IN SILICO DETECTION OF SPECIES- AND ORDER-SPECIFIC MARKERS BASED ON VIRTUAL PCR-RFLP OF THE *COI* GENE IN INSECT

SABIR, J. S. M.[1] – RABAH, S.[1] – YACOUB, H.[2] – HAJRAH, N. H.[1] – ATEF, A.[1] – HALL, N.[1,3] – BAHIELDIN, A.[1,4*]

[1]*Department of Biological Sciences, Faculty of Science, King Abdulaziz University (KAU), P.O. Box 80141, Jeddah 21589, Saudi Arabia*

[2]*Department of Biological Sciences, Faculty of Science, University of Jeddah, Dahaban 23881, Saudi Arabia*

[3]*The Genome Analysis Center, Norwich Research Park, Norwich, NR4 7UH, UK*

[4]*Department of Genetics, Faculty of Agriculture, Ain Shams University, Cairo, Egypt*

**Corresponding author*
e-mail: abmahmed@kau.edu.sa; phone: +966-50-632-9922

(Received 18th Apr 2019; accepted 1st Jul 2019)

Abstract. COI is the first subunit of the mitochondrial cytochrome C oxidase gene that is commonly used for barcoding via sequencing and molecular identification of insect orders and species. However, fast and cheap data analysis might still be required for the detection of insect orders and species to replace sequencing in large populations. The latter can be emphasized by detecting restriction enzyme site polymorphisms in COI sequences of insect orders and species. In the present study, we have utilized an in silico approach for detecting species- and order-specific molecular markers in insects collected from Saudi Arabia based on the new non-traditional open source tool, namely C*is*SERS. A number of 237 restriction enzymes were checked for the presence of their recognition sites within the *COI* gene sequence of the different insect species. Generated data was displayed in predicted agarose gels as virtual PCR-RFLP patterns. The results indicated the occurrence of 14 species-specific and three order-specific molecular markers representing 10 species and two orders of insect. The highest number of species-specific markers (3 markers) was generated for *Idolus picipennis* (Order: Coleoptera), followed by the two hemipteran *Macrosteles* sp. and *Osbornellus auronitens* (two markers each). Three order-specific markers were generated for orders Lepidoptera (two markers) and Diptera (one marker). The overall results demonstrated that C*is*SERS efficiently allowed for the detection of species- and order-specific markers that can further be utilized in screening large insect populations.
Keywords: *PCR-RFLP, CisSERS, genetic characterization*

Introduction

Subunit I (658 nt) of the mitochondrial cytochrome C oxidase (COX) gene (namely *COI*) started in 2003 to be the most universal marker for species identification in animal (Hebert et al., 2003, 2004). Recent concerns were raised for the insufficiency of insect biodiversity data in Saharo-Arabian region including Saudi Arabia and the lack of conducted traditional surveys (Ashfaq et al., 2018). The latter makes it vital to gain insights on such a virgin area in terms of insect classification and biodiversity.

Our recent study (Ashfaq et al., 2018) proved the feasibility and efficiency of utilizing *COI* gene in barcoding insects collected from Saudi Arabia. In the present study, we further utilized virtual PCR-RFLP based on the barcode sequences targeting the recovery of fast and cheap approach for characterizing insect orders and species in

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1705_1140911418

large populations. Methods based on restriction enzyme digestion (ex., PCR-RFLP) are easy to perform, reproducible, and cost-effective and require less equipment and expertise (Agarwal et al., 2008). In the past, restriction enzyme-based molecular markers were commonly employed, but with the emergence of high-throughput sequencing technologies, they facilitated more in-depth genomic analysis and characterization. PCR-based restriction fragment length polymorphic markers (or PCR-RFLP) are used for gene mapping and molecular breeding (Shu et al., 2011). To detect the efficacy of re-utilizing PCR-RFLP for molecular characterization, we were required to use efficient in silico tools suitable for analyzing large genomic datasets towards the recovery of a fast and cost-effective approach in characterizing large insect populations in the future.

Access to computational tools that processes massive amounts of DNA sequence information from multiple samples is useful for addressing difficult biological questions. Several current restriction site analysis tools were designed to handle only few sequences for targeted analysis (Neff et al., 2002; Vincze et al., 2003; Ilic et al., 2004; Thiel et al., 2004; Zhang et al., 2005; Taylor and Provart, 2006; Chang et al., 2010). *Cis*SERS (or customizable in silico sequence evaluation for restriction sites) tools were designed specifically for PCR-RFLP marker detection (Sharpe et al., 2016). Such a new graphical user interface helps in detecting enzyme restriction sites in large DNA sequence datasets and enables high-throughput analysis of multiple sequences for restriction sites with a possibly embedded dynamic visualization functionality. It can predict places the enzyme cuts during gel electrophoresis and save time and effort of running unrequired several enzymatically-digested DNA samples. Such information can help in deciding which enzymes to be selected in digesting DNA for restriction-site associated DNA (or RAD) sequencing libraries and further cloning. C*is*SERS can also be utilized in detecting methylated DNAs as the original PCR-RFLP will vary from the virtual pattern for any methylation susceptible restriction enzyme site. C*is*SERS is expected to pave the way between sequence acquisition and implementation of diverse wet-lab approaches to solve biological questions more easily.

In this work, we utilized the COI subunit for specimens of six insect orders collected from Saudi Arabia to detect species- and order-specific molecular markers based on the use of C*is*SERS surface tools in order to generate virtual PCR-RFLP. This information will be useful in the future analysis of large insect populations via PCR-RFLP.

