

# GHEP – Ejercicio de parentesco 2024: nivel avanzado

Autor: Magnus Dehli Vigeland

## Instrucciones generales

Este es un test de opción múltiple que contiene 20 preguntas. En cada pregunta, sólo una de las alternativas es correcta. Puede usar el software que prefiera, pero tenga en cuenta que algunos programas tienen convenciones integradas (por ejemplo, redondeo) que pueden afectar al resultado. Si su respuesta no coincide exactamente con ninguna de las opciones, elija la que más se aproxime.

### Archivos necesarios para realizar el test

- *cousins-data.txt* / *cousins-ibd.txt*. Datos para la Parte II del ejercicio.
- *cousins-data.txt* / *cousins-ibd.txt*. Datos para la Parte II del ejercicio.
- *db.txt*. Frecuencias alélicas para 23 marcadores STR. (Si su software requiere el tamaño de la base de datos, use  $N = 1000$ ).

### A lo largo del ejercicio, asuma las siguientes simplificaciones

- Ausencia de ligamiento entre los marcadores, ausencia de desequilibrio de ligamiento, no hay desviaciones del equilibrio HW.
- Ausencia de drop-out, drop-in, alelos silentes o mutaciones.
- Los fundadores de los pedigrís no son consanguíneos ni están emparentados entre sí.
- La longitud genética total del autosoma (cromosomas 1–22) es de 3391 cM.

### Algunas definiciones

Los alelos homólogos son *idénticos por descendencia* (IBD) si tienen el mismo origen dentro de un pedigrí dado. Los *coeficientes IBD* ( $\kappa_0, \kappa_1, \kappa_2$ ) de dos individuos A y B no consanguíneos, son las probabilidades de compartir respectivamente 0, 1 y 2 alelos IBD en un locus autosómico al azar. Estos coeficientes están relacionados con el coeficiente de parentesco  $\varphi$  mediante la fórmula  $\varphi = \kappa_1/4 + \kappa_2/2$ .

El *triángulo IBD* (Figura 1) es una herramienta útil para visualizar los coeficientes IBD. Nótese que, como  $\kappa_0 + \kappa_1 + \kappa_2 = 1$ , son suficientes dos coeficientes cualesquiera de los tres para deducir el tercero; la elección de los coeficientes  $\kappa_0$  y  $\kappa_2$  en el triángulo es simplemente mi preferencia personal. Para calcular los coeficientes IBD y situarlos en el triángulo IBD puede ser de utilidad la herramienta online [QuickPed](#).

Los coeficientes tradicionales como  $\kappa$  y  $\varphi$  miden la proporción **esperada** de IBD compartido en función del pedigrí. Por el contrario el *parentesco real* (o *parentesco genómico*) entre A y B se refiere a la cantidad **real** de segmentos IBD compartidos entre ellos, como resultado de la recombinación (Figura 2). Llamamos  $(k_0, k_1, k_2)$  a las proporciones reales del autosoma, en términos de longitud genética, en las que se comparten 0, 1 y 2 alelos IBD respectivamente. El *coeficiente de parentesco real* viene dado por  $\varphi_R = k_1/4 + k_2/2$ .

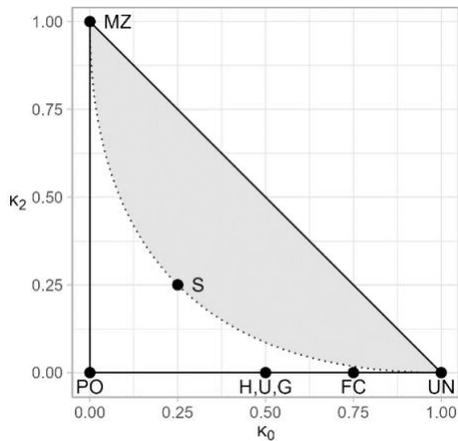


Figura 1. El triángulo IBD. FC=first cousins (primos primeros); G=grandparent-grandchild (abuelo-nieto); H=half sibs (medio hermanos); MZ=monozygous twins (gemelos monocigóticos); PO=parent-offspring (progenitor-hijo); S=full sibs (hermanos completos); U=uncle-nephew (tío-sobrino); UN=unrelated (no emparentado)

