signature science LLC

A Proteos Program Retrospective: Establishing Protein Sequencing for Forensic Analysis

Curt Hewitt, PhD

Leader, Human Identification R&D
Signature Science Center for Advanced Genomics

signature® science LLC

Outline

- IARPA Proteos Program Overview
- Technical Approach
- Project Highlights
- Limitations
- Future Research Considerations

IARPA PROTEOS Program

- Intelligence Advanced Research Projects Activity (IARPA)
- Seeks to Develop Novel Methods for Human Identification by Protein Sequencing of Touch Samples
- Develop Methods to Efficiently Co-Extract Protein and DNA in Parallel

Office of the Director of National Intelligence

IARPA

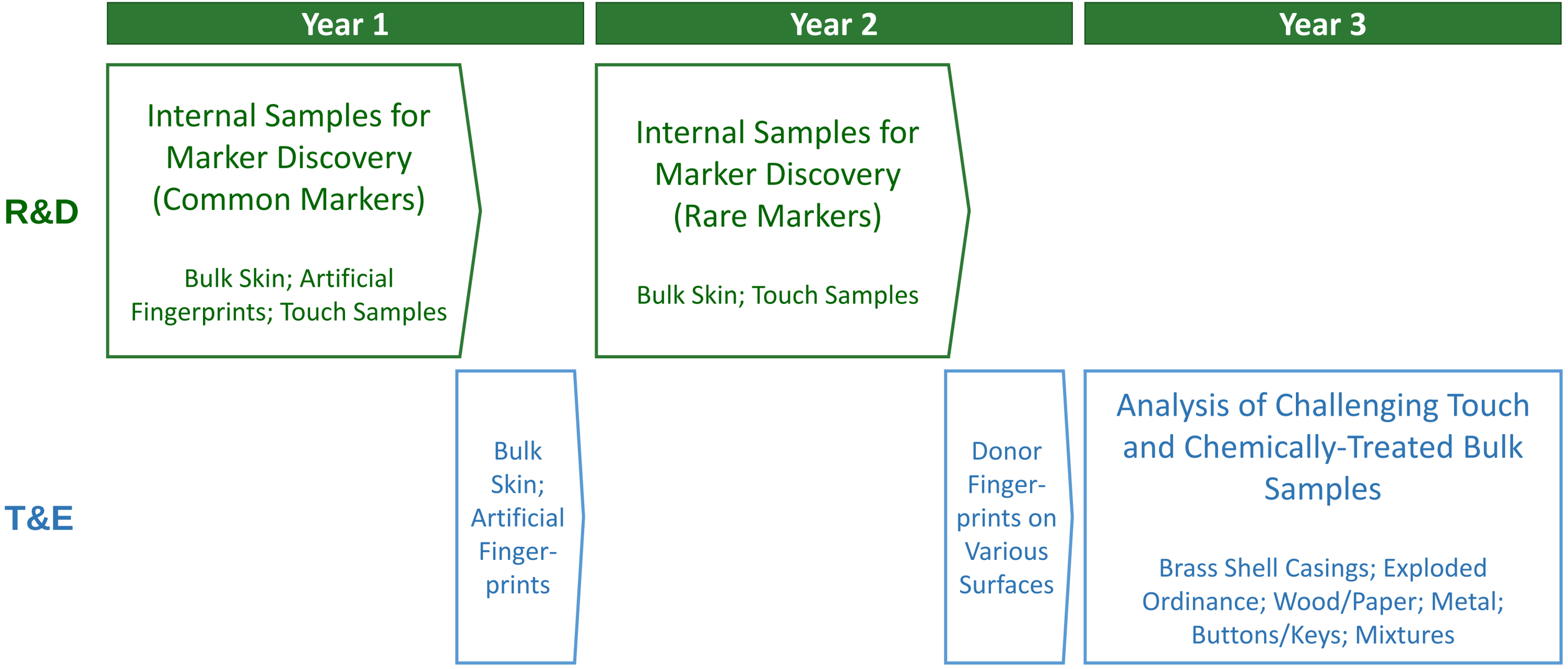
BE THE FUTURE

- SigSci teamed with UNT CHI and The Ohio State University for Proteos

THE OHIO STATE UNIVERSITY

CAMPUS CHEMICAL INSTRUMENT CENTER

IARPA PROTEOS Program

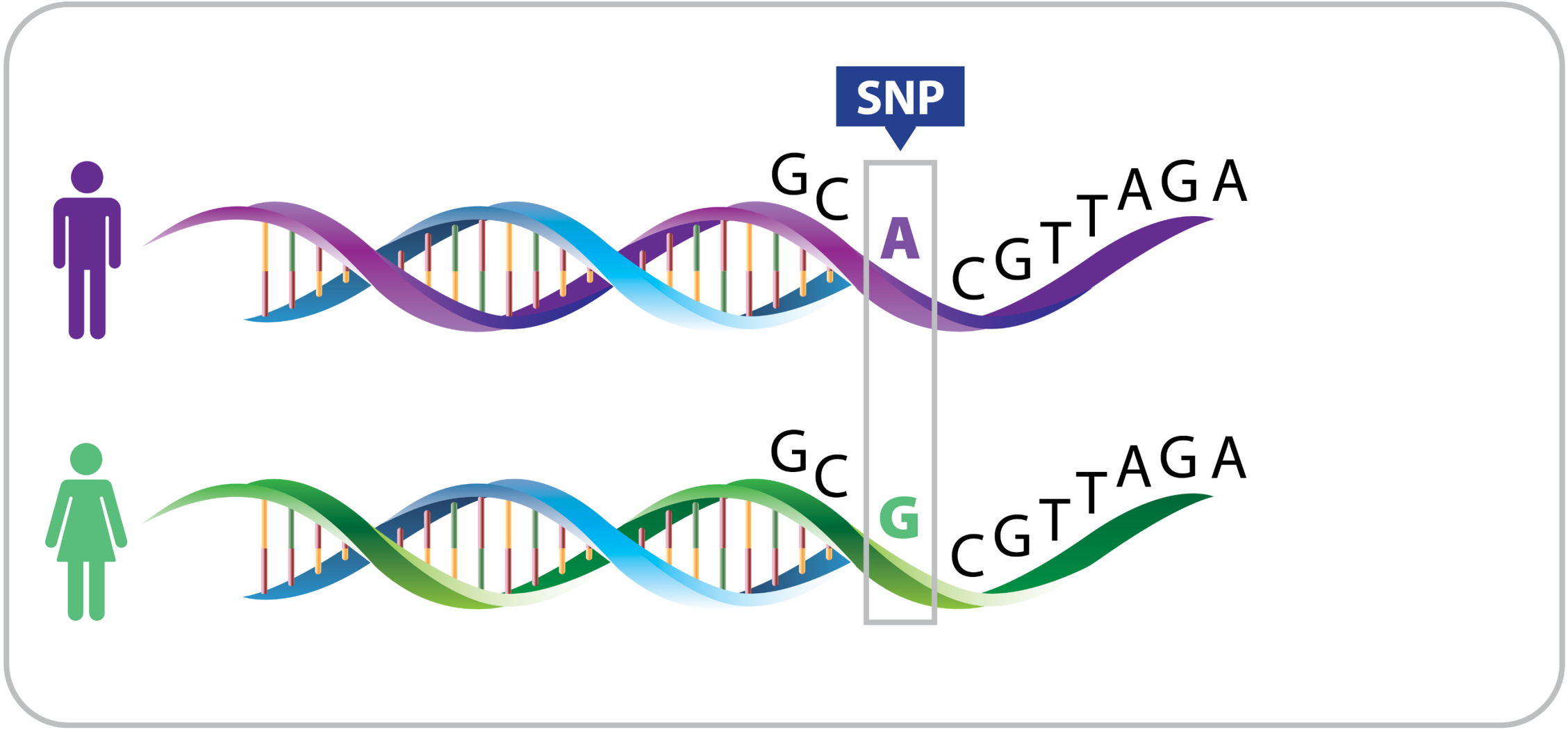


Challenges in Human Forensic Identification

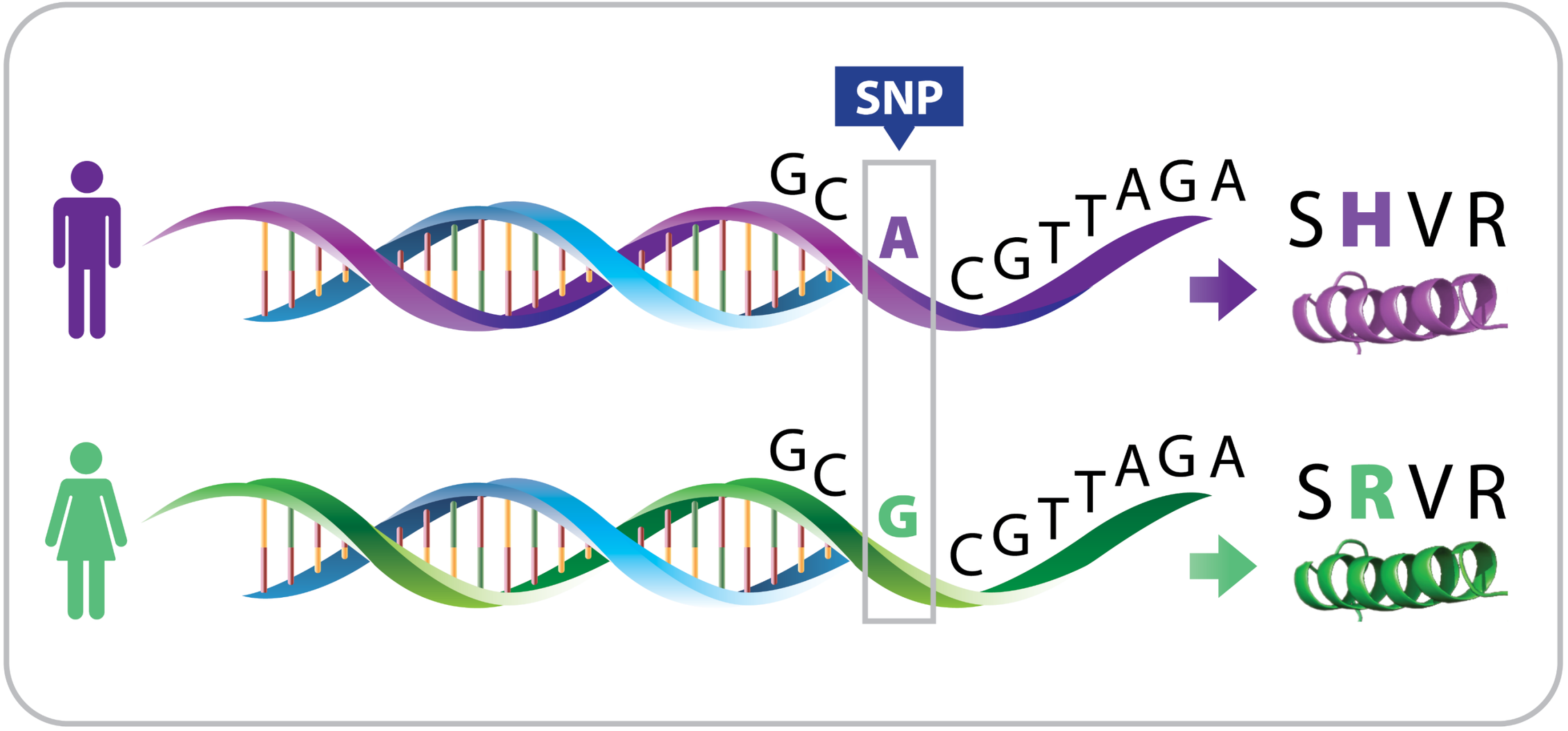
Components of Touch Samples

- Ridge Pattern
- Sebaceous Fluids (Oils, Sweat)
- Extracellular DNA
- Keratinized Skin Cells

