Chapter 35

Creating a User-Friendly and Open-Access Gene Expression Database for Comparing Embryonic Development and Regeneration in *Nematostella vectensis*

Olivier Croce and Eric Röttinger

Abstract

The sea anemone *Nematostella vectensis* has emerged as a powerful research model to understand at the gene regulatory network level, to what extend regeneration recapitulates embryonic development. Such comparison involves massive transcriptomic analysis, a routine approach for identifying differential gene expression. Here we present a workflow to build a user-friendly, mineable, and open-access database providing access to the scientific community to various RNAseq datasets.

Key words Database, Data mining, Open-access, RNAseq, Cnidarian, *Nematostella vectensis*, Regeneration, Embryonic development

1 Introduction

The anthozoan cnidarian *Nematostella vecten*sis (Fig. 1) has initially been developed as a research model organism to gain insights into the evolution of developmental mechanisms and novel cell types/biological features as it is easily cultivable under laboratory conditions [4]. The first cnidarian genome to be sequenced was the one from *Nematostella* that has revealed astonishing similarities with the ones from mammalians [5, 6]. Since then, a wealth of resources and tools have been developed ranging from embryonic RNAseq datasets [1, 2] to meganuclease-induced transgenesis [7], as well as functional approaches for gene know-downs [8–11] and CRISPR/CAS9 mediated knock-outs and knock-ins [11–13].

More recently, *Nematostella* is emerging as a powerful and complementary whole-body regeneration model, as it is able to regrow missing body parts within days after amputation [14–18]. In combination with its historical use as an embryogenesis research model, *Nematostella* is thus very well suited to compare embryonic development and whole-body regeneration within the

Simon Blanchoud and Brigitte Galliot (eds.), *Whole-Body Regeneration: Methods and Protocols*,
Methods in Molecular Biology, vol. 2450, https://doi.org/10.1007/978-1-0716-2172-1_35, © The Author(s) 2022

