

Experimentos genéticos en laboratorio

Enunciado de la Práctica de PRO2
Primavera 2017

29 de marzo de 2017

1. Introducción

Por claridad recordamos algunos conceptos básicos de la herencia biológica, aunque no todos se usan en esta práctica:

- Rasgo o carácter hereditario: es una característica morfológica, estructural o fisiológica presente en un ser vivo y transmisible a la descendencia.
- Fenotipo: conjunto de rasgos o características que muestra un individuo de una especie.
- Cromosoma: Orgánulo en forma de filamento que se halla en el interior del núcleo de la célula y que contiene el material genético. El número de cromosomas es constante para las células de una misma especie.
- Gen o factor hereditario: es la unidad de transmisión genética (en realidad un fragmento de ADN localizado en un cromosoma) que determina a priori un carácter hereditario; posteriormente, dicho carácter puede ser alterado por factores ambientales y de desarrollo del individuo.
- Genotipo: conjunto de genes que posee un individuo.
- Alelo: cada una de las posibles formas de un gen; en genética mendeliana, cada gen sólo tiene dos posibles alelos, que representaremos por los valores 0 y 1.
- Herencia dominante: cuando en un par de genes del mismo tipo existe un alelo dominante (digamos el 1) que no deja manifestarse al otro, que se denomina recesivo (digamos el 0, respectivamente).

- Herencia intermedia: cuando en un par de genes del mismo tipo pero alelos diferentes (un 0 y un 1) el rasgo resultante no corresponde a los rasgos asociados con los alelos “puros” 0 y 1, sino a un resultado “intermedio” (por ejemplo, los colores puros de una flor pueden ser “rojo” y “blanco”, y un resultado intermedio sería color “rosa”).
- Relación entre fenotipo y genotipo: si nos abstraemos de las modificaciones posteriores causadas por factores ambientales, cada rasgo del fenotipo depende de un único par de genes del mismo tipo que aparece en todas las células del organismo, concretamente depende de la combinación de sus dos alelos (11, 10, 01 o 00) y del tipo de herencia (dominante o intermedia; en ambos casos, la presencia de dos alelos iguales produce el rasgo “puro” asociado a dicho alelo).

1.1. Definición genética de una especie

Consideremos una especie con reproducción sexual y genética mendeliana. Las células de cada individuo u organismo de dicha especie contienen un par de cromosomas sexuales y N pares de cromosomas “normales” (realmente llamados autosómicos en genética).

Los cromosomas contienen una secuencia lineal de genes consecutivos. Para cada par p_i de cromosomas “normales” ($1 \leq i \leq N$), los dos cromosomas tienen la misma longitud l_i y sus genes están completamente alineados, de manera que el mismo tipo de gen aparece en la misma posición en los dos cromosomas del par. Por lo tanto el par de cromosomas p_i está compuesto por dos cromosomas c_{i1} y c_{i2} y cada uno de estos dos cromosomas consta de una secuencia de l_i genes. Se podría designar un gen concreto de un cromosoma concreto de un par concreto mediante la notación c_{ij}^k ($1 \leq i \leq N$, $j \in \{1, 2\}$, $1 \leq k \leq l_i$). Los genes c_{i1}^k y c_{i2}^k son del mismo tipo.

El par de cromosomas sexuales se distingue del resto en que o bien está formado por dos cromosomas del mismo tipo (llamado X y con longitud l_X) o bien está formado por un cromosoma X y un cromosoma diferente (llamado Y y con una longitud l_Y , normalmente $l_Y < l_X$). En el caso de tener un par XX el sexo del individuo es femenino y en el caso de tener un par XY el sexo es masculino. Los cromosomas X e Y tienen una parte común de genes alineables de longitud l_0 , mientras que el resto de sus genes son propios de cada cromosoma X o Y y se dice que están ligados al sexo.

