Article

CHARACTERIZATION OF TWO CLOSELY RELATED α-*AMYLASE* PARALOGS IN THE BARK BEETLE, *Ips typographus* (L.)

Ivana Viktorinova
Institute of Entomology of the Academy of Sciences, Czech Republic and Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany

Lucie Kucerova and Marta Bohmova
Institute of Entomology of the Academy of Sciences and Faculty of Natural Sciences of the University of South Bohemia, Czech Republic

Ian Henry
Max Planck Institute of Molecular Cell Biology and Genetics, Dresden, Germany

Marek Jindra, Petr Dolezal, Martina Zurovcova, and Michal Zurovec
Institute of Entomology of the Academy of Sciences and Faculty of Natural Sciences of the University of South Bohemia, Czech Republic

Ips typographus (L.), the eight-spined spruce bark beetle, causes severe damage throughout Eurasian spruce forests and suitable nuclear markers are needed in order to study its population structure on a genetic level. Two closely related genes encoding α-amylase in I. typographus were characterized and named AmyA and AmyB. Both α-amylase paralogs consisted of six exons and five introns. AmyA encodes a polypeptide of 483 amino acids, whereas AmyB has two alternative

Grant sponsor: Ministry of Education, Czech Republic; Grant numbers: 229518; MSM 60076605801; Grant sponsor: Entomology Institute Project; Grant number: Z50070508; Grant sponsor: European Community's Seventh Framework Programme (FP7/2007-2013); Grant number: 229518.
Additional Supporting Information may be found in the online version of this article.
Ivana Viktorinova and Lucie Kucerova contributed equally to this study.
Correspondence to: Michal Zurovec, BC AVCR, Branišovská 31, 370 05 České Budějovice, Czech Republic.
E-mail: zurovec@entu.cas.cz

ARCHIVES OF INSECT BIOCHEMISTRY AND PHYSIOLOGY, Vol. 77, No. 4, 179–198 (2011)
Published online in Wiley Online Library (wileyonlinelibrary.com).
 DOI: 10.1002/arch.20433

transcripts encoding polypeptides of 483 and 370 amino acids. The expression levels of both genes were high during larval stage and adulthood. The AmyB transcripts were absent in the pupal stage. A modification of the allozyme staining method allowed us to detect two clusters of bands on the electrophoretic gel that may correspond to the two α-amylase genes. There was a correlation between the lack of AmyB expression in pupa and the absence of the fast migrating isozyme cluster at this stage, suggesting that the faster migrating isoforms are products of the AmyB gene, whereas the slowly migrating bands are derived from the AmyA. © 2011 Wiley Periodicals, Inc.

Keywords: bark beetle; intron polymorphism; α-amylase; gene duplication; unusual splicing; isozymes

INTRODUCTION

The Eurasian spruce bark beetle *Ips typographus* (L.) (*Coleoptera*: *Curculionidae*) is the most destructive forest pest in Central Europe (Gregoire and Evans, 2004; Wermelinger, 2004). During normal conditions, the beetles feed on injured *Picea abies* trees and infest the starch-rich phloem under the bark (Fuehrer et al., 1991). However, during epidemic outbreaks, healthy trees can also succumb to massive beetle attack/infestation (Joensuu et al., 2008). For example, it was reported that an *I. typographus* outbreak killed more than 75% of the spruce trees in an area of about 3,500 ha at the "Bayerischer Wald" National Park in Germany (Netherer et al., 2002). The observed pattern of this pest's population dynamics appears relatively simple. Most of the time *I. typographus* populations are small and fragmented, but these relatively quiet durations are periodically interrupted by extreme outbreaks with damaging consequences. However, the conditions triggering these epidemics are rather complex and need to be studied at all biological levels, including the ecological, physiological and molecular biological.

One line of research was aimed at the population genetics level, with a goal of uncovering the genetic structure underlying *I. typographus* population dynamics. The first molecular data were received with the use of allozymes (Stauffer et al., 1992; Gruppe, 1997; Pavlicek et al., 1997). It was concluded that the populations are not separated, but the observed pattern was not entirely consistent for all the allozymes used, many of which showed various departures from the Hardy–Weinberg equilibrium. Later experiments with the employment of a mitochondrial-encoded marker, a portion of the cytochrome oxidase 1, indicated that there is a genetic subdivision among regional populations in Europe influenced by events, which took place during and after the last glaciation (Stauffer et al., 1999). However, mitochondrial markers were not variable enough to elucidate more recent processes. Further studies conducted with five microsatellite loci did not confirm the small isolated demes of *I. typographus* populations, but suggested the frequent exchange of migrants that effectively homogenize their genetic variability (Salle et al., 2003). Moreover, studies employing DNA markers did not elucidate the differences among allozymes, which were suggested to result from distinct selective forces acting on various metabolic pathways. Some of the enzymes, including α-amylase and esterase-2, displayed significant deficiencies of heterozygotes (Stauffer et al., 1992, 1999; Pavlicek

et al., 1997). The lack of geographic signal and the disruption of Hardy–Weinberg equilibrium was ascribed to a sign of selective forces, but this hypothesis requires further investigation.

The α-amylases play an important biochemical role in insect growth, as they catalyze the hydrolysis of internal 1,4-alpha-glycosidic linkages found in various glucose polymers. Its enzymatic activity was detected in the salivary glands and guts of various insect species and seems to have distinct regulation in both tissues (Grossman et al., 1997; Charlab et al., 1999; Jacobson and Schlein, 2001). The genes encoding α-*amylase* are often duplicated and may even form a multigene family comprising up to seven members in some species (Kikkawa, 1953; Ogita, 1968; Levy et al., 1985; Pope et al., 1986; Da Lage et al., 2002; Zhang et al., 2003; Sugino, 2007). Diverse patterns of genomic organization indicate that the amylase genes are highly adaptive (Hickey et al., 1987; Grossman et al., 1997). These genes usually display a considerably high level of polymorphism within many species. In *I. typographus*, α-amylase showed a high number of alleles assigned to a single α-amylase gene named *AmyI* (Stauffer et al., 1992). Interestingly, besides the well-stained isozyme bands on the electrophoretogram Staufer et al. (1992) occasionally observed some rare, faster bands that might not belong to the same locus.

The aim of this study was to find out whether α-amylase could be encoded by more than one locus in *I. typographus*. Here we describe the cloning and characterization of two closely related genes from the bark beetle. The putative proteins encoded by both paralogs display 90% amino acid identity and they are also very similar to other insect α-amylases. We also used a modified allozyme staining method and detected two clusters of bands that might belong to these two α-*amylase* loci. These results may considerably improve allozyme analysis, as well as bring novel DNA-based markers for the study of *I. typographus* population genetics.

MATERIALS AND METHODS

Experimental Insects

The bark beetles used in most experiments come from a temporary laboratory culture, which was established from individuals collected in the Šumava Mountains of the Czech Republic. Some experiments also involved beetles from the laboratory culture of Dr. Axel Gruppe (Munich, Germany). For isoenzyme analysis, beetles were dissected to check the presence of parasites.

Total RNA Isolation and cDNA Synthesis

The poly $(A)^+$ RNA was extracted from 10–15 beetles by using the Amersham Pharmacia Biotech QuickPrep Micro mRNA Purification Kit. SuperScript II Reverse Transcriptase was used for first strand cDNA synthesis with oligo(dT) primers following the manufacturer's protocol (Invitrogen, Carlsbad, CA). The central gene region was amplified with the degenerate A (forward) and B (reverse) primers (Table 1), which were designed according to the conserved region based on alignment of amino acid sequences from *Drosophila melanogaster* (CAA28238), *Drosophila virilis* (U02029), *Drosophila pseudoobscura* (X76240), *Tribolium castaneum* (Amy1—AAA03708 and Amy2—AAA03709) and *Tenebrio molitor* (P56634). The PCR reactions were performed using the Ex Taq DNA Polymerase (Takara Biochemicals, Osaka, Japan)

