Genetic tests for alleles of complementary-sex-determiner to support honeybee breeding programmes

Otto Hyink, Frans Laas, Peter Dearden

Apidologie (2013) 44:306–313

DOI: 10.1007/s13592-012-0181-6

Original article

Genetic tests for alleles of *complementary-sex-determiner* to support honeybee breeding programmes

Otto Hyink[1], Frans Laas[2], Peter K. Dearden[1]

[1]Laboratory for Evolution and Development, Genetics Otago and the National Research Centre for Growth and Development, Biochemistry Department, University of Otago, P.O. Box 56, Dunedin, New Zealand
[2]Betta Bees Research Ltd, P.O. Box 229, South Dunedin, New Zealand

Received 19 April 2012 – Revised 24 October 2012 – Accepted 6 November 2012

Abstract – The honeybee haplodiploid sex determination system depends on genetic variation at the complementary sex-determiner (*csd*) locus. In closed populations of honeybees, especially those undergoing selective breeding, the number of *csd* alleles can drop such that brood viability is affected. Here we present two polymerase chain reaction tests that allow the discrimination of *csd* alleles. Such tests should find utility in bee breeding programmes allowing the tracking and maintenance of *csd* alleles through successive generations.

genetic test / honeybee breeding / sex determination / high-resolution melting

1. INTRODUCTION

The honeybee sex-determination system makes them particularly sensitive to in-breeding. Honeybees employ a haplodiploid sex determining system; females are diploid, males—haploid (Mackensen 1951). In the honeybee, the primary gene regulating this process has been identified as complementary sex-determiner (*csd*) (Beye et al. 2003).

To determine sex, honeybees count the number of alleles of *csd* they have (Woyke 1963; Beye et al. 2003). Bees with two different alleles at *csd* must be diploid and thus female. Bees with one allele are usually haploid and thus male. Problems occur when a bee has two identical alleles at *csd*. These diploids should develop into females, but because they have only one allele of *csd*, male development is triggered (Woyke 1963). Diploid drones are detected by worker bees and killed (Woyke 1963), causing decreased brood viability, an undesirable effect in either breeding programmes or production bees (Woyke 1980, 1981).

Different alleles of *csd* are very common, this being a highly diverse and, presumably, fast evolving locus (Beye et al. 2003; Hasselmann and Beye 2004; Hasselmann et al. 2008b; Wang et al. 2012). This diversity in *csd* alleles means that, in normal populations, brood viability remains high. However, the worldwide reduction in feral bees due to the spread of *Varroa destructor*, and other diseases, implies a loss of genetic diversity in managed stocks that may cause decreased brood viability due to the loss of diversity in sex alleles.

The effect of decreased brood viability due to reduction in the number of *csd* alleles is a particular problem in closed populations of bees, especially those undergoing selective breeding (Woyke 1980, 1981). Due to the costs, mainly in personnel time, of artificial insemina-

Electronic supplementary material The online version of this article (doi:10.1007/s13592-012-0181-6) contains supplementary material, which is available to authorized users.

Corresponding author: P.K. Dearden,
peter.dearden@otago.ac.nz
Manuscript editor: Marina Meixner

tion, most closed populations are small and go through yearly bottlenecks as small numbers of bees are selected for the next generation. In these situations, the loss of *csd* alleles can very quickly lead to decreases in brood viability (Woyke 1980, 1981).

Selective breeding programmes may be better able to manage this severe inbreeding effect if they can monitor, and manage, the numbers of *csd* alleles in their populations. Rare *csd* alleles in a population could be rated as a positive factor in calculating the breeding value of individual queens, thus helping to maintain *csd* diversity in a closed population. To this end, we have developed, and tested, two polymerase chain reaction (PCR) tests to rapidly identify *csd* alleles.

