

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias Biológicas y Ambientales

Caracterización de polimorfismos en el gen LDC (Lysine Decarboxylase) y su influencia en la concentración de alcaloides en chocho (*Lupinus mutabilis*).

Andrea Marina Lazzarini Hermida

Ingeniería en Biotecnología

Trabajo de fin de carrera presentado como requisito
para la obtención del título de
Ingeniera en Biotecnología

Quito, 17 de diciembre de 2021

UNIVERSIDAD SAN FRANCISCO DE QUITO USFQ

Colegio de Ciencias Biológicas y Ambientales

HOJA DE CALIFICACIÓN DE TRABAJO DE FIN DE CARRERA

Caracterización de polimorfismos en el gen LDC (Lysine Decarboxylase) y su influencia en la concentración de alcaloides en chocho (*Lupinus mutabilis*).

Andrea Marina Lazzarini Hermida

Nombre del profesor, Título académico

María de Lourdes Torres, PhD.

Quito, 17 de diciembre de 2021

© DERECHOS DE AUTOR

Por medio del presente documento certifico que he leído todas las Políticas y Manuales de la Universidad San Francisco de Quito USFQ, incluyendo la Política de Propiedad Intelectual USFQ, y estoy de acuerdo con su contenido, por lo que los derechos de propiedad intelectual del presente trabajo quedan sujetos a lo dispuesto en esas Políticas.

Asimismo, autorizo a la USFQ para que realice la digitalización y publicación de este trabajo en el repositorio virtual, de conformidad a lo dispuesto en la Ley Orgánica de Educación Superior del Ecuador.

Nombres y apellidos: Andrea Marina Lazzarini Hermida

Código: 00201450

Cédula de identidad: 0106909989

Lugar y fecha: Quito, 17 de diciembre de 2021

ACLARACIÓN PARA PUBLICACIÓN

Nota: El presente trabajo, en su totalidad o cualquiera de sus partes, no debe ser considerado como una publicación, incluso a pesar de estar disponible sin restricciones a través de un repositorio institucional. Esta declaración se alinea con las prácticas y recomendaciones presentadas por el Committee on Publication Ethics COPE descritas por Barbour et al. (2017) Discussion document on best practice for issues around theses publishing, disponible en <http://bit.ly/COPETheses>.

UNPUBLISHED DOCUMENT

Note: The following capstone project is available through Universidad San Francisco de Quito USFQ institutional repository. Nonetheless, this project – in whole or in part – should not be considered a publication. This statement follows the recommendations presented by the Committee on Publication Ethics COPE described by Barbour et al. (2017) Discussion document on best practice for issues around theses publishing available on <http://bit.ly/COPETheses>

RESUMEN

Los alcaloides quinolizidínicos (AQ) son metabolitos secundarios tóxicos que confieren un sabor amargo a las semillas de chocho (*Lupinus mutabilis*). Para reducir la concentración de AQ en las semillas al valor umbral de inocuidad alimentaria (<0.02%) se suelen realizar costosos procesos de desamargado, limitando la rentabilidad y expansión del cultivo. Por ello, es fundamental investigar los genes involucrados en la ruta biosintética de los AQ. Entre los genes de interés, se destaca el gen codificante para la Lysine Decarboxylase (LDC). En la especie *Lupinus angustifolius*, el gen LDC ha sido identificado como precursor en la biosíntesis de AQ y variaciones en sus secuencias de ADN han sido asociadas a una alta y baja producción de estos compuestos. Respecto a *L. mutabilis*, hasta el momento el gen LDC no ha sido estudiado, por tanto, el objetivo del presente estudio fue la caracterización de polimorfismos en el gen LDC en accesiones de *L. mutabilis* con un alto y bajo contenido de AQ. Para esto, se analizó el gen LDC en 7 accesiones con alto y 7 accesiones con bajo contenido de AQ en *L. mutabilis* a través de secuenciación por nanoporo. Las secuencias consenso obtenidas para cada accesión se alinearon y tradujeron para identificar polimorfismos y/o cambios en la secuencia de aminoácidos. En total se identificaron 9 polimorfismos y todos correspondieron a polimorfismos de un solo nucleótido (SNP). De estos, 3 SNPs resultaron ser de interés ya que causaron cambios no sinónimos en la secuencia de aminoácidos (posición 92: Lys por Ser/Thr y posición 120: Cys por Phe). Adicionalmente se encontró 2 SNPs causantes de mutaciones sinónimas de aminoácidos (posición 236: no sustitución de Asp y posición 303: no sustitución de Gly). Estos resultados podrían ser de gran utilidad para determinar la funcionalidad de estos SNPs y así validar su uso como posibles marcadores moleculares asociados a una baja acumulación de AQ en las semillas, acelerando así el mejoramiento del cultivo de *L. mutabilis*.

Palabras clave: Alcaloides quinolizidínicos, lisina descarboxilasa (LDC), SNPs, *Lupinus mutabilis*, secuenciación por nanoporo.

ABSTRACT

Quinolizidine alkaloids (QA) are toxic secondary metabolites that confer a bitter taste to lupine seeds (*Lupinus mutabilis*). To reduce the concentration of QA in the seeds to values below the food safety threshold (<0.02%), expensive debittering processes are usually carried out, thus reducing the profitability of the crop and limiting its expansion. Therefore, it is essential to investigate the genes involved in the biosynthetic pathway of QA. Among the genes of interest, the coding gene for Lysine Decarboxylase (LDC) stands out. In the species *Lupinus angustifolius*, the LDC gene has been identified as a precursor in the biosynthesis of QA and variations in its DNA sequences have been associated with a high and low production of these compounds. Regarding *L. mutabilis*, to date the LDC gene has not been studied, therefore, the objective of the present study was the characterization of polymorphisms in the LDC gene in accessions of *L. mutabilis* with a high and low content of QA. For this, the LDC gene was analyzed in 7 accessions of *L. mutabilis* with high content and 7 accessions with low content of QA through nanopore sequencing. The consensus sequences obtained for each accession were aligned and translated to identify polymorphisms and / or changes in the amino acid sequence. A total of 9 polymorphisms were identified and all corresponded to single nucleotide polymorphisms (SNP). Of these, 3 SNPs were of interest as they non-synonymous changes in amino acids (position 92: substitution of Lys by Ser or Thr; position 120: substitution of Cys by Phe). In addition 2 SNPs caused 2 synonymous mutations (position 236: no Asp substitution; position 303: no Gly substitution). These results could be very useful to determine the functionality of these SNPs and thus validate their use as possible molecular markers associated with a low accumulation of QA in the seeds, thus accelerating the improvement of the *L. mutabilis* crop.

Key words: Quinolizidine alkaloids, lysine decarboxylase (LDC), SNPs, *Lupinus mutabilis*, Nanopore sequencing.

TABLA DE CONTENIDOS

1.	Introducción	11
2.	Metodología	15
	2.1 <i>Accesiones de L. mutabilis</i>	15
	2.2 <i>Cuantificación de alcaloides quinolizidínicos</i>	15
	2.3 <i>Diseño de cebadores para PCR</i>	15
	2.4 <i>Amplificación del gen LDC por PCR</i>	16
	2.5 <i>Secuenciación de amplicones</i>	16
	2.6 <i>Procesamiento de las lecturas generadas</i>	17
	2.7 <i>Identificación de polimorfismos</i>	17
	2.8 <i>Identificación de variaciones en la secuencia de aminoácidos</i>	17
3.	Resultados	18
	3.1 <i>Cuantificación de alcaloides quinolizidínicos</i>	18
	3.2 <i>Diseño de cebadores y amplificación del gen LDC</i>	18
	3.3 <i>Secuenciación de amplicones mediante Nanopore Sequencing</i>	19
	3.4 <i>Procesamiento de las lecturas</i>	19
	3.5 <i>Identificación de polimorfismos</i>	20
	3.6 <i>Identificación de cambios en la secuencia de aminoácidos</i>	21
4.	Discusión.....	22
5.	Conclusiones	26
6.	Tablas	27
7.	Figuras	29
8.	Referencias bibliográficas.....	33
	Anexos.....	37

ÍNDICE DE TABLAS

Tabla 1. Porcentaje de AQ s en 14 accesiones de <i>L.mutabilis</i> provenientes de una colección del INIAP	27
Tabla 2. Secuencia del par de cebadores diseñado y parámetros de calidad evaluados.....	27
Tabla 3. Rendimiento del ensayo de secuenciación con <i>Nanopore Sequencing MinION™</i> ...	27

ÍNDICE DE FIGURAS

Figura 1. Geles de agarosa con la amplificación del gen LDC en las 14 accesiones de <i>L. mutabilis</i>	29
Figura 2. Histograma del tamaño y el número de lecturas generadas en el ensayo de secuenciación	30
Figura 3. Comparación entre la calidad promedio y el tamaño promedio de las lecturas antes y después del procesamiento con <i>Nanoplot</i>	31
Figura 4. Alineación de las secuencias de aminoácidos obtenidas para el gen LDC en las 14 accesiones de <i>L. mutabilis</i> analizadas	31
Figura 5. Alineación de las secuencias de aminoácidos obtenidos tras la traducción de las secuencias de nucleótidos del gen LDC en las 14 accesiones de <i>L.mutabilis</i>	32

ÍNDICE DE ANEXOS

Anexo 1: Histograma con la distribución de valores de cuantificación de alcaloides quinolizidínicos para las 54 accesiones de chocho	37
Anexo 2: Secuencias consenso generadas del gen LDC para las 14 accesiones evaluadas. .	37
Anexo 3: Posición de los SNPs identificados en la alineación de las secuencias consenso del gen LDC en las 14 accesiones de <i>L. mutabilis</i>	41
Anexo 4: Alineación de la secuencia de aminoácidos derivados de la traducción de la secuencia de nucleótidos del gen LDC entre las 14 accesiones de <i>L. mutabilis</i>	44

1. INTRODUCCIÓN

Lupinus mutabilis, comúnmente conocido como tarwi, chocho o lupino, es una leguminosa de Los Andes, originaria al norte de Perú y el Sur de Ecuador (Jacobsen and Mujica, 2006; Atchison et al., 2016). Debido a su valor nutricional, sus semillas son principalmente destinadas a la dieta humana y son consumidas como snack directo, o utilizados en la elaboración de ensaladas, harinas, sustitutos de leches, entre otros. Adicionalmente, el chocho es muy utilizado para alimentación de animales, fines ornamentales, como pesticida o con fines medicinales por la presencia de alcaloides (Villacrés et al., 2003; Camarena et al., 2012). El interés en esta planta radica en la calidad nutricional de su semilla y llega a ser comparable o incluso superior al resto de lupinos y otros granos como la soya, pues, se ha reportado que mantiene el mayor porcentaje de proteínas y lípidos en sus semillas, alcanzando niveles máximos de 50% y 24% en peso seco respectivamente (Calgari, 2000; Güemes-Vera et al., 2008). Además, es la única especie del género *Lupinus* que posee isoflavonoides que son reconocidos por su actividad antioxidante (Ranilla et al., 2009). A pesar de esto, en el país la producción de esta leguminosa es relativamente baja en comparación otros países, teniendo alrededor de 6,000 hectáreas (Mercado et al., 2018). Una de las principales limitaciones para la expansión de su cultivo a gran escala es la presencia de compuestos anti-nutricionales en su semilla, siendo los principales los alcaloides quinolozidínicos (AQ). Estos compuestos confieren a las semillas un sabor amargo y resultan ser tóxicos (Khan et al., 2015).

Hasta el momento, la ruta de biosíntesis de los AQ ha sido dilucidada parcialmente y se conoce que la molécula precursora de la ruta es la lisina, la cual es descarboxilada por la acción de la enzima lisina descarboxilasa (LDC), generando como producto cadaverina (Bunsupa et al., 2012). La cadaverina es luego modificada por deshidrogenación, oxigenación, glicosilación o esterificación para producir cientos de estructuras de AQ (Ohmiya et al., 1995). La síntesis

de los AQ sucede en los cloroplastos de los tejidos aéreos y se transportan por el floema al sitio de acumulación que son las semillas (Wink & Witte, 1984). Los AQ son importantes para las plantas ya que les permite repeler a predadores al ser directamente tóxicos o mutagénicos para ellos (Wink, 2003). El chocho de forma natural puede tener hasta 70 AQ diferentes que son tóxicos, especialmente la lupanina y la esparteína (Ruiz, 1978). Precisamente, cada especie de lupino presenta un perfil específico de AQ y las semillas serán más o menos amargas en dependencia del tipo y la concentración de AQ que esté presente. Es importante mencionar que la concentración de AQ también depende del genotipo, el estrés biótico/abiótico y las condiciones ambientales (Gulisano, 2019). Para que la semilla sea segura y pueda ser destinada para alimentación, la concentración de AQ debe mantenerse bajo el umbral industrial fijado de 0.02% (peso seco) (Frick, 2018). Una variedad representativa de chocho en el Ecuador (INIAP 450 Andino) posee un porcentaje de AQ de 3.92% que supera el umbral seguro, por tanto, se deben llevar a cabo un proceso de eliminación de alcaloides que consta de un paso de hidratación de 14 horas, seguido de cocción por 45 minutos y desamargado con agua corriente durante 3 días. Esto eleva el costo de producción (\$1,363 por hectárea en el 2010) lo cual limita su escala a producción industrial (Caicedo et al., 2010).

Para comprender los factores que causan una alta o baja concentración de AQ en las semillas, se han realizado una serie de investigaciones para identificar los genes responsables que causan esta variación fenotípica (Abraham et al., 2019). Los esfuerzos se centraron en identificar a los genes involucrados en la biosíntesis de AQ en *L. angustifolius* y *L. albus*. En este proceso se caracterizaron 3 genes candidatos: *L. angustifolius* lisina/ornitina descarboxilasa (LaL/ODC), *L. angustifolius* amina de cobre oxidasa (LaCAO) y *L. albus* 13 α -hidroxilupanina O-tigloiltransferasa (LaHMT) (Cowling, Huyghe & Swiecicki, 1998). En estudios posteriores se encontró que en las accesiones amargas (altas en AQ) había una alta expresión del gen LaL/ODC y un elevado nivel de cadaverina, mientras que en las accesiones

dulces (bajas en AQ) la expresión de este gen y los niveles de cadaverina eran reducidos, sugiriendo que este gen actúa como regulador de la síntesis de AQ (Bunsupa et al., 2012; Frick, 2018). Estos hallazgos vuelven extremadamente interesante al gen LDC para ser investigado en especies cercanas, pues, se considera que el gen LaL/ODC es un buen indicador del contenido de AQ entre accesiones amargas y dulces en *L. angustifolius* (Frick, 2018). Hasta la fecha, este tipo de estudios no se han realizado en *L. mutabilis* debido a que los recursos genéticos para esta especie son muy limitados ya que en las bases de datos se cuenta principalmente con dianas moleculares de ARN ribosómico (Gulisano, 2019). Es así como, la información disponible para *L. angustifolius* (mapas genéticos, conjuntos de transcriptomas y proteomas, locus de rasgos cuantitativos y marcadores moleculares) puede potencialmente explorarse para la mejora de *L. mutabilis* (Wolko et al., 2011; Abraham et al., 2019).

