Development of a Novel and Sensitive DNA Analysis Multiplex Based on INNUL Markers for Highly Degraded Forensic DNA Samples

Sudhir K. Sinha, Ph.D.

Anne H. Montgomery, M.S. [1], Gina Pineda, M.S. [1], Robyn Thompson, M.S. [1], Jonathan King, M.S.[2], Bobby L LaRue, Ph.D. [2], Jianye Ge, Ph.D.[2], Ranajit Chakraborty, Ph.D.[2], and Bruce Budowle, Ph.D.[2]

[1]InnoGenomics Technologies, LLC; 1441 Canal Street, Suite 307; New Orleans, LA 70112

[2]Institute of Applied Genetics; Department of Molecular and Medical Genetics; University of North Texas Health Science Center; 3500 Camp Bowie Blvd., Fort Worth, Texas 76107

InnoGenomics
Innovation in Forensic Genetics

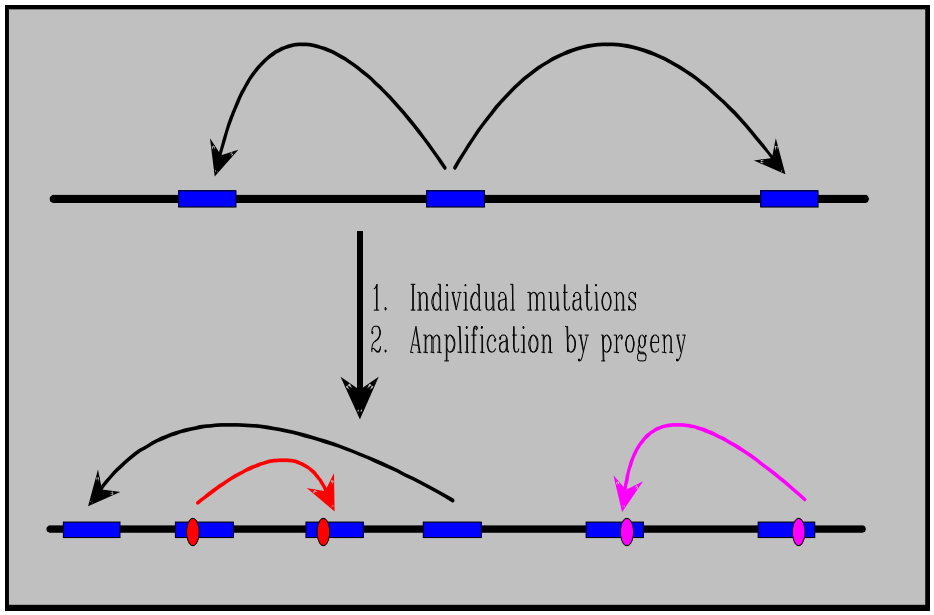
Jumping Genes

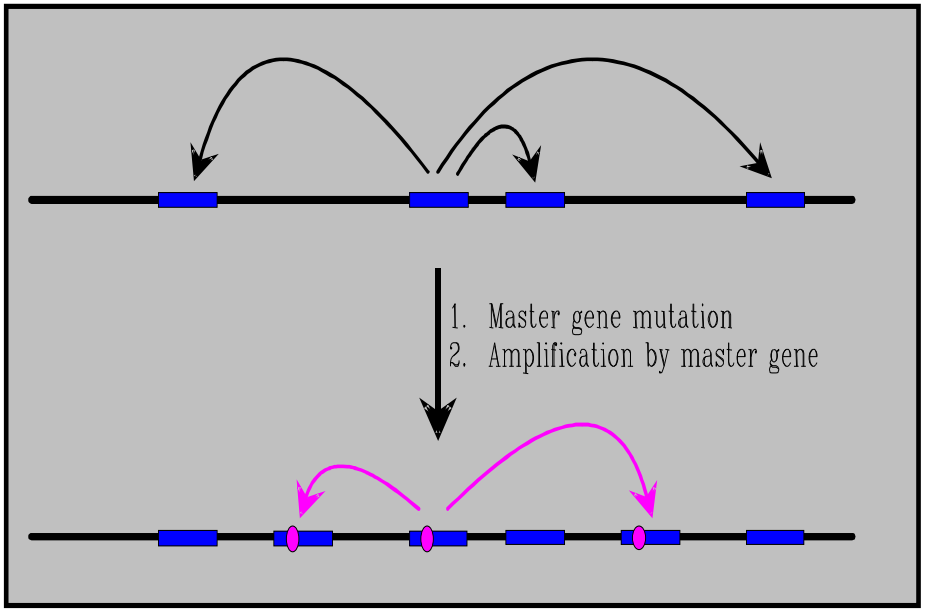
- SINE: Short Interspersed Elements: ALU 300 bp
- LINE: Long Interspersed Elements: Could be up to 6000 bp
- SVA : Composite of ALU-like Element-VNTR- SINE R (retroviral component)
- Total content in the genome 30-40%

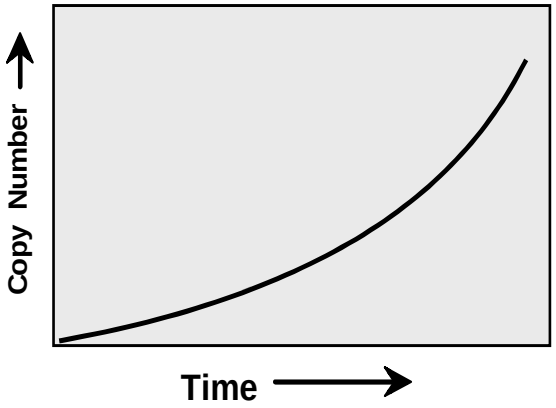
Alu Elements

- 300 bp long
- RNA polymerase III transcribed
- 3' oligo dA-rich tail
- found only in primates
- 500,000 copies in human genome
- most amplification 40 mya
- similar copy # in other primates
- dimer-like structure
- poorly transcribed

