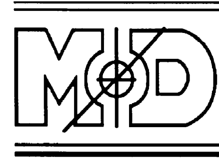
Mechanisms of Development 121 (2004) 971–976



www.elsevier.com/locate/modo

Technical report

Large-scale expression screening by automated whole-mount in situ hybridization

Rebecca Quiring[a], Beate Wittbrodt[a], Thorsten Henrich[a], Mirana Ramialison[a], Carola Burgtorf[b], Hans Lehrach[b], Joachim Wittbrodt[a,*]

[a]*Developmental Biology Programme, European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69012 Heidelberg, Germany*
[b]*Max-Planck-Institut für Molekulare Genetik, Ihnestrasse 73, D-14195 Berlin, Germany*

Received 22 December 2003; received in revised form 12 March 2004; accepted 31 March 2004

Abstract

Gene expression profiling is an important component of functional genomics. We present a time and cost efficient high-throughput whole-mount in situ technique to perform a large-scale gene expression analysis in medaka fish (*Oryzias latipes*) embryos. Medaka is a model system ideally suited for the study of molecular genetics of vertebrate development. Random cDNA clones from an arrayed stage 20 medaka plasmid library were analyzed by whole-mount in situ hybridization on embryos of three representative stages of medaka development. cDNA inserts were colony PCR amplified in a 384-format. The PCR products were used to generate over 2000 antisense RNA digoxigenin probes in a high-throughput process. Whole-mount in situ hybridization was carried out in a robot and a broad range of expression patterns was observed. Partial cDNA sequences and expression patterns were documented with BLAST results, cluster analysis, images and descriptions, respectively; collectively this information was entered into a web-based database, 'MEPD' (http://www.embl-heidelberg.de/mepd/), that is publicly accessible.
© 2004 Elsevier Ireland Ltd. All rights reserved.

Keywords: Oryzias latipes; Partial cDNA sequencing; Expressed Sequence Tags; Screening; Automated; Whole-mount in situ hybridization; Gene expression; High-throughput; Database; Expression profile; Embryogenesis; Organogenesis

1. Introduction

The understanding of the molecular mechanisms underlying a given developmental process has profited from gene expression studies by in situ analysis. This approach has been instrumental in the isolation of a large number of novel genes that are expressed in a specific spatial and temporal manner. It is a complementary approach to systematic mutagenesis screens since genes with no obvious phenotype or functionally redundant genes will be largely inaccessible to mutant screens. In situ analysis is an important contribution to micro-array analysis as well, since the temporal and quantitative data set is complemented by the spatial expression profile.

Furthermore, novel region-, tissue-, organ- and cell-specific markers are identified, facilitating a variety of studies in wild-type embryos and are especially valuable as molecular markers for the analysis of mutant embryos from mutagenesis screens. Genes with interesting expression patterns are readily sequenced and cloned from the cDNA library. In addition, the combination of cDNA sequence and expression pattern provides the scientific community with a detailed analysis of orthologs of genes identified in other organisms.

Large-scale in situ hybridization screens have been carried out for a number of organisms (Gawantka et al., 1998; Kawashima et al., 2000; Kudoh et al., 2001; Lynch et al., 1995; Neidhardt et al., 2000; Thisse et al., 2000). In some cases, these screens were designed to focus on particular aspects (e.g. secreted proteins) (Christiansen et al., 2001; Crosier et al., 2001; Nguyen et al., 2001; Thut et al., 2001; Yoda et al., 2003).

* Corresponding author. Tel.: +49-6221-387-576; fax: +49-6221-387-166.
E-mail address: jochen.wittbrodt@embl-heidelberg.de (J. Wittbrodt).

0925-4773/$ - see front matter © 2004 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.mod.2004.03.031

The high fecundity and the external development of the completely transparent embryos make small fresh water teleosts particularly suited for large-scale in situ analysis (Kudoh et al., 2001; Thisse et al., 2000). Since, even at later stages of development, the embryos are still highly transparent, a unique opportunity is given to look at genes expressed in internal organs during organogenesis.

In medaka (Wittbrodt et al., 2002), in situ screens have been reported (Henrich and Wittbrodt, 2000; Nguyen et al., 2001) and large-scale whole-mount in situ screens are in progress (reviewed in Shima et al., 2003 and the present work).

Here, we describe an efficient strategy for screening a large number of genes for their expression pattern by whole-mount in situ hybridization in medaka embryos. Using cDNA libraries arrayed in microtiter plates, high-throughput processing of RNA probes and automated whole-mount in situ hybridization, this technique is fast, reliable and cost efficient.

