

RESEARCH ARTICLE

Open Access

# A bacterial genome in transition - an exceptional enrichment of IS elements but lack of evidence for recent transposition in the symbiont *Amoebophilus asiaticus*

Stephan Schmitz-Esser<sup>1,3\*†</sup>, Thomas Penz<sup>1†</sup>, Anja Spang<sup>2</sup> and Matthias Horn<sup>1</sup>

## Abstract

**Background:** Insertion sequence (IS) elements are important mediators of genome plasticity and are widespread among bacterial and archaeal genomes. The 1.88 Mbp genome of the obligate intracellular amoeba symbiont *Amoebophilus asiaticus* contains an unusually large number of transposase genes ( $n = 354$ ; 23% of all genes).

**Results:** The transposase genes in the *A. asiaticus* genome can be assigned to 16 different IS elements termed ISCaa1 to ISCaa16, which are represented by 2 to 24 full-length copies, respectively. Despite this high IS element load, the *A. asiaticus* genome displays a GC skew pattern typical for most bacterial genomes, indicating that no major rearrangements have occurred recently. Additionally, the high sequence divergence of some IS elements, the high number of truncated IS element copies ( $n = 143$ ), as well as the absence of direct repeats in most IS elements suggest that the IS elements of *A. asiaticus* are transpositionally inactive. Although we could show transcription of 13 IS elements, we did not find experimental evidence for transpositional activity, corroborating our results from sequence analyses. However, we detected contiguous transcripts between IS elements and their downstream genes at nine loci in the *A. asiaticus* genome, indicating that some IS elements influence the transcription of downstream genes, some of which might be important for host cell interaction.

**Conclusions:** Taken together, the IS elements in the *A. asiaticus* genome are currently in the process of degradation and largely represent reflections of the evolutionary past of *A. asiaticus* in which its genome was shaped by their activity.

**Keywords:** insertion sequence element, endosymbiont, *Bacteroidetes*, genome evolution

## Background

Mobile genetic elements such as phages, plasmids and transposable elements play a vital role in horizontal gene transfer and genome rearrangement in bacteria and archaea [1]. Among transposable elements, insertion sequence (IS) elements are particularly widespread within bacterial and archaeal genomes, and are considered the most abundant and ubiquitous genes in nature [2-6]. IS elements can have profound effects on chromosome structure and evolution. Due to their ability to

disrupt genes and to induce rearrangements such as inversions, duplications and deletions they are key mediators of genome plasticity [2,3,7-9]. Although IS elements are perceived primarily as genomic parasites, their activity can also be beneficial. As composite transposons IS elements are able to mobilize adjacent genes, thereby mediating the spread of antibiotic resistance genes and genes involved in the catabolism of complex xenobiotics [10,11]. IS elements may also promote adaptation of their host genomes as demonstrated in experimental evolution experiments [12-15]. In addition, IS elements can influence or activate the expression of adjacent genes, e.g. by forming hybrid or fusion

\* Correspondence: stephan.schmitz-esser@vetmeduni.ac.at

† Contributed equally

<sup>1</sup>Department of Microbial Ecology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria

Full list of author information is available at the end of the article

promoters or by containing outward-directed promoters [16-22].

IS elements are usually less than 2.5 kbp in length and have a relatively simple genetic organization. Most IS elements are flanked both by inverted and direct repeats and generally encode no function other than those involved in mobility, which is mediated by transposases [16]. IS elements have been classified into several families based on the degree of sequence conservation of their transposases and its catalytic site, similar genetic organization such as size, number of open reading frames (ORFs) and potential coding sequences (CDSs), inverted repeats, and genome target sites [2,16]. The majority of IS elements encode transposases containing the so-called DDE-motif consisting of the three amino acids aspartic acid, aspartic acid, and glutamic acid. These residues form the catalytic triad necessary for transposition. They are found in three regions (N2, N3, and C1) of the transposase amino acid sequence separated by spacers of various lengths [2,16].

Although IS elements are found in the majority of sequenced bacterial and archaeal genomes [2-4], their distribution is patchy, and their occurrence within single genomes is usually below 3% [2,6]. IS elements are very rare in the genomes of most ancient host-restricted symbionts or pathogens such as mutualistic insect and clam symbionts or chlamydiae [23,24]. On the other hand, elevated numbers of IS elements has been observed in the genomes of bacteria which adapted only recently to an intracellular or pathogenic lifestyle [8,25-27]. However, this view has been challenged by the recent detection of IS element-rich genomes in ancient symbionts such as *Wolbachia* spp. or *Orientia tsutsugamushi* [9,28-31]. Interestingly, the genomes containing the highest percentages of IS elements are from obligate intracellular bacteria: *Orientia tsutsugamushi* [28,29], the  $\gamma$ 1 symbiont of the marine oligochaete *Olavivus algarvensis* [32], the symbionts of grain weevils [26,33], and the amoeba symbiont *Amoebophilus asiaticus* 5a2 [34].

*Amoebophilus asiaticus* is a Gram-negative, obligate intracellular symbiont, which has been discovered within an amoeba isolated from alkaline lake sediment [35]. Highly similar *A. asiaticus* strains have been recovered from various sources worldwide [35-38]. *A. asiaticus* shows highest 16S rRNA similarity to '*Candidatus Cardinium hertigii*', an obligate intracellular parasite of arthropods able to manipulate the reproduction of its hosts [39]. Both organisms belong to the phylum *Bacteroidetes* and form a monophyletic lineage in 16S rRNA-based phylogenetic trees [35], consisting only of symbionts and sequences retrieved from coral samples [40]. The *A. asiaticus* genome is only moderately reduced in size compared to many other obligate intracellular

bacteria [41,42] but nevertheless, its biosynthetic capabilities are extremely limited [34]. The *A. asiaticus* genome encodes a hitherto unparalleled high number of proteins with eukaryotic domains such as ankyrin repeats, TPR/SEL1 repeats, leucine-rich repeats and domains from the eukaryotic ubiquitin system, and it contains an unusually large number of transposase genes ( $n = 354$ ) corresponding to 23% of all genes [34].

Here, we report on the in-depth analysis of the IS elements in the *A. asiaticus* genome. We classified them and describe their main characteristics. We demonstrated that other symbionts closely related to *A. asiaticus* contain highly similar IS elements, and we could show that although they are transcribed, they exhibited no transpositional activity on a population level during a time period of almost 1,000 days. Taking into account evidence that no major rearrangements have occurred recently in the *A. asiaticus* genome, this suggests that the IS elements are evolutionary older components of the *A. asiaticus* genome, which likely played an important role during genome reduction and adaptation to an obligate intracellular life style.

## Results

### Diversity of IS elements in the *A. asiaticus* 5a2 genome

IS elements make up 183 kbp (10%) of the *A. asiaticus* genome. In total, 354 transposase genes (corresponding to 23% of all CDSs) were identified in the detailed and manually curated analysis performed here (Tables 1, 2). Compared to other sequenced prokaryotic genomes, the percentage of IS elements as well as the number of IS elements per megabase genome is among the highest in *A. asiaticus* (Additional file 1, Figures S1, S2). We were able to assign the vast majority of these transposase genes ( $n = 329$ , 93%; including partial IS element copies) to 16 different IS elements (ISCaa1 to ISCaa16), which belong to eight different IS element families, with IS5 family IS elements being most abundant in the *A. asiaticus* genome (Table 2). Each of the 16 IS elements is present in 2 to 24 full-length copies in the *A. asiaticus* genome, the only exception being ISCaa1, which was identified earlier by the ISFinder website [43] and is only present as a single full-length copy (Table 2). This results in a total copy number of 122 full-length IS elements that are evenly spread across the *A. asiaticus* genome [34]. A high number of IS elements in *A. asiaticus* is truncated ( $n = 143$ ), and in some cases (e.g. ISCaa5, ISCaa6 and ISCaa11) there are more truncated than full-length copies present (Table 2). Truncation sites were generally not conserved, i.e. truncations occurred in different regions, and truncated IS elements show varying lengths (Additional file 1, Figure S3). For most of the full-length IS element copies ( $n = 101$ , 83%) we could not identify direct repeats (Tables 1, 2). In the

**Table 1 IS element statistics for the genome of *A. asiaticus***

No. of protein coding genes	1557
No. of transposase encoding genes	354 (23% of all protein coding genes)
No. of transposase encoding genes assigned to IS elements	329 (93% of all transposase genes)
No. of full-length IS element copies*	122
No. of partial IS element copies*	143
No. of full-length IS element copies with functional transposase gene	106 (87% of all full-length IS element copies)
No. of full-length IS element copies without direct repeats	101 (83% of all full-length IS element copies)

\* Note that IS elements can consist of more than one transposase gene

following sections we shortly describe few selected IS elements of *A. asiaticus* in more detail.

#### ISCaa4

ISCaa4 is the most abundant IS element in *A. asiaticus*. It is present in 24 full-length copies, 21 of these copies should be able to produce an intact, functional transposase. ISCaa4 belongs to the IS1 family and shows a typical IS1 family DDE-motif [2,44]. Similar to other IS1 family members, the ISCaa4 transposase is encoded by two overlapping ORFs, which are probably translated into a 226 amino acid transposase by -1 ribosomal frameshifting (Table 2, Additional file 1, Figure S4). Translational frameshifting is often found in IS elements and represents an important mechanism regulating the expression of the transposases at a translational level [16,45]. Translation starts at the first ORF (*orfA*) and shifts to the -1 reading frame at the so-called slippery site and continues in a second overlapping ORF (*orfB*) resulting in a transframe ORFAB protein. The predicted frameshift site in ISCaa4 (AAAAAAG) is highly shift-prone in bacteria such as *A. asiaticus* that have only a single tRNA<sup>Lys</sup> (anticodon: UUU) and lack the tRNA<sup>Lys</sup> with the anticodon UUC [45,46]. In ISCaa4 five nucleotides downstream of the putative slippery site a stem-loop structure is predicted ( $\Delta G$  -6.3 kcal/mol) (Additional file 1, Figure S4). Such stem-loop structures have been shown to be stimulatory for -1 ribosomal frameshifting [45,46]. Interestingly, ISCaa4 shows highest amino acid sequence identity (46 to 51%) to uncharacterized IS elements from methanogenic archaea of the family *Methanosarcinaceae*; the similarity to other transposases is lower than 40%. In phylogenetic trees, ISCaa4 forms a stable monophyletic group with these archaeal transposases, indicating interdomain horizontal gene transfer between methanogenic archaea and *A. asiaticus* (Figure 1, Additional file 1, Figure S5).