Transposon vs. Master Gene Models

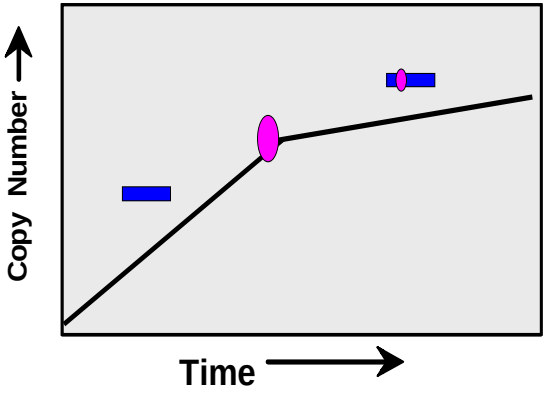




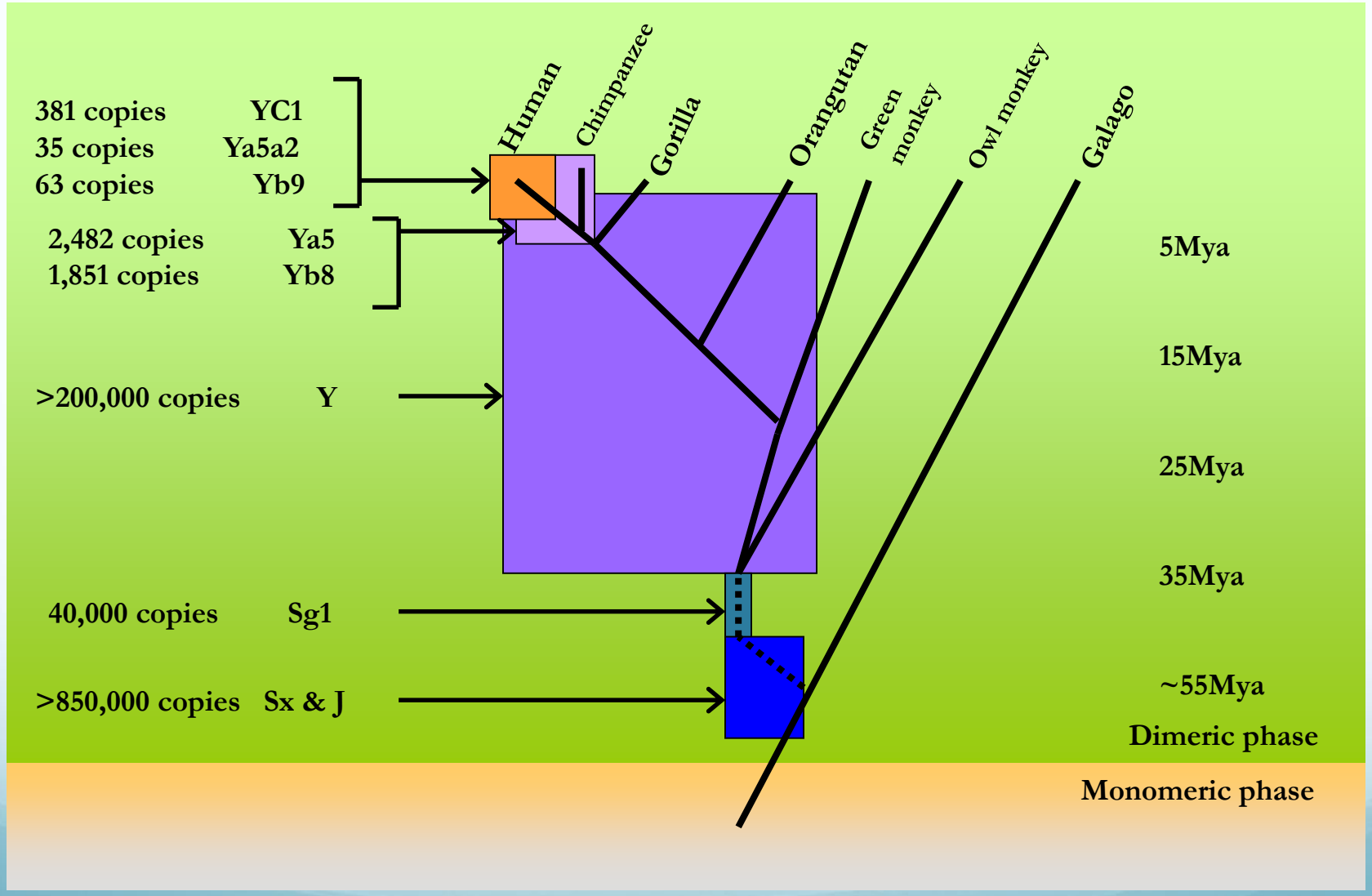


Differences:

- Amplification rate
- Subfamily structure
- Timing of subfamilies



Mobile Elements

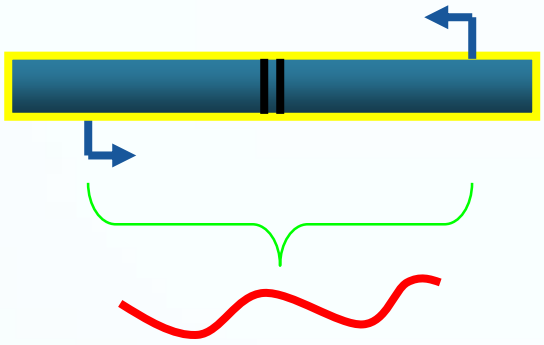


Batzer and Deininger (2002) *Nature Genetics*

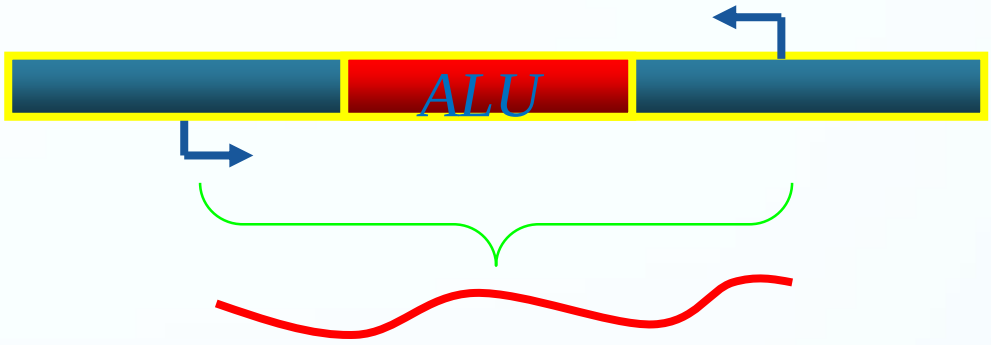
Properties of Mobile Element Insertions

1. Stable polymorphisms that are not deleted
2. Known ancestral state
3. Identical by descent
4. Population specific alleles
5. Neutral genetic loci
6. Parallel independent insertion is essentially zero

