

Current knowledge on genes and genomes of phytophagous beetles (Coleoptera: Chrysomeloidea, Curculionoidea): a review

JESÚS GÓMEZ-ZURITA and JOSÉ GALIÁN

Área de Biología Animal, Departamento de Zoología y Antropología Física, Universidad de Murcia, 30100 Murcia, Spain;
e-mail: jgzurita@um.es

Keywords. Cerambycidae, Chrysomelidae, Curculionidae, development genes, diapause genes, gene databases, genomics, insect resistance genes, phytophagy genes

Abstract. Fifteen years after the publication of the first genomic sequence of a phytophagous beetle, we review here the wealth of genetic and genomic information generated so far for the species-rich beetle superfamilies Chrysomeloidea (longhorn, seed and leaf beetles) and Curculionoidea (weevils and bark beetles). In this review we briefly describe the most common methods used to investigate the beetle genomes and also compile the nucleotide sequence information stored in public gene databases until December 2004. The motivations and relevance of these research initiatives are described in certain detail, distinguishing among structural and population studies, phylogenetic research, the study of genes involved in development and diapause, energetic metabolism, vegetal tissue digestion, and genes for insect resistance and defence.

INTRODUCTION

There is a considerable effort from public and private sources to sequencing genomes, partially or completely, to know their structure, function and evolution. These economical efforts, mainly addressed to human genomics or economically important species, have given rise to technological developments that have produced a parallel increase of information in other organisms. In arthropods and particularly in insects, these efforts have been concentrated on model systems (such as *Drosophila* Fallen), or economically (*Apis* Linnaeus, *Bombyx* Linnaeus) or medically (*Aedes* Meigen, *Anopheles* Meigen) important organisms (reviewed in Heckel, 2003; and see GenomeWeb at <http://www.hgmp.mrc.ac.uk/GenomeWeb/insect-gen-db.html>).

Drosophila has been used as a model organism for genetic and genomic research. However, whereas *Drosophila* is a good choice for analysing fundamental processes, the huge diversification reached by the arthropods and particularly insects in terms of number of species and ecological strategies, needs incorporating new models to study this diversification from an evolutionary point of view both at structural and functional levels. At present, most of the genomic information available comes from partial descriptions of genomes, and very few arthropod genomes have been targeted as a whole.

The beetles have repeatedly been described in the literature as the paradigm of biodiversity because of the huge amount of species known and the predicted number of species remaining to be discovered (Hammond, 1992). Among beetles, the tenebrionid *Tribolium* MacLeay has been used as successful model organism for genetic research, and not surprisingly it is at the head of the genomic initiative in beetles (BeetleBase at Kansas University; <http://www.bioinformatics.ksu.edu/BeetleBase/index.html>). One of the undoubtedly most successful

beetle lineages is the so-called Phytophaga, a supposedly monophyletic radiation of plant-feeding Coleoptera with over 120,000 species including two major radiations: the Curculionoidea and the Chrysomeloidea (Lawrence, 1982). The Curculionoidea include the weevils, bark beetles and other less species-rich families with an estimate of 50,000 species. The Chrysomeloidea are a heterogeneous group of beetles of uncertain interrelationships including the longhorn-beetles (Cerambycidae), the leaf-beetles (Chrysomelidae sensu lato), and the seed-beetles (Bruchidae; nested within the Chrysomelidae with subfamilial rank by some authors). The Cerambycidae include some 20,000 species, most of them wood-borers (Crowson, 1981), the Chrysomelidae in the order of 40,000 species feeding on green parts of plants (Jolivet & Verma, 2002), and the Bruchidae over 1,300 species specialised in seed- and stem-boring (Southgate, 1979). These two huge lineages include some of the most noxious pests, but in spite of their economical importance and the challenge that they represent for evolutionary studies because of their species richness, they have not been comparatively as well studied as other groups from a genetic point of view. This fact is one of the main motivations to revise what have been the genomic approaches to these beetles so far and what are the perspectives for future research.

The first DNA sequence data published for the Phytophaga were for the cotton boll weevil (Curculionoidea), *Anthonomus grandis* Boheman, precisely belonging to a protamine gene expressed in the beetle testis (Trewitt et al., 1990), and the same species rendered the complete genomic sequence, including introns and regulatory sequences, of the vitellogenin gene (Trewitt et al., 1992). In the case of the Chrysomeloidea the first characterised nuclear gene actually corresponded to a “foreign”-in origin genomic element, a transposable element of the

mariner family that was described in the galericine *Ceratomya trifurcata* (Forster) (Robertson, 1993), followed shortly after by the characterisation of a diapause protein in *Leptinotarsa decemlineata* (Say) (de Kort & Koopmanschap, 1994). Since the publication of these pioneering papers, the leading organisms within Phytophaga in genomic research have been those that represent a threat to economy because of being a pest. Among these, it is worth mentioning the western corn rootworm *Diabrotica virgifera* LeConte, the major pest of maize in the US with a 17% of potential production compromised by the beetle (Branson et al., 1980), the infamous Colorado potato beetle *Leptinotarsa decemlineata*, the bruchid *Callosobruchus maculatus* (Fabricius), known as Mexican bean weevil, or important forest pests within the bark beetles such as *Dendroctonus jeffreyi* Hopkins or several species of *Ips* De Geer, among others. These studies have usually targeted specific genes involved in particular biological processes of interest, e.g. reproduction, nutrition or defence, using an ample spectrum of molecular biology tools. A minor number of studies have used a rather eclectic approach by constructing and analysing small genomic libraries of the fraction of the genome expressed in particular tissues or the entire organism. This alternative, producing a relatively large number of sequences (~500–3500) of expressed genes, the so-called expressed sequenced tags (ESTs), has been applied to three species within the Chrysomeloidea and five within the Curculionoidea. These are *Timarcha balearica* Gory (AJ537611-AJ538039; Theodorides et al., 2002; Gómez-Zurita et al., 2004), *Callosobruchus maculatus* (CB377223-CB377725, CK594665-CK594758; Pedra et al., 2003; Moon et al., 2004) and *Diabrotica undecimpunctata howardi* Barber (CO036822-CO036850; Liu et al., 2004) in the former, and *Curculio glandium* Marsham (BQ476162-BQ476740; Theodorides et al., 2002), *Platystomus albinus* (Linnaeus) (BQ475779-BQ476161; Theodorides et al., 2002), *Ips pini* (Say) (CB407466-CB409136; Eigenheer, 2003), *Diaprepes abbreviatus* (Linnaeus) (CN488395, CN472512-CN476056; Lapointe et al., unpubl.) and *Sitophilus zeamais* Motschulsky (CN612372-CN612484; Heddi et al., 2005) in the latter.

One important branch in genomics, in particular for phylogenetic and evolutionary studies, is mitogenomics, the sequencing and comparative analysis of full mitochondrial genomes. These studies, that have proved very informative in mammals (Curole & Kocher, 1999; Arnason & Janke, 2002; Arnason et al., 2002), are still in development for insects (Whiting laboratory, Dept. Integrative Biology, Brigham Young Univ., Provo, UT). Within Phytophaga, one single mtDNA genome has been fully sequenced to date, belonging to the chrysomelid *Crioceris duodecimpunctata* (Linnaeus), one species of the so-called asparagus beetles (Stewart & Beckenbach, 2003).

In this paper, we will describe what is the wealth of genomic information that has been generated in the past fifteen years for one of the most diversified animal lineages today, the Phytophaga. Also, we will discuss what

are the main tools and goals of genomic research in these beetles and the most remarkable achievements. Finally, we will examine the future and potential applications of these organisms in genomic studies.

METHODOLOGIES USED TO EXPLORE BEETLE GENOMES

In the recent years there has been a fast development of molecular biology techniques allowing the analysis of genomic material reliably and to some extent within reach of laboratories without extremely sophisticated technology or large budgets. Several methodological approaches have been applied in the study of phytophagous beetle genomes, depending on the specific use designed for the genomic information.

Structural sequences and genomic landmarks

A great proportion of the eukaryote genomes is constituted by non-coding genetic material which can have a structural or regulatory function, as is well known for the centromeric heterochromatin or gene introns, regulatory non-coding RNAs, among others (e.g., Eddy, 2001). The knowledge based on these structural and non-coding or anonymous sequences in phytophagous beetles has been applied to the description of the organism genomic organisation, as would be the case of the study of satellite DNA (Lorite et al., 2001, 2002) and telomeric sequences (Okazaki et al., 1993; Sahara et al., 1999; Frydrychová & Marec, 2002), to the characterisation of hypervariable microsatellite loci (e.g., Alvarez et al., 2003; Sembène et al., 2003), and the elaboration of a genetic sequence tagged sites (STS) map based on AFLP data for the Colorado potato beetle (Hawthorne, 2001).

Each specific genomic object required a well-established particular methodological approach. The analysis of abundant satellite DNA sequences is usually carried out by partial and total digestions of isolated genomic DNA using particular restriction endonucleases. The existence of restriction targets in these highly repetitive sequences in tandem is typically visualised in electrophoresis gels as patterns of regularly size-decreasing DNA fragments (DNA ladder). The recognised monomeric fragment can be extracted from the gel, cloned into an appropriate vector and subsequently sequenced and characterised. The analysis of satellite DNA, mainly localised around the centromere in beetles, can be of interest for comparative analysis and study of centromere evolution in these insects. Incompatibilities in the structure of centromere DNA have been postulated as factors responsible for incorrect gamete segregation in hybrids, potentially promoting reproductive isolation and speciation (Henikoff et al., 2001).

Telomeres are protein-DNA complexes at the ends of eukaryotic chromosomes that prevent their fusion and protect their terminal degradation, among others (Blackburn, 1991). Telomeric DNA usually consists of multiple repeats of a simple motif which is typically (TTAGG)_n in the case of insects, although it can vary (Sahara et al., 1999). The standard methodology for the study of insect telomeric DNA is based on Southern hybridisation of a

labelled specific telomere probe with digested total DNA of the species of interest blotted onto a nylon membrane. A positive hybridisation signal is interpreted as the presence of the specific motif in the insect genome. The physical location of these DNA sequence motifs at the terminal ends of the chromosomes is identified at a later stage using fluorescent in situ hybridisation with fluorescently labelled motif probes onto chromosome spreads of the species of interest (e.g., Sahara et al., 1999; Frydrychová & Marec, 2002).

The microsatellites are commonly used genetic markers for population genetic studies (Goldstein & Schlotterer, 1999). These are repetitions in tandem of short DNA motifs, typically di-, tri- or tetranucleotides or combinations of them, which are polymorphic in the populations regarding the number of repetitions of the specific motif, generally presenting a high number of alleles. The individuals can be characterised for their allelic constitution at various microsatellite loci defining a specific genetic profile; this can be compared or integrated in the context of the population to make inferences about the population dynamics and genetic structure. The most extended methodology to isolate and characterise particular microsatellite loci consists in generating a partial genomic library of the organism of interest, which is screened for the presence of clones incorporating a repeated motif by probing with oligonucleotides specific for that motif. These clones are isolated, their inserted sequence determined and the presence of the microsatellite confirmed. Flanking primer pairs for the region containing the microsatellite are subsequently designed and used for a PCR-based screening of individuals from different populations to score for length polymorphisms detectable by highly resolvable polyacrylamide electrophoresis gels.

Large-scale genome projects as those conducted for model organisms required an initial effort to create genomic landmarks using anonymous and already identified loci that could serve as the backbone for the construction of the genomic maps of ever-increasing resolution and density. A similar initial step has been carried out in one of the better-studied leaf beetles, the Colorado potato beetle, by a thorough analysis and linkage mapping of amplified fragment length polymorphisms (AFLP) markers (Hawthorne, 2001). The AFLP method consists of a PCR-based amplification of a random collection of enzymatically restricted genomic fragments (Vos et al., 1995). In brief, the procedure starts by restricting isolated DNA from the organism of interest with a combination of six-cutter and four-cutter restriction enzymes. The DNA fragments generated are added to their ends specific short double stranded sequences (adapters) by an enzymatic T4 DNA ligase reaction. The obtained restriction fragment plus terminal adapters can be amplified by subsequent PCR using primers complementary to the adapter sequences. The amplification step is usually carried out in two successive rounds, the first using the primers complementary to the adapter sequences and the second using the same primers with specific 2–3 nucleotides overhangs on the 3'-end for a selective and more

restrictive PCR amplification. The obtained pattern of amplified bands is visualised on polyacrylamide gels. Finally, the comparison of individual AFLP and segregation patterns for a sufficiently high number of members of a known organism pedigree with backcrosses from an original genetically distinct mating couple, allows the identification of linkage groups and the construction of the corresponding linkage map. One potentially very useful derivation of these analyses consists in the identification of particular AFLP markers co-segregating with traits of interest in the organism, e.g., resistance to insecticide, as a tool for the discovery of genes involved with the trait in particular. The associated DNA band can be recuperated from the electrophoresis gel, cloned into a vector and sequenced.

Use of “universal” primers

The recovery and analysis of nuclear genetic data from newly studied organisms can benefit from the information generated on the ongoing genome projects or other genetic studies on model organisms. In the case of insects, the advances in the knowledge of the fruit fly genome (Adams et al., 2000) as well as those of butterflies and moths have been critical to allow a development of similar approaches to other organisms, such as the beetles, including the Phytophaga. In particular, our knowledge on primary sequence data for conserved areas of nuclear genes of interest can be used to generate PCR primers able to amplify the homologue region in a broad taxonomic range of organisms. This approach has been mainly exploited in phylogenetic studies, but not necessarily restricted to them (e.g., Krauss & Reuter, 2000; Labeyrie & Dobler, 2004). The results of a primer-based approach are usually erratic and the utility of a given set of primers for a lineage of interest can only be assessed empirically, the same as the degree of phylogenetic information allowed by the nuclear marker. Other problems associated with nuclear markers, including paralogy, allelism, inconsistent amplifications, and others (see Gómez-Zurita et al., 2004 for discussion on the problems of nuclear markers) have jeopardised the successful implantation of these markers routinely in phylogenetic analyses.

Genomic pools, libraries and cDNA sequencing

The genome of *Drosophila* has been estimated to contain about 13,379–13,601 genes as a conservative estimate (Adams et al., 2000; Misra et al., 2002) or much larger figures in studies focusing on specific tissues (Andrews et al., 2000), and it is reasonable to suppose that the genomes of phytophagous beetles will be in the same range or significantly higher. This enormous amount of genes and their interspersed among vast non-coding regions, make the isolation of genes related to specific functions in these organisms, from which our genomic knowledge is still in its initial stages, a very difficult venture. However, there is a range of elegant methodologies that make it possible to find the “needle in the stack”.

The first choice in the studies designed to isolate and characterise specific genes involves deciding the source of genetic information. There are genes that show higher levels of expression in certain tissues than in others or in specific stages of the beetle life cycle, so that studies focused on these genes benefit from a tissue or stage specific approach. This is the case of many studies interested in the characterisation of digestive enzymes, which use gut tissue as starting point (Girard & Jouanin, 1999a,b; Pedra et al., 2003; Sugimura et al., 2003; Zhu-Salzman et al., 2003; Bown et al., 2004; Gruden et al., 2004), or those centred in metabolic enzymes, focusing on the insect fat body, an organ involved in many homeostatic mechanisms and active centre of metabolic activity, where food reserves are stored in the insect and the maintenance of the balance between resources and requirements takes place (Smith et al., 1994; Vermunt et al., 1997, 1998; Saito et al., 2004). In the case of genes specific of life cycle stages, those involved in the diapause or dormancy are a good example (Yocum, 2001, 2003). The study of genes and proteins of constitutive expression, like cuticle proteins for instance (Kim et al., 2003b), is usually carried out from entire or parts of larval or adult specimens.

