TITLE: Distinct mechanisms of microRNA sorting into cancer cell-derived extracellular vesicle subtypes

Morayma M. Temoche-Diaz[1], Matthew J. Shurtleff[1,2], Ryan M. Nottingham[3], Jun Yao[3], Raj P. Fadadu[2], Alan M. Lambowitz[3], Randy Schekman[2]

[1]Department of Plant and Microbial Biology, University of California, Berkeley, United States; [2]Department of Molecular and Cellular Biology, Howard Hughes Medical Institute, University of California, Berkeley, United States, [3]Institute for Cellular and Molecular Biology and Department of Molecular Biosciences, University of Texas, Austin, United States.

ABSTRACT

Extracellular vesicles (EVs) encompass a variety of vesicles secreted into the extracellular space. EVs have been implicated in promoting tumor metastasis, but the molecular composition of tumor-derived EV sub-types and the mechanisms by which molecules are sorted into EVs remain mostly unknown. We report the separation of two EV sub-populations from a metastatic breast cancer cell line, with biochemical features consistent with different sub-cellular origins. These EV sub-types use different mechanisms of miRNA sorting (selective and non-selective), suggesting that sorting occurs via fundamentally distinct processes, possibly dependent on EV origin. Using biochemical and genetic tools, we identified the Lupus La protein as mediating sorting of some selectively packaged miRNAs. We found that two motifs embedded in miR-122 are responsible for high-affinity

binding to Lupus La and sorting into vesicles formed in a cell-free reaction. Thus, tumor cells can simultaneously deploy multiple EV species using distinct sorting mechanisms that may enable diverse functions in normal and cancer biology.

INTRODUCTION

Extracellular vesicles (EVs) are membranous compartments that consist of a lipid bilayer with a unique set of transmembrane proteins enclosing soluble contents that include nucleic acids and proteins (Colombo, Raposo, and Thery 2014). EVs are found in biofluids as well as in conditioned media of cultured cells (Pisitkun, Shen, and Knepper 2004; Caby et al. 2005; Admyre et al. 2007; Vella et al. 2007). Cells release an array of EV sub-populations, and it has been shown that different EV sub-populations carry different protein (Kowal et al. 2016) and RNA signatures (Tosar et al. 2015). Moreover, distinct EV sub-populations have been implicated in mediating various physiological responses (Willms et al. 2016). Broadly, EV sub-populations can be classified into two categories: exosomes and shedding vesicles, distinguished by the cell membrane of origin. Exosomes are 30-150 nm vesicles that originate in the endocytic pathway. Their secretion to the extracellular space occurs upon multivesicular body (MVB) fusion with the plasma membrane resulting in the release of intraluminal vesicles (ILVs) to the extracellular space (Harding, Heuser, and Stahl 1983; Pan et al. 1985). Shedding vesicles, or microvesicles, refer to a more heterogeneous group of EVs, with sizes ranging from 30 to 1,000nm, which originate by budding directly from the plasma membrane (Cocucci, Racchetti, and Meldolesi 2009). Despite the current knowledge of EV heterogeneity, most current methods used to isolate EVs do not distinguish between sub-populations, nor do they

48 separate EVs from protein and RNA-protein complexes (RNPs) that are not EV
49 encapsulated (Shurtleff, Temoche-Diaz, and Schekman 2018). New methods to facilitate the
50 purification of EV sub-populations are needed in order to evaluate the physiological role
51 and mechanism of macromolecular sorting into each species.
52
53 The lipid membrane of EVs confers protection to its soluble contents against
54 extracellular hydrolases, thus increasing their stability in the extracellular space. The EV
55 soluble content consists of a variety of proteins as well as small RNA molecules (Raposo
56 and Stoorvogel 2013). Among the EV resident transcripts, miRNAs have garnered special
57 attention. MiRNAs are ~22 nucleotide transcripts that modulate gene expression at the
58 post-transcriptional level (Bartel 2004). Although the functional importance of EV-miRNAs
59 as signaling molecules has received some support, including in immunologic response and
60 metastatic tumor cell growth (Mittelbrunn et al. 2011; Montecalvo et al. 2012; Pegtel et al.
61 2010; Fong et al. 2015; Dickman et al. 2017; Zhou et al. 2014; Tominaga et al. 2015; Hsu et
62 al. 2017), the molecular mechanisms and regulation of sorting miRNAs into EVs remain
63 poorly understood. Both non-selective (Tosar et al. 2015) and selective (Shurtleff et al.
64 2016; Santangelo et al. 2016; Villarroya-Beltri et al. 2013) mechanisms of miRNA sorting
65 into EVs have been described. One study using two breast epithelial cell lines found that
66 non-selective miRNA sorting occurs in one EV sub-population, referred to as "large-
67 exosomes", whereas a selective mechanism of sorting occurs in the smaller EV sub-
68 population (Tosar et al. 2015). The cell membrane origin of these two species was not
69 investigated.
70

Cancer cells produce more EVs than their non-transformed counterparts (Szczepanski et al. 2011; Rodriguez et al. 2014). Thus, EVs have been the focus of attention in the context of cancer biology. Several studies have implicated EVs in cancer metastasis (Peinado et al. 2017). EVs have been proposed to increase the metastatic potential of cancer cells and to prepare the pre-metastatic niche prior to tumor cell arrival. Importantly, some studies suggest these effects are due to exosomal-mediated miRNA transfer (Fong et al. 2015; Zhou et al. 2014; Tominaga et al. 2015; Rana, Malinowska, and Zoller 2013; Hsu et al. 2017). However, due to the heterogeneity of EVs and the procedures use to isolate them, it is unclear which vesicle subtypes may mediate these processes.

To gain knowledge of the molecular mechanisms behind miRNA sorting into distinct EV sub-populations, we used as a model the highly invasive breast cancer cell line MDA-MB-231. We used buoyant density gradient centrifugation to resolve two distinct EV sub-populations and found they contain different subsets and levels of enrichment of various miRNAs. With the aid of a cell-free reaction that recapitulates the sorting of miRNAs into vesicles formed *in vitro* (Shurtleff et al, 2016), we found that the Lupus La protein (La protein) is required for the sorting of two different miRNAs found in the high buoyant density EV species, a conclusion that was confirmed in cells repressed for the expression of the La protein. These two La-dependent miRNAs were previously reported to be selectively sorted into EVs to promote cancer metastasis (Fong et al. 2015; Dickman et al. 2017). Further, using the cell-free reaction, we identified two La-binding motifs that are required for the selective sorting of miR-122 *in vitro*. Overall, our use of biochemical fractionation and *in vitro* reconstitution enabled us to identify La as another one of the

RNA-binding proteins required for the selective packaging of some miRNAs into EVs.

RESULTS

Two biochemically distinct extracellular vesicle sub-populations are released by MDA-MB-231 cells

To examine the miRNA content of EVs released by the metastatic breast cancer cell line MDA-MB-231, we first sought to separate EVs from contaminating RNA and protein species that co-sediment during common EV isolation procedures (Shurtleff, Temoche-Diaz, and Schekman 2018) . To achieve this goal, we developed an EV purification strategy involving differential ultracentrifugation followed by buoyant density flotation on a linear iodixanol gradient (Figure 1a and figure 1 – figure supplement 1). This approach allowed the resolution of two distinct EV species, termed the low density (LD) and high density (HD) sub-populations (Figure 1b). Nanoparticle tracking analysis of the two EV sub-populations demonstrated that vesicle sizes largely overlapped, with HD vesicles (114.1 nm mean diameter) being slightly smaller on average than LD vesicles (121.8 nm mean diameter) (Figure 1c). The HD sub-population had densities ranging from 1.13-1.15 g/ml as was reported for exosomes (Colombo, Raposo, and Thery 2014). Furthermore, CD63, a tetraspanin enriched in MVBs and an exosomal marker, was present in the HD fraction (Figure 1b) as were other endosome-associated proteins identified by mass spectrometry (Figure 2 – source data 1). The LD sub-population had a density ranging from 1.08-1.10 g/ml and lacked the CD63 signal (Figure 1b) as well as endosome-associated proteins as judged by mass spectrometry (Figure 2 – source data 1).

Gene ontology analysis for sub-cellular localization of the proteins detected in the HD sub-population by mass spectrometry clustered the terms exosomes and endosomes together (Fig 2a). The LD-detected proteins showed exosomal marker proteins, but endosomal markers were not found at statistically significant levels (Fig 2b). Thus we hypothesize that the HD and LD vesicles have different sub-cellular origins.

The Rab27a protein is involved in the fusion of MVBs to the plasma membrane, regulating in this way exosomal secretion (Ostrowski et al. 2010). We created a Rab27a knock out cell line by using CRISPR-Cas9 technology (Fig 2c). Extracellular vesicles were isolated from conditioned medium of WT and Rab27a knock out cells and analyzed by immunoblot for the levels of two tetraspanins defining the two EV sub-populations: CD63 as a protein marker defining the HD sub-population and CD9 defining the LD sub-population (Figure 1b). EVs derived from the Rab27a KO background showed a dramatic decrease of CD63 signal, while secreting seemingly similar levels of CD9-positive vesicles (Figure 2d) in comparison to the WT levels. Our results are in accordance with a previous study showing that the secretion of CD63-positive, but not of CD9-positive vesicles, is Rab27a dependent (Bobrie et al. 2012).

The Rab35 protein has also being implicated in mediating exosomal secretion in oligodendrocytes (Hsu et al. 2010). Thus, we created a Rab35 knock out cell line (Figure 2c). When EVs derived from the Rab35 KO background were probed for CD63 and CD9 signals, no apparent difference was observed in comparison to EVs derived from WT cells (Figure 2d). Thus, Rab27a but not Rab35, is involved in the secretion of CD63-positive

vesicles derived from MDA-MB-231 cells. Collectively, our data are most consistent with HD vesicles having their origin in the endocytic pathway, whereas LD vesicles may bud directly from the plasma membrane.

We next probed the small RNA content of the separated vesicle fractions. An initial qualitative assessment of RNA profiles using a Bioanalyzer showed that both LD and HD sub-populations share a predominant peak at ~70 nucleotides (Figure 1d), which later proved to correspond to tRNAs (Figure 3 – figure supplement 2d). However, there was an apparent enrichment of transcripts smaller than 70 nucleotides in the HD sub-population (Figure 1d) that we hypothesized represented miRNAs. As such, we considered the possibility that the distinct EV sub-populations might differ in their miRNA content.

Distinct molecular mechanisms of miRNA sorting govern the discrete extracellular vesicle sub-populations.

To further analyze the miRNA content of the HD and LD vesicles, we used two orthogonal approaches: targeted miRNA profiling and an unbiased sequencing approach. Using the commercially available Firefly particle technology, Discovery Panel (Abcam), we compared the abundance of 408 miRNAs in the LD and HD EV sub-populations. Out of the ~400 tested miRNAs, 312 showed a detectable signal in the HD sub-population and 217 were detected in the LD counterpart. We found 21 miRNAs present in the HD vesicles that were enriched at least 10-fold in comparison to their intracellular levels (Figure 3a). In contrast, only 2 miRNAs detected in the LD sub-population showed over a 10-fold change relative to cell lysate (Figure 3b). Both of the apparent LD-enriched miRNAs were also

present as highly enriched in the HD sub-population (Figure 3a). Their enrichment in the LD sub-population was much lower than in the HD sub-population (Figure 3 – figure supplement 2), suggesting their detection in the LD subpopulation may have been the result of small amounts of HD vesicles co-fractionating with the LD vesicles during flotation. Validation by RT-qPCR suggested that these enriched miRNAs were highly specific to the HD sub-population (Figure 3c), confirming our initial suspicion.

As an orthogonal and unbiased approach to profile EV miRNA, we used thermostable group II intron reverse transcriptase sequencing (TGIRT-seq). The TGIRT enzyme is a thermostable, highly processive reverse transcriptase with a proficient template-switching activity used for RNA-seq adapter addition (Mohr et al. 2013). This enzyme's properties allowed us to generate comprehensive cDNA libraries that included highly structured and modified transcripts. Also due to its properties, TGIRT allows the generation of cDNA libraries from low input material (Qin et al. 2016b; Shurtleff et al. 2017). The data obtained from TGIRT-seq was normalized to total number of small non-coding RNA reads per sample. As previously described for EVs using TGIRT-seq (Shurtleff et al. 2017), we found that miRNAs represented a relatively low proportion of total reads (3.4% and 1% of the mapped reads for HD and LD, respectively), resulting in detection of fewer miRNAs by TGIRT-seq than by Firefly profiling. Nevertheless, TGIRT-seq recapitulated the general trends observed by Firefly: multiple miRNAs enriched by at least 10-fold in the HD sample compared to cells (Figure 3 – figure supplement 2a) and fewer enriched miRNAs in the LD sample compared to cells (Figure 3 – figure supplement 2b). Out of the 4 apparent LD enriched miRNAs detected by TGIRT-seq, 3 were also detected as

highly enriched in the HD sub-population, suggesting again that their presence was a result of carryover material overlapping the buoyant density gradient fractions. The total number of uniquely detected HD miRNAs by TGIRT-seq (detectable in the HD sample, but absent in the intracellular lysate) was 21. Comparatively, only 5 miRNAs were uniquely found in the LD sub-population (Figure 3 – figure supplement 2c).

In agreement with our previous findings (Shurtleff et al. 2017), we found that the most abundant small non-coding transcripts in EVs are tRNAs (representing 88.6% of HD and 96.3% of LD transcripts – Figure 3 – figure supplement 2d). We found that most of the tRNA reads in both EV sub-populations start at approximately position 16 (Figure 3- figure supplement 2e), which may represent a reverse transcription stop at the same unidentified EV-enriched D-loop modification described previously (Shurtleff et al. 2017). Thus, both EV sub-populations share a modified tRNA version as their most abundant small non-coding transcript.

We next validated the top EV-enriched/unique miRNAs per vesicle sub-population using RT-qPCR. Relative levels for each EV sub-population were compared to intracellular levels. We found that all of our top HD candidates were validated by this method. As expected, the tested HD miRNAs were highly enriched in the HD EV sample relative to the intracellular levels. The HD/intracellular lysate ratio ranged from ~50-fold to ~1,000-fold (Figure 3c). In contrast, none of our top LD miRNA candidates were validated as being highly enriched in comparison to their intracellular levels. The LD/intracellular lysate ratio ranged from ~0.5-fold to ~5-fold (Figure 3d). We then compared relative levels for each

EV sub-population to each other in order to test for miRNA vesicle specificity. HD-enriched miRNAs were highly depleted in the LD vesicles, as seen by >10-fold increase in their HD/LD ratio (Figure 3c). In contrast, the tested LD miRNAs had roughly the same abundance or were depleted in comparison to their relative levels in HD vesicles (Figure 3d). This proved to be true not only for the few miRNAs tested by RT-qPCR, but was also a general trend for multiple miRNAs detected by Firefly profiling, which demonstrated that many miRNAs were more abundant in the HD vesicles than in their LD counterpart (Figure 3e). All of this together suggested that a selective mechanism of miRNA sorting occurs in the HD sub-population with much less or no miRNA selective sorting in the LD pool.

Highly enriched (>50-fold change HD/cellular lysate) and specific (>50-fold change HD/LD) HD miRNAs were then tested to confirm that they were bona fide EV residents, and not simply EV-associated. The high-speed pellet fraction was exposed to ribonuclease in the absence or presence of non-ionic detergent. All of the highly enriched, specific HD miRNAs tested proved to be EV encapsulated: they were resistant to RNase in the absence but not the presence of detergent (Figure 3f).

These results suggest that a selective miRNA sorting mechanism into EVs occurs in the biogenesis of the HD sub-population but less so or not at all in the LD sub-population. Thus, we identified two distinct EV sub-populations that are released by MDA-MB-231 cells that utilize fundamentally different miRNA sorting mechanisms (selective vs non-selective).

232 **Selective miR-122 sorting into vesicles is recapitulated *in vitro***

233 We selected miR-122 for further biochemical analysis as it has been shown that
234 circulating miR-122 can serve as a prognostic biomarker for metastasis in patients with
235 breast cancer (Wu et al. 2012) and very specifically that MDA-MB-231 cell-derived
236 exosomal miR-122 may promote metastasis by reprogramming glucose metabolism in the
237 premetastatic niche (Fong et al. 2015). Because miR-122 is present in the fetal bovine
238 serum (FBS) used to culture cells (Wei et al. 2016), we employed clarified FBS depleted of
239 sedimentable particles to confirm that the miR-122 was released by MDA-MB-231 cells. We
240 observed an accumulation of miR-122 over time in conditioned media exposed to MDA-
241 MB-231 cells, whereas the level of miR-122 slightly decreased over time when conditioned
242 media was incubated in the absence of cells (Figure 4 – figure supplement 1). This
243 suggested that newly synthesized MDA-MB-231 cell-derived miR-122 was secreted into the
244 extracellular space.
245
246 To explore the mechanism(s) by which miR-122 is selectively sorted into the MDA-
247 MB-231 HD EV sub-population, we first sought to recapitulate miR-122 packaging *in vitro*.
248 An assay for the cell-free packaging of a miRNA into putative exosomal vesicles was
249 previously developed in our laboratory (Shurtleff et al. 2016). The reaction includes
250 sedimented membranes and clarified cytosol, obtained from mechanically ruptured
251 HEK293T cells, along with synthetic miRNA and ATP. After an incubation of 20 min at
252 30°C, RNA incorporated into vesicles was monitored by protection against degradation by
253 exogenous RNase. Protected RNA was then extracted and quantified by RT-qPCR. The
254 outcome of the reaction is represented as the percentage of initial input miRNA that has

255 become protected over the incubation period (Figure 4a). Using a modified version of our
256 previously published reaction, we found that maximum protection of miR-122 in lysates of
257 MDA-MB-231 cells required membranes, cytosol and incubation at a physiological
258 temperature (Figure 4b), as was previously observed for miR-223 in lysates of HEK293T
259 cells (Shurtleff et al. 2016). Therefore, miR-122 packaging into vesicles in lysates of MDA-
260 MB-231 cells may be recapitulated *in vitro.*
261
262 **Identifying *in vitro* miR-122 RNA-binding protein partners**
263 Having detected miR-122 packaging *in vitro*, we next sought to utilize this approach to
264 study the mechanism of miR-122 sorting into MDA-MB-231 EVs. We first tested if miR-122
265 packaging was YBX1-dependent. YBX1 is a RNA-binding protein (RBP) previously found to
266 be required for miR-223 packaging into vesicles *in vitro* and into EVs secreted by HEK293T
267 cells (Shurtleff et al. 2016). Using membranes and cytosol from WT and *ybx1* null HEK293T
268 cells, we repeated the demonstration that miR-223 packaging was YBX1-dependent, but
269 found that miR-122 was packaged nearly normally in lysates devoid of YBX1 protein
270 (Figure 4c).
271
272 Several RBPs have been implicated in miRNA sorting into EVs from different cell types
273 (Shurtleff et al. 2016; Mukherjee et al. 2016; Santangelo et al. 2016; Villarroya-Beltri et al.
274 2013). As such, in order to study the RBP(s) that might mediate miR-122 packaging in
275 MDA-MB-231 cells, we performed an *in vitro* packaging reaction employing a 3'-
276 biotinylated form of miR-122 to allow the capture of the miRNA and any bound proteins.
277 Briefly, following *in vitro* packaging, reactions were treated with RNase, the RNase activity

was quenched and the membranes solubilized with Triton X-100. Once the luminal content was released, biotinylated miR-122, along with its protein interactors, was captured with streptavidin beads. Proteins were eluted with Laemmli buffer, extracted from a SDS-polyacrylamide gel and the eluted fraction used for mass spectrometry. The proteins detected by mass spectrometry were curated for RBPs, except that any ribosomal proteins were excluded (Figure 4d). We decided to focus on the top three candidates, nucleolin (NCL), Lupus La (La) and nucleophosmin (NPM1). These three RBPs have been reported to be present in crude high-speed pellet preparations from conditioned media from different carcinoma cell lines (Liang et al. 2013; Demory Beckler et al. 2013), including from our model cell line MDA-MD-231 (Skottvoll et al. 2018).

Notably, Ago2 was not detected bound to miR-122 in our mass spectrometry results. This was not simply an artifact of our *in vitro* system, as Ago2 was also undetectable in immunoblots of the buoyant density fractionated EV membranes (Figure 4 – figure supplement 2). Both Ago2 and Dicer were present in the high-speed pellet, but not as buoyant species, suggesting they are associated with co-purifying RNP complexes that are not vesicle-associated. This finding is in accordance with other published data (Shurtleff et al. 2016; Van Deun et al. 2014) where Ago2 is detected in the high-speed pellet but absent in the vesicle sample after more stringent purification methods are used.

To test the relevance of the three RBPs in miR-122 packaging into EVs, we used CRISRPR interference (CRISPRi) (Gilbert et al. 2013; Horlbeck et al. 2016) to systematically knock down each protein in MDA-MB-231 cells. CRISPRi promotes gene silencing by

repressing transcription of the target gene. Importantly, unlike siRNA or shRNA, CRISPRi silences genes independently of the RISC machinery. Because the RISC machinery binds to miRNAs and is responsible for miRNA-mediated gene silencing (Bartel 2004), we avoided any artificial overload of the RISC machinery that might result in unpredictable effects on the miRNA sorting in our system.

