**Drosophila P transposons of the urochordata Ciona intestinalis**

Stefanie Kimbacher · Ingrid Gerstl · Branko Velimirov · Sylvia Hagemann

Received: 29 January 2009 / Accepted: 21 April 2009
© Springer-Verlag 2009

**Abstract**

P transposons belong to the eukaryotic DNA transposons, which are transposed by a cut and paste mechanism using a P-element-coded transposase. They have been detected in *Drosophila*, and reside as single copies and stable homologous sequences in many vertebrate species. We present the P elements *Pcin1*, *Pcin2* and *Pcin3* from *Ciona intestinalis*, a species of the most primitive chordates, and compare them with those from *Ciona savignyi*. They showed typical DNA transposon structures, namely terminal inverted repeats and target site duplications. The coding region of *Pcin1* consisted of 13 small exons that could be translated into a P-transposon-homologous protein. *C. intestinalis* and *C. savignyi* displayed nearly the same phenotype. However, their P elements were highly divergent and the assumed P transposase from *C. intestinalis* was more closely related to the transposase from *Drosophila melanogaster* than to the transposase of *C. savignyi*. The present study showed that P elements with typical features of transposable DNA elements may be found already at the base of the chordate lineage.

**Keywords**
P transposon · *Ciona* · *Drosophila* · THAP9 · Urochordata · Evolution

**Introduction**

P elements were first discovered in *Drosophila melanogaster* and belong to the DNA transposons (Kidwell et al. 1977). The genome of a typical P strain of *D. melanogaster* contains 50–60 P element copies dependent on geographical location (Kidwell 1983; Anxolabéhère et al. 1985a). In addition to complete autonomous elements, there are a number of internally deleted and thus non-autonomous elements (O’Hare et al. 1992).

Full-sized autonomous *D. melanogaster* P elements have a length of 2907 bp (O’Hare and Rubin 1983). The four exons of the element code for the transposase, an 87-kD protein that catalyzes transposition in the germ line (Rio et al. 1986; Laski et al. 1986). For the transposition event, 31-bp terminal inverted repeats (TIRs) are essential and 11-bp subterminal inverted repeats (STIRs) function as transpositional enhancers. P element transposition occurs exclusively in the germ line because that is only where the transposase is produced. The mechanism behind this is the tissue-specific, alternative splicing of intron 3 in the pre-mRNA. This intron is only spliced in germ line cells, which leads to assembly of the 87-kD transposase that is encoded by all four exons. In somatic cells (and in the germ line), the third intron is not spliced, which gives rise to the 66-kD repressor protein (Laski et al. 1986; Misra and Rio 1990).

P element sequences can be found in almost all species groups of the subgenus *Sophophora* (Anxolabéhère et al. 1985b; Lansman et al. 1985; Daniels and Strausbaugh 1986), but they are absent in species most closely related to *D. melanogaster* (Brookfield et al. 1984) and in geographically isolated species. These observations, together with the fact that P elements are completely absent from old laboratory strains of *D. melanogaster*, indicate that the latter has obtained its P element by horizontal transmission from
another species. This transfer must have occurred within the last century, and was followed by a rapid spread of $P$ elements through the new host (Anxolabéhère et al. 1988). $P$-homologous sequences have also been found outside the genus Drosophila, for example, in the Australian sheep blowfly Lucilia cuprina (Perkins and Howells 1992), the house fly Musca domestica (Lee et al. 1999) and in Scaptonyza pallida (Simonelig and Anxolabéhère 1991; Hagemann et al. 1996; Haring et al. 1998).

