

## INTRODUCING MPS TO A FORENSIC WORKFLOW – WHERE TO START?

Dr Dadna Hartman | Manager Molecular Biology

Warning: contains images of deceased persons

HIDS, June 2019

### **VICTORIA - AUSTRALIA**



- Coroners Court of Victoria (CCoV) is required to investigate any reportable death
  - Annually, over 6,500 cases are admitted as coronial cases
- □ Victorian Institute of Forensic Medicine (VIFM)
  - Independent forensic medical services to the Victorian public and justice system
- Identification must be established to enable the legal interment of the deceased
  - Most cases are identifiable using visual inspections
  - ~ 10 % of reportable cases required ID by scientific means
    - DNA
    - Fingerprints
    - Dental





### **MOLECULAR BIOLOGY LABORATORY (MBL)**

#### Coronial identification

- Nuclear DNA (nDNA) 24 markers (including sex determination markers)
- Mitochondrial DNA (mtDNA)
- Disaster Victim Identification (DVI)
  - Two or more persons killed in an event
  - Multiple fatality car crash to 2009 Victorian Bushfires

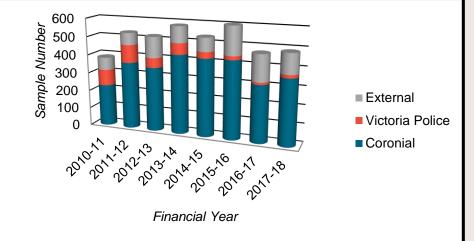
#### □ Missing persons investigations

Unidentified remains reconciled with missing persons

#### DNA testing services

- mtDNA analysis
- ID verification







### **DNA ANALYSIS PIPELINE**

#### PM samples

- Blood: liquid; stain collected on FTA card or swab
- **Tissue**: deep red muscle tissue (~ 1g)
- Toenails: few toe nail clippings or whole toenail
- Bone: compact bone preferably head of femur
- Bladder swabs: cells lining the bladder wall are collected on a swab



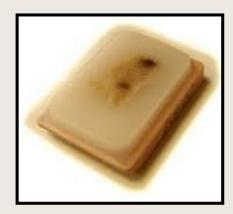




### **DNA ANALYSIS PIPELINE**

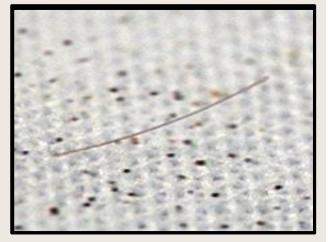
#### □ Other samples

- Fixed tissue: paraffin embedded or slides
- Urine: fresh or frozen
- Hair: with or without the hair root

