

Universidad de Costa Rica

Facultad de Ciencias

Escuela de Biología

Tesis de graduación para optar por el grado de Licenciatura en

Biología Molecular y Biotecnología:

**Caracterización de la diversidad genética de maíz criollo (*Zea mays* L.) de las
regiones Brunca y Chorotega, Costa Rica**

Sofía Carvajal Rojas

A91389

Ciudad Universitaria Rodrigo Facio

2018

Miembros del tribunal

Ph.D. Eric Fuchs Castillo

Director de tesis

Msc. Griselda Arrieta Espinoza

Asesora

Dr. Federico Albertazzi Castro

Asesor

Dra. Laura Solís Ramos

Representante del Director de la Escuela de Biología

Miembro del tribunal

Dr. Javier Trejos Zelaya

Decano de la Facultad de Ciencias

Presidente del Tribunal

Bach. Sofía Carvajal Rojas

Postulante

Ciudad Universitaria Rodrigo Facio, 7 de junio de 2018

A mi familia.

AGRADECIMIENTOS

Le agradezco a mi familia por el apoyo incondicional, la paciencia y el entendimiento durante este proceso y a lo largo de mi vida.

Al proyecto de maíz criollo por ser una propuesta interesante y darme la oportunidad de formar parte de esta.

A mi director de tesis Eric Fuchs por la guía y la ayuda con la redacción de la tesis y con los análisis estadísticos; a Griselda Arrieta Espinoza por toda la ayuda con el proyecto de maíz, por las enseñanzas, las sugerencias y la guía durante la tesis y el trabajo en el laboratorio. También le agradezco a Federico Albertazzi por los comentarios y por haberme dado la oportunidad de empezar a conocer el campo de la biología molecular de plantas.

A Genuar Núñez por el apoyo con la planeación del diseño experimental, el trabajo en laboratorio, el análisis de los datos, y por la paciencia para enseñarme a usar todos los softwares para la realización de este trabajo.

Le agradezco al personal del Centro de Investigación en Biología Celular y Molecular y a los compañeros del laboratorio del Programa de Mejoramiento Genético de Cultivos, sobre todo a Adriana Orozco y a María Fernanda Miranda por la ayuda con la extracción del ADN.

ÍNDICE GENERAL

1. INTRODUCCIÓN.....	1
1.1. Antecedentes	1
1.2. Justificación.....	7
1.3 Marco Teórico	9
II OBJETIVOS	13
2.1: Objetivo general	13
2.2: Objetivos específicos	13
III METODOLOGÍA	14
IV RESULTADOS	26
4.1- Estandarización de Microsatélites para maíz criollo de Costa Rica	26
4.2- Caracterización de la diversidad genética de maíz criollo.....	32
V. DISCUSIÓN	38
VI. CONCLUSIONES	45
VII. RECOMENDACIONES	46
VIII REFERENCIAS.....	47
IV ANEXOS	55

Índice de cuadros

Cuadro 1: Información sobre los 10 SSR dinucleotídicos utilizados para caracterizar la diversidad genética de maíz criollo de las regiones Brunca y Chorotega de Costa Rica. Se muestra el marcador SSR, el Contenido de Información Polimórfica (PIC), el motivo de repetición (Motif), los imprimadores Forward (F-) y Reverse (R-), cromosoma al que pertenecen, el Rango alélico y la referencia de donde se tomaron los SSR y su información respectiva.	19
Cuadro 2: Información sobre los 10 SSR tetranucleotídicos utilizados para caracterizar la diversidad genética de maíz criollo de las regiones Brunca y Chorotega de Costa Rica. Se muestra el marcador SSR, el Contenido de Información Polimórfica (PIC), el motivo de repetición (Motif), los imprimadores Forward (F-) y Reverse (R-), cromosoma al que pertenecen , el Rango alélico y la referencia de donde se tomaron los SSR y su información respectiva.....	20
Cuadro 3: Protocolo de amplificación base (Qiagen, 2010) utilizado para el PCR multiplex de los 20 SSR seleccionados para el estudio de diversidad genética de las regiones Brunca y Chorotega de Costa Rica.....	21
Cuadro 4: Resumen de experimentos de estandarización de condiciones de PCR multiplex para 20 SSR seleccionados para el estudio de diversidad genética de las regiones Brunca y Chorotega de Costa Rica. <i>Ta</i> = Temperatura de annealing (hibridación) en °C, <i>vf</i> = volumen final.....	22
Cuadro 5: Resultados de prueba de Wilcoxon para la comparación de promedios de las estimativas de diversidad genética por microsatélite. Se muestran los valores de la prueba (W) la significancia (p) y los intervalos de confianza.....	32
Cuadro 6: Estimativas de diversidad genética por microsatélite, luego de la limpieza de datos.....	33
Cuadro 7: Estimativas de diversidad genética por región socioeconómica Brunca y Chorotega y en promedio para Costa Rica, obtenidas con el análisis de 20 SSR.....	35

Cuadro 8: Riqueza alélica de cada microsatélite según la región socioeconómica. Se muestra el tamaño poblacional utilizado (N) para cada región y la estimativa de riqueza (Riqueza), el promedio por región y la desviación estándar.....36

Cuadro 9: Estimativas de diversidad genética por sitio de muestreo dentro de cada región, según el promedio de los 20 SSR.....37

Índice de figuras

Figura 1: Mapa de Costa Rica con las regiones Chorotega (Norte) y Brunca (Sur) resaltadas en gris. En las regiones se muestran de color los sitios de colecta. Dentro de la Región Chorotega cada círculo corresponde a una accesión en cada sitio de colecta; en la Región Brunca, las accesiones se representan con cuadrarros.16

Figura 2: Distribuciones de fragmentos para definir un alelo en *MsatAllele*. El eje y son las frecuencias de individuos y en el eje x el tamaño de los fragmentos. Cada distribución de fragmentos conformará un solo alelo.24

Figura 3: Electroferogramas visualizados en GeneMarker, de 4 muestras de maíz criollo para el marcador bnlgl1890, según la prueba de $Ta=55^{\circ}\text{C}$, 30ciclos, una concentración de ADN de 50ng/uL y con un volumen final de reacción de 15 mL. Se denotan muchas amplificaciones inespecíficas.27

Figura 4: Electroferogramas visualizados en GeneMarker para una muestra de maíz criollo con el marcador bnlgl1890. Se observa que no existe mayor diferencia entre utilizar un volumen final de 10uL (panel A) o uno de 15uL (panel B).28

Figura 5: Electroferogramas visualizados en GeneMarker para una muestra de maíz criollo con el marcador phi078. Se observa la diferencia entre utilizar diferentes temperaturas de hibridación con el mismo protocolo. A $Ta=57^{\circ}\text{C}$, B $Ta=60^{\circ}\text{C}$ y C $Ta=63$29

Figura 6: Electroferogramas visualizados en GeneMarker para 4 muestras de maíz criollo con el marcador phi078. Se observan las diferencias entre realizar una PCR a 27 ciclos (A) que a 30 ciclos (B) ambas con una Ta de 60°C30

Índice de anexos

Anexo 1: Protocolo de extracción de ADN utilizado para maíz.55

Anexo 2: Conformación de los múltiplex (filas) y color del fluorocromo (columna) que se utilizaron para la caracterización de la diversidad genética de maíz criollo de las regiones Brunca y Chorotega de Costa Rica con 20 SSR.56

Anexo 3: Reacción de PCR multiplex (Qiagen, 2010) utilizada para la amplificación de 20 SSR seleccionados para el estudio de diversidad genética de las regiones Brunca y Chorotega de Costa Rica.56

Anexo 4: Estimativas de diversidad genética para cada microsatélite con la base de datos de genotipos obtenida con el paquete MsatAllele de R y sin verificación manual de los alelos. ... 57

Anexo 5: Rango de Bins obtenidos para los 20SSR con MsatAllele y limpieza manual. Se muestran los microsatélites, el valor del bin mínimo (Min) y el valor máximo (Max).58

ABREVIATURAS

ADN: ácido desoxirribonucleico.

F: índice de endogamia (fijación).

He: heterocigocidad esperada.

Ho: heterocigocidad observada.

Máx: alelo máximo.

Mín: alelo mínimo.

N: número de individuos.

Na: número de alelos.

Ne: número de alelos efectivos.

PCR: Reacción en cadena de la polimerasa (Polymerase Chain Reaction por sus siglas en inglés).

PIC: Contenido de información polimórfico.

SSR: Secuencias simples repetidas (Simple Sequence Repeats por sus siglas en inglés).

RFU: Unidades relativas de fluorescencia.

Ta: Temperatura de hibridización (annealing).

v.f: volumen final.

RESUMEN

El maíz criollo se caracteriza por tener una gran variedad de coloraciones en el grano, por ser cultivado y conservado por pequeños agricultores que lo utilizan para autoconsumo y preparación de comidas tradicionales. Además, está adaptado a condiciones agroecológicas según el sitio donde se cultiva, por lo que constituye un germoplasma valioso para su conservación. Sin embargo, la variabilidad genética de este recurso, en Costa Rica, aún no había sido descrita. El objetivo de la investigación fue caracterizar la diversidad genética del maíz criollo de las regiones Brunca y Chorotega, zonas donde este cultivo predomina en Costa Rica. Para ello, se emplearon 20 Secuencias de Simple Repetición (SSR) utilizados en estudios sobre origen, evolución y diversidad de maíz en América. Se analizaron muestras de maíz criollo provenientes de 20 poblaciones entre ambas regiones socioeconómicas. Se estandarizó el protocolo de PCR tipo Múltiplex y la obtención de genotipos utilizando la librería *MsatAllele* de R. Se determinó que los SSR utilizados son polimórficos ($PIC=0.84$) y eficaces para estimar la diversidad genética del germoplasma criollo. La diversidad genética del maíz de la región Chorotega ($H_e=0,82$) y Brunca ($H_e=0,80$) fue muy similar. No obstante, se encontraron 3.9 (± 0.75) alelos privados para Chorotega y 1.3 (± 0.32) para Brunca. Los resultados sugieren, que el maíz criollo costarricense es genéticamente diverso, que la diversidad es alta y que es comparable con la diversidad reportada para México, América del Sur y el Caribe.

Palabras clave: *Zea mays*, maíz criollo, microsatélites, variabilidad genética, germoplasma.

1. INTRODUCCIÓN

1.1. Antecedentes

1.1.1. Información general sobre la biología el maíz

El maíz, *Zea mays spp mays* L. ($2n=20$), pertenece a la familia Poaceae (Gramíneas), tribu Maydeas. El maíz es una planta anual, monoica (Paliwal *et al.* 2001) y generalmente protándrica, es decir que madura flores masculinas antes que las femeninas (OGTR, 2008). La espiga masculina es apical y se le llama panoja o racimo plurirrámico (Cortés, 1994). La espiga femenina o mazorca, tiene disposición axilar, consta de un eje grueso cilíndrico (coronta) y de espiguillas que se ubican de manera individual en cada una de las cavidades de la coronta formando hileras (Molina, 2011). La polinización es cruzada y se da principalmente por viento. El polen puede viajar hasta 40m, sin embargo, la mayor parte se asienta entre 6 y 15 m desde su origen (Hoffman *et al.*, 2014). El fruto es una cariopse o grano, mide de 5 a 25 mm de largo y desarrolla un endospermo que puede ser amiláceo, dextrinoso o ceroso (Paliwal *et al.* 2001). El color del grano es variable, y determinado por diferentes pigmentos o la falta de ellos como el caso de los granos blancos. Las antocianinas confieren coloraciones de grano azul, rojo, morado, negro y diferentes intensidades de estos colores (Salinas *et al.*, 2012). Los carotenoides producen coloraciones amarillas y rojas (Espinosa-Trujillo *et al.*, 2009). La mazorca se mantiene permanentemente encerrada dentro de muchas brácteas, por esta razón la planta no es capaz de dispersar sus semillas como una planta silvestre, si no que depende de la intervención humana para el desgrane de semillas y su propagación (OGTR, 2008). Lo anterior es consecuencia de los procesos de mejoramiento y selección de dicho cultivo (Doebley *et al.*, 2006). El maíz se siembra en una gran variedad de regiones agroecológicas que van de altitudes de 0 m snm

hasta cerca de los 4000 m snm, en regiones de precipitación pluvial desde menos de 400 mm hasta los 3000 mm, y en suelos y climas muy variables (Kato *et al.* 2009).

La clasificación del maíz es compleja. Las primeras clasificaciones se basan en la morfología del grano de maíz, principalmente por el color y la contextura (Acosta, 2009). Se ha propuesto incluir también caracteres genéticos, por lo que los miembros de una raza de maíz tienen semejanzas marcadas no sólo en el fenotipo y en la distribución geográfica, sino también en términos genéticos, citológicos, fisiológicos y características agronómicas (Vigorous *et al.*, 2008). Prasanna (2012) define las razas criollas como una población dinámica de una planta cultivada que tiene origen histórico, identidad distinta, carece de mejoramiento tecnológico de cultivos, que es genéticamente diversa, adaptada localmente y que se asocia con los sistemas de agricultura tradicional. Dentro de estas razas también existen distintas variedades. Todas las variedades de maíz siguen el mismo patrón de desarrollo, sin embargo, el tiempo y los intervalos entre los estadios, el número de hojas, el número y el tamaño de mazorcas y la morfología de los granos, puede variar entre diferentes estaciones, el momento y la localización del cultivo. (Fuentes, 2012).

1.1.2 Origen y evolución del maíz

Los parientes silvestres del maíz son llamados colectivamente “teocintes” y están representados por especies anuales y perennes (Bedoya & Chávez, 2010). Existen diferentes hipótesis sobre el posible origen del maíz cultivado, sin embargo, la más aceptada indica que un teocinte anual *Zea mays ssp. parviglumis*, fue el ancestro del maíz cultivado actual, producto de varios procesos de domesticación humana (Doebley, 2004;

Matsouka *et al.*, 2002; Kato *et al.*, 2009). La mayoría de los investigadores establecen que el centro de origen del maíz está en las zonas altas de México (Doebley, 1990; Matsouka *et al.*, 2002, Bedoya &Chávez, 2010), donde el maíz y el teocinte han coexistido desde la antigüedad y donde ambas especies presentan una diversidad genética y morfológica alta (Cortés, 1994; Paliwal, *et al.* 2001). El hallazgo de polen fósil y de mazorcas de maíz en zonas arqueológicas mexicanas apoyan la hipótesis de que éste cultivo se originó en México (Horn & Kennedy, 2001; Paliwal *et al.* 2001) y según esta evidencia, en conjunto con análisis filogenéticos, se recalca que la domesticación de este surgió hace 9000 años como producto de un único evento en el sur de México (Matsouka *et al.*, 2002). Además, análisis filogenéticos con microsatélites sugieren que ocurrió una diversificación temprana en las zonas altas del país (Matsouka *et al.*, 2002) y que desde allí se propagó hacia las zonas de menor altitud en México, hacia el norte de América y hacia el sur pasando por Guatemala, el Caribe, Centroamérica, y de las zonas bajas de Suramérica hasta los Andes (Matsouka *et al.*, 2002; Oliveira *et al.*, 2003; Vigoroux *et al.* 2008), sitios en los cuales el maíz se continúa diversificando. La propagación del maíz alrededor del mundo, principalmente a Europa y Asia se dio entonces luego del descubrimiento de América en el siglo XV, después de la conquista se hicieron introducciones desde los Estados Unidos hacia Sudamérica, especialmente a Chile, Argentina y Brasil (Sánchez, 2011).

1.1.3 Marcadores moleculares y diversidad genética del maíz

La caracterización de la diversidad genética del maíz criollo se ha basado fundamentalmente en características fenotípicas relacionadas con la mazorca y su uso (Becerra & Paredes, 2000; Sánchez *et al.*, 2000 Sharma *et al.*, 2010). Esto ha permitido

definir variedades de maíz. Sin embargo, las principales limitantes de esta caracterización son la influencia ambiental en la expresión de las características de la planta y el tiempo requerido para coleccionar los datos (Becerra & Paredes, 2000).