In either case, tissue/stage specific or non-specific strategies, depending on the previous knowledge on the gene or genes of interest, there are two potential analytical approaches. When the target of the study is a gene that has been previously characterised in other organisms, a PCR-based approach can be undertaken. In this case the PCR is made selective by using degenerate primers designed from the back-translation of the amino acid sequence of the protein in those areas conserved across several taxa. This strategy benefits greatly from the public genetic databases and the genome initiatives. However, sometimes the expectations about protein sequence conservation do not apply to the studied organism and this approach renders unfruitful, with PCRs not working (e.g., Gómez-Zurita et al., 2004) or working suboptimally, either producing multiple bands or faint PCR products (e.g., Zhu & Clark, 1995). Alternatively to the use of degenerate primers, these can be made gene and species specific by a prior analysis of the mature protein of interest, in those cases that the protein is isolated with relative ease and characterised by amino terminal sequencing (Smid et al., 1997; Sugimura et al., 2003).

In many occasions, the relevant genes have not been characterised in related taxa or the product of their expression does not show conserved areas to allow a PCR-based approach. When this happens, or simply because it is a more versatile strategy, the investigation of the genome is done through a previous construction of a cDNA library. A cDNA is a double stranded DNA sequence complementary to a messenger RNA (mRNA). A cDNA library is therefore a collection of clones each containing the DNA version of a mature, intronless RNA coding for a peptide or a protein in the studied organism. Briefly, the construction of such a library first requires the isolation of the mRNAs from the insect, which can be

done in a two-steps process by first isolating total RNA from the tissue (including abundant ribosomal RNA and other non-protein coding RNAs) with a subsequent step of isolation of the mRNA by affinity of the poly-A tracts of these molecules to oligo-T adsorbed to a column or other substrates. The single stranded poly(A) RNAs are transformed into double stranded cDNA following a two-steps procedure. Poly(A) RNA is used as substrate to synthesise a complementary first DNA strand using a reverse-transcriptase (RT) and oligo-dT primers. This step generates a hybrid RNA-DNA molecule. The subsequent addition of a ribonuclease (RNase H), and sometimes as a by-product of the RT enzyme, degrades the RNA strand of the hybrid generating small RNA fragments that will prime the second strand DNA synthesis from the single stranded DNA (ssDNA) thanks to the activity of a bacterial DNA polymerase I. In this process a few nucleotides of the 5' end of the original poly(A) RNA are not preserved, which is not usually a big problem, since all eukaryotic mRNA molecules have about 40–80 nucleotides of 5' noncoding leader sequence (Kozak, 1983).

The products of cDNA synthesis can serve as a mixed pool of the expressed fraction of the genome or subsequently cloned into an appropriate vector, usually phages, but also plasmids, to constitute a cDNA library, with advantages for storage and posterior analysis of specific clones containing a cDNA of interest.

Depending on the specific application of the cDNA library and the availability of previous information on the genes studied, there are two different approaches to extract the information from the library. One is through random clone sequencing, which is usually a fast and productive strategy when a library is tissue/stage specific and there is an active and abundant expression of the genes of interest. The chances to spot by random selection of clones some containing useful genes are usually high. For example, 21 out of 76 randomly selected cDNA clones included digestive enzymes in a gut specific library in *Phaedon cochleariae* (Fabricius) (Girard & Jouanin, 1999b), and 15 out of 70 contained cysteine proteinases in a library from the same tissue in *Diabrotica virgifera virgifera* (Bown et al., 2004).

The second strategy opposed to random sequencing is a targeted sequencing of the clone containing the gene of interest using different methods of screening. An obvious way of screening is using a selective PCR-based approach similarly as described above, either by using degenerate primers based on areas of amino acid conservation or by designing species/gene specific primers after a previous characterisation of the amino terminal region of the studied protein. This approach can be applied both to pools of synthesised cDNAs or to cDNA libraries, and generally produces partial cDNA sequences, their size depending on the locations of the primers used for the amplification.

Two other library-screening strategies can be used, one by clone selection by Southern blot analysis with homologous probes obtained for instance with the partial cDNA

sequences generated by selective PCR as above, or with heterologous DNA probes obtained from other organisms. A second screening method is immunological and involves an initial isolation and purification of the protein target of the study. This protein is used to immunise an organism to obtain antibodies. The labelled antiserum can be exposed to the cDNA library to detect the reaction associated to the antibody recognition of the protein expressed in the clone containing the respective cDNA.

The above-mentioned techniques usually allow a partial characterisation of the cDNA sequence with the 5'-, the 3'- or both ends missing. In those cases where a full characterisation of the cDNA sequence is required, particularly on studies of gene structure or gene expression, a very useful method can be applied, known as rapid amplification of cDNA ends through PCR (RACE-PCR; Frohmann et al., 1988). 5'RACE requires the use of an antisense gene specific primer designed from the partial sequence of the cDNA of interest. This primer is used with RT on poly(A) RNA to generate ssDNA complementary to the gene of interest and spanning from the priming site to the 5'-end in full, which will be in the 3'-end of the ssDNA. A homopolymeric A-tail is added to this 3'-end and a first round of PCR on the ssDNA is carried out using an oligo-dT primer including a specific 5' anchor sequence and a nested gene specific primer. Finally, a second round of PCR is done using as primers an oligonucleotide complementary to the oligo-dT anchor sequence and a third nested gene specific primer. This final PCR product can be cloned or sequenced and will contain the complete 5'-end of the cDNA. 3'RACE is generally simpler and requires using an oligo-dT anchor primer to synthesise ssDNA from mRNA and a subsequent step of PCR using as primers the anchor sequence and a gene specific primer.

The availability of gene specific DNA probes obtained with the methods outlined above can help in the investigation of the gene structure in the genome. These probes can be used to screen genomic libraries, with clones that contain big fragments of genomic DNA (i.e., 15–45 kb), to select those with the gene of interest, including control regions, regulatory elements, and introns. This strategy has been followed to fully characterise the diapause protein 1 of *Leptinotarsa decemlineata* (Koopmanschap et al., 1995) or the vitellogenin gene in *Anthonomus grandis* (Trewitt et al., 1992).

INFORMATION STORED IN GENE DATABASES

The importance of gene databases to assist and speed up the genomic research in phytophagous beetles has been mentioned in previous sections. Databases do not only contribute to our knowledge on the available genomic information, but also contribute to new findings as different specialists can analyse existing data under different perspectives, usually with different insight than their intended original purpose.

When finishing writing this review, there are genetic data in the gene databases including DNA sequences

TABLE 1. Annual entries in public gene databases corresponding to genetic data for the Chrysomeloidea. The entries are organised depending on their genomic origin, including mitochondrial sequences, nuclear ribosomal sequences, genomic sequences (protein coding and non-coding) and EST data.

	No. genera	mtDNA	nrDNA	genomic	EST data
1993	1	–	–	6	–
1994	2	–	–	2	–
1995	12	52	–	4	–
1996	1	3	–	–	–
1997	7	63	3	2	–
1998	10	110	–	2	–
1999	16	220	10	10	–
2000	115	307	82	13	–
2001	37	291	119	38	–
2002	18	266	20	34	–
2003	146	457	359	155	932
2004	138	584	310	74	121
Total	373	2353	903	340	1053

from all genomic compartments for 373 genera of Chrysomeloidea (14.2% Cerambycidae; 82.0% Chrysomelidae; and 3.8% Bruchinae; April 7, 2005). The equivalent stored information for the Curculionoidea consists of data for 209 genera (45.0% Scolytinae, including bark and ambrosia beetles).

In Chrysomeloidea, from 1993 (year of publication of the first nuclear genomic sequence for this superfamily) until the end of 2004, a total of 3,596 annotated sequences have been deposited in the nucleotide sequence databases, together with 1,053 EST sequences (Table 1). From these, only 9.5% correspond to annotated nuclear protein coding genes (30% considering EST data), with most of the sequence production (65.4%) corresponding to mitochondrial sequences used for phylogenetic studies. In the case of the Curculionoidea, the studies on these organisms lagged slightly behind in the accumulation of sequence data in gene databases, but have recently experienced a dramatic increase, particularly in the production of genomic data, mostly thanks to the study of economically important species (Table 2). There are currently

TABLE 2. Annual entries in public gene databases corresponding to genetic data for the Curculionoidea. The entries are organised depending on their genomic origin, including mitochondrial sequences, nuclear ribosomal sequences, genomic sequences (protein coding and non-coding) and EST data.

	No. genera	mtDNA	nrDNA	genomic	EST data
1993	3	1	1	2	–
1994	–	–	–	–	–
1995	3	80	–	–	–
1996	2	33	–	1	–
1997	5	28	–	–	–
1998	4	21	–	4	–
1999	36	40	2	4	–
2000	92	222	47	77	–
2001	72	66	107	123	–
2002	95	192	83	122	1102
2003	35	299	36	154	1671
2004	27	457	83	132	3659
Total	209	1439	359	619	6432

TABLE 3. Nucleotide gene database entries for nuclear protein coding and non-coding (excluding ribosomal) sequences in Phytophaga. The name of the annotated gene is given, as well as the organisms for which the gene has been characterised, their accession numbers and the associated references.

Gene	Organism	GenBank Acc. No.	Reference
3-hydroxy-3-methylglutaryl CoA reductase	<i>Ips paraconfusus</i> , <i>Dendroctonus jeffreyi</i> , <i>Ips pini</i> , <i>Anthonomus grandis</i>	AF071750, AF159136-AF159138, AF166002, AF304440, AF162705 AY357301	Tittiger et al., 1999, 2000, 2003; Hall et al., 2002; Blake et al., unpubl. Yocum, 2004
5-aminoimidazole-4-carboxamide ribonucleotide fomytransferase/IMP cyclohydrolase	<i>Leptinotarsa decemlineata</i>		
90 kDa heat shock protein	<i>Callosobruchus maculatus</i>	CK594692	Moon et al., 2004
acaloleptin A	<i>Acalolepta luxuriosa</i>	AB094343	Imamura et al., unpubl.
acetylcholinesterase	<i>Leptinotarsa decemlineata</i>	L41180	Zhu & Clark, 1995
acidic ribosomal protein P0	Chrysomelinae spp.	AJ534424-AJ534438	Gómez-Zurita et al., 2004
actin	<i>Apriona germari</i>	AY817141	Gui et al., unpubl.
AgSP-1 arylphorin	<i>Anthonomus grandis</i>	AF512519	Lewis et al., 2002
AlCRP	<i>Acalolepta luxuriosa</i>	AB104817	Saito et al., 2004
alpha-amylase	<i>Zabrotes subfasciatus</i> , <i>Phaedon cochleariae</i> , <i>Diabrotica virgifera</i> , <i>Anthonomus grandis</i> , <i>Callosobruchus maculatus</i>	AF255722, Y17902, AF208002- AF208003, AF527876-AF527877, CK594677	Grossi-de-Sá & Chrispeels, 1997; Girard & Jouanin, 1999b; Titarenko & Chrispeels, 2000; Oliveira-Neto et al., 2003, Moon et al., 2004
alpha-glucosidase	<i>Callosobruchus maculatus</i>	CK594678-CK594679	Moon et al., 2004
antimicrobial peptides	<i>Callosobruchus maculatus</i> , <i>Acalolepta luxuriosa</i>	CK594684-CK594688, AB191477	Moon et al., 2004; Ueda & Sato, unpubl.
apolipophorin-III	<i>Derobrachus geminatus</i>	L25278	Smith et al., 1994
arylphorin-like hexameric storage protein 2 (arylphorin-like hexamerin)	<i>Apriona germari</i>	AY293287	Kim et al., 2004
arylphorin-like hexamerin	<i>Apriona germari</i>	AF509880	Kim et al., 2003c
beta-galactosidase	<i>Callosobruchus maculatus</i>	CK594680-CK594681	Moon et al., 2004
cadherin-like protein	<i>Diabrotica virgifera</i>	AY769085	Siegfried et al., 2005
carboxypeptidase	<i>Callosobruchus maculatus</i>	CK594758	Moon et al., 2004
cathepsin B	<i>Apriona germari</i>	AF483623	Kim et al., unpubl.
cathepsin B-like proteinase	<i>Diabrotica virgifera virgifera</i> , <i>Callosobruchus maculatus</i>	AJ583509, AJ583513, CK594673- CK594676, AY429465	Bown et al., 2004; Moon et al., 2004; Moon et al., unpubl.
cathepsin D	<i>Apriona germari</i>	AF454831	Kim et al., 2001
cathepsin L-like cysteine proteinase CAL1	<i>Diabrotica virgifera virgifera</i> , <i>Callosobruchus maculatus</i>	AF190653, CK594665-CK594672	Koiva et al., 2000; Bown et al., 2004; Moon et al., 2004
cathepsin L-like cysteine proteinase precursor	<i>Sitophilus zeamais</i> , <i>Callosobruchus maculatus</i> , <i>Hyperba postica</i> , <i>Anthonomus grandis</i> , <i>Acanthoscelides obtectus</i>	D82884-D82886, AF544834-AF544844, AF157961, AY336947, AY345219	Matsumoto et al., 1997; Zhu-Salzman et al., 2003; Wilhite et al., 2000; Oliveira-Neto et al., 2004b; Silva et al., unpubl.
cellulase	<i>Apriona germari</i>	AY162317, AY741064	Lee et al., 2005
cellulase II	<i>Apriona germari</i>	AY451326	Lee et al., unpubl.
ceropin	<i>Acalolepta luxuriosa</i>	AB188500	Saito et al., unpubl.
chitinase	<i>Phaedon cochleariae</i>	Y18011	Girard & Jouanin, 1999a
chymotrypsin	<i>Phaedon cochleariae</i>	Y17904	Girard & Jouanin, 1999b
coated vesicle membrane protein	<i>Callosobruchus maculatus</i>	CK594696	Moon et al., 2004
COP9 signalosome complex subunit I	<i>Callosobruchus maculatus</i>	CK594706	Moon et al., 2004
cytochrome P450 variant 3	<i>Diabrotica virgifera virgifera</i>	AF257100	Scharf et al., 2001
cytochrome P450	<i>Callosobruchus maculatus</i>	CK594690-CK594691	Moon et al., 2004

TABLE 3 continued.

