Libertas Academica
FREEDOM TO RESEARCH

Open Access: Full open access to this and thousands of other papers at http://www.la-press.com.

Biomarkers in Cancer

Deep Sequencing of Serum Small RNAs Identifies Patterns of 5′ tRNA Half and YRNA Fragment Expression Associated with Breast Cancer

Joseph M. Dhahbi[1,2], Stephen R. Spindler[1], Hani Atamna[3], Dario Boffelli[2] and David I.K. Martin[2]

[1]Department of Biochemistry, University of California at Riverside, Riverside, CA, USA. [2]Center for Genetics, Children's Hospital Oakland Research Institute, Oakland, CA, USA. [3]Department of Medical Education, California Northstate University, Elk Grove, CA, USA.

ABSTRACT: Small noncoding RNAs circulating in the blood may serve as signaling molecules because of their ability to carry out a variety of cellular functions. We have previously described tRNA- and YRNA-derived small RNAs circulating as components of larger complexes in the blood of humans and mice; the characteristics of these small RNAs imply specific processing, secretion, and physiological regulation. In this study, we have asked if changes in the serum abundance of these tRNA and YRNA fragments are associated with a diagnosis of cancer. We used deep sequencing and informatics analysis to catalog small RNAs in the sera of breast cancer cases and normal controls. 5′ tRNA halves and YRNA fragments are abundant in both groups, but we found that a breast cancer diagnosis is associated with changes in levels of specific subtypes. This prompted us to look at existing sequence datasets of serum small RNAs from 42 breast cancer cases, taken at the time of diagnosis. We find significant changes in the levels of specific 5′ tRNA halves and YRNA fragments associated with clinicopathologic characteristics of the cancer. Although these findings do not establish causality, they suggest that circulating 5′ tRNA halves and YRNA fragments with known cellular functions may participate in breast cancer syndromes and have potential as circulating biomarkers. Larger studies with multiple types of cancer are needed to adequately evaluate their potential use for the development of noninvasive cancer screening.

KEYWORDS: breast cancer, circulating small RNAs, tRNA derivatives, tRNA halves, YRNA fragments, Y RNA, serum

CITATION: Dhahbi et al. Deep Sequencing of Serum Small RNAs Identifies Patterns of 5′ tRNA Half and YRNA Fragment Expression Associated with Breast Cancer. *Biomarkers in Cancer* 2014:6 37–47 doi:10.4137/BIC.S20764.

RECEIVED: October 2, 2014. **RESUBMITTED:** November 14, 2014. **ACCEPTED FOR PUBLICATION:** November 17, 2014.

ACADEMIC EDITOR: Barbara Guinn, Editor in Chief

TYPE: Original Research

FUNDING: Authors disclose no funding sources.

COMPETING INTERESTS: Authors disclose no potential conflicts of interest.

CORRESPONDENCE: jdhahbi@ucr.edu

Paper subject to independent expert blind peer review by minimum of two reviewers. All editorial decisions made by independent academic editor. Upon submission manuscript was subject to anti-plagiarism scanning. Prior to publication all authors have given signed confirmation of agreement to article publication and compliance with all applicable ethical and legal requirements, including the accuracy of author and contributor information, disclosure of competing interests and funding sources, compliance with ethical requirements relating to human and animal study participants, and compliance with any copyright requirements of third parties. This journal is a member of the Committee on Publication Ethics (COPE).

Introduction

Small noncoding RNAs complex with proteins to regulate key cellular functions.[1,2] The potential of small noncoding RNAs to influence pathophysiological processes is best exemplified by miRNAs, which function in cell differentiation, cell signaling, tumorigenesis, and the pathogenesis of multiple diseases.[3,4] Recently, noncoding RNAs with well-known functions, such as tRNA, rRNA, snoRNA, and YRNA, have been reported to generate smaller RNAs whose characteristics imply specific processing and function.[5,6] A variety of stresses can induce the cleavage of tRNAs and YRNAs inside the cell to yield shorter noncoding RNA species.[7–9] We have reported that tRNA- and YRNA-derived small RNAs circulate in the blood as components of larger particles, and that circulating tRNA derivatives can be physiologically regulated.[10,11] More recently, others have reported tRNA- and YRNA-derived small RNAs in another biological fluid, semen.[12] Their functions in biological fluids are not yet known, but precedent for such small RNAs carrying out signaling functions is provided by miRNAs that circulate in mammalian blood, enter target cells, modulate gene expression, and serve as markers of diseases.[13–16] In particular, tumor-derived miRNAs circulate in serum and plasma, and are emerging as novel blood-based markers for cancer detection and monitoring.[13,17–20]

tRNA-derived molecules termed 5′ tRNA halves result from cleavage of full-length tRNAs by the ribonuclease angiogenin in response to stress.[21,22] They interact with components of the translation initiation complex, promote assembly of stress granules carrying stalled preinitiation complexes, and inhibit mRNA translation.[22–25] YRNA-derived fragments remain largely unexplored. Full-length YRNAs are components of the autoantigenic Ro ribonucleoproteins, play a role in chromosomal DNA replication and quality control of noncoding RNA,[26–28] and are overexpressed in various cancers.[29] YRNA-derived fragments were first observed in cells exposed to apoptotic stimuli,[30] or poly(I:C).[9] YRNA fragments were also detected in non-stressed cells,[9] and in vesicles released by immune cells.[31]

Our group and others have recently reported the presence of YRNA fragments and 5′ tRNA halves in serum.[10,11,32] The tRNA- and YRNA-derived molecules are not contained in exosomes or microvesicles, but circulate as particles of 100–300 kDa, indicating they may be actively secreted through an RNA-binding protein-dependent pathway. YRNA fragments are dramatically more abundant in human than in mouse serum,[10] possibly reflecting the much greater copy number of YRNA genes and pseudogenes in humans.[29] While the cells that produce circulating YRNA fragments are not yet known, circulating 5′ tRNA halves may originate from blood cells and hematopoietic tissues, where they are detected at significant levels.[11] The functions of circulating tRNA and YRNA derivatives are not yet known; however, we found that serum levels of specific subtypes of 5′ tRNA halves change markedly with age, and that these changes can be prevented by calorie restriction.[11]

Our recent observation that tRNA- and YRNA-derived fragments circulate in the serum and the evidence that circulating tumor-derived miRNAs may participate in signaling prompted us to explore the possibility that serum levels of 5′ tRNA halves or YRNA fragments may change in response to cancer. We used small RNA sequencing to assess the abundance of 5′ tRNA halves and YRNA fragments circulating in the sera of a small group of women with breast cancer, as well as normal controls; the results prompted us to study datasets obtained from a larger group of breast cancer cases. Our findings suggest that breast cancer may be associated with significant alterations in the serum levels of fragments derived from specific tRNAs and YRNAs.

Results and Discussion

Characterization of small RNAs circulating in serum of normal subjects and breast cancer patients. We used breast cancer to explore the possibility that changes in circulating 5′ tRNA halves and 5′ YRNA fragments, besides circulating miRNAs, can be associated with a disease state. We deep sequenced serum small RNAs and compared levels of the various types of circulating small RNAs in five breast cancer and five control subjects. The breast cancer subjects were a diverse group, and some samples were obtained after treatment (Table 1). Sequencing reads from each serum sample were aligned to the human genome. The size distribution of aligned reads revealed three peaks, at 20–24 nt, 25–29 nt, and 30–33 nt, in both normal and cancer groups (Fig. 1A and B). The peak pattern is similar to the pattern we recently reported in the human and mouse sera.[10,11]

Sequencing reads from both normal and cancer samples were pooled to inspect the length distribution of reads, the pattern and size of peaks, and the types and proportions of small RNAs from which the reads are derived from. Pooling is used only to examine the general characteristics of the reads, and not to measure the differential expression of small RNAs between control and cancer groups. A combined total of 122,490,155 pre-processed reads were aligned to the hg19 human genome to generate a dataset of 110,652,186 mapped reads (90.34%), ranging in size from 18 to 49 nt. Annotation and length distribution analyses revealed that reads in the 20–24 nt and 25–29 nt peaks were derived from miRNAs and YRNAs, respectively, while the 30–33 nt peak consists of reads mapping to both YRNA and tRNA genes (Fig. 1C). Further annotation analysis showed that of the total 110,652,186 reads mapping to the human genome, 98,999,062 (89.5%) map to known small RNAs, of which 41% were annotated as miRNAs, 38% as YRNAs, 21% as tRNAs, and <1% as rRNA, snRNA, and snoRNA (Fig. 1D). The significant presence in serum of tRNA- and YRNA-derivatives in addition to miRNA is consistent with our previous findings.[10,11]

These datasets also produced evidence of a novel type of circulating YRNA fragments. We recently reported that in

Table 1. Clinical characteristics of breast tumors.

