Forensic Mathematics and 9/11

Jonathan Hoyle

February 22, 2012

Nazareth COLLEGE

Overview

- Introduction
- World Trade Center Project
  - 9/11 and NYC
  - Direct Matching (STR Analysis)
  - Kinship Analysis
  - Mitochondrial DNA
  - SNP's
- Summary
- Q & A

Nazareth COLLEGE

Introduction

- My name is Jonathan Hoyle
- Both my Undergraduate (University of Delaware) and Graduate (University of Michigan) studies were in Mathematics with a Computer Science minor
- From 2001-2005, Mathematician and Software Engineer with Gene Codes Corp in Ann Arbor, MI
- Involved with ***M-FISys*** (pronounced "emphasis"), the forensic identification software used to identify the victims of the World Trade Center attacks
- Currently with Eastman Kodak as Macintosh Software Architect for Consumer Inkjet Printing

Nazareth COLLEGE

9/11 and NYC

Ground Zero

- Two 110 story towers
- 15 buildings over 16 acres
- Six basement levels and four subway lines
- 24,000 gallons of jet fuel
- Fires burned at 1800°F for over 3 months
- 2 billion pounds of rubble
- Existing DNA tools incapable of handling this magnitude

Nazareth COLLEGE

Nazareth
COLLEGE


Barclay St.
West Broadway St.
Washington St.
Verizon Building
7 WTC (47 Stories)
Old Post Office Building
Vesey St.
6 WTC (9 Stories)
5 WTC (9 Stories)
American Express Building
Fulton St.
Millenium Hotel
Church St.
West St.
1 WTC (110 Stories)
Austin J. Tobin Plaza
Dey St.
Merrill Lynch Building
3 WTC (22 Stories)
2 WTC (110 Stories)
4 WTC (9 Stories)
Cortlandt St.
1 Liberty Plaza Building
Liberty St.
Washington St.
Greenwich St.
90 West Building
Bankers Trust
Credit: NOAA
September 23, 2001
Nazareth COLLEGE

The Victims

- Unknown number of casualties early on
- Some family members afraid to come forward
- 20,000 total remains
- Some victims found in up to 200 fragments
- Majority of remains required DNA analysis
- 2,753 total victims

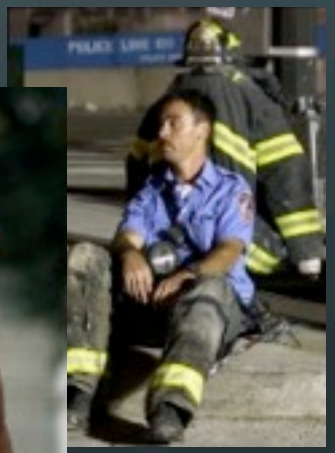
Nazareth COLLEGE

The Recovery

Thousands of rescue workers work around the clock from 9/11/01 through 5/30/02 in the recovery effort

Forensic DNA Identification Project with NYC Chief Medical Examiner's Office continued for three years

Nazareth COLLEGE



Nazareth
COLLEGE

Staten Island Triage

Trucks ship tons of debris from Ground Zero were sent to the Staten Island Recovery Site

Forensic anthropologists examine the debris to determine if it contains any human remains

Human remains found were sent to the Forensic Investigation Center in Albany, NY

Nazareth COLLEGE

Staten Island Recovery Site

Nazareth COLLEGE

Victim samples are typed using many DNA fingerprinting techniques, such as STR, MitoDNA & SNP to match against a personal effect

Family members are cheek swabbed for their DNA so that Kinship identification can be made when direct matching is not available

Nazareth COLLEGE

Software Development

- **September 17:** Armed Forces DNA Identification Lab [AFDIL] asked Gene Codes to update *Sequencher*™ for the Pentagon and Shanksville crashes
- **September 28:** Office of the Chief Medical Examiner [OCME] in New York contacts Gene Codes for new software for the World Trade Center project
- **October 15:** Development of *M-FISys* (Mass Fatality Identification System) underway
- **December 13:** *M-FISys* first release to OCME, followed by weekly releases thereafter
- Over the next three years, *M-FISys* is used to identify victims

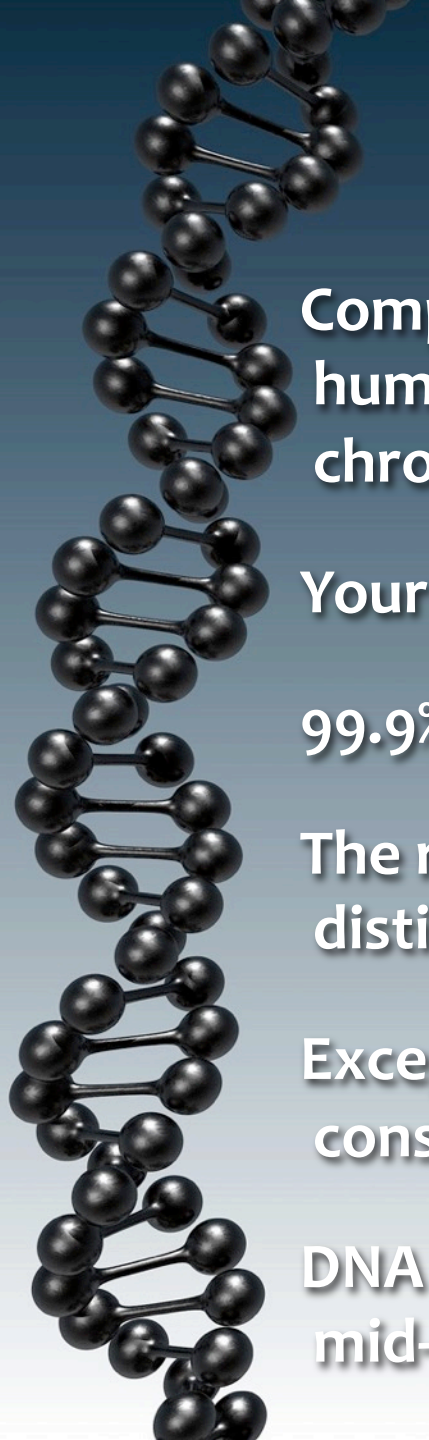
M-FISys Team Meeting

Direct Matching

STR Analysis

Nazareth COLLEGE

DNA



Composed of an alphabet of four chemicals: A, C, G, T, human DNA consists of 3.5 Billion base pairs across 23 chromosomes

Your DNA is inherited from your parents

99.9% of your DNA is shared with all of humanity

The remaining 0.1% (3.5 million base pairs) are what distinguishes us

Except for identical twins, each person's DNA is considered unique

DNA began to be used for forensic analysis in the mid-1980's

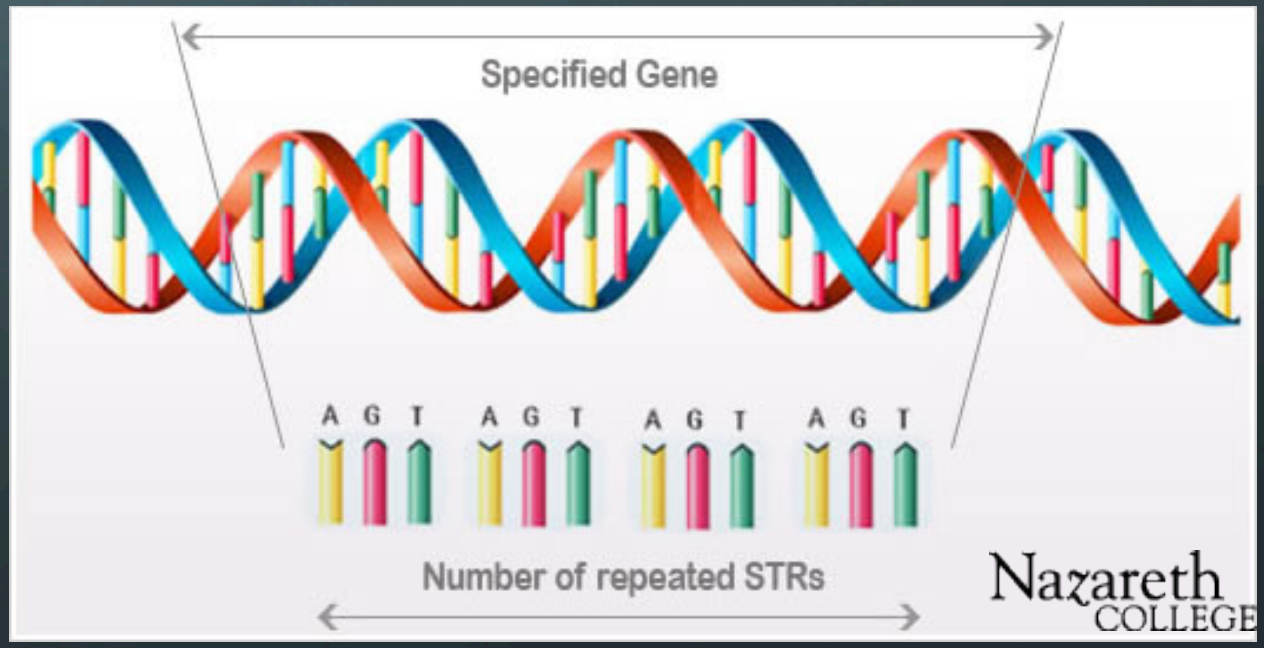
Nazareth COLLEGE

STR: Short Tandem Repeats

- A repeat of a short sequence of bases (usually 4 or 5):