#### ISCaa3

ISCaa3 is present in ten full-length copies in *A. asiaticus* and belongs to ISL2 group within the IS5 family (based

on the presence of a typical DDE-motif) whose transposases typically consist of a single ORF [2,16]. The transposase of ISCaa3 however, is encoded by two overlapping ORFs most likely translated into a 275 amino acid protein by -1 ribosomal frameshifting. In contrast to other IS elements with canonical slippery sites like ISCaa4 and ISCaa9, no stimulatory stem-loop structure possibly enhancing ribosomal frameshifting is predicted downstream of the slippery site in ISCaa3 (Additional file 1, Figure S6). ISCaa3 shows highest amino acid sequence identity (57 to 66%) to ISCaa2 and IS elements found in the intracellular bacteria *Orientia tsutsugamushi*, *Legionella drancourtii*, *Regiella insecticola*, and *Parachlamydia acanthamoebae*. In phylogenetic analyses ISCaa2, ISCaa3 and related IS elements from intracellular bacteria consistently group together in all treeing methods applied, suggesting horizontal transfer of IS elements between these intracellular bacteria (Figure 1, Additional file 1, Figure S7). Interestingly, a number of cyanobacterial IS elements form a sister group with the ISCaa3-related IS elements.

#### ISCaa9

ISCaa9 is an 881 bp IS element which is present in 15 almost identical copies (the differences occur only in the inverted repeats). ISCaa9 belongs to the IS5 family and shows highest amino acid sequence identity (45%) to ISMac15 from *Methanosarcina acetivorans* C2A, and 40% amino acid identity to ISWpi1, an IS element found in many *Wolbachia* strains [47,48]. The ISCaa9 transposase is encoded by three consecutive and overlapping ORFs which are translated into a 253 amino acid protein (Table 2, Additional file 1, Figure S8). We propose a stop codon read-through to occur at the stop codon (UGA) at nucleotide positions 263 to 265, which is supported by the presence of the stop codon in all 15 ISCaa9 copies in *A. asiaticus*, the absence of a stem-loop structure indicative of a terminator downstream of the stop codon, and the observation that UGAA is a weak stop codon quartet [49,50]. We predict that the

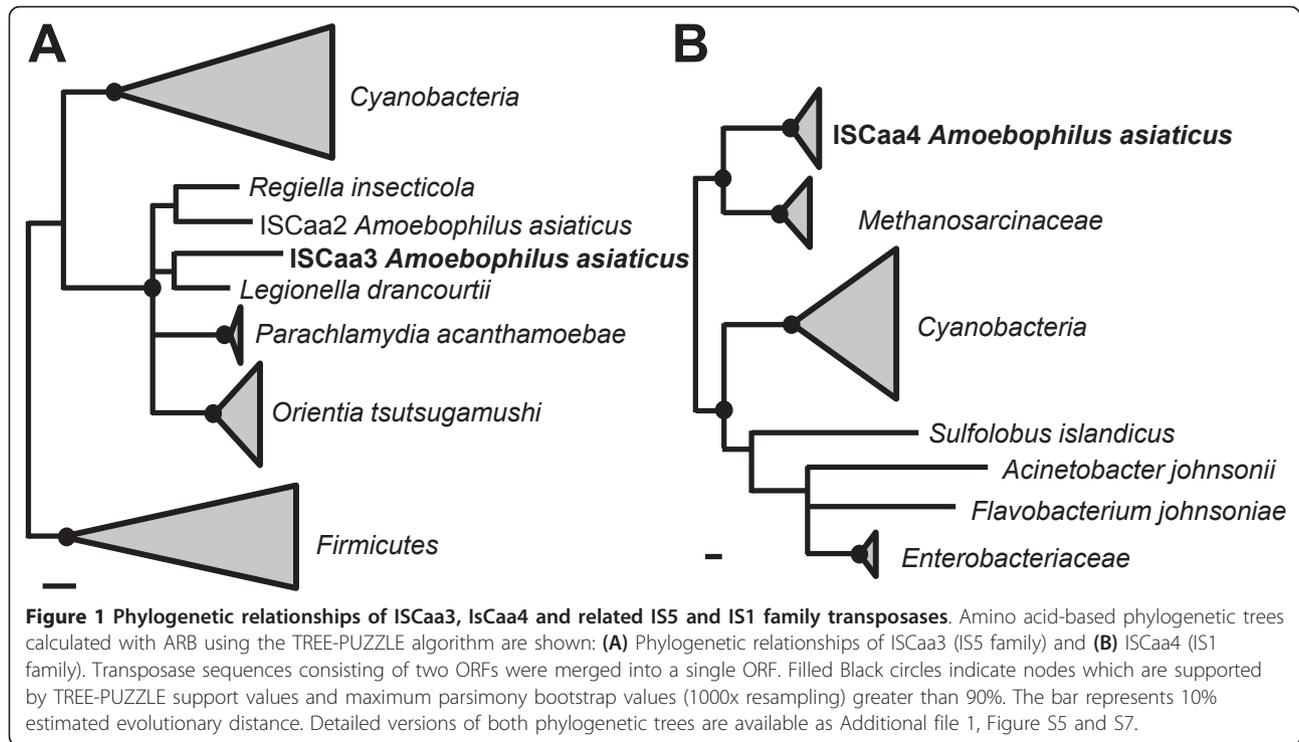
**Table 2 IS elements in the *A. asiaticus* 5a2 genome**

IS element	IS family	Number of ORFs (predicted translational frameshift)	Length of IS element [bp]	Inverted repeats [bp] <sup>a</sup>	Direct repeats [bp] <sup>b</sup>	G+C content of IS element [%] (range)	Length of transposase [amino acids]	Number of full-length IS element copies (conservation on DNA level)	Number of partial IS element copies	Number of full-length IS element copies with intact transposase genes (conservation on protein level)
ISCaa1 <sup>c</sup>	IS1	1	759	17/21	0	35.7	232	1	6	1
ISCaa2	IS5, ISL2 group	2 (-1)	916	19/20	0	36.3	275	3 (99-100%)	3	3 (99-100%)
ISCaa3	IS5 ISL2 group	2 (-1)	914	22/23	0	36.5	275	10 (99-100%)	5	9 (99-100%)
ISCaa4	IS1	2 (-1)	732	17/22	8/10	37.4 (36.2 - 38.0)	226	24 (85-100%)	8	21 (96-100%)
ISCaa5	IS982	1	932	18/21	0	38.2	274	10 (99-100%)	24	8 (99-100%)
ISCaa6	IS5, ISL2 group	1	991	15/19	0	36.6 (35.3 - 36.6)	275	18 (86-100%)	41	17 (88-100%)
ISCaa7	IS110	1	1483	0	0	31.9	343	3 (100%)	4	3 (100%)
ISCaa8	IS5, IS1031 group	2 (+1)	893	18/22	0	39.5	264	6 (99-100%)	6	6 (99-100%)
ISCaa9	IS5	3 (-1 ORFBC)	881	18/21	0	38.8	253	15 (100%)	3	15 (100%)
ISCaa10	IS200/IS605 IS200 group	1	527	0	0	38.5	147	7 (99-100%)	2	7 (99-100%)
ISCaa11	IS481	1	1031	10/11	6/3	38.9 (37.4 - 39.7)	314	10 (83-100%)	27	3 (87-100%)
ISCaa12	IS481	1	1210	29/34	6/2	37.6	364	3 (100%)	3	3 (100%)
ISCaa13	IS5, IS427 group	2 (+1)	860	17/21	0	40.7	253	2 (99%)	7	1
ISCaa14	IS110	1	1256	0	0	38.1	326	2 (97%)	0	1
ISCaa15	IS1182	1	1434	18/18	4/2	35.9	457	3 (100%)	3	3 (100%)
ISCaa16	IS6	1	837	15/18	0	37.4 (34.4 - 37.4)	235	5 (82-100%)	1	5 (85-100%)

<sup>a</sup> number of base pairs conserved between left and right end repeats/length of the repeat

<sup>b</sup> length of direct repeat/number of isoforms with direct repeat

<sup>c</sup> ISCaa1 was identified by the ISFinder website <http://www-is.biotoul.fr/>



stop codon is recoded into tryptophane (UGG), a common feature of UGAA stop codon quartets [50]. In addition, the majority of ISCaa9-related transposases encodes a tryptophane at the position of the stop codon read-through in ISCaa9 (Additional file 1, Figure S9). We predict a translational -1 frameshifting at a slippery site (AAAAAAG) between *orfB* and *orfC* (Additional file 1, Figure S7). Five nucleotides downstream of this putative slippery site, a stem-loop structure ( $\Delta G$  -12.6 kcal/mol) is predicted in ISCaa9 (Additional file 1, Figure S8). The ISCaa9 transposase contains a DDE-motif, which is most similar to the IS1031 group within the IS5 family, the transposases of this group however, are usually encoded by a single ORF [2].

#### ISCaa10

ISCaa10 is with a length of 527 bp a very short IS element that contains a single ORF encoding a 147 amino acid transposase. It belongs to the IS200/IS605 family and IS200 group of IS elements comprising the shortest known transposases [2]. Members of the IS200 group are unusual IS elements because their transposases do not contain the DDE-motif found in most transposases. Instead they belong to the Y1 transposases with a catalytic tyrosine residue and a conserved HuH motif (consisting of a histidine, a hydrophobic amino acid, and another histidine) [51,52]. Interestingly, this motif is present only in two of seven ISCaa10 copies; in the others, the second histidine is replaced by tyrosine, which

might render these copies nonfunctional. Other unusual features of IS200 IS elements, that are also found in ISCaa10, are the absence of both direct and terminal inverted repeats and the presence of secondary structures leading to low transcriptional and transpositional activity [51-53]. For example, IS200 from *Salmonella typhimurium* LT2 forms two stem-loop structures: The first is a transcriptional repressor terminating impinging transcripts, the second acts at the translational level and occludes the ribosome binding site [53]. Similarly, a stem-loop structure is predicted ten nucleotides upstream of the start codon of the ISCaa10 transposase and close to the 3' end of ISCaa10 ( $\Delta G$  -13 kcal/mol and -20.3 kcal/mol, respectively). ISCaa10 shows highest amino acid sequence identity (76%) to (uncharacterized) IS200 family transposases from *Xenorhabdus nematophila* (GenBank accession no: YP\_003712757).

#### Unclassified IS elements

Twenty-five transposase genes could not be assigned to either of the 16 *A. asiaticus* IS elements under the criteria applied here. Among these unclassified full-length transposases two transposases belong to the IS110 family (Aasi\_1379 and Aasi\_1284); and to the Tn3 family (Aasi\_0096, Aasi\_0545); one belongs to the IS3 family (probably consisting of the two consecutive ORFs Aasi\_1748 and Aasi\_0907); and two belong to the YhgA-like family of putative transposases (Aasi\_0894, Aasi\_1306; PFAM-family PF04754).

