

Una guía para

**GenALEX6**®

# Análisis genéticos en Excel

Escrito por Nicola S. Flanagan

Traducido por Angela Mina-Vargas



<http://www.anu.edu.au/BoZo/GenALEX/>

# Tabla de contenido

<i>Tabla de contenido</i> .....	1
<i>Introducción</i> .....	3
GenAIEx 6 .....	3
Registro de usuarios y cita de GenAIEx 6 .....	3
Instalación de GenAIEx 6.....	4
Sobre este Manual .....	4
Estilos del texto en el Manual .....	6
El Manual en Español .....	6
Descargo de responsabilidad .....	6
<i>El Entorno de GenAIEx</i> .....	7
Resumen .....	7
Limite de Datos .....	7
Datos de Ingreso (Input) .....	7
Resultados (Output) .....	8
<i>Formato de los datos para GenAIEx</i> .....	9
Formato Numérico .....	9
Datos faltantes.....	9
Marcas de las muestras .....	9
Estructura de los datos.....	9
Parámetros y marcas de datos.....	10
Formato de los datos.....	11
<i>Opciones de GenAIEx para Formatear Conjuntos de Datos.</i> .....	18
Template (Modelo) .....	18
Create (Creando) .....	19
Parameters (Parámetros).....	20
Data (Datos) .....	21
<i>Procedimientos Estadísticos</i> .....	22
Cuadro Normal de Parámetros de Datos .....	22
Frequency (Frecuencia) .....	23
HWE (Equilibrio Hardy-Weinberg) .....	28
Distance (Distancia).....	29
Genetic Distance (Distancia Genética).....	29
Geographic Distance (Distancia Geográfica) .....	32
AMOVA.....	36

<b>Mantel</b> .....	<b>40</b>
<b>PCA</b> .....	<b>41</b>
<b>Assignment (Asignación)</b> .....	<b>43</b>
<b>Pop Assign (Asignación de poblacion)</b> .....	<b>43</b>
<b>Sex Bias (Efecto Sesgado del sexo)</b> .....	<b>46</b>
<b><i>Procedimientos estadísticos avanzados</i></b> .....	<b>47</b>
<b>Spatial Autocorrelation (Autocorrelación Espacial)</b> .....	<b>47</b>
<b>Single Pop...(Población Única)</b> .....	<b>48</b>
<b>Multiple loci... (Loci Múltiples)</b> .....	<b>52</b>
<b>Multiple Pops... (Poblaciones múltiples)</b> .....	<b>54</b>
<b>Multiple Dclass ... (Clases de distancia múltiples)</b> .....	<b>56</b>
<b>Pops as Dclass ... (Poblaciones como clases de distancia)</b> .....	<b>59</b>
<b>2D LSA... (Anàlisis espacial local en 2D)</b> .....	<b>61</b>
<b>NN Dist... (Distancia de Vecinos más Cercanos)</b> .....	<b>65</b>
<b>Relatedness (Relaciones genéticas o parentesco)</b> .....	<b>66</b>
<b>Pairwise (En Parejas)</b> .....	<b>66</b>
<b>Pops Mean (Promedio poblacional)</b> .....	<b>68</b>
<b>Multilocus (Loci múltiples)</b> .....	<b>70</b>
<b>Clonal (Análisis de clones)</b> .....	<b>72</b>
<b><i>Edición de Datos en Bruto</i></b> .....	<b>75</b>
<b>Import Data (Importar datos)</b> .....	<b>75</b>
<b>Raw data (Datos crudos)</b> .....	<b>81</b>
<b>Merge Pops (fusión de poblaciones)</b> .....	<b>87</b>
<b>Edit Raw Data (Editar datos crudos)</b> .....	<b>92</b>
<b>Export Data (Exportar Datos)</b> .....	<b>92</b>
<b>Graph (Gráficas)</b> .....	<b>93</b>
<b>Stats (Estadísticas)</b> .....	<b>93</b>
<b>Options (Opciones)</b> .....	<b>93</b>

## Introducción

### GenAIEx 6

GenAIEx 6 está escrito en Visual Basic para Aplicaciones (VBA) con Excel. En PC es compatible con Excel 97 en adelante con Windows 95, 98, 2000 y XP. En Macintosh es compatible con Excel 98 en adelante. GenAIEx 6 ha sido extensamente evaluado y aplicado en Excel 2002 y Excel 2003 bajo Windows XP, y en Excel X y Excel 2004 en Macintosh OS X.



#### Dr Rod Peakall

School of Botany and Zoology  
The Australian National University  
Canberra ACT 0200 Australia

#### Dr Peter Smouse

Dept. of Ecology, Evolution & Natural  
Resources  
Rutgers University, Cook College  
New Brunswick NJ 08901-8551 USA



© Peakall and Smouse, 1996-2006  
Population Genetic Analysis

Peakall R and Smouse PE (2006) GENALEX 6: genetic analysis in Excel.  
Population genetic software for teaching and research. *Molecular Ecology Notes* 6, 288-295.

The Australian National University, Canberra, Australia.  
<http://www.anu.edu.au/BoZo/GenAIEx/>

Logo created by Sasha Peakall

## Registro de usuarios y cita de GenAIEx 6

### Por favor regístrese

El sitio Web de GenAIEx (<http://www.anu.edu.au/BoZo/GenAIEx/>) proporciona un formulario opcional de registro, que le invitamos a completar. Con este registro usted puede estar seguro de recibir por correo electrónico información sobre errores corregidos y sobre nuevas versiones.

### Por favor citar GenAIEx como se indica a continuación:

Peakall, R. and Smouse P.E. (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*. 6, 288-295.

Este artículo está disponible en:

<http://www.blackwell-synergy.com/doi/abs/10.1111/j.1471-8286.2005.01155.x>

Por favor lea el artículo en conjunto con esta guía y los otros documentos de soporte para GenAIEx.

## Instalación de GenAIEx 6

---

La aplicación de GenAIEx esta proporcionada como un complemento adicional en Excel, un modulo compilado y el menú de **GenAIEx** asociado, que puede ser configurado para cargar automáticamente cuando se inicia Excel, o cargado solo cuando sea necesaria.

### **Instalación para carga automática cuando se inicia Excel.**

1. Copie el ejecutable (llamado GenAIEx.xla) en la ubicación que desea en su computador. Es preferible que cree un fólder exclusivo para el programa.
2. Abra el Excel, y seleccione (*Add-Ins*) desde el menú de **Herramientas (Tools)**. Para localizar el complemento adicional GenAIEx (Add-in) pulse el botón de búsqueda (*Browse*). Una vez que el GenAIEx aparezca cargado en la lista de dialogo, pulse el botón de verificación para GenAIEx, luego pulse aceptar (*Ok*). En unos segundos el menú de **GenAIEx** aparecerá en la barra de menú de Excel, justo antes del menú de ayuda.

*Recomendación: GenAIEx puede ser instalado y desinstalado para carga automática desde el menú de **GenAIEx**. Ubique el submenú *Options* y allí seleccione *Install*. Una vez GenAIEx esta instalado de forma automática siempre aparecerá en la lista de complementos adicionales de Excel. Se recomienda enfáticamente que si hay versiones anteriores de GenAIEx éstas sean removidas. Esto debe hacerse de forma manual: enviándolas a la papelera, y seleccionándolas y eliminándolas en la lista de complementos de Excel, según las instrucciones de Excel. La opción *List Add-Ins* proporcionada por **GenAIEx: Options**, facilita la identificación de versiones anteriores.*

### **Descargue GenAIEx solo como usted lo requiere.**

1. Copie el ejecutable (llamado GenAIEx.xla) en la ubicación que desea en su computador. Es preferible que cree un fólder exclusivo para el programa.
2. Abra Excel. Seleccione *abrir* desde el menú de **Archivo**, localice el GenAIEx ejecutable, luego pulse el botón *Aceptar*. Dependiendo del nivel de la configuración de su programa de Excel, Excel puede advertirle que GenAIEx contiene macros. Pulse *habilitar* macros para continuar. En unos segundos el menú de **GenAIEx** aparecerá en la barra de menú de Excel, justo antes del menú de **Ayuda**.

Alternativamente, usted puede abrir GenAIEx y Excel simultáneamente pulsando directamente en el ejecutable de GenAIEx.

## Sobre este Manual

---

Esta guía asume un nivel de conocimientos previos de genética de poblaciones equivalente a la información que tiene un estudiante graduado. Los cálculos realizados por GenAIEx se detallan por separado en el Apéndice 1: Los métodos y estadísticas en GenAIEx 6, fueron compilados por Rod Peakall y Peter Smouse. Para más información sobre los cálculos, se aconseja a los usuarios consultar el Apéndice 1, junto con las referencias proporcionadas en el.

En esta guía se asume que los usuarios de GenAIEx están familiarizados con los procedimientos normales de operación de un sistema computacional (PC o Macintosh). También asume que hay familiaridad con el entorno de trabajo en Excel, incluyendo cómo crear y manipular nuevos cuadernos, nuevas hojas de cálculo dentro de los cuadernos y cómo entrar y manipular los datos.

Un objetivo principal en GenAIEx 6, ha sido hacer la interacción con los usuarios de GenAIEx tan directa como sea posible. Así, donde ha sido viable, se han estandarizado los cuadros de dialogo de GenAIEx. Opciones que requieren mayor explicación, se describen con más detalle en esta guía.

La guía busca proporcionar:

1. Una descripción del tipo de datos que se manejan y el formato apropiado en que deben estar.
2. Una descripción de la interfase de usuario, e instrucciones paso a paso por realizar los análisis individuales.
3. Una descripción de las diferentes opciones de análisis, donde es pertinente, con consejos para ayudar a usar más las funciones de GenAIEx.

Las ilustraciones de los cuadros de dialogo usados en esta guía fueron extraídos usando GenAIEx en Excel X, en Mac OS 10.2. La interfase puede aparecer ligeramente diferente cuando se utilizan otras versiones de Excel o otros sistemas operativos.

## **Apéndices**

Varios apéndices compilados por los autores del programa, proporcionan información adicional a los usuarios de GenAIEx. De ser necesario, estos deben citarse bajo la cita del programa. Estos apéndices están disponibles en la página Web de GenAIEx y son:

Apéndice 1: Métodos y Estadística en GenAIEx 6.

Apéndice 2: Autocorrelación Espacial en GenAIEx 6.

Apéndice 3: Resumen de respuestas y nombres de hojas de GenAIEx 6.

Apéndice 4: Diagramas de flujo que ilustran procedimientos y rutas para realizar los análisis en GenAIEx 6.

Apéndice 5: Opciones de Importación y Exportación de datos.

Apéndice 6: El análisis 'TwoGener' en GenAIEx 6.

## Estilos del texto en el Manual

---

En el texto se utilizaron estilos diferentes para referirse a las acciones, menús y cuadros del programa:

### **Nombres de Menú (eg. GenAIEx)**

*Opciones de Menú (e.g. Distance)*

*Subopciones de Menú (e.g. Genetic)*

Nombre de cuadros de diálogo (e.g. Genetic Distance Options)

*Opción de cuadro de diálogo (e.g. Binary)*

*Recomendaciones son escritas en itálica.*

<b>Notas para los usuarios con respecto a los procedimientos son escritas en cuadros de texto.</b>
--

## El Manual en Español

---

Para facilitar el manejo en la interfase del usuario, en la guía en español se han mantenido en inglés los nombres de las opciones de menú, los cuadros de diálogo y las acciones a realizar para cada análisis. Sin embargo, la guía ofrece una interpretación de cada uno, para que puedan ser fácilmente identificados y ejecutados por el usuario.

## Descargo de responsabilidad

---

*Esta guía de GenAIEx 6 es proporcionada de forma gratuita por las autoras. Ha sido escrita con consentimiento y en consulta constante con los autores del programa. Aunque se ha tenido sumo cuidado en mantener la exactitud del texto, las autoras no se hacen responsables por errores involuntarios o problemas que puedan encontrar los usuarios. No se proporciona soporte y /o apoyo particular para los usuarios.*

*Por favor observe que esta guía fue realizada en el 2005, basada en la versiones de prueba de GenAIEx 6. Pueden presentarse algunos cambios menores en los cuadros de dialogo y opciones ofrecidas en la versión oficial de GenAIEx 6 en Marzo de 2006.*

## El Entorno de GenAIEx

### Resumen

---

GenAIEx lee la información contenida en una hoja de trabajo de Excel, que consta de parámetros esenciales, etiquetas, etiquetas opcionales y los propios datos.

Hay disponibles varias opciones para que los usuarios estructuren sus datos apropiadamente, desde datos en una hoja de trabajo preexistente, hasta opciones para la importación automática, edición y estructuración de datos resultantes de un sistema de genotipado o secuenciación.

En el diseño de GenAIEx, el objetivo ha sido, en lo posible, hacer el manejo de datos tan fácil como eficiente. No obstante, hay varias restricciones impuestas por el entorno Excel. Estas son esbozadas abajo, con algunos indicadores útiles de cómo lograr mejores resultados de GenAIEx en el entorno de Excel.

### Limite de Datos

---

GenAIEx esta limitado por Excel a 256 columnas de datos. Esto equivale a 254 loci binarios o haploide o 127 loci codominantes. El número máximo de muestras es aproximadamente 65,500. En la práctica las limitaciones de memoria de su computador puede poner el límite al tamaño del conjunto de datos que usted desee correr.

Cuando se tienen más de 254 loci binarios (ejemplo. conjuntos grandes de datos de AFLP) pero con menos de 254 muestras, una opción de 'Transposición' esta disponible (*Options: Generic: Show Transposed Menu*).

El límite para las matrices triangulares de distancia es de 254 por 254 muestras. Para conjuntos de datos más grandes use la opción de *Distance Matrix as Column (Matriz de Distancia como Columna)* en el cuadro de diálogo pertinente.

### Datos de Ingreso (Input)

---

Los Inputs son los datos en bruto o matrices de distancia en el formato apropiado para GenAIEx (ver pag 11). Para hacer un análisis, la hoja de trabajo que contiene los datos debe ser activada (visible). Algunos análisis y procedimientos se hacen usando como datos de ingreso varias hojas de trabajo. A menos que se explique en otra parte, es necesario que estas hojas estén al lado izquierdo del cuaderno de Excel, en orden de 1 a n.

En lo posible, GenAIEx ofrece dos opciones para ayudar a mantener el registro de datos y los resultados de los análisis. En el cuadro de dialogo inicial para procedimientos estadísticos, (*Data Parameters*) el usuario puede ingresar: 1. Un prefijo para la hoja de trabajo, este prefijo ayuda a identificar la hoja de respuesta de un determinado análisis. 2. Un título para la hoja de respuestas que puede proporcionar detalles específicos del análisis realizado. Este título aparecerá en la parte superior de cada hoja de respuestas. Se recomienda enfáticamente usar ambas opciones.

## Resultados (Output)

---

GenAlEx puede generar muchas hojas de trabajo en un análisis rutinario, por esto es importante tener habilidad para manipular nuevos cuadernos y hojas de trabajo. Cada hoja de trabajo producida por GenAlEx tiene un nombre específico, dependiendo del análisis actualizado. Este es particularmente útil en los análisis que resultan en varias hojas de trabajo. En el manual los nombres de las hojas de trabajo se identifican usando llaves (ejemplo [GD]). Una lista de nombres para las hojas de trabajo es proporcionada en el Apéndice 3. El usuario puede definir prefijos y agregarlos a los nombres de las hojas de trabajo para tener más claridad (ver sección anterior).

La generación de las hojas de trabajo de GenAlEx está diseñada de forma que las hojas de datos aparecen siempre al extremo de la mano izquierda del cuaderno de trabajo. Así las hojas de resultado para muchas opciones pueden aparecer al extremo de la mano derecha. Sin embargo los resultados para Distancia Genética pueden aparecer en el extremo izquierdo, porque la matriz de distancia es usada como dato de ingreso para los análisis siguientes.

Los gráficos son generados en formato normal de Excel y puede ser necesario redimensionarlos para ver toda la información. Todos los gráficos pueden ser editados usando las funciones normales de Excel.

## Formato de los datos para GenAIEx

### Formato Numérico

GenAIEx requiere que todos los datos estén codificados y formateados como dato numérico de Excel. Sobre todo tenga cuidado para evitar usar las opciones de formato de texto, y desactivar todas las opciones de auto formatos. Hay disponibles opciones avanzadas para procesar datos de secuencias de ADN, para encontrar polimorfismos y haplotipos y convertir estos datos a formato numérico (ver pag 76 o 84).

*Recomendación: para verificar si sus valores numéricos son considerados como números por Excel; en la barra de herramientas de formato presione los botones de aumento o disminución de decimales. Si Excel es incapaz de mostrar los decimales, sus números no están en formato numérico.*

### Datos faltantes

Virtualmente todas las opciones de GenAIEx se ocupan de datos faltantes. Los datos faltantes pueden ser particularmente problemáticos para la distancia en parejas de individuos, por ejemplo como se encuentra en el análisis de AMOVA, Mantel y correlación espacial. GenAIEx proporciona una opción única para interpolar datos faltantes, en el cálculo de las distancias para cada pareja. Esta acción insertará como distancia entre una pareja de individuos, la distancia genética promedio de las poblaciones contrastadas (por ejemplo dentro de la población 1 o entre las poblaciones 1 y 2). No obstante para evitar el prejuicio excesivo, para los cálculos de distancia basados en individuos, deben minimizarse los números grandes de datos faltantes.

Datos codominantes y haplotipicos son codificados como '0'. Datos binarios faltantes son codificados como '-1'.

### Marcas de las muestras

Si su plan es aprovechar todas las ventajas de GenAIEx, es esencial que a cada muestra se la de un único identificador numérico. Los nombres de las muestras pueden llevar un prefijo de carácter alfa, pero este debe ser el mismo para todas las muestras en un mismo conjunto de datos. En este caso es importante conocer que, al ordenar en datos alfanuméricos, GenAIEx usa las reglas de orden de Excel, ordenando carácter por carácter, (ejemplo A11 vendrá después de A100). Para facilitar la ordenación nosotros recomendamos usar el formato A001...A199 para cuando se usan prefijos.

*Recomendación: si sus muestras no están en este formato, es posible crear rápidamente un único identificador numérico, usando la opción VLOOKUP o HLOOKUP en Excel (refiérase a la ayuda de Excel para más detalles).*

### Estructura de los datos

Para todos los análisis poblacionales con GenAIEx, los genotipos de todas las muestras de una misma población deben ingresarse como un bloque de filas continuas - una muestra por

fila. Para análisis por regiones en AMOVA, todas las poblaciones pertenecientes a una región también deben ingresarse como un bloque continuo. Las poblaciones son codificadas en la Columna 2 (ver p11).

Para datos del análisis *TwoGener*, los genotipos de cada madre y su respectiva descendencia se ingresan como un bloque continuo, siendo la madre el primer individuo de cada bloque. Las agrupaciones de la madre son codificadas en la Columna 2, como si fueran poblaciones.

## Parámetros y marcas de datos

---

Las marcas y parámetros de los datos son cruciales en GenAlEx contundentes para leer y analizar los datos. GenAlEx guarda todos los parámetros y marcas en las filas 1, 2 y 3 de las hojas de trabajo. Las columnas A y B son usados para las marcas de las muestras y poblaciones respectivamente. Los datos actuales empiezan en la celda C4 de una hoja de trabajo.

Los parámetros y marcas de los datos se puedan ingresar por diferentes maneras.

1. Una hoja de trabajo que contiene datos puede ser formateada de forma manual, para proporcionar parámetros apropiados.
2. La opción *Template* en el menú de **GenAlEx** puede ser usada para proporcionar los parámetros a través de un cuadro de diálogo, generando una hoja de trabajo formateada, en la cual se pueden ingresar los datos que se van a analizar (para más información vea la sección siguiente).
3. La opción *Parameters* en el menú de **GenAlEx** puede ser usada para obtener los parámetros relevantes de un conjunto de datos previos, e insertarlos en la ubicación apropiada (para mas información vea la sección siguiente). Esta opción requiere que sus datos inicien en la fila 4, y estén delimitados por columnas y filas en blanco.
4. Al iniciar un análisis, GenAlEx genera un cuadro de diálogo para los parámetros relevantes. Cambiando los parámetros en este cuadro proporcionan una manera fácil de seleccionar subconjuntos de datos para el análisis.
5. Si los datos fueron importados usando las opciones de GenAlEx, los parámetros esenciales y las marcas pueden ser insertados automáticamente, sin embargo marcas para los nombres de los loci se deben ingresar manualmente.

### Ubicación de parámetros

Los parámetros esenciales son insertados en la fila 1. Estos son: No Loci (celda A1); No. de muestras (celda B1); No. de Poblaciones (celda C1); Tamaño para cada población (celda D1..a celda n1).

B1 : No. Muestras

A1 : No. Loci

C1: No. Pops.

D1 – F1 : Tamaño de cada población

1	2	8	3	2	3	3
2	Example Dataset					
3	CODE	SITE	C2	CAM5	CAMM	MD
4	RF0707	CAM5	148	158	132	134
5	RF0708	CAM5	150	158	138	144
6	RF0861	CAMM	148	158	130	134
7	RF0862	CAMM	148	162	130	134
8	RF0863	CAMM	150	162	126	130
9	RF1195	MD	156	158	130	132
10	RF1196	MD	158	160	138	144
11	RF1197	MD	146	158	132	134

A2: titulo opcional.

D2 – F2: Marcas de Pop.

Fila 3: Marcas opcionales, incluyendo nombre de locus

Col. B con marcas pop. en bloques continuos

Col. A con las marcas de muestras iniciando en A4. Cada muestra tiene un único identificador numérico.

Datos codominantes con 2 columnas por locus, iniciando at C4.

Si se requiere información regional (solo para AMOVA), los parámetros para el No. de regiones se inserta dentro de la celda inmediatamente después del último tamaño poblacional, y el tamaño de cada región sigue entonces en las celdas subsecuentes (ver ejemplos sobre datos codominantes abajo).

## Formato de los datos

GenAlEx acepta 3 tipos de datos codificados numéricamente:

1. Datos codominantes con 2 columnas por locus.
2. Datos Dominantes, Haploides (incluyendo Haplotipos), o Secuencias codificados numéricamente con 1 columna por locus/base.
3. Datos geográficos con 2 columnas para coordenadas X y Y.

*Recomendación: Ejemplos de todos los formatos de datos de GenAlEx pueden ser creados usando la opción **Create**. Esta es una vía útil para explorar todo el rango de opciones de GenAlEx.*

### Formato para datos codominantes

Los datos codominantes se presentan en dos columnas por locus como en la figura de abajo. Los alelos se pueden codificar con códigos numéricos simples (1, 2, 3 etc). Alternativamente, y preferiblemente para datos de microsatélites, los alelos se pueden codificar como tamaño total en pares de bases (pb) o como el número inferido de repeticiones de secuencias simples. Estos dos últimos formatos son esenciales para el cálculo de la medida de distancia,  $R_{st}$ . No hay un límite para el número de alelos codificados numéricamente. Alelos codificados en tamaño de pb son aceptados hasta un tamaño máximo de 999. Los alelos codominantes no necesitan ser numerados consecutivamente.