Actualmente, los marcadores moleculares se han utilizado para describir la diversidad y la estructura genética de poblaciones, por tener una alta confiabilidad, reproducibilidad y automatización (Jiménez, 2014). La ventaja sobre caracteres fenotípicos es que los marcadores moleculares dan información sobre la variación alélica en un determinado locus (Schlötterer, 2004), no están influenciados por el ambiente y se distribuyen a lo largo de todo el genoma (Acosta, 2009). La información que se obtiene de estos marcadores ha permitido caracterizar a nivel genético materiales mejorados, complejos raciales, variedades de polinización abierta y materiales criollos (Hartings *et al.*, 2008; Vigoroux *et al.*, 2008; Jiménez, 2014). Así como ha permitido estudios complejos sobre la historia evolutiva de los cultivos (De Barba *et al.*, 2017).

En maíz, se han utilizado, isoenzimas, polimorfismos de longitud de fragmentos de restricción (RFLPs), polimorfismos de un único nucleótido y secuencias de repetición simple (SSR) o microsatélites. Éstos últimos, son los más utilizados ya que, presentan altos niveles de polimorfismo, pueden ser aplicados en cualquier estado de desarrollo de una planta y son codominantes por lo que permiten distinguir entre heterocigotos y homocigotos (Jiménez, 2014). Sin embargo, presentan la desventaja de que la identificación de los alelos puede ser confusa por problemas de la técnica. Entre estos, las amplificaciones inespecíficas en la PCR y la migración diferencial de los productos en el analizador (De Barba *et al.*, 2017). Ambos errores pueden sobre estimar el número de alelos al clasificar el mismo alelo como dos diferentes. También se pueden dar fallos en la amplificación de uno

o ambos alelos de un locus (*allelic dropout*) reduciendo la heterocigosidad observada (Wang *et al.*, 2012).

La estimación de la diversidad genética en maíz criollo ha generado resultados variables en cuanto su magnitud. En estudios realizados en Norteamérica, Guatemala y Suramérica utilizando SSRs, se han encontrado niveles de heterocigosidad (H_e) altos que varían entre 0.58 (Bravo, 2009, Jiménez 2014) y 0.83 (Vigorous *et al* 2008). Además, la variación genética estimada en maíces criollos es consistentemente mayor dentro de poblaciones (intrapoblacional) que entre las poblaciones (interpoblacional) (Acosta *et al*, 2003; Pressoir & Berthaud, 2004; Cruz, 2012). La mayoría de los estudios concuerdan que el maíz criollo es una fuente importante de genes para programas de mejoramiento genético y que el conocimiento de la diversidad y la distribución de ésta, tanto dentro como entre razas criollas, puede proporcionar una guía para predecir el grado de herencia, variación, y el nivel de heterosis; todas esenciales para el mejoramiento genético del maíz y su conservación (Salami *et al.*, 2007; Yao *et al.*, 2007; Yao *et al.*, 2008; Bravo, 2009; Vigorous, *et al.*, 2008; Warbuton *et al.* 2008).

1.1.4. El maíz en Costa Rica

En Costa Rica se han encontrado semillas de maíz y olotes carbonizados en sedimentos de la Laguna del Arenal (Guanacaste) que datan aproximadamente de 2000 a. C., así como cerámicas que registran prácticas agrícolas de esa época (Bradley, 1994). También, se encontraron granos de polen en sedimentos de Laguna Martínez en Guanacaste (Aford & Horn, 2004) y en Laguna Zoncho en el Pacífico Sur (Clement & Horn, 2001) de aproximadamente 3000 A.C., dando evidencia de la agricultura milenaria del maíz en

nuestro país (Taylor *et al.* 2013). Se cree que la expansión histórica hacia Suramérica se dio por el intercambio de maíz entre diferentes pueblos de horticultores de Costa Rica y Panamá (Horn & Kennedy, 2001; Oliveira *et al.* 2003; Smalley & Blake, 2003).

En Costa Rica se cultiva maíz comercial (blanco y amarillo) y el maíz criollo (Bonilla, 2009). El maíz blanco constituye un ingrediente importante en la dieta, la principal industria de maíz blanco está en manos de una empresa dedicada a la comercialización, molienda y fabricación de tortillas (Delgado, 2012). El maíz amarillo es el principal ingrediente en la fabricación de concentrados para la alimentación animal (Cortés, 1994) y se dejó de sembrar masivamente desde 1993, ya que era más barato importarlo (Delgado, 2012).

En relación con las razas criollas, aún no se ha elaborado una clasificación clara en Costa Rica, pero, se identifican visualmente por la coloración del grano, la cual puede ser morada, color vino, rosada, amarilla o blanca. El maíz criollo lo cultivan y conservan pequeños agricultores para autoconsumo y preparación de comidas tradicionales. Se ha documentado que la mayoría de los maíces criollos que se encuentran en las tierras altas del centro de Costa Rica y en la Vertiente Atlántica, son más afines a los maíces suramericanos (CNRF, 1996). Mientras que Guanacaste recibe la influencia mesoamericana, que se muestra en ciertas variedades de granos rojos y morados (CNRF, 1996). Además, un estudio molecular reciente, que incluyen una muestra de maíz criollo lo relacionan con otros maíces de tierras bajas mesoamericanas del centro de México y del sur de los Andes (Bedoya *et al.*, 2017).

Las regiones socioeconómicas costarricenses con mayor producción de maíz son la Chorotega, la Huetar Norte y la Brunca. La Región Chorotega, consta con un 34% del total

del área nacional de siembra (Bonilla, 2009). Además, el 25 de julio de 2014 las variedades de maíz criollo que se cultivan en la provincia de Guanacaste fueron declaradas patrimonio cultural de la república, incluyendo las tradiciones, prácticas agro-culturales, usos, saberes, sabores y colores asociados con este grano (ACAN-EFE, 2014; MCJ, 2014). La Región Brunca es la tercera región con mayor área nacional de siembra de maíz, un 13% (Bonilla, 2009). Dentro de esta región existen reservas indígenas que cultivan maíz criollo como su principal alimento para la subsistencia (Alvarado, 2003). Sin embargo, la producción nacional de maíz, en general, ha disminuido en los últimos años y sólo un 7% de los campos cultivados utilizan semillas de materiales criollos (IICA, 2014). Esta situación puede deber al cambio en el uso de la tierra, incluida la urbanización y la sustitución por variedades mejoradas de maíz que, aunque son más productivas, a la vez demandan más insumos agrícolas (Serratos-Hernández, 2009). Así como el cambio climático que incide negativamente en la producción de cosechas, el abastecimiento de granos de maíz criollo y pone en peligro la pérdida de este valioso recurso (Fuentes, 2012; Rivas *et al.*, 2013).

1.2. Justificación

El maíz criollo es un recurso fitogenético importante por su resistencia a condiciones climáticas específicas y a enfermedades y plagas (Carvahlo *et al.*, 2008; Pineda-Hidalgo *et al.*, 2013). Además, este cultivo, es aprovechado por los agricultores como fuente de alimento y como moneda de cambio por otros productos que no producen ellos mismos. A su vez, los productores mantienen la riqueza genética y preservan la identidad de cada región a través de prácticas tradicionales de cultivo (Rivas *et al.*, 2013). Debido a esta importancia sociocultural, económica y biológica; el maíz criollo es

considerado un recurso de interés para la conservación. Como paso inicial para desarrollar estrategias de conservación de este recurso, se requiere de un análisis de la diversidad genética. Específicamente se necesita entender su distribución dentro y entre poblaciones. Esto permitiría determinar si existen variedades o líneas de maíz criollo genéticamente independientes, y la correcta identificación de material evita su replicación en los bancos de germoplasma (Piñero *et al.*, 2008). Además, aquellos materiales más diversos serán de mayor importancia para su conservación y posible uso para mejoramiento, por el potencial adaptativo que podrían presentar.

La diversidad genética del maíz criollo costarricense no ha sido estudiada ampliamente. Solo en un estudio se ha incluido una población de maíz criollo de Costa Rica, que provenía de un banco de germoplasma internacional. En este estudio encontraron que se agrupa con maíces de México y los Andes (Bedoya *et al.*, 2017). Lo que propone a Costa Rica como un sitio interesante de estudio para complementar las hipótesis de dispersión del maíz en el continente americano (Vigorous *et al.*, 2008). Por lo tanto, caracterizar la diversidad genética aportaría información valiosa acerca del germoplasma de maíz criollo costarricense para un futuro plan de manejo y de conservación de este, así como contribuir con el vacío de información de este cultivo en Costa Rica. Para esta estimación, las regiones costarricenses productoras de maíz, la Brunca y la Chorotega, se destacan, por cultivo tradicional a manos de pequeños agricultores e indígenas y por la variabilidad morfológica presente.

Los microsatélites por sus características ya mencionadas son los predilectos para el estudio de la diversidad genética. Sin embargo, la correcta asignación e identificación de genotipos es importante para evitar sesgos en estimación de la diversidad. Por ejemplo, el

allelic dropout podría mostrar una reducida heterocigosidad observada y un aumento en la endogamia. Además, la asignación de genotipos debe calibrarse en cada investigación pues depende mucho de la técnica (PCR) y de los equipos utilizados. Lo anterior para asegurar la reproducibilidad de los datos e incluso favorecer el intercambio y el uso de bases de datos de SSR para estudios a largo plazo o a escala internacional de una manera confiable (Ellis *et al.*, 2007). Por lo que, para estimar la diversidad genética del maíz criollo costarricense, es importante automatizar la identificación de genotipos.

1.3 Marco Teórico

La diversidad genética en sentido amplio se define como las variaciones heredables que ocurren en cada organismo, entre los individuos de una población y entre las poblaciones dentro de una especie como producto de los procesos evolutivos que operan sobre esas variaciones (Piñero *et al.*, 2008). Entre estos procesos se encuentran las mutaciones, la selección natural y la artificial, y la deriva genética que pueden todas generar/modificar la variación (Hilje, 2008). La diversidad genética dentro de las especies permite la posibilidad de evolucionar bajo condiciones cambiantes del ambiente y presiones de selección; así mismo, el conocimiento de la diversidad genética es indispensable para diversificar las fuentes de germoplasma, tratar de minimizar los riesgos de vulnerabilidad genética e incrementar las probabilidades de detectar alelos favorables (Sánchez, 2011). Esto debido a que la caracterización de variedades de cultivos con mucha diversidad genética o con variación única, por ejemplo, la presencia de alelos raros es útil para establecer aquellas poblaciones que tienen mayor potencial adaptativo y con características agronómicas importantes (Vigoroux *et al.*, 2008; Bellon *et al.*, 2011)

Cuando existen diferencias genéticas entre subpoblaciones, poblaciones y dentro de cada una de éstas, se dice que las poblaciones están estructuradas (hay estructura genética espacial). Como consecuencia se pueden generar líneas genéticas independientes dentro de una misma especie, agrupaciones aisladas, entre otras (Jiménez, 2014). La estructura genética de una población también se ve afectada por fuerzas evolutivas y barreras geográficas que afectan el flujo de génico (Ochoa, 2013). Dentro de ellas se encuentran la selección natural y la deriva genética que favorecen la diferenciación y, por otro lado, el flujo de genes que permite el intercambio de material genético entre poblaciones y que puede llegar a minimizar la diversidad genética entre ellas (Hilje, 2008).

La caracterización molecular es uno de los métodos más apropiados para llevar a cabo los estudios de identificación de genotipos y similitud genética. Existen diversos tipos de marcadores moleculares, entre éstos los microsatélites o secuencias de simple repetición (SSR). Ellos son secuencias de dos o más pares de bases repetidas en tándem. Este marcador involucra la identificación de fragmentos específicos de ADN a través de imprimadores, la amplificación de los SSR por medio de una reacción en cadena de la polimerasa (PCR) y la visualización o la obtención de la medida de estas repeticiones en pares de bases para cada una de las regiones, es decir, obtener alelos. La variabilidad en el número de estas secuencias repetidas entre individuos se interpreta como diferencias genéticas entre ellos (Bedoya, 2012). Los SSR permiten el cálculo de frecuencias alélicas en las poblaciones (Bedoya, 2012). De esta manera se pueden estimar las distancias genéticas entre las poblaciones o entre individuos, así como realizar análisis filogenéticos (Matsouka *et al.*, 2002) y de estructura poblacional (Ochoa, 2013).

Durante la expansión desde México hacia el resto de América, el maíz logró adaptarse a diversas condiciones climáticas y edáficas, relacionadas a distintos pisos altitudinales (Vigouroux *et al.*, 2008), originando variedades adaptadas a diferentes condiciones de estrés biótico (Kato *et al.*, 2009). El maíz criollo ha sido cultivado generación tras generación por agricultores locales, que han mantenido su diversidad y han afectado su estructura genética, a través de procesos de manejo de semilla. Uno de éstos es la selección de la variedad por la cual, los productores seleccionan mazorcas con las “mejores” características para la siembra y que generen mejor cosecha; además eliminan aquellas mazorcas cuyos granos muestren mezcla o contaminación de otras variedades o razas. Esta práctica aumenta la similitud dentro de las variedades. Por otro lado, el intercambio de semilla entre los agricultores fomenta el flujo génico, y por lo tanto se tiende a disminuir la estructura genética (Pressoir & Berthaud 2004).

El maíz criollo es considerado un recurso fitogenético importante. En países como México y Brasil, el maíz criollo ha demostrado una alta resistencia a enfermedades y plagas, tolerancia a sequía, buen desempeño en suelos con pocos nutrientes (Carvahlo *et al.*, 2008; Pineda-Hidalgo *et al.*, 2013). También, a diferencia de variedades comerciales, tienen un mejor sabor y bajo condiciones de almacenamiento puede conservarse por más tiempo; características muy apreciadas por los agricultores (Laborda *et al.*, 2005; Carvahlo *et al.*, 2008; Pineda-Hidalgo *et al.*, 2013; Kato *et al.*, 2009). Además, se ha propuesto al maíz criollo como un cultivo idóneo para afrontar el cambio climático progresivo que está ocurriendo en la actualidad (Hellin *et al.*, 2014). Principalmente, porque, a diferencia de maíces híbridos, el maíz criollo presenta una gran diversidad genética lo que le da un potencial para adaptarse a condiciones agroecológicas variables (Bellon *et al.*, 2011; Hellin *et al.*, 2014).

1.3.1 Hipótesis y predicciones

El maíz criollo es una especie alógama, de amplia distribución y adaptada a una variedad de condiciones agroecológicas, por lo que se espera que en Costa Rica se detecten altos niveles de diversidad genética; comparables con otros maíces criollos en Norte, Sur y Mesoamérica. También que la diversidad estimada para cada SSR sea alta y que los valores de las estimativas sean mayores para los marcadores con menor motivo de repetición, pues estos tienen una mayor tasa de mutación.

Se espera encontrar que la diversidad genética en las regiones Brunca y Chorotega, sea alta y similar a la reportada para otros países americanos (Vigorous *et al.*, 2008). Además, por el manejo y el intercambio de semillas de maíz entre los productores, se espera que exista poca estructura entre las poblaciones y entre ambas regiones socioeconómicas.

II OBJETIVOS

2.1: Objetivo general

Caracterizar la diversidad genética de maíz criollo (*Zea mays*) de las regiones Brunca y Chorotega de Costa Rica mediante marcadores SSRs.

2.2: Objetivos específicos

- Estandarizar el protocolo de amplificación de 20 microsatélites de maíz en muestras de Costa Rica.
- Estimar la diversidad genética del maíz criollo de dos regiones socioeconómicas de Costa Rica utilizando microsatélites.

III METODOLOGÍA

3.1. Proyecto maíz criollo.

El presente estudio se realizó dentro del marco del proyecto colaborativo entre la Universidad de Costa Rica (UCR), la Universidad Nacional (UNA) y el Instituto Tecnológico (ITCR) “Estudio y conservación de la diversidad genética del maíz criollo en dos regiones de Costa Rica” VI 801-B4-651 financiado por FEES-CONARE. La investigación se realizó en el laboratorio del Programa de Mejoramiento Genético de Cultivos del Centro de Investigación en Biología Celular y Molecular y en la Escuela de Biología. La recolecta de las muestras se llevó a cabo por investigadores de la UNA y la UCR en la región Chorotega y del ITCR en Brunca.

