Both endo-siRNAs and tRNA-derived small RNAs are involved in the differentiation of primitive eukaryote Giardia lamblia.

Permalink
https://escholarship.org/uc/item/8q50h5wn

Journal
Proceedings of the National Academy of Sciences of the United States of America, 111(39)

ISSN
0027-8424

Authors
Liao, Jian-You
Guo, Yan-Hua
Zheng, Ling-Ling
et al.

Publication Date
2014-09-15

DOI
10.1073/pnas.1414394111

Peer reviewed
Both endo-siRNAs and tRNA-derived small RNAs are involved in the differentiation of primitive eukaryote* \textit{Giardia lamblia}

Jian-You Liao\textsuperscript{a,b}, Yan-Hua Guo\textsuperscript{a}, Ling-Ling Zheng\textsuperscript{a}, Yan Li\textsuperscript{a}, Wen-Li Xu\textsuperscript{a}, Yu-Chan Zhang\textsuperscript{a}, Hui Zhou\textsuperscript{b}, Zhao-Rong Lun\textsuperscript{a,c,1}, Francisco J. Ayala\textsuperscript{d,1}, and Liang-Hu Qu\textsuperscript{a,1}

\textsuperscript{a}State Key Laboratory of Biocontrol and Key Laboratory of Gene Engineering of the Ministry of Education, School of Life Sciences, Sun Yat-Sen University, Guangzhou 510275, People's Republic of China; \textsuperscript{b}Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Research Center of Medicine, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou 510120, People's Republic of China; \textsuperscript{c}Key Laboratory of Tropical Diseases Control of the Ministry of Education, Zhongshan School of Medicine, Sun Yat-Sen University, Guangzhou 510080, People's Republic of China; and \textsuperscript{d}Department of Ecology and Evolutionary Biology, University of California, Irvine, CA 92697

Contributed by Francisco J. Ayala, August 15, 2014 (sent for review May 21, 2014; reviewed by Hugo D. Luján and Weixiong Zhang)

Small RNAs (sRNAs), including microRNAs and endogenous siRNAs (endo-siRNAs), regulate most important biologic processes in eukaryotes, such as cell division and differentiation. Although sRNAs have been extensively studied in various eukaryotes, the role of sRNAs in the early emergence of eukaryotes is unclear. To address these questions, we deep sequenced the sRNA transcriptome of four different stages in the differentiation of \textit{Giardia lamblia}, one of the most primitive eukaryotes. We identified a large number of endo-siRNAs in this fascinating single-cell parasitic protozoan and found that they were produced from live telomeric retrotransposons and three genomic regions (i.e., endo-siRNA generating regions [eSGRs]). eSGR-derived endo-siRNAs were proven to target mRNAs in \textit{trans}. Gradual up-regulation of endo-siRNAs in the differentiation of \textit{Giardia} suggested that they might be involved in the regulation of this process. This hypothesis was supported by the impairment of the differentiation ability of \textit{Giardia} when GLDCER, essential for the biogenesis of endo-siRNAs, was knocked down. Endo-siRNAs are not the only sRNA regulators in \textit{Giardia} differentiation, because a great number of tRNAs-derived sRNAs showed more dramatic expression changes than endo-siRNAs in this process. We totally identified five novel kinds of tRNAs-derived sRNAs and found that the biogenesis in four of them might be correlated with that of stress-induced tRNA-derived RNA (siRNA), which was discovered in our previous studies. Our studies reveal an unexpected complex panorama of sRNA in \textit{G. lamblia} and shed light on the origin and functional evolution of eukaryotic sRNAs.

Significance

Small RNAs (sRNAs) are most important regulators in eukaryotes. Although different kinds of sRNAs have been extensively studied in higher eukaryotes, their role remains largely unknown in protozoa. We have systematically investigated the full genome the sRNAs of \textit{Giardia lamblia}, the most primitive eukaryote known. Surprisingly, we have found that two major types of sRNAs (i.e., endogenous siRNAs and tRNA-derived sRNAs) are largely encoded in the genome of \textit{G. lamblia}, whereas canonical microRNAs could not be identified in this parasite. Additional studies showed that both sRNAs might be involved in the differentiation regulation of \textit{G. lamblia}. This study indicates that these two kinds of eukaryotic sRNAs emerged in the early evolution of eukaryotes.
cells (14, 15). However, although several sRNAs have been identified in *Giardia* (16–18) and the antigenic variation in this parasite has been found to be controlled by RNAi pathway (19), little is known regarding the sRNA regulation system in *Giardia*. In fact, the sRNA repertoire of *Giardia* is, at present, very incomplete. A systematic identification of *Giardia* sRNA is, thus, required to facilitate the functional study of the *Giardia* sRNA regulation system, which may also provide insight into the origin and evolution of eukaryotic sRNAs.

