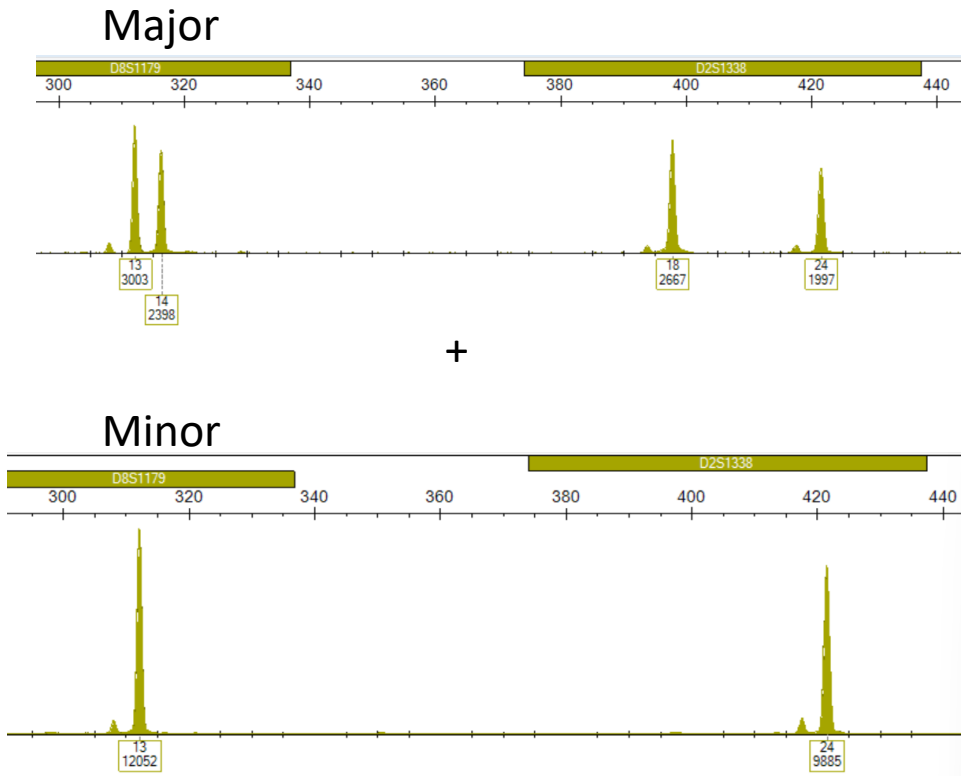
The Evaluation of the ForenSeq™ MainstAY Kit for Mixed Forensic DNA Samples

Damani Johnson, BS.*, Lucio Avellaneda, BS., Rachel Houston, PhD., Tim Kalafut, PhD.

Department of Forensic Science
Sam Houston State University
Huntsville, TX 77340

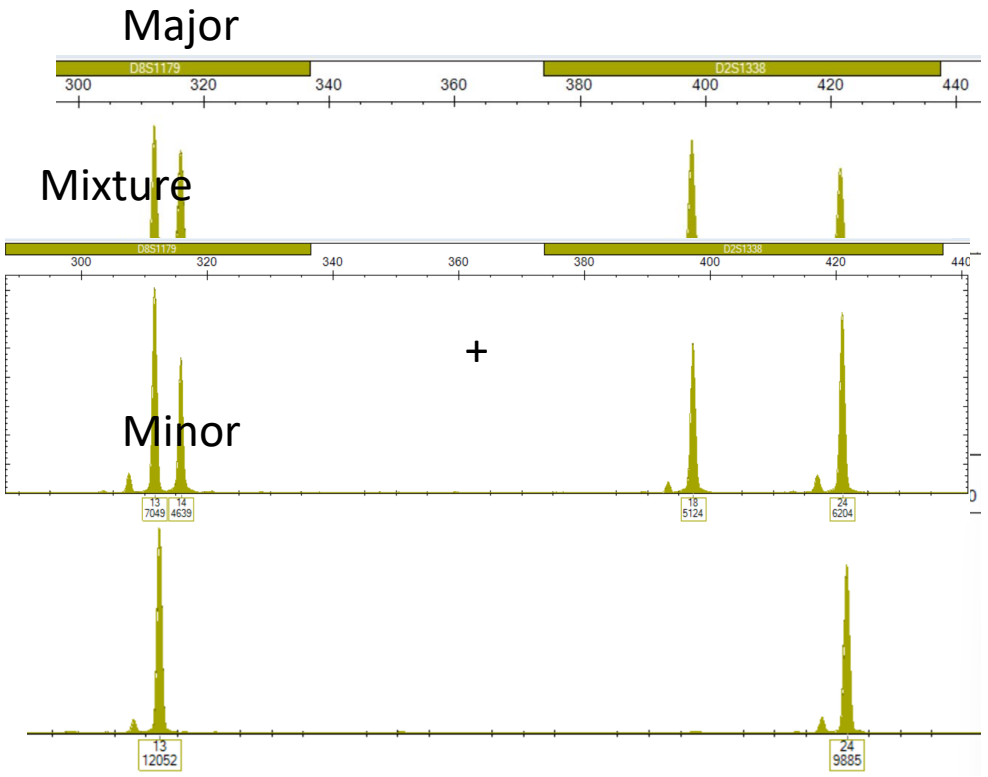
STR Analysis

- Common difficulties on CE:
  - Multiplexing limited by dye space and marker length
  - Degradation/low quant
  - Complex mixtures with allele masking