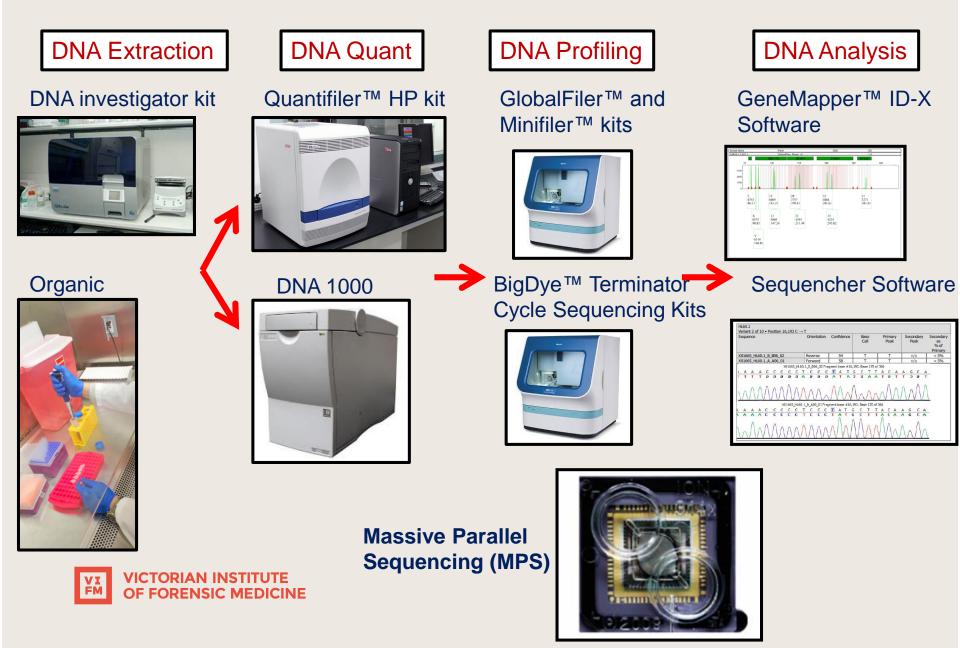




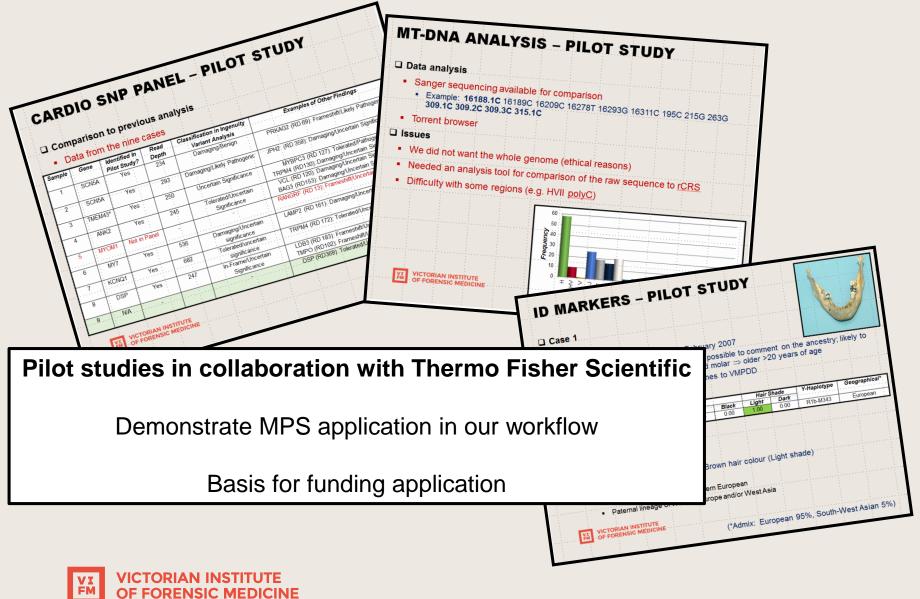




### WORKFLOW



### **MASSIVE PARALLEL SEQUENCING (MPS)**

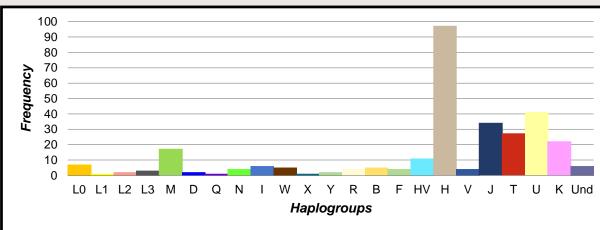


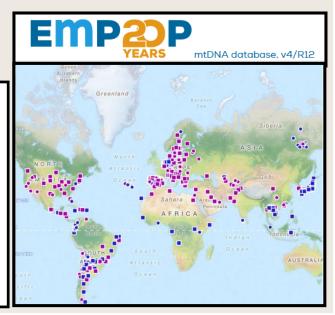
#### Victorian mtDNA haplotype database project

- Frequency of the reported mtDNA profile (haplotype) in the population
- Use EMPOP (EDNAP mtDNA Population Database) database did not contain Australian haplotype information
- mtDNA analysis of ~ 300 volunteer samples

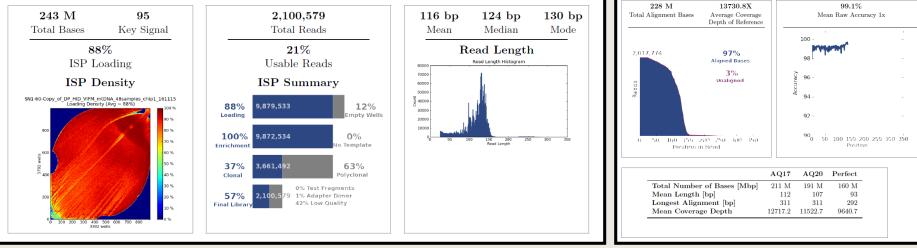
#### Early Access HID-Ion AmpliSeq Mitochondrial Tiling Panel

- 10 samples at full and half AmpliSeq reactions
- 58 samples analysed in total
- Using Ion Chef to generate templates
- Sequenced on 318v2 chips



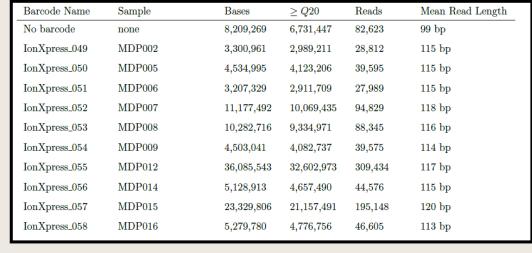


#### Quality of library construction



#### □ Bioinformatics

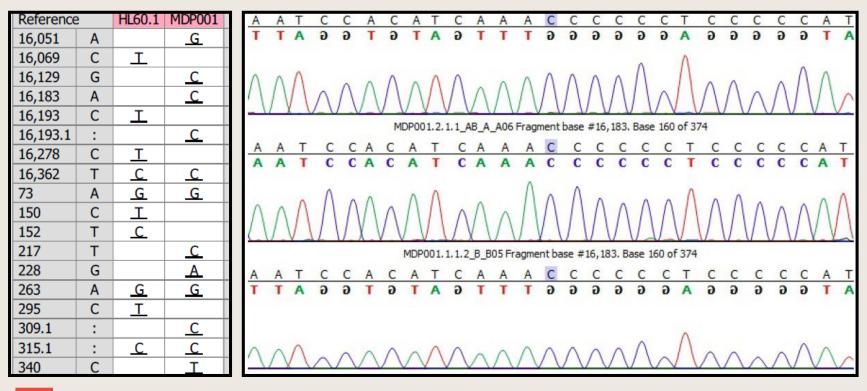
- Ion Torrent Ion Reporter
- Sequencher 5.4
- IGV 2.3





