

INTERPRETACIÓN DE PERFILES COMPLEJOS: EL MODELO DROP-OUT & DROP-IN

Lourdes Prieto
Comisaría General de Policía Científica
Laboratorio de ADN, Madrid, España



Modelo drop-out & drop-in

- Introducción
- Valor del LR e integración con otros medios de prueba
- Desventajas del LR
- Ventajas del LR
- El LR binario
- Cálculo de LR en muestras problemáticas
- Otras herramientas útiles:
 - Matches en bases de datos (Mixture deconvolution)
 - GeneMapper to Familias

Introducción

- Recomendaciones ISFG (2006): se prefieren LR
- Recomendaciones SWGDAM (2010): no afirma la preferencia por uno u otro método
- Comparación de ambos métodos:
 - Buckleton and Curran (2008), A discussion of the merits of Random Man Not Excluded and Likelihood Ratios *FSI-G* 343-348

Introducción: concepto de LR

- Se basa en el Teorema de Bayes (probabilidades condicionadas)
- El teorema de Bayes sirve para conocer las probabilidades finales de un suceso a partir de las probabilidades iniciales, dada cierta información o informaciones adicionales
 - ¿cuál es la probabilidad del hallazgo científico si suponemos culpabilidad?
 - Y ¿cuál es la probabilidad del hallazgo científico si suponemos inocencia?

Introducción: concepto de LR

• Con el LR expresamos el mayor o menor grado de incertidumbre de los resultados de nuestros análisis en forma de un cociente entre la posición de la acusación y la posición de la defensa

• A ese cociente lo llamamos LR (likelihood ratio):

$$LR = \frac{\text{prob. de la evidencia bajo } H_p}{\text{prob. de la evidencia bajo } H_d}$$

Introducción: concepto de LR

• Para usar este abordaje es necesario enunciar al menos dos hipótesis sobre los hechos, por ejemplo:

- H_p (hipótesis del fiscal) = la mancha de sangre hallada en la escena del crimen pertenece al acusado
- H_d (hipótesis de la defensa) = la mancha de sangre hallada en la escena del crimen NO pertenece al acusado.

• Este enunciado se ha simplificado en este ejemplo con el fin de hacer entendible la explicación, pero en realidad, si el acusado NO dejó la sangre en la escena, en la hipótesis de la defensa se debe definir con más precisión quién dejó la mancha en la escena: ¿un individuo al azar de la población?, ¿de qué población?, ¿un individuo relacionado familiarmente con el acusado?

Introducción: concepto de LR

• Evidencia ↔ Indubitada

Muestra	TH01	TPOX	CSF1	D3S13	VWA	FGA	D8S11	D21S1	D18S5	D5S81	D13S3	D7S820
Eviden.	7-9.3	8-10	11-12	14-15	16-19	23-25	12-12	30-30.2	10-18	11-12	8-13	9-12
Indubit.	7-9.3	8-10	11-12	14-15	16-19	23-25	12-12	30-30.2	10-18	11-12	8-13	9-12

$$LR = \frac{P(E/H_p)}{P(E/H_d)} = \frac{1}{\text{Frec. perfil}} = \frac{1}{7,20 \times 10^{-18}} = 138.868 \text{ billones}$$

Introducción: concepto de LR

• Valor de la coincidencia

ES APROXIMADAMENTE 138.000 BILLONES DE VECES MÁS PROBABLE ENCONTRAR ESTE PERFIL EN LA EVIDENCIA SI PROCEDE DEL SOSPECHOSO QUE SI PROCEDE DE UN INDIVIDUO AL AZAR DE LA POBLACIÓN ESPAÑOLA



Alto valor probatorio

En los casos más simples el LR coincide numéricamente con la RMP: proporcionalmente, 1 de cada 138.000 billones de individuos españoles presenta este perfil genético
Pero LR y RMP son dos conceptos diferentes

Modelo drop-out & drop-in

- Introducción
- Valor del LR e integración con otros medios de prueba
- Desventajas del LR
- Ventajas del LR
- El LR binario
- Cálculo de LR en muestras problemáticas
- Otras herramientas útiles:
 - Matches en bases de datos (Mixture deconvolution)
 - GeneMapper to Familias

