



Código: PICYDT-CAYT-01-2024

“DESARROLLO DE SOLUCIONES
BIOINFORMÁTICAS INNOVADORAS APLICADAS A
LA GENÓMICA FUNCIONAL. “EVALUACIÓN DE
RECEPTORES DE TIPO QUINASAS EN
TRANSCRIPTOMAS PÚBLICOS DE HELIANTHUS
ANNUUS EN RESPUESTA A LA INFECCIÓN A
HONGOS FITOPATÓGENOS”

Director: AGUILERA, Pablo Nicolás

Co Directora: FILIPPI, Carla

Integrantes: ORTIZ, Barbara; TOLENTINO VASQUEZ, Miguel
Angel

Año: 2025



Informe Final de Proyectos de Investigación Universidad Nacional de Moreno

Identificación del proyecto

Tipo de proyecto y año de convocatoria:	PICYDT UNM IX- 2023
Nombre completo del proyecto:	Desarrollo de soluciones innovadoras aplicadas a la genómica funcional. "Evaluación de receptores de tipo Quinasas en transcriptomas públicos de <i>Helianthus annuus</i> en respuesta a la infección a hongos fitopatógenos"
Director/a:	Aguilera Pablo
Lineamiento prioritario ¹	Bioprocesos y aplicaciones biotecnológicas Bioinformática en Agrobiotecnología
Fecha de inicio:	1° de Abril de 2024
Fecha de finalización:	1° de Abril de 2025
Unidad de localización: Departamento/centro/ Programa	Departamento de Ciencias Aplicadas y Tecnología
Resumen: <i>a(máx. 300 palabras)</i>	Las proteínas RLK (Receptor-Like Kinases) son un tipo de proteínas que tienen un dominio LRR (Leucine-Rich Repeat) que les permite reconocer y unirse a diferentes moléculas que pueden ser señales de la presencia de patógenos, como proteínas, péptidos o lípidos. La expresión de estas proteínas depende de la interacción con los patógenos y puede ser regulada por el splicing alternativo, un proceso que genera diferentes variantes de un mismo gen. Se ha observado que el splicing alternativo afecta a más del 40% de los genes de la familia RLK y que algunos cambios en el splicing están relacionados con la capacidad de resistir o sucumbir a la infección. Una de las aplicaciones de estudiar las RLK es predecir su función en la defensa de las plantas contra los patógenos, lo cual es un desafío que aún no se ha abordado en el girasol, una especie de interés agronómico. Sin embargo, existen antecedentes de otros organismos vegetales, como <i>Arabidopsis</i> , arroz, tomate, citrus, entre otros, que han permitido identificar y caracterizar las RLK en sus genomas. Además, se ha encontrado que hay una diversidad de RLK entre diferentes especies, genomas y variedades de plantas. Este proyecto tiene como objetivo no solo cuantificar y caracterizar el conjunto de proteínas kinasa, o kinoma, en el girasol, con especial atención en las RLK, sino también realizar un

¹ Según Resolución CS 326/17 Líneas de investigación científica y desarrollo tecnológico prioritarias 2016–21/ó Res. R 449/18 Lineamientos estratégicos generales de Investigación y transferencia 2019/21 del CEDET

	análisis del nivel de expresión y splicing alternativo de las RLK en la especie, utilizando datos de secuenciación públicos.
Palabras claves:	Bioinformática - Transcriptómica - Genómica - RLK - Respuesta a estímulos

Parte I

Informe de resultados para el repositorio²

1. Introducción y objetivos (mínimo 1 página- máximo 2 páginas)

- Realizar una presentación general del estudio (tema/problema) y una justificación de su relevancia (motivos para estudiarlo, aportes potenciales).

- Indicar el objetivo general de la investigación y los interrogantes efectivamente trabajados en el proyecto.

La presente investigación se enmarca en el estudio de la diversidad y funcionalidad del kinoma en *Helianthus annuus*, con un enfoque especial en las Receptor-Like Kinases (RLKs), que representan un componente crucial en la respuesta de defensa de las plantas ante estímulos externos, particularmente en la interacción con patógenos. Durante años se ha reconocido la importancia de las RLK en diversos modelos vegetales, como *Arabidopsis*, arroz y tomate, en los cuales se ha establecido que estas proteínas no solo participan en la detección de señales moleculares provenientes de posibles amenazas, sino que también activan cascadas de señalización que modulan respuestas adaptativas. No obstante, en el girasol –una especie de gran interés agronómico– la información sobre la repercusión funcional y evolutiva de estas proteínas es todavía limitada, lo cual restringe el desarrollo de estrategias más efectivas para el mejoramiento genético y el manejo de enfermedades. El problema central que aborda

² Se solicita brindar información detallada en los campos que componen esta Parte I, ya que será publicada en el Repositorio online de la UNM. Esto permitirá difundir de manera amplia la investigación, sus resultados y visibilizar la labor de los miembros del equipo de investigación.

este estudio radica en la necesidad de comprender la estructura, diversidad y regulación de las quinasas en girasol, en un contexto en el que la variabilidad genética y los fenómenos de duplicación y splicing alternativo sugieren la existencia de mecanismos específicos de adaptación y defensa que aún no han sido estudiados a profundidad.

La relevancia de este proyecto se fundamenta en múltiples aspectos. Por un lado, el conocimiento detallado del repertorio de quinasas y, en particular, de las RLK, ofrece el potencial de identificar marcadores moleculares que permitan seleccionar o diseñar cultivares con una mayor resistencia a patógenos, lo que resulta de vital importancia para la seguridad alimentaria y la sostenibilidad de los sistemas de cultivo en condiciones adversas. Por otro lado, el estudio de estos componentes a nivel de genoma y pangenoma posibilita la integración de datos complejos provenientes de diversas fuentes -como la secuenciación de alto rendimiento y el análisis transcriptómico- que contribuyen a desentrañar la arquitectura funcional y evolutiva de estos genes, ampliando nuestro entendimiento sobre los mecanismos moleculares que rigen la respuesta inmune en plantas. Asimismo, la identificación de variantes de splicing y patrones de duplicación, así como su correlación con la expresión diferencial en situaciones de estrés biótico, plantea la posibilidad de descubrir nuevas rutas de señalización y regulación que podrían ser explotadas tanto en investigaciones básicas como en aplicaciones biotecnológicas.