PATIENT	CHARACTERISTICS
Patient 1	63 years old; infiltrating ductal carcinoma, grade 2, ER+, PR+.
Patient 2	47 years old; infiltrating ductal carcinoma, grade 2, ER+, PR+.
Patient 3	47 years old; Infiltrating ductal carcinoma with lymph node metastasis, ER+, PR+.
Patient 4	58 years old; breast carcinoma treated with surgery and chemotherapy with subsequent widespread osseous metastases.
Patient 5	55 years old; right mastectomy with pathology revealing a pathologic stage IIIA (ypT3ypN2@4 out of 6@M0) invasive carcinoma and received radiation therapy and adjuvant letrozole.

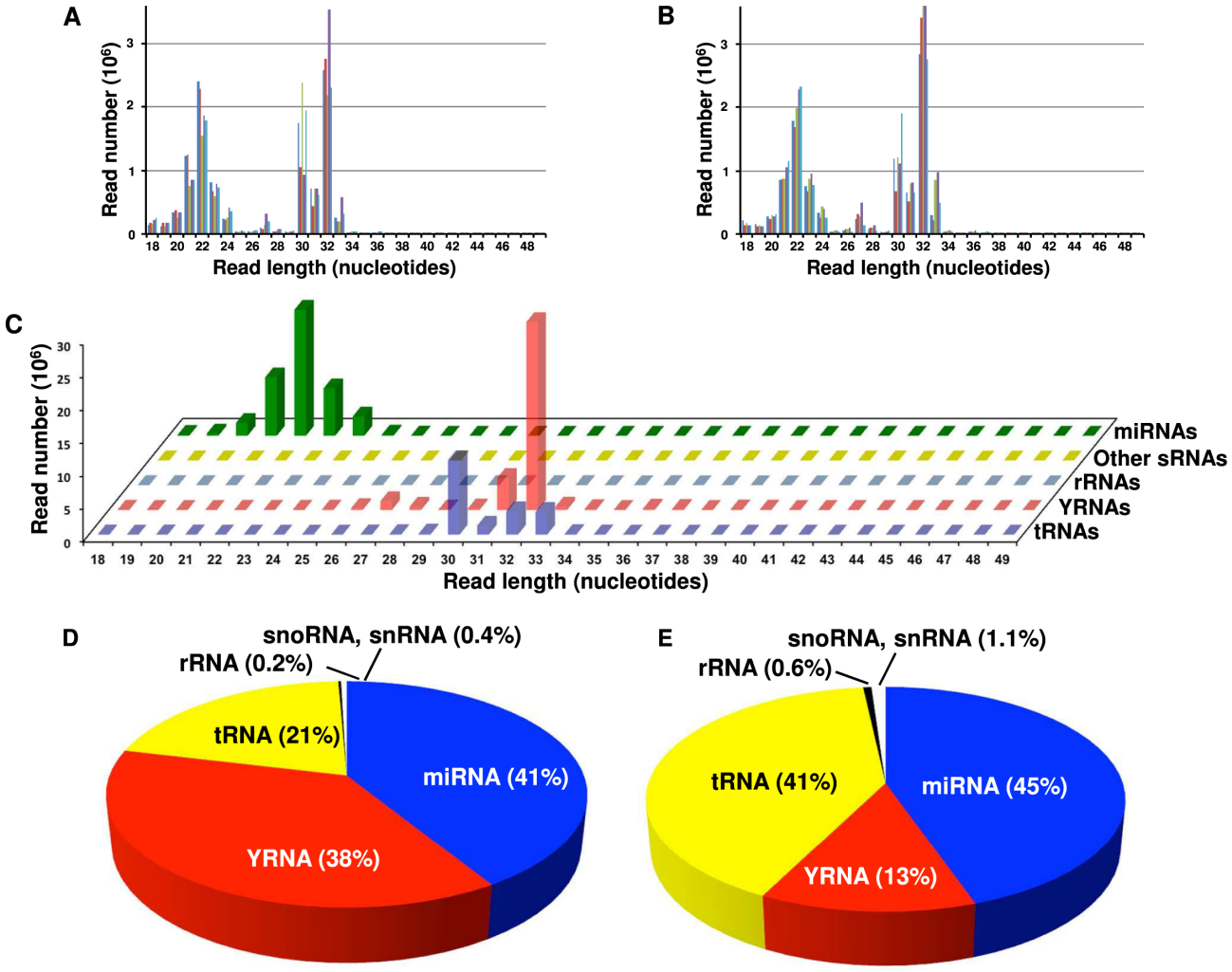

Figure 1. Length distribution and annotation of reads from serum small RNAs of five healthy females and five females diagnosed with breast cancer. Length of mapped reads from five normal (**A**) and five cancer samples (**B**) is plotted against read abundance. Colors of the bars denote the individual sources of the serum small RNAs (normal and breast cancer samples). (**C**) Length distribution of pooled sequencing reads from all 10 samples plotted against abundance of reads according to their annotation. Length distribution is plotted against abundance of the reads annotated as miRNAs, YRNAs, tRNAs, rRNAs, or other sRNAs (snRNAs and snoRNAs). (**D**) Pie chart showing the percent of reads (pooled from both normal and breast cancer) mapping to the indicated specific types of small RNAs. (**E**) Pie chart showing the percent of reads mapping to the indicated types of small RNAs in pooled datasets obtained by sequencing of small RNAs in the sera of 42 newly diagnosed breast cancer cases.

human serum, >95% of YRNA-aligned reads map to the 5′ ends of YRNA genes and >99% of tRNA-aligned reads map to the 5′ ends of tRNA genes.[10,11] In the present study, we took a closer look at the YRNA-derived reads that align to the 3′ ends of YRNA genes. In this dataset, among the total pooled reads that map to YRNA genes, 5.3% and 94.7% of YRNA reads align to the 3′ and 5′ ends of YRNA genes, respectively (Fig. 2A). Since YRNA-derived reads are found in both the 25–29 nt and the 30–33 nt peaks (Fig. 1C), we calculated the proportions of the reads mapping to the 3′ end versus the 5′ end of YRNA genes in the 25–29 nt and the 30–33 nt peaks. Almost all of the reads in the 30–33 nt peak align to the 5′ ends of YRNAs (Fig. 2B). In contrast, only 13.0% of the reads in the 25–29 nt peak align to the 5′ end of YRNAs, while 87% align to the 3′ end (Fig. 2B). This means that the circulating 5′ YRNA fragments are of two sizes, 25–29 nt and 30–33 nt, while the circulating 3′ YRNA fragments are of only one size, 25–29 nt. This is supported by Northern blot analysis of RNA extracted from serum (Fig. 2C). A probe derived from the 5′ end of the RNY4 gene (Fig. 2C, left panel) detected a major band at the 30-nt marker (consistent with the 30–33 nt peak), and a minor band below the 30-nt marker (consistent with the 25–29 nt peak). An oligonucleotide probe complementary to the 3′ end of the RNY4 gene detected only a minor band below the 30-nt marker, consistent with the 25–29 nt peak (Fig. 2C, right panel). These results indicate the presence in human serum of a major population of 30–33 nt YRNA fragments mostly derived from the 5′ end of YRNAs, and a minor population of 25–29 nt YRNA fragments almost exclusively derived from the 3′ end of YRNAs. The biogenesis and

