BMC Genomics

CrossMark

Global accumulation of circRNAs during aging in *Caenorhabditis elegans*

Mariela Cortés-López[†], Matthew R. Gruner[†], Daphne A. Cooper, Hannah N. Gruner, Alexandru-Ioan Voda, Alexander M. van der Linden[*] and Pedro Miura[*]

Abstract

Background: Circular RNAs (CircRNAs) are a newly appreciated class of RNAs that lack free 5′ and 3′ ends, are expressed by the thousands in diverse forms of life, and are mostly of enigmatic function. Ostensibly due to their resistance to exonucleases, circRNAs are known to be exceptionally stable. Previous work in *Drosophila* and mice have shown that circRNAs increase during aging in neural tissues.

Results: Here, we examined the global profile of circRNAs in *C. elegans* during aging by performing ribo-depleted total RNA-seq from the fourth larval stage (L4) through 10-day old adults. Using stringent bioinformatic criteria and experimental validation, we annotated a high-confidence set of 1166 circRNAs, including 575 newly discovered circRNAs. These circRNAs were derived from 797 genes with diverse functions, including genes involved in the determination of lifespan. A massive accumulation of circRNAs during aging was uncovered. Many hundreds of circRNAs were significantly increased among the aging time-points and increases of select circRNAs by over 40-fold during aging were quantified by RT-qPCR. The expression of 459 circRNAs was determined to be distinct from the expression of linear RNAs from the same host genes, demonstrating host gene independence of circRNA age-accumulation.

Conclusions: We attribute the global scale of circRNA age-accumulation to the high composition of post-mitotic cells in adult *C. elegans*, coupled with the high resistance of circRNAs to decay. These findings suggest that the exceptional stability of circRNAs might explain age-accumulation trends observed from neural tissues of other organisms, which also have a high composition of post-mitotic cells. Given the suitability of *C. elegans* for aging research, it is now poised as an excellent model system to determine whether there are functional consequences of circRNA accumulation during aging.

Keywords: circRNA, *C. elegans*, Aging, RNA-seq, Splicing, Age-accumulation, Gene expression

Background

Circular RNAs (circRNAs) have recently been identified as a natural occurring family of widespread and diverse endogenous RNAs [1, 2]. They are highly stable molecules mostly generated by backsplicing events from protein-coding genes. The expression trends of circRNAs are only recently emerging thanks to RNA-seq library preparation methods that deplete ribosomal RNA (ribo-depletion) rather than enrich for polyadenylated RNA. Most circRNAs are derived from protein-coding genes, and thus one challenge in mapping and quantifying circRNAs is to distinguish reads that can be uniquely ascribed to circular molecules versus linear RNAs emanating from the same

gene. Elements located within introns flanking circularizing exons play a role in promoting circRNA biogenesis [3–6], and several RNA binding proteins and splicing factors have been shown to influence circRNA expression [4, 7–10].

Despite the current interest in circRNAs, their functions are only beginning to emerge [2]. Recent reports have identified roles for circRNAs in regulating transcription, protein binding, and sequestration of microRNAs [11–14]. Some circRNAs can be translated via cap-independent mechanisms to generate proteins [15–17]. Moreover, circRNAs have been implicated in antiviral immunity [18, 19], and expression patterns of circRNAs in the brain suggest that they might serve important functions in the nervous system [20].

Several RNA-seq studies have found that circRNAs are differentially expressed during aging. Over 250 circRNAs increased in expression within *Drosophila* head tissue

* Correspondence: avanderlinden@unr.edu; pmiura@unr.edu
[†]Equal contributors
University of Nevada, Reno, Department of Biology, 1664 N. Virginia St, Reno, NV 89557, USA

Cortés-López *et al. BMC Genomics* (2018) 19:8

Page 2 of 12

between 1 and 20 days of age [21]. Trends for increased circRNA expression have also been identified during embryonic/postnatal mouse development [10, 22, 23], suggesting that circRNA accumulation might begin early in development. We recently reported that circRNAs were biased for age-accumulation in the mouse brain [24]. In hippocampus and cortex, ~5% of expressed circRNAs were found to increase from 1 month to 22 months of age, whereas ~1% decreased [24]. This accumulation trend was independent of linear RNA changes from cognate genes and thus was not attributed to transcriptional regulation. CircRNA accumulation during aging might be a result of the enhanced stability of circRNAs compared to linear RNAs [13, 25]. Age-related deregulation of alternative splicing [26, 27] leading to increased circRNA biogenesis might also play a role.

C. elegans is a powerful model organism for studying aging. Previously, thousands of circRNAs were annotated from RNA-seq data obtained from *C. elegans* sperm, oocytes, embryos, and unsynchronized young adults [4, 13]. Here, we annotated circRNAs from very deep total RNA-seq data obtained from *C. elegans* at different aging time points and uncovered 575 novel circRNAs. A massive trend for increased circRNA levels with age was identified. This age-accumulation was independent of linear RNA changes from shared host genes. Our findings suggest that circRNA resistance to degradation in post-mitotic cells is largely responsible for the age-upregulation trends identified both here in *C. elegans,* and possibly in neural tissues of other animals.

Results
Genomic features of circRNAs in *C. elegans*
We set out to map *C. elegans* circRNAs genome-wide and quantify their expression at different ages using RNA-seq. We performed an aging paradigm of wild-type Bristol N2 worms that involved treatment with 5-fluoro-2′-deoxyuridine (FUdR) to inhibit DNA synthesis during embryogenesis and prevent egg-hatching. RNA from whole worms from three independent biological replicates corresponding to four aging time-points were collected: L4-larval stage (L4), Day-1 (D-1), Day-7 (D-7), and Day-10 (D-10) adults (Additional file 1: Figure S1). Ribo-depleted total RNA-seq library preparation was performed, followed by sequencing using paired-end 125 nt reads. De novo mapping of circRNAs from these total RNA-seq datasets was performed using the find_circ algorithm [13] (see Methods), with the added restriction of only annotating circRNAs that shared known exonic splice sites (ce11 UCSC genome). This strategy for mapping circRNAs requires the use of only "back-spliced reads" (Fig. 1a), which represent a very low percentage of a typical RNA-seq run (Additional file 2: Table S1). This approach was required to distinguish reads corresponding to circRNAs versus

their linear counterparts that share the same exons. Of the ~1.9 billion paired-end reads generated, only 111,895 reads (0.006% of total) mapped to circRNA junctions after removal of PCR duplicate reads (Additional file 2: Table S1). To annotate circRNAs with high confidence, we used a cut-off of 12 unique aligned reads per circRNA across the 12 libraries. This minimum read cut-off was more stringent compared to previous circRNA annotations in *C. elegans* [4, 13]. Using our annotation pipeline (Additional file 3: Figure S2), we confidently identified a total of 1166 circRNAs. In this high confidence list, 591 circRNAs were previously annotated [4, 13], and 575 were novel (Fig. 1b).

