Immediate early proteins of herpes simplex virus transiently repress viral transcription before subsequent activation.

Laura E.M. Dunn[1,2], Claire H. Birkenheuer[1,2], Joel D. Baines[1,2*]

[1]Baker Institute for Animal Health, College of Veterinary Medicine, Cornell University, Ithaca, New York, USA

[2]Department of Pathobiology, School of Veterinary Medicine, Louisiana State University, Baton Rouge, Louisiana, USA

*Corresponding author

Email: jdb11@cornell.edu

Abstract

Herpes simplex virus 1 (HSV-1) utilizes the host-cell RNA polymerase II (Pol) to transcribe its genes in one of two phases. In the latent phase, viral transcription is highly restricted but during the productive lytic phase, more than 80 genes are expressed in a temporally coordinated cascade. In this study, we used precision nuclear Run On followed by deep Sequencing (PRO-Seq) to characterize early viral transcriptional events using HSV-1 immediate early (IE) gene mutants, corresponding genetically repaired viruses, and wild type virus. Unexpectedly, in the absence of the IE genes ICP4, ICP22 or ICP0 at 1.5 hpi we observed high levels of aberrant transcriptional activity across the mutant viral genomes, but substantially less on either wild type or the congenic repaired virus genomes. This feature was particularly prominent in the absence of ICP4 expression. Cycloheximide treatment

during infection with both the ICP4 mutant and repair led to increased Pol activity across both viral genomes. However, the repair retained some ability to repress activity on IE genes, thus indicating that both virion components and at least some *de novo* protein synthesis were required for full repression. Overall, these data reveal that prior to their role in transcriptional activation, IE gene products first function to repress transcription and that the HSV-1 lytic transcriptional cascade is mediated through subsequent de-repression steps.

Importance

Herpes simplex virus 1 (HSV-1) transcription in the productive phase is believed to comprise a series of activation steps leading to a specific sequence of gene expression. Here we show that immediate early (IE) gene products ICP0, ICP4 and ICP22 first repress viral gene transcription to varying degrees before subsequently activating specific gene subsets. It follows that the entire HSV transcriptional program involves a series of steps to sequentially reverse this repression. This previously uncharacterised repressive activity of IE genes very early in infection may represent an important checkpoint allowing HSV-1 to orchestrate either the robust lytic transcriptional cascade or the more restricted transcriptional program during latency.

Introduction

Herpes simplex virus 1 (HSV-1) expresses its double-stranded DNA genome using cellular RNA polymerase II (Pol) in two phases. During the latent phase in which the genome is maintained as an episome in sensory neurons, viral transcription is mostly limited to the production of the latency transcript (LAT). In the productive phase which usually occurs in epithelial cells, transcription of its more than 80 genes occurs in a cascade involving sequential activation of different gene subsets: from pre-immediate early, followed

sequentially by immediate early (IE), early (E), leaky late (LL), and true late (L) genes (1,2). To initiate productive replication, HSV commandeers Pol for viral transcription by introducing VP16 from the virion into the newly infected cell. VP16 binds a specific motif in viral IE promoters and recruits cellular transcription factors, HCF and Oct-1, to these sites to drive Pol initiation (3,4). Similar to cellular genes, viral gene expression is regulated through promoter proximal pausing (PPP) of Pol, followed by release into the gene body (5).

The IE genes include α0, α4 and α22/US1, encoding ICP0, ICP4 and ICP22, respectively, and are expressed by 1 hour post-infection (6). ICP0 and ICP4 are also virion components that are introduced into the cell upon the initiation of infection. ICP4 is both an essential activator of later gene sets (7) and a repressor of IE genes (8), while ICP0 is an E3 ubiquitin ligase that leads to degradation of a set of cellular proteins that would otherwise silence the viral genome early in infection (9). ICP22 negatively regulates Pol processivity and release to elongation through reduction of serine 2 phosphorylation on the Pol C-terminal domain (CTD) (10). This activity of ICP22 is important for enhancing promoter proximal pausing (PPP) on viral IE genes, and controlling anti-sense transcription on the compact HSV-1 genome (11).

The current studies were undertaken to further clarify the roles of ICP4, ICP0, and ICP22 in early viral transcription. We used Precision nuclear Run On followed by deep Sequencing (PRO-Seq) to identify sites of Pol activity to high resolution on the genomes of viruses bearing mutations in α0, α4 or α22, and corresponding viruses in which these loci were genetically restored.

Results

In the absence of ICP4, Pol activity is dysregulated throughout HSV-1 infection. We first used the ICP4 truncation mutant n12, derived from the KOS HSV-1 strain (1) to investigate the role of ICP4 during transcription. To ensure the control virus had a genetically identical background to the mutant, a repair virus was generated from n12 by recombination in infected cells with wild type α4 DNA. Expression of full length ICP4 was confirmed by immunoblotting (Fig. S1A). HEp-2 cells were then infected at a multiplicity of infection (MOI) of 5 with either n12 or repaired virus, nuclei harvested at 1.5, 3 or 6 hours post infection (hpi), and PRO-Seq performed.

Visualization of the data on a genome browser revealed substantial differences in the pattern of Pol activity on the n12 genome compared to the repair genome (Fig. 1A, first panel). Consistent with the strict temporal expression of herpesvirus genes, Pol activity on the repair virus genome was restricted to IE genes at 1.5 hpi. In contrast, strong Pol activity was distributed across the entire n12 genome. At 3 hpi, the distribution of Pol activity on the two viral genomes was more similar, but activity on the IE genes UL54 (α27), α4, α0, US1 (α22), US12 and L/ST (Fig. 1A, second panel) and many late genes remained higher in n12 than the repair. By 6 hpi, intense Pol activity of n12 was limited almost entirely to IE genes (and UL39), while activity on the repair virus genome was more evenly distributed among genes of all temporal classes (Fig. 1A, third panel).

Quantification of sequencing reads aligned to individual HSV-1 genes confirmed a significant increase in Pol activity in n12 over the repair virus at 1.5 and 3 hpi (Fig. 1B). At 6 hpi the total viral reads remained higher in n12, but this was due solely to intense IE gene activity (Fig. 1C) that did not occur in repair virus. Over time, activity on later gene classes of n12 (Fig. 1D, E, F) decreased, while in repair infection Pol activity increased on all gene

classes, except for IE genes between 3-6 hpi (Fig. 1C). Overall, these data indicate that ICP4 decreases Pol activity on all viral genes immediately after infection, and increases E, LL and L activity later in infection.

PRO-Seq profiles of Pol occupancy on most HSV-1 genes follow that of cellular genes, with peaks at the 5' PPP site, and at the 3' cleavage/Poly(A) site (5). The promoters of the IE genes α4 and US1 showed clear PPP at all time points on the repair genome (Fig. 2A). High read counts on n12 IE genes obscured patterns of Pol activity on the same scale so the browser was adjusted to a comparable read count depth between the two viruses. This revealed strong PPP on α4 and US1 throughout infection in n12. Other transcriptionally active regions in n12 lacked apparent PPP at 1.5 hpi including E genes UL29 and UL30 (Fig. 2B), the LL genes UL18, UL19 and UL20 and the L genes UL20.5 and UL22 (Fig. 2C). The presence of a promoter peak was also assessed by examining the relative read density over a gene and using bootstrap confidence of fit to indicate statistical differences. To remove any uncertainty of whether a read was from a promoter region or gene body, only genes with no overlapping transcripts and with a defined TSS were included in the analysis. We also separately analysed genes that were robustly transcribed in repair infection at each time point as it was assumed that these gene were transcribed in a temporally correct fashion. This set of appropriate positively regulated genes was termed "repair-active" and comprised only the IE genes at 1.5 hours, with the inclusion of more genes over the course of infection. Genes outside of this set (i.e., that were transcribed robustly in n12 but not repair at a given time point) comprised a "repair-repressed" subset.

Despite a large difference in total read counts on viral genes at 1.5 hpi, the distributions of reads across entire repair-active genes (i.e., relative read density) of n12 and repair were

similar (Fig. 2D, top panel). Analysis of the promoter regions of active genes showed strong PPP on both viruses but with significantly higher PPP in n12 than in repair. Read density and PPP peaks were significantly lower in the repair-repressed gene set compared to repair-active genes. At 3 hpi (Fig. 2D, middle panel) a similar pattern was seen, with a small but discernible promoter peak on n12 repair-repressed genes. When all genes in repair were included in the repair-active subset due to robust transcription at 6 hpi, both viruses had a similar pattern of read density across individual genes. Pause indices confirmed this trend and revealed that the patterns of pausing changed similarly throughout infection (Fig. S2A-C). These data indicate that the ICP4-mediated repression of most genes at IE times is due to a reduction in total Pol occupancy across entire genes rather than an increase in Pol pausing.

PRO-Seq visualisation indicated extensive Pol activity on the antisense strands of genes in n12 at 1.5 hpi. To examine the regulation of sense-to-antisense transcription, the proportion of sense reads was calculated on isolated genes that lacked genes on the opposite strand. At 1.5 hpi there was no significant difference between n12 and repair in the overall proportion of sense transcription (Fig. 2E). Although a low level of transcription was detected across the repair viral genome at 1.5 hpi, robust transcription was limited to the IE genes and preliminary analysis indicated that the most transcriptionally active genes had the highest proportion of sense transcription (Fig. S3). Therefore, we analysed repair-active and repair-repressed subsets separately. On repair-active genes there was no significant difference in the levels of sense transcription between the two viruses, and both had a high proportion of sense reads (Fig. 2F). However, the repair-repressed genes had a significant increase in antisense transcription at 1.5 and 3 hpi in n12. While n12 did have a significantly reduced global level of sense transcription at 6 hpi, this was largely due to high levels of antisense transcription on US8, US8A, US9, UL41 and UL56. Closer analysis of these regions revealed the presence of

an upstream ICP4-independent promoter (US12, UL54, US1 and UL39) potentially driving read-through antisense Pol activity (Fig. S4A-D). This activity was also visible on the repair genome, indicating that it is a feature of HSV-1 transcription. The IE α4 and α0 genes, with promoters in the repeat regions, showed less evidence of extensive read-through (Fig. S4E-F). These data indicate that much of the aberrant transcription observed in the absence of ICP4 involves high levels of antisense transcription on viral genes, except for the IE genes.

To determine whether detected intergenic Pol activity could be a result of read-through of a transcriptional termination signal (TTS) (AAUAA), a downstream-read index was calculated by dividing the reads per bp 150 bp downstream of a TTS, by the reads per bp in the upstream ORF. For reads to be accurately assigned, only singular unnested genes with a defined TTS more than 150bp away from a transcription start site (TSS) were included. At 1.5 hpi, there was no significant difference in the downstream index on repair-active genes between n12 and repair (Fig. 2G). However, there was a significantly higher proportion of downstream transcription on repair-repressed genes at 1.5 hpi. By 3 and 6 hpi, there were no significant differences in intergenic transcription on the gene sets between viruses, suggesting proper termination was restored at these time points even in the absence of ICP4.

Overall, the PRO-Seq data at 3 and 6 hpi was consistent with ICP4 mutant phenotypes, primarily consisting of a failure to repress IE genes and to promote transcription of later gene temporal classes (12,13). However, the finding that ICP4 was required to repress Pol activity on virtually all viral genes at 1.5 hpi was unexpected. Pol activity on n12 IE genes was aberrantly high at 1.5 hpi, suggesting accelerated initiation. However, these genes displayed low levels of antisense and downstream transcription, and strong PPP peaks, suggesting pausing and termination approached normal levels. In contrast, Pol activity on genes of later

kinetic classes was indiscriminate and included antisense and intergenic transcription. Restoration of features consistent with proper termination and PPP returned in n12 by 3 and 6 hpi indicating that proteins other than ICP4 are utilised to improve transcriptional control after 1.5 hpi.

Similar results as above were obtained in PRO-Seq experiments performed at 1.5 hpi of primary human foreskin fibroblasts (HFF) showing that ICP4-mediated repression of viral transcription occurs in multiple cell types (Fig. S5A-F). In addition, genomic DNA extracted from the infected HFF nuclei in this experiment was quantified by qPCR. The results indicated that differences in genome copy number between the viruses could not account for increased numbers of PRO-Seq reads on the mutant genome (Fig. S5H).

Virion-associated and *de novo* expressed ICP4 are required for full transcriptional repression at 1.5 hpi. We next asked whether the role of ICP4 to repress Pol early in infection required *de novo* protein synthesis or could be mediated by ICP4 entering the cell as a virion component. It should be noted that n12 virions do incorporate some ICP4 from the E5 cell line used to support n12 growth (17). HEp-2 cells were infected in the presence or absence of the protein synthesis inhibitor cycloheximide (CHX), followed by PRO-Seq at 1.5 hpi. Overall, CHX treatment led to a significant increase in normalised viral reads aligning to both the repair virus genome and the n12 genome (Fig. 3A), indicating that proteins expressed *de novo* are required to fully repress Pol early in infection.

DeSeq2 (14) was used to identify specific genes affected by CHX treatment. CHX led to a general upregulation of Pol activity on all genes of the repair virus, and 21 genes reached a statistically significant increase (Fig. 3B). LAT and L/ST had the lowest increase of all genes

after CHX treatment, indicating that virion components were able to repress most activity on these genes. Previous reports have revealed ICP4 to be a strong repressor of L/ST (15,16) and therefore suggest ICP4 as the strongest candidate responsible for this repression. In n12 infected cells, CHX treatment led to a strong increase on IE genes relative to other classes (Fig. 3C). Interestingly, many of the E and L genes of n12 that had significant increases in Pol activity are in genomic regions close to or nested with IE genes; (e.g., US10/11 nested with US12, UL55 immediately downstream of UL54 and UL39 has an ICP4-independent promoter). These data support the conclusion that virion-associated ICP4 is primarily involved in repression of L/ST and other IE genes. This was visible in the PRO-Seq data tracks as CHX treatment of repair yielded no visible increase in activity on L/ST and only negligible increases on the IE genes, α0, α4, and US1 (Fig. 3D). In contrast, CHX treatment of n12 led to extensive increases of Pol activity across all IE genes.

The PRO-Seq visualisation also revealed strong PPP on IE genes in both n12 and repair (Fig. 3E). To assess whether the CHX-induced increase in reads of IE genes was a result of changes in Pol pausing, the relative read density across these genes was calculated. This indicated that in the presence of CHX, both n12 and repair had a significant reduction in PPP suggesting that the increase in reads was at least partially due to increased release of Pol into the gene body (Fig. 3E). These data further suggest that proteins other than ICP4 must be synthesized to regulate early transcription, and particularly to regulate PPP and release to elongation.

Viral IE genes ICP0 and ICP22 are also required for early Pol repression. To investigate possible roles of other IE genes on early Pol repression we used HSV IE mutants n212 containing a nonsense mutation in α0 derived from the KOS HSV-1 strain (18), and a

224 ΔICP22 virus lacking the entire US1 coding region along with its genetically restored repair,
225 derived from the HSV-1(F) strain (19,20). To ensure the ICP0 control virus had a genetically
226 identical background to the mutant, a repair virus was generated by recombination with α0
227 DNA, and immunoblotting was used to confirm repair of the ICP0 locus (Fig. S1B). HEp-2
228 cells were infected with the mutants, their respective repair viruses, or wild type HSV-1(F)
229 virus and PRO-Seq performed on nuclei harvested at 1.5 hpi.
230
231 As with the n12 ICP4 mutant, both n212 and ΔICP22 bore increased Pol activity across their
232 genomes compared to their genetically repaired counterparts at 1.5 hpi (Fig. 4) and this was
233 statistically significant for all viruses (Fig. 5A). However, each mutant displayed different
234 PRO-Seq patterns. As seen in PRO-Seq profiles of repair and WT viruses, E genes UL29 and
235 UL30 are not normally active at 1.5 hpi, (Fig. 5B). Although all mutant viruses showed
236 increased activity on these genes over wild type viruses, the PRO-Seq pattern on the n212
237 genome included features indicative of proper regulation including prominent PPP peaks and
238 minimal antisense/intergenic activity. In contrast, ΔICP22 had a PRO-Seq pattern more like
239 that of the n12 ICP4 mutant, covering the entire genome on both strands (Fig. 4). Closer
240 examination at the α4/US1 region showed extensive antisense and intergenic transcription on
241 the ΔICP22 genome (Fig. 5C), and PRO-Seq patterns consistent with regulatory features were
242 obscured in the ΔICP22 PRO-Seq profile. ΔICP22 had a significant increase in Pol activity
243 relative to its parent wild type strain HSV-1(F) virus, indicating the presence of a secondary
244 mutation in the repair that affects repression. Nevertheless, ICP22 clearly played an important
245 role in repression as activity was substantially lower in the ΔICP22 repair virus than in the
246 congenic ΔICP22 mutant.
247

Analysis of sense/antisense transcription revealed that n212 had almost identical levels of sense transcription compared to repair virus on repair-active genes (Fig. 5D). There was a reduced proportion of sense transcription on the repair-repressed gene set but the proportion of sense transcription was higher than either n12 or ΔICP22. In addition, antisense transcription occurred heavily on US8, US8A, US9 and UL41, which bore high levels of antisense transcription in n12 and its repair (at 3 and 6 hpi) (Fig. 2F), further indicating that the antisense activity on these genes is a consistent feature of active HSV-1 transcription. Relative to its repair, ΔICP22 had significantly lower levels of sense transcription on repair-active genes and was not significantly different compared to repair-repressed genes. Thus, ΔICP22 lacked regulation of sense-to-antisense transcription across the entire genome. Downstream-read index calculations also indicated a deficiency in transcriptional termination of the ΔICP22 virus, with significantly increased levels of reads downstream of the TTS on the repair-repressed genes (Fig. 5E). In contrast, the downstream-read index was not significantly different between repair-active and repair-repressed genes on the n212 genome indicating normal transcriptional termination.

