The ANISEED database: Digital representation, formalization, and elucidation of a chordate developmental program

Olivier Tassy,1,13,14 Delphine Dauga,1,13 Fabrice Daian,1,13 Daniel Sobral,1,13,15 François Robin,1,16 Pierre Khoueiry,1,17 David Salgado,1,18 Vanessa Fox,1 Danièle Caillol,1 Renaud Schiappa,1,19 Baptiste Laporte,1 Anne Rios,1,18 Guillaume Luxardi,1 Takehiro Kusakabe,2 Jean-Stéphane Joly,3 Sébastien Darras,1 Lionel Christiaen,4,20 Magali Contensin,1,5 Hélène Auger,3 Clément Lamy,1 Clare Hudson,6 Ute Rothbächer,1 Michael J. Gilchrist,7,21 Kazuhiro W. Makabe,8 Kohji Hotta,9 Shigeki Fujiwara,10 Nori Satoh,11,22 Yutaka Satou,11,12 and Patrick Lemaire1,23

1–12 [A complete list of author affiliations appears at the end of the paper before the Acknowledgments section.]

Developmental biology aims to understand how the dynamics of embryonic shapes and organ functions are encoded in linear DNA molecules. Thanks to recent progress in genomics and imaging technologies, systemic approaches are now used in parallel with small-scale studies to establish links between genomic information and phenotypes, often described at the subcellular level. Current model organism databases, however, do not integrate heterogeneous data sets at different scales into a global view of the developmental program. Here, we present a novel, generic digital system, NISEED, and its implementation, ANISEED, to ascidians, which are invertebrate chordates suitable for developmental systems biology approaches. ANISEED hosts an unprecedented combination of anatomical and molecular data on ascidian development. This includes the first detailed anatomical ontologies for these embryos, and quantitative geometrical descriptions of developing cells obtained from reconstructed three-dimensional (3D) embryos up to the gastrula stages. Fully annotated gene model sets are linked to 30,000 high-resolution spatial gene expression patterns in wild-type and experimentally manipulated conditions and to 528 experimentally validated cis-regulatory regions imported from specialized databases or extracted from 160 literature articles. This highly structured data set can be explored via a Developmental Browser, a Genome Browser, and a 3D Virtual Embryo module. We show how integration of heterogeneous data in ANISEED can provide a system-level understanding of the developmental program through the automatic inference of gene regulatory interactions, the identification of inducing signals, and the discovery and explanation of novel asymmetric divisions.

With the exponential growth of biological data, databases have taken center stage in the biological sciences. The first biological databases were “monotype,” publishing a single type of information (e.g., PubMed for literature, GenBank for sequences). More recently, systems started hosting heterogeneous data sets related to the developmental biology of most major model organisms, e.g., FlyBase (The FlyBase Consortium 1994), WormBase (Harris et al. 2004), Zfin (Sprague et al. 2001), MGD (Blake et al. 1997), Dytchbase (Chisholm et al. 2006), and TAIR (Huala et al. 2001). These databases integrate the anatomy of the organism and its evolution at different scales into a global—and computable—view of the embryonic developmental program. Transparent cross-querying of heterogeneous types of data is virtually nonexistent. Integration of the results of “gene-centric” experiments into higher order gene regulatory networks (GRNs) would constitute a first step toward a global representation of the regulatory code underlying...
metaozan development (Davidson 2009). The representation of embryonic anatomy is also unsatisfactory in current systems. In particular, they do not integrate the relative positions and shape of anatomical structures, which can now be efficiently imaged (Tassy et al. 2006; Keller et al. 2008). This restricts the extent to which morphogenesis and cell communication can be digitally represented.