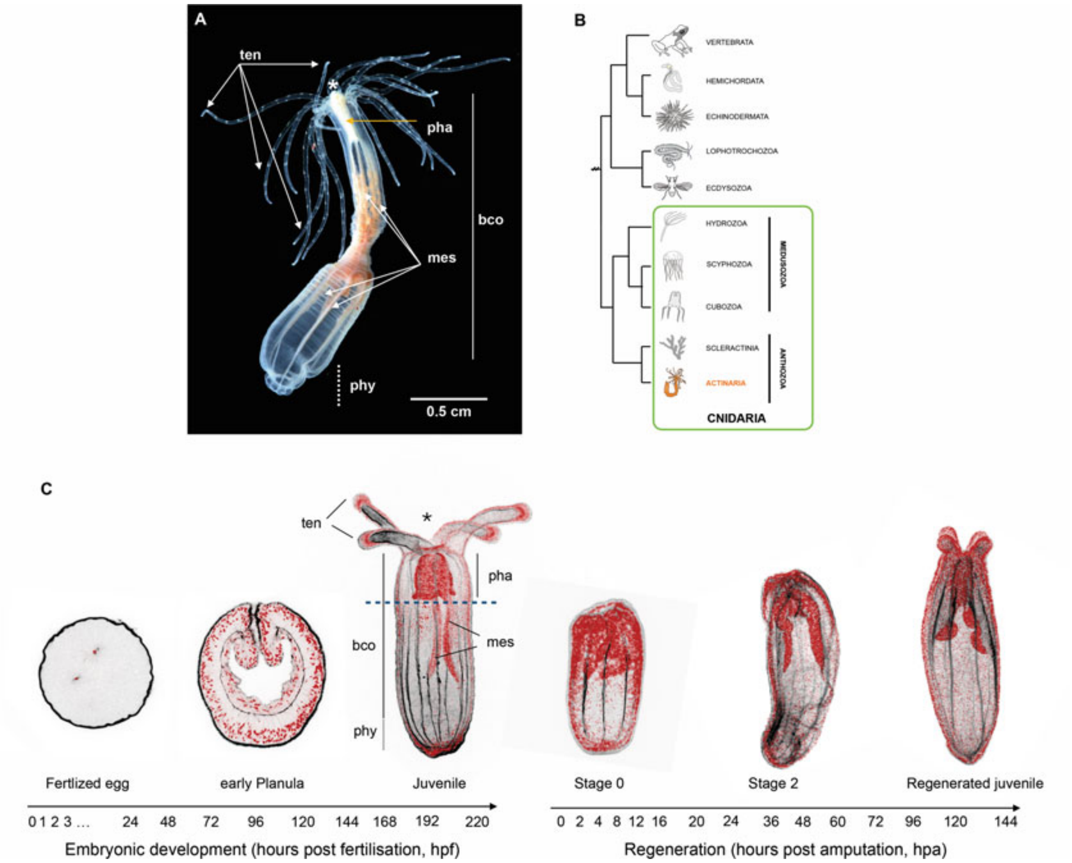


Fig. 1 *Nematostella vectensis*, a research model to assess the relationship between embryonic development and regeneration. (**a**) General anatomy of the sea anemone *Nematostella*. The body of this small anthozoan cnidarian (<5 cm) is organized along an oral/aboral axis. The oral region is formed by tentacles (ten) that surround the mouth (*) and a pharynx (pha). The body column (bco) contains internal structures called mesenteries (mes) that end at the aboral most region with the so-called physa (phy). (**b**) Schematic representation of the phylogenetic position of cnidarians (including *Nematostella*, indicated in orange) within the metazoan tree of life. (**c**) Confocal images (phalloidin/actin filaments in black, DAPI/nuclei in red) of representative stages of *Nematostella* embryonic development and regeneration. Depicted below are the time points for which published RNAseq samples [1–3] are present in the NvERTx database

same organism [3, 17, 19]. This is particularly convenient not only to determine to what extent regeneration recapitulates the gene regulatory program deployed during embryonic development—addressing one of the long-lasting questions in regeneration biology [20], but importantly to highlight if regeneration is controlled by specific toolkits only present in differentiated tissues.

Several studies have performed global gene expression analysis during embryonic development [1–3, 6, 21] or regeneration [3, 22]. In the present chapter, we describe the methods we used to create *NvERTx*, a gene expression database for comparing embryonic development and regeneration gene expression data in

Nematostella vectensis [3, 23]. This database features transcriptomic data combining 22 time points during embryonic development and 16 time points during regeneration (Fig. 1c). In addition, we have developed a web interface to facilitate the manipulation of the data contained in this database. This website simplifies data mining and can generate some figures such as graphs to compare groups of sequences. This interface also integrates various common online tools like Blast to retrieve all transcripts homologous for a given sequence entered by the user. Building such in silico tool is a mean to make RNAseq data accessible to the larger scientific community and enable additional usage of the vast amount of data that were initially produced for a given scientific question.

The workflow presented in this protocol is only a general framework for developing an open-access scientific database. More in-depth knowledge of various aspects of programming and of computer science theory in general might be necessary to implement such tool. Although collaborations will bring together knowledge about the biology to be incorporated and the technology to deploy it online, skills to develop such database can be acquired through a wealth of documentation and literature.

2 Materials

The protocol and material presented here is the one that was used to develop NvERTx, a website to access and mine quantitative transcriptomic data from the sea anemone *Nematostella vectensis* [3]. The protocol will need to be adapted to develop websites presenting other types of data as well as to take into account the unavoidable evolution of software and data format. Yet, this protocol provides a generic and flexible framework to make data accessible and usable to the larger scientific community. In addition to programming skills, we here detail the basic material/data/tools required for the development of a user-friendly transcriptome data mining website. Data implemented in the developed database and website is assumed existent. Refer to [23] for a recent protocol describing a standard workflow for quantifying and performing initial analysis of next-generation sequencing (NGS) reads from an RNA-seq analysis.

2.1 Data Resource Requirements

1. Reference genome (e.g., Nemve1, Table 1).
2. Genome annotations (*see* **Note 1**): Annotations should be provided as a single file in GTF or GFF3 format. These files can be downloaded from each genome-related repository (Genbank or specific website, e.g., Nemve1, Table 1) and do not intend to be generated (or re-generated) by the user. However, it is also possible to perform a gene prediction

Table 1
Genomic and transcriptomic resources for *Nematostella vectensis*

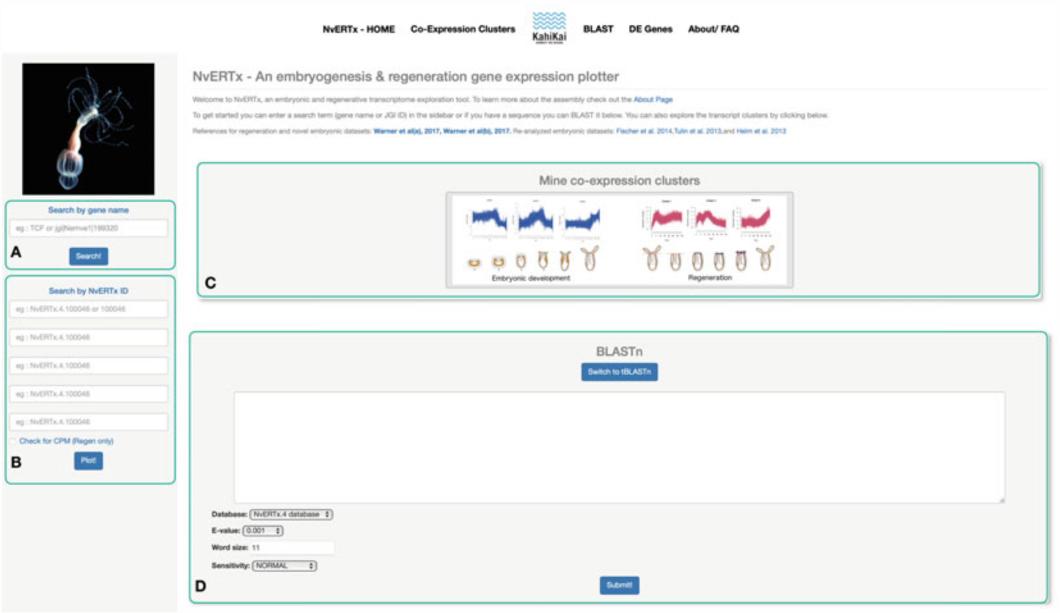
Name	Type	Version	URL	References
Nemve1	Genome, gene models, transcript models, annotations	2.0	https://mycocosm.jgi.doe.gov/pages/search-for-genes.jsf?organism=Nemve1	[5]
Nvec200	Genome, gene models, transcript models, annotations	2.0	https://simrbase.stowers.org/starletseaanemone	[24]
NVE	Gene models, transcript models	2.0	https://figshare.com/articles/Nematostella_vectensis_transcriptome_and_gene_models_v2_0/807696	[25]
NvERTx	Transcript models, annotations	1.0	http://nvertx.kahikai.org or nvertx.ircan.org	[3]

