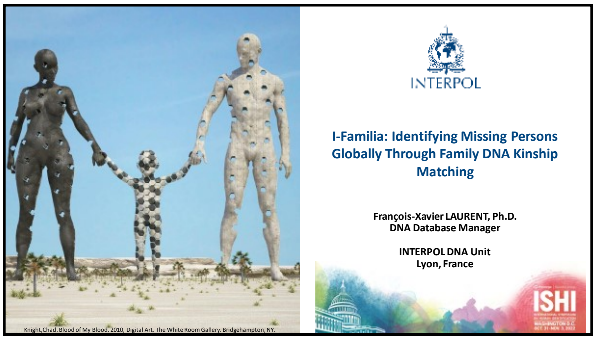


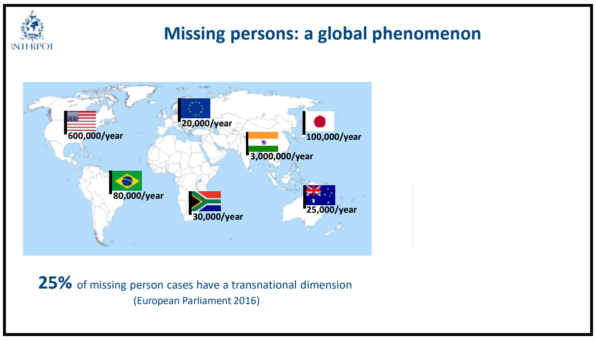
INTERPOL

# I-Familia: Identifying Missing Persons Globally Through Family DNA Kinship Matching

**François-Xavier LAURENT, Ph.D.**
**DNA Database Manager**

**INTERPOL DNA Unit**
**Lyon, France**

Knight,Chad. Blood of My Blood. 2010, Digital Art. The White Room Gallery. Bridgehampton, NY.



## Missing persons: a global phenomenon

**25%** of missing person cases have a transnational dimension
(European Parliament 2016)

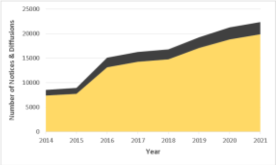
## International police cooperation via INTERPOL

- International organization established in 1923
- INTERPOL General Secretariat in Lyon, France
- Legal framework and technical infrastructure for secure international police data exchange (I-24/7)
- National Central Bureau in each member country

## Notices and Diffusions

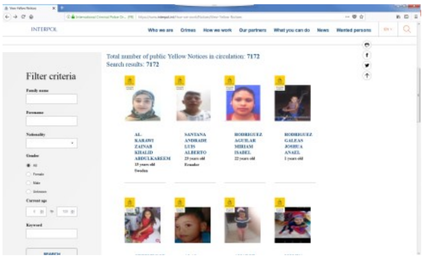
**INTERPOL YELLOW NOTICE**: To help locate **missing persons**, often minors, or to help identify persons who are unable to identify themselves.

**INTERPOL BLACK NOTICE**: To seek information on **unidentified bodies**



| | Valid Notices and Diffusions | Countries |
|---|---|---|
| ■ | 2,493 | 75 |
| 🟨 | 19,883 | 151 |

**Public Website :**
https://www.interpol.int/How-we-work/Notices/Yellow-Notices



Total number of public Yellow Notices in circulation: 7172
Search results: 7172

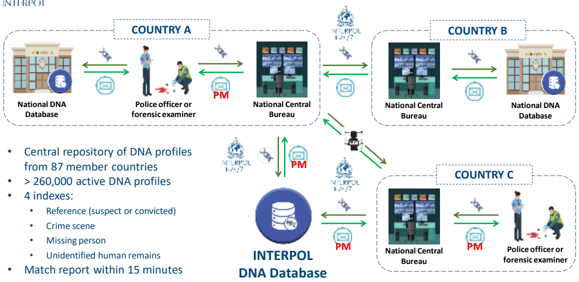
AL-KARAWI ZAINAB KHALID ABDULKAREEM
13 years old
Sweden

SANTANA ANDRADE LUIS ALBERTO
23 years old
Ecuador

RODRIGUEZ AGUILAR MIRIAM ISABEL
22 years old

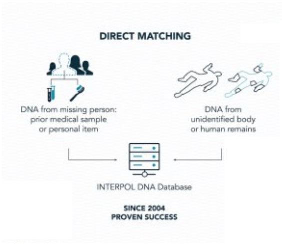
RODRIGUEZ GALEAS JOSHUA ANAEL
1 years old

## International police cooperation via INTERPOL



- Central repository of DNA profiles from 87 member countries
- > 260,000 active DNA profiles
- 4 indexes:
  - Reference (suspect or convicted)
  - Crime scene
  - Missing person
  - Unidentified human remains
- Match report within 15 minutes

## International DNA matching in INTERPOL Databases

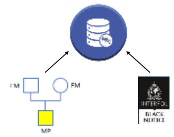


DNA of the missing perso f
international requests (Laur
- Case is too old
- No personal item/med
- No DNA capabilities in


INTERPOL
I-Familia
Identifying missing persons globally through family DNA matching
May 2021

