

# An Untranslated RNA Sequence Encoded in the Putative Regulatory Region of *fruitless*, a Major Courtship Regulator Gene of *Drosophila*

Sakino Takayanagi, Gakuta Toba & Masayuki Koganezawa  
Division of Neurogenetics  
Tohoku University Graduate School of Life Sciences  
Sendai, Miyagi 980-8577, Japan

Tamas Lukacsovich  
Developmental Biology Center  
Department of Developmental and Cell Biology  
University of California Irvine  
Irvine, CA 92697-2300, USA

Daisuke Yamamoto (Corresponding author)  
Division of Neurogenetics  
Tohoku University Graduate School of Life Sciences  
Sendai, Miyagi 980-8577, Japan  
Tel: 81-22-217-6218 E-mail: daichan@m.tokoku.ac.jp

Received: November 4, 2011 Accepted: November 18, 2011 Published: January 1, 2012  
doi:10.5539/ijb.v4n1p20 URL: <http://dx.doi.org/10.5539/ijb.v4n1p20>

*This study was supported by Grants-in-Aid for Scientific Research from the Japanese Government Ministry of Education, Culture, Sports, Science, and Technology (23220007, 1802012) to D.Y., by the Strategic Japanese-French Cooperative Program from the Japan Science and Technology Agency to D.Y., by a Life Science Grant from the Takeda Science Foundation to D. Y., by the Japan Society for the Promotion of Science (JSPS) Invitation Program (short-term) to T. L., and by a Grant-in-Aid for JSPS fellows to S. T. (Sponsoring information)*

## Abstract

At about 13 kb upstream from the P1 promoter of the *Drosophila melanogaster fruitless (fru)* gene, we identified a putative non-coding RNA (ncRNA) gene of ~150 bp (*fru*-upstream ncRNA: *fuR*). *fuR* is transcribed from the DNA strand complementary to that encoding the *fru* transcripts in both female and male adults. Of the 10 species of the subgenus *Sophophora* examined, 9 have the highly conserved *fuR* sequence. The exception is *D. willistoni*, in which a 38 bp insertion splits *fuR* into two segments. Three species of subgenus *Drosophila* examined also have *fuR* with minor insertions or deletions. It remains to be determined whether or not *fuR* modulates *fru* functions and whether or not the diversification in this sequence correlates to species differences in *fru* functions.

**Keywords:** Non-coding RNA, *D. melanogaster*, *fruitless*, Evolution

## 1. Introduction

It has long been recognized that there exist non-coding RNAs, the best characterized of which are the chromatin regulators for gene dosage compensation, such as *Tsix* and *Xist* in mammals and *RoX* in *Drosophila* (Straub & Baker, 2007). Recently, enormous efforts have been made to comprehensively characterize transcribed sequences

and chromatin dynamics across the entire genome, leading to the discovery of nearly 2000 previously undescribed genes, 90% of which appear to be non-protein-coding RNA (ncRNA) genes (modENCODE Consortium, 2010). These include the major classes of structural RNAs, snoRNAs, and Argonaute-associated small RNAs such as miRNAs, siRNAs, and piwi-associated piRNAs (Sayed & Abdellatif, 2011). Although some of these small RNAs are known to be involved in splicing (snRNAs), rRNA modification (snoRNAs), RNA interference (siRNAs), and transposon silencing (piRNAs), the functions of many other ncRNAs remain to be elucidated (Sayed & Abdellatif, 2011). Recent global run-on-sequencing assays revealed the existence of low-level-bi-directional transcripts in the transcriptionally active regulatory region of many *Drosophila* genes (Kharchenko et al., 2011). These ncRNAs are analogous to the enhancer RNAs (eRNAs) found to be transcribed from the neuronal activity-regulated enhancers in mice; the level of eRNA expression at neuronal enhancers is positively correlated with the level of mRNA synthesis at nearby genes (Kim et al., 2010).

