



UNIVERSIDAD PERUANA
CAYETANO HEREDIA
ESCUELA DE POSGRADO

Búsqueda de genes asociados a la tolerancia
de heladas a partir del análisis de genotecas de
ADNc obtenidas por hibridación sustractiva
por supresión (SSH) de la papa nativa
Solanum juzepczuckii

TESIS PARA OPTAR EL GRADO DE
MAESTRO EN BIOQUÍMICA Y BIOLOGÍA
MOLECULAR

MÓNICA CAROLA CORREA GUERRERO

LIMA – PERÚ

2018

ASESOR DE TESIS: Luis Julio César De Stefano Beltrán

DEDICATORIA

*Este trabajo se lo dedico a mis padres, hermanas y sobrinos; por su amor, con el que
puedo contar siempre para seguir adelante.*

*De igual forma se lo dedico a Luis, por ser ejemplo de lucha contra la adversidad y
por todos esos valiosos consejos que me brindó y que me reconfortaron en los
momentos difíciles de este enredado camino.*

AGRADECIMIENTOS

- A mi asesor Luis De Stefano, por brindarme la oportunidad de realizar mi tesis en el Laboratorio de Genómica Funcional – LID - UPCH.
- A Pati y Antoine, por su valioso apoyo desinteresado para culminar esta tesis.
- A la Profesora Emma Urrunaga, por mostrarme cómo ser perseverante con los objetivos planteados.
- A mis amigos de la UPCH: Mari, Nora, Laura, Savina, Pamela y L. Serquén; por acompañarme con su valiosa amistad en toda esta etapa.

FUENTES DE FINANCIAMIENTO

Fondos del Consejo de Ciencia y Tecnología (CONCYTEC) en marco del proyecto:
USO DE GENOTECAS SUBTRACTIVAS DE ADN COMPLEMENTARIO
PARA EL DESCUBRIMIENTO DE GENES INVOLUCRADOS EN LA
TOLERANCIA A HELADAS EN PAPAS SILVESTRES Y PAPAS NATIVAS Y
SU APLICACIÓN EN EL MEJORAMIENTO PARA AFRONTAR LOS EFECTOS
CAUSADOS POR EL CAMBIO CLIMÁTICO.

RESUMEN

Las bajas temperaturas extremas son un serio problema para la producción de cultivos a nivel mundial. La papa, *Solanum tuberosum*, que se cultiva mayoritariamente para el consumo, es catalogada como una especie susceptible a heladas. Los genes expresados en la respuesta, ante bajas temperaturas, de la papa nativa resistente a heladas *Solanum juzepczukii* podrían ser aprovechados en programas de mejoramiento genético mediante el uso de herramientas biotecnológicas. Con el objetivo de identificar esos genes se construyeron dos genotecas de ADNc recíprocas con la metodología de la hibridación sustractiva por supresión (SSH) a partir de ARNm de hojas de plantas sometidas a tratamiento de temperaturas bajo cero comparándolos con los obtenidos de tratamiento a condiciones normales. Los ADNc de las genotecas fueron secuenciados con la tecnología 454 y aproximadamente el 80% de los *contigs* obtenidos de ambas genotecas presentaron algún término de ontología de gen (GO) asociado significativamente. A través de búsquedas de la información de su anotación, se encontraron genes asociados a respuestas de resistencia frente a algún estrés abiótico, entre ellos, los factores de transcripción de las familias WRKY, AP2/ERF y bHLH. La validación futura del rol de estos genes contribuiría a establecer candidatos para ser usados en programas de mejoramiento genético de la papa.

Palabras clave: Hibridación sustractiva por supresión (SSH), tolerancia a heladas, *Solanum juzepczukii*, secuenciamiento 454

ABSTRACT

Extreme low temperatures are a serious problem for crop production worldwide. The potato, *Solanum tuberosum*, which is grown mainly for consumption, is classified as a species susceptible to frost. *Solanum juzepczukii* is a native frost-resistant species. Genes expressed in response to low temperatures from *S. juzepczukii* could be used in breeding programs through the use of biotechnological tools. In order to identify these genes, two reciprocal cDNA libraries were constructed using the suppressive subtractive hybridization (SSH) methodology on leaf mRNA from *S. juzepczukii* with and without treatment of temperatures below zero. The cDNA libraries were sequenced using the 454 pyrosequencing technology. Gene ontology (GO) terms significantly associated to 80 % approximately of the contigs obtained of the libraries. Genes associated with resistance response to some abiotic stress were identified, based on their gene ontology information. Among them were identified transcription factors from the WRKY, AP2/ERF and bHLH families. Validating the function of these genes is recommended in order to establish a list of candidate genes to be used in potato breeding programs.

Key words: Suppressive Subtractive Hybridization (SSH), frost tolerance, *Solanum juzepczukii*, 454 pyrosequencing

TABLA DE CONTENIDOS

	Pág.
1 INTRODUCCIÓN.....	1
2 PLANTEAMIENTO DE LA INVESTIGACIÓN.....	4
2.1 Planteamiento del problema.....	4
2.2 Marco teórico.....	6
2.2.1 Las Heladas.....	6
2.2.2 La papa, <i>Solanum juzepczukii</i>	14
2.2.3 Estrés y estrés abiótico.....	18
2.2.4 Respuesta de la planta al estrés por bajas temperaturas.....	19
2.2.5 Hibridación sustractiva por supresión (SSH).....	22
2.2.6 Secuenciamiento GS FLX Titanium.....	26
2.2.7 Ontología génica (GO).....	29
2.2.8 Análisis de enriquecimiento de términos GO entre 2 grupos de secuencias.....	34
2.3 Justificación del estudio.....	35
2.4 Objetivos.....	36
3 METODOLOGÍA.....	37
3.1 Cultivo y tratamientos de las plantas.....	38

	Pág.
3.2 Extracción de ARN	43
3.3 Síntesis de ADNc y procedimiento de sustracción.....	44
3.4 Secuenciamiento y ensamblaje.....	49
3.5 Anotación funcional y análisis de enriquecimiento de términos de GO	50
3.6 Búsqueda de genes expresados diferencialmente.....	51
3.7 Selección de genes de factores de transcripción interesantes.....	51
4 RESULTADOS	53
4.1 Cultivo y tratamiento de las plantas	53
4.2 Extracción de ARN	54
4.3 Síntesis y digestión de ADNc.....	55
4.4 Hibridación sustractiva por supresión (SSH)	57
4.5 Secuenciamiento y ensamblaje.....	57
4.6 Anotación funcional	60
4.7 Enriquecimiento de términos GO - Prueba exacta de Fisher	69
4.8 Búsqueda de genes interesantes expresados diferencialmente.....	73
5 DISCUSIÓN	79
6 CONCLUSIONES	95
7 RECOMENDACIONES.....	96
8 REFERENCIAS BIBLIOGRÁFICAS	97

LISTA DE FIGURAS

	Pág.
Figura 1. Mapa global de bajas temperaturas y heladas.....	7
Figura 2. Tipos de heladas	9
Figura 3. Mapa de frecuencia de heladas, según datos del año 2000.....	12
Figura 4. Distribución altitudinal de los cultivos en relación a las regiones naturales	13
Figura 5. Representación esquemática de la hibridación sustractiva por supresión (SSH).....	24
Figura 6. Representación esquemática de la PCR supresora con secuencias terminales repetidas e invertidas.....	25
Figura 7. Esquema del pirosecuenciamiento.....	28
Figura 8. Ejemplo de ontología génica	33
Figura 9. Diagrama de flujo general de los pasos metodológicos.....	37
Figura 10. Cámara de crecimiento KBW 400 y plantas de la especie <i>Solanum juzepeczukii</i>	40
Figura 11. Esquema de las condiciones de tratamientos para la obtención de las muestras.....	41
Figura 12. Plantas crecidas en la cámara de cultivo KBW 400	54
Figura 13. Gel de electroforesis con muestras de ARN extraído de hojas de <i>Solanum juzepeczukii</i>	55
Figura 14. Gel de electroforesis con ADNc del procedimiento de la sustracción	56
Figura 15. Gel de electroforesis con productos de la PCR del procedimiento de la sustracción.....	58

	Pág.
Figura 16. Electroforesis en el equipo BioAnalyzer 2100	59
Figura 17. Distribución de la longitud de las secuencias ensambladas.....	60
Figura 18. Distribución del E-value luego del blastX.....	62
Figura 19. Distribución de las secuencias de acuerdo a la homología por especie.....	63
Figura 20. Distribución de las especies top-hit de las secuencias.....	64
Figura 21. Distribución del número de GO luego del mapeo de las genotecas	65
Figura 22. Gráficos circulares de la distribución de los datos luego del BlastX, mapeo y anotación	66
Figura 23. Top 20 de la distribución GO en 2° nivel	68
Figura 24. Términos de GO sobre-/sub representados en la categoría proceso biológico.....	70
Figura 25. Términos de GO sobre-/sub representados en la categoría función molecular.....	71
Figura 26. Términos de GO sobre-/sub representados en la categoría componente celular	72
Figura 27. Modelo propuesto de interacciones de WRKY26 y WRKY33	88
Figura 28. Modelo propuesto de interacciones de WRKY40 con otros FTs	89
Figura 29. Modelo propuesto de interacciones de CRF4 con genes DE, CBF y CORs	91

LISTA DE TABLAS

Pág.

Tabla 1. Intervalo de temperaturas críticas (°C) por estadio fenológico en algunos cultivos	10
Tabla 2. Especies de papas cultivadas consideradas por el CIP y su ploidía.....	15
Tabla 3. Programación en la cámara KBW 400 de los tratamientos	42
Tabla 4. Secuencias de los oligonucleótidos usados en la síntesis de ADNc por SMART y par su amplificación	44
Tabla 5. Componentes de la reacción de transcripción reversa	45
Tabla 6. Componentes de la reacción de la PCR con ADNc	46
Tabla 7. Secuencias de los oligonucleótidos usados en la preparación de las genotecas por SSH.....	47
Tabla 8. Componentes de la PCR primaria.....	48
Tabla 9. Componentes de la PCR secundaria	48
Tabla 10. Genes en la genoteca sentido anotados con el término GO: actividad de factor de transcripción.....	73
Tabla 11. Genes en la genoteca antisentido anotados con el término GO: actividad de factor de transcripción.....	75

1 INTRODUCCIÓN

El estrés abiótico; ocasionado principalmente por la sequía, la alta salinidad del suelo y las temperaturas extremas, es la principal causa de pérdida de la producción de cultivos alimenticios, llegando incluso a perderse hasta un 50% (Rodziewicz *et al.*, 2014). Frente al estrés ocasionado por las bajas temperaturas, las especies de plantas varían su habilidad para sobrevivir empleando distintos mecanismos para adaptarse; los cuales involucran respuestas coordinadas tanto a nivel morfológico, fisiológico, bioquímico como molecular (Wang *et al.*, 2013). Uno de esos mecanismos es la aclimatación al frío, que acontece cuando la tolerancia de la planta a la helada (temperaturas de congelamiento) aumenta luego de una previa exposición de la planta a bajas temperaturas (de no congelamiento) por un corto periodo de tiempo. Se ha comprobado que este fenómeno se manifiesta en varias especies de plantas, como por ejemplo *Arabidopsis thaliana* y *Brassica napus* (Chinnusamy *et al.*, 2010).

Estudios recientes han demostrado que los genes regulados por el frío contribuyen a la tolerancia a heladas en plantas. Por ejemplo, la ruta de respuesta al frío CBF (C-repeat binding factor) desencadenada en el proceso de aclimatación, es considerada como la ruta de tolerancia al frío mejor entendida hasta ahora. Tiene entre sus principales actores a CB1, CBF2 y CBF3 que codifican miembros de la familia de factores de transcripción (en adelante FTs) de AP2/ERF (APETALA 2/ethylene-responsive element binding factor). La sobreexpresión de cualquiera de estos genes,

por transgénesis, en ausencia de algún estímulo de frío, causa un incremento en la tolerancia al frío en plantas como *A. thaliana*, *B. napus*, álamos y papa (Knight *et al.*, 2004; Thomashow, 1999).

La papa, junto al trigo, arroz y maíz, es uno de los alimentos que más se cultivan y consumen en todo el mundo (FAO, 2008). La papa que se cultiva comercialmente es principalmente de la especie *Solanum tuberosum* y la helada está entre los principales factores adversos que afecta su cultivo (Hijmans y Spooner, 2001). Esta papa es clasificada como sensible a bajas temperaturas, debido a que sufre daño a -3 °C, y porque es incapaz de aclimatarse al frío (Chen y Li, 1980). Entre las principales consecuencias de la helada está el daño por congelamiento del follaje, lo cual repercute en la reducción del rendimiento del tubérculo (Chang *et al.*, 2014). Con el fin de aumentar la resistencia de la papa a factores ambientales adversos, las técnicas de mejoramiento pueden hacer uso de los recursos genéticos de las variedades existentes de las papas, que son altamente diversas y cultivadas en un largo rango de condiciones agroecológicas. De particular interés, son las variedades nativas andinas, que por su crecimiento y evolución en condiciones secas y frías del altiplano han desarrollado una alta tolerancia al frío (Hijmans y Spooner, 2001; Ritter *et al.*, 2008). Sin embargo, en los esfuerzos de aprovechar esos recursos genéticos con las técnicas tradicionales, como la técnica de introgresión, se han presentado grandes dificultades en el entrecruzamiento, como por ejemplo la incompatibilidad del número de balance de endospermo (EBNs, por sus siglas en inglés) o el arrastre de caracteres no deseados (por ejemplo, alto contenido de alcaloides y largos estolones) debido a la complicada herencia que presenta la papa (Distl y Wink, 2009; Kikuchi *et al.*, 2015).

Las variedades nativas de papa pueden proporcionar genes que hagan frente no solo a las heladas, si no a otros factores de estrés abiótico. En consecuencia, en la era de los estudios genómicos y transcriptómicos, se hace necesario identificarlos y caracterizarlos para evaluar primero su mecanismo de acción y darles, luego, utilidad con las herramientas biotecnológicas que complementen a las tradicionales. Entre las variedades nativas, se encuentran las papas cultivadas localmente, como las papas amargas, que en su mayoría pertenecen a la especie *Solanum juzepczuckii* (Condori *et al.*, 2014).

En el presente trabajo, se estudió a esta papa nativa debido a su comprobada tolerancia a heladas -puede llegar a soportar hasta los -5 °C (Hijmans y Spooner, 2001). Para tal fin, se construyeron genotecas por hibridación sustractiva por supresión (SSH, por sus siglas en inglés) (Diatchenko *et al.*, 1996), que presentaron secuencias de genes expresados diferencialmente luego del tratamiento de heladas. Los ADNc obtenidos fueron secuenciados y luego del alineamiento se obtuvieron secuencias consenso que se sometieron al análisis *in silico* que permitió la identificación y selección de genes interesantes. Mediante bibliografía se analizó si estos genes han sido asociados previamente a la tolerancia a algún estrés abiótico, de preferencia osmótico en alguna especie de planta.

2 PLANTEAMIENTO DE LA INVESTIGACIÓN

2.1 Planteamiento del problema

Se calcula que para abastecer a la población mundial se debe incrementar considerablemente la producción de alimentos. Este reto será enfrentado con una ampliación de la frontera agrícola y con un crecimiento revolucionario de la productividad. En nuestro país debido al cambio climático y a las limitaciones que nuestra geografía impone sobre la agricultura lo más factible es que este incremento en la producción se logre por una mejora sustancial de la productividad y no por una significativa expansión en la frontera agrícola (por lo menos no por encima de los 2,500 m.s.n.m). En nuestro país, para aumentar la productividad de algún cultivo, en general, se recurre a las prácticas agronómicas tradicionales como por ejemplo la introgresión de caracteres por métodos clásicos. Estos involucran el movimiento de genes de una planta a otra a través de cruzamientos dirigidos, con la finalidad de dotar a la planta de mayor resistencia a los factores adversos que puedan afectar su desarrollo, como son los factores abióticos (sequías, heladas, alta irradiación lumínica, etc.) y los factores bióticos (bacterias, hongos, virus, nemátodos, etc.).

La intensidad con la que se presentan los factores abióticos va en aumento debido al cambio climático. Existe un gran interés por la búsqueda de variedades de cultivos que cuenten con los recursos genéticos necesarios para afrontar las condiciones extremas del clima. Las especies de plantas que logran crecer en ambientes extremos, cuentan con esos recursos, debido a su adaptación en la evolución.

Existe toda una corriente de investigación en la exploración de los genes implicados en la respuesta a factores abióticos en plantas resistentes a las condiciones extremas del clima. Se han dado grandes aportes en este tema a partir de la utilización de técnicas moleculares para comprender los mecanismos subyacentes. Esta información ha permitido inducir, en el laboratorio, el aumento de resistencia al frío por transgénesis a plantas que no presentaban ese nivel de resistencia. Por lo tanto, se plantea que aún queda por descubrir valiosa información de la resistencia en plantas, sobre todo proveniente de los cultivos de alimentación, que será de mucha utilidad para afrontar el desabastecimiento a causa del cambio climático.

La papa es uno de los cultivos más importantes en la producción alimenticia mundial. La mayor parte de la papa que se consume pertenece a *Solanum tuberosum*. Esta especie presenta poca tolerancia a las bajas temperaturas, como consecuencia, en nuestro país su cultivo se ve muy afectado por las heladas que se presentan anualmente entre los meses de mayo a setiembre llegándose a perder gran parte de la producción, dejando pérdidas económicas considerables a los agricultores de las zonas alto-andinas. Este problema se viene incrementado debido a que las heladas se están prolongando, llegando a permanecer incluso hasta el mes de octubre. Urge por lo tanto, contar con variedades mejoradas de papa que a su vez dependan de un *pool* de genes que brinden las características adecuadas de resistencia. Es así, que se revela el siguiente problema: existe un desconocimiento de los genes implicados en el mecanismo de resistencia a heladas y por lo tanto desaprovechamiento de esa biodiversidad de genes que podrían aportar hacia la implementación de técnicas más eficaces que permitan obtener las variedades de papa que se necesita.

2.2 Marco teórico

2.2.1 Las Heladas

Una helada es la ocurrencia de una temperatura del aire de 0 °C o inferior, medida a una altura entre 1,25 y 2,0 m por encima del nivel del suelo, dentro de una garita meteorológica adecuada. A este término usualmente se le confunde con el de congelación, que más bien describe el fenómeno cuando el agua extracelular dentro de la planta se congela (i.e. cambia de líquido a hielo). Por lo tanto, un evento de helada puede producir un evento de congelación (Snyder *et al.*, 2010). En la figura 1 se muestra un esquema general de las temperaturas de heladas a nivel global. Se observan regiones que experimentan episódicamente heladas por debajo de -10 °C, las que se extienden por Australia, sur de África, sur de EE.UU., partes de Sudamérica y Asia. También, regiones con una temperatura mínima anual entre -10 °C y -40 °C que son menos extensivas y que se presentan por grandes áreas del hemisferio norte. (Beerling *et al.*, 2002).

Las heladas pueden clasificarse por su origen climatológico, época de ocurrencia o aspecto visual (Matías-Ramírez *et al.*, 2012).

Por su origen climatológico las heladas pueden ser:

- Heladas por advección, la temperatura disminuye cuando llegan grandes masas de aire frío para reemplazar el aire más caliente que estaba presente. Está asociada con condiciones de nubosidad, vientos de moderados a fuertes, cuando no hay inversión de temperatura (i.e no aumenta la temperatura con la

altitud) y hay baja humedad. La temperatura puede ser inferior a 0 °C durante el día.

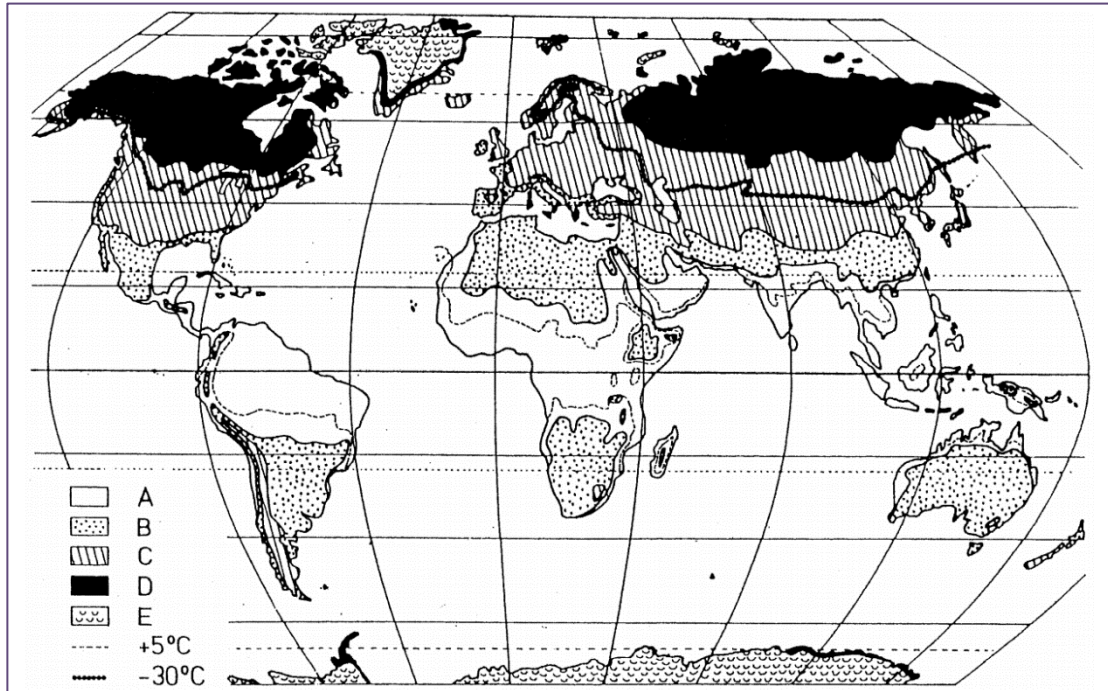


Figura 1. Mapa global de bajas temperaturas y heladas

- A. Regiones libres de helada; B. Regiones con helada episódica hasta -10 °C; C. Regiones con promedio anual mínimo entre -10 °C y -40 °C; D. Regiones con promedio anual mínimo por debajo de -40 °C y E. Regiones de hielo permanente. Fuente: Beerling, *et al.*, 2002

- Heladas por radiación, se dan cuando disminuye la temperatura cerca al suelo debido al traspaso de calor del suelo al aire mediante el fenómeno conocido como radiación. Hay mayor radiación del suelo al aire por las noches debido a que se presenta cielos despejados. Estas heladas se caracterizan por acontecer cuando existe poco viento y con inversión de temperatura (i.e. la temperatura aumenta con la altura). Las temperaturas del aire normalmente caen por debajo de 0 °C durante la noche pero están por encima de 0 °C durante el día.

- Heladas por evaporación, la temperatura disminuye por las evaporaciones intensas del agua que se halla en las plantas, esto se produce en condiciones de baja humedad relativa y en presencia de corrientes de aire seco. Este tipo de heladas son poco frecuentes.

Por la época de su ocurrencia las heladas pueden ser heladas primaverales, otoñales, e invernales.

Por su efecto visual. Esta clasificación se basa por la apariencia de los cultivos expuestos a las heladas. En las heladas blancas se observa la formación de capas de hielo sobre las plantas, comúnmente denominadas escarchas, por eso el nombre. Este tipo de heladas se producen de noche cuando hay alta humedad y un intenso enfriamiento que produce la condensación del vapor de agua del ambiente, a la vez de su congelamiento. En las heladas negras no hay formación de escarchas porque la humedad es bastante baja, por lo que el frío daña directamente a las plantas, haciendo que presenten una coloración negruzca como una de las consecuencias de la congelación del agua de los tejidos, en consecuencia los efectos de estas heladas son más dañinas que las blancas. En la figura 2 se muestran los tipos de heladas, así como las posibles combinaciones de tipos de heladas con diferente criterio.

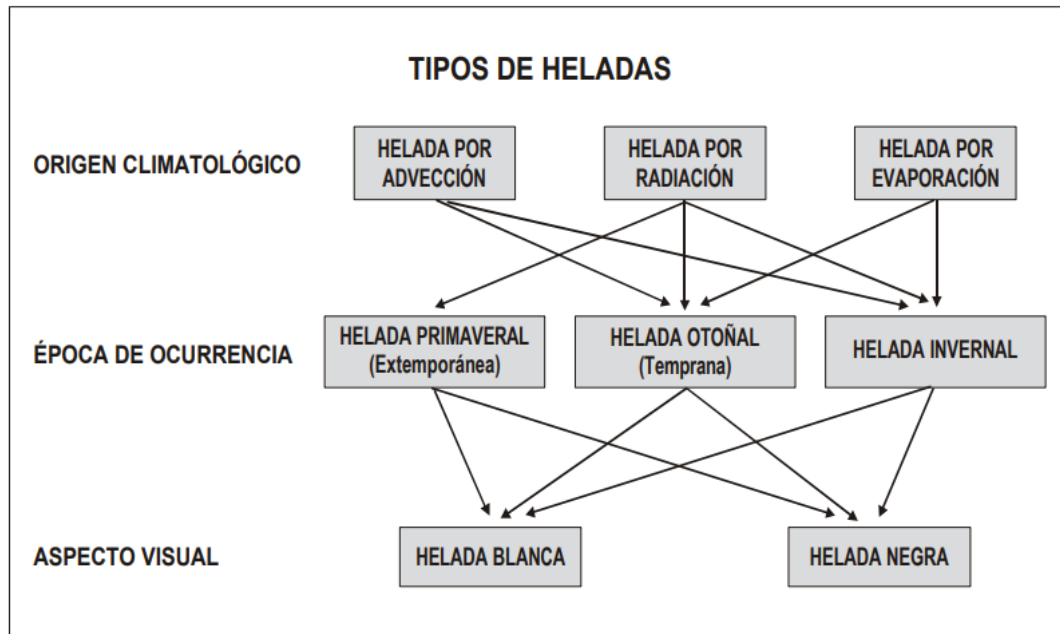


Figura 2. Tipos de heladas
Fuente: Matías-Ramírez, L. G., 2012

La temperatura correspondiente a un nivel específico de daño en una especie se denomina “Temperatura crítica” o “Temperatura crítica de daño”, y se expresa por el símbolo T_c . No todos los cultivos se dañan con la misma intensidad, ya que depende de la especie y el estadio de la planta (en general los estadios más sensibles se presentan después del desarrollo inicial de las yemas hasta el estadio de fruto pequeño), es más, los órganos de una misma planta no se dañan con la misma intensidad. En la Tabla 1 se muestra una lista de cultivos con sus temperaturas críticas relacionados con sus estadios fenológicos (Snyder *et al.*, 2010).

Tabla 1. Intervalo de temperaturas críticas (°C) por estadio fenológico en algunos cultivos

CULTIVO	GERMINACION	FLORACION	FRUCTIFICACION
Trigo de primavera	-9; -10	-1; -2	-2; -4
Avenas	-8; -9	-1; -2	-2; -4
Cebada	-7; -8	-1; -2	-2; -4
Guisantes	-7; -8	-2; -3	-3; -4
Lentejas	-7; -8	-2; -3	-2; -4
Almorta	-7; -8	-2; -3	-2; -4
Cilantro (Coriander)	-8; -10	-2; -3	-3; -4
Amapolas	-7; -10	-2; -3	-2; -3
Diente de león	-8; -10	-3; -4	-3; -4
Altramuz	-6; -8	-3; -4	-3; -4
Veza de primavera	-6; -7	-3; -4	-2; -4
Judías	-5; -6	-2; -3	-3; -4
Girasol	-5; -6	-2; -3	-2; -3
Cártamo	-4; -6	-2; -3	-3; -4
Mostaza blanca	-4; -6	-2; -3	-3; -4
Lino	-5; -7	-2; -3	-2; -4
Cáñamo	-5; -7	-2; -3	-2; -4
Remolacha azucarera	-6; -7	-2; -3	-
Remolacha pienso	-6; -7	-	-
Zanahoria	-6; -7	-	-
Nabo	-6; -7	-	-
Col	-5; -7	-2; -3	-6; -9
Soja	-3; -4	-2; -3	-2; -3
Mijo de Italia	-3; -4	-1; -2	-2; -3
Altramuz amarillo	-4; -5	-2; -3	-
Maíz	-2; -3	-1; -2	-2; -3
Mijo	-2; -3	-1; -2	-2; -3
Pasto del Sudan	-2; -3	-1; -2	-2; -3
Sorgo	-2; -3	-1; -2	-2; -3
Patatas	-2; -3	-1; -2	-1; -2
Tabaco rústico	-2; -3	-	-2; -3
Trigo sarraceno	-1; -2	-1; -2	-0,5; -2
Castor (higuerillo)	-1; -1,5	-0,5; -1	-2
Algodón	-1; -2	-1; -2	-2; -3
Melón	-0,5; -1	-0,5; -1	-1
Arroz	-0,5; -1	-0,5; -1	-0,5; -1
Sésamo	-0,5; -1	-0,5; -1	-
Kenaf	-0,5; -1	-	-
Cacahuete	-0,5; -1	-	-
Pepino	-0,5; -1	-	-
Tomate	0; -1	0; -1	0; -1
Tabaco	0; -1	0; -1	0; -1

Fuente: Snyder *et al.*, 2010

Se pueden describir dos tipos de daño causados por las heladas a nivel celular, el primero es un daño directo causado por la baja temperatura, en la que se afecta por ejemplo los componentes o procesos celulares como la fluidez de ácidos grasos de las membranas celulares y la actividad enzimática. El otro tipo de daño, el indirecto, se da por la formación de cristales de hielo y su expansión en los compartimentos extracelulares que propicia la concentración de solutos, debido a que el agua se moviliza desde adentro hacia afuera de la célula, por lo que los solutos y componentes celulares adquieren una alta concentración teniendo como consecuencia una deshidratación celular (Di Toppi y Pawlik-Skowrońska, 2003). La deshidratación celular también se produce por otros tipos de estrés como la sequía y la alta salinidad, por lo tanto, no es raro suponer que los componentes de la respuesta de las plantas sean compartidos entre estos tres tipos de estrés (Hirt y Shinozaki, 2003).

2.2.1.1 Las heladas en el Perú

El altiplano peruano se sitúa entre los 3800 a 4500 m.s.n.m., la altitud modifica el clima tornándolo frío, debido a que la gradiente de aire húmedo disminuye 0.6 °C por cada 100 m de altitud y por lo tanto, la temperatura promedio del altiplano está entre 8 a 9 °C. Las heladas son recurrentes en el altiplano, y en su mayoría son causadas por radiación, por lo que el aire se enfría rápidamente por las noches despejadas y la temperatura ambiente desciende a niveles inferiores a los 0 °C (Fairlie y Ortega, 2016).

En la figura 3 se muestra un mapa de frecuencia de heladas en el Perú. Se observa una mayor ocurrencia en la zona sur central, en el que zonas como Cusco, Puno,

Arequipa y Cerro de Pasco, son las más afectadas. En junio del año 2016, las temperaturas más bajas se registraron en las estaciones de Chuapalca (-20.5 °C) y Vilacota (-18.8 °C), en el departamento de Tacna, así como en Mazo Cruz (-18.4 °C), en Puno, e Imata (-15,8 °C) y Caylloma (-10.2 °C), en Arequipa (INEI, 2016).

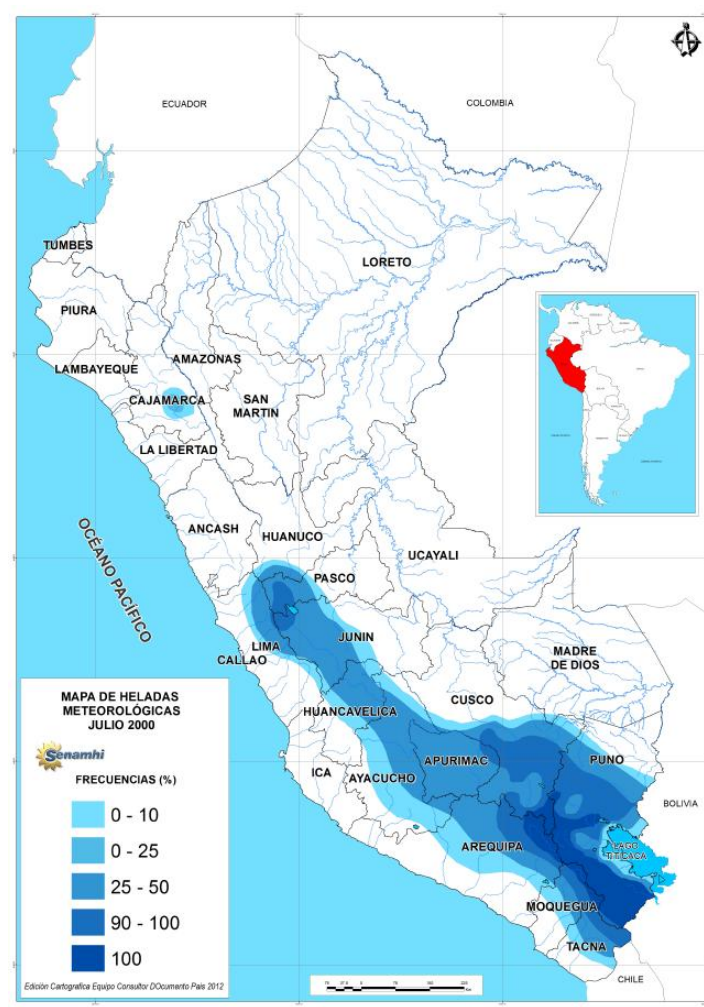


Figura 3. Mapa de frecuencia de heladas, según datos del año 2000

Fuente: <http://www.predes.org.pe/mapas/heladas/>

La cordillera de los Andes es el principal factor que condiciona los distintos cultivos en relación a la temperatura. Las actividades agrícolas en la sierra del Perú se desarrollan principalmente desde los 500 a 3800 m.s.n.m. Para hacerse una idea del

alcance de las heladas en la agricultura se ha establecido un perfil de la distribución altitudinal de los cultivos agrícolas en base a las ocho regiones naturales se muestra en la figura 4. La presentación de los cultivos en la figura se hace a manera de productos límite, es decir, se tiene en cuenta aquellas especies vegetales que no pueden ser cultivadas en la región más alta, aun cuando puedan producirse en las regiones más bajas (SENHAMI, 2010). Se observa por ejemplo que el cultivo de la papa al establecerse en la zona altiplánica se encuentra afectado por la ocurrencia de las heladas y es que se ha registrado que el 85 a 90 % del área sembrada de papa en la región altiplánica está expuesta a heladas (Arcos *et al.*, 2016).

