ASSESSING CLONAL DIVERSITY OF *PLASMODIUM MEXICANUM* USING MICROSATELLITE MARKERS: VERIFICATION OF TECHNIQUES

Anne M. Vardo
November 8, 2006

The calibration and verification of new molecular techniques that are used to explore a novel system is a critical initial goal of any new research effort. For a study on the clonal diversity of the malaria parasite *Plasmodium mexicanum* in lizard hosts, new molecular markers- microsatellites -were employed. Prior to this study, only a few investigators have used these microsatellites in studies of genetic diversity for the *Plasmodium* of humans, so 'ready to go' protocols that could cross-over for use with a lizard malaria parasite have not yet been perfected. In fact, none of the published microsatellite primers for *Plasmodium falciparum* were successful in amplifying the DNA of a lizard malaria parasite, so new markers had to be developed. After these markers were developed, they had to be tested to insure that results were repeatable and that all clones present in an infection could be detected.

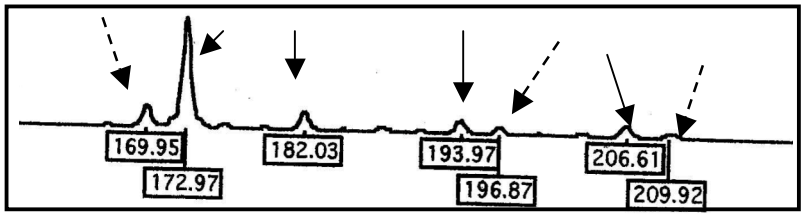
The following section describes the different tests used to verify the method's ability to detect clones and to repeat results for one to three of the most polymorphic loci for this species.

Test 1: Detecting Clones

1.1 Detecting mixed clone infections: Trial 1

For an infection of a malaria parasite, it is often common for a single infection to harbor more than one genetically distinct form of the parasite (clones). Detecting these clones, and assessing the number of clones within an infection is critical for any study on the biology of clonal diversity. That is, testing hypotheses requires an accurate estimate of the number of parasite clones present in an infection. Being able to detect a multi-clonal infection, and to determine the maximum number of clones that could be distinguished, was the first goal. One well-performing marker, Pmx306 (Schall & Vardo, 2006) was used for the first test.
DNA was extracted via the Qiagen DNeasy kit, from 6 different lizards which preliminary study found harbored 1 – 2 clones (alleles of the Pmx 306 locus). Combining 5 υl of each DNA extraction in a 1.5 ml tube, the mix of DNA extraction products contained 7 different genotypes. For the PCR, 2 υl of the extraction was used as the template DNA. The product was diluted 1:20 H2O for genotyping on an ABI sequencing instrument. The data were examined with Genotyper 2.0 software (ABI). The resulting pherogram showed 6 clear peaks, indicating different clones. These were scored based on knowledge of the actual genotypes in the mixture of extracted DNA. However, some of the peaks were only 3 bp difference in size, so it would have been difficult to determine if these were true alleles or stutter peaks without prior knowledge of the clonal makeup of the sample. Therefore, if an observer had no prior knowledge of the true clonal diversity within the sample, it would have been scored as having "4 or more" clones. Note that the larger fragment sizes in general produced small peaks. This is a normal result for the ABI instrument so infections containing many clones may well have some alleles missed if they were longer microsatellite repeats.

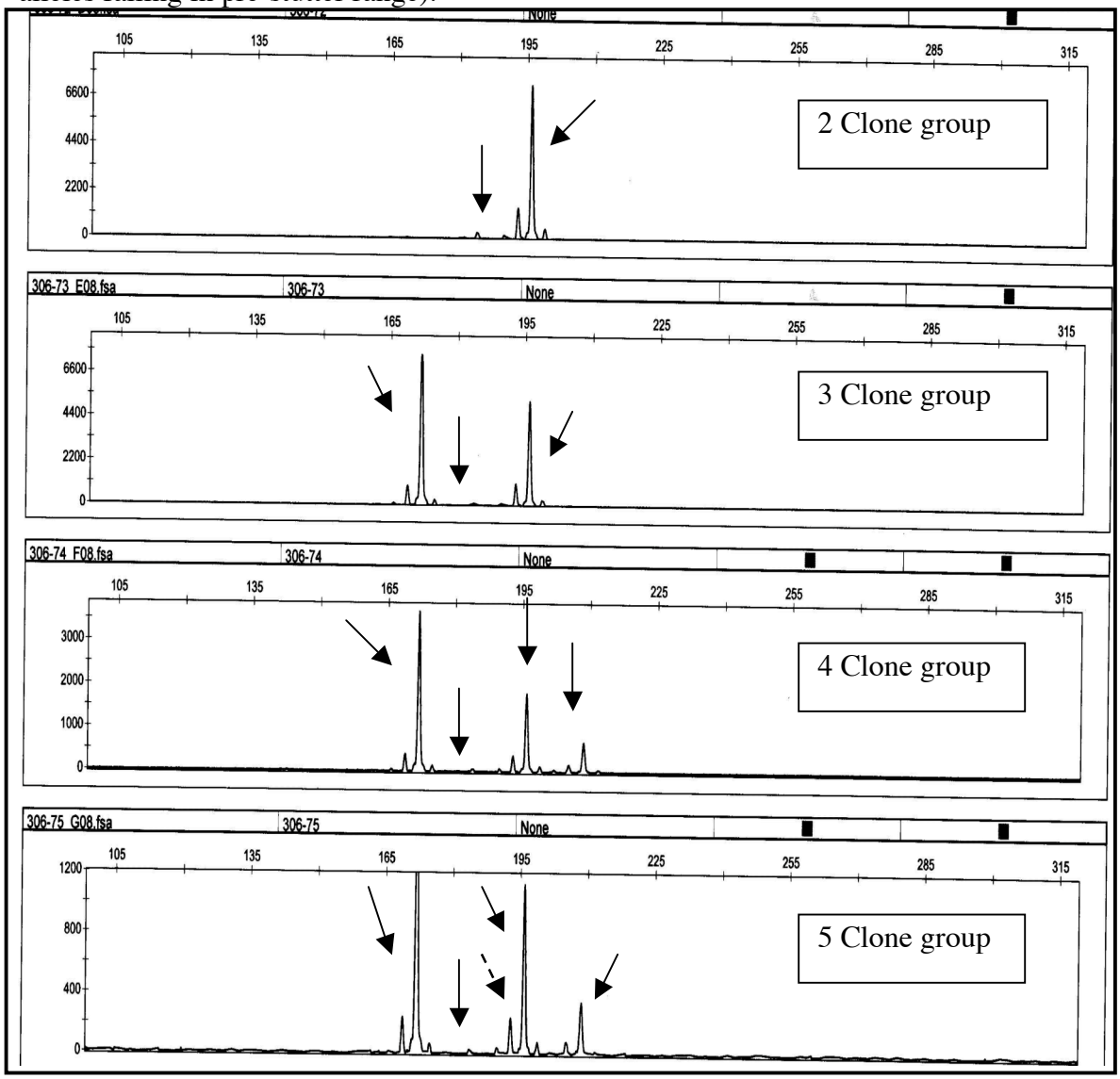
Figure 1. A pherogram resulting from mixing 7 known clones (alleles) of *P. mexicanum*. The sizes of the known clones are indicated. Note that each has at least a small peak. However, without prior knowledge of the clonal diversity present, only four peaks would have been identified (indicated with arrows). The other three alleles fell into pre and post stutter ranges, indicated by dashed arrows.