Most of the 1166 circRNAs mapped to coding-sequence (CDS) regions of exons (82.1%), followed by circRNAs mapping to exons encompassing 5′ UTR regions and CDS (14.8%) (Fig. 1b). We found that 797 genes express at least one circRNA. As shown for *haf-4* (Fig. 1c), most genes that produced circRNAs expressed a single circRNA (576/797 genes). On the other hand, some genes were found to generate a large number of circRNAs. For instance, the *afd-1* gene was found to generate 8 different circRNAs (Fig. 1d). Overall, 221 out of the 797 genes generated two or more circRNAs (Fig. 1e). The number of exons within circRNAs ranged from 1 to 13, but it was most common for them to harbor 2 exons (Fig. 1f). Only 6.6% of the 1166 circRNAs contained 5 or more exons. The reliance of this analysis on back-spliced reads precludes the determination of whether these multi-exon circRNAs have introns or particular exons removed. As previously found for *Drosophila* and mice [21, 24] there was a bias for circRNAs to emanate from the 5′ end of genes (Fig. 1g).

Base pairing between introns that flank circularizing exons is thought to bring 5′ and 3′ splice sites in close proximity to promote circRNA biogenesis over linear-splicing (Fig. 1h). We used Basic Local Alignment Search Tool (BLAST) alignment of introns that flank circRNA-forming exons to identify reverse complementary matches (RCMs). We found that RCMs flanking circRNA loci were strongly enriched ($P < 0.0001$, Kruskal-Wallis test with Dunn's post-hoc test for multiple comparisons) compared to analogous introns flanking non-circularizing exons (Fig. 1h). Thus, consistent with previous reports [4], our analysis shows that *C. elegans* circRNAs tend to be flanked by introns that pair with one another.

Experimental validation of circRNAs
We next performed experimental validation of individual circRNAs annotated from our pipeline. One validation method was to prepare complementary DNA (cDNA) using random hexamers from total RNA, and then perform PCR using outward facing primers that should only amplify a back-spliced circRNA (Fig. 2a). The presence
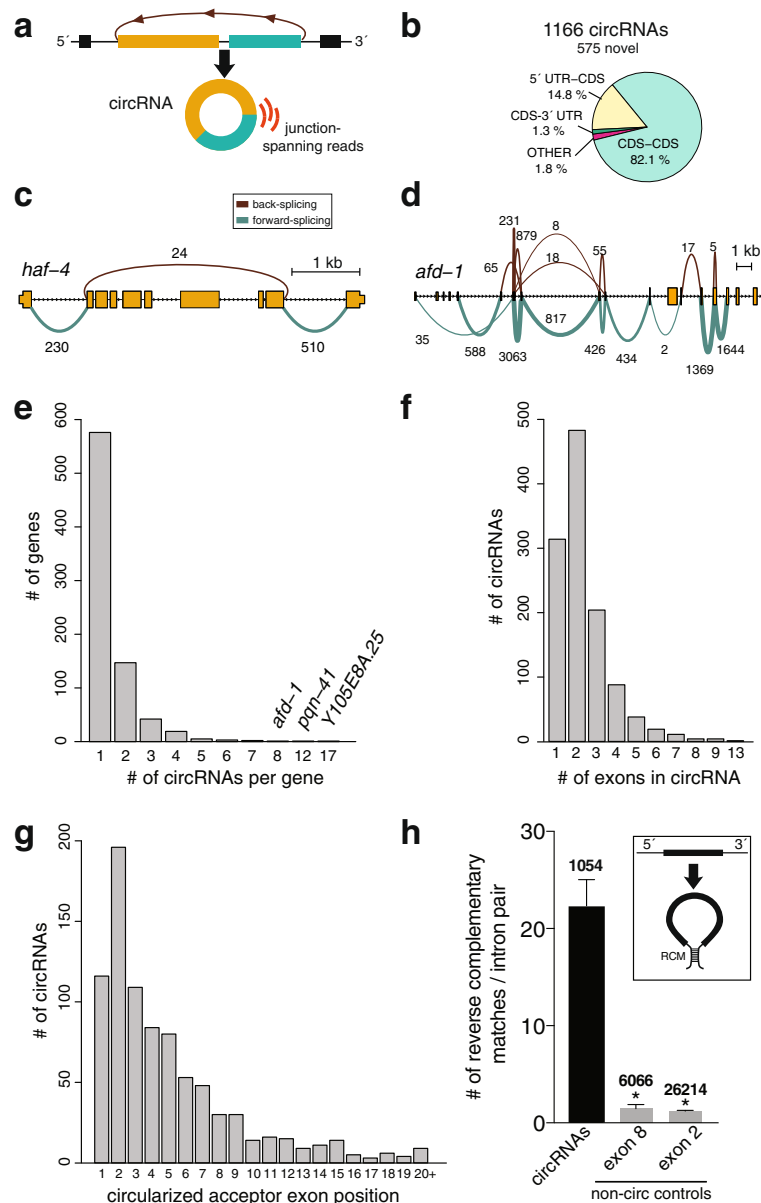
Fig. 1 Genomic features of *C. elegans* circRNAs. **a** Schematic showing a circRNA generated by backsplicing of exons, and the mapping of reads to the back-spliced junction. **b** Distribution of circRNAs in the *C. elegans* genome. Data was mapped from 12 total RNA-seq libraries of N2 worms, including L4 larvae (L4), Day 1 (D-1), Day 7 (D-7), and Day 10 (D-10). CDS, protein coding sequence. **c** Forward and reverse splicing patterns for the *haf-4* gene. Linear spliced read count (green) and back-splicing read count (brown) are shown. Numbers correspond to the number of spliced reads detected in the D-10 datasets. Only reads corresponding to the junctions included in circRNAs are shown. The gene *haf-4* generates a single circRNA that extends across 8 exons. **d** *afd-1* generates 8 circRNAs. **e** Bar plot showing the number of expressed circRNAs per gene. **f** Number of exons contained within exonic circRNAs. **g** Ranked position of circRNA first exon for circRNAs containing more than 1 exon. **h** Presence of Reverse Complementary Matches (RCM) in introns flanking circRNA exons is greater than non-circRNA generating exon controls. Number above bars correspond to # of loci. *, $P <$ 0.0001 on Kruskal-Wallis test with Dunn's post-hoc test for multiple comparisons

of a back-spliced junction was confirmed for 10/10 circRNAs tested by Sanger sequencing of RT-PCR products (Additional file 4: Table S2). In addition, we confirmed a subset of circRNAs by treating total RNA with the exoribonuclease RNase R, which is known to preferentially degrade linear RNAs over circRNAs [5, 13]. RT-qPCR experiments show that linear RNA *cdc-42* was susceptible to degradation by RNase R, whereas 4/4 circRNAs tested were enriched upon RNase R treatment (Fig. 2b).