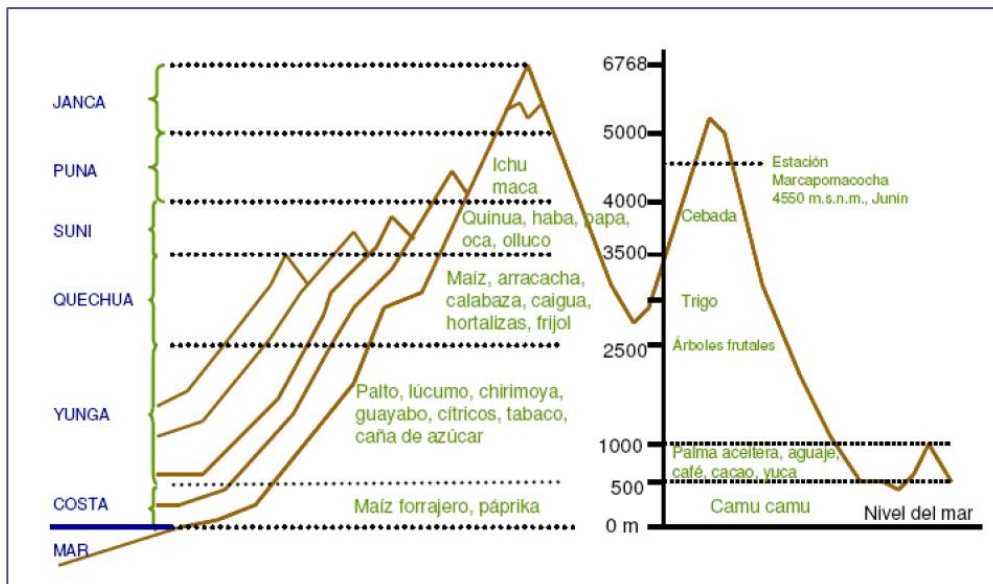


Figura 4. Distribución altitudinal de los cultivos en relación a las regiones naturales
Fuente: SENHAMI, 2010.

Las heladas en el Perú se registran con mayor intensidad durante el periodo de mayo a septiembre y el grado de afectación en los cultivos depende de los estados fenológicos en los que se encuentran. En la sierra, se dan dos tipos de campañas de agricultura; una de ellas, es denominada “campaña agrícola grande”, que se inicia en

agosto, en seco y presentan la mayor área de cultivos; la segunda es denominada “campana agrícola chica”, que se inicia en zonas donde no se dispone de agua para riego. La temporada de cosecha de la campana grande se da en el mes de mayo, por ello la afectación de los cultivos se relaciona a aquellos que no fueron recogidos antes de la helada y que están en su etapa de floración – maduración. Las heladas afectan a los de campana chica en estado de plántulas, periodo de mayor vulnerabilidad de los cultivos (PMC, 2012).

2.2.2 La papa, *Solanum juzepczukii*

La papa ocupa el cuarto lugar mundial en importancia como alimento, después del maíz, el trigo y el arroz (FAO 2008a). El número de especies de la papa cultivada sigue siendo controversial. Spooner *et al.*, (2007) sugieren la reclasificación de las papas cultivadas en 4 especies, en base al genotipaje con marcadores moleculares de 742 variedades nativas de todas las especies cultivadas y sus progenitores silvestres. Las cuatro especies que proponen son: 1) *S. tuberosum*, con dos grupos de cultivares (el grupo Andigenum que presenta genotipos altiplánicos que pueden ser diploides, triploides y tetraploides) y el grupo Chilotanum con variedades locales chilenas tetraploides; 2) *S. ajanhuiri* (diploide); 3) *S. juzepczukii* (triploide); y 4) *S. curtilobum* (pentaploide).

Las diferentes bases de datos de cultivares de papa utilizan la clasificación que considera el Centro Internacional de la Papa (CIP) (Tabla 2), en la que se tiene en cuenta 8 especies cultivadas, estas son: *S. tuberosum* con las subespecies *tuberosum* y

andigena, *S. goniocalyx*, *S. stenotomum*, *S. chaucha*, *S. phureja*, *S. juzepczukii*, *S. curtilobum* y *S. ajanhuiri* (Machida-Hirano, 2015; CIP y FEDECH, 2006).

Tabla 2. Especies de papas cultivadas consideradas por el CIP y su ploidía

Especie	Subespecie	Ploidía
<i>Solanum tuberosum</i>	<i>tuberosum</i>	2n=4x=48
	<i>andigena</i>	2n=4x=48
<i>Solanum goniocalyx</i>	-	2n=2x=24
<i>Solanum stenotomum</i>	-	2n=2x=24
<i>Solanum chaucha</i>	-	2n=3x=36
<i>Solanum phureja</i>	-	2n=2x=24
<i>Solanum juzepczukii</i>	-	2n=3x=36
<i>Solanum curtilobum</i>	-	2n=5x=60
<i>Solanum ajanhuiri</i>	-	2n=2x=24

Fuente: CIP y FEDECH, 2006

La papa que se consume en su mayoría está representada por *S. tuberosum*. La variedad más cultivada en los Andes es *Solanum tuberosum* subsp. *andígena*. (Condori, *et al.*, 2014). El potencial impacto del cambio climático en la producción de papa es crítico para la futura seguridad alimentaria, y las heladas es uno de los principales factores que afecta su producción (Raymundo, 2017).

La especie *S. tuberosum* es sensible a temperaturas inferiores a -2 °C, ya que a esa temperatura se observan daños visibles en las hojas, disminución en la producción, así como en la calidad de los tubérculos (Hijmans *et al.*, 2003). Se ha reportado que la máxima tolerancia a temperaturas bajas a la que llega esta especie es -3 °C. Por tal

motivo, durante mucho tiempo a nivel mundial, se realizan numerosos esfuerzos por dotar de una mejor resistencia a *S. tuberosum* a través de distintas técnicas. Además se ha puesto énfasis en la investigación de especies con alta tolerancia al frío. Entre ellas se ha reportado como las más tolerantes a las especies silvestres *S. acaule* y *S. commersonii*, que pueden llegar a tener una tolerancia de hasta $-9\text{ }^{\circ}\text{C}$ y $-11.5\text{ }^{\circ}\text{C}$, respectivamente (Chen y Li, 1980). Las variedades nativas también constituyen una fuente considerable de características agronómicas útiles para afrontar las bajas temperaturas que afectan el cultivo de la papa, debido a que son el resultado de un proceso de domesticación, selección y conservación ancestral. La alta resistencia de estas variedades se debe a la adaptación que han adquirido al ser cultivadas a gran altitud (Hijmans *et al.*, 2003).

A más de 4000 m.s.n.m. la tierra está cubierta principalmente por pastizales naturales y solo se la utiliza para el cultivo de variedades locales de papa que son denominadas papas amargas (Condori *et al.*, 2014). Los tubérculos de estas variedades tienden a ser amargas (de ahí el nombre), debido a los altos niveles de glicoalcaloides que presentan, por lo tanto, requieren un procesamiento antes del consumo (Hijmans, *et al.*, 2003). Para este fin, las poblaciones que habitan en las zonas altoandinas utilizan los procesos tradicionales, que consiste en la exposición de los tubérculos a varias heladas nocturnas y luego al intenso sol, para luego ser pelados, hidratados y secados. De esta forma se obtiene el chuño negro. Las papas amargas de mayor tamaño se destinan de preferencia a la elaboración del chuño blanco, llamado también tunta (aymara) o moraya (quechua) (FAO, 2008b). Estos productos alimenticios deshidratados constituyen una importante fuente de carbohidratos y calorías; sobre

todo en los años de escasez de alimentos, y es una de las formas de conservación de la papa para períodos largos de tiempo y que ha sido mantenida desde hace siglos (Arcos *et al.*, 2016).

Las papas amargas pertenecen principalmente a dos especies: la triploide *S. juzepczukii* y la pentaploide *S. curtilobum*. Aunque el área cultivada de papas amargas es variable, se estima que existen unas 15000 ha en el Perú. La época de siembra se extiende desde septiembre hasta noviembre según el inicio temprano o tardío de las lluvias (FAO, 2008b). Se estima que al menos el 25% de la superficie total plantada con papa en el altiplano es con papas amargas, de las cuales 60% son *S. juzepczukii*, y 33% son *S. curtilobum* (Hijmans *et al.*, 2003).

S. juzepczukii son papas estériles por su ploidía impar ($2n = 3x = 36$), su origen se encuentra en el altiplano por hibridación natural de la especie cultivada diploide *S. stenotomum* ($2n = 2x = 24$) y la especie altoandina silvestre *S. acaule* ($2n = 4x = 48$). La primera especie aportó importantes características como la alta tuberización y rendimiento mientras que la segunda aportó con los genes de alta resistencia a heladas y a los virus X e Y de la papa (PVX y PVY, por sus siglas en inglés) (Chávez, 2004). Está documentado que *S. juzepczukii* representa una alta resistencia a las heladas, generalmente resiste temperaturas de hasta $-5\text{ }^{\circ}\text{C}$ (Hijmans *et al.*, 2003).

La resistencia a las heladas en varias especies tuberíferas se las asocia con algunos caracteres, entre ellos morfológicos, como: menor tamaño de células estomáticas, paredes celulares gruesas, menor área de estomas, hábito arrossetado de la planta, dos o más capas en empalizada en el tejido parenquimatoso de las hojas; factores fisicoquímicos como: el contenido de lípidos en las células, capacidad de

transformación de almidones hacia azúcares, capacidad de formar hielo extracelular, pigmentos protectores, habilidad para soportar deshidratación celular y mantener la cohesión bajo la presión de congelación. En distintos trabajos se ha establecido por ejemplo la alta resistencia de *S. juzepzuckii* con la triple capa en empalizada en las células del parénquima que presenta (Gabriel *et al.*, 2009).

2.2.3 Estrés y estrés abiótico

El estrés es un estado en el que la planta crece en condiciones no ideales, por lo que la planta experimenta un cambio que demanda respuesta. Si el estrés está dentro de los límites de tolerancia de la planta, la respuesta de la planta puede restablecer el estado normal, pero si el estrés excede la capacidad de tolerancia de la planta, ocurre un daño permanente que puede llegar hasta la muerte (Mosa *et al.*, 2017; Shao *et al.*, 2009). El estrés también puede ser definido como cualquier condición desfavorable o sustancia que afecta o bloquea el metabolismo de la planta afectando su crecimiento y/o desarrollo (Kranter *et al.*, 2010).

El estrés puede ser clasificado, de acuerdo al tipo de factor que lo causa, en dos: estrés biótico y estrés abiótico. El primero, es causado por los otros organismos que comparten el ambiente e interactúan con las plantas, tales como patógenos. El estrés abiótico es ocasionado por los diferentes factores ambientales tales como la luz, agua y temperatura. El estrés abiótico, por lo tanto, es el resultado de cambios de factores no biológicos, siendo cambios ambientales o nutricionales. Entre los mayores factores de estrés abiótico se encuentran: la sequía, la alta salinidad, las temperaturas extremas

(de congelamiento y altas temperaturas). Estos tipos de estrés abiótico conllevan a la reducción de la disponibilidad de agua, por lo que los tres se les agrupan como estrés osmótico, que está caracterizado por disminución de la turgencia y pérdida de agua (Duque *et al.*, 2013).

2.2.4 Respuesta de la planta al estrés por bajas temperaturas

A diferencia de los animales, las plantas no pueden moverse a ambientes más favorables como respuesta a cambios extremos del ambiente, por lo que deben tolerarlo mediante cambios internos efectivos que le permitan rápidamente percibir y responder (Boyer, 1982). Las plantas emplean mecanismos regulatorios que implican cambios morfológicos, fisiológicos, bioquímicos y moleculares, para activar respuestas efectivas contra el estrés. Las fitohormonas juegan un rol central en la señalización del estrés. El ácido salicílico (AS), ácido jasmónico (AJ) y etileno (ET) tienen roles centrales en la señalización del estrés biótico; y el ácido abscísico (ABA) en cambio está involucrado en la respuesta a estrés abiótico (Wasilewska *et al.*, 2008). Un ejemplo de respuesta fisiológica se presenta cuando el estrés osmótico promueve la síntesis de ABA, que activa cambios principalmente en hojas. Los niveles incrementados de ABA aumentan la respuesta a estrés por agua e inducen el cierre de estoma, lo cual resulta en una reducción de pérdida de agua a través de la transpiración, decayendo así la tasa de fotosíntesis. Estos cambios tienen un efecto a corto plazo como mejorar la eficiencia de la utilización del agua por la planta. Luego,

cuando los niveles de ABA caen, los estomas se abren para disponer del agua nuevamente (Mosa *et al.*, 2017).

Dentro de la célula se activan rutas de señalización múltiples frente al estrés, que presentan interacciones complejas. El mecanismo molecular implicado está lejos de ser entendido, sobre todo acerca de los puntos convergentes de las diferentes rutas. Sin embargo, en términos generales, la red de respuesta al estrés por frío en plantas empieza con la percepción de la señal, donde el estrés abiótico es “percibido” al modificarse las propiedades físicas de la membrana (mediante la composición de lípidos, composición de ácidos grasos, etc.) donde la disminución de temperatura produce cambios conformacionales de la membrana haciéndola más rígida. Estas modificaciones inducen luego cambios en las reacciones metabólicas. La señal es más tarde transducida mediante segundos mensajeros. El calcio y las especies reactivas de oxígeno (ROS, por sus siglas en inglés) son los principales segundos mensajeros en la temprana respuesta al estrés abiótico. Los niveles del calcio citosólico se incrementan, lo que hace que distintas rutas simultáneamente se activen a través de proteínas que interactúan con calcio tales como proteínas quinasas dependientes de calcio (CDPKs), calcineurina, proteína de unión a calcineurina B (CBL). Estos “sensores de calcio” interactúan con otras proteínas, en la que la modulación final resulta en cascadas de fosforilación dirigidas a genes implicados en la respuesta al estrés y a sus factores que los regulan (Xiong y Zhu, 2001; Mahajan y Tuteja, 2005; Fraire-Velázquez *et al.*, 2011).

Se ha establecido que la regulación de la expresión génica ante la respuesta a estrés osmótico es mantenida por dos rutas, designadas de acuerdo a su relación con ABA;

la primera es una ruta dependiente en la cual elementos de respuesta de ABA (ABRE), proteínas de unión a ABRE (AREB) o factores de unión ABRE (ABFs), así como los elementos que actúan de forma *cis* son los principales componentes. La segunda es independiente de ABA, en el que Dehydration-Responsive Element/C-Repeat (DRE/CRT), DREB (Dehydration-Responsive Element-Binding Protein), son los principales componentes (Yoshida *et al.*, 2014).

Existen plantas tolerantes al frío y otras que aumentan su tolerancia al frío, las cuales primero pasaron por un proceso llamado “aclimatación al frío”, que se define como el mecanismo por el cual ciertas plantas incrementan su tolerancia a heladas en respuesta a una previa y gradual exposición de bajas temperaturas de no congelamiento (mayor a 0 ° C) (Fowler y Thomashow, 2002).

La ruta de regulación génica en la respuesta al frío, que comprende a los componentes DRE/CRT, es la más estudiada y caracterizada y se armó a partir de la identificación de varios genes que participan en condiciones de aclimatación en *Arabidopsis thaliana* (Shinozaki *et.al*, 2009; Chinnusamy *et.al*, 2007). El estudio de la regulación de estos genes permitieron descubrir a los CBFs (C – repeat Binding Factors) o también conocidos como DREBS (dehydration responsive element binding factors). *Arabidopsis thaliana* codifica 3 genes CBFs inducibles por el frío: CBF1, CBF2, CBF3; también referidos como DREB1b, DREB1c, y DREB1a respectivamente (Liu *et al.*, 1998) estos genes codifican FTs que son miembros de la familia de proteínas de unión al ADN AP2/ERF. Estas proteínas se unen a los elementos regulatorios CRT/DRE presentes en los promotores de los genes diana (que son referidos como “regulones CBF”) regulando su transcripción. Estudios de sobreexpresión de AtCBF1

o AtCBF3 inducidos por el estrés en plantas transgénicas demuestran que aumentan la tolerancia a congelamiento, salinidad y sequía en Brassica, tomate, tabaco, trigo y arroz (Welling y Palva, 2008; Jaglo *et al.*, 2001).

Estudios recientes han explorado los actores moleculares y han logrado identificar nuevos genes, incluyendo quinasas y FTs, los que están involucrados en el *crossstalk* entre las cascadas de señalización involucrados en las respuestas contra dos o más tipos de estrés (Mosa *et al.*, 2017).

Uno de los principales objetivos de la investigación relacionado a la agricultura es entender cómo la planta tolera el estrés debido al impacto económico asociado. Herramientas moleculares poderosas que incluyen el transcriptoma, análisis proteómico, secuenciamiento de genomas enteros, análisis bioinformático y estudios funcionales, permiten la identificación de nuevos componentes (Fraire-Velázquez, S., *et al.*, 2011).

2.2.5 Hibridación sustractiva por supresión (SSH)

La hibridación sustractiva por supresión (SSH, por sus siglas en inglés), es una tecnología basada en PCR que amplifica selectivamente ADNc, proveniente de ARNm expresados diferencialmente en una de las dos muestras contrastantes que se desea comparar. La muestra blanco, es decir de donde se desea obtener la información de los ADNc, es denominada *tester* y la muestra control, *driver*. Esta tecnología combina los pasos de normalización y de supresión en uno solo, en la que se equilibra la abundancia de los ADNc en la población *tester*, permitiendo la

representatividad de ADNc raros, además se excluye las secuencias comunes entre la población *tester* y *driver*. Una representación esquemática de la tecnología se muestra en la figura 5.

Para utilizar la tecnología, ADNc se sintetiza a partir de ARNm extraído de las muestras *tester* y *driver*. Luego, este ADNc se digiere con enzimas de restricción para generar extremos romos, lo que permitirá ligar adaptadores que se unirán solo al terminal 5' del ADNc. Para comenzar con el procedimiento de la SSH propiamente dicho, el ADNc *tester* se divide en 2 porciones; a cada una de ellas se le liga adaptadores diferentes (denominados adaptador 1 y adaptador 2R. A continuación se lleva a cabo dos tipos de hibridaciones. En la primera, se realizará una desnaturalización y se hibridarán ambas porciones, por separado, con ADNc *driver* que estará en exceso. En este paso ocurre la normalización de los ADNc de baja copia con los de alta copia, debido a que las hebras más abundantes se reasocian con mayor velocidad (por la 2da ley de la cinética de hibridación). Estas moléculas están representadas como moléculas b en la figura 5. De esta hibridación se formarán, en total, los productos "a", "b", "c" y "d" en cada porción.

Se realiza luego, una segunda hibridación en la que ahora ambas porciones son hibridadas entre sí, sin desnaturalizar, en la que hebras simples provenientes de cada porción se hibridarán generando moléculas con un adaptador diferente en cada extremo (hebras provenientes del ADNc *tester*), aquí se origina el producto "e".

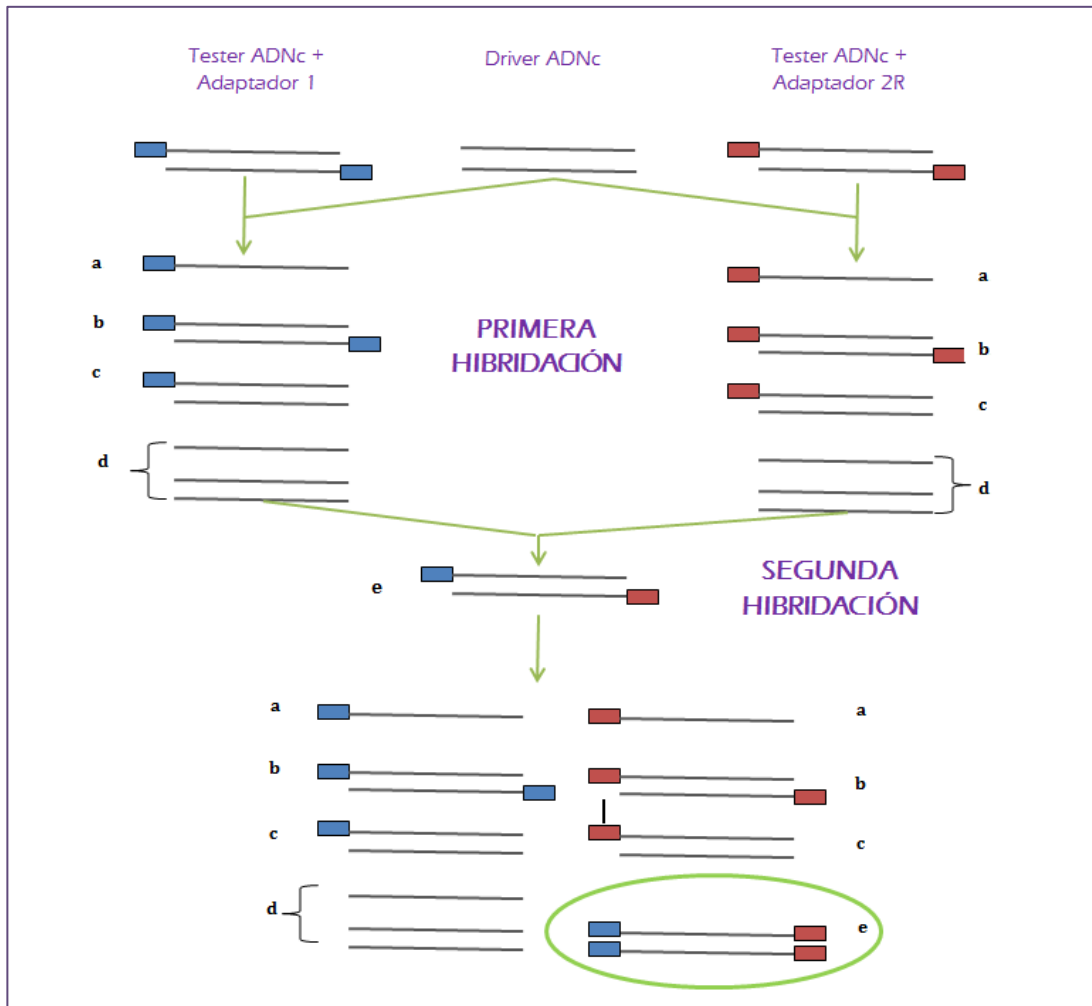


Figura 5. Representación esquemática de la hibridación sustractiva por supresión (SSH)

Luego, se da la polimerización de las hebras para dar lugar a la PCR supresora, en la que los cebadores anidados se complementan a los adaptadores 1 y 2R. En este paso las únicas moléculas que se amplificarán exponencialmente son las moléculas “e”. Las moléculas “a” y “c” presentan un solo adaptador y se amplificarán linealmente, las “d” no presentan adaptador y por lo tanto no se le une ningún cebador. Las moléculas “b” sufren el proceso de supresión en la etapa de desnaturalización de la PCR, debido a que presentan el mismo adaptador en los extremos. Estos adaptadores

contienen secuencias terminales repetidas e invertidas (LITR, siglas en inglés de *Long Inverted Terminal Repeats*) que incentivan la formación de estructuras secundarias al complementarse, denominadas “panhandle - like”, dificultando la polimerización. (Figura 6). Por lo tanto, el producto final de la SSH, básicamente serán moléculas presentes en el *tester* y ausentes en el *driver*. Finalmente los ADNc expresados diferencialmente pueden ser recuperados, clonados o secuenciados directamente. (Badapanda, 2013; Zhu *et al.*, 2001; Diatchenko, *et al.*, 1996).

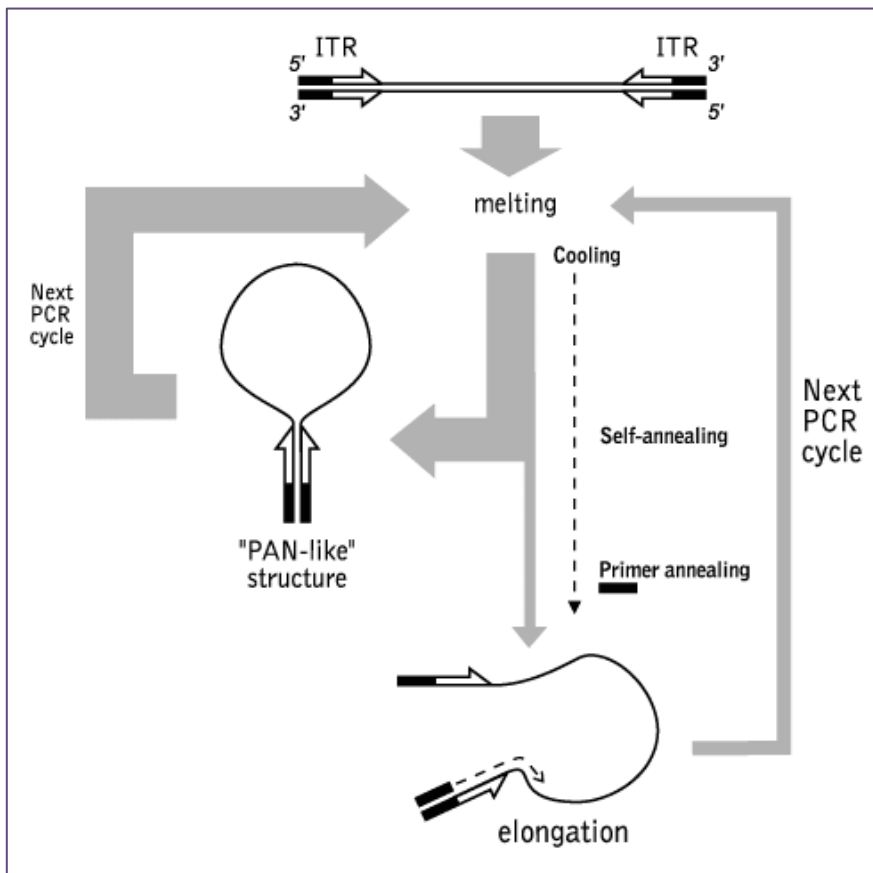


Figura 6. Representación esquemática de la PCR supresora con secuencias terminales repetidas e invertidas

Fuente: <http://evrogen.com/technologies/SSH.shtml>

2.2.6 Secuenciamiento GS FLX Titanium

La disponibilidad de la información de las secuencias de muchos organismos ha sido posible gracias a las técnicas de secuenciamiento de nueva generación (NGS, por sus siglas en inglés). La empresa 454 Life Science desarrolló el primer instrumento comercial NGS basado en el pirosecuenciamiento masivo paralelo, el cual fue liberado al mercado por la compañía Roche en el 2005 (Life Science fue adquirido por Roche). Tres años después, la compañía presentó la serie de equipos GS FLX Titanium, que se actualizó en el 2011 a FLX+. Éste último fue capaz de secuenciar fragmentos de ADN de hasta 1000 pb en longitud de lectura promedio y por corrida hasta 700 Mb, comparado con los 400 pb y 400 Mb de la serie anterior (Gilles *et al.*, 2011; Hakeem *et al.*, 2016).

Como la mayoría de las tecnologías NGS, el pirosecuenciamiento se da por síntesis (SBS, del inglés *sequencing by synthesis*), debido a que se desarrolla en el proceso de síntesis de la cadena complementaria al fragmento de ADN a secuenciar. Este secuenciamiento se basa en la detección de la luz emitida por la incorporación de un nuevo nucleótido complementario, la luz se produce bajo una serie de reacciones enzimáticas acopladas que suceden por la liberación de pirofosfato (un anión doble fosfato $P_2O_7^{4-}$, comúnmente abreviado como PPi) al incorporarse el nucleótido. El proceso se inicia con la fragmentación y polimerización del ADN para generar genotecas en el que se tienen copias de los fragmentos de ADN a secuenciar. De acuerdo con los objetivos del experimento, las genotecas se encuentran debidamente marcadas con MIDs (del inglés multiplex identifiers). El secuenciamiento se lleva a

cabo en micropocillos, de escala nanométrica, presentes en una placa denominada PicoTiterPlate (PTP), en donde se encuentra el mix de la reacción de polimerización que presenta una perla a la que se une los fragmentos de ADN para su inmovilización (figura 7). La polimerización se lleva a cabo mediante una PCR en emulsión (PCRem) en el que están presentes 4 tipos de enzimas: ADN polimerasa, ATP sulfurilasa, Luciferasa y Apirasa, y los sustratos: adenosin 5´fosfosulfato (APS, del inglés adenosine 5´phosphosulfate) y luciferina. En cada ciclo de la PCR se añaden los 4 nucleótidos en un orden establecido; al coincidir uno de ellos, la ATP sulfurilasa convierte el PPI liberado en ATP en presencia de APS, la luciferasa oxida la luciferina en presencia del ATP generado produciendo oxiluciferina y luz, la cual es detectada por una cámara CCD (del inglés charge couple device) del equipo de secuenciamiento. Los nucleótidos no incorporados y el ATP restante son degradados por la Apirasa, posibilitando que todo el proceso vuelva a repetirse en cada ciclo. La información captada por el equipo es correlacionada para generar un formato de información llamado flujograma. La intensidad de la luz emitida determina si más de un nucleótido específico ha sido añadido (por ejemplo un dinucleótido AA). Entre las principales ventajas que se adjudicó a este secuenciamiento fue la de su utilidad con organismos no modelos, ya que al producirse relativamente lecturas largas se facilitaba el proceso de ensamblaje para generar los *contig* (secuencias consenso) en ausencia de una genoma de referencia. Entre sus desventajas se señaló la alta tasa de error que presentaba, sobre todo cuando se encontraba varias bases repetidas consecutivas (Owen-Hughes y Engeholm, 2007; Rowe *et al.*, 2017; Hakeem *et al.*, 2016; El-Metwally *et al.*, 2014).

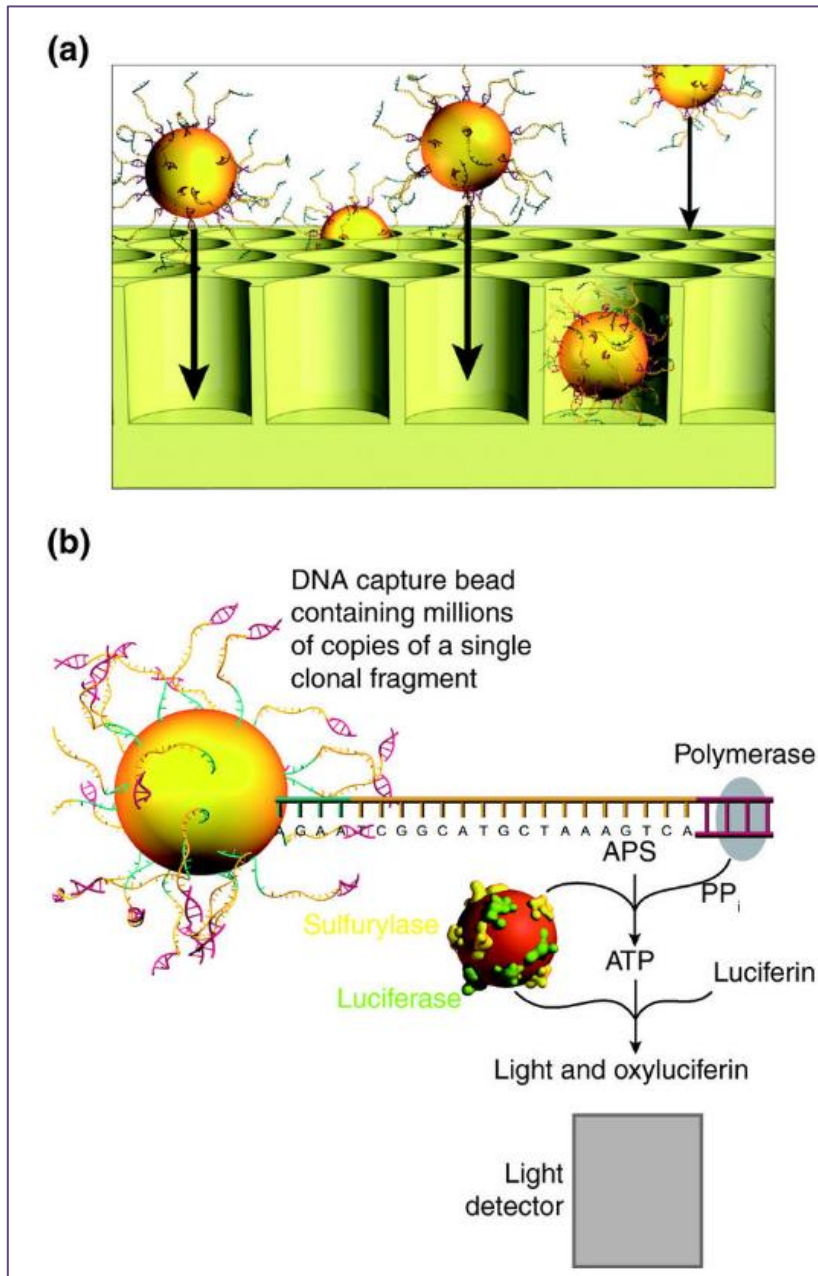


Figura 7. Esquema del pirosecuenciamiento

- a) PicotiterPlate (PTP) con perlas que tienen unido ADN de simple cadena que está fragmentado y amplificado, b) proceso del pirosecuenciamiento. APS (Adenosin fosfosulfato), ATP (Adenosin trifosfato), PP_i (Pirofosfato)

Fuente: Owen-Hughes y Engeholm, 2007

2.2.7 Ontología génica (GO)

En los experimentos de una investigación molecular es frecuente comparar el estado de un tratamiento frente al de condiciones normales para analizar los cambios producidos en un organismo. En este proceso se identifican genes y/o productos génicos (entendido como alguna entidad codificada por el gen, por ejemplo ARNs funcionales y proteínas), que se expresan diferencialmente. Si se desea explorar cuál es el cambio molecular a una mayor escala, se hace uso de un análisis de todo el sistema biológico de acuerdo a la información de los procesos y/o funciones biológicas conocidas de los genes. La información de los genes, por lo tanto, debe ser descrita de forma estructurada que posibilite su análisis con herramientas computarizadas de manera más generalizadas.