Gene	Organism	GenBank Acc. No.	Reference
defensin 1 precursor	<i>Acalolecta luxuriosa</i>	AF311352	Ueda et al., unpublished
diapause protein 1	<i>Leptinotarsa decemlineata</i>	X76080, X86074	de Kort & Koopmanschap, 1994; Koopmanschap et al., 1995
diapause-associated transcripts	<i>Leptinotarsa decemlineata</i>	AF532862-AF532864	Yocum, 2003
diapausin	<i>Gastrophysa atrocyanea</i>	AB070558	Tanaka et al., 2003
digestive cysteine protease (intestaim)	<i>Leptinotarsa decemlineata</i>	AY159368- AY159378	Gruden et al., 2004
dihydropyrimidine dehydrogenase	<i>Callosobruchus maculatus</i>	CK594697	Moon et al., 2004
elongation factor 1-alfa	Galerucinae spp., Xyleborini, <i>Coccotrypes</i> spp., <i>Curculio</i> spp.	AF186657-AF186696, AF308396-AF308432 and AF308513, AF259863-AF259893, AF375264-AF375278, AY040313-AY040318, AF397612-AF397649, [AF508866-AF508887; AF508923-AF508928; AH011660-AH011661], AF444071-AF444078, AF439740-AF439743, AF466310-AF466345, AY171449-AY171471, AY377062-AY377123, AY327652-AY327684	Normark et al., 1999; Sequeira et al., 2000; Jordal et al., 2000; Farrell et al., 2001; Sequeira & Farrell, 2001; Cognato & Vogler, 2001; Jordal, 2002; Jordal et al., 2002a,b; Duckett & Kjer, 2003; Kim et al., 2003a; Jordal et al., 2004; Hughes & Vogler, 2004
endoglucanase	<i>Phaedon cochleariae</i>	Y17907	Girard & Jouanin, 1999b
endopolygalacturonase	<i>Callosobruchus maculatus</i>	CK594682	Moon et al., 2004
engrailed	<i>Prionus californicus</i>	AF275783	Whiting, unpubl.
enolase	Scolytidae, Tomiicini	AF375279-AF375293 and AF375332-AF375347, AY040298-AY040312	Farrell et al., 2001; Sequeira & Farrell, 2001
ferritin	<i>Callosobruchus maculatus</i> , <i>Leptinotarsa decemlineata</i>	CK594689, AY297687	Moon et al., 2004; Qiu et al., unpubl.
ferritin 1	<i>Apriona germari</i>	AF509878	Kim et al., 2003c
ferritin 2	<i>Apriona germari</i>	AF509879	Kim et al., 2003c
formate dehydrogenase	<i>Callosobruchus maculatus</i>	CK594698	Moon et al., 2004
GABA receptor subunit (Rdl)	<i>Leptinotarsa decemlineata</i>	AF070936	Glueck et al., unpublished (1998)
glycerol-3-phosphate dehydrogenase (Gpdh)	<i>Hypothenemus hampei</i> , <i>Callosobruchus chinensis</i>	AF037324, AF323983-AF323984	Andreev et al., 1998; Park et al., 2001
heterochromatin protein Su(var)3-9	<i>Chytus arietis</i> , <i>Leptinotarsa decemlineata</i>	AJ290961-AJ290963, AJ290965	Krauss & Reuter, 2000
IDGF (putative)	<i>Diaprepes abbreviatus</i>	AY821658	Huang et al., unpubl.
inducible heat shock 70 KDa protein	<i>Leptinotarsa decemlineata</i>	AF322911, AF288978	Yocum, 2001
insecticide resistance-associated cytochrome P450	<i>Diabrotica virgifera virgifera</i>	AF243506-AF243507	Scharf et al., 2001
juvenile hormone esterase-related protein (JHRE)	<i>Leptinotarsa decemlineata</i>	AF039135	Vermunt et al., 1998
Lambda cDNA library	<i>Callosobruchus maculatus</i>	CB377223-CB377725	Pedra et al., 2003
larval cuticle protein 8.7 (LCP8.7)	<i>Apriona germari</i>	AY769984	Gui et al., unpubl.
larval cuticle precursor (LCP10.7)	<i>Apriona germari</i>	AF428097	Kim et al., 2003b
larval cuticle protein 12.3 (LCP12.3)	<i>Apriona germari</i>	AF518323	Kim et al., 2003b
larval cuticle protein 12.6 (LCP12.6)	<i>Apriona germari</i>	AF518324	Kim et al., 2003b
lysosomal thiol reductase	<i>Callosobruchus maculatus</i>	CK594694-CK594695	Moon et al., 2004
mariner transposase pseudogene	<i>Ceratomyia trifurcata</i> , <i>Megacyllene robiniae</i>	L10455-L10459	Robertson, 1993; Robertson et al., 1998

TABLE 3 continued.

Gene	Organism	GenBank Acc. No.	Reference
microsatellite loci	<i>Pissodes strobi</i> , <i>Gerstaeckeria galapagoensis</i> , <i>Caryedon serratus</i> , <i>Acanthoscelides obveatus</i> , <i>Coccotrypes carpophagus</i> and <i>C. dactyliperda</i> , <i>Tomicus piniperda</i> , <i>Ips typographus</i> , <i>Hypothenemus hampei</i> , <i>Amphitimetus transversus</i> , <i>Acanthoscelides obtectus</i> , <i>Monolepta vinca</i> , <i>Diabrotica virgifera</i> , <i>Oreina elongata</i> <i>Callosobruchus maculatus</i>	AF348150-AF348155, AY083353-AY083357, AY159474- AY159483, AF527375- AF527379, AF549465-AF549494, AY165173-AY165177, AY243326-AY243330, BV102444-BV102450, AY430192-AY430197, AY681082-AY681087, AY575860-AY575868, AY738552-AY738548, AY380123- AY380131 CK594693 AY703482 AJ549437- AJ617746 AY185203	Liewlaksaneeyanawin et al., 2001; Dhuyvetter et al., 2002; Sembène et al., 2003; Álvarez et al., 2003; Berg et al., 2003; Kerdelhué et al., 2003; Sallé et al., 2003; Gauthier & Rasplus, 2004; Patt et al., 2004a; Alvarez et al., 2004; Patt et al., 2004b; Kim & Sappington, 2005; Margraf & Rahier, unpubl. Moon et al., 2004 Gui et al., unpubl. Labeysrie & Dobler, 2004 Hong et al., unpubl.
mucin	<i>Callosobruchus maculatus</i>	CK594693	Moon et al., 2004
muscle LIM protein	<i>Apriona germari</i>	AY703482	Gui et al., unpubl.
Na/K-transporting ATPase alpha	<i>Platycorynus sauteri</i> , <i>Chrysochus</i> spp.	AJ549437- AJ617746	Labeysrie & Dobler, 2004
nicotinic acetylcholine receptor alpha subunit	<i>Anoplophora glabripennis</i>	AY185203	Hong et al., unpubl.
odorant-binding proteins Rpal-OBP2' and RpalOBP4'	<i>Rhynchophorus palmarum</i>	AF139912, AF141865, AY182011, AY394703	Nagnan-Le Meillour et al., 2004
p53 tumor suppressor	<i>Leptinotarsa decemlineata</i>	BD250011	Buchman et al., unpubl.
P8 apoptotic factor	<i>Callosobruchus maculatus</i>	CK594700	Moon et al., 2004
pectin methyltransferase	<i>Sitophilus oryzae</i>	AY841894	Shen et al., unpubl.
peptidoglycan recognition protein	<i>Callosobruchus maculatus</i>	CK594683	Moon et al., 2004
PhEG cellulase	<i>Psacotheca hilaris</i>	AB080266	Sugimura et al., 2003
phosphoenolpyruvate kinase	<i>Desmogramma subtropica</i> , <i>Platyphora</i> spp.	AY055489- AY055510	Termonia et al., 2002
phosphoglyceromutase	<i>Curculio</i> spp.	AY327631-AY327651	Hughes & Vogler, 2004
polygalacturonase	<i>Phaedon cochleariae</i> , <i>Sitophilus oryzae</i>	Y17906, AF207068	Gitraud & Jouanin, 1999b; Shen et al., 2003
protamine	<i>Anthonomus grandis</i>	X52058	Trewitt et al., 1990
putative juvenile hormone esterase JHE	<i>Leptinotarsa decemlineata</i>	AF035423	Vermunt et al., 1997
reverse transcriptase	<i>Callosobruchus maculatus</i>	CK594699	Moon et al., 2004
ribosomal protein S5e	<i>Timarcha balearica</i>	AJ783869	Longhorn et al., unpubl.
ribosomal protein S7e	<i>Timarcha balearica</i>	AJ783871	Longhorn et al., unpubl.
ribosomal protein S8e	<i>Curculio glandium</i>	AJ783873	Longhorn et al., unpubl.
ribosomal protein S10e	<i>Curculio glandium</i>	AJ783877	Longhorn et al., unpubl.
ribosomal protein S12e	<i>Curculio glandium</i>	AJ783879	Longhorn et al., unpubl.
ribosomal protein S13e	<i>Timarcha balearica</i>	AJ783881	Longhorn et al., unpubl.
ribosomal protein S14e	<i>Curculio glandium</i>	AJ783883	Longhorn et al., unpubl.
ribosomal protein S18e	<i>Timarcha balearica</i>	AJ783890	Longhorn et al., unpubl.
ribosomal protein S25e	<i>Platystomus albinus</i>	AJ783896	Longhorn et al., unpubl.
satellite DNA	<i>Chrysolina americana</i> , <i>Xanthogaleruca luteola</i>	AJ299723- AJ299732, AJ493398- AJ493415	Lorite et al., 2001, 2002
serine protease inhibitor	<i>Callosobruchus maculatus</i>	CK594705	Moon et al., 2004
serine proteinase	<i>Anthonomus grandis</i>	AY536260-AY536264	Oliveira-Neto et al., 2004a
STS data	<i>Leptinotarsa decemlineata</i> , <i>Hypothenemus hampei</i>	AY028388- AY028397, BV102444-BV102450	Hawthorne, 2001; Gauthier & Rasplus, 2004
thiol protease	<i>Phaedon cochleariae</i>	Y17903	Gitraud & Jouanin, 1999b
thymidilate	<i>Callosobruchus maculatus</i>	CK594703-CK594704	Moon et al., 2004

TABLE 3 continued.

Gene	Organism	GenBank Acc. No.	Reference
translation initiation factor 2 gamma (eIF-2g)	<i>Clytus arietis</i> , <i>Leptinotarsa decemlineata</i> ,	AJ290961, AJ290964	Krauss & Reuter, 2000
trypsin	<i>Phaedon cochleariae</i> , <i>Diaprepes abbreviatus</i>	Y17905, AF072690	Girard & Jouanin, 1999b; Yan et al., 1999
trypsin-like proclotting enzyme precursor	<i>Anthonomus grandis</i>	AF377992	Batista et al., unpubl.
USP-RXR	<i>Hypera postica</i>	AY157933	Bonneton et al., 2003
vitellogenin	<i>Anthonomus grandis</i> , <i>Pissodes strobi</i>	M72980, U63328	Trewitt et al., 1992; Leal et al., 1997
voltage sensitive sodium channel	<i>Leptinotarsa decemlineata</i>	AF114489, AY212030- AY212032	Lee et al., 1999; Hawthorne, 2001, 2003
xylanase	<i>Phaedon cochleariae</i>	Y17908	Girard & Jouanin, 1999b

2,417 annotated gene sequences in the public databases, and 6,432 EST sequences. Only considering the former, 25.6% actually correspond to nuclear protein coding sequences, but this proportion rises to 79.7% if EST data are included.

In Table 3 we present a comprehensive list of genes characterised in particular species within the Phytophaga, with their associated accession number in the public gene databases and the published reference. The most remarkable aspect from this account is the reduced number of species which are the object of genomic scrutiny, and basically six taxa concentrate most of the research in this field: *Acalolepta luxuriosa* (Bates), *Apriona germari* (Hope) (Cerambycidae), *Leptinotarsa decemlineata*, *Diabrotica virgifera* (Chrysomelidae), *Callosobruchus maculatus* (Bruchidae) and *Anthonomus grandis* (Curculionidae). These organisms, because they represent a head start for genomic investigation in Phytophaga and because of their economical importance, could be proposed as model organisms for future genome initiatives in this large beetle assemblage.

GENOME DATA IN PHYTOPHAGA

Structural DNA, population genetics markers and genetic maps

Satellite DNA associated to heterochromatin does not constitute a conspicuous fraction of the Phytophaga genomes as suggested by the results of C-banding experiments on chromosome preparations of several taxa (e.g., Rožek et al., 2004). However, its presence has been reported so far in two leaf beetles belonging to different subfamilies, including the chrysomeline *Chrysolina americana* (Linnaeus) (Lorite et al., 2001) and the galeucine *Xanthogaleruca luteola* (Müller) (elm leaf beetle; Lorite et al., 2002). Although totally unrelated evolutionarily, these two species-specific structural genomic elements share a similarly small monomer size (189 and 149 bp, respectively), a similarly high A+T content (~59%), the lack of methylation and their predominantly pericentromeric distribution in the chromatin (Lorite et al., 2001, 2002).

The characteristics and localisation of telomeric sequences in Phytophaga have been studied in the long-horn beetles *Spondylis buprestoides* (Linnaeus) and *Arhopalus coreanus* (Sharp), the chrysomelid *Leptinotarsa decemlineata*, the scolytid *Ips typographus* (Linnaeus) and the weevils *Sitophilus granarius* (Linnaeus) and *Sipalinus gigas* (Fabricius) (Okazaki et al., 1993; Sahara et al., 1999; Frydrychová & Marec, 2002). Interestingly, all these taxa except for the weevils show the standard (TTAGG)_n insect telomeric DNA motif. This telomere motif is supposedly the ancestral of insects, although it has been lost independently in different groups (Sahara et al., 1993). In the Coleoptera, the loss of (TTAGG)_n and its replacement by alternative motifs seems to have occurred also multiple times in several families within the Adephaga and Polyphaga suborders, and it appears clearly heterogeneous in the Curculionoidea (Frydrychová & Marec, 2002).

Regarding to the analysis of microsatellite markers in the Phytophaga beetles, it is still restricted to the characterisation of a number of microsatellite loci in two leaf beetles (*Monolepta vinca* Gerstaecker and *Diabrotica virgifera*; Patt et al., 2004; Kim & Sappington, 2005), three seed beetle species that constitute serious pests of common beans in Central America (*Acanthoscelides obvelatus* Bridwell; Alvarez et al., 2003) and worldwide [*A. obtectus* (Say); Alvarez et al., 2004] and of groundnuts in Central Africa (*Caryedon serratus* Olivier; Sembène et al., 2003), five bark beetle species pests of palm seeds [*Coccotrypes carpophagus* (Hornung) and *C. dactyliperda* (Fabricius); Berg et al., 2003], pine forests [*Tomicus piniperda* (Linnaeus); Kerdelhué et al., 2003], Norway spruce forests (*Ips typographus*; Sallé et al., 2003), and coffee (*Hypothenemus hampei* Ferrari; Gauthier & Rasplus, 2004), and the weevils *Pissodes strobi* (Peck) (Liewlaksaneeyanawin et al., 2001), *Gerstaeckeria galapagoensis* Van Dyke (Dhuyvetter et al., 2002), and *Amphitmetus transversus* (Kolbe) (Patt et al., 2004). Genetic variability has been assessed also in the scolytid *Hypothenemus hampei*, the major insect pest of coffee, using sequences of the γ -aminobutyric acid receptor subunit (*Rdl*), which together with ITS2 rDNA and mitochondrial *cox1* allowed to report the lack of nuclear nucleotide variability compared to the mitochondrial genome in this beetle (Andreev et al., 1998).

The only genomic map currently available for a phytophagous beetle is that of *Leptinotarsa decemlineata*, the Colorado potato beetle (Hawthorne, 2001). The map consists of 172 AFLP marker loci (only 96 discussed) inferred to belong to 18 linkage groups, a result consistent with the known karyotype of the species, with 17 autosomes and a single X-chromosome in males (Hsiao & Hsiao, 1983). The markers appeared to be spread throughout the beetle genome with an average distance of 11.1 cM, which is a high marker density for mapping purposes, particularly for one of the originally intended uses of the map, quantitative trait locus (QTL) positioning in *Leptinotarsa* Stål (Hawthorne, 2001). One first application of the genomic map in *Leptinotarsa* was the mapping on the X-chromosome of a candidate gene for a voltage-sensitive sodium channel (*LdVssc1*), involved in the sensitivity response of several insects, including the Colorado potato beetle, to the pyrethroid insecticides (Hawthorne, 2001). Later, the same strategy was used to find additional loci involved in pyrethroid resistance (Hawthorne, 2003).