(using AUGUST, Prodigal, or common ORFing prediction tools) if no annotation is available in public databases.

3. Assembled reference transcriptome (e.g., NvERTx, Table 1): Preferably download and use a transcriptome from a public repository (*see* **Note 2**).
4. Transcriptome annotations: We encourage to provide the annotations as GTF or GFF3 files (*see* **Note 3**).
5. Quantified RNA-seq data (e.g., different time-points of WBR, NvERTx, Table 1): The results data from RNA-seq experiments must be easily readable by a script to be stored as records of the database. This implies that the values should have to be sorted in rows and columns in a CSV or text file format.

2.2 Technical Requirements

1. Web server: We used the Apache web server under a Linux computer. A valid domain name and an openly accessible internet access are mandatory. We recommend integrating the website and the database into a Docker container for a reliable security and for more flexibility.
2. Database management system: Database management should follow a relational model. The model organizes data into many tables with a unique key identifying each row. Queries can be performed using the SQL that is the reference language since many decades. We encourage to use MySQL (*see* **Note 4**).
3. Web framework: A modern web development requires to use a robust and reusable syntax. We used Django which is a high-level Python web framework that encourages rapid development and clean, pragmatic design (https://www.djangoproject.com/). Moreover, Django facilitates the communication with the database directly or through additional



NvERTx - An embryogenesis & regeneration gene expression plotter

Welcome to NvERTx, an embryonic and regenerative transcriptome exploration tool. To learn more about the assembly check out the About Page

To get started you can enter a search term (gene name or JGI ID) in the sidebar or if you have a sequence you can BLAST it below. You can also explore the transcript clusters by clicking below.

References for regeneration and novel embryonic datasets: Warner et al(a), 2017, Warner et al(b), 2017. Re-analyzed embryonic datasets: Fischer et al. 2014,Tulin et al. 2013,and Helm et al. 2013

eg : TCF or jgi|Nemve1|199320

eg : NvERTx.4.100046 or 100046

eg : NvERTx.4.100046

eg : NvERTx.4.100046

eg : NvERTx.4.100046

eg : NvERTx.4.100046

□ Check for CPM (Regen only)

Database: NvERTx.4 database

E-value: 0.001

Word size: 11

Sensitivity: NORMAL

Fig. 2 Screenshot of the NvERTx homepage. (**a**) Users can search for genes using the gene name, Nemve1 accession number, NCBI GenBank accession number. (**b**) NvERTx IDs can also be used to directly obtain the temporal expression profiles for the genes of interest. Multiple, up to five, transcripts can be queried simultaneously. (**c**) Users can also directly explore co-expression clusters from embryonic development and regeneration to identify groups of co-expressed genes. (**d**) The transcriptome can also be searched using BLASTn or tBLASTn

libraries such as SQL-Alchemy (https://www.sqlalchemy.org) (*see* **Note 5**). For the design of the site, we have used "Bootstrap" (https://getbootstrap.com) which is an additional framework. It facilitates the creation of web interfaces (HTML + JavaScript) and avoids display problems that can occur between web browsers.

4. Data mining software suite: The web interface of NvERTx includes online software to perform basic manipulations of the RNA-seq sequences (Fig. 2). We added the BLAST suite to facilitate the search of all transcripts of the NvERTx database homologous to a given sequence entered by the user. The aligner MUSCLE was also added to visually compare a group of similar sequences (SNPs, indels, etc.). The integration of these tools in the Django code is facilitated by some specific API such as PyBlast (https://pypi.org/project/pyblast/) or django-blastplus (https://github.com/michalstuglik/django-blastplus).

3 Methods

3.1 Defining the Scope and Application of the Tool

An important aspect for developing a database for internal use or for the scientific community is to define the scope and intended use of the tool. Some preliminaries are therefore required to be clearly defined such as:

1. List the scientific goals of the tool (e.g., Mining transcriptomic data).
2. Define the community for which the tool is intended (e.g., scientists, students, laymen).
3. Define the level of expertise required to use this tool as a user and/or database manager.
4. Determine the time and budget allocated for this project.
5. Determine the functionalities that the users might want to use on the website to query the data (*see* **Note 6**).

