Zebrafish mRNA sequencing deciphers novelties in transcriptome dynamics during maternal to zygotic transition

Håvard Aanes,[1,9] Cecilia L. Winata,[2,9] Chi Ho Lin,[3] Jieqi P. Chen,[4] Kandhadayar G. Srinivasan,[2] Serene G.P. Lee,[2] Adrian Y.M. Lim,[2] Hajira Shreen Hajan,[2] Philippe Collas,[5,6] Guillaume Bourque,[3] Zhiyuan Gong,[7] Vladimir Korzh,[8] Peter Aleström,[1] and Sinnakaruppan Mathavan[2,10]

[1]BasAM, Norwegian School of Veterinary Science, 0033 Oslo, Norway; [2]Stem Cell and Developmental Biology, Genome Institute of Singapore, 138672 Singapore; [3]Computational and Mathematical Biology, Genome Institute of Singapore, 138672 Singapore; [4]Research Computing, Genome Institute of Singapore, 138672 Singapore; [5]Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, 0317 Oslo, Norway; [6]Norwegian Center for Stem Cell Research, 0316 Oslo, Norway; [7]Department of Biological Sciences, National University of Singapore, 117543 Singapore; [8]Laboratory of Fish Developmental Biology, Institute of Molecular and Cell Biology, 138673 Singapore

Maternally deposited mRNAs direct early development before the initiation of zygotic transcription during mid-blastula transition (MBT). To study mechanisms regulating this developmental event in zebrafish, we applied mRNA deep sequencing technology and generated comprehensive information and valuable resources on transcriptome dynamics during early embryonic (egg to early gastrulation) stages. Genome-wide transcriptome analysis documented at least 8000 maternal genes and identified the earliest cohort of zygotic transcripts. We determined expression levels of maternal and zygotic transcripts with the highest resolution possible using mRNA-seq and clustered them based on their expression pattern. We unravel delayed polyadenylation in a large cohort of maternal transcripts prior to the MBT for the first time in zebrafish. Blocking polyadenylation of these transcripts confirms their role in regulating development from the MBT onward. Our study also identified a large number of novel transcribed regions in annotated and unannotated regions of the genome, which will facilitate reannotation of the zebrafish genome. We also identified splice variants with an estimated frequency of 50%–60%. Taken together, our data constitute a useful genomic information and valuable transcriptome resource for gene discovery and for understanding the mechanisms of early embryogenesis in zebrafish.

[Supplemental material is available for this article.]

Zebrafish (*Danio rerio*) has contributed significantly to the understanding of vertebrate development and functional genomics primarily due to the availability of tools and resources on various areas such as mutagenesis (Golling et al. 2002; Alestrom et al. 2006), transgenesis (Stuart et al. 1988; Gong et al. 2001), genomic resources, expression data (Mathavan et al. 2005), and mutant databases (Sprague et al. 2008). Forward genetics of zebrafish has enabled discoveries on developmentally regulated genes (Alestrom et al. 2006); large-scale screens have recovered numerous zebrafish mutants exhibiting clear phenotypes similar to human congenital diseases (Amsterdam and Hopkins 2006). Approximately 30% of zebrafish genes are represented by two copies due to genome duplication (Amores et al. 1998; Postlethwait et al. 1998), and about 14,600 protein-coding genes have been identified by manual annotation in Vega (Wilming et al. 2008), of which about two-thirds of the genes have mammalian homologs (Barbazuk et al. 2000). However, despite these recent advances in understanding zebrafish genomic organization and function, several aspects of gene regulation during development have not been fully characterized, including mRNA splicing patterns and post-transcriptional regulatory events.

Maternally stored mRNAs have been shown to direct early embryonic development prior to activation of the zygotic genome (Newport and Kirschner 1982; Yasuda and Schubiger 1992; Korzh 2009). Although several studies have tried to elucidate the mechanisms regulating maternal transcripts as well as zygotic genome activation (ZGA), most were based on focused analysis of a single transcript or small subset of transcripts or were hindered by the availability of sufficient embryos. Many useful traits of zebrafish make this vertebrate a robust model for large-scale analysis of early developmental events as featured in several recent studies (Giraldez et al. 2006; Lindeman et al. 2010; Vastenhouw et al. 2010). The first cohort of zygotic genome is activated at the 10th cell division (~3.5 h post-fertilization, hpf), known as the mid-blastula transition (MBT) (Kimmel et al. 1995). Several maternal transcripts have been characterized in this model organism (Abrams and Mullins 2009; Lindeman and Pelegri 2010), although little is known about their abundance and the mechanisms regulating their maturation and translation. A role of microRNAs (miRNAs) as mediator of maternal RNA clearance has been documented for miR-430 (Giraldez et al. 2006; Bushati et al. 2008; Takeda et al. 2009); however, to what extent miRNAs regulate

maternal mRNAs in the embryo remains unknown. Moreover, progressive polyadenylation, and hence translational activation, of maternal mRNAs prior to ZGA has also been reported in other organisms (Wilt 1973; Dworkin et al. 1985; Wormington 1993; Kuge and Richter 1995; Wang and Latham 1997; Oh et al. 2000; Aoki et al. 2003; Tadros and Lipshitz 2005). An extensive analysis of transcriptome dynamics during zebrafish pre-MBT and MBT stages would provide valuable insights into the regulation of maternal transcripts and ZGA.

Advances in high-throughput sequencing technologies have had an immense impact on genomics (Green et al. 2010), transcriptomics (Sultan et al. 2008), and stem cell biology (Tang et al. 2010). Speed and accuracy of data generated have made next-generation sequencing a powerful tool to study biological events at the nucleic acid level. Here we apply mRNA deep sequencing (mRNA-seq) to gain comprehensive understanding of transcriptional processes occurring from the unfertilized egg to early gastrulation. Our data provide a wealth of information on transcriptomics before and around the time of ZGA. We report a phenomenon of delayed polyadenylation of maternal transcripts and suggest a major role of this process in developmental events linked to the MBT. We determine, with the highest resolution possible, levels of maternal and zygotic transcripts and identify novel transcripts and splice variants. Our work builds a valuable resource for zebrafish developmental biology, functional genomics, and genome reannotation. This comprehensive analysis will be useful for a wide community who are using zebrafish and other vertebrate models to study early embryonic development.

Results

Mapping and analysis of mRNA-seq reads

To generate a thorough profile of the zebrafish early embryonic transcriptome, mRNA-seq libraries were generated for the following six developmental stages: unfertilized eggs, 1-cell, 16/32-cell, 128/256-cell, 3.5-hpf (MBT, high-oblong), and 5.3-hpf (post-MBT, 50% epiboly) stages (Fig. 1A). These libraries were sequenced (ABI SOLiD; 50-bp tags) to a depth of 25–55 million reads that were mapped to the zebrafish genome assembly version 2010 (Zv9).

About 13–25 million reads could be mapped to the genome (Fig. 1B; Supplemental Table S1) (GEO database; accession number GSE22830), representing 46%–61% of all generated reads. Of these, ~90% mapped to Ensembl annotated regions. With a threshold of >30 reads in a given sample, a total of 11,581 protein-coding Ensembl genes were detected (with gene status "known"). Reads were mapped in a strand-specific manner that allowed precise determination of transcript origin from either the positive or negative strand (Fig. 1C). Reads were mapped with high resolution (Fig. 1D) (tags mapped to 57 exons in *htt* gene) and spanned exon–exon junctions (Fig. 1E). Analysis of tags mapping to 5′ and 3′ untranslated regions (UTRs) of known genes suggested alternative transcription termination and poly(A) sites (resulting in an extended 3′ UTR), or alternative transcription start sites (TSSs; resulting in an extended 5′ UTR) (Fig. 1F). These include well-characterized genes such as *zic3*, *foxa3*, and other less characterized genes (Fig. 1F; Supplemental Fig. S1A).

Statistical testing to detect significant gene expression changes

After normalization of read counts (see Supplemental Methods), we used the R-package DEGseq (Wang et al. 2010) and identified 10,062 genes as differentially expressed between at least two developmental stages (q-value < 0.001, FC > 2, and absolute change > 50). We allowed each gene only to be present once, retaining the observation with the lowest q-value. These observations were unevenly distributed over the pairwise comparisons (Supplemental Fig. S2). We found more than 4000 genes (~40%) that changed most significantly between the 1-cell and 16-cell stages, and about 5500 (~55%) between 3.5 hpf and 5.3 hpf. We also observed about 245 (~0.25%) genes changing most significantly between the 128-cell and 3.5-hpf stages. There was little change between the other stages. This demonstrates a two-step process in which the transition from the 1-cell to the 16-cell stage and 3.5 hpf to 5.3 hpf exhibited the most dramatic (~90%) change in transcriptome profile.