#### Example MDP001

- Sanger sequencing available for comparison
  - HVI (16024-16404) 16051G 16129C 16183C 16193.1C 16362C
  - HVII (18-413) 73G 217C 228A 263G 309.1C 315.1C 340T





#### □ Example MDP001

HVI (16024-16404) HVII (18-413) **16051G** 16129C 16183C **16193.1C** 16362C 73G 217C 228A 263G **309.1C 315.1C** 340T

#### Torrent browser

HALF REACTION													
Chrom	Position	Ref	Variant	Allele Call	Filter	Frequency	Quality	Filter	Туре	Allele Source	Allele	Gene ID	Region
											Name		Name
chrM	73	А	G	Homozygous	-	100	7601.75	-	SNP	Novel		unknown	mt_1
chrM	217	Т	C	Homozygous	-	100	18465.5	-	SNP	Novel		unknown	mt_2
chrM	228	G	А	Homozygous	-	100	18337.9	-	SNP	Novel		unknown	mt_2
chrM	263	A	G	Homozygous	-	100	2210.39	-	SNP	Novel		unknown	mt_3
chrM	311	-	С	Homozygous	-	100	2213.88	-	INS	Novel		unknown	mt_4
chrM	340	С	T	Homozygous	-	100	2184.29	-	SNP	Novel		unknown	mt_4
chrM	16051	A	G	Homozygous	-	96.8	14691.1	-	SNP	Novel		unknown	mt_156
chrM	16052	С	G	leterozygous	-	20.7	316.78	-	SNP	Novel		unknown	mt_156
chrM	16129	G	С	Homozygous	-	100	13476.7	-	SNP	Novel		unknown	mt_158
chrM	16183	A	С	Heterozygous	-	87.6	1848.3	-	SNP	Novel		unknown	mt_158
chrM	16190	-	С	Homozygous	-	98.1	2224.15	-	INS	Novel		unknown	mt_158
chrM	16362	Т	C	Homozygous	-	100	13455.5	-	SNP	Novel		unknown	mt_160
FULL REACTION		_											
Chrom	Position	Ref	Variant	Allele Call	Filter	Frequency	Quality	Filter	Туре	Allele Source	Allele	Gene ID	Region
											Name		Name
chrM	73	A	G	Homozygous	-	100	32137	-	SNP	Novel		unknown	mt_1
chrM	217	Т	С	Homozygous	-	100	19586.9	-	SNP	Novel		unknown	mt_2
chrM	228	G	A	Homozygous	-	100	19719	-	SNP	Novel		unknown	mt_2
chrM	263	A	G	Homozygous	-	100	7749.24	-	SNP	Novel		unknown	mt_3
chrM	311	-	C	Homozygous	-	100	5787.6	-	INS	Novel		unknown	mt_4
chrM	340	C	Т	Homozygous	-	100	3770.41	-	SNP	Novel		unknown	mt_4
chrM	16051	A	G	Homozygous	-	100	21164.2	-	SNP	Novel		unknown	mt_156
chrM	16052	С	G	Heterozygous	-	26.1	953.52	-	SNP	Novel		unknown	mt_156
chrM	16129	G	С	Homozygous	-	100	38364.7	-	SNP	Novel		unknown	mt_158
chrM	16183	А	С	Homozygous	-	93.2	4572.08	-	SNP	Novel		unknown	mt_158
chrM	16190	-	C	Homozygous	-	100	5002.64	-	INS	Novel		unknown	mt_158
chrM	16362	Т	C	Homozygous	-	100	26152.8	-	SNP	Novel		unknown	mt_160
		- T			-			-					

#### □ Comparison: non-concordance

- 58 of 58 samples incorrect calling of cytosine homopolymeric regions in HVI and HVII
- Non-SWGDAM nomenclature used by Ion Reporter
  - i.e. Sample 49 variants "16187T 16188G 16189C" reported as "16187 TGC"
- 1 missed insertion 291.1(non C repeat region)
- 3 missed deletions at 249
- 2 samples with missed variants at 150 and 152
- 2 samples with missed variant at 150 only
- 1 sample with missed variants at 263, 295
- 1 sample with missed variant at 16296
- □ Amendment to Torrent Variant Caller Analysis parameter settings
  - realignment-threshold = 0.9



#### □ Comparison: non-concordance

- 58 of 58 samples incorrect calling of cytosine homopolymeric regions in HVI and HVII
- Non-SWGDAM nomenclature used by Ion Reporter
  - i.e. Sample 49 variants "16187T 16188G 16189C" reported as "16187 TGC"
- I missed insertion 291.1(non c repeat region)
- 3 missed deletions at 249
- I sample with missed variants at 263, 295
- □ Further work
  - Re-run on second chip
  - IGV analysis

			_			e:256-:	10			_	_	69	Π	1	*	40		×	2		_	_	_	_	_	_	_	_		eakin	g: Call				
	-			260 bp				<u>.</u>			_	270	the .					1 bp	-		290 bg	,				1			ą	290 tap	8				2
press_000barn Coverage	8	ŧ.,																																10	_
рнац (190, 19, 2012, 09, 193, vala, (1915-96 серг. (1917) любла, Фазано (1912) град, Бана, 23 ало (1																																			000000000000000000000000000000000000000
ansa 🔫		A C	A	0	c	c	A C	1	T	T	c	c /		A	c	A	6	A	C A	Ţ	c	A	I	A A		C A	A	A	A	A	A	1	1	r c	c