Promoter peaks for the repair-active genes in n212 were unchanged relative to repair (Fig. 5F). However, ΔICP22 had a significant reduction in the promoter peak on these genes, matching its previously noted phenotype at 3 hpi (11). Analysis of the relative read density across the repair-repressed genes indicated that n212 maintained strong PPP on these genes, similar to n12 (Fig 5G). Calculation of gene pause indices indicated that all repair-repressed genes in n212 had an equivalent level of pausing. However, in n12 there was a large amount of variation with notably strong pauses on specific genes: UL48, UL10 and UL42, (Fig. S2A, D). In contrast, ΔICP22 lacked any evidence of pausing on repair-repressed genes.

While n212 and ΔICP22 genomes bore significant overall increases of Pol activity on all gene classes relative to their repair viruses except for IE genes with ΔICP22 (Fig. 6A, B), DeSeq2 analysis of individual genes indicated that n212 displayed the strongest increases in activity on specific E genes. There was no apparent trend for upregulation of a specific gene class over others in n12 and ΔICP22 (Fig. 6C-E). One interpretation of these data is that in the absence of ICP0, transcription progresses more rapidly through the temporal transcriptional program. PRO-Seq data using n212 at 3 and 6 hpi was consistent with this hypothesis, as the distribution of Pol across the genome in n212 at 1.5 hpi matched that of the repair at 3 hpi (Fig. 6F). By 6 hpi, n212 and its repair had virtually matching PRO-Seq profiles, although n212 continued to have higher read counts (Fig. 6G). This differed from ΔICP22 which had reduced Pol activity on IE genes at 3 hpi and loss of Pol activity from most genes at 6 hpi (11).

Discussion

We have shown that ICP0, ICP4 and ICP22 are individually required for early transcriptional repression on the HSV-1 genome that precedes the activation of the temporal transcriptional cascade. The role of each in transcriptional repression is likely distinct because the PRO-Seq patterns of each mutant differed. The ICP4 mutant, n12, displayed increased Pol activity on all regions of the genome but retained features of proper transcriptional regulation on the IE genes. The ΔICP22 mutant genome also exhibited high levels of Pol activity but with a lack of PPP and extensive anti-sense transcription across the whole genome. The ICP0 mutant n212 led to increased Pol preferentially on E genes and the pattern of transcriptional activity on these genes was equivalent to those of wild type viruses, albeit with a faster progression through the transcriptional cascade. Overall, it is apparent that prior to initiating activation of

297 transcription, IE genes first mediate transcriptional repression. We have termed this novel
298 repression **T**ransient **I**mmediate **E**arly **G**ene **R**epression or **TIEGR**.
299
300 TIEGR is a previously unrecognized process preceding steps outlined in the current paradigm
301 of HSV transcription which proposes that IE gene products function primarily to promote
302 viral gene expression (21). These data suggest that the current paradigm be modified to
303 consider the entire cascade of viral gene expression as a series of de-repression steps on
304 different genes.
305
306 High Pol activity might be expected early in infection because HSV-1 genomes enter the
307 nucleus free of nucleosomes and contain densely packed, mostly intronless genes encoded on
308 both strands. This gene arrangement necessitates a high concentration of promoter elements
309 such as TATA boxes and initiator sequences recognised by cellular Pol (22,23), and multiple
310 GC-rich regions that can bind cellular transcription factors including Sp1 and Egr-1 (24,25).
311 In addition, VP16 co-introduced with the genome is a powerful transcription factor with
312 potent DNA binding activity for viral IE promoters, and a transcriptional activation domain
313 (TAD) that interacts with many transcription factors such as TFIIA, TFIIB, TBP and TFIIH
314 (26). The effectiveness of these elements is supported by the observation that by 3 hours after
315 infection, the viral genome bears 1/3 of all Pol activity of the cell (27).
316
317 It is likely that the newly introduced HSV-1 genome is rapidly engaged by not only virion-
318 delivered VP16, ICP0 and ICP4, but also by cellular histones. Previous studies have shown
319 that histones associate with the viral genome during lytic infection (reviewed in (28)) and that
320 histone dynamics are important in the regulation of lytic transcription (29). Other studies
321 emphasize the ability of ICP4 to coat the viral genome in a form of "viral chromatin" (30).

ICP4 in this the viral chromatin was proposed to facilitate recruitment of Pol and components of the pre-initiation complex. However, because Pol is efficiently recruited to the viral genome in the absence of full-length ICP4 (and ICP0 and ICP22) we propose that a major function of ICP4 and virally orchestrated chromatin is to transiently repress transcription.

This repressive role of ICP4 is consistent with ICP4's known ability to repress IE gene transcription and was supported by our 6 hpi PRO-Seq data, in which n12 was unable to repress IE gene Pol activity. ICP4 repression is associated with ICP4-binding sites (15,16,31), that have a relatively loose consensus sequence with more than 100 copies throughout the viral genome (30). n12 expresses a truncated form of ICP4 (Fig. S1A), containing most of the N-terminal activation domain but without the DNA-binding domain (32) , indicating that ICP4 DNA binding is required for repression. How ICP4 can switch from a repressor to an activator is unknown. The repressive function likely occurs through forming a complex with TFIIB and TBP (33). Activation could be linked to modifications of these proteins; for example, both ICP4 (34) and TFIIB (35) are phosphorylated. Alternatively, the switch to activation could be linked to recruitment of other HSV-1or cellular proteins such as TAF1/TFIID, which the ICP4 C-terminal domain is known to help recruit (36).

The result with ΔICP22 likely reflects ICP22's known role in reducing processivity of RNA Polymerase (11), yet it was surprising that this occurred over the entire viral genome and was not limited to IE genes. One possible explanation is that transcription is initiated at VP16 sites, but that the increased Pol processivity in the absence of ICP22 leads to substantial read through into neighbouring genes (37). ICP22 has been shown to interact with several elongation regulatory factors including FACT (38), p-TEFb (39) and CDK9 (40) and it has

346 been proposed that these interactions leads to selective repression of cellular genes (41). A
347 similar mechanism might account for repression of the HSV-1 genome early in infection.
348
349 The increased transcriptional activity in the ICP0- mutant was surprising as ICP0 is mostly
350 viewed as a promiscuous transactivator that acts by causing degradation of cellular
351 transcriptional silencing factors (42). Unlike the ICP4 and ICP22 mutants, the pattern of Pol
352 activity on the ICP0 mutant genome reflected normal regulatory features such as PPP and
353 termination. It is known that ICP0 mutants infected at high MOI grow equivalently to wild-
354 type (42) virus, and that these mutants produce a large number of defective particles at very
355 early stages of infection (18). One possibility is that the large number of defective particles
356 deliver more abundant virion components such as the incoming VP16/ICP4 to help activate
357 the n212 genome. Our CHX experiment also highlighted the importance of virion
358 components in early transcriptional repression. In addition, ICP0 and ICP4 physically interact
359 (43), suggesting a potential role for ICP0 during the switch in viral chromatin from repressive
360 to activating.
361
362 As each IE protein affects production of the others, we cannot currently determine whether
363 the observed repression was due to the specific loss of one, or a combinatorial effect and the
364 full mechanism of this repression remains to be elucidated. TIEGR may represent an
365 important checkpoint for a virus that can orchestrate one of two potential transcriptional
366 programs: one that is robust during productive replication in which TIEGR is reversed, or one
367 during latency in which TIEGR, or a version of it, is maintained.
368

369 **Materials and Methods**

370 **Cells**

HEp-2 (human epithelial lung cancer), Vero (African green monkey), RSC (rabbit skin) and the Vero derived ICP4-complementing, E5, cells (44) were maintained in Dulbecco's modified Eagle's medium (DMEM) containing 10% new born calf serum (NBS), 100 units/ml penicillin, 100µg/ml streptomycin (pen/strep) and maintained at 37□C with 5% CO_2. U2OS (human bone osteosarcoma) cells were maintained in McCoy's 5A medium containing foetal bovine serum (FBS), pen/strep and maintained at 37□C with 5% CO_2. HFF (human foreskin fibroblasts) were maintained in DMEM containing 15% FBS, pen/strep and at 37□C with 5% CO_2. S2 (*Drosophila melanogaster*) cells were grown in Schneider's medium containing 10% FBS and maintained at 23°C. E5 cells were a gift from Dr. Neal DeLuca, University of Pittsburgh and HFF cells were a gift from Dr. Luis Schang, Cornell University.

Viruses

Mutant HSV-1 strains n12 (44) and n212 (18), both HSV-1 KOS derived, were gifts from Dr. Neal DeLuca, University of Pittsburgh. n12 virus stocks were prepared on E5 cells and n212 were prepared on U2OS cells. The ΔICP22 virus and its repair, generated from a HSV-1 (F) bacterial artificial chromosome, were gifts from Dr. Yasushi Kawaguchi (19,20). Stocks of both were prepared on Vero cells.

Repair viruses of n12 (α4 mutant) and n212 (α0 mutant) were generated through homologous recombination. Plasmids containing WT HSV-1 DNA spanning 1000 bp either side of the mutation were synthesized, linearized, then transfected using TransIT-X2 (Mirus Bio) into RSC cells. 6 hours after transfection, cells were infected at a MOI of 1 with the corresponding mutant and when all cells showed CPE, harvested by scraping into media. Viruses were released from cells through 3x cycles of freeze-thaw in LN_2 and 37°C water

bath. Samples were then serially diluted onto Vero cells with an agarose overlay and underwent 3 rounds of plaque purification. Repair virus plaques were selected by size due to the growth restriction of the mutants on Vero cells. Western blotting was used to confirm successful repair of full-length protein expression (Fig. S1). Stocks were subsequently prepared on Vero cells.

Virus infections and drug treatments

For assessment of IE protein expression, monolayers of $1.2x10^6$ HEp-2 cells were infected with HSV-1 at a MOI of 5 in 199V medium (+1% NBS). After 1h, inoculum was replaced with DMEM (+2% NBS).

For PRO-Seq, monolayers of $2x10^7$ cells were infected with HSV-1 at a MOI of 5 as above. Infections were allowed to continue for 1.5, 3 or 6 hpi – hpi refers to the time after the viral inoculum was first added to cells. For CHX treatments, 10μm of CHX was added to media 1h prior to infection and this CHX concentration was maintained in the media throughout infection.

Western blotting analysis

Lysates were prepared from infected cells by washing cells x2 in ice-cold PBS before addition of an appropriate volume of NP-40 lysis buffer (150mM NaCl, 1% NP-40, 50mM Tris pH8, protease inhibitors). Cells were then scraped into suspension and incubated on ice for 30 min. 20μg of protein was mixed 2X Protein Sample Loading Buffer (LiCor) and heated to 98□C for 5 min. Samples were separated on an 8% polyacrylamide resolving gel, layered with a 5% stacking gel. Proteins were transferred to a 0.45 μm nitrocellulose membrane using a wet transfer technique. Membranes were first blocked for 1 h at room

421 temperature in 5% BSA +0.1% Tween®20, then incubated with the primary antibodies
422 diluted in blocking buffer at 4□C overnight. Membranes were washed 4 x for 5 min in PBS
423 containing 0.1% Tween before being incubated with the secondary antibodies diluted in
424 blocking buffer at room temperature for 45 min. Membranes underwent a further 4x 5 min
425 washes in PBS-T and were then visualized on an Odyssey Scanner (LiCor). Primary
426 antibodies were anti-ICP4 (sc-69809, Santa Cruz), used at 1:200 and anti-ICP0 (sc-53070,
427 Santa Cruz), used at 1:100. Secondary antibodies were DyLight 680 and 800 (Cell
428 Signalling).
429
430 **Nuclei isolation and PRO-Seq**
431 Nuclei were isolated from infected cells, nuclear run-on with biotinylated nucleotides
432 performed and sequencing libraries generated as described previously (5,11,27,45). Nuclei
433 from Drosophila S2 cells were spiked into infected-cell nuclei at a ratio of 1:1000 prior to
434 run-on. Libraries were sequenced on an Illumina NextSeq 500, performed by the GeneLab at
435 Louisiana State University School of Veterinary Medicine or by the Biotechnology Resource
436 center (BRC) Genomics facility *(RRID:SCR_021727)* at the Cornell Institute of
437 Biotechnology.
438
439 **DNA isolation**
440 The lower interphase and organic layers of the initial TRIzol LS RNA extraction from the
441 run-on was saved. The DNA was subsequently isolated following the manufacturer's protocol
442 for TRIzol DNA isolation (Invitrogen, MAN0016385). To clean-up the isolated DNA, 3x
443 phenol:chloroform:isoamyl alcohol purification was performed before ethanol precipitation
444 using sodium acetate. DNA was resuspended in DEPC water.
445

446 **Quantitative PCR for viral genome copy**

447 Extracted DNA was quantified using NanoDrop spectrophotometer (Thermo Scientific).
448 HSV-1 genome copy number was determined in 50ng of extracted DNA as previously
449 described (5,27). Brilliant III SYBR® Green Master Mix with ROX (Agilent) was used and
450 qPCR performed on QuantStudio™ 3 (Applied Biosystems).
451
452 **Read processing**
453 FastQ files were processed using the PRO-Seq pipeline developed by the Danko lab (Cornell)
454 https://github.com/Danko-Lab/utils/tree/master/proseq. Reads were aligned to a concatenated
455 genome file containing hg38, dm3, HSV-1 genomes. HSV-1 genome builds had the external
456 repeats deleted to aid sequencing alignment; the modified genome files are available:
457 https://github.com/Baines-Lab/Public/tree/main/HSV-1. Seqmonk software (46) was used to
458 probe reads from the individual genomes in the output .bam files. Drosophila spike-in
459 normalisation was used to account for variation in sequencing depth between libraries.
460 Libraries were normalised relative to the library with the largest drosophila read count, with
461 the scale factor for this set to 1. The output .bw files (containing only the 3' final read
462 position) were visualized using the IGV genome browser (47). Individual reads for each bp
463 across the HSV-1 genome was extracted from the .bw files using multiBigWigSummary from
464 deepTools (48) and normalized as above. The resulting .txt files were converted to .bed for
465 analysis in Seqmonk or .bedgraph for visualization in IGV. The normalized .bed/.bedgraph
466 files were used for all subsequent data analysis. Principal component analysis of normalized
467 HSV-1 gene reads from each sequencing experiment was used to confirm clustering of
468 replicates (Fig. S6).
469
470 **Identification of transcriptionally active genes**

The robust transcription threshold to identify repair-active genes was determined on data from repair-virus infection at 1.5 hpi due to restricted transcriptional activity on these samples. Genes were classified to be transcriptionally active (repair-active) if the mean read per bp (TSS -TTS) was greater than 1 standard deviation of the mean of all genes. Genes below this threshold in repair infections, but above this threshold in mutant infections were classified as repair-repressed. The same threshold was used at all time points and was calculated individually for each sequencing experiment due to variation in sequencing depth. Details of normalized HSV-1 gene reads, reads per bp and robust transcription thresholds are given in Appendix 1.

Promoter proximal pause analysis

The relative distribution of reads across gene sets was determined using the Seqmonk probe trend plot on .bed files. Genes were divided into 100 bins to get relative distance across gene. Data was plotted using the smoother function in SAS JMP Pro software and the bootstrap confidence of fit calculated using 300 iterations of the data. Pause index calculations for individual genes was determined by the formula: mean read per bp (TSS +150bp)/mean read per bp rest of gene (to TTS). Only viral genes with no overlapping transcripts and with a defined TSS were included in PPP analysis to allow reads to be accurately assigned.

Sense transcription proportion calculation

The proportion of sense transcription on each HSV-1 gene was calculated by formula: reads sense strand of gene (TSS-TTS)/reads antisense strand of gene (TSS-TTS). Only isolate HSV-1 genes with no gene on opposite strand were included in final analysis to allow reads to be accurately assigned. Regression analysis was used to assess correlation between sense

proportion and read per bp (TSS-TTS) of each HSV-1 gene. The square of the correlation coefficient (R^2) was calculated using SAS JMP Pro software.

Downstream-read index calculation

The level of reads downstream of TTS was calculated by formula: reads per bp 150bp downstream of TTS/reads per bp in upstream ORF (TSS-TTS). Only singular unnested genes with a defined TTS more than 150bp away from any other TSS were included to allow reads to be accurately assigned.

Fold change analysis

Fold change analysis was performed using the R package DeSeq2 (14). DeSeq2 for HSV-1 genes was performed incorporating Hg38 genes data to account for the library size correction step. Full DeSeq2 fold change and p-values are given in Appendix 2.

Statistical analysis

Statistical significance was determined by unpaired Student's t-test for pairwise comparisons of normally distributed data, Mann-Whitney U test was used for pairwise comparison of nonparametric data and Kruskal-Wallis for multiple comparisons. Calculations were performed using SAS JMP Pro software. Each independent experiment consisted of 2-3 biological replicates.

Data availability

The data will be publicly available upon publication on the GEO database under the accession number GSE202363.

Acknowledgements

We thank Dr. Neal DeLuca, Dr. Luis Schang and Dr. Yasushi Kawaguchi for their kind gifts of cells/viruses. We thank Thaya Stoufflet and Dr. Vladimir Chouljenko at the LSU SVM GeneLab and the Genomics Facility (*RRID:SCR_021727*) of the Biotechnology Resource Center of Cornell Institute of Biotechnology for their help with sequencing experiments. Portions of this research were conducted with the high-performance SuperMic supercomputer provided by Louisiana State University (http://www.hpc.lsu.edu) and we thank Dr Le Yan for his help in maintaining the PRO-Seq pipeline. These studies were supported by National Institutes of Health grants R01 AI 141968 and R21 AI 148926 to J.D.B.

