CORRIGENDUM

Genes & Development 28: 2498–2517 (2014)

Corrigendum: Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes

Søren Lykke-Andersen, Yun Chen, Britt R. Ardal, Berit Lilje, Johannes Waage, Albin Sandelin, and Torben Heick Jensen

While revisiting next-generation sequencing (NGS) data analyses from the above-mentioned article for another project, we discovered four unintentional errors.

After correction of the errors, we reapplied the reported methods and found the originally stated conclusions to be unchanged. However, several figure panels and reported numbers were altered. Corrected Figures 1–7 are shown below, and corrected Supplemental Figures S2, S3, S5, S6, and S7 and Supplemental Tables S2, S3, and S4 are accessible on the journal site online.

The following errors were found and are corrected here (a flow chart illustrating the errors and how figures/tables were affected is also available in the Revised Supplemental Material):

(1) Due to a programming error during the processing of “5′ end-seq” data, reads spanning exon–exon junctions were erroneously omitted. This effectively led to a subsampling of the data, leaving out roughly a quarter of all “5′ end-seq” reads. In this Corrigendum, we now present analyses based on all of the reads. The increased amount of data influences the calculation of nonsense-mediated RNA decay (NMD) thresholds as well as most of the downstream computational analyses (see Box 1 in the flow chart). This affected Figure 1, E and G; Figure 2, C and D; Figure 3, A and B; Figure 4, A and B; Figure 5, A, B, and D; Figure 6B; Supplemental Figure S2, A, C, E, and H; Supplemental Figure S3, A, D, and G; Supplemental Figure S5, A–H; Supplemental Figure S7D; Supplemental Table S2; Supplemental Table S3; and Supplemental Table S4.

(2) The plots displaying NGS data mapping to specific genes or splice junctions were based on an alternative mapping procedure (RUM) (Grant et al. 2011) rather than the one reported in the Materials and Methods section (TopHat version 2). The two procedures gave very similar mapping results. In the corrected Figures 1, E–G, and 7 (below) as well as revised Supplemental Figures 2, A–E, and 3, A, B, D, E, and G (Revised Supplemental Material), we now present plots based on the reported mapping procedure (see Box 2 in the flow chart).

(3) In the top panel of Figure 6C and middle panel of Supplemental Figure S7D, we compared the number of transcript isoforms produced from snoRNA host genes with similarly expressed normal protein-coding genes. Although we reported the use of de novo transcripts based on our own RNA sequencing (RNA-seq) data from HEK293 cells, in fact, Gencode version 17 annotated transcripts were used. In the corrected Figure 6 (below) and Supplemental Figure S7 (Revised Supplemental Material), we now present analyses based on de novo assembled transcripts as stated in the corresponding legends (see Box 3 in the flow chart).

We sincerely apologize for any inconvenience caused by this negligence and stress once more that all original conclusions are left intact after the conducted modifications of the article.

The text changes reflecting the corrected data are noted below:

p. 2499, column 2, paragraph 3:

By combining global identification of nonsense RNAs and their corresponding decay intermediates, we identified primary NMD-responsive isoforms from up to 15% of all expressed genes.

(Continued on next page)

p. 2500, column 2, paragraph 3:

To investigate whether this observation was supported by our global data, we characterized NMD-specific endocleavage events in a reference set of annotated NMD substrates, consisting of 801 transcripts with a total of 1544 potential endocleavage sites (“NMD reference set” derived from annotation) (see Materials and Methods; Supplemental Table S2; Harrow et al. 2012), by mapping their positions relative to the annotated termination codons.

p. 2503, legend to Figure 2D:

(D) Box plots illustrating the distribution of RNA-seq-based expression levels of the transcripts corresponding to the 5′ end-seq signals plotted in C. (***) P -value ≤ 0.001 .

p. 2505, column 2, paragraph 1:

Finally, the NMD-sensitive transcripts identified independently by 5′ end-seq and RNA-seq (Supplemental Table S3) were compared, which yielded 3223 NMD-sensitive transcripts arising from 1563 assembled genes (Fig. 4B, right panel; Supplemental Table S4).

p. 2505, column 2, paragraph 1:

When applying a less stringent (“relaxed”) cutoff for the initial identification of peaks in the 5′ end-seq data (noncorrected P -value ≤ 0.005 based on negative binomial fitting), we could detect 9060 transcripts corresponding to 4245 genes in the overlap with the RNA-seq data (Supplemental Fig. S5D; Supplemental Table S4).

p. 2505, column 2, paragraph 1:

