A Locus with Restriction Fragment-Length Polymorphisms Characteristic of African and European Honey Bee (Hymenoptera: Apidae) Groups of Subspecies

ALONSO SUAZO,[1] MYEONG-LYEOL LEE,[2] AND H. GLENN HALL[1,3]

Ann. Entomol. Soc. Am. 95(1): 115–124 (2002)

ABSTRACT Within an anonymous region of honey bee DNA (locus 227) digested with *Alu*I, informative restriction fragment-length polymorphisms (RFLP) were found in Southern blots with a cloned honey bee DNA probe. The probe was subcloned, so that smaller sections of the locus could be analyzed with the polymerase chain reaction (PCR). Further screening of these amplified sections revealed additional useful RFLPs with *Hin*fI. The informative *Alu*I and *Hin*fI polymorphic sites were mapped to a narrow section of the original probe. A total of 14 sub-alleles was found in this region of which five were found only or predominantly in our African samples (*Apis mellifera scutellata* Lepeletier), three were found only or predominantly in our east European samples (*A. m. ligustica* Spinola' *A. m. carnica* Pollman, and *A. m. caucasica* Gorbachev), one was found predominantly in our west European samples (*A. m. mellifera* L. and *A. m. iberica* Goetze). Significant associations were found between the *Alu*I and *Hin*fI sub-alleles, reinforcing their subspecies group specificity.

KEY WORDS African and European honey bees, restriction fragment-length polymorphisms, polymerase chain reaction, allele frequencies

RESTRICTION FRAGMENT-LENGTH polymorphisms (RFLP) have been found in honey bee mitochondrial DNA (mtDNA) and nuclear DNA (nDNA) that distinguish African and European groups of subspecies (Hall 1986, 1990, 1992a, 1992b; Smith and Brown 1988; McMichael and Hall 1996). These markers have been valuable in identifying African bee swarms intercepted at Florida deep-sea ports (Hall 1992c) and in revealing processes involved in African bee spread in the New World (Hall and Muralidharan 1989, Smith et al. 1989; Hall 1990; reviewed in Hall 1991, 1992a, 1999).

The nDNA RFLPs were first detected in Southern blots with cloned, radiolabeled probes. The technical complexity of this method (Southern 1975, Hall 1995) makes it difficult to analyze the numbers of samples involved in population studies. Methods based on the PCR (Saiki et al. 1988) have been used to facilitate testing for both mtDNA and nDNA markers. We and other investigators have developed ways to identify effectively honey bee mtDNA types from RFLPs in PCR-amplified regions (Crozier et al. 1991, Hall and Smith 1991, Garnery et al. 1993, Nielsen et al. 2000). Within nuclear DNA, we have identified random amplified polymorphic DNA (RAPD) (Williams et al. 1990) and amplified fragment-length polymorphisms (AFLPs) (Vos et al. 1995) that appear to be specific to either African or European bees (Suazo et al. 1998, Suazo and Hall 1999). Other investigators have used microsatellites to establish phylogenetic relationships among honey bee subspecies (Estoup et al. 1995). As with mtDNA, we are amplifying valuable nDNA regions to make RFLPs, first detected in Southern blots, easier to analyze (Hall 1998). As discussed below, the different PCR-based methods have advantages and disadvantages to be considered, depending on the application (Sunnucks 2000). PCR-RFLP analysis serves our purpose as a reliable, consistent, and technically easy way to distinguish African and European honey bees.

Previously, we reported on one nDNA polymorphic locus in which African-specific RFLP alleles were first detected with Southern blots and of which smaller sections were subsequently amplified with the PCR (Hall 1998). We have found other useful loci with alleles of different specificity. As described in this report, subregions of a honey bee nDNA locus, 227, were amplified. From within these subregions, RFLP patterns generated with *Alu*I and *Hin*fI, referred to as sub-alleles, were found to be characteristic of different groups of honey bee subspecies. The informative sites for these enzymes were mapped. Frequency data were obtained for these sub-alleles in Old and New World African and European samples. The PCR-RFLPs at this locus increase the number and broaden the specificity of available DNA markers to identify African and European bees and to follow introgression between the two.

[1] Department of Entomology and Nematology, University of Florida, P.O. Box 110620, Gainesville, FL 32611.
[2] National Institute of Agricultural Science and Technology, Suwon 441–100, South Korea.
[3] E-mail: hgh@gnv.ifas.ufl.edu.

0013-8746/02/0115–0124$02.00/0 © 2002 Entomological Society of America

Table 1. Honey bee samples

Subspecies/Population	Country	Location	No. of colonies	No. of bees	Reference
A. m. scutellata	South Africa	Johannesburg	3	6	McMichael and Hall 1996
		Pretoria	21	42	
		Warm Baths	13	26	
		Louis Trichardt	7	14	
		White River	10	20	
A. m. iberica	Spain	Oviedo	7	7	Smith et al. 1991
		Lugo	2	2	
		Córdoba	6	6	
A. m. mellifera	France	Landes	5	9	
		Montfavet	4	4	
A. m. ligustica	Italy	Bologna	12 + 5[a]	17	
		Milano	7	8	
		Morbegno	12	12	
		Sondrio	5	5	
A. m. carnica	Austria	Graz	4 + 3[a]	7	Smith and Brown 1990
		Klagenfurt	4	4	
		Lunz-am-see	3	3	
		Eastern Austria	3	3	
	Germany	Hamburg	3	3	
	Former Yugoslavia	Split (province of Dalmatia)	3	3	
		Medvode (province of Slovenia)	3	3	
A. m. caucasica	Rep. Of Georgia	Ile d'Yeu[a,b]	7	20	
New World European	U.S.A.	Kansas	21	105D[c]	
		Tucson	23	48	Severson et al. 1986, Page et al. 1982

[a] Bees maintained in research apiaries.
[b] Colonies maintained in an experimental apiary in Ile d'Yeu, France.
[c] Five drones per colony were used to determine the queen genotype from which the sub-allele frequencies were calculated.

Materials and Methods

Honey Bee Samples and DNA Isolation. Honey bee samples were obtained either as larvae, pupae, or adults of drones (haploid males) and workers (diploid females). The origins and numbers of colonies are listed in Table 1. Samples from Europe were grouped as either west European (*A. m. mellifera* and *A. m. iberica*) or east European (*A. m. ligustica*, *A .m. carnica*, and *A. m. caucasica*) as suggested by mitochondrial and nuclear DNA markers (Garnery et al. 1992, Arias and Sheppard 1996, McMichael and Hall 1996, Smith et al. 1997, Franck et al. 2000). Samples from Tucson, AZ, were from a closed breeding population representing bee stocks from across the United States (Page et al. 1982, Severson et al. 1986). South African samples were *A. m. scutellata*. One to three individuals per colony were tested. DNA was isolated from drone pupae and isolated thoracic muscles of adult workers as described in Hall (1986, 1990) and with anionic columns as suggested by the manufacturer (Qiagen, Palo Alto, CA). For Southern blot analysis, drone samples were used, primarily because of their simple haploid pattern of restriction fragments.