### Conservation of IS elements among different *A. asiaticus* strains

In order to analyze whether the IS elements found in the genome of *A. asiaticus* 5a2 are also present in closely related *A. asiaticus* strains, we performed PCR using primers targeting the 13 most abundant IS elements (Additional file 2, Table S1) with genomic DNA from *A. asiaticus* strain EIDS3 [35] as well as from two novel *A. asiaticus* isolates, *A. asiaticus* US1 and *A. asiaticus* WR. These strains show 98.9%, 99.2%, and 98.5% 16S rRNA sequence similarity to *A. asiaticus* 5a2, respectively, corresponding to strain and species level diversity, respectively. Six out of the 13 IS elements analyzed here were detected in all four *A. asiaticus* isolates. Cloning and sequencing of PCR products obtained from *A. asiaticus* EIDS3 revealed nucleic and amino acid sequence identities to consensus sequences of the *A. asiaticus* 5a2 IS elements of 87% to 98% (Table 3). The lack of PCR products for some IS elements indicates either the absence of these IS elements in the investigated *A. asiaticus* strains or a low degree of conservation and hence the absence of or mismatches with the primer target sites.

### Transcription but lack of transpositional activity of the *A. asiaticus* IS elements

The large copy number and the high degree of conservation of some IS elements identified in the *A. asiaticus*

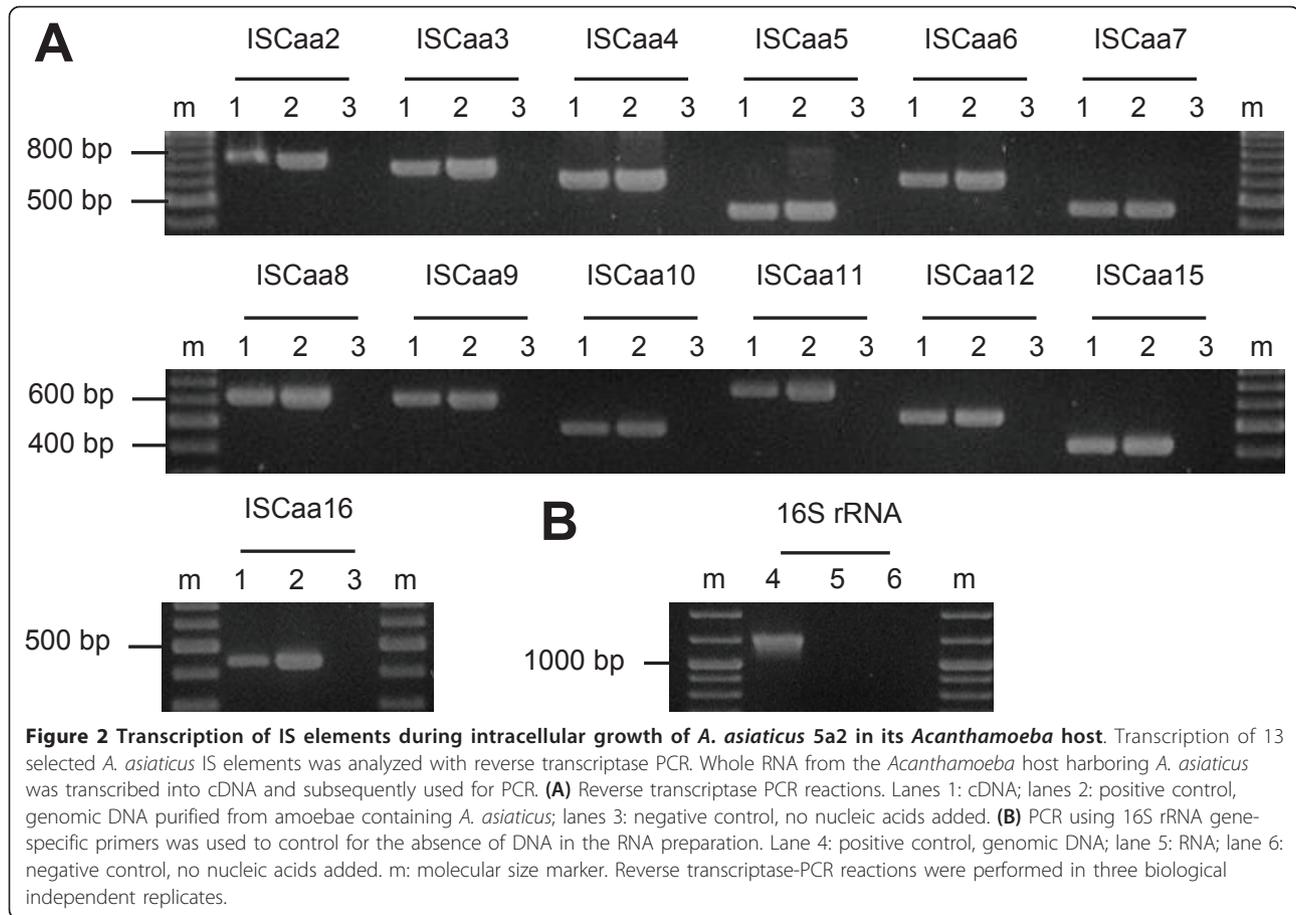
5a2 genome might indicate that they are transpositionally active. To investigate this, we first asked whether the IS elements are transcribed during intracellular replication of *A. asiaticus* in its amoeba host. Using reverse transcriptase (RT)-PCR, we analyzed the transcription of those 13 IS elements that are present in at least three copies in the genome (ISCaa2 to ISCaa12, ISCaa15 and ISCaa16). The detection of transcripts of all 13 IS elements demonstrates that at least one copy each is actively transcribed (Figure 2). Next, we used Southern hybridizations to check for chromosomal rearrangements resulting from transposition events [12,14,54]. We analyzed the same 13 IS elements for which we could show transcription and compared DNA from the same *A. asiaticus* culture isolated in November 2006 and in July 2009, respectively, a period of 984 days. We could not detect differences in the banding pattern indicative for chromosomal rearrangements in Southern hybridizations for any of the IS elements tested (Figure 3).

### Contiguous transcription of IS elements and their downstream genes

Some of the *A. asiaticus* IS elements are in close proximity to their downstream genes (with distances less than 50 bp). As previous reports have shown that IS elements can influence the transcription of neighboring

**Table 3 Occurrence of IS elements in four different *A. asiaticus* strains based on PCR.**

IS element in <i>A. asiaticus</i> 5a2	<i>A. asiaticus</i> EIDS3 (amino acid identity to <i>A. asiaticus</i> 5a2 element)	<i>A. asiaticus</i> WR	<i>A. asiaticus</i> US1
ISCaa2	+ >(95%)	-	-
ISCaa3	+ >(97%)	+	+
ISCaa4	-	+	-
ISCaa5	+ >(94%)	-	+
ISCaa6	+ >(92%)	-	+
ISCaa7	-	-	-
ISCaa8	+ >(91%)	+	+
ISCaa9	+ >(94%)	+	+
ISCaa10	+ >(98%)	+	+
ISCaa11	+ >(90%)	+	+
ISCaa12	+ >(98%)	+	+
ISCaa15	+ >(87%)	+	-
ISCaa16	-	-	-

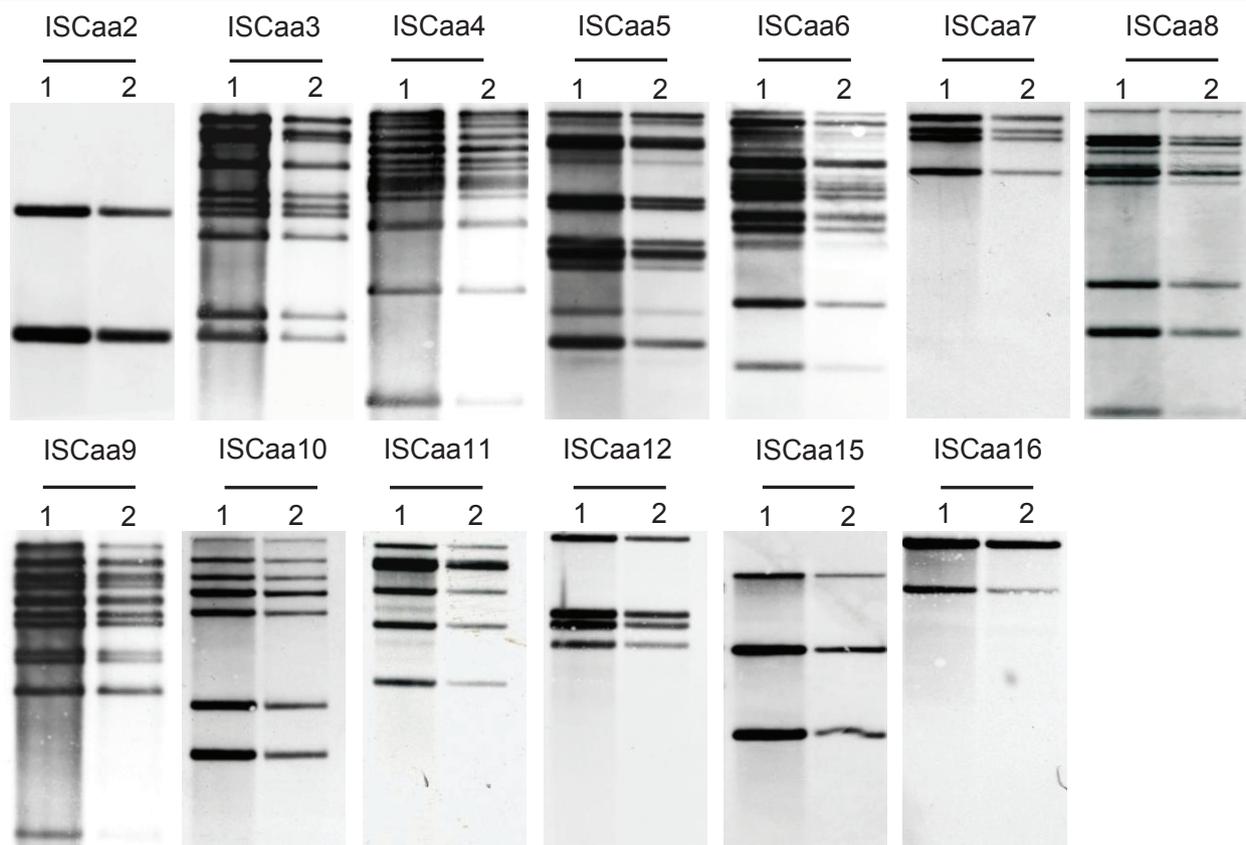


genes [17-21], we investigated whether contiguous transcripts between *A. asiaticus* IS elements and downstream genes occur. We analyzed ten selected loci where IS elements and their downstream genes are encoded on the same strand and have the same orientation (Figure 4). Using RT-PCR we could show contiguous transcripts of the investigated IS elements with their downstream genes at 9 out of 10 analyzed loci (Figure 5). We performed two control experiments in order to exclude that the observed transcripts from RT-PCR derive from unspecific background noise transcriptional read-through. One control targeted an unlikely contiguous transcript between two genes located on different strands and oriented in opposite directions (Aasi\_1200/1201, Figure 4). We could not detect transcripts in this control reaction (Figure 5), indicating that the observed transcripts from the nine loci of IS elements and their downstream genes are above unspecific read-through transcription. This is further supported by a second, semi-quantitative control experiment in which we compared RT-PCR products (using the same conditions) from contiguous transcripts between IS elements and their downstream genes with the products from RT-PCR reactions targeting only the downstream genes

(Additional file 1, Figure S10). In all cases the obtained bands were of similar intensity, providing further evidence that the observed contiguous transcripts are above unspecific transcriptional read-through.

