Habitat distribution, population status, and genetic diversity of American ginseng

John Young, Frank van Manen, Tim King, Mike Eackles
USGS Biology
Leetown Science Center
Kearneysville, WV

David DeViney
Appalachian State University
Department of Biology
Boone, NC

Project Objectives

1. Develop habitat distribution models for American ginseng for portions of 18 states.

2. Determine demographic parameters and population characteristics.

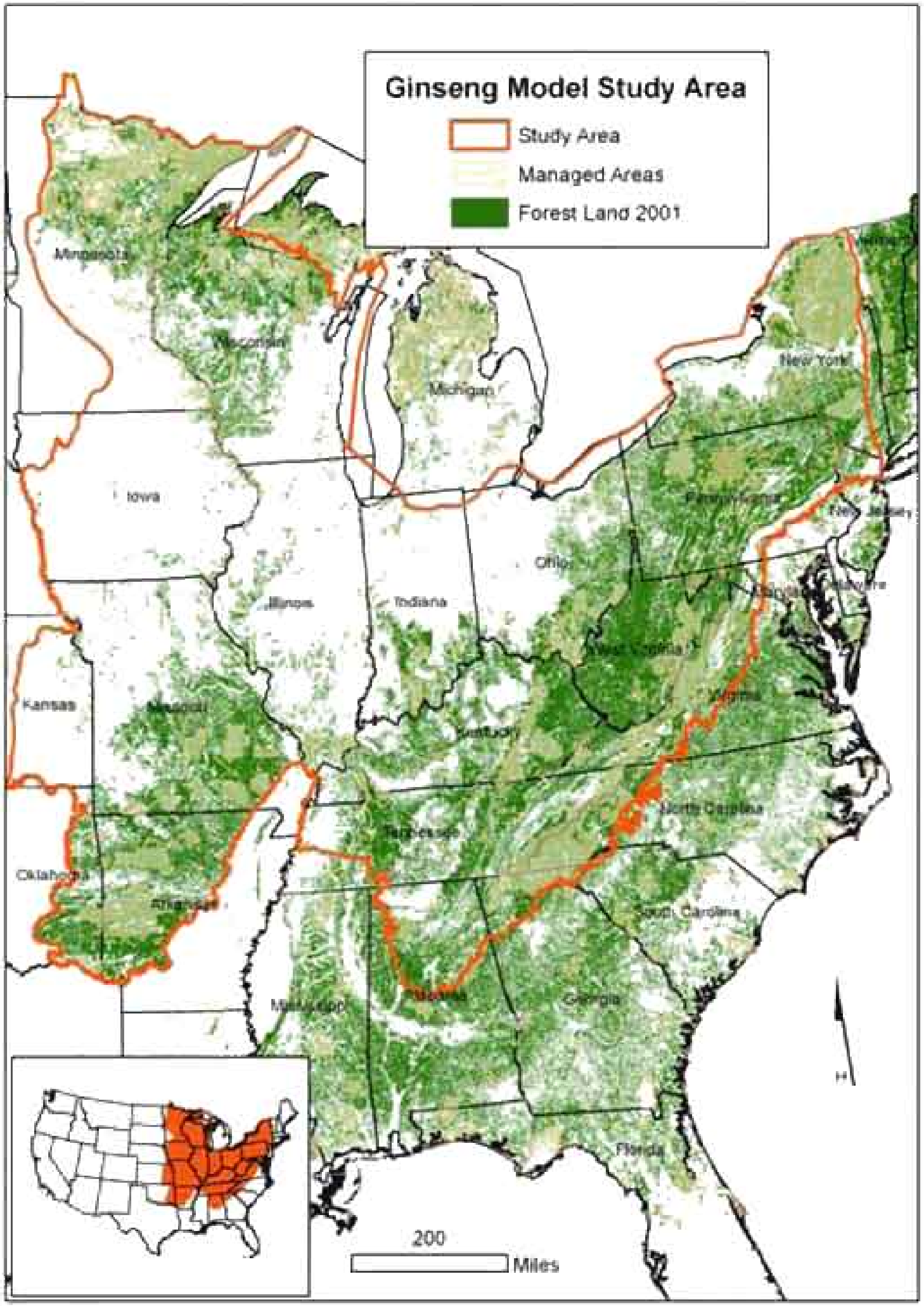
3. Determine genetic diversity and structure within and among field sampled populations.

Goals of Analysis

1. Contrast American ginseng distribution, population structure, and genetic diversity by land ownership

Public conservation - National & State Parks, Wildlife areas, Wilderness	**Private conservation** - Private conservancy lands
Public multiple use -National and State Forests, Game Lands, etc.	**Other private**

2. Assess relationship between habitat suitability, connectivity, and genetic diversity


Ginseng Model Study Area
Study Area
Managed Areas
Forest Land 2001
Minnesota
Wisconsin
Michigan
New York
Vermont
Iowa
Illinois
Indiana
Ohio
Pennsylvania
New Jersey
Maryland
Delaware
West Virginia
Virginia
Kentucky
Kansas
Missouri
Oklahoma
Arkansas
Tennessee
North Carolina
South Carolina
Mississippi
Alabama
Georgia
Florida
200
Miles
N

Habitat modeling (previous study)

Regression equation to predict relative probability of occurrence:

-1.2931 (intercept) + (slope * 0.0589)
– (elevation * 0.0012) + (% deciduous forest * 0.0181 - (average solar insolation * 0.0129)

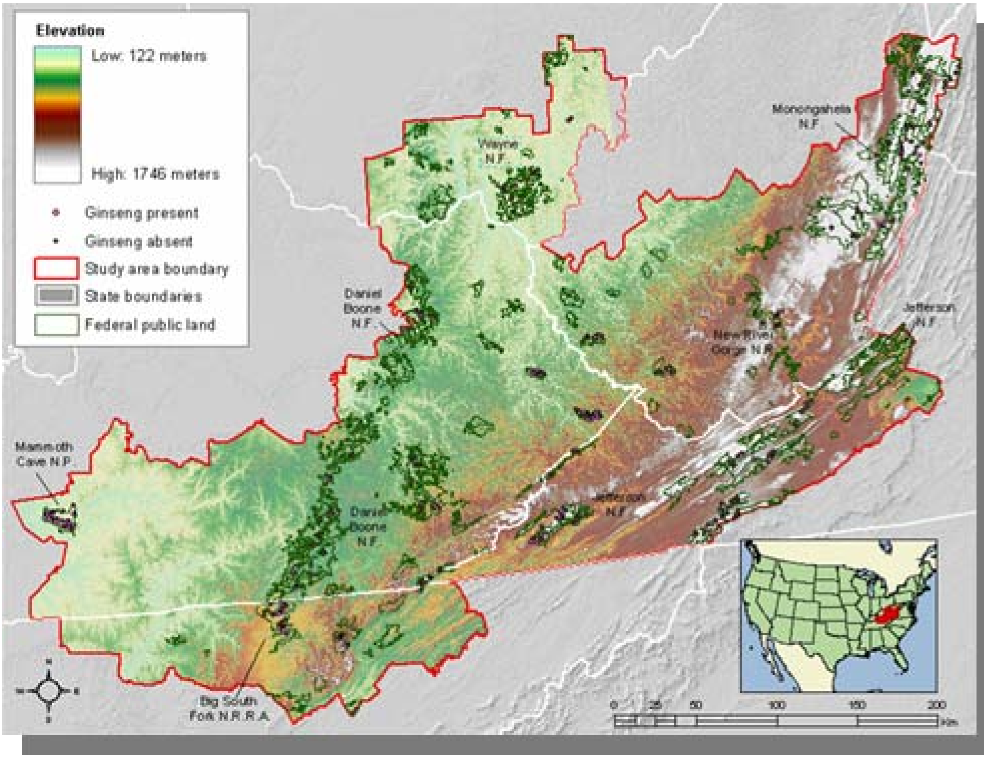
Probability of occurrence increases with:

