Appendix A. Genome resources of all 131 invasive species.

Species Name	**GenBank ID**	**RefSeq ID**	**Data Source**
Acanthaster planci	GCA_001949145.1	GCF_001949145.1	NCBI
Acyrthosiphon pisum	GCA_000142985.2	GCF_000142985.2	AphidBase
Aedes aegypti	GCA_002204515.1	GCF_002204515.2	VectorBase
Aedes albopictus	GCA_001876365.2	GCF_001876365.2	VectorBase
Aethina tumida	GCA_001937115.1	GCF_001937115.1	NCBI
Agrilus planipennis	GCA_000699045.1	GCF_000699045.1	InsectBase
Alnus glutinosa	GCA_003254965.1	n/a	NCBI
Amaranthus tuberculatus	GCA_000180655.1	n/a	NCBI
Ammotragus lervia	GCA_002201775.1	n/a	NCBI
Anopheles stephensi			VectorBase
Anoplophora glabripennis	GCA_000390285.2	GCF_000390285.2	InsectBase
Aphis glycines			AphidBase
Apis florea	GCA_000184785.1	GCF_000184785.2	InsectBase
Apis mellifera	GCA_000002195.1	GCF_000002195.4	InsectBase
Asclepias syriaca	GCA_002018285.1	n/a	NCBI
Azadirachta indica	GCA_000439995.3	n/a	NCBI
Bactrocera cucurbitae	GCA_000806345.1	GCF_000806345.1	InsectBase
Bactrocera dorsalis	GCA_000789215.2	GCF_000789215.1	InsectBase
Bactrocera latifrons	GCA_001853355.1	GCF_001853355.1	i5K
Bactrocera oleae	GCA_001188975.2	GCF_001188975.1	i5K
Bactrocera tryoni	GCA_000695345.1	n/a	InsectBase
Bemisia tabaci	GCA_001854935.1	GCF_001854935.1	WhiteflyDB
Berberis thunbergii	GCA_003290165.1	n/a	NCBI
Blattella germanica	GCA_003018175.1	n/a	InsectBase
Bombus impatiens	GCA_000188095.2	GCF_000188095.1	InsectBase
Bombus terrestris	GCA_000214255.1	GCF_000214255.1	InsectBase
Bos taurus	GCA_000003205.6	GCF_000003205.7	NCBI
Brachypodium distachyon	GCA_000005505.4	GCF_000005505.3	NCBI
Brassica nigra	GCA_001682895.1	n/a	NCBI
Bursaphelenchus xylophilus	GCA_000231135.1	n/a	NCBI
Capra hircus	GCA_001704415.1	GCF_001704415.1	NCBI
Cardiocondyla obscurior			Fourmidable
Castor canadensis	GCA_001984765.1	GCF_001984765.1	NCBI
Casuarina glauca	GCA_003255045.1	n/a	NCBI
Cephus cinctus	GCA_000341935.1	GCF_000341935.1	InsectBase
Ceratitis capitata	GCA_000347755.4	GCF_000347755.3	InsectBase
Cervus elaphus	GCA_002197005.1	n/a	NCBI
Cimex lectularius	GCA_000648675.1	GCF_000648675.1	InsectBase
Ciona intestinalis	GCA_000224145.2	GCF_000224145.3	NCBI
Ciona savignyi	GCA_000149265.1	n/a	NCBI
Conyza canadensis	GCA_000775935.1	n/a	NCBI