The embryos of land nematodes and ascidians develop with a very small number of cells and according to an invariant lineage (Conklin 1905; Sulston et al. 1983; Nishida 1987), which makes them particularly suitable to the development of integrated digital systems. We focused on ascidians. These marine invertebrate chordates are closely related to vertebrates (Delsuc et al. 2006). They are particularly suited to the study of GRNs (Imai et al. 2006, 2009; Kubo et al. 2010) and their function in the control of cellular behavior (Christiaen et al. 2008). Ascidian early embryology has been extensively studied in its cellular and molecular details (for review, see Lemaire 2009), and the functions of several hundred genes have been analyzed using loss- or gain-of-function approaches (e.g., Yamada et al. 2003; Wada et al. 2008). Central to the project described here, embryonic expression patterns by in situ hybridization are available for several thousand genes (Satou et al. 2002), including most transcription factor genes and signaling ligands (Imai et al. 2004; Miwata et al. 2006). Several specialized databases cover specific aspects of ascidian biology: genomes and coding genes (Ensembl, Birney et al. 2006; Hubbard et al. 2007; JGI, Dehal et al. 2002), embryonic anatomy (FABA, Hotta et al. 2007), gene expression (GHOST, Satou et al. 2002; MAGEST, Kawashima et al. 2002), or cis-regulatory sequences (DBTGR, Sierro et al. 2006).

The system described here is a first attempt at combining and extending available information into a generalist model organism database, with capabilities that extend the classical "gene-centric" approach used by most model organism databases. As a proof of principle of the power of this system, we identified novel asymmetric cell divisions during early embryogenesis, identified the correct neural inducer among many possible candidates, and automatically reconstructed GRNs to a significant extent.

Results

General organization of the ANISEED system

Figure 1 presents the general organization of the ANISEED system, the ascidian implementation of a generic system called NISEED (Network for in situ Expression and Embryological Data). The content of the ANISEED database can be explored via the Developmental Browser (http://aniseed-ibdm.univ-mrs.fr), which uses hub pages to organize key information about embryo anatomy, developmental stage. As such description was lacking for ascidians, we created an Open Biomedical Ontologies (OBO)-compliant ascidian anatomical ontology for each stage of the Ciona intestinalis developmental stage (Hotta et al. 2007). Development was split into three periods, for which we used different ontology logics: cleavage stages, gastrula to larval stages, and post-metamorphosis stages (Supplemental Fig. S3A). The resulting 25 ontologies for Ciona intestinalis describe 2240 anatomical territories. In addition to Ciona, four other solitary ascidian species are used in the community: Ciona savignyi, Phallusia mammillata, Halocynthia roretzi, and Boltenia villosa. The strong developmental conservation between ascidian species was used to create parallel tentative embryonic and adult anatomical ontologies for these species. Because of the phylogenetic proximity of C. savignyi and P. mammillata to C. intestinalis, their ontologies were inferred to be identical to those of Ciona intestinalis. The ontologies of the more distantly related H. roretzi and B. villosa slightly deviate from those of Ciona. Supplemental Figures S3B,C and S4 describe the deviations that were incorporated as a result of new findings (division patterns of A7.6 and B7.6; fates of B7.5) and interspecies variability (b8.19) using the dedicated ontology editing tools of the NISEED-manager.

Integration of ontologies at successive stages was achieved via lineage links between mother and daughter territories. Around 2000 such lineage links were established manually for each supported species, via a dedicated interface of the ANISEED-manager. Besides classical studies, this work integrated three recent studies of the Ciona intestinalis tail epidermis lineage (Pasini et al. 2006) and posterior neural plate lineages (Nicol and Meinertzhagen 1988; Cole and Meinertzhagen 2004). Lineage relationships were used to associate one or several larval fates with each embryonic territory.

