Supplementary Data for

"A multi-dimensional platform for the purification of non-coding RNA species"

Yok Hian Chionh[1,2], Chia-Hua Ho[1], Dumnoensun Pruksakorn[3,5,†], Ramesh Indrakanti[3], Chee Sheng Ng[1,4], Fabian Hia[1], Megan E. McBee[3], Dan Su[3], Yan Ling Joy Pang[3], Chen Gu[3], Hongping Dong[6], Erin Prestwich[3], Pei-yong Shi[6], Peter Preiser[4],Sylvie Alonso[2], Peter C. Dedon[1,4*]

[1]Singapore-MIT Alliance for Research and Technology, Campus for Research Excellence and Technological Enterprise, Singapore

[2]Department of Microbiology and Immunology Programme, Center for Life Sciences, National University of Singapore, Singapore

[3]Department of Biological Engineering and Center for Environmental Health Sciences, Massachusetts Institute of Technology, Cambridge, MA

[4]School of Biological Sciences, Nanyang Technological University, Singapore

[5]Applied Biological Science Program, Chulabhorn Graduate Institute, Bangkok, Thailand

[6]Novartis Institute for Tropical Disease, Singapore

[†] Current address: Dumnoensun Pruksakorn, Department of Orthopedics, Faculty of Medicine, Chiang Mai University, Tambon Sriphum, Chiang Mai, 50200, Thailand.

* Corresponding author: PCD, Department of Biological Engineering, 56-787, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139; Tel.: +1-617-253-8017; Fax: +1-617-324-7554; email: pcdedon@mit.edu.

Contents:

Supplementary Table 1
Supplementary Table 2

Supplementary Figure 1
Supplementary Figure 2
Supplementary Figure 3
Supplementary Figure 4
Supplementary Figure 5
Supplementary Figure 6
Supplementary Figure 7

Supplementary Tables

Supplementary Table 1: Modified ribonucleosides detected by mass spectrometric analysis of BCG tRNA hydrolysates[1]					
Identity[2]	**Short name[2]**	**Retention Time, min**	**Precursor ion, m/z**	**Product ion, m/z**	**Signal intensity[3]**
Pseudouridine	Ψ	1.41	245.1	125.1	1179
5-Methylcytidine	m^5C	2.52	258.1	126.1	76
3-Methylcytidine	m^3C	2.87	258.1	126.1	786
1-Methyladenosine	m^1A	2.87	282.1	150.1	1452735
Inosine	I	4.39	269.1	137.1	328908
7-Methylguanosine	m^7G	4.53	298.1	166.1	977212
2′-O-Methylcytidine	Cm	4.98	258.1	112.1	70613
4-Demethylwyosine	*imG-14*	*5.68*	*322*	*190*	*866*
N^6-(cis-Hydroxyisopentenyl)adenosine	io^6A	5.89	352	220	10166
5-Methyluridine	m^5U	9.13	259.1	127.1	134
2-Methyladenosine	m^2A	9.49	282.1	150.1	1577
2′-O-Methylguanosine	Gm	12.5	298.1	152.1	2326
1-Methylguanosine	m^1G	12.52	298.1	166.1	93811
6-Methyladenosine	m^6A	13.16	282.1	150.1	52194
2-Methylguanosine	m^2G	13.62	298.1	166.1	772
N^6,N^6-Dimethyladenosine	$m^6{}_2A$	20.14	296.1	164.1	28173
N^6-Threonylcarbamoyladenosine	t^6A	21.94	413.1	281.1	26426
2-Methylthio-N^6-isopentenyladenosine	*ms^2i^6A*	*23.25*	*382.1*	*250*	*40357*

[1] Multiple reaction monitoring (MRM) experiments were performed with a triple quadrupole mass spectrometer, as described in the Methods section.

[2] RNA modifications noted in *italics* are tentative identifications based on collision-induced dissociation (CID) fragmentation on a quadrupole time-of-flight mass spectrometer. All other modified ribonucleosides were corroborated with synthetic standards.

[3] Mean intensity from analysis of 3 biological replicates.

Supplementary Table 2A: ANCOVA of linear regressions of CCRF-SB RNA species-specific RiboGreen responses.

RNA Species	Slope[1], AU.ml.ng^{-1}	Goodness of fit, r^2	Y-intercept[1], AU	F[2], DFn, DFd	p	Correction[3], %
E. coli total RNA	4.645 (4.51, 4.779)	0.9992	-47.67 (-102.4, 7.063)	N/A	N/A	0
23S rRNA	4.915 (4.876, 4.955)	0.9999	-0.9157 (-16.8, 14.97)	17.2139 (1, 60)	0.00011	+6
16S rRNA	4.454 (4.392, 4.515)	0.9998	5.472 (-19.44, 30.38)	8.55627 (1, 60)	0.0049	-4
tRNA	4.048 (3.934, 4.161)	0.9992	35.18 (-10.99, 81.35)	77.5287 (1, 60)	<0.0001	-13

