The Dugesia ryukyuensis Database as a Molecular Resource for Studying Switching of the Reproductive System

AUTHOR(S):
Ishizuka, Hideyuki; Maezawa, Takanobu; Kawauchi, Junpei; Nodono, Hanae; Hirao, Yukako; Nishimura, Osamu; Nakagawa, Haruka; ... Kobayashi, Kazuya; Sakakibara, Yasubumi; Matsumoto, Midori

CITATION:

ISSUE DATE:
2007-01

URL:
http://hdl.handle.net/2433/85330

RIGHT:
(c) 日本動物学会 / Zoological Society of Japan
The Dugesia ryukyuensis Database as a Molecular Resource for Studying Switching of the Reproductive System

Hideyuki Ishizuka1,2†, Takanobu Maezawa1,3†, Junpei Kawauchi1, Hanae Nodono1, Yukako Hirao4, Osamu Nishimura4, Haruka Nakagawa1, Kiyono Sekii1, Kenta Tasaka1, Hiroshi Tarui4, Kiyokazu Agata4,5, Motonori Hoshi1, Kazuya Kobayashi1, Yasubumi Sakakibara1 and Midori Matsumoto1*

1Department of Biological Sciences and Informatics, Keio University, 3-14-1, Hiyoshi, Kouhoku-ku, Yokohama 223-8522, Japan
2Department of Computational Biology, Graduate School of Frontier Sciences, University of Tokyo, Kashiwanoha 5-1-5, Kashiwa 277-8561, Japan
3Okazaki Institute for Integrative Bioscience, National Institute for Basic Biology, National Institutes of Natural Science, Higashiyama Myodaiji, Okazaki 444-8787, Japan
4Genome Resource and Analysis Subunit, RIKEN Center for Developmental Biology, 2-2-3, Minatojima-minamimachi, Chuo-ku, Kobe 650-0047, Japan
5Department of Biophysics, Graduate School of Science, Kyoto University, Kitashirakawa-Oiwake, Sakyo-ku, Kyoto 606-8502, Japan

The planarian Dugesia ryukyuensis reproduces both asexually and sexually, and can switch from one mode of reproduction to the other. We recently developed a method for experimentally switching reproduction of the planarian from the asexual to the sexual mode. We constructed a cDNA library from sexualized D. ryukyuensis and sequenced and analyzed 8,988 expressed sequence tags (ESTs). The ESTs were analyzed and grouped into 3,077 non-redundant sequences, leaving 1,929 singletons that formed the basis of unigene sets. Fifty-six percent of the cDNAs analyzed shared similarity (E-value<1E -20) with sequences deposited in NCBI. Highly redundant sequences encoded granulin and actin, which are expressed in the whole body, and other redundant sequences encoded a Vasa-like protein, which is known to be a component of germ-line cells and is expressed in the ovary, and Y-protein, which is expressed in the testis. The sexualized planarian expressed sequence tag database (http://planaria.bio.keio.ac.jp/planaria/) is an open-access, online resource providing access to sequence, classification, clustering, and annotation data. This database should constitute a powerful tool for analyzing sexualization in planarians.

Key words: planarian, EST analysis, sexualization, asexual-sexual switch, gene ontology

INTRODUCTION

Living organisms have established unique reproductive systems to multiply and sustain their species. Two types of reproduction evolved, namely, asexual and sexual. Asexual reproduction is much simpler than sexual reproduction, and it produces offspring that are genetically identical to the parent and in which no genotypic variation occurs. In contrast, sexual reproduction reshuffles genes and results in offspring genetically different from parents and siblings. Hence, considerable genotypic variation occurs in the case of sexual reproduction (Bell, 1982).

It is well known that planarians possess remarkable regenerative abilities requiring the activity of neoblasts, which are the only known proliferating cells (Agata and Watanabe, 1999; Orii et al., 2005; Agata et al., 2006). It is also known that some planarians switch between asexual and sexual reproduction in response to environmental signals (Curtis, 1902, Hyman, 1939). To understand the mechanism of switching reproductive systems in planarians, we established an experimental bioassay system for sexual switching. Dugesia ryukyuensis of the OH strain, which have been maintained under laboratory conditions for more than 15 years through asexual reproduction, were induced to reproduce sexually by feeding them sexually mature individuals of Bdell-
locephala brunnea, an exclusively oviparous species (Kobayashi et al., 1999). Following induction, the D. ryukyuensis OH strain gradually developed sexual organs over time, and after a given point (the point-of-no-return), could not return to an asexual state (Kobayashi and Hoshi, 2002).