3.2 Sitios de estudio.

Para el estudio se utilizaron muestras de maíz criollo provenientes de las regiones socioeconómicas costarricenses: Chorotega (noroeste del país) y Brunca (Pacífico Sur). La Región Chorotega limita al norte con Nicaragua, Upala y con Guatuso de Alajuela, al este con San Carlos, San Ramón de Alajuela y Puntarenas, al sur con el Golfo de Nicoya y al oeste con el Océano Pacífico. Esta región tiene un relieve variado y el clima es distinto de acuerdo con éste: en las faldas occidentales de la cordillera de Guanacaste es templado, con lluvias abundantes y en las llanuras del río Tempisque se registran altas temperaturas hasta de 35°C según la época del año. (Alvarado, 2003).

La Región Brunca limita al norte con Paraíso y Turrialba de Cartago y con Talamanca de Limón, al este con Panamá, al sur con Golfo Dulce y al oeste con Dota (San José) y Aguirre de Puntarenas (Alvarado, 2003). Presenta grandes contrastes de relieve desde la Cordillera de Talamanca hasta la planicie costera. El clima se ve influenciado por

la acción de los vientos húmedos del Pacífico que provocan abundantes lluvias y por la latitud que provoca altas temperaturas que, van disminuyendo conforme se asciende por la montaña hasta llegar al páramo (Alvarado, 2003). Además, dada la gran variedad de tipos de relieve como de altura y pluviometría, la región es la más rica en zonas de vida (CNRF, 1996).

3.3 Muestras de maíz

Las muestras de maíz criollo corresponden a variedades, no registradas en la oficina nacional de semilla (ONS) ni conservadas en bancos de germoplasma, usualmente cultivadas por agricultores de manera tradicional y para consumo local. Por lo tanto, los sitios de muestreo dentro de cada región consisten en pequeñas fincas o parcelas de agricultores dedicadas al cultivo de maíz criollo, donde usualmente sólo se planta una variedad por parcela. Estas parcelas se definirán como poblaciones para efectos de este estudio (Fig.1). Estas fueron recolectadas en distintos sitios dentro de cada región, denominados como “sitios de recolecta”. Estos sitios fueron definidos según la localidad (Cantón de la provincia) y la cercanía entre las poblaciones en estudio. Para la región Chorotega los sitios de recolecta corresponden a Cañas, Hojancha, Liberia, Santa Cruz, La Cruz Las Brisas y La Cruz Santa Cecilia. Para Brunca los sitios de recolecta son Valencia, Concepción, Boruca y Laurel (Fig.1). Se analizaron 12 poblaciones de maíz criollo para la Región Chorotega y 8 para la Brunca. La disponibilidad del material por cada región y el costo del análisis molecular determinaron el tamaño de muestra. Por cada población se seleccionaron tres mazorcas, al azar, y cuatro granos por cada mazorca. El fin de la selección de los cuatro granos por mazorca es para facilitar la asignación de genotipos asegurándose que, al menos un alelo se comparta entre hermanos y la madre. Lo anterior

permite identificar correctamente los alelos y detectar errores de amplificación o *allelic dropout* pues, al ser el primer estudio en maíz criollo de Costa Rica, no existía información disponible sobre los alelos que podrían encontrarse en estos materiales en el país, la cual si existe para otras localidades del continente americano.

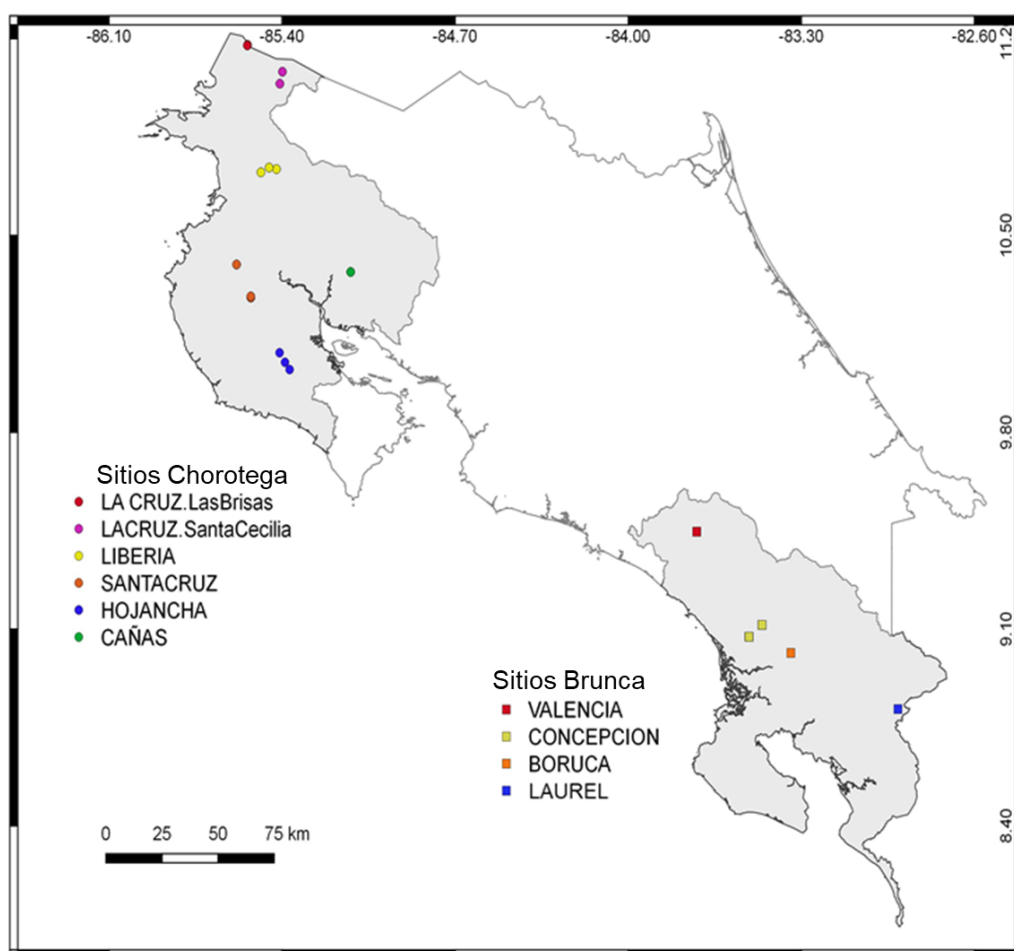


Figura 1: Mapa de Costa Rica con las regiones Chorotega (al norte) y Brunca (al sur) resaltadas en gris. En las regiones se muestran utilizando símbolos de color los sitios de colecta. Dentro de la Región Chorotega cada círculo corresponde a una población en cada sitio de colecta; en la Región Brunca, las accesiones se representan con cuadrados.

3.4 Siembra y extracción

La extracción del ADN se llevó a cabo por medio del método CTAB (Doyle & Doyle, 1990) modificado y estandarizado para maíz (Anexo 1). Se utilizaron tres distintos tipos de material vegetal:

i- Foliar:

De cada mazorca se sembraron cuatro semillas, escogidas al azar, en el invernadero del Centro de Investigación de Biología Celular y Molecular (CIBCM). Cuando la planta mostró las dos primeras hojas, aproximadamente 15 días después de la siembra, se extrajo del tejido foliar. Sin embargo, el porcentaje de germinación era muy bajo para la cantidad de muestras que se necesitaban. Por lo que se recurrió a otros tipos de material vegetal.

ii- Grano completo:

De cada mazorca se escogieron cuatro semillas al azar. Se maceraron con el molino mezclador MM400 (Retsch) a 30 frecuencias por 10 minutos. Se utilizaron aproximadamente 0,5g de masa y se extrajo ADN. Sin embargo, las muestras presentaron una masa blanca al enfriarse, que corresponde a almidón, el cual no se logró eliminar por completo.

iii-Embrión:

De cada mazorca se escogieron cuatro semillas al azar y se pusieron en remojo, con agua destilada, un día antes de la extracción. De cada semilla se extrajo el embrión con estereoscopio y navajilla. Cada embrión se cortó y luego se maceró con el molino mezclador MM400 (Retsch) a 30 frecuencias por 10 minutos y se extrajo el ADN con el método CTAB.

Todas las muestras de ADN se cuantificaron para medir su calidad y concentración por medio de espectrofotometría con NanoDrop (Thermo Scientific). Las muestras se identificaron de acuerdo con la región Brunca (B) o Chorotega (C), la universidad que realizó la colecta: UCR (C), UNA (U), TEC (E); el número de accesoión y un número consecutivo único. Todas las muestras se almacenaron a -20°C.

3.5 Microsatélites

3.5.1 Selección de marcadores moleculares

Se seleccionaron en total 20 SSR, diez con motivos de repetición de dos nucleótidos y diez de cuatro nucleótidos, que han sido utilizados en estudios sobre diversidad genética y origen del maíz a nivel Norte y suramericano (Cuadros 1 y 2; Matsuoka *et al*, 2002; Vigoroux *et al.*, 2008). Se escogieron aquellos marcadores cuyas posiciones lograran cubrir los 10 cromosomas de maíz. Además, donde se reportaron valores altos de Contenido de Información Polimórfica (PIC), y que las temperaturas de hibridación (T_a) fuesen similares para facilitar la amplificación en reacciones de múltiplex. Las secuencias “Forward” de los 20 SSR se marcaron con los fluorocromos 6-FAM, VIC, PET y NED (Romero-Severson *et al.*, 2001).

Cuadro 1: Información sobre los 10 SSR dinucleotídicos utilizados para caracterizar la diversidad genética de maíz criollo de las regiones Brunca y Chorotega de Costa Rica. Se muestra el marcador SSR, el Contenido de Información Polimórfica (PIC), el motivo de repetición (Motif), los imprimadores Forward (F-) y Reverse (R-), cromosoma al que pertenecen, el Rango alélico y la referencia de donde se tomaron los SSR y su información respectiva.

Marcador	PIC	Motif	Imprimadores	Cromosoma	Rango	Referencia
<i>bnlg1890</i>	0.93	AG	F- ACCGGAACAGACGAGCTCTA R- GTCCTGCAAAGCAACCTAGC	4	102-238	Vigorous <i>et al.</i> 2008
<i>bnlg1523</i>	0.91	AG	F- GAGCACAGCTAGGCAAAAGG R- CTCGCACGCTCTCTCTTCTT	3	172-277	Vigorous <i>et al.</i> 2008
<i>bnlg1740</i>	0.93	AG	F- TTTTCTCCTTGAGTTCGTTCG R- ACAGGCAGAGCTCTCACACA	6	115-222	Vigorous <i>et al.</i> 2008
<i>bnlg1917</i>	0.92	AG	F- ACCGGAACAGACGAGCTCTA R- TTTGCTTCCAACCTCACATGC	6	76-196	Vigorous <i>et al.</i> 2008
<i>bnlg1449</i>	0.92	AG	F- AGTCAACGTAGCTGGCGAGT R- TTCACGACGGGTCTCTCTCT	3	95-208	Vigorous <i>et al.</i> 2008
<i>bnlg1191</i>	0.97	AG	F- AATCATGCGTAGGCGTAGCT R- GCCAGAGGAAAAAGAAGGCT	10	173-251	Vigorous <i>et al.</i> 2008
<i>bnlg1265</i>	0.95	AG	F- GGTGTCCGTAAAGGCAAGA R- TGTGAAGGCCAGACAGTCAG	2	193-264	Vigorous <i>et al.</i> 2008
<i>bnlg2305</i>	0.93	AG	F- CACCTTGAAAGCATCCTCGT R- GTATCACACCCTCTGCTGCA	5	159-243	Vigorous <i>et al.</i> 2008
<i>bnlg1046</i>	0.96	AG	F- TGAGCCGAAGCTAACCTCTC R- GATGCAAAGGAGGTTTCAGGA	5	170-259	Vigorous <i>et al.</i> 2008
<i>bnlg1194</i>	0.96	AG	F- GCGTTATTAAGGCAAGCTGC R- ACGTGAAGCAGAGGATCCAT	8	120-234	Vigorous <i>et al.</i> 2008

Cuadro 2: Información sobre los 10 SSR tetranucleotídicos utilizados para caracterizar la diversidad genética de maíz criollo de las regiones Brunca y Chorotega de Costa Rica. Se muestra el marcador SSR, el Contenido de Información Polimórfica (PIC), el motivo de repetición (Motif), los imprimadores Forward (F-) y Reverse (R-), cromosoma al que pertenecen, el Rango alélico y la referencia de donde se tomaron los SSR y su información respectiva. Los rangos alélicos no reportados se muestran como NR.

Marcador	PIC	Motif	Imprimadores	Cromosoma	Rango	Referencia
<i>phi078</i>	0.68	AAAG	F- CAGCACCAGACTACATGACGTGTAA R- GGGCCGCGAGTGATGTGAGT	6	128-208	Warburton 2002, Vigorous <i>et al.</i> 2008, Bravo 2009
<i>phi064</i>	0.79	ATCC	F- CCGAATTGAAATAGCTGCGAGAACCT R- CAATGAACGGTGGTTATCAACACGC	1	68-128	Matsuoka <i>et al.</i> 2002, Warburton 2002, Vigorous <i>et al.</i> 2008
<i>phi116</i>	0.82	ACTG	F- GCATACGGCCATGGATGGGA R- TCCCTGCCGGGACTCCTG	8	149-183	Warburton 2002, Vigorous <i>et al.</i> 2008
<i>phi072</i>	0.69	AAAC	F- ACCGTGCATGATTAATTTCTCCAGCCTT R- GACAGCGCGCAAATGGATTGAACT	7	117-173	Matsuoka <i>et al.</i> 2002, Warburton 2002, Vigorous <i>et al.</i> 2008
<i>phi093</i>	0.75	AGCT	F- AGTGCGTCAGCTTCATCGCCTACAAG R- AGGCCATGCATGCTTGCAACAATGGATACA	4	284-307	Matsuoka <i>et al.</i> 2002, Warburton 2002, Vigorous <i>et al.</i> 2008
<i>phi96100</i>	0.67	ACCT	F- AGGAGGACCCCAACTCCTG R- TTGCACGAGCCATCGTAT	2	NR	Matsuoka <i>et al.</i> 2002, Warburton 2002
<i>phi015</i>	0.68	AAAC	F- ACGCTGCATTCAATTACCGGGAAG R- GCAACGTACCGTACCTTTCCGA	8	NR	Matsuoka <i>et al.</i> 2002, Warburton 2002
<i>phi089</i>	0.53	ATGC	F- GAATTGGGAACCAGACCACCCAA R- ATTTCCATGGACCATGCCTCGTG	6	89-104	Warburton 2002, Vigorous <i>et al.</i> 2008
<i>phi031</i>	0.71	GTAC	F- GCAACAGGTTACATGAGCTGACGA R- CCAGCGTGCTGTTCCAGTAGTT	9	186-240	Warburton 2002, Vigorous <i>et al.</i> 2008.
<i>phi109188</i>	0.74	AAAG	F- AAGCTCAGAAGCCGGAGC R- GGTCATCAAGCTCTCTGATCG		NR	Matsuoka <i>et al.</i> 2002, Warburton 2002

3.5.2 Selección de los múltiplex

Los microsatélites se amplificaron por el método de PCR tipo múltiplex. Para esto se conformaron cinco múltiplex diferentes, compuestos de cuatro SSR cada uno. Se utilizaron dos criterios para seleccionar los cuatro imprimadores. El primero fue que los imprimadores estuvieran marcados con un fluorocromo distinto. El segundo criterio consistió en que los rangos alélicos reportados para los marcadores no tuvieran tamaños similares (Cuadros 1 y 2). El Anexo 2 muestra los cinco múltiplex conformados.

3.5.3 Estandarización de la reacción de PCR en multiplex:

La reacción (Anexo 3) y el protocolo de amplificación (Cuadro 3) que se utilizaron como punto de partida fueron los recomendados por el protocolo de Multiplex de Qiagen (2010). Éstos se modificaron para optimizarlos, probando distintas concentraciones de ADN, temperaturas de hibridación y número de ciclos.