Materials and methods

For sample collection, a Malaise trap (Hutcheson and Jones, 1999; Hill et al., 2005) was installed at Hada Al-Sham station, King Abdulaziz University (KAU) located in the western region (near Makkah) of Saudi Arabia (21.795°N, 39.711°E). Sample collection, morphological classification up to Sanger sequencing were done as indicated in our most recent study (Ashfaq et al., 2018).

C*is*SERS was used to develop fasta files for restriction site analysis and generate predicted gel images as described (Sharpe et al., 2016). Outputs describe the cut counts and locations for each restriction site and dynamically created predicted gel images. The master list of enzymes (237 in the present study) is retrieved from the REbase database (Vincze et al., 2003). A set of formulas used via C*is*SERS is retrieved in order to produce fragment size (Neff et al., 2002; Ilic et al., 2004; Agarwal et al., 2008; Shu et

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1705_1140911418

al., 2011). The motif detection feature of C*is*SERS was previously validated using *Nostoc* and *ATPC1* genes (Sharpe et al., 2016).

For validating the in silico datasets, DNAs from the different 30 insect species were isolated and subjected to PCR to recover the *COI* gene fragment (634 bp). Restriction analysis including five selected restriction enzymes was done following standard protocols to validate five species-specific markers of four insect species. The products of restriction analysis were detected using electrophoresis on agarose gels (1.2% in 1x TBE buffer), stained with ethidium bromide (0.3 ug/ml), then visually examined with UV transilluminator and photographed using a CCD camera (UVP, UK).

Results and discussion

Generating a DNA barcode based on the common *COI* gene fragment involves sequencing, then BLAST towards the identification of sequence homology down to the species level. Conventional DNA sequencing is time-consuming and costly, which makes it difficult to be universally utilized. The latter approach generates no particular species- or order-specific markers that can be solely used in preliminary screening of any insect population. In PCR-RFLP, the restriction sites are identified in the target gene fragment (e.g., *COI*), and polymorphism is based on the number and sizes of the cleaved fragment of any given species. This method is simple, cheap and time-saving, thus meets many applications (Liu et al., 2001, 2003; Um et al., 2001; Ding et al., 2003; Wang et al., 2005, 2007; Ying et al., 2007).

In our previous study, *COI* barcodes were analyzed at the DNA level for a number of 560 insect samples collected during four-week time course in Saudi Arabia. Recovered sequences were trimmed to 634 nucleotides and number of samples was narrowed to 175 representing 30 species of six insect orders. Sequences of these species were utilized in generating the virtual PCR-RFLP patterns via C*is*SERS surface tools. A number of 237 restriction enzymes were checked for the presence of recognition sites within the *COI* gene sequence (634 bp) of the different insect species. The consensus sequences of the 30 insect species were fed to the program and restriction sites for enzymes that either uniquely cut DNA of one single species or resulted in the occurrence of a DNA fragment specific to a single insect order are shown in *Figure A1* in the *Appendix*. The figure indicates the positions of the 17 selected enzymes that cut within the COI sequence that meet the above mentioned criterion. Interestingly, no restriction sites were shown in the first 181 nt of the gene (*Fig. A1*; *Table 1*). All selected restriction sites are located downwards. The resulted species- and order-specific markers are shown in the virtual PCR-RFLP patterns shown in *Figures 1* and *2*, respectively, while summarized in *Figure 3*. Species-specific markers were recovered for 10 out of the 30 insect species. The largest number of species-specific markers are shown for *Idolus picipennis* (3 markers) of order Coleoptera, followed by *Macrosteles sp.* and *Osbornellus auronitens* (2 markers each) of order Hemiptera (*Fig. 1*). The single species-specific markers were detected for seven other species, e.g., *Monomorium junodi*, *Camponotus maculatus*, *Balta vilis*, *Belenois aurota*, *Ephysteris subdiminutella*, *Carpomya vesuviana* and *Tabanus striatus*. Numbers of 2, 4, 1, 2, 3 and 1 species-specific markers were shown for orders Hymenoptera, Hemiptera, Blattodae, Lepidoptera, Coleoptera and Diptera, respectively (*Figs. 1* and *A1*). The results at the order level indicate the recovery of three order-specific markers.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1705_1140911418