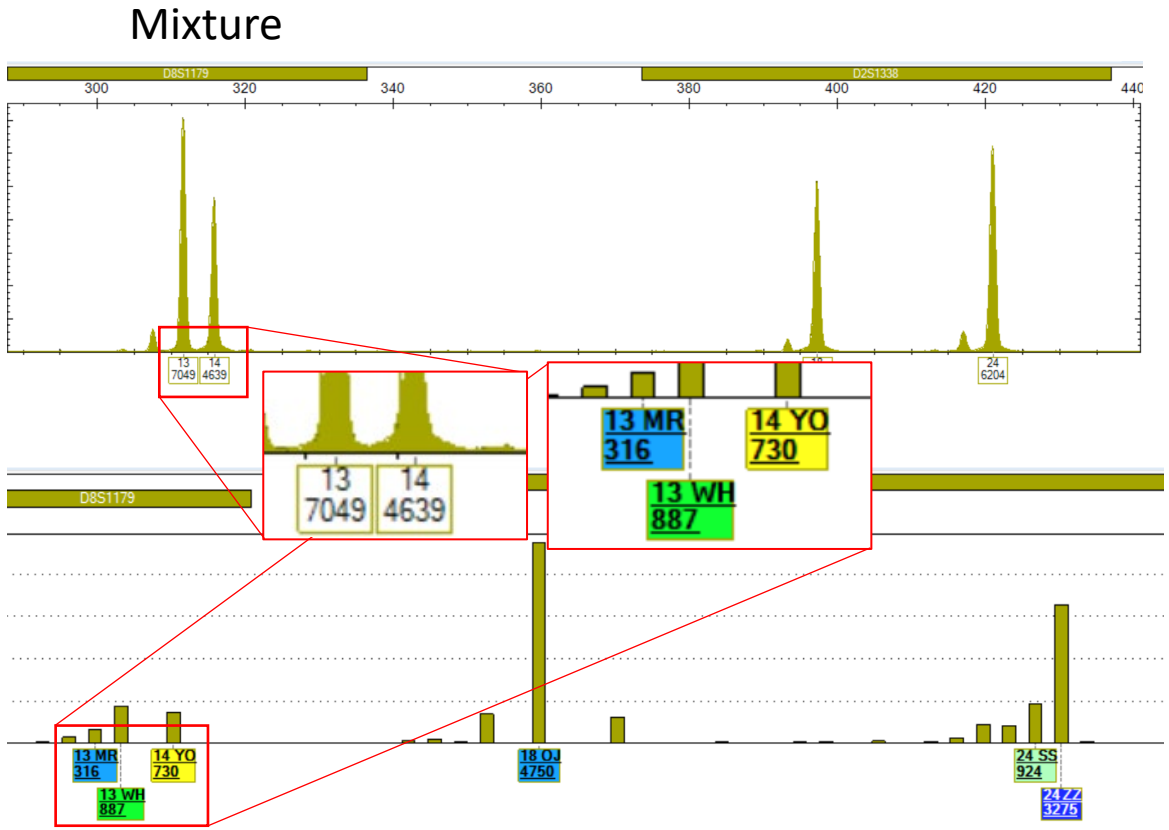
STR Analysis

- Common difficulties on CE:
  - Multiplexing limited by dye space and marker length
  - Degradation/low quant
  - Complex mixtures with allele masking



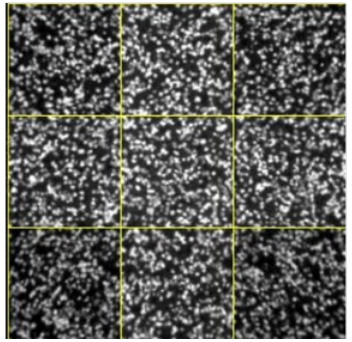
STR Analysis

- Next-Generation Sequencing
  - More markers
  - Greater sensitivity
  - Isoalleles
    - Distinguish alleles of the same length by sequence differences
- MixtureAce (NicheVision)
  - Displays sequence data to mimic CE electropherogram format
  - Separates isoallele peaks



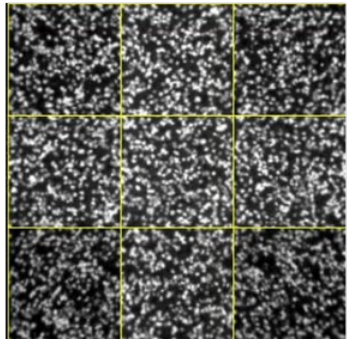
MiSeq FGx Sequencing System (Verogen)