Cuadro 3: Protocolo de amplificación base (Qiagen, 2010) utilizado para el PCR multiplex de los 20 SSR seleccionados para el estudio de diversidad genética de las regiones Brunca y Chorotega de Costa Rica.

Condiciones de ciclos	Tiempo	Temperatura
Activación inicial	15 min	95°C
Número ciclos 30-45		
Desnaturalización	30 s	94°C
Annealing	90s	57-63 °C
Extensión	90 s	72 °C
Extensión final	10min	72°C

Para las pruebas se seleccionaron cuatro muestras de ADN, tres de la región Chorotega y una de la Brunca. La primera prueba consistió en $Ta=55^{\circ}\text{C}$, 30ciclos, una concentración de ADN de 50ng/uL y con un volumen final de reacción de 15uL. Con el fin de reducir el uso de reactivos, se realizó la misma prueba con las mismas condiciones mencionadas, pero con un volumen final de 10uL. Como siguiente paso se modificó la temperatura de hibridación o de annealing (Ta). Se probaron Ta de 57°C , 60°C y 63°C con 30 ciclos. Además, se compararon 25 ciclos con $Ta= 60^{\circ}\text{C}$ y 25 ciclos con tiempos de extensión 60s y una extensión final de 30 min. Finalmente se probó reducir la concentración del ADN de 50 ng/uL a 10ng/uL realizando alícuotas. El Cuadro 4 muestra resumen de las pruebas realizadas.

Cuadro 4: Resumen de experimentos de estandarización de condiciones de PCR multiplex para 20 SSR seleccionados para el estudio de diversidad genética de las regiones Brunca y Chorotega de Costa Rica. Ta = Temperatura de annealing (hibridación) en $^{\circ}\text{C}$, vf= volumen final en uL.

Experimento	Condiciones de PCR
1	Ta 55°C , 30ciclos vf 15
2	Ta 55°C , 30ciclos, vf 10
3	Ta 57°C y 60°C con 30ciclos vf 10
4	Ta 60°C , 25 ciclos y Ta 60°C 25ciclos, temperatura extensión 60s y 30 min
5	Ta 60°C , 27 ciclos y Ta 63°C , 30ciclos
6	Ta 60°C . 30ciclos (se hizo con pipeta multicanal)
7	Touch Down 68-60, bajando un grado cada ciclo y Ta 60°C ADN 1/10 diluido
8	Ta 60°C , 30ciclos y ADN diluído a 10ng/uL (prueba final)

3.5.4- Caracterización de los genotipos:

Los SSR se analizaron por medio de una electroforesis capilar en un secuenciador automático Genetic Analyzer 3130 (Applied Biosystems). Los genotipos se determinaron mediante el análisis de electroferogramas con el programa GeneMarker V1.97 (SoftGenetics). Este software hace estimaciones de los tamaños de los fragmentos amplificados por PCR. Los tamaños de los fragmentos (alelos) en los electroferogramas se identificaron, visualmente, como señales bien definidas (una sola señal sin amplificaciones inespecíficas) con una alta fluorescencia. Sin embargo, estas estimaciones variaron en fracciones de una base (1pb) para el mismo alelo (ej: 124.2; 124.3; 124.4). Lo anterior no tiene interpretación biológica, y confundió la asignación de genotipos ya que el mismo alelo puede ser interpretado como dos fragmentos de diferente tamaño. Por lo tanto, para incrementar la precisión y repetibilidad en la asignación de genotipos, se utilizó la librería *MsatAllele* (Alberto, 2014) de R (R Core Team, 2013). *MsatAllele* utiliza la distribución de los alelos observados y codifica los tamaños de los fragmentos reportados por el secuenciador con decimales (alelos crudos) en alelos reales (codificados como números enteros). Esta codificación la realiza filtrando los datos, de manera que los fragmentos observados que sean similares (± 0.8 de desviación estándar) los acomoda juntos, generando una distribución de fragmentos (Alberto, 2014). El valor del alelo entonces se define como la mediana de todos los fragmentos de dicha distribución, redondeada a un número entero (Fig. 2).

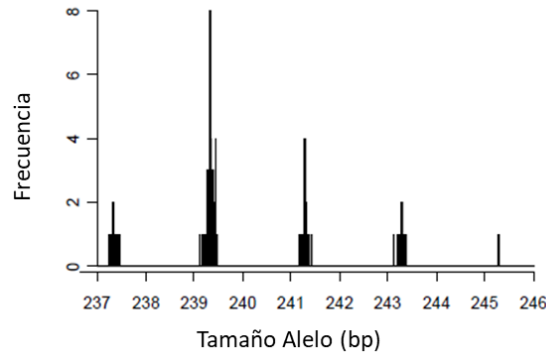


Figura 2: Ejemplo de distribuciones de tamaños de fragmento utilizados por *MsatAllele* para definir probabilísticamente un alelo real. En el eje y se observan las frecuencias de los individuos, en el eje x el tamaño de los fragmentos. Cada distribución de fragmentos (barras negras) conforma un *bin* un solo alelo. Tomado de Alberto (2014).

En el proceso de genotipo se generan *bins* (en español, contenedor), que son un rango de tamaños de fragmento donde se ubica, probablemente, un alelo. *MsatAllele* permitió definir para cada SSR los *bins*, el rango de tamaños de fragmento que conforman cada *bin* y el número de individuos que se clasificaron en cada *bin*. Con ayuda de los histogramas, que muestran la distribución de alelos para cada SSR, se corroboraron manualmente los alelos separados por 1pb y aquellos *bins* que presentarán una cantidad de individuos menor a 4. En ambos casos se revisaron los rangos alélicos por *bin* y el electroferograma de las muestras sospechosas. Lo anterior para corroborar si el alelo en cuestión se debería tomar en cuenta o no, pues podría ser producto de errores de genotipo humano. La decisión de eliminar el alelo se basó en la intensidad de la señal, la cual debía ser superior a 800 RFU (Unidades Relativas de Fluorescencia) y en la comparación con los medios hermanos y con el resto de los individuos para el mismo marcador. Con esto último, se determinaba con mayor certeza si el alelo en cuestión se trataba de una señal inespecífica

o no. Esto permitió reducir y definir los *bins* para las muestras de las regiones Brunca y Chorotega.

3.6 Análisis estadístico de los datos

Para cada microsatélite se calculó el Contenido de Información Polimórfica (PIC), para indicar cuan informativo es el SSR con base en las frecuencias alélicas de cada *locus*. El PIC toma valores entre 0 (no informativo) y 1 (informativo). La estimación de la diversidad genética del maíz criollo se realizó a través del cálculo del número de alelos (N_a), número de alelos efectivos (N_e), alelos privados, heterocigosidad esperada (H_e) y observada (H_o), y el índice de endogamia (F). Estas estimativas se obtuvieron para cada SSR, para cada región socioeconómica, para cada sitio de muestreo y en promedio para Costa Rica. Además, se calculó el G_{st} de Hedrick (Hedrick, 1999) como estimativa de la diferenciación genética entre las dos regiones. Las estimativas de diversidad genética se obtuvieron con el programa GenAlEx 6.5 (Peakall & Smouse, 2012). El porcentaje de valores perdidos por microsatélite y el G_{st} se obtuvieron con la librería *poppr* de R (R Core Team, 2013). Las estimativas de diversidad genética se calcularon antes y después de la corroboración manual de los *bins*, para analizar el efecto que tiene la optimización del genotipo sobre las estimativas de diversidad. Se utilizó la prueba no paramétrica de Wilcoxon de la librería básica de R para comparar las estimativas antes y después de la corroboración.

IV RESULTADOS

4.1- Estandarización de Microsatélites para maíz criollo de Costa Rica

4.1.2: Estandarización de la reacción y amplificación de PCR tipo Múltiplex:

Con respecto a la extracción de ADN, el material vegetal que presentó mayor concentración de ADN y calidad fue el foliar, seguido del embrión y grano completo. El uso de los tres distintos tejidos no afectó la obtención de los genotipos.

Con la primera prueba de estandarización utilizando una $Ta=55^{\circ}\text{C}$, 30 ciclos, una concentración de 50ng/uL de ADN y un volumen final de reacción de 15uL; se observaron muchas amplificaciones inespecíficas (Fig.3). Con las mismas condiciones, pero con un volumen final de 10uL, se obtuvieron resultados similares a 15uL (Fig. 4), por lo que el volumen final que se escogió para las reacciones de multiplex fue de 10uL. Entre las temperaturas de Ta que se probaron (57°C , 60°C y 63°C), se notaron diferencias marcadas. En las mismas muestras, a 57°C aparecen muchos picos inespecíficos (Fig. 5) que no aparecen a 60°C , mientras que a 63°C la mayoría de las muestras no amplificaron. Por lo que se fijó para los siguientes experimentos una Ta de 60°C . Sin embargo, se continuaban presentando amplificaciones inespecíficas y señales alélicas un poco bajas para algunos SSR.

Al comparar programas de PCR de 25 ciclos, uno con $Ta= 60^{\circ}\text{C}$ y otro con tiempos extensión de 60s y una extensión final de 30 min, no se encontraron diferencias en los electroferogramas. Cuando se aumentaron los ciclos a 27 y a 30 ciclos, los electroferogramas mostraron mejoría al poder visualizar patrones más interpretables, especialmente a 30 ciclos (Fig. 6). Finalmente, al reducir la concentración del ADN de 50

ng/uL a 10ng/uL se logró una mejoría en la identificación de los alelos. Por lo tanto, como condiciones finales se utilizaron: 30ciclos, T_a 60°C y una concentración de ADN de 10ng/uL con un volumen final de 10uL por reacción.

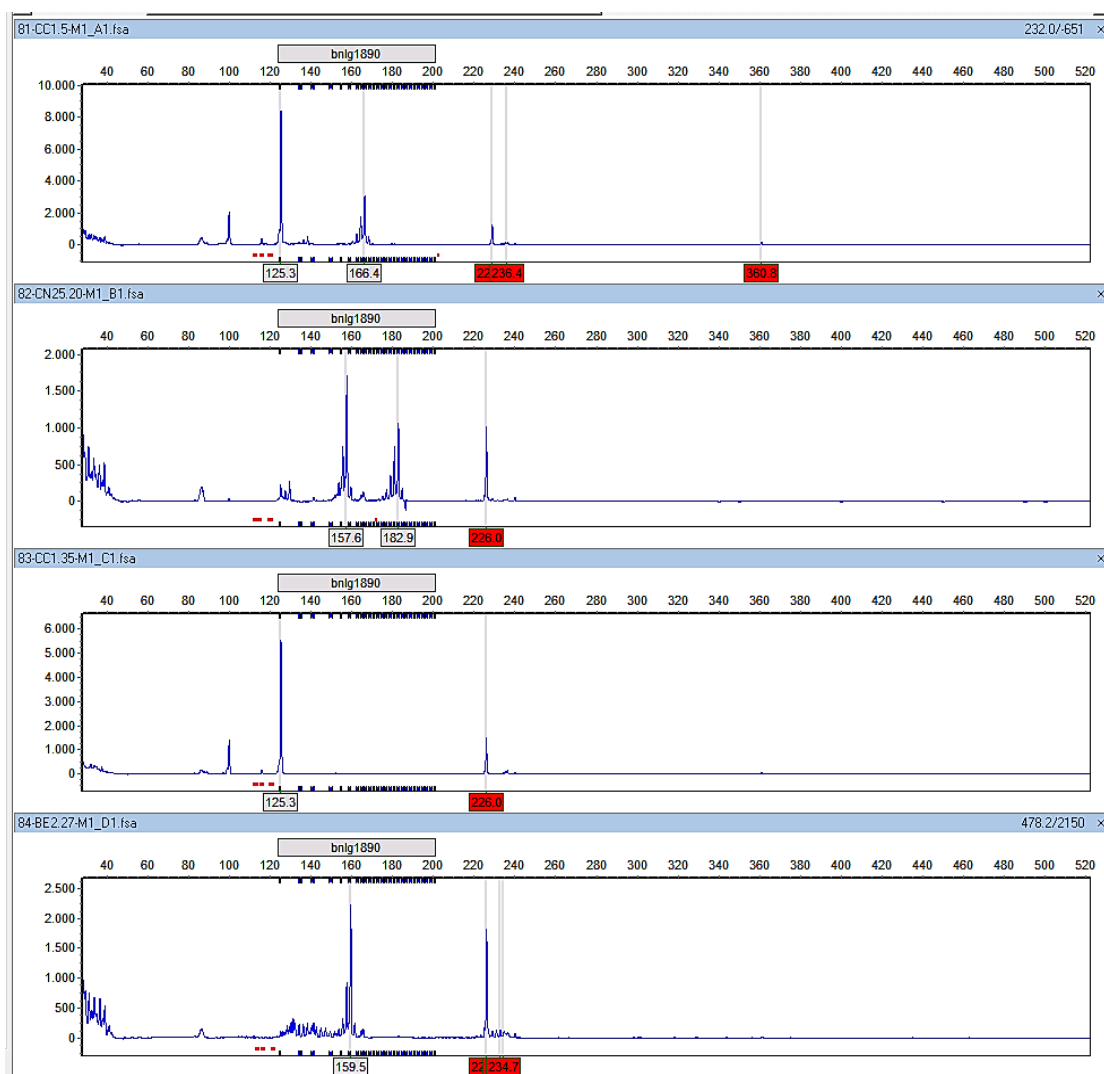


Figura 3: Electroferogramas visualizados en GeneMarker, de cuatro muestras de maíz criollo para el marcador bnlg1890, según la prueba de $T_a=55^{\circ}\text{C}$, 30ciclos, una concentración de ADN de 50ng/uL y con un volumen final de reacción de 15 uL. Se denotan muchas amplificaciones inespecíficas lo que confunde la interpretación de los alelos en todas las muestras.

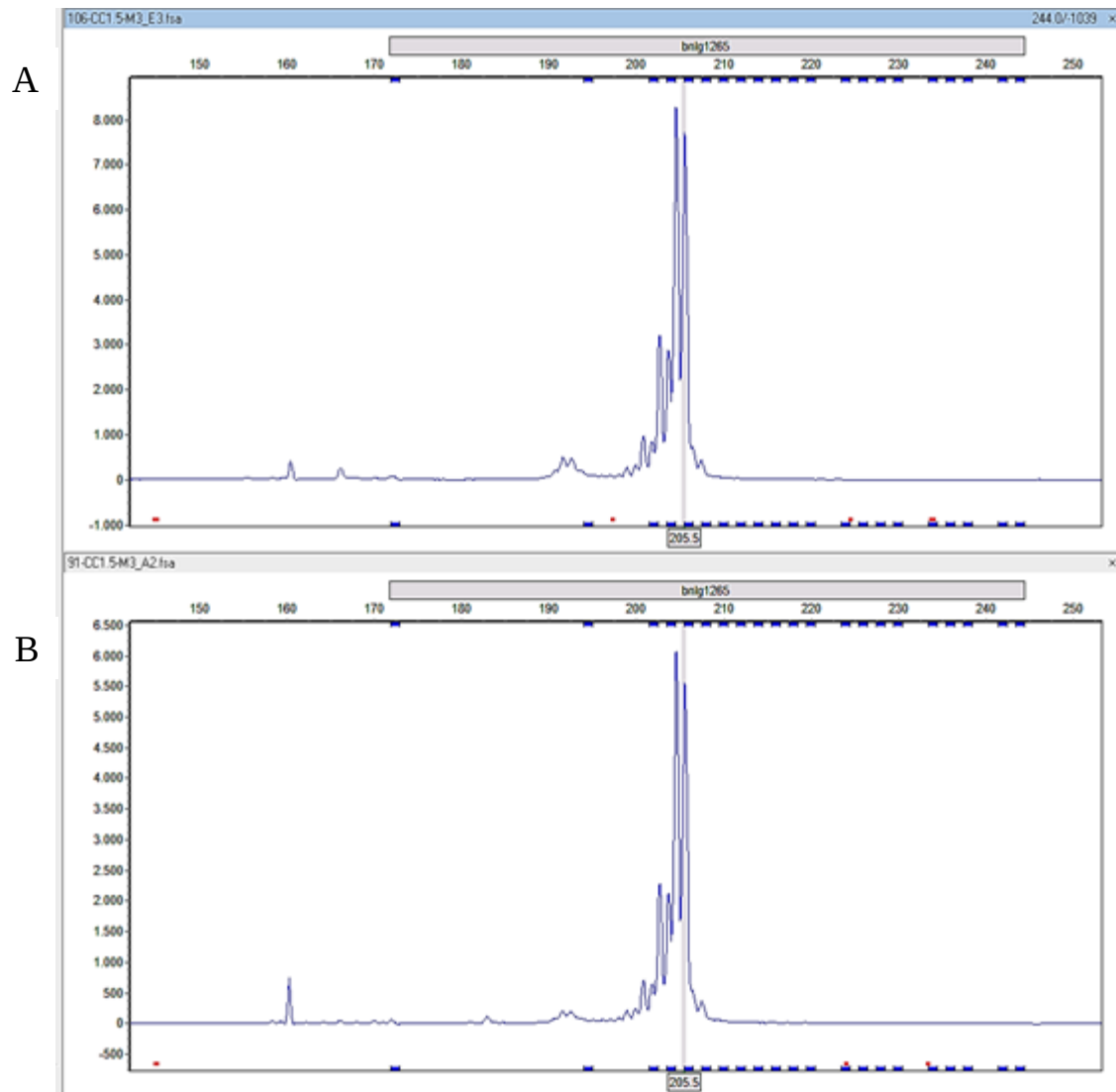


Figura 4: Ejemplo de electroferogramas visualizados en GeneMarker para una muestra de maíz criollo costarricense con el marcador *bnlg1890*. El panel A se muestra una reacción con un volumen final de 15uL, el panel B una reacción de PCR de 10uL. Se denotan resultados muy similares por lo que es indiferente utilizar un volumen final de 10uL o uno de 15uL.