2. Results

The cDNA library was prepared by directional cloning of cDNA inserts from stage 20 medaka embryos into a plasmid vector and bacterial transformation. Single clones were arrayed by a robot into 384-well microtiter plates. The initial complexity of the library was two million independent clones. The cDNA inserts were flanked by promoters for SP6 and T7 polymerase allowing antisense probe preparation and cDNA sequencing, respectively; the PCR primers were designed to include both promoters in the PCR product. 4224 inserts were PCR amplified by colony PCR and confirmed by gel electrophoresis. Two percent of the analyzed clones contained no inserts; the average insert length was between 1 and 2 kb.

Over 2000 digoxigenin-labeled RNA antisense probes were transcribed directly from the non-purified PCR products and purified using a commercially available 96-format kit.

For the whole-mount in situ analysis the embryonic stages 18, 24 and 32 (Iwamatsu, 1994) were chosen to represent key events during vertebrate development namely: neurulation, somitogenesis and organogenesis, thus covering early as well as late embryonic development.

The in situ hybridization of 96 probes was carried out simultaneously in a robot (InsituPro, Intavis AG), and all three stages of medaka embryos were hybridized together for each probe. To date all of the probes have been hybridized.

A broad range of staining patterns was observed ranging from examples where the staining is observed exclusively in one organ or a few cells, to highly complex patterns in which several but discrete tissues express the associated gene (Fig. 1).

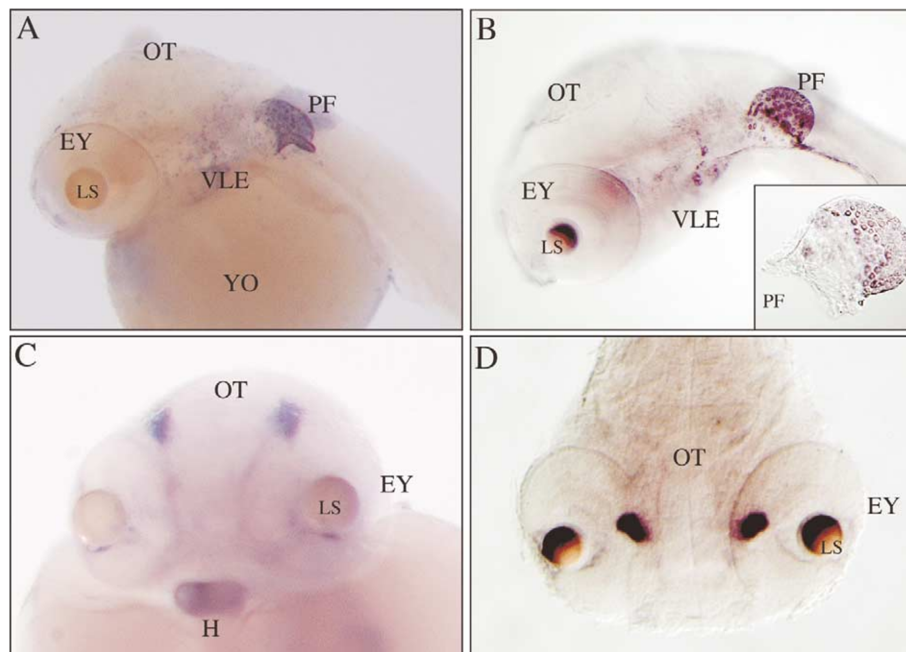


Fig. 1. Examples for the expression patterns observed. (A) and (B) Lateral view, anterior is to the left. (C) Frontal view. (D) Dorsal view of the embryo. (A) and (C) Show original pictures taken from the MEPD. (B) and (D) Show pictures of manual whole-mount in situ, these embryos were cleared in glycerol and removed from yolk. (A) and (B) Region specific expression of clone 631-135-03-D in the epidermis of the pectoral fin blade as well as in ventrolateral epidermis (stage 32). Insert in (B) shows dissected pectoral fin; note the staining in the epidermis of the blade. (C) and (D) Expression of clone 631-136-05-D in the heart as well as in a few cells at the rostralmost pole of the optic tectum in the area where optic fibers of the retinal ganglion cells enter the tectum (stage 32). EY, eye; H, heart; LS, lens; OT, optic tectum; PF, pectoral fin; VLE, ventrolateral epidermis; YO, yolk.