Author contributions

J.D.B conceived the study, L.E.M.D performed experiments involving n12 and n212 HSV-1 viruses and C.H.B performed experiments involving ΔICP22 HSV-1. L.E.M.D performed data analysis and made the figures. All authors discussed results L.E.M.D and J.D.B wrote and edited the manuscript.

References

1. Honess RW, Roizman B. Regulation of herpesvirus macromolecular synthesis. I. Cascade regulation of the synthesis of three groups of viral proteins. J Virol. 1974;14(1):8–19.

2. Ibáñez FJ, Farías MA, Gonzalez-Troncoso MP, Corrales N, Duarte LF, Retamal-Díaz A, et al. Experimental Dissection of the Lytic Replication Cycles of Herpes Simplex Viruses in vitro. Front Microbiol. 2018;9:2406–2406.

543 3. Johnson KM, Mahajan SS, Wilson AC. Herpes simplex virus transactivator VP16
544 discriminates between HCF-1 and a novel family member, HCF-2. J Virol. 1999
545 May;73(5):3930–40.

546 4. O'Hare P, Goding CR. Herpes simplex virus regulatory elements and the
547 immunoglobulin octamer domain bind a common factor and are both targets for virion
548 transactivation. Cell. 1988;52(3):435–45.

549 5. Birkenheuer CH, Baines JD. RNA Polymerase II Promoter-Proximal Pausing and
550 Release to Elongation Are Key Steps Regulating Herpes Simplex Virus 1 Transcription. J
551 Virol. 2020;94(5):e02035-19.

552 6. Harkness JM, Kader M, DeLuca NA. Transcription of the herpes simplex virus 1 genome
553 during productive and quiescent infection of neuronal and nonneuronal cells. J Virol.
554 2014;88(12):6847–61.

555 7. DeLuca NA, McCarthy AM, Schaffer PA. Isolation and characterization of deletion
556 mutants of herpes simplex virus type 1 in the gene encoding immediate-early regulatory
557 protein ICP4. J Virol. 1985;56(2):558–70.

558 8. Rivera-Gonzalez R, Imbalzano AN, Gu B, Deluca NA. The Role of ICP4 Repressor
559 Activity in Temporal Expression of the IE-3 and Latency-Associated Transcript
560 Promoters during HSV-1 Infection. Virology. 1994 Aug 1;202(2):550–64.

561 9. Gu Haidong, Roizman Bernard. Herpes simplex virus-infected cell protein 0 blocks the
562 silencing of viral DNA by dissociating histone deacetylases from the CoREST–REST
563 complex. Proc Natl Acad Sci. 2007 Oct 23;104(43):17134–9.

564 10. Fraser KA, Rice SA. Herpes simplex virus type 1 infection leads to loss of serine-2
565 phosphorylation on the carboxyl-terminal domain of RNA polymerase II. J Virol. 2005
566 Sep;79(17):11323–34.

567 11. Birkenheuer Claire H., Dunn Laura, Dufour Rachel, Baines Joel D., Sandri-Goldin
568 Rozanne M. ICP22 of Herpes Simplex Virus 1 Decreases RNA Polymerase Processivity.
569 J Virol. 96(5):e02191-21.

570 12. Dixon RA, Schaffer PA. Fine-structure mapping and functional analysis of temperature-
571 sensitive mutants in the gene encoding the herpes simplex virus type 1 immediate early
572 protein VP175. J Virol. 1980 Oct;36(1):189–203.

573 13. Watson RJ, Clements JB. A herpes simplex virus type 1 function continuously required
574 for early and late virus RNA synthesis. Nature. 1980 May 1;285(5763):329–30.

575 14. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for
576 RNA-seq data with DESeq2. Genome Biol. 2014 Dec 5;15(12):550.

577 15. Lee LY, Schaffer PA. A virus with a mutation in the ICP4-binding site in the L/ST
578 promoter of herpes simplex virus type 1, but not a virus with a mutation in open reading
579 frame P, exhibits cell-type-specific expression of gamma(1)34.5 transcripts and latency-
580 associated transcripts. J Virol. 1998 May;72(5):4250–64.

581 16. Yeh L, Schaffer PA. A novel class of transcripts expressed with late kinetics in the
582 absence of ICP4 spans the junction between the long and short segments of the herpes
583 simplex virus type 1 genome. J Virol. 1993 Dec;67(12):7373–82.

584 17. Dembowski JA, DeLuca NA. Temporal Viral Genome-Protein Interactions Define
585 Distinct Stages of Productive Herpesviral Infection. mBio. 2018;9(4):e01182-18.

18. Cai W, Astor TL, Liptak LM, Cho C, Coen DM, Schaffer PA. The herpes simplex virus type 1 regulatory protein ICP0 enhances virus replication during acute infection and reactivation from latency. J Virol. 1993 Dec;67(12):7501–12.

19. Maruzuru Yuhei, Fujii Hikaru, Oyama Masaaki, Kozuka-Hata Hiroko, Kato Akihisa, Kawaguchi Yasushi. Roles of p53 in Herpes Simplex Virus 1 Replication. J Virol. 2013 Aug 15;87(16):9323–32.

20. Maruzuru Y, Shindo K, Liu Z, Oyama M, Kozuka-Hata H, Arii J, et al. Role of herpes simplex virus 1 immediate early protein ICP22 in viral nuclear egress. J Virol. 2014/04/16 ed. 2014 Jul;88(13):7445–54.

21. Sandri-Goldin RM. Initiation of transcription and RNA synthesis, processing and transport in HSV and VZV infected cells. In: Arvin A, Campadelli-Fiume G, Mocarski E, Moore PS, Roizman B, Whitley R, et al., editors. Human Herpesviruses: Biology, Therapy, and Immunoprophylaxis. Cambridge; 2007.

22. Mavromara-Nazos P, Roizman B. Delineation of regulatory domains of early (beta) and late (gamma 2) genes by construction of chimeric genes expressed in herpes simplex virus 1 genomes. Proc Natl Acad Sci U S A. 1989 Jun;86(11):4071–5.

23. Kim DB, Zabierowski S, DeLuca NA. The initiator element in a herpes simplex virus type 1 late-gene promoter enhances activation by ICP4, resulting in abundant late-gene expression. J Virol. 2002 Feb;76(4):1548–58.

24. Jones KA, Tjian R. Sp1 binds to promoter sequences and activates herpes simplex virus ‘immediate-early’ gene transcription in vitro. Nature. 1985 Sep 1;317(6033):179–82.

607 25. Thompson RL, Sawtell NM. Targeted Promoter Replacement Reveals That Herpes
608 Simplex Virus Type-1 and 2 Specific VP16 Promoters Direct Distinct Rates of Entry Into
609 the Lytic Program in Sensory Neurons in vivo. Front Microbiol [Internet]. 2019;10.
610 Available from: https://www.frontiersin.org/article/10.3389/fmicb.2019.01624

611 26. Hirai H, Tani T, Kikyo N. Structure and functions of powerful transactivators: VP16,
612 MyoD and FoxA. Int J Dev Biol. 2010;54(11–12):1589–96.

613 27. Birkenheuer CH, Danko CG, Baines JD. Herpes Simplex Virus 1 Dramatically Alters
614 Loading and Positioning of RNA Polymerase II on Host Genes Early in Infection. J
615 Virol. 2018;92(8):e02184-17.

616 28. Knipe DM, Cliffe A. Chromatin control of herpes simplex virus lytic and latent infection.
617 Nat Rev Microbiol. 2008 Mar 1;6(3):211–21.

618 29. Hu M, Depledge DP, Flores Cortes E, Breuer J, Schang LM. Chromatin dynamics and the
619 transcriptional competence of HSV-1 genomes during lytic infections. PLOS Pathog.
620 2019;15(11):e1008076.

621 30. Dremel SE, DeLuca NA. Herpes simplex viral nucleoprotein creates a competitive
622 transcriptional environment facilitating robust viral transcription and host shut off. eLife.
623 2019;8:e51109.

624 31. Michael N, Roizman B. Repression of the herpes simplex virus 1 alpha 4 gene by its gene
625 product occurs within the context of the viral genome and is associated with all three
626 identified cognate sites. Proc Natl Acad Sci U S A. 1993 Mar 15;90(6):2286–90.

627 32. Tunnicliffe RB, Lockhart-Cairns MP, Levy C, Mould AP, Jowitt TA, Sito H, et al. The
628 herpes viral transcription factor ICP4 forms a novel DNA recognition complex. Nucleic
629 Acids Res. 2017 Jul 27;45(13):8064–78.

630 33. Kuddus R, Gu B, DeLuca NA. Relationship between TATA-binding protein and herpes
631 simplex virus type 1 ICP4 DNA-binding sites in complex formation and repression of
632 transcription. J Virol. 1995 Sep;69(9):5568–75.

633 34. Xia K, DeLuca NA, Knipe DM. Analysis of phosphorylation sites of herpes simplex
634 virus type 1 ICP4. J Virol. 1996 Feb;70(2):1061–71.

635 35. Wang Y, Fairley JA, Roberts SGE. Phosphorylation of TFIIB links transcription
636 initiation and termination. Curr Biol CB. 2010/03/11 ed. 2010 Mar 23;20(6):548–53.

637 36. Carrozza MJ, DeLuca NA. Interaction of the viral activator protein ICP4 with TFIID
638 through TAF250. Mol Cell Biol. 1996 Jun;16(6):3085–93.

639 37. Jonkers I, Lis JT. Getting up to speed with transcription elongation by RNA polymerase
640 II. Nat Rev Mol Cell Biol. 2015/02/18 ed. 2015 Mar;16(3):167–77.

641 38. Fox HL, Dembowski JA, DeLuca NA. A Herpesviral Immediate Early Protein Promotes
642 Transcription Elongation of Viral Transcripts. mBio. 2017 Jun 13;8(3):e00745-17.

643 39. Guo L, Wu W juan, Liu L ding, Wang L chun, Zhang Y, Wu L qiu, et al. Herpes Simplex
644 Virus 1 ICP22 Inhibits the Transcription of Viral Gene Promoters by Binding to and
645 Blocking the Recruitment of P-TEFb. PLOS ONE. 2012 Sep 24;7(9):e45749.

646 40. Zaborowska J, Baumli S, Laitem C, O'Reilly D, Thomas PH, O'Hare P, et al. Herpes
647 Simplex Virus 1 (HSV-1) ICP22 protein directly interacts with cyclin-dependent kinase

(CDK)9 to inhibit RNA polymerase II transcription elongation. PloS One. 2014 Sep 18;9(9):e107654–e107654.

41. Isa NF, Bensaude O, Aziz NC, Murphy S. HSV-1 ICP22 Is a Selective Viral Repressor of Cellular RNA Polymerase II-Mediated Transcription Elongation. Vaccines. 2021 Sep 22;9(10):1054.

42. Hagglund R, Roizman B. Role of ICP0 in the strategy of conquest of the host cell by herpes simplex virus 1. J Virol. 2004;78(5):2169–78.

43. Yao F, Schaffer PA. Physical interaction between the herpes simplex virus type 1 immediate-early regulatory proteins ICP0 and ICP4. J Virol. 1994 Dec;68(12):8158–68.

44. DeLuca NA, Schaffer PA. Physical and functional domains of the herpes simplex virus transcriptional regulatory protein ICP4. J Virol. 1988 Mar;62(3):732–43.

45. Mahat DB, Kwak H, Booth GT, Jonkers IH, Danko CG, Patel RK, et al. Base-pair-resolution genome-wide mapping of active RNA polymerases using precision nuclear run-on (PRO-seq). Nat Protoc. 2016 Aug;11(8):1455–76.

46. Babraham Institute. SeqMonk.

47. Thorvaldsdóttir H, Robinson JT, Mesirov JP. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013 Mar;14(2):178–92.

48. Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, et al. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res. 2016/04/13 ed. 2016 Jul 8;44(W1):W160–5.

669

670 **Figure Legends**

671 **Figure 1: In the absence of ICP4, Pol activity is dysregulated throughout HSV-1**

672 **infection.**

673 HEp-2 cells were infected with either the ICP4- mutant, n12, or its genetically restored repair

674 and PRO-Seq performed at 1.5, 3 and 6 hpi. **(A)** Genome browser view of distribution of

675 PRO-Seq reads (mean of 2 biological replicates, normalised to drosophila spike-in) across the

676 HSV-1 genome. External repeat sequences were deleted for the sequencing alignment. IE

677 gene peaks are noted. **(B)** Log10 normalised read counts of all HSV-1 genes. Mann-Whitney

678 U test was used to estimate statistical significance. Log10 normalised read counts of **(C)**

679 Immediate Early, **(D)** Early, **(E)** Leaky Late and **(F)** Late genes. Black diamonds indicate

680 mean. Kruskal-Wallis test was performed to estimate statistical significance. ns = $P > 0.05$.

681

682 **Figure 2: Aberrant transcription occurs on the HSV-1 genome at 1.5 hpi in the absence**

683 **of ICP4.** HEp-2 cells were infected with either the ICP4- mutant, n12, or its genetically

684 restored repair and PRO-Seq performed at 1.5, 3 and 6 hpi. High resolution genome browser

685 view PRO-Seq tracks HSV-1 regions (mean of 2 biological replicates, normalised to

686 drosophila spike-in): **(A)** α4 and US1, shown scaled to equal sequencing depth and equal

687 viral read counts, **(B)** Ul28, UL29 and UL30, **(C)** UL18, UL19, UL20, UL20.5, UL21, UL22.

688 IE genes: red, E genes: pink, LL genes: blue, L genes: yellow. **(D)** The relative distribution of

689 reads across repair-active genes (genes that are robustly transcribed in both repair and n12 at

690 each time point) and repair-repressed genes (genes that are robustly transcribed only in n12 at

691 each time point). A closer view of the promoter region is also shown. The bootstrap

692 confidence of fit is shown in the shaded area. Proportion of reads mapping to the sense strand

693 on **(E)** all HSV-1 genes and on **(F)** repair-active and repair-repressed genes. **(G)**

694 Downstream-read index of repair-active and repair-repressed genes. Mann-Whitney U test
695 was used to estimate statistical significance between viruses or gene sets at each time point.
696 ns = P > 0.05.
697
698 **Figure 3: Both virion-associated ICP4 and *de novo* synthesis of ICP4 are required for**
699 **early transcriptional repression on the HSV-1 genome at 1.5 hpi.** HEp-2 cells were
700 infected with either the ICP4- mutant, n12, or its genetically restored repair in the presence or
701 absence of the protein synthesis inhibitor cycloheximide (CHX) and PRO-Seq performed at
702 1.5 hpi. n=2 for each treatment **(A)** Log10 normalised read counts of all HSV-1 genes.
703 DeSeq2 Log2 fold change comparison of HSV-1 gene reads, separated by gene class of **(B)**
704 repair CHX treated relative to repair untreated, **(C)** n12 CHX treated relative to n12
705 untreated. IE: immediate early, E: early, LL: leaky late, L: late. Genes with a fold change
706 adjusted p-value ≤ 0.05 are shown in red. **(D)** PRO-Seq tracks of the HSV-1 α0-US1 IE gene
707 region (mean of 2 biological replicates, normalised to drosophila spike-in). **(E)** The relative
708 distribution of reads across IE genes. The bootstrap confidence of fit is shown in the shaded
709 area.
710
711 **Figure 4: The absence of viral IE genes ICP0 and ICP22 also leads to increased**
712 **transcriptional activity on the HSV-1 genome at 1.5 hpi.** HEp-2 cells were infected with;
713 the ICP0- mutant, n212, its genetically restored repair, ΔICP22 mutant, its genetically
714 restored repair and WT (F) and PRO-Seq performed at 1.5 hpi. Genome browser view of
715 distribution of PRO-Seq reads (n12 and n212 experiments represent mean of 2 biological
716 replicates, ΔICP22 experiments represent mean of 3 biological replicates normalised to
717 drosophila spike-in) across the HSV-1 genome. Previous data from n12 and its repair at 1.5
718 hpi is also shown. The black lines indicate separate infection experiments/library

719 preparations. External repeat sequences were deleted for the sequencing alignment. IE gene
720 peaks are noted.
721
722 **Figure 5: The viral IE genes ICP0 and ICP22 are involved in early transcriptional**
723 **repression the HSV-1 genome at 1.5 hpi.** HEp-2 cells were infected with; the ICP0- mutant,
724 n212, its genetically restored repair (n=2), ΔICP22 mutant, its genetically restored repair and
725 WT (F) (n=3) and PRO-Seq performed at 1.5 hpi. Previous data from n12 and its repair at 1.5
726 hpi is also shown **(A)** Log10 normalised read counts of all HSV-1 genes. High resolution
727 genome browser view PRO-Seq tracks HSV-1 regions: **(B)** UL28, UL29, UL30 and **(C)** α0
728 and US1. IE genes: red, E genes: pink, LL genes: blue. **(D)** Proportion of reads mapping to
729 the sense strand on repair-active (genes that are robustly transcribed in both repair and
730 mutant) and repair-repressed genes (genes that are robustly transcribed only in mutant). **(E)**
731 Downstream-read index of repair-active and repair-repressed genes. Mann-Whitney U test
732 was used to estimate statistical significance between mutant/repair or gene sets. ns = P >
733 0.05. **(F)** The relative distribution of reads across the promoter region on repair-active genes
734 on n212 and repair and ΔICP22 and repair. **(G)** The relative read distribution across repair-
735 repressed genes in IE mutant viruses: n212 (ICP0-), ΔICP22 and n12 (ICP4-). A closer view
736 of the promoter region is also shown. The bootstrap confidence of fit is shown in the shaded
737 area
738
739 **Figure 6: The absence of different IE genes leads to an increase of transcriptional**
740 **activity with varying patterns across the HSV-1 genome.** Log10 normalised PRO-Seq read
741 counts of HSV-1 genes, separated by temporal class at 1.5 hpi post infection with **(A)** ICP0-
742 mutant and its genetically restored repair (n=2) and **(B)** ΔICP22 and its genetically restored
743 repair (n=3) . Mann-Whitney U test was used to estimate statistical significance between

