba polyphaga mimivirus. PLoS One 3, e92.
## STAR METHODS

### KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|---------------------|--------|------------|
| **Bacterial and virus strains** | | |
| Escherichia coli BL2-codon plus(DE3)-RIL | Agilent Technologies | 230245 |
| **Chemicals, peptides, and recombinant proteins** | | |
| Micrococcal Nuclease | Sigma | N3755 |
| methidiumpropyl-EDTA-Fe(II) (MPE) | James Kadonaga | www.pnas.org/cgi/doi/10.1073/pnas.1424804112 |
| Nonident™ P 40 Substitute (NP-40) | Sigma | 74385 |
| Bathophenanthroline | Sigma | 133159 |
| **Critical commercial assays** | | |
| KAPA HyperPrep Kit | Roche | KK8504 |
| HighPrep PCR Clean-up System beads | MagBio Genomics | AC-60005 |
| **Deposited data** | | |
| MNase-seq and MPE-seq data and metadata | This paper | GEO: GSE193224 |
| Figure 2A gel | This paper | https://data.mendeley.com/datasets/x3rdxz89h9/1 |
| Figure 2C gel | This paper | https://data.mendeley.com/datasets/kc29z9n3jc/1 |
| Figure 7D gel | This paper | https://data.mendeley.com/datasets/sg6j5w795/1 |
| Figure S1A gel | This paper | https://data.mendeley.com/datasets/cp88jmb3f/1 |
| Mass spectrometry peptide analyses of gel slices (Figure 2) | This paper | https://data.mendeley.com/datasets/j8t5z4X9b/1 |
| Marseillevirus G648 Genome Assembly | This paper | https://data.mendeley.com/datasets/zhk9nx8v5g/1 |
| Marseillevirus G648 ORFs | This paper | https://data.mendeley.com/datasets/smncgck78m/1 |
| Widom 601 arrays | This paper | https://data.mendeley.com/datasets/vcconverted/6md/1 |
| **Experimental models: Cell lines** | | |
| Drosophila melanogaster S2 cells | Invitrogen | 10831-014 |
| **Experimental models: Organisms/strains** | | |
| Marseillevirus marseillevirus Strain T19 | Bernard La Scola | Boyer et al. (2009) |
| Marseillevirus marseillevirus Strain G648 | Bernard La Scola | Sahmi-Bounsiar et al. (2021) |
| **Recombinant DNA** | | |
| 601-147x12 plasmid | K. Luger | Dyer et al. (2004) |
| pET-24B plasmid | Valencia-Sánchez et al. (2021) | Addgene_127852 |
| 3_187_widom_601 plasmid | Valencia-Sánchez et al. (2021) | Valencia-Sánchez et al. (2021) |
| 3_147_widom_601 no linker array | Genscript, Inc. | This paper |
| **Software and algorithms** | | |
| Bowtie2 | Langmead and Salzberg, 2012 | http://bowtie-bio.sourceforge.net/bowtie2/index.shtml |
| Bedtools | Quinlan and Hall (2010) | https://bedtools.readthedocs.io/en/latest |
| deepTools v. 3.3.1 | https://deeptools.readthedocs.io/en/develop/content/installation.html | Ramírez et al. (2016) |
| BLAT | https://kentinformatics.com | RRID: SCR_011919 |
RESOURCE AVAILABILITY

Lead contact
Requests for materials should be directed to and will be fulfilled by the lead contact, Steven Henikoff (steveh@fredhutch.org).

Materials availability
Materials are available from Bernard La Scola, Karim-Jean Armache or Steven Henikoff upon request.

Data and code availability
- Sequencing data have been deposited in GEO and are publicly available as of the date of publication. Accession numbers are listed in the key resources table. Original gel images have been deposited at Mendeley and are publicly available.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Plasmid propagation was performed in Escherichia coli DH5α. Protein expression was performed in Escherichia coli BL2-codon plus(DE3)-RIL. Control chromatin for MNase-seq and MPE-seq profiling was obtained using Drosophila melanogaster S2 cells.

METHOD DETAILS

Viruses
Marseillevirus Strains T19 (Boyer et al., 2009) and G648 (Sahmi-Bounsiar et al., 2021) were cultured in Acanthamoeba polyphaga as described (Pagnier et al., 2013).

