Project title

Developing a High-Throughput System of Quantifying and Contextualizing Genetic Diversity in Beekeeping Operations

PI information

David R. Tarpy
Professor and Extension Apiculturist
Department of Entomology & Plant Pathology
Campus Box 7613
North Carolina State University
Raleigh, NC 27695-7613

TEL: (919) 515-1660
FAX: (919) 515-7746
LAB: (919) 513-7702
WEB: http://ncsuapiculture.net
EMAIL: david_tarpy@ncsu.edu

Objectives

1) Establish a high-throughput system of quantifying operation-level genetic diversity and allelic diversity at the sex locus
2) Invite up to 25 beekeeping operations to analyze their colonies each to establish a baseline level of genetic diversity

Methods

Objective 1: We aim to create a new system of determining if a given beekeeping operation has sufficient genetic diversity, as we know from previous research that genetic bottlenecks can significantly reduce productivity and facilitate the spread of disease. We will do so by employing the established methods within the Queen & Disease Clinic for genotyping of colonies but modifying the sampling protocol to quantify genetic diversity at the population level rather than at the colony level. We will sample ~25 workers from each of 16 different colonies within an operation by collecting them into standard 50 ml sampling tubes. We will extract the genomic DNA from 12 individuals from each colony then subject the DNA to a new PCR analysis at 16-24 multiplexed microsatellite loci to generate standard population genetic metrics (e.g., average heterozygosity, allelic number and diversity, F-statistics to estimate inbreeding coefficients, etc…).

In addition to quantifying the population genetic diversity among different operations at these neutral genetic markers, we also aim to develop new markers that are association with the critical *csd* gene or "sex locus." Since homozygosity is effectively lethal, the greater the number of sex alleles within an operation the better. We wish to employ newly developed markers, tightly linked to the *csd* locus, to putatively quantify the number of sex alleles within a given breeding population. This will then enable us to provide very powerful information, particularly to queen producers, about their needs to outcross their stock or import new genetics into their operation to maintain optimal genetic health.

Objective 2: We will solicit shipments of bees from all beekeepers who wish to utilize this new function of the Queen & Disease Clinic on a first-come first-served basis. We will advertise through various venues, particularly through the Bee Informed Partnership and their participating commercial beekeeping operations. To simplify the logistics, we will schedule one shipment of 16 colonies from a single operation per week (and thus accept a maximum of 25 shipments throughout the season). This will enable us to determine the distribution in genetic diversity among different operations, which will be enable us to contextualize any subsequent analysis (i.e., higher or lower than average). Once this baseline is established, we will then offer this analysis as a fee-based service for any beekeeper in the future, thus making this a self-funded sustainable initiative for the industry.

Progress and preliminary results

Objective 1: We developed three *csd* -linked microsatellites that have sizes ~300bp, and they can be added directly to the microsatellite multiplex primers that are used for non-linked loci. This will enable us to generate data on genetic diversity at the sex locus for minimal additional cost when we run any sample as we do for a mating number genotyping analysis. The *csd* microsatellites we have right now are ~23-30kb away from the *csd* coding

sequence, so while they are not directly amplifying the sex allele, they are adequate proxies for estimating the diversity at the sex locus. Over time, if we have enough genetic data, we may be able to identify indels in the *csd* coding introns and exons, then we should be able to generate molecular markers that will be able to detect the actual mutations in the *csd*.

Locus	Forward primer sequence	Reverse primer sequence
CSD03	GTAATCGCGCCGAAGATGAAC	CGCTATACCAGTTCCCGCT
CSD_2901_AGC7	GTTACCGACAGCGAGAACGAGACCA	CGATCACGTCCATTTCCTTCAACGC
CSD04-4	ACAATATCCACAGTTTTATAGCAAACACA	GCAGTAATGAATTTTAGTGCAGAAAGC

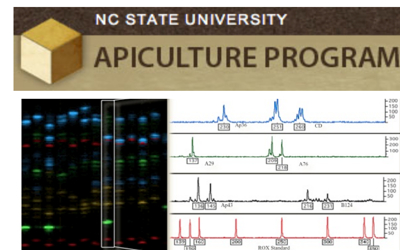
We have also developed a generic report for future screenings and analyses (**Figure 1)**. We have a very similar workflow and process for other bioassays available in the Queen & Disease Clinic (e.g., insemination analysis). Our procedure is to secure live (or frozen) samples of 8 workers from 12 different colonies (n=96) in a beekeeping operation so that we can extract the DNA from each individual. We then subject the DNA to two standard multiplexed PCR reactions using 12 microsatellite primer sets, 3 of which are linked to the *csd* locus (see above). Once scored using GENEMAPPER® software, the data are automatically analyzed and statistically compared to the baseline population so that the report is generated instantaneously. This report is then shared with the beekeeper, and a follow-up phone call is offered to help walk through the information in the report and place the results into a larger context.

