Forensic Genetics and Legal Medicine
2019-2020

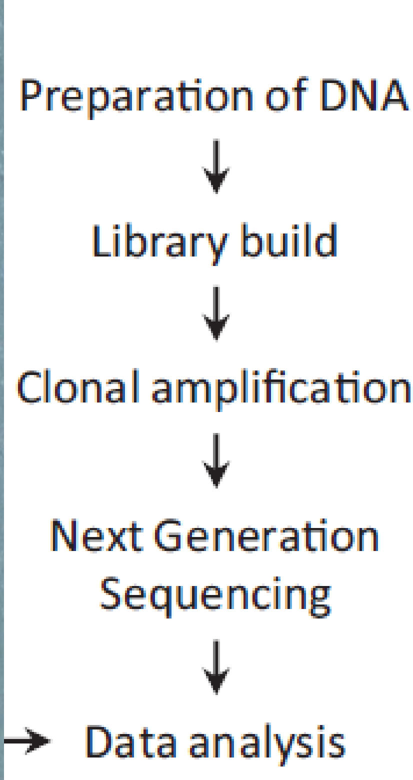
22th April 2020

Next Generation Sequencing (NGS) in forensics

"A revolution is not a bed of roses. A revolution is a struggle between the future and the past."

(Fidel Castro)

✓ Next-generation sequencing refers to non-Sanger-based high-throughput DNA sequencing technologies, with millions to billions of DNA strands sequenced in parallel (massive parallel sequencing, MPS)



Preparation of DNA

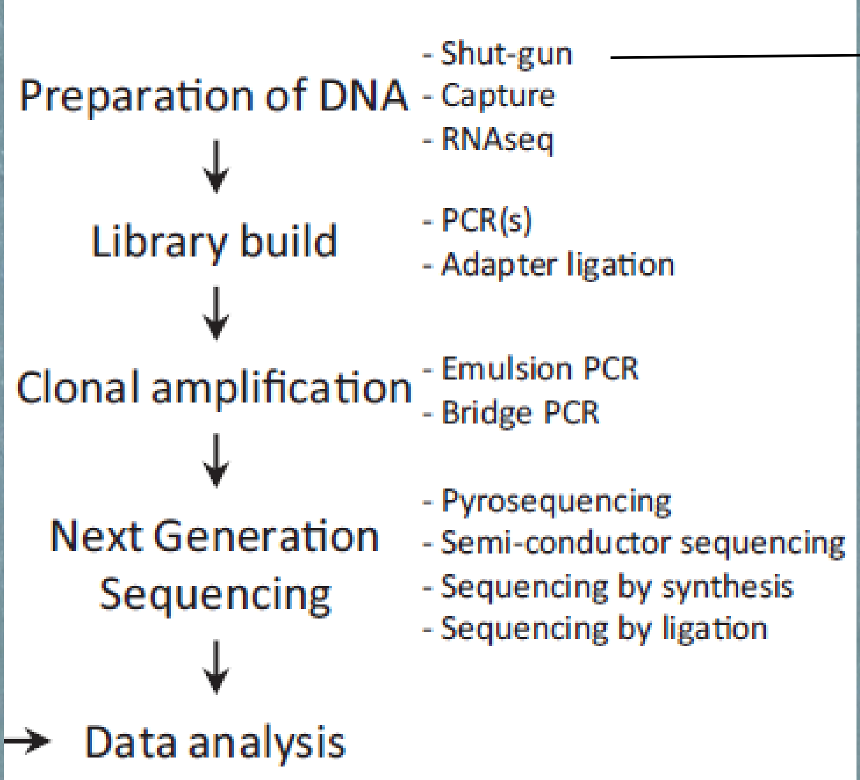
Library build

Clonal amplification

Next Generation Sequencing

Data analysis

- ✓ Next-generation sequencing refers to non-Sanger-based high-throughput DNA sequencing technologies, with millions to billions of DNA strands sequenced in parallel (massive parallel sequencing, MPS)



Preparation of DNA
- Shut-gun
- Capture
- RNAseq

↓

Library build
- PCR(s)
- Adapter ligation

↓

Clonal amplification
- Emulsion PCR
- Bridge PCR

↓

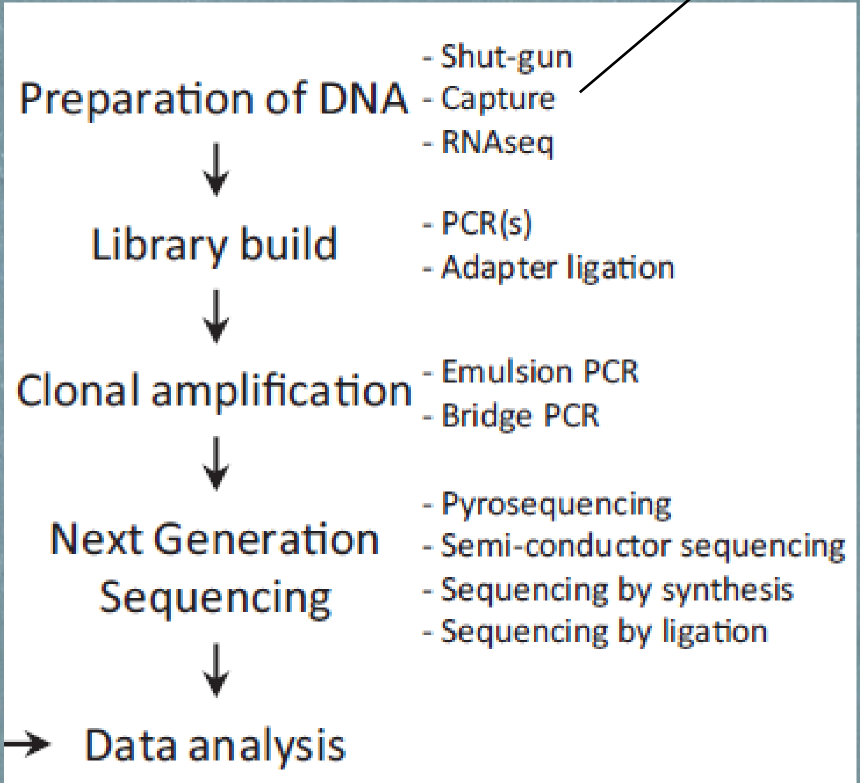
Next Generation Sequencing
- Pyrosequencing
- Semi-conductor sequencing
- Sequencing by synthesis
- Sequencing by ligation

↓

→ Data analysis

Sequencing of every double stranded DNA molecule in the sample material without any prior selection of targets

- 100 ng – 1 µg needed
- Low coverage
- Issues with STRs
- Bionformatically challanging



Preparation of DNA
- Shut-gun
- Capture
- RNAseq

↓

Library build
- PCR(s)
- Adapter ligation

↓

Clonal amplification
- Emulsion PCR
- Bridge PCR

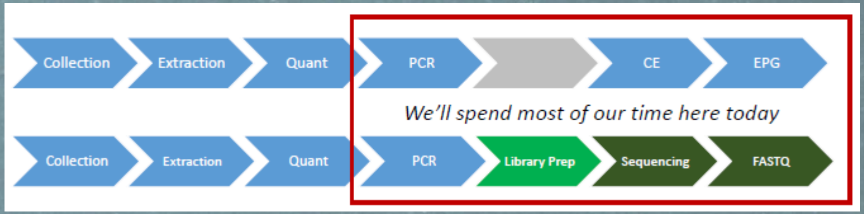
↓

Next Generation Sequencing
- Pyrosequencing
- Semi-conductor sequencing
- Sequencing by synthesis
- Sequencing by ligation

↓

→ Data analysis

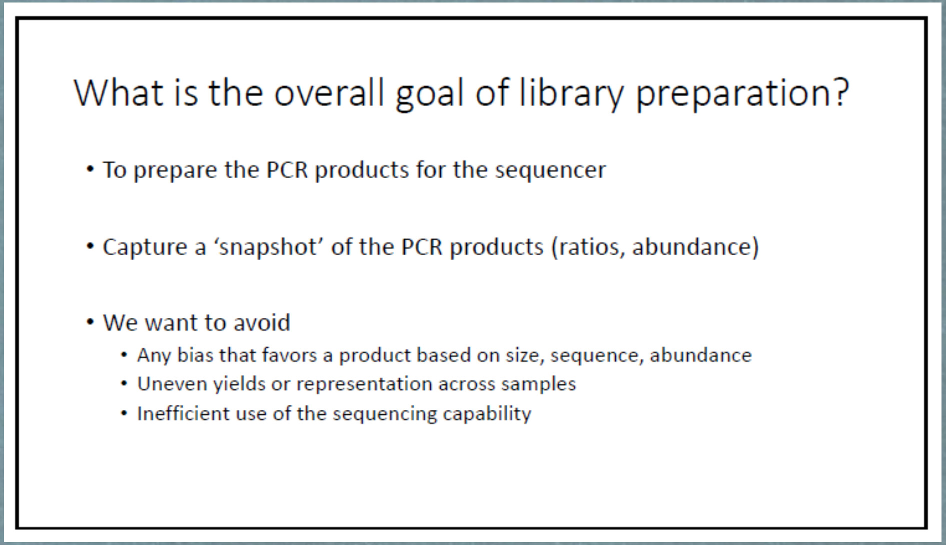
PCR continues to be the method that approaches the level of sensitivity required for forensic genetic case work (≤ 1 ng)



We'll spend most of our time here today

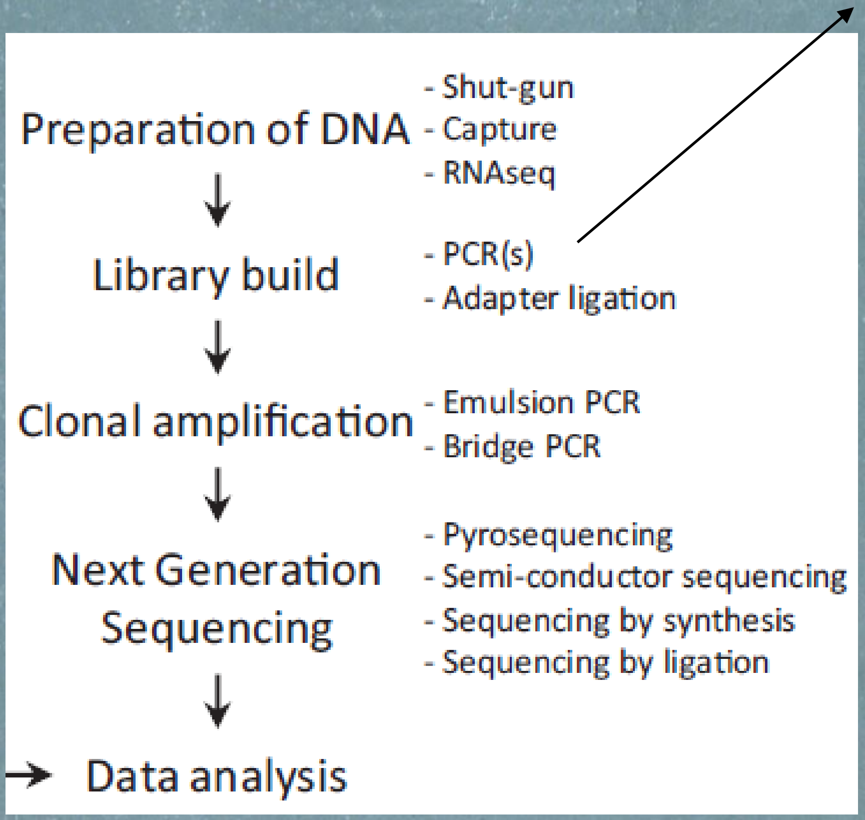
CE

NGS



What is the overall goal of library preparation?