In this study, we systematically identify the *Giardia* sRNA and investigate whether sRNAs are involved in the differentiation of *Giardia*, we deeply sequenced the sRNA transcriptomes of trophozoites (6- and 24-h encysting trophozoites) as well as cysts of *G. lamblia*. Through careful analysis of the sequencing data and experimental verification, we have found two kinds of previously unidentifed endo-siRNAs, and have identified five novel kinds of tRNA derived sRNAs. All of the sRNAs identified in this work are up-regulated in the differentiation of the parasite, suggesting that they might play important roles in this process. This hypothesis is further supported by the observation that the differentiation ability of *G. lamblia* was impaired after the knockdown of *Giardia* DICER (GLDICER), which is essential for the bio-genesis of endo-siRNAs.

**Results**

Deep Sequencing Reveals a Drastic Alteration of the sRNA Transcriptome in the Differentiation of *Giardia*. To determine whether the sRNAs are involved in the differentiation of *Giardia*, we deep sequenced the sRNA transcriptomes of trophozoites (6- and 24-h encysting trophozoites) as well as cysts of *G. lamblia*. Through careful analysis of the sequencing data and experimental verification, we have found two kinds of previously unidentified endo-siRNAs, and have identified five novel kinds of tRNA derived sRNAs. All of the sRNAs identified in this work are up-regulated in the differentiation of the parasite, suggesting that they might play important roles in this process. This hypothesis is further supported by the observation that the differentiation ability of *G. lamblia* was impaired after the knockdown of *Giardia* DICER (GLDICER), which is essential for the bio-genesis of endo-siRNAs.

Endo-siRNAs Are Involved in the Differentiation of *Giardia*. mRNA-derived sRNAs and retrotransposon-derived sRNAs are the two most abundant sRNAs in trophozoites (Fig. 1C). It is interesting that they have similar characteristics. (i) Their length distributions are all unimodal (25–28 nt) in the trophozoite library and gradually change to bimodal (20–22 and 25–28 nt) when trophozoites differentiate into cysts (Fig. 2B and Fig. S2E). (ii) A large portion of these two kinds of sRNAs contains posttranscriptionally added 3′-end untemplated nucleotides (Fig. 2B, Fig. S3B, and Dataset S2), and they have length distributions concentrated around 25–28 nt in all of four libraries (Figs. S2E and S3F). Therefore, we first analyzed these two kinds of sRNAs.

Identification of Three sRNA-Generating Regions in *Giardia* Genome. We aligned all mRNA-derived sRNAs to each mRNA and analyzed the characteristics of sRNAs derived from each mRNA

![Figure 1](image1)

**Fig. 1.** Characteristics of *G. lamblia* sRNAs. (A) The composition of the last three 3′-end nucleotides of sRNAs with different numbers of untemplated nucleotides; 1, 2, and 3 mm and Perf refer to sRNAs with 1, 2, 3, and 0 untemplated nt. (B) The length distribution of sRNAs. (C) Composition of the sRNA library of four *G. lamblia* differentiation stages. Trop refers to trophozoite; E6h and E24h refer to encystation 6 and 24 h, respectively.