2. METHODS

2.1. Sampling and DNA extraction methods

Samples were collected from the breeding population of Betta Bees Research Ltd (http://www.bettabees.co.nz/). Forty-two colonies were sampled in October 2010 and 65 colonies in October 2011. Six purple eye drone pupae were removed from each hive and placed directly into a 15-mL tube containing 10 mL 95 % ethanol. DNA was extracted from the heads of purple eye pupae. To extract DNA, the heads were removed from pupae using forceps and individually placed into 2-mL tubes containing ceramic beads (MagNAlyser Green Beads, Roche catalogue number 03 358 941 001). Seven hundred fifty microliters RIPA buffer (0.15 M NaCl, 0.05 M Tris pH 7.5, 0.5 % sodium desoxycholate, 1 % NP40, 0.1 % SDS) was added to each sample followed by homogenization in a Roche MagNA Lyser for 60 s. Samples were incubated at 75 °C for 30 min followed by centrifugation at 17,000×*g* for 10 min. Four hundred-microliters of supernatant was collected into a fresh 1.5-mL microfuge tube. Two hundred-microliters of 3-M sodium acetate was added to each sample; samples were mixed and incubated on ice for 20 min. The samples were centrifuged at 17,000×*g* and 400 μL supernatant collected into a fresh 1.5-mL microcentrifuge tube. The DNA was precipitated by adding 1 mL of ethanol and vortexing for 10 s to mix. DNA was pelleted by centrifugation at 17,000×*g* for 10 min, washed with 200 μL 70 % ethanol, air dried for 5 min and then resuspended in 100 μL TE buffer. All samples were stored at −20 °C. These DNA preparations were used for all the PCR tests carried out in this study.

2.2. *csd* isolation and sequencing

Two sets of primers were designed to amplify part of the *csd* gene for sequencing (see Figure 1 for primer locations on the *csd* gene). The primers were modified from Hasselmann et al. (2010) as the melting temperatures of their primers were too low to allow them to be used for direct sequencing. Flanking sequences were added to the 5′ end of primers to enable direct sequencing of PCR products. Primer sequences were M13FConscsdrev 5′-GTAAAACGACGGCCAGTCATCTCATWTTTCATTATTCAAT-3′, AD1genoRfw 5′-TCGCTGTCGGTGAAGACRATATGAAAAATTACACAATGA-3′ and M13FS-40 5′-GTAAAACGACGGCCAGTACTATGTGCATCAATATAAATTC-3′, AD2S-39 5′-TCGCTGTCGGTGATATAATGAAAAAGAAAAATTTTTAGAAG-3′. Sequencing primers used were M13F 5′-GTAAAACGACGGCCAGT-3′, AD1 5′-TCGCTGTCGGTGAAGAC-3′ or AD2 5′-TCGCTGTCGGTGATATAAT-3′.

PCR were performed using Platinum Taq DNA polymerase (Invitrogen) using 0.5 μL template DNA, 5 pmol primer and 2 mM $MgCl_2$ in 25 μL reactions. The optimized PCR protocol used was 95 °C/3 min, 35 cycles of 95 °C/30 s, 48 °C/30 s, 60 °C/15 s, 65 °C/60 s followed by a final step of 65 °C/5 min and 4 °C hold.

To distinguish between the two alleles from a single queen, 10 μL of each PCR product was digested with the restriction endonuclease ApoI and visualized on a 2 % agarose gel. A representative of each restriction profile identified from a hive was sequenced; 1.5 μL of each PCR product and 5 pmol of sequencing primer (total volume 5 μL) were used for sequencing at Genetic Analysis Services (University of Otago). No cleanup of PCR products was required for either the restriction digest or sequencing.

DNA sequences were translated into the CSD amino acid sequence, and the hyper-variable regions of the amino acid sequences were aligned to identify alleles. Those sequences that differed in amino acid sequence through the hyper-variable region were called as different alleles.