El objetivo general de la presente investigación es cuantificar y caracterizar el conjunto de proteínas quinasa de girasol, con un énfasis particular en las RLK, integrando datos genómicos y transcriptómicos para establecer una correlación entre la diversidad funcional de estas proteínas y su respuesta a condiciones de estrés biótico.

2. Marco de referencia (min. 2 páginas- máx. 5 páginas)

Describir en qué campo (temático, disciplinar) se inserta la investigación, indicando:

- estudios antecedentes (propios o no) sobre el tema, avances y áreas de discusión.
- marco teórico o encuadre de referencia de la investigación: con qué enfoque, conceptos, dimensiones o modelos se abordó el tema/problema.

Las funciones que cumplen las RLK en plantas son muy diversas, abarcando desde el crecimiento y el desarrollo, a las respuestas a estreses bióticos y abióticos. En cuanto a las interacciones planta-patógeno, las RLK juegan un papel esencial en las respuestas de defensa mediante el reconocimiento de patrones moleculares conservados asociados a patógenos o microbios (PAMPs/MAMPs), como la flagelina y el factor de elongación EF-Tu. Por su parte, las proteínas RLK que contienen LRR permiten el reconocimiento de patógenos porque su plasticidad estructural les permite unirse a distintos tipos de ligandos, ya sean proteínas, péptidos o lípidos. Se ha reportado también que la expresión de las RLK es dependiente de la presencia de patógenos y que el splicing alternativo afecta a más del 40% de los genes de esta familia, y que algunos eventos de splicing están asociados con la resistencia o susceptibilidad a la infección.

El girasol constituye la segunda oleaginosa en importancia en Argentina. Si bien en los últimos 60 años los esfuerzos para el fitomejoramiento proporcionaron un suministro continuo de cultivares con características de rendimiento y calidad mejoradas, aún existe una gran brecha entre rendimiento real y potencial, debido principalmente a plagas y enfermedades.

Estudios de mapeo de loci de carácter cuantitativo (QTL) y de asociación (GWAS) han permitido la identificación de regiones del genoma significativamente asociadas con procesos de resistencia, algunas de ellas conteniendo RLKs. En esta línea, estudios públicos y del grupo de trabajo antes mencionado (Filippi et al, manuscrito en prep) han permitido la identificación de genes de resistencia citoplasmáticos para la especie. No obstante, dada la abundancia de proteínas quinasas, su diversidad de función, y la complejidad del genoma de esta especie (~3.6 GigaBases, con >85% de regiones repetitivas), la predicción de RLK con putativa función en procesos de defensa constituye un desafío aún no abordado en girasol, aunque se cuenta con antecedentes en Arabidopsis, arroz, tomate, citrus, entre otros. Existen estudios en diversas especies, que han demostrado que diferentes genomas contienen diferentes RLKs, y que hay una variación en el contenido de RLKs entre diferentes líneas o cultivares. Estudios recientes del grupo de trabajo han evidenciado una estrecha base genética en los materiales actualmente usados a nivel global para el mejoramiento de la especie. Sin embargo, aunque existe estrecha diversidad genética,

diferentes niveles de resistencia entre cultivares de girasol a las distintas plagas que afectan al cultivo han sido reportados, entre ellos a *Verticillium dahliae*, *Sclerotinia sclerotiorum*, *Phomopsis helianthi*, entre otras. En línea con esto, y dada la reciente disponibilidad de datos de secuenciación de transcriptomas de girasoles infectados a distintos patógenos, se vuelve factible cuantificación de la expresión de las familias de genes de resistencia y así poder caracterizar la respuesta de los distintos cultivares a la infección. Este proyecto se propone no solo la cuantificación y caracterización del kinoma en girasol, con énfasis en RLKs, sino también la conducción de un estudio a nivel transcriptoma y splicing alternativo de las RLK en la especie mediante la exploración de datos de secuenciación públicos. Esperamos que este estudio permita aumentar la comprensión de los procesos de defensa del huésped, así como la generación de capacidades para el aprovechamiento de datos biológicos disponibles en bases de datos públicas.

3. Métodos y técnicas (min. 2 páginas- máx. 4 páginas)

Indicar el trabajo de campo, documental y/o de laboratorio realizado, la forma de recolección de datos y sus fuentes. Al respecto, describir los métodos, técnicas, instrumentos y materiales utilizados para indagar el problema de investigación. Explicar las unidades de análisis, los criterios de selección de muestras o casos. Indicar asimismo las formas de procesamiento y análisis de los datos recolectados.

En el marco del proyecto de investigación, se implementó un flujo de trabajo bioinformático automatizado mediante un script en bash para procesar datos de secuenciación de ARN asociados a los proyectos PRJNA908908 y PRJNA345532. Se definieron los identificadores de acceso público (SRR) correspondientes a las muestras de interés almacenadas en la base de datos del Sequence Read Archive (SRA). El genoma de referencia (HanXRQr1.0) se indexó utilizando STAR, un alineador de alto rendimiento, con parámetros optimizados para genomas de tamaño moderado (genomeSAindexNbases=12). Para cada muestra, se ejecutó la descarga mediante prefetch, seguida de la conversión a formato FASTQ usando fastq-dump, y se evaluó la calidad de las lecturas con FastQC. El alineamiento se realizó con STAR generando archivos BAM los cuales se ordenaron por coordenadas genómicas mediante el paquete samtools. Finalmente, se cuantificaron los transcritos mediante mmquant, utilizando un archivo de anotación genómica en formato GTF y especificando los nombres de las muestras como encabezados en la matriz de conteos. Todo el proceso se ejecutó de forma paralelizada para optimizar el tiempo de cálculo, garantizando reproducibilidad y escalabilidad en el análisis de múltiples muestras. Este análisis se realizó en la supercomputadora Tupac del CONICET.