-Decreasing elevation

-Increasing slope

-Deciduous forest cover

-Low sun exposure

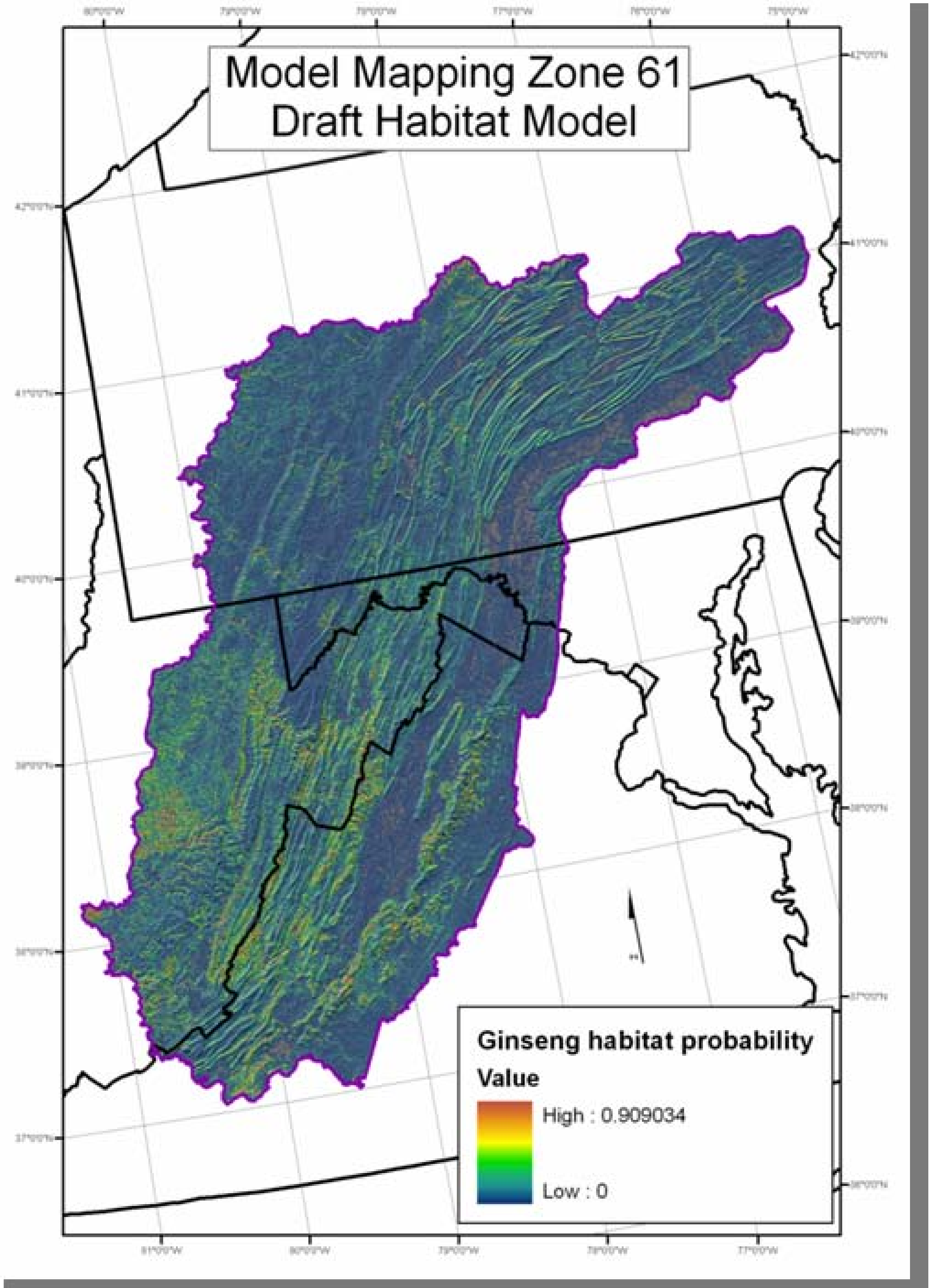


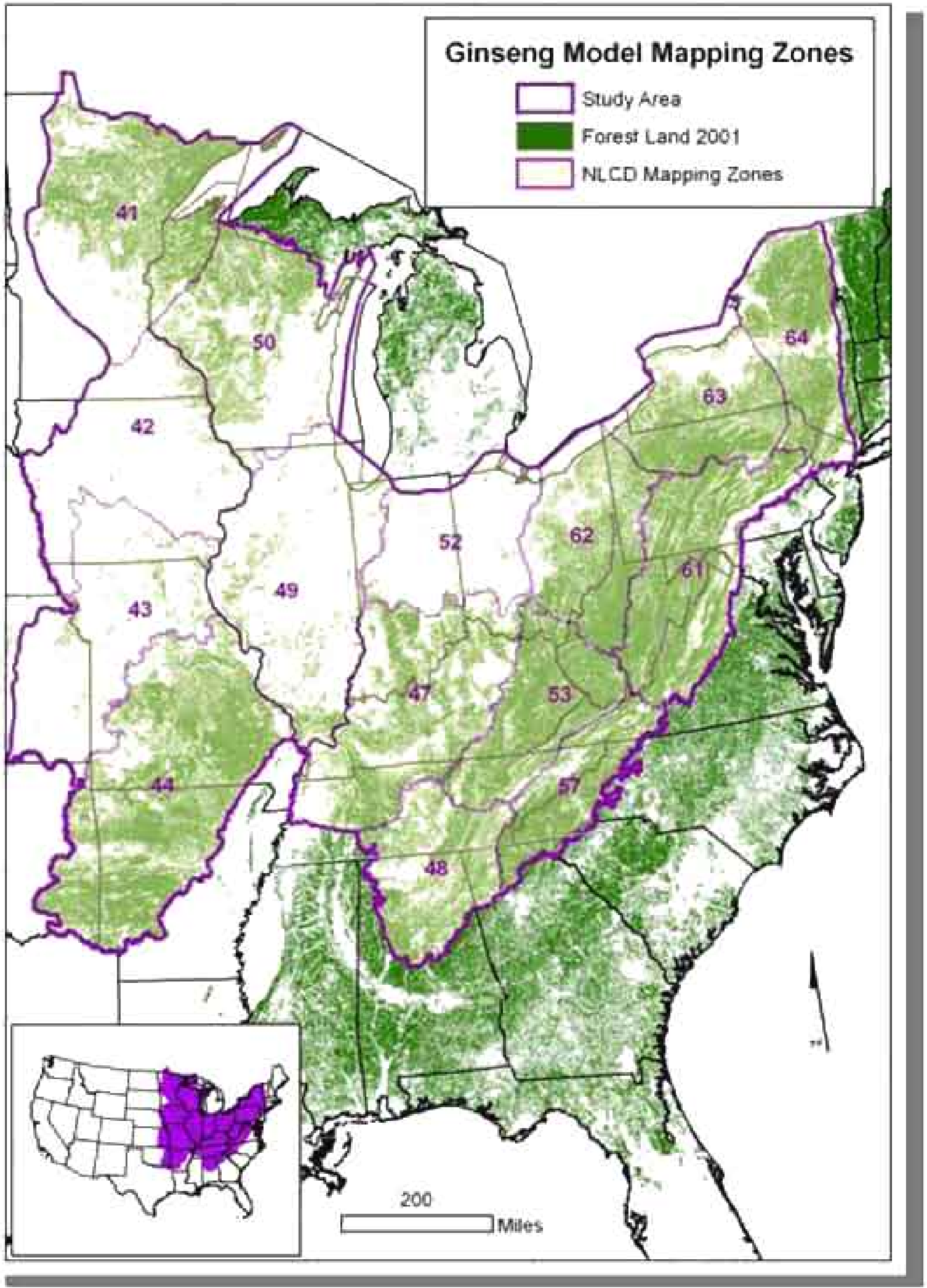
2004-2005 model study area

Habitat modeling (current study – 1st draft)

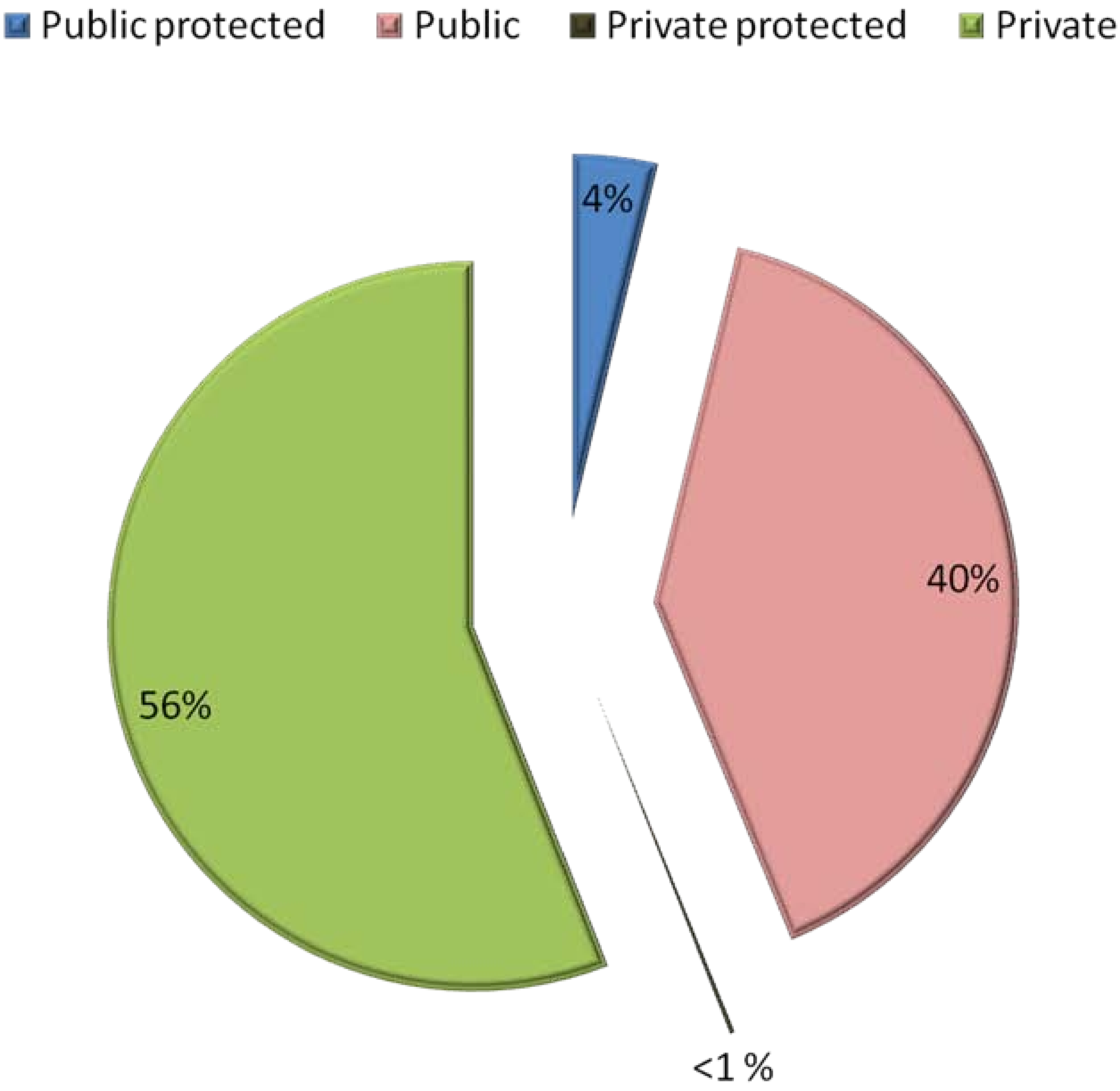
Extrapolated logistic regression parameters from previous models to 15 “mapping zones” using GIS

Zones correspond to regional land cover mapping area of the National Land Cover Database (USGS/EPA)




Ginseng Model Mapping Zones
Study Area
Forest Land 2001
NLCD Mapping Zones
41
50
42
43
49
44
52
47
48
62
61
53
57
63
64
200
Miles

Prime Habitat by Land Ownership



* Preliminary, for Mapping Zone 61 only, using prime habitat cutoff from 2004-2005 study