3.2 Setup the Relational Database

3.2.1 Database Design

Prior to starting the building of the database and the website per se, the probably most important aspect of the process is to take time to conceptualize the database regarding its intended usage. Taken this into account will make this online tool user-friendly and as useful as possible for the scientific community.

To do so, we encourage you to make an exhaustive inventory of all information and meta-information that will have to be integrated into the database (*see* **Note 7**). This information consists of the biological resources acquired from experiments such as RNA-seq results, but also include every data or metadata that could be related to the numerical values: annotations (genes, go-terms, pathways, etc.), information of the experiments protocol itself (samples sources, extraction method, library preparation, sequencer device used), general or user comments, image descriptions. The data inventory should also include the external sources of information that could be useful to associate: GSEA, ontologies, PubMed, websites of collaborators, etc.

It is also necessary to consider other information that are not directly related to the biological data, but which will be useful for the management of the database and the website: e.g., a table grouping the users with their respective access permissions (guest, administrator, permissions to read or modify part of the data in the database, etc.) (*see* **Note 8**).

3.2.2 Build the Database Structure

A popular method to build databases is based on a relational model that uses linked tables. Depending on the size of the expected database, it is advisable to create a model before directly building the Tables. A database model is a way to set a representation of the data and its relations understandable by a non-expert and independent from any further technical choice of development.

There are several tools or methods to make this model. For example, the Merise method is a general-purpose modeling methodology in the field of information systems development. Merise proceeds to separate treatment of data and processes, where the data-oriented view is modeled in three stages, from conceptual, logical through to physical. The UML (Unified Modeling Language) is also a general-purpose, developmental, modeling language in the field of software engineering that is intended to provide a standard way to visualize the design of a system. A model built using UML proposes a diagram structure representation that can be easily interpreted.

An intermediate solution is to use a graphical table construction and visualization. This type of software allows the biologist to visually verify that all the information is present in the database and in the right place. Also, the relationships between tables can be represented with this type of tools. Once the database schema is finalized, the tool can automatically generate the tables in SQL format for import into a MySQL system for instance. Among these tools, we can mention "MySQL Workbench" which is a free tool and both simple and powerful (https://www.mysql.com/products/workbench/).

Finally, the design of the database must prioritize the independence from the technical solutions that will be used to make the GUI. It must be sufficiently modular to not constrain the use of a specific coding language. It should be possible to interact with the database via an external website, a scripting API, or any interfacing solution that can be developed later (*see* **Note 9**).

3.2.3 Populate the Database

There are several possibilities to populate the database. The most direct is to fill the data directly in SQL command lines or through an external graphical tool like "PHP MyAdmin" (https://www.phpmyadmin.net) which facilitates the access and the management of a database.

Some recurring operations will be easier by using some dedicated scripts that will have to be developed. For example, importing data from RNA-Seq to populate the database is almost impossible to do without using a script that retrieves the values from a raw file and converts them into SQL entries. It will then be necessary to develop a series of scripts, in Python or in a language familiar to the database manager.

If the database is extended, it will be necessary to develop several scripts and to check that the logic of filling the database is respected: some tables must be filled before others, some records must be mandatory, others optional (*see* **Note 10**).

3.3 Build the Website

3.3.1 Design and Create the Main Webpages

In addition to the classical sections such as a Home page describing the project and an About page displaying the various contact details, the important parts of the website will have to be decided in close collaboration with the main end-users. In our example, NvERTx presents a dedicated page for displaying the data of co-expression clusters of embryogenesis versus regeneration and a dedicated page displaying RNA-Seq differential expression by sample type. Each of these webpages are linked to selection menus allowing to display specific information.

Forms to search information must be clearly visible. The entry boxes to search using terms like gene names or by IDs will be often used. They must be accessible directly, e.g., from the left or top main menu. The development of the frontend will be done via the chosen framework (*see* **Note 11**).

3.3.2 Implementing Data Visualization and Mining Tools

It is necessary to establish the list of features that will be required to achieve the project (e.g., count tables, expression plots, expression clusters). These functionalities can be classified into two sets: tools to visualize the data and tools to mine the data.