$P$ elements had been thought to occur exclusively in the genomes of dipteran insects until our group described a $P$-homologous sequence in vertebrates (Hagemann and Pinsker 2001). Primarily, it was found by BLAST search using the cDNA sequence of a 3′-truncated $P$ element from Drosophila subsilvestris in the human genome, and it was named Phsa ($P$ homolog of Homo sapiens). Phsa corresponds to a sequence designated as THAP9 that was described 2 years later by Roussigne et al. (2003a). The THAP domain was originally identified as a DNA-binding domain in the human protein THAPI, which is a nuclear pro-apoptotic factor (Roussigne et al. 2003b). Phsa has a length of 19 533 bp and consists of six exons. It is a stationary, single-copy gene, located on the long arm of chromosome 4 and encodes a hypothetical protein of 903 amino acids, with still unknown function (Hammer et al. 2005).

Pgga is the single-copy, stationary and presumably coding $P$-homologous sequence in Gallus gallus, in which it is located at an orthologous position to Phsa. Such stationary, single-copy $P$-homologous sequences have been detected in several other vertebrate species, but only as rudiments in rodents (Hammer et al. 2005). These rudiments are also located at orthologous positions to Phsa and Pgga. In contrast, several $P$-homologous sequences have been found in the genome of zebrafish, named Pdre ($P$ homolog of Danio rerio). Some of these have structural features typical of transposable elements and large open reading frames (ORFs) that suggest coding capacity (Hammer et al. 2005; Hagemann and Hammer 2006).

To improve our knowledge about the evolution of $P$ elements within the vertebrates, we proceeded down the phylogenetic lineage to the base of the chordate ancestry and chose Ciona species for our further studies, from which THAP9-like sequences have been reported (Quesneville et al. 2005). Ciona species belong to the so-called urochordata and are located phylogenetically at the base of the chordate lineage. Therefore, they constitute a suitable model system for studying chordate genome organization and evolution. Their free-swimming, tadpole-like larvae show typical chordate features, such as a prominent notochord, a dorsal hollow nerve chord and branchial segments. The present study focused on Ciona intestinalis, an ascidian species that is widely distributed throughout coastal regions around the world (Hoshino and Nishikawa 1985; Dehal et al. 2002). Its $P$ elements were compared with those from Ciona savignyi, which is distributed in the Pacific Ocean and has recently spread from Japan to the West coast of the US. These two species are often described as being closely related (Jiang and Smith 2005). Morphologically, they are quite similar and they also share essential mechanisms of early development (Johnson et al. 2004). Systematically, C. intestinalis and C. savignyi are members of the class Ascidiaeae or sea squirts and belong to the subphylum Tunicata or Urochordata. Urochordata is part of the phylum Chordata, together with Cephalochordata and Vertebrata, and diverged from the last common ancestor of all chordates at least 520 million years ago (Johnson et al. 2004). Recent studies, based on phylogenetic analyses of genome sequences, have shown that the Tunicata are the closest living relatives of the Vertebrata (Delsuc et al. 2006).

In this study, we investigated $P$ transposons from Ciona sp. and showed that they exhibited typical features of DNA-transposons, such as TIRs and TSDs. The finding of $P$ transposons in two species of Ciona provides compelling evidence that they existed already as jumping genes at the base of vertebrate evolution, and that their stable integration into the genome in higher vertebrates resulted from a molecular domestication event during evolution (Miller et al. 1992).

Materials and methods

All laboratory techniques were carried out using standard methods.

In silico analysis

The in silico analysis before and after the laboratory experiments were carried out using BLAST (http://www.ncbi.nlm.nih.gov) and BLAT (http://genome.ucsc.edu/) database searches (Kent 2002), and MegAlign (DNASTAR LaserGene) to create alignments. The $P$ elements from C. savignyi were found in the Repbase Update (Jurka et al. 2005), which is the most commonly used database of repetitive DNA elements. Phylogenetic analysis was carried out with the PhyML online web server (Guindon et al. 2005; http://atgc.lirmm.fr/phyml/) and MrBayes (Ronquist and Huelsenbeck 2003; http://mrbayes.csit.fsu.edu/). The combination of in silico and molecular biological studies was carried out because the TIR-containing ends of the C. intestinalis $P$-elements could not be detected by in silico analysis alone.