Although most circRNAs were of lower abundance compared to linear RNAs from the same gene, some circRNAs annotated were of relatively high abundance. We
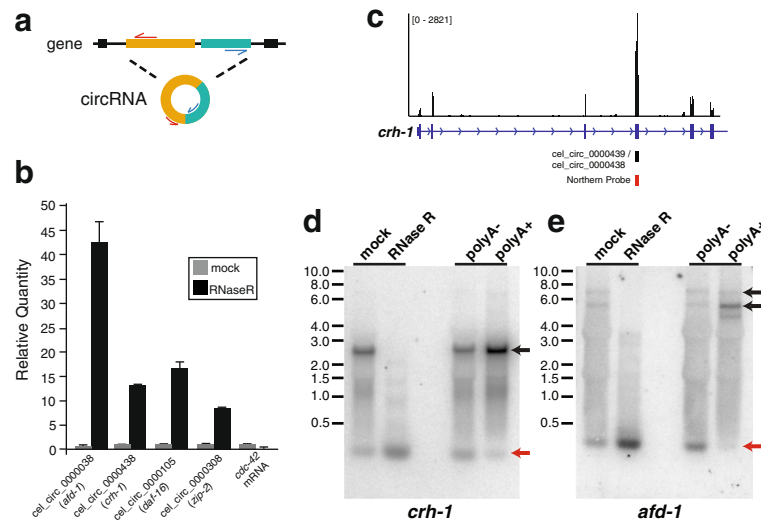
Cortés-López *et al. BMC Genomics* (2018) 19:8

Page 4 of 12

Fig. 2 Experimental validation of circRNAs. **a** RT-PCR strategy to detect circRNAs exclusively using outward facing primer sets. Sanger sequencing of PCR products confirmed 10/10 circRNAs tested (Additional file 4: Table S2). **b** RT-qPCR experiments on RNase R treated mixed age adult worms. Equal amounts of mock-treated and RNase R treated RNA were used for cDNA preparation prior to qPCR. Note the enrichment of circRNAs with RNase R treatment, whereas linear *cdc-42* mRNA is not enriched. **c** RNA-seq track visualized using Integrated Genomics Viewer from D-7 worms showing read pileup at the *crh-1* gene. Note the increased read number overlapping the circularized exon. cel_circ_0000438 and cel_circ_0000439 differ by 6 nucleotides in length at the 5′ end of the exon. **d** Northern blot using a probe overlapping the circularized exon of *crh-1* (see panel **c**) detects bands corresponding to circRNA and mRNA from mixed age adult RNA. Relative circRNA to mRNA abundance is enriched in RNase R treated and polyA- samples compared to polyA+ samples. **e** Northern blot performed using a probe that detects *afd-1* circRNA and mRNA. Red arrows denote circRNA bands. Black arrows denote likely linear RNA bands

set out to confirm two of these high abundance circRNAs using Northern blot analysis. In the case of the *crh-1* gene, an increased abundance of reads aligning to an exon harboring a circRNA was clearly evident from linear RNA alignment tracks on Integrated Genomics Viewer [28] (Fig. 2c). We performed Northern analysis using a probe targeting this exon. This probe should detect both linear and circular transcripts of the *crh-1* gene. As expected, an abundant circRNA migrating at the predicted size was detected (Fig. 2d). The expression of higher molecular weight linear RNAs was found to be diminished by RNase R treatment, whereas the circRNA bands were unaffected by RNase R treatment (Fig. 2d). We prepared polyA+ RNA from a column-based preparation, and collected and precipitated the unselected RNA (polyA-depleted). We found that polyA+ RNA had depleted circRNA levels relative to linear RNA. In contrast, polyA- samples showed enhanced levels of circRNA relative to linear RNA (Fig. 2d). Analogous results were obtained using a probe for an abundant circRNA from the *afd-1* gene (Fig. 2e). Together, these validations provide experimental support that our annotation pipeline detected bona fide circRNAs in *C. elegans*.
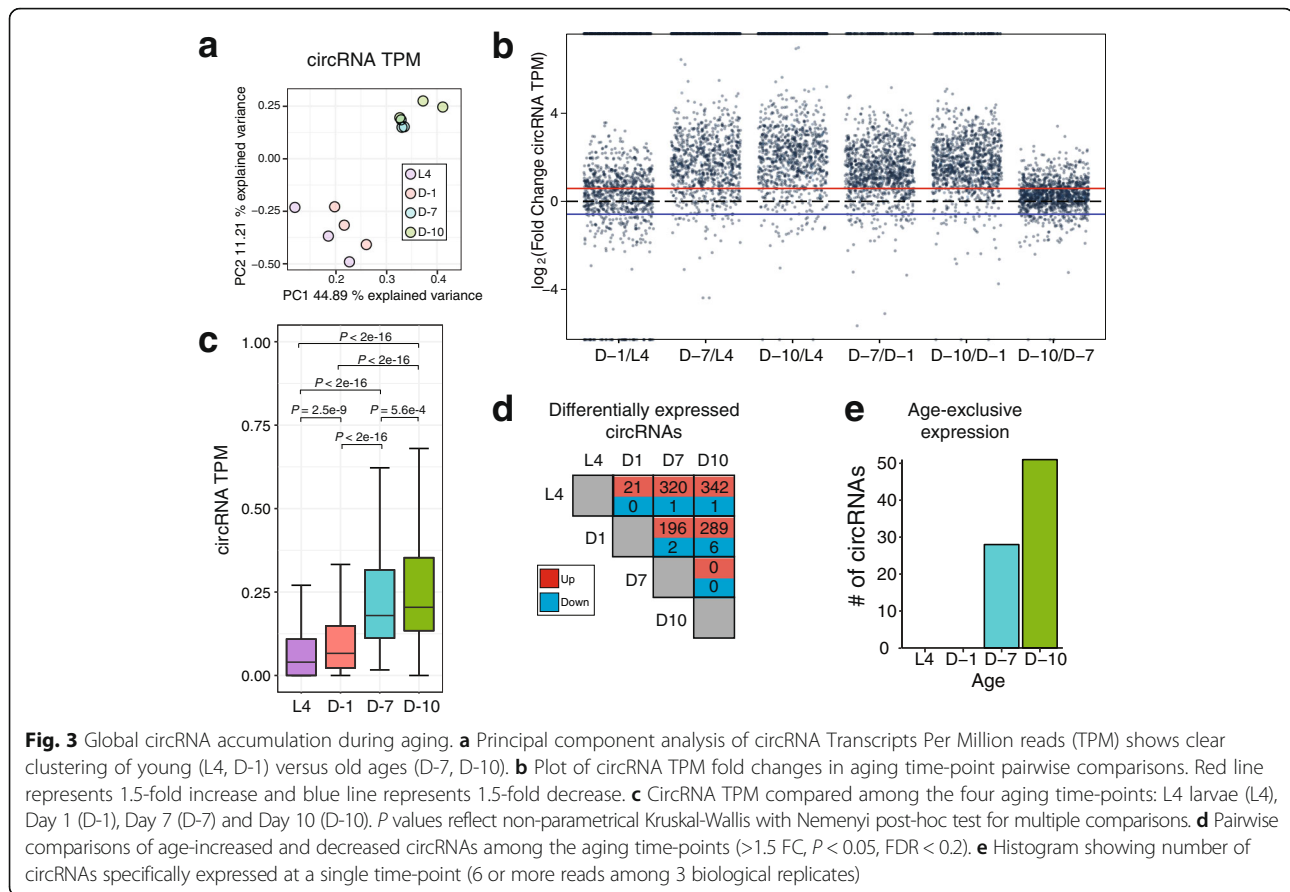
Global circRNA levels dramatically increase during aging

We next quantified the abundance of circRNAs from the different aging time-points. CircRNA read counts were

normalized to their corresponding library size to obtain Transcripts Per Million reads (TPM) (Additional file 5: Table S3). Principal Component Analysis on the circRNA TPM values was performed on the 12 RNA-seq libraries (Fig. 3a). A close clustering of L4 to D-1, and of D-7 to D-10 was observed, suggesting that global circRNA expression levels reflect the age of the animal.