Modeling issues

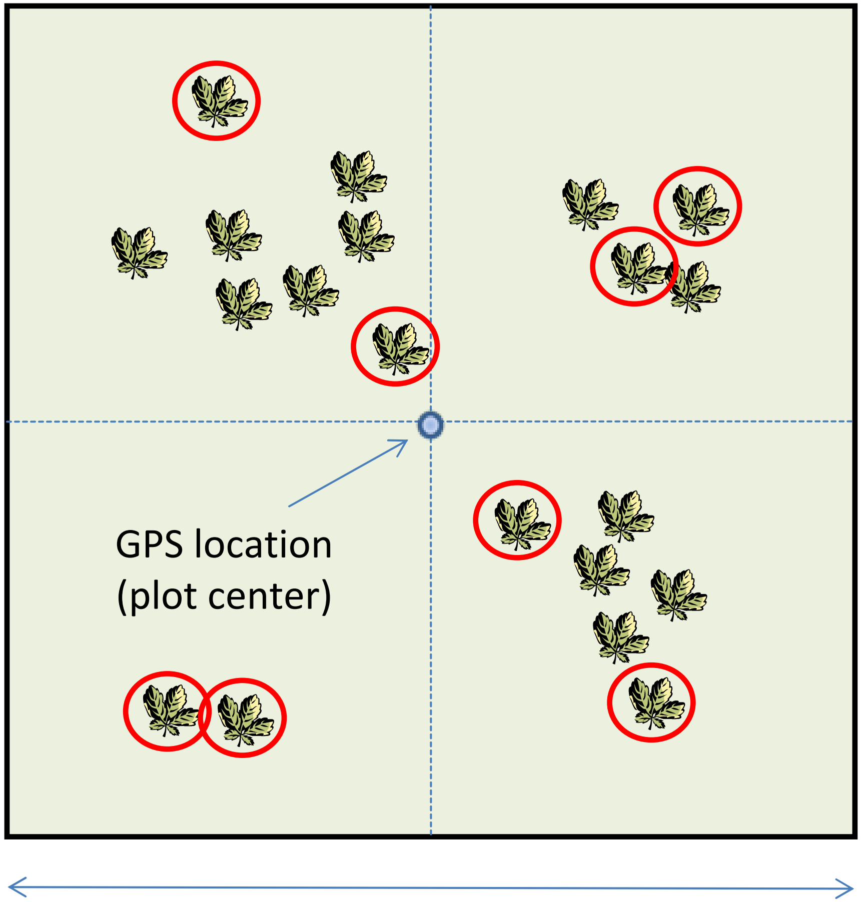
- Draft regional models used to guide field sampling
  - Models based on topography, solar radiation, and forest cover only
  - Models did not extrapolate well outside of previous study area

- New model development
  - Additional variables (i.e., geology [soils], precipitation)
  - Additional data (known plant locations) from 2007-2008 field surveys, herbaria records, and other investigators records

Field Sampling

- Sampled 196 plots during 2007-2008 in 11 states (NY, PA, OH, WV, VA, TN, KY, NC, SC, IN, MO)
- 30-m sampling plots
  - Searched for, and counted ginseng plants by size class
  - Measured elevation, aspect, slope, slope position, canopy cover
  - Recorded associated species and disturbance
  - Collected site photos and GPS coordinates
- Sampled DNA from up to 8 plants/plot
  - 802 genetics samples from 11 states
  - Measured height, # prongs, reproductive status

Field Sampling



Plot size = 30x30 m (0.09ha)

20 plants = 222 plants/ha

○ Plants sampled for genetics

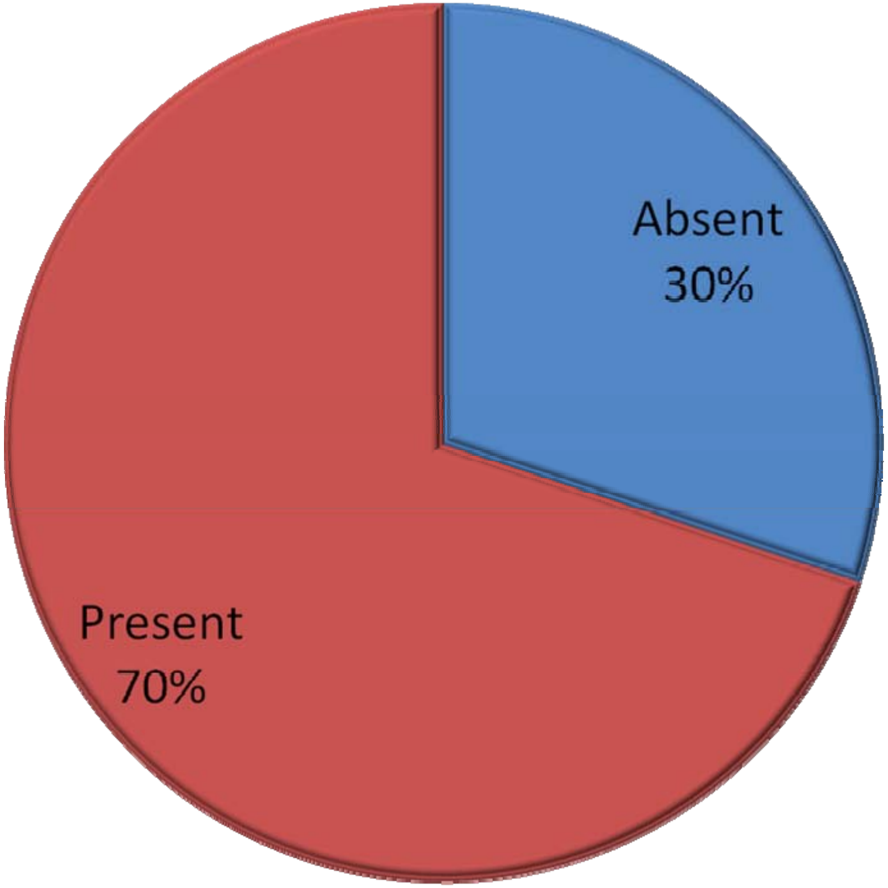

PEEL COVER HERE
FTA®
Plant Card
Whatman®


FTA®
Plant Card
Whatman®


FTA®
Plant Card

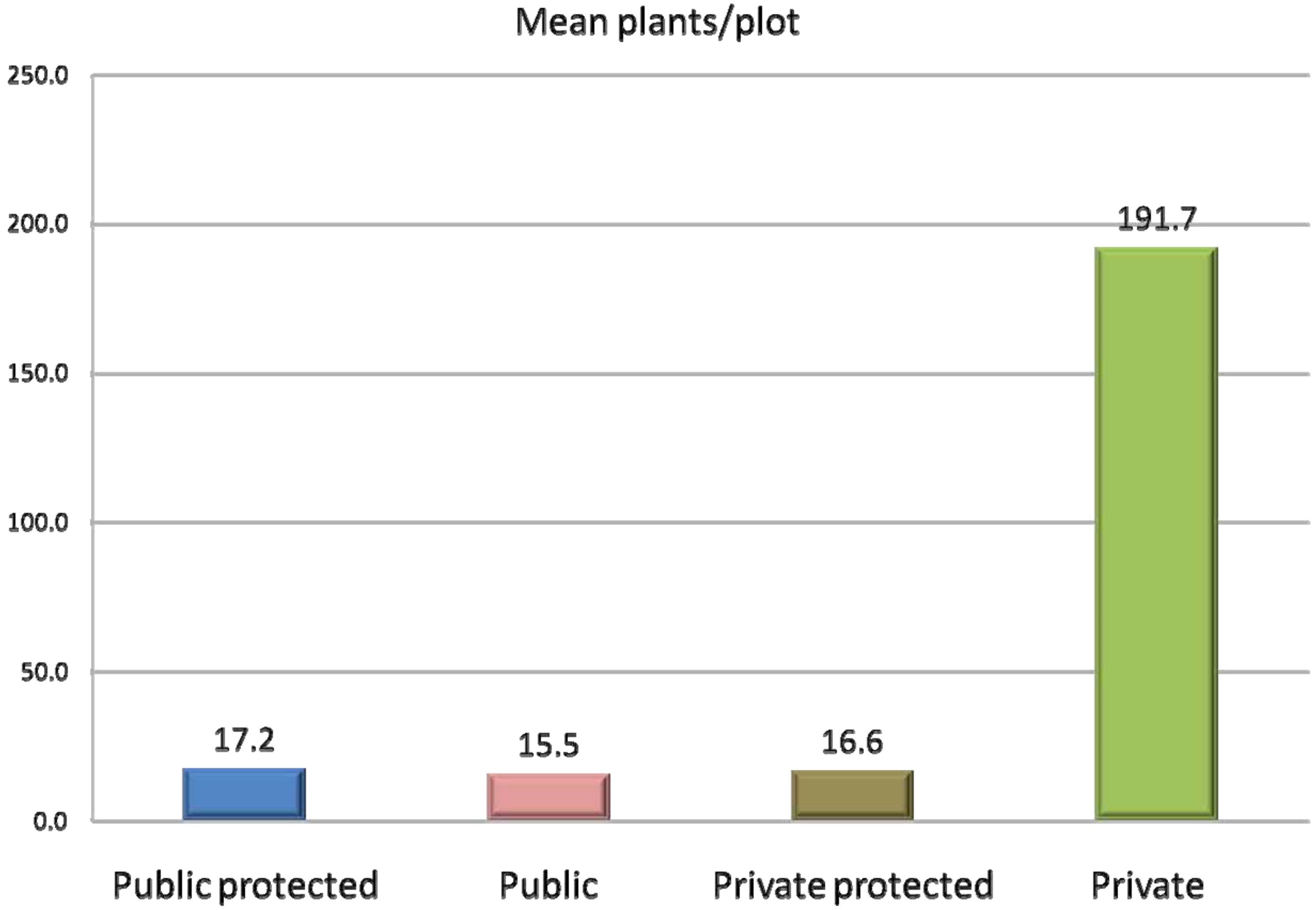

Whatman®
Lot No. FT6802308
WV-007a
TUCK COVER HERE
FTA® Plant Card
USGS 2007-2009 American Gin
(Example: WV-001)
Co-occurring herbaceous species:

Ginseng occupancy in surveyed plots



Plots sampled 2007-2008, N=196

Population status at field plots



* 33.8 plants/plot across all sites sampled 2007-2008, N=196

Genetic Analysis in Plants

- Long lived, out-crossing, late successional species retain most of their genetic variability within populations[1]

- Annual, selfing, and/or early successional species allocate relatively more genetic variability among populations[1]

- American ginseng ??
  - *Mixed mating system,* self compatible
  - long-lived
  - late successional forest species

