



# Self-circularizing 5'-ETS RNAs accumulate along with unprocessed pre ribosomal RNAs in growth-stressed *Entamoeba histolytica*

Abhishek Kumar Gupta<sup>1</sup>, Sunil Kumar Panigrahi<sup>1\*</sup>, Alok Bhattacharya<sup>2</sup> & Sudha Bhattacharya<sup>1</sup>

<sup>1</sup>School of Environmental Sciences, Jawaharlal Nehru University, New Delhi, India, <sup>2</sup>School of Life Sciences, Jawaharlal Nehru University, New Delhi, India.

SUBJECT AREAS:  
TRANSCRIPTION  
PARASITOLOGY  
BIOTECHNOLOGY  
NON-CODING RNA'S

Received  
25 November 2011

Accepted  
15 February 2012

Published  
6 March 2012

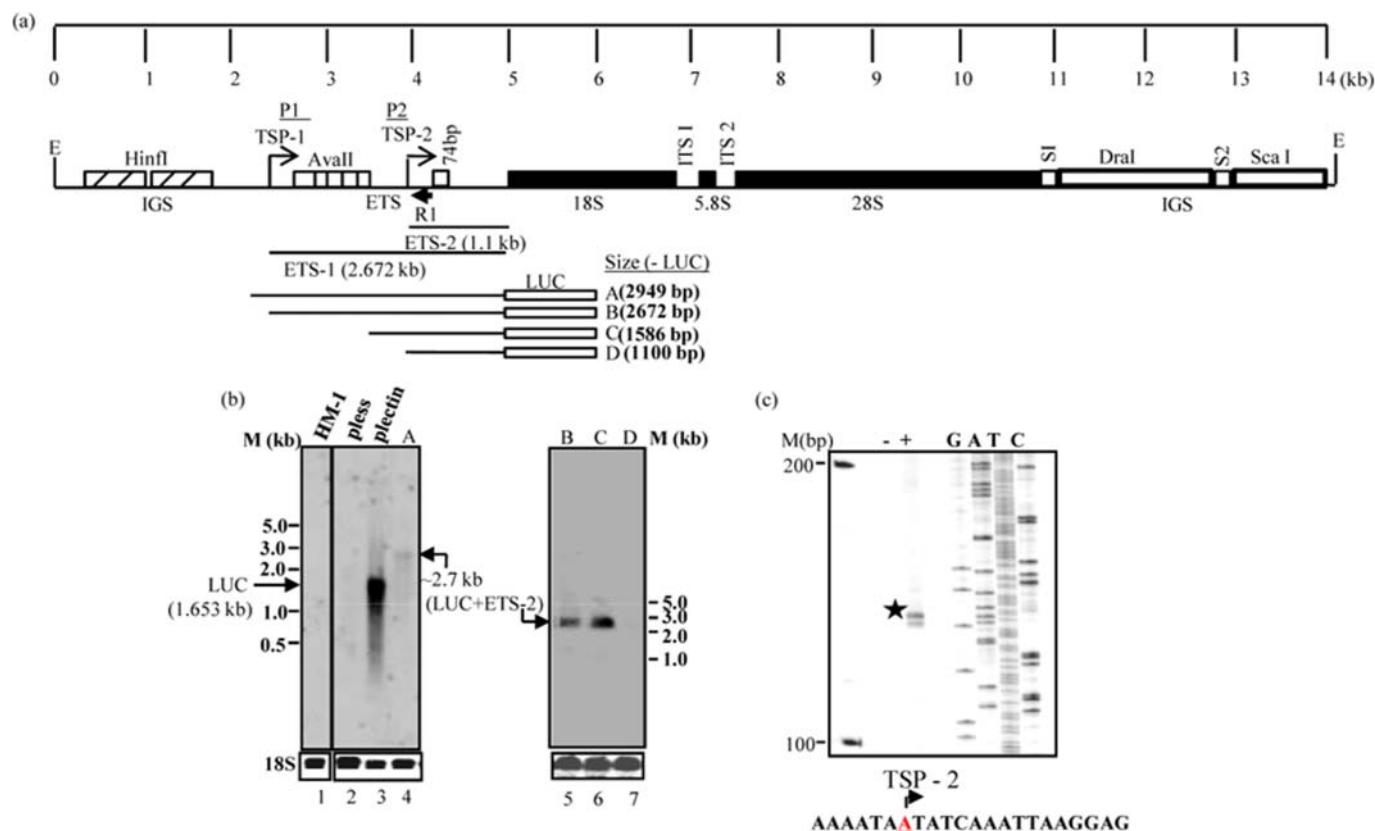
Correspondence and  
requests for materials  
should be addressed to  
S.B. (sbjnu110@  
gmail.com)

\* Current address:  
Department of  
Genetics and  
Development,  
Columbia University  
Medical Center, New  
York, USA.

The primary transcript of rRNA genes is a large pre rRNA which is precisely processed to release the mature rRNAs. The 5'-external transcribed spacer (ETS) of rRNA genes contains important sites for pre rRNA processing. Once the processing is accomplished the ETS is rapidly degraded. We show that in growth-stressed cells of the human parasitic protist *Entamoeba histolytica* the A'-A<sub>0</sub> sub-fragment of the 5'-ETS accumulates to high levels as a family of RNA molecules of size 666 to 912 nt. These etsRNAs are circular *in vivo* and can spontaneously self-circularize *in vitro*. The accumulation of etsRNAs is accompanied by accumulation of unprocessed pre rRNA, indicating a possible role of etsRNAs in inhibition of processing during growth stress. Our data shows for the first time that processed etsRNA is not a mere by-product destined for degradation but is stabilized by circularization and could play a regulatory role as noncoding RNA.

Ribosome biogenesis in any cell type is a highly energy consuming process and responds to general metabolism and to specific environmental challenges in a tightly regulated manner<sup>1-3</sup>. Ribosomal RNA genes are transcribed by RNA polymerase I as precursor molecules (pre rRNAs) which are processed to remove the external (5'- and 3'-ETS) and internal (ITS 1 and 2) transcribed spacers to yield the mature rRNA species (18S, 5.8S and 28S)<sup>4-7</sup>. The 5'-ETS is generally the longest spacer in the pre rRNA. It contains conserved regions of sequence complementarity with the box C+D snoRNA U3<sup>8</sup> which, in combination with a large number of protein factors, forms modules that sequentially assemble on the nascent pre rRNA to give rise to the 90S pre ribosomal particle<sup>9-11</sup>. This dynamic interaction leads to pre 18S rRNA compaction followed by cotranscriptional cleavage within ITS1 and release of the pre 40S subunit. Defects in this hierarchical process target the misassembled pre rRNAs for selective 3'-5' exoribonucleolytic digestion by the exosome<sup>12</sup>. Thus the 5'-ETS has a crucial role in initiating the correct processing and assembly of pre ribosomal particles. Following the assembly process, the excised 5'-ETS is rapidly degraded, presumably by the 3'-5' exonuclease activity of the exosome<sup>13,14</sup>, as experiments conducted in mouse<sup>15</sup>, yeast<sup>16</sup> and *Arabidopsis thaliana*<sup>17</sup> showed that specific fragments of the 5'-ETS accumulated in cells depleted of exosomal components.