Table 1. List of Primers Used for Cloning of AmyA and AmyB cDNA and Genes

Primer name	*Direction*	*Sequence (5′→3′)*	*Position*
A	F	CATYTVTTCGARTGGAARTGG	103–124
B	R	TGGCRTTRACRTGAATRGC	3239–3258
C	F	TACTGTGACGTTATTTCCGGT	3123–3143
D	R	GCACTCGTTAGCAATATCCG	125–145
E	F	GTATCTAAACATTATGAAGTGG	3902–3923
F	R	CCNSWDATNACRTCRCARTA	7312–7327
G*	F	GGCCACGCGTCGACTAGTAC	Adaptor
H	R	TACTGTCGGATAGCTTTTGGTGTCC	1109–1133
I	R	ACCCCTCACGTATTCTTG	2194–2211
J	R	GCCTTCCACAGTAATAG	7253–7270
K	F	CTGTGGTTTCGGGCACTG	7167–7184
L	F	TTGACAACAGAAGTGGAG	908–925
M	R	ATCAACCAAGCTACCGGAAATAAC	3131–3154
N	F	AAATCGCCATACTTGTAACCACTC	3783–3806
O	R	GTAAGCTAGTTTCAATATCAG	7281–7301
P	F	GAAATTAAACACCATGAA	1–18
Q	R	TCTAATGGAAGCAGCATCGG	1180–1199
R	F	GGAGTGGCAGGATTTCGTGTAGA	6088–6110
S	R	CAGGCCGTAGAGTGTTGA	7581–7599
Seq1	F	CAAGAAAGCATATGCTGGTGT	160–180
Seq2:	R	TCAGACATTATTGTAAGGCAGTG	363–385
Seq3:	F	GTGTTTCCAATAGGATACAAC	603–623
Seq4	F	CAAAGTGAGTTGCCTGGAAG	1408–1428
Seq5	F	CAGAGTCTGCGAGTTCAAATAC	2523–2544
Seq6	R	CTCACGTATTCTTGACTTTGA	2187–2207
Seq7	F	AAAAATGTGTTTTAGATTTCTCCC	5484–5507
Seq8	R	GTTACTGTTTAACTATTCCCAG	3494–3515
Seq9	R	GTGGATACTCGCCGCTGC	3919–3937
Seq10	F	CACTCGATTAGATAAAACAACTG	3902–3824
Seq11	R	CTTTGGTCTAAGTCCTTCAGAGC	6009–6032
Seq12	F	ATTCCAGTCCGTATTTGA	6726–6744
Seq13	F	TTCGGTTCTTCATTATGCTTTTAT	4251–4274
Seq14	R	GCAATTTTTATTTTTGTCTAGTGG	5160–5183
Seq15	R	CACTTTAGGATTTTCACTTGGG	5507–5528
AmyAF	F	TGCGAGTTCAAATACGGACTAA	2530–2551
AmyAR	R	CTCATATTCTTTATAGTTCATGTAGGATG	2681–2709
AmyBF	F	TGTGAATTCAAATACGGACTGG	6720–6741
AmyBR	R	CATATTCTTTGTAGTTCATGTAAGACTCA	6869–6897
AmyB2F	F	TGTATACAATACTGACTACGTTGGAAG	6682–6707
AmyB2R	R	ATTTGTTAGTTCAGTGCCCGA	7174–7195

The primers are shown together with their orientation and location in the final contig. The numbering starts from the first known base of *AmyA*, in the *AmyA-AmyB* contig (see Fig. S1). The primers A–S were used for cloning of *AmyA* and *AmyB* cDNA and genomic fragments (see Fig. 1); primers Seq1–15 were used for sequencing, primers AmyAF, AmyAR, AmyBF, AmyBR, AmyB2F, AmyB2R were used for quantitative PCR. Primer G was an adaptor used for 5′ RACE, AmyBF was an intron-spanning primer with the first 25 nucleotides derived from the exon 5 and the last two nucleotides complementary to the exon 6.

and the following reaction conditions: 94°C for 1 min, followed by 40 cycles denaturation at 94°C for 30 sec, 45°C annealing for 30 sec and extension at 72°C for 1 min, with a final extension step at 72°C for 10 min. The 5′-Rapid amplification of cDNA end was performed using the *5′ RACE* System, version 2.0, from Invitrogen (Carlsbad, CA). After reverse transcription, the cDNA was 5′ tailed with a dCTP

oligomer using an enzyme terminal transferase. The PCR was performed using deoxyinosine-containing anchor primer G; 5′-GGCCACGCGTCGACTAGTACGG-GIIGGGIIGGGIIG-3′ that hybridized to the oligo(dC) tail, adaptor primer G (GGCCACGCGTCGACTAGTAC) and two nested gene-specific reverse primers (primers H and I, Table 1).

Isolation of Genomic DNA and PCR Amplification

Genomic DNA from *I. typographus* adults was isolated by using the Qiagen (Hilden, Germany) genomic DNA purification kit. The sequences were analyzed using DNASTAR software (Lasergene, Madison, WI). To amplify the terminal parts of *AmyA* and *AmyB* genes, the iPCR protocol according to Ochman et al. (1988) was employed (Fig. 1, Steps 2 and 5). The primer combinations used were: C+D and J+K (Table 1). The amplification of central parts of *AmyA* and *AmyB* was performed using four pairs of specific primers L+M, N+O, P+Q and R+S (Fig. 1, steps 6 and 7). The primer sequences are listed in Table 1.

DNA Cloning and Sequencing

The PCR products were subcloned into the pGEM-T Easy Vector (Promega, Madison, WI) and sequenced. The Beckman Coulter sequencer with the CEQ 2000 Dye Terminator Cycle Sequencing Kit and the ABI Prism 310 sequencer with the AbiPrism

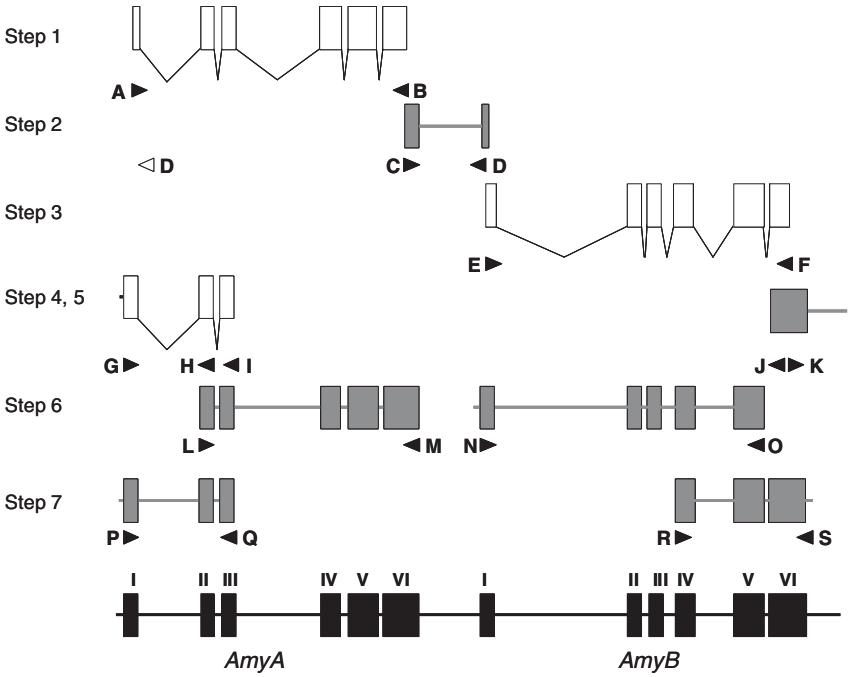

Figure 1. Schematic drawing of the step-by-step procedure for cloning *AmyA* and *AmyB* from *I.typographus*. Open boxes depict cDNA, gray boxes and lines represent genomic DNA and black boxes and lines at the bottom show the combined genomic organization of the two α-*amylase* genes. Filled triangles represent primers (see Table 1 for primer sequences). Roman numerals indicate exon numbers. Cloning steps 1 and 3 contain RT-PCR, steps 2 and 5 inverse PCR, step 4 indicates 5′ RACE PCR and steps 6 and 7 genomic PCR. The primer sequences are shown in Table 1.

BigDyeTerminator Sequencing Kit with AmpliTaq DNA polymerase (Perkin Elmer, Waltham, MA) were used for sequencing.

Real-Time Quantitative Reverse-Transcriptase Polymerase Chain Reaction (Real-Time qRT-PCR)

RNA was isolated by using RNA Blue (Top-Bio, Prague, Czech Republic), cleaned with RNAeasy Mini Kit (Qiagen, Hilden, Germany) and treated with RQ1 RNase-Free DNase (Promega, Madison, WI). 1,000 ng of total RNA was applied for reverse transcription using the SuperScript II Reverse Transcriptase (Invitrogen, Carlsbad, CA). The obtained cDNA was diluted and used for the Syber Green reaction in triplets: 5 ng (form A) or 50 ng (form B1 and B2) of RNA per 20 μl reaction, using a Rotor Gene 3000 (Corbett Research, Sydney, Australia) PCR cycler. The *AmyA* reaction ran for 40 cycles, *AmyB1* and *AmyB2* 50 cycles. Statistical analyses were performed directly in the Rotor-Gene 6 program. The quantitative RT-PCR (qRT-PCR) primers for the amplification of the three transcripts were designed using the PrimerSelect 8.0.2. (Lasergene 8, Madison, WI) and are shown in Table 1.

Cytosolic actin control primers were designed according to *Ips paraconfusus* (DQ471873) and *T. castaneum* actin mRNA sequences (XM_962212, XM_970991), forward: 5′-AACAGGGAAAAGATGACTCAAAT-3′, reverse: 5′-TTCGGTTAAGATT-TTCATCAAGTA-3′. Subsequent sequencing of the obtained PCR fragments for all *AmyA*, *AmyB1* and *AmyB2* proved their specificity.

Allozyme Staining

The protocol according to (Siciliano and Shaw, 1976; Stauffer et al., 1992) was used. Individual beetles were homogenized in extraction buffer (0.05 M Tris HCl, pH 7.0; 20% glycerol, 0.05% Triton X-100; 0.01% bromphenol blue) and samples were separated by electrophoresis in native 8% starch-polyacrylamide gels. To identify the weak bands after a short (20–40 sec) potassium iodide staining, the gels were incubated overnight in distilled water with shaking. Fresh or frozen samples kept up to 6 months at −80°C were used to avoid degradation of the α-amylase proteins.

cDNA and Protein Sequence Alignment, Phylogenetic Analysis

The evolutionary history of the α-*amylase* genes was inferred using the Neighbor-Joining method with 5,000 bootstrap replicates. The tree is drawn to scale, with branch lengths reflecting the evolutionary distances. The units of evolutionary distance were computed using the Jukes–Cantor model and refer to the number of base substitutions per site. All positions containing alignment gaps and missing data were eliminated only in pairwise sequence comparisons (Pairwise deletion option). Phylogenetic analyses were conducted using the MEGA4 program (Tamura et al., 2007).

