Small RNA Sequencing in Cells and Exosomes Identifies eQTLs and 14q32 as a Region of Active Export

Emily K. Tsang,*,† Nathan S. Abell,†,‡ Xin Li,† Vanessa Anaya,† Konrad J. Karczewski,*
David A. Knowles,‡,§ Raymond G. Sierra,** Kevin S. Smith,† and Stephen B. Montgomery†,‡,1
*Biomedical Informatics Program, †Department of Pathology, ‡Department of Genetics, and §Department of Radiology, Stanford University, California 94305, and **Stanford PULSE Institute, SLAC National Accelerator Lab, Menlo Park, California 94025

ABSTRACT Exosomes are small extracellular vesicles that carry heterogeneous cargo, including RNA, between cells. Increasing evidence suggests that exosomes are important mediators of intercellular communication and biomarkers of disease. Despite this, the variability of exosomal RNA between individuals has not been well quantified. To assess this variability, we sequenced the small RNA of cells and exosomes from a 17-member family. Across individuals, we show that selective export of miRNAs occurs not only at the level of specific transcripts, but that a cluster of 74 mature miRNAs on chromosome 14q32 is massively exported in exosomes while mostly absent from cells. We also observe more interindividual variability between exosomal samples than between cellular ones and identify four miRNA expression quantitative trait loci shared between cells and exosomes. Our findings indicate that genomically colocated miRNAs can be exported together and highlight the variability in exosomal miRNA levels between individuals as relevant for exosome use as diagnostics.

Most cell types, including those grown in culture, produce and release exosomes (Raposo and Stoorvogel 2013), and compelling evidence that exosomes mediate intercellular communication exists in multiple contexts including immune modulation (Zitvogel *et al.* 1998; Kim *et al.* 2005; Segura 2005; Admyre *et al.* 2007; Alexander *et al.* 2015), cancer proliferation (Meckes *et al.* 2010; Hood *et al.* 2011; Peinado *et al.* 2012; Tadokoro *et al.* 2013; Boelens *et al.* 2014; Costa-Silva *et al.* 2015; Fong *et al.* 2015), and neuronal activity (Frühbeis *et al.* 2013; Chivet *et al.* 2014). These extracellular vesicles range from 30 to 100 nm in diameter and are secreted by exocytosis into biological fluids and culture medium when an endosome harboring multiple vesicles fuses with the plasma membrane (Harding *et al.* 1983; Pan *et al.* 1985). They enable

communication by shuttling heterogeneous cargo, including DNA, RNA, proteins, and lipids, from their cell of origin to targeted recipients (Harding *et al.* 1983; Valadi *et al.* 2007; Mittelbrunn *et al.* 2011; Thakur *et al.* 2014). For instance, *in vitro* studies have revealed that exosomal miRNAs can repress known target genes in recipient cells (Hergenreider *et al.* 2012; Montecalvo *et al.* 2012; Umezu *et al.* 2013; van Balkom *et al.* 2013).

Studies in diverse cell types have demonstrated that the RNA content of exosomes differs from that of their parent cells (Nolte-'t Hoen *et al.* 2012; Villarroya-Beltri *et al.* 2013; Koppers-Lalic *et al.* 2014; Squadrito *et al.* 2014). However, these studies compared small numbers of samples, which limited their power to detect differentially expressed genes. Modest sample sizes have also precluded the investigation of interindividual variability in exosome cargo, which may include genetically driven differences. We can identify such individual-specific differences by identifying genetic variants associated with gene expression, known as expression quantitative trait loci (eQTLs). Most protein-coding genes, as well as some miRNAs, have at least one known eQTL in cells (Borel *et al.* 2011; Lappalainen *et al.* 2013; Battle *et al.* 2014; Huan *et al.* 2015). However, it remains unknown whether exosomes mirror cellular expression differences between individuals, and understanding interindividual variability is crucial as the field moves toward using exosomes in diagnostics.

Most eQTL studies rely on cohorts of ≥60 unrelated individuals to detect significant effects, but we have previously shown that we can leverage large nuclear families to identify eQTLs from smaller samples (Li *et al.* 2014). Here, we sequenced the small RNA from the lymphoblastoid cell lines (LCLs) and associated exosomes of a 17-member family spanning three generations. This represents, to our knowledge, the largest set of paired cell and exosome samples analyzed to date. We aimed to comprehensively quantify differences in small RNA between cells and exosomes, to assess interindividual variability, and to establish whether genetic variants influence exosome content. While it has previously been shown that miRNAs are differentially present in cells and exosomes, we discovered that exosomes can export entire genomic clusters of miRNAs. Furthermore, by using publicly available whole genome sequences, we performed the first eQTL study in exosomes to elucidate the impact of genetic variation on the small RNA in exosomes.

MATERIALS AND METHODS

Samples and genotype data
CEPH/UTAH family EBV-transformed peripheral blood B LCLs (catalog no. XC01463) were purchased from the Coriell Institute for Medical Research. The samples are from a complete three-generation pedigree that includes four grandparents, two parents, and 11 children. Variant calls from whole genome sequencing data for the 17 individuals were obtained from Complete Genomics (Analysis Pipeline v.2.0.0).

Cell culture and exosome isolation
LCLs were grown in RPMI 1640 supplemented with 10% fetal calf serum (Gibco, Life Technologies) and 1× penicillin/streptomycin (Life Technologies) in humidified 5% CO_2. Cell cultures were initiated at densities of $500–750 \times 10^5$ cells/ml and allowed to grow to 150×10^6 cells, achieving a maximum density of 2×10^6 cells/ml at collection. Fetal calf serum was depleted of bovine exosomes before use by overnight centrifugation at $120,000 \times g$. The exosome isolation procedure is depicted in Figure 1A and described in detail here. LCLs were pelleted at $300 \times g$ for 10 min and saved for cell miRNA isolation. The remaining growth medium was centrifuged at $16,500 \times g$ for 25 min and then filtered through a 200 nm Acrodisc (PAL Corp., Ann Arbor, MI). The filtered supernatant was centrifuged at $120,000 \times g$ for 70 min to pellet exosomes (Ti45 rotor). Pelleted exosomes were washed 1× in cold PBS and centrifuged again at $120,000 \times g$ for 70 min (TLA 100.3 rotor). All centrifugations were performed at 4°.

Exosome characterization
Exosomes isolated from LCLs were characterized by Western blot, electron microscopy, and NanoSight nanoparticle tracking analysis (Malvern, Worcestershire, UK). For Western blotting, ∼50 μg total cell and exosome proteins were denatured with 1× sample buffer, reduced with β-mercaptoethanol, and boiled for 5 min. Proteins were resolved on 10% PAGE gels, blotted to nitrocellulose, and probed with antibodies recognizing HSP70 and calnexin. The Stanford University Cell Sciences Imaging Facility performed transmission electron microscopy of isolated intact exosomes. Nanoparticle tracking analysis software (update 2.3) was used to determine the average size of isolated exosomes using a NanoSight LM10-HS at the Stanford PULSE Institute.