Continued Appendix A

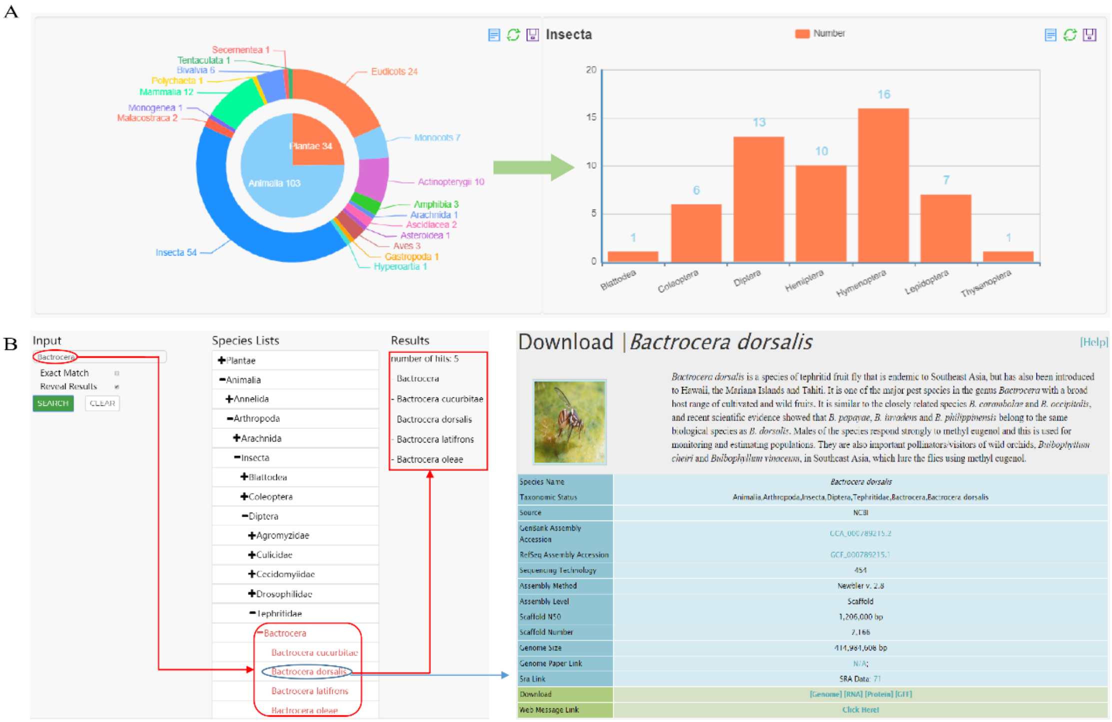
Species Name	GenBank ID	RefSeq ID	Data Source
Corbicula fluminea	GCA_001632725.1	n/a	NCBI
Crassostrea gigas	GCA_000297895.1	GCF_000297895.1	NCBI
Crassostrea virginica	GCA_002022765.4	GCF_002022765.2	NCBI
Culex quinquefasciatus	GCA_000209185.1	GCF_000209185.1	InsectBase
Cynara cardunculus	GCA_001531365.1	GCF_001531365.1	NCBI
Cyprinus carpio	GCA_000951615.2	GCF_000951615.1	NCBI
Dactylis glomerata	GCA_002892645.1	n/a	NCBI
Diaphorina citri	GCA_000475195.1	GCF_000475195.1	InsectBase
Dioscorea alata	GCA_002904275.2	n/a	NCBI
Diuraphis noxia	GCA_001186385.1	GCF_001186385.1	AphidBase
Dreissena polymorpha	GCA_000806325.1	n/a	NCBI
Drosera capensis	GCA_001925005.1	n/a	NCBI
Drosophila suzukii	GCA_000472105.1	GCF_000472105.1	SpottedWingFly Base
Echinochloa crusgalli	GCA_900205405.1	n/a	NCBI
Echium plantagineum	GCA_003412495.1	n/a	NCBI
Eichhornia paniculata	GCA_001647135.1	n/a	NCBI
Elaeis guineensis	GCA_000442705.1	GCF_000442705.1	NCBI
Equus asinus	GCA_001305755.1	GCF_001305755.1	NCBI
Eriocheir sinensis	GCA_003336515.1	n/a	NCBI
Eucalyptus camaldulensis	GCA_000260855.1	n/a	NCBI
Euphorbia esula	GCA_002919075.1	n/a	NCBI
Felis catus	GCA_000181335.4	GCF_000181335.3	NCBI
Ficus carica	GCA_002002945.1	n/a	NCBI
Frankliniella occidentalis	GCA_000697945.2	n/a	InsectBase
Gambusia affinis	GCA_003097735.1	n/a	NCBI
Gyrodactylus salaris	GCA_000715275.1	n/a	NCBI
Halyomorpha halys	GCA_000696795.1	GCF_000696795.1	InsectBase
Harmonia axyridis	GCA_003402655.1	n/a	NCBI
Helicoverpa armigera	GCA_002156985.1	GCF_002156985.1	i5K
Hydroides elegans	GCA_001703475.1	n/a	NCBI
Hyphantria cunea	GCA_003709505.1	n/a	NCBI
Ictalurus punctatus	GCA_001660625.1	GCF_001660625.1	NCBI
Leptinotarsa decemlineata	GCA_000500325.2	GCF_000500325.1	InsectBase
Limnoperna fortunei	GCA_003130415.1	n/a	NCBI
Linepithema humile	GCA_000217595.1	GCF_000217595.1	InsectBase
Liriomyza trifolii	GCA_001014935.1	n/a	i5K
Maconellicoccus hirsutus	GCA_003261595.1	n/a	i5K
Manduca sexta	GCA_000262585.1	n/a	InsectBase
Mayetiola destructor			InsectBase

