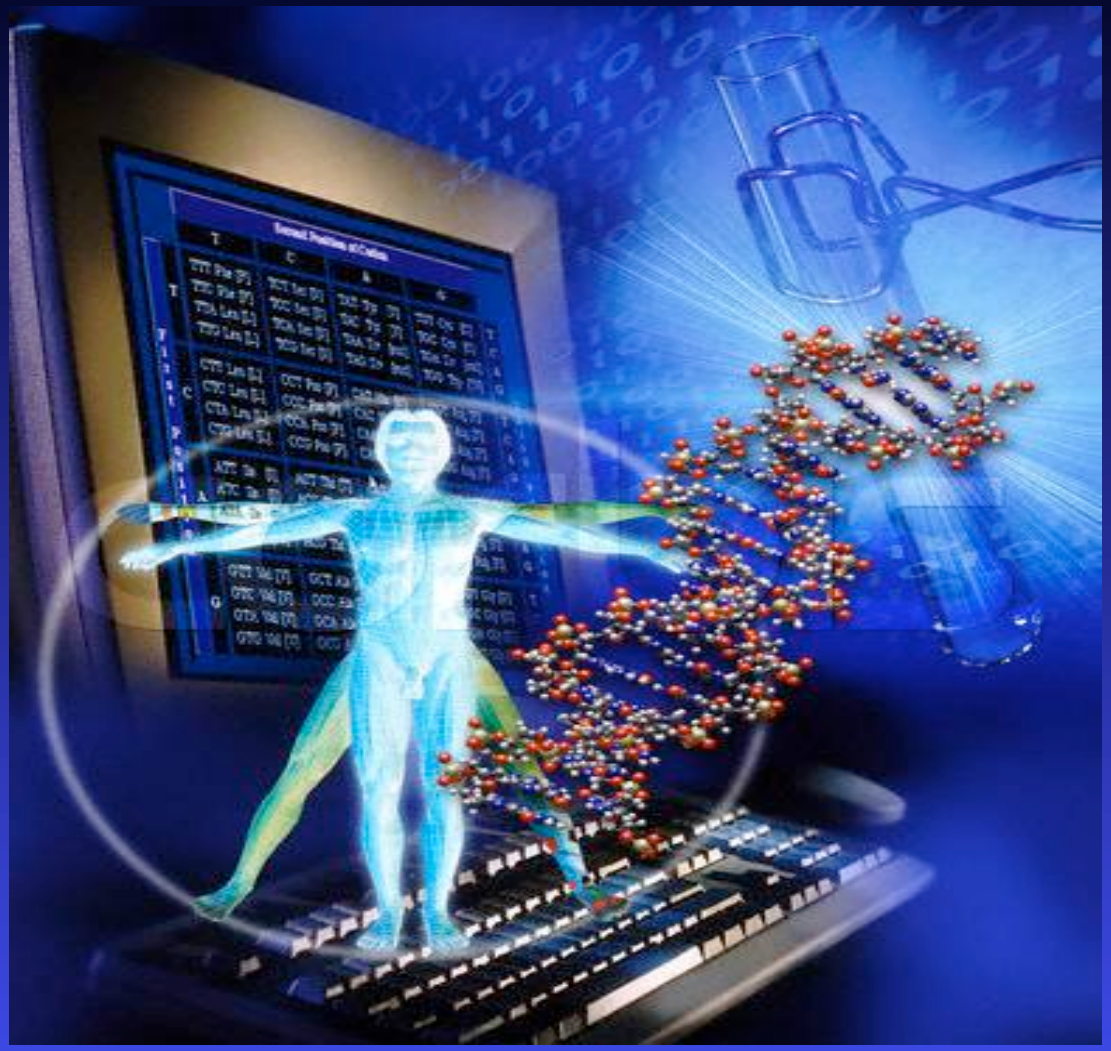
NGS: Possible Applications for Forensic DNA Analysis

What does the Person of Interest Look Like ?



Thomas Callaghan & James Robertson
FBI Laboratory, Quantico VA.

Can NGS Technology Be Used For Next-Generation Human DNA Identification ?