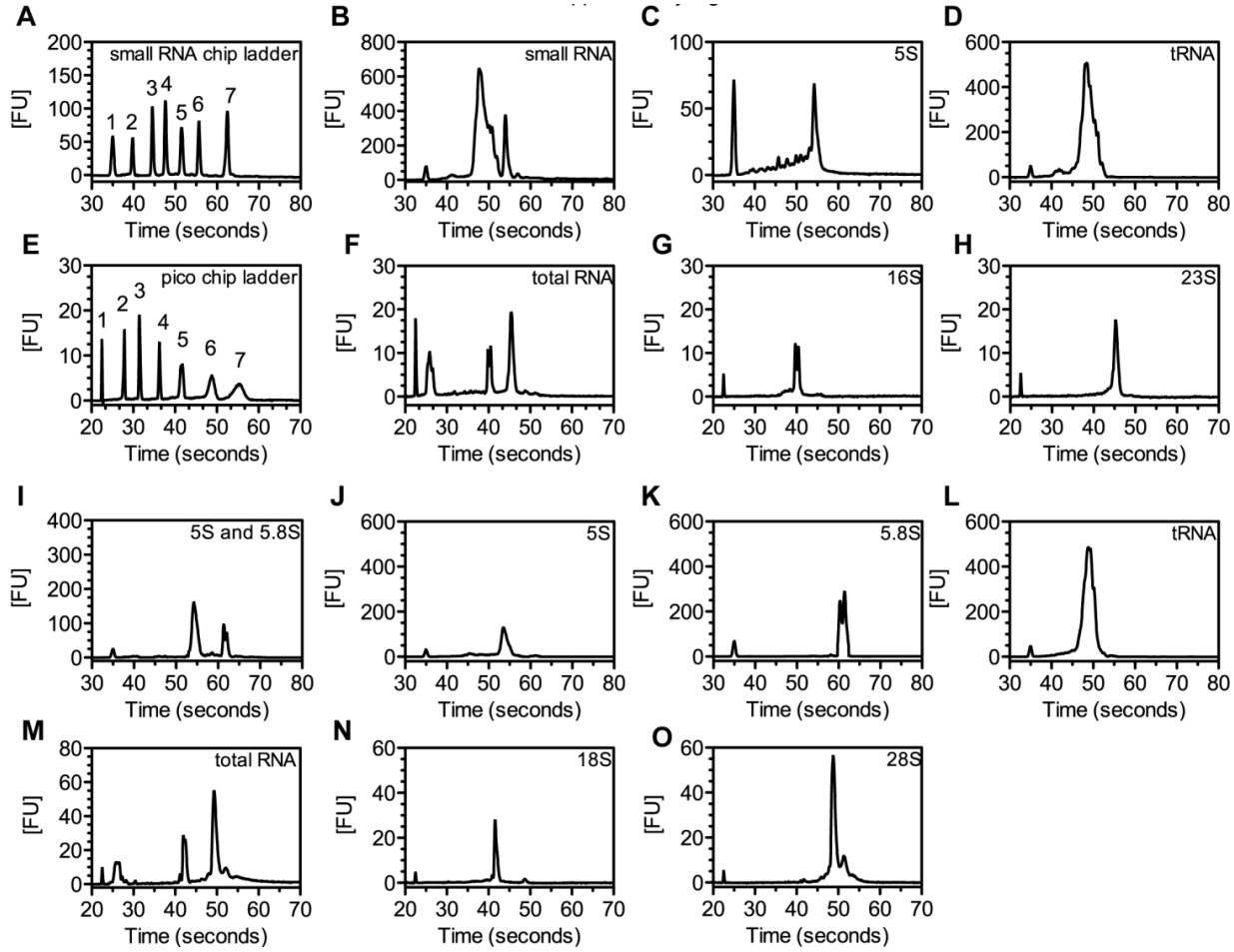
Supplementary Table 2B: ANCOVA of linear regressions of the *E. coli* RNA species-specific RiboGreen responses.

RNA Species	Slope[1], AU.ml.ng^{-1}	Goodness of fit, r^2	Y-intercept[1], AU	F, DFn, DFd[2]	p	Correction[3], %
CCRF-SB total RNA	4.554 (4.423, 4.684)	0.9992	-13.33 (-66.41, 39.74)	N/A	N/A	0
28S rRNA	5.021 (4.948, 5.093)	0.9998	-10.03 (-39.63, 19.58)	41.312 (1, 52)	<0.0001	+10
18S rRNA	4.345 (4.293, 4.397)	0.9999	-2.469 (-23.73, 18.80)	8.159 (1, 52)	0.006	-5
tRNA	4.137 (4.012, 4.261)	0.9991	26.01 (-24.61, 76.63)	35.511 (1, 52)	<0.0001	-9