1. Tools to visualize the data are a way to make the content of the database user-friendly. This interface will be important to enhance the data (gene expression profiles at different timepoints, co-expression clusters, images, annotations, etc.), but also for editing the data. Through interactive forms, this will be a graphical way for a user to add or modify certain information contained in the database as additional annotations or new entries. For NvERTx, expression profiles and volcano plots are generated dynamically using the Plotly javascript library (https://plot.ly/javascript/), a free and open-source graphing library (Figs. 3 and 4).
2. Mining tools are a way to extend the use of the data. A user can experiment two kind of actions: deep information searches, usually for nucleotides/proteins sequences, and comparisons between sets. Sequence searches are done either by entering keywords, usually annotations related to the sequence searched, or by homology by entering a given sequence and asking the system to match all the sequences like it. This last type of search is performed using tools that evaluate the degree of homology of sequences pairwise. BLASTn (used by NvERTx, Fig. 2d) is the most popular tool for this kind of applications, but there are other tools like BLAT, USEARCH, GHOSTX, or GHOSTM which can be used, and which are generally faster. For sequence comparisons, there is a multitude of tools that can be integrated to a web GUI. The simplest ones can consist in multiple sequence aligners like MUSCLE but can be more elaborate like using specific suite as Shiny (https://shiny.rstudio.com) or Ploty (https://plot.ly/javascript/)

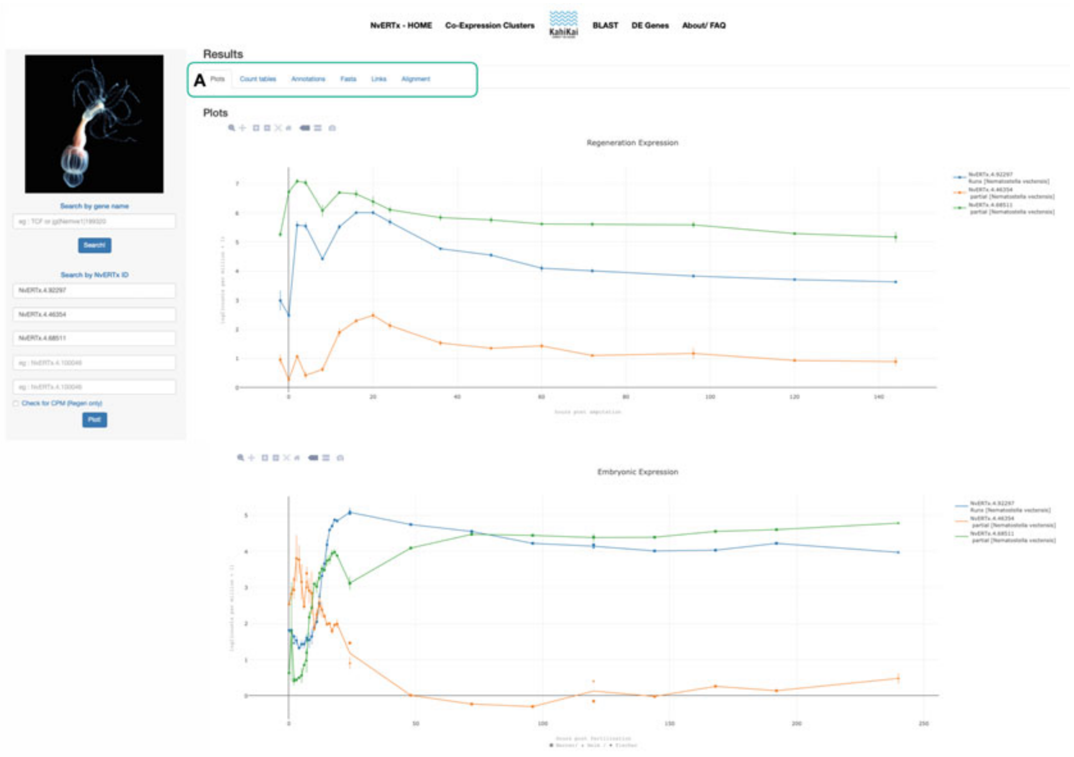


Fig. 3 Screenshot of the expression plots. Shown are plots for regeneration (top) and embryonic development (bottom) for *Nematostella runx* (NvERTx.4.92297)*, tcf/lef* (NvERTx.4.46364)*,* c-ets1B (NvERTx.4.68511). Selecting the tabs in (**a**) enables the users to obtain (1) count data from each of the data sets, (2) transcript annotations, (3) sequences in FASTA format, (4) bibliographical resources including PubMed links and PaperBlast queries (if available), and (5) MUSCLE alignment to compare similar transcripts

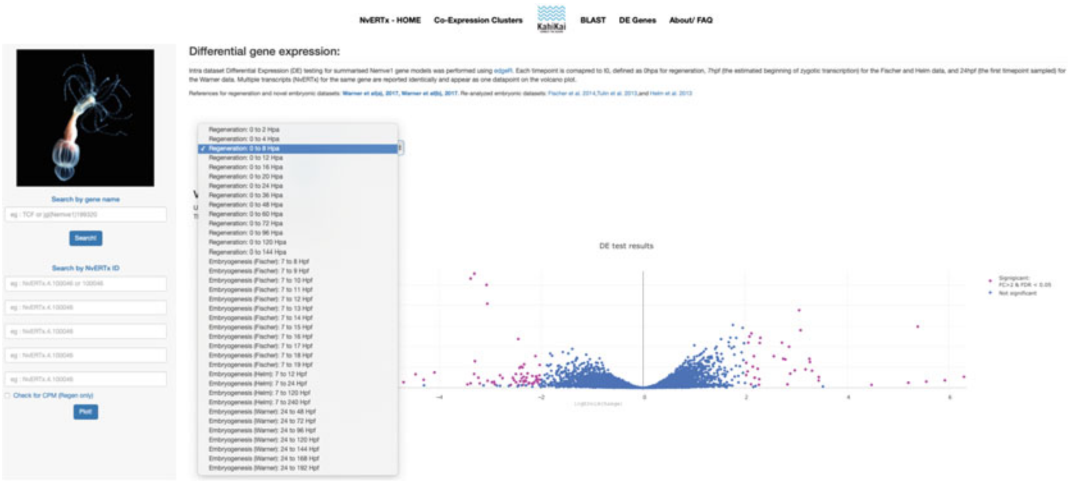
which allows to mix different sets of value to generate online graphics (3D histograms, heatmaps, or more sophisticated representations).

