Cell-Surface RNA Forms Ternary Complex with RNA-Binding Proteins and Heparan Sulfate to Recruit Immune Receptors

Zeshi Li[1, 2, 3, §,*], Bhagyashree S. Joshi[2, §], Hongbo Yin[3], Ruud H. Wijdeven[4, 5], Rue A. Koç[3], Dick Zijlman[6], Abhijeet Pataskar[7], Irene Santos-Barriopedro[6], Hailiang Mei[8], Wei Wu[3], Milad Shademan[1], Filip M. Zawisza[3], Eric Bos[9], Pradeep Chopra[10], Marvin Tanenbaum[2, 11], Thomas Sharp[12], Michiel Vermeulen[6, 13], Vered Raz[1, *] & Chirlmin Joo[2, 14, *]

[1] Human Genetics department, Leiden University Medical Centre, 2333ZC Leiden, The Netherlands
[2] Department of BioNanoScience, Kavli Institute of Nanoscience Delft, Delft University of Technology, 2629 HZ, Delft, The Netherlands
[3] Division of Chemical Biology & Drug Discovery, Utrecht Institute for Pharmaceutical Sciences, Utrecht University, 3584 CG Utrecht, The Netherlands
[4] Department of Cell and Chemical Biology, Oncode Institute, Leiden University Medical Center, 2333 ZA Leiden, Zuid-Holland, the Netherlands
[5] Department of Functional Genomics, Center for Neurogenomics and Cognitive Research (CNCR), Vrije Universiteit Medical Center, 1007 MB Amsterdam, Noord-Holland, the Netherlands
[6] Department of Molecular Biology, Faculty of Science, Radboud Institute for Molecular Life Sciences, Oncode Institute, Radboud University Nijmegen, Nijmegen, The Netherlands
[7] Division of Oncogenomics, Oncode Institute, The Netherlands Cancer Institute, 1006 BE Amsterdam, The Netherlands
[8] Sequencing Analysis Support Core, Department of Biomedical Data Sciences, Leiden University Medical Center, Einthovenweg 20, 2333 ZC, Leiden, The Netherlands
[9] Department of Cell and Chemical Biology, Leiden University Medical Center, Leiden 2300 RC, The Netherlands
[10] Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602, U.S.A.
[11] Oncode Institute, Hubrecht Institute-KNAW and University Medical Center Utrecht, Uppsalalaan 8, 3584 CT Utrecht, the Netherlands
[12] School of Biochemistry, University of Bristol, Bristol BS8 1TD, U.K.
[13] Division of Molecular Genetics, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[14] Department of Physics, Ewha Womans University, Seoul 03760, Republic of Korea
§: equal contribution
*: corresponding authorship

Abstract

Recent discoveries have shown the presence of RNA molecules on the cell surface, defying the traditional view that RNA only functions intracellularly. However, it is not well understood how cell-surface RNA (csRNA) is stably present on the plasma membrane and what functions it performs. We answer the pressing questions in the emerging field by taking integrated omic-wide approaches and multiple orthogonal validatory methods. Firstly, we exploited the RNA-sensing ability of TLR7 as a specific recombinant probe to detect csRNA. Coupling it with a genome-wide CRISPR-Cas9-knockout screening, we identified heparan sulfate (HS) as a crucial factor for RNA presentation on cells. Using the TLR7 probe, cell surface proximity labelling revealed that these HS-associated csRNAs (hepRNAs) are in vicinity with a plethora of RNA-binding proteins. The compelling observation led us to a molecular model where HS, RNA and RBP form ternary complexes at cell surface. A photochemical RNA-protein crosslinking technology termed SCOOPS were then established to validate the termolecular model in a TLR7-orthogonal manner. Moreover, enabled by SCOOPS, we unveiled identities of hepRNA using next-generation sequencing, and identified traits in RNA primary structures that facilitate HS association. We further show that hepRNA binds to killer cell immunoglobulin-like receptor 2DL5 (KIR2DL5), recruiting the protein to cell surface and potentially enhancing receptor-ligand interactions. Our findings provide a foundation for exploring how cell-surface ribonucleoproteins contribute to immune modulation.

Introduction

Ribonucleic acid (RNA) performs many critical cellular functions in both prokaryote and eukaryote cells. Although the localization of RNA was considered primarily intracellular in mammalian cells, emerging studies have indicated that RNA can be displayed on the cell surface. A previous study found that bacterial genome-encoded RNA can be anchored on their cell surface via binding to a protein.[1] Very recent studies have shown the presence of RNA on the extracellular part of mammalian cell plasma membranes.[2-5] These cell-surface-associated RNAs (csRNAs) include fragments of messenger and long non-coding RNAs and putative glycosylated small RNAs. There are indications that csRNAs can participate in molecular recognition, thereby modulating cell-cell interactions. A recent study has shown csRNA mediated neutrophil recruitment by interacting with P-selectin on the endothelial cells.[6] Despite the aforementioned studies, mechanistic understanding as to how RNAs are stably presented on the mammalian cell surface remains limited. It is also unclear if csRNA may exhibit other modes of function besides presenting glycan moieties to interacting cells.[2,6]

As most RNA molecules are associated with RNA-binding proteins (RBPs), it can be speculated that csRNA is similarly associated with RBPs, providing a mechanism for cell surface anchoring. In line with this hypothesis, both earlier and recent studies have demonstrated RBPs, previously thought to be localized only intracellularly, are also present on the plasma membrane.[7-10] Thus, the potential binding partners are indeed available for and in proximity to RNA on the cell surface. However, these RBPs are generally soluble proteins that do not possess a transmembrane domain and are therefore unlikely to be directly anchored on the lipid bilayer like many cell-surface glycoproteins. For example, a classical nuclear RBP, U5 SNRNP200, has been evaluated as an acute myeloid lymphoma-associated cell surface antigen for immunotherapy.[9, 10] This RBP was shown to be associated with the Fcγ receptor IIIA. A more recent study revealed that a plethora of RBPs can localize to the cell surface, as possible carriers of csRNA.[11] However, it is largely unknown how csRNA and RBPs become stably presented on the plasma membrane, and if so, whether they are associated with each other.

Animal cells employ highly negatively charged linear carbohydrate polymers, such as glycosaminoglycans, as a scaffold to enrich soluble proteins and organize membrane bound proteins on the plasma membrane.[12] Heparan sulfate (HS) is an important class of glycosaminoglycans ubiquitously found on plasma membrane and in extracellular matrix (ECM).[13] Portions of HS are heavily sulfated, and these sulfated domains serve as docking sites for a broad array of proteins, including growth factors, morphogens, amyloid proteins, lipoproteins, cytokines, and chemokines, as well as many microbial proteins. HS is characterized by having an extension of repeating disaccharide units, glucosamine-α1,4-glucuronic acid (GlcN-GlcA) or glucosamine-α1,4-iduronic acid (GlcN-IdoA). An HS chain can contain over a hundred monosaccharide units. Glycan modifications such as N- and O-sulfation introduce negative charges to portions of the HS chain and are known to be crucial for the biological functions of HS. The pattern of N- and O-

91 sulfation, also known as the sulfation code, critically determines what proteins interact with HS.
92 These features make HS among a main director of cell surface interactions.[14]
93 Herein we demonstrate that HS functions as a scaffold for presenting RNAs and RBPs on
94 the cell surface. A key tool in our study was toll-like receptor 7 (TLR7), which we leveraged as a
95 probe for csRNA. We established that the recombinant TLR7-Fc fusion protein can effectively
96 detect and capture RNA on living cells, enabling a genome-wide knockout screening to identify
97 genes regulating csRNA localization. The screening revealed HS biosynthetic enzymes as
98 essential factors. Using the same probe, we also identified the csRNA-proximal proteome, which
99 includes both classical and non-canonical RBPs. These findings led us to propose a model in
100 which RNA, RBP, and HS form ternary complexes for csRNA presentation. To further validate this
101 ternary configuration, we developed an TLR7-independent approach—Spatio-selective
102 Crosslinking followed by Orthogonal Organic Phase Separation (SCOOPS). SCOOPS employs
103 local photocatalytic generation of singlet oxygen (SO) to crosslink RNA with bound proteins,
104 facilitating selective isolation of csRNA. By combining SCOOPS with next-generation sequencing,
105 we identified HS-associated csRNAs and uncovered primary structural imprints contributing to their
106 localization. Finally, we demonstrated that HS-associated csRNAs can recruit killer cell
107 immunoglobulin-like receptor 2DL5 (KIR2DL5) to the cell surface, suggesting a potential role for
108 csRNA in modulating immune interactions between KIR2DL5 on immune cells and its cognate
109 ligand on target cells.

Results

Establishment of TLR7-Fc as a probe for csRNA

The sequence of RNA molecules on the plasma membrane likely varies between different cells, and RNA molecules present at low copy numbers may elude detection.[15] The state-of-the-art csRNA labeling approaches either require fixation or rupturing of the plasma membrane[3, 15] and are therefore not compatible with live cells. To overcome this, we reasoned that the use of an RNA-recognizing probe with low sequence specificity, rather than a hybridizing probe which requires prior knowledge of csRNA sequence, should boost the chance for csRNA detection, and yield strong signals. We, therefore, exploited nature's RNA-sensing machinery: toll-like receptor 7 (TLR7),[16] an endosome-localized, pattern recognition receptor in select immune cells to sense both foreign and endogenous single-stranded RNA molecules.[17-19] TLR7 requires two consecutive uridine moieties for strong RNA binding, but requires only few other sequence features.[20] We used the ectodomain of TLR7 (Ala27-Asn839) fused to the human Fc tag (denoted hereafter TLR7-Fc). The former is expected to recognize csRNA on intact (live) cells, and the latter will allow for the detection by a secondary antibody for different downstream applications.

We employed confocal microscopy to examine the efficiency of TLR7-Fc to detect csRNA. TLR7-Fc is applied onto intact, fixed cells without cell permeabilization, followed by washing and the addition of a fluorophore-conjugated, anti-human IgG secondary antibody. To check for localization of TLR7-Fc staining, we labeled the cell surface with biotin-conjugated wheat germ agglutinin (WGA-biotin),[21] a lectin that recognizes N-acetylneuraminic acid and N-acetylglucosamine residues present in cell surface glycans, and the biotin moieties were detected by fluorescent streptavidin. We observed that bright puncta given rise to by TLR7-Fc probe in complex with the fluorescent secondary antibody align with the rims of HeLa, U2OS, and Mel526 cells (**Figure 1a** and **Supplementary Figure 1a** and **b**). This is akin to many cell surface proteins, which manifest in clusters.[22, 23] WGA signal remained confined to the cell surface. csRNA signal was absent in controls devoid of TLR7-Fc but only including the secondary antibody (**Supplementary Figure 1c**). A treatment with RNaseA/T1 cocktail in fixed, non-permeabilized cells, severely disrupted the TLR7 signal (**Figure 1a**), while DNase treatment did not (**Supplementary Figure 1d**), validating the RNA specificity of TLR7 binding. Moreover, the pre-saturation of TLR7-Fc with an exogenous, synthetic single-stranded RNA competitively inhibited TLR7 binding, further supporting the specificity of TLR7 to probe RNA on cells (**Figure 1a**).

Live-cell staining with TLR7 showed the same puncta pattern, indicating that the clustering of our probe on the cell surface is unlikely to be caused by cell fixation (**Supplementary Figure 1e**).[24-26] Interestingly, the monoclonal antibody J2, previously used for intracellular detection of endogenous dsRNA[27] and viral dsRNA,[28] as well as for cell-surface glycoRNA,[2] did not give detectable signals on HeLa or U2OS cell surface (**Supplementary Figure 1f**). Next, to observe csRNA in a more detailed subcellular context, we employed electron microscopy (EM), leveraging the TLR7-Fc probe and Protein-A-gold (10 nm) as a secondary probe (**Supplementary Figure 1g**

148 and **h**). We observed csRNA signals along the cellular periphery, cell-cell junctions, and
149 endocytosed vesicles, confirming the findings with fluorescent microscopy.
150 We further used our TLR7-Fc probe to quantify csRNA and its sensitivity to RNase
151 treatment with flow cytometry (**Figure 1b** and **Supplementary Figure 2a**). Corroborating the
152 fluorescence imaging results, when live cells were pre-treated with RNase A, the signal intensity
153 given rise by TLR7-Fc in complex with the fluorescent secondary antibody dropped dramatically
154 close to the background. To further confirm that TLR7-Fc binds to RNA on the cell surface, the
155 prior csRNA-depleted cells by RNase were incubated with exogenous, purified total cellular RNA
156 (**Supplementary Figure 2b**). We found that TLR7 binding could be partly restored by exogenous
157 RNA in a concentration-dependent manner (**Figure 1c**).
158 To further validate that TLR7 binds to RNA on the cell surface, we performed in-situ RNA-
159 protein crosslinking with UVC after TLR7-Fc probe addition, followed by RNA precipitation and
160 RNA 3'-end labeling (**Figure 1d**). Upon UVC irradiation, csRNA is expected to be covalently bound
161 to the TLR7-Fc probe and can be co-precipitated with the probe using protein A-functionalized
162 magnetic beads. The precipitated RNA was then labeled at the 3'-end with biotin-cytidine using
163 RNA ligase 1. This was followed by SDS-PAGE and Western blot for biotin. As a negative control,
164 cells were UVC irradiated before TLR7-Fc incubation and subjected to the same procedure. UVC-
165 crosslinked biotinylated csRNA is expected to remain bound to TLR7 throughout the entire
166 procedure, whereas the csRNA not crosslinked to TLR7 will dissociate upon denaturation, and
167 therefore will not remain on the membrane.[29] We observed a strong chemiluminescent signal
168 above TLR7-Fc molecular weight for crosslinked csRNA-TLR7 (**Figure 1e**), but not for non-
169 crosslinked TLR7 probe. The results demonstrate TLR7 probe indeed binds to RNA on the cell
170 surface.
171 To confirm TLR7-Fc binding to RNA is mediated by TLR7 ectodomain, but not other
172 portions of the protein (such as the Fc tag), we performed an in-vitro RNA crosslinking and co-
173 precipitation assay using the TLR7-Fc and human IgG control. Fragmented, purified total cellular
174 RNA was incubated and crosslinked (irradiated with UVC, 254 nm) with TLR7-Fc and IgG, pulled
175 down on beads and intensively washed, labeled with Cy5 at 3'-end, and finally released from beads
176 by proteinase K digestion. Purified precipitated RNA were then analyzed in agarose gel
177 electrophoresis. Strong fluorescent bands were observed for TLR7-Fc captured RNA, but not for
178 IgG control (**Supplementary Figure 2c**). The results indicated the Fc tag did not non-specifically
179 bind to RNA, demonstrating the RNA specificity of our probe.
180
181 **Genome-wide screening identified essential factors for csRNA presentation**
182 Previous studies have only postulated that csRNA can be presented on the plasma membrane,
183 yet a detailed mechanistic understanding of this process is lacking. With TLR7-Fc as a general
184 csRNA probe in hand, we next sought to uncover the molecular underpinnings of csRNA
185 presentation. We performed a pooled genome-wide, CRISPR-Cas9-mediated knockout (KO)

186 screening[30] to identify essential cellular factors for csRNA stable presentation. TLR7-Fc in complex
187 with a fluorescent secondary antibody was used to enrich the csRNA^{low} population with
188 fluorescence-assisted cell sorting (FACS). Such FACS-based enrichment was performed twice.
189 The enriched cells together with the non-enriched input cells were subjected to deep sequencing
190 to identify the guide RNAs of CRISPR-Cas9 giving rise to the csRNA^{low} phenotype. (**Figure 2a**,
191 see **Supplementary Figure 3** for gating strategies). Likewise, the csRNA^{high} phenotype was also
192 enriched and sequenced (See **Supplementary Text 1** for details, and **Supplementary Table 1**
193 for full results).
194 Data analysis revealed that enzymes involved in heparan sulfate (HS) biosynthesis[31] were
195 among the most highly scored hits, suggesting HS is a pivotal molecular factor for the stable cell-
196 surface display of RNAs (**Figure 2b**, full list of hits in **Supplementary Table 2**). Our hits include
197 enzymes for the core tetrasaccharide synthesis (XYLT2, B4GALT7, B3GALT6, and B3GAT3),
198 glycan chain polymerizing enzymes (EXT1, EXT2, and EXTL3), as well as glycan chain modifying
199 enzymes such as sulfotransferases (NDST1, HS2ST1, and HS6ST1) and a xylose kinase
200 (FAM20B) (**Figure 2c**). These hits suggest the stable presentation of csRNA requires a fully
201 extended HS chain and appropriate N- and O-sulfation pattern.
202 Apart from the HS-related glycosyltransferases and glycan-modifying enzymes, we also
203 identified regulators in HS biosynthesis. For example, UGDH oxidizes uridine diphosophate linked
204 glucose (UDP-Glc) to UDP-GlcA, a glycosyl donor for HS polymerization; SLC35B2 is a Golgi
205 membrane transporter of the sulfo-donor, adenosine 3'-phospho-5'-phosphosulfate;[32] SLC10A7 is
206 crucial for cellular calcium regulation and was implicated in HS biosynthesis;[33] and C3orf58 has
207 been recently shown to regulate HS biosynthesis.[34] The genes related to the N-glycan biosynthetic
208 pathway were not among the candidates, suggesting the csRNA detected by TLR7 is independent
209 from glycoRNA, and it does not require N-glycosylation to access plasma membrane.
210
211 **Validation of genome-wide screening**
212 Next, our attention was focused on validating that HS is necessary for a display of the
213 csRNA detected by the TLR7-Fc probe. csRNA location closely correlated with the heparan sulfate
214 staining on the cell surface showing their tight proximity (**Figure 2d**). We then generated HS-
215 deficient mutant cells lacking either the elongation enzyme EXT2 or sulfotransferase HS6ST1 via
216 CRISPR-Cas9 mediated KO, each with two high-scoring guide RNAs, and used flow cytometry to
217 quantify TLR7-Fc binding. Both guide RNAs for EXT2 resulted in an almost complete loss of TLR7
218 binding. One guide RNA for HS6ST1 led to a similar phenotype as did EXT2 KO, whereas the
219 other guide RNA reduced TLR7 binding down to approximately 30% of the wildtype (**Figure 2e**).
220 In consistency with the flow cytometry experiments, confocal microscopy also showed a dramatic
221 reduction in csRNA signal around the cell rim after knocking out EXT2, HS6ST1, EXTL3, and
222 C3orf58 (**Supplementary Figure 4a**). The loss of TLR7 binding to the cell rim was also observed
223 in Chinese hamster ovary (CHO) cell mutants (pgsD-677 and pgsE-606) deficient in HS

224 biosynthesis (**Supplementary Figure 4b**). The wildtype CHO cell showed a similar punctate
225 pattern as did Mel526.
226 An alternative scenario which might have given rise to similar hits in our genome-wide
227 screening would be that TLR7-Fc directly binds to HS on cell surface (**Supplementary Figure 4c**).
228 Thus, the loss of TLR7 binding to cell surface would have been caused by the RNase non-
229 enzymatically and competitively suppressing TLR7-Fc binding to HS. To exclude this, we
230 compared the TLR7-Fc to well-studied HS-binding proteins using flow cytometry, including
231 fibroblast growth factor-2 (FGF2) and an anti-HS IgM 10E4.[35] All proteins bound to wildtype cell
232 surface strongly, but hardly to EXT2-KO cells, confirming the HS dependency of the tested proteins
233 (**Supplementary Figure 4d**). However, the two HS binders differed from TLR7 by that they were
234 refractory towards RNase treatment. In addition, we performed surface plasmon resonance binding
235 studies of the RNase (Purelink RNase A) used in the above studies with heparin-coated chips
236 (**Supplementary Figure 4e**). No signal was detected across all RNase concentrations, suggesting
237 the RNase used in the study did not bind to HS, therefore unlikely to competitively suppress other
238 proteins for HS binding. These results strengthened TLR7 as a csRNA probe and demonstrated
239 its loss of binding to cell surface was likely due to the RNase-mediated digestion of csRNA.
240 To further validate the genetic screen results, we selected two HS-deficient mutant cells
241 for TLR7 binding rescue experiment by exogenous total RNA. Cells were first treated with RNase
242 A, and then with total RNA fragments at varying concentrations (**Supplementary Figure 4f**). The
243 treated cells were then incubated with TLR7-Fc followed by a fluorescent secondary antibody. The
244 TLR7 binding remained close to the background throughout all experimental groups regardless of
245 the RNA concentrations. These results demonstrated that cells lacking an extended HS chain or
246 an appropriate sulfation pattern become refractory to exogenous RNA incubation and can no
247 longer rescue TLR7 binding, suggesting that csRNA presentation requires intact HS. Hereafter,
248 we refer the csRNA detected by the TLR7-Fc probe as heparan sulfate-associated RNA (hepRNA).
249
250 **RNA, RBP and HS are in vicinity on cell surface**
251 Both RNA and HS are highly negatively charged biopolymers. How are these two macromolecules
252 associated? We were struck by the observation that upon proteolytic detachment of adherent cells
253 by TryPLE before flow cytometry, TLR7-Fc no longer detectably bound to the cell surface
254 (**Supplementary Figure 5a**), while non-tryptic detachment methods such as EDTA allowed for
255 hepRNA detection by TLR7-Fc (**Figure 1b**). The observation led us to hypothesize that hepRNA
256 may require an additional factor, such as proteins, to be presented on the cell surface HS
257 (**Supplementary Figure 5b**). We, therefore, applied a peroxidase-mediated proximity labeling
258 approach[36] to identify the hepRNA-proximal proteome using TLR7-Fc. Peroxidase-mediated
259 proximity labeling has been shown to exhibit particularly high efficiency within the radius of ~20 nm
260 due to the direct contact of peroxidase catalytic center with proximal proteins. Decreased, but
261 considerable labeling does occur within a radius of ~270 nm,[37, 38] due to the diffusion of peroxidase-

262 activated biotin-phenol radical.[38] Thus, peroxidase-based proximity labeling is expected to afford
263 protein candidates that are direct binders of hepRNA, and those that do not physically interact with
264 hepRNA, but are spatially close to it.

265 We pre-complexed TLR7-Fc with horse radish peroxidase (HRP)-conjugated protein A,
266 which was then incubated with untreated and RNase A-pretreated Mel526 cells (**Figure 3a**).
267 RNase treatment is expected to digest unprotected regions of all csRNA, whereas the portions
268 proximal to or bound by a protein are protected from digestion (**Figure 3b**), an effect commonly
269 exploited to study proteins' footprint on an RNA.[29, 39] As an isotype control, human IgG was used
270 instead of TLR7-Fc. Biotinylation of proteins on the cell surface was confirmed by fluorescence
271 imaging and Western blot (**Supplementary Figure 5c** and **d**). Following proximity labeling and
272 non-denaturing cell lysis, biotinylated cell-surface proteins were enriched with streptavidin-
273 functionalized agarose resin. Bound proteins were then on-bead digested with trypsin and
274 subjected to LC-MS/MS and label-free quantitative proteomics analysis.

275 By comparing the TLR7-proximal proteome of cells without RNase treatment (denoted as
276 'TLR7') to that of the IgG isotype control, we identified 315 proteins as significantly enriched hits
277 (fold change ≥ 4 AND p value < 0.01. Full list of identified proteins in **Supplementary Table 3**).
278 Of the significantly enriched hits, 137 of them carry the gene ontology (GO) term 'cell surface'
279 (GO:0009986) or 'extracellular region' (GO:0005576), including known glycoproteins and
280 glycosylphosphatidylinositol-anchored proteins (**Figure 3c**). 76 of the cell-surface (glyco)proteins
281 were not anymore enriched over IgG control in the RNase-treated cell dataset (denoted as
282 'RNase') (**Supplementary Figure 6a**), which suggested upon extracellular RNase digestion,
283 hepRNA is no longer proximal to the (glyco)proteins.

284 Across the TLR7 and RNase datasets, a considerable number of proteins remained
285 consistently enriched, with a similar fold change (vs. IgG control) in both datasets (**Figure 3d** and
286 **Supplementary Figure 6b**. See full list in **Supplementary Table 4**). In the RNase sample, these
287 proteins likely had been proximity-biotinylated by the residual TLR7-Fc:protein A-HRP complex
288 bound to the RNase-truncated hepRNA (**Supplementary Figure 6c**). We queried RBP2GO, a
289 database documenting validated and putative RBPs with the consistent hits from both TLR7 and
290 RNase samples. It returned with a match of 137 proteins. Among these RBP2GO proteins, 34
291 carried the GO term 'RNA binding' (GO:0003723). The RBP2GO-documented hits include classical
292 RNA-binding proteins such as heterogeneous nuclear ribonucleoprotein A1 and A2B1 (HNRNPA1
293 and HNRNPA2B1, respectively), and Sjögren syndrome antigen B (SSB), all of which bear a high
294 RBP2GO score, indicating high confidence for direct RNA binding. Non-canonical RNA binding
295 proteins that do not possess a classical RNA-binding domain were also among the hits, many of
296 which are metabolic enzymes or chaperones with moonlighting RNA-binding functions, such as
297 enolase 1 (ENO1)[40] and calreticulin (CALR).[41] These proteins normally are scored moderately in
298 RBP2GO database.

299 We next asked if the detection of the RBPs had been because they were on the cell surface
300 or a result of non-specific labeling/enrichment of intracellular components. We selected both
301 classical and non-canonical RBP hits spanning a wide range of RBP2GO scores for
302 immunofluorescence imaging to check their localization (**Figure 3e – h**). All proteins tested,
303 including HNRNPA1, La protein, PA2G4 and ENO1 were found on cell surface, and they were
304 partially colocalizing with TLR7-Fc, suggesting hepRNA and a fraction of these RBPs are in vicinity.
305 Flow cytometry analysis of cell-surface HNRNPA1 and ENO1 (**Supplementary Figure 6d**) with
306 the EXT2-KO mutant cells showed that HS deficiency resulted in a loss of the cell-surface localized
307 RBPs (csRBPs). Western blot of whole cell lysate from wildtype and EXT-KO cells confirmed the
308 absence of csRBP was due to the lack of HS as a scaffold on cell surface, but not a global reduction
309 of protein expression because of HS deficiency (**Supplementary Figure 6e**).
310 The spatial proximity of csRBP and hepRNA prompted us to examine the formation of
311 ribonucleoprotein complexes (RNPs). Recent years have seen methodological innovation on facile
312 isolation of RNA-bound proteins coupled with MS proteomics to study their dynamics in cells.
313 Several acid guanidinium thiocyanate-phenol-chloroform (AGPC) phase separation strategies[42-44]
314 leverage the amphiphilicity of UV-crosslinked RNPs for isolation and are independent of RNA
315 sequence. Although UV crosslinking is applied to the whole cell, selectively tagging the surface
316 proteome on intact cells should enable enriching RNA-bound proteins at cell-surface after AGPC
317 phase separation (**Supplementary Figure 6f**). If the RBPs are not bound by RNA, the former
318 should be washed away during the phase separation procedure. We employed a cell-impermeable,
319 amine-reactive biotinylation agent to tag the surface proteome, after which we performed UVC
320 crosslinking and orthogonal organic phase separation (OOPS)[42] to isolate the crosslinked whole-
321 cell RNPs. The total RNA-bound proteins were then released from the cleaned interphase by
322 RNase digestion and a subsequent APGC extraction into a new organic phase and finally
323 subjected to biotin-based affinity pulldown for label-free quantitative proteomics. A non-
324 biotinylated, but crosslinked sample was taken along to control for non-specific bead binding by
325 the AGPC-denatured proteins.
326 649 proteins were significantly enriched over the no-biotin control (fold change $\geq$ 8 AND p
327 value $<$ 0.01) (**Supplementary Table 5**). Enriched GO molecular function terms were heavily
328 related to RNA, demonstrating the effectiveness of the procedure to isolate RBPs (**Supplementary**
329 **Figure 6h**). Despite the vast difference in the working principles and chemo-selectivity between
330 the TLR7-based proximity labeling and biotinylation-crosslinking-OOPS methods, the latter
331 validated 41 proteins in the TLR7-proximal datasets (TLR7 vs. IgG) as RNA-bound proteins on cell
332 surface (**Supplementary Figure 6g**) in a high-throughput manner. Taken together, the series of
333 proteomics and validatory studies pointed to a molecular configuration of hepRNA, csRBP and HS,
334 in which these three types of biomolecules form ternary complexes.

