Leiden University Medical Center

Now You See Me, Now You don't: fDNA and privacy

Peter de Knijff, Dept. of Human Genetics, Leiden University Medical Center

Body in a suitcase, Rotterdam 1927

Forensic DNA Research

crucial *a-priori* questions

- Do the crime-scene samples contain biological traces?
- What is the cellular origin of these biological traces?
- Can we make age-estimates of these biological traces?
  - Time of deposition
  - Age of the trace donor
- Who is the donor of these biological traces?
- Is there a relation between the traces and the crime?

Forensic DNA Research

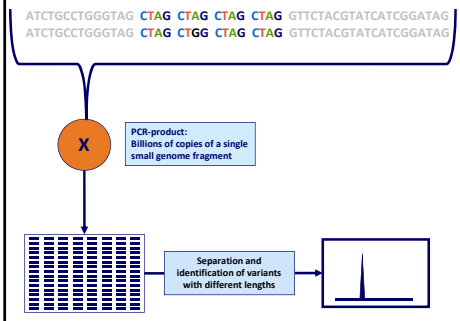
crucial *a-priori* questions

- Do the crime-scene samples contain biological traces?
- What is the cellular origin of these biological traces?
- Can we make age-estimates of these biological traces?
  - Time of deposition
  - Age of the trace donor
- **Who is the donor of these biological traces?**
- Is there a relation between the traces and the crime?

source level

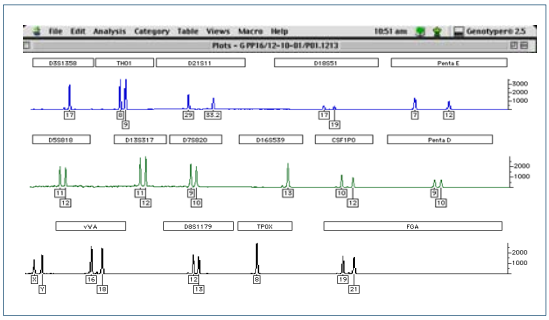
Forensic DNA Research

capillary electrophoresis-STR genotyping: the essence at one glance



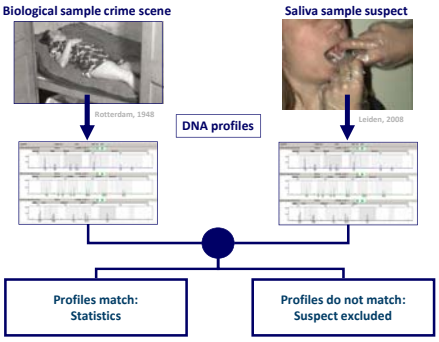
Forensic DNA Research

a multiplex STR profile



Forensic DNA Research

essence of fDNA profiling



Forensic DNA Research

crucial *a-priori* questions

- Do the crime-scene samples contain biological traces?
- What is the cellular origin of these biological traces?
- Can we make age-estimates of these biological traces?
  - Time of deposition
  - Age of the trace donor
- Who is the donor of these biological traces?
- Is there a relation between the traces and the crime?

activity level

Forensic DNA Research

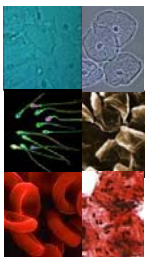
FGCN: towards fDNA 2.0

Forensic DNA Research

FGCN: towards fDNA 2.0

Research by Manfred Kayser (EMC) and Titia Sijen (NFI)

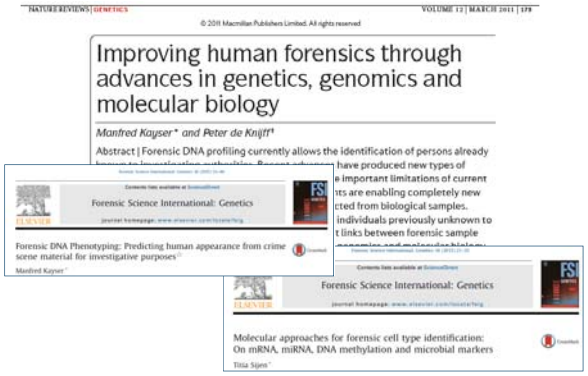
Who



What

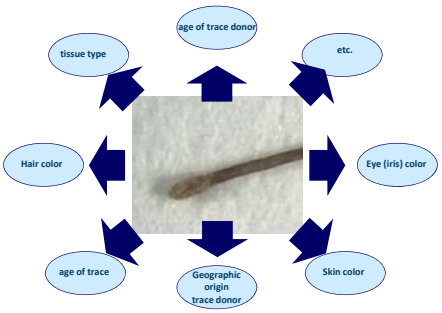
Forensic DNA Research

there is some progress



Forensic DNA Research

many tests, no DNA



Forensic DNA Research

some restrictions

- Contra-expertise requires retaining 50% of the trace / extract
- Some tests require RNA, not DNA
- Classical DNA profiling should always remain possible
- Legal restrictions

Forensic DNA Research

towards a one-for-all solution?