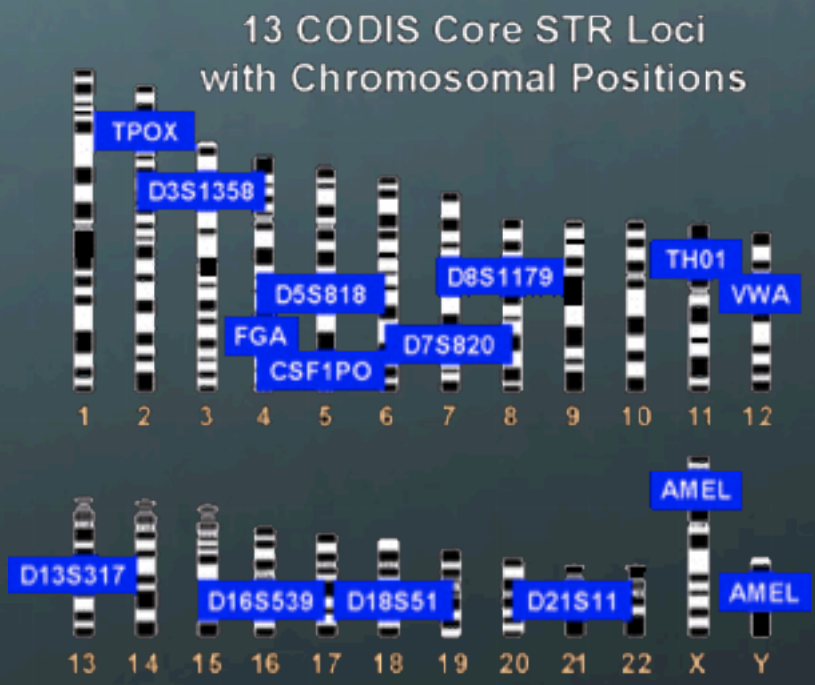
`...gcctggatagatagatagatagatagatgttta...`

- The above is repeated 5 times with a partial 3 bases
- The value for this STR locus is `5.3` (called its *allele*)
- Each locus contains a pair of alleles (inherited one from each parent), eg: `5.3 / 8`



STR Profiles

- In 1997, the FBI standardized on 13 core STR loci for its national database, CODIS
- STR analysis is the forensic standard for identification
- Includes two PowerPlex loci: Penta D and Penta E
- When both allele values are the same, it is called *homozygous*; otherwise, it is called *heterozygous*
- Gender: XX or XY
- These loci are "unlinked" and so independent



Nazareth COLLEGE

Allele Frequencies

TABLE 1—*U.S. Caucasian allele frequencies for 15 autosomal STR loci (N =302).*

Allele	CSF1PO	FGA	TH01	TPOX	VWA	D3S1358	D5S818	D7S820	D8S1179	D13S317	D16S539	D18S51	D21S11	D2S1338	D19S433
5	--	--	0.002	0.002	--	--	--	--	--	--	--	--	--	--	--
6	--	--	0.232	0.002	--	--	--	--	--	--	--	--	--	--	--
7	--	--	0.190	--	--	--	0.002	0.018	--	--	--	--	--	--	--
8	0.005	--	0.084	0.535	--	--	0.003	0.151	0.012	0.113	0.018	--	--	--	--
8.1	--	--	--	--	--	--	--	0.002	--	--	--	--	--	--	--
9	0.012	--	0.114	0.119	--	--	0.050	0.177	0.003	0.075	0.113	--	--	--	--
9.3	--	--	0.368	--	--	--	--	--	--	--	--	--	--	--	--
10	0.217	--	0.008	0.056	--	--	0.051	0.243	0.101	0.051	0.056	0.008	--	--	0.002
10.3	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
11	0.301	--	0.002	0.243	--	0.002	0.361	0.207	0.083	0.339	0.321	0.017	--	--	0.005
12	0.361	--	--	0.041	--	--	0.384	0.166	0.185	0.248	0.326	0.127	--	--	0.081
12.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0.002
13	0.096	--	--	0.002	0.002	--	0.141	0.035	0.305	0.124	0.146	0.132	--	--	0.253
13.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0.007
14	0.008	--	--	--	0.094	0.103	0.007	0.002	0.166	0.048	0.020	0.137	--	--	0.369
14.2	--	--	--	--	--	--	--	--	--	--	--	0.002	--	--	0.018
15	--	--	--	--	0.111	0.262	0.002	--	0.114	0.002	--	0.159	--	0.002	0.152
15.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0.035
16	--	--	--	--	0.200	0.253	--	--	0.031	--	--	0.139	--	0.033	0.050
16.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0.015
17	--	--	--	--	0.281	0.215	--	--	--	--	--	0.126	--	0.182	0.008
17.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	0.002
18	--	0.026	--	--	0.200	0.152	--	--	--	--	--	0.076	--	0.079	--
18.2	--	--	--	--	--	--	--	--	--	--	--		--	--	0.002
19	--	0.053	--	--	0.104	0.012	--	--	--	--	--	0.038	--	0.114	--
19.2	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
20	--	0.127	--	--	0.005	0.002	--	--	--	--	--	0.022	--	0.146	--
21	--	0.185	--	--	0.002	--	--	--	--	--	--	0.008	--	0.041	--
21.2	--	0.005	--	--	--	--	--	--	--	--	--	--	--	--	--
22	--	0.219	--	--	--	--	--	--	--	--	--	0.008	--	0.038	--
22.2	--	0.012	--	--	--	--	--	--	--	--	--	--	--	--	--
22.3	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--
23	--	0.004	--	--	--	--	--	--	--	--	--	--	--	0.118	--
23.2	--	0.003	--	--	--	--	--	--	--	--	--	--	--	--	--
24	--	--	--	--	--	--	--	--	--	--	--	--	--	0.123	--

J Forensic Sci, July 2003, Vol. 48, No. 4

http://www.cstl.nist.gov/strbase/pub_pres/Butler2003a.pdf

Allele Frequencies

- **According to the Hardy-Weinberg Principle:**
  - p^2 for homozygous alleles, *p* = frequency of allele
  - $2pq$ for heterozygous alleles, *p,q* = frequency of alleles
- **This assumes an sufficiently large population**
- **Since the population is relatively small, we must introduce the inbreeding coefficient θ:**
  - $p^2 + p(1-p)\theta$ for homozygous alleles
  - $2pq(1-\theta)$ for heterozygous alleles
- **Because θ is very small (0.03), we round on the side of being conservative:**
  - $p^2 + p(1-p)\theta$ for homozygous alleles
  - $2pq$ for heterozygous alleles

Profile Frequency

Locus	Victim	Sample	Equation	Prob	Likelihood
Gender	XY	XY	1/2	0.5000	2.00
D3S1358	14/16	14/16	2pq	0.0650	15.38
vWA	15/16	-			1.00
FGA	20/24	20/24	2pq	0.0401	24.95
D8S1179	12	12	$p^2+p(1-p)\theta$	0.0224	44.68
D21S11	28/31.2	28/31.2	2pq	0.0330	30.31
D18S51	14/17	-			1.00
D5S818	8/11	8/11	2pq	0.0106	94.47
D13S317	8	8	$p^2+p(1-p)\theta$	0.0108	92.62
D7S820	10/13	10/13	2pq	0.0172	58.13
D16S539	9	9	$p^2+p(1-p)\theta$	0.0117	85.12
TH01	6/9	-			1.00
TPOX	8/10	-			1.00
CSF1PO	10/12	10/12	2pq	0.1650	6.06
Penta D	9	-			1.00
Penta E	8/12	-			1.00

2.7E+14

Nazareth COLLEGE

Allelic Dropout

Locus	Victim	Sample	Equation	Prob	Likelihood
Gender	XY	XY	1/2	0.5000	2.00
D3S1358	14/16	14/16	2pq	0.0650	15.38
vWA	15/16	-			1.00
FGA	20/24	20/24	2pq	0.0401	24.95
D8S1179	12	12	$p^2+p(1-p)\theta$	0.0224	44.68
D21S11	28/31.2	28/31.2	2pq	0.0330	30.31
D18S51	14/17	-			1.00
D5S818	8/11	8	2p	0.3205	3.12
D13S317	8	8	$p^2+p(1-p)\theta$	0.0108	92.62
D7S820	10/13	10/13	2pq	0.0172	58.13
D16S539	9	9	$p^2+p(1-p)\theta$	0.0117	85.12
TH01	6/9	-			1.00
TPOX	8/10	-			1.00
CSF1PO	10/12	10/12	2pq	0.1650	6.06
Penta D	9	-			1.00
Penta E	8/12	-			1.00

9.0E+12

Nazareth COLLEGE

Likelihood Threshold

- How good is *good enough*?
- OCME wanted a minimum likelihood threshold set such that a chance of *any* mismatch would be less than one in a million
- What does this mean mathematically?
- Choose n such that identifications are satisfied when the likelihood value of a sample is $\geq 10^n$
- The probability of a fortuitous match of such a sample is thus $p = 10^{-n}$, no mismatch $q = 1 - 10^{-n}$
- Unknown population size, but early estimates assumed a population as high as 5000