A PCR-primer based approach has been also exploited in Chrysomeloidea to identify and characterise a very interesting feature of gene organization in eukaryotes: the dicistronic unit containing the heterochromatin-associated Su(var)3-9 protein and the γ subunit of the translation initiation factor eIF-2, functionally unrelated and resolved by alternative splicing. This dicistronic unit with most of the Su(var)3-9 open reading frame within the second intron of eIF-2 and first described in *Drosophila melanogaster* Meigen (Tschiersch et al., 1994) was found to be common to other holometabolous insects, including

the chrysomelid *L. decemlineata* and the cerambycid *Clytus arietis* (Linnaeus) (Krauss & Reuter, 2000). A combined approach was followed in this study using degenerated primers, reverse transcriptase-PCR from cDNAs with gene specific primers and inverse genomic PCR with combinations of internal primers specifically designed for each studied organism using the sequence information generated in the initial stages. The conservation of this genomic peculiarity among *Drosophila*, two species of Coleoptera and the moth *Scoliopteryx libatrix* (Linnaeus), not shared by the centipede *Lithobius forficatus* (Linnaeus) or any other organism studied so far, suggests the origin of the gene fusion in the lineage of holometabolous insects and has been proposed as a useful feature for evolutionary studies in arthropods (Krauss & Reuter, 2000).

Genomic markers for phylogenetic analyses

Up till now, most phylogenetic analyses in Phytophaga have relied on mtDNA markers or less frequently on nuclear ribosomal sequences, including the ribosomal genes and their spacer sequences (see Tables 1 and 2). Very little has been done using other nuclear markers, and only five different loci appear in the phylogenetic literature: partial sequence of the enolase gene (Sequeira & Farrell, 2001), a partial sequence containing one intron (430–572 bp) of the phosphoenolpyruvate carboxykinase gene (*PEPCK*; Termonia et al., 2002), a fragment with 1–2 introns of the acidic ribosomal protein P0 (*RpP0*; Gómez-Zurita et al., 2004), a fragment of the phosphoglyceromutase (Hughes & Vogler, 2004), and partial sequences of the elongation factor 1-alpha gene (*EF1a*) that have been used in a variety of phylogenetic studies (Duckett & Kjer, 2003; Kim et al., 2003a). Except in Gómez-Zurita et al. (2004), the origin of the primers used to amplify the corresponding nuclear marker is not indicated. However, their design possibly followed a similar procedure by comparing available primary sequences from the marker of interest in other organisms to identify conserved domains, with ulterior modifications to the primer sequence based on preliminary and fragmentary results with the original set of primers on the organism of interest. Gómez-Zurita et al. (2004) followed this strategy to generate primers able to amplify a 500–541 bp fragment of the *RpP0* gene in a sample of representatives of several genera in the subfamily Chrysomelinae (Chrysomelidae) to illustrate the potential of a genomic cDNA library as a source of information for primer design useful in phylogenetic research. In the other studies, the nuclear phylogenetic markers were used in combination with other mitochondrial and nuclear ribosomal markers to establish the phylogeny of the group of interest. Sequeira & Farrell (2001) and Farrell et al. (2001) introduced the use of enolase sequences as phylogenetic marker in beetles to test for the time of the association of certain lineages of Tomicini scolytids to *Araucaria* hosts, both with Gondwanan distributions, and for the study of the evolution and shifts between free-living and obligate mutualism with ambrosia fungi life histories in Scolytidae, respectively. The goal of the phylogenetic hypothesis in

Termonia et al. (2002) was to understand the evolution of chemical defence in chrysomeline beetles of the genus *Platyphora* Gistel by analysing the reconstruction of transitions among different chemical compounds sequestered by the beetles from their host-plants. The derived acquisition of the “ability” to sequester more than one secondary metabolite from the plants is hypothesised to broaden the range of host-plant affiliations. *EF1a* has been used as phylogenetic marker successfully in a great variety of studies in bark beetles and weevils. This marker in combination with others has been particularly important in several systematic and evolutionary studies in bark beetles at different taxonomic levels and focusing on relevant biological questions involved in this group, like the evolution and genetic consequences of haplo-diploidy and sib-mating/inbreeding, male neoteny, or host-plant associations (Normark et al., 1999; Jordal et al., 2000, 2002a; Sequeira & Farrell, 2001; Jordal, 2002) or in the study of systematic relationships (Cognato & Vogler, 2001; Sequeira et al., 2000; Jordal et al., 2002b, 2004). In the Chrysomeloidea, this marker has been used by Kim et al. (2003a) to investigate the systematic relationship between two so-called subfamilies within the Chrysomelidae, Alticinae (flea beetles) and Galerucinae, to offer molecular support for the lumping of both taxa into a single subfamily with a paraphyletic flea beetle lineage containing a monophyletic galerucine clade. Finally, Duckett & Kjer (2003) proposed a molecular phylogeny of the Neotropical Oedionychina, a group of flea beetles particularly diversified in the New World.

Genes for insect development and diapause

Most applied and genomic research in phytophagous beetles is carried out on model systems that constitute agricultural pests of serious economical importance, and the objective of the research is usually to understand metabolic or physiological aspects of the beetles that can eventually help in combatting the pests by disrupting their normal function. Two important areas of investigation along these lines are those that concentrate on the study of critical developmental stages of the beetle, particularly in the metamorphosis and the regulation of diapause programs, and on the study of the energetic metabolism and the use of energetic reserves (next section).

The study of genes acting or controlling diapause, which involves molecular mechanisms similar to those active during the metamorphosis stage in the beetle life cycle, has been almost exclusively carried out on the Colorado potato beetle, *Leptinotarsa decemlineata*. Diapause or dormancy is a stage of inactivity of adult beetles conditioned by environmental stimuli, such as a short photoperiod, but also high temperatures, that prepare and allow the beetle to go through periods of metabolic stress, including desiccation or denutrition, because of adverse environmental conditions.

Up till now, very little is known about the genes acting on diapause at the molecular level. In *L. decemlineata* only six classes of these genes have been identified and characterised, including the diapause protein 1 (Koopmanschap et al., 1995), juvenile hormone esterases (Ver-

munt et al., 1997, 1998), 70 kDa heat-shock proteins (Yocum, 2001), a protein similar to desiccation protein Dsp28 of *Tenebrio molitor* Linnaeus (Yocum, 2003), and two other proteins of unknown function upregulated during diapause or overwintering (LdDAT-1 and LdDAT-3; Yocum, 2003). Additional related genes encoding for the 3-hydroxy-3-methylglutaryl coenzyme A synthase and reductase (*HMG-S* and *HMG-R*, respectively), important in the regulation of juvenile hormone, but also in the production of aggregation pheromones, have been characterised in three species of scolytids, *Ips paraconfusus* Lanier, *I. pini* and *Dendroctonus jeffreyi* (Tittiger et al., 1999, 2000, 2003; Hall et al., 2002). Another diapause specific peptide gene of unknown function was isolated in the chrysomeline beetle *Gastrophysa atrocyanea* Motschulsky, with the peculiarity of sharing high similarity with an insect iridescent virus (Tanaka et al., 2003). The knowledge about the regulation of these genes and others involved in this important process in the life of an insect are extremely relevant in agricultural pests as serious as *L. decemlineata*, because they will eventually allow to increase the effectiveness in predicting the need for and the optimal timing to apply control measures.

L. decemlineata diapause protein 1 is an arylphorin-type storage hexamer that accumulates in the haemolymph in diapaused beetles and in their last larval stage, disappearing during metamorphosis. This hormone is present correlating with an absence of juvenile hormone. The full-length of the gene has been characterised to be around 9.0 kb, containing a coding region of 2256 bp structured as five exons and four introns. This gene is much larger than its homologue in Lepidoptera, although the exon/intron structure is fairly conserved, except for the 3'-end of the gene. The 5'-region of the gene contains a number of potential regulatory sequences that might be involved in the hormonal control of gene expression (Koopmanschap et al., 1995). A putative functional homologue for this gene (*AgSP-1*) has been studied by Lewis et al. (2002) in the boll weevil, *Anthonomus grandis*, the most noxious pest of cotton after its introduction in the US. The main utility for this protein is as an indicator of diapause in *A. grandis*, since other methods based on the analysis of the physiological or metabolic status of the beetles proved unsatisfactory.

Juvenile hormone esterases have a similar expression pattern to that of diapause protein 1, with expression peaks before metamorphosis or diapause, and it is responsible for the reduction of juvenile hormone levels in the beetle haemolymph. This enzyme is regulated by the photoperiod and its gene was first described in the Colorado potato beetle by Vermunt et al. (1997) as coding for a protein of 515 amino acids with several functional esterase motifs, but missing other typical motifs and lacking significant similarity with other known insect esterases. The same authors later proved the existence of two different genes encoding for juvenile hormone esterases in *Leptinotarsa*, the previously identified *JHE.A* and a second gene, *JHE.B*, without apparent function in

the haemolymph. These genes showed a 77% identity and were demonstrated to lay relatively close in the beetle genome, both being mapped on a restriction genome fragment of 5.4 kb (Vermunt et al., 1998).

The application of library suppression protocols of diapaused and non-diapaused beetles used by Yocum (2003) allowed discovering other genes up-regulated during dormancy stages. One such gene was *LdDAT-2*, encoding for a protein 229 amino acids with 27% identity and 40% similarity with Dsp28 of *T. molitor*, which is expected in the context of diapause given the importance of preparation of the organism to prevent desiccation naturally occurring during the winter. Two other genes of unknown function were discovered using this strategy: *LdDAT-3* with no similarity whatsoever to any known protein and *LdDAT-1*, coding for a protein of 286 amino acids with low identity with proteins characterised by leucine-rich domains, indicating that this protein could form protein-protein complexes in vivo.

Heat shock proteins (HSP) are also expressed as part of the dormancy developmental program in insects and they have a double function (i) as chaperones assisting in protein folding and assemblage of protein complexes in unstressed cells and (ii) as inhibitors and disruptors of protein aggregates, refolding proteins and targeting proteins for degradation under stress (Denlinger et al., 2001). Two genes for these proteins, *LdHSP70A* and *LdHSP70B*, similar to other HSPs described from Diptera (~70% similarity) have been characterised in the Colorado potato beetle, which show different expression patterns developmentally regulated, but also dependent on the thermal history which the beetles were exposed to (Yocum, 2001). It has been suggested that individual members of the 70 KDa HSP family may have varying roles in overwintering survival in *L. decemlineata* (Yocum, 2001).

The double role of *HMG-R* and *HMG-S* in juvenile hormone metabolism and pheromone production in beetle pests has been carefully studied at the molecular level in three pine forests pests, *Ips paraconfusus*, *I. pini* and *Dendroctonus jeffreyi*. The *HMG-R* gene was first characterised in *I. paraconfusus* at the cDNA level using a PCR-based approach and degenerated primers, yielding a complete cDNA including an ORF of 720 bp with high similarity to other insect *HMG-R* and up-regulated by juvenile hormone (Tittiger et al., 1999). The same gene in *I. pini* was described to contain an ORF for a predicted protein of 866 amino acids (Hall et al., 2002). The complete structure of the gene on a related species, *D. jeffreyi*, showed an exon-intron structure conserved in part even with the homologue gene in vertebrates, and a similar expression pattern to that observed in *I. paraconfusus* (Tittiger et al., 2003). The cDNA for *HMG-S* was isolated and described for *D. jeffreyi* as a gene encoding for a 457 amino acid protein with high identity to the homologue genes in *Blattella germanica* (Linnaeus) and *D. melanogaster* (Tittiger et al., 2000). Both genes, *HMG-R* and *HMG-S*, the latter present at least in two copies in the beetle genome, express near the metathoracic-abdominal border, particularly in the anterior midgut, where phero-

me precursors are synthesised as identified in other bark beetles (Tittiger et al., 2000; Hall et al., 2002).

A last group of genes investigated in Phytophaga related to development are those involved in molting and metamorphosis. The gene cascade regulating larval moults and the changes during metamorphosis in insects is dependent on the steroid ecdysone (Henrich & Brown, 1995). This hormone binds to the ecdysone receptor, a heterodimer of two nuclear receptors, ecdysone receptor (ECR) and ultraspiracle (USP-RXR) (Yao et al., 1993). Precisely, one of these genes, the one encoding for the ultraspiracle homologue, has been studied in the weevil *Hypera postica* (Gyllenhal) in the context of the evolution of the ecdysone receptor in insects (Bonneton et al., 2003). Finally, the cuticle proteins, the major component of insect integument, have been one important research target. In particular, the genes for three of these proteins, AgLCP9.2, AgLCP12.3 and AgLCP12.6, differentiated by the predicted molecular mass of the inferred protein, were characterised in the longhorn beetle *Apriona germari*, the mulberry longicorn beetle, which is one major pest of mulberry trees in Asia (Kim et al., 2003b). These three proteins were confirmed to express specifically in the epidermis, showing different intervals of expression after ecdysis and during larval growth (Kim et al., 2003b).

Genes involved in the energetic metabolism

The main interest in energetic metabolism of phytophagous beetles revolves around the study of proteins that mobilise reserves during special conditions in the life cycle of the insect. Hexamerin proteins, such as diapause protein 1 in *L. decemlineata* and AgSP-1 in *A. grandis*, also belong into this group. Particular interest has been devoted to larval storage proteins (LSPs), proteic hexamers synthesised during larval development in the fat body where they act as sources of amino acids used by pupae and adults during metamorphosis and reproduction, respectively, being fundamental for the insect development (Telfer & Kunkel, 1991). This work has been carried out on the cerambycid *A. germari* for which three hexameric storage proteins, SP1, SP2 and SP3, had been previously purified from the haemolymph and characterised (Yoon et al., 2001). The gene for the first hexameric LSP of *A. germari* (*AgeHex*) was identified by Kim et al. (2003c) and confirmed to be a single copy gene in the beetle genome, associated through immunological analyses to the previously identified SP2 protein. *AgeHex* belongs to the family of insect LSPs, including hexamerins, arylphorins, diapause proteins, methionine-rich storage proteins and juvenile hormone-suppressible proteins, and a phylogenetic analysis using *AgeHex* sequence information clustered it with high support with the coleopteran related sequences present in the gene database: diapausin 1 of *L. decemlineata* (Koopmanschap et al., 1995) and hexamerin2 and early-staged encapsulation protein of the mealworm *Tenebrio molitor* (Cho et al., 1999). A later study based on sequencing of randomly selected clones from a *A. germari* cDNA library yielded information on an additional LSP gene, named *AgeHex2*,

confirmed to correspond to the previously identified LSP SP3 (Kim et al., 2004). *AgeHex2* was deduced to have 679 amino acids (the identified ORF encoded a protein of 696 amino acid residues, with a secretion signal peptide of 20 amino acids) with high sequence identity with other beetle hexamerins and confirmed by Northern blotting to express specifically in the larval fat body (Kim et al., 2004).

