Title:

Maternal transcripts in good and poor quality eggs from Japanese eel, *Anguilla japonica* – their identification by large-scale quantitative analysis

Key words (3-5):

Maternal transcript, development, Japanese eel, qPCR

Author:

HIKARI IZUMI[1¶], KOICHIRO GEN[2], P MARK LOKMAN[3], SEISHI HAGIHARA[1], MOEMI HORIUCHI[1], TOSHIOMI TANAKA[4], SHIGEHO IJIRI[1*], SHINJI ADACHI[1]

1 Division of Marine Life Science, Graduate School of Fisheries Sciences, Hokkaido University, Hakodate, Hokkaido 041-8611, Japan

2 Seikai National Fisheries Research Institute, Fisheries Research Agency, Taira-cho, Nagasaki 851-2213, Japan

3 Department of Zoology, University of Otago, PO BOX 56, 340 Great King Street, Dunedin 9016, New Zealand

4 Hamanako Branch, Shizuoka Prefectural Research Institute of Fishery, Hamamatsu, Shizuoka 431-0214, Japan

¶ Present address:
Institute for Regional Innovation, Hirosaki University, Yanagawa, Aomori 038-0012, Japan

* Corresponding author:
Division of Marine Life Science, Graduate School of Fisheries Sciences, Hokkaido University, Hakodate, Hokkaido 041-8611,

Japan. Tel: (+81) 138-40-5617, Fax: (+81) 138-40-5545
e-mail: ijiri@fish.hokudai.ac.jp

Grant information:
grant sponsor: JSPS KAKENHI, grant number: 21228005
grant sponsor: JSPS Grant-in-Aid for JSPS Fellows, grant number: 13J01576

ABSTRACT

Our understanding of maternal control of development in vertebrates remains incomplete. In this study, we investigated levels of maternal transcripts in good and poor quality eggs from artificially matured Japanese eel, using RNA-Seq and quantitative PCR (qPCR), to identify candidate maternal transcripts related to development. *De novo* assembly or mapping of reads to the eel draft genome yielded 619,029 contigs and 85,906 transcripts, respectively; normalized read counts to these assemblies were calculated using reads (RPKM) or fragments (FPKM) per kilobase of transcript per million mapped reads. *In silico* screening identified 1594 contigs and 150 transcripts with lower RPKM or FPKM in poor than in good quality eggs, 245 contigs and 85 transcripts of which could be annotated by BLASTx, respectively. From selected contigs or transcripts, 6 genes (*dnajb4*, *gnpat*, *card14*, *pdp1*, *fcgbp*, *ttn*) had significantly lower mRNA levels in poor than in good quality eggs by qPCR. Multiple regression analysis showed that 5 genes (*gnpat*, *b4galnt1*, *acsl6*, *rtkn*, *trim24*) significantly correlated with hatchability. Taken together, 10 genes were identified as candidate maternal transcripts, regulating development in Japanese eel. Our results contribute to understanding the molecular basis for maternal control of development in vertebrates.

1. INTRODUCTION

In vertebrates, maternally inherited transcripts (maternal transcripts) accumulate in the egg but are translationally suppressed during oogenesis. Upon fertilization, these transcripts are promptly translated to regulate early embryonic development in the transcriptionally quiescent zygote. Subsequently, a set of maternal transcripts is eliminated, coinciding with activation of *de novo* transcription from the genome of the zygote, referred to as zygotic genome activation (ZGA). This transfer of the control of development from products of maternal origin to those of an activated zygotic genome is known as the maternal-zygotic transition (MZT) (reviewed in Tadros & Lipshitz, 2009; Lee et al., 2014). A recent report has demonstrated that some maternal products are required to initiate the zygotic developmental program (Lee et al., 2013). Additionally, several maternal products continue to function into developmental stages beyond ZGA (see review in Marlow, 2010). Thus, maternal transcripts are essential not only for early, but also for late development.

Maternal-effect mutations make for a powerful tool to help dissect the maternal control of vertebrate embryogenesis (Dosch et al., 2004; Wagner et al., 2004; Pelegri et al., 2004). Accordingly, numerous maternal-effect mutants with morphological defects have been identified in zebrafish, a model vertebrate species that is highly suitable for study of developmental biology. Such studies have contributed to understanding of maternal control of development. Take, for example, maternal *lymphoid-restricted membrane protein* (*lrmp*), which plays an essential role in nucleus-centrosome attachment and pronuclear congression during fertilization. *Lrmp* has been identified from *futile cycle* zebrafish mutants which fail to undergo pronuclear congression and fusion (Lindeman & Pelegri, 2012). Another maternal transcript, *serine-threonine kinase mitogen activated protein kinase-activated protein kinase 2* (*mapkapk2*), which regulates progression of epiboly (gastrulation-specific movement of cells leading to an epithelial monolayer over the entire yolk mass) through modulation of the actin cytoskeleton, has been identified from *betty boop* mutants that display defects in epiboly (Holloway et al., 2009). Although several causative genes of maternal-effect mutants have accordingly been identified, in many cases their molecular identity remains unknown, so that understanding of the mechanisms for maternally controlled aspects of development in vertebrates remains incomplete.

Japanese eel, *Anguilla japonica*, is one of the most important aquaculture species, and

establishment of artificial propagation techniques is required to stably supply seedlings and conserve this natural resource in Japan (Tanaka, 2015). In captivity, fertilized eggs are obtained by exogenous hormone treatment, but the quality of eggs produced under these conditions is variable (Chai et al., 2010) – not unexpectedly, abnormal embryos appear at high frequencies from batches of poor quality eggs. It is empirically known that the phenotype of embryos produced from poor quality eel eggs resemble those of maternal-effect mutants in zebrafish (ex. Dosch et al., 2004; Wagner et al., 2004; Pelegri et al., 2004). In view of these observations, the study of maternal transcripts in poor quality eggs from Japanese eel could prove an effective novel approach to understanding the role of maternal transcripts during vertebrate development.

Previously, we investigated the maternal mRNA levels of 74 cell division-related genes in eggs from artificially matured Japanese eel (Izumi et al., 2016) and documented that the abundance of several of these mRNA species correlated with egg quality. Similarly, other studies on fish have identified differentially expressed maternal transcripts between good and poor quality eggs (reviewed in Sullivan et al., 2015; Zarski et al., 2017). Together, these reports have indicated that levels of accumulated maternal transcripts in the egg were associated with development competence. Therefore, it is plausible that maternal transcripts that are present in abnormal composition or abundance in poor quality eggs play an essential role in development.

In this study, we aim to identify candidate maternal transcripts that play a role in development in eggs from Japanese eel on the basis of their quantitative signature. We focused on transcripts that had lower abundance in poor than in good quality eggs. First, we subjected RNA from Japanese eel eggs of good and poor quality to next-generation sequencing (NGS) and then conducted a large-scale RNA-Seq analysis by comparing the number of mapped reads in egg samples to either *de novo* assembled contigs or to the eel draft genome sequence. We complemented these analyses by comprehensive quantitative PCR (qPCR) assays. Finally, we searched for maternal transcripts whose abundances were lower in poor quality eggs compared to eggs of good quality, as reflected in hatchability.

2. MATERIALS AND METHODS

2.1 Animals – methods, sampling and ethical conduct

Details on the source, maintenance and experimental manipulations of Japanese eels used in this study have been described in detail previously (Izumi et al., 2016); different subsamples from the same experiment were used in the present study. In brief, wild-caught elvers from Japanese eel were purchased and reared before being feminized by female sex hormone in experimental tanks. After a further two years of on-growing, feminized eels (500–1000 g) received weekly injections of exogenous pituitary homogenates to attain sexual maturity, as described in Chai et al. (2010). After 17,20β-dihydroxy-4-pregnen-3-one-induced ovulation (Kagawa et al., 1997), eggs from 13 artificially matured female Japanese eels were collected by strip-spawning and subsamples flash-frozen (n=12) in liquid nitrogen or preserved in RNALater (Thermo Fisher Scientific, Waltham, MA; n=1) until analysis. Parts of the egg batches (2 g) were dry-fertilized and incubated; fertilization and hatching rates were scored and these data used to consider batch quality as “good” or “poor” (see Table 1), exactly as described in Izumi et al. (2016). Eggs with fertilization rate >80% and hatching rate >80% were classified as good quality eggs (developmentally competent); eggs with fertilization rate >80% and hatching rate ≤20% were classified as poor quality eggs (eggs producing viable morula embryos without hatching). All fish handling, husbandry and sampling methods were approved by the Institutional Animal Care and Use Committee of Hokkaido University.

2.2 RNA extraction and removal of ribosomal RNA for NGS

Total RNA was isolated from 6 egg samples (0.2–0.4 g per sample), 3 representing good eggs (Eel nos. 1, 2, 5; Table 1) and 3 representing eggs of poor quality (Eel nos. 10, 12, 13), after homogenization in ISOGEN reagent (Nippon Gene, Tokyo, Japan). Any remaining genomic DNA was digested on-column with DNase and RNA eluted using the Nucleospin RNA II Kit (Macherey-Nagel GmBH, Düren, Germany) in accordance with the manufacturer’s instructions. RNA quantity and quality were subsequently evaluated using a Nano-Drop 2000 (Thermo Fisher Scientific) and an Agilent RNA Pico chip, run on a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) – resulting RNA integrity numbers exceeded 8.0 for all samples.

148 Selective depletion of ribosomal RNA was achieved as described previously (Izumi et
149 al., 2015; in Japanese): using Thermo Fisher Scientific's RiboMinus™ Eukaryote System v2,
150 > 98% rRNA was removed after incubation of total RNA from eel eggs with biotinylated
151 antisense probes to mammalian rRNA.
152
153 **2.3 Library construction and sequencing**
154 Template RNA was quality-assessed on a 2100 Bioanalyzer to confirm the successful
155 removal of rRNA (rRNA abundance < 2% of that in non-depleted samples, data not shown;
156 also see Izumi et al., 2015). The Ovation® RNA-Seq System (NuGEN, Redwood City, CA)
157 was then used to synthesize double-stranded cDNA. Paired-end sequencing of the cDNA
158 libraries was performed using a HiSeq 2000 sequencer (Illumina, San Diego, CA), contracted
159 to Hokkaido System Science Co Ltd, according to the manufacturer's instructions.
160
161 **2.4 Assembly and mapping**
162 CASAVA version 1.8.2 (Illumina) was used for base calling, read filtering and cleaning of
163 sequence data. Adapters and low quality bases were trimmed in CLC Genomics Workbench
164 (Qiagen, Valencia, CA) with default-set parameters. To reduce bias, the trimmed reads were
165 mapped to rRNA and mitochondrial genome sequences (GenBank Accession No.
166 CM002536.1) using CLC Genomics Workbench as described in Izumi et al. (2015);
167 unmapped-reads were used for subsequent analyses.
168 Two approaches, i.e., *de novo* assembly (Analysis 1) and reference-based assembly
169 (Analysis 2), were performed in keeping with Martin & Wang (2011). Figure 1 shows a flow
170 diagram of the experimental design. In Analysis 1, unmapped-reads shorter than 50 b were
171 discarded and the longer reads were used for the *de novo* assembly to obtain contigs. *De*
172 *novo* assembly was carried out using CLC Genomics Workbench at default settings.
173 Thereafter, the reads from each of the 6 libraries were mapped to the *de novo* assembled
174 contigs using CLC Genomics Workbench and converted into reads per kilobase of exon per
175 million mapped reads (RPKM; Mortazavi et al., 2008). In Analysis 2, unmapped-reads from
176 all samples were mapped to the draft genome sequence of the Japanese eel (Henkel et al.,
177 2012) using TopHat (Trapnell et al., 2009), and the relative levels of each transcript were then
178 calculated using the *cufflinks*, *cuffmerge* and *cuffdiff* commands in Cufflinks on the basis of

FPKM (fragments per kilobase of transcript per million mapped reads), as described in Trapnell et al. (2012).

