Species sympatry and horizontal transfers of Mariner transposons in marine crustacean genomes

N Casse, Q T Bui, V Nicolas, S Renault, Y Bigot, M Laulier

Available online at www.sciencedirect.com
SCIENCE DIRECT
Molecular Phylogenetics and Evolution 40 (2006) 609–619
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Species sympatry and horizontal transfers of *Mariner* transposons in marine crustacean genomes

N. Casse [a], Q.T. Bui [a], V. Nicolas [c], S. Renault [b], Y. Bigot [b,*], M. Laulier [a]

[a] *Laboratoire de Biologie et Génétique Evolutive, EA 3265, Université du Maine, Avenue Olivier Messiaen, 72085 Le Mans cedex, France*
[b] *Laboratoire d'Etude des Parasites Génétiques, EA 3868, Université François Rabelais, U.F.R. des Sciences et Techniques, Avenue Monge, 37200 Tours, France*
[c] *Museum National d'Histoire Naturelle, Département de Systèmatique et Evolution, USM 601, UMR 5202, Laboratoire Mammifères et Oiseaux, 55 rue Buffon, 75005 Paris, France*

Received 19 September 2005; revised 16 January 2006; accepted 6 February 2006
Available online 11 May 2006

Abstract

Mariner-like elements (*MLEs*) have been widely detected in terrestrial species. The first complete *MLE* isolated from a marine invertebrate was detected in the genome of the hydrothermal crab *Bythograea thermydron* by Halaimia-Toumi et al. [Halaimia-Toumi, N., Casse, N., Demattei, M.V., Renault, S., Pradier, E., Bigot, Y., Laulier, M., 2004. The GC-rich transposon *Bytmar1* from the deep-sea hydrothermal crab, *Bythograea thermydron*, may encode three transposase isoforms from a single ORF. J. Mol. Evol. 59, 747–760] and called *Bytmar1*. Here, we report the isolation of three new *Bytmar1* relatives from the genomes of one hydrothermal amphipod *Ventiella sulfuris* (*Vensmar1*) and two coastal crustacea, *Maia brachydactila* (*Maibmar1*) and *Cancer pagurus* (*Canpmar1*). Like *Bytmar1*, these *MLEs* have an unusually high GC content, a high CpG ratio, and a low TpA ratio. Their consensus sequence encodes a transposase that is preceded by an N-flag, as in *Bytmar1*, which could be a marine feature. Only one of the 19 clones obtained, *Vensmar1.3*, encoded for a full-length transposase. The phylogenetic analyses revealed that all these *Bytmar1*-related elements can be differentiated into two clusters, corresponding to the coastal or hydrothermal origin of their hosts. They also confirmed that the *irritans* sub-family comprises at least four lineages that seem to depend on the taxonomical position and habitat of their hosts. Finally, we observed that elements coding for two potentially complete transposases exhibiting 99.5% similarity, *Bytmar1.11* and *Vensmar1.3*, were present in the genome of two distantly related hydrothermal crustacea, one Amphipod and one Decapod. The hypothesis of horizontal transfers is discussed in the light of the sequence similarities observed.

Keywords: Transposable element; *Mariner*; Crustacea

1. Introduction

Transposable elements (TEs)[1] are discrete DNA fragments that are able to move within prokaryotic and eukaryotic genomes. They are currently viewed as being genetic parasites that can damage the host genome by creating mutations. However, they can also be beneficial, because their mobility provides a source of genetic variability for the host genome (Kidwell and Lisch, 2001). TEs are divided into two main classes depending on their structure and transposition mechanism. Class I elements move via an RNA copy that is replicated as an RNA copy by a reverse transcriptase before being re-inserted at a new target site in the host genome. Class II elements, known as transposons, move via a DNA intermediate, and encode the enzyme, or transposase, that allows the transposition to occur.

Several families of Class II transposons have been identified and they are commonly thought to be universal genetic

* Corresponding author. Fax: +33 02 47 36 70 35.
E-mail address: bigot@univ-tours.fr (Y. Bigot).

[1] *Abbreviations used:* TEs, transposable elements; ITR, inverted terminal repeat; *MLE*, *mariner*-like element, ORF, open reading frame; SDS, sodium dodecyl sulfate; SKS, sarkosyl sulfate.

1055-7903/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2006.02.005

parasites. *Mariner* elements, and *mariner*-like elements (*MLEs*), are transposons that occur in eukaryotic genomes, and belong to a superfamily, known as *ITm* (as *IS630-Tc1-mariner*), which includes six main families: *Tc1*, *mariner*, *ITmD37D* or *maT*, *ITmD41D*, and the "plant *MLE*" (Claudianos et al., 2002; Feschotte et al., 2005). To date, the *mariner* family has been shown to consist of five sub-families known as *cecropia*, *elegans/briggsae*, *irritans*, *mauritiana*, and *mellifera/capitata* (Bigot et al., 2005). The recent discovery of *Tvmar1* in the genome of *Trichomonas vaginalis* has suggested the possible existence of a sixth sub-family (Silva et al., 2004).

MLEs consist of a 1200–2000 base pair (bp) DNA fragment that contains a single gene without introns, and which encodes for the transposase. This gene is flanked by two short inverted terminal repeats (ITRs) 19–40 bp in length (Halaimia-Toumi et al., 2004), except in the case of *Mcmar1*, where the ITRs are 355 bp long (Leroy et al., 2003). Their ubiquity in the genomes of continental organisms has been well documented (for review, see Robertson, 2002). In contrast, there has been little attempt to search for *MLEs* in aquatic organisms, and they have only been identified in the genomes of two marine flatworms, *Stylochus zebra* and *Bdelloura candida* (Robertson, 1997), of the Cnidera *Hydra littoralis*, the jellyfish *Aurelia aurita*, and of a few Rotifera species (Arkhipova and Meselson, 2000). Over the past 5 years, our work has helped to enlarge the spectrum of marine species investigated, and *MLEs* have been detected in the genomes of the Norway lobster, *Nephrops norvegicus*, and in several deep-sea invertebrates, such as the mussel *Bathymodiolus thermophilus*, the crab *Bythograea thermydron*, the shrimp *Rimicaris exoculata*, and the gutless worm *Riftia pachyptila* (Casse et al., 2000, 2002). Comparison of the *Bytmar1* elements found in the genome of *B. thermydron* to those found in the genome of continental species (Halaimia-Toumi et al., 2004) has revealed several differences in their nucleic acid sequence, and in the traits of their transposase. Indeed, the GC content of the *Bytmar1* elements is significantly higher ($54.35 \pm 1.05\%$) than that in continental *MLEs* ($40.5 \pm 3.45\%$). This difference is mainly due to codon usage in the transposase ORF and could theoretically interfere with some DNA properties, such as the curvature propensities of the nucleic acid sequence, with possible consequences for *Bytmar1* transposition efficiency. It is also noteworthy that some *Bytmar1* elements have an ORF able to encode for three transposase isoforms characterized by having N-terminal tags of different sizes.

The unusual properties of *Bytmar1* raised questions about the selection constraints that led to the evolution to retain them as valuable solutions, and the presence of *Bytmar1*-related elements was further investigated in the genomes of several crustacean species occurring in coastal and deep-sea environments. Here, we describe the presence of elements very closely related to *Bytmar1* in an Amphipod living near hydrothermal vents, *Ventiella sulfuris*, and of more distant *Bytmar1* relatives in two Decapods living in coastal waters, the sea shoe *Maia brachydactyla*, and *Cancer pagurus*. Our findings confirm that some of the characteristics of *Bytmar1*-related elements can be linked to their presence in the genome of marine species and this raises questions about the possibility of horizontal transfers occurring between deep-sea crustacean species.

2. Materials and methods

2.1. Biological materials

The coastal crustacea studied here comprise two Decapod Brachyra: *C. pagurus* and *M. brachydactyla* collected in the English Channel (off the north coast of Brittany). The hydrothermal crustacean *V. sulfuris* (Amphipoda, Lysianassidae) was collected during the Hope 99 cruise from a depth of 2600 m during submersible dives (Nautile) from the East Pacific Ridge at 103°56′43W, 12°48′67N and 103°56′81W, 12°49′84N.