Probe Description and Screening for Polymorphisms. Probe 227 (P227) is a 9.2 kilobase pair (kb) fragment of anonymous, single copy, honey bee DNA inserted into the *Pst*I site of plasmid pBR322. The methods used to construct the library, from which P227 was selected, and the Southern blot procedures used to initially seek RFLPs are described by Hall (1986, 1995).

Subcloning. For analysis by the PCR, P227 was subcloned. Restriction sites of enzymes contained within the polycloning site of vector pGEM3Z (Promega, Madison, WI) were mapped in P227 using standard double digests (Sambrook et al. 1989). Three enzymes were selected, *Xho*I, *Sph*I, and *Hin*dIII, that divided the probe into four suitably sized sections amplifiable by standard PCR (Fig. 1): section 1 (S1), from the terminal *Pst*I site (position 0) to the *Xho*I site (position 2500) (2.5 kb), section 2 (S2) from the *Xho*I site to the *Sph*I site (position 3900) (1.4 kb), section 3 (S3) from the *Sph*I site to the *Hin*dIII site (position 7100) (3.2 kb), and section 4 (S4) from the *Hin*dIII site to the second terminal *Pst*I site (position 9200) (2.1 kb). The vector and fragments of P227, digested with the appropriate pairs of enzymes, were separated by agarose gel electrophoresis, excised from the gel, electroeluted from the agarose (in dialysis bags, in concentrated Tris borate buffer; Sambrook et al. 1989), purified on anionic resin columns (Qiagen), precipitated and washed with ethanol, and dissolved in water. The digested and purified vector and the P227 fragments were ligated by standard procedures and used to transform competent DH5α cells (Life Technologies, Gibco BRL, Gaithersburg, MD). White colonies (with inserts) grown on medium with X-gal were selected. Insert sizes were verified by electrophoresis after restriction of the isolated plasmids. Subclones corresponding to each section were used separately as radioactive probes with the original blots to find the sections that contained the informative polymorphic sites. Subclones revealing informative polymorphisms were sequenced $\approx$500 bases from their terminal ends (University of Florida, Interdisciplinary Center for Biotechnology Research, ICBR). From these terminal