Valor del LR

- Si $LR > 1$ → la evidencia apoya la hipótesis del fiscal
- Si $LR = 1$ → la evidencia es neutra
- Si $LR < 1$ → la evidencia apoya la hipótesis de la defensa

Quando el perfil genético de la evidencia coincide con el de la muestra de referencia los resultados del LR suelen ser abrumadores. Pero se puede calcular también un LR si ambos perfiles no coinciden

CALCULAR UN LR NO SIGNIFICA IDENTIFICAR

Integración con otros medios de prueba

El LR permite la integración de la prueba genética con otros medios de prueba:

- En general, hay otro tipo de pruebas distintas a la genética en la investigación de los hechos delictivos (testigos, etc.)
- El juez puede tener una idea sobre la culpabilidad o inocencia de un individuo, es lo que llamamos probabilidad *a priori*
- La probabilidad *a priori* se relaciona con el LR de la siguiente forma:

$$\text{prob. } a \text{ posteriori} = \text{prob. } a \text{ priori} \times LR$$

Dato genético
Dato no genético

Integración con otros medios de prueba

- Ejemplo de LR elevado, a priori bajo:
 - filtro de cigarrillo en la escena del delito
- Ejemplo de LR bajo, a priori alto:
 - ADN degradado (perfil genético parcial) en una identificación cadavérica
- La prueba genética no da la respuesta sobre la culpabilidad o inocencia de un individuo

Integración con otros medios de prueba

- El uso del LR rompe con la tendencia errónea de los peritos de decidir por el juez y obliga a éste y a los peritos a asumir su papel específico.

<i>a priori</i>	LR	<i>a posteriori</i>
1000 a 1 a favor de inocente (es decir 0,1% culpable y 99,9% inocente)	100	10 a 1 a favor de inocencia (0,001 x 100 = 10)
1 a 1 (mismas posibilidades de culpabilidad o inocencia, es decir 50% culpable y 50% inocente)	100	100 a 1 a favor de culpabilidad (1 x 100)
1000 a 1 a favor de culpable (es decir, 10% inocente y 90% culpable)	100	100000 a 1 a favor de culpabilidad (1000 x 100 = 100000)

Integración con otros medios de prueba

Casos especiales de probabilidades *a priori*

- En casos de paternidad:
 - Misma probabilidad *a priori* a la hipótesis de "paternidad" que a la hipótesis de "no paternidad": 1/1
 - Porb. *a posteriori* = 1 x LR; W = LR / LR+1
- En casos masivos de identificación:
 - Calculadas por el antropólogo / investigadores
 - Desaparecidos en Chipre: 2700 individuos (*prior* = 1/2700)
 - Se puede acotar: por lugares?, por sexo, por edad...

Modelo drop-out & drop-in

- Introducción
- Valor del LR e integración con otros medios de prueba
- Desventajas del LR
- Ventajas del LR
- El LR binario
- Cálculo de LRs en muestras problemáticas
- Otras herramientas útiles:
 - Matches en bases de datos (Mixture deconvolution)
 - GeneMapper to Familias

Desventajas del LR

- La valoración de la prueba genética desde el punto de vista bayesiano no es ni mucho menos intuitiva
- Se cometen errores y malinterpretaciones
- La transposición de la condicional es el error más común y se produce al expresar el LR con palabras (ej. LR = 1000)

CORRECTO	INCORRECTO
Es mil veces más probable evidenciar este perfil genético (el resultante de los análisis) en la mancha de la escena si el perfil procede del acusado que si procede de otra persona al azar de la población española	Es mil veces más probable que la mancha proceda del acusado en comparación con que proceda de otra persona al azar de la población española
prob (E/Hp) / prob (E/Hd) = LR	prob (Hp) / prob (Hd) ≠ LR

