

## ORIGINAL RESEARCH

**The extracellular RNA complement of *Escherichia coli***

Anubrata Ghosal<sup>1</sup>, Bimal Babu Upadhyaya<sup>1</sup>, Joëlle V. Fritz<sup>1</sup>, Anna Heintz-Buschart<sup>1</sup>, Mahesh S. Desai<sup>1</sup>, Dilmurat Yusuf<sup>1</sup>, David Huang<sup>2</sup>, Aidos Baumuratov<sup>1</sup>, Kai Wang<sup>3</sup>, David Galas<sup>1,2</sup> & Paul Wilmes<sup>1</sup>

<sup>1</sup>Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Campus Belval, 7 avenue des Hauts-Fourneaux, Esch-sur-Alzette L-4362, Luxembourg

<sup>2</sup>Pacific Northwest Diabetes Research, Seattle, Washington

<sup>3</sup>Institute for Systems Biology, Seattle, Washington

**Keywords**

*Escherichia coli*, extracellular, outer membrane vesicle, RNA.

**Correspondence**

Paul Wilmes, Luxembourg Centre for Systems Biomedicine, University of Luxembourg, Campus Belval, 7 avenue des Hauts-Fourneaux, Esch-sur-Alzette L-4362, Luxembourg. Tel: +352 46 66 44 6188; Fax: +352 46 66 44 6949; E-mail: paul.wilmes@uni.lu

**Funding Information**

The present work was supported by an ATTRACT programme grant (A09/03) CORE programme grant (CORE11/BM/1186762), European Union Joint Programming in Neurodegenerative Diseases grant (INTER/JPND/12/01), and Proof-of-Concept grant (PoC/13/02) to PW, funded by the Luxembourg National Research Fund (FNR). Further financial support for the work was through FNR AFR fellowships to JF (Ref. no. 5821107) and DY (Ref. no.1180851).

Received: 21 August 2014; Revised: 21 November 2014; Accepted: 1 December 2014

*MicrobiologyOpen* 2015; 4(2): 252–266

doi: 10.1002/mbo3.235

**Introduction**

While growing either in a natural ecosystem or artificial conditions, bacteria secrete intracellular products into their extracellular milieu (Tseng et al. 2009). Secretory

**Abstract**

The secretion of biomolecules into the extracellular milieu is a common and well-conserved phenomenon in biology. In bacteria, secreted biomolecules are not only involved in intra-species communication but they also play roles in inter-kingdom exchanges and pathogenicity. To date, released products, such as small molecules, DNA, peptides, and proteins, have been well studied in bacteria. However, the bacterial extracellular RNA complement has so far not been comprehensively characterized. Here, we have analyzed, using a combination of physical characterization and high-throughput sequencing, the extracellular RNA complement of both outer membrane vesicle (OMV)-associated and OMV-free RNA of the enteric Gram-negative model bacterium *Escherichia coli* K-12 substrain MG1655 and have compared it to its intracellular RNA complement. Our results demonstrate that a large part of the extracellular RNA complement is in the size range between 15 and 40 nucleotides and is derived from specific intracellular RNAs. Furthermore, RNA is associated with OMVs and the relative abundances of RNA biotypes in the intracellular, OMV and OMV-free fractions are distinct. Apart from rRNA fragments, a significant portion of the extracellular RNA complement is composed of specific cleavage products of functionally important structural noncoding RNAs, including tRNAs, 4.5S RNA, 6S RNA, and tmRNA. In addition, the extracellular RNA pool includes RNA biotypes from cryptic prophages, intergenic, and coding regions, of which some are so far uncharacterised, for example, transcripts mapping to the *fimA-fimL* and *ves-spy* intergenic regions. Our study provides the first detailed characterization of the extracellular RNA complement of the enteric model bacterium *E. coli*. Analogous to findings in eukaryotes, our results suggest the selective export of specific RNA biotypes by *E. coli*, which in turn indicates a potential role for extracellular bacterial RNAs in intercellular communication.

products are not only involved in bacterial social behavior, usually referred to as quorum sensing (Molloy 2010) but also in pathogenicity (Lee and Schneewind 2001) and inter-kingdom communication (Hughes and Sperandio 2008; Shen et al. 2012; Furusawa et al. 2013). A large

diversity of molecules including small hormone-like autoinducer molecules, peptides, proteins, glycoproteins, and DNA have already been identified as mediators in both intra- and inter-kingdom communication (Waters and Bassler 2005; Molloy 2010; Shen et al. 2012) and in some cases as effectors in pathogenicity (Tseng et al. 2009; Jun et al. 2013). Bacteria disseminate secreted products either through their secretory systems via continuous or discontinuous passages across the bacterial membrane or by the release of outer membrane vesicles (OMVs) (Tseng et al. 2009; Shen et al. 2012). Gram-negative bacteria are also able to export their secretory products into the cytosol of host cells using their penetrating appendages of type III, type IV, and type VI secretion systems, whereas OMV-mediated transport allows bacteria to deliver secretory products remotely without direct contact with host cells (Tseng et al. 2009; Shen et al. 2012; Casson et al. 2013). OMVs may be taken up by host cells and may contribute to pathogenesis, as already demonstrated for OMVs derived from *Helicobacter pylori*, which are taken up by gastric epithelial cells (Parker et al. 2010) and enhance the carcinogenic potential of this specific bacterium (Chitcholtan et al. 2008).

The release of OMVs is a common phenomenon observed for many Gram-negative bacteria, including *Escherichia coli* (Wai et al. 2003; Kesty et al. 2004; Park et al. 2010; Roy et al. 2011). In addition, a recent study has reported that vesicular secretion is also common in Gram-positive bacteria (Thay et al. 2013). Single membrane OMVs comprised outer membrane proteins and periplasmic components (Carmen Schwechheimer and Kuehn 2013), whereas bilayer forms of OMVs also contain inner membrane and cytoplasmic constituents (Pérez-Cruz et al. 2013). In terms of nucleic acids, the DNA content of OMVs has been characterized for *Shewanella vesiculosa*, *Neisseria gonorrhoeae*, and *Pseudomonas aeruginosa* (Dorward et al. 1989; Renelli et al. 2004; Pérez-Cruz et al. 2013). So far, no studies have reported OMVs produced by heterotrophic bacteria to contain RNA, except a suggestion in one study (Dorward et al. 1989). A very recent report, however, has identified RNA in OMVs released by marine photoautotrophs (Biller et al. 2014). Eukaryotic cells also secrete cellular products by shedding of extracellular vesicles (Théry et al. 2002; El Andaloussi et al. 2013), which are analogue to OMVs. Vesicles derived from mammalian cells have recently been extensively characterized for their RNA content, since the RNA is thought to be involved in intercellular regulation (Kogure et al. 2011; Chen et al. 2012; Hoen et al. 2012). Similarly, fungal extracellular vesicles are also known to contain RNA (Nicola et al. 2009).

RNA secreted by bacteria may play important roles in microbe–microbe and host–microbe interactions. In this

study, we aimed to characterize, using a combination of physical characterization and high-throughput sequencing, the extracellular RNA complement (both OMV-associated and OMV-free) of the enteric Gram-negative bacterium *E. coli* K-12 substrain MG1655 and compared it to its intracellular RNA complement.

## Materials and Methods

### Bacterial strain, extraction of OMVs, and OMV-free supernatant

OMVs were derived from the enteric Gram-negative bacterium *E. coli* K-12 substrain MG1655. Bacteria were grown in Luria-Bertani (LB) broth or M9 minimal medium at 37°C. OMVs were isolated as previously described (Wai et al. 2003) with slight modifications to the protocol as detailed below. Media were inoculated with a single bacterial colony and grown until early stationary phase. The bacterial cultures were then centrifuged at 4700g for 30 min at 4°C. After centrifugation, the supernatants were separated from the cell pellets and concentrated using a tangential flow filtration device (QuixStand Benchtop System; GE HealthCare) with a 100 kDa hollow fiber membrane (GE HealthCare Europe GmbH, Freiburg, Germany). The concentrated supernatants were then further passed through a 0.8/0.2 µm filter (PALL Life Sciences, Ann Arbor, MI, USA). As a control, filtered supernatants were checked for bacterial contamination by plating 1 mL of supernatant on a LB-agar plate followed by overnight incubation at 37°C. The filtered supernatants were then ultracentrifuged at 150,000g for 3 h at 4°C in order to separate the bacteria-free media into an OMV containing fraction (pellet) and OMV-free fraction (supernatant). Thus, OMVs were contained in the pellet fractions, which were then washed with Phosphate buffered saline (PBS) (1X), and finally resuspended in PBS and stored at –80°C. The OMV-depleted supernatants were also stored at –80°C for extraction of RNAs. Unless otherwise stated, all experiments were performed with samples prepared by the method described above.

The resulting OMV fractions were then further separated into low- and high-density fractions (HDFs) by gradient ultracentrifugation as previously described (Horstman et al. 2000) with slight modifications as detailed below. In brief, the OMV fractions were precisely adjusted to a concentration of 45% Optiprep (Sigma-Aldrich, Diegem, Belgium) in 2 mL of volume and layered at the bottom of ultracentrifugation tubes, followed by 2 mL of 40%, 35%, 30%, 25%, and 20% Optiprep layered on top of the sample in sequence. The prepared tubes were then subjected to ultracentrifugation

at a speed of 100,000 for 16 h at 4°C. After centrifugation, 12 fractions were collected from the top (fractions were labeled 1 to 12 according to the order of collection). Fractions 2–6 (termed the low-density fraction, LDF) and 7–12 (termed the high-density fraction) were pooled for subsequent RNA extractions.

Similarly, OMV and OMV-free fractions were also obtained from exponentially growing bacteria in LB up to ODs between 0.4 and 0.5.

### Transmission electron microscopy

OMV samples were fixed in 2.5% glutaraldehyde and 2.5% paraformaldehyde in PBS (1X) and washed with PBS. Samples were placed on Formvar-coated nickel grids, air dried, and examined at 80 kV with a Morgagni 268D transmission electron microscope. Images were captured digitally by a Mega View III camera (Olympus Soft Imaging solutions GmbH, Münster, Germany).