When the person of interest is not in the DNA database, can we take advantage of advances in genetic analysis to provide investigative leads on the person of interest ?

What do they look like?

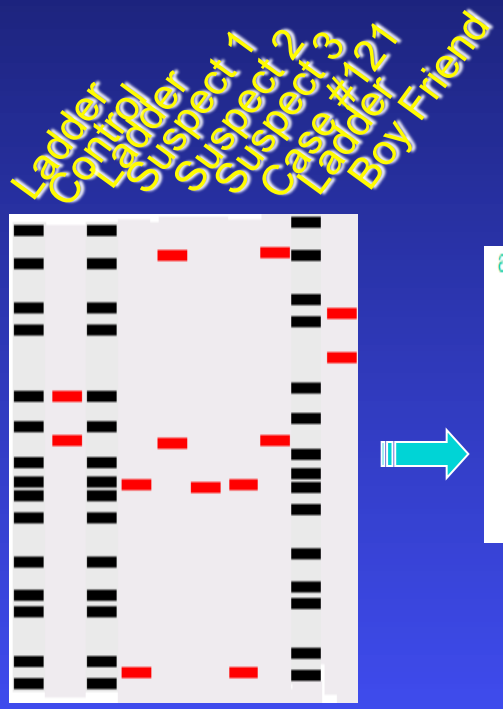
The FBI is **not** conducting human genomic sequencing on reference or crime scene samples.

Evolution Of DNA Identification

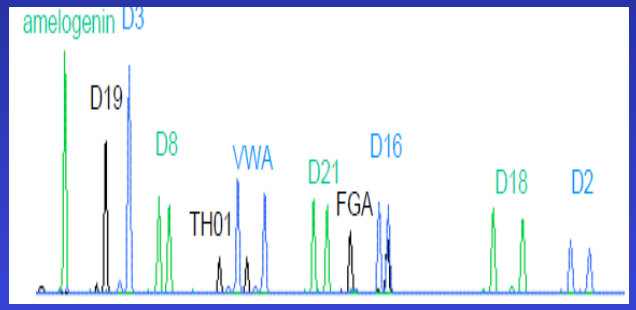
PAST

PRESENT

FUTURE
NGI



RFLP (≈ 6-8 Weeks)



13 core STRs (≈ 1 day-1wk)

RAPID DNA (≈ 1 hr)

Sequencing

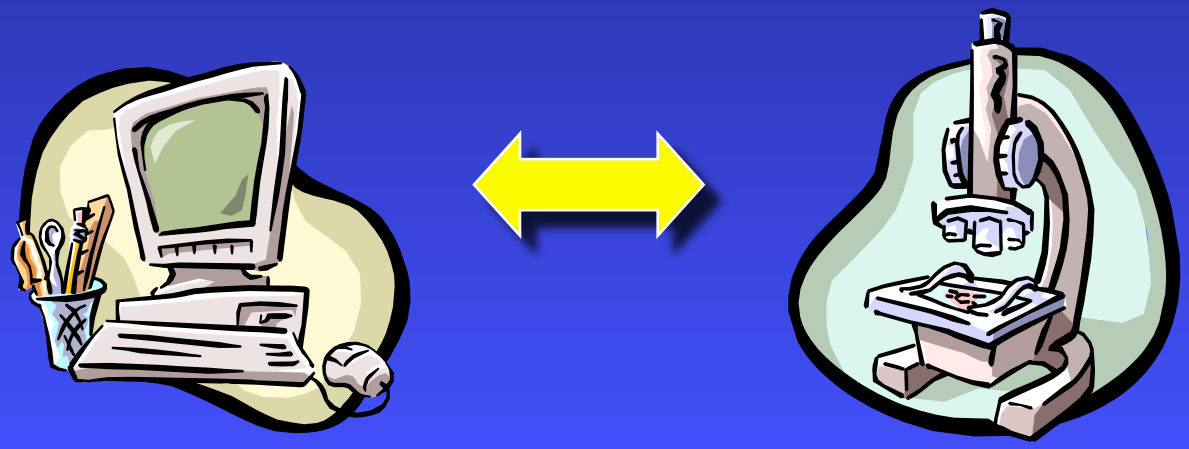
STR, mtDNA

Y-STR, SNP

Combined DNA Index System

Mission

CODIS blends forensic science and computer technology into an effective tool for solving violent crimes. CODIS enables Federal, State, and local crime laboratories to exchange and compare STR DNA profiles electronically.