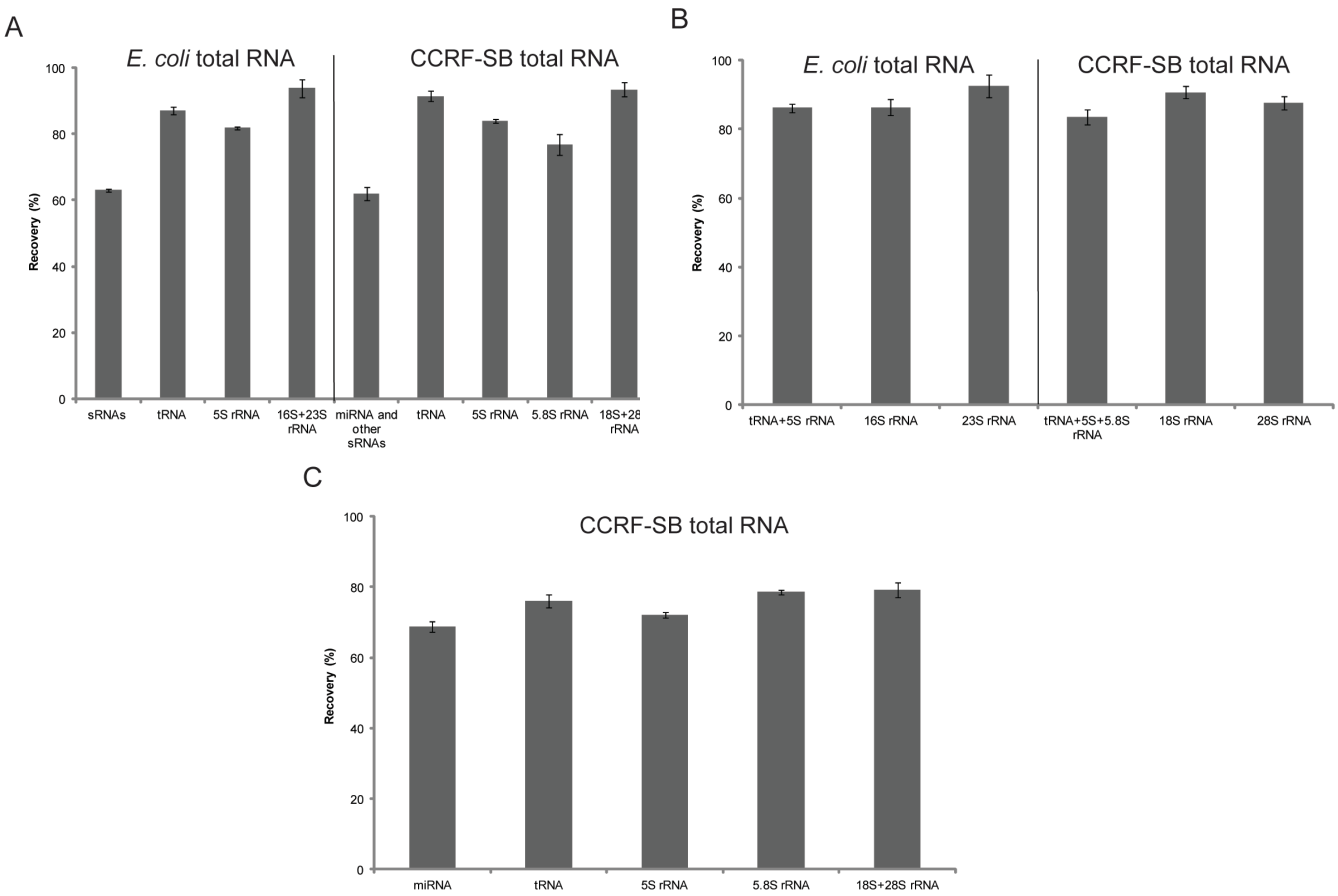
[1] 95% confidence intervals are displayed in parenthesis.

[2] Degrees of freedom for calculation of F are displayed in parenthesis.

