FARG 2004 Study: Comparison of Genotyping Results Across Instrument Platforms Using a Mock Forensics Investigation

Caprice Rosato[1] †, Yongde Bao[2], Doug Bintzler[3], Stephen A. Bobin[4], Joe Forrester[5], Robert Keefe[6], Rebecca Scholl[7], Katia Sol-Church[8], Tom Stelick[9]

[1] Oregon State University, [2] University of Virginia School of Medicine, [3] University of Cincinnati, Chair FARG, [4] Dartmouth College, [5] University of Missouri-Columbia,

[6] Wadsworth Center/NYS DOH, [7] University of Utah, [8] Alfred I. duPont Hospital for Children, [9] Cornell University † to whom correspondence should be directed

Abstract

Almost all platforms used for genotyping analysis depend upon the determination of the size (in base pairs) of DNA fragments (amplicons) generated by the polymerase chain reaction (PCR). A good number of these platforms are capable of accurately distinguishing differences in amplicon size – translated here as *allele* size – as small as a single base pair. The instruments available today that are used for genotyping are generally quite precise, but often inaccurate (with the exception of mass spectrometers) due to variation of factors that influence the calculation of DNA fragment size – like electrophoresis run conditions, the allele calling software used, the DNA sequence of the amplicon, and the covalent attachment of various fluorescent dyes to amplicons. How scientifically feasible is it then to conduct a genotyping project using more than one kind of instrument or platform - e.g., as might occur in a collaborative effort between separate labs? One aim of this study was to address this question by looking at size calling variability between instruments, or within a particular instrument model, when amplicons of known size were distributed to different labs for analysis.

Introduction

The Fragment Analysis Research Group (FARG) organized its 2004 study with an on-going interest in learning about the types of instruments and platforms being used in laboratories to carry out genotyping analysis (fragment analysis). A previous study (FARG 2000) looked into the effects of using different run conditions, reagents, and instrumentation on DNA fragment size determination. The present study is similar, seeking to determine whether the introduction over the past four years of new genotyping technology has affected the feasibility to carry out genetic variability analysis across platforms. FARG's survey last year presented evidence that the use of capillary electrophoresis instruments was increasing versus slab-gel based instruments – although it wasn't clear whether the former was replacing the latter, or being used in addition to them. Researchers from different labs wishing to collaborate on genotyping projects must keep in mind the differences or biases that exist amongst different analysis platforms in their ability to correctly call allele sizes, and decide how to accommodate such differences in their analysis.

To address these concerns, and to make the survey more interesting for participants, FARG designed the present survey as a forensic study where two samples with different genetic profiles were used to represent a crime "victim" and evidence ("hair") found at a fictitious crime scene. Seven plant genetic markers were chosen for amplification by PCR. The various DNA samples used as PCR template representing nine "suspects", the "victim", and a "hair" found at the crime scene were all derived from sunflower (*Helianthus annuus*) , and were selected so that none of the amplicons was heterozygous for any of the seven genetic markers. Platform-specific instructions (for slab gel or capillary instruments) for preparing and analyzing the DNA samples were included with mailed samples. Responses were summarized by FARG members to determine the extent any amplicon's calculated size deviated from its expected size. Also, for each of the seven genetic loci (markers) used in the study, FARG members determined the difference (in base pairs) between the "victim's" allele size and that of the crime scene sample, and looked to see how consistent this *allelic differential* was for each marker when determined by different instruments.