Finally, the ANISEED description of ascidian anatomy integrates the three-dimensional (3D) topology of the cells and tissues of the embryo, computed from reconstructed 3D embryo models (Supplemental Fig. S4A; Tassy et al. 2006). This information is currently available in Ciona for all cleavage stages, up to the early gastrula stage. Phallusia mammillata stages between the 64-cell and the early gastrula stages are covered. For each territory, the system indicates the physical distances separating cells or
structures, a measure of the surface of contacts between adjacent structures, as well as a quantitative description of the territory’s geometry (volume, external surface, sphericity, squareness, convexity, elongation, flatness, and entropy) (Fig. 2A). The search for cells most likely to communicate and the study of the evolution of cell contacts is facilitated by cell neighbor graphs available for all pregastrula stages in *Ciona intestinalis* (Fig. 2B).

In summary, for each stage, an anatomical structure is represented in ANISEED by six sets of parameters: its position in the hierarchical anatomical ontology defined at this stage (“is part of”), its lineage (“is progeny of”—a structure from the previous stage dictionary), its fate at larval stages, its contacts to neighboring structures, and its shape and size. Each of these parameters can be individually explored using dedicated interfaces organized in the “Explore anatomy” section of the Developmental Browser (see Supplemental Fig. S2).

**Figure 2.** Representation of embryonic anatomy. (A) Screenshot of an anatomical territory card representing the lineage, fate, position/contacts with neighbors, and geometry of the a6.5 cell at the early 32-cell stage. Note the tabs at the top of the screen capture that lead to the whole anatomical ontology for the stage of interest (“Anatomical Ontology”), to the precursors and progeny of the territory of interest (“Lineage”), to expression profiles restricted to this territory (“Molecular Markers”), and to the regulatory interactions that take place in the lineage leading to this territory (“Regulatory Network”). (B) Neighborhood graph showing which cells of the late 32-cell embryo contact each other. The nodes of the graph represent individual cells, the edges represent contacts. The thickness of the edge reflects the area of contact between adjacent cells. The indicated values are in square microns. Supplemental Figures S3 and S4 give more details about the description of ascidian anatomy in ANISEED.

**Representation of genomic features**

ANISEED defines a set of 20,631 *Ciona intestinalis* gene models obtained by clustering transcripts models predicted from JGI (Dehal et al. 2002), Ensembl (Hubbard et al. 2007), and the Kyoto genome consortium (Satou et al. 2008). These gene models are functionally annotated by running a dedicated automatic annotation pipeline, based on protein domain detection and evolutionary inference of function and biological name (Supplemental Fig. S5; Supplemental Methods). Gene models are linked to all available public ESTs and cDNAs and to experimentally validated cis-regulatory regions. All sequence features can be visualized in their genomic context in the ANISEED Genome Browser, which also displays the conservation profiles between the two sequenced *Ciona* genomes (*Ciona intestinalis* and *Ciona savignyi*) and predictions of local nucleosome occupancy (Khoueiry et al. 2010).

A sophisticated representation of the structure and activity of cis-regulatory elements and of their upstream regulators was designed, as these elements play a major role in the control of the developmental program and are central to the reconstruction of GRNs. To faithfully represent the cis-regulatory logic, all experimentally tested regions at one locus are organized hierarchically (Fig. 3). The type of activity of each region in in vivo reporter assays is described using a controlled vocabulary. Eight classes of activity were thus defined, five of which match Sequence Ontology terms (Supplemental Fig. S6; see Methods; Eilbeck et al. 2005). When experimentally described, individual functional transcription factor binding sites are shown and linked to the corresponding experimentally verified trans-acting factors (Fig. 3). The precision and traceability of this representation is an advance over model organism databases such as WormBase, MGI, FlyBase, or Zfin, or dedicated cis-regulatory element databases such as DBTGR (Siervo et al. 2006), ORegAnno (Montgomery et al. 2006), RedFly (Gallo et al. 2005; Halfon et al. 2008), or the VISTA enhancer browser (Visel et al. 2007).

ANISEED currently describes 528 published or unpublished regulatory regions controlling the transcription of 158 genes. For 85 regions, functionally tested binding sites for specific transcription factors are indicated. These numbers are comparable to those of the Redfly project, which annotates the regulatory regions of the much more studied *Drosophila*.