#### □ Things to consider

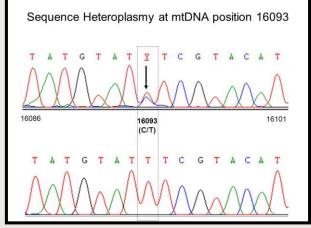
- % of heteroplasmy
  - Reporting thresholds
  - Inheritance of %
- Length vs point heteroplasmy
- Detection of sequencing errors

#### 

- Time effective
- Increased coverage

#### Bioinformatics

Better tools needed



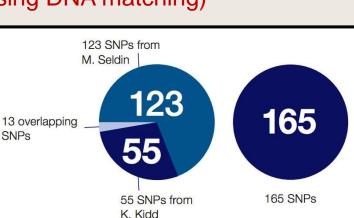




### **ID MARKERS – PILOT STUDY**

#### □ Missing persons and cold case investigations

- Have ~ 40 UHRs in our care
- Varying success of DNA profiling (STR and/or mtDNA)
- □ DNA to predict physical appearance
  - EVCs; e.g. eye and hair colour
- □ DNA to predict ancestry
  - BGA; e.g. European, Asian
- □ Ion AmpliSeq<sup>™</sup> ID panels
  - 30 of the UHR cases
  - Two control cases (long term MP cases IDed using DNA matching)







### **ID MARKERS – PILOT STUDY**

#### Case 1

- UHR
  - Jaw bone located at a beach in February 2007
  - Anthropology: absence cranium not possible to comment on the ancestry; likely to have been female; presence of a third molar ⇒ older >20 years of age
  - Conventional DNA profiling; No matches to VMPDD

#### Phenotypic and ancestry analysis

	Eye Colour			Hair (	Colour		Hair S	Shade	Y-Haplotype	Geographical*
Blue	Intermediate	Brown	Blond	Brown	Red	Black	Light	Dark		
0.01	0.05	0.94	0.09	0.62	0.29	0.00	1.00	0.00	R1b-M343	European

- mtDNA profiling
  - HV (South- Eastern Europe)
- Information provided to MP Unit
  - Male with Brown eye colour and Brown hair colour (Light shade)
  - European ancestry
    - Maternal lineage of South-Eastern European
    - Paternal lineage of Western Europe and/or West Asia





(\*Admix: European 95%, South-West Asian 5%)

### **MASSIVE PARALLEL SEQUENCING (MPS)**



**OF FORENSIC MEDICINE** 

### **MPS VALIDATION**

#### □ Where to begin?

- Genetic screening
- Mitochondrial DNA analysis
- BGA and EVC analysis

#### □ For consideration

- Limited resources (staff and funding for consumables)
- Build on existing expertise
- Uptake once validated



MASSIVE PARALLEL SEQUENCING (MPS)

PILOT STUDY

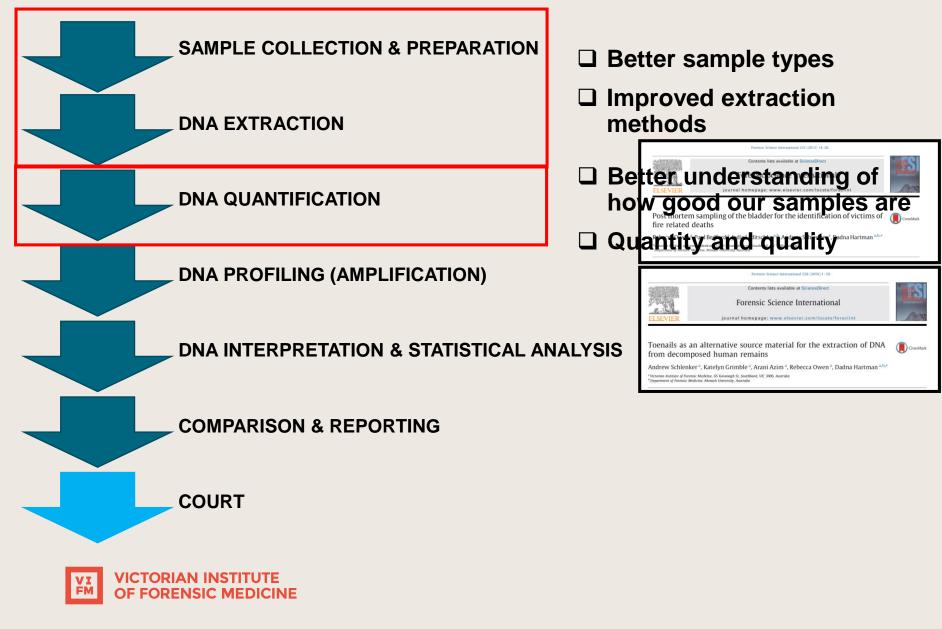
Pilot studies in collaboration with ThermoFisher Scientific

Demonstrate MPS application in our workflow

asis for funding applic:



