ASAR lncRNAs control DNA replication timing through interactions with multiple hnRNP/RNA binding proteins.

Mathew J. Thayer[1,*,#], Michael B. Heskett[2,4,#], Leslie G. Smith[1], Paul T. Spellman[2,3,5], and Phillip A. Yates[1].

[1]Department of Chemical Physiology and Biochemistry
[2] Department of Molecular and Medical Genetics
[3]Cancer Early Detection Advanced Research Center, Knight Cancer Institute
Oregon Health & Science University
3181 S. W. Sam Jackson Park Road
Portland, Oregon 97239, USA
[4]Stanford Cancer Institute
291 Campus Drive, Stanford CA, 94305.
[5]Current Address: Departments of Medicine and Human Genetics
University of California Los Angeles
Box 951722, 300 Geffen Hall
Los Angeles, CA 90095-1722

***Corresponding author.**

ORCID: 0000-0001-6483-1661

TEL: (503) 494-2447

FAX: (503) 494-7368

Email:thayerm@ohsu.edu

[#]These two authors contributed equally.

Abstract

ASARs are a family of very-long noncoding RNAs that control replication timing on individual human autosomes, and are essential for chromosome stability. The eight known ASAR genes express RNAs that remain closely associated with their parent chromosomes. Analysis of RNA-protein interaction data (from ENCODE) revealed numerous RBPs with significant interactions with multiple ASAR RNAs, with several hnRNPs as abundant interactors. An ~7kb domain within the ASAR6-141 RNA shows a striking density of RBP interaction sites. Genetic deletion and ectopic integration assays indicate that this ~7kb RNA binding protein domain contains functional sequences for controlling replication timing of entire chromosomes in *cis*. shRNA-mediated depletion of HNRNPA1, HNRNPC, HNRNPL, HNRNPM, HNRNPU, or HNRNPUL1 results in dissociation of ASAR RNAs from their chromosome territories, and disrupts the synchronous replication that occurs on all autosome pairs, recapitulating the effect of individual ASAR gene knockouts on a genome-wide scale. Our results further demonstrate the role that ASARs play during the temporal order of genome-wide replication, and that ASARs function as essential RNA scaffolds for the assembly of hnRNP complexes that help maintain the structural integrity of each mammalian chromosome.

46 **Introduction:**

47 The vast majority of mammalian DNA replicates in homologous regions of
48 chromosome pairs in a highly synchronous manner [1-3]. However, genetic disruption of
49 non-protein coding ASAR ("ASynchronous replication and Autosomal RNA") genes
50 causes a delay in replication timing on individual human chromosomes in *cis*, resulting in
51 highly asynchronous replication between pairs of autosomes [4-7]. There are eight
52 genetically validated ASAR genes located on human chromosomes 1, 6, 8, 9 and 15 [4-7].
53 All of the known ASAR genes share several distinctive characteristics, including: 1)
54 contiguously transcribed regions of >180 kb, 2) epigenetically regulated Allelic Expression
55 Imbalance (AEI; including mono-allelic expression); 3) Variable Epigenetic Replication
56 Timing (VERT; including asynchronous replication timing); 4) high density of LINE1 (L1)
57 sequences (>30%); and 5) retention of the RNA within their parent chromosome territory.
58 Recently, we used these notable characteristics to identify 68 ASAR gene candidates on
59 every human autosome [5,6].

60 In addition to genetic disruption assays, ectopic integration of ASAR transgenes into
61 mouse chromosomes has been used as a second functional assay to help define the
62 critical sequences within ASAR genes that control chromosome-wide replication timing
63 [8,9]. For example, ASAR6 RNA expressed from transgenes, ranging in size from an ~180
64 kb BAC transgene to PCR products (as small as ~3 kb) containing the critical sequences,
65 remain associated with the chromosome territories where they are transcribed [8,9]. These
66 same transgenes cause delayed replication timing (DRT) and delayed mitotic
67 condensation (DMC) of entire chromosomes in *cis* [8,9]. Furthermore, ASAR6 RNA was
68 found to control chromosome-wide replication timing using oligonucleotide mediated RNA

69 degradation of ASAR6 RNA that was expressed from a BAC transgene integrated into
70 mouse chromosome 3 [9]. These findings suggest that ASARs represent ubiquitous,
71 essential, *cis*-acting elements that control mammalian chromosome replication timing via
72 expression of chromosome associated RNAs. In this report, we identify RNA binding
73 proteins (RBPs) that are critical for ASAR RNA localization and function.

74 Chromosome-associated long non-coding RNAs are known to be important for
75 numerous aspects of chromosome dynamics in diverse eukaryotes from yeast to
76 mammals [10-14]. The mammalian nucleus is rich in heterogenous nuclear RNA (hnRNA),
77 which represents >95% of the total RNA Polymerase II output [12,15]. The vast majority of
78 hnRNA is poorly defined, has been referred to as the "Dark Matter" of the genome [15], and
79 includes spliced and unspliced intronic sequences, circular RNAs [16], very long intergenic
80 non-coding RNAs (vlincRNAs; [17]), as well as poorly defined RNAs that include chromatin
81 associated RNAs (caRNAs; [18]) and repeat rich RNAs (repRNAs [19]) that can be detected
82 by C0T-1 DNA (which is comprised primarily of LINE and SINE repetitive sequences;
83 [13,20,21]). Some of these nuclear RNAs comprise abundant species, yet lack clearly defined
84 functions, while others are thought to be the by-products of other nuclear processes and
85 have been collectively called "RNA debris" [12]. The relationship between ASAR RNAs and
86 these other chromosome-associated RNAs remains poorly defined.

87 Previous studies have suggested that long nascent transcripts that are detected by
88 C0T-1 DNA play a dynamic structural role that promotes the open architecture of active
89 chromosome territories [13,20]. The RNA species detected by C0T-1 DNA are predominantly
90 L1 sequences, and these L1 RNAs remain associated with the chromosome territories
91 where they are transcribed [20]. A link between C0T-1 RNA and ASAR RNAs is suggested

92 by the observation that ASAR RNAs contain a high L1 content and remain associated
93 with their parent chromosome territories [4-7,9]. In addition, deletion and ectopic integration
94 assays demonstrated that the critical sequences for controlling chromosome-wide
95 replication timing within *ASAR6* and *ASAR15* map to the antisense strand of L1
96 sequences located within the ASAR6 and ASAR15 RNAs [9]. In this report we found that
97 human C0T-1 DNA can detect the chromosome territory localization of ASAR6 RNA that
98 is expressed from a BAC transgene integrated into a mouse chromosome, indicating that
99 at least some of the RNA species detected by C0T-1 DNA represents ASAR RNA.

100 Heterogeneous nuclear ribonucleoproteins (hnRNPs) represent a large family of
101 abundant RBPs that contribute to multiple aspects of nucleic acid metabolism. Many of
102 the proteins that are in hnRNP complexes share general features, but differ in domain
103 composition and functional properties (reviewed in [22]). The functions of hnRNPs vary
104 according to their subcellular localization. Most of the hnRNPs possess conventional
105 nuclear localization signals and are predominantly present in the nucleus during steady
106 state, where they function in various RNA metabolic processes including transcription,
107 splicing, export, and 3' end processing [22,23], and more recently have been implicated in
108 DNA replication [24]. A direct connection between chromosome territory associated C0T-1
109 RNA and hnRNPs was established by the observation that siRNA mediated depletion, or
110 forced expression of dominant interfering mutants, of HNRNPU results in dissociation of
111 C0T-1 RNA from chromosome territories [13,20,21]. In addition, recent studies have proposed
112 that abundant nuclear proteins such as HNRNPU nonspecifically interact with "RNA
113 debris" that creates a dynamic nuclear mesh that regulates interphase chromatin
114 structure [12,18]. Here, we used publicly available enhanced Cross Link Immuno-

115 Precipitation (eCLIP) data from ENCODE [25] to identify RBPs, including HNRNPA1,
116 HNRNPC, HNRNPL, HNRNPM, HNRNPU, and HNRNPUL1, that interact with the known
117 ASAR RNAs. We also found that an ~7 kb region within the ~185 kb ASAR6-141 RNA
118 has a surprisingly high density of interacting RBPs, and that genetic deletion of this ~7kb
119 "RBP domain" from the endogenous *ASAR6-141* gene results in a delayed replication
120 timing phenotype that is comparable to deletion of the entire ~185 kb *ASAR6-141* gene
121 located on human chromosome 6. In addition, ectopic integration of transgenes,
122 expressing the ~7 kb RBP domain, into autosomes or into the inactive X chromosome
123 results in retention of the RNA within the chromosome territory of the integrated
124 chromosomes, and results in DRT/DMC, and chromosome structure instability in *cis*.
125 Taken together, these results indicate that the ~7 kb RBP domain represents a critical
126 region within the *ASAR6-141* gene that controls replication timing of human chromosome
127 6.

128 Because ASARs were identified as *cis*-acting elements that control chromosome-wide
129 replication timing, we tested if ASAR associated RBPs also control replication timing. We
130 found that shRNA mediated depletion of 9 different ASAR-associated RBPs (HNRNPA1,
131 HNRNPC, HNRNPL, HNRNPM, HNRNPU, HNRNPUL1, HLTF, KHSRP or UCHL5)
132 dramatically altered the normally synchronous replication timing program of all autosome
133 pairs, recapitulating the effect of individual ASAR gene knockouts on a genome-wide
134 scale. These results demonstrate the role that ASARs play during the temporal order of
135 genome-wide replication, and indicate that ASAR RNAs serve as essential RNA scaffolds
136 for the assembly of multiple hnRNP/RBP complexes that function to maintain the
137 structural integrity of each mammalian chromosome.

138 **Results**

139 All of the eight genetically validated ASAR genes are subject to AEI and express
140 lncRNAs that remain associated with the chromosome territories where they are
141 transcribed [4-7]. Examples of mono-allelic expression and chromosome territory
142 localization of ASAR6-141 RNA are shown in Figure 1A and 1B, where ASAR6-141 RNA
143 is localized within one of the chromosome 6 territories in two different cell types, male
144 HTD114 cells and female GM12878 cells, respectively. We note that the size of the RNA
145 hybridization signals detected by all ASAR probes are variable, ranging in size from large
146 clouds that occupy entire chromosome territories, with similarity in appearance to XIST
147 RNA "clouds" (Fig. 1B), to relatively small sites of hybridization that remain tightly
148 associated with the expressed alleles [4-7].

149 We first sought to identify RBPs that interact with the known ASAR RNAs. Using
150 publicly available eCLIP data for 150 RBPs in K562 and HepG2 cells (ENCODE; [25,31]),
151 we identified RBP interactions within 4 ASAR genes (*ASAR1-187, ASAR6-141, ASAR8-*
152 *2.7,* and *ASAR9-23*). Utilizing the eCLIP peaks previously identified by ENCODE [25] we
153 found significant RBP binding peaks for 99 different RBPs that map within the four ASAR
154 RNAs (Table S1). We also utilized a "region-based" method [26] to identify significant
155 enrichments of eCLiP reads against matched controls in 10 kb windows across the four
156 ASAR genes (Table S2). An example of this analysis is shown in Figure 1C, where
157 *ASAR6-141* displays significant RBP interactions for 35 different RBPs across the
158 transcribed region. Given the contiguous expression (RNAseq from total RNA) across the
159 ~185 kb *ASAR6-141* gene in both K562 and HepG2 cell lines (Fig. 1G) and the abundant
160 eCLIP reads from 35 different RBPs in 10 kb windows across the *ASAR6-141* gene (Fig.

161 1C and 1D; see Table S2), we were surprised to detect an ~7 kb region that contained
162 virtually all of the significant eCLIP peaks in both cell lines (Fig.1G; and Table S1). The
163 zoomed in view in Figure 1G shows alignment of 58 independent eCLIP peaks of RBP
164 interactions within the ~7kb RNA Binding Protein Domain (RBPD). Figure 1E shows the
165 23 RBPs with the most abundant peaks of eCLIP reads within the ~7kb RBPD. The eCLIP
166 peaks associated with the RBPs chosen for shRNA knockdown experiments (see below)
167 are highlighted in the zoomed in view in Figure 1G.
168

169 **Disruption of the ~7kb RBPD within ASAR6-141 results in delayed replication.**

170 To determine if the ~7kb RBPD within the *ASAR6-141* gene contains replication timing
171 activity, we used CRISPR/Cas9 mediated engineering to delete the ~7 kb sequence from
172 the endogenous locus on human chromosome 6. For this analysis we designed single
173 guide RNAs (sgRNAs) to unique sequences on either side of the ~7kb RBPD (see Fig.
174 1G; and Fig. S1). We expressed these sgRNAs in combination with Cas9 in human
175 HTD114 cells and screened clones for deletion of the ~7kb RBPD (see Fig. S1). We
176 chose HTD114 cells for this analysis because they maintain mono-allelic expression and
177 asynchronous replication of both imprinted and random monoallelic genes, they have a
178 stable karyotype that does not change significantly following transfection, drug selection
179 and sub-cloning, and knockouts of the eight known ASAR genes, including *ASAR6-141*,
180 result in DRT/DMC and chromosome structure instability in *cis* [4-9,27,28]. Because *ASAR6-*
181 *141* expression is mono-allelic in the HTD114 cells (see Fig. 1A; and [5]), we isolated clones
182 that had heterozygous or homozygous deletions of the ~7kb RBPD. We determined which

183 alleles were deleted based on retention of different base pairs of a heterozygous SNP
184 located within the ~7kb RBPD (see Fig. S1 and Table S3).
185 To assay replication timing of homologous chromosome pairs we quantified DNA
186 synthesis in mitotic cells using a BrdU terminal label assay (Fig. 2A; and [29,30]). In addition,
187 the HTD114 cells contain a centromeric polymorphism on chromosome 6, which allows
188 for an unambiguous distinction between the two chromosome 6 homologs ([7-9]; also see
189 Fig. 2B). For simplicity, we refer to the chromosome 6 with the larger centromere as
190 CHR6A and to the chromosome 6 with the smaller centromere as CHR6B. From our
191 previous studies we knew that the chromosome 6 with the smaller centromere is linked
192 to the expressed allele of *ASAR6-141* [5]. Figure 2B and 2C shows the replication timing
193 analysis on a mitotic cell where the ~7kb RBPD was deleted from the expressed allele
194 (CHR6B). Note that CHR6B contains more BrdU incorporation than CHR6A, indicating
195 later replication of CHR6B (Fig. 2D). Quantification of the BrdU incorporation in CHR6B
196 and CHR6A in multiple cells indicated that deletion of the ~7kb RBPD from the expressed
197 allele resulted in a significant delay in replication timing (Fig. 2E). This is in contrast to
198 cells prior to deletion, or in cells containing a deletion of the ~7kb RBPD from the silent
199 allele (CHR6A), where the BrdU incorporation is comparable between CHR6A and
200 CHR6B (Fig. 2E). In addition, an asynchronous replication pattern was also present in
201 cells with homozygous deletions of the 7kb RBPD, with CHR6B representing the later
202 replicating allele (Fig. 2E). Furthermore, we found that the delayed replication associated
203 with deletion of the ~7kb RBPD on CHR6B is comparable to the delayed replication
204 associated with deletion of the entire *ASAR6-141* gene from the same chromosome
205 (CHR6B; Fig. 2E; also see [5]). These results indicate that the ~7kb RBPD contains

