ISSR Genotyping of Endangered Plants Using an Optimized Workflow

Applied Biosystems 3500xL Genetic Analyzer GeneMapper® v4.1 Software

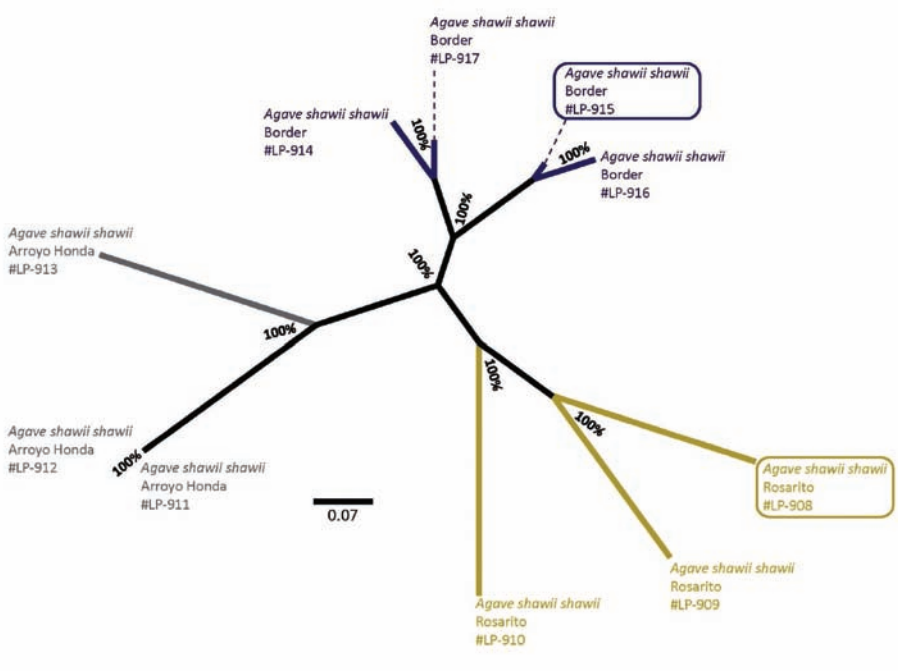


Figure 1. Phylogram Showing Three Distinct Populations of Agave. Individuals collected from Rosarito, Arroyo Honda, and Border are shown in gold, grey, and purple, respectively. Highlighted individuals correspond to the data presented in Figure 4. Nodes in phylogram with posterior probability values above 95% are considered to be informative in Monte Carlo Markov Chain (MCMC) Bayesian analysis (MrBayes).

Introduction

Inter-simple sequence repeat (ISSR) PCR is a fast and inexpensive genotyping technique with a wide range of uses, including the characterization of genetic relatedness among populations. Despite its utility, the ISSR-PCR method has typically suffered from poor reproducibility and time-consuming data analysis. The improved workflow presented here addresses each of these concerns. Reproducibility is improved, and a rapid and automated workflow, from DNA extraction to phylogram (Figure 1), is described. Improvements in the workflow were realized through optimization of thermal cycling using Applied Biosystems® AmpliTaq Gold® 360 Master Mix and by leveraging the customizable, automated AFLP® analysis method in GeneMapper® v4.1 Software. The improved ISSR-PCR workflow (Figure 2) was used to genotype individuals from populations of threatened and endangered plants, with the goal of applying the genotyping data to support ongoing direct conservation efforts in southern California and Baja California.

ISSR-PCR

ISSR-PCR is a genotyping technique based on variation found in the regions between microsatellites. It has been used in genetic fingerprinting [1], gene tagging [2], detection of clonal variation [3], cultivar identification [4], phylogenetic analysis [5], detection of genomic instability [6], and assessment of hybridization [7] in many plant and animal species. The versatility of this genotyping technique makes ISSR useful for researchers interested in diverse fields such as conservation biology and cancer research.