Likelihood Threshold

- The probability of no mismatches is thus: q^{5000}
- The probability of any mismatch in the population:

$$1 - q^{5000} = 1 - (1 - 10^{-n})^{5000}$$

- For this to be a *"less than one in a million chance"* occurrence yields the equation:

$$1 - (1 - 10^{-n})^{5000} < 0.000001$$

- Solving for n we get:

$$n > log_{10}(1 - \sqrt[5000]{0.999999}) = 9.6989\ldots$$

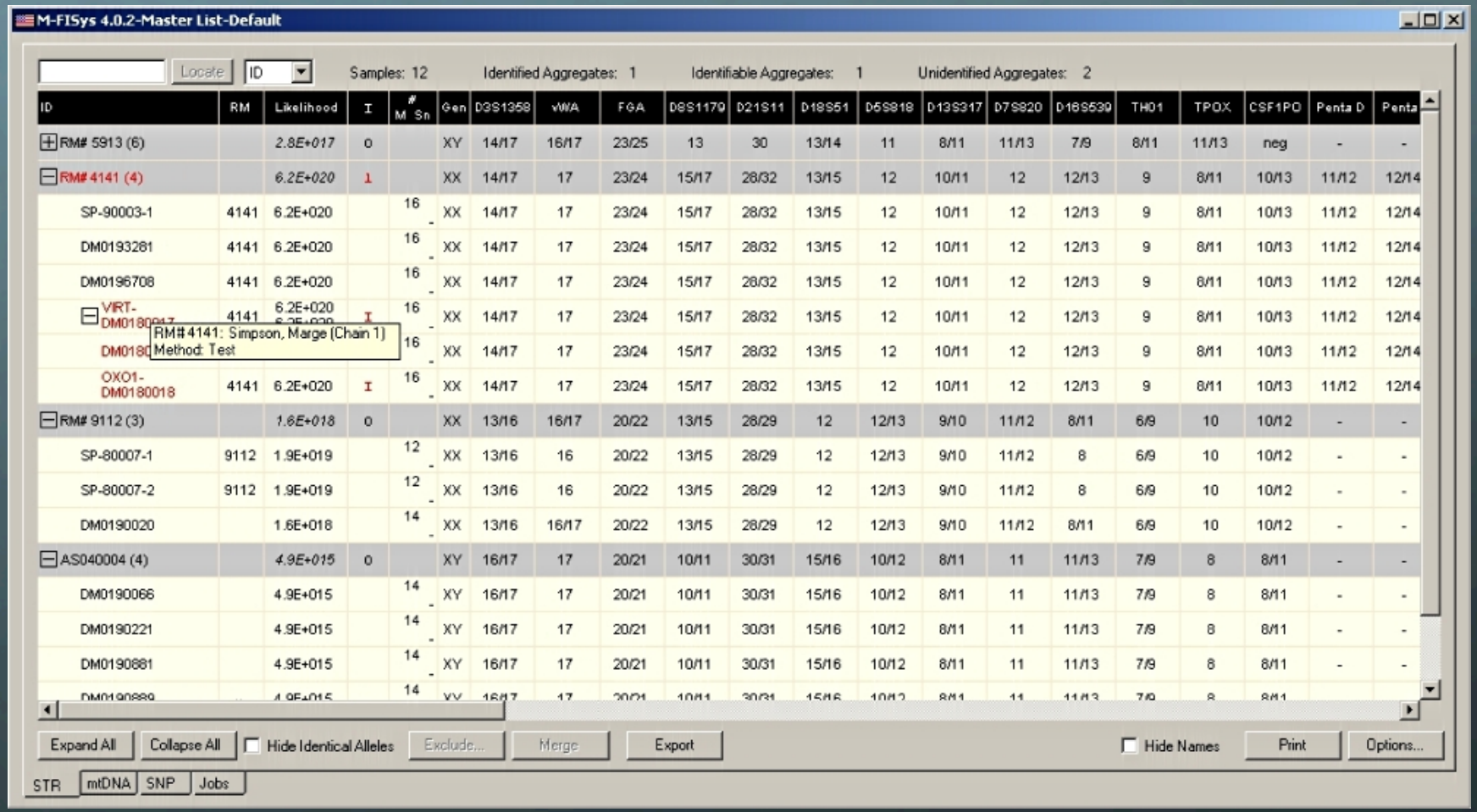
- Thus we choose $n = 10$

DNA Matching

- 12,000 personal effects were collected from families
- A sample can be identified to a personal effect if:
  - ✓ Has at least 7 common alleles
  - ✓ No more than one mismatch due to allelic dropout
  - ✓ Likelihood value $\geq 10^{10}$
- ~30% of the victim samples had complete profiles
- ~20% had partial profiles with likelihoods $\geq 10^{10}$
- ~20% had partial profiles with likelihoods $< 10^{10}$
- ~30% of the STR profiles had no data at all
- STR analysis alone would not be sufficient

Nazareth COLLEGE

M-FISys STR Form†



ID	RM	Likelihood	I	# M Sn	Gen	D3S1358	vWA	FGA	D8S1179	D21S11	D18S51	D5S818	D13S317	D7S820	D16S539	TH01	TPOX	CSF1PO	Penta D	Penta
RM# 5913 (6)		*2.8E+017*	0		XY	14/17	16/17	23/25	13	30	13/14	11	8/11	11/13	7/9	8/11	11/13	neg	-	-
RM# 4141 (4)		*6.2E+020*	1		XX	14/17	17	23/24	15/17	28/32	13/15	12	10/11	12	12/13	9	8/11	10/13	11/12	12/14
SP-90003-1	4141	6.2E+020		16	XX	14/17	17	23/24	15/17	28/32	13/15	12	10/11	12	12/13	9	8/11	10/13	11/12	12/14
DM0193281	4141	6.2E+020		16	XX	14/17	17	23/24	15/17	28/32	13/15	12	10/11	12	12/13	9	8/11	10/13	11/12	12/14
DM0196708	4141	6.2E+020		16	XX	14/17	17	23/24	15/17	28/32	13/15	12	10/11	12	12/13	9	8/11	10/13	11/12	12/14
VIRT-DM0180017	4141	6.2E+020	1	16	XX	14/17	17	23/24	15/17	28/32	13/15	12	10/11	12	12/13	9	8/11	10/13	11/12	12/14
DM0180				16	XX	14/17	17	23/24	15/17	28/32	13/15	12	10/11	12	12/13	9	8/11	10/13	11/12	12/14
OXO1-DM0180018	4141	6.2E+020	1	16	XX	14/17	17	23/24	15/17	28/32	13/15	12	10/11	12	12/13	9	8/11	10/13	11/12	12/14
RM# 9112 (3)		*1.6E+018*	0		XX	13/16	16/17	20/22	13/15	28/29	12	12/13	9/10	11/12	8/11	6/9	10	10/12	-	-
SP-80007-1	9112	1.9E+019		12	XX	13/16	16	20/22	13/15	28/29	12	12/13	9/10	11/12	8	6/9	10	10/12	-	-
SP-80007-2	9112	1.9E+019		12	XX	13/16	16	20/22	13/15	28/29	12	12/13	9/10	11/12	8	6/9	10	10/12	-	-
DM0190020		1.6E+018		14	XX	13/16	16/17	20/22	13/15	28/29	12	12/13	9/10	11/12	8/11	6/9	10	10/12	-	-
AS040004 (4)		*4.9E+015*	0		XY	16/17	17	20/21	10/11	30/31	15/16	10/12	8/11	11	11/13	7/9	8	8/11	-	-
DM0190066		4.9E+015		14	XY	16/17	17	20/21	10/11	30/31	15/16	10/12	8/11	11	11/13	7/9	8	8/11	-	-
DM0190221		4.9E+015		14	XY	16/17	17	20/21	10/11	30/31	15/16	10/12	8/11	11	11/13	7/9	8	8/11	-	-
DM0190881		4.9E+015		14	XY	16/17	17	20/21	10/11	30/31	15/16	10/12	8/11	11	11/13	7/9	8	8/11	-	-
DM0190889		4.9E+015		14	XY	16/17	17	20/21	10/11	30/31	15/16	10/12	8/11	11	11/13	7/9	8	8/11		

†presented in *The Mathematics of DNA Identification*, American Academy of Forensic Science, 2003

Nazareth COLLEGE

Kinship Analysis

Nazareth COLLEGE

Kinship Analysis

- Many personal effects lacked sufficient DNA
- Others were contaminated by external DNA
- Cheek swabs from family members were taken at Pier 94, so that a pedigree tree could be generated
- DNA profiles of victims are compared using the *Symbolic Kinship Program* algorithm (C. Brenner)
- A product of common loci can be used to produce kinship likelihood ratios (identifications $\geq 10^6$)
- A likelihood ratio is the ratio of the probability that the sample is a member of the given pedigree (H_1) over the probability that it is unrelated (H_0)