3.4 Data and Services Availability

3.4.1 Accessibility

If the website is hosted on a local web server, unless for a strict internal use, you will need external access permissions for the site to be visible to the rest of the world. The database itself does not necessarily need to have an open access to the outside because it will go through the website. It is recommended to have a real domain name associated with the website to improve its visibility.

The NvERTx website is freely and openly accessible to the community (http://nvertx.kahikai.org). The source code for the website can be found at https://github.com/IRCAN/NvER_plotter_django. Data sets from the database can be found at http://nvertx.ircan.org/ER/ER_plotter/about.


NvERTx - HOME
Co-Expression Clusters
KahiKai
BLAST
DE Genes
About/ FAQ
Search by gene name
eg : TCF or jgi|Nemve1|199320
Search!
Search by NvERTx ID
eg : NvERTx.4.100046 or 100046
eg : NvERTx.4.100046
eg : NvERTx.4.100046
eg : NvERTx.4.100046
eg : NvERTx.4.100046
□ Check for CPM (Regen only)
Plot
Regeneration: 0 to 2 Hpa
Regeneration: 0 to 4 Hpa
✓ Regeneration: 0 to 8 Hpa
Regeneration: 0 to 12 Hpa
Regeneration: 0 to 16 Hpa
Regeneration: 0 to 20 Hpa
Regeneration: 0 to 24 Hpa
Regeneration: 0 to 36 Hpa
Regeneration: 0 to 48 Hpa
Regeneration: 0 to 60 Hpa
Regeneration: 0 to 72 Hpa
Regeneration: 0 to 96 Hpa
Regeneration: 0 to 120 Hpa
Regeneration: 0 to 144 Hpa
Embryogenesis (Fischer): 7 to 8 Hpf
Embryogenesis (Fischer): 7 to 9 Hpf
Embryogenesis (Fischer): 7 to 10 Hpf
Embryogenesis (Fischer): 7 to 11 Hpf
Embryogenesis (Fischer): 7 to 12 Hpf
Embryogenesis (Fischer): 7 to 13 Hpf
Embryogenesis (Fischer): 7 to 14 Hpf
Embryogenesis (Fischer): 7 to 15 Hpf
Embryogenesis (Fischer): 7 to 16 Hpf
Embryogenesis (Fischer): 7 to 17 Hpf
Embryogenesis (Fischer): 7 to 18 Hpf
Embryogenesis (Fischer): 7 to 19 Hpf
Embryogenesis (Helm): 7 to 12 Hpf
Embryogenesis (Helm): 7 to 24 Hpf
Embryogenesis (Helm): 7 to 120 Hpf
Embryogenesis (Helm): 7 to 240 Hpf
Embryogenesis (Warner): 24 to 48 Hpf
Embryogenesis (Warner): 24 to 72 Hpf
Embryogenesis (Warner): 24 to 96 Hpf
Embryogenesis (Warner): 24 to 120 Hpf
Embryogenesis (Warner): 24 to 144 Hpf
Embryogenesis (Warner): 24 to 168 Hpf
Embryogenesis (Warner): 24 to 192 Hpf
DE test results
Significant: FC>2 & FDR < 0.05
Not significant
-4
-2
0
2
4
6
Log2(fold change)

Differential gene expression:

Intra dataset Differential Expression (DE) testing for summarised Nemve1 gene models was performed using edgeR. Each timepoint is comapred to t0, defined as 0hpa for regeneration, 7hpf (the estimated beginning of zygotic transcription) for the Fischer and Helm data, and 24hpf (the first timepoint sampled) for the Warner data. Multiple transcripts (NvERTx) for the same gene are reported identically and appear as one datapoint on the volcano plot.

References for regeneration and novel embryonic datasets: Warner et al(a), 2017, Warner et al(b), 2017. Re-analyzed embryonic datasets: Fischer et al. 2014,Tulin et al. 2013,and Helm et al. 2013

Fig. 4 Screenshot of the differential gene expression (DE Genes) feature from NvERTx. A scroll down menu indicates all possible differential gene expression analysis, whose results are represented by volcano plots. Selecting a given dot will provide additional information such as the NvERTx ID, nr_hits, and fold change values

3.4.2 Maintenance

It is highly recommended to use a Docker container system (https://www.docker.com) or similar. This system allows to encapsulate the database, the site, and all the technical requirements in a single file. This makes the project independent of the native server environment and avoids compatibility problems between different versions or other websites hosted on the same server. Moreover, a container is easily exportable and installable on another server without complicated technical modifications. NvERTx uses a Docker container (https://hub.docker.com/repository/docker/ircan/nvertx).