We have been investigating the regulation of rRNA synthesis under growth stress in the primitive parasitic protist *Entamoeba histolytica*. In this organism the rRNA genes are located exclusively on extrachromosomal circular molecules<sup>18,19</sup>. Each circle may contain either one copy of the rDNA transcription unit, or two copies organized as inverted repeats. In *E. histolytica* strain HM-1:IMSS, the 14 kb rDNA circle designated EhR2<sup>20</sup> contains one rDNA unit (Figure 1a). The 2.672 kb 5'-ETS precedes the 1.921 kb 18S rRNA gene, followed by ITS1 (149 bp), 5.8S rRNA gene (150 bp), ITS2 (123 bp), 28S rRNA gene (3.544 kb), and the intergenic spacer (5.183 kb) (Figure 1a). Here we report that in this system the pre rRNA begins to accumulate when cells are subjected to growth stress. Concomitantly there is accumulation of a large sub fragment of the 5'-ETS. This RNA fragment undergoes spontaneous circularization and the circular RNA species accumulate intracellularly. To our knowledge this is the first report of a circular non coding RNA species derived from the ETS.

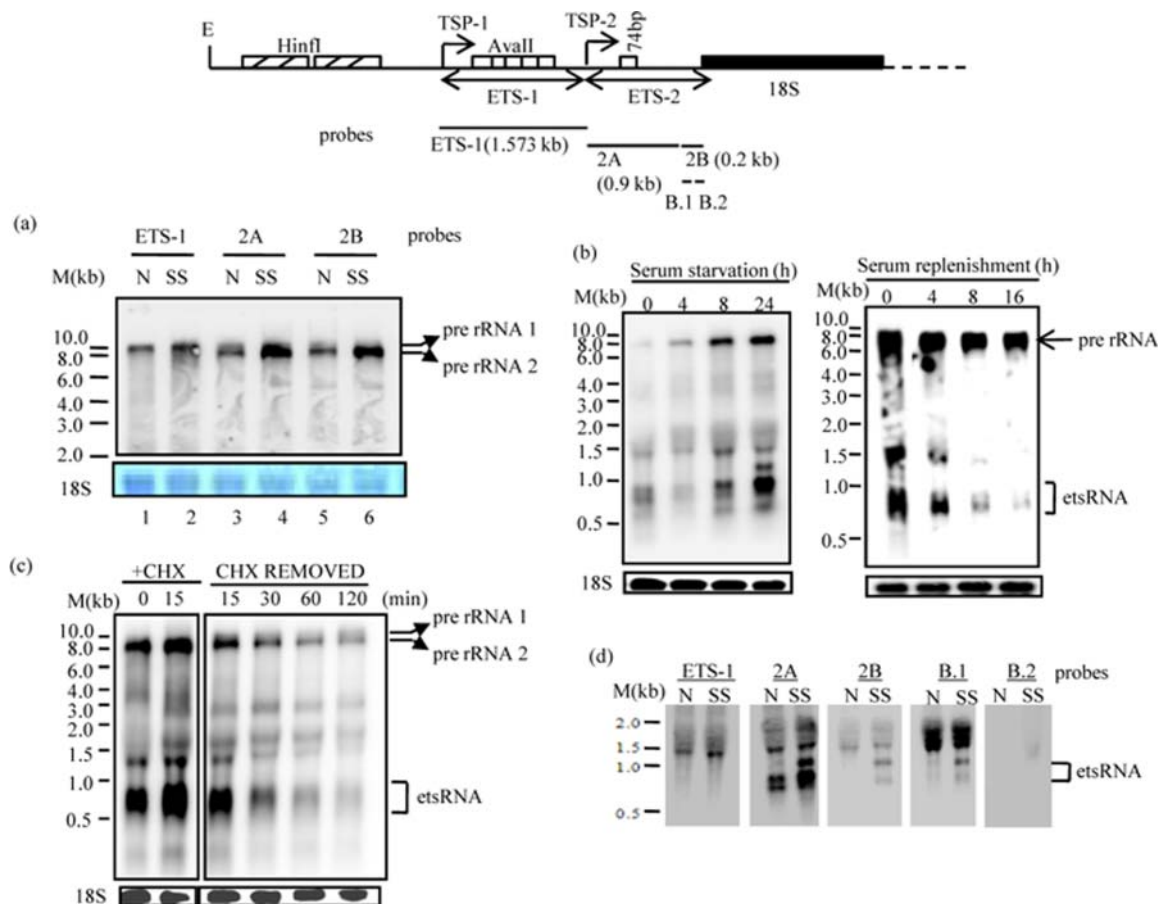


**Figure 1 | Two promoters (P1 and P2) drive transcription of rDNA.** (a) Schematic linear view of rDNA transcription unit and flanking sequences of the circular rDNA, EhR2<sup>20</sup>. Different classes of repeats present in IGS and ETS are marked as HinfI, AvaII, 74 bp, DraI and ScaI. S1 and S2 are spacer sequences present between the repeats. TSP-1, TSP-2, ETS-1 and ETS-2 corresponding to the two promoters are shown. The region cloned in each luciferase reporter construct is indicated. (b) Mapping of promoter P2. Luciferase reporter constructs (A to D) indicated above were transfected to obtain stable *E. histolytica* cell lines. Luciferase expression was measured by northern hybridization using luciferase gene as probe. Controls include RNA from untransfected cells (lane 1), cells transfected with promoter-less construct (lane 2), lectin promoter (lane 3, which gives 1.653 kb band of luc). Transcripts originating from P-2 correspond to a fusion transcript of ETS-2 and luciferase (2.7 kb). 18S RNA was used as a loading control. (c) Primer extension was performed with primer (R1) located downstream of expected TSP of P2. Products were analyzed on 6% urea-PAGE along with sequencing reaction with same primer, followed by autoradiography; star indicates the longest extended product of size 134 nt. The sequence is indicated below. IGS (Intergenic Spacer), ETS (External Transcribed Spacer), ITS (Internal Transcribed Spacer), TSP (Transcription Start Point), E-EcoRI site.

## Results

**Mapping the second promoter of rDNA in *E. histolytica*.** The rDNA promoter had earlier been mapped 2.672 kb upstream of the 5'-end of 18S rRNA<sup>21,22</sup>. We now show the presence of a second stronger promoter downstream to the previously mapped promoter. This was discovered when the expression of luciferase reporter gene (1.6 kb) cloned downstream of the full length 5'-ETS (construct A spanning -277 to +2.672 kb) was checked by northern blot analysis in a stably transfected cell line (Figure 1a, b). The size of transcript expected from the promoter (marked P1) was 4.27 kb but the observed size was 2.7 kb (Figure 1b, lane 4). Stable lines were made with the indicated deletion constructs B, C and D. Northern analysis (Figure 1b, lanes 5–7) showed that the 2.7 kb band did not originate from promoter P1 and a second promoter (P2) existed downstream. The transcripts from promoter P1 could, however, be detected by RT-PCR with luciferase primers (see supplementary Figure S1). Since pre rRNA originating from promoter P1 is readily detected in total cellular RNA (as shown in Fig. 2), we believe that the low level transcription seen with this promoter in construct A may be due to the lack of upstream IGS sequences which may be required for efficient transcription. The tsp of P2, mapped by primer extension using primer R1, was located 1.1 kb upstream of the 5'-end of 18S rRNA (Figure 1c). Most of the subsequent experiments pertain to the transcript from promoter P2, as it responded in a dramatic manner to growth stress.

