



CONFERENCIA NO. 9

## DIFERENCIACION GEOGRAFICA Y FLUJO GENETICO

Jean Pierre Dujardin

### DISTANCIA GENETICA DE NEI (Ds)

$$D_s = -\ln(\text{Identidad})$$

$$2 \cdot n \cdot l \cdot t = -\ln(I) \quad ; \quad \text{donde: } n : \# \text{ de aa de la proteína}$$

$l$  : tasa de mutación anual

$t$  : tiempo de separación de las poblaciones

La identidad ( $I$ ), es la probabilidad de encontrar las poblaciones sin diferencia. Lo que se pretende conocer generalmente es el valor de  $t$ , ya que los otros parámetros pueden ser medidos (conocidos). Para encontrar la probabilidad de que 2 poblaciones sean idénticas puede usarse la ley de probabilidad de dos eventos independientes. La medida de  $I$  sería la probabilidad de encontrar dos genes idénticos en estas poblaciones.

Cuando las poblaciones son idénticas, no siempre la sumatoria de las probabilidades de los genes idénticos es igual; para esto Nei propone la siguiente fórmula :

$$(\sum x \cdot y) \div (\sqrt{\sum x^2 \cdot \sum y^2})$$

con  $x$  = frecuencia de un gen en la población  $x$ ;  $y$  = frecuencia del mismo gen en la población  $y$ .

Esta determina la proporción de probabilidad de identidad entre poblaciones  $y$  dentro de estas; con una variación entre 0 y 1.

Cuando la identidad es nula, la distancia es infinita (la distancia de Nei no es euclidiana), es decir, especies que no presentan nada en común.

Por estudios moleculares podemos conocer  $n$  e  $l$  para estimar finalmente el  $t$  como una medida para buscar el tiempo de separación entre 2 especies o poblaciones locales.

## CONCEPTOS

**Efecto Wahlund:** Cuando se suman poblaciones en equilibrio, se tiene una población total en desequilibrio, siendo éste falta de heterocigotas. Este déficit de heterocigotas se usa para hacer medición del grado de diferencia genética entre estas poblaciones, o lo que es lo mismo, del grado de endogamia (inbreeding) dentro de las poblaciones.

Hipótesis que explican el desequilibrio de Hardy-Weinberg en una población:

- Mala lectura de los geles: Interpretación correcta o incorrecta de las variaciones en los geles; cuando se presentan dudas es necesario repetir la corrida de las muestras.

En la práctica (triatomíneos), se puede intentar hacer cruces de parejas que presentan la diferenciación y ver si en la descendencia se nota la heterocigosidad

- Cantidad de individuos utilizados : Lo ideal es tener 5 o más efectivos esperados.
- Alelo nulo: Es un alelo (mutación) que existe en la población y cuando el individuo es homocigoto para el, no hay actividad enzimática de la proteína.

El heterocigoto presenta actividad similar a la del homocigoto que tiene los alelos funcionales, lo que puede presentarse como un caso de dominancia aunque menos intensa.  $Aa$  y  $AA$  son idénticos como fenotipos, por lo que se cuentan todos, como homocigotos, provocando así déficit de heterocigotos en las frecuencias esperadas.

- Estructuración: Es el efecto Wahlund (subdivisión temporal o espacial).
- Selección: Produce más frecuentemente un exceso de heterocigotos.
- Especies gemelas: Se piensa haber muestreado una misma especie y en realidad hay dos especies gemelas (más frecuente en dípteros que

en triatomíneos). Para tener seguridad en esta interpretación es necesario encontrar:

### 1. Déficit de heterocigotos

2. Alelos alternos: Es otra forma de presentar el déficit de heterocigotos. Son individuos que no tienen la misma migración, pero entre ellos no se encuentra ninguna forma heterocigota. Con este alelo alternativo se identifican las 2 especies.

3. Distancia genética de Nei  $> 0,10$  (hasta la fecha la distancia genética máxima encontrada entre poblaciones intraespecíficas es de  $0,05$ ). Pero hay que saber que la distancia de Nei entre especies puede ser nula o muy baja.

- Muestreo de poblaciones mezcladas (migración): Al presentar frecuencias diferentes, hay déficit de heterocigotos. Pero este factor es poco probable ya que en una sola generación todo vuelve al equilibrio.
- Mutación: Es descartable por su baja probabilidad.

## F- STATISTICS

Propone tener una medición del grado de equilibrio dentro de la población total (FIT), dentro de cada subpoblación (FIS) y del grado de diferenciación entre subpoblaciones (FST). Pueden calcularse tres factores que tienen una relación entre ellos, así:

FIT: Indica lo que pasa con la población total (desequilibrio global).

FIS: Mide el grado de desequilibrio dentro de cada subpoblación.

FST: Grado de diferencia entre las frecuencias de cada subpoblación (indicio de distancia génica).

$$(1-FIT) = (1-FIS).(1-FST)$$

Para medir estos 3 parámetros, se necesita medir: heterocigosidad observada en cada subpoblación (promedio), heterocigosidad esperada (promedio), y lo esperado en la población total.

En ensayos realizados con *T. infestans* y *R. prolixus*, se tiene que en el primero las subpoblaciones son muy diferenciadas; para la segunda las poblaciones no presentan siempre diferencias génicas. Las estructuras poblacionales de estas dos especies son muy diferentes.

### **RAPD (Random Amplified Polimorphic DNA)**

Con solo esta técnica no se puede hacer una estimación confiable de frecuencias génicas por lo cual no puede aplicarse la ecuación de H-W, ni la "F" statistics, ni la distancia de Nei.

Es práctico en los RAPDs, reconocer los niveles de migración de las bandas y transformar esta información en tablas para buscar coeficientes entre individuos y llevar esta información hasta árboles UPGMA (por ej.), el cual es construido con base en una matriz de distribución.

### **MORFOMETRIA**

Se usa como método de estimación de diferenciación poblacional o específica. Cuando se tienen varias medidas como variables diferentes, pueden compararse las poblaciones por medio de dos enfoques:

1. Análisis univariado: Se analiza cada variable separadamente y se comparan las poblaciones. Para las comparaciones se tienen 2 tipos de pruebas paramétricas y no paramétricas. Las primeras suponen igual varianza y una distribución normal, (se aconseja transformar los datos a log). Las segundas no necesitan las condiciones de las pruebas paramétricas, no importan los valores o su transformación, hace un estudio sobre los rangos de los valores.
2. Análisis multivariado: Generalmente los caracteres métricos son muy relacionados entre si, hay correlación estadística y biológica entre las medidas. El interés del análisis multivariado, es sacar la correlación estadística para obtener nuevas variables no correlacionadas entre si (al menos estadísticamente).

Estas nuevas variables se llaman factores y son una combinación lineal de todas las variables utilizadas. En este análisis hay simplificación del número de variables y generalmente se trabaja con dos factores. El análisis multivariado puede mostrar cual es el porcentaje explicado por cada factor.

**Análisis en componentes principales:** Se introducen los datos sin la información de la pertenencia de cada individuo a una población.

**Análisis discriminante:** Se introducen los datos con la información de su población y el análisis mejora la calidad de los factores para permitir la mejor separación entre las poblaciones.

El análisis multivariado presenta las siguientes ventajas: Permite sacar las correlaciones, simplificar el número de variables, presentar los resultados en forma gráfica y por último se pueden sacar las diferencias de tamaño para estudiar la forma.

Para reducir el efecto de la correlación biológica entre variables, es aconsejable utilizar medidas de órganos diferentes del cuerpo (cabeza y ala, por ejemplo).