We recommend depositing all the developments of the project (database structure, website code, API, documents) on a source code development (SCM) platform such as GitHub. This has the advantage of making the code open-source and usable by anyone. Moreover, this kind of sharing platform facilitates collaborative development of IT projects: people who want to improve the project can contribute by modifying the code. The latest version of the project can be made available from this platform. Note that this does not concern the data of the database itself, but only the main source code related to the project.

3.4.3 Documentation

Several documents must be provided. These documents can be text files included at the root level of the GitHub project for example. There will be a README file giving the main lines of the project, an INSTALL detailing the technical procedure to install the database and the associated code of this project, a CHANGELOG specifying the modifications made since the last update of the code. Moreover, a document specific to the practical use of the database and the website is strongly recommended. It could be a pdf that can be downloaded directly from the site detailing the data available and all the items/tools available on the site or an "how to" tutorial.

3.4.4 Evaluating the Impact of the Database

User feedback is the best way to evaluate the usage and user friendliness of a given database. Those feedbacks may not only help to improve the database usefulness itself but also for the development of any potential upgrade or future in silico database. Several tools exist to gain information on the statistics such as number, origin, session durations, etc. of the database users (e.g., GOOGLE analytics). However, certainly the most valuable evaluation factor for scientific databases is its referenced usage by the larger scientific community, i.e., the resulting citations or collaborations. Thus, providing information on how to cite the database is crucial and should be implemented in the webpage (e.g., the FAQ section).

4 Notes

1. A unique file containing genome sequence + annotations (.gbk or .embl) could also be provided instead, but is less useful since the integrated webtools usually needs fasta or gff as separate files, i.e., the blast database indexing requires fasta format only.
2. If no transcriptome is available or if the database manager prefers to use its own transcriptome, the file must be a single fasta file. Each entry in the file has to be a single transcript sequence and the related description should preferably contain some useful IDs to be used for crossing with other databases (e.g., Ensembl, ReSeq, nt).
3. It is possible to add usual short annotations directly into the description lines of the transcriptome fasta file: GeneSymbol name, ortholog names/IDs, GO term ID, or other information specific to the transcript.
4. MySQL is an open-source relational database management system widely used in the biology field. Its capabilities allow to support a high number of entries as well as a powerful engine for complex queries. For small databases is also possible to use SQLite instead, having the advantage to be lighter and using an independent file system for storing data. Like the webserver, prefer to install and run your database server into a separate Docker container.
5. Another option is to use "Ruby on Rails" (https://rubyonrails.org) instead of Django. This framework is similar to Django

with some differences like a fully object-oriented language with a more powerful frontend part. Ruby is probably slightly more difficult to handle for beginner developers.

6. Start from your personal / the labs' needs and exchange with the community to determine additional functionalities to be added. Often additional functionalities arise during the test phase of the database/website. Once launched, consider taking users' feedback for a future update of the data mining tools.
7. Some data require high amount of space storage such as raw sequence or high-definition images. They can be stored outside the database, in files. Exporting this big data will keep the database quite light; however, it will be required to add links to these external files into the database (system directory paths, external URLs, etc.).
8. The database should be modular enough to allow additional tables or fields in the future.
9. Apply the traditional rules of database design, i.e., to avoid redundancy, to propose unique keys for each table, to organize the data in the different tables in a way that is neither too large nor too small, and to create intermediate tables containing the keys of the tables whose records are to be crossed. We advise to carefully read up some documentation on relational database conception to avoid certain errors because it can be very difficult to modify the structure of the database once it contains the data and starts to be used.
10. The development of these small pipelines can be organized in such a way that it will constitute a real Application Programming Interface (API), which can be reused during the construction of the website.
11. The code and the SQL queries generated when validating the forms must be clear and modular. The possible APIs developed to fill the database can be reused here. The developer can use free existing templates and CSS files to build the site, or he can decide to develop the site from scratch. He can also use a content management system (CMS) to set up the website (such as Zope for Django), but we do not advise this solution as these CMS are not appropriate for very specialized sites like NvERTx.

Acknowledgments

The authors thank Jacob Warner and Vincent Guerlais for the development of NvERTx as well as Boris Meyer for technical support after the launch of the online database. This work was supported by the French Government (National Research Agency, ANR) through the grant RENEW (ANR-19-PRC) to E.R.