- Simultaneously detects all samples and markers using a flow cell
  - Generates clusters on flow cell
    - Each cluster represents a Distinct PCR Target (DPT)
  - Sequences each cluster as a read
    - Read counts ≈ RFUs
  - Software assigns sequence results to individual samples



Plexity

- Flow cell has a physical capacity determined by surface area
- MiSeq FGx Reagent Micro kit allows ~5 million reads
- ForenSeq MainstAY
  - 27 Autosomal STRs
  - 25 Y-STRs
  - Amelogenin
  - Optional SE33 addition
- Maximum recommended batch of 96 samples

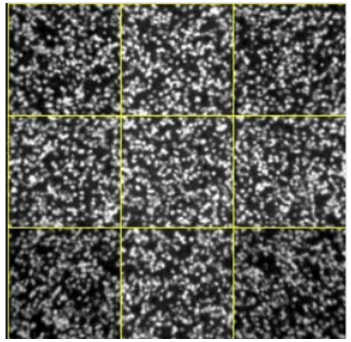


Plexity

- Flow cell has a physical capacity determined by surface area
- MiSeq FGx Reagent Micro kit allows ~5 million reads
- ForenSeq MainstAY
  - 27 Autosomal STRs
  - 25 Y-STRs
  - Amelogenin
  - Optional SE33 addition
- Maximum recommended batch of 96 samples

$$\frac{5{,}000{,}000\ total\ reads}{(29\ aSTRs * 2\ possible\ alleles) + 25\ YSTRs} = 60{,}241\ reads\ per\ sample$$

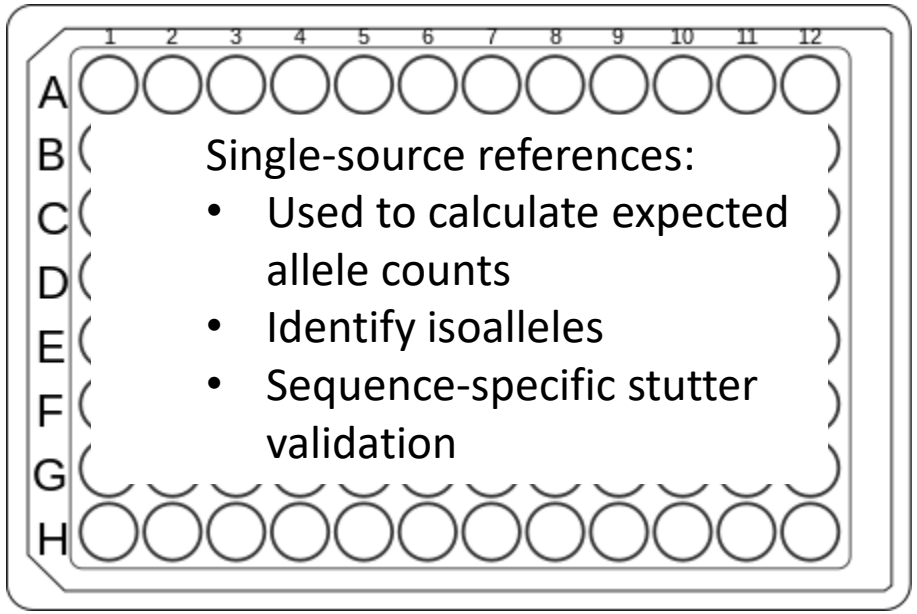
$$\frac{60{,}241\ reads\ per\ sample}{96\ samples} = 627\ reads\ per\ DPT$$