1.2. Detecting mixed clone infections: Trial 2

Shortly after the first trial, the instrumentation at the local sequencing facility (Vermont Cancer Center) was upgraded to an ABI 310 and a new software package was available for use (GeneMapper 3.5 vs. Genotyper 2.0). A trial was therefore repeated to test the sensitivity of the new instrument and software. Four different combinations of known clones (from single or two-clone infections) were used (again mixing DNA extraction product). These combinations were: 2 clones (2 υl of each of infections from two lizards), 3 clones (1 υl from each of three lizards), 4 clones (1 υl from each of four lizards) and 5-6 clones (1 υl each of five lizards). Again, 2 υl from each mix was used as DNA template for PCR amplification of locus Pmx306. The product was diluted 1:10 H2O for genotyping. For the 2, 3 and 4 clone mixtures, I was able to distinguish all alleles; however, some were much smaller than others (primarily allele size 185). For the 5-clone mixture, I again saw 4 clear alleles, but one of the alleles added into the mixture fell in the stutter range of a previous allele.

Figure 2. The pherograms resulting from GeneMapper analysis of genotyping for fragment sizes of mixed alleles of *P. mexicanum*. For the 2, 3, and 4 clone mix, all alleles were clear, although the height of the peaks varied substantially. For the 5-clone mix, only 4 were clear because one clone fell within the stutter of another allele. (Allele 185 was very small, and could only be seen when the pherograms were expanded by the GeneMapper software; dashed arrows represent alleles falling in pre-stutter range).



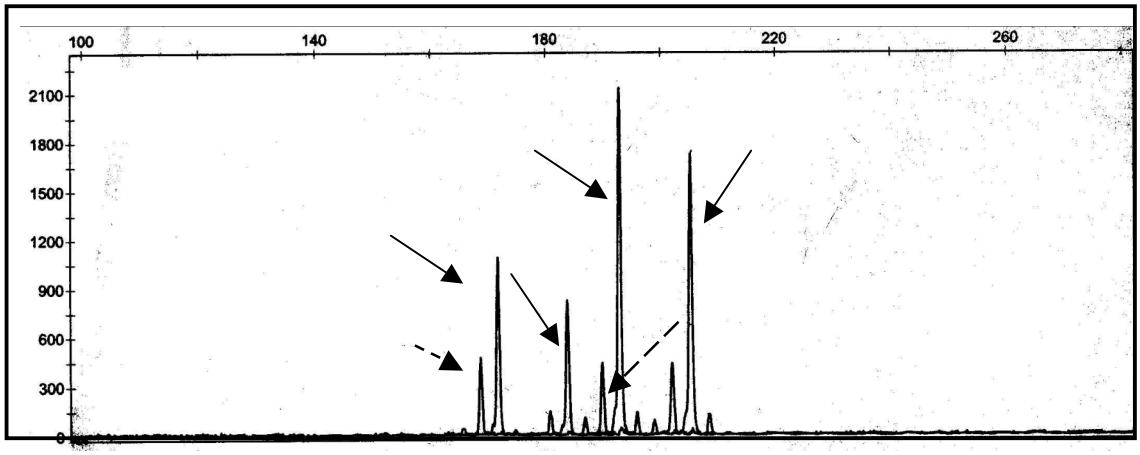
1.3: Detecting mixed clone infections: Trial 3

There were two problems with the two tests given above. 1. Alleles were used that could fall within the stutter range of other alleles in the mixture, resulting in a reduction in the number of alleles scored, and 2. The abundance of each clone within the infections differed from lizard to lizard. Although there is no clear way to measure the abundance of each clone within an infection, it can be estimated using the fluorescent units on the pherogram (Anderson et al, 1999; ABI Genescan manual). By combining DNA with differing proportions of clones, replication of

the most abundant clones was likely favored in the PCR. Therefore, this test was performed again, using two additional loci, and lizards that harbored clones with similar fluorescent unit readings on the pherograms (and thus, similar density of parasites in the blood).