Desventajas del LR

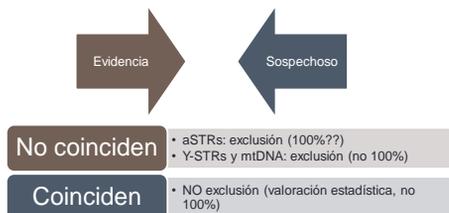
- Difícil de transmitir al juez
- Significado del valor del LR
 - $0 \leq LR \leq +\infty$
 - **No existe ningún estándar científico para determinar cuándo un valor de probabilidad concreto equivale a casi certeza**
 - Paternidad, por doctrina del TS español: probabilidad de paternidad superior al 99,73% (LR ~ 400, si *prior* es 50%) en la prueba biológica equivale a una paternidad prácticamente probada (Hummel, 1971)
 - Este umbral es considerado muy bajo por parte de los genetistas
 - Ejemplo moneda trucada

Modelo drop-out & drop-in

- Introducción
- Valor del LR e integración con otros medios de prueba
- Desventajas del LR
- Ventajas del LR
- El LR binario
- Cálculo de LRs en muestras problemáticas
- Otras herramientas útiles:
 - Matches en bases de datos (Mixture deconvolution)
 - GeneMapper to Familias

Ventajas del LR

- No es necesario tomar una decisión sobre inclusión, exclusión o inconcluyente
 - Muy importante en mezclas y en perfiles *low level DNA*
- Los usuarios del LR estamos acostumbrados a usarlo desde un punto de vista binario, pero no necesariamente hay que usarlo así:



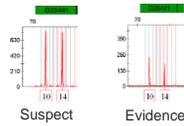
Modelo drop-out & drop-in

- Introducción
- Valor del LR e integración con otros medios de prueba
- Desventajas del LR
- Ventajas del LR
- ELLR binario
- Cálculo de LRs en muestras problemáticas
- Otras herramientas útiles:
 - Matches en bases de datos (Mixture deconvolution)
 - GeneMapper to Familias

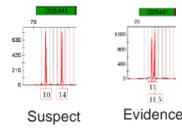
LR binario

LR binario:

- Inclusión: Match
 $LR = 1/2(\text{freq}10)(\text{freq}14)$

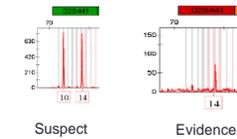


- Exclusión: No match
 $LR = 0/2(\text{freq}11)(\text{freq}11.3)$



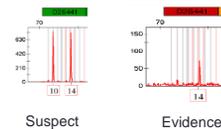
LR binario

- Pero ¿que ocurre cuando tenemos dudas?



- Lo que hacemos habitualmente es intentar análisis adicionales (minis, otros kits, etc)
- Eliminar un marcador dudoso de un perfil es un error grave (es como si el LR = 1 para ese marcador)
- ¿Cua es el valor del LR en esta situación?

LR binario



- Hasta ahora, este tipo de resultados se trataban de la siguiente forma:

- Sospechoso 10-14
- Mancha 14-14

Regla 2p
 $LR = 1/2fr(14)$

$$LR = 1 / 2fr(14)fr(1-14) + fr(14)^2$$

- Ninguno de estos métodos tiene en cuenta la NO coincidencia en el numerador, se continúa asumiendo que hay match bajo Hp
- Buckleton et al., 2006 ha demostrado que la regla 2p no es siempre conservativa (no siempre está a favor de la defensa)

Modelo drop-out & drop-in

- Introducción
- Valor del LR e integración con otros medios de prueba
- Desventajas del LR
- Ventajas del LR
- El LR binario
- Cálculo de LRs en muestras problemáticas
- Otras herramientas útiles:
 - Matches en bases de datos (Mixture deconvolution)
 - GeneMapper to Familias

Cálculo de LR en muestras problemáticas

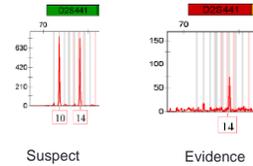
- Nuevas recomendaciones de la ISFG: Gill et al., 2012, FSI Genetics 6: 679-688

The Likelihood Ratio is the preferred approach to mixture interpretation
 DNA Commission 2006

Probabilistic approaches and Likelihood Ratio principles are superior to classical methods
 DNA Commission 2012

Cálculo de LRs en muestras problemáticas

- Este hallazgo no es neutral ($LR \neq 1$)
- En casos de parentesco no eliminamos el marcador que no es compatible, contemplamos la posibilidad de mutación
- Hoy en día se pueden incorporar otro tipo de probabilidades en el LR: probabilidad de drop-out, de drop-in, etc. (el numerador del LR ya no será 1 ó 0, sino un número entre 0 y 1)