1.2. Reproducción sexual

Para la reproducción sexual se necesita un organismo con cromosomas sexuales XY , que llamaremos padre, y uno con cromosomas sexuales XX , que llamaremos madre. Tanto en el padre como en la madre, la célula sexual correspondiente (espermatozoide y óvulo) sólo contiene uno de los dos cromosomas de cada par, elegido aleatoriamente con igual probabilidad entre los dos de cada par. Como estamos haciendo experimentos en un laboratorio, para cada uno de las pares de cromosomas, escogeremos uno concreto.

Durante la fecundación, cuando el espermatozoide y el óvulo se unen, los cromosomas correspondientes, los que tienen el mismo identificador i entre 0 y N , se emparejan. Pero además, para cada par de cromosomas, se produce un cruzamiento entre los cromosomas del padre y de la madre. En el cruzamiento, dado un par de cromosomas p_i “normales”, se selecciona aleatoriamente un punto de corte (entre dos genes) k entre 0 y l_i , de manera que, si antes del cruzamiento teníamos dos cromosomas c_{i1} y c_{i2} , el resultado del cruzamiento son dos cromosomas c'_{i1} y c'_{i2} , tales que c'_{i1} tiene los mismos alelos que c_{i1} en las posiciones entre 1 y k y los mismos alelos que c_{i2} en las posiciones entre $k+1$ y l_i , mientras que c'_{i2} tiene los mismos alelos que c_{i2} en las posiciones entre 1 y k y los mismos alelos que c_{i1} en las posiciones entre $k+1$ y l_i . Los cromosomas c'_{i1} y c'_{i2} formarán por tanto el par de cromosomas p_i en el organismo resultado de la reproducción.

En el caso del par de cromosomas sexuales, la madre siempre transmite un cromosoma X y el padre puede transmitir un cromosoma X o Y , lo que determina el sexo del nuevo organismo (femenino o masculino, respectivamente). Además el cruzamiento sólo se realiza en la parte común a ambos cromosomas (entre 1 y l_0 , con el punto de corte k entre 0 y l_0) mientras que el resto de genes de los cromosomas c_{01} y c_{02} pasan directamente a los cromosomas c'_{01} y c'_{02} . De nuevo, como estamos haciendo experimentos en un laboratorio, para cada uno de las pares de cromosomas, escogeremos el punto de corte.

1.3. Arbol genealógico de un individuo

Dado un individuo concreto, definimos su árbol genealógico como aquel árbol binario que tiene el identificador del propio individuo como raíz, los identificadores del padre (izquierda) y de la madre (derecha) en el primer nivel del árbol, los identificadores de los abuelos en el segundo nivel (de izquierda a derecha primero los abuelos paternos y después los maternos) y así sucesivamente hasta representar los identificadores de todos los antecesores.

sores del individuo registrados en el sistema.

Diremos que un árbol binario no vacío a_p es un *árbol parcial* de un árbol genealógico a_g si la estructura de nodos de a_p coincide con la estructura de algún subconjunto conexo de nodos de a_g que incluya a su nodo raíz y, además, los identificadores de los nodos de a_p son los mismos que los de los nodos correspondientes de a_g . Un ejemplo de árbol parcial de un árbol genealógico se da cuando sólo se conocen algunos de los antecesores del individuo.

Dado un árbol genealógico a_g y un árbol parcial del mismo, a_p , la operación de completar el árbol parcial a_p producirá una réplica estructural de a_g , llamémosle a_c , en la que a los identificadores de los nodos presentes en a_g pero no en a_p se les añadirá alguna marca especial con respecto a su valor en a_g , mientras que el resto de nodos de a_c tendrá el mismo identificador que en a_p y en a_g .

2. Sistema a simular mediante la práctica

Inicializaremos el sistema con los datos genéticos de la especie: los valores N , l_i ($0 \leq i \leq N$), l_X y l_Y . Seguidamente leeremos los datos de una población inicial de organismos de la especie: el número inicial de individuos sin ascendientes M , un nombre para cada individuo que lo identificará (string con letras y dígitos) y la composición genética del individuo: una X o una Y y el contenido de sus cromosomas. Primero se leeran los cromosomas sexuales y el primer cromosoma del par será siempre un cromosoma X , mientras que el segundo puede ser un cromosoma X o Y . Los M nombres leídos serán diferentes. Los M individuos no tienen ningún parentesco.