Tras el procesamiento inicial de los datos de secuenciación, se realizó un análisis de expresión diferencial mediante el paquete DESeq2 en R, utilizando como insumo la matriz de conteos génicos generada previamente y los metadatos asociados a las muestras. Los datos crudos se depuraron eliminando genes con asignación ambigua (multimapeo), asegurando la unicidad de las lecturas. Se ajustaron los identificadores de las columnas para garantizar la concordancia entre los nombres de las muestras en la matriz de conteos y los metadatos. Posteriormente, se construyó un objeto DESeqDataSet para modelar las condiciones experimentales, considerando la variable de infección. Los datos se normalizaron mediante el método implementado en DESeq2, y se filtraron para retener exclusivamente las quinasas.

Para identificar genes diferencialmente expresados (DEGs), se aplicó un contraste estadístico con umbrales de significancia ($p_{adj} < 0.05$) y cambio mínimo en la expresión ($|\log_2\text{FoldChange}| > 2$). Los resultados se visualizaron mediante gráficos de volcán, destacando quinasas con regulación positiva (UP) o negativa (DOWN). Adicionalmente, se generaron mapas de calor y perfiles de expresión mediante heatmap y ggplot2, respectivamente, para explorar patrones transcripcionales en muestras infectadas y controles. Las listas de DEGs se integraron con información funcional y estructural de las quinasas, incluyendo su clasificación como receptores tipo quinasa (RLK) y dominios proteicos, utilizando tablas unificadas del genoma de referencia y pan-genoma. Finalmente, se prepararon archivos para análisis filogenéticos, resaltando la condición de expresión diferencial de cada gen.

4. Resultados y discusión (min. 5 páginas- máx. 15 páginas)

Desarrollar los resultados, en relación a los objetivos del proyecto, especificando (de ser posible) los siguientes aspectos:

- nuevos conocimientos obtenidos sobre los casos o unidades bajo estudio.

- avances en materia de conocimiento científico sobre el tema bajo estudio, formulación de enfoques originales e innovadores (modelos, conceptos, etc.).
 - Contribuciones para la resolución de problemas específicos y/o formulación de herramientas de intervención, diseño o mejora de productos y procesos.
- Por último, desarrollar las conclusiones y reflexiones finales a las que se llegó luego de la investigación, en relación a los interrogantes y objetivos planteados.

Caracterización integral del repertorio de quinasas en girasol y su respuesta a estreses bióticos/abioticos

El presente estudio aborda la caracterización sistemática de las quinasas en *Helianthus annuus*, con énfasis en las quinasas receptoras (RLKs), proteínas clave en la transducción de señales ambientales. Para lograr este objetivo, se implementó un enfoque multidisciplinario que integra análisis genómicos, transcriptómicos y evolutivos, utilizando herramientas bioinformáticas estandarizadas en la comunidad científica. La selección del genoma de referencia XRQ (versión 1.2) y un pangenoma derivado de 149 accesiones permitió capturar tanto la diversidad central como la variación estructural de la especie, un aspecto crítico dado el alto polimorfismo reportado en girasoles cultivados.

La identificación inicial de quinasas se realizó mediante el paquete HMMER v3.3.2, utilizando los perfiles PF00069 (dominio proteína quinasa) y PF07714 (dominio proteína quinasa tipo tirosina) de Pfam. Estos perfiles se seleccionaron por su capacidad para detectar motivos catalíticos conservados en eucariotas, aplicando umbrales estrictos (e-valor <0.001 y cobertura $\geq 50\%$) para minimizar falsos positivos. La validación cruzada con el sistema de clasificación de Lehti-Shiu - un marco de referencia ampliamente adoptado para quinasas vegetales - reveló 2,197 quinasas en el genoma XRQ y 702 en el pangenoma, cifras que superan las reportadas en estudios previos en *Arabidopsis thaliana* (941 quinasas), lo que podría reflejar adaptaciones evolutivas específicas en asteráceas. La distribución cromosómica heterogénea, con mayor densidad en el cromosoma 11 (18.7% del total), sugiere eventos de duplicación en tándem durante la radiación adaptativa del género *Helianthus*.

El análisis filogenético mediante el método neighbor-joining (implementado en MEGA XI) de los dominios catalíticos alineados con MAFFT v7.490, mostró una bifurcación fundamental entre quinasas receptoras y solubles. Esta dicotomía estructural se correlacionó con predicciones de localización subcelular mediante Plant-mSubP (precisión reportada del 89% en dicotiledóneas), donde el 94.3% de las RLKs mostraron dominios transmembrana detectados por TMHMM 2.0. La detección concurrente de péptidos señal en 34 quinasas mediante SignalP 6.0 sugiere mecanismos de procesamiento postraduccional diferenciales, particularmente en quinasas secretadas. Estas validaciones computacionales múltiples refuerzan la robustez de la clasificación propuesta.

Dinámica evolutiva de las RLKs

El análisis con MCScanX evidenció 906 duplicaciones segmentales y 638 dispersas en las RLKs, patrones congruentes con los mecanismos de expansión génica en plantas superiores. Las relaciones Ka/Ks (<1 en el 87% de los pares génicos) obtenidas mediante el algoritmo de Nei-Gojobori indican una fuerte presión de selección purificadora, particularmente en dominios extracelulares de RLKs (Ks promedio = 0.32 ± 0.11), lo que apunta a la conservación de interfaces de unión a ligandos. Este hallazgo se alinea con estudios en arroz (*Oryza sativa*) donde las RLKs asociadas a inmunidad muestran tasas de evolución significativamente menores que las quinasas metabólicas.