### Ejemplo de datos codominantes, numéricamente codificados, con parámetros de región.

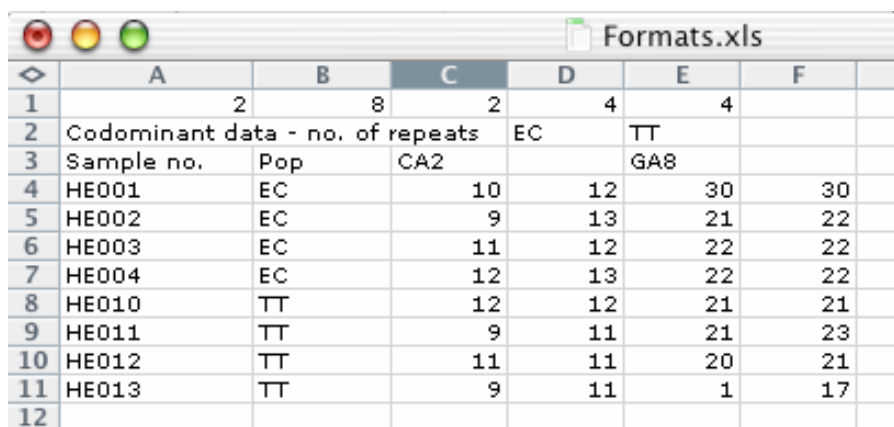
	A	B	C	D	E	F	G	H	I	J
1		2	8	4	2	2	2	2	4	4
2	Example of Codominant Data			Pop 1	Pop 2	Pop 3	Pop 4		Region 1	Region 2
3	Sample N(	Pop.	Locus 1		Locus 2					
4	1	Pop 1	3	3	1	4				
5	2	Pop 1	2	3	2	4				
6	3	Pop 2	2	4	3	4				
7	4	Pop 2	1	4	1	3				
8	5	Pop 3	3	4	2	2				
9	6	Pop 3	1	2	3	3				
10	7	Pop 4	4	4	2	2				
11	8	Pop 4	2	4	4	4				
12										

En este ejemplo las cuatros poblaciones están divididas en dos regiones, con las poblaciones 1 y 2 en la región 1 y las poblaciones 3 y 4 en la región 2.

### Ejemplo de datos codominantes de microsatélites, con alelos codificados como tamaño de fragmento,

	A	B	C	D	E	F	G
1		2	8	2	4	4	
2	Codominant data - fragment size			EC	TT		
3	Sample no.	Pop	CA2		GA8		
4	HE001	EC	294	298	274	274	
5	HE002	EC	292	300	256	258	
6	HE003	EC	296	298	258	258	
7	HE004	EC	298	300	258	258	
8	HE010	TT	298	298	256	256	
9	HE011	TT	292	296	256	260	
10	HE012	TT	296	296	254	256	
11	HE013	TT	292	296	214	248	
12							

### Ejemplo de datos de microsatélites, con alelos codificados como el número inferido de repeticiones.



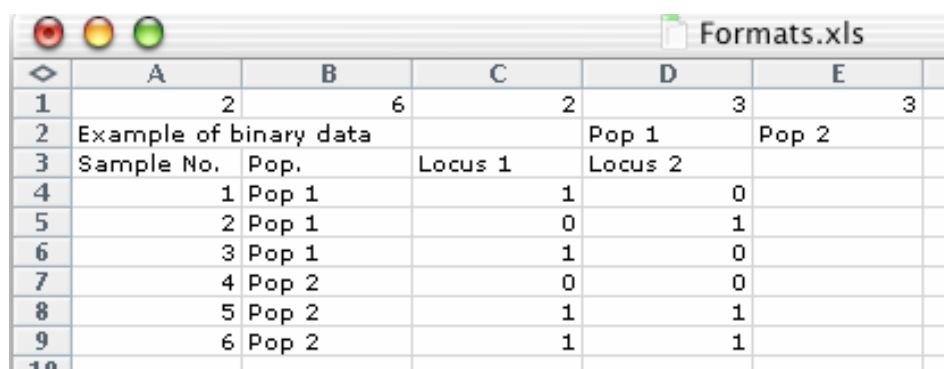
	A	B	C	D	E	F
1	2	8	2	4	4	
2	Codominant data - no. of repeats			EC	TT	
3	Sample no.	Pop	CA2		GA8	
4	HE001	EC	10	12	30	30
5	HE002	EC	9	13	21	22
6	HE003	EC	11	12	22	22
7	HE004	EC	12	13	22	22
8	HE010	TT	12	12	21	21
9	HE011	TT	9	11	21	23
10	HE012	TT	11	11	20	21
11	HE013	TT	9	11	1	17
12						

Estos son los mismos datos del ejemplo anterior, para loci con repeticiones de secuencias simples de 2pb.

### Formato para datos dominantes, haploides o secuencias.

Datos dominantes, haploides (incluidos haplotipos) o secuencias son presentados en una sola columna por locus. Los datos haploides pueden ser codificados numéricamente desde 1...n, o pueden ser representados por múltiples sitios variables. En este caso cada sitio (locus) puede tener estados múltiples, y se coloca en las columnas 3...n. Para secuencias o datos de SNP las bases se codifican numéricamente como sigue: A=1, C=2, G=3, T=4,=5; -=5, todos los otros caracteres = 0. GenAlEx proporciona varias opciones para importar los datos de secuencias y auto conversión a números.

### Ejemplo de datos dominantes, o datos binarios.



	A	B	C	D	E
1	2	6	2	3	3
2	Example of binary data			Pop 1	Pop 2
3	Sample No.	Pop.	Locus 1	Locus 2	
4	1	Pop 1	1	0	
5	2	Pop 1	0	1	
6	3	Pop 1	1	0	
7	4	Pop 2	0	0	
8	5	Pop 2	1	1	
9	6	Pop 2	1	1	
10					

**Ejemplo de datos de secuencia, codificados numéricamente como sitios polimórficos múltiples.**

	A	B	C	D	E
1	3	6	2	3	3
2	Example of haplotype data			Pop 1	Pop 2
3	Sample No.	Pop.	bp42	bp67	bp114
4	1	Pop 1	1	1	2
5	2	Pop 1	1	1	2
6	3	Pop 1	3	1	2
7	4	Pop 2	1	1	4
8	5	Pop 2	1	3	4
9	6	Pop 2	1	3	4
10					
11					

**Ejemplo de datos haplotípicos, con haplotipos individuales codificados numéricamente.**

	A	B	C	D	E	F
1	1	6	2	3	3	
2	Example of haplotype data			Pop 1	Pop 2	
3	Sample No.	Pop.	cpDNA			
4	1	Pop 1	1			
5	2	Pop 1	1			
6	3	Pop 1	2			
7	4	Pop 2	3			
8	5	Pop 2	4			
9	6	Pop 2	4			
10						
11						

Estos haplotipos corresponden a las secuencias presentadas en el ejemplo anterior.

### **Formato para datos regionales.**

Para análisis basados en regiones solo en la AMOVA, todas las poblaciones que pertenecen a una misma región deben ingresarse como un bloque continuo. Los parámetros para los datos regionales se ingresan en la Fila 1, inmediatamente después del último tamaño poblacional.

**Ejemplo de datos con parámetros regionales.**

	A	B	C	D	E	F	G	H	I	J
1	2	8	4	2	2	2	2	2	4	4
2	Example of Codominant Data			Pop 1	Pop 2	Pop 3	Pop 4		Region 1	Region 2
3	Sample No.	Pop.	Locus 1		Locus 2					
4	1	Pop 1	3	3	1	4				
5	2	Pop 1	2	3	2	4				
6	3	Pop 2	2	4	3	4				
7	4	Pop 2	1	4	1	3				
8	5	Pop 3	3	4	2	2				
9	6	Pop 3	1	2	3	3				
10	7	Pop 4	4	4	2	2				
11	8	Pop 4	2	4	4	4				
12										

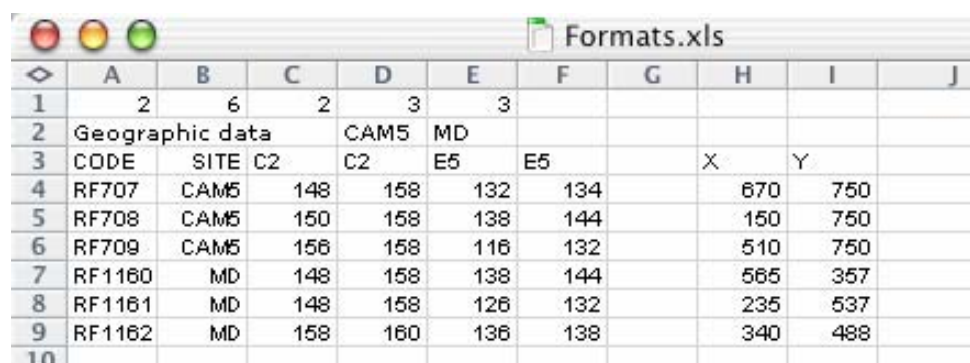
En este ejemplo las 4 poblaciones están divididas en 2 regiones (Celda H1) con las poblaciones 1 & 2 en la región 1 y las poblaciones 3 & 4 en la región 2. Ambas regiones tienen 4 individuos (Celdas I1 y J1).

**Formato para datos geográficos.**

Por conveniencia, ambas distancias geográfica y genética, pueden ser calculadas con un solo análisis. Las coordenadas pueden ingresarse como números enteros o decimales.

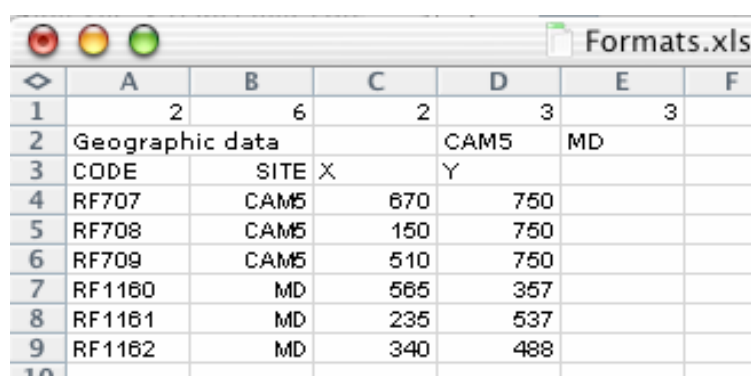
Las coordenadas X y Y pueden ser leídas por GenAlEx desde tres diferentes formatos.

1. X / Y datos están ubicados en la misma hoja de trabajo de los datos genéticos, y separados de los datos genéticos por una sola columna en blanco. Este formato es usado por GenAlEx para varios análisis, incluyendo los análisis Clonal y TwoGener.

**Ejemplo de datos geográficos después de los datos genéticos.**


	A	B	C	D	E	F	G	H	I	J
1	2	6	2	3	3					
2	Geographic data			CAM5	MD					
3	CODE	SITE	C2	C2	E5	E5		X	Y	
4	RF707	CAM5	148	158	132	134		670	750	
5	RF708	CAM5	150	158	138	144		150	750	
6	RF709	CAM5	156	158	116	132		510	750	
7	RF1160	MD	148	158	138	144		565	357	
8	RF1161	MD	148	158	126	132		235	537	
9	RF1162	MD	158	160	136	138		340	488	
10										

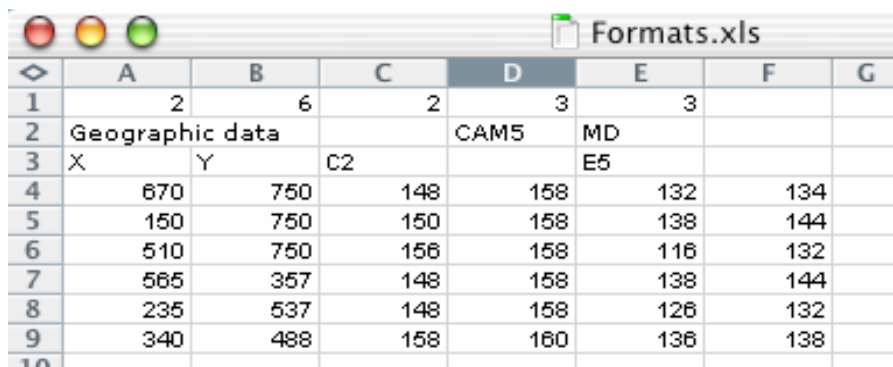
2. En una hoja separada, en las columnas C y D. En este caso, las muestras y las marcas de poblacionales en las columnas A & B corresponderán exactamente con los datos genéticos. Este formato es también apropiado si solo se requiere la distancia geográfica. Este formato se requiere para análisis como Autocorrelación Espacial 2D.

**Ejemplo de datos geográficos en las columnas C & D.**


	A	B	C	D	E	F
1	2	6	2	3	3	
2	Geographic data			CAM5	MD	
3	CODE	SITE	X	Y		
4	RF707	CAM5	670	750		
5	RF708	CAM5	150	750		
6	RF709	CAM5	510	750		
7	RF1160	MD	565	357		
8	RF1161	MD	235	537		
9	RF1162	MD	340	488		
10						

3. Antes de los datos genéticos, en las columnas A y B. Este formato se mantiene para permitir la compatibilidad con conjuntos de datos de versiones anteriores de GenAEx pero no es recomendado.

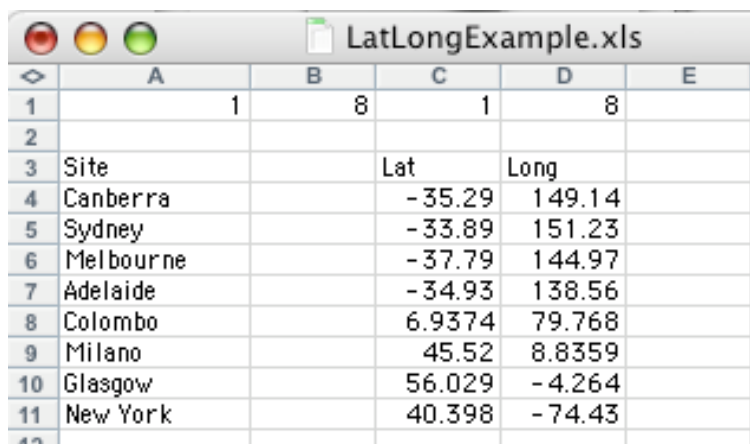
### Ejemplo de datos geográficos en las columnas A & B.



	A	B	C	D	E	F	G
1	2	6	2	3	3		
2	Geographic data			CAM5	MD		
3	X	Y	C2		E5		
4	670	750	148	158	132	134	
5	150	750	150	158	138	144	
6	510	750	156	158	116	132	
7	565	357	148	158	138	144	
8	235	537	148	158	126	132	
9	340	488	158	160	136	138	
10							

### Ejemplo de Datos de latitud / longitud decimal.

Datos de Latitud / Longitud pueden contener valores negativos. De acuerdo con las normas internacionales, los valores de latitud se presentan primero. Estos valores son transformados apropiadamente cuando están mapeados (usando la opción *Graph: Lat/Long*). Datos de Latitud / Longitud pueden ingresarse en alguno de los formatos arriba presentados.



	A	B	C	D	E
1	1	8	1	8	
2					
3	Site		Lat	Long	
4	Canberra		-35.29	149.14	
5	Sydney		-33.89	151.23	
6	Melbourne		-37.79	144.97	
7	Adelaide		-34.93	138.56	
8	Colombo		6.9374	79.768	
9	Milano		45.52	8.8359	
10	Glasgow		56.029	-4.264	
11	New York		40.398	-74.43	
12					

### Formato para ingresar datos a TwoGener.

Los datos para ingresar en el *TwoGener* consisten en bloques continuos para cada grupo de madre y descendencia, con los bloques codificado por madres en la columna 2. La madre debe estar colocada como el primer individuo en cada grupo. Las coordenadas X Y deben colocarse después de los datos genéticos en la misma hoja de trabajo y separados por una columna en blanco.

### Ejemplo de datos para *TwoGener*

	A	B	C	D	E	F	G	H	I	J	K
1	3	10	3	3	3	4					
2	TwoGener data format			1	2	3					
3	Sample	Mother	Locus1		Locus2		Locus3			X	Y
4	1	1	3	3	4	4	2	4		6	95
5	2	1	2	3	1	4	2	4		6	95
6	3	1	3	4	4	4	2	4		6	95
7	4	2	2	3	4	4	1	4		11	100
8	5	2	3	3	4	4	1	3		11	100
9	6	2	2	2	1	4	1	3		11	100
10	7	3	1	4	2	2	1	3		59	9
11	8	3	1	1	1	2	2	3		59	9
12	9	3	1	2	2	2	2	3		59	9
13	10	3	4	4	2	3	1	2		59	9
14											

En este ejemplo las muestras en filas 4, 7 y 10 son las madres. Las muestras en las filas 5 y 6 son los hijos de la primera madre y así sucesivamente.

## Opciones de GenAEx para Formatear Conjuntos de Datos.

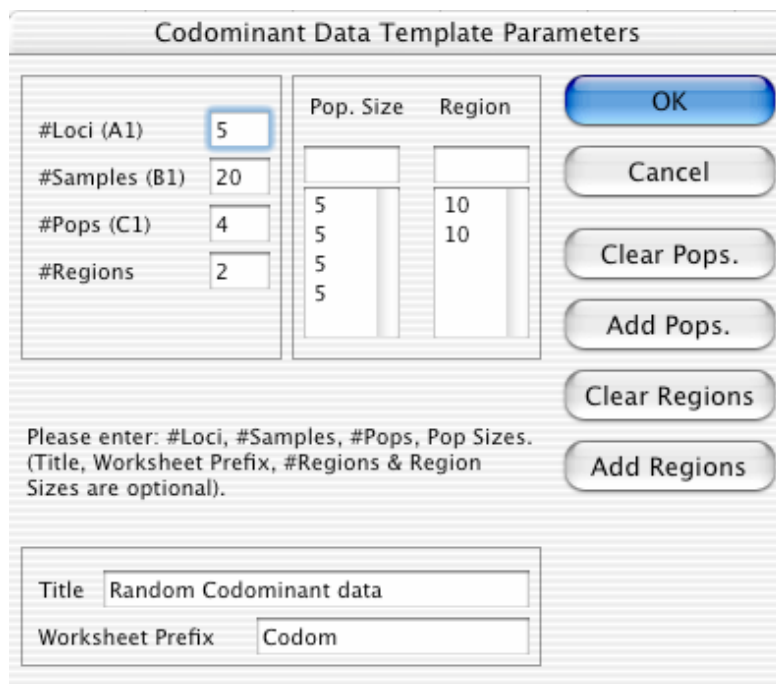
### Template (Modelo)

La opción de modelo (*Template*) hace fácil formatear un conjunto de datos para GenAEx. La opción coloca una nueva hoja de trabajo con los parámetros y marcas apropiadas en las que usted puede ingresar sus datos.

#### Procedimiento

1. Abra un libro de trabajo, escoja la opción *Template* desde el menú de **GenAEx**, y seleccione cualquiera *Codominant*, *Binary* o *Haplloid* de acuerdo con sus requerimientos.
2. En el cuadro de diálogo **Data Parameters** ingrese el número (#) de loci, # muestras (samples), # poblaciones (pops), y si se requiere, # de regiones en el panel del lado izquierdo. Estos parámetros son ingresados en la celda apropiada (A1, B1 etc) en la hoja de trabajo [D].
3. Ingrese el tamaño de cada población abajo en el cuadro de edición ('Pop. Size'), y agregue a la lista de poblaciones abajo usando la opción *Add Pops*. Use la opción *Clear Pops* para borrar la lista. La información de regiones se ingresa de forma similar cuando se requiere (solo para AMOVA).
4. Ingrese un título y un prefijo para sus datos y pulse Ok. El resultado es una hoja de trabajo [D].

Use este modelo como una base para ingresar sus conjuntos de datos.



**Codominant Data Template Parameters**

#Loci (A1)	5	Pop. Size	Region
#Samples (B1)	20		
#Pops (C1)	4	5	10
#Regions	2	5	10
		5	
		5	

Please enter: #Loci, #Samples, #Pops, Pop Sizes.  
(Title, Worksheet Prefix, #Regions & Region Sizes are optional).

Title: Random Codominant data  
Worksheet Prefix: Codom

## Create (Creando)

---

El menú *Create* da opciones para crear ejemplos al azar de todos los formatos de GenAlEx, ambos Genéticos y Geográficos. Estos conjuntos de datos son útiles para explorar el rango de procedimientos de GenAlEx.

Las sub-opciones para datos codominantes y haplotípicos (*Codominant* y *Haploid*), crean datos con alelos numéricamente codificados. Datos binarios (*Binary*) se codifican como 1 / 0. Los dos tipos de datos de secuencia generan datos de secuencias de ADN con nucleótidos de codificación-Alpha (A, C, G & T). Los datos de secuencia son creados con baja tasa de polimorfismo para posibilitar el encuentro de haplotipos en el análisis siguiente. La sub-opción *Raw Sequence* puede crear una hoja de datos con la longitud total de la secuencia en una celda, por tanto la sub-opción *Sequence* insertará cada base nucleotídica dentro de la secuencia en una celda separada hasta un máximo de 254 pb.

Si las opciones avanzadas, *TwoGener*, *Clonal* o *Transposed* están activados por medio del cuadro de texto de la opción *Generic*, aparecen opciones para crear conjuntos de datos apropiados. La opción *TwoGener* crea un conjunto de datos donde cada descendencia tiene por lo menos un alelo de la madre, quien se representa como la primera muestra de cada “grupo materno”.

*Recomendación: Create puede usarse para crear el conjunto de datos de prueba para la enseñanza. Para datos codominantes la frecuencia genotípica se aproximara a lo esperado bajo apareamiento al azar, y puede usarse para presentar una población con patrones genéticos típicos de apareamiento al azar.*

*Recomendación: Los conjuntos de datos creados usando esta opción, están en el formato correcto para GenAlEx y pueden usarse como un evaluador de errores inesperados. En este caso use Create para generar un conjunto de datos de igual tamaño al suyo, y reevalúe el procedimiento problemático. Si funciona, el problema debe estar en su conjunto de datos.*

### Procedimiento

1. Con un archivo de Excel abierto, escoja la opción *Create* desde el menú **GenAlEx** y seleccione el tipo de datos requerido.
2. En el cuadro de diálogo **Create Data Parameters** ingrese el número (#) de loci, # de muestras (Samples), # de poblaciones (pops), y si se requiere, # de regiones en el panel ubicado a mano izquierda. Estos parámetros se insertan dentro de la celda apropiada en la hoja de trabajo [D]. Para datos codominantes y haploides se debe ingresar el máximo número de alelos requeridos.
3. Ingrese el tamaño de cada población (pop) en el cuadro de edición que aparece abajo ‘Pop. Size’, y adicione la lista de poblaciones abajo usando la opción *Add Pops*. Use la opción *Clear Pops* para borrar la lista. La información con respecto a las regiones se ingresa de forma similar, si se requiere (solo en AMOVA).
4. Ingrese un título y un prefijo para la hoja de trabajo de sus datos y pulse *Ok*. El resultado es una hoja de trabajo [D].

**Create Codominant Data Parameters**

#Loci (A1)	10	Pop. Size	Region	OK
#Samples (B1)	100			Cancel
#Pops (C1)	4	25		Clear Pops.
#Regions		25		Add Pops.
Max #Alleles	10	25		Clear Regions
		25		Add Regions

Please enter: #Loci, #Samples, Max #Alleles, #Pops & Pop Sizes. (Title, Worksheet Prefix, #Regions & Region Sizes are optional).

Title: Codom data, max 10 alleles  
Worksheet Prefix: Codom

## Parameters (Parámetros)

---

La opción de Parámetros es un medio rápido para obtener los parámetros requeridos en GenAlEx a partir de un conjunto de datos previos, además de ubicarlos correctamente. Los datos deben estar en la estructura normal de GenAlEx, con las marcas de las muestras en la columna A, las marcas para la población en la columna B y los datos a partir de la celda C4. El conjunto de datos debe estar limitado por filas y columnas vacías porque GenAlEx usa celdas vacías para definir el límite de los datos. Todas las muestras de una población deben tener la misma marca de la población y estar en filas contiguas.

Hay una opción disponible para insertar las dos filas de parámetros, dentro de un conjunto de datos sin formato (*Insert Header Rows*).

Para cada opción el menu GenAlEx , examina la columna(s) reelevant e inserta los parámetros correspondientes en los lugares correctos.

## Data (Datos)

---

Las opciones de menú de datos (*Data*) ofrecen varios comandos para manipular rápidamente sus datos. Los datos deben estar en el formato apropiado de GenAIEx (incluyendo los parámetros).

*Sort on Sample (Orden en las muestras)*: Ordena el conjunto total de datos acorde con la marca de las muestras.(en la Columna A), según las reglas de ordenar en Excel (ver sección Marca de muestras, (p10).