Continued Appendix A

Species Name	GenBank ID	RefSeq ID	Data Source
Megachile rotundata	GCA_000220905.1	GCF_000220905.1	InsectBase
Microplitis demolitor	GCA_000572035.2	GCF_000572035.2	InsectBase
Mnemiopsis leidyi	GCA_000226015.1	n/a	NCBI
Monomorium pharaonis	GCA_000980195.3	GCF_000980195.1	InsectBase
Mus musculus	GCA_000001635.8	GCF_000001635.26	NCBI
Mytilus galloprovincialis	GCA_001676915.1	n/a	NCBI
Myzus persicae	GCA_001856785.1	GCF_001856785.1	AphidBase
Neovison vison	GCA_900108605.1	n/a	NCBI
Nicotiana glauca	GCA_002930595.1	n/a	NCBI
Oncorhynchus mykiss	GCA_002163495.1	GCF_002163495.1	NCBI
Onthophagus taurus	GCA_000648695.2	GCF_000648695.1	InsectBase
Ooceraea biroi	GCA_000611835.1	GCF_000611835.1	i5K
Oreochromis niloticus	GCA_001858045.2	GCF_001858045.1	NCBI
Oryctolagus cuniculus	GCA_000003625.1	GCF_000003625.3	NCBI
Osmerus eperlanus	GCA_900302275.1	n/a	NCBI
Paracoccus marginatus	GCA_900065295.1	n/a	i5K
Passer domesticus	GCA_001700915.1	n/a	NCBI
Passiflora edulis	GCA_002156105.1	n/a	NCBI
Penaeus monodon	GCA_002291185.1	n/a	i5K
Petromyzon marinus	GCA_002833325.1	n/a	NCBI
Pieris rapae	GCA_001856805.1	GCF_001856805.1	i5K
Plutella xylostella	GCA_000330985.1	GCF_000330985.1	DMB-DB
Poecilia reticulata	GCA_000633615.2	GCF_000633615.1	NCBI
Polistes dominula	GCA_001465965.1	GCF_001465965.1	i5K
Pomacea canaliculata	GCA_003073045.1	GCF_003073045.1	NCBI
Pseudomyrmex gracilis	GCA_002006095.1	GCF_002006095.1	i5K
Psidium guajava	GCA_002914565.1	n/a	NCBI
Psittacula krameri	GCA_002870145.1	n/a	NCBI
Pygocentrus nattereri	GCA_001682695.1	GCF_001682695.1	NCBI
Rana catesbeiana	GCA_002284835.2	n/a	NCBI
Raphanus raphanistrum	GCA_000769845.1	n/a	NCBI
Rattus norvegicus	GCA_000001895.4	GCF_000001895.5	NCBI
Rhinella marina	GCA_900303285.1	n/a	NCBI
Ricinus communis	GCA_000151685.2	GCF_000151685.1	PlantGDB
Rosa multiflora	GCA_002564525.1	n/a	NCBI
Saccharum spontaneum	GCA_003544955.1	n/a	NCBI
Salmo salar	GCA_000233375.4	GCF_000233375.1	SalmonDB
Silene latifolia	GCA_001412135.1	n/a	NCBI
Silybum marianum	GCA_001541825.1	n/a	NCBI
Sisymbrium irio	GCA_000411075.1	n/a	NCBI
Solenopsis invicta	GCA_000188075.1	GCF_000188075.1	InsectBase