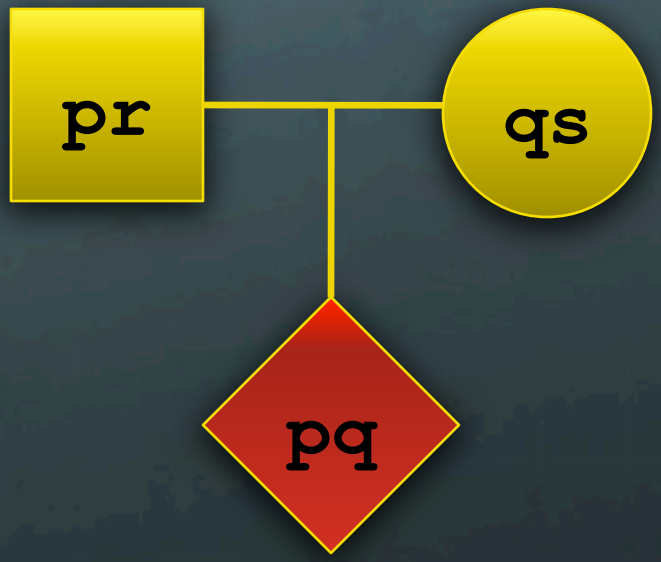
Nazareth COLLEGE

Kinship Example #1

Let **p**,**q**,**r**,**s** represent alleles and let p, q, r, s represent the probabilities of these alleles. (Let $p = 0.005, q = 0.02$)

A victim sample with allele **pq** and a pedigree containing two parents: father **pr** and mother **qs**

$$LR = P(H_1) \div P(H_0) = P(\mathbf{pq} \mid \mathbf{pr} + \mathbf{qs}) \div P(\mathbf{pq} \mid unrelated)$$



$$P(H_1) = \tfrac{1}{2} \times \tfrac{1}{2} \times 2pr \times 2qs = pqrs$$

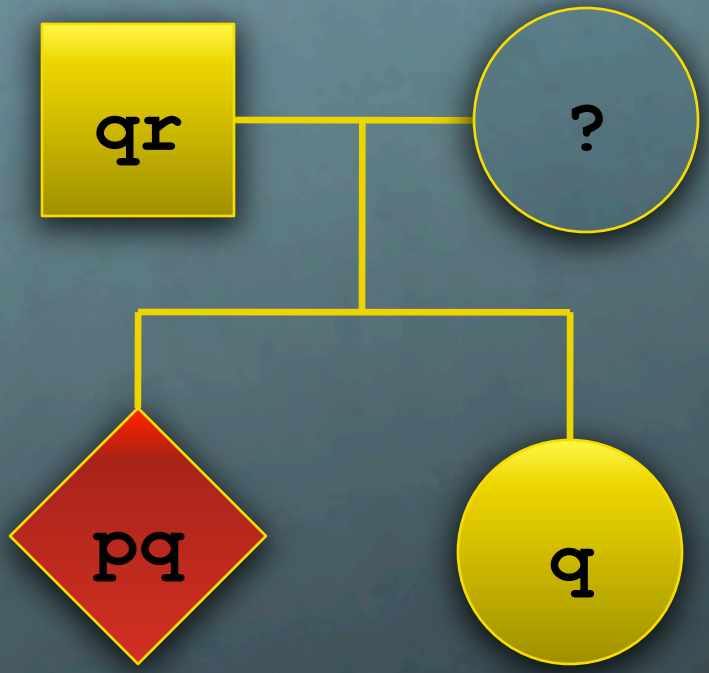
$$P(H_0) = 2pq2pr2qs = 8p^2q^2rs$$

$$LR = pqrs \div 8p^2q^2rs = 1/8pq$$

$$= \mathbf{1250}$$

Nazareth COLLEGE

Kinship Example #2



The same victim sample with Pedigree #2 containing father **qr** and sister **q**

For the **pq** victim sample to fit, the mother must be **pq** for H_1

$$P(H_1) = \frac{1}{4} \times \frac{1}{4} \times 2pq2qr = \frac{1}{4}\, pq^2r$$

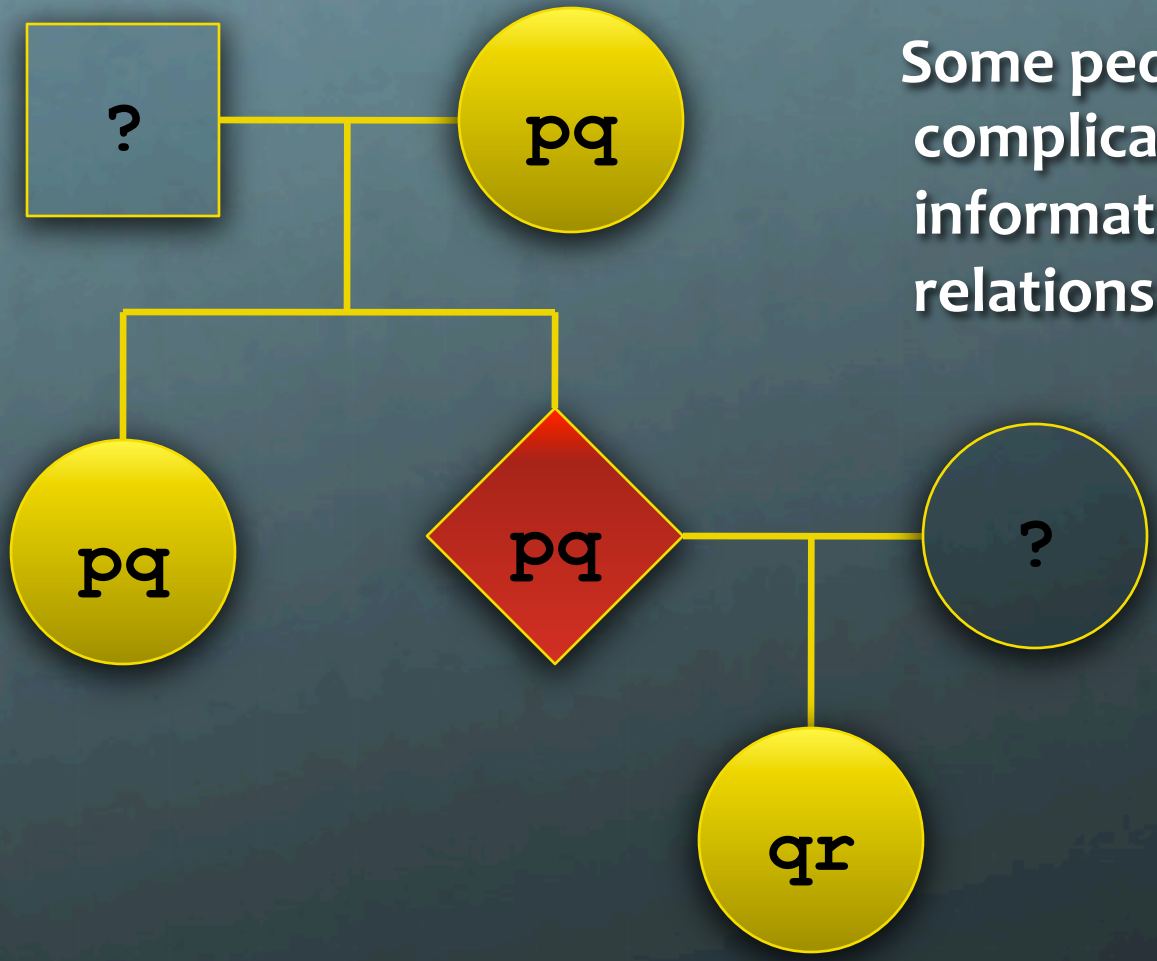
In H_0, mother may be **q** or **qx**, thus $P(H_0) = P(H_q) + P(H_{qx})$

$$P(H_q) = 2pq^4r \quad P(H_{qx}) = 2pq^3(1-q)r \rightarrow P(H_0) = 2pq^3r$$

$$LR = P(H_1) \div P(H_0) = 1/8q = \mathbf{6.25}$$

Nazareth COLLEGE

Kinship Example #3



Some pedigrees can be complicated, with partial information and extended relationships