We have been characterizing the enhancer region of the *fruitless (fru)* gene that specifies the major part of neuronal circuitry for the production of male courtship behavior in *Drosophila* (Dickson, 2008; Yamamoto, 2008). The male courtship ritual is highly stereotypic and species-specific (Bastock and Manning, 1955; Spieth, 1952; Yamamoto et al., 1997), and *fru* was suggested to determine some of its species-specific characteristics by QTL analysis (Gleason and Ritchie, 2004). The *fru* gene encodes a group of putative transcription factors with the BTB and zinc-finger motifs, and the transcripts from the most distal promoter, P1, are subjected to sexually dimorphic splicing (Ito et al., 1996; Ryner et al., 1996), leading ultimately to the production of Fru protein in the male but not the female nervous system (Lee et al., 2000; Usui-Aoki et al., 2000). The Fru protein thus produced masculinizes the nervous system (Kimura et al., 2005). The Fru protein structure is well conserved, whereas its expression changes markedly across species, implying the *fru cis* elements play important roles in courtship behavior diversified among species (Usui-Aoki et al., 2005). To evaluate this possibility, we compared the putative regulatory regions of *fru* among 13 *Drosophila* species, including 12 species whose genomic sequences have been reported (*Drosophila* 12 Genomes Consortium, 2007) and *D. subobscura*, whose genomic sequence we determined 5' to the *fru* P1 promoter. In the course of this analysis, we identified, within the putative *fru* enhancer region, a highly conserved ncRNA sequence, *fru*-upstream ncRNA (*fuR*), which is transcribed in the direction opposite that of *fru*. Here we report the sequence comparisons of *fuR* across species and discuss the evolutionary implications of the results of their phylogenetic analysis.

## 2. Materials and Methods

### 2.1 Flies

The *D. melanogaster* Canton-S strain was a gift of R. Ueda (National Institute of Genetics, Mishima, Japan). The transgenic fly lines were generated according to the method described previously (Pfeiffer et al., 2008). *D. subobscura* was obtained at the San Diego Stock Center. The fly stocks were raised on standard cornmeal yeast medium supplemented with (*subobscura*) or without (*melanogaster*) soybean fiber (183g/l) at 18°C (*subobscura*) or 25°C (*melanogaster*) under a 12:12 light-dark cycle.

### 2.2 PCR primer design for amplification of genomic *fru* of *D. subobscura*

As no genomic sequence data were available for *D. subobscura*, we cloned and sequenced the putative regulatory region upstream from the most distal promoter of *fru* of this species. We amplified the genomic DNA of this region as seven consecutive segments that cover an entire 31 kb region by PCR. The primers for amplification were designed based on the sequences conserved across three species, *D. melanogaster*, *D. yakuba*, and *D. pseudoobscura*, whose genomic sequences were available (*Drosophila* 12 Genomes Consortium, 2007).

### 2.3 Genomic DNA preparation, PCR, and sequencing

Genomic DNA extracted from whole adult bodies of *D. subobscura* was prepared for PCR template according to the method described previously (Pfeiffer et al., 2008). For PCR amplification of the seven genomic segments, the Long Range PCR Kit (Qiagen) was used. In the subsequent cloning of the amplified sequence, the PrimeSTAR GXL enzyme (TaKaRa) or PrimeSTAR MAX (TaKaRa) enzyme was used. Sequencing was performed on the CEQ8000 Genetic Analysis System in conjunction with the GenomeLab DTCS-Quick Start Kit (Beckman Coulter) according to the manufacturer's instructions.

### 2.4 Reverse transcription polymerase chain reaction (RT-PCR)

For RT-PCR, total RNA was prepared from 20 whole adult Canton-S males and females, 3–5 days old, using an RNeasy Mini Kit (Qiagen). The RNA samples were reverse transcribed by the Superscript II reverse transcriptase (Invitrogen) and a primer, Forward: CCAACAGCTTTATTGCATTTATTTGG (forward) or TGTGAAACGCACTGTCGTCGAG (reverse) after DNase I treatment, and the resulting cDNAs were used as

templates for PCRs. For PCR, the Ex Taq polymerase (Takara) and the forward and reverse primers as described were used.