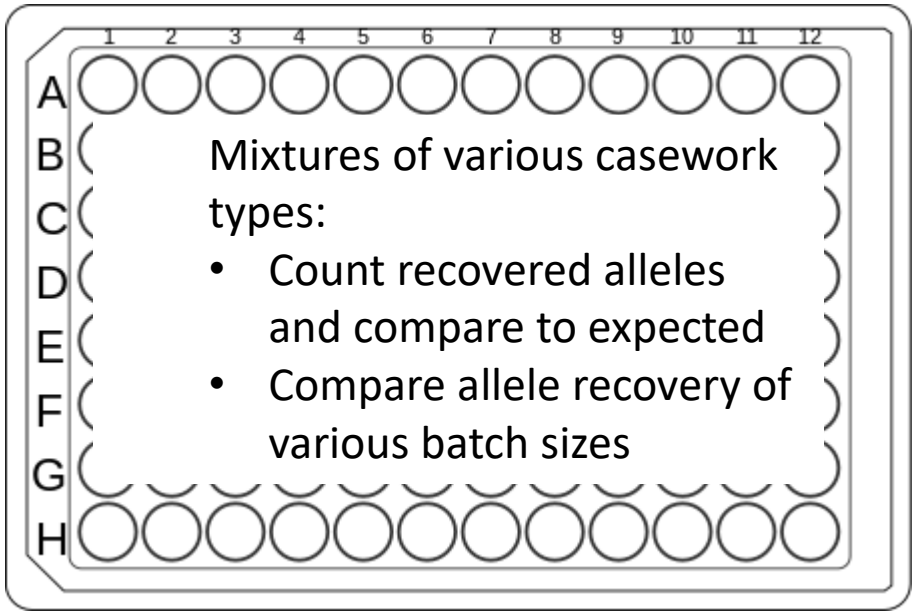


More contributors = More DPTs = Reduced read coverage

Experimental Setup

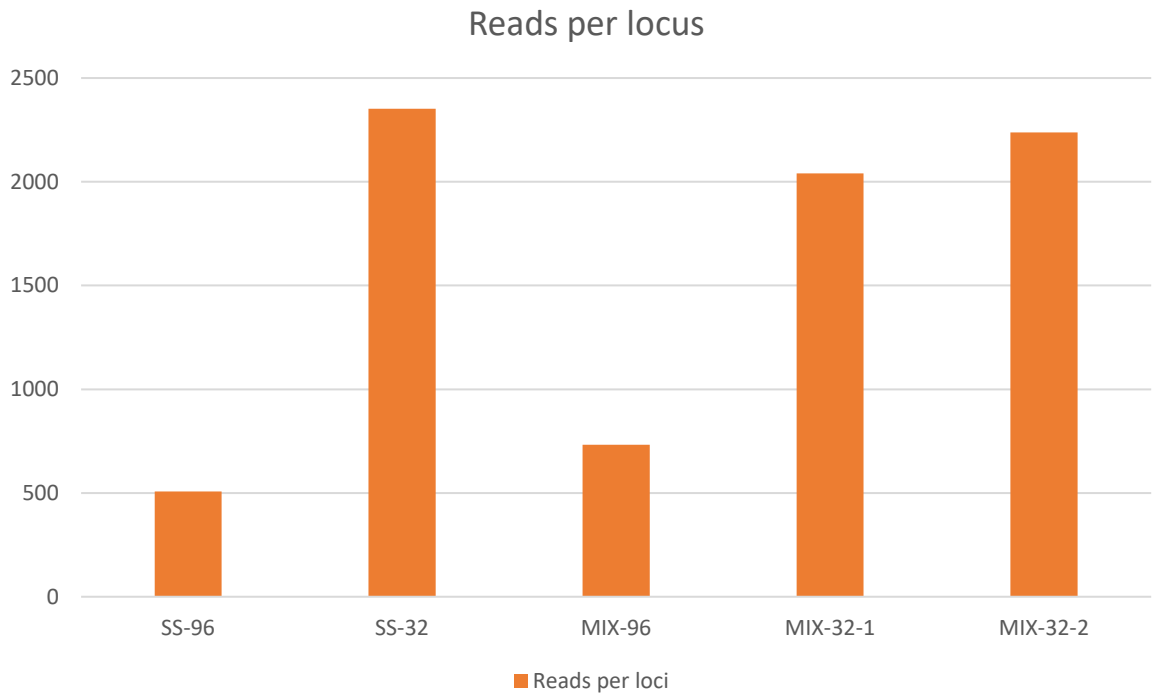
- ForenSeq MainstAY SE Kit
- Two 96 sample plates:





Single-Source References

- Two batches of single-source reference samples
- 96 sample batch
  - Low read counts
  - NGS prep experience may play a role
- 32 sample batch
  - Higher read counts and allele recovery
  - Only 3 samples fully re-prepped, remaining were rerun from NLP



Mixture Plate

- 96 samples of various types encountered in casework:
  - 2 single-source dilution series from 1ng to 8pg
  - 15 challenging samples consisting of bone and tissue (+ Positive control)
  - 16 three-person mixtures of various contributor ratios at 1ng input
  - 4 dilution series of three-person mixtures at 500pg, 125pg, 62 pg, and 31pg
  - 16 four-person mixtures of various contributor ratios at 1ng input
  - Extreme major/minor ratios mimicking differential extractions over time
- Ran multiple batches of different sample numbers
  - All 96
  - 32 – only three-person mixtures (5-8)
  - 32 – Dilution series and extreme ratio (1, 2, 11, 12)

	1	2	3	4	5	6	7	8	9	10	11	12
Type	Single Source Dilution		Challenging (Bone and Tissue)		3p 1ng		3p dilutions		4p 1ng		Mock Epithelial Fraction	Mock Sperm Fraction
A	1ng	1ng			**20:20:1**	**6:6:1**	0.5ng 10:10:1	0.063ng 10:10:1	**20:20:20:1**	**6:6:6:1**	10:1	1:50
B	0.5ng	0.5ng			**20:10:1**	**6:3:1**	0.5ng 10:1:1	0.063ng 10:1:1	**20:20:1:1**	**6:6:3:1**	20:1	1:20
C	0.25ng	0.25ng			**20:1:1**	**6:1:1**	0.5ng 10:6:2	0.063ng 10:6:2	**20:1:1:1**	**6:6:1:1:**	40:1	1:10
D	0.125ng	0.125ng			**10:10:1**	**3:3:1**	0.5ng 1:1:1	0.063ng 1:1:1	**10:10:10:1**	**6:3:1:1**	50:1	1:5
E	0.063ng	0.063ng			**10:6:1**	**3:1:1**	0.125ng 10:10:1	0.031ng 10:10:1	**10:10:5:1**	**6:1:1:1**	80:1	1:2
F	0.031ng	0.031ng			**10:3:1**	**2:2:1**	0.125ng 10:1:1	0.031ng 10:1:1	**10:10:1:1**	**3:3:3:1**	100:1	1:1
G	0.016ng	0.016ng			**10:2:1**	**2:1:1**	0.125ng 10:6:2	0.031ng 10:6:2	**10:5:1:1**	**3:3:1:1**	150:1	10:1
H	0.008ng	0.008ng		+	**10:1:1**	**1:1:1**	0.125ng 1:1:1	0.031ng 1:1:1	**10:1:1:1**	**1:1:1:1**	200:1	50:1