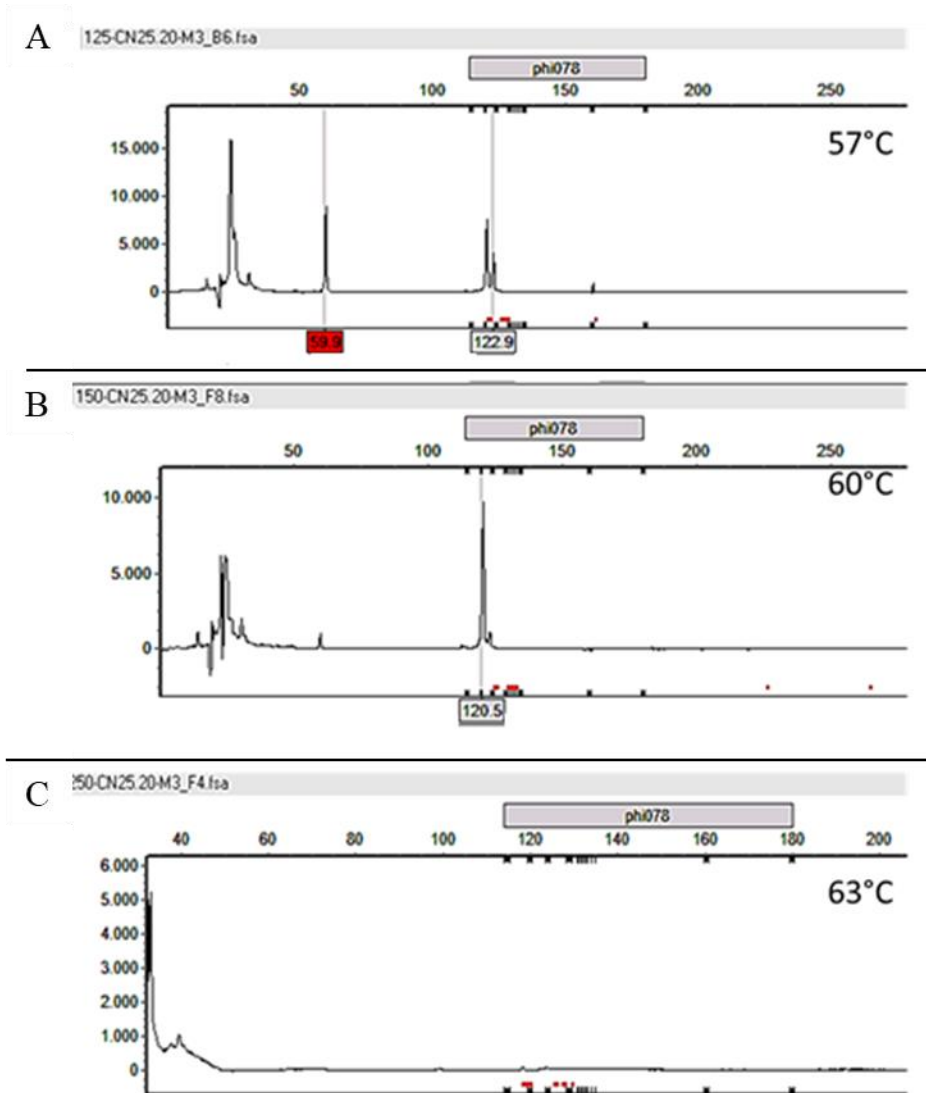
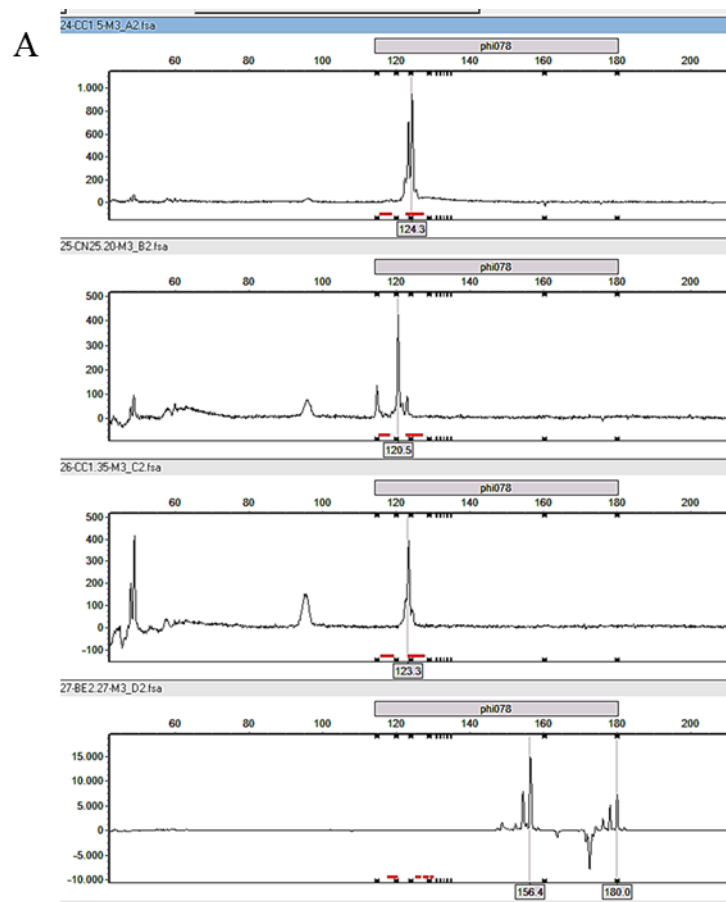
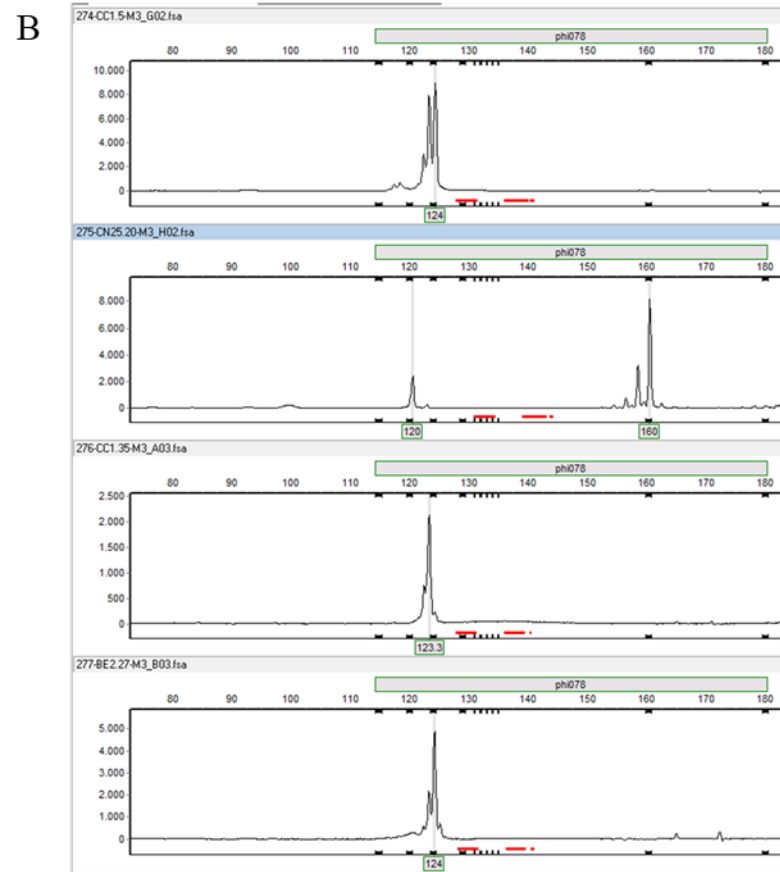


Figura 5: Electroferogramas visualizados en GeneMarker para una muestra de maíz criollo costarricense con el marcador *phi078*. Se demuestran la diferencia al utilizar distintas temperaturas de hibridación con el mismo protocolo. Panel superior $T_a=57^{\circ}\text{C}$, centro $T_a=60^{\circ}\text{C}$ y panel inferior $T_a=63^{\circ}\text{C}$. Se observan varias señales inespecíficas a 57°C que no están a 60°C , mientras que a 63°C no hubo amplificación.



Ta=60°C, 27 ciclos



Ta=60°C, 30 ciclos

Figura 6: Electroferogramas visualizados en GeneMarker para cuatro muestras de maíz criollo con el marcador *phi078*. Se observan las diferencias entre realizar una PCR a 27 ciclos (panel A) que a 30 ciclos (panel B) ambas con una *Ta* de 60°C.

4.1.4 Obtención de los genotipos, *bins* y rangos alélicos para Costa Rica:

MsatAllele, facilitó la definición de los alelos para cada SSR, y se estimó la diversidad genética. Sin embargo, se produjeron muchos alelos separados por 1pb y de baja frecuencia que podrían corresponder a errores de genotipo. Con ayuda de los histogramas y la revisión de electroferogramas, se eliminaron manualmente los alelos falsos y se volvió a estimar la diversidad genética. Al comparar las estimativas obtenidas antes y después de la verificación manual de los alelos se determinó que el número de alelos (N_a) y de alelos efectivos (N_e) por SSR fueron mayores antes de la rectificación de los datos (Cuadro 5). Mientras que el valor del PIC 0.85 (± 0.08) y 0.84 (± 0.09) y la heterocigosidad 0.84 (± 0.08) y 0.82 (± 0.09) antes y después de la verificación manual, respectivamente, dieron valores similares en ambos casos.

Por último, los rangos alélicos, luego de la corroboración, en general se disminuyeron. Por lo tanto, los nuevos resultados consistieron, principalmente, en la reducción en el número de alelos por cada marcador, pero no así la estimación de la diversidad genética. Sin embargo, estos cambios en las estimativas no resultaron ser significativos (Cuadro 5). Las estimativas de diversidad genética por cada SSR obtenidas luego de usar *MsatAllele* sin corroboración manual están en el Anexo 4. Los *bins* finales (utilizando *MsatAllele* y luego de la corroboración manual) para cada SSR se encuentran en el Anexo 5

Cuadro 5: Comparación de estimativas de diversidad genética antes y después de la corroboración manual al utilizar *MsatAllele*. Se muestra el promedio y la desviación estándar para el Contenido de información polimórfico (PIC), el Rango alélico (alelos mínimo y máximo), el número de alelos (Na), el número de alelos efectivos (Ne), la heterocigosidad observada (Ho) y la esperada (He) y la endogamia (F). Además de los resultados de la prueba de Wilcoxon para la comparación de promedios de las estimativas de diversidad genética por microsatélite antes y después de la corroboración manual.

Estimativa	Antes	Después	W	p
PIC	0.85(±0.08)	0.84(±0.09)	178.5	0.5691
Na	28.90(±12.43)	17.55(±10.31)	143	0.1255
Ne	9.29(±5.43)	7.99(±5.54)	173.5	0.4818
Ho	0.54(±0.12)	0.52(±0.12)	180.5	0.6067
He	0.84(±0.08)	0.82(±0.11)	180	0.5967
F	0.36(±0.13)	0.35(±0.15)	200.5	1

4.2- Caracterización de la diversidad genética de maíz criollo

4.2.1: Estimativas de diversidad por marcador molecular

Se estimó la diversidad genética del maíz criollo para cada SSR, utilizando los genotipos obtenidos con *MsatAllele* y corroborados manualmente (Cuadro 6). Se obtuvieron el PIC (contenido de información polimórfica, promedio del número de alelos (Na), número de alelos efectivos (Ne) y el rango alélico (mínimo, máximo), la heterocigosidad esperada (He), la heterocigosidad observada (Ho) y el índice de endogamia (F). El marcador más informativo fue bnlg1194 (PIC=0,95) y el menos informativo phi089 (PIC= 0,57). Asimismo, el marcador bnlg1194 permitió determinar el mayor número de alelos y de alelos efectivos fue con Na= 37 y Ne= 24.26. Mientras que el menor número de alelos y de alelos efectivos se obtuvo con phi089 Na= 4 y Ne=1.95.

Para todos los marcadores, la heterocigosidad observada es menor que la esperada (Cuadro 6). Además, los SSR con motivos dinucleotídicos presentaron mayores valores de

diversidad que los tetranucleotídicos. El marcador con mayor diversidad genética fue bnl1194 $H_e = 0.96$ y el menor phi089 con $H_e = 0.49$.

Por otra parte, el SSR que permite identificar el mayor índice de endogamia fue bnl1046 $F = 0.60$ y el menor se obtuvo para phi031 con $F = 0.12$. El microsatélite con un mayor porcentaje de valores perdidos fue bnl1449 (16%) y el menor fue phi015 (5%).

Cuadro 6: Estimativas de diversidad genética por microsatélite para muestras de maíz criollo de las regiones Brunca y Chorotega de Costa Rica, utilizando los genotipos generados con *MsatAllele* y la corroboración manual de los datos.