*Sort on Pop. (Orden en la población)* Ordena el conjunto total de datos de acuerdo con la marca de la población (en la Columna B), según las reglas de orden de Excel (ver sección Marca de muestras (p 10).

*Select Data Rows (Selección de las filas de datos)*: Habilita la selección rápida de todas las filas de datos. Esto es útil para después hacer ordenaciones sobre cualquier columna usando la opción de Orden (*Sort*) de Excel en el menú de Datos (**Data**)

*Select Data Rows + Labels (Selección de las filas de datos y Marcas)*: Habilita la selección rápida de todas las filas de datos y la fila de marcas (Fila 3). Esto es útil para después hacer ordenaciones sobre una columna usando la opción ordenar (*Sort*) de Excel en el menú de Datos (**Data**)

*Split by Pop (división de datos por poblaciones)*: Divide los datos de poblaciones múltiples contenidas en un solo conjunto de datos. Cada conjunto de datos de una sola población es separado y ubicado en una hoja de trabajo aparte con el nombre de la población.

*Tri -> Table*: Convierte una matriz de distancia triangular a un formato de tabla. Asegúrese de tener seleccionada la hoja de trabajo que contiene la tri matriz de interés y en formato GenAIEx. La tabla es generada en una hoja de trabajo [TB].

*Col -> Table*: Convierte una distancia de parejas en formato de columnas a un formato de tabla. Asegúrese de tener activada la hoja de trabajo que contiene la distancia por parejas y en formato de GenAIEx. La tabla es generada en una hoja de trabajo [TB].

## Procedimientos Estadísticos

### Cuadro Normal de Parámetros de Datos

Para facilitar la interacción del usuario con GenAEx, El cuadro de diálogo inicial (**Data Parameters**) para los diferentes procedimientos estadísticos ha sido estandarizado en lo posible. Generalmente el cuadro tiene el siguiente formato. Para algunas aplicaciones se han hecho adaptaciones necesarias.

La sección de arriba del cuadro de diálogo proporciona cuadros de edición para ingresar los parámetros esenciales de locus, muestra y población. Si los parámetros se presentan en la hoja de datos, son ingresados automáticamente. Si solo se requiere un subconjunto de datos aquí se pueden cambiar los parámetros.

Después se genera una sección en la cual el formato de los datos es seleccionado.

Finalmente, abajo de cuadro de diálogo se dan dos opciones para ayudar a mantener la pista de los resultados de los análisis:

1. Un título (Title) para los resultados que puede tener detalles específicos de los análisis que se están realizando (máximo 30 caracteres). Este título aparecerá en lo alto de cada hoja de trabajo, junto con el nombre de la hoja de datos original.
2. Un prefijo (worksheet prefix) para la hoja de trabajo ayuda a identifica el resultado de un análisis particular.

#### Un ejemplo de Cuadro Normal de Parámetros de Datos para procedimientos estadísticos

#### Opciones del Cuadro de Parámetros de Datos

##### Cuadro Editor de Parámetros

Ingrese el número de loci, muestras y poblaciones en cada cuadro.

Ingrese el tamaño de población requerido en el cuadro editor y pulse el botón **Add Pops.** Ingrese los tamaño en orden de población desde 1 a n, como aparecen en la hoja de datos

**Clear Pops.** Utilice este botón para limpiar la lista de tamaños poblacionales.

##### Data Format

Seleccione el formato adecuado para sus datos.

Ingrese el título y un prefijo.

## Frequency (Frecuencia)

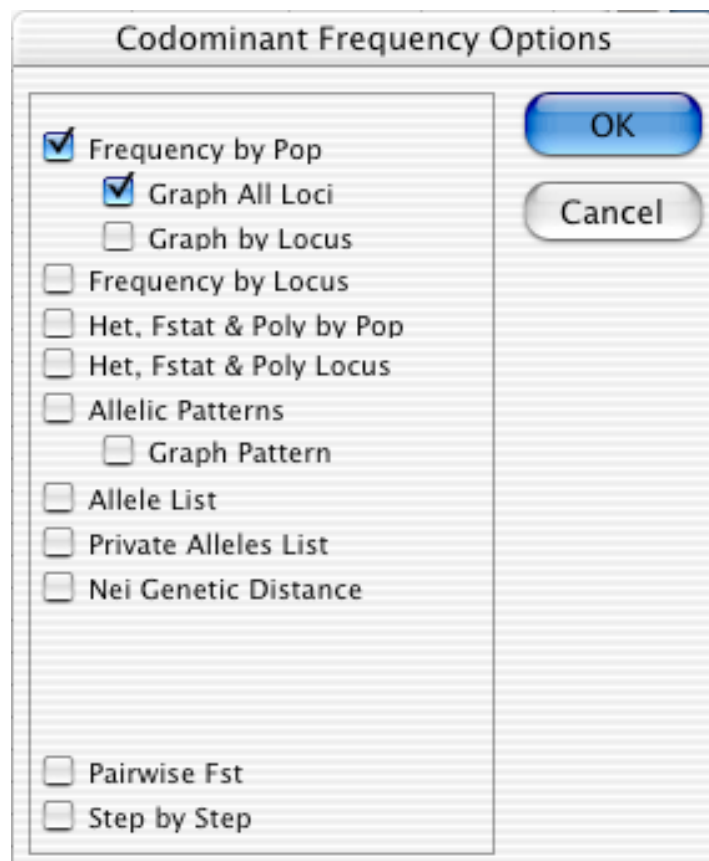
---

Este menú de opciones proporciona un rango de resúmenes estadísticos para datos codominantes, haplotípicos y dominantes.

### Procedimiento

1. Escoja la opción *Frequency* desde el menú de **GenAlEx**.
2. Ingrese toda la información apropiada en el cuadro de diálogo Allele Frequency Data Parameters.
3. Seleccione las opciones de frecuencia requeridas desde el cuadro de diálogo Frequency Options (las opciones disponibles depende del tipo de datos). Para información sobre estas opciones ver abajo.

### Opciones de Frecuencias Codominantes



**Frequency by Pop (Frecuencia por población) [AFP]:** Genera las frecuencias alélicas de cada locus por población.

**Graph All Loci (Gráfico de todos los loci) [AFP]:** Proporciona un solo gráfico de los datos de frecuencia alélica locus por locus. Para un conjunto de datos grande este resultado puede tardar un tiempo y puede ser preferible saltarse esta opción.

**Graph by Locus (Gráfico por locus) [AFG]:** Proporciona un gráfico de los datos de frecuencia alélica por cada locus individual. Para un conjunto de datos grande este resultado puede tardar un tiempo y puede ser preferibles saltarse esta opción.

**Frequency by Locus (Frecuencias por locus) [AFL]:** Genera las frecuencias alélicas de cada población con los loci en columnas. Para conjuntos de datos de microsatélites (con alelos codificados por tamaños en pares de bases) esta opción produce una tabla con el número de filas igual al número de alelos distintos encontrados en todo el conjunto de datos. Para conjuntos de datos con ciertos tamaños de alelos faltantes el resultado puede tardar un tiempo.

*Recomendación: La tabla de datos proporciona una buena indicación visual de la distribución de los tamaños de alelos, y el traslape entre loci y puede ser una herramienta útil para planear PCRs múltiplex de diferentes loci.*

**Het, Fstat & Poly by Pop (Heterocigosidad, estadísticos F y polimorfismo por población) [HFP]:** Genera para cada población: el número efectivo de alelos ( $N_e$ ), el índice de información (I), la heterocigosidad observada ( $H_o$ ) y esperada ( $H_e$ ), y el índice de Fijación (F), los estadísticos F y el porcentaje de loci polimorficos, todos en filas. Este es el formato normal para muchas publicaciones.

**Het, Fstat & Poly by Locus (Heterocigosidad, estadísticos F y polimorfismo por locus) [HFL]:** Genera la misma información de la opción anterior, pero con los loci en columnas no en filas.

**Allelic Patterns (Patrones alélicos) [APT]:** Resume las siguientes estadísticas para cada población: Número de alelos ( $N_a$ ),  $N_a$  con frecuencia >5%,  $N_e$ , I, Número de alelos únicos, Número de alelos menos comunes ( $\leq 25\%$ ), Número de alelos menos comunes ( $\leq 50\%$ ), y  $H_e$ .

**Graph Pattern (Gráfico de patrones) [APT]:** Genera una gráfica con la información anterior.

**Allele list (Lista de alelos) [ALI]:** Cuenta la ocurrencia de alelos por cada locus sobre el conjunto total.

*Recomendación: Esta es una herramienta útil para planear PCRs múltiplex de diferentes loci.*

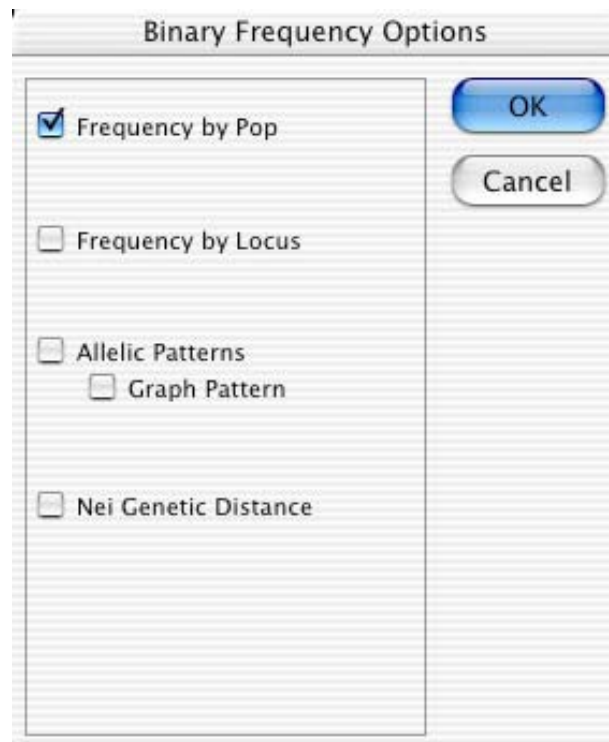
**Private alleles list (Lista de alelos únicos) [PAS] & [PAL]:** Genera una hoja PAS con una lista de alelos únicos por población y una hoja PAL, con una lista de muestras que contienen uno o mas alelos únicos. Estas están en el formato estándar de GenAlEx, por si se requieren para más análisis.

**Nei Genetic Distance (Distancia genética de Nei) [NEI] & [NEIT]:** Genera una hoja NEI con las tri-matrices de ambas distancias genéticas (Genetic Distance) de Nei e identidad genética (Genetic Identity) de Nei, por parejas de poblaciones. Estas matrices son dadas en el formato GenAlEx por si se requieren para más análisis. La distancia genética de Nei es también generada en formato de tabla en una hoja NEIT.

**Pairwise Fst (Parejas Fst) [FST] & [FSTT]:** Genera los valores de Fst de parejas poblacionales como una tri-matriz en la hoja FST, y también en formato de tabla en la hoja FSTT.

**Step by Step (Paso a paso):** Muestra paso por paso los cálculos, incluyendo las estadísticas alélicas, el promedio de heterocigosidad observada y esperada y la heterocigosidad total calculada para los estadísticos F.

### Opciones para Frecuencias Binarias



**Frequency by Pop (Frecuencia por población) [AFP]:** Genera en filas: la frecuencia de bandas,  $p$ ,  $q$  y  $H_e$  para cada locus por población. Al final, el % de bandas polimórficas,  $P$ , es producido para cada población.

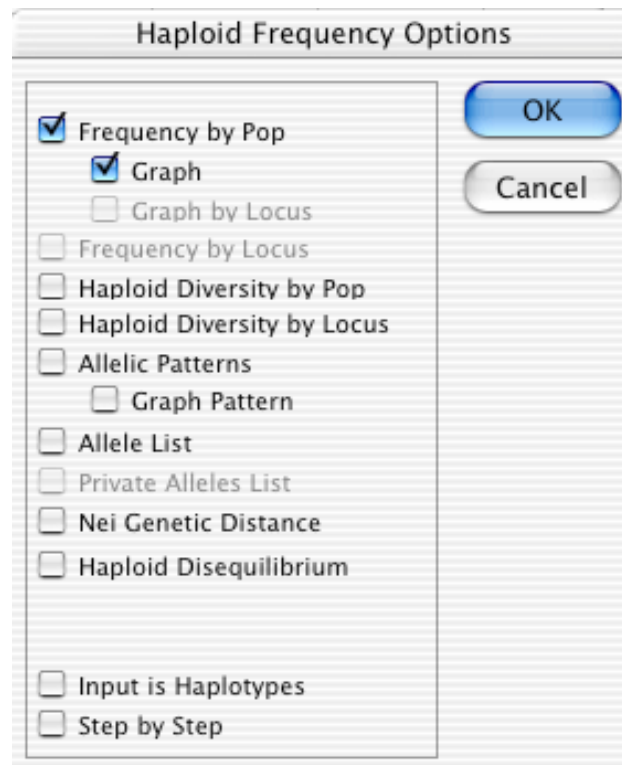
**Frequency by Locus (Frecuencia por Locus) [AFL]:** Genera en columnas: la frecuencia de bandas,  $p$ ,  $q$  y  $H_e$  para cada locus en cada población. Al final, el % de bandas polimórficas,  $P$ , es producido para cada población.

**Allelic Patterns (Patrones alélicos) [APT]:** Resume el Número de bandas de los loci, No. de bandas con frecuencia >5%, No. de bandas únicas, No. de bandas menos comunes (Lcomm.) ( $\leq 25\%$ ), No. de bandas LComm. ( $\leq 50\%$ ), el promedio y el SE de la heterocigosidad.

**Graph Pattern (Gráficas de patrones)[APT]:** Proporciona gráficos con los resultados de la información anterior.

**Nei Genetic Distance (Distancia genética de NEI) [NEI] & [NEIT]:** Genera en la hoja NEI dos tri-matrices poblacionales: distancia genética (Genetic Distance) de Nei e identidad genética (Genetic Identity) de Nei. Estas son dadas en el formato GenAlEx por si se requieren para más análisis. La distancia genética de Nei es también generada en formato de tabla en la hoja NEIT.

### Opciones para Frecuencias Haploides



**Frequency by Pop (Frecuencia por población) [AFP]:** Genera la frecuencia alélica de cada locus por población.

**Graph (Gráficas) [AFP]:** Proporciona resultados gráficos de la información anterior. Para un conjunto de datos grande los resultados pueden tardar algún tiempo y puede ser mejor saltarse esta opción.

**Graph by Locus (Gráfico locus por locus) [AFG]:** Proporciona un gráfico de los datos de frecuencia alélica para cada locus individual. Para un conjunto de datos grande este resultado puede tardar algún tiempo y puede ser mejor saltarse esta opción.

**Frequency by Locus (Frecuencia por locus) [AFL]:** Genera la frecuencia alélica de cada población con los loci en columnas.

**Haploid diversity by Pop (Diversidad haploide por población) [HDP]:** Genera el No. de alelos ( $N_a$ ) y diversidad Haploide ( $h$ ) por locus por población y también el promedio de diversidad poblacional ( $H$  &  $V_e$ ) sobre el total de los loci.

**Haploid diversity by Locus ([HDL]:** Proporciona el No. de alelos ( $N_a$ ) y  $h$  por locus por población y también el promedio de diversidad poblacional sobre el total de los loci.

**Allelic Patterns (Patrones alélicos) [APT]:** Resume para cada población las siguientes estadísticas: No. de alelos ( $N_a$ ),  $N_a$  con frecuencia  $>5\%$ ,  $N_e$ ,  $I$ , No. alelos únicos, No de alelos menos comunes ( $\leq 25\%$ ), No. de alelos menos comunes ( $\leq 50\%$ ), y  $H_e$ .

**Graph Pattern (Gráfico de patrones) [APT]:** Proporciona una gráfica de la información anterior.

**Allele List (Lista de alelos) [ALI]:** Cuenta la ocurrencia de alelos por cada locus sobre el conjunto total.

*Recomendación: Esta es una herramienta útil para planear PCR en múltiplex de diferentes loci.*

**Private Alleles List (Lista de alelos únicos) [PAS] & [PAL]:** Agrega a la hoja PAS una lista de alelos únicos por población, y a la hoja PAL una lista de muestras que contienen uno o más alelos únicos. Estas están en el formato estándar de GenAEx, por si se requieren para más análisis.

**Nei Genetic Distance (Distancia genética de Nei) [NEI] & [NEIT]:** Agrega a la hoja NEI las dos tri-matrices poblacionales, de distancia genética (Genetic Distance) de Nei e identidad genética (Genetic Identity) de Nei. Esta es dada en el formato GenAEx por si se requieren para más análisis. La distancia genética de Nei es también generado en formato de tabla en la hoja NEIT.

**Haploid disequilibrium (Desequilibrio haploide) [HDE] & [FDHDE]:** Genera el análisis de disequilibrio haploide ( $H_a$ ) en la hoja HDIS, y la distribución de frecuencia en la hoja FDHDIS. Esta opción pide el número de aleatorización requeridos para evaluar la significancia del disequilibrio haploide (0, 99, 999). Esta opción solo aplica con una sola población.

**Input is Haplotypes (Datos de haplotipos):** marcando esta opción cuando un haplotipo es codificado como un solo locus, asegura que en el resultado se anote como datos haplotipicos. Esta opción esta disponible solo con datos de un solo locus.

**Step by Step (Paso a paso):** Presenta paso por paso los cálculos, incluyendo cuentas de alelos, promedio de heterocigosidad observada y esperada y heterocigosidad total para el cálculo de estadísticos F.

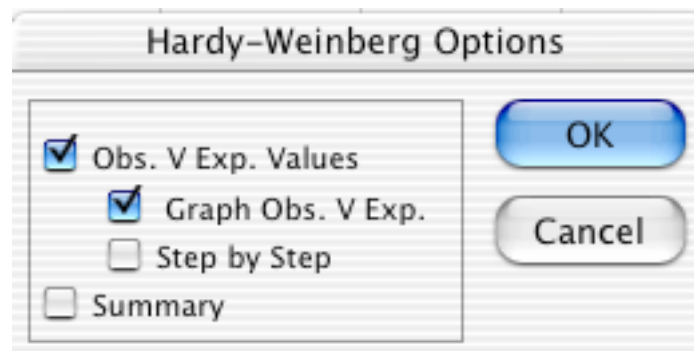
## HWE (Equilibrio Hardy-Weinberg)

---

Esta herramienta solo es aplicable cuando se tienen datos codominantes. La opción de HWE en GenAEx esta proporcionada principalmente para la enseñanza y exploración preliminar de datos. El programa GenePop proporciona pruebas exactas de HWE, y es recomendado para investigación. GenAEx ofrece una opción para exportar datos en el formato de GenePop.

### Procedimiento

1. En el menú **GenAEx**, señale la opción *HWE*.
2. Ingrese en el cuadro de diálogo de HWE Data Parameters todos los datos requeridos.
3. En la siguiente ventana, Hardy-Weinberg Options, escoja las opciones que desea, luego pulse *Ok*. Ver opciones abajo para mas detalles y los nombres de las hojas de repuestas.



### Opciones para Hardy-Weinberg

**Obs. V Exp. Values [HW]:** Presenta las frecuencias observadas y esperadas para cada genotipo y la prueba de Chi-Cuadrado para cada locus en cada población.

**Graph Obs. V Exp. [HW]:** Presenta un gráfico con los datos genotípicos antes mencionados.

**Step by Step [HW]:** Presenta paso por paso los cálculos realizados para la prueba de Chi-Cuadrado.

**Summary [HWS]:** Brinda un resumen de la estadística del Chi-Cuadrado, grados de libertad, y probabilidades para cada locus en cada población.

## Distance (Distancia)

---

El menú de *Distance (Distancia)* tiene dos sub-opciones: *Genetic...* and *Geographic...*. La sub-opción *Genetic...* permite calcular la distancia genética entre pares de individuos, para datos binarios, haplotípicos y codominantes. La sub-opción *Geographic* puede calcular distancias geográficas usando varios diferentes formatos.

Las distancias genética y geográfica pueden calcularse de forma simultánea. Para esto se ingresa al cuadro de diálogo opciones de distancia genética y se marca la opción *Geographic Options* para seguir con las opciones de distancia geográfica.

## Genetic Distance (Distancia Genética)

---

Esta opción genera matrices con la distancia genética entre pares de individuos, en el formato apropiado para análisis posteriores en GenAlEx. Una matriz de distancia genética es el primer paso para varios de los análisis disponibles en GenAlEx, incluyendo Análisis de Varianza Molecular (AMOVA). Las opciones de Distancia Genética y la de AMOVA comparten el mismo cuadro de diálogo inicial. Sin embargo dos cálculos para distancia genética codominante que son aplicables solo en la AMOVA se encuentran inactivos en el Cuadro de Distancia Genética.

### Procedimiento

1. Escoja la opción *Distance (Distancia)* desde el menú de **GenAlex** y luego seleccione *Genetic* desde el submenú.
2. Asegúrese que los parámetros del locus y muestras estén correctos en el cuadro de diálogo de *Genetic Distance Options*.
3. Seleccione el cálculo de distancia (*Distance Calculation*) apropiado y las opciones de resultados (output) que requiere (ver abajo).
4. Ingrese el título y prefijo de la hoja de trabajo, luego haga clic en *Ok*. Obtendrá la hoja de respuesta de Distancia Genética [GD].

## Opciones de Distancia Genética

### Distance Calculation (Calculo de la Distancia)

Escoja el cálculo apropiado para su tipo de datos (binarios, haploides o codominantes). Los detalles de los cálculos son presentados en el Apéndice 1. Para datos codominantes hay disponible solo un tipo de cálculo, los otros dos cálculos son específicos para la AMOVA; y solo accesibles en ese menú. La distancia genotípica que se haya aquí, es la base para muchos análisis posteriores, incluyendo Mantel, PCA, y un conjunto completo de análisis espaciales. Esta medida de la distancia genética también facilita la comparación entre datos codominantes, haplotípicos y/o binarios.

**Interpolate missing: (Interpolación de datos faltantes):** Cuando hay datos faltantes en un locus, para una comparación entre un par de individuos, GenAlEx puede interpolar la distancia genética promedio de este locus (calculada sobre todas las distancias de parejas individuales), contrastando parejas de poblaciones (por ejemplo dentro de la población 1 o entre las poblaciones 1 y 2). Esta opción no está disponible si los parámetros de las poblaciones no se especifican.

**List missing (Listado de datos faltantes) [GDML]:** Identifica las muestras que presentan datos faltantes por locus, y muestra los valores interpolados para cada locus.

**Linear Genetic (Genética Lineal):** Genera distancias genéticas linealizadas en lugar de cuadrados. Esta opción podría ser útil cuando se correlacionan la distancia genética con la distancia geográfica (la prueba de Mantel) Para otros propósitos asegurarse de no seleccionar esta opción.

**Geographic Options (Opciones Geográficas):** Esta opción trae el cuadro de diálogo de Geographic Distance Options (distancia geográfica).

**Distance Output Options (Opciones de resultados de distancia).**

**Output Total Distance Only (Hallar solo la distancia total):** Genera una matriz de distancia genética sumando todos los loci. Para los tres métodos codominantes, las matrices de distancia para cada locus suma entre loci asumiendo el principio de independencia.

**Output Distance All Loci (Hallar la distancia de todos los loci):** Genera una matriz para cada locus de forma individual, más la matriz de distancia genética total, cada uno separado en una hoja de trabajo. Esta opción es útil para analizar cada locus de forma específica, y para el análisis de auto-correlación espacial (Multiple locus Spatial Autocorrelation).

**Output (Resultados)**

**To Worksheet (Hoja de trabajo):** La matriz se presenta en una hoja de trabajo. Esta hoja es requerida para todas los análisis posteriores que requieren matriz de datos de distancia como la prueba de Mantel y análisis espaciales.

**As Tri Matrix:** Genera una matriz de distancia genética como una matriz triangular. Se recomienda su uso cuando se tienen datos con menos de 254 muestras.