![Figure 2](image2)

**Fig. 2.** Characteristics of SGR- and retrotransposon-derived sRNAs. The length and 3′-end nucleotide distribution of (A) SGR-derived sRNAs and (B) retrotransposon-derived sRNAs. (C) Northern blot analysis of SGR- and retrotransposon-derived sRNAs. (D) The composition of retrotransposon-derived sRNAs in four sRNA libraries. (E) The sRNA distribution pattern of Gt, Gt, and Gt. Trop refers to trophozoite. E6h, E12h, E24h, and E48h refer to encystation 3, 6, 12, 24, and 48 h, respectively. SGR-sRNAs and RT-sRNAs refer to SGR-derived sRNAs and retrotransposon-derived sRNAs, respectively. RT, retrotransposon.
(Fig. S3C). We found two types of mRNAs that produced sRNAs with totally different characteristics. One type of mRNAs (type I mRNAs) produced most of the mRNA-derived sRNAs (95.4%) in the trophozoite sRNA library, and both of their strands can produce sRNAs. These sRNAs have two characteristics listed above. In contrast, sRNAs derived from the other mRNAs (type II mRNAs) are the SGR-derived sRNAs, and all of them are mainly produced from the plus strand of mRNAs (Fig. S3D and Dataset S2) and have a broad length distribution in the four libraries (Fig. S2F). Also, they do not have posttranscriptionally added 3′ untemplated nucleotides (Fig. S3D and Dataset S2). All of these characteristics indicate that they are not the degradation products of mRNAs. Intriguingly, we found that all type I mRNAs were clustered in the genome and form three clusters (Fig. S3A). Most of them (34 of 39; 87.1%) were in one cluster (Dataset S2), which spans a broad region of 40.55 kb. Notably, the entire region, including the intergenic region of these three mRNA clusters, produced sRNAs with the same characteristics as type I mRNA-derived sRNAs. We named these three regions sRNA generating region I (SGRI), SGRII, and SGRIII (type II mRNAs) (Fig. 2A), according to the workflow shown in Fig. S3C. The overall characteristics of SGR-derived sRNA are the same as type I mRNA-derived sRNA (Fig. 2A and Fig. S3B). SGRI and SGRII produced high abundances of sRNAs (Fig. 2A and Fig. S3 C and E), whereas SGRIII only produced a few sRNA. It is interesting that almost no mRNA signatures [referred to as reads sequenced by high throughput mRNA sequencing technology (mRNA-Seq)] could be detected in SGRI and SGRII by a deep coverage (about 1,000×) sequencing of mRNA, which was performed by Franzén et al. (23), whereas high levels of mRNA signatures could be shown in SGRIII.

We tried to validate the expression of SGR-derived sRNAs using Northern blot. However, we failed to detect the SGR-derived sRNA with the highest RP (0.04%) indicating a relatively low abundance of single SGR-derived sRNA. Thus, we tried a mixed probe, which contained 15 probes designed to target the top 15 (Dataset S3) most abundant SGR-derived sRNAs, and succeeded in detecting the expression of SGR-derived sRNAs using Northern blot. We found that, although the RP of SGR-derived sRNAs was gradually dropped in the differentiation process of *Giardia* (Fig. 1C), their expression levels measured by Northern blot were gradually up-regulated in this process (Fig. 2C). The elevated expression level but dropped RP of SGR-derived sRNAs in encysting and cyst libraries might be caused by the substantial up-regulation of other sRNAs, such as tasRNAs (Figs. 1C and 4D and Fig. S4D), during the differentiation of *Giardia*.

**All Alive but Not Dead Telomeric Retrotransposons Produce Large Numbers of sRNAs.** There are three retrotransposons in the *Giardia* genome: two retrotransposons (GilT and GilM) are located in the telomeric region, whereas the third retrotransposon (GilD) is dead (24). Ullu et al. (16) claimed that GilT but not GilM could produce 26-nt sRNAs with unclear function. Through analysis of our sRNA libraries, we found that both GilT and GilM could generate a large number of sRNAs, which occupied 24.2% of reads in the trophozoite library. This kind of sRNA was the second highest abundance sRNA in that library (Fig. 1C and Fig. S3E). The abundance of GilT-derived sRNAs is much higher than that of GilM-derived sRNAs (Fig. 2D). The dead retrotransposon GilD does not generate any sRNA (Fig. 2E). The sRNAs were uniformly distributed throughout the whole regions of GilT and GilM, and their minus strands generated slightly more sRNAs than the plus strands (Fig. 2E). Because the RP of single retrotransposon-derived sRNA is similar to that of single SGR-derived sRNA, we also used a mixed probe (Dataset S3) to verify the expression of retrotransposon-derived sRNA using Northern blot. The Northern blot result showed that retrotransposon-derived sRNAs were also up-regulated in the differentiation of *Giardia*, which is similar to that of SGR-derived sRNAs (Fig. 2C).