Las ontologías, constituyen en sí, bases de datos que almacenan de forma estructurada la información existente o conocida de los productos de genes de manera independiente a las especies, ya que a medida que se fueron secuenciando los primeros genomas se tuvo la evidencia que las ontologías, constituyen en sí, bases de datos que almacenan de forma estructurada la información existente o conocida de los productos de genes de manera independiente a las especies, de los genes ortólogos, sobretodo de los que participan en el proceso biológico núcleo (Gaudet *et al.*, 2017; Ashburner *et al.*, 2000). En las ontologías se incluyen las relaciones lógicas que conectan diferentes conceptos relacionados a un producto génico. Por lo general las ontologías que se utilizan para el análisis funcional contienen información sobre funciones, procesos, o las localizaciones donde actúan. Existen varias ontologías,

entre las más usadas y aceptadas se encuentran la ontología GO (del inglés Gene Ontology), (Ashburner *et al.*, 2000) y KEGG (del inglés Kyoto Encyclopedia of Genes and Genomes), (Kanehisa y Goto, 2000).

GO se define como un vocabulario controlado de términos para representar la biología de una manera estructurada (Gaudet y Dessimoz, 2017). Se desarrolló en el Consorcio de Gene Ontology (www.geneontology.org) fundado en 1998 y surgió de la necesidad de la descripción estandarizada de la función de genes y de los productos de genes entre las bases de datos de eucariotas. El proyecto comenzó con la colaboración de las 3 primeras bases de datos de los organismos modelo, que fueron el de *Drosophila* (FlyBase), de *Saccharomyces* (SGD de Saccharomyces Genome Database) y el de ratón (MGD de Mouse Genome Database). En la actualidad, cuenta con la incorporación de muchas bases de datos, incluyendo los mayores repositorios de genomas de animales, plantas y microbios (Ashburner *et al.*, 2000; Gene Ontology Consortium, 2015). Una ontología génica comprende un conjunto de términos bien definidos así como de las relaciones que se establecen entre esos términos. Debido a la naturaleza de la imprecisión de cómo describir una función, en la que se puede referir a una actividad bioquímica, o una estructura celular, como por ejemplo la función de la proteína tubulina puede referirse como una “GTPasa” o como un “constituyente del huso mitótico” respectivamente.

GO se encuentra estructurada en tres ontologías independientes:

- Componente celular (CC)

Describe el lugar en la célula donde un producto génico es activo, dando entender que es la parte de algo más estructurado, como por ejemplo: “retículo endoplasmático rugoso”, “ribosoma”, “proteosoma”.

- Función molecular (FM)

Hace referencia a la actividad bioquímica de un producto génico, desde descripciones generales hasta específicas, como por ejemplo: “Transporte”, “Actividad catalítica”, “Actividad de transcripción”; es decir, describe las actividades más que a las entidades que lo realizan, sin especificar dónde, cuándo, o en qué contexto la acción toma lugar.

- Proceso biológico (PB)

Se refiere a un objetivo biológico al cual contribuyen un gen o algunos de los productos génicos asociados a él. Describe una serie de procesos realizados por uno o un conjunto de funciones moleculares. Como ejemplo está el “Metabolismo de pirimidinas”, o “Transducción de señal”. Los procesos, usualmente involucran una transformación química o física, de modo que algo ingresa a un proceso y algo diferente se obtiene a su salida. Para distinguir entre un proceso biológico y una función molecular se toma en cuenta que el primero presenta más de un paso.

En la GO la información de los genes está en estructuras con jerarquía, denominadas grafos dirigidos acíclicos (GDA, por sus siglas en inglés), un ejemplo se muestra en la figura 8. Las flechas que se presentan en estos gráficos pueden significar “es un”, “parte de”, o “regulan”. En cada GDA, los términos GO se representan como nodos

en la estructura, en cada nodo está asociado los genes que se relacionan con el concepto. A medida que se dirigen a nodos inferiores, aumenta la especificidad de la información. Por lo tanto si un gene está asignado a un determinado nodo, con seguridad lo estará en nodos que lo anteceden (el nodo que antecede es denominado “padre” y el que sale es denominado “hijo”). Cada nodo puede vincular a más información, incluyendo las palabras claves de los genes y proteínas presentes en las principales bases de datos como SwissPROT, GenBank, EMBL, etc. GO es dinámico, ya que se encuentra en constante cambio, debido a nueva evidencia de la información de los genes que se obtiene constantemente. En la figura 8 se observa una ontología de proceso biológico. Se ilustra una porción del proceso biológico del metabolismo de ADN. Nótese que un nodo puede tener más de un “padre”, por ejemplo “DNA ligation” tiene 3 “padres”: “DNA –dependent DNA replication”, “DNA repair”, “DNA recombination” (Gaudet, *et al.*, 2017; Gene Ontology Consortium, 2015).

La ontología permite que se pueda dar la anotación GO, que es el proceso de asignar un término GO, que consiste en un único identificador alfanumérico, a un producto de gen, de la forma: GO:nnnnnnn. Para hacer una anotación GO de ESTs (del inglés Expressed Sequence Tag), lo recomendable es realizar primero un BLAST y tener presente la anotación anotada manualmente de la secuencia homóloga.

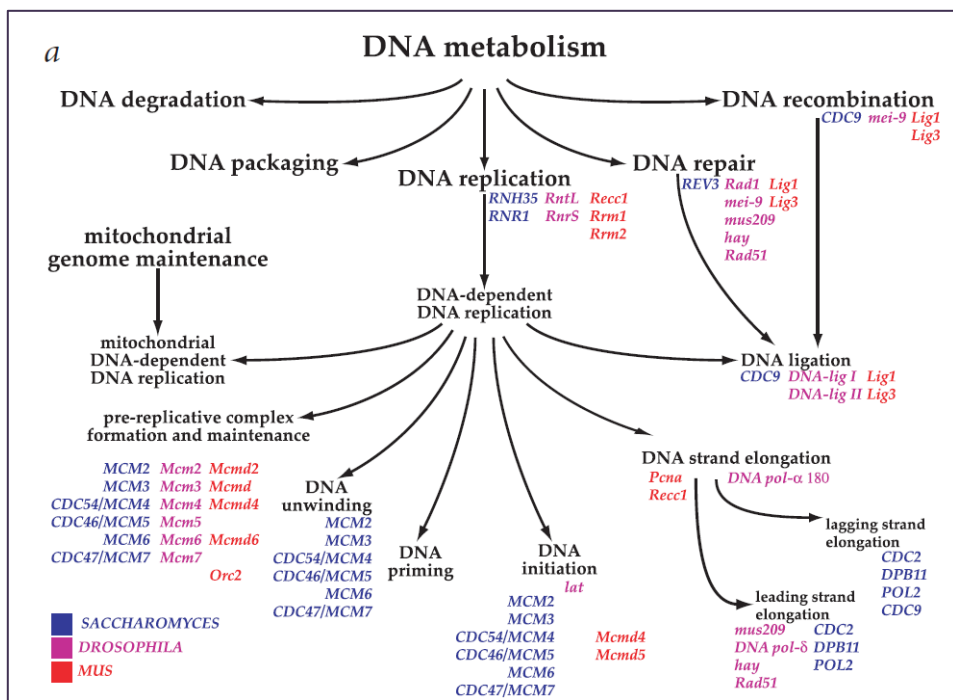


Figura 8. Ejemplo de ontología génica

Fuente: Ashburner *et al.*, 2000

Cada anotación incluye un código de evidencia para indicar el respaldo de la asociación a un término GO, este respaldo puede ser una referencia literaria, análisis computacional, u otras bases de datos. La mayoría de las anotaciones de un grupo de secuencias, son electrónicamente inferidas (IEA, por sus siglas en inglés de inferred from electronic automatic), como con todos los tipos de anotaciones, IEA cambia todo el tiempo debido a la actualización de la información. Una ventaja de la anotación automática es que es rápida, facilitando la anotación de gran cantidad de información de productos génicos conocidos o predichos, esta anotación presenta menos detalle que la manual. Normalmente la evidencia es más robusta para una anotación en general que una anotación específica (Ashburner *et al.*, 2000; Huntley *et al.*, 2014; Gaudet y Dessimoz, 2017).

2.2.8 Análisis de enriquecimiento de términos GO entre 2 grupos de secuencias

Es el análisis que se realiza al comparar 2 grupos de genes identificados por estudios genómicos de alto rendimiento para analizar las diferencias de términos GO entre los dos grupos de secuencias, con lo cual se identifica estadísticamente el término GO que está sobre-/sub-representado. El modelo estadístico que más se utilizan con este fin es la prueba exacta de Fisher, la cual está basada en tablas de contingencia para estudiar la significancia de la asociación (Tipney y Hunter, 2010).

Dado 2 tratamientos en el que se desea averiguar los términos GO que se encuentran significativamente diferentes, se recomienda primero realizar un alineamiento múltiple de las secuencias con BLAST, identificar el GO y realizar el análisis de enriquecimiento. Uno de los paquetes informáticos de mayor uso para este fin es el Blast2go (Conesa *et al.*, 2005) el cual usa el blast local o el blast remoto para encontrar las secuencias similares a la secuencia ingresada. El programa extrae el término GO asociado de cada hit obtenido y evalúa la anotación GO para la secuencia a través de la herramienta FATIGO, que usa el Prueba exacta de Fisher para la evaluación estadística de las diferencias de GO y corrige la evaluación multiple mediante el método estadístico FDR (del inglés de false discovery rate) (Al-Shahrour *et al.*, 2004).

2.3 Justificación del estudio

La FAO calcula que en los próximos 20 años se necesitará un aumento considerable en la producción de alimentos. A nivel mundial este reto será enfrentado con una ampliación de la frontera agrícola y con un crecimiento revolucionario de la productividad. En nuestro país debido a la inminencia del cambio climático y a las limitaciones que nuestra geografía imponen sobre la agricultura lo más factible es que este incremento en la producción se logre por una mejora substancial de la productividad y no por una significativa expansión en la frontera agrícola (por lo menos no por encima de los 2,500 msnm). Uno de los cultivos donde este reto es mucho más relevante es la papa cuya producción a nivel nacional se ha mantenido más o menos constante en los últimos años. Entre los factores limitantes más importantes para el rendimiento encontramos a los agentes bióticos (bacterias, hongos, virus, nemátodos, polillas de la papa, etc) y abióticos (sequías, heladas, alta irradiación lumínica, etc).

Se sabe que las especies del género *Solanum* que producen tubérculos presentan una gran variación genética tanto en la tolerancia a las heladas en el estado no aclimatado como en la capacidad de aclimatación y, debido al desconocimiento de los genes implicados en ese mecanismo existe un desaprovechamiento de esa biodiversidad de genes relacionados a la resistencia a heladas presentes en las papas silvestres y nativas para su uso en el mejoramiento genético de las papas cultivadas y otros cultivos relacionados.

2.4 Objetivos

2.4.1 Objetivo general

- Identificar genes expresados diferencialmente durante condiciones de heladas en hojas de *Solanum juzepczukii* (papa nativa) para ser candidatos a una futura aplicación en el mejoramiento genético de la papa.

2.4.2 Objetivos específicos

- Estandarizar las condiciones de cultivo de las plantas de *Solanum juzepczukii*, en cámaras de crecimiento, para la obtención de las muestras.
- Obtener genotecas de ADNc con el método de hibridación sustractiva por supresión (SSH), a partir de hojas de plantas de *Solanum juzepczukii* sometidas a tratamiento de helada.
- Identificar genes anotados como factores de transcripción, regulados “hacia arriba” y “hacia abajo” , expresados durante tratamientos de heladas en hojas de *Solanum juzepczukii* y que estén asociados a la tolerancia a algún estrés osmótico.

3 METODOLOGÍA

La figura 9 se muestra un diagrama de flujo de la metodología propuesta.

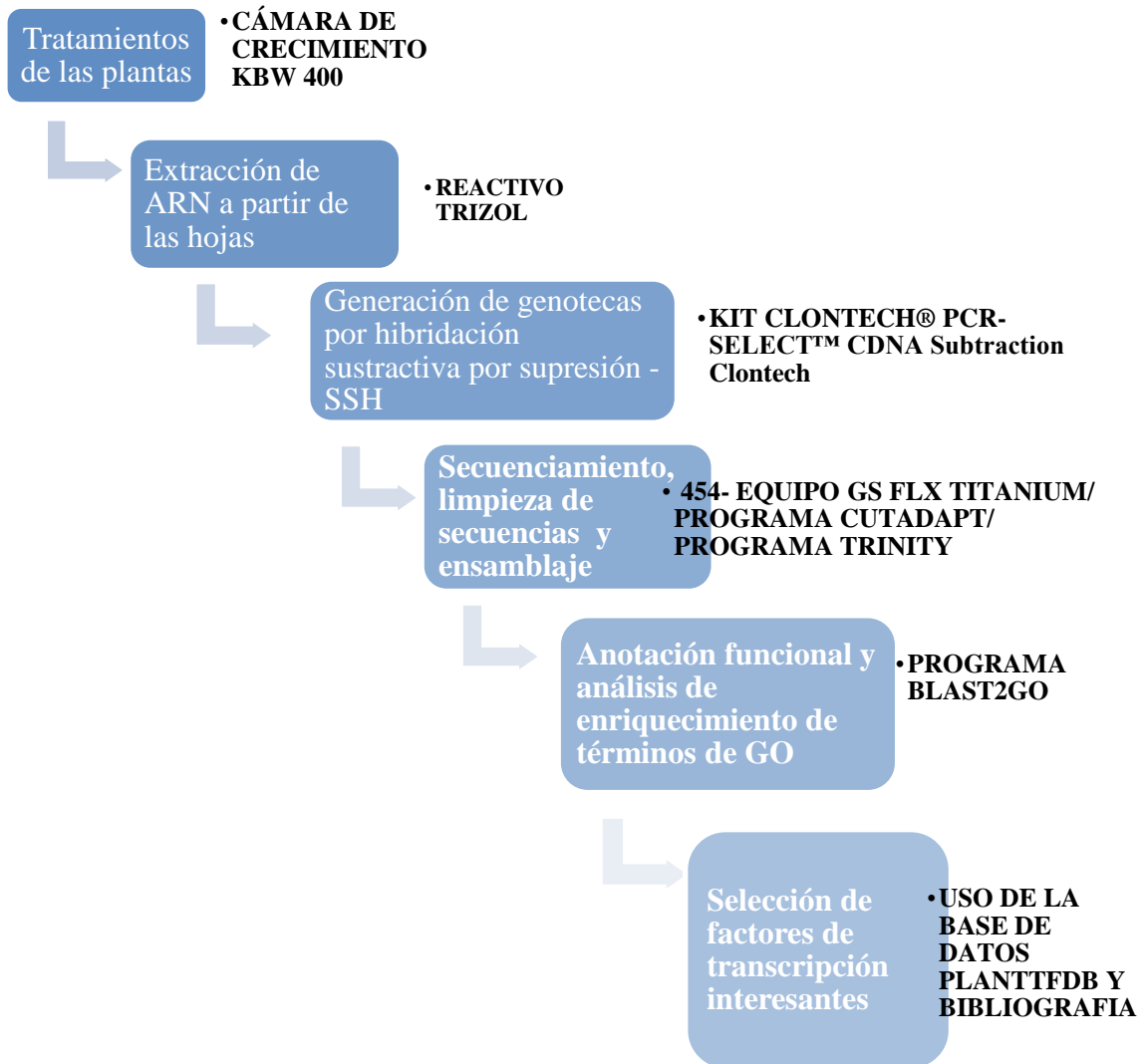


Figura 9. Diagrama de flujo general de los pasos metodológicos
Se muestra la metodología usada o programa bioinformático.

3.1 Cultivo y tratamientos de las plantas

Plántulas de dos semanas de crecimiento de la especie *Solanum juzepczukii* fueron donados por el Centro Internacional de la Papa (CIP), perteneciente al número de accesión CIP 703258 *S. juzepczukii* “Laram K'anchali”. Esta especie está catalogada como resistente a la helada en la evaluación del banco de germoplasma del CIP. Las plántulas fueron trasladadas a macetas de plástico N° 16 que presentaban tierra preparada. Se las colocó en tierra de cultivo, que presentaba la siguiente composición: 6 partes de musgo, 2 tierra de chacra y 1 de arena de playa lavada. Las macetas con las plántulas fueron colocadas en la cámara de crecimiento KBW 400 (Binder, Alemania), como se observa en la figura 10, bajo las siguientes condiciones de crecimiento: 20 °C x 12 h luz y 15 °C x 12 h oscuridad. Semanalmente las plantas fueron regadas con agua autoclavada. En el primer riego se le adicionó el fungicida Farnathe (FARMAGRO S.A. PERÚ) a concentración de 1g por cada litro de agua. Las plantas crecieron en la cámara por aproximadamente dos meses y medio. Pasado este tiempo, 2 tipos de tratamientos: T. Control y T. Helada, se tomaron en cuenta para la obtención de las muestras para las genotecas.

Las muestras del T. Control fueron hojas de plantas que crecieron en condiciones normales que simulan las condiciones ideales en las que las plantas crecen, y las de T. helada, fueron hojas de plantas que pasaron por condiciones que simulan el fenómeno climatológico de las heladas, luego de colocarlas previamente por un periodo de aclimatación. Para ambos tratamientos se colectaron hojas de al menos 3 plantas.

Un esquema de las condiciones de los tratamientos se muestra en la figura 11. Para el T. control, las condiciones de crecimiento fueron: 20 °C x 12 h luz y 15 °C x 12h oscuridad. Las demás plantas, las que no fueron consideradas para la colecta, fueron sometidas a condiciones de aclimatación a 5 °C x 12 h luz durante dos semanas, luego se procedió con el T. helada, en el que las plantas estuvieron a una temperatura de – 5 °C en 8 h sin luz. De este tratamiento se tomaron en cuenta las hojas que visiblemente no estaban necrosadas, las cuales rápidamente se colocaron en tubos Falcon de 50 ml para introducirlas en nitrógeno líquido y guardarlas a -80 °C, junto con las demás muestras, hasta su uso posterior. Los programas de condiciones para cada tratamiento en la cámara de crecimiento KBW 400, se muestra en la Tabla 3. Cada programa (P) consta de una sección (S) en la que se fija una temperatura (T°) a la cual llegar a modo de rampa (que se encuentra por defecto en la cámara) , con una velocidad de ventilación (en porcentaje), en una longitud de sección (tiempo), la iluminación está determinado por las líneas de operación (111 es para administrar luz en los tres fluorescentes de la cámara y 000, para oscuridad). También se establece la repetición de cada programa.

A.



B.

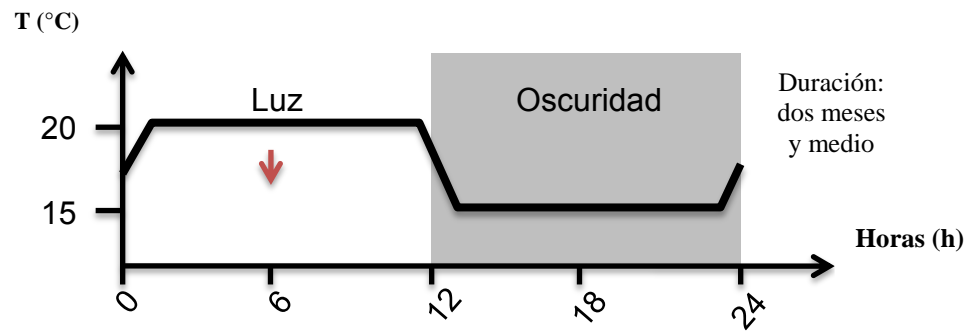


C.

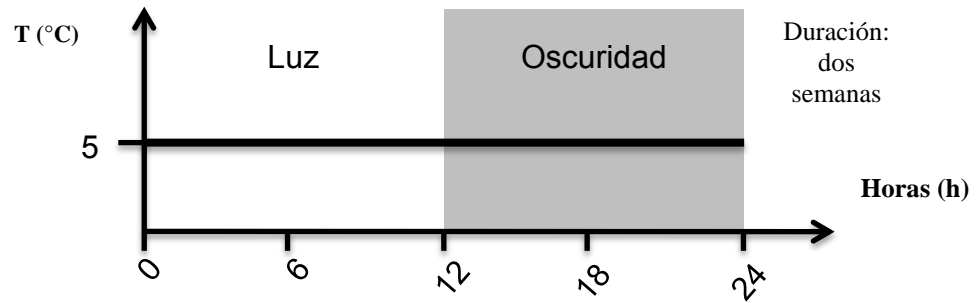


Figura 10. Cámara de crecimiento KBW 400 y plantas de la especie *Solanum juzepczukii*
A. Vista exterior, B. Vista interior y C. Plantas de la especie *Solanum juzepczukii* en macetas dentro de la cámara.

A.



B.



C.

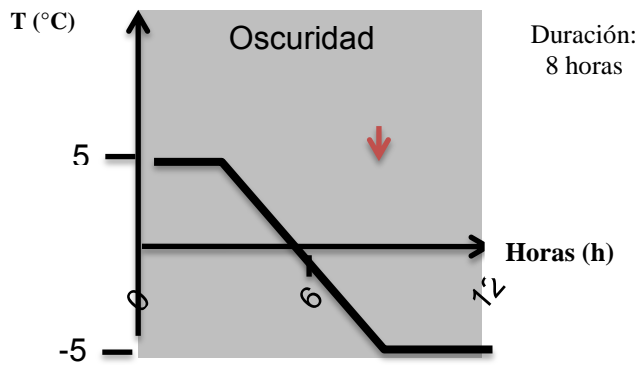


Figura 11. Esquema de las condiciones de tratamientos para la obtención de las muestras
A. Tratamiento control. B. Condiciones de aclimatación C. Tratamiento helada. La flecha roja indica el momento de la colecta de las hojas finalizado la duración del tratamiento.

Tabla 3. Programación en la cámara KBW 400 de los tratamientos

A.

Programa (P)	Sección (S)	Temperatura (°C)	Longitud de la sección (hh:mm)	Velocidad del ventilador (%)	Líneas de operación*
P01	S01	20	0:30	70%	111
	S02	20	11:30	70%	111
	S03	15	1:00	70%	0
	S04	15	11:00	70%	0

Repetición del programa 01: dos meses y medio

B.

Programa (P)	Sección (S)	Temperatura (°C)	Longitud de la sección (hh:mm)	Velocidad del ventilador (%)	Líneas de operación*
P01	S01	5	1:00	80%	111
	S02	5	11:00	80%	111
	S03	5	12:00	80%	0
P02	S01	5	12:00	80%	111
	S02	5	12:00	80%	0

Repetición del programa 01: 1, del programa 02: dos semanas

C.

Programa (P)	Sección (S)	Temperatura (°C)	Longitud de la sección (hh:mm)	Velocidad del ventilador (%)	Líneas de operación*
P01	S01	-5	8:00	100%	0
	S02	-5	1:00	100%	0

Repetición del programa 01: 1

A. Tratamiento control B. Condiciones de aclimatación C. Tratamiento helada.*Las líneas de operación refiere a la iluminación de cada uno de los tres fluorescentes de la cámara: 1, es con luz, y 0, sin luz.

3.2 Extracción de ARN

El ARN de todas las muestras se aisló con el reactivo TRIZOL (SIGMA), con algunas modificaciones de las recomendaciones del producto. Aproximadamente 100 mg de tejido fueron pulverizados con nitrógeno líquido en morteros, el polvo fue colocado en microtubos de 1.5 ml y se le agregó 1 ml del reactivo TRIZOL para luego homogenizar el contenido. Los tubos se centrifugaron a 4 °C, 14000 rpm x 10 min. y los sobrenadantes se colocaron en tubos nuevos para colocarles 200 µL de cloroformo. Se homogenizó el contenido y se centrifugaron los tubos a 4 °C, 14000 rpm x 15 min. Los sobrenadantes se colocaron en tubos nuevos y se les añadió 500 µL de isopropanol frío. Se los dejó reposar a temperatura ambiente por 10 min., luego se centrifugaron a 4 °C, 14000 rpm x 10 min. A los pélets obtenidos se les realizó dos lavados consecutivos con etanol al 80% y 70% respectivamente. Se dejó secar a temperatura ambiente por 10 min los tubos, luego los pélets fueron resuspendidos en agua tratada con DEPC (Dietilpirocarbonato) al 2%. Finalmente, la calidad del ARN fue determinada por espectrofotometría y electroforesis. El primero se realizó en el espectrofotómetro Nanodrop 1000 (Thermo Fisher Scientific) y el segundo mediante un gel al 1.5% de agarosa con 1X de tampón TAE (Tris-acetato), y tinción con bromuro de etidio y con tampón de corrida 1X TAE, la visualización del gel se realizó con luz UV. Los ARN de las hojas de al menos 3 plantas fueron mezclados entre sí para la obtención de las genotecas.

3.3 Síntesis de ADNc y procedimiento de sustracción

Para identificar genes que se expresan diferencialmente en el tratamiento helada frente al tratamiento normal, se utilizó la metodología de hibridación sustractiva por supresión (SSH) (Diatchenko *et al.*, 1996).

Se sintetizó ADNc (ADN complementario) de doble cadena a partir de ARN total obtenido de las hojas de plantas crecidas de *S. juzepczuckii* de los tratamientos helada y normal. La síntesis se realizó con la tecnología SMART (Switching Mechanism At 5' End of RNA Transcript) (Zhu *et al.*, 2001), con el Kit SMARTer™ PCR cDNA Synthesis Clontech, (Clontech, EE.UU./Canada). Se utilizó el oligonucleótido *SMART Oligo II A* y el cebador *SMART CDS II A* para la síntesis de la primera cadena y el cebador *SMART PCR IIA* para la amplificación por PCR a partir de esa primera hebra de ADNc. Las secuencias de los oligonucleótidos y cebadores utilizados para esta sección se muestran en la tabla 4.

Tabla 4. Secuencias de los oligonucleótidos usados en la síntesis de ADNc por SMART y par su amplificación

Denominación del oligonucleótido	Secuencia
SMART Oligo II A oligonucleotide	5'- AAGCAGTGGTATCAACGCAGAGTACGCrGrGrG-3'
SMART CDS PRIMER II A	5'-AAGCAGTGGTATCAACGCAGAGTA-d(T)30-3'
SMART PCR primer II A	5'-AAGCAGTGGTATCAACGCAGAGT-3'

Para la síntesis de la primera cadena de ADNc, se utilizó 0,5 µg de ARN en un volumen total de reacción de 10 µL con la enzima SMARTScribe™ Reverse Transcriptase (proveído por el kit), se siguió el protocolo recomendado por el fabricante. Los componentes de la reacción de transcripción reversa se muestran en la

tabla 5. La reacción se incubó a 42°C por 1 hora y para detener la reacción se colocó a 70°C por 10 min.

Tabla 5. Componentes de la reacción de transcripción reversa

Reactivo	Cantidad por reacción (μL)
Agua desionizada estéril	1.5
3' SMART CDS Primer II A (12 μ M)	1
ARN total (0,5 μ g en total)	2
Tampón First-Strand 5X	2
DTT (100 mM)	0.25
Mix dNTP (10 mM)	1
SMARTer II A Oligonucleotide (12 μ M)	1
RNase Inhibitor	0.25
SMARTScribe Reverse Transcriptase (100 U)	1
Volumen total	10

Luego, 1 μ L de este ADNc sintetizado y 5 veces diluído, fue utilizado para la amplificación por PCR, para lo cual se utilizó Advantage® 2 PCR Kit. (Takara Bio USA). Se siguió el protocolo recomendado por el fabricante del kit. Los componentes de la reacción de PCR se muestran en la tabla 6. Se programaron 19 ciclos de PCR. Cada ciclo se amplificó a 95 °C por 10 s, 65.5 °C por 20 s, 72 °C por 3 min. La purificación del PCR fue realizada con NucleoSpin® Extract II (Clontech, EE.UU./Canada).

Luego el ADNc amplificado fue digerido por *Rsa* I (reactivo proveído por el kit) para generar extremos romos en las secuencias y se prosiguió con la hibridación sustractiva de las muestras. La calidad del ADNc sintetizado y digerido fue evaluada por electroforesis con un gel de agarosa al 1.5% teñido con bromuro de etidio.

Tabla 6. Componentes de la reacción de la PCR con ADNc

Reactivo	Cantidad por reacción (µL)
Agua desionizada estéril	40
Tampón de reacción de PCR 10X	5
Mix dNTP (10 mM)	1
5' PCR Primer II A (10 µM)	2
50X Advantage 2 Polymerase Mix	1
ADNc	1
Volumen total	50

La sustracción de las muestras de ADNc fue realizada en ambas direcciones con el kit Clontech® PCR-Select™ cDNA Subtraction (Clontech, EE.UU./Canada). La genoteca sustraída **sentido** fue enriquecida con genes que se regulan “hacia arriba” en la condición helada (población de ADNc provenientes de transcritos presentes en mayor cantidad en la condición helada comparado a la condición normal); y la genoteca sustraída **antisentido** fue enriquecida con los genes que se regulan “hacia abajo” en la condición helada (población de ADNc provenientes de transcritos presentes en mayor cantidad en la condición normal comparado a la condición helada). Para la construcción de cada una de estas genotecas se dividió el ADNc *tester* (ADNc provenientes de las muestras del T. helada) en dos alícuotas y a cada una se le ligó los adaptadores 1 y 2R por separado, los cuales sirvieron como sitios de unión a oligonucleótidos en la amplificación de la PCR posterior. A continuación, se realizaron dos hibridaciones. En la primera, se denaturó, hibridó y renaturó las porciones de ADNc tratado con un exceso de 30 veces en cantidad de ADNc driver (ADNc provenientes de las muestras del T. control), en tubos separados. En la segunda hibridación, ambos contenidos son mezclados e hibridados juntos para

obtener ADNc sustraídos de doble cadena que presentan diferentes adaptadores. Luego este ADNc sustraído fue sometido a dos rondas de PCR (primaria y secundaria). En la PCR primaria se utilizó el oligonucleótido *Primer 1* y en la secundaria, los oligonucleótidos anidados *Nested primer 1* y *Nested primer 2R* para normalizar y enriquecer ADNc provenientes de genes diferencialmente expresados. En la tabla 7 se muestra los oligonucleótidos utilizados para sustracción.

Tabla 7. Secuencias de los oligonucleótidos usados en la preparación de las genotecas por SSH

Denominación del oligonucleótido	Secuencia
Adapter 1	5'-CTAATACGACTCACTATAGGGCTCGAGCGGCCGCCCGGGCAGGT-3' 3'-GGCCCGTCCA-5'
PCR primer 1	5'-CTAATACGACTCACTATAGGGC-3'
Nested primer 1	5'-TCGAGCGGCCGCCCGGGCAGGT-3'
Adapter 2R	5'-CTAATACGACTCACTATAGGGCAGCGTGGTCGCGGCCGAGGT-3' 3'-GCCGGCTCCA-5'
Nested primer 2R	5'-AGCGTGGTCGCGGCCGAGGT-3'

3.3.1 PCR primaria

Para la PCR primaria se utilizó 1 µL de la solución de hibridación de ADNc sustraído. Antes de iniciar el ciclaje propio de la PCR, se incubó la mezcla a 72 °C por 5 min. para complementar por polimerización la secuencia de los adaptadores. La PCR se estandarizó en 26 ciclos a 94 °C por 30 s, 66 °C por 30 s y 72 °C por 1.5 min. En la tabla 8 se muestra los componentes de la PCR.

Tabla 8. Componentes de la PCR primaria

Reactivo	Cantidad por reacción (µL)
Agua desionizada estéril	19.5
Tampón de reacción de PCR 10X	2.5
Mix dNTP (10 mM)	0.5
Primer 1 (10 µM)	1
Advantage cDNA Polymerase 50X	0.5
Volumen total	24

3.3.2 PCR secundaria

Para la PCR secundaria se utilizó 1 µL de la mezcla de la reacción corrida de la PCR primaria diluido 80 veces. La PCR se corrió en 10 ciclos a 94 °C por 30 s, 66 °C por 30 s y 72 °C por 1.5 min. En la tabla 9 se muestra los componentes de la PCR.