**As Sq Matrix:** Genera una matriz de distancia genética con simetría cuadrada. Esta matriz es muy útil como 'input' para otros programas que requieren matrices cuadradas.

**As Column:** Genera una matriz en columna. Esta opción es requerida cuando se tienen conjuntos de datos de más de 254 muestras. Debido a que el límite en Excel es de 256 columnas, las matrices triangulares o cuadradas no pueden ser usadas para un gran número de datos.

**Label Matrix:** Ubica las marcas apropiadas a su matriz. Marcas de muestras y poblaciones son útiles para los outputs gráficos subsecuentes, como un PCA.

*Recomendación: Si el conjunto de datos es de un solo locus, aparece la opción para etiquetar la matriz con los genotipos. Esta es una herramienta útil para la enseñanza.*

## **Geographic Distance (Distancia Geográfica)**

La opción para calcular distancia geográfica puede ser accesada directamente a través de la sub-opción *Distance : **Geographic*** o por el cuadro de diálogo de Genetic Distance Options seleccionando *Geographic Options*.

### **Procedimientos para distancia geográfica con distancia genética**

1. Asegurese de seleccionar la opción *Geographic options* en el cuadro de diálogo de Genetic Distance Options.
2. En el cuadro de diálogo Geographic Distance Options ingrese toda la información apropiada (para mas información sobre las opciones vea abajo). Las distancias geográficas por pares son resultados en la hoja [GGD].

### **Opciones de Distancia geográfica con distancia genética**

#### **Data Source (Fuente de datos)**

**This sheet (Esta hoja):** Marque esta opción si sus datos geográficos están en la misma hoja que los datos genéticos.

**Other (Otra):** Marque esta opción si sus datos geográficos están en otra hoja de trabajo. Seleccione la hoja desde el menú desplegable.

### **X, Y Coordinates (Coordenadas X, Y)**

**After Genetic Data (Después de datos genéticos):** Escoja esta opción si sus datos XY vienen después de sus datos genéticos, en la misma hoja de trabajo. Debe haber una columna en blanco entre los datos genéticos y los datos XY. En la ilustración de arriba los datos XY están en las columnas 6 & 7, y los datos genéticos finalizan en la columna 4.

**Cols 1 & 2:** Marque esta opción si sus datos XY estén en las columnas 1 & 2. Este formato no es recomendado pero se mantiene de las versiones anteriores de GenAlEx para asegurar la compatibilidad de datos viejos.

### **Data (Datos)**

Escoja el tipo de datos apropiados:

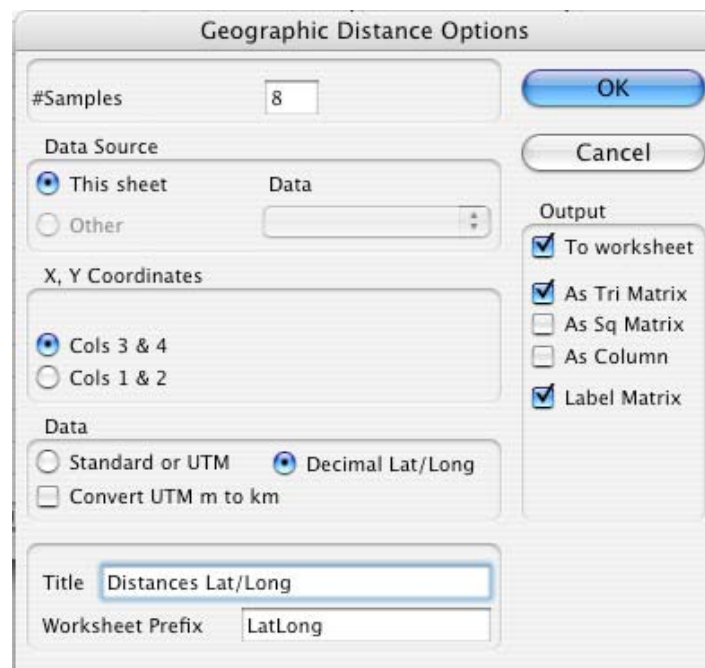
**Standard or UTM (Estandar o UTM):** Para valores dados en metros en el sistema UTM (Universal Transverse Mercator Grid), o en otras coordenadas de mapas. Estos valores deben usarse en análisis genéticos de fina escala.

**Convert UTM m to km (converter datos UTM en m a km):** Convierte valores UTM de metros a kilómetros. Esta opción es útil para un conjunto de datos de varios kms, que puede ser difícil de manejar en resultados gráficos.

**Decimal Lat/Long:** Escoja esta opción para valores decimal de latitud/longitud.

**Procedimiento solo para distancia geográfica**

1. Active la hoja de trabajo que contiene sus datos XY. Estos datos pueden ser ubicados en las columnas Cols 1 & 2, o Cols. 3 & 4.
2. Escoja la opción *Distance* desde el menú de **GenAlEx**, y luego desde el submenú seleccione *Geographic*.
3. Ingrese toda la información requerida en el cuadro de diálogo **Geographic Distance Options** (para más información sobre opciones ver abajo). Las distancias geográficas por pares son resultados en la hoja GGD.

**Opciones de distancia geográfica****Data Source (Fuente de datos)**

**This sheet:** Sus datos geográficos deben estar en la misma hoja activa.

**X, Y Coordinates (X, Y Coordenadas)**

**Cols 3 & 4:** Marque esta opción si sus datos X, Y están en las columnas 3 & 4.

**Cols 1 & 2:** Marque esta opción si sus datos X, Y están en las columnas 1 & 2.

**Data (Datos)**

Escoja el tipo de datos apropiados:

**Standard or UTM (Estandar o UTM):** Para valores dados en metros el sistema UTM. Estos valores deben usarse en análisis genéticos de fina escala.

**Convert UTM m to km (converter UTM m a km):** Convierte valores UTM en metros a kilómetros. Esta opción es útil para un conjunto de datos extenso sobre kms que podría ser difícil manejar en resultados gráficos.

**Decimal Lat/Long:** Para valores decimal de latitud/longitud.

### **Output (Resultados)**

**To Worksheet:** Genera una matriz de distancia en una hoja de trabajo. Esta es requerida para todos los análisis subsecuentes que necesitan datos de distancia geográfica, como Mantel y análisis de relaciones espaciales.

**As Tri Matrix:** Genera una matriz de distancia geográfica en forma de matriz triangular. Esta es la opción recomendada cuando se tienen conjuntos de datos de menos de 254 muestras.

**As Sq Matrix:** Genera una matriz de distancia genética en forma de matriz de cuadrado simétrico. Esta opción es útil para otros programas que requieren matrices cuadradas.

**As Column:** Genera matrices de distancia geográfica en forma de columna. Esta opción es proporcionada para conjuntos de datos con más de 254 muestras. Debido al límite de 256 columnas en Excel, matrices triangulares o cuadradas no pueden ser usadas para conjuntos de datos grandes.

**Label Matrix:** Ubica las marcas de las muestras en su matriz.

## AMOVA

El Análisis de Varianza Molecular se puede realizar con datos en bruto o datos en una matriz de distancia calculados previamente. Para datos crudos se proporcionan varios cálculos para generar matrices de distancia con datos binarios, haploides o codominantes.

### Procedimiento

1. Con la hoja de trabajo que contiene sus datos activa, escoja la opción *AMOVA* desde el menú de **GenAlEx**.
2. Ingrese todos los parámetros de datos requeridos en el cuadro de diálogo *AMOVA Data Parameters*, incluyendo poblaciones y tamaños regionales.
3. Seleccione el tipo de datos de ingreso (Input data Type). Una matriz de distancia calculada previamente puede ser un dato de ingreso en este punto.
4. Ingrese un título y un prefijo para la hoja de trabajo, luego pulse Ok.

5. En el cuadro de diálogo de *AMOVA Genetic Distance Options* seleccione el calculo de distancia apropiado y las opciones de respuesta requeridas (para más detalles ver abajo).

**Nota: Cuando se tienen conjuntos grandes el cálculo de la matriz de Distancia Genética puede tardar un tiempo, antes de que GenAlEx proceda al cuadro de diálogo de opciones de AMOVA.**

### ***AMOVA opciones de Distancia Genética***

#### **Distance Calculation (Calculo de la distancia)**

Escoja el cálculo de distancia requerido. Detalles de los cálculos son presentados en el Apéndice 1. Para datos codominantes hay tres cálculos diferentes:

**Codom - Genotypic:** Genera PhiPT, una medida que facilita la comparación entre dos codominantes, haploide/binario. Esta medida no considera la variación intra-individual.

**Codom - Allelic:** Estima los valores Fst estandares.

**Codom - Microsat:** Calcula Rst, un estimador de diferenciación genética para loci microsatélites, que asume el modelo de mutación ‘stepwise’.

#### **Distance Output Options (Opciones de respuesta de distancia)**

**Output Total Distance Only (Hallar solo la distancia total):** Genera una matriz de distancia genética sumando todos los loci. Para los tres métodos codominantes, las matrices de distancia para cada locus suma entre loci asumiendo el principio de independencia.

**Output Distance All Loci (Hallar la distancia de todos los loci):** Genera una matriz para cada locus de forma individual, más la matriz de distancia genética total, cada uno separado en una hoja de trabajo.

#### **Output (Resultados)**

Si selecciona *Output to worksheet* la distancia genética será un resultado en hojas [GD], [GDA] o [GDM], para distancias Genotípica, Alélica y Microsatélites respectivamente.

**To Worksheet (Hoja de trabajo):** La matriz se presenta en una hoja de trabajo. La generación de la matriz de distancia se recomienda solo para usuarios avanzados que quieren evaluar sus datos. Este resultado no es requerido para que GenAlEx realice la AMOVA.

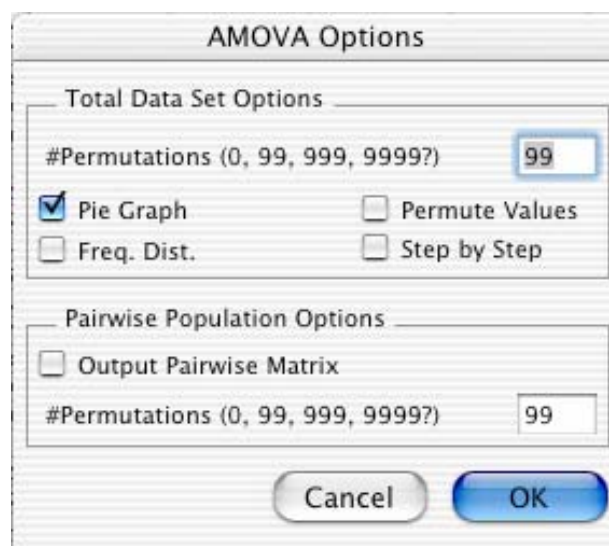
**As Tri Matrix:** Genera una matriz de distancia genética en forma de matriz triangular. Se recomienda su uso cuando se tienen datos de menos de 254 muestras.

**As Sq Matrix:** Genera una matriz de distancia genética con simetría cuadrada. Esta matriz es muy útil como 'Dato de ingreso' para otros programas que requieren matrices cuadradas.

**As Column:** Genera una matriz en columna. Esta opción es requerida cuando se tienen conjuntos de datos de más de 254 muestras. Debido a que el límite en Excel es de 256 columnas, las matrices triangulares o cuadradas no pueden ser usadas para un gran número de datos.

**Label Matrix (matriz de marcas):** Ubica las marcas apropiadas a su matriz.

6. Seleccione las opciones requeridas en el cuadro de diálogo AMOVA Options (ver detalles abajo) y pulse Ok. El análisis de AMOVA es una hoja de resultados [AM].



### **Opciones de AMOVA**

#### **Total Data Set Options (Opciones para el conjunto de datos total)**

**#Permutations (Permutaciones) [AM]:** Ingrese el número de permutaciones requeridas para evaluar significancia.

**Pie Graph [AM]:** Genera una gráfica de torta ilustrando la distribución de la varianza.

**Freq. Dist. [FDAM]:** Genera una gráfica con la frecuencia de distribución de los valores permutados de  $\Phi_{PT}/F_{st}/R_{st}$  vs el valor observado.

**Permute Values [PVAM]:** Genera los valores de AMOVA de cada permutación.

**Step by Step [SbyS AM]:** Presenta cada paso para el cálculo de la AMOVA.

**Pairwise Population Options (Opciones para poblaciones en parejas)**

**Output Pairwise Matrix [PW] & [PWT]:** Genera los valores de PhiPT/Fst/Rst entre todos los pares de poblaciones en formato de tri-matriz en hoja PW, y en formato de tabla en hoja PWT.

**# Permutations (# Permutaciones) [AM]:** Ingrese el número de permutaciones requeridas para evaluar significancia.

**Nota: Para un conjunto de datos grande las permutaciones pueden demorarse. Observe la barra de estado para ver el progreso de la acción. Para propósitos de publicación el número de permutaciones debe ser 999 o 9999.**

## Mantel

La opción de la prueba de Mantel en GenAEx permite la comprobación de una relación estadística entre los elementos de dos matrices. Por lo tanto para realizar el test de Mantel usted requiere dos matrices de distancia como datos de ingreso. Ejemplo: una matriz de distancia genética y una matriz de distancia geográfica. (Asegúrese de tener una matriz de Distancia Genética lineal y no cuadrada, (ver en *Distance* páginas anteriores)).

### Procedimiento

1. Primero calcule las matrices de distancia apropiadas por medio de la opción *Distance* en el menú GenAEx.
2. Asegurese de tener activada la hoja de trabajo que contiene su matriz X (ejemplo una matriz de Distancia Geográfica). Escoja la opción *Mantel* desde el menú **GenAEx**.
3. Ingrese toda la información requerida en el cuadro de diálogo **Mantel Parameters** (para más información sobre opciones de respuesta ver abajo).
4. Especifique una hoja de trabajo para la matriz de distancia y (ejemplo una matriz de Distancia Genética), ingrese un título y un prefijo para la hoja de resultados. Ingrese el número de permutaciones que puede ser 0, 99, 999, o 9999 y pulse *Ok*. El resultado es una hoja [MT].

### Opciones de Mantel

**XY Graph (Gráfico de XY) [MT]:** Genera un gráfico de los datos XY.

**Permute Values (valores permutados) [PVMT]:** Utilice esta opción para tener un resultado con los valores de Mantel de cada permutación.

**Freq. Dist. [FDMT]:** Esta opción genera resultados con la frecuencia de distribución de los valores Rxy permutados vs el valor Rxy observado.

## PCA

---

El análisis de coordenadas principales (en inglés Principal Coordinates Analysis), (PCA) es una técnica multivariada que permite encontrar y trazar los patrones dominantes en un conjunto de datos multivariados, ejemplo múltiples loci y múltiples muestras. GenAlEx puede trazar la relación entre los elementos de las matrices de distancia basado en sus dos primeras coordenadas principales, sin embargo hay opciones disponibles para trazar otras combinaciones de ejes (ver abajo).

Son aceptadas todas las matrices de distancia generadas con GenAlEx.

### Procedimiento

1. Primero calcule la matriz de distancia apropiada por medio de alguna de las siguientes opciones: *Distance*, *AMOVA*, o *Nei's Distance* (bajo *Frequency*). Si su matriz tiene etiquetas estas pueden ser trazadas en el plano PCA.
2. Asegurese de tener activada su hoja de trabajo que contiene la matriz de Distancia Genética. Escoja la opción *PCA* desde el menú de **GenAlEx**.
3. En el cuadro de diálogo *PCA Parameters*, seleccione el tipo de datos (Input Data Type) apropiado.
4. Seleccione su método preferido y las opciones de respuesta requeridas (para más detalles ver abajo).
5. Ingrese un título y un prefijo opcional y pulse OK. Un plano esparcido de las primeras dos coordenadas es resultado en la hoja de trabajo [PCA].

The image shows a dialog box titled "PCA Parameters" with the following fields and options:

- Input Data Type:** Radio buttons for "Tri Distance Matrix" (selected) and "Distance Matrix as Columnn".
- #Samples:** Text box containing "95".
- PCA Method:** Radio buttons for "Covariance-Standardized" (selected), "Covariance-Not Standardized", "Distance-Standardized", and "Distance-Not Standardized".
- Graph Options:** Checkboxes for "Data Labels" (checked) and "Color Code Pops" (unchecked).
- Title:** Text box containing "PCA excluding Cval".
- Worksheet Prefix:** Text box containing "ex Cval".

Buttons for "OK" and "Cancel" are located on the right side of the dialog.

**Nota: PCA es un procedimiento iterativo que puede demorarse cuando se tienen conjuntos de datos grandes. Observe la barra de estado para ver el progreso de la acción.**

### ***PCA methods (Métodos PCA)***

Como opciones se proporcionan cuatro métodos en PCA, diferentes pero relacionados. Todos los métodos producen esencialmente el mismo patrón, pero algunas agrupaciones pueden resolverse mejor que otras, dependiendo de los datos subyacentes. A usted podría gustarle experimentar con las opciones. Los dos primeros métodos son basados en la matriz de covarianza y los otros dos en la matriz de distancia.

### ***Graph options (Opciones Graficas)***

**Data labels (Marcos de datos):** Con esta opción pueden salir las marcas de las muestras en el grafico.

**Color Code Pops (Codificar poblaciones en color):** Con esta opción se puede codificar las poblaciones en color en el gráfico.

### **Trazar Primera vs Tercera y Segunda vs Tercera coordenadas.**

Excel todavía no proporciona la opción para trazar un gráfico en plano 3D. Sin embargo, las opciones de GenAlEx permiten gráficos de los segunda y tercera, y primera y tercera coordenadas.

1. Seleccione la hoja de trabajo que contiene un resultado PCA correcto.
2. Escoja los submenús requeridos (*PCA Axes 1 vs 2, PCA Axes 1 vs 3, PCA Axes 2 vs 3*) en la opción *Graph* del menú **GenAlEx**.

## Assignment (Asignación)

La opción de Asignación (*Assignment*) proporciona dos submenús, *Pop Assign* y *Sex Bias*. Estos análisis solo aplican a datos codominantes.

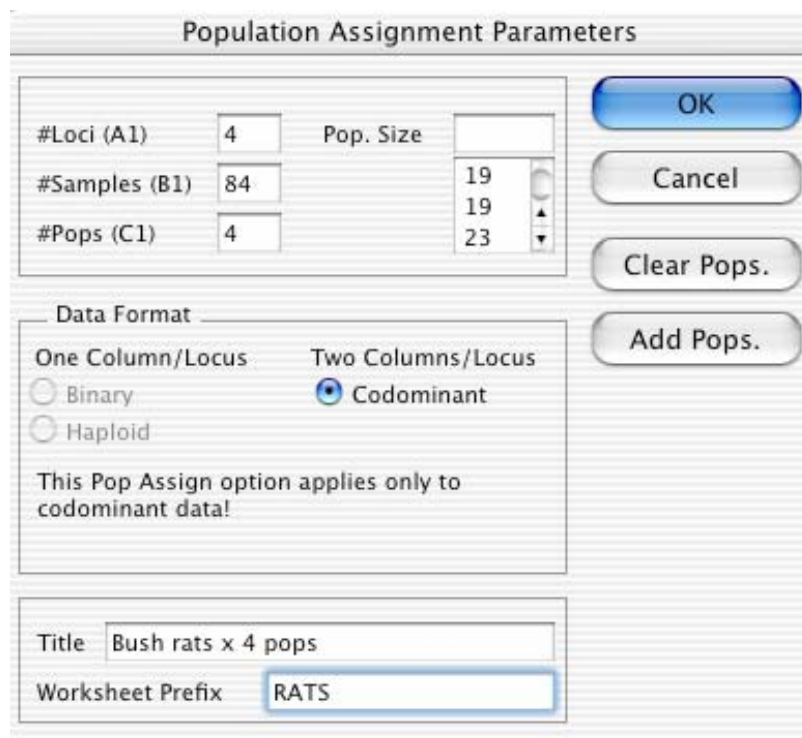
### Pop Assign (Asignación de población)

Esta opción es proporcionada principalmente para la enseñanza. Sin embargo las opciones de gráficos, únicamente disponible en GenAlEx, son útiles para la exploración previa de datos antes de analizarlos en otros programas de asignación.

*Recomendación: Los planos proporcionan una herramienta gráfica ideal para evaluar el poder de asignación de las pruebas. Ver apéndice 1 para más información y referencias.*

#### Procedimiento

1. Active la hoja de trabajo que contiene su conjunto de datos codominantes en formato GenAlEx. Escoja la opción *Assignment* desde el menú **GenAlEx**, y luego seleccione *Pop Assign* desde el submenú.
2. Asegúrese de tener correctos los parámetros locus y muestra en el cuadro de diálogo Population Assignment Parameters.
3. Ingrese un título y prefijo para la hoja de trabajo y pulse *Ok*.



**Population Assignment Parameters**

#Loci (A1) 4      Pop. Size 19

#Samples (B1) 84      19

#Pops (C1) 4      23

**Data Format**

One Column/Locus      Two Columns/Locus

Binary       Codominant

Haploid

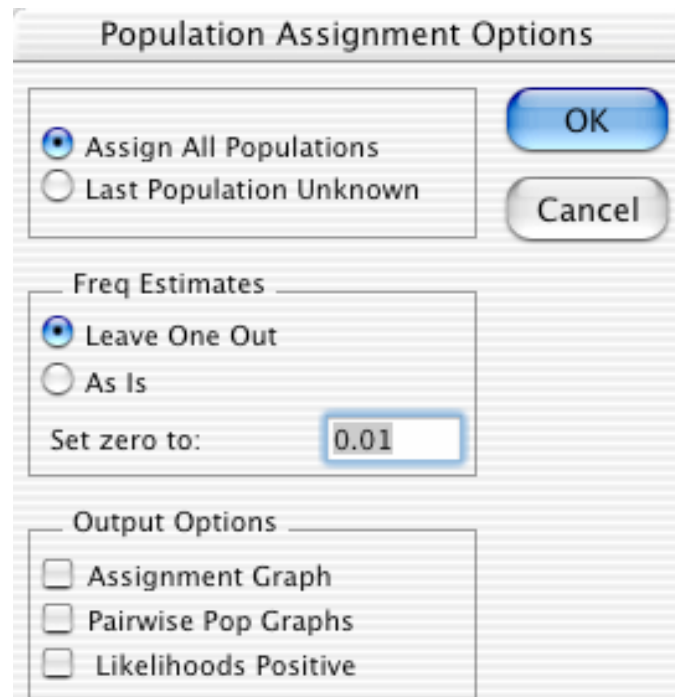
This Pop Assign option applies only to codominant data!

Title Bush rats x 4 pops

Worksheet Prefix RATS

Buttons: OK, Cancel, Clear Pops., Add Pops.

4. En el cuadro de diálogo subsiguiente Population Assignment Options verifique la opción requerida (para más detalles ver abajo), y pulse ok. El resultado es la hoja [ASS].



### ***Opciones de Asignación de poblaciones.***

**Assign All Populations (Asignar todas las poblaciones):** Calcula la asignación de todas las muestras.

**Last Population Unknown (Última población desconocida):** Trata la última población como muestras desconocidas y calcula su asignación a las poblaciones precedentes.

### **Freq Estimates (Estimación de frecuencias)**

**Leave One Out (Omitir uno):** Este es el procedimiento recomendado que incluye la corrección del sesgo para la frecuencia de la población. En este caso, el individuo en cuestión es removido del conjunto de datos, antes de calcular las frecuencias ajustadas que se usaran para estimar la probabilidad asignada.

**As Is (como es):** Incluye la muestra en cuestión cuando se calcula la frecuencia. Esta opción se mantiene para usar en clases, y para proporcionar compatibilidad con el procedimiento de sesgo sexual que no hace esta corrección.

**Set Zero to (Ajustar el cero):** La prueba de asignación no puede aceptar una frecuencia de cero, por esto requiere un valor. Ingrese un valor en el rango de 0.01 a 0.00001. Lo sugerido por GenAlEx es 0.01, siguiendo la recomendación de Paetkau et al (2004).