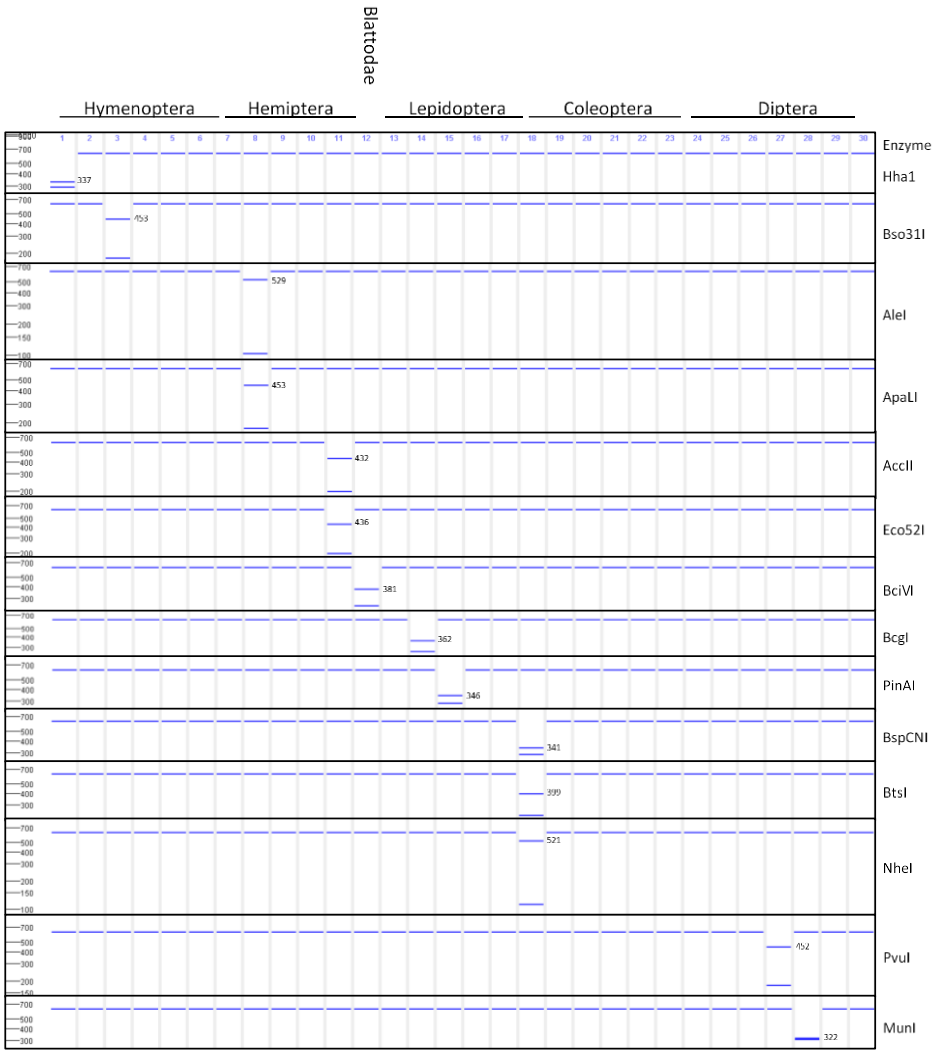


Figure 1. *Virsual electrophregrams generated from CisSERS tools to describe specific PCR-RFLP markers in COI gene in 10 insect species belonging to six insect orders. Based on ach restriction enzyme has a unique site in the gene for a given insect species. Species numbers (1-30) refer to insect species shown in Table 1*

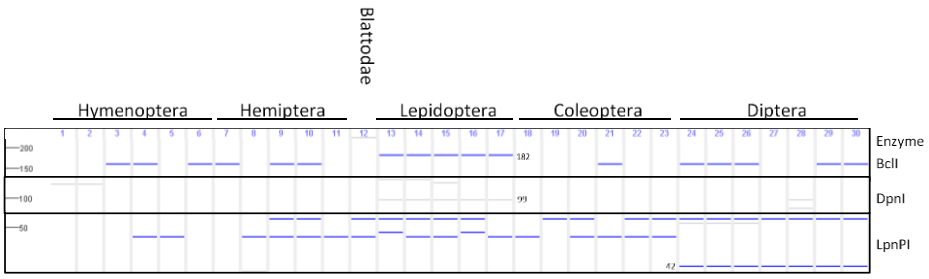


Figure 2. *Virsual electrophregrams generated from CisSERS tools to describe specific PCR-RFLP markers in COI gene in two out of six insect orders. Each restriction enzyme has multiple recogintion sites in the gene for the different insect species and orders. One fragment for each of the enzymes BclI (182 bp), DpnI (99 bp) are specific for order Lepidoptera, while one fragmente of the enzyme LpnPI (42) is specific for order Diptera. Species numbers (1-30) refer to insect species shown in Table 1*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1705_1140911418

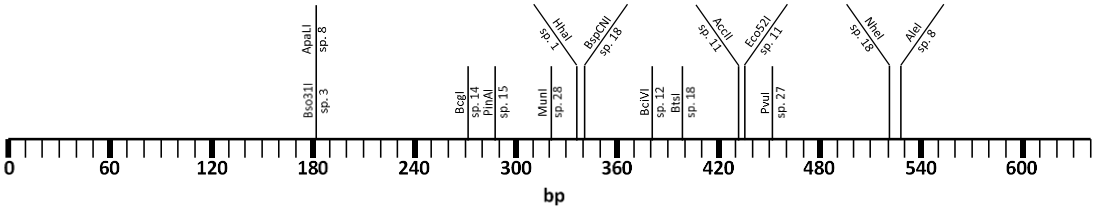


Figure 3. *Restriction map for the enzymes uniquely cut the COI gene in a given insect species. The map involves 14 enzymes specific for 10 species of the six insect orders. Species numbers refer to insect species shown in Table 1*

Table 1. *List of insect species and orders along with restriction enzymes with recognition sites within COI gene*