The Accession Numbers for Protein and cDNA Sequences

Sequences used for the phylogenetic tree were: *Anthonomus grandis* Amy1 (AF527876) and Amy2 (AF527877), *Sitophilus oryzae* Amy (HQ158012), *Callosobruchus chinensis* Amy (AB110483), *Zabrotes subfasciatus* Amy (AF255722), *Phaedon cochleariae* Amy (Y17902), *Diabrotica virgifera* Amy (AF208002) and Amy2 (AF208003), *Blaps mucronata* Amy1 (AF462603) and incomplete Amy2 (AF468013), *T. molitor* (P56634), *T. castaneum* Amy1 (NM 001114376.1), Amy2 (XM 970392.1) and other identified copies XM 964141.2

and XM 964287.2 (only functional α-amylase genes from *Tribolium* were included), *Bombyx mori* Amy (NM_001173153) and *Blattella germanica* Amy (DQ355516).

Statistical Analyses and DNA Software

Statistical analyses of qRT-PCR were done directly in the Rotor-Gene 6 program (Corbett Research, Sydney, Australia). The double-sided Student's *t*-test with 95% confidence interval was used for the comparison of α-*amylase* expression levels. The sequences were analyzed using DNASTAR software (Lasergene, Madison, WI).

RESULTS

Cloning of α-Amylase Genes in I. typographus

To isolate a part of *I. typographus* α-amylase cDNA, regions with conserved amino acids were identified in the alignment of amino acid sequences (data not shown) of several insect α-amylases and used to design degenerate primers for PCR (A, B, Fig. 1 and Table 1). Total RNA was isolated from adult beetles and used as a template for reverse transcription; PCR amplification using the resulting cDNA was performed with the two degenerate primers. A single PCR fragment of expected size (1.35 kb) was obtained and subcloned (Fig. 1, step 1). The resulting sequence was used as a query for a similarity search using the NCBI tblastx program (Altschul et al., 1990). Significant similarity was observed with insect, especially beetle, α-*amylase* genes, indicating that the α-*amylase* homolog of *I. typographus* was isolated. The alignment of our sequence with the α-*amylase* coding sequence of the closely related Coleopteran *T. castaneum* suggested that the *I. typographus* cDNA lacked approximately 96 bp from the 5′-end and 7 bp from the 3′-end.

Our first attempts to obtain the full-length α-*amylase* by 5′ and 3′ RACE PCR failed. We therefore isolated genomic DNA and attempted inverse PCR (iPCR) using reverse direction primers derived from both ends of the original cDNA fragment. Surprisingly, the inverse primer pair of C and D (Table 1 and Fig. 1, step 2) produced a single 0.9 kb band even in the control reaction containing genomic DNA without restriction digestion and ligation. Subcloning and sequencing of the PCR product revealed a sequence that contained 144 bp from 3′ end of the putative α-*amylase* open reading frame (ORF), followed by 648 bp of a noncoding region and 129 bp sequence from the 5′ end of a second, unknown gene, most probably also encoding α-amylase (Fig. 1, step 2). We therefore concluded that there are two closely related α-amylase genes in *I. typographus*, which reside in tandem in the genome. We named the first detected α-amylase gene *AmyA*, whereas the second, newly detected gene was called *AmyB*.

The missing 5′ end of the *AmyA* cDNA was amplified by using 5′- RACE. Reverse transcription was followed by attachment of an oligo(dC) chain to the 5′ cDNA end by an enzyme terminal transferase and PCR with two novel reverse *AmyA*-specific primers (C and D, Table 1) together with an adapter–primer complementary to the dC-tailed cDNA (Fig. 1, steps 4 and 5). A single 5′ RACE product of 430 bp was cloned and sequenced. Sequencing revealed that the fragment contained 300 bp of the known DNA from the *AmyA* 5′ end together with 120 bp of flanking sequence containing the putative initiation codon ATG. The resulting combined nucleotide sequence of *AmyA* spanned the entire ORF region including the start and stop codons (Fig. 2).

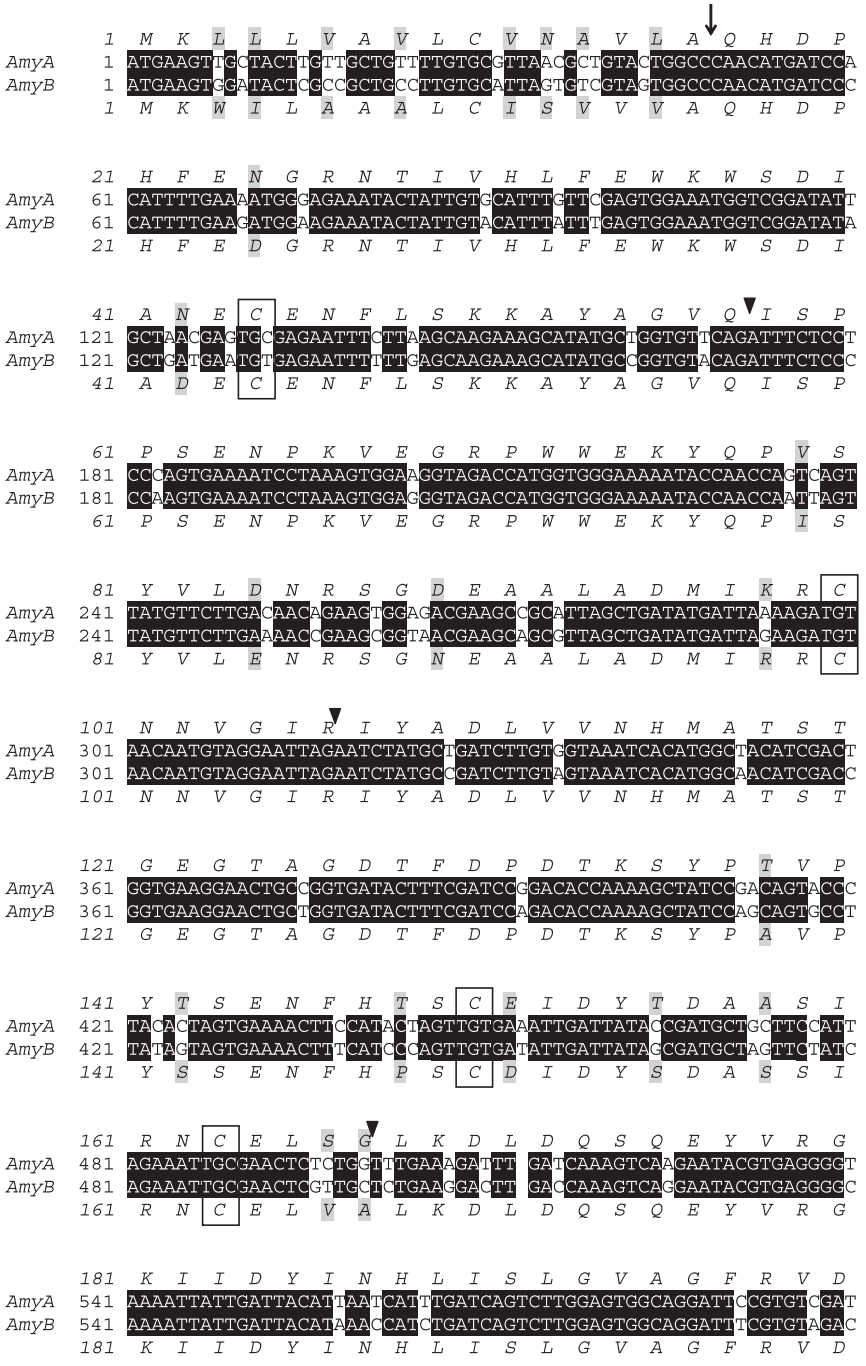

Figure 2. Comparison of cDNA nucleotide and deduced amino acid sequences of *AmyA* and *AmyB*. Nucleotides are represented by regular letters and deduced amino acid by italics. The upper two lines in the sequence alignment show *AmyA* and lower two lines *AmyB*. Identical nucleotides (cDNA) are shaded in black, amino acids differences are shaded in gray. The *AmyB* region spliced-out in *AmyB2* transcript is underlined. The positions of introns 1–5 are indicated by arrow-heads: (171–172 bp, 317–318 bp, 499–500 bp, 721–722 bp, 1042–1043 bp, respectively). The arrow indicates predicted signal peptide cleavage site. Conserved cysteine residues are boxed.

