NIST Update: What's new? What's going on?

Erica Romsos
Peter Vallone, Becky Steffen, Lisa Borsuk, Kevin Kiesler, Sarah Riman, Katherine Gettings

NIST APPLIED GENETICS

NIST National Institute of Standards and Technology U.S. Department of Commerce

25th Annual CODIS TL Summit
Norman, OK
November 19, 2019

1

What's new?

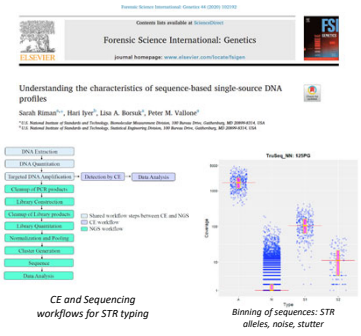
NIST

SRM 2391d: PCR-based DNA Profiling Standard

On sale as of July 9, 2019
2300 units were vialed

Developed as a successor to SRM 2391c
SRM 2391c is no longer available for sale at NIST
The expiration date was recently extended by **2 years** to **Feb. 3, 2022**
We will not further extend the expiration date after this date

2



CE and Sequencing workflows for STR typing

Binning of sequences: STR alleles, noise, stutter

Sarah Riman

Characteristics of sequencing noise in DNA profiles

The characterization of sequencing noise and artifacts is critical for assessing thresholds for data interpretation

Developed a framework using statistical tools to systematically interpret the characteristics of single-source DNA profiles generated by targeted sequencing

3

Population Sample Sequencing

Illumina MiSeq FGx instrument, ForenSeq
- 27 autosomal STRs + 24 YSTR + 7 XSTR + Amel
- 94 HID-SNPs + 56 ancestry SNPs + 22 phenotype SNPs

- 1036 Samples
- Sequenced in batches of 24 or 32
- 41 total sequencing runs in 2016

4

Population Sample Sequencing

2018
27 auSTR

Forensic Science International: Genetics

Sequence-based U.S. population data for 27 autosomal STR loci

24 Y-STR
in progress

Tunde Huszar PostDoc

2018
SE33

Sequence-based US population data for the SE33 locus

Platinum quality mitogenome haplotypes from United States populations 2020 mitogenome

Sequence-based U.S. population data for 7 X-STR loci 2020 7 X-STR

Lisa A. Borsuk[a*], Carolyn R. Steffen [a], Kevin M. Kiesler [a], Peter M. Vallone [a], and Katherine B. Gettings [a]

a. U.S. National Institute of Standards and Technology, Biomolecular Measurement Division, 100 Bureau Drive, Gaithersburg, MD 20899, USA

*Corresponding author at: U.S. National Institute of Standards and Technology, Biomolecular Measurement Division, 100 Bureau Drive, Gaithersburg, MD 20899, USA

5

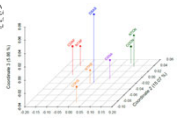
Sequencing of mitochondrial genomes

Kevin Kiesler

genes

Platinum-quality mitogenome haplotypes from United States metapopulations

Principal coordinate analyses of the eight datasets



Dataset	# of Samples Attempted	# of Samples in Finalized Data	Pooling		Removed			
			Long-range	Four-amplicon	Failed	Mixed	Duplicates	Related
CDAF	124	112	112	-	8	2	2	-
CDCN	119	112	112	-	5	2	-	-
CMH	113	109	109		1	3		
MIAF	258	254	253	1	1	1	-	-
NECN	262	260	258	2	1	-	-	1
NTHS	139	138	138		-	-	-	1
PKAS	175	169	165	4	3	3	-	-
RNAA	175	171	156	13	1	3	-	-
Total	1365	1327	1303	24	20	14	2	2

In **collaboration with the Armed Forces DNA Identification Laboratory** 1,327 mitochondrial genomes were sequenced

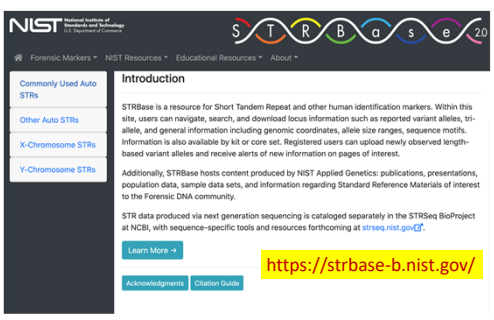
656 NIST U.S. population samples were sequenced at NIST on the Illumina MiSeq platform

Protocols optimized to reduce error and allow for 'platinum quality' complete mtGenomes

Data was submitted to EMPOP

6

STRBase



Lisa Borsuk

Soft release in 2021

https://strbase-b.nist.gov/

Project and training pages updated!

