

**UNIVERSIDAD NACIONAL AGRARIA
LA MOLINA**

FACULTAD DE AGRONOMÍA



**“DIVERSIDAD GENÉTICA DEL LOCHE (*Cucurbita moschata*
Duchesne ex Lam.) CULTIVADO EN LAMBAYEQUE, PERÚ
MEDIANTE MARCADORES MICROSATÉLITES”**

Presentada por:

CARLOS IRVIN ARBIZU BERROCAL

**TESIS PARA OPTAR EL TÍTULO DE
INGENIERO AGRÓNOMO**

Lima - Perú

2019

**UNIVERSIDAD NACIONAL AGRARIA LA
MOLINA**

FACULTAD DE AGRONOMÍA

**DIVERSIDAD GENÉTICA DEL LOCHE (*Cucurbita moschata*
Duchesne ex Lam.) CULTIVADO EN LAMBAYEQUE, PERÚ
MEDIANTE MARCADORES MICROSATÉLITES**

Presentada por:

CARLOS IRVIN ARBIZU BERROCAL

**TESIS PARA OPTAR EL TÍTULO DE:
INGENIERO AGRÓNOMO**

Sustentada y aprobada ante el siguiente jurado:

.....
Ing. M. S. Andrés Casas Díaz
PRESIDENTE

.....
Dr. Raúl Blas Sevillano
ASESOR

.....
Ing. Mg. Sc. Julián Chura Chuquiya
MIEMBRO

.....
Ing. Mg. Sc. Roberto Ugás Carro
MIEMBRO

Lima – Perú

2019

AGRADECIMIENTOS

Deseo brindar mi agradecimiento a mi asesor de tesis, Dr. Raúl Blas, por brindarme la oportunidad de unirme a su grupo de investigación. Además, por su guía en el presente estudio de investigación.

Agradezco también al comité de mi tesis por su tiempo y dedicación en la lectura del presente estudio, y por sus sugerencias de mejora. Deseo enfatizar mi agradecimiento al Ing. Roberto Ugás por brindarme la oportunidad de iniciar este estudio de investigación.

Gracias a Luis Gutiérrez por su apoyo técnico y sus enseñanzas en las técnicas moleculares. Estas me sirvieron de mucho durante mi estadía en Wisconsin, EE.UU. Asimismo, gracias a la amistad brindada por Joel Flores. Espero seguir trabajando junto a él realizando más investigación.

Finalmente, agradezco a los miembros de mi familia, especialmente a mis padres por su constante sacrificio y aliento en brindarme oportunidades para mejorar siempre. Gracias también a mis hermanos por ser mis amigos siempre. Mis agradecimientos a Jessica y Vania pues con ellas encontré una nueva etapa en mi vida, la cual me encanta, y me incentiva a seguir adelante siempre.

INDICE

| | |
|--|----------|
| AGRADECIMIENTOS | i |
| INDICE | ii |
| RESUMEN | vii |
| ABSTRACT | vii |
| | |
| I. INTRODUCCIÓN | 1 |
| | |
| II. REVISIÓN DE LITERATURA | 4 |
| 2.1. CLASIFICACIÓN Y NOMBRES COMUNES DE <i>Cucurbita. moschata</i> | 4 |
| 2.2. CARACTERÍSTICAS MORFOLÓGICAS DEL LOCHE | 4 |
| 2.3. ORIGEN Y DISTRIBUCIÓN | 7 |
| 2.4. DIVERSIDAD | 11 |
| 2.5. USOS | 12 |
| 2.6. AGRONOMÍA | 13 |
| 2.7. POTENCIAL DEL CULTIVO | 15 |
| 2.8. CARACTERIZACIÓN MOLECULAR | 16 |
| 2.8.1. COLECTA DE MATERIAL VEGETAL | 16 |
| 2.9. EXTRACCIÓN DE ADN | 17 |
| 2.9.1. Solubilización de membranas | 17 |
| 2.9.2. Inhibición enzimática | 17 |
| 2.9.3. Complejos de ADN | 18 |
| 2.10. CUANTIFICACIÓN DE LA CALIDAD DEL ADN | 18 |
| 2.11. ELECTROFORESIS | 19 |
| 2.12. REACCIÓN EN CADENA POLIMERASA | 19 |
| 2.13. CARACTERIZACIÓN DEL GENOMA | 22 |
| 2.13.1. Diversidad genética | 22 |
| 2.14. MARCADORES MORFOLÓGICOS | 24 |
| 2.15. MARCADORES MOLECULARES | 25 |
| 2.16. CLASIFICACIÓN DE LOS MICROSATÉLITES | 31 |
| 2.17. APLICACIONES DE LOS MICROSATÉLITES | 34 |
| 2.18. ANÁLISIS ESTADÍSTICO | 37 |

| | | |
|--------------|---|-----------|
| III. | MATERIALES Y MÉTODOS | 47 |
| 3.1. | MATERIAL VEGETAL | 47 |
| 3.2. | MATERIAL PARA LA COLECCIÓN DE LOCHE | 47 |
| 3.3. | MATERIAL PARA LA EXTRACCIÓN DE ADN | 48 |
| 3.4. | MATERIAL PARA LA ELECTROFORESIS | 49 |
| 3.5. | MATERIAL PARA PCR | 51 |
| 3.6. | COLECCIÓN DE ACCESIONES | 52 |
| 3.7. | EXTRACCIÓN DE ADN | 53 |
| 3.8. | DETERMINACIÓN DE CALIDAD Y CONCENTRACIÓN DE ADN | 53 |
| 3.9. | MARCADORES MICROSATÉLITES (SSRs) | 53 |
| 3.10. | ANÁLISIS ESTADÍSTICO | 62 |
| IV. | RESULTADOS Y DISCUSIONES | 66 |
| 4.1. | ANÁLISIS MOLECULAR | 66 |
| 4.2. | ANÁLISIS DE POLIMORFISMO | 67 |
| 4.3. | ANÁLISIS DE SIMILARIDAD | 68 |
| 4.4. | ANÁLISIS DE AGRUPAMIENTO MOLECULAR | 69 |
| 4.5. | ÍNDICE DE DIVERSIDAD DE NEI O HETEROCIGOSIDAD ESPERADA (H_e) Y OBSERVADA (H_o) | 70 |
| 4.6. | ANÁLISIS DE COORDENADAS PRINCIPALES | 71 |
| 4.6.1. | Análisis entre zapallos (<i>Cucurbita</i> sp.) | 71 |
| 4.6.2. | Análisis entre loches (<i>Cucurbita moschata</i>) | 72 |
| 4.7. | ANÁLISIS DE VARIANZA MOLECULAR (AMOVA) | 73 |
| 4.7.1. | AMOVA entre zapallos (<i>Cucurbita</i> sp.) | 74 |
| 4.7.2. | AMOVA entre loche (<i>Cucurbita moschata</i>) | 74 |
| V. | CONCLUSIONES | 78 |
| VI. | RECOMENDACIONES | 79 |
| VII. | REFERENCIAS BIBLIOGRÁFICAS | 80 |
| VIII. | ANEXOS | 91 |

INDICE DE FIGURAS

| | |
|----------|---|
| Figura 1 | Loche cultivado en Lambayeque, Perú. A. Flores masculinas de loche. B. Planta de loche junto a inflorescencias. C. Planta de loche sin inflorescencia. D. Flores femeninas de loche6 |
| Figura 2 | Frutos de loche. A. Pulpa y semilla. B. Frutos provenientes de Lambayeque. C. Frutos provenientes de Bagua. D. Pulpa y semilla7 |
| Figura 3 | Diagrama que muestra el polimorfismo de los microsatélites debido a las diferencias en su longitud30 |
| Figura 4 | Dendograma de <i>Cucurbita moschata</i> (A) cultivadas en Lambayeque Amazonas y Pasco, <i>C. maxima</i> (B) y <i>C. pepo</i> (C) cultivados en Ica y Lima respectivamente empleando el algoritmo UPGMA con el coeficiente de similitud Dice70 |
| Figura 5 | Análisis factorial de coordenadas principales en especies de zapallo basado en una matriz de distancia de Dice generada con 69.7% de la variancia total72 |
| Figura 6 | Análisis factorial de coordenadas principales en loches basados en una matriz de distancia de Dice generada con 38.4% de la variancia total73 |

INDICE DE TABLAS

| | | |
|----------|--|----|
| Tabla 1 | Restos arqueológicos de <i>C. moschata</i> | 8 |
| Tabla 2 | Potencial medicinal de <i>Cucurbita moschata</i> | 13 |
| Tabla 3 | Composición bromatológica de <i>C. moschata</i> (100g) | 16 |
| Tabla 4 | Clasificación de los microsátélites | 31 |
| Tabla 5 | Secuencia de los 21 iniciadores microsátélites que se usaron para el estudio de la variabilidad genética del loche | 54 |
| Tabla 6 | Protocolo de amplificación de marcadores SSR | 57 |
| Tabla 7 | Programa de amplificación de marcadores SSR | 57 |
| Tabla 8 | Número y porcentaje de bandas polimórficas y monomórficas detectadas en los patrones SSR | 66 |
| Tabla 9 | Índice de contenido polimórfico de 18 iniciadores SSR utilizados en el presente estudio | 68 |
| Tabla 10 | Índice de diversidad genética de Nei o heterocigosidad esperada (He) y heterocigosidad observada (Ho) de los loches cultivados en Lambayeque, Amazonas y Pasco | 71 |
| Tabla 11 | Análisis de varianza molecular (AMOVA) para 108 genotipos de zapallo basados en 85 bandas de SSR | 74 |
| Tabla 12 | Análisis de varianza molecular (AMOVA) para 98 genotipos de loche basados en 40 bandas de SSR | 75 |

INDICE DE ANEXOS

| | | |
|---------|--|----|
| Anexo 1 | Lista de accesiones de <i>Cucurbita</i> sp. empleados en el presente estudio | 91 |
| Anexo 2 | Protocolo de extracción de ADN genómico vegetal Doyle & Doyle. 1990 | 96 |

RESUMEN

La presente investigación contribuye al conocimiento de las relaciones genéticas entre los cultivares de loche (*Cucurbita moschata*) cultivados en Lambayeque, Perú. Noventa y ocho accesiones de loche cultivadas en Lambayeque y Amazonas, Perú, que representan la distribución geográfica del cultivo fueron tomadas de forma selectiva y aleatoria para estudios de diversidad genética mediante el uso de 21 marcadores microsátélites. Adicionalmente, se incluyeron 10 accesiones de *C. pepo* y *C. maxima* en este estudio para determinar su relación con el loche. Se detectaron un total de 85 bandas. La diversidad genética de Nei (0,10) indica que el loche cultivado en Lambayeque y Amazonas tiene una baja diversidad genética. El análisis de los resultados de la varianza molecular demuestra que la diferencia genética entre el loche cultivado en Lambayeque y el Amazonas es pequeño. La propagación vegetativa mediante estacas y el cultivo del loche en una zona geográfica muy restringida explicaría la escasa diversidad del cultivo. Esto a su vez sugiere que la aparente variación observada en la forma del fruto puede explicarse por la mutación somática y / o factores ambientales.

Palabras clave: *Cucurbita*, loche, marcadores moleculares, diversidad genética.

ABSTRACT

The current research contributes to the knowledge of genetic relationships among loche cultivars (*Cucurbita moschata*) cultivated in Lambayeque, Peru. Ninety eight accessions of loche grown in Lambayeque, and Amazonas, Peru representing the geographical distribution of the crop were selectively and randomly taken for genetic diversity studies by means of 21 Simple Sequence Repeats (SSRs) markers. Additionally 10 accessions of *C. pepo* and *C. maxima* were included in this study to determine their relationship with loche. A total of 85 bands were detected. Nei's genetic diversity (0.10) indicates that loche grown in Lambayeque and Amazonas has a low genetic diversity. Analysis of molecular variance results demonstrate genetic difference between loche cultivated in Lambayeque and Amazonas is little. Vegetative propagation by means of stem cuttings and cultivation in a very restricted geographical area would explain the rather low diversity of the crop. This in turn would suggest that the apparent variation observed in fruit shape may be explained by somatic mutation and/or environmental factors.

Keywords: *Cucurbita*, loche, molecular markers, genetic diversity.

I. INTRODUCCIÓN

El género *Cucurbita* es bien conocido por sus especies cultivadas, las cuales en los países de habla hispana son comúnmente conocidas con nombres como “calabazas”, “zapallos” o mediante numerosos nombres en lenguas indígenas, mientras que en los países anglosajones se les denomina “squashes”, “pumpkins” o “gourds”. Los frutos inmaduros y maduros y las semillas de las especies cultivadas de este género han tenido y tienen gran importancia alimenticia en una gran parte del mundo, además de que en muchas regiones de Latinoamérica, las flores y algunas partes vegetativas (puntas tiernas de los tallos o guías) también son apreciadas como verdura (Lira et al., 1995).

Los zapallos (*Cucurbita spp.*) son uno de los cultivos domesticado en América y dispersados por todo el mundo (Whitaker et al, 1950; Cruces 1987; Harlan, 1992; Merrick, 1995). Es cultivado para la producción de vegetales frescos y frutos maduros. La especie *Cucurbita moschata* es una de los cultivos más importantes en las áreas tropicales. Los frutos maduros y tiernos, flores machos, semillas y guías tiernas son consumidas por los pobladores (National Research Council, 1989).

El zapallo loche (*Cucurbita moschata Duchesne ex Lam.*) es una hortaliza que se cultiva exclusiva y tradicionalmente en la costa norte del Perú, principalmente en el departamento de Lambayeque y prácticamente no es conocido en otros lugares (Andres et al, 2006). Sin embargo, se requiere fomentar su difusión, ya que constituye uno de los cultivos de los moches y puede convertirse en un producto conquistador de mercados internacionales, teniendo en cuenta el reconocimiento de la gastronomía lambayecana en el mundo. Actualmente, los cultivos de agro-exportación amenazan con desplazar a este excepcional zapallo de su zona de producción. (Andres et al, 2006).

En la actualidad, no se tienen estudios de la diversidad genética del zapallo loche en el Perú. Este estudio constituye un factor fundamental para entender la forma como se estructuran las poblaciones del loche. Además, la estimación de la diversidad genética es de vital importancia para el mejoramiento genético, conservación de las especies y ganancia en selección. Al presente no sabemos cuál es la diversidad genética de los zapallos loche

cultivados en el norte del Perú. Tampoco sabemos su diversidad fenotípica, es decir, no sabemos cuántos cultivares son sembrados por los agricultores del norte peruano.

Las técnicas moleculares se han convertido en herramientas muy útiles para los estudios biológicos de plantas relacionados a la biodiversidad y evolución en general, y diversidad en particular. Además, el uso de marcadores moleculares acelera los programas de mejoramiento genético, en el sentido que, los marcadores pueden suplantar los tradicionales marcadores morfológicos o fenotípicos, en especial cuando la expresión de esta característica es ambientalmente inestable o difícil de observar (Varshney et al., 2005).

El análisis de la variación en la secuencia del ADN es de gran importancia en los estudios de genética. En este contexto, los marcadores moleculares son una herramienta útil para analizar la variación genética, y han mejorado bastante en el análisis genético de diversos cultivos. Existe una variedad de marcadores moleculares, incluyendo RFLPs (Restriction Fragment Length Polymorphisms, RAPDs (Random Amplification of Polymorphic DNAs), AFLPs (Amplified Fragment Length Polymorphisms) y Microsatélites o SSRs (Simple sequence repeats). Estos marcadores han sido desarrollados en diferentes cultivos agrícolas. Dentro de los diferentes tipos de marcadores moleculares, los SSRs son útiles para una gran variedad de aplicaciones en genética de plantas y cruzamientos por su alta reproducibilidad, naturaleza multialélica, herencia codominante, abundante y buena cobertura del genoma.

OBJETIVOS

OBJETIVO GENERAL

- Establecer las diferencias moleculares del zapallo *loche* cultivado en la región de Lambayeque, Perú.

OBJETIVOS ESPECÍFICOS

- Obtener cultivares de zapallo *loche* en la costa norte del Perú.
- Determinar la variabilidad genética de los cultivares de zapallo *loche* cultivados en el Perú mediante el uso de marcadores microsatélites (SSRs).

II. REVISIÓN DE LITERATURA

2.1. CLASIFICACIÓN Y NOMBRES COMUNES DE *Cucurbita moschata*

Posición taxonómica de Cucurbita moschata Duchesne ex Lam.:

El zapallo loche es citado en páginas web y en otras literaturas como *Cucurbita maxima* Duchesne (Soukup, 1987; Custer, 2000 citado por Andres et al., 2006) o *Cucurbita moschata* Duchesne, pero hay un entendimiento común en el Perú que todos los loches son *C. moschata* (Diaz, 1957; Ugás et al, 2000 citado por Andres et al., 2006)

Según Lira y Montes (1994) el zapallo loche tiene la siguiente clasificación botánica:

| | |
|----------|---|
| Reino | Plantae |
| División | Magnoliophyta |
| Clase | Magnoliopsida |
| Orden | Cucurbitales |
| Familia | Cucurbitaceae |
| Género | <i>Cucurbita</i> |
| | <i>Cucurbita moschata</i> (Duchesne ex Lam.) |
| | Duchesne ex Poir. |
| | <i>Pepo moschata</i> (Duchesne ex Lam.) Britt |

Nombres comunes:

Soukup (1970) indica que la especie *Cucurbita moschata* es conocida con los siguientes nombres comunes: lacayote y lacahuite. Por otro lado, Lira y Montes (1994) mencionan que a esta especie se le conoce también con los nombres de anco, anquito, auyama, calabaza, calabaza moscada, tamalayote o zapallito coreano. Lira (1995) menciona que este especie recibe nombres de acuerdo al lugar de producción, tales como: náhuatl: tamalayota [México, Colombia (Guerrero)]; castellano: calabaza (México), ayote (Guatemala a Costa Rica), auyama (Panamá a Venezuela), zapallo (Ecuador, Perú), joko (Bolivia).

2.2. CARACTERÍSTICAS MORFOLÓGICAS DEL LOCHE

La especie *Cucurbita moschata* (Duchesne ex Lam) es una planta herbácea, comúnmente rastrera a trepadora o algunas formas con hábito subarborescente, anuales. Tiene partes vegetativas y flores suave y densamente pubescentes a vellosas, con tricomas cortos y largos uniseriados, septados y no septados, algunas veces estrigosos; ápices vegetativos más o

menos reflexos (Lira, 1995). *C. moschata* es una hierba anual caulescente, reptante o trepadora (Lira y Montes, 1994).

Sus tallos son rígidos, ligeramente angulosos apareciendo sulcados al secar (Lira, 1995). Lira y Montes (1994) indican que el tallo es anguloso y densamente pubescente, con zarcillos apicales con los que se fija a la vegetación y al suelo.

Las hojas se encuentran sobre peciolo de 30 o más cm. de largo; láminas anchamente ovado-cordadas a suborbiculares, 20-25 cm o más de largo, 25-30 cm o más de ancho, ligeramente 3-5 lobadas, lóbulos ovados o triangulares, ápice obtuso, cortamente apiculado o mucronado, márgenes serrado-denticulados, superficie adaxial frecuentemente con manchas blancas en la intersección de las venas (Lira, 1995). Las hojas son anchas, cordadas a ovadas, ligeramente tri o pentalobuladas, de hasta 25 por 30 cm. de superficie, de márgenes serrados, ubicadas al cabo de un peciolo elongado de hasta 30 cm (Lira y Montes, 1994).

Las flores no son aromáticas. Las flores estaminadas se encuentran sobre pedicelos robustos, 16-18 cm de largo; receptáculo anchamente campanulado, 3-10 mm de largo, 8-20 mm de ancho, nada a muy ligeramente constreñido por debajo de los sépalos; sépalos linear-lanceolados, frecuentemente ligeramente a muy expandidos o foliáceos hacia el ápice, 10-40 (-50) mm de largo, 1-3 mm de ancho o un poco más en las proporciones apicales foliáceas; corola amarilla a anaranjada en el ápice, 5-13.5 cm de largo, 5-dividida hasta poco menos de la mitad de su longitud total, lóbulos triangulares, agudos a acuminados, con los márgenes enteros aunque levemente ondulado y doblados hacia adentro, el tubo generalmente angostándose hacia la base, algunas veces muy levemente ensanchado; filamentos 9.5-16 mm de largo, gradualmente angostándose de la base hacia el ápice, glabros a esparcidamente pulberulentos en la base; columna de las anteras angosta, de longitud muy variable 12-35 mm de largo, 4-6 mm de ancho. Las flores pistiladas se encuentran sobre pedicelos engrosados, 2-4 (-8) cm de largo, angulosos, pubescentes; ovario de muy diversas formas (globoso, ovoide, oblado, cilíndrico, piriforme, cónico, lageniforme, etc. Pero nunca turbaniformes), pubescente, glabrescente con la edad; perianto como en las estaminadas, pero la corola más grande, el receptáculo muy reducido y los sépalos más frecuentemente foliáceos y generalmente más largos, algunas veces de hasta 7.5 cm de largo; columna de estilos 15-18 mm de largo; estigmas bilobados (Lira, 1995).

Lira y Montes (1994) mencionan que la planta es invariablemente monoica; las flores son solitarias, axilares y pentámeras, de pétalos carnosos y succulentos. Las masculinas alcanzan los 18 cm de largo, de forma campanulada, con el cáliz corto y expandiéndose hacia el ápice; las femeninas tienen un pedicelo ancho y robusto, y el ovario globoso a cónico. Los estigmas son tres, lobulados.



Figura 1: Loche cultivado en Lambayeque, Perú. A. Flores masculinas de loche. B. Planta de loche junto a inflorescencias. C. Planta de loche sin inflorescencia. D. Flores femeninas de loche.

Los frutos son de tamaño muy variable y de formas diversas, aunque prácticamente siempre conservando la forma del ovario, lisos o más comúnmente con costillas redondeadas, algunas veces verrucosos, cáscara tanto engrosada, rígida y durable como suave y percedera, con patrones de coloración muy variables, verde claro a verde oscuro liso o con manchas crema, pardo claro a oscuro, liso o con manchas, hasta completamente blanco, nunca rojos, rosados o azules, frecuentemente con una cubierta cerosa blanquecina; pulpa abundante, totalmente de color amarillo-anaranjado pálido o anaranjado brillante o algunas veces anaranjada con un tinte verdoso-oscuro a negro en las placentas, de sabor ligeramente dulce a muy dulce en condiciones normales, de consistencia suave, granulosa y usualmente no fibrosa; pedúnculo rígido, leñoso, usualmente anguloso, con costillas obtusas o redondeadas que tienden a extenderse hacia el ápice del fruto, de hasta 20 cm de largo, por lo general notablemente

ensanchado en la unión con el fruto; semillas elípticas u ovado-elípticas, más o menos planas o muy ligeramente tumescentes, 8-21 mm de largo, 5-13 mm de ancho, algunas veces totalmente blancas o pardo claras, usualmente con el centro blanquecino-amarillento a pardo claro u oscuro y generalmente con los márgenes bien diferenciados en color y textura, ondulados y comúnmente fibrillosos o fimbriados, usualmente de color amarillento, dorado a pardo claro u oscuro, ápice obtuso a truncado (Lira, 1995).

Lira y Montes (1994) indican que la forma del ovario determinará la del fruto, un pepónide (baya modificada) de buen tamaño; el pedúnculo que la sujeta tiene una sección pentagonal característica. Puede ser lisa o segmentada, pocas veces con la superficie granulosa, y ovoide a cilíndrica, muchas veces con un engrosamiento basal. La corteza es verde a blanca o naranja claro, a veces irregular. La pulpa es blanquecino-verdosa cuando inmaduro, virando al naranja a medida que alcanza la madurez. Es raramente fibrosa, dulce y firme. En el interior del fruto hay hasta 200 semillas elípticas, achatadas, blancos grisáceos o amarillentos, de hasta 1 x 2 cm, con un núcleo blanco, dulce y rico en aceite.

Los zarcillos son 3-5- ramificados (Lira, 1995). Además, las raíces son fibrosas (Lira, 1995).



Figura 2: Frutos de loche. A. Pulpa y semilla. B. Frutos provenientes de Lambayeque. C. Frutos provenientes de Bagua. D. Pulpa y semilla.

2.3. ORIGEN Y DISTRIBUCIÓN

En alguna época se pensó que *Cucurbita moschata* tenía un origen asiático, al igual que *C. ficifolia*. Sin embargo, desde hace casi medio siglo es claro que se trata de una especie

domesticada en América Latina (Whitaker, 1947 citado por Lira, 1995; Ferriol et al., 2004b; Restrepo y Vallejo. 2008; Robinson et al., 2006), pero, no existe acuerdo acerca del área precisa de su domesticación. Robinson et al. (2006) indican que *C. moschata* fue domesticada en el área mesoamericana, incluyendo a América del Sur, y más específicamente a Colombia, como su centro de origen (Lira, 1995). Los restos arqueológicos más antiguos de esta especie (4900-3500 A.C.) han sido recuperados en las cuevas de Ocampo, Tamulipas, en el noreste de México (Lira, 1995). Por otro lado, Decker-Walters and Walters (2000) mencionan que los restos arqueológicos encontrados con mayor antigüedad fueron en el sur de México (7000 B.P) y en la costa del Perú (5000 BP). El análisis electroforético de isoenzimas no ha aportado pruebas importantes. No obstante, ha permitido reafirmar la fuerte relación entre esta especie y los taxones del grupo *C. argyrosperma* (Lira et al., 1991).

En la tabla 1 se presentan la síntesis de los registros arqueológicos de la especie *Cucurbita moschata*. Información obtenida de Kay et al. (1980), Turner y Miksicek (1984), Whitaker (1981, 1983) citado por Lira (1995).

Tabla 1: Restos arqueológicos de *C. moschata*.

| Lugar | Antigüedad |
|---|-----------------------|
| Puebla | 4900-3500 A.C. |
| Huaca Prieta | 2700-300 A.C. |
| Norte de Belice y Tikal, Guatemala | 2000 A.C.-850 D.C. |
| Valle Casma, Perú | 500 A.C. |
| Suroeste de Estados Unidos | 300 A.C. |
| Valle de Virú, Perú | 600-1100 D.C. |
| Chincha, Perú | 1310-1530 D.C. |
| Great Plains, Oeste de río Mississippi | 1700 D.C. |
| Valle de Oaxaca | 8750 A.C.-700 D.C. |
| Ocampo, Tamaulipas | 7000-5000 A.C. |

FUENTE: Lira (1995).

La Fundación “Kokefelli Seed” menciona que los vestigios de esta especie se han descubierto al hacer investigaciones arqueológicas no permiten afirmar con certeza que esta especie es original de América central o de América del sur. Los vestigios más antiguos han sido descubiertos en las grutas de Ocampo, Tamaulipas, al noroeste de México. Éstos datan

de un periodo que va de 4900 a 3500 años antes de Cristo. Se han descubierto también algunos en Huaca Prieta en Perú (3000 antes de Cristo), en Guatemala (de 2000 antes de Cristo a 850 después de Cristo). Siendo extremadamente variable en cuanto a la morfología de sus frutos y de semillas, no podemos llegar a ninguna conclusión en cuanto a la determinación de un centro del origen genético. La diversidad genética de esta especie es considerable en cuanto a la forma de sus frutos y de sus semillas, en cuanto a los ciclos de crecimiento, en cuanto a la resistencia a las enfermedades virales y en cuanto a la capacidad de conservación. Podemos mencionar la existencia de variedades muy resistentes a los virus cultivados por los pueblos Mayas o variedades con ciclos de crecimiento muy diferentes cultivadas en la península del Yucatán.

Considerando la compatibilidad genética de *C. moschata* con especies silvestres del género, se ha propuesto a *C. lundelliana* L.H. Bailey como la especie más relacionada (Whitaker, 1959, 1980 citado por Lira, 1995). De hecho, como resultado de la fuerte compatibilidad genética encontrada entre estas dos especies, Whitaker (1959) propuso dos hipótesis acerca del origen de *C. moschata*. La primera de ellas sugiere que este cultivo se pudo haber originado a partir de poblaciones de *C. lundelliana* por aislamiento y posteriores procesos de selección humana, mientras que la segunda propone que ambas especies pudieran proceder de un ancestro común y que debido a las presiones de selección humana, *C. moschata* se separó rápidamente tanto de dicho ancestro como de *C. lundelliana*. Uno de los principales argumentos en contra de estas hipótesis es la capacidad de *C. moschata* para producir híbridos altamente fértiles con los taxos silvestres del grupo *C. argyrosperma* (Merrick, 1990 citado por Lira, 1995). Aunado a esto algunas diferencias morfológicas entre *C. lundelliana* y *C. moschata* (color azul-grisáceo-verdoso de las semillas de la primera no se ha encontrado en *C. moschata*), y las evidencias derivadas de estudios de biología molecular, han contribuido a que se descarte a esta especie silvestre como el posible ancestro de *C. moschata* (Puchalski & Robinson, 1990; Wilson col 1992 citado por Lira, 1995).

