

Strategies for Genotyping: Effectiveness of Tailing Primers to Increase Accuracy in Short Tandem Repeat Determinations

Doug Bintzler, Chair, University of Cincinnati

Pamela Scott Adams, Trudeau Institute

Yongde Bao, University of Virginia

Lynn Petukhova, The Rockefeller University

Duane Bartley, Past Chair, Johns Hopkins University

Linda Wood Ballard, University of Utah

Laura Kasch, Johns Hopkins University

Caprice Rosato, Oregon State University

INTRODUCTION

A potential source of genotyping errors is introduced by the tendency of *Taq* DNA polymerase to incorporate a nontemplated nucleotide at the 3' end of PCR products. Particularly when the short tandem repeat (STR) under analysis is a dinucleotide, this activity can produce a population of products both with and without the additional nucleotide. The resulting data is difficult to interpret and thus error prone. A common way to address this issue is to force the reaction to the nucleotide addition by adding a short sequence that has been empirically determined to drive the reaction in that direction. This process is known as tailing primers.

The Fragment Analysis Research Group devised a study to demonstrate the usefulness of this method. Participating laboratories were sent two samples of pooled PCR products for electrophoresis. Each sample contained three fluorescently labeled PCR products: a dinucleotide STR amplified with ordinary primers, a dinucleotide STR amplified with a tailed primer, and a tetranucleotide STR. Each dinucleotide STR amplified two alleles that differed in size by four base pairs. As in the past, participants were asked to fill out a web-based survey regarding run conditions and results. The results were analyzed to determine the accuracy and precision of the allele calling and to identify factors that may affect sizing, including the type of primers used, as well as run conditions. Additionally, we have attempted to assess the current trend of genotyping by compiling demographics of survey results.

EXPERIMENTAL CONDITIONS

- Two study samples
- Each sample had 3 fragments (6 alleles)
- Size of each fragment determined by database
- Tetranucleotide repeat – HEX labeled
Forward: 5' CTGCCAGAGAATAGGGTGAA 3'
Reverse: 5' TTGTAATCTTGGTCTGCC 3'
- Dinucleotide repeat with untailed marker – HEX labeled
Forward: 5' TTGCACAAAGCATCAGCC 3'
Reverse: 5' TCAGGAACCTCAGCCAT 3'
- Dinucleotide repeat with tailed marker – FAM labeled
Forward: 5' TTGCACAAAGCATCAGCC 3'
Reverse: 5' GCTTCTCAGGAACCTCAGCCAT 3'