Brief Chronology

- *CODIS* - pilot project 1990 - 14 state and local labs
- *DNA Identification Act of 1994* - FBI's authority to establish a National DNA Index System (NDIS). **Required FBI to establish National Standards for Forensic and Database DNA Analysis. Samples in NDIS must meet FBI QAS. FBI establishes NDIS Custodian.**
- *October 1998* - FBI's *NDIS* became operational with 9 states participating
- *DNA Analysis Backlog Elimination Act of 2000* – Authorizes collection of DNA samples from Federal convicted offenders (FCO)
- *Justice for All Act of 2004* – Indicted Persons at NDIS, One-Time Search Capability, Accreditation & Audit, All Felonies for FCO
- *DNA Fingerprint Act of 2005* – Arrestees & Legally Collected Samples at NDIS, Elimination of One-Time Search, Arrestees and Detainees for FCO

CODIS Participants

CODIS/NDIS Laboratories

190 in 50 states, FBI, US Army Crime Laboratory, Puerto Rico and Washington, DC.

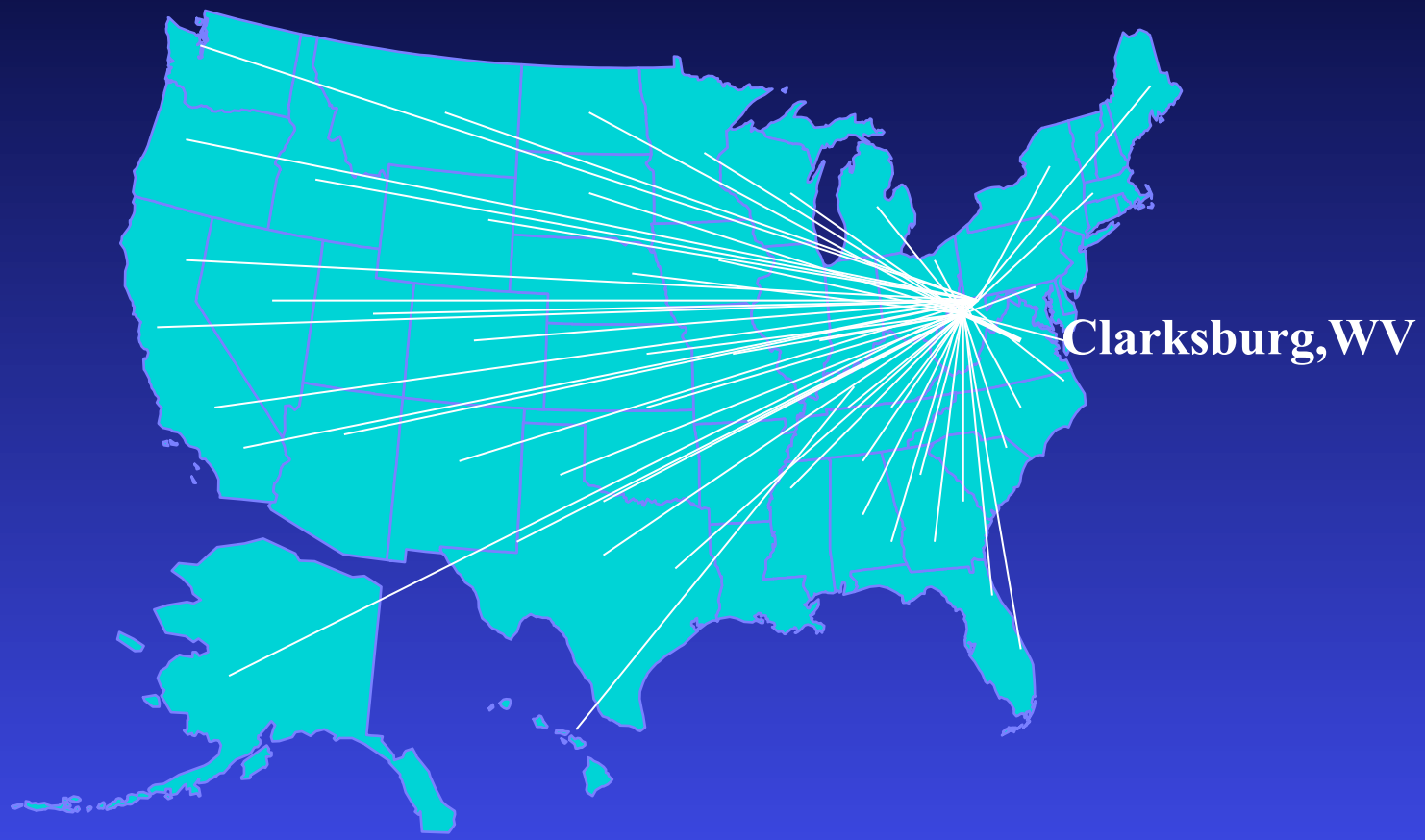
Users

Over 2,000 authorized users