SSR	PIC	Mín.	Máx.	Na	Ne	Ho	He	F	%Valores perdidos
bnlg1046	0.88	169	217	21	7.48	0.34	0.87	0.60	0.16
bnlg1191	0.94	179	235	24	15.69	0.66	0.94	0.29	0.08
bnlg1194	0.95	152	224	37	24.26	0.57	0.96	0.40	0.12
bnlg1265	0.94	194	246	28	17.25	0.71	0.94	0.24	0.07
bnlg1449	0.89	87	166	16	8.79	0.42	0.89	0.53	0.17
bnlg1523	0.91	139	269	34	9.61	0.59	0.90	0.34	0.10
bnlg1740	0.91	115	189	27	10.09	0.47	0.90	0.47	0.14
bnlg1890	0.89	94	214	32	7.98	0.55	0.87	0.37	0.06
bnlg1917	0.84	88	192	26	5.29	0.34	0.81	0.58	0.18
bnlg2305	0.89	165	208	16	8.43	0.61	0.88	0.31	0.12
phi015	0.78	77	100	10	4.21	0.66	0.76	0.14	0.05
phi031	0.71	185	222	4	2.96	0.59	0.66	0.12	0.08
phi064	0.90	72	112	16	9.54	0.58	0.90	0.35	0.16
phi072	0.76	141	166	10	3.76	0.57	0.73	0.22	0.06
phi078	0.79	114	131	10	3.98	0.31	0.75	0.59	0.12
phi089	0.57	81	95	4	1.95	0.33	0.49	0.32	0.09
phi093	0.82	283	299	10	4.71	0.40	0.79	0.49	0.1
phi109188	0.79	151	174	9	4.47	0.61	0.78	0.21	0.04
phi116	0.82	151	177	8	4.92	0.51	0.80	0.36	0.09
phi96100	0.80	269	301	9	4.33	0.65	0.77	0.16	0.08
Promedio	0.84	145	196	17.55	7.99	0.52	0.82	0.35	0.10
Error estándar	0.093	58.870	59.270	10.313	5.543	0.125	0.111	0.151	0.041

4.2.2: Estimativas de diversidad genética por Región

Se analizaron en total 224 muestras de maíz criollo, 80 provenientes de la región Brunca y 144 de la Chorotega. Se estimó la diversidad genética por cada región (Cuadro 7). Evaluando los 20 SSR, en general, se encontró que las estimativas de diversidad son mayores en la región Chorotega que para Brunca. Se encontró para la región Brunca, en promedio $N_a = 13.65 (\pm 1.72)$ y $6.82 (\pm 0.98)$ N_e . Para la Chorotega $16.25 (\pm 2.03)$ N_a y $7.55 (\pm 1.01)$ N_e . Asimismo, la riqueza alélica promedio fue mayor para la región Chorotega que la Brunca (Cuadro 8), sin embargo, la diferencia no es significativa (Wilcoxon=172, $p=0.4612$). En cuanto a diversidad genética, la heterocigosidad es mayor para la región Chorotega ($H_e = 0.82 \pm 0.026$) que para la región Brunca ($H_e = 0.80 \pm 0.025$). Además, se determinó un mayor índice de endogamia en la región Brunca $F=0.37 (\pm 0.041)$ que en la región Chorotega $F= 0.33 (\pm 0.034)$. Se encontraron en promedio $1.3 (\pm 0.32)$ alelos privados para Brunca y $3.9 (\pm 0.75)$ para la región Chorotega. Por último, se encontró que ambas regiones están genéticamente poco diferenciadas, sin embargo, esta leve estructura genética es significativa ($G_{st}=0.018$, $p<0.05$).

4.2.3: Estimativas de diversidad genética por sitio de muestreo

En la región Chorotega se analizaron cinco sitios de muestreo que corresponden a Liberia, Cañas, La Cruz las Brisas, La Cruz Santa Cecilia y Santa Cruz. En la región Brunca se analizaron cuatro sitios de muestreo: Boruca, Laurel, Concepción y Valencia. Entre los sitios de Chorotega se analizaron 13 poblaciones de maíz criollo y entre los de Brunca ocho. En relación con las estimativas de diversidad genética por sitio de muestreo

(Cuadro 9), se encontró que Santa Cruz presentó un mayor número de alelos con 11.5 (± 1.22) Na y 6.38 (± 0.72) Ne. El menor número de alelos se encontró en Valencia con 3.7 (± 0.27) Na y 2.7 (± 0.2) Ne. En todos los sitios de muestreo, la heterocigosidad observada fue menor que la esperada. La mayor diversidad genética se encontró en Santa Cruz $H_e = 0.8$ (± 0.02) y la menor en Valencia $H_e = 0.57$ (± 0.04). Boruca fue el sitio con el mayor índice de endogamia $F = 0.36$ (± 0.04) y Valencia el de menor índice de endogamia $F = 0.12$ (± 0.09). Liberia fue el sitio en el que se encontró, en promedio un mayor número de alelos privados 0.85 (± 0.21), mientras que en Cañas no se encontró ningún alelo privado. Valencia fue el sitio en el que se obtuvo el mayor porcentaje de valores perdidos (28%) y Concepción el menor porcentaje (3%) (Cuadro 9). Se encontró también una leve pero significativa estructura genética entre los sitios de muestreo ($G_{st} = 0.04$, $p < 0.05$).

Cuadro 7: Estimativas de diversidad genética por región socioeconómica Brunca y Chorotega y en promedio para Costa Rica, obtenidas con el análisis de 20 SSR. Se muestran los promedios con respectiva desviación estándar del promedio de individuos muestreados por alelo (N), número de alelos (Na), número de alelos efectivos (Ne) la heterocigosidad esperada (H_e), la heterocigosidad observada (H_o), el índice de endogamia (F) y el número de alelos privados (Na.privados).

Región	N	Na	Ne	H_o	H_e	F	Na. privados
Brunca	74.25 (± 0.80)	13.65 (± 1.73)	6.82 (± 0.98)	0.50 (± 0.032)	0.80 (± 0.025)	0.37 (± 0.041)	1.3 (± 0.32)
Chorotega	126.75 (± 1.44)	16.25 (± 2.03)	7.55 (± 1.11)	0.54 (± 0.028)	0.82 (± 0.026)	0.33 (± 0.034)	3.9 (± 0.75)
Costa Rica	100.5 (± 2.28)	14.95 (± 1.33)	7.185 (± 0.73)	0.519 (± 0.021)	0.81 (± 0.018)	0.809 (± 0.018)	---

Cuadro 8: Riqueza alélica de cada microsatélite según la región socioeconómica. Se muestra el tamaño poblacional utilizado (N) para cada región y la estimativa de riqueza (Riqueza), el promedio por región y la desviación estándar.

SSR	N_Brunca	Riqueza_Brunca	N_Chorotega	Riqueza_Chorotega
bnlg1046	142	13.69	236	18.58
bnlg1191	158	21.24	256	20.24
bnlg1194	146	29.22	248	29.74
bnlg1265	156	22.26	260	22.92
bnlg1449	146	11.57	228	14.03
bnlg1523	148	21.75	256	25.46
bnlg1740	138	16.94	246	21.09
bnlg1890	156	21.1	266	22.42
bnlg1917	134	15.04	234	19.10
bnlg2305	148	13.4	248	11.59
phi015	152	8.87	274	9.73
phi031	152	4.00	258	3.99
phi064	144	11.7	232	14.72
phi072	156	7.63	266	9.47
phi078	150	8.29	246	8.89
phi089	160	3.77	248	3.00
phi093	152	8.00	252	9.34
phi109188	158	6.55	270	6.84
phi116	150	6.00	256	7.16
phi96100	160	6.77	254	8.39
Promedio	150.30	12.89	251.70	14.33
Error estándar	1.61	1.60	2.82	1.71

Cuadro 9: Estimativas de diversidad genética por sitio de muestreo dentro de cada región, según el promedio de los 20 SSR. Se muestran los promedios con respectiva desviación estándar de número de muestra(N), número de alelos (Na), número de alelos efectivos (Ne) la heterocigosidad esperada (He), la heterocigosidad observada (Ho), el índice de endogamia (F) y el número de alelos privados (Na.privados).

Sitio muestreo	N	Na	Ne	Ho	He	F	Na.privados
BORUCA	26.45 (± 0.36)	8.05 (± 0.85)	4.56 (± 0.46)	0.47 (± 0.04)	0.73 (± 0.03)	0.36 (± 0.04)	0.2 (± 0.12)
LAUREL	21.3 (± 0.23)	7.85 (± 0.79)	4.84 (± 0.53)	0.5 (± 0.03)	0.74 (± 0.03)	0.32 (± 0.05)	0.25 (± 0.1)
CONCEPCIÓN	23.1 (± 0.46)	8.25 (± 0.81)	4.65 (± 0.63)	0.54 (± 0.05)	0.73 (± 0.03)	0.27 (± 0.06)	0.3 (± 0.13)
LIBERIA	33.5 (± 0.5)	9.75 (± 0.87)	5.39 (± 0.61)	0.53 (± 0.04)	0.76 (± 0.03)	0.3 (± 0.05)	0.85 (± 0.21)
CAÑAS	10.6 (± 0.26)	5.7 (± 0.42)	4.06 (± 0.29)	0.54 (± 0.04)	0.72 (± 0.03)	0.24 (± 0.05)	0 (± 0)
LA CRUZ L.B	23.1 (± 0.52)	8.5 (± 0.8)	4.9 (± 0.46)	0.51 (± 0.04)	0.75 (± 0.03)	0.32 (± 0.04)	0.3 (± 0.15)
LA CRUZ S.C	21.4 (± 0.36)	8.75 (± 1.02)	5.46 (± 0.6)	0.5 (± 0.05)	0.76 (± 0.03)	0.35 (± 0.06)	0.5 (± 0.15)
SANTA CRUZ	35.7 (± 0.47)	11.5 (± 1.22)	6.38 (± 0.72)	0.58 (± 0.03)	0.80 (± 0.02)	0.27 (± 0.05)	0.45 (± 0.21)

V. DISCUSIÓN

5.1- Estandarización de Microsatélites para maíz criollo de Costa Rica

Una de las ventajas de utilizar microsatélites, es la posibilidad de comparar resultados entre estudios que utilizan el mismo tipo de marcador, lo que permite estudios colaborativos o comparativos a grandes escalas geográficas (Ellis *et al.*, 2011). Sin embargo, las condiciones para amplificar los mismos SSRs puede variar considerablemente entre estudios, por lo que es necesario estandarizar (Bravo, 2009). La selección de los microsatélites, la optimización de la reacción y la obtención de genotipos son clave para evitar falsas estimativas de frecuencias alélicas, que podrían resultar en un exceso de homocigotas, sobreestimación de endogamia e inclusive inferencias poco confiables sobre la estructura de la población (Bonin *et al.*, 2004).

Las amplificaciones inespecíficas durante la reacción de PCR son parte de los artefactos que pueden inducir a errores de genotipo. Estas son consecuencia de un mal apareamiento (“miss-match”) entre los imprimadores con otras secuencias no esperadas. Varios factores pueden influir en el mal apareamiento, entre estos, la cantidad de ADN, la temperatura de hibridación y el número de ciclos (Pompanon *et al.*, 2005). Es importante lograr una interpretación electroferogramas de tal forma que se distingan correctamente los alelos verdaderos de señales no alélicas o inespecíficas, ya que se puede incurrir en errores que inciden sobre la reproducibilidad de resultados y generan sobre (o sub)-estimaciones de la diversidad genética (Ellis *et al.*, 2011).

Con las condiciones iniciales para la reacción de PCR recomendadas por la casa comercial, se obtuvieron muchas amplificaciones inespecíficas. Por lo tanto, la optimización de la reacción en el termociclador inició con la modificación la temperatura

de hibridación (*Ta*). La *Ta* influye en la especificidad de los imprimadores con el ADN. Entre más baja la *Ta*, la hibridación es menos específica, entre más alta, se obtiene mayor especificidad, sin embargo, temperaturas muy altas evitan que algunos imprimadores hibridicen con el ADN meta/blanco (Sipos *et al.*, 2007). El rango teórico de temperaturas de los SSR utilizados era entre 55-61°C. De acuerdo con los experimentos, la *Ta* 60°C generó mejores resultados. Ésta es cinco grados mayores que la utilizada por Matsouka *et al.* (2002) con los mismos SSR y coincide con la utilizada por Bravo (2009) para los marcadores tetranucleotídicos, con muestras de Ecuador.

En cuanto al número de ciclos, los mejores resultados se dieron con 30 ciclos, número también utilizado por Bravo (2009). El número de ciclos influye en las uniones inespecíficas, de manera que entre mayor sean, incrementa la posibilidad de encontrar estas uniones, sin embargo, pocos ciclos disminuyen la cantidad de producto de PCR (Don *et al.*, 1991). Los 30 ciclos generaron amplificaciones con señal alta y pocas señales inespecíficas.

La concentración y calidad del ADN, también puede generar errores en el genotipo. Si la concentración y la calidad es muy baja o si el ADN está diluido se incrementando el riesgo de contaminación y la aparición de alelos falsos, mientras que, en concentraciones altas, la presencia de bandas inespecíficas corresponde al exceso de moléculas de ADN (Pompanon *et al.*, 2005). Se determinó que reducir la concentración de 50ng/uL a 10ng/uL disminuyó la presencia de señales inespecíficas. Cabe destacar que, entre los tres tipos de tejido utilizados en la extracción del ADN, la calidad y la concentración fue distinta. Los mejores resultados en cuanto a esos factores se obtuvieron con hoja, seguido de embrión y de grano entero. Sin embargo, la obtención de los genotipos no se vio afectada.

La automatización de la obtención de los *bins* permite la reducción de la intervención humana y el aumento en la reproducibilidad de los resultados. Para esto, recientemente los autores recomiendan utilizar programas o lenguajes estadísticos versátiles como R (*Grünwald et al.*, 2015). Este lenguaje ha permitido la creación de librerías o programas para el análisis de datos genéticos (Fuchs & Barrantes, 2015). *MsatAllele*, es una herramienta de R para generar *bins* que además tiene la ventaja de poder visualizar las distribuciones de los fragmentos en los *bins* para cada locus (Alberto, 2014). Utilizar esta herramienta no solo permitió definir los alelos de una forma rápida y automática, sino que se lograron detectar aquellos alelos que fueron falsamente asignados y corregir estos errores. Muchos de estos se identificaron por las diferencias de 1pb entre sí. El problema con los alelos consecutivos recae en que, al utilizar SSR dinucleotídicos y tetranucleotídicos, se espera una diferencia mínima de dos o cuatro nucleótidos respectivamente.

Cabe destacar que la corroboración manual efectuada, afectó el número de alelos por SSR, así como el número de alelos privados para cada población. Sin embargo, no afectó sustancialmente, otras medidas de diversidad genética como la heterocigosidad (Cuadro 5). La presencia de una mayor cantidad de alelos de la que debería, sobre estima la diversidad en ese sentido (más alelos) y modifica frecuencias alélicas (Pompanon *et al.*, 2005). La designación de los alelos dependerá de la química de la reacción de PCR, del secuenciador, de los softwares utilizados en el análisis, del marcador de peso molecular (*size standard*) y de la persona que realiza el genotipo. Es importante hacer optimizaciones o estandarizar este tipo de protocolos para aumentar la reproducibilidad y confiabilidad de los resultados. Se comprobó que *MsatAllele* resultó ser una herramienta diseñada para disminuir el error en la asignación de genotipos.

Por otro lado, los valores de heterocigosidad no se vieron tan afectados debido a que, si bien se modificaron los alelos, la proporción de heterocigotas debería mantenerse en la mayoría de los casos, con la excepción de falsos heterocigotas donde amplificaciones inespecíficas parecían un alelo. Pues la heterocigosidad se estima como la fracción de heterocigotas presente en la población, independientemente de la codificación de los alelos. Las diferencias entre las estimativas antes y después de la corroboración manual sugieren que la frecuencia de falsos heterocigotas es mínima (o despreciable). Probablemente se deba al gran tamaño de muestra, pues entre mayor sea para *MsatAllele* la definición de los alelos es más certera, pues existen más datos para generar las distribuciones de cada alelo (Alberto, 2014). Además, de que los cambios, principalmente en el número de alelos, si bien fue importante, no fueron cambios significativos. Lo cual indica la solidez de los datos producidos en la investigación.

Utilizar granos de la mazorca permitió aumentar la precisión de la asignación de alelos, al descartar aquellos que no se compartían entre los medios hermanos. Lo anterior resulta valioso, reduciendo errores de genotipo. El procedimiento del genotipo permitió obtener los rangos alélicos de cada SSR para Costa Rica, información que no estaba en la base de datos de Vigoroux *et al* (2008).