Los nuevos avances en las tecnologías de secuenciación ahora brindan una oportunidad realista para superar la escasez de información genómica y podrían proporcionar información valiosa sobre las variantes en *L. mutabilis* al permitir detectar polimorfismos presentes (Abraham et al., 2019). *Oxford Nanopore Technologies* (ONT) es una herramienta de secuenciación atractiva por el hecho de que permite implementar sistemas de *multiplexing* que reduce considerablemente el costo, además, se destaca por generar lecturas de largo tamaño, a tiempo real, y con la precisión necesaria para detectar variantes (Heather & Chain, 2016). El funcionamiento del sistema se basa en nanoporos, donde el ADN monocatenario es conducido a través de una membrana por canales iónicos, de forma que la interrupción de la corriente es proporcional a la longitud del ácido nucleico, asegurando una precisión del sistema del 99.17% a una profundidad de 60x, lo cual es suficiente para detectar incluso polimorfismos simples, como es el caso de los polimorfismos de un solo nucleótido (SNP) (ONT, 2021). Hay gran interés en detectar SNPs por el hecho de que son la clase de polimorfismo más abundante en

los genomas de plantas y pueden llegar a ser los responsables de causar variación fenotípica (Kage et al., 2016).

El objetivo de este estudio fue analizar las secuencias del gen LDC en accesiones de *L. mutabilis* con alto y bajo contenido de AQ para identificar la posible presencia de polimorfismos que puedan estar relacionados con la concentración de AQ. Para lograr este objetivo se amplificó la región del gen LDC en accesiones de *L. mutabilis*. Los amplicones obtenidos fueron secuenciados mediante *nanopore sequencing* (*Oxford Nanopore Technologies*), lo cual generó una secuencia consenso por cada accesión. Estas secuencias consenso fueron comparadas entre las accesiones con alto y bajo contenido de AQ para determinar polimorfismos y verificar si estos causan un cambio en la secuencia de aminoácidos. Esta información puede ser útil para el desarrollo de marcadores moleculares funcionales para el gen LDC en *L. mutabilis* que podrán facilitar la selección de líneas parentales en cruces y permitir una discriminación entre accesiones amargas o dulces. De esta manera, se podría acelerar la selección de las mejores progenies que presenten concentraciones reducidas de AQ y que se ajusten al umbral industrial que requiere semillas con concentraciones de AQ iguales o inferiores a 0.02% (Kroc et al., 2019).

2. METODOLOGÍA

*2.1 Accesiones de *L. mutabilis**

El material vegetal empleado forma parte del Banco de Germoplasma de Chocho del Instituto Nacional de Investigaciones Agropecuarias (INIAP). En un estudio previo, realizado en el Laboratorio de Biotecnología Vegetal de la USFQ, se cuantificó el porcentaje de AQ en semillas de 54 accesiones de chocho por titulación ácido-base. Los resultados generaron una curva de distribución normal (Anexo 1). Para identificar variaciones en las secuencias de ADN asociadas a una alta y baja acumulación de AQ, se seleccionó a las 10 accesiones con el mayor y menor porcentaje de AQ (ubicadas en los extremos contrapuestos de la curva de distribución). Adicionalmente, se incluyeron 4 accesiones catalogadas por el INIAP como altas (FR9 P3 y FR11 P8) y bajas en contenido de AQ (Proinpa P5 y Proinpa P8). La cuantificación de AQ de estas 4 accesiones se realizó en este estudio. De esta manera, un total de 14 accesiones fueron seleccionadas para la secuenciación del gen LDC y posterior detección de polimorfismos.

2.2 Cuantificación de alcaloides quinolizidínicos

Para la cuantificación porcentual de AQ en semilla de las 4 accesiones (FR9 P3, FR11 P8, Proinpa P5 y Proinpa P8) se empleó el método de titulación ácido-base. Para ello, se pesó 0.2 mg de semillas secas pulverizadas y se mezcló con 10 ml de agua destilada. Los tubos se llevaron a cocción a 60°C durante 1 hora y se filtró el precipitado. Se adicionó 0.25 ml de fenolftaleína (indicador) y se añadió la solución de hidróxido de sodio (NaOH 0.1N) hasta observar el viraje a un color rosa intenso. Se registró el volumen utilizado de NaOH y se empleó en la fórmula de estequiometría para obtener el porcentaje de AQ:

$$\% \text{ Alcaloides (p/p)} = \frac{\text{Normalidad NaOH (0.1 N)} \times \text{Volumen gastado NaOH} \times \text{peso equivalente lupanina (24.8)}}{\text{peso de la muestra (0.2 g)}}$$

2.3 Diseño de cebadores para PCR

Con la finalidad de amplificar el gen LDC en *L. mutabilis* se diseñaron cebadores específicos para la especie. Para ello, se emplearon secuencias del transcriptoma de *L. mutabilis* disponibles en la base de datos del NCBI (ID: SRR3423049; SRR3423050; SRR3423051). Estas secuencias se utilizaron para reconstruir el gen LDC y la secuencia resultante se introdujo en *Primer3* v.4.0 para generar opciones de cebadores. Se comprobó la calidad de los cebadores con *mFold* (Markham, Zuker & Keith, 2008) y *Primer Blast* (NCBI, USA) y se seleccionó el par de cebadores con las mejores características para PCR convencional (Borah, 2011).

2.4 Amplificación del gen LDC por PCR

La reacción de PCR del gen LDC en *L. mutabilis* consistió en: PCR Buffer (1X), cloruro de magnesio MgCl₂ (2.0 mM), dNTPs (0.3 mM), enzima Polimerasa *Taq Platinum* (Invitrogen, USA) (1.0 U), cebadores (0.5 μM), ADN (20 ng). El volumen final de la reacción fue de 30 μl. El termociclado se realizó a las siguientes condiciones: 94 °C por 5 minutos, seguido de 35 ciclos de 94 °C por 30 segundos, 64 °C por 30 segundos y 72 °C por 2 minutos, con una extensión final de 72 °C por 5 min. Los amplicones se visualizaron por electroforesis (100V, 30 minutos) en gel de agarosa (1%). Los productos de PCR se almacenaron a -20°C.

2.5 Secuenciación de amplicones

Los amplicones de las 14 accesiones se secuenciaron por *multiplexing* mediante tecnología de *Oxford Nanopore* (ONT). El protocolo utilizado fue *Native barcoding amplicons* (*EXP-NBD104*, *EXP-NBD114* and *SQK-LSK 109*) de ONT y se siguieron las especificaciones del fabricante. Con la finalidad de maximizar el rendimiento del ensayo de secuenciación, adicional al gen LDC, la reacción incluyó 4 genes más involucrados en la biosíntesis de AQ (RAP2-7, DHDPS, LaAT y LaCAO). El presente estudio se centrará en el gen LDC, mientras que los 4 genes restantes serán analizados como parte de otros proyectos de investigación. Para cada accesión, se creó un pool con los productos de PCR de los 5 genes amplificados. Los pools fueron purificados con *beads* magnéticas (NEBNext, USA). Se añadió un *barcode* de secuencia

única a cada pool para preparar la librería. El cargado de la librería se realizó en las celdas de flujo R9.4.1 en el dispositivo *MinION™ Mk1C* con software *MinKNOW™*. La cantidad de ADN cargada fue 65.51 fmol cuantificado por *Qubit* (ThermoFisher, USA). La secuenciación se ejecutó por 24 horas y se seleccionó el algoritmo *high-accuracy basecalling*.

2.6 Procesamiento de las lecturas generadas

Con la finalidad de separar las lecturas de cada accesión marcada por *barcodes* en carpetas diferentes, se realizó el procedimiento de *demultiplexing* empleando el software de *MinKNOW™*. A continuación, los adaptadores y *barcodes* fueron removidos de las lecturas utilizando *Porechop* v.0.2.4, para lo cual se fijó un umbral de identidad del 100%. De esta manera, se logró que el procedimiento de reconocimiento y remoción de *barcodes* y adaptadores fuera mucho más preciso. Adicionalmente, se realizó un filtrado de las lecturas por calidad y tamaño utilizando el software de *BBMap* (JGI, 2020). Con ello, se logró eliminar secuencias de calidad $< q10$ y de tamaño < 200 pb. Finalmente, se realizó una evaluación de calidad y de tamaños de las lecturas antes y después del procesamiento con *Nanoplot* v.1.30.1.

2.7 Identificación de polimorfismos

Las lecturas procesadas se introdujeron en *Geneious prime* para ser mapeadas tomando como referencia la secuencia del gen LDC de *L. angustifolius* (LOC109327937) y así obtener una secuencia consenso para cada accesión. La profundidad de cobertura fue obtenida utilizando el algoritmo de *Geneious prime*. Como paso final, las secuencias consenso generadas fueron alineadas en *UGENE* v.40.1 para la identificación de polimorfismos.

2.8 Identificación de variaciones en la secuencia de aminoácidos

Las 14 secuencias consenso generadas se introdujeron en *Expasy* (SIB, 2020) para traducirlas hacia una secuencia de aminoácidos. Posteriormente, las 14 secuencias de aminoácidos se alinearon en *UGENE* para identificar variaciones. Se realizó una comparación con la secuencia de aminoácidos de la referencia de *L. angustifolius* (LOC109327937).

3. RESULTADOS

3.1 Cuantificación de alcaloides quinolizidínicos

La técnica de cuantificación de AQ mediante titulación ácido-base usando hidróxido de sodio permitió determinar exitosamente el porcentaje de AQ por gramo de muestra en las 4 accesiones cuantificadas (Proinpa P5, Proinpa P8, FR9 P3 y FR11 P8). Por un lado, Proinpa P5 y Proinpa P8, catalogadas por el personal del INIAP como accesiones dulces, mostraron un porcentaje de AQ de 3.47% y 3.85% respectivamente, mientras que las accesiones FR9 P3 y FR11P8, catalogadas como accesiones amargas, mostraron valores más elevados de 4.55% y 4.57% respectivamente. La Tabla 1 muestra los valores promedio y la desviación estándar del porcentaje de AQ de las 4 accesiones cuantificadas en el presente estudio, así como los valores correspondientes a las 10 accesiones cuantificadas en el estudio previo. El rango del porcentaje de AQ del total de las 14 accesiones va desde un valor mínimo de 2.84% hasta un valor máximo de 4.91%. Del total de las 14 accesiones, se clasificaron a 7 accesiones como bajas en AQ (706, 707, 712, 723, 742, Proinpa P5 y Proinpa P8) y a 7 accesiones como altas en AQ (176, 188, 190, 192, 193, FR9 P3 y FR11 P8). El rango del porcentaje en las accesiones bajas en AQ (2.84% - 3.85%; media: 3.39%) es menor al rango las accesiones altas en AQ (4.55% - 4.91%; media: 4.70%).

3.2 Diseño de cebadores y amplificación del gen LDC

El par de cebadores diseñados logró amplificar la región del gen LDC en las 14 accesiones seleccionadas. La secuencia del cebador *forward* y el cebador *reverse*, junto con su longitud (bp), temperatura de *melting* (Tm) y los parámetros de calidad evaluados, se presentan en la Tabla 2. Los parámetros de calidad indican que los cebadores tienen un porcentaje de GC ideal (entre 40-60%) (Borah, 2011) y además muestran muy baja o nula estabilidad para formar

dímeros, provocar emparejamiento de bases en el extremo 3' o formar *hairpins* en el extremo 3'. Adicionalmente, el valor ΔG para cada cebador tiene un valor positivo o poco negativo, representando baja estabilidad para que los cebadores formen estructuras secundarias no deseadas (Zuker, 2003). En base a esto, los cebadores propuestos presentaron muy buenas características para la amplificación del gen LDC. Los resultados de amplificación y electroforesis en geles de agarosa se muestran en la Figura 1. Las bandas obtenidas se encuentran en el tamaño esperado (1.2 Kb). En ningún caso hubo amplificación del control negativo.

3.3 Secuenciación de amplicones mediante Nanopore Sequencing

La reacción de secuenciación fue de tipo *multiplex*, en la cual los productos de PCR de las 14 accesiones fueron marcados con *barcodes* específicos y secuenciados en la misma corrida. El rendimiento del ensayo de secuenciación generó 4.23 Gb de información en un tiempo de 24 horas. Los detalles adicionales del rendimiento de la corrida se detallan en la Tabla 3. El histograma de la cantidad de lecturas generadas y el tamaño de las mismas se muestra en la Figura 2. Cabe mencionar que el tamaño esperado del gen LDC es de 1.2 Kb, por lo tanto, las lecturas del gen LDC se ubican en el rango entre 60 pb hasta 1.2 Kb aproximadamente. Las lecturas con tamaños superiores a 1.2 Kb corresponden a los otros genes incluidos en la reacción de secuenciación (RAP2-7, DHDPS, LaAT y LaCAO), que serán analizados en estudios posteriores.

3.4 Procesamiento de las lecturas

El algoritmo de *demultiplexing* utilizado por el *MinION™ Mk1C* logró identificar correctamente a los 14 *barcodes* utilizados y clasificó correctamente a todas las lecturas de cada *barcode* en carpetas separadas. Posteriormente, la remoción de *barcodes* y adaptadores fue realizado con *Porechop*. El reconocimiento de dichas secuencias fue preciso y altamente eficiente, dado que únicamente alrededor del 2% de la información suministrada fue clasificada

como secuencias sin *barcodes* o quiméricas (amplicones unidos en tándem, separados por *barcodes* intermedios).

Luego del filtrado de calidad con *BBMap* se conservó únicamente las lecturas con calidad $\geq q10$, teniendo un promedio de recuperación de información del 88% por cada muestra. La depuración por tamaño con *BBMap* se conservó únicamente las lecturas con longitud ≥ 200 pb, a excepción de las accesiones 177 y 188 donde se conservaron lecturas de con longitud ≥ 150 pb. De esta manera, se garantizó que la recuperación de información sea superior al 70% en todos los casos.

El *input* total de lecturas para ser procesadas fue de 4.1 millones de lecturas y el output obtenido tras el procesamiento fue de 3.1 millones de lecturas, teniendo así una recuperación del 76% de la información. La comparación del tamaño de las lecturas contra la calidad promedio antes y después del procesamiento fue realizada con *Nanoplot* y se presenta en la Figura 3. El procesamiento de las lecturas permitió la eliminación de lecturas muy cortas (< 200 pb) y de baja calidad ($< q10$), lo cual incrementó la calidad promedio de 12.9 a 14.5 Phred, alcanzando valores máximos de 26 Phred.

3.5 Identificación de polimorfismos

Con el objetivo de separar las secuencias del gen LDC del resto de genes amplificados para cada accesión, las lecturas procesadas fueron mapeadas frente a la secuencia consenso del gen LDC de *L. angustifolius*. En promedio, se introdujo 222,048 lecturas por accesión a *Geneious prime* para ser mapeadas, de las cuales, alrededor de 104,000 lecturas (47%) se alinearon con la secuencia de referencia de *L. angustifolius*. A continuación, las lecturas recuperadas fueron utilizadas para el ensamblado de las secuencias consenso (profundidad de cobertura: 18,433x - 47,271x; media: 31,506x; DS: 7,971.24). Las 14 secuencias consenso generadas se incluyen en el Anexo 2. La alineación de estas secuencias consenso en *UGENE* permitió identificar la presencia de 9 polimorfismos (Anexo 3). Todos los polimorfismos

detectados correspondieron a polimorfismos de un solo nucleótido (SNP). La posición donde se encuentran estos SNPs y al grupo al que corresponden se muestra en la Figura 4.