Small RNA isolation and cDNA library construction
Small RNA from TRIzol-lysed LCLs and exosomes was isolated from 1 μg total RNA and analyzed for integrity using a Bioanalyzer 2100 Total RNA 6000 Nano Kit. cDNA libraries were produced using the TruSeq Small RNA Sample Prep Kit (Illumina, San Diego, CA) per manufacturer's protocol, with the following modifications: (1) final PCR amplifications were 16 cycles for exosome cDNA and 14 cycles for cell cDNA, (2) instead of gel purification using PAGE, PCR-amplified small RNA libraries were gel-purified with 3% NuSieve GTG low-melt agarose (Lonza, Rockland, NY). Target bands of 147–157 nt, containing an adapter sequence of 125 nt, were excised and column purified (Qiagen, Germantown, MD). All cDNA libraries were indexed with Illumina adapters.

Transcriptome sequencing
Small RNA was sequenced on two lanes (one with cell samples and the other with exosome samples) of a single Illumina HiSeq 2000 flow cell as 36 bp single-end reads. This sequencing setup was chosen to maximize our power for eQTL discovery but meant that lane effects were confounded with the effect of interest for differential expression. Therefore, we resequenced the libraries to lower depth on two runs of an Illumina MiSeq, where each run had both the cells and exosome samples for half the individuals.

Quantification and cluster generation
Transcript quantification steps are depicted in Supplemental Material, Figure S1 and described here. Sequencing adapters were trimmed with Trimmomatic (v0.27) and were also provided to Novoalign (v3.00.02, www.novocraft.com), which handled incomplete adapter trimming while mapping. The reads were mapped in two steps where they were first mapped to the mature sequences in miRBase (release 20) (Kozomara and Griffiths-Jones 2014) and piRNABank (Sai Lakshmi and Agrawal 2007), then unmapped reads were aligned to the human reference genome (hg19). Novoalign was used for both mapping steps, with the parameters -l 18 -s 1 -a <adapter> -R 0 -r Random -t 0. The reads that mapped to the reference genome were assigned to transcripts annotated in Gencode (version 17), miRBase, or piRNA-Bank using HTSeq (v0.5.4p3) (Anders *et al.* 2014). The read counts from the first mapping step were added to those generated by HTSeq to get a single read count per transcript.

An miRNA cluster is defined as a set of miRNAs where every cluster member is within 10 kb of at least one other member. We identified all such clusters in the human genome using miRBase annotations. For each cluster, we summed the read counts of all mature miRNAs mapping to that cluster, excluding mature miRNAs that could have originated from multiple clusters. We used the resulting matrix for cluster-level differential expression. Notably, since many miRNAs do not cluster with any others, their mature and cluster counts are identical. As such, mature and cluster MA plots are visually similar, though not identical.

Differential expression analysis
For each sample, we computed the proportion of reads mapping to each class of RNA. We compared the small RNA composition of cells and exosomes by averaging the proportions of the 17 individuals for each compartment. We tested for a difference between the compartments with a paired two-sided *t*-test and corrected for multiple testing using the Bonferroni method. When comparing the quantities of different small RNA types, we considered miRNAs that were annotated in either Gencode or miRBase, but for all other analyses we only used the higher confidence set of miRNAs annotated in miRBase, which accounts for >99.5% of the reads mapped to miRNAs. Transcript- and cluster-level differential expression between cells and exosomes was performed separately for miRNA, miRNA clusters, and piRNA using DESeq2 (v1.6.3) (Love *et al.* 2014), which uses a negative-binomial distribution
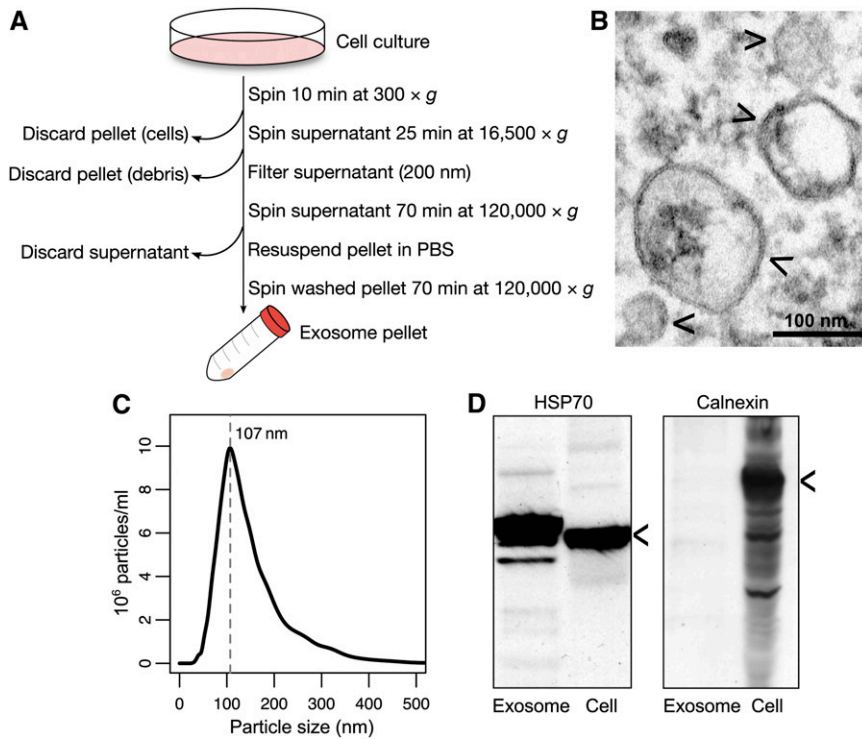
Figure 1 LCL exosome isolation procedure yields vesicles characteristic of exosomes. (A) Flow diagram of the exosome isolation procedure. All centrifugations were performed at 4°. (B) Transmission electron microscopy of isolated LCL exosomes. Four arrowheads denote isolated examples. Bar, 100 nm. (C) Example NanoSight tracing of LCL exosomes from a representative sample. For that sample, the maximal concentration of exosomes was at 107 nm diameter, as indicated by the dashed line. (D) Western blots of 50 μg total protein lysates from LCLs or their isolated exosomes hybridized with HSP70-specific (left) or calnexin-specific (right) antibodies. Arrowheads indicate expected bands at 70 and 90 kDa, respectively.

to directly model counts obtained from RNA sequencing. We modeled the mean expression as a function of whether the sample came from cells or exosomes and controlled the false discovery rate (FDR) at 1% with the Benjamini–Hochberg adjusted *p*-values (Benjamini and Hochberg 1995) reported by DESeq2. To confirm that the differential expression between cells and exosomes was not due to lane effects, we compared the initial HiSeq2000 sequencing data to the resequencing data (see *Transcriptome sequencing*) both with and without the sequencing run explicitly included in the DESeq2 model. Because the resequencing data had lower sequencing depth and therefore lower power to detect differential expression, we also compared with downsampling the initial data to 10%, which matched the depth of the resequencing data.

Replication of chromosome 14q32 miRNA cluster differential expression

We processed the raw reads counts from an exosome sequencing study in HeLa cells (Honegger *et al.* 2015) following the same procedure we used on our data to identify differentially expressed miRNA. We also downloaded the differentially expressed results produced by EdgeR from another sequencing study in B cell lines (Koppers-Lalic *et al.* 2014). Since the replication data sets used miRBase releases 18 and 19 whereas we used release 20, we confirmed that the annotations of the 14q32 cluster hairpins did not change between releases 18 and 19.