5.3- Caracterización de la diversidad genética de maíz criollo

El PIC indica la calidad de la información de los marcadores. En general, valores mayores a 0.5 se consideran muy informativos (Jiménez, 2014). Los valores obtenidos con los SSR seleccionados, son mayores a 0.5, lo que nos significa que son buenos para estimar la diversidad genética del maíz criollo de ambas regiones. Las diferencias encontradas entre las estimativas según el motivo de repetición se deben a que los SSR dinucleotídicos presentan una mayor tasa de mutación que los tetranucleotídicos (Becerra & Paredes, 2000,

Vigoroux *et al.*, 2002). Por lo tanto, los dinucleotídicos presentan mayor variación en el tamaño de los fragmentos y es posible encontrar un mayor número de alelos para éstos. El promedio de número de alelos y de alelos efectivos por marcador, es similar a lo encontrado en estudios de maíz criollo a nivel global como el de Vigoroux *et al.* (2008) y el de Matsouka *et al* (2002). En estos, se utilizaron los mismos 20 SSR de este estudio, además de otros 76 y 26 microsatélites, respectivamente. De manera comparable a los resultados de esta investigación Matsouka *et al.*, 2002; Vigoroux *et al.* 2008 encontraron que el marcador bnlgl1194 es el más diverso y phi089 el menos diverso. Contar con SSR que pueden estimar distintos niveles de diversidad es utilizado para evitar sesgos, pues escoger marcadores muy variables sobreestima la diversidad, caso contrario la subestima.

Todos los marcadores presentaron valores altos de heterocigosidad esperada. Sin embargo, los valores de heterocigosidad observados son menores (Cuadro 6). Lo cual resulta en coeficientes de endogamia significativos ($F > 0$), representando un déficit de heterocigotas, no esperado para una especie alógama. Esto puede indicar que las poblaciones en estudio están más aisladas y reciben poco flujo de genes externo o en algunos casos de autopolinización (Salami *et al.*, 2007; Jiménez 2014). Sin embargo, es posible que sea por eventos de apareamiento entre parientes. El proceso de selección de semilla por los agricultores aumenta la proporción de plantas medias hermanas pues de una misma mazorca con características deseables se almacena semilla para la próxima siembra. Lo anterior propicia el apareamiento entre plantas relacionadas lo que incrementa la endogamia de los cultivares. Este déficit de heterocigotas, también lo encontraron Vigoroux *et al.* (2008), con una reducción del 3.5% en la heterocigosidad esperada con respecto a la observada. Además, el diseño experimental propuesto para determinar los genotipos correctos también pudo influir en el déficit de heterocigotas. Al tener cuatro granos de una misma mazorca, la

presencia de individuos emparentados es mayor, lo que sobreestima la endogamia. En cuanto a los valores perdidos, en total se obtuvieron entre 5%-17%. Para valores mayores a 15% estos, se recomienda excluir el SSR del estudio (Vigouroux *et al*, 2008), o de ser posible repetir la muestra, sin embargo, por motivos de presupuesto esta última no se logró hacer para todas las muestras.

En general, las diferentes estimativas de diversidad genética son mayores para la región Chorotega que para la Brunca, especialmente en el número de alelos. Este resultado puede reflejar diferencias en el tamaño de muestra, ya que se espera un mayor número de alelos en muestras grandes por un efecto de muestreo (Slatkin, 1985). En este caso, para la región Chorotega se encontró una mayor cantidad de agricultores cultivando maíz criollo.

La presencia de alelos privados en ambas regiones es indicio de diferencias entre germoplasma presente en Chorotega y en Brunca, pues representan aquellos alelos que son únicos para cada una de éstas. Lo anterior podría estar influyendo en la leve pero significativa estructura genética entre ambas regiones ($G_{st}=0.018$, $p<0.05$). Sin embargo, los valores de heterocigosidad, son bastante similares lo que podría representar similitud entre las regiones, como ha sido reportado en México (Pressoir & Berthaud, 2004). Esto último, estaría relacionado al manejo de la semilla, a través de la selección y mantenimiento de materiales por parte de los productores de maíz criollo que diferencia los materiales pero que el intercambio de semillas entre productores de diferentes zonas homogeniza los cultivares y mantiene su diversidad genética (Pressoir & Berthaud, 2004).

El valor de la heterocigosidad esperada de $H_e=0.81$ total para ambas regiones demuestra que el maíz criollo es genéticamente diverso y es comparable con lo reportado para maíz criollo en otros países y regiones como México con $H_e=0.86$, el Caribe $H_e=0.80$, América del Sur $H_e=0.86$ y Norteamérica $H_e=0.82$ (Vigouroux *et al.*, 2008). Es importante

destacar que, debido a los procesos de domesticación del maíz cultivado, se esperaría que el maíz criollo tuviera una reducción en la diversidad en comparación con sus parientes silvestres y en comparación con el maíz criollo de los centros de origen (Vigorous *et al.*, 2002). Esto como consecuencia de la fijación de alelos de interés agronómico (por selección) y a la reducción de los individuos que contribuyen a las siguientes generaciones (cuellos de botella), producto también de la selección. Sin embargo, el que la diversidad del maíz criollo sea tan similar entre los diferentes sitios, sugiere que las tasas de mutación del genoma han hecho que las poblaciones de maíz se recuperen de los cuellos de botella producto de la domesticación (Eyre-Walker *et al.*, 1998; Vigorous *et al.*, 2005).

Los sitios de muestreo genéticamente más diversos se encuentran en la región Chorotega. Santa Cruz y Liberia fueron los sitios con mayor número de alelos y con heterocigosidad más alta (Cuadro 8). En la región Chorotega fue donde se logró coleccionar una mayor variabilidad de materiales, a nivel morfológico y de sitios, además de un mayor número de muestra que en la región Brunca. Es posible que lo anterior influyera en las diferencias entre las estimativas de diversidad. La existencia de una estructura genética leve pero significativa entre los sitios de muestreo indica la existencia de flujo génico entre los sitios y las regiones, probablemente producto del intercambio de semilla entre los agricultores de maíz criollo.

VI. CONCLUSIONES

El paquete *MsatAllele* resultó ser muy útil para la obtención de los alelos. Facilitó la asignación de los genotipos, así como la visualización de errores de genotipo. Esta herramienta y la optimización de la amplificación de los 20 SSR, permitió obtener los rangos alélicos de cada SSR para Costa Rica, información que no estaba en la base de datos de Vigoroux *et al* (2008). Estos *bins* y rangos alélicos, permitirán obtener genotipos más precisos en futuros estudios que se utilicen los mismos microsatélites.

Los 20 SSR utilizados para el análisis del maíz criollo de las regiones Brunca y Chorotega, mostraron altos niveles de diversidad genética en ambas regiones, como se esperaba. Esto sugiere la existencia de poblaciones efectivas grandes y migración de alelos entre las poblaciones, por el intercambio de semilla que mantiene la alta diversidad genética. Además, la presencia de alelos privados hace referencia a las posibles diferencias entre germoplasma presente en Chorotega y en Brunca. También las diferencias en los niveles de diversidad genética, aunque leves pero significativos, encontrados en los sitios de muestreo podría estar indicando diferenciación genética dentro de cada región. Todo lo anterior reitera la importancia del maíz criollo como un recurso fitogenético diverso y de interés para la conservación en nuestro país.

VII. RECOMENDACIONES

Se recomienda analizar la estructura genética dentro y entre las regiones Brunca y Chorotega, para determinar la diferencia o similitud genética entre las poblaciones y si efectivamente existen poblaciones de maíz criollo genéticamente independientes. También, ampliar el tamaño de muestra para ambas regiones en al menos, principalmente la Brunca e incluir poblaciones de otras regiones del país y hacer una caracterización morfológica, con el fin de relacionar la diversidad fenotípica con la genotípica. Además, comparar los datos obtenidos con bases de datos internacionales con el fin de completar la historia de la domesticación del maíz en el continente americano.

VIII REFERENCIAS

- ACAN-EFE (2014). Decretan al maíz como Patrimonio Cultural de Costa Rica. La Nación, 25 de julio de 2014. Recuperado de http://www.nacion.com/vivir/patrimonio/Decretan-Patrimonio-Cultural-Costa-Rica_0_1428857212.html?print=1 [Consulta 11 octubre de 2014].
- Acosta, R. (2009). El cultivo del maíz, Su origen y clasificación. EL MAIZ en Cuba. *Cultivos Tropicales*, 30(2), 113-120.
- Acosta, R., Ríos, H., Verde, G. & Pomagualli, D. (2003). Evaluación morfo agronómica de la diversidad genética de variedades locales de maíz (*Zea mays*, L.) en la Palma, Pinar del Río. *Cultivos Tropicales*, 24(4), 61-67.
- Aford, M. R., & Horn, S. P. (2004). Pollen evidence of the earliest maize agriculture in Costa Rica. *Journal of Latin American Geography*, 3(1), 108-115.
- Alberto F. (2014). MsatAllele_1.0: an R package to visualize the binning of microsatellite alleles. *Journal of Heredity*, 100(3): 394-397
- Alvarado, R. (2003). Regiones y cantones de Costa Rica, Serie cantones de Costa Rica N° 2. IFAM, Costa Rica. 55p
- Becerra, V. & Paredes, M. (2000). Uso de marcadores bioquímicos y moleculares en estudios de diversidad genética. *Agricultura Técnica*, 60(3), 270-281
- Bedoya, C. (2012). Estudios de diversidad genética en poblaciones de maíz (*Zea mays* L.) evaluadas con microsatélites. Tesis doctoral. Universidad de las Islas Baleares, España. 223p.
- Bedoya, C. & Chávez, V. (2010). Teocintle: el ancestro del maíz. *Claridades agropecuarias*, 201 (1), 32-42.
- Bedoya, C. A., Dreisigacker, S., Hearne, S., Franco, J., Mir, C., Prasanna, B. M., ... & Warburton, M. L. (2017). Genetic diversity and population structure of native maize populations in Latin America and the Caribbean. *PloS one*, 12(4), e0173488.
- Bellon, M. R., Hodson, D., & Hellin, J. (2011). Assessing the vulnerability of traditional maize seed systems in Mexico to climate change. *Proceedings of the National Academy of Sciences*, 108(33), 13432-13437.

- Bonilla, N. (2009). Cultivo de maíz (*Zea mays*): Manual de recomendaciones técnicas cultivo de maíz. INTA, Costa Rica. 75p.
- Bonin, A., Bellemain, E., Bronken Eidesen, P., Pompanon, F., Brochmann, C., & Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. *Molecular ecology*, 13(11), 3261-3273.
- Bradley, J. E. (1994). Tronadora Vieja: An Archaic and Early Formadve site in die Arenal region. In *Archaeology, Volcanism, and Rsmote Sensing in the Arenal Region, Costa Rica*. P. D. Sheets and B. R. McKee (eds.), pp. 73-86. Ausdn: University of Texas Press.
- Bravo, A. (2009). Caracterización morfológica y molecular de accesiones de maíz negro (*Zea mays* L.) mediante Análisis de Secuencias Simples Repetidas. Tesis para optar por el título en título de Ingeniería en Procesos Biotecnológicos. Universidad San Francisco de Quito, Ecuador. 61p.
- Carvalho, M., Teixeira, J., Abreu, I., Sousa, N., Marques, T.,Vieiria, M. & Motto, M. (2008). Evaluation of the maize (*Zea mays* L.) diversity on the Archipelago of Madeira. *Genetic Resource Crop Evolution*, 55 (2), 221-233.
- Clement, R. M., & Horn, S. P. (2001). Pre-Columbian land-use history in Costa Rica: a 3000-year record of forest clearance, agriculture and fires from Laguna Zoncho. *Holocene*, 11(4), 419-426.
- Cortés, G. (1994). Atlas agropecuario de Costa Rica 1° edición. EUNED. San José, Costa Rica. 513p.
- CNRF. (1996). Costa Rica: Informe Nacional para la Conferencia Técnica Internacional de la FAO sobre los Recursos Filogenéticos (Leipzig, 1996). Comisión Nacional de Recursos Fitogenéticos (CNRF). San José, Costa Rica. 52p.
- Cruz, H. (2012). Diversidad de maíces criollos en cinco localidades del municipio de Tacotalpa, Tabasco. (Tesis Licenciatura). Universidad Intercultural del Estado de Tabasco, México. 69p.
- Delgado, J. (2012). Ficha técnica de maíz blanco Período agrícola 2012-2013. Recuperado de: <http://www.sima.go.cr/index.php/mercados-de-maiz/item/286-ficha-tecnica-maiz> [Consulta 12 octubre de 2014].

- De Barba, M., Miquel, C., Lobréaux, S., Quenette, P. Y., Swenson, J. E., & Taberlet, P. (2017). High-throughput microsatellite genotyping in ecology: improved accuracy, efficiency, standardization and success with low-quantity and degraded DNA. *Molecular ecology resources*, 17(3), 492-507.
- Doebley, J. (1990). Molecular Evidence and the Evolution of Maize. *Economic Botany*, 44 (3), 6-27.
- Doebley, J. (2004). The Genetics of Maize Evolution. *Annual Review of Genetics*, 38, 37-59.
- Doebley, J. F., Gaut, B. S., & Smith, B. D. (2006). The molecular genetics of crop domestication. *Cell*, 127(7), 1309-1321.
- Don, R. H., Cox, P. T., Wainwright, B. J., Baker, K., & Mattick, J. S. (1991). 'Touchdown'PCR to circumvent spurious priming during gene amplification. *Nucleic acids research*, 19(14), 4008.
- Doyle, J.J. & Doyle, J.L. (1990). Isolation of plant DNA from fresh tissue. *Focus*, 12, 13-15.
- Ellis, J. S., Gilbey, J., Armstrong, A., Balstad, T., Cauwelier, E., Cherbonnel, C., ... & Dillane, E. (2011). Microsatellite standardization and evaluation of genotyping error in a large multi-partner research programme for conservation of Atlantic salmon (*Salmo salar L.*). *Genetica*, 139(3), 353-367.
- Espinosa-Trujillo, E., Castillo, M., del Carmen, M., Castillo-González, F., Ortiz Cereceres, J., Delgado-Alvarado, A., & Carrillo Salazar, A. (2009). Acumulación de antocianinas en pericarpio y aleurona del grano y sus efectos genéticos en poblaciones criollas de maíz pigmentado. *Revista fitotecnica mexicana*, 32(4), 303-309.
- Eyre-Walker, A., Gaut, R. L., Hilton, H., Feldman, D. L., & Gaut, B. S. (1998). Investigation of the bottleneck leading to the domestication of maize. *Proceedings of the National Academy of Sciences*, 95(8), 4441-4446.
- Fuchs, E. & Barrantes, G. (2015). El lenguaje estadístico R aplicado a las ciencias biológicas. Editorial UCR, San José, Costa Rica. 197p.
- Fuentes, M. (2012). Manejo, conservación y uso de la agrobiodiversidad en maíz. Programa Colaborativo de Fitomejoramiento Participativo en Mesoamérica. Guatemala. 13p.

- Grünwald, N., Kamvar, Z. & Everhart, S. (2015). Population genetics in R. Corvallis, Oregon, USA. Recuperado de http://grunwaldlab.github.io/Population_Genetics_in_R/Data_Preparation.html [Consulta 02 febrero 2017].
- Hartings, H., Berardo, N., Mazzinelli, G. F., Valoti, P., Verderio, A., & Motto, M. (2008). Assessment of genetic diversity and relationships among maize (*Zea mays* L.) Italian landraces by morphological traits and AFLP profiling. *Theoretical and Applied Genetics*, 117(6), 831-842.
- Hedrick, P. W. (1999). Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution*, 53(2), 313-318.
- Hellin, J., Bellon, M. R., & Hearne, S. J. (2014). Maize landraces and adaptation to climate change in Mexico. *Journal of Crop Improvement*, 28(4), 484-501.
- Hilje, I (2008). Diversidad genética del árbol *Annona purpurea* Moc. & Sessé ex Dunal (Annonaceae) utilizando microsatélites como marcadores moleculares. Tesis para optar por el grado de Licenciatura en Genética y Biotecnología, Universidad de Costa Rica. 80 p.
- Hofmann, F., Otto, M., & Wosniok, W. (2014). Maize pollen deposition in relation to distance from the nearest pollen source under common cultivation-results of 10 years of monitoring (2001 to 2010). *Environmental Sciences Europe*, 26(1), 24.
- Horn, S. & Kennedy, L. (2001). Pollen Evidence of Maize Cultivation 2700 B.P. at La Selva Biological Station, Costa Rica. *Biotropica*, 33 (1), 191-196.
- IICA. (2014). Las cadenas de valor de maíz blanco y frijol en Centroamérica: actores, problemas y acciones para su competitividad. IICA, Red SICTA, Cooperación Suiza en América Central. Costa Rica. 2014. 127 p.
- Jiménez, R. (2014). Caracterización de las razas criollas e indígenas de maíz colombiano por medio de marcadores moleculares SSR. Tesis para optar por el grado de Magister en Ciencias Biológicas. Universidad Nacional de Colombia Facultad de Ciencias Agropecuarias. 70p.
- Kato, Y., Mapes, C., Mera, L., Serratos, J. & Bye, R. (2009). Origen y Diversificación del Maíz Revisión Analítica. Universidad Nacional Autónoma de México, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México, D.F., 116 p.