206 functional sequences for controlling replication timing, and further supports the conclusion
207 that the expressed allele but not the silent allele controls replication timing of human
208 chromosome 6 [5].
209
210 **Ectopic integration of transgenes expressing the ~7kb RBPD into autosomes.**
211 To determine if the ~7 kb RBPD is sufficient to induce delayed replication in an
212 ectopic integration assay, we generated two transgenes containing the ~7kb RBPD, one
213 in the sense orientation and one in the antisense orientation with respect to the promoter
214 (Fig. 2F). These transgenes were introduced into mouse cells and individual clones
215 isolated. We initially screened individual clones for expression of the transgenes using
216 RNA FISH with the ~7kb RBPD as probe. Because the mouse cells that we used are
217 female, we included an RNA FISH probe for Xist to serve as positive control for the RNA
218 FISH assays. This two-color RNA FISH assay allowed us to identify clones with single
219 sites of transgene expression, and allowed us to directly compare the RNA FISH signals
220 from the transgenes to that of the endogenous Xist RNA expressed from the inactive X
221 chromosome. Figure 2 shows examples of the RNA FISH signals in cells expressing
222 either the sense (Fig. 2G) or antisense (Fig. 2H) transgenes. Large clouds of RNA FISH
223 hybridization signals that are comparable to Xist RNA hybridization signals were detected
224 in 4 out of 12 clones expressing the sense transgene (Fig. 2G). In contrast, 0 out of 12
225 clones containing the antisense transgene expressed RNA FISH hybridization signals
226 comparable to Xist RNA signals, but instead only small pinpoint sites of RNA FISH
227 hybridization were detected (Fig. 2H). Quantitation of the nuclear area occupied by the
228 sense and antisense ~7 kb RNAs is shown in Figure 2I. Note that the mouse cells are

229 tetraploid for the X chromosome and contain two Xist RNA hybridization signals,
230 representing two inactive X chromosomes.

231 To identify the mouse chromosomes that contain the transgenes, we used DNA FISH
232 using the ~7kb RBPD DNA as probe in combination with the inverted DAPI banding
233 pattern to identify the mouse chromosomes in metaphase spreads. We then confirmed
234 the identity of the mouse chromosomes using DNA FISH on metaphase spreads using
235 the ~7kb RBPD in combination with BAC probes from the chromosomes of interest. For
236 example, clone ~7kb(+)A5, which contains the sense-strand transgene integrated in
237 mouse chromosome 5 is shown in Figure S2A. Next, to determine if the RNA expressed
238 from the ~7kb RBPD sense strand transgene is retained within the chromosome territory
239 that expresses the transgene, we used RNA-DNA FISH with the ~7kb RBPD as RNA
240 probe plus a chromosome paint to detect mouse chromosome 5 DNA in clone ~7kb(+)A5.
241 Figure 2I shows that the ~7kb RBPD RNA is localized within one of the chromosome 5
242 DNA hybridization signals, indicating that the ~7kb RBPD RNA is retained within the
243 chromosome 5 territory.

244 Next to determine if the ~7kb RBPD transgene alters replication timing of mouse
245 chromosome 5, we visualized DNA synthesis in mitotic cells using the BrdU terminal label
246 assay (see Fig. 2A; and [29,30]). Cells were exposed to BrdU for five hours, harvested for
247 mitotic cells, and processed for BrdU incorporation and for DNA FISH using the ~7kb
248 RBPD transgene to identify the integrated chromosome. Figure 2J shows an example of
249 this replication-timing assay in a mitotic cell containing the ~7kb RBPD sense strand
250 transgene integrated into mouse chromosome 5. Note that the chromosome containing
251 the ~7kb RBPD transgene is delayed in mitotic condensation and contains more BrdU

252 incorporation than any other chromosome within the same cell. In contrast, integration of
253 the transgene with the ~7kb RBPD in the antisense orientation did not result in delayed
254 condensation nor in delayed replication of mouse chromosomes (Fig. 2K).
255
256 **Depletion of RBPs disrupts the chromosome territory localization of ASAR RNAs.**
257 The results described above identified 99 RBPs expressed in two different cell lines
258 that interact with four different ASAR RNAs (see Table S1). Next, to determine if the
259 ASAR associated RBPs function in the chromosome territory localization of ASAR RNAs,
260 we analyzed ASAR RNA localization in cells with shRNA depletion of 14 different RBPs.
261 Figure 3A-J shows examples of antibody staining for HNRNPU, HNRNPUL1, HNRNPC,
262 HNRNPM, and HLTF in K562 cells expressing either empty vector or vectors expressing
263 shRNAs to individual RBPs. Quantitation of the immunofluorescence for each RBP in
264 empty vector versus shRNA expressing cells, indicated a significant depletion of each
265 RBP (Fig. 3K; also see Fig. S3 for Western blots).
266 Next, to determine if the nuclear localization of ASAR6-141 RNA was affected by RBP
267 depletions we used RNA FISH assays in K562 cells expressing shRNAs against the 14
268 RBPs. For this analysis we used RNA FISH to detected ASAR6-141 RNA, and because
269 K562 cells are female we also included a probe for XIST RNA as control for the RNA
270 FISH. Figure 3L shows the expected RNA FISH patterns for both ASAR6-141 and XIST
271 RNAs in cells expressing empty vector. In contrast, cells expressing shRNAs directed at
272 HNRNPU, HNRNPUL1, HNRNPC, HNRNPM, HLTF, HNRNPA1, HNRNPL, KHSRP, or
273 UCHL5 showed a nearly complete absence of the highly localized nuclear RNA
274 hybridization signals for ASAR6-141 RNA, and the appearance of large punctate

275 cytoplasmic hybridization signals (Fig. 3M-U). We also noticed an overall increase in a
276 diffuse ASAR6-141 RNA nuclear and cytoplasmic hybridization signals. In contrast, cells
277 expressing shRNAs directed at PTBP1, PTBP2, or MATR3 showed only the typical
278 nuclear localized hybridization signals for both ASAR6-141 and XIST RNAs (Fig. 3V; and
279 Fig. S4A-D). In addition, localized nuclear XIST RNA hybridization signals were detected
280 in all shRNA treated cells (Fig. 3M-V and Fig. S4A-D). However, we detected cytoplasmic
281 foci of XIST RNA hybridization that colocalized with the ASAR6-141 RNA, and a
282 significant decrease in the size of the nuclear XIST RNA hybridization signals (Fig. 3W).
283 These observations indicate that at least some of the XIST RNA was dissociated from the
284 inactive X chromosome and colocalized with ASAR6-141 RNA in large cytoplasmic foci.
285

286 **Ectopic integration of the ~7kb RBPD transgene into the inactive X chromosome.**

287 The results described above indicate that shRNA depletion of 9 different RBPs resulted
288 in dissociation of ASAR6-141 RNA from its parent chromosome, and resulted in a
289 decreased nuclear XIST RNA hybridization signal with concomitant colocalization of both
290 RNAs in cytoplasmic foci (see Fig. 3), suggesting that ASAR6-141 and XIST RNAs are
291 retained on their respective chromosome territories via similar mechanisms, which is
292 supported by the large overlap in RBPs associated with both RNAs (Fig. 1E and 1F; and
293 Table S1). In addition, ectopic integration of *Xist* transgenes into autosomes has been an
294 instrumental tool in characterization of *Xist* functions, including the ability to delay
295 replication timing and induce gene silencing [31-33]. Therefore, we next sought to determine
296 the consequences of ectopic integration of the ASAR6-141 ~7kb RBPD transgene into
297 the inactive X chromosome. The ectopic integration assays described above involved

298 random integration of transgenes followed by an RNA FISH screen for clones expressing
299 the ~7kb RBPD, which included an RNA FISH probe for Xist RNA as positive control.
300 Therefore, to identify transgene integrations into the inactive X chromosome, we simply
301 screened an additional 50 clones for colocalization of RNA FISH hybridization signals for
302 both Xist and the ~7kb RBPD transgene. Figure 4A shows this RNA FISH assay on cells
303 where the Xist and ~7kb RBPD hybridization signals are colocalized. We note that the
304 intensity of the Xist RNA FISH signal was enhanced in the region where the ~7kb RBPD
305 hybridization signal was detected (Fig. 4B), suggesting that the presence of the ~7kb
306 RBPD RNA promoted preferential localization of Xist RNA in the same region of the X
307 chromosome territory that contains the ~7kb RBPD RNA.

308 To confirm that the mouse X chromosome contains the transgene, we used DNA FISH
309 using the ~7kb RBPD DNA as probe in combination with an X chromosome paint probe
310 on metaphase spreads. An example of this analysis is shown in Figure S2B and S2C,
311 and indicates that the ~7kb RBPD transgene is integrated into an X chromosome. During
312 this analysis we detected DMC on X chromosomes, and the X chromosomes with DMC
313 contain the ~7kb RBPD transgene (Figs. 4C and 4D).

314 Next, to determine if the ~7kb RBPD transgene alters replication timing of the inactive
315 X chromosome, we visualized DNA synthesis in mitotic cells using our BrdU terminal label
316 assay (see Fig. 2A; and [29,30]). Cells were exposed to BrdU for five hours, harvested for
317 mitotic cells, and processed for BrdU incorporation and for FISH using the ~7kb RBPD
318 transgene plus an X chromosome paint probe to identify the integrated chromosome.
319 Figure 4E-4H shows an example of delayed replication and delayed mitotic condensation
320 of the X chromosome in a mitotic cell containing the ~7kb RBPD sense transgene

321 integrated into the inactive X chromosome. We conclude that integration of the 7kb RBPD
322 transgene is sufficient to cause DRT/DMC on the inactive X chromosome.
323
324 **Disruption of the chromosome associated localization of ASAR6 and ASAR1-187**
325 **RNAs.**
326 To determine if the chromosomal localization of additional ASAR RNAs was also
327 affected by RBP knockdowns we carried out RNA-DNA FISH for ASAR1-187 and ASAR6
328 RNAs. For this analysis we used Fosmid probes to detect the ASAR RNAs plus BAC
329 probes to detect the DNA near the ASAR genes located on chromosomes 1 and 6 (see
330 Table S3). Figure 5A-5C shows this RNA FISH assay on K562 cells expressing empty
331 vector, and indicates that the RNA FISH signals for both ASARs are closely associated
332 with the DNA FISH signals on their respective chromosomes, i.e. ASAR1-187 RNA shows
333 close association with the chromosome 1 BAC DNA, and ASAR6 RNA shows close
334 association with the chromosome 6 BAC DNA. In contrast, cells with shRNA depletion of
335 HNRNPA1, HNRNPC, HNRNPL, HNRNPM, HNRNPU, HLTF, KHSRP, or UCHL5
336 showed a dispersed punctate nuclear pattern for both ASAR RNAs (Fig. 5D-5K). In
337 addition, large punctate cytoplasmic foci were detected, with the two ASAR RNA
338 hybridization signals colocalized.
339 Next, to determine if depletion of RBPs would result in loss of the ASAR RNA
340 chromosome localization in a second cell type, we carried out RNA-DNA FISH assays on
341 shRNA depleted HTD114 cells. We chose HTD114 cells because we previously used
342 these cells to detect the chromosome territory localization of all eight ASAR RNAs and to
343 measure the synchronous replication timing of five autosome pairs both before and after

344 ASAR gene disruptions ([4-8,28]; and see Figure 2). Figure 5L-5N shows the RNA-DNA FISH
345 assay in HTD114 cells expressing empty vector. We again note the tight association of
346 the ASAR RNAs with the genomic loci where they are transcribed. In contrast, cells
347 expressing shRNA against HNRNPU show loss of the chromosome associated RNA
348 hybridization signals for both ASAR RNAs, and instead show diffuse nuclear and
349 cytoplasmic hybridization, with the presence of large punctate cytoplasmic foci that
350 hybridize to both ASAR RNA FISH probes (Fig. 5O-5Q). Similar results were obtained
351 following shRNA depletion of HNRNPA1, HNRNPC, HNRNPL, HNRNPM, HNRNPUL1,
352 KHSRP, HLTF, or UCHL5 (Fig. S5). In contrast, expression of shRNAs against SAFB,
353 SAFB2, SAFB plus SAFB2, CTCF, or MATR3 did not alter the ASAR1-187 or ASAR6
354 RNA nuclear hybridization signals in HTD114 cells (Fig. S4E-S4J).
355
356 **Depletion of RBPs causes asynchronous replication on all autosome pairs.**
357 Next, to determine if ASAR associated RBPs function during the normally synchronous
358 replication timing program that occurs on pairs of homologous chromosomes, we
359 analyzed DNA synthesis in HTD114 cells depleted for the 9 RBPs implicated in ASAR
360 RNA localization (HNRNPU, HNRNPA1, HNRNPC, HNRNPL, HNRNPM, HNRNPUL1,
361 KHSRP, HLTF, or UCHL5) and for five RBPs that showed no effect on ASAR RNA
362 localization (MATR3, PTBP1, PTBP2, SAFB, or SAFB2). Using the BrdU terminal label
363 assay [29,30] in cells expressing the empty vector, we found that each autosome pair
364 displays a stereotypical and synchronous BrdU incorporation pattern (Fig. 6A). In
365 contrast, following HNRNPU depletion, individual mitotic spreads displayed dramatic
366 differential incorporation of BrdU into pairs of homologous chromosomes, examples of