ISSR-PCR uses a single fluorescently labeled primer to target the region between identical microsatellites. An ISSR-PCR primer (Figure 3) comprises three parts: a fluorescent tag, 8 dinucleotide repeat units (or 6 trinucleotide repeat units), and one or more anchor nucleotides designed with a dual purpose: to target the end of a microsatellite region and to prevent primer dimerization. There are over 100 primers developed for use in ISSR techniques [8].

Since ISSRs are dominant markers, the amplified regions in an ISSR-PCR are scored as diallelic. Between individuals within a population, changes in the amplified products can arise through structural changes to the region (insertions or deletions) or the loss of primer binding sites.

The ISSR workflow is faster and requires a lower startup investment than other genotyping methodologies such as amplified fragment length polymorphism (AFLP) and restriction fragment length polymorphism (RFLP). In addition, since there is no need to clone and characterize microsatellites, ISSR is also considerably less expensive and less time-consuming than microsatellite-based genotyping. Several studies have compared AFLP and ISSR results and have found ISSRs preferable because of the reduced number of protocol steps required and the smaller amounts of DNA consumed. Additionally, using capillary electrophoresis (CE) to detect amplified DNA fragments helps deliver significantly higher resolution than traditional agarose gel electrophoresis, thus increasing the amount of information obtained from each experiment.

Despite all of the potential advantages of ISSR-PCR, there are some drawbacks to the technique that highlight the need for improvements in the workflow. The primers that are designed to anneal to the di- or trinucleotide repeats can lack specificity in PCR and are a major contributor to the lack of reproducibility. Also, the lack of complexity of the ISSR primers can lead to nonspecific amplification, particularly if coupled to poor-quality gDNA extraction methods and suboptimal PCR amplification conditions.

In collaboration with Linda Prince of the Rancho Santa Ana Botanic Garden, we have optimized every step of the ISSR workflow on CE during an investigation of the population distribution of the endangered plant *Monardella linoides* and the threatened plant *Agave shawii shawii*.

With the optimized workflow described here, ISSR is especially amenable to high-throughput analysis and only requires primer selection as a prerequisite step.

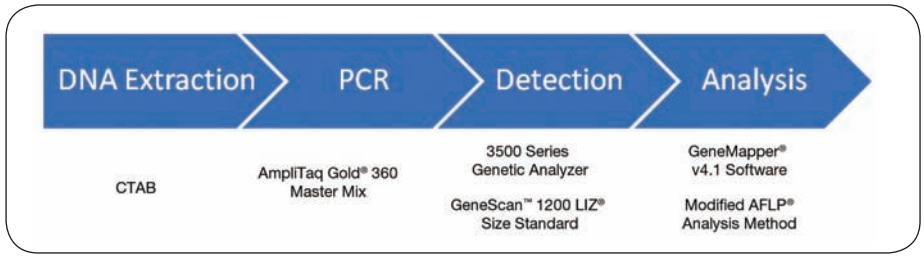


Figure 2. Workflow for an Improved ISSR-PCR Method.

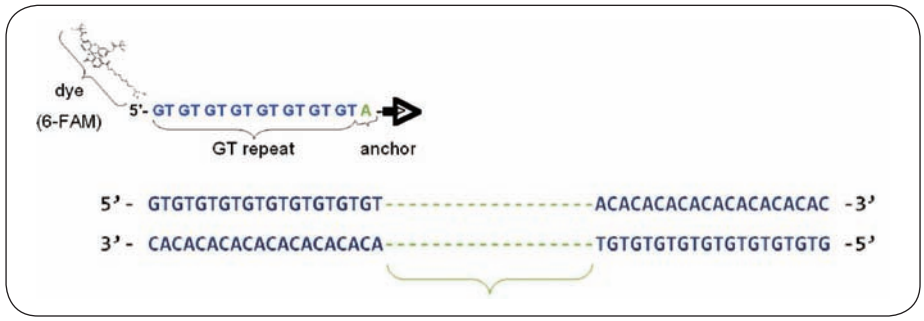


Figure 3. An Inter-Simple Sequence Repeat Region.