2.2. Nucleic acid analyses

For each specimen, total genomic DNA (gDNA) was purified as described by Laulier et al. (1995). An additional purification step was performed by ultracentrifugating the gDNA on a cesium chloride gradient (Sambrook et al., 1989).

2.3. Southern blot/hybridization

For the Southern blot, the purified gDNA was digested by restriction enzymes (*Apa*I/*Pvu*II and *Pst*I/*Eco*RI for *M. brachydactyla*; *Apa*I/*Pvu*II and *Dde*I/*Pst*I for *C. pagurus*; *Pst*I/*Eco*RI and *Dde*I for *V. sulfuris*), run on a 0.8% agarose gel and then blotted on Nylon membranes (Hybond-N, Amersham) according to Southern (1975). Membranes were hybridized overnight at 65 °C in 0.5 M Na_2HPO_4–NaH_2PO_4 pH 7.4, 7% SDS, with a probe labeled with the random priming method using Nonaprimer Kit (Appligene) and [α^{32}P]dCTP (Amersham). The membranes were then washed at 65 °C for 30 min in 2× SSC–0.1% SKS, for 30 min in 1× SSC–0.1% SKS, and finally for 30 min in 0.2× SSC–0.1% SKS (1× SSC = 0.15 M NaCl, 0.015 M tribasic sodium citrate). X-ray films were exposed overnight on the membranes with an intensifying screen at −80 °C. The probes used were *Maibmar1.4* (AJ507241) and *Vensmar1.3* (AJ507234) cloned in the plasmid pGEM-T easy vector system I (Promega). Plasmid preparations were purified using Wizard plus a miniprep Kit (Promega). The *MLE* fragments were isolated from the plasmid by double digestion, using *Apa*I/*Pst*1 and then eluted from a 1% agarose gel using Wizard SV Gel and PCR Clean-up System (Promega).

2.3.1. MLE amplification

Full-length *MLEs* were amplified using the first 25 base pairs of the 5′ ITR obtained from *B. thermydron* by Halaimia-Toumi et al. (2004): 5′-TACGAGGGGCGGTCAGA

AAGTTATG-3′. The reactions were performed with 100 ng of gDNA, placed in a total volume of 50 µl, using 0.1 µM of the primer, 150 µM of each dNTP, 1 U of *Taq* polymerase (Promega), and 10× *Taq* buffer. Amplifications were performed in an Eppendorf Mastercycler personal. The PCR conditions were as follows: denaturing at 94 °C for 2 min followed by 35 cycles of three steps: denaturing at 94 °C for 1 min, annealing for 1 min at 55 °C, extension at 72 °C for 1 min, and a final extension at 72 °C for 5 min.

2.4. Cloning and sequencing

The PCR products approximately 1300 bp in length were eluted from the agarose gel using Wizard SV Gel and PCR Clean-up System (Promega), and cloned in the plasmid pGEM-T easy vector system I (Promega). Plasmids were purified using Wizard plus miniprep Kit (Promega) and sequenced by Genome Express (http://www.gexbyweb.com/).

The sequences reported here have been deposited in the DDBJ/EMBL/GenBank sequence database under Accession Nos. AJ507232–AJ507237 (*V. sulfuris*); AJ507238–AJ507244 (*M. brachydactyla*); AJ507245–AJ507250 (*C. pagurus*).

2.5. Sequence analyses

Infobiogen facilities were used for Databank searches, sequence alignments, and calculations. The alignment deposited in the EMBL database under Accession No. DS36877 (Augé-Gouillou et al., 2000) was updated with recently reported sequences of *MLE* transposases (Sumitani et al., 2002; Krieger and Ross, 2003; Lampe et al., 2003; Leroy et al., 2003; Mandrioli, 2003; Robertson and Walden, 2003). Phylogenetic analyses of the aligned sequences were developed using the Parsimony and Neighbor-Joining programs in the PHYLIP package, version 3.5c (Felsenstein, 1993). Sequence similarities and distance matrices were investigated using Edtaln (Multiple alignment editor) at http://www.infobiogen.fr/deambulum/liste.php?page=pga_alnn.

The GC content of the nucleic acid sequences corresponding to each full-length *MLE*, to each ORF encoding an *MLE* Tnp, and to those at positions 1, 2, and 3 (GC1, GC2, and GC3) of the codons used in the *MLE* ORF were calculated at the Web site http://www.kazusa.or.jp/codon/countcodon.html. GC contents were calculated at the Web site http://www.infobiogen.fr/services/analyseq/cgi-bin/freqbnk_in.pl. The XpX dinucleotide frequencies were checked by calculating the observed XpX frequency/expected XpX frequency (XpX obs/exp) ratio at the Web site http://www.infobiogen.fr/services/analyseq/cgi-bin/freqsq_in.pl. For example, the expected TpA frequency in each investigated sequence was calculated from the T and A contents. Thus, the TpA frequency could be expected to be 21% × 15% = 3.1% for a DNA with a T content of 21% and an A content of 15%. As recommended by Karlin and Burge (1995), the XpX obs/exp ratio was considered to be under-represented if it was below 0.78 and over-represented if higher than 1.23. The analyses of differences in the GC contents or XpX frequency between *MLE* groups were performed using the non-parametric Wilcoxon/Kruskal–Wallis test at 0.05 or 0.01 significance, after confirming the normality of each sample with a Shapiro–Wilk test at 0.05 significance.

3. Results

Our searches for elements closely related to *Bytmar1* in crustacean genomes were carried out by PCR, using as primer the outer sequence of the *Bytmar1* ITR and gDNA samples purified from 11 species living in marine coastal environments (*Atelecyclus undecidentatus*, *C. pagurus*, *Galathea strigosa*, *M. brachydactyla*, *N. norvegicus*, *Necora puber*, *Pachygrapsus marmoratus*, *Palaemon serratus*, *Pisa tetraodon*, *Pilumnus hirtellus*, *Pinnotheres pisum*) and from seven species living in deep-sea hydrothermals (*B. thermydron*, *Chorocaris chacei*, *Cyanagraea preadator*, *Mirocaris fortunata, R. exoculata*, *Segonzacia mesatlantica*, *V. sulfuris*) (Fig. 1). Using this procedure, we were able to amplify, clone, and sequence *Bytmar1*-related elements from eight new species: *A. undecidentatus*, *C. pagurus*, *C. preadator*, *M. ortunata*, *M. brachydactyla*, *P. marmoratus*, *R. exoculata*, and *V. sulfuris*. Here we only report the data obtained with *MLEs* isolated from *C. pagurus*, *M. brachydactyla*, and *V. sulfuris*. The quantity and/or quality of the few individual gDNA samples available in our laboratories for the other five species were not sufficient to confirm the presence of these transposons in the chromosomal DNA of these species by Southern blot hybridization and washing at high stringency. Further investigations using new individual gDNA samples will therefore be required to confirm the presence of *Bytmar1*-related elements in these species.

Each type of *Bytmar1*-related element described in this study was named using a nomenclature similar to that proposed by Robertson and Asplund (1996). We modified this system slightly, whereas Robertson and Asplund had used only a single letter, we used the first three letters of the genus name. For example, the first complete sequence detected in *C. pagurus* was named *Canpmar1.1* for *mariner* 1 clone 1, *Maibmar1.1* in *M. brachydactyla*, and *Vensmar1.1* in *V. sulfuris*, respectively.