3.6 Identificación de cambios en la secuencia de aminoácidos

La alineación en *UGENE* de las 14 secuencias de aminoácidos traducidas a partir de las secuencias consenso generadas se incluye en el Anexo 4. La alineación permitió evaluar si los polimorfismos identificados anteriormente generaron cambios en la secuencia de aminoácidos. En total se detectaron 3 posiciones donde ocurrió una sustitución de aminoácidos. En la posición 92 hubo una sustitución del aminoácido Lisina (K) por el aminoácido Serina (S) en 3 accesiones con alto contenido de AQ y la sustitución de Lisina (K) por Treonina (T) en 1 accesión con alto contenido de AQ. En la posición 120 hubo una sustitución del aminoácido Cisteína (C) por el aminoácido Fenilalanina (F) en dos accesiones con bajo contenido de AQ. En la posición 394 hubo una sustitución del aminoácido Leucina (L) por el aminoácido Fenilalanina (F) en una accesión con alto contenido de AQ. Con respecto a la comparación de la secuencia de aminoácidos de referencia de *L. angustifolius*, se encontró que del total de los 416 aminoácidos alineados tan solo hubo una variación en 28 ubicaciones de aminoácidos, denotando así una conservación del 93.3% entre *L. mutabilis* y *L. angustifolius*. La posición de las variaciones y el tipo de aminoácido sustituido se muestra en la Figura 5.

4. DISCUSIÓN

El alto contenido de AQ disminuye el valor del chocho para ser utilizado como fuente de alimento, siendo esta una de sus principales limitantes para expandir su cultivo a escala industrial (Villacrés et al., 2009). Los programas de mejoramiento para el chocho se enfocan en identificar polimorfismos que estén asociados a una baja o alta acumulación de AQ (Frick, 2018), por ende, es crucial analizar accesiones de esta planta que presenten estos fenotipos diferentes, es decir, fenotipos dulces con bajo contenido de AQ y fenotipos amargos con alto contenido de AQ en sus semillas.

La cuantificación del porcentaje de AQ en una muestra del banco de germoplasma de chocho del INIAP permitió seleccionar las 7 accesiones con el menor contenido de AQ (promedio de 3.39%) y las 7 accesiones con el mayor contenido de AQ (promedio de 4.70%). Estos valores exceden el umbral industrial establecido como seguro (0.02%) (Frick et al., 2017), no obstante, se encuentran dentro del rango esperado, pues, las semillas de *L. mutabilis* de forma natural presentan un rango de AQ entre 0.02 a 4.45% (Mori & Paz, 2008). Al considerar la correlación entre la concentración de AQ y el sabor amargo que estas sustancias confieren a las semillas, se consideraría que todas las accesiones evaluadas en este estudio presentarían el fenotipo de sabor amargo, pues, un sabor dulce se obtiene a una concentración entre 0.01 y 0.05% (Jiménez-Martínez et al., 2007; Cortés-Avenidaño et al., 2020). Sin embargo, al realizar una prueba cualitativa de degustación, las accesiones PROINPA tuvieron un sabor dulce a pesar de su alto porcentaje de AQ. Esto podría explicarse por las diferencias en el perfil de AQ que pueden diferir dentro de la misma especie (Hatzold et al., 1983; Ramírez-Betancourt et al., 2021). Precisamente, el AQ que tiene un mayor aporte hacia un sabor amargo es la esparteína (Muzquiz et al., 2004), por lo que se podría pensar que las accesiones PROINPA posiblemente son accesiones dulces que mantienen concentraciones bajas de este compuesto. La técnica de

cuantificación por titulación ácido-base limita este tipo de análisis por lo que se sugiere que se implemente un método de cuantificación como cromatografía de gas líquido acoplado a espectrometría de masas para poder identificar y cuantificar los tipos de AQ presentes en las semillas de chocho (Ramírez-Betancourt et al., 2021). Adicionalmente, se propone analizar otras variedades como “Inti” o “2150” que contienen porcentajes de AQ reducidos (0.0075% y 0.015%, respectivamente) (Gross et al., 1988) para así identificar los diferentes tipos de polimorfismos asociados a causar ese fenotipo (Gulisano et al., 2019).

En cuanto al método de secuenciación, para detectar polimorfismos de un solo nucleótido o SNPs, se requiere que el sistema de secuenciación provea una alta precisión y baja tasa de error para poder construir secuencias consenso de alta fidelidad (Heather and Chain, 2016). En este caso, se realizó secuenciación de amplicones por *multiplexing* y este método resultó adecuado para la identificación de SNPs. Esto se debe principalmente a que la reconstrucción de cada una de las 14 secuencias consenso empleó múltiples lecturas alineadas entre sí, logrando una cobertura promedio de 31,506x lo que aumenta la confiabilidad de la secuencia obtenida y disminuye la probabilidad de errores (Wang et al., 2012). Incluso, se ha reportado que este método puede ser considerado como una alternativa a métodos más precisos como la secuenciación por Sanger que mantiene una precisión del 99.99% (Loiu et al., 2020). Otro parámetro que se debe evaluar para determinar la fidelidad de las secuencias consenso generadas es la profundidad de cobertura alcanzada en el rendimiento del ensayo de secuenciación. En base a la literatura, una profundidad de cobertura de 100x es adecuada para establecer un genotipado preciso, e incluso una cobertura de 80x de profundidad se considera suficiente para generar una secuencia consenso fiable (Liao et al., 2019; Whitford et al., 2021). En el presente estudio, la profundidad de cobertura obtenida para cada secuencia consenso varió entre 18,433x y 47,271x, con un promedio de 31,506x ($DS = 7,971.24$). Estos valores sobrepasan la profundidad de cobertura mínima sugerida para generar secuencias consenso

fiables, de forma que se considera que el genotipado de las secuencias consenso generadas para cada accesión es preciso y confiable.

La alineación de las 14 secuencias consenso generadas permitió la detección de 9 SNPs en el gen LDC entre las accesiones con alto y bajo contenido de AQ de *L. mutabilis* (Anexo 3). Hay gran interés en detectar SNPs por el hecho de que se considera que son un tipo de polimorfismo que desempeñan un papel importante en la activación de la variación fenotípica (Thornsberry et al., 2001). Pues, los SNPs que son funcionales pueden contribuir directamente a la variación fenotípica y por ende son de suma importancia para el desarrollo de marcadores funcionales (Kage et al., 2016). Es bastante conocido que los SNPs que están involucrados en cambios no sinónimos de aminoácidos pueden conducir a variaciones fenotípicas; sin embargo, se ha demostrado que los SNPs que están involucrados en cambios sinónimos, donde no hay cambio o sustituciones de aminoácidos, también pueden contribuir a una variación del fenotipo (Hunt et al., 2009; Goymer, 2007).

Del total de 9 SNPs identificados en este estudio, se encontró que 3 de ellos están relacionados a 2 sustituciones no sinónimas de aminoácidos (posición 92 sustitución del aminoácido Lys por Ser o Thr y posición 120 sustitución del aminoácido Cys por Phe). La primera sustitución de aminoácidos (posición 92) es específica para 4 de las 7 accesiones altas en AQ y la segunda sustitución de aminoácidos (posición 120) es específica para las accesiones PROINPA únicamente. Estas sustituciones no sinónimas son de gran relevancia por el hecho de que pueden influir en la expresión génica, la conformación del ARNm o inclusive en la localización de las proteínas, dando como resultado un fenotipo diferente (Shastry, 2009). Sin embargo, para poder relacionar a estos SNPs con cambios fenotípicos se requiere de estudios más detallados, en especial sobre la estructura y funcionalidad de las proteínas resultantes (Bunsupa et al., 2012). Se ha demostrado que el cambio en los aminoácidos puede cambiar el patrón de unión entre los aminoácidos, afectando a las estructuras secundarias y terciarias de

las proteínas provocando alteraciones funcionales, como, por ejemplo, la abolición completa de la función de la proteína (Schaefer & Rost, 2012). Por tanto, para obtener conclusiones más informativas es necesario realizar modelamientos moleculares y predicciones de las estructuras de las proteínas (Bunsupa et al., 2012).

Por otro lado, del total de los 9 SNPs identificados en este estudio, se encontró 2 SNPs que están relacionados a 2 sustituciones sinónimas (posición 236 sin sustitución de Asp y posición 303 sin sustitución de Gly). La primera sustitución (posición 236) es específica para las 7 accesiones altas en AQ y la segunda sustitución (posición 303) es específica para 6 accesiones altas en AQ. Varios estudios han demostrado que los polimorfismos sinónimos, es decir, que no causan un cambio en la secuencia de aminoácidos, también pueden influir en la estabilidad y empalme del ARNm e incluso en el plegamiento de las proteínas, lo cual a su vez puede producir cambios significativos en el funcionamiento de las mismas (Hunt et al., 2009). Además, los SNPs que residen fuera de las regiones codificantes también deben ser analizados, como es el caso de la región promotora y los sitios de unión de los factores de transcripción (Hunt et al., 2009). Precisamente, en un estudio realizado por Frick en el 2018, se reportó que la región promotora del gen LDC en *L. angustifolius* difería entre las accesiones amargas y dulces con variaciones que incluyen SNPs, inserciones y deleciones; dando como resultado alteraciones en los sitios de anclaje de la maquinaria de transcripción específicos para cada accesión (Frick, 2018). Estas mutaciones “silenciosas” pueden tener impactos significativos sobre la expresión de genes y conformación de las proteínas, de forma que se propone que en futuros análisis se incluya la identificación de polimorfismos en regiones no codificantes.

5. CONCLUSIONES

En el presente estudio, se logró identificar polimorfismos de tipo SNP en el gen LDC en 7 accesiones de *L. mutabilis* con alto contenido y 7 accesiones con bajo contenido de AQ mediante *nanopore sequencing* (Oxford Nanopore Technologies). El sistema de secuenciación proporcionó una profundidad promedio de cobertura de 31,506x para generar secuencias consenso, de forma que se considera que este método permitió un genotipado preciso y confiable, siendo eficaz para la identificación de SNPs. En total se detectaron 9 SNPs, de los cuales 3 SNPs causaron 2 sustituciones no sinónimas en aminoácidos y 2 SNPs causaron 2 sustituciones sinónimas en aminoácidos. Estos hallazgos son un punto de partida valioso para comprender los factores genéticos asociados a la biosíntesis de AQ en *L. mutabilis*. Los SNPs identificados son un recurso útil para explorar con mayor profundidad estudios funcionales que permitan el desarrollo de marcadores moleculares. Para alcanzar esto, se sugiere ampliar la muestra analizada y evaluar la presencia de estos SNPs en accesiones que presenten fenotipos amargos, dulces e intermedios. Así se podrá establecer el grado de ligación entre el SNP y el rasgo de interés. Además, se propone llevar a cabo estudios de predicción de las estructuras tridimensionales y análisis de la localización de las proteínas resultantes para validar si los SNPs reportados causan cambios estructurales o funcionales en la enzima LDC. Y no menos importante, se debe considerar que este rasgo está dado por un sistema poligénico, de forma que es crucial investigar otros genes involucrados en la biosíntesis de AQ para poder generar mapas genéticos de alta resolución en *L. mutabilis* que se asocien con la concentración de AQ. Este enfoque dará paso al desarrollo de marcadores moleculares funcionales, los cuales representan una importante herramienta para mejorar la eficiencia de la selección y creación de líneas con genotipo mejorado, en este caso, semillas dulces que mantengan una baja acumulación de AQ. Esto permitirá el mejoramiento acelerado del cultivo de chocho.

6. TABLAS

Tabla 1. Porcentaje de AQ en 14 accesiones de *L. mutabilis* provenientes de una colección del INIAP. El valor promedio y la desviación estándar para cada accesión se obtuvo considerando 3 réplicas realizadas. *M* = media; *DS* = Desviación Estándar.

Accesiones bajas en alcaloides			Accesiones altas en alcaloides		
Muestra	<i>M</i> [%]	<i>DS</i>	Muestra	<i>M</i> [%]	<i>DS</i>
707	2.84	0.085	FR9 P3	4.55	0.068
742	3.15	0.111	FR11 P8	4.57	0.109
706	3.34	0.344	176	4.65	0.325
712	3.47	0.099	190	4.68	0.190
Proinpa P5	3.47	0.047	188	4.75	0.036
723	3.61	0.416	193	4.79	0.278
Proinpa P8	3.85	0.245	192	4.91	0.196

Tabla 2. Secuencia del par de cebadores diseñado y parámetros de calidad evaluados. *T_m* = Temperatura de *melting*; %GC = Guanina-Citocina; *Any_{th}* = Estabilidad de formación de dímeros; *3'_{th}* = Estabilidad de emparejamiento en el extremo 3'; *Hairpin* = Estabilidad de formación de estructuras secundarias; ΔG = Energía espontánea para estructuras secundarias.

	Longitud	<i>T_m</i>	gc%	<i>any_{th}</i>	<i>3'_{th}</i>	<i>hairpin</i>	ΔG_1	ΔG_2	Secuencia
Cebador forward	25 pb	60.22	44.00	0.00	0.00	0.00	-0.04	0.30	CTTCACTACTTCTCACAGAGGGTTT
Cebador reverse	25 pb	59.71	44.00	0.00	0.00	0.00	0.52	0.62	GATGTCGTATAAGCACCCATATTCG

Nota: El tamaño del producto es 1,234 pb

Tabla 3. Rendimiento del ensayo de secuenciación con *nanopore sequencing MinION™*.

Rendimiento del ensayo de secuenciación	
Tiempo de corrida	24 horas
Cantidad de información generada	4.23 Gb
Cantidad de lecturas	6.21 M
Longitud de lecturas N50	817 pb
Calidad promedio	12.9 Phred

7. FIGURAS

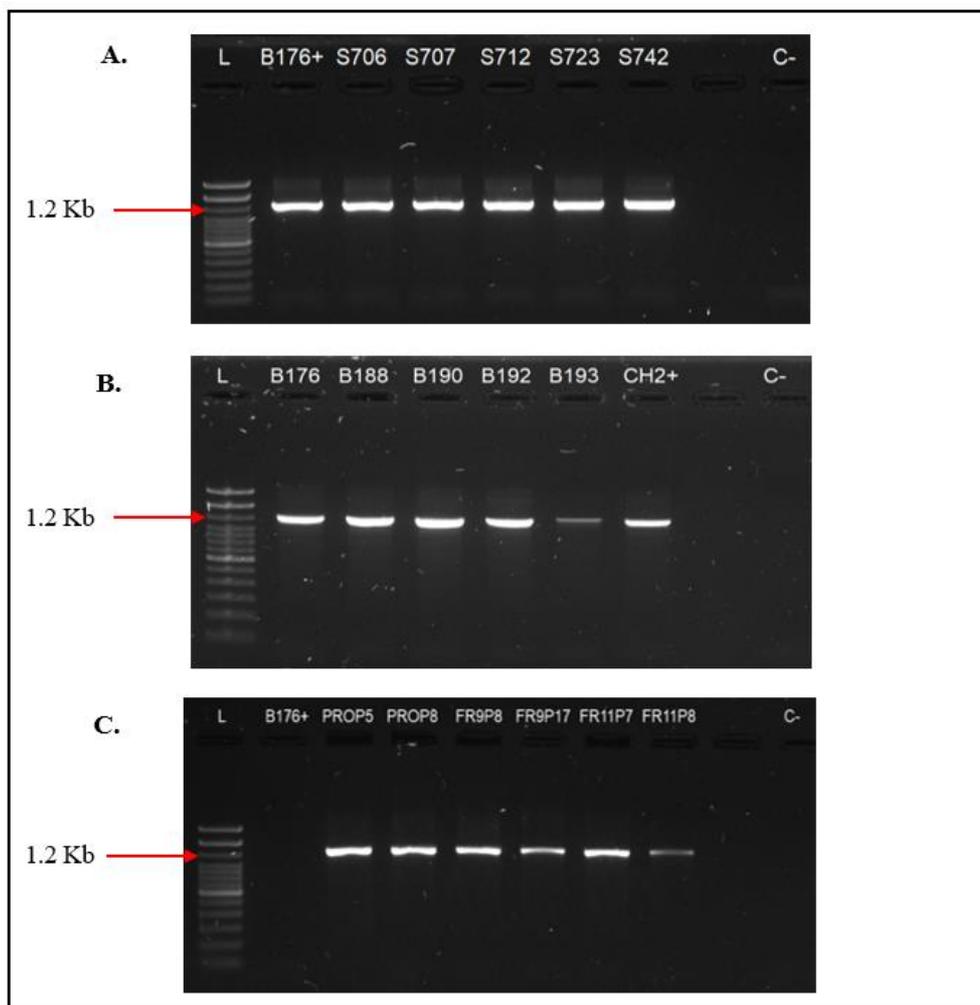


Figura 1. Geles de agarosa con la amplificación del gen LDC en las 14 accesiones de *L. mutabilis*. A) Amplificación del gen LDC en accesiones bajas en AQ (706, 707, 712, 723 y 724) y control positivo (B176+). B) Amplificación del gen LDC en accesiones altas en AQ (176, 188, 190, 192 y 193) y control positivo (CH2+). C) Amplificación del gen LDC en muestras PROINPA P5, PROINPA P8, FR9 P8, FR9 P17 y FR11 P8. La flecha roja indica la banda del Ladder correspondiente a 1.2 Kb. Todas las bandas obtenidas se encuentran en el tamaño esperado.