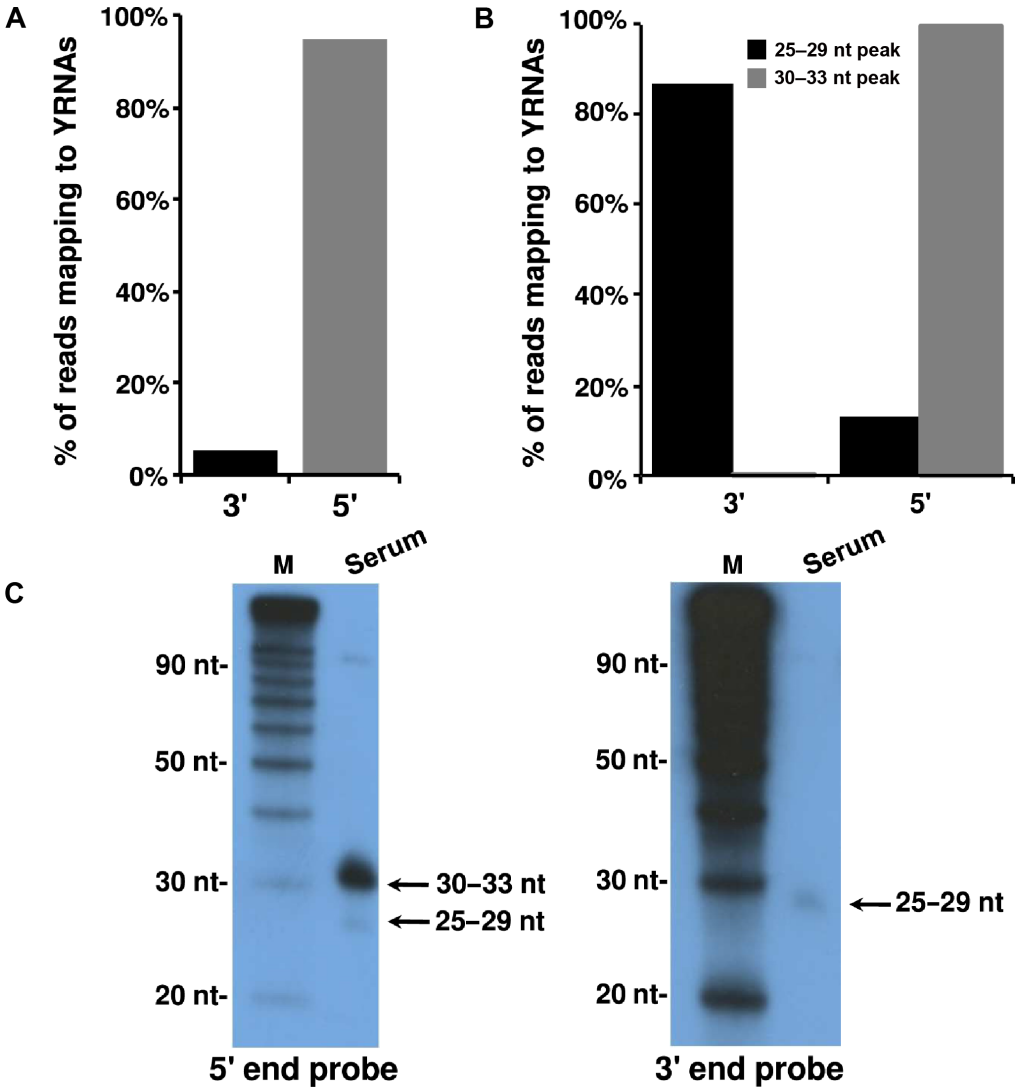

Figure 2. Characteristics of sequencing reads that map to YRNA genes. (**A**) Pooled reads from all 10 samples were used to compare the number of reads that align with 5′ or 3′ ends of YRNA genes. The *y*-axis represents the percentage of the total YRNA reads that map to YRNA genes. (**B**) Percentage of reads mapping to 5′ and 3′ ends of YRNAs in the 25–29 nt peak (black) and the 30–33 nt peak (gray). Almost 100% of the reads in the 30–33 nt peak align to the 5′ end of YRNAs, while a significant majority of reads in the 25–29 nt peak align to the 3′ end of YRNAs. (**C**) Northern blot analysis of RNA extracted from 0.2 mL of human serum. The blot was hybridized to ^{32}P-end-labeled oligonucleotide probes complementary to the 5′ end (left panel) or 3′ end (right panel) of RNY4. Arrows indicate the positions of the major (30–33 nt) and minor (25–29 nt) bands. M: Decade markers.

biological significance of these two types of circulating YRNA fragments remain to be elucidated.

Altered abundance of specific circulating 5′ tRNA halves and YRNA fragments in breast cancer patients.

5′ tRNA halves. We compared serum abundance of 5′ tRNA halves in datasets from individual breast cancer and normal subjects using the Bioconductor package edgeR.[33] There are 625 tRNA genes, each of them potentially capable of producing 5′ tRNA halves. Our analysis suggests that a breast cancer diagnosis is associated with alterations—either increase or decrease—in the circulating levels of 5′ tRNA halves derived from specific tRNA isoacceptors (Table 2). The individuals with breast cancer had significantly increased circulating levels of 5′ tRNA halves derived from the isoacceptors of tRNA-Arg, -Asn, -Cys, -Gln, -Gly, -Leu, -Ser, -Trp, and -Val, and decreased circulating levels of 5′ tRNA halves derived from tRNA-Asp, and -Lys.

YRNA fragments. We analyzed serum levels of YRNA fragments with the Bioconductor package edgeR[33] as for the 5′ tRNA halves. Here also a breast cancer diagnosis is associated with alterations—either increase or decrease—in levels of circulating YRNA fragments (Table 3). Among the 20 types of increased YRNA fragments, 19 were derived from the 3′ end, while only one was derived from the 5′ end of YRNA genes. All five YRNA fragments that decreased in abundance were derived from the 5′ ends of YRNA genes. In this study, and in a previous one,[10] we found that serum 5′ YRNA fragments are much more abundant than 3′ YRNA fragments (a ratio

Table 2. Breast cancer-associated changes in the serum levels of 5′ tRNA halves.