Relations may involve half siblings, cousins and any number of combinations

$$LR = (1+p+q) / 8pq = \mathbf{1281.25}$$

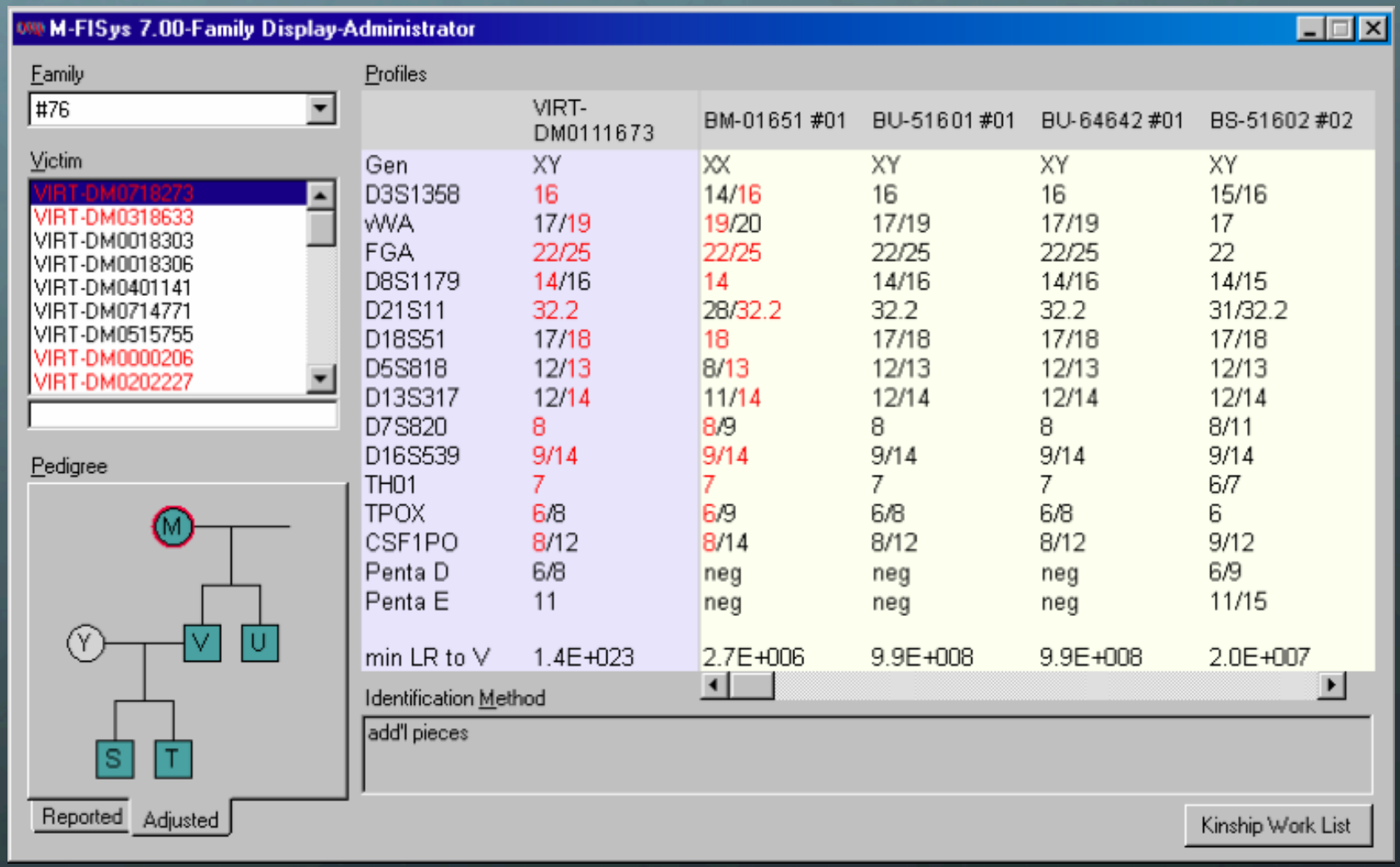
Nazareth COLLEGE

Kinship Equations

	VIRT-DM8180705	☑ BM-50527 #...	☑ BU-50527 #03	☑ BD-05721 #02	VIRT-DM8180705
Gen	XY	XX	XX	XX	-
D3S1358	15/16	15	14/15	15/16	1/4p
vWA	15/18	14/15	14/18	16/18	(1+q)/8pq
FGA	23/24	24	23/24	21/24	(1+p)/4pq
D8S1179	10/13	10/13	10/13	13/14	(p+q+pp+2pq+qq)/(8ppq+8pqq)
D21S11	30/31	30/33.2	30/31	27/31	(1+q)/8pq
D18S51	12/17	12/13	12/16	12/14	(1+q)/8pq
D5S818	10	10/12	10/12	9/10	(1+p+q)/(4pp+4pq)
D13S317	8/11	8/11	11/13	11	(p+q)/8pq
D7S820	10/12	11/12	11/12	10/11	1/8q
D16S539	10/12	12	12	12	1/4q
TH01	8/9.3	7/9.3	7/10	8/9.3	1/8p
TPOX	8/9	8/9	8/9	8/11	(p+q+pp+2pq+qq)/(8ppq+8pqq)
CSF1P0	9	9/11	9/11	9/13	(1+p+q)/(4pp+4pq)
Penta D	-	-	-	-	-
Penta E	-	-	-	-	-
					-
Likelihood	1.01e+18	1.54e+17	4.33e+16	2.01e+17	99.990883%
Kinship LR	99.990883%	2.59e+5	4.59e+5	5.73e+5	