### Discussion

Mobile genetic elements such as IS elements move within and between genomes. Owing to its intracellular lifestyle in free-living amoebae *A. asiaticus* is, however, largely shielded from other bacteria. Although horizontal gene transfer seems unlikely to occur under these circumstances, previous studies proposed that amoebae may serve as hot spots for horizontal gene transfer among intracellular bacteria [34,55], and according to the 'intracellular arena' hypothesis genetic material may move in and out of communities of obligate intracellular bacteria that co-infect the same intracellular host environment [23]. We identified four IS elements in *A. asiaticus* that were likely involved in horizontal gene transfer although the direction of the transfer cannot be inferred (ISCaa2, ISCaa3, ISCaa4, ISCaa12; Figure 1, Additional file 1, Figure S5, S7). Three of these IS elements group with IS elements from several other intracellular bacteria related to rickettsiae, legionellae and

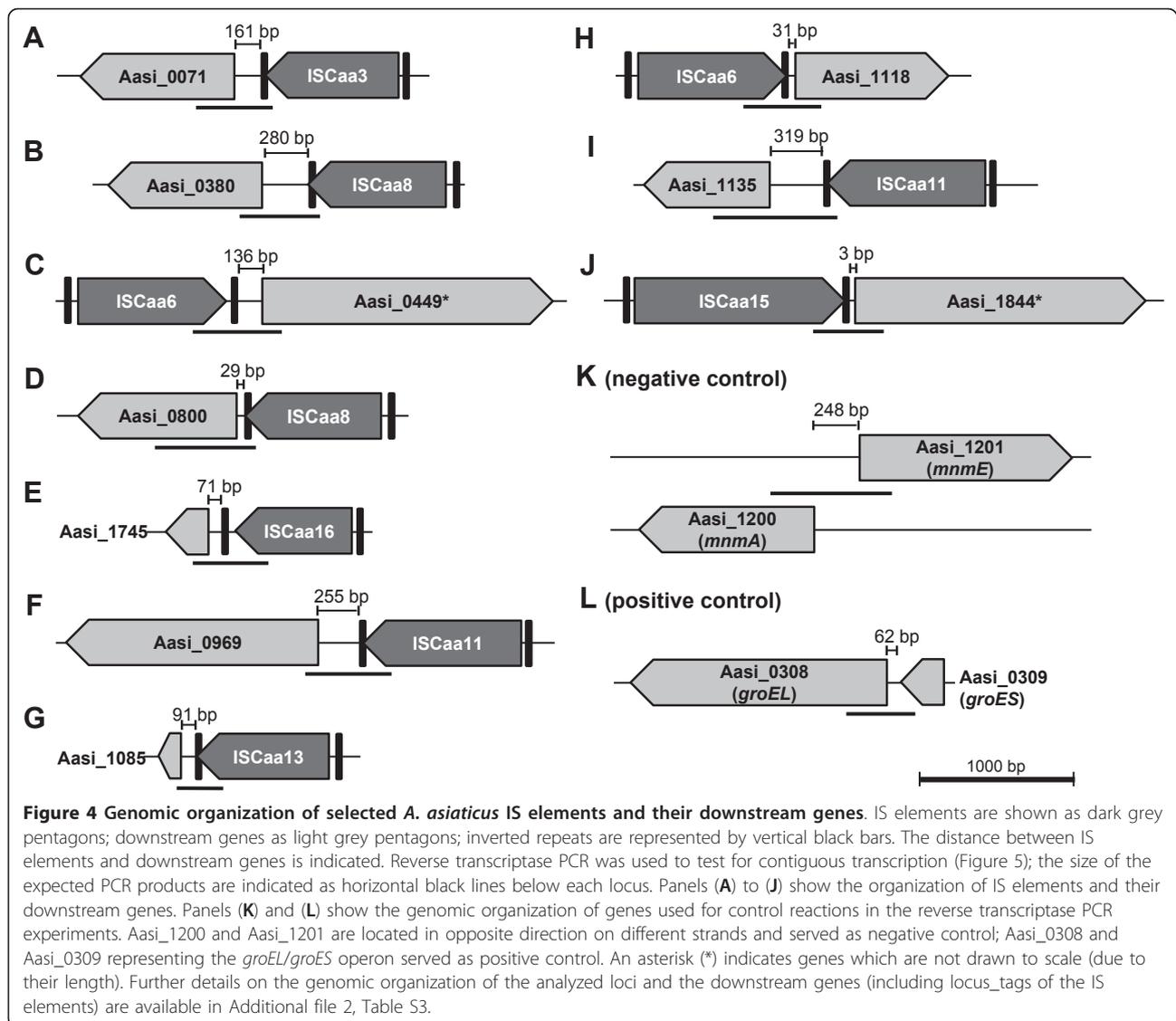


**Figure 3 Analysis of transpositional activity of the most abundant IS elements of *A. asiaticus* 5a2.** Transposition of IS elements was analyzed with Southern hybridizations using IS element-specific probes and DNA purified from the same *A. asiaticus* 5a2 culture in November 2006 and July 2009, respectively. DNA was digested with Eco32I (except for ISCaa2, where HindIII was used). Each visible band corresponds to at least one IS element copy on the respective DNA fragment, as the restriction endonucleases do not cut within the IS elements. IS elements are indicated above each hybridization; lanes 1: DNA isolated November 2006; lanes 2: DNA isolated July 2009. The absence of changes in the banding patterns between both time points indicates that no (major) chromosomal rearrangements due to IS element transposition has occurred.

chlamydiae, and consistent with previous findings it is conceivable that amoebae or other protozoa served as a common habitat for these microbes. One IS element of *A. asiaticus* is most closely related to IS elements found in free-living methanogenic archaea (*Methanosarcinaeae*). Anoxic aquatic sediments, where free-living amoebae and methanogenic archaea can be found, might represent a possible shared habitat facilitating horizontal gene transfer [56-58]. Horizontal gene transfer of IS elements between distantly related organisms is rather rare [59]. Hence the discovery of related IS elements in three different bacterial phyla (*Bacteroidetes*, *Proteobacteria*, *Chlamydiae*) and the *Archaea* might be surprising. However, a recent study based on the analysis of 800 bacterial and archaeal genomes showed that although the majority of horizontal gene transfer events occur between closely related organisms there is a considerable number of large-distance horizontal gene transfer events [60]. Our observations expand our view on the

extent of horizontal gene transfer of IS elements among distantly related microbes, and they provide a glimpse into past interactions of *A. asiaticus* with other microbes during its evolutionary history.

Several lines of evidence point to an ancient origin of many IS elements in *A. asiaticus*. First, ISCaa4, ISCaa6, ISCaa11, and ISCaa16, which together make up 46% of all full-length *A. asiaticus* IS elements, show a remarkably low degree of sequence conservation among their different copies (Table 2). This is in contrast to high sequence similarities expected if IS elements have entered a genome and spread only recently [4,61]. Second, the high number ( $n = 143$ ) of truncated IS element copies suggests that these IS elements have been present in the *A. asiaticus* genome for extended time periods during which they disintegrated slowly. Third, the GC-content of the *A. asiaticus* IS elements (37.3% on average, range: 31.9 to 40.7%) is similar to the overall GC content of the *A. asiaticus* genome (35.0%). This

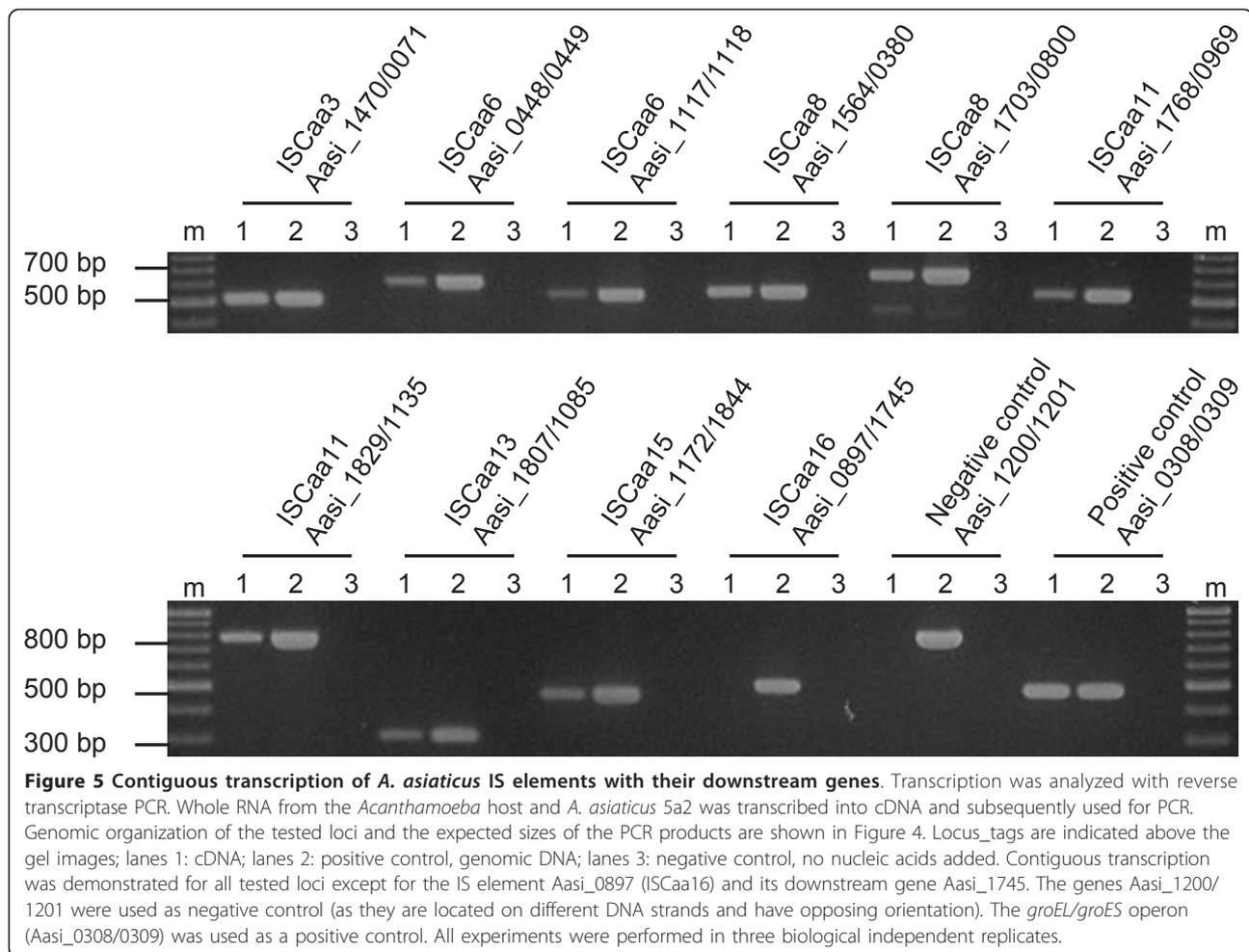


suggests that considerable time has elapsed to allow the base composition of the IS elements to adapt towards the general base composition of the *A. asiaticus* genome [62]. Finally, at least six IS elements are conserved among four different *A. asiaticus* strains, some of which show a relatively high divergence (Table 3). Taking into account that our PCR based screening likely underestimates the actual number of shared IS elements (due to mismatches at the primer binding sites in more diverged homologs), this indicates that many - if not most - *A. asiaticus* IS elements were already present in the last common ancestor of the *A. asiaticus* strains investigated here. Taken together, there is compelling evidence that the IS elements have been residing in the *A. asiaticus* genome for considerable evolutionary time periods.