No.	Species	Enzyme	Species-specific marker (bp) 5`…….3`	Order	Enzyme	Order-specific marker (bp)
1	*Monomorium junodi*	HhaI	337//297	Hymenoptera	-	-
2	*Cataglyphis ibericus*	-	-			
3	*Camponotus maculatus*	Bso31I	181//453			
4	*Hymenoptera sp.*	-	-			
5	*Nomioides facilis*	-	-			
6	*Tachytes crassus*	-	-			
7	*Batracomorphus angustatus*	-	-	Hemiptera	-	-
8	*Macrosteles sp.*	AleI, ApaLI	529//105, 181//453			
9	*Colladonus tahotus*	-	-			
10	*Sacphytopius vaccinium*	-	-			
11	*Osbornellus auronitens*	AccII, Eco52I	202//432, 436//198			
12	*Balta vilis*	BciVI	381//253	Blattodae	-	-
13	*Tarucus theophrastus*	-	-	Lepidoptera	BclI DpnI	182 99
14	*Belenois aurota*	BcgI	242//362			
15	*Ephysteris subdiminutella*	PinAI	288//346			
16	*Tuta absoluta*	-	-			
17	*Batrachedra amydraula*	-	-			
18	*Idolus picipennis*	BspCNI, BtsI, NheI	341//293, 399//235, 521//113	Coleoptera	-	-
19	*Melanotus villosus*	-	-			
20	*Algarobius prosopis*	-	-			
21	*Exomala pallidipennis*	-	-			
22	*Xylotrupes siamensis*	-	-			
23	*Copris tripartitus*	-	-			
24	*Lispe albitarsis*	-	-	Diptera	LpnPI	42
25	*Chrysomya chani*	-	-			
26	*Anthomyiopsis nigrisquamata*	-	-			
27	*Carpomya vesuviana*	PvuI	452//182			
28	*Tabanus striatus*	MunI	322//312			
29	*Leptogaster cylindrica*	-	-			
30	*Asyndetus sp.*	-	-			

Two of which are shown for order Lepidoptera and one is shown for order Diptera (*Fig. 2*). Exact position of each marker is shown in *Figure A1*. Sizes of these markers are 182 and 99 nt for order Lepidoptera and 42 for order Diptera. Enzymes that generated these markers are BclI, DpnI and LpnPI, respectively. Interestingly, DpnI

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1705_1140911418

enzyme has a 4-nt recognition site (GATC) that exists within the 6-nt recognition site (TGATCA) of BclI enzyme. The TGATCA site is located at 453 nt, therefore, the two enzymes cut DNA at this site. DpnI has another site at 552 nt, which is located 99 nt downstream the first, while BclI has no further sites downstream the first (*Fig. A1*). Accordingly, the cut with BclI resulted in the occurrence of 182 nt marker, while the cut with DpnI enzyme resulted in the occurrence of 99 nt marker. The results generated from *Cis*SERS tools were validated via the restriction analysis of PCR products involving five randomly selected species-specific markers and electrophoregram results supported those generated by *Cis*SERS tools (*Fig. A2*). Concerns on the reproducibility of PCR were repeatedly raised. Besides, dry-lab restriction analysis requires validation via wet-lab experimentation. Therefore, we consider the data shown in *Fig. A2* can validate both the PCR approach as such as well as the utility of *Cis*SERS tools. Other universal markers, ex., 16S rRNA gene, were successfully utilized in detecting PCR-RFLP patterns of the causative bacteria of neonatal sepsis, e.g., *Staphylococcus aureus* and *Klebsiella* spp. (Rohit et al., 2016). The results indicated that PCR-RFLP is useful for the early diagnosis of culture negative neonatal sepsis in human.

There are many restriction site analysis tools designed to handle only few/short sequences (Neff et al., 2002; Vincze et al., 2003; Ilic et al., 2004; Thiel et al., 2004; Zhang et al., 2005; Taylor and Provart, 2006; Chang et al., 2010). Of which, NEBcutter handles a less than 300 kb, besides, it is not set up for multiple sequence analysis (Vincze et al., 2003). The other tools are web-server based with limited functionality for high-throughput analysis. However, *Cis*SERS was developed to enable high-throughput analysis of multiple/long sequences for restriction sites. In addition, *Cis*SERS solely allows for custom motif detection to identify conserved sites, e.g., *cis*- and *trans*-acting element binding sites. Sharpe et al. (2016) indicated the effectiveness and multiple distinctive functionalities of *Cis*SERS in analyzing mature sequence data of *cfq* gene. They also indicated that utilizing *Cis*SERS in transcriptomic or genomic analysis enables a guidance during subsequent restriction enzyme-based experimentation. PCR-based approaches could make identification of insect at species level easier, faster and more accurate compared with traditional morphological approaches. We think this work in the first of its kind in detecting species- or order-specific molecular markers in insects as well as in the utility of *Cis*SERS in discriminating insect orders and species. In conclusion, *Cis*SERS is a fast, cost-effective surface tools enable bridging the gap between sequence acquisition and implementation of wet-lab approaches towards addressing important biological clues.

Conclusion

Data generated in the present study from virtual PCR-RFLP via *Cis*SERS provided useful information. Firstly, it narrows the number of restriction enzymes to be used in the future in studying any insect population. Second, it supports the utility of the wet-lab PCR-RFLP as a fast, cheap, requires less equipment and cost-effective approach for characterizing insect populations. Third, it provides species- and order-specific markers that can be used as a first layer of detection at the species or order levels. Such an approach can be adopted for other types of universal markers in prokaryotic as well as eukaryotic organisms.

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1705_1140911418

Acknowledgments. This project was funded by the Deanship of Scientific Research (DSR), King Abdulaziz University, Jeddah, under Grant no. (99-130-35-HiCi). The authors, therefore, acknowledge with thanks DSR technical and financial support.

Author contributions. All authors have read and approved this manuscript. Conceived and designed the experiments: JS, SR, NH, AB. Performed the experiments: HY, NH, AA. Analyzed the data: HY, NH, AA, KA. Wrote the paper: JS, KA, NH, AB.