tRNA[1]	GENOMIC COORDINATES	NORMAL (cpm)[2]	TUMOR (cpm)[2]	FC[3]	*P*-VALUE
Arg-CCG	chr17:66016013-66016085	49	145	3.0	0.003
Arg-CCT	chr7:139025446-139025518	23	47	2.1	0.006
Arg-TCT	chr1:159111401-159111474	6	17	2.7	0.005
	chr1:94313129-94313213	1714	4033	2.4	0.004
	chr9:131102355-131102445	6	13	2.3	0.007
	chr17:8024243-8024330	12	22	1.9	0.018
Asn-GTT	chr1:148248115-148248188	26	72	2.8	0.008
Cys-GCA	chr4:124430005-124430076	85	164	1.9	0.008
	chr17:37023898-37023969	553	1073	1.9	0.010
	chr17:37025545-37025616	87	181	2.1	0.003
	chr17:37309987-37310058	83	163	2.0	0.007
	chr17:37310744-37310815	80	163	2.0	0.006
Gln-TTG	chr6:145503859-145503930	3	7	2.1	0.036
Gly-TCC	chr1:161432166-161432237	2	5	2.5	0.016
Leu-AAG	chr5:180528840-180528921	5	12	2.2	0.007
	chr14:21078291-21078372	2	5	2.4	0.023
Ser-GCT	chr6:26305718-26305801	33	55	1.7	0.032
	chr6:27265775-27265856	3	7	2.3	0.019
Trp-CCA	chr6:26331672-26331743	23	39	1.7	0.041
Val-AAC	chr3:169490018-169490090	15896	30605	1.9	0.017
	chr5:180591154-180591226	16231	31153	1.9	0.017
	chr5:180596610-180596682	16125	31093	1.9	0.016
	chr5:180615416-180615488	3021	5485	1.8	0.021
	chr5:180645270-180645342	4304	8291	1.9	0.016
	chr6:27618707-27618779	4259	8194	1.9	0.017
	chr6:27648885-27648957	4344	8315	1.9	0.017
	chr6:27721179-27721251	4232	8154	1.9	0.017
Val-CAC	chr1:149298555-149298627	4331	8279	1.9	0.017
	chr1:149684088-149684161	4391	8347	1.9	0.018
	chr1:161369490-161369562	4414	8422	1.9	0.015
	chr5:180524070-180524142	16273	31118	1.9	0.017
	chr5:180529253-180529325	4466	8487	1.9	0.018
	chr5:180600650-180600722	16731	31644	1.9	0.018
	chr5:180649395-180649467	4395	8333	1.9	0.017
	chr6:26538282-26538354	16594	31522	1.9	0.018
	chr6:27173867-27173939	272	516	1.9	0.020
	chr6:27248049-27248121	8352	17964	2.2	0.018
	chr6:27696327-27696399	25	45	1.8	0.044
Val-TAC	chr6:27258405-27258477	200	354	1.8	0.035
Asp-GTC	chr6:27471523-27471594	94	56	–1.7	0.039
	chr12:125411891-125411962	29	15	–1.9	0.020
	chr12:125424193-125424264	31	16	–1.9	0.015
	chr17:8125556-8125627	30	14	–2.2	0.009
Lys-TTT	chr6:27559593-27559665	115	61	–1.9	0.016
	chr6:28918806-28918878	4322	2407	–1.8	0.042

Notes: [1]tRNA isoacceptor identity with corresponding genomic positions in the human hg19 genome. [2]Average tRNA read count for the indicated experimental group reported as counts per million (cpm) reads in the sequenced library. [3]Fold change calculated by EdgeR from comparison between the normal and breast cancer serum samples.

Table 3. Breast cancer-associated changes in the serum levels of YRNA-derived fragments.

YRNA[1]	GENOMIC COORDINATES	NORMAL (cpm)[2]	TUMOR (cpm)[2]	FC[3]	*P*-VALUE	END[4]
Y_RNA.400	chr8:98784541-98784653	7.4	4.1	–1.8	0.014	5′
Y_RNA.353	chr1:155092966-155093074	7.8	4.6	–1.7	0.041	5′
Y_RNA.31	chrX:135653864-135653974	103.4	63.8	–1.6	0.024	5′
Y_RNA.112	chr3:164840501-164840611	108.4	71.2	–1.5	0.039	5′
Y_RNA.639	chr14:56535161-56535245	9.9	6.5	–1.5	0.035	5′
Y_RNA.367	chrX:19394892-19394993	4.5	8.3	1.8	0.014	5′
Y_RNA.166	chr20:16651286-16651387	10.4	18.4	1.8	0.014	3′
Y_RNA.597	chr2:206890317-206890421	20.2	35.6	1.8	0.029	3′
Y_RNA.535	chr12:42848522-42848623	25.0	43.3	1.7	0.013	3′
Y_RNA.180	chr15:59867827-59867922	10.8	18.2	1.7	0.044	3′
Y_RNA.168	chr14:100049354-100049455	28.3	47.3	1.7	0.007	3′
Y_RNA.450	chr6:34789222-34789319	14.9	24.5	1.6	0.017	3′
Y_RNA.212	chr11:107955640-107955741	24.5	40.1	1.6	0.024	3′
Y_RNA.511	chr11:64063509-64063610	21.8	35.3	1.6	0.022	3′
Y_RNA.481	chr15:30965953-30966046	20.4	33.0	1.6	0.021	3′
Y_RNA.505	chr15:52454948-52455049	21.2	34.0	1.6	0.040	3′
Y_RNA.148	chr6:106902703-106902804	26.8	43.0	1.6	0.032	3′
Y_RNA.469	chr20:431307-431406	14.1	22.6	1.6	0.037	3′
Y_RNA.170	chr4:158689165-158689265	10.0	15.6	1.6	0.031	3′
Y_RNA.595	chr2:113337061-113337161	5.7	8.9	1.6	0.040	3′
RNY4P18	chr9:113859605-113859693	206.0	328.7	1.6	0.037	3′
RNY4P25	chr1:151411476-151411571	507.1	777.0	1.5	0.039	3′
Y_RNA.218	chr17:43148810-43148911	16.8	25.6	1.5	0.035	3′
Y_RNA.699	chrX:41175741-41175842	5.7	8.6	1.5	0.017	3′
Y_RNA.492	chr12:123252646-123252747	15	23	1.5	0.017	3′

Notes: [1]YRNA identity with corresponding genomic positions in the human hg19 genome. [2]Average YRNA read count for the indicated experimental group reported as counts per million (cpm) reads in the sequenced library. [3]Fold change calculated by EdgeR from comparison between the normal and breast cancer serum samples. [4]Indicates whether the sequencing reads map to the 5′ or 3′ end of YRNAs.

of ~20 to 1) (Fig. 2A). However, despite the preponderance of the 5′ YRNA fragments, many more 3′ than 5′ YRNA fragments displayed altered serum levels associated with breast cancer. 3′ end YRNA fragments are 25–29 nt, while 5′ end YRNA fragments can be 25–29 nt or 30–33 nt (Fig. 2B). Thus, there may be a specific association between breast cancer and the minor 25–29 nt population of circulating 3′ YRNA fragments. The functional significance of this specificity in size and YRNA end is not clear, especially since there is little known about the biological role of YRNA fragments.

Thus, we find that the abundance of specific circulating 5′ tRNA halves and 3′ and 5′ YRNA fragments is different when normal individuals are compared with a group of breast cancer patients. This might imply that these processed circulating small RNAs are somehow involved in the disease process. However, the breast cancer group is small, and the characteristics of the individual cases are fairly disparate in terms of stage and treatment status, making it difficult to draw any strong conclusions from this information. For this reason, we examined circulating small RNAs in a larger group of breast cancer cases.

Breast cancer clinical characteristics are associated with changes in abundance of 5′ tRNA halves and YRNA fragments. The findings described above suggest that breast cancer may, either directly or indirectly, exert multifarious effects on levels of circulating small RNAs. While these changes could potentially serve as biomarkers of disease, they might also provide clues to the pathophysiology of systemic effects during cancer progression. To further explore this finding, we made use of existing datasets of serum small RNAs from a group of 42 individuals diagnosed with breast cancer (Gene Expression Omnibus [GEO] Series GSE49035). These specimens were obtained prior to treatment, with information on tumor stage, ER, PR, and HER2 expression, and presence of inflammation. Cases in which relapse subsequently occurred were recorded. The datasets

have previously been analyzed with regard to miRNA expression;[34] we used them to search for associations between characteristics of the tumors and levels of specific 5′ tRNA halves and YRNA fragments.

We analyzed the 42 datasets with the same methods described for the pilot dataset above. We first pooled the datasets and assessed the composition of the various circulating small RNAs in the group as a whole. As in the pilot dataset (Fig. 1D) and in our previous study,[10] we find that the great bulk of reads are derived from miRNAs (45%), tRNAs (41%), and YRNAs (13%), with negligible fractions of other types (Fig. 1E). We then asked if, within this group of breast cancer cases, clinical characteristics were associated with changes in circulating 5′ tRNA halves or YRNA fragments.