### Extraction of proteins and SDS-PAGE electrophoresis

Protein samples were dissolved in lysis buffer (25 mmol/L of 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) pH 7.4, 5 mmol/L of Ethylenediaminetetraacetic acid (EDTA), 1 mmol/L of Dithiothreitol (DTT), 1 mg/mL of Lysozyme) and passed through three cycles of freezing and thawing. Finally, MgSO<sub>4</sub> was added to the sample and incubated on ice for 30 min. Samples were centrifuged at a speed of 14,000g for 30 min at 4°C. The supernatant was then separated from the samples and used for protein analyses.

Protein concentrations were determined using a Qubit Protein Assay Kit (Invitrogen/Life Technologies Europe: Ghent, Belgium) and separated on a 4–15% of polyacrylamide gel (Bio-Rad Laboratories N.V., Temse, Belgium). Gels were stained with either Imperial Protein Stain (Thermo Scientific/Piercenet, Erembodegem-Aalst, Belgium) or the SilverQuest™ Silver Staining Kit (Invitrogen/Life Technologies Europe: Ghent, Belgium).

### Extraction of RNAs and RNAseq

RNA was extracted from *E. coli* cells, OMVs, OMV-free supernatant, low-density OMV fractions (fractions 2–6), high-density OMV fractions (fractions 7–12), and culture media using the All-in-One Purification Kit (Norgen Biotek/LabOmics, Nivelles, Belgium). Samples were precipitated with 20% (w/v) trichloroacetic acid (TCA) followed by 80% (v/v) acetone washing before the extraction of RNA (Bauman and Kuehn 2006). Total RNA from each sample was collected in two fractions,

the large RNA fraction included RNA of the size greater than 200 nucleotides in length and the small RNA fraction (“microRNA” fraction) contained RNAs shorter than 200 nucleotides in length. Extracted RNAs were treated with DNase I (Thermo Scientific) to remove genomic DNA contamination. The size distributions of RNAs in both the large and small RNA fractions were determined using a Bioanalyzer (Agilent Technologies, Diegem, Belgium) instrument with the Agilent RNA 6000 Nano Kit and Agilent Small RNA kit (Agilent Technologies), respectively. The quantity of extracted RNA for the individual fractions ranged from 500 ng to 3 μg.

The RNAseq deep-sequencing library was prepared from the small RNA fractions according to a previously described protocol (Wang et al. 2012), and the library was sequenced using a single-end sequencing strategy on an Illumina (San Diego, California, USA) Genome Analyzer where the maximum read length was 50 nt.

### Bioinformatic analysis

Raw small RNA sequence libraries of different samples were subjected to preprocessing for quality control and adapter removal using the FastX toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). A primer contamination table was generated by analyzing the sequence libraries. Contaminant adapters (RNA 3' adapter and RNA PCR primer) were subsequently trimmed using the FastX toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). This step was repeated three times in order to achieve a high accuracy for adapter trimming. Following adapter trimming, the sequence reads were subjected to quality trimming to remove noise using a Phred quality scores threshold of 25, which corresponds to an error probability of 0.003 per base. In addition, sequence reads failing to retain all bases with Phred quality score of 25 per base throughout the read were removed. Reads shorter than 14 nucleotides were removed. Identical read sequences after adapter trimming and quality filtering were then collapsed while maintaining the count of the identical reads in the library. Finally, for further analyses of the sequence libraries, identical reads present in the libraries obtained from cell-free growth media and cell-free growth media incubated overnight were subtracted from the reads in all libraries.

Processed read libraries were mapped to the *E. coli* strain K-12 substrain MG1655 reference genome (accession number: NC\_000913.2) from the RefSeq database (<http://www.ncbi.nlm.nih.gov/refseq/>) using novoalign (<http://www.novocraft.com>) with zero mismatches allowed. Unmapped reads from the library were filtered using SAMtools (Li et al. 2009). Reads mapped to more

than one genomic location were given proportional weights using the strategy and algorithm described in De Hoon et al. (2010). Mapped reads were annotated using BEDTools (Quinlan et al. 2010) the gene feature file of *E. coli* strain K-12 substrain MG1655 available from the RefSeq database.

### Staining of OMVs and confocal laser scanning microscopy

OMVs were stained as described previously (Nicola et al. 2009) with slight modifications detailed below. Briefly, DNase I-treated OMV samples were incubated with 1  $\mu\text{mol/L}$  of SYTO<sup>®</sup> RNaselect<sup>™</sup> Green Fluorescent Cell Stain (Life Technologies) and 5  $\mu\text{mol/L}$  of red fluorescent Lipophilic Tracer DiD (Life Technologies) for an hour at 37°C. Samples were washed once in PBS (1X) followed by centrifugation at a speed of 150,000g for 1 h at 4°C and mounted onto a glass slide with mounting medium, Fluoroshield<sup>™</sup> (Sigma).

Confocal images were acquired on a laser scanning confocal microscope (LSM Zeiss 710) and processed using the ZEN software package (Zeiss, Zaventem, Belgium). The background was adjusted with dye and unstained control samples. For each sample, images were taken separately in rapid succession. The excitation and emission spectra were set at a wavelengths of  $\lambda_{638\text{nm}}$  and  $\lambda_{700\text{nm}}$  for DiD, and  $\lambda_{490\text{nm}}$  and  $\lambda_{530\text{nm}}$  for SYTO RNaselect, respectively. Images were adjusted for brightness.

### RT-qPCR

Equal amounts of DNase I-treated RNA were polyadenylated and converted to cDNA using the qScript microRNA cDNA Synthesis Kit (Quanta BioSciences, Gaithersburg, MD, USA). About 1 ng of cDNA was amplified in combination with sequence-specific forward primers such as: tRNA-Asn: GTATGTCACTGGTTCGAGTC, 23S rRNA: GTTAAGCGACTAAGCGTACAC, tRNA-met: TGCTCTAACCAACTGAGCTAC, tRNA-Thr: CTCA GTAGGTAGAGCAACTG, tRNA-Ile: CTTGTAGCTCAG GTGGTTAGAG, tRNA-His: CTATAGCTCAGTTGGTAGAG, tRNA-Val: TCGAGTCCACTCGGACGCAC, tmRNA: GCTGATTCTGGATTCGACG, 4.5S RNA: CTCTGTTGG TTCTCCCGCAAC or 6S RNA: TCTCTGAGATGTTCCG-CAAGC (obtained from Eurogentec) and PerfeCTa Universal PCR Primer (Quanta BioSciences) in a total volume of 10  $\mu\text{L}$  of reaction mixture where primers were added at a concentration of 0.2  $\mu\text{mol/L}$ , 5  $\mu\text{L}$  of PerfeCTa SYBR Green SuperMix (Quanta BioSciences) and the remaining volume was adjusted for cDNA amount and water. Primer efficiency was calculated using a 10-fold dilution series of the template cDNA,

and for all primer sets efficiency values were close to 100%. Real-time PCR was run on a LightCycler<sup>®</sup> 480 Real-Time PCR System (Prophac/Roche, Luxembourg-Howald, Luxembourg) using the thermal cycles of 95°C for 2 min and 40 cycles of 95°C for 10 sec, 57–62°C for 5 sec, and 72°C for 10 sec.  $C_t$  values were obtained using automatic baseline and threshold settings provided by the LightCycler<sup>®</sup> 480 Software, Version 1.5 (Prophac/Roche, Luxembourg-Howald, Luxembourg). Data were analyzed manually using the comparative  $C_T$  method ( $\Delta\Delta C_T$ ). Individual targets were analyzed in three biological replicates and represented as a mean.

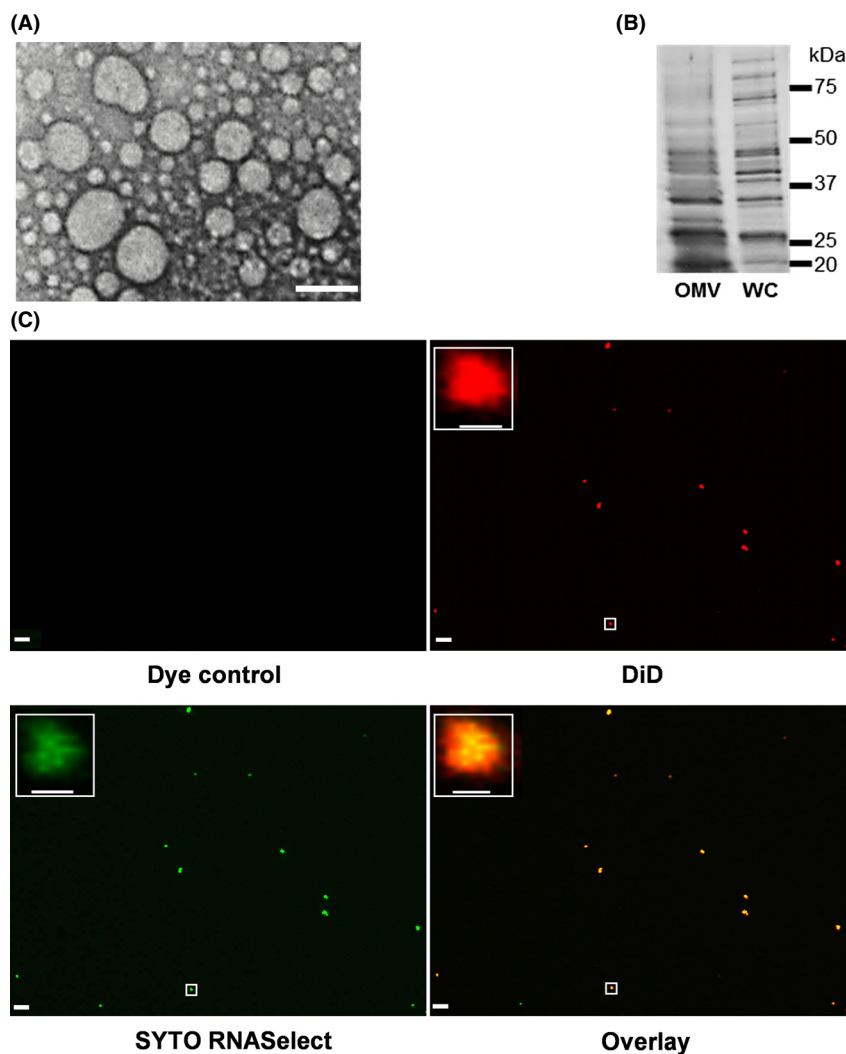
## Results and Discussion

### *Escherichia coli*-derived OMVs contain RNA

We hypothesized that bacteria release intracellular RNAs into their extracellular milieu, either through the shedding of OMVs or by using secretory pathways, in ways analogous to eukaryotic cells. In order to test our hypothesis, we first confirmed the presence of OMVs in the standard growth medium (LB) of *E. coli* K-12 substrain MG1655 using standard isolation procedures and transmission electron microscopy (Fig. 1A; Materials and Methods section). The enriched LB medium was chosen over a synthetic medium to obtain high growth yields of *E. coli* for subsequently being able to extract sufficient amounts of RNA. In addition, we analyzed the protein banding profiles of the OMV fractions and found these to be distinct from whole-cell protein extracts (Fig. 1B) (Wai et al. 2003). These commonly used analyses confirmed the presence of OMVs.