Las evidencias lingüísticas y de distribución de la variación bajo cultivo tampoco son muy claras. *Cucurbita moschata* es conocida con nombres nativos tanto en la región mesoamericana (principalmente en México), como en América del Sur, lo cual, por otra parte, apoya la observación de que ambas regiones corresponden a importantes centros de diversificación de este cultivo. La variación morfológica de *C. moschata* tampoco puede ser

un factor que permita sugerir alguna región en particular como centro de origen, pues esta especie se ha reportado como sumamente variable en cuanto a las características de sus frutos y semillas para varias zonas dentro de sus área de distribución bajo cultivo en América (Bukasov, 1981, Lira, 1985, 1988, 1991 citado por Lira, 1995). Nee (1990) indica de la existencia en Bolivia y Colombia de poblaciones silvestres de *Cucurbita* que pudieran estar relacionadas con *C. moschata*, así como otros del primer tercio de este siglo que describen una importante diversidad morfológica en las razas locales de esta especie para el norte de Colombia. Bukasov (1981) sugiere que estas áreas pueden corresponder al centro de origen de este cultivo. Sin embargo, estas regiones del continente no han sido lo suficientemente exploradas con este objetivo como para dar una conclusión definitiva, aunque en la actualidad hay instituciones como el Herbario Nacional de Colombia, la Universidad de Santa Cruz y el Banco de Germoplasma de Pairumani en Bolivia y los Jardines Botánicos de Missouri y New York que llevan a cabo inventarios florísticos en diferentes zonas de Bolivia y Colombia y es muy probable que de estos estudios se derive alguna información al respecto.

La distribución geográfica de los restos arqueológicos de *C. moschata*, indican que esta especie es cultivada desde hace más de 5000-6000 años en toda América Latina (Lira, 1995). Es frecuente encontrar en la literatura que se trata de un cultivo que se maneja principalmente en zonas de baja altitud y de clima cálido con alta humedad (Esquinas-Alcazar y Gulick, 1983; Whitaker, 1968 citado por Lira, 1995). Sin embargo, aunque ciertamente esta especie es preferentemente cultivada dentro de dichos límites, estos no parecen ser tan estrictos. Por ejemplo, Bukasov (1981) fija los 2,200-2,300 m.s.n.m. como el límite altitudinal superior para el cultivo de esta especie en México y Colombia. Esto ha sido corroborado mediante la colecta de frutos y semillas de variantes que se cultivan a más de 2200 m.s.n.m. dentro de la región Mixteca Alta en el estado de Oaxaca, México (Lira, 1995).

La amplia adaptación a las más diversas condiciones ecológicas de las distintas variantes o razas de *C. moschata*, seguramente facilitó su temprana difusión hacia otros países fuera de Latinoamérica, tanto dentro como fuera del continente americano, tal como lo sugieren la existencia de la variante llamada “Seminola Pumpkin” cultivada desde tiempos precolombinos por grupos indígenas de Florida (Erwin y Lana, 1956), así como las ilustraciones botánicas de sus frutos realizadas en Europa en el siglo XVII (Whitaker, 1947

citado por Lira, 1995). Esta difusión debe haber sido lo suficientemente continua e intensa para poder alcanzar muchos otros puntos del mundo durante el siglo XIX, cuando *C. moschata* aparentemente ya se cultivaba en países como la región oriental de la India, Java, Angola y Japón, lo cual puede ser inferido por las numerosas colecciones de estos países del Viejo Mundo citadas por Cogniaux (1881). De acuerdo con estos datos al parecer *C. moschata* llegó a Europa procedente de Oriente, lo cual en cierta medida seguramente contribuyó a acuñar la teoría de su origen asiático. En la actualidad referencias bibliográficas y reportes de colecciones de germoplasma, indican que los frutos de *C. moschata* son muy apreciados como alimento humano y animal en Estados Unidos, India, Perú y Nigeria, además de que también se le cultiva en muchas otras regiones de África, Asia y Europa (Andres, 1990;; Chakravarty, 1990; Robinson, 1988a; Walters, 1989; Yang & Walters, 1993 citados por Lira (1995), Andres et al., 2006).

2.4. DIVERSIDAD

La amplia gama de altitudes en que *C. moschata* se cultiva dentro del continente americano, la notable diversidad morfológica de sus semillas y frutos (colores, formas, grosores y durabilidad de la cáscara del fruto), la existencia de variedades con ciclos de vida de diferente duración, así como la de numerosos cultivares desarrollados en otras partes del mundo y de variedades locales con características agronómicas sobresalientes, indican claramente que la variación genética de esta especie es vastísima (Lira, 1995).

Ejemplos de algunas variedades regionales interesantes para América Latina son las existentes en la península de Yucatán (y posiblemente en otras regiones de América Latina), con dos ciclos de vida de diferente duración, así como también las cultivadas en Guanajuato y Chiapas, en las que recientemente se ha encontrado resistencia a algunas enfermedades virales. Dentro de las primeras, es de gran interés la de ciclo corto cultivada comúnmente en los huertos mayas, pues de ella seguramente se derivó la variedad más importante a nivel comercial en la región; en cuanto a las segundas, cabe señalar que actualmente están siendo empleadas en programas de mejora genética (Lira, 1995).

Respecto a las fuentes de variación que presentan los cultivares de *C. moschata* desarrollados fuera de su área de origen, el mejor ejemplo es el de un cultivar nativo de Nigeria que representa la fuente única de resistencia a ciertas enfermedades virales. Las posibilidades de

hibridación que ha mostrado *C. moschata* con otras especies cultivadas (por ejemplo, *C. maxima*) permiten afirmar que existen buenas perspectivas respecto a la mejora de estos cultivos (Lira, 1995).

Otra parte del acervo genético de *C. moschata* son los numerosos cultivares comerciales que se han desarrollado principalmente en los Estados Unidos y en menor medida en Brasil. Entre ellos destacan 'Butternut Squash', 'Golden Cushaw', 'Large Cheese', 'Tennessee Sweet Potato', 'Kentucky Field', 'Menina Brasileira' y otros. Algunos de estos cultivares comerciales también presentan diferentes niveles de resistencia y/o susceptibilidad a ciertas enfermedades, lo que es indicativo de la gran variación genética de esta especie (Lira, 1995).

Cucurbita moschata es la especie de *Cucurbita* mejor representada en los bancos genéticos de América, donde se hallan depositadas más de 2,000 accesiones. Estas proceden principalmente de México y América Central, y en menor grado de América del Sur y otras regiones del mundo. Las accesiones más importantes son las de Estados Unidos y Costa Rica. En su totalidad corresponden a materiales americanos, principalmente de América Central. Por su parte, la colección del CIFAP en México es posiblemente la más representativa de la variación de la especie en ese país (Lira, 1995).

2.5. USOS

Cucurbita moschata se cultiva por sus brotes tiernos, consumidos como verdura, su flor carnosa y comestible, y sobre todo por sus frutos verdes que se emplean del mismo modo que el zucchini (*C. pepo*), cocidos de varias maneras o aún crudos en ensalada; maduros, se utilizan hervidos, asados u horneados, formando también parte de numerosos platos dulces. Las semillas se tuestan y consumen, o muelen para ser usadas como espesante o complemento para la harina (Lira y Montes, 1994). Por otro lado, Lira et al (1991) menciona que en la mayor parte del área nativa de *C. moschata*, sus flores, tallos jóvenes, frutos tiernos y fruto maduros son consumidos como verdura. Estos últimos, además, son comúnmente empleados para la elaboración de dulces, y como forraje. Las semillas son consumidas enteras, asadas o tostadas, y molidas en diferentes guisos. Presentan altos contenidos de aceites y proteínas (similares a los observados en *C. argyrosperma*) y su consumo en zonas urbanas también es bastante común.

El potencialidad medicinal de *C. moschata* has sido explorada ampliamente en toda su área de distribución bajo cultivo, principalmente en cuanto a su capacidad de aliviar una serie de padecimientos comunes (Tabla 2). El loche goza de excelentes propiedades terapéuticas en las enfermedades agudas del aparato digestivo, especialmente en la inflamación de los intestinos, en la fiebre tifoidea y en la disentería (FAO, 1989). Para calmar los dolores de cabeza se aplica tajadas de loche cruda en la frente, varias veces. Contra las mordeduras de los perros y otros animales ponzoñosos, se usa cataplasmas tibias de loche rallado o molido (García, 2006). La pulpa se destaca por su efecto diurético, suavizando y protegiendo la mucosa del estómago, indicado su consumo en forma de crema en casos de acidez de estómago, gastritis, mala digestión y úlcera gastroduodenal. También favorece a la cicatrización de la piel por quemaduras. Las hojas se utilizan en estados febriles, recomendadas durante el embarazo. También resulta eficiente para disminuir la diarrea fuerte.

Tabla 2: Potencial medicinal de *Cucurbita moschata*.

| País o región | Partes usadas | Usos o propiedades |
|----------------------|----------------------------|--|
| Brasil | Fruto, semillas | Diurético, tenífugo y vermífugo |
| China | Rapiz, flores, semillas | Dolores de dientes |
| Colombia | Raíz | Febrífugo |
| Jamaica | Semillas | Diurético |
| Península de Yucatán | Raíz, hojas, flores, fruto | Curación de mordedura de serpiente y enfermedades de la piel, jugo para curar granos, erupciones de la piel, estimulantes del apetito. El fruto se considera pectoral, refrescante y útil en la curación de enfermedades del cuero cabelludo |
| México en general | Semillas | Antihelmíntico |
| Venezuela | Semillas | Fiebres eruptivas y el aceite para las úlceras |

Fuente: Lira (1995)

2.6. AGRONOMÍA

Las plantaciones de *C. moschata* se alternan tradicionalmente con el maíz (*Zea mays*) y alguna fabácea, normalmente los porotos (*Phaseolus vulgaris*), y a veces con alguna otra cucurbitácea, como el melón (*Cucumis melo*), que tiene requisitos muy similares. Se planta a comienzos de la estación lluviosa; el plazo para la recolección varía según el uso al que se

destine y las características del cultivar. Va de tres meses para los brotes y frutos verdes hasta siete para el fruto completamente maduro. (Lira, 1995).

Los principales centros de desarrollo de nuevos cultivares son Estados Unidos, Brasil y Argentina, donde el Instituto Nacional de Tecnología Agropecuaria ha llevado a cabo un intenso trabajo a lo largo de los años 1990 por la fácil conservación y moderados requisitos de cultivo de la especie (Lira, 1995).

Es probable que el cultivo de variedades como las descritas, y posiblemente otras más, sea más común en el continente americano de lo que se piensa o se conoce. Existen antiguas referencias de una notable variación en Colombia, pero su situación actual ha de ser debidamente documentada y evaluada (Lira, 1995).

El loche es una planta que no necesita muchos insumos para cultivarse, prefieren suelos sueltos, moderadamente tolerantes a la acidez en un pH 5,5-6,8. Es bastante resistente a la sequía pues el sistema radical puede llegar hasta 1,5 m de profundidad. Son medianamente invulnerables a la salinidad del suelo. Para la producción temprana se facilitan los suelos arenosos, ricos en materia orgánica (Anónimo, 2007).

La cosecha se lleva a cabo a los 3-5 meses de la siembra. Cuando los frutos maduran, cambian de color y su piel se endurece, estarán listos para su recolección. La recolección del loche se realiza en forma manual dejando siempre un pedúnculo de unos pocos centímetros, sobre todo si se pretenden almacenar. Los que se cultivan para primicia se los cosecha antes de llegar a plena madurez, a mitad o 3/4 de cáscara. Las calabazas que van a conservarse, se cosechan cuando el follaje se ha secado y la corteza es bien dura (Anónimo, 2007).

Los principales problemas que sufren las plantas de loche son la de los hongos como el “Mildiu Velloso” (*Pseudoperonospora*) y “Mildiu polvoriento” (*Erysiphe cichoracearum*) que son los más graves para el cultivo (Anónimo, 2007). Mosca Blanca (*Bemisia* sp.) el agente causal de la transmisión del plateado, por lo general necesita control en las primeras fases del cultivo (Anónimo, 2007).

2.7. POTENCIAL DEL CULTIVO

Como se observa en la tabla 3 el componente principal de la pulpa del loche es el agua, lo que, unido a su bajo contenido en hidratos de carbono y a su casi inapreciable cantidad de grasa, hace que sea un alimento con un escaso aporte calórico, proporcionando solamente 50 calorías por 100 gramos.

En relación con las vitaminas, el loche es rico en beta-caroteno o pro vitamina A y vitamina C. Presenta cantidades apreciables de vitamina E, folatos y otras vitaminas del grupo B tales como la B1, B2, B3 y B6. La vitamina A es esencial para la visión, el buen estado de la piel, el cabello, las mucosas, los huesos y para el buen funcionamiento del sistema inmunológico, además de tener propiedades antioxidantes (Moreiras et al., 2001).

La vitamina C se encuentra en cantidades apreciables, con 100 gramos de loche, se cubre el 20% de las ingestas diarias recomendadas, interviene en la formación de colágeno, glóbulos rojos, huesos y dientes. También favorece la absorción del hierro de los alimentos y aumenta la resistencia frente las infecciones (Moreiras et al., 2001).

Es buena fuente de fibras solubles que ofrece valor de saciedad y mejora el tránsito intestinal por la alta presencia de mucílagos. Éstos son fibra soluble que tiene la capacidad de suavizar las mucosas del tracto gastrointestinal es aconsejable su uso en casos de obesidad y estreñimiento (Casper, 2001).

En cuanto a su riqueza mineral, el loche es un alimento rico en potasio. También contiene otros minerales como fósforo y magnesio, pero en menores cantidades. El potasio es un mineral necesario para la transmisión y generación del impulso nervioso y para la actividad muscular normal, además de intervenir en el equilibrio de agua dentro y fuera de la célula (Moreiras et al., 2001).

Tabla 3: Composición bromatológica de *C. moschata* (100g).

| Compuesto | Semilla | Pulpa |
|-------------------------|---------|-------|
| Agua (g) | 7.7 | 88.3 |
| Proteína (g) | 30.5 | 1.2 |
| Grasa (g) | 27.9 | 0.3 |
| Carbohidratos tot.(g) | 30.8 | 30.8 |
| Fibra cruda (g) | 4.9 | 0.6 |
| Ceniza (g) | 3.1 | 0.4 |
| Calcio (mg) | 35 | 12 |
| Fósforo (mg) | 725 | 27 |
| Hierro (mg) | 1.1 | 0.7 |
| Act.de vitamina A (ug) | 0 | 1055 |
| Tiamina (mt) | 0.13 | 0.05 |
| Riboflavina (IPA) | 0.12 | 0.04 |
| Niacina (mg) | 2.2 | 0.6 |
| Ácido ascórbico (mg) | 0 | 42 |
| Valor energético (kcal) | 465 | 40 |

Fuente: Brack, 1999.

2.8. CARACTERIZACIÓN MOLECULAR

2.8.1. Colecta de material vegetal

Una de las consideraciones más importantes en cualquier procedimiento de extracción de ADN de plantas es la manera utilizada para colectar y preservar el tejido vegetal. En la colecta se debe cumplir con algunas condiciones: el material debe ser lo más tierno posible, no debe maltratarse, cortarse, ni pulverizarse. Con ello se evita generar una mayor superficie expuesta al oxígeno y desencadenar proceso oxidativos de polifenoles (Couch y Fritz, 1990). La cantidad y calidad de ADN son afectadas en general por la condición del tejido antes de la extracción. Esto es particularmente importante para especies que producen grandes cantidades de metabolitos secundarios que en general interfieren con una extracción exitosa de ADN.

El material fresco puede ser colectado y mantenido en hielo o en un frigorífico durante algunos días hasta ser procesado. Si fuera necesario un almacenamiento prolongado, el material puede ser congelado a -20°C durante algunas semanas, o -80°C indefinidamente y descongelado justo antes de la extracción (Ferreira y Grattapaglia, 1998).

2.9. EXTRACCIÓN DE ADN

2.9.1. Solubilización de membranas

Se emplea un detergente con el propósito de liberar el contenido nuclear y disociar la membrana celular. Un procedimiento que presentan algunas modificaciones para resolver problemas específicos de especies con gran contenido de metabolitos secundarios es el método que utiliza un detergente catiónico CTAB (cetyl trimethyl ammonium bromide) de carga positiva, muy adecuado para el aislamiento de ADN de alto peso molecular, el cual se pega fuertemente al ácido nucleico, desplazando la proteína previniendo la degradación del mismo. El CTAB se remueve mediante extracciones con cloroformo y el ADN permanece en fase acuosa listo para ser precipitado con etanol (Valadez y Kahl, 2000).

El CTAB utiliza un sistema Tris.HCL pH 8.0, para mantener el pH constante. Este además de solubilizar las membranas lipoproteicas, dependiendo de la concentración de NaCl forma un complejo con el ADN, y por lo tanto, puede ser utilizado para precipitar el mismo selectivamente en presencia de alcohol. La concentración utilizada de NaCl en el tampón de extracción es de 1.4 M., siendo ésta necesaria para la solubilidad y la acción del CTAB.

Otro tipo de detergente aniónico es el SDS (*sulfato dodecil de sodio*), este tiene la capacidad adicional de desnaturalizar proteínas globulares en sus subunidades (Lehninger et al., 1993); siendo las enzimas la mayor parte de este tipo de proteínas. También se puede incluir este detergente en el tampón sin excluir en CTAB, ya que pueden usarse en el mismo protocolo (Dellaporta et al., 1983).

2.9.2. Inhibición enzimática

El ADN debe ser protegido de la acción degradativa de las enzimas (nucleasas) o compuestos secundarios liberados con la rotura de las células. Por eso se utilizan componentes esenciales como el β – mercaptoetanol, que es un agente reductor que inhibe las actividades enzimáticas, protegiendo el ADN de las peroxidasas y polifenoloxidasas (Ferreira y Grattapaglia, 1998). Así como éste existen otros como el acomplejante de los cofactores enzimáticos, EDTA (ethylen diamin tetra acetato), que es un quelante por excelencia, que capta cationes bivalentes tales como Mg^{2+} y Ca^{2+} , también se utiliza la proteinasa K, o incubar a 65°C (Ferreira y Grattapaglia, 1998).

2.9.3. Complejos de ADN

El CTAB junto con iones Na^+ forma un complejo con el ADN (Dellaporta et.al., 1983), estabilizándolo y protegiéndolo de ataques enzimáticos y/o químicos. La espermina y la espermidina, dos poliaminas catiónicas que se encuentran asociadas al ADN de bacterias y virus forman complejos con aquél estabilizándolo, ventaja (Cheng et.al. 1997) para precipitar selectivamente a la macromolécula.

2.10. CUANTIFICACIÓN Y DETERMINACIÓN DE LA CALIDAD DEL ADN

La cuantificación y determinación de la calidad del ADN se precisa de la observación de bandas íntegras que correspondan a un ADN no degradado; cuando se observa un barrido (smear), este corresponde a un ADN degradado de mala calidad, no apto para diversas técnicas moleculares.

Las técnicas utilizadas en la cuantificación del ADN, dependen de los equipamientos muchas veces costosos, como espectrofotómetros, mientras que otras menos precisas, sencillas y baratas. La opción por uno u otro método depende de la condición del laboratorio. Dos técnicas bastante utilizadas en la cuantificación del ADN son las realizadas mediante fluorímetro o por comparación de los fragmentos en geles coloreados con bromuro de etidio.

La cuantificación mediante fluorímetro se basa en la adición de un compuesto (DAPI: 4', 6-diamino-2fenilindol) a la solución de ADN. El DAPI es un compuesto que se une a la molécula del ADN como una sonda y que emite fluorescencia al ser excitado por el haz de luz del fluorímetro. Una vez obtenida una curva de calibración de valor del fluorímetro respecto a la concentración de muestra patrón de ADN se prosigue a la lectura de las muestras de ADN que se desea cuantificar.

La cuantificación de ADN, utilizando el análisis comparativo de muestras coloreadas con Bromuro de Etidio en geles de agarosa, consiste en comparación de la intensidad de la banda del ADN con la primera banda del marcador de peso molecular (ADN del fago lambda λ cortado con la enzima Pst I), cuyo tamaño es de 14,800 pb y su intensidad corresponde a una concentración de 280 ng/ul.

El cálculo de las concentraciones de las bandas se hace de imagen digitalizada del propio gel de agarosa. Es de extrema importancia evitar la cuantificación por observación directa de las bandas bajo luz ultravioleta, incluso con máscaras de protección. (Ferreira y Grattapaglia, 1998).

2.11. ELECTROFORESIS

La electroforesis es la migración de partículas cargadas a través de un soporte, las cuales están sometidas a un campo eléctrico (Freifelder, 1991). Los materiales para la matriz más utilizados en la electroforesis son la poliacrilamida, la agarosa y el almidón. Los geles de agarosa y poliacrilamida pueden prepararse en una variedad de formas, tamaños y porosidades y pueden ser corridos en varias configuraciones diferentes. La elección de estos parámetros depende particularmente del tamaño de fragmentos que se van a separar.

Los geles de agarosa (de 0.8 a 2.0%) son muy útiles para separar fragmentos de ADN cuyo tamaño está comprendido en el intervalo de 300 a 10,000 pb. Tiene menor poder de resolución que geles de poliacrilamida pero tienen un mayor rango de separación. Los geles de agarosa se corren usualmente en una configuración horizontal en un campo eléctrico de carga y dirección constante.

Los geles de poliacrilamida (de 3.5 a 20%) son muy útiles para fragmentos cuyo tamaño oscile entre 20 y 1,000 pb. Su poder de resolución es extremadamente alto y fragmentos de ADN que difieren en tamaños más pequeños como 1 pb pueden ser separados el uno del otro. Aunque pueden correrse muy rápidamente y pueden ser colocados numerosas muestras de ADN, los geles de poliacrilamida tienen la desventaja de ser más difíciles de preparar y manejar que los geles de agarosa. Generalmente se corren en una configuración vertical en un campo eléctrico constante.

Los geles se someten a un campo eléctrico que hacen migrar los fragmentos de ADN según su tamaño, de modo que los fragmentos grandes migran más lentamente que los fragmentos cortos. Las moléculas de ADN, que están cargadas negativamente a pH neutro, migran hacia el ánodo.

La distancia de migración es proporcional al logaritmo del número de bases. Por consiguiente, el tamaño real de los fragmentos obtenidos puede calcularse partiendo de la movilidad de los fragmentos de ADN de tamaño conocido.

2.12. REACCIÓN EN CADENA POLIMERASA

La PCR fue descubierta por Mullis et al. (1986), este proceso consiste en hacer múltiples copias de un segmento de ADN, se dan en tres etapas: denaturación, unión de iniciadores o “annealing” y polimerización. Chien et al. (1976) desarrollaron la técnica de purificación de

la polimerasa de *Termus aquaticus* cuya característica más peculiar es que tiene una temperatura óptima de 80°C y es termoestable para un amplio rango de temperatura. El uso de esta polimerasa para la PCR simplificó grandemente el procedimiento y permitió que la reacción se realizara a mayores temperaturas mejorando la especificidad, la cantidad de producto, la sensibilidad y la longitud de los productos que pueden ser amplificados (Saiki et al., 1988). Actualmente los marcadores moleculares basados en la PCR son los más ampliamente usados por su fácil aplicación y porque pueden aplicar a gran escala (Ferreira y Grattapaglia, 1998).

Erlich (1989) menciona que la reacción en cadena de la polimerasa conocido como PCR es una técnica in vitro que amplifica fragmentos específicos de ADN a partir de cantidades mínimas del ADN patrón, mediante el cual se obtiene millones de copias de secuencias específicas del genoma de un organismo (Coombs, 1986; Fleschar y Kimball, 1993).

Es un proceso exponencial, desde la amplificación de los productos en cada ciclo, que sirve como modelo para el próximo ciclo de amplificación, haciendo así de él una técnica extremadamente sensible (Santos et al 2004).

El proceso de amplificación involucra la repetición de varios ciclos, cada uno de los cuales consta de 3 etapas.

Desnaturalización

La solución de reacción contiene moléculas de ADN blanco, polimerasas (enzima que copia el ADN), iniciadores o (“primers”), que sirven en la polimerización de del DNA, y nucleótidos (Adenina, Timina, Citosina, y Guanina). La reacción se calienta a altas temperaturas (90-95°C), para reducir la doble hélice del ADN a cadenas simples. Estas cadenas se tornan accesibles a los iniciadores.

Alineamiento o hibridación

Se enfría la mezcla de reacción, la temperatura desciende (45-65°C). Estas condiciones favorecen la hibridación de los iniciadores (“primers”) con las regiones complementarias a las cadenas de ADN patrón, formándose nuevas cadenas dobles entre los iniciadores y las secuencias. Las polimerasas entonces comienzan a añadir nucleótidos complementarios en estos sitios fortaleciendo la unión entre los primers y el ADN en estudio.

Extensión

En este paso la temperatura es otra vez incrementada, esta vez a 72°C, temperatura ideal para las polimerasas usadas, una de ellas la Taq polimerasa que se denominó así por la bacteria *Thermus aquaticus* aislada originalmente de aguas termales, vive a 75°C siendo estable a 94°C, lo que ha permitido la automatización del procedimiento por que la enzima se aplica una sola vez. Conocida también como ADN polimerasa, que usa una porción de la cadena simple del DNA como molde para la síntesis de la cadena complementaria nueva en dirección 5' a 3'. Puede resistir a altas temperaturas necesarias para separar las cadenas del DNA permaneciendo en el tubo de reacción. La enzima lee la secuencia de la cadena opuesta y extiende los iniciadores agregando nucleótidos en el orden en que pueden emparejarse; al mismo tiempo, cualquier unión pérdida que se forma entre los primers y segmentos de ADN que no son plenamente complementarias se rompen.

El ciclo de calentamiento y enfriamiento se repite una y otra vez, lo que estimula a los “primers” a unirse a las secuencias originales y a las secuencias recién sintetizadas. La síntesis se inicia en el “primer”, que es una porción de DNA de cadena corta con una longitud de 10 - 25 pb. La parte que amplifica depende del “primer” específicos que dirige la síntesis de una región específica de DNA. Diferentes “primers” crean diferentes patrones de bandas en el gel de acrilamida o agarosa, la diversidad del sistema marcador resulta de la variación en la longitud y secuencia de los “primers”. (Sevilla y Holle, 2004).

Luego, la reacción se desnaturaliza por calor para separar la cadena original de la nueva, la cual queda disponible para nuevos ciclos de hibridación, síntesis de DNA y separación de cadenas. Luego se separa la porción nueva de la original y se usa nuevamente.

Este ciclo de temperaturas produce copias y luego copias de copias, y así sucesivamente, lo que lleva a un aumento exponencial del número de copias de determinada secuencia. Dado que la cantidad de ADN colocado en el tubo al comienzo es muy pequeña casi todo el ADN presente al final de los ciclos de reacción pertenece a secuencias copiadas.

Los productos de la reacción se separan mediante electroforesis. Según la cantidad producida y el tamaño del fragmento amplificado, los productos de reacción se pueden visualizar

directamente mediante tinciones con bromuro de etidio, plata, o mediante radioisótopos y autoradiografía.

Las etapas de la PCR se llevan a cabo, una tras otra, en episodios cíclicos. Un ciclo (1) es como sigue:

1. Durante la desnaturalización (cerca de 1 min a 95°C), las cadenas del ADN se separan para formar cadenas sencillas.
2. Durante la hibridación (cerca de 1 min a temperaturas que oscilan entre 45°C y 60°C), un iniciador se une a una cadena de ADN y otro se une a la cadena complementaria. Los sitios de hibridación de los iniciadores se han elegido para que fomenten la síntesis del ADN en la región de interés durante la extensión.
3. Durante la extensión (cerca de 1 min a 72°C), la síntesis del ADN se lleva a cabo en la región de interés y con distancias variables en la región flanqueante, produciendo fragmentos de longitudes variables.

