

**UNIVERSIDAD NACIONAL AGRARIA
LA MOLINA**

**ESCUELA DE POSGRADO
DOCTORADO EN CIENCIAS E INGENIERÍA
BIOLÓGICAS**



**“IDENTIFICACIÓN DE ARN NO CODIFICANTE DE CADENA
LARGA EN TARWI (*Lupinus mutabilis* Sweet)”**

Presentada por:

JOSÉ ERNESTO MANUEL HIDALGO RODRÍGUEZ

**TESIS PARA OPTAR EL GRADO DE DOCTOR
DOCTORIS PHILOSOPHIAE EN CIENCIAS E INGENIERÍA
BIOLÓGICAS**

Lima - Perú

2024

final

INFORME DE ORIGINALIDAD

9%

INDICE DE SIMILITUD

7%

FUENTES DE INTERNET

9%

PUBLICACIONES

7%

TRABAJOS DEL
ESTUDIANTE

ENCONTRAR COINCIDENCIAS CON TODAS LAS FUENTES (SOLO SE IMPRIMIRÁ LA FUENTE SELECCIONADA)

11%

★ dfzljdn9uc3pi.cloudfront.net

Fuente de Internet

Excluir citas

Activo

Excluir coincidencias < 1%

Excluir bibliografía

Activo

**UNIVERSIDAD NACIONAL AGRARIA
LA MOLINA
ESCUELA DE POSGRADO
DOCTORADO EN CIENCIAS E INGENIERÍA
BIOLÓGICAS**

**“IDENTIFICACIÓN DE ARN NO CODIFICANTE DE CADENA
LARGA EN TARWI (*Lupinus mutabilis* Sweet)”**

**TESIS PARA OPTAR EL GRADO DE DOCTOR
DOCTORIS PHILOSOPHIAE**

Presentada por:

JOSÉ ERNESTO MANUEL HIDALGO RODRÍGUEZ

Sustentada y aprobada ante el siguiente jurado:

Ph.D. Ana Kitazono Sugahara

PRESIDENTE

Dr. Gastón Zolla Benites

ASESOR

Dra. Rosa Espejo Joya

MIEMBRO

Dr. Raúl Blas Sevillano

MIEMBRO

Dr. Caspar Christian Cedrid Chater

MIEMBRO EXTERNO

DEDICATORIA

A Daniel Alejandro y Sofía Cristal. Su ímpetu de vida y sensibilidad aseguran que lo mejor de nosotros quedará para el futuro.

AGRADECIMIENTOS

A Cynthia por acompañarme en este camino de superación.

Al Dr. Gastón Zolla por su asesoría, consejos y perseverancia para realizar este trabajo.

A Prociencia-UNALM, contrato 177-2015-FONDECYT, por el financiamiento de mis estudios doctorales.

A Innovate Perú-UNALM, convenio 451-PNICP-BRI-2014, por el financiamiento que permitió iniciar este proyecto.

A Vladimir Camel, Ángel Vera y Susan Linares por su apoyo invaluable en la parte experimental.

Al Dr. Jorge Jiménez y a todos mis amigos del Programa de Investigación y Proyección Social en Cereales y Granos Nativos por hacer del laboratorio un segundo hogar.

ÍNDICE GENERAL

I. INTRODUCCIÓN	1
II. MARCO TEÓRICO	3
2.1. Mecanismos de regulación genética	3
2.2. El ARN en la regulación genética	4
2.3. Síntesis y mecanismos de acción de los <i>ARNlncs</i>	7
2.4. <i>ARNlncs</i> en floración y sequía	10
2.4.1. <i>ARNlncs</i> en floración	10
2.4.2. <i>ARNlncs</i> en sequía	13
2.5. Predicción de <i>ARNlncs</i> : Calculadora de Potencial Codificante (<i>CPC</i>)	13
2.6. El potencial de los estudios transcriptómicos en tarwi	14
III. MATERIALES Y MÉTODOS	16
3.1. Lugar de ejecución	16
3.2. Datos transcriptómicos, selección de <i>reads</i> y ensamblaje	16
3.3. Identificación de <i>ARNlncs</i> putativos	17
3.4. Obtención de <i>ARNlncs</i> de especies vegetales	18
3.5. Anotación de <i>ARNlncs</i> putativos	18
3.6. Determinación de la localización subcelular	18
3.7. Caracterización termodinámica de los <i>ARNlncs</i>	19
3.8. Conservación de los transcritos	20
3.9. Validación de patrones de expresión diferencial por <i>RT-PCR</i>	20
IV. RESULTADOS Y DISCUSIÓN	22
4.1. Identificación de <i>ARNlncs</i>	22
4.2. Anotación de <i>ARNlncs</i> putativos	26
4.3. Localización Subcelular de <i>ARNlncs</i>	27
4.4. Caracterización termodinámica y estructural de <i>ARNlncs</i>	29
4.5. Conservación estructural de <i>ARNlncs</i>	33
4.6. Expresión y validación por <i>RT-PCR</i>	36
V. CONCLUSIONES	40
VI. REFERENCIAS BIBLIOGRÁFICAS	41
VII. ANEXOS	51

ÍNDICE DE TABLAS

Tabla 1. Parámetros para el cálculo de la energía libre mínima	19
Tabla 2. Descripción de los primers para la validación por <i>RT-PCR</i>	21
Tabla 3. Localización subcelular de los <i>ARNncs</i> de tarwi	28

ÍNDICE DE FIGURAS

Figura 1. Los <i>ARNlncs</i> y su relación con el dogma de la biología molecular....	5
Figura 2. Procesamiento de los <i>ARNlncs</i>	6
Figura 3. Mecanismos de acción de los <i>ARNlncs</i>	8
Figura 4. <i>ARNlncs</i> regulatorios producidos a partir del locus <i>FLC</i>	9
Figura 5. Flujo de trabajo para la identificación de los <i>ARNlncs</i> en tarwi	17
Figura 6. Predicción de los <i>ARNlncs</i> en tarwi	25
Figura 7. Anotación de los <i>ARNlncs</i> de tarwi	27
Figura 8. Energía libre y estructura secundaria del <i>ARNlnc</i> S03 de tarwi	32
Figura 9. Conservación estructural del <i>ARNlnc</i> S03 de tarwi	35
Figura 10. Validación del protocolo de <i>RT-PCR</i> aplicado en tarwi	39

ÍNDICE DE ANEXOS

Anexo 1. Resumen de la secuenciación <i>HiSeq</i> de Macrogen	51
Anexo 2. Resultados de los algoritmos de detección de los <i>ARNlncs</i>	52
Anexo 3. Elemento <i>SECIS</i> en eucariotas	66
Anexo 4. Datos de conservación estructural por <i>Crossalign</i>	67

RESUMEN

El tarwi (*Lupinus mutabilis*) es una leguminosa con un alto potencial agronómico y aunque su genoma aún no ha sido secuenciado, cuenta con importante información transcriptómica disponible. Sin embargo, no se han estudiado los ARN de cadena larga no codificante (*ARNlncs*) en esta especie. Por lo tanto, el objetivo de este trabajo fue identificar los *ARNlncs* en 12 librerías de *ARN-seq* de tarwi. Para lograrlo, se empleó un enfoque multifacético que involucró la predicción, anotación, ubicación celular, caracterización termodinámica, conservación estructural y validación de *ARNlncs*. Se identificaron 590 *ARNlncs* por al menos dos algoritmos de identificación de *ARNlncs*. La anotación con la base de datos *PLncDB* mostró 571 *ARNlncs* únicos y 19 *ARNlncs* con homología en 28 familias botánicas, incluyendo Solanaceae (19), Fabaceae (17), Brassicaceae (17), Rutaceae (17), Rosaceae (16), y Malvaceae (16), entre otras. En total, 12 *ARNlncs* mostraron homología en más de 40 especies. Según cuatro algoritmos de ubicación celular, el 67 por ciento de los *ARNlncs* se ubicaron en el citoplasma, y 33 por ciento en exosomas. La caracterización termodinámica permitió encontrar el *ARNlnc* S03 con una estructura secundaria estable (-105,67 kcal/mol). Esta estructura incluyó tres regiones, con un bucle con ramificaciones múltiples que contiene una horquilla con un elemento similar a la secuencia de inserción de selenocisteína (*SECIS*). La evaluación de la conservación estructural por *Crossalign* reveló similitudes parciales entre tarwi (S03) y *L. lycopersicum* (Solyc04r022210.1). La validación por *RT-PCR* demostró que S03 se sobreexpresa en condiciones de estrés por sequía y en el estadio E4 de desarrollo floral en tarwi. Finalmente, estos resultados resaltan la importancia de los *ARNlncs* en el mejoramiento de tarwi durante el desarrollo reproductivo y bajo sequía.

Palabras clave: *Lupinus mutabilis*, tarwi, ARN de cadena larga no codificante, ARN largo no codificante, *lncRNA*, sequía, floración

ABSTRACT

Tarwi (*Lupinus mutabilis*) is a legume with high agronomic potential and even though its genome has not been sequenced, there is important transcriptomic information. However, long non coding RNAs (lncRNAs) have not been studied in this species. Therefore, the objective of this research was to identify *lncRNAs* in 12 *ARN-seq* libraries of *L. mutabilis*. Hence, a multilevel approach was used, based on *lncRNA* prediction, annotation, subcellular location, thermodynamic characterization, structural conservation, and validation. Thus, 590 *lncRNAs* were identified by at least two algorithms of *lncRNA* identification. Annotation with the *PLncDB* database showed 571 *lncRNAs* unique to tarwi and 19 *lncRNAs* with homology in 28 botanical families including Solanaceae (19), Fabaceae (17), Brassicaceae (17), Rutaceae (17), Rosaceae (16), and Malvaceae (16), among others. In total, 12 *lncRNAs* had homology in more than 40 species. 67 per cent of the *lncRNAs* were located in the cytoplasm and 33 per cent in exosomes. Thermodynamic characterization enabled the identification of S03 with a stable secondary structure (-105.67 kcal/mol). This structure included three regions, with a multibranch loop containing a hairpin with a Selenocysteine Insertion Sequence-like (*SECIS-like*) element. The evaluation of the structural conservation by *Crosslign* revealed partial similarities between tarwi (S03) and *S. lycopersicum* (Solyc04r022210.1). *RT-PCR* validation demonstrated that S03 was upregulated under drought conditions and in the E4 stage of flower development in tarwi. Finally, these results highlighted the importance of *lncRNAs* in tarwi improvement during the reproductive development stage and under drought conditions.

Keywords: *Lupinus mutabilis*, tarwi, long non coding RNA, *lncRNA*, drought, flowering.

I. INTRODUCCIÓN

Las nuevas herramientas biotecnológicas han permitido entender las interacciones entre los sistemas biológicos de orden superior, así como su afectación por el medio ambiente. De esta manera, en las últimas décadas, se ha avanzado notablemente en el conocimiento de la función de diversos elementos genómicos, transcriptómicos y metabolómicos (Aroca 2013). El análisis de estos datos ha permitido identificar secuencias que, sin ser convertidas en proteínas, cumplen diversos roles dentro del genoma, como los ARN no codificantes (Han et al. 2018).

Las secuencias de ARN no codificantes se caracterizan por no ser traducidas a proteínas y constituyen la mayor parte del transcriptoma en la mayoría de organismos eucariotas (Khemka et al. 2016). Estas secuencias incluyen elementos diversos como micro ARNs (*miARN*), ARN cortos de interferencia (*siARN*), ARN que interactúan con piwi (*piARN*) y ARN largos no codificantes (*ARNlncs*; Rajewsky et al. 2018). Estos elementos cumplen funciones esenciales al regular a secuencias codificantes del genoma mediante mecanismos diversos (Han et al. 2018). Entre los ARN no codificantes más abundantes, se encuentran los *miARNs*, seguidos por los *ARNlncs* (Wei et al. 2004).

Los *ARNlncs* son conocidos por su gran capacidad de regulación genómica con un potencial inmenso que va desde la transcripción, catálisis y remodelación, al silenciamiento de ARN de cromosomas enteros (Goff y Rinn 2015). En tanto que su desregulación en el ser humano está relacionada con diversos tipos de afecciones como diabetes, enfermedad cardiovascular, asma, Alzheimer, enfermedad renal, desórdenes neurológicos y cáncer (Akerman et al. 2017, Chen y Zhou 2017, Goff y Rinn 2015, Sun y Wong 2016). Mientras que el estudio en organismos vegetales ha permitido conocer que también cumplen un rol esencial en la regulación de diversos procesos (Wekesa et al. 2019).

Los genomas vegetales, también presentan un gran número de *ARNlncs*, que realizan un amplio rango de funciones en respuesta a diferentes estímulos (Xuan et al. 2015). Los procesos regulados varían desde el silenciamiento génico, maduración, reproducción, diferenciación, y tolerancia a estrés (Jarroux et al. 2017), siendo dianas con mucho potencial en el mejoramiento y en el entendimiento de los mecanismos de regulación

genético vegetal (Han et al. 2018). A pesar de que se han realizado pocos estudios relacionados con *ARNlncs* en plantas, la disponibilidad de datos transcriptómicos hace posible su búsqueda y análisis.

Sin embargo, estos elementos no han sido estudiados a fondo, ni menos identificados en especies nativas con gran potencial, como el tarwi. Esta especie es una leguminosa de los Andes que ha sido considerada como un súper alimento por su alto valor nutricional (Repo-Carrasco-Valencia et al., 2021). Sus semillas están biofortificadas con niveles óptimos de hierro, zinc y boro (Vera-Vega et al., 2022) y tiene alcaloides bioactivos con propiedades antimicrobianas (25). Además, se ha realizado el ensamblaje de novo de los transcriptomas de tarwi en condiciones de respuesta a estrés (Hidalgo 2021) y en diferentes fases de desarrollo floral (Ramos 2021), lo cual abre la oportunidad de entender el panorama evolutivo de los *ARNlncs*, su conservación y funcionalidad durante el desarrollo y estrés. Por esta razón, el presente trabajo se realizó con el objetivo de identificar los ARN de cadena larga no codificante presentes en 12 librerías de *ARN-seq* de tarwi (*Lupinus mutabilis* Sweet).

II. MARCO TEÓRICO

2.1. MECANISMOS DE REGULACIÓN GENÉTICA

De forma tradicional, el estudio de los mecanismos de regulación genética se ha enfocado en la evaluación de genes específicos, poniendo énfasis en las secuencias génicas codificantes, es decir, aquellas capaces de producir proteínas (Amudha y Balasubramani 2011). El análisis de estos genes se ha realizado de manera individual, y su validación se ha hecho mediante la generación de organismos *knock-out* (Aroca 2013) y organismos recombinantes que expresan una proteína de manera heteróloga, lo cual ha permitido descubrir cómo el gen en cuestión regula diversos procesos biológicos (Ashraf 2010). Esta información ha permitido generar cultivos mejorados de forma biotecnológica, introduciendo genes específicos que brindan características deseables (Zhou et al. 2016).

Sin embargo, la evaluación de genes de manera individual ha sido una limitante para comprender el panorama global que gobierna diversos elementos de regulación genética. De manera que, mediante este enfoque, no se han podido estudiar los mecanismos complejos relacionados con procesos fisiológicos, bioquímicos y moleculares que involucran la participación de decenas de genes (Farooq et al. 2009). En una célula, estos genes interactúan entre sí formando redes de regulación que involucran a moléculas de ADN, ARN y proteínas, permitiendo responder específicamente a estímulos determinados, como el ajuste osmótico ante el estrés, o la diferenciación de brotes florales al alcanzar un estado de desarrollo (Jarroux et al. 2017).

El estudio de estas redes de regulación y los miles de elementos moleculares que las componen se han hecho posibles por la secuenciación de próxima generación (*NGS*), que ha permitido análisis genómicos a gran escala, demostrando la existencia de secuencias codificantes de proteínas y sus elementos regulatorios esenciales para la transcripción (Ashraf 2010). De esta manera, las herramientas genómicas han permitido encontrar cientos de genes relacionados a maduración, respuesta a estrés, producción de metabolitos, entre otros (Akinci 2013). Gracias a esta información, en la última década el entendimiento de los mecanismos de regulación se ha realizado de forma más eficiente.

Entre los datos más relevantes obtenidos del análisis de los trabajos *NGS*, se encuentra el hallazgo de una gran cantidad de secuencias no codificantes que son transcritas a ARN (Jarroux et al. 2017). De hecho, los estudios de anotación de genoma en eucariotas, indican que menos del dos por ciento de los genomas es traducido a proteínas; en tanto que las regiones no codificantes pueden formar diferentes tipos de ARN, además de los ARN de transferencia (ARNt) y ribosomal (ARNr). Estos ARN no codifican proteínas y pueden ser de secuencia corta, como los *miARNs* y *siARNs*, o de secuencia larga, como los *ARNlncs* (Wei et al. 2004). Estos últimos son bastante abundantes en la célula, lo cual constituye un claro indicador de su importancia en la regulación de los procesos celulares.

2.2. EL ARN EN LA REGULACIÓN GENÉTICA

Se ha postulado que los mecanismos de regulación más importantes en eucariotas han sido generados por la expansión del transcriptoma no codificante. Esta expansión ha llevado a la evolución de redes de regulación mediadas por ARN en casi cada aspecto del funcionamiento celular (Morris 2016). De hecho, se ha observado que el aumento en la regulación mediada por ARN está relacionado con la evolución de la complejidad (Taft et al. 2007). En particular, los ARN no codificantes, parecen tener un rol mucho más esencial, puesto que están presentes en mayor cantidad en organismos más complejos (Mattick 2004). De manera que estos ARNs pueden regular la expresión a nivel de ADN, ARN o proteínas, generando un cambio en el dogma de la biología molecular tal como se muestra en la Figura 1a (Lai et al. 2019).

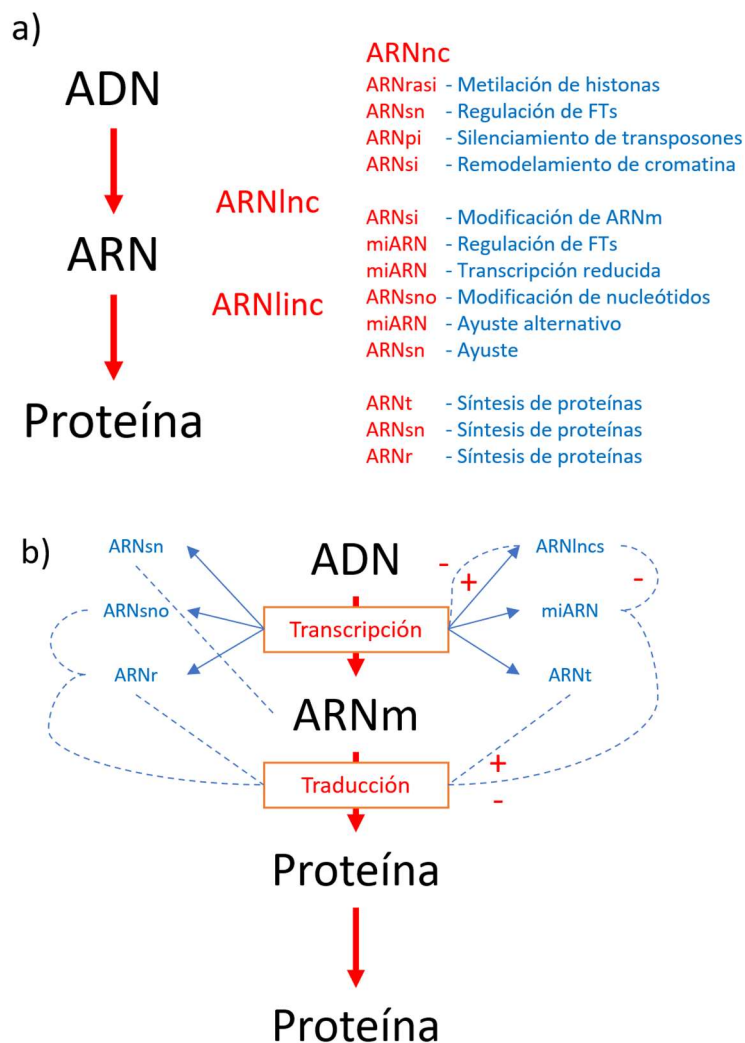


Figura 1. Los *ARNlncs* y su relación con el dogma de la biología molecular. **a)** Funciones de los diferentes tipos de ARN no codificantes (Lai et al. 2019). **b)** Esquema del modelo dinámico actual del dogma central de la biología molecular, con énfasis en el rol de los ARN no codificantes (Adaptado de Lai et al. 2019, Jarroux et al. 2017).

Como entidad biológica, el ARN puede regular la transcripción en diversos puntos (Figura 1b), puesto que tiene muchas propiedades únicas que complementan a las proteínas y las hacen particularmente adecuadas para muchas funciones celulares. Por ejemplo, los ARNs pueden reconocer y unirse fácilmente a otras moléculas de ácidos nucleicos (Morris 2016). Un ARN regulatorio puede secuestrar a su ARN diana, o estabilizarlo, o enmascarar su secuencia de otros elementos regulatorios como proteínas y otros ARNs. A través de la interacción simultánea con ADN y proteínas, los ARNs pueden ayudar a incrementar la concentración de proteínas en un locus genómico determinado (Rinn 2014). De esta manera, pueden llegar a crear redes celulares de

regulación, en las cuales los ARN no codificantes pueden actuar como competidores o activadores, indicando la localización o acción de las ribonucleoproteínas (Ugarkovic 2011).

Estas estrategias regulatorias mediadas por ARN son particularmente beneficiadas por el hecho de que la síntesis de ARN presenta un requerimiento energético más bajo en comparación con la síntesis de proteínas, la cual involucra el costo energético de la traducción y el plegamiento (Yang y Zhang 2015). Esto cobra mayor importancia en etapas críticas de desarrollo y condiciones de deficiencias nutricionales o respuesta a estrés (Chen et al. 2021). Además, algunos ARNs regulatorios, como los *ARNlncs* intrónicos, se originan de productos de expresión génica degradados, lo cual reduce aún más el costo energético de la regulación mediada por *ARNlncs* (Ayupe et al. 2015). A ello, se suma el hecho de que las moléculas de ARN son capaces de adquirir estructuras secundarias y terciarias a partir de una cadena simple, presentando dominios estructurales que interactúan con elementos específicos formando la red del interactoma de ARN (Figura 2; Zampetaki et al. 2018).

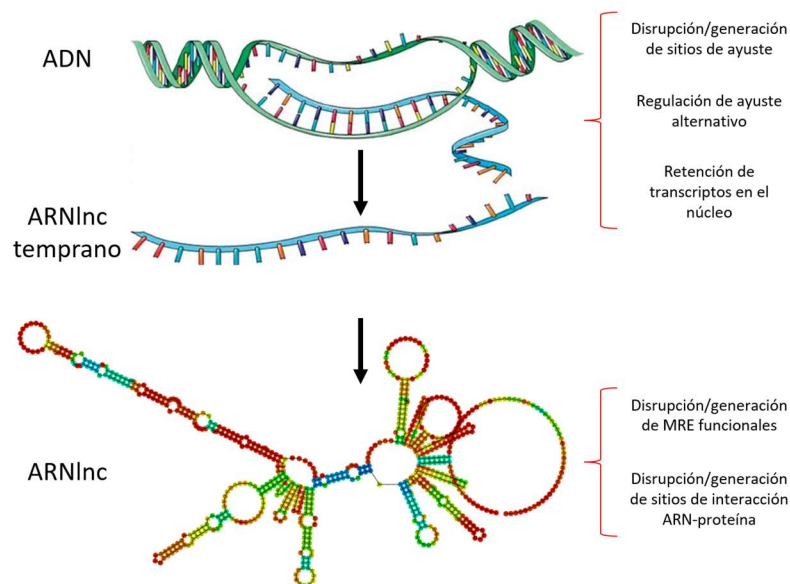


Figura 2. Procesamiento de los *ARNlncs*. Edición del ARN y conformación de estructuras complejas secundarias. Lai et al. (2019). MRE: Elemento de reconocimiento de micro-ARN.

2.3. SÍNTESIS Y MECANISMOS DE ACCIÓN DE LOS *ARNlncs*

Los *ARNlncs*, constituyen alrededor del 70 por ciento de los ARN no codificantes (Han et al. 2018). Estos elementos son definidos como secuencias transcritas de ARN mayores a 200 nucleótidos (nt) que no codifican proteínas (Rinn y Chang 2012). La mayor parte de los *ARNlncs* son transcritos por la ARN polimerasa II, y poseen una caperuza en el extremo 5' metilo, y una cola de poliadenina (poli A). Además, dependiendo de la dirección en la cual se transcriben, pueden ser definidos como *ARNlncs* sentido (transcritos de la hebra molde, al igual que el ARNm); *ARNlncs* antisentido (transcritos de la hebra codificante, complementaria a la hebra que produce el ARNm); *ARNlncs* intrónicos (se encuentran dentro del intrón de otro gen); y *ARNlncs* intergénicos (transcritos de una región alejada de los genes codificantes) (Fallis 2011; Figura 2).

Existen miles de *ARNlncs* predichos en el genoma eucariota. Estas secuencias comparten regiones de homología significativa con los ARNs de proteínas codificantes (Kiegle et al. 2018), pudiendo ser encontrados en exones e intrones que son cortados y empalmados selectivamente (Kiegle et al. 2018). Por esta razón, se sostiene que los *ARNlncs* participan en mecanismos de competencia por corte y empalme, así como en la síntesis de otros factores necesarios para el procesamiento del ARN (Kiegle et al. 2018). La existencia y develación de estos mecanismos, revela la complejidad de la expresión génica en eucariotas (Li et al. 2017).

Los *ARNlncs* son capaces de adquirir estructuras secundarias complejas que permiten regular la expresión génica (Figura 2; Yu y Zhu 2019) mediante diversos mecanismos (Figura 3). De esta manera, los *ARNlncs* pueden actuar como elementos trampa para factores de transcripción, como esponjas de *miARN*-ribonucleoproteínas (*RNPs*), como componentes de *RNPs* que modifican la cromatina, o como factores de transcripción que realizan el reclutamiento y anclaje de complejos que modifican la cromatina a sus secuencias diana, entre otros (Han et al. 2018). De esta manera, en *Arabidopsis thaliana* se ha encontrado que los *ARNlncs* afectan los patrones de metilación (Chen 2021). Además, se han encontrado *ARNlncs* que pueden impedir el silenciamiento génico al actuar como dianas de *miARNs* y evitar la degradación del ARNm (Li et al. 2017). Otros *ARNlncs* pueden dirigir la formación de complejos proteína-ácidos ribonucleicos,

influyendo en la regulación de la expresión genética y participando como reguladores en diversos procesos metabólicos (Figura 3; Rao 2017; Yu y Zhu 2019).