Drosophila cells
Drosophila S2 cells were grown in HyClone SFX Insect Cell Culture Media (Cytiva SH30278.02) supplemented with 18mM L-Glutamine, seeded at 2x10⁶/mL three times per week, and harvested with >95% viability at mid-log phase. A total of 1x10⁷ cells were centrifuged in a swinging-bucket rotor for 4 min at 700xg at 25°C and washed twice in cold 1x PBS. The cell pellet was resuspended in 1 mL TM2+PI (10mM Tris pH 8, 2mM MgCl₂ + Protease Inhibitor, Sigma 11836170001) and chilled in ice water for 1 min. NP-40 was added to 0.5% and vortexed gently at half maximum speed for ~3 sec and returned to the ice water slurry. Release of nuclei was ascertained by microscopic observation of aliquots until ≥80% of cellular membranes were disrupted (~3 min). Nuclei were centrifuged 10 min at 150xg at 4°C, washed twice in 1.5 mL TM2, 1 mM PMSF and finally resuspended in 200 μL TM2. Each digestion reaction contained either 50K or 150K nuclei per timepoint.

Viral capsid opening and permeabilization
We followed a viral opening procedure developed for Mimivirus (Schrad et al., 2020), with minor modifications. Each sample from a purified Marseillevirus culture was applied to the membrane of a 7K MWCO Dialysis Unit (Thermo 69562) and dialyzed against 250 mL of 20 mM sodium phosphate buffer adjusted to pH2, 2 mM MgCl₂, 1 mM PMSF for 2 h at 25°C, followed by dialysis against 250 mL TM2, 1 mM PMSF for 3 h. In some experiments viral suspensions were diluted with TM2 prior to dialysis. To equalize the amounts of DNA from Marseillevirus and Drosophila control cells, DNA was extracted from a range of volumes of purified Marseillevirus culture in parallel with a known number of S2 cells. Either 0.5 μL or 1.5 μL of Marseillevirus culture was used per timepoint.

To improve recovery, we included the non-ionic detergent NP-40, which is widely used for chromatin release from cells. Recovery from Marseillevirus T19 particles was vastly improved, ~85-90% in 0.5% NP-40 using the maximum MNase digestion conditions that had resulted in mostly mono-nucleosomes in the previous Drosophila experiments. High yields and fragment size distributions were obtained from samples treated with 0.5% NP-40 over a time course ranging from 1 minute to 18 hours. In contrast, 0.1% of the non-ionic detergent digitonin had no effect. As NP-40 permeabilizes membranes by sequestering lipids, whereas digitonin permeabilizes cell membranes by displacing membrane sterols, our results suggest that viral chromatin is enclosed within a lipid membrane that must be permeabilized for MNase to access chromatin.

MNase-seq
MNase-seq was performed as previously described (Chereji et al., 2019). Briefly, Micrococcal Nuclease (MNase, Sigma N3755) was reconstituted to a concentration of 1 unit per 5 μL with nuclease-free water, aliquoted and stored at -20°C. MNase was thawed on ice and diluted stepwise 1:1 from 1x to 1/512x with TM2 buffer (10mM Tris pH 8, 2 mM MgCl₂) where 1x is 20 mU/μL. The volume of each Drosophila or Marseillevirus timepoint was adjusted to 166 μL with TM2+PI, and 2.5 μL of each MNase dilution was added with mixing, then heated to 37°C for 1 min. MNase was activated by addition of 3.5 μL of 100 mM CaCl₂ and incubated for 5 min at 37°C. Reactions were stopped with 172 μL of 2xSTOP solution (10 mM Tris pH 8, 2 mM MgCl₂, 340 mM NaCl, 20 mM EDTA, 4 mM
EGTA, 100 μg/mL RNase, DNase-free (Sigma11119915001) and DNA was extracted as described below. For reconstituted nucleosomes, we digested with 0.2 U MNase/μg DNA for 7 min at room temperature (Bhardwaj et al., 2020). We performed digestions over a time course and concentration range that we had previously found to be sufficient for digesting chromatin from Drosophila S2 cells into mono- and oligo-nucleosomes (Chereji et al., 2019). In that study, 1 minute digestions at 2.5 U/million cells had yielded an electrophoretic ‘ladder’ dominated by oligonucleosomes, 5 minute digestions yielded mostly mono-nucleosomes, and 30 minute digestions resulted in partially degraded mono-nucleosomes. We prepared MNase-seq Illumina sequencing libraries from the resulting MNase-digested fragments for both NP-40-treated and untreated libraries, performed paired-end DNA sequencing, and mapped the resulting read pairs to the annotated Marseillevirus T19 genome assembly. On average 97% of fragments mapped to this assembly, confirming the purity of our viral sample. Fragment lengths inferred from sequencing data are smaller than what we observed by Agilent 4200 TapeStation analysis of purified DNA following MNase digestion (Figure 1C), which reflects the selection for smaller fragments during end-polishing and PCR.