- To prepare the PCR products for the sequencer
- Capture a 'snapshot' of the PCR products (ratios, abundance)
- We want to avoid
  - Any bias that favors a product based on size, sequence, abundance
  - Uneven yields or representation across samples
  - Inefficient use of the sequencing capability



Preparation of DNA
- Shut-gun
- Capture
- RNAseq

Library build
- PCR(s)
- Adapter ligation

Clonal amplification
- Emulsion PCR
- Bridge PCR

Next Generation Sequencing
- Pyrosequencing
- Semi-conductor sequencing
- Sequencing by synthesis
- Sequencing by ligation

Data analysis

primers are tagged with sequences needed for the downstream reactions

Use of "barcode" sequences in PCR primers or adapters allow for simultaneous sequencing of multiple samples in a single run

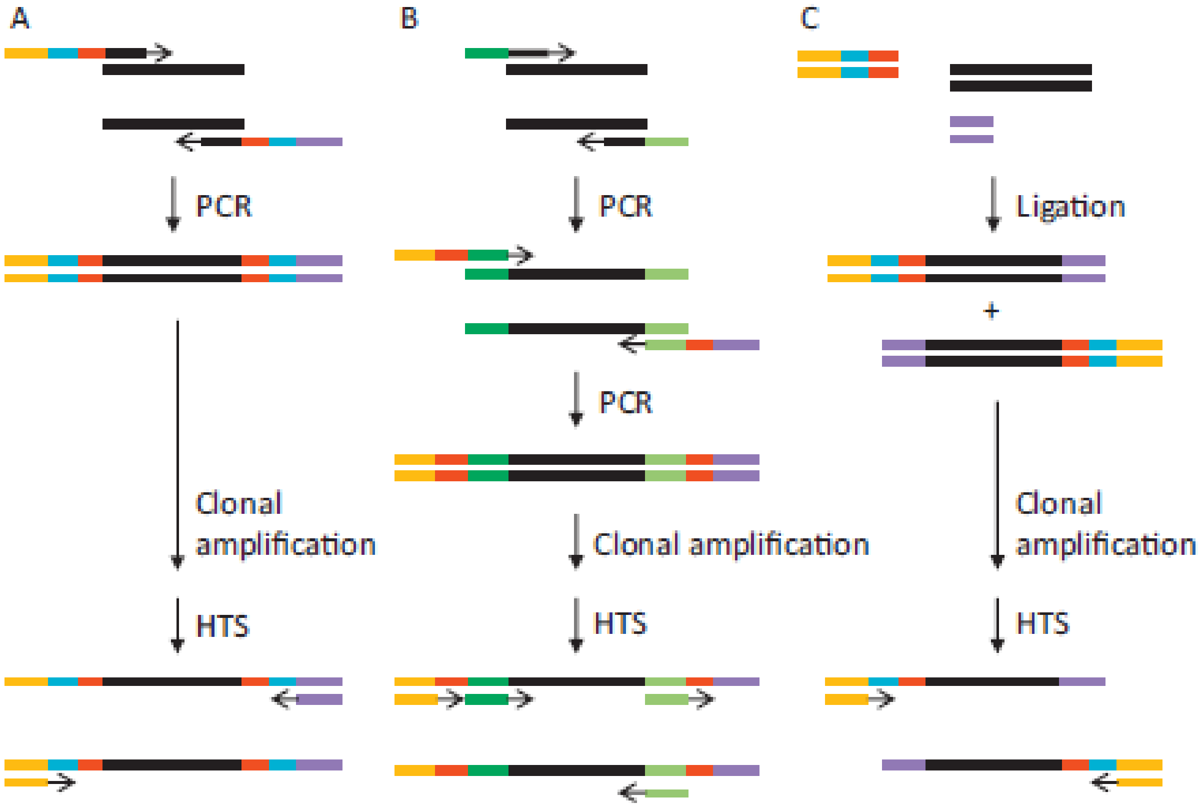
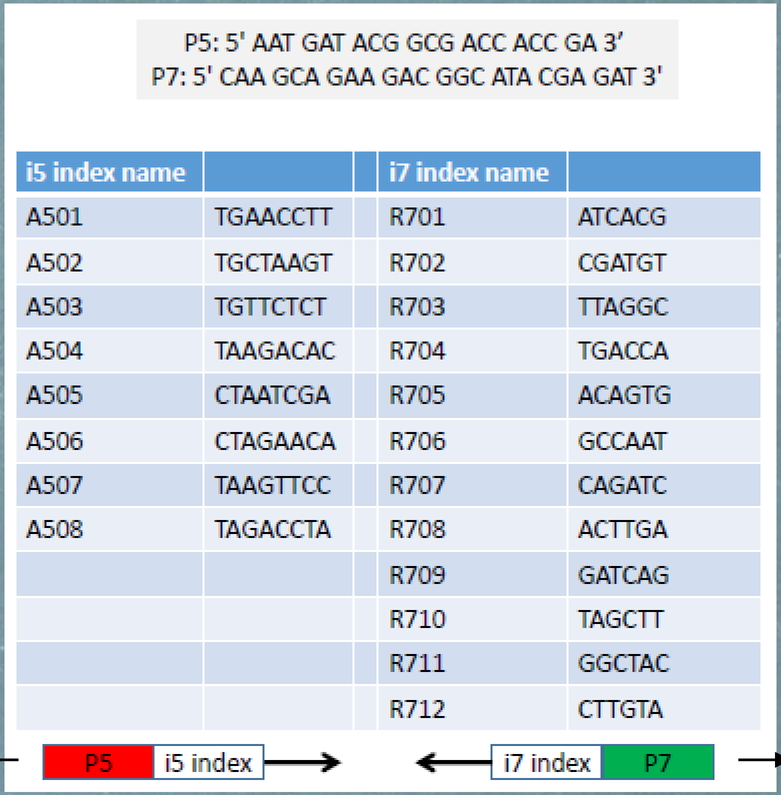


Fig. 3. Examples of library building and sequencing strategies. (A) The library is generated by one PCR reaction. The PCR primers include five elements; the target sequence (in black), the barcode for sample identification (in red), the key sequence for sequence quality control (in blue) and sequencing targets (in orange and purple). One of the sequencing targets is also used to hybridize the library to the solid surface during the clonal amplification step. With two sequencing targets, it is possible to perform directional sequencing of only one strand by choosing a sequencing primer complementary to either the orange or the purple sequencing target. With only one sequencing target (when the orange and the purple sequences are the same), both strands would be sequenced in the NGS reaction. (B) The library is generated by two PCRs. In the first PCR, the primers include the target sequence (in black) and the sequencing targets (in two shades of green). In the second PCR, the primers hybridize to the sequencing targets and include tags with the barcode (in red) and sequences for hybridization to the solid surface used for the clonal amplification. The target sequence (in black) is sequenced via the two sequencing targets (in green) whereas the barcodes are sequenced in separate reactions. (C) The library is generated by ligation of adapters to the fragmented genomic DNA. One adapter includes the barcode for sample identification (in red), the key sequence for sequence quality control (in blue) and the sequencing target (in orange). The second adapter includes the sequence for hybridization to the solid surface used for the clonal amplification. Four different products will be generated by the ligation; the two products shown in the figure, where two different adapters are ligated to the DNA fragment, and two products where the same adapter ligates to both ends. The later products cannot be used in the downstream reactions. Sequencing is conducted from hybridization of a sequencing primer complimentary to the sequencing target (in orange). Both strands will be sequenced because the adapter with the sequencing target ligates to either the forward or the reverse strand in equal numbers. HTS (high throughput sequencing).

Børsting et al Forensic Sci Int Genet 2015



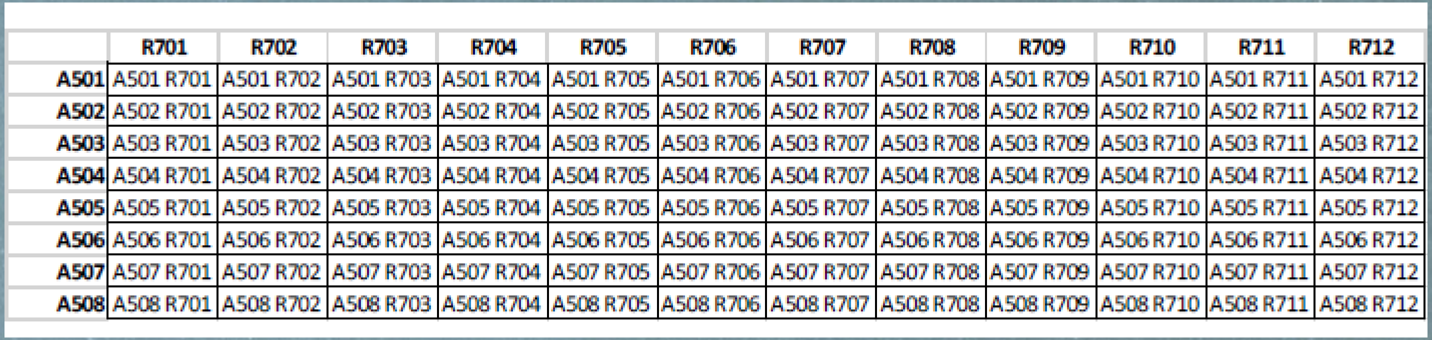
P5: 5' AAT GAT ACG GCG ACC ACC GA 3'
P7: 5' CAA GCA GAA GAC GGC ATA CGA GAT 3'

i5 index name			i7 index name	
A501	TGAACCTT		R701	ATCACG
A502	TGCTAAGT		R702	CGATGT
A503	TGTTCTCT		R703	TTAGGC
A504	TAAGACAC		R704	TGACCA
A505	CTAATCGA		R705	ACAGTG
A506	CTAGAACA		R706	GCCAAT
A507	TAAGTTCC		R707	CAGATC
A508	TAGACCTA		R708	ACTTGA
			R709	GATCAG
			R710	TAGCTT
			R711	GGCTAC
			R712	CTTGTA

P5
i5 index
i7 index
P7

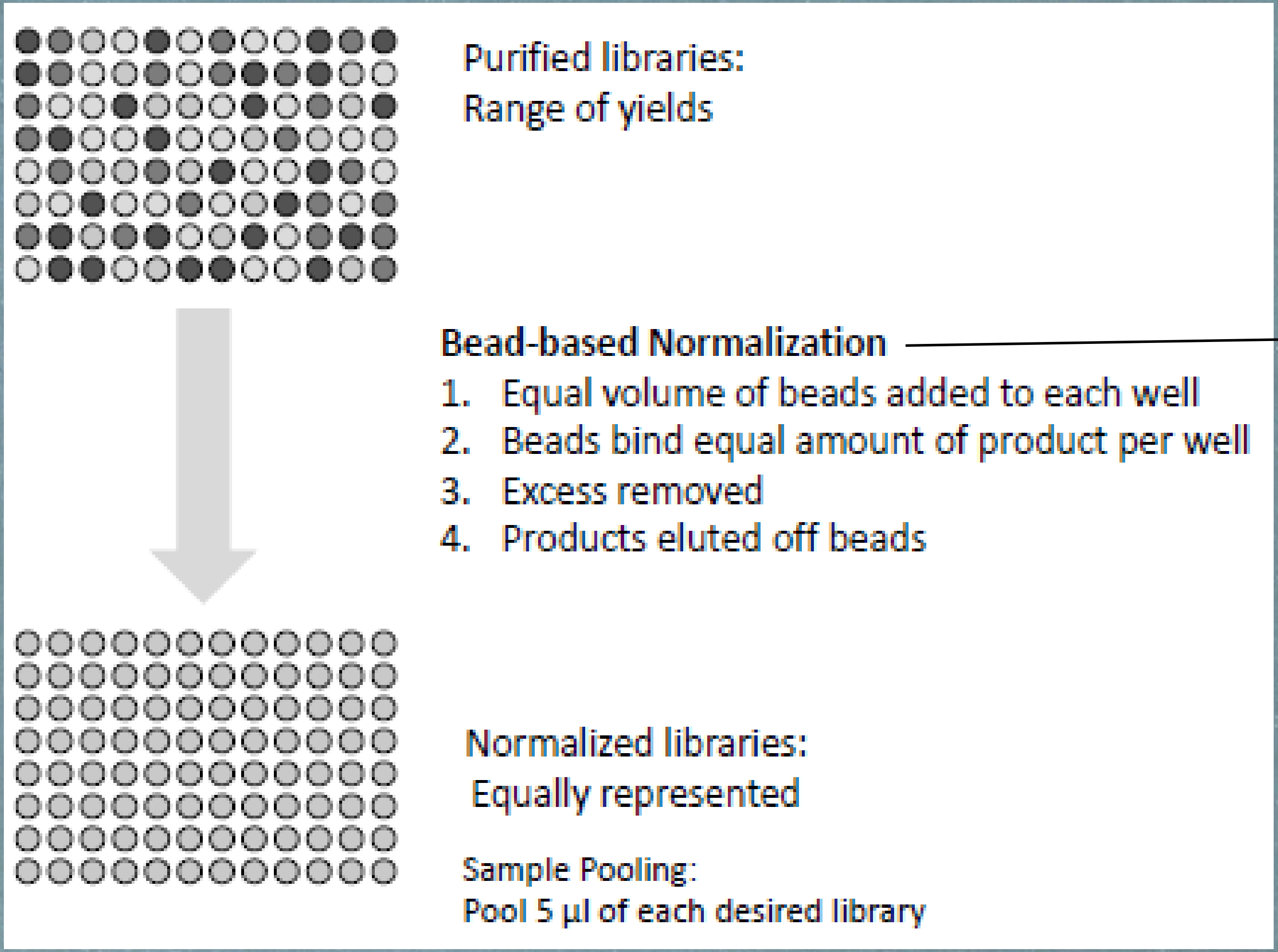
Adapter sequence used to link the DNA fragments on solid surface (flow cell)

Adapter sequence used to link the DNA fragments on solid surface (flow cell)



	R701	R702	R703	R704	R705	R706	R707	R708	R709	R710	R711	R712
A501	A501 R701	A501 R702	A501 R703	A501 R704	A501 R705	A501 R706	A501 R707	A501 R708	A501 R709	A501 R710	A501 R711	A501 R712
A502	A502 R701	A502 R702	A502 R703	A502 R704	A502 R705	A502 R706	A502 R707	A502 R708	A502 R709	A502 R710	A502 R711	A502 R712
A503	A503 R701	A503 R702	A503 R703	A503 R704	A503 R705	A503 R706	A503 R707	A503 R708	A503 R709	A503 R710	A503 R711	A503 R712
A504	A504 R701	A504 R702	A504 R703	A504 R704	A504 R705	A504 R706	A504 R707	A504 R708	A504 R709	A504 R710	A504 R711	A504 R712
A505	A505 R701	A505 R702	A505 R703	A505 R704	A505 R705	A505 R706	A505 R707	A505 R708	A505 R709	A505 R710	A505 R711	A505 R712
A506	A506 R701	A506 R702	A506 R703	A506 R704	A506 R705	A506 R706	A506 R707	A506 R708	A506 R709	A506 R710	A506 R711	A506 R712
A507	A507 R701	A507 R702	A507 R703	A507 R704	A507 R705	A507 R706	A507 R707	A507 R708	A507 R709	A507 R710	A507 R711	A507 R712
A508	A508 R701	A508 R702	A508 R703	A508 R704	A508 R705	A508 R706	A508 R707	A508 R708	A508 R709	A508 R710	A508 R711	A508 R712