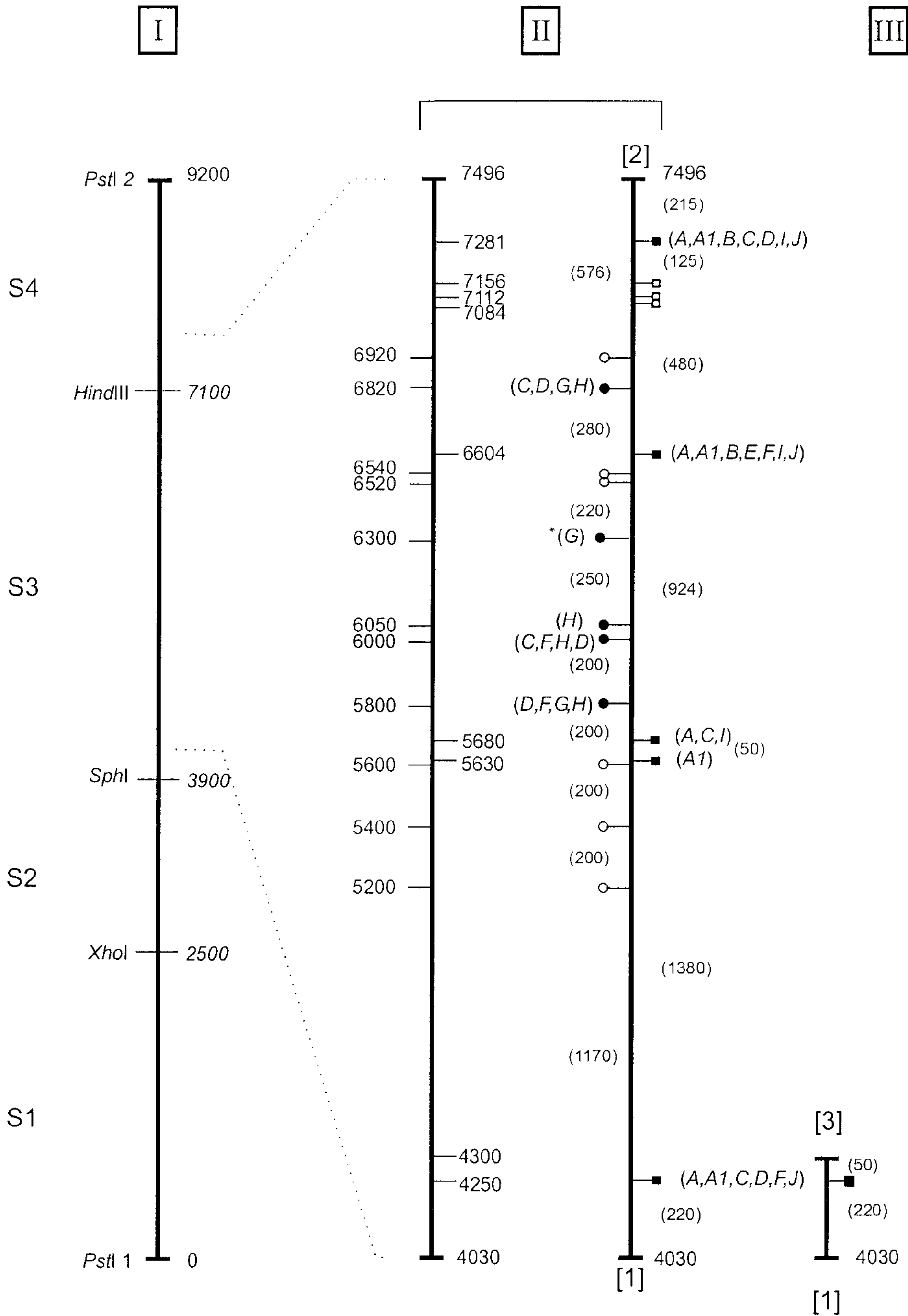

Fig. 1. Restriction enzyme cleavage site map for locus 227 and the amplifiable polymorphic sections. (I) General map of locus 227 with the enzyme sites used to subclone the different sections. (II) Position of *Alu*I and *Hin*fI sites in the polymorphic section 3 extended (P227-S3xt). (III) P227-S3ter used to distinguish sub-alleles *Alu*I-*B* and *Alu*I-*J*. Locus 227 was divided into four sections, S1 to S4, defined by the restriction sites *Xho*I, *Sph*I, and *Hin*dIII. Squares and circles correspond to *Alu*I and *Hin*fI sites, respectively. Open and closed symbols represent monomorphic and polymorphic sites, respectively. Letters in parentheses are the sub-alleles that have indicated *Alu*I and *Hin*fI sites. Position of restriction sites, in base pairs, are indicated to the right (*Alu*I) or left (*Hin*fI) side of the map. Primers used to amplify each region are indicated in brackets above and below each diagram. A single individual star indicates the polymorphic *Hin*fI site at position 6300, resulting in allele *Hin*fI-*G* that is west European-predominant (*A. m. mellifera* and *A. m. iberica*). P227-S3ter (far right) includes the polymorphic site that distinguishes sub-alleles *Alu*I-*B* (common) and *Alu*I-*J* (African).

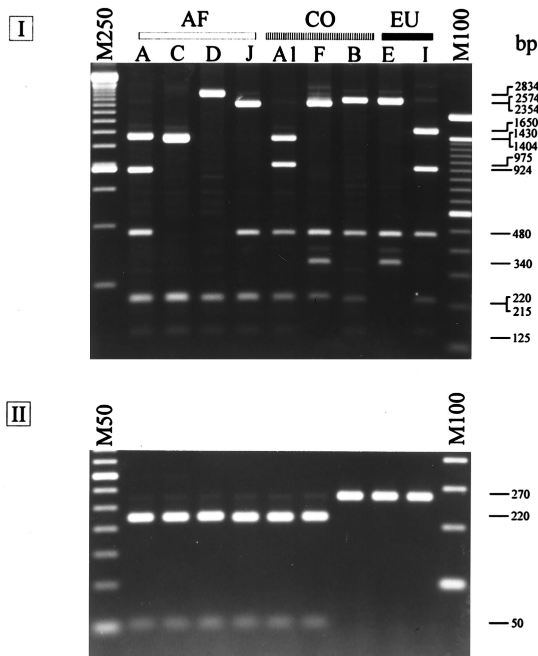

Fig. 2. Restriction digest profiles of P227-S3xt (I), and P227-S3ter (II) with *Alu*I. AF = African sub-alleles, CO = common sub-alleles and EU = European sub-alleles. Alleles in the upper panel are the same as the lower panel. M50, M100, and M250 are molecular size standard, 50, 100, and 250 bp ladders, respectively (Life Technologies). Fragment sizes, in base pairs, are shown to the right. DNA samples amplified were all from drones. See text for explanation of the fragment patterns. Sub-alleles *Alu*I-*B* and *Alu*I-*J* are difficult to distinguish in RFLP patterns of P227-S3xt (I) but are clearly distinguishable in *Alu*I digests of P227-S3ter (II).

sequences, primer sequences were selected using the computer software package OLIGO 5.0 (National Bioscience, Plymouth, MN).

Amplification and RFLP Analysis of PCR Products. Almost the entire S3 of P227 with the adjacent end of S4 was amplified as one segment, referred to as section 3 extended (P227-S3xt). Primers used were [1] 5′-AGAAGGAAAGAAGAAACGGATGAAC-3′ within S3 (position 4030) near the *Sph*I site and [2] 5′-CGGAGGAGTGGTAATAATGGAAGC-3′ within S4 (position 7496) near the *Hin*dIII site (Fig. 1). Amplifications were performed in 25 μl of 50 mM Tris-HCl pH 9.2, 16 mM $(NH_4)_2SO_4$, and 1.8 mM $MgCl_2$, containing 200 μM of each dNTPs, 250 nM of each primer, with 1.12 U of *Taq* DNA polymerase and 0.12 U of *Pwo*I DNA polymerase (taken from a stock mixture containing nine units of *Taq* DNA polymerase and one unit of *Pwo*I DNA polymerase; Boheringer Mannheim, Indianapolis, IN). Fifty to 250 ng of genomic DNA were used per reaction (the highest amount of DNA was used when DNA was extracted from thoracic muscles of adult honey bees). Amplifications were performed in a PTC-100 or PTC-200 thermocycler (MJ Research, Watertown, MA), and the following profile was used: 95°C for 2 min followed by 33 cycles each at 94°C for 45 s, 57°C for 45 s, and 68°C for 4 min with a final extension at 68°C for 10 min. A second section near the *Sph*I terminus of S3 (position 4030–4300), referred to as P227 section 3 terminal (P227-S3ter), was amplified using primer [1] from P227-S3xt and primer [3] 5′-CACTTTTTCAAAAGAGCGTGCAAAG-3′ (position 4300) (Fig. 1). The solution for amplifying P227-S3ter was as for P227-S3xt, except that 1.0 mM $MgCl_2$ and one unit of *Taq* DNA polymerase were used (Life Technologies GIBCO BRL). The amplification profile for P227-S3ter was as follows: 95°C for 2 min followed by 35 cycles of 94°C for 45 s, 60°C for 45 s, and 72°C for 30 s with a final extension of 5 min at 72°C. Following the protocols in Hall 1998, P227-S3xt was screened for 22 restriction endonucleases, some of which had been used in the initial screening with southern blots: *Hae*III, *Hin*fI, *Msp*I, *Dra*I, *Ava*II, *Bam*HI, *Eco*RI, *Dde*I, *Afl*II, *Bgl*I, *Nru*I, *Stu*I, *Bcl*I, *Nhe*I, *Bst*XI, *Eco*RV, *Spe*I, *Hpa*I, *Nsi*I, *Bgl*II, *Nco*I, and *Sty*I. Promising polymorphisms found between nine African and nine European unrelated drones were tested in a larger sample size consisting of 48 unrelated African and European drones. The restriction enzymes *Alu*I and *Hin*fI were found to generate informative polymorphisms.

The *Alu*I and *Hin*fI restriction sites for the different sub-alleles were mapped using the procedure described by Her and Weinshilboum (1995). Primers were radioactively end labeled with $^{\gamma 32}$P and used as probes in Southern blots containing partially digested DNA from PCR products of samples corresponding to each sub-allele. Autoradiography of these blots revealed the position of the restriction sites with respect to the primer used as probe.

