# PROTOCOL FOR MICROSATELLITE GENOTYPING BY UNLINKED MARKERS

**Introduction**

Molecular genotyping of *Plasmodium falciparum* using polymorphic markers is commonly used to distinguish among different clones of malarial parasites and is useful in population genetic studies. Small sequence repeat markers (SSRs) or microsatellite markers (2-3 base pair repeats) have been identified in *P. falciparum* and provide an attractive target for strain identification (ref. 2.1, 2.2). Since the repeat units are small (See Appendix 7.1), these markers typically involve post PCR amplification detection methods using capillary electrophoresis (CE), thus discrimination between two closely sized alleles is much easier. Additionally, CE can be multiplexed using different fluorescently labeled primers and are amenable to high throughput analyses.

## 1.0 Purpose

This protocol is suitable for genotyping unlinked *P. falciparum* microsatellite markers by PCR amplification of genomic DNA extracted from dried blood spots on Whatman 3MM filter paper (Whatman cat# 3030-866). Amplified fragments are then analyzed by capillary electrophoresis on an automated DNA sequencer and sized using suitable software. This protocol has been developed for the Applied Biosystems (ABI) 3730 Capillary Electrophoresis system and may require adaptation for use with other instruments.

## 2.0 References


2.1 Anderson TJ, Su XZ, Bockarie M, Lagog M, Day KP. Twelve microsatellite markers for characterization of Plasmodium falciparum from finger-prick blood samples. Parasitology 1999 Aug;119 ( Pt 2):113-25

2.2 Su XZ, Ferdig MT. Microsatellite analysis in Plasmodium falciparum. Methods Mol.Med. 2002;72:131-6

2.3 GeneMapper® Manual (ver. 4.0), Applied Biosystems, Foster City CA. Available online at: http://mvz.berkeley.edu/egl/resources/manuals/GeneMapper®_4.0_Guide.pdf


## 3.0 Materials: Reagents required for primary and secondary PCR

3.1 Water- Ultra Pure (Molecular Biology Grade) (Quality Biologicals Inc, Cat #: 351-029-101. Gaithersburg, MD)

3.2 Taq Polymerase (Invitrogen, Cat #: 10342-020. Carlsbad, CA)

3.3 10X PCR Buffer: 200 mM Tris-HCl (pH 8.4), 500 mM KCl. PCR Buffer is supplied with the above referenced reagent (Invitrogen, Cat #: 10342-020. Carlsbad, CA)

3.4 $MgCl_2$ (50 mM) is also supplied with the above referenced reagent (Invitrogen, Cat #: 10342-020. Carlsbad, CA)

3.5 dNTPs are purchased at 100 mM each dNTP and combined to produce a 25 mM stock. (Invitrogen Cat#: 10297-018. Carlsbad, CA)

3.6 Sample Templates (DNA): QIAamp (QIAGEN cat # 51161) extracted DNA (see DNA extraction protocol), 1 µl per reaction

3.7 Genomic DNA controls: 1 µL of 3D7 (MR4 Cat #: MRA-102G, Manassas, VA) and HB3 (MR4 Cat #: MRA-155G, Manassas, VA) at a concentration of 25 pg/µL are run on each plate of samples. All genomic controls are purchased from MR-4 (http://www.mr4.org/)

3.8 No Template Controls (NTC). Two water blank (no template) controls are run with each plate of samples

3.9 PCR Primers are purchased from Integrated DNA Technologies (Coralville, IA) with the exception of PET labeled primers which are purchased from Applied Biosystems (Foster City,

CA). Primers are initially reconstituted to 100 µM and then diluted to make 10 µM working stocks suitable for PCR set-up. Primer sequences follow at the end of this protocol (Appendix 7.1).