367 asynchronous replication timing between every pair of autosomes are shown in Figure
368 6B. Note that each chromosome pair contains dramatic differential BrdU incorporation
369 into chromosome homologs, which represents considerable asynchronous replication
370 timing. Figure 6C and 6D show examples for chromosome pairs 4 and 18 (also see Fig.
371 S6C). Quantification of the BrdU incorporation in chromosomes 1 and 6 (see Fig. S6) in
372 multiple cells depleted for HNRNPU, HNRNPUL1, HLTF, KHSRP, UCHL5, or HNRNPA1
373 indicated that both pairs of chromosomes display significant asynchrony (Fig. 6E and 6F).
374 In addition, the DRT/DMC phenotype was detected in ~30% of mitotic spreads following
375 shRNA-mediated depletions (two examples are shown in Fig. 6G and 6H; also see Fig.
376 S7). Furthermore, shRNA-mediated depletion of HNRNPC, HNRNPL and HNRNPM
377 resulted in dramatic alterations in chromosome morphology that included asynchronous
378 replication and abnormal mitotic condensation, including ~10% of mitotic spreads with
379 one or more chromosomes that displayed the DRT/DMC phenotype (Fig. S7).
380 Next, to determine the relationship between the RNA species detected by C0T-1 DNA
381 and ASAR RNAs in RNA FISH assays, we used human C0T-1 DNA as an RNA FISH
382 probe on mouse cells expressing an *ASAR6* BAC transgene [8]. Because human C0T-1
383 DNA does not cross hybridize with mouse RNA or DNA in FISH assays [20], we were able
384 to detect the chromosome territory localized expression of human ASAR6 RNA
385 expressed from the transgene that is integrated into one of the mouse chromosome 3
386 homologs using human C0T-1 DNA as RNA FISH probe (Fig. 7A and 7B), indicating that
387 C0T-1 DNA detects ASAR6 RNA.
388
389

390 **Discussion:**

391 We recently identified 68 ASAR candidate genes, and genetically validated five out of five
392 as controlling replication timing of individual human autosomes, bringing the total of
393 genetically validated ASAR genes to eight [4-7]. These eight ASAR genes (*ASAR1-187,*
394 *ASAR6, ASAR6-141, ASAR8-2.7, ASAR9-23, ASAR9-24, ASAR9-30 and ASAR15*)
395 display striking epigenetically controlled differential allelic expression and replication
396 timing, and their RNAs remain localized to their parent chromosome territories [4-7]. In this
397 report, we used publicly available RNA-protein interaction data from ENCODE to identify
398 RBPs that interact with multiple ASAR RNAs. One unanticipated observation that came
399 from this analysis was that an ~7 kb domain within the ~185 kb ASAR6-141 RNA
400 contained virtually all of the significant peaks of RBP interactions in two different cell lines.
401 Using genetic deletion and ectopic integration assays we found that this ~7kb RBPD
402 contains functional sequences for controlling replication timing of entire chromosomes in
403 *cis*. We also found that shRNA-mediated depletion of 9 different ASAR-associated RBPs
404 results in loss of the chromosome territory association of multiple ASAR RNAs. In
405 addition, we found that depletion of HNRNPA1, HNRNPC, HNRNPL, HNRNPM,
406 HNRNPU, HNRNPUL1, HLTF, KHSRP, or UCHL5 alters the normally synchronous
407 replication timing that occurs between pairs of homologous chromosomes resulting in the
408 DRT/DMC phenotype, recapitulating the effect of individual ASAR gene knockouts on a
409 genome-wide scale.

410 Previous studies have suggested that C0T-1 RNA plays a dynamic structural role that
411 promotes the open architecture of active chromosome territories [13,20]. C0T-1 RNA is
412 predominantly L1 sequences, and these L1 RNAs remain associated with the

413 chromosome territories where they are transcribed [20]. Additional studies have proposed
414 that repRNA is an important regulator of the dynamic organization of genomic loci into
415 membrane-less subcompartments with distinct nuclear functions (reviewed in [19]). Other
416 studies have proposed that HNRNPU nonspecifically interacts with "RNA debris" to create
417 a dynamic nuclear mesh that regulates interphase chromatin structure [12,18]. One
418 important distinction that can be made between ASAR RNAs and these other hnRNAs
419 (i.e. C0T-1 RNA, repRNA, caRNA, and "RNA debris" [12,13,18-21]) is that ASARs have been
420 defined genetically, and these other RNAs have not. Thus, we have used both loss of
421 function and gain of function assays to define the roles that ASAR genes play in
422 chromosome dynamics, which includes control of chromosome-wide replication timing
423 and mitotic chromosome condensation [4-8]. We also found that C0T-1 DNA, when used as
424 an RNA FISH probe, can detect ASAR6 RNA that is expressed and localized to an
425 individual chromosome territory, indicating that at least some of the RNA FISH signal
426 detected by C0T-1 DNA represents ASAR RNA. Given our recent discovery of ASAR
427 RNAs expressed from every autosome, and encoded by ~3% of the human genome [6],
428 we propose that ASAR RNAs represent functional "chromosome associated RNA"
429 species that control chromosome dynamics within mammalian nuclei. Whether or not
430 there are nuclear functions of other chromosome associated hnRNAs that are
431 independent of ASARs will require genetic and/or functional validation.

432 Here we identified several hnRNPs (HNRNPA1, HNRNPC, HNRNPL, HNRNPM,
433 HNRNPU, and HNRNPUL1) as important interaction partners for the chromosome
434 territory localization of multiple ASAR RNAs. In addition, we found that the SWI/SNF
435 related transcription factor HLTF, which was recently shown to regulate replication fork

436 reversal, prevent alternative mechanisms of stress-resistant DNA replication, and
437 mediates cellular resistance to replication stress [34], is required for the chromosome
438 territory localization of multiple ASAR RNAs and for the synchronous replication of every
439 autosome pair. Furthermore, we found that the KH domain (HNRNPK Homology domain)
440 containing protein KHSRP, and the ubiquitin hydrolase UCHL5 also contribute to the
441 chromosome territory localization of ASAR RNAs. We found that depletion of HNRNPA1,
442 HNRNPC, HNRNPL, HNRNPM, HNRNPU, HNRNPUL1, HLTF, KHSRP, or UCHL5
443 proteins resulted in a dramatic genome-wide disruption of the normally synchronous
444 replication timing program on all autosome pairs. Taken together our results indicate that
445 the association of ASAR RNAs with multiple hnRNPs in combination with other nuclear
446 RBPs is an essential step in the regulation of the synchronous replication timing program
447 on mammalian chromosomes.

448 While these studies have demonstrated a role for RBPs in genome-wide chromosome
449 replication timing, one limitation of these studies is that we only tested 12 out of the ~100
450 RBPs shown by eCLIP to associate with multiple ASAR RNAs. It is likely that additional
451 insights into the network of RBP/ASAR interactions could be gained by extending our
452 replication timing analyses to knockdowns of additional RBPs. Another limitation of these
453 studies is that all of the RBPs implicated in control of replication timing are known to play
454 various roles in essential nucleic acid metabolic pathways, including DNA replication,
455 transcription, splicing, mRNA export, and RNA degradation [22,23]. It is possible that
456 knocking down these RBPs has an indirect effect on replication related to these other
457 functions of the RBPs, rather than a direct effect on ASAR function. However, the RBP
458 knockdowns that disrupt genome-wide synchronous replication also disrupt chromosome

459 association of multiple ASAR RNAs. Our previous work demonstrated that the ability of
460 an ectopically integrated ASAR6 transgene to cause delayed replication timing of a
461 mouse chromosome required the ASAR6 RNA [9]. Hence, we expect that displacement of
462 ASAR RNA from chromosome territories upon RBP knockdown is a primary cause of the
463 genome-wide replication timing asynchrony, regardless of whether the displacement of
464 ASAR RNA from the chromosome is a direct or indirect effect of RBP knockdown. A final
465 limitation of these studies is that the precise mechanism by which ASARs and RBPs
466 interact to control replication timing remains unknown. One intriguing possibility that is
467 consistent with our data is that the ASAR RNAs serve as scaffolds for the assembly of
468 extensive RNA-protein and protein-protein complexes that form a chromosome-tethered
469 liquid-liquid phase that facilitates the recruitment of the requisite chromatin modifiers and
470 DNA replication machinery. In support of this notion, HNRNPA1 and HNRNPC are
471 components of the 40S hnRNP particle, which forms a nuclear biomolecular condensate
472 [23]. In addition, HNRNPU and HNRNPUL1 interact with each of the 40S hnRNP particle
473 subunits [35], suggesting the possibility of extensive protein-protein interactions. ASAR
474 RNAs normally remain associated exclusively with the chromosome territory from which
475 they are transcribed, and ASAR RNAs on different chromosomes remain separate. RBP
476 knockdowns that disrupt synchronous replication timing also release ASAR RNAs from
477 their chromosome territories, allowing colocalization of ASAR RNAs from different
478 chromosomes (i.e. ASAR6 and ASAR1-187, see Fig. 5O-Q), often in nuclear and/or
479 cytoplasmic foci reminiscent of the mixing of liquid-liquid phases. Additional support for
480 this model is the observation that Xist RNA participates in the assembly of a nuclear
481 heteromeric condensate essential for gene silencing on the inactive X chromosome [36].

482 Our observation that the ~7kb RBPD RNA induces DRT/DMC on the inactive X
483 chromosome, and associates with the Xist RNA cloud (Fig. 4A and 4B), suggests that the
484 ~7kb RBPD RNA interferes with Xist RNA function within the heteromeric condensate.
485 The formation of liquid-liquid phases via ASAR RNA-RBP interactions is speculative at
486 this point and remains to be tested directly. A full understanding of the mechanisms by
487 which ASARs control chromosome replication timing will require an extensive biochemical
488 and biophysical characterization of the ASAR RNA-protein and protein-protein complexes
489 and how they interact with the DNA replication machinery.

References:

1. Mukhopadhyay, R., Lajugie, J., Fourel, N., Selzer, A., Schizas, M., Bartholdy, B., Mar, J., Lin, C.M., Martin, M.M., Ryan, M., et al. (2014). Allele-specific genome-wide profiling in human primary erythroblasts reveal replication program organization. PLoS Genet *10*, e1004319. 10.1371/journal.pgen.1004319.
2. Dileep, V., and Gilbert, D.M. (2018). Single-cell replication profiling to measure stochastic variation in mammalian replication timing. Nature communications *9*, 427. 10.1038/s41467-017-02800-w.
3. Rivera-Mulia, J.C., Dimond, A., Vera, D., Trevilla-Garcia, C., Sasaki, T., Zimmerman, J., Dupont, C., Gribnau, J., Fraser, P., and Gilbert, D.M. (2018). Allele-specific control of replication timing and genome organization during development. Genome Res *28*, 800-811. 10.1101/gr.232561.117.
4. Donley, N., Smith, L., and Thayer, M.J. (2015). ASAR15, A cis-Acting Locus that Controls Chromosome-Wide Replication Timing and Stability of Human Chromosome 15. PLoS Genet *11*, e1004923. 10.1371/journal.pgen.1004923.
5. Heskett, M.B., Smith, L.G., Spellman, P., and Thayer, M.J. (2020). Reciprocal monoallelic expression of ASAR lncRNA genes controls replication timing of human chromosome 6. RNA (New York, N.Y.) *26*, 724-738. 10.1261/rna.073114.119.
6. Heskett, M.B., Vouzas, A.E., Smith, L.G., Yates, P.A., Boniface, C., Bouhassira, E.E., Spellman, P.T., Gilbert, D.M., and Thayer, M.J. (2022). Epigenetic control of chromosome-associated lncRNA genes essential for replication and stability. Nature communications *13*, 6301. 10.1038/s41467-022-34099-7.
7. Stoffregen, E.P., Donley, N., Stauffer, D., Smith, L., and Thayer, M.J. (2011). An autosomal locus that controls chromosome-wide replication timing and mono-allelic expression. Hum Mol Genet *20*, 2366-2378.
8. Donley, N., Stoffregen, E.P., Smith, L., Montagna, C., and Thayer, M.J. (2013). Asynchronous Replication, Mono-Allelic Expression, and Long Range Cis-Effects of ASAR6. PLoS Genet *9*, e1003423.
9. Platt, E.J., Smith, L., and Thayer, M.J. (2018). L1 retrotransposon antisense RNA within ASAR lncRNAs controls chromosome-wide replication timing. J Cell Biol *217*, 541-553. 10.1083/jcb.201707082.
10. Ding, D.Q., Okamasa, K., Katou, Y., Oya, E., Nakayama, J.I., Chikashige, Y., Shirahige, K., Haraguchi, T., and Hiraoka, Y. (2019). Chromosome-associated RNA-protein complexes promote pairing of homologous chromosomes during meiosis in Schizosaccharomyces pombe. Nature communications *10*, 5598. 10.1038/s41467-019-13609-0.
11. Faber, M.W., and Vo, T.V. (2022). Long RNA-Mediated Chromatin Regulation in Fission Yeast and Mammals. Int J Mol Sci *23*. 10.3390/ijms23020968.
12. Nozawa, R.S., and Gilbert, N. (2019). RNA: Nuclear Glue for Folding the Genome. Trends Cell Biol *29*, 201-211. 10.1016/j.tcb.2018.12.003.
13. Creamer, K.M., Kolpa, H.J., and Lawrence, J.B. (2021). Nascent RNA scaffolds contribute to chromosome territory architecture and counter chromatin compaction. Mol Cell. 10.1016/j.molcel.2021.07.004.