Sub-Allele Counts and Data Analysis. The different sub-allelic fragment patterns generated by *Alu*I and *Hin*fI were identified in haploid drones and were deciphered in diploid workers. Each pattern identified from the amplified sections from the original P227 is considered a sub-allele. Sub-alleles were named using the enzyme name followed by a letter; for example, *Alu*I-*A*. Drones were used to determine the queen genotypes according to Hall (1998), from which sub-allele frequencies were obtained. The haplotype of each drone was assigned a combination of two letters corresponding to the sub-alleles produced by *Alu*I and *Hin*fI. These combinations are referred to as composite sub-alleles. Composite sub-alleles in drone samples were used to determine the queen's genotype from which composite sub-allele frequencies were obtained, and to generate a contingency table in which sub-allele frequency associations for *Alu*I and *Hin*fI was determined. Sub-allele association was evaluated for independence using a Fisher exact test with a contingency table (under the null hypothesis of no association) based on the metropolis algorithm from the computer software package RxC (M. Miller, University of Arizona: http://www.public.asu.edu/≈mmille8/rxc.htm). Frequency data from workers were obtained from direct counts of *Alu*I and *Hin*fI sub-alleles. Observed and expected heterozygosities (H_{obs} and H_{exp}) were calculated as described in Nei

Table 2. Relative frequencies of *Alu*I sub-alleles of locus 227 in African and European populations

Sub-allele	African	United States	West European		East European		
	South Africa (*A. m. scutellata*)		France (*A. m. mellifera*)	Spain (*A. m. iberica*)	Italy (*A. m. ligustica*)	Austria (*A. m. carnica*)	Rep of Georgia (*A. m. caucasica*)
(N)	108	60	10	13	42	24	19
A^a	0.1600		0.0500				
A1	0.3090	0.0250	0.3000	0.6923		0.1667	
B	0.0570	0.0583				0.0417	0.0789
$C^{a,b}$							
$D^{a,b}$							
E^c		0.1000					0.9211
F	0.0410	0.0667					
I^c	0.0260	0.7417	0.6500	0.3077	1.0000	0.7917	
J^a	0.4070	0.0083					
C (*A,J*)	0.5670	0.0083	0.0500				
C_{EE} (*I,E*)	0.0260	0.8417	0.6500	0.3077	1.0000	0.7917	0.9211
H_{obs}	0.4567	0.3000	0.3000	0.4615	0.0000	0.1667	0.1579
H_{exp}	0.7080	0.4314	0.4850	0.4260	0.0000	0.3438	0.1454

H_{obs} and H_{exp}: Observed and expected heterozygosities. *N*, number of workers tested. C_{AF} and C_{EE}: Collective sub-allele frequencies in African and east European honey bees, respectively.
[a] Sub-alleles predominantly found in *A. m. scutellata* samples.
[b] Sub-alleles not detected in workers and found only in two drone samples.
[c] Sub-alleles predominantly found in east European bees.

(1978, 1987), with the computer software package GENETIX version 4.02 (Belkhir et al. 2001).

Recombination and Sub-Allele Identity Changes. Recombination between polymorphic sites was evaluated for all pairwise sub-allele combinations by overlaying the restriction site profile of two sub-alleles. Restriction site profiles of sub-alleles generated from recombination between every polymorphic site were compared with the paired profiles for changes in sub-allele identity or formation of new sub-alleles. For five paired inter-site regions for each enzyme, *Alu*I and *Hin*fI, 36 and 10 possible sub-allele combinations, respectively, were evaluated.

Results

Initial screening (as in Hall 1986) using P227 as a probe in Southern blots revealed informative polymorphisms with *Alu*I. Nine alleles were found of which six were found only in our African samples and three were found only in our European samples. P227 was divided in four amplifiable sections defined by the positions of the *Xho*I, *Sph*I, and *Hin*dIII restriction sites: section 1 (S1) from position 0–2500; section 2 (S2) from position 2500–3900; section 3 (S3) from position 3900–7100; and section 4 (S4) from position 7100–9200. All polymorphic bands resulting from *Alu*I digests were detected when both section 3 (S3) and section 4 (S4) subclones of P227 were used as probes. S4 detected only one polymorphic *Alu*I site located at the terminal end adjacent to S3 (determined from the sequence). RFLP analysis of the amplifiable P227-S3xt also revealed polymorphisms with *Hin*fI.

***Alu*I Polymorphisms.** *Alu*I digestions of P227-S3xt region revealed nine sub-alleles (Fig. 2). Sub-alleles *A, C, D,* and *J* were found at a collective frequency of 56.7% in the South African *A. m. scutellata* samples (Table 2). Sub-allele *A* was found in the *A. m. scutellata* and the west European *A. m. mellifera* samples at frequencies of 16.0 and 5.0%, respectively. Sub-allele *J* was found at a 40.7% frequency in the African samples. Sub-allele *J* was absent from both east and west European samples but present at a very low frequency (0.8%) in the U.S. samples. Sub-alleles *C* and *D* were found only in two African drones and were absent in all our worker samples. Two sub-alleles, *I* and *E,* were predominantly European. Sub-allele *I* was found at frequencies of 100.0 and 79.1% in the east European samples of *A. m. ligustica* and *A. m. carnica*, respectively, and 65.0 and 30.7% in the west European samples of *A. m. mellifera* and *A. m. iberica*, respectively. Sub-allele *I* was absent from our samples of *A. m. caucasica* but found at a low frequency (2.6%) in the *A. m. scutellata* samples. Sub-allele *E* was found in our *A. m. caucasica* samples at a frequency of 92.1% and was absent in all the African and other European samples. Sub-alleles *I* and *E* were found at a collective frequency of 84.2% in the U.S. samples. Sub-alleles *B* and *A1* were common to African and European samples. Sub-allele *F* was found in South African and U.S. samples.