3.1. Nucleic sequence analysis

Nineteen distinct sequences of *Bytmar1*-related elements were cloned and sequenced from the PCR products obtained from *C. pagurus*, *M. brachydactyla*, and *V. sulfuris* (Table 1). In *C. pagurus*, six *MLEs* (*Canpmar1.1* to *Canpmar1.6*; Accession Nos. AJ507245–AJ507250) were characterized. Their lengths were comprised between 1296 and 1353 bp and their nucleic acid sequences had an average similarity level of 94.08% (±3.66). Seven *MLEs* (*Maibmar1.1* to *Maibmar1.7*; Accession Nos. AJ507238–AJ507244) were identified in *M. brachydactyla*. Their

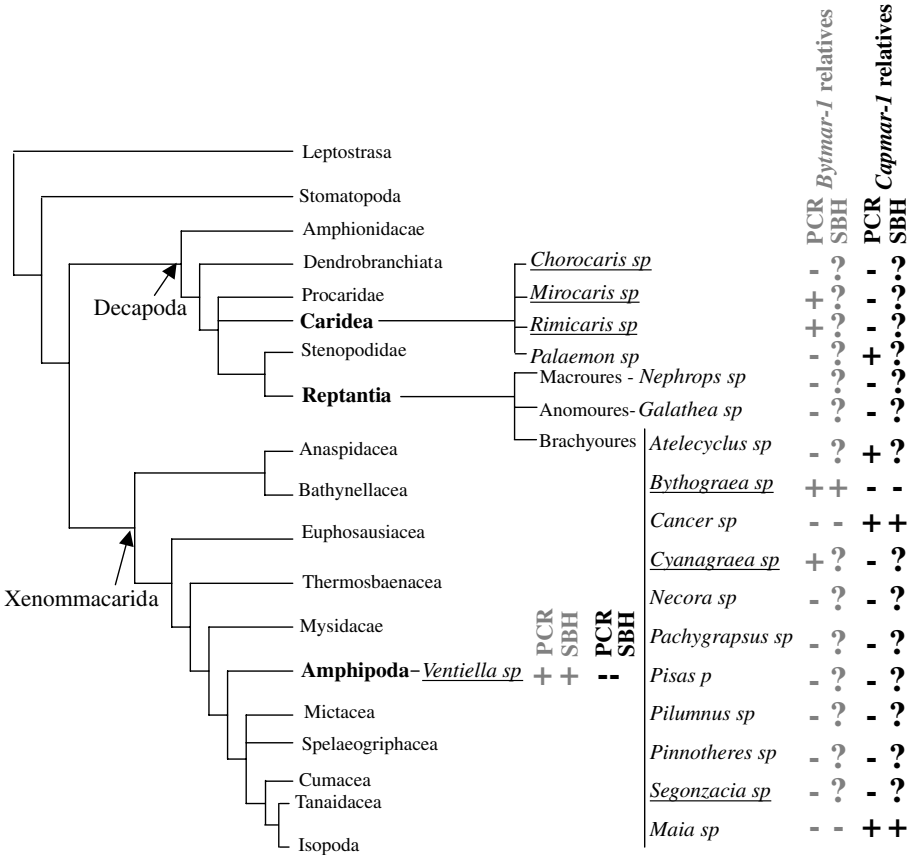

Fig. 1. Evolutionary relationships between *Malacostraca crustacean* species. Based on data at the Web site: http://sn2000.taxonomy.nl/Taxonomicon/TaxonTree.aspx?id = 33722&syn = 1. Underlined names correspond to deep-sea hydrothermal species. Tables located beside *Venturia* sp. and in the right margin describe the presence (+) or the absence (−) of detection by PCR and Southern blot/hybridization (SBH) of close relatives of hydrothermal *Bytmar1* (data typed in gray) and coastal *Bytmar1* (data typed in black). Question mark indicates missing data.

Table 1
Average identity and standard deviation calculated from pairwise comparisons of nucleic acid sequences in *C. pagurus*, *M. brachydactyla*, *V. sulfuris*, and *B. thermydron*

	C. pagurus	*M. brachydactyla*	V. sulfuris	B. thermydron
C. pagurus	94.08 ± 3.66	92.62 ± 2.92	67.24 ± 1.63	66.84 ± 2.33
M. brachydactyla		95.20 ± 2.11	69.48 ± 0.71	69.2 ± 2.12
V. sulfuris			90.45 ± 3.30	89.83 ± 3.68
B. thermydron				89.22 ± 4.13

Results in % of homology.

sequences were 1282–1303 bp long and had an average nucleic similarity level of 95.20% (±2.11). In *V. sulfuris*, six elements were obtained (*Vensmar1.1* to *Vensmar1.6*, Accession Nos. AJ507232–AJ507237). Their sequences ranged from 1288 to 1302 bp in length and they had an average nucleic acid sequence similarity level of 90.45% (±3.30).

In an attempt to determine the characteristics of the ITR of the inner part and the 5′ and 3′ UTRs (untranslated terminal regions), consensus sequences were calculated for each type of *Bytmar1*-related element isolated from *C. pagurus*, *M. brachydactyla*, and *V. sulfuris*. The ITR lengths were 31 bp in *Canpmar1*, 28 bp in *Maibmar1*, and 31 bp in *Vensmar1*. In the *Vensmar1* element, degeneracy at position 24 in our data set made it impossible to define the nature of this nucleotide in the consensus ITR. The 5′ UTR was 145 bp long in *C. pagurus* and *M. brachydactyla*, and 147 bp long in *V. sulfuris*. The 3′ UTR was 32 bp long in all three species. The ITRs and UTRs found in these elements had a cardinal *MLE* length.

The nucleotide properties of each type of *Bytmar1*-related element were investigated using all the sequences obtained. The average GC content values found were 50.95% (±0.53) in *Canpmar1*, 50.92% (±0.47) in *Maibmar1*, and 53.64% (±0.88) in *Vensmar1*. Statistical analyses of the GC content differences between the three types of *Bytmar1*-related elements showed that those found for *Vensmar1* were significantly higher than those in *Maibmar1* or *Canpmar1* (non-parametric Wilcoxon/Kruskal–Wallis test with a confidence limit of 95%). Similarly, we found that the GC content of *Vensmar1* was similar to that of *Bytmar1* (54.55 ± 0.87%; Halaimia-Toumi et al., 2004). Finally, we

observed that GC contents of the *Bytmar1* relatives isolated from marine coastal or deep-sea hydrothermal species were both significantly different from those of the *MLEs* occurring in continental species (40.5 ± 3.45%); average value established from 38 *MLEs* isolated from genomes of continental species (Table 1 in Halaimia-Toumi et al., 2004).

The GC content of the transposase ORF and of the third codon position were then calculated in an attempt to find out whether they were due to a codon bias, as previously observed for *Bytmar1* (Halaimia-Toumi et al., 2004). The percentages for *Vensmar1* ORF (54.03 ± 0.77) and *Bytmar1* ORF (55.13 ± 1.18) were significantly higher than those calculated for *Canpmar1* ORF (51.46 ± 0.57) and *Maibmar1* ORF (51.38 ± 0.54). A GC content of 54–55% therefore seems to be characteristic of a hydrothermal environment, whereas in the coastal milieu this value is about 51.5%. We observed that the nucleotide composition at the third position of each codon in the transposase ORF was 58.41% (±1.84) in *Maibmar1*, 58.32% (±0.44) in *Canpmar1*, 66.92% (±1.55) in *Vensmar1*, and 66.86% (±1.77) in *Bytmar1*, and this value was significantly higher in deep-sea hydrothermal *MLEs* than in the coastal *MLEs*. Moreover, the GC contents of all these *MLEs* were also significantly higher than those of the continental *MLEs*. In line with our previous observations (Halaimia-Toumi et al., 2004), no significant differences were found between the *MLE* categories at positions 1 and 2.