Using CRISPRi, we prepared cytosols from MDA-MB-231 cells depleted of nucleophosmin and La that were then used in the *in vitro* packaging reactions. Both nucleophosmin and La were efficiently knocked down using this system (Figure 4 – figure supplement 3a). However, nucleolin knock-down resulted in apparent cellular arrest with subsequent cell death. Thus, nucleolin functional analysis was not possible by knock-down. Nucleophosmin-depleted cytosol produced a mild packaging phenotype only at reduced cytosol concentrations (3 mg/ml) (Figure 4 – figure supplement 3b). In contrast, La-depleted cytosol showed severely reduced miR-122 packaging efficiency at cytosolic concentrations of 3 and 6 mg/ml in the cell-free reactions (Figure 4 – figure supplement 3b and Figure 5a). Thus, these experiments suggested a role for La in the packaging of miR-122 into vesicles formed *in vitro.*

Packaging of miR-122 *in vitro* is La-dependent

La is an abundant RNA-binding protein that shuttles between the nucleus and cytoplasm. Within the nucleus, La helps stabilize RNA polymerase III (PolIII) transcripts by binding to poly-uridine tails in their 3' termini (Stefano 1984; Rinke and Steitz 1985; Wolin and Cedervall 2002). In the cytoplasm, La has been suggested to bind and regulate the

translation of 5' terminal oligopyrimidine (TOP) mRNAs (Intine et al. 2003; Cardinali et al. 2003; Crosio et al. 2000). Moreover, in the context of cancer, La has been shown to shuttle from the nucleus to the cytoplasm at higher efficiencies in invasive cells undergoing epithelial to mesenchymal transition in order to control the translation of mRNAs bearing internal ribosome entry sites (IRES) (Petz, Them, Huber, Beug, et al. 2012; Petz, Them, Huber, and Mikulits 2012). La has also been indirectly implicated in the exosomal secretion of a small Epstein-Barr viral RNA, EBV-EBER1. EBV-EBER1 interaction with La in the cytoplasm masks the viral RNA from recognition by cytoplasmic surveillance machinery and facilitates its later EV-mediated secretion (Baglio et al. 2016).

In order to establish whether La itself, rather than a La-associated protein, is necessary for miR-122 incorporation into EVs *in vitro* (Figure 5a), recombinant His-tagged human La was expressed and purified from insect Sf9 cells (Figure 5 – figure supplement 1). La-depleted cytosol was reduced ~4-fold in the packaging of miR-122 but was restored to normal on addition of purified, recombinant La to the combination of membranes and La-depleted cytosol (Figure 5b). However, addition of La alone with or without membranes and no cytosol resulted in only background protection of miR-122 (Figure 5b), suggesting that additional cytosolic proteins are necessary for *in vitro* vesicle biogenesis and/or miRNA sorting process. The level of exogenously added La protein in the *in vitro* reaction was adjusted to approximate the level found in an aliquot of the normal cytosol (Figure 5c).

We next examined the requirements for co-packaging of La and miR-122 in the cell-free reaction. An immunoblot was used to test the relative amount of La co-isolated with

biotinylated miR-122 in conditions that sustain or fail to result in miR-122 packaging. We observed maximal La recovery in incubations conducted with cytosol and membranes at 30°C but much less so at 4°C (Figure 5d). Thus, the sorting of La bound to miR-122 requires the presence of both, membranes and cytosol, and a physiological temperature of incubation.

MiR-122 packaging into EVs *in vivo* is La-dependent

Our *in vitro* data suggested that La was at least one of the RBPs responsible for miR-122 packaging. However, in order to confirm this data, we first examined the effects of depletion of the three RBP candidates *in vivo*. Previous work has demonstrated that selected miRNAs accumulate in the cytoplasm in cells deficient in RBPs required for RNA sorting into EVs (Shurtleff et al. 2016; Santangelo et al. 2016; Villarroya-Beltri et al. 2013). We assessed the level of intracellular miR-122 and of another highly enriched miRNA (miR-142) upon CRISPRi depletion of each RBP candidate. Greater accumulation of miR-122 was seen upon La depletion (2.7-fold change La CRISPRi/WT) compared to nucleophosmin depletion (1.7-fold change NPM1 CRISPRi/WT) (Figure 6a). CRISPRi of nucleolin resulted in less cellular miR-122, which could be an indirect result of the more toxic effects of this depletion. As a control for selectivity of miRNA sorting and export, we assed the intracellular levels of non-selectively sorted miRNAs miR-574 and miR-320a (Figure 3d). In contrast to the selectively sorted miRNAs, we did not observe intracellular accumulation of miR-574 and miR-320a under any of the knock-down conditions (Figure 6a). At least in relation to La and nucleophosmin, the depletion results in cells confirmed our previous observations on the role of these proteins in our cell-free reaction.

We next examined the relative content of miR-122 and miR-142 and of a non-selectively sorted miRNA (miR-320a – Figure 3d) in EVs in relation to their accumulation in cells depleted of La. WT and La-depleted cells had approximately the same number of purified EVs as quantified using a Nanosight particle tracking instrument. A significant reduction in secretion of miR-122 and miR-142 with no change in miR-320a was accompanied by intracellular accumulation of miR-122 and miR-142 in the La-depleted background (Figure 6b). Thus, La plays a role in the selective sorting of at least two selectively packaged EV miRNAs without affecting EV biogenesis.

To test for the presence of La in EVs isolated from cultured MDA-MB-231cells, we first purified vesicles by buoyant density flotation. The high-speed pellet was subjected to flotation in a step sucrose gradient. Buoyant membranes were permeabilized with non-ionic detergent and soluble proteins were precipitated/concentrated with acetone in order to facilitate detection by immunoblot (Figure 6c). We successfully detected the presence of La in the buoyant sample. Its detection was improved upon acetone precipitation/concentration (Figure 6d).

In order to confirm that La was a soluble protein residing inside EVs, and not simply vesicle-associated, we performed proteinase K protection assays on crude high-speed pellet material. The high-speed pellet fraction from conditioned medium was exposed to proteinase K in the presence or absence of detergent and the exposure of La was assessed by immunoblot. The La protein was resistant to proteinase K digestion and rendered

393 sensitive upon membrane permeabilization with non-ionic detergent (Figure 6c). Flotillin-
394 2, a membrane protein associated with the inner leaflet of the EV, served as a positive
395 control and was degraded only in the presence of detergent. In contrast, Dicer, previously
396 shown to be present in the high-speed pellet but not as an EV-encapsulated RNP (Figure 4 –
397 figure supplement 2), was susceptible to degradation independent of the addition of non-
398 ionic detergent (Figure 6c). We conclude that the La protein is packaged into vesicles in our
399 cell-free reaction as well as into EVs secreted by cells and is required for packaging of miR-
400 122 into vesicles *in vitro* and *in vivo*.
401
402 **La directly interacts with miR-122 *in vivo* and *in vitro***
403 Given our results on the role of La in miR-122 sorting in cells and in the cell-free
404 reaction, we next examined the intracellular distribution of La and the selectivity of its
405 interaction with miR-122 versus a control miRNA that was detected, but not enriched, in
406 LD and HD vesicles (miR-182; Figure 3c). At steady-state, La was reported to be largely
407 confined to the nucleus (Wolin and Cedervall 2002). However, other evidence supports a
408 role for La function in the cytoplasm and its shuttling between the nucleus and cytoplasm
409 (Cardinali et al. 2003; Intine et al. 2003; Petz, Them, Huber, Beug, et al. 2012; Petz, Them,
410 Huber, and Mikulits 2012). Our observation of La as an EV resident protein (Figure 6d and
411 6e) is consistent with at least some residence time in the cytoplasm.
412
413 Endogenous La was visualized by standard and structured illumination fluorescence
414 microscopy (SIM) in fixed MDA-MB-231 cells. Immunofluorescence confirmed the presence
415 of nuclear and cytoplasmic La (Figure 7 – figure supplement 1a). The specificity of the

antibody was affirmed by a significant depletion of the immunofluorescence signal in cells depleted of La by CRISPRi (Figure 7 – figure supplement 1a and 1b). At a higher level of resolution, SIM revealed cytoplasmic puncta of La, some of which co-localized with a late endosomal marker, Rab7 (Figure 7 – figure supplement 1c).

We evaluated the interaction of La with two miRNAs by co-immunoprecipitation from mechanically ruptured MDA-MB-231 cells. A post-nuclear supernatant fraction of cells was mixed with La antiserum and the IP sample was split for protein and RNA analysis by immunoblot and RT-qPCR, respectively (Figure 7a). Approximately 15% of the La protein and miR-122 was precipitated (Figure 7b and 7c). MiR-182, which is present but not enriched in LD and HD EVs, was used as a negative control and did not co-immunoprecipitate with La (Figure 7c).

In order to explore the interaction of La and miR-122 directly, we used purified, recombinant La incubated with miR-122 and evaluated the binding affinity by electrophoretic mobility shift assays (EMSAs). Purified La was titrated, mixed with 5' fluorescently-labeled miR-122 and incubated at 30°C. After incubation, complexes were separated by electrophoresis and detected by in-gel fluorescence (Figure 7d). As negative and positive controls for La interaction we used 22 nt RNAs comprised of alternating purines or poly-uridine, respectively (Figure 7 – figure supplement 2a and 2b). The measured K_d for La:miR-122 was 4.8 nM (Figure 7e). Notably, this measurement indicates that the affinity of miR-122 for La is greater than the affinity previously reported for Ago2

438 and miRNAs, with K_ds from 10-80 nM (Tan et al. 2009). We therefore conclude that the
439 La:miR-122 complex displays high affinity and specific interaction *in vitro.*
440
441 **Finding the miR-122 motif responsible for its packaging**
442 Lastly, we explored the miR-122 motif responsible for its association with La and
443 packaging into EVs in the cell-free reaction. La has been shown to bind to UUU sequences at
444 the 3' end of PolIII transcripts (Stefano 1984) and miR-122 contains a UUUG at that
445 terminus (Figure 8a). We designed a variant of miR-122 where A residues replaced the 3'U
446 bases and found that this specie is poorly packaged into vesicles in the cell-free reaction
447 (3% efficient incorporation vs. 15% for WT) and displays a greatly reduced (~100-fold)
448 affinity for La in the EMSA analysis (WT K_d 4.8 nM vs 3'mut K_d 538 nM) (Figure 8a, b and c
449 and figure supplement 1a). We conclude that the UUU sequence at the 3' end of miR-122 is
450 necessary for a high affinity interaction with La, which may correlate to the requirement
451 for these residues in the packaging of miR-122 into vesicles in the cell-free reaction.
452
453 To identify possible motifs beyond the literature-based UUU candidate, we also
454 performed a multiple expectation maximization (EM) for motif elicitation (MEME) analysis
455 for 49 HD EV miRNAs. We considered miRNAs that were at least 10-fold enriched, as
456 detected by profiling and sequencing, as well as those unique miRNAs in the HD sub-
457 population that were detected by sequencing. This approach identified "UGGA" as an EV-
458 enriched miRNA motif. This motif was present in 13 (26%) of the 49 analyzed miRNAs,
459 including the 5' sequence of miR-122 (Fig. 8a), and it was by far the most overrepresented
460 motif detected by this analysis. The UGGA sequence appeared to be EV specific, as it did not

emerge when similar analyses were performed with exclusively cellular miRNAs (miRNAs detected in cellular lysates but absent in the EV samples). We then manually assessed the pool of exclusively cellular miRNAs for this motif where it appeared in 5 out of 97 analyzed miRNAs (5%). A variant of this 5' sequence motif impaired miR-122 incorporation in the *in vitro* packaging reaction (Figure 8b) and decreased the affinity ~10-fold for La compared to WT (Figure 8c and figure supplement 1b). Thus, we conclude that a bipartite, or possibly even more complex motif directs recognition of miR-122 by La and consequently results in its selective sorting into EVs and secretion from MDA-MB-231 cells.

DISCUSSION

Different molecular mechanisms of miRNA sorting

Here we document the separation of two EV populations, likely of different sub-cellular origin, which appear to employ different molecular mechanisms of miRNA sorting (Figure 9). We found that EVs with a density ranging from 1.08-1.10 g/ml (LD) may originate in the plasma membrane, whereas the EVs with a density ranging from 1.13 to 1.15 g/ml (HD) are more characteristic of exosomes, originating in the endocytic pathway. Some miRNAs are non-selectively sorted into both EV sub-populations, as shown by their similar relative abundances in the cellular lysate and both EV sub-populations (Figure 3d). We observed a slight enrichment for non-selectively sorted miRNAs in the HD sub-population (Figure 3a) and a minor depletion of non-selectively sorted miRNAs in the LD sub-population (Figure 3b).

Non-selective sorting may correlate to the fraction of miRNAs unbound to the RISC complex. Association of miRNAs with the RISC complex inside cells varies greatly: >1,800-fold for different miRNAs present in the same cell line (Flores et al. 2014). As a result, some miRNAs may exist in the cytoplasm as mostly RISC-free species. Our data show that Ago2 and Dicer are not present in EVs derived from MDA-MB-231 cells (figure 4 - figure supplement 2). Thus, we suggest that the non-selectively sorted miRNAs may represent the cytosolic pool of RISC-free miRNAs, not actively engaged in silencing, and their presence in EVs is the result of a passive mechanism of sorting.

P-bodies have been shown to be closely associated with MVBs (Gibbings et al. 2009) and the levels of target mRNAs can influence the levels of miRNA association with the RISC complex (Flores et al. 2014), as well as miRNA secretion as detected in the high-speed pellet fraction sedimented from conditioned media (Squadrito et al. 2014). Thus, miRNAs released from P bodies and freed from the RISC complex may be proximal to endosomes engaged in the production of HD vesicles and distant from the site of LD vesicle biogenesis, thus resulting in a slight enrichment and slight depletion of these miRNAs in the HD and LD vesicles, respectively.

We focused the bulk of our investigation on a set of miRNAs selectively packaged in the HD sub-population. These selectively sorted miRNAs may belong to the pool of pre-existing RISC-free miRNAs, or may be displaced from the RISC complex prior to or concomitant with sorting into HD vesicles. Importantly, we found that this mechanism of selective miRNA sorting is specific for the HD sub-population. We speculate that the sorting

machinery required to enrich certain miRNAs in exosomes is specific for MVBs, and it is absent at the site of origin of the LD vesicles. However, enveloped viruses that bud from the cell surface clearly do sort their RNA genomes in a selective manner (D'Souza and Summers 2005; Swanson and Malim 2006) and the same may be true of other RNA biotypes in the LD sub-population, especially in other cell lines. We previously showed that tRNAs with an unknown post-transcriptional modification in the D-loop are highly enriched in EVs in comparison to the cells from the same culture (Shurtleff et al. 2017). Similarly, we detected apparently similarly modified tRNA species in both EV sub-populations secreted from MDA-MB-231 cells (Figure 3-supplemental 2e). The machinery for the active sorting of the modified tRNA may be shared for both EV sub-populations.

It is widely suggested that EV miRNAs play a role in intercellular communication (Fong et al. 2015; Zhou et al. 2014; Tominaga et al. 2015; Rana, Malinowska, and Zoller 2013; Hsu et al. 2017). Thus, purifying EV sub-populations, with differential miRNA content, is critical in order to accurately assess their role in intercellular communication. One study purified different EVs based on their buoyant densities and showed that the LD sub-population triggered the differential expression of 257 genes in recipient cells, whereas the HD sub-population caused the differential expression of 1,116 genes (Willms et al. 2016). Although this study did not determine if any gene expression differences were the result of miRNA transfer, the effect could be partially explained by the distinct miRNA composition of the two EV sub-populations. Further studies using purified EV sub-populations are needed to understand the physiological consequence of selective vs. non-

528 selective sorting mechanisms. Additionally, it needs to be determined if the effects are
529 directly attributable to miRNAs or some other constituents of the two EV populations.
530
531 **The role of La in selective sorting of miRNAs**
532 Here we show *in vitro* and *in vivo* evidence for the requirement of La for packaging
533 of miR-122 into HD EVs derived from MDA-MB-231 cells. The La protein is an abundant
534 RBP that is upregulated in some cancers (Sommer, Rossa, et al. 2011; Trotta et al. 2003; Al-
535 Ejeh, Darby, and Brown 2007; Sommer, Dittmann, et al. 2011). Its presence in the
536 cytoplasm has been linked to the translational regulation of mRNAs bearing IRES (Petz,
537 Them, Huber, Beug, et al. 2012; Petz, Them, Huber, and Mikulits 2012), as well as 5'TOP
538 mRNAs (Intine et al. 2003; Cardinali et al. 2003; Crosio et al. 2000). Interestingly, it was
539 previously reported that 5'TOP mRNAs are enriched in EVs derived from HEK293T cells in
540 comparison to their relative cellular levels (Shurtleff et al. 2017), although we have not yet
541 evaluated the abundance of 5'TOP mRNAs in EVs isolated from MDA-MB-231 cells.
542
543 Little is known about how cytoplasmic RNA binding proteins, such as La, along with
544 bound RNA molecules are targeted and recognized for secretion. The phosphorylation
545 status of La has been suggested to be important for its intracellular localization (Intine et
546 al. 2003). In addition, sumoylation of La has been shown to be important for La-dependent
547 mRNA transport in neurons (van Niekerk et al. 2007) and sumoylation of another RNA-
548 binding protein in EVs (hnRNPA2B1) has been shown to be important for export
549 (Villarroya-Beltri et al. 2013). Post-translational modifications such as phosphorylation,

ubiquitylation and sumoylation may serve to direct RNPs to the ESCRT machinery and thus into membrane buds that are internalized into MVBs.

The La protein binds to and stabilizes poly-uridine stretches of PolIII transcripts that mature in the nucleus (Wolin and Cedervall 2002). Likewise in the cytoplasm, La may serve a similar function in the stabilization of polyuridine tracks in 5'TOP mRNAs (Intine et al. 2003). Here we show that La directly binds to, and promotes the vesicle packaging of miR-122 in a cell-free reaction. We obtained evidence that a 3' poly-uridine sequence of miR-122 is required for this interaction. This role of La may extend to miRNAs secreted by other cell lines. B cell-derived EVs are enriched for 3' uridylylated miRNA isoforms in comparison to their intracellular counterparts (Koppers-Lalic et al. 2014), and La has been detected by mass spectrometry in the high-speed pellet fraction from B cell conditioned medium (Meckes et al. 2013).

In addition to the known poly (U) binding site, sequence motif analysis of the most highly enriched miRNAs in the HD vesicles isolated from MDA-MB-231 conditioned medium identified a second, previously unreported motif required for La:miR-122 high affinity interaction and sorting into EVs in our cell-free reaction. A 5' UGGA sequence resembles the motif shown to be required for miRNA sorting into hepatocyte EVs, dependent on the SYNCRIP protein (Santangelo et al. 2016). However, we did not detect SYNCRIP in our mass spectrometry analysis for RBP partners responsible for miR-122 sorting *in vitro* (Figure 4d). Thus, it is possible hepatocytes and breast cancer epithelial cells differ in their mechanisms of miRNA sorting. The presence of both motifs, UUU and

573 UGGA, in miR-122 may allow for an exceptionally high affinity interaction with La. The
574 measured K_d for the La:miR-122 complex is 4.8 nM, indicating tighter binding than that
575 reported for Ago2 to miRNAs using a similar assay (Tan et al. 2009). This high affinity
576 interaction may explain how selected miRNAs are sequestered into exosomes away from
577 the RISC complex.
578
579 As we have shown with La, other RBPs (YBX1 (Shurtleff et al. 2016), hnRNPa2b1
580 (Villarroya-Beltri et al. 2013) and SYNCRIP (Santangelo et al. 2016)) are involved in the
581 selective sorting and secretion of miRNAs in different cell lines. Conceivably, this process
582 could be organized and controlled in a cell-specific manner by the assembly of RNA
583 granules including other more abundant RNAs such as tRNA, Y-RNA and vault RNA
584 (Shurtleff et al., 2017). In this respect, we find that most of the La protein in the cytoplasm
585 is localized in punctate structures, some of which closely associate with the late endosomal
586 marker Rab7 (Figure 7 – figure supplement 1a and 1c). It will be instructive to examine
587 the cytoplasmic co-localization of La with the other RBPs implicated in miRNA secretion in
588 EVs.
589
590 **Significance of selective miRNA sorting into EVs in cancer**
591 Intercellular communication mediated by miRNAs is a subject of considerable
592 interest, but in most cases, it remains unproven. Many studies using cancer cells suggest
593 EV-mediated miRNA transfer to recipient cells (Fong et al. 2015; Zhou et al. 2014;
594 Tominaga et al. 2015; Rana, Malinowska, and Zoller 2013; Hsu et al. 2017), possibly to
595 promote metastasis. In the context of breast cancer for example, EV resident miR-122 has

been suggested to promote metastasis through reprogramming glucose metabolism in the resident cells of the premetastatic niche (Fong et al. 2015). In another example, EV-associated miR-105 may promote invasiveness in metastatic breast cancer by disruption of tight junctions, thus promoting vascular permeability and tumor cell invasion (Zhou et al. 2014). For EV miRNAs to be functionally transferred to recipient cells, fusion of the EV membrane with a recipient membrane (either plasma membrane or endosomal membrane) must precede the presentation of an exogenous miRNA to the target cell cytoplasm. We find that the highly-selected miRNAs in exosomes are not bound to Ago2. Thus, the usual function of a miRNA would require loading from the EV interior to a cytoplasmic RISC complex. Unfortunately, nothing is known about the cytoplasmic presentation of miRNAs derived from EVs.

Alternatively, tumor cell invasiveness may depend on the selective secretion and disposal of tumor suppressor miRNAs. Such a suggestion has been made for miR-23b (Ostenfeld et al. 2014) and miR-100 (Cha et al. 2015), which are secreted at high efficiency by cancer cells. These two mechanisms, intercellular communication and disposal, are not mutually exclusive. It is possible that EVs containing detrimental miRNAs have been repurposed to serve as shuttles for intercellular communication, as has been suggested for the secretion of miR-142 (Dickman et al. 2017). Of relevance to the work we report here, miR-122 has been suggested as a tumor suppressor with cells developing greater invasiveness on loss of this miRNA (Coulouarn et al. 2009; Wang, Wang, and Yang 2012). Nonetheless, the importance of EV resident miR-122 as a molecule to convey a message to promote metastasis has also been suggested (Fong et al. 2015).