335 **HS's essential role in presenting hepRNA-csRBP complexes**
336 We have demonstrated the RNA's dependency on and the colocalization with HS, and that these
337 hepRNAs are in the state of RNP complexes. Next, we sought to further validate that the three
338 biomolecules form ternary complexes on cell surface. We introduce herein a facile strategy for cell-
339 surface RNP isolation, termed Spatio-selective Crosslinking followed by OOPS (SCOOPS).
340 SCOOPS allows for one-step isolation of crosslinked RNPs within a subcellular region, without
341 separate tagging and affinity pulldown steps. It is based on our original, unexpected finding that
342 singlet oxygen (SO), a highly reactive oxygen species, can lead to RNA-protein crosslinking
343 (**Supplementary Figure 7**. See **Supplementary Text 2** for a detailed description), an effect
344 analogous to conventional UVC irradiation. However, unlike UVC crosslinking, which can only be
345 applied to whole cells, numerous reports found SO can be generated locally within a subcellular
346 region due to its transient nature (half-life in water ~4 μs[45]) to achieve spatio-selective chemistry.[46-
347 48] Therefore, we envisioned that by tethering a SO generator to cell surface, locally produced SO
348 could lead to crosslinking of RNA to bound proteins selectively at cell surface. Subsequently,
349 crosslinked cell-surface RNA-protein complexes were brought into the AGPC interphase for
350 isolation by OOPS.
351 To establish SCOOPS for cell surface biomolecules, we tethered a small-molecule SO
352 generator, eosin Y (EY) onto streptavidin (strep-EY). Strep-EY retained full binding capacity to
353 biotin (**Figure 4a** and **b**) and can be recruited to live cell surface by WGA-biotin (**Figure 4c**). We
354 expect the AGPC interphase to contain crosslinked csRNA-protein only when csRNA is intact AND
355 when both WGA-biotin and strep-EY are present, while missing out any component would afford
356 little, if any, csRNA after SCOOPS (**Figure 4d**). To perform SCOOPS, cells were incubated
357 sequentially with WGA-biotin and strep-EY and were subjected to AGPC immediately after green
358 light irradiation.
359 The crosslinked csRNA (pink RNA in **Figure 4e**) was then released by proteinase digestion
360 of the AGPC interphase after repetitive washes. The released csRNA was then fluorescently
361 ligated at 3'-end and subjected to agarose gel electrophoresis. A small portion of total RNA (blue
362 RNA in the first aqueous phase in **Figure 4e**) was also taken along and processed in the same
363 fashion to ensure the inputs for AGPC are consistent across different samples. Strong fluorescent
364 signal of released interphase RNA was observed only when both strep-EY and WGA-biotin are
365 present (the fourth lane of **Figure 4g**). Leaving out either component afforded much weaker signals
366 (the second and third lanes). Pretreating live cells briefly with extracellular RNase to degrade
367 csRNA (first lane) resulted in a significantly weakened signal compared to untreated, csRNA-intact
368 cells.
369 We then employed SCOOPS to validate HS's essential role in presenting hepRNA-RBP
370 complexes at cell-surface. We expected cells deficient in HS biosynthesis should afford less
371 SCOOPS-isolated RNA due to the absence of the intact polysaccharide chain as a scaffold for
372 hepRNA presentation (**Figure 4f**). Indeed, we observed a significant reduction of SCOOPS-

373 isolated RNA fluorescence from HS-deficient cells as compared to the wildtype in the agarose gel
374 (**Figure 4h**). To ensure the reduced signal was caused by the biology rather than differences in
375 SO generator recruitment on cell surface, we performed flow cytometry assays to quantify recruited
376 strep-EY on wildtype and HS-deficient cells (**Figure 4i**). We found strep-EY introduction to cell
377 surface was at a comparable level between wildtype and HS-deficient cells. Taken together, the
378 results demonstrated that there is a portion of SCOOPS-isolated RNA which becomes largely
379 absent on HS-deficient cells, suggesting hepRNA's dependency on the intactness of HS for a
380 stable presentation on cell surface. Thus, SCOOPS validated the observation made with the TLR7
381 probe in an orthogonal manner.
382
383 **hepRNA identification by sequencing of SCOOPS-isolated RNAs**
384 Provided that hepRNA could be potentially captured by SCOOPS, we sought to identify these RNA
385 species using next-generation sequencing. Bioanalyzer traces revealed that SCOOPS-isolated
386 RNA in wildtype cells had a wide length distribution (**Supplementary Figure 8a**), with a bulge in
387 electropherograms spanning a large range of recorded length. Such bulge was much less
388 prevalent in the HS-deficient mutant. The electropherograms of SCOOPS-isolated RNA in wildtype
389 cells were vastly different from total RNA, as the characteristic peak pattern as observed in the
390 latter was almost lost. To cover both polyadenylated and non-polyadenylated RNA biotypes in the
391 sequencing library, we took a ligation-based strategy to introduce sequencing adapters
392 (**Supplementary Figure 8b**). The resulted product was reverse-transcribed, ligated with a second
393 adapter, amplified using polymerase chain reaction (PCR). The PCR product then underwent
394 polyacrylamide gel-based size selection (**Supplementary Figure 8c**). The size-selected libraries
395 were subjected to Illumina platforms for sequencing.
396 The analysis of sequencing data revealed large difference between SCOOPS-isolated
397 RNA and the total transcriptome (see **Supplementary Table 6** for a complete RPM table), as
398 demonstrated by the principal component analysis (**Supplementary Figure 9e**). Comparing
399 SCOOPS-isolated RNA with total RNA (SCOOPS_WT vs. total_WT) can identify the transcripts
400 which were particularly well retained in the interphase due to crosslinking. These transcripts are
401 colored as yellow dots in the scatter plot (**Figure 5a**). It is important to note that not all RNAs in the
402 interphase were crosslinked ones on the cell surface: as demonstrated in the biochemical assays
403 (**Figure 4g**) and bioanalyzer results (**Supplementary Figure 8a**), there were also background
404 RNA species retained in the interphase even without crosslinking (see **Supplementary Text 3** for
405 a detailed explanation).
406 We asked if any hepRNA had been retained in the SCOOPS interphase. We define
407 hepRNA as transcripts that are substantially reduced or absent in the SCOOPS interphase of
408 EXT2-KO mutant (SCOOPS_KO), when compared to those in the wildtype SCOOPS interphase
409 (SCOOPS_WT). As illustrated in the connected box plot (**Figure 5b**), of the well-retained
410 transcripts in the SCOOPS_WT interphase (yellow dots in **Figure 5a**), many had substantially

411 reduced counts in SCOOPS_KO samples. These transcripts include mainly mRNA and lncRNA,
412 which have been major csRNA biotypes in previously reported datasets.[3, 4] In particular, MALAT1,
413 as one of the few validated cell-surface lncRNA, was also highly enriched in the SCOOPS
414 interphase as hepRNA. The results suggested SCOOPS effectively captures RNA at cell surface,
415 and hepRNA is indeed present in the SCOOPS interphase.

416 To comprehensively identify hepRNA candidates, we first normalized the read counts of
417 SCOOPS_WT and SCOOPS_KO on the background RNA species (see **Supplementary Text 4**
418 for definitions and rationale). It is crucial to note that while SCOOPS is designed to crosslink all
419 RNA-protein complexes at the site of SO generation, it is not specific to hepRNA. The key to
420 ensuring hepRNA specificity was the use of HS-deficient cells as a negative control. In this setup,
421 RNA species that are 1) present at the cell surface but independent of HS or 2) non-specifically
422 crosslinked or isolated during SCOOPS would show minimal enrichment when comparing
423 SCOOPS_WT with SCOOPS_KO datasets.

424 A higher fold change (FC > 4, SCOOPS_WT over _KO) indicates that a transcript's cell-
425 surface presentation is strongly dependent on HS rather than other mechanisms, classifying it as
426 a high-confidence hepRNA candidate. This criterion applies to many mRNA, lncRNA, and snoRNA
427 transcripts (**Figure 5c** and **d**, see **Supplementary Table 7** for a full candidate list). In contrast,
428 transcripts with a lower fold change (2 < FC < 3) are considered low-confidence hepRNA
429 candidates, as they may associate with HS only partially or could partly result from non-specific
430 crosslinking or isolation during SCOOPS. These transcripts span nearly all investigated biotypes.
431 However, mitochondrial tRNA (Mt_tRNA) and vault RNA are unlikely to be hepRNA, as their
432 transcripts were barely enriched in the SCOOPS_WT vs. SCOOPS_KO comparison.

433 While mRNA and lncRNA biotypes dominate hepRNA candidates, which is consistent with
434 previous studies, strikingly, snoRNAs also constitute a considerable portion of hepRNA. Among
435 the high-confidence snoRNA-derived hepRNA candidates, the C/D box type outnumbers the
436 H/ACA type (21 SNORD vs. 4 SNORA), both in transcript count and total read coverage (**Figure**
437 **5e**). Remarkably, we identified defining structural traits that distinguish high-confidence SNORD
438 hepRNAs (hepSNORDs). Firstly, high-confidence hepSNORDs contain significantly less
439 guanosine (G) and cytidine (C) compared to low-confidence/non-hepSNORDs (**Figure 5f**). Such
440 nucleotide composition suggests high-confidence hepSNORDs may harbor less duplex or
441 structured region intra- or intermolecularly.[49, 50] Furthermore, in high-confidence hepSNORDs, the
442 occurrence of consecutive uridines (UU), a motif known for strong TLR7 binding,[20] is significantly
443 more frequent than low-confidence/non-hepSNORDs (**Figure 5g**). This compelling observation
444 reveals that specific RNA primary structures underlie the ability of this subset of RNA to associate
445 with cell-surface HS. Beyond structural implications, our findings based on a TLR7-independent
446 method provide an explanation at RNA primary structure level for TLR7's selectivity for hepRNA.

447 **hepRNA recruits immune receptor on cell surface**

448 The vicinity of hepRNA to cell surface glycoproteins cell surface motivated us to explore the
449 possibility that ligands or receptors to these proteins may be regulated by hepRNA. Our hepRNA-
450 proximal datasets harbor many membrane glycoproteins which were no longer enriched in
451 response to RNase treatment (**Figure 6a**). In search for an appropriate hit to follow up on from our
452 datasets, we found glycoprotein poliovirus receptor (PVR, or CD155) had a decrease of ~28 folds
453 in its intensity in RNase dataset compared to that in TLR7 one, suggesting PVR is spatially close
454 to hepRNA. PVR was also a common hit between the hepRNA-proximal and our biotinylation-
455 OOPS datasets (**Supplementary Figure 6g**). Consistently, in a previous study, the mobility of
456 PVR from cell lysates through a sucrose density gradient dramatically changed upon RNase
457 treatment.[51] Additionally, PVR is documented in RBP2GO database, despite with a low RBP2GO
458 score.[52]

459 Given the RNA proximity or direct interaction, PVR was thus selected for further
460 investigation. PVR overexpresses in different cancers and can bind to T cell immunoreceptor with
461 immunoglobulin and ITIM domain (TIGIT),[53] CD96,[54] and Killer cell immunoglobulin-like receptor
462 2DL5 (KIR2DL5),[55] all of which are inhibitory receptors able to suppress the killing by T cells or NK
463 cells. PVR also interacts with an activating receptor, CD226.[56] The engagement of PVR with these
464 proteins forms immune checkpoints, and they are thus new promising targets for cancer
465 immunotherapy. We have envisioned the possible modes of action of HepRNA on PVR recognition
466 (**Figure 6b**): 1) hepRNA may sterically hinder the binding; 2) hepRNA may facilitate the binding of
467 specific proteins to PVR by functioning as a co-receptor, or by priming PVR for optimal recognition;
468 and finally, 3) hepRNA may be just a bystander, and has no effect on binding.

469 We used flow cytometry to quantify the RNase responsiveness of cell-surface binding by
470 the soluble, Fc-tagged version of ectodomains from the four known PVR-binders with and without
471 RNase pre-treatment (**Figure 6c**). TIGIT and CD226 exhibited little change in response to RNase
472 treatment, indicating that HepRNA does not modulate their recognition. CD96 did not detectably
473 bind to either cell surface in our experimental settings, suggesting additional factors may be
474 required. KIR2DL5 did show a dramatic reduction in fluorescent signals on the cell surface when
475 the live cells were pretreated with RNase. In line with the genetic screening data, the EXT2 KO
476 mutant cells also exhibited a strong reduction in KIR2DL5 binding, but not for CD226 or TIGIT
477 (**Figure 6d**). Neither the RNase treatment nor EXT2 KO negatively affected cell-surface PVR levels
478 (**Supplementary Figure 10a** and **b**), indicating that the decrease in KIR2DL5 binding was due to
479 hepRNA removal, rather than an effect on PVR expression. Exogenous RNA added to prior
480 hepRNA-depleted cells restored partly KIR2DL5 binding (**Supplementary Figure 10c**), suggesting
481 RNA-KIR2DL5 interaction played an important role in recruiting the latter on cell surface.

482 To demonstrate KIR2DL5 binds to hepRNA on living cells, we transiently overexpressed
483 the full-length protein fused with FLAG tag on wildtype and HS-deficient mutant (pgsD-677) CHO
484 cells (**Supplementary Figure 10e**. See **Supplementary Figure 4e** for hepRNA detection). We

485 expect the expression of human KIR2DL5 in such a non-human background should minimize
486 protein-protein interactions. Upon KIR2DL5 overexpression, the cells were exposed to UVC
487 irradiation to crosslink RNA with bound proteins in situ. After cell lysis, we performed
488 immunoprecipitation against FLAG tag to pull down KIR2DL5, and if any, the crosslinked RNA
489 bound to the protein. The co-precipitated KIR2DL5-bound RNA was enzymatically ligated with
490 biotin at 3'-end on-bead and finally subjected to Western blot for the detection of biotin. Strong
491 biotin signal as a smear was found above recombinant FLAG-tagged KIR2DL5 band, suggesting
492 RNA co-precipitation (**Supplementary Figure 10e**). Such smear disappeared in HS-deficient CHO
493 cells, which indicated an HS dependency of KIR2DL5-bound RNA. To confirm it was the
494 extracellular domain of KIR2DL5 that mediated RNA binding, we performed in-vitro assays in which
495 human Fc-fused KIR2DL5 extracellular domain or IgG isotype control was incubated with 3'-end
496 biotinylated total cellular RNA fragments, UVC crosslinked, and subjected to Western blot. Strong
497 signal was found only for KIR2DL5 crosslinked with biotinylated RNA, whereas that in IgG control
498 was hardly detectable (**Supplementary Figure 10f**). In addition to experimentally demonstrating
499 hepRNA can interact with KIR2DL5's extracellular domain, we also postulated a structural model
500 for the interaction using AlphaFold 3[57] (see **Suplementary Text 5** and **Supplementary Figure 11**
501 for a detailed explanation).
502 To confirm the binding of KIR2DL5 proteins to the cell surface is at least partly dependent
503 on PVR, we generated PVR KO cells using CRISRP-Cas9. Upon treatment with extracellular
504 RNase, the residual KIR2DL5 binding was significantly lower in PVR KO cells than in the wild type
505 (**Supplementary Figure 10d**). This indicated that KIR2DL5-PVR interaction accounted for a
506 portion of fluorescence signals detected on the cell surface. Taken together, these results
507 demonstrate that hepRNA alone can recruit KIR2DL5 on the cell surface, whereas PVR alone only
508 weakly binds to soluble KIR2DL5. The data supports the model that hepRNA functions as a co-
509 receptor for KIR2DL5 and thereby facilitates its engagement with PVR by increasing local KIR2DL5
510 concentration at cell surface.

511 **Discussion**

512 **A hepRNA-csRBP-HS ternary complex model**. Taking integrated omic-wide approaches and
513 using multiple orthogonal validatory methods, our study answered the pressing question in the
514 emerging field of RNA cell-surface localization: how the biomacromolecule is associated with the
515 plasma membrane. The study was commenced by an exploitation of nature's own RNA sensor,
516 TLR7, as a probe to detect RNA at the cell surface. The use thereof in a genome-wide CRISPR-
517 Cas9 knockout screening led to the discovery of HS polysaccharide as an essential molecular
518 factor for RNA's association with the cell surface. While direct interactions between carbohydrates
519 and nucleic acids have been reported,[58] our data supports the presence of a number of RBPs at
520 cell surface as crucial adapter molecules to bridge the two highly negatively charged biopolymers,
521 forming hepRNA-csRBP-HS ternary complexes (**Figure 7**).

522 In addition to the routine validatory experiments, such as immunofluorescence, flow
523 cytometry, and Western blot assays, we have established multiple orthogonal methods to validate
524 such ternary molecular configuration. We considered that the csRBPs had been an essential
525 element collectively for RNA's association with HS, but functionally redundant as individual
526 proteins, particularly as none of the csRBPs from proteomics study were hits in the genetic
527 screening. This was because the knockout of one RBP during the pooled screening may have
528 been compensated by others, such that little, if any, apparent decrease of global hepRNA could
529 be observed in flow cytometry using the TLR7 probe. Due to this collectively essential but
530 individually redundant nature of csRBPs for hepRNA, knocking out these RBPs altogether until a
531 global reduction in hepRNA is seen would fall experimentally impractical. We, therefore, broke
532 down the validation of the termolecular model into separate steps.

533 Firstly, we sought to validate, in a high-throughput manner, the involvement of hepRNA-
534 proximal csRBP candidates in RNA-RBP complex formation on plasma membrane. This was done
535 by a TLR7-independent cell-surface biotinylation method, combined with UVC crosslinking and the
536 isolation of RNA-bound proteins using OOPS. While 41 hepRNA-proximal proteins were validated
537 as RNA-bound csRBPs, the other hits uniquely identified in biotinylation-crosslinking-OOPS
538 experiment suggested there are more csRBPs than hepRNA-bound ones, because the
539 biotinylation chemistry and UVC crosslinking are not hepRNA-specific. These proteins may be
540 presented on cell surface via an HS-independent mechanism, such as protein-protein
541 interactions.[9, 59] A more comprehensive mechanistic understanding of csRBPs will likely unveil new
542 biology at cell surface, and warrants future studies.

543 We then introduced SCOOPS, a photocatalytic spatio-selective crosslinking technology to
544 selectively capture cell-surface RNA-protein complexes. SCOOPS is independent of any RNA-
545 binding probes such as TLR7. Thus, the essential role of HS as a scaffold for hepRNA-csRBP
546 complexes could be validated using an orthogonal approach. Furthermore, SCOOPS enabled
547 unveiling the identities of hepRNA for the first time. It revealed that decreased G and C content
548 and a more frequent UU motif occurrence are traits in the RNA primary structure related to HS

549 association. Apart from sequence motifs, more factors such as their secondary and higher-order
550 structures and the transcript-specific interacting proteome should be examined, and how they
551 function in an intertwined manner to regulate RNA cell-surface localization remains to be
552 elucidated.[60]

553 While our data supports a ternary complex model of hepRNA-csRBP-HS, the structural
554 basis for the interaction between RBPs and HS remain elusive. Thus, future work will be needed
555 on investigating the molecular details how RBPs bind to HS. HS is known to be a molecular mimicry
556 of nucleic acids and there have been reports showing several nucleic acid-binding proteins can
557 indeed interact with HS. Generally, the interaction between RBP and HS may be mediated by a
558 designated HS-binding domain present on the RBP, or a promiscuous RNA-binding domain,[61]
559 which typically harbors positively charged side chains for the electrostatic interaction with HS. SSB,
560 for example, contains a putative HS-binding motif (GRRFKG) at its C-terminus disordered region,
561 implying potential HS-binding activity.[62] Conversely, HNRNPA1 does not contain any known HS-
562 binding motif, despite its presence on the cell surface of wildtype, but not on HS-deficient mutant
563 cells. It suggests a yet-to-be-characterized mode of interaction between HS and the RNA-binding
564 domains or other portion of the protein.

565 **hepRNA as a co-binder for KIR2DL5**. We further demonstrate that once displayed on the
566 cell surface, the RNA can enrich glycoprotein immune receptors, likely to facilitate the interaction
567 with their cognate ligands. In the case of KIR2DL5, hepRNA functions as a co-binder, such that
568 local concentration of KIR2DL5 becomes much higher, thereby facilitating PVR engagement, and
569 compensating for the weak affinity of PVR for KIR2DL5.[55] The co-binder effect in modulating cell
570 surface receptor-ligand interactions is not uncommon. For example, the engagement of fibroblast
571 growth factor 2 (FGF-2) with FGFR requires the former to engage with HS, likely resulting in a
572 liquid-liquid phase separation.[63]

573 KIR2DL5, together with a few other KIR family members, has been previously shown to
574 bind to heparin in-vitro via the d0 domain, which is rich in positively charged amino acids.[64] Heparin
575 is often considered a molecular mimicry of nucleic acids because of its heavy negative charges
576 along the polysaccharide chain and has been used to purify select nucleic acid-binding proteins.
577 Although binding to heparin in-vitro, we demonstrated that KIR2DL5 could directly bind to RNA in-
578 vitro and, at least partly, on living cell surface, for which a structural basis was predicted using AF3.
579 Notably, such promiscuity in binding has been described for a cell-surface-localized protein,
580 receptor for advanced glycation endproducts (RAGE). It was shown previously to bind to HS,[65] and
581 more recently also to DNA.[66] Similar observation has also been made for certain antibodies
582 directed against DNA, which can cross-react with HS.[67, 68]

583 **Origin of hepRNA**. One key question remains as to the origin of the hepRNA and the
584 associated csRBPs. It is plausible that the RNPs are transported from cytosol to the plasma
585 membrane of the same cell ("inside-out"). However, the "inside-out" model requires the
586 translocation of cytosolic RNA and RBPs to the luminal side of specific membraneous organelles.

587 Interestingly, a recent study[69] showed a distinct set of RNAs from the cytosolically localized
588 transcripts can be found in the ER lumen, implying an unidentified pathway for RNA luminal
589 translocation. In our genetic screening study, TMED10 was identified as an essential factor for
590 hepRNA presentation. This multifunctional protein forms channels on the ERGIC membrane and
591 mediates the translocation of select cytosolic proteins without leader peptide to the luminal side for
592 secretion.[70] Of particular interest, TMED10 was reportedly not relevant to HS biosynthesis[71] and
593 has been documented to potentially bind to RNA in the RBP2GO database,[52] implicating their
594 possible roles in vesicular RNA transport.

595 An equally likely scenario would be that they are secreted from live cells or released from
596 dead/dying cells and directly deposited on the HS chains ("outside-on"). There are numerous
597 biological processes in which intracellular components including RNA (protein unbound or in
598 complex) are released to extracellular space. The release can be from dying/dead cells or
599 damaged tissues, but also living cells.[72] Extracellular RNAs or their protein complexes are common
600 damage-associated molecular patterns (DAMPs) which can activate innate and adaptive immunity
601 and are associated with many disorders,[73] such as sepsis,[74] cancer,[75] and autoimmune diseases.[76]
602 The observation that HS, but not other proteoglycans, can selectively capture extracellular RNA-
603 protein complexes is strong evidence of a specific and biologically relevant interaction, suggesting
604 that HS plays a unique role in regulating extracellular RNA dynamics and immune activation: On
605 the one hand, it may limit RNA-based DAMPs locally by capturing them to slow down their lateral
606 diffusion; on the other hand, it may also serve as a scaffold to present the RNPs to immune cells
607 for activation. Further exploration of HS-RBP-RNA interactions could uncover novel regulators of
608 immune signalling in tissue injury, offering new insights into inflammation-related diseases.

609 **Supplementary Text 1. Additional insights from genome-wide screening**

610 *RBPs as hits in TLR7^{low} subpopulation*. Among the hits from the genome-wide screening,
611 there are a few putative RBPs (see **Supplementary Table 2** for a complete list). ATP-dependent
612 RNA helicase **DHX29** (RBP2GO score 43.1) is a poorly understood classical RBP involved in
613 translation initiation.[77] Transmembrane emp24 domain-containing protein 10 (**TMED10**) is a
614 multifunctional protein in the cytoplasm. It functions as a cargo receptor involved in protein
615 vesicular trafficking and quality control in ER and Golgi. It can also oligomerize and act as a protein
616 channel to facilitate the post-translational entry of leaderless cytoplasmic cargo into the ER-Golgi
617 intermediate compartment (ERGIC) for extracellular secretion.[70] TMED10 is also a non-canonical
618 RNA-binding protein with an RBP2GO score of 9.7. The function of the hit appears interesting, as
619 the ribonucleoproteins do need to be transported from cytosolic side to luminal side of organelles
620 or vesicles to become topologically connected to plasma membrane. UDP-glucose 6-
621 dehydrogenase (**UGDH**) is an enzyme which creates sugar building blocks for proteoglycan
622 synthesis, and is a non-canonical RNA-binding protein (RBP2GO score 3.8). The RNA-binding
623 activity is expected to be unrelated to csRNA biology. Transmembrane 9 superfamily member 3
624 (**TM9SF3**) is a poorly understood protein regulating glycosyltransferase function in Golgi, and is a
625 non-canonical RNA-binding protein (RBP2GO score 5.7).

626 Interestingly, CD44, a glycoprotein known to be carrier protein of HS polysaccharide,[78] was
627 a hit (criteria, $p < 0.01$ AND at least two guide RNA enriched) in one biological replicate of the
628 genome-wide screening. In this replicate, it had an enrichment score of 4.40, with 2 out of 4 guide
629 RNAs significantly enriched. In the other replicate, CD44 had one guide RNA significantly enriched,
630 the gene had a moderate enrichment score of 1.27. Moreover, CD44 has been documented in the
631 R-DeeP database[51] to be potentially associated with RNA, either directly as an RBP or indirectly
632 as a complex with other RBPs.

633 *Hits from TLR7^{high} subpopulation*. After the two-round phenotypic selection using
634 fluorescence-assisted cell sorting (FACS), we have observed a clear bimodal distribution for TLR7-
635 bound subpopulation, with a small but distinct TLR7-negative subpopulation (**Supplementary**
636 **Figure 3**). However, there was no such distinct subpopulation with particularly high TLR7 signal
637 even after the second phenotypic enrichment. The significantly enriched hits consistent in two
638 biological replicates included N-acylneuraminate cytidylyltransferase (**CMAS**), Uroporphyrinogen
639 decarboxylase (**UROD**) and C-type lectin domain family 17, member A (**CLEC17A**). Despite CMAS
640 and UROD are documented in the RBP2GO database, their relevance to csRNA remain clear.