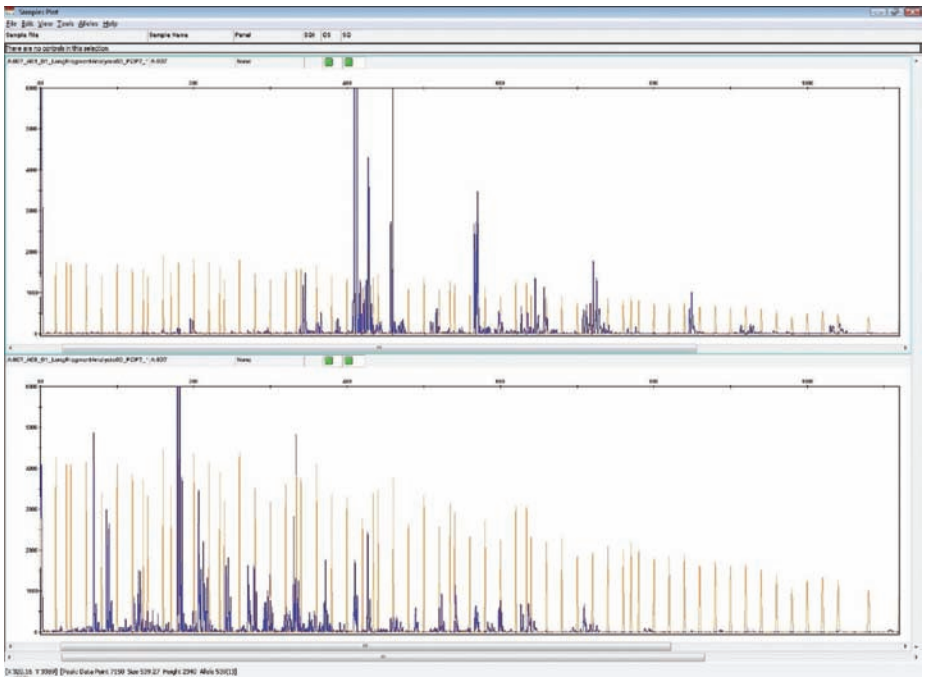


Figure 4. Example of Data Generated Using the 3500 Genetic Analyzer. The individual in the top pane is *Agave shawii shawii*, primed with primer 807, from Rosarito (#908). The individual in the bottom pane is *Agave shawii shawii* from Border (#915). Note the distinct peak patterns of these two individuals; the same two individuals are highlighted in Figure 1.

Workflow Optimization

DNA Extraction

Five gDNA extraction protocols were quantitatively and qualitatively assessed via real-time PCR (7500 Real-Time PCR System) using 18S rDNA as the amplification target. The cetyltrimethyl ammonium bromide (CTAB) gDNA isolation method delivered the highest and most consistent amplification [9] and was effective on leaf tissue from both Agave, which has a mucilaginous cytoplasm, and Monardella, which has a "typical" cytoplasm.

DNA Amplification

Four primers were selected for each species tested. Nine PCR master mixes were tested, and the Applied Biosystems AmpliTaq Gold® 360 Master Mix demonstrated the greatest robustness and consistency when amplifying ISSR targets in both species. Subsequent refinements of the thermal cycling conditions focused on primer annealing temperature and primer annealing time. The VeriFlex™ Blocks functionality of the Veriti® thermal cycler allowed the rapid optimization of the primer annealing temperature.

Fragment Detection on the 3500xL Genetic Analyzer

Detection was performed using POP-7™ polymer, a 50 cm capillary array, and standard run conditions for the long fragment analysis run module with the 3500xL Genetic Analyzer. Because ISSR regions can be long, the GeneScan™ 1200 LIZ® Size Standard was used—sizing with the GeneScan™ 1200 LIZ® Size Standard enabled the detection of fragments up to 1,200 nucleotides in length (Figure 4). This puts CE on par with agarose gels in terms of utilizing long fragments for accurate ISSR analysis. The increased sensitivity of the 3500 system over traditional analysis methods routinely allows the detection of an order of magnitude more peaks, and this increased resolution results