Tabla 9. Componentes de la PCR secundaria

Reactivo	Cantidad por reacción (µL)
Agua desionizada estéril	18.5
Tampón de reacción de PCR 10X	2.5
Mix dNTP (10 mM)	0.5
Nested primer 1 (10 µM)	1
Nested primer 2R (10 µM)	1
Advantage cDNA Polymerase 50X	0.5
Volumen total	24

Finalmente, para el análisis de los pasos del procedimiento de sustracción se tomó 5 µL de los productos de PCR primaria y secundaria y se cargó en un gel de agarosa al 2% , teñido con bromuro de etidio, que corrió con tampón 1X TAE en electroforesis, y su visualización se realizó con luz UV.

3.4 Secuenciamiento y ensamblaje

Para el secuenciamiento, los productos de la amplificación de la segunda PCR del procedimiento de SSH fueron enviados al Centro de Secuenciamiento de Genomas (GSC) de Washington University of Saint Louis (WUStL, Missouri – EE.UU.). Para lograr el secuenciamiento el ADNc tuvo que ser fragmentado con el uso de uso del equipo ultrasonificador S220 (COVARIS, Massachussetts – EE.UU.), Se utilizó la tecnología del pirosecuenciamiento con el equipo GS FLX Titanium System (ROCHE, Basel - Suiza). La genoteca fue secuenciada en media placa “picotiter” (PTP). Cada una de las dos poblaciones sustraídas fue diferenciada con etiquetas de 10 nucleótidos llamados MID (Multiplex Identifiers). El secuenciamiento fue entregado en archivos en formato SFF (Standard Flowgram Format).

Antes del ensamblaje las lecturas obtenidas del secuenciamiento pasaron por un proceso de limpieza de secuencias no deseadas, el cual fue realizado usando el programa CUTADAPT (v 1.8) (Martin, 2011), (disponible en: <https://cutadapt.readthedocs.io/en/stable/>). Se removió en los extremos de las lecturas, las secuencias poli A, las de baja calidad, las de baja complejidad, así como las pertenecientes a los adaptadores y las del proceso de sustracción. Las lecturas limpias fueron ensambladas *de novo* con TRINITY (V 2.3.2) (Haas *et al.*, 2013), (disponible en : <https://github.com/trinityrnaseq/trinityrnaseq/releases>) bajo los parámetros por defecto. Las secuencias consenso obtenidas (*contigs*) fueron usadas para los demás análisis.

3.5 Anotación funcional y análisis de enriquecimiento de términos de GO

Para la anotación se debió realizar la búsqueda de productos génicos conocidos homólogos, mapeo y anotación. Para tal fin se utilizó la plataforma bioinformática BLAST2GO (4.0) (disponible en: <http://blast2go.org>). Se compararon las secuencias consenso con la base de datos no redundante del Centro Nacional para Información de Biotecnología (NCBI) mediante el algoritmo de blastx (Altschul *et al.*, 1990), usando como límite un $E\text{-value} \leq 10^{-5}$ y con los demás parámetros por defecto. Se realizó el mapeo (búsqueda de los términos GO asociados a los hits) y la anotación con los parámetros por defecto.

En cuanto a la anotación se clasificó los transcritos ensamblados en las tres principales categorías de ontología génica (GO): proceso biológico (BP), función molecular (MF) y componente celular (CC). Cada término GO fue asignado basado en el mejor *hit* de blast X y con un $E\text{-value} \leq 10^{-5}$. Con la misma herramienta se hizo los análisis KEGG (Kyoto Encyclopedia of Genes and Genomes) para identificar los EC (Enzyme Commission) designados a los homólogos de los transcritos.

Se realizó un análisis de enriquecimiento de términos de GO para los transcritos que son representados en cada genoteca, con la Prueba exacta de Fisher presente en el software Blast2GO con un $p < 0.05$. Este análisis sirvió para determinar los términos GO que son sobre-/sub- representados en cada genoteca.

3.6 Búsqueda de genes expresados diferencialmente

Para el hallazgo de genes en las genotecas sustractivas, se hizo búsquedas con la herramienta BLAST2GO de acuerdo al término GO asociado a los transcritos anotados. Se llevaron a cabo tres búsquedas. La primera búsqueda, objetivo principal de esta investigación, fue la de identificar genes anotados como factores de transcripción (FTs); el término GO usado para esta búsqueda fue “actividad de factor de transcripción”. La segunda búsqueda, se dirigió a identificar los principales genes que intervienen en la transducción de la señal activando rutas de señalización, como son las proteínas quinasas dependiente de Ca^{2+} (CDPKs), proteínas quinasas activadas por mitógeno (MAPKs), calcineurina, proteína de unión a calcineurina B (CBL), calmodulina, entre las más conocidas. La búsqueda se realizó con el término “actividad quinasa”. La tercera búsqueda fue más generalizada, y se realizó para hallar los genes asociados por su término GO de “respuesta al estrés abiótico”.

3.7 Selección de genes de factores de transcripción interesantes

Para la selección de FTs interesantes asociados a la tolerancia al frío o a algún estrés osmótico, para una futura evaluación de su expresión, se tomó en cuenta que sean parte de las familias de FTs con dominios conocidos que intervienen en la respuesta estrés abiótico como son los FTs: bZIP (basic leucine zipper, AP2/ERF, (APETALA 2/ethylene-responsive element binding factor), NAC (NAM/ATAF1/CUC₂), WRKY, MYB, C2H2 (Cys2 (C2) His2 (H2)-type zinc fingers) y bHLH (basic helix-loop-helix), además de contar con algún reporte de ese FT y/o de su homólogo en

Arabidopsis thaliana, que lo asocie a la tolerancia. Para recuperar la información de ese FT se busco el número de accesoión del mejor hit del blastx obtenido, en la base de datos de FTs en plantas: PlantTFDB (<http://planttfdb.cbi.pku.edu.cn>), que incluye TFs predichos de organimos con genoma y secuencias EST (Expressed Sequence Tag) disponibles (Guo *et al.*, 2007). La información del dominio funcional y las referencias cruzadas a varias bases de datos públicas bien conocidas están disponibles para cada TF identificado.

4 RESULTADOS

4.1 Cultivo y tratamiento de las plantas

Para analizar el transcriptoma expresado en condiciones de helada en hojas de plantas de *Solanum juzepzuckii*, se sintetizaron dos genotecas recíprocas de ADNc con la metodología SSH. Se obtuvieron dos tipos de muestras para la preparación de las genotecas, la primera muestra consistió de plantas cultivadas a 20 °C por 2 meses y medio (Tratamiento control) y la otra, de plantas que estuvieron bajo un régimen de temperatura inicial de 5 °C que bajó a – 5 °C en el transcurso de 8 horas (Tratamiento helada), luego de pasar por un periodo de aclimatación de 2 semanas a 5 °C. Las plantas de ambos tratamientos crecieron en condiciones adaptadas a la cámara de crecimiento KBW 400. Para que la cámara pueda llegar a temperaturas por debajo de 0°C fue necesario utilizar la ventilación al 100 %. En la figura 12 se muestran algunas plantas adultas luego de pasar por los tratamientos. A la izquierda se observa las plantas crecidas bajo el tratamiento control, se puede observar las hojas y los tallos en buen estado; y a la derecha se muestran las plantas que pasaron por el tratamiento helada, se puede observar la inclinación que sufrió los tallos de las plantas como consecuencia de la baja temperatura.

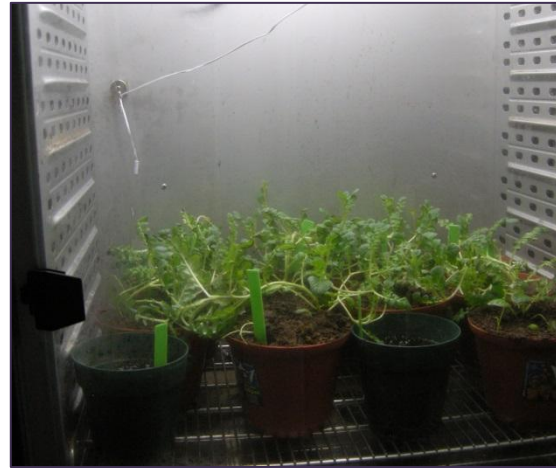


Figura 12. Plantas crecidas en la cámara de cultivo KBW 400
A. Plantas del tratamiento control. B. Plantas del tratamiento helada

4.2 Extracción de ARN

El ARN se aisló a partir de las hojas en buen estado provenientes de plantas de ambos tratamientos. Por cada 100 mg de tejido de hoja se logró recuperar 3 µg de ARN total, en promedio. La calidad y cantidad del ARN fue evaluado por espectrofotometría y electroforesis. No se obtuvo inconvenientes para extraer el ARN utilizando el reactivo de TRIZOL. Los ARN que se extrajeron presentaban la relación de absorbancia A_{260}/A_{280} y A_{260}/A_{230} cercanos a 2. Las muestras en las que se observaron las dos bandas de ARN ribosomal de 28s y 18s con un ratio de intensidad aproximado de 1.5 a 2, a partir de las fotos de electroforesis, fueron seleccionadas para la síntesis de las genotecas. La foto del gel de electroforesis de algunas muestras de ARN extraídos se muestra en la figura 13. El método de extracción escogido fue

reproducibile para las demás muestras en las que se pudo obtener al menos 20 µg para los siguientes pasos de la obtención de genotecas.

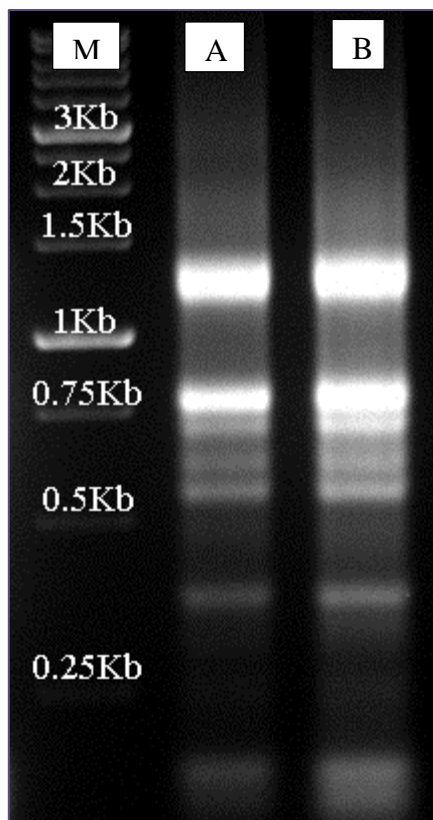


Figura 13. Gel de electroforesis con muestras de ARN extraído de hojas de *Solanum juzepczuckii*
A y B. Muestras de ARN extraído de hojas de *Solanum juzepczuckii*. M. Marcador 1Kb (Fermentas)

4.3 Síntesis y digestión de ADNc

Los ARNs seleccionados fueron utilizados para la síntesis de las dos genotecas. Como primer paso, se sintetizó ADNc con la metodología SMART, técnica que sirvió para generar ADNc con alta representación del extremo 5' de los ARNm, es decir, ARNm con la mayor longitud posible (*full-length*). A partir de la primera hebra de ADNc generada, se sintetizaron copias por PCR, el número de ciclos de la PCR se

estandarizó a 19 para la polimerización adecuada. Luego, el ADNc obtenido de estos pasos fue sometido a digestión con la enzima *Rsa I*. La calidad del ADNc luego de la digestión se evaluó mediante un gel de electroforesis. En la figura 14 se observa que el ADNc antes de la digestión (A y B) presenta alto peso molecular, es decir presenta poca degradación. Y en la sección C y D se muestra el barrido típico luego de someter una muestra a digestión, en la que se observa la reducción de la distribución del tamaño de las bandas, es decir el ADNc quedó digerido en múltiples fragmentos.

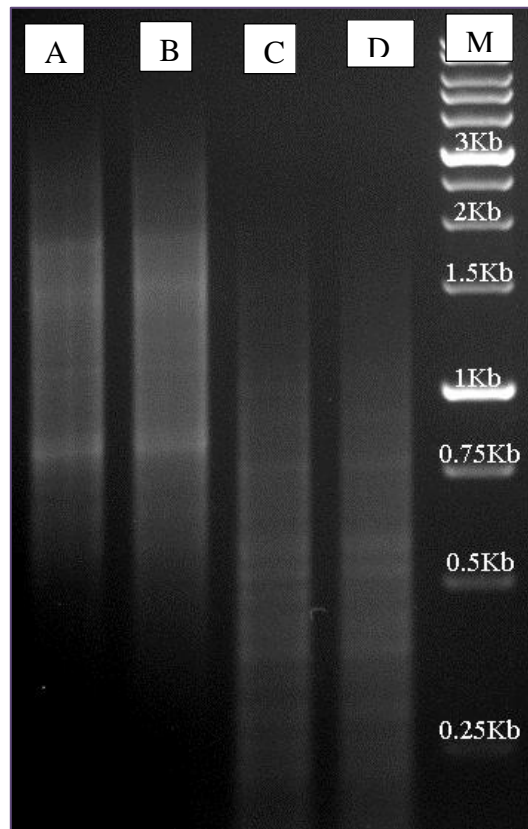


Figura 14. Gel de electroforesis con ADNc del procedimiento de la sustracción
A y B. Muestras de ADNc luego de la polimerización con la metodología SMART del tratamiento control y helada, respectivamente. C y D. Muestras de ADNc digeridas con *Rsa I* de tratamiento control y helada, respectivamente. M. Marcador 1Kb (Fermentas)

4.4 Hibridación sustractiva por supresión (SSH)

Luego de la digestión se realizó la sustracción. Se construyeron dos genotecas con la tecnología SSH en la que se normalizaron y enriquecieron los ADNc provenientes de transcritos expresados diferencialmente entre los dos tratamientos. La genoteca **sentido** fue construida con los transcritos expresados en la condición helada sustrayéndoles los de la condición normal. La genoteca **antisentido**, con los transcritos de la condición normal sustrayéndoles los de la condición helada. Los productos de la sustracción fueron corridos en un gel de electroforesis. En la figura 15 se observan los ADNc luego de la primera y segunda PCR de sustracción, frente a los controles. Se puede observar que existe una uniformidad del contenido de las genotecas comparadas con las muestras no hibridadas, así como un enriquecimiento de zonas específicas en los productos de la sustracción, lo que podría evidenciar el éxito de la sustracción.

4.5 Secuenciamiento y ensamblaje

Para evaluar la calidad de los ADNc que fueron sometidas al secuenciamiento, una muestra de las genotecas fue cargada en el equipo BioAnalyzer 2100 (AGILENT, California – EE.UU) en el Centro de Secuenciamiento de Genomas (GSC) de Washington University of Saint Louis (WUStL, Missouri – EE.UU.). (Figura 16). Como se observa, con mayor claridad, los ADNc en su mayoría, presentan tamaños mayores a los 500 pb.

Luego del secuenciamiento se registraron 155,345 y 151,599 lecturas de la genoteca de ADNc **sentido** y **antisentido**, respectivamente. La longitud promedio de las lecturas de ambas genotecas, luego de la limpieza de las secuencias no deseadas fue de 250 pb.

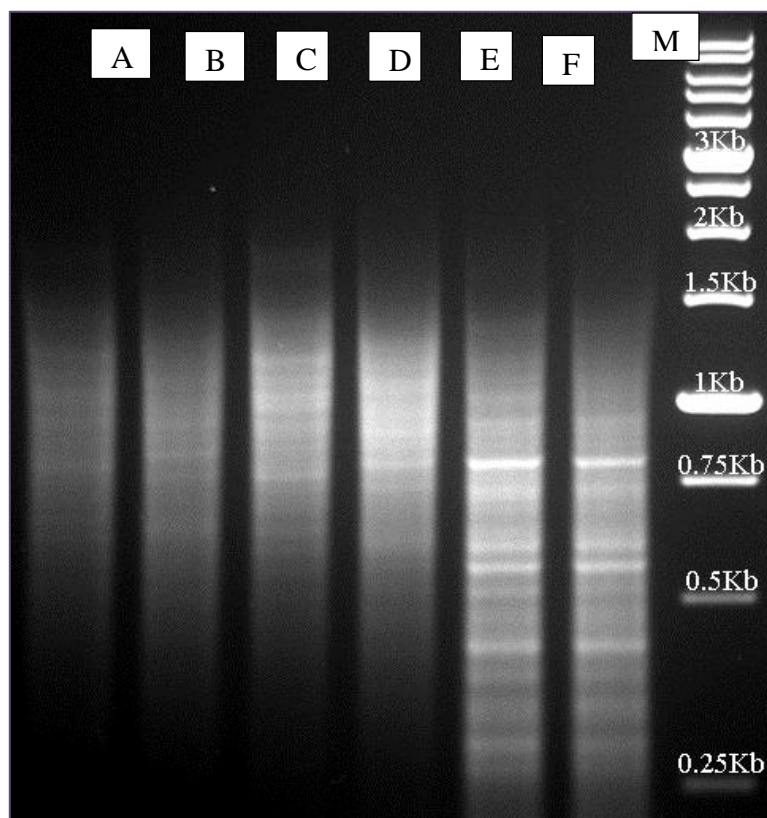


Figura 15. Gel de electroforesis con productos de la PCR del procedimiento de la sustracción
A y B. Muestras de la PCR primaria de la genoteca **sentido** y **antisentido**, respectivamente. C y D. Muestras de la PCR secundaria de la genoteca **sentido** y **antisentido**, respectivamente. E y F. ADNc no sustraídos de la genoteca **sentido** y **antisentido**, respectivamente. M. Marcador 1Kb (Fermentas)

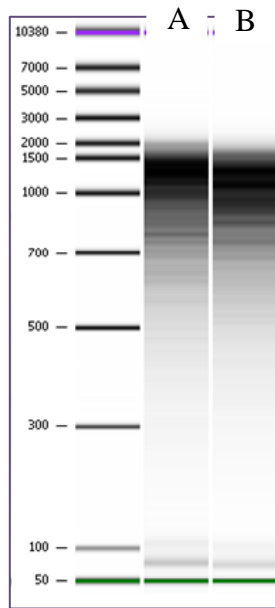
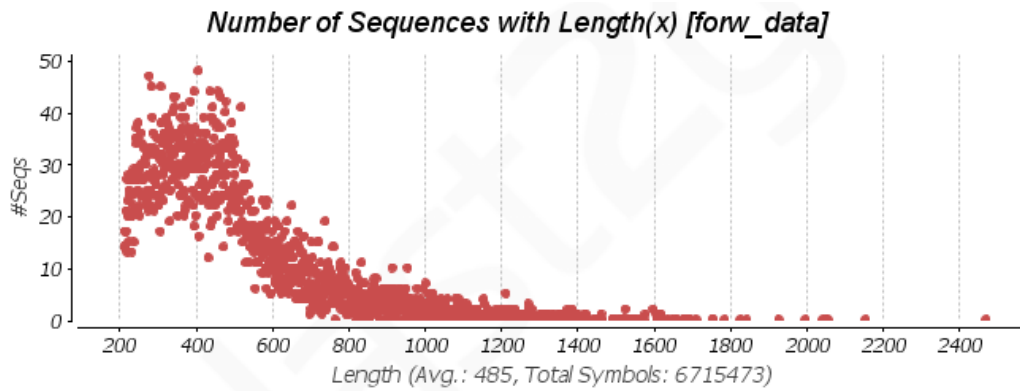


Figura 16. Electroforesis en el equipo BioAnalyzer 2100
Genotecas **sentido** (A) y **antisentido** (B).

De la genoteca **sentido**, el 99,8 % de lecturas se ensamblaron, obteniéndose 13603 *contigs* (secuencia consenso del ensamblaje de las lecturas) y 235 lecturas no ensambladas, con un tamaño promedio de 485,3 pb y un N50 de 515 de los *contigs*; de la genoteca **antisentido**, se ensamblaron también el 99,8 % de las lecturas y se obtuvieron 12535 *contigs* y 204 lecturas no ensambladas, con un tamaño promedio de 491,4 pb y N50 de 521 de los *contigs*. En la figura 17 se muestra la distribución de la longitud de las secuencias ensambladas.

A.



B.

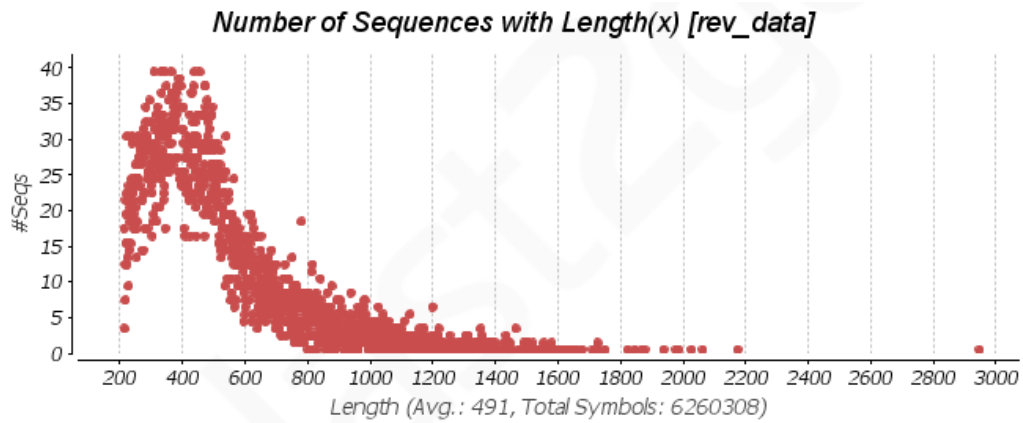


Figura 17. Distribución de la longitud de las secuencias ensambladas
De las genotecas **sentido** (A) y **antisentido** (B)

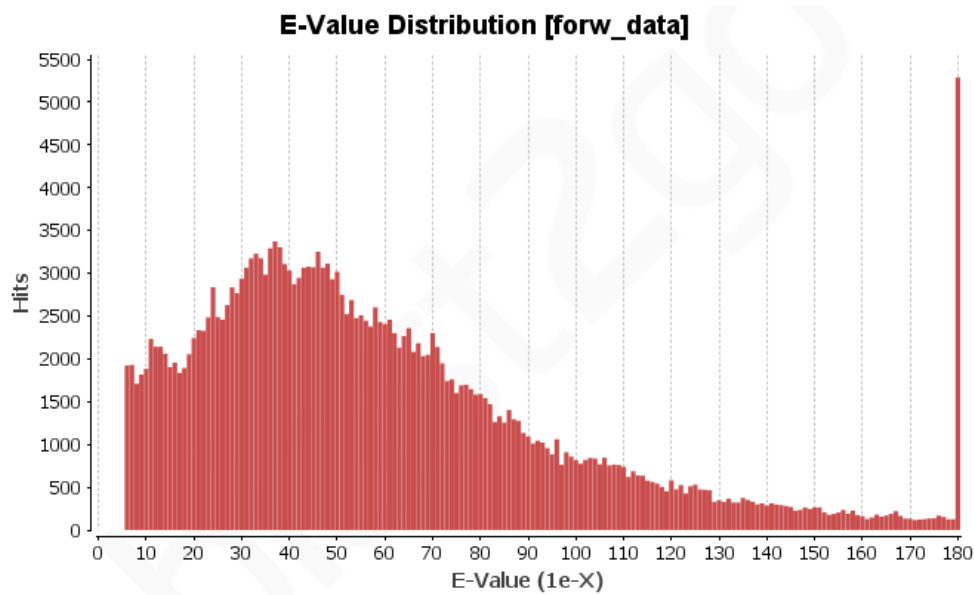
4.6 Anotación funcional

Para predecir y analizar la función de los transcritos ensamblados, las secuencias fueron sometidas a un blastX contra la base de datos no redundante (NR) del NCBI, luego a un mapeo y anotación GO con la herramienta BLAST2GO.

4.6.1 Blast X

De los 13,838 *contigs* en la genoteca **sentido** y 12,739 en la **antisentido**, 12,507 (90, 4%) y 11,892 (93, 4%) secuencias presentaron *blast hit* significativo, respectivamente. La distribución del *E-value* con que los *hits* fueron asignados para ambas genotecas se muestra en la figura 18. Se observa que la mayor parte de las secuencias presenta un *E-value* de entre 10^{-30} y 10^{-50} .

A.



B.

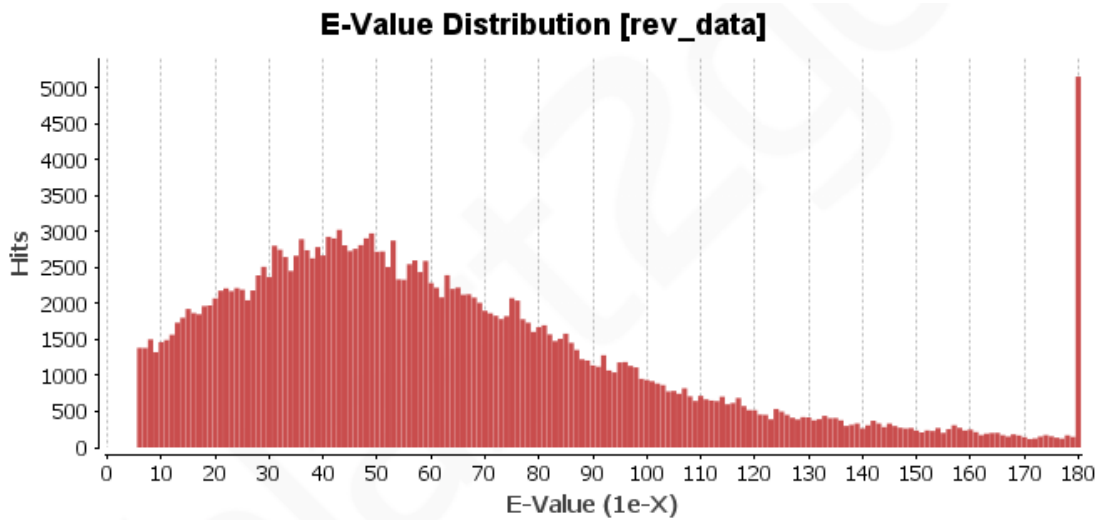


Figura 18. Distribución del E-value luego del blastX
De las genotecas **sentido** (A) y **antisentido** (B)

Las genotecas presentaron una proporción similar de secuencias homólogas con cada una de las más de 29 especies que fueron consideradas por el algoritmo de BlastX. Las especies que presentaron mayor cantidad de secuencias homólogas fueron *Nicotiana tabacum*, *Solanum tuberosum* y *Solanum lycopersicum* (figura 19). En la genoteca **sentido** 33,993 secuencias dieron mayor homología con *Nicotiana tabacum* y 23,530 fueron con *Solanum tuberosum*; en la genoteca **antisentido** 32,447 y 22,293 secuencias dieron homología con las dos especies, respectivamente.

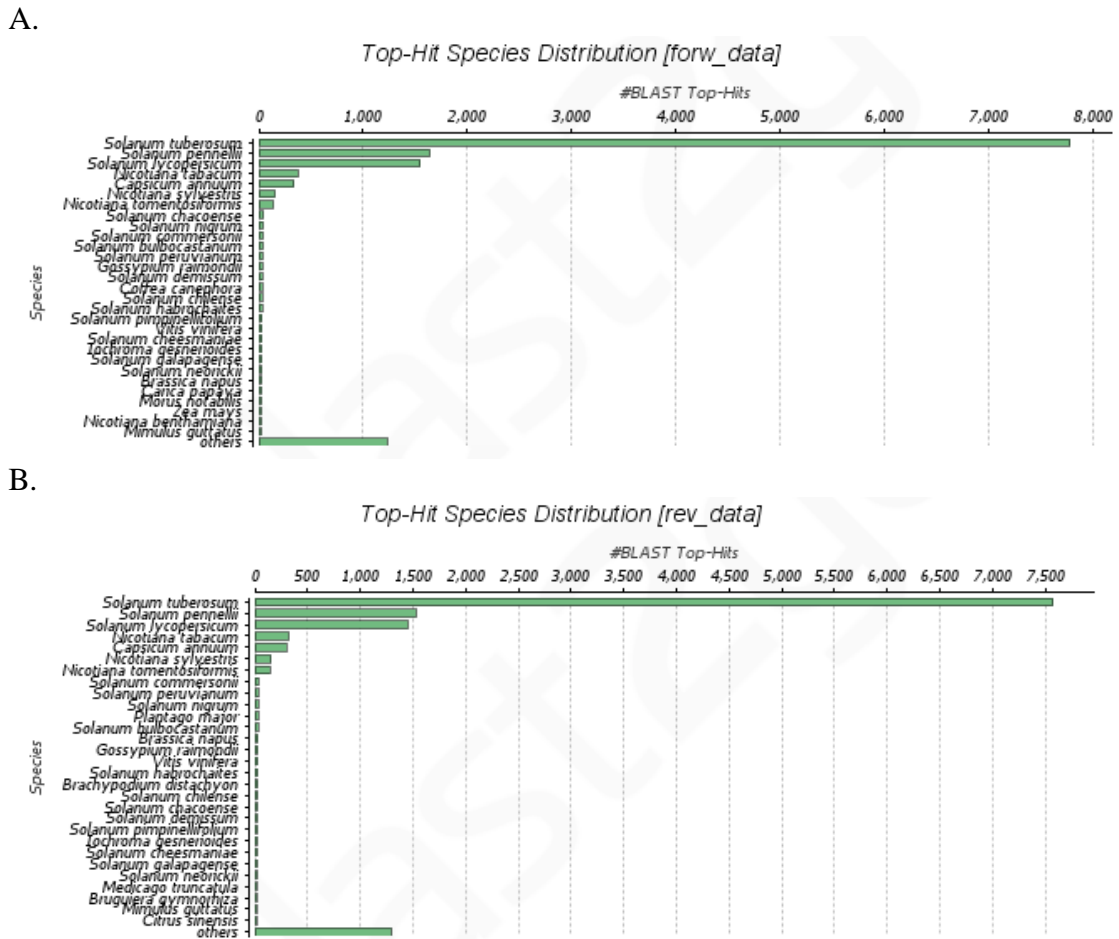
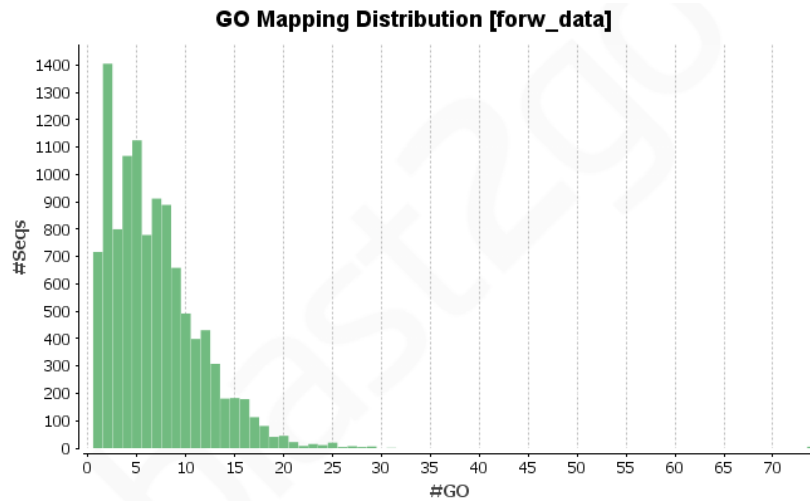


Figura 20. Distribución de las especies top-hit de las secuencias
De las genotecas **sentido** (A) y **antisentido** (B)

4.6.2 Mapeo

De los *contigs* con blastX, 10,910 (87, 2%) y 10,662 (89, 7%) secuencias, presentaron GO recuperados por el mapeo, para la genoteca **sentido** y **antisentido**, respectivamente. La base de datos que en su mayoría proporcionó los GOs fue UniPROTKB. La distribución del número de GO, se muestra en la figura 21. Se puede observar que el algoritmo de BLAST2GO recuperó de 1 a 10 términos GO a la mayor parte de las secuencias de las genotecas.

A.



B.

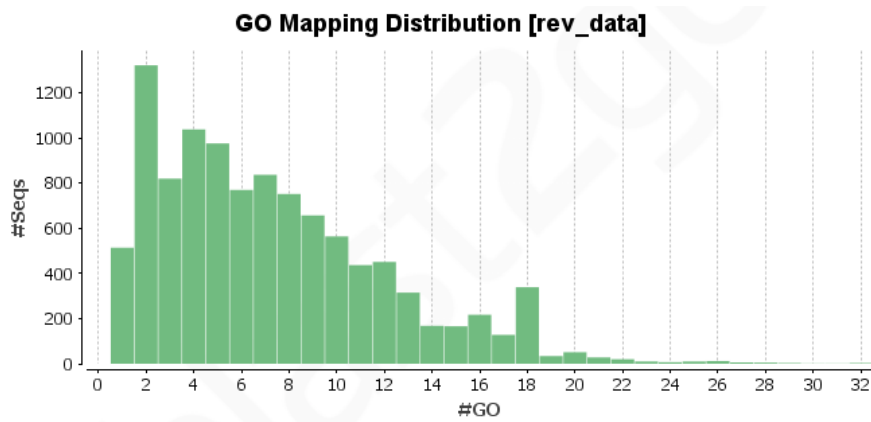


Figura 21. Distribución del número de GO luego del mapeo de las genotecas
De las genotecas **sentido** (A) y **antisentido** (B)

4.6.3 Anotación

Para categorizar funcionalmente los transcritos ensamblados, se asignaron términos de GO con la herramienta BLAST2GO, de acuerdo a los GO mapeados. De los transcritos con mapeo de la genoteca **sentido** y **antisentido**, 10,793 (98,9 %) y 10,553 (98,9 %) tienen un GOslim, respectivamente, asociado significativamente en

alguna de las 3 principales categorías GO: proceso biológico, componente celular y función molecular. El GOslim consiste de un subgrupo del vocabulario de GO que abarca términos claves propios del grupo filogenético.

En resumen, de un análisis global se tiene que en la genoteca **sentido** el 78% (10,793 de 13,838 *contigs* obtenidos) están anotados y tiene asignado un GOslim, en cuanto a la genoteca **antisentido**, 82,8% (10,553 de 12,739) de los *contigs* presenta anotación GOslim (figura 22).