641 **Supplementary Text 2. Observation of singlet oxygen-mediated RNA-protein crosslinking.**

642 The small molecule photosensitizer, eosin Y (EY), is known to passively diffuse into living
643 cells, and has been used to generate singlet oxygen (SO) in cells with the irradiation with green
644 light (**Supplementary Figure 7a**).[48] When performing acid guanidinium thiocyanate-phenol-
645 chloroform (AGPC) extraction of total RNA (from the blue 1st aqueous phase in figure panel **a**,
646 without repeated OOPS), we observed substantial thickening of the interphase, and a dropped
647 yield of RNA isolated from the aqueous phase (**Supplementary Figure 7b**). We reasoned the
648 RNA must have been brought to the interphase (off-white phase in panel **a**), a similar effect which
649 occurs when the cells are exposed to UVC irradiation.[42] UVC irradiation creates covalent bonds
650 between the RNA and the bound proteins, making the complexes into an amphiphilic entity. This
651 physical nature of crosslinked RNA-protein complexes retains them in the AGPC interphase.
652 Based on our observation, EY-generated SO may have led to RNA-protein crosslinking, creating
653 the amphiphilic species.

654 We performed OOPS on EY-treated, green light irradiated cells (denoted EY in figures), in
655 parallel with UVC irradiated cells (UV) and a no crosslinking control (NC). The OOPS-cleaned
656 interphase can be treated with proteinase K to degrade bound proteins, releasing the crosslinked
657 RNA back to the aqueous phase. The released protein-bound RNA was isolated from aqueous
658 phase by another round of AGPC. Agarose gel electrophoresis (SybrGold stain) showed that there
659 was indeed RNA in the interphase of UV and EY samples, with a corresponding reduction of the
660 amount of RNA in the first aqueous phase (**Supplementary Figure 7c**). Most total cellular RNA
661 could be isolated from the first aqueous phase of NC samples.

662 If RNA were crosslinked to bound proteins by EY-generated SO, we should also expect
663 the presence of RNA-binding proteins (RBP) in the interphase (**Supplementary Figure 7d**).
664 Treating the OOPS-cleaned interphase with an RNase cocktail releases RNA-bound proteins to
665 organic phase (pink phase). We performed SDS-PAGE for the released RNA-bound proteins from
666 EY and UV samples, followed by a Coomassie gel stain (**Supplementary Figure 7e**). Stained
667 bands from EY samples were substantially stronger than those of UV, suggesting more
668 crosslinking events may have occurred for EY than UV. We also performed Western blot to detect
669 known RBPs, including heterogeneous nuclear ribonucleoprotein A1 (HNRNPA1), enolase 1
670 (ENO1), Sjögren syndrome antigen B (SSB) (**Supplementary Figure 7f**). HNRNPA1 gave strong
671 bands in EY samples, and moderate signals in UV. ENO1 and SSB did afford strong to moderate
672 signals in EY, but only weak to undetectable signals in UV. The results suggest EY can crosslink
673 RNA to proteins and the amphiphilic entities can be brought to the AGPC interphase for isolation.
674 Take together, these results serve as technical basis for the development of SCOOPS.

675 **Supplementary Text 3. Detailed discussion on the bioanalyzer electropherograms.**

676 The electropherograms of SCOOPS_KO samples indicate the SCOOPS interphase
677 contains more RNA components than exclusively hepRNA. Two prevalent peaks observed in the
678 bioanalyzer traces of SCOOPS_KO samples include one in small RNA region (fewer than 200
679 nucleotides), and the other aligned well with 28S ribosomal RNA (rRNA) (**Supplementary Figure**
680 **8a**). The peaks may be given rise to by RNA species exhibiting an intrinsic physical property to be
681 retained in the interphase regardless of whether SO-crosslinked. The SCOOPS interphase could
682 also contain originally intracellular RNA from, for instance, cell debris, which could be randomly
683 crosslinked by singlet oxygen generated at cell surface. In addition, the interphase may also
684 contain cell-surface localized RNA species whose presentation is not dependent on HS, which is
685 out of the scope of this study. Importantly, SCOOPS_KO samples can provide indications for
686 hepRNA-irrelevant RNA species intrinsically associated with the SCOOPS procedure. These RNA
687 species can serve as background in RNA sequencing data analysis (discussed in detail in
688 **Supplementary Text 4**).

689 The bioanalyzer traces of SCOOPS_WT samples revealed that SCOOPS interphase
690 contains RNA species with a wide length distribution. The content of the interphase became much
691 less when intact heparan sulfate (HS) chain is missing. This is in consistency with the biochemical
692 assays using in-gel fluorescence. While the bioanalyzer traces gave a more global view of the
693 length profile of SCOOPS-isolated RNA, the major difference in RNA in-gel fluorescence (**Figure**
694 **4h**) between SCOOPS_WT and _KO was observed only for small RNA region. This is likely due
695 to substantially higher numbers of exposed 3'-end in small RNAs for fluorophore ligation catalyzed
696 by T4 RNA ligase 1, while long transcripts have much fewer ligation sites available for the enzyme.

697 **Supplementary Text 4. Read count normalization on background RNA species.**

698 *Rationales*. As indicated by in-gel fluorescence (**Figure 4h**) and Bioanalyzer
699 electropherograms (**Supplementary Figure 8a**), SCOOPS_KO contains background RNA
700 species intrinsically associated with the SCOOPS procedure, whereas SCOOPS_WT contains
701 hepRNA as well as the background RNA. Thus, the background RNA could serve as an internal
702 standard and provide a means to normalize sequencing data. This is similar to the idea of
703 normalization on a minor set of genes/transcripts that remain unchanged between biological
704 conditions, such as housekeeping genes and spike-in controls.[79] Unlike the routinely used median
705 of ratios[80] or trimmed mean of M values[81] methods, which presumes only a small fraction of
706 genes/transcripts differs between conditions, normalization on minor unchanging components is
707 performed when a global shift is expected, which is the case for SCOOPS_WT and SCOOPS_KO.

708 If SCOOPS were performed with comparable cell numbers for WT and KO, the background
709 RNA in SCOOPS interphase, which is intrinsic to the technique, should be of similar quantity in the
710 two samples. Due to the presence of hepRNA, WT had substantially more RNA in the SCOOPS
711 interphase (Stage 1, **Supplementary Figure 9a**). However, equitizing masses of RNA input for
712 sequencing library preparation led to the reduced background RNA present in SCOOPS_WT than
713 in SCOOPS_KO (Stage 2). Such difference in background RNA species in WT and KO were kept
714 throughout the entire process, particularly as both samples underwent the same number of PCR
715 cycles (16 cycles), and were pooled in equimolar for sequencing (Stage 3). Upon normalization by
716 library size (read count per million mapped reads, or RPM), the background RNA species in
717 SCOOPS_WT should consistently have smaller RPMs than those in SCOOPS_KO (Stage 4). To
718 make it a fair comparison between SCOOPS_WT and _KO, the background RNA in the former
719 should be scaled to the same level as in the latter. The scaling factors for normalizing on
720 background RNA should then be applied to all mapped genes/transcripts in each sample (Stage
721 5).

722 *Scaling factors for normalization*. Although external scaling factors could have been
723 derived from comparing the amount of sequencing library input (150 ng) as fraction out of the total
724 amount of SCOOPS isolated RNA in WT and KO samples, we deliberately avoided normalization
725 by these external factors. Instead, we sought to derive the scaling factors internally, by comparing
726 background RNA transcripts between SCOOPS_WT and KO datasets. We define background
727 RNA as those with more than 2 times decrease in their mean RPMs in SCOOPS_WT compared
728 to the _KO sample ($\log_2$FoldChange[WT/KO] < -1, with a $\log_2$RPM $\geq$ 5 filter. See *Rationales*). Of
729 this subset of transcripts, Pearson correlation analysis revealed high similarity (R = 0.97) between
730 SCOOPS_WT and _KO (**Supplementary Figure 9b**), which fits with the hypothesis that the
731 background RNA is intrinsically associated with the SCOOPS procedure and remains unchanged
732 between biological conditions. Library-specific scaling factors were then derived from performing
733 median of ratios normalization for each SCOOPS sample (**Supplementary Figure 9c**).
734 Remarkably, such internal scaling factors correlate well with the external factors (**Supplementary**

735 **Figure 9d**). The RPMs for all transcripts in each individual library were then uniformly corrected
736 by the corresponding internal scaling factor.

Supplementary Text 5. AlphaFold3-predicted model for KIR2DL5-RNA interaction.

Structural basis of KIR2DL5-RNA interaction. To provide a plausible structural basis for KIR2DL5-RNA interaction, we employed Alpha Fold 3 (AF3) in generating a structural model for the bimolecular complex. We used stretches of eight oligonucleotide repeats and the extracellular region of KIR2DL5 as the input. The prediction demonstrated that KIR2DL5 ectodomain folded into two distinct domains (namely, d0 and d2) at overall high local confidence (plDDT score) (**Supplementary Figure 11a**), which is consistent with previous reports.[64] Because the RNA sequences KIR2DL5 binds to have been elusive, we used stretches of repetitive oligonucleotide repeats as RNA input for model prediction. AF3 predicted a moderate to high confidence interaction between KIR2DL5 and RNA (ipTM score) with a well-defined spatial arrangement (low PAE), suggesting that the binding interface is reliable. (**Supplementary Figure 11b** and **c**). Octaguanosine (G_8) gave 0.88 (between 0 and 1) as the highest ipTM score, suggesting guanosine residues may be important for KIR2DL5-RNA interaction. The predicted model complex shows RNA oligos likely make contacts with the N-terminus d0 domain, but not the C-terminus d2 domain. Consistently, d0 domain harbors a patch of positively charged residues such as lysine and arginine (**Supplementary Figure 11d**), which features are often associated with RNA binding, but further studies would be needed to determine their precise role in this interaction.

KIR2DL5-RNA interaction specificity. To investigate if KIR2DL5 binding to cell surface may be mediated by any of the cell-surface RBP hits, we also employed AF3 to predict if complexation is likely between KIR2DL5 extracellular domain and several cell-surface RBP hits (see Materials & Methods). Among all tested cell-surface RBPs, none gave an acceptable ipTM score (all below 0.6). Furthermore, we also observed a decrease of local plDDT scores for KIR2DL5 at the predicted interface compared to KIR2DL5 alone. Such worsened plDDT scores and low ipTM values indicate KIR2DL5 is unlikely to interact, at least, any of the screened cell-surface RBPs. Taken together, the computational data, along with the biochemical assay, supports KIR2DL5 directly binds to RNA.

RNA's role in regulating oligomerization. We further employed AF3 to gain insights into the molecular function of KIR2DL5-RNA association. KIR2DL5 belongs to the immunoglobulin superfamily (IGSF), and numerous members within this superfamily undergo homophilic interactions.[82, 83] We asked if homophilic interactions between KIR2DL5 monomers may occur and if RNA might regulate the oligomer/monomer states of KIR2DL5. When KIR2DL5 ectodomains alone as two separate protein entities were input in AF3, the predicted model had an ipTM score of 0.14 (**Supplementary Figure 11e**). We observed substantially reduced local plDDT values at the dimer interface. The high distance errors in the portion off the PAE matrix diagonal line (circled in purple, **Supplementary Figure 11f**) indicate low confidence in the predicted relative positions of the residues across the two separate KIR2DL5 monomers. These results suggest low likelihood for KIR2DL5 ectodomain *alone* to dimerize. However, a stretch of guanosine repeats with varying lengths (G_8, G_{12}, G_{16} and G_{20}) as another separate entity in addition to the two protein entities

775 greatly boosted the ipTM score to 0.77 (G_{16}) (**Supplementary Figure 11g**). The RNA oligos were
776 predicted to retain contacts with d0 domain of both KIR2DL5 monomers. The local plDDT scores
777 were larger than 90 at the protein dimer interphase, suggesting high-confidence prediction. In stark
778 contrast to protein dimer alone, the off-diagonal portion in the PAE matrix (circled in purple,
779 **Supplementary Figure 11h**) showed low errors, indicating AF3 had higher confidence in
780 predicting the relative positions of the residues across the two separate KIR2DL5 monomers within
781 this ternary complex. Based on the AF3 prediction, we postulate that RNA may facilitate the
782 dimerization of KIR2DL5.

783 **Materials and Methods**

784 **Cell culture**

785 Mel526, U2OS and HeLa cells were cultured in DMEM (Thermo Fisher Scientific, 31966047)
786 supplemented with 10% Fetal Bovine Serum (FBS, HyClone, SH30071.03HI) and 1x Penicillin-
787 Streptomycin-Glutamine (Thermo Fisher Scientific, 10378016, 100x) at 37℃ with 5 % CO_2. Wild-type
788 Chinese Hamster Ovarian (CHO) (CHO-K1, ATCC, CCL-61) and heparan sulfate deficient mutant cell
789 lines - CHO pgsE-606 and pgsD-677 (ATCC, CRl-2242 and -2244, respectively) were cultured in Ham's
790 F-12K medium (Thermo Fisher Scientific, 21127022) supplemented with 10% FBS and 1x Penicillin-
791 Streptomycin and Glutamine at 37℃ with 5 % CO_2. To minimize possible influence from components in
792 bovine serum, Mel526 cells and the mutant cells were also adapted into 1x Penicillin-Streptomycin-
793 Glutamine-supplemented DMEM containing 1% FBS and 9% chemically defined serum replacement
794 (Panexin CD, Pan-Biotech, P04-93100), following the protocol suggested by the supplier. Prior to
795 experiments, cell culture flasks or dishes were replenished with DMEM containing 10% Panexin-CD
796 without FBS.

797

798 **Immunocytochemistry**

799 Cells were fixed 24 hours post-seeding in µ-Slide 8 Well Glass Bottom ibidi chambers (ibidi, 80827,
800 ~30,000 cells/well) with 4% Paraformaldehyde (PFA) aqueous solution (Electron Microscopy Sciences,
801 15710-S) in PBS for 10 min at RT and washed with PBS twice. Blocking was performed with blocking
802 buffer containing 3% Bovine Serum Albumin in PBS (Sigma, A7638) for 1 hour at RT in RNase-free
803 conditions by adding SUPERaseIn™ RNase Inhibitor (Thermo Fisher Scientific, AM2696, 1:200). For
804 RNA competition, a synthetic RNA oligo (Dharmacon™) with a sequence 5'-
805 pGUCUUCAAAACUAGGUCGUUUUAGA-3'/biotin/ was added at 10 µM during TLR7-Fc incubation.
806 For RNase-treated samples, RNase Cocktail Enzyme Mix (Thermo Fisher Scientific, AM2296, 1:100)
807 was added to the blocking buffer. Post-blocking, recombinant proteins or primary antibodies were added
808 to cells prepared in the same respective blocking buffer and incubated at 21℃ for 1 hour. Cells were
809 washed 3 times with PBS (2 minutes per wash) followed by the addition of secondary antibodies
810 prepared in the blocking buffer along with DAPI (Roche, 10236276001, 1 µg/mL final concentration)
811 and incubated at 21℃ for 1h. The cells were washed with PBS and imaged with confocal microscopy
812 (Olympus A1R SiM, oil immersion, filter settings for Red (Alexa-546 and others: excitation 551nm,
813 emission 565 nm); Blue (DAPI: excitation 358 nm, emission 463 nm); Green (Alexa-488 and others:
814 excitation 490nm, emission 544nm) and Far-red (Alexa-647 and others: excitation 650 nm, emission
815 671 nm).

816

817 **Electron microscopy**

818 U-2 OS cells were cultured in a 6 cm petri dish at 90-100% confluency. To avoid fixation-induced
819 artefacts, live cells were first stained for csRNA by incubating on ice for 30minutes and then applying
820 TLR7-Fc (3% Bovine Serum Albumin in PBS, final conc 2.5 µg/mL) for 1 hour on ice followed by
821 thorough washing and incubation with PAG-10 (Protein-A-gold 10nm, PAG10 nm/S, OD50 ,Cell
822 Microscopy Core, UMC Utrecht) solution in PBS (1:50) on ice for 30minutes. Immunogold-labelled cell

823 monolayers were fixed in 1.5% glutaraldehyde in 0.1 M sodium cacodylate buffer for 2 hours before
824 being successively incubated in 1% osmium tetroxide in 0.1 M cacodylate buffer for 1 hour and in 1%
825 uranyl acetate in water for 1 hour. The cells were then dehydrated through a series of incubations in
826 ethanol (70-100%) for 90 minutes and embedded in Epon. The flat embedded cells were sectioned with
827 an ultramicrotome (UC6, Leica, Vienna) using a 35 degrees diamond knife (Diatome, Biel, Switzerland)
828 at a nominal section thickness of 70 nm. The sections were transferred to a formvar and carbon coated
829 200 mesh copper grid and stained for 20 minutes with 7% uranyl acetate in water and for 10 minutes
830 with lead citrate. EM images were recorded using a Tecnai 12 electron microscope (Thermo Fisher
831 Scientific) equipped with an EAGLE 4k×4k digital camera. For navigation on EM images, montages of
832 images at 11,000× were generated using stitching software.[84] The stitched images were imaged and
833 annotated using Aperio ImageScope (Leica Biosystems).
834
835 **Flow cytometry analysis of csRNA on live cells**
836 Cells cultured to 80-90% confluency were washed once with 1x DPBS after removing culture media,
837 and then lifted with 10 mM EDTA (diluted from 0.5M sterile filtered stock solution) in 1x DBPS at 37 ℃
838 for 5 minutes. Lifted cells were dispensed into a 96-well FACS plate with 2-2.5 x10^5 cells per well. For
839 extracellular RNase treatment, PureLink RNase A (Thermo Fisher Scientific, 12091021) were added to
840 the lifting buffer at a final concentration of 40 µg/mL. After spinning down at 300 g for 5 minutes at 4℃
841 and supernatant removal, cells in each well were resuspended in 30 µL of TLR7-Fc or other Fc-tagged
842 proteins (final concentration 5 µg/mL) in 0.5% BSA in 1x DPBS containing 200U/mL RNasein (Promega,
843 N2511), and incubated on ice for 45 minutes. For secondary only control, cells were incubated with 30
844 µL 0.5% BSA in 1x DPBS. After incubation, 200 µL cold 1x DBPS were added to each well and the
845 plate was spun down at 300 g for 5 minutes at 4℃. After supernatant removal, cells were resuspended
846 in 30 µL of goat-anti-human IgG AlexaFluor647 conjugate (Thermo Fisher Scientific, A-21445) in 0.5%
847 BSA in 1x DPBS, and incubated in dark for 30 minutes on ice. Upon completion, 200 µL cold 1x DBPS
848 were added to each well and the plate was spun down at 300 g for 5 minutes at 4℃. Finally, cells were
849 resuspended in 200 µL cold 0.5% BSA in 1x DPBS containing 0.1 µg/mL DAPI. The cells were then
850 analyzed on a flow cytometer using HTS settings. Typically, 150 µL cell suspension per well was infused
851 in the system for analysis, with a flow rate of 1 mL/minute.
852
853 **TLR7-Fc binding rescue by exogenous RNA**
854 Cells were processed as was described above, except for that after detachment with EDTA buffer
855 containing Purelink RNase A, cells were spun down, resuspended in 1x DPBS containing 200 U/mL
856 RNasein, with varying concentration of total cellular RNA (2 – 0.2 mg/mL), incubated at 4 ℃ for 30
857 minutes, and washed with cold 1x DPBS, before incubating with TLR7-Fc. The rest of the steps followed
858 exactly the above flow cytometry experiment.
859
860 **TLR7-RNA crosslinking, precipitation, and on-bead RNA labelling**
861 The procedure after TLR7-Fc binding follows that of eCLIP with modifications.[29] Cells in a 15 cm dish
862 at 90% confluency were washed once with 1x DPBS following culture media removal, and was then

863 incubated with 3 mL per dish 10 mM EDTA in 1x DPBS at 37℃ for 5 minutes. Another 7 mL 1x DPBS
864 was added to the dish to wash cells off the dish. The total of 10 mL cell suspension was transferred to
865 a 15 mL Falcon tube. The dish was washed further with 5 mL 1x DPBS, which was then combined in
866 the same tube. The cells were spun down at 300 g for 5 minutes at 4℃, and the supernatant was
867 removed. The cell pellets were taken up in 250 µL TLR7-Fc (final concentration 5 µg/mL) in 0.5% BSA
868 in 1x DPBS containing 200U/mL RNasein, and the suspension was put on ice for 30 minutes, with
869 gentle tapping every 10 minutes. The tube was topped up to 15 mL with cold 1x DPBS, and spun down
870 at 300 g for 5 minutes at 4℃. The pellets were washed twice with 10 mL cold 1x DPBS. The cells in the
871 second wash were directly transferred to a 10 cm dish for UVC crosslinking on ice without the lid using
872 Stratalinker (400 mJ setting). After UVC exposure, cells were transferred back to the tube and spun
873 down. For TLR7 non-crosslinked control, cells after lifting were transferred into a 15 mL Falcon tube
874 and spun down, resuspended in 10 mL 1x DPBS, and transferred to a 10 cm dish for crosslinking. After
875 crosslinking, cells were transferred back to the 15 mL tube and spun down, and incubated with the
876 TLR7-Fc solution, which was followed by two washes.
877 The pellets from above were lysed in 400 µL lysis buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl,
878 1% v/v Igepal CA-630, 0.1% v/v SDS, 0.5% w/v sodium deoxycholate) containing 1x cOmplete protease
879 inhibitor (Roche, 11836153001) cocktail and 1mM PMSF. The solution was sonicated at 20% power for
880 5 minutes (10s on, 10s off) on ice. To the sonicated lysate was added 2 µL TurboDNase (Thermo Fisher
881 Scientific, AM2238), and 4 µL diluted RNase I (Thermo Fisher Scientific, AM2294, prediluted 1:20 in 1x
882 DPBS). The lysate was then incubated at 37℃ for 5 minutes, followed by the addition of 4 µL
883 SUPERaseIn. To prepare magnetic beads for TLR-Fc precipitation, 10 µL per dish protein A Dynabeads
884 (Thermo Fisher Scientific, 10001D) was washed twice with lysis buffer before added to the above RNA-
885 fragmented lysate. The bead suspension was incubated on an end-to-end rotator at 4 ℃ for 2 hours.
886 The beads were then wash twice with 200 µL CLIP high salt buffer (50 mM Tris-HCl pH 7.4, 1 M NaCl,
887 1% Igepal CA-630, 1 mM EDTA, 0.1% SDS, 0.5% sodium deoxycholate), three times with 200 µL CLIP
888 low salt buffer (20 mM Tris-HCl pH 7.4, 10 mM $MgCl_2$, 0.2% Tween-20, 5 mM NaCl). The precipitated
889 RNA-TLR7-Fc was 3'-end repaired on-bead by incubating with FastAP (Thermo Fisher Scientific,
890 EF0654, 1 U), T4 PNK (NEB, M0201, 20 U), TurboDNase (0.4 U) and SUPERaseIn (5 U) in 10 µL 1x
891 PNK buffer without DTT at 37 ℃ for 30 minutes, followed by the same wash procedure. The end-
892 repaired RNA was subjected to on-bead 3'-end biotinylation by incubating with 20 nM pCp-biotin, T4
893 RNA ligase high-concentration (NEB, M0437M, 30 U) and SUPERaseIn (5 U) in 1x RNA ligase buffer
894 without DTT, containing 1 mM ATP, 3.6% DMSO, 0.025% Tween-20 and 18% w/v PEG-8000 16 ℃
895 overnight, which was followed the same wash procedure. After the final wash, beads were directly taken
896 up in 1x LDS sample loading buffer containing 1x sample reducing solution, denatured at 75 ℃ for 10
897 minutes. The samples were resolved on an NuPAGE™ 4 -12% bis-Tris mini protein gel, wet transferred
898 onto a nitrocellulose (NC) membrane at 30 V for 4 hours at 4 ℃, and subjected to chemiluminescent
899 biotinylated nucleic acid detection. The same NC membrane was finally subjected to Western blot for
900 TLR7 detection using rabbit anti-human TLR7 (1:1000) in combination with horse radish peroxidase
901 (HRP)-conjugated goat-anti-ribbit secondary antibody (1:5000).
902

903 **Proximity biotinylation of csRNA-proximal proteome**

904 The experiment was performed using three independent cell cultures for each treatment. Cells in a 15
905 cm dish at 90% confluency were washed twice with 1x DPBS after removing culture media, and was
906 lifted with 3 mL per dish 10 mM EDTA in 1x DPBS. For csRNA-digested sample, RNase A was added
907 to the lifting buffer at a final concentration of 40 μg/mL. 1x DPBS was used to bring cell suspension to
908 10 mL, which was transferred to a 15 mL Falcon tube. The dish was washed with 5 mL 1x DPBS, which
909 was combined in the same tube. The cells were spun down at 300 g for 5 minutes at 4℃, and pellets
910 were incubated for 30 minutes at 4 ℃ with per dish 250 μL solution of precomplexed TLR7-Fc (5 μg/mL
911 final concentration) and protein-A-HRP conjugate (Thermo Fisher Scientific, 101023, 7 μg/mL final
912 concentration) in 0.5% BSA in 1x DPBS containing 200U/mL RNasein. For an IgG isotype control,
913 human IgG (Thermo Fisher Scientific, 02-7102) was premixed with protein-A-HRP at the same
914 concentration. Cells were washed three times with cold 1x DPBS, and spun down at 200 g for 2 minutes
915 at 4 ℃ each time. After the final wash, the pellets were resuspended in 1 mL per sample 500 μM biotin-
916 phenol in 1x DPBS, and let sit for 1 minute at room temperature. 1 mL 2 mM hydrogen peroxide in 1x
917 DPBS was added to each sample and mixed well. This mixture was incubated for 2 minutes at room
918 temperature and 10 mL cold quenching buffer (5 mM Trolox + 10 mM sodium ascorbate + 10 mM
919 sodium azide in 1x DPBS) was directly added to the sample. The cells were washed two more times
920 with quenching buffer, and were lysed in 300 μL per sample RIPA buffer (50 Mm Tris pH 7.5, 150 mM
921 NaCl, 0.5% sodium deoxycholate, 0.1% SDS, 1% NP40) containing 1x cOmplete™ protease inhibitor
922 cocktail and 1mM PMSF. The lysate was sonicated to clarity, and was spun down at 21,000 g for 10
923 minutes at 4℃. An aliquot of the supernatant was taken for WB to detect biotin signal, and the rest was
924 subjected to enrichment of biotinylated proteins.

925 The affinity pulldown follows the previous report by Santos-Barriopedro et al. The pellets were
926 taken up in RIPA buffer. The supernatant from the sonicated lysates was subjected to bead enrichment
927 with 12.5 μL per sample streptavidin sepharose high-performance beads (Cytiva, 15511301) on an end-
928 to-end rotator for 2 h at room temperature. The beads were washed 5 times with 500 μL RIPA buffer
929 and 4 times with 500 μL PBS, prior to LC-MS/MS sample preparation.