To understand the selection of reproductive systems and the differentiation of sexual organs, we isolated the key genes that function in sexual induction and that are expressed specifically during the sexualizing process. We detected the gene Dryg, which is expressed specifically in yolk glands of worms after the point-of-no-return (Hase et al., 2002). Dryg expression during the sexualizing process is limited to a single cell type possessing characteristics of neoblasts, which are totipotent somatic cells. We also detected other genes expressed in sexual organs (unpublished observations). However, the eventual goal is to identify all genes involved in sexual induction.

To obtain a comprehensive catalog of genes responsible for sexual induction, we are currently conducting a large-scale EST analysis of cDNA obtained from D. ryukyuensis.

**Fig. 1.** Gene-expression profile in sexualized planarians (D. ryukyuensis). Frequency of occurrence is plotted for each non-redundant clone. Most (63%) of the clones were expressed as singletons.

**Fig. 2.** BLASTX analysis of genes from sexualized planarians (D. ryukyuensis). The distribution of the BLASTX matches is according to E value. The number of ESTs for the E-value ranges is indicated above each bar.
Here, we describe this database and provide an overview of the results obtained thus far. Updated results are available on the website (http://planaria.bio.keio.ac.jp/planaria/).

**MATERIALS AND METHODS**

**Animals**
A clonal line of the asexual strain of *D. ryukyuensis* was established by Dr. S. Ishida (Hirosaki University, Aomori, Japan). This strain was named the OH strain. The animals were maintained at 20°C.

**Construction of a cDNA library**
The OH strain was fully sexualized by using the feeding procedure of Kobayashi et al. (1999). Total RNA was extracted from more than 30 fully sexualized planarians. Poly (A) RNA was purified using

### Table 1. High-ranking redundant clusters of EST sequences.

<table>
<thead>
<tr>
<th>cluster_id</th>
<th>length</th>
<th>num</th>
<th>representation</th>
<th>id</th>
<th>description</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DrC_0023</td>
<td>665</td>
<td>146</td>
<td>gb</td>
<td>AAP44511.1</td>
<td>progranulin-b [<em>Danio rerio</em>]</td>
<td>6.00E-26</td>
</tr>
<tr>
<td>DrC_00499</td>
<td>1433</td>
<td>116</td>
<td>dbj</td>
<td>BAC4866.1</td>
<td>actin [Galaxea fascicularis]</td>
<td>&lt;1.0E-170</td>
</tr>
<tr>
<td>DrC_00802</td>
<td>759</td>
<td>87</td>
<td>embi</td>
<td>CAD21525.1</td>
<td>hypothetical protein [Taenia solium]</td>
<td>1.50E-02</td>
</tr>
<tr>
<td>DrC_00591</td>
<td>1065</td>
<td>85</td>
<td>ref</td>
<td>NP217231.1</td>
<td>hatching gland gene 1 [<em>Danio rerio</em>]</td>
<td>3.00E-91</td>
</tr>
<tr>
<td>DrC_00208</td>
<td>1418</td>
<td>79</td>
<td>gb</td>
<td>AA92882.1</td>
<td>fimbrin</td>
<td>2.00E-43</td>
</tr>
<tr>
<td>DrC_00505</td>
<td>2152</td>
<td>73</td>
<td>dbj</td>
<td>BAD12572.1</td>
<td>heat shock protein [Numida meleagris]</td>
<td>&lt;1.0E-170</td>
</tr>
<tr>
<td>DrC_00493</td>
<td>607</td>
<td>66</td>
<td>emb</td>
<td>CAB62280.1</td>
<td>hydroxyproline-rich glycoprotein DZ-HRGP [Volvox carteri f. nagariensis]</td>
<td>3.30E-01</td>
</tr>
<tr>
<td>DrC_00181</td>
<td>853</td>
<td>62</td>
<td>dbj</td>
<td>BA25398.1</td>
<td>CsCA1 [<em>Ciona savignyi</em>]</td>
<td>3.00E-27</td>
</tr>
</tbody>
</table>

