Microarray-Based Analysis of Early Development in Xenopus laevis

Curtis R. Altmann,* Esther Bell,* Alex Sczyrba,† ‡ Jason Pun,* Stefan Bekiranov,† Terry Gaasterland,† and Ali H. Brivanlou*1

*Laboratory of Molecular Vertebrate Embryology, †Laboratory of Computational Genomics, The Rockefeller University, 1230 York Avenue, New York, New York 10021; and ‡Faculty of Technology, Bielefeld University, 33594 Bielefeld, Germany

In order to examine transcriptional regulation globally, during early vertebrate embryonic development, we have prepared Xenopus laevis cDNA microarrays. These prototype embryonic arrays contain 864 sequenced gastrula cDNA. In order to analyze and store array data, a microarray analysis pipeline was developed and integrated with sequence analysis and annotation tools. In three independent experimental settings, we demonstrate the power of these global approaches and provide optimized protocols for their application to molecular embryology. In the first set, by comparing maternal versus zygotic transcription, we document groups of genes that are temporally regulated. This analytical approach resulted in the discovery of novel temporally regulated genes. In the second, we examine changes in gene expression spatially during development by comparing dorsal and ventral mesoderm dissected from early gastrula embryos. We have discovered novel genes with spatial enrichment from these experiments. Finally, we use the prototype microarray to examine transcriptional responses from embryonic explants treated with activin. We selected genes (two of which are novel) regulated by activin for further characterization. All results obtained by the arrays were independently tested by RT-PCR or by in situ hybridization to provide a direct assessment of the accuracy and reproducibility of these approaches in the context of molecular embryology. © 2001 Academic Press

Key Words: microarray; development; Xenopus laevis; embryology; vertebrates.

INTRODUCTION

The development of high throughput approaches in molecular biology, where a large number of genes can be analyzed simultaneously, has provided researchers with the unique opportunity to look at biological responses globally as opposed to one gene or one pathway at a time (Schena et al., 1996b). This approach complements genetic approaches (when available), and opens powerful genome-wide analysis to be applied to nongenetic systems. While a great deal of effort and interest has been spent applying these technologies to human and mouse models as well as invertebrate systems, the utility of this type of approach has not yet been demonstrated for a vertebrate developmental model system. We built a robotic device for preparing DNA microarrays (Brown and Botstein, 1999) and prepared a prototype Xenopus laevis microarray. Xenopus embryos provide an ideal system for the study of early vertebrate development because of the abundance of biological material (up to a few thousand embryos/day/female), and because embryonic development can be followed from fertilization onward. Xenopus laevis offers advantages for the study of the molecular basis of embryonic cell-fate decisions, and it is amenable to microsurgery and microinjections.

During embryonic development, signals from one group of cells influence cell-fate decisions of other cells in a process known as induction. These inductive signals act within the embryo both in the context of time and space to induce differentiation of various cell types. Differentiation of cells is the result of stable changes in gene expression (which in most circumstances is not reversible) and the expression of cell type-specific genes. Current methods for analyzing induction and differentiation rely on a variety of molecular techniques, including reverse transcription of the mRNA message and polymerase chain reaction (RT-PCR) using primers that amplify previously defined cell type-specific genes or “markers” (such as the Neural Cell...
Adhesion Molecule, NCAM, for neural fates and cytokeratin for epidermal fates). There are about 100 cell typespecific molecular markers reported to date and employed by various laboratories to study embryonic induction. While extremely sensitive and useful, the number of markers usually assayed in a typical experiment is about 20. More importantly, for a given assay, the researcher must make a subjective selection of markers to test and thus may miss potentially relevant genes. While this limitation may not be as important when examining for the formation of a particular tissue type, it can become a critical factor when a gene of unknown function is being studied. The use of a microarray-based approach provides an objective method for examining gene function globally and represents an obvious choice of using information generated by the different ongoing EST (expressed sequence tag) projects.

In the studies presented here, we sought to demonstrate and evaluate the applicability of microarray-based approaches to molecular embryology. Initially, we have selected three applications of this technology to explore in Xenopus development: changes in gene expression with temporal, spatial, and inductive events during early embryogenesis.

**TABLE 1**

Summary of Clones on Xenopus Array

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Full-length clones</td>
<td>297</td>
<td>34%</td>
</tr>
<tr>
<td>Partial</td>
<td>159</td>
<td>18%</td>
</tr>
<tr>
<td>Duplicate</td>
<td>2× 42</td>
<td>&lt;5%</td>
</tr>
<tr>
<td>3× 7</td>
<td>&lt;1%</td>
<td></td>
</tr>
<tr>
<td>&gt;3× 5</td>
<td>&lt;1%</td>
<td></td>
</tr>
<tr>
<td>EST</td>
<td>317</td>
<td>36%</td>
</tr>
<tr>
<td>Xenopus EST</td>
<td>210</td>
<td>24%</td>
</tr>
<tr>
<td>Unique EST</td>
<td>107</td>
<td>12%</td>
</tr>
</tbody>
</table>

**MATERIALS AND METHODS**

**Clone Preparation**

Bacteria transformed by plasmids encoding Xenopus early gastrula library (Weinstein et al., 1998) were plated at low density. The vector in these libraries is pCS2+ (Turner and Weintraub, 1994). Individual clones were isolated manually and grown overnight in 1.2 ml Terrific Broth in Qiagen 96-well deep-well blocks. Plasmids were purified by using Qiagen Turbo 96 kit on a Qiagen Biorobot 9600. Purified plasmids were eluted in 150 μl buffer EB. The DNA concentration in each sample was ~0.2 μg/μl as determined by random sampling.