To further investigate trends in circRNA levels during aging, we plotted circRNA log_2 fold changes in TPM for all pairwise comparisons of the aging time-points (Fig. 3b). We found that 1052 circRNAs (90.2%) were at least 1.5-fold greater in D-10 versus L4 time-points, whereas only 37 circRNAs were 1.5-fold greater in L4. Similar trends were found in other pairwise comparisons between older (D-7, D-10) versus younger (L4, D-1) time-points (Fig. 3b). For instance, when comparing D-7 versus D-1, 80.8% of circRNAs were 1.5-fold or higher in the older time-point.

To gain statistical support for these dramatic aging trends we performed several additional analyses. The global expression of circRNA TPM values was compared across ages by non-parametrical Kruskal-Wallis test with Nemenyi post-hoc test for multiple comparisons (Fig. 3c). Comparisons between more distant aging time-points (D-10/L4, D-7/L4, D-10/D-1, and D-7/D-1) yielded the lowest *P* values ($P < 2e-16$). In contrast, the D-1/L4 and D-10/D-7 comparisons had less significant *P* values ($P =$

Fig. 3 Global circRNA accumulation during aging. **a** Principal component analysis of circRNA Transcripts Per Million reads (TPM) shows clear clustering of young (L4, D-1) versus old ages (D-7, D-10). **b** Plot of circRNA TPM fold changes in aging time-point pairwise comparisons. Red line represents 1.5-fold increase and blue line represents 1.5-fold decrease. **c** CircRNA TPM compared among the four aging time-points: L4 larvae (L4), Day 1 (D-1), Day 7 (D-7) and Day 10 (D-10). *P* values reflect non-parametrical Kruskal-Wallis with Nemenyi post-hoc test for multiple comparisons. **d** Pairwise comparisons of age-increased and decreased circRNAs among the aging time-points (>1.5 FC, *P* < 0.05, FDR < 0.2). **e** Histogram showing number of circRNAs specifically expressed at a single time-point (6 or more reads among 3 biological replicates)

2.5e-9 and 5.6e-4, respectively), which might reflect the ages being closer together.

In order to identify the individual circRNAs with statistically significant changes in expression during *C. elegans* aging, we performed *t*-tests on TPM values between each time-point (*P* < 0.05, > 1.5 Fold Change (FC), False Discovery Rate (FDR) < 0.2). An overwhelming bias for upregulation of circRNAs during aging was uncovered. For instance, in the comparison of D-7 versus D-1, a total of 196 circRNAs were upregulated whereas only 2 were downregulated (Fig. 3d). Comparing D-10 versus L4 age time-points, 342 circRNAs were upregulated, whereas only 1 was downregulated (Fig. 3d). We next sought to identify circRNAs exclusively detected in one aging time-point. Using a 6 unique back-spliced read minimum cutoff, we plotted age-specific circRNAs (Fig. 3e). We detected 28 and 51 age-specific circRNAs in D-7 and D-10 libraries, respectively. In contrast, no age-specific circRNAs were found for the L4 and D-1 time-points. Note that the D-7 and D-10 specific circRNAs might be expressed in the younger worms, but at levels below the detection limits of the RNA-seq analysis. Together, these data demonstrate that circRNAs show an overwhelming bias for age-accumulation in *C. elegans*.

Experimental validation of circRNA age-accumulation trends

We next set out to confirm circRNA expression trends for particular circRNAs that are generated from genes with interesting functions. We performed RT-qPCR validation for 11 individual circRNAs, including circRNAs generated from genes that are involved in lifespan determination (*akt-1, crh-1, daf-16, daf-2*). We selected circRNAs with a variable range of overall expression levels for validation (Fig. 4a). Of these 11 circRNAs, 7 met the statistical threshold for increased expression between at least one old versus young time-point. On the other hand, two of the circRNAs were predicted to be significantly decreased during aging. Three of these circRNAs did not meet statistical significance for differential expression (D-10 versus L4) from the RNA-seq data, including two circRNAs of low abundance (Fig. 4a). Quantification of these same circRNAs by RT-qPCR revealed that 7/7 circRNAs predicted to increase during aging were significantly upregulated in D-10 versus L4 (Fig. 4b). Interestingly, the three circRNAs predicted to not increase during aging from the RNA-seq data (including two low abundance circRNAs from the *daf-2* and *daf-16* genes) were found to robustly increase during aging by RT-qPCR. Finally, the two circRNAs
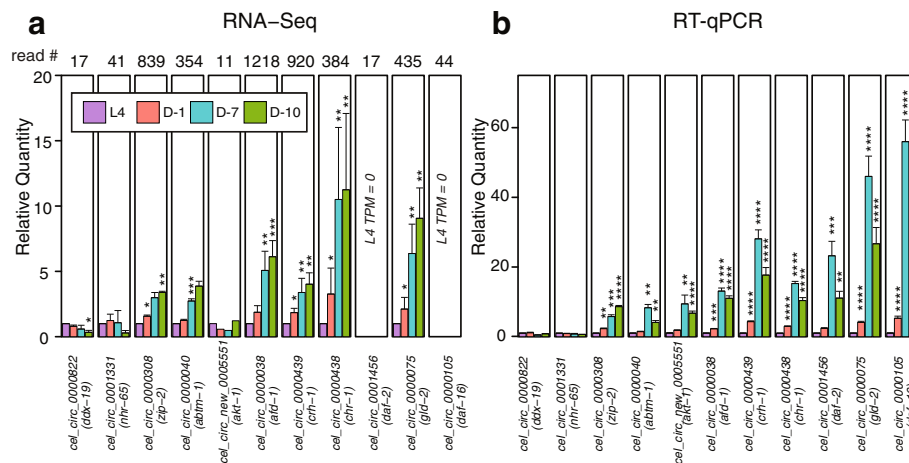
Cortés-López *et al. BMC Genomics* (2018) 19:8

Page 6 of 12

Fig. 4 Validation of circRNA age-accumulation. **a** RNA-seq quantification of select circRNAs during aging (TPM fold-change with L4 set at 1). Total number of reads across all libraries for each circRNA is noted above graph. Labels display circRNA names with host gene in brackets. **b** RT-qPCR data for the same selected circRNAs as in A). Data is normalized to *cdc-42* mRNA. Note the greater magnitude of age-related expression changes reported by RT-qPCR versus RNA-seq for all circRNAs. Note that for cel_circ_0001331, there was a significant reduction ($P < 0.05$) between D-1 and D-10 (Additional file 5: Table S3). Error bars represent SEM. *, $P \leq 0.05$; **, $P \leq 0.01$; ***, $P \leq 0.001$; ****, $P \leq 0.0001$, two-tailed *t*-test compared to L4

predicted to decrease during aging (from *ddx-19* and *nhr-65* genes) were found to be unchanged when tested by RT-qPCR analysis (Fig. 4b). RT-qPCR analysis also showed that the tested circRNAs did not continue to increase in D-10 animals, consistent with the RNA-seq results (Fig. 4b). Notably, RT-qPCR validation for most circRNAs showed fold-changes greater than those detected by RNA-seq differential expression analysis. Most of these RT-qPCR quantified changes were >10-fold between L4 and D-7. Remarkably, changes in circRNAs from the *gld-2* and *daf-16* genes were > 40-fold increased between L4 and D-7 (Fig. 4b). Overall, these RT-qPCR validations strongly support the trend of greater circRNA abundance in old (D-7, D-10) versus young (L4, D-1) animals. These confirmations also suggest that the actual number of circRNAs that increase during aging might be greater than what was found to significantly change from the RNA-seq analysis (Fig. 3d).

