

Procedures and Recommendations for Genotyping

Plant-Microbe Genomics Facility Ohio State University

420 Biological Sciences Building,
484 W. 12th Ave., Columbus OH 43210

Telephone: 614/247-6204

FAX: 614/247-8696

E-mail: pmgf@osu.edu

Internet: www.biosci.ohio-state.edu/~pmgf

Version 2.0

July, 2008

1. Introduction

- 1.1. *Overview:* Genotyping at the Plant Microbe Genomics Facility utilizes the Applied Biosystems 3730 Genetic Analyzer to separate and visualize DNA fragments that have been labeled with a fluorescent dye. The 3730 instrument has 48 capillaries, and requires the use of size standards in each capillary, which provides a resolution of at least 1 base pair. The samples are subsequently analyzed with the GeneMapper® software in order to determine the size and pattern of the DNA fragments. The Facility can perform many genotyping techniques, e.g. microsatellites (SSR), amplified fragment length polymorphism (AFLP), and single nucleotide polymorphism (SNP).
- 1.2. *Pre-experimental Considerations:* Prior to the Facility starting any genotyping work, we request two things: (1) a free face-to-face consultation or, at the minimum, a direct phone/email consultation and (2) a completed genotyping order form for each new ordering of services. This initial communication is necessary to clarify the customer's goals. Furthermore, genotyping is an exacting technique in so far as it requires precision and uniformity to maximize results. In order for the Facility to best serve the customer, it is essential that the customer have a clear understanding of the genotyping service's capabilities and the input required. A critical point that must be made clear is that **optimization for each sample type is almost inevitable** due to the specific characteristics of diverse samples and the requirements of individual customers. There are, however, steps that can be taken during sample preparation to minimize the need for lengthy optimization. The following are guidelines and recommendations, as well as specific information, provided to maximize the value of the Facility's genotyping services. The guidelines have been provided for the benefit of the customer; they are, of course, not strict rules.

2. Sample Guidelines

- 2.1. *Sample Preparation:* Due to the sensitivity of the 3730, sample preparation is extremely important. It is strongly recommended that at least a random subset of samples be checked on a gel prior to submission in order to determine 1) whether the

amplification/assay was successful and 2) the approximate concentration of the product. Although the amount of product to be loaded onto the instrument must be empirically determined, a gel photo can aid in deciding how much dilution is necessary prior to running the samples on the analyzer. DNA Fragment purity is also important, as the 3730 is sensitive to the presence of salts and buffers. Excess salts, proteins, detergents, etc., diminish signal strength and can significantly alter electrophoresis rates. In most cases dilution of the sample in loading solution is sufficient to reduce contaminants and maintain detectable levels of DNA. Sample dilution is occasionally not sufficient, so samples will need to be purified using a standard PCR cleanup method such as dialysis, spin columns or solid phase extraction kits (e.g. Qiagen QIAquick). Purification can be performed by the client or at the Facility (http://www.biosci.ohio-state.edu/~pmgf/services_robotics.html) by one of many different methods.