224 Contributors Total
268 Reads per DPT

Sensitivity

- Two single-source dilution series
  - 1ng, 500pg, 250pg, 125pg, 63pg, 31pg, 16pg, 8pg
- 96 sample batch dropped alleles at higher concentrations
- 32 sample batch reduced number of dropped alleles
  - Compared to 96 and CE
- Dependent on analytical threshold as well
  - Used 10 read threshold
  - Recommended 1.5% of locus reads, 650 minimum per locus

Dropped Alleles

Sample input	1.00ng	0.500ng	0.250ng	0.125ng	0.063ng	0.031ng	0.016ng	0.008ng
Sample 1 CE	0	0	0	0	0	1	9	23
Sample 1 NGS 96	0	0	0	0	5	10	14	29
Sample 1 NGS 32	0	0	0	0	0	5	8	18
Sample 2 CE	0	0	0	0	4	17	23	35
Sample 2 NGS 96	0	1	5	6	9	36	22	33
Sample 2 NGS 32	0	1	3	3	5	32	18	24

Key	0	5	10	20	30	40

Extreme Major/Minor

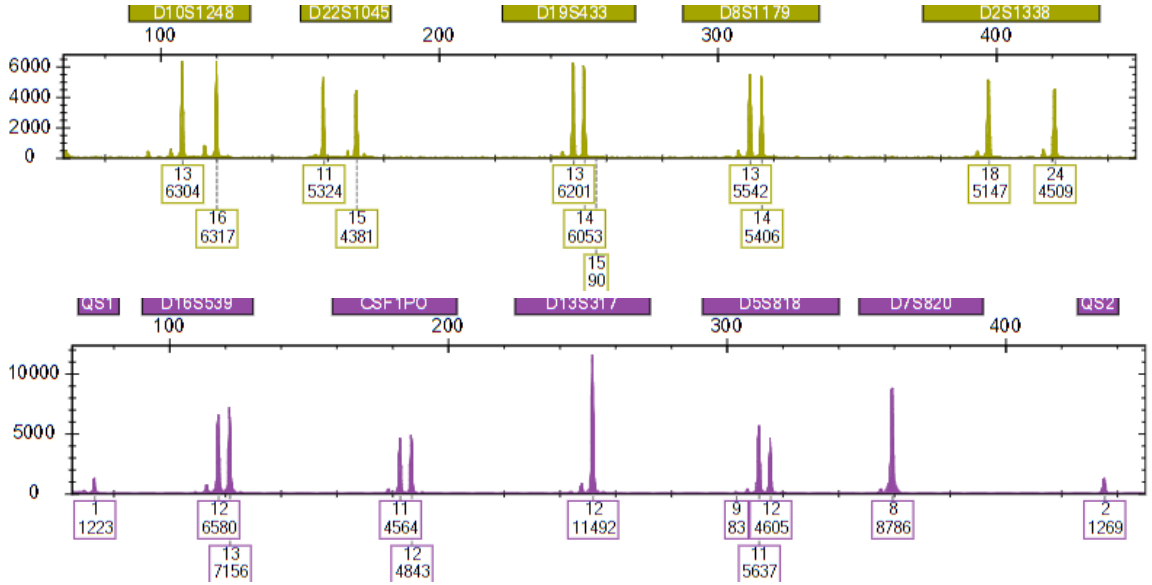
- Mimic differential epithelial/sperm fractions
- 200:1 ratio

Allele Counts	CE	NGS 96	NGS 32
Minor (1)	37	37	37
Major (200)	39	41	41
Combined Expected	61	66	66
Combined Observed	46 (75.4%)	57 (86.4%)	59 (89.4%)
Unshared Minor Alleles	9 (24.3%)	16 (43.2%)	18 (48.6%)