Expression clustering of genes

A total of 5278 Ensembl genes that showed dynamic expressions during early development were included in a clustering analysis (see Supplemental Methods). The analysis generated three "superclusters," designated as maternal, pre-MBT, and zygotic (Fig. 2A), comprising a total of seven distinct subclusters (Fig. 2; Supplemental File S1). Genes in the maternal supercluster (914 genes) exhibited high expression levels in the egg and were subsequently degraded (Fig. 2B). These genes formed two large cohorts, with one starting to decline between the 1-cell and 16-cell stages (Degradation 1; 555 genes) and the other from the 3.5-hpf stage onward (Degradation 2; 359 genes).

The pre-MBT supercluster (3111 genes) genes prominently increased in abundance before 3.5 hpf (Fig. 2C) and could be divided into two subclusters: those declining from 3.5 hpf onward (Pre-MBT1; 2413 genes) and those continuously increasing in abundance (Pre-MBT2; 698 genes).

The zygotic supercluster (1253 genes) genes were activated at 3.5 hpf or later (Fig. 2D) and had three subclusters. The MBT (164 genes) and post-MBT (649 genes) clusters were characterized by very low (less than 20 reads in all samples before the onset of increase) expression at the pre-MBT stages and increased abundance from 3.5 hpf or 5.3 hpf, respectively (Fig. 2D). These represent the presumed earliest cohort of zygotic genes initiated during ZGA in zebrafish. The maternal–zygotic cluster (440 genes) consisted of maternal transcripts, with stable levels during the pre-MBT stages and a subsequent increase at the 3.5-hpf/5.3-hpf stage (Fig. 2D). Collectively, clustering analysis highlights the dynamics of transcript levels during early development.

Validation of expression pattern by real-time RT-PCR

Expression patterns identified by mRNA-seq were validated by real-time RT-PCR (RT-qPCR) analysis of selected genes from each cluster (Supplemental Table S2). To account for possible differences in relative polyadenylation levels across developmental stages, RT-qPCR was performed using both random primers and oligo d(T) primers. In addition, we applied whole-mount in situ hybridization (WISH) to independently validate the expression pattern observed by the other methods (Fig. 3A–E).

mRNA-seq and RT-qPCR data sets correlated well for transcripts in the maternal and zygotic clusters (Fig. 3A–C). Validation of the pre-MBT cluster using oligo d(T) primers matched the mRNA-seq data; however, validation using random primers differed from the pattern observed in mRNA-seq, particularly during pre-MBT stages. Notably, mRNA-seq and oligo d(T) RT-qPCR methods failed to detect the actual levels of transcript abundance
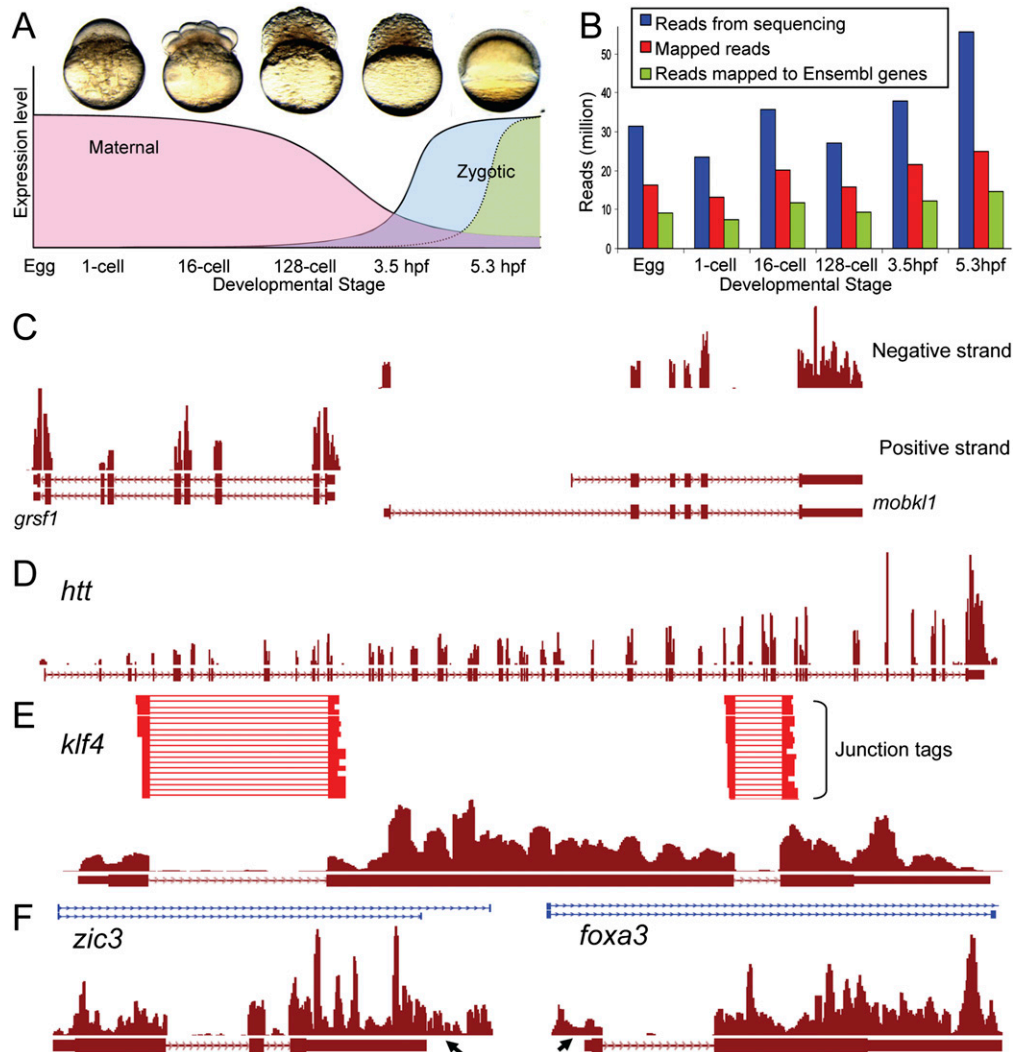
Figure 1. Overview of mRNA-seq data and mapping to the zebrafish genome. (*A*) Zebrafish embryonic developmental stages analyzed in our study. Libraries were generated from eggs, 1-cell, 16-cell, 128-cell, 3.5-hpf, and 5.3-hpf embryos. The graph depicts general expression patterns during these developmental stages. (*B*) Total number of mRNA-seq reads mapped in each developmental-stage library. Differences in the number of generated reads were observed, but the proportions of mapped reads were similar between all stages. (Blue bars) The number of reads generated from sequencing; (red bars) the number of reads mapped to the genome; (green bars) the number of reads mapped to Ensembl-annotated genes. (*C–F*) Mapping of mRNA-seq reads as viewed in the UCSC Genome Browser. Horizontal red bars at the *bottom* of each panel represent annotated exons (tallest boxes), UTRs (half-sized boxes), and introns (lines with arrowheads). (Arrowheads on bars) Orientation of transcript from 5′ to 3′. Mapped reads are represented by red histograms. (*C*) Strand-specific mapping in a region of chromosome 5, showing reads mapping to *mobkl1aa* in the negative, and *grsf1* in the positive strands. (*D*) Reads mapping to the *htt* gene, containing 67 exons. Note the sharp exon–intron boundary of the mapping. (*E*) Mapping of exon–exon junction spanning reads (red horizontal lines) to *klf4* locus, allowing the derivation of splice junctions. (*F*) Mapping of reads to regions beyond annotated 3′ (*zic3*) or 5′ UTRs (*foxa3*), indicated by arrows. Di-tags (blue horizontal bars at *top* of panel) provide support for this mapping.

in the early stages for pre-MBT cluster (Fig. 3D; Supplemental Fig. S3). WISH of several pre-MBT genes confirmed the transcript abundance in early stages for which transcripts could not be detected by mRNA-seq (Fig. 3E; Supplemental Fig. S3; Supplemental Table S3). These observations indicate that although the transcripts of pre-MBT cluster genes are present at the egg/1-cell stage, they were highly under-represented by poly(A) selection-based techniques, suggesting that these transcripts lack sufficient polyadenylation at early developmental stages. Poly(A) tail length measurements for selected pre-MBT transcripts at different developmental stages confirmed the possibility of a delayed polyadenylation mechanism. The pre-MBT transcripts had very short poly(A) tails in the egg and progressively increased their poly(A)

tail length toward 3.5-hpf and 5.3-hpf stages (Fig. 3F; Supplemental Table S4). This suggests that transcripts in this cluster are maternal transcripts that undergo polyadenylation during the pre-MBT stages.