Hybridization methods: library screening and Southern blotting

The genomic C. intestinalis library was screened with a Digoxigenin-labeled, PCR-amplified and cloned fragment
of 1,011 bp (Pcin+: 5'-CAAATACCTGCTGCTAATGC
AACC-3' and Pcin−: 5'-TAACCAGACTGAAACACAT
CTGGG-3') and genomic DNA from C. intestinalis as a template. Positive plaques were detected by the enzyme-induced color reaction using BCIP (5-bromo-4-chloro-3-indolyl phosphate; Roche Diagnostics; Mannheim, Germany) and NBT (Nitro blue tetrazolium chloride; Roche Diagnostics) as substrates. For the isolation of phage DNA, we used the Lambda Mini Kit from Qiagen (Germany). Southern blot hybridizations were carried out re-using the hybridization solution from the library screening.

Cloning, subcloning and sequencing

The PCR-amplified hybridization probe was cloned using the pGEM®-Teasy Vector System (Promega, USA). SmaI-digested, phosphatase-treated pGEM®-3Z DNA (Promega, USA) was used for subcloning blunt-end fragments eluted from agarose gels by the GFX PCR DNA and Gel Band Purification Kit from GE Healthcare (USA). The (recombinant) plasmids were transformed into competent Escherichia coli JM109 cells (Promega, USA). Sequencing was carried out in our laboratory using the SequiTherm Excel II DNA Sequencing Kit (Epicentre Biotechnologies, USA) and an LI-COR 4200, or by 4base lab (Germany).

Results

P-homologous sequences in C. intestinalis

With the BLAST search using the P-homologous amino acid sequence deduced from Pdre2 of zebrafish (BX890548) and the program tblastn (Table 1), the most significant hit was a short cDNA sequence of 700 bp from C. intestinalis (BW487041). Using the translated sequence, P transposase similarity was reflected by its homology to different THAP9 proteins (e-values up to 3e-41), which corresponded to the P-homologous proteins in vertebrates. We identified the corresponding genomic C. intestinalis sequence (AABS01002034) and the cDNA full insert sequence clone of 1,673 bp (AK113587). A BLAT search with this cDNA sequence led to multiple hits that allowed us to assume that repeated P-homologous sequences exist within the genome of C. intestinalis, but we could not determine their terminal sequences. Therefore, at this point, the laboratory experiments were started to obtain P element sequences not represented in the available genomic sequence data.

Library screening and characterization of obtained clones

Based on the alignment of the genomic sequence (AABS01002034) and the cDNA sequence (AK113587), we designed the primer pair Pcin+ and Pcin−, which bound to exonic regions and amplified a 1,011-bp genomic probe. Screening a genomic library, we received more than 30 hybridization signals and took six clones (λ7, λ11, λ12, λ13, λ17 and λ19), which were chosen randomly for further analysis. The clones were double restricted with ScaI and PvuII and the fragments electrophoresed, blotted and hybridized, to identify those with P-homologous sequences. After hybridization, all clones exhibited one fragment of ~3,500 bp (fragment 1) and another of ~1,700 bp (fragment 2), although four of the six analyzed clones originated from different genomic positions (Supplementary S1). We concluded that the two fragments, which were subcloned and partially sequenced, were parts of a larger repeat unit of ~5,200 bp.

In silico analyses based on the obtained sequence data

The regions outside the already sequenced parts had to be analyzed, in order to identify the presumed ends of the repeat unit, which were thought to display the P element termini. To this end, we combined fragments 1 and 2 from clone λ13 (Fig. 1), which resulted in a 3,148-bp screening probe (Supplementary S2). The correct orientation of the two fragments was given by the homology between them and the 1,011-bp hybridization probe, which overlapped the 3' region of fragment 1 and the 5' region of fragment 2 (Fig. 1). This artificial in silico screening probe included

Table 1: Most significant results of BLAST search using the amino acid sequence deduced from Pdre2