2.5 *In silico* screening

To identify transcripts that had lower abundance in poor quality than in good quality eggs, we carried out *in silico* screening as our first approach, and then followed this up with qPCR. The screening design that was used in this study is described in Fig 1. In Analysis 1, contigs whose RPKM values were significantly down-regulated (at least two-fold; $P < 0.05$) in poor eggs (nos. 10, 12, 13) compared to eggs of good quality (nos. 1, 2, 5) were identified. In addition, we identified contigs (> 200 bp in length) of which no reads were found among poor egg quality samples (RPKM = 0), whereas RPKM > 0 among the three batches of good eggs. These sequences were subjected to BLAST (blastx) against NCBI database (nr) using BLAST2GO software (Götz et al., 2008) and were abbreviated using the gene names used in accord with the Human Genome Organization Gene Nomenclature Committee (HGNC). To derive the potential function of a gene, individual contigs were manually annotated with the Gene Ontology (GO) term (biological process) in ZFIN (zebrafish database) and UniProtKB (human database). The GO term with the highest-ranked GO Evidence Code and involvement in embryogenesis was assigned to the respective contig.

In Analysis 2, we identified those transcripts for which FPKM from poor quality eggs = 0 and that from good eggs > 0. Additional stringent differential FPKM selection criteria were imposed, such that *i)* good egg FPKM > 1 and *ii*) the ratio between FPKM of good eggs (lowest FPKM from among No. 1 and No. 2) and poor quality eggs (using the best sample amongst these, i.e, No. 10) < 0.65. These were annotated by BLASTx (Ref-Seq), and their potential function was presumed to be the same as described in the above methods.

2.6 qPCR analysis

To reduce the number of candidate genes to a manageable number for follow-up qPCR analyses, we performed small-scale qPCR analyses (n=3) against the contigs identified by *in silico* screening in Analysis 1. For this purpose, cDNA was synthesized from rRNA-depleted RNA (i.e., that used for NGS) and subsequently amplified using the Ovation® Pico WTA System V2 (NuGEN).

For large-scale qPCR analysis (n=6 for each egg quality category), total RNA (nos.1–5, 7–13 in Analysis 1 and nos. 1–4, 6–13 in Analysis 2) was reverse-transcribed into single-stranded cDNA with random hexamer primers as described in Izumi et al. (2016). Genes that were identified by *in silico* screening and small-scale qPCR as being differentially expressed were analyzed by qPCR using THUNDERBIRD SYBR® qPCR Mix (Toyobo, Osaka, Japan) or Thermo Fisher Scientific's Universal SYBR® Select Master Mix. Primers used for qPCR analyses and designed using the software package Primer Express v3.0.1 (Thermo Fisher Scientific) are listed in Supporting Information, Table S1, S2. All qPCR data were normalized over expression of β-actin (Izumi et al., 2016) using the ΔCt method (Pfaffl, 2001).

2.7 Statistical analysis

Hierarchical clustering using Ward's method and Euclidean distance was conducted using the RPKM or FPKM data from three good quality egg samples (nos. 1, 2, 5) and three poor quality egg samples (nos. 10, 12, 13) used for NGS. Expression data from small-scale experiments (n=3), both from qPCR and from *in silico* analysis (only Analysis 1), were analyzed with Student's *t*-test; differences were considered significant for $P < 0.05$. Mann-Whitney *U*-test was employed to analyze qPCR data from experiments with larger sample sizes (n=6), and differences were considered significant for $P < 0.01$. The Benjamini-Hochberg false discovery rate (FDR = 0.3) was adopted for multiple testing correction of qPCR data in Analysis 2. Cluster analysis and comparison of treatment means were done using R version 3.1.0. To determine whether maternal transcripts contributed to hatchability, stepwise multiple regression analysis was carried out using the Statistical Package for Social Sciences (IBM). The dependent variables were the transcript abundances for 36 and 72 genes quantified by qPCR analysis (n=6) in Analyses 1 and 2, respectively. The hatchability was used as independent variable. Predicted hatchability was subsequently estimated by multiple regression equations for Analyses 1 and 2, respectively. The value of qPCR analysis (n=6) was again used as dependent variable. Average of the calculated hatchability value was defined as the final predicted hatchability value. To test for relationships between the predicted hatchability and observed hatchability, Spearman's correlation coefficient was estimated using R. Significance was defined as $P < 0.05$.

3. RESULTS

3.1 NGS statistics and cluster analysis

An average of around 60 million reads, not including those for rRNA and mtRNA, of nearly 100 bp in length was obtained for each library (35.53 Gb of sequence data; see Supporting Information, Table S3). These reads were *de novo* assembled into 619,029 contigs. The maximum contig length was 43,411 bp, the N50 equated to 296 bp, the average length to 260 bp and the approximate coverage depth was 81.6. When these reads were mapped to the draft eel genome - approximately 70% of all reads from all egg samples mapped to this genome – 85,906 transcripts were obtained.

In Analysis 1, the dendrogram derived from cluster analysis of relative levels of all transcripts showed that good and poor quality eggs clustered together (Fig. 2 a). In contrast, clustering on the basis of Analysis 2 produced two distinct clusters, one consisting of good quality eggs (nos. 1, 2, 5) and the other of poor quality eggs (nos. 10, 12, 13) (Fig. 2 b).

3.2 Screening for qPCR analysis

Normalized read counts (RPKM) of 994 contigs were at least two-fold higher in good than in poor quality eggs from Japanese eel. When listing all contigs that yielded mapped reads in good eggs and no reads in batches of poor quality eggs, a total of 600 contigs was identified. Of these contigs, 148 and 97 contigs could be annotated by BLASTx, respectively (Supporting Information, Table S4, S5); however, 10 and 11 contigs, respectively, coded for unknown protein. Altogether, primers could be designed for 152 contigs for subsequent qPCR screening on reverse-transcribed total RNA from the same 6 samples as those used for NGS. Accordingly, the mRNA levels of 26 among these 152 candidate genes were significantly lower in poor quality relative to good quality eggs ($P < 0.05$) (Table 2). A further 10 candidate genes tended to be differently expressed ($0.059 < P < 0.168$). The resulting set of 36 candidate genes (Table 2) was analyzed for subsequent qPCR analysis on 6 batches of good (Nos. 1–5, 7) and poor quality eggs (Nos. 8–13).

When adopting the presence-absence approach of transcripts in good vs poor quality eggs against the draft Japanese eel genome, 150 transcripts, of which 85 could be annotated, were identified (Table 3). Three transcripts coded for unknown protein. In Analysis 2, 72 transcripts were selected as target for subsequent qPCR analysis (n=6).

3.3 Transcripts differentially expressed in the good and poor quality egg groups

Contig Nos. 107048, 71472, 64061, 112812, 156849, annotated as *card14*, *pdp1*, *gnpat*, *dnajb4* and *fcgbp*, respectively (see Table 2), were the only 5 among 36 contigs from Analysis 1 that displayed significantly lower mRNA levels in poor quality eggs than in good quality eggs ($P < 0.01$) (Fig 3).

From the list of 72 transcripts from Analysis 2, only one transcript (transcript No. 47766, annotating *ttn*; see Table 3) was expressed at a significantly lower level in poor quality eggs as compared to good quality eggs ($P = 0.0039$, with false discovery rate < 0.3) (Fig 4).

3.4 Transcripts contributed to hatchability

Stepwise multiple regression analysis of expression data from genes listed in Tables 2 and 3 was conducted to identify genes that could influence hatchability. In Analysis 1, contig No. 64061 annotated as *gnpat* and contig No. 47139 annotated as *b4galnt1* were positively associated with hatchability and the correlation coefficient (R^2) was 0.811 ($P < 0.001$; Table 4). *Gnpat* in particular had a strongly influence on hatchability, reflected in the standardized coefficient (B) of *gnpat* being higher than that of *b4galnt1*. In Analysis 2, hatchability was positively correlated with expression of 3 genes, i.e., transcript Nos. 8852, 40658 and 69150, annotated as *acsl6*, *rtkn* and *trim24*; the correlation coefficient (R^2) was 0.921 ($P < 0.001$; Table 4). The strongest predictor of hatching was *acsl6* expression.

The multiple linear regression models to quantify the relationship between hatchability and transcripts are listed as follows:

Analysis 1: Hatchability (%) = -74.17 + 87.43 (*gnpat* mRNA level) + 70.01 (*b4galnt1* mRNA level)

Analysis 2: Hatchability (%) = -85.94 + 63.94 (*acsl6* mRNA level) + 46.57 (*rtkn* mRNA level) + 57.08 (*trim24* mRNA level)

Observed and predicted hatchability values calculated from the above two models are

303 presented in Fig. 5 and proved to be strongly and positively correlated (R = 0.986, $P < 0.001$).
304
305 4. DISCUSSION
306 In the present study, we investigated levels of maternal transcripts in good and poor quality
307 eggs (those producing a viable morula embryo without hatching) from artificially matured
308 Japanese eel using RNA-Seq and qPCR analysis. Accordingly, we aimed to identify maternal
309 transcripts that may be related to development.
310
311 **Experimental approach**
312 In this study, two RNA-seq approaches were used in keeping with Martin & Wang (2011);
313 reads were mapped to *de novo* assembled contigs or to the eel draft genome as reference
314 sequence. Notably, the number of contigs obtained from *de novo* assembly was greater than
315 the number of transcripts obtained from reference-based assembly. This suggests that the
316 former generates a number of redundant contigs due to miss-assembly or to incomplete
317 sequencing coverage. Although *de novo* assembly strategies tend to generate a number of
318 redundant contigs due to sequence variations, which affect downstream experiments, *de*
319 *novo* transcriptome analysis can recover transcripts that are transcribed from missing
320 segments of the draft genome (Martin & Wang, 2011). Indeed, the published eel genome
321 alignment (Henkel et al., 2012) was robust, but about 30% of reads were unmapped to the
322 genome sequence. Thus, we carried out both approaches. The relative levels of each
323 transcript were expressed as RPKM/FPKM in this study. It is later reported that RPKM can
324 be inconsistent among samples (c.f., Wagner et al., 2012). It is further noted that statistics
325 do not take batch effects (c.f., Leek et al., 2010) into consideration and therefore, such effects
326 may be present in our RNA-Seq analyses. Regardless, our eventual qPCR analyses
327 identified several transcripts that were significantly less abundant in poor than in good quality
328 eggs. In order to gain a more detailed understanding, or further reinforce the molecular
329 quantitative signature, of good *versus* poor quality eggs, increasing of the sample sizes and
330 additional bioinformatic analysis of RNA-Seq are needed.
331
332 **Transcripts differentially expressed in good and poor quality egg groups**
333 Upon mapping reads to the *de novo* assembled transcriptome, good and poor quality egg

groups clustered together using hierarchical clustering analysis. It is very possible that contig redundancy affected the clustering analysis. Although several approaches for removal of redundant contigs have been proposed (Haznedaroglu et al., 2012; Ono et al., 2015), these were not employed in this study. Instead, we used a reference-based approach to analyse transcriptomes and observed that good and poor quality egg groups clearly separated into two clusters. Reference-based transcriptome analysis clearly proved more effective, that is, the result of clustering shows that the maternal transcript profile significantly differed between good and poor quality eggs. The drivers for differences in egg quality between female broodstock, which were all reared under the same artificial conditions, are unknown. Regardless, this quantitative signature of maternal transcripts in unfertilized eggs from artificially matured Japanese eel appear to be associated with development during the period from morula to hatching.