930

931 **Live-cell surface biotinylation, UVC crosslinking followed by OOPS**

932 Cells in a 15 cm dish were cultured to 85 – 90% confluency. Prior to detachment, culture medium was
933 removed and the cells were washed once with 1x DPBS. The cells were lifted using 10 mM EDTA in 1x
934 DPBS and transferred into 15 ml Falcon tubes. The pellets were then taken up in 1 ml per tube
935 biotinylation solution containing 2.5 mM sulfo-NHS-SS-biotin (Thermo Fisher Scientific, 21331). For
936 non-biotinylated control, cells were treated with just 1x DPBS and were processed the same way. The
937 cells were incubated at room temperature for 15 min with gentle shaking every 5 min. Biotinylation was
938 quenched using another 8 ml per tube 20mM Tris-HCl in 1x DPBS pH 7.4 at room temperature for 10
939 min. After spin-down, the cells were taken up in 7 ml 1x DPBS and transferred into a 10 cm dish for
940 UVC crosslinking following procedures described above. The crosslinked cells were collected by
941 centrifugation and lysed in TRIzol for performing OOPS. The OOPS procedure is described in detail in

942 Section *UVC and eosin Y-mediated crosslinking and orthogonal organic phase separation (OOPS)*
943 below.
944 The RNA-bound RBP from organic phase was re-dissolved in cell lysis buffer (50mM pH7.5
945 Tris-HCl, 100mM NaCl , 1% (v/v) NP-40, 0.1% (v/v) SDS, 0.5% (w/v) Sodium deoxycholate), and
946 subjected to affinity purification using agarose beads functionalized with immobilized neutravidin
947 (Thermo Fisher Scientific, 29204). The bead binding was performed at 4ºC for 2 hours on an end-to-
948 end rotator. The supernatant post-enrichment was discarded. The bead was washed 5 times with lysis
949 buffer, and a final wash in 1x DPBS, prior to LC-MS/MS sample preparation.
950
951 **Proteomics sample preparation and LC-MS/MS analysis**
952 Beads from biotin affinity pulldown were incubated with 50 µL elution buffer (2 M Urea, 10 mM DTT,
953 100 mM Tris pH 8) on a thermoshaker at 1250 rpm for 20 minutes at room temperature. Iodoacetamide
954 (50 mM final concentration) was then added, and the samples were incubated in a thermoshaker at
955 1250 rpm in dark for 10 minutes at room temperature. 2.5 µL of trypsin (0.1 mg/mL stock solution) was
956 then added, followed by incubation in a thermoshaker at 1250 rpm for 2 hours at room temperature.
957 Samples were spun down (1500 x *g*, 2 min, RT) and supernatant was collected. Another 50 µL elution
958 buffer was then added to the beads, which were incubated on a thermoshaker at 1250 rpm for 5 minutes
959 at room temperature. Beads were then spun down (1500 x *g*, 2 min, RT), and the supernatant collected
960 and combined with the first elution. 1 µL trypsin (0.1 mg/mL) was added and the samples were incubated
961 overnight at room temperature. Tryptic peptides were acidified by adding 10 µL 10% (v/v) TFA, and
962 purified using C18 Stagetips[85] (3M Empore). Briefly, StageTips were prepared and washed with
963 methanol followed by a wash with buffer B (80% acetonitrile and 0.1% formic acid) and two washes
964 with buffer A (0.1% formic acid). All washes were done at 1500 × *g* for 4 min at RT. Samples were
965 loaded onto the StageTips and centrifuged at 600 × *g* for 10 min at RT and washed once with buffer A.
966 StageTips were stored at 4ºC
967 Peptides were eluted from StageTips with 30 ml of buffer B, reduced to a volume of 5 ml by
968 Speedvac centrifugation, after which 7 ml of buffer A (0.1% formic acid) was added. For the TLF7
969 proximity labeling experiments, 6 µl of each sample was loaded on an Easy LC1000 (Thermo Scientific)
970 connected to an Orbitrap Exploris 480 (Thermo Scientific). The machine was operated in data-
971 dependent acquisition (DDA) mode, with acquisition time set at 60 minutes. An acetonitrile gradient of
972 12-30% in 43 minutes was used followed by an increase of the acetonitrile concentration to 60% in 10
973 minutes and to 95% in 1 minute. The full scan of the peptides was set to a resolution of 120,000 in a
974 scan range of 350–1300m/z. For biotinylation-crosslinking-OOPS experiments, 3 µl of each sample was
975 loaded on a Vanquish Neo UHPLC connected to a Orbitrap Astral. The machine was operated in data-
976 independent acquisition (DIA) mode, with acquisition time set at 24 minutes. An acetonitrile gradient of
977 8-35% in 15 minutes was used followed by an increase of the acetonitrile concentration to 45% in 3
978 minutes and to 99% in 30 seconds. The full scan of the peptides was set to a resolution of 240,000 in
979 a scan range of 150–2000 m/z.
980
981 **Proteomics data analysis**

982 Raw mass spectrometry data were analysed using Proteome Discoverer version 3.1 (Thermo Scientific)
983 and searched against the human UniProt/SwissProt database.
984 For DDA, we used the built-in processing workflow
985 "PWF_OT_Precursor_Quan_and_LFQ_CID_SequestHT_Percolator" and the built-in consensus
986 workflow "CWF_Comprehensive_Enhanced Annotation_LFQ_and_Precursor_Quan", with default
987 settings. For the Sequest HT search, database parameters were enzymatic digestion with trypsin
988 allowing two missed cleavages, a minimum peptide length of 6 amino acids and a maximum peptide
989 length of 144 amino acids. We used a precursor mass tolerance of 10 ppm and a fragment mass
990 tolerance of 0.6 Da. Cysteine carbamidomethylation was included as a static modification (57.021 Da),
991 while methionine oxidation (15.995 Da) and protein N-terminal acetylation (42.011 Da) were included
992 as dynamic modifications. FDR filtering was performed via percolator with a strict target FDR of 0.01
993 and a relaxed FDR of 0.05. Strict parsimony was applied for protein grouping, and unique plus razor
994 peptides were used for quantification. Peptide quantification normalization was applied based on total
995 peptide amount. Imputation Mode was set to "Low Abundance Resampling". "Proteins" tab was
996 exported for downstream processing and statistical analysis with R.
997 For DIA, we used the built-in processing workflow
998 "PWF_Astral_CHIMERYSonArdia_DIA_MBR.BestPSM" and the built-in consensus workflow
999 "CWF_Comprehensive_Enhanced Annotation_LFQ_and_Precursor_Quan", with default settings. For
1000 the CHIMERYS search, database parameters were enzymatic digestion with trypsin allowing two
1001 missed cleavages, a minimum peptide length of 7 amino acids and a maximum peptide length of 30
1002 amino acids. We used a fragment mass tolerance of 20 ppm. Cysteine carbamidomethylation was
1003 included as a static modification, while methionine oxidation was included as dynamic modification.
1004 FDR filtering was performed via percolator with a strict target FDR of 0.01 and a relaxed FDR of 0.05.
1005 Strict parsimony was applied for protein grouping, and unique plus razor peptides were used for
1006 quantification. Peptide quantification normalization was applied based on total peptide amount.
1007 Imputation Mode was set to "Low Abundance Resampling". "Proteins" tab was exported for downstream
1008 processing and statistical analysis with R.
1009 For downstream processing, proteins were first filtered for having at least 2 unique peptides.
1010 Proteins containing "KRT" in their names as well as proteins with less than 3 values in at least 1
1011 condition were filtered out. Statistical analysis was done using the DEP package, which uses a
1012 Student's t-test. Criteria for enrichment are listed in the figure legends. Volcano plots were generated
1013 by plotting the –log10 of the p-value against the log2 fold change. Gene ontology term searches for
1014 identified hits were performed using RBP2GO.
1015 The proteomics datasets have been deposited in the ProteomeXchange Consortium via the
1016 PRIDE[86] partner repository with the dataset identifiers PXD061693. (Reviewer log-in username:
1017 reviewer_pxd061693@ebi.ac.uk. Password: ncGUlhn5XtSB)
1018
1019 **Genetic screening and data analysis**
1020 For knockout screening, the human CRISPR Brunello genome-wide knockout library was a gift from
1021 David Root and John Doench (Addgene 73178). For virus production, HEK 293T cells were transfected

1022 with packaging plasmids pRSVrev, pHCMV-G VSV-G and pMDLg/pRRE together with the Brunello
1023 plasmid using polyethyleneimine (Polyscience Inc.). Virus was harvested, filtered and 150 million
1024 Mel526 melanoma cells (a kind gift from Ramon Arens) were transduced in the presence of 8 µg/ml
1025 polybrene (Millipore) at an MOI of 0.3. Transduced cells were selected using puromycin (1 µg/ml) and
1026 after seven days, two batches of cells were stained for surface RNA using TLR7-Fc and the lowest 5%
1027 of expressing cells were sorted. Cells were grown out and sorted again using the same gating strategy
1028 as for the first sort. After this sort, genomic DNA was isolated for both the unsorted and sorted
1029 populations and gDNAs were amplified using the established protocol.[87] gRNAs were sequenced using
1030 the Illumina NovaSeq6000 and inserts were mapped to the reference. Analysis of gRNA enrichment
1031 was done using PinAPL-Py.[88]
1032
1033 **Genetic screening hit validation**
1034 Hits from the knockout screen were validated using the top scoring gRNAs of each gene, which were
1035 cloned into the lentiCRISPR v2 vector (a gift from Feng Zhang, Addgene plasmid no. 52961). Mel526
1036 cells were stably transduced with these gRNAs, selected using puromycin, and pooled knockout cells
1037 were used for the analysis. Guide sequences were as follows: EXT2, 5′-
1038 CACCGTGGTTAAGCACATCGATGGA-3′; EXTL3, 5′-CACCGAAATGAACCTCGGTAACACG-3′;
1039 HS6ST1, 5'-AAACCCGCGAGACTTGGCTCTTCTC-3', C3orf58, 5'-
1040 CACCGGCAGACGCACGTCGCCGTTG-3'. Mutant cells were subjected to flow cytometry analysis to
1041 check for TLR7-Fc binding, following the procedure described above.
1042
1043 **Preparation of PVR KO mutant cells**
1044 HEK293T cells (ATCC CRL-3216) were seeded into a 10 cm dish per plasmid at a confluency of 60-
1045 70% for lentiviral packaging. The utilized guide RNA was 5'-ATTGGTGCCCTCAAGCCAGG-3' (PVR,
1046 NM_006505.4, NC_000019.10. Sense, exon. Position: 44650019). For a 10 cm dish, 5.63 µg pCMV-
1047 VSVG, 8.55 µg pMDLg-RRE (gag/pol), 4.5 µg pRSV-REV and 11 µg pRRL were combined with serum-
1048 free DMEM to a total volume of 500 µl. In a second tube, 90 µg of PEI for a 10 cm dish was mixed with
1049 serum-free medium to a total volume of 500 µl. This mixture (tube two) was then slowly added to tube
1050 one and thoroughly mixed. After 15-30 minutes, the resulting mixture was added dropwise to the cells
1051 after adding 4 ml of serum-free medium. 4 hours post-transfection, 4 ml of DMEM with 20% FCS was
1052 supplemented to the plate. Subsequently, the plates were incubated overnight at 37°C with 5% CO_2.
1053 48 hours post-transfection, the supernatant containing virus was harvested by collecting the medium
1054 into a sterile 15 ml tube. The collected medium centrifugated at 3000 rpm at 4°C for 10 minutes and
1055 was subsequently filtered through a 0.45 µm low protein binding membrane (Millipore). The viruses
1056 were utilized immediately or stored at -80°C. The transduced cells were cultured using high glucose
1057 DMEM with 10% FBS and 1% Pen-Strep. Transduction was executed by adding 30% of lentivirus plus
1058 5 µg/mL Polybrene. Transduced cells were selected using puromycin (3 µg/ml) 48 hrs after
1059 transduction.
1060
1061 **UVC and eosin Y-mediated crosslinking and orthogonal organic phase separation (OOPS)**

1062 For suspension cells such as K562 (used in this study), start with ~20 million. Culture media was
1063 removed and the cells were wash twice with 10 ml DPBS. Cell pellets were resuspended in 10 ml DPBS,
1064 and transferred into a 10 cm dish. For UVC crosslinking, the cells (without lid) were then UV-crosslinked
1065 with a Stratalinker 1800 chamber at 254 nm for 5 min on ice. The crosslinked pellets were collected and
1066 directly lysed in 1 ml TRIzol per million cells. For EY-mediated crosslinking, cells after wash were
1067 resuspended in DPBS containing 50 μM eosin Y (CAS #17372-87-1). The cells were placed in dark at
1068 room temperature for 2 min and put on ice with lid open. The cells were exposed to green light (100 W
1069 chip-on-board) for 10 min. The crosslinked pellets were collected and directly lysed in 1 ml TRIzol per
1070 million cells.

1071 For adherent cells, start at least with a 10 cm dish. Culture media was removed and the cells
1072 were wash twice with 10 ml HBSS. In case of UVC crosslinking, upon removal of the wash solution, the
1073 dish without the lid was placed in the chamber on ice, and exposed to UV for 5 min. The crosslinked
1074 cells were lysed in 1 ml TRIzol per 10 cm dish, and scraped to one side of the dish and transferred to a
1075 2 ml Eppendorf tube. In case of EY-mediated crosslinking, 7 ml HBSS containing 50 μM eosin Y was
1076 added to the dish. The dish without lid was exposed to green light for 10min and washed once more
1077 with HBSS, followed by cell lysis in 1 ml TRIzol per 10 cm dish, scraping and the transfer as mentioned
1078 ealier.

1079 OOPS was performed following reported literature. Unless otherwise noted, the phase
1080 separation procedure and the amounts of enzymes and duration of incubation for releasing RNA-bound
1081 proteins and protein-bound RNA remain the same as reported. In brief, 0.2 ml chloroform per 1 ml
1082 TRIzol was added to the tube, followed by vortexing and spin-down at 4°C. The aqueous and organic
1083 phases were carefully removed with a blunt needle without touching the interphase. Such procedure
1084 was repeated 3 times for UVC crosslinking and 4 times for EY crosslinking. The cleaned interphase
1085 was then precipitated with methanol. The resulted pellets were wash once again with methanol and air-
1086 dried for 5 min.

1087 To isolate protein-bound RNA, the interphase pellet was directly treated with proteinase K
1088 solution. The protein-degraded RNA samples were then purified with TRIzol in combination of Zymo
1089 RNA Clean & Concentrator columns (Zymo Research C1008) following the product manual. The
1090 cleaned RNA samples were then subjected to 1% agarose gel electrophoresis and stained with
1091 SybrGold. The stained gel was imaged with a ChemiDoc system using SybrGold settings.

1092 To isolate RNA-bound proteins, the interphase pellet was redissolved in 1% SDS containing
1093 ammonium bicarbonate. The RNA was fragmented by sonication and Mg(II) ions at 94°C. To the
1094 resulted solution, RNase Cocktail was added. The digestion was incubated at 37°C for 20 hours,
1095 followed by another round of TRIzol phase separation. The RNA-bound proteins were collected from
1096 the organic phase and precipitated with methanol, redissolved in 1% SDS. An equal volume of each
1097 sample solution was taken for SDS-PAGE and Western blot. To check for total isolated proteins, the
1098 gel was Coomassie-stained and scanned under a ChemiDoc system. For Western blot, the PVDF
1099 membrane post-transfer was blocked with 5% BSA in TBST (blocking buffer), incubated with the
1100 corresponding primary antibody (1:1000 dilution for all) in blocking buffer, washed 3 times with TBST,
1101 and incubated with the secondary antibody (1:5000 – 1:10000) in blocking buffer. The resulted

1102 membrane was washed 3 times, treated with West Pico PLUS Chemiluminescent Substrate (Thermo
1103 Fisher Scientific, 34580), and imaged under a ChemiDoc system.
1104
1105 **Conjugation of eosin isothiocyanate to streptavidin**
1106 Eosin 5-isothiocyanate (Abcam, ab270343) was dissolved in DMSO to make a 100mM stock solution.
1107 The conjugation reaction was performed for 500 µg streptavidin in 0.5 ml 1X DPBS at pH 8, with eosin
1108 isothiocyanate at a final concentration of 2mM. The reaction was allowed to proceed for 2 hours at
1109 room temperature, with shaking and in dark. The resulting mixture was purified twice with MidiTrap G-
1110 25 (GE Life Sciences, 28922530) columns, and buffer exchanged into 1x DPBS pH 7.4. The purified
1111 conjugate was stored at 4 ºC in dark.
1112
1113 **Enzyme-linked immunosorbent assay (ELISA) of EY-conjugated streptavidin**
1114 The 96-well MaxiSorp microtiter plate was coated with either streptavidin-EY or unmodified streptavidin
1115 at 10 µg/ml in 100mM $NaHCO_3$ overnight in dark at room temperature. The wells were washed twice
1116 with PBS and then blocked with 2% BSA in PBST (blocking buffer) 1 hour in dark at room temperature.
1117 Serial dilutions (3x, starting concentration 10 µg/ml) of a biotinylated mouse monoclonal antibody
1118 (BioLegend, 317320) in blocking buffer were added to each row, and was incubated for 1 hour in dark
1119 at room temperature, followed by washing 3 times with PBST. To each well, HRP conjugated alpaca-
1120 anti-mouse (Jackson ImmunoResearch, 615-035-214) in blocking buffer was added and allowed to
1121 incubate for 45 minutes in dark at room temperature, followed by washing 3 times with PBST. Finally,
1122 1-step Ultra TMB-ELISA (Thermo Fisher Scientific, YJ4085531) and incubated for 2 mins, and was
1123 stopped by 2M sulfuric acid. The optical density at 450 nm (OD450) was measured by a microtiter plate
1124 reader.
1125
1126 **Spatio-selective crosslinking followed by OOPS (SCOOPS)**
1127 Cells at 85-90% of confluency in a 10 cm dish was used for each sample. For csRNA degradation, cells
1128 were treated with 13 µl of RNase A (Thermo Fisher Scientific, 12091021) and 3.2 µl of RNase I (Jena
1129 Biosciences) in 6.5 mL culture media under standard culturing condition (37˚C, 5% CO_2 and humidified
1130 atmosphere) for 30 mins. All cells were washed with HBSS, and were incubated with 6 ml culture media
1131 containing 5 µg/ml biotinylated WGA (Vector Laboratories, B-1025) at 4ºC for 1 hour. The corresponding
1132 negative control was with merely 6 ml culture media without biotinylated WGA. After aspiration, cells
1133 were washed twice with HBSS and then incubated with 6ml of culture media containing 2.5 µg/ml
1134 streptavidin-EY at 4ºC for 30 mins in dark, after which the cells were washed twice with HBSS. The
1135 corresponding negative controls was with unmodified streptavidin in the culture media. 6mL HBSS was
1136 added to each plate prior to green light irradiation at 4ºC for 10 mins. Supernatant was immediately
1137 removed after irradiation and the cells were lysed in 1.4ml TRIzol per dish, scraped and transferred into
1138 a 2 ml Eppendorf tube for OOPS. The OOPS procedure is the same as described earlier. Protein-bound
1139 RNA was released from interphase, purified, and subjected 3'-end Cy5 ligation, which followed the
1140 same procedure for 3'-end biotin ligation described above, except for the purification was performed
1141 using Zymo RCC-5 columns. 2% (v/v) of the first aqueous phase containing total cellular RNA was
1142 taken as loading control, and was also ligated with Cy5.

1143 **RNA sequencing library preparation**

1144 The RNA sequencing library preparation procedures are adapted from what was described for 'size-
1145 matched input' in the published reports.[29] All library preparations started with 150 ng (Qubit RNA
1146 quantification, broad range) RNA input. In brief, the RNA was first subjected to 3'-end repair (total
1147 volume 20 µl per sample): 2 µl 10x T4 PNK buffer, 2 µl FastAP, 1 µl T4 PNK, 0.5 µl SuperRasIn, 0.5 µl
1148 Turbo DNase, 37 ºC, 45 min, 900 rpm on a thermoshaker. The end-repaired RNA was purified using
1149 TRIzol in combination with Zymo RCC-5 columns following the product manual, and eluted in 10 µl
1150 nuclease-free water. The pre-adenylated 3'-adapter was ligated using the following procedure (total
1151 volume 20 µl per sample): 2 µl 10x RNA ligase buffer, 10 µM (final concentration) pre-adenylated
1152 adapter, 6 µl 50% PEG8000, 0.3 µl DMSO, 0.5 µl SuperRasIn, 2 µl T4 RNA ligase 1 high concentration,
1153 16 ºC, 3h, 1000 rpm on a thermoshaker. The excess adapter was degraded post-ligation using 1 µl 5´
1154 deadenylase (NEB, M0331S) and 2 µl RecJf (NEB, M0264S) after adding 2.5 µl 10x NEB buffer 1. The
1155 mixture was purified using TRIzol in combination with Zymo RCC-5 columns, and eluted with 10 µl
1156 nuclease-free water. Reverse transcription was performed with the following protocol: 4 µl 5x first strand
1157 buffer, 10 mM DTT, 0.8 mM dNTP, 0.5 µl SuperRaseIn, 1 µl SuperScript III (Thermo Fisher Scientific,
1158 18080093), 55 ºC, 30 min, on PCR heat block. The excess primer was degraded using 3 µl ExoSAP-
1159 IT, quenched with 1 µl 0.5 M EDTA. RNA was then removed under alkaline pH (3 µl 1 M NaOH) at
1160 elevated temperature (70 ºC, 12 min) and neutralized with 3 µl 1 M HCl. The resulted complementary
1161 DNA (cDNA) was purified using MyOne Silane beads, and eluted with 20 µl nuclease-free water. Half
1162 of the purified cDNA was subjected to PCR amplification with NEBNext Multiplex Oligos for Illumina: 30
1163 µl total reaction volume, 2x KAPA HiFi HotStart ReadyMix (Roche, 07958927001), 0.3 µM forward and
1164 reverse primers. Initial denaturation, 95 ºC, 3 min; denaturation, 98 ºC, 20 s; annealing, 65 ºC, 15 s;
1165 extension, 72 ºC, 15 s; total cycle number, 16; final extension, 72 ºC, 1 min. The amplified library was
1166 first purified with AmpureXP beads (Beckman Coulter, A63880) and then size-selected using PAGE gel
1167 electrophoresis. The resulted libraries were quantified on a Qubit tube reader and Bioanalyzer and
1168 finally mixed at equal molarity to make the 20 nM pooled library for Illumina sequencing on a Illumina
1169 Novaseq X plus platform at BMKGENE GmbH.

1170

1171 **Sequencing data analysis**

1172 UMIs (NNNNN) were extracted from the demultiplexed pair-end fastq files and appended to read
1173 identifiers by UMI-tools.[89] The adapters were trimmed by Cutadapt.[90] Read1 adapter to trim,
1174 AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC; Read2 adapter to trim,
1175 GATCGTCGGACTGTAGAACTCTGAACGTGTAGATCTCGGTGGTCGCCGTATCATT. The
1176 processed reads were pseudo-aligned and quantified at transcript level using Kallisto quant,[91] due to
1177 its outstanding ability to probabilistically assign multi-mapping reads to transcripts at high accuracy. The
1178 pseudo-alignment was done against the reference transcriptome of Gencode Release 47. The read
1179 counts for individual transcripts were normalized as RPMs.

1180

1181 **KIR2DL5 transfection, in-situ crosslinking and RNA co-precipitation**

1182 Wildtype (CHO-K1) and heparan sulfate-deficient cells (CHO pgsD-677) in a 10 cm dish at 70%
1183 confluency were transfected with 10 µg full-length KIR2DL5 plasmid (Addgene, #157623), using 2µl
1184 polyetherimide (PEI, 2mg/ml) per µg DNA in F-12K medium. 6h after transfection, the serum-free
1185 medium was replaced the one supplemented with 10% fetal bovine serum (FBS). The next day, cells
1186 were washed once with 1x DPBS following culture media removal. The cells were crosslinked in a
1187 Stratalinker for 5 min with the lid open. Once crosslinking is finished, the cells were scraped into 500µl
1188 per dish lysis buffer (50 mM Tris-HCl pH 7.4, 100 mM NaCl, 1% v/v NP-40, 0.1% v/v SDS and 0.5%
1189 w/v sodium deoxycholate) containing 1x cOmplete™ protease inhibitor cocktail (Roche, 11836153001).
1190 The cells were transferred into an Eppendorf tube and sonicated for 5 min (20% power, 5s on and 5s
1191 off). To the sonicated lysate was added 2 µl Turbo DNase, and 4 µl diluted RNase I (prediluted 1:20 in
1192 1x DPBS). The lysate was then incubated at 37 ºC for 5 minutes, followed by the addition of 2.5 µl
1193 SUPERaseIn. After centrifuging at 15,000g for 10 min at 4 °C, the supernatant was transferred to
1194 antibody-bound protein A beads (rabbit anti-DYKDDDDK tag antibody, Proteintech, 20543-1-AP; or
1195 rabbit IgG isotype control). The suspension was incubated on an end-to-end rotator at 4 °C for 2 h. The
1196 supernatant was magnetically separated and the beads were washed twice with 200 µl high-salt wash
1197 buffer (50 mM Tris-HCl pH 7.4, 1 M NaCl, 1% v/v NP-40, 1 mM EDTA, 0.1% v/v SDS and 0.5% w/v
1198 sodium deoxycholate), three times with 200 µl low-salt wash buffer (20 mM Tris-HCl pH 7.4, 10 mM
1199 $MgCl_2$, 0.2% v/v Tween-20 and 5 mM NaCl). The precipitated RNA-KIR2DL5 was 3'-end repaired and
1200 biotin ligated following the procedure described for TLR7-Fc. After the final wash, beads were directly
1201 taken up in 1x SDS sample loading buffer containing 1x sample reducing solution, denatured at 75℃
1202 for 10 minutes. The samples were resolved with SDS-PAGE, wet transferred onto a nitrocellulose (NC)
1203 membrane at 30 V for 4 hours at 4 ℃, and subjected to chemiluminescent biotinylated nucleic acid
1204 detection using streptavidin-HRP (Thermo Fisher Scientific, 89880D). For Flag tag detection, the
1205 samples were resolved with SDS-PAGE, semi-dry transferred onto a polyvinylidene fluoride (PVDF)
1206 membrane at 25 V for 30 mins, and detected using rabbit anti-DYKDDDDK tag (Proteintech, 20543-1-
1207 AP, 1:10000) in combination with HRP-conjugated goat-anti-ribbit secondary antibody.
1208
1209 **Biotinylation at total RNA fragments 3'-end (for in-vitro KIR2DL5 binding**)
1210 Isolated total RNA was treated with Turbo DNase and then proteinase K to remove residual DNA and
1211 polypeptides. The purified RNA was fragmented with 4 µl 10x RNA Fragmentation Buffer (NEB,
1212 E6186AVIAL) added to 36 µl of total RNA in DEPC-treated water. The reaction was incubated at 94 ºC
1213 for 2.5 minutes. Afterwards, 5 µl RNA Stop Solution (NEB, E6187AVIAL) was added. The fragmented
1214 RNA was kept on ice for over a minute was purified with the Zymo RCC-25 columns following the
1215 protocol provided by the supplier. To perfom end repair, 80 µg purified fragmented RNA was combined
1216 with 10 µl 10x PNK buffer, 5 µl FastAP (Thermo Fisher Scientific, EF0654), 1 µL SUPERase·In and 2
1217 µl Turbo DNase, and incuabted at 37 ºC for 30 minutes on a thermoshaker. The dephosphorylated
1218 RNA was combined with 1 ml TRIzol, from which the aqueous phase was mixed with an equivolume of
1219 ethanol and purified using Zymo RCC-25 columns. The RNA was eluted in 50 µl nuclease-free water.
1220 To the RNA solution, 14 µl 5x polymerase buffer (100 mM Tris-HCl pH 7.0, 3 mM $MnCl_2$, 0.1 mM EDTA,
1221 1 mM DTT, 0.5 mg/ml acetylated BSA, 50% glycerol (v/v)), 5 µl yeast poly(A) polymerase (Jena

1222 Bioscience, RNT-006-S), 2'-Azido-2'-dATP (Jena Biosciences, NU-976S, 50 μM final concentration)
1223 and 0.5 μl SUPERase·In were added. The solution was incubated at 37 ºC for 1 hour on a
1224 thermoshaker. The resulted RNA-3'-azide was purified as described for the end repair. Finally, 50 μl
1225 RNA-3'-azide was combined with 20 μl of 2x denaturing loading buffer (95% formamide, 25 mM EDTA).
1226 Next, the RNA was incubated with 1 mM of DBCO-biotin at 50 ºC for 10 minutes. The biotinylated RNA
1227 was purified as described above.
1228
1229 **In-vitro KIR2DL5 binding, crosslinking, and RNA co-precipitation**
1230 The 3'-end biotinylated RNA described above was denatured in DEPC-treated water at 70 ºC for 5
1231 minutes and immediately put on ice. KIR2DL5-Fc (Biotechne, 6634-KR) or IgG isotype control (0.5 μg)
1232 were incubated with 2 μg biotinylated or non-biotinylated RNA in DPBS lysis buffer (1% NP-40 (v/v),
1233 0.1% SDS (w/v), sodium deoxycholate 0.5% (w/v), 1x DPBS) at 16 ºC with gentle shaking for 30
1234 minutes. Afterwards, the samples were irradiated on ice with UVC (254nm) for 5 minutes. Next, the
1235 samples were resolved on a 12% SDS-PAGE gel, wet transferred onto a nitrocellulose (NC) membrane
1236 at 30 V for 4 hours at 4 ºC and subjected to chemiluminescent detection of biotinylated nucleic acid.
1237 The same NC membrane was subjected to Western blot for human Fc detection using AlexaFulor488-
1238 conjugated goat-anti-human IgG.
1239
1240 **AlphaFold3 prediction of structural model for KIR2DL5 and potential interacting partners**
1241 The prediction was performed on AlphaFold Server (https://alphafoldserver.com/). The sequence of
1242 KIR2DL5 extracellular domain (aa 22 – 238) was used as input. To predict nucleic acid interactions,
1243 octanucleotide repeats (A_8, U_8, G_8 and C_8) were used as input. To check for potential protein-protein
1244 interactions with cell-surface RBPs, the sequences of overlapping hits between the TLR7-proximal and
1245 biotinylaiton-crosslinking-OOPS were obtained from Uniprot and used directly as another protein entity.

