

## **CAPÍTULO 2: MATERIALES Y MÉTODOS**

### **2.1 Participantes**

Este trabajo de investigación fue realizado por la señorita Natalia del Rocío Salinas Aponte en el Laboratorio de Biotecnología Vegetal de la Carrera de Ingeniería en Biotecnología, Departamento de Ciencias de la Vida de la Escuela Politécnica del Ejército.

Los colaboradores científicos del proyecto fueron la Dra. Karina Proaño y la Dra. Patricia Jiménez, las mismas que desempeñaron las funciones de directora y co-directora respectivamente. La Dra. Claudia Segovia y el Ing. Vinicio Armijos colaboraron en el análisis estadístico de los datos. Además, la Ing. Valeria Ochoa y la Ing. Paola González aportaron a la investigación con asesoría técnica y científica.

### **2.2 Zona de Estudio**

#### **2.2.1 Fase de Campo**

La recolección del material vegetal se la realizó en el banco de germoplasma de *Jatropha curcas* L. del IASA II de Santo Domingo. Las tres poblaciones de *Jatropha curcas* L. que fueron analizadas pertenecen a la provincia de Manabí.

#### **2.2.2 Fase de Laboratorio**

La parte experimental se realizó en el Laboratorio de Biotecnología Vegetal de la Carrera de Ingeniería en Biotecnología de la Escuela Politécnica del Ejército,

ubicados en el campus politécnico de Sangolquí, cantón Rumiñahui, provincia de Pichincha.

### 2.3 Período de tiempo de investigación

La investigación se inició en el mes de Junio del año 2008 y finalizó en el mes de Julio del año 2009.

### 2.4 Diseño

El presente proyecto de investigación constituye la fase inicial de un estudio experimental de biología, en donde se dispone de poca información sobre la variabilidad genética de *Jatropha curcas* L. en la provincia de Manabí y en el Ecuador. Debido a esto, es preciso realizar un análisis exploratorio de variables (Batanero *et al.*, 1991). Para ello, se utilizó varias herramientas bioestadísticas de análisis molecular que permitirán describir las características globales del conjunto de datos obtenidos. En la siguiente tabla se muestran las variables analizadas:

**Tabla 2.1** Variables analizadas en el estudio de variabilidad genética de *Jatropha curcas* L.

<b>Variables Independientes</b>	<b>Variables dependientes</b>
Primer	Grupo de clasificación
Banda amplificada	
Localización de población de <i>Jatropha curcas</i> L.	