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsig

Shotgun metagenomics of biological stains using ultra-deep DNA sequencing

B. Brenig, J. Beck, E. Schütz

Table 2
Percentage of genomes present in the biological stain.

	% of all sequenced nucleotides
RepeatMasker database	
Repetitive elements (all mammals)	10.3
Porcine repetitive elements	10.0*
Human repetitive elements	5.8*
BLAST database	
Bacterial genomes	4.2
Viral genomes	0.3
Fungal genomes	2.7
Porcine genome	0.9
Human genome	0.13
Sequences of unidentified origin	81.5

Forensic DNA Research

towards single cell forensics

human enrichment

DNA (meta) genome identification

RNAseq eQTL profiling

Forensic DNA Research

Dutch legal basis

201

Wet van 8 mei 2003 tot wijziging van de regeling van het DNA-onderzoek in strafzaken in verband met het vaststellen van uiterlijk waarneembare persoonskenmerken uit celmateriaal

131

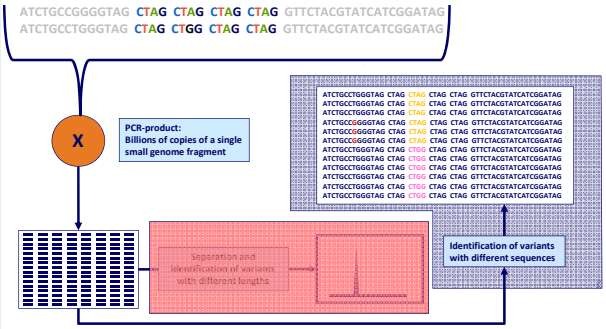
Besluit van 26 maart 2012, houdende inwerkingtreding van de wet van 24 november 2011 tot wijziging van het Wetboek van Strafvordering en de Wet DNA-onderzoek bij veroordeelden in verband met de introductie van DNA-verwantschapsonderzoek en DNA-onderzoek naar uiterlijk waarneembare persoonskenmerken van het onbekende slachtoffer en de regeling van enige andere onderwerpen (Stb. 2011, 555)

- Hair color
- Eye color
- Skin color
- Gender
- Geographic origin

- Kinship

Forensic DNA Research

introducing NGS/MPS into forensics: MPS-STR genotyping



Forensic DNA Research

the first MPS-STR papers

High-throughput sequencing of core STR loci for forensic genetic investigations using the Roche Genome Sequencer FLX platform

Sarah L. Fordyce, Maria C. Ávila-Arcos, Eszter Rockenbauer, Claus Børsting, Rune Frank-Hansen, Frederik Torp Petersen, Eske Willerslev, Anders J. Hansen, Niels Morling, and M. Thomas. P. Gilbert

Second generation sequencing allows for mtDNA mixture deconvolution and high resolution detection of heteroplasmy

ARTICLE IN PRESS

Mitchell M. Holland, Megan R. McQuillan, Katherine A. O'Hanlon

Forensic STR analysis using massive parallel sequencing

Christophe Van Neste, Filip Van Nieuwerburgh, David Van Hoofstat, Dieter Deforce

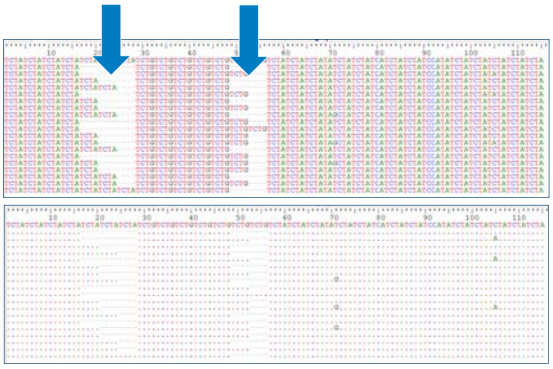
Forensic DNA Research

MPS-STR resulting data

Illumina Miseq (300bp array): up to 15.000.000 sequences for ~ € 1100

Forensic DNA Research

MPS-STR: repeat variation and SNPs



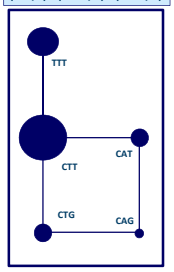
Forensic DNA Research

extra SNP variation: D7S820 as an example

D7S820	
Sequence-alleles	32
Fragment-alleles	10
Alleles with SNPs	22
Extra Alleles from SNPs	22