---

[1] Nybom (2004) review of 307 studies using DNA markers

Ginseng - Genetic Analysis Previous studies

- Allozymes:
  - More variation found among populations than within for wild populations, more within than among pop variation for cultivated (Grubbs and Case 2004)
  - Wild populations differ from cultivated pops in amount and distribution of genetic diversity, but differences in total genetic diversity not large (Grubbs and Case 2004)
  - Evidence of isolation by distance (Cruse-Sanders and Hamrick 2004)
- DNA markers [RAPD]:
  - High degree of genetic diversity within wild and cultivated ginseng (Schluter and Punja 2002, Boehm et al. 1999, Bai et al. 1997)
  - More variation detected within populations rather than among for both cultivated and wild populations (Schluter and Punja 2002)
  - Distinct populations observed (Lim et al. 1997), natural populations genetically distinct from cultivated (Schluter and Punja 2002, Boehm et al. 1999)
  - Evidence of mixing of wild populations with cultivated populations in PA (Boehm et al. 1999)

Genetic Markers

- Allozyme markers (enzymes)
  - Advantages:
    - Good for range wide analysis (broad scale variation)
    - Inexpensive
  - Disadvantages:
    - Limited to protein encoding loci
    - Underestimate genetic diversity (fewer polymorphic loci)
    - Require fresh plant material, sample preservation
- DNA markers (e.g. RAPD, Microsatellites, AFLP)
  - Advantages:
    - More resolution for detecting population size, structure, gene flow, bottlenecks (more polymorphic loci)
    - Discrimination of individuals and source populations possible
    - Minimal sample preservation needed (FTA cards, silica gel)
  - Disadvantages:
    - Reproducibility (RAPDs only)
    - Anonymous regions of DNA, dominant markers (RAPD, AFLP)
    - High level of expertise required for marker development (microsatellites)

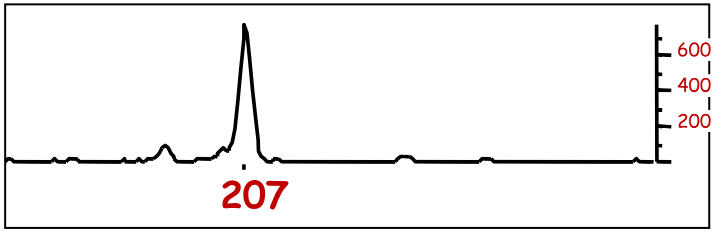
Genetic Markers - Microsatellites

- Short sections of repeated sequences of DNA
- Mutations in non-coding (neutral) DNA region
  - Some mutations common to genus, species, populations
  - Some mutations unique to individual
- Variation in length of repeats = marker
- Uses sequence information

Flanking region Flanking region

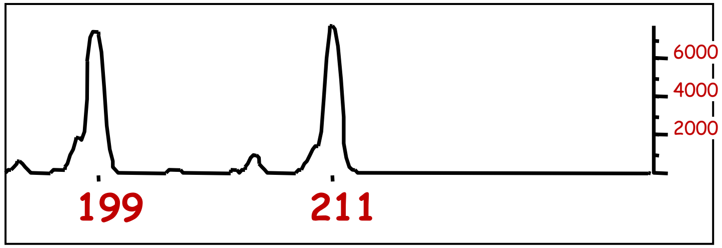
CTAGCTAG GATAGATAGATA TTGTCA

CTAGCTAG GATAGATAGATA TTGTCA



CTAGCTAG GATA TTGTCA

CTAGCTAG GATAGATAGATAGATA TTGTCA

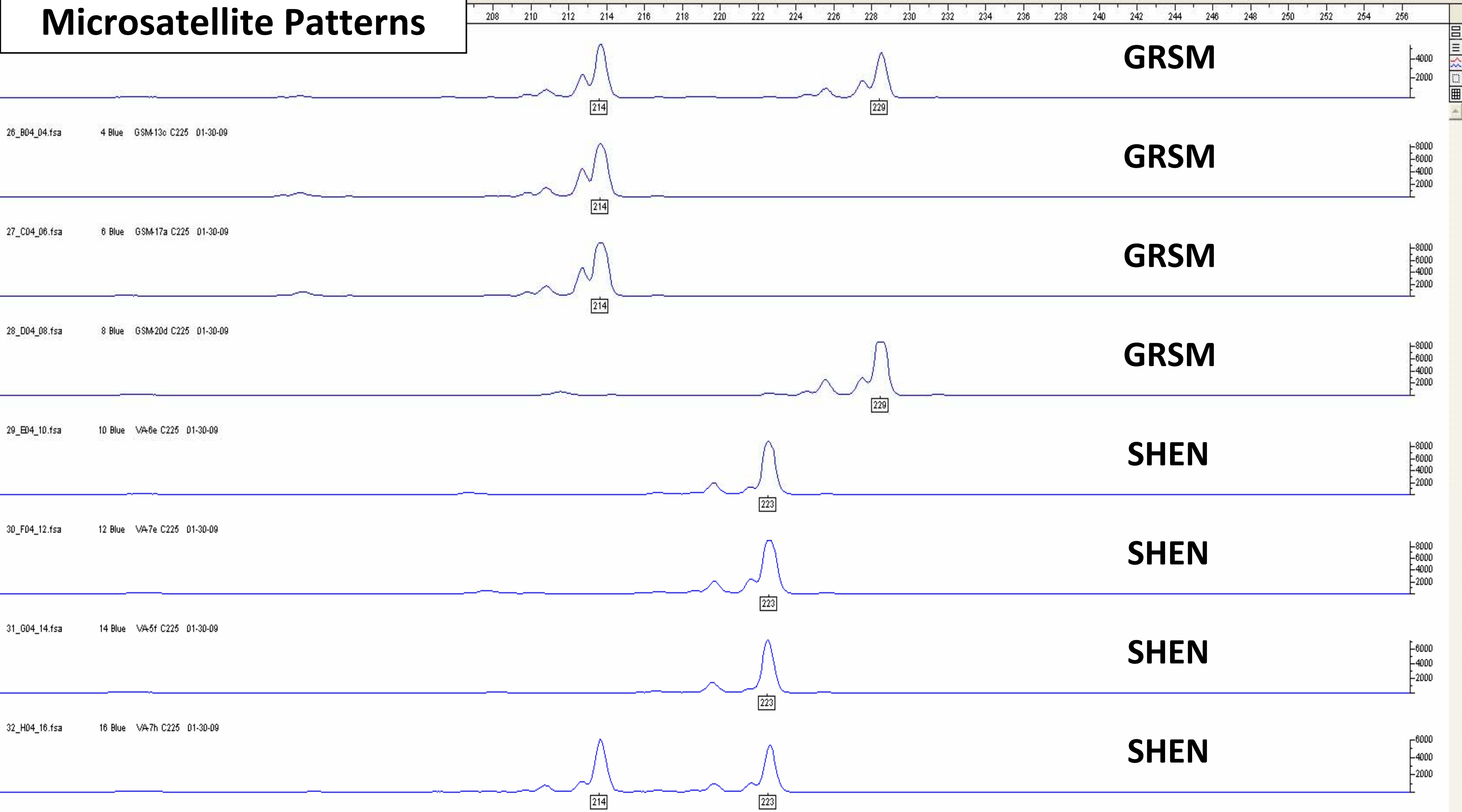


Genetic Markers – Microsatellites

King Lab

- Need 12-15 markers for population and individual discrimination
- Tested published markers for Korean ginseng (*Panax ginseng)*
  - Kim et al. 2007 [10/14], Ma et al. 2007 [14/22])
    - Amplification problems in *Panax quinquefolius*
    - Lack of polymorphism
- Developed our own set of 25 microsatellite markers for American ginseng, <u>14</u> look most promising
  - Enriched microsatellite DNA libraries created
  - Markers screened and tested in our laboratory