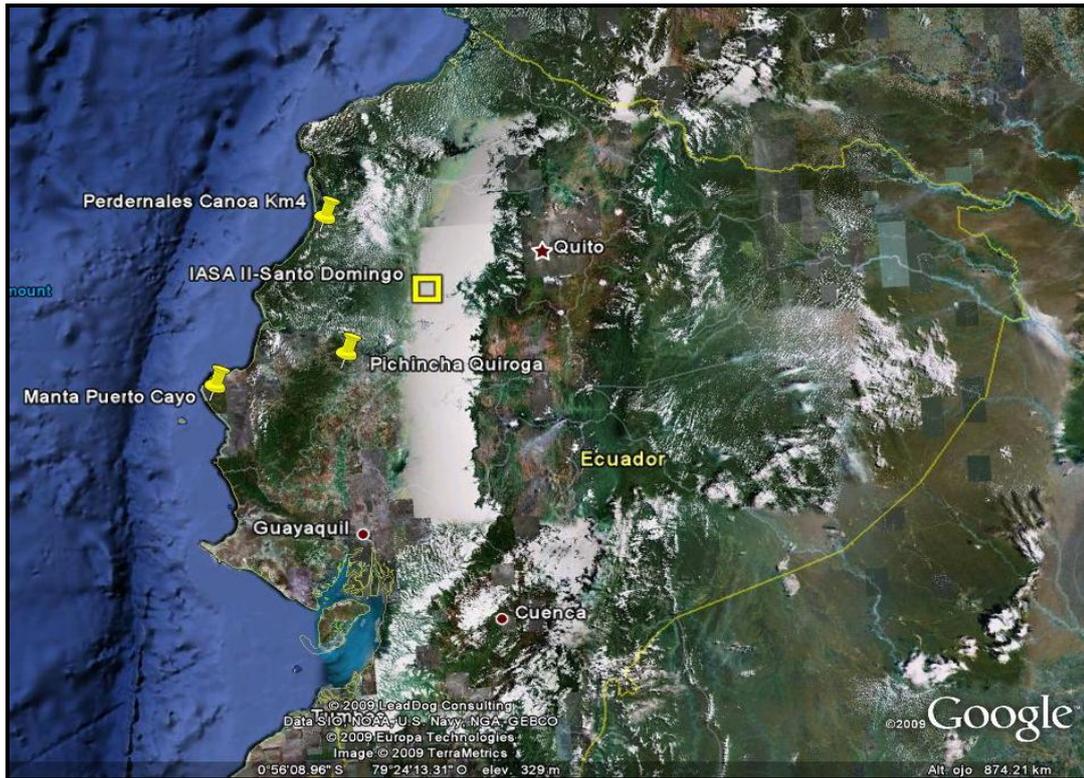
## 2.5 Recolección de material vegetal

Las muestras que se utilizaron en el presente estudio fueron recolectadas en el banco de germoplasma de *Jatropha curcas* L. en el IASA II de Santo Domingo en la Hacienda La Molestina (Figura 2.1) la cual se encuentra a una altura de 272 msnm y cuyos puntos geográficos son 0° 26' 281" S (latitud) y 79° 19' 236" O (longitud).



**Figura 2.1** Banco de germoplasma de *Jatropha curcas* L. en el IASA II de Santo Domingo en la Hacienda La Molestina.

El material vegetal proviene de semillas recolectadas en distintos sitios de la Provincia de Manabí separados físicamente de 100 a 160 km de distancia, los mismos que se encuentran en un rango de altitud de 16 a 295 msnm (Figura 2.2).



**Figura 2.2** Mapa del Ecuador donde se indican las ubicaciones de las tres poblaciones de *Jatropha curcas* L. analizadas y del banco de germoplasma de donde se obtuvieron las muestras. Foto tomada de Google Earth 2009 ([www.earth.google.es](http://www.earth.google.es)).

La información geográfica de las poblaciones analizadas de *Jatropha curcas* L. consta en la Tabla 2.2. La codificación de las poblaciones y de los individuos se realizó de acuerdo a la posición en el área del banco de germoplasma.

**Tabla 2.2** Descripción de lugares y coordenadas geográficas de las muestras recolectadas para el análisis molecular.

Grupo	Lugar	Número de individuos	Coordenadas Geográficas	Altitud (msnm)
4	Pedernales Canoa Km 4	35	0° 2' 18" N 80° 4' 7" O	16
12	Pichincha Quiroga	29	0° 58' 52" S 79° 57' 4" O	140
15	Manta Puerto Cayo	32	1° 9' 0.5" S 80° 51' 43" O	295

De cada población se recolectaron muestras de todos los individuos, para lo cual se tomaron tres hojas jóvenes en buen estado y se colocaron junto con sílica gel en sobres de papel debidamente identificados. En total se tomaron muestras de 130 plantas de dos meses de edad, las cuales fueron almacenadas en el laboratorio a -20 °C para asegurar su viabilidad en los siguientes procesos.

## **2.6 Extracción y Purificación de ADN**

La extracción de ADN se realizó de todos los individuos recolectados del banco de germoplasma y fue llevada a cabo siguiendo el protocolo modificado de Khanuja, Shasany, Darokar & Kumar (1999):

El material vegetal (1 g) se maceró con nitrógeno líquido hasta formar polvo, el cual se transfirió a tubos de 15 ml donde se añadió 3 ml de tampón de extracción (CTAB 2.5%, EDTA 25 mM, NaCl 1.5M, 100 mM Tris-Cl pH 8.0,  $\beta$ -mercaptoetanol 0.2% y PVP 1%) en cada tubo. A continuación se incubaron los tubos a 60 °C por 2 horas, agitándolos muy cuidadosamente cada 30 minutos. Luego de dejarlos enfriar se colocaron 3 ml de cloroformo-alcohol isoamílico (24:1), se agitaron y se centrifugaron a 5000 rpm por 20 minutos. Al sobrenadante se añadió 1.5 ml de NaCl 5 M y 2.7 ml de isopropanol a -20 °C y se permitió reposar la mezcla por 30 minutos a -20 °C. Nuevamente, se centrifugaron los tubos a 5000 rpm por 20 minutos. Al pellet obtenido se añadió 300  $\mu$ l de etanol al 80% a -20 °C. Se descartó el etanol y se dejó secar el pellet. Posteriormente se añadió 300  $\mu$ l de solución TE alto en sales (Tris HCl 10 mM, EDTA 1 mM, NaCl 1 M) y se almacenó toda la noche a 4 °C.