Conflict of interests. The authors declare that they have no conflict of interests.

Data availability. Authors declare no unavailable data generated through or related to the course of this research.

REFERENCES

[1] Agarwal, M., Shrivastava, N., Padh, H. (2008): Advances in molecular marker techniques and their applications in plant sciences. – Plant Cell Rep. 27: 617-631.
[2] Ashfaq, M., Sabir, J. S. M., El-Ansary, H. O., et al. (2018): Insect diversity in the Saharo-Arabian region: Revealing a little-studied fauna by DNA barcoding. – PLoS ONE 13: e0199965.
[3] Chang, H. W., Cheng, Y. H., Chuang, L. Y., et al. (2010): SNP-RFLPing 2: an updated and integrated PCR-RFLP tool for SNP genotyping. – Bmc Bioinformatics 11: Artn 173. DOI: 10.1186/1471-2105-11-173.
[4] Ding, X. Y., Wang, Z. T., Zhou, K. Y., et al. (2003): Allele-specific primers for diagnostic PCR authentication of *Dendrobium officinale*. – Planta Medica 69: 587-588.
[5] Hebert, P. D. N., Cywinska, A., Ball, S. L., et al. (2003): Biological identifications through DNA barcodes. – Proc. Biol. Sci. 270: 313-321.
[6] Hebert, P. D. N., Penton, E. H., Burns, J. M., et al. (2004): Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. – Proc. Natl. Acad. Sci. USA 101: 14812-14817.
[7] Hill, D., Fasham, M., Tucker, G., et al. (2005): Handbook of Biodiversity Methods: Survey, Evaluation and Monitoring. – Cambridge University Press, Cambridge.
[8] Hutcheson, J., Jones, D. (1999): Spatial variability of insect communities in a homogenous system: Measuring biodiversity using Malaise trapped beetles in a *Pinus radiata* plantation in New Zealand. – Forest Ecol. Manag. 118: 93-105.
[9] Ilic, K., Berleth, T., Provart, N. J. (2004): BlastDigester—a web-based program for efficient CAPS marker design. – Trends in Genet. 20: 280-283.
[10] Liu, X. H., Wang, Y. Q., Zhou, K. Y., et al. (2003): Study on allele-specific diagnostic PCR of the traditional Chinese medicines of the deers. – Acta Pharmaceutica Sinica 36: 634-635.
[11] Liu, Z., Wang, Y. Q., Zhou, K., et al. (2001): Authentication of Chinese crude drug, Gecko, by allele-specific diagnostic PCR. – Planta Medica 67: 385-387.
[12] Neff, M. M., Turk, E., Kalishman, M. (2002): Web-based primer design for single nucleotide polymorphism analysis. – Trends in Genet. 18: 613-615.
[13] Rohit, A., Maiti, B., Shenoy, S., et al. (2016): Polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) for rapid diagnosis of neonatal sepsis. – Indian J. Med. Res. 143: 72-78.
[14] Sharpe, R. M., Koepke, T., Harper, A., et al. (2016): CisSERS: Customizable *in silico* sequence Evaluation for restriction sites. – PLoS ONE 11: e0152404.
[15] Shu, Y. J., Li, Y., Zhu, Z. L., et al. (2011): SNPs discovery and CAPS marker conversion in soybean. – Mol. Biol. Rep. 38: 1841-1846.
[16] Taylor, J., Provart, N. J. (2006): CapsID: a web-based tool for developing parsimonious sets of CAPS molecular markers for genotyping. – BMC Genet. 7: Artn 27.