## The 3 components of I-Familia

### I-Familia DNA database



- No criminal DNA data
- No nominal information (only biological relationships, date and DNA profiles)
- Only accessible to the INTERPOL DNA Unit
- Retention period: 5 years for family DNA and 15 years for unidentified human remains

### DNA kinship matching software

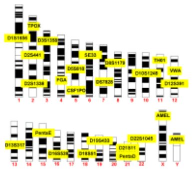
- Compare ante-mortem and post-mortem DNA data
- All pedigrees accepted
- Calculation of likelihood ratio (LR)
- Propose a list of potential biological associations

### Statistical interpretation

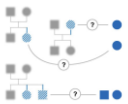
- Interpretation tables developed by INTERPOL
- Quickly evaluate the statistical relevance of the proposed associations
- Report to countries / Reject / Additional information required

## Challenges of international DNA kinship matching

- **First challenge : the choice of method to perform DNA kinship matching computations**



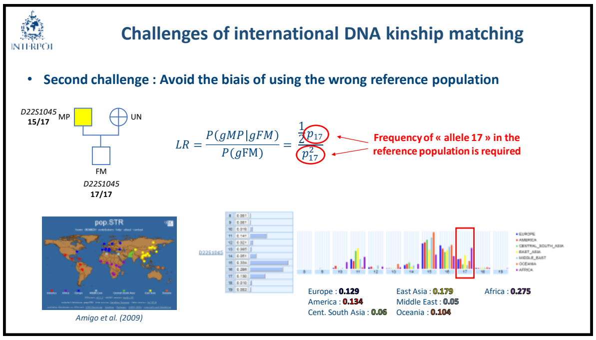
Police data (24 autosomal STRs)
Independent loci
Close relatives only



$LR = P(E|H_1)/P(E|H_2)$

BONAPARTE
SMART

- Bayesian networks
- Likelihood Ratio (LR) computed for any pedigree to measure the fit with any UHR profile
- Rank potential matches by probability of match



## Challenges of international DNA kinship matching

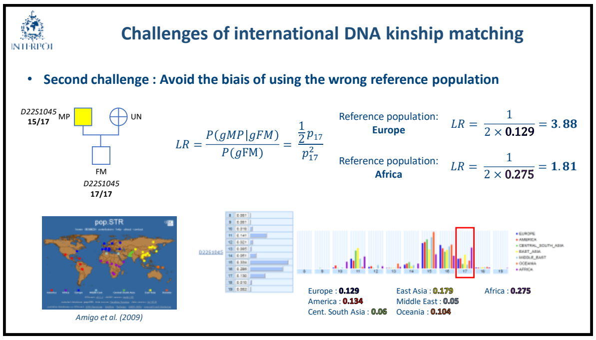
- **Second challenge : Avoid the biais of using the wrong reference population**

$$LR = \frac{P(gMP|gFM)}{P(g\mathrm{FM})} = \frac{\frac{1}{2}p_{17}}{p_{17}^2}$$

**Frequency of « allele 17 » in the reference population is required**

*Amigo et al. (2009)*

Europe : **0.129**
America : **0.134**
Cent. South Asia : **0.06**
East Asia : **0.179**
Middle East : **0.05**
Oceania : **0.104**
Africa : **0.275**



## Challenges of international DNA kinship matching

- **Second challenge : Avoid the biais of using the wrong reference population**

$$LR = \frac{P(gMP|gFM)}{P(g\mathrm{FM})} = \frac{\frac{1}{2}p_{17}}{p_{17}^2}$$

Reference population:
**Europe**

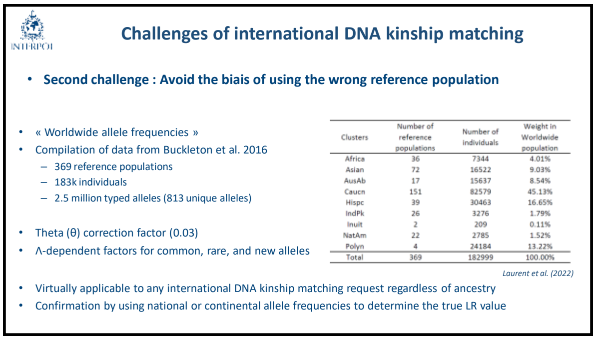
$$LR = \frac{1}{2 \times 0.129} = 3.88$$

Reference population:
**Africa**

$$LR = \frac{1}{2 \times 0.275} = 1.81$$

*Amigo et al. (2009)*

Europe : **0.129**
America : **0.134**
Cent. South Asia : **0.06**
East Asia : **0.179**
Middle East : **0.05**
Oceania : **0.104**
Africa : **0.275**