### 3. Results

In characterizing the genomic region upstream from the most distal promoter of *fru* in *D. subobscura*, we found, at about 13 kb upstream from the promoter, an extraordinarily conserved stretch of ~150 bp (i.e., *fuR*) shared by two other species, *D. melanogaster* and *D. pseudoobscura*, used for comparison (Figure 1), yet without any apparent open reading frame. Compared to the *melanogaster* sequence, the identities of the nucleotide sequences are 95% in *D. pseudoobscura* and 88% in *D. subobscura*, although the sequences are not conserved in the regions immediately distal and proximal to this ~150 bp stretch. An extended search of the genome database (Drosophila 12 Genomes Consortium, 2007) for the corresponding block of DNA in the *fru* upstream region reveals the homologous stretch in seven additional species of subgenus *Sophophora* (three aforementioned species belong to this subgenus), i.e., *D. simulans*, *D. sechelia*, *D. erecta*, *D. yakuba*, *D. persimilis*, *D. ananassae*, and *D. willistoni*, and three species of subgenus *Drosophila*, i.e., *D. mojavenensis*, *D. virilis*, and *D. grimshawi* (Figure 1). Among these 13 species, *D. willistoni* is particularly unique in that a 38 bp insertion splits the conserved block into two pieces (Figure 1). The three species of subgenus *Drosophila* carry some substitutions and insertions/deletions when compared to the sequences of *Sophophora* species (Figure 1). Two highly conserved segments are separated from each other by a low-homology region in which most of the above-mentioned variations exist (Figure 1).

The high level of conservation of this sequence block in the potential *fru* regulatory region might indicate that it serves as the binding sites for some transcription factors. Alternatively, this could be a transcribed sequence without an open reading frame. To determine whether or not this conserved sequence block bears any transcript, we carried out an RT-PCR analysis for total RNA prepared from adult whole bodies of *D. melanogaster* (Figure 2). The result demonstrates that this conserved DNA stretch yields a transcript, which is thus referred to as *fuR* (Figure 2). The cDNA template prepared with the forward primer, but not the reverse primer, yielded the product, indicating that *fuR* is transcribed from the strand opposite to that encoding the *fru* transcripts. No sex difference is detected in the abundance of *fuR* between the sexes.

To deduce how the *fuR* transcription unit has been maintained or modified in evolution, we constructed a phylogenetic tree as illustrated in Figure 3. The overall features of the tree for *fuR* match the phylogenetic relationships based on other molecular and morphological characteristics, implying that *fuR* has been maintained in different clades under similar functional constraints.

### 4. Discussion

We identified an ncRNA gene in the 5' upstream from the *fru* transcription start site in 13 species belonging to subgenera *Sophophora* and *Drosophila*. The distal limit of the *fru* regulatory region remains undefined. The *fru* mRNA expression pattern is altered in *fru*<sup>1</sup>, which has the mutagenic breakpoint of an inversion at about 10 kb upstream from the P1 promoter (Ryner et al., 1996). This means that the *fru* regulatory region likely spans across this point. Attempts to recapitulate the *fru* expression pattern with reporters driven by P1-promoter upstream regions have not been fully successful even using the genomic fragments of 30 kb or longer (K. Matsumoto, unpublished observation). This implies that the complete regulatory region of *fru* extends beyond this range. Therefore, *fuR* is judged to sit within the putative regulatory region of *fru*.