**MPE-seq**

MPE-seq was performed as described by Ishii et al. (Ishii et al., 2015). Briefly, the opened Marseillevirus and S2 nuclei were treated with hydrogen peroxide across the range of 0.001-1 mM and cleavage was induced by addition of methidiumpropyl-EDTA-Fe(ll) (MPE, a generous gift from Jim Kadonaga), which is MPE complexed with ammonium iron(II) sulfate at 10 μM or 40 μM for 5 min at room temperature. The reaction was quenched with 6 mM bathophenanthroline (Sigma 133159), followed by 2xSTOP buffer at a volume equal to the sample and total DNA extracted as described below. For reconstituted nucleosomes, we digested with 40 μM MPE using at a ratio of 1 mM H₂O₂/650 ng DNA for 5 min at room temperature.

**DNA extraction**

Sample volumes were adjusted to ~340 μL with TM2. To each sample, 3.4 μL 10% SDS and 2.5 μL Proteinase K (20 mg/ml) were added and incubated 30 min 55°C. DNA was extracted once with 350 μL Phenol:Chloroform:isoamyl Alcohol (25:24:1) in a phase-lock tube (Qiagen cat. no. 129046) and centrifuged 5 min at 16,000xg followed by extraction with 350 μL chloroform. The aqueous layer was transferred to a fresh tube containing 2 μL Glycogen (20 mg/ml Sigma cat. no. 10901393001). DNA was precipitated with three volumes of 100% ethanol and centrifuged at 16,000xg for 30 min at 4°C. The pellet was rinsed twice with 1 ml 80% ethanol, air dried and dissolved in 30 μL 10 mM Tris pH 8.0. Two μL was analyzed using either a HDS1000 or a genomic DNA ScreenTape on an Agilent 4200 TapeStation.

**Proteomics**

Viral and reconstituted protein extracts were resolved by 4-20% SDS-PAGE, and gels were silver-stained using the Pierce kit (Thermo cat. no. 24600). Protein bands were excised, destained, and proteolytically digested as described (Shevchenko et al., 1996). Proteolytic peptides were desalted and analyzed by LC-ESI-MS/MS using a ThermoScientific Orbitrap Fusion mass spectrometer. Mass spectra were analyzed with ThermoScientific Proteome Discoverer version 2.5 software using the Marseillevirus marseillevirus translated protein sequences from NCBI (ref. sequence NC_013756.1) appended to the cRAP (https://www.thegpm.org/crap/) common contaminant protein database. Peptide identification results were filtered to a false discovery rate of 1%. Mass spectrometry peptide data is available at https://data.mendeley.com/datasets/j9855z499b/1.

**Library preparation**

Sequencing libraries were prepared from DNA fragments using the KAPA HyperPrep Kit (KAPA cat. no. KK8504) following the manufacturer’s instructions. Libraries were amplified for eight cycles using a 10 sec 60°C combined annealing/extension step. Alternatively, a 72 C 1 min PCR extension step was added, with no apparent difference in the length distributions. To deplete total DNA samples of large fragments originating from insoluble chromatin prior to library preparation, samples were mixed with ½ volume of HighPrep™ PCR Clean-up System beads (MagBio Genomics cat. no. AC-60005), held 5–10 min, placed on a magnet stand, and the supernatant was collected. Based on the original sample volume, 0.8 volume of beads was added to the supernatant and held 5–10 min. Tubes were placed on a magnet stand, the supernatant discarded and the beads washed 1x with 80% ethanol prior to eluting in 15 μL 10 mM Tris-HCl pH 8. Library size distributions were resolved on an Agilent 4200 TapeStation. Barcoded libraries were achieved to equate chromosomal representation as desired aiming for a final concentration as recommended by the manufacturer for paired-end sequencing on either an Illumina HiSeq 2500 flow cell or a NextSeq flow cell.