We noted previously that the *A. asiaticus* genome shows a GC skew pattern typical for most bacterial

genomes with two major shifts at the origin and terminus of replication and only few local deviations, which are indicative of recent genome rearrangements [34]. This is remarkable because with the exception of *Lactobacillus helveticus* DPC 4571 and *Shigella sonnei* Ss046 (whose genomes contain significantly lower percentages of IS elements than *A. asiaticus*; Additional file 1, Figures S1, S2)[27,63], all other bacteria with high numbers of IS elements do not show such a regular genomic GC skew pattern (Additional file 1, Figure S11). Thus, despite of the high number of IS elements, the *A. asiaticus* genome has not been reshuffled extensively recently, which indicates that most IS elements are transpositionally inactive and also that recombination events between highly similar IS element copies have not occurred.

A mechanism by which apparently inactive, non-functional IS elements can be maintained in bacterial



genomes is gene conversion, which was described recently for the genome of *Wolbachia* wBm, a mutualistic symbiont of the nematode *Brugia malayi* whose genome contains a number of highly similar IS element copies rendered non-functional by multiple stop codons and frame shifts [64,65]. In contrast to *Wolbachia* wBm, *A. asiaticus* still encodes intact copies of each IS element, and many IS elements show relatively high sequence divergence (Table 2). In addition, the transposase genes of different non-functional IS element copies show variable pseudogenization states. This largely rules out gene conversion as the main mechanism for maintenance of IS elements in *A. asiaticus*.

Rather unexpectedly, we detected transcription of 13 *A. asiaticus* 5a2 IS elements during intracellular growth in amoebae (Figure 2). Generally, IS elements are among the lowest expressed genes due to their potentially detrimental effects on the host genome [16,61,66-70]. For several *A. asiaticus* IS elements (ISCaa2, ISCaa9, ISCaa10, ISCaa14, ISCaa15; data not shown) stable hairpin structures within the first 50 bp

of the IS elements are predicted, which might interfere with expression both at the transcriptional and the translational level, thus controlling the activity of these IS elements. In addition, evidence for programmed translational frameshifting, another regulatory mechanism, can be found in six *A. asiaticus* IS elements (Table 2). Translational frameshifting acts at the level of translation elongation between two consecutive (and partially overlapping) open reading frames where the ribosome slides one basepair up- or downstream at a so-called slippery site [16,45,66]. For several IS elements of *A. asiaticus*, the occurrence of frameshifting is supported by the presence of a canonical slippery site, of stimulatory secondary structures downstream of the slippery site and, most importantly, the merged amino acid sequences of the IS elements transposase ORFs show more significant Blast hits than the single ORFs alone (data not shown). In summary, transcription of several IS elements occurs in *A. asiaticus*, but there is evidence that many IS elements are tightly regulated both at the transcriptional and the translational level.

Southern hybridizations demonstrated the absence of major transposition events and genome rearrangements for *A. asiaticus* during a time period of 984 days (Figure 3). With an estimated generation time of *Acanthamoeba* sp. 5a2 infected with *A. asiaticus* of 19 h (data not shown), this time period corresponds to approximately 1200 generations of the *Acanthamoeba* host. Although the generation time of *A. asiaticus* is unknown, it must be shorter than that of its amoeba host (due to the high number of symbionts per amoeba cell [35,37]). The analyzed time period thus corresponds to considerably more than 1200 *A. asiaticus* generations. For *E. coli* and *Lactococcus lactis*, the first IS element-mediated genomic changes (insertions, deletions, duplications) occurred already after 400 to 500 generations [12-15]. This indicates that the time period monitored in our study should be sufficient to detect IS element-mediated genomic rearrangements. However, in contrast to our experiment, in these studies bacterial cultures were exposed to environmental stress conditions with respect to nutrient availability, temperature, or oxygen, facilitating adaptive changes. Although no genome rearrangements were observed for *A. asiaticus*, we cannot exclude the possibility that transposition events occurred in individual *A. asiaticus* cells which subsequently became not fixed at the population level and would thus be undetectable by our experimental approach. However, Southern blot is a highly sensitive method [71], and we have estimated that we should be able to monitor changes in Southern blot patterns in subpopulations consisting of only a few to a few hundred of amoeba host cells (Additional file 2, Table S2). Taking into account typical densities of *Acanthamoeba* sp. 5a2 infected with *A. asiaticus* during in vitro cultivation of  $10^5$  up to  $10^7$  cells/ml, the sensitivity of our assay should thus be sufficient to detect variations even in very small subpopulations. The IS elements in *A. asiaticus* are therefore most likely transpositional inactive. Their abundance is explained by transpositional activity in the evolutionary past of *A. asiaticus*, and while still being transcriptionally active, most IS elements are transpositionally inactive in extant *A. asiaticus*. In addition to a tight transcriptional and (post-)translational control there are several other conceivable explanations for this observation. For example, *A. asiaticus* might lack host factors required for transposition activity of IS elements although most of those are specific for certain IS elements; they act at different steps and cellular processes and their exact role in transposition is still largely unclear [16,66,72]. Alternatively, the small, reduced genome of *A. asiaticus*, which is highly adapted to the intracellular life style and optimized for host cell interactions, might not allow for major rearrangements as most transposition events would be deleterious rendering the cell nonviable.

One reason why some IS elements were retained in the *A. asiaticus* genome despite of the apparent lack of transpositional activity might be their influence on the transcription of downstream genes. Indeed, we could show contiguous transcripts of IS elements with their downstream genes at 9 out of 10 tested loci (Figure 4, Figure 5). In some cases, the distance between the IS element and the start codon of the downstream gene is too short to include known *Bacteroidetes* Shine-Dalgarno sequences, which are located at -33 and -7 bp relative to the transcription initiation site [73,74]. Expression of the respective downstream genes might thus depend on promoter sequences located within the upstream IS element (e.g. in the inverted repeats), a feature often found in IS elements [16,66], or on the endogenous promoter of the IS element. In other cases the distance between the analyzed IS elements and their downstream genes was larger (up to 300 bp). Similar polycistronic mRNAs starting from IS elements including downstream genes have been described recently for two IS elements in *Francisella tularensis* [17] and in *Mycobacterium tuberculosis* IS6110 [21]. It is striking that many of the genes whose transcription is affected by the presence of IS elements in *A. asiaticus* likely play an important role (Additional file 2, Table S3). For example, Aasi\_1844 is an uncharacterized membrane protein conserved among most *Bacteroidetes* and *Chlorobi*; Aasi\_1118 contains six TPR/SEL1 repeats, eukaryotic domains that can be involved in host cell interaction [75], and Aasi\_0380 is a ferritin homolog involved in iron storage. Furthermore, a genomic organization of IS elements and downstream genes similar to the loci analyzed in this study was found in 44 other regions on the *A. asiaticus* genome (data not shown), suggesting that contiguous transcripts between IS elements and downstream genes are even more widespread and represent a more general feature of *A. asiaticus*.

Genome reduction is an important process during the adaptation of bacteria to an obligate intracellular life style, and IS elements are considered to be important in this process [2,8]. The genome of *A. asiaticus* is only moderately reduced compared to other obligate intracellular bacteria [41,42]. Its genome size is with 1.9 Mbp notably larger than that of other related symbionts in the *Bacteroidetes* (0.2 to 1.1 Mbp), but smaller than those of free-living relatives (2.2 to 9.1 Mbp, Additional file 1, Figure S12). The genome of *A. asiaticus* thus represents a transitional stage in genome reduction. We argue that the IS elements in the *A. asiaticus* genome are evolutionary remnants. They have been present in the *A. asiaticus* genome for extended time periods and reflect the organism's evolutionary history. The IS elements proliferated and were important during the adaptation of *A. asiaticus* to the intracellular life style, but

they became increasingly redundant. The *A. asiaticus* genome thus represents a snapshot of a bacterial genome which was shaped by the activity of IS elements but whose IS elements are largely inactive and in the process of further degradation at the present stage.

## Conclusion

Analysis and characterization of the *A. asiaticus* IS elements provides evidence for an extremely IS element-rich genome, which seems to be evolutionary surprisingly stable - a feature not found in other IS element-rich genomes. The presence of contiguous transcripts between IS elements and their downstream genes indicates that these IS elements influence the transcription of their downstream genes, most of which likely play an important role for *A. asiaticus*. Proliferation of IS elements in the evolutionary past of *A. asiaticus* might thus have been an important process during the adaptation of *A. asiaticus* to an intracellular life style in which its genome was shaped by their activity.