3.10 GeneScan™ 600LIZ® size standard (Applied Biosystems Cat# 4366589, Foster City, CA)

3.11 Hi-Di™ Formamide (Applied Biosystems Cat# 4311320, Foster City, CA)

## 4.0 Equipment

4.1 Suitable Freezers (non frost-free) and refrigerators (4°C); for long term storage of DNA templates (-80°C); and storage of reagents (-20°C) when not in use

4.2 Calibrated pipettes for precision pipetting steps

4.3 Aerosol Barrier pipet tips

4.4 TempPlate™ 0.2 mL polypropylene 96 well PCR Plates (USA Scientific; cat# 1402-9700, Ocala, FL) (or comparable)

4.5 Adhesive Film for PCR plates (VWR; cat# 60941-122) (or comparable)

4.6 Easy Peel™ Heat Sealing Foil Sheets (Thermo Scientific; cat# AB-0745) (or comparable)

4.7 TempPlate® Sealing Foil Sheets (USA Scientific; cat# 2923-0100) (or comparable)

4.8 Combi Thermosealer (ABGene; model# AB-0384/110)

4.9 Bio-Rad Tetrad-2 Thermocycler (or equivalent)

4.10 Applied Biosystems (ABI) 3100 or 3700 Capillary Electrophoresis Instrumentation (Note: See Introduction and Purpose Sections)

4.11 GeneMapper® (ver 4.0), or equivalent software for data analyses

## 5.0 Safety and Precautions

5.1 Use proper laboratory etiquette when working with reagents and samples. Wear gloves, lab coats and safety glasses when appropriate.

5.2 Ensure that all equipment (pipettes, centrifuges, water baths, etc.) is properly calibrated prior to use.

5.3 Use an ice bath to hold the PCR master mix and/or PCR reagents during PCR master mix preparation.

5.4 Use aerosol barrier pipet tips to mitigate cross contamination of samples during PCR set up and template addition steps.

5.5 The Thermosealer has hot surfaces. Use with caution.

5.6 Take care when applying or removing plate sealant, to prevent cross-contamination of adjacent wells.

5.7 Do not vortex the PCR mix. Mix gently. Use light centrifugation to collect all liquid to the bottom of the PCR plate or tube.

## 6.0 Procedure

6.1 **Notes**:

6.1.1 All Primary PCR reactions are set-up in the pre-PCR area of the laboratory. All heminested PCR reactions are set up in the post-PCR area of the laboratory. **Post amplification PCR products are never taken into the pre-PCR area of the lab.**

6.1.2 Template addition occurs in the pre-PCR set up room.

6.1.3 After template addition, plates are sealed with an appropriate film and thermocycling is conducted on Bio-Rad Tetrad-2 machines. PCR conditions are specified in the Procedure section of this SOP

6.1.4 To avoid cross-contamination of endogenous templates with PCR products, the master mix for nested PCR reactions is aliquoted into suitable PCR reaction plates in the pre-PCR station; however, the primary PCR reaction serving as the template for the nested reaction is aliquoted in the post-PCR station.

6.2 **PCR Set up: Primary and Secondary (Heminested) PCR Reactions**

6.2.1 Primary PCR Reaction. Thaw all of the reagents and keep them cold in an ice bucket while compiling the reactions.

6.2.2 Forward and reverse primers for the primary PCR reactions are stored at a working concentration of 10 µM. The Master Mix for Primary PCR reactions are multiplexed using three primer sets (six primers total) and are set up as follows:

| Components | Stock Concentration | Volume(ul) | Final Concentration |
|---|---|---|---|
| Water | | 6.62 | |
| 10x PCR Buffer | 10 x | 1 | 1 x |
| MgCl2 | 50 mM | 0.6 | 3 mM |
| dNTP | 25 mM | 0.08 | 0.2 mM |
| Primer 1 F | 10 µM | 0.1 | 0.1 µM |
| Primer 1 R | 10 µM | 0.1 | 0.1 µM |
| Primer 2 F | 10 µM | 0.1 | 0.1 µM |
| Primer 2 R | 10 µM | 0.1 | 0.1 µM |
| Primer 3 F | 10 µM | 0.1 | 0.1 µM |
| Primer 3 R | 10 µM | 0.1 | 0.1 µM |
| Taq DNA polymerase | 5 U/µl | 0.1 | 0.5 U |
| Template DNA | | 1 | |
| | | **Total Volume = 10 µl** | |

6.2.3 The following primer sets have successfully been multiplexed in primary PCR reactions:

Multiplex 1: Polyα; PFPK2; and TA81

Multiplex 2: ARA2; TA87; AND TA40

Multiplex 3: TA42; 2490; and TA1

Multiplex 4: TA60; TA109; and PFG377

See Appendix for primer sequences and fluorescent labels. Note that the three different markers in these mixtures each have fluorescent labels with different absorbance and emission spectra allowing discrimination between the markers based on fluorescence as well as size.