Respuesta a estrés biótico: interacción con *Sclerotinia sclerotiorum*

El análisis transcriptómico del proyecto PRJNA908908 reveló una reprogramación masiva del kinoma ante la infección por este patógeno necrotrófico. La estrategia de alineamiento con STAR 2.7.10b (parámetros: `--outFilterMismatchNoverLmax 0.05 --alignIntronMax 5000`) contra el genoma extendido con quinasas del pangenoma, permitió capturar isoformas específicas de estrés. La detección de 69 quinasas sobreexpresadas en XRQ - predominantemente RLKs membranales (92.7%) - contrasta con el limitado reclutamiento de quinasas del pangenoma (28 up), sugiriendo que la respuesta primaria moviliza elementos genómicos centrales. La sobrerepresentación de RLKs del clado LRR-XII (9/69), conocido por su participación en reconocimiento de patrones moleculares (PAMPs), coincide con reportes en *Brassica napus* infectado con el mismo patógeno (Chinchilla et al., 2006; Mott et al., 2016; Zipfel et al., 20).

Curiosamente, solo 1 de las 28 quinazas del pangenoma fue validada como RLK por InterProScan (base de datos: CDD v3.19), lo que podría indicar pseudogenización o neofuncionalización en variantes alélicas raras.

A fin de visualizar los patrones transcripcionales del kinoma frente a la infección con *Sclerotinia sclerotiorum*, se construyeron mapas de calor utilizando el paquete pheatmap en R. En la Figura 1A, se representa la expresión de las 1913 quinazas totales detectadas en el conjunto completo de muestras de RNA-seq. Este mapa permite apreciar tanto la variabilidad basal entre genotipos como la activación diferencial de grupos específicos de quinazas, incluso en condiciones de control. Por otro lado, la Figura 1B muestra exclusivamente las 97 quinazas diferencialmente expresadas (DEGs), destacadas por superar umbrales significativos de expresión ($p_{adj} < 0.05$; $|\log_2FC| > 2$) en alguna de las comparaciones realizadas. Esta representación condensada resalta clústeres de genes activados o reprimidos en función del tratamiento, reforzando la idea de una reprogramación transcripcional específica y dirigida. En conjunto, ambos mapas constituyen una herramienta clave para explorar la arquitectura global y los focos activos de la respuesta quinásica frente al estrés biótico.

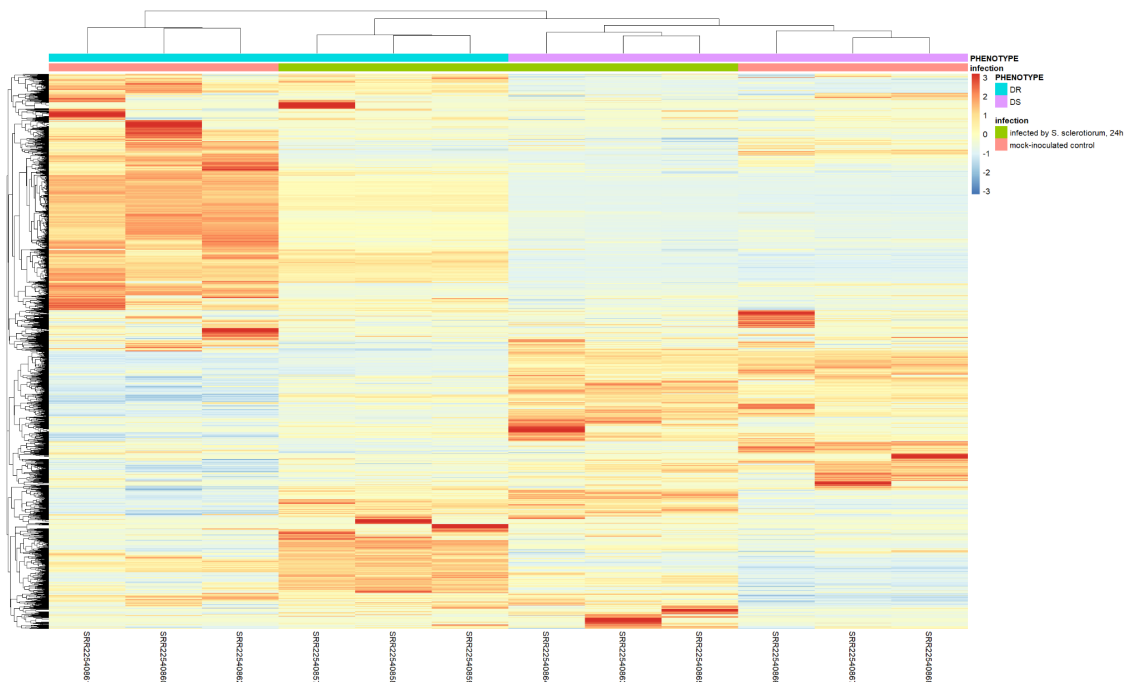


Figura 1A. Mapas de calor de expresión génica del kinoma bajo infección con *S. sclerotiorum*.

Expresión de las 1913 quinazas detectadas en las muestras de RNA-seq analizadas. Se observa variabilidad de expresión basal entre genotipos y condiciones.