**Sequencing and Sequence Analysis**

Clones were sequenced on ABI 3700 sequencers by using Big Dye chemistry. The sequencing primer used was 5′-CTT GAT TTA GGT GAC ACT ATA G-3′. Sequences were analyzed and organized by using the automated sequence-annotation tool MAGPIE (Gaasterland et al., 2000). The MAGPIE sequence-annotation system was configured to compare automatically each clone sequence with existing protein and DNA sequences from SwissProt and Genbank, and to search for Prosite (Bairoch, 1991), PFAM (Sonhammer et al., 1997), and BLOCKS (Henikoff and Henikoff, 1996) protein motifs in translated coding regions. MAGPIE synthesizes the resulting data into a comprehensive view ordered from strongest to weakest evidence. The system presents automated decisions about putative function and the supporting evidence to users for further manual annotation via web forms.

**PCR Amplification**

Individual clones were amplified by PCR using the upstream primer (defined above), and a downstream primer 5′-TAT AGT TCT AGA GGC TCG AGA G-3′. PCR conditions were: 35 rounds, 94°C for 30 s, 55°C for 1 min, and 72°C for 2 min. To confirm amplification, samples were analyzed on 1% agarose gels and the remainder isopropanol precipitated and washed with 70% EtOH and allowed to dry overnight at room temperature. The samples were resuspended in 30 μl of 3× SSC and 4 μl was transferred to a clean 96-well plate for arraying.

**Prototype Microarray Preparation**

cDNA microarrays were prepared by using quill-type pins manufactured according to Stanford specifications (DeRisi et al., 1996) onto polylysine-coated slides (Corning). Eight 96-well plates of random clones prepared as described above were printed in duplicate together with 96 previously characterized clones available in our laboratory. A complete list of all clones printed on our prototype microarray can be found on our web site: http://arrays.rockefeller.edu/xenopus. Arrays were stored at room temperature for 1 week before processing further according to previously published protocols (DeRisi et al., 1996).

**Probe Preparation**

Probes were prepared from either poly(A) − selected RNA or total RNA according to published protocols (DeRisi et al., 1996). RNA was prepared from embryos as described (Altmann et al., 1997) and...
poly(A)⁺ RNA selected by using Oligotex mRNA purification system (Qiagen). In brief, 1–2 µg of poly(A)⁺ RNA, or 15 µg of total RNA were used in RT reactions primed with oligo(dT)₁₈–₂₂ using Superscript II (Gibco/BRL) according to manufacturer's instructions in a 30-µl final volume. Either Cy³ or Cy⁵ dUTP (Amersham) were included at 15 mM, unlabeled dTTP was 10 mM, while dATP, dGTP, and dCTP were 25 mM. Reactions were incubated at 42°C for 2 h. RNA was degraded by the addition of 15 µl 0.1 N HCl and diluted to 500 µl with TE and 1 µg Poly(dA)+ was added. Unincorporated nucleotides and dyes were removed by filtration in Microcon-30 filters. Samples were washed twice with 500 µl TE before being combined, concentrated, and dried. Combined samples were resuspended in 15 µl 3× SSC containing 0.3% SDS and filtered through a prewet Millipore filter to remove particulate matter. For the analysis of dorsal and ventral marginal zones, sufficient material was dissected from Xenopus embryos to prepare 15 µg total RNA for each sample.

FIG. 1. (A) Correlation plot of the ratio of the ratios plotted against the channel intensity. The correlation of duplicate experiments is calculated as described. This plot compares the expression ratios obtained for two independent array determinations. Four similar plots can be generated from the intensity values of each of the four channels. As the intensity of the signal increases, the ratio/ratio values cluster closer to 1. (B) Array targets showing increased and decreased RNA levels grouped into different categories. The 48 up-regulated and 45 down-regulated genes were categorized as described for TIGR Rat arrays (http://www.tigr.org/docs/tigr-scripts/egad_scripts/role_report.spl) based on sequence analysis and are shown in the separate pie graphs. Color coding for the different groupings is shown on the right.
Hybridization and Scanning

The probes prepared above in a final volume of 15 μl were heated to 100°C for 3 min and applied to the Xenopus cDNA microarray, covered with a 22 × 22-mm glass coverslip (Fisher no. 12-542B) and sealed in a hybridization chamber (DeRisi et al., 1996). The samples were incubated overnight at 65°C. Following hybridization, the array was washed for 10 min each at room temperature in 1× SSC containing 0.3% SDS, followed by 0.2× SSC and a 0.05× SSC final wash. The slides were dried by centrifugation and stored in the dark at room temperature before scanning. Scanning was performed on a ScanArray 3000 confocal laser scanner (General Scanning, Inc.) to generate two 16-bit Tif images corresponding to the cy3 and cy5 channels, respectively.