Ginseng
Microsatellite Patterns



GRSM

26_B04_04.fsa 4 Blue GSM-13c C225 01-30-09

GRSM

27_C04_06.fsa 6 Blue GSM-17a C225 01-30-09

GRSM

28_D04_08.fsa 8 Blue GSM-20d C225 01-30-09

GRSM

29_E04_10.fsa 10 Blue VA-6e C225 01-30-09

SHEN

30_F04_12.fsa 12 Blue VA-7e C225 01-30-09

SHEN

31_G04_14.fsa 14 Blue VA-5f C225 01-30-09

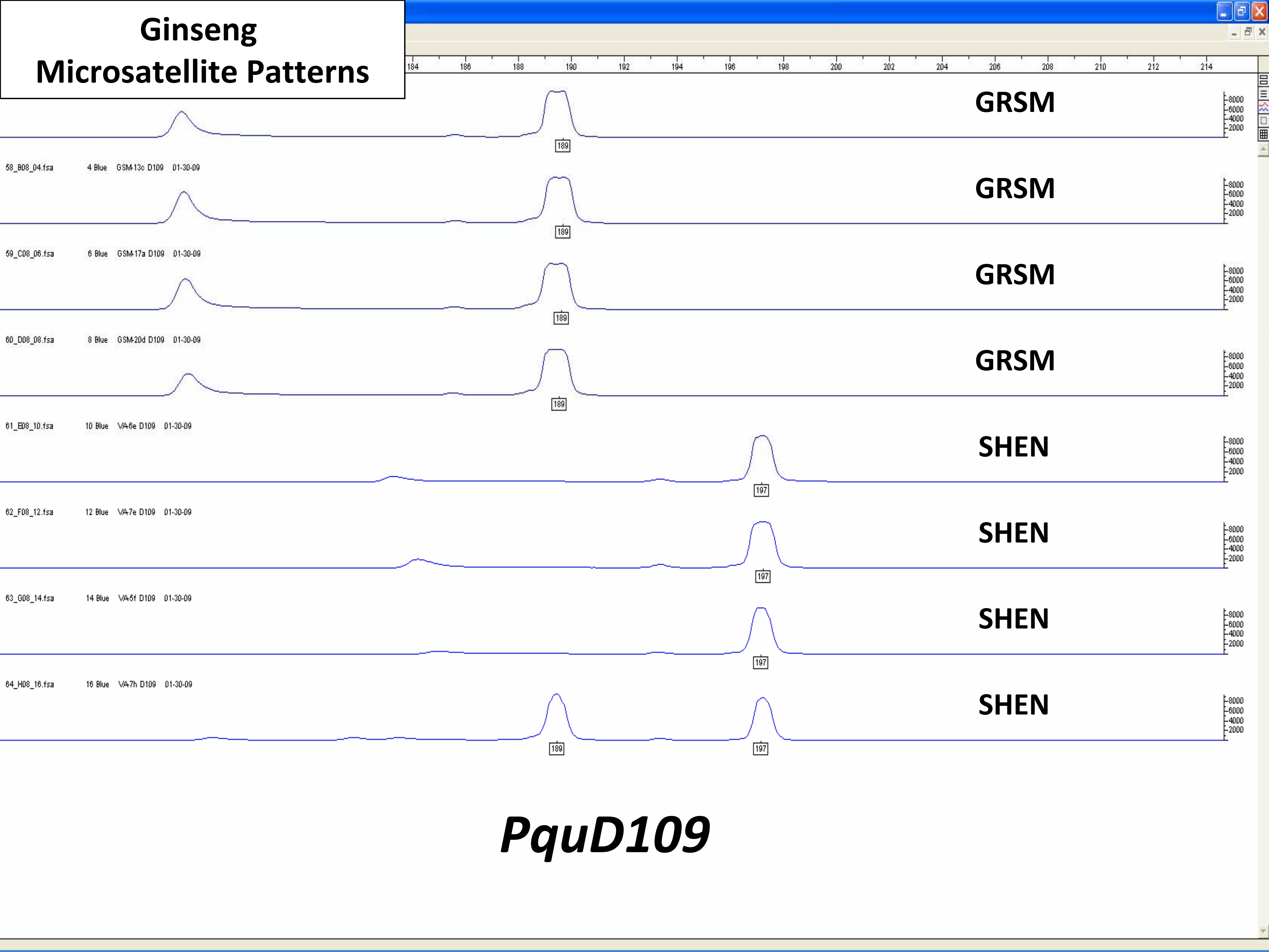
SHEN

32_H04_16.fsa 16 Blue VA-7h C225 01-30-09

SHEN

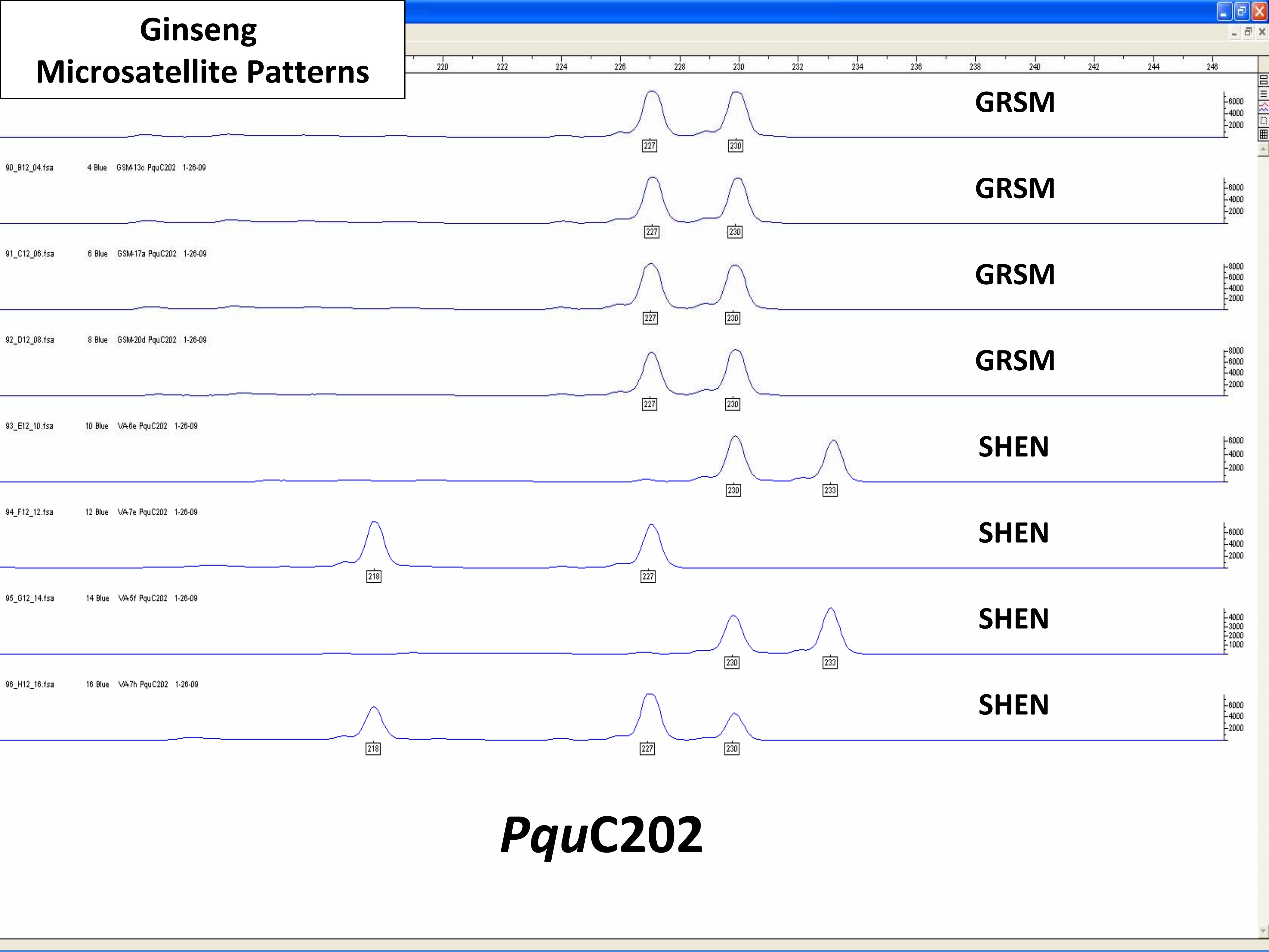
PquC225

Ginseng Microsatellite Patterns



PquD109

Ginseng
Microsatellite Patterns



GRSM

90_B12_04.fsa 4 Blue GSM-13c PquC202 1-26-09

GRSM

91_C12_06.fsa 6 Blue GSM-17a PquC202 1-26-09

GRSM

92_D12_08.fsa 8 Blue GSM-20d PquC202 1-26-09

GRSM

93_E12_10.fsa 10 Blue VA-6e PquC202 1-26-09

SHEN

94_F12_12.fsa 12 Blue VA-7e PquC202 1-26-09

SHEN

95_G12_14.fsa 14 Blue VA-5f PquC202 1-26-09

SHEN

96_H12_16.fsa 16 Blue VA-7h PquC202 1-26-09

SHEN

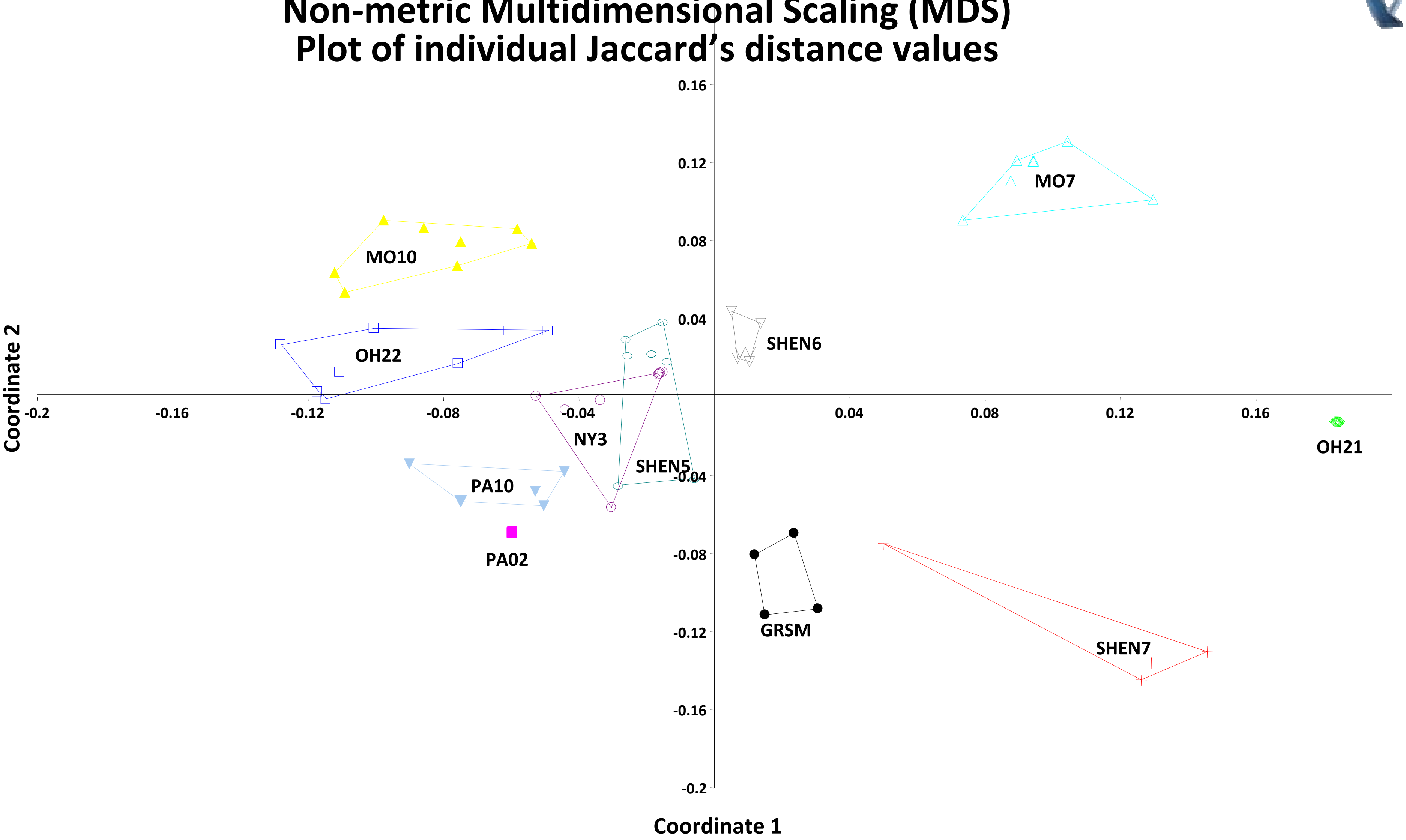
PquC202

Test of microsatellite markers

- Subset of sites across geographic range, and land ownerships (78 individuals)
- Test of allele frequency differences across 12 of 14 loci,3-9 alleles each, (66 markers)
- Sites:
    - GRSM = Great Smoky National Park (4 plants)
    - SHEN5, SHEN6, SHEN7 = Shenandoah National Park, VA (3 sites, 18 plants)
    - MO10 = Roaring River State Park, MO (8 plants)
    - MO7 = Caney Mountain Conservation Area, MO (8 plants)
    - PA02 = Allegheny National Forest, PA (8 plants)
    - PA10 = Private land, western PA (8 plants)
    - NY3 = Private land, Cairo, NY (8 plants)
    - OH22 = Cultivated ginseng farm, OH (8 plants)
    - OH21 = Edge of Appalachia Preserve, OH (8 plants)