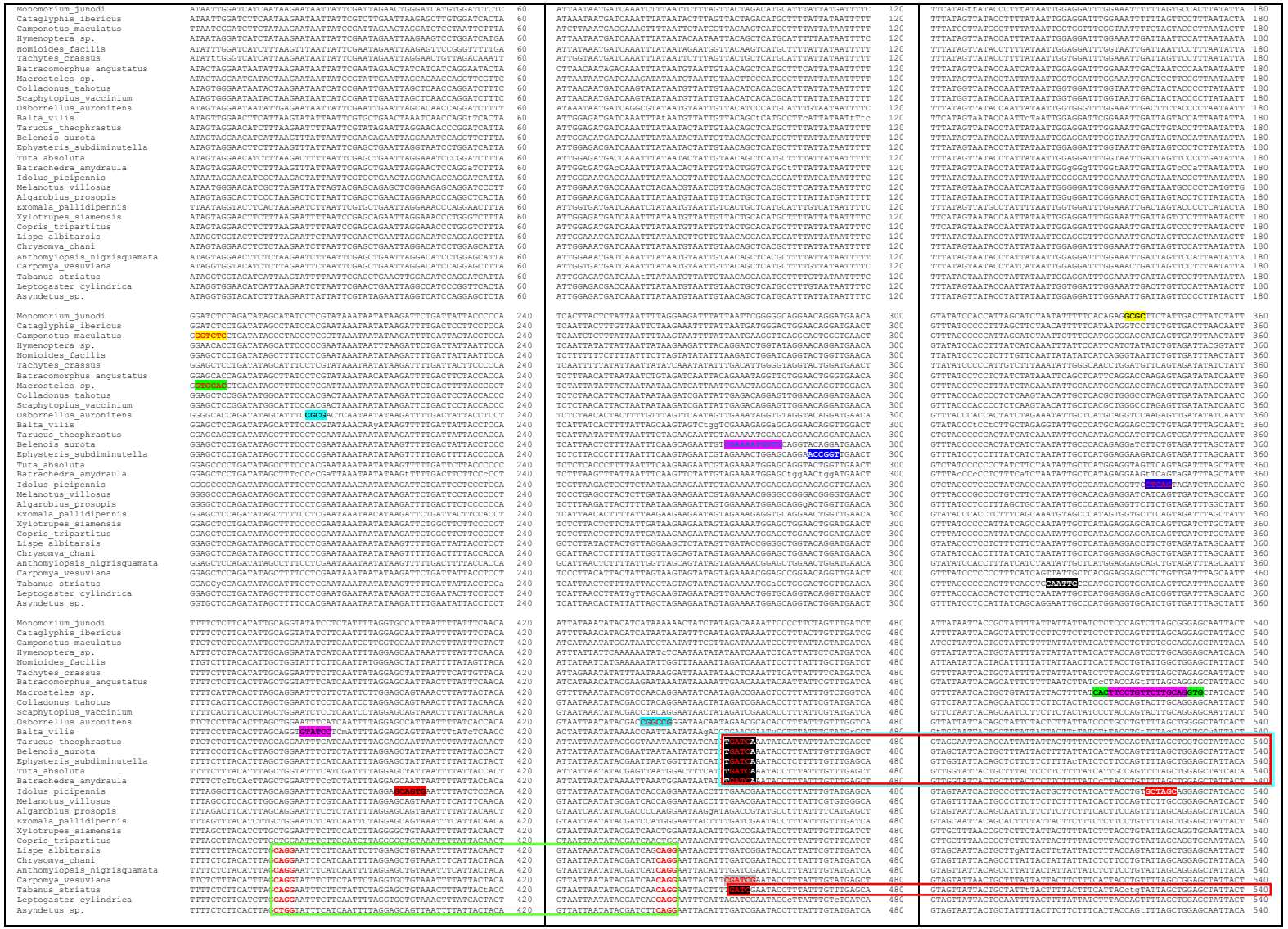
http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1705_1140911418

[17] Thiel, T., Kota, R., Grosse, I., et al. (2004): SNP2CAPS: a SNP and INDEL analysis tool for CAPS marker development. – Nucleic Acids Res. 32: ARTN e5.
[18] Um, J. Y., Chung, H. S., Kim, M. S., et al. (2011): Molecular authentication of Panax ginseng species by RAPD analysis and PCR-RFLP. – Biol. Pharma. Bull. 24: 872-875.
[19] Vincze, T., Posfai, J., Roberts, R. J. (2003): NEBcutter: a program to cleave DNA with restriction enzymes. – Nucleic Acids Res. 31: 3688-3691.
[20] Wang, C. Z., Li, P., Ding, J. Y., et al. (2005): Identification of *Fritillaria pallidiflora* using diagnostic PCR and PCR-RFLP based on nuclear ribosomal DNA internal transcribed spacer sequences. – Planta Medica 71: 384-386.
[21] Wang, C. Z., Li, P., Ding, J. Y., et al. (2007): Simultaneous identification of *Bulbus Fritillariae cirrhosae* using PCR-RFLP analysis. – Phytomedicine 14: 628-632.
[22] Ying, Y., Xu, H., Wang, Z. T. (2007): Allele-specific diagnostic PCR authentication of *Dendrobium thyrsiflorum*. – Acta Pharmaceutica Sinica 42: 98-103.
[23] Zhang, R. F., Zhu, Z. H., Zhu, H. M., et al. (2005): NP Cutter: a comprehensive tool for SNPPCR-RFLP assay design. – Nucl. Acids Res. 33: W489-W92.

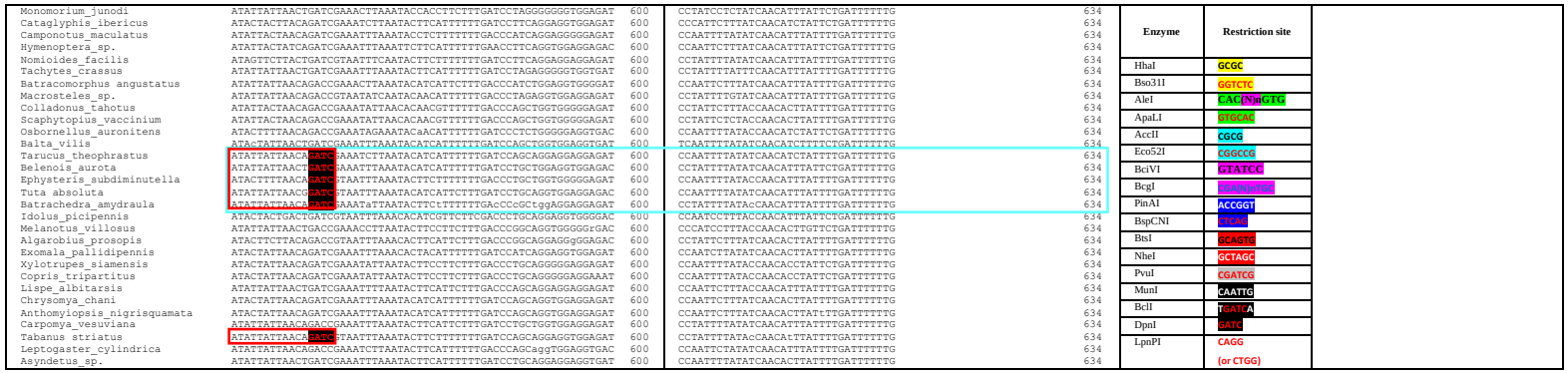
APPENDIX

Figure A1. *(See next page) Consensus GOI nucleotide sequences of the 30 insect species describing the recognition sites of the 17 restriction enzymes used in detecting PCR-RFLP markers at the species (14 markers) and order (3 markers) levels. Blue and red boxes indicate BclI (182 bp) DpnI (99 bp) fragments specific for Lepidopteran species, while green box indicates the LpnPI (42 bp) fragment specific for dipteran species*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1705_1140911418