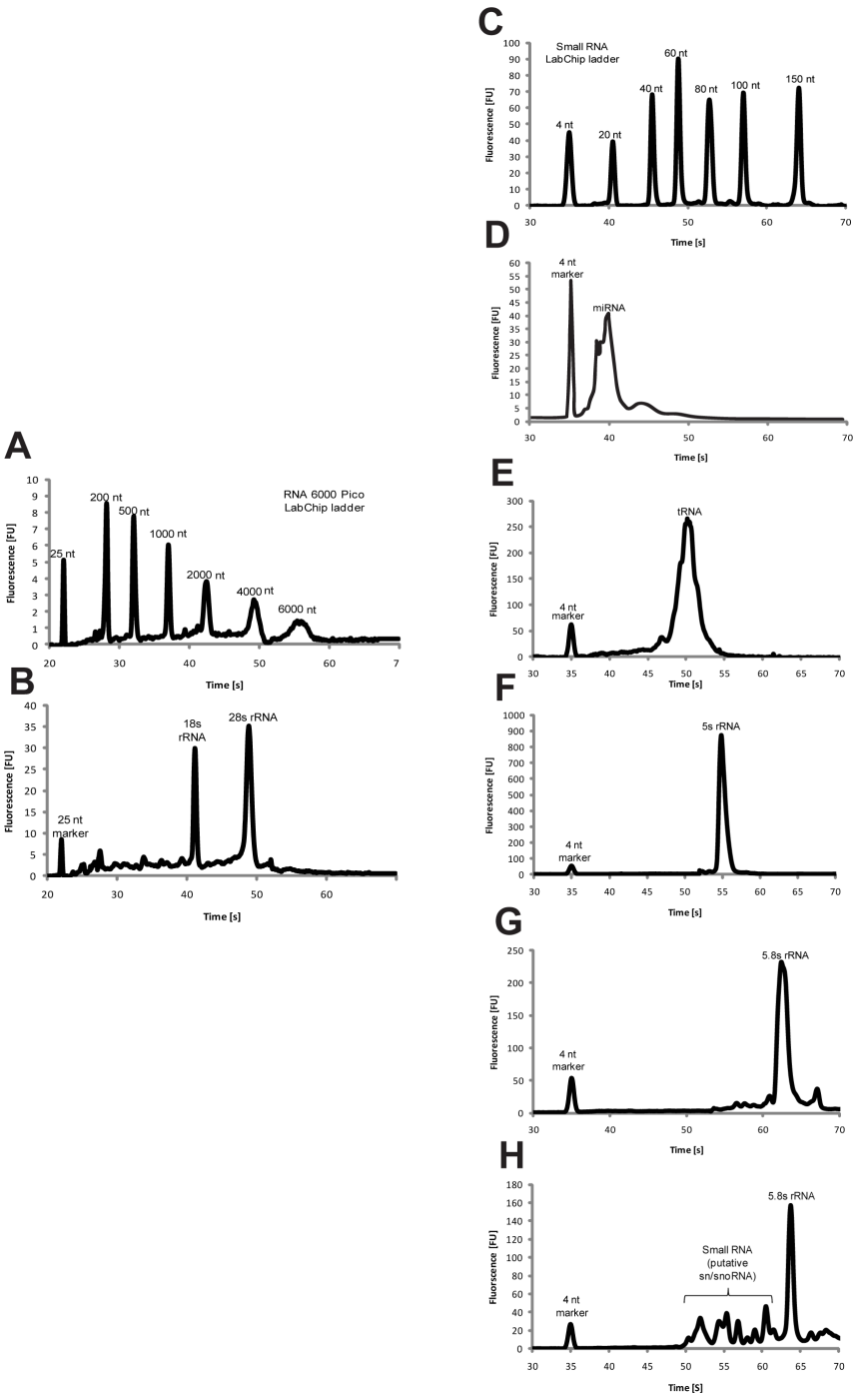
[3] Correction factors were calculated based on comparisons of slopes between each purified RNA species and *E. coli* total RNA and CCRF-SB total RNA (Figure 2C, 2D respectively). These represent the percentage change in fluorescence signals observed with the indicated purified RNA species relative to the total RNA the RNA species was isolated from.