Our analyses identified a large number of genes that had lower abundance in poor than in good quality eggs by *in silico* screening. Many of these *in silico*-identified differentially expressed genes are likely to reflect developmental competence – indeed, a leading study by Chapman et al. (2014) on striped bass clearly highlights the complex nature of developmental competence and the notion that a single or a few genes governing this trait does not withstand scrutiny. This is also evident from our qPCR-derived gene expression dataset, in which we routinely noted 1–2 outliers among the 12 assayed samples – the outlying sample(s) were inconsistent, i.e., different egg batches served as outlier for different genes, which *i)* made it difficult to detect differences in expression between batches, and which *ii)* reinforced the notion that developmental competence is a complex trait. Nonetheless, the number of genes governing developmental competence may perhaps be reduced to a subset of genes with high predictive value (c.f. Chapman et al., 2014; Sullivan et al., 2015).

Genes that were differentially expressed *in silico* in the present study were initially assayed by qPCR at low replication (n=3), before subjecting a subset to higher replication (n=6). Accordingly, we detected lower levels of 6 maternal transcripts (*dnajb4*, *gnpat*, *card14*, *pdp1*, *fcgbp*, *ttn*) in poor quality eggs as compared to good quality eggs. This result suggests that the 6 maternal transcripts appear to be associated with development during the period from morula to hatching. Although several candidate maternal transcripts related to

developmental competence were identified in some aquaculture species by comparing mRNA levels between good and poor quality eggs/embryos (reviewed in Sullivan et al., 2015), the 6 transcripts represent novel candidate genes.

Dnajb4, dnaJ heat shock protein family (Hsp40) member B4, is involved in interaction with the actin cytoskeleton (Chen et al., 2010). Based on observations on zebrafish maternal mutants, regulation of the actin cytoskeleton is important for progression of epiboly (Holloway et al., 2009). Fcgbp, Fc fragment of IgG-binding protein, was implicated in the epithelial-mesenchymal transition, a cell conversion process crucial for cell movement during gastrulation (Xiong et al., 2014, review in Thiery et al., 2009). Although phenotypic analyses of eel embryos produced from poor quality egg was not conducted, previous reports render it possible that dnajb4 and fcgbp are involved in gastrulation by regulating cytoskeletal function in eel. Card14, caspase recruitment domain family member-14 can activate NF-κB (Bertin et al., 2001). In the zebrafish embryo, blocking NF-κB activity affects mesoderm development, resulting in embryos failing to undergo gastrulation or becoming dorsalized with partial or complete lack of tail formation (Correa et al., 2004). Thus, card14 is likely to be essential for normal development in eel. Pdp1, pyruvate dehydrogenase phosphatase catalytic subunit 1, is a component of the pyruvate dehydrogenase complex which plays a crucial role in energy metabolism (Wieland, 1983). Oxygen consumption increases as gastrulation progress in fish (Boulekbache, 1981), suggesting that insufficient *pdp1* mRNA accumulation may negatively impact gastrulation. In knockout mice, it was reported that Gnpat, glyceronephosphate O-acyltransferase, plays a role in development of eye and nervous system (Rodemer et al., 2003). Observations on zebrafish have suggested that some *ttn* exons are required for sarcomere assembly in the heart and somites (Seeley et al., 2007). How these transcripts mechanistically affect development during the period from morula to hatching remains a matter of speculation – regardless, our results suggest that sufficient amounts of these transcripts in unfertilized eggs are required to ensure hatching success in eel.

Transcripts contributing to hatchability

Using multiple regression analysis, 5 transcripts (*gnpat*, *b4galnt1*, *acsl6*, *rtkn*, *trim24*) positively and significantly correlated with hatchability. This result suggests that two (*gnpat*

and *b4galnt1*) or three transcripts (*acsl6*, *rtkn*, *trim24*) co-operate to facilitate normal development and could contribute to hatchability.

Accordingly, gnpat was again highlighted for its association with hatchability. So too, was *B4galnt1,* which encodes β-1,4-N-acetyl-galactosaminyltransferase 1, an enzyme involved in the biosynthesis of glycosphingolipid essential for normal neural development and function in zebrafish (Boccuto et al., 2014). A role for acsl6 in patterning is also possible, given the recent report that demonstrated that acsl4a, a protein closely related to acsl6, regulated dorsoventral patterning in the zebrafish embryo (Miyares et al., 2013). It is reported that several maternal products regulate dorsoventral patterning of zebrafish and loss of this function brings about abnormal phenotype, characterized by failed gastrulation, ventralization, and lacking anterior structures and notochord (Reim & Brand, 2006; Nojima et al., 2010; Ge et al., 2014). Lastly, Rhotekin (Rtkn) interacts with active GTP-bound Rho proteins which regulate several cellular processes, such as cytokinesis, differentiation, cell polarity and cell movement. In zebrafish, Rho mediates cleavage furrow protein assembly during cytokinesis and cellular migration during gastrulation (Lai et al., 2005). Thus, it is likely that rtkn is an upstream regulator that plays an essential role in development of eel. Finally, Trim24, transcription intermediary factor (TIF) 1 α, modulates the first wave of transcription during early ZGA in the mouse (Torres-Padilla & Zernicka-Goetz, 2006). The ZGA is a prerequisite for gastrulation in zebrafish (Kane & Kimmel, 1993). A mouse study also has demonstrated that the *Trim24* transcript is notably abundant in the egg and embryo before ZGA. We pose that our findings - hatchability being positively correlated with mRNA levels of *trim24* - is reasonable. Indeed, it is evident that sufficiently high levels of maternal *trim24* mRNA in eggs are important for ZGA and subsequent development in fish.

Of note, our analyses showed that hatchability is adequately predicted by the levels of 5 maternal transcripts that accumulated in unfertilized eggs of artificially matured Japanese eel. These maternal transcripts, thus, are likely to be needed for development and can be used as markers of egg quality – the molecular mechanisms detailing *how* these transcripts contribute to hatchability remain to be elucidated in future studies.

Fish egg quality can be readily defined as the ability of an egg to be fertilized, to hatch and to then develop into a vital larva. However, when it comes to molecular signature (e.g. transcript abundance) or biochemical composition, egg quality is much more difficult to define

or describe. The loss of egg quality remains an ongoing problem in the artificial seedling production of aquaculture species, including Japanese eel. In aquaculture, many factors, such as environment, hormone, inappropriate timing of inducing ovulation etc, can negatively influence egg quality (Brook et al., 1997; Adachi, 2000; Bobe & Labbé, 2010), However, the complex interaction between multiple factors make it very difficult to identify the cause of compromised egg quality in many cases. Moreover, the molecular mechanisms leading to loss of egg quality remain unclear. Here, we have reported on a suite of transcripts implicated in hatchability that may prove useful as predictors of egg quality in eel. Correct prediction of egg quality in the unfertilized egg may help advance solutions to the issue of loss of egg quality in aquaculture species.

Candidate maternal transcripts associated with development

We identified 10 candidate genes whose transcripts are stored as maternal investment in the egg and whose quantitative signature may reflect developmental competence in artificially matured Japanese eel. Most of these candidate genes have not been previously studied in the context of maternal control of development. Our findings reinforce the importance of *trim24*, which was previously characterized as a key maternal transcript for development in the mouse (Torres-Padilla & Zernicka-Goetz, 2006); our results further reinforce that our experimental approach is effective in exposing maternal transcripts related to development. Table 5 shows the presumed function of these genes during embryogenesis, which were deemed to mainly concern ZGA, cell differentiation, patterning, gastrulation, and development of tissues. We propose that insufficient or reduced levels of 10 maternal transcripts in the egg may negatively affect these aspects of development, leading to hatching failure in artificially matured Japanese eel. In vertebrates, the timing of translation and the function of all identified transcripts, except *trim24*, during development have remained unclear. Further detail on the function of these transcripts will advance our understanding of maternal control of development in vertebrates and expose the molecular mechanisms responsible for reducing egg quality in species important for aquaculture.

ACKNOWLEDGMENTS

We are grateful to the members of the Laboratory of Fish Reproductive Physiology at the

Graduate School of Fisheries Sciences, Hokkaido University for assistance with the experiments. This study was partly supported by JSPS KAKENHI Grant Number 21228005 and JSPS Grant-in-Aid for JSPS Fellows Grant Number 13J01576.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interests.