Current State of the Art – DNA Forensics



Current State of the Art – DNA Forensics



Thought Experiment…

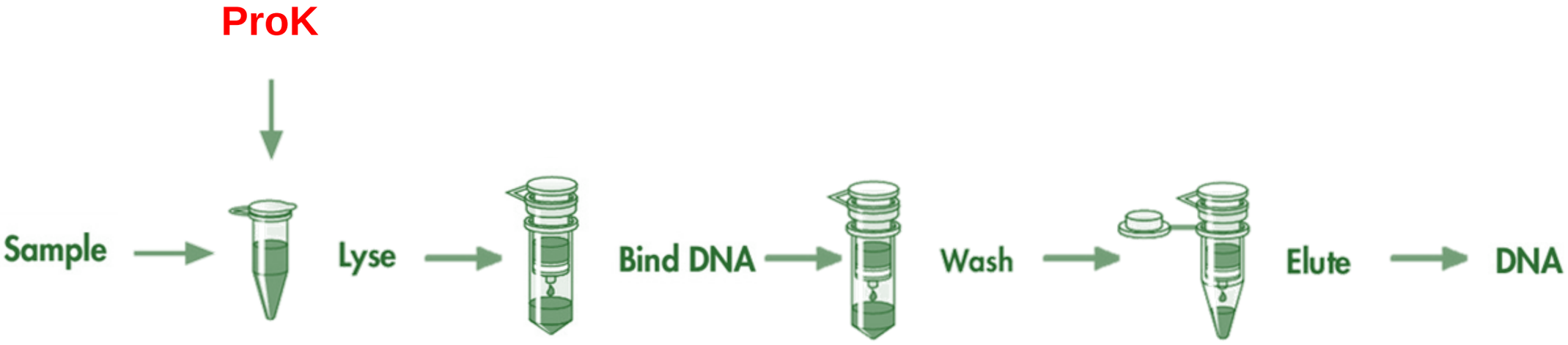
Let’s assume:

1. Protein sequencing has utility for forensic identification

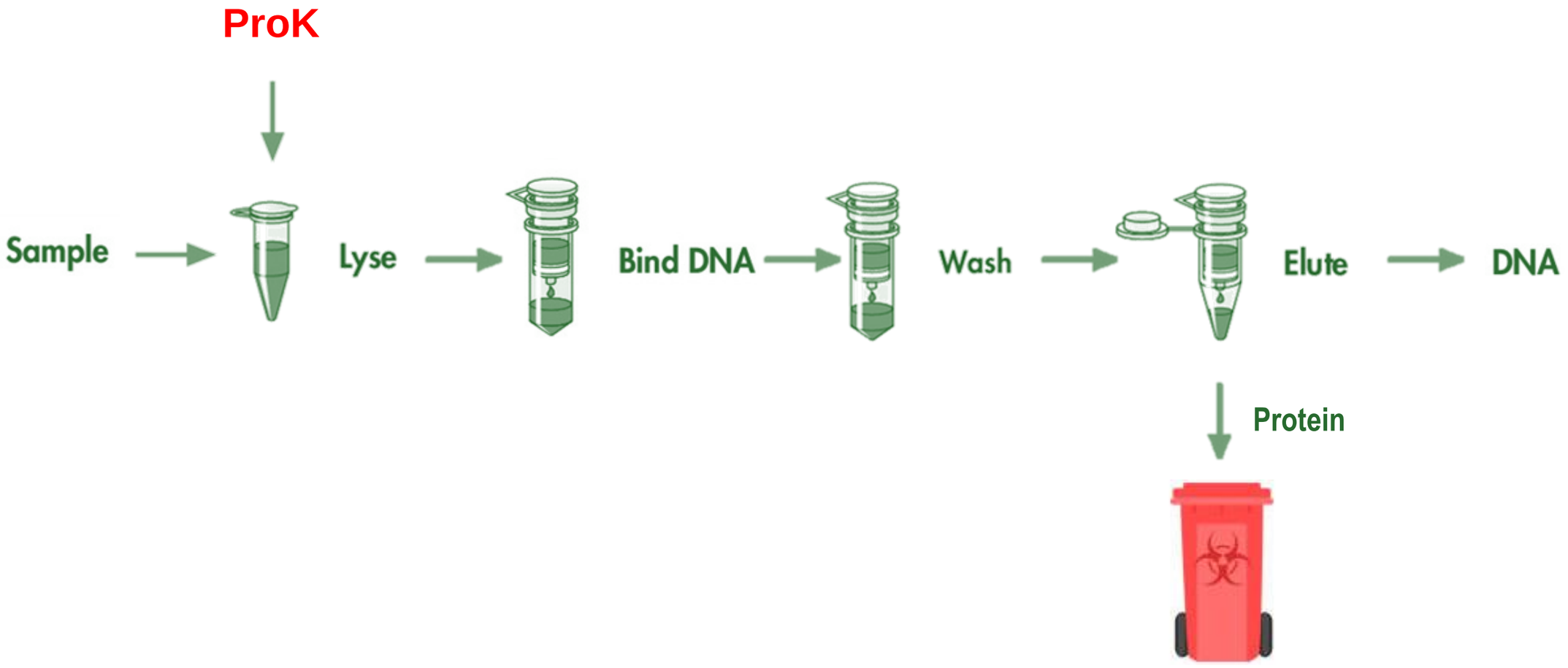
2. When you collect a touch sample, you’re collecting both DNA and protein

Extraction of Protein from Fingerprints

Extraction of Protein from Fingerprints

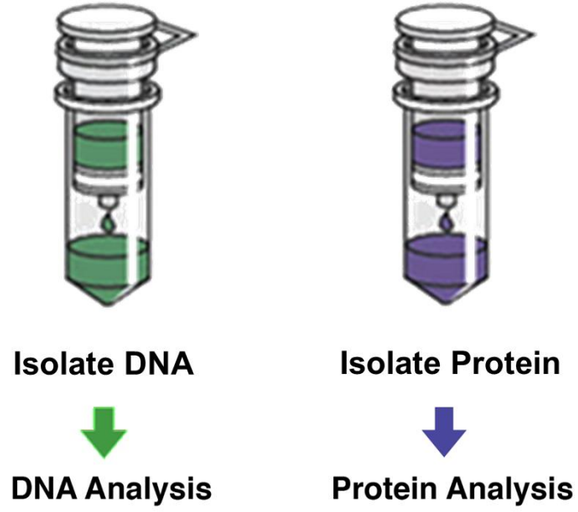


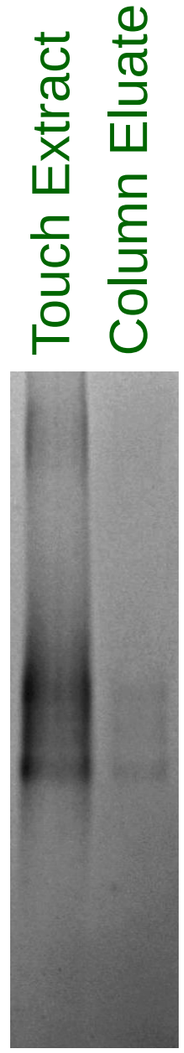
Extraction of Protein from Fingerprints



Protein Recovery from DNA Extraction Columns

- No significant protein detected in DNA column flow-through or washes
  - High salt concentration would complicate analysis
- Protein can be recovered from the column matrix following stringent denaturation/elution
  - Heat + DTT
- ~25% recovery of total protein in sample