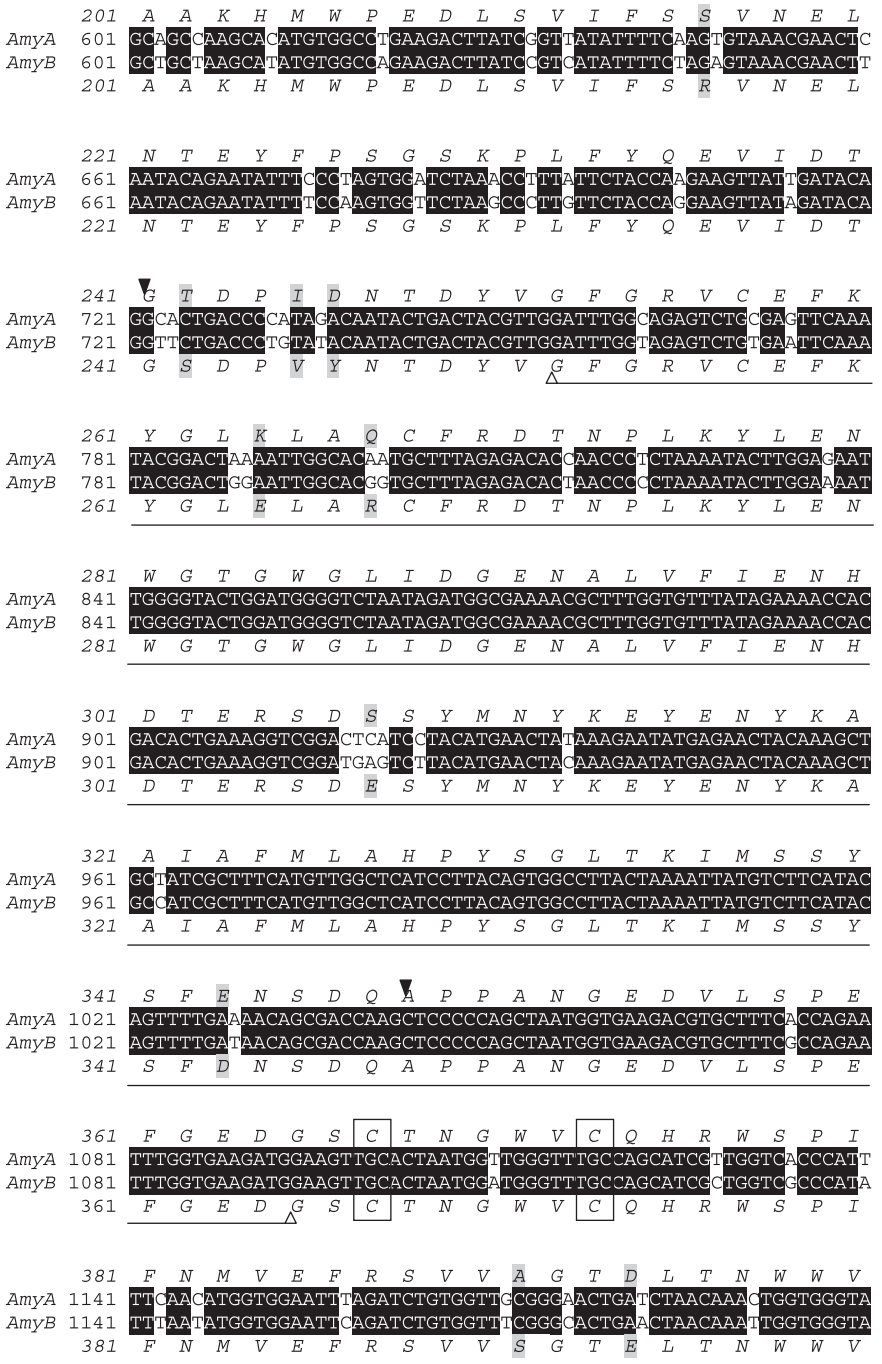

Figure 2. Continued.

We designed a pair of novel primers from the ends of *AmyA* ORF and PCR amplified the entire coding region of 1,452 nucleotides.

In order to obtain a larger piece of the *AmyB* cDNA, we used an *AmyB*-specific primer (E, Table 1) from the sequenced 5′ end of *AmyB* together with the 3′ degenerate

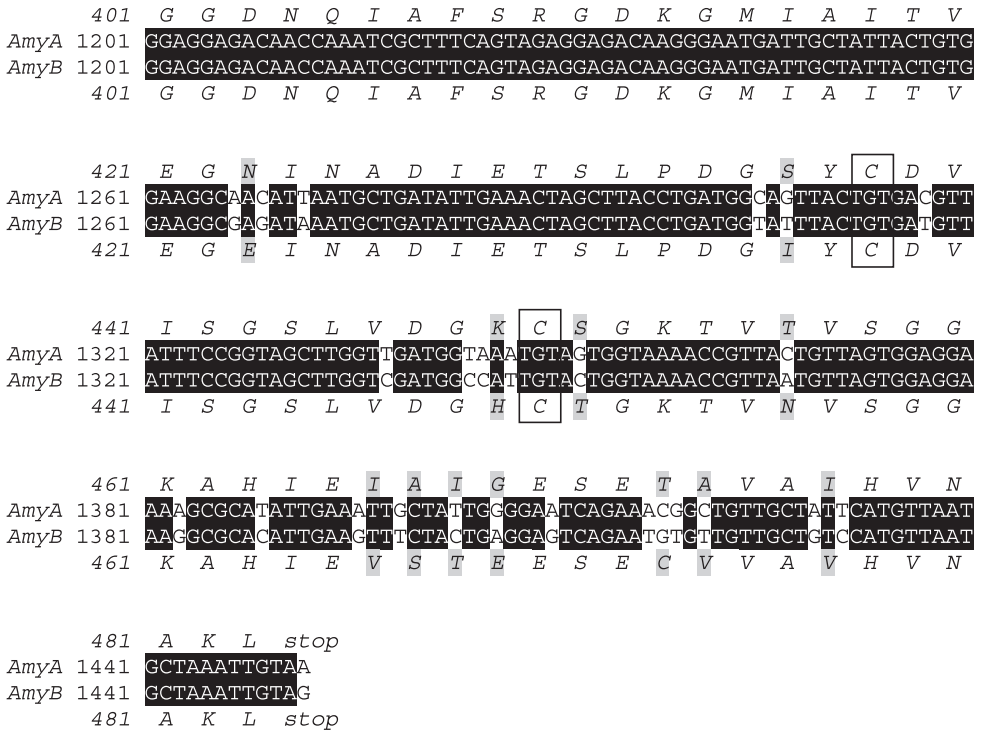

Figure 2. Continued.

primer (F, Table 1). RT PCR yielded a 1.30 kb fragment of *AmyB* (Fig. 1, step 3), containing most of the coding region and lacking the 3′ end. The missing 3′ end of the *AmyB* coding region was also acquired by iPCR using genomic DNA and primers (E and F, Table 1, Fig. 1, step 5). The resulting 0.5 kb genomic fragment was subcloned and sequenced (Fig. 1, steps 4 and 5). Sequence analysis confirmed that it contained the 3′ end of *AmyB* gene. A pair of novel primers was then designed from both ends of *AmyB* coding region and used for the amplification of the complete transcript sequence using a cDNA template. Two products corresponding to two different splicing forms *AmyB1 and AmyB2* were obtained (Fig. 2) and the total length of the ORFs were 1,452 and 1,113 nucleotides, respectively.

Eight novel primers (L-S) were designed based upon the cDNAs and genomic DNA described above (Table 1). The primers amplified a series of four overlapping genomic fragments, including 1.2 and 2.2 kb PCR products of *AmyA* and 3.0 and 1.5 kb fragments of *AmyB* (Fig. 1, Steps 6 and 7). The fragments were subcloned and sequencing was completed with the aid of additional 15 primers (see Table 1). This result confirmed that the PCR fragments covered the entire genomic regions of the *AmyA* and *AmyB* genes, linked by the intergenic sequence described above. The resulting composite sequence is 7851 bp long and it was deposited in GenBank, accession number: HQ417115 (Fig. S1).