Old Alu Multiplex Design



Amplification product
~100bp (N)

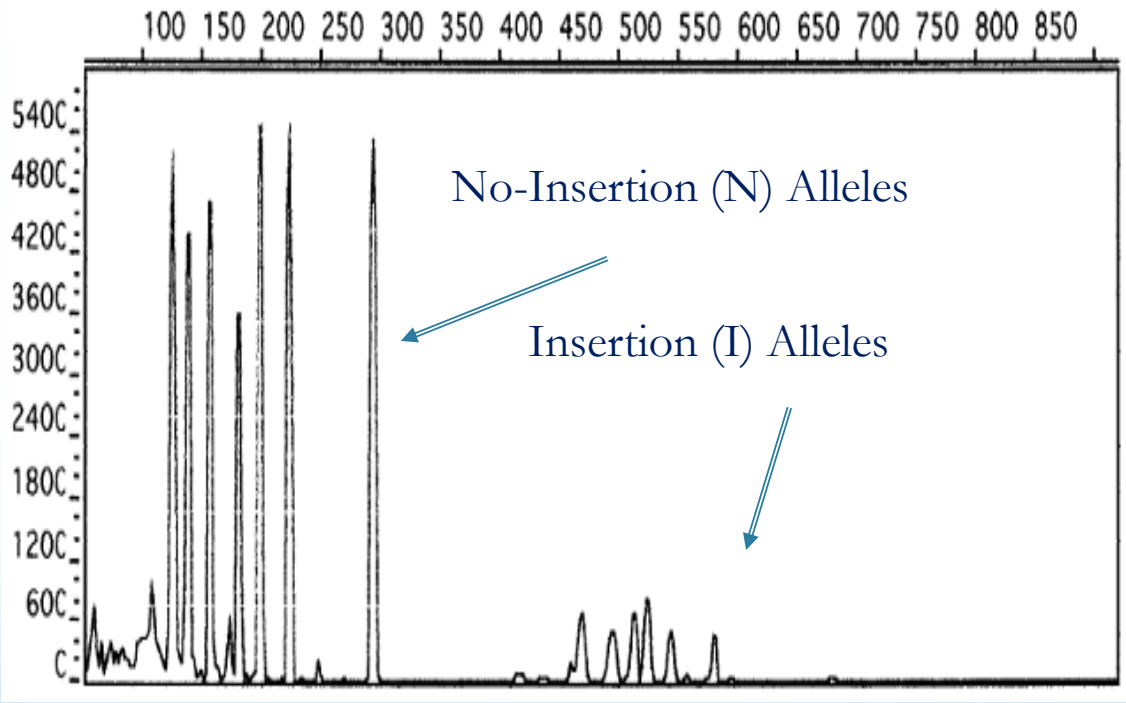


Amplification product
~400bp (I)

1. **Homozygous Insertion = I, I**
2. **Heterozygous = I, N or N, I**
3. **Homozygous No Insertion = N, N**

Caused preferential amplification of empty sites due to 300 bp allele size difference between I and N (allelic drop-out)

Alu multiplex with original primer design

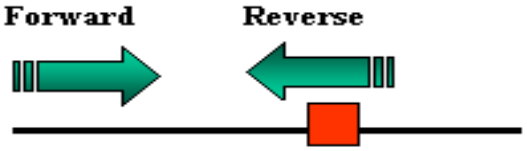


New Multiplex Design

A. Filled Site Reaction

Ccagttgttgaggggaacaaactaaata**agaagagtgaatgcacatttatga**aagtgtaaatgaacgattcTgttatgaacacaaacatg
accagggtgccgagccttatcatt**AAGAAACTG**GGCCGGGC**GCGGTGGCTCACGCCTGTAATCC**
CAGCACTTTGGGAGGCCGAGGCGGGCGGATCACGAGGTCAGGAGATCGAGACCATC
CCGGCTAAAACGGTGAAACCCCGTCTCTACTAAAAATACAAAAAATTAGCCGGGCG
TAGTGGCGGGCGCCTGTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATGGCG
TGAACCCGGGAGGCGGAGCTTGCAGTGAGCCGAGATCCCGCCACTGCACTGTCCAG
CCTGGGCGACAGAGCGAGACTCCGTCTCAAAAAAAAATAAATAAATAAATAAATAAA
AGAAACTGaattcatgactcccagtctgggggaacagaaaacattactgagctggagcacattggcc