534 14. Menon, D.U., and Meller, V.H. (2015). Identification of the Drosophila X
535 chromosome: The long and short of it. RNA biology *12*, 1088-1093.
536 10.1080/15476286.2015.1086864.
537 15. St Laurent, G., Savva, Y.A., and Kapranov, P. (2012). Dark matter RNA: an
538 intelligent scaffold for the dynamic regulation of the nuclear information
539 landscape. Frontiers in genetics *3*, 57. 10.3389/fgene.2012.00057.
540 16. Kristensen, L.S., Andersen, M.S., Stagsted, L.V.W., Ebbesen, K.K., Hansen,
541 T.B., and Kjems, J. (2019). The biogenesis, biology and characterization of
542 circular RNAs. Nat Rev Genet *20*, 675-691. 10.1038/s41576-019-0158-7.
543 17. St Laurent, G., Vyatkin, Y., Antonets, D., Ri, M., Qi, Y., Saik, O., Shtokalo, D., de
544 Hoon, M.J., Kawaji, H., Itoh, M., et al. (2016). Functional annotation of the vlinc
545 class of non-coding RNAs using systems biology approach. Nucleic Acids
546 Research. 2 *44*, 3233-3252. 10.1093/nar/gkw162.
547 18. Nozawa, R.S., Boteva, L., Soares, D.C., Naughton, C., Dun, A.R., Buckle, A.,
548 Ramsahoye, B., Bruton, P.C., Saleeb, R.S., Arnedo, M., et al. (2017). SAF-A
549 Regulates Interphase Chromosome Structure through Oligomerization with
550 Chromatin-Associated RNAs. Cell *169*, 1214-1227.e1218.
551 10.1016/j.cell.2017.05.029.
552 19. Frank, L., and Rippe, K. (2020). Repetitive RNAs as Regulators of Chromatin-
553 Associated Subcompartment Formation by Phase Separation. J Mol Biol *432*,
554 4270-4286. 10.1016/j.jmb.2020.04.015.
555 20. Hall, L.L., Carone, D.M., Gomez, A.V., Kolpa, H.J., Byron, M., Mehta, N.,
556 Fackelmayer, F.O., and Lawrence, J.B. (2014). Stable C0T-1 repeat RNA is
557 abundant and is associated with euchromatic interphase chromosomes. Cell *156*,
558 907-919. 10.1016/j.cell.2014.01.042.
559 21. Kolpa, H.J., Creamer, K.M., Hall, L.L., and Lawrence, J.B. (2021). SAF-A
560 mutants disrupt chromatin structure through dominant negative effects on RNAs
561 associated with chromatin. Mamm Genome. 10.1007/s00335-021-09935-8.
562 22. Geuens, T., Bouhy, D., and Timmerman, V. (2016). The hnRNP family: insights
563 into their role in health and disease. Hum Genet *135*, 851-867. 10.1007/s00439-
564 016-1683-5.
565 23. Domanski, M., Dedic, E., Pérez, M.E., Cléry, A., Campagne, S., Uldry, A.C.,
566 Braga, S., Heller, M., Rabl, J., Afanasyev, P., et al. (2022). 40S hnRNP particles
567 are a novel class of nuclear biomolecular condensates. Nucleic Acids Research.
568 2 *50*, 6300-6312. 10.1093/nar/gkac457.
569 24. Connolly, C., Takahashi, S., Miura, H., Hiratani, I., Gilbert, N., Donaldson, A.D.,
570 and Hiraga, S.I. (2022). SAF-A promotes origin licensing and replication fork
571 progression to ensure robust DNA replication. J Cell Sci *135*.
572 10.1242/jcs.258991.
573 25. Moore, J.E., Purcaro, M.J., Pratt, H.E., Epstein, C.B., Shoresh, N., Adrian, J.,
574 Kawli, T., Davis, C.A., Dobin, A., Kaul, R., et al. (2020). Expanded
575 encyclopaedias of DNA elements in the human and mouse genomes. Nature
576 *583*, 699-710. 10.1038/s41586-020-2493-4.
577 26. Van Nostrand, E.L., Pratt, G.A., Shishkin, A.A., Gelboin-Burkhart, C., Fang, M.Y.,
578 Sundararaman, B., Blue, S.M., Nguyen, T.B., Surka, C., Elkins, K., et al. (2016).

579 Robust transcriptome-wide discovery of RNA-binding protein binding sites with
580 enhanced CLIP (eCLIP). Nat Methods *13*, 508-514. 10.1038/nmeth.3810.
581 27. Chang, B.H., Smith, L., Huang, J., and Thayer, M. (2007). Chromosomes with
582 delayed replication timing lead to checkpoint activation, delayed recruitment of
583 Aurora B and chromosome instability. Oncogene *26*, 1852-1861.
584 28. Breger, K.S., Smith, L., and Thayer, M.J. (2005). Engineering translocations with
585 delayed replication: evidence for cis control of chromosome replication timing.
586 Hum Mol Genet *14*, 2813-2827.
587 29. Smith, L., Plug, A., and Thayer, M. (2001). Delayed Replication Timing Leads to
588 Delayed Mitotic Chromosome Condensation and Chromosomal Instability of
589 Chromosome Translocations. Proc Natl Acad Sci U S A *98*, 13300-13305.
590 30. Smith, L., and Thayer, M. (2012). Chromosome replication timing combined with
591 fluorescent in situ hybridization. J Vis Exp *10*, e4400.
592 31. Minks, J., and Brown, C.J. (2009). Getting to the center of X-chromosome
593 inactivation: the role of transgenes. Biochem Cell Biol *87*, 759-766.
594 32. Pintacuda, G., Wei, G., Roustan, C., Kirmizitas, B.A., Solcan, N., Cerase, A.,
595 Castello, A., Mohammed, S., Moindrot, B., Nesterova, T.B., and Brockdorff, N.
596 (2017). hnRNPK Recruits PCGF3/5-PRC1 to the Xist RNA B-Repeat to Establish
597 Polycomb-Mediated Chromosomal Silencing. Mol Cell *68*, 955-969.e910.
598 10.1016/j.molcel.2017.11.013.
599 33. Lee, J.T., Strauss, W.M., Dausman, J.A., and Jaenisch, R. (1996). A 450 kb
600 transgene displays properties of the mammalian X-inactivation center. Cell *86*,
601 83-94.
602 34. Bai, G., Kermi, C., Stoy, H., Schiltz, C.J., Bacal, J., Zaino, A.M., Hadden, M.K.,
603 Eichman, B.F., Lopes, M., and Cimprich, K.A. (2020). HLTF Promotes Fork
604 Reversal, Limiting Replication Stress Resistance and Preventing Multiple
605 Mechanisms of Unrestrained DNA Synthesis. Mol Cell *78*, 1237-1251.e1237.
606 10.1016/j.molcel.2020.04.031.
607 35. Britton, S., Dernoncourt, E., Delteil, C., Froment, C., Schiltz, O., Salles, B., Frit,
608 P., and Calsou, P. (2014). DNA damage triggers SAF-A and RNA biogenesis
609 factors exclusion from chromatin coupled to R-loops removal. Nucleic Acids
610 Research. 2 *42*, 9047-9062. 10.1093/nar/gku601.
611 36. Pandya-Jones, A., Markaki, Y., Serizay, J., Chitiashvili, T., Mancia Leon, W.R.,
612 Damianov, A., Chronis, C., Papp, B., Chen, C.K., McKee, R., et al. (2020). A
613 protein assembly mediates Xist localization and gene silencing. Nature *587*, 145-
614 151. 10.1038/s41586-020-2703-0.
615 37. Kruskal, J.B. (1964). Multidimensional scaling by optimizing goodness of fit to a
616 nonmetric hypothesis. Psychometrika *29*, 1–27.
617 38. Zhu, Y., Bye, S., Stambrook, P.J., and Tischfield, J.A. (1994). Single-base
618 deletion induced by benzo[a]pyrene diol epoxide at the adenine
619 phosphoribosyltransferase locus in human fibrosarcoma cell lines. Mutat Res
620 *321*, 73-79.
621 39. Van Nostrand, E.L., Freese, P., Pratt, G.A., Wang, X., Wei, X., Xiao, R., Blue,
622 S.M., Chen, J.Y., Cody, N.A.L., Dominguez, D., et al. (2020). A large-scale
623 binding and functional map of human RNA-binding proteins. Nature *583*, 711-
624 719. 10.1038/s41586-020-2077-3.

40. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics (Oxford, England) *29*, 15-21. 10.1093/bioinformatics/bts635.
41. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics (Oxford, England) *25*, 2078-2079. 10.1093/bioinformatics/btp352.

Figure legends

Figure 1. RNA binding proteins interact with an ~7 kb domain within ASAR6-141 RNA. A) RNA-DNA FISH images of ASAR6-141 expression in six individual HTD114 cells. ASAR6-141 RNA (green; green arrows) and chromosome 6 DNA (chromosome paint, red), and DNA was stained with DAPI. B) RNA-DNA FISH of ASAR6-141 RNA (red, red arrows), CHR6 DNA (chromosome paint, white), and XIST RNA (green, green arrows) visualized within individual GM12878 cells, top and bottom panels represent the same three cells with the nuclear outline drawn in white. C) RNA-protein interaction data (eCLIP from ENCODE) for 120 RBPs expressed in K562 cells for the genomic region that contains ASAR6-141, and ~200 kb of upstream non-transcribed DNA (chromosome 6 140.8 mb-141.6 mb). Each point represents the FDR-corrected p-value of the z-score of log2-ratio between eCLiP vs control for each RBP within a 10kb sliding window. D) Number of 10kb regions with significant enrichments of eCLiP reads vs control within ASAR6-141. E) Histogram of eCLiP peaks count per RBP within ASAR6-141 7kb RBPD. F) Histogram of eCLiP peaks count per RBP within lncRNA XIST. G) Genome browser view of *ASAR6-141* with RNA-seq expression and eCLiP peaks shown in K562 and HepG2. The zoomed in view shows the ~7kb RBPD with the location of the peaks of eCLIP reads that map within the region (see Table S1). The eCLIP peaks for the RBPs used in the shRNA knockdown experiments are indicated and highlighted with arrows and red bars. The location of sgRNAs are shown with arrows (see Table S3), and the asterisks mark the sgRNAs from [5].

Figure 2. Replication timing in cells with *ASAR6-141 ~7kb* RBPD deletion or ectopic integration. A) Schematic illustration of the BrdU terminal label protocol [30]. Cells were treated with BrdU (green) for 5 hours and then harvested for mitotic cells. B) BrdU incorporation in HTD114 cells containing a heterozygous deletion of the ~7kb RBPD. Cells containing a deletion of the ~7kb RBPD from the expressed allele of *ASAR6-141* were exposed to BrdU, harvested for mitotic cells, and subjected to DNA FISH using a chromosome 6 centromeric probe (red; B). The larger centromere resides on the chromosome 6 with the silent *ASAR6-141* allele (CHR6A), and the smaller centromere resides on the chromosome 6 with the expressed *ASAR6-141* allele (CHR6B). DNA was stained with DAPI (blue). C and D) DAPI staining and BrdU incorporation were quantified by calculating the number of pixels (area x intensity) and displayed as a ratio of BrdU incorporation in CHR6B divided by the BrdU incorporation in CHR6A. E) Quantification of BrdU incorporation in multiple cells with heterozygous (Δ7kb-6B, expressed allele; or Δ7kb-6A, silent allele) or homozygous (Δ7kb-6B/6A) deletions of the ~7kb RBPD. Also

671 shown is the quantification of BrdU incorporation in heterozygous deletions of the entire
672 ~185 kb *ASAR6-141* gene from chromosome 6B (Δ185-6B) or 6A (Δ185-6A) [5]. Box plots
673 indicate mean (solid line), standard deviation (dotted line), 25th, 75th percentile (box) and
674 5th and 95th percentile (whiskers) and individual cells (single points). P values were
675 calculated using the Kruskal-Wallis test [37]. Values for individual cells are shown as dots.
676 **P = 0.00027 and *P = 0.03. F) Schematic view of the ~7kb RBPD in both the sense and
677 antisense orientation. The promoter used to drive expression is from the ASAR6 gene [9].
678 G and H) Two color RNA FISH assay for expression of the ~7kb RBPD transgenes.
679 Individual clones were screened for RNA expressed from the ASAR6-141 (~7kb RBPD)
680 transgenes (red, arrows), and RNA hybridization using an Xist probe (green, arrows) was
681 used as positive control. G and H represent examples of cells containing the sense and
682 antisense transgenes, respectively. I) Quantitation of the size (pixels = area x intensity)
683 of the RNA FISH signals expressed from the 7 kb sense (7 kb S) or antisense (7 kb AS)
684 transgenes are shown. Box plots indicate mean (solid line), standard deviation (dotted
685 line), 25th, 75th percentile (box) and 5th and 95th percentile (whiskers) and individual
686 cells (single points). P value of 0.006 was calculated using the Kruskal-Wallis test [37].
687 Values for individual cells are shown as dots. J) RNA from the ~7kb RBPD remains
688 localized to mouse chromosomes. RNA-DNA FISH using the ~7kb RBPD as RNA FISH
689 probe (red, arrows), plus a mouse chromosome 5 paint to detect chromosome 5 DNA
690 (green, brackets). K) Cells containing the sense ~7kb RBPD transgene integrated into
691 mouse chromosome 5 were exposed to BrdU (green), harvested for mitotic cells, and
692 subjected to DNA FISH using the ~7kb RBPD (red, arrows). Note that the chromosome
693 that contains the transgene shows delayed mitotic condensation and more BrdU
694 incorporation than any other chromosome within the same cell. L) Cells containing
695 the antisense ~7kb RBPD transgene integrated into mouse chromosomes were exposed
696 to BrdU (green), harvested for mitotic cells, and subjected to DNA FISH using the ~7kb
697 RBPD (red). Note that the chromosome that contains the transgene shows normal mitotic
698 condensation and only a small amount of BrdU incorporation. Scale bars are 10 uM (B,
699 G, H) 5 uM (K, L) and 2 uM (J).
700
701 **Figure 3. Depletion of RBPs results in disruption of the chromosome territory**
702 **localization of ASAR6-141 RNA.** A-J) shRNA mediated depletion of RBPs. K562 cells
703 were transfected with empty vector (A, C, E, G, and I) or vectors expressing shRNAs
704 directed against HNRNPU (B), HNRNPUL1 (D), HNRNPC (F), HNRNPM (H), or HLTF
705 (J). K) Cells were stained with the appropriate antibodies and quantitation of each RBP
706 was determined in >25 individual cells. Box plots indicate mean (solid line), standard
707 deviation (dotted line), 25th, 75th percentile (box), 5th and 95th percentile (whiskers), and
708 individual cells (single points). P values were calculated using the Kruskal-Wallis test [37].
709 L-V) RNA-DNA FISH for ASAR6-141 (green) and XIST (red) RNA. K562 cells were
710 transfected with empty vector (L) or vectors expressing shRNAs against HNRNPU (M),
711 HNRNPUL1 (N), HNRNPC (O), HNRNPM (P), HLTF (Q), HNRNPA1 (R), HNRNPL (S),
712 KHSRP (T), UCHL5 (U), or PTBP1 (V). The red arrows mark the RNA FISH signals for
713 XIST. The brackets mark cytoplasmic regions that hybridized to both RNA FISH probes.
714 DNA was stained with DAPI, and Bars are 10 *u*M (A-J and L) or 5 *u*M (M-V). W)
715 Quantitation of the XIST RNA FISH signals. K562 cells transfected and processed for
716 RNA FISH as in L-V were analyzed for quantitation of XIST RNA cloud size (pixels: area

X intensity) in >25 individual cells. Box plots indicate mean (solid line), standard deviation (dotted line), 25th, 75th percentile (box), 5th and 95th percentile (whiskers) and individual cells (single points). P values were calculated using the Kruskal-Wallis test [37].