5′ tRNA halves. We found that tumor characteristics were associated with changes in circulating 5′ tRNA halves (Fig. 3 and Supplementary Table 1). ER+ tumor cases showed a decline in abundance of 26 specific circulating tRNA halves in comparison to ER– tumors. Similarly, PR+ and HER2+ status was associated with declines in 41 and 8 specific 5′ tRNA halves, respectively. Inflammatory breast cancer, on the other hand, was associated with increases in five specific 5′ tRNA halves, while eventual relapse was associated with increased levels of 18 specific 5′ tRNA halves at the time of diagnosis. Tumor stage was not associated with significant differences in serum levels of any 5′ tRNA halves.

YRNA fragments. We carried out the same comparative analysis for circulating YRNA fragments (Fig. 4 and Supplementary Table 2). Again, tumor characteristics, with the exception of inflammation, were associated with differences in serum levels of some YRNA fragments. The direction of change was consistent in the ER+ and PR+ groups (down) and in the relapse group (up), although not in the HER2+ group. The ER+, PR+, and HER2+ tumor cases were associated with a decrease in abundance of 15, 4, and 13 specific circulating YRNA fragments, respectively. The HER2+ status was also associated with an increase in abundance of three specific circulating YRNA fragments. Relapse was associated with decreased level of only one but the increased levels of 31 circulating YRNA fragments. Tumor stage was associated with decreased level of one and the increased levels of two circulating YRNA fragments.

Inter-individual differences. We asked if the highly significant differences between groups assorted by clinical characteristics were a feature of the groups as a whole, or were caused by a subset of individuals within a group. When levels of the 5′ tRNA halves in individual cases were inspected, we noted that the significant change in expression within a group was driven by a large difference in a few individuals (Fig. 5). Comparison of the 5′ tRNA half expression levels in the ER+ and ER– groups, and in the relapse and nonrelapse groups, shows that there is significant inter-individual variation. Nevertheless,

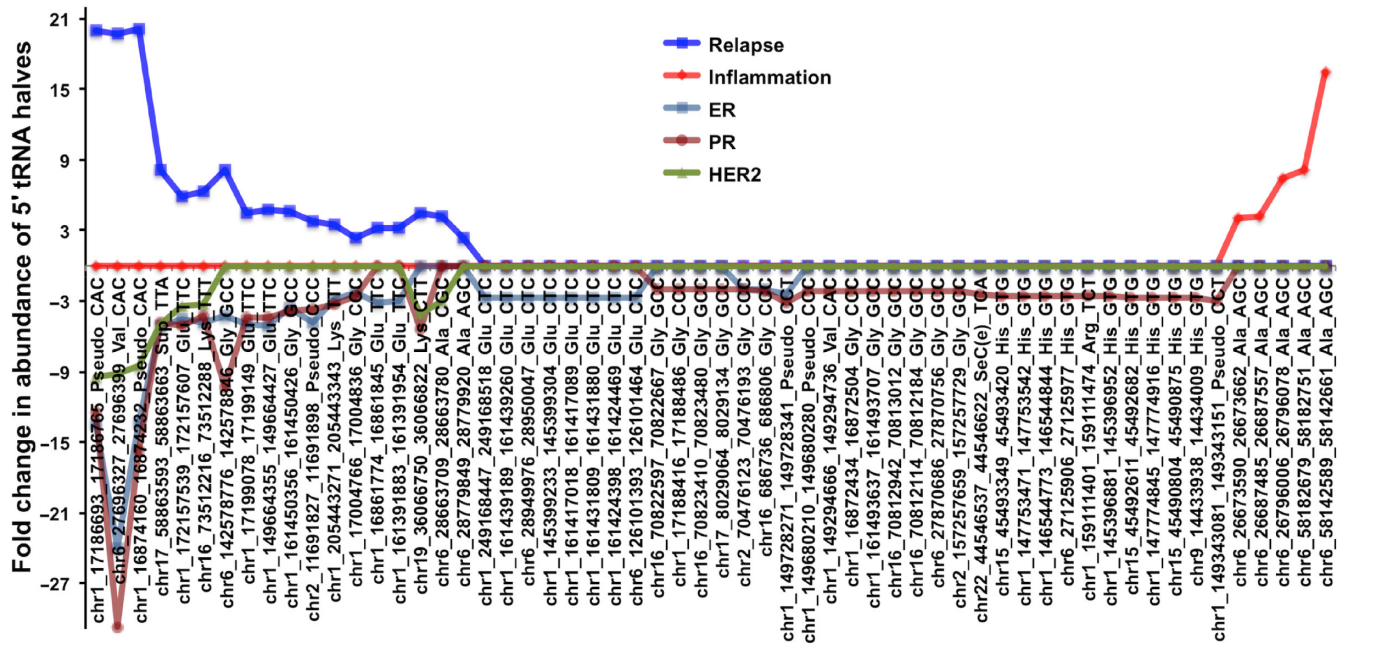

Figure 3. Changes in serum abundance of specific 5′ tRNA halves in breast cancer cases sorted by clinicopathologic characteristics. Shown in the graph are all 5′ tRNA halves found to have significant differences in abundance when 42 newly diagnosed breast cancer cases were compared for clinicopathologic characteristics. The *y*-axis denotes fold change in abundance of a 5′ tRNA half when groups are compared. Eventual relapse is associated with increased serum levels of a small set of 5′ tRNA halves, while ER expression is associated with decreased serum levels of a broader group of 5′ tRNA halves. Inflammatory disease is associated with increased serum levels of a separate group of 5′ tRNA halves. Tumor stage is not shown because no significant differences were found between Stage 2 and Stage 3 cases. An adjusted *P*-value <0.05 was set as the level of statistical significance for the difference between two possible alternatives of each clinical characteristic (stage [2 versus 3], ER [positive versus negative], PR [positive versus negative], HER2 [positive versus negative], presence of inflammation [yes versus no], and subsequent relapse [yes versus no]).