- ✓ Libraries are then purified and normalized in order to ensure that libraries of varying yields are equally represented within the sequencing run
- ✓ Once normalized libraries can be pooled together thanks to barcodes that will allow to precisely identify each library/sample in the following sequencing reaction



Purified libraries:
Range of yields

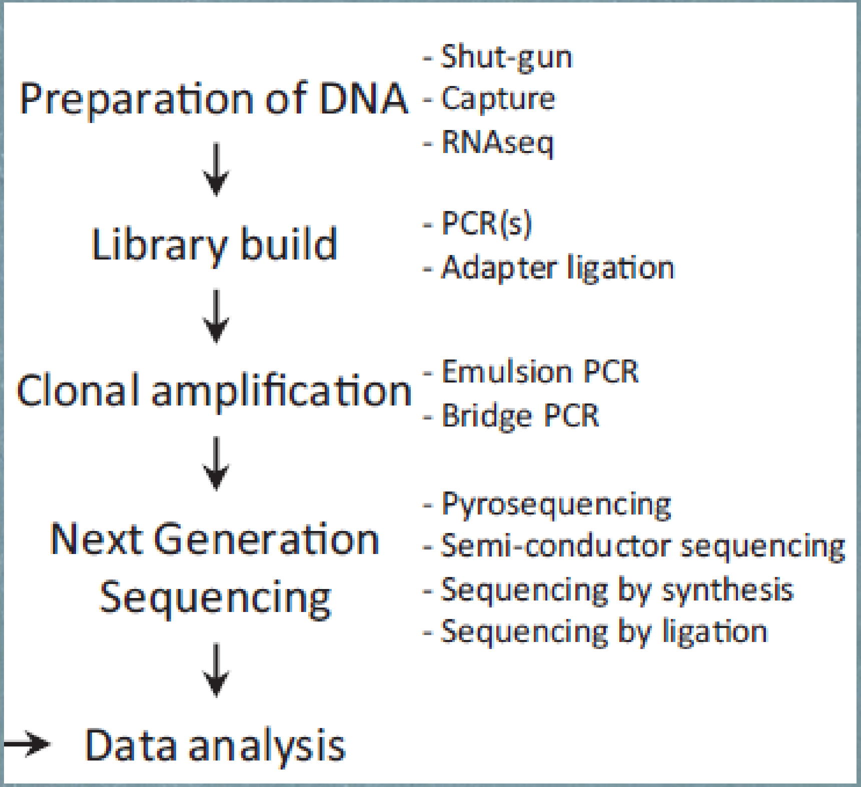
Bead-based Normalization

1. Equal volume of beads added to each well
2. Beads bind equal amount of product per well
3. Excess removed
4. Products eluted off beads

Normalized libraries:
Equally represented

Sample Pooling:
Pool 5 µl of each desired library

Alternatively libraries can be quantified by real-time PCR and proprly diluted so that equal amounts of each library are then pooled



Preparation of DNA
- Shut-gun
- Capture
- RNAseq

↓

Library build
- PCR(s)
- Adapter ligation

↓

Clonal amplification
- Emulsion PCR
- Bridge PCR

↓

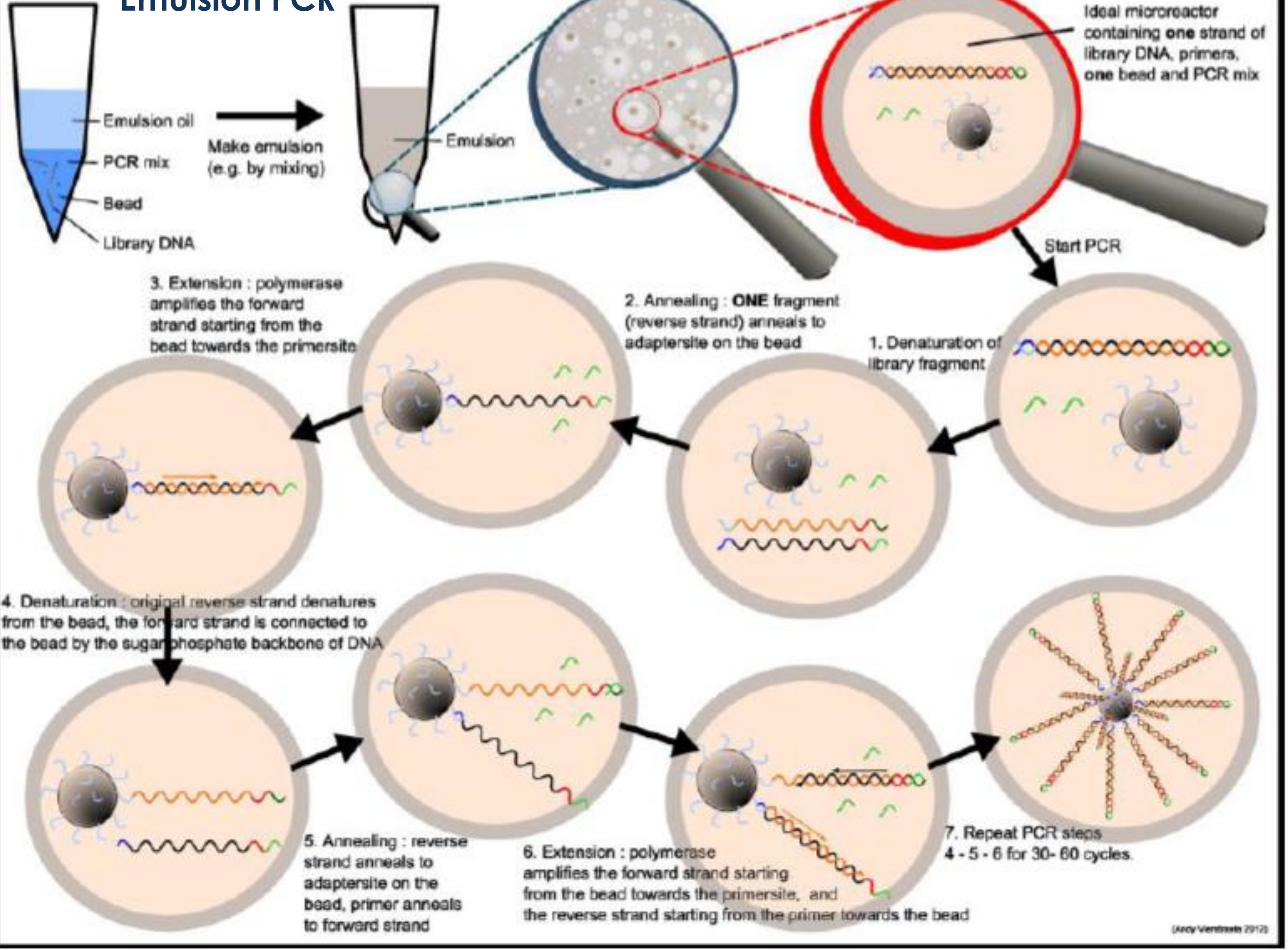
Next Generation Sequencing
- Pyrosequencing
- Semi-conductor sequencing
- Sequencing by synthesis
- Sequencing by ligation

↓

→ Data analysis

Thousands of copies of each original DNA molecule form an immobilized "cluster of DNA" on a bead (emulsion PCR) or a flow cell (bridge PCR)

Emulsion PCR



Ideal microreactor containing **one** strand of library DNA, primers, **one** bead and PCR mix

1. Denaturation of library fragment

2. Annealing : **ONE** fragment (reverse strand) anneals to adaptersite on the bead

3. Extension : polymerase amplifies the forward strand starting from the bead towards the primersite

4. Denaturation : original reverse strand denatures from the bead, the forward strand is connected to the bead by the sugar phosphate backbone of DNA

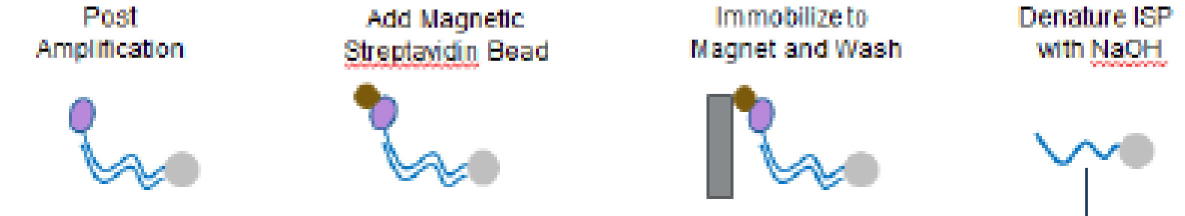
5. Annealing : reverse strand anneals to adaptersite on the bead, primer anneals to forward strand

6. Extension : polymerase amplifies the forward strand starting from the bead towards the primersite, and the reverse strand starting from the primer towards the bead

7. Repeat PCR steps 4 - 5 - 6 for 30- 60 cycles

(Andy Vierstraete 2012)

- Streptavidin beads
- Binds to Adapter X only
  - Template Positive ISPs have Adapter X at the ends

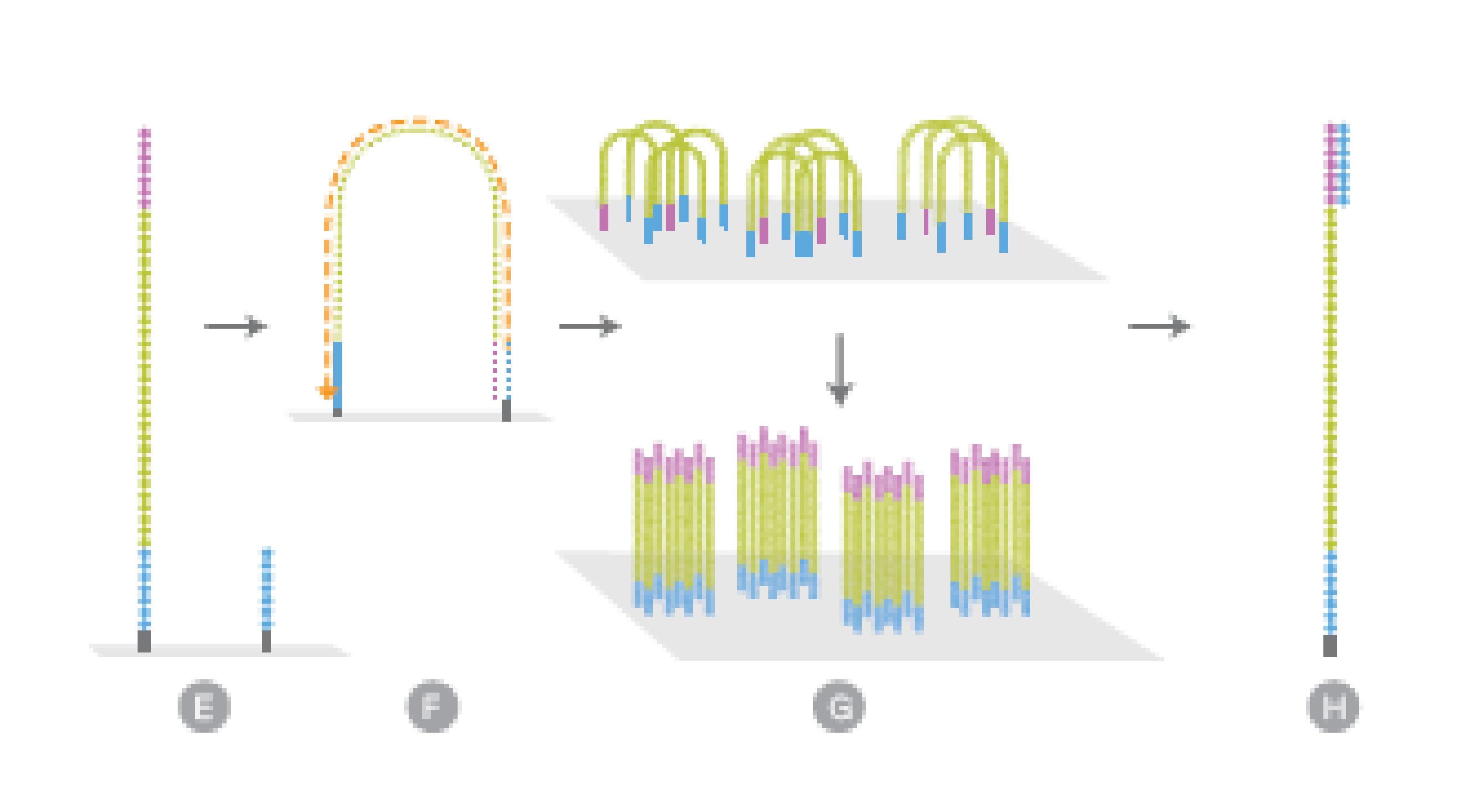


3-6 M reads per chip
Up to 600 bp

15-20 M reads per chip
Up to 600 bp

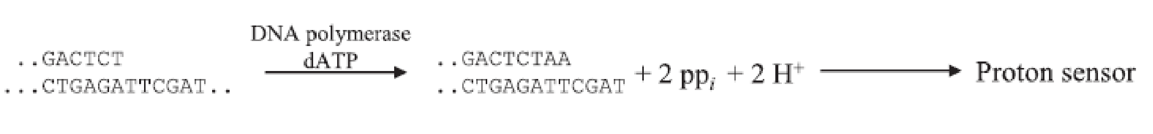
60-80 M reads per chip
Up to 200 bp

Bridge PCR



https://www.youtube.com/watch?annotation_id=annotation_228575861&feature=iv&src_vid=womKfikWlxM&v=fCd6B5HRaZ8