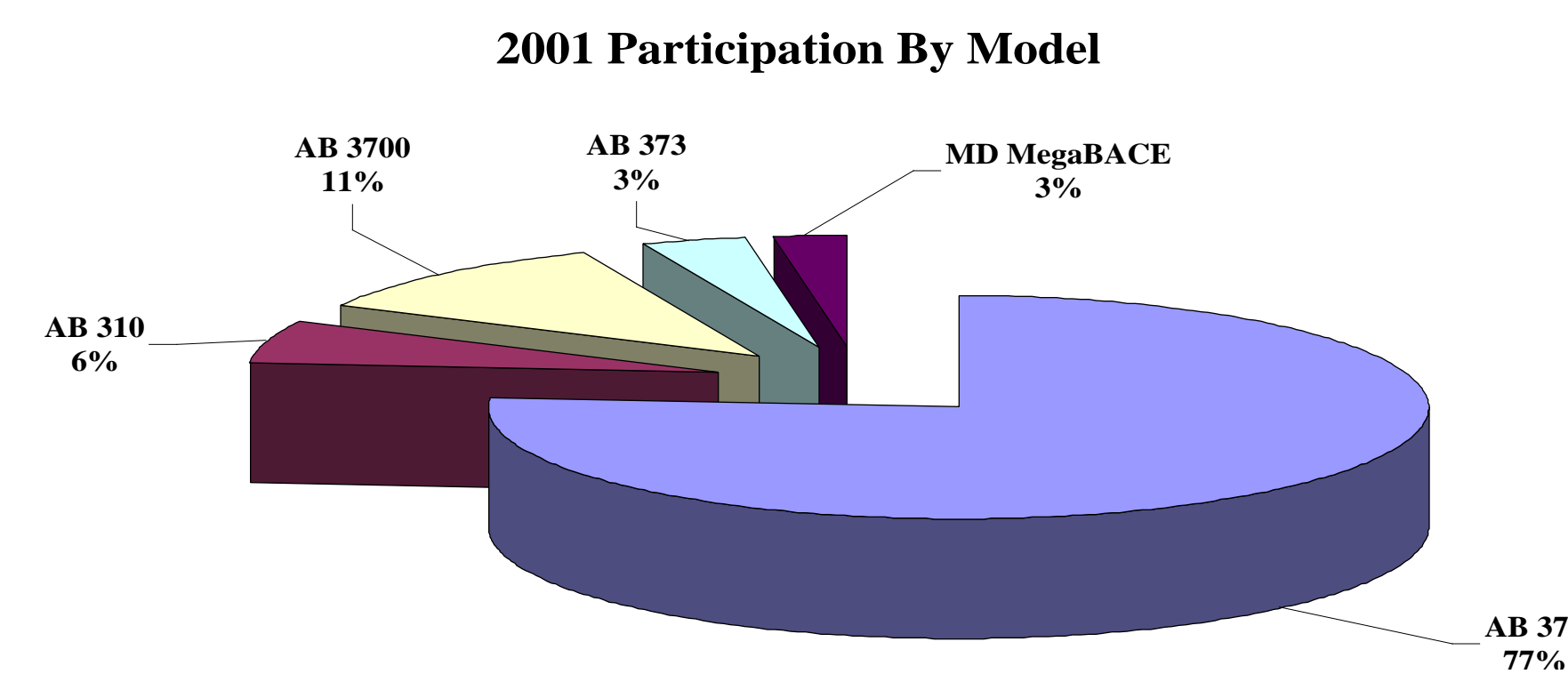


Figure 1. Participation by Model: This figure summarizes the number of samples submitted to the 2001 ABRF FARG Study. Machines are differentiated by manufacturer and model type. In total, 116 samples were received with data from 23 different machines representing 5 different models.