744 mutant/repair. DeSeq2 Log2 fold change comparison of HSV-1 gene reads at 1.5 hpi of **(C)**
745 ICP0- mutant n212 infected relative to repair infected **(D)** ΔICP22 infected relative to repair
746 infected and **(E)** ICP4- mutant n12 infected relative to repair infected. IE: immediate early, E:
747 early, LL: leaky late, L: late. Genes with a fold change adjusted p-value ≤ 0.05 are shown in
748 red. **(F)** Genome browser view of distribution of PRO-Seq reads (mean of 2 biological
749 replicates, normalised to drosophila spike-in) across the HSV-1 genome at 1.5, 3 and 6 hpi in
750 HEp-2 cells infected with either ICP0- mutant n212 or its repair. External repeat sequences
751 were deleted for the sequencing alignment. IE gene peaks are noted. **(G)** Log10 normalised
752 PRO-Seq read counts of HSV-1 genes at 1.5, 3 and 6 hpi in HEp-2 cells infected with either
753 ICP0- mutant n212 or its repair. Mann-Whitney U test was used to estimate statistical
754 significance between mutant/repair at each time point.
755
756 **Supplementary Information**
757 **Figure S1: Successful generation of n12 and n212 repair viruses. (A)** Western blot of
758 mock, n12 (ICP4-), n12 repair and WT (F) infected HEp-2 lysates using anti-ICP4 antibody.
759 **(B)** Western blot of mock, n212 (ICP0-), n212 repair and WT (F) infected HEp-2 lysates
760 using anti-ICP0 antibody.
761
762 **Figure S2: Pause index values for individual HSV-1 genes after infection with IE**
763 **mutants and their corresponding repairs.** Pause index) values calculated from PRO-Seq
764 data of repair-active (genes that are robustly transcribed in both repair and mutant) and
765 repair-repressed (genes that are robustly transcribed only in mutant) HSV-1 genes in **(A)**
766 ICP4- mutant n12 and n12 repair infection at 1.5 hpi (n=2), **(B)** 3 hpi (n=2) and **(C)** 6 hpi
767 (n=2). **(D)** ICP0- mutant, n212 and its repair at 1.5hpi (n=2) and **(E)** ΔICP22 and its repair at

1.5 hpi (n=3). Data is mean ± standard error. An unpaired student t-test was used to estimate statistical significance between mutant and repair on each gene.

Figure S3: During early HSV-1 infection, genes with a high level of Pol activity have the strongest regulation of sense-to-antisense transcription. Scatterplots comparing the proportion of reads mapping to the sense strand to the mean read per bp of the gene of **(A)** ICP4- mutant n12 at 1.5 hpi, **(B)** repair at 1.5 hpi, **(C)** n12 at 3 hpi, **(D)** repair at 3 hpi, **(E)** n12 at 6 hpi and **(F)** repair at 6 hpi. The R^2 correlation value is shown.

Figure S4: Antisense of transcription of specific genes occurs at regions downstream of ICP4-independent promoters. High resolution PRO-Seq tracks (mean of 2 biological replicates, normalised to drosophila spike-in) of regions containing ICP4-independent promoters (shown by green arrows) on the HSV-1 genome after infection with ICP4- mutant n12 and its repair at 1.5, 3 and 6 hpi. **(A)** The UL54 ICP4-independentt promoter and UL56 antisense transcription. **(B)** The US1 ICP4-independent promoter and US2 antisense transcription. **(C)** The US12 ICP4-independent promoter and US8/US8A/US9 antisense transcription. **(D)** The Ul39 ICP4-independent promoter and UL41 antisense transcription. **(E)** The α0 ICP4-independentt promoter and **(F)** The α4 ICP4-independent promoter. Dashed lines define regions of potential read-through transcription.

Figure S5: Aberrant transcription occurs on the HSV-1 genome at 1.5 hpi in the absence of ICP4 in HFF cells. HFF cells were infected with either the ICP4- mutant, n12, or its genetically restored repair and PRO-Seq performed at 1.5 hpi. **(A)** Genome browser view of distribution of PRO-Seq reads (mean of 2 biological replicates, normalised to drosophila spike-in) across the HSV-1 genome. External repeat sequences were deleted for the

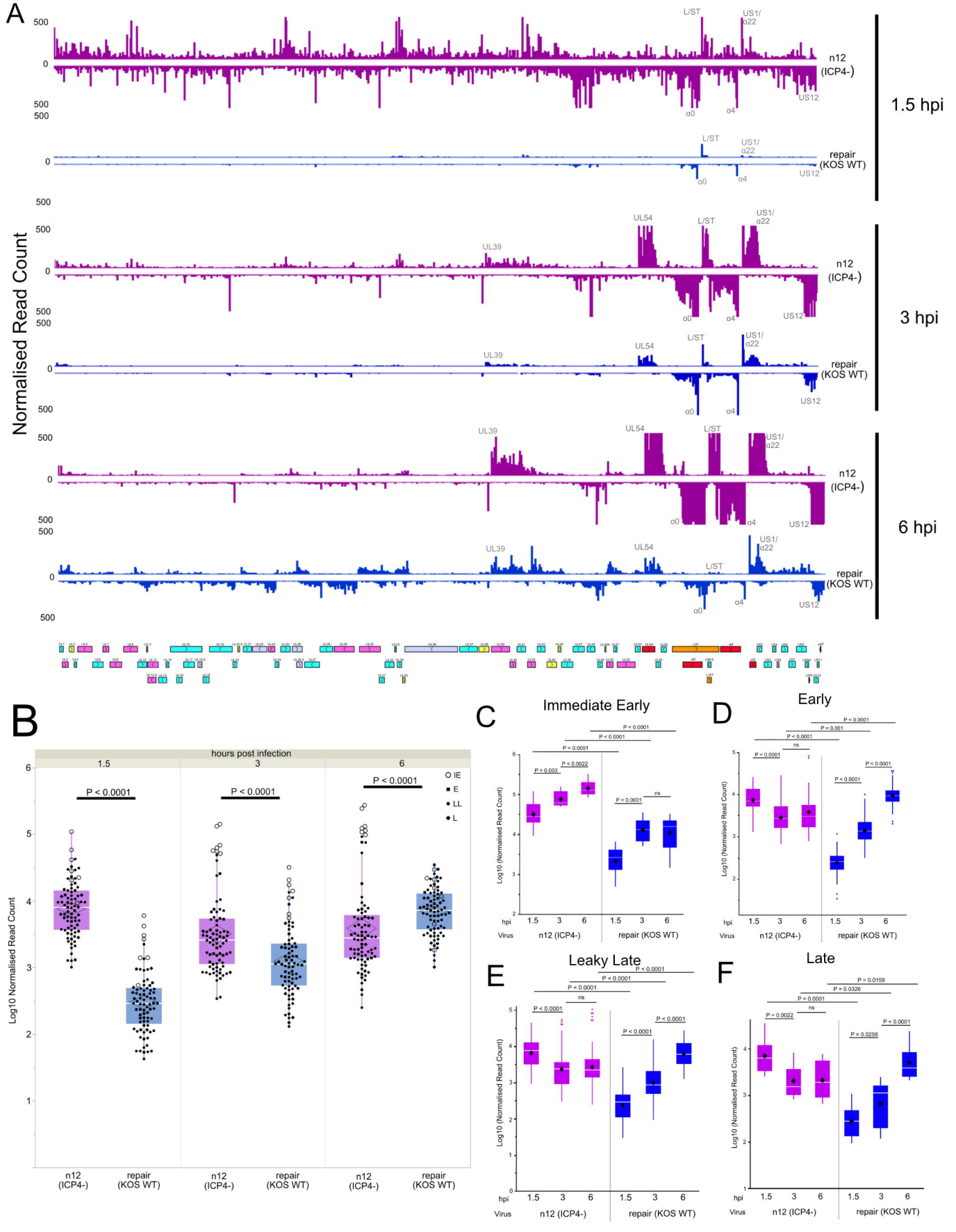
sequencing alignment. IE gene peaks are noted. **(B)** Log10 normalised read counts of all HSV-1 genes and **(C)** of genes separated by temporal class. **(D)** Proportion of reads mapping to the sense strand on repair-active (genes that are robustly transcribed in both repair and n12) and repair-repressed (genes that are robustly transcribed only in n12). genes. **(E)** Downstream-read index of repair-active and repair-repressed genes. Mann-Whitney U test was used to estimate statistical significance between mutant/repair or gene sets. ns = P>0.05. **(F)** The relative distribution of reads across repair-active genes and repair-repressed genes A closer view of the promoter region is also shown. The bootstrap confidence of fit is shown in the shaded area. **(G)** Pause index values of repair-active and repair-repressed HSV-1 genes. Data is mean of 2 biological replicates ± SEM**. (H)** The number of HSV-1 genomes in the nuclei at 1.5 hpi determined via UL51 plasmid standard-curve qPCR. Data is mean of 2 biological replicates ± standard error. An unpaired student t-test was used to estimate statistical significance, ns = P>0.05.

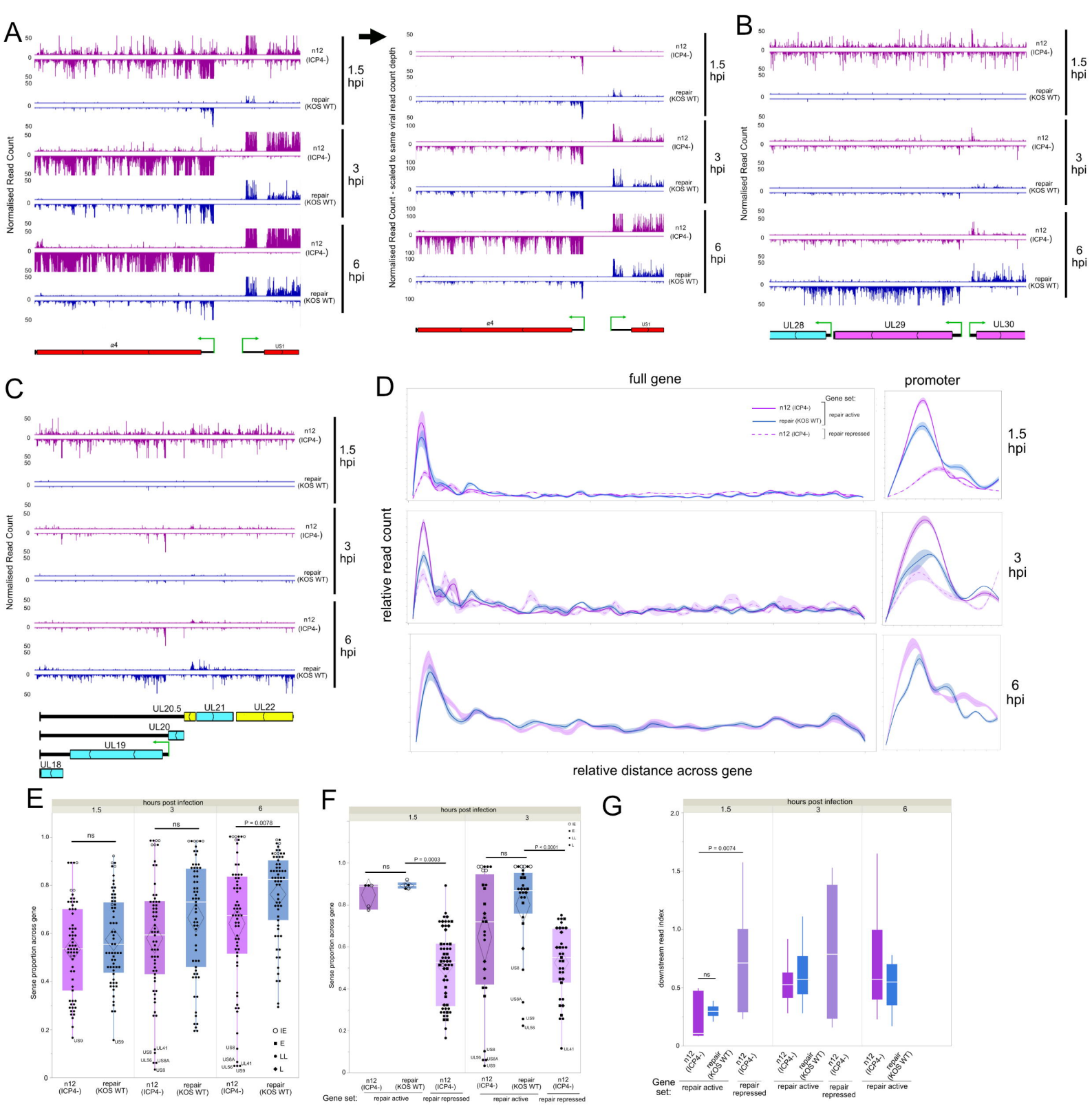
Figure S6: Confirmation of replicate clustering for each PRO-Seq experiment. Principal component analysis (PCA) plots of Log10 normalised HSV-1 gene reads (TSS-TTS) of each replicate from individual PRO-Seq sequencing experiments. **(A)** n12 (ICP4 null) and repair infected HEp-2 PRO-Seq at 1.5, 3 and 6 hpi. **(B)** n12 (ICP4 null) and repair infected HFF PRO-Seq at 1.5 hpi. **(C)** n12 (ICP4 null) and repair infected HEp-2, -/+ CHX treatment at 1.5 hpi. **(D)** n212 (ICP0 null) and repair infected HEp-2 PRO-Seq at 1.5, 3 and 6 hpi. **(E)** ΔICP22, repair and WT (F) infected HEp-2 PRO-Seq at 1.5 hpi.

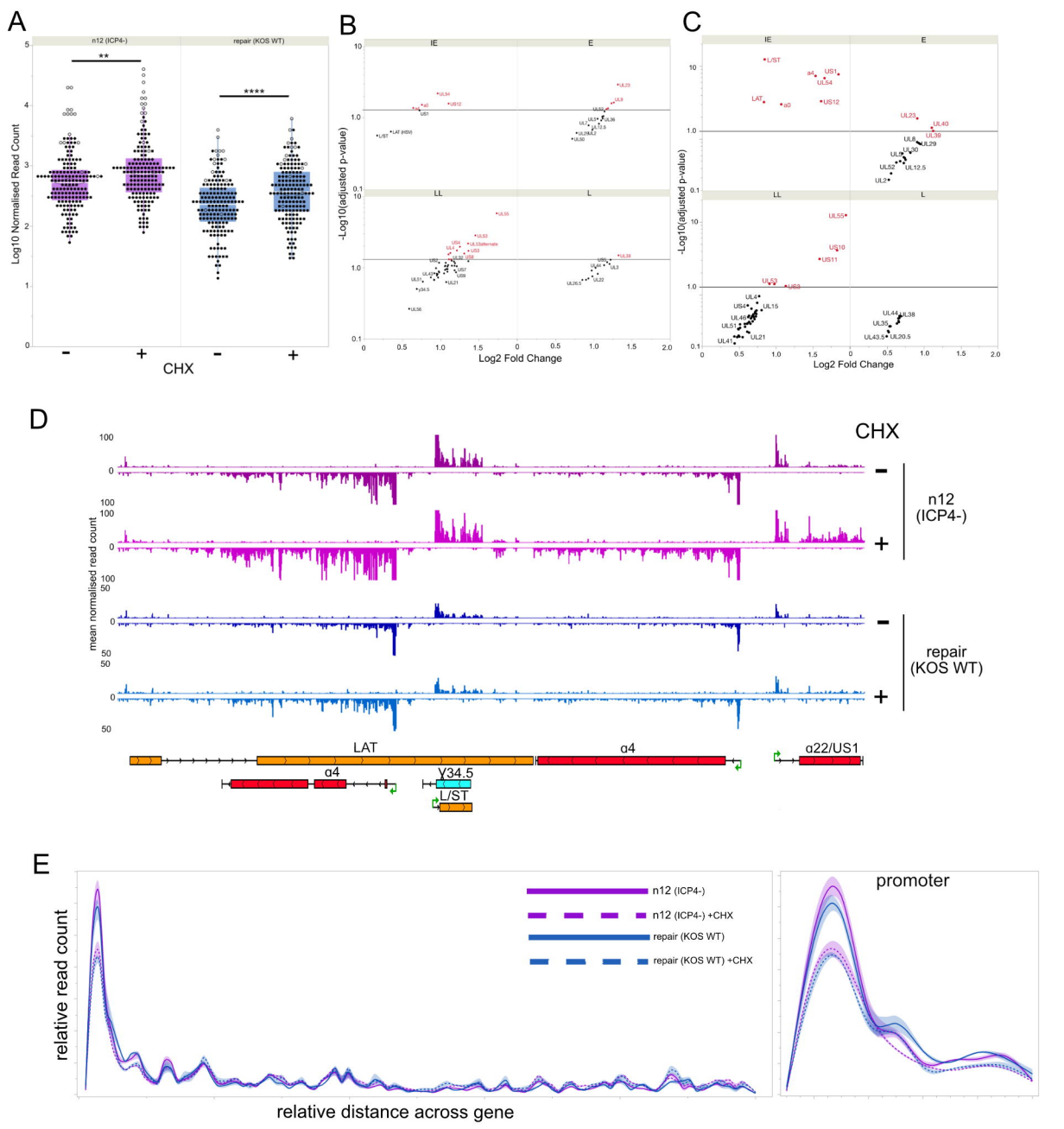
Appendix 1: Normalized HSV-1 gene reads, reads per bp and robust transcription thresholds