Cytoplasmic polyadenylation of maternal transcripts is necessary for MBT progression

The identification of a large number of polyadenylated maternal transcripts during pre-MBT suggests that polyadenylation of maternal transcripts may play a role in regulating development from the MBT onward. To test this possibility, we inhibited polyadenylation in early embryos by cordycepin treatment (3′-dA)
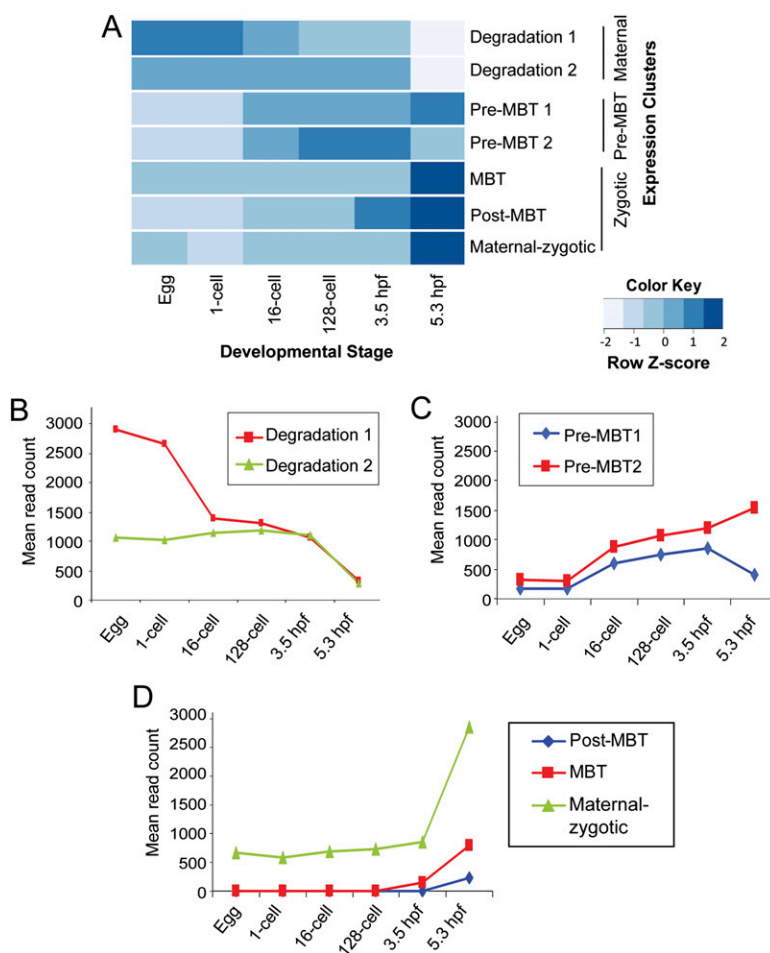
Figure 2. Expression clusters of early developmental genes. (*A*) Heatmap showing distinct expression profiles of different clusters at each developmental stage. Values were scaled for each cluster; color intensity represents expression level relative to its own cluster's average. (*B–D*) Graphs showing mean expression of all genes in each cluster from the egg to the 5.3-hpf stage. (*B*) Two clusters of the maternal supercluster, showing two different degradation patterns. (*C*) Two clusters of the pre-MBT supercluster, both characterized by a significant increase at the 16-cell stage, with either decrease (pre-MBT1) or a continued increase after 3.5 hpf (pre-MBT2). (*D*) Zygotic supercluster consisting of genes activated at 3.5 hpf (MBT) and 5.3 hpf (post-MBT), or maternal–zygotic cluster genes that showed stable accumulation of transcripts during pre-MBT and with a sudden increase at MBT and post-MBT.

(Fig. 4A; Kuge and Richter 1995; Aoki et al. 2003) and confirmed that exposure of the embryos to 3′-dA inhibited polyadenylation (Supplemental Fig. S4). Embryos continuously exposed to 3′-dA from 1-cell stage up to 3.5 hpf were morphologically similar to controls at 3.5 hpf (Fig. 4B,E). However, these embryos showed significant delay in epiboly compared to control at 6 hpf (Fig. 4C,F) and 8 hpf (Fig. 4D,G). In particular, when control embryos were at shield stage at 6 hpf (100%, $n = 21$) (Fig. 4C), treated embryos were at the oblong stage (100%, $n = 19$) (Fig. 4F) and only reached 50% epiboly by 8 hpf (Fig. 4G); these embryos died prior to 10 hpf (data not shown). These results indicate that cytoplasmic polyadenylation of maternal transcripts prior to MBT is critical for normal morphogenesis post-MBT.

Cytoplasmic polyadenylation elements are found in the 3′ UTR of pre-MBT transcripts

It is generally accepted that regulation of polyadenylation is mediated through *cis* elements in the 3′ UTR of mRNAs. We used

programs provided in the MEME Suite (Bailey and Elkan 1994) to find *cis* elements involved in delayed cytoplasmic polyadenylation in pre-MBT transcripts (see also Supplemental Methods and Text).

Through de novo motif discovery we found a long U-rich motif (20-mer) as the most significant (Fig. 4H); this resembles an embryonic CPE (eCPE) reported in *Xenopus* (Simon et al. 1992; Simon and Richter 1994). We tested the frequency of the Hex and CPE motifs in the pre-MBT and two control groups (the maternal–zygotic cluster and a group of randomly chosen genes). Surprisingly, these two elements were present in all, without over-representation in any of the groups. We also looked for the eCPE (a 12-mer U-rich motif). When we allowed one mismatch, 33%, 24%, and 15% of the transcripts harbored this in the pre-MBT, maternal–zygotic, and random group, respectively.

It has been reported that distance between the Hex and CPE motifs affects the timing and extent of polyadenylation (Simon et al. 1992; Simon and Richter 1994; Oh et al. 2000). We found that transcripts in the control groups had a significantly shorter distance between these two motifs ($p < 0.05$ for both comparisons; two-sided Wilcoxon rank-sum test) (Fig. 4I). The maternal–zygotic cluster had a median distance of 67 nt ($p = 2.9 \times 10^{-3}$) and the random group 58 nt ($p = 2.5 \times 10^{-4}$), this in contrast to the median distance of 112 nt in the pre-MBT group (two-sided Wilcoxon rank-sum test) (Fig. 4I). In summary, our results suggest that the distance between the Hex and CPE motifs as well as the presence of the eCPE may influence the regulation of embryonic cytoplasmic polyadenylation.

Functional annotation of gene clusters

To provide biological significance to the mRNA-seq data, we determined the enrichment of biological functions associated with genes in the various clusters according to Gene Ontology (GO) classification (*P*-value < 0.01). Maternal transcripts under Degradation 1 and 2 clusters were generally enriched for genes associated with cell cycle, DNA replication, recombination, and repair functions, as well as intracellular protein trafficking (Supplemental Figs. S5A, S6; Supplemental File S2).