## Challenges of international DNA kinship matching

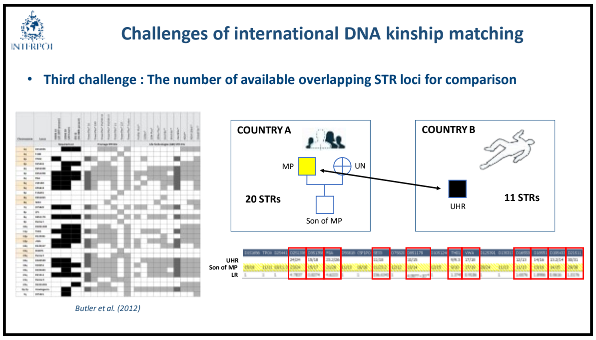
- **Second challenge : Avoid the biais of using the wrong reference population**

- « Worldwide allele frequencies »
- Compilation of data from Buckleton et al. 2016
  - 369 reference populations
  - 183k individuals
  - 2.5 million typed alleles (813 unique alleles)
- Theta (θ) correction factor (0.03)
- Λ-dependent factors for common, rare, and new alleles

| Clusters | Number of reference populations | Number of individuals | Weight in Worldwide population |
|---|---|---|---|
| Africa | 36 | 7344 | 4.01% |
| Asian | 72 | 16522 | 9.03% |
| AusAb | 17 | 15637 | 8.54% |
| Caucn | 151 | 82579 | 45.13% |
| Hispc | 39 | 30463 | 16.65% |
| IndPk | 26 | 3276 | 1.79% |
| Inuit | 2 | 209 | 0.11% |
| NatAm | 22 | 2785 | 1.52% |
| Polyn | 4 | 24184 | 13.22% |
| Total | 369 | 182999 | 100.00% |

*Laurent et al. (2022)*

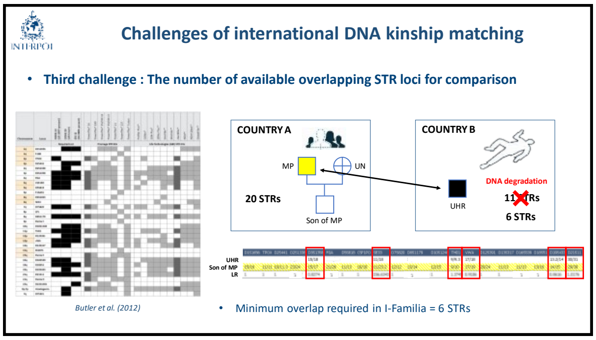
- Virtually applicable to any international DNA kinship matching request regardless of ancestry
- Confirmation by using national or continental allele frequencies to determine the true LR value



## Challenges of international DNA kinship matching

- Third challenge : The number of available overlapping STR loci for comparison

*Butler et al. (2012)*

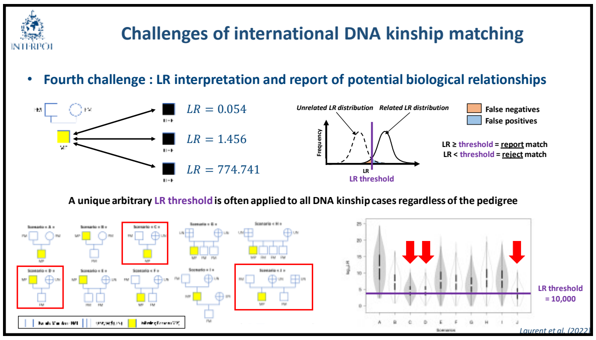


## Challenges of international DNA kinship matching

- Third challenge : The number of available overlapping STR loci for comparison

*Butler et al. (2012)*

- Minimum overlap required in I-Familia = 6 STRs



## Challenges of international DNA kinship matching

- Fourth challenge : LR interpretation and report of potential biological relationships

$LR = 0.054$

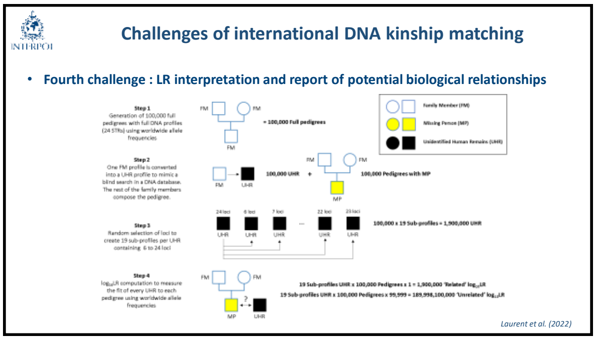
$LR = 1.456$

$LR = 774.741$

LR ≥ threshold = <u>report</u> match
LR < threshold = <u>reject</u> match

A unique arbitrary LR threshold is often applied to all DNA kinship cases regardless of the pedigree

*Laurent et al. (2022)*



## Challenges of international DNA kinship matching

- **Fourth challenge : LR interpretation and report of potential biological relationships**

**Step 1**
Generation of 100,000 full pedigrees with full DNA profiles (24 STRs) using worldwide allele frequencies

**Step 2**
One FM profile is converted into a UHR profile to mimic a blind search in a DNA database. The rest of the family members compose the pedigree.