Next, we examined whether the OMV fraction from *E. coli* contains RNA, by applying a fluorescent labeling method which was adapted for staining of the OMV fraction (Nicola et al. 2009). Using this method, OMV-associated lipids were labeled with the lipid tracer dye DiD and RNAs with SYTO RNaselect dye, respectively (Fig. 1C). In most cases, RNA is found to be associated with OMV-associated lipids but the observed patterns also suggest the presence of nucleoprotein complexes and/or free extracellular RNA (Fig. 1C). Each individual color dot in the images likely represents the aggregation of several OMVs, as the typical sizes of OMVs are below the limit of resolution of confocal microscopy. As a negative control, a mixture of both dyes was also analyzed using the same procedure and no particles were identified (Fig. 1C). Through application of protocols similar to those used for the characterization of eukaryotic extracellular RNA, we were subsequently able to isolate both free and OMV-associated RNA from cultures of *E. coli* K-12 substrain MG1655.



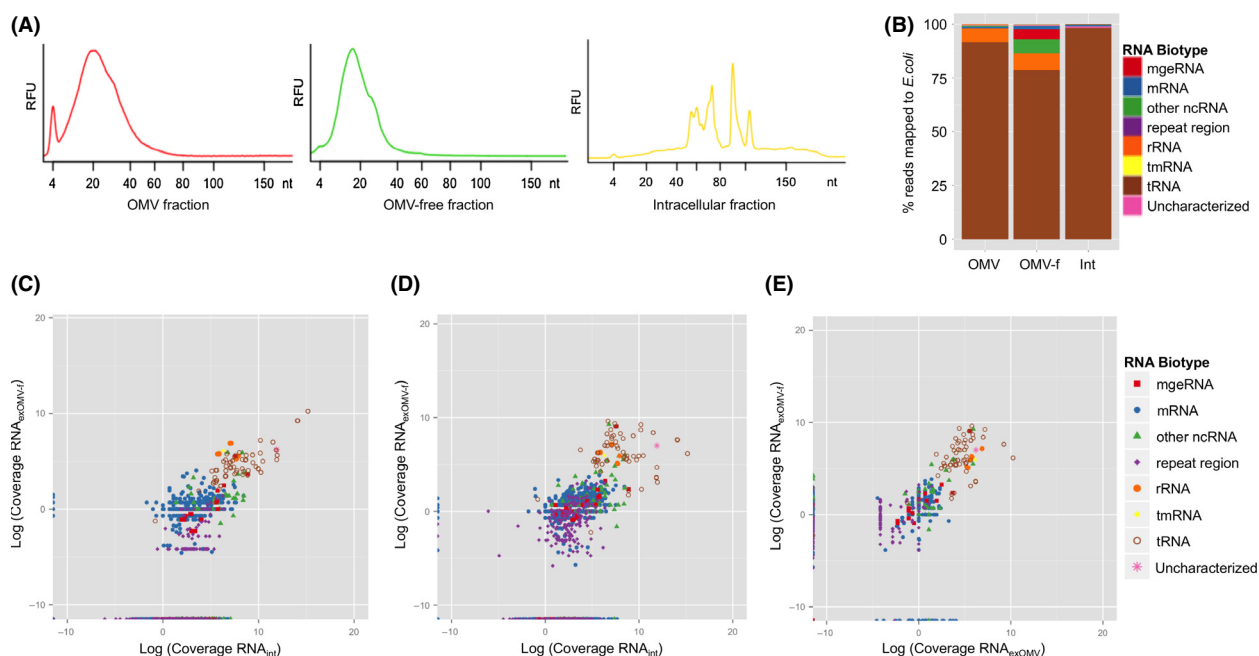


**Figure 1.** Characterization of outer membrane vesicular fractions isolated from *Escherichia coli* MG1655. (A) Visualization of isolated outer membrane vesicles (OMVs) using transmission electron microscopy (TEM). The size of the scale bar is equivalent to 50 nm. (B) SDS-PAGE gel image of protein samples from the OMV fraction and whole cells. Molecular weight markers are provided at the right side of the figure. (C) Confocal microscopy analysis of OMVs, stained with lipid tracer dye, DiD (red), and RNA-specific dye, SYTO RNaselect (green). The area highlighted within the white rectangular box is magnified in the inset. The scale bar is equivalent to 5  $\mu$ m in the main images and equivalent to 500 nm in the magnified images.

### Characterization of extracellular RNAs

Following OMV isolation, RNA was extracted from the extracellular OMV fraction ( $\text{RNA}_{\text{exOMV}}$ ), extracellular OMV-free fraction ( $\text{RNA}_{\text{exOMV-f}}$ ), and from whole cells (subsequently referred to as the intracellular RNA fraction;  $\text{RNA}_{\text{int}}$ ). To minimize potential confounding effects of the presence of RNA in the  $\text{RNA}_{\text{exOMV-f}}$  fraction due to the lysis of cells and/or OMVs, RNA was always isolated from the respective fractions from early stationary growth phase. The presence of distinct RNAs in the  $\text{RNA}_{\text{exOMV-f}}$  fraction compared to  $\text{RNA}_{\text{int}}$  (please also

see below) suggests the involvement of unknown mechanisms in the secretion of bacterial intracellular RNA into the extracellular milieu in addition to their release through OMVs. Each RNA fraction was subdivided into a <200 nt long RNA fraction (henceforth referred to as the “small RNA fraction”) and >200 nt long RNA fraction (“large RNA fraction”) and analyzed using an Agilent Bioanalyzer (Materials and Methods section). The majority of the RNA obtained from the  $\text{RNA}_{\text{exOMV}}$  and  $\text{RNA}_{\text{exOMV-f}}$  has a length less than 60 nt with an enrichment in the size range between 15 and 40 nucleotides (Fig. 2A, Fig. S1).



**Figure 2.** Characterization of the extracellular RNA complement of *Escherichia coli*. (A) Electropherograms of obtained RNA fractions below the length of 200 nucleotides. (B) Annotated profile of RNA<sub>exOMV</sub> (OMV), RNA<sub>exOMV-f</sub> (OMV-f), and RNA<sub>int</sub> (Int). (C), (D) and (E) represent plots of the inferred abundance levels of different RNA biotypes in different RNA fractions, that is, RNA<sub>int</sub>, RNA<sub>exOMV</sub>, and RNA<sub>exOMV-f</sub>, against each other. Mobile genetic element- mge.

Given the distinct enrichment in small RNA in the extracellular RNA complements, we further characterized these by deep RNAseq (Materials and Methods section). In order to compare the extracellular RNA profiles with the intracellular RNA profiles, the extracted <200 nt long intracellular RNA fraction was also sequenced. To identify and remove possible contaminant RNAs present in the LB growth medium (Pavankumar et al. 2012) from the subsequent analyses, RNA was extracted from freshly prepared cell-free LB medium and cell-free LB medium incubated overnight (without bacterial inoculation), and sequenced in the same way (Materials and Methods section). Analysis of the medium-derived reads highlight that the vast majority of these map to the yeast genome which can be explained by the presence of yeast extract in the LB medium. Prior to a detailed analysis of the sequence data mapping to the *E. coli* MG1655 genome, RNAseq read sequences in the extracellular and intracellular libraries, which were identical to those in the RNAseq libraries of cell-free LB media samples, were removed from the libraries to strictly limit potential contamination. This resulted in the removal of about 3% and 90% of the reads from intracellular and extracellular RNA read libraries, respectively (Table S1). Reads which mapped to both the media RNA libraries and the *E. coli* genome were termed “cross-mapping reads.” The majority of the cross-mapping reads from RNA<sub>int</sub> map to tRNA and pro-

tein coding genes in *E. coli*, whereas protein coding genes dominate the cross-mapping reads in RNA<sub>exOMV</sub>, and tRNAs in RNA<sub>exOMV-f</sub>, respectively (Fig. S2 and Table S2).