Scanning and Data Analysis

The two 16-bit greyscale TIF images from each scanned hybridized microarray slide were loaded into Scanalyze version 2.44 (M. Eisen, Stanford University) and gridded according to software instructions. The image files and Scanalyze data file were uploaded via a web-based interface into the TANGO system (Gaasterland et al., manuscript in preparation). For each uploaded hybridization, TANGO carries out a preconfigured automated analysis of the data. After saving the original intensity data in a database, the system computes and saves normalized ratios of intensity values for each microarray spot, generates a normalized scatterplot of channel 1 versus channel 2 intensity for each spot, and combines the two TIF files into a single color image. Each point in the scatterplot and each spot in the image corresponds to a spotted clone and is hyperlinked to the MAGPIE sequence analysis annotation display for the clone. The TANGO relational database stores intensities, ratios, functional annotations of clone sequences, locations of clones on each type of printed microarray, locations of clones in plates from which arrays were printed, vocabulary-controlled and free text descriptions of probe preparation, tissue or cell source, hybridization conditions, and other attributes relevant to the analysis of each hybridization dataset. The TANGO relational schema is adapted from M. Cherry (Stanford); data definitions are consistent with the international standards in preparation (A. Brazma, EBI). Users can query the system to retrieve single-experiment, two-experiment, and multi-experiment reports, expression level reports for single and multiple genes, and datasets for input to clustering algorithms.

The TANGO system enables users to quantify reproducibility of repeated experiments with three operations: a novel ratio-of-intensity-ratios plot, Pearson correlation, and fractional difference. Plotting the ratio of intensity ratios for each spot in two hybridization experiments versus all four-channel intensities elucidates spots that have a ratio-of-ratios unequal to 1 and reveals the intensity range, if any, above which most spots are close to 1. The Pearson correlation coefficient (Gordon, 1981) is calculated according to the following formula:

\[ r = \frac{1}{n-1} \sum \left( X_i - \langle X \rangle \right) \left( Y_i - \langle Y \rangle \right) \]

where \( i \) denotes a spot on an array with \( N \) spots; \( X_i \) is the normalized intensity ratio for spot \( i \) in one experiment and \( Y_i \) is the normalized intensity ratio for spot \( i \) in the other experiment.

TABLE 2

<table>
<thead>
<tr>
<th>Sample id</th>
<th>Description</th>
<th>Array ratio</th>
<th>PCR ratio</th>
<th>Sign agreement</th>
<th>Magnitude agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>S10-1-B11</td>
<td>MAP kinase</td>
<td>3.5</td>
<td>3.3</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>S10-1-E7</td>
<td>Xenopus EST</td>
<td>6.1</td>
<td>5.8</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>S10-1-H7</td>
<td>heterogeneous nuclear RNP A2/B1</td>
<td>0.4</td>
<td>0.6</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-2-A1</td>
<td>glycerol-3-phosphate dehydrogenase</td>
<td>0.4</td>
<td>1.2</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>S10-2-A5</td>
<td>geminin H</td>
<td>3.2</td>
<td>2.8</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>S10-2-B10</td>
<td>pituitary tumor-transforming 1</td>
<td>4.5</td>
<td>1.6</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-2-E12</td>
<td>Nucleolar protein (KKE/D repeat)</td>
<td>0.3</td>
<td>0.8</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>S10-2-E7</td>
<td>translocon-associ. Prot., β subunit</td>
<td>0.2</td>
<td>1.2</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>S10-2-F11</td>
<td>NADH dehydrogenase subunit 2</td>
<td>4.0</td>
<td>1.8</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-3-G6</td>
<td>hnRNP G</td>
<td>0.3</td>
<td>0.5</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-4-C1</td>
<td>Xenopus EST</td>
<td>0.4</td>
<td>0.4</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-4-C2</td>
<td>XGata4</td>
<td>0.3</td>
<td>0.1</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-4-D3</td>
<td>met-10+ protein</td>
<td>3.2</td>
<td>1.1</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>S10-4-F7</td>
<td>hsp-90</td>
<td>0.3</td>
<td>0.7</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-5-D5</td>
<td>cold-inducible RNA-binding protein</td>
<td>0.3</td>
<td>0.4</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-6-G4</td>
<td>X1hnRNPL</td>
<td>0.3</td>
<td>0.8</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-6-H3</td>
<td>Nucleolin</td>
<td>0.3</td>
<td>0.5</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-8-A11</td>
<td>Sialomucin</td>
<td>3.6</td>
<td>1.4</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-8-C11</td>
<td>KIAA0592 related</td>
<td>0.3</td>
<td>4.3</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>S10-8-D4</td>
<td>Xenopus EST</td>
<td>2.5</td>
<td>2.3</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-8-F9</td>
<td>hnRNP G</td>
<td>0.3</td>
<td>0.6</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>S10-8-H10</td>
<td>cytochrome c oxidase subunit II</td>
<td>3.8</td>
<td>1.4</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Note. Ratio = maternal/stage 10, values >1 indicate higher maternal expression.
normalized intensity ratio for spot \( i \) in the second experiment; \(<X>\) denotes the average of the \( X \) values (likewise for \( Y \)), \( \sigma_x \) and \( \sigma_y \) are the standard deviations for \( X_i \) and \( Y_i \), respectively. The average fractional difference, \( f_d \), between repeated experiments is calculated according to the following formula:

\[
(f_d) = \frac{1}{N} \sum_{i=1}^{N} \frac{|X_i - Y_i|}{\frac{1}{2} (X_i + Y_i)}
\]