The African sub-alleles *A* and *C* share a 1430 bp fragment (Fig. 2). Sub-allele *A* is distinguished from sub-allele *C* by having a 924 bp and a 480 bp fragment rather than a 1404 bp fragment. African sub-allele *D* has a unique 2834 fragment. African sub-allele *J* shares a 2354 bp and a 220 bp fragment with the common sub-allele *F.* Loss of an *Alu*I site between these fragments (position 4250) results in a 2574 bp fragment in the east European (*A. m. caucasica*) sub-allele *E* and common sub-allele *B,* which migrates close to the 2354 bp fragment. Sub-allele *J* is distinguished from sub-allele *F,* and more clearly from sub-allele *E,* by having a 215 bp (which co-migrates with the 220 bp fragment) and a 125 bp fragment rather than a 340 bp fragment. However, African sub-allele *J* and common sub-allele *B* are difficult to distinguish because of the close migration of the 2354 bp and 2574 bp fragments and the

Table 3. Relative frequencies of *Hin*fI sub-alleles of locus 227 in African and European populations

Sub-Allele	African	United States	West European		East European		
	South Africa (*A. m. scutellata*)		France (*A. m. mellifera*)	Spain (*A. m. iberica*)	Italy (*A. m. ligustica*)	Austria (*A. m. carnica*)	Rep of Georgia (*A. m. caucasica*)
(N)	108	60	12	14	41	24	19
C[a]							
D	1.0000	0.2583			0.4390	0.1458	1.0000
F[b]		0.4167			0.5610	0.6667	
G[c]		0.2083	1.0000	1.0000		0.1875	
H		0.1167					
C_{EE} (*F*)		0.4167			0.5610	0.6667	
C_{WE}(*G*)		0.2083	1.0000	1.0000		0.1875	
H_{obs}	0.0000	0.5000	0.0000	0.0000	0.8780	0.4167	0.0000
H_{exp}	0.0000	0.7026	0.0000	0.0000	0.4926	0.4991	0.0000

H_{obs} and H_{exp}: Observed and expected heterozygosities. *N*, number of workers tested. C_{EE} and C_{WE}: Collective sub-allele frequencies in east and west European honey bee, respectively.
[a] Sub-alleles not detected in workers and found only in two drone samples.
[b] Sub-alleles predominantly found in east European bees.
[c] Sub-alleles predominantly found in west European bees.

220 bp and 215 bp fragments. Sub-alleles *B* and *J* can be distinguished by amplifying and digesting a shorter section (P227-S3ter), which contains the polymorphic *Alu*I site at position 4250. The presence of a 270 bp fragment (uncut) identifies the *B* sub-allele, whereas its absence identifies the *J* sub-allele. The European sub-allele *I* has a unique 1650 bp fragment.

Genotypes that cannot be distinguished include: *A*/*A* from *A*/*C*, *A1*/*A1* from *A1*/*C*, and *A*/*I* from *C*/*I*. The inability to distinguish these genotypes can cause some inaccuracy in the frequencies determined for both sub-alleles *A* and *A1*. Sub-allele *C* was found at such low frequencies (found only in one drone among our samples, corresponding to a frequency of 0.3%) that it would not result in many uncertain genotypes.

***Hin*fI Polymorphisms.** Five sub-alleles were revealed with *Hin*fI, of which two, *F* and *G*, were found only in our European samples. Sub-allele *F* was found only in east European samples at frequencies of 56.1% in *A. m. ligustica* and 66.7% *in A. m. carnica* (Table 3). Sub-allele *G* was fixed in the west European samples, a 100.0% frequency in both *A. m. mellifera* and *A. m. iberica*, and was found at an 18.8% frequency in the *A. m. carnica* samples. Sub-allele *H* was found only in the U.S. samples at an 11.7% frequency. Sub-alleles *F*, *G* and *H* were found at a collective frequency of 74.2% in the U.S. samples. One sub-allele (*C*) was found only in our South African *A. m. scutellata* samples at a frequency of <1.0%. Sub-allele *D* was common to African and European samples.

Sub-alleles *C*, *F*, *G*, and *H* each have unique fragments of 400, 380, 250 and 225, and 470 bp, respectively (Fig. 3). Indistinguishable genotypes include *C*/*C* from *C*/*D* and *G*/*G* from *G*/*D*.

A strong association between *Alu*I and *Hin*fI sub-alleles was found ($P < 0.0001 \pm 0.000$ Fisher exact test) (Table 4). Although the *Alu*I and *Hin*fI sites are present along the same stretch of DNA, which is the same locus, the sites are at independent positions and may have formed at different evolutionary stages. Some strong associations are to be expected but not necessarily. Where strong associations are found, they mutually reinforce their discriminating power and, hence, the value of these markers for diagnostic purposes. For example, *Alu*I-*I*, a predominantly east European sub-allele, was generally found associated with *Hin*fI-*F*, a sub-allele that may be specific to east European bees. African *Alu*I sub-alleles were found associated with the single African or single common *Hin*fI sub-alleles.

*Alu*I and *Hin*fI sub-alleles together, along with the other undetected polymorphisms, define the actual alleles or haplotypes. However, if the patterns from the two enzymes were not dealt with separately, the composite patterns would often be too complex to decipher, particularly in diploid, heterozygous, work-

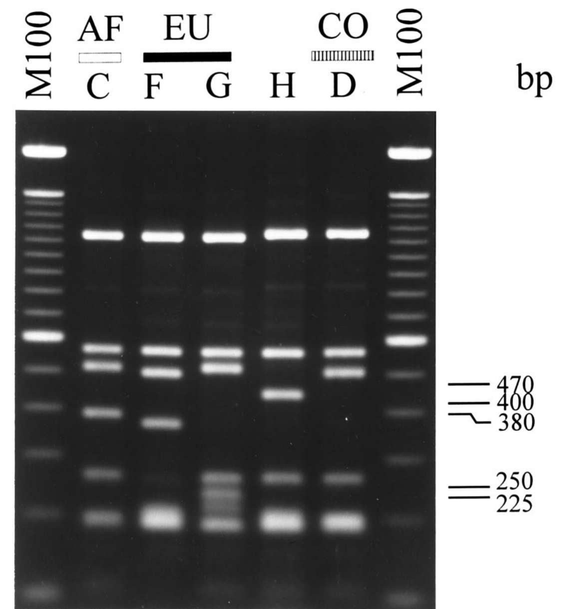

Fig. 3. Restriction digestion profile for P227-S3xt with *Hin*fI. AF, CO, EU, and M100 are as in Fig. 2. See text for explanation of the fragment patterns.