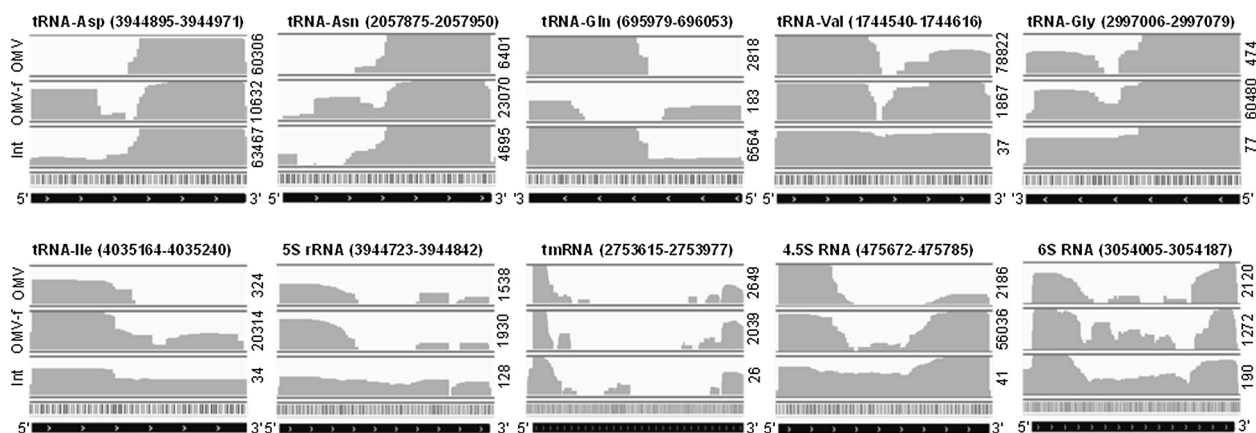
The reads not subjected to cross-mapping to RNAs in the growth medium were mapped to the *E. coli* MG1655 genome and annotated based on known RNA biotypes using the RefSeq annotation (<http://www.ncbi.nlm.nih.gov/refseq/>). Reads mapped to *E. coli* but not to known RNA biotypes, were designated as uncharacterized (Fig. 2B, Table S3). The distributions of different *E. coli* RNA biotypes in RNA<sub>exOMV</sub>, RNA<sub>exOMV-f</sub> and RNA<sub>int</sub> are distinct. However, in all fractions, the majority of the reads mapped to tRNA genes. In order to confirm the *E. coli* origin of these reads, all tRNA reads mapped to *E. coli* were again blasted against the yeast transcriptome obtained from the NCBI database and none of the identified *E. coli* tRNA reads aligned significantly to the yeast transcripts. In RNA<sub>exOMV</sub> and RNA<sub>exOMV-f</sub> the second major RNA biotype is represented by fragments of rRNA, which comprise 6.4% and 7.7% of the total sequenced reads mapped to the *E. coli* genome, whereas only 0.14% are found in RNA<sub>int</sub>. Other noncoding RNAs (ncRNAs) and RNAs encoded by mobile genetic elements (including cryptic prophage) comprise 0.54% and 0.19% of the total mapped sequence reads in RNA<sub>exOMV</sub>, compared to 6.5% and 4.5% in RNA<sub>exOMV-f</sub> and 0.19% and 0.05% in

RNA<sub>int</sub>, respectively. Furthermore, the mapping of reads corresponding to individual RNA biotypes reinforces the intracellular origin of extracellular RNAs while discrete differences in coverage may indicate discrete processing of these RNAs according to their localization (Fig. 2C and D). There are numerous biotypes that are unique to RNA<sub>int</sub>, whereas very few unique biotypes, represented mainly by mRNA fragments, are present in the RNA<sub>ex</sub> (Table S4). On the one hand, the observation of numerous biotypes unique to RNA<sub>int</sub> underpins the notion of selective export of RNA biotypes, analogous to observations in eukaryotes (Hoen et al. 2012). On the other hand, the detection of unique biotypes in RNA<sub>ex</sub> may be due to the intracellular forms of these fragments being of longer length than 60 nt and, thus, not represented in the intracellular smallRNAseq data. Moreover, the fractions of RNA<sub>exOMV</sub> and RNA<sub>exOMV-f</sub> are distinct in their RNA profile as both fractions hold unique biotypes (Fig. 2E, Table S5, S6) leading to distinct distribution patterns of RNA biotypes (Fig. 2B). More specifically, a significant number of unique reads in RNA<sub>exOMV-f</sub> map to mobile genetic elements, ncRNA genes, and protein coding regions, whereas they mainly map to protein coding regions for RNA<sub>exOMV</sub>.

### Analysis of dominant extracellular RNA biotypes

The majority of the sequenced reads represent tRNA fragments (tRFs). These map mostly to the 5' and 3' ends of mature tRNAs (Fig. 3). Despite their high coverage in both RNA<sub>ex</sub> and RNA<sub>int</sub>, the coverages of reads mapping

to tRFs vary among the three different fractions (Table 1). High abundance of full-length tRNAs is only apparent in RNA<sub>int</sub> (Figs. 2A, 3). Based on read coverage, clear differences in the levels of tRFs between RNA<sub>int</sub>, RNA<sub>exOMV</sub>, and RNA<sub>exOMV-f</sub> are apparent. More specifically, the tRFs, which are highly abundant in RNA<sub>int</sub>, are partially reflected in RNA<sub>exOMV</sub>, whereas RNA<sub>exOMV-f</sub> appears significantly different, as determined by comparing the distinct abundance ranks in each fraction. For example, in RNA<sub>exOMV</sub> and RNA<sub>exOMV-f</sub>, tRFs from tRNA-Glu are ranked 1st and 28th, tRFs from tRNA-Asp are ranked 2nd and 8th, tRFs from tRNA-Gln are ranked 5th and 109th, tRFs from tRNA-Asn are ranked 3rd and 3rd, tRFs from tRNA-Ser are ranked 24th and 57th, and tRFs from tRNA-Val are ranked 20th and 2nd, respectively. Comparison among the tRFs obtained from RNA<sub>exOMV</sub>, RNA<sub>exOMV-f</sub> and RNA<sub>int</sub> suggests that the majority of the 5' and 3' tRFs are exported into the extracellular milieu (Fig. 3) as previously observed in eukaryotes (Hoen et al. 2012). The abundances of tRFs in the RNA<sub>ex</sub> inferred from read counts (Table 1) and their read mapping patterns (Fig. 3) indicate that the export of tRFs into the extracellular milieu may involve specific processing as well as selection. Most tRFs are in the size range between 15 and 38 nt, and tRFs in that range have been reported to be functional in bacteria (Haiser et al. 2008) and eukaryotes (Thompson et al. 2008; Lee et al. 2009). In a recent study on human prostate cancer cells, tRFs of length between 17 and 26 nt were identified to constitute the second largest pool of total intracellular RNA, and more specifically tRFs originating from tRNA-Asp, tRNA-Leu, tRNA-Ala, tRNA-Gly, and tRNA-Glu constitute a significant portion of total



**Figure 3.** Coverage of *Escherichia coli* genomic regions by the intra- and extracellular RNA complement. Images were exported from the Integrative Genome Viewer software (<http://www.broadinstitute.org/igv/>). Histograms represent the abundances of RNA reads mapped over the length of the respective genomic regions. Gene names are indicated below the respective mappings. Int, OMV-f and OMV represent the histograms of RNA<sub>int</sub>, RNA<sub>exOMV-f</sub> and RNA<sub>exOMV</sub>, respectively. For each individual genomic region, CPM (counts per million) values were obtained for each type of sample and these values are reported on the right hand side of the images.

**Table 1.** List of abundant tRNA fragments (tRFs).

Fractions	Gene product	Genomic region	No. of reads <sup>2</sup>	Most abundant read (raw counts <sup>3</sup> )
OMV RNA	tRNA-Glu	4	111,349	GGCGGTAACAGGGGT (110,432)
	tRNA-Asp	3	30,817	GCAGGGGGTCGCGGGTTCGAGTCCCCTCCGTTCCGCCA (23,916)
	tRNA-Asn	4	4368	TCCGTATGTCAGTGGTTCGAGTCCAGTCAGAGGAGCCA (2118)
	tRNA-Thr	4	1879	AGTTCGATTCCGGTAGTCGGCACCA (128)
	tRNA-Gln	4	1517	ACCGGCATTCCCTGGTTCGAATCCAGGTACCCAGCCA (465)
OMV-free RNA	tRNA-Thr	4	28,299	GGGTGAGGTCGGCAGT (3732)
	tRNA-Val	7	19,834	TCGAGTCCACTCGGACGCAC (4069)
	tRNA-Asn	4	17,691	GTATGTCAGTGGTTCGAGTCCAGTCAGAGGAGC (2772)
	tRNA-Gly	6	13,466	GCGGGCGTAGTTC AATGGTAGAACGAGAGCTTCCC (3571)
	tRNA-Ala	5	11,776	GGGGCTATAGCTCAGCTGGGAGAGCGCTGCTT (2074)
	tRNA-Ile	3	11,683	AGGCTTGTAGCTCAGGTGGTTAGAGC (4330)
	tRNA-Leu	8	7273	GCCGAAGTGCGCAAATCGGTAGACGCAGTTGATT (1202)
	tRNA-Asp	3	6168	GGAGCGGTAGTTCAGTCGGTTAGA (899)
	tRNA-His	1	5920	GGTGGCTATAGCTCAGTTGGTAGAGCCC (1084)
	Intracellular RNA <sup>1</sup>	tRNA-Glu	4	15,469,939
tRNA-Asp		3	3,908,936	GCAGGGGGTCGCGGGTTCGAGTCCCCTCCGTTCCGCCA (3,049,567)
tRNA-Gln		4	591,759	TTCCGGCATTCCGAGGT (244,848)
tRNA-Asn		4	399,219	TCCGTATGTCAGTGGTTCGAGTCCAGTCAGAGGAGCCA (231,638)
tRNA-Ser		1	148,018	ACCGGCGACCCGAAAGGGTTC (138,124)
tRNA-Val		7	105,509	GGAGGGGGTCGCGGGTTCGATCCCCTCATCACCACCACT (56,389)

<sup>1</sup>The intracellular RNA fraction contains full-length tRNAs along with tRFs.

<sup>2</sup>Total number of reads mapped to the genomic region.

<sup>3</sup>Coverage of individual reads.

tRFs (Lee et al. 2009). During stress, especially oxidative stress, the levels of tRFs have been shown to increase in yeast (Thompson et al. 2008). Here, increased amounts of tRFs are typically associated with specific cleavage products of tRNAs and such findings have also been reported for plants and humans (Thompson et al. 2008). In *Streptomyces coelicolor*, the abundances of distinct tRFs follow a growth pattern, and abundant tRFs originate from tRNA-Asn, tRNA-Gln, tRNA-Gly, tRNA-Val, and others (Haier et al. 2008). Analogous to these previous findings in human, yeast, and bacterial cells, our results suggest that specific tRFs may play a range of important functional roles including possible intercellular interaction.