Objective 2: Unfortunately, the clinic technician who had developed the new microsatellite primers is no longer in our lab, and he was unable to apply the new technique to the broader population in order to establish a baseline to which to compare any future samplings. However, we are in the process of doing so this upcoming field season, supported by other funds, so that we will be able to roll out this new bioassay to beekeepers in 2017.

OPERATIONAL GENETIC DIVERSITY

Report date: February 10, 2017

Source	Received date	No. colonies
SS	02/01/17	12



Diversity at the sex locus (linked to csd)

A+	CSD03	CSD_2901	CSD04-4
Average	3.7	5.0	4.8
Maximum	5	8	10
Minimum	1	1	1
Percentile	88.0%	91.9%	75.8%

Autosomal loci (unlinked)

C	A24	A76	A88	A113	Ap43	Ap81	ApJC2	B124	Control
Average	Not measured	Not measured	Not measured	Not measured	Not measured	Not measured	Not measured	Not measured	Not measured
Maximum	Not measured	Not measured	Not measured	Not measured	Not measured	Not measured	Not measured	Not measured	Not measured
Minimum	NA	NA	Not measured	Not measured	Not measured	Not measured	Not measured	Not measured	Not measured
Percentile	NA	NA	Not measured	Not measured	Not measured	Not measured	Not measured	Not measured	Not measured

Overall composite

B-	FST	FIT	FIS	Allelic diversity	Unique alleles
Average	0.73	0.55	0.13	6.22	3.12
Maximum	0.00	0.00	0.00	1.00	0.00
Minimum	1.00	1.00	1.00	12.21	4.00
% GLOBAL	88%	33%	24%	76%	97%

OVERALL ASSESSMENT

Significantly higher than average

No further action recommended

Comments and recommendations

Genetic diversity at the sex locus (*csd* gene) is critical to prevent inbreeding, as homozygosity leads to diploid drones, inviable offspring, and shotbrood. These genetic markers are closely linked to this locus and thus are proxies for genetic diversity at this gene, but they are not direct measures of the number of sex alleles. There seems to be very good allelic variation at the sex locus in this population, and as such it is likely to be very resiliant to inbreeding depression. Other measures of genetic diversity not associated with the sex locus are also important for understanding the overall genetic diversity in the population. The current population demonstrates moderate genetic diversity compared to other populations. Overall, the population structure suggests low levels of inbreeding and sufficient genetic diversity from the global honey bee breeding population, and thus an increase in outcrossed genetic stock is not warranted at this time.

Figure 1. An example report for our new analysis. Based on our standard 'queen insemination' analysis through the Queen & Disease Clinic, this report automatically generates pertinent information of a given sample (population) compared to the global population (baseline) for genetic diversity at the sex locus, non-linked loci, and overall. The report also provides a more detailed explanation and recommendations based on the results.

Additional side projects

The choice over which larvae to select for emergency queen rearing is complicated by the extreme multiple mating of honey bee queens. Honey bee queens are reported to have an average mating number of ~12 drones, thereby giving rise to colonies consisting of many patrilineal subfamilies. The selection of larvae during emergency queen rearing pits the self-interest of individual worker subfamilies (to have a full "super-sister" as the next queen) against that of the overall colony (to have the highest quality queen, regardless of subfamily). This conflict presents a rare opportunity in honey bees, where nepotistic inclusive-fitness for workers could overcome the group-level selection that has given rise to most eusocial traits.