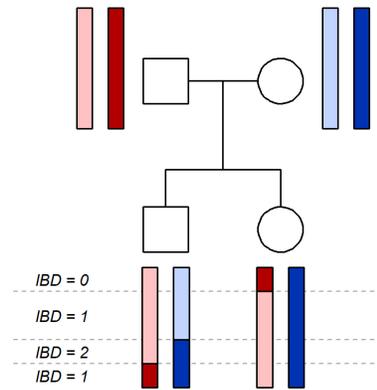


Figura 2. Ejemplo de IBD real compartido entre hermanos. El cromosoma se divide en segmentos con estatus IBD de 0, 1 o 2

## Preguntas

### Parte I: Repaso

Considere una situación en la que los individuos, A y B, se genotipan para un marcador tri-alelético. Los alelos se denominan 1,2,3, y sus frecuencias son  $p_1, p_2, p_3$ , respectivamente.

- Suponga que el marcador se localiza en una región en la que A y B tienen estatus IBD de 0. Si A tiene el genotipo 1/1, el genotipo de B es
  - 2/3
  - 2/2 o 3/3
  - 2/2, 2/3 o 3/3
  - cualquier genotipo excepto 1/1
  - cualquier genotipo
- Suponga que el marcador se encuentra en una región con IBD = 1. Si A tiene el genotipo 1/2, el genotipo de B es
  - 1/3 o 2/3
  - 1/1, 2/2, 1/3 o 2/3
  - cualquier genotipo excepto 1/2
  - cualquier genotipo excepto 3/3
  - cualquier genotipo

En las tres preguntas siguientes, suponga que se desconocen los genotipos de ambos individuos.

- Sabiendo que el marcador está en una región con IBD = 2, la probabilidad de que A y B sean homocigotos para el mismo alelo, es
  - 0
  - $p_1^2 + p_2^2 + p_3^2$
  - $p_1(1 - p_1) + p_2(1 - p_2) + p_3(1 - p_3)$
  - $p_1(1 - p_1)^2 + p_2(1 - p_2)^2 + p_3(1 - p_3)^2$
  - 1

4. Sabiendo que el marcador se encuentra en una región con IBD = 1, la probabilidad de una *coincidencia perfecta* (es decir, *full match*, A y B tienen el mismo genotipo) es
- a) 0
  - b)  $p_1^2 + p_2^2 + p_3^2$
  - c)  $p_1(1 - p_1) + p_2(1 - p_2) + p_3(1 - p_3)$
  - d)  $p_1(1 - p_1)^2 + p_2(1 - p_2)^2 + p_3(1 - p_3)^2$
  - e) 1
5. Sabiendo que el marcador está en una región con IBD = 1, la probabilidad de que sólo compartan un alelo (*partial match*) es
- a) 0
  - b)  $p_1^2 + p_2^2 + p_3^2$
  - c)  $p_1(1 - p_1) + p_2(1 - p_2) + p_3(1 - p_3)$
  - d)  $p_1(1 - p_1)^2 + p_2(1 - p_2)^2 + p_3(1 - p_3)^2$
  - e) 1
-

## Parte II: Un caso de primos

Emma y Carlos se van a casar, pero sospechan que pueden estar emparentados entre sí. Consultan a un genetista que los genotipa para 15 marcadores STR estándar. Los genotipos resultantes se encuentran en el archivo *cousins-data.txt*, junto con la localización física de cada marcador. Las frecuencias alélicas se pueden encontrar en el archivo *db.txt*.  
*Nota: Recuerde que se debe ignorar el ligamiento en los cálculos del LR. Las localizaciones se usarán solo en el Ejercicio 8.*

6. Use los resultados de los 15 marcadores para calcular el LR comparando la hipótesis de que Emma y Carlos son primos primeros, con la hipótesis alternativa de que no estén emparentados. El LR total, aproximado a dos decimales, es
- 0.10
  - 0.75
  - 1.00
  - 3.14
  - 13.14

En un abordaje más exhaustivo, el genetista compara las 5 hipótesis siguientes, ilustradas en la Figura 3:

- H1: first cousins (primos primeros)
- H2: first cousins once removed (tío-sobrino segundo)
- H3: second cousins (primos segundos)
- H4: third cousins (primos terceros)
- H5: unrelated (no emparentados)

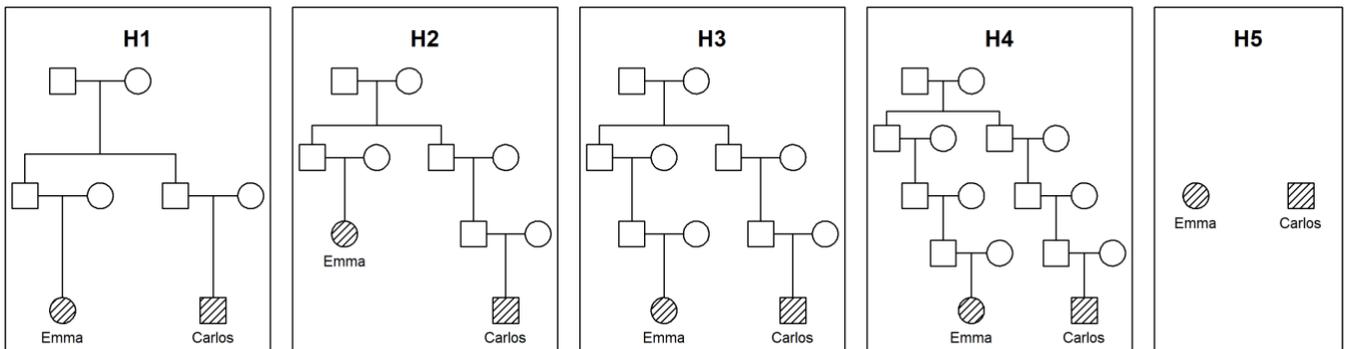


Figure 3. Hipótesis para la Cuestión 7

7. El LR que compara la hipótesis más probable con la segunda hipótesis más probable, es aproximadamente
- 1.00
  - 1.12
  - 1.45
  - 3.61
  - 10.18

No satisfechos completamente con el análisis de STRs, Emma y Carlos se someten a un análisis de secuenciación del genoma completo, y el genetista usa esto para identificar regiones IBD compartidas entre ellos. Los resultados se encuentran en el archivo *cousins-ibd.txt*, reproducido abajo en su totalidad. En todos los segmentos Emma y Carlos comparten un haplotipo, es decir, IBD = 1.

Chr	startMB	endMB	startCM	endCM
3	55.28	98.00	71.71	108.40
4	187.98	189.44	199.80	202.89
5	170.68	180.75	174.66	197.08
9	101.98	112.07	101.00	112.06
11	37.09	98.83	55.18	99.72
11	122.33	131.01	127.17	143.25
12	8.23	15.26	20.31	32.13

8. De los 15 marcadores STR, el número que se encuentra en una región IBD es
- 0
  - 1
  - 2
  - 3
  - 4
9. La proporción  $k_1$  observada del autosoma con status IBD 1, es aproximadamente
- 1.1%
  - 1.8%
  - 2.7%
  - 4.3%
  - 6.3%
10. De los siguientes parentescos, el que presenta un  $\kappa_1$  más próximo al  $\kappa_1$  observado para Emma y Carlos, es
- first cousins (primos primeros, ver Figura 3, H1)
  - first cousins once removed (tío-sobrino segundo, ver Figura 3, H2)
  - second cousins (primos segundos, ver Figura 3, H3)
  - third cousins (primos terceros, ver Figura 3, H4)
  - unrelated (no emparentados, ver Figura 3, H5)
11. De acuerdo a la herramienta DNA painter, <https://dnainter.com/tools/sharedcmv2>, la partición IBD entre Emma y Carlos *no es compatible* con (es decir, está fuera del intervalo reportado de)
- first cousins (primos primeros, ver Figura 3, H1)
  - first cousins once removed (tío-sobrino segundo, ver Figura 3, H2)
  - second cousins (primos segundos, ver Figura 3, H3)
  - third cousins (primos terceros, ver Figura 3, H4)
  - varios de los anteriores
12. La Figura 4 muestra la distribución de segmentos IBD de 200 simulaciones de primos primeros, segundos y terceros. Basándose en este gráfico, los datos observados son solo compatibles (en el sentido de estar dentro de la elipse del 95% de los datos) con
- primos primeros (first cousins, FC)
  - primos segundos (second cousins, SC)
  - primos terceros (third cousins, TC)
  - primos primeros y segundos (first and second cousins)
  - primos segundos y terceros (second and third cousins)

