9-14-06 TABASHNIK'S ANSWERS TO MATTEN'S QUESTIONS RE: PCR
FROM SHARLENE 9-8-06
Bruce --

Here are some questions of clarifications (I send most of these to Tish yesterday, who by the way, has been very helpful.)

1. Is the Molecular analysis method in the submission to EPA, the one published in Tabashnik et al. 2006? If so, I need an e-mail back to me identifying that this is indeed the method used.

Yes. The molecular analysis method in the submission to EPA is the one published in Tabashnik et al. 2006. Key portions of the paper describing the method are provided below. Please note that initial tests use only a small portion of each field-sampled individual (Tabashnik et al. 2005b), so that re-testing of individuals is possible if desired.

INTRODUCTION
"In laboratory-selected strains of pink bollworm and at least two other major lepidopteran pests of cotton, mutations in a cadherin gene are tightly linked with recessive resistance to Cry1Ac (Gahan et al. 2001, Morin et al. 2003, Xu et al. 2005). In several laboratory-selected strains of pink bollworm, three mutant alleles (*r1*, *r2*, and *r3*) of a cadherin gene (*BtR)* are associated with resistance to Cry1Ac and survival on Bt cotton (Morin et al. 2003, 2004; Tabashnik et al. 2004, 2005b). Each *r* allele has a deletion predicted to eliminate at least eight amino acids upstream of the putative Cry1Ac-binding region of cadherin protein (Morin et al. 2003). We previously developed a PCR-based method for detecting the *r1*, *r2* and *r3* alleles in pink bollworm (Morin et al. 2004). We isolated, cloned and sequenced the genomic region spanning the mutation in each *r* allele and designed allele-specific PCR primers for each region. The method can detect any of the three *r* alleles in a single heterozygote (*r1s*, *r2s*, or *r3s*) pooled with DNA from the equivalent of 19 susceptible (*ss*) individuals (Morin et al. 2004)."

METHODS: "DNA Preparation and PCR. Insects collected from bolls and traps were stored in ethanol at –20^{o}C. DNA was extracted using DNAzol (Tabashnik et al. 2005b) and PCR was done as described by Morin et al. (2004). The maximum number of individuals tested per pool was 5 for samples from 2001-2003 and 11 for samples from 2004-2005."

2. The late season sampling method developed by Dennehy and Tabashnik -- I don't have the specific protocol. Please send it or perhaps it is in the Tabashnik et al. 2006 manuscript. I see a couple of sentences describing the plan, is this it? Similar to Tabashnik et al. (2006), details coming soon.

3. What is the method for estimating false negatives? false positives? What is the likelihood of non-detection? Please clarify.

The methods for estimating false negatives, false positives, and the likelihood of non-detection are detailed in Morin et al. (2004) and Tabashnik et al. (2006), as well as below:

A. False negatives.
False negatives are possible from three causes: i) The PCR reaction is not working properly, ii) The cadherin DNA of field-sampled insects is not amplified, iii) The PCR is working and cadherin DNA is amplified from field samples, yet *r* alleles are present and are not detected.

i) To determine if the PCR reaction is working properly, we use known positive controls in every set of samples tested. This is a standard method. Known positive controls are samples of DNA from our laboratory-reared strains that contain *r* alleles, which are known to the person running the PCR reaction. For example, every test of field-sampled insects for the presence of an *r1* allele includes a gel lane in which DNA from one or more laboratory-reared individuals with the *r1* allele is run simultaneously with the field samples.

If the known sample of *r1* DNA does not yield a positive result for *r1*, the test of the field sample is not valid and must be repeated. In this case, PCR reaction conditions are corrected until the known controls yield positive results with the simultaneously tested field samples. Such corrections usually involve systematic replacement of reagents (primers, Taq, etc.) to ensure all are working properly. Because only tests yielding positive results for known positive controls are included in our analysis of the data, this source of false positives has an effective rate of 0% in the data analysis.

ii) To determine if the cadherin DNA of field-sampled insects is amplified, we test for amplification of a conserved region of the cadherin gene that occurs in all known susceptible and resistant alleles (Morin et al. 2004). As described in Tabashnik et al. (2006), "We checked all pools using this approach and >99% tested positive. Because as few as one amplifiable allele from a pool of insects could yield a positive result for this control reaction, we also tested a subset of insects individually from each of the 59 field samples. Of the 835 individuals tested, 98.6% were positive."