Figure 4. Ectopic integration of the *ASAR6-141 ~7kb RBPD* transgene into an inactive X chromosome. A) Two color RNA FISH assay for expression of the sense ~7kb RBPD transgene plus Xist RNA. Examples of three different cells showing colocalization of the ASAR RNA and Xist RNA. Individual cells were processed for expression of the sense strand ~7kb RBPD transgene (left panels, red, arrow) in combination with RNA FISH for Xist (middle panels, green, arrow). The right panels show the merged images. The white lines show the path for the pixel intensity profiles used for panel B. B) Pixel intensity profiles across the Xist RNA hybridization domain showing enhanced signal intensity over the ~7kb RBPD RNA hybridization domain. C and D) Delayed mitotic chromosome condensation of the X chromosome that contains the ~7kb RBPD transgene. An example of a mitotic spread processed for DNA FISH using the ~7kb RBPD as probe (magenta, panel C) plus a mouse X chromosome paint (red, panel D). The arrows mark the location of the transgene hybridization signal. E-H) Cells containing the ~7kb RBPD transgene integrated into the mouse X chromosome were exposed to BrdU (G, green), harvested for mitotic cells, and subjected to DNA FISH using the ~7kb RBPD (E, magenta) plus a mouse X chromosome paint (F, red). The merged images are shown in panel H.

Figure 5. Depletion of RBPs results in disruption of the chromosome territory localization of ASAR RNAs. shRNA mediated depletion of RBPs in K562 cells (A-K) or HTD114 cells (L-Q). Cells were transfected with empty vector (A-C and L-N) or vectors expressing shRNAs directed against HNRNPA1 (D), HNRNPC (E), HNRNPL (F), HNRNPM (G), HNRNPU (H, and O-Q), HLTF (I), KHSRP (J), or UCHL5 (K). Cells were processed for RNA FISH with probes for ASAR6 (red, arrows) and ASAR1-187 (green, arrows). Cells were also processed for DNA FISH using BAC DNA to detect chromosome 1 (CHR1 BAC DNA, magenta) and chromosome 6 (CHR6 BAC DNA, orange; A-C and L-Q). Arrows mark the sites of RNA FISH hybridization and arrowheads mark the sites of DNA hybridization. RNA-DNA FISH for ASAR1-187 (RNA in green, arrows; DNA in magenta and arrowheads), ASAR6 (RNA in red, arrows; DNA in orange and arrowheads). The brackets mark the cytoplasmic regions that hybridized to both RNA FISH probes. DNA was stained with DAPI, and scale bars are 10 *u*M.

Figure 6. Depletion of RBPs results in asynchronous replication on autosome pairs. A) BrdU incorporation in parental HTD114 cells. Cells were exposed to BrdU for five hours and processed for BrdU incorporation using an antibody against BrdU. This panel represents chromosomes from multiple mitotic spreads showing representative BrdU incorporation in chromosome pairs. Both homologs of autosome pair were captured from the same mitotic cell, and each pair displays a typical BrdU incorporation pattern that is consistent with synchronous replication timing. B) HTD114 cells were transfected with the HNRNPU shRNA expression vector, exposed to BrdU for five hours, and processed for BrdU incorporation. This panel represents chromosomes from multiple mitotic spreads showing representative BrdU incorporation in pairs of autosomes. Both

763 homologs of autosome pairs were captured from the same mitotic cell, and each pair
764 displays a differential BrdU incorporation pattern that is consistent with asynchronous
765 replication timing. C) Shows two chromosome 4 homologs (CHR4A and CHR4B) side by
766 side. D) Shows the BrdU incorporation in chromosome 18 homologs (CHR18A and
767 CHR18B). For each chromosome pair the inverted DAPI staining (black and white), BrdU
768 incorporation (green), and DAPI staining (blue) are shown. E and F) Quantification of
769 BrdU incorporation in multiple cells depleted for HNRNPU, HNRNPUL1, HTLF, KHSRP,
770 UCHL5, HNRNPA1, MATR3, PTBP1, PTBP2, SAFB and SAFB2. Box plots indicate mean
771 (solid line), standard deviation (dotted line), 25th, 75th percentile (box), 5th and 95th
772 percentile (whiskers) and individual cells (single points). P values were calculated using
773 the Kruskal-Wallis test [37]. G) DRT/DMC on chromosome 6 following HNRNPU depletion.
774 HTD114 cells were transfected with the HNRNPU shRNA expression vector, exposed to
775 BrdU for 5 hours, harvested for mitotic cells, and processed for BrdU incorporation (green)
776 and DNA FISH using a chromosome 6 centromeric probe (red). H) DRT/DMC on
777 chromosome 1 following UCHL5 depletion. HTD114 cells were transfected with the
778 UCHL5 shRNA expression vector, exposed to BrdU for 5 hours, harvested for mitotic
779 cells, and processed for BrdU incorporation (green) and DNA FISH using a chromosome
780 1 paint as probe (red). The arrow marks the chromosome 1 with DRT/DMC, and the inset
781 shows only the DAPI staining of the chromosome 1 highlighting the DMC.
782
783 **Figure 7. Human C0T-1 DNA detects ASAR6 RNA.** A and B) RNA-DNA FISH on mouse
784 cells containing an ASAR6 BAC transgene integrated into mouse chromosome 3 [8]. A)
785 RNA-DNA FISH using an ASAR6 fosmid probe to detect ASAR6 RNA (green; arrows)
786 plus a mouse chromosome 3 paint probe to detect DNA (red; arrowheads). B) RNA-DNA
787 FISH using human C0T-1 DNA to detect RNA (green; arrows) plus the *ASAR6* BAC (red
788 arrowheads) to detect DNA.
789
790 **Figure S1. CRISPR/Cas9 deletion of the ~7kb RBPD.** A) Integrative Genomics Viewer
791 view showing the genomic location of *ASAR6-141*. The genomic locations of RNAseq
792 and eCLIP reads from HepG2 and K562 (from ENCODE) are shown. The zoomed in view
793 shows the ~7kb RBPD with the location where all of the peaks of eCLIP reads map within
794 the region (see Table S1). The location of sgRNAs that flank the ~7kb RBPD are shown.
795 The location of PCR primers (1-6) and the heterozygous SNP rs17070386 are also shown
796 (see Table S3). B) Sanger sequencing traces on PCR products generated from genomic
797 DNA using primers 5 and 6 (see Table S3). In previous studies we generated mouse
798 monochromosomal hybrid cells containing the two chromosome 6 homologs from
799 HTD114 [7], and sequencing traces from these two hybrids, contain either chromosome 6A
800 (CHR6A) or chromosome 6B (CHR6B) are shown below the HTD114 panel. Sequencing
801 traces from PCR products generated from 4 heterozygous deletions, two with CHR6A
802 deletions {D6A(7-82-29) and D6A(7-82-13)} and two with CHR6B deletions {D6B(7-18-1)
803 and D6B(7-63-1)} are shown. C) Junction PCR (primers 1 and 4) on genomic DNA
804 isolated from heterozygous {D6A(7-82-29), D6A(7-82-13), D6B(7-18-1) and D6B(7-63-1)}
805 or homozygous deletions {D6B/6A-1 and D6B/6A-2} are shown. Parental HTD114 DNA
806 was used as control. D) Internal PCR products (primers 5 and 6) on genomic DNA isolated
807 from heterozygous {D6A(7-82-29), D6A(7-82-13), D6B(7-18-1) and D6B(7-63-1)} or

homozygous deletions {D6B/6A-1 and D6B/6A-2} are shown. Parental HTD114 DNA was used as control.

Figure S2. Identification of mouse chromosomes containing the ~7kb RBPD transgene. A) DNA FISH on mouse chromosomes from clone ~7kb(+)A5. Mitotic spread showing DNA FISH using the ~7kb RBPD (magenta; arrow) and a mouse chromosome 5 BAC (RP24-107D17; red; arrowheads). B and C) DNA FISH on mouse chromosomes from clone ~7kb(+)X1. Mitotic spread showing DNA FISH using the ~7kb RBPD (magenta; arrows) and a mouse X chromosome paint (red).

Figure S3. RBP protein levels in cells expressing shRNAs. K562 or HTD114 cells were transfected with empty vector (-) or shRNA against RBPs and cell extracts were processed for Western Blots using antibodies (Table S4) as indicated: A and B) HNRNPU (green); C) UCHL5 (green); D and E) HNRNPC (green); F and I) HNRNPL (green); G and H) HNRNPM (green); J and K) KHSRP (green); L) HNRNPUL1 (green); M) HLTF (green); N) SAFB2 (green); O) PTBP1 (red); or P) HNRNPA1 (green). Tubulin (red) was used as loading control, and size markers are indicated.

Figure S4. RNA FISH on cells transfected with shRNAs that did not disrupt ASAR localization. A-D) K562 cells were transfected with empty vector (A) or shRNA expression vectors for MATR3 (B), PTBP1 (C), or PTBP2 (D). Cells were processed for RNA FISH using probes for ASAR6-141 (green; arrows) or XIST (red; arrows). E-J) HTD114 cells were transfected with empty vector (E) or shRNA expression vectors for SAFB (F), SAFB2 (G), SAFB plus SAFB2 (H), CTCF (assayed for only ASAR6 RNA) (I), or MATR3 (J). DNA was stained with DAPI, and scale bars are 10 *u*M.

Figure S5. Disruption of the chromosome territory localization of ASAR1-187 and ASAR6 RNAs following shRNA depletion of RBPs. HTD114 cells were transfected with empty vector (A) or shRNA expression vectors for HNRNPA1 (B), HNRNPC (C), HNRNPL (D), HNRNPM (E), HNRNPUL1 (F), KHSRP (G), HLTF (H), or UCHL5 (I). Cells were processed for RNA FISH using Fosmid probes (see Table S3) for ASAR1-187 (green, arrow) and ASAR6 (red, arrow). A) DNA FISH hybridization signals for chromosome 1 (CHR1 BAC DNA; magenta) and chromosome 6 (CHR6 BAC DNA; orange) are indicated by arrowheads. The brackets mark cytoplasmic regions that hybridized to both ASAR RNA FISH probes. DNA was stained with DAPI, and scale bars are 10 *u*M.

Figure S6. Replication timing in cells transfected with shRNA expression vectors. A and B) HTD114 cells were transfected with PTBP1 shRNA expression vector, exposed to BrdU for 5 hours, harvested for mitotic cells, and processed for BrdU incorporation (green) and DNA FISH using a chromosome 1 (A, red) or 6 (B, magenta) centromeric probe. Cells were exposed to BrdU for five hours, processed for BrdU incorporation using an antibody against BrdU. The ratio of BrdU incorporation in chromosome 1 and chromosome 6 homologs are shown in the bottom panels of A and B, respectively. C) HTD114 cells were transfected with shRNA expression vector for HNRNPU, exposed to BrdU for 5 hours, harvested for mitotic cells, and processed for BrdU incorporation (green)

854 and DNA FISH using centromeric probes for chromosome 1 (red) or 6 (magenta). The
855 ratio of BrdU incorporation in chromosome homologs for 1, 6, 12, and 18 are shown in
856 the bottom panels. DNA was stained with DAPI, and scale bars are 10 *u*M.
857
858 **Figure S7. Replication timing alterations in shRNA depleted cells.** HTD114 cells
859 were transfected with shRNA expression vectors for HNRNPA1 (A), HNRNPC (B),
860 HNRNPL (C), HNRNPM (D), HLTF (E), KHSRP (F), HNRNPU (G), or UCHL5 (H),
861 exposed to BrdU for 5 hours, harvested for mitotic cells, and processed for BrdU
862 incorporation (green) and DNA FISH using a chromosome 1 centromeric probe (A and G;
863 red, arrows) or chromosome 1 paint (B, C, D, E, F and H; red, arrows) plus a chromosome
864 6 centromeric probe (A-H; magenta, arrows). The brackets mark chromosomes with
865 DRT/DMC. DNA was stained with DAPI, and scale bars are 10 *u*M.
866
867
868 **Table S1. The location of significant peaks of eCLIP reads that map within ASAR**
869 **and XIST genes.** The location of significant, fold enrichment and log10(P value), eCLIP
870 reads, from K562 and HepG2, that mapped within ASAR1-187, ASAR6-141, ASAR8-2.7,
871 ASAR9-23, and XIST.
872
873 **Table S2. eCLIP reads for 120 RBPs in 10 kb windows within ASAR genes.** The FDR-
874 corrected p-value of the z-score of log2-ratio between eCLiP vs control for each RBP in
875 10kb windows within ASAR1-187, ASAR6-141, ASAR8-2.7, and ASAR9-23, from K562
876 and HepG2 cells.
877
878

RESOURCE AVAILABILITY

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Mathew Thayer (thayerm@ohsu.edu)

Materials availability statement

All reagents generated in this study will be made freely available upon request.

Acknowledgments:

M.J.T. was supported by NIH NIGMS (R01GM114162 and R01GM130703).

M.B.H was supported by NIH NCI 4K00CA245677-03

Author contributions:

Declaration of interests

The authors declare no competing interests.

895 **Methods**
896
897 *Cell culture*
898 HTD114 cells are a human *APRT* deficient cell line derived from HT1080 cells [38]. K562
899 cells were obtained from ATCC, and cultured in RPMI 1640 (Gibco) supplemented with
900 10% fetal bovine serum (Hyclone). Mouse C2C12 cells were used for the transgene
901 integration assays, transfected using Lipofectamine 2000 (Invitrogen) according to the
902 manufacturer's recommendations, and clones containing transgene integrations were
903 isolated following selection in media containing G418. All cells were grown in DMEM
904 (Gibco) supplemented with 10% fetal bovine serum (Hyclone), and were grown in a
905 humidified incubator at 37°C in a 5% carbon dioxide atmosphere.
906
907 *eCLIP analysis*
908 eCLIP data was obtained from [39]**.** Given that non-coding genes and TLs are expressed
909 at a lower level than coding genes, a sensitive filtering approach was used as
910 recommended by the authors [39]. We filtered out peaks containing an eCLIP peak p-value
911 greater than p=0.05, and filtered out peaks containing a fold enrichment value less than
912 two.
913
914 *Identification of TL*
915 Nuclear-enriched, strand-specific, ribo-minus, total-RNA sequencing libraries were first
916 aligned to the hg19 reference genome using default parameters with STAR [40]. Next
917 SAMtools [41] was used to remove duplicate and low quality (<=MAPQ20) reads. TLs were
918 defined using a strategy of sequential merging of strand-specific, contiguous intergenic
919 reads, as performed in [6]. Utilizing strand-specific reads, stranded reads separated by
920 1,000 bp or less were merged to create contigs, and the contigs merged again while
921 allowing for gaps of 7 kb to allow for regions that are not uniquely mappable due to
922 presence of full-length LINE elements. TL were then classified above a minimum cutoff
923 length of 50 kb of strand specific, contiguous expression.
924
925 *shRNA depletion of RBPs*
926 Using Lipofectamine 2000 (Invitrogen), according to the manufacturer's
927 recommendations, we co-transfected HTD114 cells with plasmids encoding puromycin
928 resistance (PGK-PURO) and shRNAs against 14 different RBPs (Table S4). 24h after
929 transfection, cells were plated in media containing 10 ug/ml puromycin (Millipore/Sigma),
930 and 24h-48h following puromycin selection cells were processed for Western blot, RNA-
931 DNA FISH or for replication timing assays (see below).
932
933 *Western blot analysis.*
934 Cells were pelleted, rinsed once in PBS, and lysed by resuspension in RIPA Solubilization
935 Buffer (25 mM Tris-HCl pH 7.6, 150 mM NaCl, 5 mM EDTA,1% NP-40 or 1% Triton X-
936 100, 1% sodium deoxycholate, 0.1% SDS). Protein concentration was determined via the
937 Qubit Protein BR Assay kit (Invitrogen, A50669) on a Qubit 4 Fluorometer. Lysates from
938 HTD114 (20 μg) or K562 (25 μg) cell lines in 1X Laemmli Sample Buffer (Bio-Rad, 161-
939 0737) with 5% β-mercaptoethanol were loaded onto 10% Mini-PROTEAN TGX Precast
940 Tris-Glycine-SDS polyacrylamide gels, and run in 1X SDS PAGE running buffer (Bio-Rad,