Cálculo de LRs en muestras problemáticas

- ISFG recommendations 2012



- Desde el punto de vista del fiscal (H_p) esta situación sólo puede explicarse si el alelo b ha sufrido drop-out y el alelo a no ha sufrido drop-out. No hay drop-in.
- Podemos calcular la probabilidad de que esto ocurra
- No hay necesidad de definir exclusión / inclusión

Cálculo de LRs en muestras problemáticas

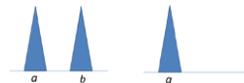
- H_p : la evidencia procede el sospechoso:
 - El alelo a no ha sufrido drop-out: $\Pr(\bar{D})$
 - El alelo b ha sufrido drop-out: $\Pr(D)$
 - No hay drop-in: $\Pr(\bar{C})$

SOSPECHOSO = ab
 EVIDENCIA = aa

$$\Pr(\bar{D}) = 1 - \Pr(D)$$

$$\Pr(\bar{C}) = 1 - \Pr(C)$$

$$\Pr(\bar{D}) \times \Pr(D) \times \Pr(\bar{C})$$



Cálculo de LR en muestras problemáticas

SOSPECHOSO = ab
EVIDENCIA = aa

• Hd: la evidencia procede de otro. Posibilidades:

- Si el otro es aa:
 - Los aa aparecen en la población con frecuencia a^2
 - No ha ocurrido drop-out: $\Pr(\bar{D}_2)$
 - No ha ocurrido drop-in: $\Pr(\bar{C})$
- Si el otro es aQ:
 - Los aQ aparecen en la población con frecuencia $2aq$
 - El alelo a no ha sufrido drop-out: $\Pr(D)$
 - El alelo Q ha sufrido drop-out: $\Pr(D)$
 - No hay drop-in: $\Pr(C)$

aa
No drop-out
No drop-in

aQ
No drop-out de a
Drop-out de Q
No drop-in

Cálculo de LR en muestras problemáticas

SOSPECHOSO = ab
EVIDENCIA = aa

• Hd: la evidencia procede de otro. Posibilidades:

- Si el otro es QQ:
 - Los QQ aparecen en la población con frecuencia q^2
 - Dos alelos idénticos han sufrido drop-out: $\Pr(D_2)$
 - El alelo a es un drop-in: $a \Pr(C)$
- Si el otro es QQ':
 - Los QQ' aparecen en la población con frecuencia $2qq'$
 - Dos alelos distintos han sufrido drop-out: $\Pr(D)^2$
 - El alelo a es un drop-in: $a \Pr(C)$

QQ
Drop-out de 2
alelos idénticos
Drop-in de a

QQ'
Drop-out de 2
alelos distintos
Drop-in de a

Cálculo de LR en muestras problemáticas

SOSPECHOSO = ab
EVIDENCIA = aa

• Hp: la evidencia procede el sospechoso:

$$\Pr(\bar{D}) \times \Pr(D) \times \Pr(\bar{C})$$

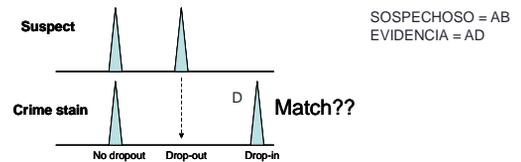
• Hd: la evidencia procede de otro:

$$a^2 \Pr(\bar{D}_2) \Pr(\bar{C}) + 2aq \Pr(\bar{D}) \Pr(D) \Pr(\bar{C}) + q^2 \Pr(D_2) a \Pr(C) + 2qq' \Pr(D)^2 a \Pr(C)$$



LR en mezclas problemáticas: drop-out y drop-in

• Gill and Buckleton, FSI Genet. 4 (2010) 221-227

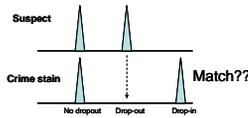


SOSPECHOSO = AB
EVIDENCIA = AD

- Desde el punto de vista del fiscal (Hp) esta situación sólo puede explicarse si el alelo B ha sufrido drop-out y el alelo D es un drop-in
- Podemos calcular la probabilidad de que esto ocurra
- No hay necesidad de definir exclusión / inclusión

