# Development of PowerSeq™ Systems for Forensic Identification Using Next Generation Sequencing

**Lotte Downey, Jaynish Patel, Spencer Hermanson, Leta Steffen, Cynthia Sprecher, Robert S. McLaren, and Douglas R. Storts.**

**Promega Corporation, 2800 Woods Hollow Road, Madison, WI 53711, U.S.A.**

**AAFS 2015, Poster # B4**

## Introduction

STR-based systems validated for forensics generate fluorescently-labeled amplicons that are separated by capillary electrophoresis (CE). Allele calling is based on the size and fluorescent label of the amplified alleles. Next generation sequencing (NGS) based methods detect the actual sequence.

NGS offers several advantages over CE methods:

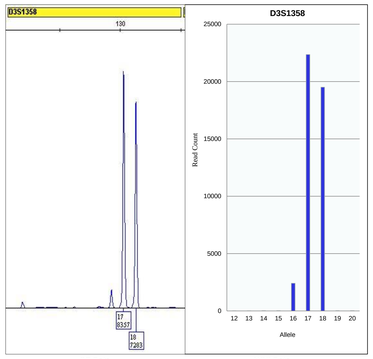
- identification of sequence polymorphisms
- more accurate quantitation of alleles
- the ability to analyze more loci in one reaction
- the ability to make all the amplicons smaller

The ability to determine the sequence of each allele in a sample makes it possible to distinguish two DNA samples which have the same number of repeats at a given locus. In the example shown, two DNA samples possess alleles containing 17 and 18 repeats at D3S1358. These would not be distinguishable by CE. NGS shows "DNA 2" has a G to A transition in one of its repeats in the 17 allele, making it possible to discriminate between these two DNA samples.

The read count, displayed as a histogram, is analogous to the peak height generated on a CE. Because the read count is more quantitative than peak height, it should allow more reliable mixture interpretation, especially when used in combination with sequence polymorphisms.

The ability to assay more loci can enable generation of autosomal STR, Y-STR, single nucleotide polymorphism, and mitochondrial DNA (mtDNA) data from the same sample in a single experiment.

Smaller amplicons allows for more data generation from degraded DNA samples.




DNA 1 CTTGCATGTATCTA[TCTG]$_2$[TCTA]$_{13}$TGAGACAGGG 17
CTTGCATGTATCTA[TCTG]$_2$[TCTA]$_{14}$TGAGACAGGG 18

DNA 2 CTTGCATGTATCTA[TCTG]$_2$TCTA[TCTA]$_{13}$TGAGACAGGG 17
CTTGCATGTATCTA[TCTG]$_2$[TCTA]$_{14}$TGAGACAGGG 18

## PowerSeq™ Systems

The PowerSeq™ Systems consists of:

- 5X PowerSeq™ Master Mix
- 5X Multiplexed Primer Pair Mix
- Control DNA (2800M)

All three systems use the same PCR cycling conditions. Amplicons for each locus and mtDNA analysis were designed to be in a range of 140-300bp. PowerSeq™ Systems cycling time is a little over 1 hour.

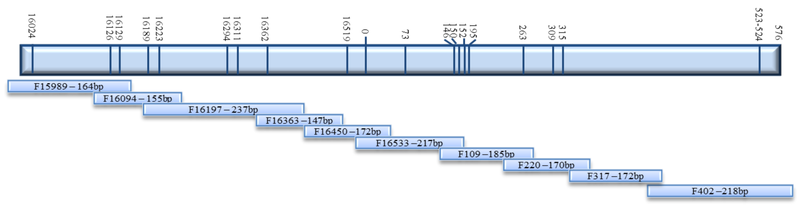
96ºC 1min
94ºC 10sec
59ºC 1min
72ºC 30sec
} 30 Cycles
60ºC 10min
4ºC ∞

PowerSeq™ Systems Cycling Conditions

PowerSeq™ Auto includes 23 STR loci and Amelogenin.

The PowerSeq™ Mito System generates 10 small amplicons (adapted from Eichmann and Parson) covering the control region of the mitochondrial genome.

PowerSeq™ Auto/Mito/Y combines both sets of amplicons in one multiplex plus 23 Y-STR loci.




Schematic of PowerSeq™ Mito amplicons across control region with mutational hotspots indicated.

## Library Preparation Workflow Using Illumina's TruSeq® DNA PCR-Free Sample Prep Kit

Multiplexed PCR amplification,≤ 500pg input DNA

↓

Qiagen® MinElute purification of amplification products, elute using 30µl Elution Buffer (EB) – Low Ethanol carryover critical for consistent library yields, quantify using NanoDrop®

↓

End repair, ~300-400ng of purified amplification product

↓

Qiagen® MinElute purification of end repaired products, elute using 20µl resuspension buffer

↓

A-tail & adaptor ligation steps

↓

Purify library using SPRI beads – Purification reactions performed at 1.1X final ligation volume (i.e. 1st purification step uses 47µl and 2nd uses 55µl), elute in 20µl resuspension buffer

↓

Quantify libraries using Kapa Biosystems Library Quantification kit and normalize to 4nM using Library Dilution Buffer (10mM Tris.HCl pH 8.0; 0.1% Tween 20)

↓

Pool libraries using equal volumes of each library and denatured according to manufacturer's protocol and spike in PhiX (10pM) ~15% of total volume.