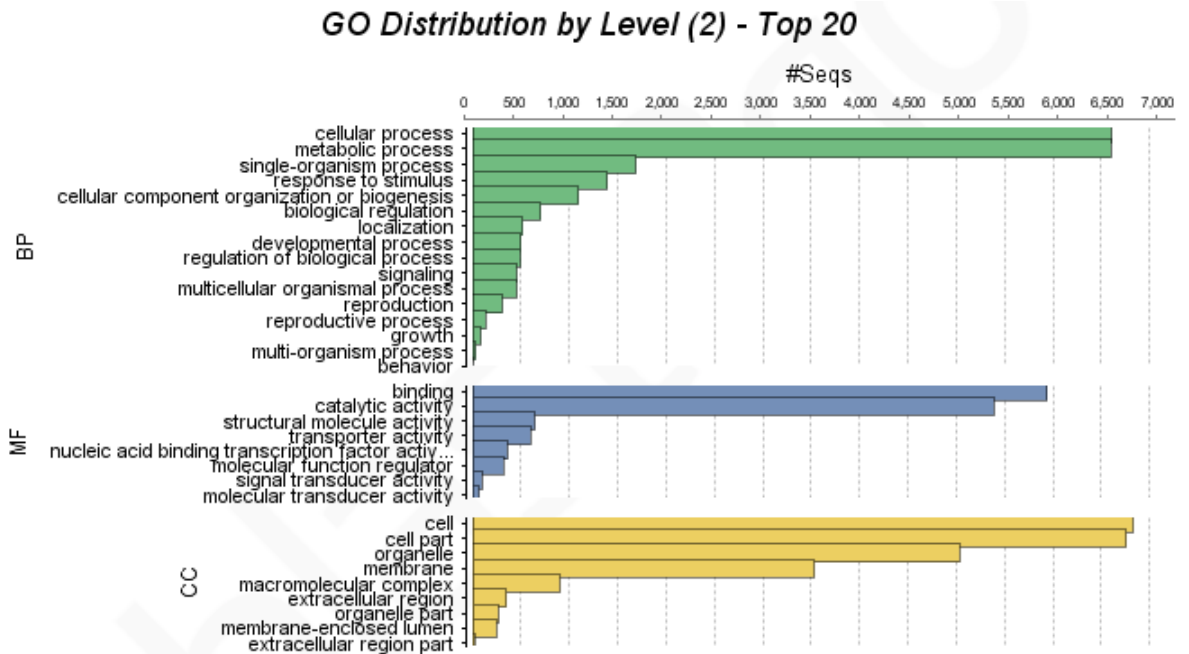


Figura 22. Gráficos circulares de la distribución de los datos luego del BlastX, mapeo y anotación
De las genotecas **sentido** (A) y **antisentido** (B)

Con BLAST2GO se tiene alto soporte en las anotaciones debido a la contribución con la que cuenta de varias bases de datos entre ellas, KEGG (Kyoto Encyclopedia of Genes and Genomes). Según los análisis de KEGG, 4188 y 4432 transcritos de la genoteca **sentido** y **antisentido**, son anotados con la correspondiente Enzyme Commission (EC), con las cuales se identificaron 130 y 140 diferentes rutas KEGG, respectivamente. La distribución de la cantidad de secuencias pertenecientes al tipo de enzimas de ambas genotecas fue similar, y las de mayor cantidad fueron el grupo de Hidrolasas con 1393 y 1489 secuencias; y Transferasas con 1273 y 1178 secuencias, en la genoteca **sentido** y **antisentido**, respectivamente.

La anotación GO de las secuencias (en 2do nivel) dio como resultado que las secuencias de la genoteca **sentido** y **antisentido** en la categoría de proceso biológico (BP), están relacionados principalmente al proceso celular y al proceso metabólico. En términos de función molecular (MF) son asignados a unión y actividad catalítica. Bajo la categoría de componente celular (CC), célula y parte celular ocupan la mayoría de términos GO (figura 23).

A.



B.

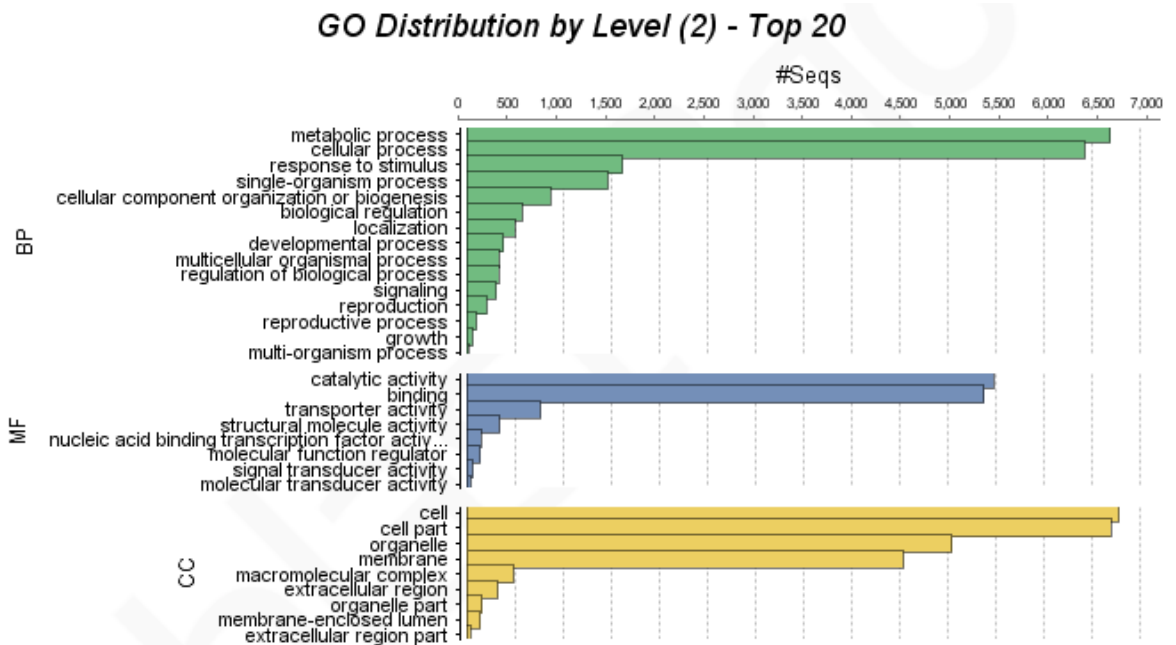


Figura 23. Top 20 de la distribución GO en 2° nivel
De las genotecas **sentido** (A) y **antisentido** (B)

4.7 Enriquecimiento de términos GO - Prueba exacta de Fisher

Para analizar las funciones específicas que muestran diferencia de expresión entre ambas genotecas, se realizó un análisis de enriquecimiento GO usando la Prueba exacta de Fisher, disponible en BLAST2GO. Las figuras 24, 25 y 26 muestran la sobre-/sub- expresión de los GOs, determinados con un $p < 0.05$, en la categoría de proceso biológico (BP), función molecular (MF) y componente celular (CC) respectivamente. En los resultados del análisis de enriquecimiento (con la opción “más específico” en BLAST2GO, se observó que 30 términos GOs en total fueron enriquecidos significativamente en la genoteca **sentido** frente a los 11 de la genoteca **antisentido**. Interesantemente, el estrés causado por la helada resulta en el enriquecimiento de los siguientes GOs en la genoteca **sentido**: actividad de factor de transcripción (MF), unión a ADN de secuencia específica (MF), expresión de genes (BP), actividad de factores de traducción, unión a ARN (BP), señalización (BP), respuesta celular a estímulo (BP), transducción de señal (BP), actividad quinasa (MF); y en la genoteca **antisentido** en: fotosíntesis (BP) y proceso catabólico (BP).

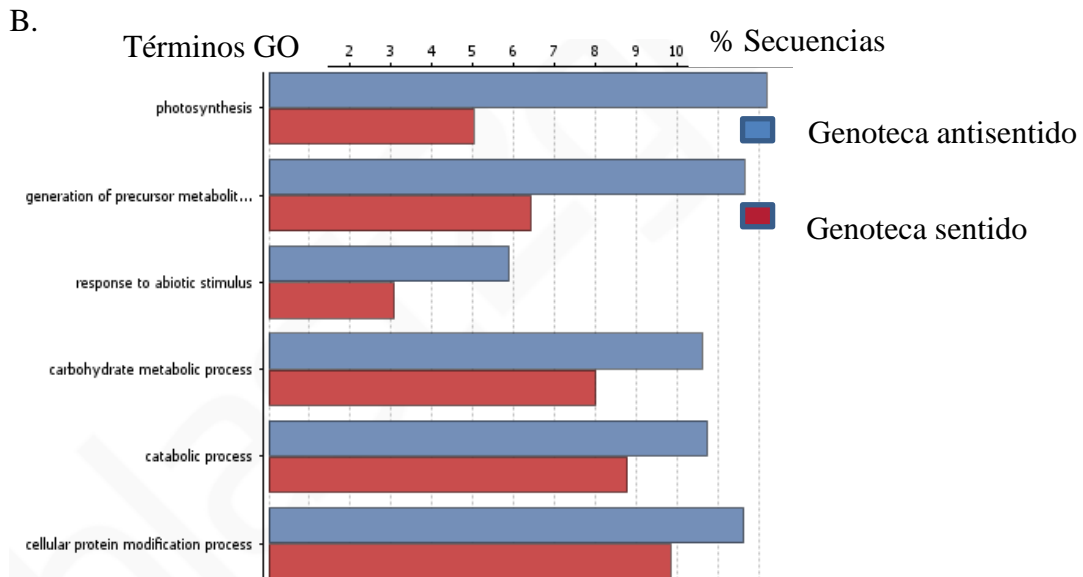
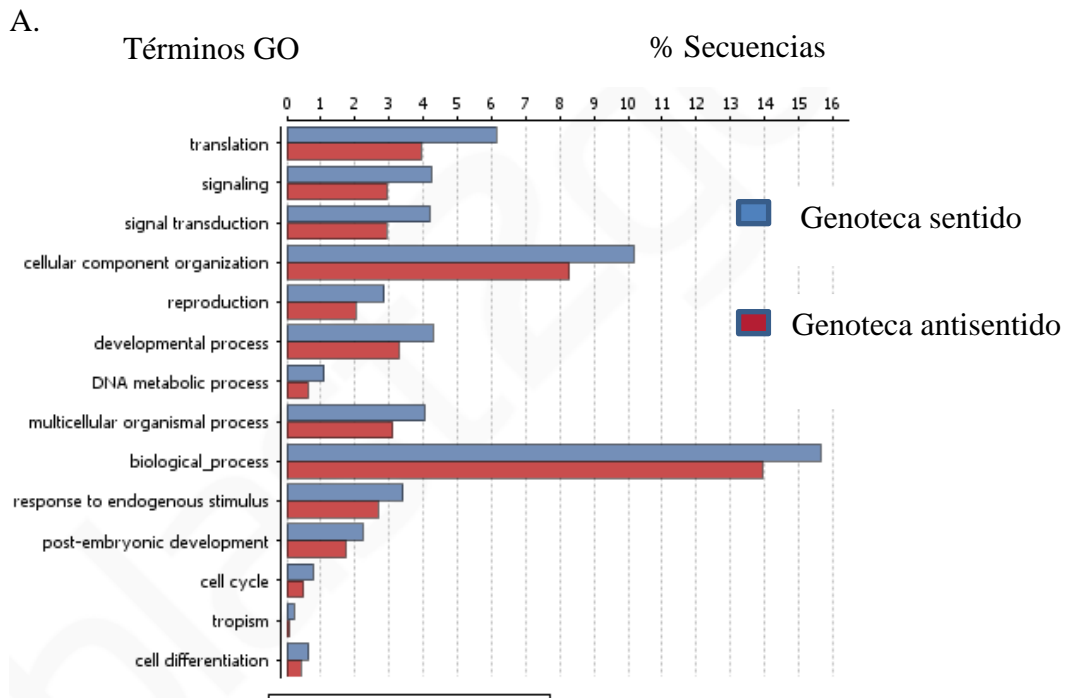


Figura 24. Términos de GO sobre-/sub representados en la categoría proceso biológico De las genotecas sentido (A) y antisense (B)

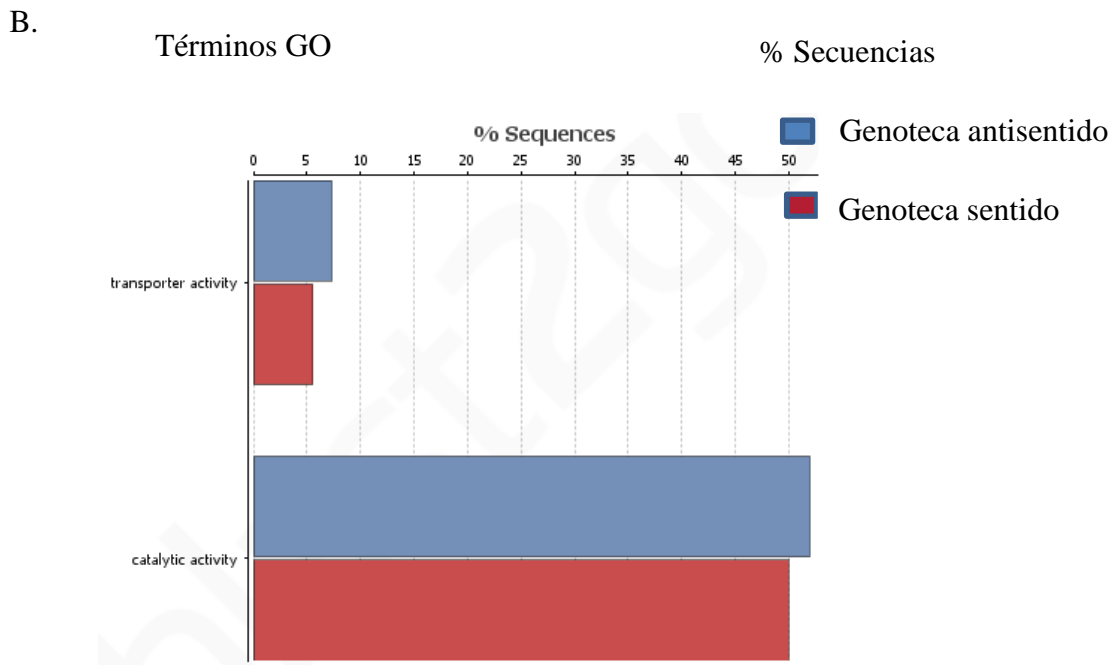
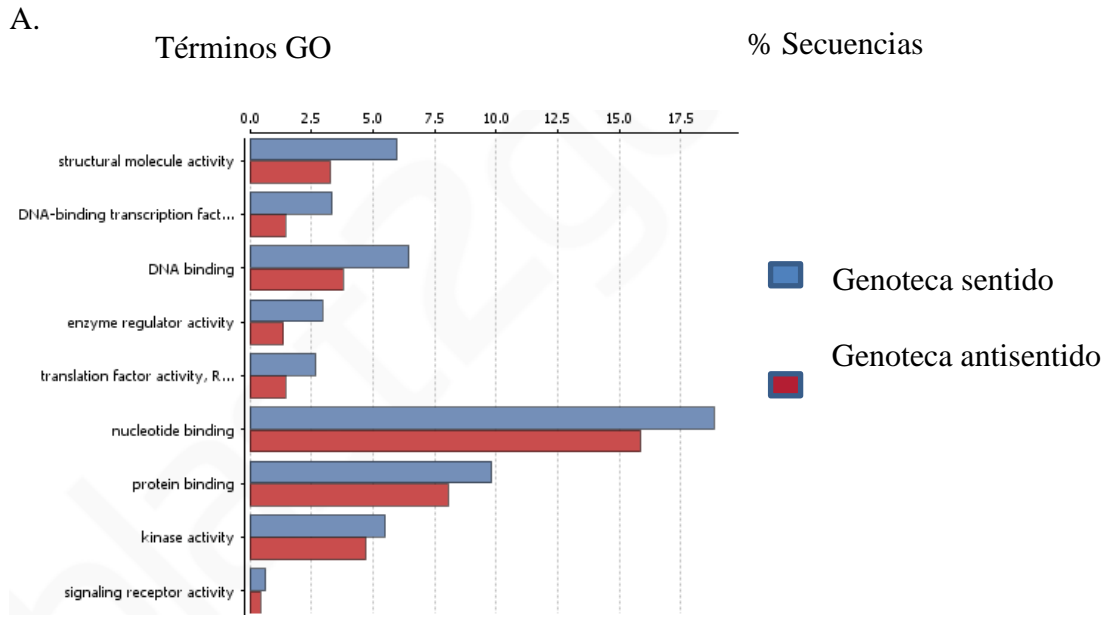


Figura 25. Términos de GO sobre-/sub representados en la categoría función molecular De las genotecas sentido (A) y antisense (B)

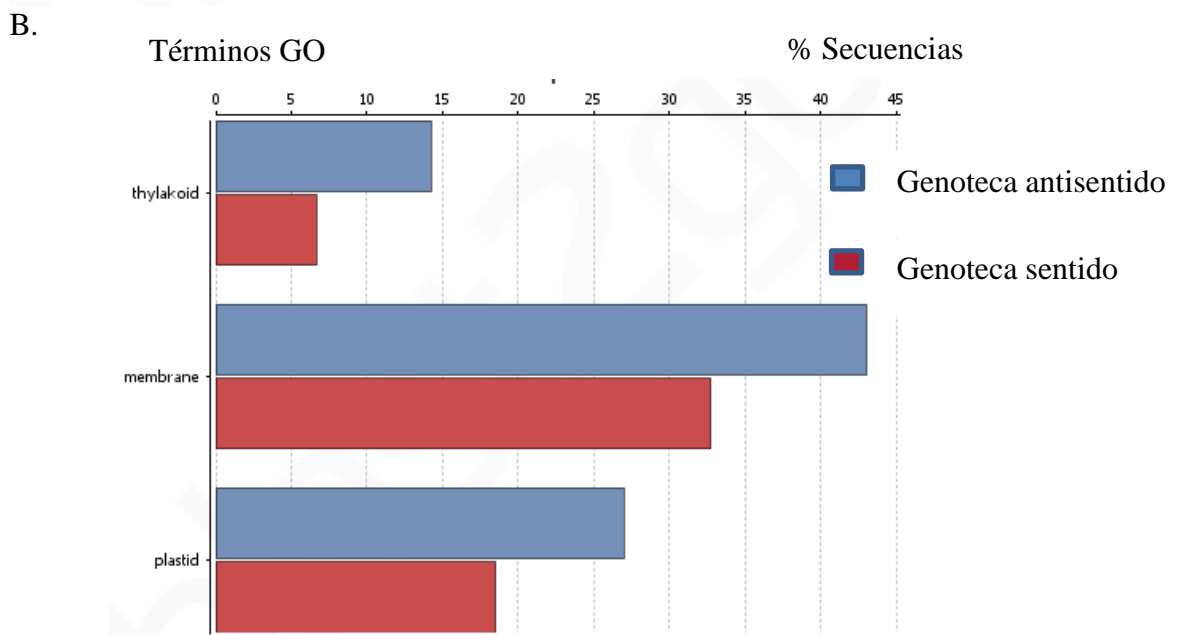
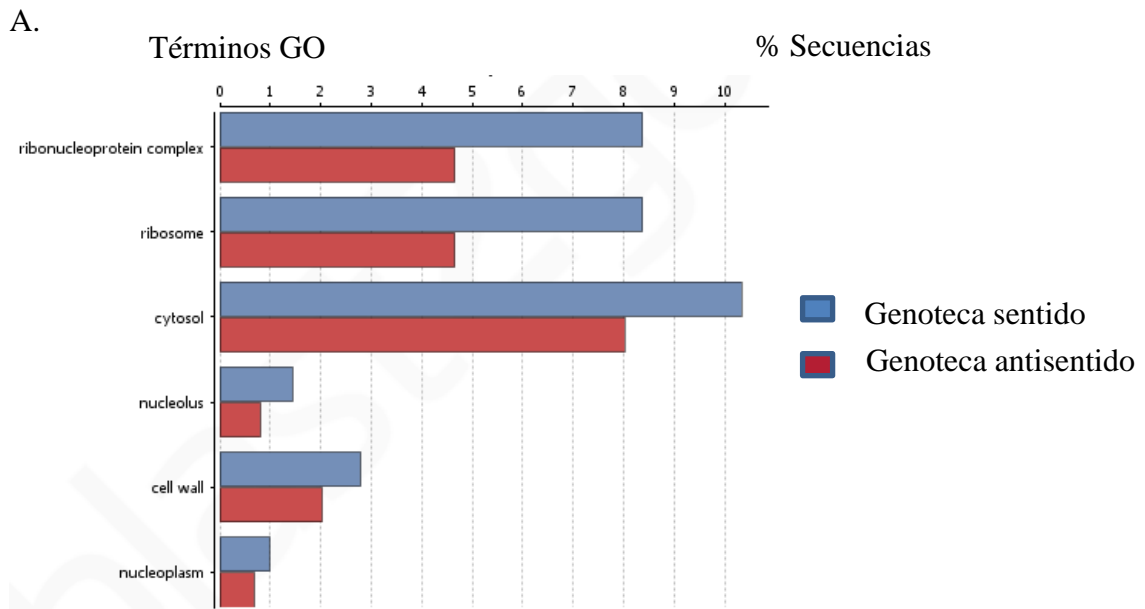


Figura 26. Términos de GO sobre-/sub representados en la categoría componente celular
De las genotecas **sentido** (A) y **antisentido** (B)

4.8 Búsqueda de genes interesantes expresados diferencialmente

Para identificar los conjuntos de genes específicos en las genotecas sustractivas, se realizaron búsquedas de genes de acuerdo a su término GO asociado. La principal búsqueda fue realizada para identificar los genes anotados como factores de transcripción (FTs), para lo cual buscamos los genes que presentan el término GO “actividad de factor de transcripción” con la herramienta BLAST2GO. Entre los 145 genes anotados con ese término GO, 74 en la genoteca **sentido** y 71 en la **antisentido**, hallamos en su mayoría a FTs (Tablas 10 y 11). De acuerdo a la descripción de la secuencia, en ambas genotecas se identificaron a familias de FTs con dominio conocido. WRKY, AP2/ERF, bHLH y C2H2 fueron las familias de dominio conocido que presentaron la mayor cantidad de FTs. De ellas AP2/ERF contiene la mayor cantidad de FTs que se hallaron. Todas las secuencias de los factores de transcripción seleccionados como interesantes de ambas genotecas se muestran en la tabla A5.

Tabla 10. Genes en la genoteca sentido anotados con el término GO: actividad de factor de transcripción

Sequence name	Sequence desc.	Hit ACC	E-Value
F9336_c0_g1_i1	ABSCISIC ACID-INSENSITIVE 5 2	XP_006339213	1.55E-18
F2566_c0_g1_i1	ABSCISIC ACID-INSENSITIVE 5 7	BAL04566	2.43E-43
F632_c0_g1_i1	activator of 90 kDa heat shock ATPase homolog	XP_006362210	2.33E-63
F2545_c0_g1_i1	agamous-like MADS-box AGL27 isoform X2	XP_015170242	8.20E-38
F1092_c0_g1_i1	AP2 ERF and B3 domain-containing transcription factor RAV1	XP_006362535	3.04E-44
F9122_c0_g1_i1	AP2 ERF and B3 domain-containing transcription factor RAV1-like	XP_015074490	2.68E-29
F3295_c0_g1_i1	AUX IAA	NP_001265888	6.18E-67
F4026_c0_g1_i1	auxin response factor 1	XP_006349280	4.02E-177
F2865_c0_g1_i1	auxin response factor 3	XP_006364312	1.71E-113
F1543_c3_g8_i1	B3 domain-containing transcription factor NGA1-like	XP_006366067	1.52E-30

Sequence name	Sequence desc.	Hit ACC	E-Value
F1543_c3_g34_i1	B3 domain-containing transcription factor NGA3-like	XP_006366067	7.12E-33
F6669_c0_g1_i1	basic leucine zipper 23-like	XP_006359698	3.25E-110
F3444_c0_g1_i1	basic leucine zipper 61-like	XP_006348841	2.47E-56
F9328_c0_g1_i1	BES1 BZR1 homolog 4 isoform X1	XP_015067386	2.32E-26
F7680_c0_g1_i1	BES1 BZR1 homolog 4-like isoform X1	XP_006358087	5.14E-67
F6552_c0_g1_i1	BRASSINAZOLE-RESISTANT 1-like	XP_015160841	4.36E-71
F3788_c0_g1_i1	bZIP transcription factor 53-like	XP_006341974	9.06E-66
F9005_c0_g1_i1	bZIP transcription factor partial	XP_015062252	1.40E-30
F203_c0_g1_i1	calmodulin-binding transcription activator 4	XP_006352173	2.62E-58
F502_c0_g1_i1	calmodulin-binding transcription activator 4-like isoform X1	XP_006341901	1.79E-47
F5791_c0_g1_i1	double WRKY type transfactor	NP_001274836	2.01E-38
F3689_c0_g1_i1	ethylene-responsive transcription factor 4	NP_001305628	1.49E-28
F3010_c0_g1_i1	ethylene-responsive transcription factor 4-like	NP_001305628	3.70E-25
F1743_c0_g1_i1	ethylene-responsive transcription factor 7-like	NP_001275619	1.25E-21
F1543_c8_g1_i1	ethylene-responsive transcription factor ABR1-like	XP_006363442	1.61E-13
F8720_c0_g1_i1	ethylene-responsive transcription factor CRF2	XP_006343893	1.64E-75
F8972_c0_g1_i1	ethylene-responsive transcription factor CRF4-like	XP_006360182	7.31E-19
F5801_c0_g1_i1	ethylene-responsive transcription factor ERF118-like	XP_006344128	7.42E-35
F5989_c0_g1_i1	ethylene-responsive transcription factor RAP2-12-like	XP_006342907	0
F1927_c0_g1_i1	FAR-RED ELONGATED HYPOCOTYL 3 isoform X1	XP_006357871	9.05E-43
F509_c0_g1_i1	GATA transcription factor 11-like	XP_006340186	2.66E-146
F3855_c0_g1_i1	GATA transcription factor 15-like	XP_006345974	6.72E-94
F240_c0_g1_i1	GATA transcription factor 1-like	XP_010326807	3.10E-124
F5370_c0_g1_i1	GATA transcription factor 28-like	XP_006351719	1.77E-21
F7330_c0_g1_i1	GATA transcription factor 5-like	XP_015063162	5.65E-48
F6746_c0_g1_i1	GATA transcription factor 8-like	XP_004239665	1.12E-62
F5781_c0_g1_i1	heat shock factor HSF8-like	XP_015160909	1.01E-44
F105_c0_g1_i1	heat stress transcription factor A-4a-like	XP_015168157	0
F3043_c0_g1_i1	heat stress transcription factor A-5	XP_015059195	7.94E-79
F5590_c0_g1_i1	heat stress transcription factor A-8	XP_004247072	8.85E-40
F5622_c0_g1_i1	homeobox-leucine zipper ATHB-13	AAT40488	1.24E-128
F7419_c0_g1_i1	homeobox-leucine zipper HAT22-like	NP_001311712	2.16E-24
F9083_c0_g1_i1	homeobox-leucine zipper HAT4-like	XP_006367464	1.13E-116
F5809_c0_g1_i1	probable WRKY transcription factor 17	XP_006342466	3.86E-107
F6602_c0_g1_i1	probable WRKY transcription factor 26	ABI95141	4.55E-11
F1935_c0_g1_i1	probable WRKY transcription factor 3	NP_001305593	5.72E-147
F2719_c0_g2_i1	probable WRKY transcription factor 33	AEA86307	8.45E-99
F1440_c0_g6_i1	probable WRKY transcription factor 40	XP_006356251	4.14E-111
F6527_c0_g1_i1	rab11 family-interacting 3 isoform X2	XP_006355873	1.74E-43

Sequence name	Sequence desc.	Hit ACC	E-Value
F1750_c0_g1_i1	RNA polymerase sigma factor sigB-like	XP_006359117	1.75E-67
F8292_c0_g1_i1	RNA-directed DNA methylation 4 isoform X3	XP_015165721	9.65E-23
F255_c0_g1_i1	scarecrow 1	KJB29093	1.18E-117
F3899_c0_g1_i1	scarecrow 14	XP_006350781	1.11E-93
F3578_c0_g1_i1	scarecrow 15	XP_006363143	1.68E-117
F139_c0_g2_i1	scarecrow 30	XP_015058087	8.51E-23
F139_c0_g1_i1	scarecrow 30 isoform X1	XP_015058083	4.04E-23
F9090_c0_g1_i1	scarecrow 32	XP_006339936	1.05E-111
F8064_c0_g1_i1	SHORT-ROOT	XP_006363828	2.09E-50
F7764_c0_g1_i1	single-stranded DNA-binding chloroplastic	NP_001275155	0
F114_c0_g3_i1	threonine dehydratase chloroplastic	XP_006366786	2.25E-91
F402_c0_g1_i1	transcription factor BEE 3-like	XP_006341598	5.29E-105
F7766_c0_g1_i1	transcription factor bHLH110	XP_006341970	2.18E-79
F6535_c0_g1_i1	transcription factor bHLH113-like	XP_015083639	2.93E-63
F8036_c0_g1_i1	transcription factor bHLH122-like	XP_006364245	1.98E-49
F4835_c0_g1_i1	transcription factor bHLH130-like	XP_015069400	1.47E-25
F3220_c0_g1_i1	transcription factor bHLH68	XP_006341272	7.24E-71
F3283_c0_g1_i1	transcription factor HBI1-like	XP_015067911	4.35E-62
F6409_c0_g1_i1	transcription factor HBP1-b(c38)-like	XP_010313245	1.38E-90
F1117_c0_g1_i1	transcription factor RF2b-like	XP_006357296	1.15E-148
F7546_c0_g1_i1	transcription factor VIP1-like	XP_006363906	2.10E-118
F3814_c0_g1_i1	truncated transcription factor CAULIFLOWER A-like	AGT21640	4.19E-77
F7129_c0_g1_i1	WRKY transcription factor 6-like	XP_006350473	5.81E-29
F1168_c0_g1_i1	zinc finger ZAT10-like	XP_006342662	1.05E-111
F2212_c0_g1_i1	zinc finger ZAT5-like	XP_006347206	1.37E-131

Tabla 11. Genes en la genoteca antisentido anotados con el término GO: actividad de factor de transcripción

Sequence name	Sequence desc.	Hit ACC	E-Value
R3293_c0_g1_i1	ABSCISIC ACID-INSENSITIVE 5 5	NP_001274925	1.20E-119
R1297_c0_g1_i1	agamous-like MADS-box AGL8 homolog	NP_001275142	0
R484_c0_g1_i1	auxin-responsive IAA8-like	NP_001265888	1.25E-111
R5410_c0_g1_i1	BES1 BZR1 homolog 2-like	XP_004233505	3.84E-18
R633_c0_g1_i1	bZIP transcription factor 60	XP_006342192	2.17E-82
R544_c0_g1_i1	CHAPERONE-LIKE PROTEIN OF chloroplastic	XP_015070515	1.46E-45
R1168_c0_g1_i1	Chlorophyll A-B binding family	XP_006362479	3.14E-79
R7789_c0_g1_i1	dehydration-responsive element-binding 3-like	XP_006359762	2.29E-22
R515_c0_g1_i1	DELLA GAI	NP_001305514	3.13E-64
R245_c0_g1_i1	ethylene-responsive transcription factor 5-like	XP_015169356	1.98E-73

Sequence name	Sequence desc.	Hit ACC	E-Value
R3627_c0_g1_i1	ethylene-responsive transcription factor ERF106-like	NP_001234313	5.27E-57
R2254_c0_g1_i1	ethylene-responsive transcription factor RAP2-13-like	XP_006355363	5.64E-69
R4916_c0_g1_i1	GATA transcription factor 16-like	XP_015059269	3.73E-28
R7560_c0_g1_i1	GATA transcription factor 22	XP_006346565	1.02E-26
R2566_c0_g1_i1	GATA transcription factor 24-like	XP_006351719	8.84E-59
R8272_c0_g1_i1	GATA transcription factor 24-like isoform X2	XP_009612623	4.03E-29
R3958_c0_g1_i1	GATA transcription factor 26-like	XP_006362478	6.60E-48
R4350_c0_g1_i1	G-box-binding factor 1-like isoform X1	XP_006346935	3.14E-23
R7947_c0_g1_i1	G-box-binding factor 4-like isoform X2	XP_015170057	1.37E-85
R423_c0_g1_i1	heat stress transcription factor C-1	XP_006353456	3.13E-165
R8551_c0_g1_i1	homeobox-leucine zipper ATHB-12	XP_006339757	7.17E-30
R984_c0_g1_i1	homeobox-leucine zipper ATHB-12-like	XP_006367352	1.54E-111
R3153_c0_g1_i1	homeobox-leucine zipper ATHB-6-like	XP_015158857	6.10E-100
R8478_c0_g1_i1	homeobox-leucine zipper HAT5	XP_006338548	2.71E-13
R462_c0_g1_i1	homeobox-leucine zipper HAT5-like	XP_006367068	6.10E-66
R3829_c0_g1_i1	light-inducible CPRF2	XP_015169057	2.14E-177
R6261_c0_g1_i1	MADS-box SOC1-like	XP_006339502	2.60E-38
R2870_c0_g1_i1	MPF1 partial	XP_006351733	4.61E-112
R7155_c0_g1_i1	MPF2-like- partial	XP_010320206	1.75E-57
R7929_c0_g1_i1	NF-X1-type zinc finger NFXL1	XP_006341494	3.82E-56
R4069_c0_g1_i1	nuclear transcription factor Y subunit A-1-like isoform X1	XP_006345444	4.13E-70
R1259_c0_g4_i1	nuclear transcription factor Y subunit A-3-like	XP_006351613	7.82E-105
R3490_c0_g1_i1	pathogenesis-related genes transcriptional activator PTI6-like	XP_006350890	2.86E-30
R8485_c0_g1_i1	probable transcription factor 21	XP_015165666	2.62E-18
R6221_c0_g1_i1	probable WRKY transcription factor 2 isoform X1	NP_001305535	1.19E-38
R7861_c0_g1_i1	probable WRKY transcription factor 20	NP_001305542	1.37E-29
R7861_c1_g1_i1	probable WRKY transcription factor 20 isoform X2	AAC49529	1.87E-50
R5929_c0_g1_i1	probable WRKY transcription factor 31	XP_006364165	4.05E-85
R2704_c0_g1_i1	probable WRKY transcription factor 31 isoform X1	NP_001305626	7.17E-70
R8164_c0_g1_i1	probable WRKY transcription factor 4	XP_006350164	8.83E-130
R547_c0_g1_i1	probable WRKY transcription factor 41	XP_006339686	2.20E-71
R7714_c0_g1_i1	probable WRKY transcription factor 51	XP_015160883	1.72E-51
R3422_c0_g1_i1	probable WRKY transcription factor 65 isoform X1	XP_015055738	2.39E-24
R6925_c0_g1_i1	probable WRKY transcription factor 65 isoform X2	NP_001274847	8.57E-28
R2867_c0_g1_i1	probable WRKY transcription factor 70	XP_015086412	1.56E-61
R3540_c0_g1_i1	rab11 family-interacting 3 isoform X1	XP_006355873	9.51E-30
R1382_c0_g1_i1	scarecrow 21	XP_006363507	6.96E-173
R7937_c0_g1_i1	scarecrow 22	XP_015062026	3.89E-113