B. Empty Site Reaction

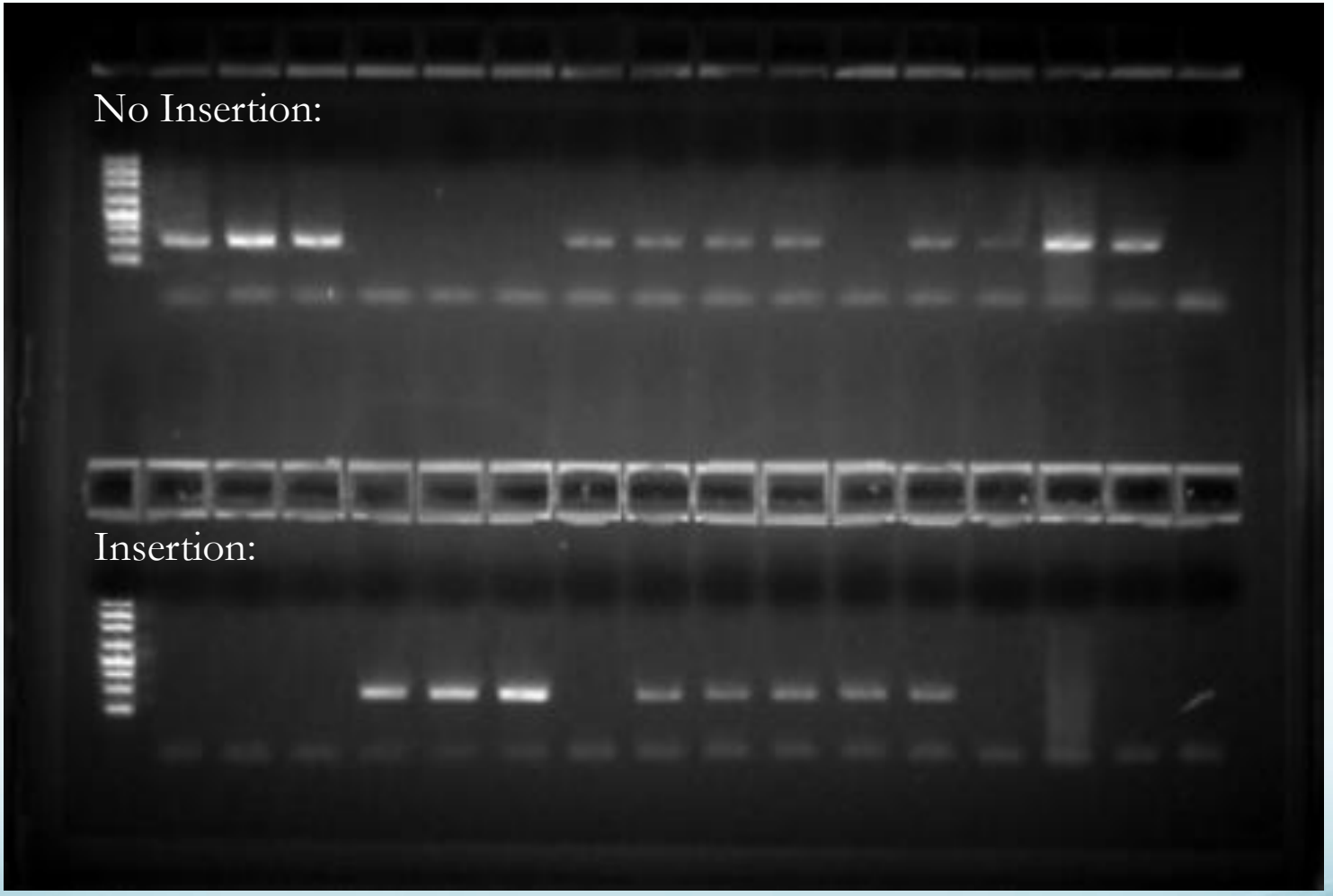


Ccagttgttgaggggaacaaactaaata**agaagagtgaatgcacatttatga**aagtgtaaatgaacgattctgttatgaacacaaacatg
accagggtgccgagccttatcatt**AAGAAACTG**aattcatgactcccagtctgggggaacagaaaacattactgagctggagca
cattggcccattcagttctgaaatgcatgtccaacaggtaggtggagccctgagcaagaattgctactgtgttgaagtcacagagccag

Carter and Sinha, US Patent Application #60/902,850

Method for Genetic Detection using Interspersed Genetic Elements

Primer Design

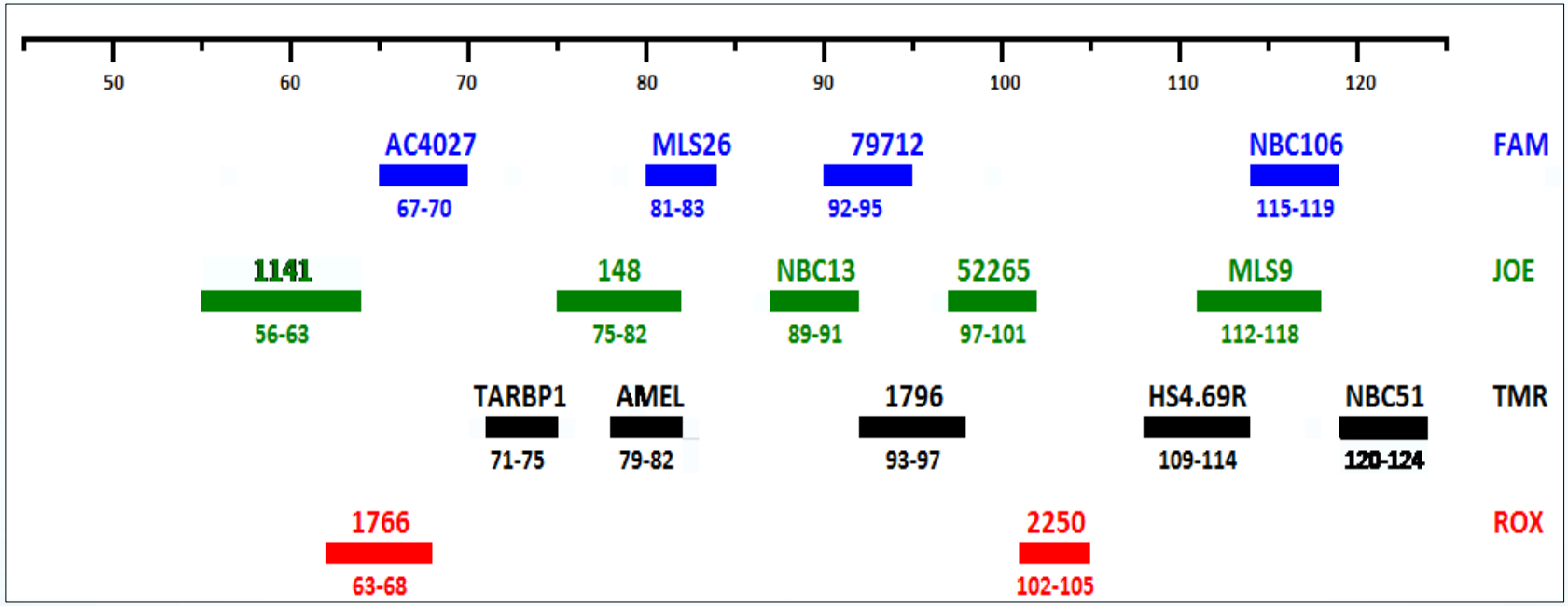


Panels showing variation within the Marker Chr20-79712 for 15 samples of Caucasian and African American population groups