<table>
<thead>
<tr>
<th>Acc. no.</th>
<th>Description of the sequences</th>
<th>e-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BW487041</td>
<td>Nori Satoh unpublished cDNA library, C. intestinalis, mRNA sequence</td>
<td>6e-44</td>
</tr>
<tr>
<td>DNS52037</td>
<td>1316550 MARC 7BOV Bos taurus cDNA 3’, mRNA sequence</td>
<td>8e-38</td>
</tr>
<tr>
<td>BW502377</td>
<td>Nori Satoh unpublished cDNA library, C. intestinalis, mRNA sequence</td>
<td>8e-38</td>
</tr>
<tr>
<td>CR551045</td>
<td>Bos Taurus, embryonic and extraembryonic tissues</td>
<td>7e-37</td>
</tr>
<tr>
<td>1262874</td>
<td>MARC 7BOV Bos taurus cDNA 5-, mRNA sequence</td>
<td>2e-36</td>
</tr>
<tr>
<td>1257998</td>
<td>MARC 7BOV Bos taurus cDNA 3-, mRNA sequence</td>
<td>3e-36</td>
</tr>
</tbody>
</table>
stretches of non-determined nucleotides and spanned the 5', as well as the 3' end of the internal 5,200-bp fragment of the assumed *C. intestinalis* *P* element. The result of the BLAT search is shown in Table 2. Three sequences, namely those located on the scaffolds 202 and 272 and on chromosome 01q were considered as especially interesting for further analysis, because their homology with the 3,148-bp screening probe started at position 1 and ended at position 3,148. This confirmed our assumption that the two subcloned fragments were parts of a longer repeat unit. The spanned regions had a length of 5,134 bp (on scaffold 202), 4,401 bp (on scaffold 272), and 1,393 bp (on chromosome 01q). Thus, the sequence located on scaffold 202 was supposed to be the most complete.

The *P* elements of *C. intestinalis*

The nucleotides located upstream and downstream of the sequences located on scaffolds 202 and 272 and chromosome 01q were aligned in order to analyze the borders of the repetitive sequence, and hence the *P* element terminal sequences. It was possible to determine three of the *C. intestinalis* *P* elements (Table 3), which were designated as *Pcin1* (located on scaffold 202; Supplementary S3),

### Table 2 Result of BLAT search using an artificial probe of 3148 bp, deduced from sequence data of clone A13

<table>
<thead>
<tr>
<th>Score</th>
<th>Start</th>
<th>End</th>
<th>Identity (%)</th>
<th>Location</th>
<th>Strand</th>
<th>Start</th>
<th>End</th>
<th>Span</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,937</td>
<td>1</td>
<td>3,148</td>
<td>99.1</td>
<td>202</td>
<td>-</td>
<td>65,901</td>
<td>71,034</td>
<td>5,134</td>
</tr>
<tr>
<td>1,453</td>
<td>1</td>
<td>2,369</td>
<td>93.3</td>
<td>59</td>
<td>+</td>
<td>114,569</td>
<td>116,511</td>
<td>1,943</td>
</tr>
<tr>
<td>1,350</td>
<td>847</td>
<td>3,148</td>
<td>98.4</td>
<td>138</td>
<td>-</td>
<td>118,386</td>
<td>122,613</td>
<td>4,228</td>
</tr>
<tr>
<td>1,263</td>
<td>1</td>
<td>2,972</td>
<td>96.8</td>
<td>181</td>
<td>-</td>
<td>121,117</td>
<td>122,977</td>
<td>1,861</td>
</tr>
<tr>
<td>1,126</td>
<td>868</td>
<td>3,147</td>
<td>94.2</td>
<td>275</td>
<td>-</td>
<td>20,179</td>
<td>23,056</td>
<td>2,878</td>
</tr>
<tr>
<td>1,121</td>
<td>1</td>
<td>3,148</td>
<td>97.9</td>
<td>272</td>
<td>-</td>
<td>46,173</td>
<td>50,573</td>
<td>4,401</td>
</tr>
<tr>
<td>1,112</td>
<td>1</td>
<td>2,365</td>
<td>98.2</td>
<td>297</td>
<td>+</td>
<td>12,169</td>
<td>16,400</td>
<td>4,232</td>
</tr>
<tr>
<td>1,090</td>
<td>1</td>
<td>3,148</td>
<td>97.0</td>
<td>01q</td>
<td>+</td>
<td>1,359,987</td>
<td>1,361,379</td>
<td>1,393</td>
</tr>
</tbody>
</table>