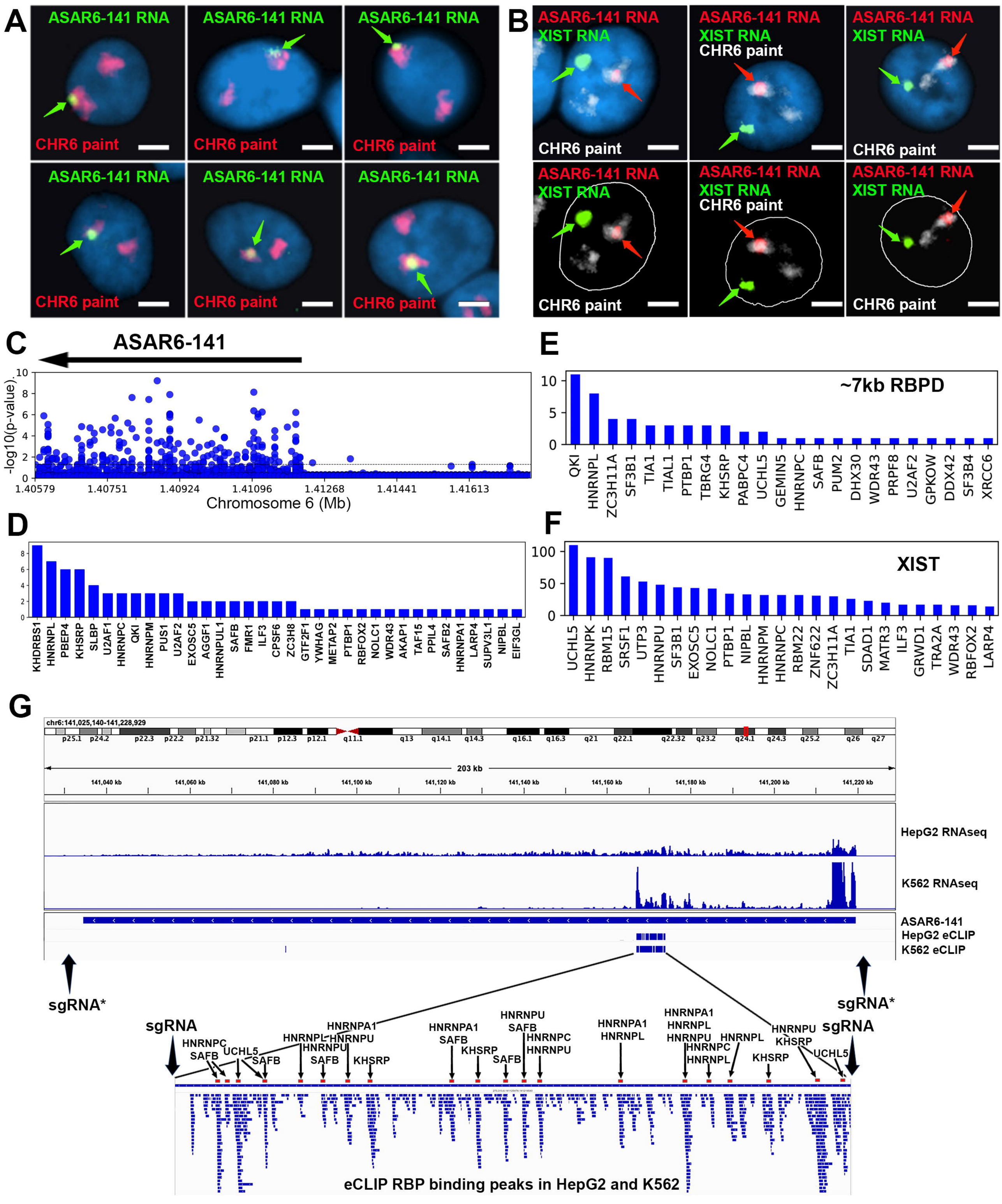
941 161-0732) at 100 Volts for 1.25 hours. Proteins were transferred to either NitroBind
942 nitrocellulose (GE, EP4HY00010) or Immobilon-FL PVDF (Millipore, IPFL00010)
943 membranes in Tris-Glycine transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol
944 pH 8.3) at 100 V for one hour. All primary antibodies are described in Table S4.
945 For protein detection, membranes were blocked for one hour in 50% AquaBlock (Arlington
946 Scientific, PP82P-500-QT-1582) in PBS (AquaBlock-P) and probed with specific
947 antibodies (Table S4) in AquaBlock-P with 0.1% Tween 20 (AquaBlock-PT) at 4 °C
948 overnight. The membranes were subjected to four 5-minutes washes in PBS with 0.1%
949 Tween 20 (PBS-T) at room temperature. Membranes were then incubated with secondary
950 antibodies in AquaBlock-PT for one hour at room temperature, and subjected to four
951 washes in PBS-T as described above. Antibody labeling was detected using an Azure
952 Sapphire Biomolecular Imager.
953
954 *Immunofluorescence:*
955 Cells were allowed to attach to poly-L lysing coated glass slides for 2 hours. Cells were
956 fixed with 4% paraformaldehyde in PBS pH 7.4 for 10 min at room temperature, then
957 washed 3X with ice-cold PBS. Cells were Permeabilized for 10 min with PBS containing
958 either 0.1% Triton X-100 (PBST), and washed 3X with PBS at room temperature. Prior to
959 antibody addition, slides were incubated for 30 minutes with 10% serum from the species
960 the secondary antibody was raised in. Slides were incubated in diluted primary antibody
961 in 1% BSA in PBST in a humidified chamber overnight at 4°C. Slides were washed three
962 times in PBS, 5 min each wash. Slides were incubated with secondary antibody in 1%
963 BSA for 1 h at room temperature in the dark. Slides were wash three times with PBS for
964 5 min each in the dark. All primary antibodies are described in Table S4.
965
966 *CRISPR/Cas9 engineering*
967 Using Lipofectamine 2000 (Invitrogen), according to the manufacturer's
968 recommendations, we co-transfected HTD114 cells with plasmids encoding GFP,
969 sgRNAs and Cas9 endonuclease (Origene). Each plasmid encoded sgRNA was designed
970 to bind at the indicated locations (Fig. S1). 48h after transfection, cells were plated at
971 clonal density and allowed to expand for 2-3 weeks. The presence of deletions was
972 identified by PCR using the primers described in Figure S1. The single cell colonies that
973 grew were analyzed for heterozygous deletions by PCR (Fig. S1). We used retention of a
974 heterozygous SNP to identify the disrupted allele, and homozygosity at the SNP confirmed
975 that cell clones were homogenous.
976
977 *PCR analysis*
978 Genomic DNA was isolated from tissue culture cells using TRIZOL Reagent (Invitrogen).
979 PCR was performed in a 12.5μL volume using 50-100ng of genomic DNA, 1x Standard
980 Taq Buffer (New England Biolabs, Inc.), 200μM each deoxynucleotide triphosphates,
981 0.2μM of each primer, and 3 units of Taq DNA Polymerase (New England Biolabs, Inc.)
982 under the following reaction conditions: 95°C for 2 minutes, followed by 30-40 cycles of
983 95°C for 30 seconds, 55-62°C for 45 seconds, and 72°C for 1 minute, with a final
984 extension time of 10 minutes at 72°C. PCR products were separated on 1% agarose gels,
985 stained with ethidium bromide, and photographed under ultraviolet light illumination.

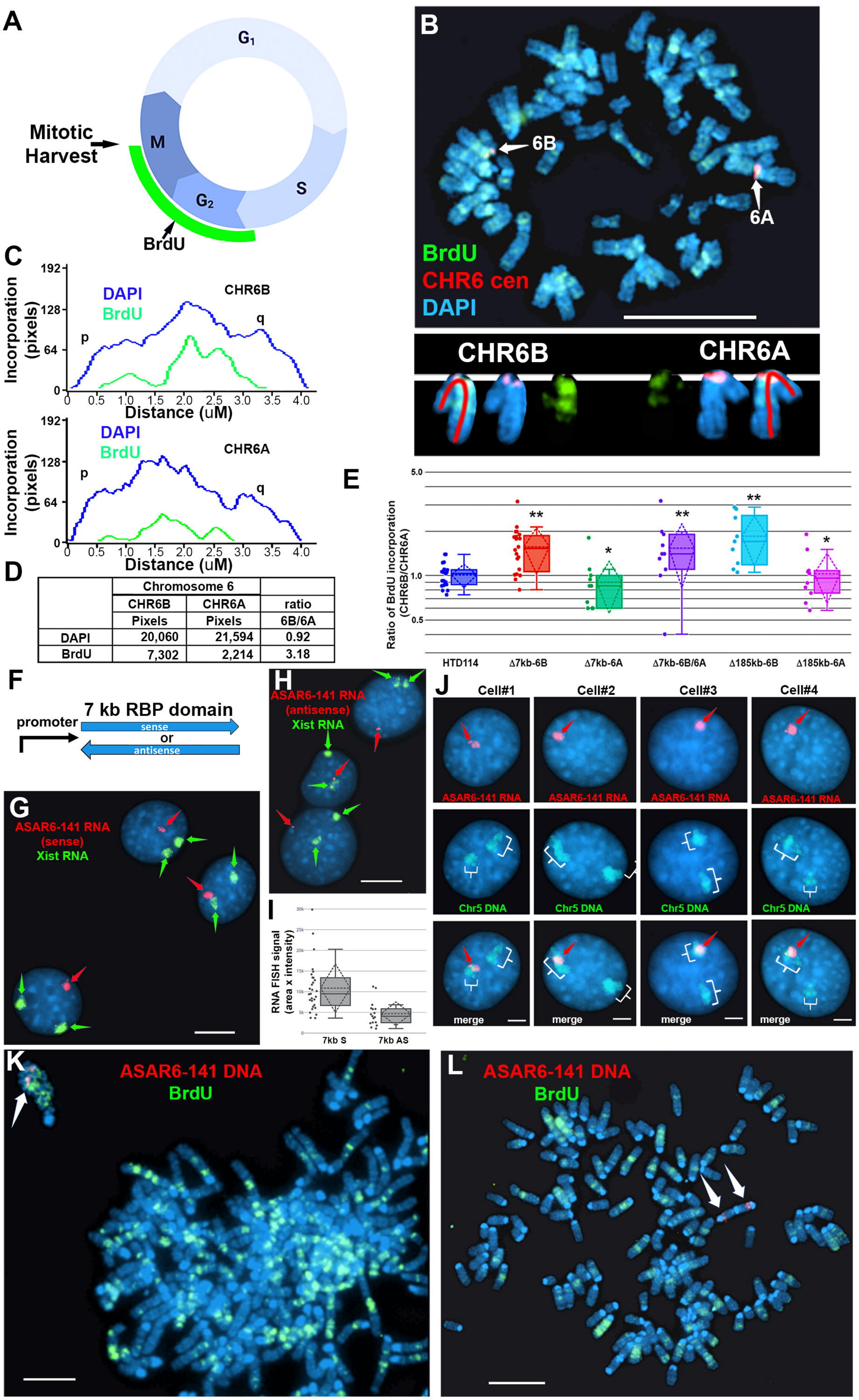
986 Sequencing of PCR products was carried out at the Vollum Institute DNA Sequencing
987 Core facility. All PCR primers are described in Table S3.
988
989 *DNA FISH*
990 Mitotic chromosome spreads were prepared as described previously [29]. After RNase
991 (100µg/ml) treatment for 1h at 37°C, slides were washed in 2XSSC and dehydrated in an
992 ethanol series and allowed to air dry. Chromosomal DNA on the slides was denatured at
993 75°C for 3 minutes in 70% formamide/2XSSC, followed by dehydration in an ice-cold
994 ethanol series and allowed to air dry. BAC and Fosmid DNAs were labeled using nick
995 translation (Vysis, Abbott Laboratories) with Spectrum Orange-dUTP, Spectrum Aqua-
996 dUTP or Spectrum Green-dUTP (Vysis). Final probe concentrations varied from 40-60
997 ng/µl. Centromeric probe cocktails (Vysis) and/or whole chromosome paint probes
998 (Metasystems) plus BAC or Fosmid DNAs were denatured at 75°C for 10 minutes and
999 prehybridized at 37°C for 10 minutes. Probes were applied to denatured slides and
1000 incubated overnight at 37°C. Post-hybridization washes consisted of one 3-minute wash
1001 in 50% formamide/2XSSC at 40°C followed by one 2-minute rinse in PN (0.1M Na_2HPO_4,
1002 pH 8.0/2.5% Nonidet NP-40) buffer at RT. Coverslips were mounted with Prolong Gold
1003 antifade plus DAPI (Invitrogen) and viewed under UV fluorescence (Olympus). All BAC
1004 and Fosmid probes are described in Table S3.
1005
1006 *RNA-DNA FISH*
1007 Cells were plated on Poly-L-Lysine coated (Millipore Singa) glass microscope slides at
1008 ~50% confluence and incubated for 4 hours in complete media in a 37°C humidified CO_2
1009 incubator. Slides were rinsed 1X with sterile RNase free PBS. Cell Extraction was carried
1010 out using ice cold solutions as follows: Slides were incubated for 30 seconds in CSK buffer
1011 (100mM NaCl/300mM sucrose/3mM $MgCl_2$/10mM PIPES, pH 6.8), 10 minutes in CSK
1012 buffer/0.1% Triton X-100, followed by 30 seconds in CSK buffer. Cells were then fixed in
1013 4% paraformaldehyde in PBS for 10 minutes and stored in 70% EtOH at -20°C until use.
1014 Just prior to RNA FISH, slides were dehydrated through an EtOH series and allowed to
1015 air dry. Denatured probes were prehybridized at 37°C for 10 min, applied to non-
1016 denatured slides and hybridized at 37°C for 14-16 hours. Post-hybridization washes
1017 consisted of one 3-minute wash in 50% formamide/2XSSC at 40°C followed by one 2-
1018 minute rinse in 2XSSC/0.1% TX-100 for 1 minute at RT. Slides were then fixed in 4%
1019 paraformaldehyde in PBS for 5 minutes at RT, and briefly rinsed in 2XSSC/0.1% TX-100
1020 at RT. Coverslips were mounted with Prolong Gold antifade plus DAPI (Invitrogen) and
1021 slides were viewed under UV fluorescence (Olympus). Z-stack images were generated
1022 using a Cytovision workstation. After capturing RNA FISH signals, the coverslips were
1023 removed, the slides were dehydrated in an ethanol series, and then processed for DNA
1024 FISH, beginning with the RNase treatment step, as described above. All BAC and Fosmid
1025 probes are described in Table S3.
1026
1027 *Chromosome replication timing assay*
1028 The BrdU replication timing assay was performed as described previously on
1029 exponentially dividing cultures and asynchronously growing cells [30]. Mitotic chromosome
1030 spreads were prepared and DNA FISH was performed as described above. The
1031 incorporated BrdU was then detected using a FITC-labeled anti-BrdU antibody (Roche).