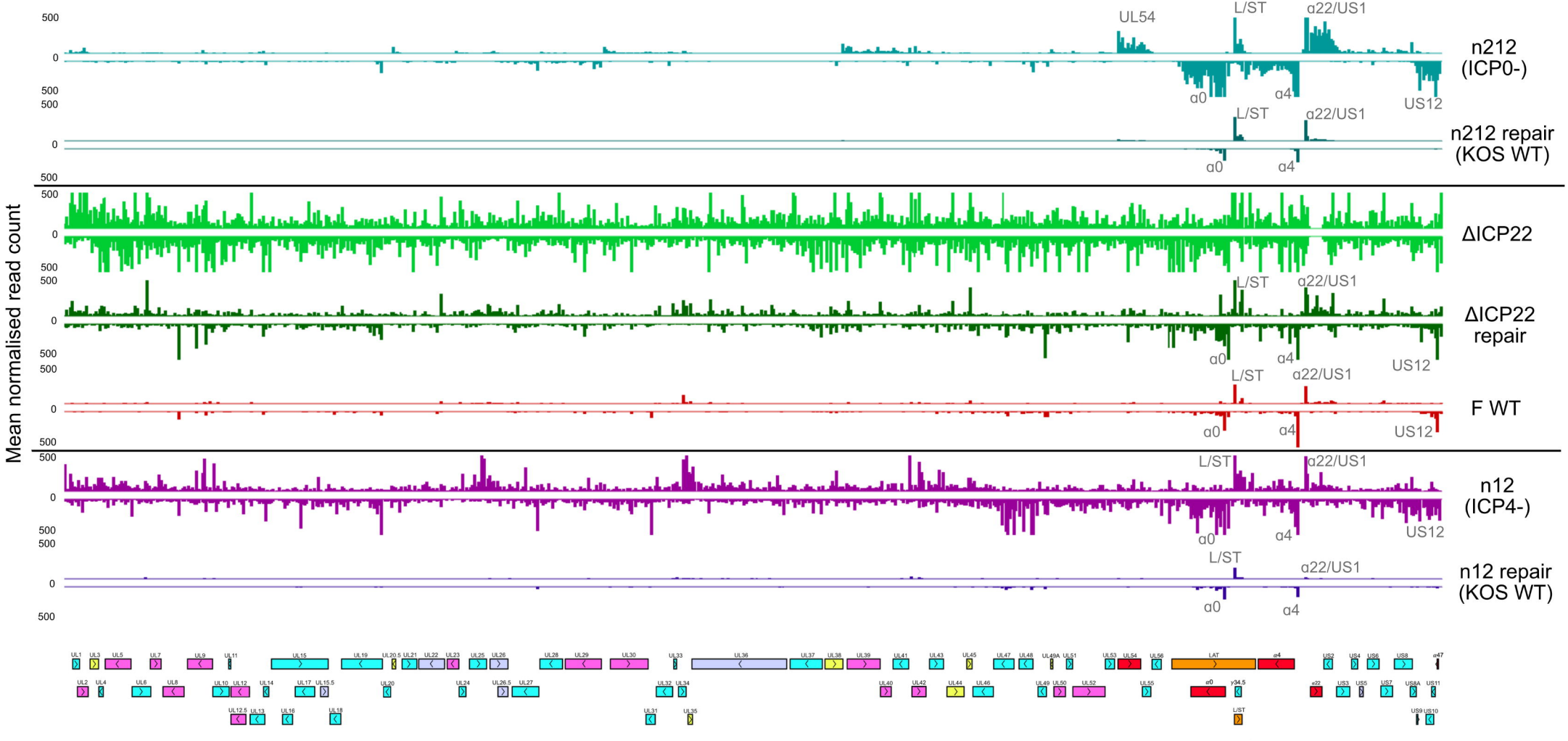
Appendix 2: DeSeq2 fold change and p-values

818

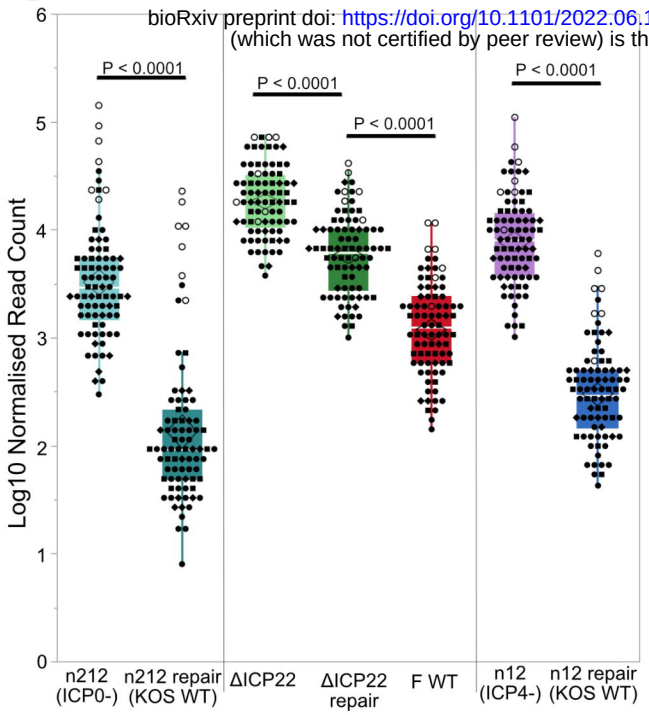

A
Normalised Read Count
1.5 hpi
3 hpi
6 hpi
n12 (ICP4-)
repair (KOS WT)
L/ST
US1/ α22
US12
α0
α4
UL39
UL54
B
hours post infection
1.5
3
6
IE
E
LL
L
P < 0.0001
Log10 Normalised Read Count
n12 (ICP4-)
repair (KOS WT)
C
Immediate Early
D
Early
E
Leaky Late
F
Late
P = 0.003
P = 0.0022
ns
P = 0.001
P = 0.0159
P = 0.0326
P = 0.0258
hpi
Virus


A
B
C
D
E
F
G
Normalised Read Count
Normalised Read Count - scaled to same viral read count depth
n12 (ICP4-)
repair (KOS WT)
1.5 hpi
3 hpi
6 hpi
α4
US1
UL28
UL29
UL30
UL20.5
UL21
UL22
UL20
UL19
UL18
full gene
promoter
Gene set:
repair active
repair repressed
relative read count
relative distance across gene
hours post infection
Sense proportion across gene
downstream read index
ns
P = 0.0078
P = 0.0003
P < 0.0001
P = 0.0074
IE
E
LL
L

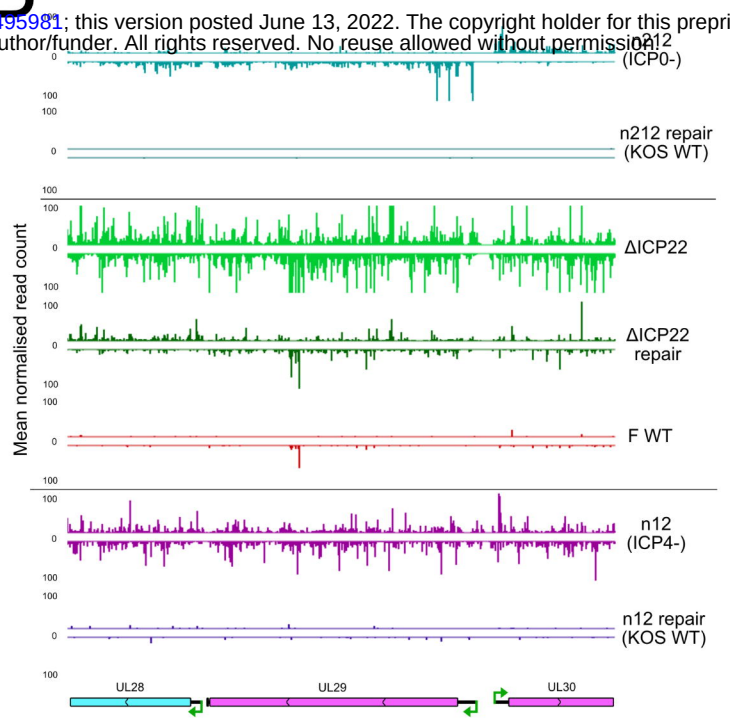

A
n12 (ICP4-)
repair (KOS WT)
Log10 Normalised Read Count
CHX
B
IE
E
LL
L
-Log10(adjusted p-value)
Log2 Fold Change
C
IE
E
LL
L
-Log10(adjusted p-value)
Log2 Fold Change
D
CHX
n12 (ICP4-)
repair (KOS WT)
mean normalised read count
LAT
α4
α22/US1
α4
V34.5
L/ST
E
relative read count
relative distance across gene
n12 (ICP4-)
n12 (ICP4-) +CHX
repair (KOS WT)
repair (KOS WT) +CHX
promoter


Mean normalised read count
500
0
500
500
0
500
UL54
L/ST
α22/US1
α0
α4
US12
n212
(ICP0-)
n212 repair
(KOS WT)
ΔICP22
ΔICP22
repair
F WT
n12
(ICP4-)
n12 repair
(KOS WT)