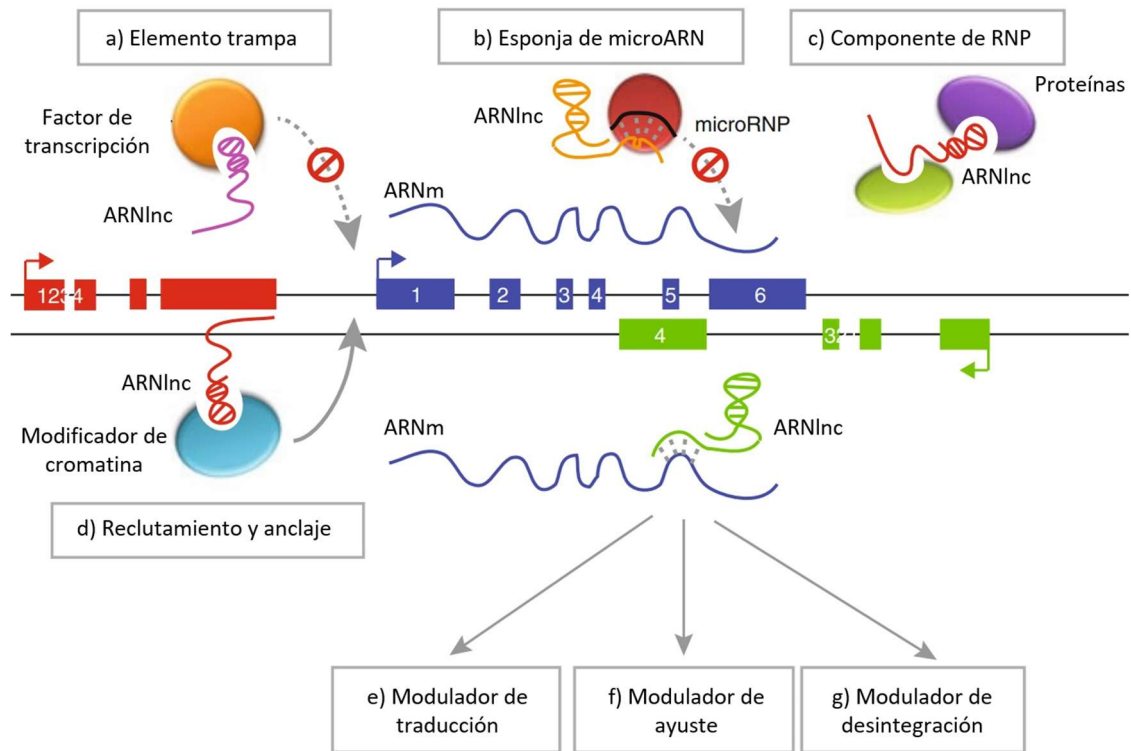


Figura 3. Mecanismos de acción de los *ARNlnc*. **a)** Los elementos trampa se unen a los factores de transcripción. **b)** Las esponjas de *miARNs* evitan la degradación del *ARNm*. **c)** Los componentes de *RNPs* se unen a combinaciones específicas de ribonucleoproteínas. **d)** Los *ARNlncs* pueden reclutar y anclar a los complejos modificadores de la cromatina. **e)** Regulación postranscripcional del *ARNm* a nivel traduccional. **f)** Regulación *ARNm* a nivel de corte y empalme. **g)** Regulación del *ARNm* a nivel de su desintegración. Adaptado de Khalil y Coller (2013).

Los *ARNlncs* regulan diversos procesos, como la respuesta a estrés y el desarrollo (Yu y Zhu 2019). En condiciones de estrés, constituyen una fracción significativa de los genes que se expresan de forma diferencial, y una vez inducidos, contribuyen a la regulación de procesos celulares corriente abajo (Morris 2016). Los estudios realizados incluyen el hallazgo de *ARNlncs* que son transcritos en respuesta a estrés hídrico en *A. thaliana* (Qin et al. 2017). Un estudio en yuca ha demostrado la presencia de *ARNlnc* que son dianas miméticas de *miARNs* en respuesta a estrés por frío o por sequía (Li et al. 2017). En *Pyrus betifolia*, se han encontrado hasta 251 *ARNlncs* de respuesta a estrés, los cuales según la

comprobación experimental estuvieron relacionados a procesos de desecación (Wang et al. 2018).

Al realizarse análisis de regulación cruzada con genes codificantes se ha encontrado que muchos *ARNncs* están asociados con las rutas de transducción en la señalización hormonal, biosíntesis de metabolitos secundarios y metabolismo de la sacarosa (Li et al. 2017), siendo dianas con mucho potencial en el mejoramiento de plantas, que aún puede ser explorada (Han et al. 2018). Por ejemplo, Khemka et al. (2016) identificaron 2248 *ARNncs* intergénicos (*ARNncis*) en garbanzo al caracterizar ocho estadios de desarrollo floral sucesivos y tres tejidos vegetativos. En estudios en *A. thaliana*, se ha encontrado que los *ARNncs* cumplen un rol esencial en la regulación de la floración al controlar la expresión de FLOWERING LOCUS C (*FLC*; Csorba et al. 2014). En tanto que un estudio en el endospermo de maíz mostró la presencia de *ARNncs* que exhibían diferencia espacio-temporal, participando en redes de regulación de factores de transcripción e impresión de genes (Kim et al. 2017). En particular, destacan los *ARN COOLAIR* y *COLDAIR* que participan en el silenciamiento epigenético de *FLC* (Figura 4).

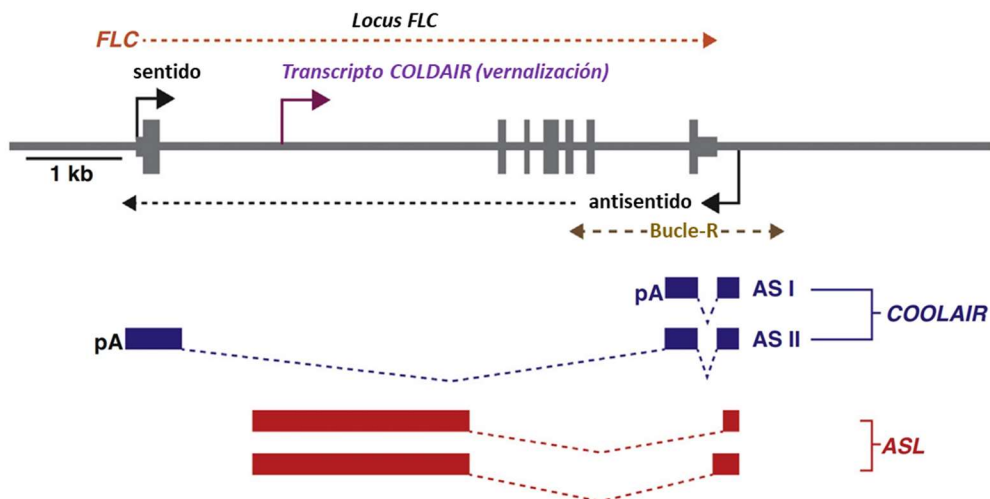


Figura 4. *ARNncs* regulatorios producidos a partir del locus *FLC*. El silenciamiento epigenético se realiza mediante la generación de tres transcritos. En azul, *COOLAIR* y *ASL* se transcriben en dirección antisentido. En morado, *COLDAIR* se transcribe del primer intrón de *FLC* en dirección sentido (Chekanova 2015).

2.4. Funciones de los *ARNlncs* EN FLORACIÓN Y SEQUÍA

Los procesos de desarrollo y respuesta a estrés abiótico se encuentran íntimamente ligados, con muchos elementos que interactúan a nivel molecular, para desencadenar los procesos de maduración floral en respuesta a estímulos ambientales (Vicente et al. 2017). Esto se debe a la presencia de mecanismos moleculares de control que actúan a nivel de factores de transcripción, proteínas y *ARNlncs* (Ying et al. 2014).

2.4.1. Funciones de los *ARNlncs* en floración

Se han descrito diversos *ARNlnc* en especies como *A. thaliana*, *Oryza sativa*, *Brassica rapa*, *Gossypium hirsutum*, *Zea mays*, *Solanum lycopersicum*, entre otros (Liu et al. 2012; Zhang et al. 2014; Huang 2018; Wang et al. 2015; Li et al. 2014; Zhu et al. 2015). Por ejemplo, los *ARNlncs* denotados como *COOLAIR*, *COLDAIR*, *MAS*, *Ef-cd*, *AG-incRNA4*, *LDMAR* están involucrados en el control del tiempo de floración y desarrollo floral (Chen et al. 2021).

- *COLDAIR*

Algunos de los *ARNlncs* más estudiados están relacionados con los procesos de desarrollo. Por ejemplo, el *ARNlnc Cold of Winter-Induced Non-coding RNA (COLDAIR)* es un transcripto intrónico de 1098 nt, metilguanilado y no poliadenilado (Kim et al. 2017). Este *ARNlnc* es producido a partir del intrón I de *FLC* en la misma dirección de *FLC* empezando por el gen *Vernalization Response Element (VRE)*. La expresión de *COLDAIR* se incrementa en respuesta a estrés por frío (Akter et al. 2018). Este *ARNlnc* tiene la capacidad de asociarse a la cromatina en el locus *FLC*, favoreciendo la represión del represor floral *FLC* y dando como resultado la respuesta de vernalización en *A. thaliana* (Kim et al. 2017). Este *ARNlnc* es capaz de regular la expresión genética de *FLC* en forma *trans*, habiéndose demostrado que su sobreexpresión exógena resulta en un fenotipo de floración tardía (Liu et al. 2020).

- *COOLAIR*

Otro *ARNlnc* que participa en la regulación del locus *FLC* se denomina cold-induced *FLC* antisense (*COOLAIR*; Jiao et al.; 2019). Este *ARNlnc* presenta metilguanósina y poliadenina, siendo además capaz de sufrir corte y empalme alternativo en transcriptos de 417 y 750 nt (Hawkes et al. 2016). *COOLAIR* es un transcripto natural antisentido

(*NAT*) de acción cis, que actúa en la regulación transcripcional y postranscripcional de su gen codificante correspondiente, es decir *FLC* (Waseem et al. 2021). La estructura secundaria de *COOLAIR* ha sido estudiada en *A. thaliana*, demostrándose su conservación evolutiva en diferentes miembros de la familia Brassicaceae (Hawkes et al. 2016).

COOLAIR ha sido relacionado con la ruta de floración autónoma por activación de su sitio proximal de corte y empalme por *FCA*, *FY*, *CstF64*, y *CstF77* (Liu et al. 2010). Asimismo, este transcripto está involucrado en la respuesta a vernalización, siendo inducido por exposición prolongada al frío y actuando de forma independiente con complejos *Polycomb* para silenciar el locus *FLC* de forma epigenética (Csorba et al. 2014).

Además, la expresión de *COOLAIR* ha sido validada en especies dicotiledóneas y monocotiledóneas como *Brachypodium*, lo cual resalta la importancia de regulación de la transcripción por *ARNlncs* antisentido (Jiao et al. 2019).

- ***COLDWRAP***

El transcripto denominado *Cold of Winter-Induced Non-coding RNA from the Promoter (COLDWRAP)* tiene 316 nt que se transcribe en dirección sentido con su inicio de transcripción ubicado 225 pares de bases (pb) corriente arriba de *FLC* (Kim y Sung 2017). Este transcripto tiene la capacidad de reprimir de forma estable a *FLC* (Yu y Zhu 2019).

COLDWRAP actúa en conjunto con *COLDAIR* para formar un bucle intragénico represor en el locus de *FLC* durante la vernalización (Kim y Sung 2017). Además, el *ARNlnc COLDWRAP* es necesario para la asociación con *PRC2* y la mutación dentro de esta región impide la asociación a *PRC2* in vitro e in vivo (Kim y Sung 2017). El promotor del *FLC COLDWRAP*, el cual podría interactuar con la represión estable de *FLC* durante la vernalización (Kim et al. 2009).

De esta manera, *COLDWRAP*, como *COLDAIR* tienen un rol importante en el proceso de vernalización (Yu & Zhu 2019). Lo cual ocurre por su asociación con *PRC2*, el cual puede ser reclutado por activación cotranscripcional de ambos *ARNlncs*, o la presencia de una

interacción directa ADN-ARN entre los loci diana y estos *ARNlncs* (Rajewsky et al., 2018).

- **MAS**

Se ha identificado el *ARNInc MAS*, que regula la expresión de *MADS Affecting Flowering4 (MAF4)*, el cual regula el tiempo de floración. El *ARNInc MAS* es un transcrito antisentido *MAF4* que contiene un sitio de poliadenilación y se localiza dentro del núcleo (Waseem et al. 2021). El extremo 5' de *MAS* se inicia en el sitio de terminación de *MAF4*, mientras que el extremo 3' se extiende al intrón 1 de *MAF4* (Zhao et al. 2018). Por esta razón, este transcrito también es considerado un NAT, que se sobreexpresa en condiciones de bajas temperaturas y regula de forma positiva la expresión de *MAF4*, reprimiendo la floración (Zhang et al. 2019). Esta regulación no está ligada a un mecanismo por mediación de ARN pequeños (*sARN*), ni a la modificación de la estabilidad del transcrito de *MAF4*, si no a la expresión de *MAF4* a nivel transcripcional por unión del *ARNInc* a WD40 containing repeat 5a (*WDR5a*), necesarios para la activación de *MAF4* y la floración (Zhao et al. 2018).

- **LINC-AP2**

LINC-AP2 es un *ARNIncis* cercano al FT regulatorio de desarrollo floral de *APETALA2 (AP2)*; Gao et al. 2016). En su trabajo, Gao et al. (2016) encontraron que el locus *LINC-AP2* con la presencia de dos potenciadores, se une a múltiples factores de transcripción reguladores maestros, incluyendo *LEAFY (LFY)*, *APETALA1 (API)* y *SEPALLATA (SEP3)*. La unión de este *ARNInci* a *API*, se ha asociado con un incremento en la expresión de *ARNIncis*, lo cual permite la accesibilidad incrementada a los dos potenciadores correspondientes, contribuyendo a la activación de *AP2* (Chen 2021). Estos hallazgos, soportan lo expuesto por Azofeifa et al. (2018), quienes mencionan que la actividad de los *ARNIncis* asociada a potenciadores, podría ser utilizada de una manera predictiva para distinguir las actividades represoras de activadores en los factores de transcripción.

Por su parte, en dos estudios independientes, Gao et al. (2016) y Waseem et al. (2021) encontraron que la sobreexpresión de *LINC-AP2* se correlaciona con estructuras florales anormales y la incapacidad de apertura de los brotes florales de *A. thaliana* durante la infección con el virus *Turnip Crinkle Virus (TCV)*. Estos efectos podrían deberse al efecto

de que este *ARNlnci* podría ejercer sobre dos estados de cromatina (eucromatina y heterocromatina), lo cual podría proporcionar un modelo único para entender la transición del estado de la cromatina durante el desarrollo floral (Chen 2021).

2.4.2. *ARNlncs* en sequía

El estrés abiótico es capaz de inducir mecanismos moleculares de respuesta, capaces de activar la tolerancia o susceptibilidad. Estos cambios incluyen la activación de cascadas de señalización, cambios bioquímicos y diferentes rutas de transducción de señal (Waseem et al. 2021). Estos cambios incluyen la producción de *ARNlncs* específicos de respuesta a estrés. Estos han sido validados de forma experimental en 43 especies vegetales, habiéndose obtenido un número de 1187 *ARNlncs* relacionados a estrés, los cuales han sido depositados en PLNlncRbase (Xuan et al. 2015).

Los *ARNlncs* regulan la respuesta a sequía en diferentes especies. Por ejemplo, en *A. thaliana* se ha encontrado la presencia de un *ARNlnc* nuclear que mejora la respuesta a estrés hídrico (Qin et al. 2017). En yuca se ha demostrado la presencia de *ARNlncs* como dianas miméticas de *miARNs* en respuesta a estrés por frío o por sequía (Li et al. 2017). En *Pyrus betifolia*, se han encontrado hasta 251 *ARNlncs* de respuesta a estrés y relacionados a procesos de desecación (Wang et al. 2018). Los análisis de regulación cruzada sugieren que muchos *ARNlncs* están asociados con las rutas de transducción en la señalización hormonal, biosíntesis de metabolitos secundarios y metabolismo de la sacarosa (Li et al. 2017), siendo dianas con mucho potencial en el mejoramiento de plantas, que aún puede ser explorada (Han et al. 2018).

2.5. PREDICCIÓN DE *ARNlncs*: CALCULADORA DE POTENCIAL CODIFICANTE (CPC)

Dado que cada vez hay más datos transcriptómicos generados por las tecnologías *NGS* (Qin 2019), ha surgido la necesidad de distinguir los elementos codificantes de los no codificantes. De esta manera, se han desarrollado algoritmos que permiten determinar el potencial codificante de una secuencia, basándose en sus características. Al integrar las características de secuencia con la significancia biológica, *CPC* ha demostrado tener una

gran precisión en la validación cruzada de *ARNIncs* (Kong et al. 2007). El algoritmo *CPC* es más rápido que herramientas como *CONC* (Liu et al. 2006), y está preparado para análisis de alto rendimiento. *CPC* utiliza menos características, obteniendo un desempeño comparable e incluso mejor en la evaluación (Kong et al. 2007). Los resultados demuestran que las características de secuencia utilizadas por *CPC* tienen un alto poder de discriminación y podrían reflejar las propiedades intrínsecas de los transcritos codificantes. (Kong et al. 2007).

CPC proporciona evidencia detallada sobre la razón por la que un transcripto es clasificado como codificante o no codificante, así como otras características de secuencia en la página de evidencias. La página de evidencias muestra las seis características del transcripto, codificado según color para una mejor visualización. También muestra gráficamente el marco abierto de lectura (*ORF*) putativo identificado por las dianas en *Framefinder* y *BLASTX*, así como el detalle de cada *ORF* y diana en *BLASTX*. La página de evidencia también proporciona opciones para comparar el transcripto problema con bases de datos bien anotadas como la base de datos de dominios funcionales *Pfam* (Bateman et al. 2004), *SMART* (Letunic et al. 2006), *SuperFamily* (Madera et al. 2004), *UTRdb* (Mignone et al. 2005) y la base de datos de *ncARNs* *RNAdb* (Pang et al. 2005).

2.6. EL POTENCIAL DE LOS ESTUDIOS TRANSCRIPTÓMICOS EN TARWI

La identificación de elementos regulatorios como los *ARNIncs* requiere la disponibilidad de datos transcriptómicos que permitan anotarlos y relacionarlos con sus respectivas funciones biológicas. Los datos transcriptómicos se han podido obtener en tarwi gracias trabajo de mejoramiento genético realizado en el Programa de Investigación y Proyección Social (PIPS) en Cereales y Granos Nativos, Facultad de Agronomía, Universidad Nacional Agraria La Molina (UNALM). Así, se ha permitido la identificación y evaluación de accesiones promisorias. De manera que en las accesiones P11 y P03 se han realizado estudios de expresión global a nivel de tejido vegetativo y reproductivo en condiciones normales y de estrés por sequía utilizando la técnica de *ARN-seq*.

En el estudio de Ramos (2021), la accesión P11 fue seleccionada por tener precocidad a la floración, mayor peso de 100 semillas, mayores niveles de sobrevivencia a estrés hídrico, adecuado contenido de microelementos (Zn, Mn, Fe) y elevado contenido de

proteínas solubles totales a nivel de semilla. A nivel morfológico, esta accesión desarrolló sus brotes florales en seis estadios, presentando diferencias entre los estadios E4 y E6, en relación a la maduración de sus verticilos florales y aparición de polen. Estas diferencias fueron evaluadas a nivel transcriptómico, encontrándose 2340 genes diferencialmente expresados mediante *ARN-seq* de Illumina, con 1443 genes sobreexpresados y 897 genes subexpresados relacionados a las rutas *KEGG* de biosíntesis de metabolitos secundarios, biosíntesis de fenolpropanoides, interconversiones de pentosas y glucuronato, entre otros.

Por otra parte, el estudio de Hidalgo (2021), permitió seleccionar la accesión P03, una accesión precoz, con tolerancia a sequía, mayor acumulación de biomasa y menor contenido de alcaloides, que es capaz de adaptarse a sequía a nivel foliar modificando sus niveles de prolina. Al analizar las diferencias en la expresión genética entre el tejido vegetativo sometido a estrés hídrico y en condiciones normales, mediante *ARN-seq*, se encontraron 7813 genes diferencialmente expresados, con 4588 genes sobreexpresados y 3225 genes subexpresados. Estos genes fueron clasificados en genes funcionales y regulatorios. Los genes funcionales incluyeron canales y transportadores, genes de detoxificación, genes relacionados a síntesis fosfolípidos, proteasas, fotosíntesis, osmolitos y factores de protección. Los genes con función regulatoria fueron los genes de transducción de señal, hormonas, factores de transcripción, y modificación de la cromatina.

Los datos transcriptómicos de ambos estudios proporcionan una oportunidad para realizar investigación avanzada que permita entender las bases moleculares de algunas características de importancia agronómica como el desarrollo floral y la respuesta a estrés abiótico gobernadas por los *ARNlncs*.

III. MATERIALES Y MÉTODOS

3.1. LUGAR DE EJECUCIÓN

El trabajo bioinformático se realizó en el Laboratorio de Fisiología Molecular de Plantas del PIPS en Cereales y Granos Nativos de la UNALM. Se utilizó un ordenador de escritorio con sistema operativo LINUX versión 5.13.8, procesador Intel Core i7 8700, tarjeta gráfica Gigabyte GTX 2060 DDR6 6Gb, memoria RAM de 64Gb DDR4, almacenamiento SSD NVME de 4Tb y un ordenador portátil Asus ROG Flow x13, con SO Windows 11 Pro, procesador AMD Ryzen 9 5900HS, memoria RAM de 16Gb DDR4 y almacenamiento SSD NVME de 2Tb. La validación se realizó en el Laboratorio de Marcadores Moleculares del PIPS en Cereales y Granos Nativos de la UNALM.

3.2. DATOS TRANSCRIPTÓMICOS, SELECCIÓN DE *READS* Y ENSAMBLAJE

Se utilizaron dos experimentos de *ARN-seq* en las accesiones P3 y P11 de tarwi, provenientes de la Estación Experimental Agraria Santa Ana del Instituto Nacional de Innovación Agraria bajo el Acuerdo de Transferencia de Material Genético *ATMG 001-2015*. El primer experimento analizó el transcriptoma de los brotes florales en dos estadios de desarrollo (Ramos 2021), y el segundo estudió la respuesta del tejido vegetativo a estrés por sequía (Hidalgo 2021).

Doce librerías fueron preparadas por MacroGen Inc., Seúl, Corea (<http://macrogen.com>). La secuenciación total mediante *ARN-seq Illumina* se realizó con el secuenciador *HiSeq 2500* para obtener *reads* pareados. Cada experimento consistió de dos tratamientos con tres replicados biológicos cada uno. Los *reads* de baja calidad con puntaje Phred<30, *reads* cortos (<20 pb), nucleótidos vacíos (N al final de los *reads*), y las secuencias de adaptadores fueron recortados utilizando el software *CutAdapt* para obtener *reads* de alta calidad. El control de calidad antes y después del filtrado se realizó utilizando *FastQC* (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Los datos en crudo comprendieron 598479350 *reads* iniciales que produjeron 581249244 *reads* de alta calidad. Finalmente, estos *reads* de alta calidad fueron ensamblados utilizando TRINITY para obtener el total de transcriptos, produciendo contigs con un porcentaje de GC=38,62 por ciento. Los contigs tuvieron un N50 de 1879, con un promedio de 1184,59 pb. El

ensamblaje TRINITY produjo 172370 transcritos y 83145 genes totales. El reporte de calidad se presenta en el Anexo 1.

3.3. IDENTIFICACIÓN DE *ARNlncs* PUTATIVOS

Los transcritos totales fueron filtrados según la metodología de Khemka et al. (2016). Para ello, se procedió a filtrar los transcritos para obtener los *ARNlnc* putativos. El flujo de trabajo es mostrado en la Figura 5. Los transcritos con una longitud ≤ 200 nt y los *ORFs* putativos con una longitud > 300 nt fueron descartados como transcritos potencialmente codificantes en *A. thaliana*, *Glycine max*, *Medicago truncatula*, y *Lupinus angustifolius* (Hane et al. 2017) con SwissProt en Pfam.

Posteriormente, se evaluó el potencial codificante de los transcritos restantes utilizando cuatro algoritmos diferentes (Anexo 2): *CPC* (Kong et al. 2007), *CPC2* (Yu-Jian et al. 2017), *CNIT* (Jin-Cheng et al. 2019) y *lncFinder* (Han et al. 2018). Los transcritos con valores de *CPC* < 0 ; *CPC2* $< 0,5$; *CNIT* < 0 ; y *lncFinder* = 0, fueron considerados como *ARNlncs*. La selección de los *ARNlncs* putativos se realizó por la coincidencia de por lo menos dos algoritmos.

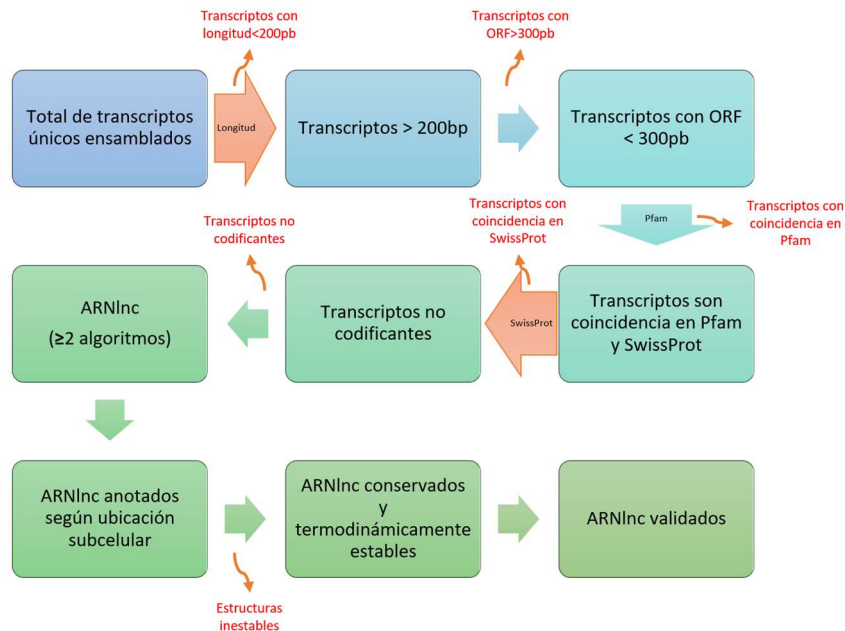


Figura 5. Flujo de trabajo para la identificación de los *ARNlncs* en tarwi.

3.4. OBTENCIÓN DE *ARNlncs* DE ESPECIES VEGETALES

Se utilizó la base de datos *PLncDB* v2.2 (<https://www.tobacodb.org/plncdb/Download>), que contiene la información transcriptómica de 80 especies, con un total de 1246372 *ARNlncs* (Jingjing et al. 2021) mediante un flujo de trabajo que utiliza los algoritmos *CPC* (Kong et al. 2007) y *PLEK* (Li et al. 2014). Estos datos fueron complementados con la información de las bases de datos *ARNlnc*, *GreenC*, *CANTATAdb*, *RNAcentral* y *Ensembl*.

3.5. ANOTACIÓN DE LOS *ARNlncs* PUTATIVOS

Para realizar la anotación de estos transcritos se procedió a la anotación funcional de los *ARNlncs*, por lo cual se realizó una comparación múltiple de secuencias siguiendo la metodología de Hezroni et al. (2015). Así, se utilizó BLASTN de la suite de *BLAST+* 2.9.0, modificando los parámetros "-task blastn -word_size 8 -strand plus". Utilizando los criterios de Khoei et al. (2021), los transcritos con alineamientos con un E-value $<10^{-5}$; identidad >70 por ciento; y longitud de alineamiento >100 nt. fueron considerados como similares en secuencia. Los *ARNlncs* de tarwi, con homología con más de 40 especies de las bases de datos del punto 3.4 fueron tomadas para su evaluación posterior.

3.6. DETERMINACIÓN DE LA LOCALIZACIÓN SUBCELULAR

Se utilizaron las bases de datos de localización subcelular *RNAlocate* (Cui et al. 2022; <http://www.rnallocate.org/iLoc-mRNA.html>), *iLoc-LncRNA* 2,0 (Zhang et al. 2022; [http://lin-group.cn/server/iLoc-LncRNA\(2.0\)/predictor.php](http://lin-group.cn/server/iLoc-LncRNA(2.0)/predictor.php)), *DeepLncLoc* (Zeng et al. 2022; <http://bioinformatics.csu.edu.cn/DeepLncLoc/>) y *lncLocator* 2.0 (Lin et al. 2021; <http://www.csbio.sjtu.edu.cn/bioinf/lncLocator2/index.html>) utilizando los parámetros por defecto. De esta manera, se pudo integrar, analizar e identificar la información de localización subcelular en las secuencias anotadas como *ARNlncs* putativos obtenidas en el paso de anotación de *ARNlnc* putativos. El consenso de los resultados de ubicación subcelular de estas bases de datos se utilizó para predecir la ubicación de cada transcrito de *ARNlnc* en tarwi (Yang et al. 2022).