The recent discoveries of eRNAs transcribed from active enhancers of certain “neuronal” genes in mice (Kim et al., 2010) suggest that *fuR* might be analogous to eRNAs. However, known eRNAs are transcribed bi-directionally, unlike *fuR*, which is uni-directionally transcribed. Our preliminary experiment indicates that overexpression of *fuR* in the *fru*-expressing neurons by means of the GAL4-UAS system does not discernibly affect male courtship behavior (S. Takayanagi, unpublished observation). The overexpression of *fru* in the developing eye disc as driven by *GMR-GAL4* results in the distortion of eye morphogenesis (Goto et al., 2011). This *fru*-dominant phenotype was also unaffected by simultaneous overexpression of *fuR* (S. Takayanagi, unpublished observation). Further *in vivo* studies are required to elucidate the functions of *fuR*.

Sequence comparisons among 13 *Drosophila* species revealed that *fuR* is conserved in all species examined. The *fru* gene seems to play a central role in organizing the neural circuit for male courtship even in non-Drosophilid insects such as the cockroach and locust (Boerjan et al., 2011; Clynen et al., 2011). The ability of a *fru*-expressing motoneuron to induce a male-specific muscle is also conserved between *Drosophila* and the mosquito *Anopheles* (Gailey et al., 2006). These observations suggest that *fru* expression is rigidly regulated and exposed to strong selective pressure (Sobrinho & de Brito, 2010). It is tempting to speculate that *fuR* has a role in modulating *fru* functions for species-specific behavior, differential selection of which has shaped this ncRNA

sequence in species-specific patterns. Functional analysis of *fuR* in the genetic model *D. melanogaster* will provide insights into the evolutionary constraint involved in the diversification of this fascinating RNA sequence.

### Acknowledgments

We thank Manabu Ote for his participation in the early phase of this study, and Hiromi Sato for secretarial assistance.