Participation by Year

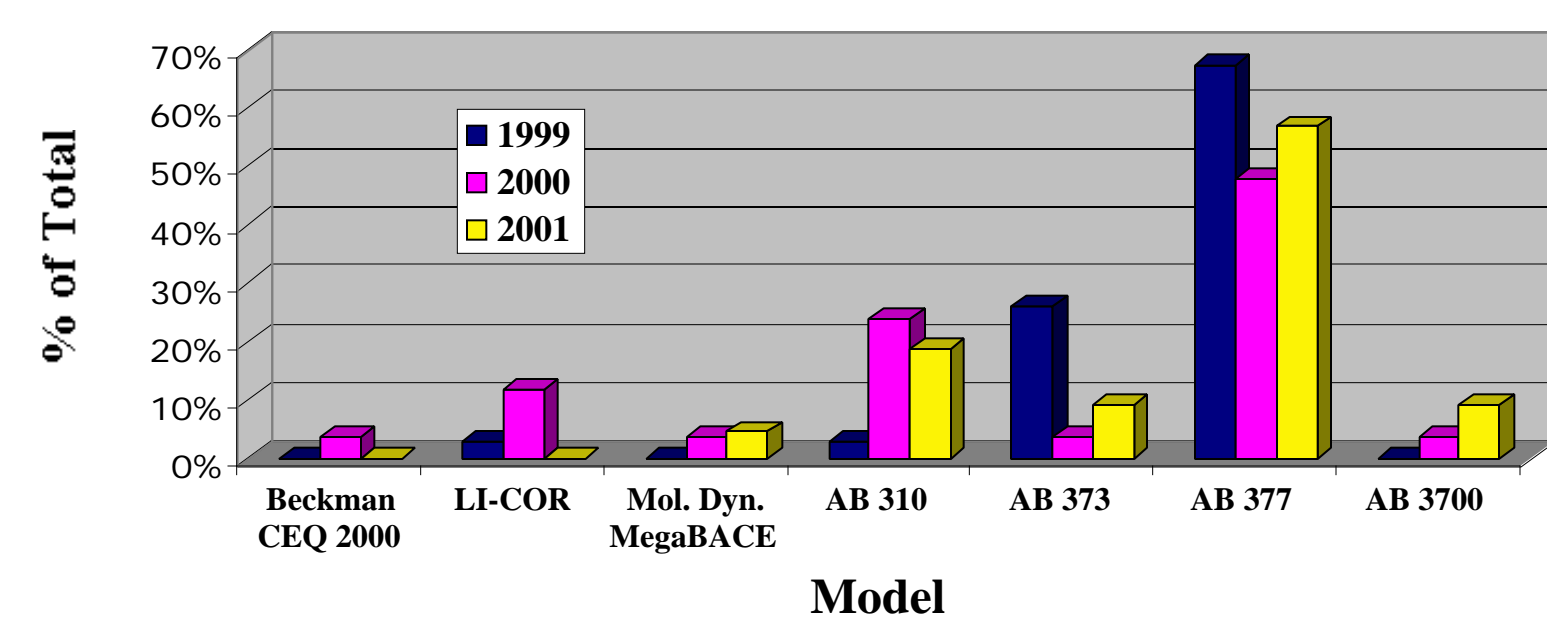


Figure 2. Participation by Model by Year: This figure summarizes the participation by models of instrument in ABRF FARG Studies over the past 3 years. Machines are differentiated by manufacturer and model type.

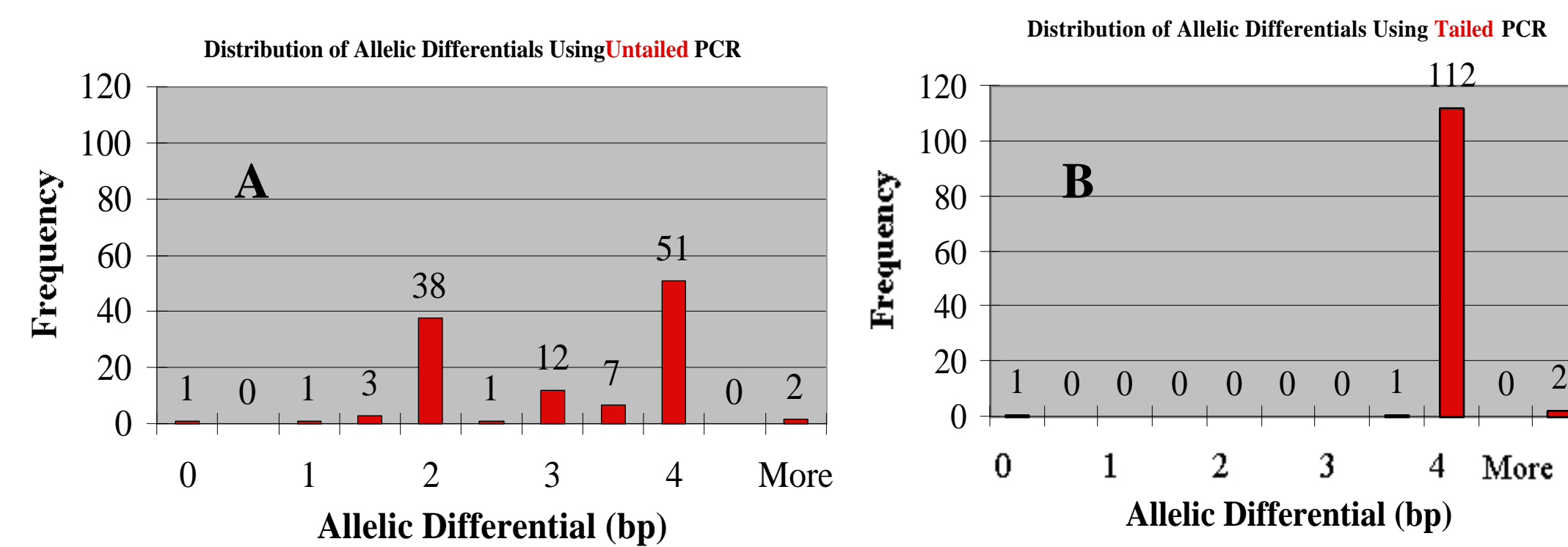


Figure 3. The effectiveness of the tailed primer in aiding automatic allele calling in dinucleotide repeats is assessed by comparing the allelic differentials obtained from both untailed (Panel A) and tailed (Panel B) primers. As can be clearly seen, 112 allelic differentials (97%) were correctly identified to be 4 bp, which is expected. While in the case of the untailed primer, only 51 (44%) were correctly called in terms of allelic differentiation.