Motif enrichment

For cells and exosomes separately, we identified mature miRNA sequences that were significantly enriched in that compartment (differential expression FDR < 1%). Then, for the separate exosome-enriched and cell-enriched sets, we defined all other annotated miRNAs as background (exosome-enriched *vs.* all nonexosome-enriched and cell-enriched *vs.* all noncell-enriched, separately). To identify cell- or exosome-enriched miRNA motifs, we applied the MEME and DREME algorithms from the MEME Suite toolset (v4.11.1) (Bailey *et al.* 2009) with default parameters, using windows of 9–13 and 3–8 bp motifs, respectively.

Nontemplated nucleotide addition analysis

To profile nontemplated nucleotide additions (NTAs) from sequencing data, we remapped reads to miRBase (release 20) hairpins using Bowtie2 (v2.2.4) (Langmead and Salzberg 2012) in local alignment mode with sensitive parameters (-D 50 -R 5 -N 0 -L 10 -i C,1,0--score-min = C,32,0), requiring a minimum alignment of 16 bp, without penalizing soft clipping. We extracted 5′- and 3′-NTAs separately from the resulting BAM files using the CIGAR string. For every sample, we counted the number of occurrences of each NTA (*e.g.*, "-AG") for each miRNA hairpin (*e.g.*, "hsa-miR-21"), extracted counts for each 1- or 2- bp NTA, and computed the percentage of all miRBase-aligned reads that possessed each NTA. To further analyze a given NTA or set of NTAs, for each sample we computed the percentage of all reads with that NTA originating from each miRNA hairpin.

eQTL mapping

From the variants identified in the Complete Genomics whole genome sequencing data, we removed any variants with Mendelian inconsistencies and used the haplotype blocks we had previously inferred for the 11 children (Li *et al.* 2014). Briefly, the haplotype blocks were defined as segments between any two recombination points in the children. For each child, we then determined which of the two possible maternal and two possible paternal haplotypes that child received.

To map miRNA and piRNA *cis*-eQTLs, we used the linear model described in Li *et al.* (2014) to test for an association between a transcript's expression level in the children and the parental haplotypes they inherited at the transcript's locus:

$$\log(Y_i + 1) \sim \mu + \beta_p \, p_i + \beta_m \, m_i$$

where Y_i is the DESeq2-normalized expression in child i, μ is the intercept, and p_i and m_i are the paternal and maternal haplotypes (encoded as 0 or 1) of the child at the locus containing the transcript being tested. The coefficients β_p and β_m represent the effect sizes of
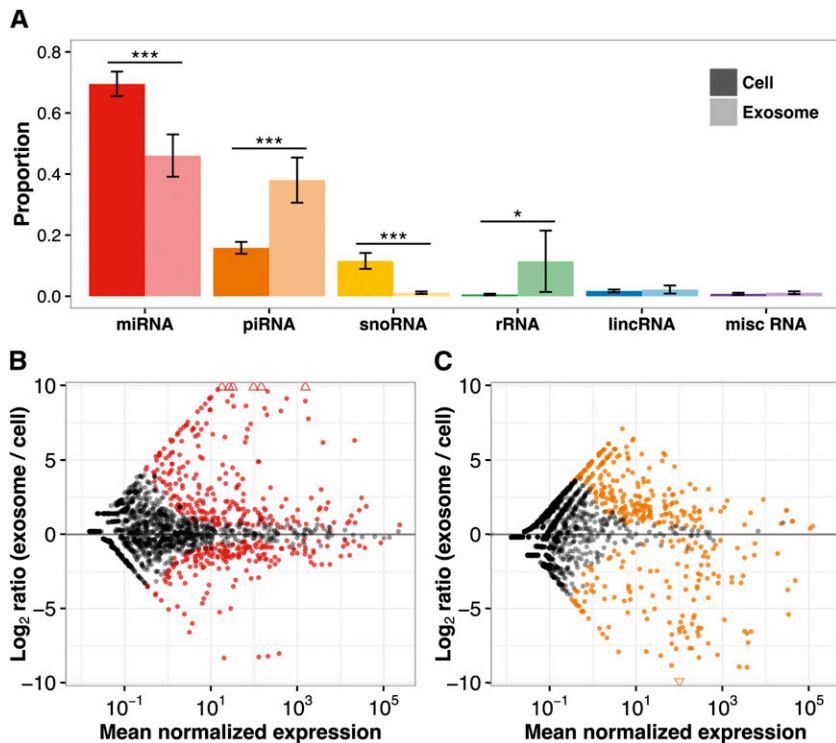
Figure 2 Cells and exosomes differ in their small RNA profiles. (A) Small RNA composition of cells and exosomes averaged over the 17 individuals. Error bars show the SD. Paired two-sided *t*-tests were used to compare the cell and exosome proportions for each miRNA type and the *p*-values for the six tests were corrected by the Bonferroni method. Asterisks denote the significance of the corrected *p*-values: *** $p < 1 \times 10^{-7}$; * $p = 0.002$. (B and C) MA plots for miRNA (B) and piRNA (C) where each point represents one transcript. Above the center horizontal line are transcripts that were relatively more abundant in exosomes and below are the ones that were present in relatively higher quantities in cells. The transcripts that were significantly differentially expressed at a 1% FDR are colored. Triangles represent points that fall outside the plotted area. lincRNA, long intergenic noncoding RNA; miRNA, micro RNA; misc. RNA, miscellaneous other RNA; piRNA, Piwi-interacting RNA; rRNA, ribosomal RNA; snoRNA, small nucleolar RNA.

the parental haplotypes. Each association was tested separately in cells and exosomes and the model fit (R^2) was used to obtain *p*-values for the associations. Only transcripts with a median normalized expression of at least 10 in the given compartment were tested. Transcripts mapping to multiple loci, as well as ones containing variants (which could bias mapping given the short read length), were also excluded.

The Pearson correlation of the effect sizes in cells and exosomes was used to assess the overall sharing of genetic effects between cells and exosomes. To identify specific instances of shared miRNA eQTLs, we intersected the cell and exosome miRNA eQTLs discovered at a 20% FDR.

Accounting for nongenetic factors in miRNA eQTL discovery

Several factors were assessed to rule out likely nongenetic explanations for the shared miRNA eQTLs to increase our confidence that we were detecting true genetic effects. We confirmed that the haplotypes of the candidate-shared miRNA eQTLs were not confounded with sex and were not explained by mapping differences. Spurious eQTLs driven by different mapping rates for the two haplotypes would be detectable by differential quantities of k-mers in the unmapped reads. Therefore, for each sample, 10-mers in the unmapped reads were counted with Jellyfish (v2.1.4) (Marçais and Kingsford 2011) and normalized by the number of sequenced reads from that sample. We verified that none of the candidate-shared eQTLs had 10-mer counts that differed significantly between the haplotypes in the opposing direction of the eQTL effect (one-sided *p*-value ≤ 0.05, Wilcoxon rank sum test). Finally, we checked whether the haplotypes of shared miRNA eQTLs were associated with many miRNAs genome-wide in both cells and exosomes. An excess of associations may indicate the haplotype is correlated with a technical factor influencing the expression of many miRNAs. We ranked haplotypes by their number of significantly associated transcripts (nominal *p*-value ≤ 0.05) to assess which haplotypes are most likely to correlate with a technical factor. In our data, we noticed that the haplotype containing the hsa-miR-151a products is ranked 13th

among 697 haplotypes in terms of the number of miRNAs genome-wide that it is associated with in cells (at a nominal significance level of 0.05). It is possible that this haplotype's segregation pattern among the children is correlated with a technical factor influencing the expression levels of many miRNAs, but we think this is unlikely since the eQTL replicates both in exosomes and in a different cohort.