A bigger challenge in the field of cancer and EV biology is that claims for a role of EV-associated miRNAs are based on crude preparations of sedimentable particles or co-culture experiments. Here we show that Ago2 is secreted to the extracellular space as a RNP, not vesicle-associated, that co-isolates with EVs in the high-speed pellet. Thus, clear demonstration that the implicated miRNAs in cancer development are enclosed within a membrane vesicle is still missing. Studies purifying, characterizing, understanding the cellular biology behind different EV sub-populations and their functional effects in the context of cancer will prove important in the future.

MATERIALS AND METHODS

Cell lines, media and general chemicals

MDA-MB-231 cells were cultured in DMEM with 10% FBS (Thermo Fisher Scientific, Waltham, MA). MDA-MB-231 cells were confirmed by short tandem repeat profiling (STR) and tested negative for mycoplasma contamination. For EV production, cells were seeded at ~10% confluency in exosome-depleted medium (System Biosciences, Palo Alto, CA) in 150mm CellBIND tissue culture dishes (Corning, Corning, NY) containing 30ml of growth medium. EVs were collected when cells reached ~80% confluency (~72h). Unless otherwise noted, all chemicals were purchased from Sigma Aldrich (St. Louis, MO).

Extracellular vesicle purification

Conditioned medium (720 ml for TGIRT-seq and miRNA profiling and 420 ml for all other

642 experiments) was harvested from 80% confluent MDA-MB-231 cultured cells. All
643 subsequent manipulations were performed at 4°C. Cells and large debris were removed by
644 centrifugation in a Sorvall R6+ centrifuge (Thermo Fisher Scientific) at 1,000xg for 15 min
645 followed by 10,000xg for 15 min in 500 ml vessels using a fixed angle FIBERlite F14-
646 6X500y rotor (Thermo Fisher Scientific). The supernatant fraction was then centrifuged at
647 ~100,000xg (28,000 RPM) onto a 60% sucrose cushion in buffer C (10mM HEPES pH 7,4,
648 0.85% w/v NaCl) for 1.5 h using two SW-28 rotors. The interface over the sucrose cushion
649 was collected and pooled for an additional ~130,000xg (32,500 RPM) centrifugation over a
650 60% sucrose cushion in a SW41 rotor for 15 h. The collected interface from the first
651 sucrose cushion should not exceed a sucrose concentration of 21%, as measured by
652 refractometry, for the second centrifugation in the SW41 to be successful. Higher
653 concentrations of sucrose cause the EVs to equilibrate at the ambient buoyant density,
654 impeding sedimentation. For purification of EV sub-populations based on their distinct
655 buoyant density, the cushion-sedimented vesicles were collected and mixed with 60%
656 sucrose to a final volume of 4ml (sucrose final concentration ~40%). Layers of 1.5 ml of
657 25%, 20%, 15%, 10% and 5% iodixanol (Optiprep) solutions in buffer C were sequentially
658 overlaid and samples were centrifuged at ~160,000xg (36,000 RPM) for 15 h in a SW41
659 rotor. Fractions (400ul) from top to bottom were then collected and mixed with Laemmli
660 buffer for immunoblot analysis or RNA was extracted using a mirVana miRNA isolation kit
661 (Thermo Fischer Scientific). This method produced a linear density gradient from 1.036 to
662 1.24 g/ml (figure 1 – figure supplement 1 – source data 1). For detection of Ago2 and Dicer
663 (Figure 4 - figure supplement 2), vesicles were not sedimented over a 60% sucrose
664 cushion. Sedimented EVs from the SW28 centrifugation were resuspended in 800ul of 30%

665 iodixanol and layers of 800ul of 25%, 20%, 15%, 10% and 5% iodixanol were layered on
666 top and centrifuged in a SW55 rotor at ~140,000xg (38,000RPM) for 15 h. Fractions
667 (200ul) were collected from top to bottom and mixed with Laemmli buffer for
668 immunoblots analysis.
669 For EV purification in bulk (without discriminating among EV sub-populations), the EVs
670 collected after the first SW41 centrifugation were mixed with 60% sucrose to a final
671 volume of 9 ml (sucrose concentration ~50%). Layers of 1.5 ml of 40% and 10% sucrose
672 were overlaid and the sample was centrifuged at ~160,000xg (36,000 RPM) for 15 h in a
673 SW41 rotor. Buoyant vesicles spanning from the 40% sucrose region and the 10/40%
674 interface were collected and mixed with Laemmli buffer or RNA was extracted with a
675 mirVana miRNA isolation kit. For further processing to improve protein detection, buoyant
676 vesicles were first diluted in buffer C (10mM HEPES pH 7,4, 0.85% w/v NaCl) to a final
677 volume of 5 ml and soluble content was released by adding 10% Triton X-100 (TX-100) to a
678 final concentration of 1%. Samples were then homogenized by the use of a vortex mixer for
679 1 min, cooled on ice for 10 min and homogenized again for an extra min. To precipitate
680 proteins, we added 4 volumes of cold acetone (previously stored at -20°C) and the mixture
681 was incubated at -20°C overnight. Precipitated proteins were collected using a Sorvall RC
682 6+ centrifuge by centrifugation at 16,000xg (11,600 RPM) for 30 min at 4°C in a FIBERlite
683 F211-8X50Y rotor (Thermo Fischer Scientific). The pellet fraction was then resuspended in
684 Laemmli buffer for analysis by immunoblotting.
685 For proteinase K protection assays, the EVs were collected by centrifugation at ~140,000xg
686 (28,000 RPM) for 1.5 h using two SW-28 rotors. Pellet fractions resuspended in buffer C
687 were pooled and centrifuged at ~160,000xg (36,000 RPM) in a SW55 rotor for 1 h. The

688 pellet was resuspended in buffer C and split into 3 equal aliquots. One sample was left
689 untreated, another sample treated with 5ug/ml proteinase K on ice for 20 min, and the last
690 was mixed with TX-100 to a final concentration of 1% prior to proteinase K treatment.
691 Proteinase K was inactivated with 5mM phenylmethane sulfonyl fluoride (PMSF) (5 min on
692 ice) and samples were then mixed with Laemmli buffer for immunoblot analysis.
693
694 Mass spectrometry analysis of EV sub-populations
695 EV sub-populations purified from the iodixanol gradient (as described above) were diluted
696 with buffer C to a final volume of 30ml. The diluted samples were then centrifuged at
697 ~100,000xg (28,000RPM) in a SW28 rotor for 1.5 h. Pellet samples were resuspended in
698 Laemmli buffer without bromophenol blue and electrophoresed in a 4-20% acrylamide
699 Tris-Glycine gradient gel (Life Technologies) for ~3 min. The bulk of proteins were stained
700 with Coomassie and the stained band was excised from the gel with a new razor blade.
701 Samples were submitted to the Vincent J. Coates Proteomics/Mass Spectroscopy laboratory
702 at UC Berkeley for in-gel tryptic digestion of proteins followed by liquid chromatography
703 and mass spectrometry analysis according to their standards.
704 For sub-cellular localization analysis of the detected proteins, the list of proteins detected
705 by mass spectrometry was analyzed first using the GoTermFinder developed at the Lewis-
706 Sigler Institute at Princeton (Boyle et al. 2004) followed by REVIGO analysis (Supek et al.
707 2011) to obtain the scatterplots of sub-cellular localization along with their organelle
708 association.
709
710 Nanoparticle tracking analysis

711 Extracellular vesicles purified on linear iodixanol gradients were diluted 1:100 with PBS
712 filtered with a 0.02um filter (Whatman GmbH, Dassel, Germany). The mixture was drawn
713 into a 1 ml syringe and inserted into a Nanosight LM10 instrument (Malvern, UK). Particles
714 were tracked for 60 s using Nanosight nanoparticle tracking analysis software. Each sample
715 was analyzed 5 times and the counts were averaged.
716
717 Immunoblots
718 Exosomes were prepared by mixing sedimented vesicles with 1X Laemmli buffer, or 1X
719 Laemmli buffer without DTT when detection of CD63 was performed using the BD
720 Bioscience antibody (below). Cell lysates were prepared by adding lysis buffer (10 mM
721 Tris-HCl, pH 7.4, 100 mM NaCl, 0.1% sodium dodecyl sulfate, 0.5% sodium deoxycholate,
722 1% Triton X-100, 10% glycerol) to cell pellets. Protein was quantified using a BCA Protein
723 Assay Kit (Thermo Fischer Scientific), and the selected amount was mixed with Laemmli
724 buffer. Samples were heated at 95°C for 5 min and separated on 4-20% acrylamide Tris-
725 Glycine gradient gels (Life Technologies). Proteins were transferred to polyvinylidene
726 difluoride membranes (EMD Millipore, Darmstadt, Germany), blocked with 5% bovine
727 serum albumin in TBST and incubated overnight with primary antibodies. Blots were then
728 washed with TBST, incubated with anti-rabbit or anti-mouse secondary antibodies (GE
729 Healthcare Life Sciences, Pittsbugh, PA) and detected with ECL-2 reagent (Thermo Fisher
730 Scientific). Primary antibodies used in this study were anti-CD9 #13174S (Cell Signaling
731 Technology, Danvers, MA), anti-alpha integrin alpha 3 #ab190731 (Abcam, Cambridge,
732 MA), anti-Mfg-e8 #MAB2767 (R&D systems, Minneapolis, MN), anti-flotillin-2 #610383 (BD
733 Biosciences, San Jose, CA), anti-TSG101 #GTX70255 (Genetex, Irvine, CA), anti-CD63 #sc-

15363 (Santa Cruz Biotechnology, Dallas, TX), anti-CD63 (BD Biosciences), anti-Ago2 #2897 (Cell Signaling Technology), anti-Dicer #sc-30226 (Santa Cruz Biotechnology), anti-Lupus La #TA500406 (Origene Technologies, Rockville, MD), anti-vinculin #ab129002 (Abcam), anti-Rab27a ab55667 (Abcam), anti-nucleophosmin1 #ab10530 (Abcam), anti-nucleolin #ab22758 (Abcam), anti-tubulin #ab7291 (Abcam), and anti-Hsc-70 #ab19136 (Abcam).

Quantitative real-time PCR

Cellular and EV RNAs were extracted with a mirVana miRNA isolation kit (Thermo Fischer Scientific), unless otherwise specified. Taqman miRNA assays for miRNA detection were purchased from Life Technologies. Assay numbers were: hsa-miR-193b, #002366; hsa-miR-29b-1, #002165; hsa-miR-486, #001278; hsa-miR-574-3p, #002349; hsa-miR-320a, #002277; hsa-miR-142-3p, #000464; hsa-miR-126, #000451; hsa-mir-145-5p, #002278; hsa-mir-122-5p, #002245; hsa-mir-451a, #001141; hsa-mir-182-5, 002334; hsa-miR-429, #001024; hsa-miR-675, #121124_mat; hsa-miR-193a-5p, #002281. As there are no well-accepted endogenous control transcripts for EVs, the relative quantification was normalized to equal amounts of starting RNA material. RNA was quantified by an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) according to the manufacturer's instructions. RNA (1ng) was used for reverse transcription according to the manufacturer's instructions. Relative quantification was calculated from the expression $2^{-(Ct_{(control)}-Ct_{(experimental)})}$. Taqman qPCR master mix with no amperase UNG was obtained from Life Technologies and quantitative real-time PCR was performed using an ABI-7900 real-time PCR system (Life Technologies).

757

758 Immunofluorescence

759 MDA-MB-231 cells on 12mm round coverslips (Corning) were fixed by adding 4% EM-
760 grade formaldehyde (Electron Microscopy Sciences, Hatfield, PA) for 20 min at room
761 temperature. Subsequently, cells were washed 3 times with PBS and
762 permeabilized/blocked by adding blocking buffer (0.1% TX-100 in 2 FBS% for validation
763 experiments, or 0.02% saponin in 2% FBS for superresolution structured illumination
764 microscopy (SIM)) for 20 min. Cells treated with 0.02% saponin retain intact nuclear
765 envelopes and endoplasmic reticulum membranes (Sirkis, Aparicio, and Schekman 2017;
766 Gorur et al. 2017), which allowed us to visualize cytoplasmic La. Cells were then incubated
767 with 1:100 dilution of anti-La (Origene Technologies, #TA500406) and 1:100 dilution of
768 anti-Rab7 (Santa Cruz, sc-6563) in blocking buffer for 1.5 h at room temperature,
769 extensively washed and incubated in secondary antibodies diluted 1:1,000 in blocking
770 buffer and Alexa Fluor 488 and 647 (Thermo Fischer Scientific), for 1.5 h. Cells were
771 extensively washed, rinsed briefly in dH_20 and mounted on slides with ProLong Gold with
772 DAPI (Thermo Fischer). For validation experiments, WT and La-depleted cells were imaged
773 keeping all the settings constant in an Axio Observer Z1 (Zeiss, Oberkochen, Germany). For
774 superresolution microscopy, images were taken with an Elyra P.1 (Zeiss)

775

776 miRNA profiling

777 We used the Discovery Panel from Firefly Service at Abcam to profile 408 miRNAs.
778 Aliquots of 40 µl of 270pg/µl of RNA (samples included HD, LD and CL) were shipped to the
779 Abcam service and the profiling was done according to their standards. Samples were

780 normalized to total amount of RNA. The data used for the analysis was raw data with
781 background subtracted. Only data with more than 0.1 arbitrary units of fluorescence (AUF)
782 was considered for further analysis.
783
784 RNA extraction and TGIRT-seq library preparation
785 Cellular and EV RNAs were extracted by using a mirVana miRNA Isolation Kit. Cellular RNA
786 was extracted using a modified version of the mirVana protocol, enriching for <200nt
787 transcripts, in order to reflect EV RNA composition by transcript size. EV RNA was
788 extracted using the standard miRVana protocol. TGIRT-seq libraries were prepared from 2-
789 10ng of starting material.
790
791 TGIRT-seq libraries were prepared essentially as described (Qin et al. 2016a), with a
792 modification in the starting molecule (see below). Reverse transcription reactions
793 contained purified RNAs, buffer (20 mM Tris-HCl, pH7.5, 450 mM NaCl, 5 mM $MgCl_2$), 5 mM
794 DTT, 100 nM starting molecule (see below) and 1 μM TGIRT-III (Ingex). Reverse
795 transcription by TGIRT-III is initiated by template switching from a starting molecule
796 consisting of a DNA primer (5'-GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTATTAN-3')
797 encoding the reverse complement of the Illumina Read2 sequencing primer binding site
798 (R2R) annealed to a complementary RNA oligonucleotide (R2) such that there is a single
799 nucleotide 3' DNA overhang composed of an equimolar mixture of A, G, C and T. The RNA
800 oligonucleotide is blocked at its 3' end with C3Sp (IDT) to inhibit template switching to
801 itself. Reactions were pre-incubated at room temperature for 30 min and then initiated by
802 addition of 1 mM dNTPs. Reactions were then incubated at 60°C for 15 min and terminated

by adding 5 N NaOH to a final concentration of 0.25 N and incubated at 95°C for 3 min to degrade RNAs and denature protein. The reactions were then cooled to room temperature and neutralized with 5 N HCl. cDNAs were purified by using a Qiagen MinElute Reaction Cleanup Kit and then ligated at their 3' ends to a modified DNA oligonucleotide (5'Phos-CCTGTTATCCCTAGATCGTCGGACTGTAGAACTCTGAACGTGTAC-3'-C3Sp) encoding the reverse complement of the Illumina Read1 primer binding site (R1R) using Thermostable 5' AppDNA/RNA Ligase (New England Biolabs). Ligated cDNAs were re-purified with MinElute Reaction Cleanup Kit and amplified by PCR for 12 cycles using Phusion DNA polymerase (Thermo Fisher Scientific) with overlapping multiplex and barcode primers that add sequences necessary for Illumina sequencing. PCR reactions were purified using a Select-a-size DNA Clean and Concentrator (Zymo) to remove adapter dimers. Libraries were sequenced on a NextSeq 500 instrument (75-nt, single end reads) at the Genome Sequencing and Analysis Facility at the University of Texas at Austin.

RNA sequencing analysis

Illumina TruSeq adapters and PCR primer sequences were trimmed from the reads with cutadapt 1.16 (Martin 2011) (sequencing quality score cut-off at 20) and reads <15-nt after trimming were discarded. Reads were then mapped using HISAT2 v2.0.2 (Kim, Langmead, and Salzberg 2015) with default settings to the human genome reference sequence (Ensembl GRCh38 Release 76) combined with additional contigs for 5S and 45S rRNA genes and the *E. coli* genome sequence (Genebank: NC_000913) (denoted Pass 1). The additional contigs for the 5S and 45S rRNA genes included the 2.2-kb 5S rRNA repeats from the 5S rRNA cluster on chromosome 1 (1q42, GeneBank: X12811) and the 43-kb 45S rRNA

826 repeats that contained 5.8S, 18S and 28S rRNAs from clusters on chromosomes
827 13,14,15,21, and 22 (GeneBank: U13369). Unmapped reads from Pass 1 were re-mapped to
828 Ensembl GRCh38 Release 76 by Bowtie 2 v2.2.6 (Langmead and Salzberg 2012) with local
829 alignment to improve the mapping rate for reads containing post-transcriptionally added 5'
830 or 3' nucleotides (e.g., CCA and poly(U)), short untrimmed adapter sequences, or non-
831 templated nucleotides added to the 3' end of the cDNAs by the TGIRT enzyme (denoted
832 Pass 2). The mapped reads from Passes 1 and 2 were combined using Samtools 1.8 (Li et al.
833 2009) and intersected with gene annotations (Ensembl GRCh38 Release 76) supplemented
834 with the RNY5 gene and its 10 pseudogene sequences, which were not annotated in this
835 release, to generate the counts for individual features. Coverage of each feature was
836 calculated by Bedtools (Quinlan and Hall 2010). To avoid mis-mapping reads with
837 embedded sncRNAs, reads were first intersected with sncRNA annotations and the
838 remaining reads were then intersected with the annotations for protein-coding genes,
839 lincRNAs, antisense, and other lncRNAs. To further improve the mapping rate for tRNAs
840 and rRNAs, we combined reads mapped to tRNAs or rRNAs in the initial alignments and re-
841 mapped them to tRNA reference sequences (Genomic tRNA Database, and UCSC genome
842 browser website) or rRNA reference sequences (GeneBank: X12811 and U13369) using
843 Bowtie 2 local alignment. Because similar or identical tRNAs with the same anticodon may
844 be multiply mapped to different tRNA loci by Bowtie 2, mapped tRNA reads were combined
845 according to their anticodon (N = 48) prior to calculating the tRNA distributions.
846 For miRNA analysis, sequences were mapped to miRbase using miRdeep2. Reads were
847 normalized by dividing the numbers or reads per miRNA to the total number of miRNA
848 reads in each sample and then this value was multiplied by one million (reads per million

849 miRNA mapped reads – RPM).

850

851 CRISPR/Cas9 genome editing

852 A pX330-based plasmid expressing Venus fluorescent protein (Shurtleff et al. 2016) was
853 used to clone the gRNAs targeting Rab27a and Rab35. Two CRIPSR guide RNAs targeting
854 each gene were selected using the CRISPR design tool (Hsu et al. 2013). For Rab27a, gRNAs
855 targeting exon 4 and exon 5 were selected. For Rab35 both gRNAs targeted exon 1. Both
856 gRNAs targeting each gene were cloned simultaneously in the pX330-Venus plasmid
857 according to the PrecisionX Multiplex gRNA Kit instructions (SBI). MDA-MB-231 cells at
858 25% confluency were transfected for 48h, and then, they were trypsinized and sorted for
859 single Venus-positive cells in a 96 well plate using a BD Influx cell sorter. Wells containing
860 clones were allowed to expand (30 clones for Rab27a and 25 for Rab35) and knockouts for
861 each gene were confirmed by immunoblots.

862

863 CRISPR interference

864 MDA-MB-231 cells expressing dCas9-KRAB, as in (Gilbert et al. 2013), were generated
865 using lentivirus. A modified version of the transfer plasmid, as in (Gilbert et al. 2013),
866 UCOE- EF1α-dCas9-BFP-KRAB, was kindly provided by Jonathan Weissman (UCSF). Cells
867 were bulk sorted for BFP signal 3 d post transduction and selected cells were expanded by
868 growth for a few generations, and then frozen and stored as parental cells (these cells are
869 referred to as WT throughout the manuscript). Sequences for gRNAs targeting the
870 promoter of the genes of interest were extracted from (Horlbeck et al. 2016). gRNAs were
871 cloned in plasmid pu6-sgRNA EF1Alpha-puro-T2A-BFP (Gilbert et al. 2014), plasmid

#60955 obtained from Addgene, following the cloning protocol as in (Gilbert et al. 2014). The 3 top gRNAs from the V.2 library (Horlbeck et al. 2016) were chosen per gene of interest. Lentiviruses with the gRNAs targeting the genes of interest were use to transduce the parental cells. Three days post transduction cells were selected with 2ug/ml puromycin for 3 d. Post puromycin selection, cells were collected for up to 3 generations (~ 72 h) for best levels of protein depletion. More doubling times showed reduced protein depletion.