### **DNA ANALYSIS PIPELINE**



## **QUANTITY AND QUALITY OF N-DNA**

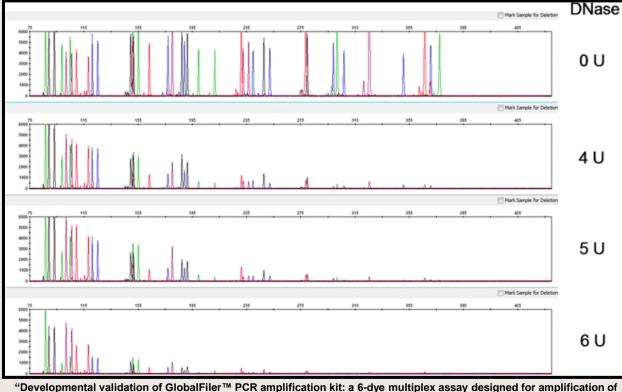
### Quantifiler HP kit

Quantitative and qualitative assessment of total human DNA

#### □ Casework

- How degraded is my sample?
- Include this in the decision matrix used to determine our amplification strategy with GlobalFiler kit

casework samples" International J. Legal Med. (2018) 132: 1555







### DEGRADATION

#### Experimental design

- 10 x 11 x 10 experiment
  - 10 samples with 11 descending concentrations with 10 increasing degradation levels

						С	Concentration (	ng/ul)				
		0.067	0.05-0.067	0.03-0.05	0.02-0.03	0.012-0.02	0.077-0.012	0.005-0.077	0.032-0.005	0.002-0.0032	0.0014-0.002	<0.0014
	<1	C1D1	C2D1	C3D1	C4D1	C5D1	C6D1	C7D1	C8D1	C9D1	C10D1	C11D1
	1-1.25	C1D2	C2D2	C3D2	C4D2	C5D2	C6D2	C7D2	C8D2	C9D2	C10D2	C11D2
ĕ	1.25-1.5	C1D3	C2D3	C3D3	C4D3	C5D3	C6D3	C7D3	C8D3	C9D3	C10D3	C11D3
Inde	1.5-2	C1D4	C2D4	C3D4	C4D4	C5D4	C6D4	C7D4	C8D4	C9D4	C10D4	C11D4
ation	2-3	C1D5	C2D5	C3D5	C4D5	C5D5	C6D5	C7D5	C8D5	C9D5	C10D5	C11D5
adat	3-5	C1D6	C2D6	C3D6	C4D6	C5D6	C6D6	C7D6	C8D6	C9D6	C10D6	C11D6
egra	5-7.5	C1D7	C2D7	C3D7	C4D7	C5D7	C6D7	C7D7	C8D7	C9D7	C10D7	C11D7
Ď	7.5-10	C1D8	C2D8	C3D8	C4D8	C5D8	C6D8	C7D8	C8D8	C9D8	C10D8	C11D8
	10-20	C1D9	C2D9	C3D9	C4D9	C5D9	C6D9	C7D9	C8D9	C9D9	C10D9	C11D9
	>20	C1D10	C2D10	C3D10	C4D10	C5D10	C6D10	C7D10	C8D10	C9D10	C10D10	C11D10

- Test the performance of GlobalFiler kit
  - Ideal (input concentration with no sample degradation)
  - Additional 109 concentrations-degradation categories
- Genemapper ID-X Software (v1.4) with 45RFU as the analytical threshold

## DEGRADATION

#### **Degradation matrix**

- GlobalFiler amp worksheet
  - Input small and large autosomal data
  - Sample is assigned C-D category
  - Most appropriate amp strategy based on required completeness of profile entered (e.g. 75%)
  - The expected % completeness at the selected amplification strategy
- Verification
  - Quantification and profiling data from all casework samples processed since the implementation of GlobalFiler kit
  - High rate of correct predictions (>98%) at 1 x amp



%		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
75%	D1	x1	x1	x1	x1	x1	x2	x2	x2	x4	x4
уı	D2	x1	x1	x1	x2	x2	x2	x2	x4	x4	x4
recovery	D3	x1	x1	x1	x2	x2	x2	x4	x4	x4	x4
rec	D4	x1	x1	x2	x2	x2	x2	x4	x4	x4	x4
ta I	D5	x1	x2	x2	x2	x2	x4	x4	x4	x4	x4
da	D6	x2	x2	x2	x4						
m	D7	x2	x2	x4							
in	D8	x4									
Minimum data	D9	x4									
	D10	x4									
		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
%06	D1	x1	x1	x1	x1	x2	x2	x2	x4	x4	x4
<u>7</u> 9	D2	x1	x1	x1	x2	x2	x2	x4	x4	x4	x4
Minimum data recovery	D3	x1	x1	x1	x2	x2	x2	x4	x4	x4	x4
) SCO	D4	x1	x2	x2	x2	x4	x4	x4	x4	x4	x4
a re	D5	x2	x2	x2	x2	x4	x4	x4	x4	x4	x4
dat	D6	x2	x4								
Ĕ	D7	x4									
nm	D8	x4									
lini	D9	x4									
2	D10	x4									
<b>%</b>		C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
recovery 100%	D1	x1	x1	x1	x2	x2	x2	x4	x4	x4	x4
/ 1(	D2	x1	x2	x2	x2	x2	x4	x4	x4	x4	x4
/er/	D3	x1	x2	x2	x2	x2	x4	x4	x4	x4	x4
co	D4	x1	x2	x2	x4						
	D5	x4									
ata	D6	x4									
Minimum data	D7	x4									
nur	D8	x4									
inin	D9	x4									
Σ	D10	x4									