InnoTyper-16

- 16 marker multiplex
- System amplifies *Alu* and *LINE* sequences less than 125 bp
- This system provides:
  - High sensitivity
  - High tolerance to degradation & inhibition
  - High discrimination power (greater than Mt DNA)

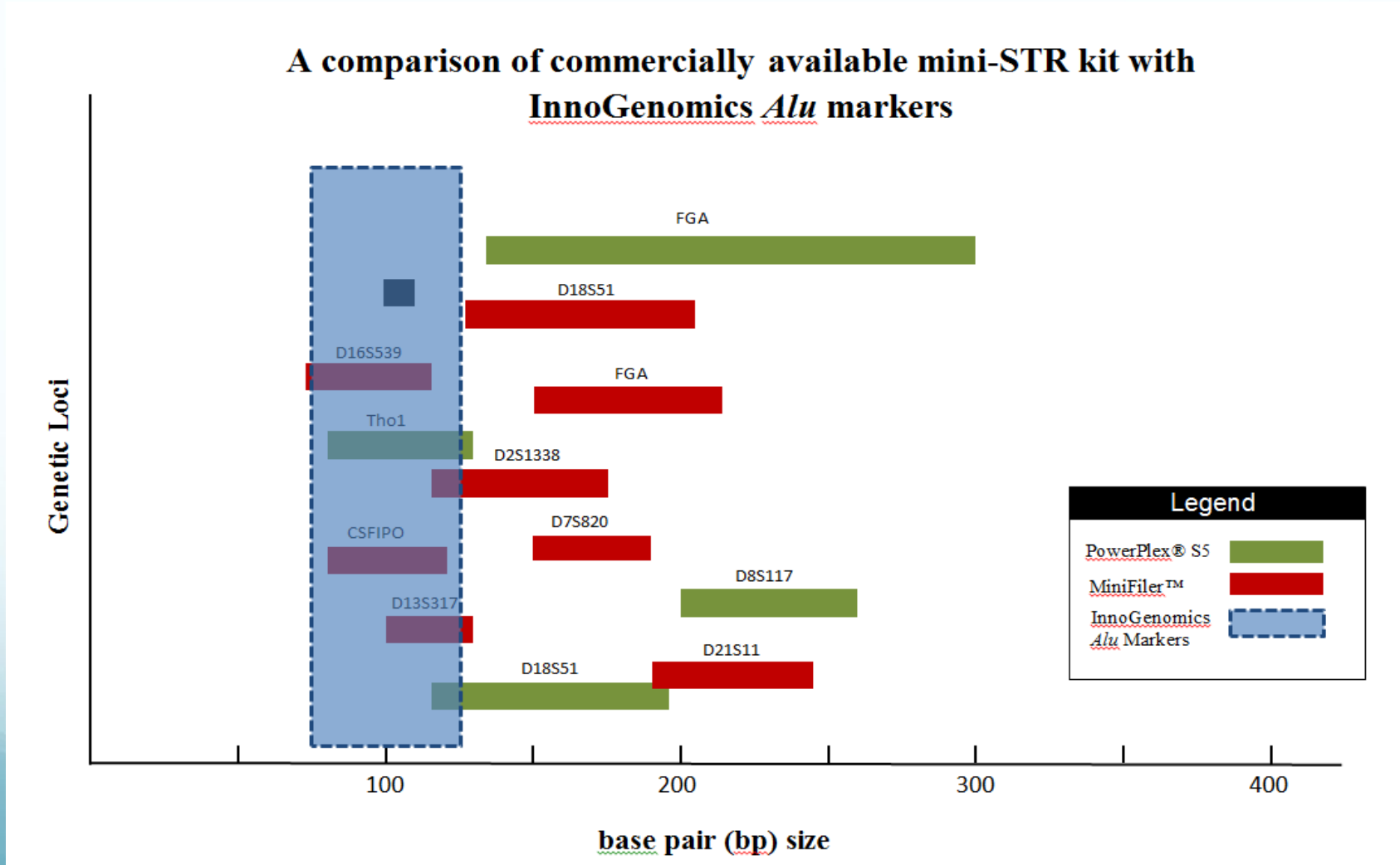
InnoTyper-16



Applications include highly degraded and inhibited DNA samples such as missing persons & mass disasters.