Selection of Clones for Further Analysis and RT-PCR

To examine the change of gene expression during development, two reciprocal hybridization experiments were performed. In the first, RNA isolated from embryos at the 32-cell stage (stage 6, (Nieuwkoop and Faber, 1956)) representing the maternal pool of RNA was labeled with Cy3 dUTP and RNA from midgastrula, when zygotic transcription has begun (stage 11), was labeled with Cy5 dUTP. In the second experiment, the dyes were reversed and stage 6 RNA was labeled with Cy5 and gastrula (stage 11) RNA was labeled with Cy3. Each sample pair was then hybridized to Xenopus microarrays containing two spots of each clone under identical conditions, washed, and scanned. This generates up to four data points (two each for each probe pair) for each spotted clone.

To select clones for further analysis, the results were sorted by intensity ratios into three groups: two-fold up-regulated, two-fold down-regulated, and less than two-fold changes. Clones that were two-fold regulated for all four data points and had small standard deviations (\(<5\%) were selected and PCR primers designed. The top 10 up-regulated and 10 down-regulated genes were initially selected for quantitation by RT-PCR (Wilson and Hemmati-Brivanlou, 1995). In addition, most of the target sequences without homologies to genes in public databases were analyzed. To ensure that PCR analysis was performed in the linear range, the number of PCR cycles was varied from 15 to 25. The PCR products were separated on 6% nondenaturing polyacrylamide gels and exposed and examined on a Molecular Dynamics PhosphorImager to ensure a linear readout of the radioactive signal. In order to normalize the signal between the two samples, we also amplified histone H4. Since similar amounts of starting material were used for the RT reactions, we observed similar levels of histone expression (ratio of 1.6). Using the ratio of histone expression, we normalized the values for the other samples.

Embryonic Explants and in Situ Hybridization

Tissue explants were isolated from blastula (stage 8.5–9.0) embryos and cultured in the presence or absence of recombinant activin protein (Thomsen et al., 1990). Explants were collected and RNA isolated at midgastrula stage as described previously (Altmann et al., 1997). Array analysis was as described above. In situ hybridization was performed as described (Harland, 1991). Explants for sectioning were embedded in 20% gelatin/PBS, fixed overnight in 4% paraformaldehyde and sectioned using a vibratome at either 50 or 100 \( \mu m \).

RESULTS

The goal of the work presented here was to examine the utility of microarray-based approaches to gene discovery and the study of global gene expression during early vertebrate development.

Clone and Sequence Analysis of Genes Printed on the Prototype Microarray

A total of 768 random clones obtained from a directionally cloned gastrula expression library (Weinstein et al., 1998) were sequenced from the 5’ end. The sequences of each clone were compared via MAGPIE (Gaasterland and Sensen, 1996) with the sequence of every other clone and with the public sequence databases. The quality of the clone set was computed by detecting full-length clones, duplicated clones, and unique clones. Clones were considered duplicates if they aligned with 90% identity over the full length of one of the sequences; a clone was considered full length if an alignment between the translated coding region and a protein started at the beginning of the aligned...
protein. After further inspection by the user, we identified 297 full-length clones, 159 partial clones, and 55 duplicated genes. Of the remaining sequences, 107 had no significant homologies to public databases, while 210 show significant homology to Xenopus EST sequences (Table 1).

To further analyze the sequences and to identify potentially interesting clones (i.e., novel or highly conserved), we searched a variety of databases separately by blastn and blastx. This analysis revealed a number of genes from plants and fungi that had not been previously found in animals and genes that had been cloned only in invertebrates. All of these sequences (as well as microarray data) are publicly available at http://arrays.rockefeller.edu/xenopus/magpie.html and have been submitted to other public databases.

Application of the Prototype Microarray in Temporal Assays

To assess the utility of Xenopus microarrays for studies of early vertebrate development and to determine the accuracy of microarray-based measurement, we compared gene expression from embryos prior to the onset of zygotic transcription at the midblastula transition (pre-MBT; Kimelman et al., 1987) to gastrula-stage embryos undergoing zygotic transcription (post-MBT). For this first set of studies, since the source of biological material is abundant, we used 1.5 μg of poly(A)+ RNA isolated from 100 32-cell embryos (stage 6; Nieuwkoop and Faber, 1956) and early gastrulae (stage 11) to generate probes labeled with either Cy3 or Cy5. These probes were hybridized competitively against the genes on the array. To minimize effects resulting from differences in Cy dye incorporation, the dyes during probe labeling were reversed for a duplicate hybridization. Thus, for each gene, four data points were generated.