**Step 3**
Random selection of loci to create 19 sub-profiles per UHR containing 6 to 24 loci

**Step 4**
$\log_{10}$LR computation to measure the fit of every UHR to each pedigree using worldwide allele frequencies

= 100,000 Full pedigrees

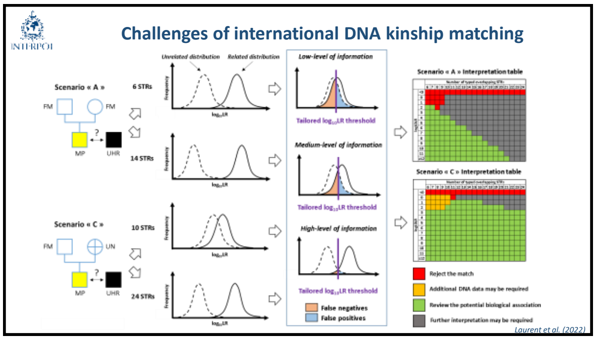
100,000 UHR + 100,000 Pedigrees with MP

100,000 x 19 Sub-profiles = 1,900,000 UHR

19 Sub-profiles UHR x 100,000 Pedigrees x 1 = 1,900,000 'Related' $\log_{10}$LR
19 Sub-profiles UHR x 100,000 Pedigrees x 99,999 = 189,998,100,000 'Unrelated' $\log_{10}$LR

Family Member (FM)
Missing Person (MP)
Unidentified Human Remains (UHR)

*Laurent et al. (2022)*



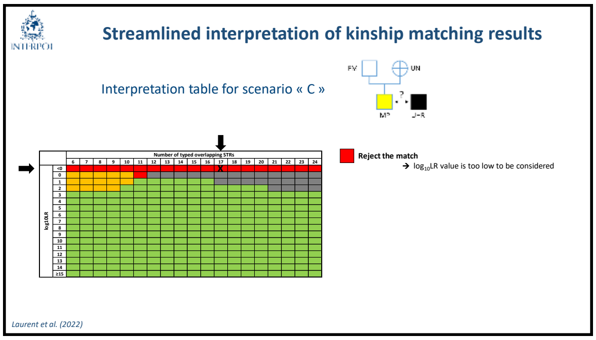
## Challenges of international DNA kinship matching

Tailored $\log_{10}$LR threshold

🟧 False negatives
🟦 False positives

🟥 Reject the match
🟧 Additional DNA data may be required
🟩 Review the potential biological association
⬛ Further interpretation may be required

*Laurent et al. (2022)*



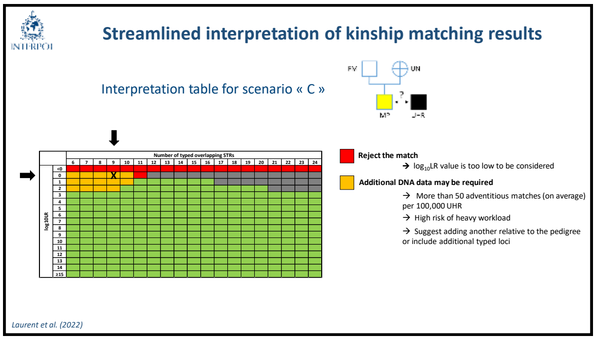
## Streamlined interpretation of kinship matching results

Interpretation table for scenario « C »

| | | Number of typed overlapping STRs | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| log10LR | <0 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | X | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 |
| | 0 | 🟧 | 🟧 | 🟧 | 🟧 | 🟧 | 🟥 | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ |
| | 1 | 🟧 | 🟧 | 🟧 | 🟧 | 🟧 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ | ⬛ |
| | 2 | 🟧 | 🟧 | 🟧 | 🟧 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | ⬛ | ⬛ | ⬛ | ⬛ |
| | 3 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 4 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 5 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 6 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 7 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 8 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 9 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 10 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 11 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 12 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 13 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 14 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | ≥15 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |

🟥 **Reject the match**
→ $\log_{10}$LR value is too low to be considered

*Laurent et al. (2022)*



## Streamlined interpretation of kinship matching results

Interpretation table for scenario « C »

| | | Number of typed overlapping STRs | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| log10LR | <0 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 |
| | 0 | 🟧 | 🟧 | 🟧 | 🟧 X | 🟧 | 🟥 | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ |
| | 1 | 🟧 | 🟧 | 🟧 | 🟧 | 🟧 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ |
| | 2 | 🟧 | 🟧 | 🟧 | 🟧 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | ▨ | ▨ | ▨ | ▨ |
| | 3 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 4 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 5 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 6 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 7 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 8 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 9 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 10 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 11 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 12 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 13 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 14 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | ≥15 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |

🟥 **Reject the match**

→ $\log_{10}$LR value is too low to be considered

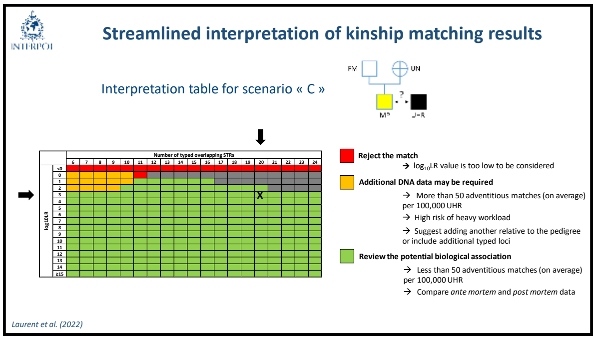
🟧 **Additional DNA data may be required**