Continued Appendix A

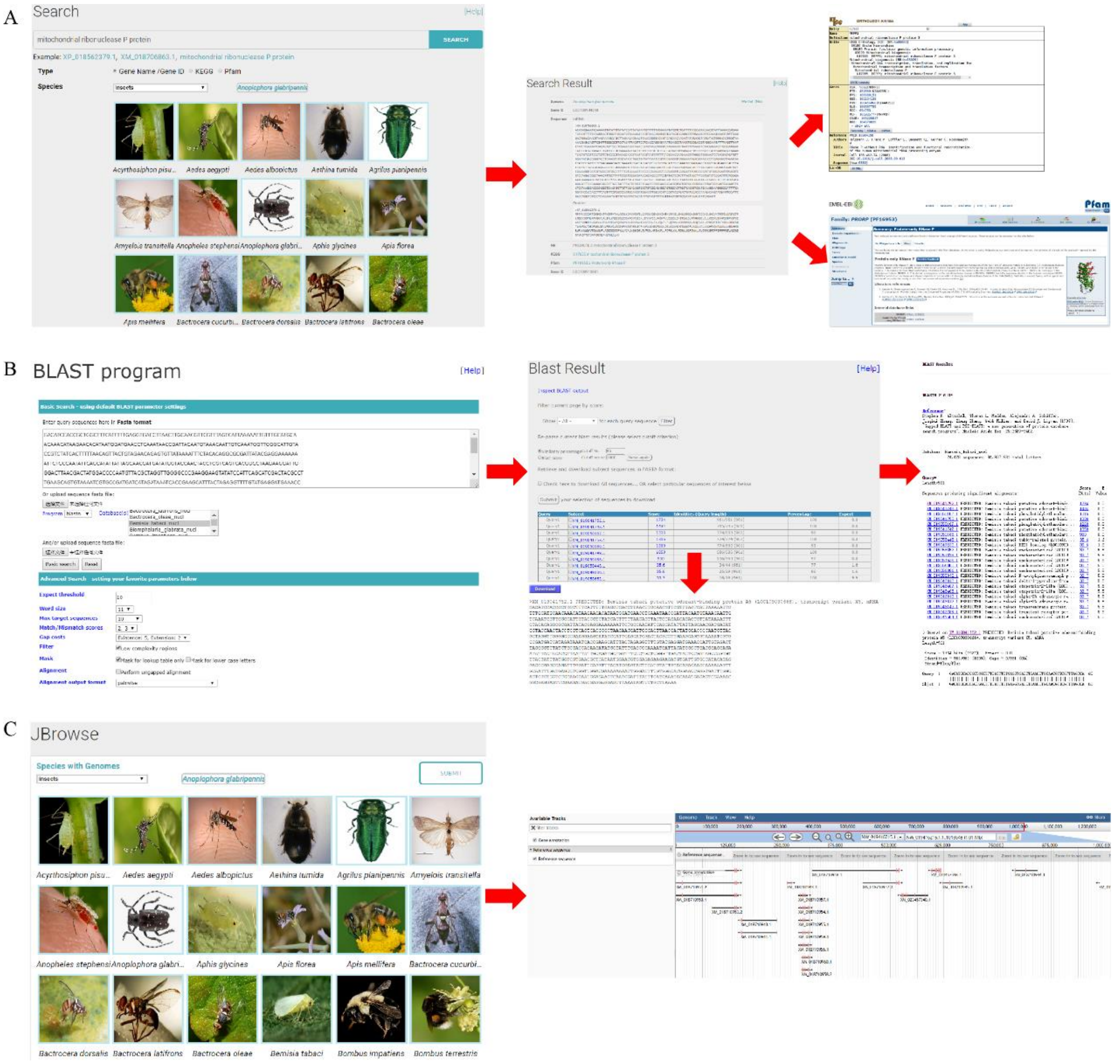
Species Name	**GenBank ID**	**RefSeq ID**	**Data Source**
Spodoptera frugiperda	GCA_002213285.1	n/a	InsectBase
Spodoptera litura	GCA_002706865.1	GCF_002706865.1	i5K
Sturnus vulgaris	GCA_001447265.1	GCF_001447265.1	NCBI
Utricularia gibba	GCA_002189035.1	n/a	NCBI
Varroa destructor	GCA_002443255.1	GCF_002443255.1	i5K
Vollenhovia emeryi	GCA_000949405.1	GCF_000949405.1	InsectBase
Vulpes vulpes	GCA_003160815.1	GCF_003160815.1	NCBI
Wasmannia auropunctata	GCA_000956235.1	GCF_000956235.1	InsectBase
Xenopus laevis	GCA_001663975.1	GCF_001663975.1	NCBI
Xiphophorus maculatus	GCA_002775205.2	GCF_002775205.1	NCBI