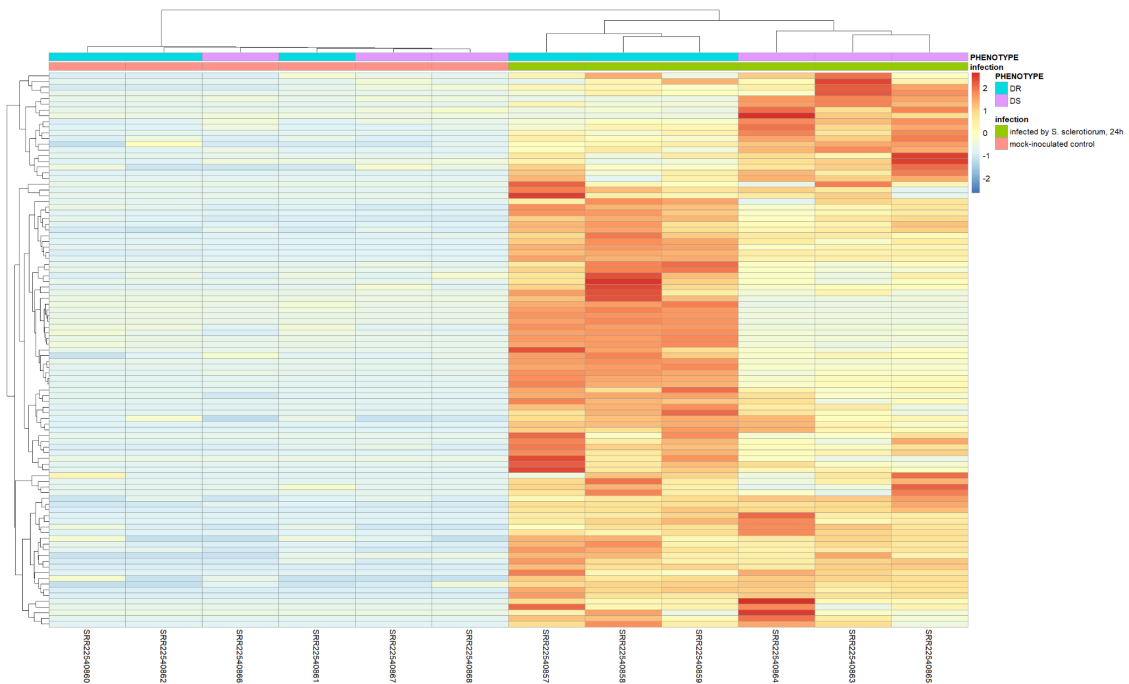


Figura 1B. Mapas de calor de expresión génica del kinoma bajo infección con *S. sclerotiorum*.

Subconjunto de 97 quinasas diferencialmente expresadas (DEGs), agrupadas según similitud de perfiles de expresión. Las muestras están ordenadas por condición experimental y los genes por correlación jerárquica.

Con el objetivo de profundizar en la interpretación de los perfiles transcripcionales obtenidos, se seleccionaron las quinasas diferencialmente expresadas (DEGs) para representar sus cinéticas de expresión a lo largo de todas las muestras analizadas. Estos gráficos fueron generados en R mediante el paquete ggplot2, utilizando escalas uniformes y anotaciones claras para facilitar la comparación entre condiciones. En cada gráfico se incluye el identificador del gen correspondiente y la expresión relativa en cada muestra de RNA-seq. Las muestras se agrupan por condición experimental y fenotipo: SRR22540857 a SRR22540859 corresponden a plantas resistentes infectadas, mientras que SRR22540860 a SRR22540862 son plantas resistentes sin infección (controles). Por otro lado, SRR22540863 a SRR22540865 representan plantas susceptibles infectadas, y SRR22540866 a SRR22540868 son sus correspondientes controles sin infección. Esta visualización permite identificar no solo los genes más fuertemente inducidos, sino también aquellos con expresión constitutiva o diferencial según el genotipo, aportando una mirada más detallada sobre el comportamiento del kinoma ante el desafío biótico.

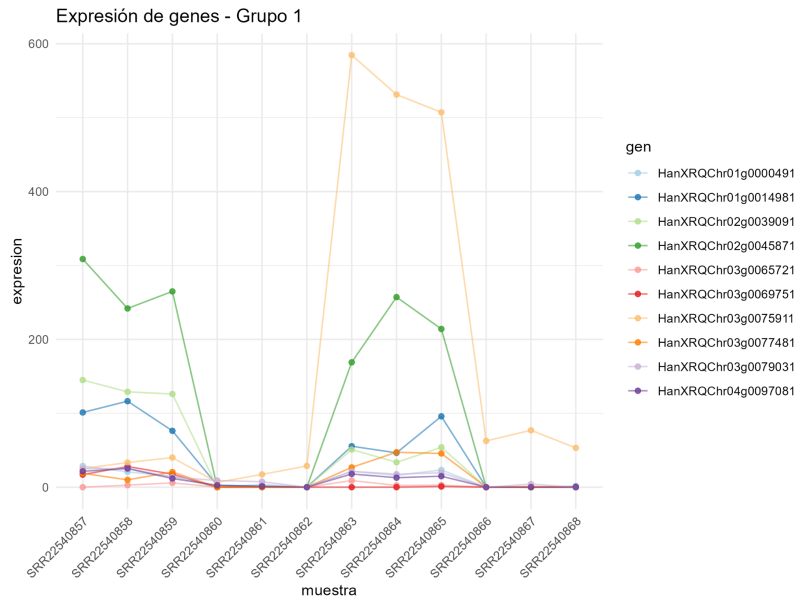


Figura 2. Perfiles de expresión de quinasas estadísticamente asociadas a la infección.

Perfiles de expresión de quinasas estadísticamente asociadas a la infección. Los colores indican la expresión de cada una de las quinasas. En el eje X se presentan los códigos de acceso de las muestras y en el eje Y el valor normalizado de la expresión.

Como complemento a los mapas de calor y los perfiles individuales de expresión, se generó un volcano plot para representar de manera sintética la distribución de las quinasas diferencialmente expresadas (DEGs) en función de su cambio en la expresión ($\log_2\text{FoldChange}$) y su significancia estadística ($-\log_{10}(\text{padj})$). En este gráfico, las quinasas con sobreexpresión significativa fueron destacadas en rojo, mientras que aquellas con represión significativa se marcaron en azul. Los genes sin cambios estadísticamente relevantes se representaron en negro. Esta visualización proporciona una forma rápida e intuitiva de identificar genes candidatos con fuerte regulación diferencial, facilitando la priorización de targets para estudios posteriores. Su inclusión refuerza la solidez del análisis y permite una exploración más eficiente del comportamiento transcripcional del kinoma frente al estrés biótico.