SGR- and Retrotransposon-Derived sRNAs Are both Endo-siRNAs and Might Be Involved in the Differentiation Process of *Giardia*. The length distribution of SGR- and retrotransposon-derived sRNAs is similar to that of the in vitro processing products of GLDICER (Fig. 2A and B) (25), indicating that they might be endo-siRNAs. To verify this hypothesis, we knocked down the expression of GLDICER by constitutive expression of part of its antisense transcripts in trophozoites (19). The substantial down-regulation (68.4% decrease) of GLDICER (Fig. 3A) and up-regulation of variant-specific surface proteins (VSPs) (Fig. 3B) were previously reported targets of the RNAi pathway, confirmed the successful construction of the GLDICER knockdown strain. The expression levels of both retrotransposon- and SGR-derived sRNAs were dramatically decreased when GLDICER was knocked down (Fig. 3C). We further constructed a 3xHA-GLAGO (*Giardia* AGO) overexpressing strain and investigated whether these sRNAs were bound to 3xHA-GLAGO using Northern blot. The Northern blot results showed that both SGR- and retrotransposon-derived sRNAs were bound to 3xHA-GLAGO (Fig. 3D and E). The GLDICER-dependent biogenesis and GLAGO binding of SGR- and retrotransposon-derived sRNAs proved that both of them are endo-siRNAs. Therefore, SGR should be renamed endo-SGR (eSGR).

The gradual up-regulation of endo-siRNAs in the differentiation of *Giardia* suggested that they might have roles in regulation of this process. To verify this hypothesis, we induced the differentiation of trophozoites with GLDICER knocked down to cysts in vitro. We found that the transcription activity of cyst wall protein 1 gene, which is usually used as a marker of *Giardia* differentiation (13), is much lower in the GLDICER knockdown strain than the control one (Fig. 3F), and the cyst formation ability was obviously impaired in the GLDICER knockdown (Fig. 3G). These
evidences supported our hypothesis that endo-siRNAs are involved in the differentiation of *Giardia*. To investigate how endo-siRNAs regulate *Giardia* differentiation, we tried to predict the potential targets of endo-siRNAs. The prediction mainly focused on the eSGR-derived endo-siRNAs, because the retrotransposon-derived endo-siRNAs in other eukaryotes always target the transcripts of retrotransposon or retrotransposon-encoding DNA regions. We found that the expression of mRNAs located in the eSGRIII region was not affected by the GLDICER knockdown (Fig. S4B), indicating that eSGR-derived endo-siRNAs might regulate gene expression predominantly in *trans*. We then predicted the targets of endo-siRNAs through directly aligning them to *Giardia* mRNAs. It is interesting that, when allowing two mismatches, eSGRs-derived endo-siRNAs could target ~36.2% of the *Giardia* genes (Fig. 3H). Gene ontology analysis on potential target genes with single mismatch revealed that eSGR derived endo-siRNAs may regulate the expression of genes with various functions (Fig. S4C).

**Six Kinds of tRNA-Derived sRNAs Are Specifically Induced in the Differentiation of *Giardia***. The deep sequencing data showed that tasRNAs undergo a far more dramatic expression change than endo-siRNAs (Figs. 1C and 4A and D and Fig. S4C) when *Giardia* differentiated from trophozoites to cysts, indicating that they might also play an important role in the differentiation of *Giardia*.

**Identification of Four Kinds of *Giardia* tasRNA.** Interestingly, the length distribution of tasRNAs have four peaks (Fig. 4B) around 20–22, 24–26, 28–30, and 36–38 nt, indicating that there might be four distinct categories of tasRNAs. We mapped the tasRNA sequences to the mature tRNA sequences and found that 20–22, 24–26, and 28–30 nt peaks corresponded to tasRNAs derived from the middle region (most of the anticodon stem loop), 3′ end, and 5′ end of mature tRNAs, respectively (Fig. 4B), which were named tRNA anticodon stem loop-associated sRNAs (actasRNAs), tRNA 3′ end-associated sRNAs, and tRNA 5′ end-associated sRNAs (stasRNAs), respectively. The 36–38 nt peak corresponded to one sRNA that covers the whole 5′- end exon of pseudotRNA-Gln (TTG) (Fig. 4E). This sRNA was named as 5′-end exon of pseudotRNA-Gln (TTG) derived sRNA (5EsRNA). Other than these four kinds of sRNAs, there are also other kinds of tRNA-derived sRNAs in all of four libraries, which might be the degradation products of tRNAs, because their abundances are low and do not have obvious length distribution pattern. The RPs of these tRNA degradation products gradually drop during the differentiation of *Giardia* (Fig. 4C).