Microsatellite DNA Collection Differentiation

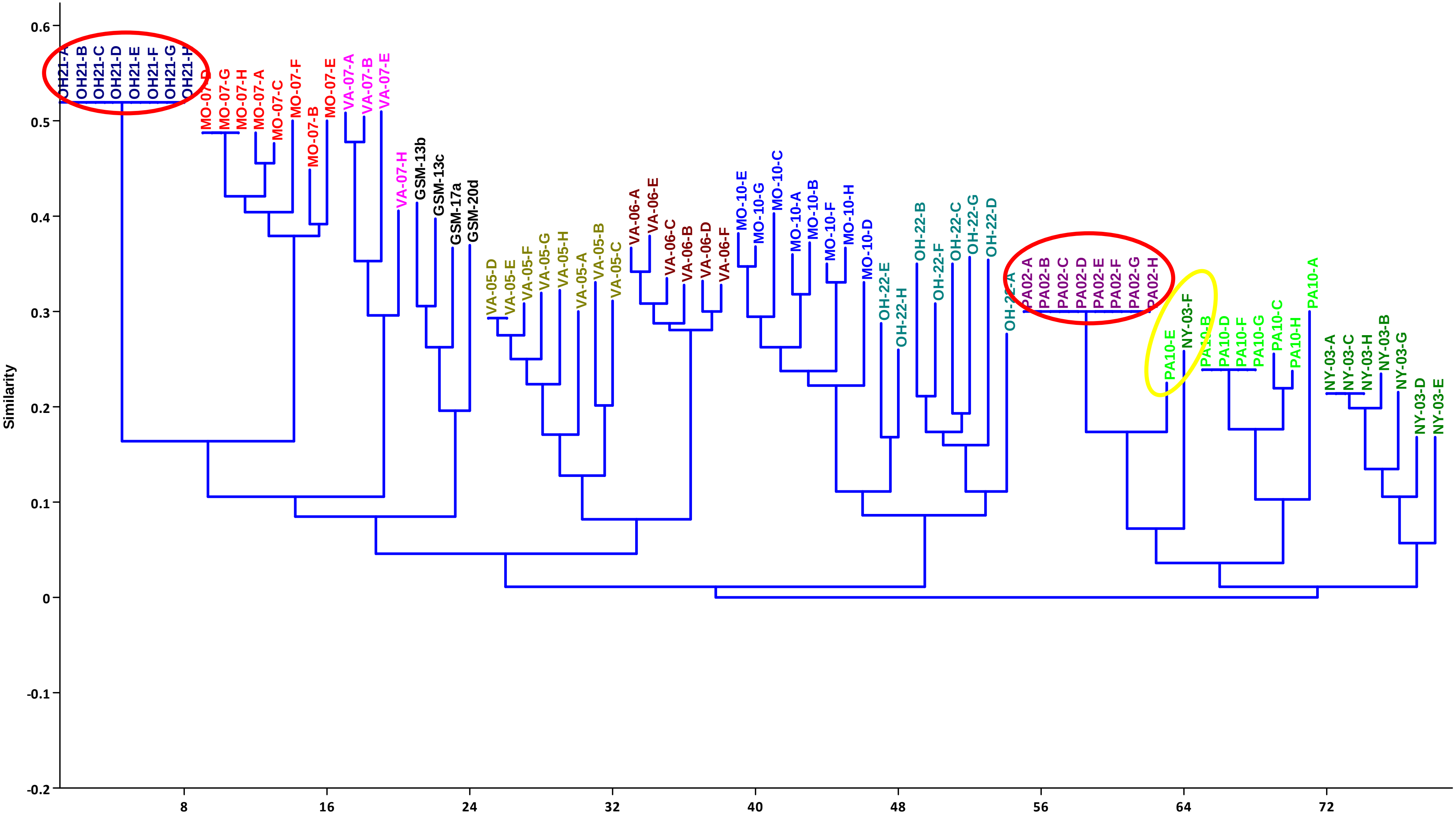
Non-metric Multidimensional Scaling (MDS) Plot of individual Jaccard's distance values



Polygons depict inter- and intra-collection variation in Jaccard's Distance

Microsatellite DNA Collection Differentiation

Neighbor Joining Tree on Jaccard's distance values



Φ_{PT} - Test of Genetic Differentiation

(F_{ST} analog for binomial data)

Collection	GRSM	MO7	MO10	NY	OH22	OH21	PA02	PA10	SHEN5	SHEN6	SHEN7	Prob.
GRSM	0.000	0.002	0.002	0.002	0.003	0.007	0.004	0.001	0.005	0.007	0.023	**GRSM**
MO7	0.806	0.000	0.001	0.001	0.003	0.001	0.002	0.001	0.001	0.002	0.003	**MO7**
MO10	0.747	0.804	0.000	0.001	0.001	0.001	0.001	0.002	0.001	0.001	0.002	**MO10**
NY	0.651	0.760	0.680	0.000	0.001	0.001	0.001	0.001	0.002	0.001	0.006	**NY**
OH22	0.639	0.714	0.483	0.455	0.000	0.001	0.001	0.001	0.001	0.001	0.004	**OH22**
OH21	0.889	0.909	0.889	0.859	0.788	0.000	0.122	0.002	0.001	0.002	0.004	**OH21**
PA02	0.848	0.923	0.855	0.771	0.691	0.000	0.000	0.001	0.001	0.001	0.003	**PA02**
PA10	0.681	0.838	0.752	0.531	0.542	0.899	0.771	0.000	0.001	0.002	0.004	**PA10**
SHEN5	0.667	0.776	0.675	0.427	0.529	0.854	0.783	0.667	0.000	0.002	0.001	**SHEN5**
SHEN6	0.768	0.851	0.765	0.637	0.648	0.958	0.944	0.812	0.576	0.000	0.002	**SHEN6**
SHEN7	0.657	0.830	0.777	0.734	0.672	0.890	0.898	0.769	0.679	0.771	0.000	**SHEN7**
Φ_{PT}	**GRSM**	**MO7**	**MO10**	**NY**	**OH22**	**OH21**	**PA02**	**PA10**	**SHEN5**	**SHEN6**	**SHEN7**	

Pair-wise Φ_{PT} (below diagonal) and probability Φ_{PT} value is > 0 (above).
Comparison in yellow is not statistically significant.

Allocation to Collection

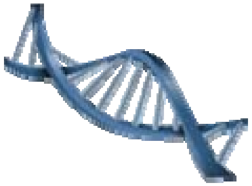
allocated to	GRSM	MO7	MO10	NY3	OH22	OH21	PA02	PA10	SHEN5	SHEN6	SHEN7
GRSM	4	0	0	0	0	0	0	0	0	0	0
MO7	0	8	0	0	0	0	0	0	0	0	0
MO10	0	0	8	0	0	0	0	0	0	0	0
NY3	0	0	0	7	0	0	0	0	0	0	0
OH22	0	0	0	0	8	0	0	1	0	0	0
OH21	0	0	0	0	0	8	0	0	0	0	0
PA02	0	0	0	0	0	0	8	1	0	0	0
PA10	0	0	0	1	0	0	0	6	0	0	0
SHEN5	0	0	0	0	0	0	0	0	8	0	0
SHEN6	0	0	0	0	0	0	0	0	0	6	0
SHEN7	0	0	0	0	0	0	0	0	0	0	4
Correct Assignment (%)	100.0%	100.0%	100.0%	87.5%	100.0%	100.0%	100.0%	75.0%	100.0%	100.0%	100.0%