La solución obtenida se traspasó a tubos de 1.5 ml y se colocó 3  $\mu$ l de RNasa (10mg/ml). Se incubó a 37 °C por 1 hora, luego se añadió 3  $\mu$ l de Proteínasa K (20mg/ml) y se incubó a 37 °C por 30 minutos. A continuación se colocó 300  $\mu$ l de

cloroformo-alcohol isoamílico (24:1) y se centrifugó a 12 000 rpm por 10 minutos. La capa acuosa obtenida se transfirió cuidadosamente a un tubo nuevo y se realizó otra purificación con 300 ul de cloroformo-alcohol isoamílico (24:1). Luego se añadió 600 ul de etanol a 95% a -20 °C, manteniendo esta solución en reposo por 30 minutos a -20 °C. Tras una nueva centrifugación se dejó secar el pellet al ambiente, para luego añadir 300 ul de etanol al 80% a -20 °C. Se eliminó el etanol y se dejó secar el pellet nuevamente. Finalmente se colocó 300ul de solución TE (Tris HCl 10 mM, EDTA 1 mM) en cada tubo, almacenando por último a -20 °C las muestras de ADN obtenidas.

## **2.7 Cuantificación de ADN**

Las soluciones de ADN obtenidas en el proceso de extracción fueron cuantificadas tanto en geles de agarosa por electroforesis como por fluorimetría. La electroforesis en geles de agarosa se realizó para todas las soluciones de ADN obtenidas. Para ello, se prepararon geles a una concentración de agarosa de 0.8 % con la solución TBE 1X (Tris base 90 mM, Ácido bórico 90 mM, EDTA 2 mM) agitándola durante 10 minutos para que se hidrate (Kumar *et al.*, 2008). Luego, esta mezcla se calentó y agitó nuevamente. Este proceso se repitió por lo menos 3 veces para permitir que el polímero se disuelva adecuadamente, consiguiendo una solución traslúcida. Cuando la temperatura alcanzó 50 °C se colocó el bromuro de etidio (5 ul/100 ml). Esta mezcla se transfirió a la cubeta previamente preparada para que se solidifique.

Para la cuantificación de ADN se colocó una mezcla de 8 ul de solución de ADN junto con 2 ul de solución de carga Blue Juice 5X INVITROGEN® en cada uno de los posillos. Además, se colocó 4 ul del marcador del peso molecular High DNA Mass Ladder INVITROGEN® para comparar la intensidad de las bandas y estimar la concentración de ADN de las soluciones analizadas según la información descrita en la Tabla 2.3. Luego de haber cargado todos los posillos se realizó la corrida electroforética a 120 V y 300 mA por 1 hora y 30 minutos. La visualización de los geles se los realizó en el fotodocumentador Dolphin-View Image System. A partir de estos geles se analiza

la calidad y se estima la cantidad de ADN y de contaminación que se tiene en cada una de las soluciones.

**Tabla 2.3** Cantidad de ADN de acuerdo a la intensidad de banda del marcador High DNA Mass Ladder INVITROGEN®.

Tamaño del fragmento	High DNA Mass Ladder	
	4 ul	Concentración de ADN (ng/ul)
10 000	200ng	50
6 000	120ng	30
4 000	80ng	20
3 000	60ng	15
2 000	40ng	10
1 000	20ng	5

Para la cuantificación por fluorometría se siguió el protocolo del kit (Quant-iT™ dsDNA BR Assay Kit, Qubit INVITROGEN®). Para ello se etiquetó tubos PCR limpios de 0.5 ml para los estándares (2) y para las muestras (n). Se preparó la solución de trabajo Quant-iT, diluyéndola 200 veces. A continuación se colocó 190 ul de la solución de trabajo Quant-iT en cada uno de los tubos estándares. Se adicionó en los tubos respectivos 10 ul de cada solución estándar. El volumen de la reacción final en todos los tubos fue de 200 ul. Luego se colocó 180-199 ul de la solución de trabajo Quant-iT en cada tubo destinado para las muestras. Se adicionó un volumen de solución de ADN en cada tubo, de tal manera que se completen los 200 ul. Se permitió la incubación de todos los tubos a temperatura ambiente por 2 minutos.