3.7. CARACTERIZACIÓN TERMODINÁMICA DE LOS *ARN*ncs

Tomando en cuenta la información de Bugnon et al. (2022), la caracterización de los *ARN*ncs se realizó mediante el hallazgo de las estructuras con menor energía libre, para lo cual se emplearon los algoritmos: *RNAfold* (Lorenz et al. 2011; <http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi>), *LinearFold* (Huang et al. 2019; <https://linearfold.eecs.oregonstate.edu/>), *LinearPartition* (Zhang et al. 2020; https://academic.oup.com/bioinformatics/article/36/Supplement_1/i258/5870487) y *RNASHapes* (Steffen et al. 2006; <https://bibiserv.cebitec.uni-bielefeld.de/rnashapes>). Los valores de energía libre mínima (*ELM*; kcal/mol) fueron promediados. Luego, según el criterio de estabilidad termodinámica de Mohammadin et al. (2015), se eliminaron las estructuras con $ELM \geq -80$ kcal/mol. Los parámetros utilizados se detallan en la Tabla 1.

Tabla 1. Parámetros para el cálculo de la energía libre mínima.

Algoritmo	Parámetros
RNAfold	Fold algorithms and basic options minimum free energy (MFE) and partition function avoid isolated base pairs Energy Parameters RNA parameters (Turner model, 2004) Output options interactive RNA secondary structure plot help RNA secondary structure plots with reliability annotation (Partition function folding only) Mountain plot
LinearFold	Arguments Constraints NO (one more line as structure constraints) Zuker suboptimal structures NO Available model(s) LinearFold-C (using CONTRAfold v2.0 machine-learned model, Do et al 2006)
LinearPartition	Available model(s) LinearPartition-C (using CONTRAfold v2.0 machine-learned model, Do et al 2006) subopt: sub-optimal energy deviation relative deviation: 10.0 absolute deviation
RNASHapes	stochastic options structure probabilities compute individual structure probabilities folding options overdangle [overdangle] shape level: 5 Temperature: 37 thermodynamic model parameters Turner model, 2004 [rna_turner2004.par] lonely base pairs no lonely pairs [0]

3.8. CONSERVACIÓN DE LOS TRANSCRIPTOS

El software *Crossalign* (Delli-Ponti 2018, http://service.tartaglialab.com//new_submission/crossalign) fue utilizado para comparar la estructura secundaria de los *ARNlncs* estables ($ELM < -80$ kcal/mol) contra sus homólogos encontrados en la anotación de los *ARNlncs* putativos. Se construyó una matriz para evaluar la similitud basada en la combinación del algoritmo Computational Recognition Algorithm of Secondary Structure (*CROSS*) para predecir el perfil de estructura secundaria del ARN, con resolución de nucleótido único; y el algoritmo Dynamic Time Warping (*DTW*) para alinear perfiles de diferente longitud. La evaluación de la significancia se realizó según el criterio de Delli-Ponti (2018), según el cual un valor de distancia estructural menor que 0,08 y un valor $p < 10^{-5}$ indica similitud estructural significativa y valores $p \geq 0,10$ resulta en un puntaje no significativo. El formato *RNAFold* Ct de cada *ARNlnc* se ploteó con *RiboSketch* (<https://rnastructure.cancer.gov/ribosketch/>) para encontrar similitudes entre estructuras secundarias. Finalmente, el software de descubrimiento de motivos *MEME* v. 5.5.3 (<https://meme-suite.org/meme/tools/meme>; Bailey et al. 2015) se utilizó para predecir nuevos motivos conservados.

3.9. VALIDACIÓN DE PATRONES DE EXPRESIÓN DIFERENCIAL POR RT-PCR

Para validar los patrones de expresión, se utilizó la técnica *RT-PCR* en tarwi. La validación se realizó tomando el material procedente de brotes vegetativos en condiciones hídricas normales y de sequía, según Hidalgo (2021); así como de brotes florales en los estadios E4 y E6 detallados en el trabajo de Ramos (2021). La extracción de ARN total y la remoción de ADN genómico se realizaron según las indicaciones del kit Direct-zol™ *RNA Miniprep Plus* (ZYMO RESEARCH, Irvine, CA). El ARN total fue cuantificado con el nanofotómetro IMPLEN *NP80* y la calidad se evaluó mediante un corrido electroforético. Se eliminó el ADN utilizando el kit *RQN* de Promega, siguiendo las indicaciones del fabricante. La síntesis de ADN complementario (ADNc) se realizó según indicaciones del kit *RevertAid First Strand cDNA Synthesis Kit*, utilizando los primers hexaméricos aleatorios. Los primers específicos para amplificar los *ARNlncs* seleccionados fueron diseñados utilizando *Primer3web*, versión 4.1.0

(<https://primer3.ut.ee/>). Las secuencias son detalladas en la Tabla 2. La amplificación se realizó en el Termociclador BioRad *T100* con tres replicados biológicos y tres replicados técnicos.

La expresión diferencial de los *ARNlncs* fue evaluada tomando los genes de referencia en tarwi reportados por Linares (2023). Estos genes incluyeron *RPL11*, *RPS18*, *RPS27L*, *ERF3*, *OEP24*, *CESA*, *AT4G27220*, y *V-ATPasa*. Al realizar la validación en condiciones de sequía en la accesión P03, se encontró que el gen *OEP24* mantuvo su expresión constante en las plantas sometidas a sequía y en las plantas del grupo control. Este gen codifica para una proteína formadora de poros en la envoltura externa del cloroplasto que se expresa de forma constitutiva en plantas C4, y ha sido previamente utilizado como housekeeping para normalizar los niveles de expresión de enzimas de la biosíntesis de clorofila (Urbischek et al. 2019).

Por otra parte, al evaluar la expresión de los genes de referencia reportados por Linares (2023) en los estadios florales E4 y E6 de tarwi, se encontró expresión constante en *ERF3*. El producto de este gen está involucrado en el proceso de traducción eucariota, formando parte del complejo eucariota de liberación (*eRF1-eRF3*)- GTP (Jackson et al. 2012). Al ser constitutivo, el gen *ERF3* se ha utilizado como gen para la normalización de expresión genética en diversas especies, incluyendo *A. thaliana* (Yang et al. 2015). Ambos genes, *OEP24* y *ERF3* fueron utilizados como un control para la normalización. Las secuencias de los primers utilizados para amplificar las regiones del *ARNInc* y genes de referencia se muestran en la Tabla 2.

Tabla 2. Descripción de los primers para la validación por *RT-PCR*.

Nombre del Gen		Secuencia (5'-3')	Tamaño de producto (pb)	Temperatura de hibridación (°C)
<i>OEP24</i> ^{1a}	F:	GTCTAAGAACTCGTGGGACTTTG	201	60,9
	R:	CATGTGGTCTCTGCACTAAGTTT		
<i>ERF3</i> ^{1b}	F:	CTACAGTGCAGTGGTCAAAAAGAA	312	60,9
	R:	GTACAGTACCCGATTCCACTTTG		
<i>S03</i> ^{2c}	F:	TCACTACTAGGCTGAGCAACC	111	60,4
	R:	TGTTCCCTGCTTCTTGTG		

^a TRINITY_DN20247_c5_g3_i3

^b TRINITY_DN23955_c2_g4_i3

^c TRINITY_DN19379_c1_g3_i8

¹ *OEP24* y *ERF3* fueron utilizados como genes de referencia.

² *S03* fue el *ARNInc* validado.

IV. RESULTADOS Y DISCUSIÓN

4.1. IDENTIFICACIÓN DE LOS *ARNlncs*

Los transcriptos generados del ensamblaje TRINITY por Hidalgo (2022) y Ramos (2022), consistieron de 222217 transcriptos únicos con una longitud mayor a 200 nt (Figura 6a). Luego, se filtró para eliminar transcriptos con *ORF* superiores a 300 nt, lo cual permitió separar 120493 secuencias, que por su tamaño podrían presentar regiones codificantes (Li et al. 2020). Al realizar este filtrado, se obtuvieron 101724 transcriptos, que fueron comparados con las bases de datos *Pfam* y *SwissProt*, eliminándose 100300 transcriptos y obteniéndose 1424 secuencias no codificantes putativas. En contraste, Khemka et al. (2016) encontró mayor cantidad de transcriptos no codificantes putativos (3051) al analizar el transcriptoma de *Cicer arietinum*.

La identificación de *ARNlncs* a partir de las secuencias no codificantes putativas, requiere del uso de herramientas que permitan evaluar el potencial codificante de los transcriptos (Pinkney et al. 2020) y que en combinación permiten incrementar la capacidad de detección de *ARNlncs* (Figura 6b). El algoritmo *CPC* (Kong et al. 2007) ha sido evaluado y tiene un buen desempeño en la identificación de muestras de *ARNlncs* en eucariotas (Duan et al. 2021). Por ello, se utilizó este algoritmo, permitiendo la identificación de 625 *ARNlncs* putativos provenientes del ensamblaje de tarwi. En contraste, Aslam et al. (2022), encontraron 2028 *ARNlncs* en raíces de *Lupinus albus* sometidas a estrés por déficit de P. En tanto Khemka et al. (2016), identificaron 2248 *ARNlncs* mediante *CPC*. La mayor cantidad de *ARNlncs* putativos encontrada en estas leguminosas por Aslam et al. (2022) y Khemka et al. (2016), podría deberse a que se trabajó con ocho estadios de desarrollo en *L. albus*, y ocho estadios sucesivos de desarrollo floral con tres tejidos vegetativos en *Cicer arietinum*, respectivamente. En tanto, los datos de tarwi incluyeron dos estadios de floración y brotes vegetativos sometidos a dos condiciones hídricas diferentes.

El algoritmo *CPC* está basado en una máquina de soporte vectorial (*SVM*) que incluye seis características en su modelo predictivo, incluyendo la cobertura del *ORF* y la similitud de secuencia a genes codificantes de proteínas (Kong et al. 2007). Sin embargo,

a pesar de las ventajas de *CPC*, el desarrollo de herramientas más precisas hace necesario utilizar otros algoritmos para identificar los *ARNlncs* (Pinkney et al. 2020). En la actualidad hay un gran número de herramientas que evalúan diversas características intrínsecas de los *ARNlncs*, permitiendo una mejor caracterización del potencial codificante del ARN (Li et al. 2020). En tal sentido, Ma et al. (2020) y Aslam et al. (2022), han utilizado más de un algoritmo para la predicción de *ARNlncs*. Por ello, en este trabajo, se utilizaron los algoritmos *CNIT*, *CPC2* y *lncFinder* en la posterior identificación de los *ARNlncs* (Figura 6b).

Al utilizar el algoritmo *CNIT*, se identificaron 266 *ARNlncs*, de los cuales 243 fueron compartidos con *CPC2*, 255 *ARNlncs* fueron comunes a *lncFinder*, y un único *ARNlnc* fue identificado solo con *CNIT*. Según Guo et al (2019), *CNIT* está basado en *CNCI*, caracterizándose por realizar una identificación de diferentes organismos eucariotas incluyendo vegetales como *A. thaliana*. Este algoritmo fue usado para evaluar muestras de siete estadios de maduración de fruto y postcosecha en dos especies de *Vaccinium*, permitiendo la identificación de 919 y 1116 transcritos identificados como *ARNlncs* (Wang y Nambeesan 2022). Por otra parte, el mismo algoritmo *CNIT* permitió la identificación de 4 382 *ARNlncs* en *Solanum tuberosum* (Glushkevich et al. 2022).

Al utilizar *CPC2*, se encontraron 429 *ARNlncs* putativos en *L. mutabilis*, de los cuales, 243 fueron compartidos con *CNIT*, 387 algoritmos fueron comunes a *lncFinder*, y 32 *ARNlncs* fueron únicos a *CPC2*. Este algoritmo utiliza la longitud de cada *ORF*, la integridad de los *ORFs*, el puntaje Fickett, y el valor del punto isoeléctrico (Kong et al. 2007). Al analizar los transcritos de 25 especies de plantas con flor mediante *CPC2*, Sang et al. (2021) obtuvieron el número de *ARNlncs* únicos de cada especie, destacando la presencia de 8817 *ARNlncs* en *G. max* y 10904 *ARNlncs* en *M. truncatula*. Así mismo, Ma et al. (2021) identificaron cuatro *ARNlncs* involucrados en la biosíntesis temprana de antocianinas inducida por luz en frutos de *Malus domestica*. *CPC2* también permitió la identificación de 547 *ARNlncs* con patrones de expresión altamente dinámicos en la formación de cuerpos fructíferos en *Fusarium graminearum* (Kim et al. 2018).

En contraste, *lncFinder* permitió la identificación de 547 *ARNlncs* putativos, de los cuales 255 transcritos fueron coincidentes con *CNIT* y 387 con *CPC2*. El algoritmo *lncFinder* evalúa las características heterólogas de secuencia mediante máquina de aprendizaje (Han

et al. 2018). En *Mangifera indica*, Moh et al. (2021) utilizaron valores de *IncFinder* menores o iguales a 0,5 para identificar *ARNlncs* involucrados en crecimiento, procesos de desarrollo, defensa a patógenos y respuesta a estrés. Por otra parte, Wang et al. (2022) también utilizaron *IncFinder* para caracterizar 12029 *ARNlncs* en *Hevea brasiliensis*.

En total, se encontraron 233 transcritos comunes a todos los algoritmos utilizados. El detalle del número de transcritos según algoritmos y la distribución de transcritos se presenta en la Figura 6b, en tanto la distribución y la probabilidad codificante de los *ARNlncs* putativos para cada algoritmo usado en *L. mutabilis* se muestran en la Figura 6c.

A pesar de que los *ARNlncs* son por definición no codificantes, se han encontrado algunos péptidos pequeños que pueden ser producidos a partir de secuencias caracterizadas como *ARNlncs* (Xing et al. 2021). Así, autores como Duan et al. (2021) recomiendan un análisis funcional que permita la caracterización fisicoquímica, anotación y validación de los *ARNlncs*. Por esta razón, se tomaron los transcritos caracterizados como *ARNlncs* por dos algoritmos para iniciar la caracterización posterior.

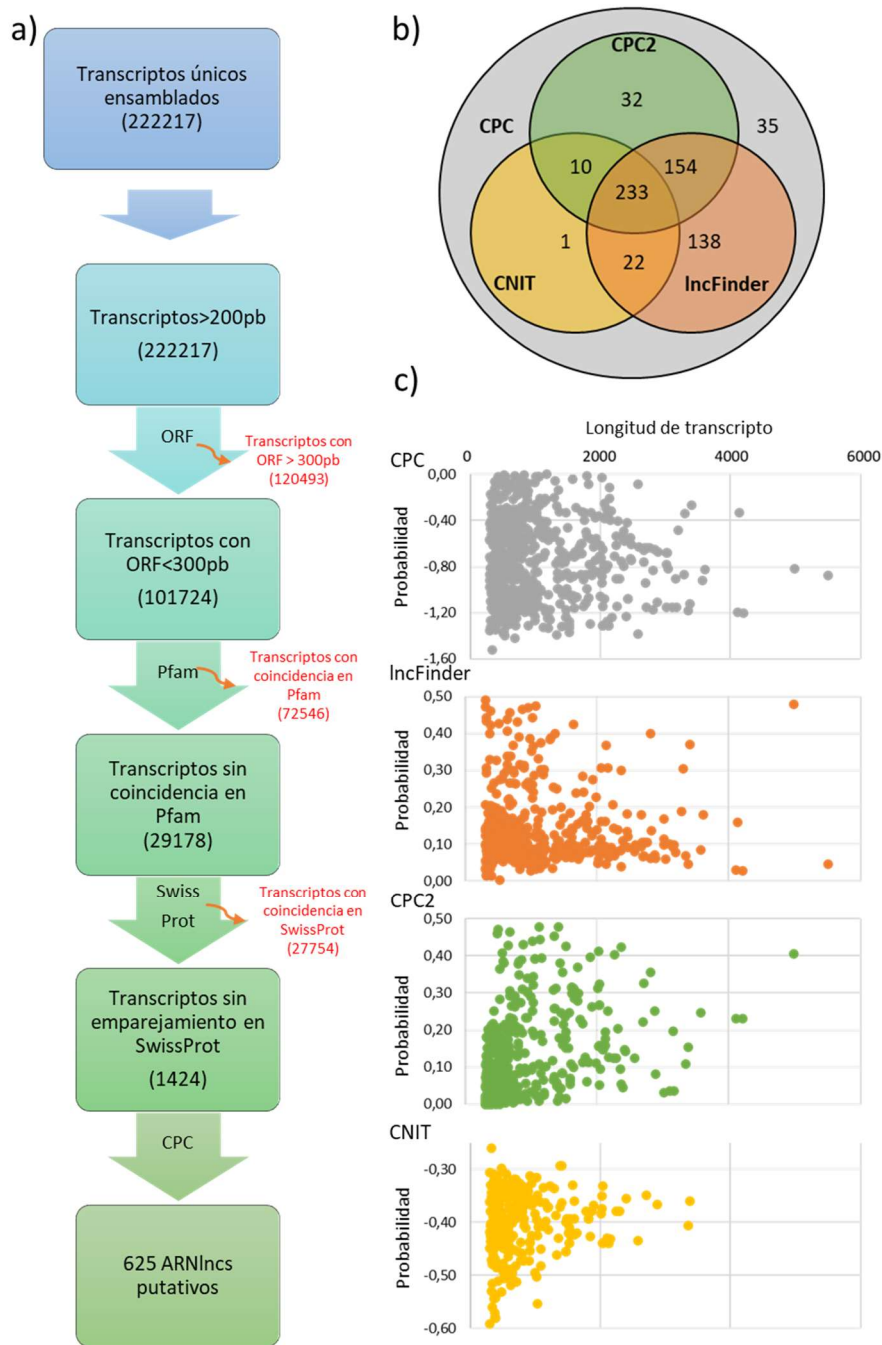


Figura 6. Predicción de los *ARNlncs* en *L. mutabilis*. **a)** Identificación de *ARNlncs*. Los filtros aplicados a los datos de *ARN-seq*, se describen en las flechas del diagrama de flujo. Los cuadros a colores detallan la información tomada en cuenta, mientras las flechas rojas describen los datos excluidos. **b)** Diagrama de Venn para los algoritmos *CPC*, *IncFinder*, *CPC2* y *CNIT* utilizados en la predicción de los *ARNlncs* putativos. **c)** Probabilidad de codificación versus la longitud del transcripto según los algoritmos *CPC*, *IncFinder*, *CPC2* y *CNIT*.

4.2. ANOTACIÓN DE LOS *ARNlncs* PUTATIVOS

La anotación directa de los *ARNlncs* ha sido difícil dada la ausencia de secuencias genómicas, anotaciones parciales de genes codificantes de proteínas y limitación de herramientas para la reconstrucción de transcritos completos de *reads* cortos provenientes de *ARN-seq* (Hezroni et al. 2015). Al reconstruir y filtrar los 590 transcritos, se realizó la anotación por comparación con los *ARNlncs* presentes en 80 especies vegetales de la base de datos *PLncDB* v2.2, siguiendo los criterios de homología de Khoei et al. (2021; E-value $<10^{-5}$; identidad >70 por ciento; y longitud de alineamiento de 100 pb). La Figura 7a muestra la agrupación de los datos en función de las 28 familias botánicas representadas en la base de datos, lo cual permitió distinguir los transcritos comunes a cada familia. Se encontraron 571 transcritos únicos para tarwi y 19 transcritos mostraron homología a nivel de diferentes familias. Las familias que presentaron mayor número de *ARNlncs* homólogos a tarwi fueron Solanaceae (19), Fabaceae (17), Brassicaceae (17), Rutaceae (17), Rosaceae (16), Malvaceae (16), entre otras. Además, se encontraron 14 transcritos comunes a más de 10 familias. El bajo número de secuencias conservadas encontradas entre las diferentes familias demuestra la pobre conservación de *ARNlnc* y la alta tasa evolutiva a la que están sometidas estas secuencias, lo que concuerda con los hallazgos de Sang et al. (2021), quienes encontraron 756 *ARNlncs* conservados presentes en *A. thaliana*.

La anotación sistemática de los *ARNlncs* ortólogos permite establecer que una función biológica putativa se ha mantenido a lo largo del proceso evolutivo (Corona-Gomez et al. 2020). Así, los *ARNlncs* conservados parecen tener funciones más importantes en el control de diversos procesos, según los hallazgos de Sang et al. (2021), quienes encontraron que los *ARNlncs* conservados tienen secuencias más largas, más exones y niveles de expresión mayores que los *ARNlncs* no conservados. Por ello, se tomó el criterio de homología de Hezroni et al. (2016), seleccionando los transcritos comunes para al menos 40 especies. Las especies con 15 o más *ARNlncs* homólogos a tarwi incluyeron a *Capsicum annuum* (19), *Citrus sinensis* (16), *Glycine max* (16), *Gossypium barbadense* (16), *Arachis ipaensis* (15), *Cicer arietinum* (15), *Elaeis guineensis* (15), *Gossypium raimondii* (15), *Medicago truncatula* (15), *Nicotiana tabacum* (15), *Solanum lycopersicum* (15), *Solanum tuberosum* (15), *Trifolium pratense* (15) y *Vigna radiata* (15) lo cual indica la baja conservación presente en las secuencias de *ARNlncs*. El mismo

patrón de conservación fue encontrado en *A. thaliana*, con muy pocos *ARNlncs* conservados (Liu et al. 2012). En contraste, la conservación de *ARNlncs* de sorgo y maíz fue de solo el 25 por ciento según Li et al. (2014). En tarwi, se encontraron 12 *ARNlncs* comunes a 40 especies, los cuales son detallados en la Figura 7b.

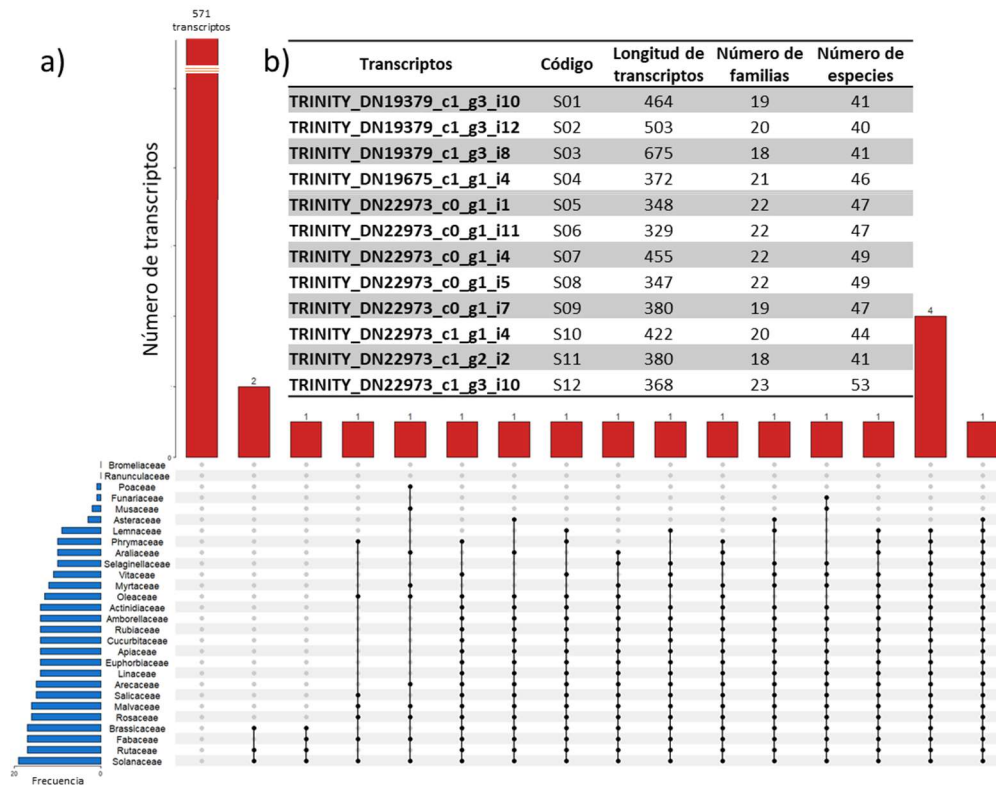


Figura 7. Anotación de los *ARNlncs* de *L. mutabilis*. **a)** Gráfico *UpSet* mostrando la abundancia de transcritos por familias de especies en *PLncDB*. **b)** Detalle de los *ARNlncs* de *L. mutabilis* con el mayor número de homólogos en *PLncDB*.

4.3. LOCALIZACIÓN SUBCELULAR DE *ARNlncs*

Los algoritmos de localización subcelular permiten el análisis de secuencia para identificar transcritos que se ubican a nivel de núcleo, citoplasma, o sistema vesicular. Para determinar la localización subcelular de cada transcritos, se utilizaron cuatro softwares incluyendo RNAlocate v2.0 (Cui et al. 2022), iLoc-lncRNA (2.0), DeepLncLoc (Zeng et al. 2022), y lncLocator 2.0 (Lin et al. 2021), cuyo consenso permitió descubrir la ubicación de los transcritos. La Tabla 3 muestra los resultados de localización de los cuatro algoritmos utilizados, brindando los resultados de localización y sublocalización de los transcritos de *ARNlnc* en tarwi. Del total de transcritos, se encontraron 67 por

ciento de *ARNlncs* presentes a nivel de citoplasma, dentro de los cuales, el 25 por ciento se ubicó a nivel de citosol, donde podrían cumplir roles diversos en la regulación de la expresión genética al interactuar con *ARNm*, ribosomas y proteínas (Noh et al. 2018); y el 75 por ciento tuvo su acción a nivel de ribosomas, donde según Jin et al. (2022), podrían activar la traducción. El 33 por ciento de los *ARNlncs* restantes fue ubicada a nivel de exosomas. Los exosomas son complejos de proteínas y ARN que participan en el metabolismo de ARN al silenciar secuencias repetitivas mediante su degradación (Shin et al. 2013). En *A. thaliana*, los *ARNlncs* presentes a nivel de exosomas son producidos a partir de potenciadores, siendo regulados por el exosoma e interactuando con otras regiones del ADN (Wang et al. 2017). Los *ARNlncs* presentes en exosomas podrían modular la interacción entre los elementos regulatorios que controlan la expresión genética y la organización nuclear al regular otros *ARNlncs* producidos a partir de estos elementos (Chekanova et al. 2015). Por ello, su presencia en el transcriptoma de tarwi podría indicar su participación en procesos de regulación genética.

Tabla 3. Localización subcelular de los *ARNlncs* de tarwi. Los números entre paréntesis indican la probabilidad de que el transcripto se ubique en la ubicación designada según el software respectivo.