LR en mezclas problemáticas: drop-out y drop-in

- Hp: la evidencia procede el sospechoso:
 - El alelo A no ha sufrido drop-out: Pr(D)
 - El alelo B ha sufrido drop-out: Pr(D)
 - El alelo D es un drop-in: Pr(C_D) d
- Hd: la evidencia procede de otro al azar no relacionado
 - Si el otro es AB
 - Si el otro es AD
 - Si el otro es AA
 - Si el otro es DD
 - Si el otro es AQ
 - Si el otros es DQ
 - Si el otro es QQ
 - Si el otro es QQ'



SOSPECHOSO = AB
EVIDENCIA = AD

Cálculo de LR en muestras problemáticas

• <http://www.isfg.org/software>



FORENSIC SOFTWARE RESOURCES

Please note that the software in this section is not validated for casework and no warranty is provided for such a purpose. The ISFG is supporting OPEN SOURCE software projects in forensic statistics, as well as freely accessible software packages offered to our scientific community.

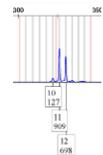
- Forensic DNA Statistics website
- Open source forensim package for R ← Hinda Haned
- MixtureCalc v1.2 Excel sheet
- Mixture Analysis Excel sheet (for deconvolution)
- ikelTLD ver. 4.1 for R
- MixCep mixture separation software for R
- DNAMIX for mixture calculation
- FamLink kinship software
- The bracket script (for replicated STR results)
- DNA commission recommendations 2012. Excel sheet for LR calculation considering dropout and dropin events

Cálculo de LR en muestras problemáticas

- Cálculo de Pr(D): aproximación cualitativa basada en características del perfil → Hinda Haned
- Cálculo de Pr(C): controles negativos de extracción
 - P. Gill et al., An investigation of the rigor of interpretation rules for STRs derived from less than 100 pg of DNA, FSI 112 (2000) 17-40.
 - Si analizamos "n" controles negativos y en ellos aparecen "y" alelos, entonces Pr(C) = y/N
 - Ejemplo: Pr(C) IDF+, 50rflus, extracción orgánica = 0,07 (y=10, N=130)

Sample file	Sample Name	D8S1179	Height	D21S11	Height	D7S820	Height	CSF1PO	Height
16.118.23-01-AGUS_H02.fsa	16.118.23-01-AGUS								
17.88-23-5-11_A03.fsa	17.88-23-5-11		30	54					
18.b3-28-02_B03.fsa	18.b3-28-02								
18.09-03211-BL_0_B03.fsa	18.09-03211-BL_0								
18.09-06302-01_B03.fsa	18.09-06302-01								
20.11-01087-BI-01-06_003.fsa	20.11-01087-BI-01-06								

Cálculo de LR en muestras problemáticas



- Alelos en posición stutter:
 - stutter real
 - stutter + alelo
 - alelo
- Si alelo en posición stutter < % stutter, se considera stutter:
 - Si la muestra de referencia tiene ese alelo: Pr(D)
 - Si la muestra de referencia no tiene ese alelo: no necesario considerar probabilidades adicionales
- Si alelo en posición stutter > % stutter, se considera alelo:
 - Si la muestra de referencia tiene ese alelo: no necesario considerar probabilidades adicionales
 - Si la muestra de referencia no tiene ese alelo: Pr(C)

Modelo drop-out & drop-in

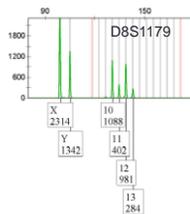
- Introducción
- Concepto de LR
- Desventajas del LR
- Ventajas del LR
- Cálculo de LRs en muestras problemáticas
- Otras herramientas útiles:
 - Matches en bases de datos (Mixture deconvolution)
 - GeneMapper to Familias