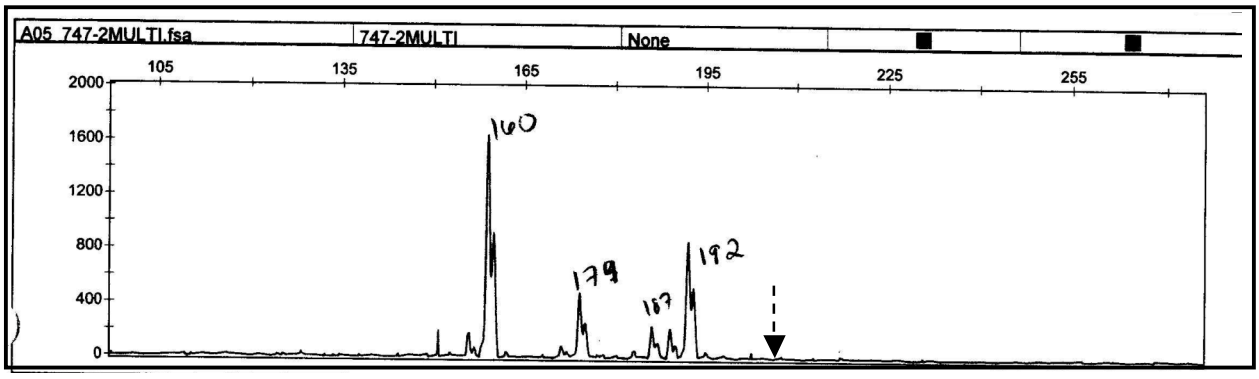
For Pmx306, 2 υl of DNA extraction product was combined from four different infected lizards. Six alleles were mixed, each with a size that should not cause stutter confusion. For the PCR mixture, 5ul of this mixture was used as DNA template and the PCR thermalcycler program appropriate for the locus was used. The product was diluted 5 υl product: 90 υl H2O for genotyping. From the resulting pherogram, 4 of the 6 alleles were clearly visible.

Figure 3. Combination of alleles for Locus Pmx306 that differed in size such that the true peaks and stutter would not overlap. Six alleles (clones) were included in the mixture of template DNA, and as seen on the pherogram, four peaks are clear. Alleles (clones) falling within the stutter range are represented by dashed arrows.



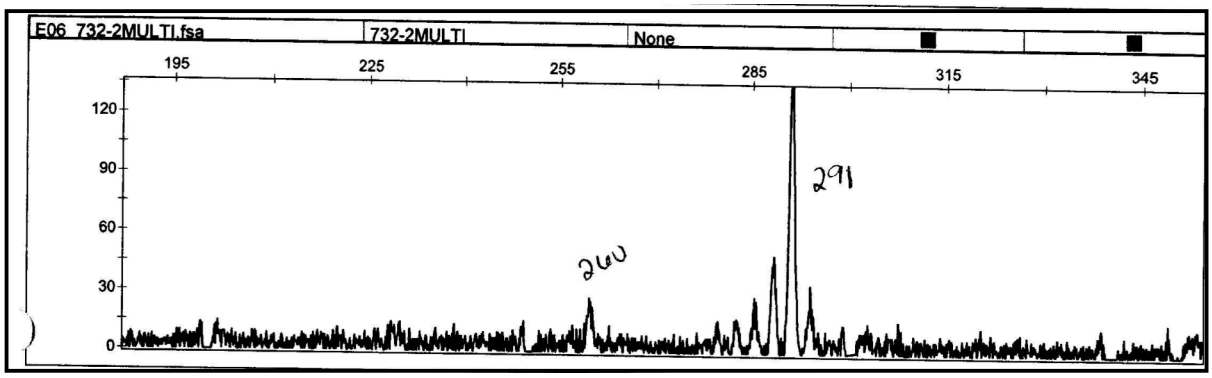
For locus Pmx747, 2υl of extracted DNA from 5 different lizards was mixed together, forming a 10υl DNA mixture with 5 different alleles. Five υl of this was used as template DNA for the PCR reaction and the program suitable for the locus was used. The product was diluted 5 υl product: 90 ul H2O for genotyping. The resulting pherogram showed 4 clear alleles.

Figure 4. A pherogram for Locus Pmx747 for a mixture of template DNA known to contain 5 alleles (clones). Note that four of these clones are obvious. The dashed arrow represents the allele (clone) that was not amplified.



For a third locus, Pmx732, 2υl of extracted DNA from 4 different lizards was mixed together, forming an 8 υl DNA mixture with 5 different alleles. 5 υl of this was used as template DNA for the PCR reaction and the thermalcycler program for Locus Pmx732 was used. The product was diluted 5 υl product: 45 υl H2O for genotyping. The resulting pherogram showed only 2 clear alleles and a very weak signal. This is a continuing problem with this locus. That is, Pmx732 tends to produce a weak signal, and most likely clones are missed. Pmx732 therefore could not be used as a single indicator of clonal diversity, but only when combined with data for other loci. This was the general protocol eventually followed: to always use several loci for any study of clonal diversity.