Fragment-allele	Seq-allele	Frequency Sequence-alleles NL	Him	Pop	Tot Freq
6	6_GATA[6]			0,005	0,002
7	7_GATA[7]	0,025		0,005	0,010
7	7_GATA[7]_+22T>A		0,005		0,002
8	8_GATA[8]	0,127	0,057	0,136	0,107
8	8_GATA[8]_+22T>A		0,005		0,002
8	8_GATA[8]_+65T>G			0,010	0,003
8	8_GATA[8]_-9C>T	0,044	0,191	0,066	0,099
9	9_GATA[10]_+44GAT>	0,005			0,002
9	9_GATA[9]	0,176	0,036	0,086	0,101
9	9_GATA[9]_+22T>A		0,010		0,003
9	9_GATA[9]_+65T>G		0,005	0,025	0,010
9	9_GATA[9]_-9C>T	0,015	0,072	0,056	0,047
10	10_GATA[10]	0,162	0,098	0,222	0,161
10	10_GATA[10]_+22T>A	0,049	0,015		0,022
10	10_GATA[10]_+22T>A_+65T>G			0,005	0,002
10	10_GATA[10]_+65T>G	0,010	0,031	0,035	0,025
10	10_GATA[10]_-9C>T		0,005	0,020	0,008
10,3	10.3_GATA[11]_22T>	0,005			0,002
11	11_GATA[11]	0,191	0,155	0,071	0,139
11	11_GATA[11]_+22T>A	0,015	0,041		0,018
11	11_GATA[11]_+22T>A_+65T>G			0,010	0,003
11	11_GATA[11]_+65T>G		0,021	0,066	0,029
11	11_GATA[11]_-9C>T		0,026	0,005	0,010
11	11_GATA[3]GGTA[1]GATA[7]			0,010	0,003
12	12_GATA[12]	0,108	0,175	0,066	0,116
12	12_GATA[12]_+22T>A	0,034	0,021		0,018
12	12_GATA[12]_+22T>A_+65T>G			0,040	0,013
12	12_GATA[12]_+65T>G			0,030	0,010
12	12_GATA[12]_-9C>T			0,035	0,012
13	13_GATA[13]	0,015	0,015		0,010
13	13_GATA[13]_+22T>A	0,010	0,005		0,005
14	14_GATA[14]	0,010		0,005	0,005

(-9 C/T) – (+22 T/A) – (+65 T/G)



Forensic DNA Research

Dutch legal basis

201

Wet van 8 mei 2003 tot wijziging van de regeling van het DNA-onderzoek in strafzaken in verband met het vaststellen van uiterlijk waarneembare persoonskenmerken uit celmateriaal

- Hair color
- Eye color
- Skin color
- Gender
- Geographic origin

131

Besluit van 26 maart 2012, houdende inwerkingtreding van de wet van 24 november 2011 tot wijziging van het Wetboek van Strafvordering en de Wet DNA-onderzoek bij veroordeelden in verband met de introductie van DNA-verwantschapsonderzoek en DNA-onderzoek naar uiterlijk waarneembare persoonskenmerken van het onbekende slachtoffer en de regeling van enige andere onderwerpen (Stb. 2011, 555)

- Kinship

Forensic DNA Research

option for MPS-STR: show repeat variation and mask SNPs

At present no concrete solutions, no International legal fine-tuning

We start offering MPS-STR genotyping per 01-01-2016, fully accredited
We will initiate a "possibilities" debate with relevant partners

Forensic DNA Research

no genetic privacy with publically available genome data

Identifying Personal Genomes by Surname Inference

www.sciencemag.org SCIENCE VOL 339 18 JANUARY 2013

Melissa Gymrek,[1,2,3,4] Amy L. McGuire,[5] David Golan,[6] Eran Halperin,[7,8,9] Yaniv Erlich[1]*

Here, we report that surnames can be recovered from personal genomes by profiling short tandem repeats on the Y chromosome (Y-STRs) and querying recreational genetic genealogy databases. We show that a combination of a surname with other types of metadata, such as age and state, can be used to triangulate the identity of the target. A key feature of this technique is that it entirely relies on free, publicly accessible Internet resources. We quantitatively analyze the probability of

Forensic DNA Research

no genetic privacy with publically available genome data

Bayesian method to predict individual SNP genotypes from gene expression data

NATURE GENETICS VOLUME 44 | NUMBER 5 | MAY 2012

Eric E Schadt[1,2], Sangsoon Woo[2,4,5] & Ke Hao[1,3,5]

in large populations. When inferring genotypes from an expression data set using eQTLs of the same tissue type (but from an independent cohort), we were able to resolve 99% of the identities of individuals in the cohort at $P_{adjusted} \leq 1 \times 10^{-5}$. When eQTLs derived from one tissue were used to predict genotypes using expression data from a different tissue, the identities of 90% of the study subjects could be resolved at $P_{adjusted} \leq 1 \times 10^{-5}$. We discuss the implications of deriving genotypic information from RNA data deposited in the public domain.

Forensic DNA Research

forensic DNA databases

Forensic DNA Research

our ultimate goal is single cell forensics

human enrichment

DNA (meta) genome identification

RNAseq eQTL profiling

Protein and metabolite profiling

Forensic DNA Research

our ultimate goal single cell forensics with encrypted matches

human enrichment

DNA (meta) genome identification

RNAseq eQTL profiling

Protein and metabolite profiling