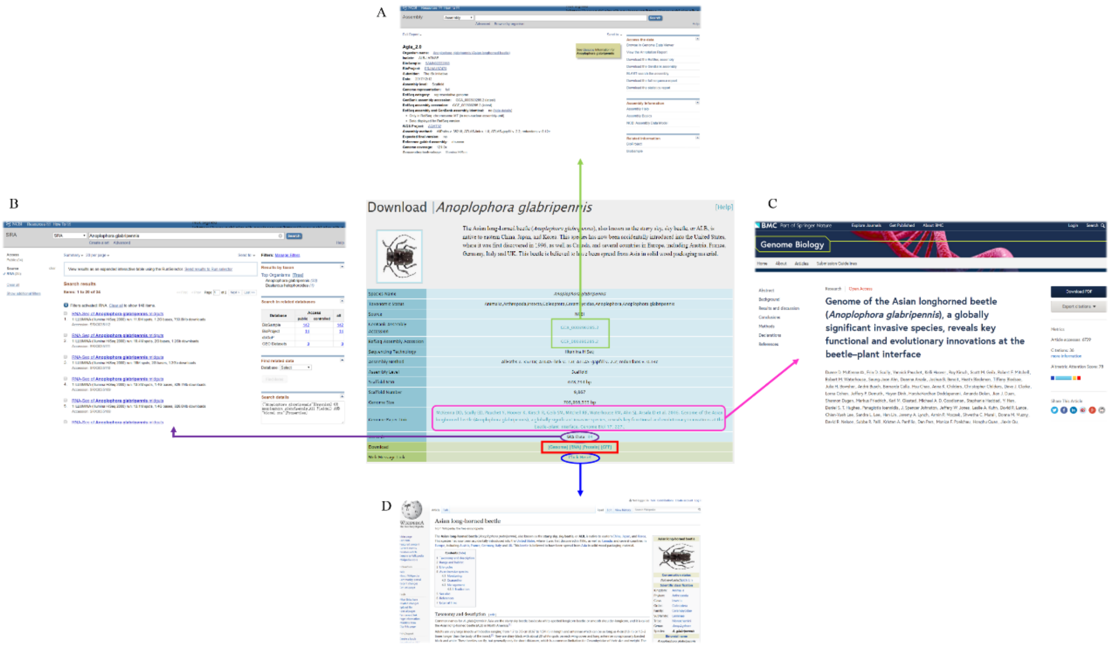
Bactrocera dorsalis is a species of tephritid fruit fly that is endemic to Southeast Asia, but has also been introduced to Hawaii, the Mariana Islands and Tahiti. It is one of the major pest species in the genus *Bactrocera* with a broad host range of cultivated and wild fruits. It is similar to the closely related species *B. carambolae* and *B. occipitalis*, and recent scientific evidence showed that *B. papayae*, *B. invadens* and *B. philippinensis* belong to the same biological species as *B. dorsalis*. Males of the species respond strongly to methyl eugenol and this is used for monitoring and estimating populations. They are also important pollinators/visitors of wild orchids, *Bulbophyllum cheiri* and *Bulbophyllum vinaceum*, in Southeast Asia, which lure the flies using methyl eugenol.