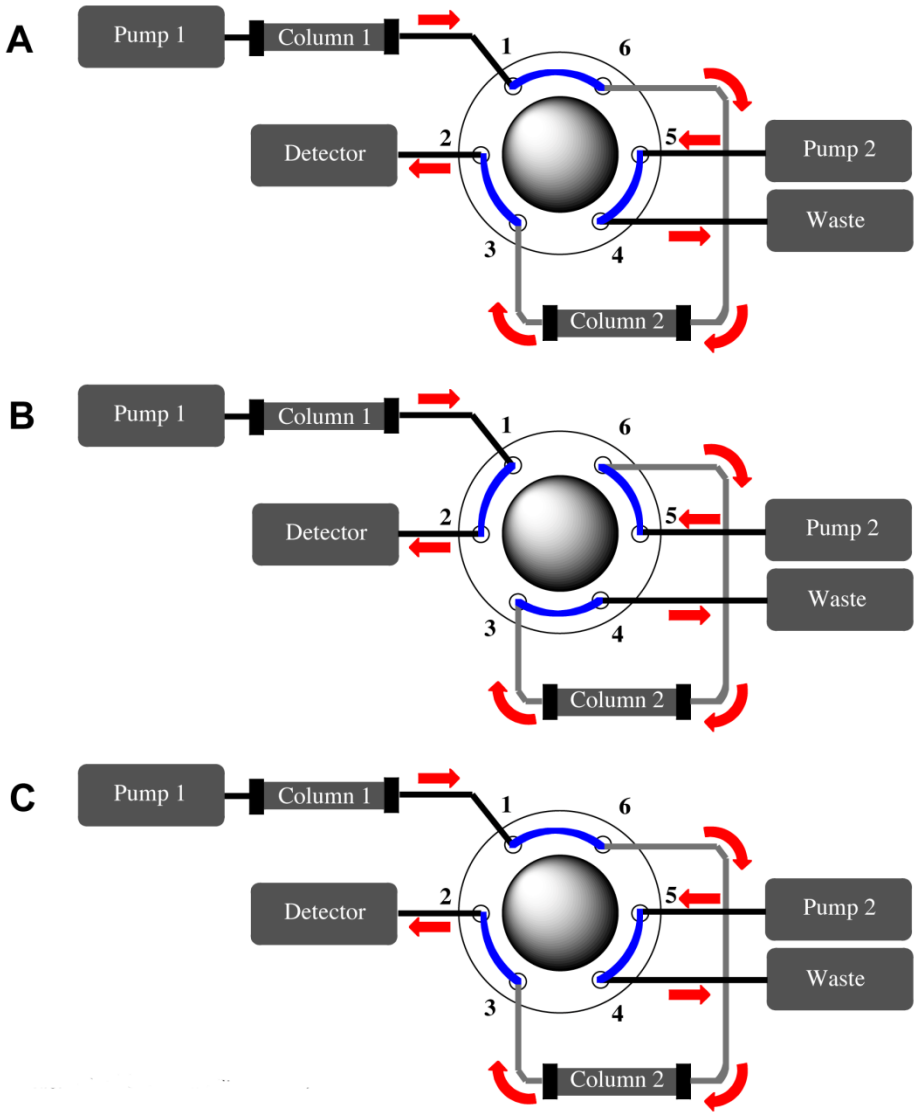
Supplementary Figure S1. Validation of RNA identity and purity of *E. coli* and CCRF-SB RNA species isolated using SEC by Bioanalyzer LabChip analysis. *Top row*: (**A**) Small RNA chip ladder standards (1, 4 nt; 2, 20 nt; 3, 40 nt; 4, 60 nt; 5, 80 nt; 6, 100 nt; 7, 150 nt). (**B**) *E. coli* small RNA consisting of 5s rRNA and tRNA isolated using a Bio SEC-3 300 Å column. (**C**) *E. coli* 5s rRNA isolated using a Bio SEC-3 300 Å column. (**D**) *E. coli* tRNA isolated using a Bio SEC-5 1000 Å column. *Second row*: (**E**) RNA 6000 Pico chip ladder standards (1, 25 nt; 2, 200 nt; 3, 500 nt; 4, 1000 nt; 5, 2000 nt; 6, 4000 nt; 7, 6000 nt). (**F**) *E. coli* total RNA. (**G**) *E. coli* 16s rRNA isolated using a Bio SEC-5 1000 Å column. (**H**) *E. coli* 23s rRNA isolated using a Bio SEC-5 1000 Å column. *Third row*: (**I**) CCRF-SB small RNA consisting of 5.8s and 5s rRNAs isolated using a Bio SEC-5 1000 Å column. (**J**) CCRF-SB 5s rRNA isolated using a Bio SEC-3 300 Å column. (**K**) CCRF-SB 5.8s rRNA isolated using a Bio SEC-3 300 Å column. (**L**) CCRF-SB tRNA isolated using a Bio SEC-3 300 Å column. *Bottom row*: (**M**) CCRF-SB total RNA. (**N**) CCRF-SB 18s rRNA isolated using a Bio SEC-5 1000 Å column. (**O**) CCRF-SB 28s rRNA isolated using a Bio SEC-5 1000 Å column.