REFERENCES

Adachi, S. (2000). Artificial control of maturation and spawning, and egg quality. *Kaiyo monthly, 32*(2), 120–126. (in Japanese)

Bertin, J., Wang, L., Guo, Y., Jacobson, M. D., Poyet, J. L., Srinivasula, S. M., ... Alnemri, E. S. (2001). CARD11 and CARD14 are novel caspase recruitment domain (CARD)/membrane-associated guanylate kinase (MAGUK) family members that interact with BCL10 and activate NF-kappa B. *Journal of Biological Chemistry, 276*(15), 11877–11882. doi:10.1074/jbc.M010512200

Bobe, J., & Labbe, C. (2010). Egg and sperm quality in fish. *General and Comparative Endocrinology, 165*(3), 535–548. doi:10.1016/j.ygcen.2009.02.011

Boccuto, L., Aoki, K., Flanagan-Steet, H., Chen, C. F., Fan, X., Bartel, F., ... Schwartz, C. E. (2014). A mutation in a ganglioside biosynthetic enzyme, ST3GAL5, results in salt & pepper syndrome, a neurocutaneous disorder with altered glycolipid and glycoprotein glycosylation. *Human Molecular Genetics, 23*(2), 418–433. doi:10.1093/hmg/ddt434

Boulekbache, H. (1981). Energy Metabolism in Fish Development. *Integrative and Comparative Biology*, *21*(2) 377–389. doi:10.1093/icb/21.2.377

Brooks, S., Tyler, C. R. & Sumpter, J. P. (1997). Egg quality in fish: what makes a good egg?. *Reviews in Fish Biology and Fisheries, 7*, 387–416. doi:10.1023/A:1018400130692

Chai, Y., Tosaka, R., Abe, T., Sago, K., Sago, Y., Hatanaka, E., … Adachi, S. (2010). The relationship between the developmental stage of oocytes in various seasons and

the quality of egg obtained by artificial maturation in the feminized Japanese eel *Anguilla japonica*. *Aquaculture Science, 58*(2), 269–278. doi:10.11233/aquaculturesci.58.269

Chapman, R. W., Reading, B. J., & Sullivan, C. V. (2014). Ovary Transcriptome Profiling via Artificial Intelligence Reveals a Transcriptomic Fingerprint Predicting Egg Quality in Striped Bass, Morone saxatilis. *Plos One, 9*(5), 10. doi:10.1371/journal.pone.0096818

Chen, C. H., Lin, H., Chuang, S. M., Lin, S. Y., & Chen, J. J. W. (2010). Acidic stress facilitates tyrosine phosphorylation of HLJ1 to associate with actin cytoskeleton in lung cancer cells. *Experimental Cell Research, 316*(17), 2910–2921. doi:10.1016/j.yexcr.2010.06.027

Correa, R. G., Tergaonkar, V., Ng, J. K., Dubova, I., Izpisua-Belmonte, J. C., & Verma, I. M. (2004). Characterization of NF-kappa B/I kappa B proteins in zebra fish and their involvement in notochord development. *Molecular and Cellular Biology, 24*(12), 5257-5268. doi:10.1128/mcb.24.12.5257–5268.2004

Dosch, R., Wagner, D. S., Mintzer, K. A., Runke, G., Wiemelt, A. P., & Mullins, M. C. (2004). Maternal control of vertebrate development before the midblastula transition: Mutants from the zebrafish I. *Developmental Cell, 6*(6), 771–780. doi:10.1016/j.devcel.2004.05.002

Ge, X. Y., Grotjahn, D., Welch, E., Lyman-Gingerich, J., Holguin, C., Dimitrova, E., ... Pelegri, F. (2014). Hecate/Grip2a Acts to Reorganize the Cytoskeleton in the Symmetry-Breaking Event of Embryonic Axis Induction. *Plos Genetics, 10*(6), 19. doi:10.1371/journal.pgen.1004422

Götz, S., Garcia-Gomez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., ... Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, *36*(10), 3420–3435. doi:10.1093/nar/gkn176

Haznedaroglu, B. Z., Reeves, D., Rismani-Yazdi, H., & Peccia, J. (2012). Optimization of *de novo* transcriptome assembly from high-throughput short read sequencing data

529 improves functional annotation for non-model organisms. *Bmc Bioinformatics*, *13*,
530 9. doi:10.1186/1471-2105-13-170
531
532 Henkel, C. V., Dirks, R. P., de Wijze, D. L., Minegishi, Y., Aoyama, J., Jansen, H. J., . . .
533 van den Thillart, G. (2012). First draft genome sequence of the Japanese eel,
534 Anguilla japonica. *Gene, 511*(2), 195–201. doi:10.1016/j.gene.2012.09.064
535
536 Holloway, B. A., Canny, S. G. D., Ye, Y., Slusarski, D. C., Freisinger, C. M., Dosch, R., . . .
537 Mullins, M. C. (2009). A Novel Role for MAPKAPK2 in Morphogenesis during
538 Zebrafish Development. *Plos Genetics, 5*(3), 14.
539 doi:10.1371/journal.pgen.1000413
540
541 Izumi, H., Gen, K., Horiuchi, M., Matsuya, N., Ijiri, S., & Adachi, S. (2016). Quantitative
542 comparisons of maternal transcripts related to cell division between good and
543 poor quality eggs from artificially matured Japanese eel *Anguilla japonica*.
544 *Aquaculture Science*, *64*(1), 29–42. doi:10.11233/aquaculturesci.64.29
545
546 Izumi, H., Hagihara, S., Gen, K., Ijiri, S., & Adachi, S. (2015). Selective depletion of
547 ribosomal RNA from total RNA for next-generation sequencing analysis of
548 maternal mRNA in eggs of the Japanese eel Anguilla japonica. *Bulletin of*
549 *Fisheries Sciences, Hokkaido University, 65*(3), 125–130.
550 doi:10.14943/bull.fish.65.3.125 (in Japanese)
551
552 Kagawa, H., Tanaka, H., Ohta, H., Okuzawa, K., & Iinuma, N. (1997). Induced ovulation by
553 injection of 17,20β-dihydroxy-4-pregnen-3-one in the artificially matured Japanese
554 eel, with special reference to ovulation time. *Fisheries Science*, *63*(3), 365–367.
555 doi:org/10.2331/fishsci.63.365
556
557 Kane, D. A., & Kimmel, C. B. (1993). THE ZEBRAFISH MIDBLASTULA
558 TRANSITION. *Development, 119*(2), 447–456.
559
560 Lai, S. L., Chang, C. N., Wang, P. J., & Lee, S. J. (2005). Rho mediates cytokinesis and
561 epiboly via ROCK in zebrafish. *Molecular Reproduction and Development, 71*(2),
562 186–196. doi:10.1002/mrd.20290
563

Leek, J. T., Scharpf, R. B., Bravo, H. C., Simcha, D., Langmead, B., Johnson, W. E., ... Irizarry, R. A. (2010). Tackling the widespread and critical impact of batch effects in high-throughput data. *Nature Reviews Genetics, 11*(10), 733–739. doi:10.1038/nrg2825

Lee, M. T., Bonneau, A. R., & Giraldez, A. J. (2014). Zygotic Genome Activation During the Maternal-to-Zygotic Transition. In R. Schekman & R. Lehmann (Eds.), *Annual Review of Cell and Developmental Biology, Vol 30* (Vol. 30, pp. 581–613). Palo Alto: Annual Reviews.

Lee, M. T., Bonneau, A. R., Takacs, C. M., Bazzini, A. A., DiVito, K. R., Fleming, E. S., & Giraldez, A. J. (2013). Nanog, Pou5f1 and SoxB1 activate zygotic gene expression during the maternal-to-zygotic transition. *Nature, 503*(7476), 360–364. doi:10.1038/nature12632

Lindeman, R. E., & Pelegri, F. (2012). Localized Products of futile cycle/lrmp Promote Centrosome-Nucleus Attachment in the Zebrafish Zygote. *Current Biology, 22*(10), 843–851. doi:10.1016/j.cub.2012.03.058

Marlow. F.L. (2010). *Maternal Control of Development in Vertebrates: My Mother Made Me Do It!*. San Rafael, CA: Morgan & Claypool Life Sciences.

Martin, J. A., & Wang, Z. (2011). Next-generation transcriptome assembly. *Nature Reviews Genetics, 12*(10), 671–682. doi:10.1038/nrg3068

Miyares, R. L., Stein, C., Renisch, B., Anderson, J. L., Hammerschmidt, M., & Farber, S. A. (2013). Long-Chain Acyl-CoA Synthetase 4A Regulates Smad Activity and Dorsoventral Patterning in the Zebrafish Embryo. *Developmental Cell, 27*(6), 635–647. doi:10.1016/j.devcel.2013.11.011

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods, 5*(7), 621–628. doi:10.1038/nmeth.1226

Nojima, H., Rothhamel, S., Shimizu, T., Kim, C. H., Yonemura, S., Marlow, F. L., & Hibi, M. (2010). Syntabulin, a motor protein linker, controls dorsal determination. *Development, 137*(6), 923–933. doi:10.1242/dev.046425

Ono, H., Ishii, K., Kozaki, T., Ogiwara, I., Kanekatsu, M., & Yamada, T. (2015). Removal of redundant contigs from *de novo* RNA-Seq assemblies via homology search improves accurate detection of differentially expressed genes. *Bmc Genomics*, *16*, 13. doi:10.1186/s12864-015-2247-0

Pelegri, F., Dekens, M. P. S., Schulte-Merker, S., Maischein, H. M., Weiler, C., & Nusslein-Volhard, C. (2004). Identification of recessive maternal-effect mutations in the zebrafish using a gynogenesis-based method. *Developmental Dynamics, 231*(2), 324–335. doi:10.1002/dvdy.20145

Pfaffl, M. W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Research*, *29*(9), 6. doi:10.1093/nar/29.9.e45

Reim, G., & Brand, M. (2006). Maternal control of vertebrate dorsoventral axis formation and epiboly by the POU domain protein Spg/Pou2/Oct4. *Development, 133*(14), 2757–2770. doi:10.1242/dev.02391

Rodemer, C., Thai, T. P., Brugger, B., Kaercher, T., Werner, H., Nave, K. A., ... Just, W. W. (2003). Inactivation of ether lipid biosynthesis causes male infertility, defects in eye development and optic nerve hypoplasia in mice. *Human Molecular Genetics, 12*(15), 1881–1895. doi:10.1093/hmg/ddg191

Seeley, M., Huang, W., Chen, Z. Y., Wolff, W. O., Lin, X. Y., & Xu, X. L. (2007). Depletion of zebrafish titin reduces cardiac contractility by disrupting the assembly of Z-discs and A-bands. *Circulation Research, 100*(2), 238–245. doi:10.1161/01.RES.0000255758.69821.b5

Sullivan, C. V., Chapman, R. W., Reading, B. J., & Anderson, P. E. (2015). Transcriptomics of mRNA and egg quality in farmed fish: Some recent developments and future directions. *General and Comparative Endocrinology, 221*, 23–30. doi:10.1016/j.ygcen.2015.02.012

Tadros, W., & Lipshitz, H. D. (2009). The maternal-to-zygotic transition: a play in two acts. *Development, 136*(18), 3033–3042. doi:10.1242/dev.033183

Tanaka, H. (2015). Progression in artificial seedling production of Japanese eel Anguilla japonica. *Fisheries Science, 81*(1), 11–19. doi:10.1007/s12562-014-0821-z

Thiery, J. P., Acloque, H., Huang, R. Y. J., & Nieto, M. A. (2009). Epithelial-Mesenchymal Transitions in Development and Disease. *Cell, 139*(5), 871–890. doi:10.1016/j.cell.2009.11.007

Torres-Padilla, M. E., & Zernicka-Goetz, M. (2006). Role of TIF1 alpha as a modulator of embryonic transcription in the mouse zygote. *Journal of Cell Biology, 174*(3), 329–338. doi:10.1083/jcb.200603146

Trapnell, C., Pachter, L., & Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics, 25*(9), 1105–1111. doi:10.1093/bioinformatics/btp120

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., ... Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols*, *7*(3), 562–578. doi:10.1038/nprot.2012.016

Wagner, D. S., Dosch, R., Mintzer, K. A., Wiemelt, A. P., & Mullins, M. C. (2004). Maternal control of development at the midblastula transition and beyond: Mutants from the zebrafish II. *Developmental Cell, 6*(6), 781–790. doi:10.1016/j.devcel.2004.04.001

Wagner, G. P., Kin, K., & Lynch, V. J. (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in Biosciences*, *131*(4), 281–285. doi:10.1007/s12064-012-0162-3

Wieland, OH. (1983). The mammalian pyruvate dehydrogenase complex: structure and regulation. *Reviews of physiology, biochemistry and pharmacology, 96*,123–170.