**Pre rRNA and etsRNA accumulate under growth stress.** We measured the levels of pre rRNA by northern blot analysis using ETS probes. In normal cells the 1.5 kb ETS-1 probe hybridized with a ~8.5 kb band which corresponds to the pre rRNA from promoter P1 (Figure 2a, lane 1). Probes from ETS-2 (and from ITS1, 5.8S and ITS2; data not shown) hybridized faintly to this band and strongly to a lower band corresponding to pre rRNA from promoter P2 (Fig. 2a, lanes 3 and 5). When cells were subjected to growth stress (serum starvation, cycloheximide (CHX) treatment), there was accumulation of pre rRNA. This was especially pronounced for pre rRNA-2 (Fig. 2a, lanes 4 and 6). The accumulation of transcript originating from promoter P2, under growth stress, was also seen in cell lines stably transfected with the LUC reporter construct driven by this promoter (see supplementary Figure S1). This accumulation was not due to slow turnover of luciferase transcript during serum starvation since no accumulation was seen under the same conditions when LUC was driven by another *E. histolytica* promoter, EhTMKB1-9<sup>23</sup>. After 24 h of serum starvation pre rRNA levels, measured by ETS-2 probe, increased by ~2.5 fold (average of 3 independent measurements) (Figure 2b). CHX treatment for 15 min also led to elevation of pre rRNA levels by ~1.5 fold (average of 3 independent measurements; Figure 2c). Upon removal of stress conditions the pre rRNA levels reverted to the levels in normal cells (Fig. 2b, c). Interestingly, the ETS-2 probe hybridized with two broad bands of 0.7 and 0.9 kb in



**Figure 2 | pre rRNA and etsRNA accumulate under growth stress.** Top panel shows the location of ETS probes used for northern analysis. (a) 10  $\mu$ g of total RNA was isolated from normal cells (N) and after 24 h serum starvation (SS). RNA samples were electrophoresed, blotted and hybridized with indicated DNA probes. (b) cells were subjected to different lengths of serum starvation, and after 24 h the serum was replenished. (c) Cells were treated with cycloheximide (100  $\mu$ g/ml) for 15 min, after which cells were transferred back to medium without CHX. In both (b) and (c) 10  $\mu$ g of total RNA from each time point was electrophoresed and hybridized with ETS-2A probe. (d) Northern analysis as in (a). Blots were hybridized with the indicated probes. Single stranded RNA ladder (Invitrogen) was used as size marker. The pre rRNA and etsRNA bands are indicated on the right.

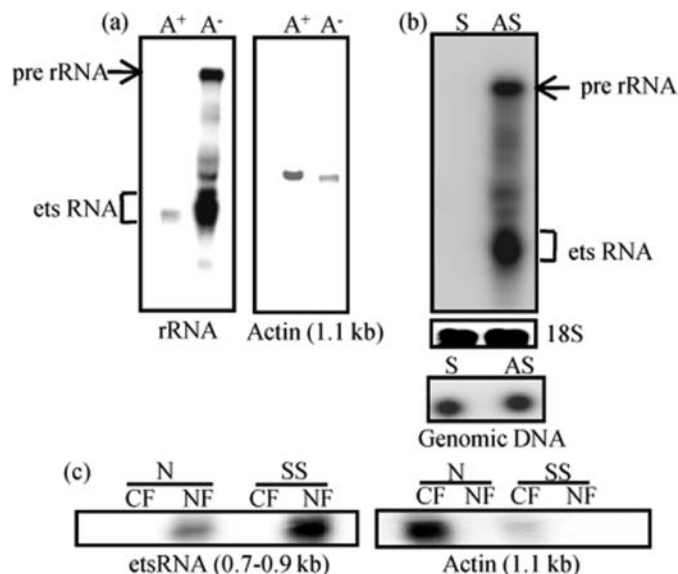
normal cells (Figure 2 b–d). The levels of this RNA (and a 1.1 kb band faintly visible in normal cells) increased greatly during stress ( $\sim 3$  fold, average of three measurements). These RNAs did not hybridize with the ETS-1 probe (Figure 2d) or with ITS and 5.8 S probes (data not shown). They also did not hybridize with the B.2 probe containing the last 89 bp of ETS-2. We call these RNA species, specific to the ETS-2, which accumulate during stress, as etsRNA. Further analysis of the etsRNA showed that it was not polyadenylated as it was detected only in the poly A<sup>-</sup> fraction (Figure 3a). It was transcribed in the same direction as rRNA as it did not hybridize with the single-stranded sense probe (Figure 3b). It was nuclear localized as it was not detected in cytoplasmic fraction (Figure 3c). The accumulation of etsRNA under stress was a significant observation since the excised ETS is generally considered to be an unstable by product of pre rRNA processing.

**etsRNA consists of circular molecules.** In order to map the ends of the etsRNA species we performed circular RT-PCR after incubation of total RNA with RNA ligase. Surprisingly we found that RT-PCR products were obtained when out-facing primers were used even in the absence of RNA ligase, suggesting that these molecules could be circular. We first performed the RT reaction with primer R1 at 37°C and the PCR with primer pair R1/F1 (Figure 4 a, b). Bands specific to ETS-2 were detected by Southern analysis; the corresponding bands were excised from the gel, further purified and sequenced. The junction points of the circles so determined showed two circular

RNA molecules of sizes 766 nt and 666 nt. Similar results were obtained when the RT reaction was done with a second reverse primer R2. The circularization event in both had occurred at the same 5'-junction (+102G), while the 3'-junction had shifted by 3 nt in the two circles (+867A and +864A). The size difference was due to an internal deletion of 97 nt in the smaller circle. This accounted for the broad 0.7 kb band of etsRNA seen by northern analysis with ETS-2 probes (Figure 2). Circles corresponding to the 0.9 kb band were visible when the RT reaction was done at 52°C using superscript III reverse transcriptase. To ensure specificity, three different forward primers (F1, F2, and F3) were used for the PCR reaction and the specific amplicons were detected by Southern hybridization (Figure 4 a, b). Sequence analysis showed that these amplicons arose from two circles of sizes 912 nt and 849 nt. They shared the same 5'- and 3'-junctions (+102G, +1013A respectively), and the shorter circle had an internal deletion of 63 nt. In fact the 5'-junction was common to all the four circular RNA species (Figure 4a). The junction nucleotides for all four circles were 5'-G.A-3'. The origin of the internal deletions of 97 nt and 63 nt is not clear at present, and further experiments are required to understand their relevance. No circles corresponding to the 1.1 kb etsRNA (Figure 2) were obtained, which appears to be a linear RNA species as indicated below.

Further evidence of the circular nature of etsRNA was obtained by digesting the RNA with exonuclease T (which requires a free 3' terminus and removes nucleotides in the 3' to 5' direction). As





**Figure 3 | etsRNA is not polyadenylated, is transcribed in the direction of rRNA and is nuclear localized.** (a) Poly A<sup>+</sup> and Poly A<sup>-</sup> RNA fractions from 24 h serum starved cells were separated on denaturing agarose gel, blotted and hybridized with ETS-2 probe. The blot was reprobed with actin as a control. (b) 10 ug of total RNA from the same cells was hybridized with sense (S) and antisense (AS) ETS-2 probe. 18S was used as loading control and quality of probes was confirmed by genomic DNA dot blot. (c) Nuclear and cytoplasmic RNA was fractionated from normal and starved cells and northern blots were hybridized with ETS-2 probe. Actin was used as control.

shown in Fig. 4c, the multiple bands constituting etsRNA (except the 1.1 kb band) were highly resistant to exo T while all of the other bands hybridizing with ETS-2 probe were rapidly digested. The etsRNA was also resistant to nicking conditions (90°C, with NaHCO<sub>3</sub><sup>24</sup>) in which linear RNA species (including the 1.1 kb band) rapidly disappeared (Figure 4d). The nucleotide sequence of ETS-2 along with the junction sequences of the various circles is given in supplementary Fig. S2.