Host gene-independent circRNA accumulation during aging

Next, we performed differential expression analysis on linear RNAs among the different age time-points. Linear RNAs previously found to be differentially regulated during aging displayed similar expression trends in our datasets. For example, between D-7 and D-1, *hsp-70* and *cht-1* were upregulated during aging, whereas *fat-7*, *ifp-1*, and *ifd-1* were downregulated (Additional file 6: Table S4). In contrast to circRNA trends, a global bias for linear RNA differential expression was not evident. Similar numbers of upregulated and downregulated linear RNAs between aging time points were identified using Cuffdiff [29] (Additional file 6: Table S4). Scatterplots comparing old

versus young time-points for linear RNA levels (Fig. 5a-d), and circRNA levels (Fig. 5e-h) exemplify the stark contrast in the age-related trends.

Although linear RNAs lacked a global bias for increased levels during aging, it was still possible that increased transcription of circRNA-hosting genes could contribute to the circRNA expression trends. Thus, we analyzed whether circRNA accumulation was independent of host gene expression. Density plots were generated to contrast circRNA total read count fold-changes versus their counterpart linear RNA changes from the same host gene (Fig. 6a-f). For this analysis, we used an expanded list of circRNAs, requiring at least 6 reads per age time-point. Using the CircTest algorithm [30], we tested for statistically significant changes in circRNA expression independent of host gene expression. From 1239 circRNAs tested, 459 showed a host gene independent significant expression (ANOVA *P* value corrected <0.05, Additional file 7: Table S6). In the old versus young time-point comparisons, a clear upward vertical shift was evident in the density plots (reflecting increased circRNA expression), and only a minor horizontal shift to the right (reflecting increased linear RNA expression) (Fig. 6b-e). This suggests that circRNA accumulation trends are largely independent of linear RNA changes. For comparisons between closer time points (D-1 versus L4 and D-10 versus D-7) the density plots lacked clear vertical or horizontal shifts (Fig. 6a, f). We next examined the host gene independent circRNAs for significant changes between time-points. We found strong trends for these circRNAs to be increased in older versus younger aging time-points (Fig. 6g). For example, 194 circRNAs were significantly increased in a host-independent manner between D-10 and L4, whereas no circRNAs were found to

Cortés-López *et al. BMC Genomics* (2018) 19:8

Page 7 of 12

Fig. 5 Comparison of circRNA and linear RNA expression during aging. **a** Scatterplot showing pairwise comparisons between aging time-points for linear RNA levels (**a-d**) and circRNAs (**e-h**). Log$_2$ linear RNA Fragments Per Kilobase per Million reads (FPKM) value scatterplots are shown for (**a**) D-7 vs L4, (**b**) D-10 vs L4, (**c**) D-7 vs D-1, and (**d**) D-10 vs D-1. Log$_2$ circRNA TPM scatterplots are shown for (**e**) D-7 vs L4, (**f**) D-10 vs L4, (**g**) D-7 vs D-1, and (**h**) D-10 vs D-1. For circRNA TPM analysis, significant changes have a fold-change >1.5, $P < 0.05$, FDR < 0.2. Significant changes for linear RNA were computed using CuffDiff (Fold change >1.5, $P < 0.05$; see Methods). Red data points show age-upregulated transcripts, whereas blue data points show age-downregulated transcripts

decrease. Together, these results show that circRNAs globally accumulate during aging in *C. elegans* independently of host gene expression.

Discussion

This study is the first to report age-accumulation of circRNAs in *C. elegans*. Previous studies have documented the bias for circRNAs to be increased during aging in neural tissues of *Drosophila* and mice [21, 24]. Interestingly, the trends uncovered here during *C. elegans* aging are much more dramatic. Many confirmed expression trends were >10-fold increased between L4 and D-7. Of the hundreds of differentially expressed circRNAs, the vast majority increased with age (Figs. 3, 4, 5).

Experimental confirmations of the age-accumulation trends by RT-qPCR were of greater magnitude than those reported from the RNA-seq analysis. Although we have higher confidence in the accuracy of circRNA quantification by RT-qPCR over RNA-seq due to the low depth of junction-spanning sequencing reads, it has been discussed that RT-qPCR might overestimate circRNA levels due to rolling circle reverse transcription [1]. It is notable that RT-qPCR experiments revealed two circRNAs to accumulate with age that were not significantly increased in the RNA-seq data, and two circRNAs predicted by RNA-seq to decrease with age were found to not change during aging (Fig. 4). Together, the

evidence presented here suggests that many circRNAs not passing statistical thresholds for increased expression during the *C. elegans* aging paradigm are of too low abundance to be accurately quantified by the limited number of back-spliced RNA-seq reads. Thus, our reported numbers of significant age-accumulated circRNAs are most likely gross underestimations.

Why is the age-accumulation trend of circRNAs in *C. elegans* much stronger than in other organisms tested so far? After completing development and a brief reproductive period, *C. elegans* spends the remainder of its adult life comprised almost exclusively of post-mitotic cells [31]. The FUdR treatment employed in this study inhibits DNA synthesis and is commonly used to prevent egg-hatching in aging and life-span studies of *C. elegans* [32], and also reduces the presence of proliferating cells. We have previously proposed that either the high stability of circRNAs or alterations in alternative splicing with age could both contribute to the increase of circRNAs during aging [2]. The trends observed here in *C. elegans* argue for a low rate of circRNA decay being the responsible mechanism. We propose that the dramatic genome-wide increase of circRNA levels during aging are a consequence of the dominance of post-mitotic cells in adult *C. elegans* combined with the high stability of circRNAs. As neurons are post-mitotic, perhaps this can also explain why age-accumulation is most
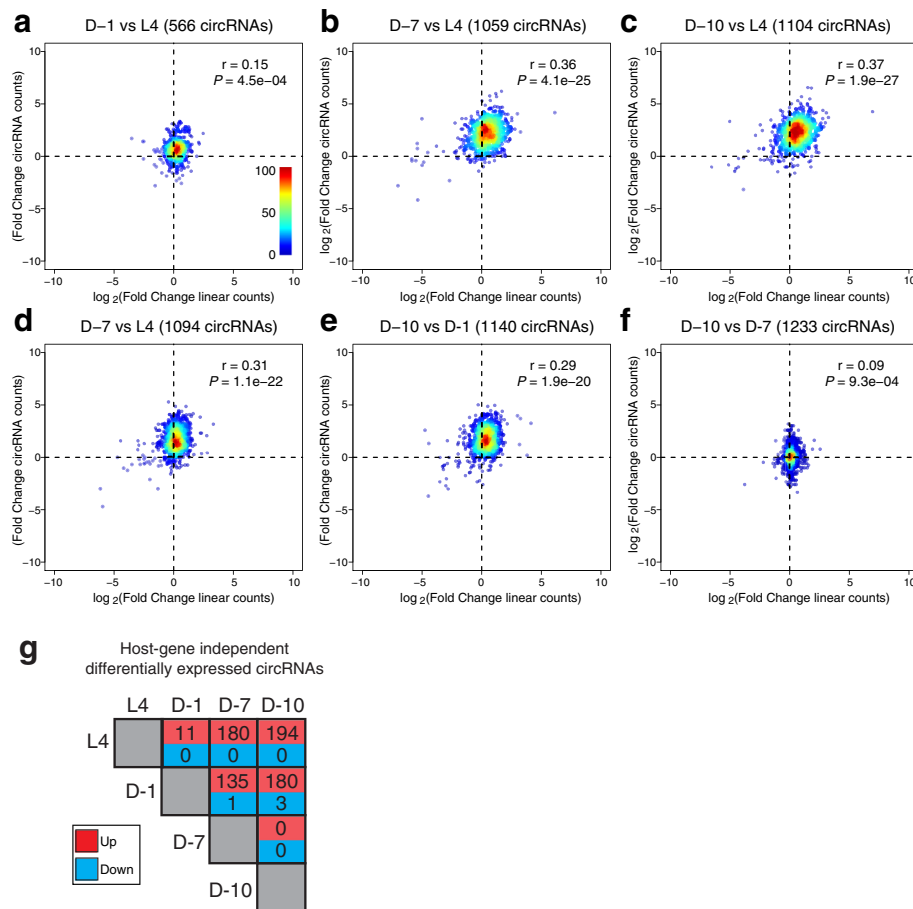
Cortés-López *et al. BMC Genomics* (2018) 19:8