After hybridization and washing as described, the array slides were scanned and intensity data extracted from the resulting raw images by using Scanalyze (M. Eisen, Stanford University). To visualize and store the data, we developed a web-based system that automated this process. This web-based system can be accessed at http://arrays.rockefeller.edu/xenopus/magpie.html and has been submitted to other public databases.

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.
Interestingly, a larger proportion of the targets that were unknown or hypothetical were found in the set of down-regulated genes (33% vs. 19%).

To independently confirm the results obtained by the array and as an assessment of confidence in the array approach, we selected a number of genes for further examination by PCR. These genes included all of the nonhomologous Xenopus ESTs as well as the most highly reproducible targets. Specific primers were used to compare the microarray results directly with RT-PCR-based approaches. To ensure that amplification was in the linear range, 18 to 26 cycles of PCR were performed, and to ensure a linear recording, the samples were analyzed on a Molecular Dynamics PhosphorImager. RT-PCR analysis of candidate genes showed that more than 80% of the genes analyzed were differentially regulated in a similar fashion to that obtained by the microarray analysis (n = 22). The results are summarized in Table 2. We further examined the data by using the correlation analysis described in the sections above. The correlation between the ratio values obtained between the two experiments is 0.46. This value reflects the fact that, while the direction of expression change is conserved (i.e., sign agreement), the magnitude can differ. Among newly discovered temporally regulated genes that were confirmed by RT-PCR were three Xenopus ESTs, one down-regulated and two up-regulated at MBT (Table 2, clones S10-1-E7, S10-8-D4, S10-4-C1, respectively). Interestingly, 6 out of 22 genes involved in RNA metabolism are down-regulated, and a number of growth and transcription factors are up-regulated (Fig. 1) and confirmed by PCR (data not shown).

**Microarray Analysis of Spatially Restricted Embryonic Gene Expression**

One of the strengths of the Xenopus system is the ease of dissecting the large embryo. We isolated cells from the dorsal and ventral regions. On the dorsal side the organizer is a source of signals responsible for the induction and patterning of the dorsal tissues including the nervous system (Harland and Gerhart, 1997). This region plays a critical role in the formation of the early vertebrate embryo body plan and has been the focus of intense study since Spemann's famous transplantation experiments (Spemann and Mangold, 1924).

In this experiment, as well as the one described below, we used 15 μg of total RNA for each type of explants', instead of using poly(A)⁺, to assess whether there is a potential
benefit of using one versus the other. The use of total RNA will obviously facilitate the microarray approach in embryological studies. As expected and in contrast with the temporal experiment, significantly fewer differentially expressed clones were identified. A number of genes displayed at least two-fold differences in expression between the two samples. Less than 10% of the genes showed to be regulated by at least two-fold between the dorsal and ventral samples, among these, the dorsal marginal zone-specific gene, goosecoid (Cho et al., 1991) showed a correct segregation to the dorsal side.

The examination of the data derived from the array suggested that the use of poly(A)⁺ instead of total RNA improves the strength of the signal and the number of spatially regulated genes might be underestimated. Nevertheless, even using suboptimal conditions, the array led to the discovery of a number of spatially regulated genes. Again, the results of the array were confirmed independently by RT-PCR and in situ hybridization. RT-PCR analysis of two novel ventrally expressed genes is presented in Fig. 2.

In addition to investigating the changes in gene expression in dorsal and ventral tissue by RT-PCR, we also examined whether the clones identified using the array demonstrated differences in mRNA transcript distribution at gastrula stages using whole-mount in situ hybridization. The in situ hybridization patterns for three of these genes are shown in Fig. 3: EIFγ (Figs. 3A–3C), EST [da68c11.yl, 24C1-1 (Gawantka et al., 1998), Figs. 3D and 3E] and glycerol-3-phosphate dehydrogenase (GPD; Figs. 3F–3H). A section through a gastrula stage embryo analyzed for EIFγ mRNA transcripts (Fig. 3A) shows stronger expression in the dorsal lip than in the ventral aspect (Fig. 3i). We see less expression in the ventral tissues at or near the blastopore lip during gastrulation. The two other genes described here showed slightly less obvious differences in dorsal/ventral expression at these early stages (Figs. 3D and 3F, and in sections iii and iv). This is presumably due to the strong ectodermal expression and saturation of the color reaction. All three genes are expressed in the ectodermal layer and are excluded from the involuted cells that become mesodermal.

We further investigated the expression patterns of these three genes by looking at later developmental stages. By st18, EIFγ is restricted to the neural plate (Figs. 3B and 3C), consistent with being expressed in the dorsal lip. At stage 18, EST [da68c11.yl] is expressed mainly in the caudal half of the embryo (see Fig. 3E), in dorsal tissues (section in Fig. 3ii). There is also weak anterior expression, in what may possibly be migrating neural crest cells (see arrow). By tadpole stages, EST [da68c11.yl] expression is restricted to the tailbud (data not shown). GPD, which appeared to be expressed throughout the gastrula-staged embryo, had very specific expression patterns later in development, becoming restricted to the dorsal neural tube (Figs. 3G and 3H). Transverse sections through the posterior end of a st18 embryo shows that the transcripts are expressed in the dorsal part of the neural tube and are not expressed ventrally, consistent with this gene being expressed more strongly dorsally than ventrally during gastrulation (Fig. 3vi). A more anterior section shows very similar dorsal expression (Fig. 3v).