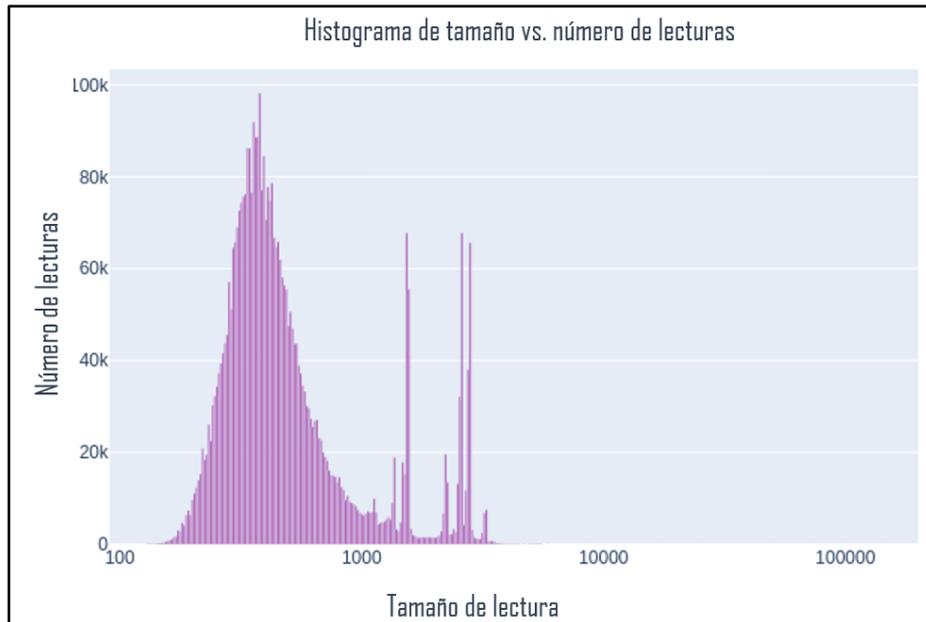


Figura 2. Histograma del tamaño y el número de lecturas generadas en el ensayo de secuenciación. Rango del tamaño de las lecturas va desde 60 pb hasta los 7,000 pb aproximadamente. Los picos observados corresponden a los tamaños esperados de los genes secuenciados (gen LDC: 1.2 Kb; RAP2-7: 3.4 Kb; LaAT: 3.8 Kb; DHPDS: 3.5 Kb; LaCAO: 6.9 Kb). Las lecturas para el gen LDC se encuentran entre el rango desde 60 pb hasta 1,200 pb aproximadamente.

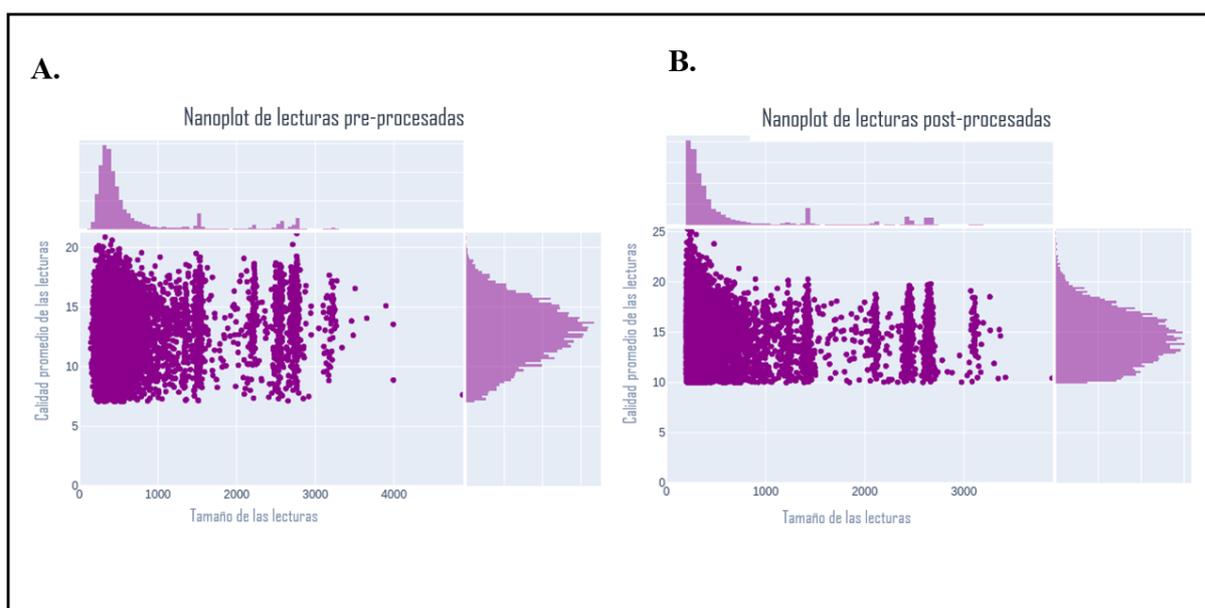


Figura 3. Comparación entre la calidad promedio y el tamaño promedio de las lecturas antes y después del procesamiento con Nanoplot. A) Resultados obtenidos antes del procesamiento (media de la calidad de las lecturas = 12.9 Phred; media del tamaño de lecturas = 662 pb). B) Resultados obtenidos después del procesamiento. Se eliminaron las lecturas con tamaño < 200 pb y con calidad < q10. (media de la calidad de las lecturas = 14.5 Phred; media del tamaño de lecturas = 643 pb).

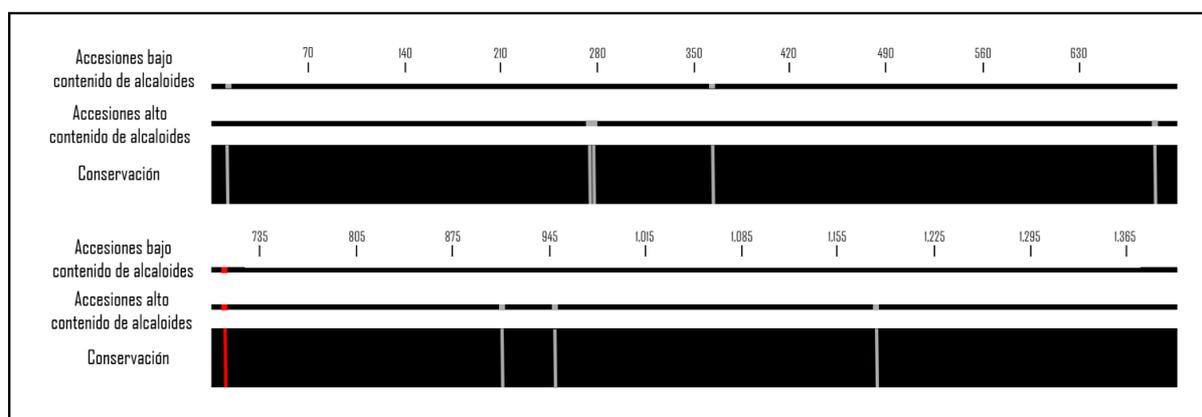


Figura 4. Alineación de secuencias del gen LDC entre accesiones con alto y bajo contenido de AQ de *L. mutabilis* para la identificación de polimorfismos. El tamaño de la secuencia de todas las accesiones es alrededor de 1,260 pb. El color oscuro denota conservación. Las regiones claras indican la posición y en qué grupo de accesiones se identificó un SNP. La región roja indica la identificación de un posible alelo específico para todas las accesiones bajas en AQ y otro específico para todas las accesiones altas en AQ.

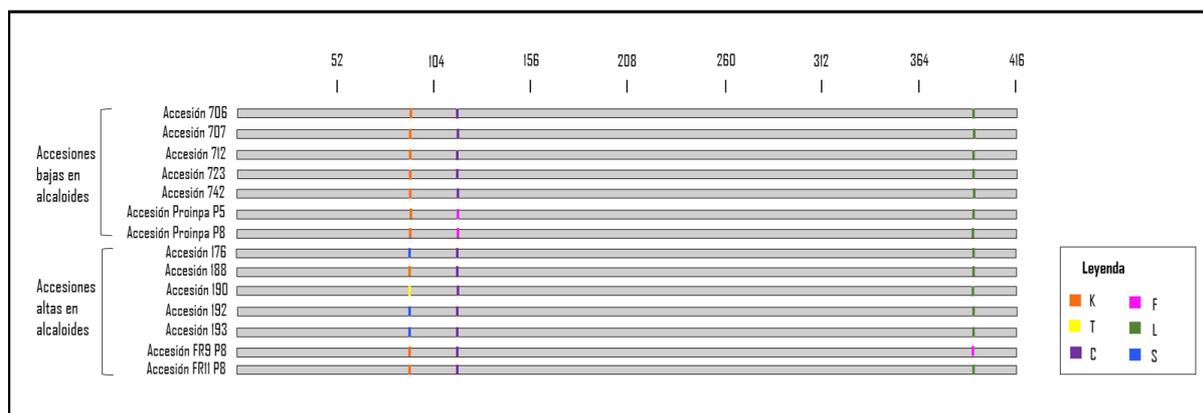


Figura 5. Alineación de las secuencias de aminoácidos obtenidas para el gen LDC en las 14 accesiones de *L. mutabilis* analizadas. Las regiones grises indican conservación. Las regiones coloridas indican la posición donde se identificó una sustitución de aminoácidos. En la leyenda se especifica el código de letras para los aminoácidos. K = Lisina; T = Treonina; C = Cisteína; F = Fenilalanina; L = Leucina; S = Serina.

8. REFERENCIAS BIBLIOGRÁFICAS

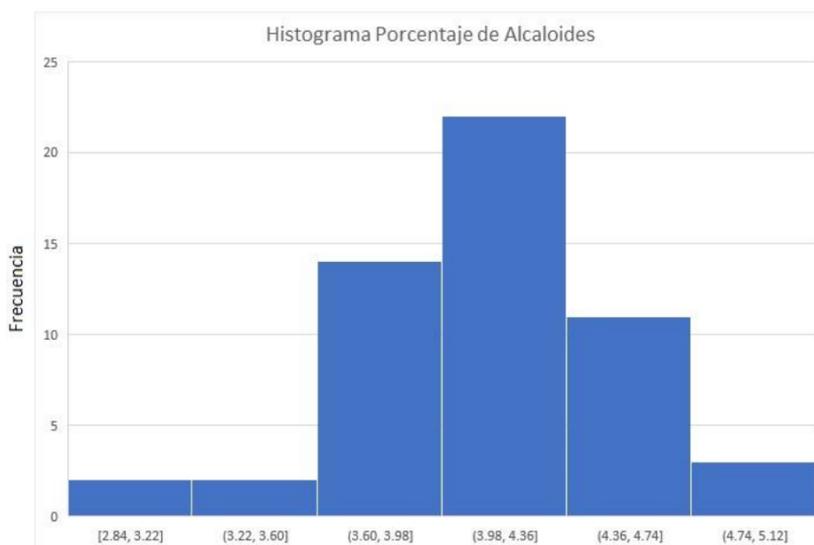
- Abraham, E. M., Ganopoulos, I., Madesis, P., Mavromatis, A., Mylona, P., Nianiou-Obeidat, I., ... & Vlachostergios, D. (2019). The use of lupin as a source of protein in animal feeding: Genomic tools and breeding approaches. *International journal of molecular sciences*, 20(4), 851.
- Atchison, G. W., Nevado, B., Eastwood, R. J., Contreras-Ortiz, N., Reynel, C., Madriñán, S., ... & Hughes, C. E. (2016). Lost crops of the Incas: Origins of domestication of the Andean pulse crop tarwi, *Lupinus mutabilis*. *American Journal of Botany*, 103(9), 1592-1606.
- Borah, P. (2011). Primer designing for PCR. *Science Vision*, 11(3), 134-136.
- Bunsupa, S., Katayama, K., Ikeura, E., Oikawa, A., Toyooka, K., Saito, K., & Yamazaki, M. (2012). Lysine decarboxylase catalyzes the first step of quinolizidine alkaloid biosynthesis and coevolved with alkaloid production in leguminosae. *The Plant Cell*, 24(3), 1202-1216.
- Caicedo, C., Murillo, A., Pinzón, J., Peralta, E., & Rivera, M. (2010). INIAP-450 Andino: Variedad de Chocho (*Lupinus mutabilis* Sweet).
- Caligari, P. D. S., Römer, P., Rahim, M. A., Huyghe, C., Neves-Martins, J., & Sawicka-Sienkiewicz, E. J. (2000). The potential of *Lupinus mutabilis* as a crop. In *Linking research and marketing opportunities for pulses in the 21st century* (pp. 569-573). Springer, Dordrecht.
- Camarena, F., Huaranga, A., Jiménez, J., & Mostacero, E. (2012). Revalorización de un cultivo subutilizado: Chocho o tarwi (*Lupinus mutabilis* Sweet). *Editorial CONCYTEC, Lima, Perú*.
- Cortés-Avenidaño, P., Tarvainen, M., Suomela, J. P., Glorio-Paulet, P., Yang, B., & Repo-Carrasco-Valencia, R. (2020). Profile and content of residual alkaloids in ten ecotypes of *Lupinus mutabilis* Sweet after aqueous debittering process. *Plant Foods for Human Nutrition*, 75(2), 184-191.
- Cowling, W. A., Huyghe, C., & Swiecicki, W. (1998). Lupin breeding. *Lupins as crop plants: biology, production and utilization*, 93-120.
- Frick, K. M. (2018). Genomic research to improve the quality of narrow-leafed lupin (*Lupinus angustifolius* L.) grain.
- Frick, K. M., Kamphuis, L. G., Siddique, K. H., Singh, K. B., & Foley, R. C. (2017). Quinolizidine alkaloid biosynthesis in lupins and prospects for grain quality improvement. *Frontiers in Plant Science*, 8, 87.