Coverslips were mounted with Prolong Gold antifade plus DAPI (Invitrogen), and viewed under UV fluorescence. All images were captured with an Olympus BX Fluorescent Microscope using a 100X objective, automatic filter-wheel and Cytovision workstation. Individual chromosomes were identified with either chromosome-specific paints, centromeric probes, BACs or by inverted DAPI staining. Utilizing the Cytovision workstation, each chromosome was isolated from the metaphase spread and a line drawn along the middle of the entire length of the chromosome. The Cytovision software was used to calculate the pixel area and intensity along each chromosome for each fluorochrome occupied by the DAPI and BrdU (FITC) signals. The total amount of fluorescent signal in each chromosome was calculated by multiplying the average pixel intensity by the area occupied by those pixels. The BrdU incorporation into human chromosomes containing CRISPR/Cas9 modifications was calculated by dividing the total incorporation into the chromosome with the deleted chromosome divided by the BrdU incorporation into the non-deleted chromosome within the same cell. Boxplots were generated from data collected from 8-12 cells per clone or treatment group. Differences in measurements were tested across categorical groupings by using the Kruskal-Wallis test [37] and listed as P-values for the corresponding plots. All BAC and Fosmid probes are described in Table S3.

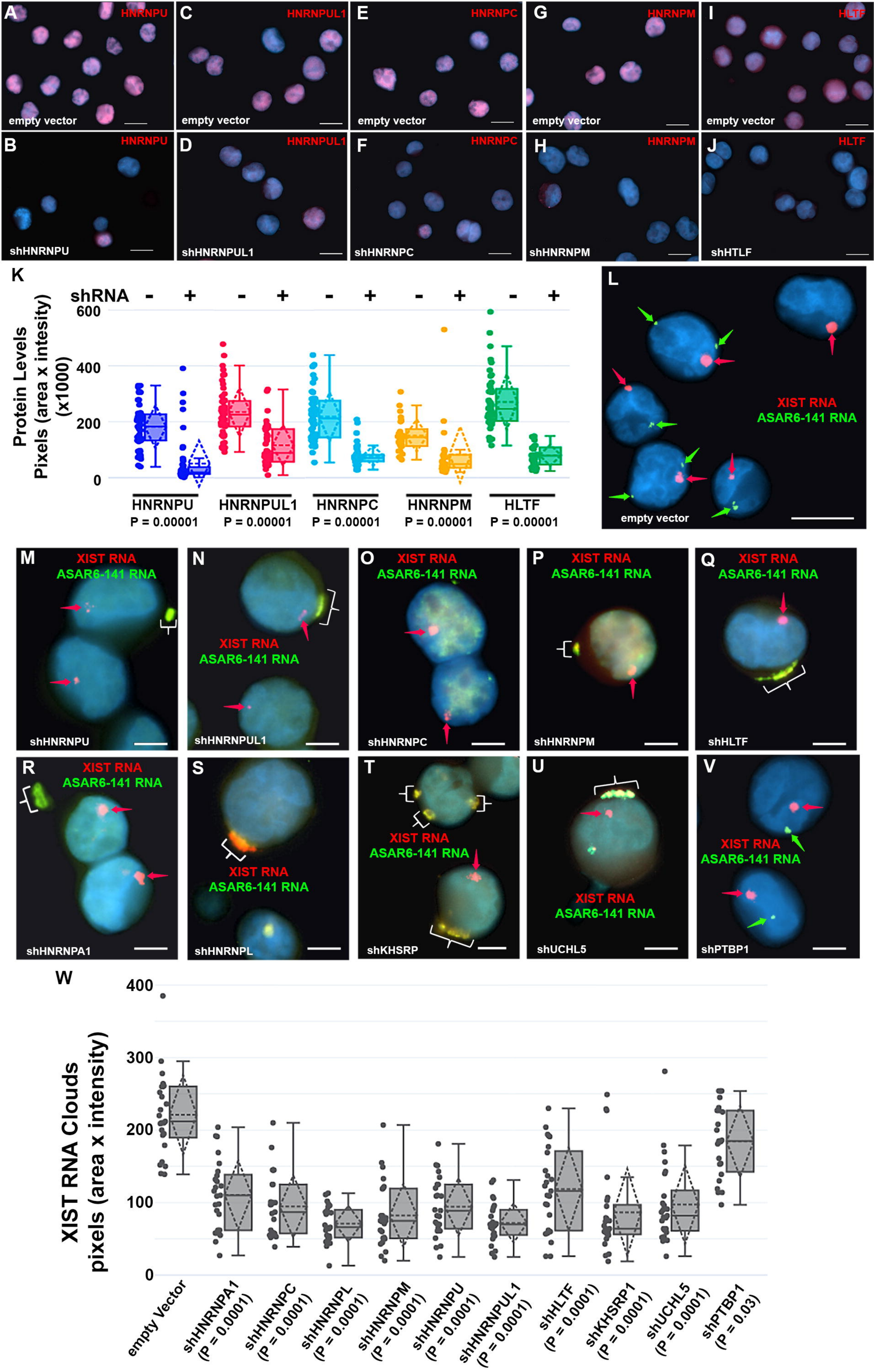
Quantification and statistical analysis

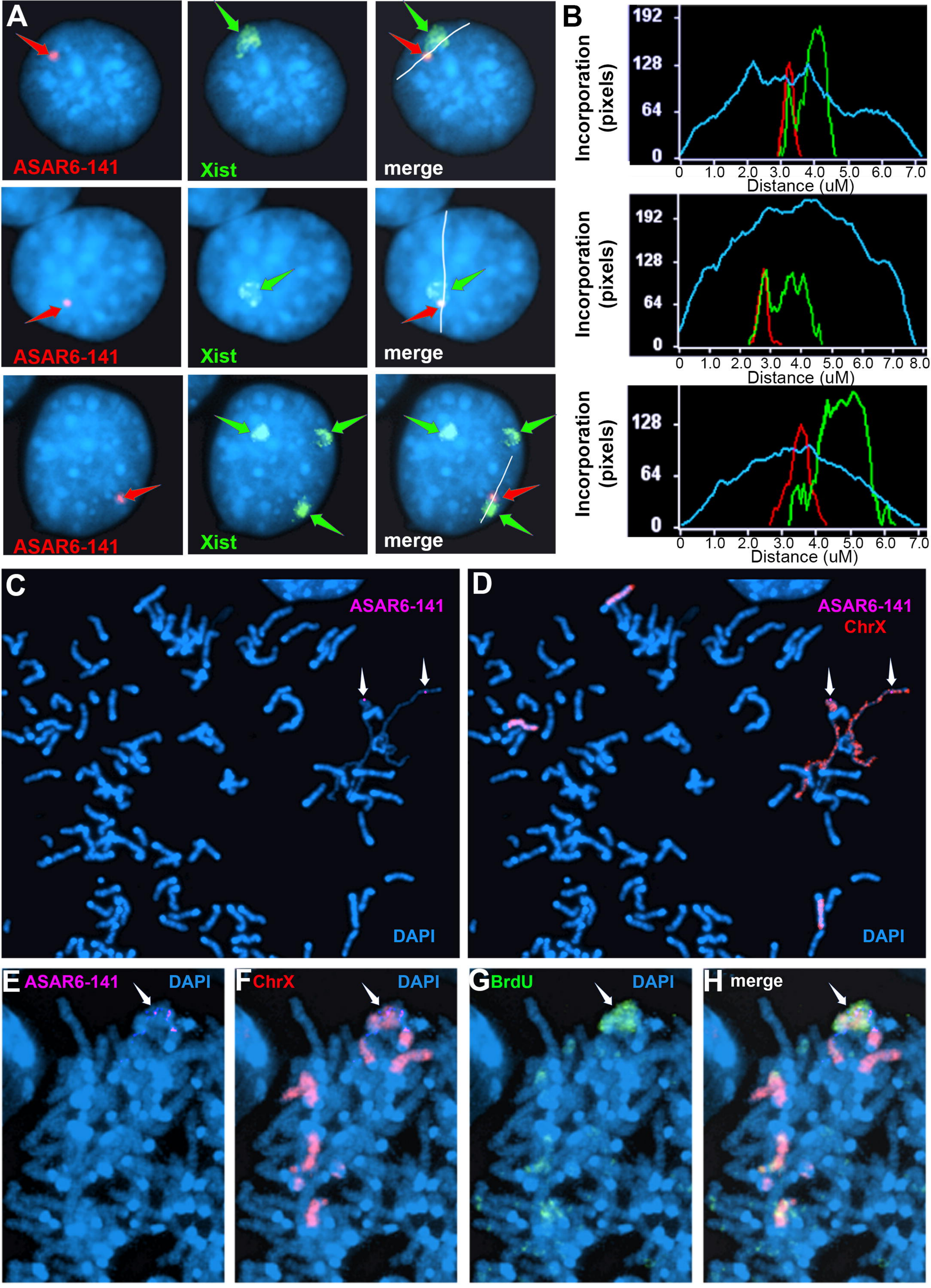
P-values were generated for eCLiP data using the “region-based” method described in [26]. For 10kb sliding windows across the genome the log2-Ratio was calculated between the number of reads in the eCliP sample and matched control, and P-values were generated using the Python scipy.stats function “zscore”. FDR correction was performed using the Python Statsmodels function “fdrcorrection”.

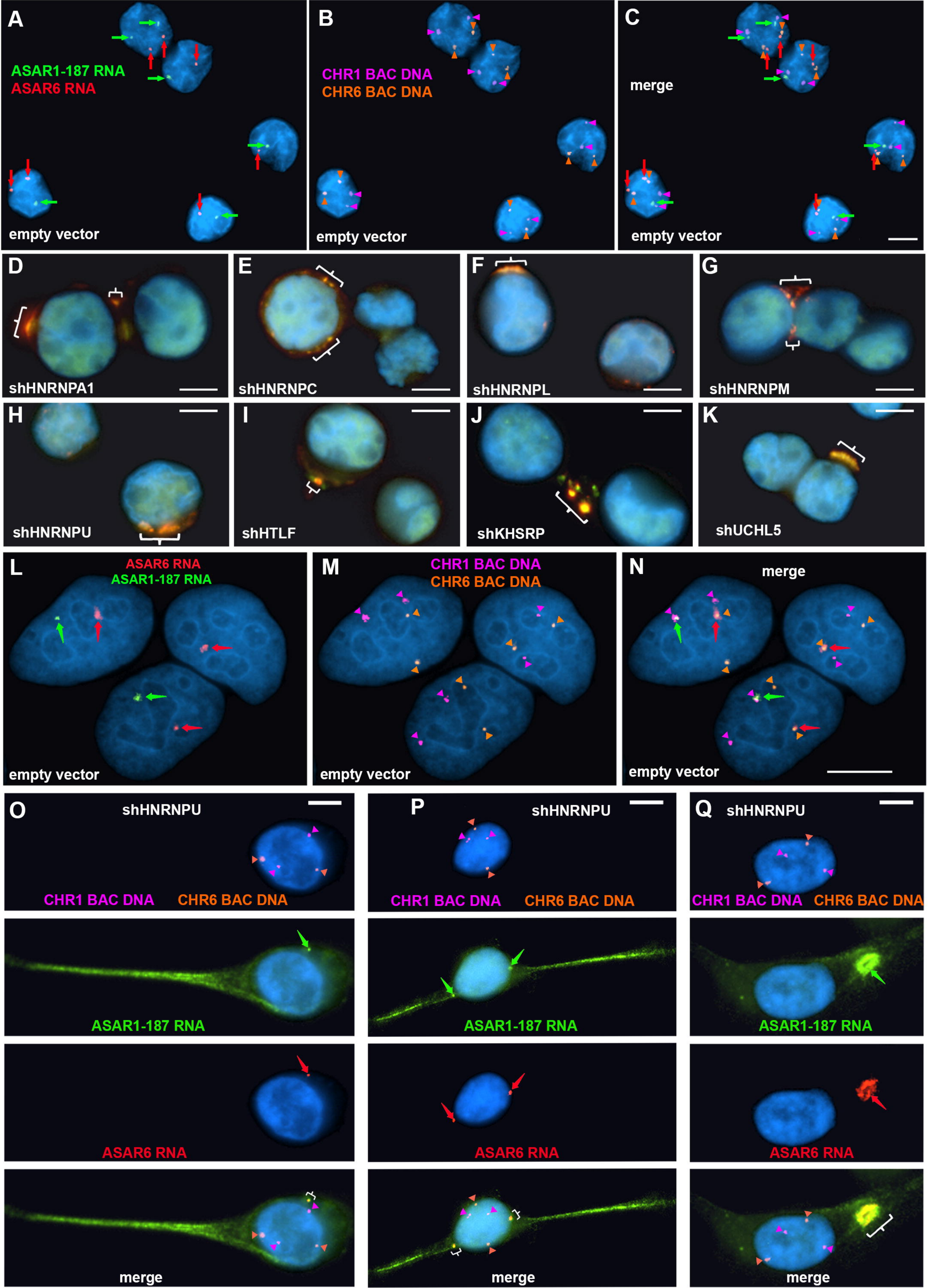

A
ASAR6-141 RNA
CHR6 paint
B
ASAR6-141 RNA
XIST RNA
CHR6 paint
C
ASAR6-141
-log10(p-value)
Chromosome 6 (Mb)
1.40579
1.40751
1.40924
1.41096
1.41268
1.41441
1.41613
D
KHDRBS1
HNRNPL
PBEP4
KHSRP
SLBP
U2AF1
HNRNPC
QKI
HNRNPM
PUS1
U2AF2
EXOSC5
AGGF1
HNRNPUL1
SAFB
FMR1
ILF3
CPSF6
ZC3H8
GTF2F1
YWHAG
METAP2
PTBP1
RBFOX2
NOLC1
WDR43
AKAP1
TAF15
PPIL4
SAFB2
HNRNPA1
LARP4
SUPV3L1
NIPBL
EIF3GL
E
~7kb RBPD
QKI
HNRNPL
ZC3H11A
SF3B1
TIA1
TIAL1
PTBP1
TBRG4
KHSRP
PABPC4
UCHL5
GEMIN5
HNRNPC
SAFB
PUM2
DHX30
WDR43
PRPF8
U2AF2
GPKOW
DDX42
SF3B4
XRCC6
F
XIST
UCHL5
HNRNPK
RBM15
SRSF1
UTP3
HNRNPU
SF3B1
EXOSC5
NOLC1
PTBP1
NIPBL
HNRNPM
HNRNPC
RBM22
ZNF622
ZC3H11A
TIA1
SDAD1
MATR3
ILF3
GRWD1
TRA2A
WDR43
RBFOX2
LARP4
G
chr6:141,025,140-141,228,929
203 kb
HepG2 RNAseq
K562 RNAseq
ASAR6-141
HepG2 eCLIP
K562 eCLIP
sgRNA*
sgRNA
eCLIP RBP binding peaks in HepG2 and K562