**etsRNA can circularize *in vitro*.** Many of the circular RNAs derived from introns have the property of spontaneous self-circularization. We determined whether etsRNA also had this ability. Linear transcripts corresponding to the 766 and 912 nt circles were obtained by *in vitro* transcription, and cRT-PCR was done to detect any spontaneous circularization. Another set of linear transcripts was also obtained which lacked the two 74-bp repeats present in the ETS-2. All four transcripts gave amplicons expected from circular RNA molecules (Figure 5), showing that these molecules could self-circularize and that the 74-bp repeats were not required for circle formation. The junction points of the circles, determined by sequence analysis, showed that the 5'-end of the linear transcript (+102G) was present at the junction in all four circles. However the 3'-junction was different in each, and the 3'-end of the linear transcript was not used for circle formation *in vitro* (Figure 5 a). Interestingly the 3'-junction in each case was, nevertheless, an adenine nucleotide. To further demonstrate the presence of circular RNA molecules, we radioactively labelled the 912 nt etsRNA *in vitro* and performed electrophoresis through different acrylamide concentrations (4% and 6%). The results presented in Figure 5c show that the bulk of RNA migrates at the expected linear size of 912 nt, while there is a slower migrating band which corresponds to the circular molecules. The migration of this band is further retarded in 6% acrylamide. The ratio of circular to linear molecules was estimated by densitometry to be ~ 9%. As a

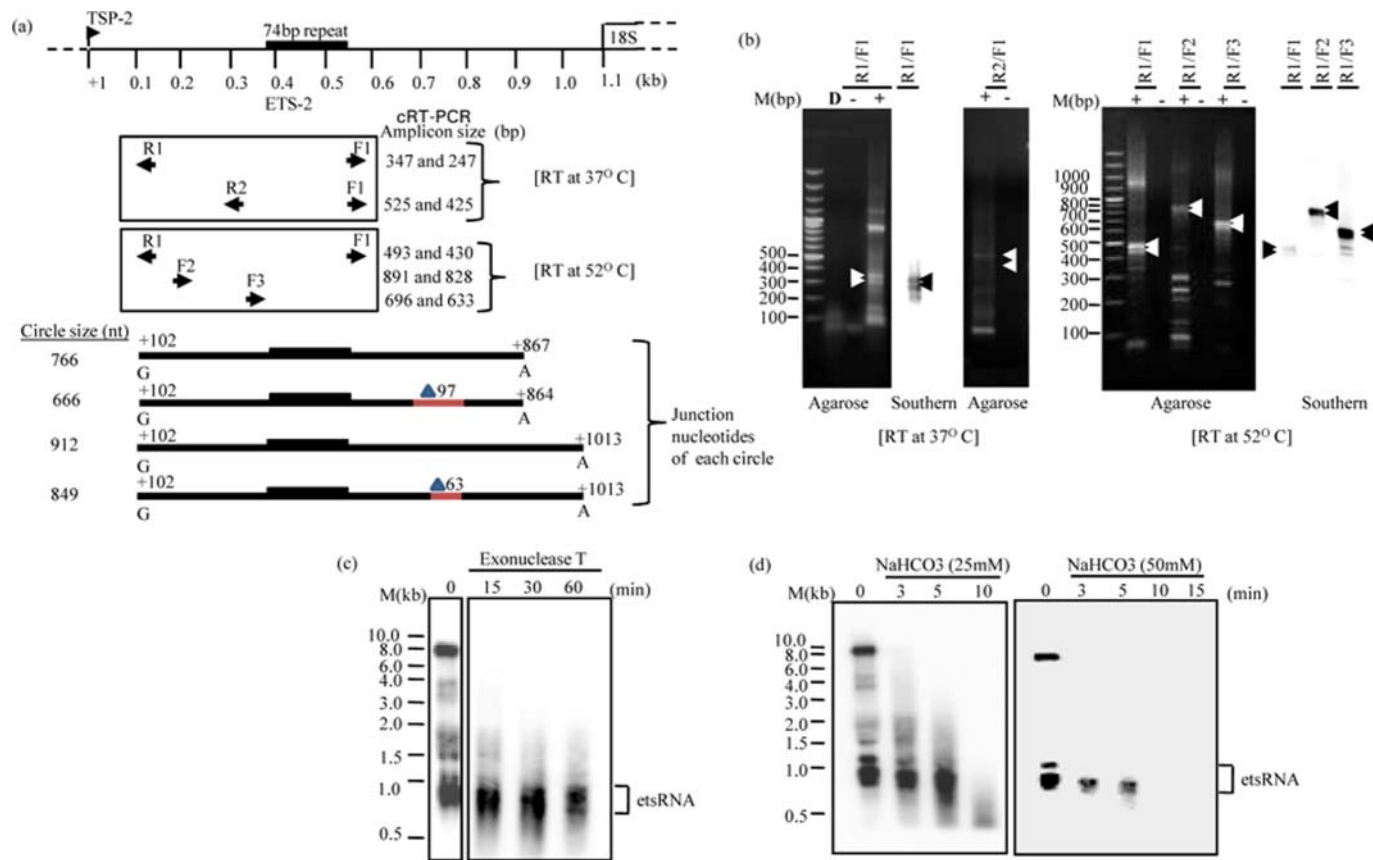
control we radioactively labelled the 937 nt etsRNA (which does not circularize, as shown in Figure 6), and this gave only a single band corresponding to the linear molecules.

**Mapping the possible linear precursors of etsRNA circles.** The full-length linear ETS-2 RNA is 1.1 kb in size, while the largest circle found by us is 912 nt. To determine the processing pathway by which the circles may be generated we performed primer extension with primer R2 from ETS-2 to map the pre rRNA processing site at 5'-end. Two major extension products were obtained- a 312 nt band corresponding to the 5'-end of pre rRNA-2, and two close bands of sizes 221 nt and 219 nt resulting from pre rRNA processing at +92 and +94 positions, respectively (Figure 6). By convention this site may be designated as A'. Primer extension with primer R3 from the 3'-end of ETS-2 gave two close bands of sizes 70 nt and 69 nt, corresponding to the 3'-processing site of pre rRNA, located at +1029 and +1030 positions respectively. This site is designated A<sub>0</sub>. We checked whether the 1.1 kb full-length ETS-2 RNA (which is visible under stress conditions, Figure 2), or its processing intermediates could self-circularize *in vitro*. To test the processing intermediates we used a 1007 nt molecule (A' to 3') and a 937 nt molecule (A' to A<sub>0</sub>) (Figure 6b). All three molecules (obtained by *in vitro* transcription) failed to circularize *in vitro*. However, a 920 nt linear molecule (A' to +1013A) ending with the junction sequence of the 912 nt circle at the 3'-end could self-circularize *in vitro* using the 5'-junction (+102G) to give the 912 nt circle. Linear molecules with the conserved 5'-end (+102G) seen in the molecules which could self-circularize *in vitro* could not be detected by primer extension, showing that this linear precursor either does not exist in the cell or is rapidly circularized. Since the 937 nt molecule ending at A<sub>0</sub> could not circularize *in vitro* while the 920 nt molecule ending at +1013A could, we checked whether further processing following the A<sub>0</sub> cleavage may generate a linear precursor of the 912 nt circle *in vivo*. RNase protection assay indeed showed the presence of this linear molecule (Figure 6c).

Our data shows that the 5'-ETS-2 in *E. histolytica* is processed at sites A' and A<sub>0</sub> which are 937 nt apart. The A'-A<sub>0</sub> region contains the junction sequences of all the circles found *in vivo* and *in vitro*. The 5'-junction sequence (+102G) is common to all the circles and is located 8 (or 10) nt downstream of processing site A'. The 3' junction of the longest circle (912 nt) is located 17 (or 16) nt upstream of site A<sub>0</sub> and the linear precursor of this circle is generated by processing after the A<sub>0</sub> cleavage. Other 3' junctions, further upstream, are also used which result in smaller circles. The propensity to circularize is an inherent property of the A'-A<sub>0</sub> sub fragments. However, the quantity of circles formed would depend on the concentration of the linear precursors, (which would be targeted by the degradation machinery); and/or the generation of suitable ends in the linear precursors to enable circularization. Once formed, the circles would be much more stable than the linear precursors. We speculate that the accumulated circular etsRNAs may serve regulatory functions, e.g. to inhibit the processing of pre rRNAs.

