available at www.sciencedirect.com

ScienceDirect

journal homepage: http://www.elsevier.com/locate/ejsobi

Original article

Construction of a Bacterial Artificial Chromosome (BAC) library and the genomic analysis of valosine-containing proteins in the earthworm *Eisenia fetida*

Kohei Suzuki[a,b], *Tomoharu Suzuki*[a], *Stephen Stürzenbaum*[c], *Shinobu Gamou*[a,*]

[a]*Department of Environment and Life Sciences, Kyorin University School of Health Sciences, Miyashita 476, Hachioji, Tokyo 192-8508, Japan*
[b]*Research Division II, STAFF Institute, Kamiyokoba, Tsukuba, Ibaraki 305-0854, Japan*
[c]*Department of Biochemistry, School of Biomedical and Health Sciences, King's College London, 150 Stamford Street, London SE1 9NH, UK*

ARTICLE INFO

Article history:
Published online 27 December 2007

Keywords:
Earthworm
BAC library
Genome structure
VCP
Spermatogenesis

ABSTRACT

In an attempt to isolate gonad-related genes from the earthworm *Eisenia fetida* (Savigny, 1826), we have isolated two genes with sequence similarities to mammalian valosine-containing proteins (*evcp*-1 and *evcp*-2) [T. Suzuki, M. Honda, S. Matsumoto, S.R. Stürzenbaum, S. Gamou. Valosine-containing proteins (VCP) in an annelid: identification of a novel spermatogenesis related factor, Gene 362 (2005) 11–18]. To understand the origin of the earthworm *vcp* genes, and to provide tools for further detailed genetic and physiological analyses, a Bacterial Artificial Chromosome (BAC) library was established, and the clones carrying the entire *evcp*-1 and *evcp*-2 genes were isolated. Genomic sequencing revealed that *evcp*-1 and *evcp*-2 occupied approximately 16 kb and 18 kb, respectively, and consisted of 14 introns and 15 exons, a number that is similar to mammalian VCPs.

1. Introduction

Over seven decades have passed since earthworms were first used to study spermatogenesis [6]. This historic interest stems from the fact that earthworms are hermaphrodites with testis and ovaries located in separate segments. Although the morphogenesis of reproductive organs is remarkably stable in the lumbricid earthworm used in this study, and the morphology of reproductive organs has been described in exquisite detail, very little is known concerning the molecular genetic background.

In an attempt to isolate gonad-related genes from the earthworm *Eisenia fetida* (Savigny, 1826), we have identified two members of AAA family (ATPase-Associated with diverse cellular Activity) with sequence similarities to mammalian valosine-containing proteins (VCP) (*evcp*-1: AB181225 and *evcp*-2: AB181226) [9]. Results from reverse transcription-polymerase chain reaction (RT-PCR) and in situ hybridization analyses suggested that one of these, *evcp*-2, is a candidate gene likely to be involved in the late stages of spermatogenesis.

Members of the VCP/P97/Cdc48 family are ubiquitously distributed in organisms ranging from prokaryotic organisms to mammals, highlighting the evolutionary conservation [1,10]. To understand the origin and evolution of the earthworm *vcp* genes and also to provide tools for further detailed genetic and physiological analyses, a Bacterial Artificial Chromosome (BAC) library carrying large genomic fragments was

* *Corresponding author*. Fax: +81 426 91 1094.
E-mail address: gamou@kyorin-u.ac.jp (S. Gamou).
1164-5563/$ – see front matter © 2007 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.ejsobi.2007.12.002

established. Using the three steps PCR screening system, genomic fragments carrying full length clones of *evcp-1* and *evcp-2* were isolated.

2. Materials and methods

2.1. BAC library

The BAC library was constructed as described before with minor modifications [3,8]. In brief, seminal vesicles were isolated from several mature *E. fetida*, and trypsinized. The dissociated cells were embedded into agarose, digested by proteinase K, partially digested with HindIII, and the resultant fragments ligated into the pCC1BAC vector (Epicenture, WI, USA). The average insert size of genomic fragments was approximately 100 kb. More than 50,000 clones were individually picked into 530 × 96-well plates. The BAC library is estimated to contain more than five billion bps, which is more than several-fold coverage of the genome [4].