→ More than 50 adventitious matches (on average) per 100,000 UHR

→ High risk of heavy workload

→ Suggest adding another relative to the pedigree or include additional typed loci

*Laurent et al. (2022)*



## Streamlined interpretation of kinship matching results

Interpretation table for scenario « C »

| | | Number of typed overlapping STRs | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| log10LR | <0 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 |
| | 0 | 🟧 | 🟧 | 🟧 | 🟧 | 🟧 | 🟥 | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ |
| | 1 | 🟧 | 🟧 | 🟧 | 🟧 | 🟧 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ |
| | 2 | 🟧 | 🟧 | 🟧 | 🟧 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | ▨ | ▨ | ▨ | ▨ |
| | 3 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 X | 🟩 | 🟩 | 🟩 | 🟩 |
| | 4 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 5 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 6 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 7 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 8 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 9 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 10 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 11 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 12 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 13 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 14 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | ≥15 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |

🟥 **Reject the match**

→ $\log_{10}$LR value is too low to be considered

🟧 **Additional DNA data may be required**

→ More than 50 adventitious matches (on average) per 100,000 UHR

→ High risk of heavy workload

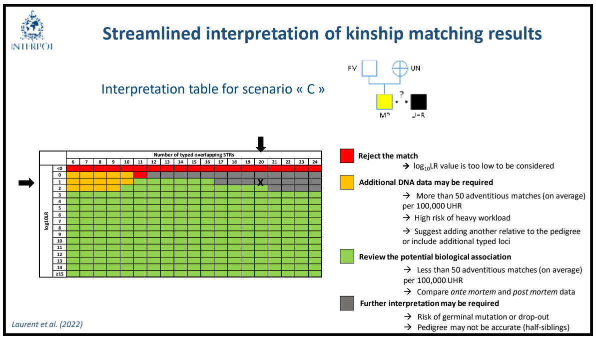
→ Suggest adding another relative to the pedigree or include additional typed loci

🟩 **Review the potential biological association**

→ Less than 50 adventitious matches (on average) per 100,000 UHR

→ Compare *ante mortem* and *post mortem* data

*Laurent et al. (2022)*



## Streamlined interpretation of kinship matching results

Interpretation table for scenario « C »

| | | Number of typed overlapping STRs | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 |
| log10LR | <0 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 | 🟥 |
| | 0 | 🟧 | 🟧 | 🟧 | 🟧 | 🟧 | 🟥 | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ | ▨ |
| | 1 | 🟧 | 🟧 | 🟧 | 🟧 | 🟧 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | ▨ | ▨ | ▨ | ▨ X | ▨ | ▨ | ▨ | ▨ |
| | 2 | 🟧 | 🟧 | 🟧 | 🟧 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | ▨ | ▨ | ▨ | ▨ |
| | 3 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 4 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 5 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 6 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 7 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 8 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 9 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 10 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 11 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 12 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 13 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | 14 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |
| | ≥15 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 | 🟩 |

🟥 **Reject the match**

→ $\log_{10}$LR value is too low to be considered

🟧 **Additional DNA data may be required**

→ More than 50 adventitious matches (on average) per 100,000 UHR

→ High risk of heavy workload

→ Suggest adding another relative to the pedigree or include additional typed loci

🟩 **Review the potential biological association**

→ Less than 50 adventitious matches (on average) per 100,000 UHR

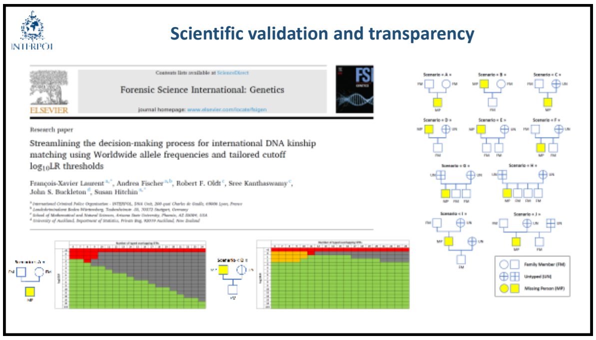
→ Compare *ante mortem* and *post mortem* data

▨ **Further interpretation may be required**

→ Risk of germinal mutation or drop-out

→ Pedigree may not be accurate (half-siblings)

*Laurent et al. (2022)*



## Scientific validation and transparency

Contents lists available at ScienceDirect

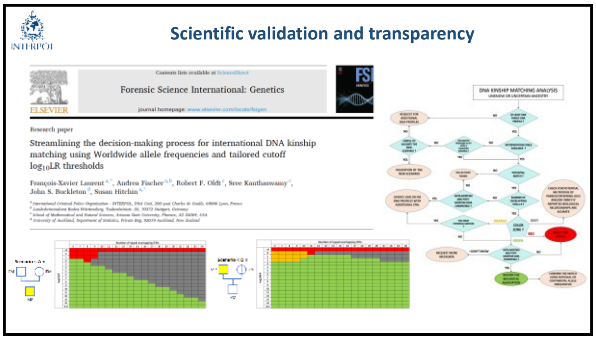
Forensic Science International: Genetics