### **Output Options (Opciones de Respuesta)**

**Assignment Graph (gráfico de asignación) [ASS]:** Genera un gráfico en dos dimensiones, asignando índices para toda la población con base en las frecuencias alélicas para las poblaciones 1 y 2. Esta es una herramienta para explorar fácilmente los datos.

**Pairwise Pop Graphs (gráfico de parejas de poblaciones) [PWASS]:** Genera gráficos separados para todas las parejas poblacionales. Estos gráficos proveen una representación visual del grado de separación genética entre las poblaciones, y aportan una vía ideal para evaluar el poder de las pruebas de asignación (Paetkau et al 2004).

**Likelihoods Positive (Máxima similitud Positiva):** Cambia el log de máxima similitud a valores positivos, donde el valor más bajo indica la población más ligada. Esto facilita la presentación y puede ser una vía más fácil para que los estudiantes interpreten el significado de los valores del log de máxima similitud.

## **Sex Bias (Efecto Sesgado del sexo)**

Esta opción calcula índices ajustados para los conjuntos de datos en los cuales el sexo del animal es identificado. Las comparaciones de la distribución de los índices entre los sexos permite la detección de sesgos al sexo que estén dispersos.

EL efecto del sexo solo puede ser calculado en conjuntos de datos que representan una sola población. El sexo de cada muestra se debe ingresar como M o F en la columna B. No hay opciones para valores de sexo faltante, entonces es necesario que los conjuntos de datos estén completos.

### **Procedimiento**

1. Active la hoja de trabajo que contiene una sola población, en el formato de datos codominantes de GenAlEx. Escoja la opción *Assignment* desde el menú **GenAlEx** y luego seleccione *Sex Bias* desde el submenú.
2. Asegurese de que los parámetros del locus y las muestras estén correctos en el cuadro de diálogo **Population Assignment Parameters**.
3. Ingrese el Título y el Prefijo de la hoja de trabajo, luego pulse *Ok*.
4. Los resultados salen en dos hojas de trabajo. Los valores de asignación de población para todos los individuos, junto con el promedio para machos y hembras, y una gráfica en la hoja de trabajo [SB] y la distribución de frecuencia del efecto de los sexos en la hoja [FDSB].

**Sex Bias Parameters**

#Loci (A1)	7	Pop. Size	46
#Samples (B1)	46		
#Pops (C1)	1		

Buttons: OK, Cancel, Clear Pops., Add Pops.

**Data Format**

One Column/Locus       Two Columns/Locus  
 Binary                       Codominant  
 Haploid

This Sex Bias option applies only to codominant data!

Title: Sex Bias for Pop. CAMM

Worksheet Prefix: CAMM

## Procedimientos estadísticos avanzados

### Spatial Autocorrelation (Autocorrelación Espacial)

#### ***Sobre los métodos de Autocorrelación espacial***

GenAlEx ofrece un amplio rango de opciones para análisis de autocorrelación espacial, empleando nuevos procedimientos multivariados desarrollados por los autores de GenAlEx. El Apéndice 2 proporciona descripciones detalladas y información bibliográfica y debe leerse en conjunto con esta guía.

Los análisis espaciales globales que ofrece GenAlEx son: *Single Pop...*, para analizar una sola población usando una matriz de distancia genética; *Multiple Loci...* para el análisis por separado de múltiples matrices de distancia genética (desde loci múltiples) con solo una matriz geográfica; *Multiple Pops...* para análisis de autocorrelación sobre poblaciones múltiples; *Multiple Dclass...* para análisis de autocorrelación de múltiples poblaciones sobre múltiples clases de distancias, *Pops as Dclass...* para comparar la autocorrelación genética entre poblaciones múltiples.

También proporciona una opción para análisis espacial “local”: el análisis espacial en 2 Dimensiones (*2D LSA*). La opción *NN Dist...*, esta disponible como un complemento de *2D LSA*, para proveer un resumen del vecino más cercano (Nearest Neighbors), y su distancia desde cada muestra, hasta un número especificado de vecinos más cercanos.

Con excepción de *Pops as Dclass* y el *NN Dist*, todas las opciones del submenú requieren las matrices de distancia genética normal en GenAlEx [GD] y distancia geográfica [GGD] en hojas de trabajo separadas como datos de ingreso. En el caso de las opciones *Multiple loci*, *Multiple pops* y *Multiple Dclass*, se requieren matrices de distancia genética múltiples. En todos los casos, los formatos de matrices y los parámetros deben estar en el formato de GenAlEx. Estos pueden ser generados por la opción *Distance* en GenAlEx.

Para algunos de los análisis espaciales en GenAlEx las hojas con los datos de ingreso deben estar en un orden específico. Esto es detallado abajo. Donde es posible, GenAlEx puede ubicar estas hojas automáticamente en el formato correcto, en el momento en que se generan las matrices de distancia.

## **Single Pop...(Población Única)**

Use esta opción para realizar una Autocorrelación Espacial para una sola población. Los hojas de datos de ingreso consisten de una sola matriz de distancia, que típicamente representa el total de la distancia genética sobre loci múltiples, y la matriz de distancia geográfica para el mismo conjunto de muestras.

### **Procedimiento**

1. Primero calcule las matrices de la Distancia Genética [GD] y la Distancia Geográfica [GGD] apropiadas, usando la opción *Distance* en el menú **GenAlEx**. Si sus datos son codominantes, asegurese de seleccionar la distancia 'Genotípica'.
2. Es recomendable que la hoja de distancia geográfica [GGD] sea la primera en el libro de trabajo, seguida por la hoja de distancia genética [GD].
3. Habilite la hoja de trabajo que contiene su matriz de Distancia Genética [GD]. Escoja la opción *Spatial* desde el menú **GenAlEx**, y luego seleccione *Single Pop...* desde el submenú.
4. En el cuadro de diálogo **Single Pop Spatial Structure Parameters**, seleccione el formato de datos apropiado, luego ingrese el número de muestras, el número de permutaciones, y el número de bootstraps requerido (0, 99, 999, 9999) (ver detalles abajo).
5. Escoja su opción de clase de distancia preferida, y seleccione la hoja de trabajo de la matriz de Distancia Geográfica. Ingrese un Título y Prefijo para la hoja de trabajo, luego pulse *Ok*. El resultado del análisis espacial es una hoja de trabajo [R], después de escoger sus opciones de clase de distancia en el cuadro de diálogo siguiente.

Single Pop Spatial Structure Parameters

Input Data Type

Tri Distance Matrix

Distance Matrix as Column

OK

Cancel

#Samples 38

#Permutations (0, 99, 999, 9999?) 99

#Bootstraps (0, 99, 999, 9999?) 99

Options

Even Distance Classes  Full output

Variable Distance Classes

Even Sample Sizes

Worksheet Names

Genetic Distance MD GD

Geographic Distance MD GGD

Title Site MD with 4 loci for spatial

Worksheet Prefix Site MD

**Nota: Cuando se tiene un conjunto de datos grande este análisis puede tardar un tiempo, mientras GenAlEx lee la matriz GGD y calcula información sobre los datos. Observe el estado de la barra de progreso**

### **Opciones de estructura espacial para una sola población**

**#Samples (#muestras):** Ingrese el número de muestras. Si la matriz esta en formato GenAlEx este se inserta utomaticamente.

**#Permutations (#permutaciones):** Ingrese el número de permutaciones requerido para la prueba de significancia. Para propositos de publicación el número de permutaciones debe ser 999 o 9999.

*Nota: Para un conjunto de datos grande, las permutaciones tardan un poco, observe el estado de la barra de progreso.*

**#Bootstraps:** Ingrese el número de bootstraps requerido para estimar un intervalo de confianza del 95% alrededor de  $r$ . Para propositos de publicación el número de bootstraps debe ser 999 o 9999.

**Full output (resultados totales):** Presenta todos los resultados estadísticos, incluso los resultados de los análisis de permutación y bootstrap en la hoja [R]. Genera los calculos del Análisis Espacial paso por paso en la hoja [SbyS R].

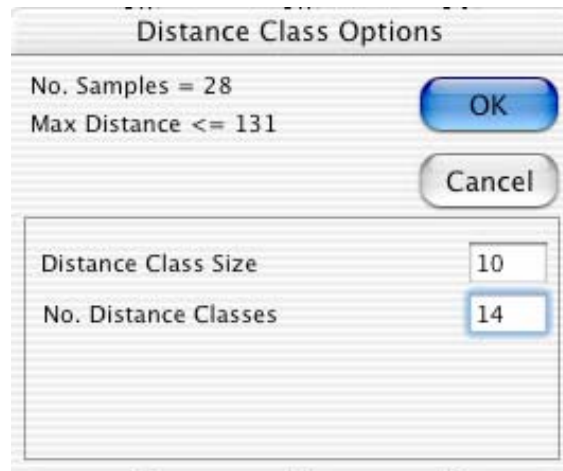
### **Distance class options (Opciones de clases de distancia)**

GenAlEx ofrece 3 diferentes opciones de clases de distancia, las cuales proveen flexibilidad para definir los tamaños y limites de las clases de distancia. Dependiendo de su escogencia para la opción de clase de distancia, pueden seguir uno o más cuadros de diálogo.

**1. Even Distance Classes (clases de distancia uniformes):** Esta opción creará clases de distancia geográfica de igual tamaño. El análisis espacial considerará todas las muestras que estan representadas por una distancia más grande que la distancia previa, y menor o igual a la clase de distancia mayor. Ejemplo, con un tamaño de clase de distancia (Distance Class size) de '1' y un número (No. of Distance Classes) de '8', son incluidas todas las muestras con una distancia geográfica de  $>1$  y  $\leq 8$ . La única excepción a esta regla es que la primera clase de distancia incluye un valor de cero hasta un valor  $\leq$  a la primera clase de distancia.

### **Procedimiento**

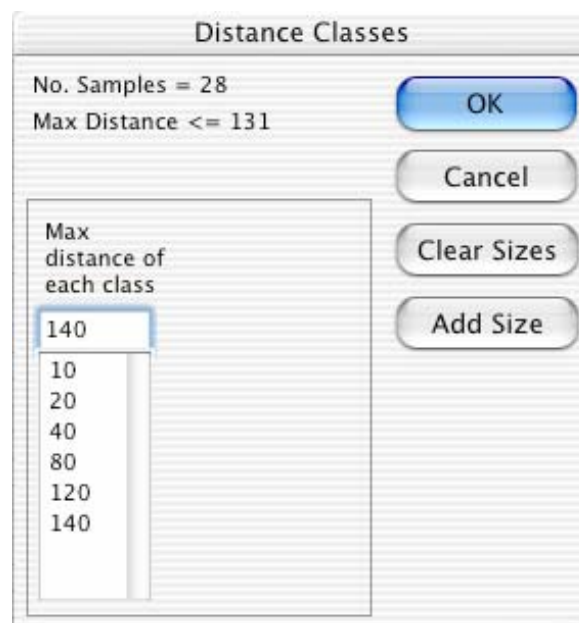
1. En el cuadro de diálogo Distance Class Options ingrese el valor numérico para la clase de distancia y el número de clases de distancia. Utilice la información proporcionada sobre la distancia máxima en la matriz, para guiar lo que usted escoja.



**2. Variable Distance Classes (clases de distancias variables):** Esta opción puede ser usada para crear manualmente clases de distancia de tamaños desiguales. Como en el análisis de clases de distancia iguales, el análisis espacial considerará todas las muestras que estén representadas por una distancia mayor que la clase de distancia previa y menor o igual a la clase de distancia superior, con excepción de la primera clase de distancia.

### Procedimiento

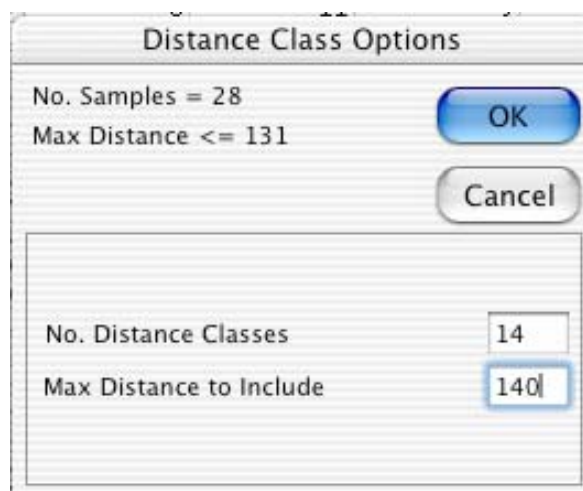
1. En el cuadro de diálogo **Distance Classes** ingrese el tamaño de la clase de distancia deseada y pulse el botón *Add Size*.
2. Repita lo anterior hasta que hayan sido ingresados todas las clases de distancia. Se pueden usar valores decimales.



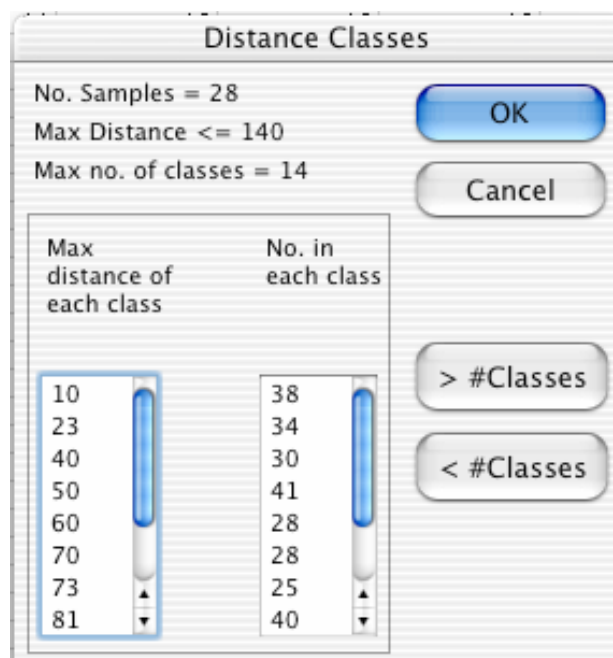
- 3. Even Sample Classes (clases de muestras iguales):** Con esta opción el análisis espacial selecciona clases de distancia que tengan en lo posible el número de muestras iguales en cada clase de distancia. Esto es particularmente útil para reducir el ruido en los límites de confianza cuando los tamaños de las muestras son muy desiguales.

### Procedimiento

1. En el cuadro de diálogo Distance Class Options proporciona el número de clases de distancia y la distancia máxima a incluir.



2. El cuadro de diálogo siguiente Distance Classes presenta una lista de clases de distancia. Pulse el botón para incrementar o disminuir el número de clases de distancia requerido.



## ***Multiple loci... (Loci Multiples)***

---

Esta opción está diseñada para el análisis automático de loci múltiples, usando el mismo conjunto de muestras. Los datos de ingreso son matrices de distancia genética, cada una deriva de un solo locus, con solo una matriz de distancia geográfica. Los resultados incluyen un gráfico de correlación con los resultados de cada locus entrecruzados. Seleccionando la opción de resultados totales se puede obtener un gráfico de correlación para cada locus individual.

### ***Procedimiento***

1. Calcule las matrices de Distancia Genética [GD] y Distancia Geográfica apropiada [GGD] por la opción *Distance* en el menú **GenAlEx**. Si sus datos son codominantes asegúrese que la matriz GD sea para distancia genotípica. Para tener matrices de distancia genética separadas para cada locus, señale la opción *Output Distance All Loci* en el cuadro de diálogo *Distance Options*.
2. Asegúrese que las hojas de trabajo de Distancia Genética por loci desde 1 a n, estén ubicados en el libro de trabajo en las posiciones 1 a n. La hoja de trabajo con la matriz de Distancia Geográfica se localizan después de las hojas de distancia genética.
3. Active la hoja de trabajo con la primera matriz de distancia [GD]. Escoja la opción *Spatial* desde el menú **GenAlEx**, y después seleccione *Multiple loci..* desde el submenú.
4. Ingrese el número de hojas de distancia genética que usted desea analizar en el cuadro de diálogo GenAlEx siguiente. (Normalmente este debe ser igual al número de loci + 1, incluyendo la matriz de distancia genética total).
5. En el cuadro de diálogo *Multiple Loci Spatial Structure Parameters*, seleccione el formato de datos de ingreso apropiado, luego ingrese el número de muestras, el número de permutaciones, y el número de bootstraps requeridos (0, 99, 999, 9999) (para más detalles ver arriba las opciones de análisis de Estructura Espacial en una Población [Single Pop]).
6. Especifique la hoja para la matriz de distancia geográfica. Ingrese un título corto y un prefijo, luego pulse *Ok*. El análisis espacial combinado está resultado en la hoja de trabajo [RML], después de escoger sus opciones de clase de distancia en el cuadro de diálogo siguiente.

<p><b>Nota: con un conjunto de datos grande y muchos loci el análisis puede tardar un tiempo. Observe el estado de la barra de progreso.</b></p>
--

### ***Opciones para estructura espacial con múltiples loci***

**#Samples (#muestras):** Ingrese el número de muestras. Este se ingresa automáticamente si la matriz de distancia esta en formato GenAlEx.

**#Permutations (#Permutaciones):** Ingrese el número de permutaciones requeridas en la prueba de significancia. Para propósitos de publicación el número de permutaciones pueden ser de 999 a 9999.

*Nota: para un conjunto de datos grande, las permutaciones pueden tardar un tiempo. Observe el estado de la barra de progreso.*

**#Bootstraps:** Ingrese el numero de bootstraps requerido para estimar un intervalo de confianza del 95% alrededor de  $r$ . Para propósitos de publicación el número de bootstraps debe ser de 999 o 9999.

**Full output (Output total):** Los outputs del análisis, incluyen gráficos de correlación, para cada locus individual en hojas de trabajo separadas. Los nombres de las hojas son asignados automáticamente por GenAlEx usando los nombres de las matrices de las hojas de trabajo de Distancia Genética [Locus1\_R].

### ***Distance class options (Opciones de clases de distancia)***

Las opciones de clases de distancia son iguales a las ofrecidas para el análisis de una población *Single Pop.*

## **Multiple Pops... (Poblaciones múltiples)**

---

Esta opción permite realizar una autocorrelación sobre múltiples poblaciones. Este análisis puede detectar si un proceso común está generando una estructura en poblaciones diferentes. Adicionalmente, los tamaños muestrales son maximizados dando poder estadístico para detectar hasta una estructura sutil, si esta existe.

*Recomendación: GenAlEx ofrece una opción fácil para dividir los conjuntos de datos que contienen múltiples poblaciones en poblaciones separadas como se requiere para este análisis, usando la opción Split by pop... (separación de poblaciones) en el menú de opciones **Data**.*

### **Procedimiento**

1. Primero calcule las matrices de Distancia Genética [GD] y Distancia Geográfica [GGD], para cada población, por medio de la opción *Distance* en el menú **GenAlEx**. Si sus datos son codominantes, asegúrese de escoger la Distancia Genotípica.
2. Estas hojas deben ser ordenadas dentro de un libro de trabajo de la siguiente manera: la matriz de distancia geográfica (GGD) para la población 1, seguido de la matriz de distancia genética (GD) para la población 1, GGD Pop 2, GD Pop 2 ... GGD Pop n, GD Pop n. GenAlEx no reviza si sus hojas están en el orden correcto, asegúrese de revizar el orden de sus hojas antes de los análisis.
3. Active la hoja que contiene su primer matriz de Distancia Genética [GD]. Escoja la opción *Spatial* desde el menú **GenAlEx**, y luego seleccione *Multiple loci.* desde el submenú.
4. Cuando entre por el cuadro de diálogo **Input for Spatial Multiple Pops**, ingrese el número de poblaciones que usted desee analizar. Cada población requiere un GGD y un GD en hojas consecutivas separadas, como se describe arriba.
5. En el cuadro de diálogo **Multiple Loci Spatial Structure Parameters**, seleccione el formato de datos de ingreso apropiado, luego ingrese el número de muestras, el número de permutaciones, y el número de bootstraps requeridos (0, 99, 999, 9999) (para más detalles ver arriba las opciones para análisis de Estructura Espacial en una Población [Single Pop]).
6. Escoja su opción de clase de distancia requerido y ingrese un título y prefijo para la hoja de trabajo, luego pulse *Ok*. El análisis de estructura espacial para cada población es un resultado en la hoja de trabajo [RMP], y el análisis combinado para todas las poblaciones es un resultado en la hoja [RC], después de escoger sus opciones de clase de distancia en el cuadro de diálogo siguiente.

<p><b>Nota: con un conjunto de datos grande y muchos loci el análisis puede tardar un tiempo. Observe el estado de la barra de progreso.</b></p>
--

**Multiple Pops Spatial Structure Parameters**

Input Data Type

Tri Distance Matrix

Distance Matrix as Column

OK

Cancel

#Samples: 46

#Permutations (0, 99, 999, 9999?): 99

#Bootstraps (0, 99, 999, 9999?): 99

Options

Even Distance Classes  Full output

Variable Distance Classes

Worksheet Names

Multi pop analysis for Sheet 1 to Sheet 4

No pops=2

Title: CAMM + MD multiple pops, even DC

Worksheet Prefix: CAMM+MD

### **Opciones de estructura espacial con múltiples loci**

**#Samples (# muestras):** Ingrese el número de muestras. Este se ingresa automáticamente si la matriz de distancia esta en formato GenAlEx.

**#Permutations (#Permutaciones):** Ingrese el número de permutaciones requeridas en la prueba de significancia. Para propósitos de publicación el número de permutaciones pueden ser de 999 a 9999.

*Nota: para un conjunto de datos grande, las permutaciones pueden tardar un tiempo. Observe el estado de la barra de progreso.*

**#Bootstraps:** Ingrese el número de bootstraps requerido para estimar un intervalo de confianza del 95% *r*. Para propósitos de publicación el número de bootstraps debe ser de 999 o 9999.

**Full output (resultados totales):** Resultados: 1. toda la estadística para el análisis de estructura espacial para cada población en una hoja de trabajo [RMP]; 2. toda la estadística para el análisis espacial de cada población individual en hojas de trabajo apropiadamente nombradas [Pop1\_R]; 3. La estructura espacial combinada para cada población para clase de distancia 1 [DC1].

### **Distance class options (Opciones de clases de distancia)**

Para una descripción de las opciones de clases de distancia vea el procedimiento de análisis espacial *Single Pop.*

## ***Multiple Dclass ... (Clases de distancia múltiples)***

---

Esta opción realiza una autocorrelación espacial sobre medidas de clases de distancia múltiples para una población o múltiples poblaciones. Este análisis es diferente al análisis espacial normal en que el análisis se realiza con un incremento automático en las medidas de clases de distancia. Así, el análisis es equivalente a recomenzar repetidamente un análisis espacial de un población sola o poblaciones múltiples con una clase de distancia diferente.

El resultado *rc* (*r combinado*) para cada clase de distancia tiene una pequeña corrección por sesgo que varía entre cada corrida. Si usted desea comparar los resultados a las correlaciones normales, asegúrese de utilizar los valores *rc* sin corrección en la tabla debajo del gráfico de autocorrelación (no el valor *r* trazado, el cual tiene aplicada la corrección).

Este análisis permite la exploración de la interacción entre el tamaño de muestra y el tamaño de la clase de distancia, permitiendo determinar hasta que punto es perceptible la estructura genética.