In vitro packaging reactions

Membrane and cytosol preparation

Fractionation of cells and membranes was done as in (Shurtleff et al. 2016) with some modifications as indicated. MDA-MB-231 cells were harvested at ~80% confluency by adding cold PBS and physically removing the cells by the use of a cell scrapper. Cells were then centrifuged at 1,000xg for 10min at 4°C and cell pellets were frozen at -80°C until use. Cells were thawed and resuspended in 2 volumes of HB buffer (20mM HEPES pH 7,4, 250mM sorbitol) containing protease inhibitor cocktail I (1mM 4 aminobenzamidine dihydrochloride, 1 mg/ml antipain dihydrochloride, 1 mg/ml aprotinin, 1 mg/ml leupeptin, 1 mg/ml chymostatin, 1 mM phenymethylsulfonyl fluoride, 50 mM N-tosyl-L-phenylalanine chloromethyl ketone and 1 mg/ml pepstatin). Cells were passed 21-25 times through a 22 gauge needle until >80% of cells were lysed as assessed by microscopy and trypan blue staining. All steps from hereon were done on ice and 4°C, unless otherwise specified. The homogenized cells were centrifuged at 1,500xg for 20 min and the supernatant fraction was subsequently centrifuged at 15,000xg for 15 min to collect donor membranes using a

FA-45-30-11 rotor and an Eppendorf 5430 centrifuge (Eppendorf, Hamburg, Germany). The supernatant fraction was centrifuged again at ~150,000xg (49,000 RPM) in a TLA-55 rotor and Optima Max XP ultracentrifuge (Beckman Coulter) to generate the cytosol fraction (5-6 mg/ml). The 15,000xg membrane fraction (pellet) was resuspended in 1 volume of HB buffer and an equal volume of 0.8M LiCl. Donor membranes were then centrifuged again as before and the pellet fraction was resuspended in 1 volume of original starting material HB buffer.

In vitro miRNA packaging reaction

The *in vitro* packaging reaction was performed as described in (Shurtleff et al. 2016). Briefly, miR-122-5p and the variant versions were purchased from Integrated DNA Technologies (IDT, Coralville, IA). Membranes and cytosols were prepared from MDA-MB-231 cells as described above, with the exception of assays using YBX1 KO cells, in which case HEK293T cells WT and YBX1 KO were used as described (Shurtleff et al. 2016). Complete miRNA packaging assays consist of 10ul membranes, 16ul cytosol (the cytosolic concentration was 5.5 mg/ml, unless otherwise specified), 4ul 10X ATP regeneration system (400mM creatine phosphate, 2mg/ml creatine phosphokinase, 10mM AT, 20mM HEPES pH 7.2, 250mM sorbitol, 150mM KOAc, 5mM MgOAc), 8ul 5X incorporation buffer (400 mM KCl, 100 mM CaCl2, 60 mM HEPES-NaOH, pH 7.4, 6 mM, MgOAc), 1ul of 10nM synthetic single stranded RNA and 1ul RNAsin (Promega, Madison, WI). For rescue experiments with purified La, ~0.1ug of total purified La was added per reaction. Reactions were incubated for 20 min at 30°C, then 1ul of RNase If (50,000units/ml) (NEB, Ipswich, MA) and 4ul of 10X NEB buffer 3 (NEB) was added to the reactions and incubated for an

extra 20 min at 30°C. Following incubation, RNA was extracted using Direct-Zol (Zymo Research, Irvine, CA) according to manufacturer's instructions and miRNA was quantified using TaqMan miRNA assays as described above. The output was represented as a percentage of protection by comparing the level of miRNA left in the RNase-treated samples relative to an ice control which was not RNase-treated. The ice, no RNase control was set to 100%.

Streptavidin pull-down of miR-122 and interacting proteins

The *in vitro* packaging reaction was performed as described above with a modified version of miR-122, 3' biotinylated miR-122 (IDT), which was used to mediate its capture with streptavidin beads. Post-incubation with RNase If, the reactions were heated to 65°C for 15 min to inactivate RNase If and then mixed with 10% TX-100 for a final concentration of 1% and kept on ice for 30 min. Dynabeads Myone Streptavidin T1 (Thermo Fischer Scientific) were washed three times with 1X incorporation buffer and then added to the reactions. The reaction was incubated for 1.5 h at 4°C with constant rotation. Beads were washed with 1X incorporation buffer 5 times and bound proteins were eluted with 1X Laemmli buffer without bromophenol blue. Proteins were electrophoresed in a 4-20% acrylamide Tris-Glycine gradient gel (Life Technologies) for ~3 min. The bulk of proteins were stained with Coomassie and the stained band was excised from the gel using a fresh razor blade. Samples were submitted to the Vincent J. Coates Proteomics/Mass Spectroscopy laboratory at UC Berkeley for in-gel tryptic digestion of proteins followed by liquid chromatography and mass spectrometry analysis according to their standards. The list of detected proteins was then curated for RNA binding proteins, excluding any structural ribosomal protein.

941

942 Protein purification

943 Human La protein was expressed in Sf9 cells as a N-terminal 6X His tagged version.
944 pLJM60-Ssb (Thoreen et al. 2012), plasmid #38241 obtained from Addgene, served as the
945 La sequence backbone. La was cloned in the pFastBac vector from the Bac-to-Bac
946 Baculovirus expression System (Thermo Fischer Scientific) according to the manufacturer's
947 specifications. The generation of bacmids and the production of baculoviruses was
948 performed as indicated by the manufacturer. For La purification, 1 liter of Sf9 cells was
949 infected with baculovirus. Cells expressing La were harvested 2 d post infection by
950 centrifugation at 1,000xg for 15min at 4°C using a Sorvall RC 6+ centrifuge with a
951 FIBERLite F10-6X500Y rotor. Cellular pellet fractions were stored at -20°C until use. The
952 cellular pellet was thawed and resuspended in 40ml of lysis buffer (20mM HEPES, pH8,
953 0.1mM EGTA, 500mM NaCl, 5mM DTT, 10mM imidazole, 10mM $MgCl_2$, 250mM sorbitol, 5%
954 glycerol, 1mM PMSF, and protease inhibitor cocktail 1mM 4-aminobenzamidine
955 dihydrochloride, 1 µg/ml antipain dihydrochloride, 1 µg/ml aprotinin, 1 µg/ml leupeptin, 1
956 µg/ml chymostatin, 1 mM phenymethylsulfonyl fluoride, 50 µM N-tosyl-Lphenylalanine
957 chloromethyl ketone and 1 µg/ml pepstatin)), homogenized and then sonicated. Post-lysis,
958 the material was centrifuged at ~40,000xg (20,000 RPM) using a FIBERlite F211-8X50Y
959 rotor for 20 min at 4°C. The supernatant fraction was centrifuged again at ~42,500xg for 1
960 h using a Ti70 rotor. In preparation for the sample, 2 ml of Ni-NTA resin beads (Thermo
961 Fischer Scientific) were washed twice with lysis buffer. The supernatant then was mixed
962 with Ni-NTA beads and the suspension incubated with constant rotation at 4°C for 2.5 h.
963 Beads were washed 3 times with washing buffer (recipe for washing buffer was the same

964 as lysis buffer (above) but with 50mM imidazole instead of 10mM imidazole). Proteins
965 were eluted with 4ml of elution buffer (elution buffer similar recipe to lysis buffer, but the
966 imidazole concentration was 250mM). Eluted proteins were applied to a Sephadex G25
967 column (Thermo Fischer Scientific). Buffer was exchanged to Buffer B (10mM Tris, pH 7,4,
968 150mM NaCl, 3mM MgCl2, 5% glycerol). Purified fractions were pooled, flash-frozen in
969 liquid nitrogen and stored at -80 C.
970
971 Electrophoretic mobility shift assays
972 In-gel fluorescence was detected in order to assess the free and protein-bound RNA.
973 Fluorescently labeled (5') RNAs (IRD800CWN) were ordered from Integrated DNA
974 Technologies (IDT, Coralville, IA). EMSAS were performed following (Rio 2014), with some
975 modifications. In brief, 1nM of fluorescently labeled RNA was mixed and incubated with
976 increasing amounts of purified La, ranging from 250pM – 2uM. Buffer E was used in this
977 incubation (25mM Tris pH8, 100mM KCl, 1.5mM MgCl2, 0.2mM EGTA, 0.05% Nonidet P-40,
978 50ug/ml heparin). Reactions were incubated at 30°C for 30 min and then chilled on ice for
979 10 min and mixed with 6X loading buffer (60mM KCl, 10mM Tris pH 7,6, 50% glycerol,
980 0.03% (w/v) xylene cyanol). Polyacrylamide gels (6%), acrylamide:bisacrylamide (29:1)
981 (Bio-Rad, Hercules, CA), made with Tris-glycine buffer were prerun for 30min at 200V in
982 the cold room. Samples were resolved in the prerun gels at 200V for 50 min in a 4°C cold
983 room. Fluorescence was detected by using an Odyssey CLx Imaging System (LI-COR
984 Biosciences, Lincoln, NE). The software of the Odyssey CLx Imaging System was used to
985 obtain quantification of fluorescence. To calculate K_ds, we fit Hill equations with quantified
986 data points. Fraction bound was calculated as a function of exhaustion of free-miRNA.

987

988 Immunoprecipitation of La and La-RNA complexes

989 Approximately 40 million cells were harvested as described in the section "Membrane and
990 Cytosol preparation." Cells were homogenized in 2 volumes of HB buffer and physically
991 disrupted as previously described. Non-lysed cells and nuclei were centrifuged at 1,500xg
992 for 20min. The supernatant fraction was used as source of cytoplasmic La for
993 immunoprecipitation and was divided into 3 equal parts: input, La-IP, beads only.
994 Dynabead Protein G (Thermo Fischer Scientific) was washed 3 times in polysome lysis
995 buffer (Peritz et al. 2006). A 5X polysome lysis buffer was made and mixed with 1,500xg
996 post-nuclear supernatant to a final 1X concentration. Anti-La (Origene Technologies,
997 #TA500406) was added to the post nuclear supernatant to a final concentration of
998 2ug/500ul and the mixture was incubated with rotation overnight at 4°C. Dynabeads
999 Protein G (Thermo Fischer) beads were added to the mixture and this mixture was
1000 incubated for an additional 3 h at 4°C with constant rotation. Beads were washed 5 times
1001 with 1X polysome buffer and the content was divided for protein or RNA analysis. Beads
1002 for protein analysis were incubated with Laemmli buffer and heated at 95°C for 10 min and
1003 beads for RNA analysis were exposed to TRI reagent (Zymo Research) for RNA extraction
1004 using Direct-Zol (Zymo Research) as indicated by the manufacturer. Protein and miRNAs
1005 were analyzed by immunoblots and TaqMan miRNA qPCR analysis, respectively.

1006

1007 Motif analysis

1008 For motif analysis, we used the MEME Suite 5.0.2 (Bailey and Elkan 1994). MiRNAs
1009 detected uniquely in the HD sub-population by TGIRT-seq were pooled together with

miRNAs that were found to be enriched by at least 10-fold (HD/cellular lysate) through TGIRT-seq and Firefly profiling. A total of 49 miRNAs met those requirements and were used for further analysis. A 0 and 1 Markov background model was used to find 3-5 nucleotide motifs. The total mature *Homo sapiens* miRNAs from miRBase were used as background.

FIGURE LEGENDS

Figure 1. Two biochemically distinct EV sub-populations are released by MDA-MB-231 cells. a) Schematic showing the two-step purification methodology. Differential ultracentrifugation was followed by buoyant density flotation in a linear iodixanol gradient (Figure 1 – figure supplement 1). b) Immunoblot across the iodixanol gradient for classical EV markers. The two discrete sub-populations are indicated. CD63, a glycosylated protein, runs as a smear. The CD63 signal is indicated by a bracket. c) Nanoparticle tracking analysis showing the size distribution of the HD and LD sub-population. The high-speed pellet is also shown. d) Bioanalyzer analysis of the HD and LD RNA. The high-speed pellet RNA is also shown.

Figure 1 – figure supplement 1. Linearity of iodixanol density gradient. Calculated iodixanol fraction densities collected from the top to bottom of the gradient (fractions 1 to 25) as measured by refractometry (figure 1 – figure supplement 1 – source data 1). Fractions were collected post 160,000g ultracentrifugation (Figure 1a). In red, fractions

corresponding to the LD sub-population. In blue, fractions corresponding to the HD sub-population.

Figure 1 – figure supplement 1 – source data 1. Refraction index and calculated densities of the fractions across the gradient (31 fractions in total).

Figure 2. The two biochemically distinct EV sub-populations co-fractionate with membranes of different sub-cellular origin. a, b) Gene ontology analysis for sub-cellular localization of membrane proteins that coincide with HD and LD proteins detected by mass spectrometry. Size of circles correlates to log10 p-value. c) Analysis of WT, CRISPR-Cas9 Rab27a KO and CRISPR-Cas9 Rab35 KO by immunoblot. Rab27a, Rab35 and vinculin are shown. d) Immunoblots for CD63 and CD9 in EVs secreted by WT, Rab27a KO cells and Rab35 KO cells. The amounts of loaded EVs were normalized by total cellular number.

Figure 2 - source data 1. List of proteins detected by mass spectrometry in both EV sub-populations.

Figure 3. MicroRNA profiling of high buoyant density and low buoyant density EV sub-populations. a, b) Scatterplots showing the relative abundance (as arbitrary units of fluorescence (AUF)) for miRNAs detected in HD and LD sub-populations relative to cellular levels, as detected by Firefly profiling and normalized per ng of total RNA. In blue, HD miRNAs that were selected for further validation by RT-qPCR. Circled miRNAs in LD represent those also found to be enriched in the HD-subpopulation. c, d) Log2 fold change

for top HD and LD candidate miRNAs in HD or LD relative to cells (HD/CL or LD/CL, respectively) and HD relative to LD or vice versa (HD/LD or LD/HD, respectively). MiRNA species were quantified by RT-qPCR and normalized per ng of total RNA. e) Scatterplots showing relative abundance of miRNAs (AUF) detected in HD and LD sub-populations normalized per ng of total RNA, as detected by Firefly profiling. f) RNase protection of highly enriched HD miRNAs quantified by qPCR. Purified EVs were treated with or without RNase If and or Triton X-100. CL: cellular lysate. Dashed red lines in a, b and e, represent 10-fold differences.

Figure 3 – figure supplement 1. MiR-122 and miR-451 are highly enriched in the HD sub-population. Fold difference HD/CL or LD/CL for miR-451a and miR-122. Data extracted from Firefly profiling. CL: cellular lysate.

Figure 3 – figure supplement 2. TGIRT-sequencing of high buoyant density and low buoyant density EV sub-populations. a, b) Scatterplots showing the relative abundance (as reads per million (RPM)) for miRNAs detected in HD and LD sub-populations relative to CL levels. Normalized per total number of miRNA reads. In blue, HD miRNAs that were selected for further validation by qPCR. In red, LD miRNAs selected for further validation by qPCR. Circled miRNAs in LD represent those also found to be enriched in the HD-subpopulation c) Venn Diagram showing the total number of HD and LD uniquely found miRNAs (miRNAs with non-detectable signal in cellular lysates) detected by TGIRT-seq. d) Bar graph showing the distribution of all small non-coding RNA transcripts found in HD and LD sub-populations. e) tRNA start site distribution showing the majority of tRNA reads

start at position 16, previously described as the site of an EV-specific D-loop modification (Shurtleff et al. 2017). The putative D-loop modification is present at relatively same levels in both the HD and LD sub-populations. CL: cellular lysate. Dashed red lines in a and b represent 10-fold differences.

Figure 4. MiR-122 packaging is recapitulated in a cell-free reaction. a) Schematic depicting the *in vitro* packaging reaction. Image reproduced from the original manuscript developing the cell-free reconstitution assay (Shurtleff et al. 2016). b) Quantification of the *in vitro* packaging reaction of miR-122. Reactions with or without membranes (15,000xg pellet), cytosol (150,000xg supernatant) prepared from MDA-MB-231 cells and incubated at 30°C or 4°C are shown. Data plotted are from 3 independent experiments, each with triplicate qPCR reactions; error bars represent standard deviation from independent samples. c) Quantification of the *in vitro* packaging of miR-122 and miR-223. Cytosols from WT or YBX1 KO HEK293T cells were used. Incubations at 30°C or 4°C are shown. Data plotted are from 2 independent experiments, each with triplicate qPCR reactions; error bars represent standard deviation from independent samples. d) List of RNA binding proteins pulled down with biotinylated miR-122 after a modified *in vitro* packaging reaction was performed (see materials and methods).

Figure 4 – figure supplement 1. MiR-122 accumulates in conditioned media (CM) over time. Quantification of total number of miR-122 molecules in CM. CM was incubated with or without MDA-MB-231 cells for a period of 48 h. Samples were collected at the indicated

time points. Cell debris and apoptotic bodies were removed by low and medium speed centrifugation (as in steps 1-2 shown in Fig. 1a).

Figure 4 – figure supplement 2. Ago2 and Dicer are not EV-associated in MB-MDA-231 cells. Immunoblot for EV markers across the linear ioxidanol gradient as depicted in Fig 1a. The presence of Ago2 and Dicer was also tested.

Figure 4 – figure supplement 3. CRISPRi efficiently depletes the RNA binding protein candidates. a) Analysis of RBP depletion by immunoblots. Immunoblots for La, NPM1 and vinculin are shown. b) Relative quantification of miR-122 *in vitro* packaging. Cytosols from WT, La or NPM1-depleted backgrounds were tested in the reactions (cytosols from (a) were used). Two cytosolic protein concentrations were tested: 3 and 6mg/ml.

Figure 5. Sorting of miR-122 into EVs *in vitro* requires La. a) Quantification of miR-122 *in vitro* packaging. WT and La-depleted cytosols were titrated from 1.5 to 6mg/ml. Data plotted are from 2 independent experiments, each with triplicate qPCR reactions; error bars represent standard deviation from independent samples. b) Heterologously expressed La rescues miR-122 *in vitro* packaging. Quantification of miR-122 *in vitro* packaging is shown. Cytosols from WT or La-depleted cells with or without complementation using purified La were used. Reactions without cytosol, with or without membranes and with added purified La are also shown. Data plotted are from 3 independent experiments, each with triplicate qPCR reactions; error bars represent standard deviation from independent samples. c) Immunoblots showing the levels of endogenous or heterologously expressed

1123 added La used in the *in vitro* reactions as in (b) are shown. d) Immunoblots for La following
1124 miR-122 *in vitro* packaging performed as for Fig 3d according to the conditions indicated.
1125
1126 Figure 5 – figure supplement 1. Purification of heterologously expressed La. Coomassie
1127 staining of purified La fractions. Fractions shown were collected post gel filtration on
1128 Sephadex G25 columns.
1129
1130 Figure 6. MiR-122 sorting into EVs requires La *in vivo*. a) Log2 fold change for intracellular
1131 levels of miRNAs of interest post-depletion of the shown RBPs by CRISPRi. MiR-122 and
1132 miR-142 were used as representatives of selectively packaged miRNAs. MiR-574 and miR-
1133 320a were used as representatives of non-selectively packaged miRNAs. Data quantified by
1134 RT-qPCR, normalized per ng of total RNA. b) Log2 fold change of intracellular and secreted
1135 levels of miRNAs of interest quantified by RT-qPCR. EVs were purified from WT and La-
1136 depleted cells. Cellular lysates were isolated at the moment of EV collection. MiR-320a was
1137 used as a control for non-selectively packaged miRNAs. Data plotted are from 3
1138 independent experiments (biological replicates) for miR-122 and miR-142 and from 2
1139 independent experiments (biological replicates) for miR-320a; each with triplicate qPCR
1140 reactions, error bars represent standard deviation from independent samples. c) Schematic
1141 represention of the flotation assay and acetone precipitation as in (d). The high-speed
1142 pellet of conditioned medium was floated in a step sucrose gradient (input) and the soluble
1143 content was released by the addition of TX-100 and then concentrated by acetone
1144 precipitation (acetone precipitated). d) EV samples post-flotation and acetone
1145 precipitation/concentration were tested. Immunoblots for flotillin-2 and La are shown.

Acetone precipitation captures mostly soluble proteins (Feist and Hummon 2015), thus flotillin-2 is only detectable in the floated sample, prior to acetone precipitation, whereas La detection improves post precipitation/concentration. e) Proteinase K protection assays in high-speed pellet fractions. Samples were treated with or without proteinase K and or Triton X-100. Immunoblots for flotillin-2, La and Dicer are shown.

Figure 7: La interacts with miR-122 *in vitro* and *in vivo*. a) Schematic representing immunoprecipitation of endogenous La shown in (b) and (c). b) Immunoblot for La post-La IP. c) Quantification of the La immunoblot post-La IP. The relative levels, as percentage of input, of miR-122 and miR-182 co-IP with La are also shown. MiR-182, a non-selectively sorted miRNA, served as a negative control for La binding. Data plotted are from 2 independent experiments (biological replicates), for qPCR data each with triplicate qPCR reactions; error bars represent standard deviation from independent samples. d) EMSA assays using 5' fluorescently labeled miR-122. Purified La was titrated from 250pM to 2uM. In gel fluorescence was detected. e) Quantification of (d) showing the calculated K_d. Fraction bound was quantified as a function of exhaustion of free miRNA.

Figure 7 – figure supplement 1. Detection of nuclear and cytoplasmic La by immunofluorescence. a) Immunofluorescence for endogenous La was performed. WT and La-depleted cells were fixed and permeabilized with 0.1% Triton X-100. La and DAPI staining are shown. b) Cells with arrows in (a) are shown side by side to confirm antibody specificity. c) La associates with Rab7-positive vesicles. Structured illumination microscopy

(SIM) was used to achieve higher resolution. Cells were fixed, permeabilized with 0.02% saponin, and antibodies against La and Rab7 were used.

Figure 7- figure supplement 2. Controls for La specificity during EMSA. a) 22nt RNA oligonucleotide consisting of alternating purines was incubated with different concentrations of purified La. Negative control for binding. b) 22nt poly-uridine oligonucleotide was incubated with different concentrations of purified La as a positive control for binding.