## Discussion

Genome-wide transcription analysis of a variety of genomes has established that organisms engage in widespread transcription that does not arise from protein-coding genes<sup>25</sup>. Transcripts that do not seem to encode protein are referred to as non coding RNA or transcripts of unknown function. A number of non coding RNAs are known to play important regulatory functions, e.g. in epigenetic control of gene expression and chromatin dynamics<sup>26</sup>. Non coding RNAs with regulatory functions have also been reported from the rDNA locus. Transcripts arising from the rDNA intergenic spacer (IGS) in mouse are processed into 150–300 nt RNAs which interact with TIP5, the large subunit of the chromatin remodelling complex NoRC. This association is essential for the epigenetic silencing of the

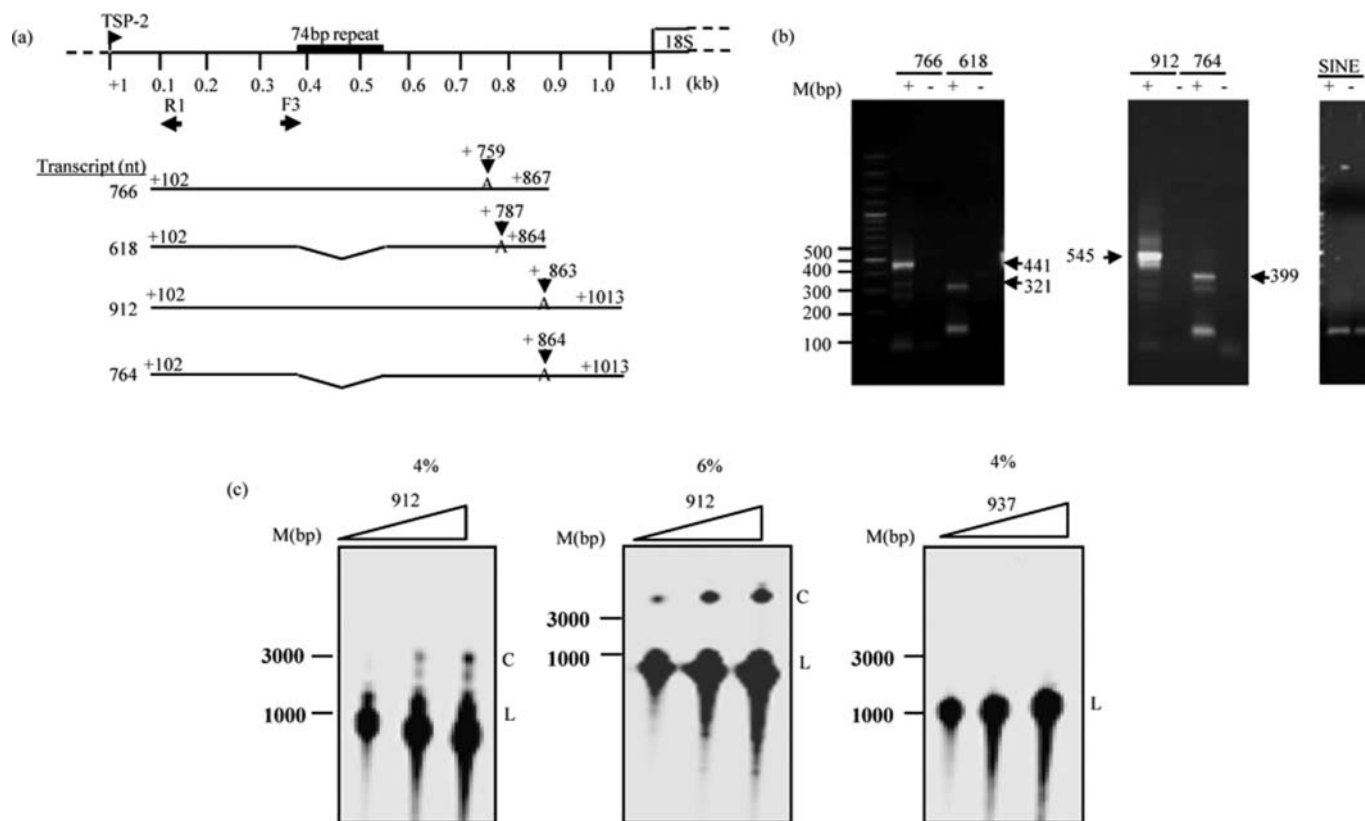


**Figure 4** | etsRNA accumulates as multiple circular transcripts. (a) ETS-2 showing position and orientation of primer pairs (R1/F1, R1/F2, R1/F3 and R2/F1) used for circular RT-PCR. Sizes of amplicons obtained with each primer pair is shown to the right. Linear view of the junction of circular RNAs deduced from sequence analysis of amplicons is shown below. The positions indicated are with respect to TSP-2.  $\Delta$  63 (+0.698 to +0.760) and  $\Delta$  97 (+0.664 to +0.760) are internal deletions. 74 bp repeats are present in all circles. (b) 5  $\mu$ g of total RNA from 24 h serum starved cells was used for reverse transcription with primer R1 and R2 followed by PCR with forward primers (F1, F2, F3). Arrowheads show the specific amplicons as confirmed by Southern using ETS-2 DNA probe. The amplicon obtained with primer R2 was confirmed by sequencing. + and - lanes are with and without reverse transcriptase respectively. Lane D is PCR with genomic DNA using R1 and F1 primers. To further demonstrate circular nature of etsRNA, 10  $\mu$ g of total RNA from 24 h serum starved cells was treated with (c) Exonuclease T for the indicated times; or (d) NaHCO<sub>3</sub> (25 mM and 50 mM) for indicated times. RNA was fractionated on 1% denaturing agarose gel and hybridized with ETS-2 DNA probe.

rDNA locus<sup>27</sup>. In another study, an IGS transcript originating 1 kb upstream of the rRNA tsp was reported in lung cancer cells. The levels of this transcript correlated negatively with those of the 45 S pre rRNA, indicating a regulatory role for this RNA<sup>28</sup>. However there is no report yet of the processed ETS RNA playing a regulatory role in rRNA synthesis or processing. Here we report a new class of RNA molecules called the etsRNAs, which are formed after processing of the 5'-ETS-2 in *E. histolytica*. They are stabilized in response to physiological stress, with simultaneous accumulation of unprocessed pre rRNA. Stabilization is a result of the property of etsRNAs to spontaneously self-circularize. Their circular nature would make them more resistant to exonucleases. Since the 5'-ETS region of pre rRNA is the primary site for assembly of the pre rRNA processing machinery, we speculate that accumulated ETS RNA may serve as a scaffold to trap or store the processing factors in an inactive state under stress conditions when synthesis of new ribosomes is stopped. *E. histolytica* living in the dynamic environment of the human gut may encounter frequent episodes of nutrient depletion. Accumulation of unprocessed pre rRNA may permit the starving trophozoite to rapidly assemble new ribosomes when normal nutrient conditions are restored. Accumulation of pre rRNA has also been reported in mitotic cells<sup>29</sup>. This pre rRNA is partially processed and is found in nucleolar derived foci in association with processing factors. It is thought that this RNA may re enter the new nuclei and assist in the rapid processing of pre rRNAs for generation of mature rRNAs.