- 2.2. Sample Dyes:** The 3730 analyzer is currently calibrated to analyze the following sample dyes: FAM (blue), NED (yellow), PET (red), VIC (green), and LIZ (orange; size standard only). In addition, the SNaPshot dye set can also be analyzed (dR110, dR6G, dTAMRA, and dROX). Other dyes, such as ROX (red) and HEX (green) can be used, but they are not recommended because the fluorescent signal they emit may be detectable in additional and/or non-corresponding color channels. The additional peaks are commonly referred to as “pull up”. For example, in addition to red, ROX is also visible in yellow (NED channel) to a lesser degree. We recommend using FAM and/or VIC as your first dyes since they are robust and commonly used.
- 2.3. Size Standards:** The Facility maintains four size standards for use on the 3730, and all use LIZ dye (orange). These are the GS120LIZ (15-120bp), GS500LIZ (35-500bp), GS600LIZ (20-600bp) and GS1200LIZ (20-1200bp). All size standards are manufactured by Applied Biosystems. Size standards with the appropriate range are added to each capillary here at the facility. The PMGF uses the GS600LIZ as the default size standard for AFLP, microsatellite and most other fragment analysis excluding SNaPshot. Customized size standards as well as those from other vendors maybe employed, but Facility staff must approve them (size range and fluorescent dye) prior to usage.
- 2.4. Sample Submission:** To place an order, go to dnaLIMS (http://www.biosci.ohio-state.edu/~pmgf/services_genotyping.html), login and follow the online instructions for submitting genotyping orders through the “96- and 48-well plate requests”. If there are any questions or problems with dnaLIMS please contact the Facility for assistance. Samples should be submitted in a 96-well plate. The plate should be at least labeled on the side with the following information: (1) Order number (from dnaLIMS), (2) last name of client and (3) date of submission. Additional information on the plate maybe useful to facilitate identification. Any 96-well plate will work since samples are transferred to a standard Applied Biosystems plate for analysis on the 3730, although thermal cycler compatible plates are strongly preferred. Plates should be covered with a secure adhesive film or strip caps to prevent cross contamination of wells and evaporation. Plates should be covered with Aluminum foil in order to reduce the light exposure to the samples. For clients not on campus, we strongly suggest placing bubble around the plate(s) and shipping them on dry ice by overnight courier. This is important to prevent the plates from being damaged by shifting dry ice. Maintaining

the samples as frozen is the **only effective** means to prevent evaporation and/or cross contamination of samples during transit. For the 48-well service (1-48 samples) the samples must be placed in the odd columns (1, 3, 5 etc.) in order to be compatible with the 3730 DNA Analyzer, whereas for the 96-well service (49-96 samples) the samples can be placed anywhere on the plate. Samples will be maintained for one month at -20°C , and at the end of this time the samples will be discarded unless the client specifically requests their return.

2.5. Optimization: Although not required, an optimization is **highly** recommended for new customers, especially if the marker(s) analyzed have not been run previously. There are three main components to genotyping optimization: (1) determination of concentration of sample and salts/contaminants, (2) determination of suitable electrophoresis conditions, and (3) multiplexing. First, a dilution series can be run in one plate in order to determine the optimal amount of PCR sample to be used. Second, successive runs can be performed on the same plate with varying electrophoresis conditions to assess fluorescent signal levels. Third, if samples are to be multiplexed or pooled, *i.e.* more than one marker per capillary, there is a need to make sure the markers don't overlap as a result of similar size, or cause pull-up peaks. Pull-up peaks are false peaks that occur when the fluorescent signal from one marker is strong enough to be detected in another channel/color. For example, a DNA fragment labeled with FAM (blue) 200bp in size is so concentrated that the fluorescence is also detected in the VIC (green) channel, and therefore the software reports a false positive DNA fragment at 200bp for the marker labeled with VIC. The expected size ranges of fragments produced by markers labeled with the same dye should be separated by at least 50 bases. It is best that even markers labeled with different dyes not be within the same size range(s). Consultation with the Facility staff is critical in order to determine an optimization protocol.

2.6. Electrophoresis: Typically samples are loaded onto the instrument in 10 μL volumes containing HiDi (formamide), sample and size standard. Routinely, the Facility staff combines the HiDi, samples and size standards immediately prior to running the samples. Although it might be possible to analyze samples that had been previously combined with HiDi and stored at 4C or below. Electrophoresis conditions are quite standardized the amount of sample injected onto the capillary can be easily manipulated to accommodate sample and contamination concentrations to optimize results.

3. Analysis of Data

3.1. Software: Analysis is performed using the GeneMapper® v4.0 software. GeneMapper® is designed to look for peaks in predefined "bins", where each bin represents a possible allele within a marker. For example, allele 100 for marker A in blue has a bin at 100bp +/- 0.5bp. The bins and ranges can be predefined for each marker or determined automatically from your data. Additionally, GeneMapper® v4.0 has the ability to analyze other genotyping applications, *e. g.* single nucleotide polymorphisms, make custom tables, display plots with custom colors as well as many other options. The GeneMapper® software is also available for an additional

cost for customer use in our facility, although training is required which Facility staff can provide, also at an additional cost.