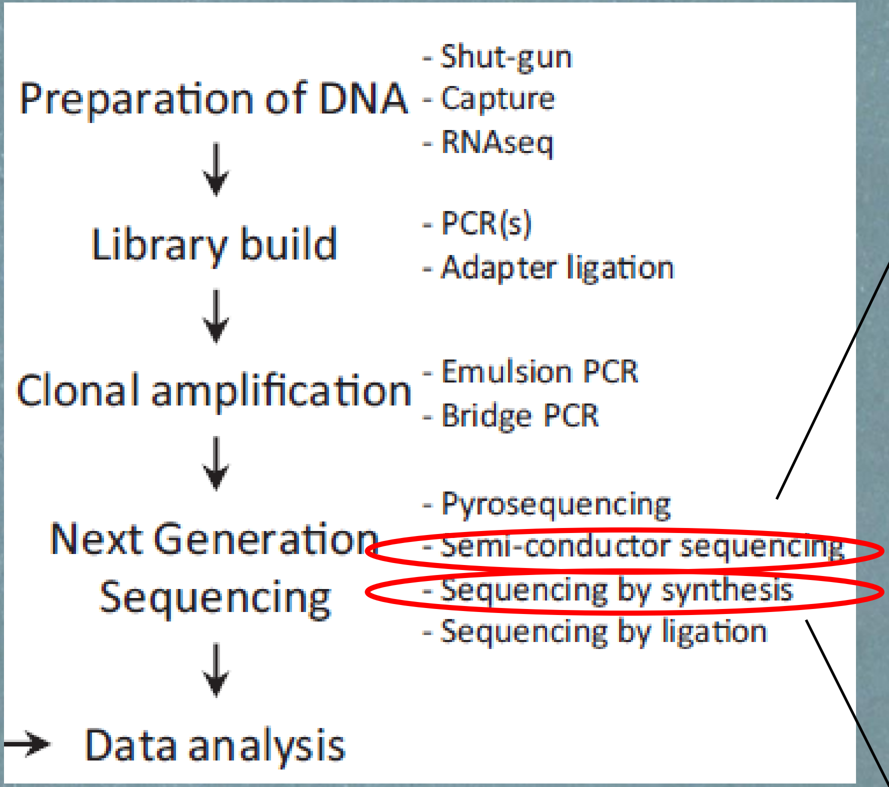
up to min 2:15



Cheap is floaded with one nucleotides type at a time. Incorporation of one or more nucleotide(s) to the growing strand release one or more hydrogen ion(s) that are detected by an ion sensor.

https://www.youtube.com/watch?v=DyijNS0LWBY

from min 1:05 to min 2:42



Preparation of DNA
- Shut-gun
- Capture
- RNAseq

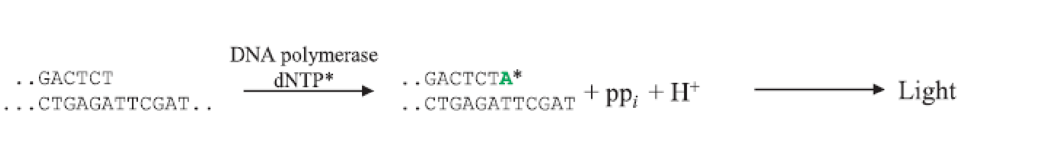
Library build
- PCR(s)
- Adapter ligation

Clonal amplification
- Emulsion PCR
- Bridge PCR

Next Generation Sequencing
- Pyrosequencing
- Semi-conductor sequencing
- Sequencing by synthesis
- Sequencing by ligation

Data analysis

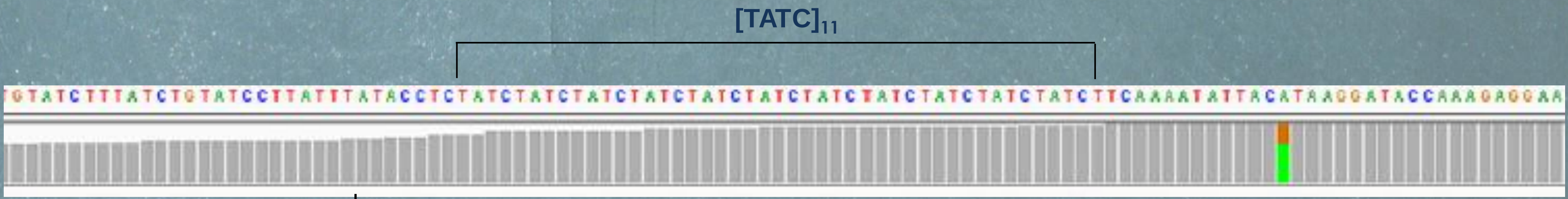
DNA synthesis is performed with fluorescently labeled dNTPs with reversible 3' terminators (marked by an asterisk). Each addition of a nucleotide to the growing strand is detected by a camera. The terminator is chemically removed allowing for the next nucleotide to be incorporated.



From min 2:15 to min 4:42

https://www.youtube.com/watch?annotation_id=annotation_228575861&feature=iv&src_vid=womKfikWlxM&v=fCd6B5HRaZ8

- ✓ The final output will be several copies of each target sequence that will be alligned to a reference sequence or tiled in order to reconstruct a longer target sequence (e.g. whole mtDNA control region)



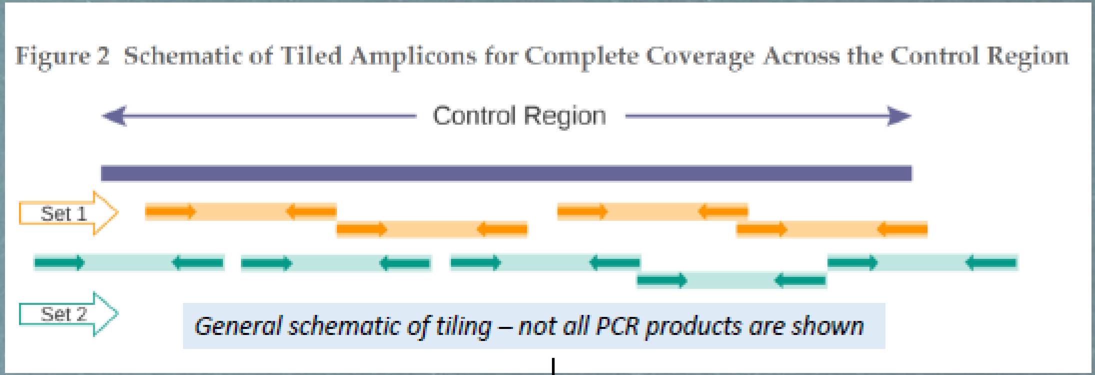
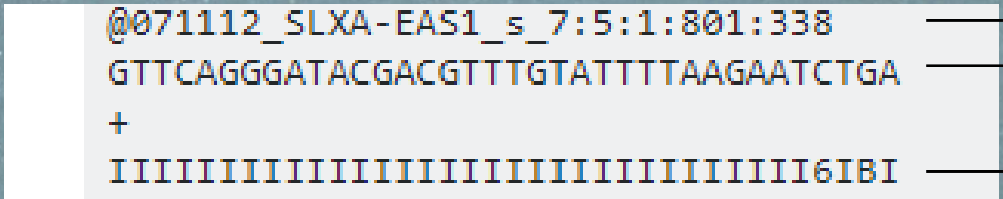


Figure 2 Schematic of Tiled Amplicons for Complete Coverage Across the Control Region

Control Region
Set 1
Set 2
General schematic of tiling – not all PCR products are shown

Short-read data are stored as FASTQ files

BAM file (binary representation of sequence alligning map SAM files) that store information about where and how a sequence maps into the reference.



```
@071112_SLXA-EAS1_s_7:5:1:801:338
GTTCAGGGATACGACGTTTGTATTTTAAGAATCTGA
+
IIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII6IBI
```

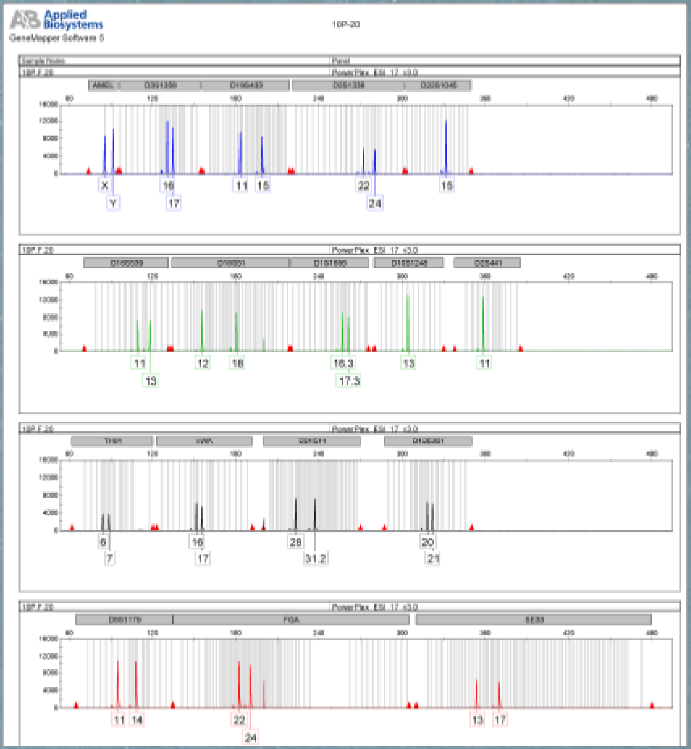
Sequence identifier and optional data

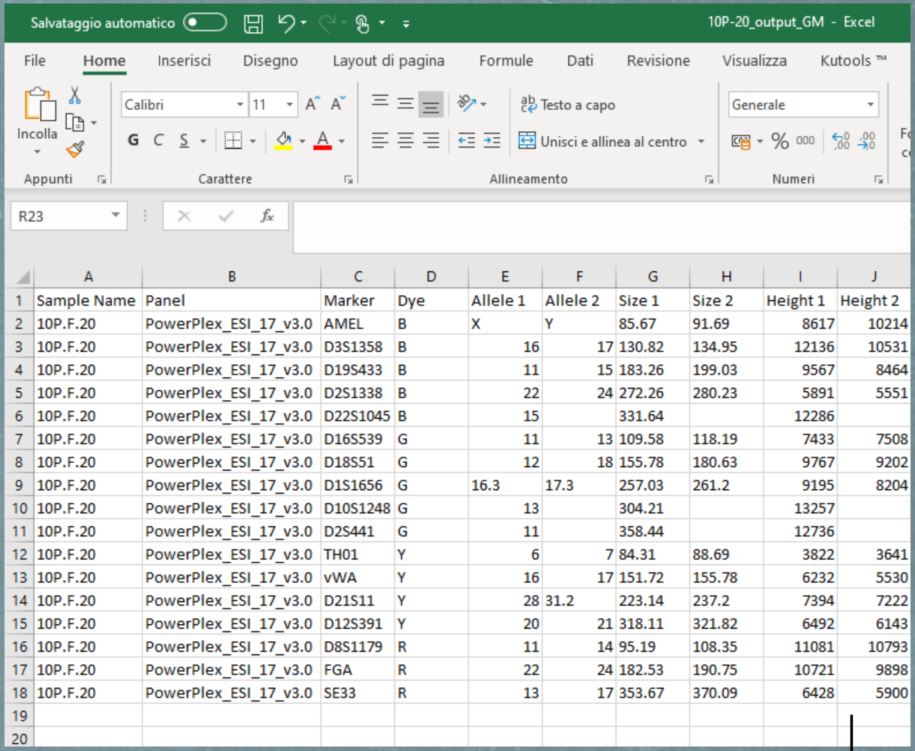
Raw sequence

Quality value of each base from line 2

Capillary electrophoresis vs NGS

- ✓ CE translates machine measured DNA-molecule migration times into DNA fragment lengths which, to further aid interpretation, are visualized in peak profiles and tables with a very simple string of numbers representing these fragment lengths