### **Procedimiento**

1. Primero calcule las matrices apropiadas de distancia Genética [GD] y Geográfica [GGD], por medio de la opción *Distance* en el menú **GenAlEx**. Si sus datos son codominantes, asegúrese que escoge la Distancia Genotípica.
2. Estas hojas deben ser ordenadas dentro de un solo libro de trabajo de la siguiente manera: la matriz de distancia geográfica (GGD) para la población 1, seguido de la matriz de distancia genética (GD) para la población 1, GGD Pop 2, GD Pop 2 ... GGD Pop n, GD Pop n. GenAlEx no revisa si sus hojas están en el orden correcto, asegúrese a revizar el orden de sus hojas antes de los análisis.
3. Active la hoja que contiene su primer matriz de Distancia Genética [GD]. Escoja la opción *Spatial* desde el menú **GenAlEx**, y luego seleccione *Multiple Dclass.* desde el submenú.
4. Cuando entre por el cuadro cuadro de diálogo Input for Spatial Multiple Dclass, ingrese el número de poblaciones que usted desea analizar. Cada población requiere un GGD y un GD en hojas consecutivas separadas, como se describe arriba.
5. En el cuadro de diálogo Multiple Dclass Spatial Structure Parameters, seleccione el formato de datos de ingreso apropiado, luego ingrese el número de muestras, el número de permutaciones y el número de bootstraps requeridos (0, 99, 999, 9999). (para más detalles ver arriba las opciones para análisis de Estructura Espacial en una Población [Single Pop]).
6. Escoja su opción de clase de distancia requerido e ingrese un título y un prefijo para la hoja de trabajo, luego pulse *Ok*. El análisis de estructura espacial sobre clases de distancia múltiple, es un resultado en la hoja de trabajo [MDC].

<p><b>Nota: con un conjunto de datos grande y muchos loci el análisis puede tardar un tiempo. Observe el estado de la barra de progreso.</b></p>
--

### ***Opciones de estructura espacial con clases distancia múltiple***

**#Samples (# muestras):** Ingrese el número de muestras. Este se ingresa automáticamente si la matriz de distancia esta en formato GenAlEx.

**#Permutations (#Permutaciones):** Ingrese el número de permutaciones requeridas en la prueba de significancia. Para propósitos de publicación el número de permutaciones pueden ser de 999 a 9999.

*Nota: para un conjunto de datos grande, las permutaciones pueden tardar un tiempo. Observe el estado de la barra de progreso.*

**#Bootstraps:** Ingrese el número de bootstraps requerido para estimar un intervalo de confianza del 95% *r*. Para propósitos de publicación el número de bootstraps debe ser de 999 o 9999.

**Full output (Resultados totales).** Además de los resultados del análisis de clases de distancia múltiples, en la hoja de trabajo [MDC], esta opción genera por separado los análisis de clase de distancia para cada población, en hojas de trabajo apropiadamente nombradas [ejemplo 0-100 DC].

**Distance class options (Opciones de clase de distancia)**

**1. Even Distance Classes (Clases de distancia uniformes):** Esta opción genera el cuadro de diálogo de las opciones de clases de distancia Distance Class Options, en el cual se ingresa el tamaño base de la clase de distancia  $s$ . Los tamaños subsecuentes de las clases de distancia para este análisis se calculan como  $run\ no * s$  para el número de corridas que usted especificó. Por ejemplo, si usted fijó el tamaño base en 50 m, para 3 corridas, el primer análisis calculará el  $rc$  ( $rc = r$  para un análisis combinado de la población) basado en todas las comparaciones en parejas, con una distancia geográfica de 0 a 50 M. En el segundo análisis, el  $rc$  será calculado para todas esas comparaciones en parejas con una distancia geográfica de 0 a 100 M. Para el tercer análisis el  $rc$  será calculado para todas esas comparaciones en parejas con una distancia geográfica de 0 a 150 M

**2. Variable Distance Class (Clases de distancia variable):** Esta opción permite que usted fije manualmente los tamaños de la clase de distancia.

## **Pops as Dclass ... (Poblaciones como clases de distancia)**

Esta opción realiza el equivalente de la opción *Multiple Dclass*, pero con las clases de distancias definidas por parámetros normales de una población en GenAlEx, en lugar de distancias en parejas geográficas. Las comparaciones también se pueden delimitar por la distancia entre muestras. En este caso, y cuando se delimita por tamaño, esta opción requiere matrices genéticas y geográficas de la distancia como ingreso. Observe que los valores de  $r$  obtenidos en este análisis se pueden diferenciar de los obtenidos por otras opciones tales como *Single Pop* o *Multiple Pops*, los cuáles utilizan clases de distancia geográficas.

*Recomendación: Esta opción es útil para producir rápidamente un gráfico que compare los valores  $r$  entre machos y hembras en una misma población. En este caso, los dos sexos son definidos como dos poblaciones diferentes.*

### **Procedimiento**

1. Primero calcule las matrices apropiadas de distancia Genética [GD] y Geográfica [GGD], por medio de la opción *Distance* en el menú **GenAlEx**. Asegurese que la matriz GD incluya los parámetros de la población.
2. Estas hojas deben ser ordenadas dentro de un solo libro de trabajo con la matriz Geográfica [GGD] en la posición inicial, seguida por la matriz de distancia genética [GD]. GenAlEx no verifica si sus hojas estan en el orden correcto, asegurese antes de analizar.
3. Active la hoja que contiene su primera matriz de Distancia Genética [GD]. Escoja la opción *Spatial* desde el menú **GenAlEx**, y luego seleccione *Pops as Dclass..* desde el submenú.
4. En el cuadro de diálogo **Pops as Dclass Data Parameters**, se deben ingresar automáticamente las muestras y los parámetros de la población. Ingrese el tipo de datos de entrada. Ingrese un título y prefijo para la hoja de trabajo y luego pulse *Ok*.

Parameter	Value
#Loci (A1)	1
#Samples (B1)	38
#Pops (C1)	2
#Regions	1

Pop. Size	Region
19	38
19	

Input Data Type:  Tri or Square Distance Matrix  Distance Matrix as Column

Title: Site MD with 4 loci for spatial  
Worksheet Prefix: sex

5. En el cuadro de diálogo **Pop Dclass Parameters**, ingrese el número de permutaciones y bootstraps requeridos (0, 99, 999, 9999).
6. Si usted desea restringir las comparaciones adentro de las poblaciones, ingrese la distancia deseada en el cuadro de edición. Para clases de distancia múltiple, ingrese el número de corridas y seleccione el método para incrementar las clases de distancia en estas corridas, doblando el tamaño en cada corrida o incrementándolo por la distancia nominada.

Pop Dclass Parameters

#Permutations (0, 99, 999, 9999?) 999

#Bootstraps (0, 99, 999, 9999?) 999

Distance Options

Distance (leave empty if not required) 50

#Runs (1 or more) 4

Double Size  Increment Size

Worksheet Names

Genetic Distance sex GD

Geographic Distance sex GGD

Cancel OK

7. Seleccione la hoja de trabajo que contiene la matriz de distancia geográfica y pulse ok. Cada corrida del análisis es un resultado en hojas separadas [RP1...n].

## **2D LSA... (Análisis espacial local en 2D)**

---

Este procedimiento realiza un análisis de autocorrelación espacial local en dos dimensiones. Para más detalles ver apéndice 2.

En adición a las matrices de distancia genética y geográfica, el análisis requiere una tercer hoja de trabajo, que contiene las coordenadas XY en el mismo orden de las muestras usadas para generar las matrices de distancia. Esta hoja de trabajo debe estar en formato GenAlEx, con las coordenadas ubicadas en las columnas 3 & 4.

### **Procedimiento**

1. Primero calcule las matrices de distancia Genética [GD] y la Distancia Geográfica [GGD] por medio de la opción *Distance* en el menú **GenAlEx**. Asegurese que la matriz [GD] incluya los parámetros poblacionales.
2. Prepare una tercera hoja que contenga los datos XY correspondientes a las mismas muestras de las matrices de distancia. Es mas conveniente ubicar estos datos en las columnas 3 & 4 con las etiquetas de muestras en la columna 1.
2. Estas hojas deben estar en el formato normal de GenAlEx y debe estar en un solo archivo con la matriz de distancia geográfica [GGD] en la primera posición, seguida por la matriz de distancia genética [GD], luego los datos XY.
3. Ubíquese en la hoja de trabajo que contiene la matriz de distancia genética [GD]. Escoja la opción *Spatial* en el menú **GenAlEx**, luego seleccione *2D LSA..* desde el submenú.
4. En el cuadro de diálogo **2D LSA Parameters**, seleccione el tipo de datos de ingreso apropiado. Ingrese el número de muestras, y el número de permutaciones requeridas. Seleccione la permutación condicional o corridas múltiples (para mas detalles ver abajo).
5. Seleccione las opciones de resultados que requiera (Para mas detalles ver abajo).
6. Seleccione las hojas de trabajo apropiada para el análisis. Ingrese un título y un prefijo y luego pulse *Ok*.

**2D LSA Parameters**

Input Data Type

Tri Distance Matrix

Distance Matrix as Column

#Samples: 74

#Permutations (0, 99, 999, 9999?): 99

Conditional Permute  Multi Runs

Output

Sort on Prob.  Sort on R Values

1-Tail Prob.  2-Tail Prob.

Optional Correlogram

Worksheet Names

Genetic Distance: GD

Geographic Distance: GGD

Sample XY Coords: XY

Title: Wrens 2D LSA

Worksheet Prefix: Wrens

### ***Opciones de análisis espacial local en 2 dimensiones (2D LSA).***

**Conditional Permute (permutación condicional):** Seleccione esta opción para las permutaciones condicionales.

**Multiple Runs (corridas múltiples):** Seleccione esta opción para realizar múltiples corridas incrementando el número de vecinos cercanos. Escogiendo esta opción cambiará el cuadro de diálogo subsecuente (ver abajo).

**Sort on Prob (clasificación por probabilidad).** Clasifica datos del más pequeño al más grande valore de probabilidad. Ejemplo, ubicar los valores de probabilidad más pequeños en el inicio. Generalmente se ubican los valores más grandes de  $r$  en el inicio, salvo por efectos del tamaño de muestra.

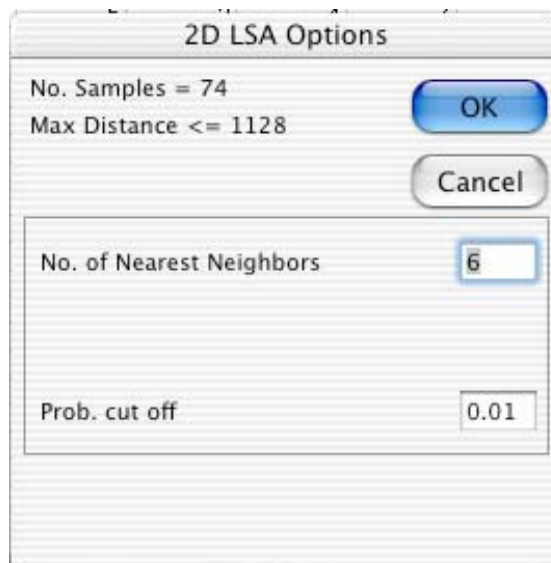
**Sort on  $r$  Values (clasificación en valores  $r$ ).** Clasifica por valores de  $r$ , moviendo los valores más grandes de  $r$  al inicio.

**1-Tail Prob. (Probabilidad de una cola).** Elija esta opción para etiquetar en los gráfico los valores mayor o igual al limite de probabilidad especificado (ingresado en el siguiente cuadro de diálogo).

**2-Tail Prob. (Probabilidad de dos colas).** Elija esta opción para etiquetar en el gráfico los valores mayores o igual al límite de probabilidad especificado dividido por dos, para los valores positivos y negativos.

**Optional Correlogram (gráfico de correlación opcional):** Demuestra la correlación con los resultados individuales de 2D LSA, enumerados desde la muestra 1 a n, sobre el eje x.

7. En el cuadro de diálogo 2D LSA Options ingrese el número de vecinos cercanos y el límite de probabilidad requerida (ver detalles abajo). El resultado es [R2D1].

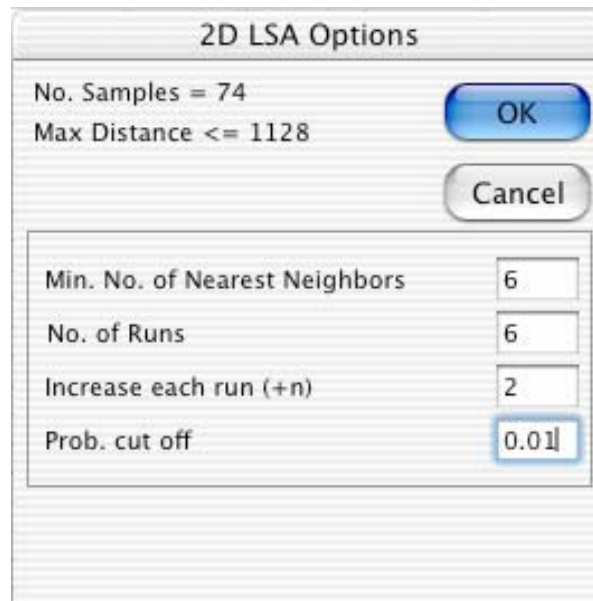


**No of Nearest Neighbors (número de vecinos cercanos):** Ingrese un valor para un número significativo de vecinos cercanos en su conjunto de datos.

*Recomendación:* Puede ser provechoso utilizar la opción de distancia de vecinos cercanos (NN Dist) previo al análisis para tener una guía de este valor.

**Prob. Cut off (límite de probabilidad):** Ingrese el límite de probabilidad deseado, menor o igual a 0.05.

8. Si usted seleccionó *Multiple Runs (corridas múltiples)*, ingrese el número mínimo de vecinos cercanos, el número de corridas requeridas, el incremento para cada corrida y el límite de probabilidad en el cuadro de diálogo subsecuente 2D LSA Options (ver detalles abajo). El resultado es [R2D1] y [R1 to Rn].



2D LSA Options

No. Samples = 74  
Max Distance <= 1128

OK

Cancel

Min. No. of Nearest Neighbors 6

No. of Runs 6

Increase each run (+n) 2

Prob. cut off 0.01

### **Opciones de 2D LSA para Corridas Múltiples**

**No of Nearest Neighbors (número de vecinos cercanos):** Ingrese un valor para un número significativo de vecinos cercanos en su conjunto de datos.

*Recomendación: Puede ser provechoso utilizar la opción de distancia de vecinos cercanos (NN Dist) previo al análisis para tener una guía de este valor.*

**No of Runs (número de corridas):** Ingrese el número de corridas. Para cada corrida, GenAlEx aumentará el número de los vecinos más cercanos en el valor incorporado abajo.

**Increase each run (+n) (incremente cada corrida en n+):** Incorpore un valor para aumentar el número de los vecinos más cercanos en cada corrida.

**Prob. cut off (límite de probabilidad):** Ingrese el límite de probabilidad deseado, menor o igual to 0.05.

## **NN Dist... (Distancia de Vecinos más Cercanos)**

Esta opción es proporcionada para complementar los resultados de 2D LSA, generando un resumen de los vecinos más cercanos y de sus distancias de cada muestra, hasta el número de vecinos más cercanos especificado por el usuario.

*Recomendación: Esta opción es útil para identificar una gama significativa de los números de los vecinos más cercanos antes de 2D LSA.*

### **Procedimiento**

1. Active la hoja que contiene sus datos XY. Es más conveniente establecer estos datos en las columnas 3 y 4 con las etiquetas de la muestra en la columna 1.
2. Escoja la opción *Spatial* desde el menú **GenAlEx**, y luego seleccione **NN Dist** desde el submenú.
3. En el cuadro de diálogo Nearest Neighbor Distance Options, seleccione la ubicación de sus datos XY.
4. Seleccione las opciones de resultados requeridas (ver detalles abajo).
5. Ingrese un título y prefijo, luego pulse *Ok*. El resultado es una hoja [NND].

### **Opciones de distancia de vecinos cercanos**

**Frequency Distribution of NN Distances (distribución de frecuencia de distancias de vecinos cercanos) [NN1]:** Esta opción produce la distribución de frecuencia de las distancias NN.

**Bin Size (tamaño del bloque) :** Ingrese el valor de bloque requerido, que sea relevante para su conjunto de datos.

## Relatedness (Relaciones genéticas o parentesco)

GenAlEx proporciona opciones en *Pairwise* para el cálculo de varios estimadores de relaciones genéticas, que son ampliamente utilizados en la literatura. Además, la opción *Pop Means* permite el cálculo del promedio de parejas emparentadas y realiza la evaluación estadística de permutaciones al azar.

*Recomendación: Las poblaciones en el contexto de relaciones genéticas también pueden ser grupos familiares o sexos.*

*Recomendación: esta opción no está restringida a matrices de parentesco, puede ser aplicada para el análisis de promedio para cualquier matriz de distancia GenAlEx.*

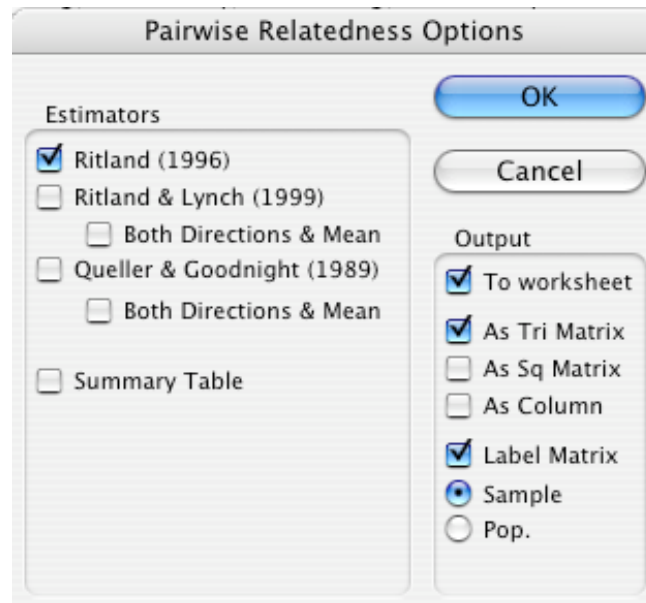
### Pairwise (En Parejas)

La opción *Pairwise* aplica solo a datos codominantes para una sola población. Si en el conjunto de datos se presentan múltiples poblaciones, el análisis ignora los parámetros de la población y trata los datos como una sola población. Sin embargo, los parámetros se mantienen para el análisis de *Pops Mean*.

#### Procedimiento

1. Active la hoja de trabajo que contiene su conjunto de datos codominantes en el formato de GenAlEx. Elija la opción *Relatedness* desde el menú **GenAlEx**, y seleccione la opción del submenú *Pairwise*.
2. Asegúrese de tener los parámetros de locus y muestras correctos en el cuadro de diálogo *Pairwise Relatedness Parameters*.
3. Ingrese el título y el prefijo de la hoja de trabajo, luego pulse *Ok*.

4. En el cuadro de diálogo siguiente, Pairwise Relatedness Options, seleccione las opciones requeridas (ver abajo). Vea abajo para los nombres de las hojas de resultado.



### ***Opciones de parentesco en pareja***

**Ritland (1996) [RI]:** Resultados con los valores para el estimador RI de Ritland (1996).

**Ritland & Lynch (1999) [LR1]:** Resultados con los valores para el estimador asimétrico 1 de Lynch & Ritland (1999).

**Both Directions & Mean:** Seleccione este cuadro para obtener el estimador asimétrico 2 [LR2] y el estimador promedio [LRM].

**Queller & Goodnight (1989) [QG1]:** Resultados con los valores para el estimador asimétrico 1 de Queller & Goodnight (1999).

**Both Directions & Mean:** Escoja este cuadro para obtener los dos estimadores, asimétrico 2 [QG2] y el promedio [QGM].

**Summary Table [RS]:** Resultados con un resumen de los valores de parejas con parentesco, de todos los estimadores seleccionados, en formato de tabla.

### **Output**

**To Worksheet:** Resultado con los estimadores de relaciones genéticas en una hoja de trabajo. Escoja una de las tres opciones de formato:

**As Tri Matrix:** Genera una matriz de distancia genética como una matriz triangular inferior. Esta opción es recomendada para conjuntos de datos pequeños con menos de 254 muestras.

**As Sq Matrix:** Genera una matriz de distancia genética como una matriz cuadrada simétrica.

**As Column:** Genera una matriz de distancia genética en columnas. Esta opción es proporcionada para conjuntos de datos con más de 254 muestras.

**Label Matrix:** Ubica las etiquetas de las muestras o poblaciones de su matriz.

## ***Pops Mean (Promedio poblacional)***

Esta opción requiere como datos de ingreso una matriz de las parejas relacionadas, con parámetros de poblaciones, con los cuales GenAlEx calculará un promedio dentro de cada población. Si usted quiere limitar el calculo de los promedios poblacionales por distancia, GenAlEx requiere una matriz de distancia geográfica, en el formato normal de GenAlEx. Esta opción es análoga a la opción *Pops as Dclass*, disponible en el menú *Spatial*.

*Recomendación: Esta opción no esta restringida a matrices relacionadas y puede ser utilizada para análisis de promedios de cualquier matriz de distancia de GenAlEx.*

### **Procedimiento**

1. Seleccione la hoja de trabajo que contiene la matriz de parejas relacionadas en formato de GenAlEx. Desde el menú **GenAlEx**, escoja la opción *Relatedness*, y seleccione la opción de submenú, *Pops Mean*.
2. Asegúrese de que los parámetros de locus y muestra estén correctos en el cuadro de diálogo Pops Mean Data Parameters.
3. Ingrese el título y el prefijo de la hoja de trabajo, luego pulse *Ok*

4. En el cuadro de diálogo siguiente, Pops Mean Options, seleccione las opciones requeridas (ver abajo).
5. Si usted desea restringir la comparación dentro de poblaciones, ingrese la distancia deseada en el cuadro de edición proporcionado. Para clases de distancia múltiple, ingrese el número de corridas y seleccione el método para incrementar las clases de distancia sobre estas corridas, doblando el tamaño en cada corrida, o incrementando por la distancia nominada.
6. Pulse *Ok*. El resultado es una hoja [PM1].

### ***Opciones de promedio poblacional***

**#Permutations (#permutaciones):** Ingrese el número de permutaciones requeridas en la prueba de significancia. Para propósitos de publicación el número de permutaciones debe ser de 999 o 9999.

*Nota: para un conjunto de datos grande, las permutaciones pueden tardar un poco. Observe el avance en la barra de progreso.*

**#Bootstraps:** Ingrese el número de bootstraps requeridos para estimar el 95% de intervalos de confianza alrededor de  $r$ . Para propósitos de publicación el número de bootstraps debe ser establecidos a 999 o 9999.

### **Distance Options (Opciones de distancia)**

**Distance (distancia):** Ingrese el valor de distancia requerido, y seleccione la matriz de distancia geográfica abajo. Dejar vacío para ignorar la distancia.

**#Runs (corridas):** Ingrese el número de veces que usted desea reiniciar el análisis con un nuevo valor de distancia.

**Double Size (tamaño doble):** Dobra la distancia en cada corrida consecutiva.

**Increment Size (tamaño de incremento):** Incrementa la distancia en un mismo valor en cada corrida consecutiva.

## Multilocus (Loci múltiples)

Este menú de opciones proporciona herramientas para marcaje genético (genetic tagging). También puede ser útil en otros contextos, como localizar muestras duplicadas en un conjunto de datos, o localizar clones del conjunto de datos de plantas o bacterias. El menú tiene cuatro subopciones: *Matches*, la cual detecta los genotipos repetidos dentro del conjunto de datos; *Genotype Prob.*, el cual calcula la probabilidad de apareamiento al azar de un genotipo específico en una población dada; *Prob. Identity*, el cual proporciona una estimación de la probabilidad promedio de que dos individuos no relacionados sacados de la misma población tengan el mismo genotipo multilocus; y *Prob. Exclusion*, el cual ofrece tres diferentes estimaciones de probabilidad de exclusión de paternidad, dependiendo de si los parentales son conocidos o no.

### Procedimiento

1. Seleccione la hoja de trabajo que contiene su conjunto de datos codominantes en el formato GenAEx. Escoja la opción *Multilocus* desde el menú **GenAEx**, y luego seleccione una de las cuatro opciones de submenú: *Matches*, *Genotype Prob.*, *Prob. Identity* o *Prob. Exclusion*. Todas las subopciones tienen el mismo cuadro de diálogo normal al iniciar.
2. Asegúrese de que los parámetros de locus y muestras estén correctos en el cuadro de diálogo *Multilocus Parameters*.
3. Ingrese título y prefijo de la hoja de trabajo y luego pulse *Ok*. Para *Matches* aparece un cuadro de diálogo a continuación. Para *Genotype Probability*, y *Probability Identity*, los resultados son las hojas [GP] y [PI], respectivamente. Para *Prob. Exclusion* la probabilidad de exclusión cuando ambos parentales son conocidos es un resultado en la hoja [P1X], la probabilidad de exclusión cuando un solo parental es conocido en la hoja [P2X], y la probabilidad para excluir dos parentales putativos en la hoja [P3X].