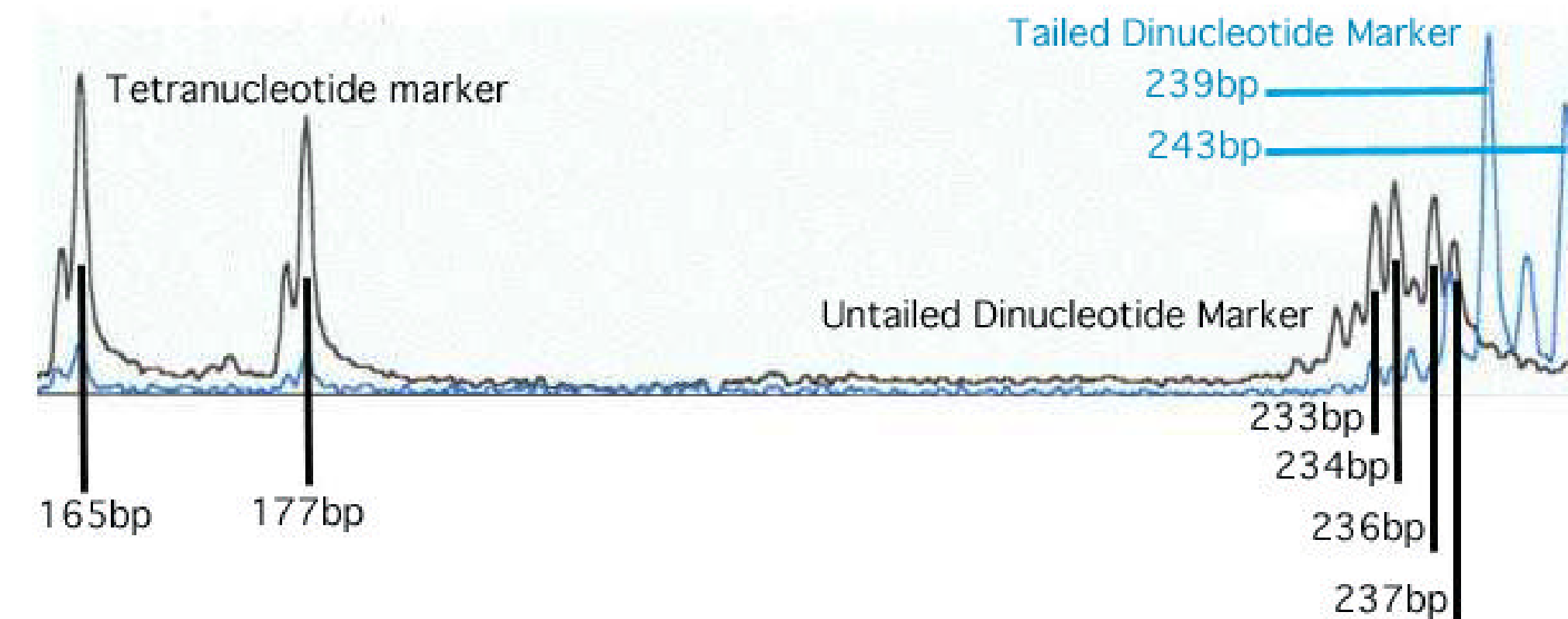


Figure 4. This figure illustrates the improvement of peak characteristics by primer tailing to aid in proper allele calling. The untailed dinucleotide marker (larger black marker) demonstrates the difficulties in calling the correct peak for each allele. Automated allele determination with Genotyper software would call the two highest peaks, 234 & 236, and although these would bin correctly for a dinucleotide repeat marker, the genotype would be incorrect. This can be seen by comparing the same marker with a 6 base pair tail (blue). The addition of a tail to the reverse PCR primer forces the addition of a non-templated "A" to the product, verifying that the 233 & 237 peaks are the true alleles. The plus "A" issue also can be worked around by using a marker that is a trinucleotide or tetranucleotide as shown in the figure at 165 & 177bp.