Systematic 5′ and 3′ end sequencing of all cDNA clones is in progress and the sequences and expression patterns are accessible on a web-based database, 'MEPD' (Medaka Expression Data Base), at http://www.embl-heidelberg.de/mepd/ (Henrich et al., 2003). The expressed sequence tags (ESTs) are clustered upon entry into the database and blasted against public databases. So far, 922 clones have been sequenced and clustered to represent 694 genes. Based on amino acid sequence comparison, 140 of these genes have no significant homology hit in the public databases.

Currently, the expression patterns from the in situ screen are being documented by representative pictures at each embryonic stage as well as by descriptions, according to a comprehensive ontology of anatomical terms; collectively this data is entered into the database. The anatomical terms following gene ontology have been submitted to: http://obo.sourceforge.net by Henrich and further details on the ontology will be published elsewhere. The database contains to date 395 sequenced genes with fully annotated expression patterns; these genes were amenable for calculations and gave the following results: 42.5% of the genes were expressed specifically in one or a few embryonic structures, in at least one of the embryonic stages; 18% were ubiquitously expressed and 2% gave no staining. The remaining 37.5% were differentially expressed; the genes in this category were expressed either ubiquitously but at different intensities or they were expressed in several structures and tissues (Fig. 2). This project is still ongoing and a final statistical overview will be presented upon completion.

3. Conclusions and perspective

In situ hybridization screening is a robust and straightforward method to examine the expression profile of a large set of genes. The resolution of the pattern analysis can reach cellular levels. Furthermore, this approach is independent of the function of the genes analyzed, therefore, functional redundancy or severe cellular or temporal restrictions which may cause them to be missed in screens for lethal mutant phenotypes, can be circumvented.

With the presented protocol for high-throughput gene identification, it is possible for one person alone to analyze 50 genes per week including probe generation, in situ hybridization, imaging and pattern annotation in the database. This in combination with the sequence information of the analyzed cDNAs should allow rapid identification of developmental regulators. The protocol can easily be adapted to other species (Arendt, personal communication).

In this study, 922 sequences from random picked cDNA clones resulted in 694 gene clusters; of these 591 were singletons. This reflects the high complexity of the arrayed library. We therefore conclude that using a highly complex

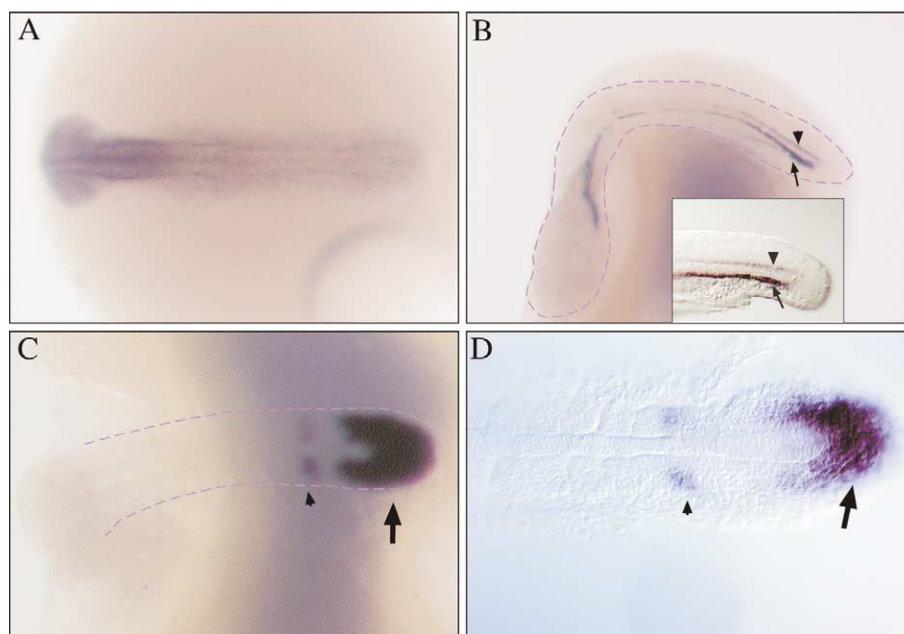


Fig. 2. Examples for differential and specific expression patterns. (A), (C) and (D) Dorsal view. (B) Lateral view, anterior is to the left. (A)–(C) Show original pictures taken from the MEPD. (D) and insert in (B) show pictures of manual whole-mount in situ from the tail region of embryos. (C) and (D) Embryos were cleared in glycerol and removed from yolk; pictures were taken from a Zeiss Axiophot using Normarski optics. (A) Differential expression of clone 631-134-09-I along body axis (stage 18). (B) Expression of clone 631-136-07-H in the hypochord (arrow) and floorplate (arrowhead) (stage 24). (C) and (D) Specific staining of clone 631-134-02-E in the tailbud (arrow); arrowhead indicates staining in the latest pair of forming somites (stage 24).

cDNA library that has not been pre-screened, normalized or subtracted, results in a low redundancy in the screen, if the number of clones analyzed is in the order of 2000–5000.