Sequence name	Sequence desc.	Hit ACC	E-Value
R1140_c0_g1_i1	scarecrow 8	XP_006353907	7.97E-129
R8584_c0_g1_i1	scarecrow-like transcription factor PAT1	ABD72963	1.07E-74
R2599_c0_g1_i1	TGACG-sequence-specific DNA-binding TGA-	XP_006351371	1.46E-43
R7225_c0_g1_i1	transcription factor	XP_006350164	4.22E-67
R7568_c0_g1_i1	transcription factor bHLH112-like	XP_015063418	2.27E-109
R6730_c0_g1_i1	transcription factor bHLH47-like	XP_006364187	1.08E-71
R8043_c0_g1_i1	transcription factor bHLH84-like	XP_015055483	1.98E-61
R174_c0_g1_i1	transcription factor IIIA	XP_006364766	0
R4537_c0_g1_i1	transcription factor LHW-like	XP_006362846	6.43E-62
R1883_c0_g1_i1	transcription factor MYB1R1-like	XP_006344987	2.69E-78
R6365_c0_g1_i1	transcription factor TGA1-like	XP_006354049	7.60E-158
R5007_c0_g1_i1	transcription factor UNE12-like	ABB16991	4.10E-47
R4452_c0_g1_i1	transcription initiation factor TFIID subunit 2 isoform X3	XP_015069896	2.57E-85
R7812_c0_g1_i1	trihelix transcription factor GT-1-like isoform X2	XP_004252697	9.48E-72
R1_c0_g1_i1	truncated transcription factor CAULIFLOWER A-like isoform X1	XP_015064042	2.80E-67
R1_c0_g2_i1	truncated transcription factor CAULIFLOWER A-like isoform X2	XP_006338337	7.86E-70
R6824_c0_g1_i1	tubby-like F-box 7	XP_015082792	9.46E-57
R3610_c0_g1_i1	two-component response regulator ARR12-like	XP_006343681	2.29E-42
R4926_c0_g1_i1	two-component response regulator ARR1-like	XP_004229281	2.30E-66
R406_c0_g1_i1	zinc finger BED domain-containing DAYSLEEPER-like	XP_015165932	1.94E-78
R5130_c0_g1_i1	zinc finger BED domain-containing RICESLEEPER 1-like	XP_015170144	1.97E-47
R4599_c0_g1_i1	zinc finger NUTCRACKER-like [Tarenaya hassleriana]	XP_006357381	5.63E-115

Se realizó una segunda búsqueda para hallar genes claves que participan en las rutas de señalización en el proceso de transducción de la señal del estrés. Es sabido que muchas de las rutas de señalización de estímulo estrés es activada por proteínas quinasas dependiente de Ca^{2+} (CDPKs), proteínas quinasas activadas por mitógeno (MAPKs), calcineurina, proteína de unión a calcineurina B (CBL), calmodulina, entre las más conocidas. En este estudio identificamos a través de la búsqueda por el término GO: “actividad quinasa”, a 8 CDPKs (5 en genoteca **sentido** y 3 en **antisentido**), 5 MAPKs (4 en genoteca **sentido** y 1 en **antisentido**), 3 calcineurina (1

en genoteca sentido y 2 en antisentido), y 2 calmodulinas (1 en genoteca **sentido** y 1 en **antisentido**), (tabla anexo A1 y A2).

Se realizó una tercera búsqueda para identificar, en general, genes asociados a la tolerancia a heladas de acuerdo a búsquedas por su término GO “respuesta al estrés abiótico”. Se identificaron 211 genes (76 en genoteca **sentido** y 135 en **antisentido**), varios de esos genes presentan función conocida en la maquinaria de respuesta al estrés, entre ellos, genes de proteínas que participan en las cascadas de antioxidantes, en la transducción de señales en la célula, en el plegamiento de proteínas, metabolitos secundarios, así como algunos FTs, entre los cuales se identificaron genes que fueron hallados anteriormente en la primera y segunda búsqueda (Tabla A3 y A4).

5 DISCUSIÓN

5.1 Las genotecas de ADNc obtenidas por SSH

Las respuestas adaptativas de las plantas frente al estrés abiótico involucran cambios moleculares. Investigar sobre la expresión de los genes implicados en esos cambios contribuye a comprender el mecanismo de tolerancia al estrés. En la actualidad se cuenta con diversas metodologías para explorar la expresión de genes; en este trabajo se optó por la construcción de genotecas por la metodología de la hibridación sustractiva por supresión (SSH), implementada por Diatchenko (2000). SSH es una técnica poderosa para obtener información de los genes expresados específicamente en una de dos condiciones biológicas contrastantes, además permite aislar transcritos raros (aquellos con muy bajo número de copias) ya que incluye un paso de normalización de la abundancia de los transcritos (Gurskaya *et al.*, 1996).

Debido a que se obtiene un número restringido de genes se puede realizar un mayor análisis (Diatchenko, 1996). Entre las principales aplicaciones para las que se usa esta técnica está la identificación de cambios en la expresión génica durante procesos de infección o para evaluar las diferencias entre tejidos contrastantes (Ji *et al.*, 2002). Ha sido bastante usado para estudiar los mecanismos moleculares de las plantas en estreses bióticos y abióticos. El artículo de revisión de Sahebi *et al.*, 2015, muestra diferentes trabajos en plantas realizados con SSH, en su mayoría para la captura de genes expresados antes distintos tipos de estrés. Como principales ventajas que se le adjudica a la técnica están la de no necesitar, *a priori*, información de la secuencia genómica, punto clave que hicieron que técnicas como las del microarreglo (que era

la que se utilizaba anteriormente normalmente para identificar genes expresados diferencialmente) ya no sean considerados en la actualidad; además, de no requerir análisis de alto rendimiento bioinformático (ya que esta técnica normalmente se usa para clonar y obtener directamente la información del gen), característica que la hace ventajosa frente al análisis que se necesita de la gran información obtenida por RNA-seq; sin embargo, esta última tiene la ventaja de mostrar muchos más genes expresados diferencialmente y calcular aproximadamente la diferencia de su expresión (Sahebi *et al.*, 2015, Wang *et al.*, 2009).

En el presente trabajo se utilizó esta metodología para obtener dos genotecas recíprocas con información de los genes expresados diferencialmente en hojas de *Solanum juzepzuckii*, bajo tratamiento en condiciones de helada comparándola con tratamiento normal. Se lograron obtener dos genotecas con el kit de Clontech® PCR-Select™ cDNA Subtraction (Clontech, EE.UU./Canada). En la figura 15 se observa un gel de electroforesis con el barrido de ADNc provenientes del procedimiento de sustracción de ambas genotecas, en el que se puede apreciar que se obtuvieron zonas normalizadas de la abundancia de ADNc, así como el enriquecimiento de otras zonas, lo que es una señal del éxito del procedimiento de los pasos de normalización y sustracción. Los resultados obtenidos en el gel son similares al de otros estudios que han empleado la misma metodología para evaluar genes expresados diferencialmente bajo estrés en plantas como *Capsicum chinense* (Gasca-González *et al.*, 2008), y *Haloxylon ammodendron* (Jiang *et al.*, 2004).

5.2 Secuenciamiento y ensamblaje

Para revelar los genes sustraídos de las genotecas SSH, luego de la hibridación sustractiva, tradicionalmente se usa la técnica del clonamiento; la cual resulta laboriosa y limitada ya que se necesita lograr un clonamiento exitoso por cada transcrito para obtener la información del gen. En este trabajo se decidió utilizar el secuenciamiento masivo con tecnología 454 FLX, con el objetivo de obtener mayor alcance en la identificación de los genes. La combinación de la técnica SSH y el secuenciamiento masivo ha sido utilizado en otros trabajos y con buenos resultados; por ejemplo los realizados en las siguientes especies: soya (Rodrigues *et al.*, 2012), molusco vieira (Llera-Herrera *et al.*, 2013) y salmón (Olsvik *et al.*, 2013); sin embargo nuestros resultados demuestran que si bien se obtuvo un mayor alcance de secuencias, también se observó algunas secuencias en común en ambas genotecas, debiendo identificarlas necesariamente antes de determinar los grupos de genes interesantes. Estos resultados se pueden deber a que las genotecas son productos del enriquecimiento de transcritos expresados diferencialmente en dos muestras de ADNc, más no de la eliminación de los que se encuentran en común, por lo que el secuenciamiento masivo logró captarlos, así hallan estado en muy baja concentración. Llera-Herrera *et al* (2013) también han reportado el hallazgo de algunas secuencias en común en genotecas SSH recíprocas secuenciadas.

Con la finalidad de no limitar el ensamblaje de las secuencias de las genotecas sustractivas con el transcriptoma del genoma del papa ya secuenciado, el cual fue realizado a partir de la especie *Solanum tuberosum* Grupo phureja (Massa *et al.*,

2011), se decidió realizar el ensamblaje sin utilizar transcriptoma de referencia, es decir un ensamblaje *de novo*. El ensamblaje se realizó con el ensamblador Trinity que normalmente se usa para ensamblar los datos del secuenciamiento por Illumina; éste como los otros ensambladores de mayor popularidad, están basados en los Grafos de De Bruijn, que ha permitido realizar el proceso con mucha mayor eficiencia. (Grabherr *et al.*, 2011), Existen pocos ensambladores de este tipo para la plataforma de Roche 454, los que se usa normalmente son Newbler y MIRA. El trabajo de Ren *et al.*, (2012) compara el trabajo de varios ensambladores entre los que se encuentra Trinity y Newbler para el tratamiento de las lecturas de la plataforma 454. Sus resultados demuestran que algunos ensambladores basados en Grafos De Bruijn obtuvieron malos resultados con los datos transcriptómicos 454, pero otros, especialmente Trinity obtuvieron incluso mejores resultados que Newbler y MIRA, mostrándose como una solución alternativa para que los investigadores analicen con mayor eficiencia los datos transcriptómicos 454.

Luego del secuenciamiento y ensamblaje de las dos genotecas, se observó que ambas presentaron similar cantidad de lecturas ensamblados, sugiriendo por lo tanto, que la cantidad de transcritos diferencialmente expresados fue similar. De acuerdo a este resultado, la búsqueda de las diferencias de la expresión entre ambas genotecas se centró en la identificación de los GO enriquecidos y de los transcritos que lo conforman. Un trabajo que refuerza este hallazgo fue el reportado por Carvallo *et al.*, (2011). Los autores compararon el transcriptoma, en condiciones de baja temperatura, de *Solanum tuberosum* y *Solanum comersonii*, en búsqueda de diferencias a macroescala, que puedan asociarse a la marcada distinta capacidad que tienen para

aclimatarse. Ellos determinaron que la expresión génica en ambas especies se daba con similar grado y cinética. Además, en ambas especies había regulones CBF (genes controlados por FTs CBF) integrado con cientos de genes. De acuerdo a sus resultados, los autores proponen que la diferencia de la capacidad para aclimatarse podría deberse más bien al grupo de genes que integraban el transcriptoma a bajas temperaturas de ambas especies y no a diferencias a nivel de macroescala, como la cantidad de transcritos diferentes.

5.3 GOs enriquecidos en las genotecas de ADNc

De acuerdo con las anotaciones y al análisis de enriquecimiento GO con la prueba exacta de Fisher, se observa curiosamente que entre los GOs enriquecidos, con mayor significancia, se encuentran, en la genoteca **sentido**, a: actividad de factor de transcripción - unión a ADN de secuencia específica (FM) con $FDR = 3.6E^{-19}$ y actividad de factor de traducción (FM) con $FDR = 1.2 E^{-9}$; y en la **antisentido** a: fotosíntesis (BP) con $FDR = 6.0E^{-78}$, proceso catabólico (BP) con $FDR = 1.2E^{-5}$ y proceso metabólico de carbohidratos (BP) con $FDR = 7.9E^{-10}$. Es decir, se observa que existe una evidente reprogramación genética con una regulación positiva en la transcripción y traducción de genes y también una alteración en la maquinaria de producción de energía, con una regulación negativa en los componentes de la fotosíntesis. En cuanto a este último punto, se sabe que, en la respuesta fisiológica temprana frente a un estrés abiótico en general, se presenta un tipo de cambio en común: el estatus bioenergético de la planta; tales cambios pueden involucrar

crecimiento o decrecimiento en la demanda de producción de energía para enfrentar el estrés (Duque *et al.*, 2013). En relación a estos procesos, el estrés abiótico tiene un impacto directo en la actividad de todos los componentes de la fotosíntesis y la rápida respuesta del metabolismo y la adaptación de la maquinaria fotosintética al estrés son factores claves para la sobrevivencia de las plantas (Nouri *et al.*, 2015). Se sabe también que el estrés abiótico, específicamente por frío, afecta el aparato de la fotosíntesis a través de la inactivación de los fotosistemas, alteraciones moleculares de las membranas de tilacoides, de pigmentos, de la actividad enzimática relacionados con la fotosíntesis, en el intercambio de gases, en la fluorescencia de clorofila, en la disminución de la asimilación de CO₂ y en la tasa de transporte de electrones en las plantas (Kazemi-Shahandashti y Maali-Amiri, 2018).

La fotosíntesis es consecuencia de múltiples pasos que involucran distintas rutas biológicas. Sus componentes están bajo el control de muchos genes que son codificados tanto en el genoma nuclear como en el cloroplasto. La expresión en ambos genomas es variable y afectada por diversos rangos de factores ambientales (Berry *et al.*, 2013), y como se observa en nuestros resultados, el impacto se da a nivel de la regulación transcripcional. En un resultado similar al nuestro, un reporte con análisis de GO, demuestra que en líneas de arroz el ser tratadas a 10 °C por 72 h hace que los genes implicados en la fotosíntesis se expresen diferencialmente negativamente (Sharma *et al.*, 2005).

5.4 Búsqueda y selección de FTs interesantes

En cuanto a la exploración de los FTs presentes en ambas genotecas, la búsqueda por el término de GO “actividad de factor de transcripción” con la herramienta de BLAST2GO, logró hallar 145 *contigs* con anotaciones diferentes en ambas genotecas, de los cuales, en su mayoría cuentan con la descripción de factor de transcripción. Se sabe que entre las familias de FTs con dominio conocido, y que se han reportado que participan en la respuesta a estrés abiótico, se tiene a las siguientes principales: bZIP (basic leucine zipper), AP2/ERF, (APETALA 2/ethylene-responsive element binding factor), NAC (NAM/ATAF1/CUC₂), WRKY, MYB, C2H2 (Cys2 (C2) His2 (H2)-type zinc fingers) y bHLH (basic helix-loop-helix), Lindemose *et al.*, 2013. Las familias que presentaron la mayor cantidad de FTs en las genotecas fueron: WRKY, AP2/ERF, bHLH y C2H2. (Tablas 8 y 9). De ellas, AP2/ERF contiene la mayor cantidad de FTs que se hallaron (9 en la genoteca **sentido** y 7 en la **antisentido**). Se observa además que la cantidad de FTs por cada familia es similar en ambas genotecas, por lo que la diferencia se mostró, en sí, en la identidad de los FTs que lo conforman.

En la selección de FTs interesantes asociados a la tolerancia al frío o a algún estrés osmótico, para una futura evaluación de su expresión, se tomó en cuenta que sean parte de las familias de FTs con dominios conocidos mencionados anteriormente, además de contar con algún reporte de ese FT y/o de su homólogo en *Arabidopsis thaliana*, que lo asocie a la tolerancia frente a algún estrés abiótico. En el caso de los FTs de la genoteca **sentido** se halló reportes con las metodologías de transgénesis y/o

mutaciones con las que se comprobaba la participación del FT en la respuesta de tolerancia a sequía (inducido por manitol) y/o salinidad (inducido por altas concentraciones de NaCl), pero no por frío. En el caso de los FTs de la genoteca **antisentido** solo se encontraron reportes de su acción negativa en la regulación por ABA, hormona crucial en la respuesta al estrés abiótico. Este criterio de selección va de acuerdo a la noción que se tiene sobre la participación de un mismo FT ante varias respuestas de tolerancia, ya que los cultivos en el campo están sometidos a distintos tipos de estrés abiótico y biótico.

En un evento de helada a elevada altura los principales distintos tipos de estrés abiótico que estarían afectando a las plantas son: la baja temperatura, la escasez de agua, así como la poca cantidad de oxígeno presente en el ambiente. La respuesta de tolerancia frente a los dos primeros se asemeja ya que comparten el efecto del estrés osmótico subyacente (Huang *et al.*, 2012). En cuanto a la respuesta a la hipoxia, por ahora se ha planteado que las plantas estarían respondiendo a través de dos mecanismos a nivel primario. El primero que concierne a la producción y *scavenging* de ROS (reactive oxygen species), que si bien es, en cierta concentración, perjudicial a la célula, juega un rol crucial en la transducción de la señal regulando la defensa a nivel sistémico (Czarnocka y Karpiński, 2018). Por otro lado se plantea la activación de genes de tolerancia, luego de medirse la cantidad de oxígeno a través de una modificación post-traducciona de un grupo de FTs de la familia ERF, en la que la presencia de un motivo conservado en la región N-terminal de estos FTs, los hace blanco de la degradación bajo condiciones de normoxia, ruta conocida como NERP

(N-end rule pathway). Frente a la hipoxia estos FTs están estables y activarían los genes de la respuesta de tolerancia. (Pucciariello y Perata, 2012).

En este trabajo se logró recopilar información asociada a la tolerancia frente algún estrés abiótico de los siguientes genes putativos de la genoteca **sentido**: WRKY26, WRKY33, WRKY40, ERF7, CRF4, RAP2.12, BHLH122 y BHLH68; y de la **antisentido**: WRKY41. En la mayoría de los casos, los reportes demuestran que estos FTs están implicados en la tolerancia mediante métodos de sobreexpresión o mutación de los genes. Estos reportes junto a otros que brindan información sobre las proteínas con las que estarían interactuando estos genes se detallan a continuación, de acuerdo a la familia a la que pertenecen:

FTs WRKY.- En el trabajo de Li *et al.*, (2011) analizaron los roles de WRKY25, WRKY26 y WRKY33 en la regulación de la resistencia al estrés por calor demostrando que la regulación cruzada entre esos FTs era esencial para promover la tolerancia al estrés por altas temperaturas, mediante trabajos de simple, doble y triple mutación (Figura 27). Mao *et al.*, 2011 demostraron además que la proteína WRKY33 regula su propia expresión a través de una retroalimentación positiva mediante la unión a su propio promotor.

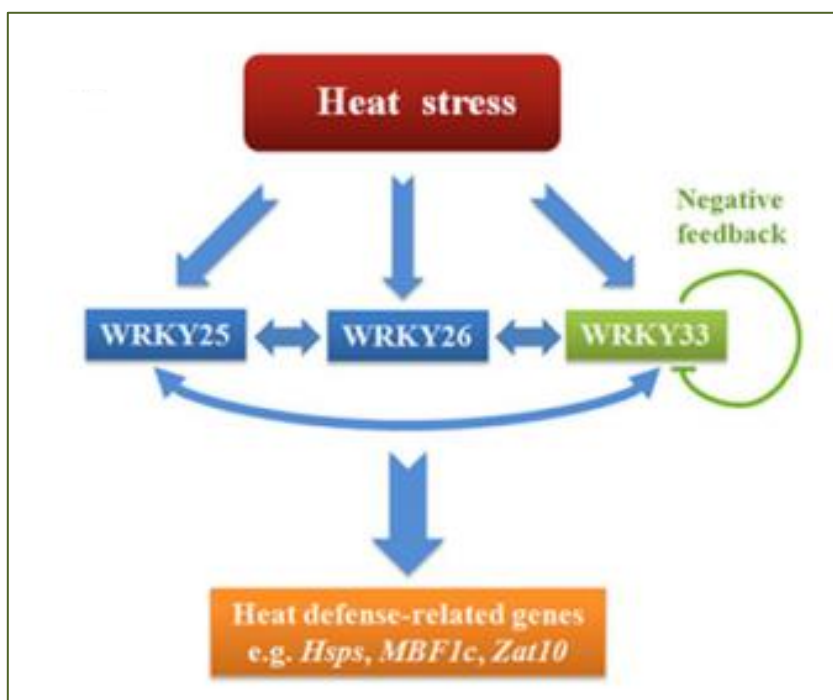


Figura 27. Modelo propuesto de interacciones de WRKY26 y WRKY33
 Fuente: Li *et al.*, 2011

Por otro lado, Xu *et al.*, 2006, demostraron la interacción física y funcional que tienen los FTs WRKY40, WRKY18 y WRKY60; más tarde Chen *et al.*, 2010, observaron que WRKY18 y WRKY60 aumentan la sensibilidad de la planta a la sequía, y que WRKY40 antagoniza los efectos de esos genes, cumpliendo el papel de un represor transcripcional a favor de la tolerancia al estrés por sequía, la manera cómo estarían actuando se muestra en la figura 28 en la que WRKY40 forma homodímeros y heterodímeros con los otros dos genes, y estando en la segunda forma estaría actuando como un represor. Estos 3 genes además son vistos como candidatos de la regulación cruzada que existe entre la respuesta del estrés biótico y abiótico ya que participan en las rutas reguladas por SA (ácido salicílico), JA (ácido jasmónico) y

ABA, las dos primeras hormonas conocidas por ser las principales hormonas involucradas en la transducción de la respuesta contra el estrés biótico y la última en el estrés abiótico (Banerjee y Roychoudhury, 2015).

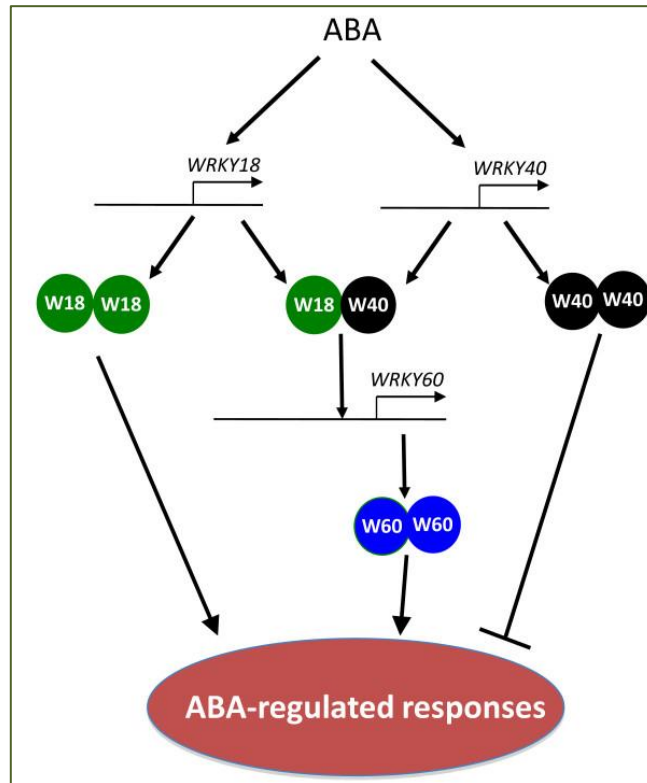


Figura 28. Modelo propuesto de interacciones de WRKY40 con otros FTs
Fuente: Chen *et al.*, 2010

Otro reporte que muestra la implicancia de WRKY40 en la tolerancia a estrés abiótico es el de Su *et al.*, 2016. En este trabajo se muestra que WRKY40 es un regulador negativo del gen ABI5, el cual regula directamente y negativamente la expresión de ABR (ABA RESPONSIVE PROTEIN) en condiciones de oscuridad. ABR es un gen que codifica una proteína LEA (LATE EMBRYOGENESIS ABUNDANT), cual es un gen de respuesta que está asociada con el rescate de la hoja en el proceso de

senescencia. También observaron que ABR es regulado positivamente en tratamientos con ABA, NaCl, manitol y oscuridad.

Para el caso de los FTs WRKY de la genoteca **antisentido**, Ding *et al.*, 2014 demostraron que WRKY41 actúa como un regulador negativo en la señalización de ABA, al analizar la germinación de semillas de *Arabidopsis*. ABA es la hormona del estrés universal y las proteínas WRKY asociadas con la señalización de ABA definitivamente puede influir en las respuestas inducidas por el estrés.

FTs AP2/ERF.- En referencia a los genes hallados en la genoteca **sentido**, Song, *et al.*, 2005 a base de trabajos de sobreexpresión y de ARN interferencia demostraron que ERF7 juega un importante rol en la respuesta con ABA, y puede formar parte de un complejo de represión transcripcional regulado por la proteína quinasa PKS3, el cual es un regulador global positivo de respuestas a ABA.

Zwack *et al.*, 2016 demostraron que el Factor de respuesta de citoquinina 4 (CRF4) muestra una inducción de su expresión en un plazo corto por frío (4 °C), y el patrón de su expresión, junto al análisis de mutantes llevó a la hipótesis de que la inducción por frío de CRF4 podría desempeñar un papel en la aclimatación en frío a corto plazo que conduce a una mayor tolerancia a la congelación. Los autores a partir de la información sobre la expresión alterada de genes de la ruta CBF y genes de la señalización de COR15A, pero no de ICE 1; sugieren que CRF4 estaría uniéndose a los elementos CIS de esos genes, sin participar del regulón de ICE1, actuando así como un positivo regulador de la tolerancia al frío (Figura 29).

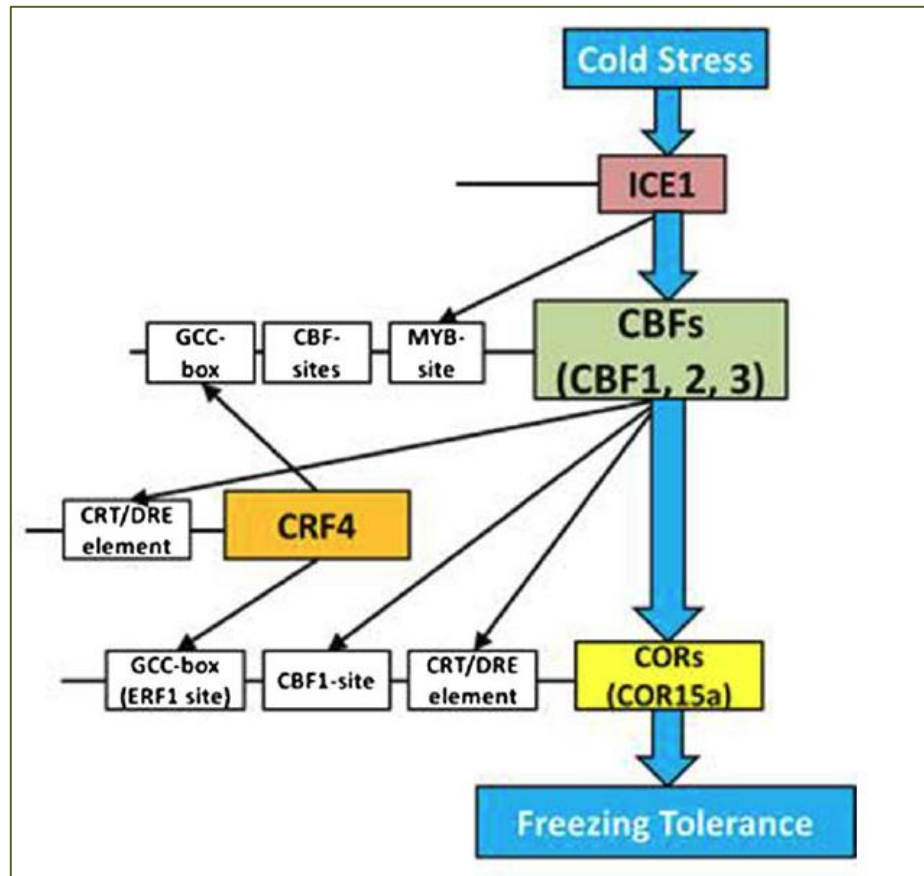


Figura 29. Modelo propuesto de interacciones de CRF4 con genes DE CBF y CORs
 Fuente: (Zwack *et al.*, 2016)

Papdi, *et al.*, 2015 mostraron que la expresión inducible de RAP2-12 junto a RAP2.2, RAP2.3 confirió tolerancia a anoxia, estrés oxidativo, estrés por sequía, y mejoró la sensibilidad al ABA. Esta proteína junto a RAP2.2, RAP2.3, HRE1, Y HRE2, son los 5 FTs miembros caracterizados de la familia ERF del grupo VII. Los TFs RAP2.2 y Rap2.12 han sido sugeridos para activar la respuesta anaeróbica a través de la ruta NERP, ruta comentada líneas arriba. Este último además controla la expresión de diferentes genes marcadores de hipoxia como PDC1, ADH1, HB1, LBD41, HUP7, HRA1 y PCO2 (Licausi *et al.*, 2011).

FTs bHLH.- Liu, *et al.*, 2014 mostraron que plantas transgénicas que sobreexpresaban bHLH122 presentaban mayor resistencia a sequía y salinidad, además el mutante con pérdida de función de bHLH122 era más sensible a esos factores comparado con la planta silvestre. Los mismo autores sugieren que bHLH122 estaría actuando como un un elemento *crossstalk* entre la señal independiente y dependiente de ABA. En cuanto a la ruta dependiente de ABA proponen un modelo de acción en la que ejercen acción sobre el gen CYP707A3.

Le Hir *et al.*, (2017) demostraron que el transgénico bHLH68 demuestra un incremento de tolerancia al estrés por sequía además que su expresión era modulada por ABA, el cual fue monitoreado al ser aplicado de manera exógena. También proponen que esta proteína estaría controlando la homeostasis de ABA.

OTRAS FAMILIAS FTs.- Se encontraron varios TFs zinc-finger en las genotecas. Estos factores se subclasifican de acuerdo a la estructura del dominio conservado, entre los que se encuentra los TFs GATA (Gupta *et al.*, 2017) los cuales se caracterizan porque reconocen a los elementos de los promotores con una secuencia central G-A-T-A. (Behringer y Schwechheimer, 2015). De estos factores identificados en las genotecas no se hallaron reportes de sobreexpresión asociados a la tolerancia a algún estrés osmótico. Poco se sabe de su participación en la respuesta frente a estrés abiótico, pero por ejemplo, se ha reportado en *A. thaliana* sobre su participación en la prevención del daño fotooxidativo a través de la biosíntesis de tetrapirrol (TPB) (Kobayashi y Masuda, 2016). Adicionalmente, otro reporte demuestra que diferentes estrés abióticos tal como salinidad, sequía y ABA exógeno

causaron regulación diferencial de formas de splicing alternativo de FTs GATA en arroz. (Gupta *et al.*, 2017).

5.5 Búsqueda de genes adicionales interesantes

En referencia a la transducción de la señal del estrés, los cambios en la concentración de iones de calcio actúan como segundos mensajeros y estos cambios son decodificados por diferentes familias de proteínas, como por ejemplo: proteínas quinasas dependiente de Ca^{2+} (CDPKs), calcineurina, proteína de unión a calcineurina B (CBL), calmodulina (CAMs), entre las más conocidas que participan regulatoriamente en el proceso de transducción de señal (Wurzinger *et al.*, 2011). Con el objetivo de hallar genes de estas proteínas de transducción de la señal en las genotecas obtenidas, se realizó la búsqueda de acuerdo término de GO “actividad quinasa”, en donde se encontraron principalmente CDPKs y MAPKs (Tablas 10 y 11). Entre los genes hallados se tiene reporte de MAPK2 de *A. thaliana* que ha sido identificada como un regulador clave de la respuesta a estrés por frío y sal (Wurzinger *et al.*, 2011). Franz *et al.*, (2011), proporcionaron evidencia de que CPK21 contribuye a la señalización de estrés abiótico y que se activa bioquímicamente *in vivo* en respuesta al estrés hiperosmótico.