The 98.6% amplification rate of the conserved region of the cadherin gene indicates that DNA was not amplified from 1.4% of field-sampled insects. We take this into account in estimating the likelihood of non-detection by adjusting the sample size accordingly. For example, if 1000 alleles are screened from 500 individuals and the amplification rate of the conserved region is 98.6%, the corrected sample size is 986 (see C below).

iii) As described in Tabashnik et al. (2006): "In addition to standard positive controls for each of the three *r* alleles in all tests, we included "blind" positive controls as follows: Two researchers analyzed each field sample. One researcher prepared DNA and added individuals with one or two *r* alleles from laboratory-selected resistant strains in zero to three (usually one) of the pools tested from each field site. The other researcher performed PCR and did not know which, if any, of the pools contained these blind positive controls. The detection rate for blind positive controls was 97% (97/100)." The rate of false negatives (3%) caused by failure to detect *r* alleles present in pools is incorporated in the estimate of the likelihood of non-detection, as described below (C).

B. False positives. To detect false positives, we use a standard technique. All tests of field samples include blanks, which are gel lanes containing all of the PCR reagents, but no DNA. If a blank yields a positive result, this indicates contamination (i.e., a false positive). In this case, PCR reaction conditions are corrected and the field samples are retested. Results are included in the data analysis only if the blanks do not yield positive results.

Of the 5,571 field-sampled insects tested in Tabashnik et al. (2006), none yielded positive results. Thus, the problem of false positives is minimal to nil. When a pool of field-sampled insects yields a positive result for an *r* allele (e.g., *r2*), each individual in the pool will be tested separately to verify the positive result and to more precisely estimate the frequency of resistance in the pool.

C. Non-detection. As described in Tabashnik et al. (2006), the likelihood of non-detection is estimated as follows:

"The probability of detecting no *r* alleles in a sample of *N* individuals was calculated as $(1-[F \text{ X } D])^{2N \text{ X } A}$, where *F* is the frequency of resistance alleles, *D* is the probability of detecting an *r* allele present in screened individuals (0.97, based on the data from blind controls), 2*N* is the number of alleles screened, and A is the probability of amplifiable cadherin DNA occurring in field-sample insects (estimated as 0.986, based on the proportion of positive results for amplification of a conserved sequence in 835 insects tested individually). We assumed that the probability of an *r* allele occurring was an independent event at each cadherin allele screened. For example, with an *r* allele frequency of 0.001, the probability of detecting no *r* alleles in the sample of 5,571 individuals (11,142 alleles) is $0.000023 = (1- [0.001 \text{ X } 0.97])^{11,142 \text{ X } 0.986}$. Analogously, with an *r* allele frequency of 0.0003, the probability of detecting no *r* alleles in the sample of 5,571 individuals is $0.041 = (1 – [0.0003 \text{ X } 0.97])^{11,142 \text{ X } 0.986}$."

The goal in 2006 is to screen 500 field-sampled individuals with PCR (i.e, *N*=500). Assuming no *r* alleles are detected and values for *D* and *A* similar to those above, the probability (P) of non-detection is estimated as:

i) for true *r* allele frequency of 0.00316 (frequency of *rr* = 0.00001),
$P = (1- [0.00316 \text{ X } 0.97])^{1000 \text{ X } 0.986} = 0.048$

ii) for true *r* allele frequency of 0.01 (frequency of *rr* = 0.0001),
$P = (1- [0.01 \text{ X } 0.97])^{1000 \text{ X } 0.986} = 0.000067$

iii) for true *r* allele frequency of 0.001 (frequency of *rr* = 0.000001),
$P = (1- [0.001 \text{ X } 0.97])^{1000 \text{ X } 0.986} = 0.38$

Below please find additional discussion of the potential for non-detection from Tabashnik et al. (2006):
"It is important to consider potential underestimation of *r* allele frequency based on DNA screening. DNA screening based solely on males caught in pheromone traps could cause underestimation if the probability of capture in traps was lower for *rr* or *rs* males than for *ss* males. However, tests conducted in large cages (64 m^3) in the field refuted this hypothesis for

pink bollworm (Carrière et al. 2006). Furthermore, DNA screening of pink bollworm from bolls, which was independent of males caught in traps, also detected no *r* alleles (n = 1,344; Table 1).

If alleles other than cadherin mutants *r1*, *r2*, and *r3* confer pink bollworm resistance to Bt cotton, the results of our DNA screening could underestimate the frequency of resistance. For example, resistance to Cry1Ac in some strains of diamondback moth is not linked with cadherin (Baxter et al. 2005). However, in four laboratory-selected Cry1Ac-resistant strains of pink bollworm tested so far, all resistant individuals screened have two copies of the known *r* alleles (i.e., *r1r1*, *r2r2*, *r3r3*, *r1r2*, *r1r3* or *r2r3)* and no other resistant alleles have been detected (Morin et al. 2003, Tabashnik et al. 2004, 2005b). Although the presence of additional resistance alleles at the cadherin locus or other loci cannot be excluded, such alleles appear to be more rare than the three known resistance alleles."