ID de Transcripto	Código	RNALOCATE	iLoc-LncRNA(2.0)	DeepLncLoc	lncLocator 2.0	Ubicación predicha	Sububicación predicha
TRINITY_DN19379_c1_g3_i10	S01	Ribosoma (0,99)	Ribosoma (0,99)	Citoplasma (0,56)	Núcleo (-1,04)	Citoplasma	Ribosoma
TRINITY_DN19379_c1_g3_i12	S02	Ribosoma (0,58)	Ribosoma (0,58)	Citoplasma (0,49)	Núcleo (-1,11)	Citoplasma	Ribosoma
TRINITY_DN19379_c1_g3_i8	S03	Citoplasma, Citosol (0,70)	Citoplasma, Citosol (0,70)	Citoplasma (0,62)	Núcleo (-0,56)	Citoplasma	Citosol
TRINITY_DN19675_c1_g1_i4	S04	Citoplasma, Citosol (0,99)	Citoplasma, Citosol (0,99)	Citoplasma (0,67)	Núcleo (-1,31)	Citoplasma	Citosol
TRINITY_DN22973_c0_g1_i1	S05	Ribosoma (0,53)	Ribosoma (0,53)	Citoplasma (0,67)	Núcleo (-0,62)	Citoplasma	Ribosoma
TRINITY_DN22973_c0_g1_i11	S06	Ribosoma (0,72)	Ribosoma (0,72)	Citoplasma (0,76)	Núcleo (-1,20)	Citoplasma	Ribosoma
TRINITY_DN22973_c0_g1_i4	S07	Ribosoma (0,82)	Ribosoma (0,82)	Citoplasma (0,63)	Núcleo (-1,18)	Citoplasma	Ribosoma
TRINITY_DN22973_c0_g1_i5	S08	Exosoma (0,83)	Exosoma (0,83)	Citoplasma (0,64)	Núcleo (-1,16)	Exosoma	Exosoma
TRINITY_DN22973_c0_g1_i7	S09	Exosoma (0,67)	Exosoma (0,67)	Citoplasma (0,65)	Núcleo (-1,14)	Exosoma	Exosoma
TRINITY_DN22973_c1_g1_i4	S10	Exosoma (0,77)	Exosoma (0,77)	Citoplasma (0,48)	Núcleo (1,32)	Exosoma	Exosoma
TRINITY_DN22973_c1_g2_i2	S11	Exosoma (0,59)	Exosoma (0,59)	Citoplasma (0,71)	Núcleo (-1,05)	Exosoma	Exosoma
TRINITY_DN22973_c1_g3_i10	S12	Ribosoma (0,83)	Ribosoma (0,83)	Citoplasma (0,68)	Núcleo (-1,04)	Citoplasma	Ribosoma

4.4. CARACTERIZACIÓN TERMODINÁMICA Y ESTRUCTURAL DE LOS *ARNlncs*

La cadena de polinucleótidos de la molécula de ARN constituye su estructura primaria y se pliega como resultado del apareamiento, generando estructuras de orden superior, como las estructuras secundarias y terciarias, y es responsable de la mayor parte de la energía termodinámica de cada estructura (Herschlag et al. 2018). También, desde una perspectiva termodinámica, la estructura más estable aparece cuando la energía libre se minimiza (Mathews et al. 1999) y se utiliza junto con otros criterios para la identificación de *ARNlncs* funcionales (Trotta 2014). En esta investigación, se utilizaron cuatro algoritmos, incluyendo RNAfold (Lorenz et al. 2011), LinearFold (Huang et al. 2019), LinearPar-tition (Zhang et al. 2020) y RNASHapes (Steffen et al. 2006), para medir la energía libre mínima (*ELM*) de cada *ARNlnc* (Figura 8a).

Las secuencias S02 y S12 mostraron valores sobre -10 kcal/mol y 9 transcriptos (S01, S04, S05, S06, S07, S08, S09, S10, S11) mostraron un valor de *ELM* de entre -10 y -50 kcal/mol. En tanto, el transcripto S03, con homología en 41 especies, tuvo un valor de *ELM* de 105,67 kcal/mol (Figura 8a), el cual puede considerarse como estable (Mohammadin et al. 2015). Las estructuras de orden superior estables contienen estructuras y bolsillos con información capaz de interactuar con otros elementos a nivel molecular (Statello et al. 2021). Por lo tanto, dado su baja *ELM*, S03 podría formar estructuras de orden superior involucrados en la regulación genética de tarwi.

Dado que la estructura del ARN está altamente limitada por la estructura bidimensional (O'Leary et al. 2022), se analizó la estructura secundaria de S03 generada por RNAFold Delta (Figura 8b). Una molécula de ARN contiene estructuras rígidas y estáticas, así como estructuras altamente dinámicas (Mailler et al. 2019), que pueden ser organizadas en regiones. Según Ding et al. (2014), estas regiones de ARN contienen motivos estructurales distintivos que se agrupan al interior de una molécula. En el caso de *ARNlnc* S03, la estructura muestra tres regiones distintas y bien definidas que incluyen dos multibucles y un tallo corto con un bucle interno (Figura 8b), las que son descritas según Bugnon et al. (2022).

Cada región tiene diferentes motivos (bucles, tallos, horquillas, protuberancias, etc.), que son patrones estructurales tridimensionales que emergen de un grupo particular de interacciones que proporciona estabilidad y funcionalidad (Butcher y Pyle 2011). La región I muestra un bucle con múltiples ramificaciones (Figura 8b-I) que da origen a seis horquillas con bucles terminales (h1, h2, h3, h4, h5 y h6) y dos tallos cortos (s1 y s2). Las horquillas h1 y h2 presentan dos bucles internos, mientras que h3, h4 y h5 tienen un bucle terminal grande, en tanto, h6 muestra dos protuberancias con un bucle terminal en su tallo. La complejidad observada en la región I indica la presencia potencial de estructuras con funciones específicas que interactúan con otros elementos (Ross y Ulitsky 2022).

La región II (Figura 8b-II), posee un tallo corto con un bucle interno. Según Ross y Ulitsky (2022), estas regiones cortas tienden a ser menos importantes a nivel funcional porque son más simples, no incluyen otros elementos y son menos rígidas. Esta flexibilidad permite que adquieran diferentes conformaciones nivel tridimensional, al reorganizar estructuras secundarias para formar diferentes estructuras terciarias (Ganser et al. 2019). Las conformaciones producidas pueden variar en el espacio, dado que ellas dependen de interacciones que podrían ocurrir a temperaturas fisiológicas, produciendo un cambio en la ELM total de la molécula (Spitale y Incarnato 2023). Estas propiedades permiten la conexión de elementos funcionales (Mailler et al. 2019), como puede verse en S03, en la cual la región II conecta las regiones más complejas I y II. Sin embargo, la configuración de S03 podría no ser variable en el tiempo, a pesar de la flexibilidad de la región II, tal como se describe por Mustoe et al. (2014). Esto podría deberse a las restricciones termodinámicas entre las posibles estructuras terciarias de los *ARNlncs* (Spitale y Incarnato 2023).

La región III (Figura 8b-III) se conecta con un bucle interno de la región II y también tiene una estructura compleja que incluye un segundo bucle con múltiples ramas con dos tallos cortos (s3 y s4) y cuatro horquillas con protuberancias y bucles internos (h7, h8, h9 y h10). Las horquillas h7, h9 y h10 tienen 2-4 protuberancias, y las horquillas h9 y h10 incluyen bucles internos. De forma similar, h7, h8, h9 y h10 tienen un bucle externo. Estos tipos de estructuras permiten la interacción de *ARNlncs* con proteínas, regulando la transcripción y traducción (Shaath et al. 2022). En particular, la horquilla h10, tiene una estructura similar a la secuencia de inserción de selenocisteína (*elemento similar a*

SECIS), que incluye una estructura con un bucle terminal que incluye una secuencia AGU en el tallo adyacente a la segunda protuberancia.

La presencia de estructuras como tallos con protuberancias y bucles internos, como los observados en las regiones I y III de S03, incluyendo h1, h2, h7, h9 y h10 (Figura 8b), podrían indicar que este transcrito tiene funcionalidad biológica. De hecho, estos tipos de estructuras han sido involucradas en la interacción entre *ARNlncs* y proteínas, y son capaces de regular la transcripción, el corte y empalme, y la traducción (Shaath et al. 2022). Además, de acuerdo a Li et al. (2016), los bucles, protuberancias y horquillas presentes en estas estructuras podrían estar involucrados en la interacción con ARNm (Li et al. 2016) y la regulación de diferentes procesos metabólicos, tal como podría ser el caso del *elemento similar a SECIS* en S03.

El elemento *SECIS* presente en algunos ARNs es un elemento estructural que contiene UGA (Svoboda y Di Cara 2006; Anexo 3) y permite la inserción de selenocisteína en una proteína. La selenocisteína es un selenoaminoácido que puede ser tomado por transportadores de membrana a nivel de las raíces (Hu et al. 2022). El elemento *SECIS* es un sitio de unión para la proteína SECIS-binding protein 2 (*SBP2*). Esta proteína redefine el codón UGA de terminación para introducir el aminoácido selenocisteína durante su síntesis (Mix et al. 2007). Este mecanismo es propio de los ARNm, pero puede ser extendido a secuencias no codificantes. Por ejemplo, el *elemento similar a SECIS* fue encontrado en el *ARNlnc* CCDC152 (Mita et al. 2021). Este *ARNlnc* puede interactuar con los ARNm que codifican selenoproteínas, inhibiendo la unión de la proteína *SECIS-binding protein 2* y reduciendo su afinidad por los ribosomas. De esta manera, CCDC152 es capaz de reducir la expresión de la selenoproteína P (Mita et al. 2021).

Fajardo et al. (2014) reportaron el elemento *SECIS* en plantas vasculares, proporcionando información sobre los mecanismos genéticos que regulan el metabolismo de selenio en plantas. De hecho, Frías et al. (2009) encontró selenometilselenocisteína en semillas de *L. angustifolius*. Sin embargo, el rol de las selenoproteínas podría ser desvirtuado por la falta de comprensión de los mecanismos moleculares que controlan diferentes procesos a nivel genético. Una de estas proteínas es Molybdenum cofactor sulfurase (*LOS5/ABA3*), la cual se expresa en tarwi en condiciones de sequía (Hidalgo 2021). Esta es una liasa de selenocisteína involucrada en la generación de un cofactor de molibdeno (Xiong et al.

2001) necesario para la actividad de aldehído oxidasa en el paso final de la biosíntesis de ABA en plantas (Watanabe et al. 2018). Por lo tanto, *LOS5/ABA3* podría influenciar la respuesta a estrés en plantas. Además, el ABA y el Se han sido reportados para disminuir el estrés por sequía en algunas plantas (Ramasamy et al. 2022). Notablemente, las selenoproteínas O están asociadas con la eliminación de las especies reactivas de oxígeno y la tolerancia a deshidratación en *A. thaliana* (Fichman et al. 2018). Estas proteínas también están involucradas en la regulación de las funciones de óxido-reducción al inactivar H₂O₂ y otros compuestos tóxicos (Bela et al. 2015). Sin embargo, se requiere más investigación para entender la homeostasis de Se bajo sequía y la regulación de su metabolismo por los *ARNlncs*.

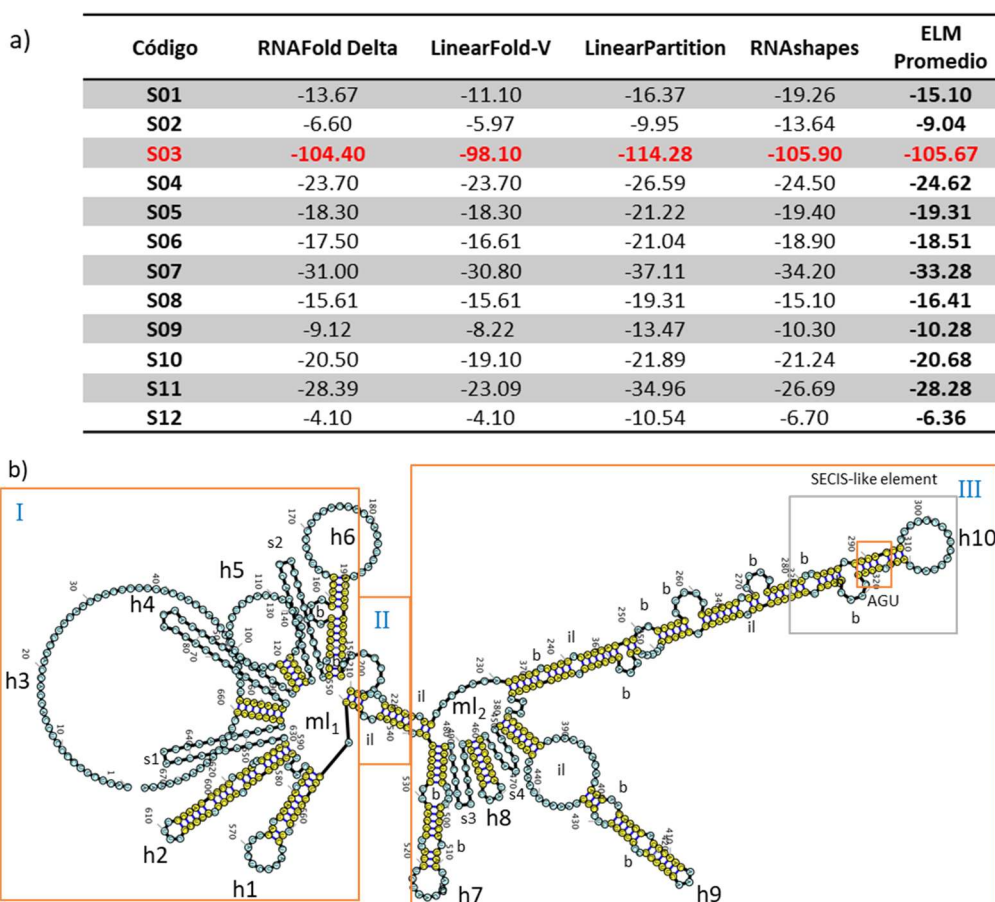


Figura 8. Energía libre y estructura secundaria del *ARNlnc* S03 de *L. mutabilis*. **a)** Energía libre para los *ARNlncs* de *L. mutabilis*. **b)** Estructura secundaria de S03 según *RNAFold*. I: Bucle multiramificado con ocho ramas. II: Tallo con bucle interno. III. Multibucle con 6 ramas. h: horquilla; ml: multibucle, hl: horquilla-bucle; el: bucle externo; il: bucle interno; b: bucle abultado; s: tallo.

4.5. CONSERVACIÓN ESTRUCTURAL DE *ARNlncs*

Los *ARNlncs* funcionales presentan relativamente baja identidad a nivel de secuencia, pero sus estructuras secundarias se conservan en miles de transcritos (Sanbonmatsu 2022). Estas estructuras dependen del apareamiento de múltiples bases nitrogenadas que generan estructuras secundarias estables que no pueden ser alteradas a temperaturas fisiológicas. Por lo tanto, la conservación estructural del *ARNlnc* S03 con sus 41 homólogos fue evaluada mediante el método *Crossalign* (Delli-Ponti et al. 2018), que predice similitudes estructurales entre dos ARN según las probabilidades de hibridación, nucleótidos vecinos y el ambiente celular (Ponti et al. 2020). Según el criterio de Delli-Ponti (2018), una distancia de *Crossalign* menor a 0,08 y un valor $p < 10^{-5}$ para determinar similitud estructural (Anexo 4). Sin embargo, en este trabajo no se encontraron transcritos que cumplan con estos criterios. Por lo tanto, se decidió estudiar los cinco *ARNlncs* que mostraron una distancia estructural normalizada $< 0,1$ y un valor $p < 0,05$ (Figura 9a) para descubrir similitudes estructurales entre los *ARNlncs*. Dado que un valor $p > 0,10$ indica un valor no significativo (Delli-Ponti et al. 2018).

Los *ARNlncs* que cumplen estas características incluyen Solyc04r022210.1 (*S. lycopersicum*), BNAP_LNC003827.4 (*B. napus*), BRAP_LNC003139.3 (*B. rapa*), MTRU_LNC004005.50 (*M. truncatula*), ACHI_LNC008871.8 (*A. chinensis*) y S03 (*L. mutabilis*). Tal como se muestra en la Figura 9a, el *ARNlnc* S03 de *L. mutabilis* comparte un grado de conservación estructural con Solyc04r022210.1 de *S. lycopersicum* con un valor de distancia estructural normalizada de perfiles de estructura secundaria de 0,095, obtenido en la región de 115-790 nt y que muestra una correlación de 0,91 con una significancia de 0,033. La matriz también permitió establecer la similitud entre BNAP_LNC003827.4 de *B. napus* y BRAP_LNC003139.3 de *B. rapa*. Estos *ARNlncs* tienen un valor de distancia estructural de 0,047 entre los nucleótidos 267-997, con una correlación de 0,94 y un valor p significativo = 0,00013. Finalmente, los transcritos BRAP_LNC003139.3 de *B. rapa* y ACHI_LNC008871.8 de *A. chinensis* mostraron un valor de distancia estructural de 0,093 en la región 265-1021, con una correlación de 0,92 y un valor $p = 0,016$.

El análisis estructural se complementó con el software MEME (Bailey et al. 2015), que fue utilizado para comparar los motivos entre *ARNlncs*. Este software predijo la presencia

de cinco motivos, en los transcriptos de las seis especies (Figura 9b). Sin embargo, el motivo 5 no fue encontrado en el *ARNlnc* S03 de *L. mutabilis*. Las similitudes entre estructuras secundarias fueron evidentes al nivel de un multibucle marcado con un círculo rojo en la Figura 9c. Este patrón topológico, considerado la suma de los arreglos físicos o lógicos de nudos y conexiones dentro de una red, es más importante que los apareamientos individuales (Huang et al. 2016). Lo que indica un rol más interactivo con diferentes elementos a nivel molecular (Li et al. 2021).

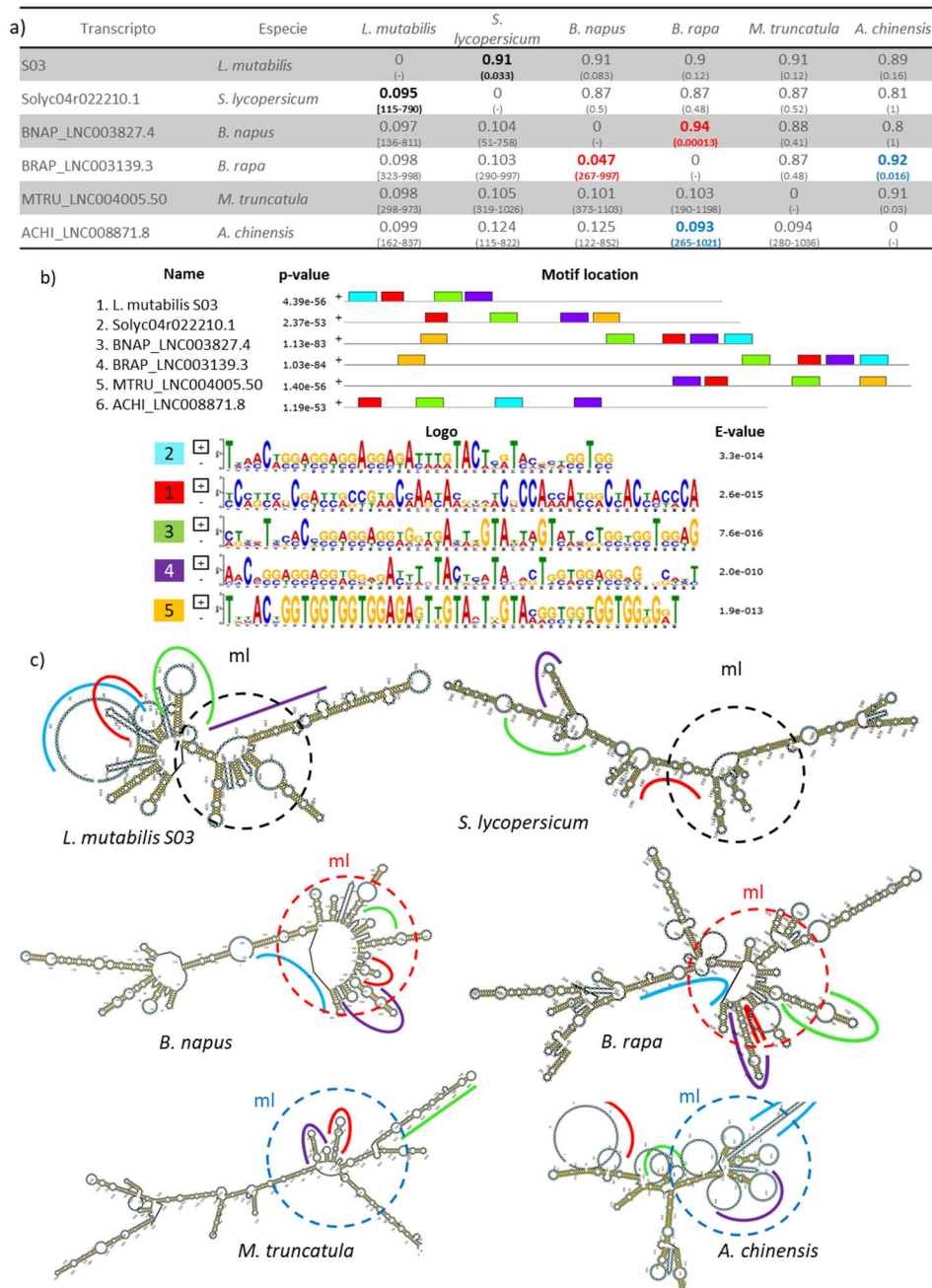


Figura 9. Conservación estructural del *ARNlnc* S03 de *L. mutabilis*. **a)** Distancias *Crossalign* entre S03 y sus homólogos. Bajo la diagonal, datos y región de las distancias estructurales normalizadas (corchetes). Sobre la diagonal, los valores de correlación y significancia (paréntesis). **b)** Motivos identificados por el software *MEME*. Los valores “e” correspondientes de *MEME* se mencionan. **c)** Las estructuras secundarias predichas de los *ARNlncs* fueron modeladas en *RNAFold*. Las líneas coloreadas muestran regiones alineadas en los perfiles completos.

4.6. EXPRESIÓN Y VALIDACIÓN POR *RT-PCR*

Los *ARNlncs* son elementos moleculares involucrados en el control de múltiples procesos a nivel molecular, incluyendo el desarrollo y la respuesta a estrés (Yu et al. 2019). Dentro de los factores de estrés, la sequía es el estrés agrícola más dañino a nivel mundial (Kumar et al. 2021) al generar la pérdida de casi el 20 por ciento de rendimientos potenciales (Ashraf 2010). En efecto, los Andes de Perú han sido afectados por sequías extremas (Poveda et al. 2020). Los eventos de sequía son especialmente nocivos al afectar el desarrollo durante la etapa reproductiva, durante las fases de floración y llenado de granos, afectando la producción de las leguminosas de grano (Khatun et al. 2021), como tarwi. Por esta razón, se procedió a validar la presencia del *ARNlnc* S03 en condiciones de estrés hídrico y en diferentes estados de maduración floral (Figura 10). La validación se realizó únicamente para el *ARNlnc* S03, dado que no pudieron generarse primers para S05, S07, S08, S09 y S10, en tanto que los *ARNlncs* S01, S02, S04, S11 y S12 no pudieron ser amplificados. De forma coincidente, el *ARNlnc* S03 validado en este estudio fue el único que alcanzó la estabilidad estructural según el criterio de $ELM < -80$ kcal/mol (Mohammadin et al. 2015).

El *ARNlnc* S03 se expresó en el estadio E4, lo cual podría señalar su importancia en la diferenciación temprana de los brotes florales. En este estadio se observaron 3 verticilos, con ausencia de maduración de polen en inflorescencias de 7,40-9,44 mm de altura, 4,9-5,9 mm de diámetro y 0,042-0,060 g de peso (Figura 10a). En contraste, en el estadio E6, no se observó expresión de S03. A nivel morfológico, la caracterización denota principalmente un incremento longitudinal (12,8-13,6 mm) ocasionada por un mayor número de verticilos (4-5), además que se observó aparición de polen en el verticilo basal y aumento de peso del brote floral (0,09-0,10 g). Estos cambios a nivel de la arquitectura de la inflorescencia podrían estar asociados a la variación en la expresión de S03 y otros transcriptos.

De forma similar, existen diversos *ARNlncs* asociados con el desarrollo reproductivo, que son capaces de regular el tiempo de floración y desarrollo floral en plantas. Entre los *ARNlncs* más caracterizados se encuentran los reguladores negativos de *FLC* (FLOWERING LOCUS C), *COOLAIR*, *COLD AIR*, *COLDWRAP*, involucrados en la ruta autónoma y de vernalización en diferentes especies (Chen et al. 2021). Estos *ARNlncs*

están relacionados con la inducción de la transición floral, es decir, la conversión de brotes vegetativos en brotes reproductivos. Durante las primeras etapas de floración, la expresión de *ARNlncs* específicos se incrementa, induciendo el proceso de floración (Wang y Chekanova 2017), luego de lo cual, disminuye. Este cambio en la expresión también se observó en los brotes florales de *L. mutabilis* (Figura 10a).

Por otra parte, en *L. mutabilis* sometido a 10 días de sequía progresiva (1,51 por ciento agua en sustrato), se encontró sobreexpresión de S03 al ser comparado con el control bajo condiciones hídricas normales (20,89 por ciento de agua en sustrato), tal como se puede observar en la Figura 10b. Dado que este *ARNlnc* se expresa en una accesión tolerante a sequía e incluye elementos capaces de interactuar con diferentes dianas a nivel molecular (Figura 8b), se propone que S03 es importante en la respuesta de *L. mutabilis* a sequía. Además, el elemento similar a *SECIS* está involucrado en la síntesis de selenoproteínas, participando en la homeostasis de selenio (Steinbrenner y Sies 2013). Este podría ser el caso de S03 de *L. mutabilis* en la respuesta a sequía.

La inserción de Se a nivel proteico parece ser un punto fundamental durante la respuesta a sequía. Esto es soportado por Ahmad et al. (2016), quienes reportan que, en plantas sometidas a estrés hídrico, el Se podría regular el estado hídrico e incrementar la biomasa a través de la actividad antioxidante de enzimas como glutatión peroxidasa. Esta inserción de Se podría ser regulada a nivel postraduccional por la presencia de algunos *ARNlncs*. De hecho, se ha observado que, durante el estrés hídrico, estas moléculas pueden afectar a todos los elementos genéticos, incluyendo promotores, regiones no traducidas, exones, intrones y terminadores, que controlan la expresión genética en múltiples niveles, incluyendo la modificación de la accesibilidad de la cromatina, transcripción, corte y empalme y traducción (Wierzbicki et al. 2021).

En tal sentido, la sobreexpresión del *ARNlnc* S03 podría estar asociada a tolerancia a sequía, como se ha observado en otras especies con *ARNlncs* específicos (Jampala et al. 2021). Por ejemplo, en *A. thaliana*, el *ARNlnc* DRIR (Drought Induced lncRNA) se sobreexpresa en condiciones de deshidratación y tratamiento con ABA, actuando como un regulador positivo de la tolerancia a estrés; a su vez, este transcripto induce tolerancia a salinidad y sequía, e incremento en la respuesta a ABA en líneas que lo sobreexpresan (Qin et al. 2017). Estos resultados también coinciden con los hallazgos de Chen et al.

(2023) en líneas transgénicas de *G. max* donde la sobreexpresión del *ARNlnc lncRNA77580*, mejoró la tolerancia a estrés por sequía, relacionándose con incremento en el rendimiento al inducir mayor número de semillas por planta. A su vez, Li et al. (2017) encontraron el *ARNlnc lincRNA340* en *Manihot esculenta*, el cual es inducido para reprimir a *miR169*, lo cual libera la expresión del factor nuclear Y (*NF-Y*) para aumentar la tolerancia a estrés por sequía actuando como una diana mimética de *miARNs* y transcritos asociados con las rutas de transducción en la señalización hormonal, biosíntesis de metabolitos secundarios y metabolismo de la sacarosa (Li et al. 2017). Por lo tanto, S03 también podría estar involucrado en los procesos de señalización en respuesta a déficit hídrico en *L. mutabilis*.

Dado que S03 se expresa de forma diferencial durante el desarrollo floral y en respuesta a sequía, su presencia podría estar relacionada con la regulación trasversa de estos procesos. Esta regulación es fundamental en las leguminosas de grano, en las cuales la producción de semillas y la seguridad alimentaria dependen del desarrollo reproductivo adecuado (Hidalgo et al. 2021). Sin embargo, la producción de las plantas puede disminuir hasta en 90 por ciento cuando están sometidas a estrés hídrico durante la etapa reproductiva (Lafitte et al. 2007). De hecho, cuando las sequías afectan a las plantas durante el estadio reproductivo, incrementa la esterilidad del polen, causa aborto de brotes florales e induce senescencia, lo que disminuye la producción de granos (Aroca 2013). Por esta razón, es necesario estudiar los *ARNlncs*, como S03, para comprender los mecanismos de regulación que operan a nivel molecular durante el desarrollo floral y en respuesta a sequía, así como las funciones de los *ARNlncs* en el control de procesos biológicos.

