Dear Dr. Lazo,

Thank you for the prompt decision on our MS. Please consider our revised manuscript for publication in PEERJ. We have carefully considered the comments of the two reviewers and your suggestions.

Below we have listed (in the order of reviewer comments) how we have addressed the points.

All of the editor’s suggestions below have been addressed in the revision and marked in the revised copy.

Page 4:

80: /De novo/de novo/.

78: CLC (website citation).

Page 5:

84: Geneious (website citation).

85: /454 Life Sciences/Roche 454 Life Sciences/(website citation).

87: (BLAST citation) needed because of extensive use further on.

101: Have Genebank accession numbers been obtained yet?

There is a release when published option for Genbank submissions.

102: /blastp/BLAST blastp/.

103: /cut-off of E<0.05/cutoff E\_value of E<0.05/.

104: /nr peptide sequence database/non-redundant 'nr' peptide sequence database/.

107: /tBLASTx/BLAST tblastx/

Page 6:

125-127: (Are there identifiers for these encoded PEPTIDES or genes?)

Page 7:

138: /The tRNA was folded using VARNA secondary structure visualization

program/ The folding of tRNA was predicted using the VARNA secondary structure visualization

program/.

Page 8:

156: /tblastn/BLAST tblastn search/; be consistent in usage of BLAST notations

(tblastn program option in this case).

174: /blastp/BLAST blastp/.

176: /greatest expect value (E= 2e-137)

/highest expect value (E= 2e-137)

obtained/.

Page 9:

186: /EST libraries but/EST libraries, but/

Page 10:

216: /another basidiomycete,/another genera of basidomycete,/

Page 12:

265: /Do these endogenous non-retroviral genes function as

proteins or are they merely transcriptional noise?

/It is not clear whether these endogenous non-retroviral genes produce any functional protein or are just expressed at the transcriptome level./.

Figure 1:

/BLASTp/BLAST blastp/

Page 14:

310: (There are no PUBLISHER or PAGES listed).

Page 16:

372: /Bmc Bioinformatics/BMC Bioinformatics/(There are no PAGES listed).

Page 18:

409: /Bmc Biology/BMC Biology/.

Comments from Reviewer 1:

No revisions requested.

Comments from Reviewer 2:

**Reviewer 2**

**Basic reporting**

“The manuscript has lot of grammatical errors all over it. This makes proofreading essential before the submission. There are more than 100 first person plural narratives in the text which need to be changed to third person passives.”

We disagree that all first person sentences in scientific writing “need to be changed”. Use of the first person often clarifies writing. There is presently no consensus on the usage of first person in scientific writing. In the revision we have retained the first person sentences in the introduction and in the results/discussion. However, we changed most of the first person sentences in the revised materials and methods to the third person passive.

 The abstract is not a good summary of the work and it doesn't show why the authors are claiming that the RNA virus is co-evolved with the host.

We assume that the reviewer meant that the abstract is not a good summary because it doesn’t show RNA virus-host co-evolution. The evidence for co-evolution is the existence of a virus in the CTG yeasts and the viral avoidance of the derived genetic code of the host. The host genus has furthered the co-evolutionary story by co-opting a viral protein. We have reworded the abstract to emphasize the evidence for reciprocal adaptations in the host and the virus.

The material and methods needs some changes in the subheadings (like PCR, RT-PCR and DNA sequencing subheading that explain more the Next generation sequencing and RNA sequencing rather than DNA sequencing). Therefore the materials and methods section should be restructured again.

We have changed the heading to specifically mention each type of sequencing.

There are some occasions that the reference or the name of the company is missed for a device, software or a kit. The result is vague and it is not clear why the authors used different molecular nalysis to confirm their hypothesis. (there used different technologies and molecular analysis in this research but it is not clear for what reason this has been done.). For example I couldn't understand using both RNA seq and also RT-PCR. Was the second one a confirmation or what?

We did mention that the Sanger sequence was a confirmation of the RNA seq assembly (Line 82 of the original; line 85 in the revision). In the revision we also now mention why we carried out RNA-seq (72-74).

Or It is not clear why both 454 FLX sequencing was done for the whole yeast genome when we needed just a confirmation of existance of some Totivirus genes in the nuclear genome. I believe there are some unnecessary works that is done without any particular reason or the result is not fully included in this manuscript for other reasons.

Starting on line 89 of the revision we state why we carried out 454 FLX sequencing:

“Because endogenous RNA viruses of fungi can be fragmented and differ greatly in their nucleotide sequences from known viruses (Taylor et al. 2009), PCR probes alone are often an ineffective tool for paleoviral discovery. We therefore carried out 454 Life Sciences (http://www.454.com) sequencing with GS FLX Titanium series reagents of a DNA library from *Scheffersomyces coipomoensis*. This form of sequencing also permitted multigene bioinformatics analysis of the host protein coding genes.”

 There is not enough discussion about results and the references need a few corrections like BMC Genomics which is written in small letters as Bmc genomics..

Capitalization in the references has been corrected.

**Experimental design**

In the materials and methods (line 91) and result (line 168) the authors used RT-PCR to find out the nuclear genomic copy of the virus. I believe they have mistaken ordinary PCR with RT-PCR which is for RNA amplification not genomic DNA. I am not sure why they have not used any RNAase prior to PCR to be sure about exogenous virus RNA degradation.

We carried out both RTPCR and PCR. In the revision (line 99) we expand our materials and methods to clarify:

“To establish that the virus was coded by exogenous RNA and not by the DNA of the host, we compared RT-PCR and PCR products. For RNA templates, DNase-treated extracts were exposed to RT-PCR using the Qiagen one step RT-PCR kit. For DNA templates, nucleic acid extracts were exposed to PCR by excluding reverse transcriptase from the RT-PCR protocol. We amplified a fragment of the single copy xylose reductase gene as a positive control for the PCR of DNA.”

The reviewer seems to suggest that the RNA might contaminate the PCR reaction for DNA. However, we detected no DNA band using the viral primers for regular PCR. There can be no contribution from viral RNA when there is no detectable product and we excluded reverse transcriptase.

**Comments for the author**

These comments are re-iterations of the above comments and are addressed above.