### Fig. 1 The strategy for in silico searching for the terminal sequences of C. *intestinalis* *P* elements was as follows. Fragment 1 (F1/3500 bp) and fragment 2 (F2/1,700 bp), which were identified as two subunits of a larger repeat unit, were partially sequenced (black regions) and joined together to construct an artificial in silico screening probe of 3,148 bp (Supplementary S2), which was used for BLAT searching. The homologous regions of three hits (scaffolds 202 and 272, and chromosome 01q; Table 2) started with the first nucleotide (1) and ended with the last nucleotide (3,148) of the screening probe. These sequences were used to determine the ends of the repetitive sequence by comparing their flanking regions. This led to the detection of the TIR-containing termini of the *C. intestinalis* *P* elements (shown as black-framed white arrows), designated as *Pcin1*, *Pcin2* and *Pcin3*.
**Table 3** Most important features of Pcin1, Pcin2 and Pcin3

<table>
<thead>
<tr>
<th>Position (nt)</th>
<th>Strand</th>
<th>Length (bp)</th>
<th>TIR (bp)</th>
<th>TSD (bp)</th>
<th>Coding region</th>
<th>Special features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pcin1 202(^a)</td>
<td>Minus</td>
<td>5,614</td>
<td>29</td>
<td>None</td>
<td>+</td>
<td>Duplication of 157 bp at the 5' end</td>
</tr>
<tr>
<td>Pcin2 272(^a)</td>
<td>Minus</td>
<td>4,877</td>
<td>29</td>
<td>7</td>
<td>?</td>
<td>Duplication of 153 bp at the 5' end</td>
</tr>
<tr>
<td>Pcin3 01q(^b)</td>
<td>Plus</td>
<td>1,716</td>
<td>29</td>
<td>7</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Scaffold
\(^b\) Chromosome

**Table 4** Genomic excision footprints

<table>
<thead>
<tr>
<th>Chromosome</th>
<th>Position (nt)</th>
<th>Strand</th>
<th>Sequence (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>05q</td>
<td>4,084,754–4,084,825 Plus</td>
<td>gatcagCATAGTACTCTCTTAagagaagagactagtgtactctcTAAGAGAGTACTATGgatcag</td>
<td></td>
</tr>
<tr>
<td>01p</td>
<td>3,837,258–3,837,299 Plus</td>
<td>ctagagCATAGTACTCTCATAAGAGAGTACTATGctagag</td>
<td></td>
</tr>
</tbody>
</table>