Figure 8. A bipartite motif in miR-122 is required for its packaging and interaction with La *in vitro*. a) The sequences of miR-122 WT and mutated versions are shown. b) Quantification of *in vitro* packaging of miR-122 WT and mutated versions. Data plotted are from 3 independent experiments, each with triplicate qPCR reactions; error bars represent standard deviation from independent samples. c) Binding affinity curves (as calculated from EMSAs) for miR-122 and the mutated versions are shown. Fraction bound was quantified as a function of exhaustion of free miRNA.

Figure 8 – figure supplement 1. EMSA for miR-122 mutated versions. a) miR-122 3' mutated oligonucleotide incubated with increasing amounts of La. b) miR-122 5' mutated oligonucleotide incubated with increasing amounts of La.

Figure 9: Diagram representing the current model of miRNA sorting into extracellular vesicles derived from MDA-MB-231 cells. HD vesicles have their origin in the endocytic

pathway, representing the classical exosomes. We propose a selective mechanism of miRNA sorting occurring at the site of HD vesicle biogenesis where La, along with bound miR-122 (which interaction is mediated by at least two RNA motifs), is targeted for capture into a bud invaginating into the interior of an endosome. LD vesicles may originate at the plasma membrane, representing shedding vesicles. There may be no selective miRNA sorting occurring at the site of LD vesicle biogenesis. Non-selective miRNA sorting occurs in both the HD and LD sub-populations. MVB: multivesicular body, PM: plasma membrane.

Acknowledgements

We thank Kathleen Collins, Arash Komeili and Donald Rio for advice. We also thank Kathleen Collins and Daniel Sirkis for reading and editing the manuscript. We would also like to thank the staff at the UC Berkeley shared facilities, the DNA sequencing facility, the Vincent J Coates Proteomics Facility, the Flow Cytometry Facility, the Cell Culture Facility and the Genomic Sequencing and Analysis Facility at UT Austin. Cartoons were created with Biorender (Toronto, Ontario, Canada).

Competing interests

Thermostable group II intron reverse transcriptase (TGIRT) enzymes and methods for their use are the subject of patents and patent applications that have been licensed by the University of Texas and East Tennessee State University to InGex, LLC. A.M.L., some former and present members of the A.M.L. laboratory, and the University of Texas are minority

equity holders in InGex, LLC and receive royalty payments from the sale of TGIRT enzymes and kits and from the sublicensing of intellectual property to other companies.

REFERENCES

Admyre, C., S. M. Johansson, K. R. Qazi, J. J. Filen, R. Lahesmaa, M. Norman, E. P. Neve, A. Scheynius, and S. Gabrielsson. 2007. 'Exosomes with immune modulatory features are present in human breast milk', *J Immunol*, 179: 1969-78.

Al-Ejeh, F., J. M. Darby, and M. P. Brown. 2007. 'The La autoantigen is a malignancy-associated cell death target that is induced by DNA-damaging drugs', *Clin Cancer Res*, 13: 5509s-18s.

Baglio, S. R., M. A. van Eijndhoven, D. Koppers-Lalic, J. Berenguer, S. M. Lougheed, S. Gibbs, N. Leveille, R. N. Rinkel, E. S. Hopmans, S. Swaminathan, S. A. Verkuijlen, G. L. Scheffer, F. J. van Kuppeveld, T. D. de Gruijl, I. E. Bultink, E. S. Jordanova, M. Hackenberg, S. R. Piersma, J. C. Knol, A. E. Voskuyl, T. Wurdinger, C. R. Jimenez, J. M. Middeldorp, and D. M. Pegtel. 2016. 'Sensing of latent EBV infection through exosomal transfer of 5'pppRNA', *Proc Natl Acad Sci U S A*, 113: E587-96.

Bailey, T. L., and C. Elkan. 1994. 'Fitting a mixture model by expectation maximization to discover motifs in biopolymers', *Proc Int Conf Intell Syst Mol Biol*, 2: 28-36.

Bartel, D. P. 2004. 'MicroRNAs: genomics, biogenesis, mechanism, and function', *Cell*, 116: 281-97.

Bobrie, A., M. Colombo, S. Krumeich, G. Raposo, and C. Thery. 2012. 'Diverse subpopulations of vesicles secreted by different intracellular mechanisms are present in exosome preparations obtained by differential ultracentrifugation', *J Extracell Vesicles*, 1.

Boyle, E. I., S. Weng, J. Gollub, H. Jin, D. Botstein, J. M. Cherry, and G. Sherlock. 2004. 'GO::TermFinder--open source software for accessing Gene Ontology information and finding significantly enriched Gene Ontology terms associated with a list of genes', *Bioinformatics*, 20: 3710-5.

Caby, M. P., D. Lankar, C. Vincendeau-Scherrer, G. Raposo, and C. Bonnerot. 2005. 'Exosomal-like vesicles are present in human blood plasma', *Int Immunol*, 17: 879-87.

Cardinali, B., C. Carissimi, P. Gravina, and P. Pierandrei-Amaldi. 2003. 'La protein is associated with terminal oligopyrimidine mRNAs in actively translating polysomes', *J Biol Chem*, 278: 35145-51.

Cha, D. J., J. L. Franklin, Y. Dou, Q. Liu, J. N. Higginbotham, M. Demory Beckler, A. M. Weaver, K. Vickers, N. Prasad, S. Levy, B. Zhang, R. J. Coffey, and J. G. Patton. 2015. 'KRAS-dependent sorting of miRNA to exosomes', *Elife*, 4: e07197.

Cocucci, E., G. Racchetti, and J. Meldolesi. 2009. 'Shedding microvesicles: artefacts no more', *Trends Cell Biol*, 19: 43-51.

1253 Colombo, M., G. Raposo, and C. Thery. 2014. 'Biogenesis, secretion, and intercellular
1254 interactions of exosomes and other extracellular vesicles', *Annu Rev Cell Dev Biol*, 30:
1255 255-89.
1256 Coulouarn, C., V. M. Factor, J. B. Andersen, M. E. Durkin, and S. S. Thorgeirsson. 2009. 'Loss
1257 of miR-122 expression in liver cancer correlates with suppression of the hepatic
1258 phenotype and gain of metastatic properties', *Oncogene*, 28: 3526-36.
1259 Crosio, C., P. P. Boyl, F. Loreni, P. Pierandrei-Amaldi, and F. Amaldi. 2000. 'La protein has a
1260 positive effect on the translation of TOP mRNAs in vivo', *Nucleic Acids Res*, 28: 2927-
1261 34.
1262 D'Souza, V., and M. F. Summers. 2005. 'How retroviruses select their genomes', *Nat Rev*
1263 *Microbiol*, 3: 643-55.
1264 Demory Beckler, M., J. N. Higginbotham, J. L. Franklin, A. J. Ham, P. J. Halvey, I. E. Imasuen, C.
1265 Whitwell, M. Li, D. C. Liebler, and R. J. Coffey. 2013. 'Proteomic analysis of exosomes
1266 from mutant KRAS colon cancer cells identifies intercellular transfer of mutant
1267 KRAS', *Mol Cell Proteomics*, 12: 343-55.
1268 Dickman, C. T., J. Lawson, J. Jabalee, S. A. MacLellan, N. E. LePard, K. L. Bennewith, and C.
1269 Garnis. 2017. 'Selective extracellular vesicle exclusion of miR-142-3p by oral cancer
1270 cells promotes both internal and extracellular malignant phenotypes', *Oncotarget*, 8:
1271 15252-66.
1272 Feist, P., and A. B. Hummon. 2015. 'Proteomic challenges: sample preparation techniques
1273 for microgram-quantity protein analysis from biological samples', *Int J Mol Sci*, 16:
1274 3537-63.
1275 Flores, O., E. M. Kennedy, R. L. Skalsky, and B. R. Cullen. 2014. 'Differential RISC association
1276 of endogenous human microRNAs predicts their inhibitory potential', *Nucleic Acids*
1277 *Res*, 42: 4629-39.
1278 Fong, M. Y., W. Zhou, L. Liu, A. Y. Alontaga, M. Chandra, J. Ashby, A. Chow, S. T. O'Connor, S.
1279 Li, A. R. Chin, G. Somlo, M. Palomares, Z. Li, J. R. Tremblay, A. Tsuyada, G. Sun, M. A.
1280 Reid, X. Wu, P. Swiderski, X. Ren, Y. Shi, M. Kong, W. Zhong, Y. Chen, and S. E. Wang.
1281 2015. 'Breast-cancer-secreted miR-122 reprograms glucose metabolism in
1282 premetastatic niche to promote metastasis', *Nat Cell Biol*, 17: 183-94.
1283 Gibbings, D. J., C. Ciaudo, M. Erhardt, and O. Voinnet. 2009. 'Multivesicular bodies associate
1284 with components of miRNA effector complexes and modulate miRNA activity', *Nat*
1285 *Cell Biol*, 11: 1143-9.
1286 Gilbert, L. A., M. A. Horlbeck, B. Adamson, J. E. Villalta, Y. Chen, E. H. Whitehead, C.
1287 Guimaraes, B. Panning, H. L. Ploegh, M. C. Bassik, L. S. Qi, M. Kampmann, and J. S.
1288 Weissman. 2014. 'Genome-Scale CRISPR-Mediated Control of Gene Repression and
1289 Activation', *Cell*, 159: 647-61.
1290 Gilbert, L. A., M. H. Larson, L. Morsut, Z. Liu, G. A. Brar, S. E. Torres, N. Stern-Ginossar, O.
1291 Brandman, E. H. Whitehead, J. A. Doudna, W. A. Lim, J. S. Weissman, and L. S. Qi.
1292 2013. 'CRISPR-mediated modular RNA-guided regulation of transcription in
1293 eukaryotes', *Cell*, 154: 442-51.
1294 Gorur, A., L. Yuan, S. J. Kenny, S. Baba, K. Xu, and R. Schekman. 2017. 'COPII-coated
1295 membranes function as transport carriers of intracellular procollagen I', *J Cell Biol*,
1296 216: 1745-59.
1297 Harding, C., J. Heuser, and P. Stahl. 1983. 'Receptor-mediated endocytosis of transferrin and
1298 recycling of the transferrin receptor in rat reticulocytes', *J Cell Biol*, 97: 329-39.

Horlbeck, M. A., L. A. Gilbert, J. E. Villalta, B. Adamson, R. A. Pak, Y. Chen, A. P. Fields, C. Y. Park, J. E. Corn, M. Kampmann, and J. S. Weissman. 2016. 'Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation', *Elife*, 5.

Hsu, C., Y. Morohashi, S. Yoshimura, N. Manrique-Hoyos, S. Jung, M. A. Lauterbach, M. Bakhti, M. Gronborg, W. Mobius, J. Rhee, F. A. Barr, and M. Simons. 2010. 'Regulation of exosome secretion by Rab35 and its GTPase-activating proteins TBC1D10A-C', *J Cell Biol*, 189: 223-32.

Hsu, P. D., D. A. Scott, J. A. Weinstein, F. A. Ran, S. Konermann, V. Agarwala, Y. Li, E. J. Fine, X. Wu, O. Shalem, T. J. Cradick, L. A. Marraffini, G. Bao, and F. Zhang. 2013. 'DNA targeting specificity of RNA-guided Cas9 nucleases', *Nat Biotechnol*, 31: 827-32.

Hsu, Y. L., J. Y. Hung, W. A. Chang, Y. S. Lin, Y. C. Pan, P. H. Tsai, C. Y. Wu, and P. L. Kuo. 2017. 'Hypoxic lung cancer-secreted exosomal miR-23a increased angiogenesis and vascular permeability by targeting prolyl hydroxylase and tight junction protein ZO-1', *Oncogene*, 36: 4929-42.

Intine, R. V., S. A. Tenenbaum, A. L. Sakulich, J. D. Keene, and R. J. Maraia. 2003. 'Differential phosphorylation and subcellular localization of La RNPs associated with precursor tRNAs and translation-related mRNAs', *Mol Cell*, 12: 1301-7.

Kim, D., B. Langmead, and S. L. Salzberg. 2015. 'HISAT: a fast spliced aligner with low memory requirements', *Nat Methods*, 12: 357-60.

Koppers-Lalic, D., M. Hackenberg, I. V. Bijnsdorp, M. A. J. van Eijndhoven, P. Sadek, D. Sie, N. Zini, J. M. Middeldorp, B. Ylstra, R. X. de Menezes, T. Wurdinger, G. A. Meijer, and D. M. Pegtel. 2014. 'Nontemplated nucleotide additions distinguish the small RNA composition in cells from exosomes', *Cell Rep*, 8: 1649-58.

Kowal, J., G. Arras, M. Colombo, M. Jouve, J. P. Morath, B. Primdal-Bengtson, F. Dingli, D. Loew, M. Tkach, and C. Thery. 2016. 'Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes', *Proc Natl Acad Sci U S A*, 113: E968-77.

Langmead, B., and S. L. Salzberg. 2012. 'Fast gapped-read alignment with Bowtie 2', *Nat Methods*, 9: 357-9.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and Subgroup Genome Project Data Processing. 2009. 'The Sequence Alignment/Map format and SAMtools', *Bioinformatics*, 25: 2078-9.

Liang, B., P. Peng, S. Chen, L. Li, M. Zhang, D. Cao, J. Yang, H. Li, T. Gui, X. Li, and K. Shen. 2013. 'Characterization and proteomic analysis of ovarian cancer-derived exosomes', *J Proteomics*, 80: 171-82.

Martin, Marcel. 2011. 'Cutadapt removes adapter sequences from high-throughput sequencing reads', *2011*, 17: 3.

Meckes, D. G., Jr., H. P. Gunawardena, R. M. Dekroon, P. R. Heaton, R. H. Edwards, S. Ozgur, J. D. Griffith, B. Damania, and N. Raab-Traub. 2013. 'Modulation of B-cell exosome proteins by gamma herpesvirus infection', *Proc Natl Acad Sci U S A*, 110: E2925-33.

Mittelbrunn, M., C. Gutierrez-Vazquez, C. Villarroya-Beltri, S. Gonzalez, F. Sanchez-Cabo, M. A. Gonzalez, A. Bernad, and F. Sanchez-Madrid. 2011. 'Unidirectional transfer of microRNA-loaded exosomes from T cells to antigen-presenting cells', *Nat Commun*, 2: 282.

Mohr, S., E. Ghanem, W. Smith, D. Sheeter, Y. Qin, O. King, D. Polioudakis, V. R. Iyer, S. Hunicke-Smith, S. Swamy, S. Kuersten, and A. M. Lambowitz. 2013. 'Thermostable group II intron reverse transcriptase fusion proteins and their use in cDNA synthesis and next-generation RNA sequencing', *RNA*, 19: 958-70.

Montecalvo, A., A. T. Larregina, W. J. Shufesky, D. B. Stolz, M. L. Sullivan, J. M. Karlsson, C. J. Baty, G. A. Gibson, G. Erdos, Z. Wang, J. Milosevic, O. A. Tkacheva, S. J. Divito, R. Jordan, J. Lyons-Weiler, S. C. Watkins, and A. E. Morelli. 2012. 'Mechanism of transfer of functional microRNAs between mouse dendritic cells via exosomes', *Blood*, 119: 756-66.

Mukherjee, K., B. Ghoshal, S. Ghosh, Y. Chakrabarty, S. Shwetha, S. Das, and S. N. Bhattacharyya. 2016. 'Reversible HuR-microRNA binding controls extracellular export of miR-122 and augments stress response', *EMBO Rep*, 17: 1184-203.

Ostenfeld, M. S., D. K. Jeppesen, J. R. Laurberg, A. T. Boysen, J. B. Bramsen, B. Primdal-Bengtson, A. Hendrix, P. Lamy, F. Dagnaes-Hansen, M. H. Rasmussen, K. H. Bui, N. Fristrup, E. I. Christensen, I. Nordentoft, J. P. Morth, J. B. Jensen, J. S. Pedersen, M. Beck, D. Theodorescu, M. Borre, K. A. Howard, L. Dyrskjot, and T. F. Orntoft. 2014. 'Cellular disposal of miR23b by RAB27-dependent exosome release is linked to acquisition of metastatic properties', *Cancer Res*, 74: 5758-71.

Ostrowski, M., N. B. Carmo, S. Krumeich, I. Fanget, G. Raposo, A. Savina, C. F. Moita, K. Schauer, A. N. Hume, R. P. Freitas, B. Goud, P. Benaroch, N. Hacohen, M. Fukuda, C. Desnos, M. C. Seabra, F. Darchen, S. Amigorena, L. F. Moita, and C. Thery. 2010. 'Rab27a and Rab27b control different steps of the exosome secretion pathway', *Nat Cell Biol*, 12: 19-30; sup pp 1-13.

Pan, B. T., K. Teng, C. Wu, M. Adam, and R. M. Johnstone. 1985. 'Electron microscopic evidence for externalization of the transferrin receptor in vesicular form in sheep reticulocytes', *J Cell Biol*, 101: 942-8.

Pegtel, D. M., K. Cosmopoulos, D. A. Thorley-Lawson, M. A. van Eijndhoven, E. S. Hopmans, J. L. Lindenberg, T. D. de Gruijl, T. Wurdinger, and J. M. Middeldorp. 2010. 'Functional delivery of viral miRNAs via exosomes', *Proc Natl Acad Sci U S A*, 107: 6328-33.

Peinado, H., H. Zhang, I. R. Matei, B. Costa-Silva, A. Hoshino, G. Rodrigues, B. Psaila, R. N. Kaplan, J. F. Bromberg, Y. Kang, M. J. Bissell, T. R. Cox, A. J. Giaccia, J. T. Erler, S. Hiratsuka, C. M. Ghajar, and D. Lyden. 2017. 'Pre-metastatic niches: organ-specific homes for metastases', *Nat Rev Cancer*, 17: 302-17.

Peritz, T., F. Zeng, T. J. Kannanayakal, K. Kilk, E. Eiriksdottir, U. Langel, and J. Eberwine. 2006. 'Immunoprecipitation of mRNA-protein complexes', *Nat Protoc*, 1: 577-80.

Petz, M., N. C. Them, H. Huber, and W. Mikulits. 2012. 'PDGF enhances IRES-mediated translation of Laminin B1 by cytoplasmic accumulation of La during epithelial to mesenchymal transition', *Nucleic Acids Res*, 40: 9738-49.

Petz, M., N. Them, H. Huber, H. Beug, and W. Mikulits. 2012. 'La enhances IRES-mediated translation of laminin B1 during malignant epithelial to mesenchymal transition', *Nucleic Acids Res*, 40: 290-302.

Pisitkun, T., R. F. Shen, and M. A. Knepper. 2004. 'Identification and proteomic profiling of exosomes in human urine', *Proc Natl Acad Sci U S A*, 101: 13368-73.

Qin, Y., J. Yao, D. C. Wu, R. M. Nottingham, S. Mohr, S. Hunicke-Smith, and A. M. Lambowitz. 2016a. 'High-throughput sequencing of human plasma RNA by using thermostable

group II intron reverse transcriptases', *RNA.*, 22: 111-28. doi: 10.1261/rna.054809.115. Epub 2015 Nov 9.

———. 2016b. 'High-throughput sequencing of human plasma RNA by using thermostable group II intron reverse transcriptases', *RNA*, 22: 111-28.

Quinlan, A. R., and I. M. Hall. 2010. 'BEDTools: a flexible suite of utilities for comparing genomic features', *Bioinformatics*, 26: 841-2.

Rana, S., K. Malinowska, and M. Zoller. 2013. 'Exosomal tumor microRNA modulates premetastatic organ cells', *Neoplasia*, 15: 281-95.

Raposo, G., and W. Stoorvogel. 2013. 'Extracellular vesicles: exosomes, microvesicles, and friends', *J Cell Biol*, 200: 373-83.

Rinke, J., and J. A. Steitz. 1985. 'Association of the lupus antigen La with a subset of U6 snRNA molecules', *Nucleic Acids Res*, 13: 2617-29.

Rio, D. C. 2014. 'Electrophoretic mobility shift assays for RNA-protein complexes', *Cold Spring Harb Protoc*, 2014: 435-40.

Rodriguez, M., J. Silva, A. Lopez-Alfonso, M. B. Lopez-Muniz, C. Pena, G. Dominguez, J. M. Garcia, A. Lopez-Gonzalez, M. Mendez, M. Provencio, V. Garcia, and F. Bonilla. 2014. 'Different exosome cargo from plasma/bronchoalveolar lavage in non-small-cell lung cancer', *Genes Chromosomes Cancer*, 53: 713-24.

Santangelo, L., G. Giurato, C. Cicchini, C. Montaldo, C. Mancone, R. Tarallo, C. Battistelli, T. Alonzi, A. Weisz, and M. Tripodi. 2016. 'The RNA-Binding Protein SYNCRIP Is a Component of the Hepatocyte Exosomal Machinery Controlling MicroRNA Sorting', *Cell Rep*, 17: 799-808.

Shurtleff, M. J., M. M. Temoche-Diaz, K. V. Karfilis, S. Ri, and R. Schekman. 2016. 'Y-box protein 1 is required to sort microRNAs into exosomes in cells and in a cell-free reaction', *Elife*, 5.

Shurtleff, M. J., J. Yao, Y. Qin, R. M. Nottingham, M. M. Temoche-Diaz, R. Schekman, and A. M. Lambowitz. 2017. 'Broad role for YBX1 in defining the small noncoding RNA composition of exosomes', *Proc Natl Acad Sci U S A*, 114: E8987-E95.

Shurtleff, Matthew J., Morayma M. Temoche-Diaz, and Randy Schekman. 2018. 'Extracellular Vesicles and Cancer: Caveat Lector', *Annual Review of Cancer Biology*, 2: 395-411.