With the presented protocol for the ongoing whole-mount in situ screen, we have identified novel genes and possible markers to study medaka embryonic development.

Within the context of the medaka genome initiative (Shima et al., 2003), we plan to systematically analyze clones from a medaka unigene library (Himmelbauer, personal communication). In a collaborative project, we are aiming to complete the expression analysis by whole-mount in situ hybridization of the medaka trancriptome within the next 3 years. Based on their expression, cDNAs will then be directly available for gain-of-function experiments by injection and over-expression of the cDNA clone or using synthetic mRNA.

Upon conclusion of data sampling, comparison of expression patterns will define groups of genes with similar complex expression domains, since these genes are often involved in a common process (Niehrs and Pollet, 1999). Such synexpression groups (Gawantka et al., 1998) have a prognostic value on the function of novel components of known pathways as well as of novel genes with no sequence similarity (Niehrs and Pollet, 1999; Onichtchouk et al., 1999; Tsang et al., 2000).

Whole genome sequencing of medaka is in progress. Once genes are identified, this type of combined approach will clearly be an important contribution to functional genomics, which in turn will contribute to an understanding of overall gene function in medaka as one vertebrate model system.

4. Experimental procedures

4.1. Embryos

Embryos were staged according to Iwamatsu (1994). For stage 32, albino (Heino) embryos were used (Loosli et al., 2000). Embryos were fixed over night in 4% paraformaldehyde/2×PTW at 4 °C and subsequently dechorionated manually, washed in PTW according to published procedures (Loosli et al., 1998) and stored in methanol at −20 °C until usage.

4.2. Probe synthesis

An oriented stage 20 cDNA library in the plasmid vector pSPORT1 (Life Technologies) was used. Individual clones were arrayed with a robot into 384-well microtiter plates and grown in liquid microculture.

Fresh replicas from the plated library were made by transferring bacteria from the 384-well microtiter plate (Nalge Nunc International) with a 384-pin replicator (Nalge Nunc International) into a fresh 384-well plate containing 70 μl LB/ampicillin/10% glycerol per well. Replicas were incubated for 16 h without shaking at 37 °C.

For PCR amplification of the cDNA inserts, small aliquots of the microculture were transferred from the replica plates into 384-well PCR-microtest plates (La-Bio-Med) using the 384-pin replicator. Plates were covered with a silicon cover. Two-step colony PCR was performed in a 10 μl volume (vector based primers: M13fwd primer: 5′GCTATTACGCCAGCTGGCGAAAGGGGGATGTG-3′ and 3/86 primer: 5′-CCGGTCCGGAATTCCCGGGT3′). PCR conditions were as follows: an initial denaturation step at 96 °C for 3 min was followed by 30 cycles of two steps: 96 °C for 10 s and 70 °C for 3 min, and a final extension at 72 °C for 8 min. PCR reactions were analyzed by gel electrophoresis in batches of 384.

The PCR products include a promoter for SP6 polymerase allowing the direct transcription of antisense RNA probes. Digoxigenin-labeled antisense RNA probes were synthesized from PCR-amplified templates using SP6 RNA polymerase (Roche). Reactions were carried out in 96-well plates with a 10 μl reaction volume containing 0.5 μl of the PCR reaction, 1 mM of ATP, CTP and GTP each, 0.65 mM UTP/0.35 mM digoxigenin–UTP mix (Roche), 15 units of RNAguard (Amersham Pharmacia), 0.4 units of thermostable inorganic pyrophosphatase (New England BioLabs), 2 mM $MgCl_2$, 20 units SP6 polymerase and 1 μl of 10× SP6 transcription buffer (Roche). Incubation was at 37 °C for 5 h; DNase I treatment was omitted. To reduce the costs of the purification of the riboprobes, the 'Quiaquick PCR purification kit 96' (QIAGEN) was used with the 'QIAvac 96' vacuum device according to manufacturers instructions; the important exception was that in the first step buffer 'RLT' from the 'RNAesy KIT' (QIAGEN) was used. To the eluate (60 μl) 150 μl of Hybridization Buffer was added.