**Description of expression data in wild-type and experimentally manipulated contexts**

Spatio-temporal gene expression is described in ANISEED with EST counts, in situ hybridizations, protein immunolocalization,
and cis-regulatory element activity. ESTs counts from 28 sequenced non-normalized cDNA libraries provide a low-resolution temporal expression information for 80% of the genes (Satou et al. 2003). A digital differential display (DDD) tool (Supplemental Fig. S2) uses this information to find genes differentially represented between conditions with its matching wild-type control in the same experiment (Fig. 4). This curation decision is unique to ANISEED: Although Zfin describes some in situ expression profiles of target genes in response to experimental perturbations, this information is not associated with the matching wild-type control.

Impact, issues, and enhancement of manual curation

Published small-scale studies are the primary source for expression profiles in deregulated contexts and their associated controls. One-hundred-sixty manually curated articles are currently publicly available in ANISEED, a data set that covers most of the molecular literature on Ciona intestinalis. These manually extracted data represent 18.5% of all ANISEED spatial expression profiles, illustrating that the manual compilation of small- and mid-scale studies can reach a scale similar to large-scale screens.

The capture of published information was streamlined by the creation of the “Article Card” concept. Each Article Card summarizes the changes in the expression patterns of 18 genes were analyzed. In contrast to ANISEED, advanced model organism databases such as FlyBase, WormBase, Zfin, or MGI describe the morphological or anatomical phenotype resulting from experimentally altered gene function, but provide little information about the transcriptional consequences of these mutations.

As developmental genes often have dynamic expression patterns (Sobral et al. 2009), wild-type and experimentally modified expression patterns should be compared within a given experiment. We thus associated each expression pattern in perturbed conditions with its matching wild-type control in the same experiment (Fig. 4). This curation decision is unique to ANISEED: Although Zfin describes some in situ expression profiles of target genes in response to experimental perturbations, this information is not associated with the matching wild-type control.

Collectively, these experiments describe the changes in the expression patterns of 232 target genes (total of 1152 patterns). Embryological perturbations, such as cell ablations or explants, are defined by the removed anatomy parts and the developmental stage of the perturbation. Eighty-seven such ablation experiments are reported, in which the expression patterns of 18 genes were analyzed. In contrast to ANISEED, advanced model organism databases such as FlyBase, WormBase, Zfin, or MGI describe the morphological or anatomical phenotype resulting from experimentally altered gene function, but provide little information about the transcriptional consequences of these mutations.

Figure 3. Representation of cis-regulatory information. Screenshot of the regulatory region card for the early minimal neural enhancer of Ciona intestinalis Otx. The precise pattern of activity of this region is accessed by clicking on the “view in situ data” link located in the “Constructs made to test this region” section. Supplemental Figure S6 presents the classification system for cis-regulatory regions used in ANISEED.
Unequal cell cleavages, leading to daughter cells of different sizes, are frequently associated with cell fate decisions (Sardet et al. 2007). We previously integrated lineage information and cell volumes to identify unequal cleavages up to the 44-cell stage in *Ciona* (Tassy et al. 2006). This approach was extended in *Ciona* embryos in which endoderm invagination had been prevented by treatment with the Rho-kinase inhibitor Y-27632 (Sherrard et al. 2010). While the orientation of the division of mother cells was unchanged, the volumes of sister cells were now equal. Thus, endoderm invagination may influence the asymmetry, but not the orientation, of animal cell cleavages (Fig. 6E).