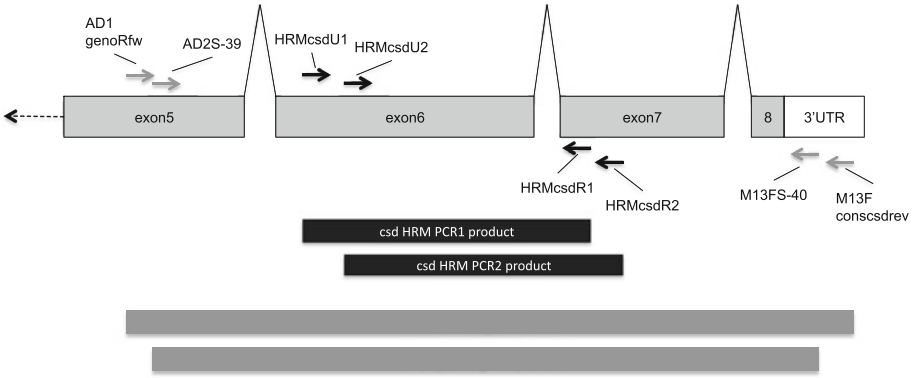

Figure 1. Placement of primers in the *csd* locus. Cartoon of the 3′ end of the *csd* locus, showing placement of the primers and the PCR products they amplify for the PCR assays developed in this study.

2.3. *csd* high-resolution melting test

Two PCR primer sets were designed to amplify the hyper-variable region of the *csd* gene from each sample (see Figure 1 for locations of primers on the *csd* gene and Fig S1 for a sequence alignment of the Betta Bees alleles with primer binding sites highlighted). The primer sets are specific for *csd* and do not amplify *feminizer*, a closely related gene (Hasselmann et al. 2008a). PCR primers were designed using an alignment of 105 alleles (99 alleles from the GenBank database and six novel alleles identified from the Betta Bees Research Ltd population). Under the assay conditions optimized here, we would expect the HRM assay to yield a specific *csd* PCR product for 103 of the 105 alleles (98 %). Primer sequences were: pair 1: HRMcsdU1 5′-TATCGAGAAASAT CGAAAGAACGAT-3′, HRMcsdR1 5′-ATTGA AATCCAAGGTCCCATTGGT-3′; pair 2: HRMcsdu2 5′-AGAACCTAAAATAATTTCATCTTTATCGAA-3′, HRMcsdR2 5′-ATGCCTAAATCTTGGTA TTTGTTCTTG-3′.

HRM assays were performed using the Meltdoctor HRM Master Mix (Applied Biosystems) on a BioRad C1000 thermal cycler with the CFX-96 real-time system attached. PCR mixes consisted of 10 μL HRM Master Mix, 0.1 μL template DNA and 5 pmol of each primer in a total reaction volume of 20 μL.

Final HRM assay conditions were 95 °C/10 min, 50 cycles (95 °C/15 s, 55 °C/30 s, 60 °C/30 s, plate read), 95 °C/60 s, 40 °C/60 s (slowed ramp rate of 2°/s), 60 °C/10 s, melt curve from 60 °C to 85 °C using 0.2 °C increments, 10 s and a plate read at each increment and then a final step of 40 °C/60 s. Results were analysed using the BioRad Precision Melt Analysis Software.

3. RESULTS

The *csd* locus has been previously identified and a large number of alleles identified around the world. We aimed to develop a simple, cost-effective test to identify alleles of *csd* to support bee-breeding programmes.

3.1. Identification of *csd* alleles by digest and sequencing of *csd* alleles

Our first test utilized a classic restriction fragment length polymorphism approach (Narayanan 1991), with amplification of the variable region of *csd* followed by digestion with the ApoI restriction enzyme to produce a diagnostic banding pattern on a gel for each allele. The first primer pair (primers M13FConscsdrev and AD1genoRfw) successfully amplified the *csd* gene from all DNA samples generated, and this DNA was suitable for restriction enzyme digest and/or sequencing without further purification. Digestion of this PCR fragment with Apo1 produces banding patterns that allow discrimination of the two alleles carried by the queen, allowing a representative of each allele to be identified and sequenced. Figure 2 shows a representative set of digests from drones from a