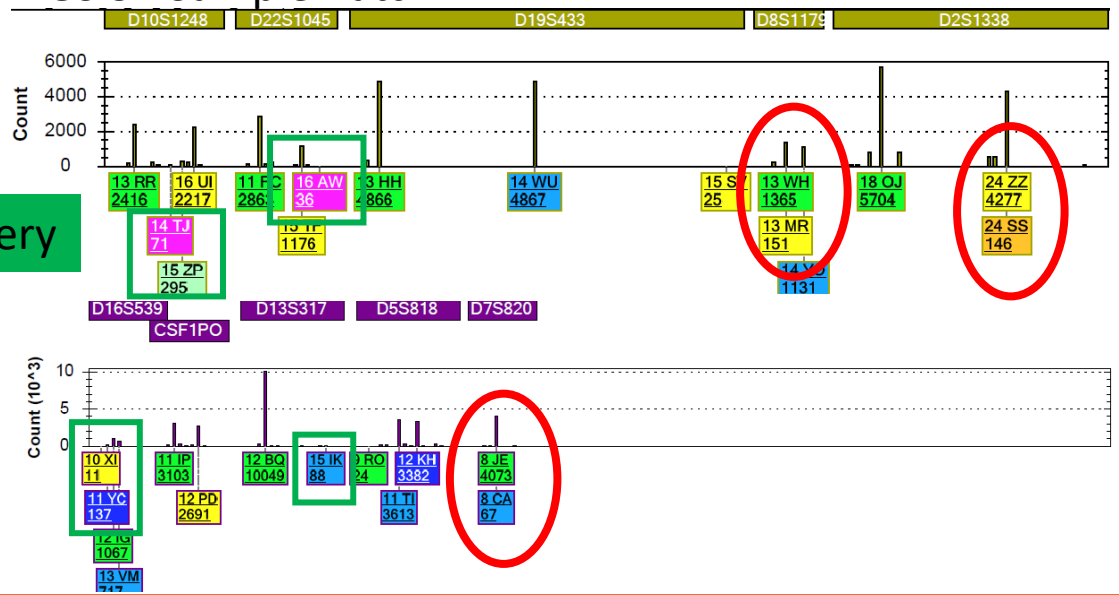
Minor allele recovery

Isoalleles

CE



NGS 32 Sample Batch

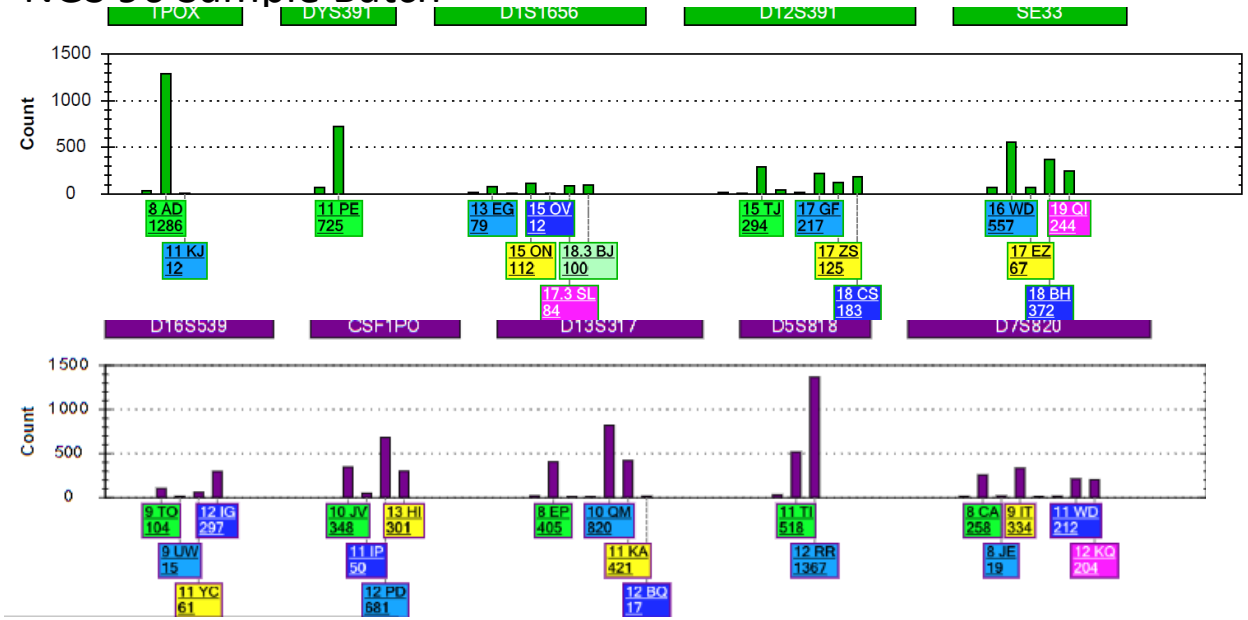


Three-Person Mixture

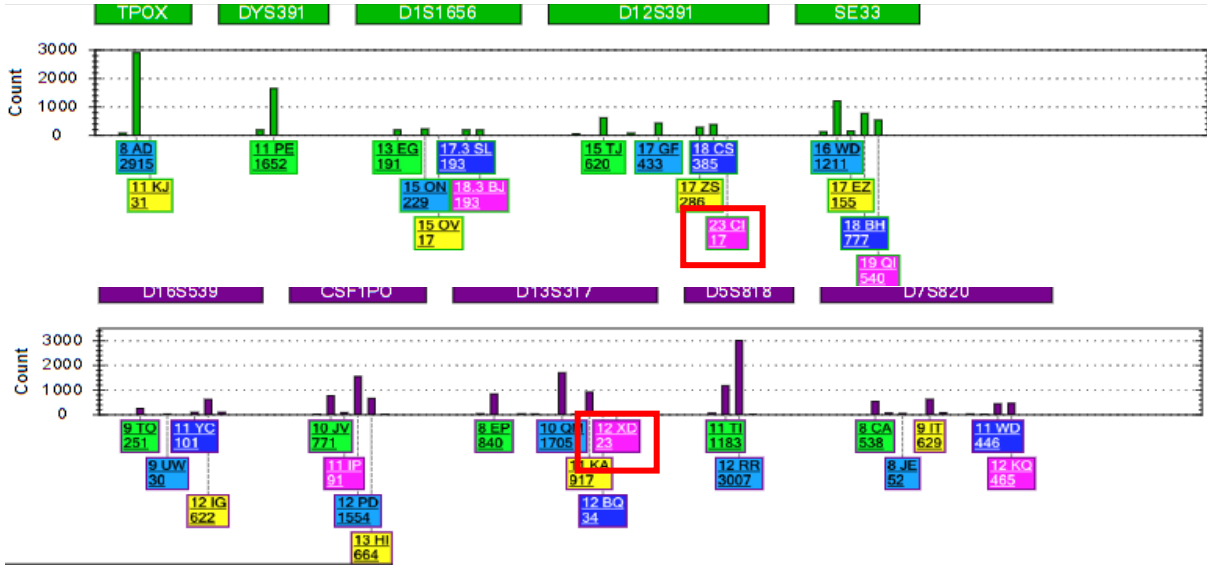
- Three-person mixtures at 1ng input
- 20:20:1 mixture proportion

Allele Counts	CE 24Plex	NGS 24Plex	NGS All loci
Major 1	41	41	51
Major 2	42	42	55
Minor	39	42	54
Combined Expected	85	94	118
Combined Observed (96)	79 (92.9%)	86 (91.5%)	107 (90.7%)
Combined Observed (32)	-	88 (93.6%)	111 (94.1%)
Minor Alleles Observed (96)	12 (30.8%)	14 (33.3%)	17 (31.5%)
Minor Alleles Observed (32)	-	16 (38.1%)	21 (38.8%)