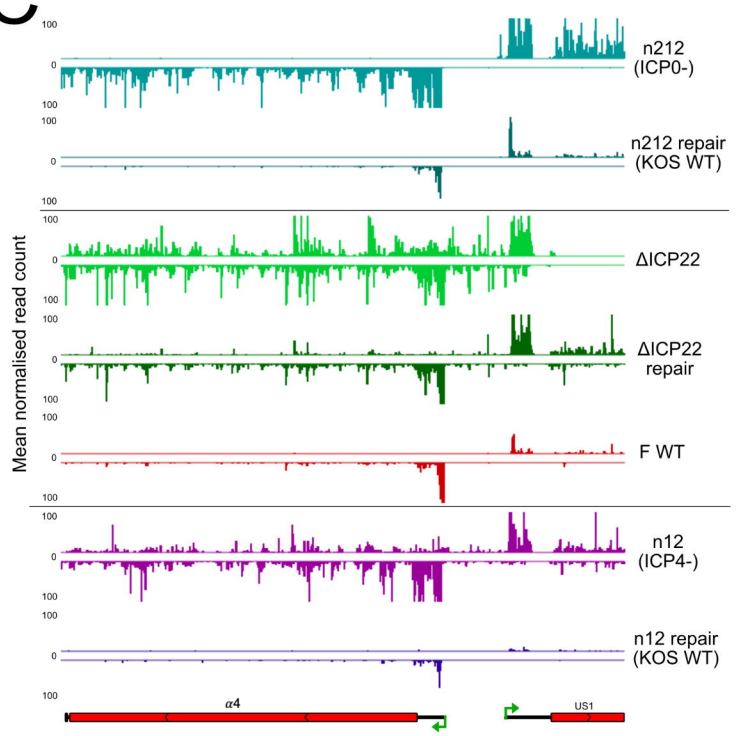
A



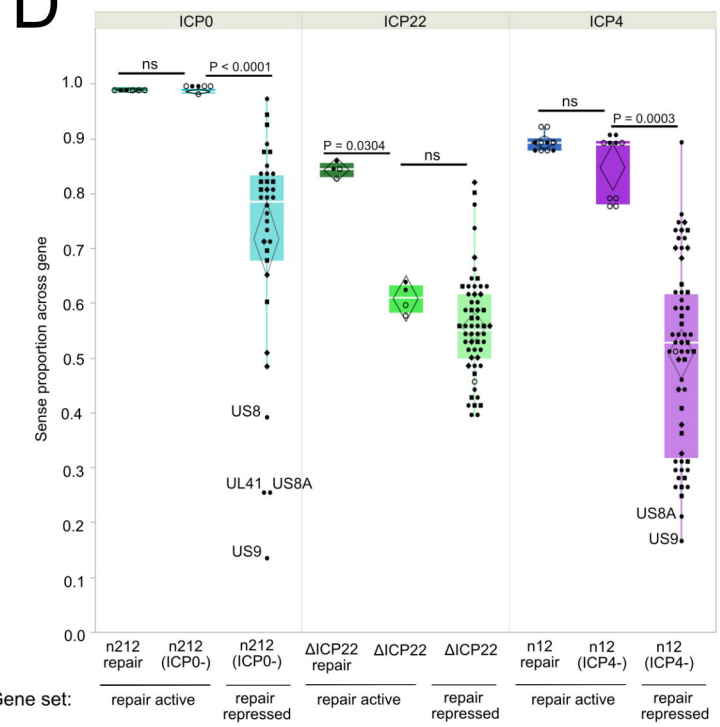
B



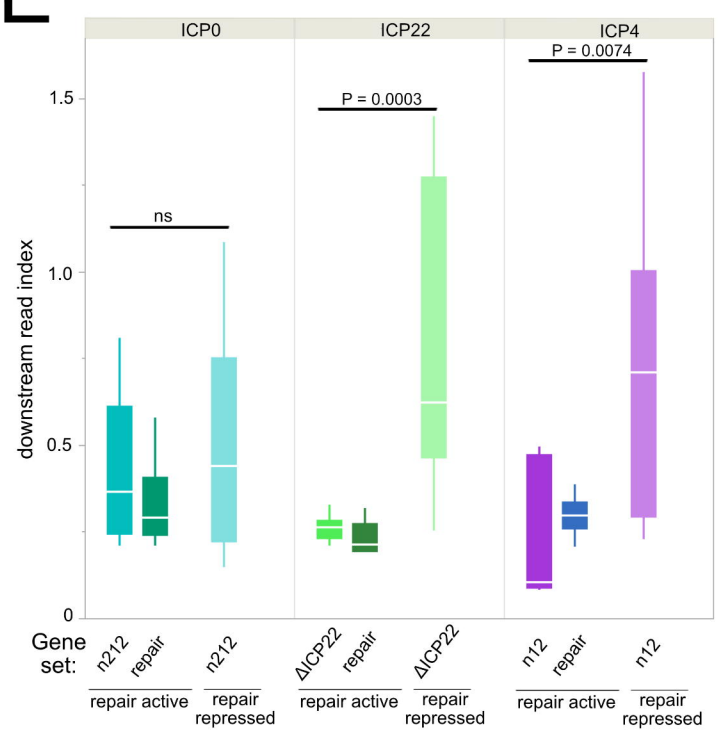
C



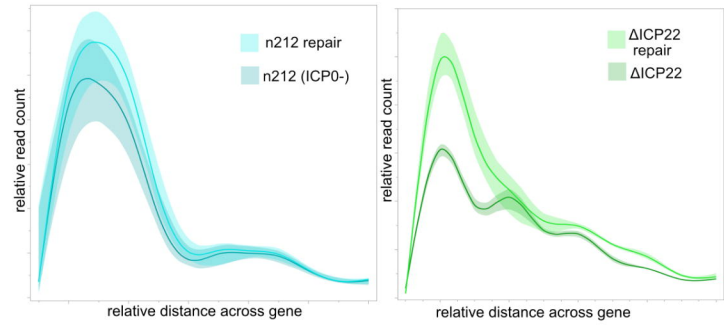
D



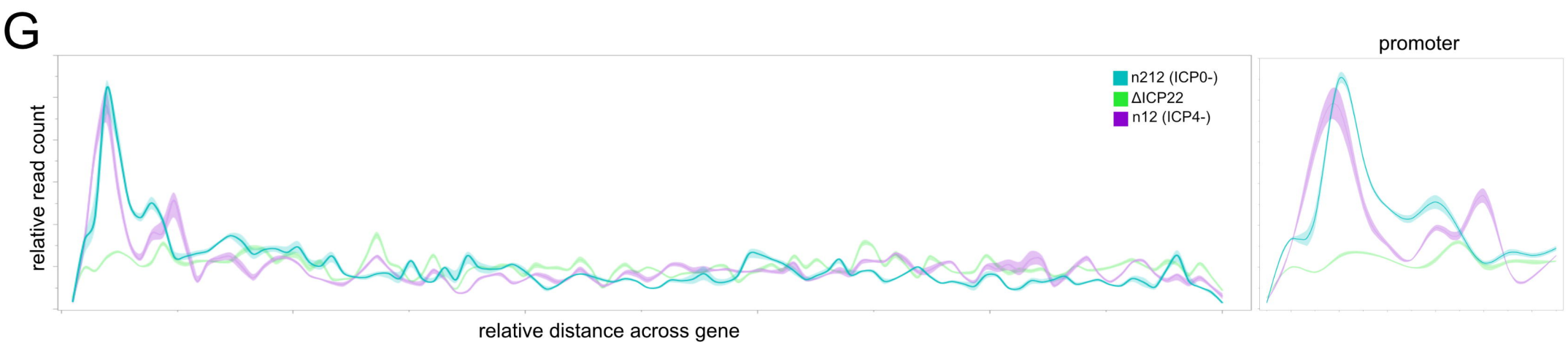
E



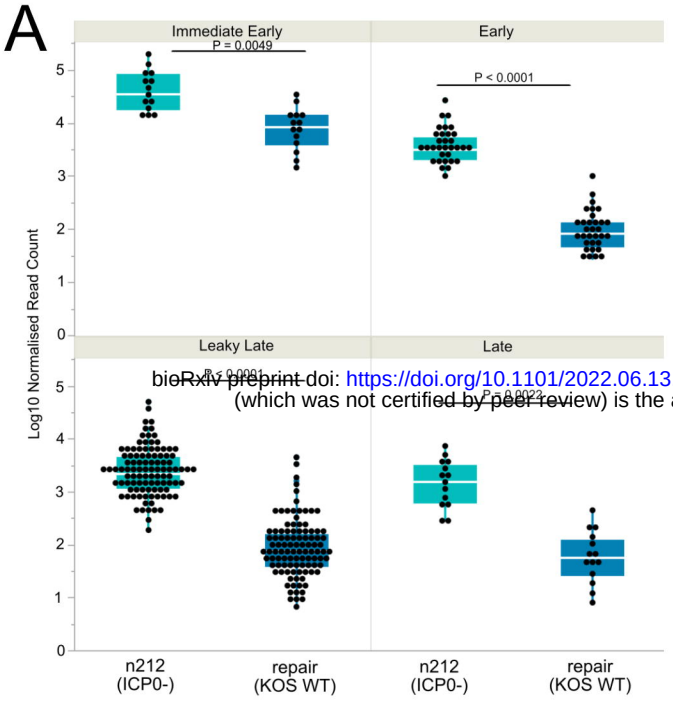
F



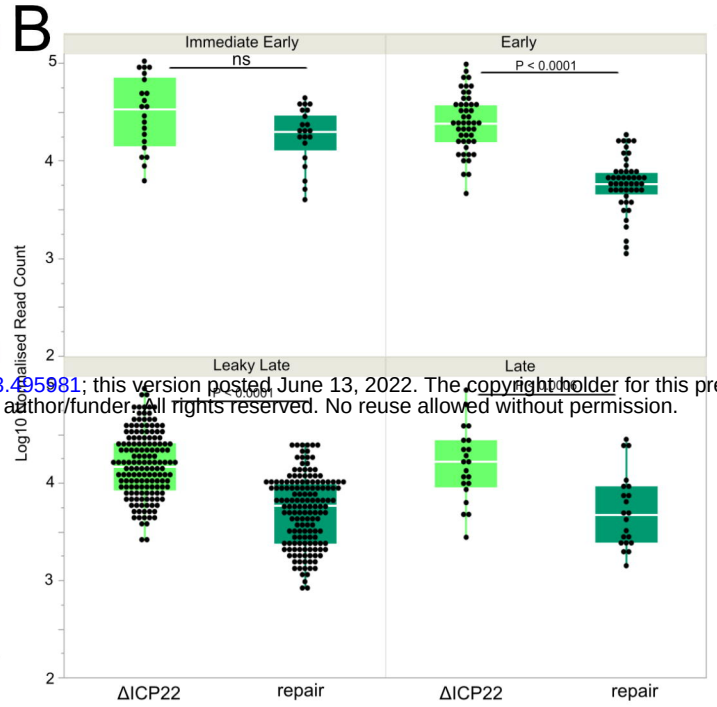
G



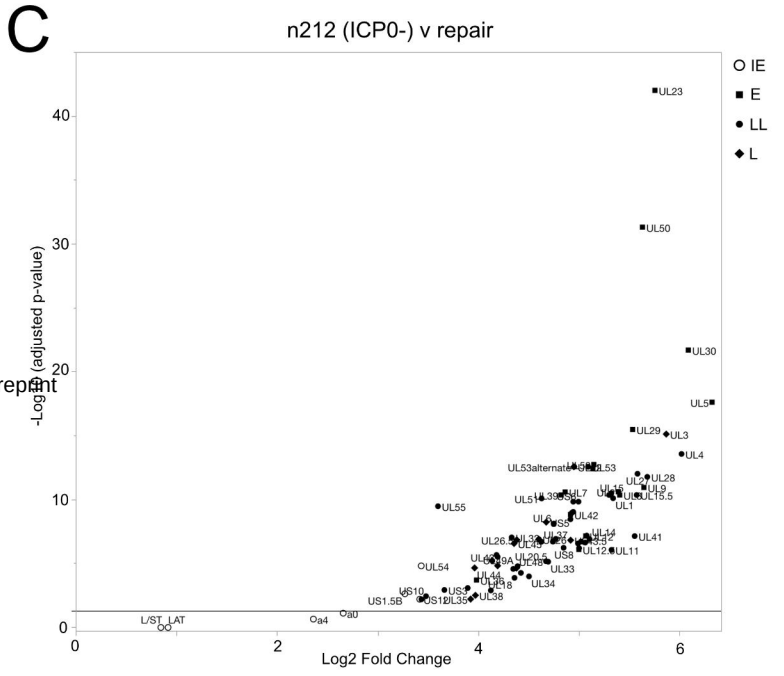
A



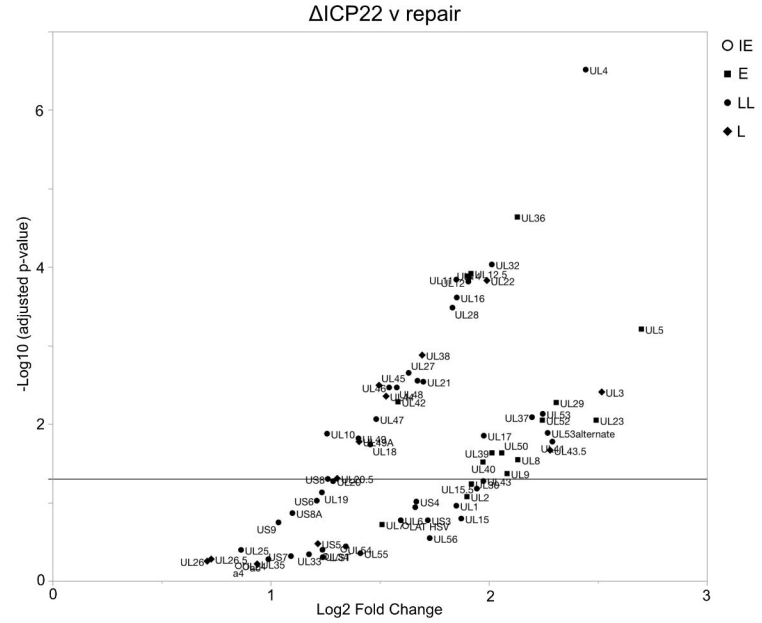
B



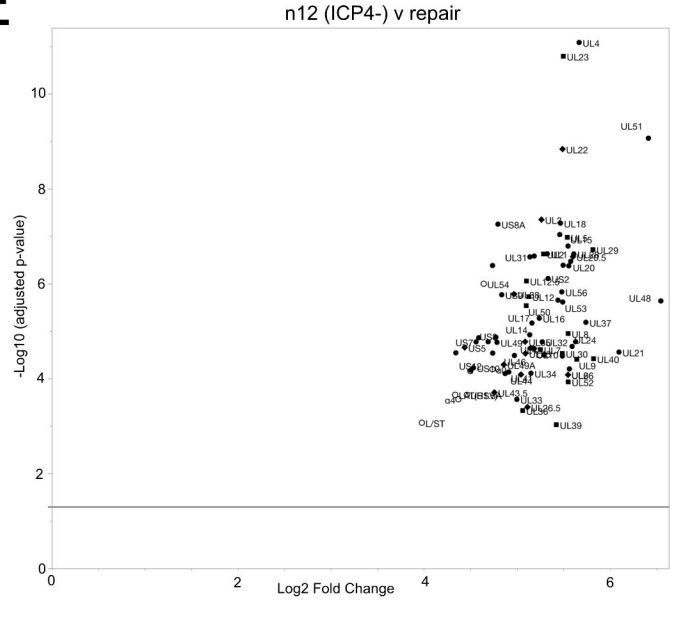