The comparison of the CpG obs/exp ratio (0.71–0.81) revealed that there were no significant differences between the four types of *Bytmar1*-related element isolated from marine invertebrates. In agreement with the findings for *Bytmar1*, these data also confirmed that this dinucleotide was significantly under-represented in the nucleic acid sequence of all *Bytmar1*-related elements. Comparison of the TpA obs/exp ratio with those of the continental *MLEs* (0.61 ± 0.13) showed that this dinucleotide was also under-represented in the nucleic acid sequence of all *Bytmar1*-related elements: 0.43 ± 0.03 in *Vensmar1*, 0.55 ± 0.03 in *Maibmar1*, 0.56 ± 0.03 in *Canpmar1* (confidence limit 99%), and 0.45 ± 0.05 for *Bytmar1* (Halaimia-Toumi et al., 2004). Moreover, they revealed that these rates were significantly lower in *Bytmar1* and *Vensmar1* isolated from hydrothermal species than in those isolated from the coastal hosts *Maibmar1* and *Canpmar1*. TpA shortages have been reported to be a general feature of transposons (Karlin and Burge, 1995). What is striking in our data is that these shortages seem to be accentuated in *MLEs* occurring in marine coastal and hydrothermal deep-sea crustacean species.

Databank searches for *ITm* elements in the genomes of marine invertebrates located a 796-bp fragment containing a fossil of transposase ORF with a GC content of 41% in the genome of the crustacean copepod *Lepeophtheirus salmonis* (AJ783358, positions 30–826). Although this element did not belong to the *mariner* family, but to the *ITmD37D* family, its nucleotide properties suggested that the GC enrichment of a nucleotide sequence is not an absolute rule for the Class II transposons in the genomes of marine invertebrates. However, in the absence of further data, the retrotransposons with a GC-rich sequences previously characterized in the genome of *B. thermydron* (Halaimia-Toumi et al., 2004) and our data for *Bytmar1* relatives suggest that the low GC content of this *ITmD37D* might be considered to be an unusual property in the genomes of the marine invertebrates or even a property specific to this transposon family.

3.2. ORF features

Consensus amino acid sequences of the transposase were determined for each type of *Bytmar1*-related element (Fig. 2). The similarity percentages were calculated between the distinct sequences and were found to be 92% for the hydrothermal transposases *Vensmar1* and *Bytmar1*, and 91% for the coastal transposases *Canpmar1*/*Maibmar1*. The similarity between the coastal and hydrothermal transposases was 75%.

In the hydrothermal *MLE*, *Bytmar1*, part of the ORF encoding the transposase encodes N-terminal "flags" of 49 and 30 amino acids, respectively, that putatively gave rise to two longer proteins (Halaimia-Toumi et al., 2004). An N-terminal 48 amino acid extension, similar to that of *Bytmar1*, was found in the consensus *Vensmar1* transposase (Fig. 2). The flags found in *Bytmar1*-related elements isolated from coastal host genomes were very different. Consensus N-terminal extensions of 30 and 29 amino acids were identified in the ORFs of *Maibmar1* and *Canpmar1*, respectively. The sequence of these flags was very similar in *Maibmar1* and *Canpmar1*, with the exception of a serine that was inserted at position 3 in the *Maibmar1* flag, but absent from the *Canpmar1* flag. It is noteworthy that a stop codon was present in the N-terminal flag in the 13 sequenced *Maibmar1* and *Canpmar1* ORFs. So far, therefore, it looks as if a putative N-terminal flag could be a distinctive characteristic of *Bytmar1*-related transposases that was also recently found and unclassified *MLE*, *Tvmar1* (Silva et al., 2004).

The amino acid sequences of the consensuses of the *Maibmar1*, *Canpmar1*, and *Vensmar1* transposases were further investigated to identify conserved motifs and structures (Fig. 2). A search for cardinal structures did not reveal HTH (helix–turn–helix), but two putative monopartite NLSs (nuclear localization signals) were found. The first motif (K(R/Q)RGKL) begins at positions 267, 268, and 286, and the second motif (H/RKAR(R/V)A) at positions 286, 287, and 304 in *Canpmar1*, *Maibmar1*, and *Vensmar1*, respectively. Unexpectedly, the putative NLSs found in these three transposases were all located in the C-terminal domain, whereas it had previously been proposed that these motifs were located in the N-terminal domain in the *Mos-1* and *Himar1* proteins (Plasterk and Van Luenen, 2002).

The WVPHEL signature (Robertson, 1993) was slightly modified here to WVPRNL in all three proteins, whereas the YSPDLAP motif (Robertson, 1993) was well conserved. In the C-terminal domain, the FLHDNARPH motif usually found in *MLE* transposases was replaced by LLHDNTPVR in *Vensmar1* and by LLHDNA[P/S]VH in *Canpmar1* and *Maibmar1*. The catalytic triad (D,D34D) was located at positions (187, 278, 312) in *Canpmar1*, (188,

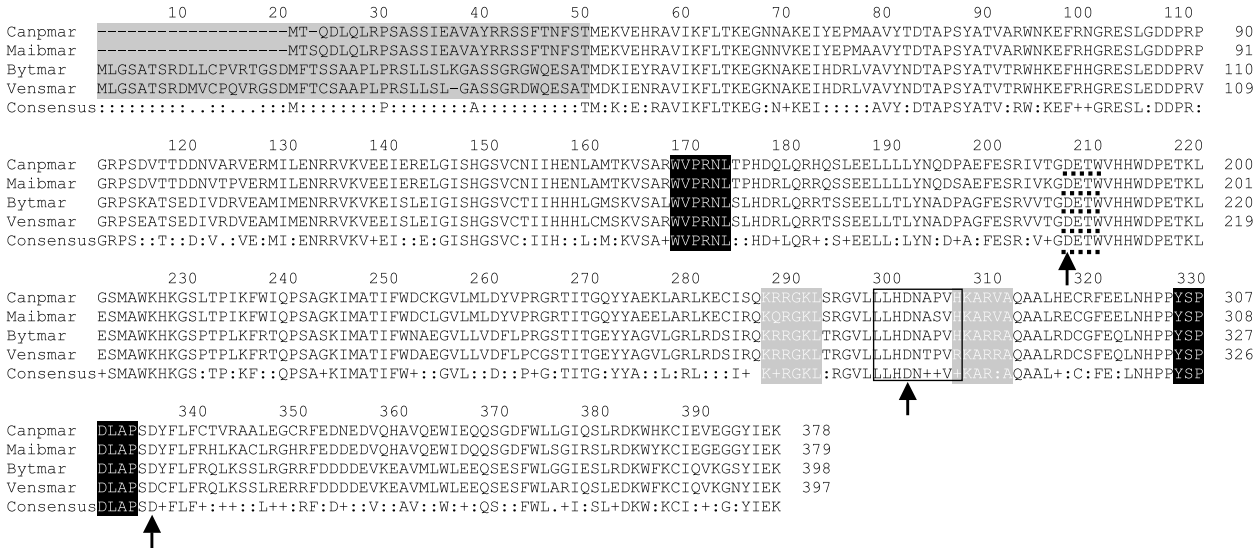

Fig. 2. Amino acid sequence alignment of transposases encoded by the ORF of the consensus element *Maibmar1*, *Canpmar1*, and *Vensmar1* (present study) and *Bytmar1* (Halaimia-Toumi et al., 2004). Consensus sequences were established using the majority rule. The conserved motifs have been located. An arrow indicated each aspartic acid in the D,D34D catalytic domain. The WVPRNL and YSPDLAP motifs are black boxed. The DETW motif is dash underlined. The putative nuclear localization signals (K(R/Q)RGKL) and (H/RKA(R/V)RA) are written in white and gray boxed. The positions of the amino acid nucleic sequences are numbered in normal type in the right margin.

279, 313) in *Maibmar1*, and (206, 297, 331) in *Vensmar1*. The first D was contained in a well-conserved DETW motif that is specific to the transposases of the *irritans* sub-family, other *MLE* transposases being characterized by DEKW motifs (Shao and Tu, 2001), apart from *Tvmar1*, which displays a DESW motif (Claudianos et al., 2002).

Direct conceptual translations of the 19 sequences showed that only one of them, *Vensmar1.3*, has an uninterrupted ORF and encodes a putative transposase of 347 amino acids. The features described above have been observed in this sequence, but no N-flag has been detected. It was striking that we also observed that it exhibited 99.5% similarity to that encoded by *Bytmar1.11*.