Multilocus Matches Parameters

#Loci (A1) 4 Pop. Size 48

#Samples (B1) 48

#Pops (C1) 1

Data Format

One Column/Locus Two Columns/Locus

Binary  Codominant

Haploid

This Match option applies only to codominant data!

Title Site MD with 4 loci

Worksheet Prefix MD

OK

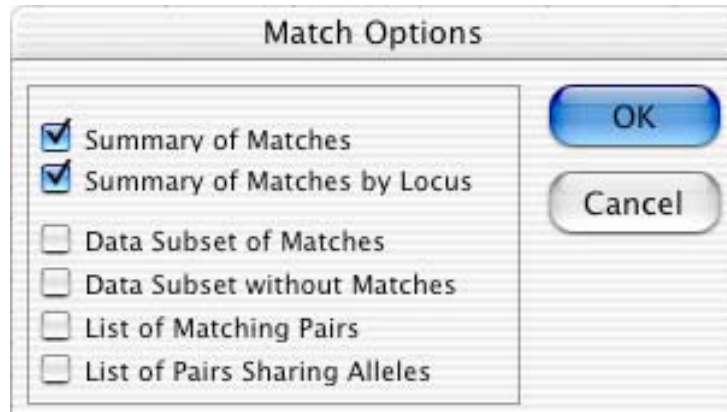
Cancel

Clear Pops.

Add Pops.

## **Matches (Apareamientos)**

En el cuadro de diálogo Match Options elija las opciones requeridas (ver abajo para mas detalles), y luego pulse *ok*. Vea las opciones para los nombres de las hojas de resultado.



### **Opciones de apareamiento**

**Summary of Matches (Resumen de apareamiento) [MS]:** Genera una lista de los genotipos multilocus presentes, con la cantidad y una etiqueta para cada genotipo. Las muestras estan clasificadas, con los genotipos repetidos presentados de primero.

**Summary of Matches by locus (Resumen de parejas por locus) [MLS]:** Genera un resumen de parejas por locus, para un número de loci creciente. El resultado también se presenta en forma gráfica.

*Recomendación: Esta opción es útil para determinar el número mínimo de loci requeridos para marcaje genético, cuando sea combinado con los resultados de probabilidad genotípica y la probabilidad de identidad estimada.*

**Data Subset of Matches (Subconjuntos de datos en parejas) [MDS]:** Genera un subconjunto que contiene aquellos individuos con genotipos pareado. El resultado incluye el número de genotipos repetidos y una etiqueta de cada genotipo.

**Data Subset Without Matches (Sunconjunto de datos sin parejas) [DS]:** Genera un subconjunto de datos que excluye las muestras que tienen genotipos repetidos. Solo se incluye el último individuo que lleva el genotipo repetido.

**List of Matching Pairs (Lista de parejas pareadas) [ML]:** Genera una lista de muestras emparejadas para todos los loci, todas menos un locus, todas menos dos loci, y todas menos tres loci.

*Recomendación: Esta opción es particularmente útil para encontrar errores en la asignación del genotipo.*

**List of Pairs Sharing Alleles (Lista de pares de alelos con un lugar compartido) [MAL]:** Genera una lista de pares de muestras que comparten por lo menos un alelo en cada locus .

## Clonal (Análisis de clones)

Este menú de opciones proporciona herramientas para calcular estimadores de probabilidad para inferir clonación usando datos codominantes, y para el análisis de estructura clonal. El menú tiene dos subopciones: *Find Clones*, el cual detecta automáticamente genotipos repetidos dentro del conjunto de datos, y calcula probabilidades; y *Clone size*, la cual estima el tamaño de clones putativos.

### Procedimiento para encontrar clones

1. Seleccione la hoja de trabajo que contiene su conjunto de datos codominante, en el formato GenAEx.
2. Escoja la opción *Clonal* desde el menú de **GenAEx**, y luego seleccione el sub-opción *Find Clones*.
3. Asegúrese que los parámetros de locus y muestra sean correctos en el cuadro de diálogo Find Clones Parameters.
4. Ingrese Título y prefijo y luego pulse *Ok*.

Find Clones Parameters

#Loci (A1) 4 Pop. Size 10

#Samples (B1) 10

#Pops (C1) 1

OK

Cancel

Clear Pops.

Add Pops.

Data Format

One Column/Locus Two Columns/Locus

Binary  Codominant

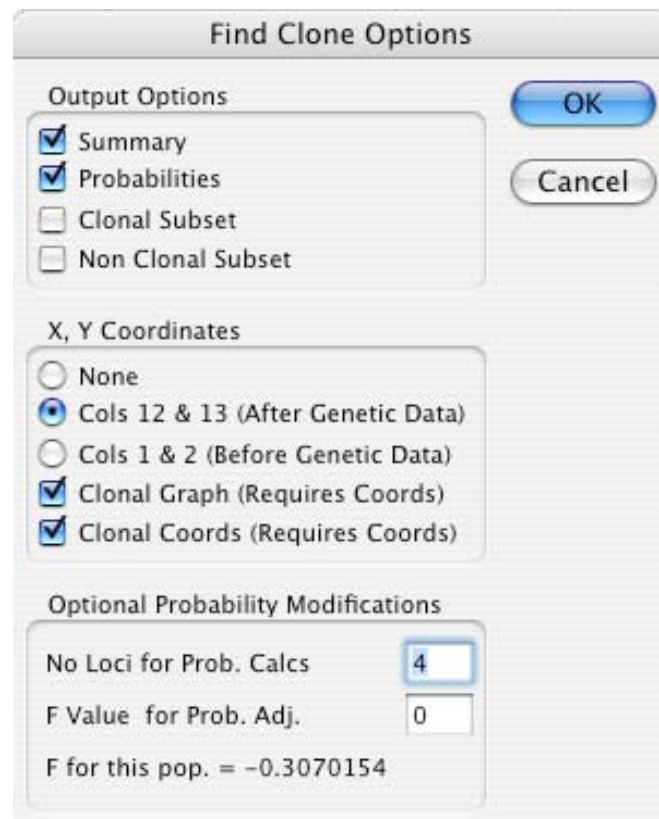
Haploid

This Find Clones option applies only to codominant data!

Title Find Clones

Worksheet Prefix

3. En el cuadro de diálogo siguiente, Find Clones Options, escoja las opciones requeridas (ver abajo para detalles) y luego pulse *OK*. Ver opciones para los nombres de los resultados.



### ***Opciones para encontrar clones***

**Summary (resumen) [CL]:** Genera una lista de clones putativos, basado en genotipos de múltiples loci repetidos, con la cantidad e una etiqueta para cada genotipo. Las muestras estan organizadas con los genotipos repetidos presentados primeros.

**Probabilities (probabilidades) [CLP]:** Genera un rango de estimadores de probabilidad de los genotipos clónales putativos.

**Clonal Subset (subconjunto clonal) [MDS]:** Genera un subconjunto de datos que contiene aquellos individuos con genotipos de loci múltiples emparejados (clones). El resultado incluye el número de genotipos repetidos y una etiqueta de cada genotipo.

**Non Clonal Subset (subconjunto no clonal) [DS]:** Genera un subconjunto de datos que excluye las muestras que tienen un genotipo repetido. El último individuo que porta el genotipo repetido es el único que aparece en el conjunto.

### **X, Y Coordinates (Coordenadas xy)**

Especifique la ubicación de sus datos XY y seleccione los resultados requeridos:

**Clonal Graph (gráfico clonal) [CL]:** Genera un gráfico de X,Y de todas las muestras, con los genotipos repetidos marcados.

**Clonal Coords (coordenadas clonales) [CLC]:** Resume las coordenadas XY de los clones putativos. Estos son los datos de ingreso requeridos para el análisis de tamaño de clones (Clone Size analysis).

### **Optional Probability Modifications (Modificaciones de probabilidades opcionales)**

**No. of loci for Prob. Calcs. (número de loci para calcular probabilidad):** Esta opción le permite escoger los loci que GenAIEx usara para calcular las probabilidades, desde el locus 1 a n.

*Recomendación: Si datos haploides son codificados como codominantes, pueden ser usados para encontrar clones. Esta opción podría facilitar entonces la exclusión de datos de los cálculos de probabilidad.*

**F value for Prob. Adj (Valor F para ajustar la probabilidad):** Esta opción permite hacer un ajuste a las probabilidades, tomando en cuenta el coeficiente de endogamia, F. Nb. Esto no sería apropiado si el F-valor está calculado de datos incluyendo clones.

### ***Procedimiento para tamaño de clones***

1. Seleccione la hoja de trabajo que contiene los datos formateados con la identificación de los genotipos en la columna 2, y las coordenadas XY en la columna 3 y 4. Este formato es el que se genera en la hoja [CLC] (ver arriba).
2. Escoja la opción *Clonal* desde el menú **GenAIEx**, y luego seleccione *Clone Size* en el submenú de opciones. Asegúrese que los parámetros de locus y muestra estén correctos en el cuadro de diálogo **Clone Size Parameters**.
3. Ingrese título y prefijo, luego pulse *Ok*. Los resultados incluyen una distribución de frecuencia para el tamaño máximo de clonación [CLMS], y para la distancia mínima entre clones [CLMD].

The image shows a software dialog box titled "Clone Size Parameters". It features several input fields and buttons. The fields include "#Loci (A1)" set to 1, "#Samples (B1)" set to 10, "#Pops (C1)" set to 1, and "Pop. Size" set to 10. On the right side, there are buttons for "OK", "Cancel", "Clear Pops.", and "Add Pops.". Below these is a "Data Format" section with a text box containing the message: "This Clone Size option requires genotype labels in the Col 2 and Coords in Cols 3 and 4!". At the bottom, there are two text input fields labeled "Title" and "Worksheet Prefix".

## Edición de Datos en Bruto

### Import Data (Importar datos)

GenAlEx ofrece varias opciones que facilitan la importación de datos genéticos a Excel. Archivos con texto delimitado por tabulación o espacios pueden ser importados directamente. Como alternativa, archivos múltiples de Genotipos o datos de Secuencias pueden ser importados simultáneamente cuando están contenidos dentro de una sola carpeta. Esto facilita la extracción de datos desde sistemas de genotipación / secuenciación.

GenAlEx también ofrece la opción para importar archivos formateados desde el programa de análisis genético poblacional, GenePop, y archivos de secuencias alineadas en formato Nexus (opción disponible en *Import Data* : Sequences).

Si estas opciones no son adecuados para su conjunto de datos, también se puede usar las herramientas de importación de Excel. Escoja la opción *Open* en el menú **File** de Excel.

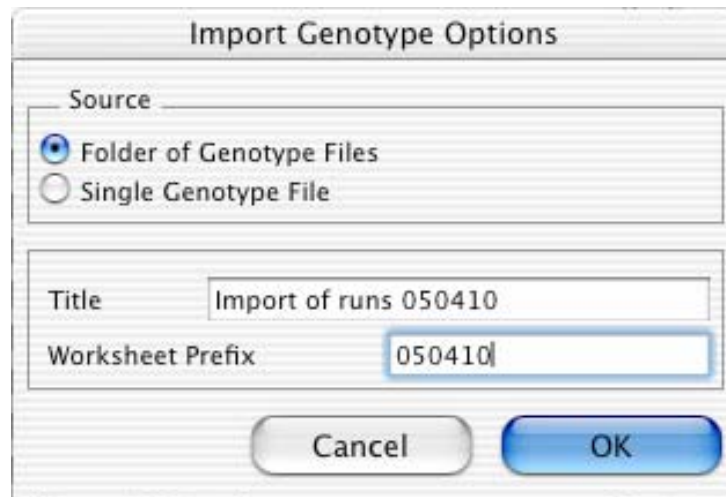
### Genotypes (Genotipos)

Esta opción importa archivos de datos genotípicos como archivos de texto delimitados por tabulaciones. Cada archivo es importado dentro del formato de GenAlEx en una hoja de trabajo separada en un solo libro de trabajo de Excel. También se proporciona una opción para importar un solo archivo.

*Recomendación: Para facilitar el rápido proceso de datos en GenAlEx, se recomienda que las opciones estén utilizadas dentro del software genotyping, e.g GeneMapper (AB), para preparar un archivo que consiste de: Columna 1 - etiqueta de la muestra; Columna 2 - información adicional tal como nombre del locus; Columna 3 - Alelo 1 para el locus 1; Columna 4 - Alelo 2 para el locus 1.*

#### Procedimiento

1. Escoja la opción *Import Data* desde el menú **GenAlEx**, y luego seleccione Genotypes desde el submenú.
2. En el cuadro de diálogo *Import Genotypes Options* seleccione la fuente de sus archivos de datos. Ingrese un título y un prefijo para sus hoja(s) y luego pulse *Ok*. El resultado es una hoja [IG].



### **Opciones para importar genotipos**

#### **Source (Fuente)**

**Folder of Genotype Files (Carpeta de archivos de genotipos):** Seleccione esta opción si usted tiene múltiples archivos individuales contenidos en un solo folder. GenAlEx necesita que identifique el folder que contiene sus archivos. Para esto usted necesita seleccionar un archivo dentro del folder identificado en el cuadro de diálogo siguiente, y luego pulsar 'Abrir'

**Single Genotype File (Un único archivo de genotipos)** Seleccione esta opción si usted tiene un solo archivo con datos genotípicos para importar. Los datos son importados en formato GenAlEx dentro de una sola hoja de trabajo con el nombre del archivo importado.

### **Sequences (Secuencias)**

Esta opción importa múltiples secuencias dentro de una sola hoja de trabajo en Excel, en formato GenAlEx. Opciones están proporcionadas para automatizar el proceso de los datos de la secuencia. Los datos de secuencia pueden ser en mayúscula o minúscula, con vacíos, y codificaciones de bases ambiguas. Como GenAlEx solo puede funcionar con datos numéricos, los caracteres alpha son convertidos a datos numéricos en la siguiente forma : A=1, C=2, G=3, T=4, :=5, -=5, otros=cero.

#### **Procedimiento**

1. Escoja la opción *Import Data* desde el menú **GenAlEx**, y luego seleccione *Sequences* desde el submenú.
2. En el cuadro de diálogo **Import Sequences Options** seleccione la fuente de sus archivos de datos. Ingrese los datos para procesar las secuencias, si lo requiere. Ingrese un título y un prefijo para sus hoja(s) output y luego pulse *Ok*. Vea las opciones abajo para los nombres de las hojas de resultado.

The image shows a dialog box titled "Import Sequence Options". It is divided into several sections:

- Source:** Contains five radio button options:
  - Folder of Multiple Sequence Files in Text Format
  - Text File of Multiple Sequences (Nexus Format)
  - Text File of Multiple Sequences (Fasta Format)
  - Text File of Multiple Sequences (Phylip Format)
  - Text File of Multiple Sequences (Mega Format)
- Sequence Processing:** Contains two text input fields:
  - Start Sequence: TTTCTCTT
  - End Sequence: GAAGAA
- Options:** Contains a checked checkbox for "Find Haplotypes".
- Title and Worksheet Prefix:** Contains two text input fields:
  - Title: Sequences 050410
  - Worksheet Prefix: 050410

At the bottom of the dialog are "Cancel" and "OK" buttons.

### ***Import Sequence Options (Opciones para importar secuencias)***

#### **Source (Fuente)**

**Folder of Multiple Sequence Files in Text Format (Archivos de secuencias Múltiples en Formato de texto):** Seleccione esta opción si tiene múltiples archivos con secuencias individuales que estén en el formato de “solo texto” y están contenidos dentro de un solo folder. Es necesario seleccionar un archivo dentro del folder desde el cuadro de diálogo siguiente, antes de pulsar ‘Abrir’.

**Text Files of Multiple Sequences (Archivos de texto de secuencias múltiples):** Seleccione el formato apropiado del archivo de texto con las múltiples alineaciones de secuencias que usted desee importar.

Usando ambas opciones se obtienen las siguientes hojas de trabajo:

**Imported Sequence [IS] (Secuencia Importada):** Importa cada dato de secuencias sin procesar dentro de una casilla en la hoja de trabajo.

**Sequence [SQ]:** Importa cada secuencia con un solo nucleótido por casilla. GenAlEx usa la longitud de la primera secuencia como una guía para procesar las secuencias siguientes. De esta forma si la secuencia siguiente es más larga que la primera, las bases extras no son procesadas.

Si el total de las secuencias consiste de más de 254 nucleótidos, esta opción no se completa porque el máximo de columnas en una hoja Excel es de 256.

**Sequence warnings [WS](Alertas):** Esta hoja hace salir todas las advertencias asociadas a la importación de la secuencia. Estas advertencias incluirán la detección de los códigos alfa con excepción de A, C, T, G: y, además facilita la comprobación de las llamadas bases ambiguas.

### **Sequence Processing (Procesando secuencias)**

**Start Sequence (Inicio de secuencia):** Las secuencias solo serán importadas desde donde inicia la secuencia nucleotídica especificada.

*Recomendación: Esto es útil para el ajuste de extremos indeseados en la secuencia, así se tienen secuencias alineadas en GenAEx.*

**End Sequence (secuencia final) :** La secuencia siguiente a la secuencia nucleotídica indicada, no será importada.

*Recomendación: Esto es útil para ajustar extremos indeseados desde la secuencia.*

### **Opciones**

**Find haplotypes (encontrar haplotipos):** Utilizando esta opción los datos pueden procesarse para encontrar haplotipos, generando varias respuestas en hojas de trabajo separadas:

**Polymorphic Sites (Sitios Polimórficos) [PS]:** Los resultados son un subconjunto de datos, que incluye solo los sitios polimórficos.

*Recomendación: Esta es una vía rápida para tener una tabla de sitios variables para un conjunto de secuencias.*

**Polymorphic Numerical (Polimorfismo numérico) [PN]:** Los resultados son un subconjunto de Sitios Polimórficos, convertidos a código numérico porque GenAEx solo puede usar datos numéricos. (A=1, C=2, G=3, T=4, :=5, -=5, otros=cero).

**Haplotype (Haplotipo) [HA]:** Proporciona un código haplotípico para cada secuencia individual, junto al haplotipo ambos caracteres alpha y numéricos.

**Haplotype Count (Conteo Haplotípico) [HC]:** Proporciona un conteo de cada haplotipo junto con su código numérico.

**Haplotype List (Listado de Haplotipos) [HL]:** Proporciona una lista de haplotipos con el código numérico de sus sitios polimórficos. También proporciona una muestra de individuo y población que contiene cada haplotipo.

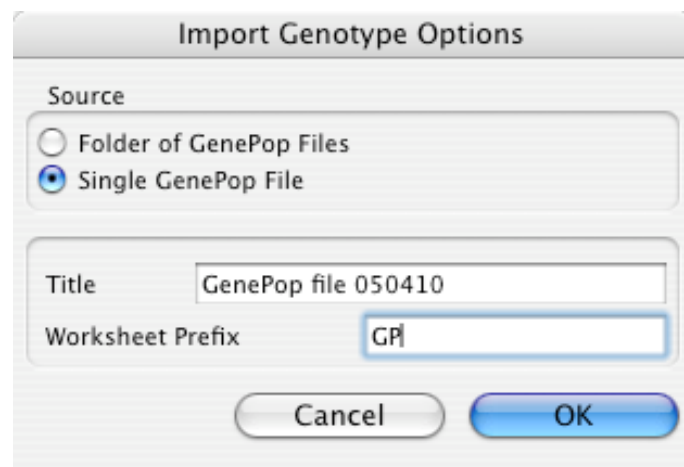
## **GenePop file (Archivo de GenePop)**

---

Esta opción permite importar un archivo de GenePop, solo cuando el formato del archivo es idéntico al formato de GenPop usado en el opción de exportar a GenPop usado por GenAlEx.

### **Procedimiento**

1. Escoja la opción *Import Data* desde el menú **GenAlEx**, y luego seleccione *GenePop* desde el submenú.
2. En el cuadro de diálogo **Import Genotypes Options** seleccione la fuente de su archivo de datos. Ingrese un título y un prefijo para sus hoja(s) de respuesta y pulse *Ok*. El resultado es una hoja [genepop].



### **Import GenePop Options (Opciones para importar desde GenePop)**

#### **Source (Fuente)**

**Folder of GenePop Files (folder de archivos de GenPop):** Seleccione esta opción si usted tiene múltiples archivos de GenePop contenidos en una sola carpeta. Es necesario identificar la carpeta que contiene sus archivos. En el cuadro de diálogo siguiente seleccione un archivo dentro de la carpeta y luego pulse 'Abrir'. Cada archivo se importa en una hoja de trabajo separada en un libro de trabajo nuevo, con el nombre del archivo importado. Los datos importados están en el formato de GenAlEx.

**Single GenePop File (Archivos únicos):** Seleccione esta opción si usted va a importar un solo archivo GenePop. Los datos son importados en el formato GenAlEx en una hoja de trabajo con el nombre del archivo impotado.

### ***Tab Delimited (delimitado por un tabulador)***

---

Importa un archivo que contiene datos delimitados por un tabulador, como una sola hoja de trabajo, en un libro de trabajo en Excel.

Para adelantar trabajo en GenAIEx, archivos importados deben ser guardados como libros de trabajo de Excel, y formateados apropiadamente en GenAIEx (hay opciones disponibles bajo la opción *Raw Data* en el menú **GenAIEx** para automatizar algunas de estas tareas).

### ***Space Delimited (Delimitado por Espacios)***

---

Importa un archivo delimitado por espacios, como una hoja de trabajo en un libro de trabajo en Excel. Para adelantar trabajo en GenAIEx, archivos importados deben ser guardados como libros de trabajo de Excel, y formateados apropiadamente en GenAIEx ( algunas opciones estan disponibles bajo la opción *Raw Data* en el menú **GenAIEx** para automatizar algunas de estas tareas).

## **Raw data (Datos crudos)**

---

GenAIEx ofrece varias opciones para ayudar a asimilar y ajustar conjuntos de datos de un único locus, importados desde sistemas de genotipado o secuencias. Estas opciones son particularmente útiles para la manipulación de conjuntos de datos grandes.

### ***Sort on Col 3 (Ordenar por columna 3)***

---

Ordena el conjunto de datos de GenAIEx (Iniciando en la Fila 4) con el contenido de la columna 3.

### ***Exclude Empty at Col 3 (Exclusión de vacíos)***

---

Esta opción excluye muestras que tienen una casilla vacía en la columna 3, y reformatea los parámetros de GenAIEx. Las muestras extraídas se mueven abajo al final del conjunto de datos, en la misma hoja de trabajo, intervenidas por una fila en blanco, que significa que GenAIEx no utiliza estas muestras en el análisis.

### ***Exclude ? at Col 3 (Exclusión de interrogantes)***

---

Esta opción excluye las muestras que contienen un '?' en la Columna 3 y reformatea los parámetros de GenAIEx. Las muestras extraídas son movidas abajo al final de los datos, en la misma hoja de trabajo, intervenidas por una fila en blanco, que significa que GenAIEx no utiliza estas muestras en el análisis.

### ***Combine Data (Combinar datos)***

---

Esta opción combina en una hoja de trabajo los datos de un mismo locus, que están en varias hojas de trabajo dentro de un mismo libro de trabajo (ejemplo, resultados de diferentes análisis genotípicos). Conjuntos de igual número de hojas de trabajo para diferentes loci pueden ser combinados simultáneamente ingresando la información apropiada en el cuadro de diálogo. Datos Geográficos, y otros datos pertenientes a la muestra, ejemplo sexo, etc, también pueden ser combinados.