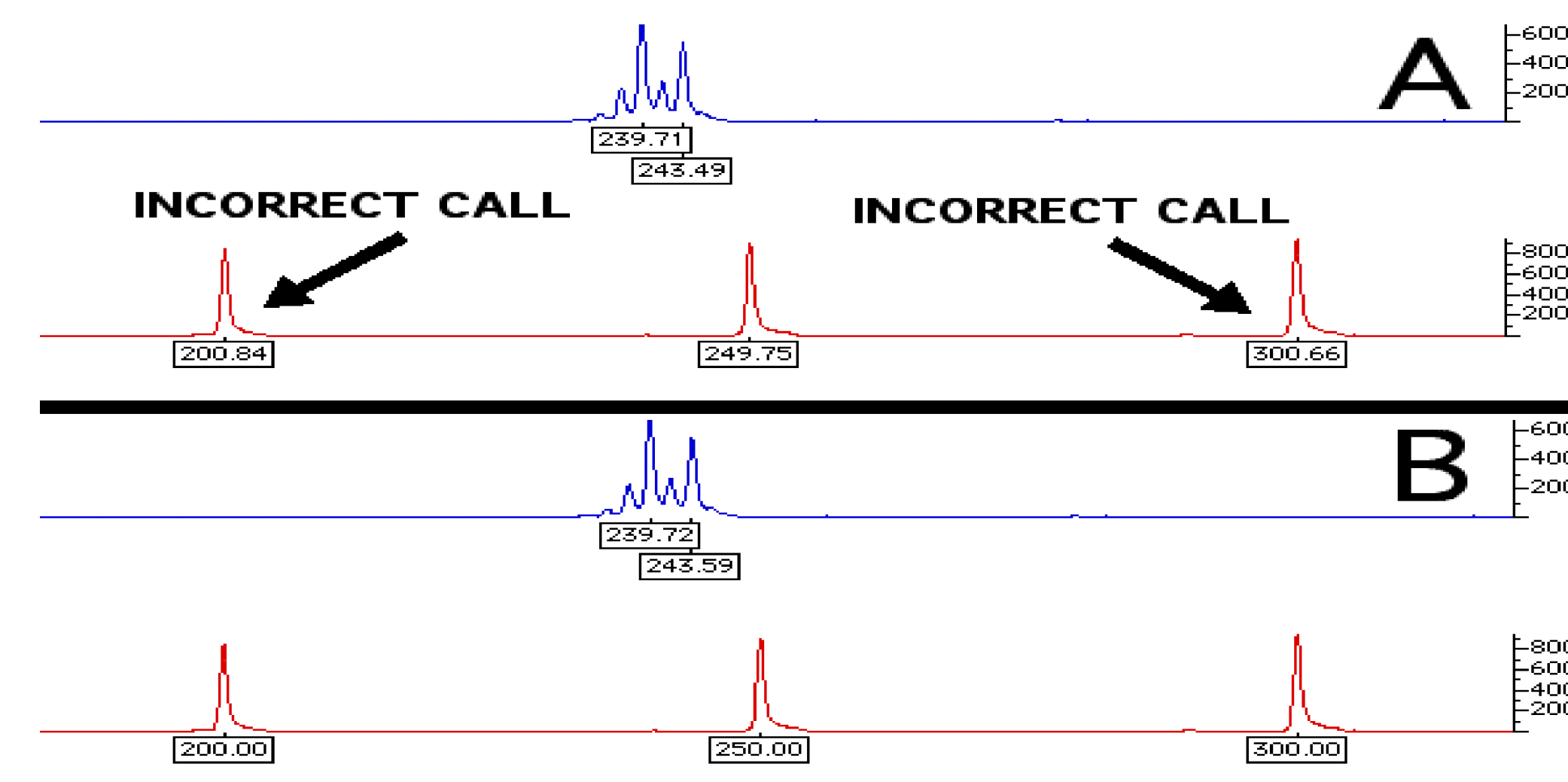


Figure 5. This figure illustrates the importance of correctly calling the size standard to obtain accurate size calls. In frame "A" the alleles for this marker would be rounded to 240 and 243bp. This is a dinucleotide repeat marker but has a spacing of 3bp between alleles. Binning this marker as a dinucleotide would cause one of the two alleles to fall outside a bin. Inspection of the size standard (red) shows that the standard points have been miscalled. In frame "B" the size standard has been corrected and the lane reanalyzed with a new standard curve. This produced allele sizes of 240 and 244bp, which would fall within a dinucleotide bin.