Phylogenetic Analysis

A phylogenetic tree was constructed based on the complete cDNA sequence of *AmyA* and *AmyB* and included all known beetle α-*amylase* sequences (Fig. 3 and Experimental

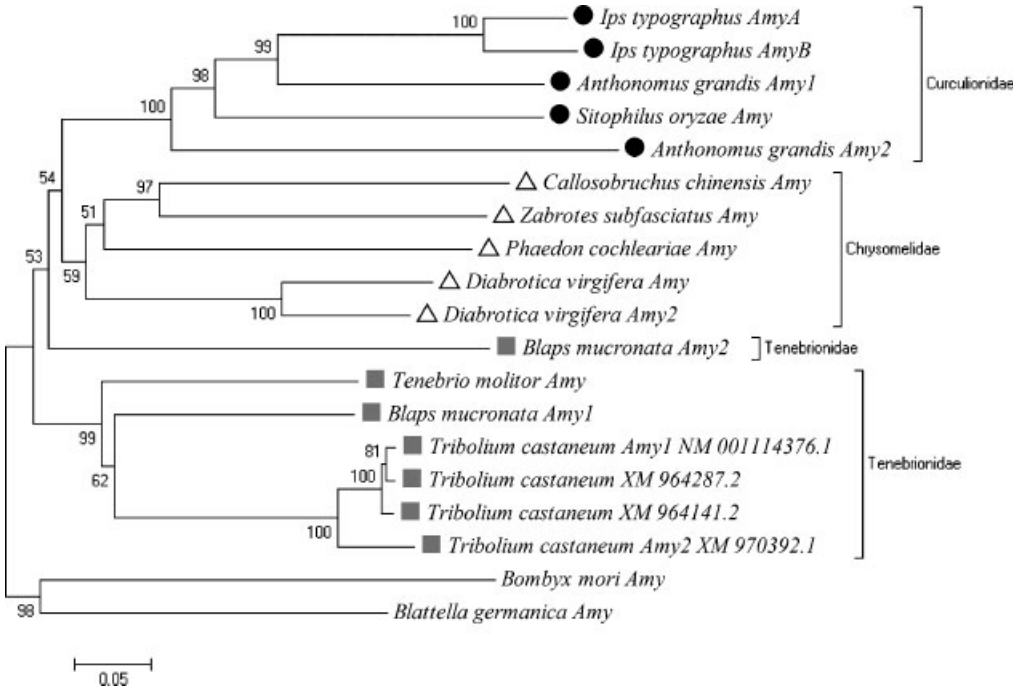

Figure 3. Phylogenetic relationship of beetle α-*amylases*. A phylogenetic tree was generated based on the cDNA sequences of indicated species using the Neighbor-Joining method with evolutionary distances based on the Jukes–Cantor model. Bootstrap values (5,000 repeats) above 50% are shown. Both *AmyA* and *AmyB* of *I. typographus* clustered within the family *Curculionidae* in the beetle group as expected. The tree was rooted with the α-*amylase* of *Bombyx mori* (Lepidoptera) and *Blattella germanica* (Orthoptera). Markers indicate beetle families: *Curculionidae* (black circles), *Chrysomelidae* (white triangles) and *Tenebrionidae* (gray squares). The branch length scale bar represents 0.05 nucleotide substitutions per site.

procedures). The α-*amylase* cDNA of *B. mori* (Lepidoptera) and *B. germanica* (Orthoptera) were used as outgroups to root the phylogeny tree. The results confirmed close relationship of both α-*amylase* paralogs of *I. typographus* to α-*amylases* from *A. grandis* and *S. oryzae*, which belong to the same family *Curculionidae*. It strongly suggests that both sequences represent *I. typographus* α-*amylase* genes. It is also obvious from Figure 3 that most of the species contain more than one α-*amylase* gene and even more paralogs might still exist in some beetles, as only the genome of *T. castaneum* was fully sequenced. The pairs of α-*amylase* paralogs within the species were usually closer to each other than to their putative orthologs from other beetles. As the genes clustered according to species, it suggested that the gene duplications occurred relatively recently. An exception was found for *B. mucronata* and *A. grandis* α-*amylase* paralogs, which were more divergent (Fig. 3).

The Structure of I. typographus α-Amylase Genes

To elucidate the exon–intron structure of both *I. typographus* α-*amylase* genes, the complete genomic sequences were aligned with cDNAs. Six exons and five introns were found in each α-*amylase* paralog (Fig. 1 and S1). The lengths of all exons were perfectly conserved between *AmyA* and *AmyB* and spanned 171, 146, 182, 222, 321 and 410 nucleotides. Also intron positions were perfectly conserved in both paralogs. The

intron sizes and sequences were not conserved, however, except for the short introns 2 and 5. The intron sizes were in the case of *AmyA*: 648, 55, 955, 90 and 53 bp, and in the case of *AmyB*: 1413, 54, 130, 440 and 53 bp, respectively (Fig. S1).

Interestingly, the mRNA of *AmyB* had also a truncated alternative splice form named *AmyB2* (Fig. 2), which was less abundant. It had the same 3′ and 5′ ends as the major *AmyB1* mRNA variant, but differed in sizes of exons five and six. Most of exon five (encoding 95 amino acids) and part of exon six (encoding 17 amino acids) were spliced out without changing the ORF of the sixth exon. The alternative intron 5 was 389 bp long. As it did not contain the canonical GT/AG ends and *AmyB2* alternative splicing site contained a duplication of the TGGA sequence, the determination of the exact exon/intron boundaries was ambiguous. There are five possible intron ends TT/GA, AT/GG, GA/TG, GG/AT and TG/GA (Fig. 4).

The comparison between the ORFs of both α-*amylase* paralogs revealed 104 synonymous and 55 nonsynonymous substitutions (Fig. 2). The GC content of the *AmyA* coding region was 41.1%, (the third codon position contained 20.9% C and 34.3% G or C) and 41.8% for *AmyB* (20.2% C and 35.5% G or C in the third codon position). The GC content of these two genes was lower than that of *Drosophila* α-*amylase*, where it was reported to be 62.8%, resulting in a high (60%) number of codons with a C in the third position (Boer and Hickey, 1986). The observed lower GC content might reflect a generally lower GC content in the *I. typographus* genome. For example, the GC content in the *Tribolium* genome is 34%, whereas in *Drosophila* it approaches 41% (Wang and Leung, 2009).

The first and third introns of *I. typographus AmyA* and the first and fourth introns of *AmyB* showed high polymorphism among several individuals from two regions

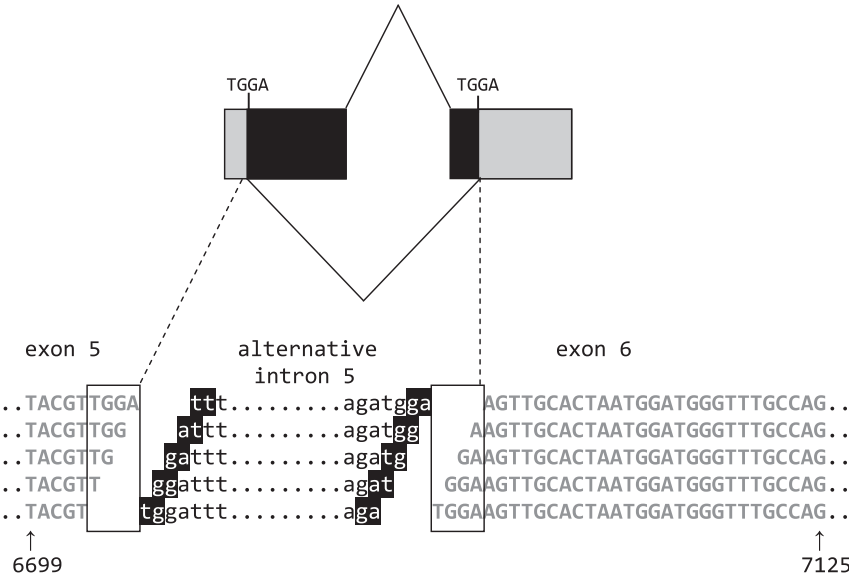

Figure 4. Possible ends of *AmyB* alternative intron 5**.** The TGGA duplication at the splicing site (boxed) and noncanonical intron ends make the correct identification of exon–intron boundaries ambiguous. Five possible 5′ and 3′ intron ends are shown in black shading. The *AmyB2* cDNA sequence is shown in capital letters, the alternative intron in lowercase letters. Numbers below the sequences refer to the corresponding positions in the *AmyA-AmyB* contig.

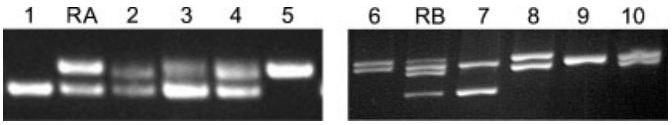

Figure 5. Polymorphism in *AmyA* introns. Electrophoretogram showing polymorphism in the *AmyA* intron. 1 (lanes 1–5 and RA) and intron 3 (lanes 6–10 and RB). RA and RB are reference samples containing mixture of several beetles) and samples 1–10 are from individual beetles. All beetles were from Šumava mountains, except for these in lines 5 and 10 that were from German population. The band sizes were: (1) 647 bp, (2–4) 730/647 bp and (5) 730 bp. (6) 955/910 bp, (RB) reference with intron length of 955, 910, 860 and 760 bp, (7) 910/760 bp, (8) 955/860 bp, (9) 910/910 bp, and (10) 910/860 bp. The primers for intron 1 (*AmyA*) amplification were as follows: 1AF: (5′-ATA TTG CTA ACG AGT GCG AG-3′), 1AR (5′-TTC GTC TCC ACT TCT GTT GT-3′). Intron 3 was amplified by 3AF (5′-CGA CAG TAC CCT ACA CTA G-3′) and 3AR (5′-GCA TCG ACA CGG AAT CCT G-3′).

(Figs. 5 and 6). Two alleles for *AmyA* intron 1 (648 and 730 bp) and four alleles (included polymorphic bands of 955, 910, 860 and 760 bp) for the *AmyA* intron 3 were observed in the electrophoretograms of PCR amplification products (Fig. 5). Three alleles were found for the *AmyB* intron one and sequenced (1145, 1413 and 1469 bp). Their alignment is shown in Figure 6.