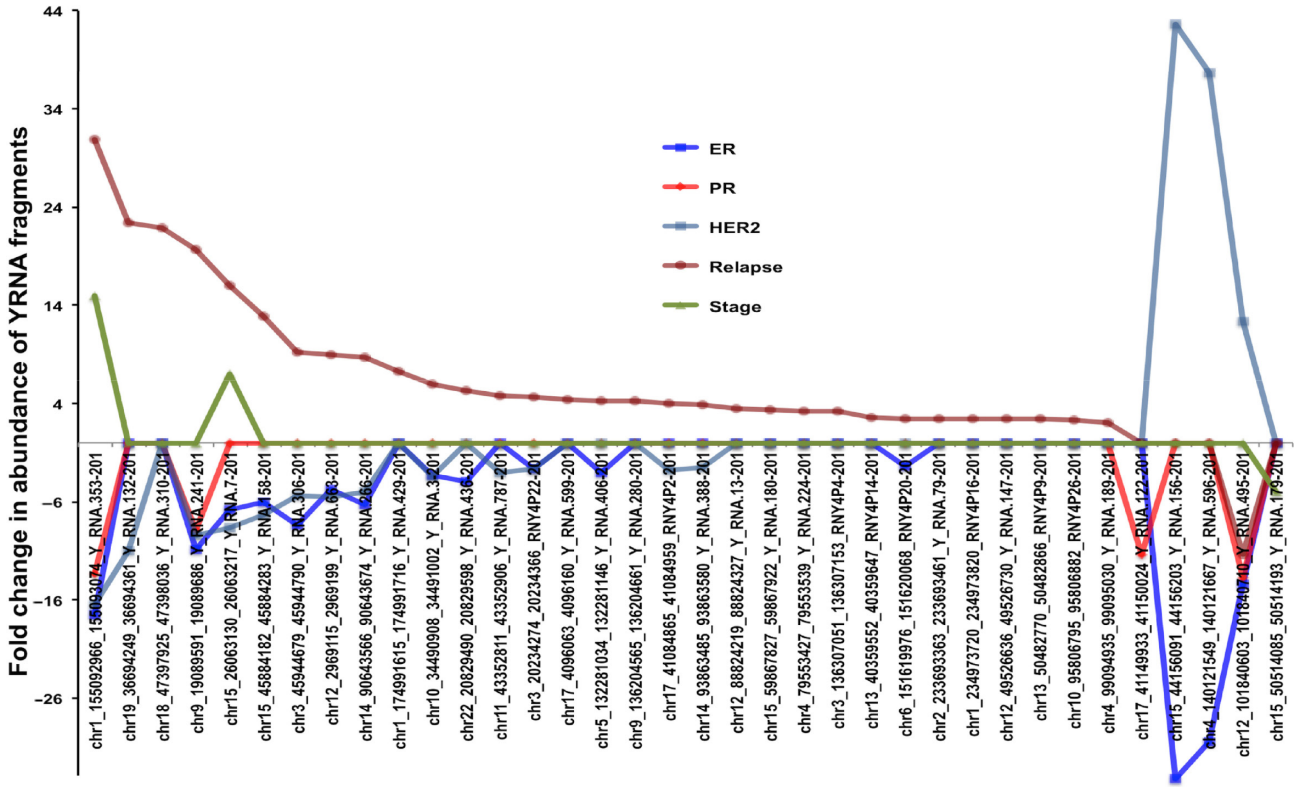

Figure 4. Changes in serum abundance of specific YRNA fragments in breast cancer cases sorted by clinicopathologic characteristics. Shown in the graph are all YRNA fragments found to have significant differences in abundance when 42 newly diagnosed breast cancer cases were compared for clinicopathologic characteristics. The *y*-axis denotes fold change in the indicated YRNA fragments. Eventual relapse was associated with increased serum levels of a broad group of YRNA fragments, while ER expression was associated with decreased serum levels of an overlapping group. Inflammatory status is not shown because no significant differences were found. An adjusted *P*-value <0.05 was set as the level of statistical significance for the difference between two possible alternatives of each clinical characteristic (stage [2 versus 3], ER [positive versus negative], PR [positive versus negative], HER2 [positive versus negative], presence of inflammation [yes versus no], and subsequent relapse [yes versus no]).

when abundance of 5′ tRNA halves changed in association with a clinicopathologic characteristic, the direction of the change was consistent among cases within a group. The most striking finding came from comparison of the relapse and non-relapse groups. Two of the 10 individuals in the relapse group account for most of the difference between these two groups (Fig. 5). Levels of 5′ tRNA halves were very high in these two individuals, while levels in the other eight in the relapse group were uniformly low, as were levels in the 31 cases in the non-relapse group. Both of these two individuals were ER–, PR–, HER2–, Stage 3 without inflammation. This pattern is highly significant; we speculate that this high expression in these individuals is in some way connected to their later relapse.

Similar patterns of inter-individual differences were observed in the levels of circulating YRNA fragments; the differences between groups assorted by clinical characteristics were caused by a subset of individuals within a group.

In summary, we find that serum abundance of both 5′ tRNA halves and YRNA fragments differs between groups of breast cancer cases sorted by tumor characteristics, and these differences seem to be accounted for by a subset of individuals within a group. Do these differences (between normal individuals and breast cancer patients, and between patients whose tumors have different characteristics) have any biological or clinical significance? It cannot be concluded from this data that the level of any specific 5′ tRNA half or YRNA fragment has strong predictive value. However, we speculate that the association between breast cancer and changes in these novel circulating RNA species reflects some aspect of either the biology of the tumor or the reaction of the individual to a tumor.

The significant association between breast cancer characteristics and alterations in serum levels of 5′ tRNA halves raises the possibility of a causal connection between some aspect of the disease process and the 5′ tRNA halves. A more definite conclusion is impeded by the rudimentary state of knowledge regarding the function of 5′ tRNA halves. Currently, the only described function of 5′ tRNA halves is inhibition of translation in cultured cells subjected to a variety of stressors.[21–25,35] Stress-induced tRNA fragments promote assembly of stress granules, and inhibit mRNA translation.[23,25] It is not known if specific subtypes of 5′ tRNA halves have different activities in this mechanism. Whether they originate from the cancer itself, or from cells affected by the cancer, it is tempting to speculate that 5′ tRNA halves may mediate systemic effects of

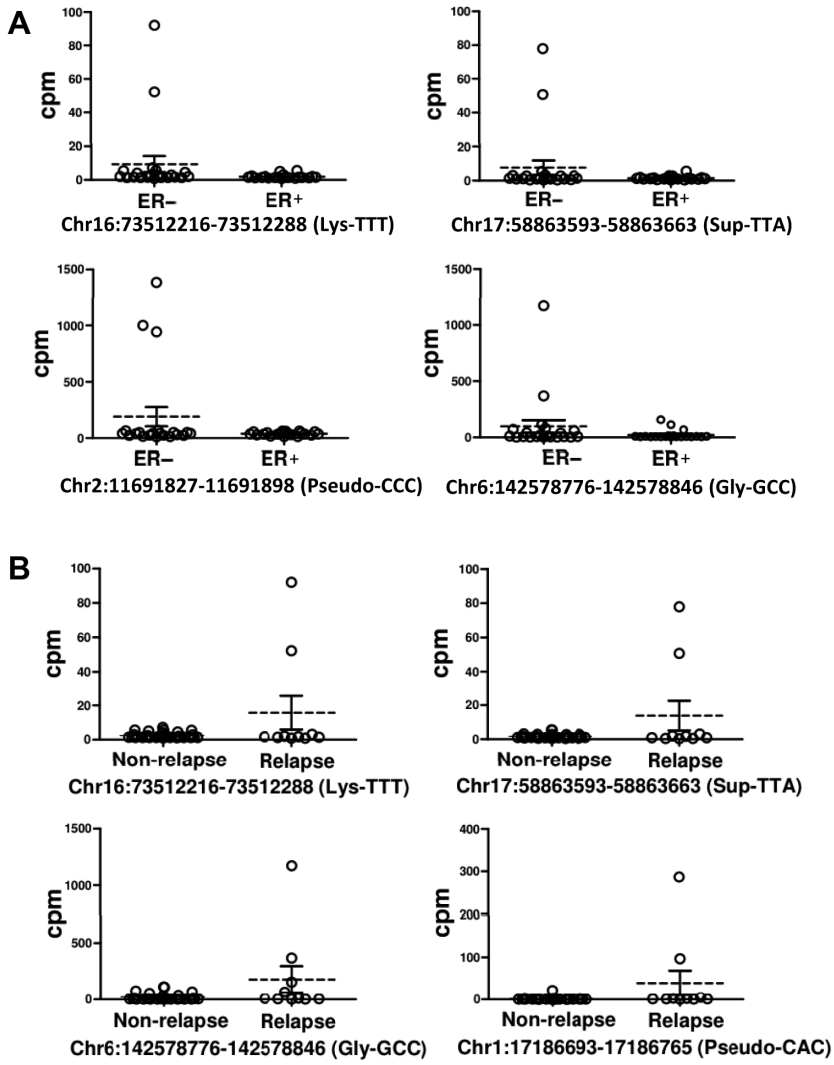

Figure 5. Inter-individual variation in expression of 5′ tRNA halves. Expression levels of the indicated 5′ tRNA halves are represented as counts per million reads (cpm) in each individual breast cancer case. For each group, the mean (dotted line) and the standard error of the mean (error bars) are indicated. Comparisons were carried out between ER– and ER+ groups (**A**), and relapsed and nonrelapsed groups (**B**). Each tRNA isoacceptor is identified by its genomic position in the human hg19 genome. In (**B**), the two individuals showing increased levels of different 5′ tRNA halves are the same in each graph.

cancer. We currently have no information on the abundance of 5′ tRNA halves in breast cancer cells. We have shown that in mice, under normal physiological conditions, 5′ tRNA halves are concentrated in hematopoietic and lymphoid tissues, and present in other tissues at very low levels that may reflect residual blood cells.[11] Although the origin and destinations of the serum particles carrying 5′ tRNA halves are not yet known, the abundance of 5′ tRNA halves in blood cells suggests that they may produce the serum particles. This raises interesting questions in regard to the current findings. Does breast cancer itself change the serum composition of 5′ tRNA halves, or is the change caused by some physiological response to the disease?