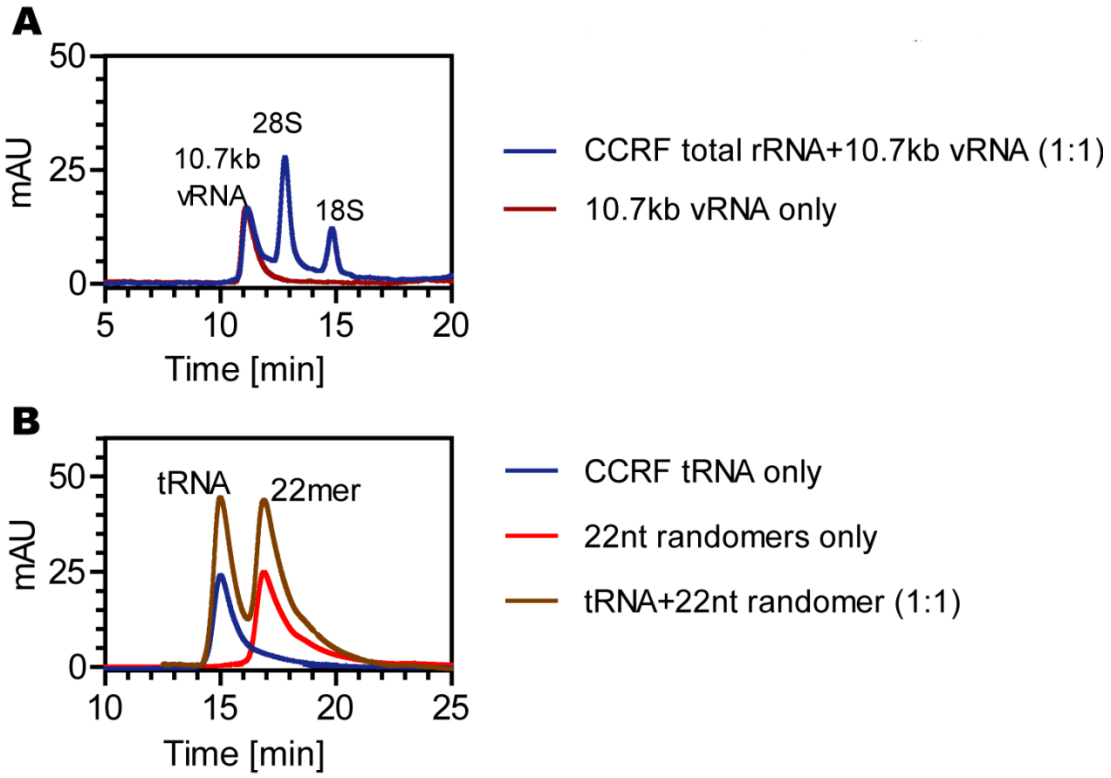
Supplementary Figure S2. Yields of ncRNA purified by SEC or IP-RPC. Fraction collection was automated by use of a fraction collector and RNA collected by custom time segments corresponding to the retention times and peak widths of each RNA species. Samples clean-up was then performed as stated in Material and Methods. (**A**) Recoveries (%) of sRNA (of lengths smaller than tRNA), tRNA, 5S rRNA and co-eluding 16S and 23S rRNA from *E. coli* total RNA (left) and miRNA and other sRNA, tRNA, 5S rRNA, 5.8S rRNA and co-eluding 18S and 28S rRNA from CCRF-SB total RNA (right) off the Bio-SEC3 300Å HPLC system. (**B**) Recoveries (%) of co-eluding 5S rRNA and tRNA, 16S rRNA and 23S rRNA from *E. coli* total RNA (left) and co-eluding 5S, 5.8S rRNA and tRNA, 18S rRNA and 28S rRNA from CCRF-SB total RNA (right) off the Bio-SEC5 1000 Å HPLC system. 10-100 µg of total RNA are injected unto the SEC systems. (**C**) Recoveries (%) miRNA, tRNA, 5S rRNA, 5.8S rRNA and co-eluding 18S and 28S rRNA from the Source 5RPC 4.6/150 HPLC system. 2-40 µg of total RNA are injected into the IP RPC system.



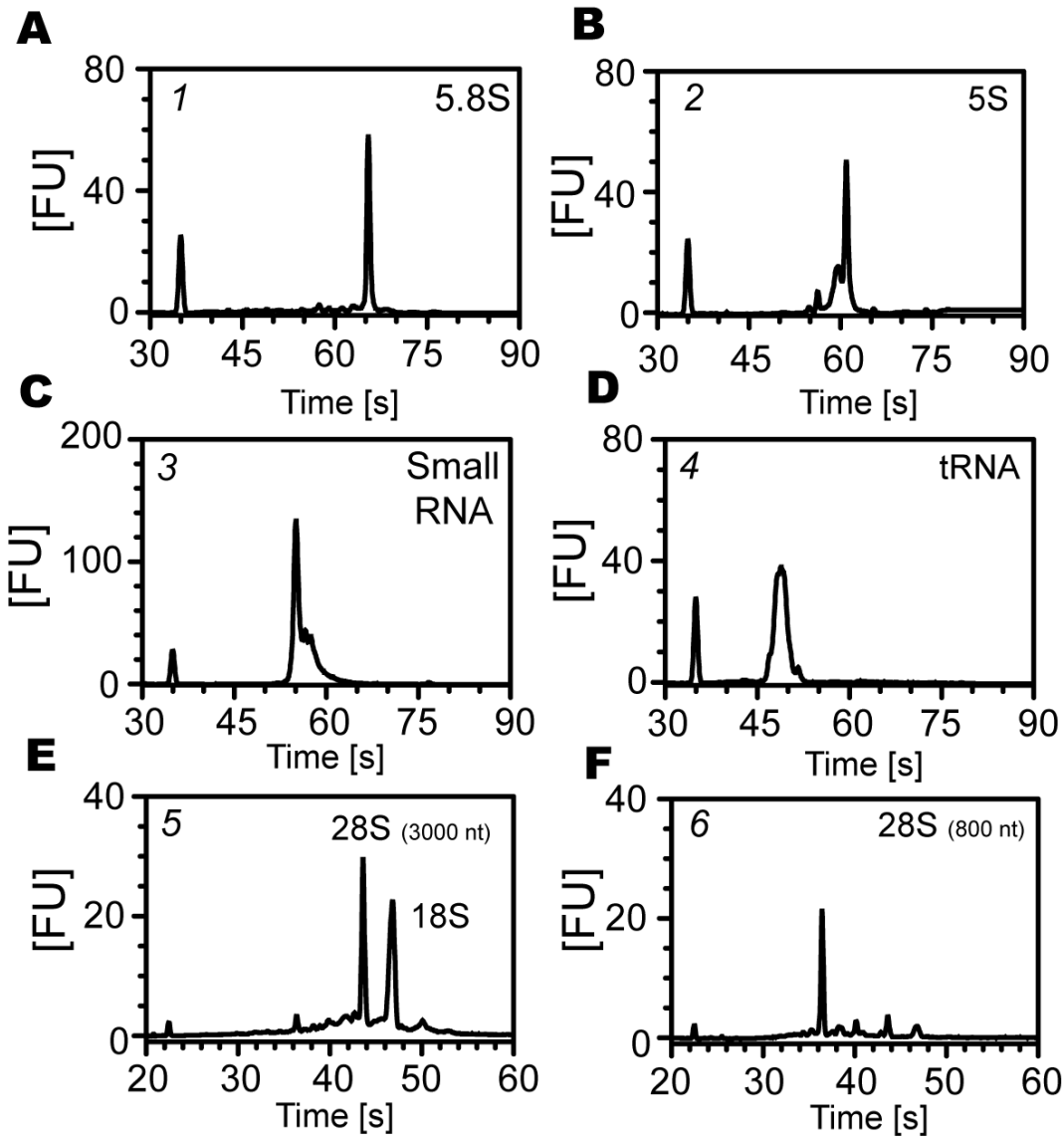
Supplementary Figure S3. Validation of RNA identity and purity of CCRF-SB RNA species isolated using IP RPC on a Source 5RPC 4.6/150 column by Bioanalyzer LabChip analysis. (**A**) RNA 6000 Pico chip ladder standards (From left to right: 25 nt; 200 nt; 500 nt; 1000 nt; 2000 nt; (**B**) Co-eluding 18S and 28S rRNA analyzed on a RNA 6000 Pico chip. 4000 nt; 6000 nt). (**C**) Small RNA chip ladder standards (From left to right: 4 nt; 20 nt; 40 nt; 60 nt; 80 nt; 100 nt; 150 nt). (**D**) miRNA, (**E**) tRNA, (**F**) 5S rRNA, (**G**) 5.8S rRNA, (**H**) putative sn/sno RNA analyzed on a small RNA chip.