Vitellogenin is the precursor of the egg storage protein vitellin in insects, which is essential during embryogenesis providing the nutrients for embryonic development. The gene for vitellogenin has been studied in two weevil species, including *A. grandis* and the white pine weevil, *Pissodes strobi* (Trewitt et al., 1992; Leal et al., 1997). The vitellogenin beetle gene shows a high degree of conservation from nematodes to vertebrates and insects, even with regard to the exon-intron structure (Trewitt et al., 1992). The interest in this gene and its regulation, apart from being relevant to understand the molecular mechanisms of oogenesis, also roots in the fact that its expression seems affected by certain host characteristics. Differential levels of this expression have been found for beetles feeding on susceptible or resistant host trees, the latter compromising severely the insect female fertility and the embryo survival (Leal et al., 1997).

Another protein characterised at the molecular level, important for the beetle energetic metabolism, is the apolipoprotein III (apoLp-III). This apolipoprotein is present in the haemolymph of insects and has been associated to the translocation of lipid stores required to fuel prolonged flight. The gene and the encoded protein were characterised in the Palo Verde beetle, the cerambycid *Derobrachus geminatus* LeConte, showing certain similarity with other apoLp-III available data suggesting a common origin and not similarity due to analogous secondary structure (Smith et al., 1994).

One last study that can be cited in this section is the molecular investigation by Smid et al. (1997) on genes for peptides produced by the male accessory glands (MAG) in *L. decemlineata*. It is believed that male accessory glands produce metabolites that are used to manipulate the behaviour of females by inducing monogamy and/or accelerated oviposition. A peptide specific of MAG in *Leptinotarsa decemlineata* (Led-MAGP) was characterised as a 74 residues peptide highly repetitive with a succession of imperfect hexarepeats and certain similarity with the N-terminal hexa-repeat section of the chicken prion protein, whereas it did not share any similarity to other previously known insect MAG peptides (Smid et al., 1997). Given the similarity with the avian prion, inducible of endocytosis, the function of Led-MAGP, once transferred to the female in the coupling, could be related to increasing the uptake of yolk protein accelerating oocyte growth increasing oviposition rates.

The genes of phytophagy

Most beetles have been proven insensitive to *Bacillus thuringiensis* Bt toxins, so that alternative strategies are needed to control pests and reduce the use of pesticides (Girard & Jouanin, 1999a). One of the natural defence

mechanisms in plants against their predators consists in the release in the attacked tissue of specific inhibitors of digestive enzymes, which reduce the nutrients intake by the insect slowing down its development and chances of survival (Shade et al., 1994; Leplé et al., 1995; Schroeder et al., 1995). Naturally, as in any system where organisms' interests are in conflict, an arm-race develops. In this case, the beetles as shown in the Colorado potato beetle or in the Mexican bean weevil are able to counter-attack the use of enzyme inhibitors in a variety of ways with an integrated strategy, for instance by overproduction of the inhibited digestive enzyme, by degrading the inhibitor molecule and/or by simply producing an alternative digestive enzyme less susceptible to the action of the inhibitor (Zhu-Salzman et al., 2003; Gruden et al., 2004). The interest in the characterisation at the molecular level of digestive enzymes in phytophagous beetles is therefore to provide the attacked plants with specific enzyme inhibitors through genetic engineering that could enhance their resistance against the pest (e.g., as suggested by Gilikin et al., 1992). Genetic engineering allows to quickly bypass the result of the interaction among organisms through evolutionary times providing in this case to the plants with a new molecular advantage taken from other plants defence mechanisms to which the beetles were never exposed. Inhibitors of digestive enzymes are good candidates not just because of their potential effectivity, but also for being an environmentally friendly alternative to pest control.

The study of the digestive enzymatic machinery in Phytophaga has centred once again onto species with a marked economical impact on agriculture. Particularly important in this research have been the citrus, alfalfa and cotton boll weevils (*Diaprepes abbreviatus*, *Hypera postica* and *Anthonomus grandis*), the bruchids *Zabrotes subfasciatus* (Boheman), the Mexican bean weevil, and *Callosobruchus maculatus*, the cowpea weevil, the cerambycids *Apriona germari*, mulberry longicorn beetle, and *Psacotheta hilaris* (Pascoe), another mulberry and fig trees pest, and the chrysomelids *Phaedon cochleariae*, pest of cruciferous crops, and the severe pests *L. decemlineata* and *D. virgifera virgifera*.

Two different methods have been used to characterise digestive enzymes, and investigate their expression patterns and regulation profiles, either by trying to survey as much enzymatic diversity as possible or by focusing on specific enzymes or enzyme types. In the first case, gut specific cDNA libraries were surveyed by random sequencing of clones or total sequencing of the library in *P. cochleariae* (Girard & Jouanin, 1999a), *C. maculatus* (Pedra et al., 2003) and *D. virgifera* (Siegfried et al., 2005), respectively. This approach readily allows the identification of many genes involved in food breakdown, including glycoside (e.g., amylases, polygalacturonases, glucanases) and peptide hydrolases (e.g., cysteine and aspartic acid proteinases, trypsin and chymotrypsin).

Glycoside hydrolases like the α -amylase have been studied at the molecular level in *Z. subfasciatus*, *D. virgifera virgifera* and *A. grandis*. In *Zabrotes* Horn the

deduced mature enzyme, product of the gene *ZsAmy*, has 466 amino acid residues and is similar to other previously characterised insect α -amylases (Grossi de Sá & Chrispeels, 1997). In *Diabrotica* Chevrolat two different α -amylase genes were discovered, *Dva1* and *Dva2*, with an 83% identity at the amino acid level and also sharing characteristics with other previously described insect and mammal α -amylases (Titarenko & Chrispeels, 2000). Interestingly, the introduction of both cDNAs for α -amylase into an expression baculovirus vector infecting Sf9 cells from *Spodoptera frugiperda* (Smith), showed that *ZsAmy* is not inhibited by the α -amylase inhibitor α AI-1 of cultivated beans, but it is inhibited by α AI-2 of some wild beans, whereas *Dva1* showed inhibition by α AI-1 as well as the inhibitor WI of *Triticum aestivum* Linnaeus (Grossi de Sa & Chrispeels, 1997; Titarenko & Chrispeels, 2000). Similarly to the case of *Diabrotica*, in *Anthonomus* Germar two different genes for α -amylases were described, *Amylag1* and *Amylag2*, sharing a 58% identity at the predicted amino acid sequence and 50–62% sequence identity with other Phytophaga sequences (Oliveira-Neto et al., 2003).

Among the glycoside hydrolases of phytophagous beetles, also digestive cellulases have deserved attention. These enzymes were supposed to be exogenous in animals; however, they have been discovered in a wide range of animals and arranged belonging to three different families probably of unrelated origin (Watanabe & Tokuda, 2001). The inferred ORF from the cDNA isolated in the longhorn beetle *Psacotha hilaris* had 325 amino acids (304 excluding a predicted signal sequence for secretion) and similarity searches against the protein database revealed higher similarity with the family 5 sub-family 2 of cellulases, so far only reported for nematode and some bacterial cellulases (Sugimura et al., 2003). Lee et al. (2004, 2005) have reported on two different but related forms of endogenous cellulases, *Ag-EGase I* and *II*, expressed specifically in the midgut of *Apriona germari*. Another related enzymatic activity is that of the polygalacturonase, which has been described both in the Curculionidae and Chrysomelidae, but again without clear evidence supporting that it is encoded by the insect genomes rather than by symbiotic microorganisms (e.g., Campbell, 1989). The isolation of a cDNA encoding for an endo-polygalacturonase (pectinase) in the rice weevil, *Sitophilus oryzae* (Linnaeus), both from control and antibiotic-treated beetles, has proven the insect origin of the putative gene encoding for the pectinase, opening the possibility that this gene was incorporated into the beetle genome through horizontal transfer (Salzberg et al., 2001; Shen et al., 2003).

Digestive proteolysis in Coleoptera is predominantly due to cysteine proteinase activity, contrary to what happens in mammals and most insects that use serine proteinases. Among these in *D. virgifera* cathepsin L-like proteinases are the major enzymes responsible for protein digestion (Koiwa et al., 2000). These enzymes in the western corn rootworm were shown to be inhibited specifically by soyacystatin N (scN), the soybean cysteine

proteinase inhibitor N (Koiwa et al., 2000; Liu et al., 2004). Random sequencing of a cDNA library from *D. virgifera* gut tissue resulted in a relatively high proportion (15%) of cysteine proteinase-like proteins, with the characteristic amino acid residues of the catalytic centre of the predicted enzymes, that belonged to two proteinase groups: cathepsins L-like (9 out of 11 clones with sequence similarities in the range 36–40%) and cathepsins B-like (2 out of 11 clones; sequence similarity 44%) (Bown et al., 2004). Similar results were obtained in a posterior study by Siegfried et al. (2005). The analysis by Gruden et al. (2004) of a gut-specific cDNA library of *L. decemlineata* allowed retrieving the genes for three classes of digestive cysteine proteinases induced by adaptation to feeding on plants releasing high levels of proteinase inhibitors. The adapted cysteine proteinases, named intestains A, B and C (within group 90% identity, between groups 40–60%) constituted about 10% of the clones of the library. The induced intestains have a predicted structure similar to papain and they could confer resistance to proteinase inhibitors inactivating them by cleavage, for instance (Gruden et al., 2004). The screening of a cDNA library with low stringency and a cysteine protease probe allowed identifying 30 different genes as multiple isoforms of cathepsin-L like proteases belonging to two families (CmCPA and CmCPB) in the Mexican bean weevil, *Z. subfasciatus* (Zhu-Salzman et al., 2003). Observed differences in key amino acids of the protein sequences may well be significant enough to cause differences in the susceptibility to scN inhibition, with CmCPB possibly responsible for scN-insensitive CP activity in *Z. subfasciatus* (Zhu-Salzman et al., 2003).

Cathepsin D is an aspartic protease included in the pepsin family, and one of the major factors contributing to lysosomal digestive activity. This enzyme was shown to express in most body tissues in the mulberry longhorn beetle, *A. germari*, and the gene was characterised to produce a protein of 386 amino acids with the typical features of the aspartic proteases and more similar to the homologue enzymes in other insects (67.2–68.2% identity with Diptera sequences) than to other animals (Kim et al., 2001).

Apart from the predominant cysteine and aspartic proteinases, serine proteinases (e.g., trypsin) can be important as well in digestive processes in some phytophagous beetles, particularly curculionids. Thus, the gene encoding trypsin has been also studied in detail in one weevil, *Diaepes abbreviatus* (Yan et al., 1999).

Biopesticides and the molecular study of insect resistance and defence

In the previous section it has been described the importance of knowing the molecular properties of beetle digestive enzymes to understand how resistance against plant defence develops, but also to learn about their susceptibility to specific plant inhibitors and how they can be used through genetic engineering to combat beetle pests. However, other genes and the product of their expression can be used as the target of biotechnological actions for pest control, either inducing the synthesis of biopesticides

by the plants or by manipulating the mechanism of molecular resistance to insecticides in the beetles.

Insect chitinases expressed in crop plant genomes could be used as biopesticides against insects because of toxicity associated to deregulation of chitin degradation in the gut (Ding et al., 1998). For this reason, the molecular analysis of genes encoding for chitinases in beetles is a first step needed to implement this strategy against agricultural pests. Girard & Jouanin (1999b) randomly sequenced a cDNA library allowing the characterisation of a gut-specific chitinase gene in the chrysomelid *Phaedon cochleariae* (pest of cruciferous plants such as oilseed rape and cabbage). Southern blot and activity gels analyses using the cDNA as a probe suggested the presence of several such genes in the genome of this species, some involved in the peritrophic matrix turnover and some highly expressed mainly during pupation, possibly involved in the degradation of larval cuticle. Other natural pesticides that act in the plant defence against insect pests are the production of cardenolides (cardiac glycosides) by the attacked plant tissues. Cardenolides such as ouabain have their toxic effect by binding to the Na^+/K^+ -ATPase, inhibiting its function (Emery et al., 1998). Some insects manage to escape to the toxicity of these compounds and even use them in their own benefit and defence by structural modification in the affected enzyme. The PCR amplification and sequencing characterisation of the Na^+/K^+ -ATPase gene in several species of the eumolpine leaf beetle *Chrysochus* Redtenbacher feeding on plants producing toxic cardenolides has shown that they have a particular amino acid in the inferred protein sequence that was associated already to resistance to the cardenolide ouabain in the monarch butterfly [*Danaus plexippus* (Linnaeus); Holzinger & Wink, 1996]; the amino acid in this position is different in other species of *Chrysochus*, the eumolpine *Platycorynus sauteri* and *Drosophila*, feeding on plants or living in places devoid of cardenolides (Labeyrie & Dobler, 2004).

For many insecticides it is well known which specific protein is targeted disrupting the normal functioning and affecting the survival of the insect pests. Thus, the molecular characterisation of these genes and proteins affected by insecticides can help understanding the mechanisms behind insecticide resistance and devise alternative methods to overcome the resistance. Organophosphates and carbamate insecticides inhibit the acetylcholinesterase (*AChE*) in insects, enzyme in charge of terminating cholinergic synaptic transmission. Resistance to these insecticides has been verified to be the result of modification of this enzyme (see Oppenoorth, 1985). A cDNA of *AChE* was identified in an insecticide-susceptible strain of *L. decemlineata* and described as an ORF of 1887 nucleotides encoding for a protein of 629 amino acids, with the first 29 corresponding to the signal peptide. The deduced characteristics of the protein match those deduced applying other techniques, including the estimated molecular weight or the amino acid composition, and it was highly conserved compared with those in other insects (Zhu & Clark, 1995). The availability of this

information can be exploited to design primers to characterise at the molecular level the *AChE* gene in resistant strains of the beetle. Cytochrome P450 monooxygenases are important enzymes in all organisms and in insects they have been associated to insecticide resistance (Scott, 1999). In *D. virgifera virgifera*, resistance to organophosphates and carbamates is in part due to oxidative P450-based insecticide metabolism, but these enzymes participate in other insecticide degradation pathways, not necessarily involving detoxification. Several forms of P450s are present in insects (up to 86 functional genes in *D. melanogaster*; Adams et al., 2000) and Scharf et al. (2001) characterised three different forms previously undescribed from twenty successfully cloned cDNAs (*CYP4AJ1*, *CYP4G18* and *CYP4AK1*). These enzymes, confirmed to have higher levels of expression in resistant strains of western corn rootworm, probably corresponded to three different enzyme subfamilies given their relatively low amino acid identity and their comparisons with previously described P450s. In another study on *Leptinotarsa*, Lee et al. (1999) found that the resistance to pyrethroid insecticides in this beetle had a molecular mechanism identical to that previously described in the house fly by a single amino acid substitution in a sodium channel gene functionally similar to the *Vssc1* sodium channel gene in *Musca domestica* (Linnaeus).

The study of beetle resistance or defence mechanisms is not just restricted to insecticide resistance. A defensin-like antibacterial peptide was characterised from the cerambycid *Acalolepta luxuriosa* showing a typical cysteine-stabilised $\alpha\beta$ motif with a cysteine rich consensus motif; however, multiple sequence comparison with other available insect defensins showed that the one reported from this beetle constituted a new class (Saito et al., 2004). The inferred peptide sequence after modifications including cleavage of the signal peptide was named *A. luxuriosa* cysteine-rich peptide (AICRP). A phylogenetic analysis of AICRP with other insect and scorpion defensins showed a distant relationship with those and was consequently proposed as a novel peptide with a cysteine-stabilised $\alpha\beta$ motif (Saito et al., 2004).