Methods

All amplicons were labeled with either HEX or FAM during the PCR using labeled primers, and were supplied by Dr. Shunxue Tang of the Department of Crop and Soil Science, Oregon State University. Some study participants possessed genotyping instruments that were incompatible with the use of HEX and FAM fluors. In these cases, Cy5 labeled amplicons were provided by Doug Bintzler at the University of Cincinnati. All amplicons were generated as single-product PCRs, diluted to give optimum instrument detection, and then multiplexed so that each of the 11 samples sent to each study participant was composed of seven labeled amplicons representing the allele sizes generated from the seven different genetic markers examined. Seven distinct genetic profiles (seven markers per profile) were used to generate nine "suspects", one victim, and a "hair" sample found as evidence at a fictitious crime scene, and study participants were asked to determine which suspect's genetic profile matched that found at the crime scene. One suspect's sample was intentionally contaminated to demonstrate the implications of how this could affect the outcome of a forensic investigation. Identical samples of desiccated, fluorescently labeled amplicons were distributed by mail to survey participants, who were instructed to determine the size of each amplicon (allele) using whatever platform was currently used in their lab. Attempts by FARG to sequence the PCR products of the seven different loci to determine true allele sizes were unsuccessful. The allele sizes submitted by study participants were compared to "expected" allele sizes that were determined by Dr. Shunxue Tang using an ABI model 377 DNA Analyzer.

After sample analysis, participants were asked to fill out an online survey (hosted by the College of Business, Oregon State University) to report conditions they used to genotype samples and the allele sizes they obtained. Also, a file transfer protocol (ftp) server was created to collect actual data files (instrument output) to allow FARG to assess if allele sizing was performed properly. Some study participants did not submit instrument files to the ftp site, and some that were submitted could not be analyzed and therefore corroborated by FARG members.

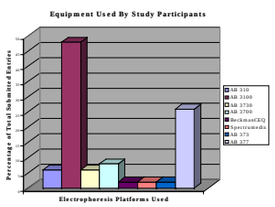

Fig. 1: Capillary instruments were the dominant form of electrophoresis platform used for study results with the AB 3100 having the largest contribution. AB 377 slab gel was the second largest contributor.

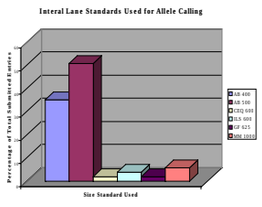

Fig. 2: Choices for internal lane standard for genotyping analysis showed a range in size ladders, the most frequent being the 400 and 500 bp.

	ABI 377	ABI 310 CE	ABI 3700 CE	ABI 3100	ABI 3730	CEQ	SEC 9610
Long Ranger	9	0	0	0	0	0	0
Amresco 29:1	1	0	0	0	0	0	0
Gene Page Plus	2	0	0	0	0	0	0
POP 4	0	2	0	18	0	0	0
POP 6	0	0	5	1	0	0	0
POP 5	0	0	0	1	0	0	0
POP 7	0	0	0	4	3	0	0
LPA	0	0	0	0	0	1	0
MFGR-180-001	0	0	0	0	0	0	1

Table 1: Polymer Types Used by Study Participants. Left-hand column in table describes the brand names of the polymers used by study participants as a substrate for electrophoresis of labeled DNA fragments. "Long Ranger" (Cambrex/BioWhittaker), "Amresco 29:1", and "Gene Page Plus" (Amresco, Inc.) were polymers used with slab gel-based platforms. All other polymers were used by capillary instruments. Top row of table lists types of genotyping instruments used by participants: "ABI", Applied Biosystems, Inc.; CEQ, Beckman-Coulter; SCE 9610, SpectruMedix.

Allelic Profile of Study Samples

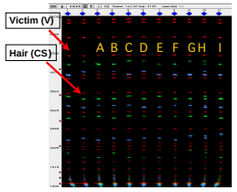

Figure 3. The figure above is a digital reproduction of a urea/polyacrylamide gel that was loaded with seven labeled PCR products in each gel lane followed by electrophoresis on an ABI 377 DNA Sequencer. PCR products were amplified from seven different plant markers and labeled with the fluorescent dyes 6-FAM (blue) or HEX (green). Red colored bands are ROX-labeled molecular weight standards run in each lane.

Allele Size Determinations Made by Participants

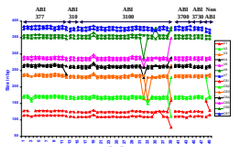

Figure 4: This graph plots the allele sizes obtained from seven different genetic markers by all the study participants for the "victim" (V) and the "Crime Scene" (CS) samples. The different electrophoresis instruments used are indicated at top of graph. The ABI 377 instrument is the only slab gel instrument. All others are capillary based. Large deviations from average values were due to missing data points, or stutter peaks and/or PCR degradation products called as alleles by instrument software.