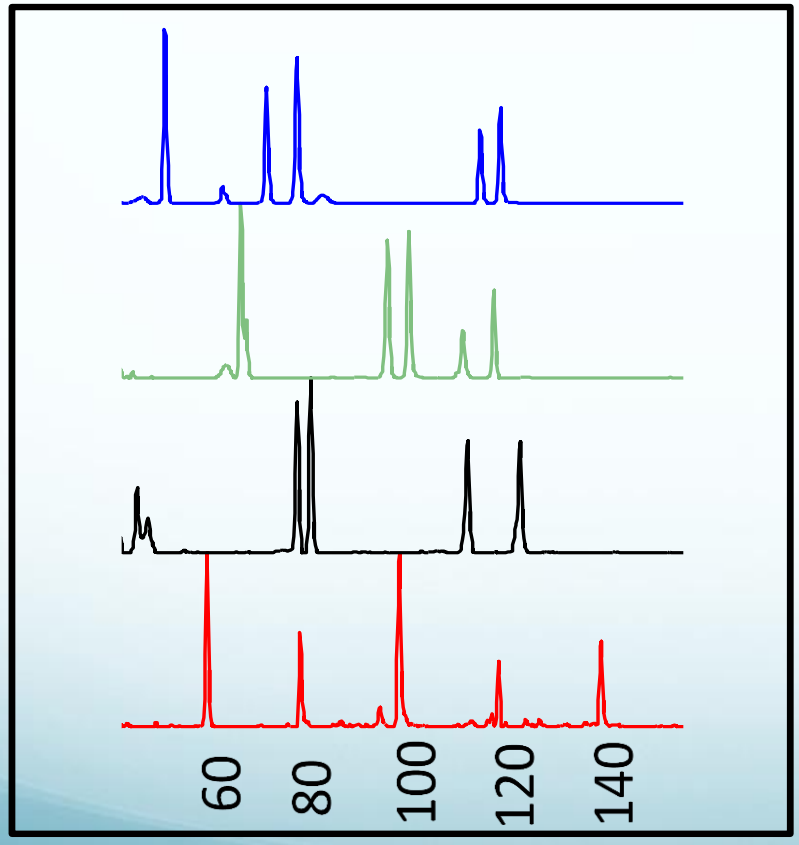
Schematic of the InnoTyper-16 Human DNA Analysis Kit

BP Size Comparison of mini-STR kits with InnoGenomics markers

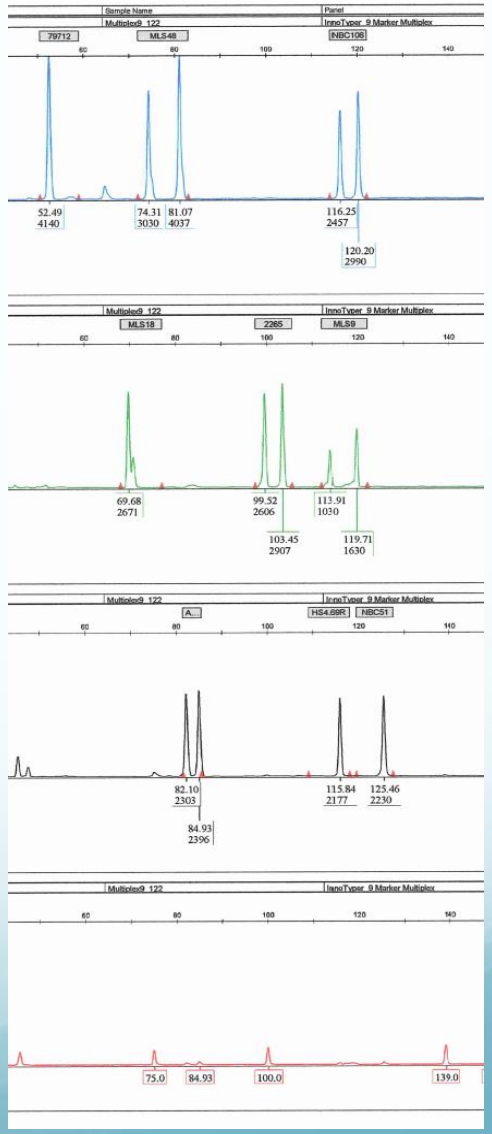


IntegenX RapidHIT™ Human DNA Identification System

RH 200 - 122



ABI 3130 - 122



Database Samples

- Four US populations: Caucasian (94), African American(90), Hispanic (92), Asian (90)
- Additional Caucasian and African American:

  100 Anonymous Paternity Trios (~200 M and AF samples) with STR data and known Exclusion and Inclusion results.

- Environmentally degraded:

Swab samples left at >90˚F in Louisiana for >5 years

Current InnoTyper16 Multiplex

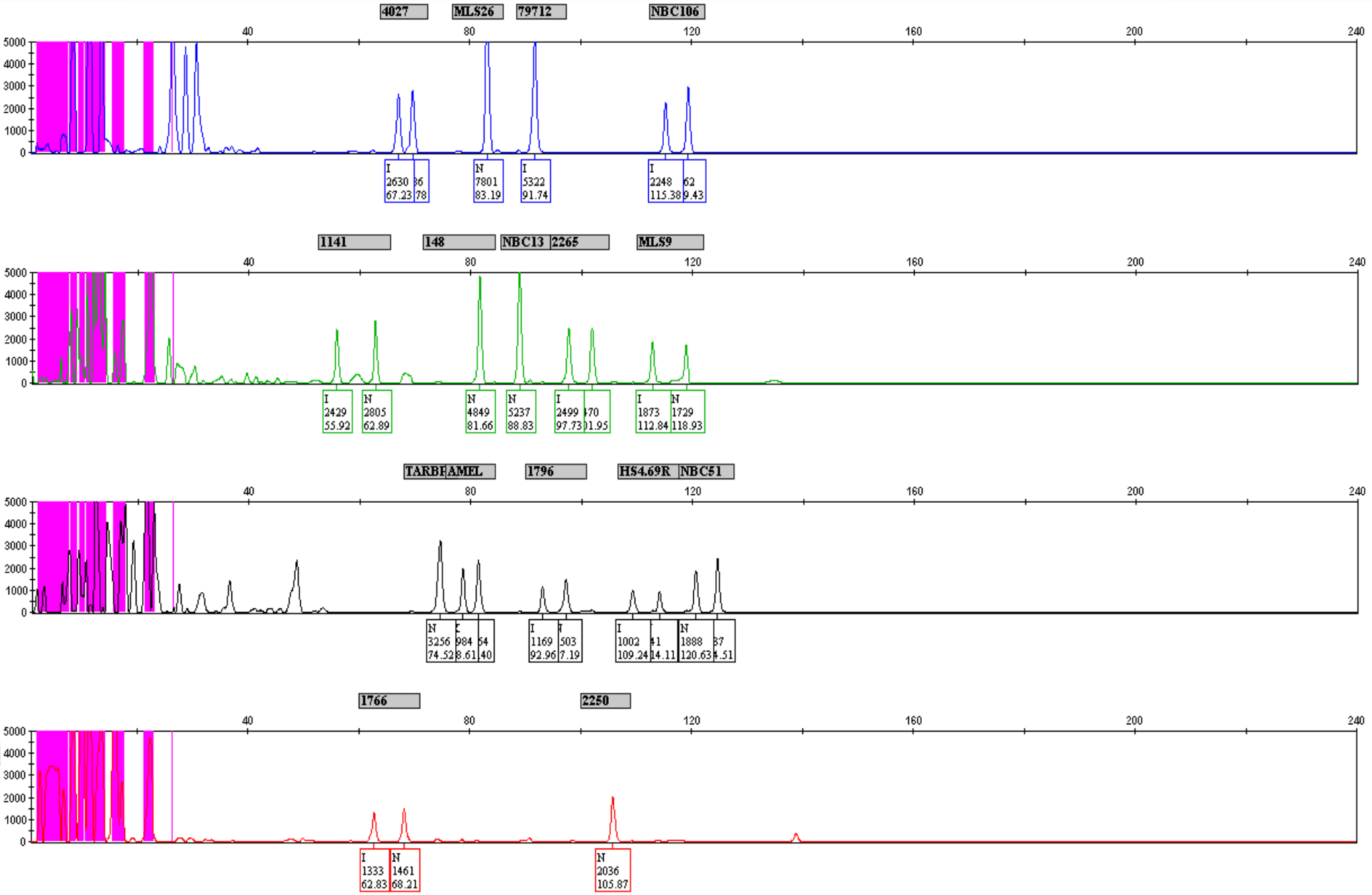


Table 1. Allele Frequencies, Test of HWE, and F_{ST} statistics of the 13 markers in Blacks and Caucasians

Marker	US Blacks (n = 140)				US Caucasians (n = 48)				F_{ST}
	Freq. Null allele[1]	Obs. Het[2] (H_{obs})	Exp. Het.[2] (H_{exp})	HWE p-value[3]	Freq. Null allele[1]	Obs. Het[2] (H_{obs})	Exp. Het.[2] (H_{exp})	HWE p-value[3]	
148	0.4679	0.5214	0.4979	0.5765	0.6979	0.4792	0.4217	0.3449	0.0544
1766	0.2750	0.4071	0.3987	0.8018	0.3750	0.4167	0.4688	0.4416	0.0114
2250	0.9179	0.1643	0.1507	0.2903	0.7500	0.3333	0.3750	0.4930	0.0509
2265	0.5893	0.4643	0.4841	0.6278	0.3021	0.2292	0.4217	0.0016	0.0835
4027	0.4643	0.5143	0.4974	0.6882	0.5417	0.5417	0.4965	0.5281	0.0060
79712	0.6679	0.4500	0.4508	0.8990	0.5000	0.5833	0.5000	0.2483	0.0290
NBC51	0.7107	0.4071	0.4113	0.9058	0.6250	0.5000	0.4688	0.6444	0.0095
MLS26	0.8250	0.2500	0.2887	0.1127	0.7500	0.4583	0.3750	0.1237	0.0084
MLS9	0.7786	0.3857	0.3354	0.0763	0.5313	0.5208	0.4979	0.7483	0.0677
NBC106	0.4500	0.4143	0.4950	0.0415	0.5933	0.5625	0.4827	0.2510	0.0206
NBC13	0.2071	0.3429	0.3284	0.6047	0.4375	0.5833	0.4922	0.1995	0.0608
HS4.69R	0.4143	0.4571	0.4853	0.4925	0.4688	0.6875	0.4979	0.0084	0.0030
TARBP1	0.7286	0.3857	0.3955	0.7693	0.4167	0.4583	0.4863	0.6910	0.0994
Average									0.0387

Population Data Analysis

[1] Allele designated as (N) in the genotype data is No Insertion and(I) is Insertion
[2] N allele frequencies, observed and expected Heterozygosity are all in proportions
[3] P-values for HWE test are based on all 3 genotypes evaluated by permutation tests (number of
replications 10,000)
Entries colored in red are the ones which do not conform to HWE

NOTE: Multiple testing Bonnferroni p-value (for 26 tests) is 0.0019, which implies that of the three significant deviations from HWE, only one (locus 2265 in US Caucasians) remains significant after multiple testing adjustment. Further, with the nominal level of significance of 0.05, the chance of observing 3 or more significant deviations from HWE in 26 tests is 0.1386, suggesting that the 3 Sporadic deviations seen can be attributed to chance alone.