THOUGHTS ABOUT THE FUTURE OF GENOMIC RESEARCH IN PHYTOPHAGA

Full-genome sequencing projects seem to be well justified when they target organisms that are of economical, medical or social importance. Therefore, it should be advisable to invest in the study of the genomes of some of the most injurious agricultural pests within the Phytophaga. The holistic approach to the study of the molecular biology of these insects from the analysis of their genomes should exponentially increase the benefits compared to the important advances achieved so far studying individual genes, as described above. Among other things, enhancing our knowledge about different genes, their structure and function, should open new paths for research complementary to those followed until now.

One alternative strategy for advancement in the study of beetle genomes would be to target a variety of tissues

in addition to those more commonly used, i.e., digestive tract or fat body. Tissue specific libraries from different organs should bear answers to precise metabolic, physiological or behavioural areas of research.

We have described how digestive enzymes have been proposed as good objectives for inhibition in pest control. However, since the beetles produce a wide range of digestive enzymes, the inhibition of a single gene is probably not deleterious to the insect, being necessary to target simultaneously other gene products, pheromones for instance (Eigenheer et al., 2003). Strategies based on comparative analysis of tissue specific libraries from different insects can assist in pinpointing those genes susceptible to be the target for insect control, because showing different expression patterns related to the specialization of the insects. In particular, Eigenheer et al. (2003) mention the example of pheromone biosynthetic genes in *Ips pini* and lipase-like ESTs in *Bombyx mori* (Linnaeus), the first not expressed in the silkworms and the second being more prevalent in the latter than in the pine bark beetle, compatible with the high fraction of lipid esters in the moth diet.

The study of signalling genes (pheromones) is indeed a promising area of research, particularly for pest control. Chemical communication in beetles has been particularly well studied in bark beetles, but studies at the molecular level are still in their initial stages, with a single published paper on the characterisation of genes potentially involved in the reception of pheromone signals (Nagnan-Le Meillour et al., 2004). In this study, the main sensorial organs of the beetles, the antennae, were the specific tissue used for the isolation of antennal proteins. These proteins showed similarity to previously identified insect odorant binding proteins, including those of other Coleoptera, mainly in the Scarabaeiformia (Nagnan-Le Meillour et al., 2004).

Phytophagous beetles show behaviour patterns that are appealing for genome-based research programs. Among these, cycloaexy (formation of larval defensive rings), association with ants, parental care and eusociality, courtship, and others, are specific behaviours whose genetic basis could be studied with a broader knowledge of the beetle genomes. The study of the molecular basis of behaviour has been attempted in insects already, and particularly in the honey bee, using a 20,000 cDNA clones tissue specific library from brain (Whitfield et al., 2002).

The application of "old" methods to new questions is one way to advance in the genomics of Phytophaga. Yet another possibility is to incorporate new technologies to this research. Two promising approaches are the use of DNA microarrays and the application of genome dissection techniques. DNA microarrays (or DNA chips) are collections of spatially ordered probes of known genes on a grid. Their most common use is to investigate and compare what genes are activated or repressed and/or what is their level of expression under particular experimental conditions (e.g., different physiological status, different tissues, different species, among others). Briefly, cDNA is isolated and labelled differentially with fluorescent

dyes (typically one green and one red) from each experimental setup. The labelled probes are mixed and hybridised with the DNA microarray which is scanned with laser beams of appropriate wave length to detect the hybridisation. Genes expressed in one of the experimental conditions will be detected in one colour, those particular of the alternative condition in the other colour, and those unaffected in the comparison in a mixed colour (e.g., yellow). DNA arraying techniques have been used successfully in Phytophaga beetles already to study the differential expression of digestive enzymes in *Leptinotarsa decemlineata* feeding on control and defence-induced leaves of potato plants (Gruden et al., 2004), and for the identification of scN responsive genes in *Diabrotica undecimpunctata* Mannerheim (Liu et al., 2004).

Microdissection of chromosomes or chromosome segments is another technology that can yield interesting results for the study of beetle genomes. A laser coupled to the light path of a microscope or a micromanipulator with microneedles can be used to isolate specific chromosomes from cell divisions onto a slide. Once a sufficient number of chromosomes are isolated, they can be used as template of unspecific DNA amplification methods and the obtained DNA used for library construction, probe design for in situ hybridisation analyses or chromosome mapping. These techniques are successfully used in other economically important groups, particularly farm animals and plants (e.g. Kubickova et al., 2002), and their implementation to beetles, at least those well known cytologically, easy to rear and important as pests or as models, is granted.

Although eminently related to phylogenetic studies, mitogenomics is another promising area of research to disentangle the evolutionary relationships within phytophagous beetles. Despite numerous efforts to find out the basal relationships within this lineage, they are still under controversy, and even its monophyly has not been proven beyond doubts. The sequence characterisation and phylogenetic analyses based on the mtDNA molecule for representative species of the major lineages within Phytophaga could help addressing this issue and possibly finding an objective solution. Beetles are the largest group of animals in the animal kingdom, and the Chrysomeloidea and Curculionoidea combined their most species-rich lineage. Their potential for speciation and adaptation has not yet been addressed from interconnected genomic and evolutionary perspectives. The study of a variety of genes in these beetles should provide insights into the general capabilities for adaptation of these insect lineages. Furthermore, comparative studies particularly of EST libraries on closely related beetle species should allow the characterisation of a number of rapidly evolving genes (orphan genes) putatively involved in the evolution of adaptive traits, therefore promoting speciation, as has been already reported in *Drosophila* (Domazet-Loso & Tautz, 2003). Population genomics, the analysis of variation in genome-wide sampling of markers and the identification of locus-specific effects (Black et al., 2001), is another promising approach

useful to identify the distribution and population dynamics of adaptive traits (Dicke et al., 2004). Similarly, the study through the implementation of molecular strategies of the interactions and performance of the beetle expressed genotypes against the environment and other elements of the ecosystem, typically host plants, but also predators, parasitoids or symbionts, provides evolutionary ecology with many opportunities for research (ecogenomics; Dicke et al., 2004).

The field is vast and the opportunities for research and the expected benefits are immense. The economical importance of some phytophagous beetles has allowed for an initial exploration of their genomes. This should provide a hotstart, together with the progressive implantation of more accessible and less expensive molecular techniques in our labs, to allow for similar and broader studies in other species with different levels of economic impact.

ACKNOWLEDGEMENTS. We are very grateful to F. Marec and two anonymous reviewers for their encouragement and suggestions to improve this work. This work has been supported by the EC Marie Curie ERG Programme Grant (Contract No. MERG-CT-2004-500004) and the project BOS2002-02870 from the Spanish Ministry of Science and Technology.

REFERENCES

- ADAMS M.D., CELNIKER S.E., HOLT R.A., EVANS C.A., GOCAYNE J.D., AMANATIDES P.G., SCHERER S.E., LI P.W., HOSKINS R.A., GALLE R.F. ET AL. 2000: The genome sequence of *Drosophila melanogaster*. *Science* **287**: 2185–2195.
- ALVAREZ N., AEBI A., RISTERUCCI A.-M., HOSSAERT-McKEY M. & BENREY B. 2003: Isolation and characterization of polymorphic microsatellite loci in *Acanthoscelides obvelatus* Bridwell (Coleoptera: Bruchidae). *Mol. Ecol. Notes* **3**: 12–14.
- ALVAREZ N., BORN C., RISTERUCCI A.-M., SOURROUILLE P., BENREY B. & HOSSAERT-McKEY M. 2004: Isolation and characterization of polymorphic microsatellite loci in *Acanthoscelides obtectus* Say (Coleoptera: Bruchidae). *Mol. Ecol. Notes* **4**: 683–685.
- ANDREEV D., BREILID H., KIRKENDALL L., BRUN L.O. & FRENCH-CONSTANT R.H. 1998: Lack of nucleotide variability in a beetle pest with extreme inbreeding. *Insect Mol. Biol.* **7**: 197–200.
- ANDREWS J., BOUFFARD G.G., CHEADLE C., LÜ J., BECKER K.G. & OLIVER B. 2000: Gene discovery using computational and microarray analysis of transcription in the *Drosophila melanogaster* testis. *Genome Res.* **10**: 2030–2043.
- ARNASON U. & JANKE A. 2002: Mitogenomic analyses of eutherian relationships. *Cytogenet. Genome Res.* **96**: 20–32.
- ARNASON U., ADEGOKE J.A., BODIN K., BORN E.W., ESA Y.B., GULLBERG A., NILSSON M., SHORT R.V., XU X. & JANKE A. 2002: Mammalian mitogenomic relationships and the root of the eutherian tree. *Proc. Natl. Acad. Sci. USA* **99**: 8151–8156.
- BERG P.R., DAWSON D.A., PANDHAL J., KIRKENDALL L.R. & BURKE T. 2003: Isolation and characterization of microsatellite loci from two inbreeding bark beetle species (Coccotrypes). *Mol. Ecol. Notes* **3**: 270–273.
- BLACK W.C., BAER C.F., ANTOLIN M.F. & DU TEAU N.M. 2001: Population genomics: Genome-wide sampling of insect populations. *Annu. Rev. Entomol.* **46**: 441–469.
- BLACKBURN E.H. 1991: Structure and function of telomeres. *Nature* **350**: 569–573.
- BONNETON F., ZELUS D., IWEMA T., ROBINSON-RECHAVI M. & LAUDET V. 2003: Rapid divergence of the ecdysone receptor in Diptera and Lepidoptera suggests coevolution between ECR and USP-RXR. *Mol. Biol. Evol.* **20**: 541–553.
- BOWN D.P., WILKINSON H.S., JONGSMA M.A. & GATEHOUSE J.A. 2004: Characterisation of cysteine proteinases responsible for digestive proteolysis in guts of larval western corn rootworm (*Diabrotica virgifera*) by expression in the yeast *Pichia pastoris*. *Insect Biochem. Mol. Biol.* **34**: 305–320.
- BRANSON T.F., SUTTER G.R. & FISHER J.R. 1980: Plant-response to stress-induced by artificial infestations of Western corn-rootworm (Coleoptera, Chrysomelidae). *Environ. Entomol.* **9**: 253–257.
- CAMPBELL B.C. 1989: On the role of microbial symbiotes in herbivorous insects. In Bernays E.A. (ed.): *Insect-Plant Interactions*. CRC Press, Boca Raton, pp. 1–44.
- CHO M.Y., CHOI H.W., MOON G.Y., KIM M.H., KWON T.H., HOMMA K., NATORI S. & LEE B.L. 1999: An 86 kDa diapause protein 1-like protein is a component of early-staged encapsulation-relating proteins in coleopteran insect, *Tenebrio molitor* larvae. *FEBS Lett.* **451**: 303–307.
- COGNATO A.I. & VOGLER A.P. 2001: Exploring data interaction and nucleotide alignment in a multiple gene analysis of *Ips* (Coleoptera: Scolytinae). *Syst. Biol.* **50**: 758–780.
- CROWSON R.A. 1981: *The Biology of the Coleoptera*. Academic Press, London, 802 pp.
- CUROLE J.P. & KOCHER T.D. 1999: Mitogenomics: digging deeper with complete mitochondrial genomes. *Trends Ecol. Evol.* **14**: 394–398.
- DE KORT C.A.D. & KOOPMANSCHAP A.B. 1994: Nucleotide and deduced amino acid sequence of a cDNA clone encoding diapause protein 1, an amylophorin-type storage hexamer of the Colorado potato beetle. *J. Insect Physiol.* **40**: 527–535.
- DENLINGER D.L., RINEHART J.P. & YOCUM G.D. 2001: Stress proteins: a role in insect diapause? In Denlinger D.L., Giebul-towicz J.M. & Saunders D.S. (eds): *Insect Timing: Circadian Rhythmicity to Seasonality*. Elsevier, Amsterdam, pp. 155–171.
- DHUYVETTER H., VERDYCK P., GAUBLOMME E., DESENDER K., MONDOR-GENSON G. & RASPLUS J.-Y. 2002: Isolation and characterization of microsatellite loci in the Galápagos Opuntia weevil *Gerstaeckeria galapagoensis* (Coleoptera, Curculionidae). *Mol. Ecol. Notes* **2**: 475–477.
- DICKE M., VAN LOON J.J.A. & DE JONG P.W. 2004: Ecogenomics benefits community ecology. *Science* **305**: 618–619.
- DING X., GOPALAKRISHNAN B., JOHNSON L.B., WHITE F.F., WANG X. & MUTHUKRISHNAN S. 1998: Insect resistance of transgenic tobacco expressing an insect chitinase gene. *Transgenic Res.* **7**: 77–84.
- DOMAZET-LOSO T. & TAUTZ D. 2003: An evolutionary analysis of orphan genes in *Drosophila*. *Genome Res.* **13**: 2213–2219.
- DOYLE K.E. & KNIPPLE D.C. 1991: PCR-based phylogenetic walking: Isolation of *para*-homologous sodium channel gene sequences from seven insect species and an arachnid. *Insect Biochem.* **21**: 689–696.
- DUCKETT C.N. & KJER K.M. 2003: Cladistic analysis of the Oedionychines of Southern Brazil (Galerucinae: Alticini) based on two molecular markers. In Furth D.G. (ed.): *Special Topics in Leaf Beetle Biology. Proceedings of the Fifth International Symposium on the Chrysomelidae*. Pensoft Publishers, Sofia, pp. 117–132.
- EDDY S.R. 2001: Non-coding RNA genes and the modern RNA world. *Nature Rev. Genet.* **2**: 919–929.
- EIGENHEER A.L., KEELING C.L., YOUNG S. & TITTIGER C. 2003: Comparison of gene representation in midguts from two phy-