### Microarray Analysis of Activin-Treated Ectodermal Explants

The animal cap explants, derived from the top of the blastula embryo, also known as ectodermal explants, are far the most commonly used assay system in Xenopus experimental embryology. Ectodermal cells at that stage are pluripotent and can respond to a variety of treatments to generate a large repertoire of embryonic cell types. The molecular dissection of inductive events in all three germ layers has relied heavily on the use of this tissue. To explore the usefulness of the arrays, in animal cap assays, we examined the changes in gene expression when these cells were exposed to activin, a member of the TGFβ superfamily of growth factors. Activin has been shown to induce both mesoderm, with the ability to act as a morphogen (Green and Smith, 1990), and endoderm in these cells. Five micrograms of total RNA, derived from 25 animal caps either exposed to activin or buffer alone, were labeled as described and inverse dye hybridization was performed. We focused on genes that showed at least two-fold differences in expression. Three categories of activin-regulated genes were obtained by these studies. The first category are genes which were previously characterized and shown to be regulated by activins. Those include the signal transducer MAPKιnase2 (Hartley et al., 1994), the homeobox genes Xlim1 (Rebert and Dawid, 1997), and Xhox-lab (Sive and Cheng, 1991) for the up-regulated genes. The second category represents genes discovered by our array approach, which were previously described but not correlated with activin signaling. Among those are the up-regulated genes, the transcription factors, XFD-12 (Solter et al., 1999), and Hairy2a (Dawson et al., 1995), the secreted factor neurotrophin-4 (Hallbook et al., 1991) and a down-regulated gene, Osteoblast specific factor-1 (Tsujimura et al., 1995). The final category is made up of novel activin responsive genes which were not previously described in Xenopus, and therefore never correlated with activin signaling. Those are: XCG1-23 for up-regulated and an uncharacterized Xenopus EST, 1F1 for down-regulated genes. The genes belonging to the second and third category described above were independently confirmed by RT-PCR (Fig. 4).

The in situ expression pattern of XFD-12 was recently published (Fetka et al., 2000). We further characterized one of the genes down-regulated by activin by whole-mount in situ hybridization, clone 1F1 (Drosophila CG5142, C. elegans F54C1.5, see Fig. 5). At gastrula stages this gene is expressed throughout the embryo (data not shown). At st18, expression of 1F1 is located within the epidermis in ciliated cells (see Figs. 5A and 5B and high power picture of the ciliated cells, Fig. 5ii). This is in a similar pattern to that...
described by Deblandre and colleagues for Xdelta-1 (Deblandre et al., 1999), suggesting that expression may be mediated by the notch/delta signaling pathway. At this stage, F1 is not expressed in the dorsal neural tube (Fig. 5A). F1 is however expressed in the floor plate, but not the notochord (see Fig. 5i, arrow). Since floor plate specification has involved signaling through the sonic hedgehog (shh) pathway (Muller et al., 1999), it is tempting to speculate that this evolutionary conserved gene participates in that pathway.

**DISCUSSION**

To address the usefulness of microarray approaches in the context of early vertebrate embryonic development, we specifically addressed four questions. First, what is the limit of detection of differentially expressed genes in the embryo? Our observations are that, in terms of the limit of detection of low-abundance mRNA species, microarray analysis is not superior to RT-PCR methods. Currently, we do not know of any assay which is qualitatively and quantitatively more sensitive than RT-PCR. However, the large number of genes arrayed in a single experiment can compensate for the disadvantage in sensitivity. Furthermore, microarrays represent an unbiased method in the assessment of mRNA expression profiles, since no gene-specific primers are required, and therefore it is a more suitable approach than traditional PCR methods for gene discovery.

The second question is whether the quantification is comparable to other techniques such as RT-PCR. We find that, while the general direction of change of gene expression is in very good agreement between the microarray and the RT-PCR analysis, the quantitation, while close, is not always the same. As noted above, our results are consistent with previously published reports (Soukas et al., 2000; Schena et al., 1996a; Coller et al., 2000). There are a number of reasons why this quantitative difference might occur, including cross-hybridization between similar sequences on the microarray. Cross-hybridization may potentially increase or decrease the observed ratio and should be dependent on the stringency of hybridization and washing of the array. When a differential expression is observed in the microarray and no difference by PCR, a closely related homologue that is differentially regulated might be contributing to the microarray signal. While the potential to amplify a closely related homologue by PCR exists, in those few cases one might observe multiple bands when the products are separated by gel electrophoresis. In contrast, it has been observed that genes which are >85% conserved have the potential to cross-hybridize under the conditions used for microarray analysis (Heller et al., 1997). Thus, the lower correlation may arise due to the difference in specificity between microarrays and PCR. Increasing the stringency of hybridization and/or microarray washing is likely to decrease the amount of cross-hybridization, although it may also decrease the overall signal. These results have important implications for microarray-based analysis both for the experiments presented here but also for the published reports in which multiple experiments are compared by using clustering approaches.