En el fluorómetro Qubit se escoge el respectivo tipo de ensayo, en este caso se utilizó el Quant-iT™ dsDNA BR, porque el ADN extraído es genómico y de doble cadena. Luego se calibró el equipo con las dos soluciones estándares. Se insertó uno por uno los tubos de las muestras para su respectiva lectura. Los valores obtenidos pueden estar en ng/ml o ug/ ml. Se realizaron tres lecturas de las soluciones, se promedió estos valores y se calculó la concentración de ADN en cada una de ellas, mediante la siguiente fórmula:

$$\text{Concentración de la muestra} = \text{valor}_{QF} \times \left(\frac{200}{x}\right)$$

En donde:

Valor QF: el valor dado por el fluorómetro.

X = el número de  $\mu\text{l}$  de solución que se colocó en el tubo de ensayo.

La ecuación genera un resultado con las mismas unidades del valor dado por el fluorómetro.

## 2.8 Amplificación de fragmentos de ADN mediante primers ISSRs

Para la amplificación de fragmentos de ADN se emplearon los primers ISSRs descritos en la Tabla 2.4 (Ochoa *et al.*, 2008). El ensamblaje de las soluciones master mix se realizó en una cámara de flujo laminar tipo II CSB 120 y se utilizaron reactivos INVITROGEN®. Cabe destacar que en todos los ensayos de PCR se trabajó con dupletas y con un control negativo que contiene todos los reactivos de PCR excepto la solución de ADN, con el fin de comprobar la ausencia de contaminación en los reactivos (Basha & Sujatha, 2009). Las reacciones de PCR se llevaron a cabo en el termociclador TECHNE-TC-512.

**Tabla 2.4** Lista de primers utilizados en el análisis molecular y sus respectivas secuencias (Ochoa *et al.*, 2008).

Nº	Primers ISSRs	Secuencia 5` a 3`
1	17899A	cacacacacacaag
2	17899B	cacacacacacagg
3	HB9	gtgtgtgtgtgtgg
4	HB11	gtgtgtgtgtgtcc
5	HB14	ctctctctctgc
6	844A	ctctctctctctctac
7	844B	ctctctctctctctgc
8	17898A	cacacacacacaac
9	17898B	cacacacacacagt
10	HB12	caccaccacgc

En los primeros ensayos se amplificaron productos con los 10 primers descritos anteriormente (Tabla 2.4) tomando en cuenta las condiciones y concentraciones de PCR (Tabla 2.5) descritas por Ochoa *et al.* (2008).

**Tabla 2.5** Concentraciones iniciales, finales y volumen final de cada reactivo necesario para la PCR en un volumen total de 25 ul. Los datos del cloruro de magnesio dependen del primer (Ochoa *et al.*, 2008).

Reactivo	Concentración Inicial	Concentración Final	Volumen Final (ul)
Buffer	10X	1X	2,50
MgCl <sub>2</sub>	50mM	2mM ó 1,8mM	1 ó 0,9
dNTP`s	10mM	0,8mM	2
Primer	100pmol/ul	0,25pmol	0,25
Taq-polimerasa	5U/ul	1U	0,2
H <sub>2</sub> O	-	-	17,05 ó 17,15
ADN	20ng/ul	1,6ng/ul	2
TOTAL			25

Los tiempos y las temperaturas de PCR para cada primer que se utilizaron en esta investigación se pueden observar en la Tabla 2.6.

**Tabla 2.6** Condiciones de tiempo y temperatura de cada fase en la PCR para los 10 primers utilizados (Ochoa *et al.*, 2008).

Fases	Primers ISSRs										Tiempo (min)
	17899A	17899B	HB9	HB11	HB14	844A	844B	17898A	17898B	HB12	
	Temperatura °C										
Denaturación inicial	94	94	94	94	94	94	94	94	94	94	2
40 ciclos	Denaturación	94	94	94	94	94	94	94	94	94	0.5
	Alineamiento	43	43	44	45	40	54	54	47	42	0.75
	Extención	72	72	72	72	72	72	72	72	72	2
Extención Final	72	72	72	72	72	72	72	72	72	72	7

Se realizaron además ensayos para determinar la concentración de ADN adecuada, la concentración de cloruro de magnesio, la temperatura de alineamiento y el número de ciclos. Luego de establecer las condiciones de PCR, se amplificaron los productos con todos los primers utilizando las soluciones de ADN de las tres poblaciones de *Jatropha curcas* L.