International

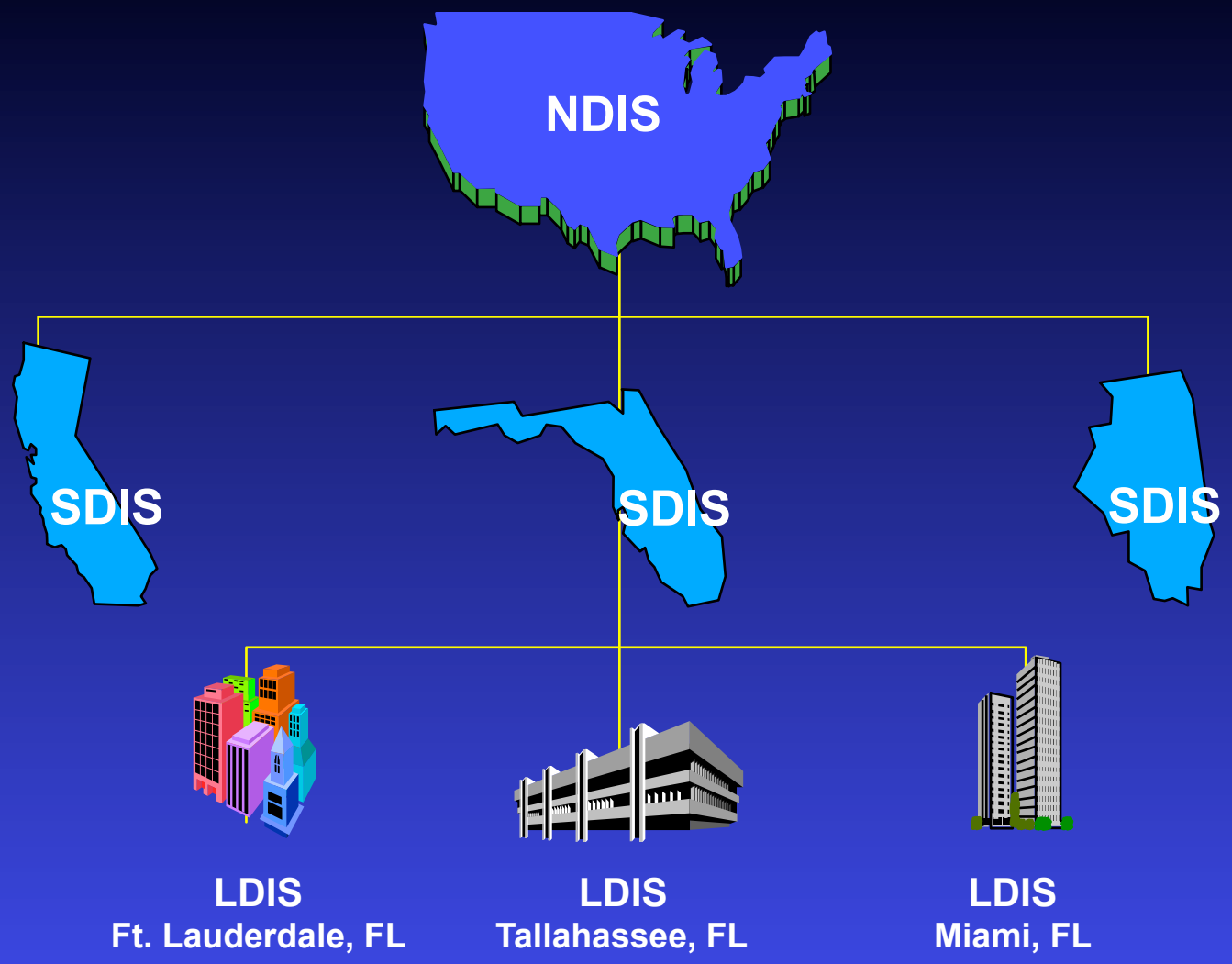
73 Labs in 37 countries

Database Connectivity



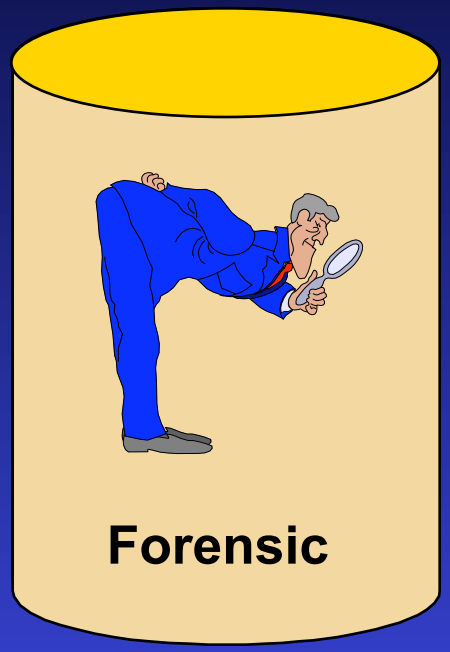
Secure network connecting all public DNA laboratories in the USA - allows searches and match messaging.

NDIS Architecture



190 Sites = 136 LDIS and 54 SDIS

NDIS Indexes



Arrestee

Legal

Unidentified Human (Remains)

Relatives of Missing Person

NDIS Searcher - Match Manager - [Match Display]

File Edit Target Profile Candidate Profile Window Help

Target Profile 1 of 1

Locus	Target DCFBIWAD7 MD1245781011Q1-2 KEYBOARD Source ID:N/A	Candidate GAGBI0073 DECATUR4800 Convicted Offender Source ID:N/A Matched High
CSF1PO	11	[H] 11
D13S317	12	[H] 12
D16S539	12, 13	[H] 12, 13
D18S51	10, 20	[H] 10, 20
D21S11	29, 30	[H] 29, 30
D3S1358	16, 18	[H] 16, 18
D5S818	12, 13	[H] 12, 13
D7S820	9, 11	[H] 9, 11
D8S1179	11, 15	[H] 11, 15
FGA	23, 24	[H] 23, 24
TH01	6, 9.3	[H] 6, 9.3
TPOX	8	[H] 8
vWA	17, 19	[H] 17, 19