2.13. CARACTERIZACIÓN DEL GENOMA

2.13.1. Diversidad Genética

La diversidad genética es un componente de la diversidad biológica y se refiere a la frecuencia y diversidad de los diferentes genes y/o genomas. Incluye las variaciones tanto dentro de una población como entre poblaciones. Según Sevilla y Holle (2004) la diversidad genética presenta dos niveles: el primero es el número de las categorías intra-específicas de la especie tales como: razas, ecotipos, variedades, morfotipos, entre otros; y el segundo se refiere a la variación en loci específicos de marcadores genéticos (Sevilla y Holle, 2004). De manera semejante Brush (1995) indica que la variabilidad se refiere a la heterogeneidad de los alelos y genotipos con sus correspondientes fenotipos y morfotipos.

La diversidad genética es el resultado de la selección, la mutación, la migración, la deriva genética o la recombinación. Estos fenómenos ocasionan cambios en las frecuencias génicas y genotípicas, lo que conduce a la evolución de las poblaciones (Sevilla y Holle, 2004).

Es importante para mantener un reservorio de condiciones de variación de respuesta al medio, que permita la adaptación y la supervivencia. El uso y manejo de la diversidad genética de los cultivos mantenida “in situ” ha ayudado a los campesinos a sobrevivir bajo

las más difíciles condiciones por lo que es considerada la principal vía para la seguridad alimentaria de las generaciones futuras (Collins y Hawtin, 1999).

El conocimiento del nivel de diversidad es fundamental en el planeamiento de actividades de conservación ya que sin monitoreo no es posible verificar la efectividad de la conservación (Brush, 1995; citado por Negri y Tosti, 2002). Estudios empíricos en ecología y evolución frecuentemente dependen de una evaluación precisa de la diversidad genética para resolver cuestiones referidas a las relaciones entre individuos, la estructura poblacional, relaciones filogenéticas y mapeo de caracteres cuantitativos (Mueller y Wolfenbarger, 1999). Los estudios de diversidad genética basados en marcadores moleculares pueden brindar excelente información sobre la diferenciación genética y la estructura genética. Además pueden ser de gran utilidad para validar la clasificación taxonómica, validar la procedencia de las accesiones, identificar redundancias y mejorar el balance en la representación de las especies en los bancos de germoplasma (Van Hintum y Van Treuren, 2002). Adicionalmente, la información de la diversidad genética en material de mejoramiento es esencial para el diseño de programas de mejoramiento óptimos. Esto implica, por ejemplo, la selección de parentales para el establecimiento de poblaciones de cruzamiento, el reconocimiento de grupos heteróticos, y la introgresión de germoplasma exótico.

Además, otro interés para los mejoradores es el registro y protección de cultivares (Muminovic et al., 2004). La pérdida o reducción de la diversidad genética, se puede definir como erosión genética. Esta pérdida puede ser natural o causada por el hombre. Los factores que contribuyen a la erosión de la diversidad biológica se pueden clasificar en las siguientes categorías fundamentales (Soulé y Mills, 1998): crecimiento poblacional, pobreza, antropocentrismo, transiciones culturales, económicas y políticas.

La caracterización del genoma se hace a nivel del cromosoma. La caracterización puede ser química, morfológica y genética o a nivel molecular. Las dos primeras son indispensables para la molecular. O sea para usar plenamente los marcadores moleculares en una especie es necesario conocer el número y forma de los cromosomas. Los genes que gobiernan características morfológicas, fisiológicas, citogenéticas y agronómicas de importancia deben estar localizados en los cromosomas.

Los marcadores se usan para localizar los genes en los cromosomas; para construir mapas genéticos. Se aplican en el estudio del origen y evolución de las especies; la estimación de la diversidad genética y las relaciones de parentesco; en la genotipificación (“fingerprinting”); identificación y diferenciación varietal; caracterización genética de poblaciones y para asegurar la transmisión de genes deseables o ligados a los marcadores durante la selección y para detectar genes deseables, o alteraciones a nivel del ADN.

Los marcadores tratan de solucionar los defectos que tiene la observación fenotípica en poblaciones segregantes, que es el método que usaba la genética clásica para definir la herencia de los caracteres.

Para que un carácter sea considerado un marcador genético debe mostrar una variación experimentalmente detectable entre los individuos de la población y un modo de herencia predecible según las leyes de Mendel. Esta variación puede ser considerada a diferentes niveles biológicos, desde cambios fenotípicos heredables, hasta la variación de un solo nucleótido. Un marcador ideal debe ser: altamente polimórfico o variable dentro y entre especies, de herencia mendeliana no epistática (sin interacción entre genes), insensible a los efectos ambientales, codominante, de rápida identificación y de simple análisis y de posible detección en los estadios tempranos del desarrollo de la planta. En los análisis genómicos, se han venido utilizando varios tipos de marcadores genéticos: los marcadores morfológicos, bioquímicos (izoenzimáticos), y los marcadores moleculares basados en ADN (Picca et al., 2004).

2.14. MARCADORES MORFOLÓGICOS

Son los caracteres observables que se presenta en los individuos de la población, y que pueden ser genéticos en la medida de que sean heredables. La desventaja es que son muy pocos los marcadores de este tipo, y su presencia puede estar influenciado por el factor ambiental. En el caso del loche, no se han establecido marcadores morfológicos.

Estos tipos de marcadores se presentan en la planta adulta, y también tienden a cambiar por condiciones externas: fisiológicas y ambientales; lo que implica dificultades para hacer una efectiva selección y en el menor tiempo. Por eso, se inventó los marcadores moleculares, que rápidamente complementan las estrategias clásicas de mejoramiento (Weising et al., 1995).

Las principales limitaciones de los marcadores morfológicos se encuentran en:

- Número reducido de marcadores disponibles en cada población.
- Bajo nivel de polimorfismo.
- Pueden producir alteraciones fenotípicas que dificulten el desarrollo de la planta.
- Varios se hallan bajo control poligénico.
- Dominancia
- Muchos de ellos se expresan en estadio de planta adulta, lo cual prolonga los tiempos de evaluación en los programas de mejoramiento.
- El hecho de que varios factores puedan modificar la expresión del gen limita su utilidad como marcador genético.

2.15. MARCADORES MOLECULARES

Marcador molecular es todo fenotipo molecular producto de la expresión de un gen correspondiente a zonas expresadas o no del genoma. Cuando se verifica su comportamiento de acuerdo con las leyes mendelianas se le llama marcador genético. (Ferreira y Gratapaglia, 1998).

González-Andrés (2001) indica que un marcador molecular es la secuencia de ADN que puede ser fácilmente detectable y cuyo mecanismo hereditario puede ser analizado. Un marcador ideal debería presentar suficiente variación para el problema en estudio, ser confiable y simple de generar e interpretar. Existen problemas asociados con todas las técnicas de marcadores moleculares existentes. Estos se pueden clasificar en problemas prácticos, de tipo de datos y de análisis. Lamentablemente hasta el momento no existe este marcador ideal y los investigadores deben optar por aquel que se ajuste mejor a sus objetivos y a sus recursos; esto último es lo que prima muchas veces en la práctica (Robinson y Harris, 1999). Actualmente los SSR y los AFLP, unidos a la información de secuenciamiento son las principales técnicas para el análisis de la variación genética (Mueller y Wolfenbarger, 1999). Asimismo, debido al avance de la tecnología de la secuenciación de genomas, es posible hoy la identificación de marcadores SNPs.

Tradicionalmente, los marcadores utilizados en estudios de genética y mejoramiento eran aquellos controlados por genes asociados a caracteres morfológicos, en general fenotipos de fácil identificación visual, que contribuyeron significativamente al desarrollo teórico del

análisis de ligamiento y en la construcción de las primeras versiones de mapas genéticos (Ferreira y Grattapaglia, 1998). Sin embargo, el bajo número de estos marcadores y la posibilidad de interferencia epistática o ambiental limitaban su uso. La revolución en este plano se inició con el descubrimiento y utilización de los marcadores isoenzimáticos. El número de marcadores genéticos disponibles fue ampliado y la aplicación de la técnica se expandió prácticamente a todas las especies de plantas (Ferreira y Grattapaglia, 1998). Con la llegada de las técnicas modernas de la biología molecular, surgieron diversos métodos de detección de polimorfismo genético directamente a nivel del DNA, los llamados marcadores moleculares, que en general, pueden acelerar los programas de mejoramiento genético.

Actualmente se han convertido en una herramienta muy útil para la conservación del germoplasma, para determinar la identidad de las accesiones y los cambios genéticos que pueden ocurrir durante la conservación y regeneración, para descartar duplicados y vacíos en los bancos de germoplasma, para mantener la identidad de las colecciones durante la regeneración, para determinar el grado de similitud entre individuos de una accesión o entre accesiones dentro de una colección; para definir la estructura genética y la variación de individuos, accesiones, poblaciones o especies; para formar las colecciones núcleo, para caracterizar y clasificar el germoplasma y para facilitar su uso y manejo. En comparación con los caracteres morfológicos los marcadores moleculares tienen muchas más ventajas, y su expresión es independiente de las condiciones ambientales (Lombard et al., 2000). El marcador molecular ideal debe tener las siguientes cualidades:

Neutralidad fenotípica

A diferencia de los marcadores morfológicos, los marcadores moleculares no presentan influencia fenotípica, esto significa que la interacción entre el medio ambiente y las regiones de ADN, empleadas como marcadores, es nula, ya que estas no son codificantes.

Ser polimórfico

Es decir presentan diferente expresión en las diferentes O.T.U. (Unidad Taxonómica Operativa por sus siglas en castellano) a analizar.

Los marcadores moleculares están teóricamente disponibles en un número ilimitado, a menudo rápido y fácil de usar, exploran una alta variabilidad y por consiguiente un nivel más perceptible de polimorfismo comparado con los descriptores morfo-fisiológicos o marcadores bioquímicos (Faccioli et al, 1999)

Esta característica es muy importante, ya que depende directamente de la variación alélica del marcador, que está condicionada a las características reproductivas del organismo, a las migraciones (flujo génico), así como a los eventos moleculares que presenta la herencia del marcador, recombinación genética, mutaciones puntuales, inserciones/deleciones.

Abundancia

Está referida a la ocurrencia de variaciones alélicas Teóricamente es posible detectar todos los poligenes de una especie.

Ampliamente distribuido en el genoma

Cuando lo que se pretende conocer es la diversidad genética entre genotipos, interesa explorar la mayor parte del genoma, y en consecuencia interesa que los marcadores estén distribuidos por todo el genoma. Obtener marcadores moleculares SNPs distribuidos en todo el genoma empleando la técnica de Genotyping-by-Sequencing (Elshire et al., 2011).

Que sean específicos de un determinado locus

Quiere decir que cada banda corresponda a un único locus, ya que con frecuencia en determinados marcadores moleculares basados en ADN existe el problema de la no homología, es decir que bandas correspondientes a fragmentos de igual movilidad electroforética, corresponde a diferentes loci.

Heredable y con herencia codominante

Cuando es un individuo se presentan dos alternativas alélicas de un locus, ambas se manifiestan. Esto permite calcular las frecuencias de cada uno de los alelos en cada uno de los loci.

Ausencia de epistasis y pleiotropía

La epistasis es el tipo de interacción génica en el cual la acción de uno o más genes se ve influenciada por la expresión de genes situados en diferentes loci. La pleiotropía es la expresión múltiple de un gen, influyendo sobre procesos bioquímicos distintos a los que tienen como objetivo principal. Los marcadores moleculares están exentos de estos efectos de interacción génica, razón por la cual es posible emplear un gran número de marcadores sobre una población para mapear poligenes. (Trujillo, 2004).

Reproducible entre laboratorios y dentro de un mismo laboratorio

Fácil, rápido y económico de detectar.

El presente estudio emplea los marcadores microsatélites o SSRs (Simple Sequence Repeats, en inglés). Estas son secuencias cortas de ADN, de 1 a 6 nucleótidos, repetidas cierto número de veces (Fig. 9a y 9b) y que se encuentran esparcidos por todo el genoma de los organismos eucariontes (Tautz, 1989) y procariontes (Zane et al., 2002). Tautz y Renz (1984) menciona que los microsatélites son regiones del genoma constituidas por secuencias repetidas en tándem, estas secuencias o motivos repetidos de 1 a 5 pb, están ampliamente distribuidos tanto en genomas de eucariotas y procariotas, son ubicuos, es decir están repartidos por todo el genoma.

Los parámetros más importantes para utilizar marcadores moleculares en la caracterización son: novedad, distinguidad, homogeneidad y estabilidad. Así deben ser capaces de demostrar un alto poder de discriminación, no deben presentar interacción con el ambiente, deben ser capaces de producir resultados equivalentes entre distintos laboratorios, deben permitir el cálculo de distancias entre genotipos en forma consistente con otros caracteres, como información de pedigree. Entre las principales técnicas de marcadores moleculares, la más adecuada a esos parámetros es la de microsatélites. Actualmente, este tipo de estudio tiene un gran impacto, pues son apuntadas como las herramientas más poderosas para identificación/caracterización de variedades (Varshney et al. 2005). Sin embargo, debido a la reducción de costos y mayor incremento en la confiabilidad de secuenciamientos de genomas, hoy es posible el común uso de marcadores SNPs, siendo de gran aceptación por parte de la comunidad científica para estudios de diversidad, genética de poblaciones, entre otros.

Cuando estas regiones de microsatélites son individualmente amplificadas por medio de la reacción en cadena de la polimerasa (PCR), utilizando un par de oligonucleótidos flanqueantes como iniciadores, muestran casi invariablemente polimorfismo debido a las diferencias en su longitud, como consecuencia de la ocurrencia de los diferentes números de las unidades de repetición o motivos (Morgante y Olivieri, 1993). El origen de tal polimorfismo está aún en debate, aunque previsiblemente este se deba al evento de

“resbalón” durante la replicación del DNA (Schlotterer y Tautz 1992, Zane et al., 2002), en el cual ya sea la cadena nueva o la cadena patrón se desliza sobre la otra en al menos un motivo, resultando en un mal apareamiento de las mismas (Moxon y Wills, 1999). Otras posibles causas para la generación de polimorfismo se deban a otros tipos de mutaciones, como las deleciones e inserciones, ya sean en el microsatélite en si o en las regiones que lo flanquean alterando el tamaño de la región microsatélite. La duplicación de una secuencia microsatélite en otra región del genoma da lugar a la generación de marcadores microsatélites multilocus. (Jarne y Lagoda, 1996).

Los motivos de repeticiones microsatélites más frecuentes son diferentes dependiendo de la región del genoma en que se encuentren (Tóth et al., 2000). Además, se ha encontrado que existe una variación de las frecuencias de distribución de los motivos de los microsatélites relacionada al grupo taxonómico (Tóth et al., 2000). Por ejemplo, el dinucleótido (AT)_n es la secuencia repetitiva más común en genomas de plantas (Morgante y Olivieri, 1993), mientras que el motivo (AAAN)_n es el más común en humanos (Katti et al., 2001).

Los microsatélites otorgan una gran cantidad de información extra comparada con otros marcadores gracias a que frecuentemente presentan un alto número de alelos en bajas frecuencias (alelos privados), es decir que se pueden encontrar alelos que se presentan en una o pocas poblaciones. Esta característica contribuye enormemente a la evaluación de las relaciones entre poblaciones e individuos (Scotti, 1999). Debido a sus altos niveles de polimorfismo han sido ampliamente usados como marcadores en estudios de tipificación, estructura de poblaciones y mapeo genético (Tóth et al., 2000) Feldman et al., 1997, recomendaron estudiar más de 15 a 20 loci para estimar relaciones interespecíficas de manera confiable.

Las secuencias que flanquean un microsatélite son las que se utilizan como zona de anclaje de los iniciadores y son generalmente conservadas dentro de una especie (Tautz, 1989) y a veces hasta niveles taxonómicos superiores (Provan et al., 1996) de ahí que se pueden usar para determinar las relaciones históricas entre microsatélites (Blankenship et al., 2002).

Los microsatélites fueron inicialmente descritos en humanos, y en poco tiempo fueron hallados en otros mamíferos como el ratón y el cerdo. Su potencial como marcadores útiles

para los estudios en plantas fue rápidamente reconocido, resultando en su aislamiento y aplicación en muchas especies (Rossetto et al., 1999). Es así, que los principales cultivos en la actualidad cuentan con una base de datos con apreciable cantidad de microsatélites como es el caso de la papa (Milbourne et al., 1998), yuca (Mba et al., 2001), arroz (Cho et al., 2000), trigo (Bryan et al., 1997), entre otros.

En un individuo diploide, el cromosoma A contiene el alelo (ACC)₈, mientras que el cromosoma A' en el locus homólogo presenta el alelo (ACC)₆. Mediante el PCR, utilizando iniciadores diseñados en las regiones flanqueantes, los alelos son amplificados, para luego ser separados mediante una corrida electroforética. Figura 3 muestra el patrón de bandas que deben determinar tanto los homocigotos como el heterocigoto para los mencionados alelos.

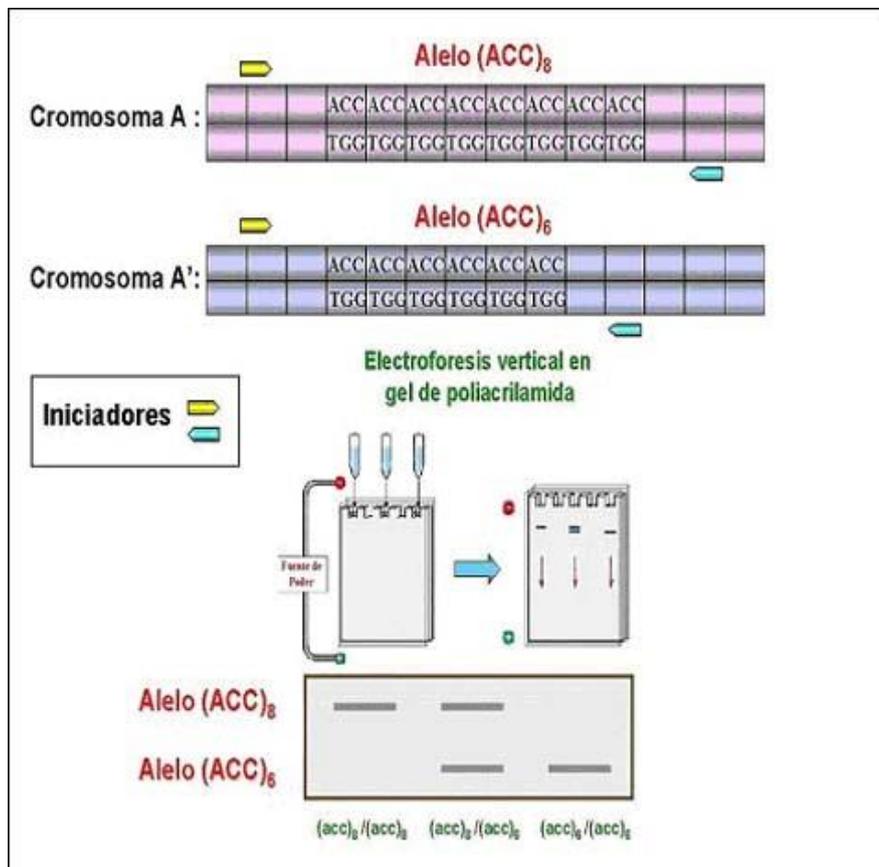


Figura 3: Diagrama que muestra el polimorfismo de los microsatélites debido a las diferencias en su longitud.

Los diferentes modelos que han intentado explicar el polimorfismo encontrado en los microsatélites ordenados de manera cronológica son: el modelo de alelos infinitos o IAM (Kimura y Crow, 1964), modelo mutacional escalonado o SMM (Ohta y Kimura, 1973) y el

modelo mutacional de dos fases o TPM (Di Rienzo et al., 1994). Sin embargo, estos modelos no han podido explicar toda la variación y distribución de los microsatélites existentes en el genoma por lo que se plantea la necesidad de buscar otros modelos que expliquen mejor estos fenómenos (Tóth et al., 2000)

2.16. CLASIFICACIÓN DE LOS MICROSATÉLITES

De acuerdo con la secuencia de los microsatélites (Tabla 4), Jarne y Lagoda (1996) clasificaron los SSR en tres tipos: puros (p.e. (CA)_n), compuestos (p.e.(C.A.n(GA)_n) e interrumpidos (p.e. (CA)_nTT(GA)_n). Estas regiones son muy comunes en genomas eucariotas, se encuentran repartidas a lo largo de todo el genoma y presentan un alto nivel de polimorfismo por lo que se emplean comúnmente como marcadores genéticos (Ferreira y Gratapaglia, 1998). Chambers y MacAvoy (2000) proponen que los términos puro (o perfecto), compuesto y complejo se utilice para especificar que uno, dos o más tipos de motivos, respectivamente.

Por otro lado, se ha demostrado que las repeticiones imperfectas son menos polimórficas (Jin et al., 1996). Las interrupciones de las secuencias repetitivas dentro de un microsatélite pueden estabilizar estas secuencias ya que pueden crear puntos de apareamiento durante la replicación, y así un mal apareamiento podría detectarse más fácilmente por las enzimas de reparación (Heale y Petes, 1995).

Tabla 4: Clasificación de los microsatélites.

| Tipo de Microsatélite | Ejemplos |
|------------------------------|--|
| SSR Puro | (ACC) ₉ |
| SSR Interrumpido puro | (ACC) ₆ —TG—(ACC) ₇ |
| SSR Compuesto | (ACC) ₅ —(TTG) ₉ |
| SSR Interrumpido compuesto | (ACC) ₈ —TG—(ACC) ₅ —GA—(TTA) ₆ |
| SSR Complejo | (ACC) ₈ —TG—(GA) ₁₂ —(TTA) ₅ —GC—(TTA) ₄₊₃ |

FUENTE: Elaboración propia.

Ventajas del uso de microsatélites

Entre las ventajas de esta técnica podemos señalar las siguientes:

- Multialélicos, transmisión codominante, facilidad de detección por PCR, relativa abundancia, cobertura extensiva del genoma y requerimiento de solo una pequeña cantidad de ADN inicial (Powell, 1996b; Milbourne et al., 1998).
- Alto polimorfismo en comparación con otras técnicas como RFLP, RAPD y AFLP (Powell et al., 1996).
- Otorgan información intermedia entre la secuencia completa y las meras diferencias electroforéticas ya que conocemos el motivo repetido y podemos estimar el número de repeticiones presentes (Scotti, 2000).
- Son considerados neutrales por lo que la selección y la presión del ambiente no influyen directamente en su expresión (Scotti, 2000).
- Los productos de PCR de diferentes loci se pueden correr en mismo gel (PCR multiplex) (Scotti, 2000).
- Si el método de corrida es el mismo, los resultados pueden ser fácilmente intercambiados y comparados con otros laboratorios (Scotti, 2000).
- Altamente informativos: presentan herencia codominante y muchos alelos son encontrados entre individuos estrechamente relacionados.
- Ser técnicamente simples: la tecnología del PCR puede fácilmente y rápidamente ser utilizada para la automatización en el uso de estos marcadores.
- Ser analíticamente simples: los datos son producidos de forma confiable y altamente reproducibles.
- Ser muy abundantes: los microsatélites están uniformemente dispersos a través del genoma, aproximadamente cada 10 Kbp.
- Ser ampliamente aplicables: los loci son frecuentemente conservados entre especies relacionadas y algunas veces entre géneros, y
- Ser de fácil intercambio de datos entre laboratorios: la información puede ser comunicada por la simple secuencia de los iniciadores, sin la necesidad de la transferencia física de los iniciadores.

Desventajas del uso de los microsatélites

Entre las desventajas de esta técnica podemos señalar las siguientes:

- Los productos de amplificación frecuentemente consisten de varios fragmentos de diferentes tamaños (Oda et al., 1997), llamadas “bandas tartamudas” (Walsh et al., 1996). La enzima Taq polimerasa así como otras polimerasas es susceptible al salto o “slippage”. Adicionalmente posee una actividad desoxinucleotidil trnasferasa terminal por lo que adiciona un nucleótido extra (adenina) en el terminal 3´ del producto de PCR, de manera dependiente de la secuencia (Oda et al., 1997).
- La electroforesis en geles de poliacrilamida es susceptible a errores de migración (Oda et al., 1997). No se puede conocer el número de alelos que hay por cada tamaño de banda a menos que se maximice el número de alelos diferentes que puede tener de acuerdo a su ploidía (Oda et al., 1997).
- Homoplasia, por este fenómeno aun cuando los tamaños de los alelos de un mismo locu sean idénticos, su estructura (secuencia) puede no serlo (Chambers y Mac Voy, 2000). El impacto que la homoplasia podría tener es la sobreestimación de la similitud (Chambers y Mac Voy, 2000). Los microsatélites producen información generalmente concordante con otros tipos de marcadores, lo cual sugiere que la homoplasia no oscurece la diferenciación de las frecuencias alélicas (Blankenship, 2002).
- La cercanía de un gen sujeto a selección puede influir grandemente en la frecuencia en que se encuentra un alelo microsatélite. Un microsatélite que estuvo ligado inicialmente a un gen con una característica ventajosa permanecería o incluso podría fijarse en la población (Oda et al., 1997).
- Los alelos nulos se pueden originar por la ausencia del microsatélite o por mutaciones en las secuencias flanqueantes en las cuales hibridizan los iniciadores, lo que resulta en la seudo muerte del alelo (Jarne y Lagoda, 1996; Chambers y Mac Avoy, 2000). Sólo es posible observar un alelo aún cuando existan más copias por lo que un supuesto individuo homocigoto podría traerse de un heterocigoto siendo el otro alelo un alelo nulo.
- Se pueden obtener patrones complejos de bandas con ciertos pares de iniciadores, debido a la amplificación de múltiples loci microsatélite usando un único par de iniciadores (Provan et al., 1996).
- La inversión inicial de recursos económicos y la experiencia técnica requerida para el clonamiento y secuenciamiento de los loci SSR.

2.17. APLICACIONES DE LOS MICROSATÉLITES

Los microsatélites se han convertido en una herramienta poderosa por su amplio polimorfismo y naturales codominante; ambas características han permitido el estudio de las poblaciones naturales con un alto grado de resolución. Actualmente estas regiones son ampliamente usadas en el estudio de genética humana, de mamíferos, y de especies vegetales (Jarne y Lagoda, 1996; Powell et al., 1996) e identificación de genotipos (Brondani et al., 1998).

En genética de poblaciones, por las asociaciones con grandes fragmentos cromosómicos cuyas frecuencias corresponden más directamente a factores demográficos que gobiernan la diferenciación de las poblaciones (Blankenship, 2002). En genética humana, las repeticiones de trinucleotidos han sido asociadas a enfermedades humanas (Mitas, 1997).

En la actualidad se han realizado gran cantidad de trabajos aplicando las ventajas de los microsatélites aún antes de ser establecida la correlación entre uno o más SSRs y los loci que controlen características de interés agronómico, la disponibilidad de una cantidad de microsatélites, por si sola, ya permite emprender inmediatamente una serie de estudios importantes para los programas de mejoramiento genético. Así tenemos que las aplicaciones acorto plazo incluyen, básicamente, la identificación y discriminación de individuos (fingerprints), información sumamente necesaria en estudios de identificación de los parentales de algunas poblaciones, control de cruzamientos, estudios de diversidad y distancia genética y, últimamente, en la identificación y protección de variedades patentadas.

Testolin et al. (2000) desarrollaron 26 iniciadores SSRs y los utilizaron para confirmar el parentesco y paternidad de diferentes cultivares de melocotón (*Prunus persica* L.). Debido a la consistencia de los resultados, recomiendan el establecimiento de esta técnica para la identificación de individuos y para discernir el origen genético de cultivares de melocotón. Asimismo, Prasad et al. (2000) utilizaron un total de 20 marcadores microsatélites para probarlos en 55 genotipos elite de trigo con el objetivo de determinar la diversidad genética entre ellos y la utilidad de los iniciadores SSRs para la identificación de genotipos. Los genotipos provenientes de 29 países, representando a todos los continentes, mostraron valores de los coeficientes de similitud genética (GS), para los 1,485 pares de genotipos posibles, de 0.05 a 0.88 con un promedio de 0.23, sugiriendo un alto nivel de diversidad

entre los genotipos. Con respecto a su poder para discriminar un genotipo entre los 55 presentes en este estudio, los resultados mostraron que un solo marcador era suficiente para discriminar 16 genotipos, 2 marcadores para discriminar 29 genotipos, 3 microsatélites discriminaron 36 genotipos, y 4 discriminaron 41 genotipos. Al final, con un total de 12 microsatélites se pudo discriminar un máximo de 48 de los 55 genotipos. Los resultados demuestran la utilidad de los microsatélites para la detección de polimorfismo, liderar la identificación de genotipos y para la estimación de la diversidad genética.