A short sequence of 648 bp was identified between the *I. typographus AmyA* and *AmyB* coding regions and consisted of the untranslated 3′ end of the former, intergenic region together with the promoter and 5′ leader of the latter (Fig. 1 and S3). Similarly, a short fragment of 1.2 kb also separates the α-*amylase* genes in *T. castaneum* (GenBank accession number UO4271). The 3′ UTR of *AmyB* contained a consensus AATAAA polyadenylation signal at position 49 bp after the stop codon. The 5′ UTR of the *AmyB* displayed several putative eukaryotic promoter elements that were appropriately spaced: CAAT box motifs, a canonical TATA box, a binding site of the general transcription factor, TFIID, positioned at −53 bp from the start of the ORF (Fig. S3). The *AmyB* CAAT box with the ACTGCAATAA motif at position −90 likely corresponds to the *Drosophila* consensus A[A/T]GCA[A/T]N[A/T]N motif (O'Connell and Rosbash, 1984). Four copies of the GATAA motif in the *AmyB* promoter might correspond to GATAAG motif, which is supposed to be a midgut-specific regulator in the *Drosophila* genus (Magoulas et al., 1993).

The mRNA Expression of α-Amylase Paralogs

To determine in which developmental stages the α-*amylase* paralogs are expressed and measure the relative abundance of each transcript, real-time qRT-PCR was performed on RNA isolated from larvae, pupae, adults and the separated adult heads and abdomens of *I. typographus*. Cytosolic actin mRNA, previously used in the *Ips* genus as a stable endogenous control for transcription experiments (Keeling et al., 2006; Sandstrom et al., 2006), was used as a normalizer to calculate the relative mRNA abundance. As shown in Figure 7, *AmyA* was expressed at a significantly higher level than *AmyB1* in imagoes ($P<0.01$), larvae ($P<0.05$) and adult abdomens ($P<0.05$).

In addition, both *AmyB1* and *AmyB2* are not expressed in pupae, indicating that the expression of *AmyB* could be connected with feeding. Interestingly, *AmyB2* showed high expression in imago heads almost lacking *AmyB1* transcript, suggesting that the splicing of the *AmyB2* transcript might be tissue specific (Fig. 7).

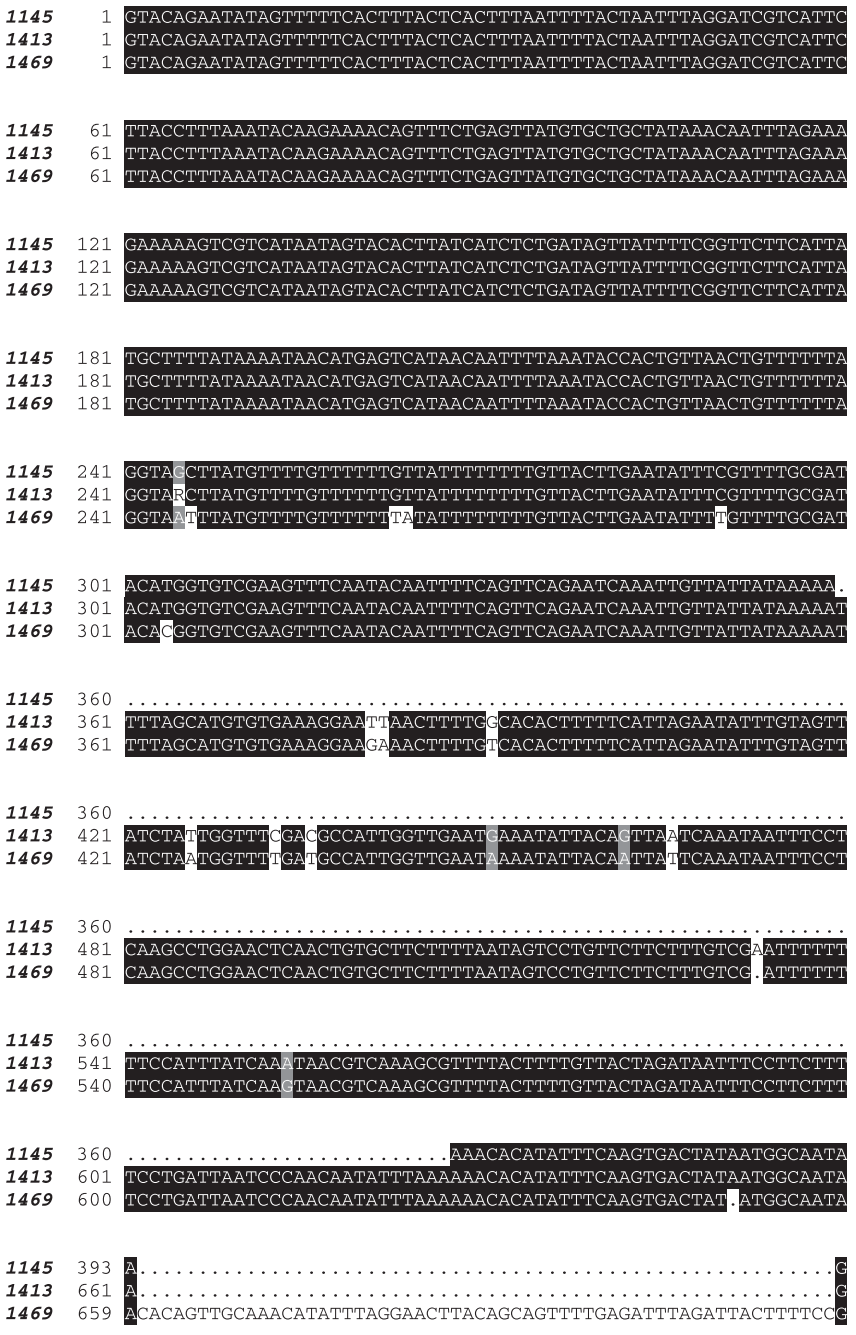

Figure 6. Polymorphism in *AmyB* intron 1. Alignment of three variants of *AmyB* intron 1 sequences comprised 1145, 1413 and 1469 bp. Numbers on the left (in bold and italic) indicate different alleles (named after the length of intron 1). Allele 1413 seemed to be the most common and was used in the final composite sequence (GenBank, accession number: HQ417115). The PCR fragments were received by using a primer pair N+O (Table 1). All beetles were from Šumava mountains.

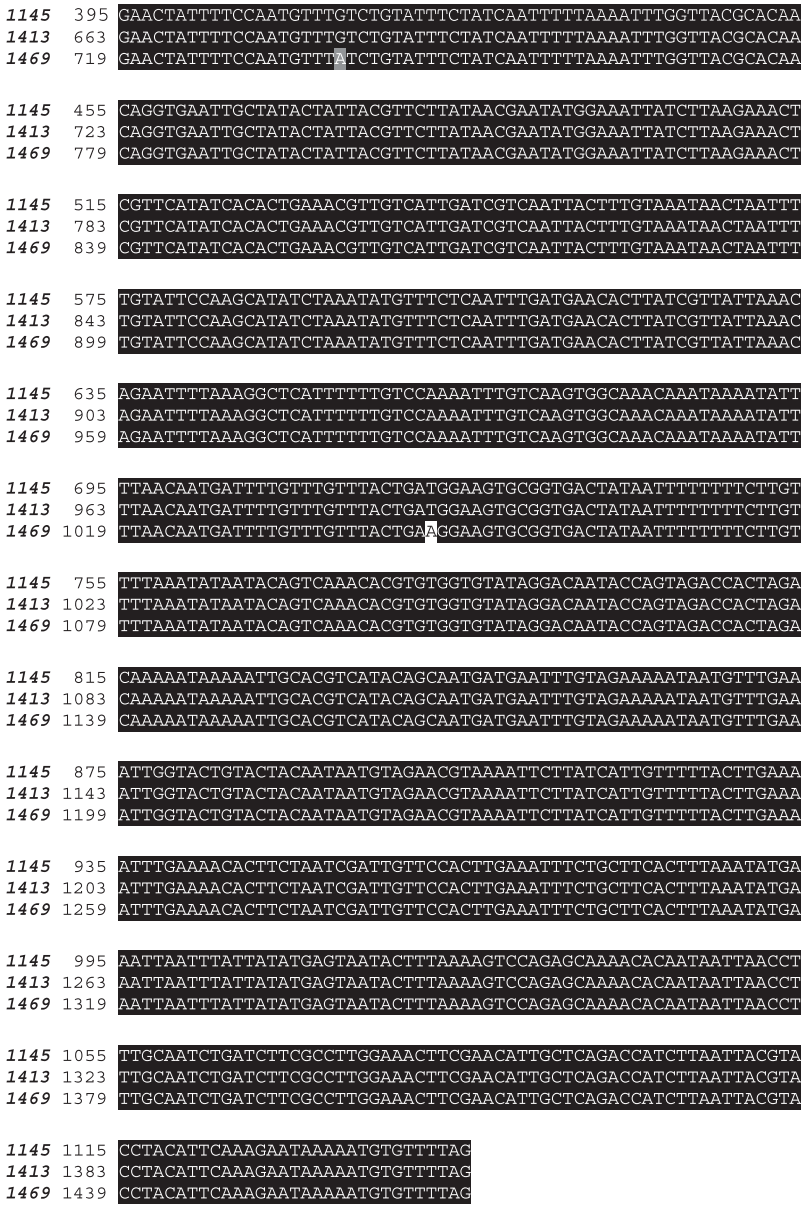

Figure 6. Continued.