Effect of In-lane Standards on AD

Loci	1	2	3	4	5	6	7
Expected Allele Sizes (bp) Victim/Crime Scene	128/134	168/173	233/238	226/232	285/349	358/358	
Expected AD `between Victim & Crime Scene	14	5	5	4	6	9	8
Average Deviation From Expected AD by standard type							
400 standards (n =18)	0.82	1.04	-0.05	1.09	0.72	0.00	0.27
500 standards (n =23)	0.15	0.87	-0.11	0.74	0.65	-0.43	-0.14
600 standards (n =4)	0.57	0.93	-0.23	1.04	0.77	-0.01	0.18
1000 standards (n =3)	0.12	0.79	0.13	1.25	0.81	0.03	0.43

*Allelic Differential

Table 2: Study participants used a total of four different in-lane molecular weight standards. The use of any one in-lane standard showed less than a 1.50 base pair shift in AD. Any one in-lane standard had an average deviation of 0.50 base pairs for all seven markers. Markers 3 and 6 scored the closest to expected ADs by all four standards. Markers 2 and 4 had the greatest deviation from expected AD.

Analysis of DNA fragment data collected from different labs

Markers	Average size call in bp		A-factor
	Slab Gel	CE	
CS1	113.56	110.48	3.1
V1	127.74	124.95	2.8
V2	167.85	165.27	2.6
CS2	173.53	171.23	2.3
V3	233.32	229.93	3.4
CS3	238.17	234.85	3.3
CS4	262.69	259.88	2.8
V4	267.48	264.86	2.6
CS5	284.82	283.25	1.6
V5	297.18	295.92	1.2
V6	348.99	348.48	0.5
CS6	357.97	357.21	0.8
V7	376.59	373.18	3.4
CS7	384.38	381.34	3.0

Table 3: Average allele calls of 7 markers amplified from Victim (V) and the Hair samples found at Crime Scene (CS). Allele sizes are averages for either slab gel (12 entries), or capillary electrophoresis (CE, 36 entries) platform types. Slab gel platforms invariably called larger allele sizes than did the CE platforms. The delta-factor is the difference in fragment sizes as determined on a slab gel vs. CE platform.

Marker	Average AD (SD)	
	Slab Gel	CE
1	14.1 (0.4)	14.5 (0.5)
2	5.7 (0.3)	5.9 (0.2)
3	4.8 (0.5)	4.9 (0.2)
4	4.8 (0.6)	4.9 (0.4)
5	6.3 (0.4)	6.7 (0.3)
6	8.9 (0.5)	8.8 (0.4)
7	7.8 (0.6)	8.1 (0.3)

Table 4: Average Allelic Differential (AD) between the V and CS alleles for each of the 7 markers tested. Standard Deviations (SD) are as indicated in parenthesis and result from 12 entries using slab gel platforms and 36 entries using CE platforms.

Effect of Polymer Type on Allelic Differential

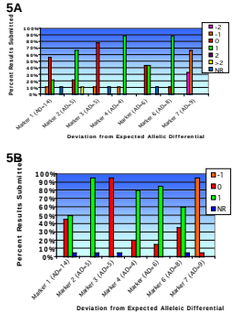

Figure 5: A comparison between slab gel-based and capillary platforms was made by looking at allelic differential (AD) values obtained using either "Long Ranger" polyacrylamide or "POP 4" (Applied Biosystems, Inc.) to represent slab gel, and capillary platforms respectively. NR = No Result. All values rounded to nearest integer. **Fig. 5A** The AD was determined for nine entries in which Long Ranger polyacrylamide was used, and the deviation from expected AD values (shown parenthetically on X-axis) was plotted. **Fig. 5B**. The AD was determined for twenty entries in which POP 4 polymer was used, and the deviation from expected AD values (shown parenthetically on X-axis) was plotted.

Results & Discussion

✓ A total of 50 data submissions were made by study participants in labs from 7 different countries representing 5 different continents. The overall data showed good concordance amongst participants' reported allele sizes of samples sent them, and also in determining which suspect's DNA matched that found at the fictitious crime scene (all but two participants deduced that Diane was the perpetrator).