6.2.4 Aliquot 9 µl of the Master Mix for Primary PCR into a 0.2 ml thin-walled PCR reaction tube or 0.2 ml polypropylene 96 well PCR Plate (as appropriate, depending upon the numbers of reactions).

6.2.5 Aliquot 1µl of sample DNA to each reaction.

6.2.6 Aliquot 1µl of control DNA to each reaction serving as a positive control (3D7 and HB3 at 25 pg/µl). **Note: 3D7 and HB3 DNA must be run on each plate as downstream analyses depend upon the correct sizing of these controls.**

6.2.7 Aliquot 1µl of Ultra Pure water to each reaction serving as a negative control (No Template Control (NTC)). **Note: an NTC must be run on each plate as downstream analyses requires the evaluation of background fluorescence in these controls. These controls are also required to detect cross-contamination during PCR set-up.**

6.2.8 The thermocycling conditions for the Primary reaction are as follows:

Program: Hemi-1
94°C 2 min

94°C 30 sec
42°C 30 sec
40°C 30 sec
65°C 40 sec
} **x 25 cycles**

65°C 2 min

6.2.9 PCR products from the primary reaction may be held at 4°C before proceeding to the next PCR step or stored at -20°C (See Appendix section 7.2).

6.2.10 Secondary PCR Reaction. **Note: The second (hemi-nested) PCR reactions are run individually for each marker such that each of the four multiplexed PCR reactions described in section 6.2.4 of this protocol will give rise to three separate PCR reactions.**

6.2.11 Thaw all of the reagents and keep them cold in an ice bucket while compiling the reactions.

6.2.12 Forward and reverse primers for the heminested (secondary) PCR reactions are stored at a working stock concentration of 10 µM. The Master mix preparations are as follows.

**Heminested (secondary) PCR Reaction Master Mix:**

| Components | Stock Concentration | Volume(ul) | Final Concentration |
|---|---|---|---|
| Water | | 10.23 | |
| 10x PCR Buffer | 10 x | 1.5 | 1 x |
| MgCl2 | 50 mM | 0.75 | 2.5 mM |
| dNTP | 25 mM | 0.12 | 0.2 mM |
| Primer F | 10 µM | 0.6 | 0.4 uM |
| Primer R | 10 µM | 0.6 | 0.4 uM |
| Taq DNA polymerase | 5 U/µl | 0.2 | 1.0 U |
| PCR Products from corresponding multiplexed Primary PCR Reaction | | 1 | |
| | | **Total Volume = 15 ul** | |

6.2.13 The DNA template used for the heminested reactions are from the corresponding multiplexed Primary PCR reaction.

6.2.14 The thermocycling conditions for the secondary reaction are as follows:

Program: Hemi-2

| | | |
|---|---|---|
| 94°C | 2 min | |
| 94°C | 20 sec | **x 25 cycles** |
| 45°C | 30 sec | |
| 65°C | 30 sec | |
| 65°C | 2 min | |

## 6.3 Preparation of Sample Plates for Capillary Electrophoresis

6.3.1 **Note:** FAM and HEX-labeled amplicons are detected more readily than PET-labeled amplicons on the capillary instrument

6.3.2 Dilute and mix the single PCR reactions from the secondary (heminested) PCR reactions. The ratio of the three PCR products in each mixture can be adjusted based on variation in peak heights caused by differences in strength of fluorescence and DNA quality. In general, mixtures of PCR product are diluted 1:10 with Ultra-Pure water.