Xiong, L., Wen, Y., Miao, X. Y., & Yang, Z. L. (2014). NT5E and FcGBP as key regulators of TGF-1-induced epithelial-mesenchymal transition (EMT) are associated with

tumor progression and survival of patients with gallbladder cancer. *Cell and Tissue Research, 355*(2), 365–374. doi:10.1007/s00441-013-1752-1

Zarski, D., Nguyen, T., Le Cam, A., Montfort, J., Dutto, G., Vidal, M. O., ... Bobe, J. (2017). Transcriptomic Profiling of Egg Quality in Sea Bass (Dicentrarchus labrax) Sheds Light on Genes Involved in Ubiquitination and Translation. *Marine Biotechnology, 19*(1), 102–115. doi:10.1007/s10126-017-9732-1

TABLE

Table 1. Quality, estimated by fertilization and hatching rates, and purpose of egg samples from artificially matured Japanese eel.

No.	Fertility (%)	Hatchability (%)	Egg quality	Purpose
1	98.9	93.5	good	NGS & qPCR
2	97.5	93.3	good	NGS & qPCR
3	100	87.8	good	qPCR
4	91.9	82.1	good	qPCR
5†	85 <	85.2	good	NGS & qPCR
6	91.4	81.5	good	qPCR
7	98.4	92.9	good	qPCR
8	95.2	20	poor	qPCR
9	86	18	poor	qPCR
10	90.5	15.5	poor	NGS & qPCR
11	85.8	10.1	poor	qPCR
12	87.4	0	poor	NGS & qPCR
13	86.7	0	poor	NGS & qPCR

†Sample preserved RNA Later

681 **Table 2.** List of 36 transcripts selected by *in silico* screening (Analysis 1) and qPCR screening.

Contig No.	Length	Description (BLASTx)	E-value	Biological process	Fold change	P value
346537	265	lamin-b receptor	5.7E-12	cell migration involved in gastrulation	0.4	0
107048	1535	caspase recruitment domain-containing protein 14	1.9E-93	regulation of apoptotic process	0.25	0.002
47139	1843	beta- n-acetylgalactosaminyltransferase 1-like	0	lipid glycosylation	0.19	0.002
81492	139	protein fam57b-like	5.2E-15	central nervous system development	0.42	0.002
186064	279	solute carrier family 40 member 1	3.2E-11	iron ion transport	0.13	0.002
180644	351	novel protein human titin	6.6E-12	skeletal muscle tissue development	0.39	0.004
113183	598	testicular haploid expressed gene	1.1E-41	spermatogenesis	0.15	0.009
35937	141	pantothenate kinase 3	1.9E-15	coenzyme A biosynthetic process	0.38	0.01
8543	174	spermidine synthase-like	7.5E-28	amine metabolic process	0.39	0.01
123154	440	von willebrand factor d and egf domain-containing protein	2.5E-62	-	0.23	0.01
44548	432	calponin-1	7.7E-25	actomyosin structure organization	0.22	0.013
36581	368	gtpase imap family member 7-like	1.8E-29	-	0.11	0.014
25493	5433	otoancorin- partial	0	cell-matrix adhesion	0.13	0.018
71472	1502	pyruvate dehydrogenase	3.6E-06	protein dephosphorylation	0.18	0.024
187717	272	roundabout homolog 3 isoform x3	8.6E-43	axon guidance	0.44	0.024
64061	187	dihydroxyacetone phosphate acyltransferase-like	1.4E-14	cellular lipid metabolic process	0.45	0.027
185635	215	astrotactin-1-like isoform x1	2.4E-43	neuron cell-cell adhesion, neuron migration	0.52	0.027

61146	306	bromodomain and wd repeat-containing protein 3-like isoform x1	3.0E-07	cytoskeleton organization, regulation of cell shape	0.06	0.03
90079	530	matrix metalloproteinase-20	7.9E-08	proteolysis	0.19	0.031
98904	956	sodium hydrogen exchanger 5-like	1.7E-155	regulation of intracellular pH	0.43	0.032
32498	249	mesothelin-like partial	4.1E-13	cell-matrix adhesion	0.12	0.034
112812	106	dnaj homolog subfamily b member 4-like	5.6E-13	protein folding	0.33	0.035
7830	4035	parathyroid hormone parathyroid hormone-related peptide receptor	0	dorsal aorta development	0.37	0.041
131862	487	scan domain-containing protein 3-like	2.9E-06	regulation of transcription, DNA-templated	0.64	0.041
69803	339	perilipin-2 isoform x3	4.3E-31	long-chain fatty acid transport	0.55	0.044
105031	344	cub and sushi domain-containing protein 1-like	4.9E-50	-	0.24	0.049
54188	1235	kelch-like protein 38-like	1.1E-151	protein ubiquitination	0.18	0.059
228141	265	low quality protein: titin	1.0E-45	skeletal muscle tissue development	0.32	0.095
184082	254	ras rap gtpase-activating isoform x1	1.4E-17	negative regulation of Ras protein signal transduction	0.31	0.113
216371	209	slit homolog 3	5.8E-33	central nervous system projection neuron axonogenesis	0.21	0.115
51153	202	tumor necrosis factor receptor superfamily member 14-like isoform x2	1.8E-14	regulation of apoptotic process, response to lipopolysaccharide	1.78	0.116
218733	270	alpha-catulin	1.4E-47	cell adhesion	0.3	0.116
40760	268	matrix metalloproteinase-19-like	1.8E-30	proteolysis	0.36	0.133
223030	480	titin-like	1.7E-62	skeletal muscle tissue development	0.31	0.138
156849	231	c-binding protein	2.5E-24	-	0.12	0.156

177733	456	junctional protein associated with coronary artery disease	1.0E-37	cell adhesion	0.17	0.168

682

683 **Table 3.** List of 85 transcripts selected by *in silico* screening (Analysis 2).

Transcript No.	Length (b)	genome locus	Description (BLASTx)	E-value	Biological process
47766	230	gi|546466767|gb|KI307198.1|:40381-42518	low quality protein: titin	3.3.E-34	skeletal muscle tissue development
32885	1587	gi|546463338|gb|KI309797.1|:9705-14172	solute carrier family 12 member 4-like isoform x1	0	chemical synaptic transmission
8852	287	gi|546447610|gb|KI322079.1|:6642-8285	long-chain-fatty-acid-- ligase 6-like	5.8.E-55	long-chain fatty acid metabolic process
51979	723	gi|546467663|gb|KI306639.1|:645-4603	spermatogenesis-associated protein 13 isoform x2	4.8.E-78	regulation of Rho protein signal transduction
74897	583	gi|546470452|gb|KI304769.1|:145862-147457	von willebrand factor a domain-containing protein 3a	1.0.E-39	-
14109	198	gi|546453374|gb|KI317089.1|:12877-13507	receptor-type tyrosine-protein phosphatase zeta isoform x3	8.5.E-25	dephosphorylation
19392	1601	gi|546457140|gb|KI314184.1|:7157-19322	ras gtpase-activating protein 1-like	0	blood vessel development
35139	191	gi|546464182|gb|KI309311.1|:321-5864	calcium uniporter mitochondrial-like isoform x1	7.2.E-38	regulation of heart contraction, convergent extension involved in gastrulation
40658	245	gi|546465458|gb|KI308266.1|:114072-116308	rhotekin-like isoform x2	2.9.E-35	signal transduction
52511	297	gi|546467781|gb|KI306555.1|:226-29854	titin isoform x5	1.0.E-41	skeletal muscle tissue development
70699	379	gi|546470120|gb|KI304996.1|:39143-44808	epidermal growth factor receptor kinase substrate 8-like protein 2-like isoform x2	9.7.E-61	positive regulation of ruffle assembly
822	145	gi|546401872|gb|KI360409.1|:0-532	tight junction protein zo-1-like isoform x2	1.9.E-22	pigmentation
74065	338	gi|546470376|gb|KI304810.1|:184478-188419	cyclin-dependent kinase-like 1	1.4.E-75	protein phosphorylation
11598	1087	gi|546451022|gb|KI319089.1|:584-14381	plexin-a1-like isoform x2	0	signal transduction
19102	2436	gi|546456976|gb|KI314297.1|:22368-32455	gem-associated protein 5-like	0	spliceosomal snRNP assembly

32022	2485	gi|546463143|gb|KI309992.1|:7634-19903	peptidyl-prolyl cis-trans isomerase fkbp10-like	0	chaperone-mediated protein folding
48754	941	gi|546467007|gb|KI307054.1|:94-9086	protein hook homolog 2-like isoform x1	3.9.E-157	endosome to lysosome transport
19151	454	gi|546456996|gb|KI314284.1|:9679-18283	hepatoma-derived growth factor	7.3.E-61	negative regulation of transcription from RNA polymerase II promoter
31311	3671	gi|546462934|gb|KI310169.1|:44205-53406	telomere length regulation protein tel2 homolog	0	protein stabilization, regulation of TOR signaling
42044	1727	gi|546465682|gb|KI308042.1|:2024-42167	focal adhesion kinase 1 isoform x1	0	lens fiber cell morphogenesis
735	218	gi|546399759|gb|KI362139.1|:15-922	phosphatidylinositol 4-kinase alpha-like isoform x1	7.5.E-39	phosphatidylinositol-mediated signaling
21621	1605	gi|546458616|gb|KI313236.1|:0-12784	zinc finger protein ozf-like	2.2.E-98	regulation of transcription, DNA-templated
23638	312	gi|546459803|gb|KI312454.1|:8718-9436	nesprin-1 isoform x4	9.4.E-32	cytoskeletal anchoring at nuclear membrane
37240	189	gi|546464742|gb|KI308900.1|:25879-27085	alkaline ceramidase 3	3.4.E-30	ceramide metabolic process
67483	1322	gi|546469875|gb|KI305212.1|:9404-47383	leukocyte elastase inhibitor-like isoform x2	2.8.E-158	neutrophil degranulation
1067	163	gi|546406455|gb|KI356682.1|:306-469	PREDICTED: cadherin-1	1.1.E-04	epiboly involved in gastrulation with mouth forming second, single organismal cell-cell adhesion, neuron migration
3199	113	gi|546429263|gb|KI337706.1|:139-473	ubiquitin carboxyl-terminal hydrolase 34- partial	2.2.E-16	positive regulation of canonical Wnt signaling pathway
10477	861	gi|546449739|gb|KI320182.1|:7419-17054	cytosolic carboxypeptidase 1-like	1.8.E-100	chordate embryonic development, T cell differentiation in thymus
11307	281	gi|546450697|gb|KI319373.1|:37205-38100	protein inscuteable homolog isoform x1	8.5.E-15	nervous system development
21147	2058	gi|546458320|gb|KI313396.1|:2066-11024	ap-1 complex subunit gamma-1-like	5.6.E-90	pectoral fin development
22694	604	gi|546459307|gb|KI312793.1|:30352-35119	rcc1 and btb domain-containing protein 1	6.1.E-131	retina vasculature morphogenesis in camera-type eye, blood vessel development
29117	290	gi|546462094|gb|KI310726.1|:24176-41307	pre-mrna 3 end processing protein wdr33 isoform x8	1.5.E-61	mRNA cleavage
30567	3871	gi|546462656|gb|KI310344.1|:44822-56454	protein o-linked-mannose beta- -n-acetylglucosaminyltransferase 1	0	protein glycosylation
64640	552	gi|546469589|gb|KI305405.1|:106647-198499	glycine receptor subunit alpha-4- partial	7.1.E-42	neuropeptide signaling pathway