Ghislain et al. (2000) han caracterizado la colección de papa (*Solanum* spp.) conservada en el Centro Internacional de la Papa (CIP) con una variedad de marcadores moleculares para aminorar las redundancias genéticas e identificar accesiones importantes para la conservación y el mejoramiento del cultivo. Para estos propósitos más de 70 microsatélites han sido probados sobre una muestra genéticamente diversa del germoplasma de papa y fueron seleccionados 18 de los microsatélites más informativos de acuerdo a la calidad de sus productos de amplificación, su localización en el genoma y el nivel de polimorfismo que detectan. Este grupo de microsatélites está ubicuo en las siete especies de papas cultivadas y serán usadas rutinariamente como una referencia para cada nueva entrada al germoplasma.

Pensando en el mediano y largo plazo muchos trabajos que aplican los SSRs están dirigidos a los programas de selección asistida por marcadores (MAS, por sus siglas en inglés), que son, en pocas palabras, los tradicionales programas de mejoramiento genético ahora apoyados en las nuevas tecnologías disponibles, como es el caso de los marcadores moleculares. Estos programas buscan la transferencia de alelos deseables para crear o mejorar cultivares de élite, seleccionados sobre la base de su genotipo, por regiones específicas asociados a características cualitativas o cuantitativas. En el caso de las plantas, la mayoría de estudios están orientados al desarrollo de cultivos resistentes a enfermedades y de mejorar sus características agronómicas.

La selección asistida por marcadores (MAS) en su forma más simple puede ser aplicada a reemplazar la evaluación de una característica que es difícil o demanda muchos recursos su evaluación. Cuando un marcador es encontrado que cosegrega con un gen que determina una característica importante, puede ser más fácil y económico definir la presencia del marcador ligado al gen, que evaluar dicha característica. Como en el caso del gen *xa-5*, el cual confiere

resistencia al arroz contra la enfermedad de la mancha bacterial (*Xanthomonas oryzae* pv. *oryzae*), donde se han reportado tres marcadores microsátélites ligados a este gen (Blair y McCouch, 1997). En consecuencia, la aplicabilidad de los microsátélites, por ser marcadores basados en el PCR, resultaría en el ahorro de tiempo y recursos, en vez de realizar una evaluación, en invernadero o en campo, de la presencia de esta enfermedad que ataca a las hojas del arroz, del mismo modo a ayudar a desterrar el alelo recesivo del gene *xa-5* sin afectar a las otras características de importancia en este cultivo.

Pero a consecuencia que la gran mayoría de las características heredables de importancia económica resultan de la acción conjunta de varios genes (herencia poligénica o cuantitativa), donde el fenotipo resultante presenta una variación continua en vez de clases genotípicas discretas, se hace necesaria la búsqueda de loci controladores de características cuantitativas, también llamados QTLs (por su acrónimo en Inglés), que similar a un gen, un QTL indica una región en el genoma, la cual podría comprender uno o más genes funcionales. Investigadores en Brasil se han utilizado los marcadores SSR y RAPD para la identificación de QTLs asociados a la resistencia al nemátodo de los quistes (*Heterodera glycines*), el cual es económicamente el patógeno más dañino que afecta al cultivo de la soya (Schuster et al., 2001).

El tizón del trigo es causado por patógenos fúngicos del género *Fusarium* y es responsable de severas pérdidas en la producción y calidad de las cosechas, pero más preocupante es la contaminación de los granos con micotoxinas producidas por estos hongos. Buerstmayr et al. (2002) probaron diferentes tipos de marcadores moleculares, pero especialmente AFLPs y SSRs, en una población de 364 individuos descendientes del cruce de una línea resistente y otra susceptible al tizón. El QTL con efecto predominante fue detectado en el brazo corto del cromosoma 3B contribuyendo con el 60% de la variancia genotípica para la resistencia a *Fusarium*.

Buscando resistencia en campo contra el patógeno fúngico *Phytophthora infestans*, causante de la enfermedad del tizón tardío, responsable de la devastación de los cultivos de papa en Irlanda, resultando en la más grande hambruna del siglo XIX, Ghislain et al. (2001) probaron diversos marcadores moleculares, especialmente marcadores RFLPs y SSRs, en una población interespecífica derivada del cruce entre *Solanum phureja* x *Solanum tuberosum*.

Se construyó un mapa genético para cada parental. De esta forma, dos importantes QTLs fueron detectados en los cromosomas VII y XII como una contribución de ambos parentales, controlando el 16% y 43%, respectivamente, de la variación fenotípica (VF). Por parte del parental *S. phureja* un QTL fue detectado en el cromosoma XI (11% de la VF) y tres QTLs fueron detectados como contribución del parental *S. tuberosum*, en el cromosoma III (13% de la VF), en el cromosoma V (11% de la VF) y en el cromosoma VIII (11% de la VF).

2.18. ANÁLISIS ESTADÍSTICO

Taxonomía numérica

La taxonomía numérica ha sido definida como la evaluación numérica de la afinidad o similitud entre unidades taxonómicas y el agrupamiento de estas en taxones mediante operaciones matemáticas, basándose en el estado de sus caracteres (Sneath y Sokal, 1963). La taxonomía numérica se utiliza para eliminar las subjetividades en el proceso de asignar individuos o clases previamente definidos. Al usar procedimientos estadísticos, la clasificación de los individuos en los grupos se hace con objetividad (Sevilla y Holle, 2004).

Unidad taxonómica operativa (OTU)

Las siglas O.T.U son el acrónimo de Operational Taxonomic Unit. Dicho de una manera práctica son las unidades en estudio. El rango taxonómico de estas dependerá del objetivo del estudio si se quieren caracterizar todas las accesiones de una misma especie existente en un banco de germoplasma, con el fin de estudiar si existen duplicados, homónimos o sinónimos, cada una de las accesiones será una O.T.U. Si por el contrario lo que se quiere analizar es la variabilidad interpoblacional de una determinada especie, las OTUs serán cada una de dichas poblaciones. En este caso como lo que interesa analizar es la variedad intrapoblacional, la O.T.U sería el individuo. (González-Andrés, 2001).

Caracteres

El carácter puede definirse como cualquier propiedad que varía en las OTUs en estudio. Los posibles valores que ese carácter pueda presentar se los consideran sus estados (Sneath y Sokal, 1973), en la clasificación biológica cuando los caracteres diferenciales están gobernados por muchos genes; en términos moleculares eso significa que el ADN de los individuos separados en grupos distintos debe ser más diferente que el ADN de individuos

localizados en el mismo grupo. No basta que los individuos sean fenotípicamente distintos porque si el carácter está gobernado por un solo gen, la diferencia en el ADN de individuos diferentes es mínima. (Sevilla y Holle, 2004).

Codificación de los caracteres

Los caracteres taxonómicos forman parte del universo denominado “datos científicos” y responden a las exigencias de éste. El científico observa hechos y los registra en datos. Los datos son representaciones simbólicas de los eventos y/o estados y se obtienen por la observación (Kneller, 1978). Según Kneller (1978) existen distintas categorías de datos y se han propuesto numerosas clasificaciones para expresar esa variabilidad.

Matriz básica de datos

La matriz básica de datos (MBD) es una matriz rectangular, colocando en filas los caracteres y en columnas las OTUs analizadas. Por tanto la matriz recoge el valor de cada carácter en cada una de las OTU. En el caso de que exista variabilidad intra-OTU se tomará como valor representativo de cada carácter en cada OTU (González-Andrés, 2001).

Coefficiente de similitud

El parecido o similitud es cuantificable aplicando un coeficiente de similitud. Con el uso de los coeficientes en operaciones matemáticas pueden calcularse las similitudes (distancia) respecto a cada par posible de OTU de una matriz básica de datos (Crisci y López. 1983). Existen diferentes tipos de coeficiente basados en fundamentos matemáticos.

Coefficientes de distancia

Se aplican sobre matrices básicas que presentan dato doble-estado o multiestados, o en las que poseen ambos tipos de datos (datos mixtos).

Coefficientes de correlación

Calcula la similitud midiendo la separación angular formada por las dos líneas que parten del origen de las coordenadas y pasan por las OTU j y k. Los coeficientes de correlación son funciones de esos ángulos. Coeficientes de asociación, el coeficiente de asociación mide la similitud y diferencias en los caracteres entre dos OTUs, cuando éstos tienen sólo dos estados para cada característica. Los estados son, presencia (codificado con 1) o ausencia (codificado

con 0) (Sevilla y Holle, 2004). Si comparamos un carácter de doble estado de dos OTUs j y k , se presentan 4 posibilidades que se observan a continuación:

1,1 (a) Ambos OTUs tienen el carácter comparado

1,0 (b) El carácter está presente en el primer OTU y ausente en el segundo.

0,1 (c) El carácter está ausente en el primer OTU y presente en el segundo.

0,0 (d) Ambos OTUs presentan ausencia del carácter comparado.

El coeficiente varía de 0 a 1, siendo el valor 1 el que expresa la máxima semejanza entre dos OTUs.

Elección del coeficiente de similitud

La elección del coeficiente de similitud dependerá del tipo de datos que constituyan la MBD. Generalmente los coeficientes de distancia y los coeficientes de correlación se usan para los datos multiestado y los coeficientes de asociación y de distancia se utilizan para los datos doble estado.

Entre los coeficientes de asociación se encuentran: Coeficiente de Roger y Tanimoto, Sokal y Sneath, Hamann, Yule, Coeficiente (ϕ) de Pearson, Russel y Rao, Kulczynski, Ochiai (Crisci y López, 1983; Rohlf, 1997).

Los coeficientes de similitud más utilizados para datos binarios tenemos: el Coeficiente de similitud de Dice y Sorensen; el Coeficiente de similitud de Jaccard (J) y el coeficiente de similitud de Sokal y Michener o Coeficiente de Concordancia simple “Simple Matching” (SM).

Coeficiente de similitud de Dice y Sorensen

Da mayor importancia a las presencias dobles; es la proporción de caracteres presentes respecto al promedio de los caracteres presentes en cada OTU

$$S_D = \frac{2a}{2a + b + c} = \frac{a}{(a + b) + (a + c)}$$

Da mayor importancia a las dobles presencias. Además, es la proporción de caracteres copresentes respecto al promedio de los caracteres presentes en cada OTU.

Coefficiente de similitud de Jaccard (Sneath)

No considera las dobles ausencias; respecto al total de caracteres presentes en las OTUs. Puede usarse con datos de marcadores dominantes (RAPD y AFLP), por cuanto las ausencias podrían corresponder a recesivos homocigóticos.

$$S_J = \frac{a}{a+b+c}$$

No considera las dobles ausencias. Es la proporción de caracteres presentes que comparten, respecto al total de caracteres presentes en las 2 OTUs. Además, puede usarse con datos de marcadores dominantes (RAPD y AFLP), por cuanto las ausencias podrían corresponder a recesivos homocigotos.

Coefficiente de similitud de Simple Matching

$$S_{SM} = \frac{a+d}{a+b+c+d}$$

Simplemente es la proporción de caracteres que comparten las dos muestras. (Winzer et al., 2004). Para considerar algunos criterios en la elección de los coeficientes de similitud se tiene que tener en cuenta la concepción del mismo.

Según Winzer et al. (2004), dos muestras son más parecidas cuantos más “unos” compartan; la coincidencia de “ceros” no aporta a la similitud, es decir, no considera dobles ausencias para no sobrestimar la similitud a partir de características que no están presentes en ninguno de los dos individuos.

Matriz de similitud

Con los valores de similitud obtenidos de cualquiera de los coeficientes de similitud para los pares posibles de OTUs ordenados en forma tabular se constituye la matriz de similitud (Crisci y López, 1983). La cual es de tipo diagonal y recoge el valor de similitud entre OTUs. En ocasiones en vez de matrices de similitud se calculan matrices de distancias entre todos los pares posibles de OTU (González-Andrés, 2001).

Análisis de agrupamientos

El análisis de agrupamientos comprende técnicas que siguiendo reglas más o menos arbitrarias, forman grupos de OTU que se asocian por su grado de similitud. Las más utilizadas son las del siguiente tipo: exclusivas, jerárquicas, aglomerativas y secuenciales (SAHN: Sequential, Agglomerative, Hierarchical and Nested). El significado de cada uno de los términos es el siguiente:

Técnicas exclusivas: Originan grupos, donde las OTU son exclusivas del grupo del cual forman parte y no pueden pertenecer a otro grupo que se halle en un mismo rango o nivel.

Técnicas jerárquicas: Originan conjuntos que presentan rangos, en los cuales las OTU o grupos de OTU subsidiarios forman parte de un grupo mayor o inclusivo.

Técnicas aglomerativas: Son las que partiendo en “n” OTU separadas, las agrupa en sucesivos conjuntos para llegar finalmente a un solo conjunto que contiene a la “n” unidades.

Técnicas secuenciales: Se forma un grupo de cada vez, hasta que se agota el conjunto total.

El procedimiento a seguir en un análisis de agrupamientos es el siguiente:

1. Examen de la matriz de similitud para localizar el valor de similitud más alto, lo que conduce a la formación del primer núcleo (conjunto formado por dos OTUs).
2. Se busca el próximo valor de mayor similitud, lo que puede suponer: (i) formación de nuevos núcleos, (ii) incorporación de una nueva OTU a un núcleo, con lo que se originaría un grupo (conjunto formado por más de dos OTU), o (iii) fusión de núcleos existentes.
3. Se repite la segunda etapa hasta agotar todas las OTU. El resultado final es la obtención de un dendograma.

Existen métodos para el cálculo de la similitud entre la OTU candidato a incorporarse y el grupo o núcleo al que se incorpora. La más utilizada es el UPGMA (Unweighted Pair-Group Method Using Arithmetic Averages) “Media aritmética no ponderada”. Según este método las OTU se incorporan a grupos o núcleos ya formados, teniendo en cuenta que el valor de similitud entre las OTU candidato a incorporarse y el grupo o núcleo es igual a una similitud promedio resultante de los valores de similitud entre el candidato y cada uno de los integrantes del grupo o núcleo. (González-Andrés, 2001).

Dendograma

La estructura taxonómica obtenida de la matriz de similitud con las técnicas de análisis de agrupamientos puede representarse gráficamente de varias formas, pero la más utilizada es

el dendograma. Este es un diagrama arborescente, en forma alterna que muestran una multibifurcación como una serie de bifurcaciones en el mismo nivel, que muestran la relación en grado de similitud entre dos OTUs o grupo de OTUs. Los valores de similitud se expresan en una escala que suele encontrarse en su extremo superior (Rohlf, 1997).

Básicamente un dendograma incluye a dos tipos de gráficos: fenogramas y cladogramas, el primero representa relaciones fenéticas, en cambio, en el segundo las ramificaciones están basadas en las conexiones filogenéticas entre las OTUs (Crisci y López, 1983).

Coefficiente de correlación cofenético (r)

Si se examina con atención las técnicas de construcción de un dendograma, se comprobará que es imposible que el dendograma sea un reflejo exacto de la matriz de similitud. Algunas de las relaciones de similitud serán necesariamente distorsionadas al realizar una representación bidimensional.

Se han propuesto varias técnicas para medir el grado en el que el dendograma representa los valores de la matriz de similitud. La técnica más conocida es la del coeficiente de correlación cofenética (r) establecido por Sokal y Rohlf (1962). El cual consiste en construir una nueva matriz de similitud a partir de los valores del dendograma y se denomina “matriz cofenética” (Crisci y López, 1983).

El coeficiente de correlación cofenética es una medida de la distorsión interna de la técnica, pero no va una evaluación de la clasificación que se construirá a partir del dendograma (Rohlf, 1970).

Los valores de la correlación cofenética pueden interpretarse de la siguiente manera:

| Rango | Interpretación |
|--------------------|-----------------------|
| $0.9 \leq r$ | Muy bueno |
| $0.8 \leq r < 0.9$ | Bueno |
| $0.7 \leq r < 0.8$ | Pobre |
| $r < 0.7$ | Muy pobre |

Valores bajos de esta correlación no necesariamente indican que el dendograma obtenido no sea de utilidad, sino estaría indicando algún tipo de distorsión al momento de obtener el dendograma (Rohlf, 1970). El método con el que se obtenga mayores valores de correlación cofenética sería el considerado como el más apropiado para los análisis e interpretación posteriores.

Análisis de contenido de índice polimórfico (PIC)

El PIC (Polymorphism Index Content) se utiliza para corregir la capacidad de cada marcador para revelar los loci polimórficos en el germoplasma. A mayor PIC, mayor información dará el marcador SSR.

Si los valores PIC de cada marcador son sumados entonces se podrá tener un índice de iniciador SSR (Ghislain et.al., 1999). El PIC es similar al valor de heterocigocidad y oscila entre 0 y 1. Este índice evalúa la informatividad de un marcador en la población de acuerdo a las frecuencias de los alelos. Para su cálculo se multiplica la probabilidad de cada posible cruzamiento (estimado a partir de las frecuencias alélicas) por la probabilidad que sean informativos, es decir, que se pueda identificar al progenitor del que procede el alelo (Botsein et al., 1980).

$$\begin{aligned} PIC &= 1 - \sum_{i=1}^n f_i^2 \\ &= 1 - p^2 - q^2 \\ &= 1 - f_{(1)}^2 - f_{(2)}^2 \end{aligned}$$

Donde

- fi : es la frecuencia del i-ésimo alelo.
- p : frecuencia de las bandas presentes
- q : frecuencia de las bandas ausentes

Índice de diversidad genética

La diversidad genética es un componente de la diversidad biológica y se refiere a la frecuencia y diversidad de los diferentes genes y/o genomas. Incluye las variaciones tanto dentro de una población como entre poblaciones. Según Sevilla y Holle (2004), la diversidad genética presenta dos niveles: el primero es el número de las categorías intra-específicas de la especie tales como: razas, ecotipos, variedades, morfotipos, entre otros. El segundo se refiere a la variación en loci específico de marcadores genéticos (Sevilla y Holle, 2004). De manera semejante Brush (1995) indica que la variabilidad se refiere a la homogeneidad de los alelos y genotipos con sus correspondientes fenotipos y morfotipos.

La diversidad genética es importante para mantener un reservorio de condiciones de respuesta al medio, que permita la adaptación y la supervivencia. El uso y manejo de la diversidad genética de los cultivos mantenida “in situ” ha ayudado a los campesinos a sobrevivir bajo las más difíciles condiciones por lo que es considerada la principal vía para la seguridad alimentaria de las generaciones futuras (Collind & Hawtin, 1999).

El conocimiento del nivel de diversidad es fundamental en el planeamiento de actividades de conservación ya que sin monitoreo no es posible verificar la efectividad de la conservación (Brush, 1999; citado por Negri y Tosti, 2002). Estudios empíricos en ecología y evolución frecuentemente dependen de una evaluación precisa de la diversidad genética para resolver cuestiones referidas a las relaciones entre individuos, la estructura poblacional, relaciones filogenéticas y mapeo de caracteres cuantitativos (Mueller y Wolfenbarger, 1999). Los estudios de diversidad genética basados en marcadores moleculares pueden brindar excelente información sobre la diferenciación genética y estructura genética. Además, pueden ser de gran utilidad para validar la clasificación taxonómica, validar la procedencia de las accesiones, identificar redundancias y mejorar el balance en la representación de las especies en los bancos de germoplasma (Van Hintum y Van Treuren, 2002). Adicionalmente, la información de la diversidad genética en material de mejoramiento es esencial para el diseño de programas de mejoramiento óptimos. Esto implica, por ejemplo, la selección de parentales para el establecimiento de poblaciones de cruzamiento, el reconocimiento de grupos heteróticos, y la introgresión de germoplasma exótico. Además, otro interés para los mejoradores es el registro y protección de cultivares (Muminovic et al., 2003).

La pérdida o reducción de la diversidad genética, se puede definir como erosión genética. Esta pérdida puede ser natural o causada por el hombre.

Existen varios índices creados para estimar la diversidad genética, el índice a usar en un trabajo determinado dependerá del tipo de marcador (dominante o codominante), del modelo mutacional que se asume y del tamaño de muestra. Entre los índices de diversidad más comunes tenemos al índice de diversidad de Shannon o Shannon & Weaver (1949), el índice de Simpson (1949) y el índice de diversidad o heterocigosidad esperada de Nei (1973). Este último es el que ha sido comúnmente empleado para los análisis de diversidad genética empleando marcadores microsatélites (Ghislain et al., 2000).

Análisis de factorial multivariado

Es un método de interdependencia de variables en el que la variabilidad de las variables observadas se explica en términos de unas pocas variables no observadas llamadas factores. Se usa con carácter exploratorio para identificar factores sin restricciones o para confirmar hipótesis previas formuladas por razonamiento teórico y/o experiencias similares. Se pueden calcular las correlaciones entre variables (análisis R) o entre individuos (análisis Q). Existen dos tipos de análisis factorial: el de componentes principales (PCA) y el de factores comunes (CFA).

Análisis de varianza molecular (AMOVA)

La estructura genética de una especie está relacionada con la distribución de la diversidad entre y dentro de las poblaciones. Su estudio se lleva a cabo normalmente mediante un análisis de varianza que permite comprobar la bondad de una estructura genética jerárquica propuesta *a priori*. La varianza molecular total es la suma de los componentes de la covarianza debido, por ejemplo, a las diferencias entre los individuos dentro de una población, a las diferencias entre las distintas poblaciones dentro de un grupo y a las diferencias entre distintos grupos de poblaciones.

El AMOVA proporciona parámetros de estructura genética en forma de estadísticos F que permiten el análisis de la estructura en poblaciones subdivididas para medir la distancia genética entre las subpoblaciones (Excoffier & Lischer, 2010). Wright aprovecha la consistencia de 3 coeficientes diferentes que estiman:

F_{ct} : La correlación de poblaciones entre grupos (comunidades entre Lambayeque, Amazonas y Pasco).

F_{sc} : La correlación de haplotipos entre poblaciones dentro de grupos (haplotipos entre comunidades dentro de Lambayeque, Amazonas y dentro de Pasco).

F_{st} : La correlación de haplotipos entre poblaciones entre grupos (haplotipos entre comunidades entre grupos).

El rango ($F_{st} = 0$), significa no existe divergencia genética, es decir las subpoblaciones tienen frecuencias alélicas idénticas. El F_{st} es una medida de diferenciación genética de las subpoblaciones y este es siempre positivo F_{sc} y F_{ct} son medidas de desviación de las proporciones de Hardy-Weinberg dentro de subpoblaciones y en el total de poblaciones,

respectivamente, donde los valores positivos indican una deficiencia de heterocigotos, y valores negativos indican un exceso de heterocigotos.

F es el índice de fijación “Fixation Indexes”, que es la probabilidad de que los dos alelos de un individuo sean los mismos. Su cálculo debe hacerse solo con marcadores codominantes. F_{st} es el índice de diferenciación génica entre las poblaciones (≤ 1), en función de las frecuencias alélicas (Hollingsworth & Ennos, 2004).

Cuando F_{st} es:

De 0 a 0.05 = La diferenciación genética es pequeña.

De 0.05 a 0.15 = La diferenciación genética es moderada

De 0.15 a 0.25 = La diferenciación genética es grande

> 0.25 = La diferenciación genética es muy grande

III. MATERIALES Y MÉTODOS

3.1. MATERIAL VEGETAL

El presente trabajo se llevó a cabo en los laboratorios de la unidad de Biología Molecular del Instituto de Biotecnología (IBT) de la Universidad Nacional Agraria la Molina ubicada en la ciudad de Lima.

Se consideraron 98 accesiones de loche en el presente estudio de investigación. Las hojas jóvenes de estas accesiones fueron colectadas aleatoriamente en estadio floral en las comunidades de Pacora, Illimo y Pitipo, Departamento de Lambayeque, Perú durante la campaña agrícola 2008; frutos de cinco accesiones de fueron adquiridos aleatoriamente en el Mercado Mayorista de Lima, y dos en el mercado de Lambayeque, Perú. Estos son llamados “loche de montaña”. Además, el fruto de una accesión de Puerto Bermúdez, Pasco fue incluido en esta investigación. Por otro lado, los frutos de diez accesiones cultivadas en Lima e Ica fueron colectados al azar (cinco *C. maxima* y cinco *C. pepo*) fueron incluidos en esta investigación para determinar la relación genética con el loche. En total se analizaron 98 accesiones de *Cucurbita moschata*, 5 de *C. maxima* y 5 de *C. pepo*, haciendo un total de 108 accesiones. Los datos pasaporte de las accesiones utilizadas se muestran en el Anexo 1.

3.2. MATERIAL PARA LA COLECCIÓN DE LOCHE

- Bolsas plásticas
- Cooler
- Guantes de goma desechables
- Papel toalla
- Recipientes de plástico con tapa hermética de capacidad de 10 litros
- Cintas para rótulos
- Libreta de campo
- Lápiz
- Altímetro
- GPS
- Bolsas de plástico
- Pizeta
- Etiquetas de colgar

- Plumones indelebles
- Prensa de herbario
- Papel periódico.

3.3. MATERIAL PARA LA EXTRACCIÓN DE ADN

Reactivos Químicos

- Ácido acético glacial
- Agua destilada
- Alcohol isoamilico (Merck)
- Bromuro de hexadeciltrimetil amonio (CTAB), (Sigma)
- B – mercaptoetanol, (Sigma)
- Cloroformo (Merck)
- Cloruro de sodio (Merck)
- EDTA (ácido etilen diamino tetra-acético (Sigma)
- Etanol absoluto 70° y 96° (Merck)
- Isopropanol (Merck)
- Nitrógeno líquido
- RNAase A, (Concert TM Invitrogen Life Technologies)
- SDS (Sodium dodecyl sulfate)
- PVP (Tris[hidroxymetyl] amino ethan), (Merck)

Materiales de laboratorio

- Agitadores magnéticos
- Espátulas, pinzas, tijeras
- Contenedor de Nitrógeno Líquido
- Gradillas y soportes para tubos
- Guantes de vinilo (Guardian TM)
- Juego completo de pipetores (BIO-RAD)
- Magnetos
- Matraces de Erlenmeyer (Pirex®)
- Morteros de porcelana con su respectivo pistilo
- Puntas plásticas de 10, 200 y 1000ul de capacidad (Axygen)

- Probeta de 10, 100, 1000 ml. (Pirex ®)
- Racks polypropylene (Axygen)
- Tubos para microcentrífuga de 0.6, 1.7 y 2 ml. (Eppendorf), (Axygen)
- Vasos de precipitación (Pirex ®)

Equipos

- Autoclave, (Market Forge, modelo Sterlimatic)
- Balanza analítica (Sartorius, modelo CP323S)
- Bomba de vacío
- Cámara de baño maría (Cimatec S.A.)
- Cámara de flujo laminar
- Centrífuga refrigerada (4°C) (Mikro 22R Hettich Zentrifugen)
- Congelador a -20°C (Electrolux)
- Congelador a -70°C (Electrolux)
- Estufa (Lab-line Instruments. Inc) modelo 3511-1
- Horno microondas (Electrolux)
- Potenciómetro (Hanna instruments pH 211)
- Refrigeradora (Electrolux)
- Vórtex (Scientific Industries Vórtex - Genie)

3.4. MATERIAL PARA LA ELECTROFORESIS

Reactivos Químicos

- Ácido acético glacial (Merck)
- Acrilamida (BIO-RAD 99.9%)
- Agarosa (Gibco BRL. Ultra Pure)
- Agua destilada
- Azul de bromofenol (Sigma)
- Bind Silano adherente, Silane A 174 (Merck)
- Bromuro de etidio (Sigma)
- Bisacrilamida (N,N` - Methyllene – bis - acrylamide) (Sigma)
- Carbonato de sodio (Merck)
- Glicerol

- Formaldehído al 37% (Sigma)
- Formamida (Applichem)
- Hidróxido de Sodio
- Marcador de peso molecular (Gene Ruler™ 100 bp DNA Ladder, Fermentas Life Sciences).
- Marcador de peso molecular, ADN del fago Lambda λ PST I (Gibco)
- Nitrato de plata (Merck)
- Persulfato de Amonio (Sigma)
- Revelador y fijador de fotografía (Kodak GXB, 5x)
- Temed (N,N,N` N` - Tetra – metil etilen di-amina) (Sigma)
- Tiosulfato de sodio (Merck)
- Sigmacote (Sigma - Aldrich)
- Úrea (Merck)
- Xilén cianol (Sigma)

Materiales de laboratorio

- Cajas de dilución de 96 pocillos, 1ml de capacidad c/u (Axygen Scientific)
- Gradillas y soportes para tubos
- Placas de vidrio para electroforesis vertical
- Papel tissue (Kimwipel®)
- Película fotográfica (Kodak APC film)
- Racks polypropylene (Axygen)
- Termómetro (Fisherbrand®)
- Vasos de precipitación (Pyrex)

Equipos

- Cámara de video, tarjeta capturadora de imágenes digital (Digital Color CCD AVC 591) y computadora
- Cámaras de electroforesis horizontal (EC Minicell y BIO-RAD SubCell - 192)
- Cámara de electroforesis vertical, Sequi - Gen® GT Nucleic Acid Electrophoresis Cell (Bio-Rad)
- Cámara digital (AVC 591 Digital Color CCD)

- Expositor de luz UV Transiluminador 2000 (BIO - RAD)
- Escáner (Epson Perfection 1260)
- Fuente de poder (BIO - RAD)
- Horno microondas (Electrolux)
- Impresora Láser (Xerox Phaser 3120)
- Purificador de agua (Milipore, Simplicity 185)
- Refrigeradora (Electrolux)
- Shaker (Modelo C10 Platform, New Brunswick Scientific)
- Bomba de vacío

3.5. MATERIAL PARA PCR

Reactivos Químicos

- Aceite mineral (Sigma)
- Agua mineral de nucleasas (Sigma)
- Alcohol 96%
- Cloruro de Magnesio 1 M solución (Sigma Lote 99H8925)
- 15 iniciadores microsátélites diseñados para *Cucurbita moschata* y 3 diseñados para *C. pepo*
- EDTA (Sigma)
- Enzima Taq ADN polimerasa y solución tampón de almacenamiento
- Solución tampón 10X PCR Buffer II (Invitrogen)

Material de laboratorio

- Micropipetas regulables de 0.5-10ul, 5-20ul, 100-1000ul (uso exclusivo para PCR), (BIO – RAD)
- Placas de policarbonato PCR de 96 pocillos (Axygen)
- Puntas para micropipeta 0.5-10ul, 20-100ul, 100-1000ul, (Axygen)
- Tapa de placas PCR de 96 pocillos (Axygen)
- Tubos en tiras y tapas en tiras para PCR (Axygen)
- Tubos Strips 0.2ml para PCR (Axygen)

Equipos

- Microcentrifuga para tubos de 1.7 ml, de un máximo hasta 10000 rpm (Denver Instruments Co.)
- Termociclador (GeneAmp® PCR System 9700)

Software

- NTSYS-pc versión 2.10 (Applied Biostatistics Inc., Setauket, Nueva York, EE.UU.)
- Arlequin 3.5.1.2 (Excoffier, 2005)
- Paint Shop Pro 5.01 (Jasc Corporate, MN, EE.UU.)