Species Name	*Bactrocera dorsalis*
Taxonomic Status	Animalia,Arthropoda,Insecta,Diptera,Tephritidae,Bactrocera,Bactrocera dorsalis
Source	NCBI
GenBank Assembly Accession	GCA_000789215.2
RefSeq Assembly Accession	GCF_000789215.1
Sequencing Technology	454
Assembly Method	Newbler v. 2.8
Assembly Level	Scaffold
Scaffold N50	1,206,000 bp
Scaffold Number	7,166
Genome Size	414,984,608 bp
Genome Paper Link	N/A;
Sra Link	SRA Data: 71
Download	[Genome] [RNA] [Protein] [GFF]
Web Message Link	Click Here!

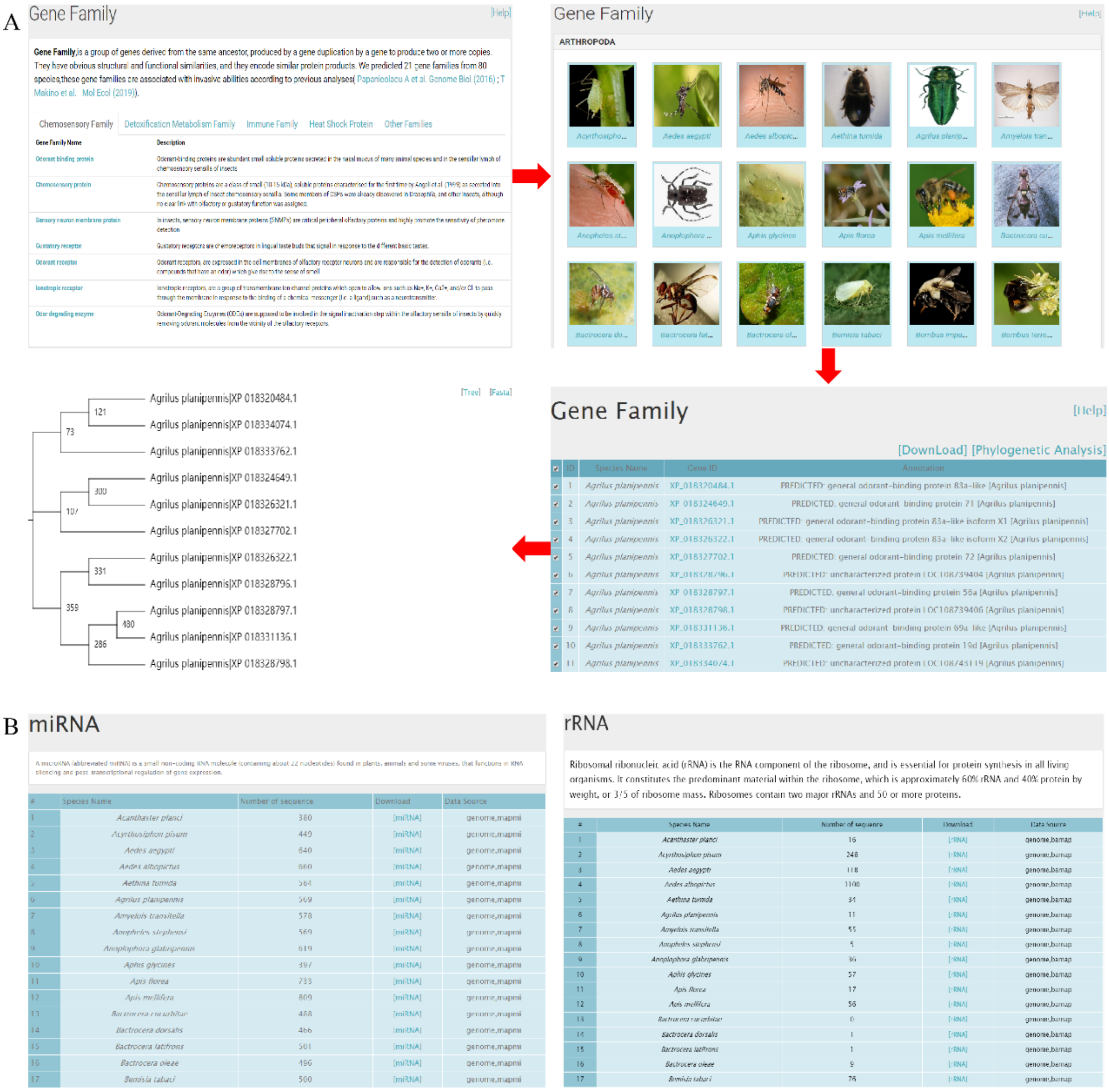
Appendix B. Species information interface. (A) The statistics of invasive species by the pie chart of different classes and histogram of different orders; (B) Quick search species by using the genus or species name.