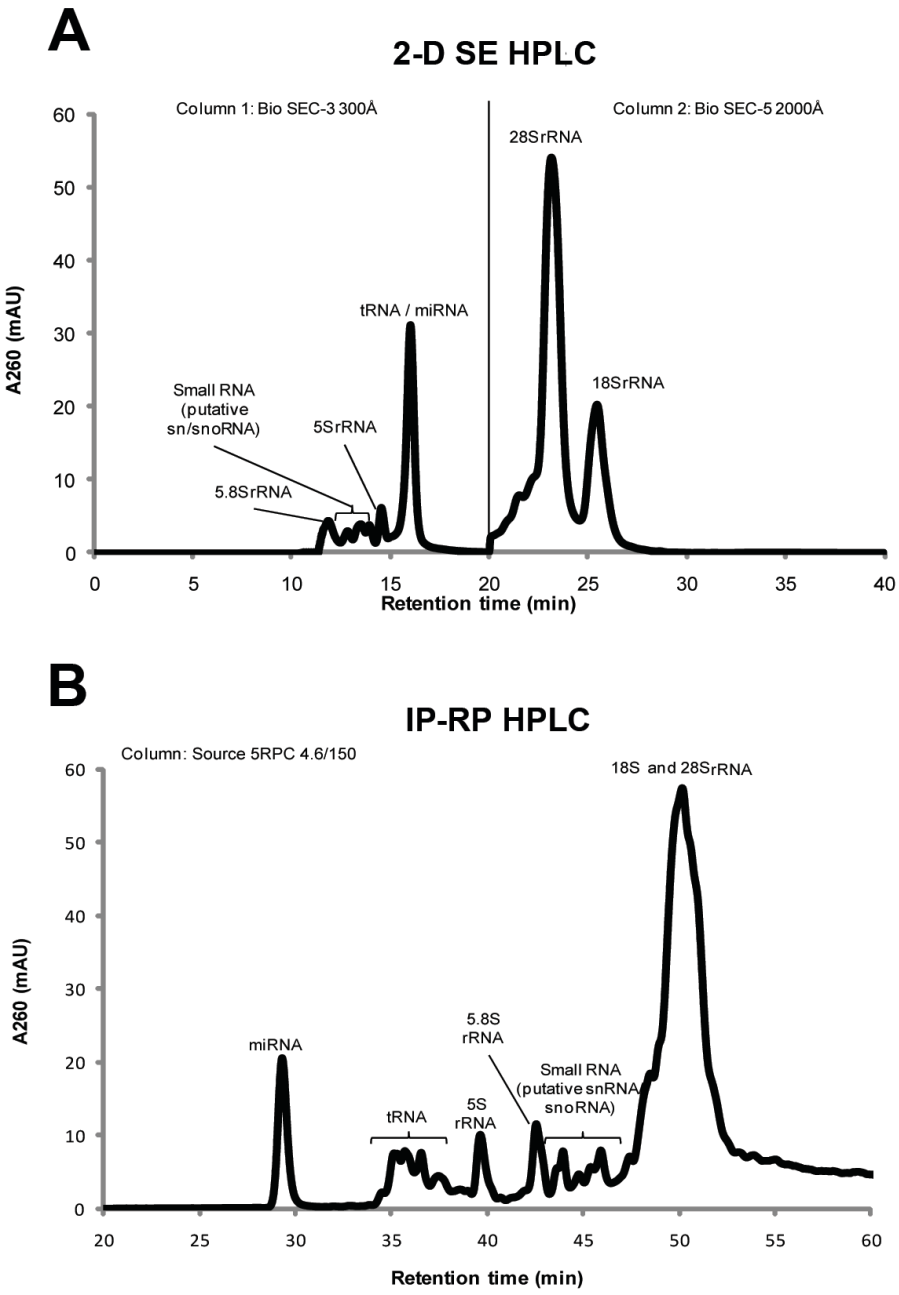
Supplementary Figure S4. Valve configuration for 2-D SEC. Either Bio SEC-3 or Bio SEC-5 columns can be connected into the column 1 or column 2 positions. (**A**) Step 1: Columns are in tandem and detector is in-line. If Bio SEC-5 is in the column 1 position, small RNAs are fractioned from large molecular weight rRNAs on the Bio SEC-5 and are eluted onto the Bio SEC-3 column (column 2) for separation. Alternatively, If Bio SEC-3 is in the column 1 position, large molecular weight rRNAs are eluted onto the Bio SEC-5 column (column 2) for separation. (**B**) Step 2: Column 2 is taken off-line (switched to pump 2) while the detector is in-line with column 1. Hence, if Bio SEC-5 is in the column 1 position, large molecular weight rRNAs are quantified by absorbance and collected. Alternatively, if Bio SEC-3 is in the column 1 position, small RNAs are quantified and collected. With column 2 in-line with pump 2, RNA separation can continue at the same or lower flow rates depending on the run time of column 1. (**C**) Step 3: Column 1 and 2 are connected in tandem and the detector is connected in-line with column 2. RNAs from column 2 (small RNA if Bio SEC-5 is in the column 1 position, and large molecular weight rRNAs if Bio SEC-3 is in the column 1 position) are quantified and collected.