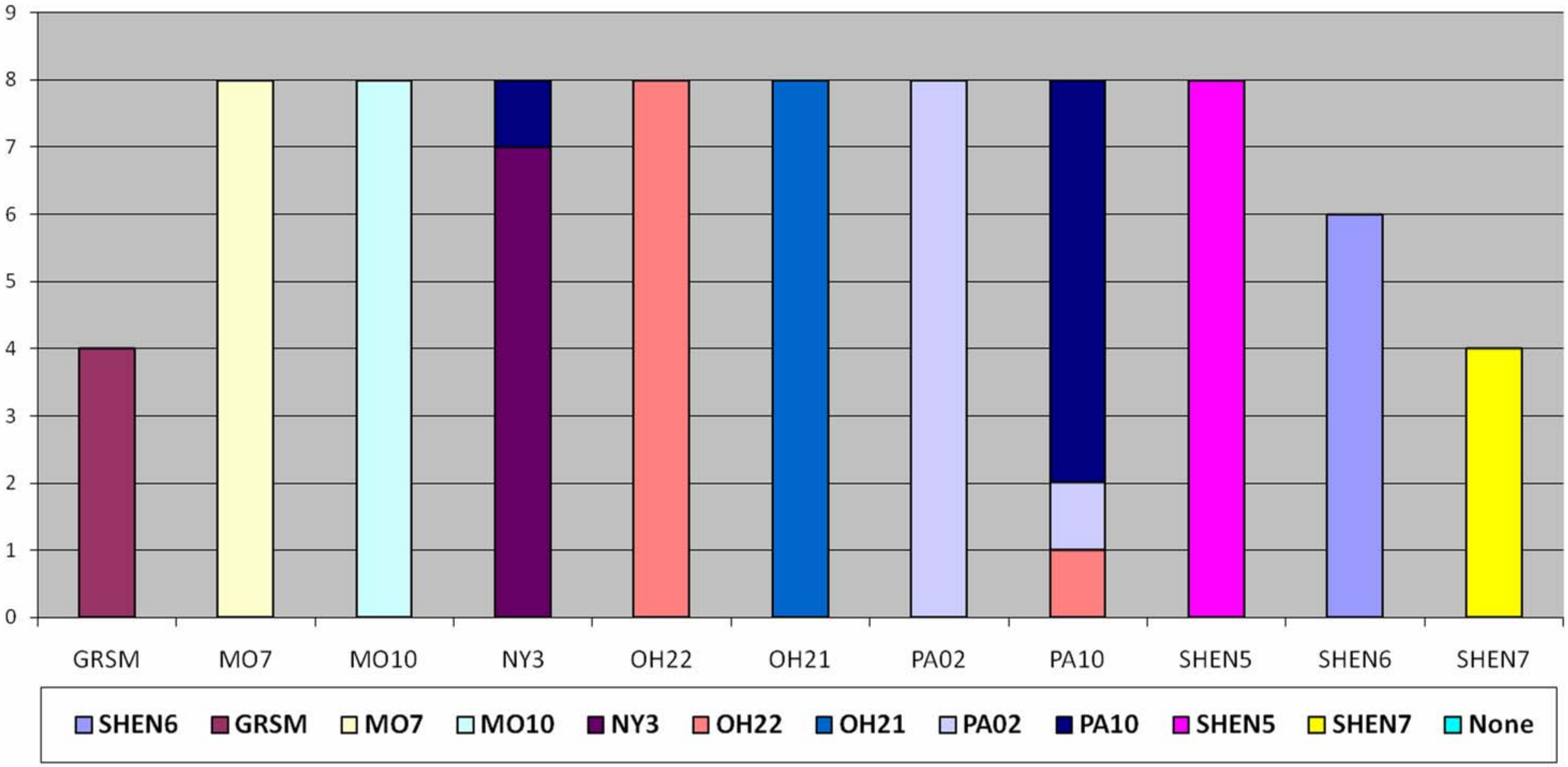
Overall Assignment Success: 96.2%

Incorrect assignments are distributed vertically and are highlighted in yellow.

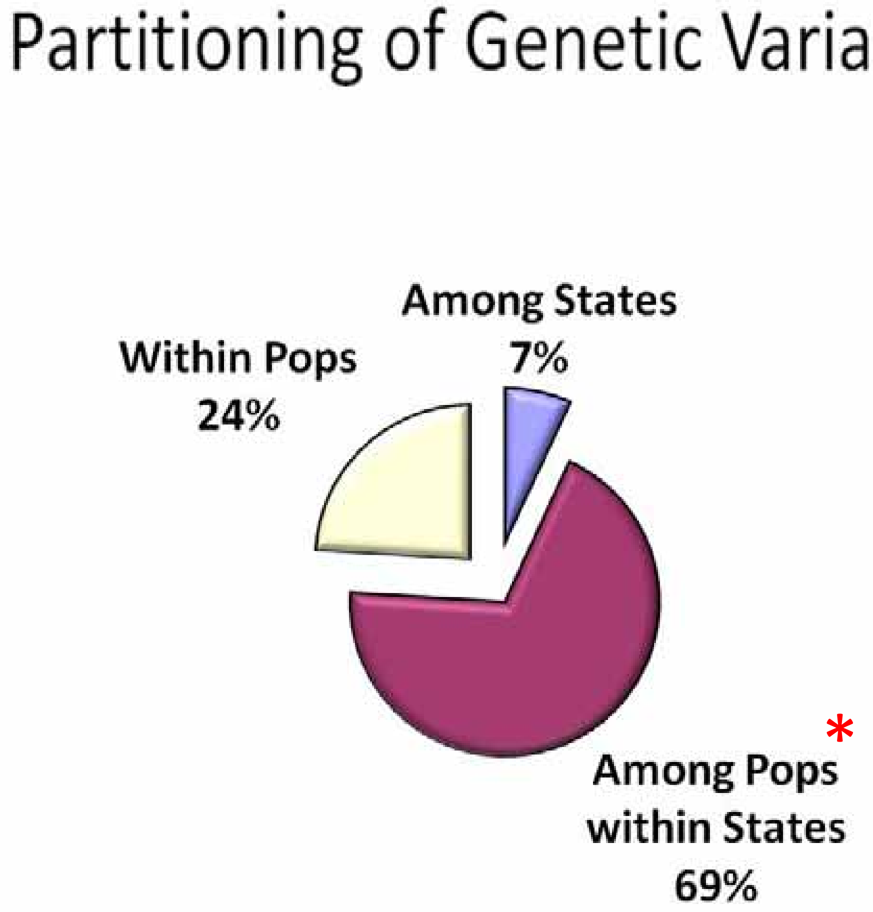
Microsatellite DNA

Allocation Graphic





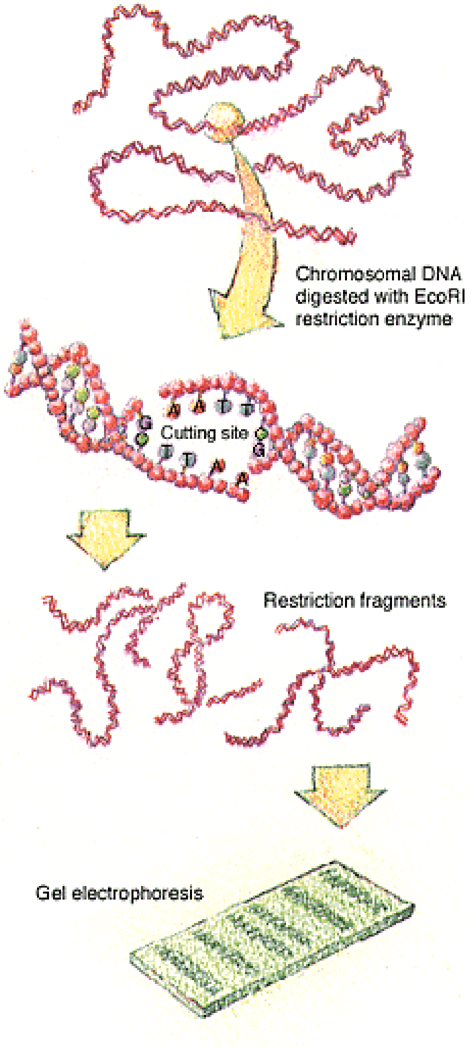
AMOVA (*A*nalysis of *MO*lecular *VA*riance)



* **Preliminary results (AMOVA, pair-wise Φ_{PT}, and assignment tests), needs validation with larger dataset**

Genetic Markers - AFLP

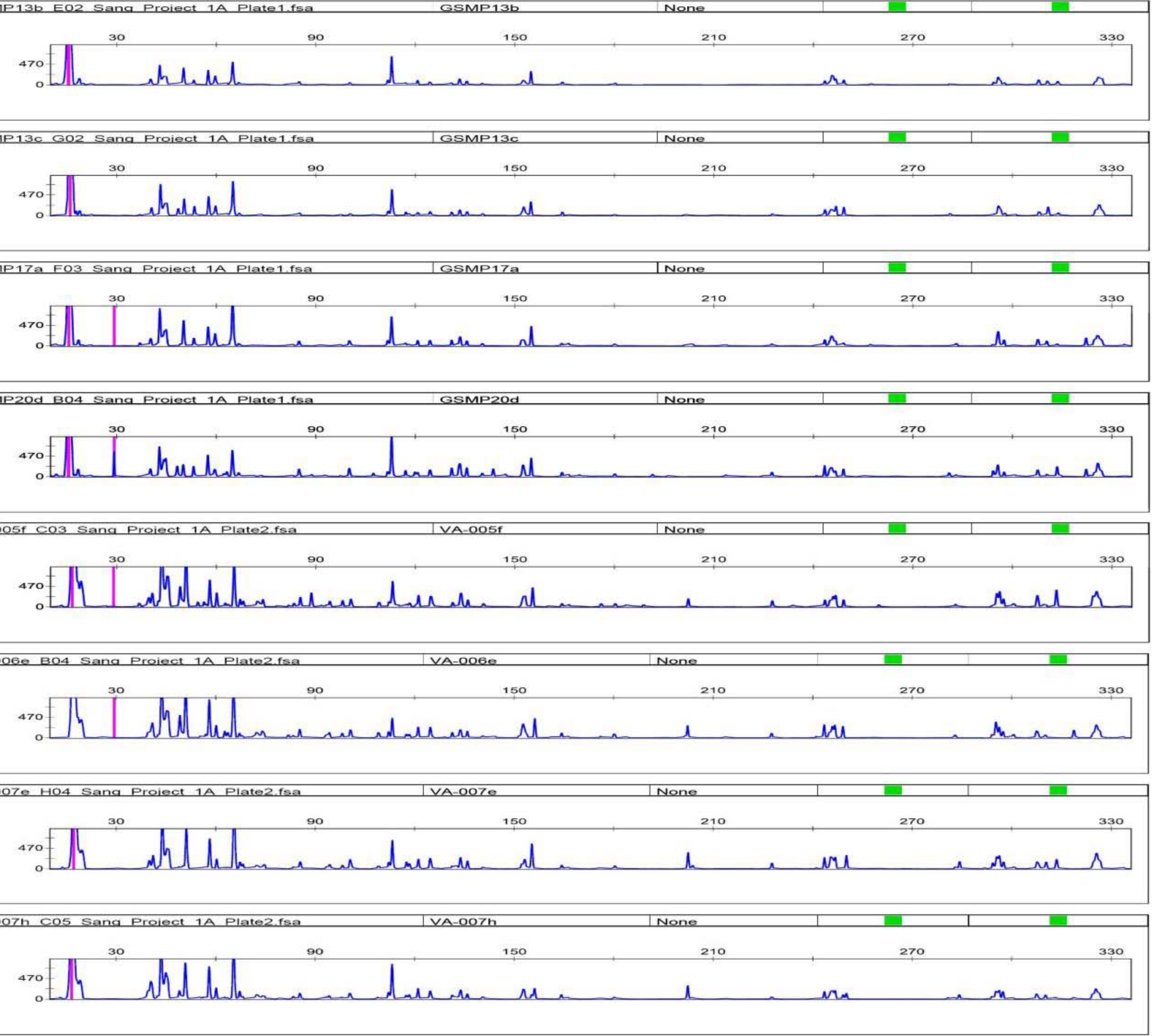
- AFLP = “Amplified Fragment Length Polymorphism”
  - – No sequence information required
  - – Highly repeatable
- Two step process:
  1. Cut DNA with restriction enzymes (primers), bind adapters to end of fragments
  2. Use PCR to amplify fragments
- Polymorphisms from:
  - – Sequence variation in and around restriction sites
  - – Insertions or deletions in amplified fragments
- High numbers of dominant (presence/absence) allelic markers produced