Table 4. Composite sub-allele frequencies for *Alu*I (first letter) and *Hin*fI (second letter) sub-alleles obtained from drone samples

Composite sub-allele	Population	
	South Africa	United States
N	20	30
A1-D	15 (0.385)	
A^a-*D*	5 (0.128)	
B-D	5 (0.128)	
C^a-*D*	1 (0.026)	
D^a-*D*	1 (0.026)	
J^a-C^a	2 (0.051)	
J^a-*D*	9 (0.231)	
B-H		7 (0.119)
E^b-*D*		8 (0.136)
F-D		2 (0.034)
F-H		5 (0.085)
I^b-F^b		28 (0.475)
I^b-G^c		7 (0.119)
I^b-*D*	1 (0.026)	2 (0.034)
Total	39	59

n, number of colonies used. Drones were used to determine the queen's genotype from which composite sub-allele frequencies were obtained.
[a] African-specific or predominant sub-alleles.
[b] East European-specific or predominant sub-allele.
[c] West European-specific or predominant sub-allele.

ers. Thus, the most feasible approach is first to collect and interpret the information from each enzyme and later to combine the data found in single individuals to identify composite alleles or sub-alleles. Fourteen composite sub-alleles were found in the African and the U.S. populations, of which five were African-specific and four were European-specific (Table 4). A composite sub-allele was African or European-specific if one of its sub-alleles was African or European-specific.

Observed heterozygosity values for the sub-alleles found with *Alu*I were higher in South African and west European bees compared with east European bees (Tables 2 and 3). Compared with the *Alu*I sub-alleles, heterozygosity values found with *Hin*fI sub-alleles were higher in east European bees compared with South African and west European bees.

Recombination and Sub-Allele Identity Changes. Sub-allele identity changes as a consequence of possible recombination between the polymorphic sites of every possible pairwise combination of sub-alleles were determined. Such events would be rare because of the close linkage between the sites, especially between the *Hin*fI sites. The five polymorphic *Alu*I sites could result in 32 sub-alleles, nine of which have been found. A total of 144 differently paired inter-site regions could exist among the nine sub-alleles. Recombination within 27 of these paired regions would result in sub-allele identity changes. Recombination within only seven regions would create African or European sub-alleles: recombination between sites 4250 and 5680 of African sub-allele *A* and common sub-allele *B* would form European sub-allele *I*, between any of the sites (3 regions) of European sub-allele *I* and common sub-allele *F* would form African sub-alleles *J* or *A*, and between any of the sites of common sub-alleles *B* and *F* would form African sub-allele *J* and European sub-allele *E*. Recombinations between other sub-alleles would result in the change in identity of an African sub-allele to another African sub-allele or in the loss of a European or African sub-allele and the creation of a common or new sub-allele. The five polymorphic *Hin*fI sites could result in 32 sub-alleles, five of which have been found. A total of 40 differently paired inter-site regions could exist among the five sub-alleles. Recombination within 19 of these paired regions would result in sub-allele identity changes. Recombination within 12 of these regions would result in the loss of a European sub-allele and the creation of a common or new sub-allele.

Discussion

Differences in DNA, detected with a number of methods, serve as valuable genetic markers. The methods differ in their level of technical difficulty and the quality and complexity of genetic data obtained (Sunnucks 2000). Multilocus methods, such as RAPD and AFLP, are technically convenient and reveal large numbers of markers. Generally, the data obtained by these methods are not of the highest quality. Most RAPD and AFLP markers are dominant, that is, homozygous and heterozygous individuals cannot be distinguished. RAPD and AFLP analyses do not target specific sequences in the genome. Thus, bands can be generated from contaminating DNA of microorganisms or other sources (Ayliffe et al. 1994, Rabouam et al. 1999). RAPD has been criticized for artifacts often generated, as well as low reproducibility and reliability, thereby making comparisons among laboratories difficult (Pérez et al. 1998). The bases for the polymorphisms detected in RAPD are not well understood.

Single-locus methods, involving RFLPs and microsatellites, can be technically difficult but produce data of high quality. Alleles are codominant, that is, both can be identified in heterozygous diploid organisms. Thus, these markers are useful for following introgression and for conducting parental analyses. Specific regions of DNA are targeted, and amplified DNA is less likely to be from contaminants. The bases for the different RFLP and microsatellite patterns are well understood. Microsatellite patterns are easy to interpret. RFLP patterns can be simple or very complex, sometimes making data interpretation difficult.

Both the technical implementation and data interpretation of RFLP analyses have been greatly simplified with the PCR. Informative DNA regions, digested with the restriction enzymes, are not visualized with probes but, rather, are amplified. The regions of the genome are generally shorter than those detected with probes, and banding patterns are less complex. Different allelic forms can be amplified, thereby facilitating the mapping of restriction sites. Consequently, more narrow informative regions can be selected and amplified, thereby further simplifying band patterns. Large numbers of samples can be tested for PCR-RFLP markers. The quality of the data remains high, except that heteroduplex artifacts can be gen-

erated from heterozygous individuals (Anglani et al. 1990), as can occur with RAPD analyses (Novy and Vorsa 1996). Also, some allelic forms can be preferentially amplified (Hare et al. 1996).

With Southern blots, we found a number of loci carrying useful RFLPs that distinguish African and European bees. To make these markers more readily available and easier to use, we have selected sequences from the original probes to amplify the regions with the PCR. One region, 227, has alleles revealed with the restriction endonucleases *Alu*I and *Hin*fI; alleles specific to or predominant in either African or European groups of honey bee subspecies. Digests with the two enzymes must be done separately, as patterns from double digests are too complex. Even with single enzymes, some complex patterns are generated. The allelic bands can be difficult to identify in diploid workers but are much easier to identify in haploid drones. The combinations of sub-allelic patterns, from separate *Alu*I and *Hin*fI digests of individual drones, identify the composite sub-alleles. From analysis of several drones from the bee colonies, the genotypes of the mother queens, from which the drones are parthenogenetically derived, are determined. Queen genotypes are used to determine the frequencies of composite sub-alleles. After the allelic patterns are initially identified in drones, the patterns can be more-easily recognized in workers from which frequencies can be obtained.

Sub-alleles of 227 generated by *Alu*I and *Hin*fI distinguish bees in the African (*A. m. scutellata*), west European (*A. m. mellifera* and *A. m. iberica*), and east European (*A. m. ligustica, A. m. carnica* and *A. m. caucasica*) groups of subspecies, making this a particularly valuable locus for identification. Frequency data of these sub-alleles in U.S. bees reflect the contributions of different subspecies introduced by European settlers to North America (Pellet 1938, Nelson 1967, Oertel 1976) and is consistent with those found for alleles at other polymorphic loci (Hall 1990, McMichael and Hall 1996). Most of the 227 sub-alleles in the United States are east European, which reflects the preference of the Italian honey bee *A. m. ligustica* among beekeepers. West European bees were introduced into North America long before east European bees. Because of their defensive behavior, they became less popular among beekeepers (Pellet 1938, Oertel 1976). The *A. m. caucasica* sub-allele *Alu*I-*E,* present at a low frequency in the U.S. samples, reveals a genetic contribution of this subspecies also.