Linda Prince is a researcher at the Rancho Santa Ana Botanic Garden, Claremont, California, USA. Her research interests are varied and include the evolution (analyzed within a phylogenetic framework) and population dynamics of rare plants, with particular interest in elucidating relationships among plant groups that have interesting geographic distributions or complicated morphologies. Her subjects include two very different groups of flowering plants: members of the dicot tea family (*Theaceae*) and two families of the monocot ginger order (the prayer plant family, *Marantaceae*, and canna lilies, *Cannaceae*). The goal of her research is to inform natural resource management decisions made by government agencies and private organizations.

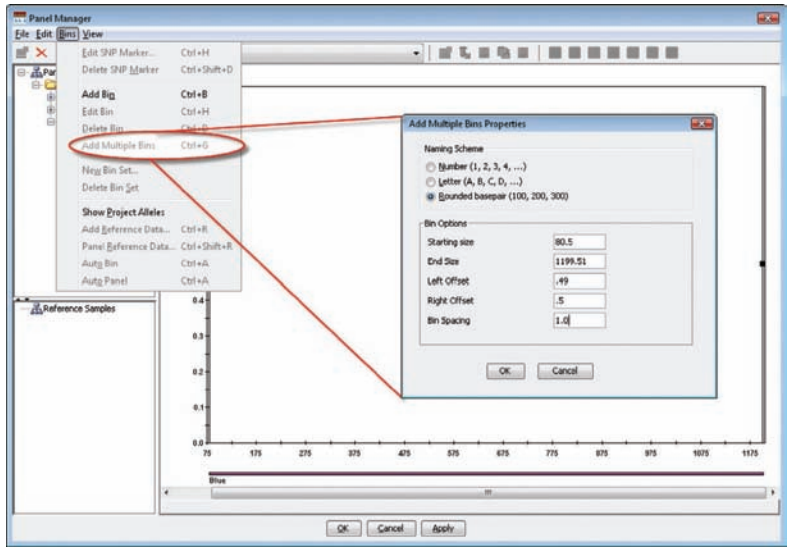


Figure 5A. Creating Multiple ISSR Bins Using the Panel Manager in GeneMapper® Software.

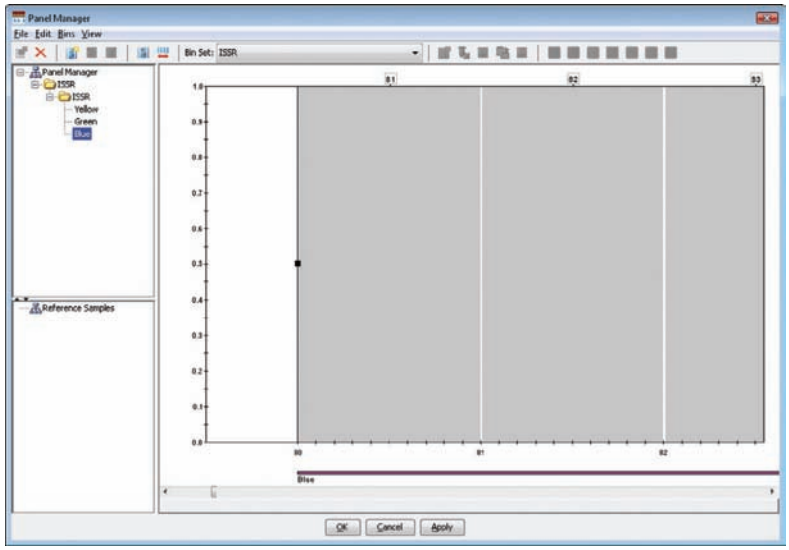


Figure 5B. An Example of Multiple Bins Centered at Whole Base Pairs for the Blue Marker in an ISSR Panel.

in better discrimination between individuals being compared in the populations. The reproducibility of ISSR-PCR fragment migration is reduced when the on-instrument life of the polymer exceeds the recommended seven days. The 3500 Series Genetic Analyzers incorporate RFID (radio frequency identification) tagging of buffers, polymer, and arrays that enables automated electronic tracking of lot number, usage, expiration date, and on-instrument information during analysis (sidebar). The RFID tracking helped to ensure that out-of-date polymer did not compromise data quality.