### Table 2. Reproduction and regeneration-related genes in sexualized planarian EST sequences.

<table>
<thead>
<tr>
<th>cluster_id</th>
<th>length</th>
<th>num</th>
<th>representation</th>
<th>id</th>
<th>description</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>DrC_0023</td>
<td>665</td>
<td>146</td>
<td>gb</td>
<td>AAP44511.1</td>
<td>progranulin-b [<em>Danio rerio</em>]</td>
<td>2.00E-25</td>
</tr>
<tr>
<td>DrC_00812</td>
<td>1085</td>
<td>22</td>
<td>emb</td>
<td>CA66879.1</td>
<td>Y-box protein [<em>Dugesia japonica</em>]</td>
<td>1.00E-138</td>
</tr>
<tr>
<td>DrC_00495</td>
<td>513</td>
<td>8</td>
<td>pirl</td>
<td>IS0247</td>
<td>class V zygote-specific protein - <em>Chlamydomonas reinhardtii</em></td>
<td>4.00E-24</td>
</tr>
<tr>
<td>DrC_00081</td>
<td>637</td>
<td>5</td>
<td>dbj</td>
<td>BAB12217.1</td>
<td>vasa homolog [<em>Ciona savignyi</em>]</td>
<td>2.00E-62</td>
</tr>
<tr>
<td>DrC_00456</td>
<td>1591</td>
<td>5</td>
<td>dbj</td>
<td>AJ439094.1</td>
<td>Y-box protein [<em>Dugesia etrusca</em>]</td>
<td>1.00E-159</td>
</tr>
</tbody>
</table>

![Fig. 3. Comparison of the deduced amino acid sequence of DrC-00456 with those of other Y-box proteins in planarians: *DeY1* in *D. etrusca* (accession no. AJ512636) and *DjY1* in *D. japonica* (accession no. AJ512635). Asterisks indicate identical residues. The multiple alignment was obtained by CLUSTAL W (1.83) multiple-sequence alignment.](image-url)
an Oligotex-dT30<sup>Super</sup> mRNA Purification Kit (Takara, Shiga, Japan) and was converted to oligo(dT)-primed cDNAs with a cDNA Synthesis Kit with oligo(dT)<sub>18</sub> anchor primer (Stratagene). Both ends of the cDNAs were linked to EcoRI adapters. The cDNAs were digested with EcoRI and ligated into the Bluescript II SK<sup>(+)</sup> plasmid digested with EcoRI and XhoI. Clones containing the inserts were then selected for nucleotide sequencing.

**EST sequences**

The cDNAs were introduced into E.coli DH10B competent cells (Invitrogen) by using a Gene Pulser electroporation system (Bio-Rad Laboratories) and plated onto LB-ampicillin plates (25 cm<sup>2</sup> x 25 cm). Purified plasmid DNA was sequenced with a 3700 ABI sequencer and a Big-Dye Terminator Sequencing Kit according to the manufacturer’s instructions (Applied Biosystems). The 5’ end one-pass sequences were read, and a total of 10,745 ESTs were obtained. Prior to the computational analysis of ESTs, contaminating subsequences such as repeat and vector sequences were identified by a BLASTN similarity search against the UniVec vector database (ftp://ftp.ncbi.nih.gov/pub/univec) and were eliminated.

**Classification, clustering and annotation**

To control EST quality, the sequence data were processed automatically to remove the polyA tail and vector sequences. For classification, sequences obtained underwent automatic BLASTX searches against NCBI’s GenBank. The redundancy of sequences was checked with the Contig Assembling Program (CAP3) (Huang et al., 2004). Sequences were annotated with gene ontology (GO) terms by searching against the Gene Ontology Database (Harris et al., 2004).

**In situ hybridization and histology**

Whole-mount in situ hybridization was performed using the protocols described by Umesono et al. (1997). Sense and anti-sense riboprobes were prepared by in vitro transcription using RNA polymerase with digoxigenin-UTP (Roche).

**RESULTS AND DISCUSSION**

**An overview of ESTs derived from the sexualized planarians**

The asexual Dugesia ryukyuensis OH strain was fully sexualized by the method of Kobayashi et al. (1999), and a cDNA library was generated. From this cDNA library, we obtained from the fully sexualized planarians 8,988 5’ EST cDNA sequences after the removal of clones with no inserts. The average length of sequences was 600 bp. Sequences longer than 100 bp were registered with GenBank (accession numbers BW632996–BW643983).