Using the present data, we confirmed that 25 of 53 SRSF protein-coding genes (Supplemental Table S5; total number of SRSF genes based on Long and Cáceres 2009) could be detected in our most stringent NMD substrate set based on the combined analysis of RNA-seq and 5′ end-seq.

p. 2505, column 2, paragraph 1:

Even more genes could be included by using the “relaxed” criteria (40 out of 53) or taking the maximum union of independent RNA-seq and 5′ end-seq approaches (51 out of 53) (Supplemental Fig. S5G).

p. 2506, column 1, paragraph 3:

We found that 59 (34%) and 104 (60%) of these genes were responsive to NMD based on the “stringent” and “relaxed” NMD substrate set, respectively (Supplemental Table S4).

p. 2506, column 1, paragraph 3:

Even more snoRNA hosts are potential targets of NMD, since the independent 5′ end-seq and RNA-seq procedures combined include up to 166 (96%) of these genes (Supplemental Fig. S5H; Supplemental Table S3).

p. 2507, legend to Figure 5A:

Three asterisks indicate that snoRNA host genes are significantly enriched compared with protein-coding genes (Fisher two-sided test, $P = 1.5 \times 10^{-13}$).

p. 2508, column 2, paragraph 1:

RNAs containing the latter splice junction constitute ~60% of the total amount of exon 2-containing RNA under control conditions (Fig. 7, histograms next to the schematics of the four isoforms show the percent of the total amount of exon 2-containing RNA within samples).

(Continued on next page)

Corrigendum

p. 2509, column 1, paragraph 1:

However, upon inhibition of NMD, this variant, although being somewhat NMD-sensitive, only constitutes ~25% of the total RNA.

p. 2509, column 1, paragraph 1:

Conversely, the isoforms encoding only *SNORD49A* comprise ~60% of the total RNA when NMD is inhibited as opposed to ~30% under normal conditions (Fig. 7, i,ii).

p. 2509, column 1, paragraph 1:

Similarly, the relative level of the transcript variants giving rise to both *SNORD49A* and *SNORD49B* increase from ~5% to ~10% upon inhibition of NMD (Fig. 7, iii).

p. 2509, column 2, paragraph 3:

This strategy revealed nonsense RNA isoforms produced from 6% and 15% of expressed gene loci, respectively.

p. 2513, column 1, paragraph 3:

This gave us a total set of 1321 transcripts corresponding to 916 genes.

p. 2513, column 1, paragraph 3:

This gave us a total set of 256 transcripts derived from 134 genes.

p. 2513, column 2, paragraph 4:

$SMG6/XRN1 > 1.56 \times CTRL$ and $UPF1/XRN1 > 1.56 \times CTRL$; $SMG6/XRN1 > 1.23 \times XRN1$, and $UPF1/XRN1 > 1.23 \times XRN1$.

p. 2514, column 1, paragraph 6:

NMD-specific endocleavage site: $XRN1 > 1.36 \times SMG6/XRN1$ and $XRN1 > 1.36 \times UPF1/XRN1$; NMD-specific decapping site: $SMG6/XRN1 > 1.93 \times XRN1$ and $UPF1/XRN1 > 1.93 \times XRN1$.

p. 2516, column 1

The reference below was also mistakenly omitted from the published article.

Grant GR, Farkas MH, Pizarro AD, Lahens NF, Schug J, Brunk BP, Stoeckert CJ, Hogenesch JB, Pierce EA. 2011. Comparative analysis of RNA-seq alignment algorithms and the RNA-seq unified mapper (RUM). *Bioinformatics* **27**: 2518–2528.

Supplemental Material, p. 12, legend to Supplemental Figure S5E:

(E) As Fig. 5A, but based on “relaxed” criteria for identification of NMD-responsive genes. *** Indicates that snoRNA host genes are significantly enriched compared to protein-coding genes (Fisher two-sided test, $P = 8.9 \times 10^{-12}$).

doi: 10.1101/gad.281881.116

(Continued on next page)

Figure 1

Figure 2

(Continued on next page)

Corrigendum

Figure 3

Figure 4

(Continued on next page)

Figure 5

Figure 6

(Continued on next page)

Corrigendum

Figure 7

Corrigendum: Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes

Søren Lykke-Andersen, Yun Chen, Britt R. Ardal, et al.

Genes Dev. 2016, **30**:

Access the most recent version at doi:[10.1101/gad.281881.116](https://doi.org/10.1101/gad.281881.116)

Related Content **Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes**
Søren Lykke-Andersen, Yun Chen, Britt R. Ardal, et al.
[Genes Dev. November , 2014 28: 2498-2517](#)

License

Email Alerting Service Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).