The 5'-ETS, excised during pre rRNA processing, is subjected to degradation by the nuclear exosome. It has been shown in a variety of organisms, namely yeast, mouse and *Arabidopsis* that the 5'-ETS accumulates in exosome-depleted cells. Interestingly, only a specific fragment of the 5'-ETS accumulated in each case. In yeast, this fragment was the 610 nt 5'-A0 fragment<sup>16,30</sup>; in mouse it was the 1 kb A'-A0 fragment<sup>15</sup> and in *Arabidopsis* it was the 450 nt P-P' fragment<sup>17</sup>. It is significant to note that both in yeast and mouse where the U3 snoRNA interaction sites in the 5'-ETS have been mapped, the accumulated 5'-ETS fragment contained the regions which base pair with U3 snoRNA<sup>15,31</sup>. This interaction is crucial for pre rRNA processing. We speculate that specific sub-fragments of the 5'-ETS, which accumulate under adverse conditions, may be designed to play a regulatory role in pre rRNA transcription or processing.

The observation that etsRNAs exist as circles in *E. histolytica* is novel since no such circular RNAs have been reported from rDNA spacers in any organism. The most well-known circular RNAs are the self splicing group I and group II introns. In group I introns the released linear intron RNA can further give rise to full-length circles, or smaller circles missing a part of the intron<sup>32</sup>. Group II intron splicing liberates the intron as a lariat, or sometimes as complete circles<sup>33</sup>. Another class of circular RNAs are derived from exons. Most of these are associated with exon skipping whereby the skipped exons are retained as circular molecules as in the CytP450 gene<sup>34,35</sup>.



**Figure 5 | etsRNAs can spontaneously self circularize.** (a) Transcripts of indicated size (with or without 74 bp repeats) were obtained by *in vitro* transcription using T7 RNA polymerase. Primer pair (R1 and F3) was used for cRT-PCR. (b) 5  $\mu$ g of *in vitro* transcribed RNA was used for cRT-PCR reactions. Expected amplicons in each case are shown by arrows. These were cloned and sequenced. The positions of new 3'-junctions of each circle are shown by vertical arrows in panel (a). *In vitro* transcribed 0.5 kb SINE 1 RNA of *E. histolytica* was used as negative control for cRT-PCR, using out-facing SINE primers. (c) Further demonstration of circular nature of 912 nt etsRNA. The 912 nt (and 937 nt) RNAs were *in vitro* transcribed in presence of 20  $\mu$ M  $\alpha$ P32-UTP, purified and electrophoresed in 4% and 6% denaturing PAGE. Slow migrating band in 912 panel corresponds to the circular form (C) which gets further retarded in 6% gel. (L) denotes the linear form. No circular form was seen with the 937 nt RNA, as expected (see Fig. 6).

the long non coding RNA, ANRIL<sup>36</sup> and the Sry transcript in mouse testis<sup>37</sup>. Although there is no clear biological function associated with exonic circular RNAs (circular ANRIL exons are correlated with atherosclerosis risk), it appears that they may not be mere by-products of splicing since they are transported to the cytoplasm and accumulate there. The etsRNA circles described here do not typically fall into either the intron or the exon class of circles. One may envision that RNAs with the propensity to circularize occur in many different contexts in the cell. In their linear state these RNAs have transient existence, but as circular molecules they are stabilized and may be recruited for a variety of functions, e.g. intron mobility in the case of full-length circular introns<sup>38</sup>, or regulatory functions yet to be assigned in other cases. The primitive parasitic protist *E. histolytica* has given the first evidence of circular RNAs associated with the 5'-ETS of rRNA genes.

## Methods

**Cell culture and growth conditions.** Trophozoites of *E. histolytica* strain HM-1: IMSS (clone 6) were axenically maintained in TYI-S-33 medium supplemented with 15% adult bovine serum (PAA laboratories, Austria), at 35.5°C<sup>39</sup>. For serum starvation, mid log phase trophozoites were collected by centrifugation and resuspended in medium containing 0.5% serum. CHX (sigma) was added at 100  $\mu$ g/ml to the mid log phase trophozoites and grown for 15 minutes, after which cells were collected by centrifugation and resuspended in medium lacking CHX.

**Luciferase reporter construct and stable transfection.** Deletion constructs (as indicated in Figure 1a) were cloned upstream of LUC gene at Sac II/ Kpn I site in promoter-less vector pEh-NEO-LUC Dra I in which the S1-DraI-S2 region of Ehr2 was cloned downstream of LUC. Constructs were transfected by electroporation and maintained in presence of G418 at 10  $\mu$ g/ml<sup>40</sup>.

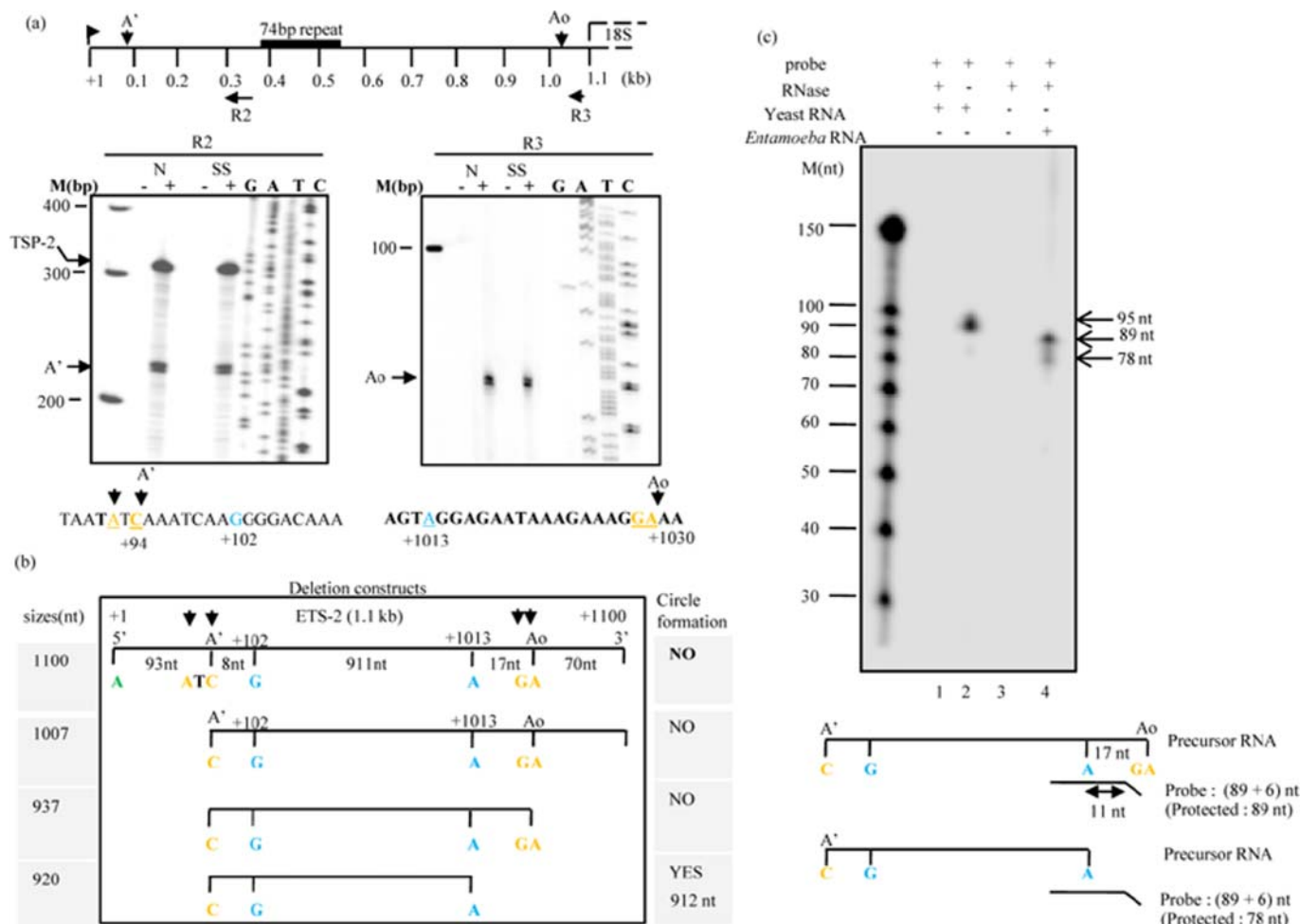
**Primer extension.** DNase I (Roche)-treated total RNA (10  $\mu$ g) was used for primer extension with end labeled oligonucleotides (listed in supplementary Table 1). Reverse transcription was carried out using superscript III (Invitrogen) at 52°C according to manufacturer's instructions.