Fragment Analysis of 80 to 1,200 Base Pairs

Sample files are imported into GeneMapper® software projects created for each primer and species. Files are analyzed with a modified AFLP Analysis Method, using a panel for each dye color (primer) with bins, centered at whole base pairs, one base pair wide covering the entire range of 80 to 1,200 base pairs (Figure 5A and 5B). This method scores each peak above a minimum peak height as an allele and applies a binary label of either 1 or 0 for the presence of a peak in a particular bin (Figure 5C). The scored alleles are exported as a genotypes file. A spreadsheet program was used to assess the consistency of allele calling for four replicate ISSR-PCR reactions for each primer analyzed on the 3500xL. Spreadsheet software was used to calculate the alleles shared between the replicates; only those alleles with 100% concordance were scored as true alleles and used in subsequent phylogenetic analyses. True allele data for each individual for each primer were concatenated into a single list of binary states. The binary data were then analyzed using the phylogenetic package MrBayes [10,11].

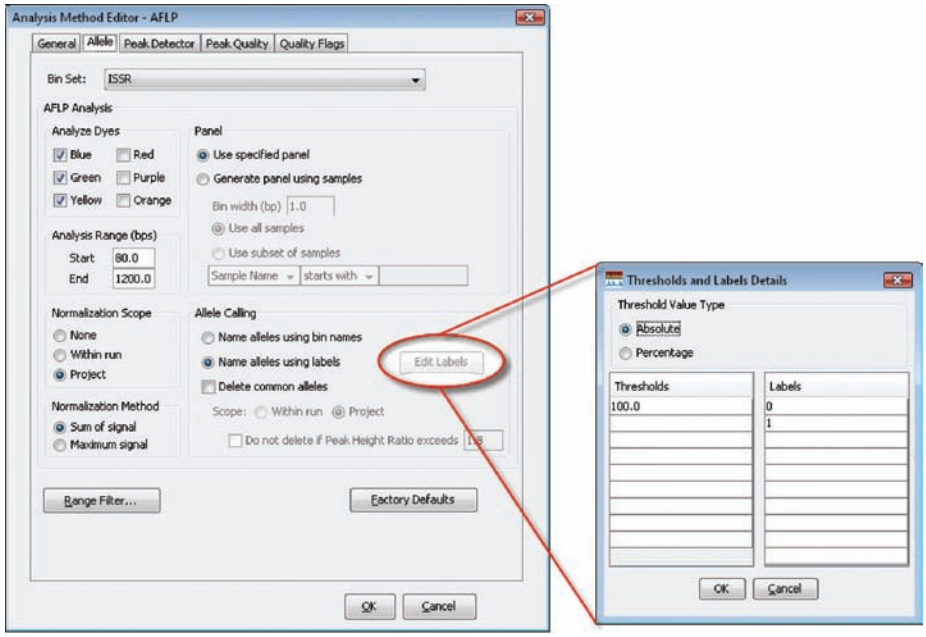


Figure 5C. An Example of the Allele Tab Settings for an ISSR Analysis Method in GeneMapper® Software.

Conclusions

The data presented here represent a workflow that can step rapidly from extraction to phylogram.