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

Research paper

Streamlining the decision-making process for international DNA kinship matching using Worldwide allele frequencies and tailored cutoff $\log_{10}$LR thresholds

François-Xavier Laurent[a,*], Andrea Fischer[a,b], Robert F. Oldt[c], Sree Kanthaswamy[c], John S. Buckleton[d], Susan Hitchin[a,*]



## Scientific validation and transparency

Contents lists available at ScienceDirect

Forensic Science International: Genetics

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

Research paper

Streamlining the decision-making process for international DNA kinship matching using Worldwide allele frequencies and tailored cutoff $\log_{10}$LR thresholds

François-Xavier Laurent[a,*], Andrea Fischer[a,b], Robert F. Oldt[c], Sree Kanthaswamy[c], John S. Buckleton[d], Susan Hitchin[a,*]

## First international confirmed match with I-Familia

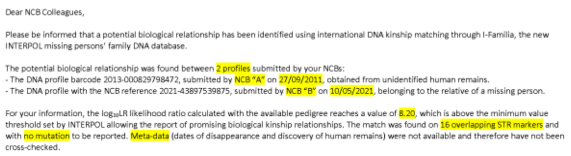
- **October 2020**
- Request from NCB Rome
- Missing Person since April 2004
- No DNA profile available for the missing person
- DNA profile from 2 children
- Upload of pedigree information to I-Familia

- **April 2004**
- Body found in Adriatic Sea by Croatian Police
- No fingerprints or face data
- DNA profile from human remains
- No match in National DNA database
- DNA profile shared by NCB Zagreb in April 2004

Potential match report

Confirmation of the missing person's identity by both NCBs

## Potential match report



Dear NCB Colleagues,

Please be informed that a potential biological relationship has been identified using international DNA kinship matching through I-Familia, the new INTERPOL missing persons' family DNA database.

The potential biological relationship was found between 2 profiles submitted by your NCBs:
- The DNA profile barcode 2013-000829798472, submitted by NCB "A" on 27/09/2011, obtained from unidentified human remains.
- The DNA profile with the NCB reference 2021-43897539875, submitted by NCB "B" on 10/05/2021, belonging to the relative of a missing person.

For your information, the $\log_{10}$LR likelihood ratio calculated with the available pedigree reaches a value of 8.20, which is above the minimum value threshold set by INTERPOL allowing the report of promising biological kinship relationships. The match was found on 16 overlapping STR markers and with no mutation to be reported. Meta-data (dates of disappearance and discovery of human remains) were not available and therefore have not been cross-checked.

As this potential association could lead to the identification of the missing person and help bring closure to the family, we highly encourage your NCB to officially confirm the identity of the human remains using your national procedures.

- **No information given on DNA profiles**
- **No information about the case / family pedigree and familial relationships**
- **Only the countries can decide if they wish to cooperate with the other source of data**

## Statistics of I-Familia

| COUNTRIES | BIOLOGICAL RELATIVES | UNKOWN BODIES | COUNTRIES | BIOLOGICAL RELATIVES | UNKOWN BODIES |
|---|---|---|---|---|---|
| ALBANIA | 2 | 0 | JORDAN | 0 | 3 |
| ALGERIA | 0 | 1 | KAZAKHSTAN | 2 | 0 |
| ARGENTINA | 4 | 0 | KUWAIT | 0 | 1 |
| AUSTRALIA | 6 | 26 | LATVIA | 23 | 4 |
| AZERBAIJAN | 1 | 3 | LIECHTENSTEIN | 0 | 1 |
| BAHRAIN | 0 | 8 | LITHUANIA | 7 | 257 |
| BELARUS | 129 | 93 | LUXEMBOURG | 0 | 4 |
| BELGIUM | 2 | 48 | MALAYSIA | 0 | 172 |
| BELIZE | 0 | 1 | MEXICO | 94 | 3 |
| BOSNIA AND HERZEGOVINA | 0 | 4 | MONTENEGRO | 13 | 10 |
| BRAZIL | 19 | 5948 | MOROCCO | 0 | 39 |
| BULGARIA | 0 | 12 | NETHERLANDS | 0 | 869 |
| CANADA | 0 | 18 | NORTH MACEDONIA | 0 | 13 |
| CHILE | 0 | 1 | NORWAY | 12 | 7 |
| COLOMBIA | 1 | 0 | PANAMA | 0 | 4 |
| COSTA RICA | 0 | 1 | PHILIPPINES | 0 | 11 |
| CROATIA | 4 | 17 | POLAND | 63 | 35 |
| CYPRUS | 4 | 8 | PORTUGAL | 0 | 4 |
| CZECH REPUBLIC | 11 | 527 | ROMANIA | 17 | 3 |
| DENMARK | 0 | 16 | RUSSIAN FEDERATION | 36 | 61 |
| ESTONIA | 3 | 7 | SERBIA | 2 | 11 |
| FINLAND | 0 | 4 | SLOVAKIA | 49 | 71 |
| FRANCE | 13 | 155 | SLOVENIA | 178 | 23 |
| GERMANY | 135 | 253 | SPAIN | 6 | 131 |
| GIBRALTAR | 0 | 2 | SWEDEN | 1 | 9 |
| GREECE | 6 | 67 | SWITZERLAND | 3 | 34 |
| HONG KONG | 0 | 1 | SYRIAN ARAB REPUBLIC | 0 | 2 |
| HUNGARY | 0 | 11 | THAILAND | 0 | 7 |
| INDIA | 0 | 10 | TURKEY | 0 | 40 |
| IRAQ | 0 | 12 | UKRAINE | 8 | 27 |
| IRELAND | 51 | 28 | UNITED ARAB EMIRATES | 0 | 39 |
| ISRAEL | 0 | 79 | UNITED KINGDOM | 3 | 218 |
| ITALY | 2 | 19 | UNITED STATES | 4 | 21 |
| JAPAN | 0 | 5 | URUGUAY | 19 | 0 |