C



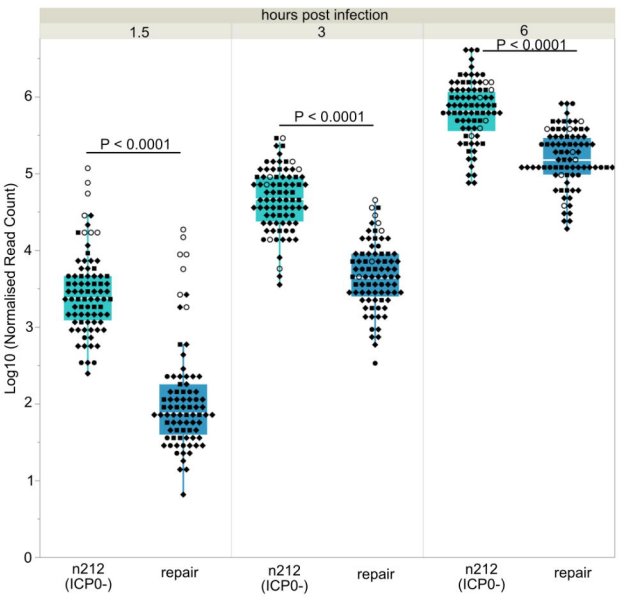
D



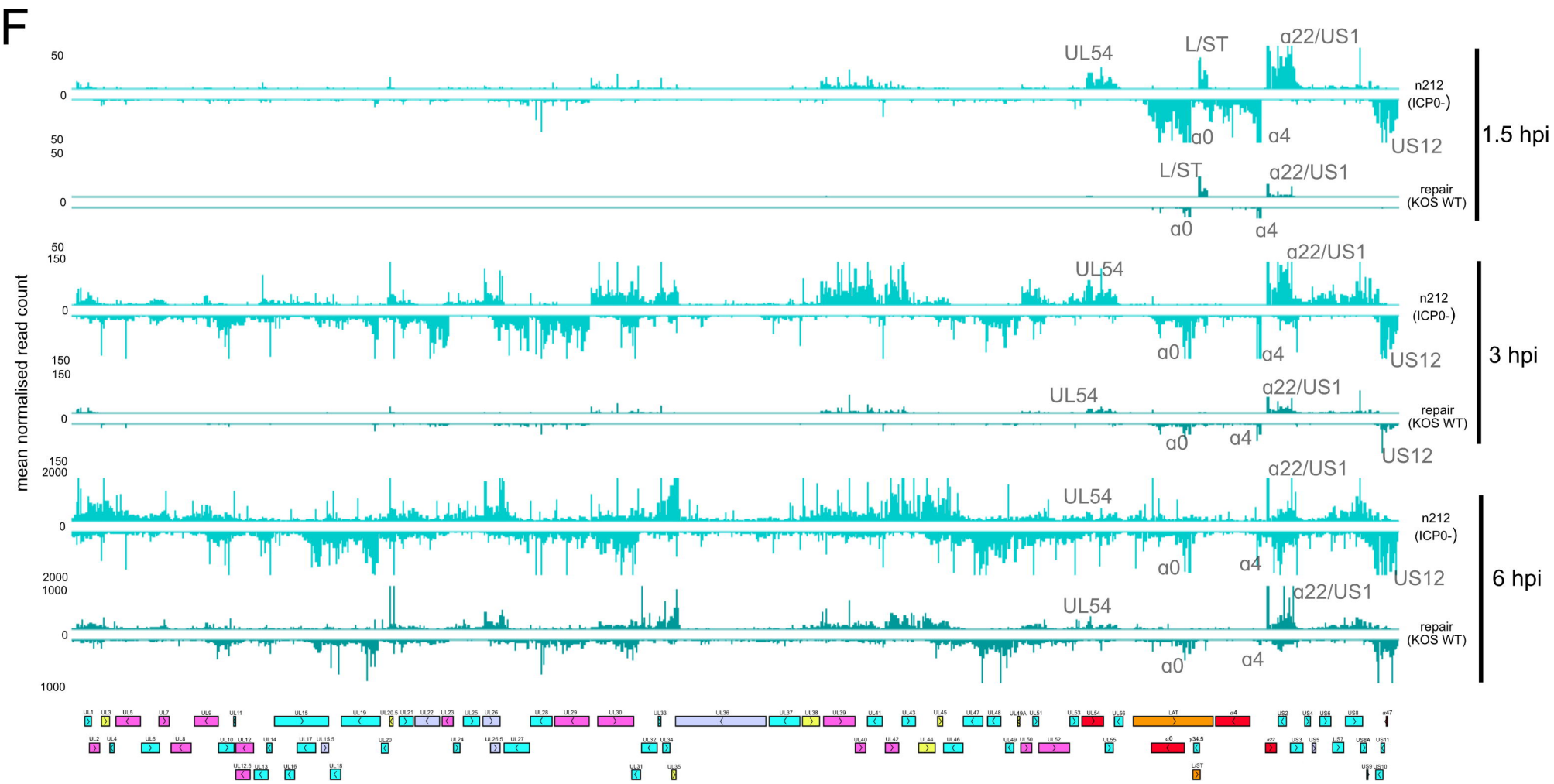
E

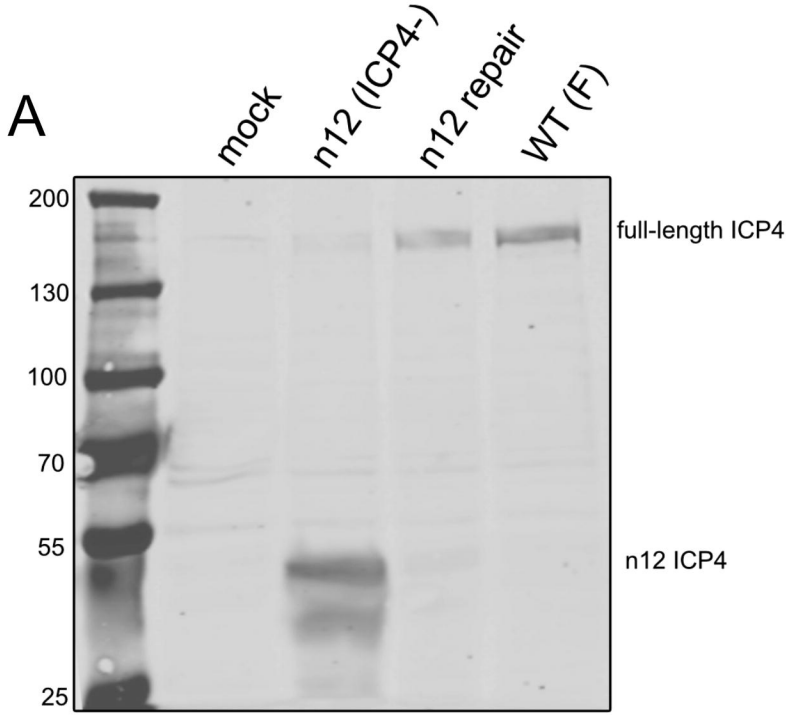


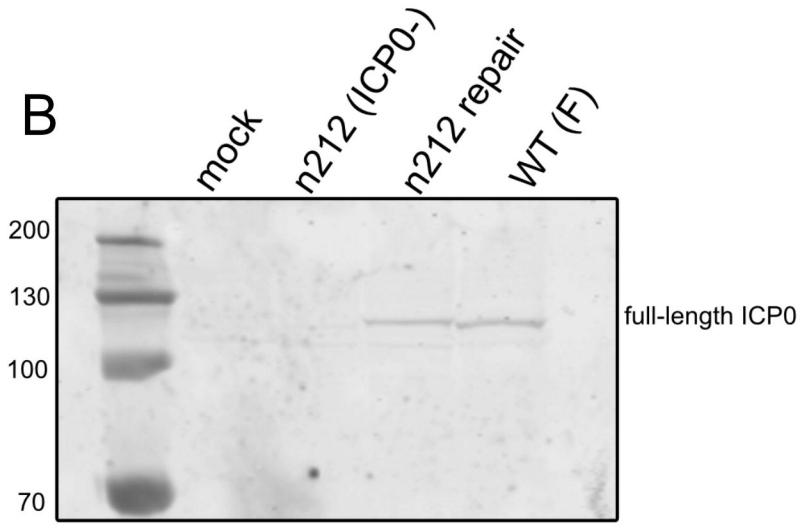
G

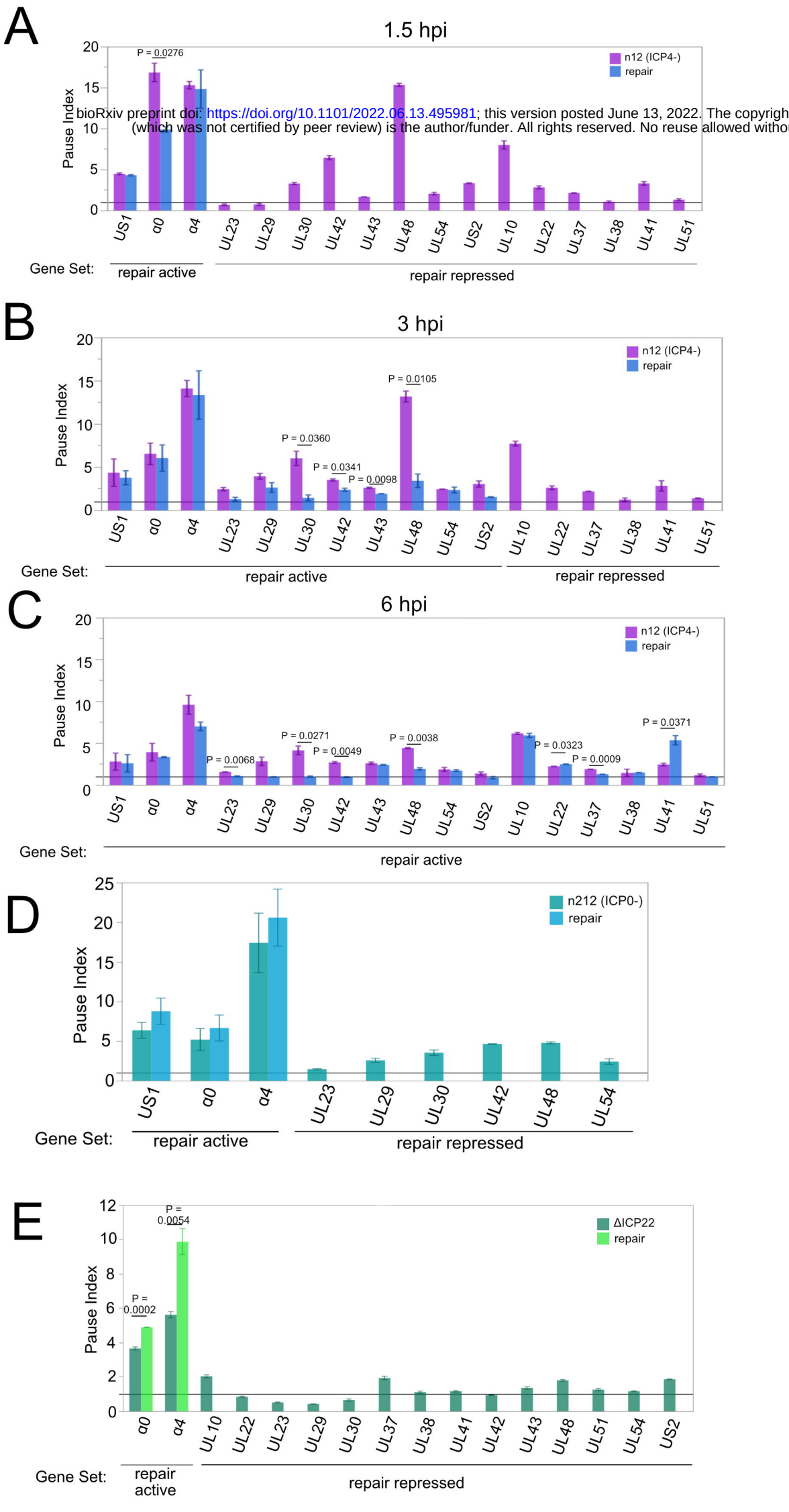


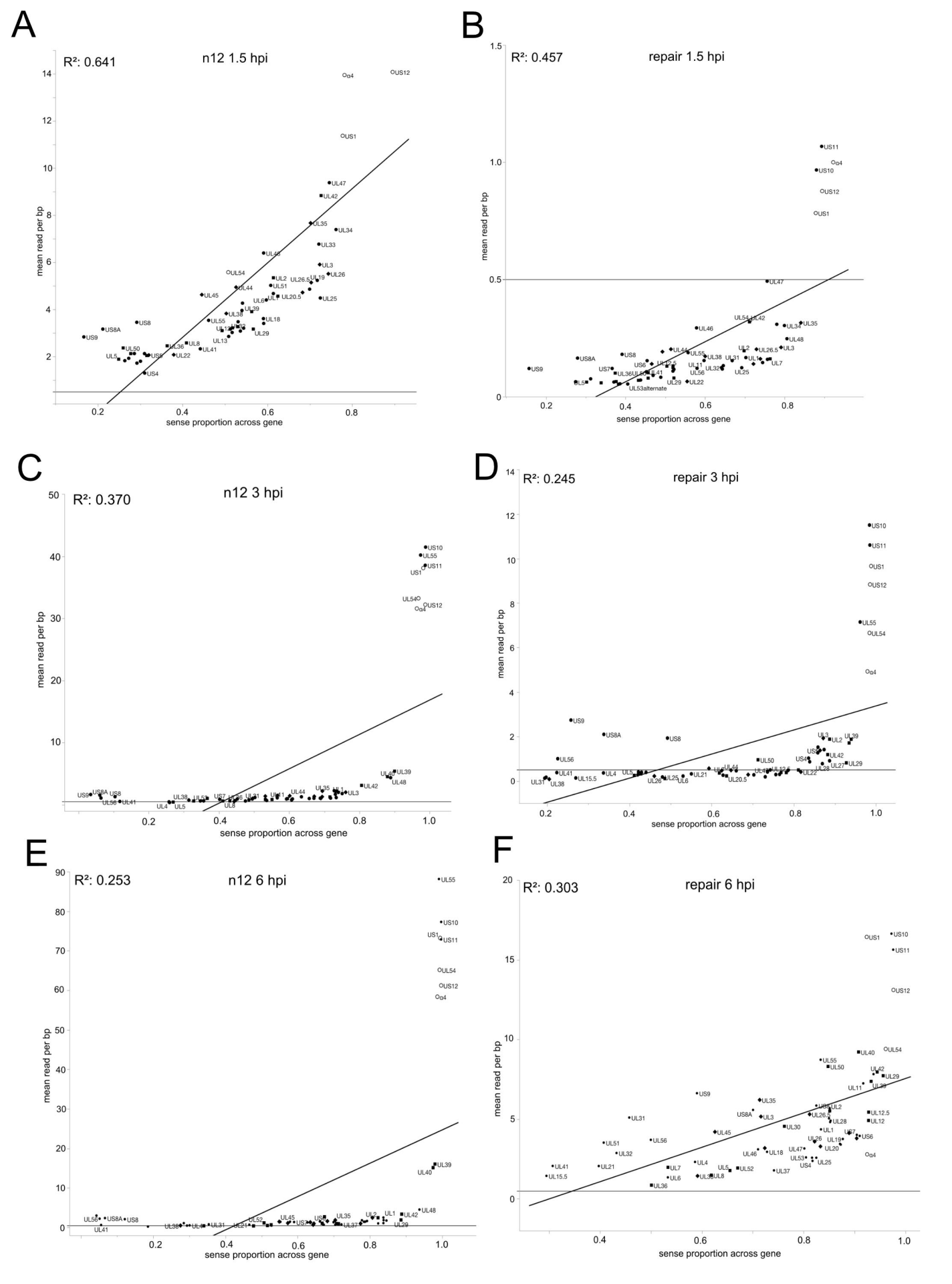
F

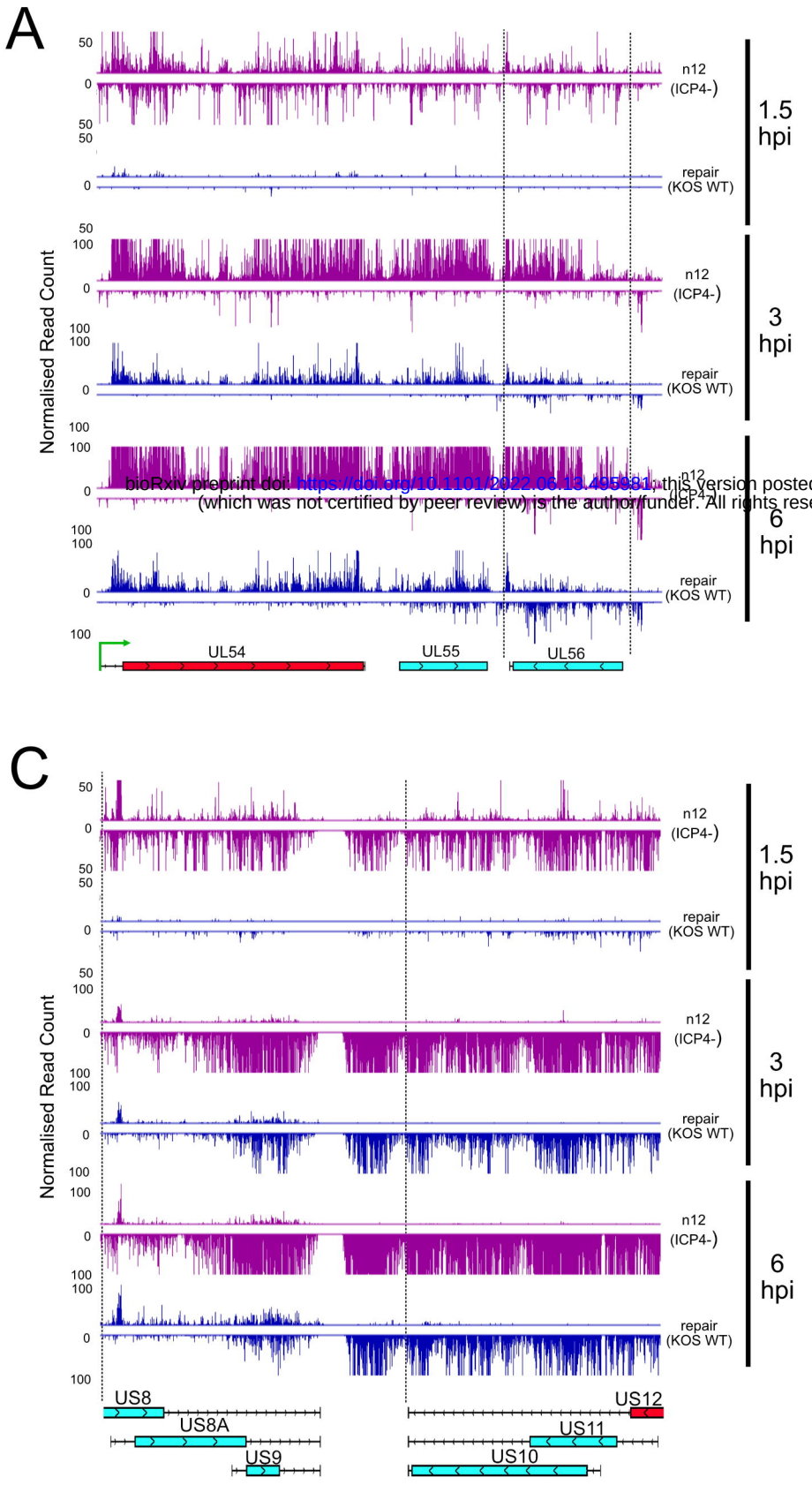


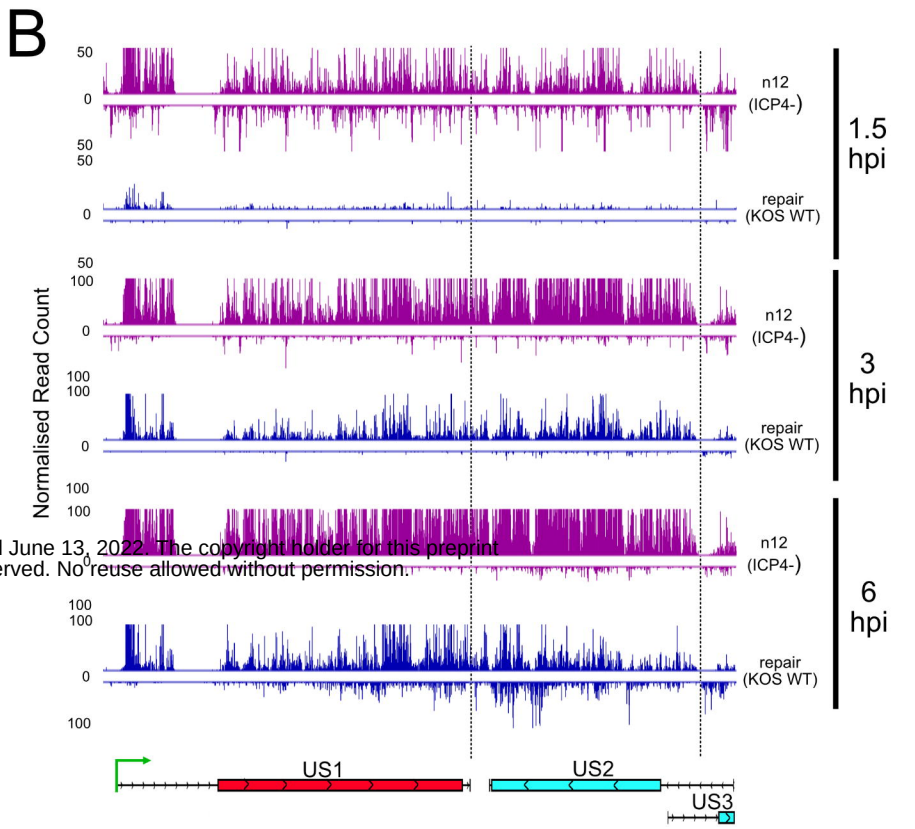

A
mock
n12 (ICP4-)
n12 repair
WT (F)
200
130
100
70
55
25
full-length ICP4
n12 ICP4