- tophagous insects, *Bombyx mori* and *Ips pini*, using expressed sequence tags. *Gene* **316**: 127–136.
- EMERY A.M., BILLINGSLEY P.F., READY P.D. & DJAMGOZ M.B.A. 1998: Insect Na⁺/K⁺-ATPase. *J. Insect Physiol.* **44**: 197–209.
- FARRELL B.D., SEQUEIRA A.S., O'MEARA B.C., NORMARK B.B., CHUNG J.H. & JORDAL B.H. 2001: The evolution of agriculture in beetles (Curculionidae: Scolytinae and Platypodinae). *Evolution* **55**: 2011–2027.
- FROHMANN M.A., DUSH M.K. & MARTIN G.R. 1988: Rapid production of full-length cDNAs from rare transcripts: amplification using a single gene-specific oligonucleotide primer. *Proc. Natl. Acad. Sci. USA* **85**: 8998–9002.
- FRYDRYCHOVÁ R. & MAREC F. 2002: Repeated losses of TTAGG telomere repeats in evolution of beetles (Coleoptera). *Genetica* **115**: 179–187.
- GAUTHIER N. & RASPLUS J.-Y. 2004: Polymorphic microsatellite loci in the Coffee Berry Borer, *Hypothenemus hampei* (Coleoptera, Scolytidae). *Mol. Ecol. Notes* **4**: 294–296.
- GILLIKIN J.W., BEVILACQUA S. & GRAHAM J.S. 1992: Partial characterization of digestive tract proteinases from Western Corn Rootworm larvae, *Diabrotica virgifera*. *Arch. Insect Biochem. Physiol.* **19**: 285–298.
- GIRARD C. & JOUANIN L. 1999a: Molecular cloning of cDNAs encoding a range of digestive enzymes from a phytophagous beetle, *Phaedon cochleariae*. *Insect Biochem. Mol. Biol.* **29**: 1129–1142.
- GIRARD C. & JOUANIN L. 1999b: Molecular cloning of a gut-specific chitinase cDNA from the beetle *Phaedon cochleariae*. *Insect Biochem. Mol. Biol.* **29**: 549–556.
- GOLDSTEIN D.B. & SCHLÖTTERER C. 1999: *Microsatellites. Evolution and Applications*. Oxford University Press, New York, 352 pp.
- GÓMEZ-ZURITA J., KOPLIKU F., THEODORIDES K. & VOGLER A.P. 2004: Resources for a phylogenomic approach in leaf beetle (Coleoptera) systematics. In Jolivet P., Santiago-Blay J.A. & Schmitt M. (eds): *New Developments in the Biology of Chrysomelidae*. SPB Academic Publishing, The Hague, pp. 19–35.
- GROSSI-DE-SÁ M.F. & CHRISPPEELS M.J. 1997: Molecular cloning of bruchid (*Zabrotes subfasciatus*) alpha-amylase cDNA and interactions of the expressed enzyme with bean amylase inhibitors. *Insect Biochem. Mol. Biol.* **27**: 271–281.
- GRUDEN K., POPOVIĆ T., CIMERMAN N., KRIŽAJ I. & ŠTRUKELJ B. 2003: Diverse enzymatic specificities of digestive proteinases, 'intestains', enable Colorado potato beetle larvae to counter the potato defense mechanism. *Biol. Chem.* **384**: 305–310.
- GRUDEN K., KUIPERS A.G.J., GUNÉAR G., SLAPAR N., ŠTRUKELJ B. & JONGSMA M.A. 2004: Molecular basis of Colorado potato beetle adaptation to potato plant defence at the level of digestive cysteine proteinases. *Insect Biochem. Mol. Biol.* **34**: 365–375.
- HALL G.M., TITTIGER C., ANDREWS G.L., MASTICK G.S., KUENZLI M., LUO X., SEYBOLD S.J. & BLOMQUIST G.J. 2002: Midgut tissue of male pine engraver, *Ips pini*, synthesizes monoterpene pheromone component ipsdienol de novo. *Naturwissenschaften* **89**: 79–83.
- HAMMOND P.M. 1992: Species inventory. In Groombridge B. (ed.): *Global Biodiversity, Status of the Earth's Living Resources*. Chapman and Hall, London, pp. 17–39.
- HAWTHORNE D.J. 2001: AFLP-based genetic linkage map of the Colorado potato beetle *Leptinotarsa decemlineata*: sex chromosomes and a pyrethroid-resistance candidate gene. *Genetics* **158**: 695–700.
- HAWTHORNE D.J. 2003: Quantitative trait locus mapping of pyrethroid resistance in Colorado potato beetle, *Leptinotarsa decemlineata* (Say) (Coleoptera: Chrysomelidae). *J. Econ. Entomol.* **96**: 1021–1030.
- HECKEL D.G. 2003: Genomics in pure and applied entomology. *Annu. Rev. Entomol.* **48**: 235–260.
- HEDDI A., VALLIER A., ANSELME C., XIN H.W., RAHBE Y. & WACKERS F. 2005: Molecular and cellular profiles of insect bacteriocytes: mutualism and harm at the initial evolutionary step of symbiogenesis. *Cell. Microbiol.* **7**: 293–305.
- HENIKOFF S., AHMAD K. & MALIK H.S. 2001: The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* **293**: 1098–1102.
- HENRICH V.C. & BROWN N.E. 1995: Insect nuclear receptors: a developmental and comparative perspective. *Insect Biochem. Mol. Biol.* **25**: 881–897.
- HOLZINGER F. & WINK M. 1996: Mediation of cardiac glycoside insensitivity in the monarch butterfly (*Danaus plexippus*): role of an amino acid substitution in the ouabain binding site of Na⁺, K⁺-ATPase. *J. Chem. Ecol.* **22**: 1921–1937.
- HSIAO T.H. & HSIAO C. 1983: Chromosomal analysis of *Leptinotarsa* and *Labidomera* species (Coleoptera: Chrysomelidae). *Genetica* **60**: 139–150.
- HUGHES J. & VOGLER A.P. 2004: The phylogeny of acorn weevils (genus *Curculio*) from mitochondrial and nuclear DNA sequences: the problem of incomplete data. *Mol. Phylogenet. Evol.* **32**: 601–615.
- JOLIVET P. & VERMA K.K. 2002: *Biology of Leaf Beetles*. Intercept Publ., Andover, 332 pp.
- JORDAL B.H. 2002: Elongation Factor 1 alpha resolves the monophyly of the haplodiploid ambrosia beetles Xyleborini (Coleoptera: Curculionidae). *Insect Mol. Biol.* **11**: 453–465.
- JORDAL B.H. & HEWITT G.M. 2004: The origin and radiation of Macaronesian beetles breeding in Euphorbia: the relative importance of multiple data partitions and population sampling. *Syst. Biol.* **53**: 711–734.
- JORDAL B.H., NORMARK B.B. & FARRELL B.D. 2000: Evolutionary radiation of an inbreeding haplodiploid beetle lineage (Curculionidae, Scolytinae). *Biol. J. Linn. Soc. Lond.* **71**: 483–499.
- JORDAL B.H., BEAVER R.A., NORMARK B.B. & FARRELL B.D. 2002a: Extraordinary sex ratios and the evolution of male neoteny in sib-mating Ozopemon beetles. *Biol. J. Linn. Soc. Lond.* **75**: 353–360.
- JORDAL B.H., NORMARK B.B., FARRELL B.D. & KIRKENDALL L.R. 2002b: Extraordinary haplotype diversity in haplodiploid inbreeders: phylogenetics and evolution of the bark beetle genus *Coccotrypes*. *Mol. Phylogenet. Evol.* **23**: 171–188.
- JORDAL B.H., KIRKENDALL L.R. & HARKESTAD K. 2004: Phylogeny of a Macaronesian radiation: host-plant use and possible cryptic speciation in *Liparthrum* bark beetles. *Mol. Phylogenet. Evol.* **31**: 554–571.
- KERDELHUÉ C., MONDOR-GENSON G., RASPLUS J.-Y., ROBERT A. & LIEUTIER F. 2003: Characterization of five microsatellite loci in the pine shoot beetle *Tomicus piniperda* (Coleoptera: Scolytidae). *Mol. Ecol. Notes* **3**: 100–101.
- KIM K.S. & SAPPINGTON T.W. 2005: Polymorphic microsatellite loci from the western corn rootworm (Insecta: Coleoptera: Chrysomelidae) and cross-amplification with other *Diabrotica* spp. *Mol. Ecol. Notes* **5**: 115–117.
- KIM S.R., YOON H.J., PARK N.S., LEE S.M., MOON J.Y., JIN B.R. & SOHN H.D. 2001: Molecular cloning of a cDNA encoding a cathepsin D homologue from the mulberry longicorn beetle, *Apriona germari*. *Int. J. Indust. Entomol.* **3**: 121–126.
- KIM S., KJER K.M. & DUCKETT C.N. 2003a: Comparison between molecular and morphological-based phylogenies of galerucine/alticine leaf beetles (Coleoptera: Chrysomelidae). *Insect Syst. Evol.* **34**: 53–64.

- KIM S.R., YOON H.J., LEE K.S., KIM I., JE Y.H., SOHN H.D. & JIN B.R. 2003b: Molecular cloning of three cDNAs encoding putative larval cuticle protein expressed differentially after larval ecdysis from the mulberry longicorn beetle, *Apriona germari*. *Comp. Biochem. Physiol. (B) Biochem. Mol. Biol.* **136**: 803–811.
- KIM S.R., YOON H.J., PARK N.S., LEE S.M., MOON J.Y., SEO S.J., JIN B.R. & SOHN H.D. 2003c: Molecular cloning, expression, and characterization of a cDNA encoding the arylphorin-like hexameric storage protein from the mulberry longicorn beetle, *Apriona germari*. *Arch. Insect Biochem. Physiol.* **53**: 49–65.
- KIM S.R., LEE K.S., YOON H.J., PARK N.S., LEE S.M., JE Y.H., JIN B.R. & SOHN H.D. 2004: cDNA cloning, expression, and characterization of an arylphorin-like hexameric storage protein, AgeHex2, from the mulberry longicorn beetle, *Apriona germari*. *Arch. Insect Biochem. Physiol.* **56**: 61–72.
- KOIWA H., SHADE R.E., ZHU-SALZMAN K., PAINO D'URZO M., MURDOCK L.L., BRESSAN R.A. & HASEGAWA P.M. 2000: A plant defensive cystatin (soyacystatin) targets cathepsin L-like digestive cysteine proteinases (DvCALs) in the larval midgut of western corn rootworm (*Diabrotica virgifera virgifera*). *FEBS Lett.* **471**: 67–70.
- KOOPMANSCHAP A.B., LAMMERS J.H.M. & DE KORT C.A.D. 1995: The structure of the gene encoding diapause protein 1 of the Colorado potato beetle (*Leptinotarsa decemlineata*). *J. Insect Physiol.* **41**: 509–518.
- KOZAK M. 1983: Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. *Microbiol. Rev.* **47**: 1–45.
- KRAUSS V. & REUTER G. 2000: Two genes become one: the genes encoding heterochromatin protein Su(var)3-9 and translation initiation factor subunit eIF-2gamma are joined to a dicistronic unit in holometabolic insects. *Genetics* **156**: 1157–1167.
- KUBICKOVA S., CERNOHORSKA H., MUSILOVA P. & RUBES J. 2002: The use of laser microdissection for the preparation of chromosome-specific painting probes in farm animals. *Chromosome Res.* **10**: 571–577.
- LABEYRIE E. & DOBLER S. 2004: Molecular adaptation of *Chrysomus* leaf beetles to toxic compounds in their food plants. *Mol. Biol. Evol.* **21**: 218–221.
- LAWRENCE J.F. 1982: Coleoptera. In Parker S.P. (ed): *Synopsis and Classification of Living Organisms. Vol. 2*. McGraw-Hill, New York, pp. 482–553.
- LEAL I., WHITE E.E., SAHOTA T.S. & MANVILLE J.F. 1997: Differential expression of the vitellogenin gene in the spruce terminal weevil feeding on resistant versus susceptible host trees. *Insect Biochem. Mol. Biol.* **27**: 569–575.
- LEE S.H., DUNN J.B., CLARK J.M. & SODERLUND D.M. 1999: Molecular analysis of *kdr*-like resistance in a permethrin-resistant strain of Colorado potato beetle. *Pesticide Biochem. Physiol.* **63**: 63–75.
- LEE S.J., KIM S.R., YOON H.J., KIM I., LEE K.S., JE Y.H., LEE S.M., SEO S.J., SOHN H.D. & JIN B.R. 2004: cDNA cloning, expression, and enzymatic activity of a cellulase from the mulberry longicorn beetle, *Apriona germari*. *Comp. Biochem. Physiol. (B) Biochem. Mol. Biol.* **139**: 107–116.
- LEE S.J., LEE K.S., KIM S.R., GUI Z.Z., KIM Y.S., YOON H.J., KIM I., KANG P.D., SOHN H.D. & JIN B.R. 2005: A novel cellulase gene from the mulberry longicorn beetle, *Apriona germari*: gene structure, expression, and enzymatic activity. *Comp. Biochem. Physiol. (B) Biochem. Mol. Biol.* **140**: 551–560.
- LEPLÉ J.C., BONADÉ-BOTTINO M., AUGUSTIN S. & PILATE G. 1995: Toxicity to *Chrysomela tremulae* (Coleoptera: Chrysomelidae) of transgenic poplars expressing a cysteine proteinase inhibitor. *Mol. Breeding* **1**: 319–328.
- LEWIS D.K., SPURGEON D., SAPPINGTON T.W. & KEELEY L.L. 2002: A hexamerin protein, AgSP-1, is associated with diapause in the boll weevil. *J. Insect Physiol.* **48**: 887–901.
- LIEWLAKSANEYANAWIN C., RITLAND C.E., NEWTON C.H. & EL-KASSABY Y.A. 2001: Characterization of microsatellite loci in white pine weevil (*Pissodes strobi*). *Mol. Ecol. Notes* **1**: 248–249.
- LIU Y., SALZMAN R.A., PANKIW T. & ZHU-SALZMAN K. 2004: Transcriptional regulation in southern corn rootworm larvae challenged by soyacystatin N. *Insect Biochem. Mol. Biol.* **34**: 1069–1077.
- LORITE P., PALOMEQUE T., GARNERÍA I. & PETITPIERRE E. 2001: Characterization and chromosome location of satellite DNA in the leaf beetle *Chrysolina americana* (Coleoptera, Chrysomelidae). *Genetica* **110**: 143–150.
- LORITE P., CARRILLO J.A., GARNERÍA I., PETITPIERRE E. & PALOMEQUE T. 2002: Satellite DNA in the elm leaf beetle, *Xanthogaleruca luteola* (Coleoptera, Chrysomelidae): characterization, interpopulation analysis, and chromosome location. *Cytogenet. Genome Res.* **98**: 302–307.
- MATSUMOTO I., EMORI Y., ABE K. & ARAI S. 1997: Characterization of a gene family encoding cysteine proteinases of *Sitophilus zeamais* (maize weevil), and analysis of the protein distribution in various tissues including alimentary tract and germ cells. *J. Biochem.* **121**: 464–476.
- MISRA S., CROSBY M.A., MUNGALL C.J., MATTHEWS B.B., CAMPBELL K.S., HRADECKY P., HUANG Y., KAMINKER J.S., MILLBURN G.H., PROCHNIK S.E., SMITH C.D., TUPY J.L., WHITFIELD E.J., BAYRAKTAROGU L., BERMAN B.P., BETTENCOURT B.R., CELNIKER S.E., DE GREY A.D.N.J., DRYSDALE R.A., HARRIS N.L., RICHTER J., RUSSO S., SCHROEDER A.J., SHU S.-Q., STAPLETON M., YAMADA C., ASHBURNER M., GELBART W.M., RUBIN G.M. & LEWIS S.E. 2002: Annotation of the *Drosophila melanogaster* euchromatic genome: a systematic review. *Genome Biol.* **3**: 1–22.
- MOON J., SALZMAN R.A., AHN J.-E., KOIWA H., ZHU-SALZMAN & K. 2004: Transcriptional regulation in cowpea bruchid guts during adaptation to a plant defence protease inhibitor. *Insect Mol. Biol.* **13**: 283–291.
- NAGNAN-LE MEILLOR P., FRANÇOIS M.-C. & JACQUIN-JOLY E. 2004: Identification and molecular cloning of putative odorant-binding proteins from the American palm weevil, *Rhynchophorus palmarum* L. *J. Chem. Ecol.* **30**: 1213–1223.
- NORMARK B.B., JORDAL B.H. & FARRELL B.D. 1999: Origin of a haplodiploid beetle lineage. *Proc. R. Soc. Lond. (B) Biol. Sci.* **266**: 2253–2259.
- OKAZAKI S., TSUCHIDA K., MAEKAWA H., ISHIKAWA H. & FUJIWARA H. 1993: Identification of a pentanucleotide telomeric sequence, (TTAGG)_n, in the silkworm *Bombyx mori* and in other insects. *Mol. Cell. Biol.* **13**: 1424–1432.
- OLIVEIRA-NETO O.B. DE, BATISTA J.A.N., RIGDEN D.J., FRANCO O.L., FALCÃO R., FRAGOSO R.R., MELLO L.V., DOS SANTOS R.C. & GROSSI-DE-SÁ M.F. 2003: Molecular cloning of alpha-amylases from cotton boll weevil, *Anthonomus grandis* and structural relations to plant inhibitors: An approach to insect resistance. *J. Protein Chem.* **22**: 77–87.
- OLIVEIRA-NETO O.B. DE, BATISTA J.A., RIGDEN D.J., FRANCO O.L., FRAGOSO R.R., MONTEIRO A.C., MONNERAT R.G. & GROSSI-DE-SÁ M.F. 2004a: Molecular cloning of a cysteine proteinase cDNA from the cotton boll weevil *Anthonomus grandis* (Coleoptera: Curculionidae). *Biosci. Biotechnol. Biochem.* **68**: 1235–1242.
- OLIVEIRA-NETO O.B., BATISTA J.A.N., RIGDEN D.J., FRAGOSO R.R., SILVA R.O., GOMES E.A., FRANCO O.L., DIAS S.C., COR-