The distribution of alleles in honey bee populations, as a result of natural events, is further complicated by frequent human transport of bees, adding to the difficulty of assigning subspecies specificity to the alleles. In the Old World, introduction of Italian queens from Italy to France brings *A. m. ligustica* into areas where *A. m. mellifera* evolved (Garnery et al. 1998a, 1998b). The composition of the Iberian bee population is already the result of a similar, but much earlier, natural migration and introgresssion of north African bees (Smith et al. 1991, Franck et al. 1998). European bees have been brought to South Africa (Fletcher 1973, 1978). In the New World, the honey bee populations came primarily from a few European and one African subspecies. However, there were other minor introductions. For example, *A. m. lamarckii* Cockerell from Egypt, and perhaps some near-eastern subspecies, were introduced into the United States in the early 18th century (Pellet 1938, Oertel 1976). Poor adaptability of European bees to tropical environments and of African bees to temperate environments probably prevented the establishment of these subspecies in the presence of the better-adapted, resident bees. Nevertheless, their genes may have introgressed and persisted to some extent. Evidence has been found in mtDNA of the non-*scutellata* African introductions into North and South America (Sheppard et al. 1999). Likewise, some nuclear DNA may persist from these early natural and artificial introductions and may explain the distribution of some of the 227 RFLP alleles. For example, the sub-allele *Alu*I-*F,* found at low frequencies in both the United States and South Africa, may reflect the early non-*scutellata* African introductions into North America. European predominant sub-alleles, such as *Alu*I-*I,* found in South Africa, may be a remnant of the unsuccessful European introductions there. Sub-allele *Alu*I-*A1* is found at a high frequency in *A. m. iberica,* at intermediate frequencies in *A. m. scutellata,* and at very low frequencies in the U.S. samples. This sub-allele may be common and specific to African subspecies, but present in the Iberian peninsula and in North America as a result of human-assisted importations.

Locus 227 is the second nDNA locus we have described to be first detected with probes and subsequently made analyzable with the PCR. The first locus had African-specific sub-alleles, whereas locus 227 has sub-alleles specific to or predominant in three major groups of honey bee subspecies. These markers can be applied to a wide range of genetic studies including the characterization of population genetic variation, the spread of African bees in the New World, and introgression of their genes into the resident European populations.

Acknowledgments

We thank Deborah R. Smith (University of Kansas), P. Vienne and M.-R. Guillevic for providing samples of Old World European honey bees; O. R. Taylor (University of Kansas), G. Loper, G. Waller, J. Martin, and E. Erickson (Carl Hayden Bee Research Center USDA-ARS) for providing honey bee samples from the United States. We also want to thank M. T. Sanford (University of Florida) for collecting samples. This research was supported by the USDA National Research Initiative Competitive Grants Program. Florida Agricultural Experiment Station Journal Series No R-08171.

References Cited

Anglani, F., L. Picci, Ch. Camporese, and F. Zacchello. 1990. Heteroduplex formation in polymerase chain reaction. Am. J. Hum. Genet. 47: 169–170.

Arias, M. C., and W. S. Sheppard. 1996. Molecular phylogenetics of honey bee subspecies (*Apis mellifera* L.) inferred from mitochondrial DNA sequence. Mol. Phylogenet. Evol. 5: 557–566.

Ayliffe, M. A., G. J. Lawrence, J. G. Ellis, and A. J. Pryor. 1994. Heteroduplex molecules formed between allelic sequences cause non-parental RAPD bands. Nucleic Acids Res. 22: 1632–1636.

Belkhir, K., P. Borsa, J. Goudet, C. Chikhi, and F. Bonhomme. 2001. Genetix v4.02 logiciel sous Windows pour la génétique des populations. Laboratoire Génome et Populations, CNRS UPR 9060. Université de Montpellier 2, France (http://www.univ-montp2.fr/~genetix/intro.htm).

Crozier, Y. C., S. Koulianos, and R. H. Crozier. 1991. An improved test for Africanized honey bee mitochondrial DNA. Experientia 47: 968–969.

Estoup, A., L. Garnery, M. Solignac, and J.-M. Cornuet. 1995. Microsatellite variation in honey bee (*Apis mellifera* L.) populations: hierarchical genetic structure and test of the infinite allele and stepwise mutation models. Genetics 140: 679–695.

Fletcher, D.J.C. 1973. Should European races of honey bees be introduced into South Africa? S. Afr. Bee J 45: 19–25.

Fletcher, D.J.C. 1978. The African bee, *Apis mellifera adansonii*, in Africa. Annu. Rev. Entomol. 23: 151–171.

Franck, P., L. Garnery, M. Solignac, and J.-M. Cornuet. 1998. The origin of west European subspecies of honeybees (*Apis mellifera*): new insights from microsatellite and mitochondrial data. Evolution 52: 1119–1134.

Franck, P., L. Garnery, M. Solignac, and J.-M. Cornuet. 2000. Molecular confirmation of a fourth lineage in honey bees from the Near East. Apidologie 31: 167–180.

Garnery, L., J.-M. Cornuet, and M. Solignac. 1992. Evolutionary history of the honey bee *Apis mellifera* inferred from mitochondrial DNA analysis. Mol. Ecol. 1: 145–154.

Garnery, L., M. Solignac, G. Celebrano, and J.-M. Cornuet. 1993. A simple test using restricted PCR-amplified mitochondrial DNA to study the population structure of *Apis mellifera* L. Experientia 49: 1016–1021.

Garnery, L., P. Franck, E. Baudry, D. Vautrin, J.-M. Cornuet, and M. Solignac. 1998a. Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*). I. Mitochondrial DNA. Genet. Sel. Evol. 30(suppl. 1): S31–S47.

Garnery, L., P. Franck, E. Baudry, D. Vautrin, J.-M. Cornuet, and M. Solignac. 1998b. Genetic diversity of the west European honey bee (*Apis mellifera mellifera* and *A. m. iberica*). II Microsatellite loci. Genet. Sel. Evol. 30(suppl. 1): S49–S74.

Hall, H. G. 1986. DNA differences found between Africanized and European honeybees. Proc. Natl. Acad. Sci. U.S.A. 83: 4874–4877.

Hall, H. G. 1990. Parental analysis of introgressive hybridization between African and European honeybees using nuclear DNA RFLPs. Genetics 125: 611–621.

Hall, H. G. 1991. Genetic characterization of honey bees through DNA analysis, pp. 45–73. *In* M. Spivak, D.J.C. Fletcher, and M. D. Breed [eds.], The "African" honey bee. Westview, Boulder, CO.