The pre-MBT1 cluster was highly enriched for genes involved in post-translational modification processes, protein transport, protein folding, and localization; many transcripts were also involved in protein degradation (ubiquitin pathway) and DNA replication (Supplemental Fig. S5B). The pre-MBT2 cluster contained genes involved in gene expression, protein synthesis, and RNA
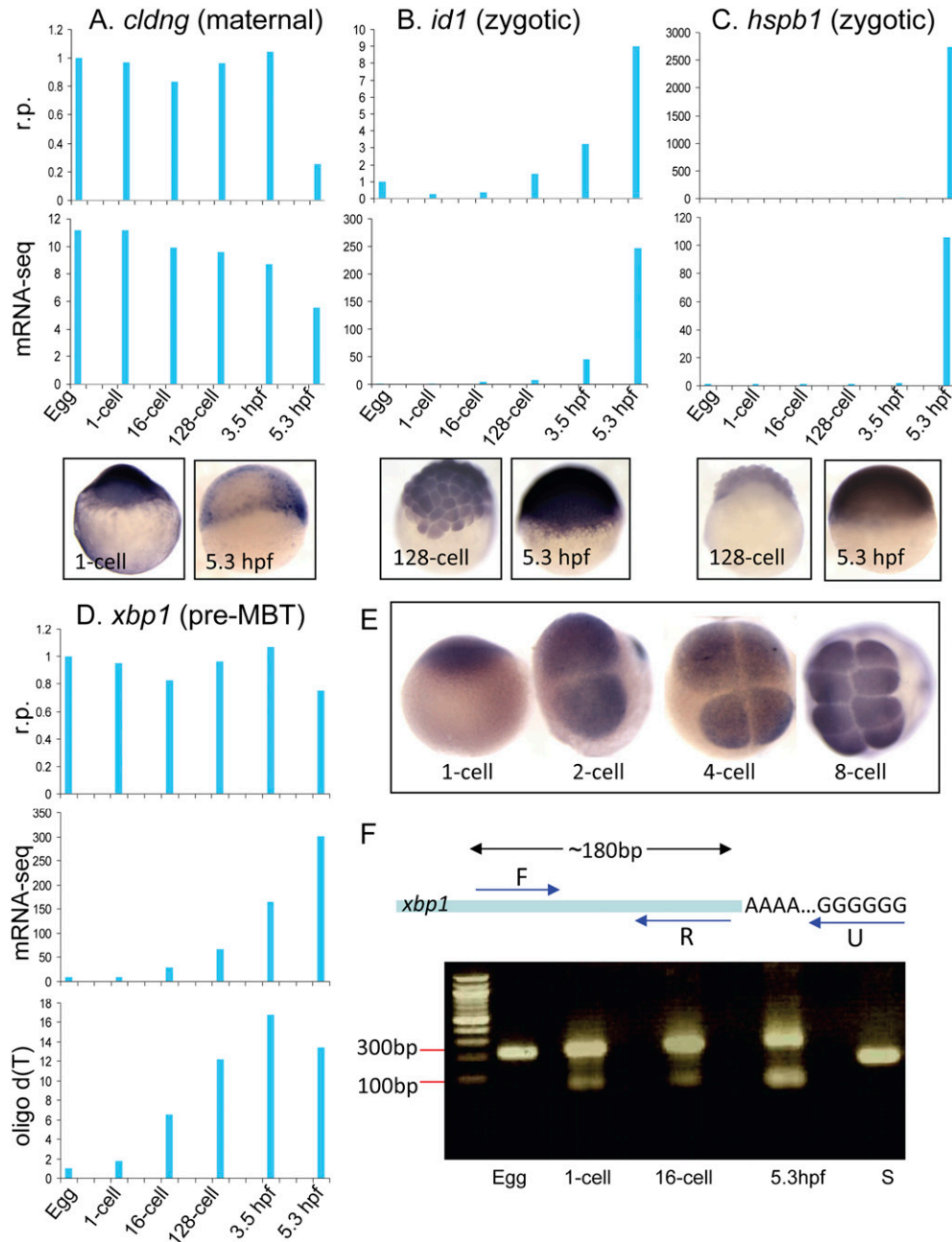
Figure 3. Validation of mRNA-seq expression clusters. (*A–C*) RT-qPCR of the maternal gene *cldng* and the zygotic genes *id1* and *hspb1* show similar expression patterns to those detected by mRNA-seq. WISH confirmed the expression pattern of these genes as observed in mRNA-seq. (*D*) RT-qPCR of the pre-MBT gene *xbp1* showed a different expression profile from that of mRNA-seq when random primers (r.p.) were used, while it showed good correlation when oligo d(T) primers were used. (*E*) WISH of *xbp1* demonstrates the presence of maternal transcript in the egg and the 1-cell stage and confirmed r.p. RT-qPCR data. (*F*) Measurement of *xbp1* poly(A) tail length by RT-PCR in the egg and 1-cell, 16-cell, and 5.3-hpf stages. Total RNA was G/I-tailed and amplified with the *xbp1*-specific forward primer (*F*) and a universal oligo d(C) reverse primer (U). A control RT-PCR with *xbp1*-specific F and R primers gave a 180-bp fragment (S). The increasing poly(A) tail appears as a smear that increases gradually from the 1-cell stage to the 5.3-hpf stage, indicating the presence of an extended poly(A) tail.

post-transcriptional modifications that include cleavage and polyadenylation (Supplemental Figs. S5C, S6; Supplemental File S2).

The MBT cluster displays enrichment for genes associated with signaling pathways linked to development (differentiation, growth and proliferation, cell morphology, and movement), as well as pluripotency and gene expression (Supplemental Figs. S5D, S6; Supplemental File S2). Post-MBT transcripts represented genes involved in organismal and anatomical structure development, cell movement, and cellular development (Supplemental Figs. S5E, S6; Supplemental File S2). IPA analysis revealed 11 transcripts involved in embryonic stem cell pluripotency (*BMP4*, *FGF4*, *FGFR2*, *FZD2*, *FZD3*, *FZD9*, *GSK3B*, *SOX2*, *TGFBR1*, *WNT5B*, *SMAD1*). Genes in the maternal–zygotic cluster were mainly involved in housekeeping functions (Supplemental Figs. S5F, S6).
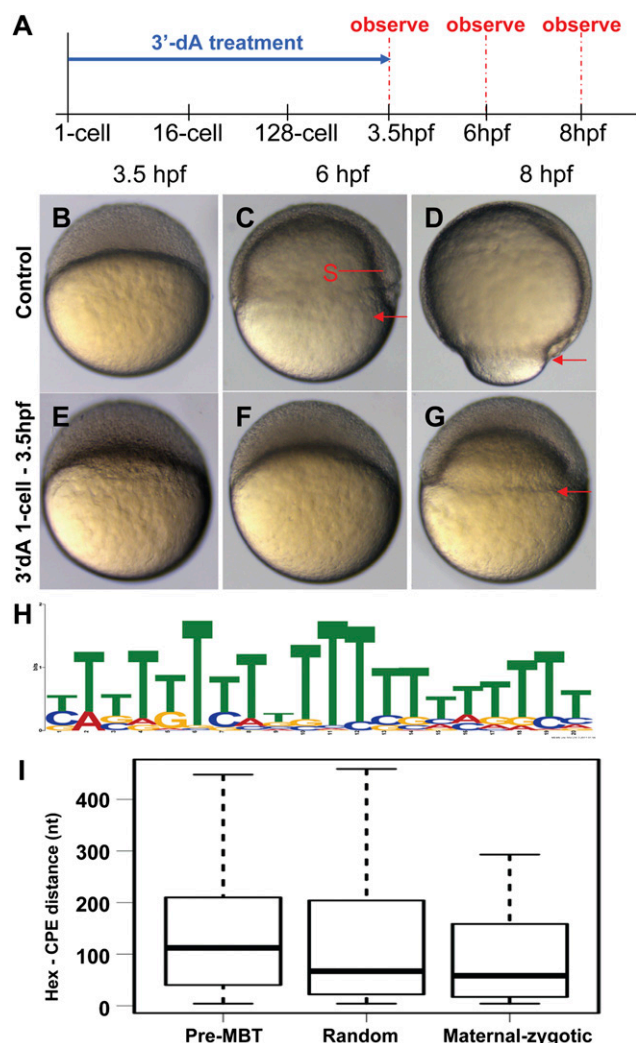
Figure 4. Inhibition of cytoplasmic polyadenylation by 3′-dA. (*A*) Schematic of treatments at various developmental time frames. Horizontal arrows represent exposure to 3′-dA. (*B–D*) Control embryos showing normal development. (Red arrows) The level of marginal cells, signifying the progress of epiboly. Note the presence of embryonic shield (S) at the 6-hpf stage. (*E–G*) Embryos treated with 3′-dA from the 1-cell to 3.5-hpf stage, showing a severe delay in development. (*H*) U-rich motif found in genes of the pre-MBT supercluster using the MEME algorithm. (*I*) Distance between Hex and CPE in different groups of genes. The boxplot shows the distribution of the shortest distance between any Hex and CPE pair in the 3′ UTR of genes in maternal–zygotic and pre-MBT clusters, as well as in a randomly chosen group of genes (random). Notice the longer distance between the two elements in the pre-MBT group compared to the other two groups.

Discovery of novel transcribed regions (NTRs)

A major part of the mRNA-seq reads did not fall within Ensembl-annotated regions. We generated a catalog of novel transcribed regions (NTRs) that mapped to regions of the genome with no previous annotation (see Supplemental Methods). An average of 4067 NTRs were found in each library, across all chromosomes (Supplemental Table S6). NTRs were either found in clusters that are likely to be novel genes or as representation in close proximity to or within annotated transcripts, probably constituting novel/new exons. Genome-wide analysis of the NTRs shows that some of the NTRs

were present at all stages analyzed, while others were present at a subset of the developmental stages. The spatio-temporal distribution of all NTRs is shown in Figure 5A (higher-resolution image in Supplemental File S3). We selected a few examples of NTRs for detailed analysis (Fig. 5B,C; Supplemental Fig. S7). In the gene *ENSDARG00000035540*, 12 NTRs were found 5′ upstream of its annotated exons, 10 of which mapped perfectly to exons of its predicted human homolog, *BRWD1* (Fig. 5B), a gene implicated in spermiogenesis and oocyte–embryo transition (Philipps et al. 2008). This provides evidence for 12 novel exons in zebrafish. PCR validation and junction mapping confirm the novel NTRs as part of the gene. Therefore, *ENSDARG00000035540* has more exons than annotated in Zv9 and is the zebrafish homolog of the human gene *BRWD1* (Fig. 5B). This example demonstrates the usefulness of our data for novel gene discovery in the zebrafish.