## **2.9 Visualización de productos amplificados**

Los productos de PCR fueron visualizados en geles de agarosa al 1.5 % (Dong, Chen, Gituru & Wang, 2007), en donde se mezclaron 8 ul de la solución master mix con 2 ul de la solución de carga Blue Juice 5X INVITROGEN® en cada uno de los posillos. Se incluyó también 5 ul del marcador molecular TrackIt 100 bp DNA Ladder INVITROGEN®. Los geles se corrieron a 120 V y 300 mA por 2 horas 30 minutos aproximadamente. Al terminar la corrida electroforética los geles fueron visualizados en el fotodocumentador Dolphin-View Image System.

## **2.10 Registro de datos**

Las fotografías de los geles fueron analizados en el programa Dolphin-View Band Tool en donde se tomó como referencia los pesos moleculares de cada banda del marcador TrackIt 100 bp DNA Ladder INVITROGEN® para determinar el peso de las bandas amplificadas. A partir de los resultados obtenidos se registró la presencia con el valor “1” o la ausencia con el valor “0” de las bandas amplificadas en una matriz binaria (Basha & Sujatha, 2007). Las bandas no muy bien definidas y las bandas que no constaban en ambas dupletas no fueron consideradas en el registro de datos (Kumar *et al.*, 2008). Aquellas bandas con la misma distancia de migración fueron tomadas como homólogas (Venkatachalam *et al.*, 2008). La matriz obtenida es compatible en los distintos ordenadores de análisis molecular.

## 2.11 Análisis de datos

La naturaleza dominante de los marcadores moleculares ISSRs exige considerar ciertos aspectos en el análisis de datos. Uno de ellos es que no se puede asumir el equilibrio de Hardy-Weinberg debido a que es difícil calcular las frecuencias alélicas a partir de los resultados. Además, la ausencia de bandas puede deberse a distintas circunstancias (genotipo homocigoto recesivo, mutaciones, error experimental) siendo éste un resultado ambiguo. Existen programas estadísticos específicos para el análisis de datos de marcadores dominantes que excluyen de sus cálculos el equilibrio de Hardy-Weinberg y toman en cuenta solo la presencia de bandas (Culley, 1999).

### 2.11.1 Características y atributos de los primers ISSRs.

A partir de la matriz binaria se encontró el porcentaje de bandas polimórficas y monomórficas para cada primer ISSR y totales. Las bandas polimórficas fueron categorizadas como: únicas (específicas de una accesión), raras (presentes en menos del 10% de los individuos), compartidas (presentes en el 70% de los individuos) y similares (presentes en más del 70% de los individuos) (Tatikonda *et al.*, 2009).

Para conocer el poder informativo o discriminatorio que tienen los 6 primers ISSRs se evaluó el poder de resolución (RP). El poder de resolución de cada primer fue calculado con la fórmula  $RP = \sum I_b$ . El valor de  $I_b$  representa la cantidad de información que muestra cada banda y se obtiene con la fórmula  $I_b = 1 - [2 \times |0.5 - p|]$ , donde  $p$  es la proporción de todos los individuos que contienen dicha banda (Tatikonda *et al.*, 2009). De este análisis se determinó el primer más y menos informativo, con los cuales se realizó un análisis de componentes principales (ACP) individual, con el fin de conocer la capacidad de cada primer para agrupar a los individuos. Se utilizó el programa MVSP (Multi-Variate Statistical Package) versión 3.1 para obtener el ACP.

### **2.11.2 Cálculo de distancias genéticas**

Las distancias genéticas pueden ser obtenidas utilizando los coeficientes de similaridad de Jaccard y Dice, los mismos que consideran para el análisis solamente la presencia de bandas excluyendo así el equilibrio de Hardy-Weinberg. La diferencia entre el coeficiente de Jaccard y de Dice, es que el coeficiente de Dice da más peso a la presencia de bandas en ambos individuos (Bonin *et al.*, 2007).

La matriz binaria obtenida fue editada con la opción “NTedit” del programa NTSYSpc (Numerical Taxonomy and Multivariate Analysis System) versión 2.0 para lograr un formato compatible con este programa (Rohlf, 1998). Posteriormente se calculó la matriz de similaridad con el coeficiente de Dice mediante la opción “Similarity - SimQual”. A partir de esta matriz cuadrática de similaridad se elabora una matriz de disimilaridad o variabilidad que resulta de restar 1 menos los valores de similaridad (variabilidad = 1-similaridad). Los valores de similaridad van desde 0 que significa que no existe similaridad entre los grupos, hasta 1 que indica que existe la máxima similaridad entre los grupos (Legendre, P & Legendre, L, 2000).