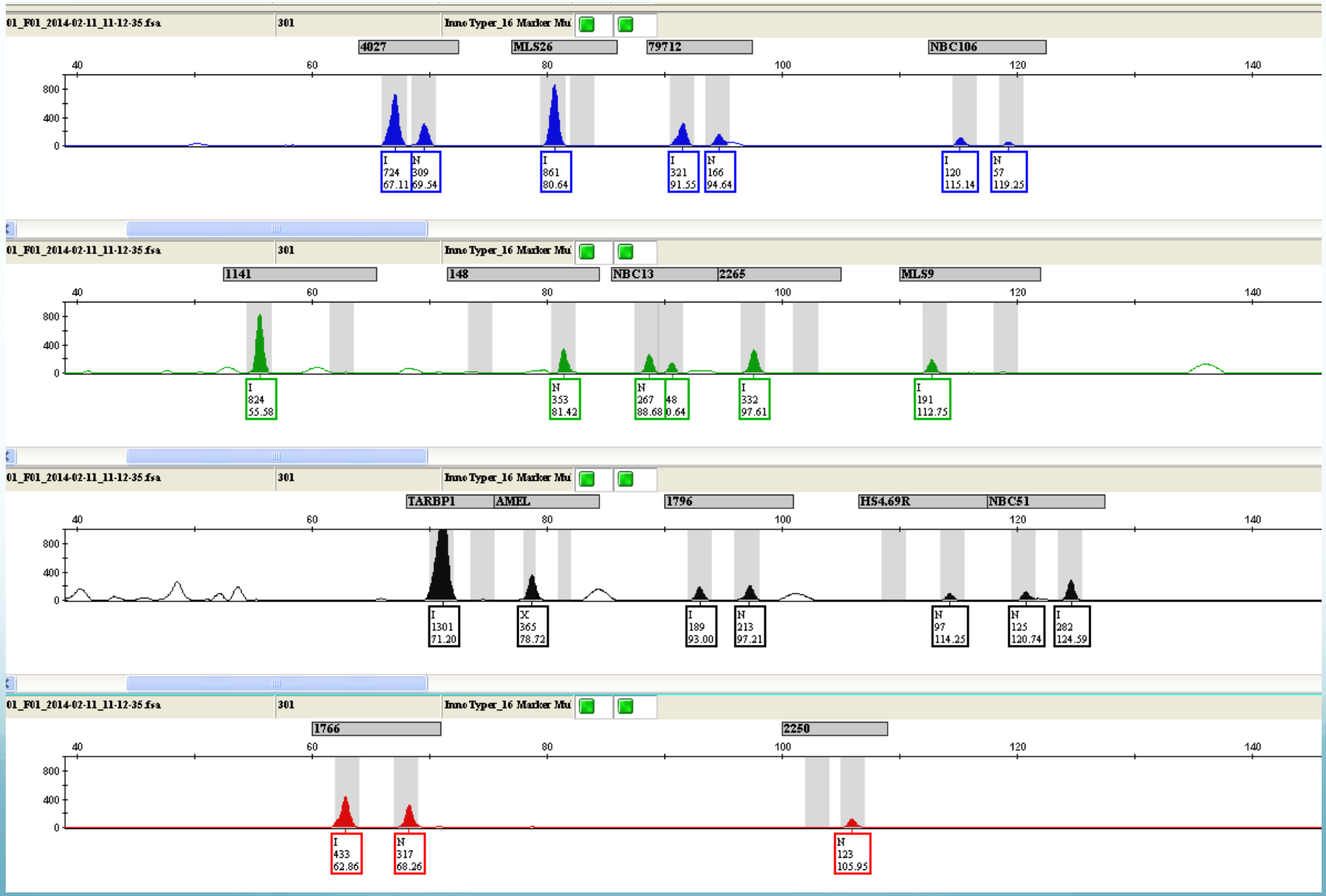
Table 2a. Estimates of Forensic and Parentage Testing Parameters of the 13 Markers in the Caucasian Population

Marker	RMP	PD	PE (Trio)	PE (Def)	PI (min)	PI (Max)
148	0.4233	0.5767	0.1664	0.0889	0.3582	3.3102
1766	0.3921	0.6079	0.1797	0.1099	0.4000	2.6667
2250	0.4609	0.5391	0.1523	0.0703	0.3333	4.0000
2265	0.4233	0.5767	0.1664	0.0889	0.2554	3.3102
4027	0.3768	0.6232	0.1866	0.1232	0.4615	2.1820
79712	0.3750	0.6250	0.1875	0.1250	0.5000	2.0000
NBC51	0.3921	0.6079	0.1795	0.1099	0.4000	2.6667
MLS26	0.4609	0.5391	0.1523	0.0703	0.3333	4.0000
MLS9	0.3761	0.6239	0.1870	0.1240	0.4705	2.1336
NBC106	0.3841	0.6110	0.1831	0.1165	0.4214	2.4588
NBC13	0.3790	0.6210	0.1855	0.1211	0.4444	2.2857
HS4.69R	0.3761	0.6239	0.1870	0.1240	0.4706	2.1331
TARBP1	0.3821	0.6179	0.1840	0.1182	0.4286	2.3998
Combined						
13 loci	6.51×10^{-6}	0.999993	0.9203	0.7708	6.67×10^{-6}	417,215

Table 2b. Estimates of Forensic and Parentage Testing Parameters of the 13 Markers in the African-American Population

Marker	RMP	PD	PE (Trio)	PE (Def)	PI (min)	PI (Max)
148	0.3761	0.6239	0.1870	0.1240	0.4698	2.1308
1766	0.4410	0.5590	0.1596	0.0795	0.3448	3.6364
2250	0.7327	0.2673	0.0697	0.0114	0.2724	12.1803
2265	0.3833	0.6167	0.1835	0.1172	0.4242	2.4349
4027	0.3763	0.6237	0.1868	0.1237	0.4667	2.1538
79712	0.4032	0.5968	0.1746	0.1016	0.3743	3.0111
NBC51	0.4312	0.5688	0.1634	0.0846	0.3518	3.3411
MLS26	0.5476	0.4524	0.1235	0.0417	0.3030	5.7143
MLS9	0.4979	0.5021	0.1396	0.0562	0.3211	4.5126
NBC106	0.3775	0.6225	0.1862	0.1225	0.4545	2.2222
NBC13	0.5050	0.4950	0.1372	0.0539	0.3128	4.8286
HS4.69R	0.3827	0.6173	0.1838	0.1178	0.4268	2.4137
TARBP1	0.4436	0.5564	0.1586	0.0782	0.3431	3.6846
Combined						
13 loci	2.69×10^{-5}	0.999973	0.9766	0.6904	2.33×10^{-6}	12.3 million

4 cm Hair Shaft – InnoTyper™ Results



In Summary…

1. Stable well characterized and published markers with a number of appealing genetic attributes, inherited by descent only.
2. Ability to analyze degraded nuclear DNA, ideal for use with samples requiring mt DNA Analysis.
3. High Power of Discrimination: greater than Mito.
4. Ideal for mass disaster testing of highly compromised samples.
5. Can be utilized with RDIS, next-gen platforms
6. Can provide information regarding bio-ancestral origin and sex of an unknown sample.
7. Like other Bi-Allelic systems, not yet suitable for mixture analysis using standard methods.

Acknowledgements

- This material is based upon work supported by the National Science Foundation under Grant No. 1230352. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Contact information:

Sudhir K. Sinha

Email:
ssinha@innogenomics.com

Thank You