**RNA isolation and northern hybridization.** Total RNA from  $\sim 5 \times 10^6$  cells was isolated using trizol (Invitrogen) as per manufacturer's instructions. Poly A<sup>+</sup> and poly A<sup>-</sup> RNA fractions were obtained using polyATtract mRNA isolation system (Promega) as per protocol prescribed by manufacturer. Nuclear and cytoplasmic RNA were fractionated following protocol from Sugdens lab. For northern analysis 10  $\mu$ g of RNA was fractionated on 1% formaldehyde agarose gel in buffer {(0.1 M MOPS (pH 7.0), 40 mM sodium acetate, 5 mM EDTA (pH 8.0))} followed by transfer on Genescreen plusR membrane (Perkin Elmer). Probe was prepared by random priming method using NEBlot kit (NEB). Sense and antisense single stranded DNA (ssDNA) probes for ets RNA were synthesized by linear PCR. 150 ng of ets DNA template was used in linear PCR either with forward or reverse primer in presence of 30  $\mu$ Ci of  $\alpha$ P32-dATP for 25 cycles. The resulting ssDNA probe was purified directly from PCR reaction mixture using nucleotide removal kit (Quiagen) as per manufacturer instructions. Hybridization was carried out at 65°C for 16 hrs followed by washing as per manufacturer instruction.

***In vitro* transcription.** Linear ETS-2 DNA templates with desired end points were PCR amplified from genomic DNA of *E. histolytica*. T7 promoter sequence was incorporated in the forward primer. 500 ng purified template was *in vitro* transcribed and purified using RiboMAX<sup>TM</sup> (Promega) according to instruction prescribed by manufacturer.

**Analysis of cyclization sites for etsRNAs.** Cyclization sites were mapped using circularized RNA (C)-RT-PCR method without RNA Ligase<sup>41</sup>. 5  $\mu$ g of total RNA or *in vitro* transcribed RNA was subjected to reverse transcription with R1 using MMLV (USB) at 37°C or superscript III (Invitrogen) at 52°C, followed by PCR with forward primers F1-F3 (position and sequence of primers used are listed in supplementary Table 1). Amplicons arising from ETS-2 were confirmed by Southern using standard method<sup>42</sup> and cloned into pGEMT-Easy vector (Promega) according to





**Figure 6 | Testing the self-circularizing ability of 5' ETS-2 processed intermediates.** (a) Primer extension analysis to map the processing sites of ETS-2 RNA. Primer extension was performed with total RNA isolated under normal and serum starved conditions with primers R2 and R3. The products obtained were analyzed on 6% urea-PAGE with sequencing reaction. The extended products obtained with primer R2 were 312 nt (corresponding to TSP-2), and 219 and 221 nt (corresponding to A'). Primer R3 extended products of 69 and 70 nt, which corresponded to Ao. (b) cRT-PCR analysis of RNAs spanning the indicated regions of ETS-2. 5  $\mu$ g of *in vitro* transcribed and purified RNA was subjected to cRT-PCR with primers R1 and F3 (Figure 4 a). +102 G and +1013 A are the junctions of the 912 nt circle *in vivo*. (c) RNase protection assay to map the 3' end of linear 912 nt etsRNA *in vivo*. 30  $\mu$ g of total RNA from serum starved cells were hybridized with antisense probe (+936 to +1024) spanning the 3' end of 912 nt RNA with an 11 nucleotide overhang. Lanes 1&2 are controls with 30  $\mu$ g of yeast RNA. Size of the full-length probe is 95 nt (indicated by arrow) due to an extra 6 nucleotides added by the T7 promoter during transcription. Lane 3 is probe incubated with RNase alone. Lane 4 shows the protected fragments of 78 and 89 nucleotides (indicated by arrows) as a result of protection by the 912 nt linear precursor and nascent pre rRNA respectively.

manufacturer's instructions. Amplicons were sequenced commercially (TCGA, India).

**RNase protection assay.** The 89 bp fragment corresponding to +936 to +1024 (see supplementary fig. S2) was PCR amplified with forward primer 5'-TAATTAGATGACAGAACAG-3' and T7 reverse primer 5'-TAATACGACTCACTATAGGGAGACTTTATTCTCTACTTTTGAC-3'. The purified fragment (Gel Extraction Kit, Qiagen) was transcribed in antisense orientation in the presence of 50  $\mu$ Ci of  $\alpha$ <sup>32</sup>-UTP as mentioned above. RNase protection assays were performed with 30  $\mu$ g of total RNA from 24 h serum starved cells hybridized with  $\sim 4 \times 10^4$  cpm of radiolabelled probe using the RPA III ribonuclease protection assay kit (Ambion) as per protocol prescribed by manufacturer. The protected fragments were precipitated and analyzed on a denaturing 8% urea-polyacrylamide gel by autoradiography.

- Grummt, I. Life on a planet of its own: regulation of RNA polymerase I transcription in the nucleolus. *Genes Dev.* **7**, 1691–702 (2003).
- Moss, T. At the crossroads of growth control; making ribosomal RNA. *Curr. Opin. Genet. Dev.* **14**, 210–7 (2004).
- Jacob, S. T. Regulation of ribosomal gene transcription. *Biochem J.* **306**, 617–26 (1995).
- Jacob, S. T. Transcription of eukaryotic ribosomal RNA gene. *Mol Cell Biochem.* **70**, 11–20 (1986).

- Sollner-Webb, B. & Tower, J. Transcription of cloned eukaryotic ribosomal RNA genes. *Annu. Rev. Biochem.* **55**, 801–30 (1986).
- Henras, A. K. *et al.* The post-transcriptional steps of eukaryotic ribosome biogenesis. *Cell Mol Life Sci.* **65**, 2334–59 (2008).
- Eichler, D. C. & Craig, N. Processing of eukaryotic ribosomal RNA. *Prog. Nucleic Acid Res. Mol. Biol.* **49**, 197–239 (1994).
- Borovjagin, A. V. & Gerbi, S. A. The spacing between functional Cis-elements of U3 snoRNA is critical for rRNA processing. *J. Mol. Biol.* **300**, 57–74 (2000).
- Krogan, N. J. *et al.* High-definition macromolecular composition of yeast RNA-processing complexes. *Mol. Cell.* **13**, 225–39 (2004).
- Pérez-Fernández, J., Román, A., De Las Rivas, J., Bustelo, X. R. & Dosil, M. The 90S preribosome is a multimodular structure that is assembled through a hierarchical mechanism. *Mol. Cell Biol.* **27**, 5414–29 (2007).
- Gallagher, J. E. *et al.* RNA polymerase I transcription and pre-rRNA processing are linked by specific SSU processome components. *Genes Dev.* **18**, 2506–17 (2004).
- Wery, M., Ruidant, S., Schillewaert, S., Leporé, N. & Lafontaine, D. L. The nuclear poly (A) polymerase and Exosome cofactor Trf5 is recruited cotranscriptionally to nucleolar surveillance. *RNA.* **15**, 406–19 (2009).
- Houseley, J., LaCava, J. & Tollervey, D. RNA-quality control by the exosome. *Nat. Rev. Mol. Cell Biol.* **7**, 529–39 (2006).
- Houseley, J. & Tollervey, D. The many pathways of RNA degradation. *Cell.* **136**, 763–76 (2009).