The annotated form of the *foxa* gene consisted of two exons (Fig. 5C). We found an additional NTR located 5′ of its annotated start site (bracket in Fig. 5C). In addition, we observed developmentally regulated splice variants based on mRNA-seq reads. The annotated 5′ exon appeared at 3.5- and 5.3-hpf stages and was absent at earlier stages; however, the NTR was present at all the developmental stages (Fig. 5C). Junction mapping analysis revealed that the newly identified NTR formed the first 5′ exon of the *foxa* gene and generated an additional splice isoform at 5.3 hpf (Fig. 5C). Sequence analysis of the NTR showed that it has a start codon, thus confirming that it formed the first exon of the *foxa* open reading frame; this is also supported by di-tag mapping. RT-PCR confirmed the presence of the two splice isoforms at 5.3 hpf, as well as the absence of the 5′ annotated exon in the egg (Fig. 5C).

Analysis of splice variants

To gain a genomic overview of splice variants, we surveyed read counts per exon in a subset of genes (see Supplemental Methods; Supplemental File S4). Of 1591 genes selected, 755–803 (47%–50%) were detected at each developmental stage as alternatively spliced (Supplemental File S5). Of the 16,544 exons tested, 2728 (from 1304 genes) had a Z-score ≤ -2 (constituting 16.4% and 82% of all exons and transcripts tested, respectively). However, only 353 exons (from 289 genes) displayed this attribute at all six stages, 1135 at one stage, and the remainder (1240) at two to five stages. Only a few exons were entirely skipped, with no reads mapping to the exons: This was observed for 167 exons, of which 100 had no reads in any of the six mRNA-seq libraries. The numbers of spliced exons and transcripts were stable across developmental stages (Fig. 6A). Through spatial investigation, we found that most spliced exons were located between other exons, suggesting that they are either part of an exon-skipping event or have alternative splice/donor sites. This constituted 53%–56% of the alternative splicing events (Fig. 6B).

We also did isoform quantification of genes with several annotated isoforms using Cufflinks (Trapnell et al. 2010). Of the genes tested, 1719 genes passed our cutoff (more than 50 reads in each sample), and of these ~60% had a major isoform fraction value <0.8, suggesting that several isoforms are present simultaneously.

Through splice junction mapping, we detected additional splice variants. Junction mapping of *msna* revealed the presence of splice variants in which up to six exons were skipped (Fig. 6C), which is supported by di-tag mapping. We observed reads mapping between exons 1 and 8, 2 and 8, and 3 and 11 (blue arrows). This implies that the zebrafish *msna* has several isoforms present
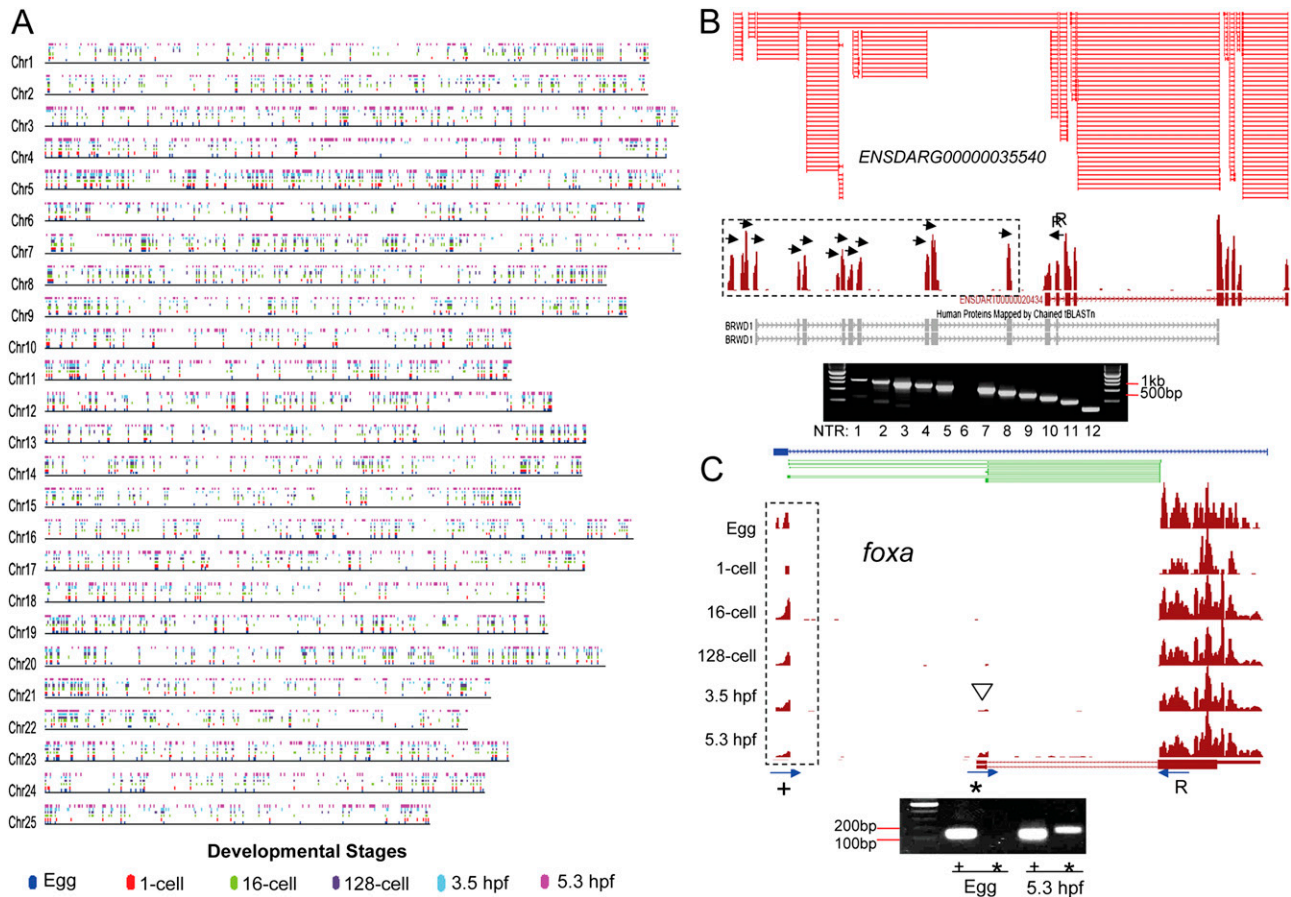
Figure 5. Discovery of NTRs. (*A*) Genome-wide distribution of NTRs plotted at their mapped chromosomal positions. The color code depicts at which stage an NTR was detected. (*B*) Novel exons of *ENSDARG00000035540*, some of which correspond to exons of its human homolog, *BRWD1*. Splice junctions were also identified among these novel exons. RT-PCR performed using reverse primer at annotated exon 2 (R) and forward primers at each of the 12 NTRs (*inset*) showed that at least 11 of these NTRs are expressed and are part of the transcription unit. (*C*) Novel exon in *foxa* found 5' of annotated transcription start site. Note the presence of the 5' annotated exon at 3.5 hpf and 5.3 hpf (white arrowheads). Di-tag (blue horizontal lines) provides support for an isoform that includes the novel exon. Junction mapping at 5.3 hpf (green horizontal lines) indicated the presence of two distinct transcript isoforms. RT-PCR using a reverse primer (R) designed on the 3' distal exon and forward primers on either 5' annotated (*) or NTR (+) confirmed the existence of a new 5' distal exon (*inset*). The 5' annotated exon (*) is skipped in the maternal mRNAs (egg to 128-cell stage).

simultaneously during development. Other examples are described in the Supplemental Text for *tp53* and *gli2a* (Supplemental Fig. S8).

Based on our analysis, we estimate that ~50%–60% of the transcripts are alternatively spliced at each developmental time point examined. Furthermore, most exons seem to be expressed, suggesting several isoforms present simultaneously. Also, it is likely that isoform-specific changes are prevalent (both pertaining to degradation and onset of transcription), as we observed low consistency across stages.