Figure 5. Pherogram for Locus Pmx732, with a known mixture of 5 alleles. Only two weak peaks are obvious.



Conclusions: From these tests, it is clear that up to 6 alleles can sometimes be scored from a multi-clonal infection. However, peaks falling within the stutter of other peaks, and very small peaks for low-density clones and/or for alleles that are large (large fragment length detected in the instrument), may easily be missed while evaluating the pherograms. One strategy to improve detection is to recognize that if an allele falls in the post stutter range (the first stutter peak after the main allele peak) it can still be detected if the first peak post-stutter has a higher peak than any pre-stutter peak. Alleles falling in the pre-stutter range, however, are not as clearly detected. Therefore, for many studies, infections were grouped as 1 clone, 2 clones, 3 clones, and > 3 clones. For simple counts of number of clones, the count will often underestimate the true clonal diversity.

1.4: Detecting rare clones: Trial 1

PCR is a competitive process that favors the most common template in the reaction mixture. Therefore, if two clones are present, with one much more common than the other, it is possible that only the common clone will be detected on the pherogram. The following tests examined the outcome if two competing DNA templates (two alleles) differed in concentration.
Again, multiple tests were completed using different loci and the two different ABI instruments.

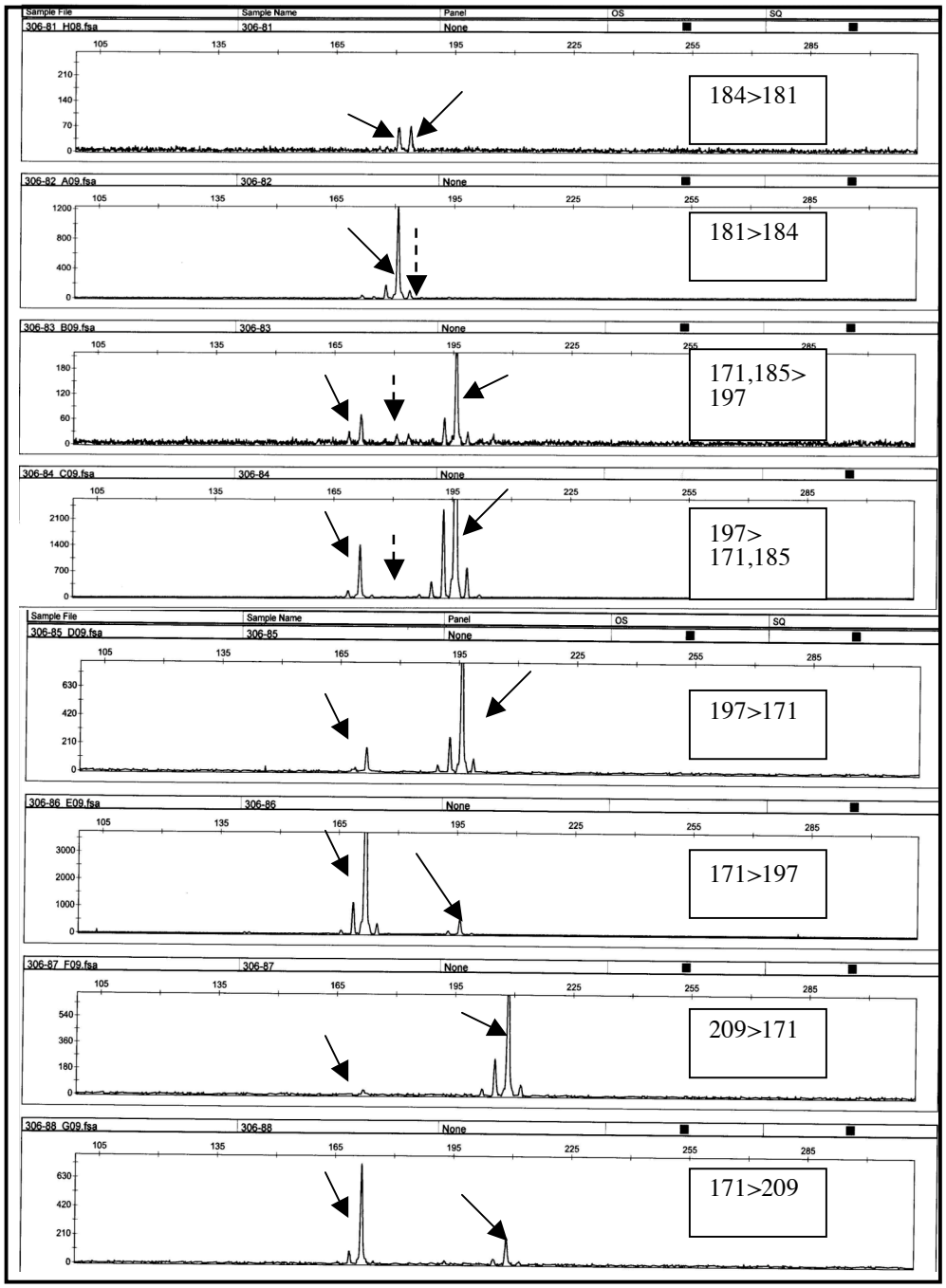
For the first trial in this group, Locus Pmx306 was used for the assays. Five alleles were chosen for mixing in various concentrations with lengths (in bp) 171, 181, 194, 197, and 209. Note that some allele pairs are similar in length, whereas others are substantially different. Table 1 shows the different concentrations of extracted DNA used in the trial. The PCR products were diluted 1:10 H2O for genotyping and were run on the new ABI machine with new software (some of the allele calls differ by 1-2 bp from those scored using the original instrument for the same lizard).