Appendix C. Functional sections provided in InvasionDB. (A) A keyword search and results; (B) BLAST interface and a example of tBLASTn results; (C) Genome sequence view in JBrowse.



Appendix D. Download interface. (A) The genome assembly and annotation information; (B) The SRA data information; (C) The genome paper information; (D) The detail information of invasive species in wikipedia.



Gene Family

Gene Family,is a group of genes derived from the same ancestor, produced by a gene duplication by a gene to produce two or more copies. They have obvious structural and functional similarities, and they encode similar protein products. We predicted 21 gene families from 80 species,these gene families are associated with invasive abilities according to previous analyses(Papanicolacu A et al. Genome Biol.(2016) ; T Makino et al. Mol Ecol (2019)).

Gene Family Name	Description
Odorant binding protein	Odorant-binding proteins are abundant small soluble proteins secreted in the nasal mucus of many animal species and in the sensillar lymph of chemosensory sensilla of insects
Chemosensory protein	Chemosensory proteins are a class of small (10-15 kDa), soluble proteins characterised for the first time by Angeli et al. (1999) as secreted into the sensillar lymph of insect chemosensory sensilla. Some members of CSPs were already discovered in Drosophila, and other insects, although no clear link with olfactory or gustatory function was assigned.
Sensory neuron membrane protein	In insects, sensory neuron membrane proteins (SNMPs) are critical peripheral olfactory proteins and highly promote the sensitivity of pheromone detection
Gustatory receptor	Gustatory receptors are chemoreceptors in lingual taste buds that signal in response to the different basic tastes.
Odorant receptor	Odorant receptors, are expressed in the cell membranes of olfactory receptor neurons and are responsible for the detection of odorants (i.e., compounds that have an odor) which give rise to the sense of smell
Ionotropic receptor	Ionotropic receptors, are a group of transmembrane ion channel proteins which open to allow ions such as Na+, K+, Ca2+, and/or Cl- to pass through the membrane in response to the binding of a chemical messenger (i.e. a ligand),such as a neurotransmitter.
Odor degrading enzyme	Odorant-Degrading Enzymes (ODEs) are supposed to be involved in the signal inactivation step within the olfactory sensilla of insects by quickly removing odorant molecules from the vicinity of the olfactory receptors.

Gene Family

☑	ID	Species Name	Gene ID	Annotation
☑	1	*Agrilus planipennis*	XP_018320484.1	PREDICTED: general odorant-binding protein 83a-like [Agrilus planipennis]
☑	2	*Agrilus planipennis*	XP_018324649.1	PREDICTED: general odorant-binding protein 71 [Agrilus planipennis]
☑	3	*Agrilus planipennis*	XP_018326321.1	PREDICTED: general odorant-binding protein 83a-like isoform X1 [Agrilus planipennis]
☑	4	*Agrilus planipennis*	XP_018326322.1	PREDICTED: general odorant-binding protein 83a-like isoform X2 [Agrilus planipennis]
☑	5	*Agrilus planipennis*	XP_018327702.1	PREDICTED: general odorant-binding protein 72 [Agrilus planipennis]
☑	6	*Agrilus planipennis*	XP_018328796.1	PREDICTED: uncharacterized protein LOC108739404 [Agrilus planipennis]
☑	7	*Agrilus planipennis*	XP_018328797.1	PREDICTED: general odorant-binding protein 56a [Agrilus planipennis]
☑	8	*Agrilus planipennis*	XP_018328798.1	PREDICTED: uncharacterized protein LOC108739406 [Agrilus planipennis]
☑	9	*Agrilus planipennis*	XP_018331136.1	PREDICTED: general odorant-binding protein 69a-like [Agrilus planipennis]
☑	10	*Agrilus planipennis*	XP_018333762.1	PREDICTED: general odorant-binding protein 19d [Agrilus planipennis]
☑	11	*Agrilus planipennis*	XP_018334074.1	PREDICTED: uncharacterized protein LOC108743119 [Agrilus planipennis]