**Integrating functional gene annotation, expression data, and anatomy: In search of the ascidian neural inducer**

In ascidians, the anterior neural tissue is induced in a6.5 blastomeres at the 32-cell stage by a signal originating from the A-line vegetal cells (for reviews, see Meinertzhagen et al. 2004; Lemaire 2009). To identify the best candidate inducer among the 54 secreted ligands from major signaling pathways expressed around this stage, we made use of the sequential query mode of ANISEED, in which the results of an initial query are used as the search space for the next one (Supplemental Fig. S8). Assuming that translation, folding, and transport of the inducer takes ~20 min and that activation of a target gene may take an additional 15 min, we looked for genes coding for secreted proteins that were expressed at the 16-cell stage in the vegetal neighbors of the progenitors of the induced cells. We formulated two sets of parallel queries: one molecular and the other anatomical. The set of molecular queries identified genes coding for secreted proteins (Gene Ontology query) and expressed during the cleavage stages (Digital Differential Display query). The set of anatomical queries identified the ancestor of the a6.5 cell pair, and then the vegetal neighbors of these animal cells. Integration of these two sets through the “Query integration by ID” web interface (Supplemental Fig. S2) yielded two genes, FGF9/16/20 and Lefty. The former encodes the demonstrated neural inducer (Bertrand et al. 2003). Lefty, an antagonist of nodal signaling...
Meno et al. (1996), does not appear to play a major role during early Ciona embryogenesis (Imai et al. 2006).

Integration of expression patterns in wild-type and deregulated conditions: Automatic inference of gene regulatory interactions

We finally sought to use the ANISEED expression data set to automatically extract transcriptional regulatory interactions underlying ascidian development. The analysis of Ci-Otx regulation (Fig. 7A) exemplifies the simple inference logic we used (described in Supplemental Methods). At the 44-cell stage, Ci-Otx is expressed in the a6.5, b6.5, and B6.4 cell pairs. When the function of the maternal ETS1/2 mRNAs is blocked by Morpholino (MO) injection, expression of Ci-Otx becomes restricted to B6.4. Comparing the two sets of expressing territories, we infer that by the 44-cell stage ETS1/2 positively regulates Otx in a6.5 and b6.5. Integration with cis-regulatory information reveals that this regulatory event is direct: The early Otx neural enhancer includes two ETS1/2 binding sites that are required for its activity (Bertrand et al. 2003).

By applying these rules to the whole ANISEED data sets, we obtained 498 distinct gene regulatory interactions, supported by loss-of-function assays involving 183 genes at different developmental stages (Supplemental Fig. S9). This network includes and significantly extends the networks from Imai et al. (2006, 2009) (195 interactions for 81 genes) and describes 20 direct interactions confirmed by cis-regulatory region mutational analysis.

Two classes of web interfaces were developed to navigate this information. For each anatomical territory, the “Anatomical Gene Network card” indicates the known regulatory interactions that have taken place in this territory by the stage studied (Fig. 7B). For each individual gene, the “Upstream regulators” page gives access to known upstream regulators in each territory. Conversely, the “Downstream targets” page lists targets of the gene of interest by territories and stages (Fig. 7C). In all cases, the interactions are linked to their supporting experimental evidence (Fig. 7D).

Discussion

ANISEED is the first integrated system for ascidians and is widely used in the growing community of ascidian laboratories. Over the past 2 yr the system received 51,000 unique visitors, mainly from France, Japan, the USA, and Italy—all countries with strong ascidian communities. In addition, the generic nature of the system and its design principles may hold lessons for the future of model organism databases.

The tight collaboration between experimentalists, biocurators, and computer scientists was a central aspect of the ANISEED project. Experimental biologists used their knowledge of ascidian development and experimental approaches to define the general scope of the project, and the most important types of data for the ascidian community. This led to a strong focus on transcriptional regulation, gene regulatory networks, and their integration into a precise three-dimensional representation of embryo development. A second major role of experimentalists was to define minimal information standards for each type of experiment represented, thereby achieving a well-defined and consistent set of data representations. Biocurators then defined the ontologies and controlled vocabularies necessary to formalize the representation of experiments, thus transforming the “biological interpretability” of the data into their “computability.” This close fit between experimental design, raw experimental data, and their
Although the stereotyped development of ascidians, based on fixed cell lineages, is an obvious advantage for the approach presented here, the system can easily be adapted to integrate the geometry of embryos that develop in a less reproducible fashion. This adaptation could involve a change of scale of analysis, measuring the generally reproducible geometry and topological arrangement of fields of cells or organ rudiments. This tissue-level representation could be complemented at the mesoscopic scale by statistical measures of individual cell geometries and arrangements within each field or rudiment (e.g., Blanchard et al. 2009; Butler et al. 2009). Implementation of parallel NISEED systems will greatly facilitate the evolutionary comparison of developmental program within a taxon or phylum (e.g., Sobral et al. 2009).