Finally, the physiological role of serum 5′ tRNA halves is as yet unknown, but there is clearly potential for them to be taken up by peripheral tissues where they may alter biological processes;[23] thus our finding might suggest a marker of physiologic changes accompanying malignancy. Further insights into this issue will be promoted by a better understanding of the origins and destinations of serum 5′ tRNA halves, and more extensive information regarding their relationship with cancer syndromes.

The findings with YRNA fragments are similar to those with 5′ tRNA halves: differences in the serum abundance of specific subtypes, related to breast cancer diagnosis or breast tumor characteristics. The finding implies that a type- and end-specific biogenesis and/or release of the circulating YRNA fragments is affected by cancer through some mechanism, possibly even direct secretion of fragments by cancer cells. Full-length YRNAs are overexpressed in some solid tumors.[29] 3′ end fragments derived from human RNY5 were detected in MCF7 mammary adenocarcinoma cells.[9] Moreover, 3′ end fragments of human RNY5 were significantly induced upon treatment of cancerous and noncancerous cell lines with the stressor poly(I:C), a double-stranded RNA mimic that is an

immunostimulant.[9] Human YRNA-derived fragments of size range 22–25 nt and 27–36 nt were also shown to be induced in cells exposed to apoptotic stimuli, where the induced YRNA fragments remain bound to Ro after they are cleaved in a caspase-dependent manner.[30]

Conclusions

We have found that, in some individuals diagnosed with breast cancer but not yet treated, specific circulating tRNA halves and YRNA fragments show altered abundance in comparison to the group as a whole. Changes in serum levels of specific 5′ tRNA halves occur in some individuals with ER– tumors, and in individuals who later relapsed. Because the numbers of cases were relatively small, and we have little information on variation in circulating tRNA halves and YRNA fragments in normal individuals, these results do not establish a clear relationship between the disease and the changes. However, we note that changes in specific 5′ tRNA halves always occurred in the same direction, and occurred only in individuals with certain clinical characteristics, consistent with a relationship between disease characteristics and the levels of specific 5′ tRNA halves. A better understanding of this relationship will require a larger number of cases with normal controls.

Materials and Methods

Blood collection and serum preparation. The study was approved by the University of California at Riverside Institutional Review Board and performed in accordance with the Helsinki Declaration II and Standards of Good Clinical Practice. Blood samples were collected from five healthy adult women between 30 and 57 years of age and five women with breast cancer between 47 and 58 years of age after obtaining informed consent. Blood was collected in BD Vacutainer SST tubes (#367985, BD), incubated for 15 minutes at room temperature to allow coagulation, and centrifuged at 5,000 g for 10 minutes. The serum supernatant was transferred to new tubes, centrifuged at 16,000 g for 15 minutes to remove any residual cells and debris, and stored at −80°C.

RNA isolation and small RNA library construction. Total RNA, including small RNA, was isolated with Qiagen miRNeasy kit (#217004) according to the manufacturer's protocol with the following alterations: 2 mL of Qiazol reagent was mixed with 0.4 mL serum and the aqueous phase was loaded onto a single column from the MinElute Cleanup Kit (#74204, Qiagen). RNA was eluted in 20 µL of RNase-free water. One-fourth (5 µL) of the RNA isolated from each serum sample was used to construct sequencing libraries with the Illumina TruSeq Small RNA Sample Prep Kit (#RS-200-0012, Illumina), following the manufacturer's protocol. Libraries' quality was validated by checking the size, purity, and concentration of the amplicons on the Agilent Bioanalyzer High Sensitivity DNA chip (#5067-4626, Genomics Agilent). Equimolar amounts of libraries were pooled and sequenced on an Illumina HiSeq 2000 instrument to generate 50 base reads.

Bioinformatics analysis of sequencing reads. FASTX-Toolkit (hannonlab.cshl.edu) was used to trim the adaptor sequences, and discard low-quality reads. The filtered reads were mapped to the human (hg19) genome with Bowtie version 0.12.8[36] using the "end-to-end k-difference (-v)" alignment mode and allowing up to two mismatches. In addition, this mode of alignment was combined with options (-k 1—best) that instructed Bowtie to report only the best alignment if more than one valid alignment exists. BEDTools[37] was used to annotate the mapped sequencing reads using noncoding RNAs from Ensembl GRCh37 release 70, miRNAs from miRBase 20 (mirbase.org), and tRNAs from Genomic tRNA Database.[38]

The Bowtie alignment files were also analyzed with BEDTools to count the reads that align to tRNA genes in the Genomic tRNA Database and YRNA genes and pseudogenes in the noncoding RNAs of Ensembl GRCh37 release 70. The tRNA and YRNA read counts were further analyzed with the Bioconductor package edgeR[33] to measure changes in the levels of circulating tRNA halves and YRNA fragments in the experimental groups.

Northern blot analysis. RNA for Northern blot analysis was extracted from 0.2 mL serum. The extracted RNA was separated on 15% denaturing polyacrylamide gels, transferred, and fixed to a membrane by chemical cross-linking.[39] Hybridization of the membranes was carried out overnight at 42°C in ULTRAhyb-Oligo Buffer (Invitrogen) with ^{32}P-5′-end-labeled oligonucleotide probes against the 5′ end (5′-AGTTCTGATA ACCCACTACCATCGGACCAGCC) or 3′ end (5′-AGC CAGTCAAATTTAGCAGTGGGGGGTTGTAT) of RNY4. Following hybridization and washing, the blots were exposed to X-ray films for detection of signals.

Analysis of breast cancer data from the GEO public repository. Raw serum small RNA sequencing data in the GEO Series GSE49035[34] was downloaded, preprocessed, mapped to the human (hg19) genome with Bowtie, and annotated with BEDTools according to the parameters used to analyze the pilot samples above. The tRNA and YRNA read counts were analyzed with the Bioconductor package edgeR to assess the associations between changes in the levels of circulating tRNA halves and YRNA fragments and the characteristics of the tumors (stage, ER, PR, and HER2 expression, presence of inflammation, and subsequent relapse). The algorithm of edgeR fits a negative binomial model to the count data, estimates dispersion, and measures differences using the generalized linear model likelihood ratio test, which is well suited for the analysis of experiments with multiple factors, such as the simultaneous analysis of the characteristics of the tumors (stage, ER, PR, and HER2 expression, presence of inflammation, and subsequent relapse) performed here. The fitted count data was analyzed by performing pairwise comparisons between the two possible alternatives of each clinical characteristic: stage (2 versus 3), ER (positive versus negative), PR (positive versus negative), HER2 (positive versus negative), presence of inflammation (yes versus no), and subsequent relapse (yes versus no). EdgeR uses the

Benjamini and Hochberg method to adjust *P*-values to control the false discovery rate. The level of statistical significance for the difference between two possible alternatives of each clinical characteristic was set for an adjusted *P*-value <0.05.

Acknowledgments

We thank Noel Guerrero, Patricia Mote, and Hanh Lam for their help.