Compare the Target Profile to the Candidate Profile - They match at the 13 CORE CODIS loci

Single Sample Target Profile

Reference: 030408099 Q2

Locus	Allele 1	Allele 2
D3S1358	17	19
vWA	15	16
FGA	23	28
D8S1179	13	
D21S11	29	30
D18S51	13	14
D5S818	9	12
D13S317	8	11
D7S820	9	13
D16S539	12	
TH01	6	
TPOX	11	

Reference = 030408099 Q2

Display Inverse Probability Summary

Summary of Probability Statistics

Locus	BLK	CAU	SEH	SWH
D3S1358	4.7600E-03	5.2103E-03	4.2523E-03	3.6518E-03
vWA	1.2721E-01	4.5217E-02	5.3760E-02	5.4947E-02
FGA	4.1750E-03	4.0499E-03	3.9090E-03	3.4538E-03
D8S1179	5.1101E-02	1.1737E-01	1.2903E-01	1.0788E-01
D21S11	6.7908E-02	8.4067E-02	1.1472E-01	1.3494E-01
D18S51	7.1057E-03	4.2508E-02	3.0840E-02	5.7800E-02
D5S818	9.8857E-03	2.1800E-02	3.1670E-02	3.1501E-02
D13S317	1.7235E-02	6.3461E-02	7.0204E-02	2.6866E-02

	BLK	CAU	SEH	SWH
Total	6.846E-23	1.846E-19	1.765E-19	4.679E-20

Configuration

Database: c:\CODISII\CODIS\POPDATA\FBI\STR

Description: FBI's U.S. STR population database for Cau Blk, SE Hispanic and SW Hispanic.