15. Kent, T., Lapik, Y. R. & Pestov, D. G. The 5' external transcribed spacer in mouse ribosomal RNA contains two cleavage sites. *RNA*. **15**, 14–20 (2009).
16. de la Cruz, J., Kressler, D., Tollervey, D. & Linder, P. Dob1p (Mtr4p) is a putative ATP-dependent RNA helicase required for the 3' end formation of 5.8S rRNA in *Saccharomyces cerevisiae*. *EMBO J.* **17**, 1128–40, (1998).
17. Lange, H. *et al.* Degradation of a polyadenylated rRNA maturation by-product involves one of the three RRP6-like proteins in *Arabidopsis thaliana*. *Mol. Cell Biol.* **28**, 3038–44 (2008).
18. Sehgal, D. *et al.* Nucleotide sequence organisation and analysis of the nuclear ribosomal DNA circle of the protozoan parasite *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* **67**, 205–14 (1994).
19. Bhattacharya, S., Som, I. & Bhattacharya, A. The ribosomal DNA plasmids of entamoeba. *Parasitol Today*. **14**, 181–5 (1998).
20. Ghosh, S., Zaki, M., Clark, C. G. & Bhattacharya, S. Recombinational loss of a ribosomal DNA unit from the circular episome of *Entamoeba histolytica* HM-1:IMSS. *Mol. Biochem. Parasitol.* **116**, 105–8 (2001).
21. Michel, B., Lizardi, P. M., Alagón, A. & Zurita, M. Identification and analysis of the start site of ribosomal RNA transcription of *Entamoeba histolytica*. *Mol. Biochem. Parasitol.* **73**, 19–30 (1995).
22. Panigrahi, S. K. *et al.* Promoter analysis of palindromic transcription units in the ribosomal DNA circle of *Entamoeba histolytica*. *Eukaryot. Cell.* **8**, 69–76 (2009).
23. Shrimal, S., Bhattacharya, S & Bhattacharya, A. Serum-dependent selective expression of EhTMKB1-9, a member of *Entamoeba histolytica* B1 Family of Transmembrane Kinase. *PLoS Pathog.* **6**, e1000929 (2010).
24. Bohjanen, P. R., Colvin, R. A., Puttaraju, M., Been, M. D. & Garcia-Blanco, M. A. A small circular TAR RNA decoy specifically inhibits Tat-activated HIV-1 transcription. *Nucleic Acids Res.* **24**, 3733–8 (1996).
25. apranov, P., Willingham, A. T. & Gingeras, T. R. Genome-wide transcription and the implications for genomic organization. *Nat. Rev. Genet.* **8**, 413–23 (2007).
26. Bernstein, E. & Allis, C. D. RNA meets chromatin. *Genes Dev.* **19**, 1635–55 (2005).
27. Mayer, C., Schmitz, K. M., Li, J., Grummt I & Santoro, R. Intergenic transcripts regulate the epigenetic state of rRNA genes. *Mol. Cell.* **22**, 351–61 (2006).
28. Shiao, Y. H. *et al.* An intergenic non-coding rRNA correlated with expression of the rRNA and frequency of an rRNA single nucleotide polymorphism in lung cancer cells. *PLoS One.* **4**, e7505 (2009).
29. Dundr, M. & Olson, M. O. Partially Processed pre-rRNA Is Preserved in Association with Processing Components in Nucleolus-derived Foci during Mitosis. *Mol. Biol. Cell.* **9**, 2407–22 (1998).
30. Mitchell, P. *et al.* Rrp47p is an exosome-associated protein required for the 3' processing of stable RNAs. *Mol. Cell Biol.* **23**, 6982–92 (2003).
31. Beltrame, M. & Tollervey, D. Identification and functional analysis of two U3 binding sites on yeast pre-ribosomal RNA. *EMBO J.* **11**, 1531–42 (1992).
32. Vicens, Q. & Cech, T. R. A natural ribozyme with 3', 5' RNA ligase activity. *Nat. Chem. Biol.* **5**, 97–9 (2009).
33. Murray, H. L. *et al.* Excision of group II introns as circles. *Mol. Cell.* **8**, 201–11 (2001).
34. Zaphiropoulos, P. G. Exon skipping and circular RNA formation in transcripts of the human cytochrome P-450C18 gene in epidermis and rat androgen binding protein gene in testis. *Mol. Cell. Biol.* **17**, 2985–93 (1997).
35. Pasmán, Z., Been, M. D. & Garcia-Blanco, M. A. Exon circularization in mammalian nuclear extracts. *RNA*. **2**, 603–10 (1996).
36. Burd, C. E. *et al.* Expression of linear and novel circular forms of an INK4/ARF-associated on-coding RNA correlates with atherosclerosis risk. *PLoS Genet.* **6**, e1001233 (2010).
37. Capel, B. *et al.* Circular transcripts of the testis-determining gene Sry in adult mouse testis. *Cell.* **73**, 1019–30 (1993).
38. Nielsen, H. *et al.* The ability to form full-length intron RNA circles is a general property of nuclear group I introns. *RNA*. **9**, 1464–75 (2003).
39. Diamond, L. S., Harlow, D. R. & Cunnick, C. C. A new medium for the axenic cultivation of *Entamoeba histolytica* and other entamoebas. *Trans.R. Soc. Trop. Med. Hyg.* **72**, 431–432 (1978).
40. Hamann, L., Nickel, R. & Tannich, E. Transfection and continuous expression of heterologous genes in the protozoan parasite *Entamoeba histolytica*. *Proc. Natl. Acad. Sci. U.S.A.* **92**, 8975–9 (1995).
41. Kuhn, J. & Binder, S. RT-PCR analysis of 5' to 3'-end-ligated mRNAs identifies the extremities of cox2 transcripts in pea mitochondria. *Nucleic Acids Res.* **30**, 439–46 (2002).
42. Sambrook, J. & Russell, D. W. *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 3rd ed. NY.

## Acknowledgement

This work was supported by a grant from the Department of Biotechnology, India to SB and fellowship from Council of Scientific and Industrial Research, India to AKG.

## Author contribution

SB proposed, designed and interpreted the experiments, wrote and drafted the final version of the manuscript; AB analyzed the data and contributed in writing all sections of paper. AKG designed, conducted all the experiments and helped in writing the paper. SKP conducted experiments related to mapping of the promoter. All authors have read and approved the final manuscript.

## Additional information

**Supplementary information** accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**License:** This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivative Works 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/>

**How to cite this article:** Gupta, A.K., Panigrahi, S.K., Bhattacharya, A. & Bhattacharya, S. Self-circularizing 5'-ETS RNAs accumulate along with unprocessed pre ribosomal RNAs in growth-stressed *Entamoeba histolytica*. *Sci. Rep.* **2**, 303; DOI:10.1038/srep00303 (2012).