3.2. Data Format. The files produced by the 3730 Genetic Analyzer (*.fsa) have a proprietary format and are provided to the client through dnaLIMS. There are only 4 software options that will access the files: SoftGenetics, GeneMapper®, PeakScanner or dnaLIMS (online viewer). All 4 programs will allow the client to view the electropherogram. Extensive analysis can be done with the first two programs, but they also have a considerable cost. Whereas the last two programs are free but offer very limited analysis, *i.e.* only peak size determination. Results from an in-house analysis with GeneMapper are provided in a choice of formats: Excel table, pdf document, and/or printouts. These files are typically provided through email or postal mail.

4. Common Types of Analysis

4.1. Microsatellites: Microsatellite analysis utilizes short repetitive regions that show a high degree of variability to study alleles of interest, for example locating a gene in a breeding population that confers resistance to a pathogen. For each marker analyzed, the marker name, allele size range, dye and repeat type (*i.e.* dinucleotide, etc.) are needed to perform the analysis. Below are potential problems with these studies. First, marker morphology, *i.e.* pattern of stutter and plus-A peaks, can vary significantly, so often times consultation with the client is required to determine guidelines for determining the alleles given the marker peak pattern. Second, when multiplexing markers, it is critical to insure that markers with identical dyes are sufficiently separated by fragment size. Also, it is strongly recommended that markers with very similar allele sizes are labeled with two different dyes (*i.e.* FAM and PET) that do not tend to cause “pull-up” (see section 2.2).

4.2. Single Nucleotide Polymorphism (SNP): The Facility can use two different proprietary systems from Applied Biosystems to explore SNPs: SNaPshot and SNPLex. The SNaPshot kit can be used to analyze any SNP from any organism. Below are potential problems with this kit. Primer design is the most important part of the SNaPshot assay. It is recommended that primers be tested individually prior to multiplexing in order to determine their efficiency and mobility. Alternatively, the SNaPshot Primer Focus kit (Applied Biosystems, PN 4329538) can be used to determine the mobility of primers to be multiplexed. The analyzed product size is not only dependent upon length of primer but also the molecular weight, dye, and sequence. Therefore it is vital to know where the products will appear so that there is no overlapping of alleles. Analysis of the primer focus kit results can be performed using our 3730 data collection and Genemapper software. The SNPLex system is a high throughput approach that can be used to analyze SNPs in humans and other select organisms. The Facility has the ability to process all of the robotics steps within the SNPLex protocol as well as analyze the data with the 3730 DNA Analyzer.

4.3. Highly Variable Short Sequences (HVSS): Examples of this type of analysis include AFLP (amplified fragment length polymorphisms), T-RFLP (terminal – restriction

fragment length polymorphism) and RAPD (randomly amplified polymorphic DNA). They all measure genetic diversity by utilizing the varied distribution of endonuclease restriction sites and/or priming sites between individuals. There are three aspects of these techniques making optimization essential. First, fragment sizes can have a very broad range, 20 to 1200 bases in size, requiring more exacting electrophoresis. Second, the density of fragments per sample can be high enough to lead to overlapping peaks. Third, within a sample the amount of DNA per fragment can vary by 10,000 fold. So GeneMapper software provides many options for analysis of the data set in terms of size range (bin selection), reference standards, and a broad dynamic range of fragment signal in order to aid analysis of these samples.

4.4. Novel Applications: The Facility is always interested in utilizing the instrumentation in the Facility for either novel applications or adapting established protocols to fluorescent and/or high throughput formats. For example, we have adapted DNA Footprint Analysis so it can be done on the Applied Biosystems 3730 DNA Analyzer. If you have an experiment or an idea in which capillary electrophoresis and fluorescent dyes can be used, then please contact us to discuss your ideas.

(jp, kt and mz; July 2008)