Table 1. Summary of set of trials in which five alleles were mixed (Locus Pmx306) in various concentrations. Results in bold are those that followed expectations based upon DNA concentration (expected lower amplification via PCR for lower concentrated DNA) and allele size bias (weaker signal on the pherogram is expected). Concentrations of the DNA extraction used were either full strength (non-diluted) or diluted 1/10 H2O.

	Mixture	Results
184>181	2 υl 184 + 2 υl 1/10 181	**181 and 184 (weak signal)**
181>184	2 υl 1/10 184 + 2 υl 181	**181 only**
171+185> 197	2 υl 171,185 + 2 υl 1/10 197	171 (weak signal) and 197 (strong signal)
197> 171+185	2 υl 1/10 171,185 + 2 υl 197	**Smaller 171, strong 197**
197>171	2 υl 197 + 2 υl 1/10 171	**Weak 171, strong 197**
171>197	2 υl 1/10 197 + 2 υl 171	**Strong 171, small 197**
209>171	2 υl 209 + 2 υl 1/10 171	**Strong 209, weak 171**
171>209	2 υl 1/10 209 + 2 υl 18042 171	**Strong 171, weak 209**

This trial demonstrated that smaller fragments (smaller alleles) and alleles with a higher concentration were more likely to be seen on the pherogram. However, in most cases, both alleles were seen. This result suggests that the “size” and “concentration” biases are at least sometimes not so great as to miss alleles.

Figure 6. Pherograms resulting from trials summarize in Table 1**.** Dashed arrows represent alleles that did not amplify.



The second and final tests performed to demonstrate the sensitivity of the PCR amplification utilized three loci (Pmx306, 747 and 732) and alleles with similar fluorescent units. For each locus, I chose three lizards with single clone infections harboring alleles fairly spaced apart. I tried to have two alleles similar in size and one of a different size, to emphasize the above results. I tested 1:1, 1:10 and 1:5 combinations. The 1:1 combinations of alleles were made by putting 4υl of each DNA extraction together; for a 1/10 dilution, 5 υl of DNA was mixed with 50 υl H2O) and for a 1/5 dilution, 5 υl DNA was mixed with 25 υl H2O. For these PCRs, 2.5 υl of each mixture was pipetted directly into the PCR vial; a separate mixture vial for each combination was not made (unlike the above trials). If a band showed on the PCR product gel, the product was diluted 5ul product : 90 ul H2O for genotyping. If no band was produced, 5 ul of product was mixed with 45 ul H2O.

Table 2. Lizards used for each locus and the alleles in the infection for each lizard. Shown are the alleles for each locus for each infected lizard. For example, the infection for Lizard A included alleles 195 and 172 (length).

Lizard	Pmx306	Pmx747	Pmx732
A	185 +172	187	271
B	193 + 199	193	280
C	169	169	206

Table 3. Ten combinations run for each locus.

Trial Label	**Contents**	**Trial Label**	**Contents**
2AB	2.5 υl A + 2.5 υl 1/10 B	1B1C	2.5 υl B + 2.5 υl C
2BA	2.5 υl B + 2.5 υl 1/10 A	.20B	2.5 υl A + 2.5 υl 1/5 B
2BC	2.5 υl B + 2.5 υl 1/10 C	.20A	2.5 υl B + 2.5 υl 1/5 A
2CB	2.5 υl C + 2.5 υl 1/10 B	.20C	2.5 υl B + 2.5 υl 1/5 C
1A1B	2.5 υl A + 2.5 υl B	.20BC	2.5 υl C + 2.5 υl 1/5 B

Table 4. Results for trials summarized in Table 4. For each cell, the size of the peak seen on the pherogram (large, medium, small, or no peak seen) and the allele(s) seen are given. Bolded cells represent assays that followed expectations.

Label	**Locus 306**	**Locus 747**	**Locus 732**
2AB	Large 184, Small 193, 199	Med 187, Large 193	Large 270, small 280
2BA	Small 172, 184, Large 193, 199	Small 187, Large 193	No 270, Large 280
2BC	Med 169, Large 193, small 199	No 169, Large 193	Small 206, Large 280
2CB	Large 169, Small 193	Small 169, Large 193	Large 206, Small 280
1A1B	Large 184, Small 193, 199	Small 187, Large 193	Small 270, Large 280
1B1C	Large 169, Small 193, 199	Small 169. Large 193	Small 206, Large 280
.20B	Small 172, Large 184, med 199	Med. 187, Large 193	No show
.20A	Small 172, 184, Large 193, 199	Small 187, Large 193	Small 270, Large 280
.20C	Large 169, Small 193, 199	No 169, Large 193	Small 206, Large 280
.20BC	Large 169, Small 193	Small 169. Large 193	Large 206, Small 280

This set of tests showed mixed results, depending upon which locus was examined. For locus 306, all tests, except .20C, resulted in the most abundant allele within a PCR reaction producing the larger allele peak on the pherogram. For this locus, allele size (in bp) did not seem to bias PCR amplification, except for the one case in test .20C, where the 169 allele was preferentially amplified and/or drawn up into the capillary of the Genescan instrument, resulting in a higher peak, despite that allele having a lower concentration in the PCR mixture. In summary, this suggests that for locus 306, clones that are 10x less abundant than another within the blood stream of the host can still be detected.

