• Evaluating the statistical power of DNA-based identification, exemplified by «The missing grandchildren of Argentina»
  – More genotyped individuals, additional markers, needed?

General problem in genetics: account for linkage disequilibrium
Argentina 1976 - 1983

- Military dictatorship
- *War* against guerrillas
- Opponents killed or disappeared

- *Children abducted*
  - kidnapped with their parents or born in captivity
  - parents killed
  - raised by police or military families
1982: Falklands war against UK
1983: Democracy restored
1985: First trial against dictatorship leaders: life imprisonment sentence
1986: Final point law
1987: Due obedience law

Exceptions: Theft, rape and *child abduction*
The missing grandchildren

- **Grandmothers of Plaza de Mayo**
  - formed in 1977
  - weekly marches ever since

- 1984: First grandchild recovered
  - HLA typing + blood groups

- 1989: National genetic data bank

- In total: 122 reunifications
  - 75 of these by BNDG
The missing Grand Children of Argentina: *Family reunification*

Press conference April 2017: Grandchild no. 122 recovered
Estela de Carlotto anuncia el encuentro de otra nieta la #125 en el acto por los #40AñosDeLucha

13:08 - 27 oct. 2017

3 439 919
Genetics of family reunification

- DNA-based evidence
  - autosomal markers
  - mtDNA
  - Y chromosome

- Simplest when
  - available DNA from the missing person
  - available DNA from parents

- Argentina: Parents usually dead/missing

Forensic markers:
- standard kits, 15 - 24 STRs
- up to 50 alleles
- unlinked (mostly)
Genetics of family reunification

POI = person of interest
MP = missing person

\[
LR = \frac{P(\text{marker data} \mid MP = POI)}{P(\text{marker data} \mid POI \text{ unrelated})}
\]
1. Founders (HW)

2. Gene dropping

3. Remove untyped

SNP: \( p_A = p_B = 0.5 \)

Unconditional simulation

1. Compute conditional distribution in child

2. Sample from this SNP:
   \( p_A = p_B = 0.5 \)

MUCH harder than unconditional

Conditional simulation

1. Compute conditional distribution in child

2. Sample from this

Algorithm in Weeks, Ott, Lathrop (1990) implemented with speed-ups by Vigeland
Do we have enough data to give a positive match if POI = MP?

Exceedance probability

Can be computed by simulation!
- Unconditional → average for all pedigrees of this type
- Conditional → probability for this particular case
- Thresholds differ depending on context and lab, we use 10000 in examples:

$$E_{10000} = P(LR > 10000|POI = MP)$$

Implemented in R library paramlink (Magnus Dehli Vigeland)
Statistical power of reunification - part 2

- Do we have enough data to **exclude** an unrelated POI?
- **Exclusion probability**
- Can be computed exactly (Egeland, Pinto, Vigeland, 2014).
  - Implemented in **paramlink**

\[ PE = P(\text{data incompat. with ped} | \text{POI unrelated}) \]
Status May 2017

- Power evaluation of 196 families in the BNDG database
  - 171 unsolved
  - 25 solved

<table>
<thead>
<tr>
<th>Typed</th>
<th>Families</th>
<th>Parent(s) typed</th>
<th>2nd degree only</th>
<th>Exclusion possible</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>5</td>
<td>15</td>
<td>14</td>
</tr>
<tr>
<td>3</td>
<td>29</td>
<td>7</td>
<td>20</td>
<td>22</td>
</tr>
<tr>
<td>4</td>
<td>41</td>
<td>6</td>
<td>33</td>
<td>33</td>
</tr>
<tr>
<td>5</td>
<td>29</td>
<td>7</td>
<td>14</td>
<td>26</td>
</tr>
<tr>
<td>6</td>
<td>26</td>
<td>5</td>
<td>14</td>
<td>21</td>
</tr>
<tr>
<td>7</td>
<td>14</td>
<td>0</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>8</td>
<td>14</td>
<td>0</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>9</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>10+</td>
<td>7</td>
<td>0</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Total</td>
<td>196</td>
<td>35</td>
<td>126</td>
<td>151</td>
</tr>
</tbody>
</table>

- For each family: Compute $E_{10000}$ and PE
Excellent power
Excellent power and PE

- 68 families with PE > 99% and $E_{10000} > 99$
- Includes all (except 1) of the 31 cases with parental data
Poor power in spite of many typed relatives
Low power despite many typed

PE = 81 %
$E_{10000} = 6 \%$
Good power, but not for exclusion
Good power - exclusion impossible

$PE = 0\%$

$E_{10000} = 82\%$

$PE = 0\%$

$E_{10000} = 89\%$
Overall

• 34% of the unsolved families have poor power with threshold 10,000

• Reasons:
  – few markers
  – few typed relatives

Ongoing actions:
• retyping 1000 individuals to 24 markers
• exhumation of 100 (!) informative relatives
Exhumation of paternal grandmother and markers increased from 15 to 24

<table>
<thead>
<tr>
<th>Family</th>
<th>Typed</th>
<th>Markers</th>
<th>Median LR</th>
<th>95% Interval</th>
<th>P(LR&gt;100)</th>
<th>P(LR&gt;1000)</th>
<th>P(LR&gt;10000)</th>
</tr>
</thead>
<tbody>
<tr>
<td>left</td>
<td>3</td>
<td>15</td>
<td>246703</td>
<td>[10, 6.5E+9]</td>
<td>0.928</td>
<td>0.856</td>
<td>0.732</td>
</tr>
<tr>
<td>right</td>
<td>4</td>
<td>24</td>
<td>1.32E+10</td>
<td>[6E+3, 2.8E+15]</td>
<td>0.998</td>
<td>0.990</td>
<td>0.969</td>
</tr>
</tbody>
</table>
Implementation and Summary

• Software: Familias
  – Original publication: Egeland, Mostad et al., 2000
  – Now maintained, extended, by Daniel Kling

• Used in all 75 identifications by BNDG

• Familias exports to R for
  ✓ Plotting
  ✓ Conditional simulation
  ✓ Probability of exclusion
  ✓ …

Currently in BNDG
• 350 families
• 10 000 POIs
2015: Forensic Anthropology Unit at BNDG
Algunos crímenes muy argentinos

La historia argentina está marcada por crímenes políticos que tardan años en resolverse o no se resuelven nunca, como las muertes de Alberto Nisman, Carlos Menem Jr. y recientemente el caso de Santiago Maldonado. Esta realidad ha originado una desconfianza muy amplia hacia la administración de justicia.