The phylograms generated from the ISSR genotyping data indicate with high confidence that three distinct populations of *Agave shawii shawii* (also known as Shaw's Agave) existed (Figure 1). The distinct separation of populations also reflects the life history of this Agave: since *Agave shawii shawii* bloom infrequently and populations bloom asynchronously throughout the early months of the year, this results in low gene flow between populations. *Agave shawii shawii* is a seriously threatened Agave native to southern California and Baja California, Mexico. In the United States, the range of Shaw's Agave consists of only four locations; of these, two locations are native, with a further two locations populated with transplanted stock. Found along 30 miles of coastline in San Diego County, the range of *Agave shawii shawii* is under serious threat due to the physical loss of habitat and individuals by human interference. Habitat loss has been a particularly acute problem during the construction of a border fence between Mexico and the United States, which just eliminated all but one individual of *Agave shawii shawii* from one of the native populations. In addition, genetic depletion is a considerable issue due to the small size of native populations, the distance between populations, and low numbers of reproductive individuals at any one time as a result of infrequent and asynchronous blooming mentioned above. The individual Agaves genotyped in this study were sampled from the location along the border between the United States and Mexico and two populations from Baja California. *Agave shawii shawii* is under threat from habitat loss on both sides of the border, and genotyping individuals within surviving populations by ISSR-PCR will help conservationists assess the genetic variability within and between individuals of native populations in the US and compare those to transplanted populations in the US and native populations in Baja California.

Individuals from three populations of Monardellas were genotyped in this study, two populations of Willowy Monardella (*Monardella linoides viminea*) located 15 km apart, and one population of Flaxleaf Monardella (*Monardella linoides linoides*), located 57 km away. The data represented in the phylogram in Figure 6 confirm the distinction between populations of different subspecies (*linoides viminea* versus *linoides linoides*) with a high degree of confidence. The same data show reasonable confidence in the genetic distinction between populations of Willowy Monardella located only 15 km apart (Carroll Canyon versus Miramar Marine Corps Air Station), whereas the ISSR analysis of the Miramar Marine Corps Air Station samples does not resolve them well, indicating potentially greater relatedness between individuals from this location. Geographically, *M. viminea* is restricted to an area with a width of ~22.5 km and a length of ~11 km in coastal San Diego County [12]. Taken together with the phylogenetic analysis of ISSR-PCR data indicating less genetic diversity of some populations, any potential habitat loss points to difficult management issues for *M. viminea*. In addition to the close geographic proximity, relatedness of the two populations of *Monardella linoides viminea* may be indicative of higher gene flow between these populations, since Monardella are perennial plants that readily disperse

pollen. *Monardella viminea* are federally endangered plants native to southern California and Baja California. ISSR-PCR genotyping of individuals from different populations could be useful for conservation efforts aimed at reestablishing populations at risk, or for selecting Monardella populations to propagate for seedbanking, reintroduction, or restoration efforts [12].

The key benefits of this modified workflow for the analysis of ISSR-PCR fragments are:

- Cycling conditions optimized using the VeriFlex™ Blocks function of the Veriti® thermal cycler alleviated the traditional stumbling block of promiscuous amplification
- Consistent amplification was achieved with AmpliTaq Gold® 360 Master Mix, compared to other master mixes
- CE fragment analysis of long fragments (1,200 bp) using GeneScan™ 1200 LIZ® Size Standard
- The increased sensitivity of the 3500xL Genetic Analyzer over traditional analysis methods routinely enables the detection of more peaks, which helps give better discrimination between individuals
- GeneMapper® software v4.1 effectively assembled the scored alleles and produced a genotypes file that was processed through spreadsheet software and finally through MrBayes phylogenetic software
- The on-instrument tracking of consumables in the 3500 Genetic Analyzers was instrumental in alerting the user that the polymer had exceeded the recommended on-instrument life, helping to ensure that data quality was consistent

In summary, this improved ISSR-PCR analysis workflow enabled significant time savings over traditional methods and helped deliver greater data accuracy. The phylograms generated from the ISSR genotyping data for both species are supported by the life histories of Shaw's Agave and Monardella, and will enable researchers in labs like the RSABG to maximize their conservation efforts.