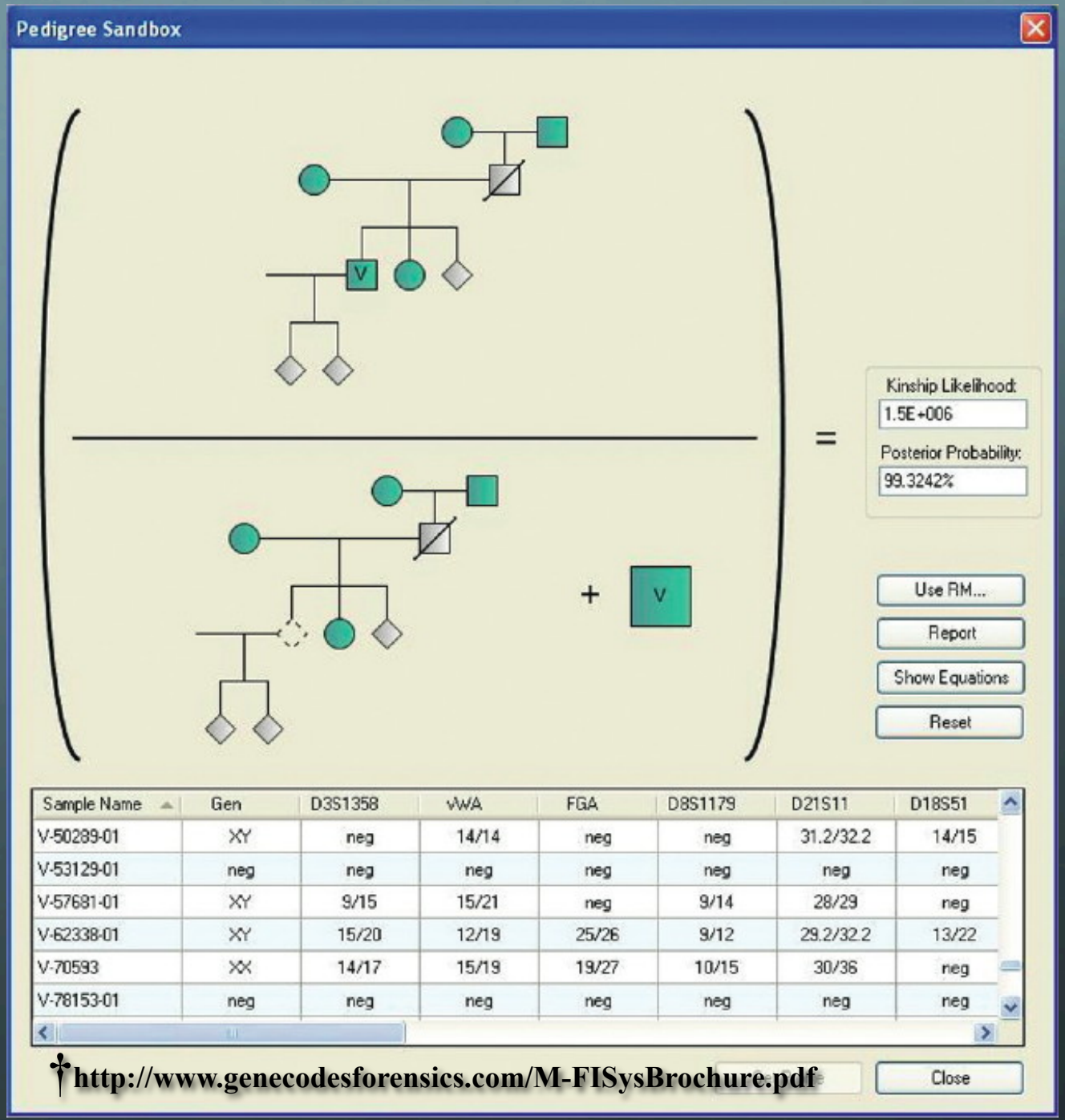
Nazareth COLLEGE

M-FISys Kinship Form†



†presented in *Bioinformatics for 9/11*, Dr. Simon Mercer, Bio IT World, 2004

Nazareth COLLEGE



M-FISys Pedigree Sandbox† displays the pedigree chart as a ratio

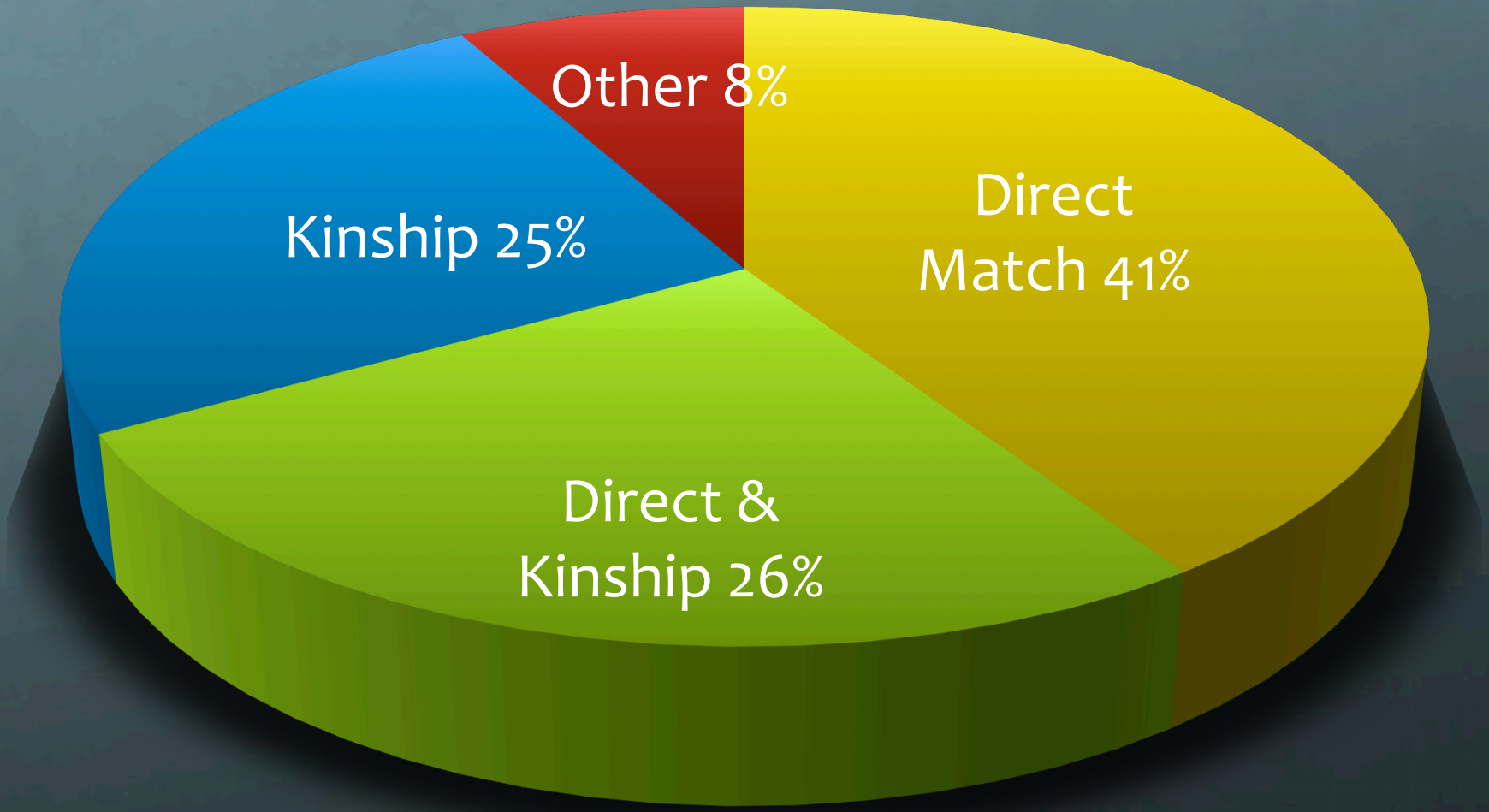
Pedigree chart is editable, making complex family relationships easy to manage

Used in other mass disaster forensics projects

†http://www.genecodesforensics.com/M-FISysBrochure.pdf

Nazareth COLLEGE

Match Methods on Remains

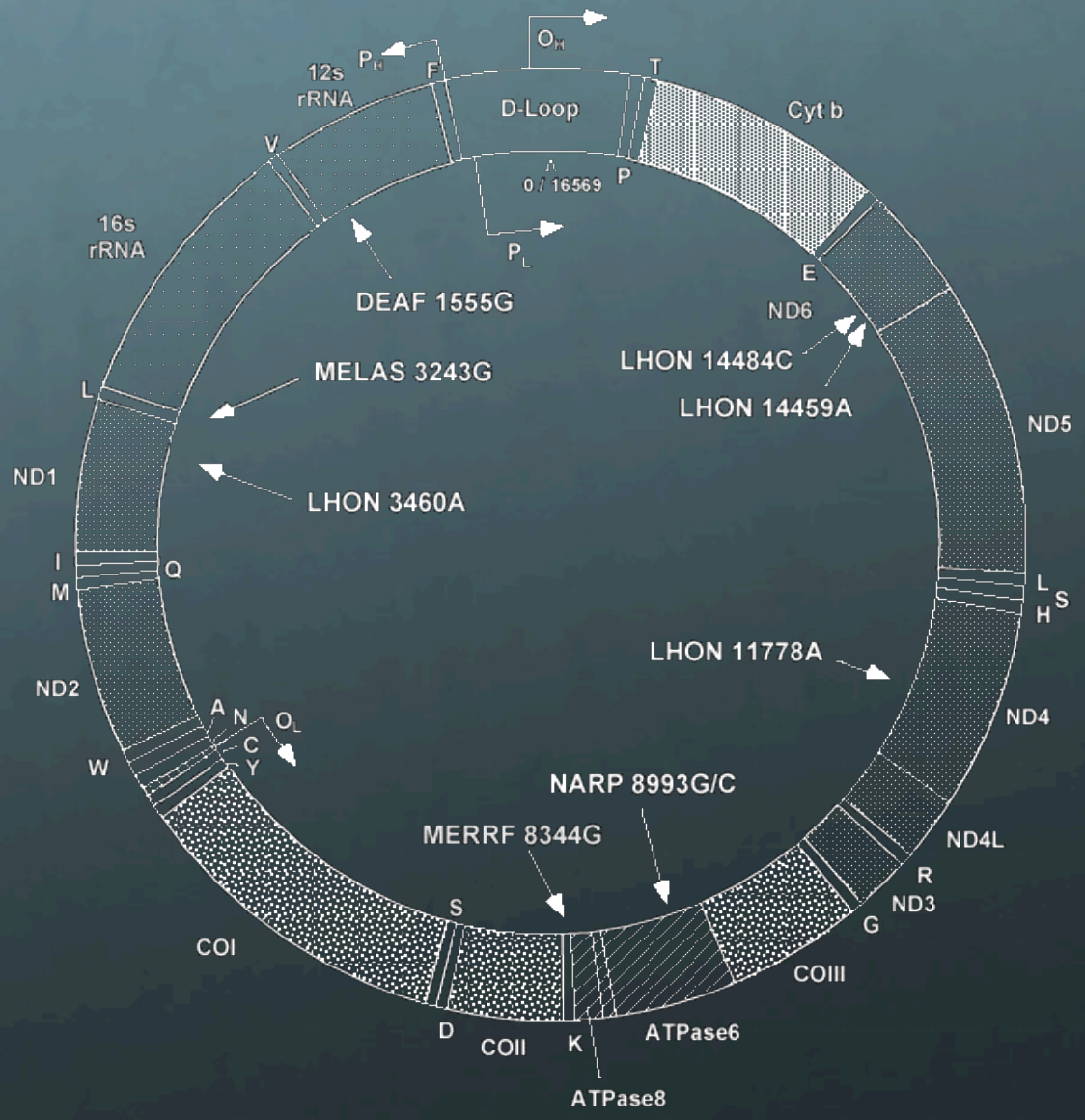


Mitochondrial DNA

Mitochondrial DNA

- Some victim samples were so degraded that STR analysis could not yield an identification
- Mitochondrial DNA (mtDNA) is heartier material, surviving under extreme conditions
- mtDNA is a 16,569-based circular genome
- Being circular (unlike the double helix of nuclear DNA), it is more stable and less prone to mutation
- Although each cell contains only two copies of nuclear DNA, it has thousands of copies of mtDNA
- mtDNA has been retrieved from ancient bones, including woolly mammoths and Neanderthals

mtDNA Map



Nazareth COLLEGE

mtDNA Typing

- Mito-typing involves direct sequencing of two highly variable regions of mtDNA (HV1, HV2)
- Differences from *the Anderson Sequence* (an internationally accepted standard) are tracked

mtDNA profile	
16093:	C
16224:	D
16311:	C
195:	C
263:	G
315.1:	C

- mtDNA is not unique, it is maternally inherited
- Thus matching can be done against a personal effect or from maternal relatives (eg: mother, full sibling, maternal half-sibs, not father or paternal half-sibs)
- 75% of the victims had maternal relatives providing sample mtDNA for potential matches