Los otros productos génicos encontrados luego de realizar la búsqueda por el término de GO “respuesta al estrés”, curiosamente incluyen proteínas que han sido previamente vinculadas en la adaptación al estrés abiótico, tales como chaperones moleculares, proteínas de choque térmico (HSP, heat shock protein), y productos

génicos con actividad enzimática. (Tabla 12). Frente al estrés por frío la superfamilia de HSP desempeña un papel crucial en la protección de la membrana plasmática manteniendo las conformaciones funcionales de las proteínas, facilitando el replegamiento de las proteínas desnaturalizadas, prohibiendo la agregación de proteínas y ayudando a la translocación de proteínas en los orgánulos. La regulación “hacia arriba” de los genes de la familia HSP70 se demostró en *A. thaliana* durante estrés por frío (Kazemi-Shahandashti y Maali-Amiri, 2018).

6 CONCLUSIONES

- Se logró obtener genotecas de ADNc por hibridación sustractiva por supresión (SSH) a partir de ARN de hojas de *Solanum juzepczukii* sometidos a tratamientos de helada y control.
- La tecnología de la hibridación sustractiva por supresión (SSH) combinada con el secuenciamento masivo fue una estrategia que permitió la búsqueda de genes expresados diferencialmente en hojas de *Solanum juzepczukii* bajo tratamientos de helada.
- Se encontraron los siguientes genes anotados como factores de transcripción: WRKY26, WRKY33, WRKY40, ERF7, CRF4, RAP2.12, BHLH122, BHLH68 y WRKY41, los cuales están involucrados en la tolerancia a algún estrés osmótico, de acuerdo a estudios de transgénesis.
- Mediante el análisis de enriquecimiento de GO se logró mostrar que el tratamiento de helada aplicado a hojas de *Solanum juzepczukii* afectaron a los procesos biológicos involucrados en la generación de energía (fotosíntesis y catabolismo) y en la regulación de la expresión genética (transcripción).

7 RECOMENDACIONES

- Validar las genotecas mediante técnicas de análisis de expresión diferencial, como *northern blot*, análisis de expresión de genes por qPCR, etc.
- Corroborar los dominios de los genes anotados como factores de transcripción, mediante análisis bioinformático.
- Analizar la expresión de los genes identificados en este trabajo, bajo diferentes condiciones como por ejemplo: tipos de tratamientos, tipos de tejidos, diferente estrés osmótico, tanto en especies resistentes como susceptibles de papa.
- Analizar la información nueva en las genotecas sustractivas, como los *contigs* que no dieron homología con la base de datos NR del NCBI

8 REFERENCIAS BIBLIOGRÁFICAS

- Al-Shahrour, F., Díaz-Uriarte, R., & Dopazo, J. (2004). FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics*, 20(4), 578-580.
- Arcos, J., Gastelo, M., & Holguín, V. (2016). INIA 317–Altiplano, variedad de papa con buena adaptación en la región altiplánica del Perú. *Revista Latinoamericana de la Papa*, 19(2), 68-75.
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... & Harris, M. A. (2000). Gene Ontology: tool for the unification of biology. *Nature genetics*, 25(1), 25-29.
- Badapanda, C. (2013). Suppression subtractive hybridization (SSH) combined with bioinformatics method: an integrated functional annotation approach for analysis of differentially expressed immune-genes in insects. *Bioinformation*, 9(4), 216.
- Banerjee, A., & Roychoudhury, A. (2015). WRKY proteins: signaling and regulation of expression during abiotic stress responses. *The Scientific World Journal*, 2015.
- Beerling, D. J., Terry, A. C., Hopwood, C., & Osborne, C. P. (2002). Feeling the cold: atmospheric CO₂ enrichment and the frost sensitivity of terrestrial plant foliage. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 182(1-2), 3-13.
- Behringer, C., & Schwechheimer, C. (2015). B-GATA transcription factors—insights into their structure, regulation, and role in plant development. *Frontiers in plant science*, 6.
- Berry, J. O., Yerramsetty, P., Zielinski, A. M., & Mure, C. M. (2013). Photosynthetic gene expression in higher plants. *Photosynthesis research*, 117(1-3), 91-120.
- Boyer, J. S. (1982). Plant productivity and environment. *Science*, 218(4571), 443-448.
- Carvalho, M. A., Pino, M. T., Jeknić, Z., Zou, C., Doherty, C. J., Shiu, S. H., ... & Thomashow, M. F. (2011). A comparison of the low temperature transcriptomes and CBF regulons of three plant species that differ in freezing tolerance: *Solanum commersonii*, *Solanum tuberosum*, and *Arabidopsis thaliana*. *Journal of experimental botany*, 62(11), 3807-3819.
- Centro de Estudios y Prevención de Desastres – PREDES. 2016. Mapa de heladas meteorológicas Julio 2000. Recuperado de: <http://www.predes.org.pe/mapas/heladas/>
- Centro Internacional de la Papa (CIP), y la Federación Departamental de Comunidades Campesinas (FEDECH), 2006. Catálogo de Variedades de papa nativa de Huancavelica – Perú. Recuperado de: <http://cipotato.org/wp-content/uploads/2014/08/003524.pdf>
- Chang, D. C., Sohn, H. B., Cho, J. H., Im, J. S., Jin, Y. I., Do, G. R., ... & Lee, Y. B. (2014). Freezing and frost damage of potato plants: a case study on growth recovery, yield response, and quality changes. *Potato research*, 57(2), 99.

- Chávez, R. (2004). Sobre el origen, evolución y diversidad genética de la papa cultivada y la silvestre. *Ciencia y desarrollo*.
- Chen, H. H., & Li, P. H. (1980). Characteristics of cold acclimation and deacclimation in tuber-bearing *Solanum* species. *Plant Physiology*, 65(6), 1146-1148.
- Chen, H., Lai, Z., Shi, J., Xiao, Y., Chen, Z., & Xu, X. (2010). Roles of *Arabidopsis* WRKY18, WRKY40 and WRKY60 transcription factors in plant responses to abscisic acid and abiotic stress. *BMC plant biology*, 10(1), 281.
- Chen, L., Song, Y., Li, S., Zhang, L., Zou, C., & Yu, D. (2012). The role of WRKY transcription factors in plant abiotic stresses. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*, 1819(2), 120-128.
- Cheng, M. C., Liao, P. M., Kuo, W. W., & Lin, T. P. (2013). The *A. thaliana* ETHYLENE RESPONSE FACTOR1 regulates abiotic stress-responsive gene expression by binding to different cis-acting elements in response to different stress signals. *Plant Physiology*, 162(3), 1566-1582.
- Chinnusamy, V., Ohta, M., Kanrar, S., Lee, B. H., Hong, X., Agarwal, M., & Zhu, J. K. (2003). ICE1: a regulator of cold-induced transcriptome and freezing tolerance in *A. thaliana*. *Genes & development*, 17(8), 1043-1054.
- Chinnusamy, V., Zhu, J. K., & Sunkar, R. (2010). Gene regulation during cold stress acclimation in plants. *Plant Stress Tolerance: Methods and Protocols*, 39-55.
- Chinnusamy, V., Zhu, J., & Zhu, J. K. (2007). Cold stress regulation of gene expression in plants. *Trends in plant science*, 12(10), 444-451.
- Condori, B., Hijmans, R. J., Ledent, J. F., & Quiroz, R. (2014). Managing potato biodiversity to cope with frost risk in the high andes: a modeling perspective. *PloS one*, 9(1), e81510.
- Condori, B., Hijmans, R. J., Ledent, J. F., & Quiroz, R. (2014). Managing potato biodiversity to cope with frost risk in the high Andes: a modeling perspective. *PloS one*, 9(1), e81510.
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21(18), 3674-3676.
- Czarnocka, W., & Karpiński, S. (2018). Friend or foe? Reactive oxygen species production, scavenging and signaling in plant response to environmental stresses. *Free Radical Biology and Medicine*.
- Di Toppi, L. S., & Pawlik-Skowrońska, B. (Eds.). (2003). *Abiotic stresses in plants*. Dordrecht: Kluwer Academic Publishers.
- Diatchenko, L., Lau, Y. F., Campbell, A. P., Chenchik, A., Moqadam, F., Huang, B., ... & Siebert, P. D. (1996). Suppression subtractive hybridization: a method for generating differentially regulated or tissue-specific cDNA probes and libraries. *Proceedings of the National Academy of Sciences*, 93(12), 6025-6030.
- Ding, Z. J., Yan, J. Y., Li, G. X., Wu, Z. C., Zhang, S. Q., & Zheng, S. J. (2014). WRKY41 controls *Arabidopsis* seed dormancy via direct regulation of ABI3 transcript levels not downstream of ABA. *The Plant Journal*, 79(5), 810-823.

- Distl, M., & Wink, M. (2009). Identification and quantification of steroidal alkaloids from wild tuber-bearing *Solanum* species by HPLC and LC-ESI-MS. *Potato Research*, 52(1), 79-104.
- Dubois, M., Skirycz, A., Claeys, H., Maleux, K., Dhondt, S., De Bodt, S., ... & Inzé, D. (2013). ETHYLENE RESPONSE FACTOR6 acts as a central regulator of leaf growth under water-limiting conditions in *Arabidopsis*. *Plant physiology*, 162(1), 319-332.
- Duque, A. S., de Almeida, A. M., da Silva, A. B., da Silva, J. M., Farinha, A. P., Santos, D., ... & de Sousa Araújo, S. (2013). Abiotic stress responses in plants: unraveling the complexity of genes and networks to survive. In *Abiotic Stress-Plant responses and applications in agriculture*. InTech.
- El-Metwally, S., Ouda, O. M., & Helmy, M. (2014). Next generation sequencing technologies and challenges in sequence assembly (Vol. 7). Springer Science & Business.
- Evrogen. Suppression Subtractive Hybridization (SSH). Recuperado de <http://evrogen.com/technologies/SSH.shtml>
- Fairlie, T. E., & Ortega, A. (2016). Efecto de la presencia de heladas simuladas en diferentes estados fenológicos y su impacto en el rendimiento de la papa cv Ccompis: estudio preliminar. *Revista Latinoamericana de la papa*, 7(1), 86-93.
- Food and Agriculture Organization - FAO. 2008a. La papa. Recuperado de: <http://www.fao.org/potato-2008/es/lapapa/index.html>
- Food and Agriculture Organization - FAO. 2008b. Cultivo Endémico: Papas Amargas (*Solanum x juzepczukii*). Tecnologías y prácticas para pequeños productores agrarios (TECA). Disponible en <http://teca.fao.org/es/read/4100>
- Food and Agriculture Organization (FAO). (2008): <http://www.fao.org/potato-2008/es/lapapa/index.html>
- Fowler, S., & Thomashow, M. F. (2002). *Arabidopsis* transcriptome profiling indicates that multiple regulatory pathways are activated during cold acclimation in addition to the CBF cold response pathway. *The Plant Cell*, 14(8), 1675-1690.
- Fraire-Velázquez, S., Rodríguez-Guerra, R., & Sánchez-Calderón, L. (2011). Abiotic and biotic stress response crosstalk in plants. In *Abiotic Stress Response in Plants-Physiological, Biochemical and Genetic Perspectives*. InTech.
- Franz, S., Ehlert, B., Liese, A., Kurth, J., Cazalé, A. C., & Romeis, T. (2011). Calcium-dependent protein kinase CPK21 functions in abiotic stress response in *Arabidopsis thaliana*. *Molecular Plant*, 4(1), 83-96.
- Gabriel, J., Magne, J., García, R., & Coca, J. 2009. Variedades nativas de papa: Tecnología tangible para afrontar el efecto causado por la helada y la sequía. *Revista Latinoamericana de la Papa*. 15(1):75-77
- Gasca-González, M., Rivera-Herrera, Y., Torres-Pacheco, I., González-Chavira, M. M., Guevara-Olvera, L., Muñoz-Sánchez, C. I., & Guevara-González, R. G. (2008). Estudio del transcriptoma en *Capsicum chinense* Jacq. resistente al virus huasteco vena amarilla del Chile. *Agrociencia*, 42(1), 107-117.
- Gaudet, P., & Dessimoz, C. (2017). Gene ontology: pitfalls, biases, and remedies. *The Gene Ontology Handbook*, 189-205.

- Gaudet, P., Škunca, N., Hu, J. C., & Dessimoz, C. (2017). Primer on the gene ontology. *The Gene Ontology Handbook*, 25-37.
- Gene Ontology Consortium. (2015). Gene ontology consortium: going forward. *Nucleic acids research*, 43(D1), D1049-D1056.
- Gilles, A., Megléc, E., Pech, N., Ferreira, S., Malausa, T., & Martin, J. F. (2011). Accuracy and quality assessment of 454 GS-FLX Titanium pyrosequencing. *BMC genomics*, 12(1), 245.
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., ... & Chen, Z. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, 29(7), 644.
- Guo, A. Y., Chen, X., Gao, G., Zhang, H., Zhu, Q. H., Liu, X. C., ... & Luo, J. (2007). PlantTFDB: a comprehensive plant transcription factor database. *Nucleic Acids Research*, 36(suppl_1), D966-D969.
- Gupta, P., Nutan, K. K., Singla-Pareek, S. L., & Pareek, A. P. (2017). Abiotic stresses cause differential regulation of alternative splice forms of GATA transcription factor in rice. *Frontiers in plant science*, 8, 1944.
- Gurskaya, N. G., Diatchenko, L., Chenchik, A., Siebert, P. D., Khaspekov, G. L., Lukyanov, K. A., ... & Sverdlov, E. D. (1996). Equalizing cDNA subtraction based on selective suppression of polymerase chain reaction: cloning of Jurkat cell transcripts induced by phytohemagglutinin and phorbol 12-myristate 13-acetate. *Analytical biochemistry*, 240(1), 90-97.
- Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., ... & MacManes, M. D. (2013). De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nature protocols*, 8(8), 1494.
- Hakeem, K. R., Tombuloğlu, H., & Tombuloğlu, G. (Eds.). (2016). *Plant Omics: Trends and Applications*. Springer.
- Hijmans, R. J., & Spooner, D. M. (2001). Geographic distribution of wild potato species. *American Journal of Botany*, 88(11), 2101-2112.
- Hijmans, R. J., Condori, B., Carrillo, R., & Kropff, M. J. (2003). A quantitative and constraint-specific method to assess the potential impact of new agricultural technology: the case of frost resistant potato for the Altiplano (Peru and Bolivia). *Agricultural Systems*, 76(3), 895-911.
- Hirt, H., & Shinozaki, K. (2003). *Plant responses to abiotic stress* (Vol. 4). Springer Science & Business Media.
- Huang, G. T., Ma, S. L., Bai, L. P., Zhang, L., Ma, H., Jia, P., ... & Guo, Z. F. (2012). Signal transduction during cold, salt, and drought stresses in plants. *Molecular biology reports*, 39(2), 969-987.
- Huntley, R. P., Sawford, T., Mutowo-Meullenet, P., Shypitsyna, A., Bonilla, C., Martin, M. J., & O'donovan, C. (2014). The GOA database: gene ontology annotation updates for 2015. *Nucleic acids research*, 43(D1), D1057-D1063
- Instituto Nacional de Estadística e Informática -INEI, (2016). Más del 30% de la población que vive en zonas afectadas por heladas está en edad vulnerable. Recuperado de: <https://www.inei.gob.pe/prensa/noticias/mas-del-30-de-la>

poblacion-que-vive-en-zonas-afectadas-por-heladas-esta-en-edad-vulnerable-9208/

- Jaglo, K. R., Kleff, S., Amundsen, K. L., Zhang, X., Haake, V., Zhang, J. Z., ... & Thomashow, M. F. (2001). Components of the Arabidopsis C-repeat/dehydration-responsive element binding factor cold-response pathway are conserved in *Brassica napus* and other plant species. *Plant physiology*, 127(3), 910-917
- Ji, W., Wright, M. B., Cai, L., Flament, A., & Lindpaintner, K. (2002). Efficacy of SSH PCR in isolating differentially expressed genes. *BMC genomics*, 3(1), 12.
- Jiang, X. C., Guo, X. H., Pan, X. L., & Song, S. Q. (2004). Construction and differential screening of a cDNA library specific to osmotic stress of *Haloxylon ammodendron* seedlings. *BMB Reports*, 37(5), 527-532.
- Jiang, Y., & Deyholos, M. K. (2009). Functional characterization of *A. thaliana* NaCl-inducible WRKY25 and WRKY33 transcription factors in abiotic stresses. *Plant molecular biology*, 69(1-2), 91-105.
- Kanehisa, M., & Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, 28(1), 27-30.
- Kazemi-Shahandashti, S. S., & Maali-Amiri, R. (2018). Global insights of protein responses to cold stress in plants: Signaling, defence, and degradation. *Journal of plant physiology*, 226, 123-135.
- Kikuchi, A., Huynh, H. D., Endo, T., & Watanabe, K. (2015). Review of recent transgenic studies on abiotic stress tolerance and future molecular breeding in potato. *Breeding science*, 65(1), 85-102.
- Knight, H., Zarka, D. G., Okamoto, H., Thomashow, M. F., & Knight, M. R. (2004). Abscisic acid induces CBF gene transcription and subsequent induction of cold-regulated genes via the CRT promoter element. *Plant physiology*, 135(3), 1710-1717.
- Kobayashi, K., & Masuda, T. (2016). Transcriptional regulation of tetrapyrrole biosynthesis in *A. thaliana thaliana*. *Frontiers in plant science*, 7.
- Kranner, I., Minibayeva, F. V., Beckett, R. P., & Seal, C. E. (2010). What is stress? Concepts, definitions and applications in seed science. *New Phytologist*, 188(3), 655-673.
- Le Hir, R., Castelain, M., Chakraborti, D., Moritz, T., Dinant, S., & Bellini, C. (2017). AtbHLH68 transcription factor contributes to the regulation of ABA homeostasis and drought stress tolerance in *A. thaliana thaliana*. *Physiologia Plantarum*.
- Li, S., Fu, Q., Chen, L., Huang, W., & Yu, D. (2011). *A. thaliana thaliana* WRKY25, WRKY26, and WRKY33 coordinate induction of plant thermotolerance. *Planta*, 233(6), 1237-1252.
- Licausi, F., Kosmacz, M., Weits, D. A., Giuntoli, B., Giorgi, F. M., Voesenek, L. A., ... & van Dongen, J. T. (2011). Oxygen sensing in plants is mediated by an N-end rule pathway for protein destabilization. *Nature*, 479(7373), 419.

- Licausi, F., Ohme-Takagi, M., & Perata, P. (2013). APETALA2/Ethylene Responsive Factor (AP2/ERF) transcription factors: mediators of stress responses and developmental programs. *New Phytologist*, 199(3), 639-649.
- Lindemose, S., O'Shea, C., Jensen, M. K., & Skriver, K. (2013). Structure, function and networks of transcription factors involved in abiotic stress responses. *International journal of molecular sciences*, 14(3), 5842-5878.
- Liu, J. H., Peng, T., & Dai, W. (2014). Critical cis-acting elements and interacting transcription factors: key players associated with abiotic stress responses in plants. *Plant molecular biology reporter*, 32(2), 303-317.
- Liu, W., Tai, H., Li, S., Gao, W., Zhao, M., Xie, C., & Li, W. X. (2014). bHLH122 is important for drought and osmotic stress resistance in *A. thaliana* and in the repression of ABA catabolism. *New Phytologist*, 201(4), 1192-1204.
- Llera-Herrera, R., García-Gasca, A., Abreu-Goodger, C., Huvet, A., & Ibarra, A. M. (2013). Identification of male gametogenesis expressed genes from the scallop *Nodipecten subnodosus* by suppressive subtraction hybridization and pyrosequencing. *PLoS One*, 8(9), e73176.
- Machida-Hirano, R. (2015). Diversity of potato genetic resources. *Breeding science*, 65(1), 26-40.
- Mahajan, S., & Tuteja, N. (2005). Cold, salinity and drought stresses: an overview. *Archives of biochemistry and biophysics*, 444(2), 139-158.
- Mao, G., Meng, X., Liu, Y., Zheng, Z., Chen, Z., & Zhang, S. (2011). Phosphorylation of a WRKY transcription factor by two pathogen-responsive MAPKs drives phytoalexin biosynthesis in *Arabidopsis*. *The Plant Cell*, tpc-111.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. journal*, 17(1), pp-10.
- Matías-Ramírez, L. G., Fuentes-Marilles, O. A., & García-Jiménez, F. (2012). Heladas. *Sistema Nacional de Protección Civil*.
- Mosa, K. A., Ismail, A., & Helmy, M. (2017). *Plant Stress Tolerance: An Integrated Omics Approach*. Springer.
- Müller, M., & Munné-Bosch, S. (2015). Ethylene response factors: a key regulatory hub in hormone and stress signaling. *Plant physiology*, 169(1), 32-41.
- Nouri, M. Z., Moumeni, A., & Komatsu, S. (2015). Abiotic stresses: insight into gene regulation and protein expression in photosynthetic pathways of plants. *International journal of molecular sciences*, 16(9), 20392-20416.
- Olsvik, P. A., Vikeså, V., Lie, K. K., & Hevrøy, E. M. (2013). Transcriptional responses to temperature and low oxygen stress in Atlantic salmon studied with next-generation sequencing technology. *BMC genomics*, 14(1), 817.
- Owen-Hughes, T., & Engholm, M. (2007). Pyrosequencing positions nucleosomes precisely. *Genome biology*, 8(6), 217
- Papdi, C., Pérez-Salamó, I., Joseph, M. P., Giuntoli, B., Bögre, L., Koncz, C., & Szabados, L. (2015). The low oxygen, oxidative and osmotic stress responses synergistically act through the ethylene response factor VII genes RAP2. 12, RAP2. 2 and RAP2. 3. *The Plant Journal*, 82(5), 772-784.

- Park, S., Lee, C. M., Doherty, C. J., Gilmour, S. J., Kim, Y., & Thomashow, M. F. (2015). Regulation of the Arabidopsis CBF regulon by a complex low-temperature regulatory network. *The Plant Journal*, 82(2), 193-207.
- Pino, M. T., Skinner, J. S., Park, E. J., Jeknić, Z., Hayes, P. M., Thomashow, M. F., & Chen, T. H. (2007). Use of a stress inducible promoter to drive ectopic AtCBF expression improves potato freezing tolerance while minimizing negative effects on tuber yield. *Plant Biotechnology Journal*, 5(5), 591-604.
- Presidencia del Consejo de Ministros - PMC. (2012). Plan nacional de intervención para enfrentar los efectos de la temporada de heladas y friaje 2012. . Recuperado de <http://minagri.gob.pe/portal/download/pdf/marcolegal/normaslegales/resolucionessupremas/2012/plan-rs174-2012-pcm.pdf>
- Pucciariello, C., & Perata, P. (2012). How plants sense low oxygen. *Plant signaling & behavior*, 7(7), 813-816.
- Raymundo, R., Asseng, S., Robertson, R., Petsakos, A., Hoogenboom, G., Quiroz, R., ... & Wolf, J. (2017). Climate change impact on global potato production. *European Journal of Agronomy*.
- Ren, X., Liu, T., Dong, J., Sun, L., Yang, J., Zhu, Y., & Jin, Q. (2012). Evaluating de Bruijn graph assemblers on 454 transcriptomic data. *PLoS One*, 7(12), e51188.
- Ritter, E., Barandalla, L., López, R., & de Galarreta, J. I. R. (2008). Exploitation of exotic, cultivated *Solanum* germplasm for breeding and commercial purposes. *Potato Research*, 51(3-4), 301-311.
- Rodrigues, F. A., Marcolino-Gomes, J., Carvalho, J. D. F. C., Nascimento, L. C. D., Neumaier, N., Farias, J. R. B., ... & Nepomuceno, A. L. (2012). Subtractive libraries for prospecting differentially expressed genes in the soybean under water deficit. *Genetics and molecular biology*, 35(1), 304-314.
- Rodziewicz, P., Swarczewicz, B., Chmielewska, K., Wojakowska, A., & Stobiecki, M. (2014). Influence of abiotic stresses on plant proteome and metabolome changes. *Acta Physiologiae Plantarum*, 36(1), 1-19.
- Rowe, G., Sweet, M., & Beebe, T. (2017). *An introduction to molecular ecology*. Oxford University Press.
- Rushton, D. L., Tripathi, P., Rabara, R. C., Lin, J., Ringler, P., Boken, A. K., ... & Chen, X. (2012). WRKY transcription factors: key components in abscisic acid signalling. *Plant biotechnology journal*, 10(1), 2-11.
- Sahebi, M., Hanafi, M. M., Azizi, P., Hakim, A., Ashkani, S., & Abiri, R. (2015). Suppression subtractive hybridization versus next-generation sequencing in plant genetic engineering: challenges and perspectives. *Molecular biotechnology*, 57(10), 880-903.
- Servicio Nacional de Meteorología e Hidrología- SENAMHI (2010). Atlas de heladas del Perú. Recuperado de www.senamhi.gob.pe/main_down.php?ub=est&id=atlasHeladas
- Shao, H. B., Chu, L. Y., Jaleel, C. A., Manivannan, P., Panneerselvam, R., & Shao, M. A. (2009). Understanding water deficit stress-induced changes in the basic metabolism of higher plants—biotechnologically and sustainably improving

- agriculture and the environment in arid regions of the globe. *Critical Reviews in Biotechnology*, 29(2), 131-151.
- Sharma, P., Sharma, N., & Deswal, R. (2005). The molecular biology of the low-temperature response in plants. *Bioessays*, 27(10), 1048-1059
- Skubacz, A., Daszkowska-Golec, A., & Szarejko, I. (2016). The role and regulation of ABI5 (ABA-Insensitive 5) in plant development, abiotic stress responses and phytohormone crosstalk. *Frontiers in plant science*, 7, 1884
- Snyder, R. L., Melo-Abreu, J. P., Matulich, S., Villar-Mir, J. M., & Millán, J. (2010). *Protección contra las heladas: fundamentos, práctica y economía*. Volume 2. *Protección contra las heladas: fundamentos, práctica y economía*, 2, 1-80.
- Song, C. P., Agarwal, M., Ohta, M., Guo, Y., Halfter, U., Wang, P., & Zhu, J. K. (2005). Role of an *A. thaliana* AP2/EREBP-type transcriptional repressor in abscisic acid and drought stress responses. *The Plant Cell*, 17(8), 2384-2396.
- Spooner, D. M., Núñez, J., Trujillo, G., del Rosario Herrera, M., Guzmán, F., & Ghislain, M. (2007). Extensive simple sequence repeat genotyping of potato landraces supports a major reevaluation of their gene pool structure and classification. *Proceedings of the National Academy of Sciences*, 104(49), 19398-19403.
- Su, M., Huang, G., Zhang, Q., Wang, X., Li, C., Tao, Y., ... & Wang, Y. (2016). The LEA protein, ABR, is regulated by ABI5 and involved in dark-induced leaf senescence in *Arabidopsis thaliana*. *Plant Science*, 247, 93-103.
- Thomashow, M. F. (1999). Plant cold acclimation: freezing tolerance genes and regulatory mechanisms. *Annual review of plant biology*, 50(1), 571-599.
- Tipney, H., & Hunter, L. (2010). An introduction to effective use of enrichment analysis software. *Human genomics*, 4(3), 202.
- Wang, X. C., Zhao, Q. Y., Ma, C. L., Zhang, Z. H., Cao, H. L., Kong, Y. M., ... & Jin, J. Q. (2013). Global transcriptome profiles of *Camellia sinensis* during cold acclimation. *BMC genomics*, 14(1), 415.
- Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews genetics*, 10(1), 57.
- Wasilewska, A., Vlad, F., Sirichandra, C., Redko, Y., Jammes, F., Valon, C., ... & Leung, J. (2008). An update on abscisic acid signaling in plants and more.... *Molecular plant*, 1(2), 198-217.
- Welling, A., & Palva, E. T. (2008). Involvement of CBF transcription factors in winter hardiness in birch. *Plant physiology*, 147(3), 1199-1211.
- Wessler, S. R. (2005). Homing into the origin of the AP2 DNA binding domain. *Trends in plant science*, 10(2), 54-56.
- Wurzinger, B., Mair, A., Pfister, B., & Teige, M. (2011). Cross-talk of calcium-dependent protein kinase and MAP kinase signaling. *Plant signaling & behavior*, 6(1), 8-12.
- Xiong, L., & Zhu, J. K. (2001). Abiotic stress signal transduction in plants: molecular and genetic perspectives. *Physiologia plantarum*, 112(2), 152-166.
- Xu, X., Chen, C., Fan, B., & Chen, Z. (2006). Physical and functional interactions between pathogen-induced *Arabidopsis* WRKY18, WRKY40, and WRKY60 transcription factors. *The Plant Cell*, 18(5), 1310-1326.

- Yoshida, T., Mogami, J., & Yamaguchi-Shinozaki, K. (2014). ABA-dependent and ABA-independent signaling in response to osmotic stress in plants. *Current opinion in plant biology*, 21, 133-139.
- Zhu, Y. Y., Machleder, E. M., Chenchik, A., Li, R., & Siebert, P. D. (2001). Reverse transcriptase template switching: A SMART™ approach for full-length cDNA library construction. *Biotechniques*, 30(4), 892-897.
- Zwack, P. J., Compton, M. A., Adams, C. I., & Rashotte, A. M. (2016). Cytokinin response factor 4 (CRF4) is induced by cold and involved in freezing tolerance. *Plant cell reports*, 35(3), 573-584.