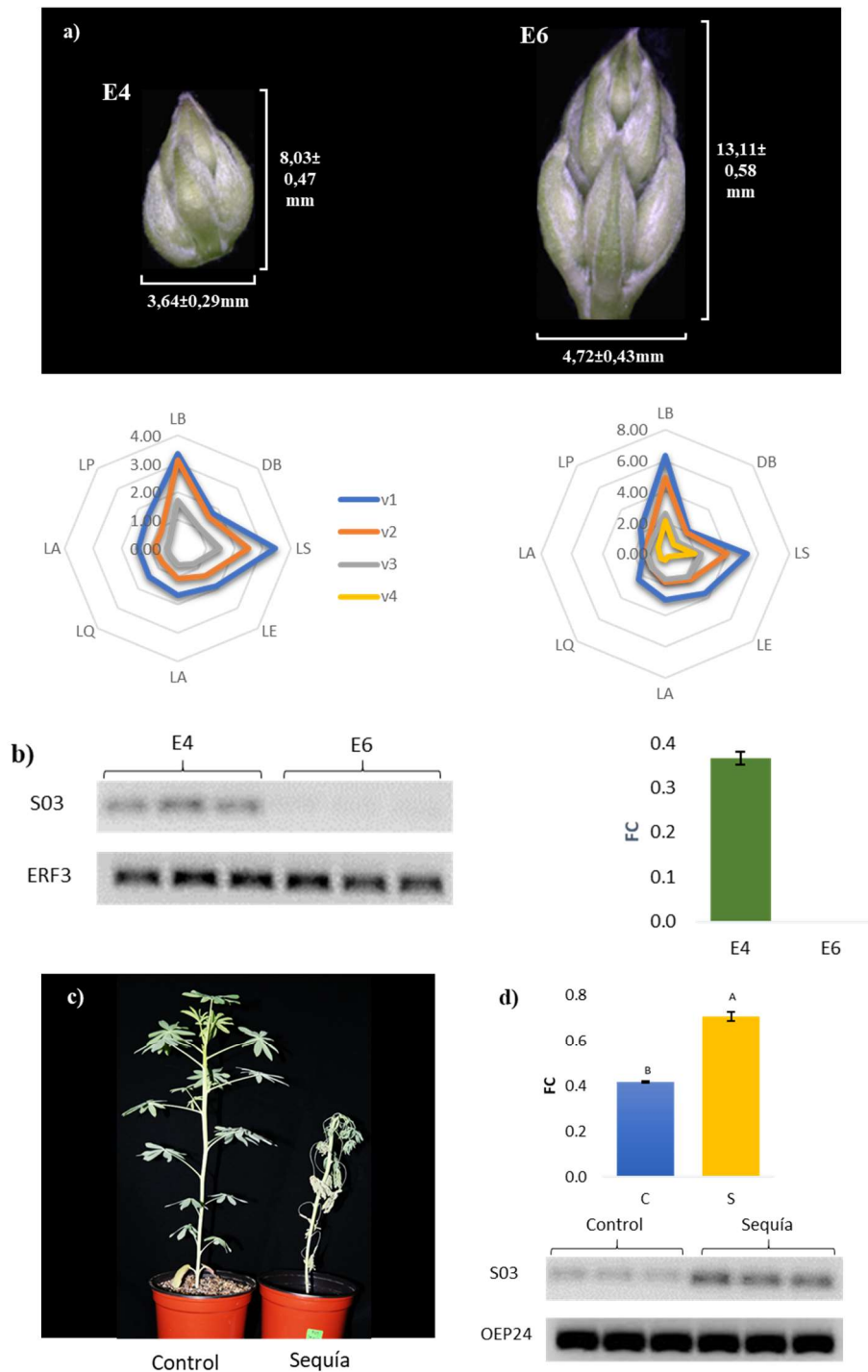


Figura 10. Validación del protocolo de *RT-PCR* aplicado en *L. mutabilis*. **a)** Caracterización morfológica de los brotes florales de *L. mutabilis*. **b)** Expresión de S03 en brotes florales en comparación con *ERF3*. **c)** *L. mutabilis* sometido a 10 días de sequía progresiva. **d)** Resultado de *RT-PCR* de S03 en condiciones de sequía en comparación con *OEP24*. Los resultados fueron obtenidos en base a tres repeticiones. Las barras de colores indican el promedio. Las líneas con remate expresan el error estándar. Letras diferentes indican diferencias según Tukey HSD ($p < 0,05$).

V. CONCLUSIONES

- Se identificaron 590 *ARNlncs* en *L. mutabilis* según dos algoritmos (*CPC*, *CPC2*, *CNIT* o *lncFinder*), utilizando datos de *ARN-seq*.
- Se identificó el *ARNlnc* S03, con homología en más de 40 especies vegetales, estabilidad termodinámica (-105,67kcal/mol) y que incluye las regiones I, II y III, con un tallo largo, dos multibucles y un elemento similar a *SECIS* en su estructura, el cual podría ser requerido para la regulación de la síntesis de selenoproteínas, que son proteínas involucradas en la homeostasis de Se y podrían tener un rol en la floración y respuesta a sequía.
- A nivel estructural, existe conservación estructural por *Crossalign* entre S03 de *L. mutabilis* y Solyc04r022210.1 de *S. lycopersicum*, lo cual se demuestra mediante el análisis de topología, por la similitud estructural en el multibucle de ambos transcritos.
- El *ARNlnc* S03 estuvo sobreexpresado en la accesión tolerante a sequía P03 de *L. mutabilis* sometida a condiciones de sequía.
- El *ARNlnc* S03 mostró expresión diferencial, apareciendo en el estadio E4 y desapareciendo en el estadio E6 de la accesión P11 de *L. mutabilis*.
- Estos resultados resaltan la importancia de los *ARNlncs* en los procesos de floración y respuesta a sequía y pueden ser utilizados en el mejoramiento de tarwi para el desarrollo de marcadores genéticos.

VI. REFERENCIAS BIBLIOGRÁFICAS

- Ahmad, A., Lin, H., Shatabda, S. 2020. Locate-R: Subcellular localization of long non-coding RNAs using nucleotide compositions. *Genomics* 112(3):2583-2589.
- Ahmad, R., Waraich, E. A., Nawaz, F., Ashraf, M. Y., Khalid, M. 2016. Selenium (Se) Improves Drought Tolerance in Crop Plants—A Myth or Fact?: Selenium (Se) and Drought Tolerance in Crop Plants. *Journal of the Science of Food and Agriculture* 96:372–380.
- Akerman, I., Tu, Z., Beucher, A., Rolando, D. M. Y., Sauty-Colace, C. 2017. Human pancreatic β cell lncRNAs control cell-specific regulatory networks. *Cell Metabolism* 25:400–411.
- Akinci, S. (Ed.). 2013. Responses of Organisms to Water Stress (1st Ed.). InTech Janeza Trdine. 190 pp.
- Akter, A., Nishida, N., Takada, S., Itabashi, E., Osabe, K., Shea, DJ, Fujimoto, R. 2018. Genetic and Epigenetic Regulation of Vernalization in Brassicaceae. In: Brassica Germplasm - Characterization, Breeding and Utilization. *Brassicaceae Research* 14(4):25-37.
- Amudha, J., Balasubramani, G. 2011. Recent molecular advances to combat abiotic stress tolerance in crop plants. *Biotechnology and Molecular Biology Reviews* 6(2):31–58.
- Aroca, R. 2013. Plant responses to drought stress: From morphological to molecular features. Berlin Heidelberg: Springer-Verlag.
- Ashraf, M. 2010. Inducing drought tolerance in plants: Recent advances. *Biotechnology Advances* 28(1):169-183.
- Aslam, M., Waseem, M., Xu, W., Ying, L., Zhang, J., Yuan, W. 2022. Global Identification of White Lupin lncRNAs Reveals Their Role in Cluster Roots under Phosphorus Deficiency. *International Journal of Molecular Sciences* 23(16):9012.
- Ayupe, A., Tahira, A., Camargo, L. 2015. Global analysis of biogenesis, stability and sub-cellular localization of lncRNAs mapping to intragenic regions of the human genome. *RNA Biology* 12:877–892.
- Azofeifa, J., Allen, M., Hendrix, J., Read, T., Rubin, J., Dowell, R. 2018. Enhancer RNA profiling predicts transcription factor activity. *Genome Research* 28(3):334–344.
- Bailey, T. L., Johnson, J., Grant, C. E., Noble, W. S. 2015. The MEME Suite. *Nucleic Acids Research* 43(W1):W39-W49.
- Bateman, A., Coin, L., Durbin, R., Finn, R., Hollich, V., Griffiths-Jones, S., Khanna, A. 2004. The Pfam protein families database. *Nucleic Acids Research* 32(1):D138-D141.
- Bela, K., Horvath, E., Galle, A., Szabados, L., Tari, I., Csiszar, J. 2015. Plant glutathione peroxidases: Emerging role of the antioxidant enzymes in plant development and stress responses. *Journal of Plant Physiology* 176:192–201.
- Bugnon, L., Edera, A., Prochetto, S., Gerard, M., Raad, J., Fenoy, E., Rubiolo, M., Chorostecki, U. 2022. Secondary structure prediction of long noncoding RNA: Review

and experimental comparison of existing approaches. *Briefings in Bioinformatics* 23(4): 45-63.

Butcher, S., Pyle, A. 2011. The molecular interactions that stabilize RNA tertiary structure: RNA motifs, patterns, and networks. *Accounts of Chemical Research* 44(1):1302–1311.

Chekanova, J. 2015. Long non-coding RNAs and their functions in plants. *Current Opinion in Plant Biology* 27(1):207-216.

Chen, L. 2021. Functional and evolutionary characterization of flowering-related long non-coding RNAs. Tesis doctoral de la Humboldt-Universität zu Berlin.

Chen, X., Jiang X., Niu F., Sun X., Hu Z., Gao F., Zhang H, Jiang Q. 2023. Overexpression of lncRNA77580 Regulates Drought and Salinity Stress Responses in Soybean. *Plants (Basel)* 12(1):181-195.

Chen, Y., Zhou, J. 2017. LncRNAs: Macromolecules with Big Roles in Neurobiology and Neurological Diseases. *Metabolic Brain Disease* 32(1):281-291.

Clark, D., Pazdernik, N. 2013. *Molecular Biology*. Academic Cell Update Edition. Second Edition. 907pp.

Corona-Gomez, J., Garcia-Lopez, I., Stadler, P., Fernandez-Valverde, S. 2020. Splicing conservation signals in plant long noncoding RNAs. *RNA* 26(7):784-793.

Csorba, T., Questa, J., Sun, Q., Dean, C. 2014. Antisense COOLAIR mediates the coordinated switching of chromatin states at FLC during vernalization. *Proceedings of the National Academy of Sciences of the United States of America* 111(45):16160-16165.

Cui, T., Dou, Y., Tan, P., Ni, Z., Liu, T., Wang, D., Huang, Y., Cai, K., Zhao, X., Xu, D., Lin, H., Wang, D. 2022. RNALocate v2.0: an updated resource for RNA subcellular localization with increased coverage and annotation. *Nucleic Acids Research* 50(D1):D333-D339.

Delli Ponti, R., Armaos, A., Marti, S., Tartaglia, G. G. 2018. A Method for RNA Structure Prediction Shows Evidence for Structure in lncRNAs. *Frontiers in Molecular Biosciences* 5:111-117.

Ding, Y., Tang, Y., Kwok, C., Zhang, Y., Bevilacqua P., Assmann S. 2014. In vivo genome-wide profiling of RNA secondary structure reveals novel regulatory features. *Nature* 505:696–700.

Duan, Y., Zhang, W., Cheng, Y., Shi, M., Xia, X. 2021. A systematic evaluation of bioinformatics tools for identification of long noncoding RNAs. *RNA* 27(1):80-98.

Fajardo, D., Schlautman B., Steffan S., Polashock J. Vorsa N., Zalapa, J. 2014. The American cranberry mitochondrial genome reveals the presence of selenocysteine (tRNA-Sec and SECIS) insertion machinery in land plants. *Gene* 536(2):336-343.

Fallis, A. 2011. Long Non-Coding RNAs. Ed. Ugarkovic D. *Journal of Chemical Information and Modeling* 13(1):34-53.

Farooq, M., Wahid, A., Kobayashi, N., Fujita, D., Basra, S. 2009. Plant Drought Stress: Effects, Mechanisms and Management. *Sustainable Agriculture* 13(1):153-188.

- Fichman, Y., Koncz, Z., Reznik, N., Miller, G. 2018. SELENOPROTEIN O Is a Chloroplast Protein Involved in ROS Scavenging and Its Absence Increases Dehydration Tolerance in *Arabidopsis thaliana*. *Plant Science* 270:278–291.
- Frias, J., Gulewicz, P., Martínez-Villaluenga, C., Pilarski, R., Blazquez, E., Jiménez, B., Gulewicz, K. y Vidal-Valverde, C. 2009. Influence of germination with different selenium solutions on nutritional value and cytotoxicity of lupin seeds. *Journal of Agricultural and Food Chemistry* 57(4):1319-1325.
- Ganser, L., Kelly, M., Herschlag, D. Al-Hashimi, H. 2019. The roles of structural dynamics in the cellular functions of RNAs. *Nature Reviews Molecular Cell Biology* 20:474-489.
- Gao, R., Liu, P., Irwanto, N., Loh, R. y Wong, S. 2016. Upregulation of LINC-AP2 is negatively correlated with AP2 gene expression with Turnip crinkle virus infection in *Arabidopsis thaliana*. *Plant Cell Reports* 35(11):2257-2267.
- Goff, L., Rinn, J. 2015. Linking RNA biology to lncRNAs. *Genome Research* 25:1456-1465.
- Guo, J., Fang, S., Wu, Y., Zhang, J., Chen, Y. 2019. CNIT: a fast and accurate web tool for identifying protein-coding and long non-coding transcripts based on intrinsic sequence composition. *Nucleic Acids Research* 47(W1):W516-W522.
- Han, S., Liang, Y., Ma, Q. 2018. LncFinder: an integrated platform for long non-coding RNA identification utilizing sequence intrinsic composition structural information and physicochemical property. *Briefings in Bioinformatics* 11(7):43-54.
- Hane, J., Ming, Y., Kamphuis, L., Nelson, M. 2017. A comprehensive draft genome sequence for lupin (*Lupinus angustifolius*) an emerging health food: insights into plant-microbe interactions and legume evolution *Plant Biotechnology Journal* 15(3):318-330.
- Hawkes, E., Hennelly, S., Novikova, I., Irwin, J., Dean, C., Sanbonmatsu, K. 2016. COOLAIR Antisense RNAs Form Evolutionarily Conserved Elaborate Secondary Structures *Cell Reports* 16(12):3087-3096.
- Herschlag, D., Bonilla, S., Bisaria N. 2018. The story of RNA folding as told in epochs. *Cold Spring Harbor Perspectives in Biology* 10:a032433.
- Hezroni, H., Koppstein, D., Schwartz, M., Avrutin, A., Bartel, D., Ulitsky, I. 2015. Principles of long noncoding RNA evolution derived from direct comparison of transcriptomes in 17 species. *Cell Reports* 11(7):1110-1122.
- Hidalgo, J. 2021. Transcriptoma de *Lupinus mutabilis* sometido a estrés hídrico. Tesis para optar el grado de Magister Scientiae en Mejoramiento Genético de Plantas. Escuela de Postgrado de la Universidad Nacional Agraria La Molina. 302pp.
- Hu, J., Wang, Z., Zhang, L., Peng, J., Huang, T., Yang, X., Jeong, B. R., Yang, Q. 2022. Seleno-Amino Acids in Vegetables: A Review of Their Forms and Metabolism. *Frontiers in Plant Science* 13(1):804368.
- Huang, L., Zhang, H., Deng, D., Zhao, K., Liu, K., Hendrix, D., Mathews, H. 2019. LinearFold: linear-time approximate RNA folding by 5'-to-3' dynamic programming and beam search. *Bioinformatics* 35(14):i295-i304.

- Huang, L. 2018. Systematic identification of long non-coding RNAs during pollen development and fertilization in *Brassica rapa*. *Plant Journal* 1: 203–222.
- Huang, J., Li, K., Gribskov, M. 2016. Accurate Classification of RNA Structures Using Topological Fingerprints. *PLoS ONE* 11: e0164726.
- Jackson, R., Hellen, C., Pestova, T. 2012. Termination and post-termination events in eukaryotic translation. *Advances in Protein Chemistry and Structural Biology* 86: 45-93.
- Jampala, P., Garhewal, A., Lodha, M. 2021. Functions of long non-coding RNA in *Arabidopsis thaliana*. *Plant Signaling and Behavior* 16(9): 1925440.
- Jarroux, J., Morillon, A., Pinskaya, M. 2017. Long Non-Coding RNA Biology. En: Rao M.R.S., ed., *Advances in Experimental Medicine and Biology*. 402 pp.
- Jiao, F., Pahwa, K., Manning, M., Dochy, N., Geuten, K. 2019. Cold-Induced Antisense Transcription of FLOWERING LOCUS C in Distant Grasses. *Frontiers in Plant Science* 10: 72.
- Jingjing, J., Peng, L., Yalong, X., Zefeng, L., Shizhou Y. 2021. PLncDB V2.0: A Comprehensive Encyclopedia of Plant Long Noncoding RNAs. *Nucleic Acids Research* 49(1): D1489-D1495.
- Khalil, A., Coller, J. 2013. Molecular Biology of Long Non-Coding RNAs. En *Molecular Biology of Long Non-Coding RNAs*.
- Khatun, M., Sarkar, S., Era, F., Islam, A., Anwar, M., Fahad, S., Datta, R., Islam, A. 2021. Drought Stress in Grain Legumes: Effects, Tolerance Mechanisms, and Management. *Agronomy* 11(12):2374.
- Khemka, N., Singh, V., Garg, R., et al. 2016. Genome-wide analysis of long intergenic non-coding RNAs in Chickpea and their potential role in flower development. *Scientific Reports* 6:33297.
- Khoei, M. A., Karimi, M., Karamian, R., Amini, S. y Soorni, A. 2021. Identification of the Complex Interplay Between Nematode-Related lncRNAs and Their Target Genes in *Glycine max*. *Frontiers in Plant Science* 12:779597.
- Kiegle, E., Garden, A., Lacchini, E. y Kater, M. 2018. A Genomic View of Alternative Splicing of Long Non-coding RNAs during Rice Seed Development Reveals Extensive Splicing and lncRNA Gene Families. *Frontiers in Plant Science* 9(February):1-12.
- Kim, D. y Sung, S. 2017a. Vernalization-triggered intragenic chromatin loop formation by long noncoding RNAs. *Developmental Cell* 40:302–312.
- Kim, D. H., Xi, Y., Sung, S. 2017b. Modular function of long noncoding RNA, COLDAIR, in the vernalization response. *PLoS Genetics* 13(7):e1006939.
- Kim, W., Miguel-Rojas, C., Wang, J., Townsend, J. Trail, F. 2018. Developmental Dynamics of Long Noncoding RNA Expression during Sexual Fruiting Body Formation in *Fusarium graminearum*. *Bio* 9(4):e01292-18.
- Kim, D., Doyle, M., Sung, S., Amasino, R. 2009. Vernalization: Winter and the Timing of Flowering in Plants *Annual Review of Cell and Developmental Biology* 25:277–299.

- Kong L., Zhang Y., Ye Z., Liu X., Zhao S., Wei L., Gao G. 2007. CPC: Assess the protein-coding potential of transcripts using sequence features and support vector machine *Nucleic Acids Research* 35(1):W345-W349.
- Kumar, M., Kumar, P., Kumar, N., Bajpai, A., Siddique, K. 2021. Metabolomics and Molecular Approaches Reveal Drought Stress Tolerance in Plants. *International Journal of Molecular Sciences* 22:9108.
- Lafitte, H., Yongsheng, G., Yan, S., Li, Z. 2007. Whole plant responses, key processes, and adaptation to drought stress: the case of rice. *Journal of Experimental Botany* 58(2):169-175.
- Lai, X., Gupta, S. K., Vera, J. (Eds.). 2019. *Computational Biology of Non-Coding RNA. In Methods in Molecular Biology.* Springer New York. 402pp.
- Letunic, I., Copley, R., Pils, B., Pinkert, S., Schultz, J., Bork, P. 2006. SMART 5: domains in the context of genomes and networks. *Nucleic Acids Research* 34(4):D257-60.
- Li, J., Zhang, X., Liu, C. 2020. The computational approaches of lncRNA identification based on coding potential: Status quo and challenges. *Computational and Structural Biotechnology Journal* 18:3666-3677.
- Li, L., Eichten, S., Shimizu, R. 2014. Genome-wide discovery and characterization of maize long non-coding RNAs. *Genome Biology* 15(2):R40.
- Li, R., Zhu, H., Luo, Y. 2016. Understanding the Functions of Long Non-Coding RNAs through Their Higher-Order Structures. *International Journal of Molecular Sciences* 17(5):702.
- Li, S., Yu, X., Lei, N., Cheng, Z., Zhao, P., He, Y., Wang, W., Peng, M. 2017. Genome-wide identification and functional prediction of cold and/or drought-responsive lncRNAs in cassava. *Scientific Reports* 7:45981.
- Li, A., Zhang, J., Zhou, Z. 2014. PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme *BMC Bioinformatics* 15(1):311.
- Li, Y., Sun, H., Feng S. 2021. Capsule-LPI: a lncRNA–protein interaction predicting tool based on a capsule network *BMC Bioinformatics* 22(1):246-255.
- Lin, Y., Pan, X., Shen, H. 2021. IncLocator 2.0: a cell-line-specific subcellular localization predictor for long non-coding RNAs with interpretable deep learning. *Bioinformatics* 2(14):127-135.
- Linares, S. 2021. Identificación de genes de referencia en *Lupinus mutabilis* Sweet para estudios cuantitativos de expresión génica. Tesis para optar el Título Profesional de Biólogo. Facultad de Ciencias de la Universidad Nacional Agraria La Molina. 89 pp.
- Liu, F., Marquardt, S., Lister, C., Swiezewski, S., Dean, C. 2010. Targeted 3' processing of antisense transcripts triggers Arabidopsis FLC chromatin silencing. *Science* 327(5961):94-97.
- Liu, J., Gough, J., Rost, B. 2006. Distinguishing protein-coding from non-coding RNAs through support vector machines. *PLoS Genetics* 2(4):e29.
- Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., Arenas-Huertero, C., Chua, N. H. 2012. Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in Arabidopsis. *The Plant Cell* 24:4333-4345.

- Liu, Z. W., Zhao, N., Su, Y. N. 2020. Exogenously overexpressed intronic long noncoding RNAs activate host gene expression by affecting histone modification in Arabidopsis. *Scientific Reports* 10:3094.
- Lorenz, R., Bernhart, S. H., Höner zu Siederdisen, C., et al. 2011. ViennaRNA Package 2.0. *Algorithms for Molecular Biology* 6(1):26.
- Ma, H., Yang, T., Li, Y., Zhang, J., Wu, T., Song, T., Yao, Y., Tian, J. 2021. The long noncoding RNA MdLNC499 bridges MdWRKY1 and MdERF109 function to regulate early-stage light-induced anthocyanin accumulation in apple fruit. *Plant Cell* 33(10):3309-3330.
- Madera, M., Vogel, C., Kummerfeld, S. K., Chothia, C., Gough, J. 2004. The SUPERFAMILY database in 2004: additions and improvements. *Nucleic Acids Research* 32(4):235-239.
- Mailler, E., Paillart, J. C., Marquet, R., Smyth, R. P., Vivet-Boudou, V. 2019. The evolution of RNA structural probing methods: from gels to next-generation sequencing. *Wiley Interdisciplinary Reviews: RNA* 10(1):e1518.
- Mathews, D. H., Sabina, J., Zuker, M., Turner, D. H. 1999. Expanded sequence dependence of thermodynamic parameters improves prediction of RNA secondary structure. *Journal of Molecular Biology* 288(1):911-940.
- Mattick, J. 2004. RNA regulation: a new genetics? *Nature Reviews Genetics* 5(1):316-323.
- Mignone, F., Grillo, G., Licciulli, F., Iacono, M., Liuni, S., Kersey, P. J., Duarte, J., Saccone, C., Pesole, G. 2005. UTRdb and UTRsite: a collection of sequences and regulatory motifs of the untranslated regions of eukaryotic mRNAs. *Nucleic Acids Research* 33(4): 141-146.
- Mita, Y., Uchida, R., Yasuhara, S., Kishi, K., Hoshi, T., Matsuo, Y., Yokooji, T., Shirakawa, Y., Toyama, T., Urano, Y., Inada, T., Noguchi, N., Saito, Y. 2021. Identification of a novel endogenous long non-coding RNA that inhibits selenoprotein P translation. *Nucleic Acids Research*, 49(12), 6893-6907.
- Mix, H., Lobanov, A. V., Gladyshev, V. N. 2007. SECIS elements in the coding regions of selenoprotein transcripts are functional in higher eukaryotes. *Nucleic Acids Research*, 35(2), 414-423.
- Moh, N., Zhang, P., Chen, Y., Chen, M. 2021. Computational Identification of miRNAs and Temperature-Responsive lncRNAs From Mango (*Mangifera indica* L.). *Frontiers in Genetics* 12(1):607248.
- Mohammadin, S., Edger, P., Pires, J., Schranz, M. 2015. Positionally-conserved but sequence-diverged: identification of long non-coding RNAs in the Brassicaceae and Cleomaceae. *BMC Plant Biology* 15(1):217.
- Morris, K. (ed). 2016. Long Non-coding RNAs in Human Disease. Ed. Springer International Publishing. 262 pp.
- Mustoe, A., Brooks, C., Al-Hashimi, H. 2014. Hierarchy of RNA functional dynamics. *Annual Review of Biochemistry* 83(1):441-466.

- Noh, J., Kim, K., McClusky, W., Abdelmohsen, K., Gorospe M. 2018. Cytoplasmic functions of long noncoding RNAs. *Wiley Interdisciplinary Reviews: RNA* 9(3):e1471.
- O'Leary, C., Van Tompkins S., Rouse W., Nam G., Moss W. 2022. Thermodynamic and structural characterization of an EBV infected B-cell lymphoma transcriptome. *NAR Genomics and Bioinformatics* 4(4):lqac082.
- Pang, K., Stephen, S., Engström, P. Tajul-Arifin K., Chen W., Wahlestedt C., Lenhard B., Hayashizaki Y., Mattick J. 2005. RNADB—a comprehensive mammalian noncoding RNA database. *Nucleic Acids Research* 33(4):D125–130.
- Pinkney, H., Wright, B., Diermeier, S. 2020. The lncRNA Toolkit: Databases and In Silico Tools for lncRNA Analysis. *Noncoding RNA* 6(4):49.
- Ponti, R., Armaos, A., Vandelli, A., Tartaglia G. 2020. CROSSalive: a web server for predicting the in vivo structure of RNA molecules. *Bioinformatics* 36:940–941.
- Qin, D. 2019. Next-generation sequencing and its clinical application. *Cancer Biology Medicine* (16):4–10.
- Qin, T., Zhao, H., Cui, P., Albeshier, N., Xiong, L. 2017. A nucleus-localized long non-coding RNA enhances drought and salt stress tolerance. *Plant Physiology* (175):1321–1336.
- Rajewsky, N., Jurga, S., Barciszewski, J. 2018. *Systems Biology* (Eds. N. Rajewsky, S. Jurga, J. Barciszewski). Cham, Springer International Publishing, (RNA Technologies), 404 pp.
- Ramasamy, S., Ganesh Thiruvengadam Nandagopal, J., Balasubramanian, M., Girija, S. 2022. Effect of Abscisic acid and Selenium foliar sprays on drought mitigation in tomato (*Solanum lycopersicum* L.). In *Materials Today: Proceedings*, Elsevier BV 48(1): 191–195.
- Ramos, C. 2021. Transcriptoma del brote floral de *Lupinus mutabilis* sometido a estrés hídrico (Tesis para optar el grado de Magister Scientae en Mejoramiento Genético de Plantas). Escuela de Postgrado de la Universidad Nacional Agraria La Molina. 95 pp.
- Rao, M. 2017. *Long Non Coding RNA Biology* (1st ed., Ed. MRS Rao). Singapore, Springer Singapore. *Advances in Experimental Medicine and Biology* 1008(1):323.
- Repo-Carrasco-Valencia, R.; Basilio-Atencio, J.; Luna-Mercado, G.I.; Pilco-Quesada, S.; Vidaurre-Ruiz, J. *Andean Ancient Grains: Nutritional Value and Novel Uses*. *Biol. Life Sci. Forum* 2021, 8, 15.
- Rinn, J., Chang, H. 2012. Genome regulation by long noncoding RNAs. *Annual Review of Biochemistry* 81(1):145-166.
- Ross, C.J., Ulitsky, I. 2022. Discovering functional motifs in long noncoding RNAs. *Wiley Interdisciplinary Reviews. RNA* 13(1):e1708.
- Sanbonmatsu, K. 2022. Getting to the bottom of lncRNA mechanism: structure-function relationships. *Mammalian Genome* 33(2):343-353.
- Sang, S., Chen, W., Zhang, D., Zhang, X., Yang, W., Liu, C. 2021. Data integration and evolutionary analysis of long non-coding RNAs in 25 flowering plants. *BMC Genomics* 22(Suppl 3):739.