- Laborda, P., Oliveira, K., Garcia, A., Paterniani, M. & de Souza, A. (2005). Tropical maize germplasm: what can we say about its genetic diversity in the light of molecular markers? *Theoretical and Applied Genetics* (2005) 111: 1288–1299.
- Matsuoka, Y., Vigouroux, Y., Goodman, M., Sanchez, J., Buckler, E. & Doebley, J. (2002). A single domestication for maize shown by multilocus microsatellite genotyping. *Proceedings of the National Academy of Sciences*, 99 (9), 6080–6084.
- MCJ. (2014). Maíz criollo integra alimentación y tradiciones por miles de años. Ministerio de Cultura y Juventud, comunicado de prensa, 7 agosto 2014. <http://www.mcj.go.cr/actualidad/noticias/2014/agosto/noticias/consecutivo368.aspx>. [Consulta 01 noviembre 2015]
- Molina, M. (2011). Estudios citogenéticos evolutivos del género *Zea*. (Tesis doctorado). Universidad Politécnica de Valencia, Departamento de Biología Vegetal.
- Ochoa, A. B. (2013). Estructura genética espacial asociada a la selección en razas de maíz de México. Tesis para optar por *Maestría en Ciencias Biológicas (Biología Evolutiva)*, UNAM, México. 117p.
- Office of the Gene Technology Regulator : OGTR. (2008). The biology of *Zea mays* L. ssp *mays* (maize corn). Recuperado de [http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/content/maize-3/\\$FILE/biologymaize08_2.pdf](http://www.ogtr.gov.au/internet/ogtr/publishing.nsf/content/maize-3/$FILE/biologymaize08_2.pdf) [Consulta 5 mayo 2014].
- Oliveira, F., Bendela, G., Allaby, R. & Brown, T. (2003). DNA from primitive maize landraces and archaeological remains: implications for the domestication of maize and its expansion into South America. *Journal of Archaeological Science*, 30 (7), 901–908.
- Paliwal, R., Granados, G., Lafitte, H., Violic, A. & Marathée, J.P. (2001). El Maíz en los trópicos: mejoramiento y producción. Organización de las Naciones Unidas para la Agricultura y la Alimentación, Roma. 392p.
- Peakall, R. & Smouse, P.E. (2012). GenAlEx 6.5: Genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28(1): 2537–2539
- Pineda-Hidalgo, K. V., Méndez-Marroquín, K. P., Alvarez, E., Chávez-Ontiveros, J., Sánchez-Peña, P., Garzón-Tiznado, J. A., Vega-García, M. & López-Valenzuela, J. A. (2013). Microsatellite-based genetic diversity among accessions of maize landraces from Sinaloa in México. *Hereditas*, 150(4), 53-59.

- Piñero, D., Caballero-Mellado, J., Cabrera-Toledo, D., Canteros, C., Casas, A., Castañeda, A., Castillo, A., Cerritos, R., *et al.* (2008). La diversidad genética como instrumento para la conservación y el aprovechamiento de la biodiversidad: estudios en especies mexicanas. pp 437-494. *In* Brush, S. & Zizumbo-Villarreal, D. (eds). Capital natural de México, Conocimiento actual de la biodiversidad. CONABIO, México, 621p.
- Pompanon, F., Bonin, A., Bellemain, E., & Taberlet, P. (2005). Genotyping errors: causes, consequences and solutions. *Nature Reviews Genetics*, 6(11), 847-846.
- Prasanna, B.M. (2012). Diversity in global maize germplasm: Characterization and utilization. *Journal of Bioscience*, 37 (5), 843–855.
- Pressoir, G., & Berthaud, J. (2004). Patterns of population structure in maize landraces from the Central Valleys of Oaxaca in Mexico. *Heredity*, 92(2), 88-94.
- QIAGEN (2010). Multiplex PCR Handbook. 48p.
- R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- Rivas, G., Rodríguez, A., Padilla, D., Hernández, L. & Suchini, J. (2013). Bancos comunitarios de semillas criollas: una opción para la conservación de la agrobiodiversidad. CATIE, Turrialba, Costa Rica, 16p.
- Romero-Severson, J., Smith, J., Ziegler, J., A Hauser, J., A Joe, L. & Hookstra, G. (2001). Pedigree analysis and haplotype sharing within diverse groups of *Zea mays* L. *Theoretical and Applied Genetics*, 103(4), 567-574.
- Sánchez, G. (2011). Diversidad del Maíz y el Teocintle. Informe preparado para el proyecto: “Recopilación, generación, actualización y análisis de información acerca de la diversidad genética de maíces y sus parientes silvestres en México”. Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. Manuscrito.
- Sanchez G., Goodman, M. & Stuber, C. (2000). Isozymatic and Morphological Diversity in the Races of Maize of Mexico. *Economic Botany*, 54 (1), 43-59.
- Salami, A., Adegoke, S. & Adegbite, O. (2007) Genetic Variability among Maize Cultivars Grown in Ekiti-State, Nigeria. *Middle-East Journal of Scientific Research*, 2 (1), 9-13.
- Salinas Moreno, Y., Cruz Chávez, F. J., Díaz Ortiz, S. A., & Castillo González, F. (2012). Granos de maíces pigmentados de Chiapas, características físicas, contenido de antocianinas y valor nutraceutico. *Revista fitotecnica mexicana*, 35(1), 33-41.

- Schlötterer, C. (2004). The evolution of molecular markers--just a matter of fashion?. *Nature reviews. Genetics*, 5(1), 63.
- Serratos-Hernández, J.A. (2009). El origen y diversidad del maíz en el continente americano. Green Peace, México. 33p.
- Sharma, L, Prasanna, B.M. & Ramesh, B. (2010). Phenotypic and microsatellite-based diversity and population genetic structure of maize landraces in India, especially from the NorthEast Himalayan region. *Genetica*, 138 (6), 619–631.
- Sipos, R., Székely, A. J., Palatinszky, M., Révész, S., Márialigeti, K., & Nikolausz, M. (2007). Effect of primer mismatch, annealing temperature and PCR cycle number on 16S rRNA gene-targeting bacterial community analysis. *FEMS Microbiology Ecology*, 60(2), 341-350.
- Slatkin, M. 1985. Estimating levels of gene flow in natural populations. *Genetics*, 99: 323-335.
- Smalley, J. & Blake, M. (2003). Sweet Beginnings Stalk Sugar and the Domestication of Maize. *Current Anthropology*, 44 (5), 675-703.
- Taylor, Z. P., Horn, S. P., & Finkelstein, D. B. (2013). Maize pollen concentrations in Neotropical lake sediments as an indicator of the scale of prehistoric agriculture. *The Holocene*, 23(1), 78-84.
- Vigouroux, Y., Glaubitz, J., Matsuoka, Y., Goodman, M., Sánchez, J. & Doebley. (2008). Population structure and genetic diversity of new world maize races assessed by DNA microsatellites. *American Journal of Botany*, 95 (10), 1240–1253.
- Vigouroux, Y., Jaqueth, J. S., Matsuoka, Y., Smith, O. S., Beavis, W. D., Smith, J. S. C., & Doebley, J. (2002). Rate and pattern of mutation at microsatellite loci in maize. *Molecular Biology and Evolution*, 19(8), 1251-1260.
- Vigouroux, Y., Mitchell, S., Matsuoka, Y., Hamblin, M., Kresovich, S., Smith, S., Jaqueth, J., Smith, O. & Doebley, J. (2005). An analysis of genetic diversity across the maize genome using microsatellites. *Genetics*, 169(4), 1617-1630.
- Wang, C., Schroeder, K. B., & Rosenberg, N. A. (2012). A maximum-likelihood method to correct for allelic dropout in microsatellite data with no replicate genotypes. *Genetics*, 192(2), 651-669.

- Warburton, M.L., Reif, J.C., Frisch, M., Bohn, M., Bedoya, C., Xia, X.C., Crossa, J., Franco, J., Hoisington, D., Pixley, K., Taba, S. & Melchinger, A.E. (2008). Genetic diversity in CIMMYT nontemperate maize germplasm: Landraces, open pollinated varieties, and inbred lines. *Crop Science*. 48, (2), 617–624
- Yao, Q., Fang, P., Kang, K. & Pan, G.T. (2008) Genetic diversity based on SSR markers in maize (*Zea mays* L.) landraces from Wuling mountain region in China. *Journal of Genetics*, 87(3), 287–291.
- Yao, Q., Yang, K., Pan, G. & Rong, T. (2007). Genetic Diversity of Maize (*Zea mays* L.) Landraces from Southwest China Based on SSR Data. *Journal of Genetics and Genomics*, 34(9), 851-860.

IV ANEXOS

Anexo 1

Protocolo de extracción de ADN de maíz

Modificado de Doyle & Doyle (1990)

- 1-Colocar de 0,1g a 0,5g del tejido tubos de 2mL. Agregar 0,5mL de Buffer de extracción: Mezclar inmediatamente invirtiendo los tubos, asegurarse de que toda la muestra quede en contacto con el buffer.
- 2-Colocar dos perlas de metal (o de cerámica) y macerar en el macerador automático MM400 (Retsch) a 30 frecuencias por 6 minutos, una vez que se detiene, volver a poner por otros 6 minutos.
- 3- Incubar en baño maría 30 min a 60°C, luego dejar enfriar a temperatura ambiente.
- 4-En la capilla agregar 500uL de Cloroformo Octanol (24:1).
- 5-Centrifugar a 5min por 8mil rpm.
- 6-En la capilla, transferir 400 µL de la fase acuosa a tubos nuevos, transferir primero 200µL y luego el resto con cuidado de no capturar parte de la fase orgánica.
- 7-Agregar 1/10 parte del volumen transferido de acetato de sodio (3M, pH=5).
- 8- Agregar 2,5 volúmenes de etanol al 95% frío. Dejar reposar en frío (4°C) por 1h.
- 9- Centrifugación diferencial 3min a 300 rpm luego 3min a 5mil rpm. Descartar el sobrenadante.
- 10- Agregar 500 µL de etanol al 76%.
- 11- Centrifugar por 6 min a 5000 rpm. Descartar el sobrenadante.
- 12- Secar los tubos en el horno a 37°C por unos 20 o 40 min, o bien hasta que se evapore todo el líquido. También se pueden centrifugar al vacío por 6min a 45°C.
- 13- Agregar 50 µL de TE e incubar por 15min a 50°C.

Anexo 2: Conformación de los múltiplex (filas) y color del fluorocromo (columna) que se utilizaron para la caracterización de la diversidad genética de maíz criollo de las regiones Brunca y Chorotega de Costa Rica con 20 SSR.

Multiplex	Marcador azul (6-FAM)	Marcador rojo (PET)	Marcador verde (VIC)	Marcador negro (NED)
1	<i>bnlg1890</i>	<i>bnlg1917</i>	<i>bnlg1449</i>	<i>bnlg1523</i>
2	<i>bnlg1191</i>	<i>bnlg1740</i>	<i>phi116</i>	<i>bnlg2305</i>
3	<i>bnlg1265</i>	<i>bnlg1046</i>	<i>bnlg1194</i>	<i>phi078</i>
4	<i>phi015</i>	<i>phi089</i>	<i>phi031</i>	<i>phi064</i>
5	<i>phi109188</i>	<i>phi96100</i>	<i>phi093</i>	<i>phi072</i>

Anexo 3: Reacción de PCR multiplex (Qiagen, 2010) utilizada para la amplificación de 20 SSR seleccionados para el estudio de diversidad genética de las regiones Brunca y Chorotega de Costa Rica.

Componente	Volumen de Reacción	Concentración final
2X Multiplex PCR Master Mix	5uL	1X
Primer Mix (F+R)	0.4uL c/u = 1.6uL	0,2uM
Agua	1.4uL	---
ADN	2uL	10ng/uL
Volumen total	10uL	

Anexo 4: Estimativas de diversidad genética para cada microsatélite con la base de datos de genotipos obtenida con el paquete *MsatAllele* de R y sin verificación manual de los alelos. Se muestran el Contenido de información Polimórfica (PIC), el Rango alélico (alelos mínimo y máximo), el número de alelos (Na), el número de alelos efectivos (Ne), la heterocigosidad observada (Ho) y la esperada (He) y la endogamia (F).

SSR	PIC	Mínimo	Máximo	Na	Ne	Ho	He	F
<i>bnlg1046</i>	0.91	124	216	32	11.36	0.41	0.87	0.53
<i>bnlg1191</i>	0.95	162	247	40	20.63	0.6	0.89	0.33
<i>bnlg1194</i>	0.95	130	257	43	18.57	0.59	0.87	0.33
<i>bnlg1265</i>	0.95	188	246	44	20.27	0.44	0.82	0.47
<i>bnlg1449</i>	0.89	82	231	19	8.82	0.68	0.94	0.28
<i>bnlg1523</i>	0.9	87	275	46	9.74	0.47	0.9	0.48
<i>bnlg1740</i>	0.92	115	200	32	12.96	0.62	0.89	0.3
<i>bnlg1890</i>	0.88	94	220	44	8.55	0.51	0.78	0.35
<i>bnlg1917</i>	0.84	88	222	44	6.08	0.37	0.9	0.59
<i>bnlg2305</i>	0.89	151	227	28	9.20	0.67	0.94	0.29
<i>phi015</i>	0.85	76	100	37	6.54	0.71	0.94	0.24
<i>phi031</i>	0.79	185	223	31	4.82	0.4	0.81	0.5
<i>phi064</i>	0.93	62	112	27	14.19	0.67	0.82	0.19
<i>phi072</i>	0.74	140	168	14	3.87	0.56	0.79	0.29
<i>phi078</i>	0.85	114	198	30	6.77	0.5	0.92	0.46
<i>phi089</i>	0.66	78	100	17	2.94	0.29	0.65	0.55
<i>phi093</i>	0.8	277	299	12	4.95	0.59	0.73	0.2
<i>phi109188</i>	0.78	138	198	12	4.58	0.41	0.79	0.49
<i>phi116</i>	0.8	153	301	11	5.03	0.59	0.76	0.23
<i>phi96100</i>	0.83	151	174	15	6.04	0.71	0.82	0.14
Promedio	0.85	129.75	210.70	28.90	9.29	0.54	0.84	0.36
Desviación estándar	0.08	50.56	57.91	12.44	5.44	0.12	0.08	0.14

Anexo 5: Rango de *Bins* obtenidos para los 20SSR con *MsatAllele* y corroboración manual. Se muestran los microsatélites, el valor del *bin* mínimo (Min) y el valor máximo (Max).

SSR	Min	Max
<i>bnlg1046</i>	169	217
<i>bnlg1191</i>	179	235
<i>bnlg1194</i>	152	195
<i>bnlg1265</i>	194	246
<i>bnlg1449</i>	82	166
<i>bnlg1523</i>	139	269
<i>bnlg1740</i>	115	174
<i>bnlg1890</i>	94	179
<i>bnlg1917</i>	88	178
<i>bnlg2305</i>	165	208
<i>phi015</i>	77	100
<i>phi031</i>	185	222
<i>phi064</i>	72	112
<i>phi072</i>	141	166
<i>phi078</i>	114	131
<i>phi089</i>	81	95
<i>phi093</i>	283	299
<i>phi109188</i>	151	174
<i>phi116</i>	138	177
<i>phi96100</i>	269	301