For locus 747, most of the experimental runs did not result in amplification of the most abundant allele within the PCR, rather, a certain larger allele, 193, was always preferentially amplified, no matter which allele it was paired with. Reduction in concentration of this allele also did not seem to have an effect on its amplification. Smaller alleles, when in weaker concentrations (2BC and .20C), were undetectable, suggesting that diversity at this locus can be extremely biased by allele size and concentration. For this locus, clonal diversity is more often than not, under estimated (this is a good reason to always use multiple loci for detecting multi-clone infections).

Locus 732, like 306, gave expected results, with the most abundant allele within a PCR reaction being preferentially amplified over alleles in lower concentration. The only discrepancy is in the 1:1 mixtures. Because PCR often amplifies smaller fragments at a higher rate than larger fragments, and because the ABI instrument preferentially draws up smaller fragments, I suspected that the smaller alleles would show a higher signal when in equal concentration with a larger sized allele. This, however did not occur for the tests on this locus, but did for locus 306.

1.5: Summary. For each locus, a clone or genotype of the parasite will be detected by a single peak on the pherogram (the parasite is haploid). Mixed-clone infections, which are common in nature, are readily detected on the pherograms. However, both the repeat length of an allele and the relative concentrations of the DNA of the clones will affect the ability to detect them via PCR and fragment analysis on the ABI instrument. PCR has been shown to bias certain fragment sizes, usually smaller fragments (as seen in the 306 examples). The ABI instrument also preferentially draws up smaller fragments, so it has been hypothesized that smaller alleles will be favored for replication and detection. This is not always the case, as shown above. If multiple loci are used to determine clonality, then the concentration of clones within a DNA sample can be as little as 1:10.

Another finding from these experiments is that detecting the number of clones within an infection can be hindered by the presence of clones represented by alleles very close together in size. Microsatellites often mutate via a stepwise model, in which one repeat unit is added or deleted with each new mutation. Stutter peaks, minor peaks observed before and after the main allele peak, are often ± 1 repeat length, so detecting whether a stutter peak is a true allele or not can be challenging. Alleles that may be present as post stutter peaks can be recognized, because these peaks are relatively high compared to pre-stutter peaks. Determining true alleles from pre-stutter peaks needs further calibration. Often times, pre-stutter peaks representing true alleles are very high compared to the true allele peak, so a 50% or even 75% rule can be used to rule out the possibility of calling an allele that is only stutter. This approach has not been tested and

therefore, I cannot account for the accuracy. Therefore, for the studies on *P. mexicanum*, any pre-stutter peaks are not counted as true alleles. Overall, clonal diversity will be underestimated, especially for infections with many parasite genotypes.

Test 2: Repeatability

To test the repeatability of the genotyping results, five infected lizards were chosen and DNA was extracted from a blood sample using the DNeasy kit (Qiagen). Using 2υl of this DNA for template, 5 separate PCRs were done using locus Pmx306. These PCRs were run on separate days with new vials of primer and water each time. After each PCR, the product was run out on a gel and then diluted to a 1:10 solution for genotyping. If a band was present on the gel for a sample, 0.5υl of the 1/10 solution was used for genotyping, but 1υl was used of the 1/10 product if no band appeared. Each PCR was genotyped on a separate day to account for day-to-day fluctuations in the accuracy of the ABI instrument. In all but 1 of the PCR/genotype trials, methods and results were the same as the original genotyping done months earlier. Thus, identical results were obtained for five samplings of the infections. One trial (the second in the series), however, produced false results. Contamination or errors in identification of the samples when run in the ABI instrument produced a pherogram with spurious alleles in all the samples for that day.

Test 3: Genotypes over time within an infection

Within an infection, multiple genotypes may be present, but not necessarily in the blood stream. Newly invading parasites replicate for a while in the liver before entering the peripheral blood. Amplification of microsatellite loci will only work on the parasite genotypes found in the blood, so any clones in the liver and/or other organs may be missed. Also, some clones may be present in the blood, but in very low densities, too low for detection during the competitive PCR (as described above). Examining the clonal diversity within a single infection over time will allow detecting if (1) Alleles mutate so rapidly as to change over the course of a three-month period (rendering the use of microsatellites for these studies to be limited); (2) if new alleles show up in the infection that weren't previously detected; and (3) if alleles within an infection remain in the same relative abundance over time.

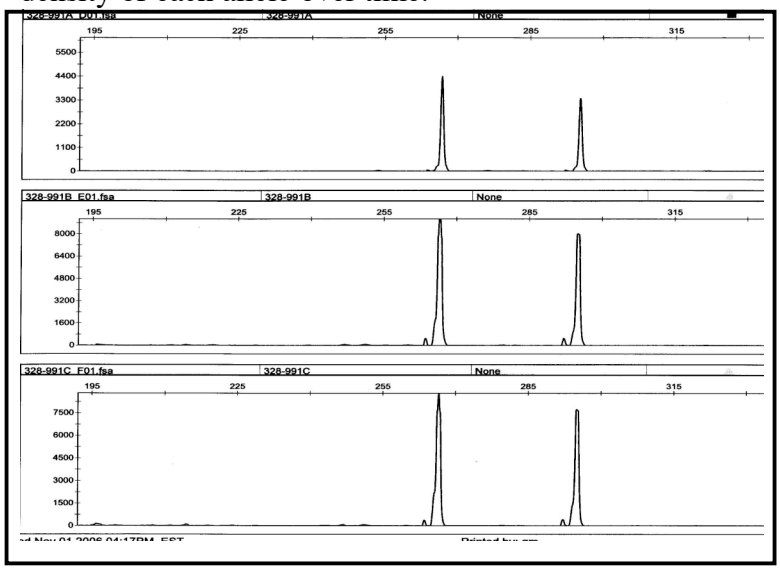
Three tests were run. For two of these, undergraduate laboratory classes in parasitology (BIOL 245, Spring 2005 and Spring 2006) participated in the laboratory duties.