At the transcriptional level, none of the nucleic acid sequences obtained exhibited appropriate CAAT and/or TATA boxes in the 5′ UTR. However, they all contained a polyadenylation signal that overlapped the 3′ UTR and 3′ ITR, rather than the stop codon of the transposase ORF, as previously reported in most of the *MLEs* so far described (Robertson and Lampe, 1995; Robertson and Martos, 1997; Robertson and Zumpano, 1997; Lampe et al., 2003).

3.3. Distribution in the host genomes

The distribution of *Canpmar1*, *Maibmar1*, and *Vensmar1* in the gDNA of each host species was verified by Southern blot hybridization (Fig. 3). The high similarity between *Canpmar1* and *Maibmar1* sequences makes it possible to use the same probe, *Maibmar1-4*, for both *M. brachydactyla* and *C. pagurus* (Figs. 3a and b). In *V. sulfuris*, the probe used was the clone *Vensmar1.3* (Fig. 3c).

In each species, the consensus sequence was used to determine two kinds of restriction enzyme digests, except in the case of *V. sulfuris*, for which only limited individual samples of gDNA samples were available due to the small size of these tiny organisms. On the one hand, double gDNA digests were performed to cut at both *MLE* extremities in an attempt to visualize full-length and putatively internally deleted forms of the transposons. On the other hand, double gDNA digests were also carried out using enzymes that did not cut within the *MLE* consensus sequence. This made it possible to separate the distinct *MLEs* present in a genome, and to estimate the approximate number of elements present. This also provides a way of checking that the patterns obtained are different in order to confirm that the presence of *Bytmar1* relatives in the gDNA samples is not due to microbial contamination.

In *M. brachydactyla*, one 1056-bp fragment strongly hybridized with the probe in all *Eco*RI/*Pst*I gDNA digests (Fig. 3a). This fragment had the expected size of the internal region of full-length *Maibmar1* digested by these enzymes and indicated that most of the elements were full-length. A smear was observed, indicating that numerous *MLEs* were probably present in the genome of *M. brachydactyla*. This interpretation was confirmed by bands and smears in the hybridization patterns obtained with *Apa*I/*Pvu*II digests, which do not cut within the *Maibmar1* sequences.

In *C. pagurus* (Fig. 3b), one 1033-bp fragment strongly hybridized with the probe in all *Ava*I/*Pvu*II gDNA digests. This fragment corresponded to the internal region of full-length *Canpmar1*. Four weaker fragments were also detected, approximately 700, 400, and less than 100-bp in size, respectively. Internally deleted forms, ranging from 300 bp to 1 kbp, had never been detected by PCR using *ITRBytmar1* as primer, so we interpreted the presence of these four bands as reflecting the presence of 5′ and/or 3′ truncated forms. The *MLE* external digestions using *Apa*I/*Pvu*II revealed only a single strong signal at approximately

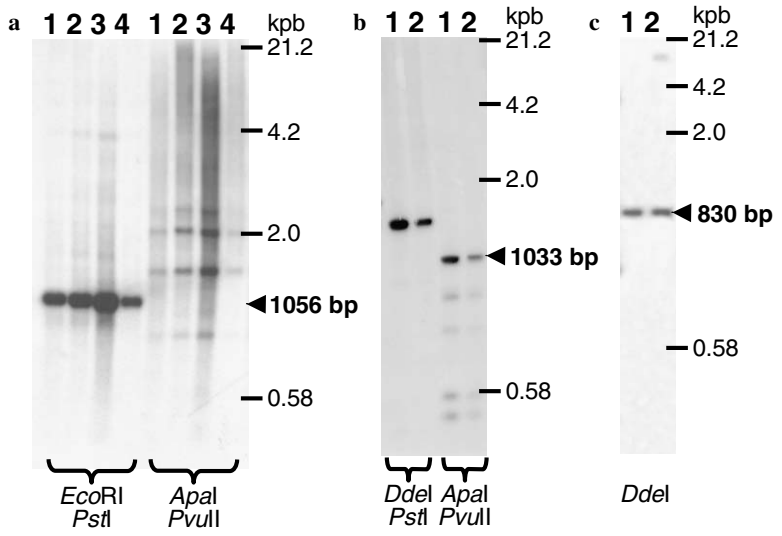

Fig. 3. Southern blot analyses of the *Bytmar1* relatives present in the genomes of *M. brachydactyla* (a), *C. pagurus* (b), and *V. sulfuris* (c). (a) 7 μg gDNA of four different *M. brachydactyla* imagoes were digested by *Eco*RI/*Pst*I or *Apa*I/*Pvu*II prior separation onto a 0.8% agarose gel and blotting. (b) 7 μg gDNA of two *C. pagurus* specimens were digested by *Dde*I/*Pst*I or *Apa*I/*Pvu*II prior separation onto a 1% agarose gel and blotting. (c) 2 μg gDNA of two *V. sulfuris* specimens were digested by *Dde*I prior separation onto a 1.5% agarose gel and blot. The blots were hybridized with *Maibmar1.4* as probe (a and b) and with *Vensmar1.3*. Final washings were done in 0.2× SSC at 65 °C. Molecular weights are indicated in the right margin. Exposure times at −80 °C were 12 h in (a), 72 h in (b), and 24 h in (c).

1300 bp. No smear was detected, indicating that few *MLEs* were present in the genome of *C. pagurus*.

In *V. sulfuris*, the hybridization pattern obtained in the *Dde*I digest revealed an 830-bp fragment with the expected size for an internal region of full-length *Vensmar1*.This confirmed the presence of *MLEs* in the host genome.

The very considerable differences in genome size (the exposure times required to obtain similar signals for the different blots were about 12 h for *M. brachydactyla*, 72 h for *C. pagurus*, and 24 h for *V. sulfuris*) made it impossible to estimate the number of elements in the three species. The differences between the hybridization patterns of the gDNA of the three species and those previously found for *B. thermydron* confirmed that the presence of *MLEs* in these species was not attributable to microbial contamination (Halaimia-Toumi et al., 2004).

3.4. Phylogenetic relationships between Bytmar1-related elements

The evolution of *Bytmar1* relatives was analyzed at two levels using consensus trees that had been developed using Neighbor-Joining and parsimony methods.

First, phylogenetic relationships between *Bytmar1* relatives were investigated, using *Hsmar2* as the outgroup (Fig. 4). The first characteristic of the consensus trees identified was the separation of the *Bytmar1* relatives into two distinct clusters, the first containing only the coastal elements, *Maibmar1* and *Canpmar1*, and the second only the hydrothermal elements, *Bytmar1* and *Vensmar1*. The second characteristic was that the elements in the groups did not segregate with their host. In the coastal group, a bootstrap value of 93% included both the *Maibmar 1.2*, *Canp-*

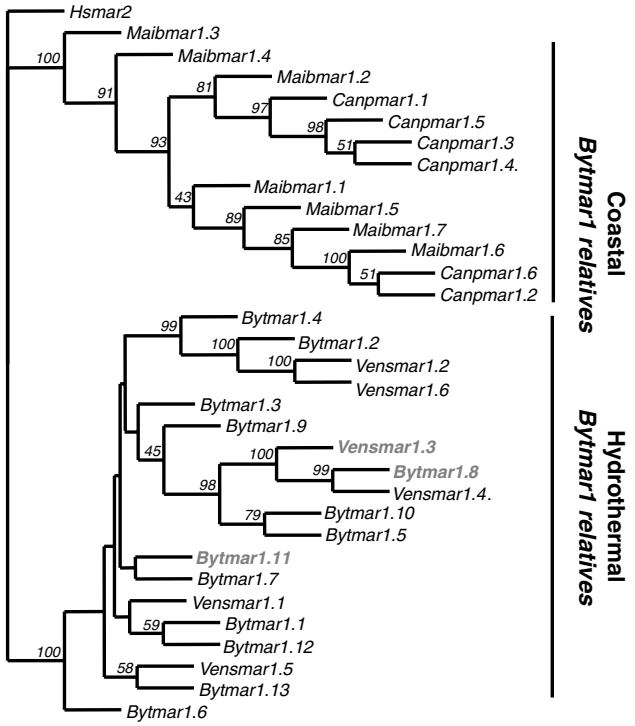

Fig. 4. Consensus phylogenetic tree of the *Bytmar1* relatives calculated with the aligned nucleic sequences of *Canpmar1*, *Maibmar1*, *Vensmar1* (Accession Nos. AJ507219–AJ50723) and *Bytmar1* using the parsimony procedure (similar results were obtained using the Neighbor-Joining procedure). *Hsmar2* (Accession No. U49974) is used as outgroup to root the tree. The numbers at the nodes of branches correspond to the bootstrap percentage calculated by sampling 1000 replicates. The environmental origins of the host species are specified in the right margin.