3.6. COLECCIÓN DE ACCESIONES

Existen aproximadamente cincuenta agricultores que se dedican a la producción del *loche* en los distritos de Illimo y Pacora, departamento de Lambayeque (Javier Moreano – alumno tesista del cultivo loche). Con el fin de capturar la mayor diversidad genética se tomaron 3-5 muestras de hojas apicales de cada variante morfológica de *loche* de todos los campos en producción en las zonas antes mencionadas, es decir, estas colectas se realizaron usando el criterio de selección de muestreo selectivo. Se colectaron hojas jóvenes, colocándose en bolsas plásticas transparentes donde se escribió el código de cada entrada para su identificación en laboratorio. Las muestras colectadas se colocaron en una caja tecnopor con el fin de evitar la deshidratación.

Además, se colectaron aleatoriamente frutos del loche denominado “loche de montaña” en el Mercado de Lambayeque y Mercado Mayorista de Lima. Por otro lado, los frutos de diez accesiones de *C. maxima* y *C. pepo* cultivados en Lima e Ica fueron colectados al azar y fueron incluidos en esta investigación para determinar la relación genética con el *loche*. El Prof. Roberto Ugás brindó una accesión de loche proveniente de Puerto de Bermúdez, Pasco. Cada muestra colectada contó con un dato pasaporte (nombre común de la variedad, lugar de colección, distrito, provincia, altitud, latitud, longitud (Anexo 1).

Las muestras de cada accesión de loche ideales para este tipo de estudio son frecuentemente las hojas más jóvenes. Al llegar al laboratorio de Biología Molecular del IBT, se almacenaron las muestras a 4°C hasta el momento de la extracción de ADN.

3.7. EXTRACCIÓN DE ADN

La extracción del ADN vegetal se realizó por el método de Doyle y Doyle (1990) con algunas modificaciones adicionales. Las modificaciones fueron usar menor concentración de etanol en los lavados (ambos a 70%) y menor tiempo de centrifugación. El protocolo de extracción y la preparación de los reactivos se detallan en el Anexo 2.

3.8. DETERMINACIÓN DE CALIDAD Y CONCENTRACIÓN DE ADN

La prueba de calidad y concentración de ADN de las muestras extraídas, fue llevada a cabo en geles de agarosa Merck 1.5%. Se usó el marcador λ /Pst-I (Genoma del bacteriófago λ digerido con la enzima Pst-I), adecuado de tal manera que la intensidad de la primera banda producida, correspondió a la concentración de 280 ng/ μ l, cuya banda de mayor peso molecular es de 14,800 pb. La determinación de la calidad del ADN se definió por observación visual identificando que las bandas sean claras, bien definidas y que la corrida electroforética no presente ningún barrido *smear*. Si la muestra presenta el *smear* significa que la muestra de ADN estaría presentando indicios de degradación. De esta manera, fue posible calcular la cantidad de buffer TE y de ADN stock requerido para preparar 300 μ l de dilución, a una concentración de 6 ng/ μ l. Para el análisis poblacional, la dilución para el análisis fue preparada en una placa de dilución. Tanto las diluciones, como la solución de ADN stock, fueron almacenadas a -20 ° C.

3.9. MARCADORES MICROSATÉLITES (SSRs)

Se usaron 21 pares de iniciadores microsatélites (18 para *Cucurbita moschata* y 3 para *C. pepo*) diseñados por Gong et al. (2008) (Tabla 5) siguiendo el protocolo (tabla 6) y el programa (tabla 7) de PCR según el manual de Protocolos de Biología Molecular del Instituto de Biotecnología (IBT). Los programas de amplificación varían con respecto a la temperatura de anclaje de los iniciadores de cada locus SSR.

Tabla 5: Secuencia de los 21 iniciadores microsatélites que se usaron para el estudio de la variabilidad genética del loche.

| Nombre | Alelos | Motivo | No. de repeticiones | Secuencia | Tamaño esperado(bp) | Temp. de anclaje (°C) |
|-------------|--------|--------|---------------------|--|---------------------|-----------------------|
| CMTp20 6 | 6 | ttc | 11 | Forward: CCACCGATT CGAAATTCA G Reverse: TTCCGAAAG AAACCGTTA CC | 114 | 59 |
| CMTp21 1 | 4 | gaa | 13 | Forward: TTGAATGGG TTGAGTTTA GG Reverse: ATGGTGAGA AGATTTGCA CT | 142 | 55 |
| CMTp24 5 | 5 | gcg | 9 | Forward:AGGT AACTGCACC CCAGCTA Reverse: CCCATCTCTC AAGCTCATC C | 134 | 60 |
| CMTm9 | 4 | aag+at | 11+4 | Forward: GCCCAGAAG ACAAAAGTT CG Reverse: TTTTTGTGTG CGTGTGTGG | 158 | 60 |
| CMTm3 9 | 5 | tttc | 6 | Forward:ATCT TTGTGCTTCT AACCTTTTTC Reverse:TAAA GCTCGAAAA AGCAACC | 153 | 57 |
| CMTm4 2 | 4 | gtc | 7 | Forward:ATTG GTGCCGAAG CTATCAC Reverse:CCCA CGTTATGGA GCAGAAT | 211 | 60 |

| | | | | | | |
|-------------|---|--------------|----|--|-----|----|
| CMTm8 6 | 4 | ttc | 8 | Forward:CCGC TCTCACCTCT TCTATG Reverse:AGAG GTGGGAAGA GAGAAGG Forward:AATA CGACGGCGT TTCACCTC Reverse:CCTA CAGCCGTGG ATGCT Forward:ACTT GCGGACTTT CACACCT Reverse:TGAA ATGAACGCA GAATTGC Forward:CCCA GCTATGAAG GTCTCTG Reverse:AAAA GCACAAAAG AAACACCA Forward:TTCC AGACCCAAT TCCCTTA Reverse:TCCA TGGAGATGG GAGGTAG Forward:CAGA CGTCGAACG TTGTAC Reverse:CTGG AGAGGAGCA TTTGGAG Forward:TCCA AATTTCAGG GAAGTGAGT Reverse:CCGA GATGAAGAA GGACGAG Forward:GACA GGACAGGTC AACACCTC Reverse:AACC CAATTGCAC AGCTTCT | 122 | 57 |
| CMTm1 12 | 3 | tagaga | 4 | Forward:CCGC TCTCACCTCT TCTATG Reverse:AGAG GTGGGAAGA GAGAAGG Forward:AATA CGACGGCGT TTCACCTC Reverse:CCTA CAGCCGTGG ATGCT Forward:ACTT GCGGACTTT CACACCT Reverse:TGAA ATGAACGCA GAATTGC Forward:CCCA GCTATGAAG GTCTCTG Reverse:AAAA GCACAAAAG AAACACCA Forward:TTCC AGACCCAAT TCCCTTA Reverse:TCCA TGGAGATGG GAGGTAG Forward:CAGA CGTCGAACG TTGTAC Reverse:CTGG AGAGGAGCA TTTGGAG Forward:TCCA AATTTCAGG GAAGTGAGT Reverse:CCGA GATGAAGAA GGACGAG Forward:GACA GGACAGGTC AACACCTC Reverse:AACC CAATTGCAC AGCTTCT | 107 | 60 |
| CMTm1 55 | 4 | agagaga c | 4 | Forward:CCGC TCTCACCTCT TCTATG Reverse:AGAG GTGGGAAGA GAGAAGG Forward:AATA CGACGGCGT TTCACCTC Reverse:CCTA CAGCCGTGG ATGCT Forward:ACTT GCGGACTTT CACACCT Reverse:TGAA ATGAACGCA GAATTGC Forward:CCCA GCTATGAAG GTCTCTG Reverse:AAAA GCACAAAAG AAACACCA Forward:TTCC AGACCCAAT TCCCTTA Reverse:TCCA TGGAGATGG GAGGTAG Forward:CAGA CGTCGAACG TTGTAC Reverse:CTGG AGAGGAGCA TTTGGAG Forward:TCCA AATTTCAGG GAAGTGAGT Reverse:CCGA GATGAAGAA GGACGAG Forward:GACA GGACAGGTC AACACCTC Reverse:AACC CAATTGCAC AGCTTCT | 140 | 60 |
| CMTm1 58 | 5 | ttc | 11 | Forward:CCGC TCTCACCTCT TCTATG Reverse:AGAG GTGGGAAGA GAGAAGG Forward:AATA CGACGGCGT TTCACCTC Reverse:CCTA CAGCCGTGG ATGCT Forward:ACTT GCGGACTTT CACACCT Reverse:TGAA ATGAACGCA GAATTGC Forward:CCCA GCTATGAAG GTCTCTG Reverse:AAAA GCACAAAAG AAACACCA Forward:TTCC AGACCCAAT TCCCTTA Reverse:TCCA TGGAGATGG GAGGTAG Forward:CAGA CGTCGAACG TTGTAC Reverse:CTGG AGAGGAGCA TTTGGAG Forward:TCCA AATTTCAGG GAAGTGAGT Reverse:CCGA GATGAAGAA GGACGAG Forward:GACA GGACAGGTC AACACCTC Reverse:AACC CAATTGCAC AGCTTCT | 80 | 57 |
| CMTm1 65 | 4 | ctaa | 5 | Forward:CCGC TCTCACCTCT TCTATG Reverse:AGAG GTGGGAAGA GAGAAGG Forward:AATA CGACGGCGT TTCACCTC Reverse:CCTA CAGCCGTGG ATGCT Forward:ACTT GCGGACTTT CACACCT Reverse:TGAA ATGAACGCA GAATTGC Forward:CCCA GCTATGAAG GTCTCTG Reverse:AAAA GCACAAAAG AAACACCA Forward:TTCC AGACCCAAT TCCCTTA Reverse:TCCA TGGAGATGG GAGGTAG Forward:CAGA CGTCGAACG TTGTAC Reverse:CTGG AGAGGAGCA TTTGGAG Forward:TCCA AATTTCAGG GAAGTGAGT Reverse:CCGA GATGAAGAA GGACGAG Forward:GACA GGACAGGTC AACACCTC Reverse:AACC CAATTGCAC AGCTTCT | 126 | 59 |
| CMTm1 77 | 3 | ttc | 13 | Forward:CCGC TCTCACCTCT TCTATG Reverse:AGAG GTGGGAAGA GAGAAGG Forward:AATA CGACGGCGT TTCACCTC Reverse:CCTA CAGCCGTGG ATGCT Forward:ACTT GCGGACTTT CACACCT Reverse:TGAA ATGAACGCA GAATTGC Forward:CCCA GCTATGAAG GTCTCTG Reverse:AAAA GCACAAAAG AAACACCA Forward:TTCC AGACCCAAT TCCCTTA Reverse:TCCA TGGAGATGG GAGGTAG Forward:CAGA CGTCGAACG TTGTAC Reverse:CTGG AGAGGAGCA TTTGGAG Forward:TCCA AATTTCAGG GAAGTGAGT Reverse:CCGA GATGAAGAA GGACGAG Forward:GACA GGACAGGTC AACACCTC Reverse:AACC CAATTGCAC AGCTTCT | 101 | 60 |
| CMTm1 86 | 4 | gccaag | 4 | Forward:CCGC TCTCACCTCT TCTATG Reverse:AGAG GTGGGAAGA GAGAAGG Forward:AATA CGACGGCGT TTCACCTC Reverse:CCTA CAGCCGTGG ATGCT Forward:ACTT GCGGACTTT CACACCT Reverse:TGAA ATGAACGCA GAATTGC Forward:CCCA GCTATGAAG GTCTCTG Reverse:AAAA GCACAAAAG AAACACCA Forward:TTCC AGACCCAAT TCCCTTA Reverse:TCCA TGGAGATGG GAGGTAG Forward:CAGA CGTCGAACG TTGTAC Reverse:CTGG AGAGGAGCA TTTGGAG Forward:TCCA AATTTCAGG GAAGTGAGT Reverse:CCGA GATGAAGAA GGACGAG Forward:GACA GGACAGGTC AACACCTC Reverse:AACC CAATTGCAC AGCTTCT | 146 | 60 |
| CMTm2 43 | 4 | ttc | 9 | Forward:CCGC TCTCACCTCT TCTATG Reverse:AGAG GTGGGAAGA GAGAAGG Forward:AATA CGACGGCGT TTCACCTC Reverse:CCTA CAGCCGTGG ATGCT Forward:ACTT GCGGACTTT CACACCT Reverse:TGAA ATGAACGCA GAATTGC Forward:CCCA GCTATGAAG GTCTCTG Reverse:AAAA GCACAAAAG AAACACCA Forward:TTCC AGACCCAAT TCCCTTA Reverse:TCCA TGGAGATGG GAGGTAG Forward:CAGA CGTCGAACG TTGTAC Reverse:CTGG AGAGGAGCA TTTGGAG Forward:TCCA AATTTCAGG GAAGTGAGT Reverse:CCGA GATGAAGAA GGACGAG Forward:GACA GGACAGGTC AACACCTC Reverse:AACC CAATTGCAC AGCTTCT | 139 | 59 |

| | | | | | | |
|-------------|---|---------|----|--|-----|----|
| CMTmC 3 | 3 | atg | 11 | Forward:CATC AGGGGTA TCCCATT Reverse:ACTT CCGGAGGTC AGAAGAC Forward:TTGC AACATTTGA CTAAATTCC | 131 | 58 |
| CMTmC 11 | 3 | atg | 6 | A Reverse:CTCC TCTTCTGCCG ATCAGT Forward:ACCT TCCTCCTCCG | 127 | 59 |
| CMTmC 44 | 3 | ttg | 6 | TAATGT Reverse:GACG AAGCCAACA ACTACAAC Forward:AAAC GTTCTGCAA TCAAGGA | 112 | 57 |
| CMTmC 64 | 6 | tc | 14 | Reverse:CAAG CTTGTGTAC AATAGGATA AGG Forward:AAAG GCCTTGGAT TTTTCTCC | 94 | 58 |
| CMTm1 03 | 3 | cct | 8 | Reverse:CCTT TCATGGTGG TGGAAGA Forward:CAGG CTATTCGCA CCCTCTA | 198 | 61 |
| CMTmC 70 | 3 | cat | 6 | Reverse:CCTC ATGCATTTTG CGTGTA Forward:TGAA GCTTTCAGA AATGAATGT | 196 | 60 |
| CMTm2 5 | 3 | ttcttct | 5 | G Reverse:CTGA CGTCGCTAC TCATAGCA | 158 | 59 |

CMTp significa pepo, CMTm moschata microsatélites.

FUENTE: Elaboración propia.

Tabla 6: Protocolo de amplificación de marcadores SSR.

| Reactivos | 1 Reacción | 110Rx (Final) |
|-------------------------|-----------------------|--------------------------|
| Agua libre de nucleasas | 2.6 µl | 286 µl |
| PCR Buffer 10X | 1 µl | 110 µl |
| Desoxirribonucleótidos | | |
| Trifosfato (dNTPs) 5mM | 0.4 µl | 44 µl |
| Forward | 0.4 µl | 44 µl |
| Reverse | 0.4 µl | 44 µl |
| ADN | 5 µl | 550 µl |
| Taq | 0.2 µl | 22 µl |
| Total | 10 µl | 1100 µl |

FUENTE: Elaboración propia.

Tabla 7: Programa de amplificación de marcadores SSR.

| Temperatura (°C) | Tiempo (minutos) | Ciclos |
|-----------------------------|-----------------------------|---------------------|
| 94 | 5 | 1 |
| 94 | 1 | |
| Anclaje | 1 | 35 |
| 72 | 1 | |
| 72 | 5 | 5 |
| 4 | ∞ | Fin del programa |

FUENTE: Elaboración propia.

Electroforesis en geles de agarosa

a) Preparación de geles

Se prepararon geles de agarosa Merk 1.5%-2.0%; utilizándose buffer TBE 1x como solvente. La mezcla agarosa-buffer TBE fue calentada empleando un horno microondas (2 minutos), una suave agitación, y 1 minuto adicional.

El gel completamente disuelto y caliente se dejó enfriar a temperatura ambiente por un tiempo aproximado de 1 minuto, hasta aproximadamente 45°C, luego fue vertido al molde para formar geles de agarosa; colocándose los peines. La temperatura muy elevada de la agarosa puede ocasionar deformaciones tanto en el molde como en los peines, en vista de que estos son generalmente de acrílico.

El previo nivelado del molde es una cuestión que se incluye en caso de que este sea grande. Pero siempre debe considerarse que la mesa de trabajo no esté sujeta a vibraciones; por

ejemplo, a consecuencia del frotamiento de vidrios en la preparación de geles de poliacrilamida.

A los 30 minutos aproximadamente luego de la vertida, el gel está listo para ser utilizado (polimerizado). Si se prefiere llevar a cabo la electroforesis después, puede colocarse un poco de buffer TE (hasta sumergir al gel) en el mismo molde o colocar el gel en la cámara electroforética que contiene buffer TBE 1x. De esta manera, el gel puede conservarse por meses.

b) Carga de muestras en geles de agarosa

Las muestras de ADN obtenidas en la extracción y con el respectivo tratamiento con TE y RNAsa, fueron colocados en una placa de PCR para luego agregarle azul de bromofenol; hasta un volumen total de 10 μ l (9 μ l de azul de bromofenol y 1 μ l de ADN). La elevada concentración le da un peso a la solución y permite que caiga totalmente dentro de cada pocillo del gel de agarosa, evitando la ocurrencia de mezclas.

La carga de las muestras puede efectuarse a cabo sobre los pocillos, sin sumergir el gel en buffer TBE, como también una vez colocado en la cámara electroforética con buffer TBE. En el primer caso, resulta más sencillo cargar, pero hay que tener cuidado al sumergir el gel con las muestras en el buffer TBE de la cámara; por cuanto las muestras pueden levantarse y suspenderse en el buffer fuera de los pocillos. En el segundo caso, el cuidado debe estar en colocar cada muestra dentro de pocillos diferentes, evitando vibraciones en las manos.

c) Tinción y visualización de los geles de agarosa

Luego de llevarse a cabo la corrida en agarosa, no es posible la observación de bandas a simple vista, sólo se visualizan dos frentes de corrida que son los componentes del colorante azul de bromofenol (que sirven de referencia para estimar la distancia de migración). Se sumerge el gel en una bandeja con solución de bromuro de etidio (5 mg/ml); el cual es una sustancia tóxica de efecto neoplásico, por lo que necesariamente se tienen que usar guantes de látex para colocar el gel en una fuente que contiene la mencionada solución. Allí se deja reposar por un tiempo aproximado de 10 a 15 minutos. Seguidamente, el gel es transferido a otra bandeja que contiene agua destilada; para enjuagar el gel de agarosa. Luego puede colocarse el gel en el transiluminador UV.

Mediante la incidencia de la luz UV en el gel con las muestras corridas, se observaron los fragmentos de ADN como una sola banda muy cerca de los pocillos. Para observar esta

imagen, es muy importante protegerse los ojos de la incidencia directa de la luz UV, mediante unos lentes de acrílico y/o mediante el protector acrílico del transiluminador.

Esta imagen es capturada mediante una cámara de video conectada a la computadora, utilizando el programa Intervideo WINDVD V. 7.0.

Electroforesis en geles de poliacrilamida

a) Preparación de geles

Se preparó geles de poliacrilamida al 6% (60 g. acrilamida, 3 g. bis-acrilamida, 420 g. urea, 50 ml de TBE 10x; enrasado a 1L) usando un vidrio grande de 38.5 x 28.5 cm y un vidrio pequeño de 36 x 28.5 cm (4 mm de espesor), ambos vidrios fueron concienzudamente limpiados empleando etanol 96% y paños de papel toalla. Seguidamente, los vidrios grandes fueron tratados con repelente de lluvia usado para evitar la retención de las gotas de agua en el parabrisas de los automóviles (“rain repell”); asimismo, los vidrios pequeños fueron tratados con solución adherente (1.5 ml de etanol absoluto, 5 μ l de bind silane, 5 μ l de ácido acético). Luego de los respectivos tratamientos, se dejó por espacio de 30 minutos para permitir una adecuada fijación de las soluciones a los respectivos vidrios. Una vez fijada las soluciones respectivas, colocar los separadores sobre el vidrio grande, y luego colocar frente a frente las superficies tratadas con el repelente y adherente (forma de sándwich), en los extremos de los vidrios en forma de sándwich colocar clamps para evitar la filtración de la acrilamida. Luego preparamos en un vaso precipitado 50 ml de acrilamida, 333 μ l de APS (1g persulfato de amonio /10 ml agua milliQ), y 33 μ l de TEMED para un gel, de esta manera, verter en el sándwich formado por los vidrios y esperar la polimerización por un espacio de 30 -60 minutos. Luego colocar el peine hacia arriba entre los vidrios para formar un molde, y dejar por un espacio de 1 hora para la polimerización. Una vez polimerizado invertir el peine para formar los pocillos para la carga de la muestra, de esta manera, el gel está listo para ser colocado en la cámara electroforética para el precorrido por un espacio de 30 minutos.

b) Sustitución de la urea convencional por urea a granel

Debido al agotamiento y por consiguiente de la carencia de la urea en el laboratorio, se planteó reemplazar la urea grado molecular por una de grado técnico y a granel (industrial). El empleo de esta urea, en la misma cantidad que la convencional, produjo algunas diferencias: en la disolución de la urea, para preparar la solución de acrilamida, esta adquiere

un aspecto más cremoso que la convencional; sin embargo, mediante el calentamiento y la agitación del frasco de vidrio transparente (para visualizar grumos), es posible la disolución completa de los componentes de la solución de acrilamida. La solución de acrilamida fue almacenada a temperatura de 4°C.

La otra diferencia tiene que ver con el tiempo de polimerización. La cantidad de APS y TEMED utilizado por 50 ml de solución de acrilamida con la urea convencional (333 µl de APS y 33 µl de TEMED) son suficientes para polimerizar el gel de manera adecuada por un lapso de tiempo de 15 a 30 minutos ocurre la polimerización completa (sin que fluya la solución al colocar los vidrios en distintas posiciones).

c) Carga y corrida de los geles

Los geles de poliacrilamida fueron colocados en cámaras electroforéticas diseñadas en Alemania, modificadas y fabricadas en el laboratorio de Biología Molecular del Instituto de Biotecnología de la Universidad Nacional Agraria la Molina.. Al igual que otras cámaras, esta consta de dos recipientes donde se coloca el buffer. Para los experimentos realizados, se empleó buffer TBE 0.5x en la parte superior y buffer TBE 1x en la parte inferior a sugerencia de Luis Gutiérrez (técnico de laboratorio, comunicación personal).

Para colocar el gel en la cámara, esta tiene que estar de tal manera que la única comunicación entre el buffer superior y el inferior sea el gel de poliacrilamida, el cual se encuentra entre los dos vidrios. Esta condición requiere que los dos vidrios, que contienen al gel, estén sujetos fuertemente a la cámara superior.

Es muy recomendable que el buffer se coloque una vez se hayan sujetado los vidrios a la cámara superior. El vertido del buffer TBE 0.5x, hasta que los pocillos queden completamente sumergidos en él, produce burbujas al interior de los pocillos.

Se utilizó agujas hipodérmicas delgadas para eliminar las burbujas y retirar algún resto de poliacrilamida. Luego se hicieron las pre corridas a 1200 V por 30 minutos, luego de eliminar las burbujas que se generan tanto en los pocillos como en la parte inferior de los vidrios.

Nuevamente se hizo una eliminación de burbujas, de algún resto de poliacrilamida y también de la urea. La urea tiende a formar una fase viscosa dentro de los pocillos, el cual tiene que

ser expulsado por completo, justo antes de cargar las muestras. Mientras tanto, en la cámara inferior, en la línea de contacto entre el buffer y el gel, también se hizo la eliminación de burbujas empleando una aguja hipodérmica doblada.

Para la reacción de denaturación de los amplificadores por PCR, se añadió 5 μ l de buffer tampón de carga (haciendo un volumen final de 15 μ l); y se llevó a 94 °C en un termociclador Gene Amp®; los productos de PCR denaturados fueron colocados de manera inmediata en hielo. Los voltajes a los que se llevó a cabo las corridas fueron variables y de acuerdo a la disponibilidad de las cámaras y fuentes de poder; sin embargo, se hizo la parada cuando el frente de corrida azul (del tampón de carga) alcanzaba el borde inferior del vidrio aproximadamente a unos 5cm.

d) Tinción con nitrato de plata

Las reacciones de tinción fueron llevadas a cabo luego de separar los dos vidrios que contienen el gel. Así, el vidrio pequeño con el gel adherido fue sumergido por 30 minutos en solución fijadora (900 ml de agua destilada, 10 ml de ácido acético), luego se hicieron 2 lavados de 4 y 3 minutos respectivamente en agua destilada fría, para luego pasar a una solución de tinción y en oscuridad (1 g AgNO₃ , 900 ml de agua destilada) por 40 min en el primer uso, luego por 5 segundos de lavado en 1/2 litro de agua destilada fría (10°C), aproximadamente 10 minutos en solución reveladora (28 g de Na(CO₃)₂ en 900 ml de agua destilada con 900 μ l de formaldehído y 25 μ l de tiosulfato de sodio), luego se pasó a la solución fijadora por 5 minutos, y finalmente un muy buen enjuague en un chorro de agua de caño. El secado completo de los geles se da entre 2 a 3 días.

Escaneado y lectura (scoreado) de bandas amplificadas

Se leyeron las bandas con mejor resolución y se enumeraron consecutivamente empezando desde el primer individuo del primer carril. Una vez hecha la lectura de las bandas, cada gel fue escaneado para almacenar su patrón de bandas.

Los diferentes fragmentos se visualizaron y analizaron para realizar la lectura de las bandas amplificadas (scoreo), esto se realizó empleando el Software Paint Shop Pro 5.01 (Jasc Corporate, MN, U.S.A.).

3.10. ANÁLISIS ESTADÍSTICO

Registro de datos

Se creó un registro de datos binarios, conteniendo los escoreos, cada banda de ADN se consideró como un locus individual, con dos posibles alelos: presencia de la banda (1) y a la ausencia (0), para cada fragmento de interés generado por cada iniciador usado. Mientras que a las bandas de presencia dudosa o faltante se le asignaron un valor de (9). Todos los alelos fueron ingresados a una tabla u hoja de cálculo *Microsoft Excel* 2003 siendo las filas los loci SSR y las columnas las entradas de loche. Con estos datos se formó una matriz de 108 entradas x 21 iniciadores de microsatélites.