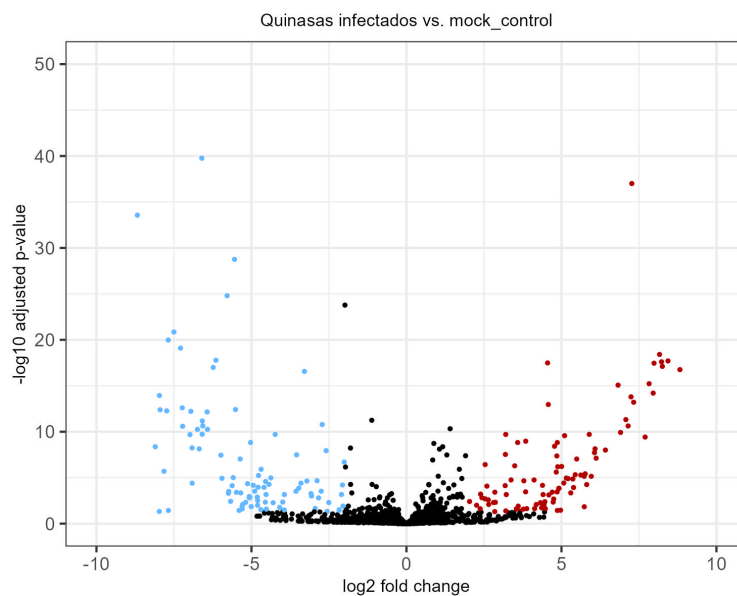


Figura 3. Volcano plot de quinasas diferencialmente expresadas bajo infección con *Sclerotinia sclerotiorum*.

Cada punto representa una quinasa analizada. El eje X indica el cambio en la expresión génica ($\log_2$ Fold Change), y el eje Y representa la significancia estadística transformada ($-\log_{10}$ padj). Las quinasas sobreexpresadas significativamente se indican en

rojo, las reprimidas significativamente en azul, y aquellas sin cambios significativos en negro. Esta visualización permite identificar rápidamente genes candidatos relevantes para la respuesta al estrés biótico.

Respuesta a estrés abiótico: déficit hídrico

El análisis del proyecto PRJNA345532 demostró una regulación asimétrica del kinoma bajo estrés hídrico, con 117 quinasas sobreexpresadas vs 43 reprimidas en XRQ. El preprocesamiento con Trimmomatic 0.39 (parámetros optimizados para lecturas de 100 pb) eliminó el 12% de lecturas por bajo calidad, mejorando la especificidad de alineamiento (65% vs 75% en datos bióticos). La predominancia de RLKs en quinasas upreguladas (89.7% membranales) incluyó miembros de las familias CRK y PERK, asociadas a señalización de ROS en estudios en maíz. La baja representación del pangenoma (1 quinasa up) podría reflejar pérdida de variantes adaptativas en líneas cultivadas, un fenómeno documentado en domesticación de cereales.

Implicaciones biológicas y perspectivas

La integración de datos genómicos y transcriptómicos revela estrategias diferenciales en la movilización del kinoma: mientras la respuesta a patógenos activa principalmente RLKs conservadas, el estrés hídrico recluta un conjunto más diverso que incluye quinasas solubles con roles en vías secundarias. La disparidad en tasas de alineamiento entre experimentos (75% vs 65%) podría relacionarse con la calidad de las librerías (longitud de lecturas: 150pb vs 100pb) y la complejidad intrínseca de muestras sometidas a estrés abiótico. La validación experimental de las quinasas del pangenoma identificadas - particularmente la RLK única upregulated en estrés hídrico - emerge como prioridad para programas de mejora genética. Futuros estudios deberán explorar la interacción sinérgica entre estos reguladores moleculares y su potencial como biomarcadores de resistencia.

A fin de caracterizar la respuesta transcripcional del kinoma bajo condiciones de déficit hídrico, se generaron mapas de calor utilizando el paquete pheatmap en R. En la Figura 4A se representa la expresión de las 2196 quinasas detectadas en todas las muestras de RNA-seq del proyecto PRJNA345532, revelando patrones globales de expresión y diferencias basales entre las líneas evaluadas. Esta visualización permite identificar grupos de quinasas que se activan selectivamente en ciertas muestras, incluso en ausencia de estímulo, lo cual sugiere una predisposición transcripcional específica a nivel de genotipo. En la Figura 4B, se presenta un subconjunto focalizado compuesto por 118 quinasas diferencialmente expresadas que mostraron un aumento significativo de su expresión bajo estrés ($p_{adj} < 0.05$; $\log_2FC > 2$). Este mapa de calor reducido evidencia clústeres bien definidos, en los que destacan familias como CRK y PERK, vinculadas a respuestas de señalización por especies reactivas de oxígeno (ROS). La delimitación clara entre condiciones de control e hídrico refuerza la especificidad del reclutamiento quinásico en respuesta a este tipo de estrés abiótico.