- Shaath, H., Vishnubalaji, R., Elango, R., Kardousha, A., Islam, Z. 2022. Long non-coding RNA and RNA-binding protein interactions in cancer: Experimental and machine learning approaches. *Seminars in Cancer Biology* 86(3):325-345.
- Shin, J., Wang, H., Lee, J., Dinwiddie, B., Belostotsky, D., Chekanova, J. 2013. The role of the Arabidopsis Exosome in siRNA-independent silencing of heterochromatic loci. *PLoS Genetics* 9(3):e1003411.
- Spitale, R., Incarnato, D. 2022. Probing the dynamic RNA structurome and its functions. *Nature Reviews Genetics*. Springer Science and Business Media 43(3):178–196.
- Steffen, P., Voss, B., Rehmsmeier, M., Reeder, J., Giegerich, R. 2006. RNASHAPES: an integrated RNA analysis package based on abstract shapes. *Bioinformatics* 22(4):500-503.
- Steinbrenner, H., Sies, H. 2013. Selenium Homeostasis and Antioxidant Selenoproteins in Brain: Implications for Disorders in the Central Nervous System. *Archives of Biochemistry and Biophysics* 536(1):152–157.
- Sun, X., Wong, D. 2016. Long non-coding RNA-mediated regulation of glucose homeostasis and diabetes. *American Journal of Cardiovascular Disease* 6(1):17–25.
- Svoboda, P., Di Cara, A. 2006. Hairpin RNA: a secondary structure of primary importance. *Cellular and Molecular Life Sciences* 63(8):901-908.
- Taft, R., Pheasant, M., Mattick, J. 2007. The relationship between non-protein-coding DNA and eukaryotic complexity. *BioEssays* 29(1):288–299.
- Trotta, E. 2014. On the Normalization of the Minimum Free Energy of RNAs by Sequence Length. *PLoS ONE* 9(11):e113380.
- Ugarkovic, D. (ed.). 2011. Long Non-Coding RNAs. Berlin, Heidelberg, Springer Berlin Heidelberg. Progress in Molecular and Subcellular Biology. 205 pp.
- Urbischek, M., Rannikmae, H., Foets, T., Ravn, K., Hyvönen, M., de la Roche, M. 2019. Organoid culture media formulated with growth factors of defined cellular activity. *Scientific Reports* 9(1):6193.
- Vera-Vega, M.; Jimenez-Davalos, J.; Zolla, G. The Micronutrient Content in Underutilized Crops: The Lupinus Mutabilis Sweet Case. *Sci. Rep.* 2022, 12, 15162.
- Rinn, J. 2014. lncRNAs: Linking RNA to Chromatin (online). Cold Spring Harbor Perspectives in Biology. Cold Spring Harbor Laboratory 23(4): a018614–a018614.
- Vicente, J., Mendiondo, G., Movahedi, M., Peirats-Llobet, M., Juan, Y. 2017. The Cys-Arg/N-End Rule Pathway Is a General Sensor of Abiotic Stress in Flowering Plants. *Current Biology* 27(20):3183–3190.
- Wang, H., Chekanova, J. 2017. Long Noncoding RNAs in Plants. *Advances in Experimental Medicine and Biology* 1008(1):133-154.
- Wang, L., Wang, J., Chen, H., Hu, B. 2022. Genome-wide identification, characterization, and functional analysis of lncRNAs in *Hevea brasiliensis*. *Frontiers in Plant Science* 13(1):1012576.
- Wang, M., Yuan, D., Tu, L. 2015. Long noncoding RNAs and their proposed functions in fibre development of cotton (*Gossypium* spp.). *New Phytologist* 207(1):1181–1197.

- Wang, Y., Nambeesan, S. 2022. Full-length fruit transcriptomes of southern highbush (*Vaccinium* sp.) and rabbiteye (*V. virgatum* Ait.) blueberry. *BMC Genomics* 23(1):733.
- Wang, J., Lin, J., Kan, J., Wang, H., Li, X., Yang, Q. 2018. Genome-Wide Identification and Functional Prediction of Novel Drought-Responsive lncRNAs in *Pyrus betulifolia*. *Genes* 9(6):311.
- Waseem, M., Liu, Y., Xia, R. 2021. Long Non-Coding RNAs, the Dark Matter: An Emerging Regulatory Component in Plants. *International Journal of Molecular Sciences* 22(1):86.
- Watanabe, S., Sato, M., Sawada, Y. 2018. Arabidopsis molybdenum cofactor sulfuryase ABA3 contributes to anthocyanin accumulation and oxidative stress tolerance in ABA-dependent and independent ways. *Scientific Reports* 8(1):16592.
- Wei, Z., Qing-jie, S., Chu-fu, Z., Yong-ze, Y., Ji, Z., Bin-bin, L. 2004. Effect of Salt Stress on Ammonium Assimilation Enzymes of the Roots of Rice (*Oryza sativa*) Cultivars Differing in Salinity Resistance. *Acta Botanica Sinica* 46(8):921-926.
- Wekesa, J., Luan, Y., Chen, M., Meng, J. 2019. A Hybrid Prediction Method for Plant lncRNA-Protein Interaction. *Cells* 8(6):521.
- Wierzbicki, A., Blevins, T., Swiezewski, S. 2021. Long Noncoding RNAs in Plants. *Annual Review of Plant Biology* 72:245-271.
- Xing, J., Liu, H., Jiang, W., Wang, L. 2021. LncRNA-Encoded Peptide: Functions and Predicting Methods. *Frontiers in Oncology* 10(1):622294.
- Xiong, L., Ishitani, M., Lee, H., Zhu, J. K. 2001. The Arabidopsis LOS5/ABA3 locus encodes a molybdenum cofactor sulfuryase and modulates cold stress- and osmotic stress-responsive gene expression. *The Plant Cell* 13(9):2063-2083.
- Xuan, H., Zhang, L., Liu, X. 2015. PLNlncRbase: A resource for experimentally identified lncRNAs in plants. *Gene* 573(1):328-332.
- Yang, Z., Song, Y., Li, Y., Mao, Y., Du, G., Tan, B., Zhang, H. 2022. Integrative analyses of prognosis, tumor immunity, and ceRNA network of the ferroptosis-associated gene FANCD2 in hepatocellular carcinoma. *Frontiers in Genetics* 13(1):955225.
- Yang, J., Zhang, J. 2015. Human long noncoding RNAs are substantially less folded than messenger RNAs. *Molecular Biology and Evolution* 32(4):970-977.
- Ying, L., Chen, H., Cai, W. 2014. BnNAC485 is involved in abiotic stress responses and flowering time in *Brassica napus*. *Plant Physiology and Biochemistry* 79(1):77-87.
- Yu, T., Zhu, H. 2019. Long Non-Coding RNAs: Rising Regulators of Plant Reproductive Development. *Agronomy* 9(2):53-57.
- Yu, Y., Zhang, Y., Chen, X., Chen, Y. 2019. Plant Noncoding RNAs: Hidden Players in Development and Stress Responses. *Annual Review of Cell and Developmental Biology* 35(1):407-431.
- Kang, Y.J., Yang, D.C., Kong, L., Hou, M., Meng, Y.Q., Wei, L., Gao, G. 2017. CPC2: A fast and accurate coding potential calculator based on sequence intrinsic features. *Nucleic Acids Research* 45(W1):W12-W16.

- Zampetaki, A., Albrecht, A., Steinhofel, K. 2018. Long non-coding RNA structure and function: Is there a link? *Frontiers in Physiology* 9(1):1-8.
- Zeng, M., Wu, Y., Lu, C., Zhang, F., Wu, F.X., Li, M. 2022. DeepLncLoc: A deep learning framework for long non-coding RNA subcellular localization prediction based on subsequence embedding. *Briefings in Bioinformatics* 23(1):360.
- Zhang, H., Zhang, L., Mathews, D., Huang, L. 2020. LinearPartition: Linear-time approximation of RNA folding partition function and base-pairing probabilities. *Bioinformatics* 36(Suppl_1):i258-i267.
- Zhang, Y., Liao, J, Li, Z. 2014. Genome-wide screening and functional analysis identify a large number of long noncoding RNAs involved in the sexual reproduction of rice. *Genome Biology* 15(1):512.
- Zhang, X., Wang, W., Zhu, W., Dong, J., Cheng, Y., Yin, Z., Shen, F. 2019. Mechanisms and Functions of Long Non-Coding RNAs at Multiple Regulatory Levels. *International Journal of Molecular Sciences* 20(22):5573.
- Zhang, Z., Sun, Z., Yang, Y. 2022. Towards a better prediction of subcellular location of long non-coding RNA. *Frontiers in Computational Science* 16(1):165-173.
- Zhao, X., Li, J., Lian, B., Gu, H., Li, Y., Qi, Y. 2018. Global identification of Arabidopsis lncRNAs reveals the regulation of MAF4 by a natural antisense RNA. *Nature Communications* 29(3):50-56.
- Zhou, L. Liu, Z., Liu, Y. 2016. A novel gene OsAHL1 improves both drought avoidance and drought tolerance in rice. *Scientific Reports* 6(1):1-15.
- Zhu, B., Yang, Y., Li, R. 2015. RNA sequencing and functional analysis implicate the regulatory role of long non-coding RNAs in tomato fruit ripening. *Journal of Experimental Botany* 66(1):4483-4495.

VII. ANEXOS

Anexo 1. Resumen de la secuenciación HiSeq de Macrogen.

Experimento	Tratamiento	ID	Reads totales (pb)	Reads de alta calidad (pb)	GC (%)	AT (%)	Q20 (%)	Q30 (%)
Floración	Estadio E4 de floración	C14	5594097908	55387108	43.46	56.54	96.81	94.84
		C15	5128560628	50777828	43.52	56.48	96.85	94.92
		C19	4843016458	47950658	43.44	56.56	96.93	95.02
	Estadio E6 de floración	C21	5135089470	50842470	43.46	56.54	96.75	94.75
		C23	5196980250	51455250	43.38	56.63	96.83	94.89
		C24	5211518796	51599196	43.30	56.70	96.81	94.91
Sequía	Control	M11	5799010546	57415946	43.82	56.18	96.84	94.89
		M17	5162300082	51111882	43.86	56.14	96.98	95.08
		M1A	4924410944	48756544	44.03	55.97	96.77	94.78
	Sequía	M23	4450645396	44065796	43.20	56.80	96.70	94.68
		M24	4397283864	43537464	43.11	56.90	96.86	94.90
		M29	4603500008	45579208	43.27	56.73	97.52	95.79

Anexo 2. Resultados de los algoritmos de detección de los *ARNlncs*.

ID	Longitud	CPC	CNIT	CPC2	lncFinder
TRINITY_DN38174_c0_g1_i1	520	-0,817097	-0,324436574	1,41129E-05	0,00272335
TRINITY_DN46326_c0_g1_i1	312	-0,877516	-0,426742171	0,0778179	0,0131862
TRINITY_DN25851_c0_g4_i1	387	-0,95815	-0,541767366	9,42897E-06	0,01390129
TRINITY_DN25006_c0_g2_i1	318	-0,837614	-0,529740629	0,113736	0,01547576
TRINITY_DN22576_c0_g1_i2	332	-0,90138		0,0785498	0,02329028
TRINITY_DN37101_c0_g1_i1	324	-0,99186	-0,471122203	0,0554415	0,0251756
TRINITY_DN8984_c0_g1_i1	355	-1,16165	-0,544252301	0,0223897	0,02529995
TRINITY_DN19570_c0_g1_i2	987	-1,16022	-0,495126049	0,289034	0,02684732
TRINITY_DN22693_c2_g1_i1	325	-0,573748	-0,26040625	0,0174088	0,02740252
TRINITY_DN25030_c1_g2_i1	4206	-1,19927		0,231128	0,02839543
TRINITY_DN52122_c0_g1_i1	313	-0,266539		0,181622	0,02881803
TRINITY_DN25030_c1_g2_i5	4108	-1,19833		0,231128	0,02917246
TRINITY_DN19638_c2_g1_i2	1068	-0,508445	-0,360983175		0,03200482
TRINITY_DN16235_c0_g1_i1	306	-1,10174		0,00843892	0,03222957
TRINITY_DN20109_c0_g1_i5	826	-0,12335		0,0137685	0,03252141
TRINITY_DN21355_c2_g3_i3	1318	-0,489358	-0,386205479		0,03370287
TRINITY_DN16914_c0_g1_i2	1127	-0,244601		0,443164	0,03395209
TRINITY_DN21775_c0_g3_i1	352	-0,903296	-0,448047673	0,0588084	0,0340674
TRINITY_DN25541_c1_g1_i7	380	-0,685103	-0,536458157	0,0388412	0,03431731
TRINITY_DN23688_c0_g5_i3	300	-0,974961	-0,44906955	0,071187	0,03570292
TRINITY_DN22693_c2_g1_i8	324	-0,473142		0,0220472	0,03823797
TRINITY_DN24537_c1_g2_i4	2361	-0,798371		0,422799	0,03852023
TRINITY_DN23684_c7_g1_i9	2137	-0,21093			0,04494627
TRINITY_DN15360_c0_g1_i1	1200	-1,0426			0,04494789
TRINITY_DN17529_c0_g1_i7	3377	-1,12587	-0,359332055	0,15317	0,04497183
TRINITY_DN22693_c2_g1_i7	325	-0,526254		0,0217504	0,04506522
TRINITY_DN50713_c0_g1_i1	409	-0,712974	-0,380177234	0,0284495	0,04558701
TRINITY_DN21682_c0_g2_i2	5507	-0,879272			0,04612798
TRINITY_DN25621_c2_g2_i1	310	-0,729537	-0,43539721	0,0840639	0,04740835
TRINITY_DN20665_c6_g2_i1	1032	-1,07402	-0,424914015	0,272767	0,04797738
TRINITY_DN23843_c0_g1_i1	2006	-0,721534	-0,384354399	0,312781	0,04815559
TRINITY_DN26367_c1_g4_i3	858	-1,03141		0,192847	0,04884591
TRINITY_DN23284_c0_g4_i4	937	-1,07821	-0,326900302		0,04895987
TRINITY_DN25011_c6_g2_i5	920	-1,00461	-0,439407922		0,04900566
TRINITY_DN20434_c0_g1_i1	750	-0,809561	-0,34465496	0,262246	0,05019427
TRINITY_DN25411_c0_g4_i1	338	-0,394777	-0,5601109	0,0892643	0,05208958
TRINITY_DN20104_c2_g3_i1	1224	-1,26372	-0,330941294		0,0525065
TRINITY_DN24123_c1_g1_i7	856	-1,14428	-0,410445075	0,390096	0,0530931
TRINITY_DN25012_c0_g4_i2	856	-0,873246			0,05320878
TRINITY_DN24470_c1_g3_i2	571	-1,04787	-0,391355895	0,161998	0,05346054
TRINITY_DN25057_c1_g1_i3	1005	-0,741337			0,05377867
TRINITY_DN24299_c2_g1_i5	941	-1,11445	-0,356658061	0,302998	0,05414455
TRINITY_DN26593_c1_g1_i8	808	-0,924693	-0,350125108	0,384867	0,05438923

TRINITY_DN25206_c0_g2_i2	1034	-0,845299	-0,55269115	0,321084	0,05460267
TRINITY_DN22116_c0_g2_i1	300	-0,870978	-0,361307564	0,049602	0,05527744
TRINITY_DN21403_c1_g2_i1	1144	-0,412026	-0,362401052		0,05588363
TRINITY_DN26051_c0_g7_i1	724	-0,398016			0,05638563
TRINITY_DN223_c0_g1_i1	366	-0,79527	-0,367239219	0,0984951	0,05741938
TRINITY_DN23737_c2_g1_i16	1031	-1,02203	-0,348226308	0,344825	0,05746842
TRINITY_DN20758_c0_g2_i3	871	-0,219829			0,05755136
TRINITY_DN19915_c0_g3_i1	698	-0,902162		0,205649	0,05772977
TRINITY_DN26574_c2_g2_i4	795	-1,05603	-0,457913207	0,0468069	0,05839152
TRINITY_DN8762_c0_g1_i2	523	-0,357834			0,05844418
TRINITY_DN21073_c5_g3_i3	835	-0,577259	-0,340671057		0,05878873
TRINITY_DN26016_c2_g1_i2	1714	-0,59857			0,05936269
TRINITY_DN2209_c0_g1_i1	464	-0,966604	-0,432576592	0,0790578	0,0593798
TRINITY_DN26279_c0_g1_i1	1374	-1,12253	-0,371062546	0,311981	0,0597662
TRINITY_DN15018_c0_g1_i3	759	-0,672651			0,05982108
TRINITY_DN6700_c0_g1_i1	825	-0,853635	-0,473086489		0,05994347
TRINITY_DN24788_c3_g2_i5	427	-1,00122	-0,511338471	0,0244024	0,06007089
TRINITY_DN15067_c0_g1_i1	2024	-0,938637	-0,34997772	0,4109	0,06027308
TRINITY_DN21682_c0_g2_i6	3037	-0,818191			0,06040701
TRINITY_DN13891_c0_g1_i2	1001	-0,0213938			0,06076996
TRINITY_DN19178_c0_g1_i10	939	-0,967722	-0,367331467	0,300074	0,0613444
TRINITY_DN24171_c1_g2_i2	874	-0,9872	-0,344403052	0,429163	0,06170115
TRINITY_DN25479_c3_g1_i29	1020	-1,10159	-0,395082331	0,196929	0,06176617
TRINITY_DN15064_c0_g1_i8	2891	-0,672409		0,0819176	0,06184489
TRINITY_DN14031_c0_g1_i4	1024	-0,329666	-0,315528522	0,102279	0,06199958
TRINITY_DN23684_c7_g1_i3	2804	-0,831846		0,355589	0,06232792
TRINITY_DN24106_c2_g1_i1	1165	-1,19715	-0,397988336	0,169621	0,06330991
TRINITY_DN23700_c0_g1_i11	2172	-1,23629		0,126948	0,06333703
TRINITY_DN22115_c1_g1_i16	1616	-0,691629		0,315183	0,06345533
TRINITY_DN18189_c0_g2_i2	1526	-1,27231		0,424854	0,06357649
TRINITY_DN24583_c3_g2_i11	1342	-0,991011		0,258716	0,06366595
TRINITY_DN9067_c0_g1_i1	500	-0,460229	-0,303908733	0,175271	0,06371832
TRINITY_DN26593_c1_g1_i4	649	-0,917866	-0,371343803	0,428861	0,06419774
TRINITY_DN22115_c1_g1_i10	1610	-0,691149		0,315183	0,06450741
TRINITY_DN23165_c0_g1_i11	799	-0,908396	-0,471928594	0,164362	0,06462667
TRINITY_DN17100_c0_g1_i6	1401	-0,728892			0,06475504
TRINITY_DN20350_c0_g1_i2	759	-1,14683	-0,326430582		0,06532682
TRINITY_DN15620_c0_g2_i1	370	-0,860499	-0,334370729	0,0489723	0,0656808
TRINITY_DN19178_c0_g1_i4	801	-1,0264	-0,380744927	0,155342	0,06576462
TRINITY_DN21252_c0_g1_i2	1577	-1,04911	-0,330535018	0,247406	0,06599518
TRINITY_DN24583_c3_g2_i3	1357	-0,992596		0,258716	0,06615982
TRINITY_DN22115_c1_g1_i22	1641	-0,693519		0,312337	0,06624581
TRINITY_DN21394_c2_g1_i5	905	-0,964402		0,13879	0,06635195
TRINITY_DN26004_c5_g1_i4	481	-1,17806	-0,530514654	0,459042	0,06645757
TRINITY_DN24829_c1_g1_i2	1525	-1,33073	-0,439083783	0,176078	0,06647598
TRINITY_DN22115_c1_g1_i4	1651	-0,694316		0,306791	0,06650014
TRINITY_DN24829_c1_g1_i11	1524	-1,30502	-0,439083783	0,195459	0,06667707
TRINITY_DN14308_c0_g2_i2	1333	-1,38827	-0,461270695	0,0659405	0,06676006

TRINITY_DN25684_c5_g2_i5	999	-1,22826	-0,360861425	0,0919126	0,06681011
TRINITY_DN25698_c1_g3_i5	1019	-0,95672	-0,502343316	0,251251	0,06681332
TRINITY_DN20710_c2_g2_i7	1698	-0,866772		0,369105	0,06681369
TRINITY_DN26019_c1_g4_i2	1592	-1,16179	-0,359977509	0,0636172	0,06710977
TRINITY_DN20743_c0_g2_i3	2695	-0,944158		0,220858	0,0671411
TRINITY_DN14780_c0_g1_i3	339	-0,45828		0,201279	0,06751814
TRINITY_DN24104_c4_g1_i6	2257	-1,14159		0,403103	0,06763372
TRINITY_DN22643_c0_g1_i11	520	-0,333157		0,205117	0,06782484
TRINITY_DN7373_c0_g1_i2	329	-1,15645		0,0928692	0,06796446
TRINITY_DN22693_c1_g1_i4	517	-0,0171533		0,0210348	0,06817524
TRINITY_DN23696_c0_g1_i3	3348	-1,17992	-0,40604857	0,107496	0,06843115
TRINITY_DN24231_c5_g1_i5	462	-0,84833	-0,397706391	0,0573044	0,06873708
TRINITY_DN24925_c0_g2_i9	1744	-0,909892		0,2275	0,06879552
TRINITY_DN26593_c1_g1_i6	817	-0,891602		0,137542	0,068822
TRINITY_DN25722_c0_g2_i2	758	-0,516888	-0,360448183	0,210814	0,06914424
TRINITY_DN24372_c2_g3_i2	987	-1,2674	-0,383209735	0,0341016	0,06938268
TRINITY_DN23252_c8_g1_i2	2712	-1,15238	-0,348129984	0,325106	0,06969373
TRINITY_DN21665_c1_g4_i1	454	-0,21548	-0,4420782	0,150779	0,07104666
TRINITY_DN22840_c1_g4_i1	615	-0,709566		0,0510914	0,07115136
TRINITY_DN24881_c2_g1_i19	1506	-1,05002	-0,419646887	0,236161	0,07131341
TRINITY_DN43828_c0_g1_i1	488	-0,53693		0,180491	0,071615
TRINITY_DN23866_c5_g3_i14	2880	-1,11015	-0,366642915	0,251705	0,07192738
TRINITY_DN19100_c2_g3_i5	1426	-0,865732		0,179653	0,07211134
TRINITY_DN11453_c0_g1_i1	444	-0,724915		0,0085948	0,0724504
TRINITY_DN33734_c0_g1_i1	687	-1,00855		0,0604796	0,07275117
TRINITY_DN25946_c0_g1_i15	873	-0,546149			0,07315632
TRINITY_DN19637_c4_g2_i1	531	-1,39967	-0,430005808	0,00409056	0,07317727
TRINITY_DN24853_c0_g1_i35	2158	-0,398946	-0,431829053	0,251201	0,07331502
TRINITY_DN20684_c5_g2_i2	1080	-0,525362			0,073375
TRINITY_DN20182_c0_g2_i5	1535	-0,901028			0,07343307
TRINITY_DN22115_c1_g1_i14	1649	-0,694139		0,284783	0,07345691
TRINITY_DN24829_c1_g1_i5	2037	-1,31234	-0,439083783	0,267501	0,07348608
TRINITY_DN21812_c0_g1_i3	697	-1,31164	-0,318349827		0,07378
TRINITY_DN25371_c1_g1_i3	1594	-0,718572	-0,423080961		0,07391218
TRINITY_DN22973_c0_g1_i11	329	-0,540285		0,0919115	0,07404765
TRINITY_DN22558_c3_g2_i8	2085	-0,814687		0,146553	0,07455191
TRINITY_DN15062_c0_g1_i5	1347	-0,906511		0,453664	0,07466518
TRINITY_DN23566_c3_g6_i2	1403	-1,31474	-0,363524205	0,00832972	0,07501373
TRINITY_DN21597_c2_g2_i5	557	-0,896834	-0,31000155	0,127755	0,07620781
TRINITY_DN24829_c1_g1_i9	2035	-1,31234	-0,439083783	0,158235	0,07636091
TRINITY_DN22828_c0_g3_i1	1085	-1,01504	-0,450754505	0,0467226	0,0765195
TRINITY_DN36500_c0_g1_i1	384	-1,34978	-0,438632624	0,00912524	0,07661065
TRINITY_DN21759_c1_g4_i2	599	-0,838002	-0,332144254	0,215356	0,07703982
TRINITY_DN23747_c0_g2_i2	404	-1,30158		0,0133874	0,07707041
TRINITY_DN24382_c1_g3_i1	1773	-0,866716	-0,378349468	0,112166	0,07723593
TRINITY_DN25644_c7_g1_i1	1893	-0,627131	-0,368678933	0,395922	0,07747011
TRINITY_DN24382_c1_g3_i8	1869	-0,875144	-0,378349468	0,112166	0,07774507
TRINITY_DN22661_c2_g2_i9	1846	-1,29454	-0,426365285	0,0616522	0,07777078