| TOTAL WORLD | 933 RELATIVES | 9,519 UNKNOWN BODIES |
|---|---|---|

**68 countries are providing DNA data**

**3 confirmed matches**

**40 matches pending confirmation**

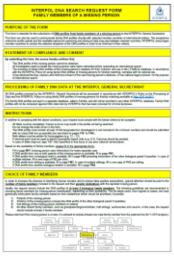
**International and national matches**

🟩 Mass uploads

🟧 Changes in the DNA legislation

## How can you help ?

- **Submit UHR and Family DNA Data to INTERPOL via your NCB**
- No nominal information required
- 1 biological relative required (2 prefered)
- Full STR profiles highly appreciated
- Contains instructions (choice of family members) and rules on the processing of data
- Autosomal DNA (Y-STR and mtDNA optional)
- Statement of compliance and consent
  - The source hereby confirms that its provision of these DNA profiles of the missing person's relatives for their inclusion and use in the I-FAMILIA database, in accordance with the INTERPOL Policy on using family DNA profiles of missing persons for kinship matching, complies with its national law. The source further confirms that it has obtained the free, express and informed consent of the said missing person's relatives, or has national legal provision, for the purpose of international search.
- Forms, brochure and policy available in 4 languages ar en es fr
- Mass-upload options available



INTERPOL DNA SEARCH REQUEST FORM
FAMILY MEMBERS OF A MISSING PERSON

PURPOSE OF THE FORM

STATEMENT OF COMPLIANCE AND CONSENT

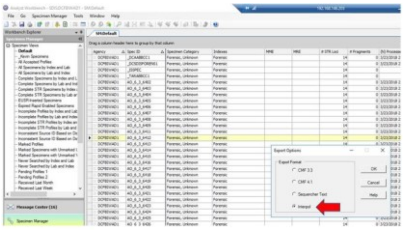
PROCESSING OF FAMILY DNA DATA AT THE INTERPOL GENERAL SECRETARIAT

INSTRUCTIONS

CHOICE OF FAMILY MEMBERS

## Extraction of national UHR DNA profiles

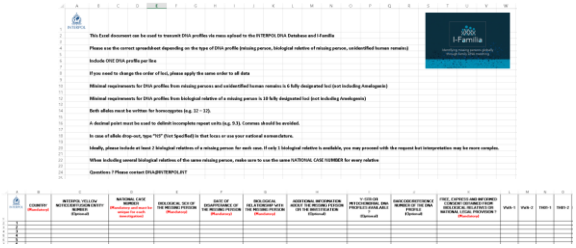
CODIS-to-INTERPOL export function



28

## Extraction of national biological relatives DNA profiles

XLSX2XML SPREADSHEET



This Excel document can be used to transmit DNA profiles via mass upload to the INTERPOL DNA Database and I-Familia

Please use the correct spreadsheet depending on the type of DNA profile (missing person, biological relative of missing person, unidentified human remains)

Include ONE DNA profile per line

If you need to change the order of loci, please apply the same order to all data

Minimal requirements for DNA profiles from missing persons and unidentified human remains is 6 fully designated loci (not including Amelogenin)

Minimal requirements for DNA profiles from biological relative of a missing person is 10 fully designated loci (not including Amelogenin)

Both alleles must be written for homozygotes (e.g. 12 – 12).

A decimal point must be used to delimit incomplete repeat units (e.g. 9.3). Commas should be avoided.

In case of allele drop-out, type "NS" (Not Specified) in that locus or use your national nomenclature.

Ideally, please include at least 2 biological relatives of a missing person for each case. If only 1 biological relative is available, you may proceed with the request but interpretation may be more complex.