TSDs lower case, underlined, TIRs upper case, bold

**Pcin2** (located on scaffold 272; Supplementary S4) and **Pcin3** (located on chromosome 01q; Supplementary S5) All three elements have 29-bp long TIRs; **Pcin2** and **Pcin3** are flanked by TSDs of 7 bp, whereas, in the sequence 202 (**Pcin1**), the TSD was missing (Table 3). **Pcin1** had a length of 5,614 bp and its 29-bp TIRs were conserved. At the 5’ end, a perfect duplication of 157 bp, including the 29-bp TIRs, was found, which corresponded to a 5’ duplication in **Pcin2**. Comparison with the cDNA showed that the coding region of **Pcin1** consisted of 13 small exons and 12 introns. All introns had intact splicing sites. The coding sequence was flanked by 5’ and 3’ non-coding sequences. **Pcin2** had a length of 4,877 bp with perfectly conserved 29-bp TIRs and a 7-bp TSD. As mentioned above, at the 5’ end, 153 bp were duplicated, with one substitution at position 152. This duplication included the 7-bp TSD and the 29-bp TIR. The internal region of this sequence was not determined in the scaffold sequence, thus its coding capacity was questionable. **Pcin3** was an internally deleted element of 1,716 bp with conserved 29-bp TIRs and a 7-bp TSD. The 5’ breakpoint corresponded with position 620 in **Pcin1** and the 3’ breakpoint with position 4355 in **Pcin1**. Hence, they were located within the first intron and the 3’ non-coding sequence. All three analyzed C. intestinalis P elements showed characteristic features of potentially active DNA transposons, namely TIRs (in **Pcin1**, **Pcin2** and **Pcin3**) and TSDs (in **Pcin2** and **Pcin3**). Neither the in silico constructed transcript from **Pcin1** nor the cDNA obtained from the database (AK113587) had an intact reading frame. Therefore, we suppose that either the coding master sequence was not identified, or that the **Pcin** elements were already silenced. The first hypothesis was supported by genomic excision footprints that were perfectly conserved and flanked by perfectly conserved target-site duplications (Table 4). **Pcin1**, **Pcin2** and **Pcin3**, as well as the two transcripts of C. intesti-

**Discussion**

The P elements from C. intestinalis are highly divergent from those of C. savignyi

Three P-homologous sequences, **P1_cis**, **P6_cis** and **P1a_cis**, from C. savignyi, are found in the Repbase Update (Jurka et al. 2005). One of these, namely **P1_cis** (Supplementary S6) can be translated into a P-homologous protein. The second element, **P1a_cis**, has multiple internal deletions with respect to **P1_cis**. **P6_cis** is depicted as a P element because of a short region (324 bp) with encoded similarity to P transposases, although this is the only indication of any relationship with P elements. All three elements were posted by Smit in the Repbase Update in 2005 (Jurka et al. 2005). **P1_cis** has a length of 4,062 bp, exhibits 26-bp TIRs (with one substitution), and is flanked by an 8-bp TSD. It has two large overlapping ORFs. ORF1 encodes a protein that contains a THAP4 domain, which is a member of the THAP superfamily. The larger ORF2 probably encodes a P-transposase homolog, **P1a_cis** is 2,614 bp long and its 26-bp TIRs are perfectly conserved.

Comparative analysis of **P1_cis** and the **Pcin** elements has shown that the P elements of C. savignyi and C. intestinalis are highly divergent. The TIRs of the C. savignyi elements **P1_cis** and **P1a_cis** are 26 bp long, whereas all three C. intestinalis elements had 29-bp TIRs and only the first (considering the left TIR) five nucleotides (CATAG…CTATG) were identical between the TIRs of the two species. The TSD in C. savignyi is 8 bp long, which is the typical length for P-transposon TSDs. The 7-bp TSDs that flank the C. intestinalis
elements $P_{cin2}$ and $P_{cin3}$ were unusual for $P$ transposons. Considering the coding regions, $P_{cin1}$ depicted 13 small exons with sizes between 34 and 299 bp (Supplementary S3), whereas there are two large ORFs in $P_{1_cis}$ (Supplementary S6), and only one of them codes for a putative transposase. Although the translated sequences of $P_{1_cis}$ and $P_{cin1}$ gave significant hits to $P$ transposases of other species (vertebrates and drosophilids) in a BLAST search, they showed a high divergence rate of 24.3% within the region used for the phylogenetic trees shown in Fig. 2. Highly divergent amino acid sequences have been reported also for the huntingtin gene product of $C. intestinalis$ and $C. savignyi$: the 27.5% divergence is comparable to that observed between mammals and fishes (Gissi et al. 2006). Analyses of 18S tRNA sequences show that the pairwise divergence of the two $Ciona$ sp. is slightly less than that between human and frog, and slightly greater than between human and chicken (Johnson et al. 2004). It is supposed that this is caused by the presence of long branches for these species (Gissi et al. 2006; Supplementary S7).