Apart from tRFs, the second major constituents of extracellular RNAs are fragments of rRNA (Fig. 2B). These rRNA fragments map to different discrete regions of the 23S rRNA, 16S rRNA, and 5S rRNA genes in both RNA<sub>ex</sub> and RNA<sub>int</sub>. These rRNA fragments may represent rRNA degradation products. However, we identified a 23 nt long fragment of the 5' end of the 5S rRNA [TGCTGGCGCCGTAGCGGGTG (+st)] and two 17 nt long fragments of the 3' end of the 5S rRNA [GTGGGGTCTCCCCATGC (+st) and AGGGAAGTGC-CAGGCAT (+st)], which seem to be exclusively exported as suggested by the mapping patterns of read sequences (Fig. 3). rRNA fragments of similar size are known to be constituents of small RNA fractions not only in bacteria

(Haier et al. 2008; Kang et al. 2013) but also in eukaryotes (Thompson et al. 2008; Hoen et al. 2012). Significant amounts of various fragments of bacterial rRNA have also been observed in human plasma (Wang et al. 2012). A recent study has demonstrated that a conserved 23S rRNA motif "CGGAAAGACC" is recognized by the TLR13 receptor in mice, and modifications in the motif allows bacteria, for example, erythromycin-resistant *Staphylococcus aureus*, to bypass host immune recognition (Oldenburg et al. 2012). Although TLR13 is not found in humans, it is possible that human pattern recognition receptors (PRRs) exist which may ensure similar functions in humans. The "CGGAAAGACC" motif was identified in RNA<sub>ex</sub> and RNA<sub>int</sub> albeit with low coverage values of 19 in RNA<sub>exOMV</sub>, 24 in RNA<sub>exOMV-f</sub> and 43 in RNA<sub>int</sub>, respectively. From our own data, it is intriguing that the tRF originating from tRNA-Glu has the sequence "CGGTAACAGG" which differs from the sequence of the motif described for fragments of the 23S rRNA by the following highlighted (bold, underlined) nucleotides: "CGG**AAAGACC**." For the tRF, thymine is substituted by adenine, whereas cytosine is substituted by guanine and vice versa. This sequence is present on a 15 nt long tRF, which was found in RNA<sub>exOMV</sub> (coverage of 111,349), RNA<sub>exOMV-f</sub> (coverage of 270), and RNA<sub>int</sub> (coverage of 15,469,939) (Table 1). Further investigation on the tRF originating from tRNA-Glu may identify new PRRs,



**Table 2.** List of abundant small non-coding RNAs.

Fractions	Gene product	No. of reads <sup>1</sup>	Most abundant read (raw counts <sup>2</sup> )	Rank <sup>3</sup>
OMV RNA	tmRNA 105a RNA ( <i>ssrA</i> )	452	GGGGCTGATTCTGGATTTCGACGGG (108)	124
	4.5S sRNA component of Signal Recognition Particle ( <i>ffs</i> )	373	GGGGGCTCTGTTGGTTCTCCCGCAACGCT (82)	95
	6S RNA inhibits RNA polymerase promoter binding ( <i>ssrS</i> )	361	CATCTCGGAGATCCCC (52)	47
OMV-free RNA	4.5S sRNA component of Signal Recognition Particle ( <i>ffs</i> )	10,743	GGGGGCTCTGTTGGTTCTCCCGCAACGCT (1383)	95
	Novel sRNA ( <i>ryeA</i> )	405	CAAGAGCCATTTCCCTGGACCG (103)	40
	tmRNA 105a RNA ( <i>ssrA</i> )	391	GGGGCTGATTCTGGATTTCGACGGGA (87)	124
	Novel sRNA ( <i>ryfD</i> )	369	AAGACGATCCGGTACGC (108)	54
	6S RNA inhibits RNA polymerase promoter binding ( <i>ssrS</i> )	244	TTTCTCTGAGATGTTTCGCAAGC (19)	47
	sRNA antisense regulator of <i>gadAB</i> transcriptional activator <i>GadX</i> mRNA ( <i>gadY</i> )	230	TCTGGAGACGGCAGACT (104)	100
	sRNA antisense regulator represses <i>oppA</i> , <i>dppA</i> , <i>gltI</i> and <i>livJ</i> expression ( <i>ryeB</i> )	124	AGAGAGCCGTGCGCTAAAAGT (61)	72
	Novel sRNA ( <i>ryeA</i> )	4907	AGAGAGCCGTGCGCTAAAAGT (477)	1
Intracellular RNA	sRNA antisense regulator blocking <i>mokB</i> ( <i>sokB</i> )	4730	ATTCGTTGGCCTCGGTTGA (3580)	2
	sRNA antisense regulator affects <i>LdrB</i> translation ( <i>rdlB</i> )	4413	GTCAGGTTTTACCTCTCAACGTGCGGGGGTTTTCTC (2180)	3
	6S RNA inhibits RNA polymerase promoter binding ( <i>ssrS</i> )	4050	TTTCTCTGAGATGTTTCGCAAGCGGGCCAGTCC (287)	4
	Novel sRNA ( <i>ryfD</i> )	2575	GGAGGAGTTATGCGTCTGGATCGTCTTACT (536)	5
	sRNA antisense regulator of OM chitoporin ( <i>chiX</i> )	2396	TCTTTGACGGGCCAATAGCGATATTGGCCATTTTTT (413)	6

<sup>1</sup>Total number of reads mapped to the genomic region.

<sup>2</sup>Coverage of individual reads.

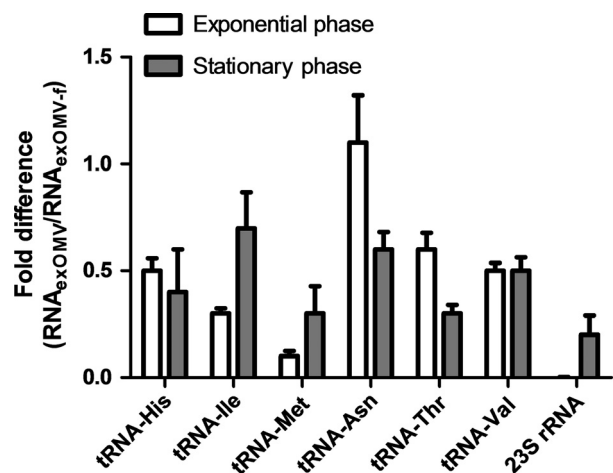
<sup>3</sup>Ranks provided for small non-coding RNAs derived from RNA<sub>exOMV</sub> and RNA<sub>exOMV-f</sub> are ranked according to their abundance in RNA<sub>int</sub>.

which may be involved in exogenous RNA recognition by hosts including human.

Other small ncRNAs such as parts of tmRNA, 4.5S RNA (SRP-RNA), 6S RNA, and others are also selectively enriched in the RNA<sub>ex</sub> (Table 2 and Fig. 3). Further analysis confirmed that 29 and 23 nt long small ncRNAs [GGGGGCTCTGTTGGTTCTCCCGCAACGCT (+st) and TGTTAGCTGGCAGGGCCCCACCC (+st)] result from the 5' and 3' ends of the 4.5S RNA, respectively. These are selectively exported in OMVs but are also found in RNA<sub>exOMV-f</sub>. In bacteria, 4.5S RNA is bound to the Ffh protein and constitutes the signal recognition particle (SRP). One of the functions of SRP involves the targeting of membrane and secretory proteins to the plasma membrane of bacteria following the recognition of a signal peptide sequence on the nascent polypeptide (Wassarman et al. 1999; Batey 2000). The SRP is also conserved in mammals (Batey 2000). In particular, immune cells are known to release fragments of SRP-RNA into the extracellular space (Hoen et al. 2012). Although the functions and the sequence information of exported fragments of SRP-RNA have not yet been described in detail, this does suggest a commonality in exported RNA in both bacteria and eukaryotes. Likewise, 22 and 16 nt long small ncRNAs [TTTCTCTGAGATGTTTCGCAAGC (+st) and CATCTCGGAGATTTCCC (+st)] from the 5' and 3' ends of

the 6S RNA, respectively, and a 25 nt long small ncRNA [GGGGCTGATTCTGGATTTCGACGGGA (+st)] from the 5' end of tmRNA are also released by *E. coli* into its extracellular milieu. 6S RNA and tmRNA are not well conserved in eukaryotes (Wassarman et al. 1999) but play an important role in transcription and translation in bacteria (Wassarman and Storz 2000; Janssen and Hayes 2013). The presence of fragments of 6S RNA and tmRNA in the extracellular milieu, similar to fragments of SRP-RNAs, suggests selective mechanisms for their export.

In our analysis, we have found that a number of sequencing reads in both the RNA<sub>ex</sub> and RNA<sub>int</sub> mapped to mobile genetic elements of the *E. coli* MG1655 genome, primarily to the cryptic prophages CP4-6, DLP12, and e14, with the majority of those exported reads in the RNA<sub>exOMV-f</sub> fraction mapping to cryptic prophage CP4-6. Cryptic prophages provide selective advantages to their hosts in different conditions (Wang et al. 2010) and they may confer protection against external bacteriophages (Casjens 2003). The two most abundant 35 nt and 28 nt long small ncRNAs [GAAAAAAGTGGTACCGCCAAGAC TACACACAGCAT (+st) and AATGACGCCCGGAGTGTGCAGTCCCGG (+st)] are part of e14 and were found in high relative abundance in both RNA<sub>int</sub> and RNA<sub>ex</sub>. Similar to the CRISPR/Cas system where transcripts of virus-originating spacer DNA sequence defend bacteria

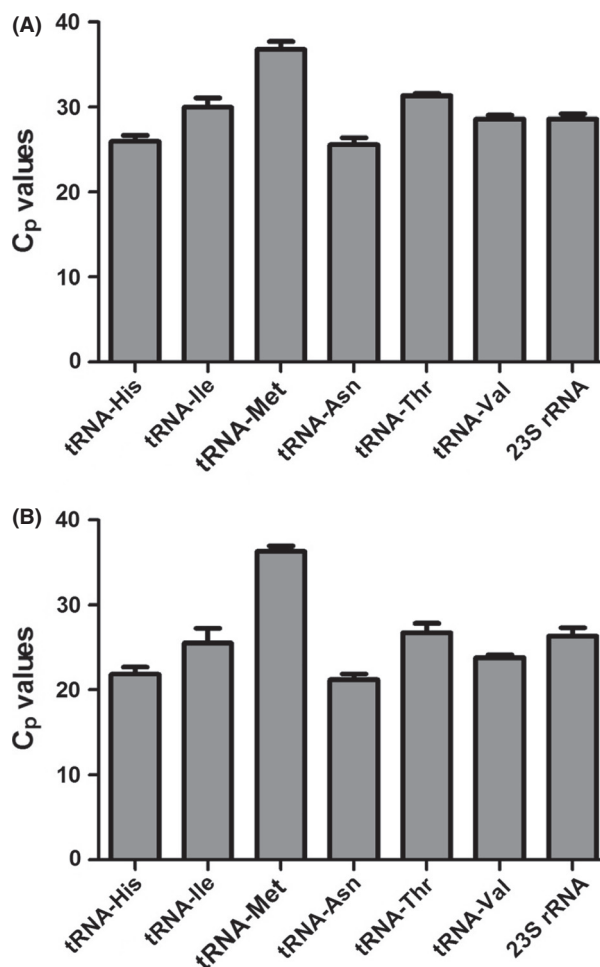


**Figure 4.** RT-qPCR validation of identified small noncoding RNAs. Extracellular RNAs quantified from exponentially growing *Escherichia coli* cultures and overnight grown cultures using RT-qPCR. Error bars represent standard error of the mean ( $n = 3$ ).

from invading viruses, small ncRNAs of cryptic prophage origin may play an analogous role in defense and/or virulence.