79333	1373	gi|546470891|gb|KI304576.1|:60912-77508	bcl2 adenovirus e1b 19 kda protein-interacting protein 2 isoform x1	0	apoptotic process
20857	190	gi|546458113|gb|KI313529.1|:11706-14316	creb-binding partial	1.5.E-33	histone acetylation
5428	292	gi|546440436|gb|KI328487.1|:3290-7399	protein jagged-1b-like	2.3.E-53	cartilage development involved in endochondral bone morphogenesis, liver development, pancreas development, thyroid gland development, face morphogenesis
10737	164	gi|546450072|gb|KI319918.1|:5654-5986	protein fam92b isoform x1	2.0.E-14	-
14719	812	gi|546453904|gb|KI316654.1|:4876-27344	gtpase imap family member 7-like	1.1.E-75	-
22105	421	gi|546458974|gb|KI313006.1|:31151-36012	calcium calmodulin-dependent protein kinase type 1d-like	2.5.E-76	chordate embryonic development
31852	444	gi|546463110|gb|KI310025.1|:68937-81387	nuclear pore complex protein nup93- partial	2.4.E-39	nuclear pore complex assembly
54956	1839	gi|546468211|gb|KI306286.1|:28060-50504	integrator complex subunit 8 isoform x1	0	snRNA processing
58632	357	gi|546468847|gb|KI305914.1|:11007-20940	tight junction protein zo-3 isoform x1	5.2.E-42	multicellular organismal homeostasis, water homeostasis
69150	1440	gi|546469987|gb|KI305100.1|:64856-91293	transcription intermediary factor 1-alpha-like isoform x1	0	regulation of apoptotic process
1479	244	gi|546412573|gb|KI351313.1|:401-962	haus augmin-like complex subunit 8	5.9.E-10	spindle assembly, centrosome organization
14712	938	gi|546453904|gb|KI316654.1|:4876-27344	gtpase imap family member 7-like	1.8.E-61	-
23549	194	gi|546459771|gb|KI312481.1|:4298-5306	serine threonine-protein kinase mtor	9.6.E-27	regulation of myelination, skin development
728	141	gi|546399452|gb|KI362415.1|:774-915	zinc finger swim domain-containing protein 7-like	8.6.E-04	double-strand break repair via homologous recombination
29131	739	gi|546462098|gb|KI310722.1|:91-35794	peptidyl-glycine alpha-amidating monooxygenase-like	2.6.E-76	oxidation-reduction process, peptide metabolic process
36664	356	gi|546464594|gb|KI309018.1|:61825-82426	slit-robo rho gtpase-activating protein 1-like	2.1.E-58	negative regulation of cell migration
39183	548	gi|546465173|gb|KI308551.1|:34818-37937	cytochrome c oxidase subunit mitochondrial-like	8.1.E-39	response to cadmium ion
59652	543	gi|546468959|gb|KI305814.1|:13123-25297	yeats domain-containing protein 2 isoform x1	6.0.E-42	regulation of transcription, DNA-templated
60426	2211	gi|546469035|gb|KI305763.1|:74515-88014	spartin-like isoform x1	0	cell division, regulation of mitochondrial membrane potential
2414	229	gi|546423577|gb|KI342752.1|:43-1073	rho guanine nucleotide exchange factor 12-like isoform x3	9.9.E-33	G-protein coupled receptor signaling pathway

12712	383	gi|546452144|gb|KI318107.1|:1691-14650	high mobility group protein hmgi-c	1.1.E-13	base-excision repair
43730	2138	gi|546465958|gb|KI307770.1|:4503-12983	non-receptor tyrosine-protein kinase tyk2 isoform x2	0	cell differentiation, regulation of cell proliferation, cell migration, innate immune response
4854	543	gi|546438664|gb|KI330226.1|:82-3897	mitogen-activated protein kinase 8	4.0.E-129	phosphorylation
16140	1182	gi|546454873|gb|KI315840.1|:2469-15234	inositol -trisphosphate receptor type 2	3.0.E-67	release of sequestered calcium ion into cytosol
17593	1341	gi|546455882|gb|KI315080.1|:2702-12533	protein mon2 homolog isoform x4	0	Golgi to endosome transport
35572	317	gi|546464323|gb|KI309213.1|:6543-21515	thioredoxin domain-containing protein 16-like	4.9.E-30	cell redox homeostasis
41542	1374	gi|546465596|gb|KI308128.1|:300-8041	tubby-related protein 4- partial	0	protein localization to cilium
9465	1258	gi|546448481|gb|KI321307.1|:1943-4142	peflin	1.1.E-16	response to calcium ion
14289	860	gi|546453504|gb|KI316980.1|:4007-10361	vacuolar protein sorting-associated protein 4a	0	endosomal transport
20462	151	gi|546457878|gb|KI313690.1|:22752-25883	unconventional myosin-xviiia	5.1.E-16	cell migration, actomyosin structure organization, Golgi organization
21232	307	gi|546458366|gb|KI313372.1|:11188-19175	sjoegren syndrome nuclear autoantigen 1 homolog	6.7.E-04	G2/M transition of mitotic cell cycle, ciliary basal body docking
55810	1030	gi|546468390|gb|KI306202.1|:38233-54470	pc-esterase domain-containing protein 1a-like	3.1.E-106	-
79079	1304	gi|546470877|gb|KI304586.1|:17093-36390	mitochondrial ubiquitin ligase activator of nfkb 1-a-like	5.7.E-81	activation of cysteine-type endopeptidase activity involved in apoptotic process, mitochondrial fission, mitochondrion localization
769	104	gi|546400551|gb|KI361484.1|:536-940	ornithine decarboxylase	6.0.E-13	eye photoreceptor cell development
18562	302	gi|546456596|gb|KI314587.1|:18319-18621	type i inositol -bisphosphate 4-phosphatase- partial	1.5.E-06	signal transduction
39088	1949	gi|546465157|gb|KI308567.1|:77779-95914	serine--trna mitochondrial-like	2.4.E-179	seryl-tRNA aminoacylation
54702	1899	gi|546468173|gb|KI306309.1|:13997-23346	alpha- and gamma-adaptin-binding protein p34-like	2.2.E-160	small GTPase mediated signal transduction
54769	655	gi|546468183|gb|KI306303.1|:76568-87792	terf1-interacting nuclear factor 2 isoform x1	6.0.E-29	negative regulation of epithelial cell proliferation
1170	215	gi|546408290|gb|KI355226.1|:131-658	casein kinase ii subunit beta	9.2.E-30	signal transduction

2419	265	gi|546423662|gb|KI342682.1|:1283-1998	protein fam184a-like isoform x1	2.9.E-21	-
5279	207	gi|546439970|gb|KI328920.1|:65-272	dual specificity protein phosphatase cdc14a-like	6.6.E-38	cilium assembly
7184	830	gi|546444665|gb|KI324638.1|:4579-9883	uncharacterized kda	2.0.E-05	-
16685	2606	gi|546455280|gb|KI315514.1|:2361-16743	rap guanine nucleotide exchange factor 2 isoform x7	0	signal transduction
35444	491	gi|546464291|gb|KI309242.1|:14932-15423	PREDICTED: uncharacterized protein LOC102076953, partial	1.4.E-05	-
43733	2099	gi|546465958|gb|KI307770.1|:4503-12983	non-receptor tyrosine-protein kinase tyk2 isoform x2	0	cell differentiation, regulation of cell proliferation, cell migration, innate immune response
56001	313	gi|546468440|gb|KI306181.1|:75455-75768	dual specificity protein phosphatase 13 isoform a-like	5.6.E-25	protein dephosphorylation
71876	1128	gi|546470215|gb|KI304925.1|:766-49335	arf-gap with ank repeat and ph domain-containing protein 3-like isoform x2	0	small GTPase mediated signal transduction
63840	203	gi|546469508|gb|KI305473.1|:136004-146252	PREDICTED: uncharacterized protein LOC104934648	4.4.E-12	-
67268	990	gi|546469861|gb|KI305226.1|:15073-50557	anoctamin-9-like	1.3.E-150	chloride transport
79332	3866	gi|546470891|gb|KI304576.1|:29882-59308	autophagy-related protein 13 isoform x3	0	autophagosome assembly
82757	1870	gi|546471130|gb|KI304459.1|:354247-368538	immunoglobulin superfamily member 21-like isoform x2	2.5.E-108	-

685

686 **Table 4.** Stepwise multiple regression analyses to identify factors influencing hatchability.

Analysis	Independent variable		Unstandardized coefficients B	Standardized coefficients β	Validity of the regression model	
	No.	Description (BLASTx)			R^2	F
1	64061	dihydroxyacetone phosphate acyltransferase-like	87.43	0.808***	0.811	24.65***
	47139	beta- n-acetyl galactosaminyltransferase 1-like	70.01	0.412*		
2	8852	long-chain-fatty-acid-- ligase 6-like	63.94	0.627***	0.921	43.64***
	40658	rhotekin-like isoform x2	46.57	0.542***		
	69150	transcription intermediary factor 1-alpha-like isoform x1	57.08	0.364**		

687 Data marked with asterisk were statistically different (*P < 0.05, **P < 0.01, ***P < 0.001).

688

689

690

691

692 **Table 5.** List of candidate maternal transcripts related to development in Japanese eel and

693 presumed gene function during embryogenesis.

Characteristic	Symbol[†]	Biological process	Presumed gene function during embryogenesis
Differentially expressed transcripts (PQ < GQ[‡])	*dnajb4*	protein folding	Gastrulation
	gnpat	cellular lipid metabolic process	Eye development, Development of nervous system
	card14	regulation of apoptotic process	Cell differentiation
	pdp1	protein dephosphorylation	Energy metabolism
	fcgbp	-	Gastrulation
	ttn	skeletal muscle tissue development	Heart development, Somite development
Transcripts contributed to hatchability	*gnpat*	cellular lipid metabolic process	Eye development, Development of nervous system
	b4galnt1	lipid glycosylation	Development of nervous system
	acsl6	long-chain fatty acid metabolic process	Patterning
	rtkn	signal transduction	Cytokinesis, Gastrulation
	trim24	regulation of apoptotic process	ZGA

694 [†]For all genes, gene symbols identical to those in humans were given when possible.