**Expression and purification of marseillevirus histone tetramers**

A pET-24B plasmid containing Marseillevirus marseillevirus doublet histones (H1–H2 with a C-terminal 6xHis tag and H3–H4) (Valencia-Sánchez et al., 2021) was transformed into Escherichia coli BL2–codon plus(DE3)–RI L competent cells (Agilent Technologies) and cultured at 37°C in 2xYT-Kanamycin medium. After reaching optical density = 0.6, the cultures were induced with 0.5 mM IPTG at 37°C for 3 h and harvested by centrifugation. Cells were resuspended in 20 mM Tris pH 7.5, 2 M NaCl, 5 mM imidazole, 1 mM β-mercaptoethanol (βme) and lysed on a cell disruptor (AvestinEmulsiflexC3). The extract was clarified by centrifugation, applied to Ni-NTA agarose beads (Qiagen) and eluted with 20 mM Tris pH 7.5, 2 M NaCl, 300 mM imidazole, 1 mM βme. The protein sample was further
purified using a Superdex 200 26/600 size-exclusion chromatography column (GE Healthcare), and fractions were collected and concentrated to 1.8 mg/ml in 10 mM Tris pH 7.5, 2 M NaCl, 1 mM EDTA, 5 mM DTT.

For purification of Marseillevirus histone variant H2C-Hc, a C-terminal Strep tag II version of the histone was expressed similarly to canonical histones, and purified through Strept-Tactin XT (IBA-Lifescience) equilibrated with high salt Strept buffer (20 mM Tris pH 8, 2 M NaCl, 1 mM EDTA and 1 mM DTT), and eluted overnight with high salt strep buffer supplemented with 50 mM Biotin (IBA-Lifescience). The histone was further purified though SP-Sepharose (GE Healthcare) in 10 mM sodium acetate pH 5.2, 200 mM NaCl, 1 mM EDTA, 1 mM DTT, and eluted with a linear gradient of 0.200 to 2.0 M NaCl and through Ni-NTA agarose beads as above, and again purified through SP-Sepharose. Each histone was quantified by its absorbance at 280 nm. Histone tetramers were assembled by mixing equal molar amounts of each histone in high salt strep buffer.

**Purification of widom 601 DNA arrays**

A plasmid (3_187_widom_601) containing three tandem copies of the Widom 601 nucleosome positioning sequence with a 40-bp linker flanked by the EcoRV restriction site was transformed into DH5α plasmid (3_187_widom_601) containing three tandem copies of the Widom 601 nucleosome positioning sequence with a 40-bp linker flanked by the EcoRV restriction site was transformed into DH5α competent cells (ThermoFisher) and cultured in 2xYT-Ampicillin medium overnight. The 3_187_widom_601 DNA insert fragment was excised using EcoRV and purified using a previously published protocol (Dyer et al., 2004).

For reconstitutions using a Widom 601 array with 12-bp linkers, we used DNA from the 601-147-12 copy pUC19 plasmid (Dyer et al., 2004) kindly provided by Pamela Dyer and Karolin Luger and prepared as previously described (Furuyama et al., 2013). After digestion with BamHI and KpnI to separate the 601 array from the vector, we estimated by TapeStation capillary gel electrophoresis that the insert from our single-colony plasmid preparation contained 7 tandem copies of the 601 sequence.

For reconstitutions using a Widom 601 array without linkers, we used a custom synthetic construct purchased from Genscript, Inc.