Data availability

RNA sequencing fastqs and raw read counts are available at GEO with the accession number GSE74759. File S1 includes the legends for Table S1, Table S2, and Table S3, which include read mapping statistics, differential expression results, and eQTL data and results, respectively.

RESULTS

Isolation and characterization of exosomes in LCLs

We isolated exosomes following the procedure outlined in Figure 1A. To validate this protocol, we subsequently assayed the size and protein content of extracted exosomes. We first imaged exosome pellets by transmission electron microscopy and confirmed that we isolated vesicles with bilayered membranes of varying sizes, averaging roughly 100 nm (Figure 1B). We next applied nanoparticle tracking to confirm the vesicle sizes assessed by microscopy and found that most vesicles had a diameter just above 100 nm (Figure 1C). Finally, we made total protein lysates from both exosomes and parent cells and probed for enriched proteins by Western blotting. We detected HSP70 in lysate fractions of both cells and exosomes, but only resolved the endoplasmic reticulum-specific protein calnexin in the cell fraction (Figure 1D), which agreed with previous biochemical characterizations of exosomes (Baietti *et al.* 2012).

Cells and exosomes differ in small RNA composition

We sequenced small RNA libraries for cells and exosomes of the 17 family members to an average depth of 5.9 million reads after
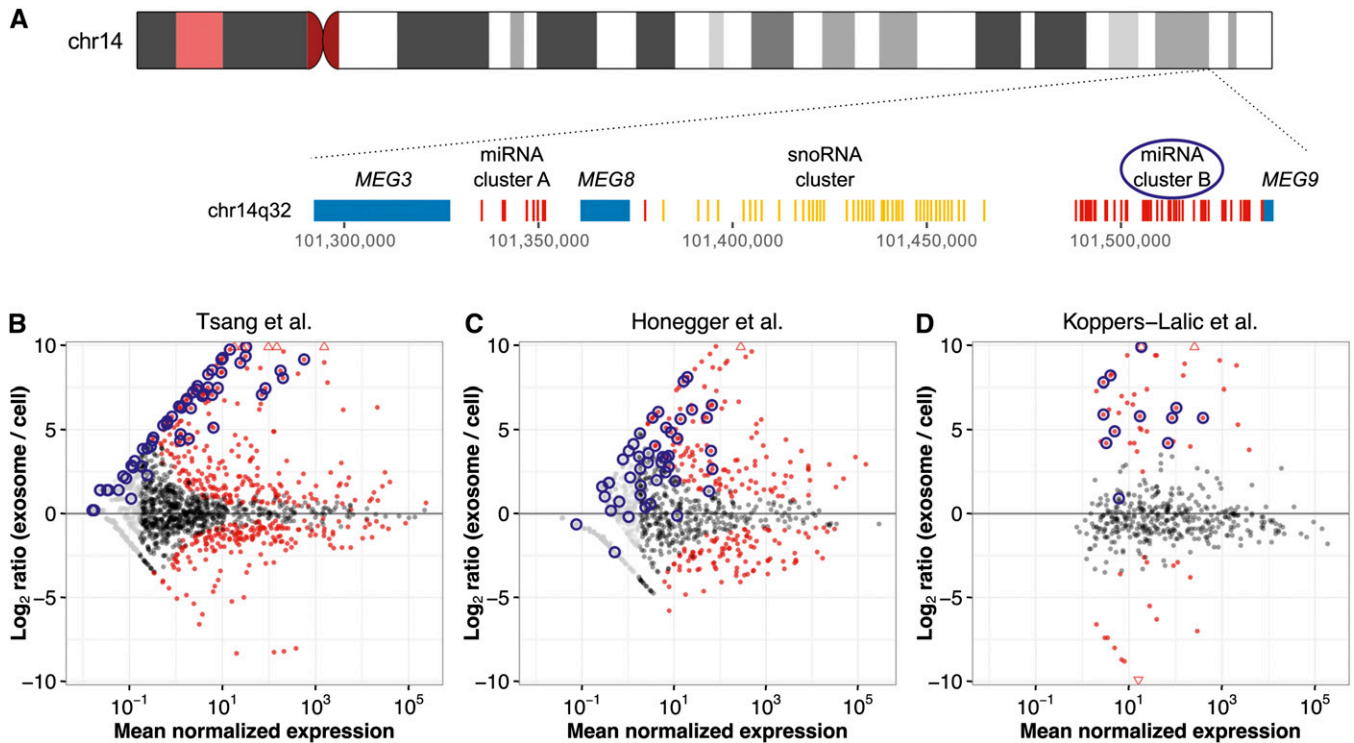
Figure 3 A large miRNA cluster on chromosome 14q32 is exported in exosomes. (A) Diagram of the 14q32 locus, which contains two miRNA clusters denoted as cluster A and cluster B that comprise 15 and 74 mature miRNAs, respectively. These miRNA clusters are flanked by lincRNAs and separated by a lincRNA and an snoRNA cluster. (B) MA plot of our miRNA differential expression results (*n* = 34) with miRNAs from the larger miRNA cluster on 14q32 circled in dark blue. DESeq2 uses independent filtering to reduce the number of explicit differential expression tests it runs (Love *et al.* 2014). miRNAs that were not differentially expressed are depicted in light gray if they were removed by independent filtering and in dark gray otherwise. Significantly differentially expressed miRNAs (FDR = 1%) are colored in red. (C and D) Replication of the overrepresentation in exosomes of miRNAs from the large cluster on 14q32 using HeLa cell data (*n* = 5) from Honegger *et al.* (2015) (C) and B cell line data (*n* = 6) from Koppers-Lalic *et al.* (2014) (D). Note that the Koppers-Lalic *et al.* data were tested for differential expression using EdgeR, so the *x*-axis is in counts per million (instead of in DESeq2 normalized counts). Since only miRNAs that were explicitly tested for differential expression were reported, no light gray points appear in D.

removing reads that were too short after adapter trimming (Table S1). By comparing cells with exosomes, we noticed broad differences in the distribution of RNA biotypes. Cells expressed proportionally more miRNA and small nucleolar RNA, while exosomes contained a greater proportion of Piwi-interacting RNA (piRNA) and ribosomal RNA (Figure 2A). The proportion of other RNAs was similar between cells and exosomes. While we observed differences between cells and exosomes for four RNA classes, we focused on miRNA and piRNA for the remainder of the analyses, since these two types of RNA are the most abundant and corresponded to the target lengths of our small RNA isolation step at ~22 and 24–30 nucleotides long, respectively (Ha and Kim 2014).