Extraction of Protein from Fingerprints

1

Visualize Fingerprint

2

Solubilize and Collect

3

Bind DNA and Protein with DNA Investigator

4

Isolate DNA

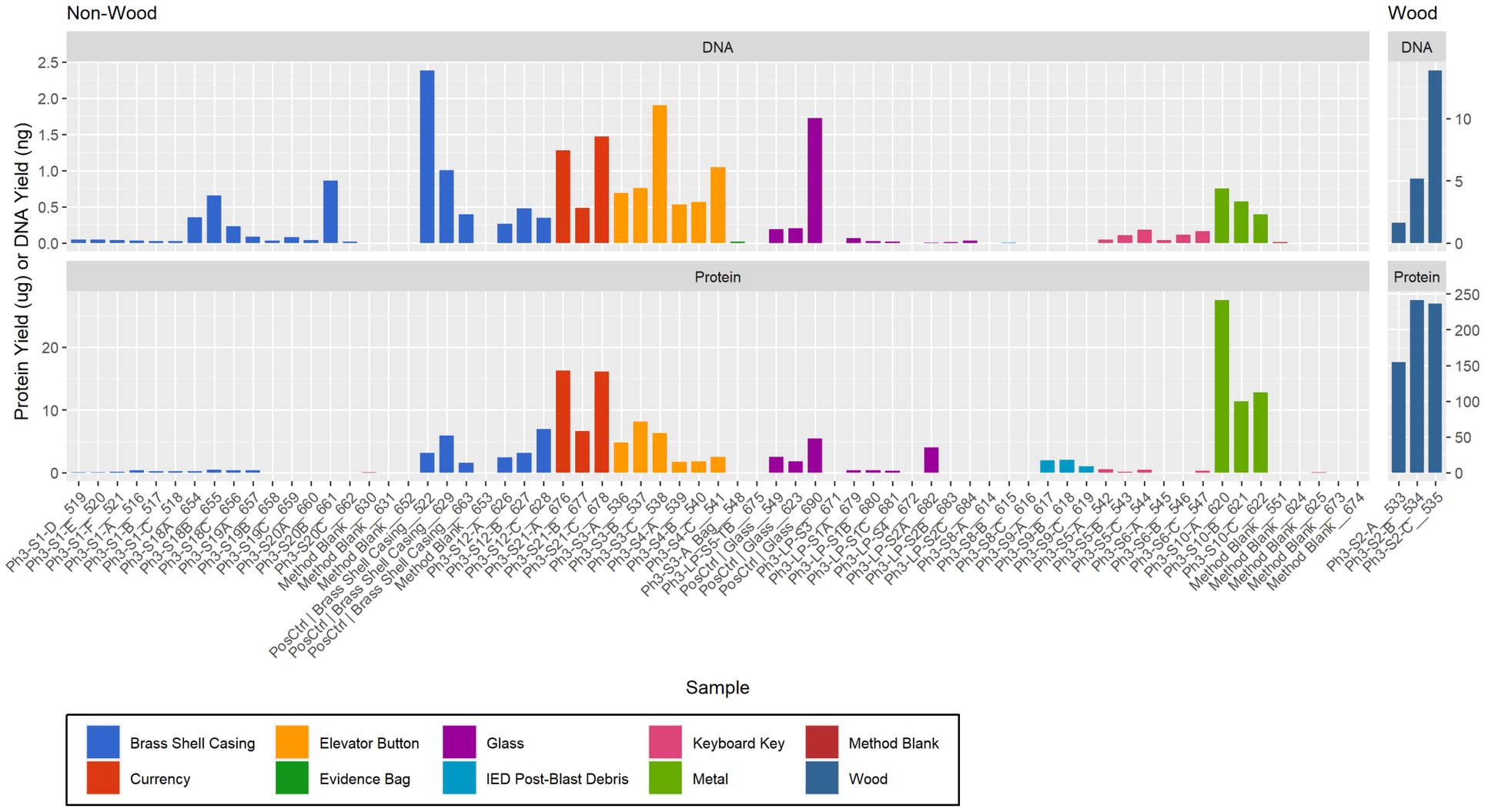
DNA Analysis

5

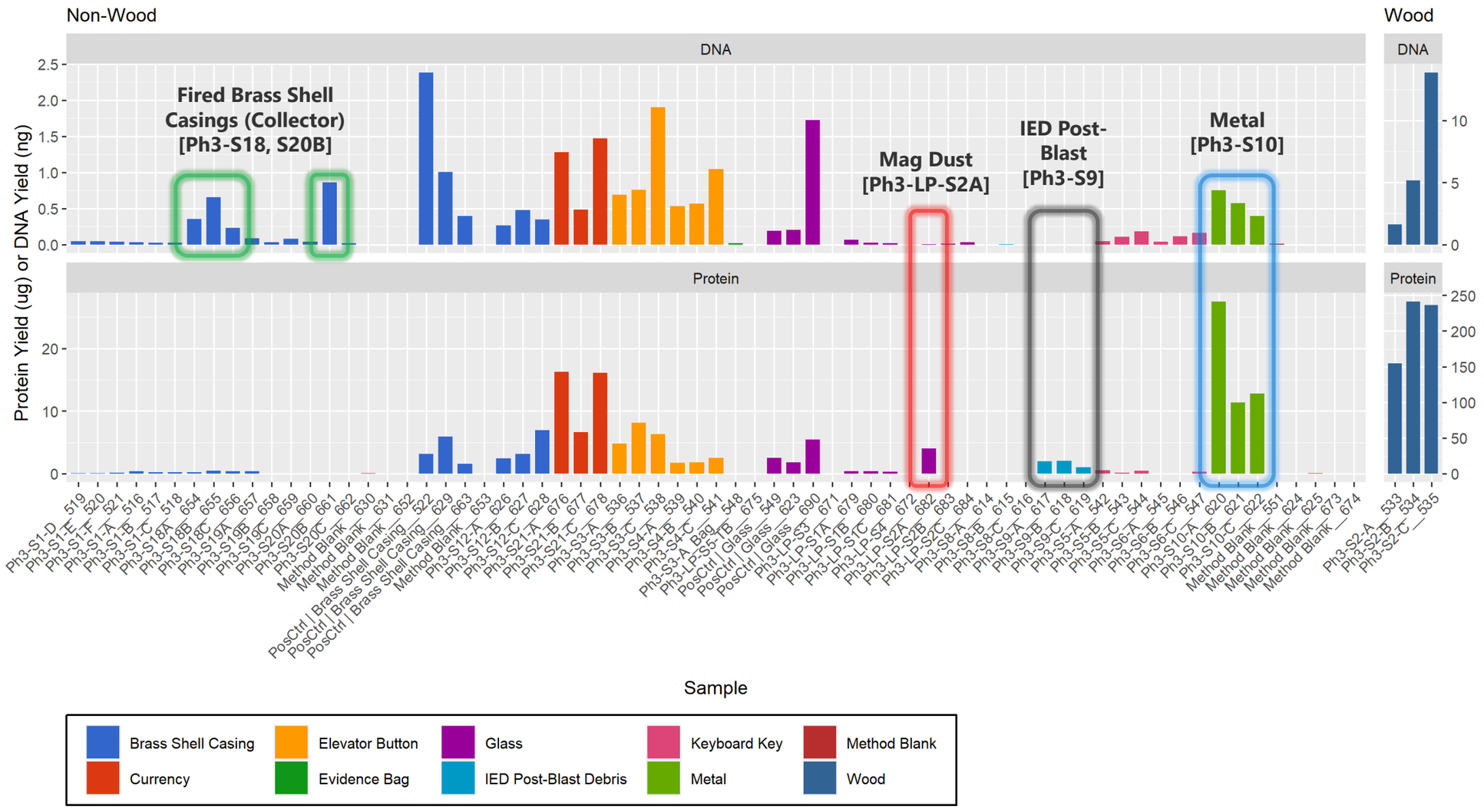
Isolate Protein

Protein Analysis

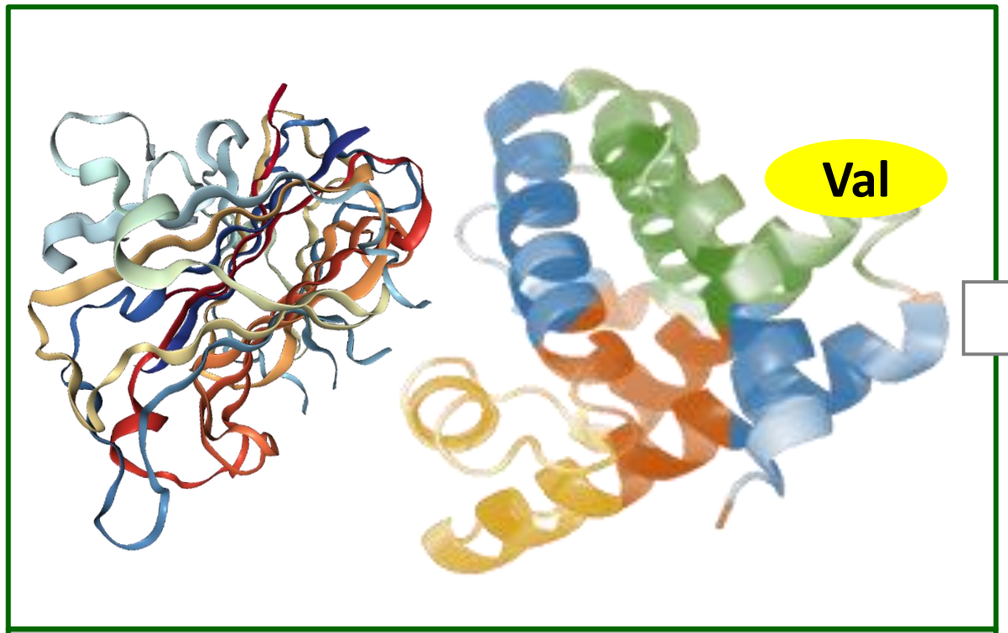
Correlation Between DNA and Protein Recovery



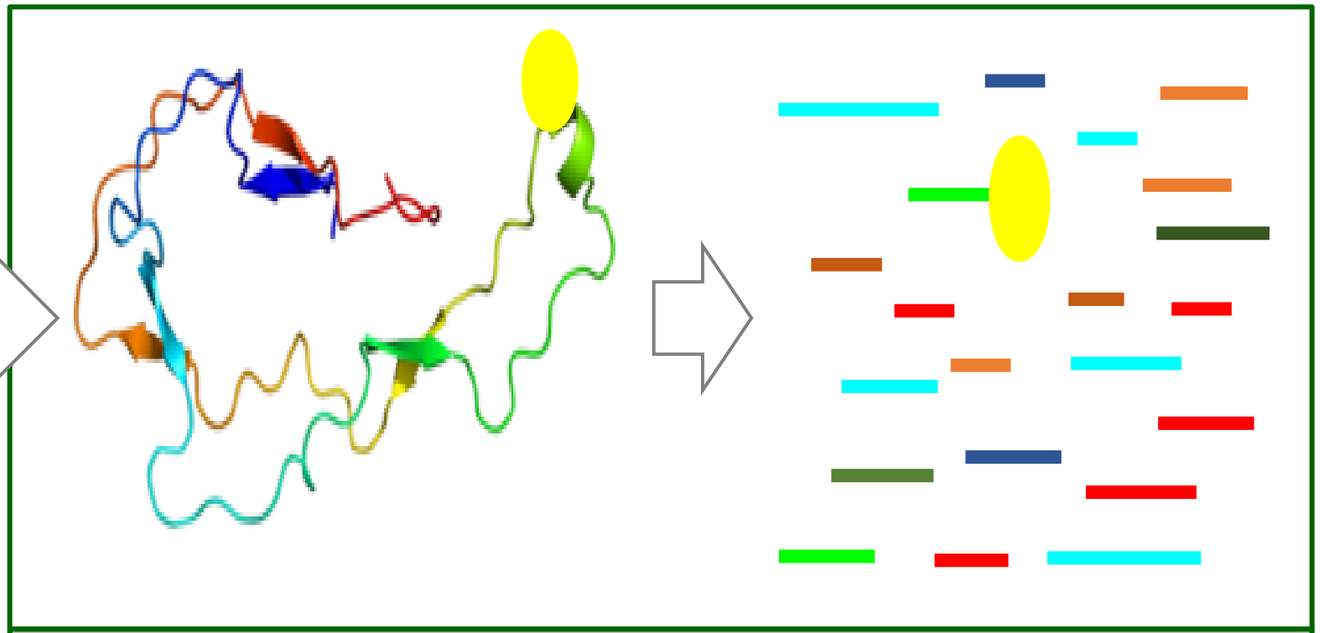
Correlation Between DNA and Protein Recovery