The types of data represented in NISEED currently limit its predictive power. It will be simple to represent mutant lines, thus opening the system to model systems with a strong history of genetics, including fly, Caenorhabditis elegans, and mouse. Genome-wide chromatin assays and RNA-seq data are rapidly accumulating in many model organisms. Their careful integration into NISEED will help reconstruct regulatory networks, provided the convergent support given by distinct data sets (e.g., chromatin immunoprecipitation analyses, cis-regulatory information, expression patterns) are adequately ranked and attributed specific confidence values. This will require an evolution of the database schema. The GMOD (Generic Model Organism Database) consortium recently proposed a flexible modular database schema, CHADO (Mungall and Emmert 2007), relying, like NISEED, on the widespread use of ontologies and controlled vocabularies. This system currently offers a sophisticated representation of sequence features, but limited representations of embryo anatomy and gene expression patterns. Its extension with NISEED modules describing the three-dimensional anatomy of the embryo and the transcriptional program of each territory will help to further reconstruct the developmental GRNs acting in each territory by integrating the results of short-read-based expression and chromatin assays.

Methods

Hardware and software

The system runs on all platforms supporting PostgreSQL, JAVA, and JAVA 3D, including Windows, UNIX, Linux, and Mac OSX 10.5. Source code is available as Supplemental source code. Web pages are correctly displayed on Browsers supporting HTML level 4.0 or higher and JavaScript, including Firefox and Internet Explorer. Some display issues were reported with current versions of Chrome or Safari.

System administration, curation, and data download

Two sets of management tools are proposed. The NISEED-manager allows de novo creation of NISEED databases, the management of users, the editing of ontologies, the functional annotation of genes, and the import of large-scale data from flat files. It also centralizes scripts to update and upload data. Through the NISEED-curator, data from small-scale experiments can be inserted and manually curated. A precise description of the curation strategy can be found in the Supplemental Methods.

Three types of data can be obtained from the download section: genomic, anatomical, and expression data. Genomic data include gene and transcript models, functional gene annotations, and cis-regulatory regions. Anatomical data include anatomical ontologies in OBO format (Smith et al. 2007), the collection of
available reconstructed embryos, each with its biometry data and the 3DVE installation package. Expression data include all in situ hybridization data in a MISFISHIE-compliant XML format (Deutsch et al. 2008), including images, and modified to associate expression patterns in experimentally manipulated and control situations (example available at http://aniseed-bdcm.univ-mrs.fr/exchange_format.php). The NISEED-manager also allows to import Misfishie-Compliant XML files describing large-scale in situ hybridizations data sets.

3D embryo reconstruction and treatments

Reconstructed two-cell to 44-cell *Ciona* embryos were obtained from Tassy et al. (2006). Reconstructed *Boltenia villosa* and some *Phallusia* embryos were obtained from (Sherrard et al. 2010). Additional *Ciona* and *Phallusia* embryos were obtained by in vitro fertilization as in Robin et al. (2010). Rho-kinase inhibition was achieved by treating embryos from the 64-cell stage with 100 μM Y-27632, a specific pharmacological inhibitor of this kinase. Live or fixed embryos of *Phallusia mammillata* and *Ciona intestinalis* between the 64-cell and early gastrula stages were reconstructed using Amira and the files processed for import in 3DVE as indicated in Robin et al. (2010).