7

Additional Things Going on…

- **Forensics@NIST talks are available:**
  - https://www.nist.gov/news-events/events/2020/11/forensicsnist-2020
- **Katherine is presenting in a workshop at AAFS February 2021**
  - STR Wars: The Rise of Sequencing
- **NIST Validation Symposium**
  - More information coming soon!
  - POC: Robert Ramotowski

8

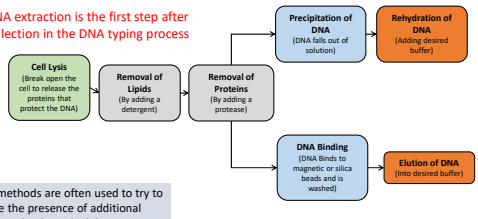
What else is going on?

Examination of "Front End" methods in DNA typing

9

DNA Extraction

DNA extraction is the first step after collection in the DNA typing process

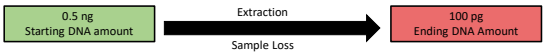


Purification methods are often used to try to eliminate the presence of additional proteins, lipids, and inhibitors

10

Examination of Front End Methods in DNA Typing

- **Problem**: Amount of sample loss during the extraction process is unknown for commonly used extraction methods
  - Low extraction efficiency cold result in overall lower sample quantity
  - May fail to yield full STR profiles or minor components in mixtures



Methods for determining extraction efficiency and sample loss vary

11

Relative Extraction Efficiency

- Measuring recovery of DNA compared to another method or technique
  - Comparing current method in a laboratory to a new method to bring online
- Reporting in the literature when comparing methods
  - Full vs. Partial profiles or allele/locus counting
    - Assumes that a full profile is equlivant to 100% efficiency
  - Total peak height measurement
    - Assumes sumation of peak heights from known sample input to be the theshold of 100% efficiency

Measurement is made when comapring known samples in the genoytping phase between extraction protocols without accounting for the quantity of material obtained or lost

12

Absolute Extraction Efficiency

- Measures the ratio of the amount of DNA recovered (quantity) to the original amount of DNA (known) after extraction

$$\frac{\text{DNA Post Extraction}}{\text{Original Amount DNA}} = \text{Absolute Extraction Efficiency}$$

Original amount **MUST** be known

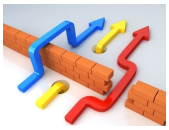
- Offers the ability to evaluate individual extraction processes and their efficiency independent of another method

13

Testing Absolute Extraction Efficiency

Placing a **known amount** of DNA into the extraction process and determining the **amount recovered**

Limitations



DNA Source: known amounts of DNA as a starting material
- Extracted/naked DNA is not an accurate surrogate (no lysis)
- Measuring starting material (cell counting, white blood cells)

Extraction Protocol: Understanding if different protocols perform differently
Organic Phenol-Chloroform is thought as the gold standard
- Differing commercial chemistries
  - Labs often "tweak" to better perform in their hands

Measurement: Downstream calculation of quantity
- qPCR is relative to a standard (need appropriate standard)
- qPCR chemistry bias and target differences

14

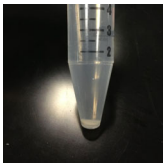
Component A of SRM 2372a: Human DNA Quantitation Standard

Known concentration of 49.8 ng/µL
Determined by ddPCR

Freshly collected whole blood

Known WBC of 4.6 x10^3 per µL
WBC reported by blood bank

Washed cell suspension in dPBS



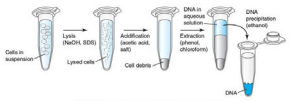
Known cell count of 1x10^6 per mL
Determined by flow cytometry

DNA Sources

15

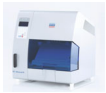
Extraction Methods: Phenol Chloroform (Organic)

- Phenol Chloroform (Organic)
  - Often referred to as the "gold standard"
  - Proteinase K digestion of the cells
  - Equal volumes of Phenol Chloroform added
  - Phase lock gel tubes used for promoting separation
  - DNA was precipitated with Ethanol and resolubilized with TE^{-4} buffer



16

Extraction Methods: Qiagen Commercial Chemistries



QIAamp Spin Columns

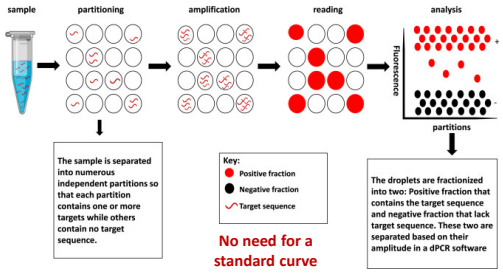
- Manual method commonly used in forensic DNA laboratories
- Silica columns for collection of DNA
- Elution in proprietary buffer
  - Similar to TE^{-4}

Qiagen EZ1 Advanced XL

- Robotic purification instrument
- Cell lysis takes place on the benchtop in a thermomixer
- Purification with paramagnetic bead collection
- Elution in TE^{-4}

17

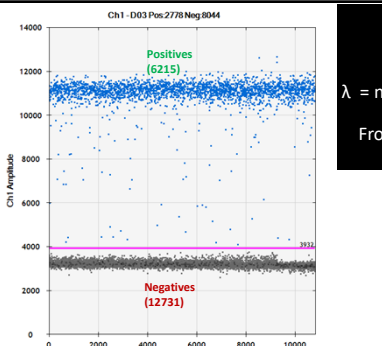
Measurement Method: Droplet Digital PCR



No need for a standard curve

dPCR is counting *accessible* amplifiable targets

18



$\lambda = -\ln(N_{neg}/N_{tot})$

λ = number of targets per partition

From here, concentration can be calculated.