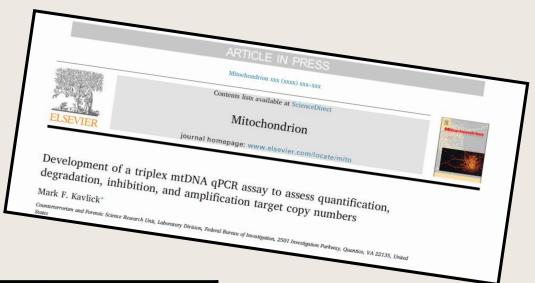
## **QUANTITY AND QUALITY OF MT-DNA**

#### Quantification kit

None commercially available

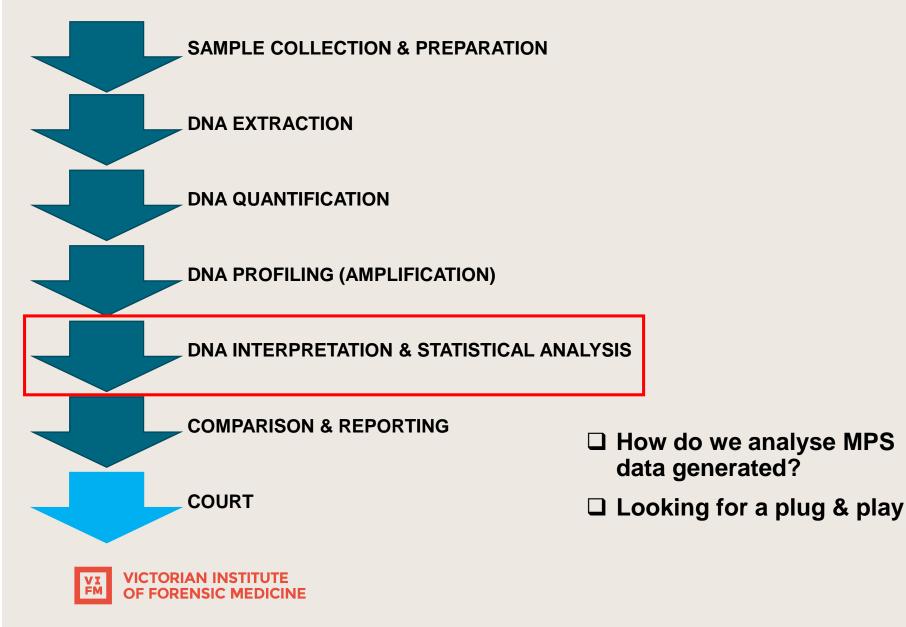
#### In-house quantification

Based on published method



Sample	Nuclear DNA	Mito DNA
	( <i>ng/</i> μ <i>L</i> )	(copies/μL)
1	0.00068	16,883
2	Undetermined	0 (Inhibited)
3	0.00096	1,179
4	Undetermined	368
5	0.00068	676
6	Undetermined	91
7	0.011	12,941

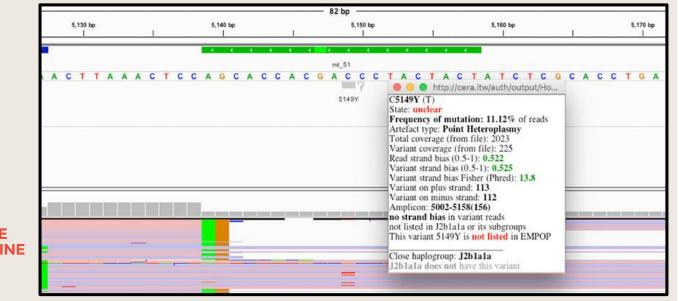
### **DNA ANALYSIS PIPELINE**



### **ANALYSIS TOOLS – CONVERGE™ SOFTWARE**

#### □ Converge<sup>™</sup>

- MPS data analysis
  - Mitochondrial DNA (mtDNA)
  - Ancestry (biogeographic) SNPs
  - Identity SNPs
  - STR markers
- CE to MPS comparison of STR profiles
- Kinship
- Case management



conver



#### □ MPS analysis

- 8 UHR cases
  - Varying degrees of DNA quantity and quality
- Panels
  - Precision ID Ancestry Panel
  - Precision ID Identity Panel
  - DNA Phenotyping Panel (Ion Ampliseq<sup>™</sup> custom pane
  - Precision ID mtDNA Control Region Panel
- Libraries & Sequencing
  - Library construction on Ion Chef<sup>™</sup> System
  - Template preparation on Ion Chef System
  - Sequencing on Ion 530<sup>™</sup> chip on Ion S5<sup>™</sup> System
    Sample

1

2

3

4

5

6

7

8

### Analysis

- Converge Software
- IrisPlex (EVC data)