Author Contributions

Conceived and designed the experiments: JMD. Analyzed the data: JMD, SRS, HA, DB, DIKM. Wrote the first draft of the manuscript: JMD. Contributed to the writing of the manuscript: DIKM. Agree with manuscript results and conclusions: JMD, SRS, HA, DB, DIKM. Made critical revisions and approved final version: JMD, SRS, HA, DB, DIKM. All authors reviewed and approved of the final manuscript.

Supplementary Data

Supplementary table 1. Fold changes in serum abundance of specific 5′ tRNA halves associated with breast cancer clinicopathologic characteristics.

Supplementary table 2. Fold changes in serum abundance of specific YRNA fragments associated with breast cancer clinicopathologic characteristics.

REFERENCES

1. Okamura K. Diversity of animal small RNA pathways and their biological utility. *Wiley Interdiscip Rev RNA*. 2012;3:351–368.
2. Esteller M. Non-coding RNAs in human disease. *Nat Rev Genet*. 2011;12:861–874.
3. Iorio MV, Croce CM. microRNA involvement in human cancer. *Carcinogenesis*. 2012;33:1126–1133.
4. Avraham R, Yarden Y. Regulation of signalling by microRNAs. *Biochem Soc Trans*. 2012;40:26–30.
5. Rother S, Meister G. Small RNAs derived from longer non-coding RNAs. *Biochimie*. 2011;93:1905–1915.
6. Baraniskin A, Nopel-Dunnebacke S, Ahrens M, et al. Circulating U2 small nuclear RNA fragments as a novel diagnostic biomarker for pancreatic and colorectal adenocarcinoma. *Int J Cancer*. 2013;132:E48–E57.
7. Sobala A, Hutvagner G. Transfer RNA-derived fragments: origins, processing, and functions. *Wiley Interdiscip Rev RNA*. 2011;2:853–862.
8. Lee YS, Shibata Y, Malhotra A, Dutta A. A novel class of small RNAs: tRNA-derived RNA fragments (tRFs). *Genes Dev*. 2009;23:2639–2649.
9. Nicolas FE, Hall AE, Csorba T, Turnbull C, Dalmay T. Biogenesis of YRNA-derived small RNAs is independent of the microRNA pathway. *FEBS Lett*. 2012;586:1226–1230.
10. Dhahbi JM, Spindler SR, Atamna H, Boffelli D, Mote P, Martin DI. 5′ YRNA fragments derived by processing of transcripts from specific YRNA genes and pseudogenes are abundant in human serum and plasma. *Physiol Genomics*. 2013;45(21):990–998.
11. Dhahbi JM, Spindler SR, Atamna H, et al. 5′ tRNA halves are present as abundant complexes in serum, concentrated in blood cells, and modulated by aging and calorie restriction. *BMC Genomics*. 2013;14:298.
12. Vojtech L, Woo S, Hughes S, et al. Exosomes in human semen carry a distinctive repertoire of small non-coding RNAs with potential regulatory functions. *Nucleic Acids Res*. 2014;42:7290–7304.
13. Allegra A, Alonci A, Campo S, et al. Circulating microRNAs: new biomarkers in diagnosis, prognosis and treatment of cancer (review). *Int J Oncol*. 2012;41:1897–1912.
14. Etheridge A, Lee I, Hood L, Galas D, Wang K. Extracellular microRNA: a new source of biomarkers. *Mutat Res*. 2011;717:85–90.
15. Cuk K, Zucknick M, Heil J, et al. Circulating microRNAs in plasma as early detection markers for breast cancer. *Int J Cancer*. 2013;132:1602–1612.
16. Brase JC, Johannes M, Schlomm T, et al. Circulating miRNAs are correlated with tumor progression in prostate cancer. *Int J Cancer*. 2011;128:608–616.
17. Shen J, Stass SA, Jiang F. MicroRNAs as potential biomarkers in human solid tumors. *Cancer Lett*. 2013;329:125–136.
18. Tanaka Y, Kamohara H, Kinoshita K, et al. Clinical impact of serum exosomal microRNA-21 as a clinical biomarker in human esophageal squamous cell carcinoma. *Cancer*. 2013;119:1159–1167.
19. Jung E-J, Santarpia L, Kim J, et al. Plasma microRNA 210 levels correlate with sensitivity to trastuzumab and tumor presence in breast cancer patients. *Cancer*. 2012;118:2603–2614.
20. Healy NA, Heneghan HM, Miller N, Osborne CK, Schiff R, Kerin MJ. Systemic mirnas as potential biomarkers for malignancy. *Int J Cancer*. 2012;131:2215–2222.
21. Fu H, Feng J, Liu Q, et al. Stress induces tRNA cleavage by angiogenin in mammalian cells. *FEBS Lett*. 2009;583:437–442.
22. Saikia M, Krokowski D, Guan BJ, et al. Genome-wide identification and quantitative analysis of cleaved tRNA fragments induced by cellular stress. *J Biol Chem*. 2012;287(51):42708–42725.
23. Ivanov P, Emara MM, Villen J, Gygi SP, Anderson P. Angiogenin-induced tRNA fragments inhibit translation initiation. *Mol Cell*. 2011;43:613–623.
24. Yamasaki S, Ivanov P, Hu GF, Anderson P. Angiogenin cleaves tRNA and promotes stress-induced translational repression. *J Cell Biol*. 2009;185:35–42.
25. Sobala A, Hutvagner G. Small RNAs derived from the 5′ end of tRNA can inhibit protein translation in human cells. *RNA Biol*. 2013;10:553–563.
26. Lerner MR, Boyle JA, Hardin JA, Steitz JA. Two novel classes of small ribonucleoproteins detected by antibodies associated with lupus erythematosus. *Science*. 1981;211:400–402.
27. Reed JH, Sim S, Wolin SL, Clancy RM, Buyon JP. Ro60 requires y3 RNA for cell surface exposure and inflammation associated with cardiac manifestations of neonatal lupus. *J Immunol*. 2013;191:110–116.
28. Christov CP, Gardiner TJ, Szuts D, Krude T. Functional requirement of noncoding Y RNAs for human chromosomal DNA replication. *Mol Cell Biol*. 2006;26:6993–7004.
29. Christov CP, Trivier E, Krude T. Noncoding human Y RNAs are overexpressed in tumours and required for cell proliferation. *Br J Cancer*. 2008;98:981–988.
30. Rutjes SA, van der Heijden A, Utz PJ, van Venrooij WJ, Pruijn GJ. Rapid nucleolytic degradation of the small cytoplasmic Y RNAs during apoptosis. *J Biol Chem*. 1999;274:24799–24807.
31. Nolte-'t Hoen EN, Buermans HP, Waasdorp M, Stoorvogel W, Wauben MH, 't Hoen PA. 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*. 2012;40:9272–9285.
32. Meiri E, Levy A, Benjamin H, et al. Discovery of microRNAs and other small RNAs in solid tumors. *Nucleic Acids Res*. 2010;38:6234–6246.
33. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2010;26:139–140.
34. Wu X, Somlo G, Yu Y, et al. De novo sequencing of circulating miRNAs identifies novel markers predicting clinical outcome of locally advanced breast cancer. *J Transl Med*. 2012;10:42.
35. Emara MM, Ivanov P, Hickman T, et al. Angiogenin-induced tRNA-derived stress-induced RNAs promote stress-induced stress granule assembly. *J Biol Chem*. 2010;285:10959–10968.
36. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol*. 2009;10:R25.
37. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. 2010;26:841–842.
38. Chan PP, Lowe TM. GtRNAdb: a database of transfer RNA genes detected in genomic sequence. *Nucleic Acids Res*. 2009;37:D93–D97.
39. Pall GS, Hamilton AJ. Improved northern blot method for enhanced detection of small RNA. *Nature protocols*. 2008;3:1077–1084.