miRNA

A microRNA (abbreviated miRNA) is a small non-coding RNA molecule (containing about 22 nucleotides) found in plants, animals and some viruses, that functions in RNA silencing and post-transcriptional regulation of gene expression.

#	Species Name	Number of sequence	Download	Data Source
1	*Acanthaster planci*	380	[miRNA]	genome,mapmi
2	*Acyrthosiphon pisum*	449	[miRNA]	genome,mapmi
3	*Aedes aegypti*	640	[miRNA]	genome,mapmi
4	*Aedes albopictus*	660	[miRNA]	genome,mapmi
5	*Aethina tumida*	584	[miRNA]	genome,mapmi
6	*Agrilus planipennis*	569	[miRNA]	genome,mapmi
7	*Amyelois transitella*	578	[miRNA]	genome,mapmi
8	*Anopheles stephensi*	569	[miRNA]	genome,mapmi
9	*Anoplophora glabripennis*	619	[miRNA]	genome,mapmi
10	*Aphis glycines*	397	[miRNA]	genome,mapmi
11	*Apis florea*	733	[miRNA]	genome,mapmi
12	*Apis mellifera*	809	[miRNA]	genome,mapmi
13	*Bactrocera cucurbitae*	488	[miRNA]	genome,mapmi
14	*Bactrocera dorsalis*	466	[miRNA]	genome,mapmi
15	*Bactrocera latifrons*	501	[miRNA]	genome,mapmi
16	*Bactrocera oleae*	496	[miRNA]	genome,mapmi
17	*Bemisia tabaci*	500	[miRNA]	genome,mapmi

rRNA

Ribosomal ribonucleic acid (rRNA) is the RNA component of the ribosome, and is essential for protein synthesis in all living organisms. It constitutes the predominant material within the ribosome, which is approximately 60% rRNA and 40% protein by weight, or 3/5 of ribosome mass. Ribosomes contain two major rRNAs and 50 or more proteins.

#	Species Name	Number of sequence	Download	Data Source
1	*Acanthaster planci*	16	[rRNA]	genome,barrnap
2	*Acyrthosiphon pisum*	248	[rRNA]	genome,barrnap
3	*Aedes aegypti*	118	[rRNA]	genome,barrnap
4	*Aedes albopictus*	1100	[rRNA]	genome,barrnap
5	*Aethina tumida*	34	[rRNA]	genome,barrnap
6	*Agrilus planipennis*	11	[rRNA]	genome,barrnap
7	*Amyelois transitella*	55	[rRNA]	genome,barrnap
8	*Anopheles stephensi*	5	[rRNA]	genome,barrnap
9	*Anoplophora glabripennis*	36	[rRNA]	genome,barrnap
10	*Aphis glycines*	57	[rRNA]	genome,barrnap
11	*Apis florea*	17	[rRNA]	genome,barrnap
12	*Apis mellifera*	56	[rRNA]	genome,barrnap
13	*Bactrocera cucurbitae*	0	[rRNA]	genome,barrnap
14	*Bactrocera dorsalis*	1	[rRNA]	genome,barrnap
15	*Bactrocera latifrons*	1	[rRNA]	genome,barrnap
16	*Bactrocera oleae*	9	[rRNA]	genome,barrnap
17	*Bemisia tabaci*	76	[rRNA]	genome,barrnap

Appendix E. Genome features associated with invasiveness. (A) Gene families related to invasiveness; (B) ncRNAs related to invasiveness.