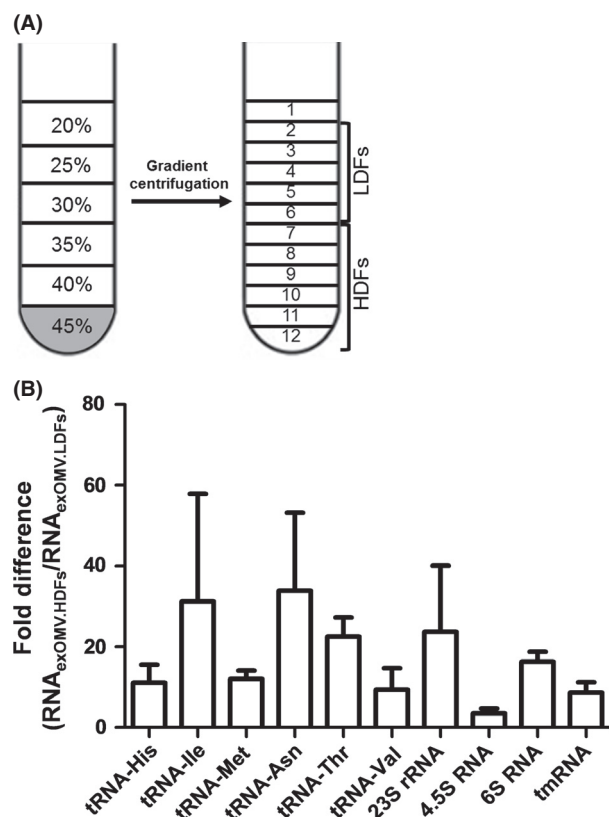
To identify the potential changes in  $RNA_{ex}$  at different growth phases and to validate the results obtained from RNAseq, equal amount of RNAs from  $RNA_{exOMV}$  and  $RNA_{exOMV-f}$  fractions, extracted from exponentially growing bacteria and overnight grown cultures were analyzed by RT-qPCR (Fig. 4). For all the samples analyzed, the amounts of analyzed small ncRNAs are higher in  $RNA_{exOMV}$  in comparison with the  $RNA_{exOMV-f}$  and the observed differences are consistent in both exponential and stationary growth phases (Fig. 4). By and large, the amount of small RNA fragments released is also consistent in both bacterial growth phases with the exception of RNA fragments originating from tRNA-Asn and 23S rRNA (Fig. 4). For additional validation of these findings, we also extracted small RNA from the OMV and OMV-f fractions obtained from bacterial cultures grown in synthetic M9 minimal medium (RNA-free) and analyzed these in an analogous way. The results from these experiments highlight the presence of the previously identified small RNAs in the M9 OMV and OMV-f fractions (Fig. 5), which lends additional independent support to the previous RT-qPCR results from samples obtained from cultures grown in LB medium. Taken together, these RT-qPCR results validate the previous observations based on deep-sequencing analyses and put forward a notion of continuous release of small RNA into the extracellular milieu.

The secretion of vesicles is a common, conserved process which exists in both eukaryotes and prokaryotes



**Figure 5.** RT-qPCR validation of small noncoding RNAs in the (A) OMV and (B) OMV-f fractions obtained from bacterial cultures grown in M9 minimal medium. Error bars represent standard error of the mean ( $n = 3$ ).

(Mashburn-Warren et al. 2008; Chen et al. 2012). Vesicles are known to contain signaling molecules (Mashburn-Warren et al. 2008; Hannafon and Ding 2013) and are thought to be important mediators of intercellular information exchange via transfer of small RNA in eukaryotes (Valadi et al. 2007). To further characterize the RNA of the *E. coli* OMV fraction, we subfractionated the OMV fraction into low- and HDFs (Fig. 6A). The LDFs contain “pure OMVs” (Horstman et al. 2000), whereas the HDFs may contain nucleic acid-protein complexes and RNA associated with the OMV surface dislodged during the fractionation process, in addition to OMVs (Fig. S3). RNA from these fractions was further analyzed by a combination of RT-qPCR and high-throughput sequencing. The amounts of previously identified extracellular RNAs comprising tRFs, fragments of rRNA, and small ncRNAs are higher in the HDFs compared to the LDFs,



**Figure 6.** Analysis of small noncoding RNAs in subfractionated OMV samples. (A) Representation of the differential centrifugation scheme used to subfractionate the obtained- *Escherichia coli* OMV fraction into low- and high-density fractions (LDF and HDF), respectively. (B) Comparison of the relative abundances of identified small noncoding RNAs between HDFs (less pure OMV fractions) and LDFs (pure OMV fractions). RNA quantities were determined by RT-qPCR. LDFs: low-density fractions, HDFs: high-density fractions. Error bars represent the standard error of the mean ( $n = 3$ ).

although the amount of individual small ncRNAs varies significantly between samples (Fig. 6B). Even though the amount of small RNAs is higher in the HDFs, all RT-qPCR analyzed small RNAs are present in the LDFs (pure OMV fraction). The RNA obtained from the pure OMV fraction was further analyzed by deep sequencing. Similar to the previous observations of RNA<sub>exOMV</sub>, tRFs, and fragments of rRNA along with other small RNAs from coding regions dominate this OMV fraction (Table S7). Although, according to these results, OMVs contain RNA, our results also indicate that RNA may be associated with the OMV surface as previously described for DNA in *Moraxella catarrhalis*, where OMV-associated DNA is recognized by toll-like receptor 9 (TLR9) inside immune cells and thereby stimulating an immune response (Vidakovics et al. 2010). Similar to the situation described for DNA, OMV surface-associated RNA may be protected

from degradation and may be involved in remote immune system stimulation through OMV-mediated transport. However, the determination of the precise localization of OMV-associated RNAs will require further in-depth studies.

### Identification of previously uncharacterized sRNAs

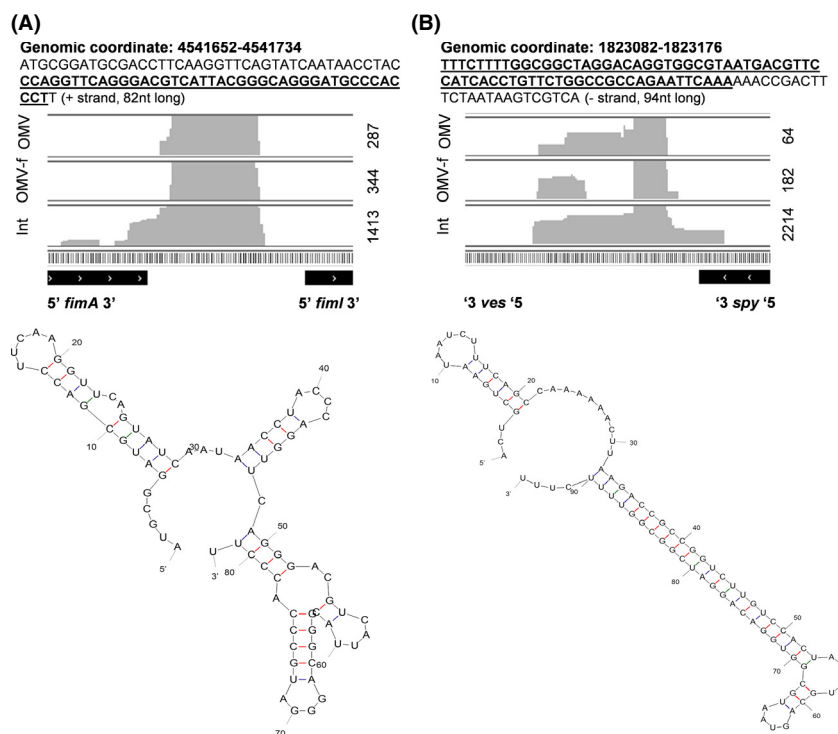
By careful inspection of read mapping profiles of short RNA, we identified genomic regions with high read coverages. These regions include two so far uncharacterized genomic regions in the *E. coli* genome (Fig. 7). Both of these intergenic regions, *fimA-fimL* and *ves-spy*, have previously been found to exhibit high levels of expression (Kawano et al. 2005). In particular, a 35 nt long transcript originating from the *fimA-fimL* intergenic region was previously found to be abundant in *E. coli* in stationary growth phase (Kawano et al. 2005). Interestingly, we have observed transcripts of lengths between 31 and 37 nt from the same genomic region, which are specifically released into the extracellular milieu (Fig. 7A). In addition, a previously uncharacterized 82 nt long transcript from the same region was identified in RNA<sub>int</sub> close to what has previously been reported in the ECSBrowser (<http://rna.iab.keio.ac.jp/>, (Shinhara et al. 2011)).