$\lambda = -\ln(12731/18946)$

$\lambda = 0.3976$

dPCR is counting *accessible* amplifiable targets

19

Converting copies per nanoliter to nanograms nuclear DNA per microliter

$$[nDNA]\frac{ng}{\mu L} = \left(\frac{\lambda \text{ copies of target}}{\text{droplet}}\right)\left(\frac{\mu L \text{ mixture}}{F\ \mu L \text{ sample}}\right)\left(\frac{\text{droplet}}{V \text{ mixture}}\right)\left(\frac{HHGE}{r \text{ target}}\right)\left(\frac{n \text{ base pairs}}{HHGE}\right)\left(\frac{\overline{w}\,g}{\text{mol base pairs}}\right)\left(\frac{\text{mol base pairs}}{6.022\,10^{23} \text{ base pairs}}\right)\left(\frac{10^3\,nL}{\mu L}\right)\left(\frac{10^9\,ng}{g}\right) \quad (1)$$

DNA in ng/µL = 3.301 x (λ/(Dilution * Droplet Vol)

DNA in ng/µL = 3.301 x (0.3976/(0.25 * 0.07349)

DNA = 71.4 ng/µL

$\lambda = 0.3976$

***Duewer** DL, Kline MC, Romsos EL, Toman B. Anal Bioanal Chem. 2018 May;410(12):2879-2887*

20

Four DNA input amounts were tested in replicates of 5 for each extraction method

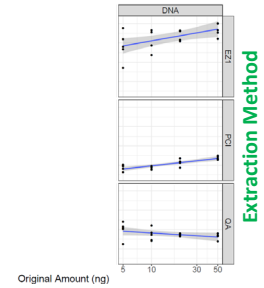
60 Samples per DNA Source

	Amount (ng)	# of Cells	Uncertainty (± # Cells)	# of Replicates
Extracted DNA	50	8,333	833	5 per amount (20 per DNA Source)
	20	3,333	333	
	10	1,667	167	
	5	781	78	
Cells	38	6,250	313	5 per amount (20 per DNA Source)
	19	3,125	156	
	9	1,563	78	
	5	781	39	
Blood	276	46,000	2,300	5 per amount (20 per DNA Source)
	138	23,000	1,150	
	28	4,667	233	
	14	2,333	117	

60 Samples per Extraction Method

21

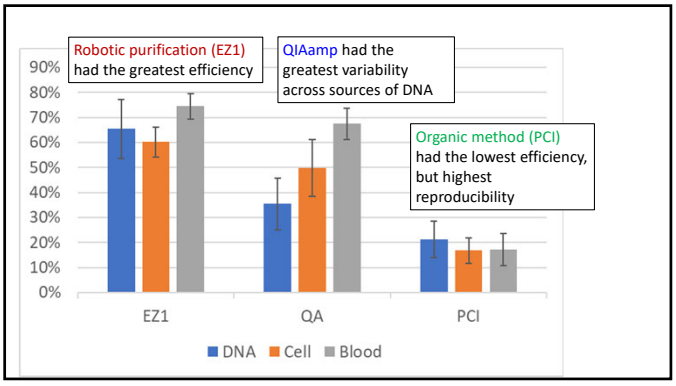
DNA Source



The efficiency of the extraction methods are different between methods but similar across DNA sources.

This demonstrates that the efficiency of extraction is ***independent of cell type*** and ***dependent on method***

22



23

Conclusions

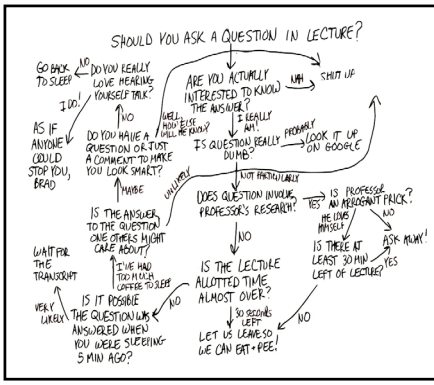
- Different methods yield different efficiencies
- Consistency across different DNA sources for each method
- Increase in DNA starting amount showed a trend toward increased efficiency

24

Work in process…

- Further experimental data to **confirm repeatability** of these measurements
- Optimization of extraction processes
  - Incubation time, purification method, addition of carrier RNA, microcon concentration (organic), varying input amounts, etc
- Examination of alternate extraction methods and operators
- Addition of a substrate to the extraction process

25



THANK YOU FOR YOUR ATTENTION

Questions?
erica.romsos@nist.gov
1-301-975-5107

Acknowledgements
Margaret Kline
David Duewer
Steven Lund

26