## Methods

### Sequence analyses

The genome sequence of *A. asiaticus* 5a2 has recently been determined and analyzed [34] and is available at GenBank under accession no. CP001102. For identification of IS elements we first compiled a list of candidate transposase genes by keyword, PFAM and InterPro domain search available in the genome annotation software Pedant [76]. We then manually inspected this list in order to verify the evidence for each gene to encode a putative transposase. In addition, further transposase genes were identified by manually analyzing each predicted gene in the *A. asiaticus* genome (e.g by using Blast against the NCBI nr dataset (provided by the annotation software Pedant), Blast against the ISfinder database <http://www-is.biotoul.fr/>) In order to classify the transposases into groups of homologs we performed Blast (BlastP, BlastN) searches against the *A. asiaticus* genome. In order to identify full-length IS elements, the gene sequences of the transposases and surrounding genomic regions were aligned and the full-length IS elements were then manually identified based on these alignments. Partial IS element copies were identified by BlastN and BlastP searches and alignment of full-length IS element copies against the *A. asiaticus* genome. Inverted repeats were identified with the EMBOSS software palindrome and einverted [77]. Nucleic acid sequences of IS elements and amino acid sequences of transposase genes were aligned with MAFFT [78]; alignments were visualized using BOXSHADE [http://www.ch.embnet.org/software/BOX\\_form.html](http://www.ch.embnet.org/software/BOX_form.html). For detection of direct repeats the nucleic acid alignments of the IS elements and their genomic neighborhood were

searched manually. We grouped and classified IS elements using Blast against the ISfinder website <http://www-is.biotoul.fr/> [43] and the following criteria: (i) a minimum amino acid sequence identity of 30% of the transposase to described transposases, (ii) the presence of flanking inverted repeats (exception: IS elements belonging to family IS110 and IS200/605, which do not have flanking inverted repeats), and (iii) the presence of at least two copies in the genome. IS element copies that shared more than 80% nucleic and amino acid sequence identity over at least 98% of their length were considered isoforms. The nomenclature suggested by the ISFinder website was used for naming of IS elements <http://www-is.biotoul.fr/> [43]. mRNA secondary structures were predicted using the Mfold web server [79]. For calculations of phylogenetic relationships of the transposases from selected IS elements, the amino acid sequences of overlapping ORFs were merged resulting in a single peptide sequence (in the case of IS elements with predicted ribosomal frameshifting), aligned with MAFFT [78] and imported into ARB [80]. Phylogenetic trees were constructed with the Phylip maximum parsimony, distance matrix (Fitch), ProML (using the JTT amino acid replacement model) methods and the TREE-PUZZLE algorithm (using the VT model of amino acid substitution) [81,82] implemented in ARB. Maximum parsimony bootstrap analysis was performed with 1000 resamplings. A filter considering only those alignment positions that were conserved in at least 10% of all sequences (resulting in a total number of 274 and 228 alignment columns for ISCa3 and ISCa4, respectively) was used for all treeing calculations. For each IS element analyzed, the overall tree topology between the different treeing methods applied was consistent, thus only trees calculated using the TREE-PUZZLE algorithm are shown.

### Cultivation and isolation of amoebae

Amoebae harboring *A. asiaticus* 5a2 (ATCC no. PRA-228) and amoebae harboring *A. asiaticus* EIDS3 (ATCC no. PRA-221) were maintained as adherent culture in 25 cm<sup>2</sup> tissue culture flasks containing 10 ml peptone-yeast-glucose medium (PYG: 20 g/l proteose peptone, 2 g/l yeast extract, 90 mM glucose, 4 mM MgSO<sub>4</sub>\*7H<sub>2</sub>O, 3.4 mM C<sub>6</sub>H<sub>5</sub>Na<sub>3</sub>O<sub>7</sub>\*2H<sub>2</sub>O, 2.5 mM KH<sub>2</sub>PO<sub>4</sub>, 1.3 mM Na<sub>2</sub>HPO<sub>4</sub>\*2H<sub>2</sub>O, 51 μM Fe(NH<sub>4</sub>)<sub>2</sub>(SO<sub>4</sub>)<sub>2</sub>\*6H<sub>2</sub>O). Cultures were incubated at 27°C and passaged at confluency by 1:10 dilution of the culture every five to ten days. Amoebae harboring *A. asiaticus* WR and amoebae harboring *A. asiaticus* US1 were isolated from soil and lake sediment (Alkaline lake "Unterer Stinker", Burgenland, Austria) samples, respectively, using non-nutrient agar plates seeded with live or heat-inactivated *Escherichia coli* as described previously [83]. Both isolates were

cultivated as described above using modified PYNFH (10 g/l bacteriological peptone, 10 g/l yeast extract, 1 g/l yeast nucleic acid, 15 mg/l folic acid, 1 mg/l hemin, 2.6 mM  $\text{KH}_2\text{PO}_4$ , 2.8 mM  $\text{Na}_2\text{HPO}_4 \cdot 2\text{H}_2\text{O}$ ).

#### DNA isolation

Amoebae harboring *A. asiaticus* 5a2, EIDS3, WR and US1 were harvested by centrifugation ( $5000 \times g$ , 10 min). The cell pellet was resuspended in 250  $\mu\text{l}$  1 $\times$  TE buffer (10 mM Tris, 1 mM EDTA, pH 8) and subsequently used for high molecular weight DNA isolation using a modified protocol from Zhou et al. [84]. Briefly, 675  $\mu\text{l}$  DNA extraction buffer (100 mM Tris/HCl, 100 mM EDTA, 100 mM sodium-phosphate, 1.5 M NaCl, 1% (w/v) cetyltrimethylammonium bromide (CTAB), 200  $\mu\text{g}/\text{ml}$  proteinase K, pH 8.0) were added to the cell pellet and incubated for 30 min at 37°C. After addition of 75  $\mu\text{l}$  20% (w/v) SDS, the samples were incubated at 65°C for 1 h. To recover the aqueous phase, the lysate was mixed with an equal volume of chloroform/isoamylalcohol (24:1, v/v) and centrifuged ( $11200 \times g$ , 10 min). Nucleic acids were precipitated with 0.6 volume isopropanol at room temperature for 1 h. The resulting pellet from centrifugation ( $16000 \times g$ , 20 min) was washed with 70% ethanol, centrifuged again ( $16000 \times g$ , 5 min), resuspended in ddH<sub>2</sub>O and stored at -20°C until use.

#### Transcription analysis

Amoebae harboring *A. asiaticus* 5a2 were harvested by centrifugation ( $7000 \times g$ , 3 min, 27°C). The resulting cell pellet was resuspended in 750  $\mu\text{l}$  TRIzol (Invitrogen Life Technologies), transferred to a Lysing Matrix A tube (MP Biomedicals) and homogenized using a BIO101/Savant FastPrep FP120 instrument (speed: 4.5 m/sec, 30 sec). RNA was extracted by phase separation, precipitation, washing and redissolving according to the recommendations of the manufacturer (TRIzol, Invitrogen Life Technologies). Remaining DNA was removed using the TURBO DNA-free Kit (Ambion). After DNase treatment RNA was resuspended in ddH<sub>2</sub>O<sub>DEPC</sub> and stored at -80°C until use. The absence of DNA contamination in the DNase-treated RNA was verified by performing a control PCR with 42 cycles using primers targeting the 16S rRNA gene of *A. asiaticus* 5a2 (Additional file 2, Table S1). DNA-free total RNA (containing host and symbiont RNA) was used to synthesize cDNA using the RevertAid™ First Strand cDNA Synthesis Kit (Fermentas) according to the recommendations of the manufacturer. cDNA was subsequently used as template in standard PCR reactions (35 cycles and annealing temperatures according to the optimal conditions for the primers listed in Additional file 2, Table S1). Negative controls (no cDNA added) and positive controls (genomic DNA) were included in all PCR reactions.

Amplification products were sequenced to ensure that amplification was specific. All experiments were performed in biologically independent triplicates.

#### PCR screening for IS elements in different *A. asiaticus* strains

A standard PCR cycling program with 35 cycles at low stringency (annealing temperature 45°C) with primers specific for different *A. asiaticus* 5a2 IS elements was used for the detection of IS elements in the *A. asiaticus* strains EIDS3, WR and US1 (see Additional file 2, Table S1 for primer sequences). Negative (no DNA added) and positive controls (genomic DNA from *A. asiaticus* 5a2) were included in all PCR reactions. The amplified fragments from *A. asiaticus* EIDS3 were cloned using the TOPO TA cloning kit and cloning vector pCRII (Invitrogen Life Technologies). Nucleotide sequences of the cloned DNA fragments were determined on an ABI 3130 XL genetic analyzer using the BigDye Terminator kit v3.1 (Applied Biosystems).

#### Southern hybridizations

Southern hybridization was performed using a modified protocol based on Sambrook et al. [71]. Two  $\mu\text{g}$  DNA (containing host amoeba and *A. asiaticus* DNA) were digested with Eco32I for all investigated IS elements, except for IS<sub>Caa2</sub>, for which DNA was digested with HindIII and subsequently separated on a 0.7% TAE agarose gel (4°C, 17 h, 30 V). The gel was depurinated for 10 min in 0.25 M HCl, denatured for 30 min in 1.5 M NaCl/0.5 M NaOH and neutralized for 30 min in 1.5 M NaCl/1 M Tris-HCl (pH 7.5). Between each of these steps the gel was briefly rinsed in ddH<sub>2</sub>O. DNA was transferred onto Hybond N<sup>+</sup> nylon membranes (GE Healthcare) with a vacuum transfer system and 20 $\times$  SSC (3M NaCl, 0.3M sodium citrate, pH 7.0) as transfer buffer for 30 min. After immobilizing the DNA by UV cross-linking (120000  $\mu\text{J cm}^{-2}$ ), the membrane was briefly rinsed in ddH<sub>2</sub>O. Pre-hybridization was carried out for 2 h at 42°C in hybridization buffer (containing 50% formamide, 5 $\times$  SSC, 2% blocking reagent (Roche), 0.1% N-lauroyl sarcosyl sodium salt, 0.02% SDS (v/v)) in a rotation hybridization chamber as the following steps. The blot was hybridized with digoxigenin (DIG)-labeled probes (synthesized using the PCR DIG Probe Synthesis Kit, Roche; each probe was specific for a single IS element; see Additional file 2, Table S1) and hybridization buffer over night at 42°C. The membrane was washed twice for 15 min each with 2 $\times$  SSC/0.1% SDS at 25°C, and twice with 0.2 $\times$  SSC/0.1% SDS at 60°C for 15 min, followed by 2 min with DIG washing buffer (0.5 M maleic acid, 0.75 M NaCl, 0.3% Tween 20, pH 7.5) at 25°C, 30 min with buffer 2 (0.5 M maleic acid, 0.75 M NaCl, 0.3% Tween 20, 20% blocking reagent) at 25°C, 30

min with buffer 2 and Anti-Digoxigenin-AP Fab fragments (1:10000) at 25°C, twice for 15 min with DIG washing buffer at 25°C and finally for 5 min in 100 mM Tris/100 mM NaCl/50 mM MgCl<sub>2</sub> (pH 9.5) at 25°C. The membrane was swayed for 1 min in 1% CSPD solution (Roche) and subsequently exposed to Amersham Hyperfilm™ ECL (GE Healthcare).

#### Amplification of 16S and 18S rRNA genes

Oligonucleotide primers targeting 16S rRNA or 18S rRNA gene signature regions were used for PCR to obtain near full-length bacterial 16S rRNA or amoeba 18S rRNA gene fragments of the novel isolates *Acanthamoeba* sp. WR (containing *A. asiaticus* WR) and *Acanthamoeba* sp. US1 (containing *A. asiaticus* US1); see Additional file 2, Table S1. Nucleotide sequences of DNA fragments were determined on an ABI 3130 XL genetic analyzer using the BigDye Terminator kit v3.1 (Applied Biosystems).