Supplementary Figure S5. Biologically relevant exclusion and permeation limit of SEC HPLC. (**A**) Exclusion limit of Bio SEC-5 1000 Å (ID 7.8mm, 300mm) assessed with 10.7 kb *in vitro* transcribed full length DENV vRNA. Chromatograms of 50 ng of DENV 10.7 kb vRNA alone (red) and 50 ng of 10.7 kb vRNA spiked into 50 ng of CCRF-SB total RNA (blue) performed with 100 mM ammonium acetate buffer (pH 7.0) at a flow rates of 0.5 ml/min. (**B**) Permeation limit of Bio SEC-3 300 Å (ID 7.8mm, 300mm) assessed with synthetic 22 nt random oligonucleotides. Chromatograms of 100 ng of 22 nt random oligonucleotides alone (red), 100 ng of CCRF-SB tRNA alone (blue) and 100 ng of 22 nt random oligonucleotides spiked into 100 ng of CCRF-SB tRNA (brown) performed with 100 mM ammonium acetate buffer (pH 7.0) at a flow rate of 0.5 ml/min.



Supplementary Figure S6. Bioanalyzer LabChip validation of RNA identity and purity of *P. berghei* RNA species isolated using 2-D SEC. (**A**) 5.8s rRNA (fraction 1). (**B**) 5s rRNA (fraction 2). (**C**) Small RNA (putative snoRNA and snRNA; fraction 3). (**D**) tRNA (fraction 4). (**E**) 28s rRNA 3000 nt fragment and 18s rRNA isolated from fraction 5 as labeled in Figure 4b. (**F**) 28s rRNA 800 nt fragment isolated from fraction 6. Fractions 1-4 are analyzed on a Bioanalyzer Small RNA chip while Fraction 5 and 6 are analyzed on a Bioanalyzer RNA 6000 Pico chip. Sizing ladders are shown in **Supplementary Figures 1A and 1E**.



Supplementary Figure S7. Orthogonal separations of TK6 total RNA by 2-D SEC and IP RPC. (**A**) 2-D SEC chromatogram of TK6 total RNA consisting of co-eluting miRNA and tRNA, putative snRNA and snoRNA, 5S, 5.8S, 18S, 28S rRNAs separated obtained on a Bio SEC-3 300 Å (column 1) and Bio SEC-5 2000 Å (column 2) arranged as an in-line two-column system. Both separations were performed in 100 mM TEAA (pH 7.0) buffer at a flow rate of 1 ml/min at 60 °C. (**B**) IP RP HPLC of TK6 total RNA consisting of miRNA, tRNA, putative snRNA and snoRNA, 5S, 5.8S rRNAs, and co-eluting 18S and 28S rRNAs separated on a Source 5RPC 4.6/150 column, performed at a flow rate of 1 ml/min at 60 °C. Buffer A: 100 mM TEAA, pH 7.0 with 2% acetonitrile. Buffer B: 100% acetonitrile. Elution gradient as stated in Materials and Methods.