

CURATION OF NISEED, AN INTEGRATIVE FRAMEW ORK FOR THE DIGITAL REPRESENTATION OF EMBRYONIC DEVELOPMENT

http://aniseed.jbdm univ-mrs.fr



Delphine Dauga, Olivier Tassy, Fabrice Daian, Daniel Schral and Patrick Lemaire

IBDM L CNRS, UM R6216 CNRS, Campusde Luminy-Case 907, Marseille Cedex 09, France

ABSTRACT

NISEED (Network for In situ Expression and Embryological Data) is a generic infrastructure

for the creation, maintenance and integration of molecular and anatomical information on

model organisms. We applied it to ascidians which are marine invertebrate chordates. These

animals constitute model organisms of choice for developmental biology because their

embryos develop with a small number of cells and an invariant lineage, allowing their study

with a cellular level of resolution. In ANISEED (Ascidian NISEED), embryogenesis of

ascidian is represented at the level of the genome via functional gene annotations, cis-

regulatory elements or gene expression data, at the level of the cell by representing its

morphology, fates, lineage, and relations with its neighbors, or at the level of the whole

embryo by representing its anatomy and morphogenesis at successive developmental stages.

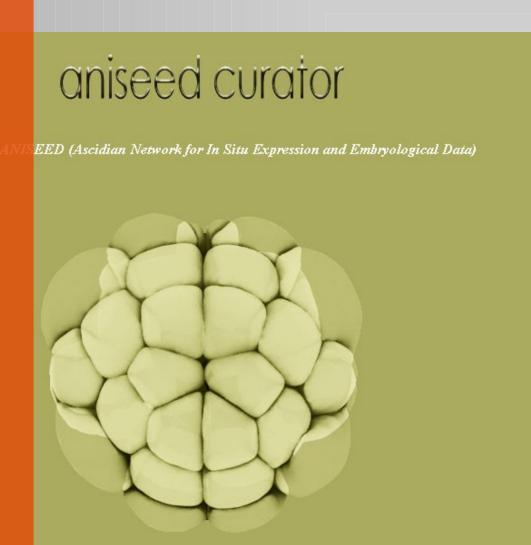
The system provides also tool and standard to enter, annotate, curate and manage data. All

results can be accessed through the ANISEED website at http://aniseed-ibdm.univ-mrs.fr

Expression curation pipeline

The NISEED-Curator website is used to enter, edit and validate molecular and expression data entered from the literature or contributed by other labs.

The "curation pipeline" allows papers of interest to be added to the system: after agreements of the authors, a team of annotators enter papers from Pubmed for detailed manual curation, using a series of annotation pages.



Genomic annotation pipeline

The gene functional annotation pipeline is handled by the NISEED-Manager website. This pipeline which integrates Interproscan, In paranoid and Psiblast allocates Gene Ontology terms to the gene models. In addition to that, this tool facilitates the creation of parallel databases for additional model organisms, manages users, and centralizes updating procedures.

aniseed

dian Network for In Situ Expression and Embryological Data)



NISEED

DATABASE

DATA STORAGE

Developmental

browser

NISEED database is

interfaced to a

developmental browser

(http://aniseed-ibdm.univ-

mrs.fr.), which textually

describes the anatomy and

functional genomics data

linked to the development of

a model organism

Rules

- * Users are requested to get an individual password before entering data
- * All new data are " *Private*" and " *Not* curated" by default

Import of external dataset

Format supported:
OBO
MISFISHIE
Fasta

Ontologies

We defined and used controlled vocabulary to describe features as accurately as possible. We based our definition according to:

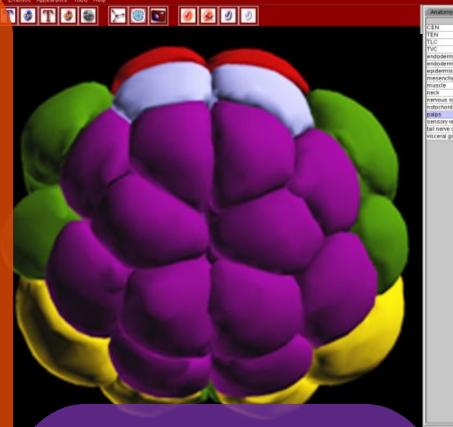
Gene ontology
Sequence ontology
Evidence ontology
Anatomical ontology
Protein ontology

data The 3D Virtual Embryo module uses interactive

module uses interactive
three dimensional digital
embryo reconstructions
to generate a quantitative
description of the
geometry of individual
embryonic
territories/cells and their
topological arrangement