Research to date suggests that this is not the case, however. Instead, it appears that workers preferentially select larvae from particular "royal" subfamilies that are rare in the overall worker population. All of these previous studies, however, include one or more empirical shortcomings—such as small number of colonies, small sample sizes per colony, or too few molecular markers to fully differentiate subfamilies—each of which can make it difficult to completely determine what pattern is present in emergency queen rearing and what factors might be responsible. We have carried out an experiment—using the genetic markers used as part of this award—to address these shortcomings and more fully investigate the purported "royal patriline" effect in emergency queen rearing. Many of these "royal" subfamilies are rare enough to be undetected in typical patriline analyses of

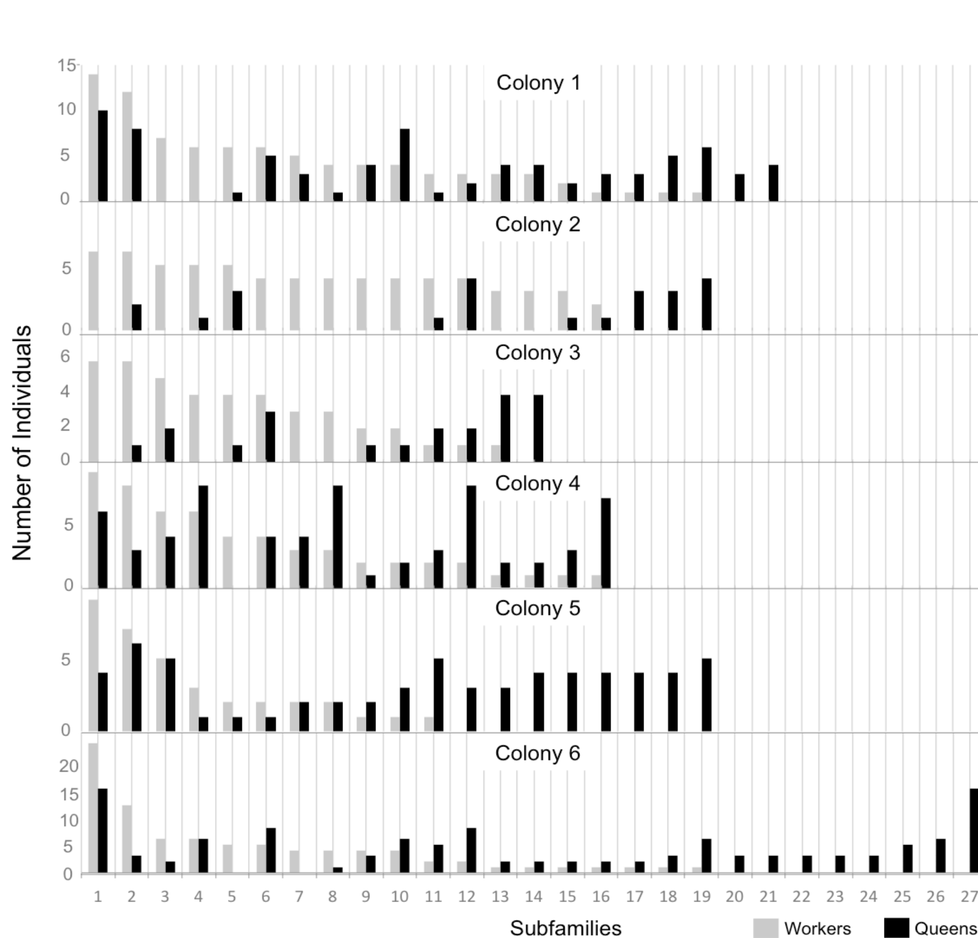


Figure 2. Subfamily distribution of workers and emergency-reared queens in six honey bee colonies using subfamilies comprising at least three observed individuals. *Grey bars* are the counts of workers in each colony. *Black bars* are the counts of emergency queens from the same subfamilies. Subfamily counts per colony range from 14-27, with many queens deriving from patrilines rare or absent in workers sampled.

honey bee colonies using only workers, indicating that queen promiscuity is even higher than previously recognized (**Figure 2**).

For the beekeeping community and apiculture industry, these findings may provide insight into which individuals make the best queens. Current queen-rearing practice is to randomly select young worker larvae and graft them into queen cups to be reared as queens. If certain colony genotypes are more likely to be raised by workers when given the choice, then absolving the workers of that choice may have implications on the types of queens that beekeepers are raising. This begs the question as to how queens from “royal” patrilines may differ from those raised following standard practices, and ongoing research will aim to answer this question.