4.3. In situ hybridization

Embryos were rehydrated and processed according to standard procedure (Loosli et al., 1998). Proteinase K digestion was according to stages between 7 min and 1.5 h. In situ hybridization was performed with 96 probes simultaneously in a robot (In situ Pro, Intavis AG). In the robot, the incubation volume for all the subsequent steps of the in situ hybridization protocol was 150 μl. Pre-hybridization was done in Hybridization Buffer (Hybridization Buffer: 50% formamide (Fluka, ultra pure), 50 μg/ml heparin, 0,1% Tween 20, 5 mg/ml torula RNA, 5 × SSC) at 65 °C for 2 h. Denatured probes (1.5 μl of the diluted digoxigenin-labeled antisense RNA probe (see above) in 150 μl Hybridization Buffer) were hybridized for 12 h at 65 °C. The different stages were analyzed together for each probe. All washing steps were performed at 65 °C. The first three washes were done with 50% formamide/2 × SSCT (two washes

for 15 min, 1 wash for 30 min). These were followed by two washes with 2 × SSCT (15 min each) and two washes with 0.2 × SSCT (15 min each). Blocking of the embryos was performed for 2 h with 5% sheep serum/PTW (PTW:PBS, pH 7.5, 0.1% Tween 20). This and all the following steps were performed at room temperature, except the pre-absorption of the anti-Dig-AP F_{ab} fragments to medaka embryonic extracts which was done at 4 °C overnight. The embryos were incubated for 2 h in pre-absorbed anti-Dig-AP F_{ab} fragments at a 1:2000 dilution in PTW. After eight washes for 10 min with PTW, two incubations of 7 min each with Staining Buffer (Staining Buffer: 100 mM Tris–Cl, pH 9.5, 100 mM NaCl, 50 mM $MgCl_2$, 0.1% Tween 20) concluded the hybridization procedure in the robot. As a positive control for the quality of the in situ hybridization procedure, a *Pax6* digoxigenin probe was included. A floppy disc with the program for the in situ robot is available upon request. For the staining, which was done manually at room temperature, the embryos were transferred into 24-well dishes (Nunc). Staining was done for a maximum of 6 h, whereby individual reactions were stopped according to the signal intensity. Staining reaction was stopped by several washes with PTW and post-fixed overnight in 1% paraformaldehyde/PTW at 4 °C. Embryos were photographed with the HC 300Z Digital Camera (Fujix) on a Stemi 2000-C microscope (Carl Zeiss).

4.4. Sequence analysis

5′ cDNA ends were sequenced with T7 and 3′ ends with SP6 primers, respectively, at the EMBL Genomics Core Facility. The average length of read was 620 bp.

4.5. Web-based database

MEPD, the web-based database is described in Henrich et al. (2003) and accessible via http://www.embl-heidelberg.de/mepd/. BLASTN, BLASTX programs (National Center for Biotechnology Information (NCBI), USA) were used to search the ‘nrdb’, (Non-Redundant DataBase, consists of the following databases: SWISSPROT, SWISSNEW, PDB, TREMBL, TREMBLNEW, PIRONLY, GENPEPT, GENPEPTNEW, WORM) and the ‘nrnd’ (Non-Redundant Nucleotide Databases of EMBL + GenBank + DDBJ, without EST’s or STS’s) as well as the ‘dbest’ (Non-Redundant Database of GenBank + EMBL + DDBJ EST Divisions). BLAST results are regularly updated by an automated script (Henrich et al., 2003). Clusters were generated with the use of the blastclust program of the standalone BLAST package.

Acknowledgements

We are grateful to Felix Loosli for helpful comments on this manuscript. We would like to thank Erika Grzebisz and Aldona Nowicka for excellent maintenance of the fish facility. We gratefully acknowledge the expert input of Franck Bourrat. This work was supported by a fellowship of the DFG (C.B.) and through grants of the DFG, European Commission and the HFSPO (JW).

References

Christiansen, J.H., Coles, E.G., Robinson, V., Pasini, A., Wilkinson, D.G., 2001. Screening from a subtracted embryonic chick hindbrain cDNA library: identification of genes expressed during hindbrain, midbrain and cranial neural crest development. Mech. Dev. 102, 119–133.