Abundant differential expression of miRNA and piRNA between cells and exosomes

We tested whether individual miRNAs and piRNAs were selectively exported in exosomes or retained in cells (Figure 2, B and C). Of the 1739 miRNAs expressed in at least one sample, cells and exosomes differentially expressed 433 (25%) at an FDR of 1%. Cells expressed 187 miRNAs at higher levels, while exosomes contained proportionally more of the remaining 246 miRNAs (Table S2A). Of the 3422 piRNAs expressed, we found 374 (11%) were differentially expressed, with 135 being more highly expressed in cells and 239 more frequent in exosomes compared with cells (Table S2B). Since the cells and exosomes were sequenced on different lanes, confounding the contrast of

interest with potential lane effects, we resequenced the libraries this time batching cells and exosomes from the same individual together. Our resequencing data confirmed the differential expression results from our initial sequencing run, establishing that any lane effects were minimal (Figure S2). Our results highlight clear and widespread differences in the levels of particular miRNAs and piRNAs between cells and exosomes.

Exosomes export a large miRNA cluster on chromosome 14q32

Many miRNAs are in close genomic proximity to other miRNAs, and are sometimes transcribed as long polycistrons with multiple hairpins (Altuvia *et al.* 2005). If clustered miRNAs have correlated counts, then analyzing cluster-level rather than mature miRNA counts may be a more powerful approach. Accordingly, we generated cluster-level counts for each sample and tested for differential cluster expression between cells and exosomes. While the cluster results broadly mimicked the mature miRNA results (as many miRNAs are in their own individual cluster), two large miRNA clusters appeared to be selectively exported in exosomes (Figure S3). Most of the 42-hairpin 14q32 miRNA cluster, which yields 74 mature miRNAs, was massively exported into exosomes while being nearly undetectable in cells (Figure 3, A and B). We additionally confirmed that miRNAs from this cluster were enriched in exosomes in data from two independent
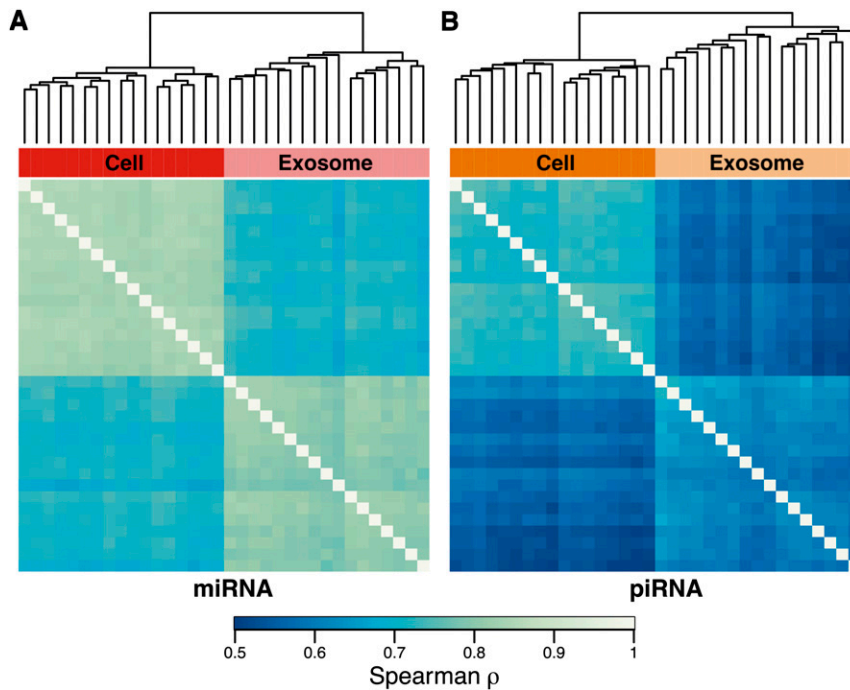
Figure 4 Cells and exosomes cluster by their miRNA and piRNA expression profiles. Hierarchical clustering of samples by the Spearman correlation coefficients of (A) miRNA and (B) piRNA expression. Samples cluster by compartment, confirming that cells and exosomes have distinct expression profiles. On average, the correlations between cell samples are higher than between exosome samples for both miRNA (0.83 *vs.* 0.80) and piRNA (0.72 *vs.* 0.61) (two-sided Wilcoxon rank sum test, $p < 1 \times 10^{-15}$ for miRNA and piRNA).

studies (Koppers-Lalic *et al.* 2014; Honegger *et al.* 2015) that sampled additional cell types, such as HeLa cells (Figure 3, C and D).

Selective miRNA export in exosomes is incompletely explained by known miRNA sorting mechanisms

The mechanism controlling coordinated export of the 14q32 cluster in exosomes is unknown. In the normal miRNA biogenic pathway, primary miRNAs (pri-miRNAs), including polycistronic miRNAs, are transcribed then cleaved into individual hairpins before nuclear export (Bartel 2004). Therefore, even if the cluster is transcribed as a polycistronic pri-miRNA—a hypothesis put forward by Seitz *et al.* (2004) that remains to be rigorously verified in humans—the component hairpins would typically be physically separated before entering the cytoplasm. We explored two mechanisms derived from recent studies (Villarroya-Beltri *et al.* 2013; Koppers-Lalic *et al.* 2014) that could explain exosome miRNA sorting for this cluster in particular, or for all differentially expressed miRNAs: (1) that a short sequence motif directs exosome loading, and (2) that NTAs to the 3′ end of the mature miRNAs favor inclusion in exosomes or retention in cells.

We searched for exosome- or cell-specific motifs using the MEME algorithm, but did not identify previously characterized motifs (Villarroya-Beltri *et al.* 2013). The 14q32 cluster dominated the exosome signal so strongly that the only significant motif was a sequence shared by most miRNAs in that cluster, and the cell signal was similarly dominated by the sequence similarities within the let-7 and mir-17 miRNA families. We then profiled and analyzed 3′ NTAs in all samples. On average, 80% of miRNA reads in both compartments lacked NTAs. This was true for both the 14q32 cluster as well as for the set of all mature miRNAs. We failed to detect substantial changes in the frequencies of NTAs between compartments, except for poly-adenylation, which was significantly higher in cells than exosomes. However, the adenylation of a small number of highly abundant miRNAs, rather than consistent changes in adenylation frequencies across many miRNAs, appeared to drive that effect (Figure S4). Taken together, these data suggest that neither known sequence motifs nor 3′-NTAs

explain 14q32-encoded miRNA export or our differential expression results more generally.

Interindividual variability in miRNA and piRNA expression profiles

We clustered samples hierarchically, according to the Spearman correlation of their miRNA and piRNA expression, and found that samples from the same compartment cluster together based on both their miRNA and piRNA expression profiles (Figure 4 and Figure S5). Notably, there was significantly more variability between exosomes samples than between cells samples for both miRNA and piRNA, though the difference was small for miRNA, as well as more variability overall for piRNA compared with miRNA (Figure 4 and Figure S5).

Cells and exosomes can share miRNA eQTL

Genetic variants associated with the expression of particular genes, known as eQTLs, have been studied in multiple cell types and tissues as a common source of gene expression variability between individuals (GTEx Consortium 2015). However, to our knowledge, it remains unknown whether genetic variation influences exosome content. To evaluate this, for each unique miRNA and piRNA, we tested for an association between the expression in the 11 children and the parental haplotypes they inherited, using a linear model (Table S3). The excess of small *p*-values for miRNA eQTL tests in cells suggests that genetic differences may explain a portion of the variability observed between individuals (Figure S6A). We observed the same but weaker signal in exosomes, but uncovered little evidence of piRNA eQTLs in either cells or exosomes (Figure S6, A and B).