#### Nucleotide sequence accession numbers

Obtained nucleotide sequences of IS elements of *A. asiaticus* EIDS3 and 16S and 18S rRNA genes of the isolates *Acanthamoeba* sp. WR (containing *A. asiaticus* WR) and *Acanthamoeba* sp. US1 (containing *A. asiaticus* US1) were submitted to EMBL/DDBJ/GenBank under accession numbers HM159367 to HM159370. The sequences of the *A. asiaticus* IS elements were deposited at EMBL/DDBJ/GenBank under accession numbers HM159371 to HM159380 and the ISFinder database [http://www-is.biotoul.fr/\[43\]](http://www-is.biotoul.fr/[43]).

#### Additional material

Additional file 1: pdf-file containing Figures S1 to S12.

Additional file 2: pdf-file containing Tables S1 to S32.

#### Acknowledgements

We are grateful to Daniela Teichmann and Christa Schleper (Department for Genetics in Ecology, University of Vienna) for providing protocols and helpful discussions. Technical assistance by Gabriele Schwammel, Diana Perez-Lopez and Christian Baranyi is greatly acknowledged. This work was supported by grants from the Austrian Science Fund (FWF) to SSE (grant no. P22703-B17) and MH (Y277-B03).

#### Author details

<sup>1</sup>Department of Microbial Ecology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria. <sup>2</sup>Department of Genetics in Ecology, University of Vienna, Althanstrasse 14, 1090 Vienna, Austria. <sup>3</sup>Institute for Milk Hygiene, University of Veterinary Medicine Vienna Veterinärplatz 1, 1210 Vienna, Austria.

#### Authors' contributions

SSE and MH designed the study. SSE, TP and AS performed sequence analyses; TP and AS carried out the molecular biology experiments. SSE and MH wrote the manuscript; all authors read, edited, and approved the final manuscript.

Received: 17 May 2011 Accepted: 26 September 2011

Published: 26 September 2011

#### References

1. Frost LS, Leplae R, Summers AO, Toussaint A: **Mobile genetic elements: the agents of open source evolution.** *Nat Rev Microbiol* 2005, **3**(9):722-732.
2. Siguier P, Filee J, Chandler M: **Insertion sequences in prokaryotic genomes.** *Curr Opin Microbiol* 2006, **9**(5):526-531.
3. Touchon M, Rocha EP: **Causes of insertion sequences abundance in prokaryotic genomes.** *Mol Biol Evol* 2007, **24**(4):969-981.
4. Wagner A, Lewis C, Bichsel M: **A survey of bacterial insertion sequences using IScan.** *Nucleic Acids Res* 2007, **35**(16):5284-5293.
5. Aziz RK, Breitbart M, Edwards RA: **Transposases are the most abundant, most ubiquitous genes in nature.** *Nucleic Acids Res* 2010, **38**(13):4207-4217.
6. Newton IL, Bordenstein SR: **Correlations Between Bacterial Ecology and Mobile DNA.** *Curr Microbiol* 2011, **62**(1):198-208.
7. Rocha EP: **Order and disorder in bacterial genomes.** *Curr Opin Microbiol* 2004, **7**(5):519-527.
8. Moran NA, Plague GR: **Genomic changes following host restriction in bacteria.** *Curr Opin Genet Dev* 2004, **14**(6):627-633.
9. Klasson L, Westberg J, Sapountzis P, Naslund K, Lutnaes Y, Darby AC, Veneti Z, Chen L, Braig HR, Garrett R, et al: **The mosaic genome structure of the *Wolbachia* wRi strain infecting *Drosophila simulans*.** *Proc Natl Acad Sci USA* 2009, **106**(14):5725-5730.
10. Tan HM: **Bacterial catabolic transposons.** *Appl Microbiol Biotechnol* 1999, **51**(1):1-12.
11. Wagner A: **Cooperation is fleeting in the world of transposable elements.** *PLoS Comput Biol* 2006, **2**(12):e162.
12. de Visser JA, Akkermans AD, Hoekstra RF, de Vos WM: **Insertion-sequence-mediated mutations isolated during adaptation to growth and starvation in *Lactococcus lactis*.** *Genetics* 2004, **168**(3):1145-1157.
13. Riehle MM, Bennett AF, Long AD: **Genetic architecture of thermal adaptation in *Escherichia coli*.** *Proc Natl Acad Sci USA* 2001, **98**(2):525-530.
14. Schneider D, Duperchy E, Coursange E, Lenski RE, Blot M: **Long-term experimental evolution in *Escherichia coli*. IX. Characterization of insertion sequence-mediated mutations and rearrangements.** *Genetics* 2000, **156**(2):477-488.
15. Treves DS, Manning S, Adams J: **Repeated evolution of an acetate-crossfeeding polymorphism in long-term populations of *Escherichia coli*.** *Mol Biol Evol* 1998, **15**(7):789-797.
16. Mahillon J, Chandler M: **Insertion sequences.** *Microbiol Mol Biol Rev* 1998, **62**(3):725-774.
17. Carlson PE Jr, Horzempa J, O'Dee DM, Robinson CM, Neophytou P, Labrinidis A, Nau GJ: **Global transcriptional response to spermine, a component of the intramacrophage environment, reveals regulation of *Francisella* gene expression through insertion sequence elements.** *J Bacteriol* 2009, **191**(22):6855-6864.
18. Ciampi MS, Schmid MB, Roth JR: **Transposon Tn10 provides a promoter for transcription of adjacent sequences.** *Proc Natl Acad Sci USA* 1982, **79**(16):5016-5020.
19. Kallastu A, Horak R, Kivisaar M: **Identification and characterization of IS1411, a new insertion sequence which causes transcriptional activation of the phenol degradation genes in *Pseudomonas putida*.** *J Bacteriol* 1998, **180**(20):5306-5312.
20. Lin H, Li TY, Xie MH, Zhang Y: **Characterization of the variants, flanking genes, and promoter activity of the *Leifsonia xyli* subsp. *cynodontis* insertion sequence IS1237.** *J Bacteriol* 2007, **189**(8):3217-3227.
21. Safi H, Barnes PF, Lakey DL, Shams H, Samten B, Vankayalapati R, Howard ST: **IS6110 functions as a mobile, monocyte-activated promoter in *Mycobacterium tuberculosis*.** *Mol Microbiol* 2004, **52**(4):999-1012.
22. Han HJ, Kuwae A, Abe A, Arakawa Y, Kamachi K: **Differential Expression of Type III Effector BteA Protein Due to IS481 Insertion in *Bordetella pertussis*.** *PLoS One* 2011, **6**(3):e17797.
23. Bordenstein SR, Reznikoff WS: **Mobile DNA in obligate intracellular bacteria.** *Nat Rev Microbiol* 2005, **3**(9):688-699.
24. Moya A, Pereto J, Gil R, Latorre A: **Learning how to live together: genomic insights into prokaryote-animal symbioses.** *Nat Rev Genet* 2008, **9**(3):218-229.
25. Song H, Hwang J, Yi H, Ulrich RL, Yu Y, Nierman WC, Kim HS: **The early stage of bacterial genome-reductive evolution in the host.** *PLoS Pathog* 2010, **6**(5):e1000922.