123 SNPs from M. Seldin 123 165 13 overlapping **SNPs** 55 165 SNPs 55 SNPs from K. Kidd 34 48 SNPforID 124 SNPs 34 upper 13 K Kidd SNPlex system Y-clade SNPs SNPs (1 shared SNP) 90 autosomal SNF Pool 1 Human mtDNA 16,569 bp Nuclear DNA Mito DNA Median Read (copies/µL)  $(ng/\mu L)$ Depth 0.00068 16,883 15,931 Undetermined 0 (Inhibited) 23,297 20.290 0.00096 1.179 Undetermined 368 26,823 0.00068 676 0

91

12,941

Not Available

27,478

19,574

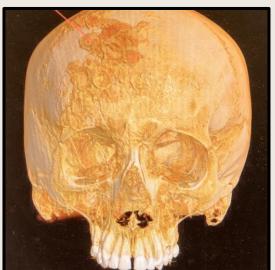
18,408

Undetermined

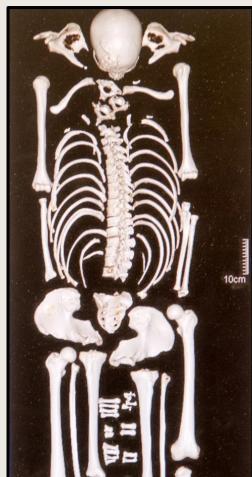
0.011

0.007

- UHR
  - December 2017
  - Sandy Point coast line; near Wilsons Promontory in Victoria
  - Complete skeleton; no fractures; few bones missing;
  - No soft tissue
  - Distinctive dental work (a gold filling in his front tooth)
  - Anthropology: Caucasoid male; in his 20s
  - Conventional DNA profiling: nDNA and mtDNA profiling
  - No matches to VMPDD

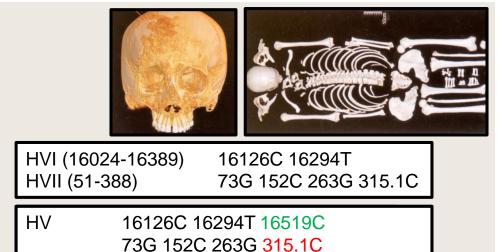


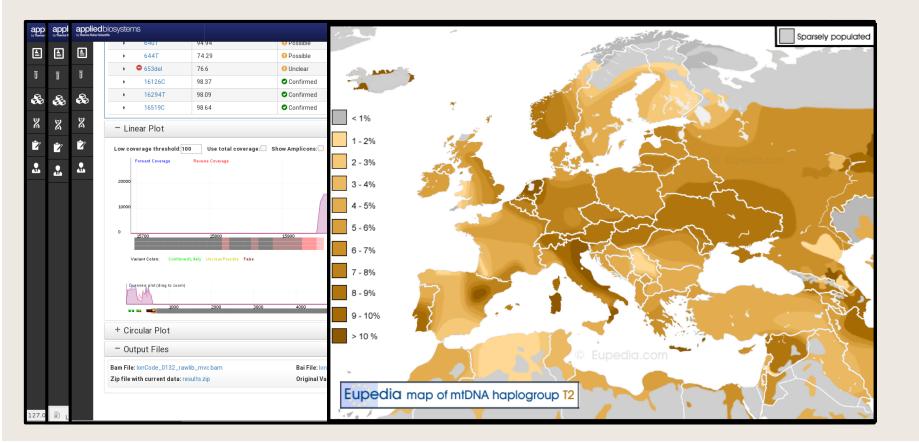




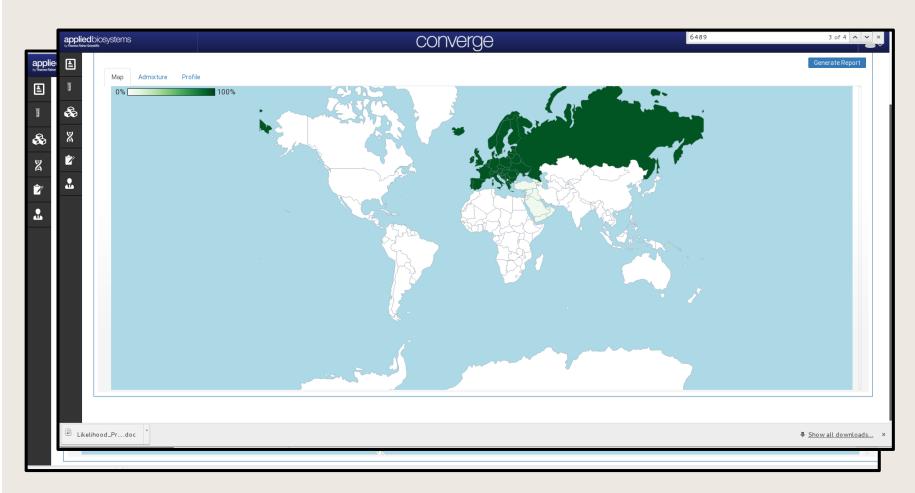


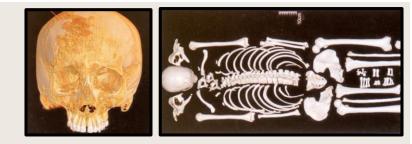
- MPS mtDNA analysis
  - T2
    - Maternal lineage: West Eurasia