### References

- Bastock, M. & Manning, A. (1955). The courtship of *Drosophila melanogaster*. *Behaviour*, 8, 85-111. <http://dx.doi.org/10.1163/156853955X00184>
- Boerjan, B., Tobback, J., De Loof, A., Schoofs, L. & Huybrechts, R. (2011). *fruitless* RNAi knockdown in males interferes with copulation success in *Schistocerca gregaria*. *Insect Biochemistry and Molecular Biology*, 41, 340-347. <http://dx.doi.org/10.1016/j.ibmb.2011.01.012>
- Clynen, E., Ciudad, L., Bellés, X. & Piulachs, M.-D. (2011). Conservation of *fruitless*'s role as master regulator of male courtship behaviour from cockroaches to flies. *Development Genes and Evolution*, 221, 43-48. <http://dx.doi.org/10.1007/s00427-011-0352-x>
- Dickson, B.J. (2008). Wired for sex: the neurobiology of *Drosophila* mating decisions. *Science*, 322, 904-909. <http://dx.doi.org/10.1126/science.1159276>
- Drosophila* 12 Genomes Consortium (2007). Evolution of genes and genomes on the *Drosophila* phylogeny. *Nature*, 450, 203-218.
- Gailey, D.A., Billeter, J.C., Liu, J.H., Bauzon, F., Allendorfer, J.B. & Goodwin, S.F. (2006). Functional conservation of the *fruitless* male sex-determination gene across 250 Myr of insect evolution. *Molecular Biology and Evolution*, 23, 633-643. <http://dx.doi.org/10.1093/molbev/msj070>
- Gleason, J.M. & Ritchie, M.G. (2004). Do quantitative trait loci (QTL) for a courtship song difference between *Drosophila simulans* and *D. sechellia* coincide with candidate genes and intraspecific QTL? *Genetics*, 166, 1303-1311. <http://dx.doi.org/10.1534/genetics.166.3.1303>
- Goto, J., Mikawa, Y., Koganezawa, M., Ito, H. & Yamamoto, D. (2011). Sexually dimorphic shaping of interneuron dendrites involves the Hunchback transcription factor. *Journal of Neuroscience*, 31, 5454-5459. <http://dx.doi.org/10.1523/JNEUROSCI.4861-10.2011>
- Ito, H., Fujitani, K., Usui, K., Shimizu-Nishikawa, K., Tanaka, S. & Yamamoto, D. (1996). Sexual orientation in *Drosophila* is altered by the satori mutation in the sex-determination gene *fruitless* that encodes a zinc finger protein with a BTB domain. *Proceedings of the National Academy of Sciences of the United States of America*, 93, 9687-9692. <http://dx.doi.org/10.1073/pnas.93.18.9687>
- Kharchenko, P.V., Alekseyenko, A.A., Schwartz, Y.B., et al. (2011). Comprehensive analysis of the chromatin landscape in *Drosophila melanogaster*. *Nature*, 471, 480-486. <http://dx.doi.org/10.1038/nature09725>
- Kim, T.-K., Hemberg, M., Gray, J.M., et al. (2010). Widespread transcription at neuronal activity-regulated enhancers. *Nature*, 465, 182-187. <http://dx.doi.org/10.1038/nature09033>
- Kimura K.-I., Ote, M., Tazawa, T. & Yamamoto, D. (2005). *Fruitless* specifies sexually dimorphic neural circuitry in the *Drosophila* brain. *Nature*, 438, 229-233. <http://dx.doi.org/10.1038/nature04229>