Figures and captions

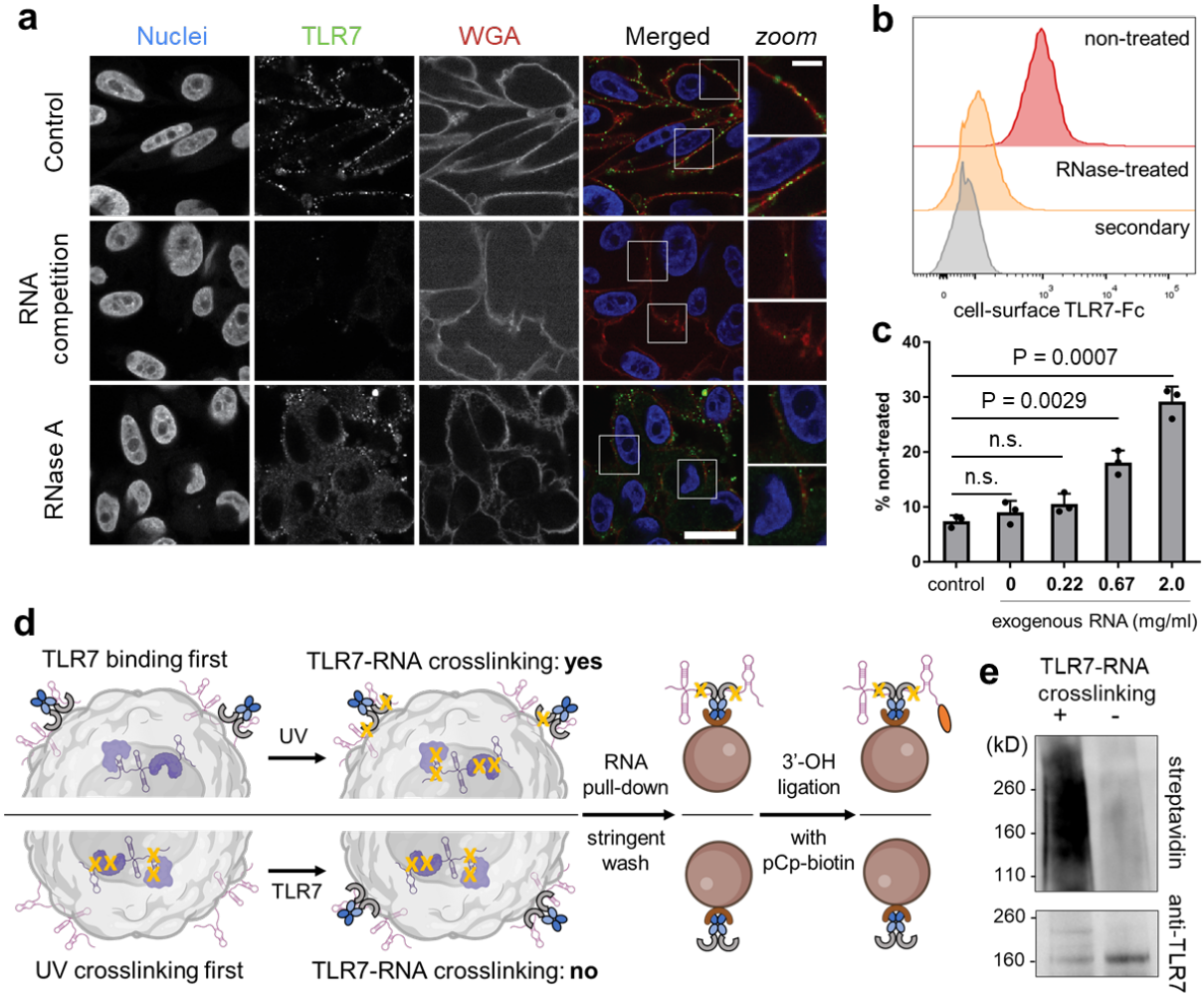


Figure 1. TLR7-Fc is a csRNA probe.

a) Confocal microscopy images of csRNA on Mel526 cells probed by TLR7-Fc and AlexaFluor488-conjugated Goat-anti-human IgG. Cell surface was stained by WGA-biotin and streptavidin-AlexaFluor594. See supplementary information for the sequence of competitor ssRNA. RNase A concentration, 40 µg/ml. Nuclear counterstaining: DAPI. Scale bar: 20µm.

b) Histograms from flow cytometry analysis of csRNA on Mel526 cells probed by TLR-Fc in complex with goat-anti-human-AlexaFluor647. 40 µg/ml RNase A was used to deplete csRNA during EDTA lifting. n = 3 independent cell cultures.

c) TLR7-Fc binding was rescued by exogenous total RNA in a concentration-dependent manner. n = 3 independent cell cultures. Exogenous RNA in 1x DPBS was added to cells at 4 ℃. 0 mg/ml RNA means 1x DPBS buffer without RNA. The control was with cells treated with RNase A but without any additional incubation. Y-axis shows relative signal intensity, with the geometric mean of TLR7 binding on non-RNase treated cells set to 100%. Two-tailed Student's t-test, mean ± s.d.

d) Schematics of in-situ crosslinking and csRNA capture experiment. pCp, cytidine-5'-phosphate-3'-(6-aminohexyl)phosphate. Orange oval represents biotin.

e) Chemiluminescent detection of TLR7-captured and biotin-labeled RNA. Rabbit-anti-TLR7 was used to check for TLR7-Fc on the same nitrocellulose membrane after blotting biotinylated RNA. n = 3 independent cell cultures for crosslinked and non-crosslinked samples.

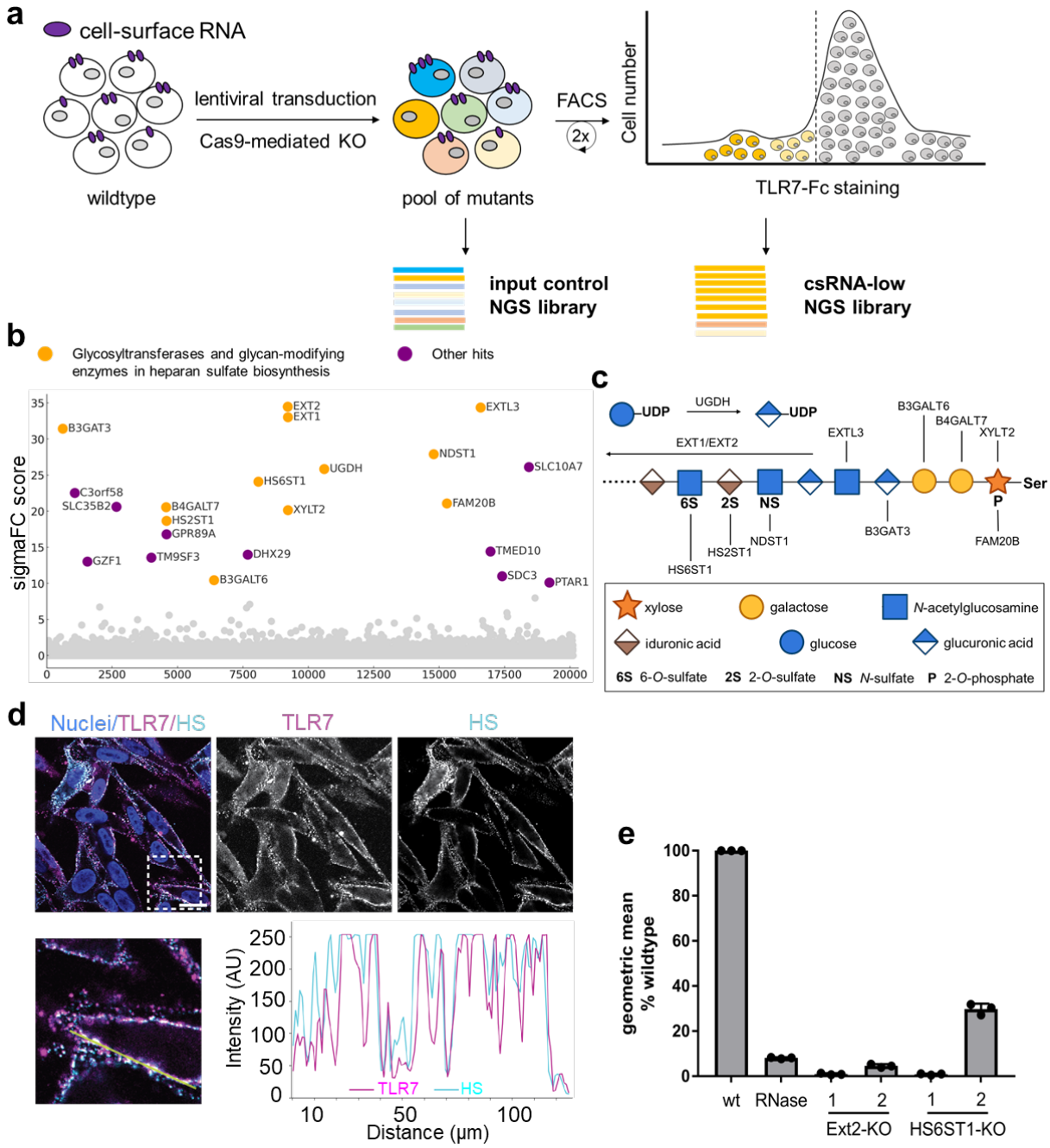


Figure 2. Genome-wide screening reveals essential factors for csRNA stable presentation.
a) Schematics of the genetic screen in Mel526 cells with csRNA stained with TLR7-Fc and AlexaFluor488-conjugated goat-anti-human IgG. Color-coded cells represent the 5% of cells with the lowest fluorescent signal, indicating the loss of csRNA.
b) Dot plot of genetic screening results. Y-axis in this plot is the geometric mean of the SigmaFC scores from the duplicates calculated using PinAPL-Py. Heparan sulfate-related glycosyltransferases and glycan-modifying enzymes are colored in yellow. Other hits with a SigmaFC larger than 10 are colored in purple.
c) HS biosynthetic pathway. Gene candidates as yellow dots in Panel **b** are shown in this schematic representation.
d) csRNA and HS colocalization on cell surface of Mel526. csRNA was probed by TLR7-Fc and AlexaFluor647-conjugated rat-anti-human IgG (magenta), while HS was labeled with mouse anti-HS (10E4 epitope) antibody and AlexaFluor488-conjugated goat anti-mouse IgM mu chain secondary antibody (cyan). The line intensity profile along the yellow line shows csRNA tightly overlaps with heparan sulfate. Nuclear staining (blue): DAPI. Scale bar: 20 µm. The frame with dashed line indicates the area where intensity profile was collected.

1283 **e**) Flow cytometry analysis showing csRNA loss in EXT2 and HS6ST1 KO mutant Mel526 cells.
1284 TLR7-Fc with goat-anti-human-AlexaFluor647 was used for csRNA staining. 40 µg/ml RNase A
1285 treated cells are used as control. The numbers above EXT2 and HS6ST1 represents different guide
1286 RNAs for KO (see Materials and Methods for sequences). n = 3 independent cell cultures.

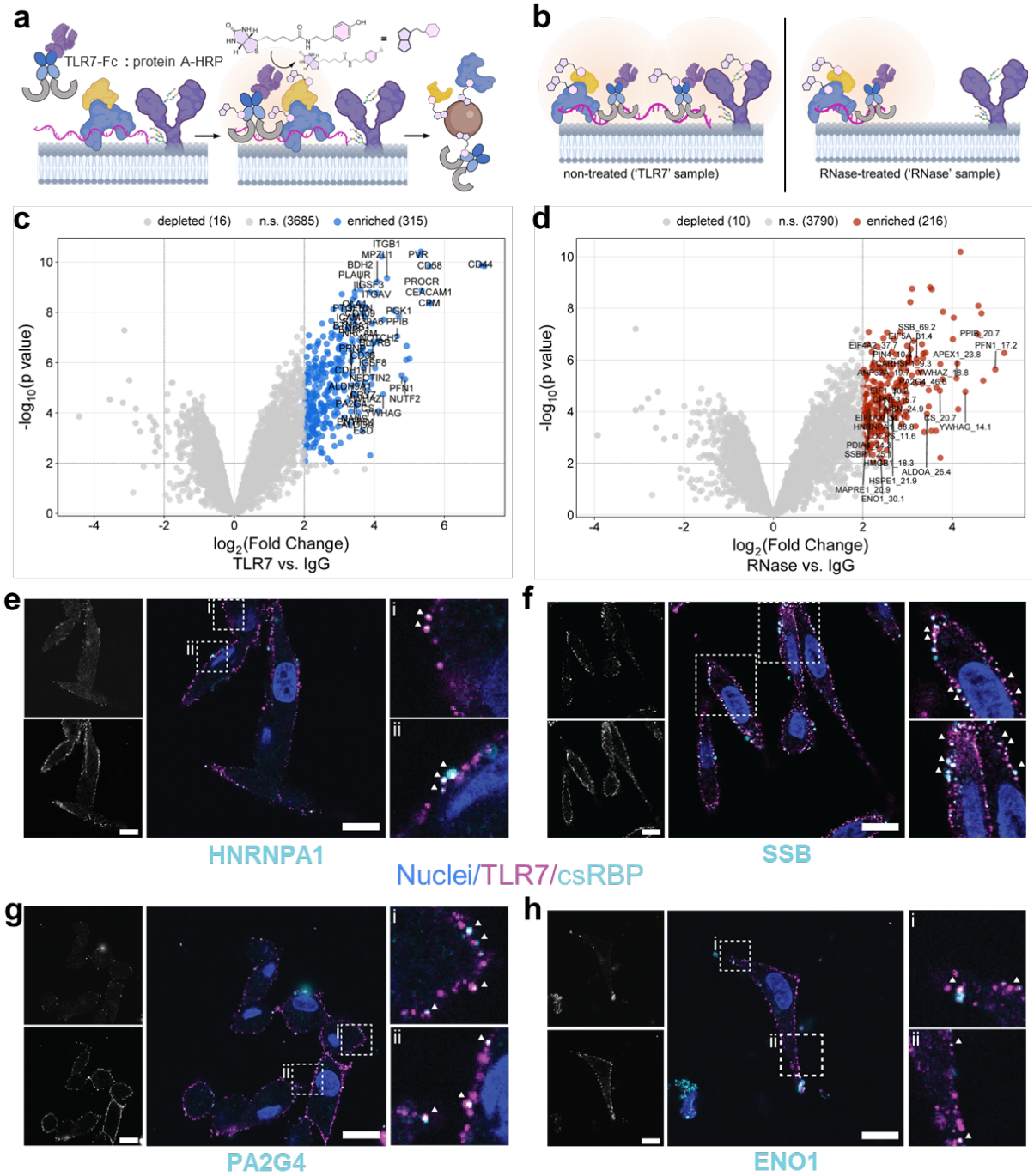


Figure 3. TLR7-proximal proteome harbors known RBPs.
a) Schematics of TLR7-mediated proximity labeling to identify hepRNA-proximal proteome. As negative control, IgG was used instead of TLR7-Fc for forming complex with protein A-HRP.
b) Extracellular RNase treated cells were also proximity-labeled in the same fashion.
c) Volcano plot from the comparison of proteome between TLR7-mediated proximity labeling and IgG control. hepRNA-proximal candidates (fold change ≥ 4, $p < 0.01$. n = 3 independent cell cultures) are shown in blue. Classically cell-surface localized proteins are labeled with their names.
d) Volcano plot from the comparison of proteome between TLR7-mediated proximity labeling post-RNase treatment and IgG control. hepRNA-proximal candidates are shown in red. RBP2GO database-documented proteins (RBP2GO score > 10) are labeled with their names and the RBP2GO scores.

1299 **e**) – **h**) Confocal microscopy was performed to confirm cell-surface colocalization with hepRNA for
1300 select csRBP candidates. The csRBPs were detected with rabbit anti-RBP antibodies and Cy3-
1301 conjugated goat anti-rabbit secondary antibody. Nuclear staining: DAPI. Scale bars: 20μm.

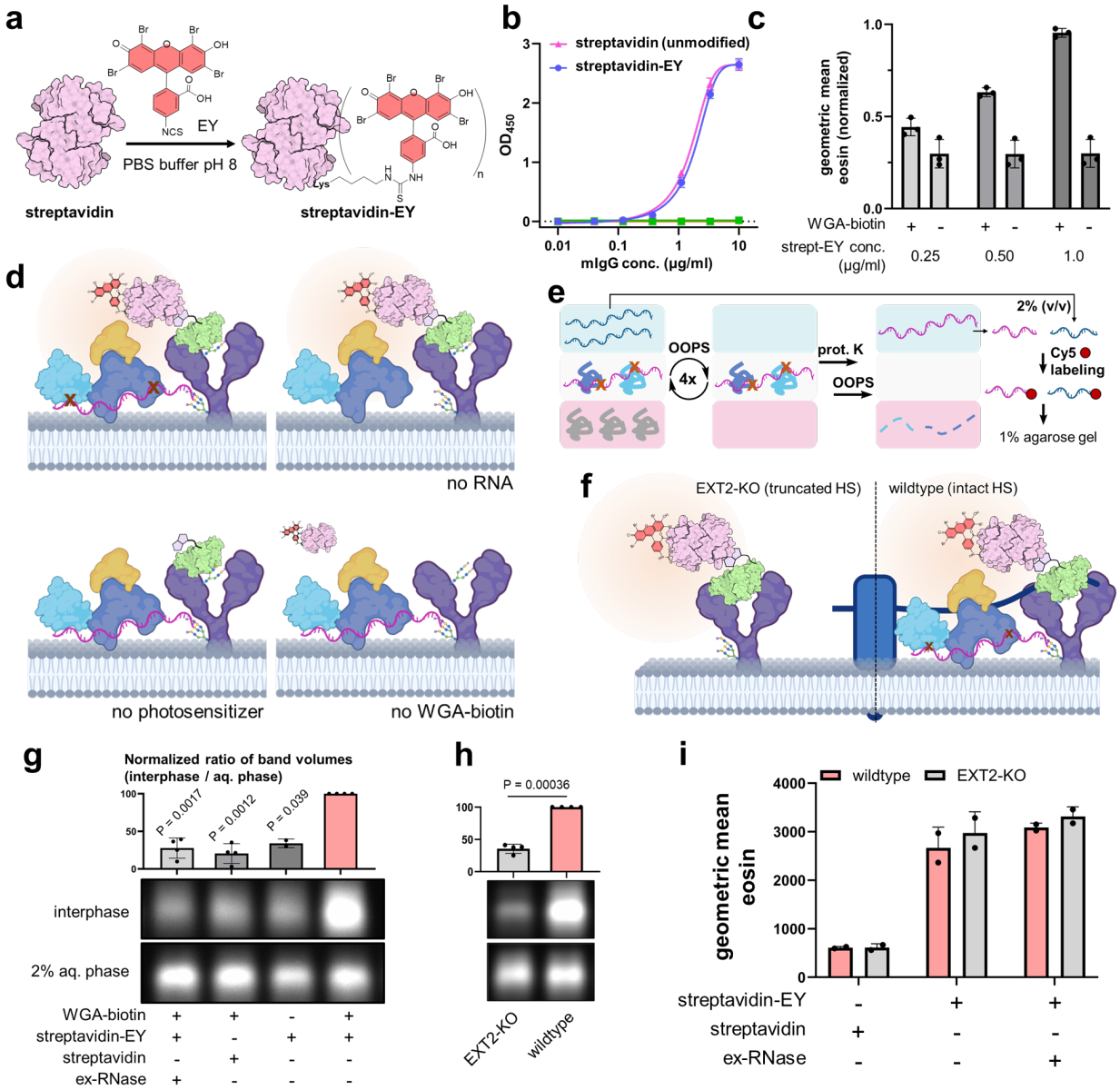


Figure 4. HepRNA HS-dependency validation by TLR7-independent cell-surface crosslinking experiments.

a) Reaction scheme for eosin isothiocyanate (EY)-streptavidin conjugation. Isothiocyanate generally react with primary amines on proteins, forming thiourea.

b) Enzyme-linked immunosorbent assay (ELISA) demonstrates modified streptavidin retains full capacity for biotin binding. Blue circle, with streptavidin-EY coated on the plate, incubated biotinylated IgG; pink triangle, unmodified streptavidin, biotinylated IgG; green square, streptavidin-EY, unmodified IgG; orange triangle (inverted), unmodified streptavidin, unmodified IgG.

c) Titration of streptavidin-EY on live cell surface with and without biotinylated WGA deposition.

d) Schematics of EY-mediated cell surface crosslinking experiment and controls. In no-RNA control, lives cells were pretreated with RNase I.

e) Sample processing after live cell surface crosslinking. Abbreviations: OOPS, orthogonal organic phase separation; prot.K, proteinase K; Cy5, for this experiment, cyanine 5-conjugated cytidine-5′-phosphate-3′-(6-aminohexyl)phosphate.

f) Schematics of EY-mediated cell surface crosslinking on wildtype and HS-deficient cells.

g) 1% agarose gel electrophoresis of 3’-end Cy5-labeled RNA from samples depicted in panel **d**. Gel image were cropped for small RNA region. Bar graphs show the normalized band volumes corrected by aqueous phase total RNA. n = 4 independent cell cultures; n = 2 for no-WGA-biotin control. Error bars, s.d.

1322 **h**) Image and bar graphs of 1% agarose gel electrophoresis of 3'-end Cy5-labeled RNA from wildtype
1323 and EXT2-KO (HS-deficient) cells. The wildtype samples were produced in the same batch as EXT2-
1324 KO samples. n = 4 independent cell cultures.
1325 **i**) Bar graphs of flow cytometry quantification of streptavidin-EY recruited on the surface of wildtype
1326 cell with or without RNase treatment and EXT2-KO cells.

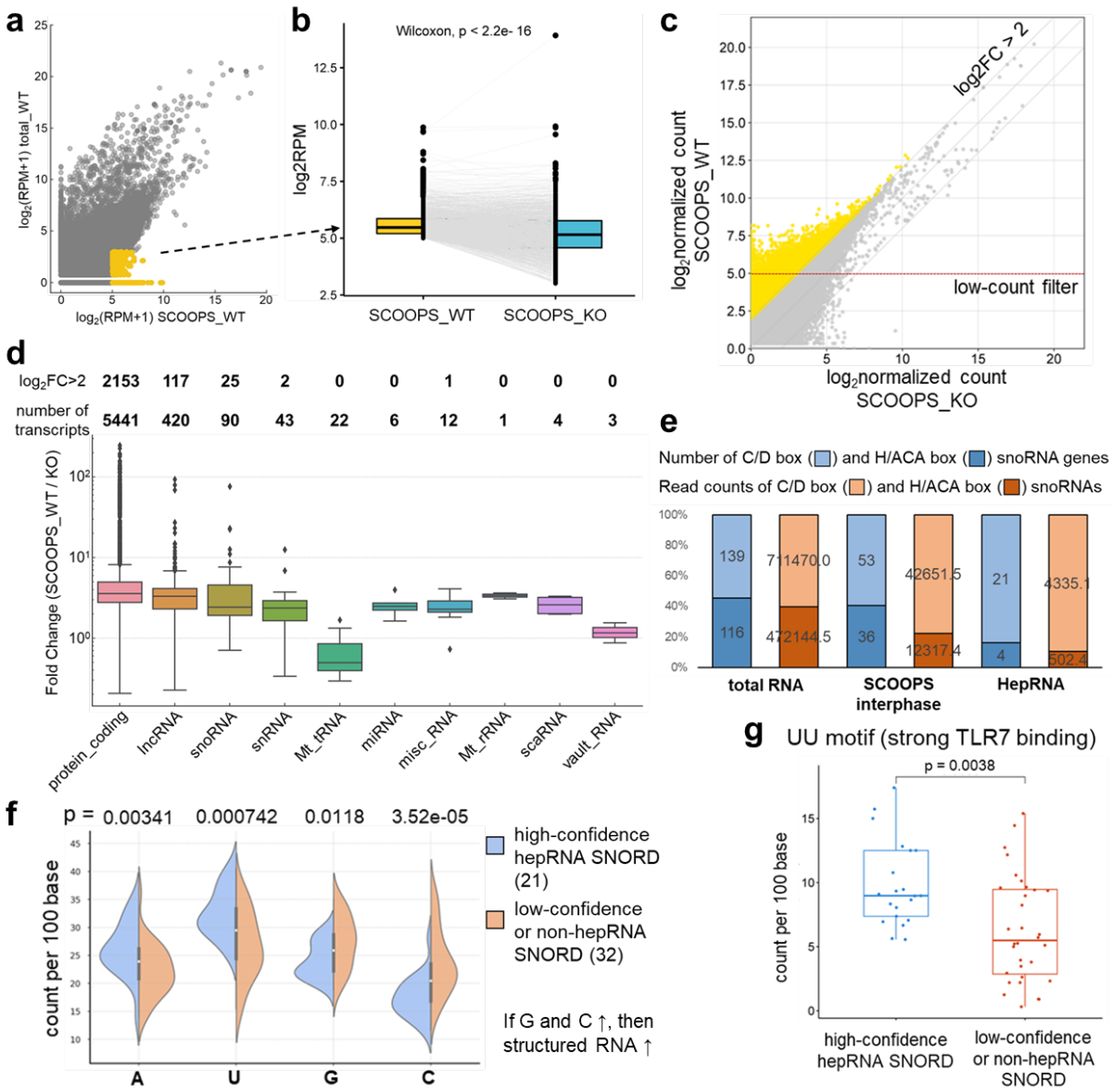


Figure 5. HepRNA identification via next-generation sequencing.
a) Scatter plot of mean RPM values. n = 2 independent SCOOPS experiments. Mean (RPM+1) was log2-transformed. Transcripts are displayed without low-count filtering. X-axis is SCOOPS_WT sample, and Y-axis is total cellular RNA from wildtype cells.
b) Connected box plot of RPM of SCOOPS interphase-retained transcripts in wildtype (SCOOPS_WT) and EXT-KO (SCOOPS_KO) samples. Grey lines connect corresponding transcripts in both datasets. Statistical significance, Wilcoxon test.
c) Scatter plot of mean read counts normalized on background RNA species, comparing SCOOPS_WT (Y-axis) with SCOOPS_KO (X-axis). High-confidence hepRNA (yellow dots) defined as the transcripts which have 4 times higher read counts post-normalization (on background RNA) in SCOOPS_WT than in SCOOPS_KO. Transcripts with normalized read counts larger than 32 in SCOOPS_WT were taken for further analysis in panel **d** – **g**.
d) Box plot summarizing fold change values per RNA biotype derived from SCOOPS_WT over _KO.
e) Bar graph illustrating relative composition of snoRNA subtypes in total transcriptome, SCOOPS interphase, and as high-confidence hepRNA.
f) Violin plot showing significant difference in nucleotide compositions between high-confidence hepRNA and low-confidence/non-hepRNA SNORD transcripts. Mann-Whitney U test was used to derive p values for each nucleotide.