The following mixtures and dilutions are optimized for the ABI 3730 system and should be used as a guideline:

6.3.2.1 Multiplex 1: Poly α- FAM (1 µl) + PFPK2- HEX (1 µl) + TA81- PET (1 µl) + 27 µl Ultra-Pure water

6.3.2.2 Multiplex 2: ARA2- FAM (1 µl) + TA87- HEX (1 µl) + TA40- PET (3 µl) + 45 µl Ultra-Pure water

6.3.2.3 Multiplex 3: TA42- FAM (1 µl) + 2490- HEX (1 µl) + TA1- PET (3 µl) + 45 µl Ultra-Pure water

6.3.2.4 Multiplex 4: TA60- FAM (6 µl) + TA109- HEX (5 µl) + TA40- PET (3 µl) + 126 µl Ultra-Pure water

6.3.3 Thaw enough Hi-Di™ Formamide (Applied Biosystems Cat# 4311320, Foster City, CA) for 8.5 µl per reaction.

6.3.4 In a separate tube mix 0.5 µl per reaction GeneScan™ 600LIZ® size standard (Applied Biosystems Cat #: 4366589) with 8.5 µl Hi-Di™ Formamide per reaction. Mix well.

6.3.5 Dispense 9 µl of Size Standard+HiDi mixture to each well.

6.3.6 Add 1 µl of sample mixture to each well.

6.3.7 Samples are then brought to the UMB Biopolymer Lab for processing on an Applied Biosystems capillary electrophoresis instrument. A formatted Excel file with sample names is also provided at this time to facilitate proper sample ID and data integrity.

## 6.4 Reporting Results using GeneMapper® ver 4.0

6.4.1 CE is performed per manufacturer's protocol on an Applied Biosystems capillary electrophoresis instrument. Raw data is then obtained and analyzed using GeneMapper® software (ver 4.0) from Applied Biosystems (Foster City, CA).

6.4.2 All raw capillary data files are imported into a new project within the Genemapper™ application and analyzed using the default settings for microsatellite analyses.

6.4.3 Size standards are then evaluated after the initial analyses and edited as appropriate to accommodate off-scale data or sizing errors.

6.4.3.1 Samples must be “re-analyzed” with the adjusted Size Standard.

6.4.4 The samples, controls, and NTCs are also evaluated for background fluorescence or artifacts from either PCR or CE.

6.4.5 **Controls** (*P. falciparum* 3D7 and HB3), which are run on each plate, are also evaluated for the presence of “companion peaks” or “stutter” to derive a peak “morphology” (Appendix 7.3).

6.4.5.1 These peaks are considered to be **artifacts** of the PCR process; are always associated with major peaks, and are **disregarded** in the analyses.

**6.4.5.1.1** **Important**: when evaluating multiple clone infections, the height and spacing of the companion peaks will often determine which peaks are included in the analyses and which are excluded as artifact.

6.4.5.2 The size of the alleles for 3D7 are used to correct run to run variation among CE runs and bin alleles appropriately. **Note: these sizes will differ slightly from the sizes expected from DNA sequencing or those sizes found *in silico* due to subtle variations in capillary electrophoresis. However only minor variation (±0.25 bp) is to be expected between CE runs on the same instrument with the same control DNA.**

6.4.6 **Raw data analyses**: several scenarios are observed that presented challenges to accurate allele scoring and sizing. (See Appendix 7.4)

6.4.6.1 Unless the sample has been contaminated with different clones, the number of alleles present for a given marker will be reflective of the number of parasite clones contributing to the infection.

6.4.6.2 The simplest scenarios for genotyping are single clone infections. In these instances, the convention for scoring data is similar to scoring controls.

6.4.6.3 Multiple clone infections (2 or 3 clones) presented more challenges, but are still relatively simple, especially where allele sizes are spread out so that peak morphology could easily be ascertained.

6.4.6.4 In samples where the allele sizes are clustered and sized at 3 base pair intervals (the size of the microsatellite motif), the companion peak will actually be under a neighboring allele. Consistent genotyping can still be achieved by taking into consideration the additive affect in peak height fluorescence of the companion peak.

6.4.6.5 An increase in the number of clones (4 or more) contributing alleles to an infection will greatly affect the ability of this system to ascertain the number of alleles present in any given infection and will confound genotyping efforts.

6.4.6.6 Since the GeneMapper® software allows the suppression of one or more dyes in the electropherogram (see GeneMapper® Manual), it may still be possible to genotype such samples by selectively suppressing individual dyes in the electropherogram and scoring each marker’s alleles separately.