Determinación del índice de contenido polimórfico

Se determinó el índice de contenido polimórfico (PIC) para cada marcador o fragmento polimórfico (Guislain et al., 1999), usando la siguiente fórmula:

$$\text{PIC} = 1 - \sum_{i=1}^n f_i^2 = 1 - p^2 - q^2$$

Donde:

f_i : frecuencia de bando del i-ésimo alelo.

p : frecuencia alélica [presencia de banda (f1)]

q : frecuencia alélica [ausencia de banda (f0)]

Análisis de similitud

Dada la naturaleza codominante de los marcadores microsatélites se deben emplear coeficientes que consideren las semejanzas de presencias de los marcadores mas no las semejanzas de ausencias. El coeficiente de similitud de Dice está definido por la siguiente fórmula:

$$S_D = \frac{2a}{2a + b + c} = \frac{a}{(a + b) + (a + c)}$$

Donde:

a: número de marcadores compartidos por dos individuos

b: número de marcadores presentes en el individuo 1

c: número de marcadores presentes en el individuo 2, pero ausentes en 1.

Los análisis con los coeficientes de similitud genética se realizaron usando el módulo SIMQUAL, obteniéndose una matriz de similitud genética para el coeficiente.

Análisis de agrupamiento UPGMA

En base a la matriz de similitud que fue obtenida del análisis de similaridad con cada uno de los coeficientes usados anteriormente se obtuvieron los dendogramas, que se realizaron según el algoritmo UPGMA (*Unweighted Pair-Group Method using Arithmetic Averages*), dentro del módulo SAHN del programa NTSYS-pc versión 2.10 (Applied Biostatistics Inc., Setauket, Nueva York, EE.UU.)

Valor cofenético

Se calculó el coeficiente o valor cofenético empleando la opción COPH del software NTSYS pc 2.10, mediante el cual se infieren los valores de similitud o disimilitud a partir de los grupos formados en un árbol. Posteriormente para evaluar que tan bien se ajusta la matriz generada sobre la base del agrupamiento y la matriz de similitud o distancia se calculó la correlación cofenética usando el análisis MXComp. De acuerdo a Rohlf (2001), los rangos de confianza de la correlación cofenética (r) son:

| Rango | Interpretación |
|--------------------|-----------------------|
| $0.9 \leq r$ | Muy bueno |
| $0.8 \leq r < 0.9$ | Bueno |
| $0.7 \leq r < 0.8$ | Pobre |
| $r < 0.7$ | Muy pobre |

Estimadores de la diversidad genética

La frecuencia de un alelo SSR específico se calculó como el número de presencias de este alelo en un locus microsatélite; luego se calcularon las frecuencias relativas respecto al número total de alelos en la población. El cálculo de las frecuencias y de los índices de diversidad se realizó empleando el programa Microsoft Excel.

a) Diversidad genética de Nei o Heterocigosidad esperada (H_e)

Este índice es aplicable a cualquier población sin importar el número de alelos por locus, el patrón de las fuerzas evolutivas tales como mutación, migración, y método reproductivo del organismo (Nei, 1973). Este índice representa la probabilidad que 2 genes seleccionados al azar dentro de una población sea diferentes (Nei, 1973). Se calculó este índice para cada locus y el promedio se consideró como una medida de la diversidad genética poblacional.

Para cada locus,

$$h_j = 1 - \sum p_i^2$$

Donde:

P_i = frecuencia relativa del i-ésimo alelo

n = número de alelos observados en el locus j

h_j = probabilidad de que 2 alelos tomados al azar de la población sean diferentes

Promedio de diversidad genética o Poblacional (Provan, 1996),

$$H = \sum h_j / L$$

Donde:

H = la heterocigosidad promedio para varios loci

L = el número total de loci

h_j = la heterocigosidad por locus

Análisis factorial de coordenadas principales (PCo)

El tipo de análisis que se empleará en el presente trabajo es el Análisis Factorial de Coordenadas Principales (PCo). Las correlaciones se determinarán a partir de la matriz de distancia genética generada con el índice de similitud de Dice. Para realizar este análisis se empleará el Programa DARwin versión 5.0 (Perrier and Jacquemoud-Collet, 2006).

Análisis de variancia molecular (AMOVA)

El análisis de AMOVA como un método para estudiar la variación molecular dentro de las especies consideró la amplitud geográfica de donde provinieron las muestras. Para el presente trabajo de investigación, se realizaron dos análisis de variancia molecular (AMOVA) usando el programa Arlequin ver. 3.5.1.2 con 10,100 permutaciones y se permitió un error de 0.05. En el primer análisis AMOVA se estimó la variación genética de zapallos existente entre las poblaciones (entre las especies de *Cucurbita moschata*, *C. pepo* y *C. maxima*) y la variación genética existente entre los zapallos dentro de cada una de las poblaciones (especies). En el segundo análisis AMOVA se estimó la variancia genética de loches existente entre los grupo (localidades: Lambayeque y Amazonas), entre poblaciones (entre las comunidades) y la variación genética existente entre los loches dentro de cada una de las comunidades.

IV. RESULTADOS Y DISCUSIONES

4.1. ANÁLISIS MOLECULAR

Determinación de calidad y concentración de ADN

Se observó ADN de buena calidad y concentración aceptable para el estudio de diversidad genética. El rendimiento promedio del ADN total estuvo entre 50-100 ng/μl, concentración suficiente para el análisis molecular con la técnica SSR.

Selección de los iniciadores SSR

De los 21 iniciadores evaluados en este estudio, se seleccionaron 18 por presentar mayor nitidez de bandas, como se indica a continuación: CMTp206, CMTp211, CMTp245, CMTm9, CMTm39, CMTm42, CMTm86, CMTm112, CMTm155, CMTm158, CMTm165, CMTm177, CMTm186, CMTm243, CMTmC3, CMTmC11, CMTmC44, CMTmC64.

Los iniciadores que generaron el mayor número de bandas totales fueron los siguientes: CMTp245, CMTm9 y CMTmC3 con 4 bandas (Tabla 8).

Tabla 8: Número y porcentaje de bandas polimórficas y monomórficas detectadas en los patrones SSR.

| Iniciadores SSR | Bandas Monomórficas | %Bandas Monomórficas | Bandas Polimórficas | % Bandas Polimórficas | Nº total de bandas |
|-----------------|---------------------|----------------------|---------------------|-----------------------|--------------------|
| CMTp206 | 1 | 33.33 | 2 | 66.67 | 3 |
| CMTp211 | 1 | 50 | 1 | 50 | 2 |
| CMTp245 | 1 | 25 | 3 | 75 | 4 |
| CMTm9 | 1 | 25 | 3 | 75 | 4 |
| CMTm39 | 1 | 100 | 0 | 0 | 1 |
| CMTm42 | 1 | 50 | 1 | 50 | 2 |
| CMTm86 | 0 | 0 | 2 | 100 | 2 |
| CMTm112 | 1 | 33.3 | 2 | 66.67 | 3 |
| CMTm155 | 1 | 100 | 0 | 0 | 1 |
| CMTm158 | 1 | 50 | 1 | 50 | 2 |
| CMTm165 | 0 | 0 | 2 | 100 | 2 |
| CMTm177 | 1 | 50 | 1 | 50 | 2 |
| CMTm186 | 1 | 100 | 0 | 0 | 1 |
| CMTm243 | 1 | 33.3 | 2 | 66.67 | 3 |
| CMTmC3 | 2 | 50 | 2 | 50 | 4 |
| CMTmC11 | 1 | 100 | 0 | 0 | 1 |
| CMTmC44 | 1 | 50 | 1 | 50 | 2 |
| CMTmC64 | 1 | 100 | 0 | 0 | 1 |
| Total | 17 | | 23 | | 40 |
| Promedio | 0.94 | 42.5 | 1.28 | 57.5 | 2.22 |

FUENTE: Elaboración propia.

Marcadores SSR detectados

Un total de 40 bandas fueron detectados dentro de los 98 genotipos analizados de *C. moschata*, donde el 57.5% de bandas (23) SSR mostraron polimorfismo.

El número de loci polimórfico por iniciador varió de 0 a 3 bandas (CMTm39, CMTm155, CMTm186, CMTmC11, CMTmC64 y CMTp245, CMTm9), respectivamente, con un promedio de 1.28 bandas polimórficas por iniciador.

4.2. ANÁLISIS DE POLIMORFISMO

Determinación del índice de contenido polimórfico (PIC)

Se calcularon los valores PIC para cada fragmento polimórfico los cuales indicaron el poder de discriminación de cada marcador (Tabla 9).

Determinación de los iniciadores más polimórficos

En la tabla 9 se muestra el poder de discriminación para cada iniciador o primer, los cuales fueron estimados con los índices de iniciador SSR, siendo los primer CMT165, CMTp206, CMTp245 los más informativos, los cuales produjeron 2 bandas polimórficas para los dos primeros primers y 3 para el último (Tabla 9).

Tabla 9: Índice de contenido polimórfico de 18 iniciadores SSR utilizados en el presente estudio.

| Iniciador SSR | Índice de Contenido Polimórfico de Iniciador SSR |
|----------------------|---|
| CMTm165 | 0.82 |
| CMTp206 | 0.42 |
| CMTp245 | 0.39 |
| CMTm86 | 0.26 |
| CMTmC3 | 0.15 |
| CMTm9 | 0.14 |
| CMTm158 | 0.1 |
| CMTm112 | 0.08 |
| CMTmC44 | 0.08 |
| CMTm243 | 0.06 |
| CMTp211 | 0.06 |
| CMTm177 | 0.04 |
| CMTm42 | 0.04 |
| CMTm155 | 0 |
| CMTm186 | 0 |
| CMTm39 | 0 |
| CMTmC11 | 0 |
| CMTmC64 | 0 |
| Promedio | 0.15 |

4.3. ANÁLISIS DE SIMILARIDAD

Similitud genética

La similitud genética entre las 98 accesiones de Cucurbita moschata, 5 de C. pepo y 5 de C. maxima se calcularon empleando el coeficiente de similitud de Dice

Coeficiente de correlacion cofenético “r”

Una vez generada la matriz de similitud original, se evaluó esta para obtener la matriz cofenética, la cual resultó de un análisis cofenético a partir del dendograma generado con el coeficiente de similitud empleado en este estudio.

Se midió la concordancia entre ambas matrices (matriz de similitud a partir de dendograma y la nueva matriz cofenética) para evaluar la existencia de diferencias significativas en la correlación de las mismas, dando como resultado el coeficiente de correlación cofenético de 0.99, el cual es considerado dentro de un rango de correspondencia de muy buena.

4.4. ANÁLISIS DE AGRUPAMIENTO MOLECULAR

Se utilizaron todos los alelos encontrados para analizar la similitud genética y luego formar los agrupamientos moleculares.

Considerando un nivel de similitud de 0.6, el dendograma resultante (Figura 4) dividió a las 98 accesiones de *Cucurbita moschata* (loche) en un solo grupo. Las 5 de *C. maxima* y 5 *C. pepo* también fueron divididas en un solo grupo cada una. Dentro del cluster de loches, este se separa a un nivel de similitud de 0.62 aproximadamente conteniendo 2 subgrupos: *i*, *ii*. El grupo *i* representa a 97 accesiones de loche cultivadas en Lambayeque y Amazonas, el grupo *ii* representa a la única accesión proveniente Puerto Bermúdez, departamento de Pasco. Respecto al cluster de *C. maxima*, a un nivel de 0.78 aproximadamente se forman dos subgrupos. El cluster de *C. pepo* contiene dos subgrupos a un nivel de 0.64.

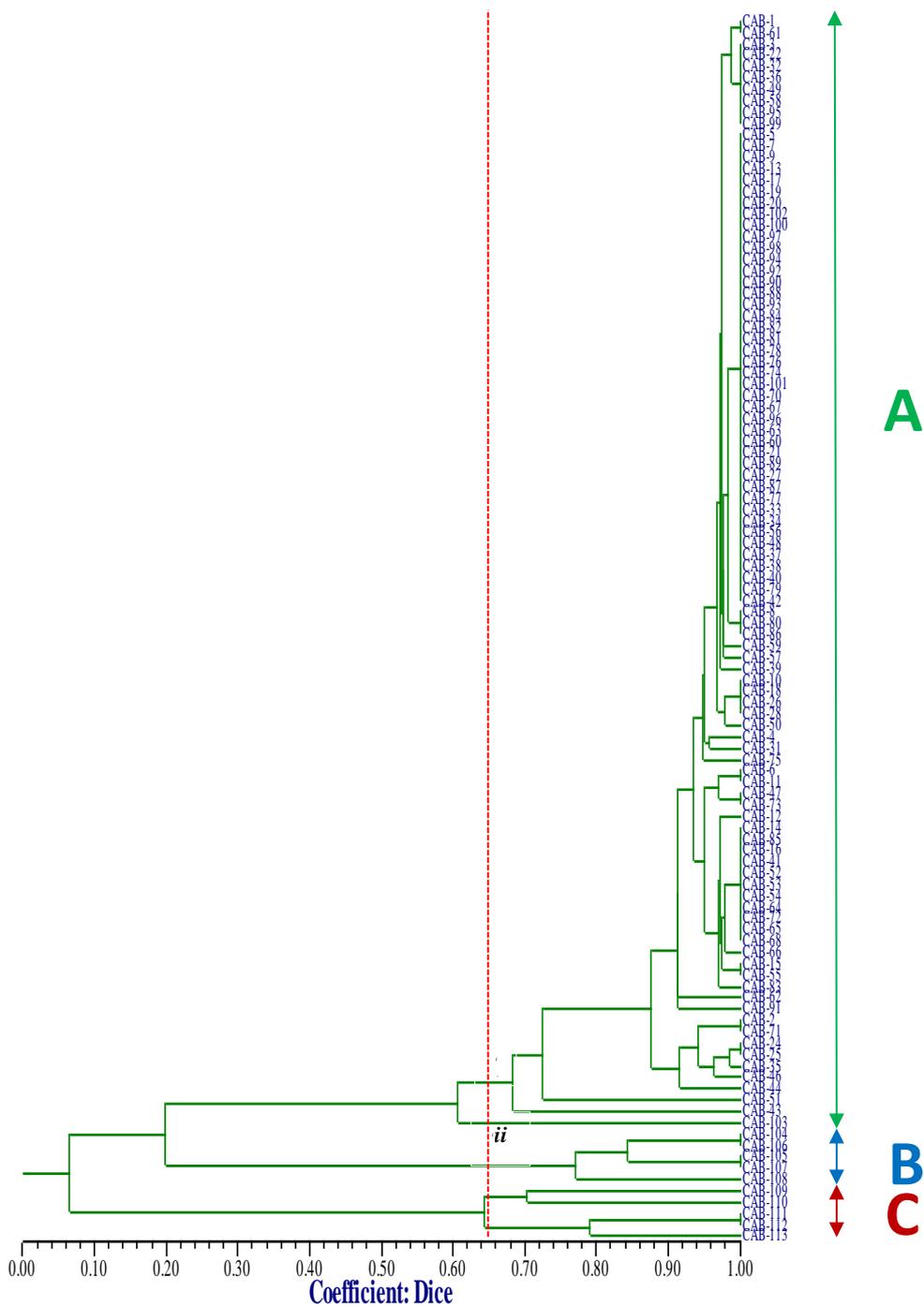


Figura 4: Dendrograma de *Cucurbita moschata* (A) cultivadas en Lambayeque Amazonas y Pasco, *C. maxima* (B) y *C. pepo* (C) cultivados en Ica y Lima respectivamente empleando el algoritmo UPGMA con el coeficiente de similitud Dice.

4.5. Índice de diversidad de Nei o heterocigosidad esperada (H_e) y observada (H_o)

En el grupo de loches analizados en el presente estudio se observó que el índice de diversidad genética de Nei o heterocigosidad esperada (H_e) estuvo en un rango entre 0 (CMTm39,

CMTm155, CMTm186, CMTmC11 y CMTmC64) y 0.42 (CMTp206) con un promedio de 0.10. Por otro lado, se tuvo resultado similares al analizar la heterocigosidad observada. Se obtuvo un rango de 0 (CMTm39, CMTm42, CMTm112, CMTm155, CMTm158, CMTm165, CMTm177, CMTm186, CMTmC11 y CMTmC64) a 0.25 (CMTp206). Se puede observar en la evaluación algunos loci con índice de diversidad igual a cero, esto significa que todos los cultivares presentaron los mismo alelos en ese locus (Tabla 10).

Tabla 10: Índice de diversidad genética de Nei o heterocigosidad esperada (He) y heterocigosidad observada (Ho) de los loches cultivados en Lambayeque, Amazonas y Pasco.

| Iniciador SSR | Heterocigosidad esperada | Heterocigosidad observada |
|--------------------------|-------------------------------------|--------------------------------------|
| CMTm39 | 0 | 0 |
| CMTm155 | 0 | 0 |
| CMTm186 | 0 | 0 |
| CMTmC11 | 0 | 0 |
| CMTmC64 | 0 | 0 |
| CMTm42 | 0.04 | 0 |
| CMTm177 | 0.04 | 0 |
| CMTp211 | 0.06 | 0.02 |
| CMTm243 | 0.06 | 0.004 |
| CMTm112 | 0.08 | 0 |
| CMTm158 | 0.1 | 0 |
| CMTmC44 | 0.1 | 0.02 |
| CMTmC3 | 0.11 | 0.24 |
| CMTm9 | 0.14 | 0.24 |
| CMTp245 | 0.2 | 0.22 |
| CMTm86 | 0.26 | 0.04 |
| CMTm165 | 0.27 | 0 |
| CMTp206 | 0.42 | 0.25 |
| Promedio | 0.1 | 0.06 |

FUENTE: Elaboración propia.

4.6. ANÁLISIS DE COORDENADAS PRINCIPALES

4.6.1. Análisis entre zapallos (*Cucurbita* sp.)

En el análisis factorial entre zapallos se graficaron los dos factores que daban la mayor información, Factor 1: 50.7% y Factor 2: 19.1%, confirmando una clara diferenciación de la especie *Cucurbita moschata* de las otras especies de zapallo (*C. maxima* y *C. pepo*) (Figura 5).

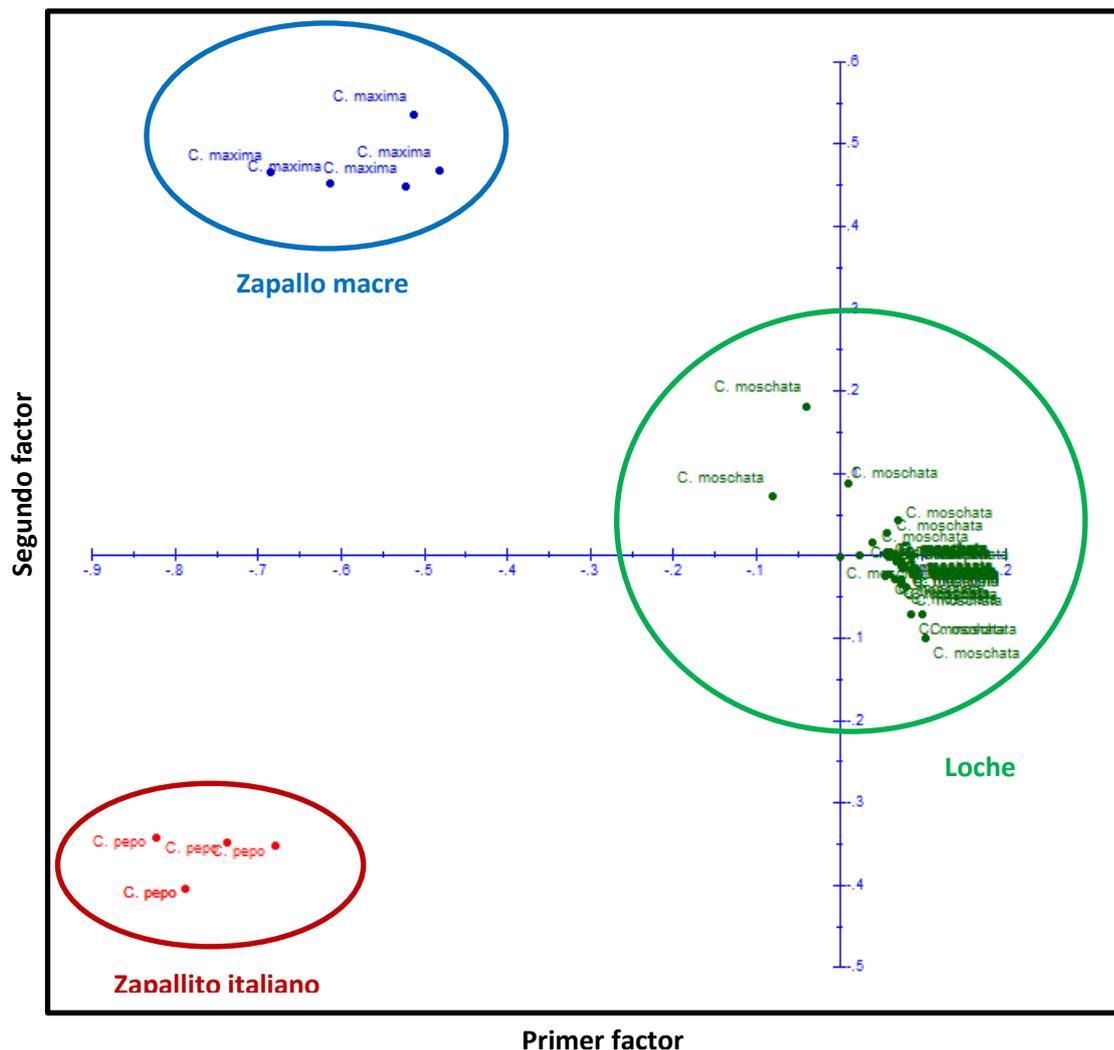


Figura 5: Análisis factorial de coordenadas principales en especies de zapallo basado en una matriz de distancia de Dice generada con 69.7% de la variancia total.

4.6.2 Análisis entre loches (*Cucurbita moschata*)

En el análisis factorial entre loches se graficaron los dos factores que daban la mayor información, Factor 1: 24.6% y Factor 2: 13.8%, confirmando una clara semejanza entre los loches cultivados en Lambayeque y Amazonas. Existen dos loches cultivados en Lambayeque que se agrupan a una pequeña distancia del grupo mayor de loches. Además, el loche cultivado en Puerto Bermúdez, Pasco se agrupa individualmente y lejano del mayor grupo de loches (Figura 6).

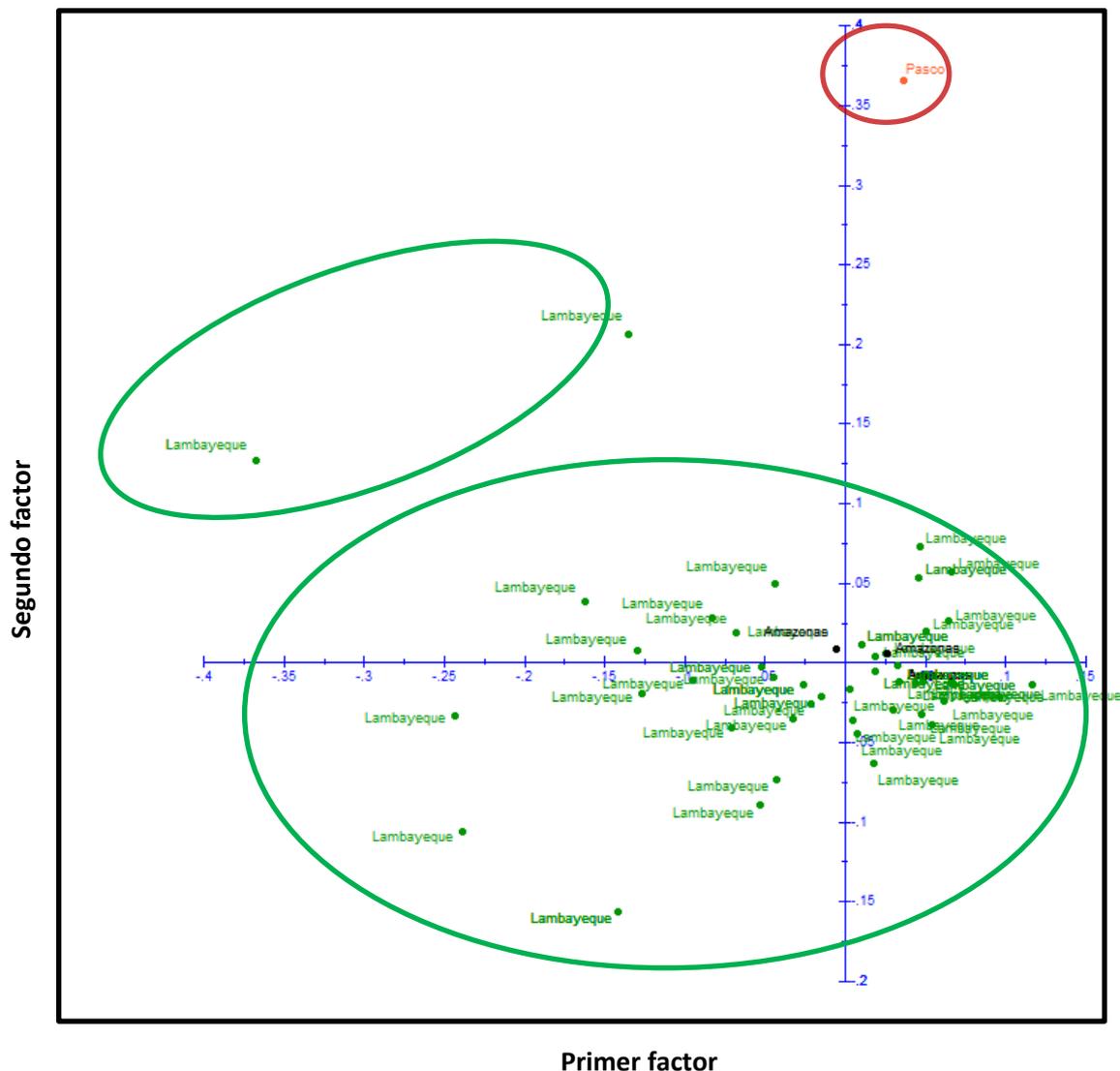


Figura 6: Análisis factorial de coordenadas principales en loches basados en una matriz de distancia de Dice generada con 38.4% de la variancia total.

4.7. ANÁLISIS DE VARIANZA MOLECULAR (AMOVA)

La estructura de la variabilidad genética de la colección fue tratada mediante el análisis de varianza molecular (AMOVA) (Excofier et al., 2010). Se realizaron dos AMOVA para estudiar la variación molecular dentro (i) de especies de *Cucurbita* sp. y (ii) de loches cultivados en Lambayeque y Amazonas.

4.7.1. AMOVA entre zapallos (*Cucurbita* sp.)

De acuerdo con los resultados, la principal fuente de variabilidad se encuentra entre las especies de zapallo analizadas (*Cucurbita moschata*, *C. maxima* y *C. pepo*) con un 93.1%, seguida por la variabilidad dentro de las poblaciones (6.9%). Ambas resultaron ser significativas. Además, se obtuvo un valor F_{st} de 0.93 lo que indica que existe una diferenciación genética muy grande entre las especies de zapallos estudiados (Tabla 11).

Tabla 11: Análisis de varianza molecular (AMOVA) para 108 genotipos de zapallo basados en 85 bandas de SSR.

| Fuente de variación | g.l . | Suma de cuadrado | Componente de la varianza | Porcentaje de variación |
|---|---------|------------------|---------------------------|-------------------------|
| Entre poblaciones (especies de zapallos) | 2 | 164.7 | 8.8 Va | 93.1 |
| Dentro de poblaciones (cultivares de zapallo) | 10 5 | 68.5 | 0.66 Vb | 6.9 |
| Total | 10 7 | 233.3 | 9.43 | |

FUENTE: Elaboración propia.

4.7.2. AMOVA entre loche (*Cucurbita moschata*)

En la tabla 12 podemos apreciar que del total de la variación genética observada en todas las comunidades y localidades estudiadas, el 95.54% de variabilidad se encontró dentro de las comunidades de Lambayeque: Illimo y Pitipo y por parte de Amazonas la comunidad de Bagua Grande. La siguiente fuente de variabilidad, con un 4.17%, se encontró entre los grupos estudiados: Lambayeque y Amazonas, seguida por una variabilidad de 0.29% entre comunidades dentro de los grupos estudiados. Se obtuvo un F_{st} de 0.04460 que nos indica que existe una pequeña diferenciación genética entre los loches cultivados en Lambayeque y Amazonas.