1347 **g**) Box plot showing strong TLR7-binding motif (UU) occurs more frequently in high-confidence
1348 hepRNA SNORD transcripts than in low-confidence/non-hepRNA ones. Statistical significance, Mann-
1349 Whitney U test.

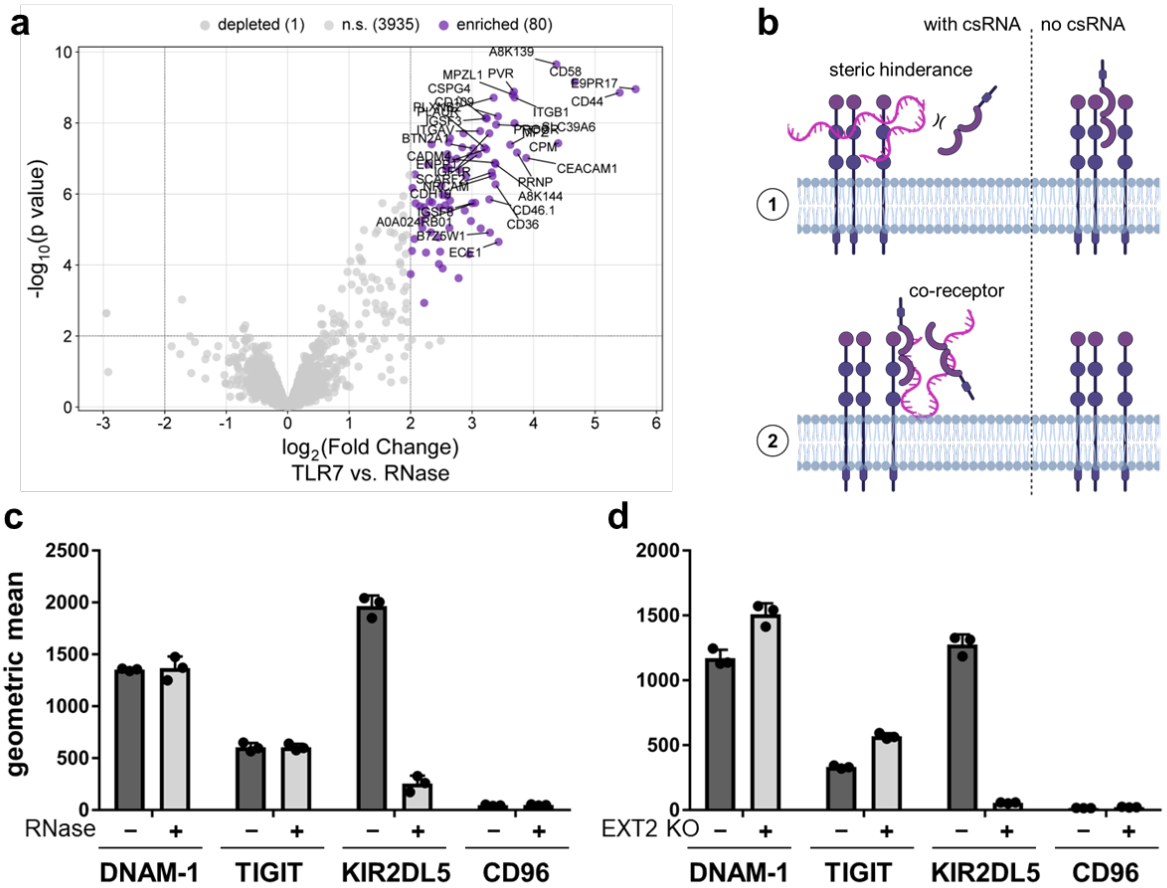


Figure 6. HepRNA recruits KIR2DL5 on cell surface.

a) Volcano plot from the comparison of proteome between TLR7-mediated proximity labeling with vs. without RNase treatment. Significantly enriched proteins with high fold change (≥ 8, $p < 0.01$. n = 3 independent cell cultures) are labeled with their names.

b) Schematic representation of how csRNA may regulate ligand-receptor interactions on the plasma membrane.

c) Bar graphs from flow cytometry analysis of live cell surface binding of PVR-binding Fc-fused recombinant proteins, in the presence or absence of csRNA. 40 µg/mL RNase A was used to deplete csRNA. AlexaFluor647-conjugated goat-anti-human was used to detect Fc-fused proteins on live cell surface. n = 3 independent cell cultures.

d) Bar graphs from flow cytometry analysis live cell surface binding of PVR-binding Fc-fused recombinant proteins. Wildtype Mel526 and the EXT-KO mutant cells were compared side-by-side. n = 3 independent cell cultures.

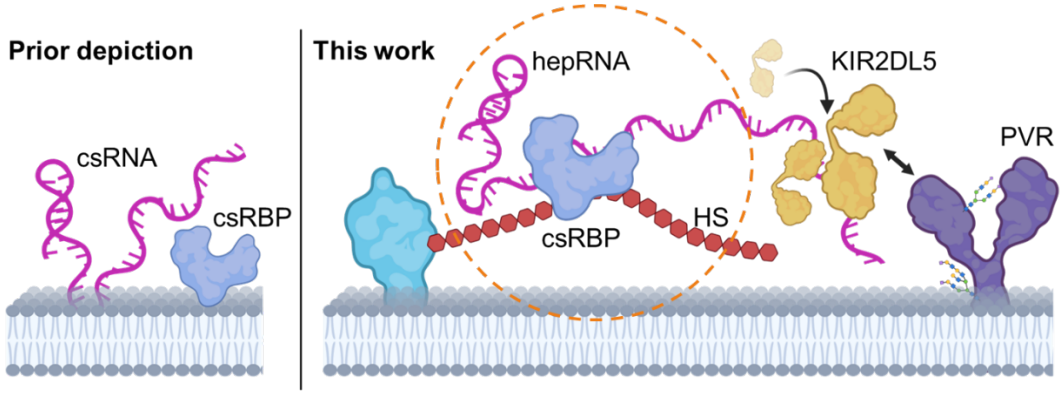
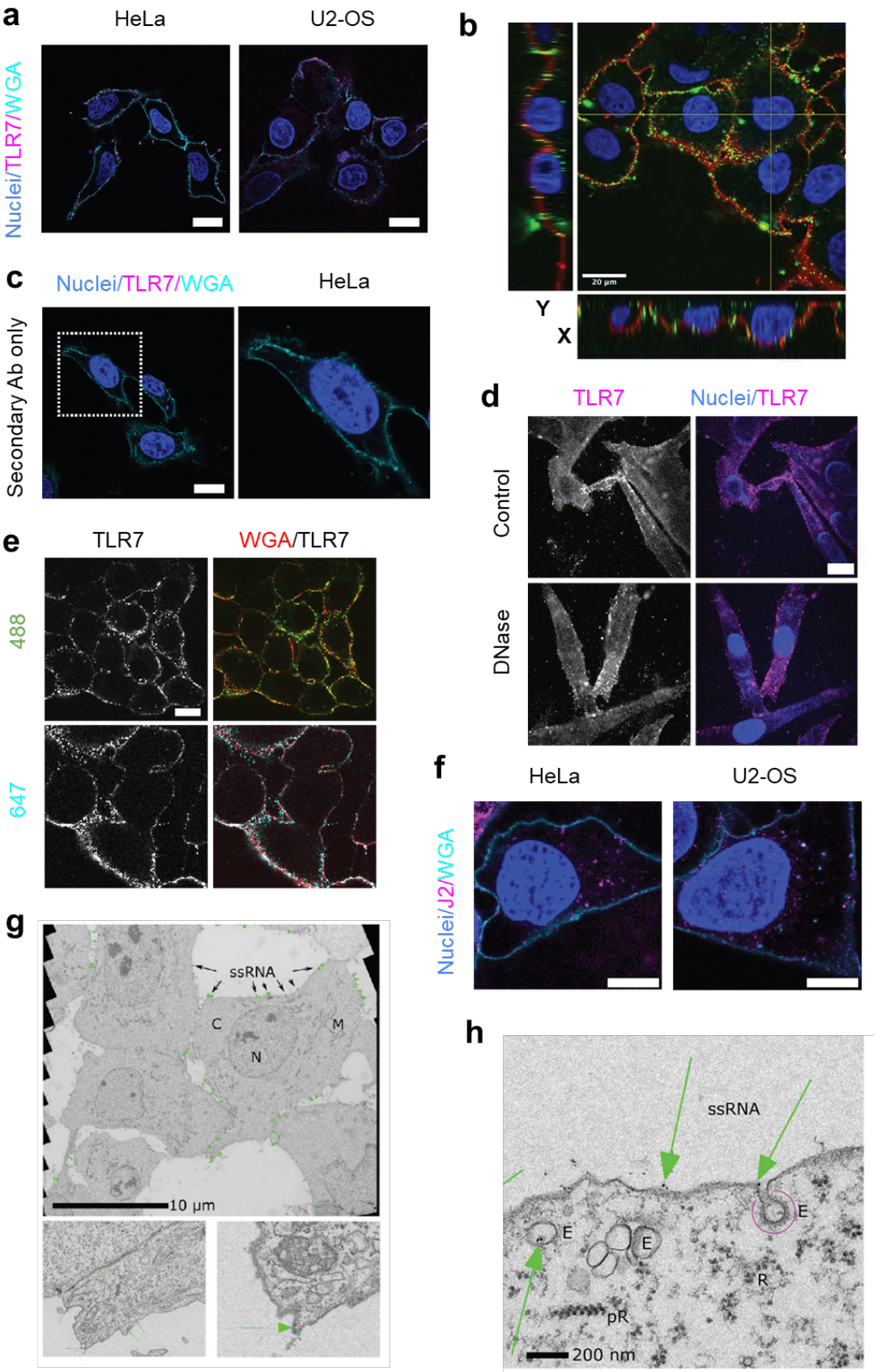
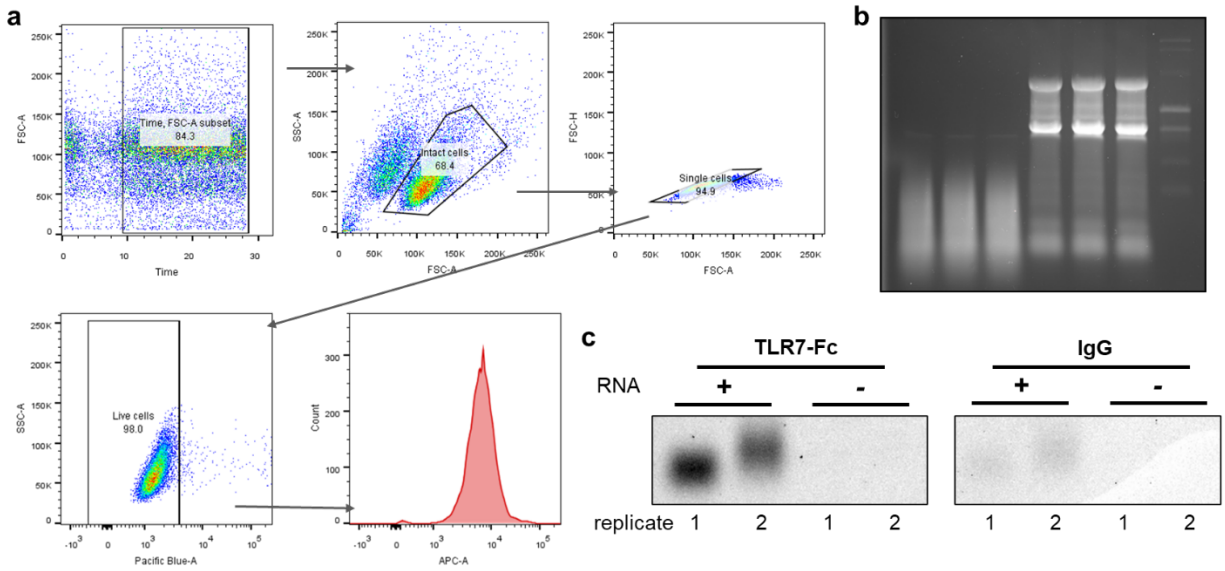


Figure 7. Schematic representation of advanced understanding cell-surface RNA biology. The hepRNA-csRBP-HS ternary complex model is shown in dashed orange circle.



Supplementary Figure 1. Immunofluorescence imaging on different cell lines using TLR7-Fc. **a**) csRNA staining on HeLa and U2-OS cells probed by TLR7-Fc and AlexaFluor647-conjugated rat-anti-human IgG Fc antibody (HeLa) or AlexaFluor488-conjugated goat-anti-human IgG Fc antibody

(U2-OS) pseudocoloured as magenta. Cell surface was stained by WGA-biotin and streptavidin-AlexaFluor488 (HeLa) or streptavidin-AlexaFluor594 (U2-OS), pseudocoloured as cyan.
b) Optical sections through U2-OS cells labeled with TLR7-Fc and AlexaFluor488-conjugated rat-anti-human IgG Fc antibody, pseudocoloured green. It shows the occurrence of csRNA on the cellular surface. Cell surface was stained by WGA-biotin and streptavidin-AlexaFluor594, pseudocoloured red.
c) Representative confocal microscopy images of secondary antibody control (U2-OS cells) i.e. without TLR7-Fc but with AlexaFluor647-conjugated rat-anti-human IgG Fc antibody. Inset shows enlarged view.
d) csRNA staining in the presence or absence of DNase I (10 U/ml) on Mel526 cells probed by TLR7-Fc and AlexaFluor647-conjugated rat-anti-human IgG Fc antibody.
e) csRNA staining on living cells incubated at 4 ℃ with TLR7 and two different secondary antibodies in U2-OS cells (AlexaFluor488- or AlexaFluor647-conjugated anti-human IgG Fc antibodies.
f) Confocal images of double-stranded RNA staining using HeLa and U2-OS cells probed by the J2 antibody and Cy3-conjugated donkey-anti-mouse IgG. Cell surface was stained by WGA-biotin and streptavidin-AlexaFluor488. Nuclear staining: DAPI. Scale bars: 10μm (**a**), 20μm (**b** - **f**).
g) The ultrastructure of csRNA (green dots, indicated with arrows) localizations in U2-OS cells. TLR7-Fc served as a primary probe. 10nm gold nanoparticles conjugated to Protein A used as a secondary probe provide the necessary contrast in electron microscopy. Bottom panels provide different fields of view.
h) Observation of csRNA internalized into the cell via endocytosis. The green arrows indicates the presence of csRNA inside endosomes with or without caveolae (pink arc). Abbreviations: C, cytosol; N, nucleus; M, mitochondria, E, endosome, R, ribosome; pR, polyribosome. Scale bars: 10μm (**g**), 200nm (**h**).

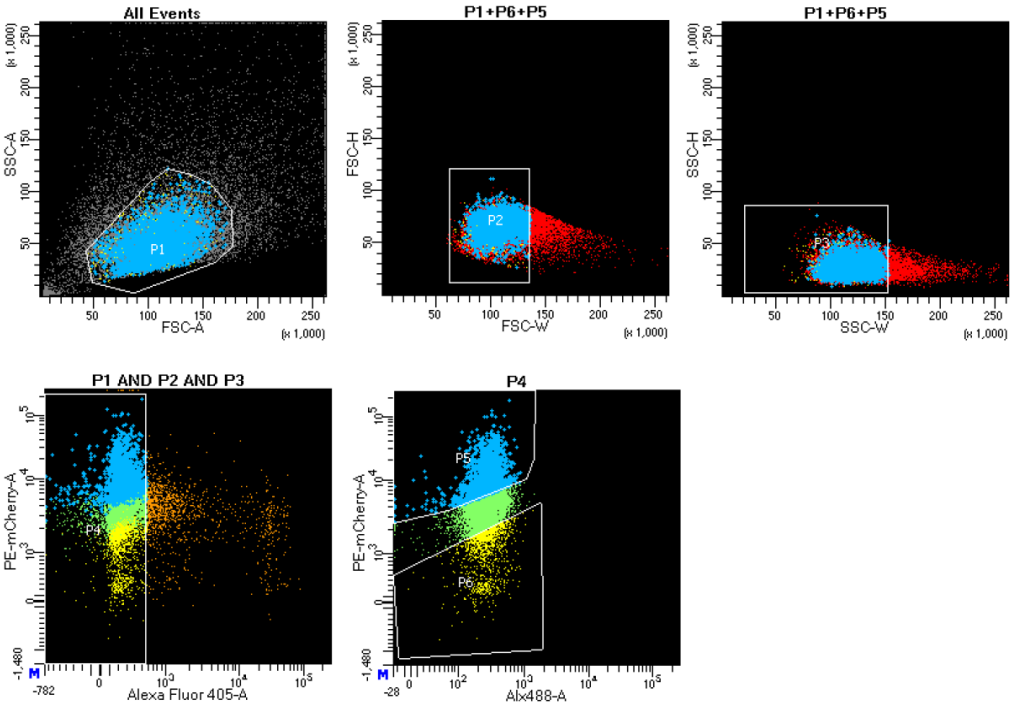


Supplementary Figure 2. Additional data on flow cytometry analysis, exogenous RNA rescue, and in-vitro TLR7-RNA crosslinking and pull-down.

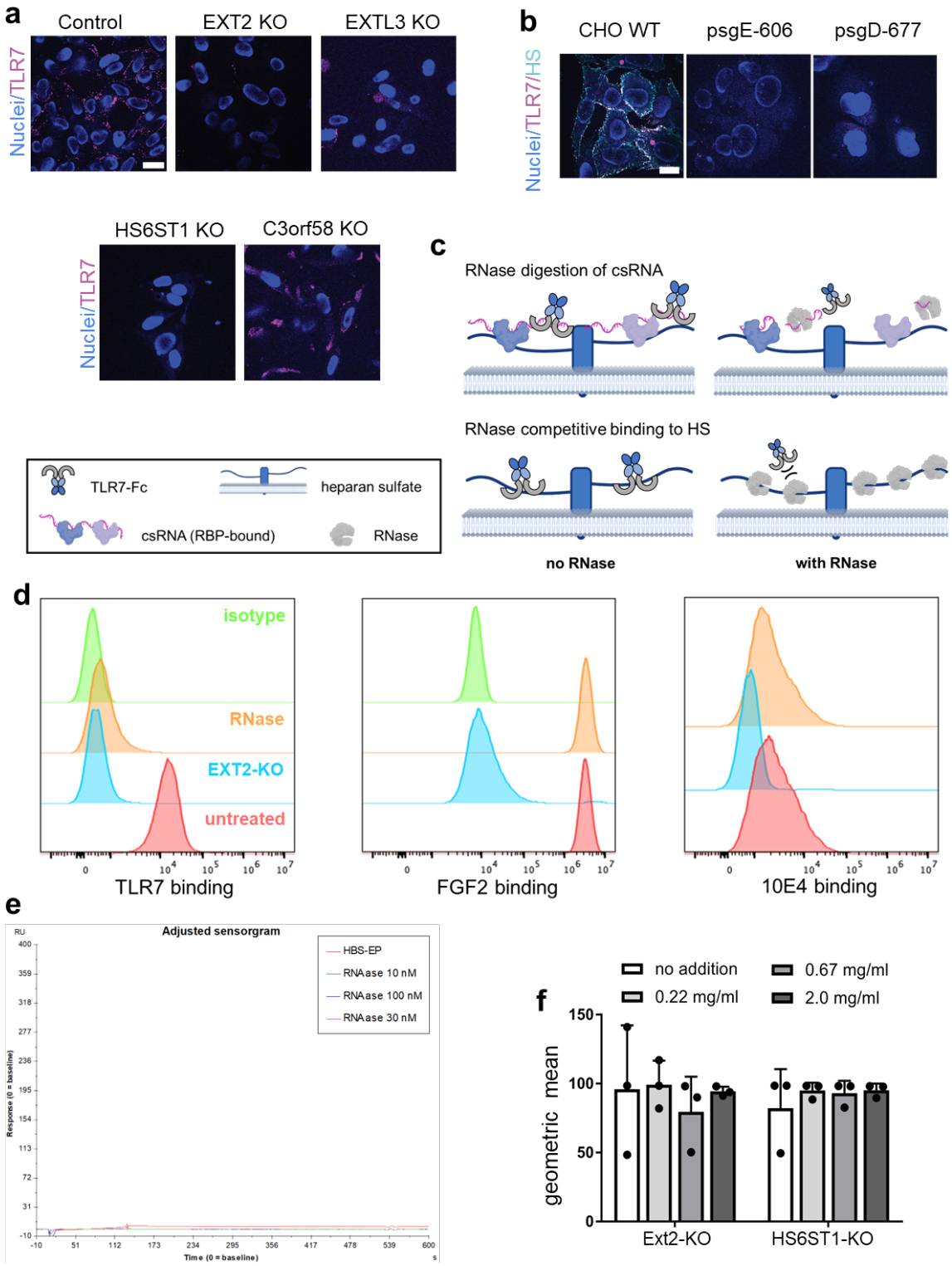
a) Gating strategy for TLR7-Fc probed cells in flow cytometry.

b) 1% agarose gel analysis of extracted total RNA. Left three lanes were 2-minute fragmented total RNA (from the RNA in the right three lanes) using magnesium-based fragmentation buffer.

c) 1% agarose gel electrophoresis of on-bead 3'-end Cy5-labeled RNA released from bound TLR7-Fc. IgG was used as an RNA non-binding control. Replicate numbers mean the crosslinking experiments were performed with different batches of fragmented RNA independently. In replicate 1, 3'-end Cy5 ligation was performed overnight, and in replicate 2 for 4 hours.

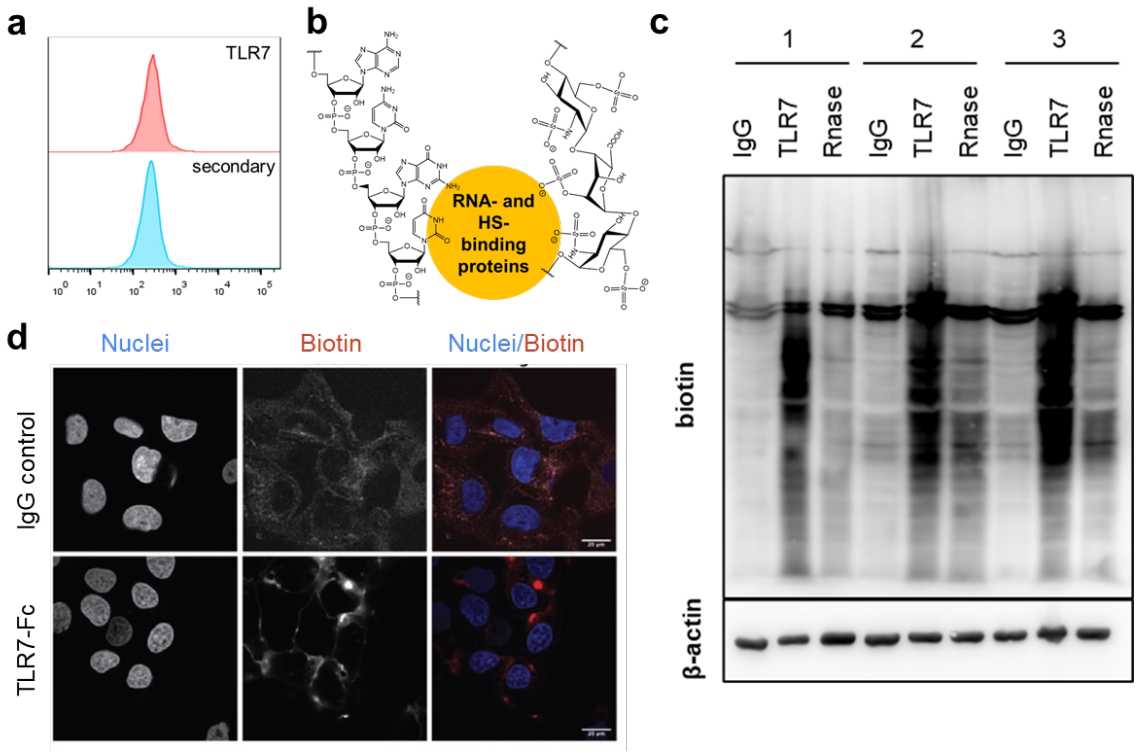


Supplementary Figure 3. Gating strategies in FACS to enrich csRNA-low and -high phenotypes among lentiviral transduced cells. P1 gated for intact cells; P2 and P3 gated for singlets; P4 gated for living cells; P5 gated csRNA-high phenotype. P6 gated csRNA-low phenotype.