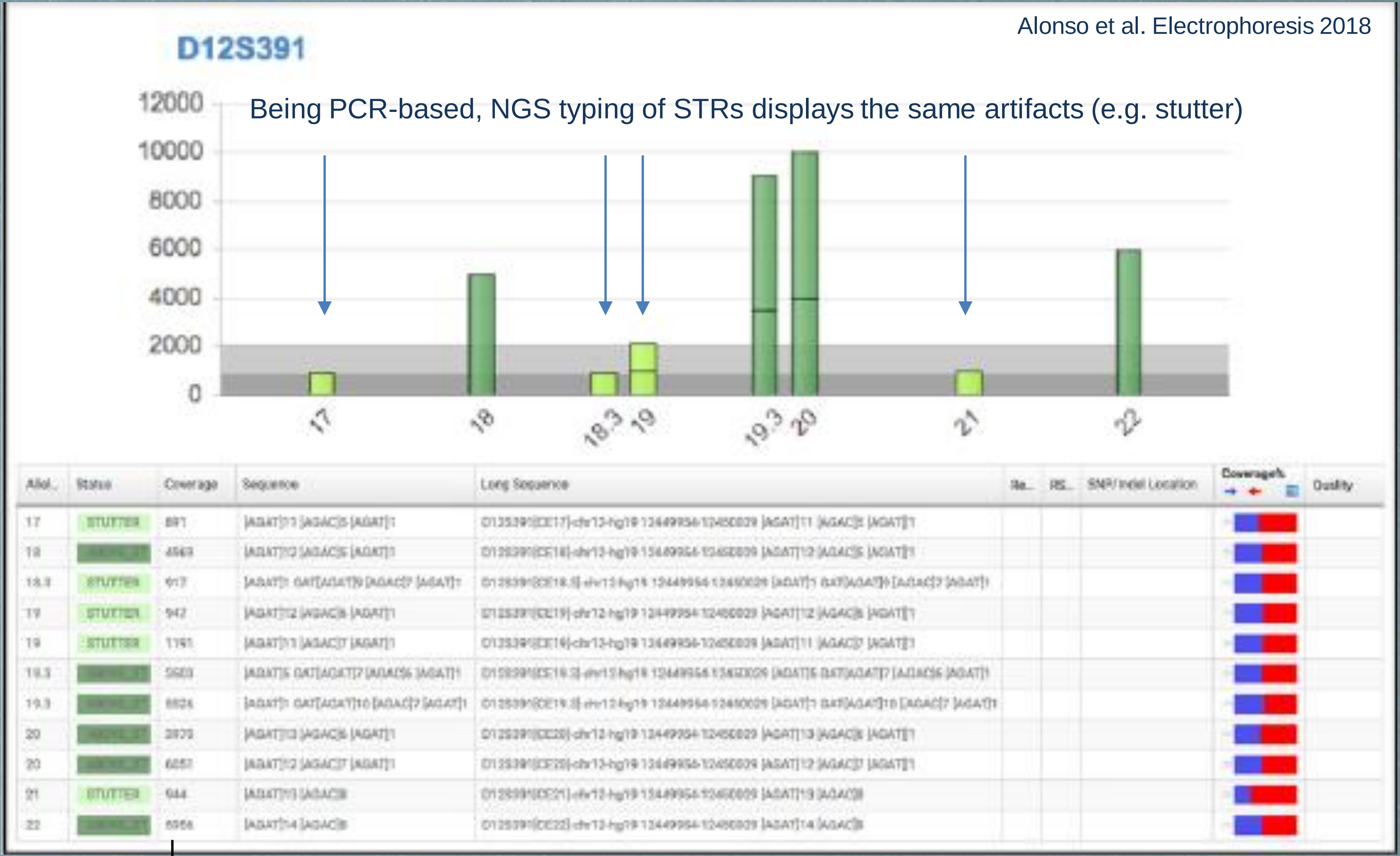




	A	B	C	D	E	F	G	H	I	J
1	Sample Name	Panel	Marker	Dye	Allele 1	Allele 2	Size 1	Size 2	Height 1	Height 2
2	10P.F.20	PowerPlex_ESI_17_v3.0	AMEL	B	X	Y	85.67	91.69	8617	10214
3	10P.F.20	PowerPlex_ESI_17_v3.0	D3S1358	B	16	17	130.82	134.95	12136	10531
4	10P.F.20	PowerPlex_ESI_17_v3.0	D19S433	B	11	15	183.26	199.03	9567	8464
5	10P.F.20	PowerPlex_ESI_17_v3.0	D2S1338	B	22	24	272.26	280.23	5891	5551
6	10P.F.20	PowerPlex_ESI_17_v3.0	D22S1045	B	15		331.64		12286	
7	10P.F.20	PowerPlex_ESI_17_v3.0	D16S539	G	11	13	109.58	118.19	7433	7508
8	10P.F.20	PowerPlex_ESI_17_v3.0	D18S51	G	12	18	155.78	180.63	9767	9202
9	10P.F.20	PowerPlex_ESI_17_v3.0	D1S1656	G	16.3	17.3	257.03	261.2	9195	8204
10	10P.F.20	PowerPlex_ESI_17_v3.0	D10S1248	G	13		304.21		13257	
11	10P.F.20	PowerPlex_ESI_17_v3.0	D2S441	G	11		358.44		12736	
12	10P.F.20	PowerPlex_ESI_17_v3.0	TH01	Y	6	7	84.31	88.69	3822	3641
13	10P.F.20	PowerPlex_ESI_17_v3.0	vWA	Y	16	17	151.72	155.78	6232	5530
14	10P.F.20	PowerPlex_ESI_17_v3.0	D21S11	Y	28	31.2	223.14	237.2	7394	7222
15	10P.F.20	PowerPlex_ESI_17_v3.0	D12S391	Y	20	21	318.11	321.82	6492	6143
16	10P.F.20	PowerPlex_ESI_17_v3.0	D8S1179	R	11	14	95.19	108.35	11081	10793
17	10P.F.20	PowerPlex_ESI_17_v3.0	FGA	R	22	24	182.53	190.75	10721	9898
18	10P.F.20	PowerPlex_ESI_17_v3.0	SE33	R	13	17	353.67	370.09	6428	5900
19										
20										

Further information can derive from peak height, measured in relative fluorescence units (rfu), that can be used in the interpretation of complex DNA profiles affected by artifacts (stutter, drop-out, ...) and DNA mixtures

- ✓ With MPS, irrespective of the underlying sequence technology, the final experimental result is represented as a long list of DNA sequence reads that reveals all underlying sequence variation in the targeted DNA sample.

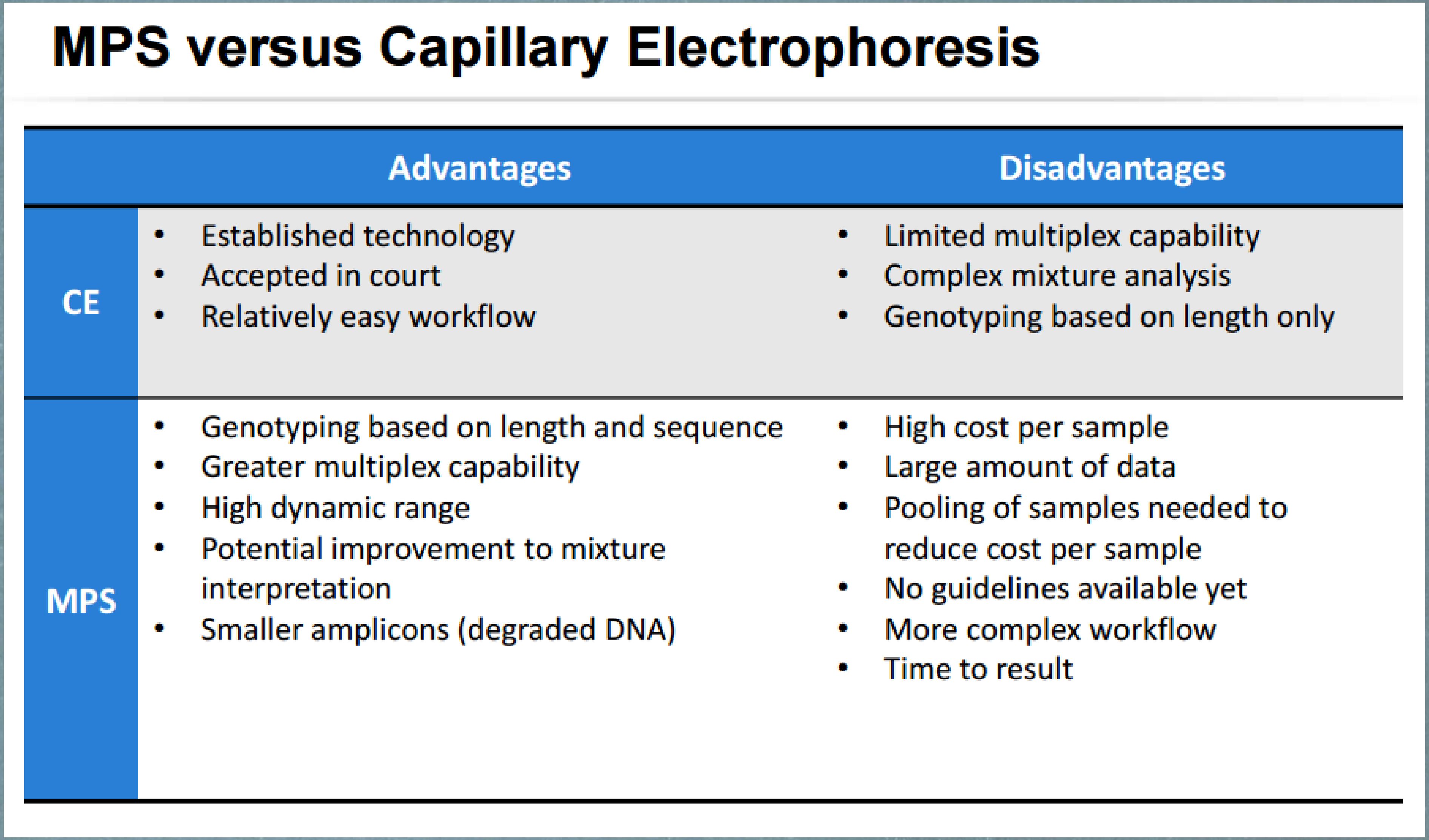


Alonso et al. Electrophoresis 2018

Being PCR-based, NGS typing of STRs displays the same artifacts (e.g. stutter)

Allel..	Status	Coverage	Sequence	Long Sequence	Re..	RS..	SNP/Indel Location	Coverage%	Quality
17	STUTTER	891	[AGAT]11 [AGAC]5 [AGAT]1	D12S391[CE17]-chr12-hg19 12449954-12450029 [AGAT]11 [AGAC]5 [AGAT]1				🟦🟥	
18	[illegible]	4969	[AGAT]12 [AGAC]5 [AGAT]1	D12S391[CE18]-chr12-hg19 12449954-12450029 [AGAT]12 [AGAC]5 [AGAT]1				🟦🟥	
18.3	STUTTER	917	[AGAT]1 GAT[AGAT]9 [AGAC]7 [AGAT]1	D12S391[CE18.3]-chr12-hg19 12449954-12450029 [AGAT]1 GAT[AGAT]9 [AGAC]7 [AGAT]1				🟦🟥	
19	STUTTER	947	[AGAT]12 [AGAC]6 [AGAT]1	D12S391[CE19]-chr12-hg19 12449954-12450029 [AGAT]12 [AGAC]6 [AGAT]1				🟦🟥	
19	STUTTER	1191	[AGAT]11 [AGAC]7 [AGAT]1	D12S391[CE19]-chr12-hg19 12449954-12450029 [AGAT]11 [AGAC]7 [AGAT]1				🟦🟥	
19.3	[illegible]	5603	[AGAT]5 GAT[AGAT]7 [AGAC]6 [AGAT]1	D12S391[CE19.3]-chr12-hg19 12449954-12450029 [AGAT]5 GAT[AGAT]7 [AGAC]6 [AGAT]1				🟦🟥	
19.3	[illegible]	6826	[AGAT]1 GAT[AGAT]10 [AGAC]7 [AGAT]1	D12S391[CE19.3]-chr12-hg19 12449954-12450029 [AGAT]1 GAT[AGAT]10 [AGAC]7 [AGAT]1				🟦🟥	
20	[illegible]	3979	[AGAT]13 [AGAC]6 [AGAT]1	D12S391[CE20]-chr12-hg19 12449954-12450029 [AGAT]13 [AGAC]6 [AGAT]1				🟦🟥	
20	[illegible]	6051	[AGAT]12 [AGAC]7 [AGAT]1	D12S391[CE20]-chr12-hg19 12449954-12450029 [AGAT]12 [AGAC]7 [AGAT]1				🟦🟥	
21	STUTTER	944	[AGAT]13 [AGAC]8	D12S391[CE21]-chr12-hg19 12449954-12450029 [AGAT]13 [AGAC]8				🟦🟥	
22	[illegible]	5956	[AGAT]14 [AGAC]8	D12S391[CE22]-chr12-hg19 12449954-12450029 [AGAT]14 [AGAC]8				🟦🟥	