Tabla 12: Análisis de varianza molecular (AMOVA) para 98 genotipos de loche basados en 40 bandas de SSR.

| Fuente de variación | g.l | Suma de cuadrado | Componente de la varianza | Porcentaje de variación |
|--|------------|-------------------------|----------------------------------|--------------------------------|
| Entre grupos (Lambayeque y Amazonas) | 1 | 1.82 | 0.05 Va | 4.17 |
| Entre comunidades dentro de grupos (Lambayeque y Amazonas) | 1 | 1.28 | 0.004 Vb | 0.29 |
| Dentro de comunidades (Illimo, Pitipo, Bagua Grande) | 94 | 107.18 | 1.140 Vc | 95.54 |
| Total | 96 | 110.28 | 1.19 | |

FUENTE: Elaboración propia.

La utilización eficiente de los recursos fitogenéticos depende de su adecuada caracterización. Tradicionalmente, estas caracterizaciones se ha realizado empleando marcadores morfológicos. Esta metodología permitió diferenciar materiales vegetales, pero en ocasiones en forma imprecisa debido a la interacción de estas variables con las condiciones del medio ambiente. Los marcadores presentan muchas ventajas en comparación con los marcadores morfológicos, especialmente por su alto número y su independencia de los efectos del medio ambiente. Ambos marcadores, morfológicos y moleculares, se usan para caracterización de germoplasma. Un ejemplo del uso individual e integrado de estos marcadores se empleó en el caso de las zanahorias, pudiendo caracterizar las zanahorias silvestres y las cultivadas colectadas en alrededor de 40 países, y mantenidas por el USDA (United States Department of Agriculture) (Arbizu et al., 2014a,b; 2016a,b; 2017).

Las regiones simples repetidas que constituyen marcadores microsatélites se encuentran en gran cantidad y bien distribuidos en el genoma de las especies biológicas (Kruglyak et al., 1998). Lo que significa una ventaja para la búsqueda de marcadores microsatélites, especialmente en especies con potencial económico, pero que carecen de información genética; como es el caso del cultivo de loche. Los marcadores microsatélites fueron empleados en cultivos de importancia económica con escaso desarrollo de herramientas moleculares, tales como en la zanahoria (Iorizzo et al., 2011).

A la fecha no existen trabajos de investigación respecto al estudio de la diversidad genética de *Cucurbita moschata* mediante el uso de SSRs (microsatélites). Sin embargo, existen trabajos previos conducidos por Jeon et al. (1994) usaron marcadores RAPD para discriminar entre cultivares de *C. moschata* y *C. pepo*. Lee et al. (1995) también mediante marcadores RAPD analizaron grupos de ligamiento en la generación F2 de híbridos de *C. pepo* X *C. moschata*. De manera similar, Youn & Chung (1998) emplearon marcadores RAPD para investigar las relaciones genéticas entre los zapallos nativos coreanos. La diversidad genética dentro de las especies y entre las especies de *C. pepo*, *C. moschata*, y *C. maxima* también fue estudiada usando marcadores RAPD (Baranek et al., 2000). Los marcadores RAPD también fueron usados por Gwanama et al. (2000) para asistir en la selección de parentales para la mejora de las características del fruto de *C. moschata*, y para el análisis de la diversidad genética en las variedades tradicionales cultivadas en el sur y centro de África.

Ferriol et al. (2004b) usaron marcadores AFLP (Vos et al. 1995) y SRAP (Li & Quiros, 2001) para estudiar la diversidad genética de una colección de germoplasma de zapallo (*C. moschata*) mantenidas en España, encontrando que las accesiones estudiadas se agrupan de acuerdo a su origen geográfico. Esto sugiere la existencia de dos centros independientes de domesticación en el continente Americano, y/o introgresión de especies relacionadas a *C. moschata*. Wu et al. (2011) estudiaron la diversidad genética de *C. moschata* empleando marcadores AFLP y 74 accesiones de China, y 15 accesiones de otros países. Las accesiones de China fueron clasificadas en dos grupos, diferenciándose de las accesiones de México, Guatemala, Honduras, y Ecuador. Los autores concluyen que las accesiones Americanas presentan un mayor número de loci que las provenientes de China. La diversidad genética de accesiones de *C. moschata* de Mesoamérica fueron estudiadas mediante secuencias de ADN del cloroplasto, encontrándose un alto nivel de diversidad genética especialmente en el germoplasma Mexicano (Barboza et al. 2012). Los resultados de los estudios de la investigación indican que la diversidad genética de las accesiones fue alta y concordante con la diversidad morfoagronómica reportada. La mayoría de la variación genética entre las colecciones de germoplasma de zapallo se atribuyó a las variaciones entre individuos dentro de las localidades respectivas.

En el presente estudio de investigación se observa que la diversidad genética del loche en el Perú es bastante baja. Estos resultados parecen ser congruentes con las poblaciones de plantas casi homogéneas observadas en campos agrícolas de las comunidades de Lambayeque. Debido a que el loche se propaga vegetativamente, muy pocos clones estrechamente relacionados están siendo cultivados en las comunidades de Pacora e Illimo, así como en otras localidades del Perú. Por lo tanto, no existe oportunidad para la recombinación genética. Muy probablemente la variación morfológica ha surgido por mutación somática, que se mantiene y propaga por medio de estacas. Por otro lado, el loche de montaña que se cultiva en Bagua, Amazonas no tiene variación genotípica del loche cultivado en Lambayeque. La razón principal de esto es que los esquejes utilizados por los agricultores de Bagua serían los mismos que los empleados en Lambayeque debido a que los agricultores que han decidido vivir en Bagua tomaron sus semillas vegetativas de loche de ahí con el fin de continuar cultivando loche. Por lo tanto, la poca diferencia morfológica que existe entre estos frutos de loche cultivado en estas dos localidades se puede explicar por factores ambientales.

La similitud de los loches a un nivel de 0.6 nos indica que se forma un solo grupo considerando todas las 98 accesiones de estos. Dentro del grupo de loches, podemos encontrar dos sub-grupos: i) loches cultivados en Lambayeque y Amazonas, y ii) loche cultivado en Puerto Bermúdez, Pasco. Por otro lado, nuestras especies control, *C. maxima* y *C. pepo* se encuentran en grupos totalmente distintos. Estos debido a que son especies distintas al loche.

Se pudo también confirmar la transferibilidad de los marcadores microsatélites dentro del género *Cucurbita*, tal como ha sido descrita para otros géneros que incluyen *Phaseolus* y *Vigna* (Yu et al., 1999) *Lycopersicon* (Alvarez et al., 2001) *Oryza* (Gao et al., 2005) *Coffea* (Cubry et al., 2007), entre otros.

V. CONCLUSIONES

- La diversidad genética del loche cultivado en Lambayeque es baja, esto debido a su propagación clonal.
- No se encontró diferencia genotípica significativa entre el loche cultivado en Lambayeque y el proveniente de Bagua, Amazonas.
- La diferencia morfológica (presencia de protuberancias en la piel del fruto) se puede explicar por factores ambientales entre estas dos zonas geográficas.

VI. RECOMENDACIONES

Aún se necesitan realizar estudios taxonómicos y agronómicos para conocer más sobre la clasificación, y manejo de cultivo del loche. Es importante realizar la caracterización agronómica de las entradas usadas en el presente estudio para entablar congruencia con el análisis molecular. Los resultados obtenidos aquí permiten un primer acercamiento a la caracterización molecular a nivel de ADN de las entradas de loche. Sin embargo, es recomendable profundizar a estudios a nivel molecular, dando soporte a programas de mejoramiento genético. De esta manera se podrán desarrollar cultivares con mejores características agronómicas. Queda pendiente aún afianzar los vínculos de investigación con las empresas privadas de semillas. Al lograr ello se puede estudiar el loche y aprovechar sus bondades nutricionales y agronómicas, desarrollando cultivares híbridos para los agricultores peruanos.

VII. REFERENCIAS BIBLIOGRÁFICAS

- Alvarez, A.E., C.C.M. van de Wiel, M.J.M. Smulders, B. Vosman. 2001. Use of microsatellites to evaluate genetic diversity and species relationships in the genus *Lycopersicon*. *Theoretical and Applied Genetics* 103:1283-1292.
- Andres T.C., R. Ugás, & F. Bustamante. 2006. Loche: A unique pre-Columbian squash locally grown in North Coastal Peru. En: Proceedings of Cucurbitaceae G.J. Holmes (eds.). Universal Press, Raleigh, North Carolina, USA. pp. 333-340.
- Anónimo. 2007. Instructivo técnico del cultivo de la calabaza. Biblioteca ACTAF.
- Arbizu, C.I., P.M. Tas, P.W. Simon, & D.M. Spooner. 2017. Phylogenetic prediction of *Alternaria* leaf blight resistance in wild and cultivated species of carrots. *Crop Science* 57:1-9.
- Arbizu, C.I., S.L. Ellison, D. Senalik, P.W. Simon, & D.M. Spooner. 2016a. Genotyping-by-sequencing provides the discriminating power to investigate the subspecies of *Daucus carota* (Apiaceae). *BMC Evolutionary Biology* 16:234.
- Arbizu, C.I., P.W. Simon, F. Martínez-Flores, H. Ruess, M.B. Crespo, & D.M. Spooner. 2016b. Integrated molecular and morphological studies of the *Daucus guttatus* complex (Apiaceae). *Systematic Botany* 41:479-492.
- Arbizu, C., K.R. Reitsma, P.W. Simon, & D.M. Spooner. 2014a. Morphometrics of *Daucus* (Apiaceae): A counterpart to a phylogenomic study. *American Journal of Botany* 101:2005-2016.
- Arbizu, C., H. Ruess, D. Senalik, P.W. Simon, & D.M. Spooner. 2014b. Phylogenomics of the carrot genus (*Daucus*, Apiaceae). *American Journal of Botany* 101:1666–1685.
- Baranek, M., G. Stift, J. Vollmann, & T. Lelley. 2000. Genetic diversity within and between the species *Cucurbita pepo*, *C. moschata* and *C. maxima* as revealed by RAPD Markers. *Cucurbit Genetics Cooperative Report* 23:73-77.
- Barboza, N., F.J. Albertazzi, J.A. Sibaja-Cordero, F. Mora-Umaña, C. Astorga, & P. Ramírez. 2012. Analysis of genetic diversity of *Cucurbita moschata* (D.) germplasm accessions from Mesoamerica revealed by PCR SSCP and chloroplast sequence data. *Scientia Horticulturae* 134:60-71.

- Blair, M.W., & S.R. McCouch. 1997. Microsatellite and sequence tagged site markers diagnostic for the rice bacterial leaf blight resistance gene xa-5. *Theoretical and Applied Genetics* 95:174-184.
- Blankenship, S.M., B. May, & D. Hedgecock. 2002. Evolution of a perfect simple sequence repeat locus in the context of its flanking sequence. *Molecular Biology and Evolution* 19:1943-1951.
- Botstein D, R.L. White, M. Skolnick, & R.W Davis. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32:314-331.
- Brack, A. 1999. Diccionario Enciclopédico de Plantas Útiles del Perú. Ediciones CBC. Cusco, Perú.
- Brondani, R.P.V., C. Brondani, R. Tarchini, & D. Grattapaglia. 1998. Development, characterization and mapping of microsatellite markers in *Eucalyptus grandis* and *E. urophylla* *Theoretical and Applied Genetics* 97: 816-827.
- Brush, S.B. 1995. In situ conservation of landraces in centers of crop diversity. *Crop Science* 35: 346-354.
- Bryan, G.J., A.J. Collins, P. Stephenson, A. Orry, J.B. Smith, & M.D. Gale. 1997. Isolation of microsatellites from hexaploid bread wheat. *Theoretical and Applied Genetics* 94:557-563.
- Buerstmayr, H., M. Lemmens, L. Hartl, L. Doldi, B. Steiner, M. Stierschneider, & P. Ruckebauer. 2002. Molecular mapping of QTL for fusarium head blight resistance in spring wheat. I. Resistance to fungal spread (type II resistance). *Theoretical and Applied Genetics* 104: 84-91.
- Bukasov, S. M. 1981. Las Plantas Cultivadas de Mexico, Guatemala & Colombia. CATIE-GTZ, Turrialba, Costa Rica. 121pp.
- Casper, W. 2001. The role of dietary fibre in the digestive physiology of the pig. *Animal Feed Science and Technology* 90:1-2.
- Chambers G.K., & E.S. Macavoy. 2000. Microsatellites: Consensus and controversy. *Comparative Biochemistry and Physiology* 126:455-476.
- Cheng, F., Brown, S. & Weeden, N., 1997. A DNA extraction protocol from various tissues in woody species. *Hortscience* 32:921-922.
- Chien, A., D.B. Edgar, & J M Trela. Deoxyribonucleic acid polymerase from the extreme thermophile *Thermus aquaticus*. *Journal of Bacteriology* 127: 1550-1557.

- Cho, Y.G., T. Ishii, S. Temnykh, X. Chen, L. Lipovich, S.R. McCouch, W.D. Park, N. Ayres, & S. Cartinhour. 2000. Diversity of microsatellites derived from genomic libraries and GenBank sequences in rice (*Oryza sativa* L.). *Theoretical and Applied Genetics* 100:713-722.
- Cogniaux, A. 1881. Cucurbitacées. En: A. De Candolle & C. De Candolle (Eds.), *Monographiae Phanerogamarum*. G. Masson, Paris, Francia.
- Collins, W.W., & G.C. Hawtin. 1999. Conserving and using crop plant biodiversity in agroecosystems. In: W.W. Collins & C.O. Qualset (eds.). *Biodiversity in agroecosystems*. CRC Press LLC, Boca Raton, Florida, USA.
- Coombs, J. 1986. *MacMillan Dictionary of biotechnology*. MacMillan Press Ltd. London and Basingtoke, UK.
- Couch, J., & P. Fritz. 1990. Isolation of DNA from plants high in polyphenolics. *Plant Molecular Biology Reporter* 8:8-12.
- Crisci, J.V., & M.F. López. 1983. Introducción a la teoría y práctica de la taxonomía numérica. Washington, D.C.: Departamento de asuntos científicos y tecnológicos de la Secretaría General de la OEA, serie de Biología, 132 p.
- Cruces, C.R. 1987. Lo Que México aportó al Mundo. *Panorama México*. Editorial Lectorum
- Cubry, P. P. Musoli, H. Legnaté, D. Pot, F. de Bellis, V. Poncet, F. Anthony, M. Dufour, & T. Leroy. 2008. Diversity in coffee assessed with SSR markers: Structure of the genus *Coffea* and perspectives for breeding. *Genome* 51:50–63.
- Decker-Walters, D.S., & T.W. Walters. 2000. Squash. En: K.F. Kiple, & K.C. Ornelas (eds.). *The Cambridge world history of food*. Cambridge University Press. Cambridge, UK.
- Dellaporta, S.L., J. Wood, & J.B. Hicks, 1983. A plant DNA miniprep: version II. *Plant. Molecular Biology Reporter* 1:19-21.
- Di Rienzo, A., A.C. Peterson, J.C. Garza, A.M. Valdes, M. Slatkin, N.B. Freimer. 1994. Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United States of America* 91:3166-3170.
- Díaz, A. 1957. Estudio de las variedades nacionales de zapallos. Tesis para Ingeniero Agrónomo. Escuela Nacional de Agricultura, Lima, Perú.
- Doyle, J.J., & J.L. Doyle. 1990. Isolation of plant DNA from fresh tissue. *Focus* 12:13-15.

- Elshire, R.J., J.C. Glaubitz, Q. Sun, J.A. Poland, K. Kawamoto, E.S. Buckler, S.E. Mitchell. 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* 6: e19379
- Erlich, H.A. 1989. Polymerase chain reaction. *Journal of Clinical Immunology* 9: 437-447.
- Erwin, A.T. & E.P. Lana. 1956. The Seminole pumpkin. *Economic Botany* 10: 33-37.
- Excoffier L, & H.E: Lischer. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources* 10: 564-567.
- Faccioli P., N. Pecchioni, A.M. Stanca, & Terzi, V. 1999. Amplified fragment length polymorphism (AFLP) markers for barley malt fingerprinting. *Journal of Cereal Science* 29:257-260.
- FAO 1989. Prevention of the Post-Harvest Food Losses. Fruits, Vegetables and Root Crops. A Training Manual. Roma, Italia
- Feldman, M.W., A. Bergman, D.D. Pollock, & D.B. Goldstein. 1997. Microsatellite genetic distances with range constraints: analytic description and problems of estimation. *Genetics* 145:207-216.
- Ferreira, M., & D. Grattapaglia. 1998. Introducción al uso de marcadores moleculares en el análisis genético. EMBRAPA-CENARGEN. Brasilia, Brasil.
- Ferriol, M., B. Picó, & F. Nuez. 2004a. Morphological and Molecular Diversity of a Collection of Cucurbita maxima Landraces. *Journal of the American Society for Horticultural Science* 129: 60-69.
- Fleschar R.;Kimball M. 1993. Glossary of biotechnology terms technomic publishing company, Inc. Lancaster, Pennsylvania, USA.
- Freifelder, D. 1991. Técnicas de Bioquímica & Biología Molecular, Editorial Reverté, Barcelona, España.
- Gao, L.Z. 2005. Microsatellite diversity and population genetic structure of an endangered wild rice, *Oryza officinalis* (Poaceae) from China. *Molecular Ecology* 14:4287-4297.
- García, C.L. 2006. Características de la especie *Cucurbita pepo* L. & su semilla en el tratamiento de la próstata. Resumen del congreso de fitoterapia.

- Ghislain, M., D. Zhang, D. Fajardo, Z. Huamán, & R. J. Hijmans. 1999. Marker-assisted sampling of the cultivated Andean potato *Solanum phureja* collection using RAPD markers. *Genetic Resources and Crop Evolution* 46:547-555.
- Ghislain M, F. Rodriguez, F. Villamon, J. Nunez, R. Waugh, & M. Bonierbale M. 2000. Establishment of microsatellite assays for potato genetic identification. *Research on Potato* 167-174.
- Ghislain, M., B. Trognitz, M. del Rosario Herrera, J. Solis, G. Casallo, C. Vásquez, O. Hurtado, R. Castillo, L. Portal, & M. Orrillo. 2001. Genetic loci associated with field resistance to late blight in offspring of *Solanum phureja* and *S. tuberosum* grown under short-day conditions. *Theoretical and Applied and Genetics* 103:433-442.
- Gong L, G. Stift, R. Kofler, M. Pachner, & T. Lelley. 2008. Microsatellites for the genus *Cucurbita* and an SSR-based genetic linkage map of *Cucurbita pepo* L. *Theoretical and Applied Genetics* 117:37-48.
- González-Andrés, F. 2001. La caracterización vegetal: Objetivos & enfoques. En: F. González-Andrés. & J. M. Pita Villamil (Eds.). Conservación y caracterización de recursos filogenéticos. Publicaciones Instituto Nacional de Educación Agrícola. Valladolid, España.
- Gwanama, C., MT. Labuschagne, & A.M. Botha. 2000., Analysis of genetic variation in *Cucurbita moschata* by random amplified polymorphic (RAPD) markers. *Euphytica* 113:19-24.
- Harlan, J. R. 1992. Crops and man. 2nd ed. Am. Soc. Agronomy, Madison, WI.
- Heale S.M., T.D. Petes. 1995. The stabilization of repetitive tracts of DNA by variant repeats requires a functional DNA mismatch repair system. *Cell* 83:539-545.
- Hollingsworth, P.M.; & R.A. Ennos. 2004. Neighbour joining trees, dominant markers and population genetic structure. *Heredity* 92: 490-498.
- Iorizzo, M., D.A Senalik, D. Grzebelus, M. Bowman, P.F Cavagnaro, M. Matvienko, H. Ashrafi, A. Van Deynze, & P.W Simon. 2011. De novo assembly and characterization of the carrot transcriptome reveals novel genes, new markers, and genetic diversity. *BMC Genomics* 12:389.
- J. Muminovic, A. E. Melchinger, & T. Lubberstedt. 2004. Genetic diversity in cornsalad (*Valerianella locusta*) and related species as determined by AFLP markers. *Plant Breeding* 123:460-466.

- Jarne, P., & P.J.L. Lagoda. 1996. Microsatellites, from molecules to populations and back. *Trends in Ecology Evolution* 11:424-429.
- Jeon, H.J., C.G. Been, K.H. Hong, Y.K. Om, B.D. Kim. 1994. Identification of cucurbitacea cultivars by using RAPD markers. *Journal of Korean Society of Horticultural Science* 35:449-456.
- Jin, L., C. Macaubas, J. Hallmayer, A. Kimura, & E. Mignot. 1996. Mutation rate varies among alleles at a microsatellite locus: Philogenetic evidence. *Proceedings of the National Academy of Sciences of the United States of America* 93: 15285-15288.
- Katti M.V., P.K. Ranjekar, V.S. Gupta. 2001. Differential distribution of simples sequence repeats in eukaryotic genome sequences. *Molecular Biology and Evolution* 18:1161-1167.
- Kay, M., F. B. King, & C. K. Robinson. 1980. Cucurbits from Phillips Spring: New evidence and interpretation. *American Antiquity* 45: 806–822.
- Kimuro, M., & J.F. Crow. 1964. The number of alleles that can be maintained in a finite population. *Genetics* 49:725-738.
- Kneller, G.F. 1978. Arte e ciência da criatividade (5. ed.). São Paulo: IBRASA.
- Kruglyak, S.R. M. Durrett, M. Schug, & C.F. Aquadro. 1998. Equilibrium distributions of microsatellite repeat length resulting from a balance between slippage events and point mutations. *Proceedings of the National Academy of Sciences of the United States of America* 95:10774-10778.
- Lee, Y.H., H.J. Jeon, K.H. Hong, & B.D. Kim. 1995. Use of random amplified polymorphic DNA for linkage group analysis in an interspecific cross hybrid F2 generation of *Cucurbita*. *Journal of Korean Society of Horticultural Science* 36:323-330.
- Lehninger, A., D.L. Nelson, & M.M. Cox. Principles of Biochemistry. Worth Publishers, New York, New York, EEUU.
- Li, G. & C. Quiros. 2001. Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in Brassica. *Theoretical and Applied Genetics* 103:455-461.
- Lira R 1991 Estudios taxonómico & ecogeográfico de las Cucurbitáceas de Latinoamérica. 3rd biannual report (Jan.-Aug. 1991). IBPGR Rome, Italy.

- Lira, R., & S. Montes. 1994. Cucurbits (*Cucurbita* spp.). In: Neglected crops: 1492 from a different perspective. J.E. Hernandez Bermejo, and J. León (eds.). Food and Agriculture Organization of the United Nations. Rome, Italy.
- Lira, R. 1995. Estudios Taxonómicos & Ecogeográficos de las Cucurbitaceae Latinoamericanas de Importancia Económica: Cucurbita, Sechium, Sicana & Cyclanthera. Systematic and Ecogeographic Studies on Crop Gene Pools. 9. International Plant Genetic Resources Institute. Rome, Italy. 281p.
- Lombard, V., C.P. Baril, P. Dubreuil, F. Blouet, & D. Zhang, 2000. Genetic relationships and fingerprinting of rapeseed cultivars by AFLP consequences for varietal registration. *Crop Science* 40: 1417-1425.
- Mba, R.E.C., P. Stephenson, K. Edwards, S. Melzer, J. Nkumbira, U. Gullberg, K. Apel, M. Gale, J. Tohme, M. Fregene. 2001. Simple sequence repeat (SSR) markers survey of the cassava (*Manihot esculenta* Crantz) genome: towards an SSR-based molecular genetic map of cassava. *Theoretical and Applied Genetics* 102:21-31.
- Merrick, L.C. 1995. Squashpumpkins, and gourds: *Cucurbita*(Cucurbitaceae) Smart, J.Simmonds, N.W.ed.s. Evolution of Crop Plants, Longman, Londres.
- Milbourne, D., R.C. Meyer, A.J. Collins, L.D. Ramsay, C. Gebhardt, & R. Waugh. 1998. Isolation, characterization and mapping of simple sequence repeat loci in potato. *Molecular Genetics and Genomics* 259: 233-245.
- Mitas, M. 1997. Trinucleotide repeats associated with human disease. *Nucleic Acids Research* 25:2245-2254.
- Moreiras, O., A. Carbajal, L. Cabrera, C. Cuadrado. 2001. Tabla de composición de alimentos. Ediciones Piramide. Madrid, España.
- Morgante, M., & A.M. Olivieri. 1993. PCR-amplified microsatellite as markers in plant genetics. *Plant Journal* 3: 175-182.
- Ferriol, M., B. Picó, P. Fernández de Córdova, & F. Nuez. 2004b. Molecular diversity of a germplasm collection of squash (*Cucurbita moschata*) determined by SRAP and AFLP markers. *Crop Science* 44:653-664.
- Moxon E.R., & C. Wills. 1999. DNA microsatellites: Agents of evolution? *Sci Am.* 280:94-99.
- Mueller, U.G., & L.L. Wolfenbarger. 1999. AFLP genotyping and fingerprinting. *Trends in Ecology and Evolution* 14:389-394.

- Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., Erlich, H. 1986. Specific Enzymatic Amplification of DNA In Vitro: The Polymerase Chain Reaction *Cold Spring Harbor Symposia on Quantitative Biology* 51: 263-273.
- National Research Council. 1989. Lost Crops of the Incas: Little-Known Plants of the Andes with Promise for Worldwide Cultivation. National Academy Press. Washington D.C.
- Nee, M. 1990. The domestication of *Cucurbita* (Cucurbitaceae). *Economic Botany* 44 (Supl.): 56-58.
- Negri, V., & N. Tosti. 2002. Phaseolus genetic diversity maintained on-farm in central Italy. *Genet Resources and Crop Evolution* 49: 511-520.
- Nei M, 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences of the United States of America* 70:3321-3323.
- Oda S, E. Oki, Y. Maehara, & K. Sugimachi. 1997. Precise assessment of microsatellite instability using high resolution fluorescent microsatellite analysis. *Nucleic Acids Research* 25:3415-3420.
- Ohta, T., & M. Kimura. 1973. A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a finite population. *Genetics Research* 22:201-204.
- Perrier, X., & J.P. Jacquemoud-Collet. 2006. Darwin software, <http://darwin.cirad.fr/darwin>
- Picca, A., M. Helguera, N. Salomón, & A. Carrera. 2004. Marcadores Moleculares. In: V. Echenique, C. Rubinstein, & L. Mroginsky (eds.). *Biotecnología & mejoramiento Vegetal*. INTA, Buenos Aires, Argentina.
- Powell W, M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey., & A. Rafalski. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for gemplasm analysis. *Molecular Breeding* 2:225-238.
- Prasad, M., R.K. Varshney, J.K. Roy, H.S. Balyan, & P. K. Gupta. 2000. The use of microsatellites for detecting DNA polymorphism, genotype identification and genetic diversity in wheat. *Theoretical and Applied Genetics* 100:584-592.
- Provan, J., W. Powell, & R. Waugh. 1996. Microsatellite analysis of relationships within cultivated potato (*Solanum tuberosum*). *Theoretical and Applied Genetics* 92: 1078-1084.

- Provan, J., W. Powell, & R. Waugh. 1996. Microsatellite analysis of relationships within cultivated potato (*Solanum tuberosum*). *Theoretical and Applied Genetics* 92: 1078-1084.
- R K. Varshney, A. Graner, & M. E. Sorrells. 2005. Genic microsatellite markers in plants:
- Restrepo J.A., & F.A. Vallejo. 2008. Caracterización molecular de introducciones colombianas de zapallo Cucurbita moschata. *Acta agronómica* 57: 9-18.
- Robinson JP, & S.A. Harris. 1999. En: Which DNA Marker for Which Purpose. Gillet EM, editor. Institut für Forstgenetik und Forstpflanzenzüchtung, Universität Göttingen; Göttingen, Germany. pp. 1-27.
- Robinson, R.W., & D.S. Decker-Walters. 1997. Cucurbits. CAB International, New York, USA.
- Rohlf, F.J. 1970. Adaptive hierarchical clustering schemes. *Systematic Zoology* 19:58-82.
- Rohlf, F.J. 2001. Comparative methods for the analysis of continuous variables geometric interpretations. *Evolution* 55:2143-2160.
- Rossetto, M., A. Mclachlan, F.C.L. Harriss, R.J. Henry, P.R. Baverstock, L.S. Lee, T.L. Maguire, & K.J. Edwards. 1999. Abundance and polymorphism of microsatellite markers in the tea tree (*Melaleuca alternifolia*, Myrtaceae). *Theoretical and Applied Genetics* 98:1091-1098.
- Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, & H.A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science* 239:487-491.
- Santos, C., V. Sakal, & M. Machado. 2003. Reverse transcription and polymerase chain reaction: principles and Applications in dentistry. *Journal of Applied Oral Science* 12: 1-11.
- Schlötterer, C., & D. Tautz. 1992. Slippage synthesis of simple sequence DNA. *Nucleic Acids Research* 20: 211-215.
- Schuster, I., R.V. Abdelnoor, S.R.R. Marin, V.P. Carvalho, R.A.S. Kiihl, J.F.V. Silva, C.S. Sedyama, E. G. Barros, & M.A. Moreira. 2001. Identification of a new major QTL associated with resistance to soybean cyst nematode (*Heterodera glycines*). *Theoretical and Applied Genetics* 102:91-96.