2.2. Three step PCR screening of evcp genomic fragments

The PCR primer pairs were selected from 5′ and 3′ regions of each gene based on the estimation of the intron–exon junctions as described by Suzuki et al. [9]. The BAC library was divided into 44 super-pools, each containing 12 × 96-well plates. This first PCR screen identified the positive super-pool, which was split into respective 12 plate-pools (second screen). The *evcp* containing clone was identified from the positive plate in a third screen comprising a column-pool (A–H) and row-pool (1–12).

2.3. Genomic sequencing and analysis

The BAC clones were used to amplify 2–3 kb fragments of genomic *evcp* DNA with primers designed on the previously identified cDNA sequences [9] using Primer 3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi). The sequence reaction was carried out using the DYE terminator sequencing Kit (GE-Healthcare, NJ, USA) and the ABI PRIZM 310 genetic analyzer (Applied Biosystems, Tokyo, Japan). Phylogenic and homology analyses were performed using the following on-line resources: http://www.ddbj.nig.ac.jp/Welcome-j.html and http://pipmaker.bx.psu.edu/pipmaker.

3. Results

3.1. Isolation of BAC clones

Utilizing the three step screening protocol genomic fragments containing *evcp-1* and *evcp-2* were identified from over 50,000 BAC clones. The positive isolates were clones 20E8 and 91F3 for *evcp-1* and 307G11 for *evcp-2* gene. The pulse field gel electrophoresis revealed that the insert sizes were approximately 120 kb, 115 kb and 120 kb, for 20E8, 91F3 and 307G11, respectively (Fig. 1). Further PCR analysis using cDNA based-primers including those for 5′ and 3′ terminal region confirmed these clones contained the entire genomic regions of the *evcps*. This

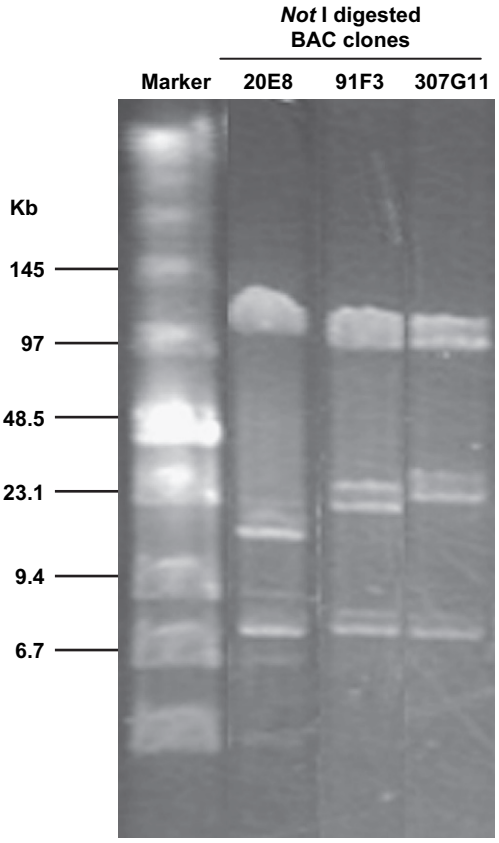

Fig. 1 – Pulse field gel electrophoresis of BAC clones. DNAs were isolated from each clone and subjected to digestion by NotI. Pulse field gel electrophoresis was carried out using the BioRad CHEF Mapper.

analysis highlighted that both *evcp-1* and *evcp-2* genes are over 10 kb in length and therefore are different to their orthologues in *Caenorhabditis elegans* (C41C4.8 and C06A1.1), which comprise a total length of 2.9 kb and 3.4 kb, respectively [11].

3.2. Genomic sequence

Genomic sequencing revealed *evcp-1* gene is approximately 16 kb long (AB284052) and consists of 14 introns and 15 exons (Fig. 2). The intron–exon boundaries are similar to mammalian VCPs including the human vcp (*hvcp*: NC_000009). *evcp-2* is approximately 18 kb in size (AB284053) and also contains 14 introns and 15 exons (Fig. 2). It is noteworthy that GAA/TTC tri-nucleotide repeats were found in the introns of *evcp-2*, indicating the possible occurrence of micro-satellites.

3.3. Intron–exon structure

The intron–exon junctions are summarized in Fig. 2. Most introns comply with the AG/gt -cag/rule. The homology analysis using the PipMaker revealed that a sequence of *evcp-1* exon 7 was found to be homologous to three exons in *hvcp*, indicative of a possible exon duplication event, and similarly some introns reverse-duplicated (Fig. 3).