**Annotation and clustering of the sexualized planarian EST sequences**

To classify the sequences obtained, ESTs were subjected to a BLASTX homology search and annotated against the Uniprot / Swissprot database maintained by the European Bioinformatics Institute (EBI). Of the 8,988 sequences obtained, 4,902 exhibited high similarity (E-Value<1E-20) to proteins in other organisms. To eliminate sequence redundancy, EST clustering was performed using CAP3. A group of sequences was considered clustered when sequence similarity was greater than 90% (-p option) and an overlap of more than 49 bp existed (-o option). The 8,988 ESTs were assembled into 3,077 clusters consisting of 1,147 contigs and 1,930 singletons. Fig. 1 shows the frequency distribution of clusters based on the number of clusters and the frequency of each cluster obtained. A unique identifier starting from DrC-00001 and continuing to DrC-03077 was assigned to each assembled sequence to provide annotations in a manner similar to EST. Annotations of consensus sequences are also available in the database.

These planarian EST sequences were subjected to BLASTX analysis for annotation. Of the 3,077 clusters, 1,646 (53%) were highly similar (E-value<1E-20) to functionally characterized genes in other organisms (Fig. 2). A total of 1,158 clusters showed low similarity (1E-20<E-value<1E-20) to functionally characterized genes in other organisms. The distribution of GO slim in the Biological Process category associated with sexualized D. ryukyuensis ESTs is shown in Table 3.

<table>
<thead>
<tr>
<th>GO slim ID</th>
<th>description</th>
<th>Number of cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0005622</td>
<td>intracellular</td>
<td>610</td>
</tr>
<tr>
<td>GO:0005623</td>
<td>cell</td>
<td>195</td>
</tr>
<tr>
<td>GO:0005575</td>
<td>cellular component</td>
<td>46</td>
</tr>
<tr>
<td>GO:0005576</td>
<td>extracellular</td>
<td>40</td>
</tr>
<tr>
<td>GO:0005941</td>
<td>unlocalized protein complex</td>
<td>10</td>
</tr>
<tr>
<td>GO:0008372</td>
<td>unknown</td>
<td>0</td>
</tr>
<tr>
<td>total</td>
<td></td>
<td>901</td>
</tr>
</tbody>
</table>

**Table 3.** The distribution of GO slim in the Biological Process category associated with sexualized D. ryukyuensis ESTs.
no similarity was observed for 283 clusters.

**High-ranking redundant clusters of EST sequences**

It has previously been shown that highly redundant clusters usually correspond to highly expressed genes. Highly redundant clusters of ESTs for *D. ryukyuensis* are shown in Table 1. The most redundant cluster, DrC-00023, is homologous to granulin (grn, E-value<2E-25), a cytotrophic factor modulating the growth of various cells in vitro (Bateman and Bennett, 1998). DrC-00023 was present in 146 copies in the EST-DB, suggesting that it is very abundantly expressed. This indicates that DrC-00023 could act as a growth factor for promoting strong regeneration in planarians. Interestingly, it has been reported that grn precursor gene expression increases in the neonatal rat hypothalamus in response to estrogen treatment during the perinatal period (Suzuki et al., 2001) and that the grn product functions as a hormone in the brain during sex differentiation in mammals. Other highly redundant clusters included housekeeping gene families such as actin and heat-shock proteins.

A list of highly redundant gene clusters involved in reproduction and regeneration is shown in Table 2. Cluster DrC-00812 exhibited high similarity to the planarian Y-box gene *DjY1* (E-value<1E-138), which is specifically expressed in regeneration blastemas in *D. japonica* (Salvetti et al., 1998). Cluster DrC-00456 was very similar to another planarian Y-box gene, *DeY1* (E-value<1E-159; Fig. 3). This gene is expressed in testicular germ cells (spermatogonia, spermatocytes, and spermatids) of the sexual planarian *D. etrusca*, whereas no expression is detected in blastemas (Salvetti et al., 2002). *DjY1* is not highly similar to DrC-00456. DrC-00495 is similar to a class V zygote-specific gene (E-value 4E-13) that functions in gametogenesis (Matters and Goodenough, 1992). DrC-00081 is similar to Vasa (E-value 1E-40), which is a member of the ATP dependent helicase family expressed specifically in stem cells in many organisms and which is regarded as a stem-cell marker (Hay et al., 1998; Shibata et al., 1999).