Min Allele Config: By Locus-Population Group

RFLP Bin Method: Fixed Bin

Max Band Size: 22621 Exclude: By Locus

Current Probability Formulas:

Homozygotes RFLP: $f = Ap^x$

PCR: $f = Ap^x + p(1-p)\theta_1$

Heterozygotes: $f = 2pq$

Bin Determination Window (1)

<--- ±8.00%

10000 <--- ±2.50%

2000 <--- ±2.50%

Match Window (1)

<--- ±8.00%

10000 <--- ±2.50%

2000 <--- ±2.50%

Inverse Summary of Probability Statistics

	BLK	CAU	SEH	SWH
Total	14,610,000,000,000,000,000,000	5,417,000,000,000,000,000	5,666,000,000,000,000,000	21,370,000,000,000,000,000

Details of Probability Statistics - D3S1358, BLK

Specimen	Allele	Lo Win	Hi Win	Lo Bin	Hi Bin	Lo Bin Freq	Hi Bin Freq	Bin Used	Count
030408099 Q2	17							8	84

Total Number of Alleles = 420

Minimum Allele Frequency = 1.19E-2

Probability f = 4.7600E-03 *

1 / f = 210

* The rounded minimum allele frequency was used in the calculation.

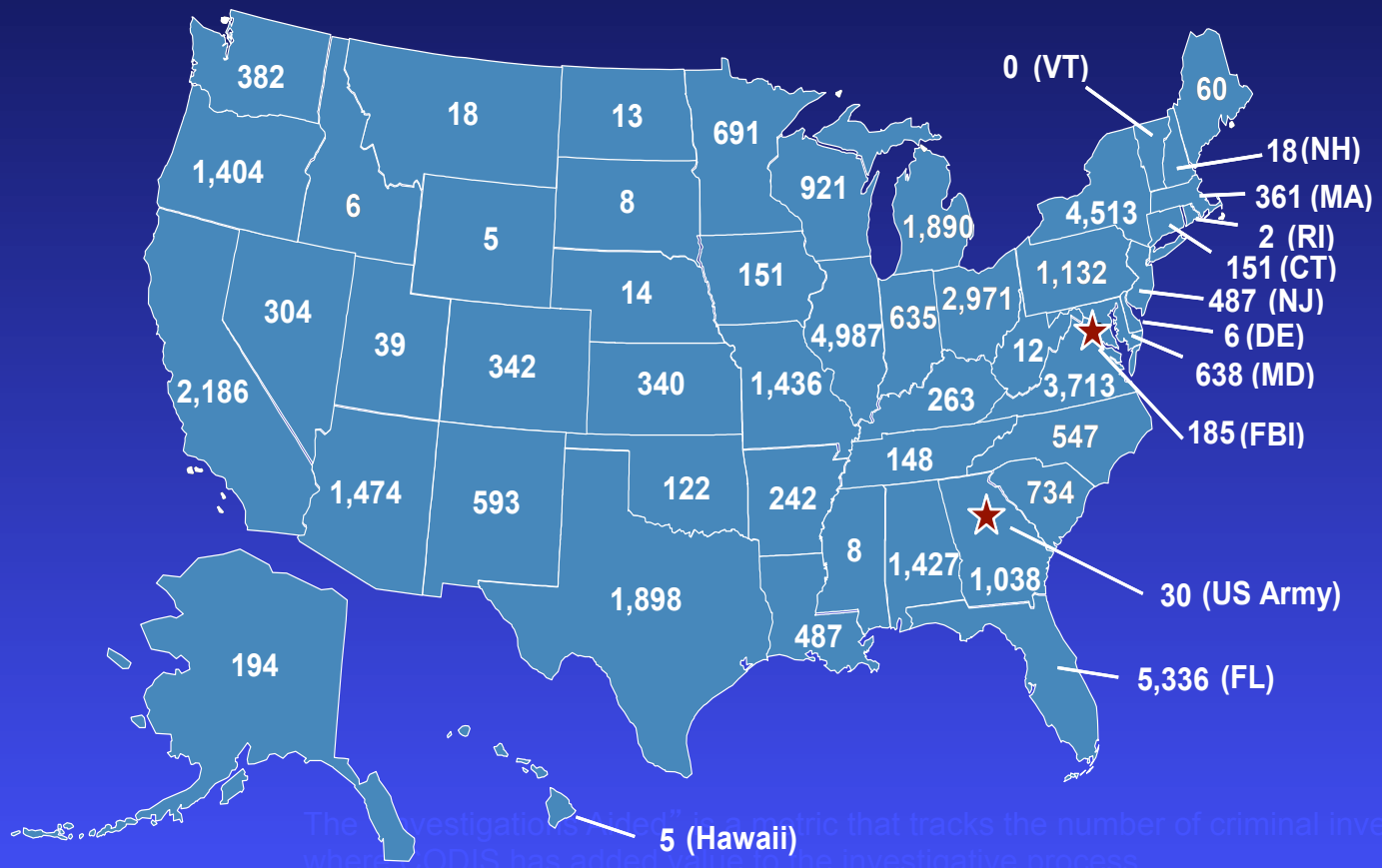
POPSTATS-

used to calculate the random match probability.

CODIS Investigations Aided

Through April 30, 2012

171, 052 Offender /Arrestee Hits
28,515 Forensic Hits



The "Investigations Aided" is a metric that tracks the number of criminal investigations where CODIS has added value to the investigative process.

NDIS Statistics

As of April 30, 2012

Category	*Total Number of Profiles*
Offender Profiles	9,629,475
Arrestee Detainee, Legal	1,077,436 4,127 & 7,681
Forensic Profiles	427,556
Missing Person	1,025
Rel. of Missing Person	5,779
UnID'd Human Remains	4,101
Pedigree Trees	2,832

NGS and Human Forensic Analysis

- Sensitivity (200-300 pg CODIS STRs)