Discussion

Irradiation of fish embryos before MBT has been shown to arrest development at late blastula stages, proving the existence of a mechanism driving development before ZGA (Neyfakh 1956; Newport and Kirschner 1982; Korzh 2009). Subsequent studies in different model organisms identified the role of maternally deposited proteins and mRNAs in this mechanism (St Johnston and Nusslein-Volhard 1992; Kane and Kimmel 1993; Wylie and Heasman 1997). Clearance of maternal mRNAs is thought to

facilitate, and is a prominent feature of, maternal to zygotic transition (Schier 2007; Stitzel and Seydoux 2007). Consistent with this, we have identified three subclusters of maternal RNAs with different degradation kinetics (represented by the degradation clusters of maternal transcripts and the pre-MBT1 cluster). Notably, a high percentage of mRNAs was degraded between 3.5 hpf and 5.3 hpf, a fraction of which has been predicted as miR-430 targets (out of 46 predicted targets, 40 were degraded in our data set) (Giraldez et al. 2006). Since miRNAs have not been detected during the pre-MBT stages (Schier and Giraldez 2006), the large number of transcripts degraded during this period suggest mechanisms other than miRNA-directed decay. In addition, the huge number of transcripts degraded between MBT and post-MBT imply the presence of many more targets for miR-430 than previously thought (Giraldez et al. 2006) or the possibility of involvement of new miRNAs in maternal RNA degradation. Protein degradation in maternal to zygotic transition has also been reported in several organisms (for review, see Stitzel and Seydoux 2007). Consistent with this, analysis of the pre-MBT supercluster revealed strong GO enrichment for post-translational modifications, including the ubiquitin pathway components for protein degradation.
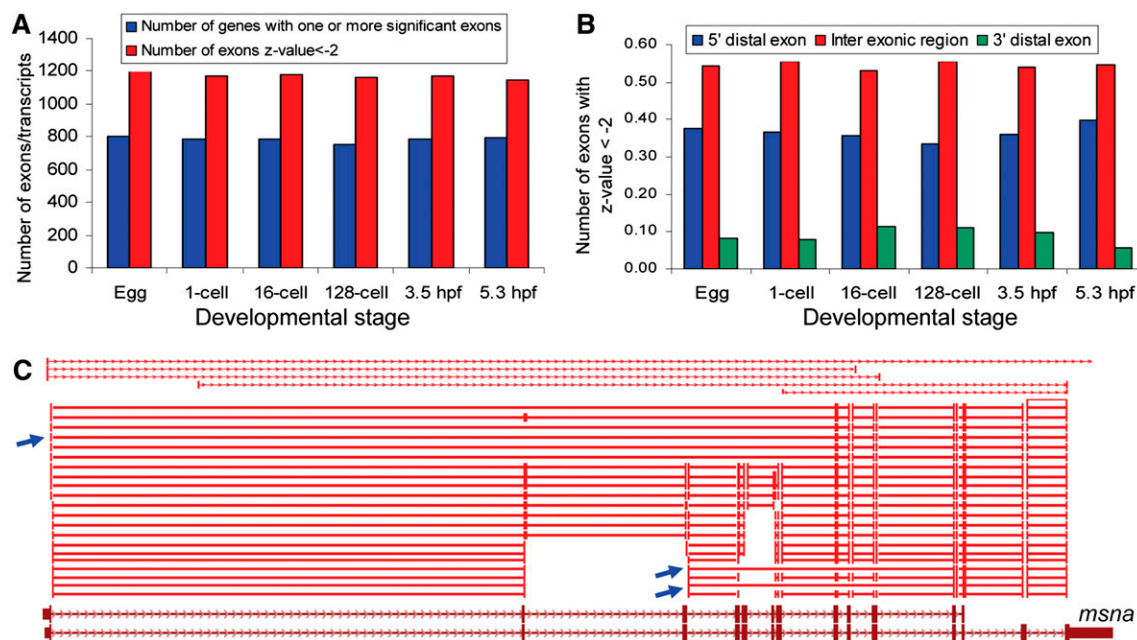
Figure 6. Analysis of splice variants. (*A*) The number of alternatively spliced transcripts (blue bars) and alternatively spliced exons (red bars) at each stage. Both were stable through all stages. (*B*) Location of alternatively spliced exons in all transcripts. Most spliced exons were located between exons, but with a substantial fraction corresponding to distal 5′ exons. (*C*) Splice variants in *msna* at 5.3 hpf. Junction mapping shows skipping of exons 1 to 8, 2 to 8, and 3 to 11 (blue arrows). Di-tags (red bars with arrowheads) support the presence of several splice isoforms.

A substantial fraction (~70%) of maternal transcripts underwent temporal control of their activation through delayed cytoplasmic polyadenylation. This addresses the issue of how maternal mRNAs are regulated prior to ZGA. In the mouse, protein synthesis initiated during the late 1-cell stage is necessary to condition the embryo for activation of zygotic transcription; this process is regulated by cytoplasmic polyadenylation initiated just prior to ZGA (Oh et al. 2000; Aoki et al. 2003). Thus, there is a temporal control for the recruitment of maternal transcripts for translation at later stages. Cytoplasmic polyadenylation has also been associated with translation of maternal transcripts in mouse (Huarte et al. 1987; Vassalli et al. 1989), *Xenopus* (Wilt 1973; Rosenthal et al. 1983; Vassalli et al. 1989; Kuge and Richter 1995), zebrafish (Zhang and Sheets 2009), and *Drosophila* (Salles et al. 1994). Analogous with these studies, our data indicate a substantial increase in transcript abundance between the 1- and 16-cell stages, which is represented by transcripts in the pre-MBT supercluster. Validation analysis by qRT-PCR and poly(A) tail measurement revealed that this increase was not due to de novo transcription of poly(A)+ RNA, but rather reflected an extension of the poly(A) tail of existing maternal mRNAs. Furthermore, results from the 3′-dA experiment strongly suggest that the two cohorts of maternal mRNAs have specific functions. The cohort of maternal RNA, which is polyadenylated prior to fertilization, appears to regulate early cleavage and blastula as inhibition of cytoplasmic polyadenylation did not affect formation of blastomers. The second cohort of delayed polyadenylated maternal mRNAs is involved in functions necessary for progression beyond the MBT.

Cytoplasmic polyadenylation is regulated by *trans*-acting factors assembled at 3′ UTR *cis*-elements (Richter 2007). We observed that several of the known *trans*-acting factors had high read counts in the egg and at the 1-cell stage, with a two-step pattern of substantial reduction, first from the 1-cell to the 16-cell stages, and then

again between the MBT and the post-MBT stages. These included the key gene *zorba* (a *CPEB1* homolog), *tacc3*, and *mos*. The RNA binding protein Dazl, previously shown to be involved in translational activation of dormant transcripts (Padmanabhan and Richter 2006) and to increase polyadenylation of mRNAs (Takeda et al. 2009), had a similar gene expression profile, suggesting that Dazl may have a critical role during the first few cell divisions of embryogenesis.

The pre-MBT supercluster contained several mRNAs encoding proteins that modify DNA, such as DNA methylases (*dnmt3*, *dnmt7*), and histone modifiers, associated both with active and repressive markers (Polycomb, Trithorax groups, and *myst2*, *myst3*, *ehmt1b*, and *hdac3*). A similar expression pattern of Polycomb and Trithorax proteins has been observed in *Drosophila* (Gan et al. 2010) and is suggestive of a flexible epigenetic state. Also present among the pre-MBT transcripts are components of the transcription machinery such as RNA polymerase II and TATA-box-associated factors, including the TATA-box-binding protein (*tbp*) and TBP-associated factor (*taf1a*). Tbp-dependent transcription is necessary for activation of a subset of zygotic genes involved in miR-430 degradation of maternal transcripts (Ferg et al. 2007). Our results suggest that the pre-MBT transcripts are maternal transcripts that are divided into two cohorts—those polyadenylated prior to fertilization and stored as poly(A)+ RNA (~30%) and those with initially short poly(A) tails that are progressively polyadenylated and translationally activated after fertilization (~70%). Thus, we identified a bi-phasic nature of maternal RNA that might share a division of labor at different stages of development—the former during the earlier part of development and the latter containing key regulators of the MBT.

Our splice variant analysis estimates the frequency of multiple splice isoforms of maternal mRNAs in the early embryo to be in the range between 50% and 60%. The lower percentage of splicing compared to humans, where estimated frequencies between 92%

and 95% have been reported (Pan et al. 2008; Wang et al. 2008), may be due to our focused analysis on early embryos, where the identities of most cells are still similar. In addition, the zebrafish represents an intermediate step in the evolution of terrestrial vertebrates, including mammals, which are reported to have many more splice isoforms (Alekseyenko et al. 2007; Keren et al. 2010). Our results extend the recent findings of a high frequency of alternative splicing to zebrafish early development.

In this study, a large number of NTRs were detected for the first time based on mRNA-seq. To date, 16,416 genes have been fully annotated (Vega database) in the zebrafish. Our sequencing identified ~75% of these genes as expressed in early embryos. It also detected an average of 108,216 NTRs in each library, of which an average of 4067 had no supporting information. Temporal and spatial identification of such a huge number of NTRs will improve the current zebrafish genome annotation. Taken together, identification of extended 5′ and 3′ UTRs, detection of NTRs, and splice variants are expected to have a major impact on the understanding of zebrafish genome complexity and annotation.

Conclusion

Our study demonstrates how mRNA-seq can be applied to address the issue of MBT regulation in zebrafish. We unravel delayed polyadenylation of a large cohort of maternal transcripts that might play distinct roles in regulation of developmental events linked to MBT. We also generate valuable genomic resources for further analysis of the early embryonic transcriptome. These results lay the foundation for the discovery and analysis of novel genes, as well as for the identification of new developmental regulatory factors in zebrafish, which are likely to advance our understanding of developmental mechanisms in other vertebrates. Identification of a large number of splice variants implies alternative splicing as a critical factor contributing to the complexity of developmental regulation, and the discovery of a large number of NTRs warrants the need for reannotation of the zebrafish genome.