#### ***Procedimiento***

1. Asegure que las hojas de trabajo que contiene sus datos para un solo locus está en las posiciones 1 a n, con los datos en formato GenAIEx. Si quiere combinar datos para más de un locus, las hojas de trabajo para cada locus necesitan estar agrupadas en conjunto en un libro de trabajo.
2. Escoga la opción *Raw Data* desde el menú **GenAIEx**, y luego seleccione *Combine Data* desde el submenú.
3. En el cuadro de diálogo *Combine Data Options* ingrese la información requerida (ver abajo para detalles). Ingrese un título y un prefijo para sus hoja(s) de resultado y luego pulse *Ok*. Los resultados son hojas [CBD].

### ***Combine data options (Opciones para combinar datos).***

#### **Data Specifications (Especificaciones de los datos)**

**No. Data Cols to Combine (Número de columnas para combinar):** Ingrese el número de columnas a combinar, iniciando por la columna 1.

*Recomendación: En un subconjunto de loci presentes en las hojas de trabajo, se puede escoger solo las columnas requeridas. Alternativamente, si usted desea incluir columnas adicionales que contengan por ejemplo datos XY, asegurese de que estas columnas estén contenidas en el número de columnas seleccionadas.*

**No. Worksheets per Set (número de hojas de trabajo para cada locus):** Ingrese el número de hojas de datos que quiere combinar para cada locus. Cuando desea combinar datos para loci múltiples, el número de hojas por locus debe ser igual, con todas las hojas de cada locus agrupadas en conjunto.

**No sets (Número de conjuntos):** Ingrese el número de conjuntos de hojas de trabajo que desee combinar ( 1 conjunto por locus).

#### **Options (Opciones)**

Deje estos cuadros vacíos si no los requiere.

**Col. No to Sort On (número de columna para ordenar ):** Ingrese un número de Columnas para simultáneamente ordenar todos los datos combinados por la columna especificada.

**Col. No to Exclude On if Missing (Número de la columna para exclusión de vacíos) :** Ingrese un número de columna, para simultáneamente excluir las muestras con datos ausentes en la columna especificada. Celdas vacías o celdas con interrogantes están considerados como datos ausentes. Celdas que contiene “?” son convertidas a celdas vacías.

## ***Check for Dups. (Revisión de Duplicados)***

Esta opción procesa datos y busca nombres de muestras duplicadas.

*Recomendación: Esta función es útil para remover controles duplicados.*

### **Procedimiento**

1. Seleccione la hoja de trabajo que contiene sus datos en formato GenAEx, escoja la opción *Raw Data* desde el menú **GenAEx**, y luego seleccione *Check for Dups* desde el submenú.
2. En el cuadro de diálogo Check For Duplicates Options ingrese la información requerida (ver detalles abajo). Ingrese un título y un prefijo para sus hoja(s) de resultado(s) y pulse *Ok*.

**Resultados:** la hoja [DS] contiene un subconjunto con las muestras duplicadas removidas, y un resumen del número de muestras emparejadas (con el mismo nombre). El parámetro de las muestras es ajustado así: en dos columnas más, separadas de los datos por una columna en blanco, se encuentra la información sobre el número de muestras duplicadas. Encontradas por cada nombre de muestra (No. matches), si las muestras con el mismo nombre tienen el genotipo idéntico. Si las muestras duplicadas tienen genotipos distintos, se mantiene la dúplica y aparece una nota “Dup does not match (Duplicado no apareado)” unida a la muestra.

### ***Check for Duplicates options (Opciones para revisar Duplicados)***

#### **Check options (Opciones de revisión)**

**Active worksheet only (solo la hoja de trabajo activada):** Busca duplicados solo en los datos ubicados en la hoja de trabajo activada.

**Multiple Worksheets (Positions 1 to n) (Hojas de trabajo múltiples):** Busca muestras duplicadas en múltiples hojas de trabajo. Las hojas de trabajo deben estar en posición 1 a n; en el cuadro de edición inserte el número de hojas para revisar.

### Advanced options (Opciones Avanzadas)

**No. of Extra Data cols to Extract (Número de columnas con datos extras):** Ingrese el número de columnas con datos extras que desea llevar a las hojas de trabajo. Estas columnas se separan de los datos por una columna en blanco.

**Data format (formato de datos):** Seleccione el formato apropiado para sus datos.

## **Merge Loci (Fusión de Loci)**

Esta opción fusiona datos de diferentes loci que están en hojas separadas dentro del mismo libro de trabajo. Opcionalmente, datos de poblaciones que estén en una hoja de trabajo específica, también pueden ser fusionados.

*Recomendación: Para facilitar el manejo de los datos deben insertarse etiquetas de locus en la fila 3 de las hojas de datos iniciales.*

### **Procedimiento**

1. Verifique que los datos de sus hojas de trabajo estén en formato GenAEx, y en las posiciones 1 a n en el libro de trabajo. Escoja la opción **Raw Data** desde el menú **GenAEx**, y luego seleccione **Merge Loci** desde el submenú.
2. En el cuadro de diálogo Merge Loci Options ingrese la información requerida (ver detalles abajo). Ingrese un título y un prefijo para sus hoja(s) de resultado(s) y pulse **Ok**.

**Output (resultados):** la hoja [MGL] con datos fusionados, con los parámetros apropiados de locus y muestra. Cuando los datos para un locus contiene muestras no presentes en conjuntos de datos de otros loci, estos datos faltantes aparecen como 0.

The screenshot shows the 'Merge Options' dialog box. It is divided into several sections:

- Merge Options:** Contains input fields for 'Optional Prefix' (CV), 'Smallest Numerical Identifier' (0010), 'Largest Numerical Identifier' (0900), and 'No Worksheets' (4). Below these is a note: 'Sample codes must be unique numerical identifiers. Data in GenAEx format. Worksheets in positions 1 to n.'
- Data Format:** Contains three radio buttons: 'Binary', 'Haploid', and 'Codominant' (which is selected).
- Advanced Options:** Contains five checkboxes: 'Tally Missing Loci on Merged Data Sheet' (checked), 'Output List of Missing by Set' (checked), 'Output List of Missing by Locus' (checked), 'Skip Output if No Pop Data for Sample' (checked), and 'Merge Pop Data Options' (unchecked).
- Output:** Contains three radio buttons: 'On Loci' (selected), 'On Pop', and 'Either'.
- Title and Worksheet Prefix:** At the bottom, there is a 'Title' field containing 'MERGED 170904' and a 'Worksheet Prefix' field containing '170904'.

'OK' and 'Cancel' buttons are located on the right side of the dialog.

**Merge loci options. (Opciones para fusión de loci)****Merge options (Opciones de fusión).**

**Optional Prefix (Prefijo Opcional):** Ingresa un prefijo de caracter alpha para las etiquetas de las muestra, si se desea. Este prefijo debe ser el mismo para todas las muestras.

**Smallest Numerical Identifier (Identificador numérico pequeño):** Ingresa el valor para la etiqueta numérica mas pequeña.

**Largest Numerical Identifier (Identificador numérico grande):** Ingresa el valor para la etiqueta numérica mas grande.

*Recomendación: Las etiquetas no tienen que ser continuas, pero los rangos definidos para estas dos opciones deben incluir el valor mínimo y máximo. Si desea fusionar solo un subconjunto de datos solo especifique el rango requerido.*

**No. Worksheets (Número de hojas):** Ingrese el número de loci que usted desea fusionar. Los datos para cada locus debe estar en hojas separadas en posición 1 a n.

**Advanced Options (Opciones Avanzadas)**

**Tally missing loci on merged datasheet (Contar datos faltantes en hojas de datos fusionados) [MGL]:** Esta opción cuenta los datos faltantes entre todas las muestras para cada locus, y el total de faltantes por muestra entre todos los loci.

**Output list of missing by set (Lista de faltantes por conjunto) [MISS]:** Esta opción genera una lista de muestras faltantes para cada locus en una sola hoja de resultados separada.

*Recomendación: Esta opción es útil para producir rápidamente una lista de muestras que posiblemente necesitan repetirse.*

**Output list of missing by locus (lista de datos faltantes por locus) [locus name\_MISS]:** Esta opción genera una lista de muestras faltantes para cada locus en hojas separadas.

**Skip Output if No Pop Data for Sample (Eliminar resultados cuando no hay datos poblacionales para la muestra):** Utilice esta opción para remover del conjunto fusionado las muestras que no tengan datos poblacionales.

**Merge Pop Data Options (Opciones de fusión de datos):** Esta opción permite que datos poblacionales se fusionen con datos genéticos. Esta opción genera un cuadro de diálogo subsecuente. Ver abajo para mas detalles.

**Data format (Formato de datos)**

Seleccione el formato apropiado para sus datos.

**Output (Resultados)**

**On loci:** seleccione esta opción para fusionar datos para los cuales la información del locus esta disponible. Si una muestra aparece en los datos de población pero no en los datos genotípicos, no aparecera en los resultados.

**On pop:** seleccione esta opción para fusionar datos para los cuales la información del población esta disponible. Si una muestra aparece en los datos de los loci pero no en los

datos poblacionales, no apareciera en los resultados. Alternativamente, Si una muestra aparece en los datos de población pero no en los datos genéticos, apareciera en los resultados .

**Recomendación:** esta opción es útil para identificar muestras que necesitan datos genotípicos

**Either (los dos):** seleccione esta opción para fusionar datos para los cuales la información del población o de loci esta disponible.

### **Merge Pop data with loci (Fusiona datos de población con datos genéticos)**

Esta opción permite que datos poblacionales incluidos en una hoja de trabajo separada sean fusionados con datos genéticos. Seleccione la hoja de trabajo que contiene los datos poblacionales, en el “Pop data’ menú. Los datos poblacionales deben estar formateados con la etiqueta de Muestra en la Columna 1, y de población y etiqueta de población en la Columna 2, y con el parámetro de muestra correcto en la Celda B1. Los resultados estan presentados en la hoja [MGLP].

### **Advanced Options (Opciones Avanzadas)**

**No Extra Data Cols to Extract (número de columnas adicionales para incluir):** Ingrese el número de columnas de datos extras, desde Columna 1 a Columna. n, que usted desee incluir en los resultados. Estas columnas deben estar separadas de los datos por una columna en blanco.

**Recomendación:** Esta opción permite extraer datos adicionales de una hoja de trabajo que contiene datos poblacionales (ejemplo coordenadas XY).

**Start Col for Outputting Extra Data (Columna de inicio de resultados de datos**

**adicionales):** Ingrese la columna, sobre la cual usted desea que sean insertados los datos extras en la hoja de resultados.

**Overwrite any existing pop data (Reemplazar datos ya existentes en la población):**

Reemplaza datos de una población ya existentes en las hojas de datos fusionados, con los nuevos datos de una población de la hoja seleccionada .

**Repeat Sample & Pop before Extra Data (repetir muestra y población antes del datos**

**extras):** Repite las etiquetas de la muestra y la población antes de cualquier dato adicional. Esto puede ser conveniente en el manejo de grandes conjuntos de datos genéticos.

## ***Unmerge Loci (Separar Loci)***

---

Esta opción produce hojas separadas en formato GenAEx para cada locus, desde un conjunto de datos multi-locus que esta en la hoja activada. Es necesario ubicarse en la hoja seleccionada. Las hojas de respuesta son llamadas de acuerdo a los nombres de los loci. Los datos originales permanecen intactos.

### ***Procedimiento***

1. Seleccione la hoja que contiene sus datos genéticos de loci múltiples, escoga la opción *Raw Data* desde el menú **GenAEx**, y luego seleccione *Unmerge Loci* desde el submenú.
2. En el cuadro de diálogo **Unmerge loci Data Parameters**, verifique que todos los parámetros estén correctos. Seleccione el formato de datos apropiado, Ingrese un título y prefijo para las hojas de trabajo y pulse “*Aceptar*”.

## ***Merge Pops (fusión de poblaciones)***

---

Esta opción permite fusionar datos de poblaciones con datos genéticos, ubicados en una sola hoja separada. Los datos de la población deben estar en el formato de *GenAEx, con el código de muestra en la Columna 1 y el código de población en la Columna 2 y tener correcto el parámetro de muestras.*

### ***Procedimiento***

1. Seleccione la hoja que contiene sus datos genéticos, escoga la opción *Raw Data* desde el menú **GenAEx**, y luego seleccione *Merge Pops* desde el submenú.
2. En el cuadro de diálogo **Merge Pop Data Options** ingrese la información requerida (ver detalles abajo).
3. Seleccione la hoja de trabajo que contiene los datos de la población desde el menú “*Pop Data*”. Los datos de la población deben estar en el formato con las etiquetas de la muestra en la columna 1 y las etiquetas de la población en la Columna 2.
3. Ingrese un título y un prefijo para sus hoja(s) de trabajo y pulse *Ok*. El resultado es una hoja [MGLP].

### ***Merge Pop options. (Opciones para fusión de loci)***

#### **Merge options (Opciones de fusión).**

**Optional Prefix (Prefijo Opcional):** Ingresa un prefijo de carácter alpha para las etiquetas de las muestra, si se desea. Este prefijo debe ser el mismo para todas las muestras.

**Smallest Numerical Identifier (Identificador numérico pequeño):** Ingresa el valor para la etiqueta numérica mas pequeña.

**Largest Numerical Identifier (Identificador numérico grande):** Ingresa el valor para la etiqueta numérica mas grande.

*Recomendación: Las etiquetas no tienen que ser continuas, pero los rangos definidos para estas dos opciones deben incluir el valor mínimo y máximo. Si desea fusionar solo un subconjunto de datos solo especifique el rango requerido.*

#### **Advanced Options (Opciones avanzadas)**

**No Extra Data Cols to Extract (número de columnas adicionales para incluir):** Ingrese el número de columnas de datos extras que usted desea incluir en los resultados. Estas columnas estan separadas de los datos por una columna en blanco.

*Recomendación: esta opción permite extraer datos adicionales desde una hoja de trabajo que contiene los datos poblacionales (ejemplo coordenadas XY).*

**Start Col for Outputting Extra Data (Columna de inicio de resultados de datos adicionales):** Ingrese la columna sobre la cual usted desea que sus datos sean insertados.

**Overwrite any existing pop data (Reemplaza datos ya existentes en la población):** Reemplaza datos de población ya existentes en la hoja de datos genéticos, con los nuevos datos de una población de la hoja seleccionada.

**Repeat Sample & Pop before Extra Data (repetir muestra y población antes de los datos extras):** Repite las etiquetas de la muestra y la población antes de cualquier dato adicional. Esto puede ser conveniente en el manejo de grandes conjuntos de datos genéticos.

## ***Process Seqs (Procesar secuencias)***

---

Esta opción procesa datos de secuencias contenidas en una sola hoja de trabajo de Excel, en formato GenAlEx, con el objetivo de detectar haplotipos.

Los datos de secuencias pueden estar en mayúsculas o minúsculas, con vacíos, y bases de codificación ambigua. Como GenAlEx solo puede funcionar con datos numéricos, los caracteres alpha son convertidos a datos numéricos de la siguiente forma: A=1, C=2, G=3, T=4 :=5, -=5, otros=cero.

### ***Procedimiento***

1. Escoja la opción *Raw Data* desde el menú **GenAlEx**, y luego seleccione *Process Seqs* desde el submenú.
2. En el cuadro de diálogo **Process Sequences Options** asegure que la fuente de sus datos este correcta. Ingrese los datos para procesamiento de secuencia, si es requerido, y pulse *Find Haplotypes*. Ingrese un título y un prefijo para sus hoja(s) de resultados y pulse "Aceptar". Ver abajo opciones para los nombres de las hojas de resultados.

The screenshot shows the 'Process Sequence Options' dialog box. It has a title bar 'Process Sequence Options'. The 'Source' section has two radio buttons: 'GenAlEx Raw Sequences (As Single Cell in Col 3)' (selected) and 'GenAlEx Sequences (As Single Nucs Each Col)'. Below are input fields for '#Samp' (7) and 'Seq. Length' (273). The 'Sequence Processing' section has 'Start Sequence' (TGTTAAA) and 'End Sequence' (AGTTTCG) input fields. The 'Options' section has a checked checkbox for 'Find Haplotypes'. The 'Title' field contains 'Sequences' and the 'Worksheet Prefix' field contains 'seqs'. At the bottom are 'Cancel' and 'OK' buttons.

## **Process Sequence Options (Opciones para procesar secuencias)**

### **Source (Fuente)**

GenAEx determina si sus datos de secuencias están incluída en una sola celda o con cada nucleótido en una celda separada, y selecciona automáticamente la fuente apropiado.

Las respuestas consisten de las siguientes hojas:

**Sequence (secuencia) [SQ]:** Si los datos de ingreso consisten de filas de secuencias en una sola celda, las secuencias son procesadas de forma que cada nucleótido está contenido en una sola celda. GenAEx usa la longitud de la primera secuencia como una guía para procesar las secuencias siguientes. De esta forma si la secuencia siguiente es más larga que la primera, las bases extras no son procesadas.

Si el total de las secuencias consiste de más de 254 nucleótidos, esta opción no se completa porque el máximo de columnas en una hoja Excel es de 256.

**Sequence warnings (advertencias) [WS]:** Esta hoja hace salir todas las advertencias asociadas a la importación de la secuencia. Estas advertencias incluirán la detección de los códigos alfa con excepción de A, C, T, G, : y - para facilitar la comprobación de las llamadas bases ambiguas.

### **Sequence Processing (Procesando secuencias)**

**Start Sequence: (Secuencia inicial):** Las secuencias solo serán importadas desde donde inicia la secuencia nucleotídica especificada.

*Recomendación: Esto es útil para el ajuste de extremos indeseados en la secuencia, así se tienen secuencias alineadas en GenAEx*

**End Sequence (Secuencia final) :** La secuencia siguiente a la secuencia nucleotídica indicada no será importada.

*Recomendación: Esto es útil para ajustar extremos indeseados desde la secuencia.*

### **Opciones**

**Find haplotypes (Encontrar haplotipos):** Utilice esta opción para encontrar haplotipos, obteniendo varios resultados en hojas de trabajo separadas:

**Polymorphic Sites (Sitios polimórficos) [PS]:** Proporciona un subconjunto de datos, que solo incluyen sitios polimórficos.

*Recomendación: Esta es una vía rápida para obtener una tabla de resultados con los sitios variables para datos de secuencia..*

**Polymorphic Numerical (Polimorfismo Numérico) [PN]:** Proporciona un subconjunto de Sitios Polimórficos, convertidos a código numérico porque GenAEx solo puede usar datos numéricos. (A=1, C=2, G=3, T=4, :=5, -=5, otros=cero).

**Haplotype [HA]:** Produce un listado de haplotipos para cada muestra en caracteres alfa y numéricos, y un código para cada haplotipo. Una 'h' es adicionada al final de los haplotipos para que no son tratados como un número por Excel.

**Haplotype Count [HC]:** Proporciona un conteo de cada haplotipo, junto con su código numérico.

**Haplotype List (Lista de haplotipos) [HL]:** Proporciona una lista de haplotipos con los códigos numéricos de sus sitios polimórficos. También proporciona un ejemplo de individuos y poblaciones que contienen cada haplotipo.

### ***Find Haplotypes (Encontrar haplotipos)***

Aplica para datos Haploides en formato numérico de GenAEx. Para datos de secuencias funciones similares son realizadas en *Process Sequences* bajo *Raw Data*, o por las opciones de *Processing* del menu de opciones *Import Sequences*.

#### **Procedimiento**

1. Active la hoja de trabajo que contiene sus datos en el formato GenAEx. Escoja la opción *Raw Data* desde el menú **GenAEx**, y luego seleccione *Find Haplotypes* desde el submenú.
2. Verifique que los parámetros de locus y muestra estén correctos en el cuadro de diálogo Find Haplotypes Options.
3. Ingrese un Título y Prefijo para su hoja de trabajo y luego pulse “*Aceptar*”. Ver opciones abajo para los nombres de las hojas de respuesta.

**Resultados** están en varias hojas:

**Polymorphic [PN]:** Resultados solo con las Posiciones Polimórficas.

**Haplotype [HA]:** Lista de haplotipos para cada muestra y proporciona un código para cada haplotipo, Una 'h' es adicionada al final de los haplotipos para que no son tratados como un número en Excel.

**Haplotype Count [HC]:** Proporciona un conteo de cada haplotipo junto con su código numérico.

**Haplotype List [HL]:** Proporciona una lista de haplotipos junto con su posición polimórfica. También proporciona un ejemplo de muestra y población que contiene cada haplotipo.

## **Edit Raw Data (Editar datos crudos)**

---

Las opciones del menú *Edit Data* proporcionan una serie de opciones que se explican así mismas. Son útiles para manipular y editar sus datos crudos y para prepararlos para exportar por la vía de las funciones Excel a otros programas. Cuando es apropiado GenAIEx requiere la selección de sus datos.

*Recomendación: Algunas de estas vías pueden ser útiles fuera del contexto de análisis genético.*

### *Empty > Zero (Vacíos a Cero)...*

Esta opción convierte celdas vacías en cero.

### *? > Zero... (interrogantes a cero)*

Esta opción convierte a cero todas las celdas que contienen un signo de interrogación ?.

### *Empty > Number (vacío a número)...*

Esta opción convierte celdas vacías a un número específico.

### *Empty > Text (vacío a texto)...*

Esta opción convierte celdas vacías en un texto específico.

### *Zero > Empty (Cero a vacío)...*

Esta opción deja vacías todas las celdas que contengan cero.

### *Text > Num (Texto a número)...*

Esta opción convierte valores que aparecen como números, pero que son tratados por Excel como texto al formato numérico verdadero para GenAIEx.

### *Num > Bin (Numérico a Binario)...*

Esta opción convierte datos numéricos a formato binario. Todas las celdas vacías son convertidas a cero, y todos los valores numéricos > 0 a 1. Valores cero no son alterados.

### *Split... (Separar)...*

Esta opción separa Genotipos codominantes cuando los dos alelos han sido importados en una sola columna.

## **Export Data (Exportar Datos)**

---

GenAIEx ofrece opciones para exportar datos formateados para una serie de otros programas listados en el menú. Para todas las opciones de exportación se proporciona un cuadro de diálogo normal *Export Parameters*, que presenta los formatos de datos disponibles para exportar y opciones específicas relevantes a los software de destino. También se proporcionan notas breves e instrucciones para familiarizarse con las opciones de formato y requerimientos del programa de destino.

Dependiendo de la opción de exportación escogida, GenAIEx proporciona resultados directamente como un archivo de texto, o como una hoja de trabajo de Excel, la cual luego necesita ser guardada manualmente como un archivo de texto delimitado por tabulador. En

ambos casos puede ser que el usuario requiere hacer mas modificaciones para los análisis en el software destino.

## Graph (Gráficas)

---

En adición a las opciones para visualización de dimensiones adicionales del análisis PCA, GenAlEx ofrece opciones para crear gráficos de coordenadas XY con datos en el fomato normal de GenAlEx. En este caso, se requiere una hoja de datos de ingreso en el formato normal de GenAlEx, con las etiquetas de muestras en la columna 1, y XY en las Columnas. 3 & 4. Una opción similar esta disponible para coordenadas de Latitud/longitud, con la longitud ubicada en el eje X.

## Stats (Estadísticas)

---

GenAlEx ofrece una serie de herramientas para estadísticas y transformacion las cuales pueden ser utiles tanto para el manejo de conjuntos de datos de GenAlEx, como para usos mas amplios del Excel. Los submenús y cuadros de diálogo son explicativos asi mismos.

## Options (Opciones)

---

El submenú *Generic* proporciona opciones para modificar el menú GenAlEx y los cuadros de diálogo según Ud requiere.

*Recomendación: Los maestros pueden usar esta opcion para eliminar la ventana de introducción de GenAlEx (Splash screen), además de algunas de las opciones avanzadas desde el menú.*

*Recomendación: Esta opción se puede también utilizar para modificar los cuadros de diálogos communes. Por ejemplo, si usted trabaja con un conjunto de datos binarios, puede reajustar sus cuadros de diálogos a binario. Después de cambiar las opciones del cuadro de diálogo durante una análisis, retorne al menú de subopciones **Generic** y pulse salvar en el cuadro de diálogo.*