We observed that the miRNA eQTL effect sizes of both parental haplotypes were significantly correlated between cells and exosomes, hinting at shared genetic effects on expression (Figure S7). To explore this further, we intersected the list of significant miRNA eQTLs from cells and exosomes at an FDR of 20% (Figure 5A). This identified four mature miRNAs, the pairs of products from hsa-miR-151a and hsa-miR-335, as putative shared miRNA eQTLs (Figure 5, B and C). The

effect sizes are comparable within each pair of products, which is consistent with a *cis*-regulatory variant controlling the transcription of the miRNA precursors. Furthermore, the hsa-miR-151a and hsa-miR-335 products were previously identified as having an eQTL in cells by Huan *et al.* (2015) and Lappalainen *et al.* (2013), respectively.

DISCUSSION

We isolated extracellular vesicles from the LCLs of a large family, then verified they were exosomes by three orthogonal methods. By sequencing the small RNA from both cells and exosomes, we not only confirmed previous observations that exosomes selectively export specific transcripts from the cells that release them (Nolte-'t Hoen *et al.* 2012; Villarroya-Beltri *et al.* 2013; Koppers-Lalic *et al.* 2014; Squadrito *et al.* 2014), but also discovered that a large miRNA cluster on chromosome 14q32 is exosome-specific. We observed that exosome expression profiles tend to be more variable between individuals than cellular ones. Finally, we provide the first evidence that exosomes reflect a subset of genetically driven cellular miRNA expression differences.

Our large sample for differential expression analysis provided ample power to detect differences between cells and exosomes, and our results support the hypothesis of directed RNA export in exosomes. Indeed, we report almost an order of magnitude more differentially expressed transcripts than a previous exosome sequencing study in LCLs (Koppers-Lalic *et al.* 2014). Overall, the relative proportions of the RNA biotypes we measured in cells and exosomes agree with previous work in immune cell types (Nolte-'t Hoen *et al.* 2012; Koppers-Lalic *et al.* 2014) when we consider the types of RNA in common between the studies. One exception is that Koppers-Lalic *et al.* (2014) observed the same relative quantity of piRNA between cells and exosomes, while we observed proportionally more piRNA in exosomes. This discrepancy is probably explained by differences in our mapping strategies. Several piRNA reside in multiple genomic loci and therefore can be challenging to map uniquely to the reference genome. Since we specifically aligned reads to piRNABank as well as to the reference genome, our method can capture more piRNA reads, allowing us to perform a more sensitive analysis of this type of RNA.

We also discovered that a large cluster of miRNAs on 14q32 is preferentially present in exosomes while being virtually undetectable in cells. This constitutes the first report of a large set of miRNAs with a shared genomic locus almost completely exported by exosomes. The maternally expressed 14q32 locus contains two clusters of eight and 42 miRNA hairpins, spanning 200 kb. Multiple studies have linked expression changes in these miRNAs to human diseases and cancers (Sarver *et al.* 2013; Kameswaran *et al.* 2014; Manodoro *et al.* 2014), and have described the effect of deleting larger regions around the 14q32 locus, including 14q uniparental disomy (Temple *et al.* 1991; Ogata and Kagami 2016). Many of these studies measured and associated cellular miRNA levels with a range of molecular and cellular phenotypes, such as epigenetic modification and disease progression. Our results suggest that cellular miRNA levels may inaccurately reflect 14q32 locus expression if exosomes export virtually all resulting miRNA products; indeed, in our data, the locus would appear silenced if evaluated solely by our cellular miRNA samples. While we observed replication of this cluster-level differential expression in HeLa cells, it remains important for future studies to replicate our observations in primary cell-derived exosomes.

Several mechanisms have been proposed to explain how cells sort small RNAs into exosomes. For example, Villarroya-Beltri *et al.* (2013) proposed that sumoylated hnRNPA2B1 shuttles miRNA with particu-

miRNA	Adjusted p-value		Replication cohort
	Cell	Exosome	
hsa-miR-151a-3p†	0.0348	0.0580	Huan *et al.*
hsa-miR-151a-5p†	0.0162	0.161	Huan *et al.*
hsa-miR-335-3p†	0.061	0.161	Lappalainen *et al.*
hsa-miR-335-5p†	0.061	0.0580	Lappalainen *et al.*
hsa-miR-378a-5p	0.114	0.997	
hsa-miR-589-5p	0.0162	0.265	Lappalainen *et al.*
hsa-miR-4326	0.181	0.902	Lappalainen *et al.*
hsa-miR-4677-3p	0.879	0.161	
hsa-miR-6840-5p	0.00142	0.275	
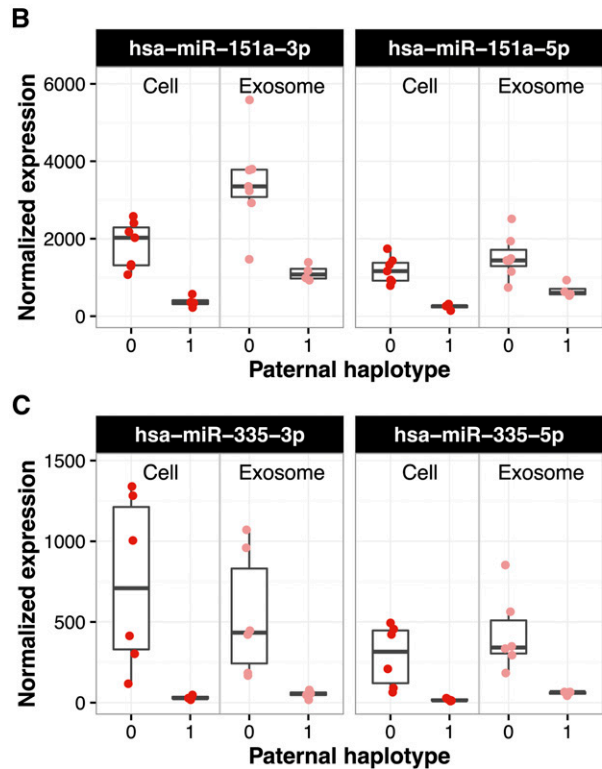
Figure 5 Shared miRNA eQTL between cells and exosomes. (A) Table of all miRNAs with an eQTL at FDR ≤ 20% in either cells or exosomes. The four miRNAs that pass the FDR threshold in both cells and exosomes are marked with a dagger and depicted in (B and C). (B and C) Putative shared miRNA eQTLs. The normalized expression levels of the 11 children are shown in cells and exosomes for both products of hsa-miR-151a (B) and hsa-miR-335 (C). The expression values are segregated by their inherited paternal haplotype, denoted as 0 or 1. The maternal haplotypes are not depicted because they did not show a strong association with expression.

lar sequence motifs into exosomes. Koppers-Lalic *et al.* (2014) reported that 3′ uridylated miRNA isoforms are preferentially loaded into exosomes compared with 3′ adenylated ones. However, neither of these mechanisms accounts for the export of the 14q32 cluster or our differential expression results more generally. While each of these mechanisms may contribute to directing miRNA export, there are clearly other mechanisms accounting for miRNA sorting into exosomes that remain unexplained. Squadrito *et al.* (Squadrito *et al.* 2014) found that both an miRNA's expression level and that of its targets influences its relative enrichment in exosomes compared with cells, but we lacked the mRNA measurements to evaluate this possibility.