6.4.6.7 In rare instances, the peak heights and morphology of minority component alleles in multiple clone infections coupled with the presence of off-scale data from majority and background noise will preclude accurate genotyping. Additionally,

some samples will exhibit excessive artifacts of PCR called stutter (See Appendix 7.5).

6.4.7 After scoring raw data, allele sizes are normalized against the 3D7 and HB3 control strains to take variations in capillary electrophoresis runs into account. “Bins” are constructed to round the allele sizes into 3 base pair intervals for data analyses.

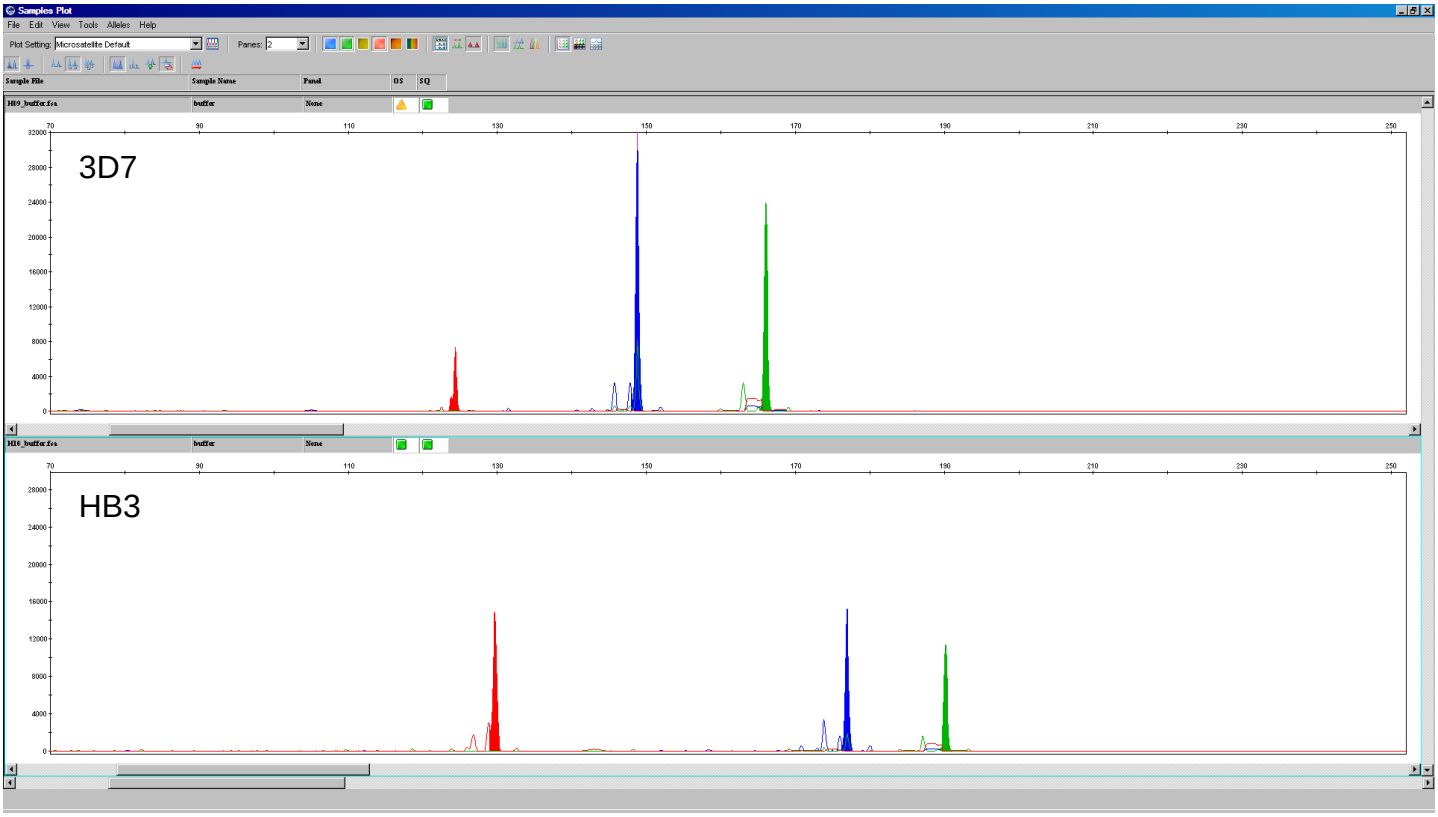
## 7.0 Appendices

7.1 Primers used to amplify Neutral Microsatellite Markers in *P. falciparum*. For each locus the first and second primers are used in the primary reaction and the second and third primers are used in the secondary reaction. Primers listed third are internal to the other two and are end-labeled. For TA-40, an error was noted in the labeled primer of the published sequence (Anderson, personal communication). This primer was redesigned and the sequence differs from that presented in Anderson *et al* (1999).