Similarly, high levels of expression were observed for the intergenic region between *ves-spy*, where the majority of the reads correspond to a ~66 nt long transcript which appears to originate from an ~94 nt long genomic region (Fig. 7B). In a previous study, high expression of a 45 nt long transcript and lower expression of 70 and 120 nt long transcripts from this intergenic region was observed at early stationary growth phase of *E. coli* (Kawano et al. 2005). In ECSBrowser, the transcript length is reported to be 86 nt between the genomic region of *ves* and *spy* genes. Interestingly, we identified a specific 17 nt long transcript [GAC-CGGCGGTCTTAAGT (+st)] encoded by this region which was present in both RNA<sub>exOMV-f</sub> and RNA<sub>exOMV</sub> (Fig. 7B).

### Conclusions

We have carried out an in-depth characterization of the extracellular RNA complement of the enteric bacterium *E. coli*. We found that the secreted RNA complement is largely smaller than 60 nt, that it is enriched in nonprotein coding RNAs, that it is distinct from the intracellular RNA pool, and that the identities and abundances of the secreted RNA differ between the extracellular milieu and outer membrane vesicular fractions.

The observations reported here raise a number of important questions. These include the following: (1) How are non-OMV-associated bacterial RNAs transported



**Figure 7.** Identification of uncharacterized noncoding RNAs by mapping RNAseq data onto the *Escherichia coli* genome. (A) *fimA*-*fimL* and (B) *ves*-*spy*. In (A) and (B) the genomic coordinates represent the region where sequence reads were mapped. The nucleotide sequence is provided below the genomic coordinate. Underlined areas within the nucleotide sequence represent their overrepresentation in the RNAseq libraries. Mappings are provided for *fimA*-*fimL* and *ves*-*spy*. Flanking regions are named with genes. Int, OMV-f, and OMV represent the histogram of RNA<sub>int</sub>, RNA<sub>exOMV-f</sub>, and RNA<sub>exOMV</sub>, respectively. For each individual genomic region, CPM (counts per million) values were obtained for each type of sample and these values are reported on the right-hand side of the images. Noncoding RNA sequences were structured in Mfold web server (<http://mfold.rna.albany.edu/?q=mfold/RNA-Folding-Form>).

across bacterial membranes and subsequently released into the extracellular space? (2) How are the RNAs selected for export? (3) Is there a difference in the extracellular RNA profile between pathogenic and non-pathogenic bacteria? Apart from seeking answers to these questions, extracellular bacterial RNAs should be investigated for their potential protein binding partners, host target cells and their potential as biomarkers. Understanding the role of bacteria-derived exogenous RNA in host–microbe interactions, in pathogenesis as well as in mutualism, will elucidate new mechanisms and perhaps allow the identification of new drug targets and/or the development of RNA-based vaccines. Further investigations in the field of extracellular bacterial RNAs are clearly needed to shed light on their potential role as mediators of microbe–microbe and host–microbe intercellular communication.

## Acknowledgments

We thank E. Lentzen (Centre de Recherche Public – Gabriel Lippmann) for experimental support. We are grateful

to C. Linster and P. Shah for helpful discussions. The present work was supported by an ATTRACT programme grant (A09/03) CORE programme grant (CORE11/BM/1186762), European Union Joint Programming in Neurodegenerative Diseases grant (INTER/JPND/12/01), and Proof-of-Concept grant (PoC/13/02) to PW, funded by the Luxembourg National Research Fund (FNR). Further financial support for the work was through FNR AFR fellowships to JF (Ref. no. 5821107) and DY (Ref. no.1180851).

## Conflict of Interest

None declared.

## References

- Batey, R. T., R. P. Rambo, L. Lucast, B. Rha, and J. A. Doudna. 2000. Crystal structure of the ribonucleoprotein core of the signal recognition particle. *Science* 287:1232–1239. Available at <http://www.sciencemag.org/cgi/doi/10.1126/science.287.5456.1232> (accessed 20 December 2013).

- Bauman, S. J., and M. J. Kuehn. 2006. Purification of outer membrane vesicles from *Pseudomonas aeruginosa* and their activation of an IL-8 response. *Microbes Infect.* 8:2400–2408.
- Billar, S. J., F. Schubotz, S. E. Roggensack, A. W. Thompson, R. E. Summons, and S. W. Chisholm. 2014. Bacterial vesicles in marine ecosystems. *Science* 343:183–186. Available at <http://www.sciencemag.org/cgi/doi/10.1126/science.1243457> (accessed 9 January 2014).
- Carmen Schwegheimer, C. J. S. A., and M. J. Kuehn. 2013. Envelope control of outer membrane vesicle production in gram-negative bacteria. *Biochemistry* 52:3031–3040.
- Casjens, S. 2003. Prophages and bacterial genomics: what have we learned so far? *Mol. Microbiol.* 49:277–300. Available at <http://doi.wiley.com/10.1046/j.1365-2958.2003.03580.x> (accessed 20 December 2013).
- Casson, C. N., A. M. Copenhaver, E. E. Zwack, H. T. Nguyen, T. Strowig, B. Javdan, et al. 2013. Caspase-11 activation in response to bacterial secretion systems that access the host cytosol. *PLoS Pathog.* 9:e1003400. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3675167&tool=pmcentrez&rendertype=abstract> (accessed 11 December 2013).
- Chen, X., H. Liang, J. Zhang, K. Zen, C. Y., Zhang. 2012. Secreted microRNAs: a new form of intercellular communication. *Trends Cell Biol.* 22:125–132. Available at <http://www.ncbi.nlm.nih.gov/pubmed/22260888> (accessed 20 December 2013).
- Chitcholtan, K., M. B. Hampton, and J. I. Keenan. 2008. Outer membrane vesicles enhance the carcinogenic potential of *Helicobacter pylori*. *Carcinogenesis* 29:2400–2405. Available at <http://www.ncbi.nlm.nih.gov/pubmed/18784355> (accessed 13 February 2014).
- De Hoon, M. J., R. J. Taft, T. Hashimoto, M. Kanamori-Katayama, H. Kawaji, M. Kawano, et al. 2010. Cross-mapping and the identification of editing sites in mature microRNAs in high-throughput sequencing libraries. *Genome Res.* 20:257–264. Available at <http://www.ncbi.nlm.nih.gov/pubmed/20051556>.
- Dorward, D. W., C. F. Garon, and R. C. Judd. 1989. Export and intercellular transfer of DNA via membrane blebs of *Neisseria gonorrhoeae*. *J. Bacteriol.* 171:2499–2505.
- El Andaloussi, S., I. Mäger, X. O. Breakefield, and M. J. Wood. 2013. Extracellular vesicles: biology and emerging therapeutic opportunities. *Nat. Rev. Drug Discov.* 12:347–357. Available at <http://www.ncbi.nlm.nih.gov/pubmed/23584393> (accessed 13 December 2013).
- Furusawa, Y., Y. Obata, S. Fukuda, T. A. Endo, G. Nakato, D. Takahashi et al. 2013. Commensal microbe-derived butyrate induces the differentiation of colonic regulatory T cells. *Nature* 504:446–450. Available at <http://www.ncbi.nlm.nih.gov/pubmed/24226770> (accessed 10 January 2014).
- Haiser, H. J., F. V. Karginov, G. J. Hannon, and M. A. Elliot. 2008. Developmentally regulated cleavage of tRNAs in the bacterium *Streptomyces coelicolor*. *Nucleic Acids Res.* 36:732–741. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2241913&tool=pmcentrez&rendertype=abstract> (accessed 30 December 2013).
- Hannafon, B. N., and W. Q. Ding. 2013. Intercellular communication by exosome-derived microRNAs in cancer. *Int. J. Mol. Sci.* 14:14240–14269. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3742242&tool=pmcentrez&rendertype=abstract> (accessed 20 December 2013).
- Hoen, E. N. M. N. T., H. P. Buermans, M. Waasdorp, W. Stoorvogel, M. H. Wauben, and P. A. Hoen. 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. Available at <http://nar.oxfordjournals.org/content/40/18/9272.abstract> (accessed 2 September 2013).
- Horstman, A. L., and M. J. Kuehn. 2000. Enterotoxigenic *Escherichia coli* secretes active heat-labile enterotoxin via outer membrane vesicles. *J. Biol. Chem.* 275:12489–12496. Available at <http://www.jbc.org/cgi/doi/10.1074/jbc.275.17.12489> (accessed 20 December 2013).
- Hughes, D. T., and V. Sperandio. 2008. Inter-kingdom signalling: communication between bacteria and their hosts. *Nat Rev Microbiol.* 6:111–20.
- Janssen, B. D., and C. S. Hayes. 2013. The tmRNA ribosome rescue system. *Adv. Protein Chem. Struct. Biol.* 86:151–191.
- Jun, S. H., J. H. Lee, B. R. Kim, S. I. Kim, T. I. Park, J. C. Lee, and Y. C. Lee. 2013. Acinetobacter baumannii outer membrane vesicles elicit a potent innate immune response via membrane proteins. *PLoS ONE* 8:e71751. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3743744&tool=pmcentrez&rendertype=abstract> (accessed 2 September 2013).
- Kang, S.M., J. W. Choi, Y. Lee, S. H. Hong, and H. J. Lee. 2013. Identification of microRNA-size, small RNAs in *Escherichia coli*. *Curr. Microbiol.* 67:609–613. Available at <http://www.ncbi.nlm.nih.gov/pubmed/23783561> (accessed 8 August 2013).
- Kawano, M., A. A. Reynolds, J. Miranda-Rios, and G. Storz. 2005. Detection of 5'- and 3'-UTR-derived small RNAs and cis-encoded antisense RNAs in *Escherichia coli*. *Nucleic Acids Res.* 33:1040–1050. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=549416&tool=pmcentrez&rendertype=abstract> (accessed 20 December 2013).
- Kesty, N. C., K. M. Mason, M. Reedy, S. E. Miller, and M. J. Kuehn. 2004. Enterotoxigenic *Escherichia coli* vesicles target toxin delivery into mammalian cells. *EMBO J.* 23:4538–4549. Available at <http://www.pubmedcentral.nih.gov>