Putative α-Amylase Proteins

Both *AmyA* and *AmyB1* contain ORFs encoding 483 amino acids (see Fig. 2) starting with typical signal sequences of 16 amino acids, which are characteristic for secretory proteins (von Heijne, 1985). The proteolytic cleavage site was predicted by the SignalP 3.0 program (Bendtsen et al., 2004) to be between alanine-16 and glutamine-17. The mature, secreted proteins are predicted to be 467 amino acids long with a molecular weight of 51.8 and 52.2 kDa and pI of 4.64 and 4.59 for AmyA and AmyB1, respectively. The *AmyB2* transcript is truncated (see Fig. 2) and encodes a polypeptide

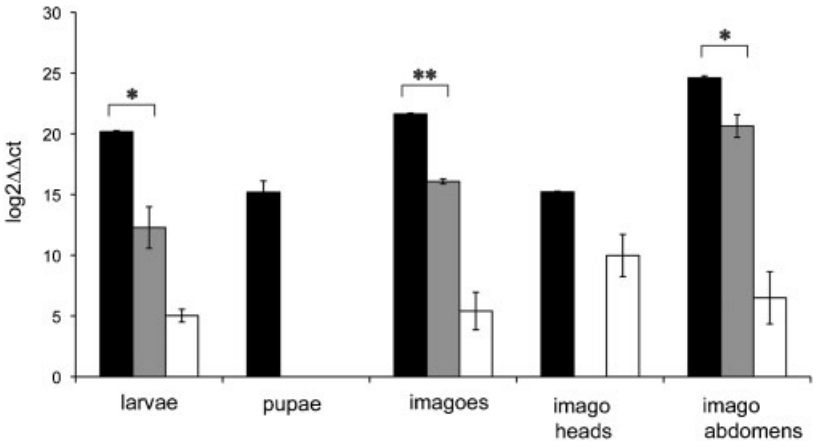

Figure 7. Expression levels of α-*amylase* mRNA during *I. typographus* development as assayed by real-time qRT-PCR. The real time $\log_2\Delta\Delta$ct values of all α-*amylase* transcription variants are shown in larvae, late pupae, imagoes and separated imago heads and abdomens. The expression data on full-length *AmyA* (black columns) is based on two independent experiments (each carried out in triplicates), whereas *AmyB1* (gray columns) and spliced variant *AmyB2* (white columns) are based on three independent experiments (each carried in triplicates). *AmyA* was expressed at a significantly higher level than *AmyB1*. One asterisk indicates $P \leq 0.05$, two asterisks, $P \leq 0.01$. Bars indicate SEM.

of 370 amino acids (including signal peptide) and secreted protein has a molecular weight of 39.3 kDa and pI of 4.60.

The conceptual translation products of *AmyA* and *AmyB1* showed 94% identity at the amino acid level and display 55–80% identity with known homologs of other beetle species, (80% with *A, grandis* amylase I). Several conserved cysteine residues with important structural roles are typical for α-amylases from many insect species (Da Lage et al., 2002). They include eight cysteine residues present in AmyA and AmyB1 as well as AmyB2 and their positions are shown in Figure 2 and S2. These three isoforms also contain additional cysteine residues that are not conserved (Fig. 2 and S2). Two of the nonconserved cysteine residues (Cys257 and Cys268 in *AmyB1*) are missing in *AmyB2*.

Native Gel Electrophoresis and α-Amylase Isoenzymes

I. typographus α-amylase isoenzymes were previously shown to be encoded by a single gene (Stauffer et al., 1992). As we isolated two genes encoding α-*amylases* in *I. typographus* both giving rise to functional proteins, we reexamined the α-amylase isoenzyme staining. Analysis of several individual beetles revealed a single or two-band pattern of well-stained bands, consistent with the homozygous or heterozygous genotypes of a single gene, respectively. Using longer destaing times allowed us to detect a novel class of α-amylase bands with weaker staining and higher electrophoretic mobility. These bands also displayed a one or two band pattern, suggesting that they also may be the products of one gene (Fig. 8A). To further examine the developmental stages of α-amylase expression, we compared the range of isozyme bands in larvae, pupae and adults. As seen in Figure 8B, the weaker and faster bands visible in adults were undetectable during the pupal stage. In contrast, slowly migrating bands could be detected in both adults and late pupae, whereas in pupae they displayed much lower intensity. Larval bands also showed two-loci pattern. Interestingly, the bands

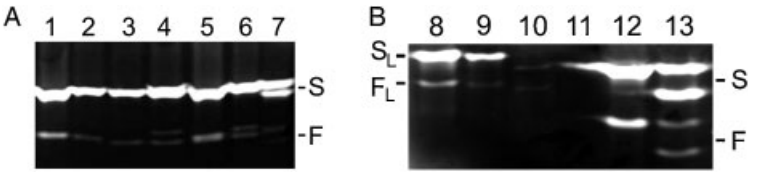

Figure 8. Zymogram of the native starch polyacrylamide gel electrophoresis. (**A**) Extracts from individual beetle abdomens were compared (lanes 1–7). The better stained class of bands migrated slowly (S), whereas weaker band cluster exhibited faster mobility (F). The 1 or 2 band pattern in both classes suggests homo- and heterozygotes, respectively. (**B**) Extracts from larvae (lanes 8 and 9), late pupae (lanes 10 and 11) and adults (lanes 12 and 13). Faster migrating bands (putative AmyB1) were not detected in pupae, whereas both types of bands were detected in adults. Faster migrating bands in larval extracts are marked as F_L and slowly migrating bands are marked as S_L. Lane 11 contains crosscontaminating bands from adjacent sample (lane 12).

consistently displayed slower mobility suggesting some stage-specific posttranslational modification (Fig. 8B).

DISCUSSION

Our data provide evidence that the α-amylase in *I. typographus* is encoded by a tandem of two closely related paralogs. Phylogenetic analysis based on cDNA sequences shows that their closest known homolog is the *Anthonomus Amy1* gene, suggesting that duplication leading to *AmyA* and *AmyB* genes in *I. typographus* occurred recently, after the split of these genera. The sequence similarity of both *I. typographus α-amylase* paralogs might be caused by concerted evolution. This process was shown to play an important role in the evolution of α-amylase genes arranged in tight clusters (Zhang et al., 2003). Alternatively, the low variability of coding sequences might be also maintained by purifying selection (Page and Holmes, 1998). The high divergence of introns and a very low ratio between nonsynonymous and synonymous substitutions detected in *I. typographus α-amylase* genes support the influence of purifying selection rather than concerted evolution.

Both *I. typographus α-amylase* paralogs have five introns. Analogous or slightly simpler exon–intron arrangement was observed in related beetles, including *B. mucronata* (3 and 5 introns in *AmyI* and *AmyII*, respectively) and *T. castaneum* (3 introns). The comparable numbers of introns were also observed in *Apis mellifera* (AF259649.1) and *B. mori* (GQ274006) *α-amylase* genes, which have 4 and 8 introns, respectively. In contrast, the *α-amylase* genes of dipteran insects (including *Drosophilids* and *Anopholes gambiae*) either lack or contain only a single intron (Da Lage et al., 1996). The variety of insect *α-amylases* with multiple introns suggests that the insect prototype gene probably contained multiple introns. The position of an intron corresponding to intron 1 in *I. typographus α-amylase* is the most conserved among holometabolous insects (Da Lage et al., 1996).

The alternative *AmyB2* transcript was detected. The alternatively spliced region was unusual as the 5′ and 3′ splicing sites were localized inside of exons 5 and 6, respectively. The exact intron ends are not known because of the absence of canonical intron flanking sequences (GT-AG) and a 4 bp duplication of the TGGA sequence at the splicing site (Fig. 4). None of the five possible intron ends has been reported, except for the GA dinucleotide found previously at the 5′ splice site of human fibroblast growth factor receptor (Brackenridge et al., 2003) and the TG dinucleotide

detected in the introns of at least 36 human genes as an alternative 3′ splice site (Szafranski et al., 2007). The splicing event is thus presumed to occur at the GA-TG splice sites (line 3, Fig. 4). There is no evidence for alternative splicing of α-amylase in other insects.

Our examination of *AmyA* and *AmyB* expression patterns suggests that they are differentially regulated. *AmyA* was expressed at high levels in larvae and adult abdomens, whereas lower expression was detected in late pupae and adult heads. *AmyB1* and *B2* transcripts were undetectable in pupal stage (Fig. 7). It is possible that the expression of *AmyA* is stage specific, as it appears in late pupa. The expression of both genes might be further influenced by feeding. Differentially regulated α-amylase paralogs were described in *Aedes aegypti*, in which one gene was reported to be tissue specific for the salivary gland (Grossman and James, 1993), or in *Phlebotomus papatasi*, in which the transcript of one paralog was upregulated in the midgut after plant feeding (Jacobson and Schlein, 2001).

Stauffer et al. (1992) described several alleles of α-amylase as products of a single *AmyI* gene. Optimization of staining method allowed us to detect two classes of isoenzyme bands. The first class was comprosed of well-stained slowly migrating bands that were most likely the same as those observed by Stauffer et al. (1992). The newly discovered second class contains faster migrating bands that most probably are products of a separate locus. These faster migrating bands were not detected in samples from pupae, correlating with the absence of *AmyB1* and *AmyB2* mRNA expression at this stage. Taken together, these data suggest that the faster α-amylase bands may be associated with *AmyB*, whereas the class of stronger slowly migrating bands may be products of the *AmyA* gene.