Page 8 of 12

Fig. 6 Age-accumulation of circRNAs is independent of host gene expression. Density plots for CircTest-derived circRNA read counts fold-change versus linear read count RNA fold-change. Log₂ fold-changes of circRNAs versus log₂ fold-changes of linear RNAs from host genes are shown. **a** D-1 vs L4, (**b**) D-7 vs L4, (**c**) D-10 vs L4, (**d**) D-7 vs L4, and (**e**) D-10 vs D-1, (**f**) D-10 vs D-7. Scale bar inset in panel A represents circRNA number and applies to all the density plots. For old versus young time-point comparisons, it is evident that upregulation of circRNAs is largely independent of linear RNA expression from the same gene (upward shift in plots). Pearson correlation values are shown in the upper right corner, indicating weak correlation between the circular and linear ratios in all comparisons. Plots include circRNAs with a minimum of 6 reads for each time-point under comparison. **g** Pairwise comparisons of CircTest-derived counts showing significant host gene independent changes (Significant expression changes: $P < 0.05$, FDR < 0.2)

notable in *Drosophila* heads (which are rich in neurons) [21], and in brain regions of mice [24].

To investigate the potential functional significance of age-accumulated circRNAs, we performed Gene Ontology (GO) analysis on the host genes of expressed *C. elegans* circRNAs, and found many significantly enriched categories, including an enrichment in the Biological Process category of "determination of adult lifespan" (Additional file 8: Figure S3, Additional file 9: Table S5). A clear bias for particular GO categories, however, not uncovered by these efforts. It is certainly possible that *trans* functions of circRNAs are completely distinct from the curated roles of their host genes.

Future work can now take advantage of the powerful genetics of *C. elegans* to delineate aging functions of individual circRNAs. Generating loss-of-circRNA mutants

in *C. elegans* by disrupting base-pairing of flanking introns could be a fruitful approach. Various RNAs found to be differentially regulated during aging were subsequently found to impact lifespan in *C. elegans* mutant analysis, including linear RNAs [33], microRNAs [34] and long non-coding RNAs [35]. Given that these mutant studies on lifespan were based on comparatively modest fold-changes during aging, the massive upregulation trends for circRNAs provide solid rationale for disrupting or overexpressing circRNAs in *C. elegans* and testing for effects on lifespan and healthspan. However, one should also consider that the aging process might be impacted generally by the total compendium of hundreds of circRNAs accumulating in cells, as opposed to individual circRNAs. Thus, non-conventional approaches to alter the expression of many circRNAs

Cortés-López *et al. BMC Genomics* (2018) 19:8

Page 9 of 12

simultaneously might be required to uncover age-related functions of circRNAs.

Conclusion

We have shown that circRNAs accumulate during aging on a genome-wide scale in *C. elegans*. Given that circRNAs are also increased during aging in *Drosophila* [21] and mice [24], age-accumulation of circRNAs in post-mitotic cells appears to be a universal phenomenon.

Methods

C. elegans maintenance and culturing

The *C. elegans* Bristol N2 wild type strain was grown and maintained as previously described [36]. To synchronize populations, gravid adults were bleached, eggs were collected and left overnight in 1X M9 buffer with rocking. Starvation arrested L1 larvae were placed on 150x100mm NGM plates with 10X concentrated *E. coli* OP50 as a primary food source, and kept at 15 °C. At the L4 larval stage, animals were collected using a 25 μm nylon mesh (Sefar) and either harvested for RNA extraction, or placed on 150x100mm NGM plates containing 75 mM FUdR (Sigma Aldrich) with 10X concentrated *E. coli* OP50 and kept at 15 °C until animals were harvested at D-1, D-7 and D-10 of adulthood (see Additional file 1: Figure S1). Three biological replicates were collected for each time-point and used for RNA extraction and RNA-seq library preparation.

RNA extraction

A 250 μl mixture of animals in 1X M9 buffer was added to 750 μl of TRIzol LS (ThermoFisher Scientific) and immediately frozen with liquid N_2. Lysates were freeze/thawed at −80 °C, disrupted with Mixer Mill 400 (Retsch) and Dounce homogenizer (Corning) to break apart the cuticle of animals. Any cellular debris was removed by low-speed centrifugation. RNA was extracted using the Purelink RNA mini-kit with DNAse I treatment (Ambion). RNA quality was assessed by Bioanalyzer (Agilent) and quantified using Quant-iT RiboGreen RNA Assay kit (ThermoFisher Scientific).

Library preparation and high-throughput sequencing

Libraries were prepared using the Illumina TruSeq Stranded Total RNA Library Prep Kit as recommended by the manufacturer (Illumina) with modified conditions to increase the size of the cloned fragments (fragmentation at 85 °C × 5 min). Barcoded libraries were sequenced at New York Genome Center (New York, NY) using the Illumina HiSeq 2500 system to obtain paired-end 125 nt reads. Raw FASTQ files from the RNA-seq data were deposited at the NCBI Sequence Read Archive (BioProject: PRJNA357503, individual accession numbers are listed in Additional file 2: Table S1).

Experimental validation of circRNAs

To confirm individual circRNAs, RNA was reverse transcribed using Superscript III with random hexamers (Invitrogen). PCR products were gel extracted then Sanger sequenced, or first cloned into the PCR 2.1- TOPO TA vector (Invitrogen) prior to Sanger sequencing (Nevada Genomics Center, University of Nevada, Reno). For qPCR analysis, we used a BioRad CFX96 real-time PCR machine with SYBR select mastermix for CFX (Applied Biosystems) using the delta-delta Ct method for quantification. These experiments were performed using technical quadruplicates. Total RNA from *C. elegans* was treated with or without 0.4 U/μL RNase R (Epicentre) with 2 U/μL RNaseOUT (ThermoFisher Scientific) for 2 h at 37 °C. RNase R reactions were terminated with 0.5% SDS buffer (0.5% SDS, 10 mM Tris–HCl [pH 7.5], 1.25 mM EDTA [pH 8], 100 mM NaCl) and RNA was purified from reactions using acid phenol chloroform extraction with isopropanol precipitation and 70% ethanol washes. Equal amounts of RNase R or mock treated RNA served as input for cDNA preparation.