The third question addresses how much biological material is required for these experiments to be useful? We have tried microarray experiments using as little as 5 μg of total RNA in settings where the inductive ability of activin was measured on a small set of animal cap explants, and this has been enough to allow the discovery of novel regulated genes. The total content of a Xenopus embryo and the end of blastula early gastrula is about 4–5 μg, which implies that, in principle, about 25 explants or a single embryo should provide enough RNA to generate a probe. Small numbers of embryos obtained by transgenic technologies or loss of function approaches in Xenopus are therefore amenable to this type of analysis. We believe, however, that this is the absolute minimum amount of unamplified biological material required to perform these assays, as the signal-to-noise ratio decreases, when limited amounts of total RNA are used. Overall, our experiments suggest that, at least in the context of embryological studies, the use of poly(A)+ RNA is preferable to total RNA.

Fourth, can new regulatory genes be identified using this type of approach? The answer is yes. By just using the prototype array, we have identified a large number of novel regulated genes. Among those are temporally regulated genes, including 20 cDNAs, which do not show significant homology to any sequence in public databases. In the temporal set, many of the up-regulated genes are involved in transcription and regulatory functions, while the down-

![FIG. 4. PCR analysis of genes identified by the array. Activin treated caps (+) and untreated caps (−) were amplified by PCR as described and quantitated by Molecular Dynamics Phosphor-Imager. Ratios are given for the RT-PCR as well as for the array analysis to compare the two.](image-url)
regulated genes included a greater number of metabolic genes. The prototype microarray also identified genes differentially expressed in the gastrula embryo independently confirmed by RT-PCR. Two of the genes had been previously identified but not shown to be differentially expressed and the other is a novel Xenopus EST sequence. Finally, we have combined array analysis with the Xenopus explant assay system to examine changes in gene expression in response to the TGFβ family member activin. We identified three new activin-regulated genes and further showed that a number of previously characterized genes are affected by activin activity.

In the studies presented here, we have not attempted to draw direct biological conclusions from the array data. Rather we observe that the arrays when used under a variety of conditions can select for regulated genes in the developmental assays. In many but not all cases, the differences can be confirmed independently. During vertebrate development, considerably different developmental outcomes can arise from highly related factors. With these observations in mind, we have tempered our interpretation of the array results. Despite the constraints on the specificity and the detection of low-abundance messages observed with microarrays, it remains a powerful tool for identifying new regulated transcripts.

Application of Microarray to Temporal Development Analysis

In the study presented here, we have demonstrated the feasibility of examining global changes in steady-state transcript levels. With these prototype arrays, we note that many but not all of the down-regulated genes are involved in metabolic processes, while many of the up-regulated genes are transcription factors, regulators, or involved in protein synthesis.

A strength of the microarray-based approach is the ability to identify both down-regulated and up-regulated genes in a single experiment. While the potential importance of down-regulated genes is appreciated, a large majority of published studies describe genes whose expression is activated. Temporally down-regulated genes might have key functions during early embryonic development. For example, down-regulated genes might play a role in embryonic competence. Competence is the ability of a tissue to respond to inductive signals, and this ability is in many cases transient and occurs in a window of time. While there is little known about the molecular basis of competence, down-regulated genes identified in a temporal assay are potential candidates for competence factors.

In Xenopus, zygotic transcription does not begin until mid-blastula stages when the embryo is made of about 4,000 cells (Kimelman et al., 1987). Thus, all embryonic events mediated before the first 13 cell divisions are under maternal control. By in vitro fertilization, it is possible to obtain large numbers of synchronized embryos, which divide every 20 min, after the first cell cycle. In yeast, the global analysis of transcription was possible due to the ability to synchronize cell division and was key to the successful analysis performed for the cell cycle and sporulation (DeRisi et al., 2000; Spellman et al., 1998a,b; Percival-Smith and Segal, 1984). The temporal experiments presented here demonstrate the feasibility of examining these early embryonic events, document changes in RNA levels, and cluster the information, with regard to each cell cycle.

Application of Microarray Approaches to Spatial Studies

We have dissected the developing embryo into dorsal and ventral regions to successfully identify differentially regulated genes. Among the genes are two factors which would generally not be expected to show such regulation, one is translational elongation factor, EIF4z (Keiper et al., 1999) and the other is glycerol-3-phosphate-dehydrogenase (GDPH; Adler and Klucznik, 1982). The case of the translation factor is of interest because these proteins have been suggested to be part of the ubiquitous translational machinery (Gingras et al., 1999) and yet show restricted spatial expression, suggesting the presence of regions within the embryo subject to specialized regulation at the translational level. Recently, two other EIF4 translation factors, elF4AIII (Weinstein et al., 1997) and elF4AII (Morgan and Sargent, 1997) were shown to be differentially expressed and sufficient to direct cell-fate decisions. It will be of interest whether elF4z is also endowed with the ability to specify cell fates. More generally, this embryological evidence suggests that we need to reconsider the regulatory role of the translation factors in vivo. The differential expression of GDPH was surprising, as this gene has been shown to be involved in the lipid biosynthesis pathway, and therefore to be expressed ubiquitously. Its localized expression may either reflect general differences in metabolic activity during development or indicate a link to regionalized embryonic signaling pathways.