A partir de aquí se leerán una serie de comandos que realizarán operaciones de modificación o consulta del sistema. Cada comando se escribirá junto con su resultado caso de haberlo o mensaje de error correspondiente. Para ver el formato exacto hay que consultar el juego de pruebas público. Los comandos son los siguientes:

1. Añadir un nuevo individuo sin ascendientes. Los datos son los mismos que para cualquiera de los M individuos iniciales sin ascendientes. Si el sistema ya contiene un individuo con el nombre dado se escribirá el mensaje **error**.
2. Reproducir: Dados tres nombres, el primero correspondiente a la madre, el segundo al padre y el tercero el que se pondría al posible descendiente, se intentará realizar su reproducción. Si alguno de los dos

primeros nombres no está entre los individuos del sistema (no existe) o el tercero sí está (ya existe) se escribirá el mensaje **error**. Si los dos primeros individuos no pueden reproducirse se escribirá el mensaje **no es posible reproduccion**. Para poder reproducirse el primer individuo ha de tener cromosomas sexuales XX , el segundo cromosomas sexuales XY , no ser hermanos ni por parte de padre ni de madre, y no ser ninguno de los dos antecesor del otro. La operación de reproducción sexual producirá un nuevo individuo cuyo nombre será el tercero leído. El genotipo del descendiente se calculará a partir de los genotipos de sus padres indicando qué cromosomas de cada par (el primero o el segundo) forman el óvulo y el espermatozoide y cuáles son los puntos de corte para cada par de cromosomas del descendiente. El sistema registrará la relación de parentesco entre los padres y el nuevo individuo.

3. Dado un individuo, escribir su árbol genealógico ordenado por niveles, e indicando de alguna manera, que concretaremos más adelante en el juego de pruebas público, los individuos que forman cada nivel. Dentro del mismo nivel, el padre y la madre asociados (es decir, que hayan formado pareja reproductora) aparecerán consecutivos, primero el padre y después la madre. Si el individuo no está en el sistema se escribirá el mensaje **error**.
4. Dado un árbol genealógico (normalmente incompleto), comprobar si es árbol parcial del árbol genealógico de un individuo y, si lo es, completar el árbol identificando y marcando los nombres que falten (la manera de marcar los nombres se indicará en el juego de pruebas público); en caso contrario (si no es árbol parcial), se escribirá el mensaje **no es arbol parcial**.
5. Escribir un listado de todos los individuos, ordenado alfabéticamente por nombre, en el que además de su nombre, se indique cuál es su sexo y cuáles son los nombres de su padre y de su madre.
6. Dado el nombre de un individuo, escribir su genotipo, por orden creciente de identificador de par de cromosomas (primero el par de cromosomas sexuales y luego los N pares de cromosomas normales); dentro de cada par, primero se escribirá el primer cromosoma y luego el segundo; dentro de cada cromosoma, se escribirán los valores de los genes por orden creciente de posición. Para cada par de cromosomas, el orden entre los dos cromosomas depende de si el individuo tiene pro-

genitores conocidos o no. Si se trata de un individuo sin ascendientes, el orden de escritura será el mismo que el orden en que se leyeron los dos cromosomas. Si se trata de un individuo con ascendientes, primero se escribirá c'_{i1} y luego c'_{i2} , resultado del cruce entre c_{i1} y c_{i2} , donde c_{i1} es el cromosoma proveniente del óvulo (célula de la madre) y c_{i2} el cromosoma proveniente del espermatozoide (célula del padre). Si el individuo no está en el sistema se escribirá el mensaje **error**.

7. Acabar. No produce ningún resultado. Se acaba la ejecución del programa.