Sirkis, D. W., R. E. Aparicio, and R. Schekman. 2017. 'Neurodegeneration-associated mutant TREM2 proteins abortively cycle between the ER and ER-Golgi intermediate compartment', *Mol Biol Cell*, 28: 2723-33.

Skottvoll, Froydis Sved, Henriette Engen Berg, Kamilla Bjorseth, Kaja Lund, Norbert Roos, Sara Bekhradnia, Bernd Thiede, Hanne Roberg-Larsen, Bo Nystrom, Elsa Lundanes, and Steven Ray Wilson. 2018. 'Comparison of ultracentrifugation and a commercial kit for isolation of exosomes derived from glioblastoma and breast cancer cell lines', *bioRxiv*.

Sommer, G., J. Dittmann, J. Kuehnert, K. Reumann, P. E. Schwartz, H. Will, B. L. Coulter, M. T. Smith, and T. Heise. 2011. 'The RNA-binding protein La contributes to cell proliferation and CCND1 expression', *Oncogene*, 30: 434-44.

Sommer, G., C. Rossa, A. C. Chi, B. W. Neville, and T. Heise. 2011. 'Implication of RNA-binding protein La in proliferation, migration and invasion of lymph node-metastasized hypopharyngeal SCC cells', *PLoS One*, 6: e25402.

Squadrito, M. L., C. Baer, F. Burdet, C. Maderna, G. D. Gilfillan, R. Lyle, M. Ibberson, and M. De Palma. 2014. 'Endogenous RNAs modulate microRNA sorting to exosomes and transfer to acceptor cells', *Cell Rep*, 8: 1432-46.

Stefano, J. E. 1984. 'Purified lupus antigen La recognizes an oligouridylate stretch common to the 3' termini of RNA polymerase III transcripts', *Cell*, 36: 145-54.

Supek, F., M. Bosnjak, N. Skunca, and T. Smuc. 2011. 'REVIGO summarizes and visualizes long lists of gene ontology terms', *PLoS One*, 6: e21800.

Swanson, C. M., and M. H. Malim. 2006. 'Retrovirus RNA trafficking: from chromatin to invasive genomes', *Traffic*, 7: 1440-50.

Szczepanski, M. J., M. Szajnik, A. Welsh, T. L. Whiteside, and M. Boyiadzis. 2011. 'Blast-derived microvesicles in sera from patients with acute myeloid leukemia suppress natural killer cell function via membrane-associated transforming growth factor-beta1', *Haematologica*, 96: 1302-9.

Tan, G. S., B. G. Garchow, X. Liu, J. Yeung, J. P. th Morris, T. L. Cuellar, M. T. McManus, and M. Kiriakidou. 2009. 'Expanded RNA-binding activities of mammalian Argonaute 2', *Nucleic Acids Res*, 37: 7533-45.

Thoreen, C. C., L. Chantranupong, H. R. Keys, T. Wang, N. S. Gray, and D. M. Sabatini. 2012. 'A unifying model for mTORC1-mediated regulation of mRNA translation', *Nature*, 485: 109-13.

Tominaga, N., N. Kosaka, M. Ono, T. Katsuda, Y. Yoshioka, K. Tamura, J. Lotvall, H. Nakagama, and T. Ochiya. 2015. 'Brain metastatic cancer cells release microRNA-181c-containing extracellular vesicles capable of destructing blood-brain barrier', *Nat Commun*, 6: 6716.

Tosar, J. P., F. Gambaro, J. Sanguinetti, B. Bonilla, K. W. Witwer, and A. Cayota. 2015. 'Assessment of small RNA sorting into different extracellular fractions revealed by high-throughput sequencing of breast cell lines', *Nucleic Acids Res*, 43: 5601-16.

Trotta, R., T. Vignudelli, O. Candini, R. V. Intine, L. Pecorari, C. Guerzoni, G. Santilli, M. W. Byrom, S. Goldoni, L. P. Ford, M. A. Caligiuri, R. J. Maraia, D. Perrotti, and B. Calabretta. 2003. 'BCR/ABL activates mdm2 mRNA translation via the La antigen', *Cancer Cell*, 3: 145-60.

Van Deun, J., P. Mestdagh, R. Sormunen, V. Cocquyt, K. Vermaelen, J. Vandesompele, M. Bracke, O. De Wever, and A. Hendrix. 2014. 'The impact of disparate isolation methods for extracellular vesicles on downstream RNA profiling', *J Extracell Vesicles*, 3.

van Niekerk, E. A., D. E. Willis, J. H. Chang, K. Reumann, T. Heise, and J. L. Twiss. 2007. 'Sumoylation in axons triggers retrograde transport of the RNA-binding protein La', *Proc Natl Acad Sci U S A*, 104: 12913-8.

Vella, L. J., R. A. Sharples, V. A. Lawson, C. L. Masters, R. Cappai, and A. F. Hill. 2007. 'Packaging of prions into exosomes is associated with a novel pathway of PrP processing', *J Pathol*, 211: 582-90.

Villarroya-Beltri, C., C. Gutierrez-Vazquez, F. Sanchez-Cabo, D. Perez-Hernandez, J. Vazquez, N. Martin-Cofreces, D. J. Martinez-Herrera, A. Pascual-Montano, M. Mittelbrunn, and F. Sanchez-Madrid. 2013. 'Sumoylated hnRNPA2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs', *Nat Commun*, 4: 2980.

Wang, B., H. Wang, and Z. Yang. 2012. 'MiR-122 inhibits cell proliferation and tumorigenesis of breast cancer by targeting IGF1R', *PLoS One*, 7: e47053.

Wei, Z., A. O. Batagov, D. R. Carter, and A. M. Krichevsky. 2016. 'Fetal Bovine Serum RNA Interferes with the Cell Culture derived Extracellular RNA', *Sci Rep*, 6: 31175.

Willms, E., H. J. Johansson, I. Mager, Y. Lee, K. E. Blomberg, M. Sadik, A. Alaarg, C. I. Smith, J. Lehtio, S. El Andaloussi, M. J. Wood, and P. Vader. 2016. 'Cells release subpopulations of exosomes with distinct molecular and biological properties', *Sci Rep*, 6: 22519.

Wolin, S. L., and T. Cedervall. 2002. 'The La protein', *Annu Rev Biochem*, 71: 375-403.

Wu, X., G. Somlo, Y. Yu, M. R. Palomares, A. X. Li, W. Zhou, A. Chow, Y. Yen, J. J. Rossi, H. Gao, J. Wang, Y. C. Yuan, P. Frankel, S. Li, K. T. Ashing-Giwa, G. Sun, Y. Wang, R. Smith, K. Robinson, X. Ren, and S. E. Wang. 2012. 'De novo sequencing of circulating miRNAs identifies novel markers predicting clinical outcome of locally advanced breast cancer', *J Transl Med*, 10: 42.

Zhou, W., M. Y. Fong, Y. Min, G. Somlo, L. Liu, M. R. Palomares, Y. Yu, A. Chow, S. T. O'Connor, A. R. Chin, Y. Yen, Y. Wang, E. G. Marcusson, P. Chu, J. Wu, X. Wu, A. X. Li, Z. Li, H. Gao, X. Ren, M. P. Boldin, P. C. Lin, and S. E. Wang. 2014. 'Cancer-secreted miR-105 destroys vascular endothelial barriers to promote metastasis', *Cancer Cell*, 25: 501-15.

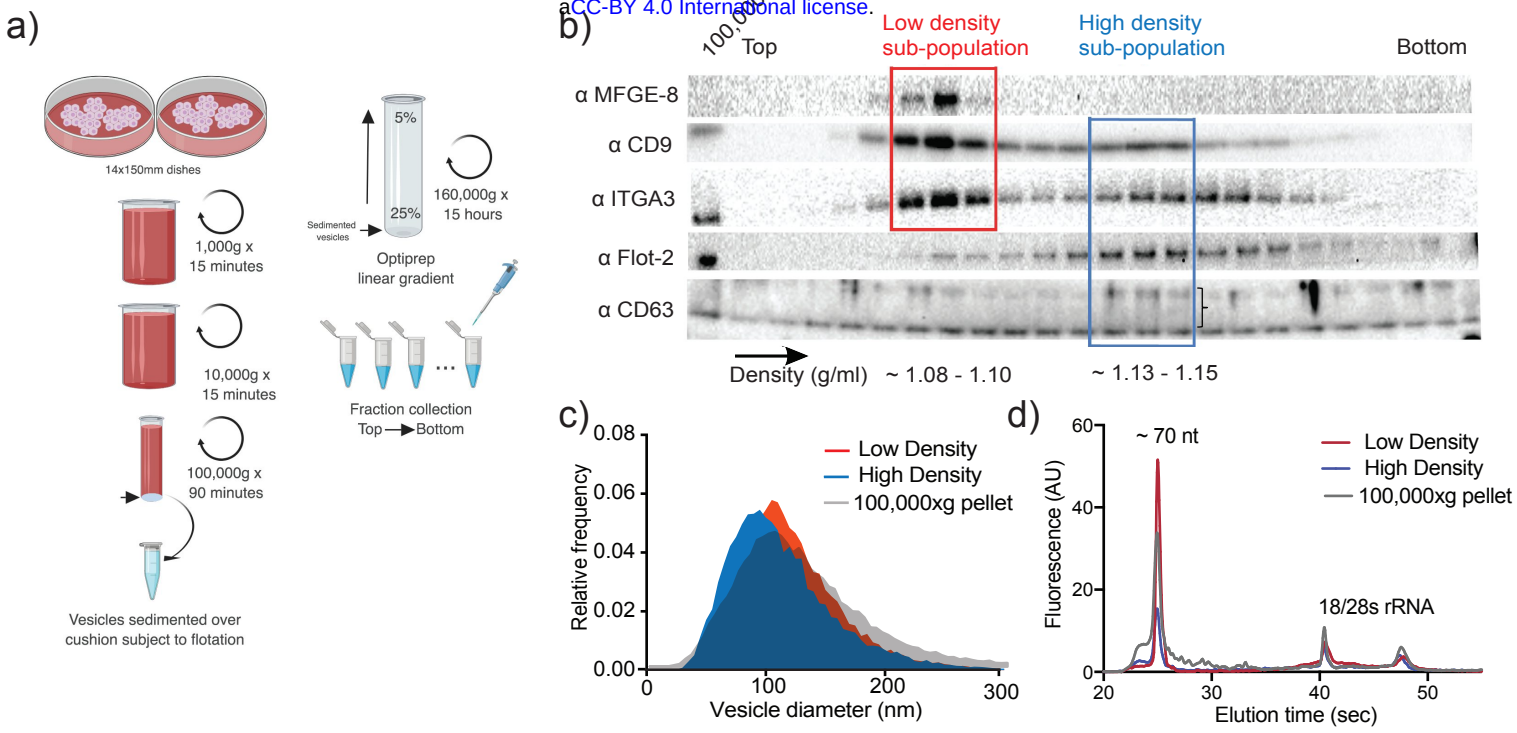


Figure 1. Two biochemically distinct EV sub-populations are released by MDA-MB-231 cells.

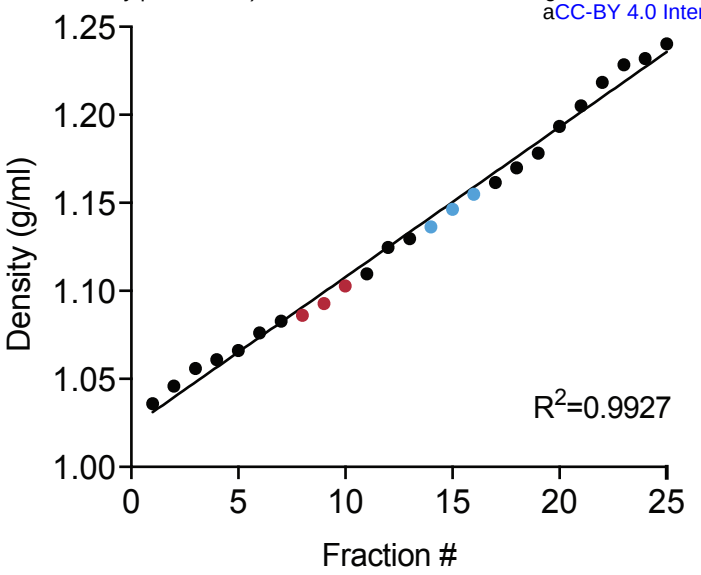


Figure 1 - figure supplement 1. Linearity of iodixanol density gradient.

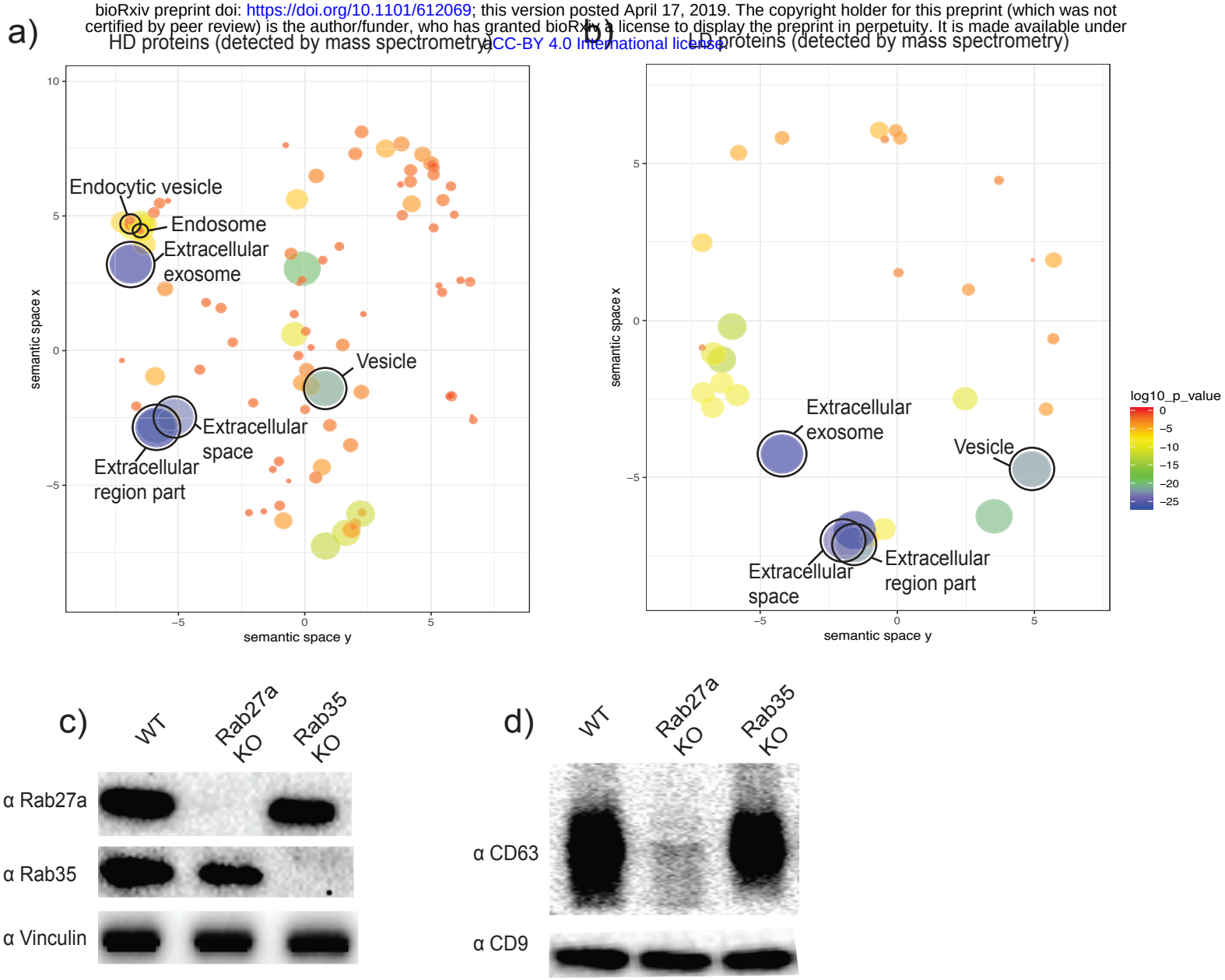


Figure 2. The two biochemically distinct EV sub-populations co-fractionate with membranes of different sub-cellular origin.

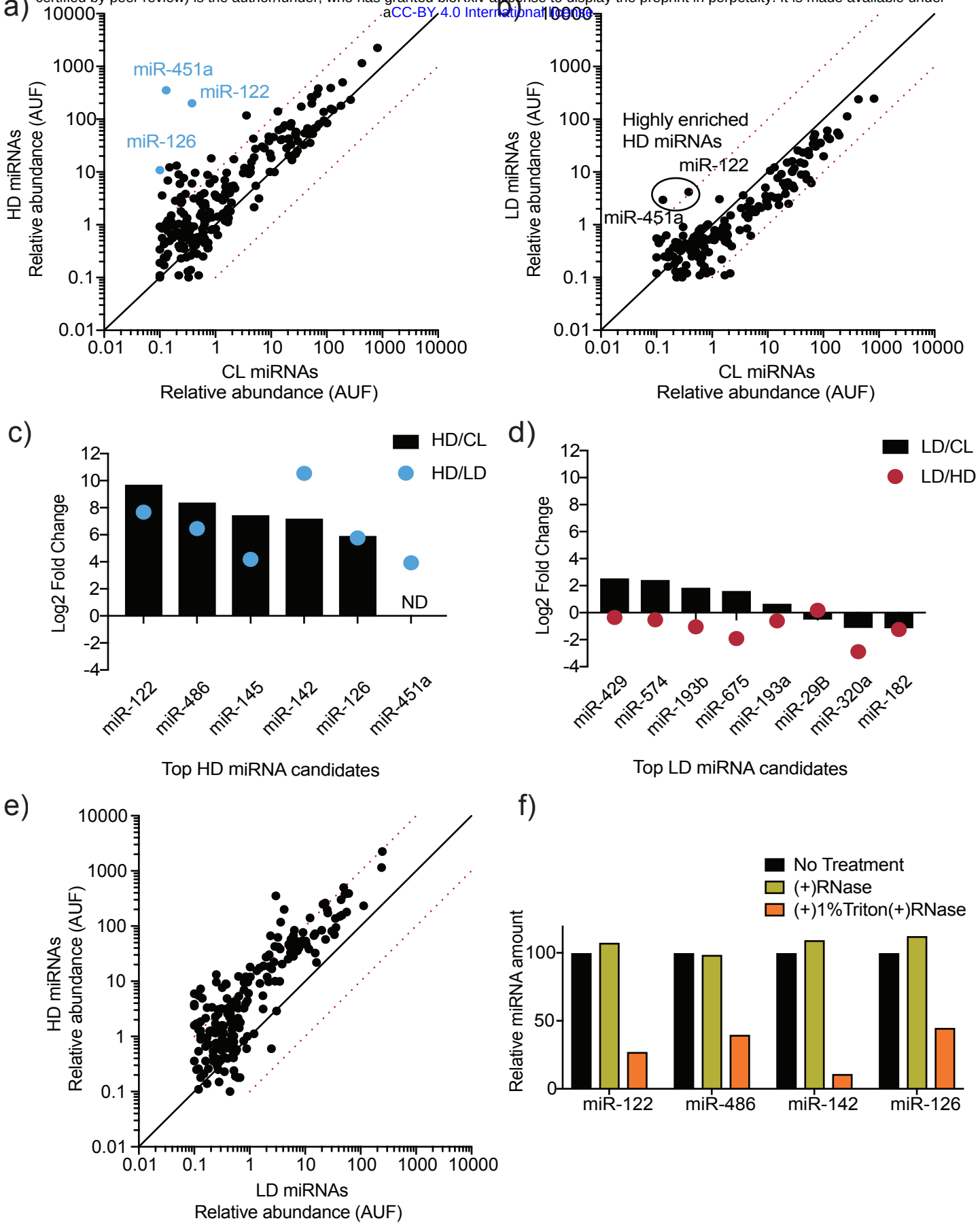


Figure 3. MicroRNA profiling of high buoyant density and low buoyant density EV sub-populations.

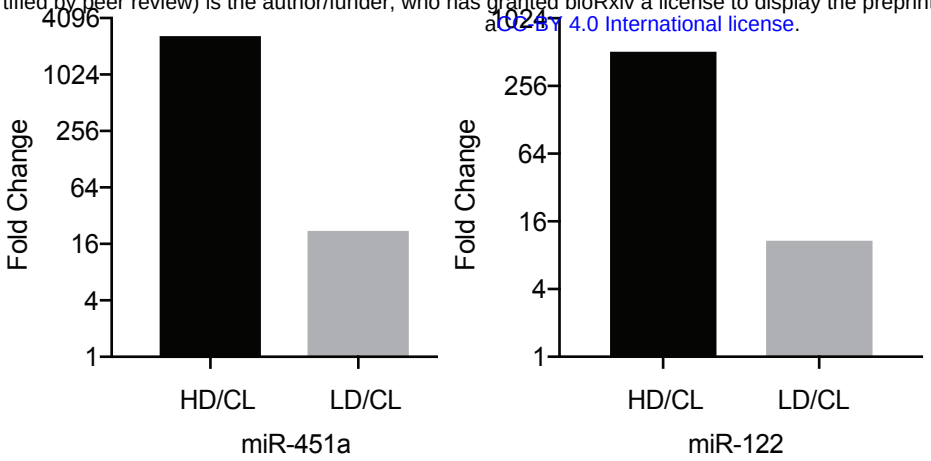


Figure 3 - figure supplement 1. MiR-122 and miR-451a are highly enriched in the HD sub-population.