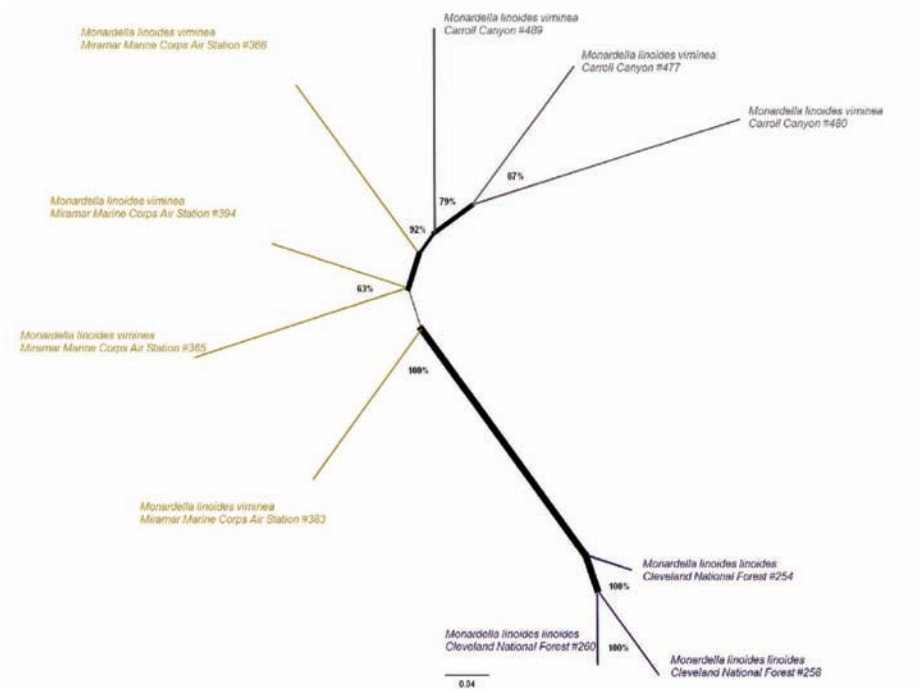


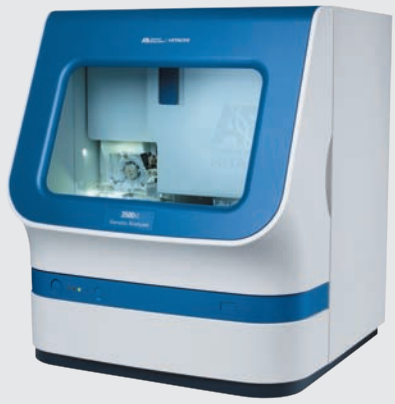
Figure 6. Phylogram Depicting the Relationship Between Three Populations of Monardella, Showing Possible Gene Flow Between the Gold and Grey Populations, Which Are Located 15 km Apart. Gold is Miramar Marine Corps Air Station, while grey is from the Carroll Canyon. About 57 km from Miramar is Cleveland National Forest population (purple), a different subspecies (*linoides linoides*) from the Miramar and Carroll Canyon subpecies (*linoides viminea*). Nodes in phylogram with posterior probability values above 95% are considered to be informative in Monte Carlo Markov Chain (MCMC) Bayesian analysis (MrBayes)

REFERENCES

1. Blair MW, Panaud O, McCouch SR (1999) Inter-simple sequence repeat (ISSR) amplification for analysis of microsatellite motif frequency and fingerprinting in rice (Oryza sativa L.). *Theor Appl Genet* 98:780–792.
2. Ammiraju JSS, Dholakia BB, Santra DK et al. (2001) Identification of inter simple sequece repeat (ISSR) markers associated with seed size in wheat. *Theor Appl Genet* 102:726–732.
3. Leroy XJ, Leon K (2000) A rapid method for detection of plant genomic instability using unanchored-microsatellite primers. *Plant Mol Biol* Rep 18 (2000) 283a–283g.
4. Wang HZ, Wu ZX, Lu JJ et al (2009) Molecular diversity and relationships among Cymbidium goeringii cultivars based on inter-simple sequence repeat (ISSR) markers. *Genetica* 136(3):391–399.
5. Gupta SK, Souframanien J, Gopalakrishna T (2008) Construction of a genetic linkage map of black gram, Vigna mungo (L.) Hepper, based on molecular markers and comparative studies. *Genome* 51:628–637.
6. Anderson GR, Brenner BM, Swede H et al. (2001) Intrachromosomal genomic instability in human sporadic colorectal cancer measured by genome-wide allelotyping and inter- (simple sequence repeat) PCR. *Cancer Res* 61:8274–8283.
7. Wolfe AD, Xiang Q-Y, Kephart SR (1998) Assessing hybridization in natural populations of Penstemon (Scrophulariaceae) using hypervariable inter simple sequence repeat markers. *Mol Ecol* 71:1107–1125.
8. UBC Primer Set 9, 2006 catalog (www.michaelsmith.ubc.ca/services/NAPS/Primer_Sets/Primers_Oct2006.pdf)
9. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull* 19:11–15.
10. Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17:754–755.
11. Ronquist F, Huelsenbeck JP (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574.
12. Prince LM (2009) The relationship of Monardella viminea to closely related taxa based on analyses of ISSRs. USFWS Report P0750003.