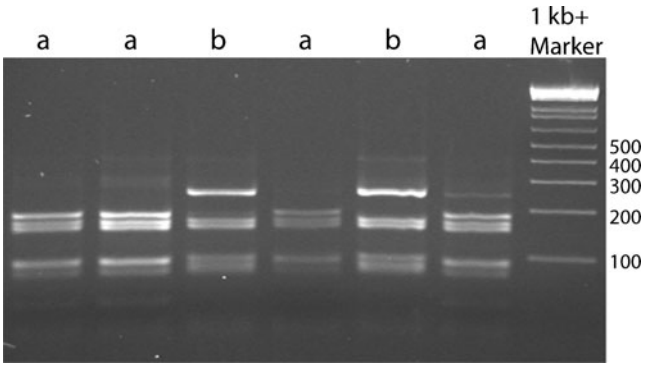

Figure 2. Restriction digest screening to differentiate alleles from one queen. PCR products from the chromosomal DNA of six drone pupae obtained from one hive were digested with the restriction enzyme ApoI and electrophoresed on a 2 % agarose gel. *Lanes 1–6*: The six digested *csd* PCR products, *a* and *b* delineate different alleles, *lane 7* 1 kb + DNA ladder (Life Technologies). The *csd* alleles were amplified by PCR using primers M13FConscsdrev and AD1genoRfw.

single hive showing the discrimination of the two alleles.

We screened the *csd* locus of drone pupae from 42 hives within the Betta Bees Ltd closed population of bees using PCR and restriction enzyme digest in October 2010. Representatives of each allele were then sent for DNA sequencing. Both allele sequences were obtained from 35 of the 42 queens while only one allele sequence was obtained from the remaining seven due to lack of drone samples. Based on amino acid sequence of the hyper-variable region, 16 different alleles and their abundance within the population were identified (Table I). Comparison of the 16 alleles sequenced from the Betta Bees breeding population to *csd* sequences in the GenBank database showed that six of these 16 alleles are new alleles that have not previously been sequenced (GenBank accessions JX915851–JX915866). The remaining ten alleles were all identical to sequences found in GenBank.

3.2. HRM test of *csd* alleles

Because of the time and cost associated with sequencing DNA fragments, we designed a second test to discriminate *csd* alleles based on high-resolution melting (HRM) and used this to determine the *csd* alleles of queens in the Betta Bees breeding population in October 2011. HRM uses differences in melting curves of DNA fragments to discriminate genetic variants (Vandersteen et al. 2007; Vossen et al. 2009; Wittwer 2009). Primers were designed as close to either side of the hyper-variable region as possible as HRM is most effective with small DNA fragments. To ensure good discrimination of as many alleles as possible, we designed two sets of HRM primers and used them to amplify *csd* alleles. HRM was then carried out on each amplified fragment.

The HRM assay was optimized using existing DNA preparations for the 16 previously identified alleles from the Betta Bees population. All 16 alleles could be clearly distinguished from each other using HRM (Table I). Based on DNA sequence, we predicted that primer pair one (HRMcsdU1 and HRMcsdR1) would amplify all Betta Bees alleles while primer pair two (HRMcsdU2 and HRMcsdR2) would not amplify alleles BB3 and BB10. Primer pair one did amplify all 16 alleles and primer pair two only failed to amplify allele BB10. This HRM test and DNA preparations containing the 16 Betta Bees alleles were then given to a diagnostic laboratory to test (Dnature Ltd, www.dnature.co.nz), giving identical results. When the identity of the hives from which bees were collected was blinded to the

Table I. Frequency of the 16 *csd* alleles identified in the Betta Bees honeybee population.

Allele	Amino acid sequence of hyper-variable region	Freq in 2010 population[b]	Freq in 2011 population[c]
BB1	IISSLSNNYKYSNYNNYNNYNKKLYYKNYI	0.05	0.07
BB2	IISSLSNNYKYSNYNNYNNYNNNNYNNYNKKLYYKNYI	0.03	0.05
BB3[a]	ITSSLSSSCNYSNNYNNYYNNNKKLYYNII	0.12	0.08
BB4[a]	IISSLSNNYKYSNYNNYNNNNYNNNNYNNNSKKLYYNII	0.05	0.05
BB5	IISSLSNNYNYNNYNNNYKPLYYNINYI	0.05	0.13
BB6[a]	IISSLSNKTIHNNNNYKYNYNNNNYKKLQYYNINYI	0.01	0.01
BB7	IISSLSNNYNYNNNKYNYNNNNYKQLCYNINYI	0.11	0.06
BB8	IISSLSNNYNYSNNYNNNNYYNNNNYKKLYYNINYI	0.04	0.02
BB9	IISSLSNNYNYSNYNNNNYKQLCYNINYI	0.05	0.15
BB10	IISSNNYNYKNYNNNYNNYKQLCYNINYI	0.05	0.03
BB11[a]	ITSSLSNNYNSNSYNNYNNNYYNNKKLQYYNINYI	0.03	0.00
BB12	IISSLSNKTIHNNNNYKPYYNINYI	0.05	0.11
BB13	IISSLSNKTIHNNNNYKYNYNNNNNNYKNYNNYKKLYYNINYI	0.11	0.08
BB14[a]	IISSLSNKTIHNNNNYKYNYNNNYNNNCKKLYYNII	0.09	0.12
BB15[a]	IISSLSNKTIHNNNNYKYNYNNNYNNNSKKLYYNINYI	0.03	0.00
BB16	IISSLSNKTIHNNNNYNNNNYNNYNNNYNNYKKLYYNII	0.11	0.04