- Lee, G., Foss, M., Goodwin, S.F., Carlo, T., Taylor, B.J. & Hall, J.C. (2000). Spatial, temporal, and sexually dimorphic expression patterns of the *fruitless* gene in the *Drosophila* central nervous system. *Journal of Neurobiology*, 43, 404-426.
- Pfeiffer, B.D., Jenett, A., Hammonds, A.S., et al. (2008). Tools for neuroanatomy and neurogenetics in *Drosophila*. *Proceedings of the National Academy of Sciences of the United States of America*, 105, 9715-9720. <http://dx.doi.org/10.1073/pnas.0803697105>
- Ryner, L.C., Goodwin, S.F., Castrillon, D.H., et al. (1996). Control of male sexual behavior and sexual orientation in *Drosophila* by the *fruitless* gene. *Cell*, 87, 1079-1089. [http://dx.doi.org/10.1016/S0092-8674\(00\)81802-4](http://dx.doi.org/10.1016/S0092-8674(00)81802-4)
- Rzhetsky, A. & Nei, M. (1992). A simple method for estimating and testing minimum evolution trees. *Molecular Biology and Evolution*, 9, 945-967.
- Sayed, D. & Abdellatif, M. (2011). MicroRNAs in development and disease. *Physiological Reviews*, 91, 827-887. <http://dx.doi.org/10.1152/physrev.00006.2010>

- Sobrinho, I.S.J. & de Brito, R.A. (2010). Evidence for positive selection in the gene *fruitless* in *Anastrepha* fruit flies. *BMC Evolutionary Biology*, 10, 293-308. <http://dx.doi.org/10.1186/1471-2148-10-293>
- Spieth, H.T. (1952). Mating behavior within the genus *Drosophila* (Diptera). *Bulletin of the American Museum of Natural History*, 99, 399-474.
- Straub, T. & Becker, P.B. (2007). Dosage compensation: the beginning and end of generalization. *Nature Reviews Genetics*, 8, 47-57. <http://dx.doi.org/10.1038/nrg2013>
- Tamura, K., Nei, M., & Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences of the United States of America*, 101, 11030-11035. <http://dx.doi.org/10.1073/pnas.0404206101>
- The modENCODE Consortium (2010). Identification of functional elements and regulatory circuits by *Drosophila* modENCODE. *Science*, 330, 1787-1797.
- Thompson, J.D., Higgins, D.G., & Gibson, T.J (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 22, 4673-4680. <http://dx.doi.org/10.1093/nar/22.22.4673>
- Usui-Aoki, K., Ito, H., Ui-Tei, K., *et al.* (2000). Formation of the male-specific muscle in female *Drosophila* by ectopic *fruitless* expression. *Nature Cell Biology*, 2, 500-506. <http://dx.doi.org/10.1038/35019537>
- Usui-Aoki, K., Mikawa, Y. & Yamamoto, D. (2005). Species-specific patterns of sexual dimorphism in the expression of Fruitless protein, a neural masculinizing factor in *Drosophila*. *Journal of Neurogenetics*, 19, 109-121. <http://dx.doi.org/10.1080/01677060591007191>
- Yamamoto, D. (2008). Brain sex differences and function of the *fruitless* gene in *Drosophila*. *Journal of Neurogenetics*, 22, 309-321. <http://dx.doi.org/10.1080/01677060802298491>
- Yamamoto, D., Jallon, J.-M. & Komatsu, A. (1997). Genetic dissection of sexual behavior in *Drosophila melanogaster*. *Annual Review of Entomology*, 42, 551-585. <http://dx.doi.org/10.1146/annurev.ento.42.1.551>