3.1: Genotypes over time: Trial 1

For the first test, I utilized blood dots form Sarah Osgood's MS thesis work. I picked 4 lizards with bleed dates spanning the course of 3 months (Osgood lizard numbers 15789, 15991, 16102, 16456). For the first two lizards, the following dates were used for DNA: 6/13/01, 7/25/01, and 9/19/01. The latter two lizards had blood sample extracted from 7/10/01, 8/8/01 and 9/19/01. These DNA samples were amplified for a rather non-polymorphic locus, PM328. I used this locus because it was highly reliable and often gave good representations of mixed vs. single genotypes. In all cases, the same alleles were present throughout the entire sampling period, but

the first PCR of 16456 did not amplify well, so the data are not reliable. Only one of these infections, 15991 had >1 clone. Both clones in that infection remained throughout the period, but the density of the 'weaker' clone (smaller peak on the pherogram) increased as time progressed, while the clone that initially had the higher peak remained at the same density. This is difficult to determine with any accuracy, however, because the last two genotypes had peaks that were higher than the sizing chart, usually indicative of too much DNA (Figure 7).

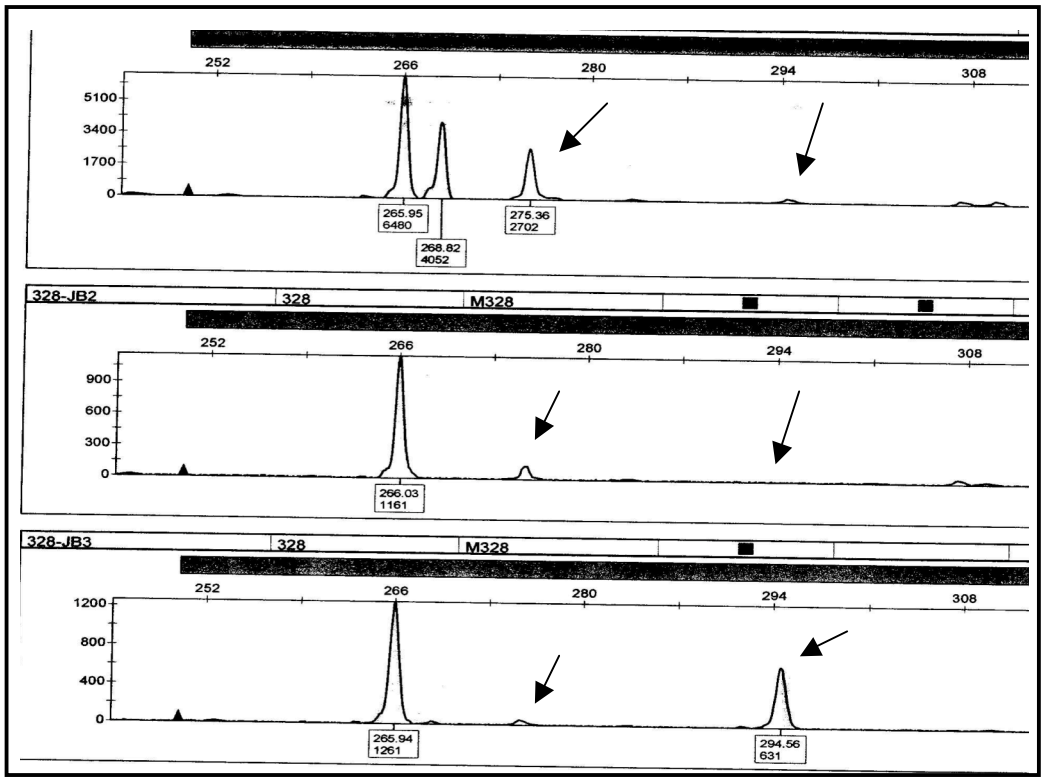
Figure 7: Pherograms of an infection followed over a three-month period. Notice the increase in density of each allele over time.