PolyA+ RNA and polyA- RNA were obtained from total RNA using the NucleoTrap mRNA kit (Machery-Nagel). RNA bound to oligo(dT) beads was carried through the complete polyA+ enrichment according to the manufacturer's protocol while RNA that remained in the column flow-through (unbound to the oligo(dT) beads) was precipitated with isopropanol and washed with 70% ethanol. Northern analysis was performed as previously described [24], with probe hybridization taking place overnight at 42 °C, and all blot washing steps at 50 °C.

CircRNA prediction and mapping

For de novo identification of circRNAs, a computational pipeline was carried out as previously described with filtering for duplicate reads and removing circRNA annotations spanning multiple genes [24] (Additional file 3: Figure S2). We obtained circRNA junction spanning FASTA sequence templates of 200 nucleotides using the *C. elegans* genome (from WBcel235/ce11 UCSC genome) as a reference. Assignment of circRNAs to their corresponding parental genes was performed using custom R scripts based on the library GenomicFeatures [37].

CircRNA normalization and quantification

A minimum of 12 reads summed from all 12 libraries was required for each circRNA in order to be considered for downstream analyses. To account for variability due to differences in library size, counts attributed to individual circRNAs were normalized to total read count to obtain circRNA TPM values. We calculated the corresponding fold changes by pair-wise comparisons of the average TPMs between time-points. Individual unpaired *t*-tests were performed for these pairwise comparisons. *P* values

were corrected for multiple hypothesis testing with FDR < 0.2. To identify stage-specific circRNAs (Fig. 3e), and for pair-wise comparisons of host gene expression independent changes using CircTest [30] (Fig. 6), we required a minimum of 6 reads per time-point.

CircRNA host gene independence test

The CircTest pipeline was implemented to test for host gene independent circRNAs [30]. DCC 0.0.4 (https://github.com/dieterich-lab/DCC) was used to quantify host gene read counts. In accordance with the DCC pipeline, FASTQ files were mapped with STAR 2.5 [38] using the recommended parameters. Picard 2.2.4 (http://broadinstitute.github.io/picard/) was used to remove duplicated reads from the BAM files. The circRNA raw read counts (using an annotation list of circRNAs meeting a 6 read minimum requirement) together with the DCC linear counts were used as input for the Circ.test function which was run with default parameters. To define a circRNA as host gene independently expressed, we set a cutoff of $P < 0.05$, FDR < 0.05.

CircRNA expression and plots

ggplot2 (Wickham, 2012) and ggrepel (https://cran.r-project.org/web/packages/ggrepel/index.html) R libraries were used for scatterplots. Gene model diagrams were generated using the Gviz package [39]. Density plots were generated with the LSD package (https://cran.r-project.org/web/packages/LSD/index.html).

Quantification of linear expression

Linear RNA-seq reads were mapped to the *C. elegans* ce11 annotation using TopHat [40] with default settings. To quantify the differential expression across time-points, we used Cuffdiff [40]. Genes with fold changes > 1.5 and a Benjamini-Hochberg corrected P value < 0.05 (default parameter used in the Cuffdiff algorithm) were considered to be differentially expressed.

GO analysis

We used the Cytoscape plugin ClueGO [41] (together with the GO Annotation (GOA EMBL-EBI) (http://www.ebi.ac.uk/GOA) released on 11/17/2016. The background set for GO was all genes with annotated GO entries for the GOA annotation released on 11/17/2016. Network specificity was set to "medium". The enrichment statistic used was a two-side hypergeometric test, correcting for multiple testing with the Bonferroni method. The cutoff for considering a term as enriched was set at $P < 0.05$. To reduce the number of redundant terms we used the GO term grouping option, which uses a Kappa score to collapse terms that share elements. We set the minimal number of elements in a group at 3. Bar graphs of significant GO terms were created using the –log (P-value).

RCM and motif analysis

To identify RCMs, pairs of intron sequences flanking the circRNAs were extracted from the *C. elegans* genome (WBcel235/ce11) using custom scripts available at: https://github.com/alexandruioanvoda/IntronPicker. The corresponding sequences were used as input for the RCM analysis using custom scripts (https://github.com/alexandruioanvoda/autoBLAST) that employed BLAST (parameters: blastn, word size 7, output format 5) to identify matches. Exons 2 and 8 from non-circRNA generating exons were used as controls to account for the possibility that intron pairing might be influenced by exon location within genes.

Additional files

Additional file 1: Figure S1. *C. elegans* aging paradigm. Protocol for collecting total RNA during *C. elegans* aging. Wild-type animals were fed *E. coli* OP50 and grown at 15°C. Gravid adults were bleached and populations were synchronized as L1 larvae and grown for an additional 4 days. At the L4 larval stage (Day 0), animals were either collected or transferred to FUdR containing NGM agar plates seeded with *E. coli* OP50, and were allowed to continue growth at 15°C. Total RNA was collected at different age time-points (L4, D-1, D-7, and D-10). (PDF 107 kb)

Additional file 2: Table S1. RNA-seq read statistics. (XLSX 11 kb)

Additional file 3: Figure S2. Pipeline for circRNA annotation. A flowchart of the computational pipeline used for circRNA identification. (PDF 138 kb)

Additional file 4: Table S2. Oligonucleotides used for experimental validation. (XLSX 75 kb)

Additional file 5: Table S3. circRNA expression data. (XLSX 385 kb)

Additional file 6: Table S4. Linear RNA differential expression analysis. (XLSX 2355 kb)

Additional file 7: Table S6. Host-independent circRNA expression analysis using CircTest. (XLSX 139 kb)

Additional file 8: Figure S3. GO analysis of circRNA host genes. Visualization of ClueGO analysis of the 797 genes harboring the 1166 expressed circRNAs. X axis represents -log(P-value) of the enrichment. A Complete GO analysis is found in Additional file 9: Table S5. (PDF 141 kb)

Additional file 9: Table S5. Gene Ontology analysis for circRNA host genes. (XLSX 47 kb)