APPLIED ECOLOGY AND ENVIRONMENTAL RESEARCH 17(5):11409-11418.
http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1705_1140911418



Monomorium_junodi	ATATTATTAACTGATCGAAACTTAAATACCACCTTCTTTGATCCTAGGGGGGGTGGAGAT	600	CCTATCCTCTATCAACATTTATTCTGATTTTTTG	634
Cataglyphis_ibericus	ATACTACTTAACAGATCGAAATCTTAATACTTCATTTTTTGATCCTTCAGGAGGTGGAGAT	600	CCCATTCTTTATCAACATCTATTCTGATTTTTTG	634
Camponotus_maculatus	ATATTACTAACAGATCGAAATTTAAATACCTCTTTTTTTGACCCATCAGGAGGGGGAGAT	600	CCAATTTTATATCAACATTTATTTTGATTTTTTG	634
Hymenoptera_sp.	ATATTACTATCAGATCGAAATTTAAATTCTTCATTTTTTGAACCTTCAGGTGGAGGAGAC	600	CCAATTCTTTATCAACATTTATTCTGATTTTTTG	634
Nomioides_facilis	ATAGTTCTTACTGATCGTAAATTCAATACTTCTTTTTTTGATCCTTCAGGAGGAGGAGAT	600	CCTATTTTATATCAACATTTATTTTGATTTTTTG	634
Tachytes_crassus	ATATTATTAACTGATCGAAATTTAAATACTTCATTTTTTGATCCTAGAGGGGGTGGTGAT	600	CCTATTTTATTTCAACATTTATTTTGATTTTTTG	634
Batracomorphus_angustatus	ATATTATTAACAGACCGAAACTTAAATACATCATTCTTTGACCCATCTGGAGGTGGGGAT	600	CCAATTCTTTATCAACATTTATTTTGATTTTTTG	634
Macrosteles_sp.	ATATTATTAACAGACCGTAAATATCAATACAACATTTTTTGACCCTAGAGGTGGAGGAGAT	600	CCTATTTTGTATCAACATTTATTTTGATTTTTTG	634
Colladonus_tahotus	ATATTACTAACAGACCGAAATATTAACACAACGTTTTTTGACCCAGCTGGTGGGGGAGAT	600	CCTATTCTTTACCAACACTTATTTTGATTTTTTG	634
Scaphytopius_vaccinium	ATATTACTAACAGACCGAAATATTAACACAACGTTTTTTGACCCAGCTGGTGGGGGAGAT	600	CCTATTCTCTACCAACACTTATTTTGATTTTTTG	634
Osbornellus_auronitens	ATACTTTTAACAGACCGAAATAGAAATACaACATTTTTTGATCCCTCTGGGGGAGGTGAC	600	CCAATTTTATACCAACATCTATTCTGATTTTTTG	634
Balta_vilis	ATACTATTAACTGATCGAAATTTAAATACATCATTTTTTGATCCAGCTGGTGGAGGTGAT	600	TCAATTTTATATCAACATCTTTCTGATTTTTTG	634
Tarucus_theophrastus	ATATTATTAACAGATCGAAATCTTAATACATCATTTTTTGATCCAGCAGGAGGAGGAGAT	600	CCAATTTTATATCAACATCTATTTTGATTTTTTG	634
Belenois_aurota	ATATTATTAACTGATCGAAATTTAAATACATCATTTTTTGATCCTGCTGGAGGTGGAGAC	600	CCTATTTTATATCAACATTTATTCTGATTTTTTG	634
Ephyseria_subdiminutella	ATACTTTTAACAGATCGTAATTTAAATACTTCTTTTTTTGACCCTGCTGGTGGGGGAGAT	600	CCAATTTTATACCAACATTTATTTTGATTTTTTG	634
Tuta_absoluta	ATATTATTAACGGATCGTAATTTAAATACATCATTCTTTGATCCCTGCAGGTGGAGGAGAC	600	CCAATTTTATATCAACATTTATTTTGATTTTTTG	634
Batrachedra_amydraula	ATATTATTAACAGATCGAAATaTTAATACTTCtTTTTTTGAcCCcGCtggAGGAGGAGAT	600	CCTATTTTATAcCAACATTTATTTTGATTTTTTG	634
Idolus_picipennis	ATACTACTGACTGATCGTAATTTAAACACATCGTTCTTCGACCCTGCAGGAGGTGGGGAC	600	CCAATCCTTTACCAACATTTATTCTGATTTTTTG	634
Melanotus_villosus	ATATTATTAACTGACCGAAACCTTAATACTTCCTTCTTTGACCCGGCAGGTGGGGGrGAC	600	CCCATCCTTTACCAACACTTGTTCTGATTTTTTG	634
Algarobius_prosopis	ATACTTCTTACAGACCGTAATTTAAACACTTCATTCTTTGACCCGGCAGGAGGGGAGAC	600	CCTATTCTTTATCAACACTTATTTTGATTTTTTG	634
Exomala_pallidipennis	ATACTATTAACAGATCGAAATTTAAACACTACATTTTTTGATCCATCAGGAGGTGGAGAT	600	CCAATCTTATATCAACACTTATTTTGATTTTTTG	634
Xylotrupes_siamensis	ATACTATTAACAGATCGAAATTTAAACACTTCCTTTTTTGACCCTGCAGGGGGAGGAGAT	600	CCAATTTTATTTCAACACCTATTCTGATTTTTTG	634
Copris_tripartitus	ATACTATTAACAGATCGAAATATTAATACTTCCTTCTTTGACCCTGCAGGGGGAGGAAAT	600	CCAATTTTATACCAACACCTATTCTGATTTTTTG	634
Lispe_albitarsis	ATATTTATTAACTGATCGAAATTTTAATACTTCATTCTTTGACCCAGCAGGAGGAGGAGAT	600	CCAATTCTTTACCAACATTTATTTTGATTTTTTG	634
Chrysomya_chani	ATACTATTAACAGATCGAAATTTAAATACATCATTTTTTGATCCAGCAGGTGGAGGAGAT	600	CCAATTCTTTATCAACACTTATTTTGATTTTTTG	634
Anthomyiopsis_nigrisquamata	ATACTATTAACAGATCGAAATTTAAATACATCATTTTTTGATCCAGCAGGTGGAGGAGAT	600	CCAATTCTTTATCAACACTTATcTTGATTTTTTG	634
Carpomya_vesuviana	ATATTATTAACAGACCGAAATTTAAATACTTCATTCTTTGATCCTGCTGGTGGAGGAGAT	600	CCTATTTTATATCAACATTTATTTTGATTTTTTG	634
Tabanus_striatus	ATATTATTAACAGATCGTAATTTAAATACTTCTTTTTTTGATCCAGCAGGAGGTGGAGAT	600	CCTATTTTATAcCAACAtTATTTTGATTTTTTG	634
Leptogaster_cylindrica	ATATTATTAACAGACCGAAATCTTAATACTTCATTTTTTGACCCAGCaggTGGAGGTGAC	600	CCAATTCTATATCAACATTTATTTTGATTTTTTG	634
Asyndetus_sp.	ATATTATTAACTGATCGAAATTTAAATACTTCATTTTTTGATCCTGCAGGAGGAGGTGAT	600	CCAATTTTATATCAACACTTATTTTGATTTTTTG	634