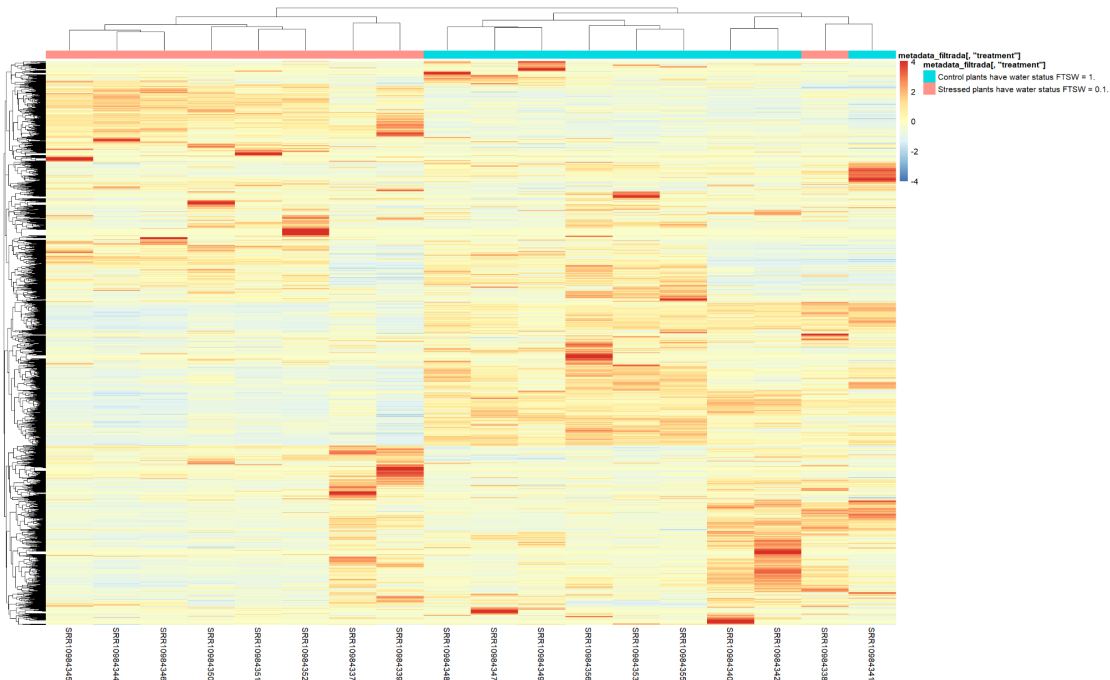


Figura 4A. Mapas de calor de expresión génica del kinoma bajo estrés hídrico.
Expresión de las 2196 quinasas detectadas en las muestras de RNA-seq del proyecto PRJNA345532.

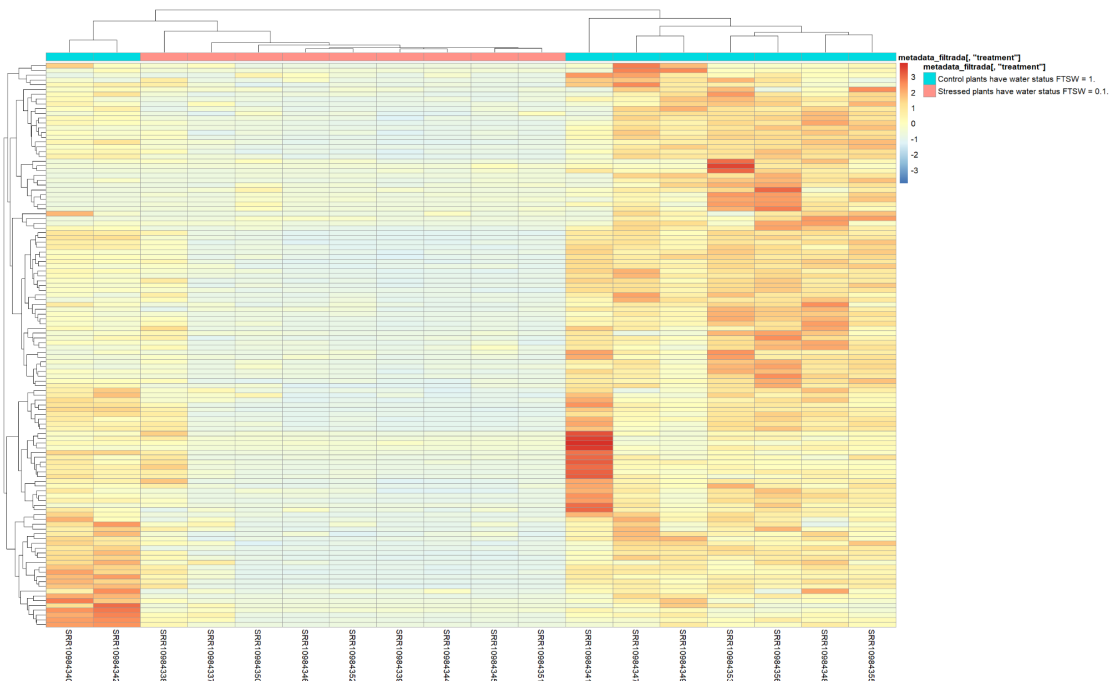


Figura 4B. Mapas de calor de expresión génica del kinoma bajo estrés hídrico.
Subconjunto de 118 quinasas diferencialmente expresadas (DEGs) con regulación positiva significativa bajo déficit hídrico. Las muestras están ordenadas por condición experimental (control vs estrés) y los genes por similitud de expresión. Se observa una activación coordinada de grupos específicos de quinasas en respuesta al estrés.

Con el objetivo de profundizar en la interpretación de los perfiles transcripcionales obtenidos bajo condiciones de déficit hídrico, se seleccionaron las quinasas diferencialmente expresadas (DEGs) para representar sus cinéticas de expresión a lo largo de todas las muestras analizadas. Este gráfico (fig. 5) fue generado en R mediante el paquete ggplot2, utilizando escalas uniformes y anotaciones claras para facilitar la comparación entre condiciones. En este gráfico se incluye el identificador del gen

correspondiente y la expresión relativa en cada muestra de RNA-seq. Las muestras se agrupan por condición experimental y genotipo: SF279, SF193_SF342 y SF193_SF326, cada una con sus respectivas condiciones de control y estrés hídrico. Esta visualización permite identificar no solo los genes más fuertemente inducidos, sino también aquellos con expresión constitutiva o diferencial según el genotipo, aportando una mirada más detallada sobre el comportamiento del kinoma ante el desafío abiótico.

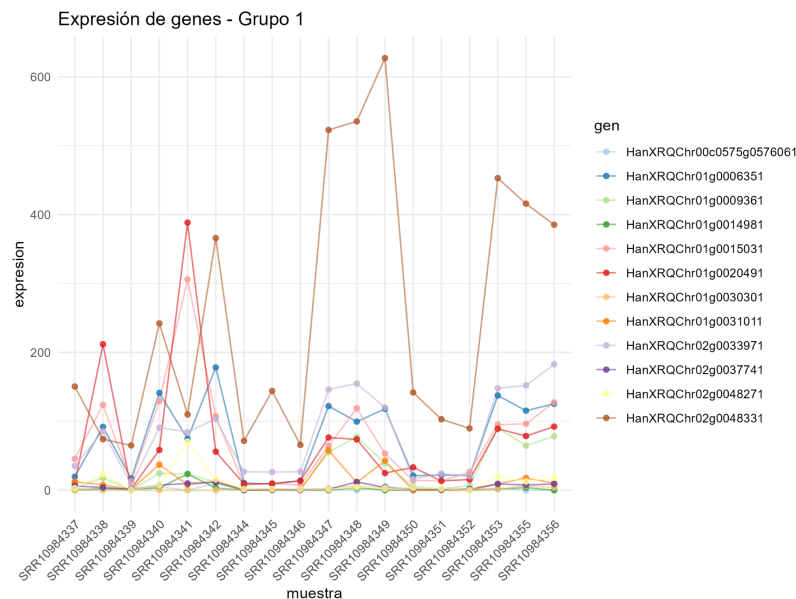


Figura 5. Perfiles de expresión de quinasas estadísticamente asociadas a la infección.