Proteogenomic Chemical Analysis



Extracted Protein



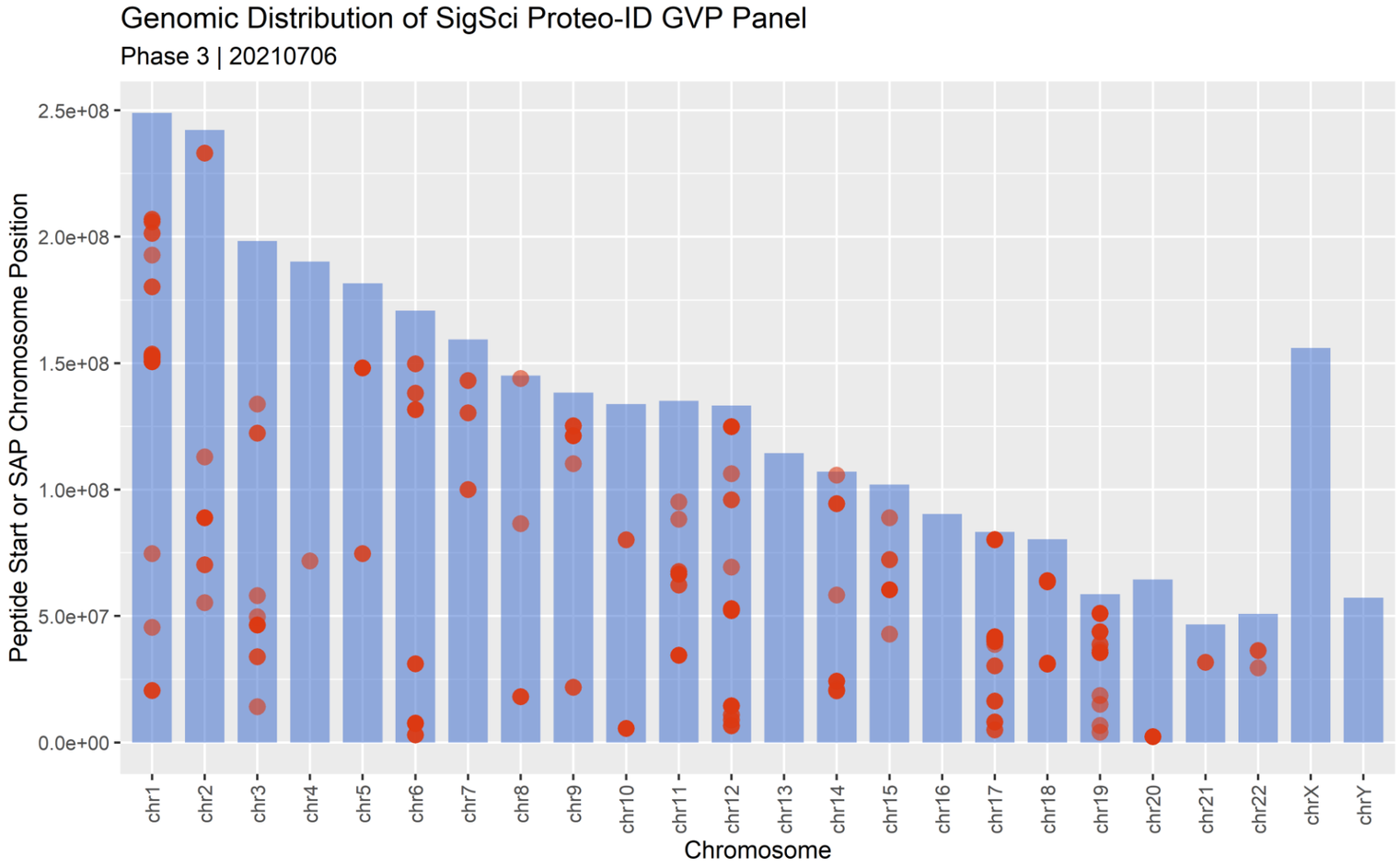
Enzymatically Digest Protein into Peptide Segments

Proteomic Analysis Method | LC-HRAM-MS2

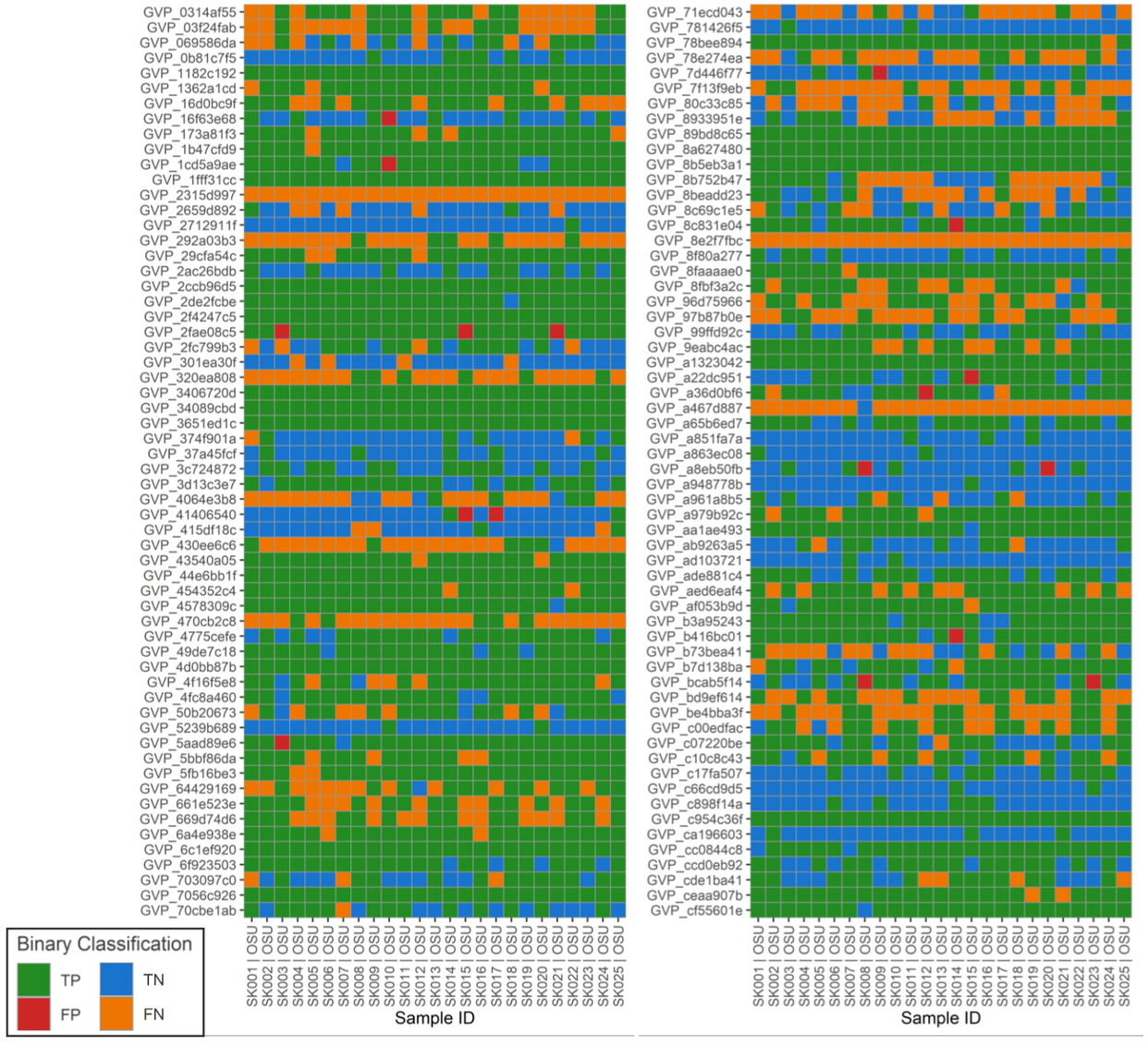
- Nano-LC
  - Thermo Scientific Ultimate 3000 RSLCNano
  - LC Run Time: 320 min (5 h gradient)
- High Resolution, Accurate MS2
  - Thermo Scientific Q Exactive Plus
  - Data Dependent MS2 or PRM Resolution = 17,500

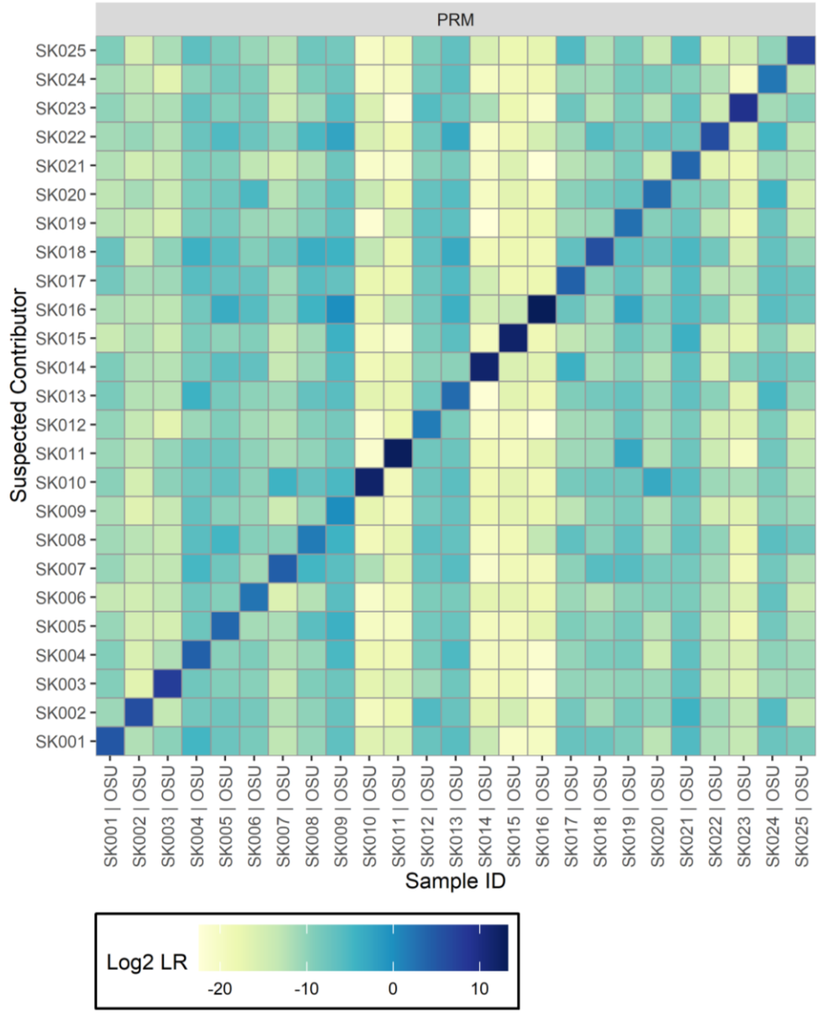
Genome-Wide Marker Distribution

Statistic	Value
Unique Proteins	445
Reference GVPs	281
Variant GVPs	191
Unique to 1 Donor	49
Unique to 1 T&E Donor	39
Unique to 1 SigSci Donor	42