Definition, expression, and annotation of ANISEED gene models and cis-regulatory regions

ANISEED supports four previously generated independent *Ciona intestinalis* transcript model sets (JGIv1.0, Ensembl v2.0, KyotoGrail2005, and KH), which were found to be of highest quality and complemented each other. Short single exon models (<300 bp) that were supported neither by EST information nor by blast hits to other species, and presumably representing erroneous ab initio predictions, were not considered. Transcript models sharing at least one full exon were grouped into 20,915 ANISEED-v3.0 gene models by applying a published clustering algorithm (Gilchrist et al. 2004). *Halocynthia roretzi* transcript models were generated assembling 60,000 expressed sequences (ESTs, cDNAs) into contigs (Huang and Madan 1999; Quackenbush et al. 2000; Gilchrist et al. 2004). This procedure generated 14,970 tentative consensus sequences (TCs) that were used as transcript models. The NISEED automatic annotation pipeline integrates Interproscan (Quevillon et al. 2005), InParanoid v2.0 (Remm et al. 2001), and BLASTP. The pipeline is detailed in the Supplemental Methods section and illustrated in Supplemental Figure S5.

Figure 7. Automatic inference of transcriptional regulatory interactions. (A) Example of the inference logic that led to the establishment of a regulatory interaction between ETS1/2 and Otx in a6.5 and b6.5 neural precursors at the late 32-cell stage. Use of an antisense Morpholino oligonucleotide is interpreted as a loss-of-function experiment. Territories in which Otx expression is lost are thus territories where ETS1/2 regulates Otx expression. (B) Screenshot of the "Anatomy Regulation" card for the B5.1 precursor at the 16-cell stage, showing the regulatory interactions that take place in the lineage leading to this blastomere. (C) Screenshot of the "Downstream Target" card of beta-catenin, showing transcriptional targets of this gene at the 16-cell and late 32-cell stages. Note that the anatomical territories in which the regulatory interactions are confirmed as indicated. (D) Screenshot of the page presenting the evidence that beta-catenin upregulates Ci-FoxD in the A-line at the 16-cell stage. Supplemental Figure S9 presents the number of inferred regulatory links at each developmental stage.
polymorphism level in ascidians, the tested sequences often depart from the published genome sequence, so that each regulatory region is associated with the specific constructs that were used to test it, with their sequence. A description of the strategy used to name, represent, and classify cis-regulatory regions and their activity can be found in the Supplemental Methods and Supplemental Figure S6.

Spatial expression data (ISH, immunohistochemistry, and reporter assays) were represented by associating the in situ probe clone (ISH) or transcript model (Immunohistochemistry) or reporter construct (cis-reg. regions) to ontology terms using a curation strategy detailed in the Supplemental Methods section.

List of Affiliations

1Institut de Biologie du Développement de Marseille-Luminy, UMR6216 CNRS, F-13288 Marseille, Cedex 9, France; 2Department of Biology, Faculty of Science and Engineering, Konan University, Higashinada, Kobe 658-8501, Japan; 3Equipe INRA U1126 “Mor-
The ANISEED database: Digital representation, formalization, and elucidation of a chordate developmental program

Olivier Tassy, Delphine Dauga, Fabrice Daian, et al.

*Genome Res.* 2010 20: 1459-1468 originally published online July 20, 2010
Access the most recent version at doi:10.1101/gr.108175.110

**Supplemental Material**
http://genome.cshlp.org/content/suppl/2010/07/20/gr.108175.110.DC1

**References**
This article cites 66 articles, 14 of which can be accessed free at:
http://genome.cshlp.org/content/20/10/1459.full.html#ref-list-1

**Open Access**
Freely available online through the *Genome Research* Open Access option.

**License**
Freely available online through the Genome Research Open Access option.

**Email Alerting Service**
Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or click here.

To subscribe to *Genome Research* go to:
http://genome.cshlp.org/subscriptions

Copyright © 2010 by Cold Spring Harbor Laboratory Press