Import

anatomical



DATA ENTRY AND

MANAGEMENT

Visualization of 3D
Embryo

The 3D Virtual Embryo module uses interactive three dimensional digital embryo reconstructions to display information from the database

<u>OniSearc</u>

ANISEED version 3:

mammillata

A model organism database for Ciona intestinalis, Ciona savignyi,

Halocynthia roretzi and Phallusia

Explore Anatomy Expression data Explore Literature Download

Aniseed is a database designed to offer a representation of ascidiar

embryonic development at the level of the genome (cis-regulator

sequences, gene expression, protein annotation), of the ce (morphology, fate, induction, lineage) or of the whole embryo (anatomy

morphogenesis). Currently, four organism models are described i Aniseed: Ciona intestinalis, Ciona savignyi, Halocynthia roretzi an

This version supports three sets of Ciona intestinalis transcript models

JGI v1.0, KyotoGrail 2005 and ENSEMBL, all functionally annotated, and

grouped into Aniseedv3.0 gene models. You can explore their expressio

profiles during normal or manipulated development, access the

regulatory regions, or get the molecular tools used to assay the

ANISEED Ascidian Network for InSitu Expression and Embryological Data

instructed early

DATA MINING

GBROWSE

DATABASE

Genome browser

JGI

cio100133709

Ci-otx

Ensembl JGI2

ENSCINT00000018028

0tx

kyoto Grail

KY0T0GRAIL2005.98.32.1

0TX2 / 0TX1

KH gene models

KH1.C4.84.v1.A.nonSL2-1

KH1.C4.84.v1.A.nonSL3-1

KH1.C4.84.v1.A.SL1-1

KH1.C4.84.v1.A.SL5-1

KH1.C4.84.v2.A.SL4-1

aniseedV3_3592

Ci-otx

The Genome
Browser allows
visualization of
molecular
information in the
genomic context

Export dataset

All data can be downloaded under formats that comply with established standards:

OBO

MISFISHIE

Fasta

Resources

*4 200 anatomical entities
*31 600 in situ hybridisation
profiles in wild type or mutant
context

*151 articles

*400 cis-regulatory sequences *4 species

*1 Genome (Ciona intestinalis)
*16 000 functionally annotated
transcript models

*230 morpholino sequences *2 millions ESTs

General characteristics of a gene (name, cis-regulatory region, functional annotation, expression data, ESTs and molecular tools)

Gene card

Expression card

General overview of an expression data, with corresponding annotation picture and references

Anatomical card

General characteristics of an anatomical territory (biometry, fate, lineage) inherited from the 3D Virtual embryo and the database

Article card

Access point to all data present in the article and available on ANISEED website (gene studied, regulatory regions tested, evidences, molecular tools, pictures...) allowing an efficient mining on ANISEED data and facilitating curation.

DISCUSSION

ANISEED provides a unique system to understand the developmental strategies of a biological model system. Its concept of virtual embryo allows to describe a species at the level of the genome, of the cell and of the whole embryo. By combining every tools from the system, ANISEED allows to search genes expressed in a given lineage or tissue precursor, find its function and its expression profile at different developmental stages, describe its expression when the function of another gene is disturbed or when blastomeres are ablated, search for genes with similar expression profiles, localize cells and structures where this gene is expressed, knowing their fate, their neighbours, their surface of contact, their volume and surface, both in a 2D and a 3D representation. Additional tools like the NISEED-curator website allows the community to contribute to the annotation data. Management and curation tools ease every aspects of the management of the system: curation, update, database loading. Because this system is entirely generic, all these tools are adaptable to your biological system.