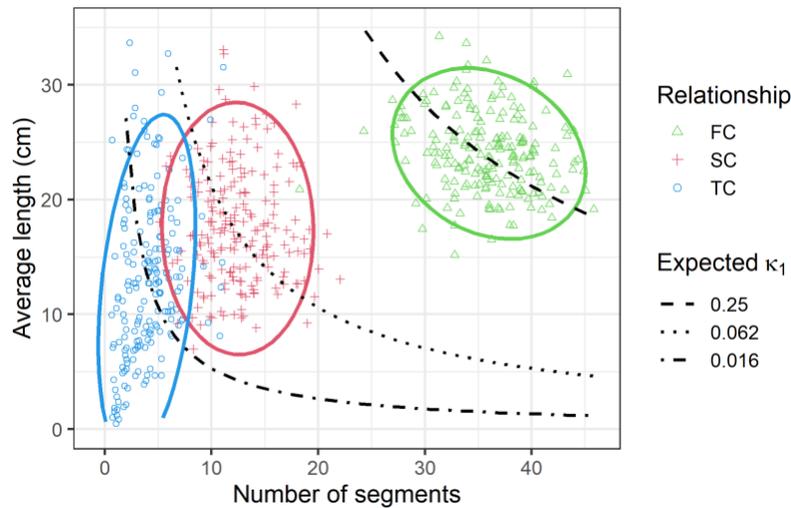


Figure 4. Distribuciones IBD para primos 1º, 2º y 3º

### Parte III: Un caso de hermanos

Este caso trata de 4 hermanos varones, denominados A, B, C y D. Se cree que los cuatro son hermanos de madre, pero las paternidades no están claras. Los resultados genotípicos de A, B, C, D para 23 marcadores forenses se pueden encontrar en el archivo *sibship-data.txt*, con las frecuencias alélicas en el archivo *db.txt* como anteriormente.

13. El LR tras comparar las hipótesis de que A y B son hermanos completos frente a que son medio hermanos, es aproximadamente
  - a) 0.68
  - b) 1.00
  - c) 431.47
  - d) 133506.3
  - e) Ninguno de los anteriores
  
14. El LR resultante de la comparación de que C y D son hermanos completos frente a que son medio hermanos, es aproximadamente
  - a) 0.68
  - b) 1.00
  - c) 431.47
  - d) 133506.3
  - e) Ninguno de los anteriores

Considere ahora las siguientes suposiciones:

- i) los cuatro tienen la misma madre
- ii) C y D son hermanos completos
- iii) para cada par de individuos entre A,B,C,D, las relaciones de parentesco son o bien de medio hermanos o de hermanos completos, sin otro parentesco adicional ni endogamia en el pedigrí

15. El número de hipótesis posibles (pedigríes) que relacionan a A,B,C,D es
- 5
  - 6
  - 7
  - 8
  - 9
16. Sea  $H^*$  la hipótesis de que los 4 son hermanos completos. Suponiendo una probabilidad a priori equiprobable para el conjunto de hipótesis de la pregunta anterior, la probabilidad a posteriori de  $H^*$  dados los datos genéticos, es aproximadamente
- 0.00
  - 0.01
  - 0.73
  - 0.99
  - 1.00

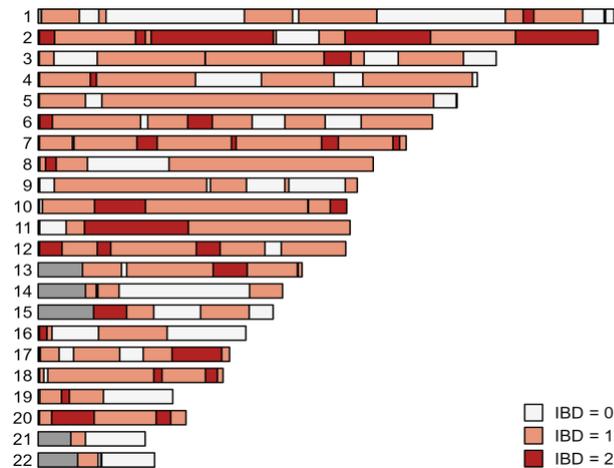


Figure 5. Segmentos de identidad por descendencia entre los supuestos hermanos A y B.