Enzyme	Restriction site
HhaI	GCGC
Bso31I	GGTCTC
AleI	CAC(N)$_4$GTG
ApaLI	GTGCAC
AccII	CGCG
Eco52I	CGGCCG
BciVI	GTATCC
BcgI	CGA(N)$_6$TGC
PinAI	ACCGGT
BspCNI	CTCAG
BtsI	GCAGTG
NheI	GCTAGC
PvuI	CGATCG
MunI	CAATTG
BclI	TGATCA
DpnI	GATC
LpnPI	CAGG (or CTGG)

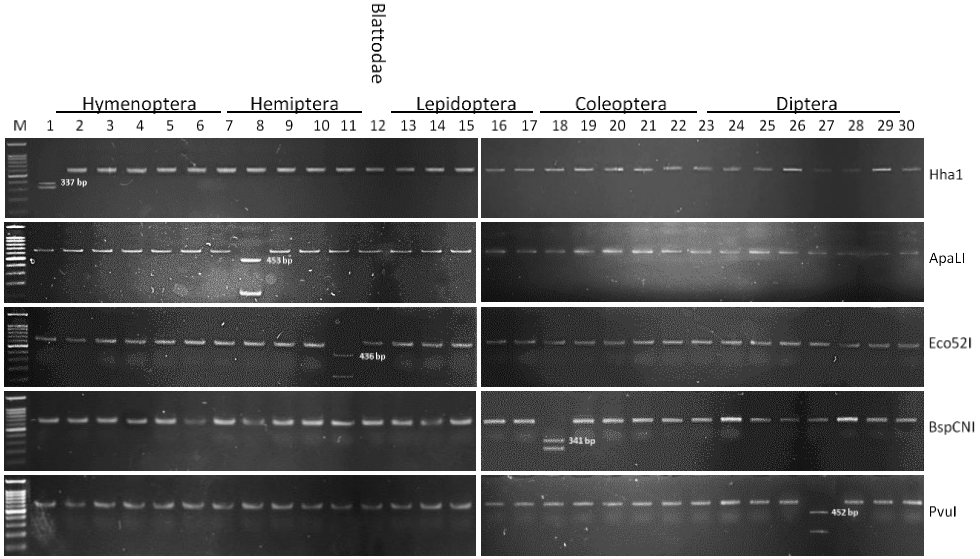


Figure A2. *Electrophregrams describing specific PCR-RFLP markers in COI gene in five insect species belonging to four insect orders used for validating restriction sites resulted from CisSERS tools. Each restriction enzyme has a unique site in the gene for a given insect species. Species numbers (1-30) refer to insect species shown in Table 1*

http://www.aloki.hu ● ISSN 1589 1623 (Print) ● ISSN 1785 0037 (Online)
DOI: http://dx.doi.org/10.15666/aeer/1705_1140911418