Microsurgical approaches can be applied at increasing levels of refinement. We focused on the dorsal–ventral axis and did not examine the differences in the animal–vegetal axis along which the three primary germ layers lie. Secondary experiments will refine this approach to examine the separate dorsal ectodermal, mesodermal, and endodermal derivatives. Further refinements will allow the examination of spatial expression patterns in the neural plate, mesoderm and other tissues. The large size of Xenopus embryos allows the collection of sufficient material to do these studies without amplification of the probe. Using improved amplification approaches, examination of individual differentiated cell types within differentiated tissues is possible when combined with cell sorting and micro laser dissection.
important to remember that this is a fast moving technol-
regulated genes under a variety of conditions. It is also
Based on the success of the prototype arrays, these
unique maternal clones to include on the next iteration of
Xenopus gastrula-stage genes. In addition, we have obtained 2,600
Future Prospects
explants is particularly well suited to
Application of Microarrays to Inductive Studies in
Intact or Manipulated Explants
For each of the approaches described above, Xenopus offers a number of unique advantages for the application of
tools. The application of array analysis in tissue explants is particularly well suited to Xenopus. In this study, we have identified new activin-regulated genes, including the first vertebrate member of an evolutionarily
conserved family of proteins. In addition, the microarray has allowed the discovery of a linkage between a number of previously described genes to the activin pathway. A similar
analysis can be done for any number of growth factors and common genes identified by various clustering
approaches. The inclusion of a consistent control, the un-
treated animal cap, allows experiments performed in separ-
ate labs examining diverse molecules to be directly and
easily compared, building toward a global understanding of embryonic signaling strategies.

Future Prospects
We have now prepared Xenopus arrays containing 5,000
gastrula-stage genes. In addition, we have obtained 2,600
unique maternal clones to include on the next iteration of
Xenopus array (Washington University, Xenopus EST project). Based on the success of the prototype arrays, these
larger scale arrays should allow the rapid identification of
regulated genes under a variety of conditions. It is also
important to remember that this is a fast moving technol-
ogy. As the robotic qualities of the array printers, scanners, and bioinformatic analysis improves, the power and resolution of these genome-wide approaches will increase. The establishment of the Xenopus EST projects, already on the
way (www.nih.gov/science/models/xenopus/index.html), also pave the way to a uniset Xenopus microarray within the
realm of proximal reality.

ACKNOWLEDGMENTS
We thank Ignacio Muñoz-Sanjuan for critical reading of this
manuscript and Tetyana Obukhanych for the activin array hybrid-
ization. This work was supported by the following grants: PHS
HD32105, PHS EY 12370, The John Merck Fund, The Steinbach
Fund to AHB. E.B. is supported by The Rockefeller University's
Women in Science Award and by an EMBO fellowship.

REFERENCES
dehydrogenase in developing chick retina and brain. J. Neuro-
chem. 38, 909–915.
Altmann, C. R., Chow, R. L., Lang, R. A., and Hemmati-Brivanlou,
185, 119–123.
Brown, P. O., and Botstein, D. (1999). Exploring the new world of
Cho, K. W., Blumberg, B., Steinebeisser, H., and De Robertis,
E. M. (1991). Molecular nature of Spemann's organizer: The role of
Coller, H. A., Grandori, C., Tamayo, P., Colbert, T., Lander, E. S.,
with oligonucleotide microarrays reveals that MYC regulates
genes involved in growth, cell cycle, signaling, and adhesion.
(1995). Specificity for the hairy/enhancer of split basic helix-
loop–helix (bHLH) proteins maps outside the bHLH domain and
suggests two separable modes of transcriptional repression. Mol.
Cell Biol. 15, 6923–6931.
DeRisi, J., Penland, L., Brown, P. O., Bittner, M. L., Meltzer, P. S.,
cDNA microarray to analyse gene expression patterns in human
DeRisi, J., van den, H. B., Marc, P., Balzi, E., Brown, P., Jacq, C., and
Goffeau, A. (2000). Genome microarray analysis of transcrip-
tional activation in multidrug resistance yeast mutants. FEBS
Lett. 470, 156–160.
Fetka, I., Doederlein, G., and Bouwmeester, T. (2000). Neuroecto-
dermal specification and regionalization of the spemann organ-
Gaasterland, T., Sczyrba, A., Thomas, E., Aytekin-Kurban, G.,
tion of the 2.9-Mb Drosophila melanogaster Adh region. Genome
Res. 10, 502–510.

Copyright © 2001 by Academic Press. All rights of reproduction in any form reserved.

Sive, H. L., and Cheng, P. F. (1991). Retinoic acid perturbs the expression of Xholab genes and alters mesodermal determina-
tion in Xenopus laevis. Genes Dev. 5, 1321–1332.