- Scotti, I., G. Paglia, F. Magni, & M. Morgante. 2000. Microsatellite markers as a tool for the detection of intra- and interpopulational genetic structure. *Bundesforschungsanstalt für Forst- und Holzwirtschaft* 198:53-58.
- Sevilla, R., & M. Holle, 2004. Recursos genéticos vegetales. Univesidad Nacional Agraria la Molina, Lima, Perú.
- Shannon, C.E., & Weaver, W. 1949. The mathematical theory of communication. Urbana: University of Illinois Press.
- Simpson, E.H. 1949. Measurement of diversity. *Nature* 163: 688.
- Sneath, P.H.A., & R.R. Sokal. 1973. Numerical taxonomy: the principles and practice of numerical classification. Ed. Freeman, San Francisco, California. 573pp.
- Sokal, R.R., & F.J. Rohlf. 1962. The comparison of dendrograms by objective methods. *International Association for Plant Taxonomy* 11, 33-40.
- Soukup, J. 1987. Vocabulario de los nombres vulgares de la flora peruana & catálogo de los géneros. Editorial Salesiana, Lima. Perú.
- Soukup, J. 1970. Vocabulary of the common names of the Peruvian flora and catalog of the genera, Editorial Salesiano, Lima, Perú.
- Soulé, M.E., & L.S. Mills. 1998. Population genetics – No need to isolate genetics. *Science* 282:1658-1659.
- Tautz, D. & M. Renz. 1984. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research* 12:4127-4138.
- Tautz, D. 1989. Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucleic Acids Research* 17: 6463-6471.
- Testolin, R, M.T. Marrazo, G. Cipriani, R. Quarta, I. Verde, M.T. Dettori, M. Pancaldi, & S. Sansavini. 2000. Microsatellite DNA in peach [*Prunus persica* (L.) Batsch] and its use in fingerprinting and testing the genetic origin of cultivars. *Genome* 43: 512-520.
- Tóth G., Z. Gáspári, & J. Jurka. 2000. Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research* 10:967-981.
- Trujillo, L.G. 2004. Desarrollo de marcadores SCAR y CAPS en un QTL con efecto importante sobre la Resistencia al Tizón tardío de la papa. Tesis para optar el título de Biólogo con mención en Genética. UNMSM, Lima, Perú.
- Turner, B.L., & C. Miksicek. 1984. Economic plant species associated with prehistoric agriculture in the Maya lowlands. *Economic Botany* 38: 179-193.

- Valadez, E., & G. Kahl. 2000. Huellas de DNA en Genomas de Plantas (Teoría & Protocolos de Laboratorio). Mundiprensa. México D. F., México.
- van Hintum, Th.J.L., & R. van Treuren. 2002. Molecular markers: tools to improve genebank efficiency. *Cellular & Molecular Biology Letters* 7:737-744
- Varshney RK, Graner A, and Sorrells ME. 2005. Genic microsatellite markers in plants: features and applications. *Trends in Biotechnology* 23:48-55.
- Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Freijters, J. Pot, J. Peleman, M. Kuiper, & M. Zabeau. 1995. AFLP: a new concept for DNA fingerprinting. *Nucleic Acids Research* 21:4407-4414.
- Walsh P.S., N.J. Fildes, R. Reynolds. 1996. Sequence analysis and characterization of stutter products at the tetranucleotide repeat locus vWA. *Nucleic Acids Research* 24: 2807-2812.
- Weising, K., H. Nybom, K. Wolff, & W. Meyer. 1995. In: DNA Fingerprinting in Plants and Fungi (ed. Arbor, A.). CRC Press, Boca Raton, Florida, USA.
- Whitaker, T.W. 1959. An interspecific cross in *Cucurbita* (*C. lundelliana* Bailey x *C. moschata* Duch.). *Madroño* 15:4-13.
- Whitaker, T.W., & G.W Bohn. 1950. The taxonomy, genetics, production and uses of the cultivated species of *Cucurbita*. *Economic Botany* 4:52-81.
- Winzer, N., M. Di Renzo, & S. Olmos. 2004. Métodos para estimar variabilidad genética. INTA, Buenos Aires, Argentina, pp.: 199-210.
- Wu, J., Z. Chang, Q. Wu, H. Zhan, & S. Xie. 2011. Molecular diversity of Chinese *Cucurbita moschata* germplasm collections detected by AFLP markers. *Scientia Horticulturae* 128:7-13.
- Youn, S.J., & H.D. Chung. 1998. Genetic relationship among the local varieties of the Korean native squashes (*Cucurbita moschata*) using RAPD technique. *Journal of the Korean Society for Horticultural Science* 39:517-521
- Yu, K., S.J. Park, & V. Poysa. 1999. Abundance and variation of microsatellite DNA sequences in beans (*Phaseolus* and *Vigna*). *Genome* 42:27-34.
- Zane L., L. Bargelloni, & T. Patarnello T. 2002. Strategies for microsatellite isolation: A review. *Molecular Ecology* 11:1-16.

VIII. ANEXOS

Anexo 1: Lista de accesiones de *Cucurbita* sp. empleados en el presente estudio.

| Entrada | Especie | Nombre común | Lugar | Distrito | Provincia | Departamento | Altitud | Latitud | Longitud |
|---------|--------------------|--------------|------------------|----------|------------|--------------|---------|------------|------------|
| CAB-01 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 66 | 06°28'29'' | 79°49'35'' |
| CAB-02 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 66 | 06°28'29'' | 79°49'35'' |
| CAB-03 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 66 | 06°28'29'' | 79°49'35'' |
| CAB-04 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 66 | 06°28'29'' | 79°49'35'' |
| CAB-05 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 66 | 06°28'29'' | 79°49'35'' |
| CAB-06 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 66 | 06°28'29'' | 79°49'35'' |
| CAB-07 | <i>C. moschata</i> | Loche | Sapamé II | Illimo | Lambayeque | Lambayeque | 66 | 06°28'26'' | 79°49'32'' |
| CAB-08 | <i>C. moschata</i> | Loche | Sapamé II | Illimo | Lambayeque | Lambayeque | 65 | 06°28'26'' | 79°49'32'' |
| CAB-09 | <i>C. moschata</i> | Loche | La Choza | Illimo | Lambayeque | Lambayeque | 54 | 06°28'40'' | 79°51'08'' |
| CAB-10 | <i>C. moschata</i> | Loche | La Choza | Illimo | Lambayeque | Lambayeque | 54 | 06°28'40'' | 79°51'08'' |
| CAB-11 | <i>C. moschata</i> | Loche | La Choza | Illimo | Lambayeque | Lambayeque | 54 | 06°28'40'' | 79°51'08'' |
| CAB-12 | <i>C. moschata</i> | Loche | La Choza | Illimo | Lambayeque | Lambayeque | 54 | 06°28'40'' | 79°51'08'' |
| CAB-13 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 55 | 06°28'38'' | 79°50'46'' |
| CAB-14 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 55 | 06°28'38'' | 79°50'46'' |
| CAB-15 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 55 | 06°28'38'' | 79°50'46'' |
| CAB-16 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 55 | 06°28'38'' | 79°50'46'' |
| CAB-17 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 55 | 06°28'38'' | 79°50'46'' |
| CAB-18 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 55 | 06°28'38'' | 79°50'46'' |
| CAB-19 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 55 | 06°28'41'' | 79°49'33'' |
| CAB-20 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 55 | 06°28'41'' | 79°49'33'' |
| CAB-21 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 55 | 06°28'41'' | 79°49'33'' |
| CAB-22 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 55 | 06°28'41'' | 79°49'33'' |
| CAB-23 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 63 | 06°28'42'' | 79°49'20'' |
| CAB-24 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 63 | 06°28'42'' | 79°49'20'' |
| CAB-25 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 63 | 06°28'42'' | 79°49'20'' |
| CAB-26 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 67 | 06°28'45'' | 79°49'24'' |
| CAB-27 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 67 | 06°28'45'' | 79°49'24'' |
| CAB-28 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 67 | 06°28'45'' | 79°49'24'' |

| | | | | | | | | | |
|--------|--------------------|-------|-----------------------------------|--------|------------|------------|----|------------|------------|
| CAB-29 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 66 | 06°28'45'' | 79°49'24'' |
| CAB-30 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 66 | 06°28'45'' | 79°49'24'' |
| CAB-31 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 65 | 06°28'37'' | 79°49'31'' |
| CAB-32 | <i>C. moschata</i> | Loche | Sapamé de Illimo | Illimo | Lambayeque | Lambayeque | 65 | 06°28'37'' | 79°45'34'' |
| CAB-33 | <i>C. moschata</i> | Loche | Bosque de Pomac-Batan Grande | Illimo | Lambayeque | Lambayeque | 78 | 06°29'08'' | 79°49'31'' |
| CAB-34 | <i>C. moschata</i> | Loche | Bosque de Pomac-Batan Grande | Illimo | Lambayeque | Lambayeque | 78 | 06°29'08'' | 79°49'31'' |
| CAB-35 | <i>C. moschata</i> | Loche | Bosque de Pomac-Batan Grande | Illimo | Lambayeque | Lambayeque | 78 | 06°29'08'' | 79°49'31'' |
| CAB-36 | <i>C. moschata</i> | Loche | San Jorge,lindero de Batán grande | Illimo | Lambayeque | Lambayeque | 67 | 06°28'12'' | 79°49'03'' |
| CAB-37 | <i>C. moschata</i> | Loche | San Jorge,lindero de Batán grande | Illimo | Lambayeque | Lambayeque | 67 | 06°28'12'' | 79°49'03'' |
| CAB-38 | <i>C. moschata</i> | Loche | San Jorge,lindero de Batán grande | Illimo | Lambayeque | Lambayeque | 67 | 06°28'12'' | 79°49'03'' |
| CAB-39 | <i>C. moschata</i> | Loche | San Jorge,lindero de Batán grande | Illimo | Lambayeque | Lambayeque | 67 | 06°28'12'' | 79°49'03'' |
| CAB-40 | <i>C. moschata</i> | Loche | San Jorge,lindero de Batán grande | Illimo | Lambayeque | Lambayeque | 67 | 06°28'12'' | 79°49'03'' |
| CAB-41 | <i>C. moschata</i> | Loche | San Jorge,lindero de Batán grande | Illimo | Lambayeque | Lambayeque | 67 | 06°28'12'' | 79°49'03'' |
| CAB-42 | <i>C. moschata</i> | Loche | San Jorge,lindero de Batán grande | Illimo | Lambayeque | Lambayeque | 67 | 06°28'12'' | 79°49'03'' |
| CAB-43 | <i>C. moschata</i> | Loche | San Jorge,lindero de Batán grande | Illimo | Lambayeque | Lambayeque | 67 | 06°28'07'' | 79°49'25'' |
| CAB-44 | <i>C. moschata</i> | Loche | San Jorge,lindero de Batán grande | Illimo | Lambayeque | Lambayeque | 67 | 06°28'07'' | 79°49'25'' |
| CAB-45 | <i>C. moschata</i> | Loche | San Jorge,lindero de Batán grande | Illimo | Lambayeque | Lambayeque | 67 | 06°28'07'' | 79°49'25'' |
| CAB-46 | <i>C. moschata</i> | Loche | San Jorge,lindero de Batán grande | Illimo | Lambayeque | Lambayeque | 67 | 06°28'07'' | 79°49'25'' |

| | | | | | | | | | |
|--------|--------------------|-------|------------------------------------|--------|------------|------------|----|------------|------------|
| CAB-47 | <i>C. moschata</i> | Loche | San Jorge, lindero de Batán grande | Illimo | Lambayeque | Lambayeque | 67 | 06°28'07'' | 79°49'25'' |
| CAB-48 | <i>C. moschata</i> | Loche | Huaca de Piedra, Tres Viudas | Illimo | Lambayeque | Lambayeque | 66 | 06°28'12'' | 79°49'38'' |
| CAB-49 | <i>C. moschata</i> | Loche | Huaca de Piedra, Tres Viudas | Illimo | Lambayeque | Lambayeque | 66 | 06°28'12'' | 79°49'38'' |
| CAB-50 | <i>C. moschata</i> | Loche | Huaca de Piedra, Tres Viudas | Illimo | Lambayeque | Lambayeque | 66 | 06°28'12'' | 79°49'38'' |
| CAB-51 | <i>C. moschata</i> | Loche | Huaca de Piedra, Tres Viudas | Illimo | Lambayeque | Lambayeque | 66 | 06°28'12'' | 79°49'38'' |
| CAB-52 | <i>C. moschata</i> | Loche | Cruz Verde | Illimo | Lambayeque | Lambayeque | 64 | 06°28'17'' | 79°50'30'' |
| CAB-53 | <i>C. moschata</i> | Loche | Cruz Verde | Illimo | Lambayeque | Lambayeque | 64 | 06°28'17'' | 79°50'30'' |
| CAB-54 | <i>C. moschata</i> | Loche | Cruz Verde | Illimo | Lambayeque | Lambayeque | 64 | 06°28'17'' | 79°50'30'' |
| CAB-55 | <i>C. moschata</i> | Loche | Cruz Verde | Illimo | Lambayeque | Lambayeque | 64 | 06°28'17'' | 79°50'30'' |
| CAB-56 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 68 | 06°28'39'' | 79°49'20'' |
| CAB-57 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 68 | 06°28'39'' | 79°49'20'' |
| CAB-58 | <i>C. moschata</i> | Loche | Sapamé | Illimo | Lambayeque | Lambayeque | 68 | 06°28'39'' | 79°49'20'' |
| CAB-59 | <i>C. moschata</i> | Loche | Calla Salas (Bagua) | Illimo | Lambayeque | Lambayeque | 35 | 06°46'18'' | 79°50'18'' |
| CAB-60 | <i>C. moschata</i> | Loche | Calla Salas (Callanca) | Illimo | Lambayeque | Lambayeque | 35 | 06°46'18'' | 79°50'18'' |
| CAB-61 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 64 | 06°25'45'' | 79°48'23'' |
| CAB-62 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 64 | 06°25'45'' | 79°48'23'' |
| CAB-63 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 64 | 06°25'45'' | 79°48'23'' |
| CAB-64 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 64 | 06°25'45'' | 79°48'23'' |
| CAB-65 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | | 06°26'07'' | 79°48'22'' |
| CAB-66 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 67 | 06°26'07'' | 79°48'22'' |
| CAB-67 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 67 | 06°26'07'' | 79°48'22'' |
| CAB-68 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 67 | 06°26'07'' | 79°48'22'' |
| CAB-69 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 67 | 06°26'07'' | 79°48'22'' |
| CAB-70 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 67 | 06°26'07'' | 79°48'22'' |
| CAB-71 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 67 | 06°26'07'' | 79°48'22'' |
| CAB-72 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 75 | 06°26'08'' | 79°48'13'' |
| CAB-73 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 75 | 06°26'08'' | 79°48'13'' |

| | | | | | | | | | |
|--------|--------------------|-------|---------------------------------|----------|-----------|------------|-----|------------|------------|
| CAB-74 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 75 | 06°26'08'' | 79°48'13'' |
| CAB-75 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 75 | 06°26'08'' | 79°48'13'' |
| CAB-76 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 71 | 06°26'25'' | 79°48'15'' |
| CAB-77 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 71 | 06°26'25'' | 79°48'15'' |
| CAB-78 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 71 | 06°26'25'' | 79°48'15'' |
| CAB-79 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 71 | 06°26'25'' | 79°48'15'' |
| CAB-80 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 71 | 06°26'34'' | 79°48'19'' |
| CAB-81 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 71 | 06°26'34'' | 79°48'19'' |
| CAB-82 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 71 | 06°26'34'' | 79°48'19'' |
| CAB-83 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 71 | 06°26'34'' | 79°48'19'' |
| CAB-84 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 75 | 06°26'42'' | 79°48'23'' |
| CAB-85 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 75 | 06°26'42'' | 79°48'23'' |
| CAB-86 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 75 | 06°26'42'' | 79°48'23'' |
| CAB-87 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 75 | 06°26'42'' | 79°48'23'' |
| CAB-88 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 72 | 06°26'02'' | 79°48'08'' |
| CAB-89 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 72 | 06°26'02'' | 79°48'08'' |
| CAB-90 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 72 | 06°26'02'' | 79°48'08'' |
| CAB-91 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 72 | 06°26'02'' | 79°48'08'' |
| CAB-92 | <i>C. moschata</i> | Loche | Poma III | Pitipo | Ferreñafe | Lambayeque | 72 | 06°26'02'' | 79°48'08'' |
| CAB-93 | <i>C. moschata</i> | Loche | Mercado 28 de Julio (Bagua) | Agustino | Lima | Lima | 173 | 12°03'40'' | 77°00'40'' |
| CAB-94 | <i>C. moschata</i> | Loche | Mercado 28 de Julio (Bagua) | Agustino | Lima | Lima | 173 | 12°03'40'' | 77°00'40'' |
| CAB-95 | <i>C. moschata</i> | Loche | Mercado 28 de Julio (Bagua) | Agustino | Lima | Lima | 173 | 12°03'40'' | 77°00'40'' |
| CAB-96 | <i>C. moschata</i> | Loche | Mercado 28 de Julio (Bagua) | Agustino | Lima | Lima | 173 | 12°03'40'' | 77°00'40'' |
| CAB-97 | <i>C. moschata</i> | Loche | Mercado 28 de Julio (Bagua) | Agustino | Lima | Lima | 173 | 12°03'40'' | 77°00'40'' |
| CAB-98 | <i>C. moschata</i> | Loche | Mercado 28 de Julio (Illimo) | Agustino | Lima | Lima | 173 | 12°03'40'' | 77°00'40'' |

| | | | | | | | | | |
|---------|--------------------|-----------------------|---------------------------------|-----------------|----------|-------|-----|------------------|------------------|
| CAB-99 | <i>C. moschata</i> | Loche | Mercado 28 de Julio (Illimo) | Agustino | Lima | Lima | 173 | 12°03'40'' | 77°00'40'' |
| CAB-100 | <i>C. moschata</i> | Loche | Mercado 28 de Julio (Illimo) | Agustino | Lima | Lima | 173 | 12°03'40'' | 77°00'40'' |
| CAB-101 | <i>C. moschata</i> | Loche | Mercado 28 de Julio (Illimo) | Agustino | Lima | Lima | 173 | 12°03'40'' | 77°00'40'' |
| CAB-102 | <i>C. moschata</i> | Loche | Mercado 28 de Julio (Illimo) | Agustino | Lima | Lima | 173 | 12°03'40'' | 77°00'40'' |
| CAB-103 | <i>C. moschata</i> | Loche | Puerto Bermúdez | Puerto Bermúdez | Oxapampa | Pasco | 280 | 10°17'33'' | 74°56'30'' |
| CAB-104 | <i>C. maxima</i> | Zapallo | Callango | Callango | Ica | Ica | 406 | 14°29'53'' | 75°37'57'' |
| CAB-105 | <i>C. maxima</i> | Zapallo | Callango | Callango | Ica | Ica | 406 | 14°29'53'' | 75°37'57'' |
| CAB-106 | <i>C. maxima</i> | Zapallo | Callango | Callango | Ica | Ica | 406 | 14°29'53'' | 75°37'57'' |
| CAB-107 | <i>C. maxima</i> | Zapallo | Ocucaje | Ocucaje | Ica | Ica | 395 | 14°20'45'' | 75°40'00'' |
| CAB-108 | <i>C. maxima</i> | Zapallo | Callango | Callango | Ica | Ica | 406 | 14°29'53'' | 75°37'57'' |
| CAB-109 | <i>C. pepo</i> | Zapallito italiano | Callango | Callango | Ica | Ica | 406 | 14°29'53'' | 75°37'57'' |
| CAB-110 | <i>C. pepo</i> | Zapallito italiano | Nuevo Imperial | Nuevo Imperial | Cañete | Lima | 110 | 13°03'03'' | 76°20'04'' |
| CAB-111 | <i>C. pepo</i> | Zapallito italiano | Nuevo Imperial | Nuevo Imperial | Cañete | Lima | 110 | 13°03'03'' | 76°20'04'' |
| CAB-112 | <i>C. pepo</i> | Zapallito italiano | Nuevo Imperial | Nuevo Imperial | Cañete | Lima | 110 | 13°03'03'' | 76°20'04'' |
| CAB-113 | <i>C. pepo</i> | Zapallito italiano | Macas | Huamantanga | Canta | Lima | 600 | 11° 41' 25.9" | 76° 50' 22.3" |

Anexo 2: Protocolo de extracción de ADN genómico vegetal *Doyle & Doyle. 1990.*

1. Moler el tejido con N₂ líquido y llevar a tubos eppendorf pre enfriados, lo que permite la ruptura de las paredes celulares.
2. Añadir a cada muestra Buffer CTAB 2X - Mercapto (700 µl - 2 µl, respectivamente). Para 200 muestras, 140 y 400 µl de CTAB 2X - Mercapto respectivamente. Con el fin de solubilizar las membranas lipoprotéicas y desnaturalizar las proteínas, preservando el DNA de las enzimas que la degradan.
3. Llevar a baño maría 60 - 65 °C por 30 min. (Dejar enfriar por 3 min aprox a temperatura ambiente). Lo que permite la solubilización y homogenización del DNA, proteínas, y polímeros contaminantes.
4. Añadir 700 µl Cloroformo - OH isoamílico (24 : 1) mezclar hasta obtener un líquido lechoso. El Cloroformo - OH isoamílico, en reposo, forma dos fases: Fase inorgánica (alcohol isoamílico, superior), fase orgánica (Cloroformo, inferior); en la fase orgánica se resuspenden las proteínas y lípidos
5. Centrifugar por 10 min a 14 000 rpm a 4 °C.
6. Añadir 50 µl de CTAB 10%-NaCl 7M a nuevos tubos y llevar la fase inorgánica de los anteriores a estos. Añadir 60 µl de CTAB (1° sobrenadante a los tubos y CTAB, luego mezclar).
7. Añadir 700 µl Cloroformo OH - Isoamílico (24 - 1) y mezclar enérgicamente.
8. Centrifugar por 10 min a 14000 rpm a 4 °C.
9. Transferir el sobrenadante a nuevos tubos y añadir igual volumen de isopropanol. Mezclar lentamente hasta observar la "medusa" de DNA. Centrifugar a 1000 rpm por 20 min a 4 °C.
10. Eliminar el isopropanol y dejar la "medusa" en el tubo. Agregar 500 µl de etanol 70% por dos veces. Centrifugar a 1000 rpm 20 min a 4 °C.
11. Eliminar el Etanol 70%. Agregar 500 µl de etanol al 90%. Centrifugar a 1000 rpm 20 min a 4 °C (se puede guardar las muestras en el congelador).
12. Eliminar el OH hasta evaporarlo.
13. Agregar Buffer TE y mezclar hasta obtener un líquido homogéneo. Almacenar a 4 ó - 20 °C.
14. La calidad y cuantificación del ADN se realiza mediante electroforesis en geles de agarosa al 1%.
15. La cuantificación del ADN extraído se determina mediante análisis comparativo de las muestras coloreadas con bromuro de etidio en geles de agarosa al 1% y un

buffer de corrida TBE 1X. Para determinar la calidad de ADN se realiza una corrida electroforética a 100 voltios por 1 hora, agregando 5 ul de azul de bromofenol a la muestra de ADN (2ul). La concentración de ADN se determina por comparación con el marcador de peso molecular ADN lambda digerido con Pst I. Determinadas las concentraciones de ADN problema luego estandarizar a 500 ng en volumen final de 10ul, para AFLPs.

Buffer de extracción (CTAB 2X)

| Reactivo | [final] | 100 ml |
|-----------------------|---------|--------------|
| CTAB | 2% | 2 g |
| NaCl 5M | 1,4 M | 28 ml |
| EDTA 0,5M (pH 8,0) | 20 mM | 4 ml |
| Tris:HCl 1 M (pH 8,0) | 100 mM | 10 ml |
| PVP | 1% | 1 g |
| Agua destilada | | hasta 100 ml |

Buffer de extracción (CTAB 10X)

| Reactivo | [Final] | 100 ml |
|-----------------------|---------|--------|
| CTAB | 10% | 10 g |
| NaCl 5M | 0,7 M | 14 ml |
| EDTA 0,5M (pH 8,0) | 20 mM | 4 ml |
| Tris:HCl 1 M (pH 8,0) | 100 mM | 10 ml |
| Agua destilada | | 72 ml |

Preparación de soluciones

Tris-HCl 1M (pH=7.5)

| | |
|----------------|---------|
| Tris base | 121.1 g |
| Agua destilada | 1000 ml |

Disolver en 800 ml de agua destilada, ajustar el pH requerido con HCl concentrado y enrasar con agua destilada hasta 1 litro.

Solución stock de Bromuro de Etidio

| | |
|-----------------------------------|---------|
| BrEt (10 mg / ml) | 1000 mg |
| Agua destilada H ₂ Odd | 100 ml |

Buffer T₁₀ E_{0.1}

10 mM Tris-HCl, pH 7.5

1 mM EDTA

Buffer TBE

Tris base 108 g

Ácido bórico 55 g

Na₄EDTA 9.3 g

Enrasar con agua destilada a 1litro

El pH es de 8.3 y no requiere de ajustes

Blue Juice 5X

Solución de 20% Ficoll, 0.01M TRIS, 0.01M EDTA, 0.09% Xylene Cyanol,
0.09% azul de bromofenol.

Tampón de carga

Dye stock:

Azul de bromofenol 0.5 g

Xilene xianol 0.5 g

10 ml de agua milliQ

Dye de corrida:

Formamida 9.6 ml

0.2 M EDTA 500 µl

Dye stock 100 µl

Solución de acrilamida 6%

Solución stock:

Acrilamida 60 g.

Bis-acrilamida 3 g

Urea 420 g

Buffer TBE 10X 50 ml

Enrasar a 1litro con agua destilada

La acrilamida y la bis-acrilamida son sustancias muy peligrosas, y más aún cuando están en polvo, el peso de los componentes indicados tiene que llevarse a cabo bajo una campana de extracción y usando mascarilla para evitar inhalar partículas tóxicas. Aunque el peso de la urea y la medida del buffer TBE pueden llevarse a cabo fuera de la campana y sin mascarilla.

Se recomienda comenzar la preparación con la dilución de la cantidad de urea indicada en 500 ml de agua destilada; por cuanto es difícil disolverlo. Para acelerar la dilución, puede colocarse en baño maría, pero no sin antes lavar muy bien con detergente la superficie del frasco (frasco vacío donde se prepara acrilamida), así como los guantes que se está usando. Colocando a temperatura de 65 °C, se hizo la dilución de la urea.

Seguidamente, se coloca la acrilamida y bis acrilamida, empleando un embudo de plástico de uso exclusivo para ellos; esto se hace en la campana extractora, utilizando mascarilla y guantes. También se vierte el volumen indicado de buffer TBE 10X.

Se agita para mezclar, se enraza a 1000 ml (de preferencia en el mismo frasco) con agua destilada, y luego se continúa agitando hasta que los componentes queden completamente disueltos.

El frasco se envuelve con papel platino, se pone el rótulo (Acrilamida 6%), y se conserva a temperatura ambiente.

Tinción argéntica de geles con carbonato:

Solución fijadora

Agua destilada 900 ml.

Ácido acético 100 ml

Solución de tinción

Nitrato de plata (AgNO₃) 2 g.

Agua destilada 900 ml

Formaldehido 1.5 ml

Solución reveladora

Carbonato de sodio 18 g.

(Na₂CO₃)

Agua destilada 900 ml

Formaldehido 900 ul

Tiosulfato de sodio 20 ul

Es recomendable que el formaldehido sea colocado justo antes de hacer el revelado.