695 [‡]GQ, good quality egg, PQ, poor quality egg.

696

FIGURE LEGEND

Figure 1. Work flow. Two approaches, i.e., *de novo* assembly (Analysis 1) and reference-based assembly (Analysis 2), were performed in keeping with Martin & Wang (2011). GQ, good quality egg, PQ, poor quality egg.

Figure 2. Dendrogram derived from hierarchical cluster analysis of RPKM (a) or FPKM values (b) of all transcripts in three batches of good quality eggs (No.1, 2, 5) and three of poor quality eggs (No.10, 12, 13). The vertical scale represents the distance between clusters.

Figure 3. Relative levels of transcripts selected by Analysis 1 screening in good quality and poor quality eggs. Open columns, good quality eggs (No. 1, 2, 3, 4, 5, 7); solid columns, poor quality eggs (No. 8, 9, 10, 11, 12, 13). Data marked with asterisk were statistically different ($P < 0.01$).

Figure 4. Relative levels of the transcripts selected by Analysis 2 screening in good quality and poor quality eggs. Open columns, good quality eggs (No. 1, 2, 3, 4, 6, 7); solid columns, poor quality eggs (No. 8, 9, 10, 11, 12, 13). Data marked with asterisk were statistically different ($P < 0.01$).

Figure 5. Relationship between predicted and observed hatchability. The predicted hatchability was estimated using Analyses 1 and 2 equations of multiple linear regression.

LEGENDS FOR SUPPLEMENTARY TABLES

Table S1. Primer sequence (Analysis 1)

Table S2. Primer sequence (Analysis 2)

Table S3. Summary of sequencing result

Table S4. List of 148 transcripts selected by *in silico* screening (Analysis 1) and results of qPCR analyses

729 **Table S5.** List of 97 transcripts selected by *in silico* screening (Analysis 1) and results of
730 qPCR analyses
731 †RPKM values among the three batches of good eggs.
732
733

Fig. 1

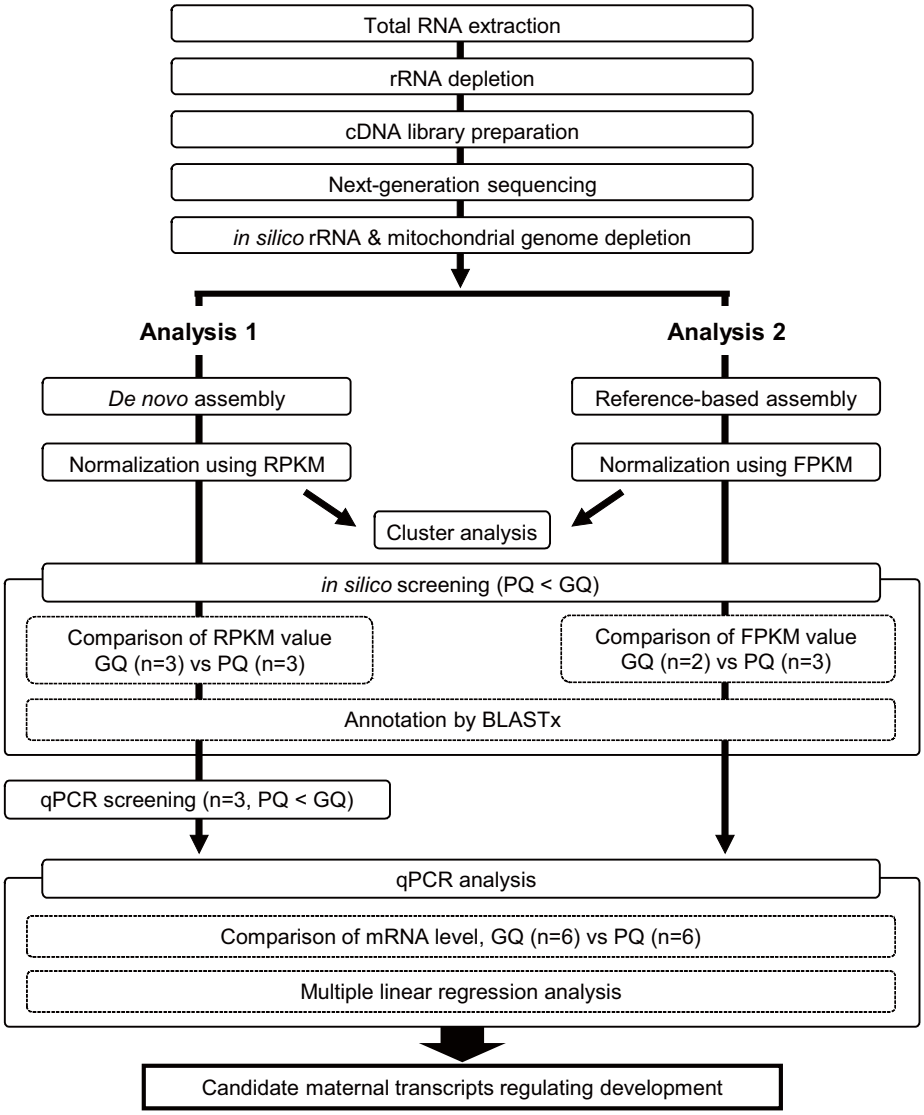

Total RNA extraction
rRNA depletion
cDNA library preparation
Next-generation sequencing
in silico rRNA & mitochondrial genome depletion
Analysis 1
Analysis 2
De novo assembly
Reference-based assembly
Normalization using RPKM
Normalization using FPKM
Cluster analysis
in silico screening (PQ < GQ)
Comparison of RPKM value
GQ (n=3) vs PQ (n=3)
Comparison of FPKM value
GQ (n=2) vs PQ (n=3)
Annotation by BLASTx
qPCR screening (n=3, PQ < GQ)
qPCR analysis
Comparison of mRNA level, GQ (n=6) vs PQ (n=6)
Multiple linear regression analysis
Candidate maternal transcripts regulating development

Fig. 2

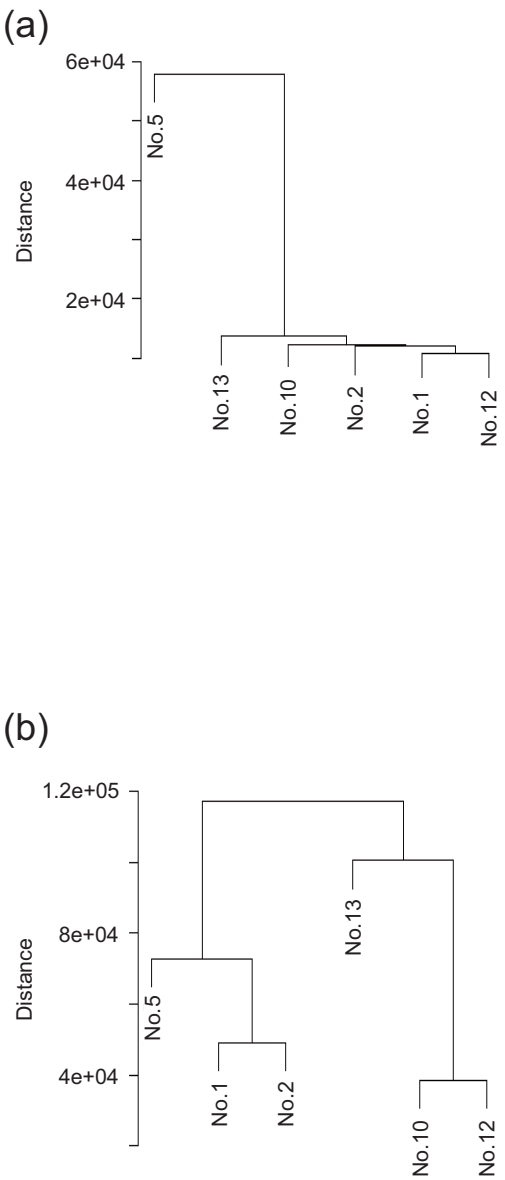

(a)
6e+04
4e+04
2e+04
Distance
No.5
No.13
No.10
No.2
No.1
No.12
(b)
1.2e+05
8e+04
4e+04
Distance
No.5
No.1
No.2
No.13
No.10
No.12