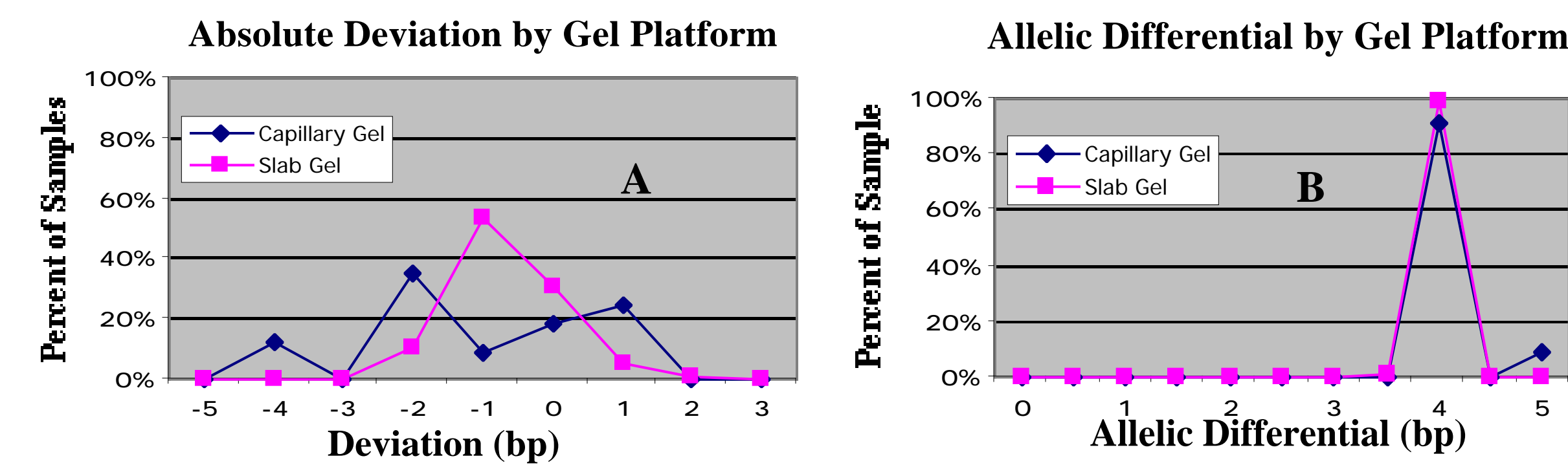


Figure 6. The Effect of Gel Platform Using Untailed and Tailed Primers. This figure shows the performance of the two different platform types of fragment analyzers when compared by the deviation in base pairs (bp) of individual alleles from the correct size (Panel A) or by the difference in base pairs (bp) of paired alleles (Panel B). Only the data obtained from using tailed primers is included. Slab gel models are the AB Models 373 and 377. The capillary based models are the AB Models 310 and 3700 and the Molecular Dynamics MegaBACE. If using tailed primers, both platforms are equally capable of determining the correct genotype.