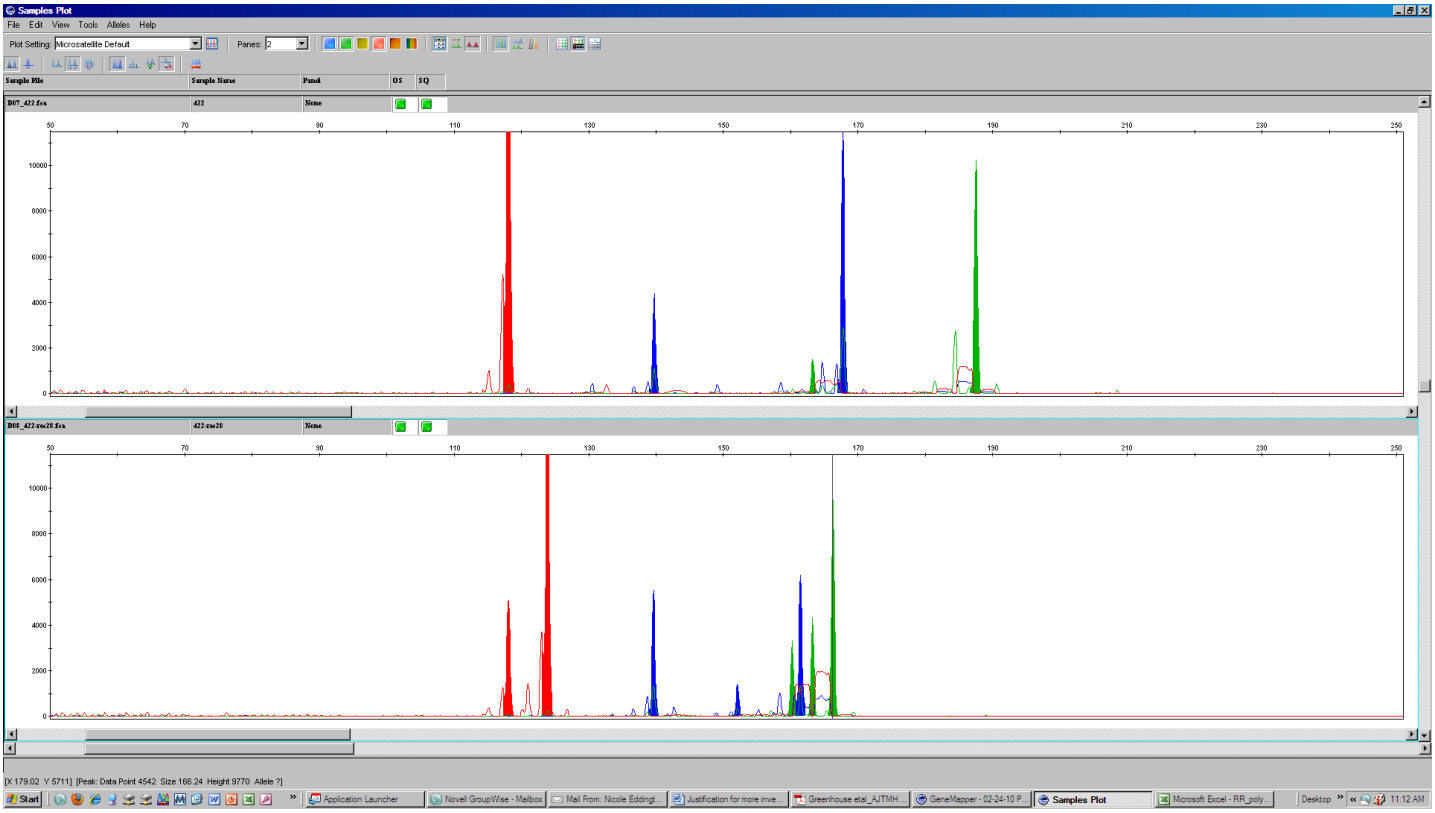
| Multiplex | Locus (Anderson 1999) | primer 5' to 3' | Fluorescent Label | size range from Anderson 1999 | Size in 3D7 reference genome |
|---|---|---|---|---|---|
| 1 | Poly α-R | ATCAGATAATTGTTGGTA | FAM | 114-201 | 151 bp |
| | Poly α-F | AAAATATAGACGAACAGA | | | |
| | Poly α-3(IR) | GAAATTATAACTCTACCA | | | |
| 1 | PFPK2-3R | CCTCAGACTGAAATGCAT | HEX | 159-192 | 172 bp |
| | PFPK2-F | CTTTCATCGATACTACGA | | | |
| | PFPK2-R | AAAGAAGGAACAAGCAGA | | | |
| 1 | TA81-3F | GAAGAAATAAGGGAAGGT | PET | 112-142 | 122 bp |
| | TA81-R | TTTCACACAACACAGGATT | | | |
| | TAA81-F | TGGACAAATGGGAAAGGATA | | | |
| 2 | ARA2-3(F) | GTACATATGAATCACCAA | FAM | 63-90 | 73 bp |
| | ARA2-R | GCTTTGAGTATTATTAATA | | | |
| | ARA2-F | GAATAAACAAAGTATTGCT | | | |
| 2 | TA87-3F | ATGGGTTAAATGAGGTACA | HEX | 90-126 | 100 bp |
| | TA87-R | ACATGTTCATATTACTCAC | | | |