Crosier, P.S., Bardsley, A., Horsfield, J.A., Krassowska, A.K., Lavallie, E.R., Collins-Racie, L.A., et al., 2001. In situ hybridization screen in zebrafish for the selection of genes encoding secreted proteins. Dev. Dyn. 222, 637–644.

Gawantka, V., Pollet, N., Delius, H., Vingron, M., Pfister, R., Nitsch, R., et al., 1998. Gene expression screening in *Xenopus* identifies molecular pathways, predicts gene function and provides a global view of embryonic patterning. Mech. Dev. 77, 95–141.

Henrich, T., Wittbrodt, J., 2000. An in situ hybridization screen for the rapid isolation of differentially expressed genes. Dev. Genes Evol. 210, 28–33.

Henrich, T., Ramialison, M., Quiring, R., Wittbrodt, B., Furutani-Seiki, M., Wittbrodt, J., Kondoh, H., 2003. MEPD: a medaka gene expression pattern database. Nucleic Acids Res. 31, 72–74.

Iwamatsu, T., 1994. Stages of normal development in the medaka *Oryzias latipes*. Zool. Sci. 11, 825–839.

Kawashima, T., Kawashima, S., Kanehisa, M., Nishida, H., Makabe, K.W., 2000. MAGEST: MAboya Gene Expression patterns and Sequence Tags. Nucleic Acids Res. 28, 133–135.

Kudoh, T., Tsang, M., Hukriede, N.A., Chen, X., Dedekian, M., Clarke, C.J., et al., 2001. A gene expression screen in zebrafish embryogenesis. Genome Res. 11, 1979–1987.

Loosli, F., Koster, R.W., Carl, M., Krone, A., Wittbrodt, J., 1998. Six3, a medaka homologue of the *Drosophila* homeobox gene sine oculis is expressed in the anterior embryonic shield and the developing eye. Mech. Dev. 74, 159–164.

Loosli, F., Koster, R.W., Carl, M., Kuhnlein, R., Henrich, T., Mucke, M., et al., 2000. A genetic screen for mutations affecting embryonic development in medaka fish (*Oryzias latipes*). Mech. Dev. 97, 133–139.

Lynch, A.S., Briggs, D., Hope, I.A., 1995. Developmental expression pattern screen for genes predicted in the *C. elegans* genome sequencing project. Nat. Genet. 11, 309–313.

Neidhardt, L., Gasca, S., Wertz, K., Obermayr, F., Worpenberg, S., Lehrach, H., Herrmann, B.G., 2000. Large-scale screen for genes controlling mammalian embryogenesis, using high-throughput gene expression analysis in mouse embryos. Mech. Dev. 98, 77–93.

Nguyen, V., Joly, J.S., Bourrat, F., 2001. An in situ screen for genes controlling cell proliferation in the optic tectum of the medaka (*Oryzias latipes*). Mech. Dev. 107, 55–67.

Niehrs, C., Pollet, N., 1999. Synexpression groups in eukaryotes. Nature (London) 402, 483–487.

Onichtchouk, D., Chen, Y.G., Dosch, R., Gawantka, V., Delius, H., Massague, J., Niehrs, C., 1999. Silencing of TGF-beta signaling by the pseudoreceptor BAMBI. Nature (London) 401, 480–485.

Shima, A., Himmelbauer, H., Mitani, H., Furutani-Seiki, M., Wittbrodt, J., Schartl, M., 2003. Fish genomes flying. Eur. Mol. Biol. Org. Rep. 4, 121–125.

Thisse, B., Fuerthauer, M., Steffan, T., Heyer, V., Thisse, C., 2000. Large-scale in situ hybridization screen: a genome project. Zebrafish Development and Genetics, Cold Spring Harbor, NY.

Thut, C.J., Rountree, R.B., Hwa, M., Kingsley, D.M., 2001. A large-scale in situ screen provides molecular evidence for the induction of eye anterior segment structures by the developing lens. Dev. Biol. 231, 63–76.

Tsang, M., Kim, R., de Caestecker, M.P., Kudoh, T., Roberts, A.B., Dawid, I.B., 2000. Zebrafish nma is involved in TGFbeta family signaling 28, 47–57.

Wittbrodt, J., Shima, A., Schartl, M., 2002. Medaka—a model organism from the far east. Nat. Rev. Genet. 3, 53–64.

Yoda, H., Momoi, A., Esguerra, C.V., Meyer, D., Driever, W., Kondof, H., Furutani-Seiki, M., 2003. An expression pattern screen for genes involved in the induction of the posterior nervous system of zebrafish. Differentiation 71, 152–162.