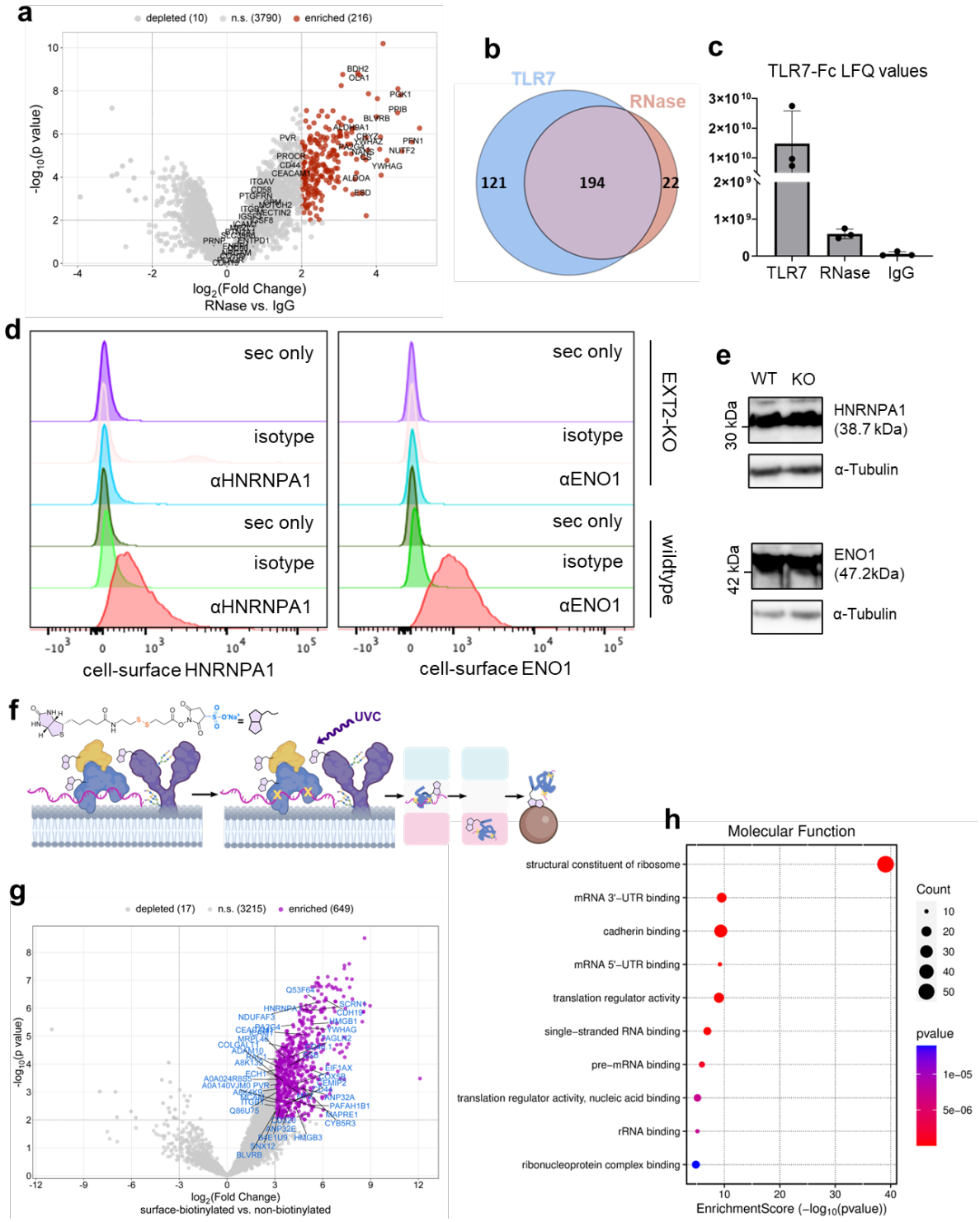


Supplementary Figure 4. Validation of HS as essential factors for csRNA presentation.
a) csRNA staining of EXT2, EXTL3, HS6ST1 and C3orf58 KO mutant cell lines (see Materials and Methods for sequences). TLR7-Fc was detected by AlexaFluor488-conjugated goat-anti-human IgG for csRNA (magenta).
b) csRNA (magenta) and HS (cyan) staining of CHO-K1 (wild type) and heparan sulfate deficient cell lines. The primary secondary antibodies were the same as used in Figure 3d. Nuclear staining: DAPI. Scale bars: 20 µm (**d** - **e**).

1419 **c**) Schematic representation of an alternative scenario which could have given rise to the observation
1420 in genetic screening using TLR7-Fc.
1421 **d**) Histograms from flow cytometry analysis of RNase responsiveness of TLR7-Fc and well-known HS
1422 binders, including fibroblast growth factor 2 (FGF2) and anti-HS IgM (10E4 epitope). n = 3
1423 independent cell cultures. Untagged recombinant FGF2 was pre-complexed with rabbit anti-FGF2.
1424 **e**) Surface plasmon resonance sensorgram of PureLink RNase A with heparin-coated chip.
1425 **f**) TLR7-Fc binding was not rescued by exogenous total RNA for HS-deficient mutant cells. n = 3
1426 independent cell cultures. None of the groups is significant compared to no-RNA control (the white
1427 bar on the left). Two-tailed Student's t-test, mean ± s.d.



Supplementary Figure 5. Rationales and supporting data for TLR7-mediated proximity labeling.
a) Histograms from flow cytometry analysis of proteolytically lifted cells, which lost TLR7-Fc binding.
b) Postulated model for RNA-HS association. Because both are negatively charged, they need a bridging molecule in between, which is hypothesized to be proteins.
c) Western blot of cell-surface proximity biotinylated proteins. IgG, wildtype cells incubated with human IgG isotype control precomplexed with protein A-HRP; TLR7, wildtype cells incubated with TLR7-Fc precomplexed with protein A-HRP; RNase, extracellularly RNase A-treated cells incubated with TLR7-Fc precomplexed with protein A-HRP. The numbers above sample names indicate replicates.
d) Fluorescence images of proximity biotinylation of extracellular RNase-treated cells. IgG was used as a control to check for background labeling. Red foci indicate biotinylation of HepRNA proximal proteins via biotin-phenol in the presence of HRP and hydrogen peroxide.



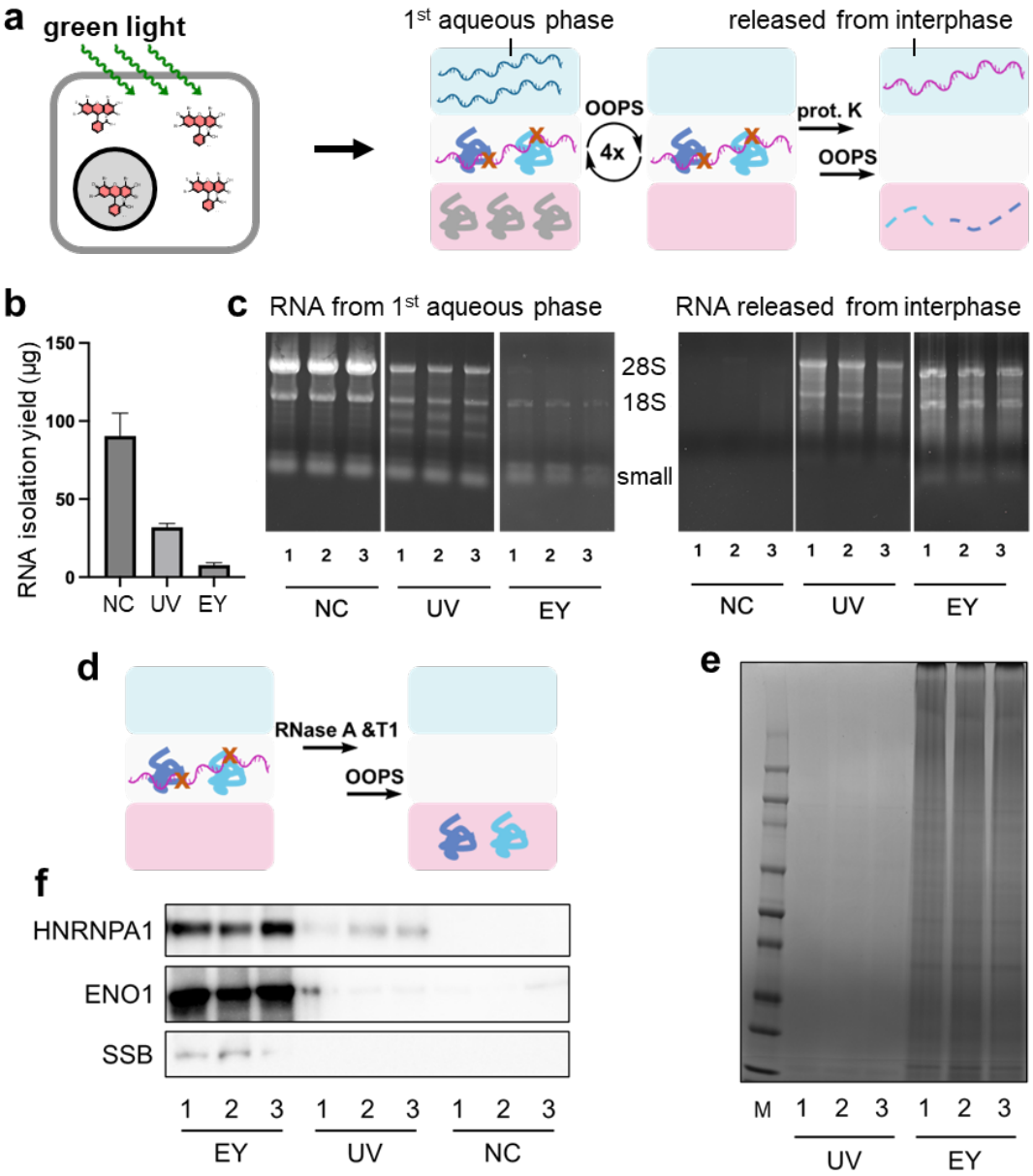
Supplementary Figure 6. Validation of cell-surface localization and RNA complexation of csRBP candidates.

a) Volcano plot from the comparison of proteome between TLR7-mediated proximity labeling post-RNase treatment and IgG control. hepRNA-proximal candidates are shown in red. Classically cell-surface localized proteins are labeled with their names to demonstrate loss of enrichment in RNase sample for many of these proteins.

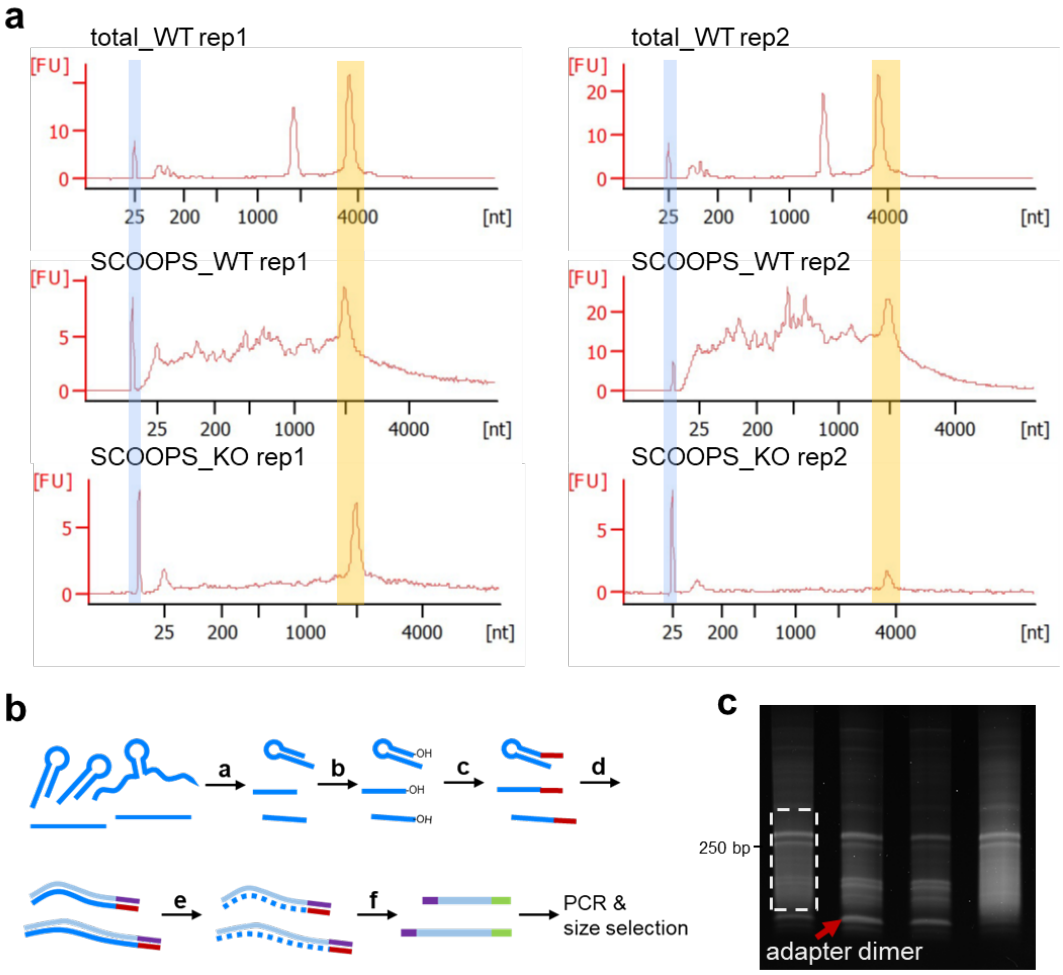
b) Venn diagrams showing consistently and uniquely enriched proteins (vs. IgG control) from TLR7 and RNase samples.

c) Label-free quantification (LFQ) values of TLR7 indicates residual TLR7 were bound to cell surface after extracellular RNase treatment.

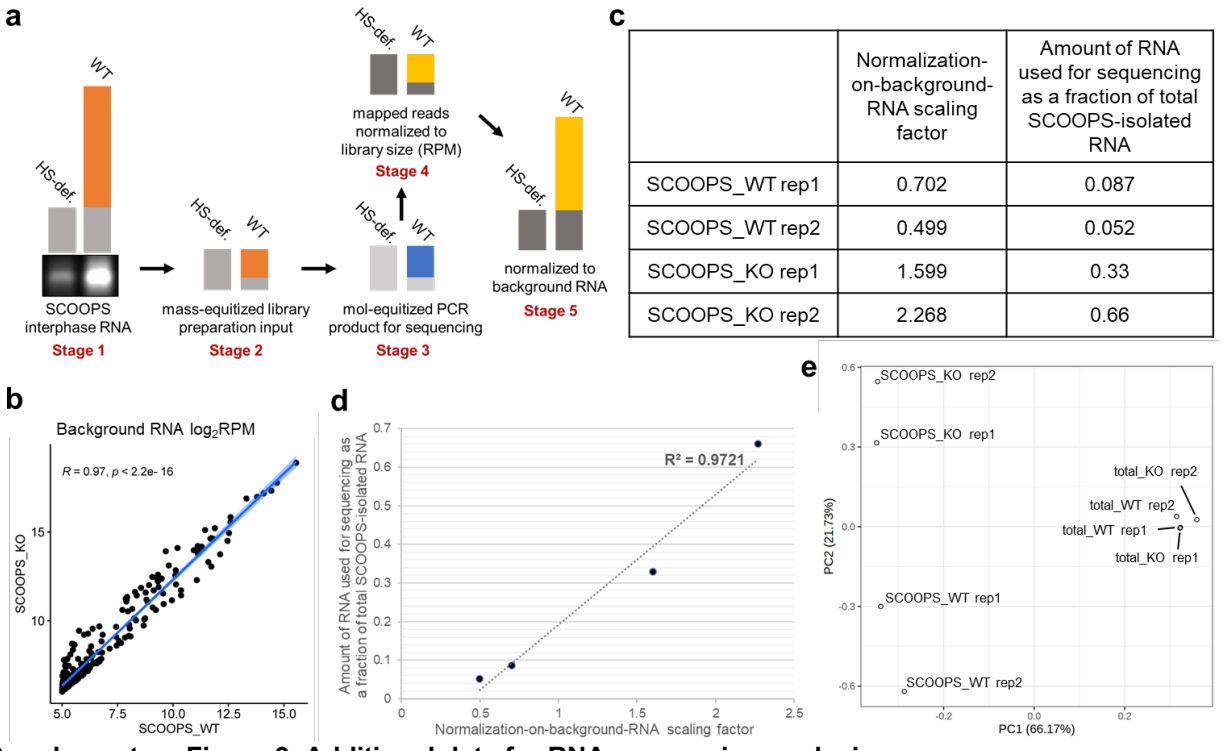
1452 **d**) Histograms from flow cytometry analysis demonstrating the presence of HNRNPA1 and ENO1 on
1453 the cell surface, and their dependency on intact HS chain.
1454 **e**) Western blot analysis of total cellular proteins indicates HS deficiency did not negatively affect the
1455 expression of HNRNPA1 or ENO1.
1456 **f**) Schematics of workflow for surface biotinylation, UVC crosslinking, OOPS and affinity pulldown to
1457 enrich csRNA-bound csRBPs. Cell-impermeable, amine-reactive biotin reagent tags the entire cell
1458 surface proteome. OOPS enriches RNA-bound proteins.
1459 **g**) Volcano plot from the biotinylation-OOPS experiment. The comparison is with biotinylated, OOPS-
1460 isolated samples over non-biotinylated, OOPS-isolated ones. Strongly enriched proteins (fold change
1461 ≥ 8, $p < 0.01$. $n = 3$ independent cell cultures) are shown in pink. csRBP candidates as TLR7-
1462 proximal proteins (**Supplementary Table 3**) are labelled with their names, and are considered to
1463 have been in complex with RNA on cell surface.
1464 **h**) Gene ontology analysis (molecular function) of strongly enriched proteins from biotinylation-OOPS
1465 experiments indicates the specificity of OOPS for isolating RNA-bound proteins.



Supplementary Figure 7. Observation of singlet oxygen-mediated RNA-protein crosslinking.
a) Eosin Y (EY)-photosensitized singlet oxygen generation in living cells and APGC isolation of crosslinked RNA-protein complexes. Abbreviation: prot.K, proteinase K; OOPS, orthogonal organic phase separation.
b) Yields of RNA isolated from the first aqueous phase decreased in EY- and UV-crosslinked cells compared to no-crosslinking (NC) control. n = 3 independent cell cultures. Error bars, s.d.
c) 1% agarose gel electrophoresis of RNA isolated from the first aqueous phase and released from interphase after OOPS. Lane numbers in each sample represent independent cell cultures.
d) Proteins crosslinked to RNA were released from the interphase and brought into the second organic phase.
e) Coomassie staining of SDS-PAGE resolved proteins from the second organic phase of EY- and UV-crosslinked samples.
f) Western blot of common RNA-binding proteins isolated in the second organic phase. HNRNPA1, heterogeneous nuclear ribonucleoprotein A1; ENO1, enolase 1; SSB, Sjögren syndrome antigen B.



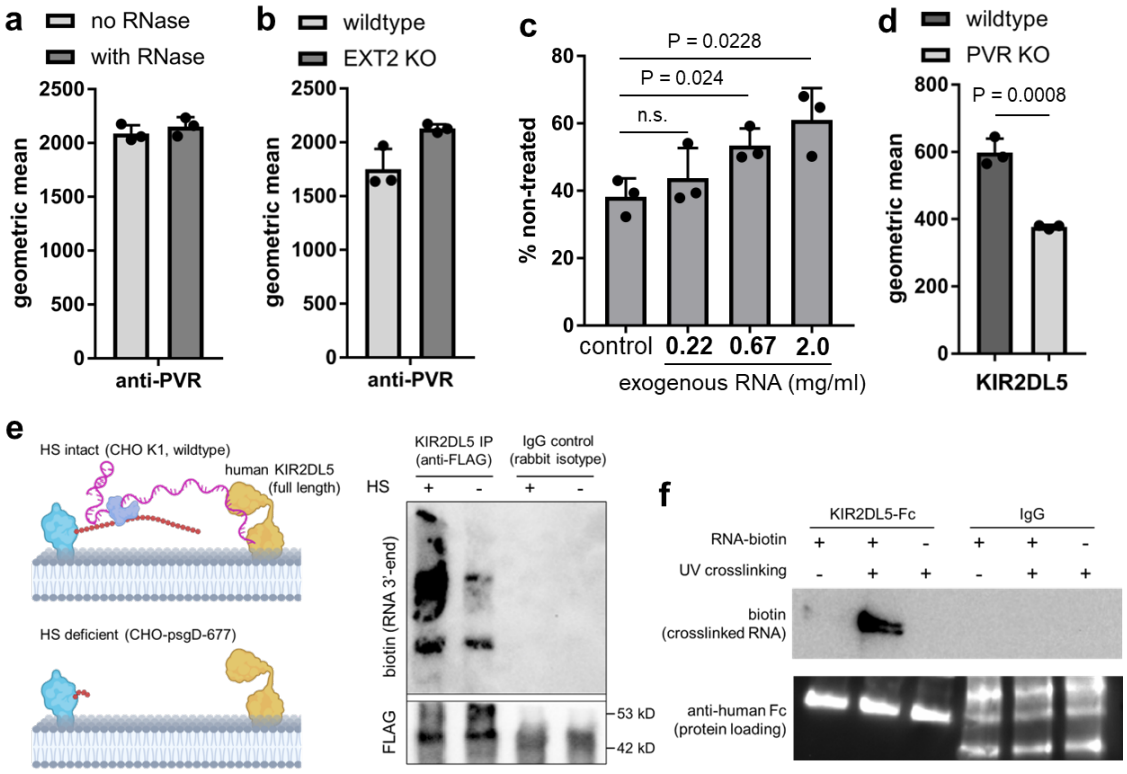
Supplementary Figure 8. Supporting data for sequencing library preparation.
a) Bioanalyzer electropherograms of all samples prior to RNA fragmentation. Blue shaded area indicates lower marker, and yellow shaded area indicates shift of 28S ribosomal RNA. For SCOOPS_WT rep1 and rep2, and SCOOPS_KO rep1, the processing software incorrectly assigned the ladder positions.
b) Schematic representation of library preparation procedures for all samples. (a) Magnesium-based fragmentation; (b) 3'-end dephosphorylation; (c) pre-adenylated adapter ligation at RNA 3'-end; (d) reverse transcription; (e) RNA removal using sodium hydroxide; (f) second adapter ligation at 3'end of complementary DNA.
c) A representative gel for PCR product size selection. Area circled by a white dashed frame right above the adapter dimer band (indicated by a red arrow) was excised in each lane. Samples from left to right, SCOOPS_WT rep1, SCOOPS_KO rep1, SCOOPS_KO rep2, SCOOPS_WT rep2.



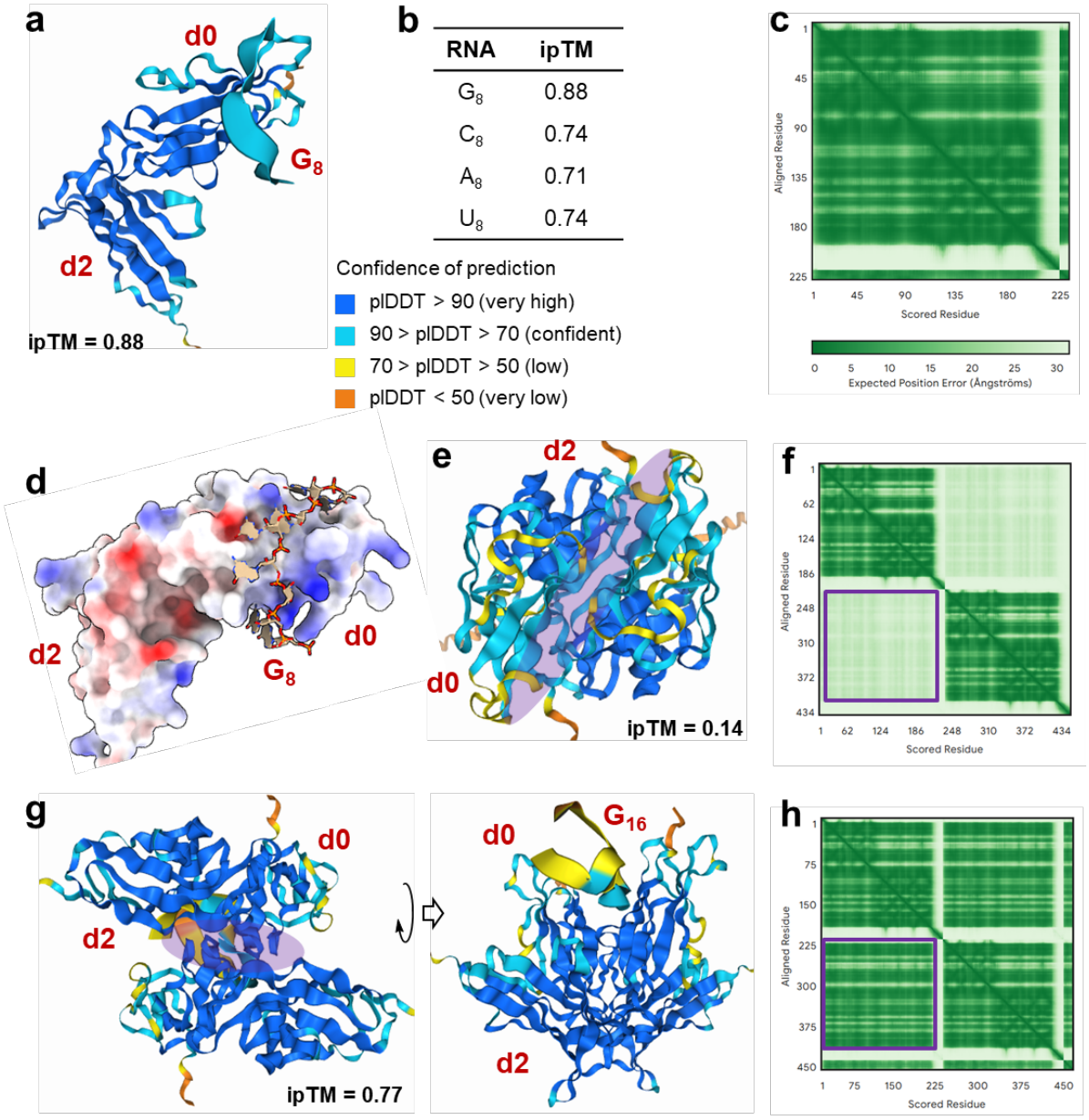
c

	Normalization-on-background-RNA scaling factor	Amount of RNA used for sequencing as a fraction of total SCOOPS-isolated RNA
SCOOPS_WT rep1	0.702	0.087
SCOOPS_WT rep2	0.499	0.052
SCOOPS_KO rep1	1.599	0.33
SCOOPS_KO rep2	2.268	0.66

Supplementary Figure 9. Additional data for RNA sequencing analysis.
a) Illustration of rationales for normalization on background RNA prior to comparing SCOOPS_WT and KO samples. See **Supplementary Text 3** for a detailed description. HS-def., heparan sulfate-deficient sample, equivalent to EXT-KO. Grey bar stands for background RNA species in all stages.
b) Pearson correlation analysis revealed background RNA species (see **Supplementary Text 3** for definition) from SCOOPS_WT and KO samples (prior to normalization on background RNA) were highly similar (R = 0.97). RPM of background RNA transcripts larger than 32 in both samples were log2-transformed prior to plotting the scatter plot.
c) and **d**) Library-specific scaling factors generated by median of ratios normalization correlates well with the amount of library preparation input (150 ng) as fractions of total SCOOPS-isolated RNA. To scale read counts for each library, RPM values were divided by the scaling factor.
e) Principal component analysis (PCA) of all samples. Normalized read counts were filtered (transcripts with read count < 10 were removed) and log2-transformed.