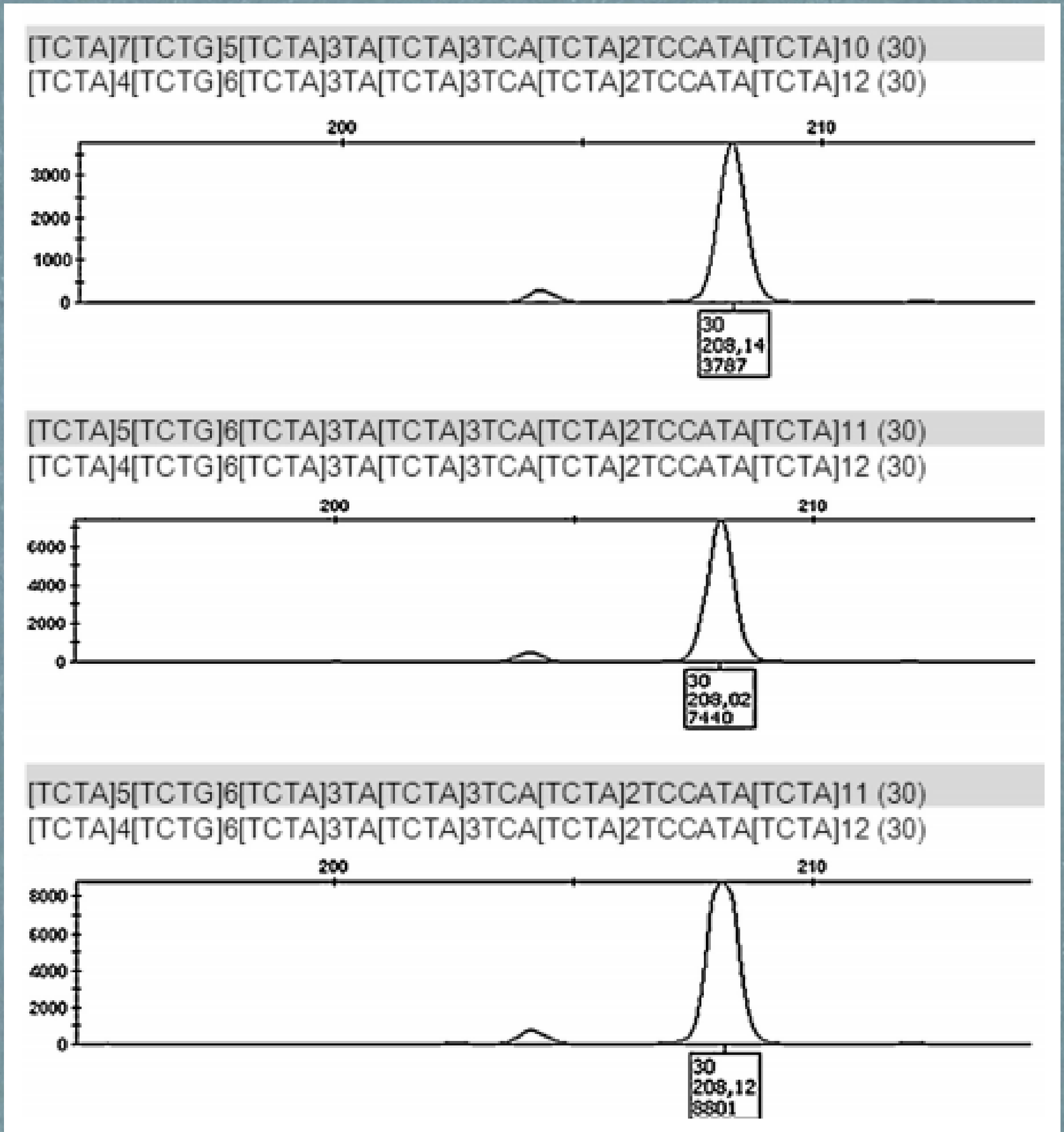
Depth of coverage or read depth (the number of times a specific target molecule is sequenced by the NGS system) can be used similarly to rfus to graphically represent in bars of different height alleles detected in a sample and combine quantitative information in data interpretation



MPS versus Capillary Electrophoresis

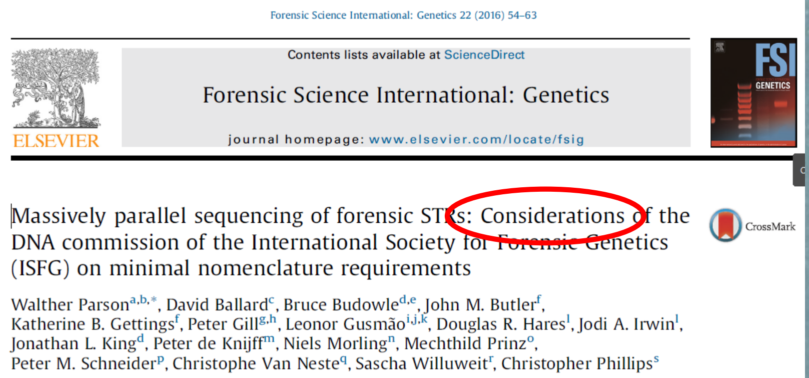
	Advantages	Disadvantages
CE	• Established technology • Accepted in court • Relatively easy workflow	• Limited multiplex capability • Complex mixture analysis • Genotyping based on length only
MPS	• Genotyping based on length and sequence • Greater multiplex capability • High dynamic range • Potential improvement to mixture interpretation • Smaller amplicons (degraded DNA)	• High cost per sample • Large amount of data • Pooling of samples needed to reduce cost per sample • No guidelines available yet • More complex workflow • Time to result

Nomenclature



A STR allele displaying the same lenght in bp in capillary electrophoresis actually consists of several sequence variants detected by NGS.

- How should we label these STR allele variants?
- How do we guarantee backward compatibility to CE generated data?



Forensic Science International: Genetics 22 (2016) 54–63

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

ELSEVIER

FSI GENETICS

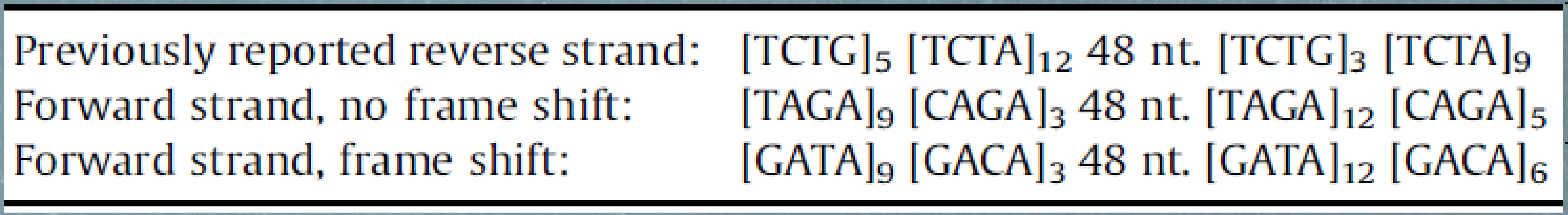
Massively parallel sequencing of forensic STRs: Considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements

CrossMark

Walther Parson[a,b,*], David Ballard[c], Bruce Budowle[d,e], John M. Butler[f], Katherine B. Gettings[f], Peter Gill[g,h], Leonor Gusmão[i,j,k], Douglas R. Hares[l], Jodi A. Irwin[l], Jonathan L. King[d], Peter de Knijff[m], Niels Morling[n], Mechthild Prinz[o], Peter M. Schneider[p], Christophe Van Neste[q], Sascha Willuweit[r], Christopher Phillips[s]

https://www.isfg.org/Publication;Parson2016

- ✓ At the time of writing, GRCh38 is the most up-to-date sequence assembly and is recommended as the framework with which to define repeat region structure for sequence alignment and for the mapping of sequence features such as SNPs.
- ✓ The forward strand direction (from 5' p-arm to 3' q-arm) assigned in the human genome has been constant for all assemblies published since the first draft in 2001 and can be used to align STR sequences.
- ✓ Out of 58 STR loci for which MPS designs have become available at the time of this writing, 23 have been designated historically on the reverse strand. Change to the forward strand for repeat region designation results in a potential shift of the reading frame, that can cause inconsistencies in allelic designation (if we respect former ISFG recommendation that the first 5'-nucleotides that can define a repeat motif should be used)



DYS389	Previously reported reverse strand:	[TCTG]$_5$ [TCTA]$_{12}$ 48 nt. [TCTG]$_3$ [TCTA]$_9$	29
	Forward strand, no frame shift:	[TAGA]$_9$ [CAGA]$_3$ 48 nt. [TAGA]$_{12}$ [CAGA]$_5$	
	Forward strand, frame shift:	[GATA]$_9$ [GACA]$_3$ 48 nt. [GATA]$_{12}$ [GACA]$_6$	30

Forensic Science International: Genetics 22 (2016) 54–63

Contents lists available at ScienceDirect

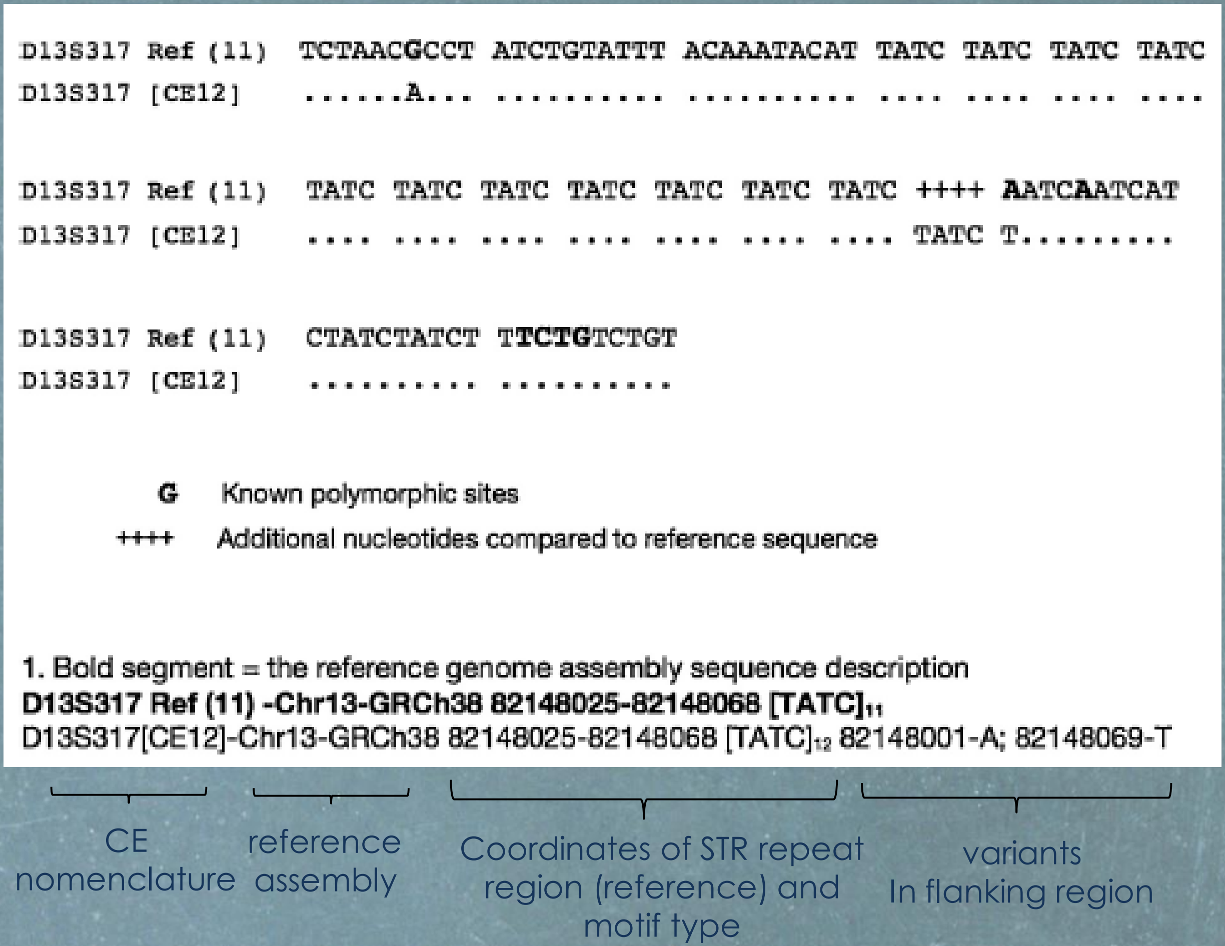
Forensic Science International: Genetics

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

Massively parallel sequencing of forensic STRs: Considerations of the DNA commission of the International Society for Forensic Genetics (ISFG) on minimal nomenclature requirements

Walther Parson[a,b,*], David Ballard[c], Bruce Budowle[d,e], John M. Butler[f], Katherine B. Gettings[f], Peter Gill[g,h], Leonor Gusmão[i,j,k], Douglas R. Hares[l], Jodi A. Irwin[l], Jonathan L. King[d], Peter de Knijff[m], Niels Morling[n], Mechthild Prinz[o], Peter M. Schneider[p], Christophe Van Neste[q], Sascha Willuweit[r], Christopher Phillips[s]

- ✓ Although simple STR nomenclature systems may be required at some point in the future to facilitate communication and data exchange, **comprehensive STR nomenclature** systems are preferred



```
D13S317 Ref (11)  TCTAACGCCT ATCTGTATTT ACAAATACAT TATC TATC TATC TATC
D13S317 [CE12]    ......A... .......... .......... .... .... .... ....

D13S317 Ref (11)  TATC TATC TATC TATC TATC TATC TATC ++++ AATCAATCAT
D13S317 [CE12]    .... .... .... .... .... .... .... TATC T.........

D13S317 Ref (11)  CTATCTATCT TTCTGTCTGT
D13S317 [CE12]    .......... ..........
```