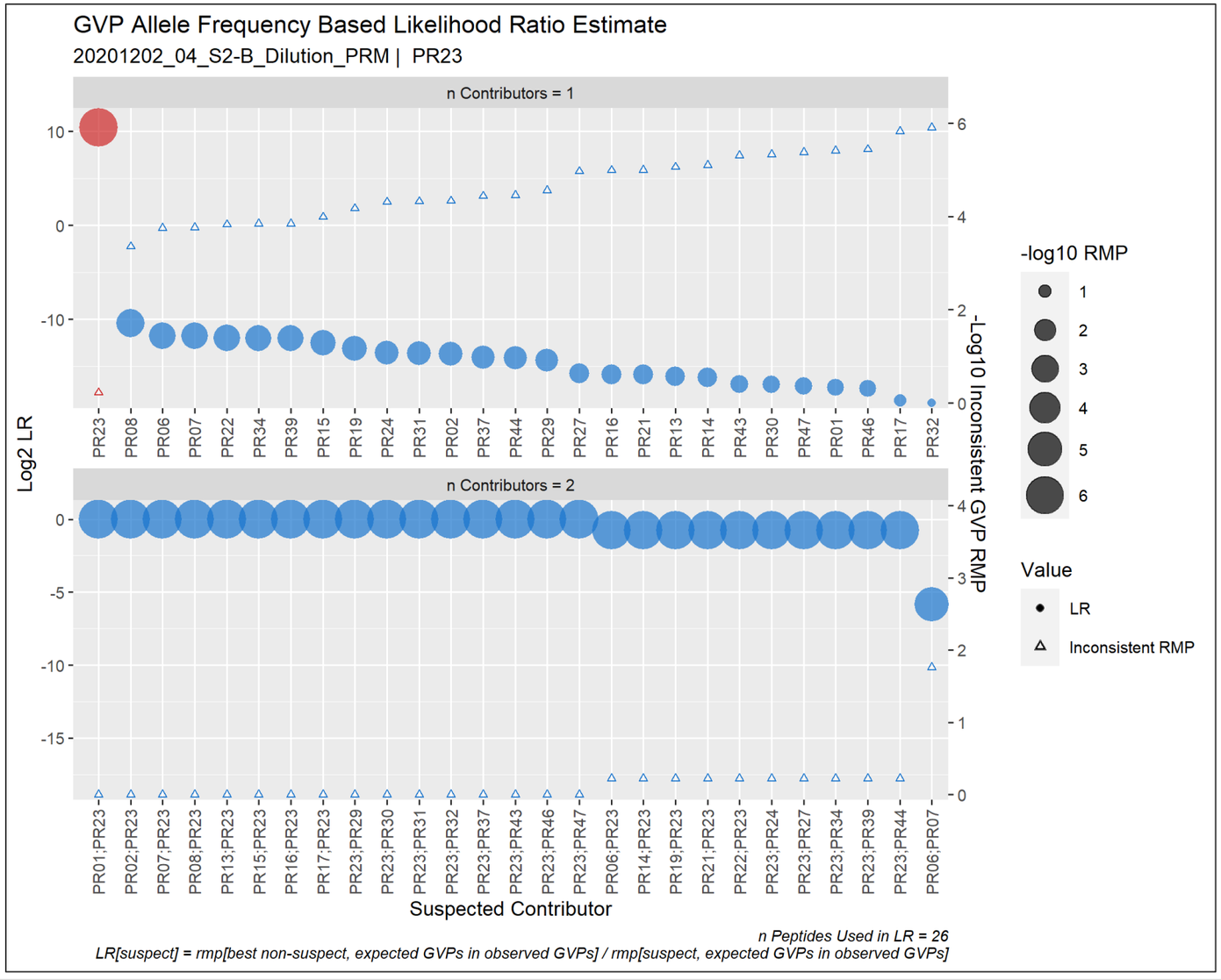


PRM Analyses of Internal Donors Using Common GVPs





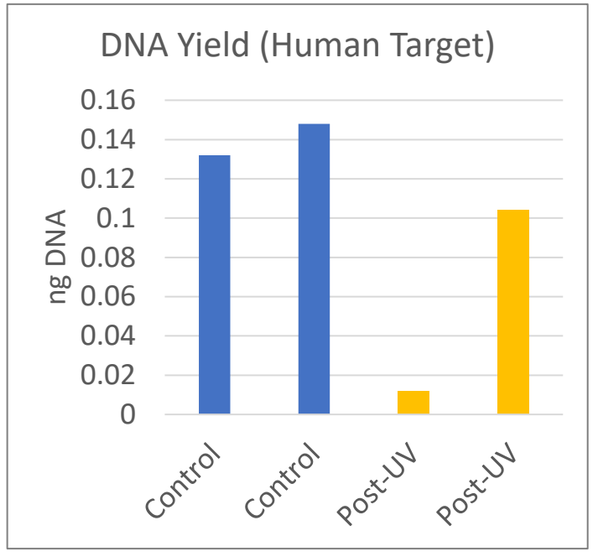
Simple Likelihood Ratio Calculations

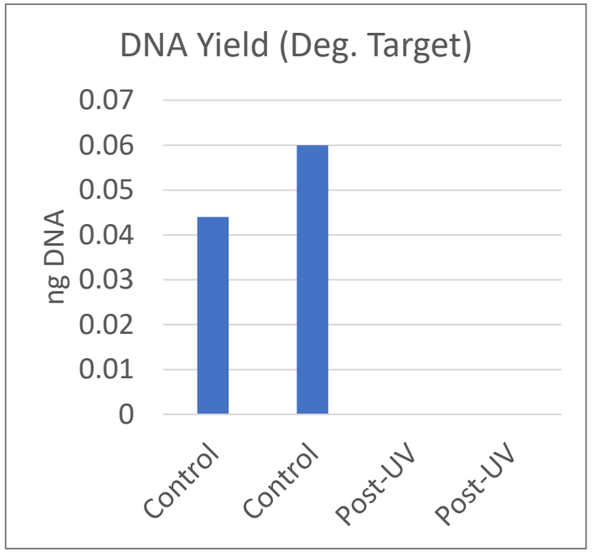


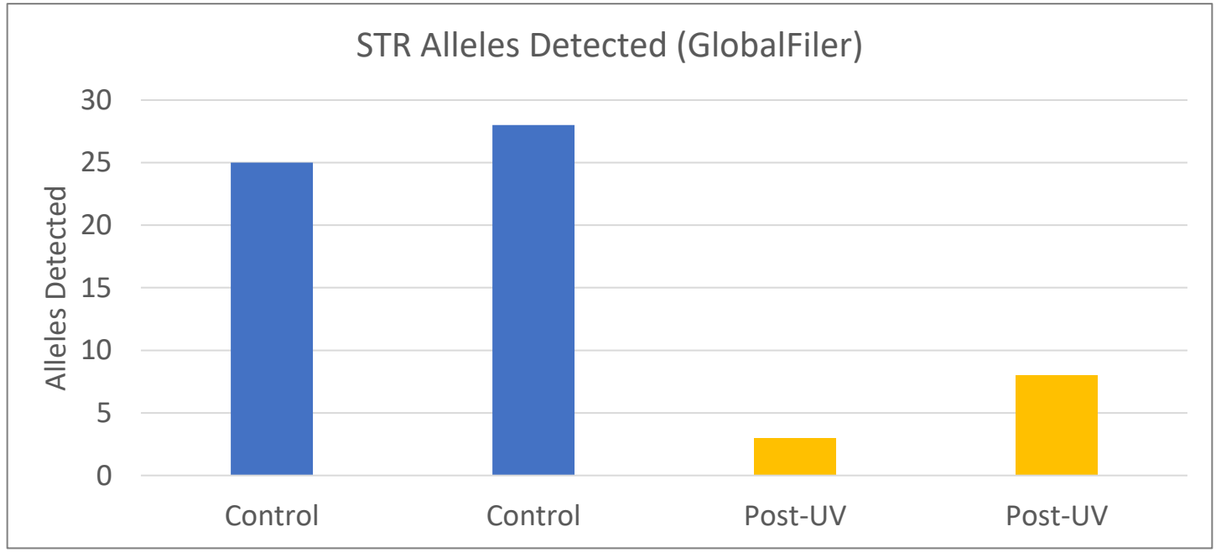
Project Highlights

DNA Degradation | Case Study

- Internal touch samples collected on plastic coupons
  - Controls (no UV)
  - UV treatment
  - Two replicates of each
- High degradation index
- Major decrease in STR alleles detected