Genetic Markers – AFLP

David DeViney, M.S. Student, ASU

- We tested 3 primer pairs for American ginseng:
  1) “Pqu1-A”: EcoRI-ACT, MseI-CAA
  2) “Pqu4-H”: EcoRI-ACC, MseI-CTT
  3) “Pqu7-A”: EcoRI-AGG, MseI-CAA
- Number of allelic markers generated for American ginseng:
  - 378 with 3 primers (1+2+3)
  - 240 with 2 primers (1+2)



GRSM

GRSM

GRSM

GRSM

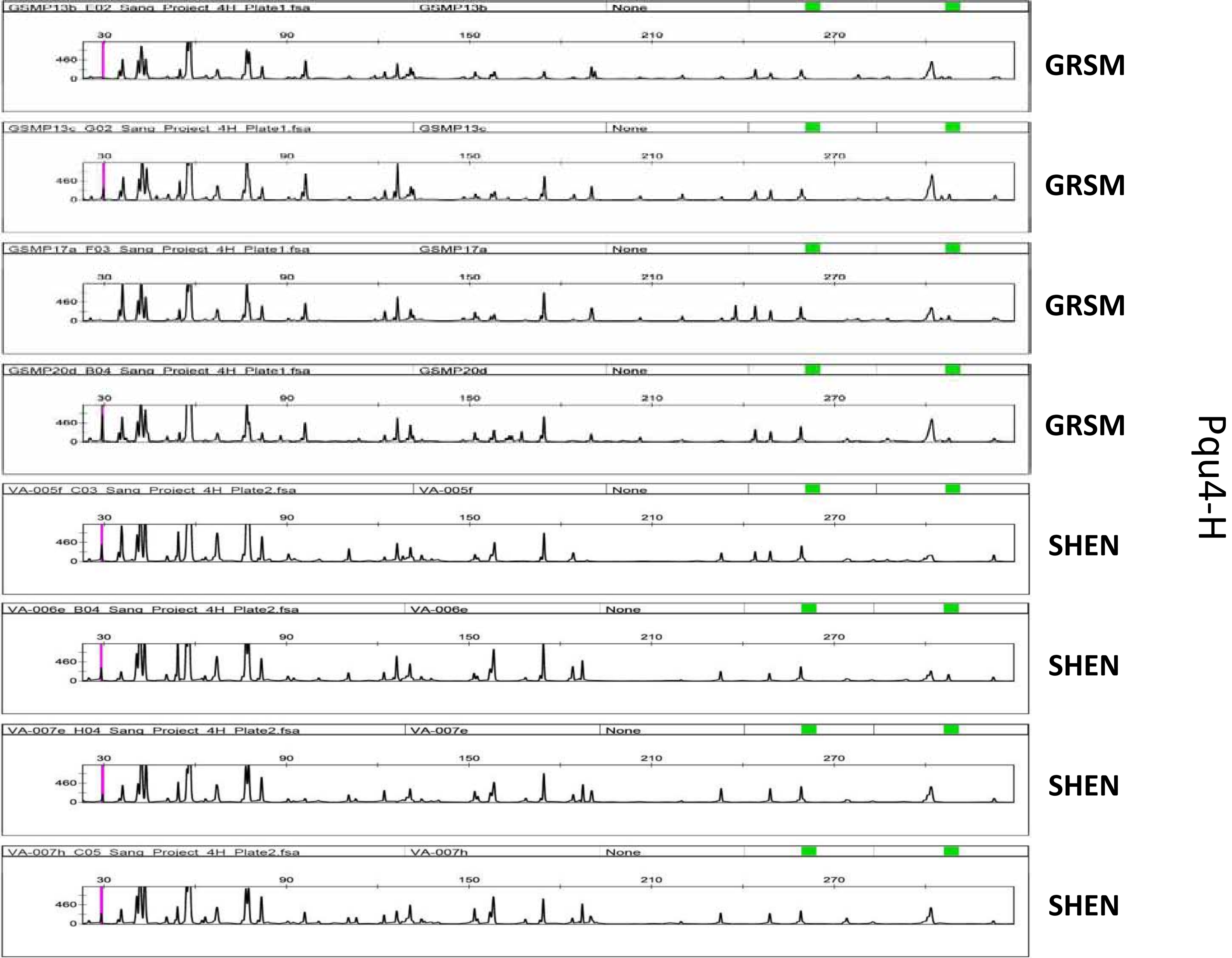
SHEN

SHEN

SHEN

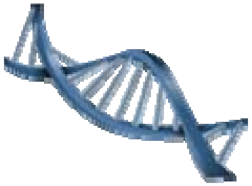
SHEN

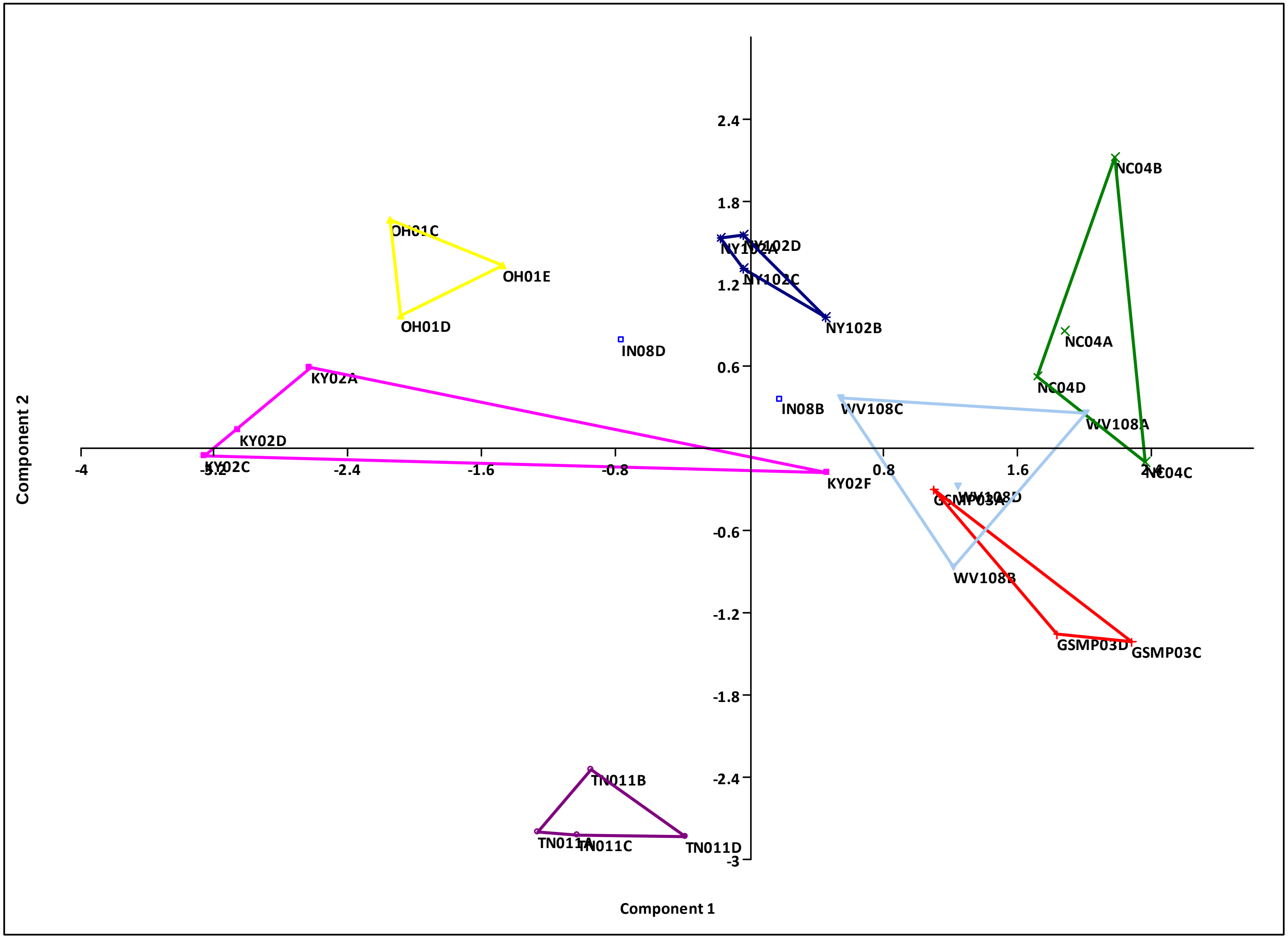
PquI-A


GSMP13b_E02_Sang_Project_4H_Plate1.fsa
GSMP13b
None
GRSM
GSMP13c_G02_Sang_Project_4H_Plate1.fsa
GSMP13c
None
GRSM
GSMP17a_F03_Sang_Project_4H_Plate1.fsa
GSMP17a
None
GRSM
GSMP20d_B04_Sang_Project_4H_Plate1.fsa
GSMP20d
None
GRSM
VA-005f_C03_Sang_Project_4H_Plate2.fsa
VA-005f
None
SHEN
VA-006e_B04_Sang_Project_4H_Plate2.fsa
VA-006e
None
SHEN
VA-007e_H04_Sang_Project_4H_Plate2.fsa
VA-007e
None
SHEN
VA-007h_C05_Sang_Project_4H_Plate2.fsa
VA-007h
None
SHEN
30
90
150
210
270
460
0
Pqu4-H

Population discrimination – AFLP

Principal Components Analysis





Remaining tasks:

- Winter-Spring 2009
  - Develop new ginseng distribution models
  - Survey all genetic samples for AFLP and microsatellite variation

- Early summer 2009
  - Collect additional genetic samples

- Late summer-early fall 2009
  - Data analysis
    - Assess phylogeographic and population structure
    - Diversity within/among sites by land ownership
    - Genetic diversity compared to habitat distribution/quality

Acknowledgements

- Marcus Springmann (USGS) – field data collection, genetic laboratory assistance
- Chris Walter (USGS) – field data collection, assistance with GIS data compilation
- Pat Ford (FWS), state ginseng coordinators – contacts with land owners
- Bob Beyfuss, Eric Burkhart – contacts with private and state landowners
- Numerous landowners and land managers who provided access to their property for sampling