TRINITY_DN19977_c2_g2_i1	3191	-0,481018			0,07777233
TRINITY_DN49223_c0_g1_i1	333	-0,204163		0,15197	0,07778018
TRINITY_DN21177_c1_g1_i4	1480	-0,998165		0,160718	0,07780569
TRINITY_DN24973_c0_g3_i11	466	-0,949563	-0,411509788	0,0631958	0,07802035
TRINITY_DN23355_c3_g1_i3	995	-0,625657			0,07809799
TRINITY_DN22073_c0_g2_i10	410	-1,03887	-0,344658848	0,077045	0,07845426
TRINITY_DN11424_c0_g1_i1	440	-1,22899	-0,418859861	0,0411425	0,07848557
TRINITY_DN16288_c0_g1_i1	546	-1,31507	-0,424925512	0,00556644	0,07859578
TRINITY_DN26219_c0_g1_i5	1187	-1,24424	-0,373144159	0,0655672	0,07873428
TRINITY_DN17038_c0_g1_i12	1745	-1,14556			0,07883014
TRINITY_DN24382_c1_g3_i11	2293	-0,904905	-0,378349468	0,123586	0,07886796
TRINITY_DN19670_c0_g2_i6	2572	-1,38438	-0,435183826	0,124467	0,0791998
TRINITY_DN24741_c2_g1_i6	1475	-1,17462	-0,398853612	0,354648	0,07978506
TRINITY_DN25371_c1_g1_i2	1528	-0,710599	-0,423080961		0,07996962
TRINITY_DN21893_c0_g1_i36	2046	-0,783246		0,0942498	0,08001713
TRINITY_DN22693_c2_g1_i9	360	-0,47672		0,0223224	0,0802722
TRINITY_DN24724_c0_g3_i1	322	-0,941849	-0,351278637	0,026147	0,08045511
TRINITY_DN24382_c1_g3_i3	2245	-0,902079	-0,378349468	0,123586	0,08089015
TRINITY_DN7289_c0_g1_i1	613	-0,720931	-0,430399325	0,0386948	0,08090072
TRINITY_DN30943_c0_g1_i1	325	-0,397937		0,0884243	0,08090624
TRINITY_DN19660_c0_g1_i1	787	-0,280169		0,128844	0,08090998
TRINITY_DN20969_c2_g2_i6	585	-1,06999		0,0713912	0,08141038
TRINITY_DN4683_c0_g1_i2	672	-0,779836		0,304643	0,08156372
TRINITY_DN20266_c0_g1_i2	405	-0,722065	-0,485624243	0,0138644	0,08161352
TRINITY_DN26384_c5_g1_i6	598	-1,08317	-0,44398708	0,0423915	0,08178994
TRINITY_DN41706_c0_g1_i1	313	-1,32263	-0,378669283	0,00870924	0,08192841
TRINITY_DN25279_c0_g1_i7	787	-0,729002			0,08215138
TRINITY_DN22164_c0_g1_i7	767	-0,697788		0,0459335	0,08225769
TRINITY_DN22046_c2_g1_i2	1469	-1,2353	-0,455103536	0,0844637	0,08244042
TRINITY_DN24829_c1_g1_i8	2126	-1,31381	-0,439083783	0,193669	0,08260331
TRINITY_DN23657_c1_g1_i5	1352	-0,744457	-0,429300428	0,0733447	0,08276309
TRINITY_DN39773_c0_g1_i1	363	-1,17498		0,133936	0,08279696
TRINITY_DN24881_c2_g1_i17	1530	-1,05205	-0,392612516	0,151178	0,0828033
TRINITY_DN24853_c0_g1_i19	2352	-0,559496		0,0942249	0,08285258
TRINITY_DN31984_c0_g1_i1	453	-1,00586	-0,366810546	0,102281	0,08330172
TRINITY_DN18587_c0_g1_i5	1923	-0,602676		0,261741	0,08356207
TRINITY_DN21893_c0_g1_i20	2368	-0,580702		0,0538402	0,0839372
TRINITY_DN25024_c0_g1_i10	646	-0,991094	-0,494090114	0,0834956	0,08401967
TRINITY_DN19379_c1_g3_i8	675	-0,578891		0,232393	0,08457681
TRINITY_DN12443_c0_g1_i6	661	-0,827786	-0,327615043	0,196563	0,08476489
TRINITY_DN22728_c1_g1_i2	333	-1,08919	-0,515552583	6,42137E-07	0,08487329
TRINITY_DN21177_c1_g1_i7	3569	-0,91912		0,246747	0,08522038
TRINITY_DN27742_c0_g1_i1	718	-0,596455	-0,357128721	0,0652902	0,08523161
TRINITY_DN23226_c0_g1_i11	706	-0,911731	-0,389857581	0,231582	0,08528463
TRINITY_DN20840_c5_g2_i2	1482	-0,993241	-0,393711391	0,214175	0,08546549
TRINITY_DN24829_c1_g1_i3	2036	-1,31234	-0,439083783	0,266244	0,08549724
TRINITY_DN11046_c0_g1_i2	569	-0,145696			0,08616497
TRINITY_DN23669_c1_g1_i3	700	-0,87824	-0,482823777	0,263761	0,08643399

TRINITY_DN22643_c0_g1_i15	1019	-0,397121		0,438731	0,08671803
TRINITY_DN23596_c0_g2_i1	1062	-1,14063	-0,402444767	0,0536771	0,08701275
TRINITY_DN21537_c0_g6_i7	611	-1,29915	-0,489100525	0,0051354	0,08709826
TRINITY_DN23580_c1_g2_i3	1225	-1,17679		0,100722	0,08727261
TRINITY_DN23442_c5_g1_i4	919	-0,760263		0,0848309	0,08766599
TRINITY_DN21893_c0_g1_i26	1383	-0,881618	-0,294035292	0,0581424	0,08789835
TRINITY_DN19190_c0_g1_i11	1576	-1,0881			0,08832816
TRINITY_DN21589_c0_g2_i5	2098	-1,27607	-0,429459536	0,149179	0,08838225
TRINITY_DN10055_c0_g1_i2	411	-1,11804	-0,388654902	0,00623743	0,08857703
TRINITY_DN18632_c0_g1_i3	1422	-1,04096		0,0425887	0,08881946
TRINITY_DN7444_c0_g1_i1	604	-0,383195		0,125586	0,08906518
TRINITY_DN16953_c0_g1_i3	1588	-0,880236		0,0698023	0,08915669
TRINITY_DN44474_c0_g1_i1	363	-1,05175	-0,401621527	0,0219923	0,08973626
TRINITY_DN14379_c0_g3_i1	397	-0,432814		0,154562	0,0897789
TRINITY_DN22338_c3_g1_i2	373	-0,622132		0,0179238	0,08996226
TRINITY_DN19163_c0_g1_i2	628	-0,658394		0,207578	0,09012831
TRINITY_DN18792_c0_g1_i2	939	-0,658875		0,232062	0,09039164
TRINITY_DN19058_c0_g2_i11	2213	-0,712422			0,09039502
TRINITY_DN21112_c6_g1_i27	2469	-0,609625			0,09045815
TRINITY_DN25988_c1_g3_i1	352	-0,886286	-0,399915193	0,111235	0,09047556
TRINITY_DN24881_c2_g1_i16	1602	-1,0579	-0,392612516	0,107894	0,09057874
TRINITY_DN20109_c0_g1_i7	505	-0,580908		0,0152673	0,09062831
TRINITY_DN21177_c1_g1_i8	3159	-0,904767		0,197308	0,09076324
TRINITY_DN21812_c0_g1_i2	669	-0,997869	-0,318349827	0,465003	0,09112245
TRINITY_DN22046_c2_g1_i3	1466	-1,23495	-0,455103536	0,0824309	0,09147158
TRINITY_DN27680_c0_g1_i1	416	-0,784727	-0,376603458	0,00459667	0,09150701
TRINITY_DN9348_c0_g1_i1	437	-0,686176		0,0624296	0,09191656
TRINITY_DN16953_c0_g1_i5	676	-0,781478		0,0717843	0,09219293
TRINITY_DN20684_c5_g5_i2	824	-0,989281	-0,393738266	0,202799	0,09228609
TRINITY_DN16107_c0_g1_i2	577	-1,04548	-0,413901538	0,0398377	0,09231821
TRINITY_DN24829_c1_g1_i4	2128	-1,31392	-0,439083783	0,158085	0,09293291
TRINITY_DN24129_c1_g2_i4	1119	-1,32301	-0,406821014	0,02107	0,09293421
TRINITY_DN21145_c0_g2_i7	1407	-0,771008		0,478298	0,09297157
TRINITY_DN18632_c0_g1_i8	1442	-1,0459		0,0425887	0,09342608
TRINITY_DN24531_c0_g1_i17	2031	-0,653182		0,324509	0,09348952
TRINITY_DN25493_c13_g1_i7	2421	-1,08003		0,146313	0,09350106
TRINITY_DN25479_c3_g1_i21	2022	-0,880353	-0,37427452	0,155453	0,09352576
TRINITY_DN24829_c1_g1_i1	2124	-1,31381	-0,439083783	0,176237	0,09354596
TRINITY_DN24729_c0_g2_i64	1720	-0,684996		0,298938	0,0936472
TRINITY_DN21112_c6_g1_i18	2670	-0,760772			0,0937631
TRINITY_DN25493_c13_g1_i3	2423	-1,08013		0,14207	0,09383667
TRINITY_DN22971_c1_g1_i8	857	-0,461107			0,09402948
TRINITY_DN22671_c3_g1_i1	378	-0,919147		0,0325036	0,0944523
TRINITY_DN18636_c0_g1_i21	1373	-0,657893			0,09450228
TRINITY_DN18037_c0_g4_i1	422	-1,05158	-0,342840242	0,0201557	0,09465239
TRINITY_DN18683_c0_g1_i13	460	-0,303182		0,141713	0,09469325
TRINITY_DN25800_c3_g4_i3	498	-0,867246	-0,477764276	0,149595	0,09501022
TRINITY_DN21925_c0_g2_i4	652	-0,215389			0,09513612

TRINITY_DN23580_c1_g2_i5	1137	-1,16219		0,0827617	0,09521792
TRINITY_DN18587_c0_g2_i2	682	-0,112948			0,095512
TRINITY_DN797_c0_g1_i1	601	-0,427105	-0,373410693	0,110513	0,09599174
TRINITY_DN24137_c0_g1_i6	299	-0,564963		0,0215161	0,09599327
TRINITY_DN21036_c1_g1_i2	740	-1,00645	-0,455172905	0,0793856	0,09624031
TRINITY_DN20743_c0_g1_i3	692	-0,687608			0,09637252
TRINITY_DN20382_c0_g1_i5	695	-0,698117	-0,331959922		0,09659878
TRINITY_DN25493_c13_g1_i12	2036	-1,14089	-0,331611697	0,051911	0,09696373
TRINITY_DN22835_c1_g1_i1	801	-0,39878			0,09711306
TRINITY_DN20109_c0_g1_i1	688	-0,537674		0,0378414	0,09728007
TRINITY_DN19807_c0_g1_i14	3012	-0,679838			0,09731175
TRINITY_DN27630_c0_g1_i1	413	-0,778244			0,09786019
TRINITY_DN22314_c0_g1_i3	821	-0,978954	-0,333295461	0,0958775	0,09788833
TRINITY_DN14031_c0_g1_i5	910	-0,345005	-0,312130621	0,0557165	0,09837685
TRINITY_DN24881_c2_g1_i15	1641	-1,06094	-0,419646887	0,100023	0,0984487
TRINITY_DN23782_c2_g1_i27	3166	-1,15225		0,0367313	0,09852117
TRINITY_DN7741_c0_g1_i1	618	-0,393916			0,09865813
TRINITY_DN50075_c0_g1_i1	392	-1,27045	-0,333125424	0,211879	0,09875695
TRINITY_DN21302_c1_g2_i1	655	-1,12917	-0,332053914	0,191562	0,09893189
TRINITY_DN23413_c0_g1_i4	2585	-0,631551			0,09899643
TRINITY_DN26425_c1_g1_i14	373	-0,590226	-0,580796842	0,0106401	0,09905538
TRINITY_DN23782_c2_g1_i6	3100	-1,15112		0,0367313	0,09926239
TRINITY_DN19807_c0_g1_i2	3024	-0,68031			0,09960255
TRINITY_DN41998_c0_g1_i1	307	-0,858379	-0,329478497	0,0147069	0,10028652
TRINITY_DN23927_c1_g1_i4	598	-1,1372	-0,367264346	0,15369	0,10052928
TRINITY_DN25623_c0_g1_i17	720	-0,419558	-0,334379263		0,10056456
TRINITY_DN24881_c2_g1_i6	1569	-1,05525	-0,419646887	0,151178	0,1008417
TRINITY_DN10580_c0_g2_i1	498	-0,790642		0,471383	0,10097514
TRINITY_DN26413_c1_g1_i2	352	-0,418097		0,102967	0,10106789
TRINITY_DN13523_c0_g1_i2	464	-0,488014			0,10130129
TRINITY_DN18792_c0_g1_i4	912	-0,650923		0,232062	0,10142267
TRINITY_DN45823_c0_g1_i1	571	-0,44877		0,383945	0,10159178
TRINITY_DN38609_c0_g1_i1	453	-0,498325	-0,418116168	0,148992	0,10173152
TRINITY_DN24881_c2_g1_i18	1813	-1,07322	-0,393671264	0,127338	0,10237857
TRINITY_DN2443_c0_g1_i1	386	-1,2414		0,0512087	0,10248816
TRINITY_DN28261_c0_g1_i1	614	-0,529319	-0,518001195	0,0539219	0,10249568
TRINITY_DN20988_c4_g5_i2	652	-0,52519		0,0288559	0,10269796
TRINITY_DN26462_c1_g1_i6	512	-0,740928		0,0595786	0,10287007
TRINITY_DN21062_c0_g1_i20	603	-0,583289	-0,432058659		0,1029755
TRINITY_DN3353_c0_g1_i1	939	-1,19592	-0,42493817	0,00872869	0,10360596
TRINITY_DN25085_c2_g5_i1	1193	-1,22797	-0,371123764	0,0315126	0,10362296
TRINITY_DN22260_c0_g1_i1	720	-1,18626	-0,392224573	0,0499504	0,10377667
TRINITY_DN23413_c0_g1_i1	2721	-0,638274			0,10378098
TRINITY_DN24881_c2_g1_i14	1813	-1,07322	-0,392612516	0,116082	0,10418304
TRINITY_DN23413_c0_g1_i11	2677	-0,636185			0,10431164
TRINITY_DN19015_c4_g6_i1	420	-1,16828		0,0605461	0,10466465
TRINITY_DN23413_c0_g1_i6	2754	-0,639797			0,10475437
TRINITY_DN27077_c0_g1_i1	332	-0,790177	-0,308401789	0,177964	0,10477831

TRINITY_DN21420_c1_g1_i8	554	-0,764546	-0,490353149	0,40816	0,10479526
TRINITY_DN12186_c0_g1_i1	514	-1,36464	-0,340582399	0,0240881	0,10589625
TRINITY_DN22578_c3_g1_i15	773	-1,07353	-0,438414161	0,0446917	0,10630493
TRINITY_DN23222_c1_g1_i1	457	-1,11883	-0,411760611	0,0247883	0,10723272
TRINITY_DN10549_c0_g1_i2	585	-0,43155		0,380247	0,108083
TRINITY_DN21281_c0_g4_i1	663	-0,906182	-0,424813762	0,15408	0,1082844
TRINITY_DN998_c0_g1_i2	442	-0,721982		0,0101907	0,1083556
TRINITY_DN21919_c0_g3_i8	2392	-1,22828	-0,354715572	0,0446643	0,10868139
TRINITY_DN21893_c0_g1_i29	1400	-0,882374	-0,294035292	0,0508011	0,10885226
TRINITY_DN23560_c0_g1_i16	1417	-1,03829		0,0198546	0,10897669
TRINITY_DN23859_c0_g2_i5	695	-1,14402	-0,355707206	0,0275037	0,10906527
TRINITY_DN35159_c0_g1_i1	587	-0,800407	-0,373596103	0,0668632	0,10911788
TRINITY_DN22681_c1_g3_i2	404	-0,628929	-0,338013646	0,0200989	0,10959396
TRINITY_DN21259_c1_g1_i4	600	-0,664407	-0,345100218		0,11034627
TRINITY_DN22956_c0_g4_i2	396	-0,911606	-0,442511685	0,0188175	0,11092114
TRINITY_DN20109_c0_g1_i3	820	-0,157026		0,087405	0,11093848
TRINITY_DN38739_c0_g1_i1	842	-1,19545	-0,366123392	0,0233295	0,11143286
TRINITY_DN21340_c5_g1_i19	642	-1,16778	-0,506431465	0,174362	0,11145562
TRINITY_DN26251_c1_g3_i1	686	-1,22956	-0,384349934	0,136979	0,11170009
TRINITY_DN25389_c1_g1_i2	826	-1,09461	-0,350202994	0,0143395	0,11241777
TRINITY_DN23441_c1_g4_i1	368	-1,0652	-0,376258737	0,0178935	0,11277657
TRINITY_DN9017_c0_g1_i1	397	-0,537592		0,0488257	0,1129167
TRINITY_DN38465_c0_g1_i1	666	-1,01621	-0,346196906	0,138527	0,11303287
TRINITY_DN20104_c2_g2_i7	537	-0,812523	-0,352998216	0,0152443	0,11321499
TRINITY_DN1775_c0_g1_i1	516	-0,957365		0,0460832	0,11331451
TRINITY_DN48656_c0_g1_i1	491	-0,465917		0,131597	0,11348048
TRINITY_DN18644_c0_g2_i1	885	-1,30932	-0,340617807	0,0155257	0,11369929
TRINITY_DN741_c0_g1_i1	402	-1,18671	-0,328586218	0,0205578	0,11426536
TRINITY_DN21043_c0_g1_i21	467	-0,79437	-0,41238217	0,13724	0,11430755
TRINITY_DN6442_c0_g1_i1	501	-0,733198			0,11433387
TRINITY_DN1882_c0_g1_i1	481	-0,87405		0,0264555	0,11480358
TRINITY_DN19190_c0_g1_i1	2057	-1,08187			0,11494539
TRINITY_DN23505_c1_g2_i1	789	-1,12632	-0,351538515	0,0421478	0,11498105
TRINITY_DN19807_c0_g1_i10	3014	-0,782856			0,11506525
TRINITY_DN34612_c0_g1_i1	345	-1,2753	-0,388553359	0,018778	0,11563444
TRINITY_DN24678_c0_g3_i3	428	-1,04138	-0,414965253	0,251035	0,11618814
TRINITY_DN26057_c1_g1_i7	471	-0,0438325			0,11667023
TRINITY_DN24116_c1_g3_i4	616	-1,00926	-0,391262696	0,214146	0,11688625
TRINITY_DN24129_c1_g2_i5	1199	-1,32909	-0,413296273	0,0241832	0,11689599
TRINITY_DN26614_c1_g4_i1	416	-0,733379		0,0132575	0,11696717
TRINITY_DN24729_c0_g2_i48	1608	-0,672786		0,298938	0,11797
TRINITY_DN13982_c0_g1_i1	806	-0,652312		0,348319	0,11838315
TRINITY_DN24919_c2_g4_i4	586	-0,996468	-0,352161917	0,0387453	0,11865258
TRINITY_DN5516_c0_g1_i1	1146	-0,545638			0,11869971
TRINITY_DN22229_c0_g1_i6	1084	-0,997687	-0,481532447	0,0212083	0,11893185
TRINITY_DN21759_c1_g3_i3	496	-0,90719	-0,388215747	0,107127	0,1192653
TRINITY_DN10797_c0_g1_i2	912	-0,542637	-0,435476329	0,0284686	0,11930488
TRINITY_DN24419_c0_g4_i7	369	-1,0103	-0,397163954	0,00867928	0,11963302

TRINITY_DN20687_c12_g1_i4	2470	-0,741858			0,12021767
TRINITY_DN23622_c0_g1_i7	1119	-0,477734		0,476751	0,12023939
TRINITY_DN21259_c1_g2_i1	685	-1,42158	-0,34258888	0,0262053	0,1205368
TRINITY_DN23280_c1_g4_i10	643	-1,08501	-0,375502658	0,098593	0,1206443
TRINITY_DN26591_c1_g1_i1	779	-1,11915	-0,471756106	0,0319417	0,12151326
TRINITY_DN25467_c2_g1_i2	643	-1,18274	-0,322108663	0,0189066	0,12162206
TRINITY_DN24788_c0_g1_i1	785	-1,16568		0,130451	0,12217934
TRINITY_DN21655_c1_g1_i15	1126	-0,550091			0,12233816
TRINITY_DN22692_c0_g1_i3	1917	-0,743549			0,12253247
TRINITY_DN22634_c0_g1_i2	604	-0,488154		0,00343698	0,12253523
TRINITY_DN20611_c3_g2_i1	302	-0,921443	-0,590863061	0,00988666	0,12266002
TRINITY_DN44713_c0_g1_i1	594	-1,06472	-0,34600849	0,0246779	0,12297484
TRINITY_DN25552_c0_g1_i8	729	-1,07791	-0,367282319	0,00516569	0,12303571
TRINITY_DN20665_c5_g2_i1	816	-0,982877	-0,370669955	0,083498	0,12317279
TRINITY_DN26486_c0_g1_i4	642	-0,852653	-0,40070389	0,121662	0,12322244
TRINITY_DN25008_c1_g1_i11	1043	-0,72393			0,12352767
TRINITY_DN12634_c0_g1_i5	550	-0,815091		0,0191379	0,12421209
TRINITY_DN21579_c0_g1_i1	500	-1,21523	-0,493415903	0,0146203	0,12426917
TRINITY_DN23119_c3_g1_i1	479	-0,928885	-0,399623726	0,0066211	0,1245265
TRINITY_DN20369_c1_g2_i1	510	-1,13227	-0,506570018	0,0337393	0,12589339
TRINITY_DN18636_c0_g1_i12	1343	-0,855817		0,223752	0,12599231
TRINITY_DN22626_c5_g3_i2	492	-0,713325		0,0609323	0,12627623
TRINITY_DN10055_c0_g1_i1	344	-1,10197	-0,388654902	0,0150483	0,12645336
TRINITY_DN25772_c2_g3_i12	711	-1,24445	-0,511573399	0,0244272	0,12655306
TRINITY_DN24099_c0_g1_i2	552	-1,23113	-0,451575188	0,0385285	0,12661738
TRINITY_DN11664_c0_g1_i1	427	-0,333816		0,0761414	0,12674749
TRINITY_DN15620_c0_g1_i1	746	-0,580844		0,00804973	0,12682262
TRINITY_DN21394_c2_g2_i1	610	-1,10203	-0,373617082	0,0341873	0,12713176
TRINITY_DN24099_c0_g1_i1	675	-1,25568	-0,451575188	0,0307685	0,12817579
TRINITY_DN23252_c6_g1_i1	347	-1,12226	-0,399133242	0,00567498	0,12822638
TRINITY_DN26057_c1_g1_i4	759	-0,0957582			0,12834901
TRINITY_DN23729_c9_g1_i1	1144	-0,679481			0,12838487
TRINITY_DN21585_c1_g6_i1	354	-1,18185	-0,46052413	0,0205388	0,12891101
TRINITY_DN21216_c0_g1_i2	1524	-1,36813	-0,409570855	0,0151921	0,12900073
TRINITY_DN36501_c0_g1_i1	467	-1,06288	-0,33935562	0,109815	0,12908129
TRINITY_DN48196_c0_g1_i1	525	-0,990804	-0,334533013	0,0678817	0,12934555
TRINITY_DN21139_c0_g3_i2	509	-1,37446	-0,448722187	0,00769684	0,12937448
TRINITY_DN26307_c1_g2_i15	559	-0,591193		0,0910693	0,12939159
TRINITY_DN23317_c0_g1_i12	1123	-0,718082		0,00817887	0,1296264
TRINITY_DN23459_c2_g2_i6	361	-1,07307	-0,424936074	0,0123547	0,13000383
TRINITY_DN24714_c0_g2_i9	565	-1,17592	-0,441594757	0,0100026	0,1309225
TRINITY_DN18643_c0_g1_i5	364	-0,820432	-0,340781301	0,177136	0,13136867
TRINITY_DN18587_c0_g2_i1	871	-0,183437			0,13172483
TRINITY_DN25276_c1_g1_i1	917	-0,929021	-0,453361747	0,0180789	0,13207127
TRINITY_DN22334_c0_g1_i42	1915	-0,580766			0,13210785
TRINITY_DN21073_c5_g2_i1	397	-0,0970164	-0,322695305	0,0604683	0,13220388
TRINITY_DN22164_c0_g1_i12	640	-0,42655		0,0739752	0,13289884
TRINITY_DN26425_c1_g1_i9	383	-0,591278	-0,580796842	0,0347895	0,13293907

TRINITY_DN1216_c0_g1_i1	492	-0,511607		0,0972782	0,13367839
TRINITY_DN18781_c0_g1_i3	376	-0,0659649			0,13444877
TRINITY_DN21713_c0_g5_i1	369	-1,0052	-0,311258889	0,00462219	0,1344661
TRINITY_DN19379_c1_g1_i4	477	-0,624348	-0,297888932	0,0148764	0,13463391
TRINITY_DN20758_c0_g1_i3	1079	-0,416071			0,1349009
TRINITY_DN19058_c0_g2_i2	2458	-0,73414			0,13512928
TRINITY_DN25008_c1_g2_i8	545	-0,894642	-0,435027374	0,193336	0,13560934
TRINITY_DN26138_c2_g1_i1	302	-1,35456	-0,417900263	0,00895457	0,13588938
TRINITY_DN24807_c1_g2_i2	375	-0,754772		0,0454971	0,13617946
TRINITY_DN24825_c1_g1_i3	473	-1,21617	-0,390727381	0,0108892	0,13679577
TRINITY_DN38114_c0_g1_i1	580	-1,25354	-0,46442102	0,0154769	0,13681054
TRINITY_DN24531_c0_g1_i2	2829	-0,653813			0,13740931
TRINITY_DN21558_c0_g1_i4	472	-0,442734	-0,380042123		0,13815269
TRINITY_DN24013_c2_g1_i2	828	-0,682748	-0,405466749	0,0143817	0,13815841
TRINITY_DN38272_c0_g1_i1	447	-0,97767	-0,356732654	0,191236	0,13847683
TRINITY_DN19484_c0_g1_i13	1630	-0,28703			0,13897023
TRINITY_DN19190_c0_g1_i13	1995	-1,05415			0,13938027
TRINITY_DN18662_c0_g1_i2	475	-0,554122		0,0260807	0,13941858
TRINITY_DN23569_c1_g1_i1	2317	-0,753288		0,204548	0,1399884
TRINITY_DN25239_c1_g1_i2	473	-1,26942	-0,34284728	0,162004	0,14056197
TRINITY_DN22707_c0_g1_i11	515	-0,685338	-0,487895441	0,365439	0,14243236
TRINITY_DN11726_c0_g2_i1	336	-1,01717	-0,3579714	0,0148506	0,14245194
TRINITY_DN35028_c0_g1_i1	361	-1,06833	-0,38771039	0,0121509	0,14266357
TRINITY_DN13126_c0_g1_i1	725	-0,675542			0,14301817
TRINITY_DN24574_c0_g1_i3	1840	-0,555881			0,14320591
TRINITY_DN22527_c1_g5_i1	856	-0,183194			0,14327214
TRINITY_DN35841_c0_g1_i1	303	-0,716873	-0,306285317	0,0196896	0,14368929
TRINITY_DN22334_c0_g1_i15	2245	-0,611362			0,14412038
TRINITY_DN10547_c0_g1_i1	936	-0,570989			0,14428801
TRINITY_DN24729_c0_g2_i33	1608	-0,672485		0,313908	0,14435012
TRINITY_DN26244_c1_g3_i2	502	-0,386081		0,164171	0,144908
TRINITY_DN23540_c2_g4_i2	938	-0,766421		0,339802	0,14611934
TRINITY_DN23579_c2_g2_i1	362	-0,358594	-0,320456022		0,1490813
TRINITY_DN47926_c0_g1_i1	363	-1,30124	-0,450915446	0,0200137	0,14937274
TRINITY_DN23622_c0_g1_i5	1090	-0,411258			0,14980791
TRINITY_DN29692_c0_g1_i1	330	-0,961853	-0,349553643	0,00491516	0,14997174
TRINITY_DN18683_c0_g1_i6	771	-0,0825042			0,15015313
TRINITY_DN18144_c0_g1_i2	479	-0,192691			0,15169606
TRINITY_DN24691_c4_g1_i7	622	-0,349106			0,151718
TRINITY_DN22277_c0_g1_i1	447	-0,734905	-0,397885395	0,0617858	0,15215434
TRINITY_DN20794_c1_g1_i4	1435	-0,679053		0,379615	0,15225853
TRINITY_DN24531_c0_g1_i11	1699	-0,603143		0,277181	0,15447468
TRINITY_DN48681_c0_g1_i1	353	-0,354191	-0,334866666		0,15580138
TRINITY_DN14664_c0_g1_i1	610	-0,0787622			0,15662313
TRINITY_DN26236_c2_g1_i3	810	-0,340128			0,15975419
TRINITY_DN24062_c2_g1_i3	4139	-0,331476			0,16027664
TRINITY_DN22129_c1_g1_i12	1851	-0,334772			0,16082814
TRINITY_DN21597_c2_g1_i4	2120	-0,336071			0,16087383