Nazareth COLLEGE

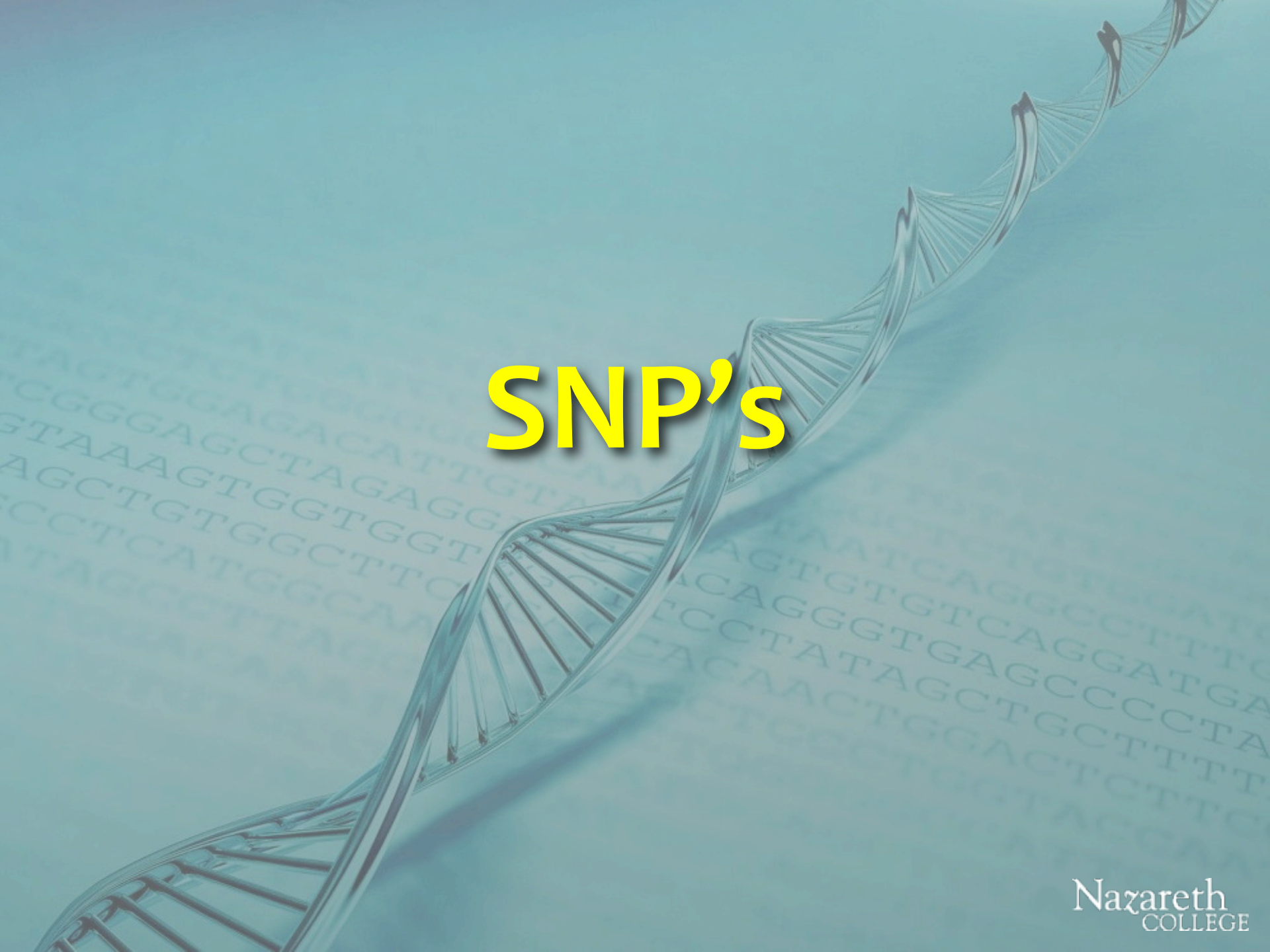
mtDNA Likelihood

- Likelihood for a given mitotype is determined by the number of hits x in the FBI's CODIS^{mt} database, of size n (~5000). Thus we have probability $p = x/n$.
- For a Binomial distribution, we have the equations: $\mu = p$ (mean) and $\sigma = \sqrt{p(1-p)}$ (standard deviation)
- The 95% confidence interval is defined by the formula:

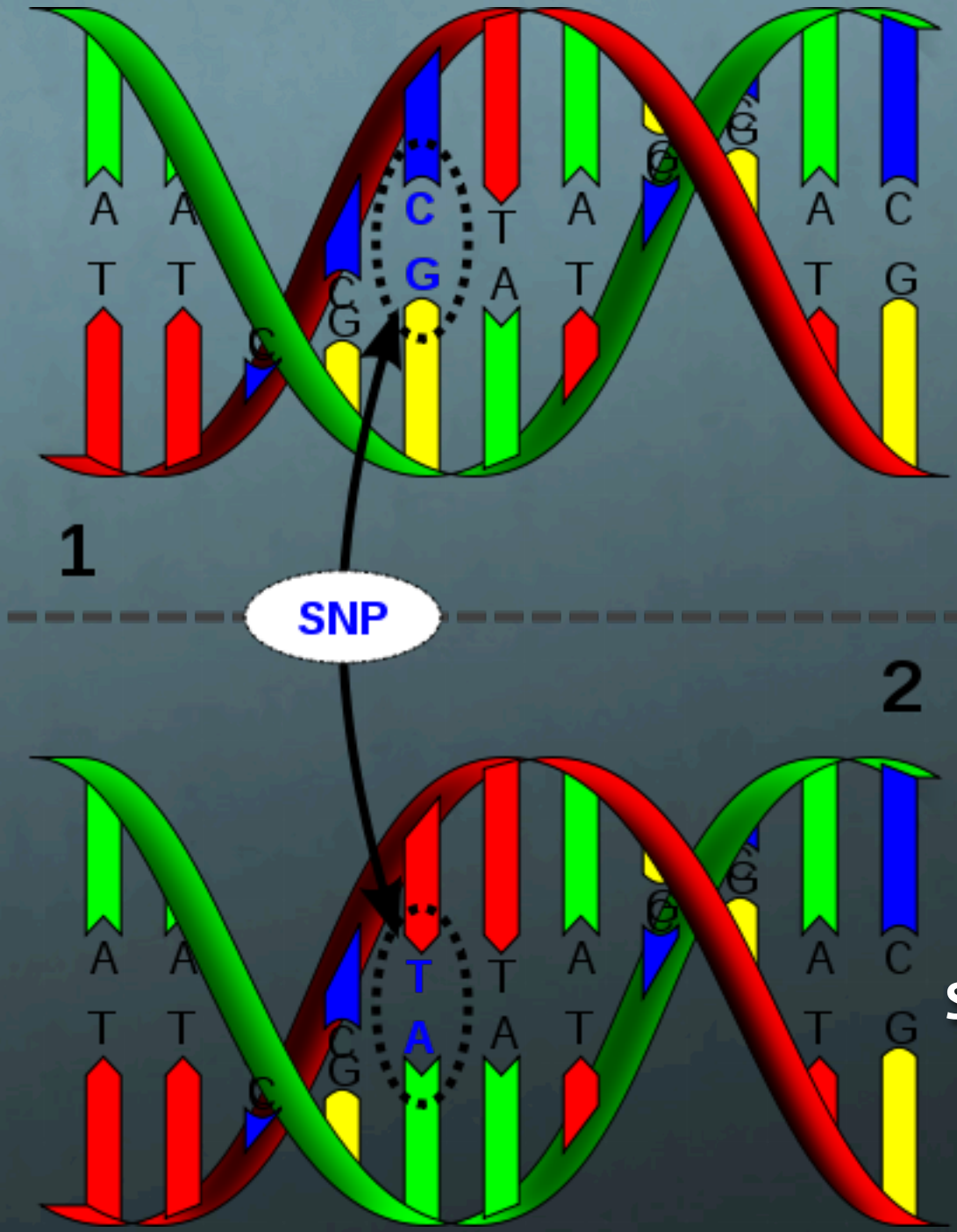
$$[\ \mu - 1.96\sigma/\sqrt{n},\ \mu + 1.96\sigma/\sqrt{n}\ ]$$

- Which reduces to an upper bound of $x/n + 2\sqrt{x(n-x)/n}$
- If no database entries, we use: $1 - \alpha^{1/n}$ with $\alpha = 0.05$
- mtDNA is independent of STR, so can be multiplied

Nazareth COLLEGE



Nazareth COLLEGE



Single Nucleotide Polymorphisms, representing single base differences from the genome

Useful for badly degraded samples

Mutation rate is 100,000 times lower than STR's

Occur on both nuclear and mitochondrial DNA

SNP's occur on average every 100 – 300 base pairs

Nazareth COLLEGE

SNP's

	Victim	PE	BF #01	BM #01	BU #02
Amel	CC	CC	-	TT	TT
65882	TC	TC	-	TC	TC
68532	-	TC	-	TC	CC
234217	CC	CC	-	TC	CC
231480	TT	TT	-	TT	TT
62059	-	-	-	TT	TT
56608	-	TC	-	TC	TC
61955	-	TT	-	TC	TC
220875	-	TT	-	TT	TT
58388	-	TT	-	TT	TT
63799	CC	CC	-	CC	TC
219561	TT	TT	-	TT	TT
60188	-	CC	-	CC	CC
182622	-	TC	-	TT	TT
85187	-	TC	-	TC	TC
212605	CC	CC	-	CC	CC
58091	-	TT	-	TT	TT
66026	-	TT	-	TC	TC
63836	-	CC	-	CC	CC
214373	TC	TC	-	TC	TT
238155	TT	TT	-	TT	TT

Two out of three SNP's involve replacing a C with a T

Of these, there is a panel of 70 chosen by Orchid BioSciences in for each C and T are equally likely

Many more SNP's are needed to reach STR likelihood levels

Used with Kinship Analysis

Nazareth COLLEGE

SNP Likelihood

- The Center for Genome Information concluded that although these 70 SNP's lack theoretical independence, allelic dependence was low enough for use in forensic identification
- Conservative likelihoods can be calculated even without the assumption of equi-probability. Heterozygous SNP's have a minimum likelihood of 2:

$$f = 2pq = 2p(1-p) \leq 0.5 \ \forall p \in [0,1]; \ \therefore L = 1/f \geq 2$$

- Thus the minimum likelihood of a SNP profile containing n heterozygous alleles is 2^n
- Average profile has ~35 heterozygous alleles, giving a minimum likelihood of $2^{35} \approx 10^{10}$