G Known polymorphic sites

++++ Additional nucleotides compared to reference sequence

1. Bold segment = the reference genome assembly sequence description

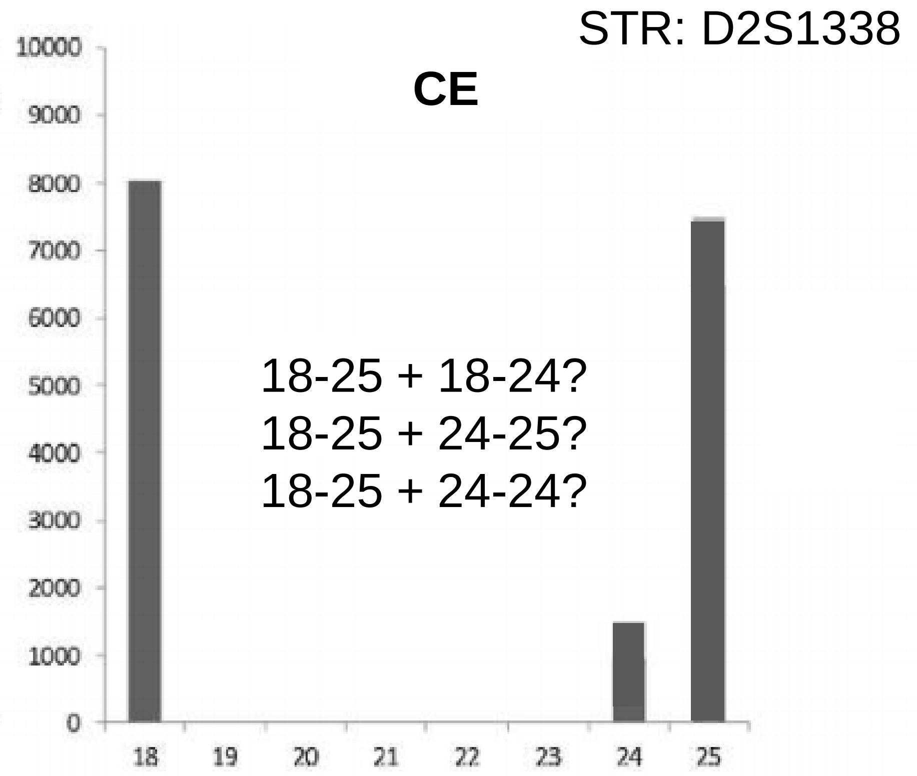
D13S317 Ref (11) -Chr13-GRCh38 82148025-82148068 [TATC]$_{11}$

D13S317[CE12]-Chr13-GRCh38 82148025-82148068 [TATC]$_{12}$ 82148001-A; 82148069-T

CE nomenclature
reference assembly
Coordinates of STR repeat region (reference) and motif type
variants In flanking region

Mixture interpretation

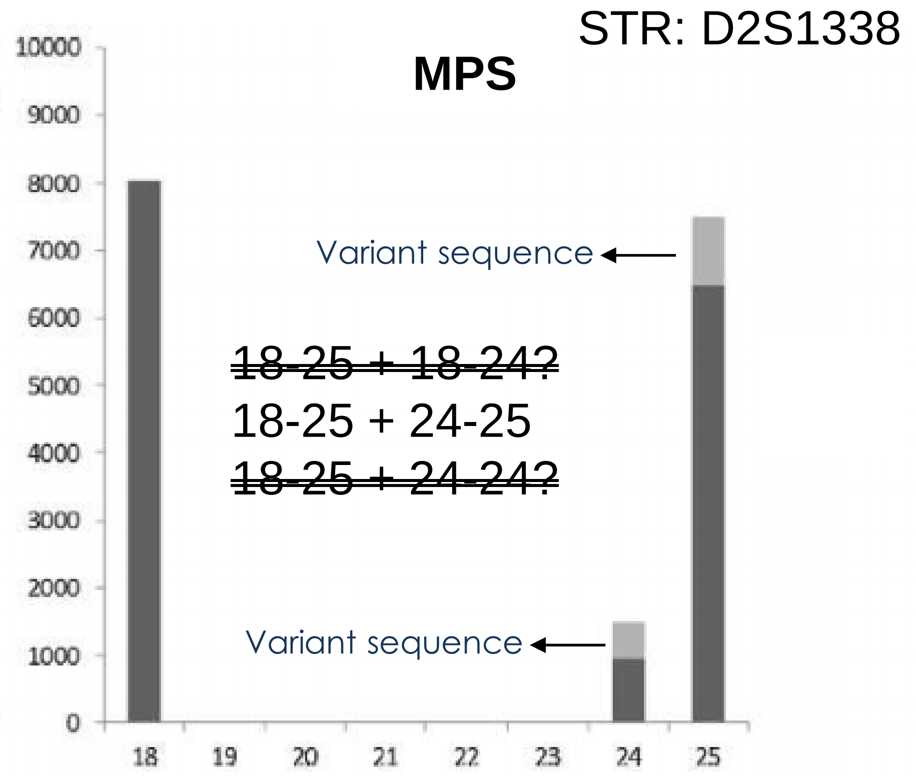
- ✓ Sequence information can often be an advantage e.g. in DNA mixtures interpretation



STR: D2S1338

CE

18-25 + 18-24?
18-25 + 24-25?
18-25 + 24-24?



STR: D2S1338

MPS

Variant sequence

~~18-25 + 18-24?~~
18-25 + 24-25
~~18-25 + 24-24?~~

Variant sequence

Multiplexing capabilities

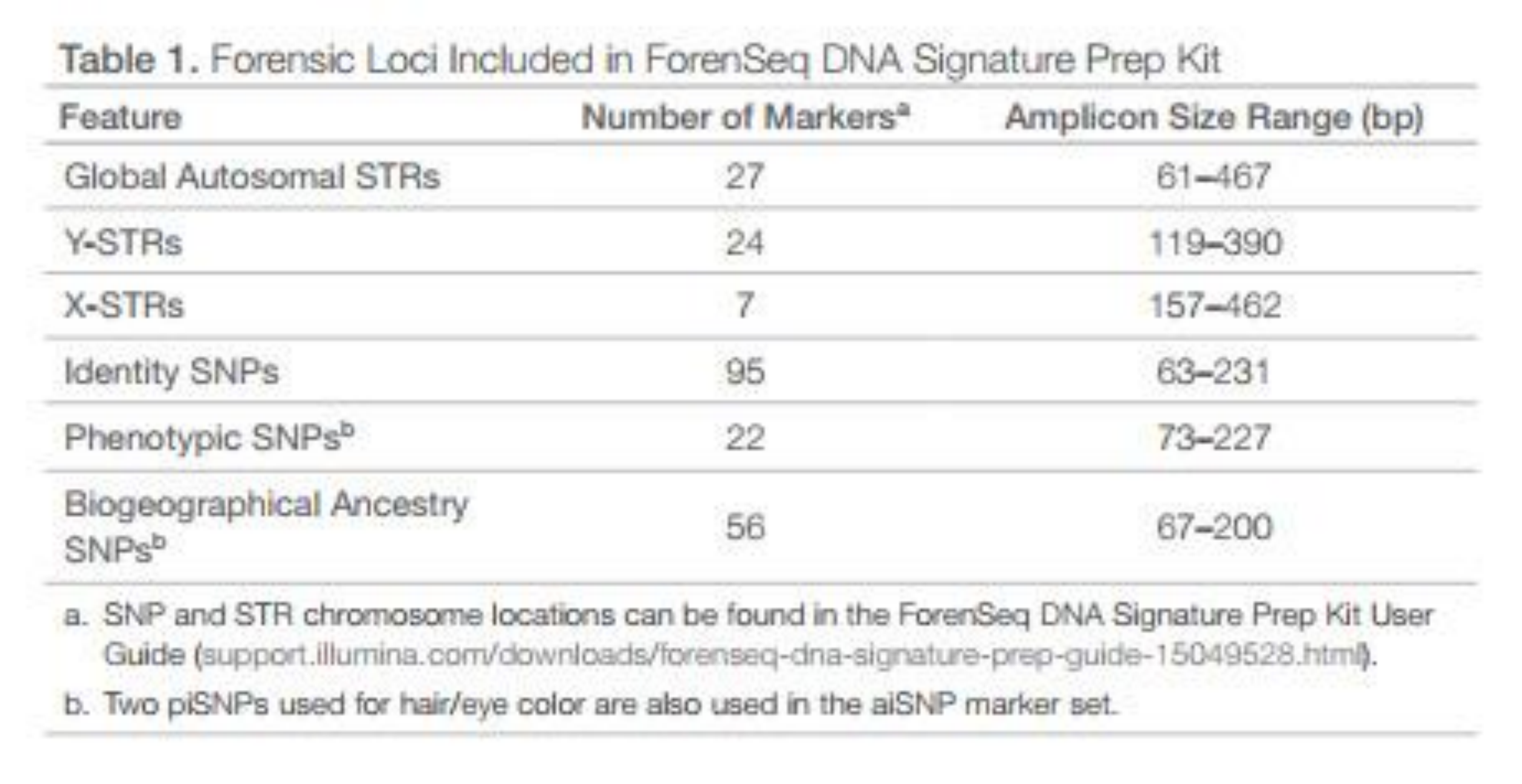


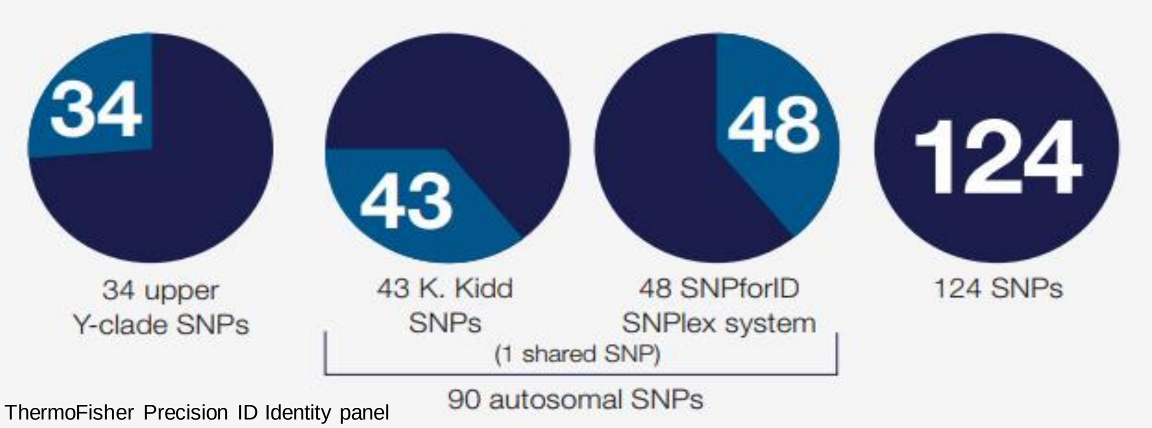
Table 1. Forensic Loci Included in ForenSeq DNA Signature Prep Kit

Feature	Number of Markers[a]	Amplicon Size Range (bp)
Global Autosomal STRs	27	61–467
Y-STRs	24	119–390
X-STRs	7	157–462
Identity SNPs	95	63–231
Phenotypic SNPs[b]	22	73–227
Biogeographical Ancestry SNPs[b]	56	67–200

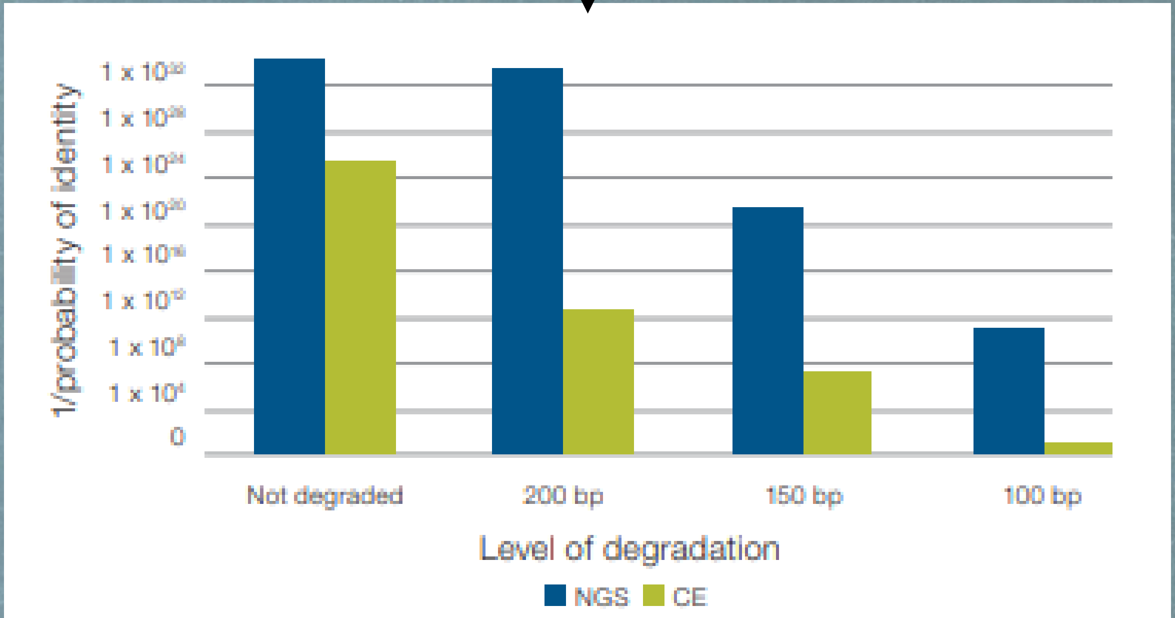
a. SNP and STR chromosome locations can be found in the ForenSeq DNA Signature Prep Kit User Guide (support.illumina.com/downloads/forenseq-dna-signature-prep-guide-15049528.html).

b. Two piSNPs used for hair/eye color are also used in the aiSNP marker set.