Hall, H. G. 1992a. DNA studies reveal processes involved in the spread of New World African honeybees. Fla. Entomol. 75: 51–59.

Hall, H. G. 1992b. Further characterization of nuclear DNA RFLP markers that distinguish African and European honeybees. Arch. Insect Biochem. Physiol. 19: 163–175.

Hall, H. G. 1992c. Suspected African honeybee colonies in Florida tested for identifying DNA markers. Fla. Entomol. 75: 257–266.

Hall, H. G. 1995. Restriction fragment length polymorphism (RFLP) analyses: distinguishing African and European honey bees, pp. 333–366. *In* J. P. Clapp (ed.), Molecular methods for species identification, methods in molecular biology, vol. 50. Humana, Totawa, NJ.

Hall, H. G. 1998. PCR amplification of a locus with RFLP alleles specific to African honey bees. Biochem. Genet. 36: 351–361.

Hall, H. G. 1999. Genetic and physiological studies of African and European honey bee hybridization: past, present and into the 21st century, pp. 52–59. *In* R. Hoopingarner and L. J. Connor [eds.], Apiculture for the 21st century. Wicwas, Cheshire, CT.

Hall, H. G., and K. Muralidharan. 1989. Evidences from mitochondrial DNA that African honeybees spread as continuous maternal lineages. Nature (Lond.) 339: 211–213.

Hall, H. G., and D. R. Smith. 1991. Distinguishing African and European honeybee matrilines using amplified mitochondrial DNA. Proc. Natl. Acad. Sci. U.S.A. 88: 4548–4552.

Hare, M. P., S. A. Karl, and J. C. Avise. 1996. Anonymous nuclear DNA markers in the American oyster and their implications for the heterozygote deficiency phenomenon in marine bivalves. Mol. Biol. Evol. 13: 334–345.

Her, C., and R. M. Weinshilboum. 1995. Rapid restriction mapping by use of long PCR. Biotechniques 19: 530–532.

McMichael, M., and H. G. Hall. 1996. DNA RFLPs at a highly polymorphic locus distinguish European and African subspecies of the honey bee, *Apis mellifera* L. and suggest geographical origins of New World honey bees. Mol. Ecol. 5: 403–416.

Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89: 583–590.

Nei, M. 1987. Genetic variation within species, pp. 176–207. *In* Molecular and evolutionary genetics. Columbia University Press, New York.

Nelson, E. V. 1967. History of beekeeping in the United Stated. *In* Beekeeping in the United States. USDA ARS Handb. 335.

Nielsen, D. J., P. R. Ebert, R. E. Page, G. J. Hunt, and E. Guzman-Novoa. 2000. Improved polymerase chain reaction-based mitochondrial genotype assay for identification of the Africanized honey bee (Hymenoptera: Apidae). Ann. Entomol. Soc. Am. 93: 1–6.

Novy, R. G., and N. Vorsa. 1996. Evidence for RAPD heteroduplex formation in cranberry: implications for pedigree and genetic-relatedness studies and source of codominant RAPD markers. Theor. Appl. Genet. 92: 840–849.

Oertel, E. 1976. Bicentennial bee. Early records of honeybees in the Eastern United States. Am. Bee J. 116: 70, 71, 114, 128, 156, 157, 214, 215, 260, 261, 290.

Page, R. E., E. H. Erickson, and H. H. Laidlaw. 1982. A close population breeding program for honey bees. Am. Bee J. 122: 350–355.

Pellet, F. C. 1938. History of American beekeeping. Collective Press, Ames, IA.

Pérez, T., J. Albornoz, and A. Dominguez. 1998. An evaluation of RAPD fragment reproducibility and nature. Mol. Ecol. 7: 1347–1357.

Rabouam, C., A. M. Comes, V. Bretagnolle, J.-M. Humbert, G. Periquet, and Y. Bigot. 1999. Features of DNA frag-

ments obtained by random amplified polymorphic DNA (RAPD) assays. Mol. Ecol. 8: 493–503.

Saiki, R. K., D. H. Gelfand, S. Stoeffel, S. Scharf, G. T. Higuchi, K. B. Mullis, and H. A. Erlich. 1988. Primer directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239: 487–491.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular cloning. A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Severson, D. W., R. E. Page, and E. H. Erickson. 1986. Closed population breeding in honey bees: a report on its practical application. Am. Bee J. 126: 93–94.

Sheppard, W. S., T. E. Rinderer, L. Garnery, and H. Shimanuki. 1999. Analysis of Africanized honey bee mitochindrial DNA reveals further diversity of origin. Genet. Mol. Biol. 22: 73–75.

Smith, D. R., and W. M. Brown. 1988. Mitochondrial DNA restriction site polymorphisms in the American and Africanized honey bees (*Apis mellifera*). Experientia 44: 257–260.

Smith, D. R., O. R. Taylor, and W. M. Brown. 1989. Neotropical Africanized honey bees have African mitochondrial DNA. Nature (Lond.) 339: 213–215.

Smith, D. R., M. F. Palopoli, B. R. Taylor, L. Garnery, J.-M. Cornuet, M. Solignac, and W. M. Brown. 1991. Geographical overlap of two mitochondrial genomes in Spanish honeybees (*Apis mellifera iberica*). J. Hered. 82: 96–100.

Smith, D. R., A. Slaymaker, M. Palmer, and O. Kaftanoglu. 1997. Turkish honey bees belong to the east Mediterranean mitochondrial lineage. Apidologie 26: 503–510.

Southern, E. M. 1975. Detection of specific sequences among DNA fragments detected by gel electrophoresis. J. Mol. Biol. 98: 503–517.

Suazo, A., and H. G. Hall. 1999. Modification of the AFLP protocol applied to honey bee (*Apis mellifera* L.) DNA. Biotechniques 26: 704–705,708–709.

Suazo, A., R. McTiernan, and H. G. Hall. 1998. Differences between African and European honey bees (*Apis mellifera*) in random amplified polymorphic DNA (RAPD). J. Hered. 89: 32–36.

Sunnucks, P. 2000. Efficient genetic markers for population biology. Trends Ecol. Evol. 15: 199–203.

Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res. 23: 4407–4414.

Williams, J. G., A. R. Kubelik, K. J. Livak, J. A. Rafalski, and S. V. Tingey. 1990. DNA polymorphisms amplified by arbitrary primers are useful genetic markers. Nucleic Acids Res. 18: 6531–6536.

Received for publication 2 January 2001; accepted 28 August 2001.