Fig. 3

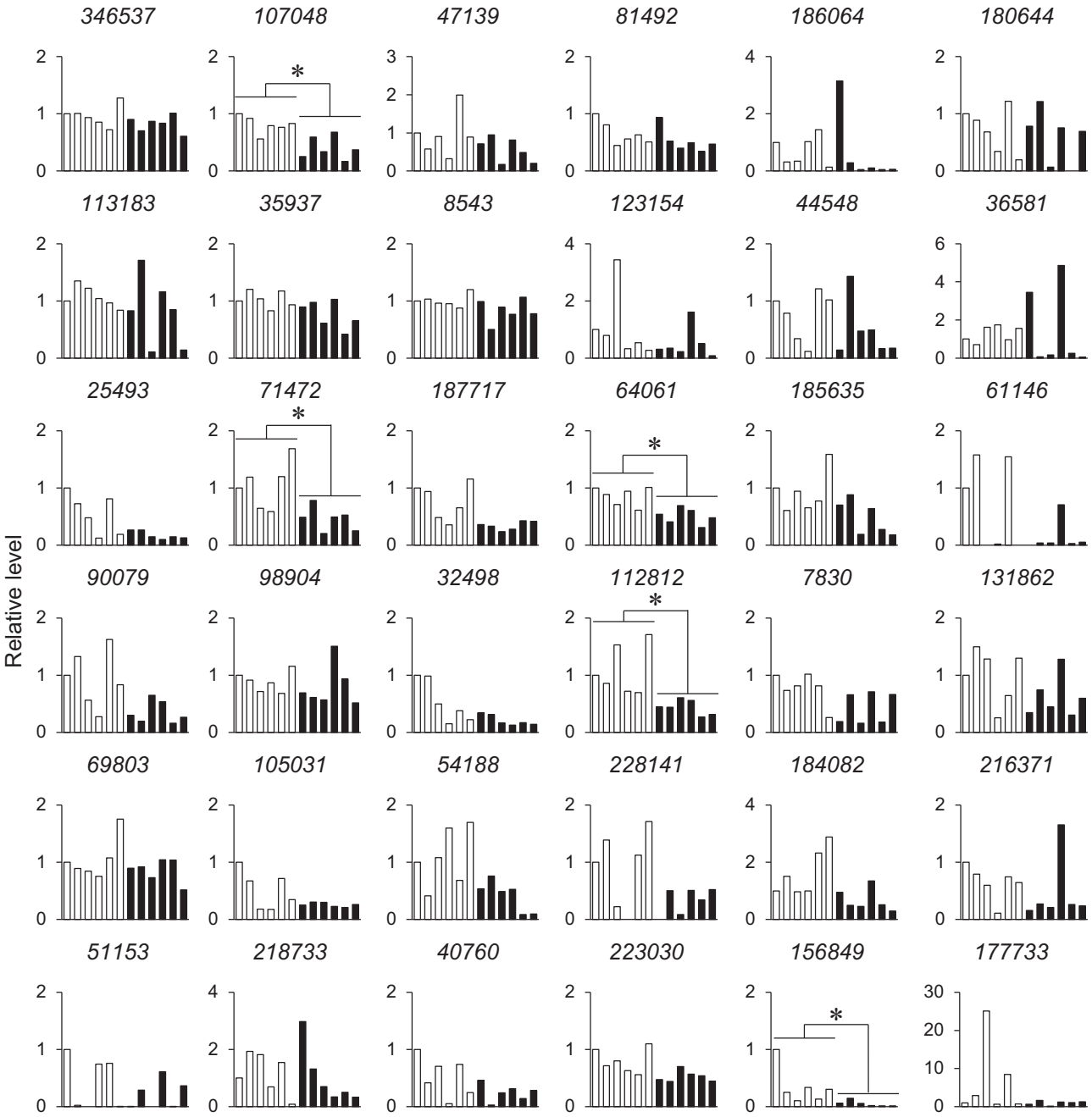
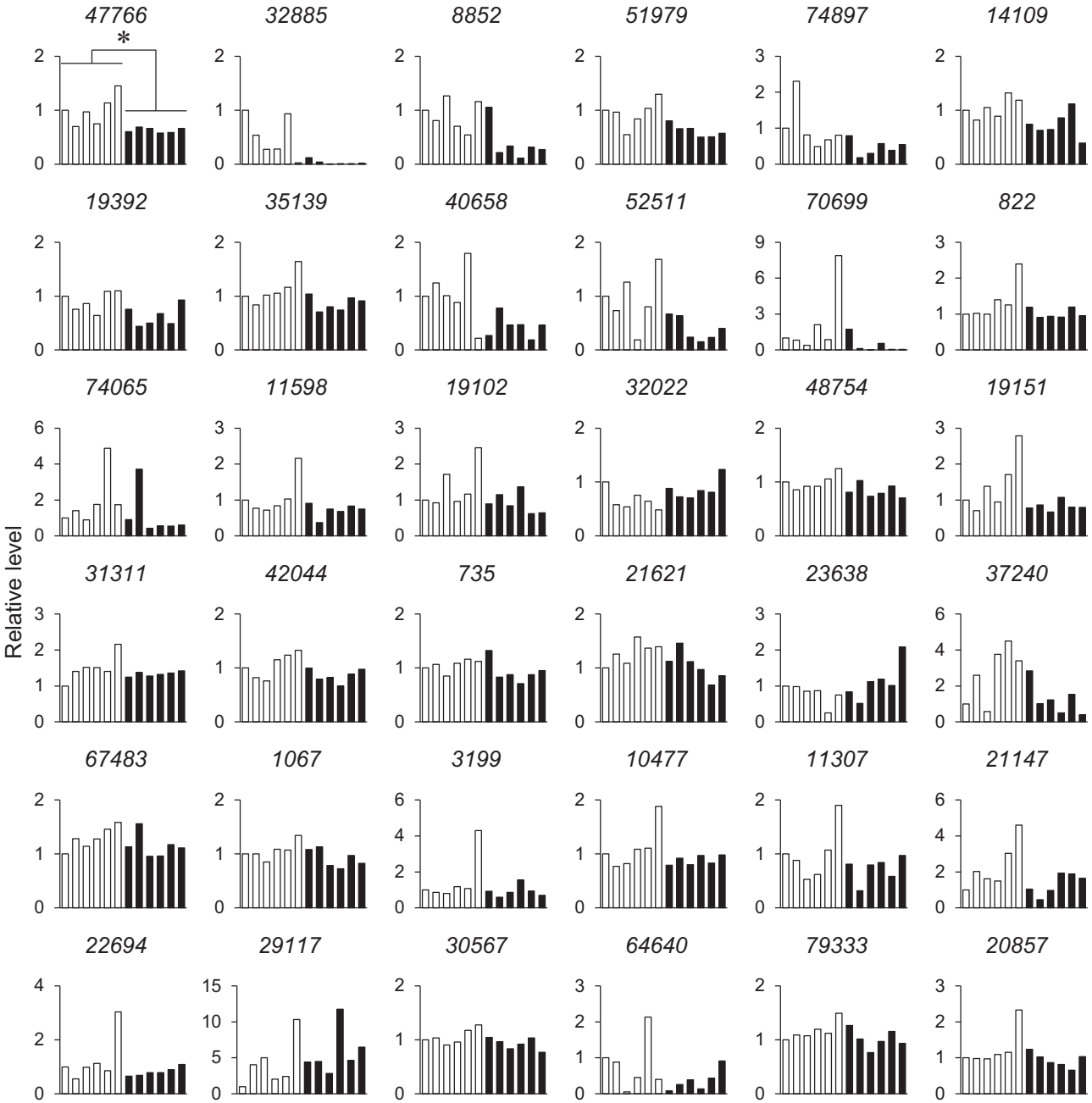

346537
107048
47139
81492
186064
180644
113183
35937
8543
123154
44548
36581
25493
71472
187717
64061
185635
61146
90079
98904
32498
112812
7830
131862
69803
105031
54188
228141
184082
216371
51153
218733
40760
223030
156849
177733
Relative level
*

Fig. 4


47766
*
32885
8852
51979
74897
14109
19392
35139
40658
52511
70699
822
74065
11598
19102
32022
48754
19151
31311
42044
735
21621
23638
37240
67483
1067
3199
10477
11307
21147
22694
29117
30567
64640
79333
20857
Relative level

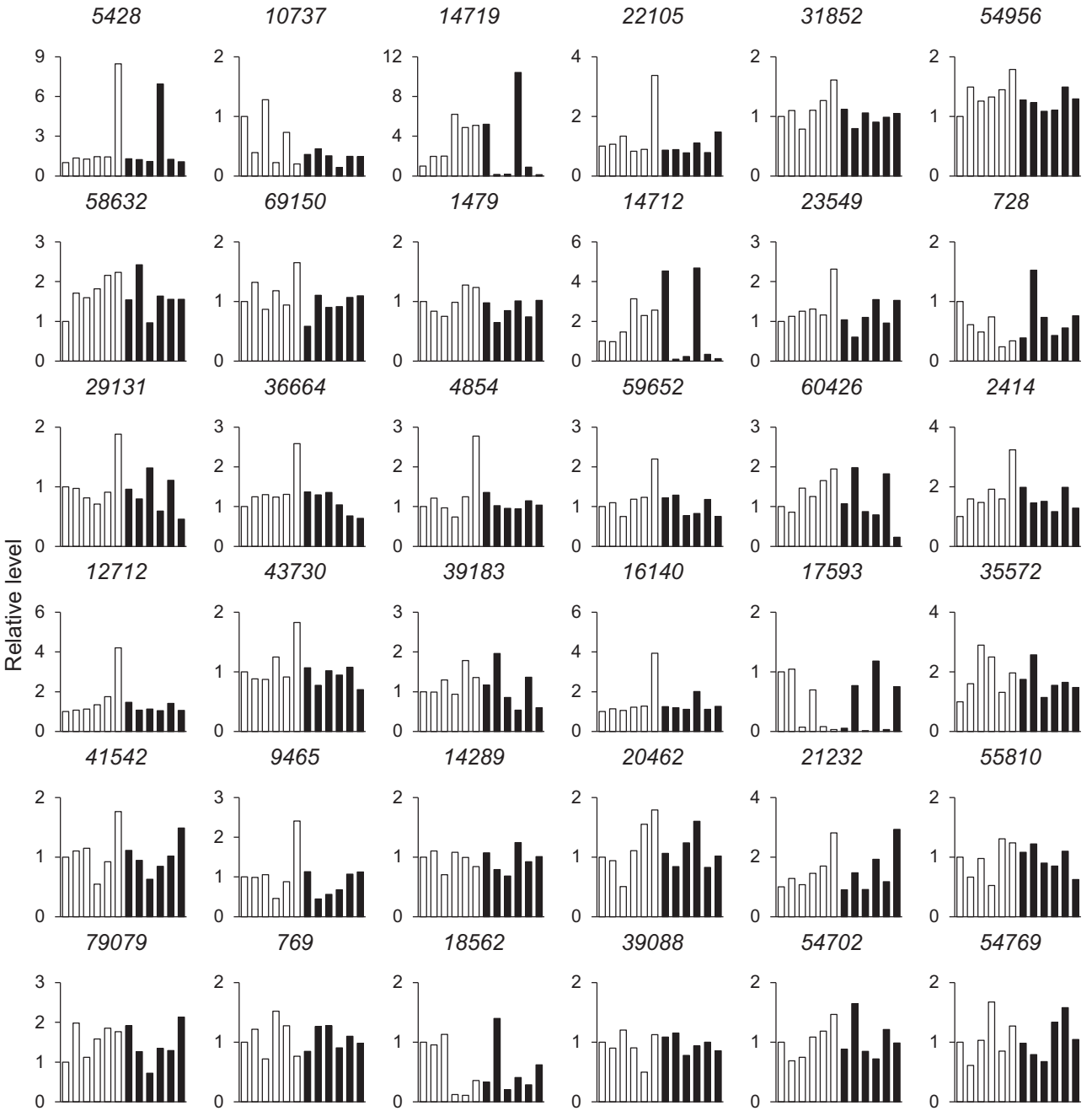
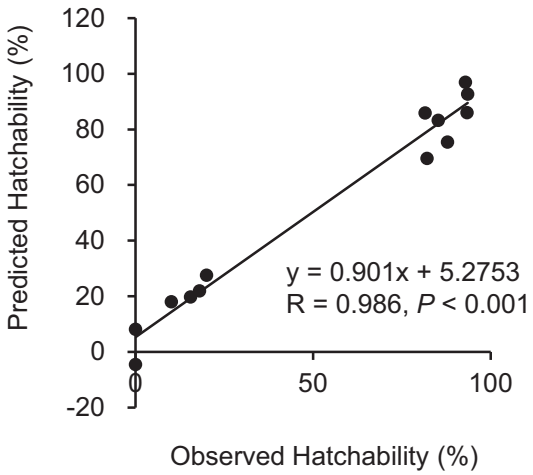

Relative level
5428
10737
14719
22105
31852
54956
58632
69150
1479
14712
23549
728
29131
36664
4854
59652
60426
2414
12712
43730
39183
16140
17593
35572
41542
9465
14289
20462
21232
55810
79079
769
18562
39088
54702
54769

Fig. 5


Predicted Hatchability (%)
120
100
80
60
40
20
0
-20
0
50
100
Observed Hatchability (%)
y = 0.901x + 5.2753
R = 0.986, P < 0.001