Cortés-López *et al. BMC Genomics* (2018) 19:8

Page 11 of 12

Publisher's Note

References
1. Barrett SP, Salzman J. Circular RNAs: analysis, expression and potential functions. Development. 2016;143(11):1838–47.
2. Cortes-Lopez M, Miura P. Emerging functions of circular RNAs. Yale J Biol Med. 2016;89(4):527–37.
3. Dubin RA, Kazmi MA, Ostrer H. Inverted repeats are necessary for circularization of the mouse testis Sry transcript. Gene. 1995;167(1–2):245–8.
4. Ivanov A, Memczak S, Wyler E, Torti F, Porath HT, Orejuela MR, Piechotta M, Levanon EY, Landthaler M, Dieterich C, et al. Analysis of intron sequences reveals hallmarks of circular RNA biogenesis in animals. Cell Rep. 2015;10(2):170–7.
5. Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF, Sharpless NE. Circular RNAs are abundant, conserved, and associated with ALU repeats. RNA. 2013;19(2):141–57.
6. Liang D, Wilusz JE. Short intronic repeat sequences facilitate circular RNA production. Genes Dev. 2014;28(20):2233–47.
7. Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S. circRNA biogenesis competes with pre-mRNA splicing. Mol Cell. 2014;56(1):55–66.
8. Conn SJ, Pillman KA, Toubia J, Conn VM, Salmanidis M, Phillips CA, Roslan S, Schreiber AW, Gregory PA, Goodall GJ. The RNA binding protein quaking regulates formation of circRNAs. Cell. 2015;160(6):1125–34.
9. Kramer MC, Liang D, Tatomer DC, Gold B, March ZM, Cherry S, Wilusz JE. Combinatorial control of drosophila circular RNA expression by intronic repeats, hnRNPs, and SR proteins. Genes Dev. 2015;29(20):2168–82.
10. Rybak-Wolf A, Stottmeister C, Glazar P, Jens M, Pino N, Giusti S, Hanan M, Behm M, Bartok O, Ashwal-Fluss R, et al. Circular RNAs in the mammalian brain are highly abundant, conserved, and dynamically expressed. Mol Cell. 2015;58(5):870–85.
11. WW D, Yang W, Chen Y, ZK W, Foster FS, Yang Z, Li X, Yang BB. Foxo3 circular RNA promotes cardiac senescence by modulating multiple factors associated with stress and senescence responses. Eur Heart J. 2016;
12. Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J, Natural RNA. Circles function as efficient microRNA sponges. Nature. 2013;495(7441):384–8.
13. Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, et al. Circular RNAs are a large class of animal RNAs with regulatory potency. Nature. 2013;495(7441):333–8.
14. Zhang Y, Zhang XO, Chen T, Xiang JF, Yin QF, Xing YH, Zhu S, Yang L, Chen LL. Circular intronic long noncoding RNAs. Mol Cell. 2013;51(6):792–806.
15. Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, Fatica A, Santini T, Andronache A, Wade M, et al. Circ-ZNF609 is a circular RNA that can be translated and functions in Myogenesis. Mol Cell. 2017;
16. Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, Hanan M, Wyler E, Perez-Hernandez D, Ramberger E, et al. Translation of CircRNAs. Mol Cell. 2017;
17. Yang Y, Fan X, Mao M, Song X, Wu P, Zhang Y, Jin Y, Yang Y, Chen L, Wang Y, et al. Extensive translation of circular RNAs driven by N6-methyladenosine. Cell Res. 2017;
18. Chen YG, Kim MV, Chen X, Batista PJ, Aoyama S, Wilusz JE, Iwasaki A, Chang HY. Sensing self and foreign circular RNAs by Intron identity. Mol Cell. 2017; 67(2):228–38. e225
19. Li X, Liu CX, Xue W, Zhang Y, Jiang S, Yin QF, Wei J, Yao RW, Yang L, Chen LL. Coordinated circRNA biogenesis and function with NF90/NF110 in viral infection. Mol Cell. 2017;67(2):214–27. e217
20. Chen W, Schuman E. Circular RNAs In brain and other tissues: a functional enigma. Trends Neurosci. 2016;39(9):597–604.
21. Westholm JO, Miura P, Olson S, Shenker S, Joseph B, Sanfilippo P, Celniker SE, Graveley BR, Lai EC. Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. Cell Rep. 2014;9(5):1966–80.
22. Szabo L, Morey R, Palpant NJ, Wang PL, Afari N, Jiang C, Parast MM, Murry CE, Laurent LC, Salzman J. Statistically based splicing detection reveals neural enrichment and tissue-specific induction of circular RNA during human fetal development. Genome Biol. 2015;16:126.
23. You X, Vlatkovic I, Babic A, Will T, Epstein I, Tushev G, Akbalik G, Wang M, Glock C, Quedenau C, et al. Neural circular RNAs are derived from synaptic genes and regulated by development and plasticity. Nat Neurosci. 2015; 18(4):603–10.
24. Gruner H, Cortes-Lopez M, Cooper DA, Bauer M, Miura P. CircRNA accumulation in the aging mouse brain. Sci Rep. 2016;6:38907.
25. Zhang Y, Xue W, Li X, Zhang J, Chen S, Zhang JL, Yang L, Chen LL. The biogenesis of nascent circular RNAs. Cell Rep. 2016;15(3):611–24.
26. Harries LW, Hernandez D, Henley W, Wood AR, Holly AC, Bradley-Smith RM, Yaghootkar H, Dutta A, Murray A, Frayling TM, et al. Human aging is characterized by focused changes in gene expression and deregulation of alternative splicing. Aging Cell. 2011;10(5):868–78.
27. Rodriguez SA, Grochova D, McKenna T, Borate B, Trivedi NS, Erdos MR, Eriksson M. Global genome splicing analysis reveals an increased number of alternatively spliced genes with aging. Aging Cell. 2016;15(2):267–78.
28. Thorvaldsdottir H, Robinson JT, Mesirov JP. Integrative genomics viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform. 2013;14(2):178–92.
29. Ghosh S, Chan CK. Analysis of RNA-Seq data using TopHat and cufflinks. Methods Mol Biol. 2016;1374:339–61.
30. Cheng J, Metge F, Dieterich C. Specific identification and quantification of circular RNAs from sequencing data. Bioinformatics. 2016;32(7):1094–6.
31. Sulston JE, Horvitz HR. Post-embryonic cell lineages of the nematode, Caenorhabditis Elegans. Dev Biol. 1977;56(1):110–56.
32. Mitchell DH, Stiles JW, Santelli J, Sanadi DR. Synchronous growth and aging of Caenorhabditis Elegans in the presence of fluorodeoxyuridine. J Gerontol. 1979;34(1):28–36.
33. Mansfeld J, Urban N, Priebe S, Groth M, Frahm C, Hartmann N, Gebauer J, Ravichandran M, Dommaschk A, Schmeisser S, et al. Branched-chain amino acid catabolism is a conserved regulator of physiological ageing. Nat Commun. 2015;6:10043.
34. de Lencastre A, Pincus Z, Zhou K, Kato M, Lee SS, Slack FJ. MicroRNAs both promote and antagonize longevity in C. Elegans. Current biology : CB. 2010; 20(24):2159–68.
35. Essers PB, Nonnekens J, Goos YJ, Betist MC, Viester MD, Mossink B, Lansu N, Korswagen HC, Jelier R, Brenkman AB, et al. A long noncoding RNA on the ribosome is required for lifespan extension. Cell Rep. 2015;
36. Brenner S. The genetics of Caenorhabditis Elegans. Genetics. 1974;77(1):71–94.
37. Lawrence M, Huber W, Pages H, Aboyoun P, Carlson M, Gentleman R, Morgan MT, Carey VJ. Software for computing and annotating genomic ranges. PLoS Comput Biol. 2013;9(8):e1003118.
38. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15–21.

Cortés-López *et al. BMC Genomics* (2018) 19:8

Page 12 of 12

39. Hahne F, Ivanek R. Visualizing genomic data using Gviz and bioconductor. Methods Mol Biol. 2016;1418:335–51.

40. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and cufflinks. Nat Protoc. 2012;7(3):562–78.

41. Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, Fridman WH, Pages F, Trajanoski Z, Galon J. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. Bioinformatics. 2009;25(8):1091–3.