**Classification of genes according to their functions**

Genes were classified according to gene ontology (GO). If the E-value of a sequence annotated by the BLASTX search against the Uniprot protein database was lower than 1E-15, the sequence was annotated with corresponding GO terms by referencing the Gene Ontology Annotation (GOA) database. Additionally, GO slims, which are slimmed-down versions of GO terms, were used for summarizing the GO term distribution of GO-annotated sequences. Table 3 shows the distribution of GO slim categories of consensus sequences. The major transcripts in the EST-DB complied here encode intracellular proteins playing catalytic roles during physiological processes. The updated results are available on the website (http://planaria.bio.keio.ac.jp/planaria/).

**Expression patterns of high-ranking redundant clusters related to development**

To identify highly redundant gene clusters specifically expressed during the sexualization process, whole-mount *in situ* hybridization was performed for the highly redundant gene clusters DrC-00023 (a granulin homolog) and DrC-
00499 (an actin homolog). Expression patterns of these genes are shown in Fig. 4A-D. DrC-00023 and DrC-00499 are expressed in the whole body of sexual and asexual worms, not specifically in the brain, suggesting that DrC-00023 may not function as a brain hormone during sexual differentiation. As granulin is known as a pleiotropically expressed growth factor that mediates cell-cycle progression and cell motility in mammals (Bateman and Bennett, 1998; He and Bateman, 2003), DrC-00023 may function as a growth factor in planarians.

Since DrC-00456 (DeY1 homolog) and DrC-00081 (Vasa homolog) have high copy numbers in the EST-DB and high E-values, whole-mount in situ hybridization was also performed for these highly redundant gene clusters. DrC-00456 was specifically expressed in the testes of sexual worms, but not expressed in asexual worms (Fig. 4E, F). Interestingly, DeY1 from D. etrusca (Salvetti et al., 2002) is likely a homolog of DrC-00456.

Asexual worms do not maintain genital organs such as the testis, yolk gland, etc., but exhibit small ovarian primordia with a few oogonia. DrC-00081 was expressed specifically in the ovary of sexual worms and in these ovarian primordia in asexual worms (Kobayashi et al., 1999). Planarian Vasa-like genes, vlg-A and vlg-B, are transcribed in the ovary and testis (Shibata et al., 1999). Since DrC-00081 mRNA was detected only in the ovary in sexual worms or the ovarian primordia in asexual worms, but not in the testis (Fig. 4G-I), DrC-00081 could be a novel Vasa-like gene.

Future directions
Several planarian databases have recently been established. An EST-DB generated from the head portion of D. japonica should facilitate understanding of the evolutionary divergence of the planarian central nervous system (Mineta et al., 2003). EST-DBs generated from asexual and sexual stages of the hermaphroditic strain of Schmidtea mediterranea (Zayas et al., 2005) enable a comparison of the asexual and sexual stages of planarians and the analysis of germ-cell specification. However, they cannot trace the sexualization process and are not appropriate for studying processes such as oogenesis or spermatogenesis.

In contrast, our EST-DB facilitates studies related to these processes by identifying genes activated in the switch from asexual to sexual reproductive mode. For example, microarray probes of sexualized planarians could be prepared from our EST-DB and used to analyze gene expression at each stage of sexualization. We can also classify genes according to their functions, such as sexual hormones, sexual hormone receptors, or hormone-synthesizing enzymes. This sexualized planarian EST-DB is therefore an excellent tool for studying the process of the sexualization in planarians.

ACKNOWLEDGMENTS
We thank Dr. Sachiko Ishida, Hirosaki University, for the kind gift of the OH strain. This work was supported in part by Ministry of Education, Culture, Sports, Science, and Technology grants to M.M. and M.H.; by Special Coordination Funds for Promoting Science and Technology to K.A.; and by Grants-in-Aid for Creative Research to K.A. and Scientific Research on Priority Areas to K.A.

REFERENCES
Hyman LH (1939) North American triclad Turbellaria. IX. The priority of Dugesia Girard 1850 over Euplanaria Hesse 1897 with notes on American species of Dugesia. Trans Amer Micros Soc 58: 264–275


(Received May 25, 2006 / Accepted September 27, 2006)