  Forensic Samples: degradation, mixtures, inhibitors

- mtDNA: Missing Person (bone/teeth) – hair
- Y-STR: male-female mix., male missing per.

- SNPs

  BioGeographic Ancestry

  Phenotype

  Lineage/Clan

  Identity By Decent

NGS and Human Forensic Analysis Issues

- Technology/Chemistry
- Assembly/Alignment
- Standards for Data Exchange - (NAS)
- BioInfomatic Hard/Software and Expertise
- Sample Prep – Labor, dedicated space

Can NGS Technology Be Used For Next-Generation Human DNA Identification ?

When the person of interest is not in the DNA database, can we take advantage of advances in genetic analysis to provide investigative leads on the person of interest ?

CODIS STR, mtDNA, Y-STR, X-STR

What do they look like?

SNPs

Just one example…

- Baton Rouge Serial Killer: Derrick Todd Lee
  - Linked to at least 7 homicides in Louisiana
  - Many arrests over 20 years
  - No felony convictions
  - Composite sketch
  - AIM SNPs analysis
  - Louisiana passes arrestee DNA law

Can NGS Technology Be Used For Next-Generation Human DNA Identification ?

Before the FBI would apply NGS to crime scene sample analysis, it would have to gain approval from FBI and DOJ executive management and establish well thought out rules and policies in order to maintain public trust in forensic DNA analysis.

Acknowledgments:

Peter Vallone	**NIST**
James Robertson	**FBI CFSR Unit**
Doug Hares	**FBI CODIS Unit**
Joseph Donfack	**FBI CFSR Unit**

Thomas F Callaghan

thomas.callaghan@ic.fbi.gov

703 632-8391

FBI Laboratory
2501 Investigation Parkway
Quantico, VA 22135