### **2.11.3 Análisis de agrupamiento**

Para conocer la estructura de agrupamiento de los datos se realizaron dos técnicas multivariantes, el método no ponderado de apareamiento de grupos a través de la media aritmética o UPGMA (unweighted pair-group method with arithmetic averages) y el análisis del vecino más cercano o NJoin (neighbor joining). El método UPGMA ejecuta una reducción progresiva de la matriz de similaridad por lo que los resultados deben considerarse para relacionar grupos más no individuos (Legendre, P & Legendre, L, 2000). El análisis Njoin puede analizarse para relacionar individuos.

Las técnicas antes mencionadas se efectuaron a partir de la matriz de similaridad obtenida con el coeficiente de Dice. Se utilizaron los programas NTSYSpc versión 2.0 y MVSP versión 3.1 para construir los dendogramas del método UPGMA y del análisis NJoin, mostrando resultados muy similares. Se analizó sin embargo, el dendograma UPGMA obtenido con el programa MVSP y el dendograma NJoin obtenido con el programa NTSYSpc.

Además se realizó un análisis de componentes principales (ACP) para conocer la distribución que tienen los individuos en las poblaciones. El ACP resume en pocas dimensiones, la variabilidad de una matriz de dispersión que tiene un gran número de variables (Legendre, P & Legendre, L, 2000). Este análisis se lo llevó a cabo en NTSYSpc versión 2.0, en FAMD (Fingerprint Analysis with Missing Data) versión 1.23 beta, en MSVP versión 3.1 y en Genalex (Genetic Analysis in Excel) versión 6.2. La distribución resultante es muy parecida en todos los casos, sin embargo se analizó el ACP obtenido con el programa MSVP versión 3.1.

#### **2.11.4 Análisis de la diferenciación genética**

El grado de diferenciación dentro de las poblaciones y entre poblaciones puede ser medido usando distancias genéticas mediante el análisis jerárquico de varianza molecular o AMOVA. El AMOVA fue originalmente diseñado para analizar datos de marcadores codominantes, pero si se tratan a los resultados de marcadores dominantes como haplotipos, el AMOVA es igualmente útil (Culley, 1999). El AMOVA se asemeja a un ANOVA anidado. Los datos de un ANOVA anidado están arreglados jerárquicamente y los cuadrados conocidos se calculan para grupos en todos los niveles jerárquicos, permitiendo probar hipótesis de diferencias entre grupos y dentro de grupos en muchos niveles jerárquicos (Dorak, 2009).

El AMOVA fue realizado en el programa FAMD versión 1.23 beta, utilizando las distancias obtenidas con el coeficiente de Dice. También se utilizaron los programas Genalex versión 6.2 y Arlequín 3.1 para calcular las distancias euclidianas y ejecutar el AMOVA. Las distancias euclidianas entre pares de vectores constituyen una escala que es igual a la distancia menor entre dos puntos (Ochoa *et al.*, 2008).

Los valores obtenidos en Genalex versión 6.2 y Arlequín 3.1 son iguales, mientras que éstos difieren relativamente poco con los resultados del programa FAMD versión 1.23 beta. Esta diferencia se debe al tipo de distancia genética utilizada (coeficiente de Dice, distancia euclidiana), debido a esto se decidió trabajar con los resultados del programa Genalex versión 6.2.

La diferenciación de poblaciones fue analizada también mediante el cálculo de los coeficientes  $F_{IS}$ ,  $F_{ST}$  y  $G_{ST}$ . El índice  $F_{IS}$  o coeficiente de endogamia fue calculado para los individuos de cada una de las tres poblaciones en el programa Geste. El índice de fijación  $F_{ST}$  fue calculado en los programas Genalex 6.2, Arlequín 3.1 y en Geste, en los cuales se obtuvieron valores muy parecidos. El valor de  $G_{ST}$ , análogo al valor  $F_{ST}$  fue calculado en el programa PopGen32, a partir del cual se calculó también el flujo génico  $Nm$ .

#### **2.11.5 Test de Mantel**

Se llevó a cabo el Test de Mantel para determinar si existe o no una relación entre las distancias genéticas y entre las distancias geográficas de las poblaciones. El test de Mantel es una regresión en donde las variables son matrices de distancia que resumen la disimilaridad entre grupos (Reynolds & Houle, 2002). Este análisis fue llevado a cabo en el programa Genalex versión 6.2 y en NTSYSpC versión 2.0

utilizando 9999 permutaciones. Los resultados de ambos programas son sumamente similares, sin embargo se analizará aquellos obtenidos con Genalex versión 6.2.