Supplementary Figure 10. Supporting data for KIR2DL5 recruitment by hepRNA.
a) Bar graphs from flow cytometry analysis of live cell-surface binding of monoclonal anti-PVR antibody, in the presence or absence of csRNA. Primary antibody was detected by AlexaFluor488-conjugated goat-anti-mouse secondary antibody. n = 3 independent cell cultures.
b) Bar graphs from flow cytometry analysis of live cell-surface binding of monoclonal anti-PVR antibody to wildtype and EXT2-KO mutant Mel526 cells. n = 3 independent cell cultures.
c) KIR2DL5-Fc binding was rescued by exogenous total RNA in a concentration-dependent manner. n = 3 independent cell cultures.
d) Bar graphs from flow cytometry analysis of live cell surface binding of KIR2DL5-Fc to wildtype and PVR-KO mutant cells. n = 3 independent cell cultures.
e) In-situ crosslinking of RNA to KIR2DL5 and immune-coprecipitation experiments in wildtype and HS-deficient cells. Precipitated KIR2DL5-bound RNA was on-bead labeled with pCp-biotin. FLAG tag was used to check for equal protein loading. Signals in IgG control were a result of polyclonal goat-anti-rabbit IgG HRP conjugate recognizing rabbit IgG released from beads.
f) Western blot analysis of in-vitro crosslinked KIR2DL5-Fc and biotnylated RNA. Human IgG isotype control was included to check for non-specific binding.



Supplementary Figure 11. AlphaFold3-predicted structural model for KIR2DL5-RNA interaction. **a**) Predicted model for ectodomain of KIR2DL5 interacting with octaguanosine (G_8) RNA. Immunoglobulin domains d0 and d2 of KIR2DL5 were indicated as red labels. The ribbon structures were colored by predicted local distance difference test (plDDT) scores, a measure of local confidence. See **Supplementary Table 8** for input sequences for AlphaFold3 prediction. **b**) Interface predicted template modeling (ipTM) scores for the KIR2DL5 ectodomain sequence and difference octanucleotide RNA oligos. Confident high-quality predictions are with values higher than 0.8. **c**) Predicted aligned error (PAE) matrix for KIR2DL5 ectodomain in complex with G_8. **d**) Electrostatic surface potential of KIR2DL5 ectodomain. Immunoglobulin domain d0 contains a patch of positively charged residues proximal to the interphase between RNA and the protein. **e**) and **f**) Predicted model for dimerization of ectodomain of KIR2DL5 (protein only) and the corresponding predicted aligned error (PAE) matrix. The interphase between the two monomers is indicated by a purple shade. **g**) and **h**) Predicted model for dimerization of ectodomain of KIR2DL5 in the presence of a 16-mer RNA of guanosine (G_{16}) and the corresponding predicted aligned error (PAE) matrix.

References

1. Block, K. F.; Puerta-Fernandez, E.; Wallace, J. G.; Breaker, R. R. Association of OLE RNA with bacterial membranes via an RNA–protein interaction. *Molecular Microbiology* 2011, **79**(1)**:** 21-34.

2. Flynn, R. A.; Pedram, K.; Malaker, S. A.; Batista, P. J.; Smith, B. A. H.; Johnson, A. G.*, et al.* Small RNAs are modified with N-glycans and displayed on the surface of living cells. *Cell* 2021, **184**(12)**:** 3109-3124.e3122.

3. Huang, N.; Fan, X.; Zaleta-Rivera, K.; Nguyen, T. C.; Zhou, J.; Luo, Y.*, et al.* Natural display of nuclear-encoded RNA on the cell surface and its impact on cell interaction. *Genome Biology* 2020, **21**(1)**:** 225.

4. Wu, E.; Guo, X.; Teng, X.; Zhang, R.; Li, F.; Cui, Y.*, et al.* Discovery of Plasma Membrane-Associated RNAs through APEX-seq. *Cell Biochemistry and Biophysics* 2021, **79**(4)**:** 905-917.

5. Hemberger, H.; Chai, P.; Lebedenko, C. G.; Caldwell, R. M.; George, B. M.; Flynn, R. A. Rapid and sensitive detection of native glycoRNAs. *bioRxiv* 2023**:** 2023.2002.2026.530106.

6. Zhang, N.; Tang, W.; Torres, L.; Wang, X.; Ajaj, Y.; Zhu, L.*, et al.* Cell surface RNAs control neutrophil recruitment. *Cell* 2024, **187**(4)**:** 846-860.e817.

7. Semenkovich, C. F.; Ostlund, R. E., Jr.; Olson, M. O. J.; Yang, J. W. A protein partially expressed on the surface of HepG2 cells that binds lipoproteins specifically is nucleolin. *Biochemistry* 1990, **29**(41)**:** 9708-9713.

8. Hee Jo, S.; Choi, J.-A.; Lim, Y.-J.; Lee, J.; Cho, S.-N.; Oh, S.-M.*, et al.* Calreticulin modulates the intracellular survival of mycobacteria by regulating ER-stress-mediated apoptosis. *Oncotarget; Vol 8, No 35* 2017.

9. Knorr, K.; Rahman, J.; Erickson, C.; Wang, E.; Monetti, M.; Li, Z.*, et al.* Systematic evaluation of AML-associated antigens identifies anti-U5 SNRNP200 therapeutic antibodies for the treatment of acute myeloid leukemia. *Nature Cancer* 2023, **4**(12)**:** 1675-1692.

10. Gillissen, M. A.; Kedde, M.; Jong, G. d.; Moiset, G.; Yasuda, E.; Levie, S. E.*, et al.* AML-specific cytotoxic antibodies in patients with durable graft-versus-leukemia responses. *Blood* 2018, **131**(1)**:** 131-143.

11. Perr, J.; Langen, A.; Almahayni, K.; Nestola, G.; Chai, P.; Lebedenko, C. G.*, et al.* RNA binding proteins and glycoRNAs form domains on the cell surface for cell penetrating peptide entry. *bioRxiv* 2023**:** 2023.2009.2004.556039.

12. Vallet, S. D.; Clerc, O.; Ricard-Blum, S. Glycosaminoglycan–Protein Interactions: The First Draft of the Glycosaminoglycan Interactome. *Journal of Histochemistry & Cytochemistry* 2020, **69**(2)**:** 93-104.

13. Esko, J. D.; Lindahl, U. Molecular diversity of heparan sulfate. *The Journal of Clinical Investigation* 2001, **108**(2)**:** 169-173.

14. Lamanna, W. C.; Kalus, I.; Padva, M.; Baldwin, R. J.; Merry, C. L. R.; Dierks, T. The heparanome—The enigma of encoding and decoding heparan sulfate sulfation. *Journal of Biotechnology* 2007, **129**(2)**:** 290-307.

15. Ma, Y.; Guo, W.; Mou, Q.; Shao, X.; Lyu, M.; Garcia, V.*, et al.* Spatial imaging of glycoRNA in single cells with ARPLA. *Nature Biotechnology* 2024, **42**(4)**:** 608-616.

16. Petes, C.; Odoardi, N.; Gee, K. The Toll for Trafficking: Toll-Like Receptor 7 Delivery to the Endosome. *Frontiers in Immunology* 2017, **8**.

17. Crawford, J. D.; Wang, H.; Trejo-Zambrano, D.; Cimbro, R.; Talbot, C. C., Jr.; Thomas, M. A.*, et al.* The XIST lncRNA is a sex-specific reservoir of TLR7 ligands in SLE. *JCI Insight* 2023, **8**(20)**:** e169344.

18. Lehmann, S. M.; Krüger, C.; Park, B.; Derkow, K.; Rosenberger, K.; Baumgart, J.*, et al.* An unconventional role for miRNA: let-7 activates Toll-like receptor 7 and causes neurodegeneration. *Nature Neuroscience* 2012, **15**(6)**:** 827-835.

19. Pawar, K.; Shigematsu, M.; Sharbati, S.; Kirino, Y. Infection-induced 5′-half molecules of tRNAHisGUG activate Toll-like receptor 7. *PLOS Biology* 2020, **18**(12)**:** e3000982.

20. Zhang, Z.; Ohto, U.; Shibata, T.; Taoka, M.; Yamauchi, Y.; Sato, R.*, et al.* Structural Analyses of Toll-like Receptor 7 Reveal Detailed RNA Sequence Specificity and Recognition Mechanism of Agonistic Ligands. *Cell Reports* 2018, **25**(12)**:** 3371-3381.e3375.

21. Monsigny, M.; Roche, A.-C.; Sene, C.; Maget-Dana, R.; Delmotte, F. Sugar-Lectin Interactions: How Does Wheat-Germ Agglutinin Bind Sialoglycoconjugates? *European Journal of Biochemistry* 1980, **104**(1)**:** 147-153.

22. Fischer, L. S.; Klingner, C.; Schlichthaerle, T.; Strauss, M. T.; Böttcher, R.; Fässler, R.*, et al.* Quantitative single-protein imaging reveals molecular complex formation of integrin, talin, and kindlin during cell adhesion. *Nature Communications* 2021, **12**(1)**:** 919.

23. Owen, D. M.; Rentero, C.; Rossy, J.; Magenau, A.; Williamson, D.; Rodriguez, M.*, et al.* PALM imaging and cluster analysis of protein heterogeneity at the cell surface. *Journal of Biophotonics* 2010, **3**(7)**:** 446-454.

24. Melan, M. A.; Sluder, G. Redistribution and differential extraction of soluble proteins in permeabilized cultured cells Implications for immunofluorescence microscopy. *Journal of Cell Science* 1992, **101**(4)**:** 731-743.

25. Tanaka, K. A. K.; Suzuki, K. G. N.; Shirai, Y. M.; Shibutani, S. T.; Miyahara, M. S. H.; Tsuboi, H.*, et al.* Membrane molecules mobile even after chemical fixation. *Nature Methods* 2010, **7**(11)**:** 865-866.

26. Helmreich, E. J. M. Environmental influences on signal transduction through membranes: a retrospective mini-review. *Biophysical Chemistry* 2002, **100**(1)**:** 519-534.

27. Dhir, A.; Dhir, S.; Borowski, L. S.; Jimenez, L.; Teitell, M.; Rötig, A.*, et al.* Mitochondrial double-stranded RNA triggers antiviral signalling in humans. *Nature* 2018, **560**(7717)**:** 238-242.

28. Weber, F.; Wagner, V.; Rasmussen Simon, B.; Hartmann, R.; Paludan Søren, R. Double-Stranded RNA Is Produced by Positive-Strand RNA Viruses and DNA Viruses but Not in Detectable Amounts by Negative-Strand RNA Viruses. *Journal of Virology* 2006, **80**(10)**:** 5059-5064.

29. Van Nostrand, E. L.; Pratt, G. A.; Shishkin, A. A.; Gelboin-Burkhart, C.; Fang, M. Y.; Sundararaman, B.*, et al.* Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nature Methods* 2016, **13**(6)**:** 508-514.

30. Sanson, K. R.; Hanna, R. E.; Hegde, M.; Donovan, K. F.; Strand, C.; Sullender, M. E.*, et al.* Optimized libraries for CRISPR-Cas9 genetic screens with multiple modalities. *Nature Communications* 2018, **9**(1)**:** 5416.

31. Marques, C.; Reis, C. A.; Vivès, R. R.; Magalhães, A. Heparan Sulfate Biosynthesis and Sulfation Profiles as Modulators of Cancer Signalling and Progression. *Frontiers in Oncology* 2021, **11**.

32. Kamiyama, S.; Suda, T.; Ueda, R.; Suzuki, M.; Okubo, R.; Kikuchi, N.*, et al.* Molecular Cloning and Identification of 3'-Phosphoadenosine 5'-Phosphosulfate Transporter. *Journal of Biological Chemistry* 2003, **278**(28)**:** 25958-25963.

33. Durin, Z.; Dubail, J.; Layotte, A.; Legrand, D.; Cormier-Daire, V.; Foulquier, F. SLC10A7, an orphan member of the SLC10 family involved in congenital disorders of glycosylation. *Human Genetics* 2022, **141**(7)**:** 1287-1298.

34. Vanderperre, B.; Muraleedharan, A.; Dorion, M. F.; Larroquette, F.; Del Cid Pellitero, E.; Rajakulendran, N.*, et al.* A genome-wide CRISPR/Cas9 screen identifies genes that regulate the cellular uptake of α-synuclein fibrils by modulating heparan sulfate proteoglycans. *bioRxiv* 2023**:** 2023.2009.2029.560170.

35. van den Born, J.; Salmivirta, K.; Henttinen, T.; Östman, N.; Ishimaru, T.; Miyaura, S.*, et al.* Novel Heparan Sulfate Structures Revealed by Monoclonal Antibodies*. *Journal of Biological Chemistry* 2005, **280**(21)**:** 20516-20523.

36. Li, J.; Han, S.; Li, H.; Udeshi, N. D.; Svinkina, T.; Mani, D. R.*, et al.* Cell-Surface Proteomic Profiling in the Fly Brain Uncovers Wiring Regulators. *Cell* 2020, **180**(2)**:** 373-386.e315.

37. Yang, Z.; Zhang, Y.; Fang, Y.; Zhang, Y.; Du, J.; Shen, X.*, et al.* Spatially barcoding biochemical reactions using DNA nanostructures unveil a major contact mechanism in proximity labeling. *bioRxiv* 2024.

38. Oakley, J. V.; Buksh, B. F.; Fernández, D. F.; Oblinsky, D. G.; Seath, C. P.; Geri, J. B.*, et al.* Radius measurement via super-resolution microscopy enables the development of a variable radii proximity labeling platform. *Proceedings of the National Academy of Sciences* 2022, **119**(32)**:** e2203027119.

39. König, J.; Zarnack, K.; Rot, G.; Curk, T.; Kayikci, M.; Zupan, B.*, et al.* iCLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. *Nature Structural & Molecular Biology* 2010, **17**(7)**:** 909-915.

40. Huppertz, I.; Perez-Perri, J. I.; Mantas, P.; Sekaran, T.; Schwarzl, T.; Russo, F.*, et al.* Riboregulation of Enolase 1 activity controls glycolysis and embryonic stem cell differentiation. *Molecular Cell* 2022, **82**(14)**:** 2666-2680.e2611.

41. Timchenko, L. T.; Iakova, P.; Welm, A. L.; Cai, Z. J.; Timchenko, N. A. Calreticulin Interacts with C/EBPα and C/EBPβ mRNAs and Represses Translation of C/EBP Proteins. *Molecular and Cellular Biology* 2002, **22**(20)**:** 7242-7257.

42. Queiroz, R. M. L.; Smith, T.; Villanueva, E.; Marti-Solano, M.; Monti, M.; Pizzinga, M.*, et al.* Comprehensive identification of RNA–protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nature Biotechnology* 2019, **37**(2)**:** 169-178.

43. Trendel, J.; Schwarzl, T.; Horos, R.; Prakash, A.; Bateman, A.; Hentze, M. W.*, et al.* The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. *Cell* 2019, **176**(1)**:** 391-403.e319.

44. Urdaneta, E. C.; Vieira-Vieira, C. H.; Hick, T.; Wessels, H.-H.; Figini, D.; Moschall, R.*, et al.* Purification of cross-linked RNA-protein complexes by phenol-toluol extraction. *Nature Communications* 2019, **10**(1)**:** 990.

45. Koh, E.; Fluhr, R. Singlet oxygen detection in biological systems: Uses and limitations. *Plant Signaling & Behavior* 2016, **11**(7)**:** e1192742.

46. Lin, Z.; Schaefer, K.; Lui, I.; Yao, Z.; Fossati, A.; Swaney, D. L.*, et al.* Multiscale photocatalytic proximity labeling reveals cell surface neighbors on and between cells. *Science*, **385**(6706)**:** eadl5763.

47. Müller, M.; Gräbnitz, F.; Barandun, N.; Shen, Y.; Wendt, F.; Steiner, S. N.*, et al.* Light-mediated discovery of surfaceome nanoscale organization and intercellular receptor interaction networks. *Nature Communications* 2021, **12**(1)**:** 7036.

48. Luo, H.; Tang, W.; Liu, H.; Zeng, X.; Ngai, W. S. C.; Gao, R.*, et al.* Photocatalytic Chemical Crosslinking for Profiling RNA–Protein Interactions in Living Cells. *Angewandte Chemie International Edition* 2022, **61**(27)**:** e202202008.

49. Chan, C. Y.; Carmack, C. S.; Long, D. D.; Maliyekkel, A.; Shao, Y.; Roninson, I. B.*, et al.* A structural interpretation of the effect of GC-content on efficiency of RNA interference. *BMC Bioinformatics* 2009, **10**(1)**:** S33.

50. Zhang, J.; Kuo, C. C. J.; Chen, L. GC content around splice sites affects splicing through pre-mRNA secondary structures. *BMC Genomics* 2011, **12**(1)**:** 90.

51. Caudron-Herger, M.; Rusin, S. F.; Adamo, M. E.; Seiler, J.; Schmid, V. K.; Barreau, E.*, et al.* R-DeeP: Proteome-wide and Quantitative Identification of RNA-Dependent Proteins by Density Gradient Ultracentrifugation. *Molecular Cell* 2019, **75**(1)**:** 184-199.e110.

52. Caudron-Herger, M.; Jansen, R. E.; Wassmer, E.; Diederichs, S. RBP2GO: a comprehensive pan-species database on RNA-binding proteins, their interactions and functions. *Nucleic Acids Research* 2021, **49**(D1)**:** D425-D436.

53. Stanietsky, N.; Simic, H.; Arapovic, J.; Toporik, A.; Levy, O.; Novik, A.*, et al.* The interaction of TIGIT with PVR and PVRL2 inhibits human NK cell cytotoxicity. *Proceedings of the National Academy of Sciences* 2009, **106**(42)**:** 17858-17863.

54. Fuchs, A.; Cella, M.; Giurisato, E.; Shaw, A. S.; Colonna, M. Cutting Edge: CD96 (Tactile) Promotes NK Cell-Target Cell Adhesion by Interacting with the Poliovirus Receptor (CD155). *The Journal of Immunology* 2004, **172**(7)**:** 3994-3998.

55. Wojtowicz, W. M.; Vielmetter, J.; Fernandes, R. A.; Siepe, D. H.; Eastman, C. L.; Chisholm, G. B.*, et al.* A Human IgSF Cell-Surface Interactome Reveals a Complex Network of Protein-Protein Interactions. *Cell* 2020, **182**(4)**:** 1027-1043.e1017.

56. Pende, D.; Bottino, C.; Castriconi, R.; Cantoni, C.; Marcenaro, S.; Rivera, P.*, et al.* PVR (CD155) and Nectin-2 (CD112) as ligands of the human DNAM-1 (CD226) activating receptor: involvement in tumor cell lysis. *Molecular Immunology* 2005, **42**(4)**:** 463-469.

57. Abramson, J.; Adler, J.; Dunger, J.; Evans, R.; Green, T.; Pritzel, A.*, et al.* Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 2024, **630**(8016)**:** 493-500.

58. Lucas, R.; Gómez-Pinto, I.; Aviñó, A.; Reina, J. J.; Eritja, R.; González, C.*, et al.* Highly Polar Carbohydrates Stack onto DNA Duplexes via CH/π Interactions. *Journal of the American Chemical Society* 2011, **133**(6)**:** 1909-1916.

59. Whitlock, J. M.; Leikina, E.; Melikov, K.; De Castro, L. F.; Mattijssen, S.; Maraia, R. J.*, et al.* Cell surface-bound La protein regulates the cell fusion stage of osteoclastogenesis. *Nature Communications* 2023, **14**(1)**:** 616.

60. Spitale, R. C.; Flynn, R. A.; Zhang, Q. C.; Crisalli, P.; Lee, B.; Jung, J.-W.*, et al.* Structural imprints in vivo decode RNA regulatory mechanisms. *Nature* 2015, **519**(7544)**:** 486-490.

61. Castello, A.; Fischer, B.; Frese, Christian K.; Horos, R.; Alleaume, A.-M.; Foehr, S.*, et al.* Comprehensive Identification of RNA-Binding Domains in Human Cells. *Molecular Cell* 2016, **63**(4)**:** 696-710.

62. Cardin, A. D.; Weintraub, H. J. Molecular modeling of protein-glycosaminoglycan interactions. *Arteriosclerosis* 1989, **9**(1)**:** 21-32.

63. Xue, S.; Zhou, F.; Zhao, T.; Zhao, H.; Wang, X.; Chen, L.*, et al.* Phase separation on cell surface facilitates bFGF signal transduction with heparan sulphate. *Nature Communications* 2022, **13**(1)**:** 1112.

64. Klein, K.; Hölzemer, A.; Wang, T.; Kim, T.-E.; Dugan, H. L.; Jost, S.*, et al.* A Genome-Wide CRISPR/Cas9-Based Screen Identifies Heparan Sulfate Proteoglycans as Ligands of Killer-Cell Immunoglobulin-Like Receptors. *Frontiers in Immunology* 2021, **12**.

65. Xu, D.; Young, J. H.; Krahn, J. M.; Song, D.; Corbett, K. D.; Chazin, W. J.*, et al.* Stable RAGE-Heparan Sulfate Complexes Are Essential for Signal Transduction. *ACS Chemical Biology* 2013, **8**(7)**:** 1611-1620.

66. Sirois, C. M.; Jin, T.; Miller, A. L.; Bertheloot, D.; Nakamura, H.; Horvath, G. L.*, et al.* RAGE is a nucleic acid receptor that promotes inflammatory responses to DNA. *Journal of Experimental Medicine* 2013, **210**(11)**:** 2447-2463.

67. Termaat, R.-M.; Brinkman, K.; van Gompel, F.; van den Heuvel, L. P. W. J.; Veerkamp, J. H.; Smeenk, R. J. T.*, et al.* Cross-reactivity of monoclonal anti-DNA antibodies with heparan sulfate is mediated via bound DNA/histone complexes. *Journal of Autoimmunity* 1990, **3**(5)**:** 531-545.

68. Park, H.; Kim, M.; Kim, H.-J.; Lee, Y.; Seo, Y.; Pham, C. D.*, et al.* Heparan sulfate proteoglycans (HSPGs) and chondroitin sulfate proteoglycans (CSPGs) function as endocytic receptors for an internalizing anti-nucleic acid antibody. *Scientific Reports* 2017, **7**(1)**:** 14373.

69. Ren, Z.; Li, R.; Zhou, X.; Chen, Y.; Fang, Y.; Zou, P. Enzyme-Mediated Proximity Labeling Identifies Small RNAs in the Endoplasmic Reticulum Lumen. *Biochemistry* 2023, **62**(12)**:** 1844-1848.

70. Zhang, M.; Liu, L.; Lin, X.; Wang, Y.; Li, Y.; Guo, Q.*, et al.* A Translocation Pathway for Vesicle-Mediated Unconventional Protein Secretion. *Cell* 2020, **181**(3)**:** 637-652.e615.

71. Luteijn Rutger, D.; van Diemen, F.; Blomen Vincent, A.; Boer Ingrid, G. J.; Manikam Sadasivam, S.; van Kuppevelt Toin, H.*, et al.* A Genome-Wide Haploid Genetic Screen Identifies Heparan Sulfate-Associated Genes and the Macropinocytosis Modulator TMED10 as Factors Supporting Vaccinia Virus Infection. *Journal of Virology* 2019, **93**(13)**:** 10.1128/jvi.02160-02118.

72. Gruner, H. N.; McManus, M. T. Examining the evidence for extracellular RNA function in mammals. *Nature Reviews Genetics* 2021, **22**(7)**:** 448-458.

73. Ma, M.; Jiang, W.; Zhou, R. DAMPs and DAMP-sensing receptors in inflammation and diseases. *Immunity* 2024, **57**(4)**:** 752-771.

74. Nofi, C. P.; Wang, P.; Aziz, M. Chromatin-Associated Molecular Patterns (CAMPs) in sepsis. *Cell Death & Disease* 2022, **13**(8)**:** 700.

75. Fischer, S.; Gesierich, S.; Griemert, B.; Schänzer, A.; Acker, T.; Augustin, H. G.*, et al.* Extracellular RNA Liberates Tumor Necrosis Factor-α to Promote Tumor Cell Trafficking and Progression. *Cancer Research* 2013, **73**(16)**:** 5080-5089.

76. Keith, M. P.; Moratz, C.; Tsokos, G. C. Anti-RNP immunity: Implications for tissue injury and the pathogenesis of connective tissue disease. *Autoimmunity Reviews* 2007, **6**(4)**:** 232-236.

77. Parsyan, A.; Shahbazian, D.; Martineau, Y.; Petroulakis, E.; Alain, T.; Larsson, O.*, et al.* The helicase protein DHX29 promotes translation initiation, cell proliferation, and tumorigenesis. *Proceedings of the National Academy of Sciences* 2009, **106**(52)**:** 22217-22222.

78. Tuhkanen, A.-L.; Tammi, M.; Tammi, R. CD44 Substituted with Heparan Sulfate and Endo-β-galactosidase-Sensitive Oligosaccharides: A Major Proteoglycan in Adult Human Epidermis. *Journal of Investigative Dermatology* 1997, **109**(2)**:** 213-218.

79. Evans, C.; Hardin, J.; Stoebel, D. M. Selecting between-sample RNA-Seq normalization methods from the perspective of their assumptions. *Briefings in Bioinformatics* 2018, **19**(5)**:** 776-792.

80. Love, M. I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 2014, **15**(12)**:** 550.

81. Robinson, M. D.; McCarthy, D. J.; Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010, **26**(1)**:** 139-140.

82. Zinn, K.; Özkan, E. Neural immunoglobulin superfamily interaction networks. *Current Opinion in Neurobiology* 2017, **45:** 99-105.

83. Aricescu, A. R.; Jones, E. Y. Immunoglobulin superfamily cell adhesion molecules: zippers and signals. *Current Opinion in Cell Biology* 2007, **19**(5)**:** 543-550.

84. Faas, F. G. A.; Avramut, M. C.; M. van den Berg, B.; Mommaas, A. M.; Koster, A. J.; Ravelli, R. B. G. Virtual nanoscopy: Generation of ultra-large high resolution electron microscopy maps. *Journal of Cell Biology* 2012, **198**(3)**:** 457-469.

85. Rappsilber, J.; Mann, M.; Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nature Protocols* 2007, **2**(8)**:** 1896-1906.

86. Perez-Riverol, Y.; Bai, J.; Bandla, C.; García-Seisdedos, D.; Hewapathirana, S.; Kamatchinathan, S.*, et al.* The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Research* 2022, **50**(D1)**:** D543-D552.

87. Joung, J.; Konermann, S.; Gootenberg, J. S.; Abudayyeh, O. O.; Platt, R. J.; Brigham, M. D.*, et al.* Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nature Protocols* 2017, **12**(4)**:** 828-863.

88. Spahn, P. N.; Bath, T.; Weiss, R. J.; Kim, J.; Esko, J. D.; Lewis, N. E.*, et al.* PinAPL-Py: A comprehensive web-application for the analysis of CRISPR/Cas9 screens. *Scientific Reports* 2017, **7**(1)**:** 15854.

89. Smith, T.; Heger, A.; Sudbery, I. UMI-tools: modeling sequencing errors in Unique Molecular Identifiers to improve quantification accuracy. *Genome Research* 2017, **27**(3)**:** 491-499.

90. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal; Vol 17, No 1: Next Generation Sequencing Data AnalysisDO - 1014806/ej171200* 2011.

91. Bray, N. L.; Pimentel, H.; Melsted, P.; Pachter, L. Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology* 2016, **34**(5)**:** 525-527.