Multiplexing capabilities



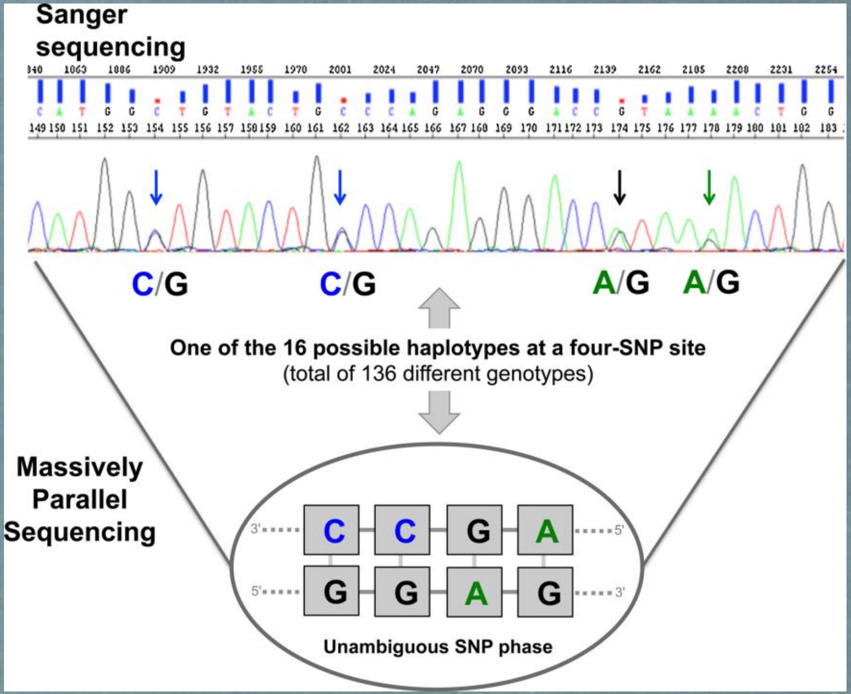
ThermoFisher Precision ID Identity panel

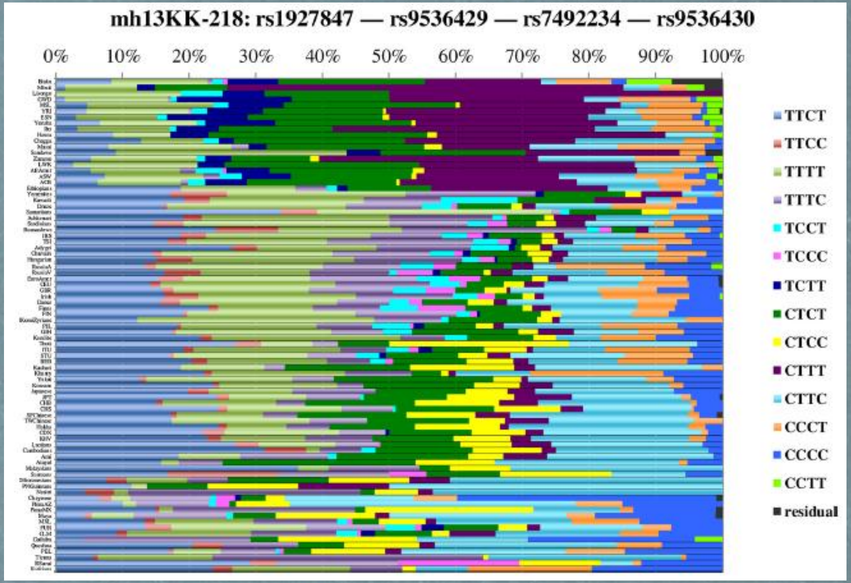


Beyond SNPs: microhaplotypes

Microhaplotypes (MH) are short segments of DNA < 300 bps characterized by the presence of two or more closely linked SNPs.

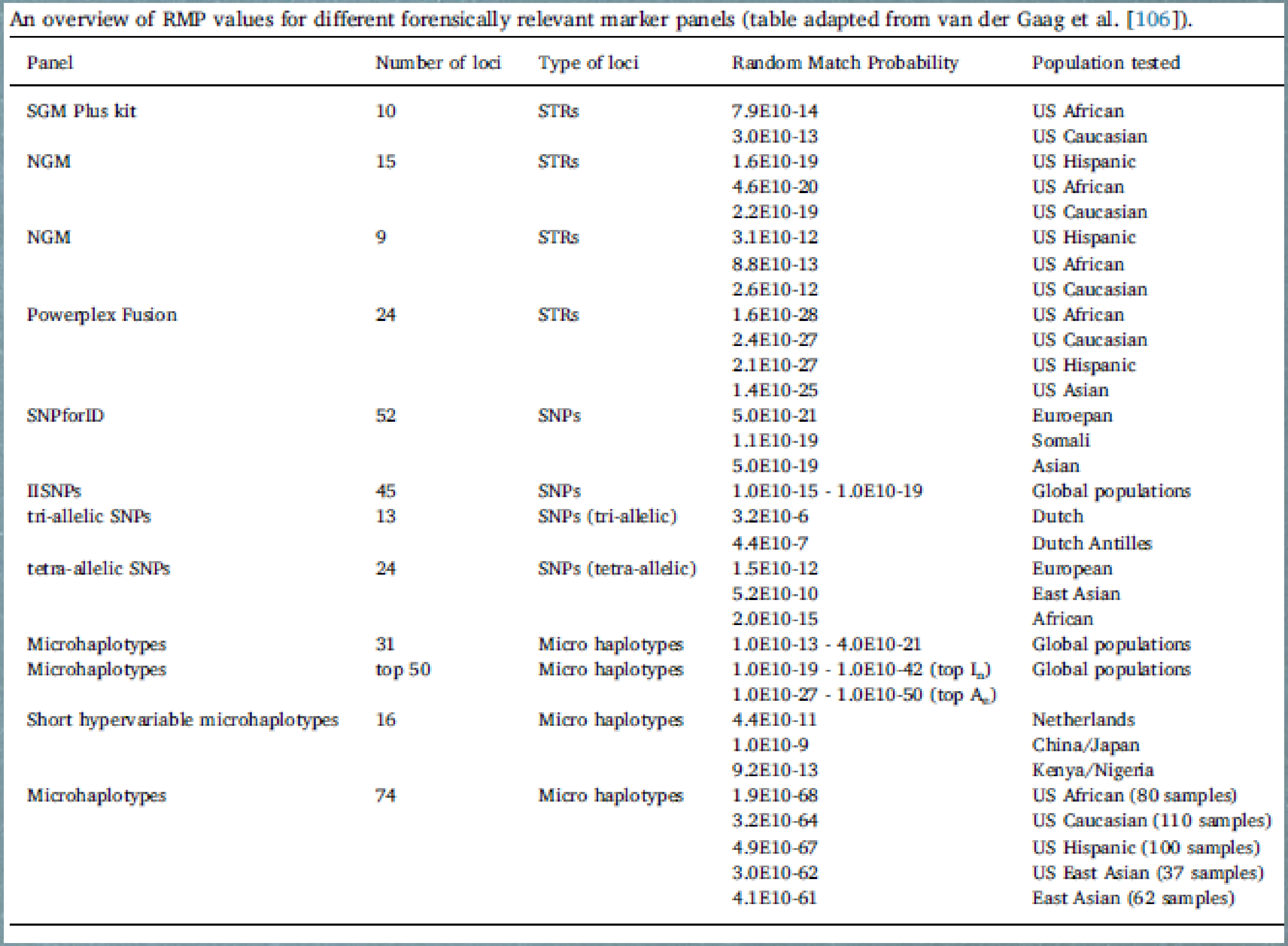
- NGS technology, through clonal amplification and sequencing of each target DNA strand separately, allows precise identification of the combination of alleles on a chromosome (haplotype). Phasing is otherwise impossible in standard Sanger sequencing
- Each MH can be considered as a multiallelic marker


Sanger sequencing
C/G
C/G
A/G
A/G
One of the 16 possible haplotypes at a four-SNP site
(total of 136 different genotypes)
Massively Parallel Sequencing
3'
C
C
G
A
5'
5'
G
G
A
G
3'
Unambiguous SNP phase



Oldoni et al. FSIG 2019

- Combinations of MH have equivalent/higher identificative power compared to available STR/SNP panels

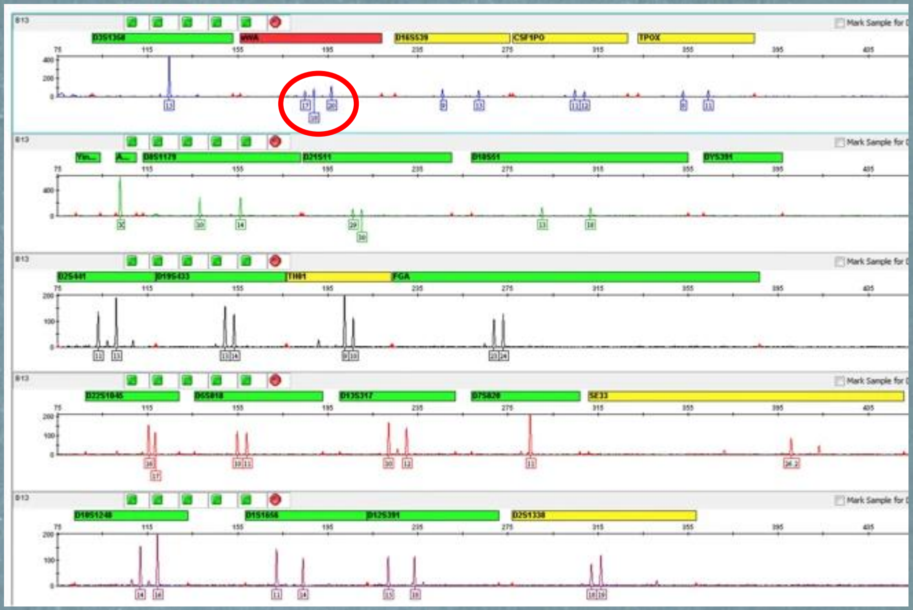


An overview of RMP values for different forensically relevant marker panels (table adapted from van der Gaag et al. [106]).

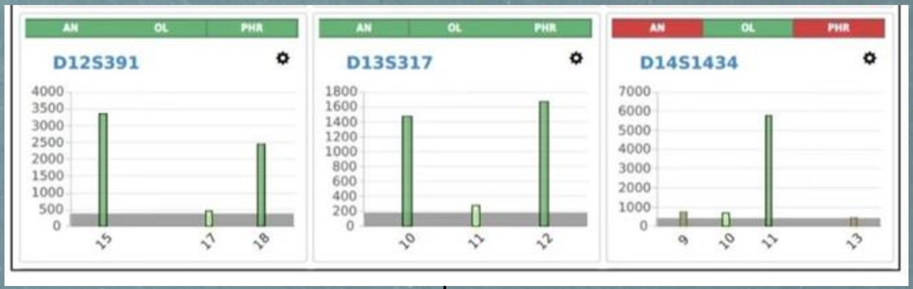
Panel	Number of loci	Type of loci	Random Match Probability	Population tested
SGM Plus kit	10	STRs	7.9E10-14	US African
			3.0E10-13	US Caucasian
NGM	15	STRs	1.6E10-19	US Hispanic
			4.6E10-20	US African
			2.2E10-19	US Caucasian
NGM	9	STRs	3.1E10-12	US Hispanic
			8.8E10-13	US African
			2.6E10-12	US Caucasian
Powerplex Fusion	24	STRs	1.6E10-28	US African
			2.4E10-27	US Caucasian
			2.1E10-27	US Hispanic
			1.4E10-25	US Asian
SNPforID	52	SNPs	5.0E10-21	Euroepan
			1.1E10-19	Somali
			5.0E10-19	Asian
IISNPs	45	SNPs	1.0E10-15 - 1.0E10-19	Global populations
tri-allelic SNPs	13	SNPs (tri-allelic)	3.2E10-6	Dutch
			4.4E10-7	Dutch Antilles
tetra-allelic SNPs	24	SNPs (tetra-allelic)	1.5E10-12	European
			5.2E10-10	East Asian
			2.0E10-15	African
Microhaplotypes	31	Micro haplotypes	1.0E10-13 - 4.0E10-21	Global populations
Microhaplotypes	top 50	Micro haplotypes	1.0E10-19 - 1.0E10-42 (top L_n)	Global populations
			1.0E10-27 - 1.0E10-50 (top A_e)	
Short hypervariable microhaplotypes	16	Micro haplotypes	4.4E10-11	Netherlands
			1.0E10-9	China/Japan
			9.2E10-13	Kenya/Nigeria
Microhaplotypes	74	Micro haplotypes	1.9E10-68	US African (80 samples)
			3.2E10-64	US Caucasian (110 samples)
			4.9E10-67	US Hispanic (100 samples)
			3.0E10-62	US East Asian (37 samples)
			4.1E10-61	East Asian (62 samples)

Oldoni et al. FSIG 2019

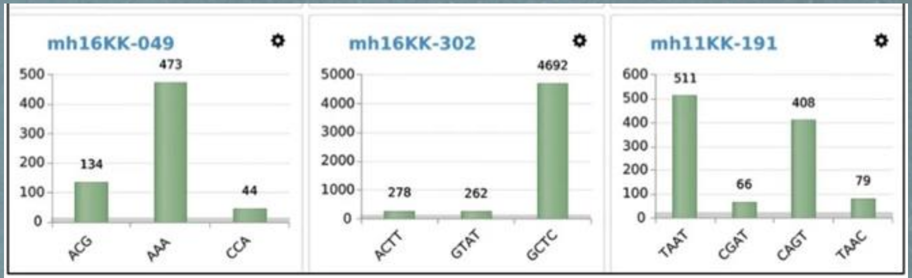
- MH, being unaffected by PCR artifacts typical of STRs can help the identification and interpretation of mixtures (example from Bennett et al IJLM 2019)



CE profile from forensic stain (cigarette butt): mixture suspected because of allelic imbalance at some loci plus third allele (in stutter position) at a single STR locus



Several minor alleles were detected by MPS-STR typing of the same sample, though most of them at stutter position



MPS-MH typing of the same sample, unambiguously identified several tri- and tetrallelic genotypes confirming a second minor contributor