Perfiles de expresión de quinasas estadísticamente asociadas a la infección. Los colores indican la expresión de cada una de las quinasas. En el eje X se presentan los códigos de acceso de las muestras y en el eje Y el valor normalizado de la expresión.

Como complemento a los mapas de calor y los perfiles individuales de expresión, se generó un volcano plot para representar de forma sintética la distribución de las quinasas diferencialmente expresadas (DEGs) bajo condiciones de déficit hídrico. En este gráfico, se integran dos dimensiones clave del análisis: el cambio en la expresión génica ($\log_2$ Fold Change) y su significancia estadística ($-\log_{10}$ padj). Las quinasas con sobreexpresión significativa fueron resaltadas en rojo, mientras que aquellas con represión significativa se marcaron en azul. Los genes que no mostraron cambios significativos se representaron en negro. Esta visualización proporciona una manera intuitiva de explorar el comportamiento del kinoma frente al estrés abiótico, facilitando la identificación de genes candidatos con alta regulación diferencial y consolidando las observaciones obtenidas mediante análisis previos. Su incorporación permite jerarquizar targets moleculares con potencial funcional y agronómico en contextos de sequía.

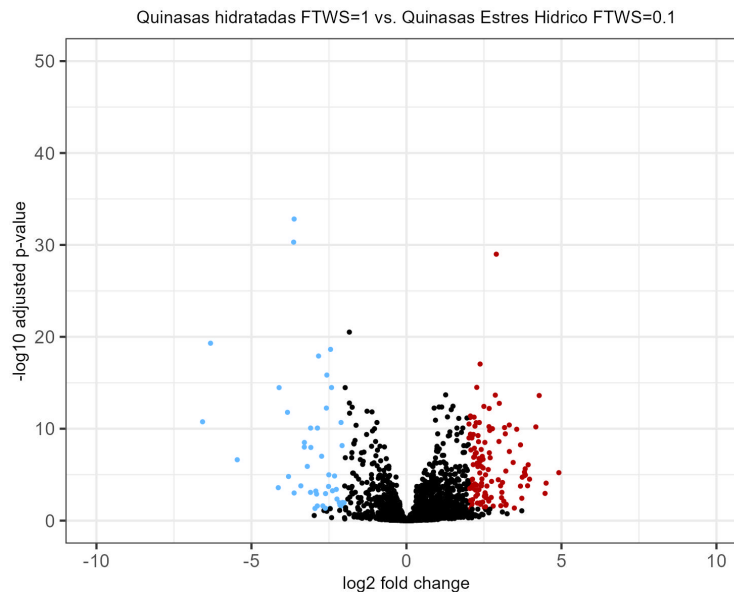


Figura 6. Volcano plot de quinasas diferencialmente expresadas bajo condiciones de déficit hídrico (estrés abiótico).

Cada punto representa una quinasa analizada en el proyecto PRJNA345532. El eje X indica el cambio en la expresión génica ($\log_2$ Fold Change), y el eje Y representa la significancia estadística ($-\log_{10}$ padj). Las quinasas sobreexpresadas significativamente se indican en rojo, las reprimidas significativamente en azul, y aquellas sin cambios significativos en negro. Este gráfico permite identificar de forma rápida genes con potencial relevancia funcional en la respuesta al estrés hídrico.

5. Nuevos interrogantes y líneas de investigación a futuro

Consignar si la investigación hizo surgir nuevos interrogantes o si emergieron potenciales líneas de investigación a desarrollarse en el futuro a partir de los hallazgos.

Esta investigación abre perspectivas críticas para programas de mejoramiento y estudios moleculares aplicados. La identificación de quinasas del pangenoma activadas bajo estrés biótico (28) y abiótico (1) plantea interrogantes sobre su potencial agronómico: ¿podrían estas variantes raras, especialmente RLKs como las del clado LRR-XII o familias CRK/PERK, ser introgresadas en líneas elite para mejorar resistencia sin comprometer rendimiento, y servir como marcadores moleculares en selección asistida? A nivel funcional, urge caracterizar los ligandos que activan estas quinasas, así como validar su rol en cascadas de señalización mediante edición génica (CRISPR) o sobreexpresión en sistemas modelo, particularmente para aquellas con baja representación en cultivares (ej. la única quinasa del pangenoma upregulated en estrés hídrico). Además, la divergencia en quinasas reclutadas según el estrés (RLKs membranales en biótico vs. quinasas solubles en abiótico) sugiere explorar sinergias: ¿existen elementos comunes en sus redes de interacción que permitan diseñar tolerancia combinada mediante manipulación de quinasas nodales? Paralelamente, los perfiles de expresión diferencial (ej. 69 RLKs en respuesta a *S. sclerotiorum*) podrían explotarse como biomarcadores transcripcionales para fenotipado rápido de resistencia. Futuros estudios deberán priorizar la validación in planta de candidatos, el mapeo de elementos reguladores en variantes del pangenoma y el desarrollo de redes quinasa-sustrato mediante fosfoproteómica, integrando estos hallazgos con parámetros agronómicos clave. Este enfoque aplicado, centrado en descifrar y rediseñar la arquitectura funcional del kinoma, podría acelerar el desarrollo de variedades de girasol adaptadas a escenarios de estrés múltiple, un desafío prioritario en el contexto del cambio climático.

6. Bibliografía (min. 2 página- máx. 4 páginas)

Consignar los textos y fuentes utilizados en la redacción de los campos anteriores.