The first and third introns of *I. typographus AmyA* and the first and fourth introns of *AmyB* seem to be much longer than the *Amy* introns of close relatives *T. castaneum* and *B. mucronata* (range 38–57 bp). We observed high polymorphism in the relatively long first and third introns of *AmyA* and first intron of *AmyB* (among several individuals from two areas, see Figs. 5 and 6). Our results offer the possibility of performing a method known as "exon-primed intron-crossing" (EPIC) PCR (Lessa, 1992) that allows an easy detection of a number of presumably neutral and co-dominant allelic variants. As the introns are noncoding regions, it is assumed that they are subjects to fewer functional constrains (Kimura, 1983). Our preliminary screening of a small set of individuals revealed considerable intron length polymorphism. The *α-amylase* genes represent potentially beneficial nuclear markers composed of the sequences encoding key metabolic enzymes and highly variable neutral introns, which can reveal valuable information on the genetic structure of *I. typographus* populations.

ACKNOWLEDGMENTS

We wish to thank Prof. František Sehnal for support of this work, Prof. Christian Stauffer for sharing his experience and knowledge of bark beetles and providing us access to his beetle collections and Prof. František Marec for advices in beetle genetics. We also thank Dr. Jiří Zelený, Dr. Mirka Uhlířová and Mgr. Pavel Trubač for their help with experiments. This work was supported by European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 229518 and MSM 60076605801 (Ministry of Education, Czech Republic) and Entomology Institute Project Z50070508.

LITERATURE CITED

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403–410.

Bendtsen JD, Nielsen H, von Heijne G, Brunak S. 2004. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 340:783–795.

Boer PH, Hickey DA. 1986. The alpha-amylase gene in *Drosophila melanogaster*: nucleotide sequence, gene structure and expression motifs. Nucleic Acids Res 14:8399–8411.

Brackenridge S, Wilkie A, Sreaton G. 2003. Efficient use of a "dead-end" GA 5′ splice site in the human fibroblast growth factor receptor genes. EMBO J 22:1620–1631.

Charlab R, Valenzuela JG, Rowton ED, Ribeiro JM. 1999. Toward an understanding of the biochemical and pharmacological complexity of the saliva of a hematophagous sand fly Lutzomyia longipalpis. Proc Natl Acad Sci USA 96:15155–15160.

Da Lage JL, Wegnez M, Cariou ML. 1996. Distribution and evolution of introns in Drosophila amylase genes. J Mol Evol 43:334–347.

Da Lage JL, Van Wormhoudt A, Cariou ML. 2002. Diversity and evolution of the α-amylase genes in animals. Biologia, Bratislava 57:181–189.

Fuehrer E, Hausmann B, Wiener L. 1991. Borkenkaeferbefall (Col., Scolytidae) und Terpenmuster der Fichtenrinde (Picea abies Karst.) an Fangbaeumen. J Appl Entomol 112:113–123.

Gregoire JC, Evans HF. 2004. Damage and control of BAWBILT organisms, an overview. In: Lieutier F, Day KR, Battisti A, Grégoire J-C, Evans HF, editors. Bark and Wood Boring Insects in Living Trees in Europe, a Synthesis. Dordrecht: Springer. p 19–37.

Grossman GL, James AA. 1993. The salivary glands of the vector mosquito, *Aedes aegypti*, express a novel member of the amylase gene family. Insect Mol Biol 1:223–232.

Grossman GL, Campos Y, Severson DW, James AA. 1997. Evidence for two distinct members of the amylase gene family in the yellow fever mosquito, *Aedes aegypti*. Insect Biochem Mol Biol 27:769–781.

Gruppe A. 1997. Isoenzymatische Variation beim Buchdrucker *Ips typographus*. Mitt Dtsch Ges Allg Angew Ent 11:659–662.

Hickey DA, Benkel BF, Boer PH, Genest Y, Abukashawa S, Ben-David G. 1987. Enzyme-coding genes as molecular clocks: the molecular evolution of animal alpha-amylases. J Mol Evol 26:252–256.

Jacobson RL, Schlein Y. 2001. *Phlebotomus papatasi* and Leishmania major parasites express alpha-amylase and alpha-glucosidase. Acta Trop 78:41–49.

Joensuu J, Heliovaara K, Savolainen E. 2008. Risk of bark beetle (Coleoptera, Scolytidae) damage in a spruce forest restoration area in central Finland. Silva Fennica 42:233–245.

Keeling CI, Bearfield JC, Young S, Blomquist GJ, Tittiger C. 2006. Effects of juvenile hormone on gene expression in the pheromone-producing midgut of the pine engraver beetle, *Ips pini*. Insect Mol Biol 15:207–216.

Kikkawa H. 1953. Biochemical genetics of *Bombyx mori* (silkworm). Adv Genet 5:107–140.

Kimura M. 1983. The neutral theory of molecular evolution. Cambridge, Cambridgeshire; New York: Cambridge University Press. xv, 367p.

Lessa EP. 1992. Rapid surveying of DNA sequence variation in natural populations. Mol Biol Evol 9:323–330.

Levy JN, Gemmill RM, Doane WW. 1985. Molecular cloning of alpha-amylase genes from *Drosophila melanogaster*. II. Clone organization and verification. Genetics 110:313–324.

Magoulas C, Loverre-Chyurlia A, Abukashawa S, Bally-Cuif L, Hickey DA. 1993. Functional conservation of a glucose-repressible amylase gene promoter from *Drosophila virilis* in *Drosophila melanogaster*. J Mol Evol 36:234–242.

Netherer S, Pennerstorfer J, Baier P, Fuhrer E, Schopf A. 2002. Monitoring and risk assessment of the spruce bark beetle, *Ips typographus*. In: McManus ML, Liebhold AM, editors. Proceedings: Ecology, Survey and Management of Forest Insects. Krakow, Poland: USDA Forest Service, Northeastern Reserach Station. p 75–79. Gen Tech Rep NE 311.

O'Connell PO, Rosbash M. 1984. Sequence, structure, and codon preference of the Drosophila ribosomal protein 49 gene. Nucleic Acids Res 12:5495–5513.

Ochman H, Gerber AS, Hartl DL. 1988. Genetic applications of an inverse polymerase chain reaction. Genetics 120:621–623.

Ogita Z-I. 1968. Genetic control of isozymes. Ann N Y Acad Sci 151:243–262.

Page R, Holmes E. 1998. Molecular evolution: a phylogenetic approach. Oxford: Blackwell Science.

Pavlicek T, Zurovcova M, Stary P. 1997. Geographic population-genetic divergence of the Norway spruce bark beetle, *Ips typographus* in the Czech Republic. Biologia 52:273–279.

Pope GJ, Anderson MD, Bremner TA. 1986. Constancy and divergence of amylase loci in four species of Tribolium (Coleoptera, Tenebrionidae). Comp Biochem Physiol B 83:331–333.

Salle A, Kerdelhue C, Breton M, Lieutier F. 2003. Characterization of microsatellite loci in the spruce bark beetle *Ips typographus* (Coleoptera: Scolytinae). Mol Ecol Notes 3:336–337.

Sandstrom P, Welch WH, Blomquist GJ, Tittiger C. 2006. Functional expression of a bark beetle cytochrome P450 that hydroxylates myrcene to ipsdienol. Insect Biochem and Mol Biol 36:835–845.

Siciliano M, Shaw CR. 1976. Separation and visualization of enzymes on gels. In: Smith I, editor. London: Heinemann Medical Books.

Stauffer C, Leitinger R, Simsek Z, Schreiber JD, Fhrer E. 1992. Allozyme variation among nine Austrian *Ips typographus* L. (Col., Scolytidae) populations. J Appl Entomol 114:17–25.

Stauffer C, Lakatos F, Hewitt G. 1999. Phylogeography and postglacial colonization routes of *Ips typographus* L. (Col., Scolytidae). Mol Ecol 8:763–773.

Sugino H. 2007. Comparative genomic analysis of the mouse and rat amylase multigene family. FEBS Lett 581:355–360.

Szafranski K, Schindler S, Taudien S, Hiller M, Huse K, Jahn N, Schreiber S, Backofen R, Platzer M. 2007. Violating the splicing rules: TG dinucleotides function as alternative 3′ splice sites in U2-dependent introns. Genome Biol 8:154.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599.

von Heijne G. 1985. Signal sequences. The limits of variation. J Mol Biol 184:99–105.

Wang Y, Leung F. 2009. Comparative genomic study reveals a transition from TA richness in invertebrates to GC richness in vertebrates at CpG flanking sites: an indication for context-dependent mutagenicity of methylated CpG sites. Genomics, Proteomics Bioinformatics 6:144–154.

Wermelinger B. 2004. Ecology and management of the spruce bark beetle *Ips typographus*—a review of recent research. Forest Ecol Manag 202:67–82.

Zhang Z, Inomata N, Yamazaki T, Kishino H. 2003. Evolutionary history and mode of the amylase multigene family in Drosophila. J Mol Evol 57:702–709.