Mixture Deconvolution

- Dos posibles abordajes en la interpretación de mezclas:
 - Contemplar todas las posibilidades: «unconstrained conditional method» (ISFG, 2006) o «unrestricted method» (SWGAM, 2010)
 - Contemplar sólo las combinaciones más probables: «constrained conditional method» (ISFG 2006) o «restricted method» (SWGAM 2010)
- Si se contemplan sólo las combinaciones más probables es necesario separar los componentes de la mezcla: «deconvolution»
- Podemos realizar esta tarea con distintas metodologías manuales o con la ayuda de software:
 - M. Bill et al. Forensic Sci. Int. 148 (2005) 181–189 (Pendulum)
 - P. Gill et al. Forensic Sci. Int. 91 (1998) 41–53
 - T. Clayton et al. Forensic Sci. Int. 91 (1998) 55–70
 - M.W. Perlin et al. Am. J. Hum. Genet. 57 (1995) 1199–1210.
 - M.W. Perlin et al. J. Forensic Sci. 46 (6) (2001) 1372–1377

Mixture deconvolution

- Estima de la Proporción de la mezcla (Mixture proportion, M_x) o radio de la mezcla (Mixture ratio, Mr):
 - Componente minoritario con respecto al total



$M_x = 0.25$

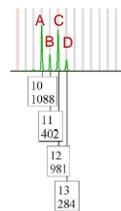
$$M_x = \frac{\phi_a + \phi_b}{\phi_a + \phi_b + \phi_c + \phi_d}$$

$$\frac{11 + 13}{10 + 11 + 12 + 13}$$

$$\frac{402 + 284}{1088 + 402 + 981 + 284}$$

Mixture deconvolution

- El balance de heterocigotos también nos indicará la combinación de genotipos más probable



Posibles combinaciones

$$B/A = 402 / 1088 = 0,37$$

$$B/C = 402 / 981 = 0,41$$

$$D/A = 284 / 1088 = 0,26$$

$$C/A = 981 / 1088 = 0,90$$

$$D/B = 284 / 402 = 0,70$$

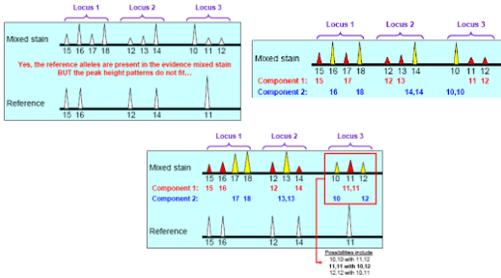
$$D/C = 284 / 981 = 0,29$$

Combinaciones más probables

↓
MAYORITARIO AC
MINORITARIO BD

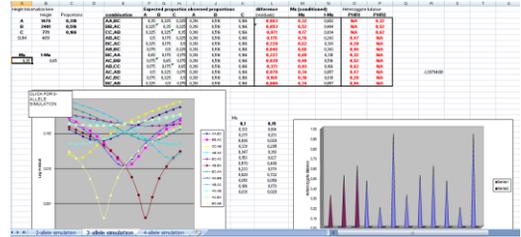
Mixture deconvolution: utilidad

- Comparación con base de datos: es muy útil tener en cuenta las alturas /áreas de los picos para decidir si una coincidencia en la base de datos puede ser real o no



Mixture deconvolution

- Peter Gill, actualizado por Guzmán Moreno



GeneMapper to Familias

- Antonio Vozmediano (CGPC):
 - <http://antonio.scienceontheweb.net/>



Genética Forense Final (GFF)

Fecha última modificación: 07-08-2013
 Software: (06-08-2013)
 Desarrollador: Software Genética Forense Final 2.562.9



- GeneMapper to Familias

GeneMapper to Familias

- Generar un archivo .txt con los genotipos del proyecto en GeneMapper (traditional allele table)
- Preparar un archivo excel con las frecuencias alélicas y los datos del proyecto (templates disponibles para 4 kits)
- Cargar el archivo excel en la aplicación GeneMapper to Familias
- Salvar el archivo txt que GeneMapper to Familias genera para cargarlo directamente en Familias

GeneMapper to Familias

- Disponible versión beta
 - Aplicación
 - Manual (disponible en Inglés sólo)
 - Templates para 4 kits
- Solicitarla a:
 - Antonio Vozmediano: a.vozme@hotmail.es
 - Lourdes Prieto: lourditasmt@gmail.com
- Compatible con Familias 2
- Thore Egeland está implementando algo similar para Familias 3 (aún no disponible)