9 ANEXOS

Tabla A1. *Genes en la genoteca **sentido** anotados con el término GO: actividad quinasa*

Sequence name	Sequence desc.	Hit ACC	E-Value
F5320_c0_g1_i1	calcineurin B 1	NP_001239047	2.55E-43
F6526_c0_g1_i1	calcium-dependent kinase 21-like	XP_0097770081	1.12E-82
F6964_c0_g1_i1	calcium-dependent kinase 3	XP_016570448	2.12E-06
F9226_c0_g1_i1	calcium-dependent kinase 32-like	XP_006351224	6.73E-72
F1078_c0_g1_i1	calcium-dependent kinase 5	NP_001274790	7.05E-91
F1383_c0_g1_i1	calmodulin-7-like isoform X1	XP_018478487	1.08E-56
F3567_c0_g1_i1	mitogen-activated kinase 19-like	XP_016481981	1.54E-38
F3999_c0_g1_i1	mitogen-activated kinase 20 isoform X1	XP_015082331	4.78E-58
F3191_c0_g1_i1	mitogen-activated kinase homolog MMK2	XP_015167490	2.47E-43
F6958_c0_g1_i1	mitogen-activated kinase partial	ABO65104	4.01E-26

Tabla A2. *Genes en la genoteca **antisentido** anotados con el término GO: actividad quinasa*

Sequence name	Sequence desc.	Hit ACC	E-Value
R942_c0_g1_i1	calcium-dependent kinase 8-like	XP_006366539	0
R2825_c0_g1_i1	calcium-dependent kinase 4	AAW31900	2.56E-56
R3044_c0_g1_i1	calcium-dependent kinase 13	XP_006367515	9.78E-12
R4313_c0_g1_i1	calcineurin B 7	XP_006350549	5.59E-137
R4316_c0_g1_i1	calcineurin B 10	XP_006360986	2.45E-71
R7847_c0_g1_i1	calmodulin 1	XP_015059721	5.80E-26
R3766_c0_g1_i1	mitogen-activated kinase homolog NTF4	ABM55743	1.19E-102

Tabla A3. *Genes en la genoteca **sentido** anotados con el término GO: respuesta al estrés abiótico*

Sequence name	Sequence desc.	Hit ACC	E-Value
F1279_c0_g1_i1	vicianin hydrolase-like	XP_006341337	0
F193_c0_g1_i1	two-component response regulator-like APRR7 isoform X1	XP_015056055	2.42E-17
F1540_c0_g3_i1	transketolase- chloroplastic	XP_018492328	7.95E-120
F79_c0_g1_i1	transcriptional corepressor LEUNIG	EEF47044	0
F2632_c0_g1_i1	transcription factor VOZ1-like	XP_016432550	2.88E-44
F4158_c0_g1_i1	sorting nexin 1	XP_006341082	1.34E-168
F1839_c0_g1_i1	scarecrow-like transcription factor PAT1	XP_004243547	1.63E-123

Sequence name	Sequence desc.	Hit ACC	E-Value
F1279_c0_g1_i1	vicianin hydrolase-like	XP_006341337	0
F410_c0_g2_i1	hsp70-Hsp90 organizing 2	XP_006355497	1.42E-147
F836_c0_g1_i1	eukaryotic translation initiation factor 3 subunit E	XP_006351291	0
F4317_c0_g1_i1	ethylene-insensitive 2	AAZ95507	3.86E-133
F8508_c0_g1_i1	dnaJ homolog mitochondrial-like	XP_004230549	1.42E-150
F8042_c0_g2_i1	cryptochrome-1 isoform X1	XP_006360269	0
F1333_c0_g1_i1	chlorophyll a-b binding chloroplastic-like	XP_004246325	2.97E-168
F3705_c0_g1_i1	autophagy-related 18a-like	XP_015875013	5.30E-104
F829_c0_g1_i1	ankyrin repeat-containing At3g12360	XP_006345920	0
F3293_c0_g1_i1	Tetratricopeptide repeat 30A	XP_006354117	1.84E-89
F8286_c0_g2_i1	TPR repeat-containing thioredoxin TTL1-like	XP_006348295	1.65E-125
F561_c0_g1_i1	SENSITIVE TO PROTON RHIZOTOXICITY 1-like	XP_006353630	1.39E-66
F5592_c0_g1_i1	EPS15 homology domain 1	KVI02296	4.56E-105
F6795_c0_g1_i1	BAG family molecular chaperone regulator 4-like	XP_006338974	7.58E-42
F1724_c0_g2_i1	B-box zinc finger 21-like	XP_006342575	2.56E-97
F9001_c0_g1_i1	ACT domain-containing ACR11-like	XP_006341438	4.89E-37
F2949_c0_g1_i1	HSP20-like chaperones superfamily	XP_015055443	5.24E-79
F2419_c0_g1_i1	peroxisome biogenesis 2	XP_015077119	1.38E-45
F2473_c0_g1_i1	J domain-containing required for chloroplast accumulation response 1	XP_006367308	4.54E-90
F2452_c0_g1_i1	phytochrome C isoform X1	XP_006357639	2.72E-38
F4878_c0_g1_i1	asparagine synthetase [glutamine-hydrolyzing]-like	XP_006344055	1.90E-19
F113_c0_g1_i1	tubulin alpha chain	NP_001275472	1.48E-79
F113_c0_g3_i1	Tubulin alpha partial	JAT42170	8.02E-106
F187_c0_g1_i1	heavy metal-associated isoprenylated plant 26-like	XP_016539965	5.06E-94
F1188_c0_g2_i1	major intrinsic 1	ABB16979	2.84E-60
F4377_c0_g1_i1	acyl- -binding domain-containing 3 isoform X1	XP_006362065	5.19E-44
F5384_c0_g1_i1	MID1-COMPLEMENTING ACTIVITY 1	XP_006343349	4.49E-48
F1246_c1_g1_i1	root phototropism 2	XP_006353728	8.04E-17
F1949_c0_g1_i1	Reticulata-related 1 isoform 3	XP_006340988	7.10E-91
F1927_c0_g1_i1	FAR-RED ELONGATED HYPOCOTYL 3 isoform X1	XP_006357871	9.05E-43
F1951_c0_g1_i1	pre-mRNA-splicing factor 38B-like	XP_006346763	1.61E-40
F7959_c0_g1_i1	TIME FOR COFFEE isoform X2	XP_015087522	1.48E-79
F7942_c0_g1_i1	ultraviolet-B receptor UVR8 isoform X2	XP_009769437	1.17E-70

Sequence name	Sequence desc.	Hit ACC	E-Value
F1279_c0_g1_i1	vicianin hydrolase-like	XP_006341337	0
F8045_c0_g1_i1	histone deacetylase complex subunit SAP18	XP_006360642	5.19E-88
F9068_c0_g1_i1	ycf3-interacting chloroplastic isoform X2	XP_015166482	1.36E-52
F9100_c0_g1_i1	auxin transport BIG	XP_004233657	7.37E-72
F2879_c0_g1_i1	mediator of RNA polymerase II transcription subunit 16 isoform X2	XP_006347066	3.45E-88
F5170_c0_g1_i1	cytoplasmic	XP_015169000	6.07E-53
F7498_c0_g1_i1	PALE chloroplastic isoform X1	XP_015087880	5.04E-53
F4124_c0_g1_i1	26S proteasome non-ATPase regulatory subunit 8 homolog A	XP_006343823	7.03E-85
F4679_c0_g1_i1	transcription termination factor chloroplastic	XP_004251918	4.85E-41
F7399_c0_g1_i1	enolase	BAD94751	6.97E-62
F7026_c0_g1_i1	TIME FOR COFFEE isoform X1	XP_006359914	1.85E-45
F4554_c0_g1_i1	peptide methionine sulfoxide reductase B5-like	XP_015055943	2.50E-32
F9125_c0_g1_i1	DNA excision repair ERCC-1	XP_006345166	1.02E-45
F9118_c0_g1_i1	CBL-interacting serine threonine- kinase 23 isoform X1	XP_008444632	2.33E-43
F4456_c0_g1_i1	beta-glucosidase 12-like	XP_016555685	3.95E-54
F6419_c0_g1_i1	respiratory burst oxidase homolog C	NP_001234271	8.36E-60
F2104_c0_g1_i1	protease Do-like 7	XP_006364016	4.22E-29
F7638_c0_g1_i1	cyclin-L1-1 isoform X1	XP_006347847	1.53E-11
F4711_c0_g1_i1	BTB POZ domain-containing NPY2	XP_006346315	6.66E-83
F3152_c0_g1_i1	cytochrome P450 724B1-like	XP_016579508	3.59E-26
F635_c0_g1_i1	zinc finger CONSTANS-LIKE 4-like	XP_015061111	1.00E-30
F8730_c0_g1_i1	probable histone H2A variant 1	CDP03683	8.60E-17
F8424_c0_g1_i1	aconitase 1	AAP30039	6.48E-38
F8478_c0_g1_i1	EARLY FLOWERING 4-like	XP_006338849	1.19E-67
F8528_c0_g1_i1	chaperone	XP_015070523	6.57E-59
F6241_c0_g1_i1	probable cadmium zinc-transporting ATPase chloroplastic	XP_006339845	4.89E-45
F8636_c0_g1_i1	glycosyltransferase-like KOBITO 1	XP_015062930	2.57E-15
F8681_c0_g1_i1	DNA repair recA homolog chloroplastic	XP_006346054	5.11E-74
F1383_c0_g2_i1	calmodulin	XP_013631799	5.37E-104
F7701_c0_g1_i1	phytochrome B-like	XP_006355734	1.48E-46
F7796_c0_g1_i1	omega-6 fatty acid chloroplastic	AGW21689	5.40E-40
F2575_c0_g1_i1	root UVB sensitive chloroplastic	XP_006361229	1.57E-119
F2503_c0_g1_i1	histone H4	XP_009126676	2.35E-48
F6168_c0_g1_i1	E3 ubiquitin- ligase HOS1-like	AEA11207	3.15E-81
F1581_c0_g1_i1	5 -3 exoribonuclease 4	XP_006340038	3.79E-55

Sequence name	Sequence desc.	Hit ACC	E-Value
F1279_c0_g1_i1	vicianin hydrolase-like	XP_006341337	0
F1750_c0_g1_i1	RNA polymerase sigma factor sigB-like	XP_006359117	1.75E-67
F1709_c0_g1_i1	Cullin-4B isoform 1	DAA51792	5.13E-48
F5276_c0_g1_i1	nuclear factor 1 A-type (DUF1005)	XP_016543514	1.42E-24

Tabla A4. Genes en la genoteca *antisentido* anotados con el término GO: respuesta al estrés abiótico.

Sequence name	Sequence desc.	Hit ACC	E-Value
R421_c0_g1_i1	zeaxanthin epoxidase	ADF28629	3.14E-105
R1086_c0_g1_i1	ycf20	XP_006352147	1.24E-51
R1260_c0_g2_i1	violaxanthin de-epoxidase	XP_006340494	3.38E-82
R1260_c0_g1_i1	violaxanthin de- chloroplastic	XP_006340494	1.26E-84
R144_c0_g1_i1	vesicle-associated membrane 714	XP_006361082	3.88E-144
R506_c0_g1_i1	tocopherol cyclase	NP_001274927	3.04E-17
R7817_c0_g1_i1	temperature-induced lipocalin	XP_015081947	2.28E-74
R4109_c0_g1_i1	squalene epoxidase 3-like	XP_015067151	9.93E-105
R7453_c0_g1_i1	sodium hydrogen exchanger 2-like	XP_006342736	8.42E-19
R121_c0_g1_i1	serine threonine- kinase SRK2E	XP_009138251	1.07E-34
R474_c0_g1_i1	probable zinc metalloprotease chloroplastic	XP_006339202	7.86E-48
R975_c0_g1_i1	probable pyridoxal 5 -phosphate synthase subunit PDX1	XP_006350840	4.31E-119
R4981_c0_g1_i1	phytochrome partial	NP_001274786	1.10E-69
R4206_c0_g1_i1	phytochrome C	XP_006357640	5.75E-85
R7943_c0_g1_i1	phytochrome B	XP_006358209	6.04E-47
R1170_c0_g1_i1	phytochrome A-associated F-box	XP_006359377	6.76E-132
R4814_c0_g1_i1	phytochrome A	XP_006339917	3.79E-59
R8025_c0_g1_i1	phosphatidylinositol 4-kinase gamma 4-like	XP_015060647	2.80E-72
R526_c0_g1_i1	nucleoside diphosphate kinase chloroplastic	XP_006351499	2.23E-86
R4391_c0_g1_i1	monothiol glutaredoxin-S17	XP_006347053	1.19E-114
R4363_c0_g1_i1	mediator of RNA polymerase II transcription subunit 32-like	XP_006359248	5.92E-73

Sequence name	Sequence desc.	Hit ACC	E-Value
R421_c0_g1_i1	zeaxanthin epoxidase	ADF28629	3.14E-105
R1048_c0_g1_i1	homogentisate phytyltransferase chloroplastic isoform X1	XP_004242907	7.57E-83
R5348_c0_g1_i1	histidine kinase 3	XP_006352176	0
R124_c0_g1_i1	heavy metal-associated isoprenylated plant 26	XP_009376948	1.89E-46
R3223_c0_g1_i1	glutamate--cysteine chloroplastic	XP_006343987	3.29E-142
R7473_c0_g1_i1	gibberellin 20 oxidase 2-like	XP_006366931	1.14E-100
R1272_c0_g1_i1	gibberellin 20 oxidase 1-like	XP_006362690	7.46E-173
R7839_c0_g1_i1	fasciclin-like arabinogalactan 4	XP_004241229	5.40E-39
R5371_c0_g1_i1	endoplasmic homolog	XP_015073391	1.07E-98
R4986_c0_g1_i1	dnaj homolog mitochondrial	XP_006364226	3.04E-28
R7912_c0_g1_i1	dnaj homolog 2-like	XP_015063195	1.02E-23
R945_c0_g1_i1	dnaj homolog	XP_006347935	3.43E-69
R1134_c0_g1_i1	cytochrome P450 90A1	XP_006338854	0
R2421_c0_g1_i1	cryptochrome-1-like isoform X1	XP_015073343	3.76E-104
R4973_c0_g1_i1	copper transport CCH-like	XP_009628580	2.34E-41
R1153_c0_g1_i1	containing PDZ a K-box and a TPR region	XP_015062563	2.05E-63
R997_c0_g1_i1	chloroplast light-harvesting complex II partial	ACF05893	4.05E-49
R1943_c0_g7_i1	chlorophyll binding partial	ACJ13833	8.48E-43
R2807_c0_g1_i1	chlorophyll a-b binding of LHCII type 1	XP_016734829	2.48E-10
R997_c0_g2_i1	chlorophyll a-b binding chloroplastic	XP_004244187	2.28E-75
R1983_c0_g1_i1	chlorophyll a-b binding CP24 chloroplastic	XP_006351799	5.36E-58
R5112_c0_g1_i1	chlorophyll a-b binding 1D	KQK04837	2.46E-40
R1615_c0_g1_i1	chlorophyll a b-binding Cab- partial	EMT23001	1.21E-35
R3608_c0_g1_i1	chaperone chloroplastic	BAE06227	1.13E-51
R4827_c0_g1_i1	calmodulin-binding 60 C	XP_006344728	4.88E-29
R1978_c0_g1_i1	calmodulin-7	KZV46024	3.18E-104
R4316_c0_g1_i1	calcineurin B 10	XP_006360986	2.45E-71
R920_c0_g1_i1	annexin D1-like	NP_001234104	2.24E-88
R145_c0_g1_i1	annexin 2	CAB92956	7.83E-10
R25_c0_g1_i1	alpha-glucan water chloroplastic isoform X1	Q9AWA5, CAA70725	2.63E-138
R2955_c0_g1_i1	alpha-glucan H isozyme	XP_006366595	6.87E-42

Sequence name	Sequence desc.	Hit ACC	E-Value
R421_c0_g1_i1	zeaxanthin epoxidase	ADF28629	3.14E-105
R498_c0_g1_i1	alpha-1,4 glucan phosphorylase L-1 chloroplastic amyloplastic	P04045, BAA00407	2.83E-33
R7986_c0_g1_i1	STABILIZED1	XP_016436523	1.39E-39
R4084_c0_g1_i1	RNA-binding KH domain-containing PEPPER-like	XP_006367746	1.96E-111
R7882_c0_g1_i1	RNA polymerase sigma factor chloroplastic isoform X2	XP_010325979	4.47E-49
R1153_c0_g2_i1	PREDICTED: uncharacterized protein LOC102590017	XP_006354117	1.04E-21
R7929_c0_g1_i1	NF-X1-type zinc finger NFXL1	XP_006341494	3.82E-56
R1145_c0_g1_i1	LSD1	XP_010316974	1.51E-66
R4867_c0_g1_i1	GDP-Man:Man(3) c(2)-PP-Dol alpha-1,2-mannosyltransferase	XP_015089286	3.97E-56
R4369_c0_g1_i1	F-box 7	XP_006351907	1.74E-80
R3264_c0_g1_i1	DNA repair RAD51 homolog 2	XP_015057217	2.48E-87
R2881_c0_g1_i1	DNA damage-binding 1 isoform X1	XP_016557038	2.90E-169
R8051_c0_g1_i1	DNA (cytosine-5)-methyltransferase CMT2 isoform X1	XP_010324482	2.70E-77
R515_c0_g1_i1	DELLA GAI	NP_001305514	3.13E-64
R963_c0_g1_i1	DEAD-box ATP-dependent RNA helicase 38-like	AIU49061	5.41E-35
R963_c1_g1_i1	DEAD-box ATP-dependent RNA helicase 38	XP_004249589	2.45E-62
R2470_c0_g1_i1	CHROMATIN REMODELING 8	XP_006351663	3.01E-39
R544_c0_g1_i1	CHAPERONE-LIKE PROTEIN OF chloroplastic	XP_015070515	1.46E-45
R3841_c0_g1_i1	CBL-interacting serine threonine- kinase 24	XP_006345226	6.09E-155
R4997_c0_g1_i1	B-box zinc finger 22 isoform X1	XP_006348330	3.12E-56
R4805_c0_g1_i1	ATP-dependent zinc metalloprotease chloroplastic	XP_006342112	8.76E-44
R148_c0_g1_i1	ATP-dependent zinc metalloprotease FTSH chloroplastic	XP_004243266	4.15E-115
R3293_c0_g1_i1	ABSCISIC ACID-INSENSITIVE 5 5	NP_001274925	1.20E-119
R3641_c0_g1_i1	ABC transporter G family member 15-like	XP_006365197	5.14E-51
R2442_c0_g1_i1	ABC transporter G family member 11-like	XP_006351519	1.81E-23
R4240_c0_g1_i1	ABC transporter G family member 11	XP_016563904	4.21E-126
R3309_c0_g1_i1	ABC transporter C family member 5	XP_006355894	3.40E-87

Sequence name	Sequence desc.	Hit ACC	E-Value
R421_c0_g1_i1	zeaxanthin epoxidase	ADF28629	3.14E-105
R4667_c0_g1_i1	B-box zinc finger 22-like	XP_006348330	2.34E-85
R3973_c0_g1_i1	PREDICTED: uncharacterized protein LOC102604646	XP_006355940	5.22E-41
R3972_c0_g1_i1	RING-H2 finger ATL38-like	XP_006360481	2.51E-64
R7345_c0_g1_i1	systemin receptor SR160	XP_006357355	8.15E-58
R7389_c0_g1_i1	histidine kinase 3 isoform X1	XP_006352176	1.05E-29
R7357_c0_g1_i1	alpha-glucan phosphorylase 1	XP_013706200	3.59E-85
R7044_c0_g1_i1	disulfide isomerase-like 1-2	NP_177875	3.37E-51
R7096_c0_g1_i1	alpha-1,4 glucan phosphorylase L chloroplastic amyloplastic	XP_004234848	3.77E-08
R1422_c0_g1_i1	amino acid selective channel	XP_004960518	1.72E-43
R1476_c0_g1_i1	cryptochrome 1 partial	XP_006360269	4.46E-77
R3732_c0_g1_i1	sodium hydrogen exchanger 3-like isoform X3	XP_010321696	4.11E-66
R6429_c0_g1_i1	probable histone H2A variant 3	XP_004251586	4.32E-22
R2115_c0_g1_i1	zinc finger CONSTANS-LIKE 16-like	XP_004235805	1.35E-25
R2181_c1_g7_i1	chloroplast light-harvesting chlorophyll a b-binding partial	AAA33700	6.76E-34
R2155_c0_g1_i1	heat shock 81-2	KVI10442	8.31E-19
R7633_c0_g1_i1	aspartate aminotransferase	XP_006343100	7.54E-132
R7644_c0_g1_i1	ABC transporter B family member 1	XP_016542793	4.81E-51
R4767_c0_g1_i1	probable serine threonine- kinase WNK5	XP_006346683	2.42E-26
R3109_c0_g1_i1	PROTON GRADIENT REGULATION chloroplastic	XP_006354442	5.40E-08
R6366_c0_g1_i1	probable tocopherol chloroplastic	NP_001274927	6.79E-36
R6331_c0_g1_i1	SAL1 phosphatase-like	XP_006345757	2.85E-70
R6525_c0_g1_i1	mediator of RNA polymerase II transcription subunit 14	XP_006347747	5.42E-43
R8566_c0_g1_i1	starch partial	NP_001275128	2.91E-55
R6221_c0_g1_i1	probable WRKY transcription factor 2 isoform X1	NP_001305535	1.19E-38
R1320_c0_g1_i1	2-aminoethanethiol dioxygenase	XP_006345075	2.20E-144
R8297_c0_g1_i1	constitutive photomorphogenesis 10	XP_006354144	1.40E-81
R8243_c0_g1_i1	mediator of RNA polymerase II transcription subunit 18	XP_015079222	4.63E-62
R896_c0_g1_i1	succinate-semialdehyde mitochondrial isoform X1	NP_001293104	2.15E-94

Sequence name	Sequence desc.	Hit ACC	E-Value
R421_c0_g1_i1	zeaxanthin epoxidase	ADF28629	3.14E-105
R804_c0_g2_i1	5-methyltetrahydropteroyltriglutamate-- homocysteine methyltransferase	XP_015055513	2.69E-91
R833_c0_g1_i1	cytoplasmic-like isoform X1	NP_001275333	1.32E-42
R2563_c0_g1_i1	nucleoside diphosphate kinase chloroplastic-like	XP_006347398	7.03E-18
R2374_c0_g1_i1	BTB POZ domain-containing POB1-like	XP_006339663	3.09E-18
R2304_c0_g1_i1	phosphatidylinositol 4-phosphate 5- kinase MSS4	XP_004248788	5.79E-16
R2037_c0_g3_i1	sodium calcium exchanger family calcium-binding EF hand family	XP_006348283	5.96E-29
R2098_c0_g1_i1	superoxide dismutase [Cu-Zn] chloroplastic	NP_001234769	5.22E-117
R2009_c0_g1_i1	late embryogenesis abundant Lea14- partial	XP_006351075	3.04E-166
R2072_c0_g1_i1	B-box zinc finger 24-like	ABC25454	6.84E-117
R2061_c1_g1_i1	leaf L-A	2210387B	2.46E-25
R6041_c0_g1_i1	phytochrome E isoform X2	XP_006355590	2.46E-97
R3422_c0_g1_i1	probable WRKY transcription factor 65 isoform X1	XP_015055738	2.39E-24
R3459_c0_g1_i1	chromo domain LHP1-like	XP_006353982	5.57E-36
R3479_c0_g1_i1	AP-3 complex subunit mu	XP_015072013	5.90E-91
R6131_c0_g1_i1	BTB POZ domain-containing POB1-like isoform X2	XP_006339663	2.92E-108
R5993_c0_g1_i1	yfc3-interacting chloroplastic isoform X1	XP_006352612	3.45E-52
R7206_c0_g1_i1	probable galactinol--sucrose galactosyltransferase 2	XP_006361829	1.29E-141
R1566_c0_g1_i1	signal recognition particle 43 kDa chloroplastic	XP_010316268	2.98E-171
R5541_c0_g1_i1	E3 ubiquitin- ligase RMA1H1-like	XP_006357485	2.38E-148
R5545_c0_g1_i1	OS-9 homolog	XP_006359509	1.27E-26
R5521_c0_g1_i1	polyadenylate-binding RBP47	XP_006346456	0
R207_c0_g1_i1	cryptochrome partial	NP_001234577	3.43E-55
R2617_c0_g1_i1	SENSITIVE TO PROTON RHIZOTOXICITY 1	XP_006351163	1.87E-46
R5868_c0_g1_i1	DELLA GAI-like	XP_015058359	2.53E-128
R5876_c0_g1_i1	XAP5 CIRCADIAN TIMEKEEPER	XP_016553369	4.13E-29
R5890_c0_g1_i1	TIME FOR COFFEE isoform X3	XP_006359914	1.21E-85
R2209_c1_g1_i1	calcium sensing chloroplastic	XP_006345067	0
R2284_c2_g3_i1	Fe superoxide dismutase 1	XP_015168847	2.87E-77

Sequence name	Sequence desc.	Hit ACC	E-Value
R421_c0_g1_i1	zeaxanthin epoxidase	ADF28629	3.14E-105
R2718_c0_g1_i1	rac-like GTP-binding 5	ABE11611	1.10E-114
R2702_c0_g1_i1	asparagine synthetase [glutamine-hydrolyzing]	XP_006344055	1.02E-125

Tabla A5. Secuencias de los factores de transcripción seleccionados como interesantes

Nombre de la secuencia	Descripción de la secuencia	Secuencia
F1743_c0_g1_i1	ethylene-responsive transcription factor 7-like	<p>>F1743_c0_g1_i1 len=585 GAGCGGCCGCCCGGGCAGGTACGGGGGATACCAAATATACCTTGTTTCCAAGAAGATTCA AAAAGAGCAAGAAAAAAGATTCAATCTTGATTTTGTTCCTTGTTAATCAAATCTTAAAGT TTTGATTTTTAATATCTGGGGTTTGTGTAAAACCTTGATTGGATTCTTGATTTGGGATA ATAAGACTATGGGGTTCAACAAAGATGAAGCAAGTTCATCTTCTCATGTTTTGAGTATCC CAAGAGAAGATACACCCTTCTTGGAAGATTCAACATCTTCTTCTCCTTCCAAGACTT TTGCTAATGTTTTTCATAGCCATAGTTGGAGCTGGAGTCTTGGTCTTCTTACACCTTTA AGAGAAGTGGATGGGTTATGGGTTCTATCATGCTCTTCTCAGTTGCCTTTCTTACTTACT ACTGTATGATGCTTTTGGTCTATTCAAGGCGTAAGGTTGAATCCCACATCAAAGCTACAA AAATTTTCATCTTTTGGTGATTTGGGGTTTGCTGTGTGTGGACCTGTTGGTAGATTGTCCG TTGATCTTATGATTATTCTATCTCAAGCTGGTTTTTGTGTTAGTT</p>
F8972_c0_g1_i1	ethylene-responsive transcription factor CRF4-like	<p>>F8972_c0_g1_i1 len=505 GAGCGGCCGCCCGGGCAGGTACCAGCTCGTAGTTTCTGACCTAGATACTGTATAGCCAAT GTAGAAGCATGACCTGAACCCAATCCAACAACCATACCATTCTTCACATATGAATCCACA GTGTGTTTGGCGGCTAGGAAGAGATCATCAGAGAGAGTAGAACGAGTAACAAAGCGATGG AGGGATTTTCTTGATGAGGTGAAATTATAAGCTTGTTTTTTTAGTAGAAGAAGAAGAAGAA AGAAATGGTGATTTTGAAGAAGGCAAAGCCATAGCCATAGCCATGTTCCCTCATTTCTGTT TCACCTTCCCCGTACCTCGGCCGTGACCACGCTGCCCTATAGTGAGTCGTATTACATGA ATTCCCAAACGAAAACCCAATAATCAAAGCAACCAAACATTGAGACCCACATGAGTAAG CGCCAAGGCAACAACAGTAACCAAGAGCCATAATCTTTCCTAGAAACAGACTTCTACGAC GATAGGCATAAAGGCATGATTAGTT</p>

Nombre de la secuencia	Descripción de la secuencia	Secuencia
F5989_c0_g1_i1	ethylene-responsive transcription factor RAP2-12-like	<p>>F5989_c0_g1_i1 len=335</p> <p>GAGCGGCGCCCGGGCAGGTCTATAGGGCTCGAGCGGCCCGCCCGGGCAGGTACCTCACTC CCGCCGTCGACCGTTACCGCCACCTCATCCTCTCAGAGCACCACCATCCCATCATCACTC CGGCGATCAATCTCACTTCATCAATACCGTTGCAAAGCCTCATCATCTCCGTCTCCGATG TCACTTCACCGCTCGCTCACGGAGTCAACGAATCCTACTCTCTTTCATACCTCCGATG GCTCCTCCGCTACATCACTGCGGGGACTGTATGGGGAGCCATGCGAGGGCTGGAGACAT TCTCACAACCTCGTGTACCTCGGCCGCGACCACGCT</p>
F6602_c0_g1_i1	probable WRKY transcription factor 26	<p>>F6602_c0_g1_i1 len=550</p> <p>TCAAGAACCCTAACTACAAGGGGAAGTGAAGGCTCCTATGATTGACAACCCAGACTTCA AGGATGACCCAGATCTATATGTTTTCCCAAAATTTGAAGTATGTCGGAGTTGAGCTGTGGC AAGTGAATCTGGAACCTTGTTTGACAATGTTGTGATATGCGATGATCCAGAGTATGCCA AGTCCATTGACAGAGGAAACATGGGGAAAGCAGAAGGATGCTGAAAAGGCTGCTTTTGAGG AAGCAGAAAAGAAGAAAGAGGAGGAGGAATCAAAGAATGCTCCAGCTGAATCTGATGCCG AGGAAGATGACGAGGCTGATGAGGCTGATTCAGATGATGCTGATGACAAGTCTGACTCCA AGGATGAGGACACACATGATGAACTGTAAGGGCCAAGCATTGATTCAACTGACAAGGA AGCTACTTGGCAGTTCCCGGAAGATTTCTCCTCCCTCCCCGTTTCTTTTTTTTCCTTA TTAGTGAATTAGCCCCTTCCAATTTTTATTGAAAAGACGGAGATTGAGGGTGCATGAGG TCTTATATGG</p>
F2719_c0_g2_i1	probable WRKY transcription factor 33	<p>>F2719_c0_g1_i1 len=602</p> <p>GTTAATCCACCTTGGCAAGATATGTTAGATCAGCTTCAATGAGGAAACCAGAAGAACC AACCAGAACAAGGTTCCCTCAGAGGCTCAGCCAAGAAAGAAAATGACCCAAATCAAGATTC ATTGTCTGCTGATCCTCAGGTGAGACTAGCCCCCTTATATAGCTATGGCTCATGCTGGAC TTGTATTCACATTTTTCATTATTTATGGGTGTTGGTAAATTGCTTGAAGAGTATTTGAGA GCTATGTTATGGGCTGTGCTTTGTTCAATTCCATTAAGGGGAATTCAACAAACCCCTGTT GCATTTTGGTCTGAACCTTTAAAGTTAGGTCTTACTGAGACTATTTTAGCTGTCCCTGTA GCTGTTTTCAGTGTTTTCTAGGGACCCTAGTTGATATTAAGAGAAGATTTCTAGGGTT GTTCTTAGGAGGCCTAAAGGTAACCTCGACGAGGCGTCATACGAGTGGATTTTCGTGTTG TTGAGGTGGCTTATGCTTTTTGGGGTGTGTTGTGATTGCTTATGAGCAGATTGGTGGGATT GGATCAGTTGCACTTCTTGCATTGGGATTTATGTTTACTGCTAATAGTGTGGATTCAACT AT</p>

Nombre de la secuencia	Descripción de la secuencia	Secuencia
F1440_c0_g6_i1	probable WRKY transcription factor 40	>F1440_c0_g6_i1 len=348 AGCGTGGTCGCGGCCGAGGTACAATATGGGGGAAATCACCAATGTGATGGAGTTTGAAGC AATTGCAAAGCAAAAAGTTGCCAAAGATGGTGTATGATTACTATGCATCTGGAGCAGAGGA CCAATGGACTCTGGCTGAGAACAGAAACGCCTTTTCAAGAATTTTGTGTTAGGCCCGTAT TCTTATTGATGTAAGCAAGATTGACTTAAGCACAACCGTGCTCGGTTTCAATATCTCAAT GCCTATCATGGTTGCACCAACAGCTATGCAGAAAATGGCTCATCCTGATGGTGAGTATGC TATAGCAAGAGCAGCATCGGCAGCAGGAACAATTATGACACTGTCATC
F8036_c0_g1_i1	transcription factor bHLH122-like	>F8036_c0_g1_i1 len=214 AGCGTGGTCGCGGCCGAGGTACAGGTTCCATGCATAATAGACAAGCACTTAGTTCATACA CTGGATGCCCTTCCGCTGCAGAACTCTCTACTATCTGGTCGTTTCTCTTAGATAACCCTT TAGGTGGTGATTCAACTTGTGAGAAGGGAATGGGTTCAATGAGCATCACTGACAACCGAC CTACGGATAGCCGAGCTCCTCAAGGAGCATTGAG
F3220_c0_g1_i1	transcription factor bHLH68	>F3220_c0_g1_i1 len=499 AGCGTGGTCGCGGCCGAGGTACTATTTTTGTTGAGCAATTTGTTTGGTTAGGAAAAGCAGG GTTGATTGATAAGAAGAATTTGAATAGTTTACAGAAAATTAGTGCGTGGTGTGAGTTTAT TGGTTATATTGGAAGTGTGAGTTTGAAGGTTAAGGAATTAAGGCAAATTAGTGAAGATGA GCAGTGTGTTGCTTTCGACTCGTGAGGTATCGATGATCAGGGGGATTGGGTTTGTGATGA AGAGGAGAAAATTGAGGAAATTGAGATTGAAGAGGTTGATGAAGAGGTTATCAGTGATTCA AGATTTTTGCTGATGGGTTAATGGCTTTAGCTGATATTAGTGATGGTAAAGGGCTGCTTTC TGCACCATTTGTTGTTGTCATCAGCAGGGTTGTTGTCAGCTCTTATTAGCACACATAAGAA TTGGATCTCTTGTGATCATCCATAGCTGGATTTGGAGGTATGGTTTGTGCTTAATCAAG TGGAAGTTTACGTTGTGGT
R245_c0_g1_i1	ethylene-responsive transcription factor 5-like	>R245_c0_g1_i1 len=273 GAGCGGCCCGCCGGGCAGGTACAATTCCTATCCCTGGAGTCAAGTCTGTAAAACAACTG AAGACAACCTAGGAGCCCTTGGTTGGCAACTCAGTTCTGATGAATTGCTCCAGTTAGAAT ACGCAGCTCTTGATTCTCCACAGAAGATGATCCAAAACGATTTTTCAGACAAAAGTGACAAC TTCTCCAAATAATAGATATGTTTAGGAAGATATACTTCTAGAAGGCAATATTCGAGCTAA AGGATACCAATGTACCTCGGCCGCGACCACGCT

Nombre de la secuencia	Descripción d ela secuencia	Secuencia
R6221_c0_g1_i1	probable WRKY transcription factor 2 isoform X1	>R6221_c0_g1_i1 len=274 GAGCGGCGCCCGGGCAGGTACATGAAATAATATCCGATCTACATATTCATAATTTGTTA CTCGTTTTTATTATCATTATCAACTATCATCATTGCAACTGCATTGTCCGTTTGACCGAG TAGACTACTCAAAGATCTAGCTTTTACCATTCTAAGCTTCTTCATCTTTTTGTCTTCGT CTTTTTCTTGTCATCTTCATCATCATTATTATCATCATCAATACCATAATATTCATGATC ATTTATCAACGCGTACCTCGGCCGCGACCACGCT
R547_c0_g1_i1	probable WRKY transcription factor 41	>R547_c0_g1_i1 len=991 GAGCGGCCCGCCGGGCAGGTACGACTTGCTATGGAAGGGGTCTGATTGCTCATAAGGATG GATCTGATACAATATGCCTGAATTGTAAAGGTAATGGAAAACCTCCTTGTGCGACTTGTG AATCTCGTGGCCTCATCAAATGTCAAACATGCCAAGGTGGTGGTTCTCTTTTGACACGGA AAGTAGCTGTTGTTTCGATGGAGGACACGTTCAACCCGGAAAGTTAATGCTACAAGCGGAG CAGCTTCTGTTCCAGACGATGTGTTTTACAGAGCAAAGGGAGTTCAGTTGTGCAATAACC AAGCTTACCAGTGCACACCAGCCTACTTTGCTGATTTCGTATTTCCCTTAACAAGTTCCTT CTGAAGTAATTGCAGAGAGACCTTCTATACCTACTACTGCAAGGATGATATGTGAGCGCC ATACTATCTCAGTTGTCCCGGTGACTCGCGTGACCATGACTGATCGTAACCGCTCCTTCA GCTTCTATATTATTGGAAATGACAGAGAGGTTTATATGAAAGACTACTATCCTTCTAGGT TCTGTTGGGGCCTTTGCCCTTGCCCTGGAATGGTTGAAGTTATAATTCAAGAACTGGATTT GTTGTGTGTCATCATTGGTCTGAGTTTGTCCCAAGTGTTCAGGGGAGAACATGTTGGTTT TTGTGTAGAACAATATGCTCTATGCTTTGCACTGTTCCCAAAAAGAAAAGAAAAGTTGT TTTCCATCAATAATGTCATAATCGGGAAATCTTAGTAGCTCAGTTTCGTTGGTTGGCTAT CTGGACTTTCATCCCGTTGGTGATCATTGGTGGTTGGCTACCTTTCATCACGTTGGTGA TCATTGGCTGGCTGGTTACCTGAACCTTAATGTTGTTCTTCATAAGTTGGTCGATTGACT ACTTGAACCTTCATCTCATTGCTTAAGGTTTGATTCTCCACCTTGTAATCTCCTTTTCTA TTACCCTTTGTACCTCGGCCGCGACCACGCT