Methods

Collection of embryos and isolation of total RNA

Wild-type zebrafish from AB background were maintained in the zebrafish facilities of the Genome Institute of Singapore and Institute of Molecular and Cell Biology. Embryos were grown in embryo medium at 28°C, staged according to standard morphological criteria (Kimmel et al. 1995), and harvested at five different developmental stages: 1-cell (0 hpf), 16-cell stage (~1.5 hpf), 128-cell stage (~2.5 hpf), MBT (~3.5 hpf), and post-MBT (~5.3 hpf). Unfertilized eggs were collected by squeezing the abdomen of spawning females, while 1-cell stage embryos were collected ~20 min after spawning. Immediately after harvesting, embryos were snap-frozen in liquid nitrogen and stored at −80°C. Total RNA was extracted from whole embryos using TRIzol (Invitrogen) according to the manufacturer's instructions. RNA concentrations were determined using NanoDrop 2000 (Thermo Scientific). The integrity of RNA samples was determined using an Agilent RNA 6000 Nano chip and size-separated using an Agilent 2100 Bioanalyzer.

Preparation of mRNA-seq library

mRNA-seq was performed by Mission Biotech. SOLiD sequencing libraries were prepared using the Whole Transcriptome Library Preparation for SOLiD Sequencing kit (ABI) according to the manufacturer's instructions. About 200–280 μg of total RNAs was used as starting materials, which were subjected to poly(A) selection using the Applied Biosystems Poly(A) Purist Kit (AM1916), fragmentation, and library construction using distinct adapters for each library (SOLiD Barcoding) (Parameswaran et al. 2007). From each library, equal volumes were pooled together and sequenced in a SOLiD3 (ABI) platform generating 50-bp tags (https://www3.appliedbiosystems.com/cms/groups/mcb_support/documents/generaldocuments/cms_065852.pdf).

Real-time RT-PCR

To verify the expression pattern in our mRNA-seq data, several genes from each expression cluster were tested using real-time RT-PCR in all developmental stages analyzed. Reverse transcription was performed using the High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) with either oligo d(T) or random primers. Real-time PCR was performed using SYBR Green PCR Master Mix (Applied Biosystems) on the ABI 7500 Real-time PCR system according to the manufacturer's instructions (see primer list in Supplemental Table S2). Using the maternal stage as a reference, fold change was calculated by normalizing Ct values in each developmental stage against endogenous control using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen 2001).

Poly(A) tail measurement and cordycepin (3′-dA) treatment

Measurement of poly(A) tails was performed using the Poly(A) Tail Length Assay Kit (USB Corporation). Measurement of poly(A) tails was performed using a gene-specific forward primer (Supplemental Table S4) and a universal reverse primer followed by running the PCR products in agarose gel. For 3′-dA treatment, embryos were placed in glass Petri dishes containing 4 mM 3′-dA (cordycepin {MW = 251.24}, Sigma-Aldrich #C3394) dissolved in embryo medium, and their chorions were partly torn to allow access of the chemical. Observations were made in at least three independent experiments.

Whole-mount in situ hybridization (WISH)

WISH was performed using digoxygenin (DIG)–labeled riboprobes as previously described (Korzh et al. 1998). Digoxigenin-labeled anti-sense probes were generated using clones obtained from the GIS collection (Supplemental Table S3). Images were captured using a Leica M205FA stereomicroscope (Leica).

Data analysis

A full description of the bioinformatics part of the analysis is given in the Supplemental Methods. In short, we mapped mRNA-seq reads using Bioscope version 1.3.1 (Applied Biosystems) and the Whole Transcriptome Analysis Pipeline version 1.2 (WTAP 1.2) to the latest available zebrafish genome assembly (Zv9). Novel transcribed regions (NTRs) were identified by the NTR Finder module in WTAP. Using uniquely mapped reads from Bioscope, read counts per exon were obtained using a Bioscope count module, and read counts per gene were calculated by summarizing the read counts for all exons within a gene. These values were subsequently normalized via a modified normalization strategy previously presented (Robinson and Oshlack 2010) to obtain data adjusted for the amount of poly(A) RNA per embryo. Statistical testing was done using the R-package DEGseq, and clustering was performed on significantly changing genes. For detection of splice variants, we implemented the strategy proposed by Richard et al. (2010) included in the R-package SolaS, and the program Cufflinks v.0.9.3 (Trapnell et al. 2010). Using BioMart (http://www.ensembl.org/

biomart/), we downloaded the 3′ UTRs of transcripts associated with the different genes and applied different programs provided in the MEME suite (Bailey and Elkan 1994).

Data access

The raw sequence data can be accessed from the Gene Expression Omnibus database (accession number GSE22830), and all the genomic resources generated in this study are provided as Supplementary Files.

Acknowledgments

References

Abrams EW, Mullins MC. 2009. Early zebrafish development: it's in the maternal genes. *Curr Opin Genet Dev* **19:** 396–403.

Alekseyenko AV, Kim N, Lee CJ. 2007. Global analysis of exon creation versus loss and the role of alternative splicing in 17 vertebrate genomes. *RNA* **13:** 661–670.

Alestrom P, Holter JL, Nourizadeh-Lillabadi R. 2006. Zebrafish in functional genomics and aquatic biomedicine. *Trends Biotechnol* **24:** 15–21.

Amores A, Force A, Yan YL, Joly L, Amemiya C, Fritz A, Ho RK, Langeland J, Prince V, Wang YL, et al 1998. Zebrafish hox clusters and vertebrate genome evolution. *Science* **282:** 1711–1714.

Amsterdam A, Hopkins N. 2006. Mutagenesis strategies in zebrafish for identifying genes involved in development and disease. *Trends Genet* **22:** 473–478.

Aoki F, Hara KT, Schultz RM. 2003. Acquisition of transcriptional competence in the 1-cell mouse embryo: requirement for recruitment of maternal mRNAs. *Mol Reprod Dev* **64:** 270–274.

Bailey TL, Elkan C. 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol* **2:** 28–36.

Barbazuk WB, Korf I, Kadavi C, Heyen J, Tate S, Wun E, Bedell JA, McPherson JD, Johnson SL. 2000. The syntenic relationship of the zebrafish and human genomes. *Genome Res* **10:** 1351–1358.

Bushati N, Stark A, Brennecke J, Cohen SM. 2008. Temporal reciprocity of miRNAs and their targets during the maternal-to-zygotic transition in *Drosophila*. *Curr Biol* **18:** 501–506.

Dworkin MB, Shrutkowski A, Dworkin-Rastl E. 1985. Mobilization of specific maternal RNA species into polysomes after fertilization in *Xenopus laevis*. *Proc Natl Acad Sci* **82:** 7636–7640.

Ferg M, Sanges R, Gehrig J, Kiss J, Bauer M, Lovas A, Szabo M, Yang L, Straehle U, Pankratz MJ, et al 2007. The TATA-binding protein regulates maternal mRNA degradation and differential zygotic transcription in zebrafish. *EMBO J* **26:** 3945–3956.

Gan Q, Chepelev I, Wei G, Tarayrah L, Cui K, Zhao K, Chen X. 2010. Dynamic regulation of alternative splicing and chromatin structure in *Drosophila* gonads revealed by RNA-seq. *Cell Res* **20:**763–783.

Giraldez AJ, Mishima Y, Rihel J, Grocock RJ, Van Dongen S, Inoue K, Enright AJ, Schier AF. 2006. Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312:** 75–79.

Golling G, Amsterdam A, Sun Z, Antonelli M, Maldonado E, Chen W, Burgess S, Haldi M, Artzt K, Farrington S, et al. 2002. Insertional mutagenesis in zebrafish rapidly identifies genes essential for early vertebrate development. *Nat Genet* **31:** 135–140.

Gong Z, Ju B, Wan H. 2001. Green fluorescent protein (GFP) transgenic fish and their applications. *Genetica* **111:** 213–225.

Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, et al. 2010. A draft sequence of the Neandertal genome. *Science* **328:** 710–722.

Huarte J, Belin D, Vassalli A, Strickland S, Vassalli JD. 1987. Meiotic maturation of mouse oocytes triggers the translation and polyadenylation of dormant tissue-type plasminogen activator mRNA. *Genes Dev* **1:** 1201–1211.

Kane DA, Kimmel CB. 1993. The zebrafish midblastula transition. *Development* **119:** 447–456.

Keren H, Lev-Maor G, Ast G. 2010. Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet* **11:** 345–355.

Kimmel CB, Ballard WW, Kimmel SR, Ullmann B, Schilling TF. 1995. Stages of embryonic development of the zebrafish. *Dev Dyn* **203:** 253–310.

Korzh V. 2009. Before maternal-zygotic transition … There was morphogenetic function of nuclei. *Zebrafish* **6:** 295–302.