*Relationship between $P$ elements from $Ciona$ sp. and other species groups*

When analyzing the relationship between $P$ elements, it is always tempting to look at their TIRs and reasonable to compare their deduced amino acid sequences. The TIRs are essential features for $P$-element transposition, and they are bound by the inverted repeat binding proteins encoded by the organism itself (Rio and Rubin 1988). The 29-bp TIRs of the three $P$ elements from $C. intestinalis$ ($P_{cin1}, P_{cin2}$ and $P_{cin3}$) were conserved perfectly, whereas the 26-bp TIR from the $C. savignyi$ $P$ element was conserved perfectly in $P_{1a_cis}$, but had one substitution in $P_{1_cis}$. There was only a weak sequence homology of ~50% between the TIRs of $P_{cin}, P_{1a_cis}$ and the canonical $P$ element from *D. melanogaster*. Only the first three nucleotides of the left TIR and last three nucleotides of the right TIR were highly conserved, which emphasized their importance for the transposition event.

For comparison at the amino acid level, the canonical $P$ transposase ($P_{mel}$) from *D. melanogaster* was aligned to the $P$ element amino acid sequences from *C. intestinalis* ($P_{cin1}$) and *C. savignyi* ($P_{1a_cis}$), and to the putative proteins from zebrafish ($P_{dre}$), chicken ($Pgga$) and human ($Phsa$) (Supplementary S8). The amino acid alignment from position 269–594 (Fig. 3) was used to calculate phylogenetic trees using the online web server PhyML and MrBayes. The trees obtained (Fig. 3) did not reflect the phylogenetic relationships and placed the $P$ element of *C. savignyi* ($P_{cis}$) adjacent to that of zebrafish ($P_{dre}$; weak bootstrap value of 64 in the left tree) and the $P$ element of *C. intestinalis* adjacent to that of *Drosophila* ($P_{mel}$; bootstrap value of 99 in the right tree).

![Fig. 2 The phylogenetic trees show the relationship of $P$ elements and $P$-homologous sequences from several organisms. The left tree was calculated by the online web server PhyML, the right one with the program Mrbayes. They were based on the protein alignment (aminoacids 269–594) shown in Fig. 3. The following sequences were used: $P_{cin1}$ from *C. intestinalis* ($P_{cin}$); $P_{1_cis}$ from *C. savignyi* ($P_{cis}$); $P_{pe25.1}$ from *D. melanogaster* ($P_{mel}$); $P_{dre2}$ from *Danio rerio* ($P_{dre}$); $Pgga$ from chicken; and $Phsa$ from human ($Phsa$). The trees do not reflect the phylogenetic relationships, and place the $P$ element of *C. savignyi* ($P_{cis}$) adjacent to that of zebrafish ($P_{dre}$; weak bootstrap value of 64 in the left tree) and the $P$ element of *C. intestinalis* adjacent to that of *Drosophila* ($P_{mel}$; bootstrap value of 99 in the right tree)](image-url)
transmitted horizontally. Further investigations will tell us whether horizontal transmission can also be considered within the chordate lineage as a mechanism for P element propagation.

Conclusion

The present study demonstrates that P elements with typical features of mobile DNA elements have been found already at the base of the chordate lineage. It can be speculated that P elements were present in a common ancestor of the urochordata and vertebrates, or considering the well-known Drosophila P elements, even within the common ancestor of protostomia and deuterostomia.

Acknowledgments

The genomic C. intestinalis library was kindly provided by Weiyang Shi (Berkley). We thank Daniela Schneider for various discussions, and the anonymous reviewers for valuable suggestions to improve the manuscript.

References


Hammer SE, Strehl S, Hagemann S (2005) Homologs of Drosophila P transposons were mobile in zebrasch but have been domesticated in a common ancestor of chicken and human. Mol Biol Evol 22:833–844