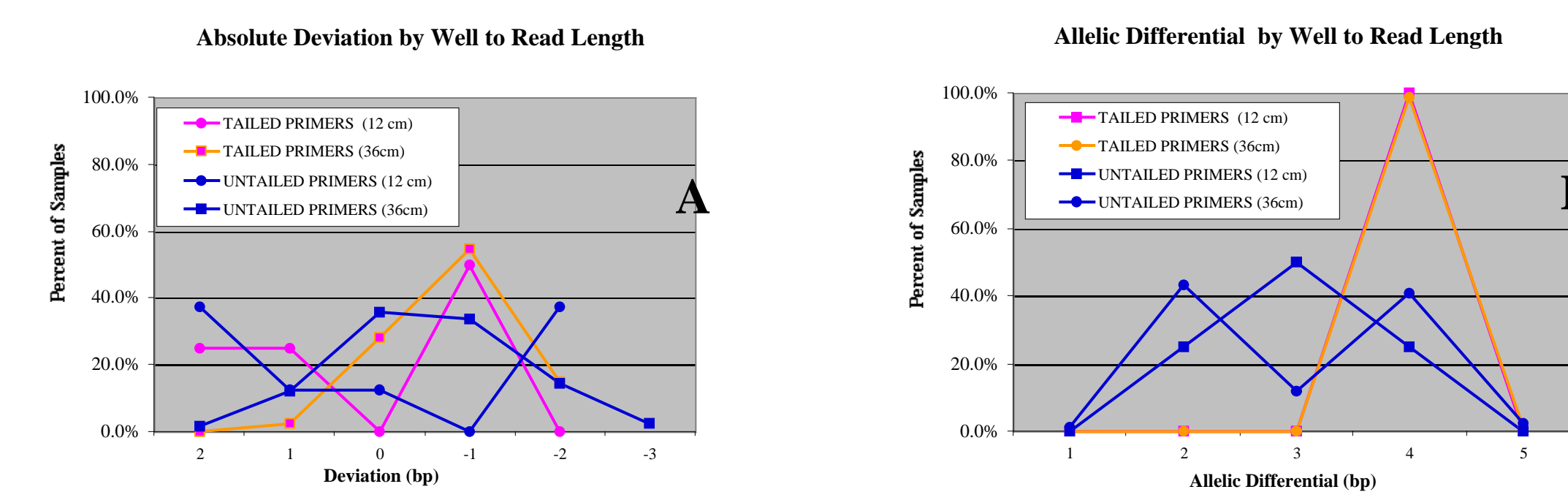


Figure 7: The Effect of Well To Read (WTR) length Using Untailed and Tailed Primers. Base pair deviation of individual alleles from the correct size (Panel A) or by the difference in base pairs (bp) of paired alleles (Panel B) for the 12 cm (WTR) length and the 36 cm (WTR) length ABI model 377. When compared by deviation, the samples amplified using an untailed reverse primer show a wide range of deviation values compared to the samples amplified with the same dinucleotide marker containing a six base tail attached to the reverse primer. When compared by allelic differential, tailing the primer eliminates the effect of WTR.

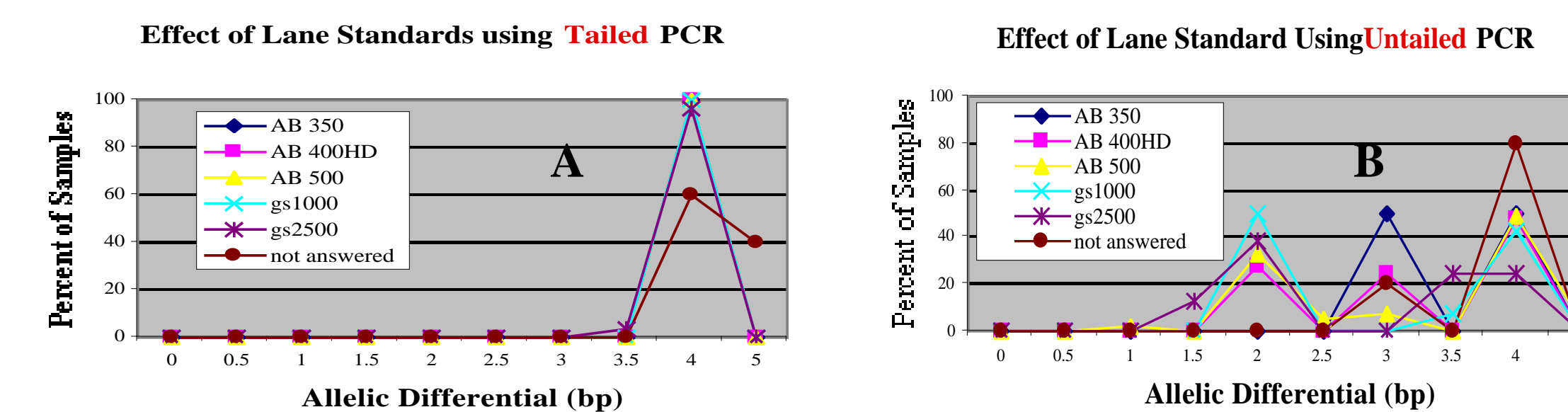


Figure 8: The Effect of Lane Standard Using Untailed and Tailed Primers. Allelic differential is used to measure sizing accuracy for the tailed and untailed primers. Figure A contains data generated with the tailed primer set. Figure B contains data generated with the untailed primer set. The results from the untailed primer set are highly variable and appear to be dependant upon the type of lane standard used. It appears that tailing the primer eliminates the effect of the lane standard.