- DEIRO C.M.T., MONNERAT R.G. & GROSSI-DE-SÁ M.F. 2004b: A diverse family of serine proteinase genes expressed in cotton boll weevil (*Anthonomus grandis*): implications for the design of pest-resistant transgenic cotton plants. *Insect Biochem. Mol. Biol.* **34**: 903–918.
- OPPENORTH F.J. 1985: Biochemistry and genetics of insecticide resistance. In Kerkut G.A. & Gilbert L.I. (eds): *Comprehensive Insect Physiology, Biochemistry and Pharmacology*. Vol. 12. Pergamon Press, New York, pp. 731–773.
- PARK K.S., BAE Y.J., YEAO S.H. & KANG S.J. 2001: Cloning and sequence analysis of Gpdh in *Callosobruchus chinensis* (Coleoptera: Bruchidae). *Mol. Cells* **11**: 405–410.
- PATT A., MISOF B., WAGNER T.H. & NAUMANN C.M. 2004a: Characterization of microsatellite loci in *Amphimetis transversus* (Kolbe, 1897) (Coleoptera, Curculionidae). *Mol. Ecol. Notes* **4**: 188–190.
- PATT A., MISOF B., WAGNER T. & NAUMANN C.M. 2004b: Isolation and characterization of microsatellite loci in *Monolepta vincta* Gerstaecker, 1871 (Coleoptera, Chrysomelidae, Galeucinae). *Mol. Ecol. Notes* **4**: 572–574.
- PEDRA J.H.F., BRANDT A., WESTERMAN R., LOBO N., LI H.M., ROMERO-SEVERSON J., MURDOCK L.L. & PITTENDRIGH B.R. 2003: Transcriptome analysis of the cowpea weevil bruchid: identification of putative proteinases and alpha-amylases associated with food breakdown. *Insect Mol. Biol.* **12**: 405–412.
- ROBERTSON H.M. 1993: The *mariner* transposable element is widespread in insects. *Nature* **362**: 241–245.
- ROBERTSON H.M., SOTO-ADAMES F.N., WALDEN K.K.O., AVANCINI R.M.P. & LAMPE D.J. 1998: The *mariner* transposons of animals: horizontally jumping genes. In Syvanen M. & Kado C. (eds): *Horizontal Gene Transfer*. Chapman and Hall, London, pp. 268–284.
- ROŽEK M., LACHOWSKA D., PETITPIERRE E. & HOLECOVÁ M. 2004: C-bands on chromosomes of 32 beetle species (Coleoptera: Elateridae, Cantharidae, Oedemeridae, Cerambycidae, Anthicidae, Chrysomelidae, Attelabidae and Curculionidae). *Hereditas* **140**: 161–170.
- SAHARA K., MAREC F. & TRAUT W. 1999: TTAGG telomeric repeats in chromosomes of some insects and other arthropods. *Chromosome Res.* **7**: 449–460.
- SAITO A., UEDA K., IMAMURA M., MIURA N., ATSUMI S., TABUNOKI H. & SATO R. 2004: Purification and cDNA cloning of a novel antibacterial peptide with a cysteine-stabilized ab motif from the longicorn beetle, *Acalolepta luxuriosa*. *Dev. Comp. Immunol.* **28**: 1–7.
- SALLÉ A., KERDELHUÉ 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.
- SALZBERG S.L., WHITE O., PETERSON J. & EISEN J.A. 2001: Microbial genes in the human genome: Lateral transfer or gene loss? *Science* **292**: 1903–1906.
- SCHARF M.E., PARIMI S., MEINKE L.J., CHANDLER L.D. & SIEGFRIED B.D. 2001: Expression and induction of three family 4 cytochrome P450 (CYP4)* genes identified from insecticide-resistant and susceptible western corn rootworms, *Diabrotica virgifera virgifera*. *Insect Mol. Biol.* **10**: 139–146.
- SCHROEDER H.E., GOLLASCH S., MOORE A., TABE L.M., CRAIG S., HARDIE D., CHRISPEELS M.J., SPENCER D. & HIGGINS T.J.V. 1995: Bean α -amylase inhibitor confers resistance to the pea weevil (*Bruchus pisorum*) in transgenic peas (*Pisum sativum* L.). *Plant Physiol.* **107**: 1233–1239.
- SCOTT J.G. 1999: Cytochromes P450 and insecticide resistance. *Insect Biochem. Mol. Biol.* **29**: 757–777.
- SEMBÈNE M., VAUTRIN D., SILVAIN J.F., RASPLUS J.Y. & DELOBEL A. 2003: Isolation and characterization of polymorphic microsatellites in the groundnut seed beetle, *Caryedon serratus* (Coleoptera, Bruchidae). *Mol. Ecol. Notes* **3**: 299–301.
- SEQUEIRA A.S. & FARRELL B.D. 2001: Evolutionary origins of Gondwanan interactions: How old are Araucaria beetle herbivores? *Biol. J. Linn. Soc. Lond.* **74**: 459–474.
- SEQUEIRA A.S., NORMARK B.B. & FARRELL B.D. 2000: Evolutionary assembly of the conifer fauna: distinguishing ancient from recent associations in bark beetles. *Proc. R. Soc. Lond. (B) (Biol. Sci.)* **267**: 2359–2366.
- SHADE R.E., SCHROEDER H.E., PUEYO J.J., TABE L.M., HIGGINS T.J.V. & CHRISPEELS M.J. 1994: Transgenic pea seeds expressing the α -amylase inhibitor of the common bean are resistant to bruchid beetles. *Bio/Technology* **12**: 793–796.
- SHEN Z., DENTON M., MUTTI N., PAPPAN K., KANOST M.R., REESE J.C. & REECK G.R. 2003: Polygalacturonase from *Sitophilus oryzae*: Possible horizontal transfer of a pectinase gene from fungi to weevils. *J. Insect Sci.* **3**: 1–9.
- SIEGFRIED B.D., WATERFIELD N. & FRENCH-CONSTANT R.H. 2005: Expressed sequence tags from *Diabrotica virgifera virgifera* midgut identify a coleopteran cadherin and a diversity of cathepsins. *Insect Mol. Biol.* **14**: 137–143.
- SMID H.M., KOOPMANSCHAP A.B., DE KORT C.A.D. & SCHOONEVELD H. 1997: A peptide from the male accessory gland in *Leptinotarsa decemlineata*: Purification, characterization and molecular cloning. *J. Insect Physiol.* **43**: 355–362.
- SMITH A.F., OWEN L.M., STROBEL L.M., CHEN H., KANOST M.R., HANNEMAN E. & WELLS M.A. 1994: Exchangeable apolipoproteins of insects share a common structural motif. *J. Lipid Res.* **35**: 1976–1984.
- SOUTHGATE B.J. 1979: Biology of the Bruchidae. *Annu. Rev. Entomol.* **24**: 449–473.
- STEWART J.B. & BECKENBACH A.T. 2003: Phylogenetic and genomic analysis of the complete mitochondrial DNA sequence of the spotted asparagus beetle *Crioceris duodecimpunctata*. *Mol. Phylogenet. Evol.* **26**: 513–526.
- SUGIMURA M., WATANABE H., LO N. & SAITO H. 2003: Purification, characterization, cDNA cloning and nucleotide sequencing of a cellulase from the yellow-spotted longicorn beetle, *Psacotha hilaris*. *Eur. J. Biochem.* **270**: 3455–3460.
- TANAKA H., SATO K., SAITO Y., YAMASHITA T., AGOH M., OKUNISHI J., TACHIKAWA E. & SUZUKI K. 2003: Insect diapause-specific peptide from the leaf beetle has consensus with a putative iridovirus peptide. *Peptides* **24**: 1327–1333.
- TELFER W.H. & KUNKEL J.G. 1991: The function and evolution of insect storage hexamers. *Annu. Rev. Entomol.* **36**: 205–228.
- TERMONIA A., PASTEELS J.M., WINDSOR D.M. & MILINKOVITCH M.C. 2002: Dual chemical sequestration: a key mechanism in transitions among ecological specialization. *Proc. R. Soc. Lond. (B)* **269**: 1–6.
- THEODORIDES K., DE RIVA A., GÓMEZ-ZURITA J., FOSTER P.G. & VOGLER A.P. 2002: Comparison of EST libraries from seven beetle species: towards a framework for phylogenomics of the Coleoptera. *Insect Mol. Biol.* **11**: 467–475.
- TITARENKO E. & CHRISPEELS M.J. 2000: cDNA cloning, biochemical characterization and inhibition by plant inhibitors of the α -amylases of the Western corn rootworm, *Diabrotica virgifera virgifera*. *Insect Biochem. Mol. Biol.* **30**: 979–990.
- TITTIGER C., BLOMQUIST G.J., IVARSSON P., BORGESON C.E. & SEYBOLD S.J. 1999: Juvenile hormone regulation of HMG-R gene expression in the bark beetle *Ips paraconfusus* (Coleoptera: Scolytidae): implications for male aggregation pheromone biosynthesis. *Cell. Mol. Life Sci.* **55**: 121–127.

- TITTIGER C., O'KEEFFE C., BENGGA C.S., BARKAWI L.S., SEYBOLD S.J. & BLOMQUIST G.J. 2000: Isolation and endocrine regulation of an HMG-CoA synthase cDNA from the male Jeffrey pine beetle, *Dendroctonus jeffreyi* (Coleoptera: Scolytidae). *Insect Biochem. Mol. Biol.* **30**: 1203–1211.
- TITTIGER C., BARKAWI L.S., BENGGA C.S., BLOMQUIST G.J. & SEYBOLD S.J. 2003: Structure and juvenile hormone-mediated regulation of the HMG-CoA reductase gene from the Jeffrey pine beetle, *Dendroctonus jeffreyi*. *Mol. Cell. Endocrinol.* **199**: 11–21.
- TREWITT P.M., HEILMANN L.J. & KUMARAN A.K. 1990: Boll weevil testis-specific cDNA. *Nucleic Acids Res.* **18**: 3646.
- TREWITT P.M., HEILMANN L.J., DEGRUGILLIER S.S. & KUMARAN A.K. 1992: The boll weevil vitellogenin gene: nucleotide sequence, structure, and evolutionary relationship to nematode and vertebrate vitellogenin genes. *J. Mol. Evol.* **34**: 478–492.
- TSCHIRSCH B., HOFMANN A., KRAUSS V., DORN R., KORGE G. & REUTER G. 1994: The protein encoded by the *Drosophila* position-effect variegation suppressor gene *Su(var)3-9* combines domains of antagonistic regulators of homeotic gene complexes. *EMBO J.* **13**: 3822–3831.
- VERMUNT A.M., KOOPMANSCHAP A.B., VLAK J.M. & DE KORT C.A. 1997: Cloning and sequence analysis of cDNA encoding a putative juvenile hormone esterase from the Colorado potato beetle. *Insect Biochem. Mol. Biol.* **27**: 919–928.
- VERMUNT A.M., KOOPMANSCHAP A.B., VLAK J.M. & DE KORT C.A. 1998: Evidence for two juvenile hormone esterase-related genes in the Colorado potato beetle. *Insect Mol. Biol.* **7**: 327–336.
- VOS P., HOGERS R., BLEEKER M., REIJANS M., VAN DE LEE T., HORNES M., FRIJTERS A., POT J., PELEMAN J., KUIPER M. & ZABEAU M. 1995: AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res.* **23**: 4407–4414.
- WATANABE H. & TOKUDA G. 2001: Animal cellulases. *Cell. Mol. Life Sci.* **58**: 1167–1178.
- WHITFIELD C.W., BAND M.R., BONALDO M.F., KUMAR C.G., LIU L., PARDINAS J.R., ROBERTSON H.M., SOARES M.B. & ROBINSON G.E. 2002: Annotated expressed sequence tags and cDNA microarrays for studies of brain and behavior in the honey bee. *Gen. Res.* **12**: 555–566.
- WILHITE S.E., ELDEN T.C., BRZIN J. & SMIGOCKI A.C. 2000: Inhibition of cysteine and aspartyl proteinases in the alfalfa weevil midgut with biochemical and plant-derived proteinase inhibitors. *Insect Biochem. Mol. Biol.* **30**: 1181–1188.
- YAN X.H., DE BONDT H.L., POWELL C.C., BULLOCK R.C. & BOROVSKY D. 1999: Sequencing and characterization of the citrus weevil, *Diaprepes abbreviatus*, trypsin cDNA. Effect of *Aedes* trypsin modulating oostatic factor on trypsin biosynthesis. *Eur. J. Biochem.* **262**: 627–636.
- YAO T.P., FORMAN B.M., JIANG Z., CHERBAS L., CHEN J.D., MCKEOWN M., CHERBAS P. & EVANS R.M. 1993: Functional ecdysone receptor is the product of *EcR* and *ultraspiracle* genes. *Nature* **366**: 476–479.
- YOCUM G.D. 2001: Differential expression of two *HSP70* transcripts in response to cold shock, thermoperiod, and adult diapause in the Colorado potato beetle. *J. Insect Physiol.* **47**: 1139–1145.
- YOCUM G.D. 2003: Isolation and characterization of three diapause-associated transcripts from the Colorado potato beetle, *Leptinotarsa decemlineata*. *J. Insect Physiol.* **49**: 161–169.
- YOCUM G.D. 2004: Environmental regulation of the purine synthesis enzyme *purH* transcript during adult diapause in *Leptinotarsa decemlineata* (Coleoptera: Chrysomelidae). *Eur. J. Entomol.* **101**: 199–203.
- YOON H.J., KIM S.R., JIN B.R., LEE S.M., MOON J.Y., MAH Y.I. & SOHN H.D. 2001: Purification and characterization of storage proteins from the mulberry longicorn beetle, *Apriona germari* Hope. *Int. J. Indust. Entomol.* **2**: 61–66.
- ZHU K.Y. & CLARK J.M. 1995: Cloning and sequencing of a cDNA encoding acetylcholinesterase in Colorado potato beetle, *Leptinotarsa decemlineata* (Say). *Insect Biochem. Mol. Biol.* **25**: 1129–1138.
- ZHU-SALZMAN K., KOIWA H., SALZMAN R.A., SHADE R.E. & AHN J.E. 2003: Cowpea bruchid *Callosobruchus maculatus* uses a three-component strategy to overcome a plant defensive cysteine protease inhibitor. *Insect Mol. Biol.* **12**: 135–145.

Received May 5, 2005; revised and accepted June 15, 2005