We speculate that the widespread differential expression of miRNA and piRNA, including the export of the 14q32 cluster, is likely driven through sequence recognition by RNA-binding protein(s) with uncharacterized binding affinities. This hypothesis, similar to what has been previously proposed (Villarroya-Beltri *et al.* 2013), could plausibly explain the massive export of 14q32 cluster miRNAs, since they share a consensus sequence untargeted by known RNA-binding proteins. Previous studies have suggested that the miRNA exported in exosomes can silence mRNA in target cells (Hergenreider *et al.* 2012; Montecalvo *et al.* 2012; Umezu *et al.* 2013; van Balkom *et al.* 2013). Given recent evidence that piRNAs, like miRNAs, can silence target mRNAs (Gou *et al.* 2014; Gebert *et al.* 2015), it is possible that both the miRNA and piRNA exported in exosomes serve a regulatory role in recipient cells. Both repression (Sarver *et al.* 2013; Kameswaran *et al.* 2014) and overexpression (Manodoro *et al.* 2014) of the miRNAs in the 14q32 cluster has been associated with disease in various cell types. It is unclear whether the cluster miRNAs have a regulatory role in the target cells or are exported to prevent their activity in the parent cells.

In addition to assessing global differences between cells and exosomes, we sought to quantify the variability seen between individuals, as this has not previously been measured in a sample of this size. We observed that miRNA and piRNA levels varied more between exosome samples than between cell samples. This implies that RNA export in exosomes is less tightly controlled than RNA regulation within cells, which has important implications for the use of exosomes in diagnostics. To investigate one possible source of variability in miRNA and piRNA expression between individuals, we performed the first eQTL analysis of exosomes and their parent cells. We identified the products of hsa-miR-151a and hsa-miR-335 as miRNA eQTLs that appear to be shared between cells and exosomes. It is unsurprising that we only identified a few miRNA eQTLs, since previous studies have suggested that they are rare relative to mRNA eQTLs (Borel *et al.* 2011; Lappalainen *et al.* 2013; Huan *et al.* 2015) and the small sample size limited our statistical power. While it appears that genetic variation can only explain a small portion of the variability in exosomal miRNA between individuals, we found evidence that genetics may play a role in exosome miRNA regulation. We expect that future genetic studies of exosome RNA content will enable the discovery of more miRNA eQTLs, improving upon our identification of genetic differences in intercellular communication and their potential consequences on human phenotypes.

As exosomes are increasingly explored as diagnostic tools and tied to different aspects of human health, it becomes important for us to understand the different levels of regulation they undergo and how their contents may vary from person to person. Our findings improve our understanding of the factors influencing miRNA export and underscore the importance of considering individual differences when studying exosomes.

LITERATURE CITED

Admyre, C., S. M. Johansson, K. R. Qazi, J.-J. Filen, R. Lahesmaa *et al.*, 2007 Exosomes with immune modulatory features are present in human breast milk. J. Immunol. 179: 1969–1978.

Alexander, M., R. Hu, M. C. Runtsch, D. A. Kagele, T. L. Mosbruger *et al.*, 2015 Exosome-delivered microRNAs modulate the inflammatory response to endotoxin. Nat. Commun. 6: 7321.

Altuvia, Y., P. Landgraf, G. Lithwick, N. Elefant, S. Pfeffer *et al.*, 2005 Clustering and conservation patterns of human microRNAs. Nucleic Acids Res. 33: 2697–2706.

Anders, S., P. T. Pyl, and W. Huber, 2014 HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics 31: 166–169.

Baietti, M. F., Z. Zhang, E. Mortier, A. Melchior, G. Degeest *et al.*, 2012 Syndecan–syntenin–ALIX regulates the biogenesis of exosomes. Nat. Cell Biol. 14: 677–685.

Bailey, T. L., M. Boden, F. A. Buske, M. Frith, C. E. Grant *et al.*, 2009 MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res. 37: W202–8.

Bartel, D. P., 2004 MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116: 281–297.

Battle, A., S. Mostafavi, X. Zhu, J. B. Potash, M. M. Weissman *et al.*, 2014 Characterizing the genetic basis of transcriptome diversity through RNA-sequencing of 922 individuals. Genome Res. 24: 14–24.

Benjamini, Y., and Y. Hochberg, 1995 Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. R. Stat. Soc. B 57: 289–300.

Boelens, M. C., T. J. Wu, B. Y. Nabet, B. Xu, Y. Qiu *et al.*, 2014 Exosome transfer from stromal to breast cancer cells regulates therapy resistance pathways. Cell 159: 499–513.

Borel, C., S. Deutsch, A. Letourneau, E. Migliavacca, S. B. Montgomery *et al.*, 2011 Identification of *cis*- and *trans*-regulatory variation modulating microRNA expression levels in human fibroblasts. Genome Res. 21: 68–73.

Chivet, M., C. Javalet, K. Laulagnier, B. Blot, F. J. Hemming *et al.*, 2014 Exosomes secreted by cortical neurons upon glutamatergic synapse activation specifically interact with neurons. J. Extracell. Vesicles 3: 24722.

Costa-Silva, B., N. M. Aiello, A. J. Ocean, S. Singh, H. Zhang *et al.*, 2015 Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. Nat. Cell Biol. 17: 816–826.

Fong, M. Y., W. Zhou, L. Liu, A. Y. Alontaga, M. Chandra *et al.*, 2015 Breast-cancer-secreted miR-122 reprograms glucose metabolism in premetastatic niche to promote metastasis. Nat. Cell Biol. 17: 183–194.

Frühbeis, C., D. Fröhlich, W. P. Kuo, J. Amphornrat, S. Thilemann *et al.*, 2013 Neurotransmitter-triggered transfer of exosomes mediates oligodendrocyte–neuron communication. PLoS Biol. 11: e1001604.

Gebert, D., R. F. Ketting, H. Zischler, and D. Rosenkranz, 2015 piRNAs from pig testis provide evidence for a conserved role of the Piwi pathway in post-transcriptional gene regulation in mammals. PLoS One 10: e0124860.

Gou, L.-T., P. Dai, J.-H. Yang, Y. Xue, Y.-P. Hu *et al.*, 2014 Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. Cell Res. 24: 680–700.

GTEx Consortium, 2015 The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science 348: 648–660.

Ha, M., and V. N. Kim, 2014 Regulation of microRNA biogenesis. Nat. Rev. Mol. Cell Biol. 15: 509–524.

Harding, C., J. Heuser, and P. Stahl, 1983 Receptor-mediated endocytosis of transferrin and recycling of the transferrin receptor in rat reticulocytes. J. Cell Biol. 97: 329–339.

Hergenreider, E., S. Heydt, K. Tréguer, T. Boettger, A. J. G. Horrevoets *et al.*, 2012 Atheroprotective communication between endothelial cells and smooth muscle cells through miRNAs. Nat. Cell Biol. 14: 249–256.

Honegger, A., D. Schilling, S. Bastian, J. Sponagel, V. Kuryshev *et al.*, 2015 Dependence of intracellular and exosomal microRNAs on viral E6/E7 oncogene expression in HPV-positive tumor cells. PLOS Pathog. 11: e1004712.