- Ranilla, L. G., Genovese, M. I., & Lajolo, F. M. (2009). Isoflavones and antioxidant capacity of Peruvian and Brazilian lupin cultivars. *Journal of food composition and analysis*, 22(5), 397-404.
- Ruiz jr, L. P. (1978). Alkaloid analysis of “sweet” lupin seed by GLC. *New Zealand Journal of Agricultural Research*, 21(2), 241-242.
- Goymer P. (2007). Synonymous mutations break their silence. *Nat Rev Genet*, 8, 92
- Gross, R., Von Baer, E., Koch, F., Marquard, R., Trugo, L., & Wink, M. (1988). Chemical composition of a new variety of the Andean lupin (*Lupinus mutabilis* cv. Inti) with low-alkaloid content. *Journal of food composition and analysis*, 1(4), 353-361.
- Güemes-Vera, N., Peña-Bautista, R. J., Jiménez-Martínez, C., Dávila-Ortiz, G., & Calderón-Domínguez, G. (2008). Effective detoxification and decoloration of *Lupinus mutabilis* seed derivatives, and effect of these derivatives on bread quality and acceptance. *Journal of the Science of Food and Agriculture*, 88(7), 1135-1143.
- Gulisano, A., Alves, S., Martins, J. N., & Trindade, L. M. (2019). Genetics and breeding of *Lupinus mutabilis*: An emerging protein crop. *Frontiers in plant science*, 10, 1385.
- Hatzold, T., Elmadfa, I., Gross, R., Wink, M., Hartmann, T., & Witte, L. (1983). Quinolizidine alkaloids in seeds of *Lupinus mutabilis*. *Journal of Agricultural and Food Chemistry*, 31(5), 934-938.
- Heather, J. M., & Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. *Genomics*, 107(1), 1-8.
- Hunt, R., Sauna, Z. E., Ambudkar, S. V., Gottesman, M. M., & Kimchi-Sarfaty, C. (2009). Silent (synonymous) SNPs: should we care about them?. *Single nucleotide polymorphisms*, 23-39.
- Jacobsen, E. & Mujica, A. 2006. El tarwi (*Lupinus mutabilis* Sweet) y sus parientes silvestres. Universidad Mayor de San Andrés, La Paz 2006 458-482.
- Jiménez-Martínez, C., Hernández-Sánchez, H., & Dávila-Ortiz, G. (2007). Diminution of quinolizidine alkaloids, oligosaccharides and phenolic compounds from two species of *Lupinus* and soybean seeds by the effect of *Rhizopus oligosporus*. *Journal of the Science of Food and Agriculture*, 87(7), 1315-1322.
- Joint Genome Institute. (2020). Recuperado desde JGI website: <https://genome.jgi.doe.gov/portal/>
- Kage, U., Kumar, A., Dhokane, D., Karre, S., & Kushalappa, A. C. (2016). Functional molecular markers for crop improvement. *Critical reviews in biotechnology*, 36(5), 917-930.

- Khan, M. K., Karnpanit, W., Nasar-Abbas, S. M., Huma, Z. E., & Jayasena, V. (2015). Phytochemical composition and bioactivities of lupin: a review. *International journal of food science & technology*, 50(9), 2004-2012.
- Kroc, M., Czepiel, K., Wilczura, P., Mokrzycka, M., & Świącicki, W. (2019). Development and validation of a gene-targeted dCAPS marker for marker-assisted selection of low-alkaloid content in seeds of narrow-leaved lupin (*Lupinus angustifolius* L.). *Genes*, 10(6), 428.
- Liao, Y. C., Cheng, H. W., Wu, H. C., Kuo, S. C., Lauderdale, T. L. Y., & Chen, F. J. (2019). Completing circular bacterial genomes with assembly complexity by using a sampling strategy from a single MinION run with barcoding. *Frontiers in microbiology*, 10, 2068.
- Liou, C. H., Wu, H. C., Liao, Y. C., Lauderdale, T. L. Y., Huang, I. W., & Chen, F. J. (2020). nanoMLST: accurate multilocus sequence typing using Oxford Nanopore Technologies MinION with a dual-barcode approach to multiplex large numbers of samples. *Microbial genomics*, 6(3).
- Markham, N. R., Zuker, M., & Keith, J. M. (2008). UNAFold: software for nucleic acid folding and hybridization., pp. 3–31. *Bioinformatics*.
- Mercado, G., Davalos, J., Ipdrs, Hivos, and Cipca., (2018). Memoria foro virtual: Los caminos del tarwi y la integración andina: Bolivia, Perú y Ecuador. Bolivia: IPDRS.
- Mori, L., & Paz, R. (2008). Eliminación de alcaloides en el tarwi (*Lupinus mutabilis*) mediante lavado con agua a diferentes pH. *Rev. CC. Biol*, 6, 48-53.
- Muzquiz, M., Hill, G. D., Burbano, C., Cuadrado, C., & Pedrosa, M. M. (Eds.). (2004). *Recent advances of research in antinutritional factors in legume seeds and oilseeds: Proceedings of the fourth international workshop on antinutritional factors in legume seeds and oilseeds* (Vol. 110). Wageningen Academic Publishers.
- Ramírez-Betancourt, A., Hernández-Sánchez, A. M., Salcedo-Morales, G., Ventura-Zapata, E., Robledo, N., Wink, M., & Bermúdez-Torres, K. (2021). Unraveling the biosynthesis of quinolizidine alkaloids using the genetic and chemical diversity of Mexican lupins. *Diversity*, 13(8), 375.
- Ruiz, L. P. (1978). Alkaloid analysis of “sweet” lupin seed by GLC. *New Zealand Journal of Agricultural Research*, 21(2), 241-242.
- Ohmiya, S., Saito, K., & Murakoshi, I. (1995). Lupine alkaloids. In *The alkaloids: Chemistry and pharmacology* (Vol. 47, pp. 1-114). Academic Press.
- Oxford Nanopore Technologies. (2021). Recuperado desde nanoporetech.com website: <https://nanoporetech.com/>
- Schaefer, C., & Rost, B. (2012). Predict impact of single amino acid change upon protein structure. In *BMC genomics* (Vol. 13, No. 4, pp. 1-10). BioMed Central.

- Shastry, B. S. (2009). SNPs: impact on gene function and phenotype. *Single Nucleotide Polymorphisms*, 3-22.
- SIB Swiss Institute of Bioinformatics. (2020). Recuperado desde Expasy website: <https://web.expasy.org/translate/>
- Thornsberry JM, Goodman MM, Doebley J, et al. (2001). Dwarf8 polymorphisms associate with variation in flowering time. *Nat Genet*, 28, 286–9
- Villacrés, E., Peralta, E. and Alvarez, M. (2003). Chochos en su punto, Vol 118. INIAP-FUNDACYT, Quito, Ecuador. In *Chochos. Recetarios. Disfrute Cocinando con Chochos*, ed. Peralta, E., Published as CD by INIAP FUNDACYT. Quito, Ecuador.
- Villacrés, E., Peralta, E., Cuadrado, L., Revelo, J., Addo, S., & Aldaz Berrones, R. F. (2009). Propiedades y aplicaciones de los alcaloides del chocho (*Lupinus mutabilis* Sweet).
- Wang, X. V., Blades, N., Ding, J., Sultana, R., & Parmigiani, G. (2012). Estimation of sequencing error rates in short reads. *BMC bioinformatics*, 13(1), 1-12.
- Whitford, W., Hawkins, V., Moodley, K., Grant, M. J., Lehnert, K., Snell, R. G., & Jacobsen, J. C. (2021). Optimised multiplex amplicon sequencing for mutation identification using the MinION nanopore sequencer. *bioRxiv*.
- Wink, M. (2003). Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry*, 64(1), 3-19.
- Wink, M., & Witte, L. (1984). Turnover and transport of quinolizidine alkaloids. Diurnal fluctuations of lupanine in the phloem sap, leaves and fruits of *Lupinus albus* L. *Planta*, 161(6), 519-524.
- Wolko, B., Clements, J. C., Naganowska, B., Nelson, M. N., and Yang, H. A.,(2011). “Lupinus,” in *Wild Crop Relatives: Genomic and Breeding Resources: Legume Crops and Forages*. Ed. Kole., C. (Berlin, Heidelberg: Springer Berlin Heidelberg), 153–206. doi: 10.1007/978-3-642-14387-8_9
- Zuker, M. (2003). Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic acids research*, 31(13), 3406-3415.

**ANEXO 1 : HISTOGRAMA CON LA DISTRIBUCIÓN DE VALORES DE
CUANTIFICACIÓN DE ALCALOIDES QUINOLIZIDÍNICOS PARA LAS 54
ACCESIONES DE CHOCHO**



Descripción: El histograma muestra la curva de distribución generada con los resultados de la cuantificación de AQ de las 54 accesiones realizado en un estudio previo en el Laboratorio de Biotecnología Vegetal de la USFQ. Con la finalidad de analizar las secuencias de ADN de accesiones que muestren fenotipos contrarios, se han seleccionado para el presente estudio las 5 accesiones con el mayor porcentaje de AQ y las 5 accesiones con el menor porcentaje de AQ que se ubican en los extremos contrapuestos de la curva de distribución.

**ANEXO 2 : SECUENCIAS CONSENSO GENERADAS DEL GEN LDC PARA LAS 14
ACCESIONES EVALUADAS.**

ID accesión	Secuencia consenso generada
706	TAATAAGGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCAACAACATGTACACCATTGAGCCTCA AGGCCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCTCA CATTTCATCAATCCATTATTTCGTAACACTCCAGAAAATTGATTCACCGTTCTTAGTGCTTGATCTGGCGTAGTTATGCACC TCATGGGCAAATGGACCACCAATCTTCTAAGGTTGAGCCTTATTATGCAGTTAAGTGAACCCCTAACGCGTCGTTGTAG GTGTACTCGCAGCACTTGGTTCGAGCTTCGACTGTGCCAGCCGAGCCGAAATCGAATCGGTTTTATCACTTGGAGTCTCAC

CGGACCGGATCATTTATGCGAACCCATGCAAAATCAGAGAGTACATTAAATATGCAGCCAGTGTGGTGTAAACGTTACA
ACGTTTCGACTCCAAGGAAGAGATTCAAAAGATTA AAAAGTGGCACCCGAAATGTGAGTTAGTTATTCGTTAATAGCCACC
ACAAGACAGTGGAGCAAGAAAATGCTTTAGGTCCTCAAATACGGTGCCTTCTGAAAGAAATTAAGCCACTGTTACAAGCCG
CAAAGGACGCAGGATTGAACGTTGTCGGTGTTCATTCCACATTTGGAAGTGGTGTGTCAGATTCTAGAACCCTACCATGGA
GCAATTTGCTCAGCTAAAAGCGTTTTTCGACATGGCTTCTAAAGAACTAGGCATGCCGAGAATGAAAATAGTGGACATTGG
AGGGCGTTTCACTTGTGGGAACCAATTTGATGCTGCTGTTTTACGTTGAATGAGGCTTTGAGCAAAATTTGGAAAAGA
AGAAGGTGTTGTGATAATCGGAGAACCTGGTCGTTATTTGTCAGAGTACCTTTTACTTTGGCTAGTAAAGTTATTGGGAA
GCGCGTGAGGGGAGAAGTGAGGGAGTATTGGATTGATGATGGGATTTACGGTTCCTTAATTGCATAATGTATGACTTTG
CAACTGTCACTTGCACGCCACTCGCGTGTGGATCAA AACACAGGAGGATCCTGGTTGCAGAAACGCGAAAGTACCCCTCA
ACTGTGTTTGGACCCACATGCGATTTCGTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATTGGAACCTGACGATTGG
CTTGTGTTCCCGAATATGGGTGCTTATACGACATCGTCAKGG

707

AATAATGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCACAACATGTACACCATTGAGCCTCAA
GGCCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCTCAC
ATTTCAATCCATTATTCGTAACACTCCAGAAAATTGATTCACCGTTCTTAGTGCTTGTATCTTGGCGTAGTTATGCACCT
CATGGGCAAATGGACCACCAATCTTCTAAGGTTACAGCCTTATTATGCAGTTAAGTGC AACCCCTAACCGCTCGTTGTAGG
TGTACTCGCAGCACTTGGTTCGAGCTTCGACTGTGCCAGCCGAAATCGAATCGGTTTTATCACTTGGAGTCTCAC
GGACCGGATCATTTATGCGAACCCATGCAAAATCAGAGAGTACATTAATATGCAGCCAGTGTGGTGTAAACGTTACAA
CGTTCGACTCCAAGGAAGAGATTCAAAAGATTA AAAAGTGGCACCCGAAATGTGAGTTAGTTATTCGTTAATAGCCACCA
CAAGACAGTGGAGCAAGAAAATGCTTTAGGTCCTAAAATACGGTGCCTTCTGAAAGAAATTAAGCCACTGTTACAAGCCG
AAAGGACGCAGGATTGAACGTTGTCGGTGTTCATTCCACATTGGAAGTGGTGGTGCAGATTCTAGAACCCTACCATGGAG
CAATTGCTCAGCTAAAAGCGTTTTTCGACATGGCTTCTAAAGAACTAGGCATGCCGAGAATGAAAATAGTGGACATTGGA
GGCGGTTTTCACTTGTGGGAACCAATTTGATGCTGCTGTTTTACGTTGAATGAGGCTTTGAGGACAATTTGGAAAAGA
GAAGGTGTTGTGATAATCGGAGAACCTGGTCGTTATTTGTCAGAGTACCTTTTACTTTGGCTAGTAAAGTTATTGGGAA
GCGGTGAGGGGAGAAGTGAGGGAGTATTGGATTGATGATGGGATTTACGGTTCCTTAATTGCATAATGTATGACTTTG
AACTGTACTTGCACGCCACTCGCGTGTGGATCAA AACACAGGAGGATCCTGGTTGCAGAAACGCGAAAGTACCCCTCA
CTGTGTTTGGACCCACATGCGATTTCGTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATTGGAACCTGACGATTGG
TTGTGTTCCCGAATATGGGTGCTTATACGACATCGTCAAG

712

AGTAATAATGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCACAACATGTACACCATTGAGCCT
CAAGGCCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCT
CACATTTCAATCCATTATTCGTAACACTCCAGAAAATTGATTCACCGTTCTTAGTGCTTGTATCTTGGCGTAGTTATGCA
CCTCATGGGCAAATGGACCACCAATCTTCTAAGGTTACAGCCTTATTATGCAGTTAAGTGC AACCCCTAACCGCTCGTTGT
AGGTGTACTCGCAGCACTTGGTTCGAGCTTCGACTGTGCCAGCCGAGCCGAAATCGAATCGGTTTTATCACTTGGAGTCT
ACCGGACCGGATCATTTATGCGAACCCATGCAAAATCAGAGAGTACATTAATATGCAGCCAGTGTGGTGTAAACGTTA
CAACGTTCCGACTCCAAGGAAGAGATTCAAAAGATTA AAAAGTGGCACCCGAAATGTGAGTTAGTTATTCGTTAATAGCCA
CCACAAGACAGTGGAGCAAGAAAATGCTTTAGGTCCTAAAATACGGTGCCTTCTGAAAGAAATTAAGCCACTGTTACAAG
CGCAAAGGACGCAGGATTGAACGTTGTCGGTGTTCATTCCACATTGGAAGTGGTGGTGCAGATTCTAGAACCCTACCATG
GAGCAATTGCTGCAGCTAAAAGCGTTTTTCGACATGGCTTCTAAAAGAACTAGGCATGCCGAGAATGAAAATAGTGGACATT
GGAGCGGTTTTCACTTGTGGGAACCAATTTGATGCTGCTGTTTTACGTTGAATGAGGCTTTGAGGACAATTTGGAAAA
GAAGAAGGTGTTGTGATAATCGGAGAACCTGGTCGTTATTTGTCAGAGTACCTTTTACTTTGGCTAGTAAAGTTATTGGG
AAGCGGTGAGGGGAGAAGTGAGGGAGTATTGGATTGATGATGGGATTTACGGTTCCTTAATTGCATAATGTATGACTTTG
TGCAACTGTACTTGCACGCCACTCGCGTGTGGATCAA AACACAGGAGGATCCTGGTTGCAGAAACGCGAAAGTACCCCT
CAACTGTGTTTGGACCCACATGCGATTTCGTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATTGGAACCTGACGATT
GGCTGTGTTCCCGAATATGGGTGCTTATACGACATCGTCAAGGA