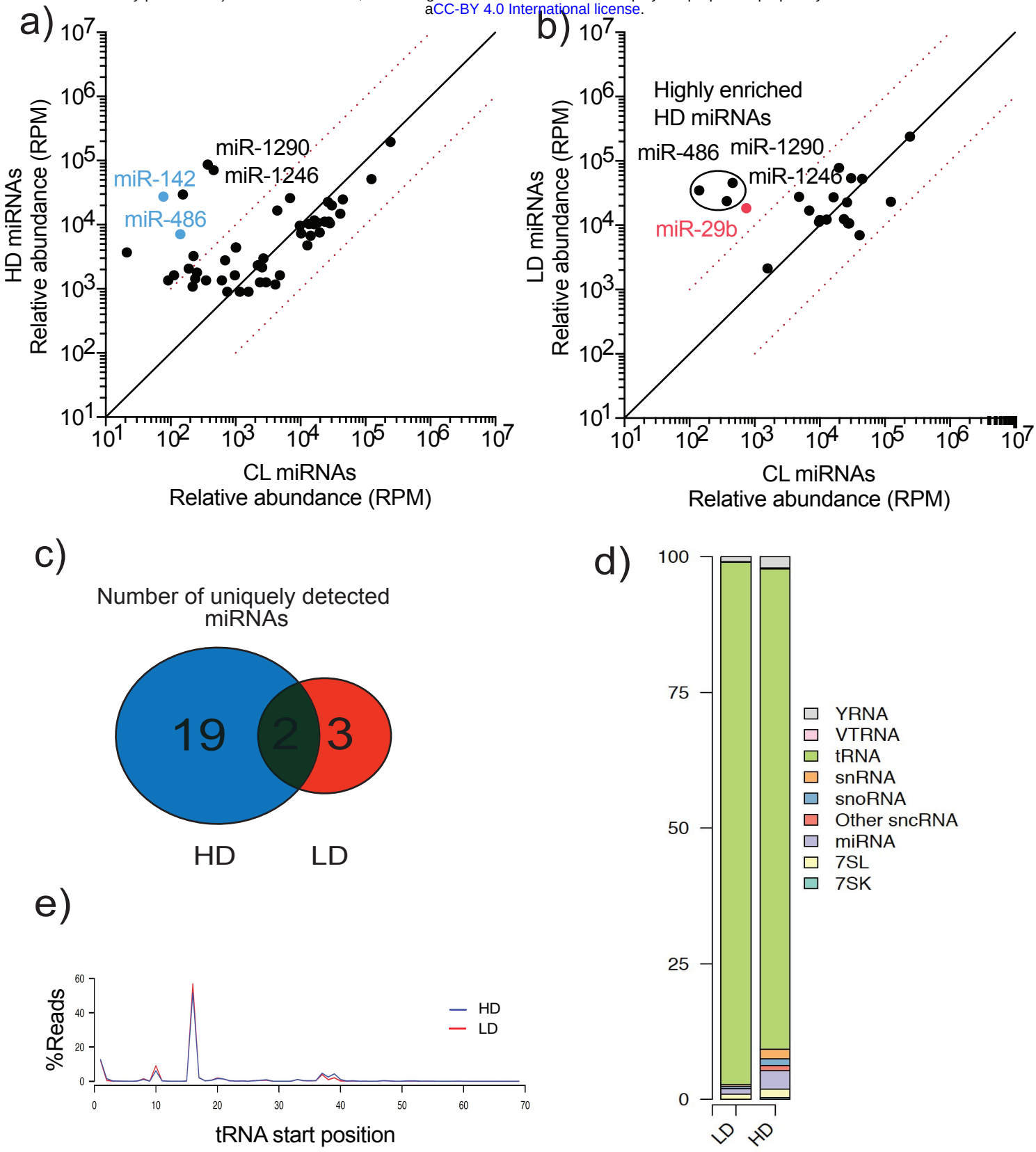
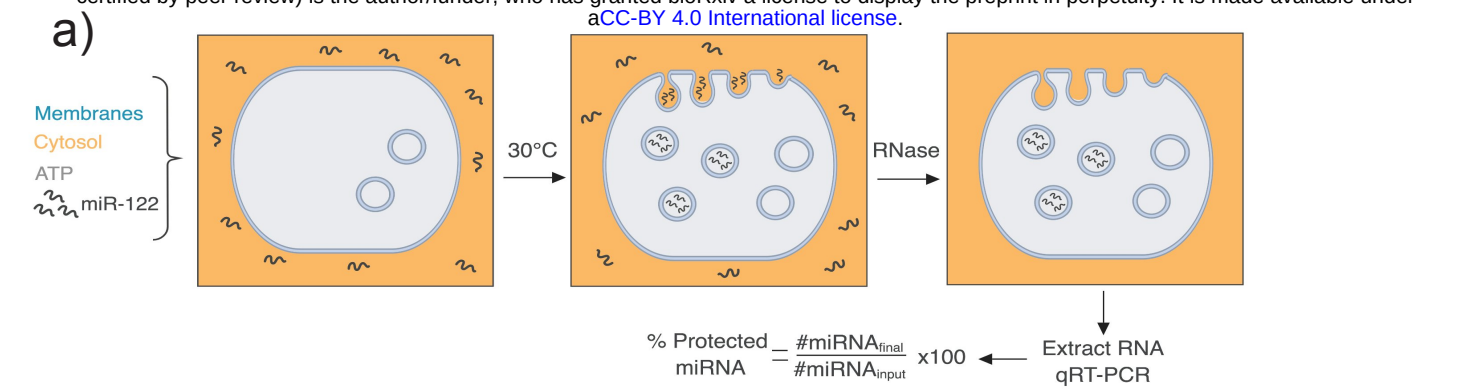
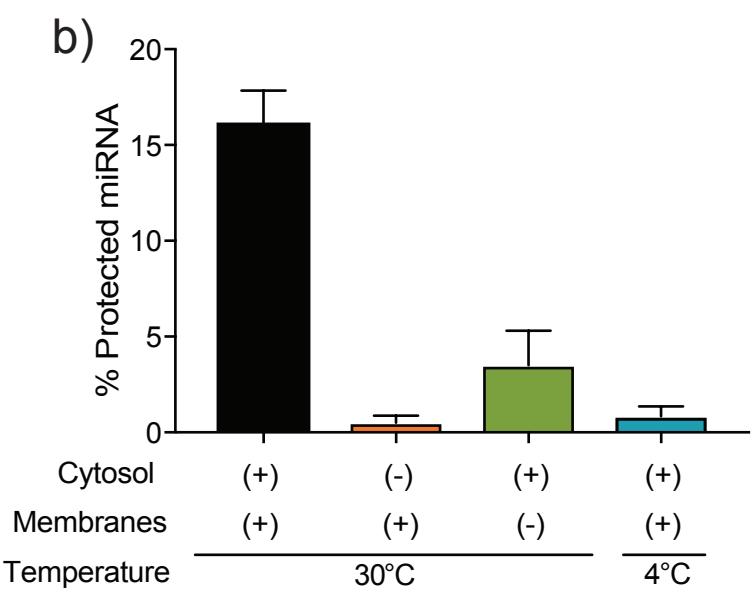
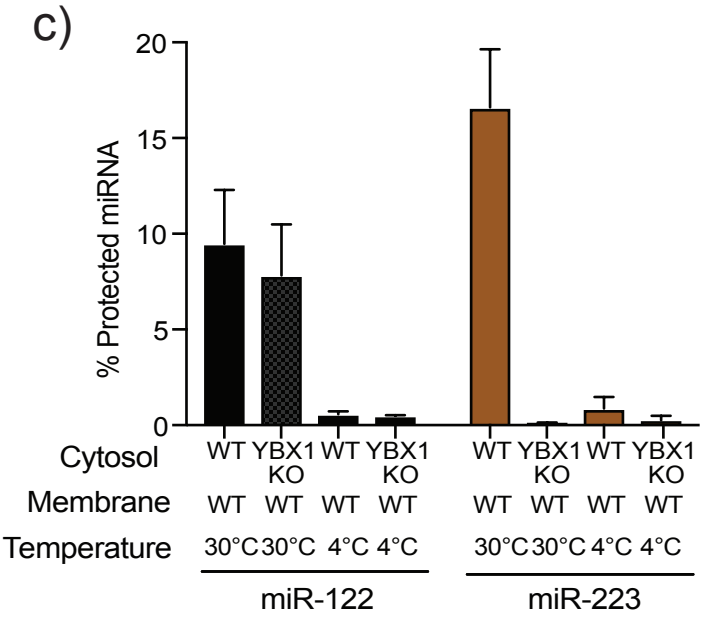


Figure 3 - figure supplement 2. TGIRT-sequencing of high buoyant density and low buoyant density EV sub-populations.







d)

Protein	# of unique peptides	Coverage
Nucleolin	13	23.5%
Lupus La	7	25.7%
Nucleophosmin1	7	35.0%
Poly (rC)-binding protein	4	14.6%
Small nuclear ribonucleoprotein Sm D1	3	25.2%

Figure 4. miR-122 packaging is recapitulated in a cell-free reaction.

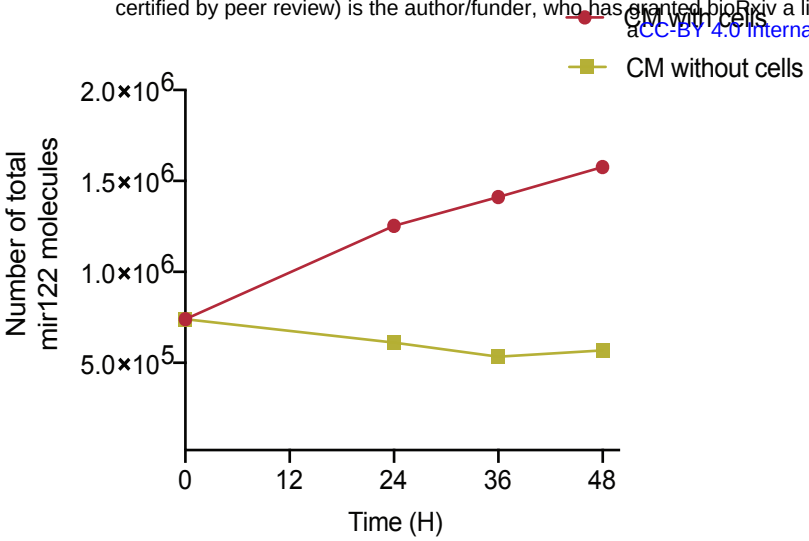


Figure 4 - figure supplement 1. MiR-122 accumulates in conditioned media over time.

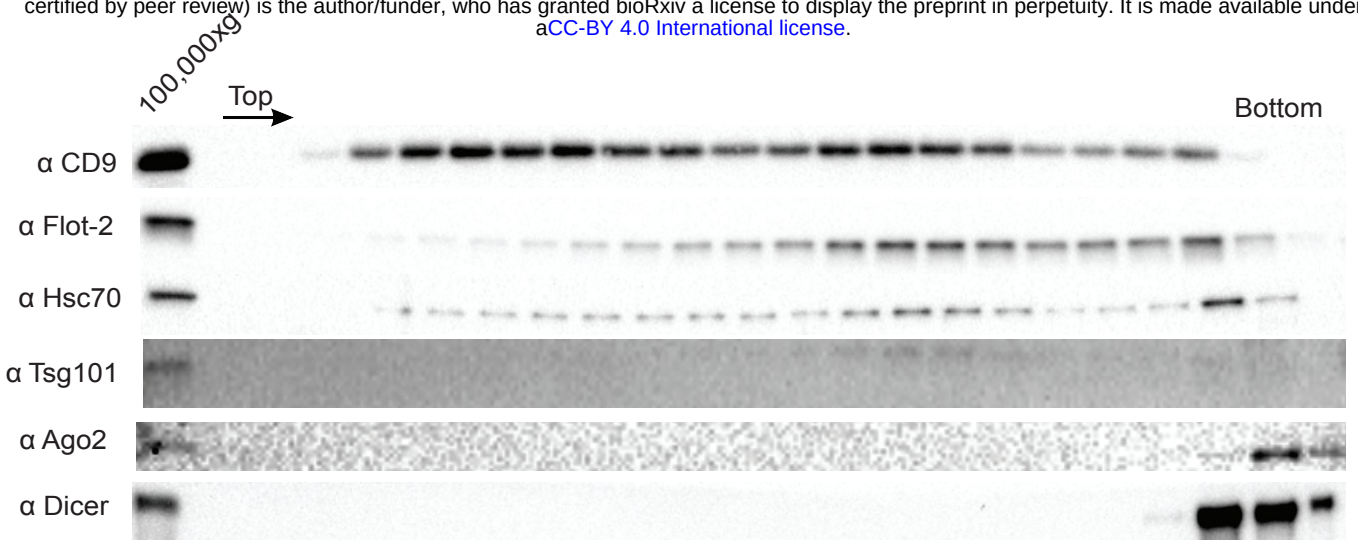


Figure 4 - figure supplement 2. Ago2 and Dicer are not EV-associated in MB-MDA-231 cells.

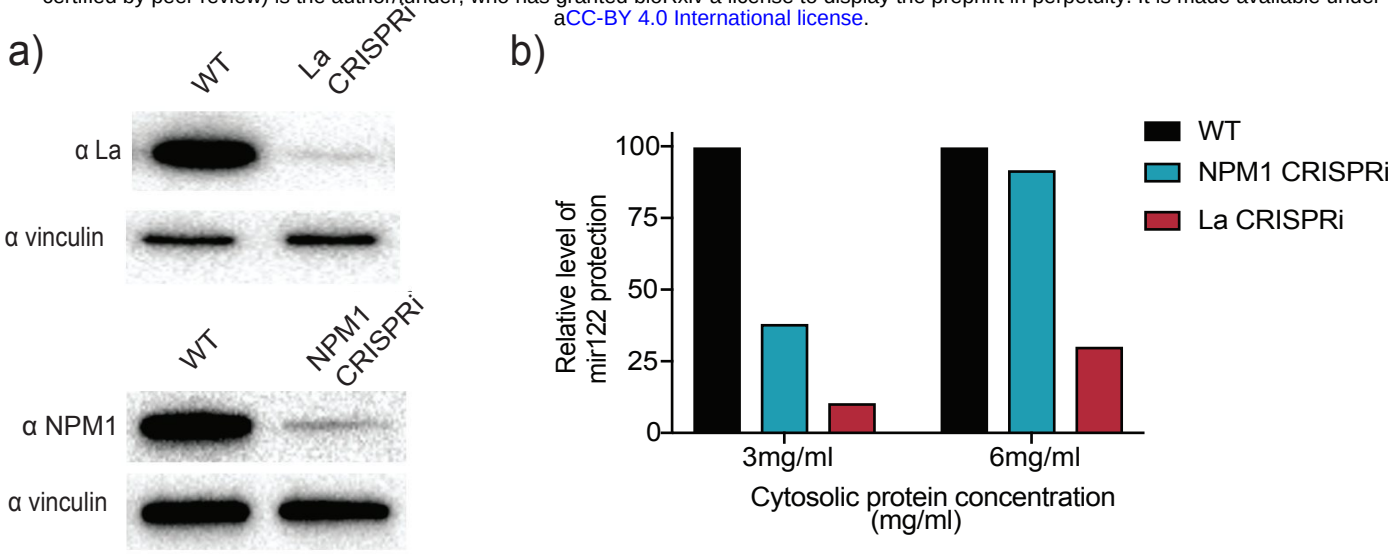


Figure 4 - figure supplement 3. CRISPRi efficiently depletes the RNA-binding protein candidates.

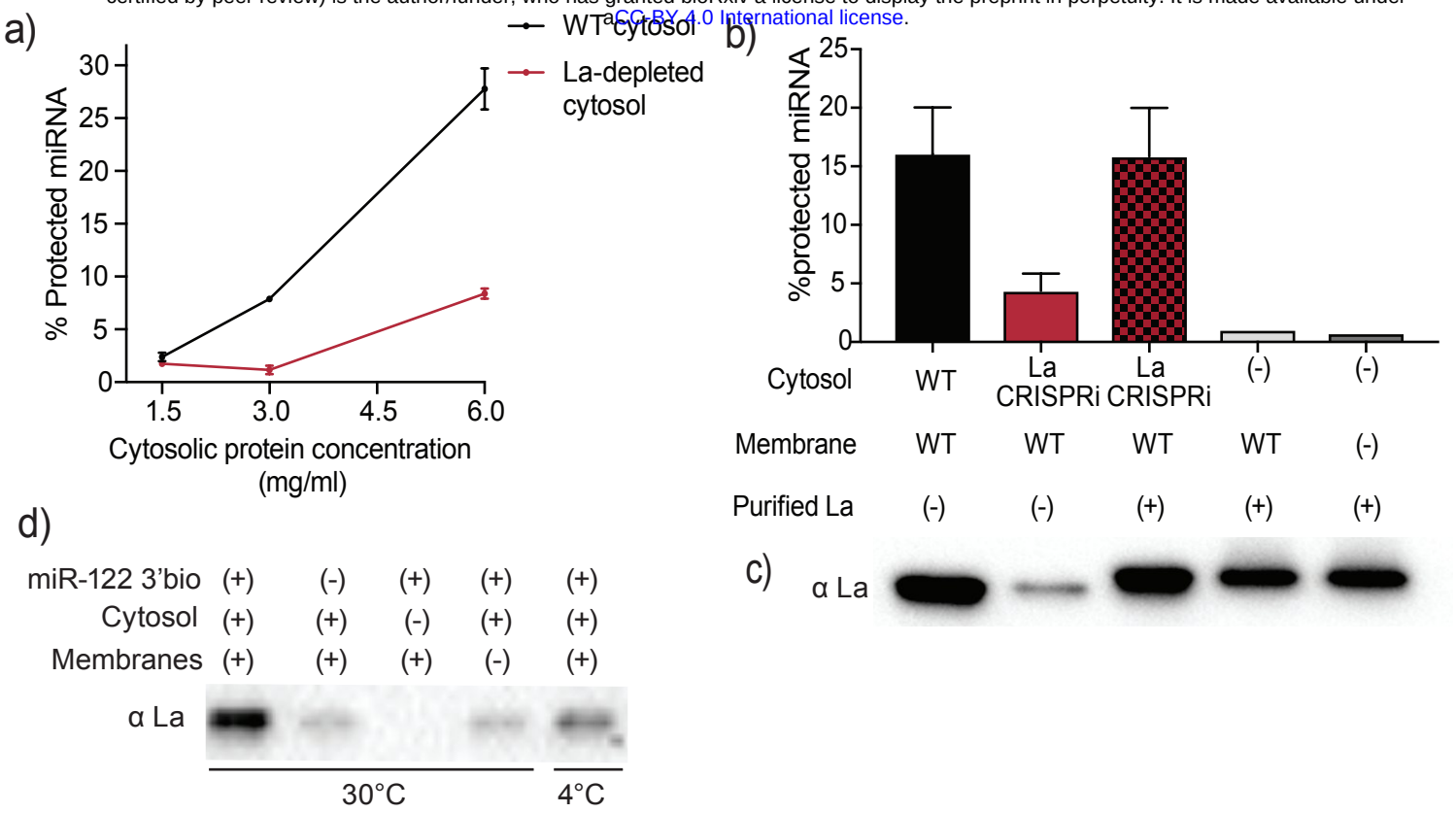


Figure 5. Sorting of miR-122 into EVs *in vitro* requires La.

a)

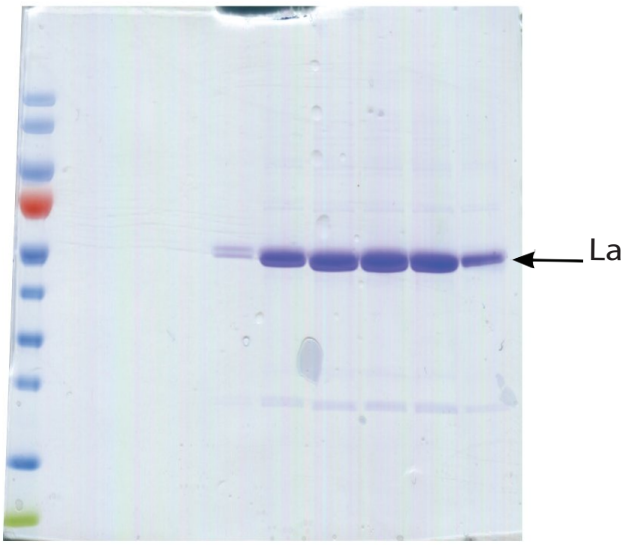


Figure 5 - figure supplement 1. Purification of heterologously expressed La.