↓

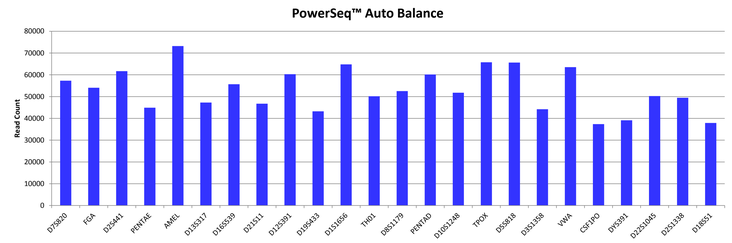
Load and sequence using MiSeq® V2 Chemistry, 1X 250-300 Cycles

↓

Analyze data using STRait Razor or commercial software (NextGENe®)

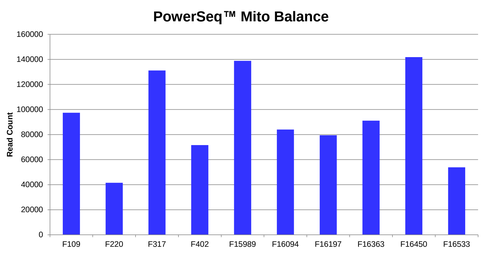
## PowerSeq™ Auto System

PowerSeq™ Auto includes 23 STR loci (D3S1358, D1S1656, D2S441, D10S1248, D13S317, D16S539, D18S51, D2S1338 CSF1PO, Penta D, Penta E, TH01, vWA, D21S11, D7S820, D5S818, TPOX, D8S1179, D12S391, D19S433, D22S1045, FGA, DYS391) and Amelogenin.

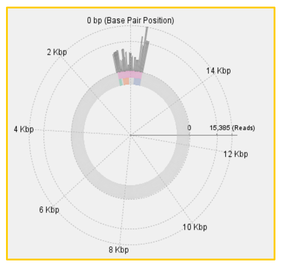



Histogram representation of read count data for the same library sequenced using MiSeq v2 chemistry, 1X 260 cycles.

## PowerSeq™ Mito System



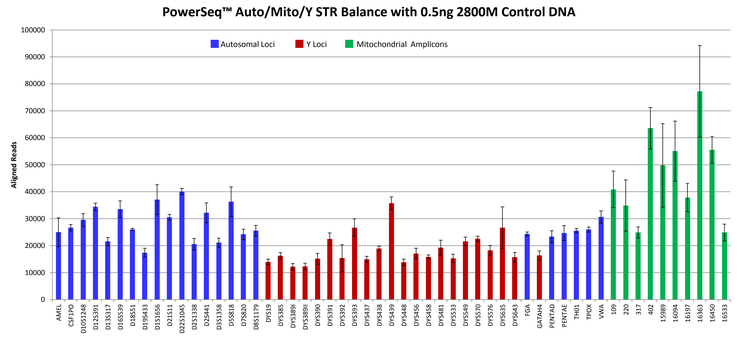

Histogram representation of read count data for 2800M DNA library sequenced using MiSeq® v2 chemistry, 1X 260 cycles.




IGV Visualization of coverage depth over mitochondrial control region.

## PowerSeq™ Auto/Mito/Y System

PowerSeq Auto/Mito/Y combines the autosomal loci from PowerSeq Auto and the 10 amplicons from the mitochondrial control region. In addition, 23 Y STR loci are included: DYS19, DYS385a/b, DYS389I, DYS389II, DYS390, DYS391, DYS392, DYS393, DYS437, DYS438, DYS439, DYS448, DYS456, DYS458, DYS481, DYS533, DYS549, DYS570, DYS576, DYS635, and DYS643.




Average number of aligned reads per locus from three independent experiments using MiSeq v2 chemistry, 1X 260 cycles. To control for sample loading, aligned reads per locus have been normalized to average of total aligned reads.

## Summary

- STR amplicons reduced to 140-300bp for sequencing on the Illumina MiSeq platform
- Single master mix (PowerSeq™ 5X Master Mix) for all PowerSeq™ Systems
- PowerSeq™ Auto – A 24-Plex kit for analyzing autosomal STR's , amelogenin and DYS391
- PowerSeq™ Mito – Sequencing of the mtDNA control region (HV1 and HV2)
- PowerSeq™ Auto/Mito/Y – Configured for the simultaneous analysis of 22 autosomal STRs, amelogenin, 23 Y STRs, and the control region of the mitochondrial genome

## References

D.M. Bornman, M.E. Hester, J.M. Schuetter, M.D. Kasoji, A. Minard-Smith, C.A. Barden, et al., Short-read, high-throughput sequencing technology for STR genotyping, BioTechniques. (2012) 1-6.

D.H. Warshauer, D. Lin, K. Hari, R. Jain, C. Davis, B. Larue, et al., STRait Razor: a length-based forensic STR allele-calling tool for use with second generation sequencing data, Forensic Sci. Int. Genet. 7 (2013) 409-417.

C. Eichmann, W. Parson., "Mitominis": multiplex PCR analysis of reduced size amplicons for compound sequence analysis of the entire mtDNA control region in highly degraded samples, Int J Legal Med. 122 (2008) 385-8.

## Acknowledgements

Thank you to our collaborators.