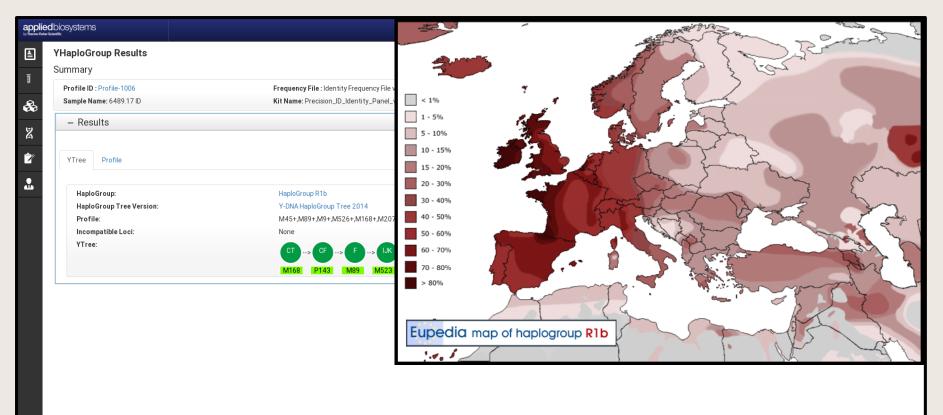


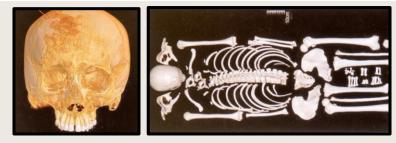
- BGA analysis
  - Admix: 95% Europe; 5% South-West Asia





- Y Haplogroup
  - R1b
    - Paternal lineage: Western Europe





#### □ Sandy Point Skeleton

- UHR
  - Complete skeleton; no fractures; few bones missing;
  - Distinctive dental work (a gold filling in his front tooth)
  - Anthropology: Caucasoid male; in his 20s
  - Conventional DNA profiling: nDNA and mtDNA profiling
- EVC & BGA

	Eye Colour			Hair (	Colour		Hair S	Shade	Y-Haplotype	Geographical*
Blue	Intermediate	Brown	Blond	Brown	Red	Black	Light	Dark		
0.55	0.15	0.30	0.02	0.56	0.00	0.42	0.16	0.84	R1b	European

#### mtDNA profiling

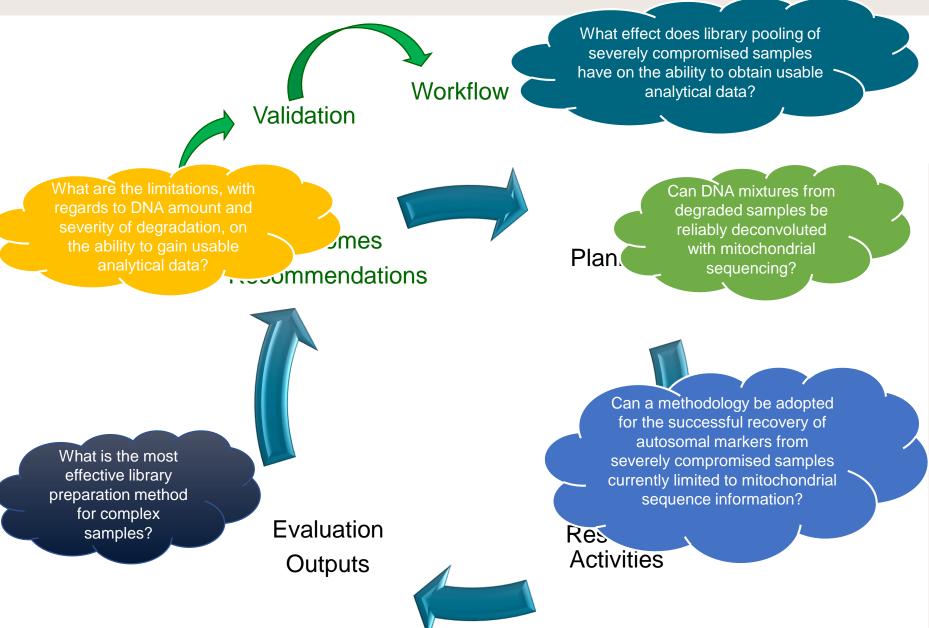
HV 73G 152C 263G 315.1C 16126C 16294T 16519C

T2 (West Eurasia)



(\*Admix: European 95%, South-West Asian 5%)

### NEXT



### ACKNOWLEDGMENTS

VIFM Linda Benton Zoe Bowman Ashil Davawala Dadna Hartman Kaitlyn Hart Andrew Schlenker Michelle Spiden April Stock

















# THANK YOU

When used for purposes other than Human Identification or Paternity Testing the instruments and software modules cited are for Research Use Only. Not for use in diagnostic procedures. Thermo Fisher Scientific and its affiliates are not endorsing, recommending, or promoting any use or application of Thermo Fisher Scientific products presented by third parties during this seminar. Information and materials presented or provided by third parties are provided as-is and without warranty of any kind, including regarding intellectual property rights and reported results. Parties presenting images, text and material represent they have the rights.