723

AGTAATAATGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCACAACATGTACACCATTGAGCCT
CAAGGCCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCT
CACATTTCAATCCATTATTCGTAACACTCCAGAAAATTGATTCACCGTTCTTAGTGCTTGTATCTTGGCGTAGTTATGCA
CCTCATGGGCAAATGGACCACCAATCTTCTAAGGTTACAGCCTTATTATGCAGTTAAGTGC AACCCCTAACCGCTCGTTGT
AGGTGTACTCGCAGCACTTGGTTCGAGCTTCGACTGTGCCAGCCGAGCCGAAATCGAATCGGTTTTATCACTTGGAGTCT
ACCGGACCGGATCATTTATGCGAACCCATGCAAAATCAGAGAGTACATTAATATGCAGCCAGTGTGGTGTAAACGTTA
CAACGTTCCGACTCCAAGGAAGAGATTCAAAAGATTA AAAAGTGGCACCCGAAATGTGAGTTAGTTATTCGTTAATAGCCA
CCACAAGACAGTGGAGCAAGAAAATGCTTTAGGTCCTAAAATACGGTGCCTTCTGAAAGAAATTAAGCCACTGTTACAAG
CGCAAAGGACGCAGGATTGAACGTTGTCGGTGTTCATTCCACATTGGAAGTGGTGGTGCAGATTCTAGAACCCTACCATG
GAGCAATTGCTGCAGCTAAAAGCGTTTTTCGACATGGCTTCTAAAAGAACTAGGCATGCCGAGAATGAAAATAGTGGACATT
GGAGCGGTTTTCACTTGTGGGAACCAATTTGATGCTGCTGTTTTACGTTGAATGAGGCTTTGAGGACAATTTGGAAAA
GAAGAAGGTGTTGTGATAATCGGAGAACCTGGTCGTTATTTGTCAGAGTACCTTTTACTTTGGCTAGTAAAGTTATTGGG
AAGCGGTGAGGGGAGAAGTGAGGGAGTATTGGATTGATGATGGGATTTACGGTTCCTTAATTGCATAATGTATGACTTTG
TGCAACTGTACTTGCACGCCACTCGCGTGTGGATCAA AACACAGGAGGATCCTGGTTGCAGAAACGCGAAAGTACCCCT
CAACTGTGTTTGGACCCACATGCGATTTCGTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATTGGAACCTGACGATT
GGCTGTGTTCCCGAATATGGGTGCTTATACGACATCGTCAAGGACT

742

ATAGTAATAATGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCACAACATGTACACCATTGAGC
CTCAAGGCCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACAT
CTCACATTTCAATCCATTATTCGTAACACTCCAGAAAATTGATTCACCGTTCTTAGTGCTTGTATCTTGGCGTAGTTATG
CACCTCATGGGCAAATGGACCACCAATCTTCTAAGGTTACAGCCTTATTATGCAGTTAAGTGC AACCCCTAACCGCTCGTTG
TTAGGTGTACTCGCAGCACTTGGTTCGAGCTTCGACTGTGCCAGCCGAGCCGAAATCGAATCGGTTTTATCACTTGGAGT
TCACCGGACCGGATCATTTATGCGAACCCATGCAAAATCAGAGAGTACATTAATATGCAGCCAGTGTGGTGTAAACGTT
TACAACGTTCCGACTCCAAGGAAGAGATTCAAAAGATTA AAAAGTGGCACCCGAAATGTGAGTTAGTTATTCGTTAATAGC
CACCACAAGACAGTGGAGCAAGAAAATGCTTTAGGTCCTAAAATACGGTGCCTTCTGAAAGAAATTAAGCCACTGTTACAA
GCCGCAAAGGACGCAGGATTGAACGTTGTCGGTGTTCATTCCACATTGGAAGTGGTGGTGCAGATTCTAGAACCCTACCA
TGGAGCAATTTGCTGCAGCTAAAAGCGTTTTTCGACATGGCTTCTAAAAGAACTAGGCATGCCGAGAATGAAAATAGTGGACA
TTGGAGGCGGTTTTCACTTGTGGGAACCAATTTGATGCTGCTGTTTTACGTTGAATGAGGCTTTGAGGACAATTTGGAA
AAGAAGAAGGTGTTGTGATAATCGGAGAACCTGGTCGTTATTTGTCAGAGTACCTTTTACTTTGGCTAGTAAAGTTATTGG
GGAAGCGGTGAGGGGAGAAGTGAGGGAGTATTGGATTGATGATGGGATTTACGGTTCCTTAATTGCATAATGTATGACT
TTTGAACCTGTCACTTGCACGCCACTCGCGTGTGGATCAA AACACAGGAGGATCCTGGTTGCAGAAACGCGAAAGTACCCCT
CAACTGTGTTTGGACCCACATGCGATTTCGTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATTGGAACCTGACGATT
TTGGCTGTGTTCCCGAATATGGGTGCTTATACGACATCGTCAAGGACTAATTTCAA

Proinpa
P5

AGTAATAATGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCAACAACATGTACACCATTGAGCCT
CAAGGCCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCT
CACATTTCAATCCATTATTCGTAACACTCCAGAAATGATTACCGTCTTAGTGCTTGATCTTGGCGTAGTTATGCA
CCTCATGGGCAAAATGGACCACCAATCTTCTAAGTTTCAGCCTTATTATGCAGTTAAGTGCAACCCTAACCGCTCGTTGTT
AGGTGACTCGCAGCACTTGGTTCGAGCTTCGACTTTGCCAGCCGAGCCGAAATCGAATCGGTTTTATCACTTGGAGTCTC
ACCGGACCGGATCATTTATGCGAACCCATGCAAAATCAGAGAGTCACTTAAATATGCAGCCAGTGTGGTGTAAACGTTA
CAACGTTGACTCCAAGGAAGAGATTCAAAAGATTAAGAAAGTTGACCCGAAATGTGAGTTAGTTATTTCGTTAAGCCA
CCACAAGACAGTGGAGCAAGAAATGCTTTAGGTCTCAAATACGGTGCCTTCTGAAGAAGTTAAGCCACTGTTACAAGC
CGAAAGGACGCAGGATTGAACGTTGTCGGTGTTCATCCACATTTGAAAGTGGTGGTGCAGATTCTAGAACCTACCATG
GAGCAATTGCTGCAGCTAAAAGCGTTTTGACATGGCTTCTAAAGAACTAGGCATGCCGAGAATGAAAATAGTGGACATT
GGAGCGGTTTTCACTTGTGGGAACCAATTTGATGCTGCTGCTTTTACGTTGAATGAGGCTCTTGAGGACAATTTGGAAAA
GAAGAAGGTGTGTGATAATCGGAGAACCTGGTCTGTTATTTGCAGAGTCACTTTTACTTTGGCTAGTAAAGTTATTGGG
AAGCGGTGAGGGGAGAAGTGAAGGAGTATTGGATTGATGATGGGATTTACGGTCTCTTAATTGCATAATGTATGACTT
TGCAACTGCTTGCACGCCACTCGCGTGTGGATCAAACAGAGGATCTGGTTGCAGAAACGCGAAAAGTGTACCCTT
CAACTGTGTTGGACCCACATGCGATTCTGTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATTTGAACTTGACGATT
GGCTTGTGTTCCCGAATATGGGTGCTTATACGACATCGTCAGGKGTAAATTTSAAT

Proinpa
P8

AATAATGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCAACAACATGTACACCATTGAGCCTCAA
GGCCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCTCAC
ATTTCAATCCATTATTCGTAACACTCCAGAAATGATTACCGTCTTAGTGCTTGATCTTGGCGTAGTTATGCACCT
CATGGGCAAAATGGACCACCAATCTTCTAAGTTTCAGCCTTATTATGCAGTTAAGTGCAACCCTAACCGCTCGTTGTTAGG
TGTAACGACAGCACTTGGTTCGAGCTTCGACTTTGCCAGCCGAGCCGAAATCGAATCGGTTTTATCACTTGGAGTCTCACC
GGACCGGATCATTTATGCGAACCCATGCAAAATCAGAGAGTCACTTAAATATGCAGCCAGTGTGGTGTAAACGTTACAA
CGTTGACTCCAAGGAAGAGATTCAAAAGATTAAGAAAGTTGACCCGAAATGTGAGTTAGTTATTTCGTTAAGCCACCA
CAAGACAGTGGAGCAAGAAATGCTTTAGGTCTCAAATACGGTGCCTTCTGAAGAAGTTAAGCCACTGTTACAAGCCG
AAAGGACGCAGGATTGAACGTTGTCGGTGTTCATCCACATTGGAAGTGGTGGTGCAGATTCTAGAACCTACCATTGGAG
CAATTGCTGCAGCTAAAAGCGTTTTGACATGGCTTCTAAAGAACTAGGCATGCCGAGAATGAAAATAGTGGACATTGGA
GGCGTTTTCACTTGTGGGAACCAATTTGATGCTGCTGCTTTTACGTTGAATGAGGCTCTTGAGGACAATTTGGAAAAAG
GAAGTGTGTGATAATCGGAGAACCTGGTCTGTTATTTGCAGAGTCACTTTTACTTTGGCTAGTAAAGTTATTGGGAAG
CGGTGAGGGGAGAAGTGAAGGAGTATTGGATTGATGATGGGATTTACGGTCTCTTAATTGCATAATGTATGACTTTGC
AACTGTCACTTGCACGCCACTCGCGTGTGGATCAAACAGAGGATCTGGTTGCAGAAACGCGAAAAGTGTACCCTTCAA
CTGTTTTGGACCCACATGCGATTCTGTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATTTGAACTTGACGATTGGC
TTGTTTCCCGAATATGGGTGCTTATACGACATCGTCAGGAACATAATTTT

176

TAATGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCAACAACATGTACACCATTGAGCCTCAAGG
CCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCTCACATT
TCATTCAATCCATTATTCGTAACACTCCAGAAATGATTACCGTCTTAGTGCTTGATCTTGGCGTAGTTATGCACCTCAT
GGGCAAAATGGACCACCAATCTTCTTCCGTTTCAGCCTTATTATGCAGTTAAGTGCAACCCTAACCGCTCGTTGTTAGGTG
ACTCGCAGCACTTGGTTCGAGCTTCGACTTGTGCCAGCCGAGCCGAAATCGAATCGGTTTTATCACTTGGAGTCTCACC
CCGGATCATTTATGCGAACCCATGCAAAATCAGAGAGTCACTTAAATATGCAGCCAGTGTGGTGTAAACGTTACAACGTT
CGACTCCAAGGAAGAGATTCAAAAGATTAAGAAAGTTGACCCGAAATGTGAGTTAGTTATTTCGTTAAGCCACCAAG
ACAGTGGAGCAAGAAATGCTTTAGGTCTCAAATACGGTGCCTTCTGAAGAAGTTAAGCCACTGTTACAAGCCGCAAG
GACGCAGGATTGAACGTTGTCGGTGTTCATCCACATTGGAAGTGGTGGTGCAGATTCTAGAACCTACCATTGGAGCAATT
GCTGCAGTAAAAGCGTTTTGACATGGCTTCTAAAGAACTAGGCATGCCGAGAATGAAAATAGTGGACATTGGAGCGCG
TTTTCACTTGTGGGAACCAATTTGATGCTGCTGCTTTTACGTTGAATGAGGCTCTTGAGGACAATTTGGAAAAAGAAAG
TGTTGTGATAATTTGGAGAACCTGGTCTGTTATTTGCAGAGTCACTTTTACTTTGGCTAGTAAAGTTATTGGGAAGCGGT
GAGGGGAGAAGTGAAGGAGTATTGGATTGATGATGGGATTTACGGTCTCTTAATTGCATAATGTATGACTTTGCAACTGT
CACTTGCACCTTGCACGCCACTCGCGTGTGGATCAAACAGAGGATCTGGTTGCAGAAACGCGAAAAGTGTACCCTTCAA
TTGGACCCACATGCGATTCTGTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATTTGAACTTGACGATTGGCTTGTG
TCCCGAATATGGGTGCTTATACGACATCGTCAGGAA

188

AGTAATAATKCCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCAACAACATGTACACCATTGAGCCT
CAAGGCCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCT
CACATTTCAATCCATTATTCGTAACACTCCAGAAATGATTACCGTCTTAGTGCTTGATCTTGGCGTAGTTATGCA
CCTCATGGGCAAAATGGACCACCAATCTTCTAAGTTTCAGCCTTATTATGCAGTTAAGTGCAACCCTAACCGCTCGTTGTT
AGGTGACTCGCAGCACTTGGTTCGAGCTTCGACTTGTGCCAGCCGAGCCGAAATCGAATCGGTTTTATCACTTGGAGTCTC
ACCGGACCGGATCATTTATGCGAACCCATGCAAAATCAGAGAGTCACTTAAATATGCAGCCAGTGTGGTGTAAACGTTA
CAACGTTGACTCCAAGGAAGAGATTCAAAAGATTAAGAAAGTTGACCCGAAATGTGAGTTAGTTATTTCGTTAAGCCA
CCACAAGACAGTGGAGCAAGAAATGCTTTAGGTCTCAAATACGGTGCCTTCTGAAGAAGTTAAGCCACTGTTACAAGC
CGAAAGGACGCAGGATTGAACGTTGTCGGTGTTCATCCACATTGGAAGTGGTGGTGCAGATTCTAGAACCTACCATTG
GAGCAATTGCTGCAGCTAAAAGCGTTTTGACATGGCTTCTAAAGAACTAGGCATGCCGAGAATGAAAATAGTGGACATT
GGAGCGGTTTTCACTTGTGGGAACCAATTTGATGCTGCTGCTTTTACGTTGAATGAGGCTCTTGAGGACAATTTGGAAAA
GAAGAAGGTGTGTGATAATCGGAGAACCTGGTCTGTTATTTGCAGAGTCACTTTTACTTTGGCTAGTAAAGTTATTGGG
AAGCGGTGAGGGGAGAAGTGAAGGAGTATTGGATTGATGATGGGATTTACGGTCTCTTAATTGCATAATGTATGACTT
TGCAACTGCTACTTGCACGCCACTCGCGTGTGGATCAAACAGAGGATCTGGTTGCAGAAACGCGAAAAGTGTACCCTT
CAACTGTGTTTTGGACCCACATGCGATTCTGTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATTTGAACTTGACGATT
GGCTTGTGTTCCCGAATATGGGTGCTTATACGACATCGTCAGGAACT