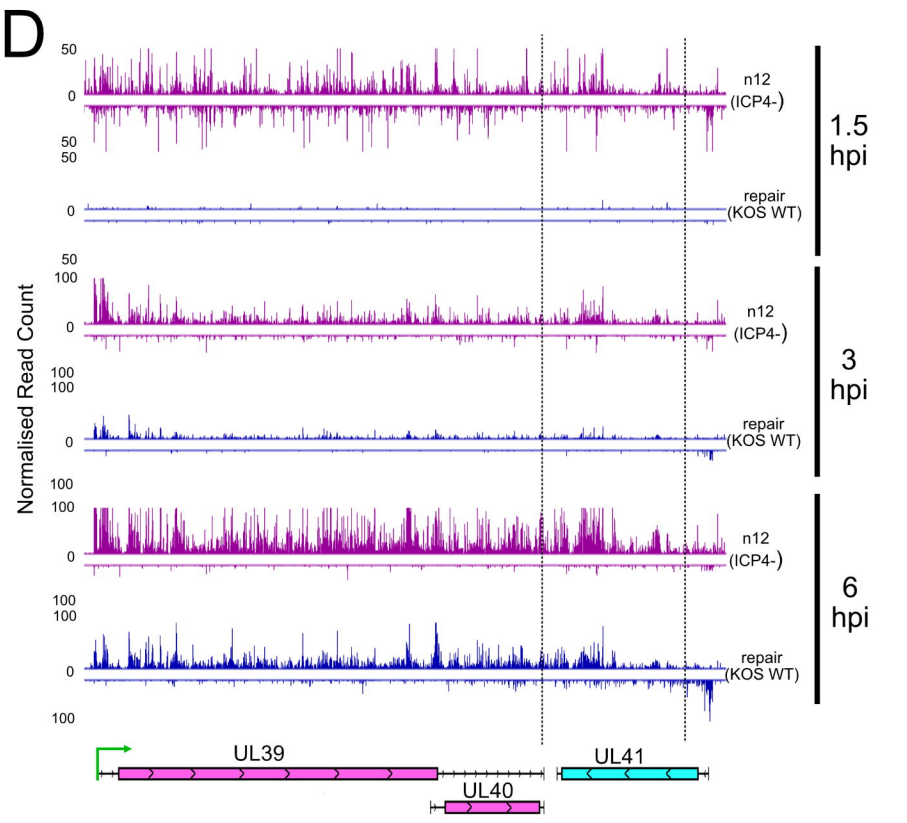

B
mock
n212 (ICP0-)
n212 repair
WT (F)
200
130
100
70
full-length ICP0

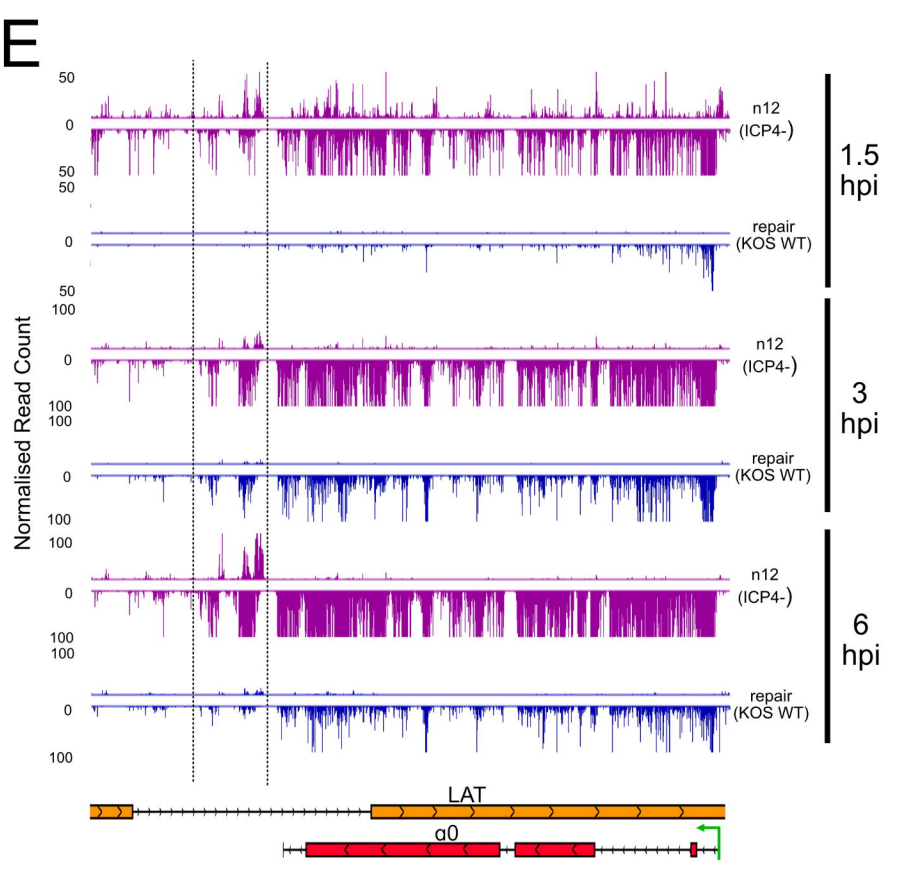


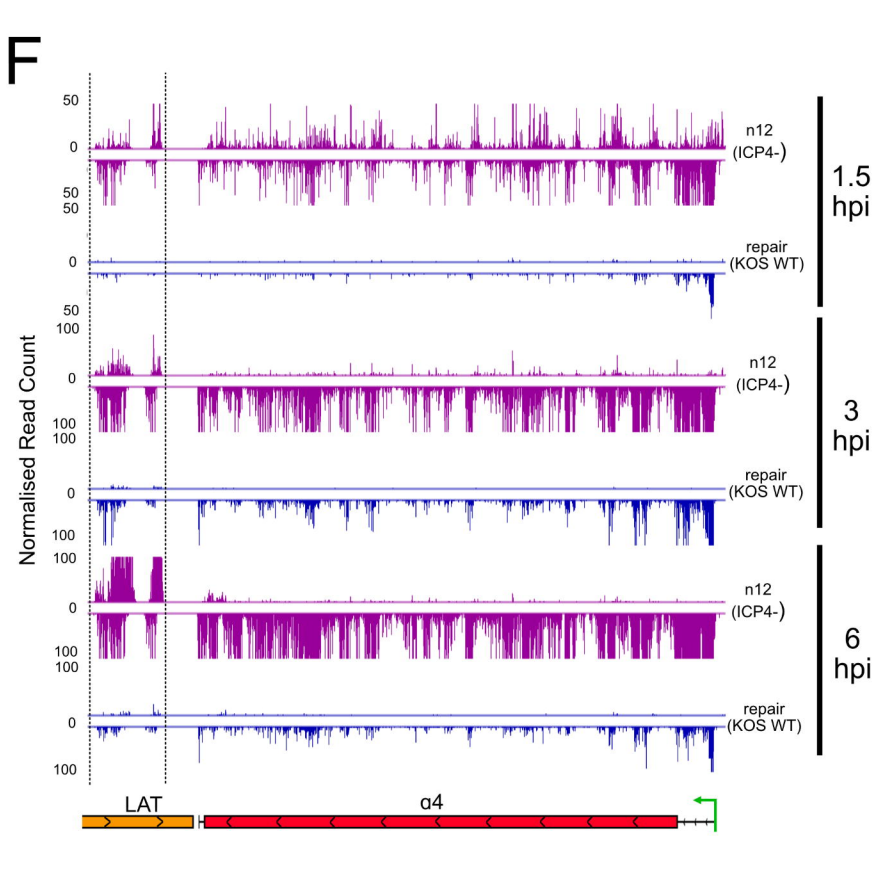

A
R²: 0.641
n12 1.5 hpi
mean read per bp
sense proportion across gene
B
R²: 0.457
repair 1.5 hpi
mean read per bp
sense proportion across gene
C
R²: 0.370
n12 3 hpi
mean read per bp
sense proportion across gene
D
R²: 0.245
repair 3 hpi
mean read per bp
sense proportion across gene
E
R²: 0.253
n12 6 hpi
mean read per bp
sense proportion across gene
F
R²: 0.303
repair 6 hpi
mean read per bp
sense proportion across gene



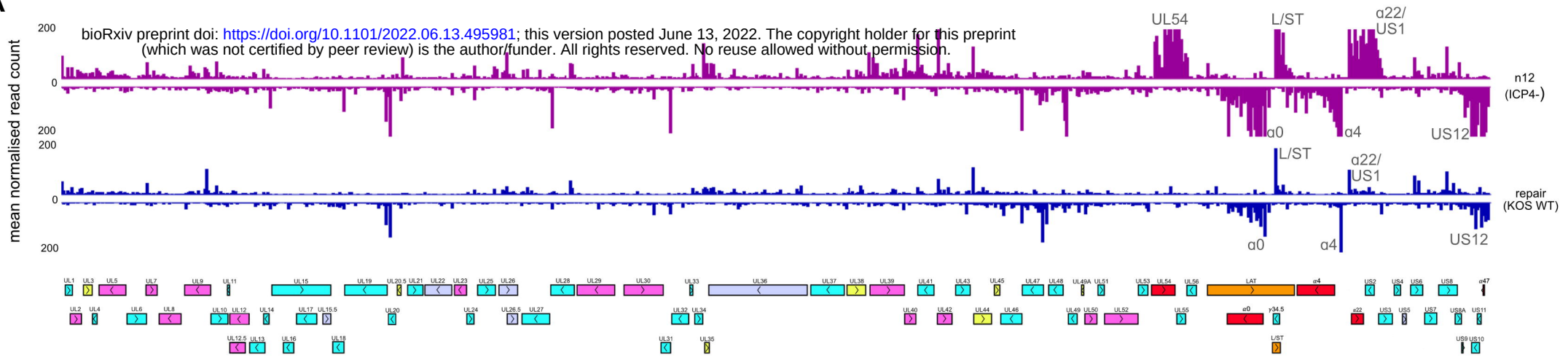




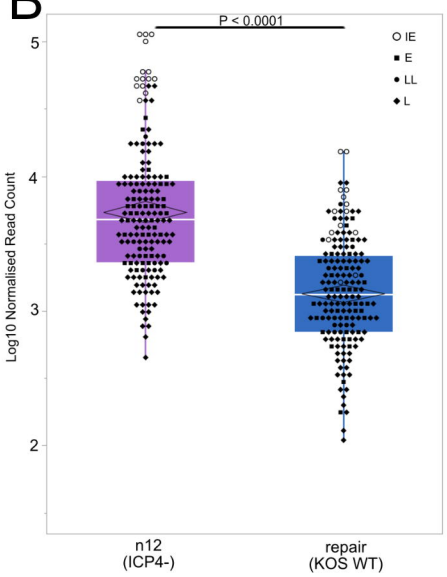




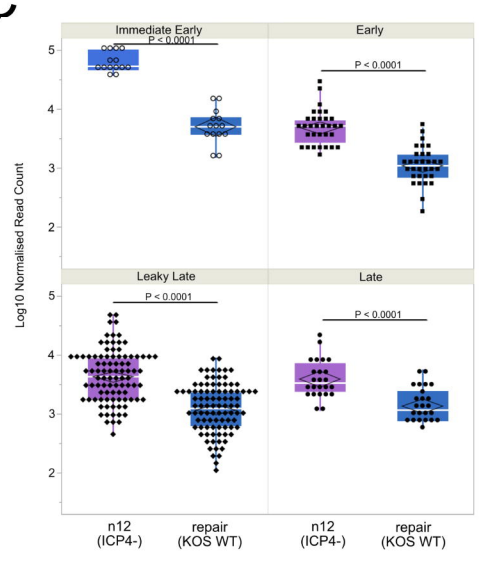
A



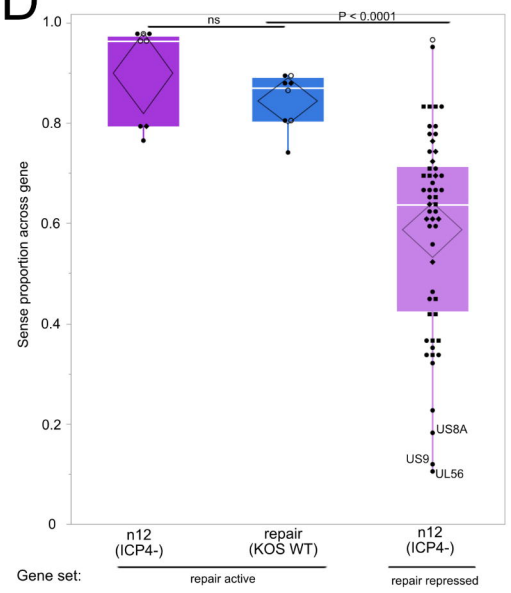
B



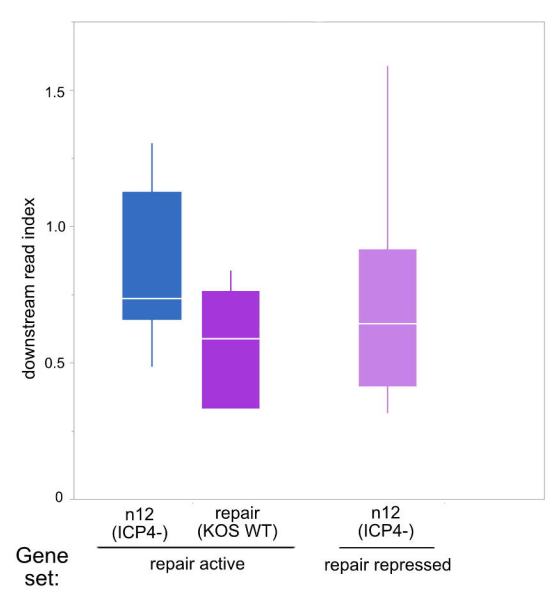
C



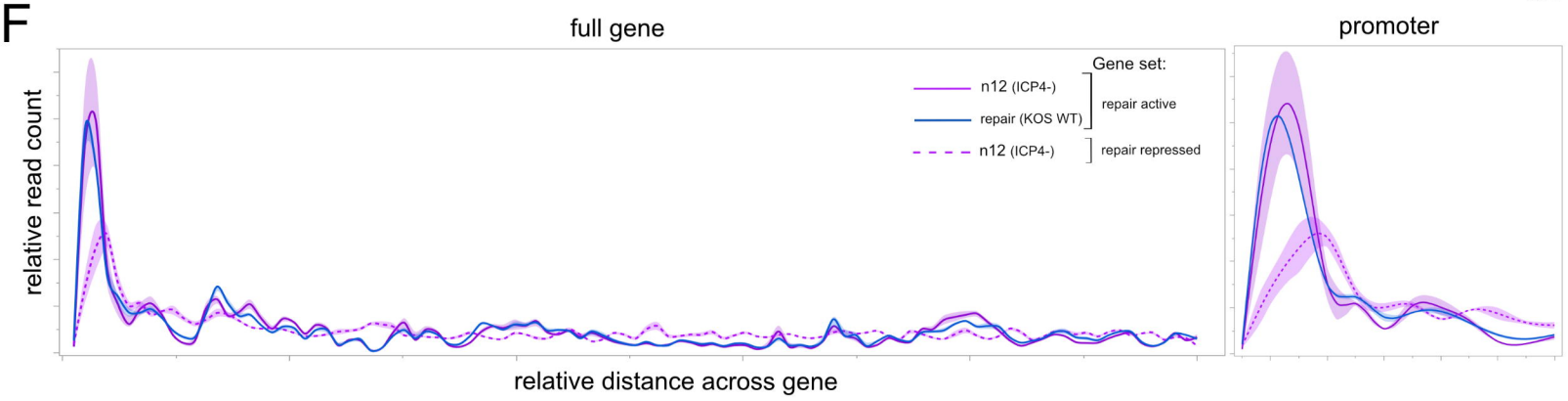
D



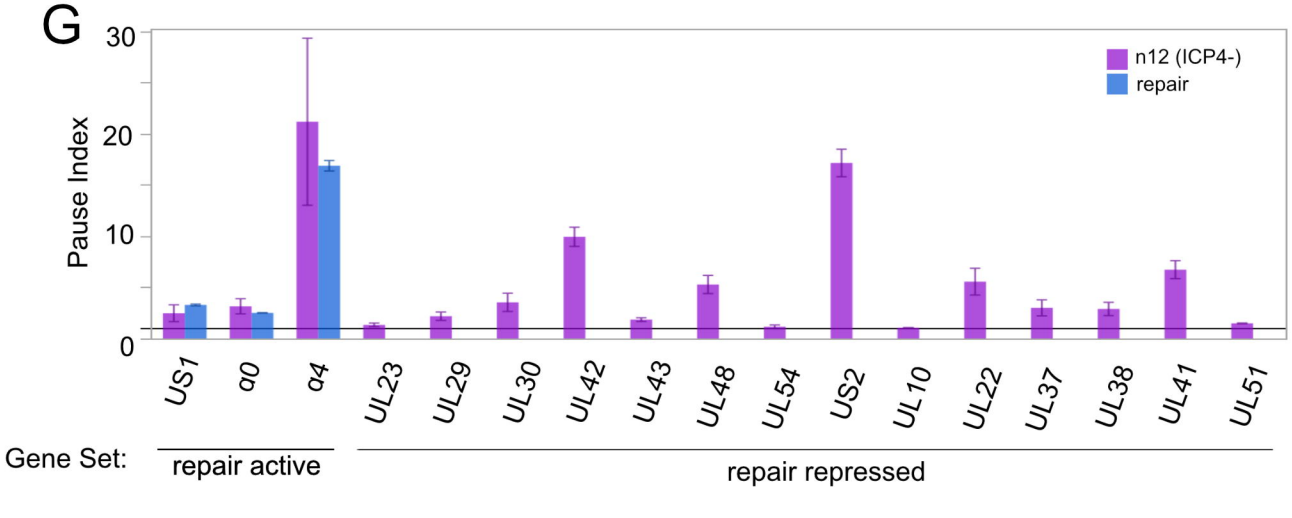
E



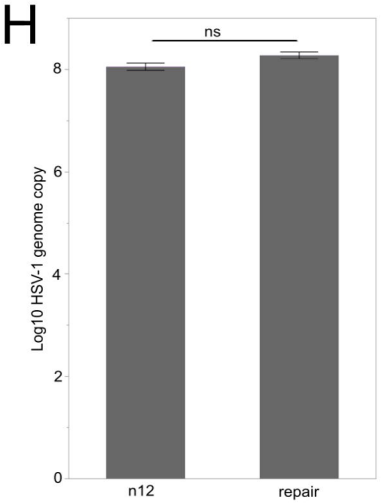
F

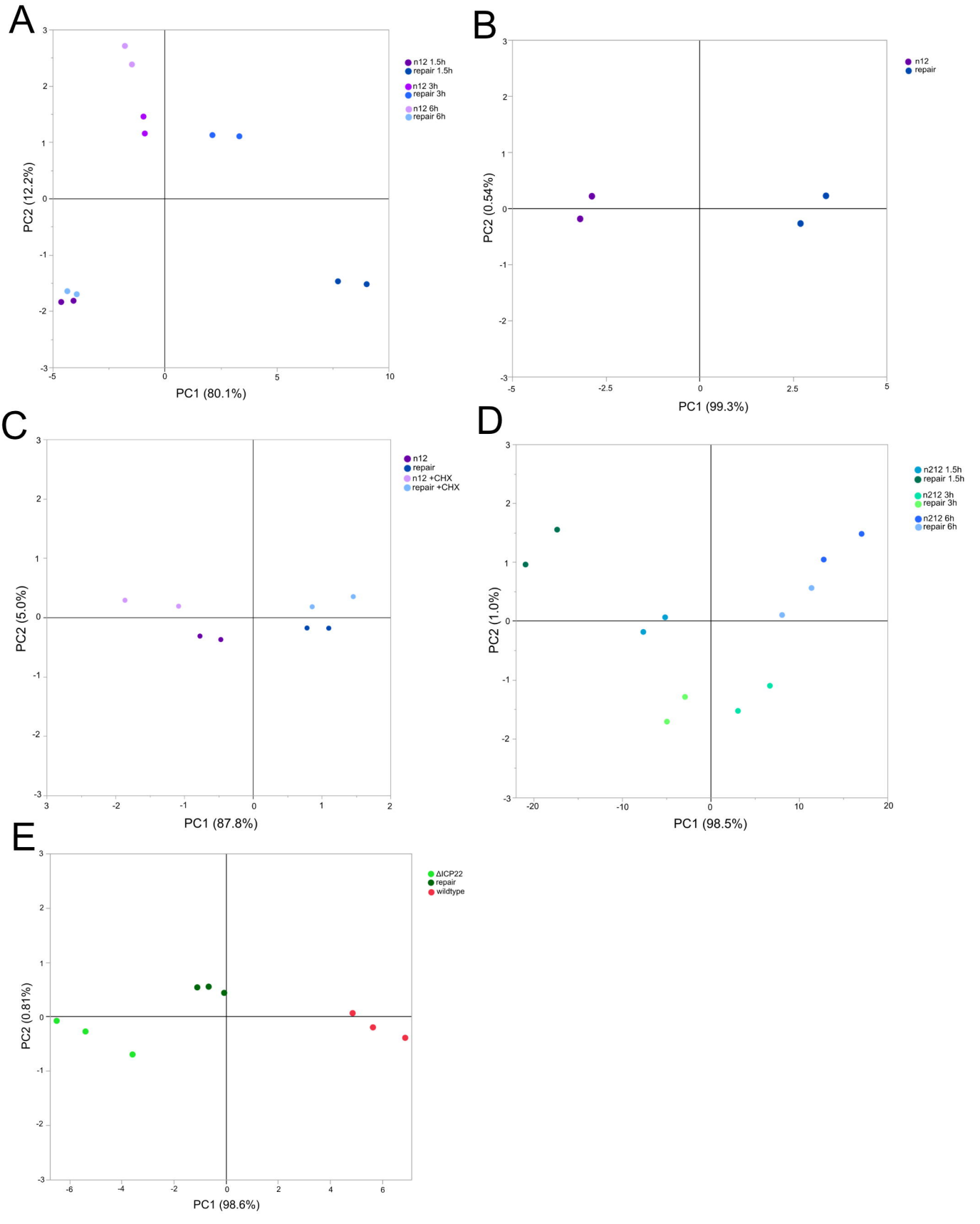


G



H




A
n12 1.5h
repair 1.5h
n12 3h
repair 3h
n12 6h
repair 6h
PC2 (12.2%)
PC1 (80.1%)
B
n12
repair
PC2 (0.54%)
PC1 (99.3%)
C
n12
repair
n12 +CHX
repair +CHX
PC2 (5.0%)
PC1 (87.8%)
D
n212 1.5h
repair 1.5h
n212 3h
repair 3h
n212 6h
repair 6h
PC2 (1.0%)
PC1 (98.5%)
E
ΔICP22
repair
wildtype
PC2 (0.81%)
PC1 (98.6%)