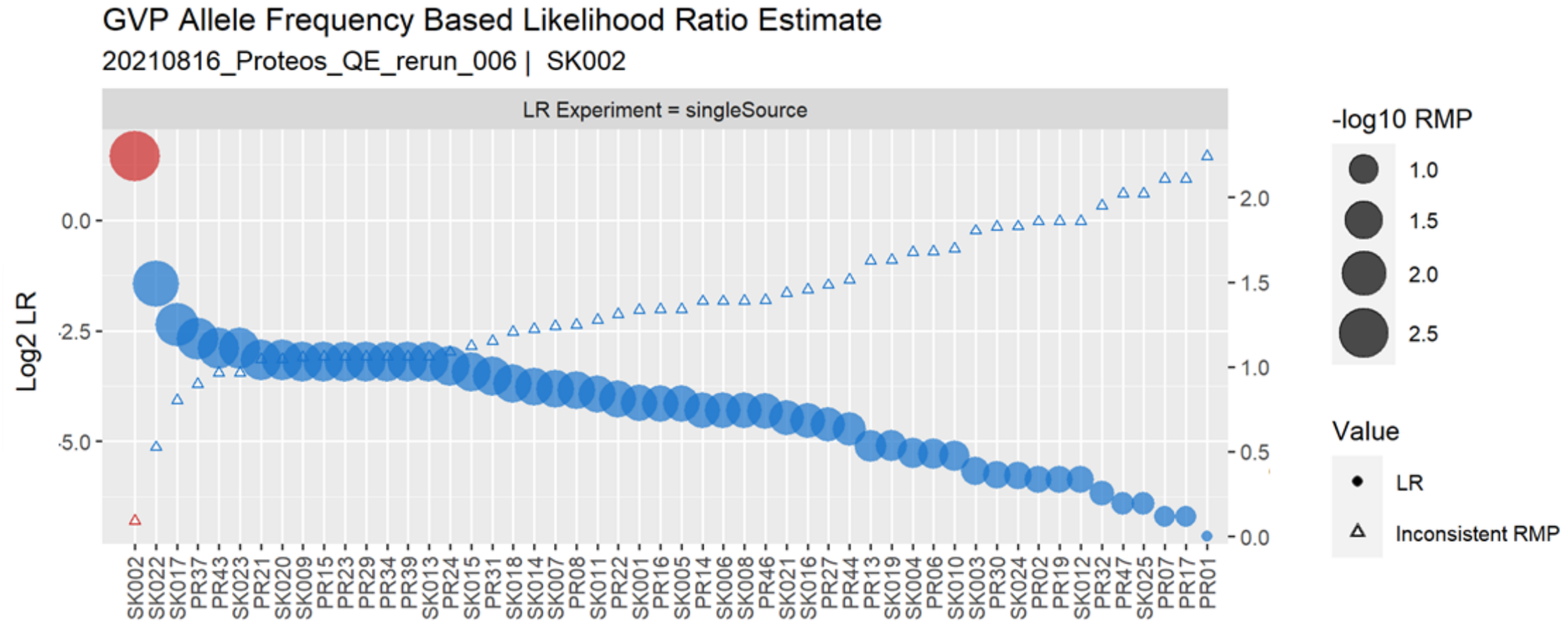






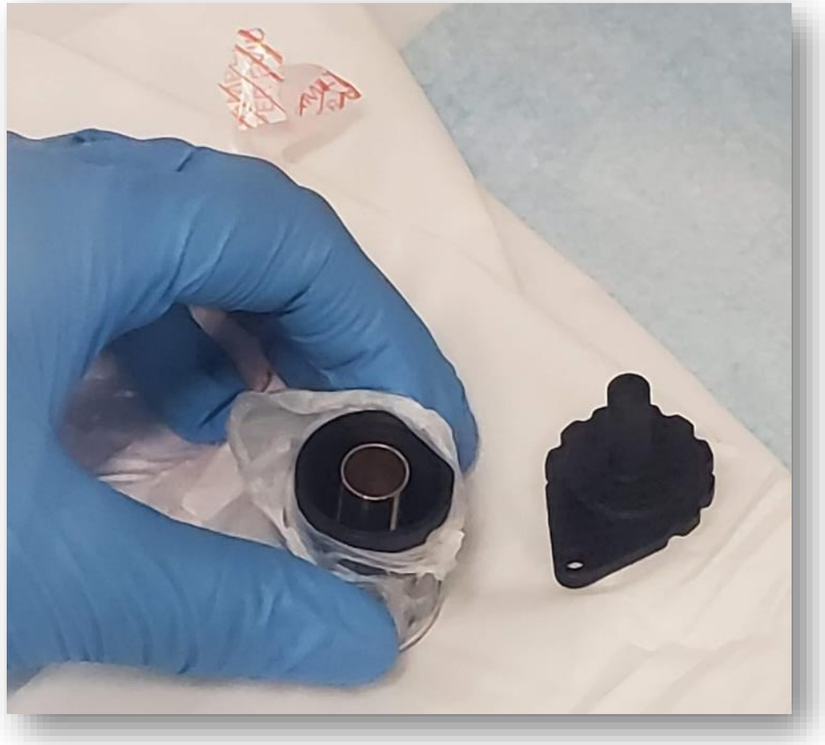
DNA Degradation | Case Study

- Following UV-exposure, 24 variable peptides identified
- Identified correct donor out of panel of 52 individuals

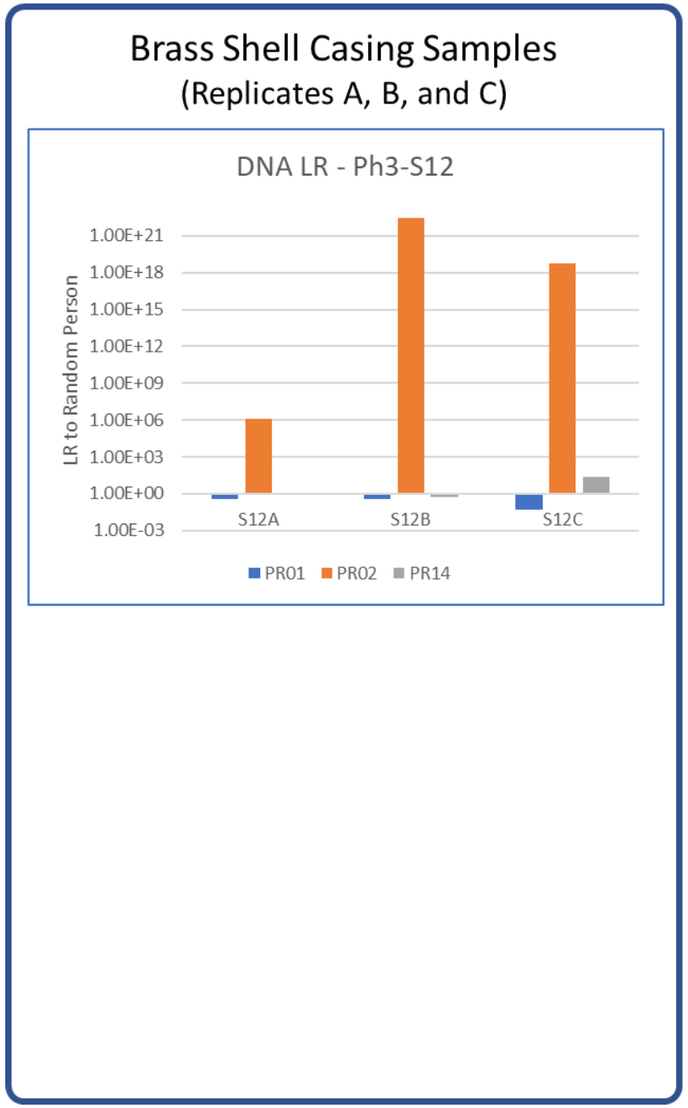


Brass Shell Casing | Case Study

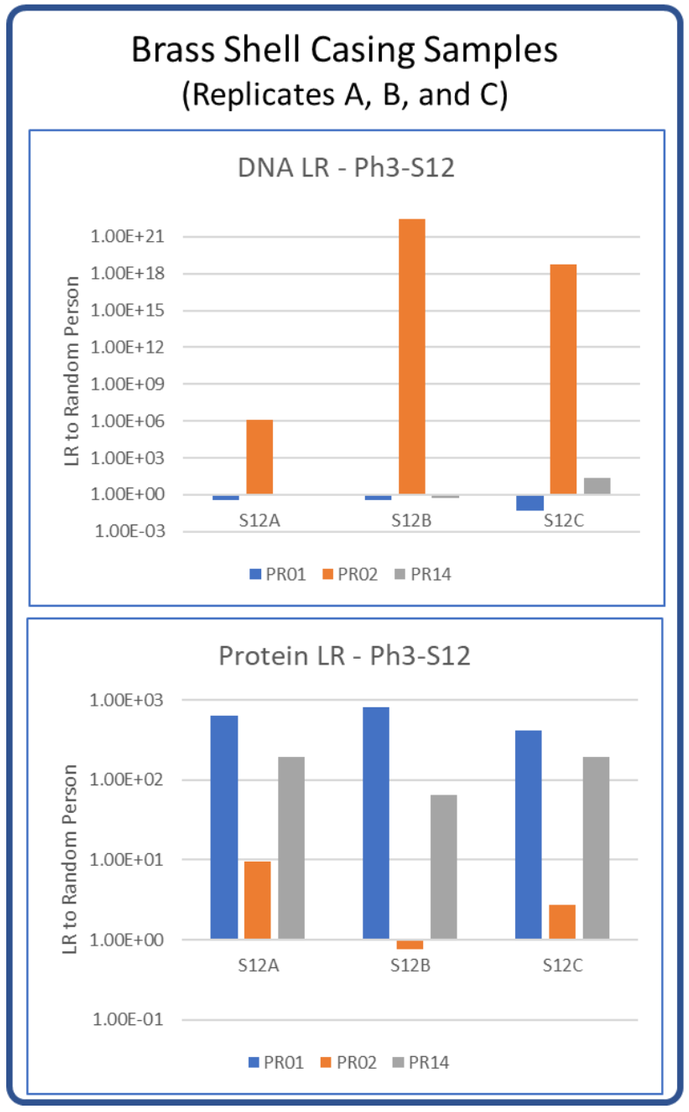
- 9mm Brass Shell Casings (Blinded):
  - Three replicates
  - Question: Who is/are the contributor(s)?
- Contradictory results:
  - DNA:
    - PR02 = major contributor
    - Probable mixture, but insufficient data for comparison to any reference profiles
  - Protein
    - PR01 = major contributor
    - PR14 = minor contributor