NGS 96 Sample Batch

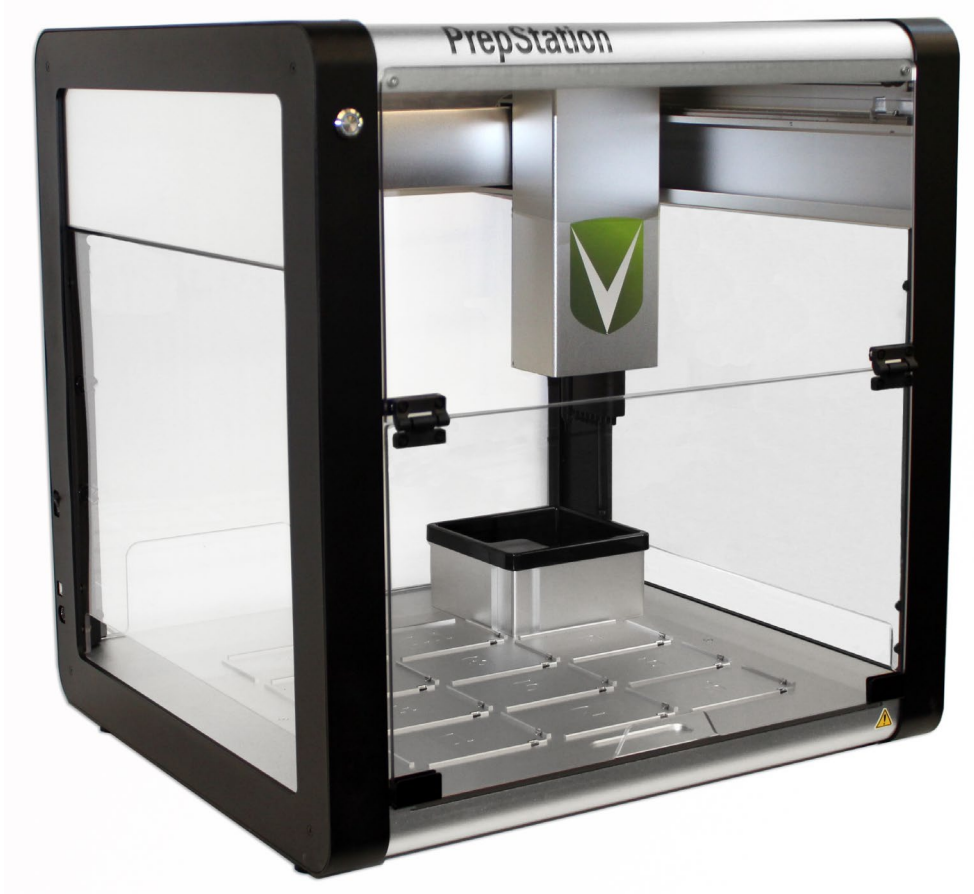


NGS 32 Sample Batch



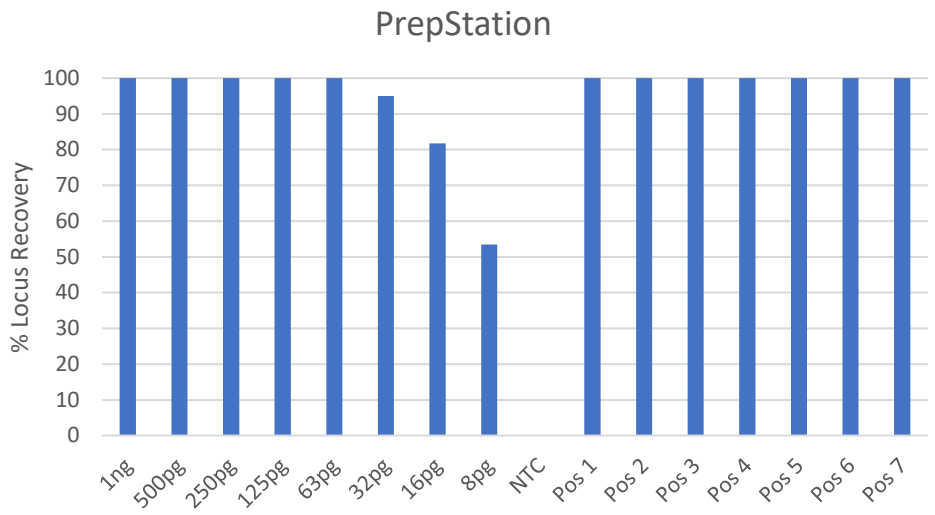
Automation

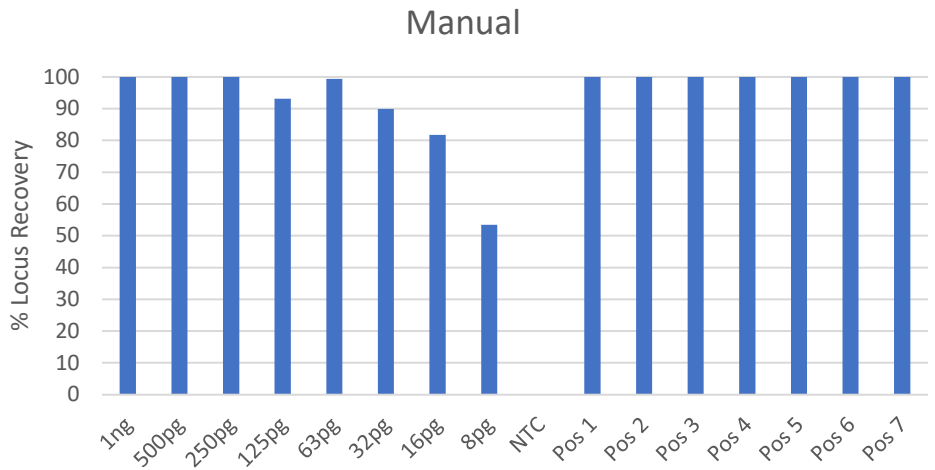
- NGS is extremely labor-intensive
  - Requires several bead-based purification and normalization steps
    - Can be time sensitive
  - Even experienced pipetters may need time to become comfortable with the workflow
- Verogen PrepStation
  - Inexpensive laboratory automation robot
  - Verogen designed protocols for their kits
  - Saves much manual labor, simple to set up and run



Automation

- NGS is extremely labor-intensive
  - Requires several bead-based purification and normalization steps
    - Can be time sensitive
  - Even experienced pipetters may need time to become comfortable with the workflow
- Verogen PrepStation
  - Inexpensive laboratory automation robot
  - Verogen designed protocols for their kits
  - Saves much manual labor, simple to set up and run





Conclusions

- Plexity should be considered when running NGS mixtures
- Smaller sample batches result in higher read counts and increased allele recovery
- Both 96 and 32 sample NGS batches saw increased recovery of minor alleles compared to CE
- Isoalleles and additional markers in MainstAY increase discrimination power
- Automation can increase NGS workflow efficiency

- Future Work:
  - Continue sorting through mountains of data
    - Y-STRs
  - Run different batch combinations with remaining flow cells
  - Develop “Intelligent Batching” recommendations
  - Develop sequence-specific stutter model
    - May need to pool community stutter values
  - Probabilistic genotyping for NGS

Acknowledgements

Center for Advanced Research in Forensic Science (CARFS)

Verogen

NicheVision

Lucio and Cesar

SHSU Department of Forensic Science

Questions?

Damani Johnson

dtj012@shsu.edu