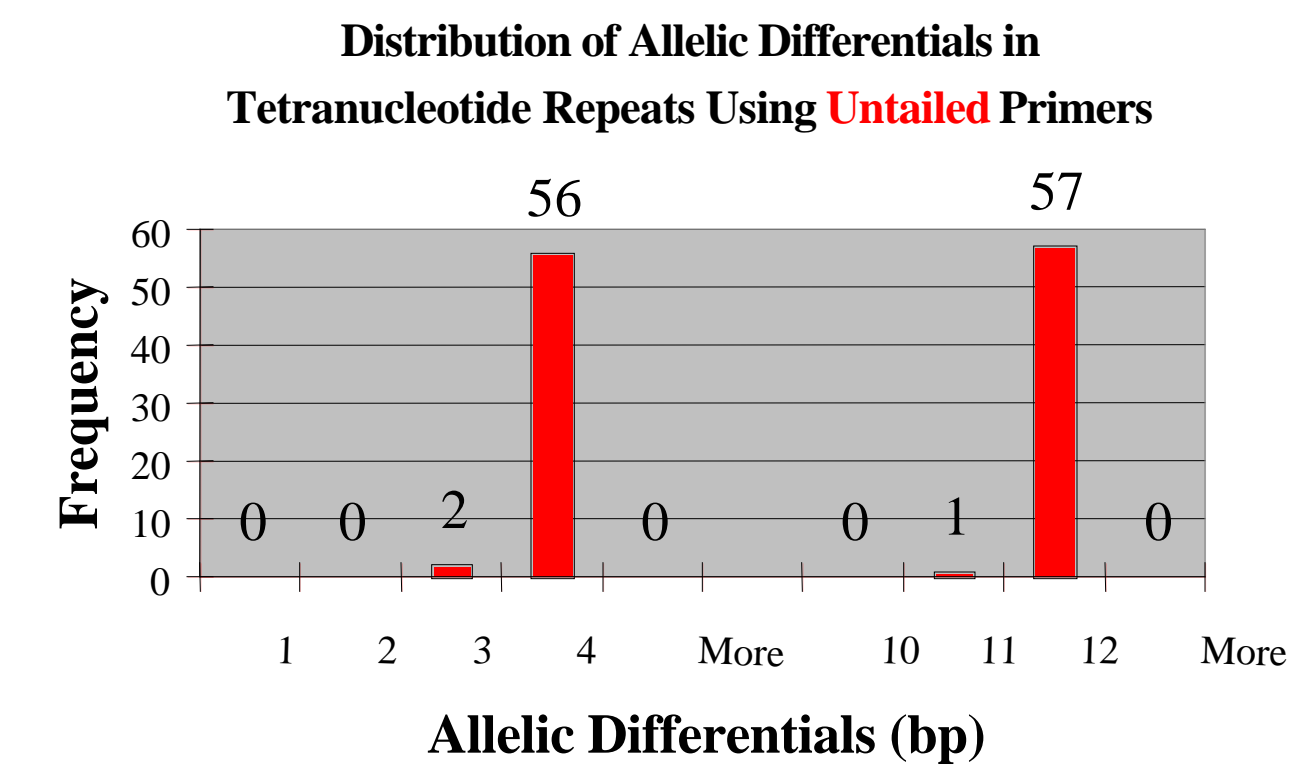


Figure 9. This figure shows the allelic differentials of the two different tetranucleotide repeats using regular primers. As clearly illustrated, the regular primers yielded the expected results (4 bp and 12 bp, respectively, for the two amplicons). Therefore it can be concluded that tailed primers are not needed with the tetranucleotide repeats.

CONCLUSIONS

1. The general performance by the participants of FARG study 2001 was excellent, with 97% of the allelic differentials called correctly.
2. The Applied Biosystems Model 377 is still used by a majority of the facilities (77%) participating in the study.
3. Adding a tail to a dinucleotide marker does help in identifying the peak for an allele and improves accuracy.
- 97 % correct allelic differential with the tailed marker
- 44 % correct allelic differential with the untailed marker
4. Lane standards must be assigned correctly to obtain accurate and precise results.
5. Capillary (91%) and Slab gel (99%) systems both performed well with tailed markers.
6. No other factors, including WTR or Lane Standards appeared to have an effect on accuracy when tailed markers are used.

REFERENCES

Smith JR, Carpen JD, Brownstein MJ, Ghosh S, Magnuson VL, Gilbert DA, Trent JM, Collins FS. 1995. Approach to genotyping errors caused by nontemplated nucleotide addition by *Taq* DNA polymerase. *Genome Res* 1995 Oct;5(3):312-7

Magnuson VL, Ally DS, Nylund SJ, Karanjawala ZE, Rayman JB, Knapp JJ, Lowe AL, Ghosh S, Collins FS. 1996. Substrate nucleotide-determined non-templated addition of adenine by *Taq* DNA polymerase: implications for PCR-based genotyping and cloning. *Biotechniques* 1996 Oct;21(4):700-9

Brownstein MJ, Carpen JD, Smith JR. 1996. Modulation of non-templated nucleotide addition by *Taq* DNA polymerase: primer modifications that facilitate genotyping. *Biotechniques* 1996 Jun;20(6):1004-6, 1008-10

Hengen PN. 1995. Fidelity of DNA polymerases for PCR. *Trends Biochem Sci* 1995 Aug;20(8):324-5

Haag E, Raman V. 1994. Effects of primer choice and source of *Taq* DNA polymerase on the banding patterns of differential display RT-PCR. *Biotechniques* 1994 Aug;17(2):226-8