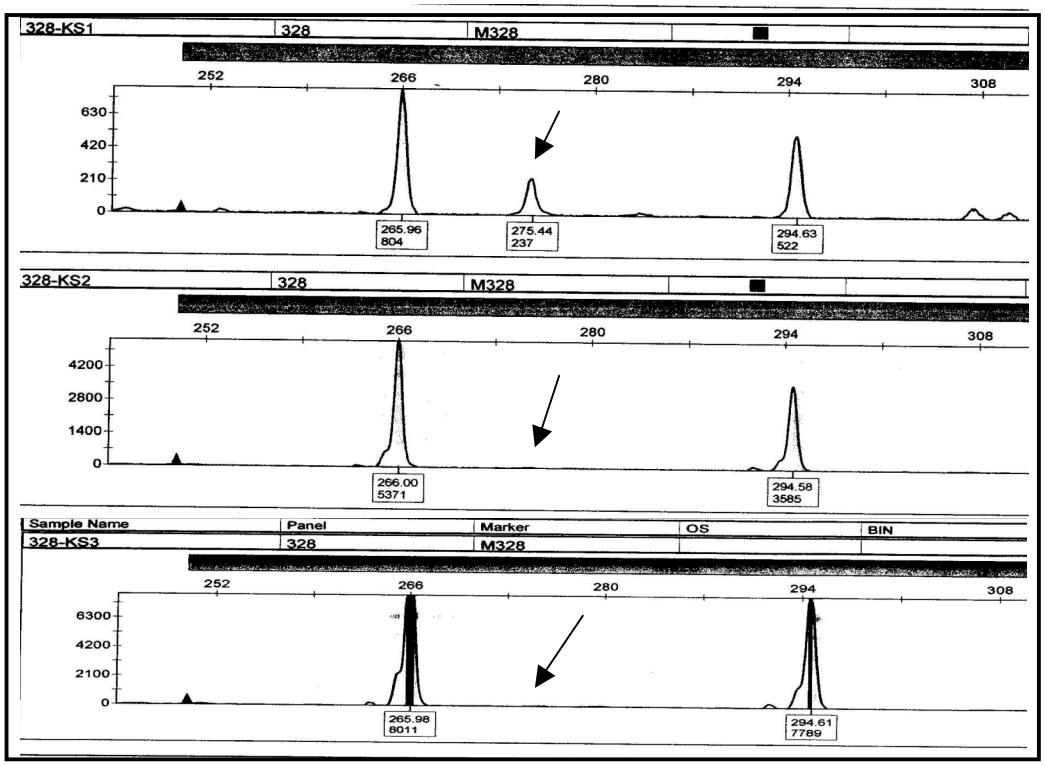
3.2: Genotypes over time: Trial 2

The next two tests were performed during the undergraduate course labs years. For the first year, 8 students followed infections of 8 different lizards, again from Sarah Osgood's data (Osgood lizard numbers 15701, 15686, 15963, 15794, 15991, 16006, 15784, 16100). These lizards were infected with blood from 1-3 donors, suggesting that diversity should differ between them. DNA was extracted for three different dates for each lizard (depending upon the lizard bleed dates and the student's choice of dates to work with). The three dates spanned 2-3 months. Locus 328 was again utilized; 5 υl of DNA template was used and PCR products were diluted 5:45 or 5:90 depending upon whether or not a band showed up on the gel. In the group, 3 lizards had single clone infections and the other 5 had 2-3 clones. Interestingly, the clones in the multi-clonal infections waxed and waned in abundance and even presence over the 2-3 month period studied. In one particular case, JB series, a clone that was not strongly present early on in the infection showed up in the last date examined, becoming a dominant clone at the end of the infection period. Two other clones observed in the first date had diminished in abundance and presence so that at the final bleed date, they were not distinguishable. A similar result was seen for the KS series (below, Figure 8).

Figure 8 (set of figures): Pherograms for JB and Ks series, respectively. Notice how allele presence and abundance can vary over three months.



J B series

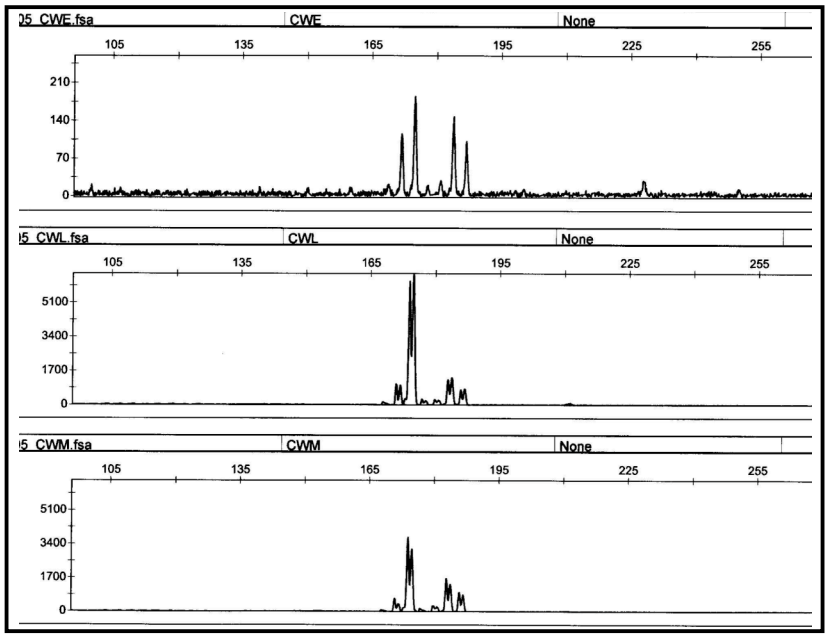


KS series

3.3: Genotypes over time: Trial 3

The second year, 9 students followed an additional 9 individual infections from the same data set (Osgood lizards 16076, 15789, 16102, 15805, 16005, 15679, 15786, 15687, 16098, 15973). Similar results as above were found. No new alleles appeared over time, but the abundance of each clone did appear to vary over the three dates examined (Figure 9).

Figure 9: Pherogram for an infection followed over three months. Again, notice the change in allele density over time.



3.4: Summary

The results of these over time studies are interesting when more than two clones appear to be present, but indicate the challenge of "defining" clonal diversity for an infection. As seen in the two examples from the 2005 Parasitology lab, clones rarely appeared and disappeared throughout a three-month period, but they did vary in their density over time. If these infections were watched longer than the three months studied, the clonality of the infection may be different than the month prior. A clone that decreases in its abundance in the blood stream may become undetectable if the concentration is low (see above tests of concentration vs. detection). However, in no case did a "new" clone appear that was $\pm$ one repeat change, indicating that the alleles were not rapidly mutating at these loci, rendering them stable for long term studies.