Hood, J. L., S. San Roman, and S. A. Wickline, 2011 Exosomes released by melanoma cells prepare sentinel lymph nodes for tumor metastasis. Cancer Res. 71: 3792–3801.

Huan, T., J. Rong, C. Liu, X. Zhang, K. Tanriverdi *et al.*, 2015   Genome-wide identification of microRNA expression quantitative trait loci. Nat. Commun. 6: 6601.

Kameswaran, V., N. C. Bramswig, L. B. McKenna, M. Penn, J. Schug *et al.*, 2014   Epigenetic regulation of the DLK1–MEG3 MicroRNA cluster in human type 2 diabetic islets. Cell Metab. 19: 135–145.

Kim, S. H., E. R. Lechman, N. Bianco, R. Menon, A. Keravala *et al.*, 2005   Exosomes derived from IL-10-treated dendritic cells can suppress inflammation and collagen-induced arthritis. J. Immunol. 174: 6440–6448.

Koppers-Lalic, D., M. Hackenberg, I. V. Bijnsdorp, M. A. J. van Eijndhoven, P. Sadek *et al.*, 2014   Nontemplated nucleotide additions distinguish the small RNA composition in cells from exosomes. Cell Reports 8: 1649–1658.

Kozomara, A., and S. Griffiths-Jones, 2014   miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. 42: D68–D73.

Langmead, B., and S. L. Salzberg, 2012   Fast gapped-read alignment with Bowtie 2. Nat. Methods 9: 357–359.

Lappalainen, T., M. Sammeth, M. R. Friedländer, P. A. C. 't Hoen, J. Monlong *et al.*, 2013   Transcriptome and genome sequencing uncovers functional variation in humans. Nature 501: 506–511.

Li, X., A. Battle, K. J. Karczewski, Z. Zappala, D. A. Knowles *et al.*, 2014   Transcriptome sequencing of a large human family identifies the impact of rare noncoding variants. Am. J. Hum. Genet. 95: 245–256.

Love, M. I., W. Huber, and S. Anders, 2014   Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 15: 550.

Manodoro, F., J. Marzec, T. Chaplin, F. Miraki-Moud, E. Moravcsik *et al.*, 2014   Loss of imprinting at the 14q32 domain is associated with microRNA overexpression in acute promyelocytic leukemia. Blood 123: 2066–2074.

Marçais, G., and C. Kingsford, 2011   A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics 27: 764–770.

Meckes, D. G., K. H. Y. Shair, A. R. Marquitz, C.-P. Kung, R. H. Edwards *et al.*, 2010   Human tumor virus utilizes exosomes for intercellular communication. Proc. Natl. Acad. Sci. USA 107: 20370–20375.

Mittelbrunn, M., C. Gutiérrez-Vázquez, C. Villarroya-Beltri, S. González, F. Sánchez-Cabo *et al.*, 2011   Unidirectional transfer of microRNA-loaded exosomes from T cells to antigen-presenting cells. Nat. Commun. 2: 282.

Montecalvo, A., A. T. Larregina, W. J. Shufesky, D. Beer Stolz, M. L. G. Sullivan *et al.*, 2012   Mechanism of transfer of functional microRNAs between mouse dendritic cells via exosomes. Blood 119: 756–766.

Nolte-'t Hoen, E. N. M., H. P. J. Buermans, M. Waasdorp, W. Stoorvogel, M. H. M. Wauben *et al.*, 2012   Deep sequencing of RNA from immune cell-derived vesicles uncovers the selective incorporation of small non-coding RNA biotypes with potential regulatory functions. Nucleic Acids Res. 40: 9272–9285.

Ogata, T., and M. Kagami, 2016   Kagami-Ogata syndrome: a clinically recognizable upd(14)pat and related disorder affecting the chromosome 14q32.2 imprinted region. J. Hum. Genet. 61: 87–94.

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: 942–948.

Peinado, H., M. Alečković, S. Lavotshkin, I. Matei, B. Costa-Silva *et al.*, 2012   Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. Nat. Med. 18: 883–891.

Raposo, G., and W. Stoorvogel, 2013   Extracellular vesicles: exosomes, microvesicles, and friends. J. Cell Biol. 200: 373–383.

Sai Lakshmi, S., and S. Agrawal, 2007   piRNABank: a web resource on classified and clustered Piwi-interacting RNAs. Nucleic Acids Res. 36: D173–D177.

Sarver, A. L., V. Thayanithy, M. C. Scott, A.-M. Cleton-Jansen, P. C. Hogendoorn *et al.*, 2013   MicroRNAs at the human 14q32 locus have prognostic significance in osteosarcoma. Orphanet J. Rare Dis. 8: 7.

Segura, E., 2005   ICAM-1 on exosomes from mature dendritic cells is critical for efficient naive T-cell priming. Blood 106: 216–223.

Seitz, H., H. Royo, M.-L. Bortolin, S.-P. Lin, A. C. Ferguson-Smith *et al.*, 2004   A large imprinted microRNA gene cluster at the mouse Dlk1-Gtl2 domain. Genome Res. 14: 1741–1748.

Squadrito, M. L., C. Baer, F. Burdet, C. Maderna, G. D. Gilfillan *et al.*, 2014   Endogenous RNAs modulate microRNA sorting to exosomes and transfer to acceptor cells. Cell Reports 8: 1432–1446.

Tadokoro, H., T. Umezu, K. Ohyashiki, T. Hirano, and J. H. Ohyashiki, 2013   Exosomes derived from hypoxic leukemia cells enhance tube formation in endothelial cells. J. Biol. Chem. 288: 34343–34351.

Temple, I. K., A. Cockwell, T. Hassold, D. Pettay, and P. Jacobs, 1991   Maternal uniparental disomy for chromosome 14. J. Med. Genet. 28: 511–514.

Thakur, B. K., H. Zhang, A. Becker, I. Matei, Y. Huang *et al.*, 2014   Double-stranded DNA in exosomes: a novel biomarker in cancer detection. Cell Res. 24: 766–769.

Umezu, T., K. Ohyashiki, M. Kuroda, and J. H. Ohyashiki, 2013   Leukemia cell to endothelial cell communication via exosomal miRNAs. Oncogene 32: 2747–2755.

Valadi, H., K. Ekström, A. Bossios, M. Sjöstrand, J. J. Lee *et al.*, 2007   Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nat. Cell Biol. 9: 654–659.

van Balkom, B. W. M., O. G. de Jong, M. Smits, J. Brummelman, K. den Ouden *et al.*, 2013   Endothelial cells require miR-214 to secrete exosomes that suppress senescence and induce angiogenesis in human and mouse endothelial cells. Blood 121: 3997–4006.

Villarroya-Beltri, C., C. Gutiérrez-Vázquez, F. Sánchez-Cabo, D. Pérez-Hernández, J. Vázquez *et al.*, 2013   Sumoylated hnRNPA2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs. Nat. Commun. 4: 2980.

Zitvogel, L., A. Regnault, A. Lozier, J. Wolfers, C. Flament *et al.*, 1998   Eradication of established murine tumors using a novel cell-free vaccine: dendritic cell-derived exosomes. Nat. Med. 4: 594–600.

Communicating editor: D. J. de Koning