TRINITY_DN25608_c0_g2_i10	357	-0,89929	-0,354146403	0,189552	0,16107662
TRINITY_DN24316_c0_g5_i1	411	-0,83464	-0,317221102		0,16145444
TRINITY_DN23489_c1_g2_i12	1762	-1,11338			0,16175841
TRINITY_DN21933_c1_g1_i1	643	-0,767178	-0,452120284	0,160472	0,16225026
TRINITY_DN25147_c3_g1_i4	2062	-0,383752			0,16252653
TRINITY_DN23489_c1_g2_i10	1777	-1,11494			0,16313997
TRINITY_DN14303_c0_g1_i3	432	-0,359121			0,16492239
TRINITY_DN15692_c0_g1_i2	1101	-0,4042			0,16579403
TRINITY_DN24818_c1_g3_i2	375	-0,818246	-0,467481348	0,00821479	0,16669166
TRINITY_DN20252_c8_g1_i3	849	-0,150221			0,16737509
TRINITY_DN23782_c2_g1_i7	3016	-1,14967		0,0302898	0,16757819
TRINITY_DN23630_c0_g1_i6	715	-0,301986			0,17059978
TRINITY_DN24691_c4_g1_i6	546	-0,53539			0,17117717
TRINITY_DN26111_c3_g1_i4	1483	-0,853685			0,17129517
TRINITY_DN26095_c3_g1_i2	889	-0,14557			0,17237483
TRINITY_DN16661_c0_g1_i4	610	-0,459812			0,17265648
TRINITY_DN22728_c1_g6_i1	372	-0,748806	-0,571043729	1,47229E-06	0,17734579
TRINITY_DN24321_c0_g5_i1	568	-1,12052		0,033494	0,17855764
TRINITY_DN22261_c0_g1_i12	3610	-0,825564			0,17879076
TRINITY_DN20969_c1_g1_i4	2668	-0,657314			0,1790192
TRINITY_DN25656_c4_g1_i1	1830	-0,783203			0,18036587
TRINITY_DN21236_c0_g2_i13	2456	-0,551135			0,18108345
TRINITY_DN22423_c0_g3_i1	430	-0,827316	-0,399220452	0,16916	0,18166527
TRINITY_DN19379_c1_g3_i10	464	-0,0296817		0,091959	0,18357367
TRINITY_DN1111_c0_g1_i1	338	-1,00154	-0,479058697	1,66064E-05	0,18377717
TRINITY_DN6700_c0_g1_i2	900	-0,986226	-0,473086489	0,0974181	0,18392961
TRINITY_DN24729_c0_g2_i2	1593	-0,690993		0,0540066	0,18497836
TRINITY_DN24853_c0_g1_i33	1157	-0,22425		0,394478	0,18600587
TRINITY_DN20053_c4_g2_i5	3283	-0,871573			0,18781337
TRINITY_DN24691_c4_g1_i3	539	-0,307398			0,19004377
TRINITY_DN25868_c1_g1_i2	1320	-0,772998		0,116496	0,19238488
TRINITY_DN21931_c3_g2_i1	1060	-0,27446		0,0261734	0,19458608
TRINITY_DN10017_c0_g1_i1	469	-0,800877	-0,344538106	0,10039	0,19759497
TRINITY_DN24231_c5_g1_i1	420	-0,380342	-0,320136388	0,21727	0,19891368
TRINITY_DN19299_c0_g1_i4	1452	-0,758063		0,246188	0,19933285
TRINITY_DN25650_c0_g3_i3	432	-0,393035		0,0779078	0,20047755
TRINITY_DN20208_c0_g2_i3	1876	-0,822359			0,20306088
TRINITY_DN6146_c0_g1_i2	676	-0,372739			0,20431125
TRINITY_DN16661_c0_g1_i5	376	-1,08438	-0,322272949	0,0904234	0,20455599
TRINITY_DN22855_c4_g2_i1	304	-0,811985	-0,408238596	5,05042E-06	0,20576317
TRINITY_DN21236_c0_g2_i12	2237	-0,53406			0,2069513
TRINITY_DN17901_c0_g2_i4	715	-0,0384327			0,21210077
TRINITY_DN25983_c0_g2_i3	643	-0,712106			0,22126282
TRINITY_DN23664_c0_g2_i19	662	-0,300772			0,22182042
TRINITY_DN24174_c0_g2_i7	1010	-0,262605			0,22397071
TRINITY_DN18655_c0_g1_i7	1986	-0,712316		0,111395	0,22756615
TRINITY_DN23105_c0_g2_i3	470	-0,442465		0,0672565	0,23649542
TRINITY_DN20182_c0_g2_i4	1676	-0,631281			0,23901489

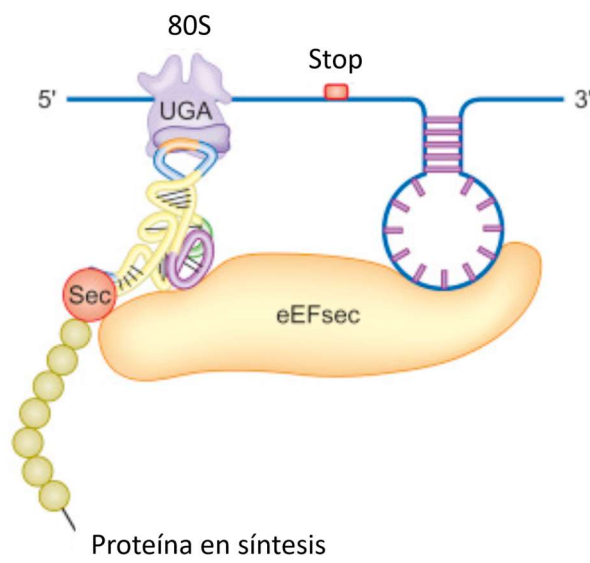
TRINITY_DN24174_c0_g2_i1	1022	-0,264709			0,24007626
TRINITY_DN17229_c0_g2_i4	1789	-0,990594			0,24020505
TRINITY_DN24729_c0_g2_i47	1386	-0,673696		0,0497849	0,25037566
TRINITY_DN9594_c0_g1_i1	388	-0,898053	-0,472287167	0,056919	0,25116714
TRINITY_DN21825_c2_g1_i6	978	-0,775578			0,25269753
TRINITY_DN20955_c0_g1_i6	1329	-0,27354		0,307775	0,25359489
TRINITY_DN18258_c0_g1_i9	963	-0,359948			0,26041396
TRINITY_DN24103_c0_g1_i4	1019	-0,362826			0,26100919
TRINITY_DN23630_c0_g1_i8	918	-0,373342			0,26188777
TRINITY_DN26221_c1_g1_i1	766	-0,606304		0,279633	0,26294871
TRINITY_DN30160_c0_g1_i1	718	-0,458267	-0,361682476	0,0640699	0,26522022
TRINITY_DN26057_c1_g1_i3	432	-0,338677		0,147335	0,26635833
TRINITY_DN24996_c5_g1_i2	672	-0,678535		0,0792293	0,26665364
TRINITY_DN25019_c0_g2_i6	955	-0,914518			0,2668913
TRINITY_DN22693_c1_g1_i8	514	-0,00568203		0,282428	0,27390312
TRINITY_DN23191_c0_g3_i2	1928	-0,55625			0,27445479
TRINITY_DN21361_c0_g1_i2	556	-0,291959		0,12149	0,27814046
TRINITY_DN22334_c0_g1_i27	1781	-0,497261		0,216567	0,2842549
TRINITY_DN20351_c0_g3_i4	1039	-0,486495	-0,409039114	0,184866	0,28593588
TRINITY_DN12489_c0_g1_i1	1186	-0,491221			0,2888223
TRINITY_DN22277_c0_g1_i4	578	-0,528227	-0,411083256	0,151662	0,28986347
TRINITY_DN23664_c0_g2_i21	732	-0,311083			0,2934927
TRINITY_DN41590_c0_g1_i1	320	-1,03375	-0,381016194	0,0436368	0,29359728
TRINITY_DN24784_c1_g3_i2	2360	-0,488275			0,29904135
TRINITY_DN15601_c0_g1_i1	720	-0,828994			0,30021086
TRINITY_DN21259_c1_g1_i5	636	-0,50836		0,143715	0,30087898
TRINITY_DN22477_c1_g2_i10	1196	-0,569654			0,30216348
TRINITY_DN20852_c2_g2_i2	1144	-0,211187			0,30254803
TRINITY_DN24062_c2_g2_i7	3301	-0,34256			0,30477017
TRINITY_DN23663_c2_g5_i2	893	-0,076945			0,30699372
TRINITY_DN18949_c1_g1_i5	321	-0,524512	-0,344426059	0,0249175	0,30705298
TRINITY_DN21597_c2_g1_i11	2162	-0,316465			0,3072019
TRINITY_DN21314_c0_g2_i8	2064	-0,253494			0,30722101
TRINITY_DN16389_c0_g1_i2	858	-0,294973			0,31048936
TRINITY_DN19047_c0_g1_i4	886	-0,484677			0,31342383
TRINITY_DN25698_c1_g3_i6	633	-0,825096	-0,502343316		0,31464355
TRINITY_DN20380_c8_g1_i4	822	-0,641491			0,31518008
TRINITY_DN25200_c3_g1_i17	472	-0,092296	-0,387766253		0,32620931
TRINITY_DN6345_c0_g1_i1	368	-0,951582	-0,472767759	0,199292	0,32982984
TRINITY_DN26642_c8_g1_i2	1100	-0,50025			0,33141704
TRINITY_DN26362_c1_g2_i2	628	-0,605002		0,151928	0,33849049
TRINITY_DN22973_c1_g1_i2	680	-0,371938			0,33877253
TRINITY_DN26201_c2_g1_i1	1003	-0,336018			0,35243782
TRINITY_DN18520_c0_g1_i2	1032	-0,735716			0,36483999
TRINITY_DN21557_c0_g2_i11	2137	-0,118222			0,36760831
TRINITY_DN22973_c1_g1_i3	726	-0,16611			0,36880171
TRINITY_DN20969_c2_g1_i14	3408	-0,264367			0,36948343
TRINITY_DN22107_c0_g2_i1	1142	-0,415866		0,186676	0,37414521

TRINITY_DN22477_c1_g2_i7	1278	-0,540959			0,38404353
TRINITY_DN16430_c0_g1_i7	978	-0,30316			0,38582076
TRINITY_DN22477_c1_g2_i14	1290	-0,54334			0,38713859
TRINITY_DN18490_c0_g1_i2	806	-0,023144			0,39092512
TRINITY_DN25030_c1_g1_i13	2811	-1,17442			0,39833656
TRINITY_DN18660_c0_g1_i5	1363	-0,656915			0,39918293
TRINITY_DN22693_c2_g3_i1	372	-0,489199		0,100383	0,39977504
TRINITY_DN3617_c0_g2_i1	639	-0,371327		0,153772	0,40728966
TRINITY_DN32461_c0_g1_i1	381	-1,24263	-0,351739666	4,796E-07	0,4213262
TRINITY_DN20208_c0_g2_i7	1636	-0,687222			0,42388637
TRINITY_DN19379_c1_g3_i12	503	-0,0522535		0,142594	0,42745992
TRINITY_DN22830_c3_g2_i5	784	-0,145627			0,43160964
TRINITY_DN25366_c1_g2_i2	316	-0,467996		0,0601864	0,43358168
TRINITY_DN27370_c0_g1_i1	637	-0,579646		0,210583	0,43783268
TRINITY_DN16430_c0_g1_i4	1010	-0,312452			0,44170279
TRINITY_DN24928_c0_g3_i4	315	-0,173155		0,0496399	0,44222252
TRINITY_DN8439_c0_g1_i2	672	-0,739822		0,292588	0,45558581
TRINITY_DN42011_c0_g1_i1	384	-1,05501	-0,399051122	0,18825	0,45991199
TRINITY_DN10547_c0_g1_i2	864	-0,240095			0,46532194
TRINITY_DN22124_c0_g2_i7	950	-0,316982			0,46932496
TRINITY_DN51276_c0_g1_i1	307	-0,725775	-0,403058037	0,105972	0,47290198
TRINITY_DN24151_c0_g1_i2	1068	-0,212138			0,47468297
TRINITY_DN21112_c6_g1_i31	4986	-0,816639		0,404479	0,47929779
TRINITY_DN17121_c0_g2_i1	300	-0,570393		1,33591E-05	0,49091742
TRINITY_DN21968_c0_g1_i2	1350	-0,136046			
TRINITY_DN21665_c1_g6_i1	321	-0,336543	-0,449091043	2,27445E-06	
TRINITY_DN22462_c3_g1_i2	2090	-0,292571			
TRINITY_DN25621_c2_g1_i3	991	-0,849904			
TRINITY_DN23038_c0_g2_i5	1758	-0,334835			
TRINITY_DN22973_c1_g3_i10	368	-0,496401		0,00322434	
TRINITY_DN17121_c0_g1_i3	587	-0,0858468		0,148632	
TRINITY_DN12933_c0_g1_i1	797	-0,0664759		0,378137	
TRINITY_DN22973_c0_g1_i5	347	-0,615491		1,89757E-05	
TRINITY_DN19795_c0_g2_i10	2110	-0,789745			
TRINITY_DN18035_c0_g1_i4	1308	-0,0541553			
TRINITY_DN25366_c1_g2_i1	298	-0,426025		0,046017	
TRINITY_DN49414_c0_g1_i1	375	-0,12227		0,00980859	
TRINITY_DN19795_c0_g2_i6	1973	-0,847636			
TRINITY_DN49933_c0_g1_i1	768	-0,430471		0,192834	
TRINITY_DN23664_c0_g2_i7	803	-0,137554			
TRINITY_DN22973_c0_g1_i7	380	-0,530732		0,00810196	
TRINITY_DN15305_c0_g1_i2	1568	-0,0402601			
TRINITY_DN22973_c0_g1_i4	455	-0,180225		1,89757E-05	
TRINITY_DN24231_c5_g4_i2	519	-0,196536	-0,39984923	0,16653	
TRINITY_DN22728_c1_g3_i9	1813	-0,0881269			
TRINITY_DN24962_c0_g2_i3	726	-0,182434		0,228174	
TRINITY_DN16422_c0_g1_i2	952	-0,534134			
TRINITY_DN25307_c1_g1_i4	2262	-0,406251			

TRINITY_DN26331_c5_g1_i1	329	-1,0125		0,0685917
TRINITY_DN22973_c1_g2_i2	380	-0,223048		0,165119
TRINITY_DN26113_c2_g1_i3	339	-1,52383	-0,399699975	7,62483E-06
TRINITY_DN11372_c0_g2_i1	419	-0,798171	-0,438940846	0,110494
TRINITY_DN3987_c0_g2_i1	310	-0,819436	-0,379100361	0,128025
TRINITY_DN23287_c0_g2_i4	1385	-0,58299		
TRINITY_DN18038_c0_g1_i1	353	-0,961744		0,182952
TRINITY_DN22973_c0_g1_i1	348	-0,541666		1,89757E-05
TRINITY_DN12651_c0_g2_i1	393	-0,561173	-0,439454126	0,00304296
TRINITY_DN24843_c0_g1_i10	463	-0,72009	-0,343239998	0,151343
TRINITY_DN20352_c4_g5_i12	2405	-0,419773		
TRINITY_DN24358_c1_g5_i7	1916	-0,376917		
TRINITY_DN17673_c0_g1_i3	1119	-0,410381	-0,335043424	0,108037
TRINITY_DN18255_c0_g1_i1	728	-0,38136		
TRINITY_DN19675_c1_g1_i4	372	-0,30678		0,167706
TRINITY_DN12528_c0_g1_i3	320	-1,01545	-0,426233797	0,0468966
TRINITY_DN25866_c0_g2_i4	412	-0,718858		0,108313
TRINITY_DN41220_c0_g1_i1	826	-0,00372846		
TRINITY_DN19440_c0_g1_i2	1795	-0,0133248		
TRINITY_DN22698_c0_g2_i1	748	-0,148648		0,291285
TRINITY_DN42111_c0_g1_i1	391	-0,966443		0,0370544
TRINITY_DN21716_c0_g1_i4	835	-0,159305		0,273697
TRINITY_DN23301_c0_g2_i1	316	-1,27061		6,09424E-07
TRINITY_DN18038_c0_g1_i5	371	-0,979162		0,176365
TRINITY_DN19807_c0_g1_i12	2101	-0,351543		
TRINITY_DN23301_c0_g2_i4	328	-1,26573		6,09424E-07
TRINITY_DN25665_c0_g2_i1	689	-0,0156171		0,201506
TRINITY_DN25146_c3_g2_i11	1792	-0,538717		
TRINITY_DN25743_c1_g1_i2	1966	-0,359862		
TRINITY_DN22986_c0_g1_i7	559	-0,144793		0,174112
TRINITY_DN31570_c0_g1_i1	315	-0,729917		0,130466
TRINITY_DN18922_c0_g1_i10	1295	-0,201953	-0,336566562	
TRINITY_DN21655_c1_g1_i13	1751	-0,398693		
TRINITY_DN26294_c0_g2_i3	1213	-0,514719		
TRINITY_DN22529_c2_g1_i9	2000	-0,614175		
TRINITY_DN20171_c3_g5_i1	1035	-0,406551		
TRINITY_DN3898_c0_g1_i1	426	-0,721222	-0,415855154	0,0932035
TRINITY_DN24871_c1_g2_i6	1173	-0,415093		0,0996803
TRINITY_DN15394_c0_g1_i2	348	-0,19934		0,150719
TRINITY_DN21366_c1_g1_i15	2580	-0,0846076		
TRINITY_DN18721_c0_g1_i2	859	-0,300355		0,31041
TRINITY_DN21716_c0_g1_i1	988	-0,0679118		
TRINITY_DN24766_c2_g2_i9	1097	-0,226964		
TRINITY_DN16664_c0_g1_i3	1152	-0,507379		
TRINITY_DN22973_c1_g1_i4	422	-0,0203997		0,169635
TRINITY_DN22164_c0_g1_i11	1501	-0,4616		
TRINITY_DN22033_c1_g2_i1	1718	-0,322185		
TRINITY_DN22693_c1_g1_i7	540	-0,0172749		0,020842

TRINITY DN24224 c1 g1 i7	1467	-0,331906	
TRINITY_DN23579_c2_g1_i6	1167	-0,000545712	
TRINITY DN22482 c0 g2 i12	1050	-0,0233143	0,392504
TRINITY_DN22482_c0_g2_i8	1053	-0,0231549	0,392504
TRINITY_DN20790_c2_g1_i2	1363	-0,322236	
TRINITY DN22913 c1 g1 i4	2025	-0,0258182	

Anexo 3. Elemento SECIS en eucariotas



* Adaptado de Clark y Pazdernik, 2013.

Anexo 4. Datos de conservación estructural mediante *Crossalign*.

Familia	Transcripto	Código	Distancia estructural normalizada	Valor p	Correlación	Aligned Region In the Longer Profile	T (U)	C	A	G	GC content	Length	Optimal Secondary Structure Minimum Free Energy (kcal/mol)	Free energy of the thermodynamic ensemble	Frequency of the MFE structure in the ensemble	Ensemble diversity	Centroid secondary structure in dot-bracket notation minimum free energy (kcal/mol)
Fabaceae	Lupinus_mutabilis	TRINITY_DN19379_c1_g3_8	0	0	1	1-676	28,1	28,3	30,8	12,7	41	676	-104,4	-117,2	0	136,7	-80,34
Solanaceae	Solanum_lycopersicum	Inc_RNA_Solyc04r02221.0.1.1 gene=Solyc04r022210.1	0,095	0,0	0,91	115-790	30,7	7,9	23,9	37,5	45	707	-145,1	-158,7	0	229,45	-85,76
Brassicaceae	Brassica_napus	BNAP_LNC003827.4 gene=BNAP_LNC003827	0,097	0,0	0,91	136-811	24,8	8,4	26,40	40,8	49	730	-154,5	-168,2	0	227,45	-103,2
Brassicaceae	Brassica_rapa	BRAP_LNC001139.3 gene=BRAP_LNC003139	0,098	0,1	0,9	323-998	24,1	8,8	23,43	43,3	52	1008	-242,1	-258,8	0	220,92	-189,8
Fabaceae	Medicago_truncatula	MTRU_LNC004005.50 gene=MTRU_LNC004005	0,098	0,1	0,91	298-973	33,2	8,4	29,9	28,5	37	1012	-220,3	-241,3	0	278,58	-139,2
Actinidiaceae	Actinidia_chinensis	ACHI_LNC008871.8 gene=ACHI_LNC008871	0,099	0,1	0,89	162-837	22,6	33,2	31,3	12,8	46	756	-79,67	-92,21	0	167,3	-28,61
Salicaceae	Populus_tremula	PTRE_LNC002321.1 gene=PTRE_LNC002321	0,101	0,2	0,9	306-981	29,8	14,4	29,26	26,7	41	1086	-274,7	-295,9	0	223,18	-242,2
Fabaceae	Vigna_radiata	VRAD_LNC00907.1 gene=VRAD_LNC00907	0,101	0,2	0,9	1334-2009	29,10	10,6	34,24	24,9	36	1878	-392,2	-429,5	0	582,43	-256,5
Arecaceae	Elaeis_guineensis	EGUI_LNC025832.7 gene=EGUI_LNC025832	0,102	0,7	0,87	324-688	25,5	5,2	25,44	44,3	49	364	-54,7	-60,07	0,02	87,54	-32,1
Euphorbiaceae	Manihot_esculenta	MESC_LNC006791.5 gene=MESC_LNC006791	0,103	0,3	0,89	232-907	25,8	11,6	28,2	34,4	46	791	-203,6	-216,4	0	162,92	-170,7
Brassicaceae	Eutrema_salsuginum	ESAL_LNC001003.35 gene=ESAL_LNC001003	0,104	0,8	0,89	301-727	26,8	6,6	25,41	41,5	48	426	-78,3	-86,62	0	126,07	-31,5
Linaceae	Linum_usitatissimum	LUSI_LNC005493.9 gene=LUSI_LNC005493	0,104	0,4	0,89	34-709	32,6	13,6	26,26	26,8	40	984	-253,8	-272,6	0	366,61	-122,2
Rosaceae	Fragaria_vesca	FVES_LNC002111.17 gene=FVES_LNC002111	0,107	0,9	0,86	398-728	27,6	7,6	28,37	37,6	44	330	-58,2	-61,62	0,39	50,51	-53,1
Fabaceae	Arachis_hypogaea	AIPA_LNC002320.2 gene=AIPA_LNC002320	0,108	0,9	0,86	408-665	27,2	8,5	17,49	49,9	55	257	-46,6	-51,28	0,05	45,79	-20
Rutaceae	Citrus_maxima	CMAX_LNC002496.5 gene=CMAX_LNC002496	0,108	1	0,85	262-676	24,2	11,4	24,9	39,6	51	414	-104,8	-111,3	0	127,54	-64,43
Fabaceae	Pisum_sativum	PSAT_LNC036744.6 gene=PSAT_LNC036744	0,108	1	0,84	278-641	29,2	5,2	23,7	41,9	47	363	-51	-57,79	0	105,25	-31,5
Solanaceae	Capsicum_annuum	CANN_LNC006489.2 gene=CANN_LNC006489	0,109	0,9	0,86	394-614	31,8	5,9	20,42	48,3	48	220	-28,3	-34,39	0,01	83,09	-5,4
Rutaceae	Citrus_sinensis	CSIN_LNC002075.6 gene=CSIN_LNC002075	0,109	1	0,81	122-482	23,3	8,3	15,53	53,1	61	360	-65,2	-72,91	0	103,18	-29,5
Apiaceae	Daucus_carota	DCAR_LNC002982.25 gene=DCAR_LNC002982	0,109	1	0,85	359-744	27,5	3,16	53,56	56,1	385	385	-43,8	-51,16	0	96,85	-30,5
Euphorbiaceae	Jatropha_curcas	JCUR_LNC00659.5 gene=JCUR_LNC00659	0,109	0,9	0,81	199-469	30,4	7,8	14,47	47,55	55	270	-53,5	-59,74	0	67,23	-42,7
Oleaceae	Olea_europaea	OEUR_LNC006518.2 gene=OEUR_LNC006518	0,109	0,9	0,85	456-707	29,5	5,6	19,45	51,9	51	251	-36,9	-43,1	0	75,13	-19,9
Brassicaceae	Arabidopsis_thaliana	ATHA_LNC000044.2 gene=ATHA_LNC000044	0,11	1	0,87	278-719	28,3	12,5	27,32	44,2	44	441	-121,7	-130,3	0	146,93	-69
Zosteraceae	Zostera_marina	ZMAR_LNC001227.10 gene=ZMAR_LNC001227	0,11	1	0,86	226-660	25,8	16,8	23,33	33,9	51	434	-131,6	-140,7	0	56,16	-121,7
Fabaceae	Lupinus_angustifolius	LANG_LNC003548.1 gene=LANG_LNC003548	0,111	1	0,85	300-745	29,7	4,26	37,45	45,3	45	445	-91	-99,07	0	142,25	-35,1
Rosaceae	Malus_domestica	MDOM_LNC008886.172 gene=MDOM_LNC008886	0,111	1	0,85	288-724	26,6	6,9	17,49	49,5	56	436	-89,6	-97,55	0	118,63	-63,35
Solanaceae	Solanum_pimpinellifolium	SPIM_LNC000101.1 gene=SPIM_LNC000101	0,111	0,9	0,86	449-672	31,8	8,5	19,7	49,9	48	223	-49,5	-52,8	0,47	77,58	-11,9
Solanaceae	Nicotiana_tabacum	NTAB_LNC005274.9 gene=NTAB_LNC005274	0,112	0,9	0,85	426-699	29,7	8,4	22,39	39,9	48	273	-53,7	-58,53	0,04	100,71	-34,64
Fabaceae	Cicer_arietinum	CARI_LNC003190.12 gene=CARI_LNC003190	0,113	1	0,83	319-733	30,9	6,5	19,42	42,8	49	414	-74,47	-83,9	0	138,21	-25,2
Malvaceae	Gossypium_barbadosense	GBAR_LNC014012.38 gene=GBAR_LNC014012	0,113	1	0,83	151-680	29,9	6,14	49,56	56,6	56	529	-98,3	-105,7	0	98,74	-83,9
Asteraceae	Lactuca_sativa	LSAT_LNC001553.1 gene=LSAT_LNC001553	0,113	0,9	0,79	191-447	29,7	4,17	48,6	4,4	53	256	-27,7	-34,83	0	91,75	-5,1
Solanaceae	Solanum_melongena	SMEL_LNC00097.18 gene=SMEL_LNC00097	0,114	1	0,85	404-714	31,3	8,4	22,6	37,7	46	310	-67,8	-73,95	0	119,22	-20,7
Solanaceae	Solanum_tuberosum	STUB_LNC005069.36 gene=STUB_LNC005069	0,114	1	0,85	172-586	31,2	7,5	21,40	40,1	48	414	-75,1	-84,22	0	154,43	-26
Cucurbitaceae	Cucumis_sativus	CSAT_LNC002832.1 gene=CSAT_LNC002832	0,115	1	0,85	322-723	34,2	7,8	22,34	34,2	43	401	-92,4	-100,4	0	102,5	-78,3
Fabaceae	Glycine_max	GMAX_LNC003740.2 gene=GMAX_LNC003740	0,116	1	0,84	30-521	30,5	9,21	39,4	41,1	48	491	-110,1	-120,7	0	167,8	-58,5
Malvaceae	Theobroma_cacao	TCAC_LNC004835.10 gene=TCAC_LNC004835	0,116	1	0,84	221-635	25,4	8,7	20,45	45,9	55	414	-87,8	-95,15	0	109,98	-47,3
Fabaceae	Trifoliumpratense	TPRA_LNC000558.12 gene=TPRA_LNC000558	0,117	1	0,83	235-647	32,8	6,19	41,9	41,3	47	412	-81,9	-90,82	0	105,38	-59,42
Malvaceae	Gossypium_raimondii	GRAI_LNC003490.2 gene=GRAI_LNC003490	0,118	1	0,84	24-627	29,3	10,24	36,7	36,7	46	603	-137	-147,8	0	173,53	-95,3
Fabaceae	Phaseolus_vulgaris	PVUL_LNC003538.18 gene=PVUL_LNC003538	0,118	1	0,82	308-826	28,2	5,8	23,42	42,3	48	518	-89,8	-99,13	0	139,86	-49,73
Amborellaceae	Amborella_trichopoda	ATRI_LNC004639.2 gene=ATRI_LNC004639	0,119	1	0,82	27-655	22,1	9,2	25,43	43,5	53	628	-128,7	-140	0	153,09	-93,8

Araliaceae	Panax_ginseng	PGIN_LNC002793.4 gene=PGIN_LNC002793	0,119	1	0,84	109- 752	30 ,5	10 ,3	25	34 ,2	45	643	-144,9	-157,1	0	187,6 3	-97,14
Rubiaceae	Coffea_arabica	CARA_LNC007267.26 gene=CARA_LNC007267	0,125	1	0,82	229- 724	30 ,1	8, 3	25 ,9	35 ,8	44	495	-104	-112,4	0	76,76	-97,7