We then used Northern blot to verify the expressions of 15 tasRNAs with various abundances from 5 tRNAs (Fig. 4D and Fig. S4D). We found that the expression patterns of tasRNAs determined by Northern blot were not quite consistent with those determined by RP value, suggesting that RPs do not faithfully reflect the expression levels of tasRNAs (Fig. 4D and Fig. S4 D and E). To obtain the overall expression patterns of tasRNAs, we used the Northern blot data of eSGR-derived endo-siRNAs, in which the expressions were up-regulated at a relatively lower rate throughout the differentiation of *Giardia* than that of tasRNAs (Fig. 2C) as an inner control to roughly normalized abundance of tasRNAs (details are in Materials and Methods). After normalization, the deep sequencing data were shown to be consistent with Northern blot results (Fig. 4D and Fig. S4 D and E), and we found that all tasRNAs were gradually up-regulated in the differentiation of *Giardia* (Fig. 4D and Fig. S4 D–G). It is worthy to note that, although all three kinds of tasRNAs of most tRNAs could be detected by deep sequencing (Fig. S4E), the RP of part of them was low and could not be detected by Northern blot (Fig. 4D and Fig. S4D). The variation of expression levels of different tasRNAs (Fig. S4E) implies that their expressions might be under stringent regulation. The expression pattern of 5EsRNA was also verified by Northern blot (Fig. 4E). It is interesting that we could not detect the mature tRNA of pseudotRNA-Gln (TTG) and could only detect its precursor. Unlike tasRNAs, 5EsRNA are highly expressed in trophozoite and rise substantially in cysts.

**Identification of a Novel Kind of Stress-Induced tRNA-Derived RNA.** Intriguingly, four probes targeting to the 5′ end of tRNAs could also detect a novel kind of sRNAs with a slightly longer length (~50 nt) than sitRNAs (stress-induced tRNA-derived sRNAs) (~46 nt), which were identified by our previous work and could not be detected by probe target to the tRNA 5′ end (17). Notably, the length of this novel sRNA (~50 nt) is approximately equal to the sum of the lengths of 5tasRNAs (~29 nt) and actasRNAs (~21 nt), suggesting that this novel sRNA shares the same 3′ end with actasRNAs. The simultaneous detection of this sRNA and sitRNA by the probe used to detect actasRNAs supports this assumption (Fig. 4D and Fig. S4D). We named this kind of sRNA derived from the 5′ end of mature tRNAs (5sitRNA). It is the opposite of the previously reported tRNA 3′ end-derived sitRNAs (3sitRNAs). In contrast to 3sitRNA, 5sitRNA is preferentially expressed in the late stage of *Giardia* differentiation.
tasRNAs and sitRNAs Might Be Produced from Two Endonuclease Cleavage Sites in Mature tRNA. The tasRNAs might be produced by the cleavage of the mature tRNAs by one or more unknown endonucleases. We found that the cleavage mainly occurred at two sites in the mature tRNA: one site is at the 5′ end of the anticodon stem [cleavage site I (CSI)], and one site is at the 3′ end of the TPC stem (CSII). The cleavage site and frequency did not change during the differentiation of Giardia. Interestingly, the CSI is the cleavage site used to generate the 3′tRNAs by Giardia, which we reported years ago (17), and the 3′ end of 5′tRNA is located at CSII, implying a connection between the biogenesis of tasRNAs and sitRNAs. It is possible that the endonuclease cleavage of CSI and CSII in the mature tRNAs was activated in the differentiation of Giardia and produced five sRNA products simultaneously (Fig. 4 F and G). It is worth noting that, although the expression levels of five tRNA cleavage products were up-regulated in the differentiation, the expressions of the various tRNAs were stable in this process (Fig. 4D and Fig. S4D), indicating that the production of various tasRNAs and sitRNAs were the result of a purposely controlled process but not the degradation of tRNAs.