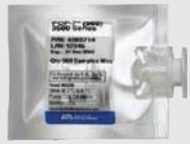
3500 Series Genetic Analyzers

The multicapillary 3500 and 3500xL Systems enable consistently reliable results, critical for demanding fragment analysis applications, and offer:



- Optimized instrument run modules and implementation of the GeneScan™ 600 LIZ® Size Standard v2.0 for improved signal normalization, especially when comparing capillary to capillary, injection to injection, or instrument to instrument within an association of investigators who run samples on multiple identical instruments
- New thermal subsystem design for improved temperature stability
- 3500 Data Collection Software with an intuitive workflow that performs size calls and applies quality control flags to alert the user to off-scale data, broad peaks, and interfering artifacts or "pull-up"

The 3500 Series Genetic Analyzers are available in two throughput options: the 8-capillary 3500 system, and the 24-capillary 3500xL system. Sample analysis is fully automated from the moment each 96- or 384-well plate is placed on the instrument and the run is initiated. Easy-to-use wizards for instrument operation and maintenance enable predictable, hassle-free performance. And, with recent improvements in the instrument run module, the majority of applications can be analyzed on a single configuration of POP-7™ Polymer with a 50 cm capillary array.



The 3500 Series Systems feature simplified and easy-to-install consumables. The Anode and Cathode Buffer Containers are supplied as ready-to-use 1X Genetic Analysis Buffer formulations. The Automated Polymer Delivery System has been improved with the introduction of polymer pouches, available in 960- and 384-sample sizes, which help to significantly reduce setup time. RFID (radio frequency identification) tagging of buffers, polymer, and arrays enables automated electronic tracking of lot number, usage, and expiration date information during analysis.

Also, by employing a single–excitation line solid-state laser, the overall footprint of the 3500 Series System is more compact than previous genetic analyzers and operates using a standard power supply. The smaller footprint and standard power supply mean that 3500 and 3500xL Genetic Analyzers fit in more places and don't require ducting for heat removal.

www.appliedbiosystems.com/3500series

For Research Use Only. Not for use in diagnostic procedures.

Trademarks of Life Technologies Corporation and its affiliated companies: AB Logo™, Applied Biosystems®, GeneMapper®, GeneScan™, Life Technologies™ (Stylized), LIZ®, POP-7™, PrepMan™, Sample-to-SNP™, SeqScape®, Variant Reporter®, Veriflex™, Veriti®. AmpliTaq Gold is a registered trademark of Roche Molecular Systems, Inc. All other trademarks are the sole property of their respective owners.

© 2010 Life Technologies Corporation. All rights reserved. Printed in the USA. 0-090404-r2 02/2010 Publication 106AP31-01

applied biosystems™
part of life technologies™

Headquarters
850 Lincoln Centre Drive | Foster City, CA 94404 USA
Phone 650.638.5800 | Toll Free 800.327.3002
www.appliedbiosystems.com

International Sales
For our office locations please call the division headquarters or refer to our Web site at
www.appliedbiosystems.com/about/offices.cfm