Para investigar mejor la relación de parentesco entre A y B, se lleva a cabo la secuenciación del genoma completo, y los segmentos IBD determinados se muestran en la Figura 5. Las coordenadas se pueden encontrar en el archivo *sibs-ibd.txt*.

17. En términos de longitud genética, la proporción  $k_1$  del autosoma con estatus IBD igual a 1, es
- 0.25
  - 0.30
  - 0.45
  - 0.50
  - 0.60
18. En el triángulo IBD de la Figura 6, el parentesco real entre A y B se corresponde con el punto
- P1
  - P2
  - P3
  - P4
  - P5

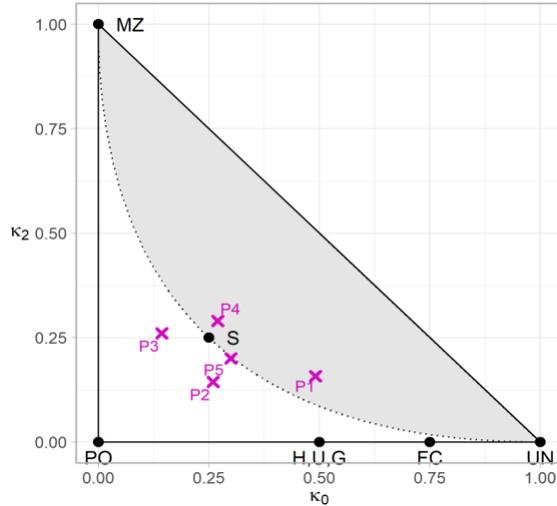
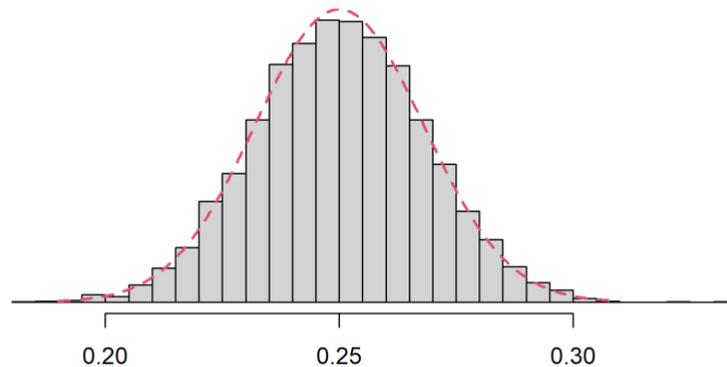


Figure 6. Triángulo IBD con las alternativas para la relación de parentesco real entre A y B.

19. Basándose en los segmentos IBD, el coeficiente de parentesco real  $\varphi_R$  entre A y B es aproximadamente
- 0.14
  - 0.19
  - 0.22
  - 0.25
  - 0.26

El histograma de la Figura 7 muestra el coeficiente de parentesco real de 5000 pares simulados de hermanos completos, el cual se aproxima cercanamente a una distribución normal de media  $\mu = 0.25$  y desviación estándar  $\sigma = 0.018$  (curva de puntos roja).

20. Comparado con la aproximación normal para hermanos completos, el  $\varphi_R$  observado está en el
- percentil 0º
  - percentil 5º
  - percentil 10º
  - percentil 15º
  - percentil 20º



Simulaciones de parentesco real entre hermanos completos

**BUENA SUERTE!**

Responsable del desarrollo del ejercicio:

Magnus Dehli Vigeland

[magnusdv@gmail.com](mailto:magnusdv@gmail.com)

Traducción:

Lourdes Prieto ([lourditasmt@gmail.com](mailto:lourditasmt@gmail.com))

Comité Ejecutivo del GHEP-ISFG

[info@ghép-isfg.org](mailto:info@ghép-isfg.org)