In summary, our data revealed that extensive tRNAcleavages were involved in the differentiation process of G. lamblia and that cleavages produce six kinds of tRNA-derived sRNAs, with expression patterns indicating that they might be involved in the differentiation of Giardia.

Discussion
sRNAs are among the most important regulators in eukaryotic species. They provide sequence specificity for regulating gene expression (2) or act as ligands of protein (26) in higher eukaryotic species. They provide sequence specificity for regulating gene expression (2) or act as ligands of protein (26) in higher eukaryotic species. It is well-known that the conservation of these two kinds of sRNAs suggests that they may play critical roles in eukaryotic cells. It is well-known that the Giardia genome encodes three key proteins of the RNAi pathway [i.e., GLAGO, GLDICER, and Giardia RNA-dependent RNA polymerase (GLRDRP)] (19). Because GLAGO and GLDICER are two key proteins of the miRNA pathway, almost all previous studies on sRNA of Giardia were focused on the identification of Giardia miRNA in trophozoite (18, 27–33). As far as we know, 166 miRNAs were reported from trophozoite of G. lamblia by different laboratories through analysis of homology searching, in silico prediction, or deep sequencing (Table S1) (18, 27–33). However, by very careful analysis and experimental verification (Figs. 2 and 3), we surprisingly found that most sRNAs (90.8%) in trophozoite are endo-siRNAs (Fig. 1C and Fig. S3E); only 24% (40) of these reported miRNAs could be detected by our deep sequencing. Among 40 Giardia miRNAs found in our trophozoite sRNA library, 12.5% (5) of them were shown from known ncRNAs, whereas 50% (20) of them were derived from the eSGR region (see below). We consider, therefore, that these eSGR-derived miRNAs might be endo-siRNAs but not miRNAs. Both plus and minus strands of the whole precursor region of these miRNAs were covered by a large number of eSGR-derived sRNAs (Fig. S44). As a matter of fact, the sRNA distribution pattern is significantly different from the canonical miRNA precursor, which only has its plus strand of the mature miRNA region covered by sRNAs (34). Unfortunately, Friedländer et al. (34), who reported the miRNAs from G. lamblia, did not provide convincing evidence to prove whether these inferred miRNA precursors had a real hairpin structure. The total reads number of 15 expressed canonical miRNAs only accounts for 0.0048% of all reads in the trophozoite sRNA library (Table S1), which is much less than the abundance of eSGR-derived endo-siRNA (60.9%) (Fig. S3E). Therefore, in our opinion, although the Giardia genome can really encode canonical miRNAs, they may play a less important function given their extremely low abundance in this parasite.

Giardia endo-siRNAs consist of two major types: one type is retrotransposon-derived endo-siRNAs, and the other type is eSGR-derived endo-siRNAs. eSGRs are three genomic regions identified in this study where endo-siRNAs are produced. They produced almost all non–retrotransposon-derived endo-siRNAs in trophozoites of G. lamblia. The expression of endo-siRNAs was up-regulated in the differentiation process of Giardia, suggesting that they might be involved in the regulation of differentiation in Giardia. This hypothesis was strongly supported by the observation that the differentiation ability of G. lamblia was impaired after the knockdown of GLDICER, which is essential for the biogenesis of endo-siRNAs. The relatively high expression level of endo-siRNAs in trophozoites indicated that they might also be involved in the regulation of basic biologic processes of Giardia. This assumption was supported by the previous excellent finding that the surface antigenic variation mechanism was disrupted after the knockdown of GLDICER or GLRDRP (19). It is worth noting that, although Prueca et al. (19) speculated that VSP-derived endo-siRNAs were related to the regulation of surface antigenic variation in cis, we only found that low levels of VSP-derived endo-siRNAs express in Giardia trophozoite (Dataset S2). Low expression level of VSP-derived endo-siRNAs might be caused by the strain of Giardia that we used, which expressed several VSPs. It is also possible that low levels of VSP-derived endo-siRNAs can regulate the antigenic variation or that some non–VSP-derived siRNAs may be involved in the regulation of antigenic variation of G. lamblia. Intriguingly, we found that eSGR-derived endo-siRNAs could largely regulate gene expression in trans, because the silencing of GLDICER did not change the expression of mRNA embedded in eSGRIII (Fig. S4B). Therefore, it is possible that VSP mRNAs can be regulated by endo-sRNAs that are not derived from VSP in trans. Recent findings support this hypothesis, because evidence indicated that RNA-4 is located in eSGRIII (Fig. S4D) (30), and we suggested that it might be an endo-siRNA and able to regulate VSP mRNA in trans.