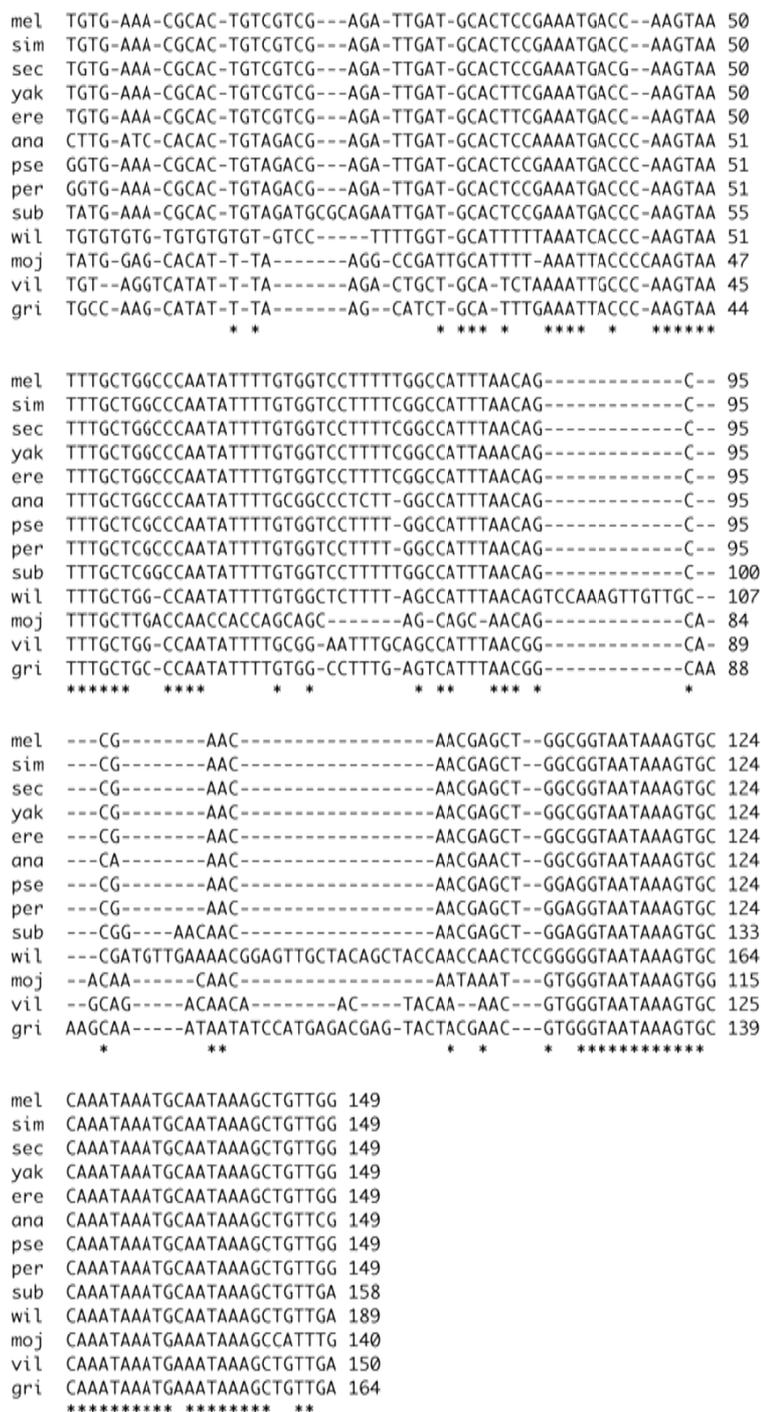


Figure 1. Aligned nucleotide sequences of the ~150 bp conserved stretch (*fuR*) in 13 *Drosophila* species. Nucleotide sequences are shown for *D. melanogaster* (mel), *D. simulans* (sim), *D. sechelia* (sec), *D. yakuba* (yak), *D. erecta* (ere), *D. ananassae* (ana), *D. pseudoobscura* (pse), *D. persimilis* (per), *D. subobscura* (sub), *D. willistoni* (wil), *D. mojavenensis* (moj), *D. virilis* (vir), and *D. grimshawi* (gri). The forward strands with respect to the *fuR* transcription (i.e., the reverse strands with respect to the *fru* transcription) are shown. The sequences were aligned by the Clustal W program (Thompson et al., 1994). The identical nucleotides are marked with asterisks at the bottom.



Figure 2. *fuR* transcript in *D. melanogaster*

Total RNAs were prepared from the whole bodies of either females or males (indicated above the lane). The cDNA template prepared with the forward (For, indicated above the lane) but not reverse (Rev) primer yielded the ~150 bp RT-PCR product.

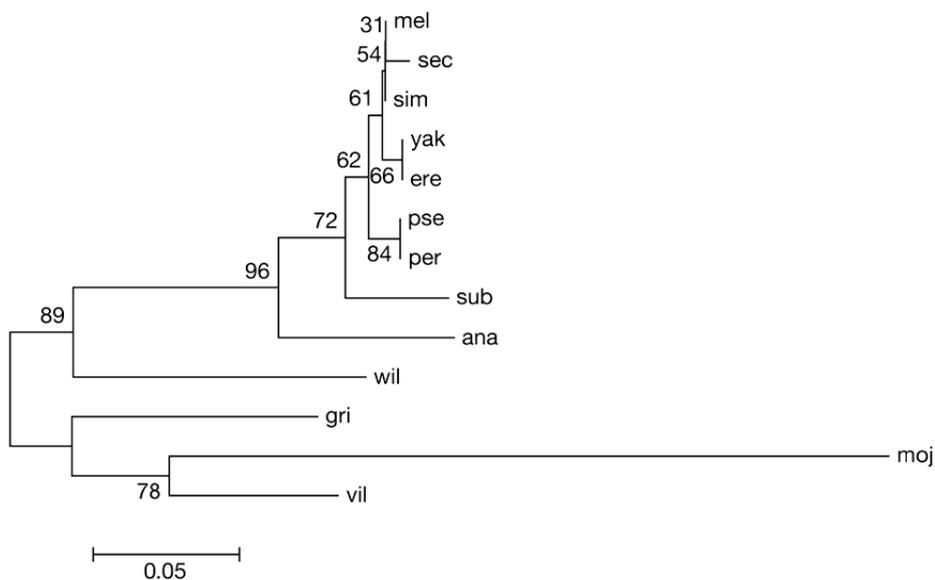


Figure 3. A phylogenetic tree of *fuR*

The evolutionary distances were computed by the maximum composite likelihood method (Tamura et al., 2004), and the tree was constructed by the minimum evolution method (Rzhetsky and Nei, 1992). The branch and scale bar lengths represent the number of nucleotide substitutions per site. The percentage in which a particular branch was supported in the bootstrap test of 1000 replicates is shown next to the branch.