- gov/articlerender.fcgi?artid=533055&tool=pmcentrez&rendertype=abstract (accessed 2 September 2013).
- Kogure, T., W. L. Lin, I. K. Yan, C. Braconi, and T. Patel. 2011. Intercellular nanovesicle-mediated microRNA transfer: a mechanism of environmental modulation of hepatocellular cancer cell growth. *Hepatology* 54:1237–1248. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3310362&tool=pmcentrez&rendertype=abstract> (accessed 9 August 2013).
- Lee, V. T., and O. Schneewind. 2001. Protein secretion and the pathogenesis of bacterial infections. *Genes Dev.* 15:1725–1752. Available at <http://www.ncbi.nlm.nih.gov/pubmed/11459823> (accessed 19 December 2013).
- Lee, Y. S., Y. Shibata, A. Malhotra, and A. Dutta. 2009. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev.* 23:2639–2649. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2779758&tool=pmcentrez&rendertype=abstract> (accessed 13 December 2013).
- Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, et al. 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079.
- Mashburn-Warren, L., R. J. C. McLean, and M. Whiteley. 2008. Gram-negative outer membrane vesicles: beyond the cell surface. *Geobiology* 6:214–219. Available at <http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=18459967&retmode=ref&cmd=prlinks> (accessed 7 August 2013).
- Molloy, S. 2010. Quorum sensing: setting the threshold. *Nat. Rev. Microbiol.* 8:388–389. Available at <http://www.nature.com/nrmicro/journal/v8/n6/full/nrmicro2374.html>.
- Nicola, A. M., S. Frases, and A. Casadevall. 2009. Lipophilic dye staining of *Cryptococcus neoformans* extracellular vesicles and capsule. *Eukaryot. Cell* 8:1373–1380. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2747831&tool=pmcentrez&rendertype=abstract> (accessed 2 September 2013).
- Oldenburg, M., A. Krüger, R. Ferstl, A. Kaufmann, G. Nees, A. Sigmund, et al. 2012. TLR13 recognizes bacterial 23S rRNA devoid of erythromycin resistance-forming modification. *Science* 337:1111–1115. Available at <http://www.ncbi.nlm.nih.gov/pubmed/22821982> (accessed 30 December 2013).
- Park, K. S., K. H. Choi, Y. S. Kim, B. S. Hong, O. Y. Kim, J. H. Kim, et al. 2010. Outer membrane vesicles derived from *Escherichia coli* induce systemic inflammatory response syndrome. *PLoS ONE* 5:e11334. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2893157&tool=pmcentrez&rendertype=abstract> (accessed 2 September 2013).
- Parker, H., K. Chitcholtan, M. B. Hampton, and J. I. Keenan. 2010. Uptake of *Helicobacter pylori* outer membrane vesicles by gastric epithelial cells. *Infect. Immun.* 78:5054–5061. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2981328&tool=pmcentrez&rendertype=abstract> (accessed 2 September 2013).
- Pavankumar, A. R., S. P. Ayyappasamy, and K. Sankaran. 2012. Small RNA fragments in complex culture media cause alterations in protein profiles of three species of bacteria. *Biotechniques* 52:167–172. Available at <http://www.ncbi.nlm.nih.gov/pubmed/22401549> (accessed 2 September 2013).
- Pérez-Cruz, C., O. Carrión, L. Delgado, G. Martínez, C. López-Iglesias, and E. Mercade. 2013. A new type of outer membrane vesicles produced by the Gram-negative bacterium *Shewanella vesiculosa* M7T: implications for DNA content. *Appl. Environ. Microbiol.* 79:1874–81. Available at <http://aem.asm.org/content/early/2013/01/08/AEM.03657-12.abstract> (accessed 2 September 2013).
- Quinlan, A. R. and I. M. Hal. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–842.
- Renelli, M., V. Matias, R. Y. Lo, and T. J. Beveridge. 2004. DNA-containing membrane vesicles of *Pseudomonas aeruginosa* PAO1 and their genetic transformation potential. *Microbiology* 150(Pt. 7):2161–2169. Available at [http://mic.sgmjournals.org/content/150/7/2161.abstract?ikey=623d428acd3930ca99099e7fa84e69ebcd0e40ff&keytype2=tf\\_ipsecsha](http://mic.sgmjournals.org/content/150/7/2161.abstract?ikey=623d428acd3930ca99099e7fa84e69ebcd0e40ff&keytype2=tf_ipsecsha) (accessed 11 August 2013).
- Roy, K., D. J. Hamilton, G. P. Munson, and J. M. Fleckenstein. 2011. Outer membrane vesicles induce immune responses to virulence proteins and protect against colonization by enterotoxigenic *Escherichia coli*. *Clin. Vaccine Immunol.* 18:1803–1808. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3209013&tool=pmcentrez&rendertype=abstract> (accessed 2 September 2013).
- Shen, Y., M. L. Giardino Torchia, G. W. Lawson, C. L. Karp, J. D. Ashwell, and S. K. Mazmanian. 2012. Outer membrane vesicles of a human commensal mediate immune regulation and disease protection. *Cell Host Microbe* 12:509–520. Available at <http://www.ncbi.nlm.nih.gov/pubmed/22999859> (accessed 18 August 2013).
- Shinohara, A., M. Matsui, K. Hiraoka, W. Nomura, R. Hirano, K. Nakahigashi, et al. 2011. Deep sequencing reveals as-yet-undiscovered small RNAs in *Escherichia coli*. *BMC Genom.* 12:428. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3175480&tool=pmcentrez&rendertype=abstract> (accessed 20 December 2013).
- Thay, B., S. N. Wai, and J. Oscarsson. 2013. *Staphylococcus aureus* a-toxin-dependent induction of host cell death by membrane-derived vesicles. *PLoS ONE* 8:e54661. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3561366&tool=pmcentrez&rendertype=abstract> (accessed 9 August 2013).

- Théry, C., L. Zitvogel, and S. Amigorena. 2002. Exosomes: composition, biogenesis and function. *Nat. Rev. Immunol.* 2:569–579. Available at <http://www.ncbi.nlm.nih.gov/pubmed/12154376> (accessed 12 December 2013).
- Thompson, D. M., C. Lu, P. J. Green, and R. Parker. 2008. tRNA cleavage is a conserved response to oxidative stress in eukaryotes. *RNA* 14:2095–2103.
- Tseng, T.T., B. M. Tyler, and J. C. Setubal. 2009. Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC Microbiol.* 9:1471–2180. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2654662&tool=pmcentrez&rendertype=abstract> (accessed 11 December 2013).
- Valadi, H., K. Ekström, A. Bossios, M. Sjöstrand, J. J. Lee, and J. O. Lötvall. 2007. Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. *Nat. Cell Biol.* 9:654–659. Available at <http://www.ncbi.nlm.nih.gov/pubmed/17486113> (accessed 20 January 2014).
- Vidakovics, M. L., J. Jendholm, M. Mörgelin, A. Månsson, C. Larsson, L. O. Cardell, and K. Riesbeck. 2010. B cell activation by outer membrane vesicles—a novel virulence mechanism. *PLoS Pathog.* 6:e1000724. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2799554&tool=pmcentrez&rendertype=abstract> (accessed 20 December 2013).
- Wai, S. N., B. Lindmark, T. Söderblom, A. Takade, M. Westermark, and J. Oscarsson, et al. 2003. Vesicle-mediated export and assembly of pore-forming oligomers of the enterobacterial ClyA cytotoxin. *Cell* 115:25–35. Available at <http://www.ncbi.nlm.nih.gov/pubmed/14532000>.
- Wang, X., Y. Kim, Q. Ma, S. H. Hong, K. Pokusaeva, J. M. Sturino, and T. K. Wood. 2010. Cryptic prophages help bacteria cope with adverse environments. *Nat. Commun.* 1:147. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3105296&tool=pmcentrez&rendertype=abstract> (accessed 20 December 2013).
- Wang, K., H. Li, Y. Yuan, A. Etheridge, Y. Zhou, D. Huang, et al. 2012. The complex exogenous RNA spectra in human plasma: an interface with human gut biota? *PLoS ONE* 7: e51009. Available at <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3519536&tool=pmcentrez&rendertype=abstract> (accessed 11 December 2013).
- Wassarman, K. M., and G. Storz. 2000. 6S RNA regulates *E. coli* RNA polymerase activity. *Cell* 101:613–623. Available at <http://www.ncbi.nlm.nih.gov/pubmed/10892648>.
- Wassarman, K. M., A. Zhang, and G. Storz. 1999. Small RNAs in *Escherichia coli*. *Trends Microbiol.* 7:37–45. Available at <http://www.ncbi.nlm.nih.gov/pubmed/10068996>.
- Waters, C. M., and B. L. Bassler. 2005. Quorum sensing: cell-to-cell communication in bacteria. *Ann. Rev. Cell Dev. Biol.* 21:319–346. Available at <http://www.ncbi.nlm.nih.gov/pubmed/16212498> (accessed 11 December 2013).

## Supporting Information

The raw sequence libraries are deposited in the Sequence Read Archive database (<http://www.ncbi.nlm.nih.gov/sra>) under experiment accession number SRX425464. Additional Supporting Information may be found in the online version of this article:

**Figure S1.** Electropherograms of *Escherichia coli* RNA fractions above the length of 200 nucleotides. (A) OMV fraction, (B) OMV-free fraction, (C) intracellular RNA fraction.

**Figure S2.** Distribution of cross-mapping reads.

**Figure S3.** Comparative analysis of protein fractions extracted from the OMV high density fractions (HDFs), the OMV low density fractions (LDFs) and whole cells (WC). Molecular markers are showed on the right side of the gel image.

**Table S1.** Analysis of sequence reads.

**Table S2.** Genomic features and corresponding read counts of cross-mapping reads.

**Table S3.** Genomic features and corresponding read counts.

**Table S4.** Unique biotypes represented by RNA<sub>ex</sub>.

**Table S5.** Unique biotypes represented by RNA<sub>exOMV</sub> in comparison with RNA<sub>exOMV-f</sub>.

**Table S6.** Unique biotypes represented by RNA<sub>exOMV-f</sub> in comparison with RNA<sub>exOMV</sub>.

**Table S7.** Genomic features and corresponding read counts in low density fractions (LDFs).