References

1. Fischer, Antje HL, Smith J (2013) Nematostella High-density RNAseq time-course, 2013-06-14.
2. Helm RR, Siebert S, Tulin S, Smith J, Dunn CW (2013) Characterization of differential transcript abundance through time during Nematostella vectensis development. BMC Genomics 14(1):266
3. Warner JF, Guerlais V, Amiel AR, Johnston H, Nedoncelle K, Röttinger E (2018) NvERTx: a gene expression database to compare embryogenesis and regeneration in the sea anemone Nematostella vectensis. Development 145(10): dev162867
4. Hand C, Uhlinger KR (1992) The culture, sexual and asexual reproduction, and growth of the sea anemone Nematostella vectensis. Biol Bull 182(2):169–176. https://www.researchgate.net/
5. Putnam NH, Srivastava M, Hellsten U, Dirks B, Chapman J, Salamov A et al (2007) Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. Science 317(5834):86–94
6. Schwaiger M, Schonauer A, Rendeiro AF, Pribitzer C, Schauer A, Gilles AF et al (2014) Evolutionary conservation of the eumetazoan gene regulatory landscape. Genome Res 24(4): 639–650
7. Renfer E, Amon-Hassenzahl A, Steinmetz PRH, Technau U (2010) A muscle-specific transgenic reporter line of the sea anemone, Nematostella vectensis. Proc Natl Acad Sci U S A 107(1):104–108
8. Layden MJ, Röttinger E, Wolenski FS, Gilmore TD, Martindale MQ (2013) Microinjection of mRNA or morpholinos for reverse genetic analysis in the starlet sea anemone, Nematostella vectensis. Nat Protoc 8(5):924–934
9. Röttinger E, Dahlin P, Martindale MQ (2012) A framework for the establishment of a Cnidarian gene regulatory network for "Endomesoderm" specification: the inputs of ß-catenin/TCF signaling. PLoS Genet 8(12):e1003164
10. Rentzsch F, Fritzenwanker JH, Scholz CB, Technau U (2008) FGF signalling controls formation of the apical sensory organ in the cnidarian Nematostella vectensis. J Embryol Exp Morphol 135(10):1761–1769
11. Servetnick MD, Steinworth B, Babonis LS, Simmons D, Salinas-Saavedra M, Martindale MQ (2017) Cas9-mediated excision of Nematostella brachyury disrupts endoderm development, pharynx formation and oral-aboral patterning. Development 144(16):2951–2960
12. Ikmi A, McKinney SA, Delventhal KM, Gibson MC (2014) TALEN and CRISPR/Cas9-mediated genome editing in the early-branching metazoan Nematostella vectensis. Nat Commun 5:5486
13. Kraus Y, Aman A, Technau U, Genikhovich G (2016) Pre-bilaterian origin of the blastoporal axial organizer. Nat Commun 7:11694
14. Bossert PE, Dunn MP, Thomsen GH (2013) A staging system for the regeneration of a polyp from the aboral physa of the anthozoan cnidarian Nematostella vectensis. Dev Dyn 242: 1320–1331
15. Amiel A, Johnston H, Nedoncelle K, Warner J, Ferreira S, Röttinger E (2015) Characterization of morphological and cellular events underlying oral regeneration in the sea anemone, Nematostella vectensis. Int J Mol Sci 16(12):28449–28471
16. Passamaneck YJ, Martindale MQ (2012) Cell proliferation is necessary for the regeneration of oral structures in the anthozoan cnidarian Nematostella vectensis. BMC Dev Biol 12(1):1
17. Reitzel A, Burton P, Krone C, Finnerty J (2007) Comparison of developmental trajectories in the starlet sea anemone Nematostella vectensis: embryogenesis, regeneration, and two forms of asexual fission. Invertebr Biol 126(2):99–112
18. Amiel AR, Röttinger E (2021) Experimental tools to study regeneration in the sea anemone Nematostella vectensis. Methods Mol Biol 2219:69–80

19. Warner JF, Amiel AR, Johnston H, Röttinger E (2019) Regeneration is a partial redeployment of the embryonic gene network. bioRxiv 121: 658930
20. Morgan TH (1901) Regeneration, Columbia University Biological series, vol 279. Macmillan, New York, NY
21. Tulin S, Aguiar D, Istrail S, Smith J (2013) A quantitative reference transcriptome for Nematostella vectensis early embryonic development: a pipeline for de novo assembly in emerging model systems. EvoDevo 4(1):16
22. Schaffer AA, Bazarsky M, Levy K, Chalifa-Caspi V, Gat U (2016) A transcriptional time-course analysis of oral vs. aboral whole-body regeneration in the sea anemone Nematostella vectensis. BMC Genomics 17(1):718
23. Warner JF, Röttinger E (2021) Transcriptomic analysis in the sea anemone Nematostella vectensis. Methods Mol Biol 2219:231–240
24. Zimmermann B, Robb SMC, Genikhovich G, Fropf WJ, Weilguny L, He S et al (2020) Sea anemone genomes reveal ancestral metazoan chromosomal macrosynteny. bioRxiv 2020.10.30.359448
25. Fredman D, Schwaiger M, Rentzsch F, Rentzsch F, Technau U (2013) Nematostella vectensis transcriptome and gene models v2.0

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.