[a] Novel CSD sequences identified as part of this study

[b] The number of times this allele was sequenced divided by the total number of alleles sequenced

[c] The number of times the melt profile for this allele was identified divided by the total number of alleles screened for in the HRM test

operator, the HRM test allowed identification of those alleles in 100 % of cases. Screening of the 2011 Betta Bees breeding population identified 14 of the 16 alleles present in 2010 (Table I).

Because of the AT-rich nature of the variable region, extension temperatures of 60 °C were required for efficient amplification. Figure 3 shows representative melting curves, as well as alignments of alleles discriminated by this technique. Figure 3a shows the melt curve for the PCR product produced with primer pair 1 for alleles 4 and 14. The melting of the PCR product starts just over 61 °C due to the AT-rich region. All the PCR products were completed melted at a temperature of around 75 °C. The melt curve was analysed in the temperature range of 60 to 85 °C to ensure that the complete dissociation pattern could be observed. Figure 3b shows an alignment of the two alleles discriminated by HRM in Figure 3a. Figure 3c shows the melt curve for the PCR product produced with primer pair 2 for alleles 4 and 14, and Figure 3d shows the alignment of these two alleles.

4. DISCUSSION

We present here two methods for the determination of, and screening for, *csd* alleles in honeybee populations. These two techniques are ideal for determining *csd* alleles in closed populations of bees and have been used to inform breeding strategies in the Betta Bees Research Ltd closed population of bees (www.bettabees.co.nz).

These tests use DNA sequence, or melting characteristics to discriminate alleles. As it is not yet clear how much sequence difference is required between *csd* alleles for them to be recognised as biologically different alleles and thus trigger female development, it is important to