✓ Submitted allele sizes obtained on slab gel platforms were closer to the expected allele sizes than those obtained on capillary platforms, and the latter platform consistently gave allele sizes smaller than the expected sizes.

✓ When the size difference between the victim's allele and the "hair" sample's allele for any marker was determined – defined here as the allelic differential, or AD – neither instrument platform consistently generated AD values that matched those of expected AD values.

✓ Only two instrument platforms - the ABI 377 DNA Analyzer and the ABI 3100 Genetic Analyzer -reported using more than one type of polymer or matrix through which DNA fragments migrated during electrophoresis. Data collected in this study was obtained from a total of eight different instruments that used a total of nine different polymers.

✓ The choice of in-lane molecular weight standard used by participants did not appear to significantly impact the consistency of AD values obtained.

✓ For the seven markers considered in this study, the use of POP 4 polymer in a capillary platform resulted in less deviation from expected AD values than when Long Ranger polyacrylamide was used in a slab gel platform.

Conclusions

➢ By sending out identical samples to many different laboratories that use a variety of instruments for DNA fragment analysis, the FARG hoped to quantitate the extent of variation in the results reported by study participants. We found that despite AD variability between platforms, almost all study participants were able to correctly determine that Diane was the perpetrator.

➢ Consistency in AD measurements in a single instrument, or within a general platform (e.g., all capillary instruments), would indicate that certain types of genotyping analyses could be carried out using more than one instrument. When all the capillary instrument submissions were grouped (36 entries) and the respective AD values between "victim" & "hair" sample determined, it was observed that the AD's deviated less than 1.0 bp (data not shown). If this pattern of AD variability within all capillary instruments was truly representative of the platform, it is conceivable that some types of genotyping - e.g., those in which the genetic variability being measured was > 2 bp - could be carried out on more than one instrument. However, before embarking on such a project, it would be imperative to test all instruments used in the analysis for AD variability (at the start and end of project) by running a larger number of markers and/or controls of known allele size.

➢ Unforeseen problems encountered during the course of this study included: (1) sample degradation after shipment of samples to at least two participants - resulting in some, or no labeled amplicons being detected by instrument. It is unclear why this occurred, and may have been a factor for why the number of responses we received was less than the number of sample sets mailed out. (2) confusion over the interpretation of genotyping results of our "contaminated" sample - which really was a mixture of Helen and that designated as "crime scene" sample. That is, for this contaminated sample, it was difficult to unambiguously say which allele sizes belonged to Helen, and which were the contaminates. And because we contaminated Helen's sample with crime scene sample (instead of with another suspect's, for example), one could reasonably argue that Helen could have been the perpetrator. (3) The increased sensitivity of newer capillary instrument models resulted in some artifactual peaks being called as alleles - thus highlighting the importance of ensuring that PCR amplicons are artifact free, and that samples analyzed are not degraded.

References

Bartley, D.A., P.S. Adams, L.W. Ballard, Y. Bao, D.A. Bintzler, G. Grills, L. Kasch, P. Morrison, L.Petukhova and C.E. Terrell. (2000). The Effects On Mobility Of Various Protocols, Reagents And Equipment Used In Fragment Analysis. ABRF 2000: From Singular to Global Analyses of Biological Systems, Journal of Biomolecular Techniques 11(1),00 (abstract).

FARG 2000 Study: The Effects on Mobility of Various Protocols, Reagents and Equipment Used in Fragment Analysis: Preliminary Results from the Association of Biomolecular Resource Facilities (ABRF) Fragment Analysis Research Group (FARG) 2000 Study http://www.abrf.org/index.cfm/group.show/FragmentAnalysis.40.htm

Acknowledgments:

Many thanks to all participants who responded to, and participated in this year's study!

The FARG wishes to thank Dr. Shunxue Tang, Department of Crop and Soil Science, Oregon Sate University, and to Doug Bintzler, University of Cincinnati, for supplying the samples used in this study.

Thanks also go out to the College of Business, Oregon State University, for providing a server with which to collect online survey responses and for other (raw) data collection.