mar 1.1, *1.3*, *1.4*, and *1.5* sequences, and the *Maibmar 1.1*, *1.5*, *1.6*, *1.7*, and *Canpmar 1.2* and *1.6* sequences. In the hydrothermal group, most of the bootstrap values were very low, and so all the elements found in both species had very similar sequences (Table 1), even though the data suggested two lineages, the first of which included *Bytmar1.5*, *1.8* and *1.10*, and *Vensmar1.3* and *1.4*, and the second of which included *Bytmar1.4*, *1.2* and *Vensmar1.2* and *1.6*.

The significance of these findings was further investigated as they suggested that the host genome contained several populations of similar ages of *Bytmar1* and *Vensmar1*, on one hand, and of *Canpmar1* and *Maibmar1*, on the other hand. However, because we had previously demonstrated that the features of the ITR regions and the palindromy of the inner regions of the *MLE* resulted from a phenomenon of concerted evolution and are subjected to selection during and between horizontal transfer events (Bigot et al., 2005; Rouleux-Bonnin et al., 2005; personal data on the palindromy of *Bytmar1*-related element), we only used the rates of sequence similarity for our analyses. Indeed, the conventional methods available to investigate the evolution of the genic and non-coding genomic sequences (distance matrixes and pattern of mutation accumulation at silent and non-silent positions) are not appropriate for our purpose, because they use models that do not include the concerted impact of evolution within functional mobile sequences. Moreover, our position was reinforced by the fact that the data set was heterogeneous, as some of the *Bytmar1*-related elements corresponded to fossil elements: non-autonomous elements requiring a transposase source *in trans* for their mobility, and autonomous mobile elements encoding a putatively active transposase.

In agreement with our phylogenetic findings, the statistical analysis of the sequence similarity rates (similar results were obtained with distance matrices calculated using the Feselstein 81, the Jukes and Cantor and the Kimura's two transversion-parameter models) between *Canpmar1* and *Maibmar1* confirmed the existence of two groups of elements of which the members were more similar to each other than to the members of the other group. The first comprised *Maibmar 1.2*, *Canpmar 1.1*, *1.3*, *1.4*, and *1.5* ($98.08 \pm 1.03\%$ of intragroup similarity; $92.36 \pm 1.69\%$ of similarity versus group 2) and the second *Maibmar 1.1*, *1.3*, *1.4*, *1.5*, *1.6*, *1.7*, and *Canpmar 1.2* and 1.6 ($94.61 \pm 3.69\%$ of intragroup similarity). The difference in sequence degeneracy between members of the group 1 and 2 indicated that they had different ages and confirmed that the clustering of these elements to form two distinct groups was independent of the species from which they were cloned.

Similar analyses of the *Bytmar1* and *Vensmar1* sequences revealed three groups of elements within which the members were also more similar to each other than to members of the other two groups. The first comprised *Bytmar1.5*, *1.8*, and *1.10*, *Vensmar1.3* and *1.4* ($95.62 \pm 2.13\%$ intragroup similarity; $89.93 \pm 1.28\%$ and $87.92 \pm 3.02\%$ similarity with group 2 and 3, respectively), the second *Bytmar1.4*, *1.2*, and *Vensmar1.2* and *1.6* ($94.01 \pm 2.08\%$ of intragroup similarity; $90.09 \pm 3.06\%$ of similarity with group 3), and the third *Bytmar1.1*, *1.3*, *1.6*, *1.7*, *1.9*, *1.11*, *1.12*, and *1.13* and *Vensmar1.1* and *1.5* ($88.22 \pm 4.03\%$ intragroup similarity). The similarity levels between groups 1 and 2 confirmed that they were distinct and similar in age. Those of group 3 indicated that it consisted of more degenerate and aged sequences, although one member, *Bytmar1.11*, contained a complete ORF encoding a putatively active transposase. Overall, these results also confirmed that the clustering of the *Bytmar1* and *Vensmar1* elements was independent of the species from which they were cloned.

3.5. Phylogenetic relationships within the irritans MLE subfamily

Recently published results (Sinzelle et al., 2006) have indicated that at least four lineages have differentiated

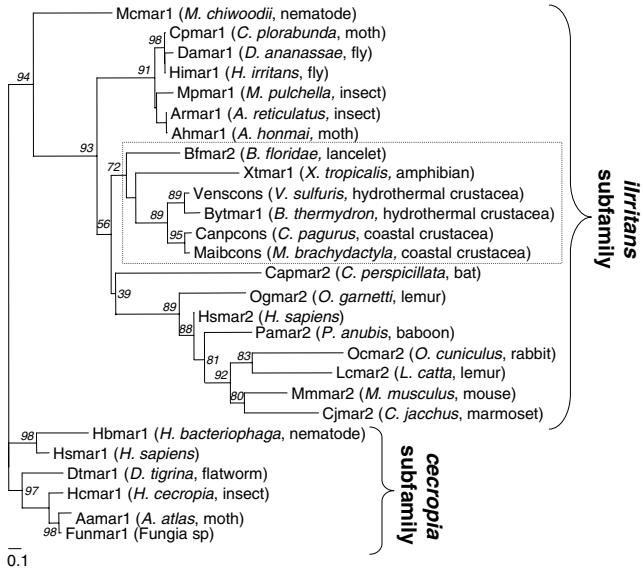

Fig. 5. Phylogenetic relationships of the transposases encoded by the *Bytmar1* relatives with those of the other *MLEs* so far identified and belonging to the *irritans* sub-family. Sub-family designations are indicated to the right of the tree. The tree was produced with the Neighbor-Joining method. The numbers at the nodes of branches correspond to the bootstrap percentage calculated by sampling 1000 replicates. *Ahmar1* (*Adoxophyes honmai*, *AB020617*); *Armar1* (*Ascogaster reticulates*, AB020618) *Bfmar2* (*B. floridae*; contig ti553931181 /ti539597141); *Capmar2* (*Carollia perspicillata*, Accession No. AC148202); *Cjmar2* (*Callitrix jacchus*, Accession No. AC150821); *Cpmar1* (*Chrysoperla plorabunda*; U11650); *Damar1* (*Drosophila ananassae*, U11646), *Himar1* (*Haematobia irritans*, U11642); *Hsmar2* (Homo sapiens, U49974); *Lcmar2* (*Lemur catta;* Accession No. AC133072); *Mmmar2* (*Mus musculus*, genomic contig NT_040837.01); *Mpmar1* (*Mantispa pulchella,* U11649); *Ocmar2* (*Oryctolagus cuniculus*, Accession No. AC147588); *Ogmar2* (*Otolemur garnetti*, Accession No. AC147523); *Pamar2* (*Papio anubis*, Accession No. AC126226); *Xtmar1* (*X. tropicalis*, sequence available at: http://genome.jgi-psf.org/Xentr3/Xentr3.home.html). The *MLE* sequences of *Mcmar1* (*Meloidogyne chiwoodi, AJ437557*) and of several members of the *cecropia* sub-family *Aamar1* (*Attacus atlas*, AB006464); *Dtmar1* (*Dugesia tigrina*, X80776); *Hbmar1* (*Heterorhabditis bacteriophaga*, U68392); *Funmar1* (*Fungia sp*, AB055188); *Hsmar1* (*Homo sapiens*, U52077); *Hcmar1* (*Hyalophora cecropia*, M63844) were used as outgroups to root the tree. Elements belonging to each of the four *irritans* lineage are boxed. In the *Bytmar1*-like lineage, elements belonging to the GC-poor and -rich sub-lineages are typed in clear and dark gray, respectively.

within the *irritans* sub-family during the course of evolution, depending on the host taxon in which they occurred. We, therefore, re-investigated the evolutionary relationships of the *Bytmar1* relatives to other *MLEs*, using *Mcmar1* and several members of the *cecropia* sub-family as outgroups (Fig. 5). Our data confirmed the existence of a *Bytmar1*-like lineage in the *irritans* sub-family that comprised the *Bytmar1*-related elements and *MLEs* isolated in the genomes of the lancelet *Branchiostoma floridae* (*Bfmar2*) and the toad *Xenopus tropicalis* (*Xtmar1*). Strikingly, all the members of this lineage were found in the genomes of organisms that occur in aquatic environments. Our analyses of the GC content and the CpG and TpA ratios also suggested that this lineage could probably be divided into two sub-lineages, as the features in *Bfmar2* and *Xtmar1* were very similar to those found in terrestrial *MLEs*, whereas they were very different from those found among *Bytmar1* relatives.