Korzh V, Sleptsova I, Liao J, He J, Gong Z. 1998. Expression of zebrafish bHLH genes ngn1 and nrd defines distinct stages of neural differentiation. *Dev Dyn* **213:** 92–104.

Kuge H, Richter JD. 1995. Cytoplasmic 3′ poly(A) addition induces 5′ cap ribose methylation: implications for translational control of maternal mRNA. *EMBO J* **14:** 6301–6310.

Lindeman RE, Pelegri F. 2010. Vertebrate maternal-effect genes: Insights into fertilization, early cleavage divisions, and germ cell determinant localization from studies in the zebrafish. *Mol Reprod Dev* **77:** 299–313.

Lindeman LC, Winata CL, Aanes H, Mathavan S, Alestrom P, Collas P. 2010. Chromatin states of developmentally-regulated genes revealed by DNA and histone methylation patterns in zebrafish embryos. *Int J Dev Biol* **54:** 803–813.

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C(T)}$ method. *Methods* **25:** 402–408.

Mathavan S, Lee SG, Mak A, Miller LD, Murthy KR, Govindarajan KR, Tong Y, Wu YL, Lam SH, Yang H, et al. 2005. Transcriptome analysis of zebrafish embryogenesis using microarrays. *PLoS Genet* **1:** 260–276.

Newport J, Kirschner M. 1982. A major developmental transition in early *Xenopus* embryos: I. Characterization and timing of cellular changes at the midblastula stage. *Cell* **30:** 675–686.

Neyfakh AA. 1956. The changes of radiosensitivity in the course of fertilization in the loach *Misgurnus fossilis*. *Dokl Akad Nauk SSSR* **109:** 943–946.

Oh B, Hwang S, McLaughlin J, Solter D, Knowles BB. 2000. Timely translation during the mouse oocyte-to-embryo transition. *Development* **127:** 3795–3803.

Padmanabhan K, Richter JD. 2006. Regulated Pumilio-2 binding controls RINGO/Spy mRNA translation and CPEB activation. *Genes Dev* **20:** 199–209.

Pan Q, Shai O, Lee LJ, Frey BJ, Blencowe BJ. 2008. Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet* **40:** 1413–1415.

Parameswaran P, Jalili R, Tao L, Shokralla S, Gharizadeh B, Ronaghi M, Fire AZ. 2007. A pyrosequencing-tailored nucleotide barcode design unveils opportunities for large-scale sample multiplexing. *Nucleic Acids Res* **35:** e130. doi: 10.1093/nar/gkm760.

Philipps DL, Wigglesworth K, Hartford SA, Sun F, Pattabiraman S, Schimenti K, Handel M, Eppig JJ, Schimenti JC. 2008. The dual bromodomain and WD repeat-containing mouse protein BRWD1 is required for normal spermiogenesis and the oocyte-embryo transition. *Dev Biol* **317:** 72–82.

Postlethwait JH, Yan YL, Gates MA, Horne S, Amores A, Brownlie A, Donovan A, Egan ES, Force A, Gong Z, et al 1998. Vertebrate genome evolution and the zebrafish gene map. *Nat Genet* **18:** 345–349.

Richard H, Schulz MH, Sultan M, Nurnberger A, Schrinner S, Balzereit D, Dagand E, Rasche A, Lehrach H, Vingron M, et al. 2010. Prediction of alternative isoforms from exon expression levels in RNA-Seq experiments. *Nucleic Acids Res* **38:** e112. doi: 10.1093/nar/gkq041.

Richter JD. 2007. CPEB: a life in translation. *Trends Biochem Sci* **32:** 279–285.

Robinson MD, Oshlack A. 2010. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol* **11:** R25. doi: 10.1186/gb-2010-11-3-r25.

Rosenthal ET, Tansey TR, Ruderman JV. 1983. Sequence-specific adenylations and deadenylations accompany changes in the translation of maternal messenger RNA after fertilization of *Spisula* oocytes. *J Mol Biol* **166:** 309–327.

Salles FJ, Lieberfarb ME, Wreden C, Gergen JP, Strickland S. 1994. Coordinate initiation of *Drosophila* development by regulated polyadenylation of maternal messenger RNAs. *Science* **266:** 1996–1999.

Schier AF. 2007. The maternal–zygotic transition: Death and birth of RNAs. *Science* **316:** 406–407.

Schier AF, Giraldez AJ. 2006. MicroRNA function and mechanism: Insights from zebra fish. *Cold Spring Harb Symp Quant Biol* **71:** 195–203.

Simon R, Richter JD. 1994. Further analysis of cytoplasmic polyadenylation in *Xenopus* embryos and identification of embryonic cytoplasmic polyadenylation element-binding proteins. *Mol Cell Biol* **14:** 7867–7875.

Simon R, Tassan JP, Richter JD. 1992. Translational control by poly(A) elongation during *Xenopus* development: differential repression and enhancement by a novel cytoplasmic polyadenylation element. *Genes Dev* **6:** 2580–2591.

Sprague J, Bayraktaroglu L, Bradford Y, Conlin T, Dunn N, Fashena D, Frazer K, Haendel M, Howe DG, Knight J, et al. 2008. The Zebrafish Information Network: the zebrafish model organism database provides expanded support for genotypes and phenotypes. *Nucleic Acids Res* **36:** D768–D772.

Stitzel ML, Seydoux G. 2007. Regulation of the oocyte-to-zygote transition. *Science* **316:** 407–408.

St Johnston D, Nusslein-Volhard C. 1992. The origin of pattern and polarity in the *Drosophila* embryo. *Cell* **68:** 201–219.

Stuart GW, McMurray JV, Westerfield M. 1988. Replication, integration and stable germ-line transmission of foreign sequences injected into early zebrafish embryos. *Development* **103:** 403–412.

Sultan M, Schulz MH, Richard H, Magen A, Klingenhoff A, Scherf M, Seifert M, Borodina T, Soldatov A, Parkhomchuk D, et al. 2008. A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* **321:** 956–960.

Tadros W, Lipshitz HD. 2005. Setting the stage for development: mRNA translation and stability during oocyte maturation and egg activation in *Drosophila*. *Dev Dyn* **232:** 593–608.

Takeda Y, Mishima Y, Fujiwara T, Sakamoto H, Inoue K. 2009. DAZL relieves miRNA-mediated repression of germline mRNAs by controlling poly(A) tail length in zebrafish. *PLoS ONE* **4:** e7513. doi: 10.1371/journal.pone.0007513.

Tang F, Barbacioru C, Bao S, Lee C, Nordman E, Wang X, Lao K, Surani MA. 2010. Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. *Cell Stem Cell* **6:** 468–478.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28:** 511–515.

Vassalli JD, Huarte J, Belin D, Gubler P, Vassalli A, O'Connell ML, Parton LA, Rickles RJ, Strickland S. 1989. Regulated polyadenylation controls mRNA translation during meiotic maturation of mouse oocytes. *Genes Dev* **3:** 2163–2171.

Vastenhouw NL, Zhang Y, Woods IG, Imam F, Regev A, Liu XS, Rinn J, Schier AF. 2010. Chromatin signature of embryonic pluripotency is established during genome activation. *Nature* **464:** 922–926.

Wang Q, Latham KE. 1997. Requirement for protein synthesis during embryonic genome activation in mice. *Mol Reprod Dev* **47:** 265–270.

Wang ET, Sandberg R, Luo S, Khrebtukova I, Zhang L, Mayr C, Kingsmore SF, Schroth GP, Burge CB. 2008. Alternative isoform regulation in human tissue transcriptomes. *Nature* **456:** 470–476.

Wang L, Feng Z, Wang X, Zhang X. 2010. DEGseq: an R package for identifying differentially expressed genes from RNA-seq data. *Bioinformatics* **26:** 136–138.

Wilming LG, Gilbert JG, Howe K, Trevanion S, Hubbard T, Harrow JL. 2008. The vertebrate genome annotation (Vega) database. *Nucleic Acids Res* **36:** D753–D760.

Wilt FH. 1973. Polyadenylation of maternal RNA of sea urchin eggs after fertilization. *Proc Natl Acad Sci* **70:** 2345–2349.

Wormington M. 1993. Poly(A) and translation: development control. *Curr Opin Cell Biol* **5:** 950–954.

Wylie CC, Heasman J. 1997. What my mother told me: examining the roles of maternal gene products in a vertebrate. *Trends Cell Biol* **7:** 459–462.

Yasuda GK, Schubiger G. 1992. Temporal regulation in the early embryo: is MBT too good to be true? *Trends Genet* **8:** 124–127.

Zhang Y, Sheets MD. 2009. Analyses of zebrafish and *Xenopus* oocyte maturation reveal conserved and diverged features of translational regulation of maternal cyclin B1 mRNA. *BMC Dev Biol* **9:** 7. doi: 10.1186/1471-213X-9-7.