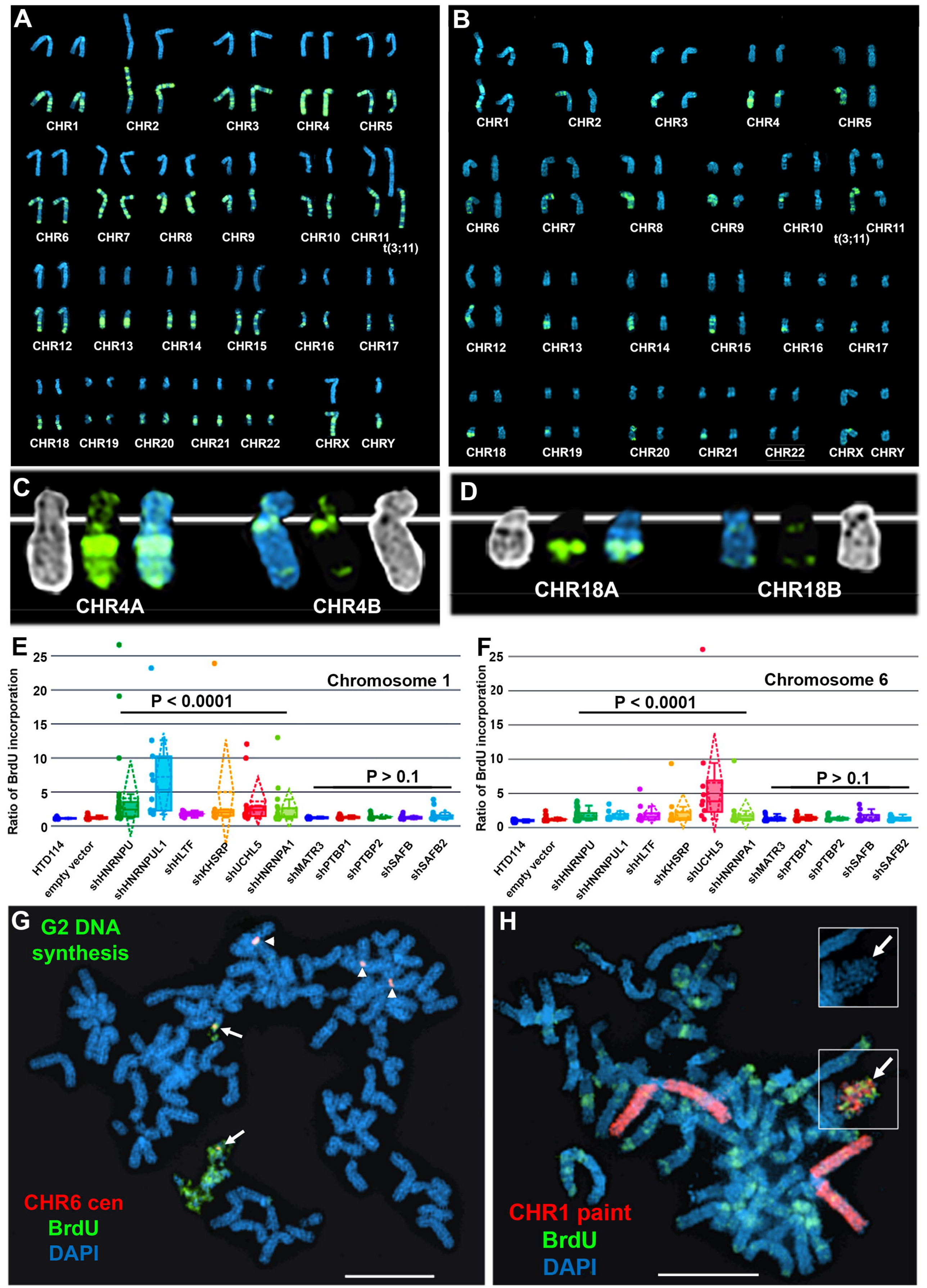


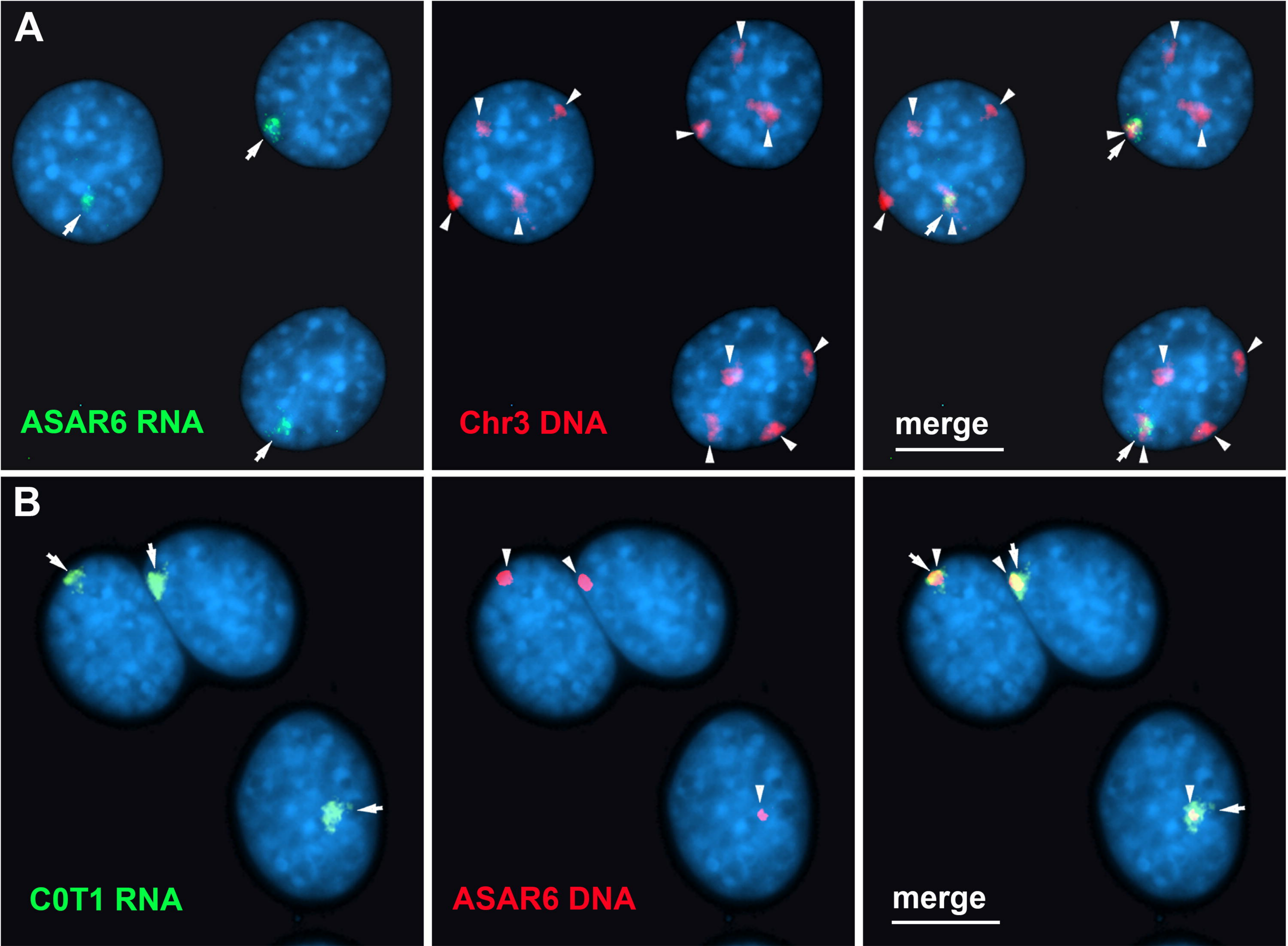
	Chromosome 6		
	CHR6B	CHR6A	ratio
	Pixels	Pixels	6B/6A
DAPI	20,060	21,594	0.92
BrdU	7,302	2,214	3.18


A
HNRNPU
empty vector
B
HNRNPU
shHNRNPU
C
HNRNPUL1
empty vector
D
HNRNPUL1
shHNRNPUL1
E
HNRNPC
empty vector
F
HNRNPC
shHNRNPC
G
HNRNPM
empty vector
H
HNRNPM
shHNRNPM
I
HLTF
empty vector
J
HLTF
shHTLF
K
shRNA - + - + - + - + - +
Protein Levels
Pixels (area x intesity)
(x1000)
600
400
200
0
HNRNPU
P = 0.00001
HNRNPUL1
P = 0.00001
HNRNPC
P = 0.00001
HNRNPM
P = 0.00001
HLTF
P = 0.00001
L
XIST RNA
ASAR6-141 RNA
empty vector
M
XIST RNA
ASAR6-141 RNA
shHNRNPU
N
XIST RNA
ASAR6-141 RNA
shHNRNPUL1
O
XIST RNA
ASAR6-141 RNA
shHNRNPC
P
XIST RNA
ASAR6-141 RNA
shHNRNPM
Q
XIST RNA
ASAR6-141 RNA
shHLTF
R
XIST RNA
ASAR6-141 RNA
shHNRNPA1
S
XIST RNA
ASAR6-141 RNA
shHNRNPL
T
XIST RNA
ASAR6-141 RNA
shKHSRP
U
XIST RNA
ASAR6-141 RNA
shUCHL5
V
XIST RNA
ASAR6-141 RNA
shPTBP1
W
XIST RNA Clouds
pixels (area x intensity)
400
300
200
100
0
empty Vector
shHNRNPA1 (P = 0.0001)
shHNRNPC (P = 0.0001)
shHNRNPL (P = 0.0001)
shHNRNPM (P = 0.0001)
shHNRNPU (P = 0.0001)
shHNRNPUL1 (P = 0.0001)
shHLTF (P = 0.0001)
shKHSRP1 (P = 0.0001)
shUCHL5 (P = 0.0001)
shPTBP1 (P = 0.03)


A
ASAR6-141
Xist
merge
ASAR6-141
Xist
merge
ASAR6-141
Xist
merge
B
Incorporation (pixels)
192
128
64
0
0 1.0 2.0 3.0 4.0 5.0 6.0 7.0
Distance (uM)
Incorporation (pixels)
192
128
64
0
0 1.0 2.0 3.0 4.0 5.0 6.0 7.0 8.0
Distance (uM)
Incorporation (pixels)
128
64
0
0 1.0 2.0 3.0 4.0 5.0 6.0 7.0
Distance (uM)
C
ASAR6-141
DAPI
D
ASAR6-141
ChrX
DAPI
E ASAR6-141
DAPI
F ChrX
DAPI
G BrdU
DAPI
H merge


A
ASAR1-187 RNA
ASAR6 RNA
empty vector
B
CHR1 BAC DNA
CHR6 BAC DNA
empty vector
C
merge
empty vector
D
shHNRNPA1
E
shHNRNPC
F
shHNRNPL
G
shHNRNPM
H
shHNRNPU
I
shHTLF
J
shKHSRP
K
shUCHL5
L
ASAR6 RNA
ASAR1-187 RNA
empty vector
M
CHR1 BAC DNA
CHR6 BAC DNA
empty vector
N
merge
empty vector
O
shHNRNPU
CHR1 BAC DNA
CHR6 BAC DNA
P
shHNRNPU
CHR1 BAC DNA
CHR6 BAC DNA
Q
shHNRNPU
CHR1 BAC DNA
CHR6 BAC DNA
ASAR1-187 RNA
ASAR1-187 RNA
ASAR1-187 RNA
ASAR6 RNA
ASAR6 RNA
ASAR6 RNA
merge
merge
merge


A
CHR1
CHR2
CHR3
CHR4
CHR5
CHR6
CHR7
CHR8
CHR9
CHR10
CHR11
t(3;11)
CHR12
CHR13
CHR14
CHR15
CHR16
CHR17
CHR18
CHR19
CHR20
CHR21
CHR22
CHRX
CHRY
B
CHR1
CHR2
CHR3
CHR4
CHR5
CHR6
CHR7
CHR8
CHR9
CHR10
CHR11
t(3;11)
CHR12
CHR13
CHR14
CHR15
CHR16
CHR17
CHR18
CHR19
CHR20
CHR21
CHR22
CHRX
CHRY
C
CHR4A
CHR4B
D
CHR18A
CHR18B
E
Chromosome 1
P < 0.0001
P > 0.1
Ratio of BrdU incorporation
0
5
10
15
20
25
HTD114
empty vector
shHNRNPU
shHNRNPUL1
shHLTF
shKHSRP
shUCHL5
shHNRNPA1
shMATR3
shPTBP1
shPTBP2
shSAFB
shSAFB2
F
Chromosome 6
P < 0.0001
P > 0.1
Ratio of BrdU incorporation
0
5
10
15
20
25
HTD114
empty vector
shHNRNPU
shHNRNPUL1
shHLTF
shKHSRP
shUCHL5
shHNRNPA1
shMATR3
shPTBP1
shPTBP2
shSAFB
shSAFB2
G
G2 DNA synthesis
CHR6 cen
BrdU
DAPI
H
CHR1 paint
BrdU
DAPI


A
ASAR6 RNA
Chr3 DNA
merge
B
C0T1 RNA
ASAR6 DNA
merge