Endo-siRNAs are not the only sRNA regulators in Giardia differentiation. We found that six different kinds of tRNA-derived sRNAs were specifically expressed during the period of differentiation in this parasite (Figs. 4 and 5), which indicates that tRNA-derived sRNAs may also play a pivotal role in the differentiation of Giardia. Although tRNA-derived sRNAs were
proven to be biologically functional in eukaryotes (9, 17), little is known about their biogenesis, characteristics, and classification. Results from this study clearly showed comprehensive tRNA-derived sRNAs in the early evolution of eukaryotes (Fig. 5), and they offer a clue to the biogenesis of various tRNA-derived sRNAs. Among six kinds of tRNA-derived sRNAs, five sRNAs are novel ncRNAs that were identified in this study. Thus, this work has significantly expanded the repertoire of functional sRNAs in G. lamblia. It is hard to figure out how these tRNA-derived sRNAs regulate G. lamblia differentiation, because each of them has a large number of members, and we currently lack appropriate methods to alter their expressions. The identification of endonucleases responsible for the biogenesis of these sRNAs may overcome this obstacle. In fact, we have identified two potential endonuclease cleavage sites in mature tRNAs of G. lamblia, which may be responsible for the generation of 5′tsRNAs, tRNA 3′ end-associated sRNAs, actaRNAs, 5iRNAs, and 3siRNAs (Fig. 4F). The features of the sequences and structures of these two cleavage sites may help to find the endonucleases involved in the process of expression of mature tRNAs into various sRNAs.

In somatic cells of higher multicellular eukaryotes, like mammals, sRNAs are also important regulators of many important biological processes such as cellular differentiation (2), and the main regulators are miRNAs but not endo-siRNAs or tRNA-derived sRNAs. Our findings suggest that endo-siRNAs and/or tRNA-derived sRNAs might be the main sRNA regulators in early evolution of eukaryotes or if canonical miRNA does not exist. This scenario is reminiscent of tRNA-derived sRNAs and/or endo-siRNAs with interacting RNAs, which might be the main regulators in germ cell and very early developmental stages, but miRNAs do not function in these stages (35–37). Therefore, to better understand the functions of siRNAs and miRNAs in the differentiation regulation of protozoan and metazoan, respectively, we will need additional insight into the evolution of these cell differentiation regulators.

In summary, we have provided the first comprehensive tRNA transcriptome panorama, to our knowledge, of the primitive eukaryote G. lamblia. The data have revealed an unexpected complex sRNA system used in the regulation of differentiation of this parasite (Fig. 5). Our work provides insight into the function, biogenesis, and evolution of endo-siRNAs and tRNA-derived sRNAs in eukaryotes.

Materials and Methods

G. lamblia isolate WB clone C5 (ATCC 50803) was cultured as described previously (38). Encystation and cyst purification were carried out as reported (38). Total RNAs were isolated using TRIzol. sRNA library preparation and deep sequencing were performed by BGI. G. lamblia genome assembly A 2.0 was used throughout this study. Northern blot and quantitative RT-PCR were performed as described (39). All primers used in this study’s analysis are listed in Dataset S3. Tophozoites were transfected using the LONZA electroporation system. Additional details of materials and methods are supplied in SI Materials and Methods.

ACKNOWLEDGMENTS. The authors thank Barbara J. Davids for providing the protocol for G. lamblia culture. We greatly appreciated the critical comments and suggestions of the two reviewers, which were very useful for improving the manuscript. J.-Y. L. thanks the Guangdong Province Key Laboratory of Computational Science and the Guangdong Province Computational Science Innovative Research Team for support. This work was supported by National Natural Science Foundation of China Grants 30870530 (to H.Z.), 31401092 (to J.-Y.L.), 31401975 (to L.-L.Z.) and 31272305 (to Z.-R.L.), Sun Yat-Sen University Grant 12gic11 (to Z.-R.L), and National Basic Research Program (973 Program) Grant 2011CB811300 (to H.-Q.).