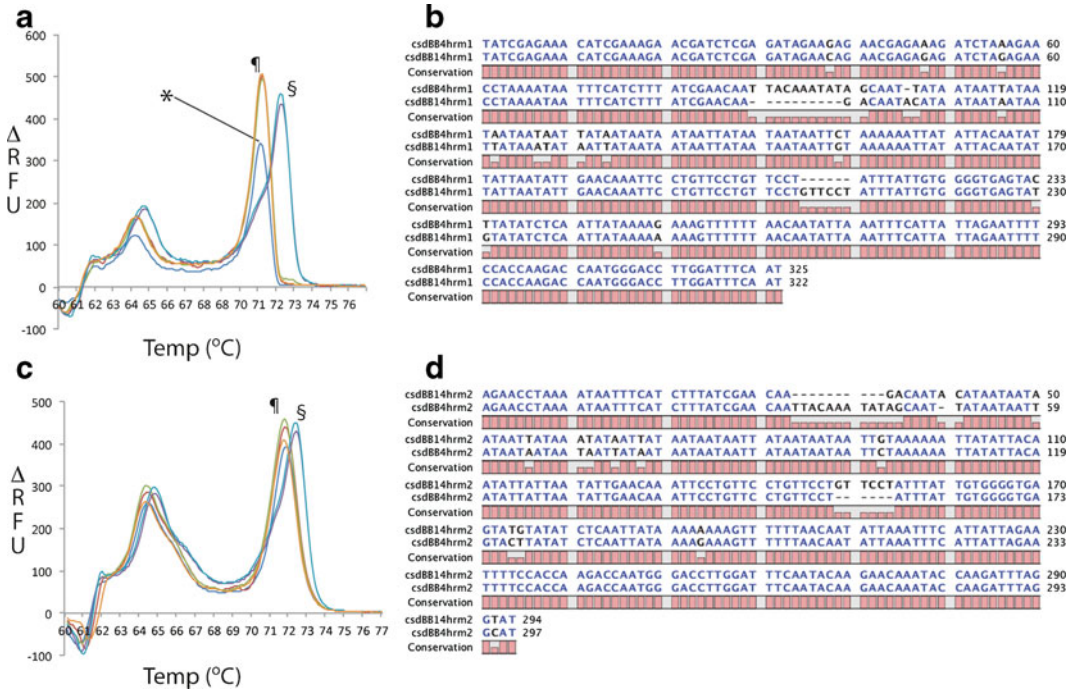

Figure 3. High-resolution melt analysis of *csd* alleles. **a** HRM analysis of PCR products, generated using primers HRMcsdU1 and HRMcsdR1, from the chromosomal DNA of six individual drone pupae collected from one hive. The melt curves shown are for alleles BB4 (*pilcrow sign*) and BB14 (*section sign*). The *asterisk* indicates that despite the total fluorescence being lower (weaker PCR product), the melt curve generated by HRM can still be used to identify the allele. **b** Alignment of the sequence of the alleles discriminated by HRM in **a**. **c** HRM analysis of the PCR products, generated using primers HRMcsdU2 and HRMcsdR2, from the same chromosomal DNA preparations used in **a**. The melt curves shown are for alleles BB4 (*pilcrow sign*) and BB14 (*section sign*). **d** Alignment of the sequence of the alleles discriminated by melting in **c**.

recognize that these PCR tests may over-estimate allele diversity. The HRM test will also discriminate alleles that differ by a silent mutation in the DNA but are identical at the amino acid level. For a large study, we would recommend sequencing a representative of each unique HRM profile identified to obtain the most accurate allele information. Determining the sequence will also allow bee breeders to identify phylogenetically distant alleles of *csd* to ensure that these alleles act functionally as different alleles in their populations. The size of the PCR fragments produced in this study would also make them amenable to determining their sequence using next-generation sequencing technologies. This would be particularly useful and cost effective in the case that determination of the range of *csd* alleles from a population or populations was required.

In these experiments, we used six drone pupae from each hive as this gives a 0.96 probability of amplifying both alleles (assuming random prevalence of the two alleles from a queen in the drone population the probability of finding both alleles is equal to $(n^2-2)/n^2$ where n is the number of drones sampled; six drones gives a probability of 31/32 or 0.96 probability of amplifying the two alleles carried by the queen). Drone pupae were used to avoid problems with differential amplification of the different alleles from one sample. In breeding programmes, the production and screening of drone bees is relatively simple. In situations where only female bees are available for screening, it will be necessary to clone each allele for sequencing. HRM assays from female bees that identify both alleles are possible, but there is a strong possibility

that amplification of the two alleles will not be equal, which may affect identification of alleles.

The tests we have developed are designed to help inform bee-breeding programmes about *csd* allele diversity in their populations and help make breeding decisions that minimize *csd* allele loss. We are aware, therefore, that the key characteristics of such a test are that they must be accurate, fast and cost-effective.

The digest and sequencing method is the most accurate, as it gives the definitive sequence of the variable region of the *csd* gene and, if the sequencing works, is not affected by variability in amplification or DNA quality. While our HRM assay does not give the sequence of the allele, we have found it very effective in discriminating different alleles (especially when both PCRs are carried out) and identifying unknown ones. Most importantly perhaps, it provides a stable melting curve even when the DNA amplification is poor (asterisk in Figure 3a), allowing identification of alleles even with low-quality/quantity DNA extractions.

The HRM test is significantly faster than the digest and sequencing test. Total time for results after DNA extraction is around 4 h. For the digest and sequencing test, PCR takes around 2 h, setting up digests and running them on a gel takes another 2 h. Sequencing the alleles is the critical time constraint and depends on access to sequencing equipment.

The total costs for the sequencing approach (not including labour) come to around US $20.50/queen (six drone samples, DNA extraction, PCR, enzyme digest and sequencing). The HRM assay costs US $11.70/queen (six DNA extractions, 12 PCR/HRM reactions). It should be noted that for the HRM technique, an HRM capable real-time PCR machine is required.