**Reconstitution of nucleosomes on widom 601 arrays**

_Xenopus laevis_ and Marseillevirus nucleosome array reconstitutions were performed as described (Grau et al., 2021). Briefly, nucleosomes arrays were assembled by mixing purified 3_187_widom_601 DNA fragment and histone octamers, followed by overnight gradient salt dialysis in 10 mM Tris–HCl pH 7.5, 1 mM EDTA, 1 mM dithiothreitol (DTT) and 2 M to 0.25 M KCl using a peristaltic pump. Nucleosomes arrays were dialyzed into TCS-50 buffer (20 mM Tris-HCl pH 7.5, 50 mM NaCl, 1 mM EDTA, 1 mM DTT), concentrated, and stored at 4°C until use. Octamer/DNA ratios were optimized using small-scale reactions, and products were verified by native PAGE. The ratio used for Marseillevirus tetramer:DNA was 3.1, and for _X. laevis_ octamer:DNA was 3.9.

**Gradient fixation (GraFix) of nucleosome arrays**

Gradient fixation of nucleosome arrays was performed using the GraFix method (Stark, 2010). Nucleosome arrays were dialyzed into Buffer A (20 mM HEPES pH 7.9, 50 mM NaCl, 5% Glycerol, 1 mM DTT) for 3 hours at 4°C. Buffer B (20 mM HEPES pH 7.0, 50 mM NaCl, 10% Glycerol, 1 mM DTT) and Buffer C (20 mM HEPES pH 7.0, 50 mM NaCl, 40% Glycerol, 1 mM DTT, 0.1% Glutaraldehyde) were mixed using a gradient maker (Gradient Master, Biocom Instruments). Then, the nucleosome array was added to the top layer of the tube gradient and centrifuged at 50K rpm 4°C for 16 hours (Optima XL-100K, Beckman Coulter). Sedimentation products were fractionated and analyzed by 4% Native-PAGE (0.2X TBE). Selected fractions were quenched with 100 mM Tris pH 7.5 and dialyzed into 20 mM Tris pH 7.5, 50 mM NaCl, 1 mM DTT, and concentrated.

**QUANTIFICATION AND STATISTICAL ANALYSIS**

For each sample the following analysis steps were performed

1. Aligned Illumina fastq files to a reference genome with bowtie2 2.4.2 using parameters –end-to-end –very-sensitive –no-mixed –no-discordant –phred33 -I 10 -X 700
   a) T19: GCF_000887095.1_ViralProj43573_genomic.fna from NCBI
   Marseillevirus marseillevirus, strain T19 taxon 1559367
   b) G648: G648_26-12-14_Genome_vD.fasta: https://data.mendeley.com/datasets/zsh9nx8v5g/1
      ORFs: https://hgdownload.cse.ucsc.edu/goldenPath/dm6/database/refFlat.txt.gz
   c) Widom 601 3x Array: Genome_601x3.fasta: https://data.mendeley.com/datasets/ll7mg9c9wz/1
   d) _Drosophila melanogaster_: dm6 from UCSC. Chronological list of _D. melanogaster_
      ORFs: https://hgdownload.cse.ucsc.edu/goldenPath/dm6/database/refFlat.txt.gz
   (2) Extracted aligned fragments from bowtie2 sam files into bed files and divided into four fragment length subgroups: 1-, 1-100, 101-200, 201-300. Steps 3-8 were done for each of the four fragment lengths subgroups.
   (3) Computed percent GC for 10 base pairs on either end of the mapped fragments.
   (4) Made normalized count bedgraph files using bedtools 2.30.0 genomewc. Normalized counts are the fraction of counts at each base pair scaled by the size of the reference sequence so that if the scaled counts were uniformly distributed there would be 1 at each position.
   (5) Obtained annotations and selected ≥ 600 bps long
De novo genome assembly
Paired-end PE50 fragments were trimmed using cutadapt version 2.9 (Martin, 2010) with parameters: cutadapt -j 8 -m 20 –nextseq-trim 20 -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCA -A AGATCGGAAGAGCGGTGTAGGGAAAGAGTGT -Z. Trimmed paired-end fastq files were used as input for SPAdes (Bankevich et al., 2012) version 3.13.0 with parameters (-k = kmers): -k 21,27,33,35,37,41. Alignments to Marseillevirus T19 (GenBank NC_013756.1) and G648 reference sequences (Sahmi-Bounsiar et al., 2021) were performed using BLAT (kentinformatics.com).

DATA AND SOFTWARE AVAILABILITY
DNA sequencing datasets have been deposited in GEO under Accession Number GSE193224.