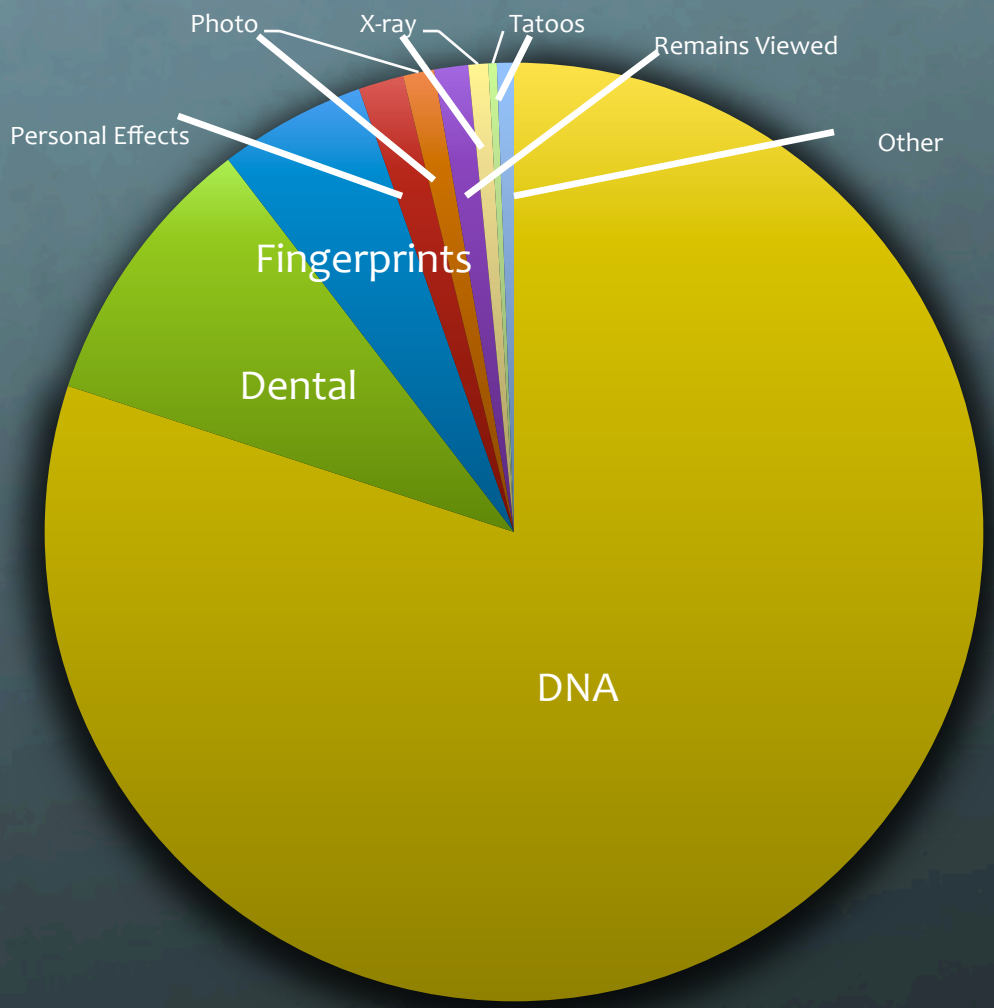
Summary

Statistics

- 2,753 victims (not including 10 hijackers)
- 21,814 total remains recovered
- 52,528 STR profiles
- 31,155 mtDNA profiles
- 16,938 SNP profiles
- Victims identified (as of 2/10/12): 1,633 (59%)
- Hijackers identified: 3 (out of 10)
- Remains identified: 12,811 (59%)

Nazareth COLLEGE

Identification Modalities



Of all the victims identified by a single modality, DNA represented 81% of the identifications

Of identifications made with multiple modalities, 87% included DNA

7/26/04 Statistics: *"Who They Were"* ©2005 Robert Shaler Forensics

Nazareth COLLEGE

Bibliography

Cash, Hoyle, Sutton, *Development Under Extreme Conditions: Forensic Bioinformatics in the Wake of the World Trade Center Disaster* (Pacific Symposium of BioComputing, 2003)

Hoyle, *Mathematics of Forensic DNA Identification* (American Academy of Forensic Science, 2003)

Mercer, *Bioinformatics for 9/11* (Bio IT World, 2004)

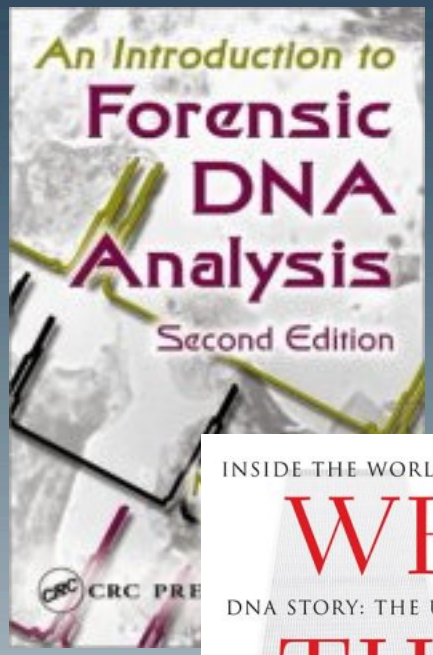
Shaler, *Who They Were: Inside the World Trade Center DNA Story* (ISBN 1416584471)

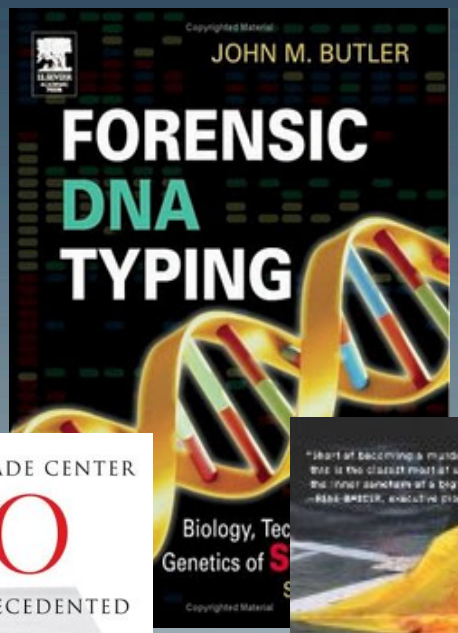
Ribowsky, *Dead Center: Behind the Scenes of the World's Largest Medical Examiner's Office* (ISBN 0061189405)

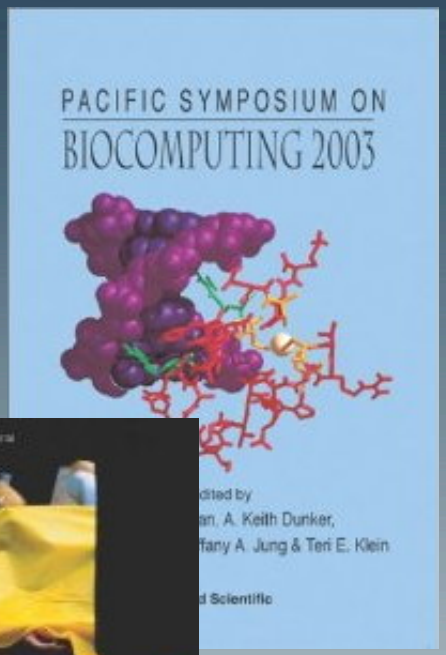
Butler, *Forensic DNA Typing* (ISBN 0121479528)

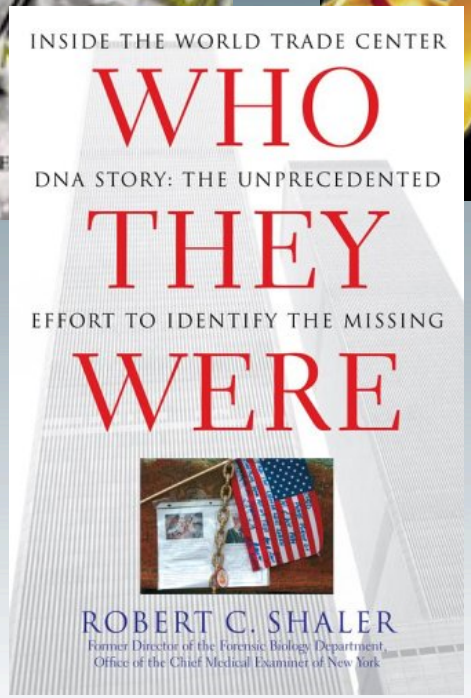
Nazareth COLLEGE

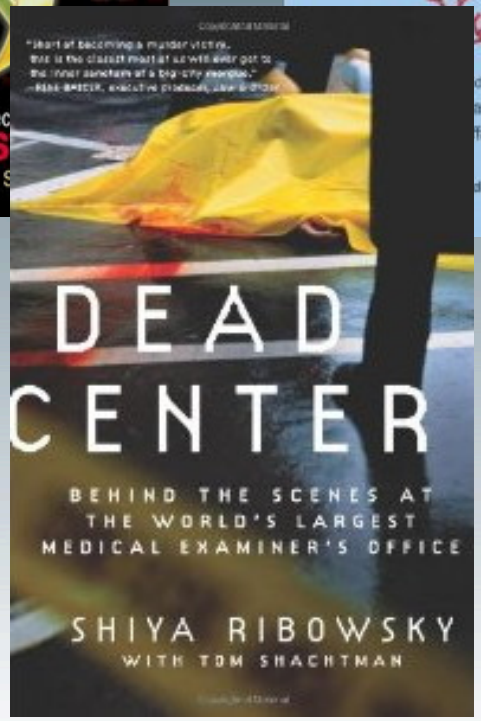
Further Reading











Nazareth COLLEGE

More Information

Web Site:

http://www.jonhoyle.com/MAASeaway

Slides:

http://www.jonhoyle.com/Presentations/ForensicMathNaz

Contact:

http://www.facebook.com/jonhoyle

jonhoyle@mac.com

Nazareth COLLEGE

Nazareth COLLEGE