190

AATAATGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCAACAACATGTACACCATTGAGCCTCAA
GGCCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCTCAC
ATTTCAATCCATTATTCGTAACACTCCAGAAATGATTACCGTCTTAGTGCTTGATCTTGGCGTAGTTATGCACCT
CATGGGCAAAATGGACCACCAATCTTCTACGGTTCAGCCTTATTATGCAGTTAAGTGCAACCCTAACCGCTCGTTGTTAGG
TGTACTCGCAGCACTTGGTTCGAGCTTCGACTTGTGCCAGCCGAGCCGAAATCGAATCGGTTTTATCACTTGGAGTCTCACC
GGACCGGATCATTTATGCGAACCCATGCAAAATCAGAGAGTCACTTAAATATGCAGCCAGTGTGGTGTAAACGTTACAA
CGTTGACTCCAAGGAAGAGATTCAAAAGATTAAGAAAGTTGACCCGAAATGTGAGTTAGTTATTTCGTTAAGCCACCA
CAAGACAGTGGAGCAAGAAATGCTTTAGGTCTCAAATACGGTGCCTTCTGAAGAAGTTAAGCCACTGTTACAAGCCG
AAAGGACGCAGGATTGAACGTTGTCGGTGTTCATCCACATTGGAAGTGGTGGTGCAGATTCTAGAACCTACCATTGGG
CAATTGCTGCAGCTAAAAGCGTTTTGACATGGCTTCTAAAGAACTAGGCATGCCGAGAATGAAAATAGTGGACATTGGA
GGCGTTTTCACTTGTGGGAACCAATTTGATGCTGCTGCTTTTACGTTGAATGAGGCTCTTGAGGACAATTTGGAAAAAGAA

GAAGGTGTTGTGATAAATTGGAGAACCTGGTCGTTATTTTGCAGAGTCACCTTTTACTTTGGCTAGTAAAGTTATTGGGAAG
CGCGTGAGGGGAGAAGTGAGGGGAGTATTGGATTGATGATGGGATTTACGGTTCTCTTAATTGCATAATGTATGACCTTTGC
AACTGTCACTTGCACGCCACTCGCGTGTGGATCAAACACAGAGGATCCTGGTTGCAGAAACGCGAAAGTTACCTTTCAA
CTGTGTTTGGACCACATGCGATTGCTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATGGAACTTGACGATTGGC
TTGTGTTCCCGAATATGGGTGCTTATACGACATCGTCAGGGA

192

AATAATGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCAACAACATGTACACCATTGAGCCTCAA
GGCCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCTCAC
ATTTCAATCAATCCATTATTCGTAACACTCCAGAAATGATTACCGGTTCTTAGTGCTTGATCTTGGCGTAGTTATGCACCT
CATGGGCAAATGGACCACCAATCTTCTTCGGTTACGCTTATTATGCAGTTAAGTGCAACCCCTAACCGCTCGTTGTTAGG
CAAGACAGTGGAGCAAAGAAATGCTTTAGGTCTCAAATACGGTGCCTTCTTGAAGAAGTTAAGCCACTGTTACAAGCCG
GGACCGGATCATTATGCGAACCCATGCAAATCAGAGAGTCACTAAATATGCAGCCAGTGTGGTGTAAACGTTACAA
CGTTCGACTCCAAGGAAGAGATTCAAAGATTAAAAAGTGGCACCCGAAATGTGAGTTAGTTATTCGTATTAAGCCACCA
CAAGACAGTGGAGCAAAGAAATGCTTTAGGTCTCAAATACGGTGCCTTCTTGAAGAAGTTAAGCCACTGTTACAAGCCG
AAAGGACGCAGGATTGAACGTTGTCGGTGTTCATTCCACATTGGAAGTGGTGGTGTGATTCTAGAACCTACCATGGAG
CAATTGCTGCAGTAAAAGCGTTTTTCGACATGGCTTCTAAAGAAGTGGCATGCCGAGAATGAAAATAGTGGACATTGGA
GGCGTTTTCACTTGTGGGAACCAATTTGATGCTGCTGTTTTACGTGAATGAGGCTTGTAGGACAATTTGGAAAAGAA
GAAGTGTGTGATAAATTGGAGAACCTGGTCGTTATTTTGCAGAGTCACTTTTACTTTGGCTAGTAAAGTTATTGGGAAG
CGCGTGAGGGGAGAAGTGGAGGAGTATTGGATTGATGATGGGATTTACGGTTCTCTTAATTGCATAATGTATGACTTTGC
AACTGTCACTTGCACGCCACTCGCGTGTGGATCAAACACAGAGGATCCTGGTTGCAGAAACGCGAAAGTGTACCTTCAA
CTGTGTTTGGACCACATGCGATTGCTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATGGAACTTGACGATTGGC
TTGTGTTCCCGAATATGGGTGCTTATACGACATCGTCAGGGA

193

AATGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCAACAACATGTACACCATTGAGCCTCAAG
CCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCTCACATT
TCATTCAATCCATTATTCGTAACACTCCAGAAATGATTACCGGTTCTTAGTGCTTGATCTTGGCGTAGTTATGCACCTCAT
GGGCAAATGGACCACCAATCTTCTTCGGTTACGCTTATTATGCAGTTAAGTGCAACCCCTAACCGCTCGTTGTTAGGTGT
ACTCGCAGCACTTGGTTCGAGCTTCGACTGTGCCAGCCGAAATCGAATCGGTTTTATCACTTGGAGTCTCACCCGGA
CCGGATCATTATGCGAACCCATGCAAATCAGAGAGTCACTAAATATGCAGCCAGTGTGGTGTAAACGTTACAACGTT
CGACTCCAAGGAAGAGATTCAAAGATTAAAAAGTGGCACCCGAAATGTGAGTTAGTTATTCGTATTAAGCCACCAAG
ACAGTGGAGCAAAGAAATGCTTTAGGTCTCAAATACGGTGCCTTCTTGAAGAAGTTAAGCCACTGTTACAAGCCGCAAG
GACGCAGGATTGAACGTTGTCGGTGTTCATTCCACATTGGAAGTGGTGGTGTGATTCTAGAACCTACCATGGAGCAATT
GCTGCAGTAAAAGCGTTTTTCGACATGGCTTCTAAAGAAGTGGCATGCCGAGAATGAAAATAGTGGACATTGGAGGCGG
TTTTCACTTGTGGGAACCAATTTGATGCTGCTGTTTTACGTGAATGAGGCTTGTAGGACAATTTGGAAAAGAAAGAA
TGTTGTGATAAATTGGAGAACCTGGTCGTTATTTTGCAGAGTCACTTTTACTTTGGCTAGTAAAGTTATTGGGAAGCGCGT
GAGGGGAGAAGTGGAGGAGTATTGGATTGATGATGGGATTTACGGTTCTCTTAATTGCATAATGTATGACTTTGCAACTGT
CACTTGCACGCCACTCGCGTGTGGATCAAACACAGAGGATCCTGGTTGCAGAAACGCGAAAGTGTACCTTCAA
TTGGACCACATGCGATTGCTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATGGAACTTGACGATTGGCTTGTGT
TCCGAATATGGGTGCTTATACGACATCGTCAGGTGCCTACTTTCAATACATGGGTTTAAAC

FR9 P3

TAATGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCAACAACATGTACACCATTGAGCCTCAAG
CCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCTCACATT
TCATTCAATCCATTATTCGTAACACTCCAGAAATGATTACCGGTTCTTAGTGCTTGATCTTGGCGTAGTTATGCACCTCAT
GGGCAAATGGACCACCAATCTTCTTAAGGTTACGCTTATTATGCAGTTAAGTGCAACCCCTAACCGCTCGTTGTTAGGTGT
ACTCGCAGCACTTGGTTCGAGCTTCGACTGTGCCAGCCGAAATCGAATCGGTTTTATCACTTGGAGTCTCACCCGGA
CCGGATCATTATGCGAACCCATGCAAATCAGAGAGTCACTAAATATGCAGCCAGTGTGGTGTAAACGTTACAACGTT
CGACTCCAAGGAAGAGATTCAAAGATTAAAAAGTGGCACCCGAAATGTGAGTTAGTTATTCGTATTAAGCCACCAAG
ACAGTGGAGCAAAGAAATGCTTTAGGTCTCAAATACGGTGCCTTCTTGAAGAAGTTAAGCCACTGTTACAAGCCGCAAG
AGCGAGGATTGAACGTTGTCGGTGTTCCTTCCACATTGGAAGTGGTGGTGTGATTCTAGAACCTACCATGGAGCAATT
GCTGCAGTAAAAGCGTTTTTCGACATGGCTTCTAAAGAAGTGGCATGCCGAGAATGAAAATAGTGGACATTGGAGGCGG
TTTTCACTTGTGGGAACCAATTTGATGCTGCTGTTTTACGTGAATGAGGCTTGTAGGACAATTTGGAAAAGAAAGAA
TGTTGTGATAAATTGGAGAACCTGGTCGTTATTTTGCAGAGTCACTTTTACTTTGGCTAGTAAAGTTATTGGGAAGCGCGT
GAGGGGAGAAGTGGAGGAGTATTGGATTGATGATGGGATTTACGGTTCTCTTAATTGCATAATGTATGACTTTGCAACTGT
CACTTGCACGCCACTCGCGTGTGGATCAAACACAGAGGATCCTGGTTGCAGAAACGCGAAAGTGTACCTTCAA
TTGGACCACATGCGATTGCTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATGGAACTTGACGATTGGCTTGTGT
CCCGAATATGGGTGCTTATACGACATCGTCAGGANKAATTTCACTGGGTGATATTAACAACAACA

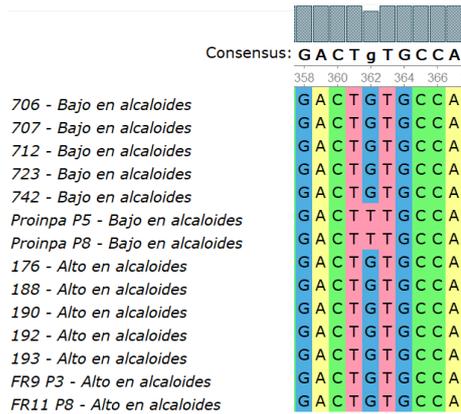
FR11 P8

ATAATGCCTTCACTACTTCTCACAGAGGGTTTATACCAGGCCAATAATGGTGCAACAACATGTACACCATTGAGCCTCAAG
GCCATTTATAATGCTTCTGGGGTGAAGGGTAAGCGAGTCACTCCATTACTCGCTAATGATGAACAAGGTGACATCTCACATT
TCATTCAATCCATTATTCGTAACACTCCAGAAATGATTACCGGTTCTTAGTGCTTGATCTTGGCGTAGTTATGCACCTCA
TGGGCAAATGGACCACCAATCTTCTAAGGTTACGCTTATTATGCAGTTAAGTGCAACCCCTAACCGCTCGTTGTTAGGTG
TACTCGCAGCACTTGGTTCGAGCTTCGACTGTGCCAGCCGAAATCGAATCGGTTTTATCACTTGGAGTCTCACCCG
ACCGGATCATTATGCGAACCCATGCAAATCAGAGAGTCACTAAATATGCAGCCAGTGTGGTGTAAACGTTACAACG
TTCGACTCCAAGGAAGAGATTCAAAGATTAAAAAGTGGCACCCGAAATGTGAGTTAGTTATTCGTATTAAGCCACCA
AGACAGTGGAGCAAAGAAATGCTTTAGGTCTCAAATACGGTGCCTTCTTGAAGAAGTTAAGCCACTGTTACAAGCCGCA
AGGACGAGGATTGAACGTTGTCGGTGTTCATTCCACATTGGAAGTGGTGGTGTGATTCTAGAACCTACCATGGAGCA
ATTGCTGCAGTAAAAGCGTTTTTCGACATGGCTTCTAAAGAAGTGGCATGCCGAGAATGAAAATAGTGGACATTGGAGG
CGGTTTCACTTGTGGGAACCAATTTGATGCTGCTGTTTTACGTGAATGAGGCTTGTAGGACAATTTGGAAAAGAAAGAA
AGGTTTGTGATAAATTGGAGAACCTGGTCGTTATTTTGCAGAGTCACTTTTACTTTGGCTAGTAAAGTTATTGGGAAGCGG
CGTGAAGGGGAGAAGTGGAGGAGTATTGGATTGATGATGGGATTTACGGTTCTCTTAATTGCATAATGTATGACTTTGCAA
CTGTCACTTGCACGCCACTCGCGTGTGGATCAAACACAGAGGATCCTGGTTGCAGAAACGCGAAAGTGTACCTTCAA
GTGTTTGGACCACATGCGATTGCTTAGATACTGTTCTAACAGATTACTTGTACCAGGAAATGGAACTTGACGATTGGCTT
GTGTTCCCGAATATGGGTGCTTATACGACATCGTCAGGACTAA

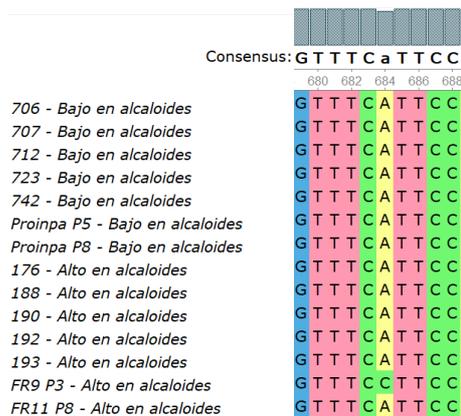
Descripción: Se incluye la secuencia consenso (formato FASTA) de cada una de las 14 accesiones evaluadas. Las secuencias consenso se obtuvieron luego del mapeado y ensamblado en *Geneious prime* de las lecturas que alineadas a la secuencia de referencia (LOC109327937). Estas secuencias consenso fueron alineadas para la identificación de SNPs.

ANEXO 3 : POSICIÓN DE LOS SNPS IDENTIFICADOS EN LA ALINEACIÓN DE LAS SECUENCIAS CONSENSO DEL GEN LDC EN LAS 14 ACCESIONES DE *L. MUTABILIS*.