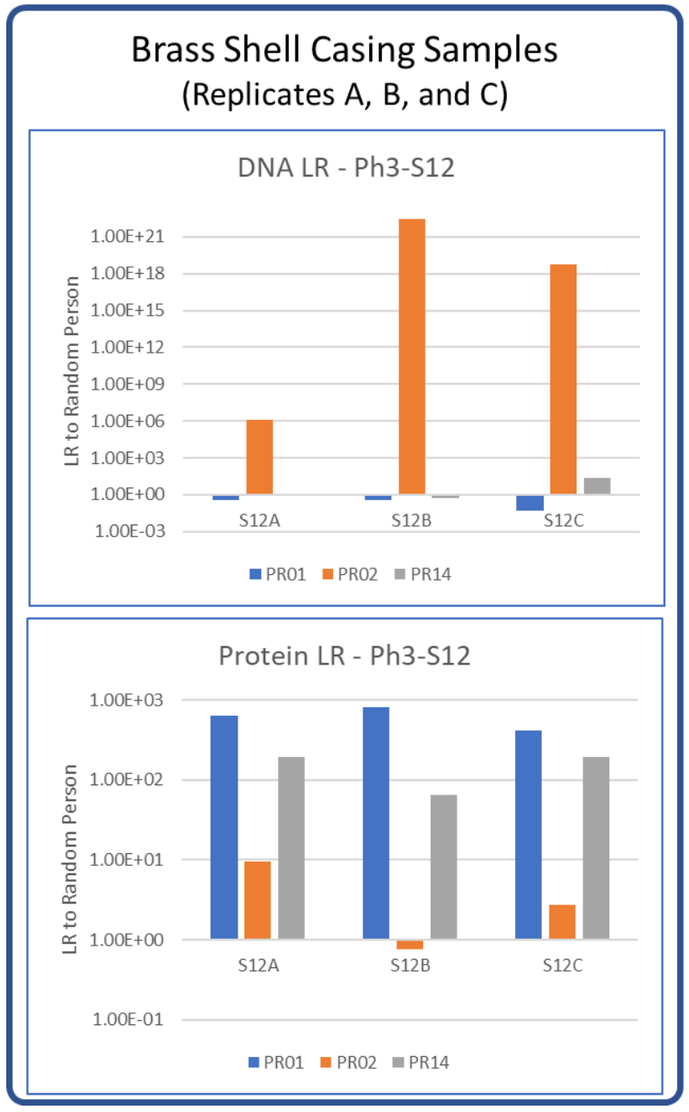
Brass Shell Casing | Case Study



Brass Shell Casing | Case Study

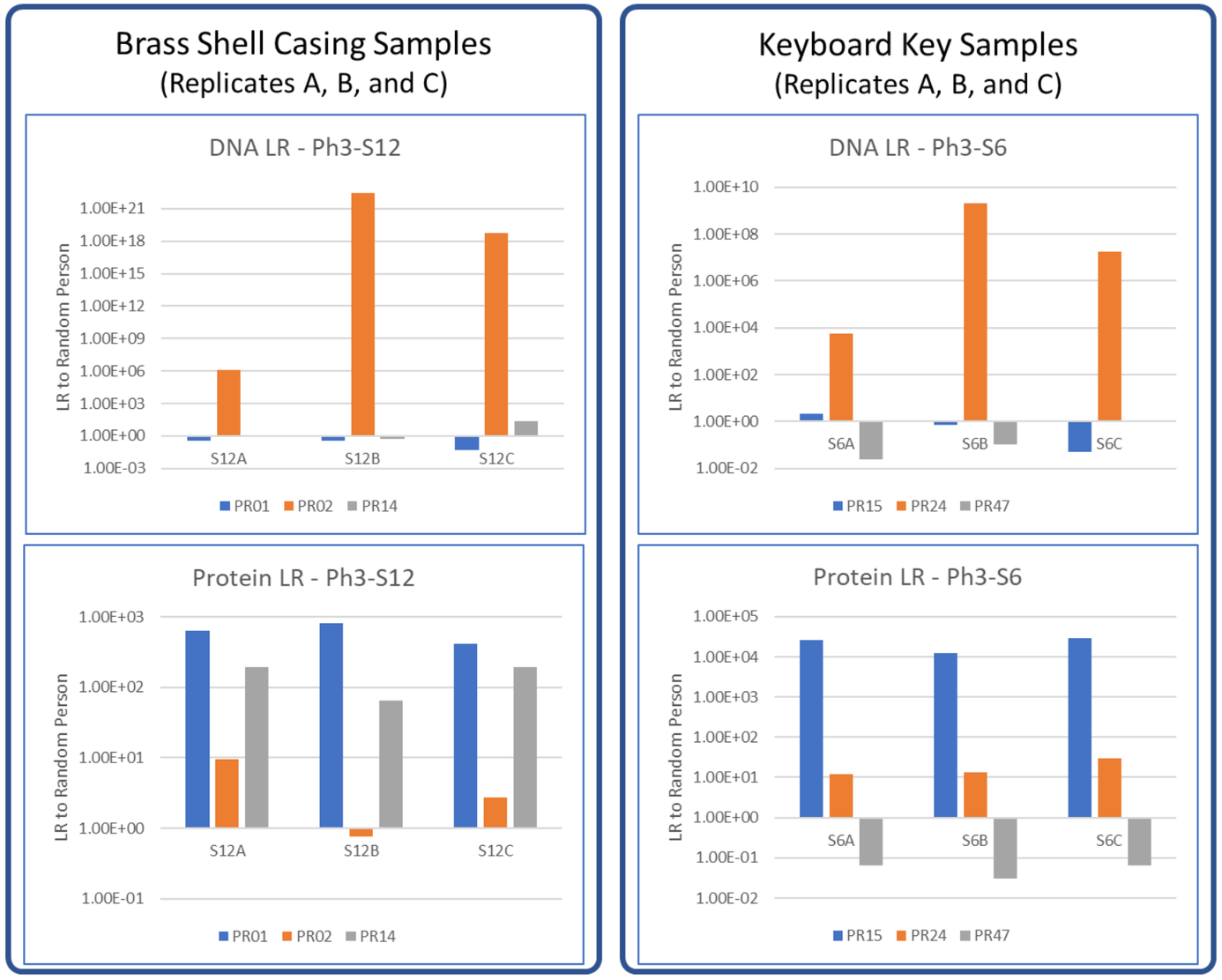


Brass Shell Casing | Case Study



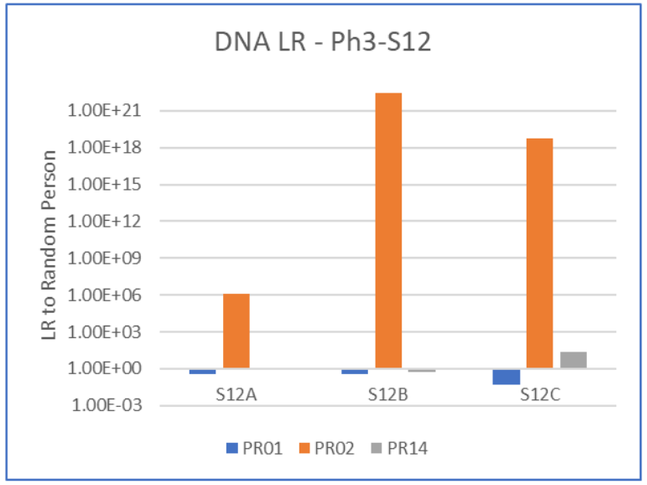
- Ground Truth:
  - Samples provided were mixtures
  - Three contributors: PR01, PR02, and PR14
- Contradictory results were actually complimentary
- How broad is the utility of combined DNA/protein analysis for mixtures?

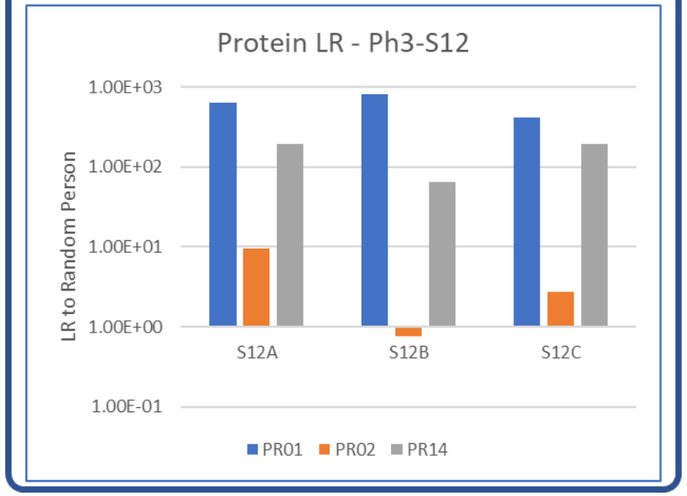
Brass Shell Casing | Case Study



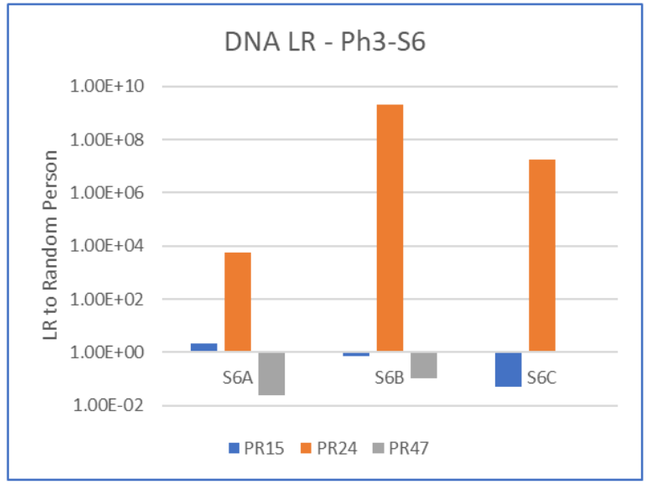
Brass Shell Casing | Case Study

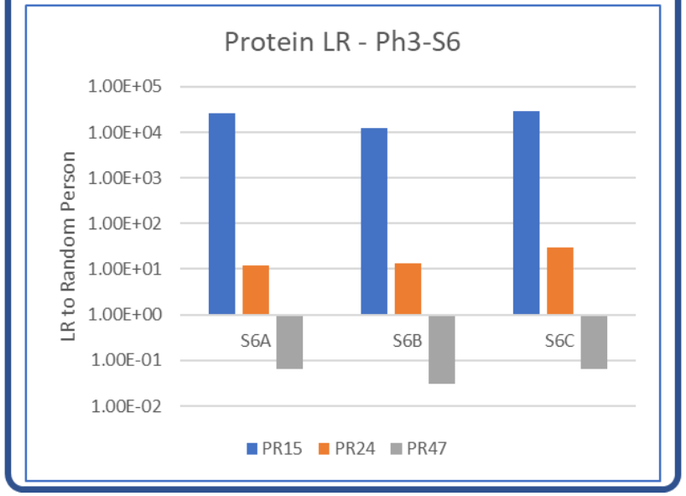
Brass Shell Casing Samples (Replicates A, B, and C)