When including several biological relatives of the same missing person, make sure to use the same NATIONAL CASE NUMBER for every relative

Questions ? Please contact DNA@INTERPOL.INT

|  | COUNTRY (Mandatory) | INTERPOL YELLOW NOTICE/DIFFUSION ENTITY NUMBER (Optional) | NATIONAL CASE NUMBER (Mandatory and must be unique for each investigation) | BIOLOGICAL SEX OF THE MISSING PERSON (Mandatory) | DATE OF DISAPPEARANCE OF THE MISSING PERSON (Mandatory) | BIOLOGICAL RELATIONSHIP WITH THE MISSING PERSON (Mandatory) | ADDITIONAL INFORMATION ABOUT THE MISSING PERSON OR THE INVESTIGATION (Optional) | Y-STR OR MITOCHONDRIAL DNA PROFILES AVAILABLE ? (Optional) | BARCODE/REFERENCE NUMBER OF THE DNA PROFILE (Optional) | FREE, EXPRESS AND INFORMED CONSENT OBTAINED FROM BIOLOGICAL RELATIVES OR NATIONAL LEGAL PROVISION ? (Mandatory) | vWA-1 | vWA-2 | TH01-1 | TH01-2 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

29

## Acknowledgments

**INTERPOL DNA Unit**

Susan Hitchin  
Claire Dennis  
Anna Nikolaeva  
Andrea Fischer

**Collaborators**

Robert F. Oldt  
Sree Kanthaswamy  
John Buckleton

3 anonymous reviewers of our published manuscript  
Member countries which provide the DNA data  
Families of missing persons

DNA@interpol.int

30

## INTERPOL's new service : I-Familia

Laurent et al. under review

**Efficiency of worldwide allele frequencies and tailored LR thresholds confirmed using simulated DNA profiles**

Forensic Science International: Genetics

Correspondence

Corrigendum to 'U.S. Population Data for 29 Autosomal STR Loci' [Forensic Sci. Int. Genet. 7 (2013) e82–e83]

Carolyn R. Steffen, Michael D. Coble, Katherine B. Gettings, Peter M. Vallone

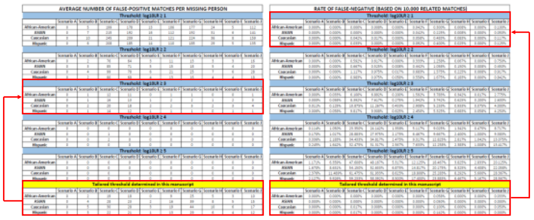
**NIST 1036 Revised U.S. Population Dataset (July 2017)** NEW

**African-American / Asian / Caucasian / Hispanic**

**Isolated populations:**
Native Americans (Ng et. al, 2016)
Mexican-Mestizo population from Northwest Mexico (Martinez-Cortes et. al, 2019)
Circassian subpopulation in Jordan (Al-Eitan et. al, 2020)

## INTERPOL's new service : I-Familia

Classification efficiency of related and unrelated matches using the a specific threshold vs. tailored thresholds



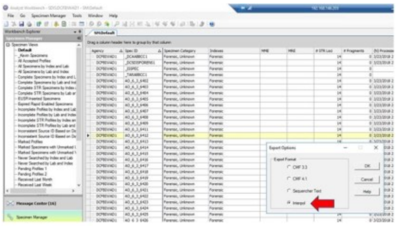
**As stringent against false-positive reports as a LR threshold of 1,000**

**Efficient identification of real matches with a false-negative rate similar to a LR threshold of 1**

Laurent et al. under review

## Mass-upload options

**CODIS-to-INTERPOL export function**

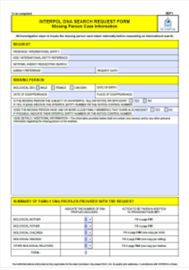


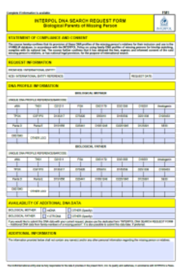
**XLSX-to-INTERPOL export function**

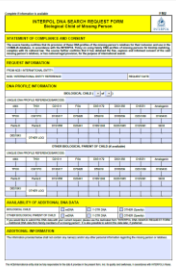
Excel spreadsheet for mass upload

34

## How can you help ?

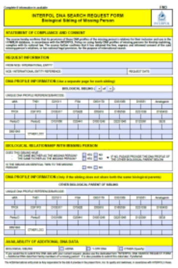

INTERPOL DNA SEARCH REQUEST FORM
Missing Person Case Information

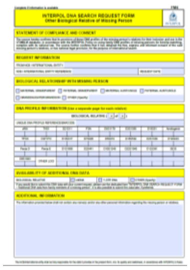

INTERPOL DNA SEARCH REQUEST FORM
Biological Parents of Missing Person

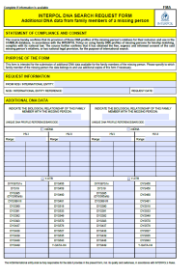

INTERPOL DNA SEARCH REQUEST FORM
Biological Child of Missing Person

35

## How can you help ?


INTERPOL DNA SEARCH REQUEST FORM
Biological Sibling of Missing Person


INTERPOL DNA SEARCH REQUEST FORM
Other Biological Relative of Missing Person


INTERPOL DNA SEARCH REQUEST FORM
Additional DNA data from family members of a missing person

36