| | TA87-F | AATGGCAACACCATTCAAC | | | |
| 2 | TA40 Rev-1 | GAAATTGGCACCACCACA | PET | 217 | 217 bp |
| | TA40 For | AAGGGATTGCTGCAAGGT | | | |
| | TA40 Rev-2 | CATCAATAAAATCACTACTA | | | |
| 3 | TA42-3F | ACAAAAGGGTGGTGATTCT | FAM | 182-251 | 201 bp |
| | TA42-R | GTATTATTACTACTACTAAAG | | | |
| | TA42-F | TAGAAACAGGAATGATACG | | | |
| 3 | 2490-3R | ATGATGTGCAGATGACGA | HEX | 78-93 | 84 bp |
| | 2490-F | TTCTAAATAGATCCAAAG | | | |
| | 2490-R | TAGAATTATTGAATGCAC | | | |
| 3 | TA1-3(F) | CTACATGCCTAATGAGCA | PET | 159-204 | 186 bp |
| | TA1-R | TTTTATCTTCATCCCCAC | | | |
| | TA1-F | CCGTCATAAGTGCAGAGC | | | |
| 4 | TA60-F | CTCAAAGAAAAATAATTCA | FAM | 69-99 | 87 bp |
| | TA60-R | AAAAAGGAGGATAAATACAT | | | |
| | TA60-3(IF) | TAGTAACGATGTTGACAA | | | |
| 4 | TA109-3F | TAGGGAACATCATAAGGAT | HEX | 154-223 | 174 bp |
| | TA109-R | CCTATACCAAACATGCTAAA | | | |