26. Plague GR, Dunbar HE, Tran PL, Moran NA: **Extensive proliferation of transposable elements in heritable bacterial symbionts.** *J Bacteriol* 2008, **190**(2):777-779.
27. Yang F, Yang J, Zhang X, Chen L, Jiang Y, Yan Y, Tang X, Wang J, Xiong Z, Dong J, *et al*: **Genome dynamics and diversity of *Shigella* species, the etiologic agents of bacillary dysentery.** *Nucleic Acids Res* 2005, **33**(19):6445-6458.
28. Cho NH, Kim HR, Lee JH, Kim SY, Kim J, Cha S, Kim SY, Darby AC, Fuxelius HH, Yin J, *et al*: **The *Orientia tsutsugamushi* genome reveals massive proliferation of conjugative type IV secretion system and host-cell interaction genes.** *Proc Natl Acad Sci USA* 2007, **104**(19):7981-7986.
29. Nakayama K, Yamashita A, Kurokawa K, Morimoto T, Ogawa M, Fukuhara M, Urakami H, Ohnishi M, Uchiyama I, Ogura Y, *et al*: **The Whole-genome sequencing of the obligate intracellular bacterium *Orientia tsutsugamushi* revealed massive gene amplification during reductive genome evolution.** *DNA Res* 2008, **15**(4):185-199.
30. Wu M, Sun LV, Vamathevan J, Riegler M, Deboy R, Brownlie JC, McGraw EA, Martin W, Esser C, Ahmadinejad N, *et al*: **Phylogenomics of the reproductive parasite *Wolbachia pipiensis* wMel: a streamlined genome overrun by mobile genetic elements.** *PLoS Biol* 2004, **2**(3):E69.
31. Klasson L, Walker T, Sebahia M, Sanders MJ, Quail MA, Lord A, Sanders S, Earl J, O'Neill SL, Thomson N, *et al*: **Genome Evolution of *Wolbachia* Strain wPip from the *Culex pipiens* Group.** *Mol Biol Evol* 2008, **25**(9):1877-1887.
32. Woyke T, Teeling H, Ivanova NN, Huntemann N, Richter M, Gloeckner FO, Boffelli D, Anderson IJ, Barry KW, Shapiro HJ, *et al*: **Symbiosis insights through metagenomic analysis of a microbial consortium.** *Nature* 2006, **443**(7114):950-955.
33. Gil R, Belda E, Gosalbes MJ, Delaye L, Vallier A, Vincent-Monegat C, Heddi A, Silva FJ, Moya A, Latorre A: **Massive presence of insertion sequences in the genome of SOPE, the primary endosymbiont of the rice weevil *Sitophilus oryzae*.** *Int Microbiol* 2008, **11**(1):41-48.
34. Schmitz-Esser S, Tischler P, Arnold R, Montanaro J, Wagner M, Rattei T, Horn M: **The genome of the amoeba symbiont "*Candidatus Amoebophilus asiaticus*" reveals common mechanisms for host cell interaction among amoeba-associated bacteria.** *J Bacteriol* 2010, **192**(4):1045-1057.
35. Schmitz-Esser S, Toenshoff ER, Haider S, Heinz E, Hoenninger VM, Wagner M, Horn M: **Diversity of bacterial endosymbionts of environmental *Acanthamoeba* isolates.** *Appl Environ Microbiol* 2008, **74**(18):5822-5831.
36. Choi SH, Cho MK, Ahn SC, Lee JE, Lee JS, Kim DH, Xuan YH, Hong YC, Kong HH, Chung DI, *et al*: **Endosymbionts of *Acanthamoeba* isolated from domestic tap water in Korea.** *Korean J Parasitol* 2009, **47**(4):337-344.
37. Horn M, Harzenetter MD, Linner T, Schmid EN, Muller KD, Michel R, Wagner M: **Members of the *Cytophaga-Flavobacterium-Bacteroides* phylum as intracellular bacteria of acanthamoebae: proposal of '*Candidatus Amoebophilus asiaticus*'.** *Environ Microbiol* 2001, **3**(7):440-449.
38. Xuan YH, Yu HS, Jeong HJ, Seol SY, Chung DI, Kong HH: **Molecular characterization of bacterial endosymbionts of *Acanthamoeba* isolates from infected corneas of Korean patients.** *Korean J Parasitol* 2007, **45**(1):1-9.
39. Zchori-Fein E, Perlman SJ, Kelly SE, Katzir N, Hunter MS: **Characterization of a '*Bacteroidetes*' symbiont in *Encarsia* wasps (Hymenoptera: Aphelinidae): proposal of '*Candidatus Cardinium hertigii*'.** *Int J Syst Evol Microbiol* 2004, **54**(Pt 3):961-968.
40. Sunagawa S, Woodley CM, Medina M: **Threatened corals provide underexplored microbial habitats.** *PLoS ONE* 2010, **5**(3):e9554.
41. Merhej V, Royer-Carenzi M, Pontarotti P, Raoult D: **Massive comparative genomic analysis reveals convergent evolution of specialized bacteria.** *Biol Direct* 2009, **4**:13.
42. Moran NA, McCutcheon JP, Nakabachi A: **Genomics and evolution of heritable bacterial symbionts.** *Annu Rev Genet* 2008, **42**:165-190.
43. Siguier P, Perochon J, Lestrade L, Mahillon J, Chandler M: **ISfinder: the reference centre for bacterial insertion sequences.** *Nucleic Acids Res* 2006, **34** Database: D32-36.
44. Ohta S, Tsuchida K, Choi S, Sekine Y, Shiga Y, Ohtsubo E: **Presence of a characteristic D-D-E motif in IS1 transposase.** *J Bacteriol* 2002, **184**(22):6146-6154.
45. Baranov PV, Fayet O, Hendrix RW, Atkins JF: **Recoding in bacteriophages and bacterial IS elements.** *Trends Genet* 2006, **22**(3):174-181.
46. Namy O, Rousset JP, Naphtine S, Brierley I: **Reprogrammed genetic decoding in cellular gene expression.** *Mol Cell* 2004, **13**(2):157-168.
47. Cordaux R: **ISWpi1 from *Wolbachia pipiensis* defines a novel group of insertion sequences within the IS5 family.** *Gene* 2008, **409**(1-2):20-27.
48. Cordaux R, Pichon S, Ling A, Perez P, Delaunay C, Vavre F, Bouchon D, Greve P: **Intense transpositional activity of insertion sequences in an ancient obligate endosymbiont.** *Mol Biol Evol* 2008, **25**(9):1889-1896.
49. Cridge AG, Major LL, Mahagaonkar AA, Poole ES, Isaksson LA, Tate WP: **Comparison of characteristics and function of translation termination signals between and within prokaryotic and eukaryotic organisms.** *Nucleic Acids Res* 2006, **34**(7):1959-1973.
50. Poole ES, Brown CM, Tate WP: **The identity of the base following the stop codon determines the efficiency of in vivo translational termination in *Escherichia coli*.** *Embo J* 1995, **14**(1):151-158.
51. Barabas O, Ronning DR, Guynet C, Hickman AB, Ton-Hoang B, Chandler M, Dyda F: **Mechanism of IS200/IS605 family DNA transposases: activation and transposon-directed target site selection.** *Cell* 2008, **132**(2):208-220.
52. Guynet C, Hickman AB, Barabas O, Dyda F, Chandler M, Ton-Hoang B: **In vitro reconstitution of a single-stranded transposition mechanism of IS608.** *Mol Cell* 2008, **29**(3):302-312.
53. Beuzon CR, Chessa D, Casadesus J: **IS200: an old and still bacterial transposon.** *Int Microbiol* 2004, **7**(1):3-12.
54. Martusewitsch E, Sensen CW, Schleper C: **High spontaneous mutation rate in the hyperthermophilic archaeon *Sulfolobus solfataricus* is mediated by transposable elements.** *J Bacteriol* 2000, **182**(9):2574-2581.
55. Ogata H, La Scola B, Audic S, Renesto P, Blanc G, Robert C, Fournier PE, Claverie JM, Raoult D: **Genome sequence of *Rickettsia bellii* illuminates the role of amoebae in gene exchanges between intracellular pathogens.** *PLoS Genet* 2006, **2**(5):e76.
56. Liu Y, Whitman WB: **Metabolic, phylogenetic, and ecological diversity of the methanogenic archaea.** *Ann N Y Acad Sci* 2008, **1125**:171-189.
57. Khan NA: ***Acanthamoeba*: biology and increasing importance in human health.** *FEMS Microbiol Rev* 2006, **30**(4):564-595.
58. Rodríguez-Zaragoza S: **Ecology of free living amoebae.** *Crit Rev Microbiol* 1994, **20**(3):225-241.
59. Wagner A, de la Chaux N: **Distant horizontal gene transfer is rare for multiple families of prokaryotic insertion sequences.** *Mol Genet Genomics* 2008, **280**(5):397-408.
60. Hooper SD, Mavromatis K, Kyrpides NC: **Microbial co-habitation and lateral gene transfer: what transposases can tell us.** *Genome Biol* 2009, **10**(4):R45.
61. Wagner A: **Periodic extinctions of transposable elements in bacterial lineages: evidence from intragenomic variation in multiple genomes.** *Mol Biol Evol* 2006, **23**(4):723-733.
62. Lawrence JG, Ochman H: **Amelioration of bacterial genomes: rates of change and exchange.** *J Mol Evol* 1997, **44**(4):383-397.
63. Callanan M, Kaleta P, O'Callaghan J, O'Sullivan O, Jordan K, McAuliffe O, Sangrador-Vegas A, Slattery L, Fitzgerald GF, Beresford T, *et al*: **Genome sequence of *Lactobacillus helveticus*, an organism distinguished by selective gene loss and insertion sequence element expansion.** *J Bacteriol* 2008, **190**(2):727-735.
64. Cordaux R: **Gene conversion maintains nonfunctional transposable elements in an obligate mutualistic endosymbiont.** *Mol Biol Evol* 2009, **26**(8):1679-1682.
65. Foster J, Ganatra M, Kamal I, Ware J, Makarova K, Ivanova N, Bhattacharyya A, Kapatral V, Kumar S, Posfai J, *et al*: **The *Wolbachia* genome of *Brugia malayi*: endosymbiont evolution within a human pathogenic nematode.** *PLoS Biol* 2005, **3**(4):e121.
66. Nagy Z, Chandler M: **Regulation of transposition in bacteria.** *Res Microbiol* 2004, **155**(5):387-398.
67. Kleckner N: **Regulating Tn10 and IS10 Transposition.** *Genetics* 1990, **124**(3):449-454.
68. Jager D, Sharma CM, Thomsen J, Ehlers C, Vogel J, Schmitz RA: **Deep sequencing analysis of the *Methanosarcina mazei* Go1 transcriptome in response to nitrogen availability.** *Proc Natl Acad Sci USA* 2009, **106**(51):21878-21882.
69. Perkins TT, Kingsley RA, Fookes MC, Gardner PP, James KD, Yu L, Assefa SA, He M, Croucher NJ, Pickard DJ, *et al*: **A strand-specific RNA-Seq analysis of the transcriptome of the typhoid bacillus *Salmonella typhi*.** *PLoS Genet* 2009, **5**(7):e1000569.

70. Wurtzel O, Sapra R, Chen F, Zhu Y, Simmons BA, Sorek R: **A single-base resolution map of an archaeal transcriptome.** *Genome Res* 2010, **20**(1):133-141.
71. Sambrook J, Russell DW: **Molecular cloning: a laboratory manual.** Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press,; 3 2001.
72. Twiss E, Coros AM, Tavakoli NP, Derbyshire KM: **Transposition is modulated by a diverse set of host factors in *Escherichia coli* and is stimulated by nutritional stress.** *Mol Microbiol* 2005, **57**(6):1593-1607.
73. Bayley DP, Rocha ER, Smith CJ: **Analysis of *cepA* and other *Bacteroides fragilis* genes reveals a unique promoter structure.** *FEMS Microbiol Lett* 2000, **193**(1):149-154.
74. Chen S, Bagdasarian M, Kaufman MG, Bates AK, Walker ED: **Mutational analysis of the *ompA* promoter from *Flavobacterium johnsoniae*.** *J Bacteriol* 2007, **189**(14):5108-5118.
75. Mittl PR, Schneider-Brachert W: **Sel1-like repeat proteins in signal transduction.** *Cell Signal* 2007, **19**(1):20-31.
76. Frishman D, Albermann K, Hani J, Heumann K, Metanomski A, Zollner A, Mewes HW: **Functional and structural genomics using PEDANT.** *Bioinformatics* 2001, **17**(1):44-57.
77. Rice P, Longden I, Bleasby A: **EMBOSS: the European Molecular Biology Open Software Suite.** *Trends Genet* 2000, **16**(6):276-277.
78. Katoh K, Toh H: **Recent developments in the MAFFT multiple sequence alignment program.** *Brief Bioinform* 2008, **9**(4):286-298.
79. Zuker M: **Mfold web server for nucleic acid folding and hybridization prediction.** *Nucleic Acids Res* 2003, **31**(13):3406-3415.
80. Ludwig W, Strunk O, Westram R, Richter L, Meier H, Buchner A, Lai T, Steppi S, Jobb G, *et al*: **ARB: a software environment for sequence data.** *Nucleic Acids Res* 2004, **32**(4):1363-1371.
81. Felsenstein J: **PHYLIP - Phylogeny inference package (version 3.2).** *Cladistics* 1989, **5**:164-166.
82. Schmidt HA, Strimmer K, Vingron M, von Haeseler A: **TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing.** *Bioinformatics* 2002, **18**(3):502-504.
83. Heinz E, Kolarov I, Kastner C, Toenshoff ER, Wagner M, Horn M: **An *Acanthamoeba* sp. containing two phylogenetically different bacterial endosymbionts.** *Environ Microbiol* 2007, **9**(6):1604-1609.
84. Zhou J, Bruns MA, Tiedje JM: **DNA recovery from soils of diverse composition.** *Appl Environ Microbiol* 1996, **62**(2):316-322.

doi:10.1186/1471-2148-11-270

**Cite this article as:** Schmitz-Esser *et al*: A bacterial genome in transition - an exceptional enrichment of IS elements but lack of evidence for recent transposition in the symbiont *Amoebophilus asiaticus*. *BMC Evolutionary Biology* 2011 **11**:270.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