In conclusion, the application of these relatively simple protocols will enable both the typing and tracking of *csd* alleles in bee populations, and we hope will prove useful to bee breeders wishing to incorporate *csd* allele frequencies into their selection matrices to minimise the drop in brood viability caused by inbreeding.

ACKNOWLEDGMENTS

We would like to thank the Tertiary Education Commission's Partnering for Innovation Scheme and the Agricultural and Marketing Research and Development Trust for funding this work. We would like to thank the National Research Centre for Growth and Development for funding the purchase of real-time PCR equipment and dnature (http://dnature.co.nz), for testing this protocol. We would also like to thank Betta Bees Research Ltd for access to their closed population of bees and the National Beekeepers Association of New Zealand for their support.

Tests génétiques d'identification des allèles du gène *csd* pour soutenir les programmes d'élevage d'abeilles

Test génétique / élevage d'abeilles / détermination du sexe / courbe de fusion à haute résolution

Genetische Tests auf Allele des *complementary-sex determiner* zur Unterstützung von Zuchtprogrammen für die Honigbiene

Genetischer Test / Zucht der Honigbiene / Geschlechtsbestimmung / hochauflösende Schmelzpunktbestimmung

REFERENCES

Beye, M., Hasselmann, M., Fondrk, M.K., Page, R.E., Omholt, S.W. (2003) The gene *csd* is the primary signal for sexual development in the honeybee and encodes an SR-type protein. Cell **114**(4), 419–429

Hasselmann, M., Beye, M. (2004) Signatures of selection among sex-determining alleles of the honey bee. Proc. Natl. Acad. Sci. U. S. A. **101**(14), 4888–4893

Hasselmann, M., Gempe, T., Schiott, M., Nunes-Silva, C.G., Otte, M., Beye, M. (2008a) Evidence for the evolutionary nascence of a novel sex determination pathway in honeybees. Nature **454**(7203), 519–522

Hasselmann, M., Vekemans, X., Pflugfelder, J., Koeniger, N., Koeniger, G., Tingek, S., Beye, M. (2008b) Evidence for convergent nucleotide evolution and high allelic turnover rates at the *complementary sex determiner* gene of Western and Asian honeybees. Mol. Biol. Evol. **25**(4), 696–708

Hasselmann, M., Lechner, S., Schulte, C., Beye, M. (2010) Origin of a function by tandem gene duplication limits the evolutionary capability of its sister copy. Proc. Natl. Acad. Sci. U. S. A. **107**(30), 13378–13383

Mackensen, O. (1951) Viability and sex determination in the honey bee (*Apis mellifera* L.). Genetics **36**(5), 500–509

Narayanan, S. (1991) Applications of restriction fragment length polymorphism. Ann. Clin. Lab. Sci. **21** (4), 291–296

Vandersteen, J.G., Bayrak-Toydemir, P., Palais, R.A., Wittwer, C.T. (2007) Identifying common genetic variants by high-resolution melting. Clin. Chem. **53** (7), 1191–1198

Vossen, R.H., Aten, E., Roos, A., den Dunnen, J.T. (2009) High-resolution melting analysis (HRMA): more than just sequence variant screening. Hum. Mutat. **30**(6), 860–866

Wang, Z., Liu, Z., Wu, X., Yan, W., Zeng, Z. (2012) Polymorphism analysis of csd gene in six *Apis mellifera* subspecies. Mol. Biol. Rep. **39**(3), 3067–3071

Wittwer, C.T. (2009) High-resolution DNA melting analysis: advancements and limitations. Hum. Mutat. **30** (6), 857–859

Woyke, J. (1963) Drone larvae from fertilised eggs of the honeybee. J. Apic. Res. **2**(1), 19–24

Woyke, J. (1980) Effect of sex allele homo-heterozygosity on honeybee colony population and on their honey production. 1. Favourable development conditions and unrestricted queens. J. Apic. Res. **19**(1), 51–63

Woyke, J. (1981) Effect of sex allele homo-heterozygosity on honeybee colony population and on their honey production. 2. Unfavourable development conditions and restricted queens. J. Apic. Res. **20**(3), 148–155