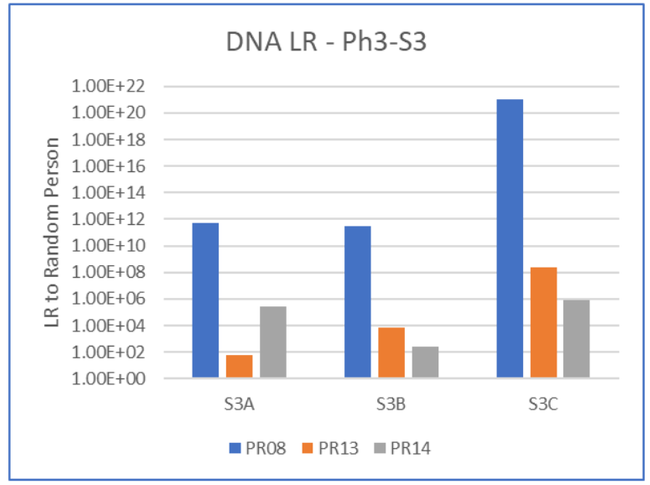


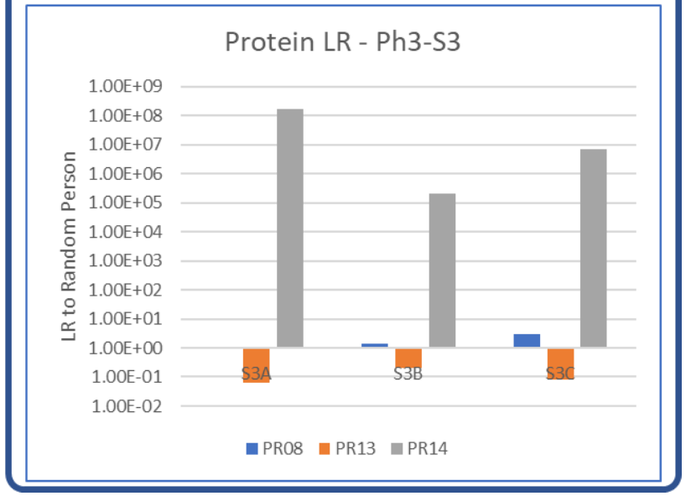
Keyboard Key Samples (Replicates A, B, and C)





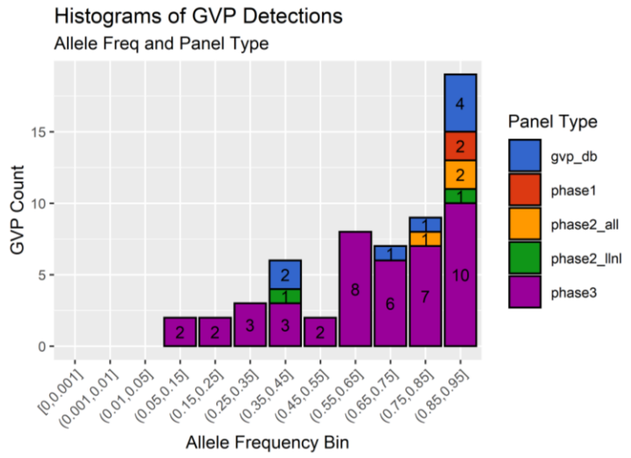
Elevator Button Samples (Replicates A, B, and C)

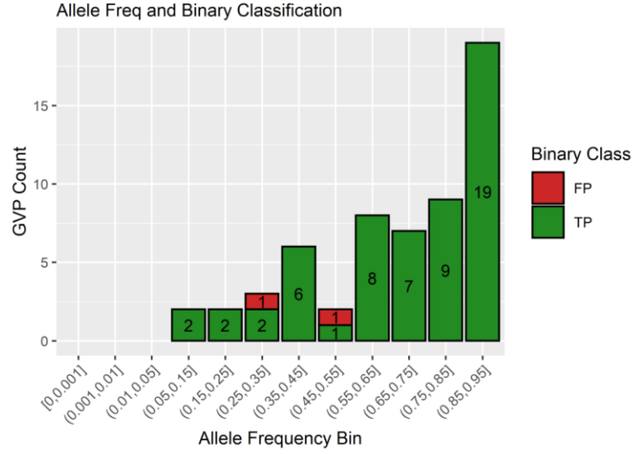


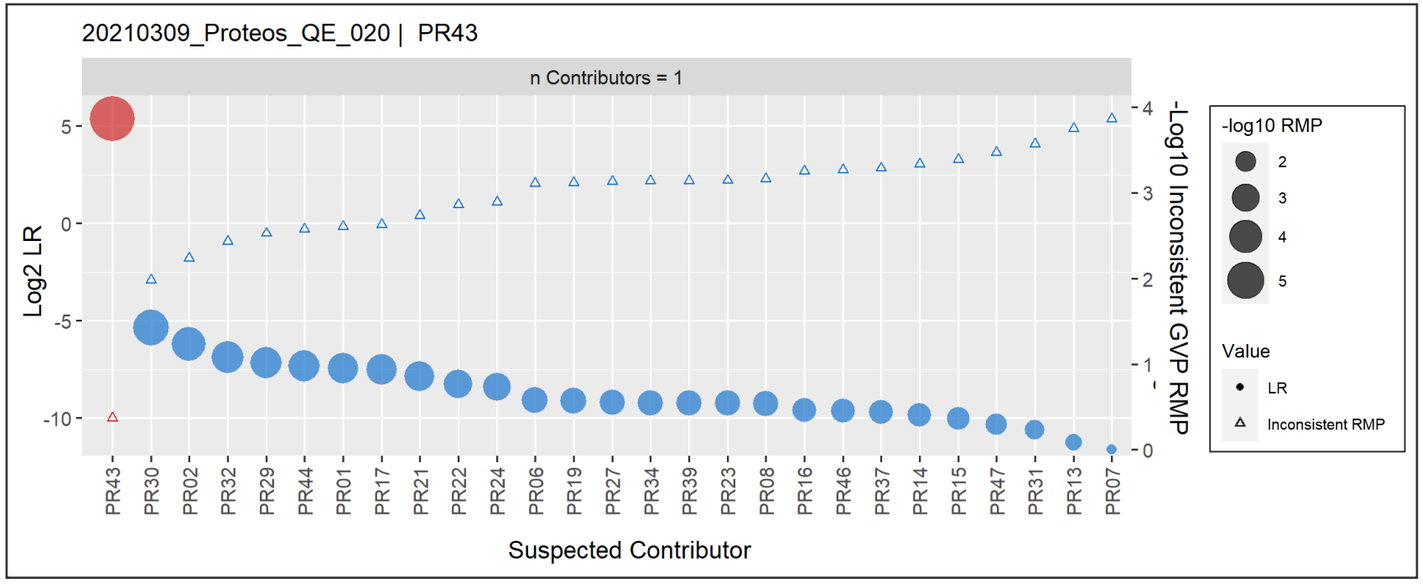


Identification of New Contributor with Known Exome

- New exome profile provided (PR43)
- Touch samples provided:
  - Three replicates
  - Question: Is PR43 present?
- Exome data processed, variable peptides identified
- Successful identification of new contributor





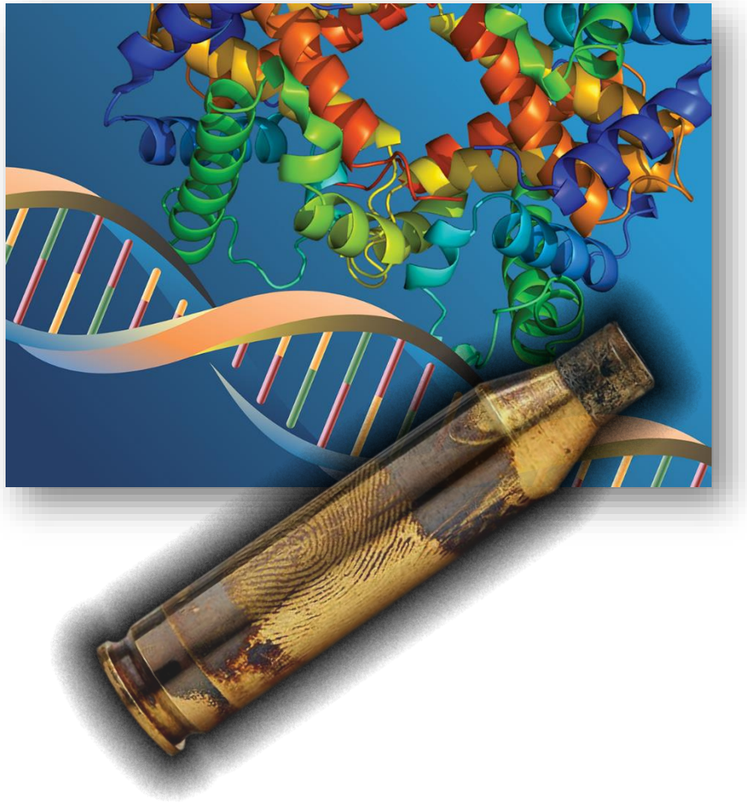


Current Challenges

- Sample collection and extraction methods not in complete alignment with most forensic labs
- Instrumentation
- No equivalent to CODIS for protein markers
  - Protein profiles can be compared to each other or a known exome/whole genome sequence
- Limited panel size and detectability can limit LR values
  - Even harder considering that some alleles may not be expressed or detectable
  - May be further complicated by kinship
- Rare protein markers
  - Extremely discriminating or just a false positive? How do you validate a marker you have never encountered?

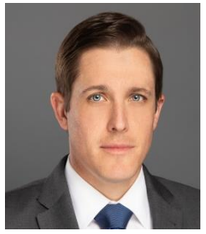
Future Directions

- Collection and extraction method optimization
- Method transfer to operational labs and independent laboratory evaluation from beginning to end
- Validation on relevant sample types
- Collection of additional data on sample matrices with known contributor(s)
- Database development (structure, format, hosting, accessibility, content)



Acknowledgements

Project Co-PI
Curt Hewitt, Ph.D.



Project Co-PI
Myles Gardner, Ph.D.

August Woerner, Ph.D.

Michael Freitas, Ph.D.

Key Contributors (SigSci)

- **Danielle LeSassier, Ph.D.**
- **Alan Smith**
- **Megan Powals**
- **Ben Ludolph**
- **Kathleen Schulte, M.S.**
- **Nicolette Albright, M.S.**
- **Michael Re, Ph.D., PMP**

Key Contributors (UNT CHI)

- **Ben Crysup, Ph.D.**
- **Bruce Budowle, Ph.D.**

Key Contributors (OSU)

- **Andrew Reed, Ph.D.**
- **Liwen Zhang, Ph.D.**
- **Maryam Baniasad**

Funding Acknowledgement:
This research is based upon work supported in part by the Office of the Director of National Intelligence (ODNI), Intelligence Advanced Research Projects Activity (IARPA), via contract number 2018-18041000003. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright annotation therein.

signature science LLC