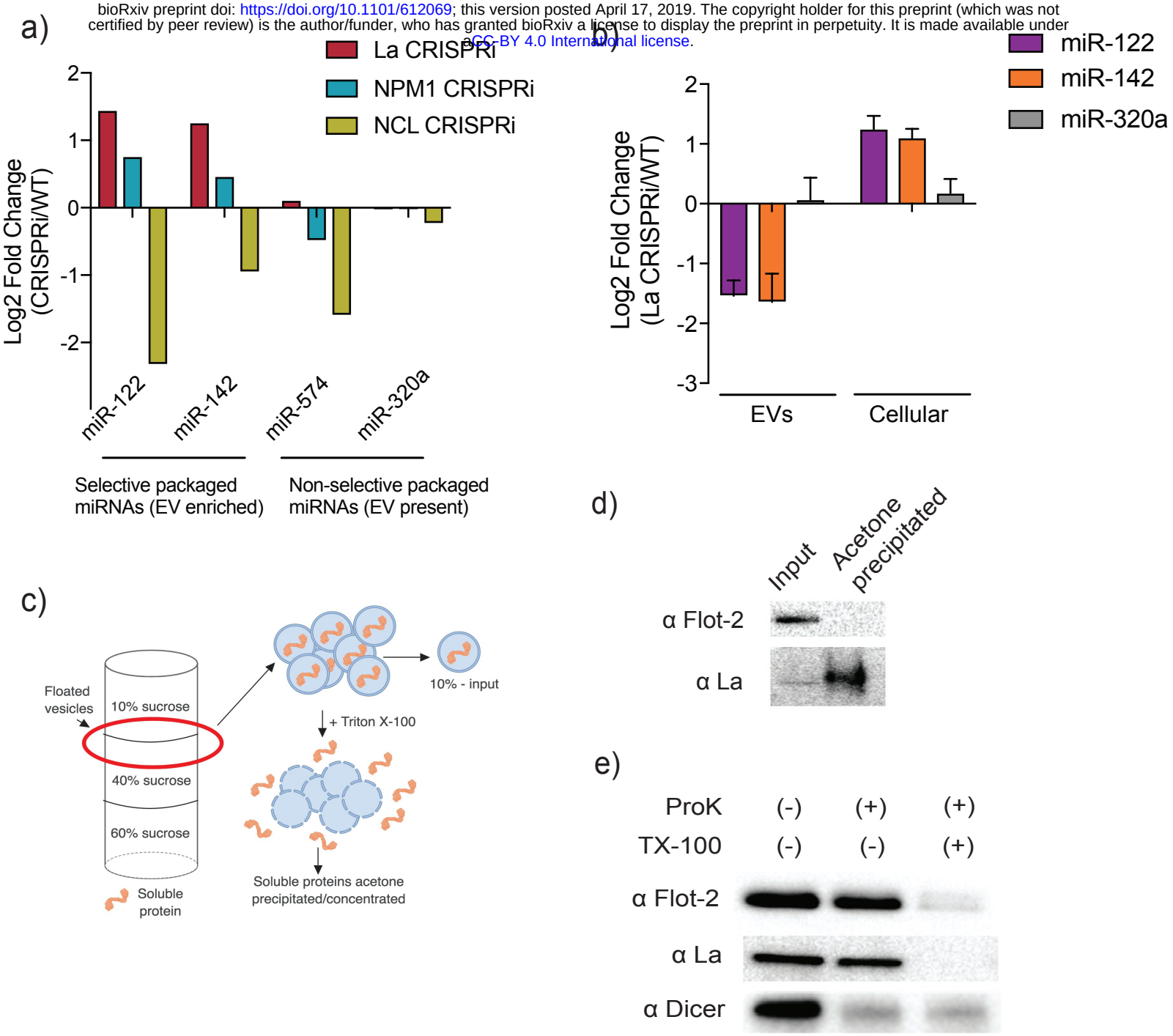


Figure 6. MiR-122 sorting into EVs requires La *in vivo.*

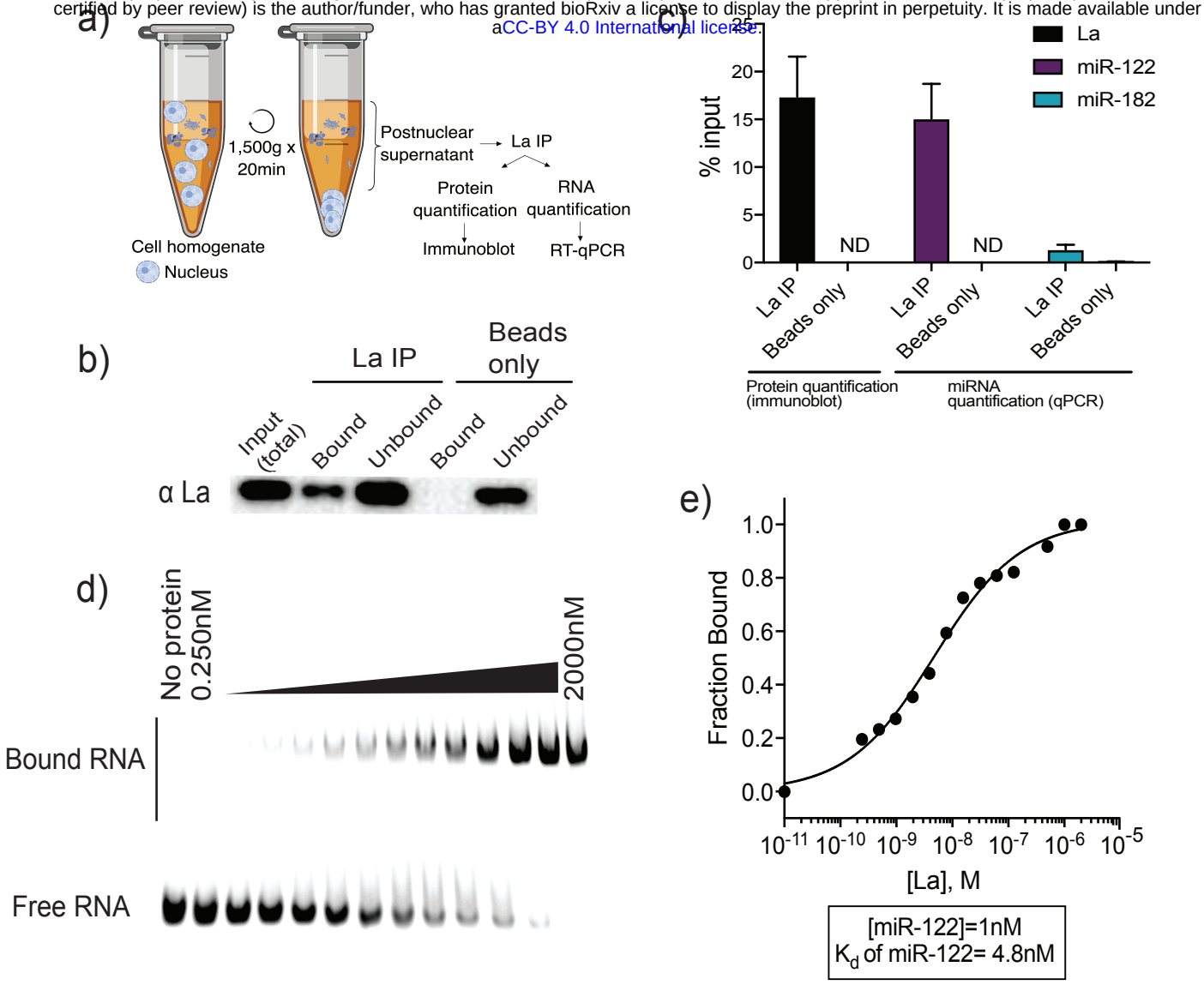


Figure 7: La interacts with miR-122 *in vitro* and *in vivo.*

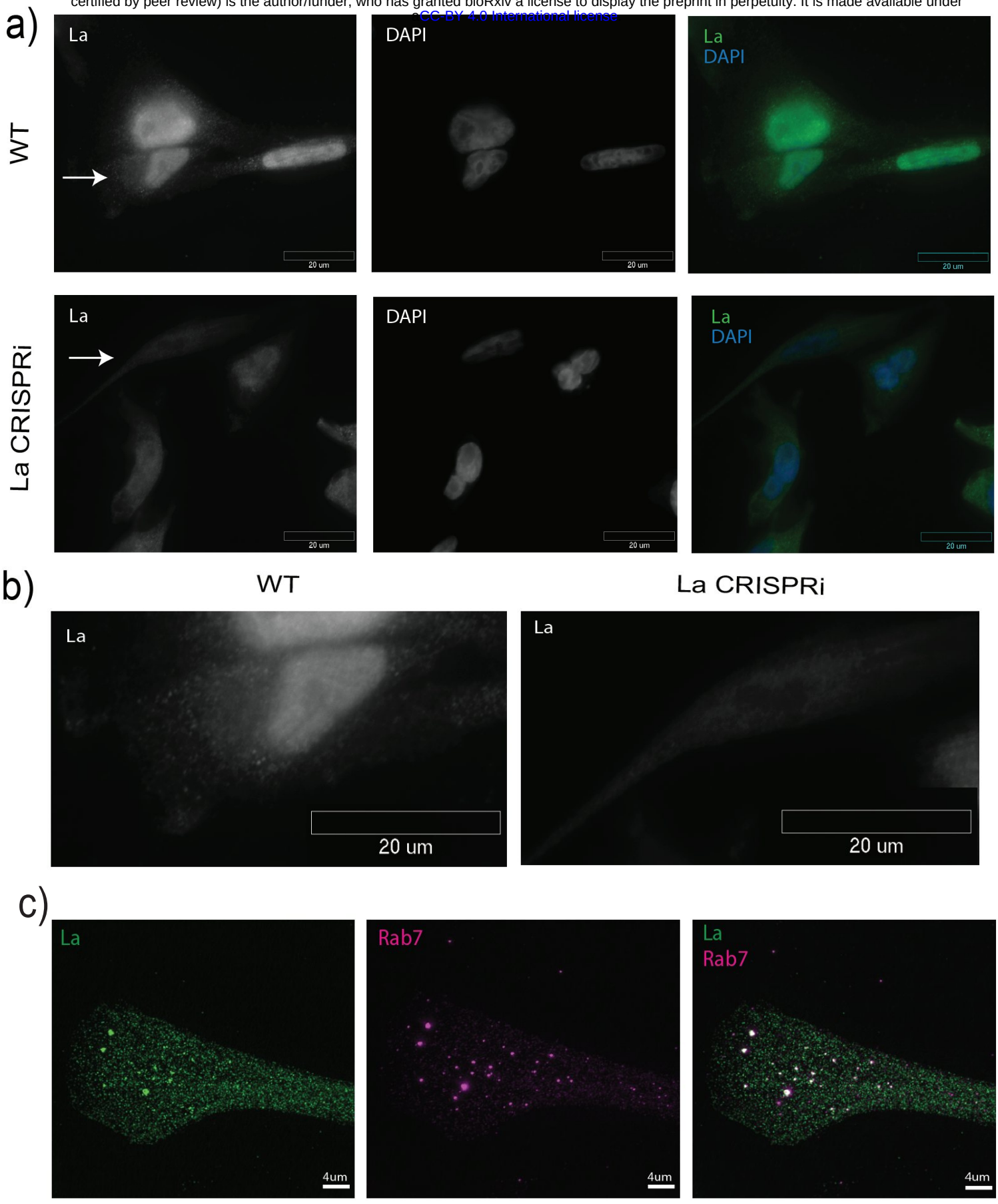


Figure 7 - figure supplement 1. Detection of nuclear and cytoplasmic La by immunofluorescence.

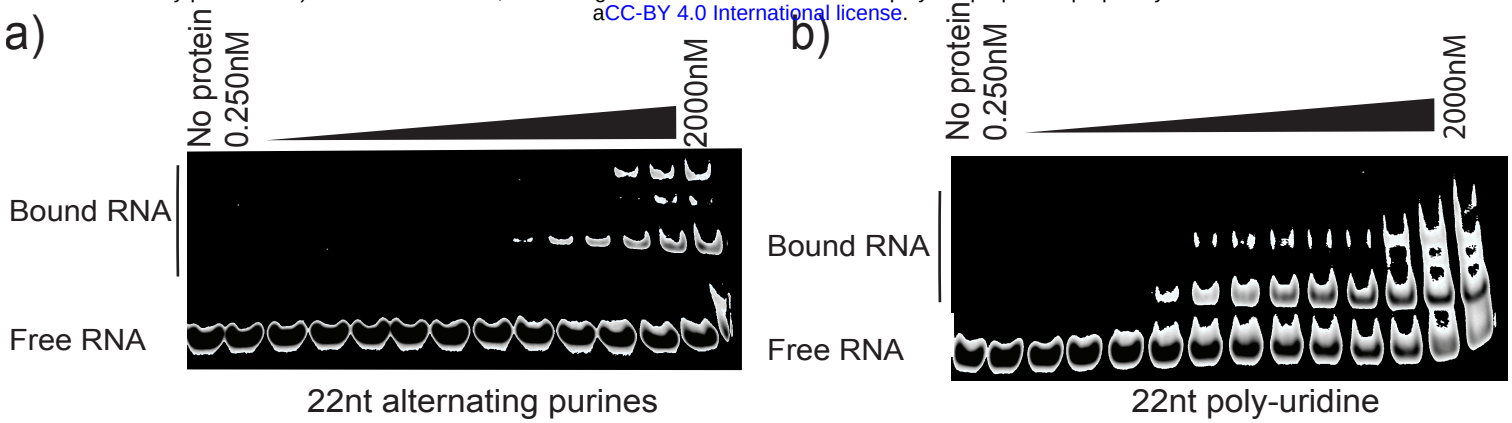


Figure 7 - figure supplement 2. Controls for La specificity during EMSA.

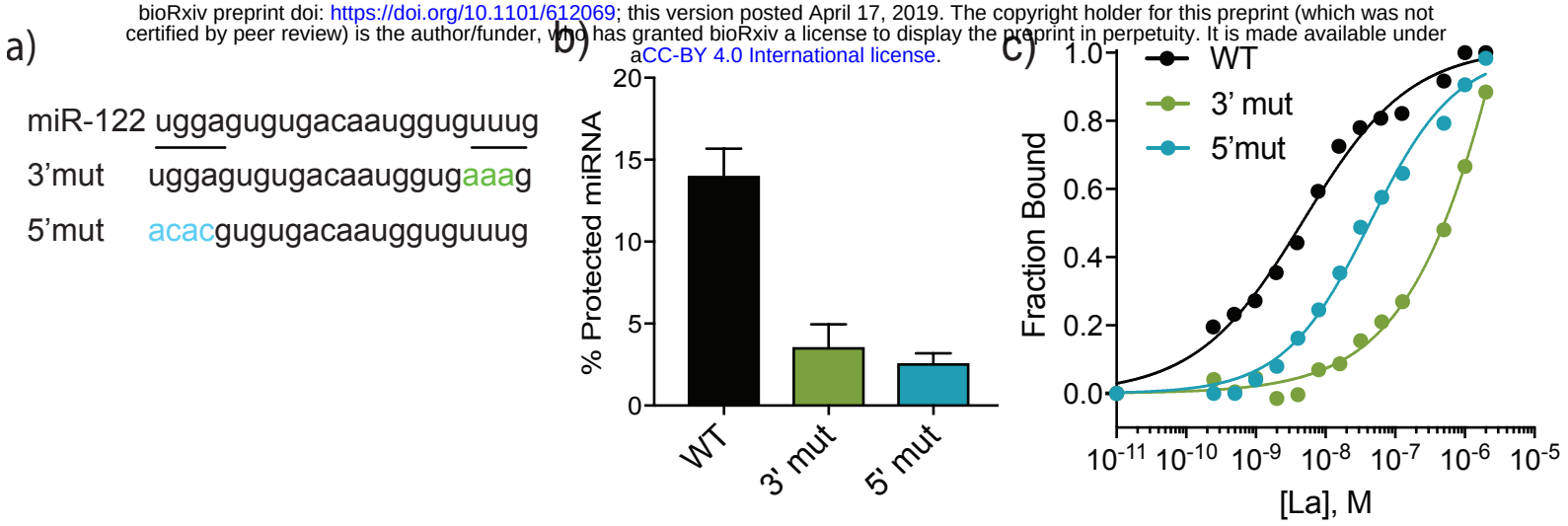


Figure 8. A bipartite motif in miR-122 is required for its packaging and interaction with La *in vitro*.

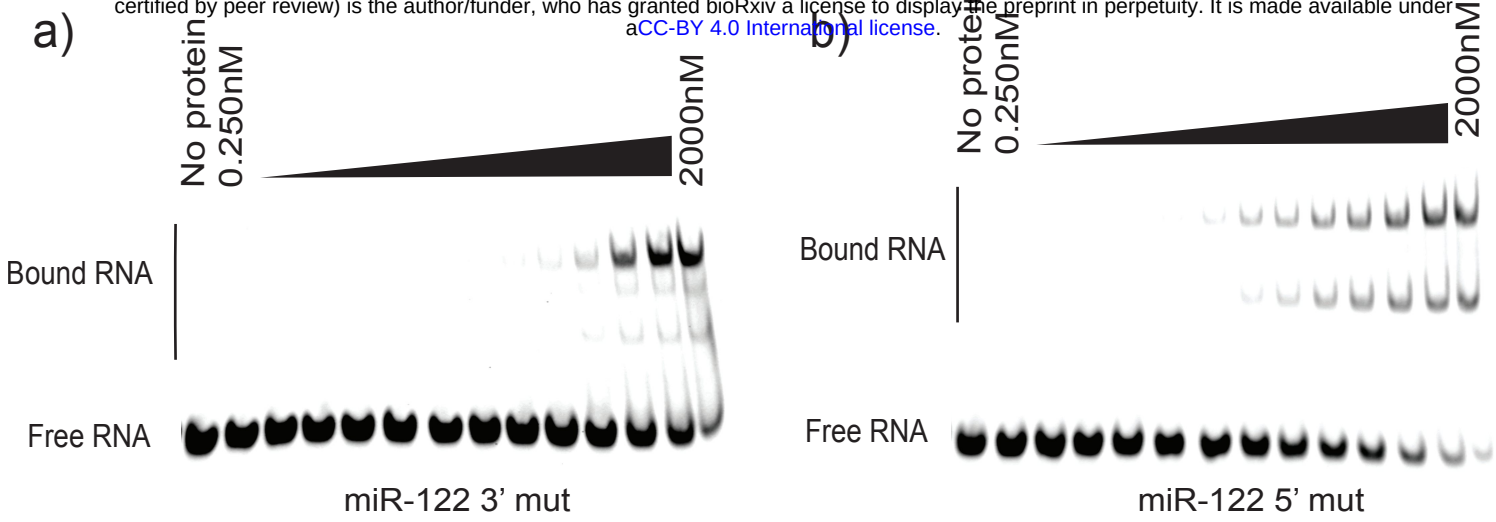


Figure 8 - figure supplement 1. EMSA for miR-122 mutated versions.

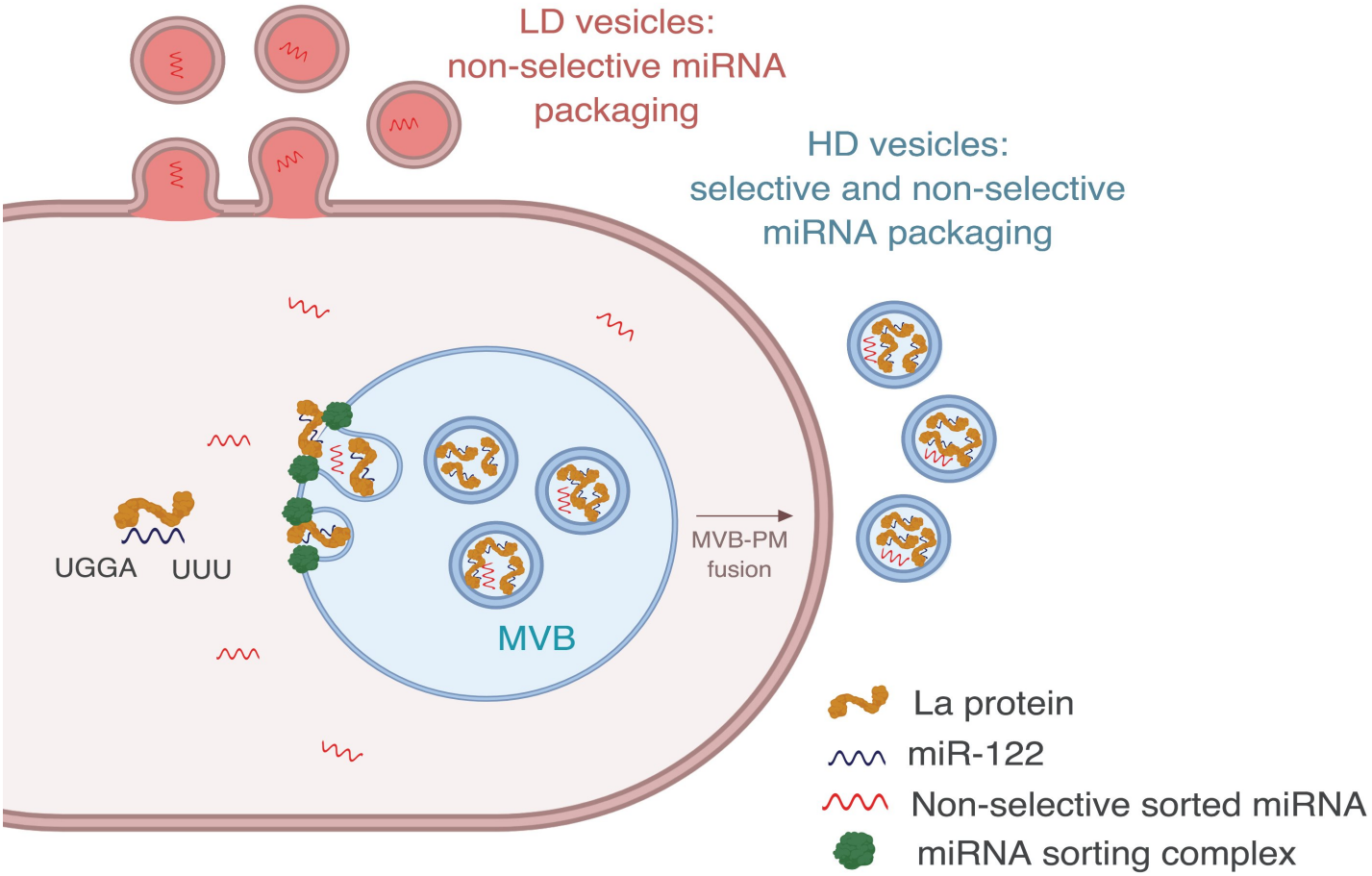


Figure 9: Diagram representing the current model of miRNA sorting into extracellular vesicles derived from MDA-MB-231 cells.