4. Discussion

The present data demonstrate that *Bytmar1*-related elements are present in the genomes of two coastal crustacean species, *C. pagurus* and *M. brachydactyla*, which belong to the same clade as *B. thermydron*, the Reptantia, and in that of one hydrothermal species, *V. sulfuris*, that belongs to a very distantly related clade, the Amphipoda (Fig. 1). Overall, the "low" levels of similarity between the nucleic acid sequences of *Canpmar1* or *Maibmar1*, on one hand, and of *Bytmar1* or *Vensmar1*, on the other hand, and the absence of close relatives of *Canpmar1* and *Maibmar1* in the *B. thermydron* and *V. sulfuris* genomes and the absence of close relatives of *Bytmar1* and *Vensmar1* in the genomes of the hosts *C. pagurus* and *M. brachydactyla* led us to investigate two distinct hypotheses.

The first hypothesis proposes that sequence and presence/absence polymorphisms of *Bytmar1*-related elements in the individuals of the ancestor species of the Decapoda and the Xenommacarida (about 500 millions years ago) might be the origin of the clustering found between coastal and hydrothermal species in our phylogenetic analyses. According to this hypothesis, it would be expected that, during evolution, the genomes of the *B. thermydron* and *V. sulfuris* ancestors would have lost the elements closely related to *Canpmar1* and *Maibmar1*, whereas in the genomes of the *C. pagurus* and *M. brachydactyla* ancestors, it would have been the elements closely related to *Bytmar1* and *Vensmar1* that have been erased. Although such an explanation is always difficult to discount in the absence of complete sequencing data for the genomes, it does not account for some very high levels of interspecies similarity observed for nucleic acid sequences, such as those between *Maibmar1.6* and *Canpmar1.6* (99.4%), and *Bytmar1.8* and *Vensmar1.3* or *Vensmar1.4* (99 and 99.5%). It also does not seem to explain how groups or populations of closely related elements can occur in different species in the absence of recurrent horizontal transfers.

The second hypothesis that could account for this kind of observation is that several horizontal transfers have occurred between these species or their recent ancestors at similar times. High rates of horizontal transfer of transposable elements between related species has recently been demonstrated to have occurred among dipteran species (Sanchez-Gracia et al., 2005; de Almeida and Carareto, 2005). In the present case, this hypothesis is supported by the geographical sympatry of the specimens: our *M. brachydactyla* and *C. pagurus* samples originated from the English Channel (North coast of Brittany), and *B. thermydron* and *V. sulfuris* had been taken from the same hydrothermal vents (Halaimia-Toumi et al., 2004). Our findings, therefore, suggests that the presence of different populations of *Canpmar1* and *Maibmar1* in some Decapod species living in coastal environments would result from recurrent horizontal transfers occurring at similar times, as our phylogenetic analyses are unable to detect any differentiation of the *Canpmar1* and *Maibmar1* elements that matches their presence in the genomes of two crustacean hosts (Fig. 4, coastal sub-lineage). These data also indicate that recurrent and recent horizontal transfers of *Bytmar1*-related elements have occurred between the two hydrothermal sympatric species investigated here, or from some other sympatric species into both these species, as suggested by the 99.5% similarity between the nucleic acid sequences of the two putatively active elements *Vensmar1-3* and *Bytmar1-8*. It will be necessary to analyze a wider range of species to confirm the occurrence and the frequency of these horizontal transfers, as our sample is not large enough to be considered to be representative of crustacean evolution. However, our preliminary sequence analyses of the *Bytmar1* relatives isolated from the closely related crab species *C. preadator* (Reptantia, Brachyurans), and the more distantly related shrimp species, *M. fortunata* and *R. exoculata* (Caridae), supported the hypothesis that very frequent horizontal transfers have occurred between sympatric crustacean species living in hydrothermal environments.

Here, we also confirm that the *irritans* sub-family has differentiated into four lineages during evolution. So far, the available data make it look as though this evolutionary scenario has occurred mainly within hosts living in similar environments or that have a "similar" biology. Indeed, the three first lineages seem to be restricted to host genomes from species corresponding to continental invertebrates or vertebrates, whereas members of the fourth lineage have only evolved in aquatic vertebrates and marine invertebrates. Some properties of the nucleic acid sequences: the codon usage in the transposase ORF, the GC content, and the CpG obs/exp and TpA obs/exp ratio also indicate that this third lineage can probably be differentiated into two distinct sub-lineages, the first of which comprises elements found in aquatic vertebrates, which have sequence properties similar to those of the "continental" *MLEs*, and the second of which comprises marine *Bytmar1* relatives. Further information about the occurrence of these *MLEs* in the genomes of other aquatic

vertebrates and marine invertebrates will be required to confirm this last proposition.

Finally, our results confirm our earlier observation that there are significant differences between the *MLEs* so far isolated from the genomes of continental species and those described here in the genomes of marine invertebrates species. To date, the genomes of two marine chordates, *C. intestinalis* and *B. floridae*, have been completely sequenced, but to the best of our knowledge, no genome from a marine invertebrate is yet available. This lack of data makes it impossible to know whether these properties are a common trait of the genes of their marine hosts. This means that it is too soon to decide whether the specificities of the marine *Bytmar1* relatives (the presence of an N-flag fused to the transposase, the usage of a GC-rich codon in the transposase ORF, a high GC content, and very low CpG obs/exp and TpA obs/exp ratio) result directly from selection pressure due to the conditions of marine and hydrothermal life. Such information should be soon available when the complete sequencing project of transcribed sequences of the hydrothermal invertebrate *Alvinella pompejana* has been achieved by the *Alvinella* consortium (http://alnitak.u-strasbg.fr/Alvinella/). However, our data already provide some information, suggesting that the marine environment may have had a major impact on the evolution of the *Bytmar1* transposase. For example, although we used electrophoretic mobility shift assays (manuscript in preparation) to confirm that there is a functional *cro* DNA-binding domain (Augé-Gouillou et al., 2000, 2005a,b; Bigot et al., 2005) with a cardinal HTH motif in its N-terminal domain of the *Bytmar1* transposase, we found that none of the HTH prediction software available on the web was able to locate it in this protein, or in those of any of its marine relatives. This would seem to indicate that some of the bioinformatic tools, which currently use matrices devised using data originating mainly from continental organisms, will probably need to be adapted to make them suitable for the accurate processing of the molecular data recovered from marine deep-sea hydrothermal organisms.

Acknowledgments

We thank the Senior Scientist, F. Lallier, the captain and the crew of the Hope 99 cruise for their support. We thank J. Rossignol and S. Amiard for their technical assistance. We thank the reviewers for their stimulating remarks, which have helped us to improve our interpretations. The English text was revised by Monika Ghosh. The studies were funded by grants from the Institut National des Sciences de l'Univers, the Ministère de l'Éducation de la Recherche et de la Technologie, the governments of the French regions Pays de Loire and Centre, the Centre National de la Recherche Scientifique, the Groupement de Recherche CNRS 2157, the Institut Fédératif de Recherche 136, the University of Le Mans and the University François Rabelais of Tours.