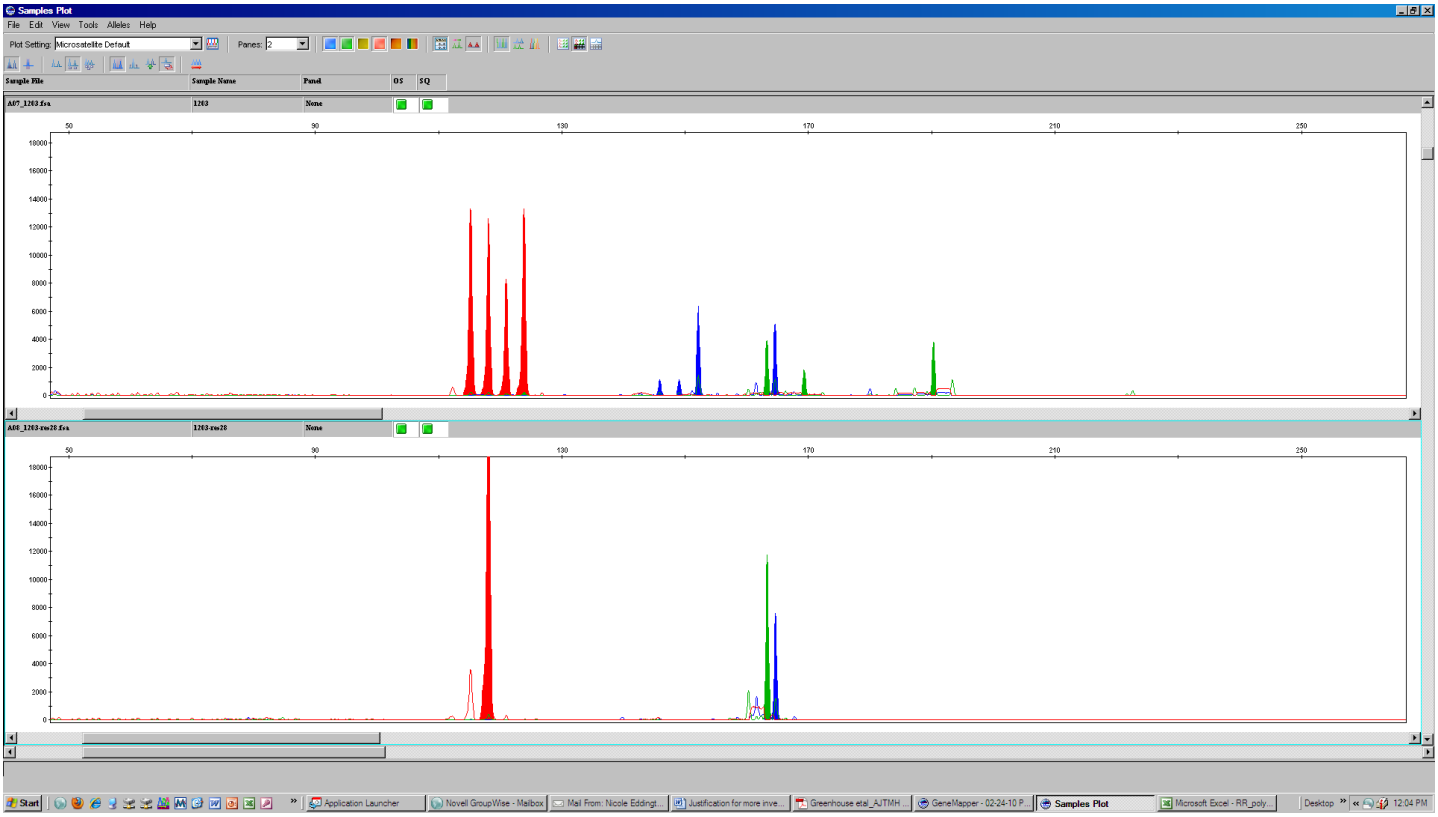
| HB3 | | | | | |
|---|---|---|---|---|---|
| | TA109-F | GGTTAAATCAGGACAACAT | | | |
| **4** | PFG377-3R | TTATGTTGGTACCGTGTA | PET | 89-113 | 98 bp |
| | PFG377-F | GATCTCAACGGAAATTAT | | | |
| | PFG377-R | TTATCCCTACGATTAACA | | | |

7.2 Electropherograms for 3D7 and HB3. Peak morphologies are depicted for two control strains (3D7 and HB3). The markers shown are Poly-α, PFPK2, and TA81 (6-FAM (blue); HEX (green); and PET labeled (red) respectively. The alleles that were scored are filled in whereas the companion peaks are not filled in.



## 7.3 Multiple clone infections



## 7.4 **Troubleshooting:** Stutter associated with a PET labeled microsatellite marker (TA40)