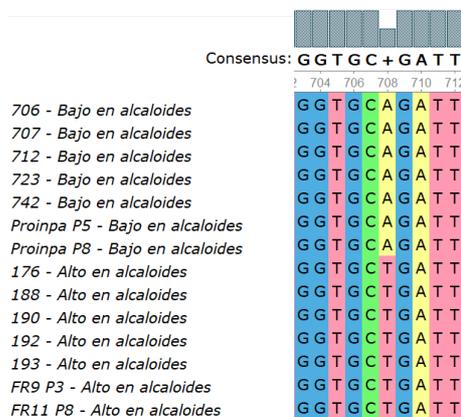
Polimorfismo identificado	Posición y grupo al que pertenece
 <p>Consensus: A A t G C C T T C A</p> <p>10 12 14 16 18</p> <p>706 - Bajo en alcaloides A A G G C C T T C A</p> <p>707 - Bajo en alcaloides A A T G C C T T C A</p> <p>712 - Bajo en alcaloides A A T G C C T T C A</p> <p>723 - Bajo en alcaloides A A T G C C T T C A</p> <p>742 - Bajo en alcaloides A A T G C C T T C A</p> <p>Proinpa P5 - Bajo en alcaloides A A T G C C T T C A</p> <p>Proinpa P8 - Bajo en alcaloides A A T G C C T T C A</p> <p>176 - Alto en alcaloides A A T G C C T T C A</p> <p>188 - Alto en alcaloides A A T G C C T T C A</p> <p>190 - Alto en alcaloides A A T G C C T T C A</p> <p>192 - Alto en alcaloides A A T G C C T T C A</p> <p>193 - Alto en alcaloides A A T G C C T T C A</p> <p>FR9 P3 - Alto en alcaloides A A T G C C T T C A</p> <p>FR11 P8 - Alto en alcaloides A A T G C C T T C A</p>	<p>Polimorfismo de un solo nucleótido en la posición 11 en accesión baja en alcaloides.</p>
 <p>Consensus: T C C T a a G G T T</p> <p>274 276 278 280 282</p> <p>706 - Bajo en alcaloides T C C T A A G G T T</p> <p>707 - Bajo en alcaloides T C C T A A G G T T</p> <p>712 - Bajo en alcaloides T C C T A A G G T T</p> <p>723 - Bajo en alcaloides T C C T A A G G T T</p> <p>742 - Bajo en alcaloides T C C T A A G G T T</p> <p>Proinpa P5 - Bajo en alcaloides T C C T A A G G T T</p> <p>Proinpa P8 - Bajo en alcaloides T C C T A A G G T T</p> <p>176 - Alto en alcaloides T C C T T C G G T T</p> <p>188 - Alto en alcaloides T C C T A A G G T T</p> <p>190 - Alto en alcaloides T C C T A C G G T T</p> <p>192 - Alto en alcaloides T C C T T C G G T T</p> <p>193 - Alto en alcaloides T C C T T C G G T T</p> <p>FR9 P3 - Alto en alcaloides T C C T A A G G T T</p> <p>FR11 P8 - Alto en alcaloides T C C T A A G G T T</p>	<p>Polimorfismo de un solo nucleótido en la posición 277 y 278 en accesiones altas en alcaloides.</p>



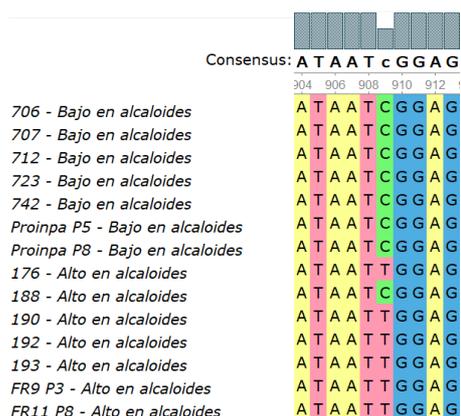
Polimorfismo de un solo nucleótido en la posición 362 en accesiones bajas en alcaloides.



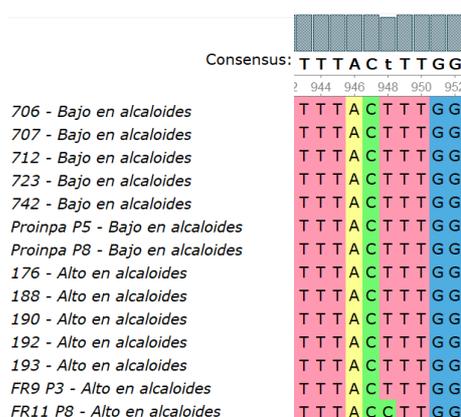
Polimorfismo de un solo nucleótido en la posición 684 en accesión alta en alcaloides.



Polimorfismo de un solo nucleótido en la posición 708 en todas las accesiones del grupo.



Polimorfismo de un solo nucleótido en la posición 909 en accesiones altas en alcaloides.



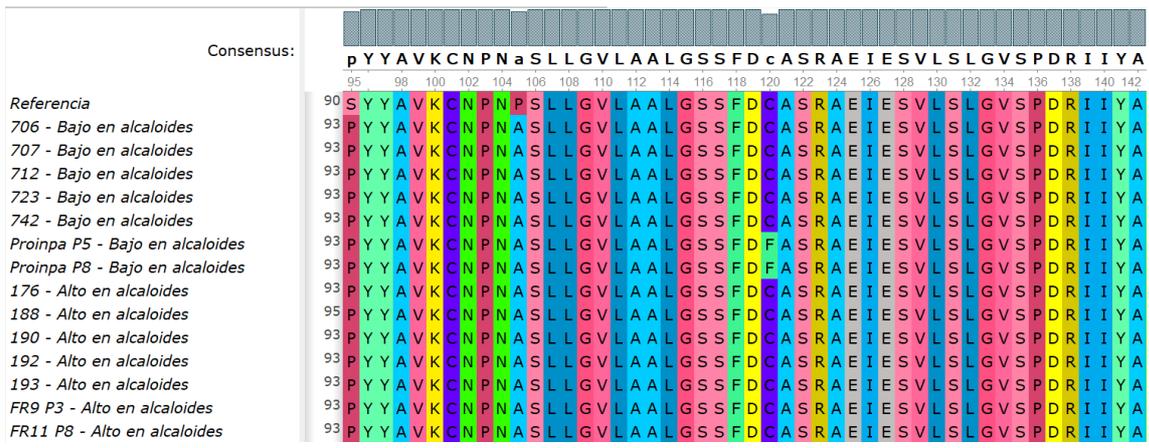
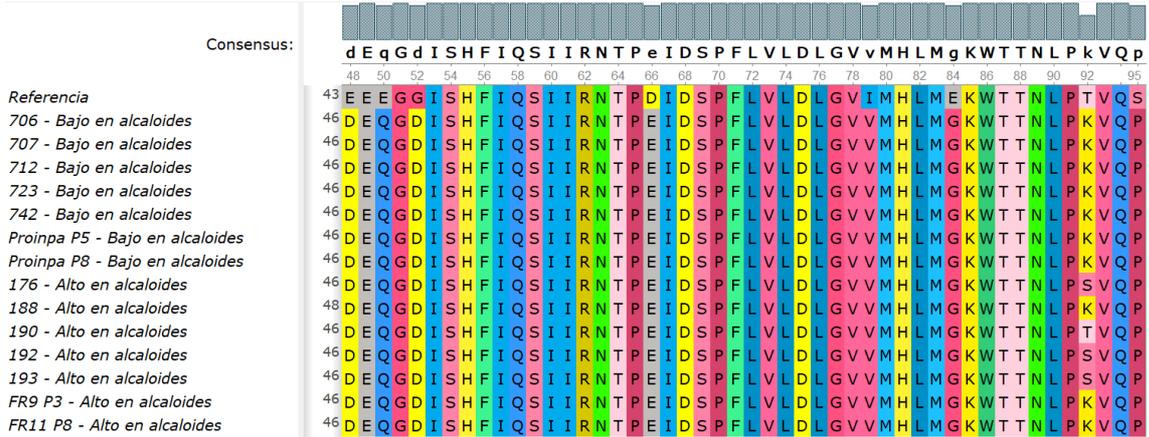
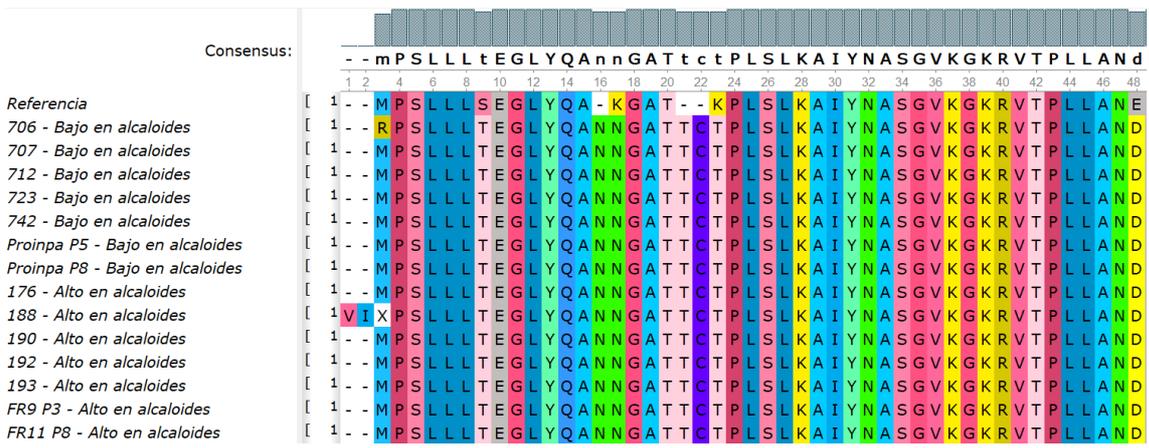
Polimorfismo de un solo nucleótido en la posición 948 en accesión alta en alcaloides.



Polimorfismo de un solo nucleótido en la posición 1185 en accesión alta en alcaloides.

Descripción: Se muestra los SNPs identificados tras la alineación de las 14 secuencias consenso generadas. En la parte izquierda se indica el ID de la accesión. En la parte derecha se indica la posición donde se encuentra cada SNP.

ANEXO 4 ALINEACIÓN DE LA SECUENCIA DE AMINOÁCIDOS DERIVADOS DE LA TRADUCCIÓN DE LA SECUENCIA DE NUCLEÓTIDOS DEL GEN LDC ENTRE LAS 14 ACCESIONES DE *L. MUTABILIS*.



Consensus: ANPCKSESHIKYAA s VGVNVTTFDSKEE I q KIKKWHPKCELvIRIKPP

Referencia

- 706 - Bajo en alcaloides
- 707 - Bajo en alcaloides
- 712 - Bajo en alcaloides
- 723 - Bajo en alcaloides
- 742 - Bajo en alcaloides
- Proinpa P5 - Bajo en alcaloides
- Proinpa P8 - Bajo en alcaloides
- 176 - Alto en alcaloides
- 188 - Alto en alcaloides
- 190 - Alto en alcaloides
- 192 - Alto en alcaloides
- 193 - Alto en alcaloides
- FR9 P3 - Alto en alcaloides
- FR11 P8 - Alto en alcaloides

Consensus: PQDSGARNALGLKYGALPEEVKPLLQAAKDAgLnVVGVSFHIGSGGAD

Referencia

- 706 - Bajo en alcaloides
- 707 - Bajo en alcaloides
- 712 - Bajo en alcaloides
- 723 - Bajo en alcaloides
- 742 - Bajo en alcaloides
- Proinpa P5 - Bajo en alcaloides
- Proinpa P8 - Bajo en alcaloides
- 176 - Alto en alcaloides
- 188 - Alto en alcaloides
- 190 - Alto en alcaloides
- 192 - Alto en alcaloides
- 193 - Alto en alcaloides
- FR9 P3 - Alto en alcaloides
- FR11 P8 - Alto en alcaloides

Consensus: DSRTYHGAIAAAKSVFDMASkELGMPRMKI VDI GGGFTCGNQFDAA a F

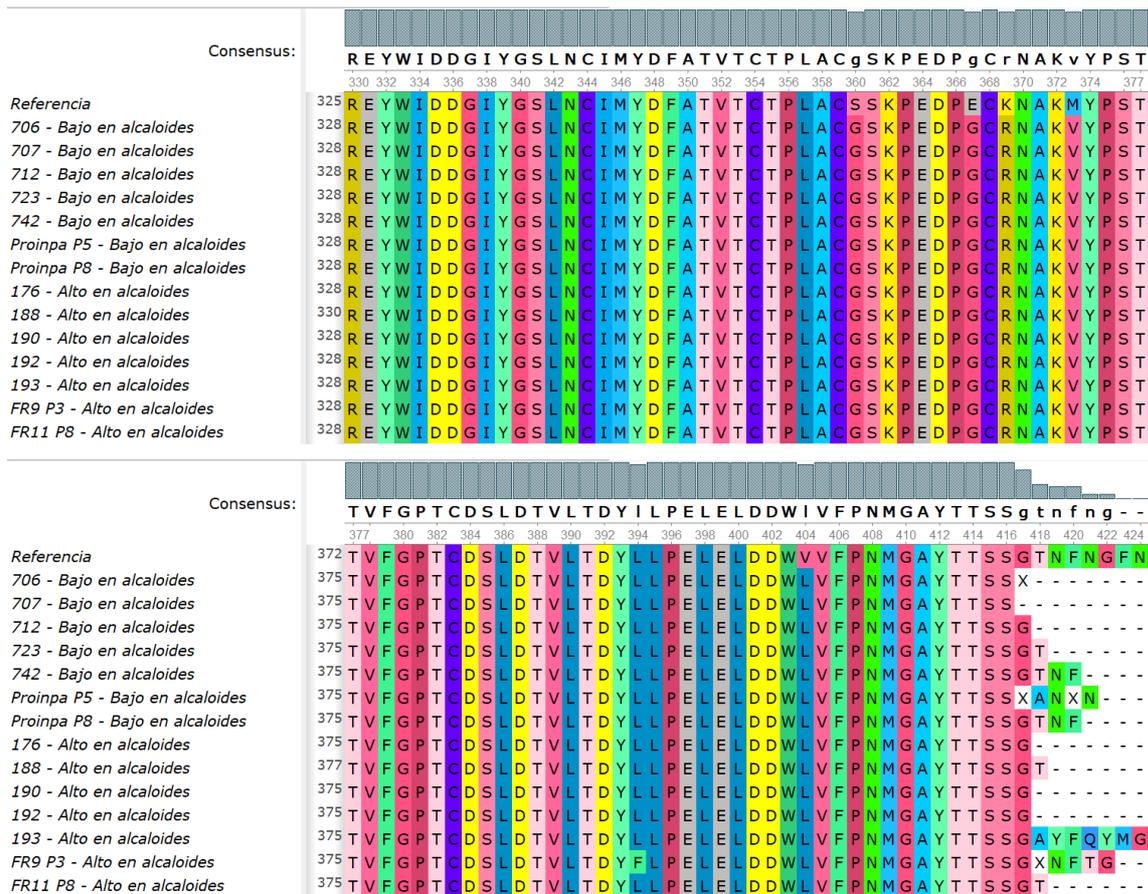
Referencia

- 706 - Bajo en alcaloides
- 707 - Bajo en alcaloides
- 712 - Bajo en alcaloides
- 723 - Bajo en alcaloides
- 742 - Bajo en alcaloides
- Proinpa P5 - Bajo en alcaloides
- Proinpa P8 - Bajo en alcaloides
- 176 - Alto en alcaloides
- 188 - Alto en alcaloides
- 190 - Alto en alcaloides
- 192 - Alto en alcaloides
- 193 - Alto en alcaloides
- FR9 P3 - Alto en alcaloides
- FR11 P8 - Alto en alcaloides

Consensus: FHVNEALEDNFGKEEGVViIGEPGRYFAESPFTLASKVIGKRVRGEVR

Referencia

- 706 - Bajo en alcaloides
- 707 - Bajo en alcaloides
- 712 - Bajo en alcaloides
- 723 - Bajo en alcaloides
- 742 - Bajo en alcaloides
- Proinpa P5 - Bajo en alcaloides
- Proinpa P8 - Bajo en alcaloides
- 176 - Alto en alcaloides
- 188 - Alto en alcaloides
- 190 - Alto en alcaloides
- 192 - Alto en alcaloides
- 193 - Alto en alcaloides
- FR9 P3 - Alto en alcaloides
- FR11 P8 - Alto en alcaloides



Descripción: Se muestra la posición de las variaciones de aminoácidos detectados luego de la alineación las secuencias de aminoácidos para las 14 accesiones de *L. mutabilis*. Estas secuencias han derivado de la traducción de las secuencias de nucleótidos del gen LDC. En la parte superior se incluye la secuencia de aminoácidos para la proteína LDC de *L. angustifolius*.