References

Arkhipova, I.R., Meselson, M., 2000. Transposable elements in sexual and ancient asexual taxa. Proc. Natl. Sci. Acad. USA 97, 14473–14477.

Augé-Gouillou, C., Notareschi-Leroy, H., Abad, P., Periquet, G., Bigot, Y., 2000. Evolution of DNA-binding and catalytic domains in *mariner*-like transposases: consequences on their phylogenetic analyses. Mol. Gen. Genet. 264, 514–520.

Augé-Gouillou, C., Brillet, B., Hamelin, M.H., Bigot, Y., 2005a. Assembly of the *mariner Mos-1* synaptic complex. Mol. Cell. Biol. 25, 2861–2870.

Augé-Gouillou, C., Brillet, B., Germon, S., Hamelin, M.H., Bigot, Y., 2005b. *Mariner Mos1* transposase dimerizes prior to ITR binding. J. Mol. Biol. 351, 117–130.

Bigot, Y., Brillet, B., Augé-Gouillou, C., 2005. Conservation of palindromic and mirror motifs within inverted terminal repeats of *mariner*-like elements. J. Mol. Biol. 351, 108–116.

Casse, N., Pradier, E., Bigot, Y., Loiseau, C., Laulier, M., 2000. *Mariner*, a mobile DNA transposon characterized in the genomes of several hydrothermal invertebrate species. Interridge 9, 15–17.

Casse, N., Pradier, E., Demattei, M.V., Bigot, Y., Laulier, M., 2002. *Mariner* transposons are genetic parasites occurring in the genome of hydrothermal invertebrates. Cah. Biol. Mar. 43, 325–328.

Claudianos, C., Brownlie, J., Russell, R., Oakeshott, J., Whyard, S., 2002. *maT*-A clade of transposons intermediate between *mariner* and *Tc1*. Mol. Biol. Evol. 19, 2019–2101.

de Almeida, L.M., Carareto, C.M.A., 2005. Multiple events of horizontal transfer of the *Minos* transposable elements between *Drosophila* species. Mol. Phylogenet. Evol. 35, 583–594.

Felsenstein, J., 1993. PHILIPS (Phylogeny Inference Package) version 3.5.c. University of Washington, Seattle.

Feschotte, C., Osterlund, M.T., Peeler, R., Wessler, S.R., 2005. DNA-binding specificity of rice *mariner*-like transposases and interactions with *Stowaway* MITEs. Nucleic Acid Res. 33, 2153–2165.

Halaimia-Toumi, N., Casse, N., Demattei, M.V., Renault, S., Pradier, E., Bigot, Y., Laulier, M., 2004. The GC-rich transposon *Bytmar1* from the deep-sea hydrothermal crab, *Bythograea thermydron*, may encode three transposase isoforms from a single ORF. J. Mol. Evol. 59, 747–760.

Karlin, S., Burge, C., 1995. Dinucleotide relative abundance extremes: a signature. Trends Genet. 11, 283–290.

Kidwell, M.G., Lisch, R., 2001. Perspective: transposable elements, parasitic DNA, and genome evolution. Evolution 55, 1–24.

Krieger, M.J.B., Ross, K.G., 2003. Molecular evolutionary analyses of *mariners* and other transposable elements in fire ants (Hymenoptera: Formicidae). Insect Mol. Biol. 12, 155–165.

Lampe, D.J., Witherspoon, D.J., Soto-Adames, F.N., Robertson, H.M., 2003. Recent horizontal transfer of mellifera subfamily *mariner* transposons onto insect lineages representing four different orders shows that selection acts during horizontal transfer. Mol. Biol. Evol. 20, 554–562.

Laulier, M., Pradier, E., Bigot, Y., Périquet, G., 1995. An easy method for preserving nucleic acids in field samples for later molecular and genetic studies without refrigerating. J. Evol. Biol. 8, 657–663.

Leroy, H., Castagnone-Sereno, P., Renault, S., Augé-Gouillou, C., Bigot, Y., Abad, P., 2003. Characterization of *Mcmar1*, a *mariner*-like element with large ITR from the phytoparasitic nematode *Meloidogyne chiwoodi*. Gene 304, 35–41.

Mandrioli, M., 2003. Identification and chromosomal localization of *mariner*-like elements in the cabbage moth *Mamestra brassicae* (Lepidoptera). Chromatogr. Res. 11, 319–322.

Plasterk, R.H.A., Van Luenen, G.A.M., 2002. The *Tc1/mariner* family of transposable elements. In: Craig, N.L., Robert Craigie, R., Gellert, M., Lambowitz, A. (Eds.), Mobile DNA II. ASM Press, Washington, DC, pp. 519–531.

Robertson, H.M., 1993. The *mariner* transposable element is widespread in insects. Nature 362, 241–245.

Robertson, H.M., 1997. Multiple *mariner* transposons in Flatworms and Hydras are related to those of *Insects*. J. Heredity 88, 195–201.

Robertson, H.M., 2002. Evolution of DNA transposons in eukaryotes. In: Craig, N.L., Robert Craigie, R., Gellert, M., Lambowitz, A. (Eds.), Mobile DNA II. ASM Press, Washington, DC, pp. 1093–1110.

Robertson, H.M., Asplund, M.L., 1996. *Bmmar1*: a basal lineage of the *mariner* family of transposable elements in the silkworm moth, *Bombyx mori*. Insect Biochem. Mol. Biol. 26, 945–954.

Robertson, H.M., Lampe, D.J., 1995. Recent horizontal transfer of a *mariner* transposable element among and between Diptera and Neuroptera. Mol. Biol. Evol. 12, 850–862.

Robertson, H.M., Martos, R., 1997. Molecular evolution of the second ancient human *mariner* transposon, *Hsmar2*, illustrates patterns of neutral evolution in the human genome lineage. Gene 205, 219–228.

Robertson, H.M., Walden, K.K.O., 2003. *Bmmar6*, a second *mori* subfamily *mariner* transposon from the silkworm moth *Bombyx mori*. Insect Mol. Biol. 12, 167–171.

Robertson, H.M., Zumpano, K.L., 1997. Molecular evolution of an ancient *mariner* transposon, *Hsmar1*, in the human genome. Gene 205, 203–217.

Rouleux-Bonnin, F., Petit, A., Demattei, M.V., Bigot, Y., 2005. Evolution of full-length and deleted forms of the *mariner*-like element, *Botmar1*, in the genome of the bumble bee, *Bombus terrestris* (Hymenoptera: Apidae). J. Mol. Evol. 60, 736–747.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, second ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanchez-Gracia, A., Maside, X., Charlesworth, B., 2005. High rate of horizontal transfer of transposable elements in *Drosophila*. Trends Genet. 21, 200–203.

Shao, H., Tu, Z., 2001. Expanding the diversity of *IS630-Tc1-mariner* superfamily: discovery of a unique DD37E transposon and reclassification of DD37D and DD39D transposons. Genetics 159, 1103–1115.

Silva, J.C., Bastida, F., Bidwell, S.L., Johnson, P.J., Carlton, J.M., 2004. A potentially functional *mariner* transposable element in the protist *Trichomonas vaginalis*. Mol. Biol. Evol. 22, 126–134.

Sinzelle, L., Chesneau, A., Pollet, N., Bigot, Y., Mazabraud, A., 2006. The *mariner* transposons belonging to the *irritans* subfamily were maintained in chordate genomes by vertical transmission. J. Mol. Evol. 62 (1), 53–65.

Southern, E.M., 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503–517.

Sumitani, M., Lee, J.M., Hatakeyama, M., Oishi, K., 2002. Cloning and characterization of Acmar1, a mariner-like element in the asiatic honey bee, *Apis cerana* japonica (Hymenoptera, Apocrita). Arch. Insect Biochem. Physiol. 50, 183–190.