

Desarrollo y optimización de un pipeline para el análisis dual del transcriptoma planta/patógeno: identificando los mecanismos de patogenicidad y defensa

Jorge Martín García



Máster en Bioinformática y Bioestadística

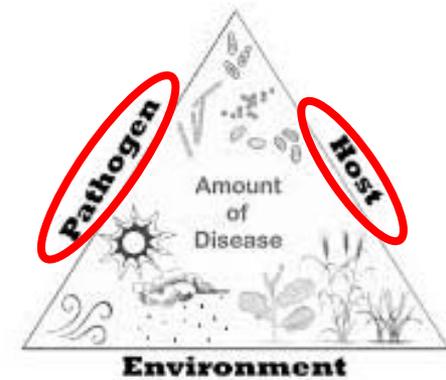
Junio, 2018

Introducción

Patología vegetal - Enfermedad



Patosistema



Introducción

Estudios de transcritos individuales (hasta finales de los 90s)

- Northern blott
- RT-qPCR



PERGAMON

Phytochemistry 57 (2001) 1187–1195

PHYTOCHEMISTRY

www.elsevier.com/locate/phytochem

Planta (2002) 214: 895–901
DOI 10.1007/s00425-001-0698-9

ORIGINAL ARTICLE

Chang-Kui Ding · Chien Yi Wang · Kenneth C. Gross
David L. Smith

Jasmonate and salicylate induce the expression of pathogenesis-related-protein genes and increase resistance to chilling injury in tomato fruit

Two cinnamoyl-CoA reductase (CCR) genes from *Arabidopsis thaliana* are differentially expressed during development and in response to infection with pathogenic bacteria

Virginie Lauvergeat^a, Christophe Lacomme^b, Eric Lacombe^a, Eric Lasserre^b,
Dominique Roby^b, Jacqueline Grima-Pettenati^{a,*}



Contents lists available at ScienceDirect

Physiological and Molecular Plant Pathology

journal homepage: www.elsevier.com/locate/pmpp



MOLECULAR PLANT PATHOLOGY

DOI: 10.1111/mpp.12346

Identification of pathogenicity-related genes in *Fusarium oxysporum* f. sp. *cepae*

ANDREW TAYLOR^{1,*}, VIKTÓRIA VÁGÁNY¹, ALISON C. JACKSON¹, RICHARD J. HARRISON²,
ALESSANDRO RAINONI^{1,†} AND JOHN P. CLARKSON¹

Relative expression of seven candidate genes for pathogen resistance on *Pinus radiata* infected with *Fusarium circinatum*

Andrea Donoso^a, Victoria Rodriguez^a, Angela Carrasco^a, Rodrigo Ahumada^b,
Eugenio Sanfuentes^a, Sofia Valenzuela^{a,*}

LA ERA DE LAS ÓMICAS - TRANSCRIPTOMICA

El transcriptoma representa un **pequeño porcentaje del código genético** (< 5% humanos; Frith et al., 2005)

Transcriptoma:

- Conecta el genoma con la **función del gen**
- **Identifica genes dianas** relacionados con patogenicidad
- Base de los programas de mejora genética (**breeding**)



MICROARRAYS

The Plant Journal (2003) 36, 177–188

doi: 10.1046/j.1365-313X.2003.01867.x

Transcriptome analysis of *Arabidopsis* colonized by a plant-growth promoting rhizobacterium reveals a general effect on disease resistance

Fabienne Cartieaux^{1,†,‡}, Marie-Christine Thibaud^{1,†}, Laurent Zimmerli², Philippe Lessard^{1,‡}, Catherine Sarrobert¹, Pascale David¹, Alain Gerbaud^{1,§}, Christophe Robaglia^{1,3}, Shauna Somerville² and Laurent Nussaume^{1,*}

The Plant Cell, Vol. 21: 1273–1290, April 2009, www.plantcell.org © 2009 American Society of Plant Biologists

Interaction Transcriptome Analysis Identifies *Magnaporthe oryzae* BAS1-4 as Biotrophy-Associated Secreted Proteins in Rice Blast Disease

Gloria Mosquera,^{a,1,2} Martha C. Giraldo,^{a,2} Chang Hyun Khang,^a Sean Coughlan,^{b,3} and Barbara Valent^{a,4}

EVOLUCIÓN DE LOS MÉTODOS TRANSCRIPTÓMICOS

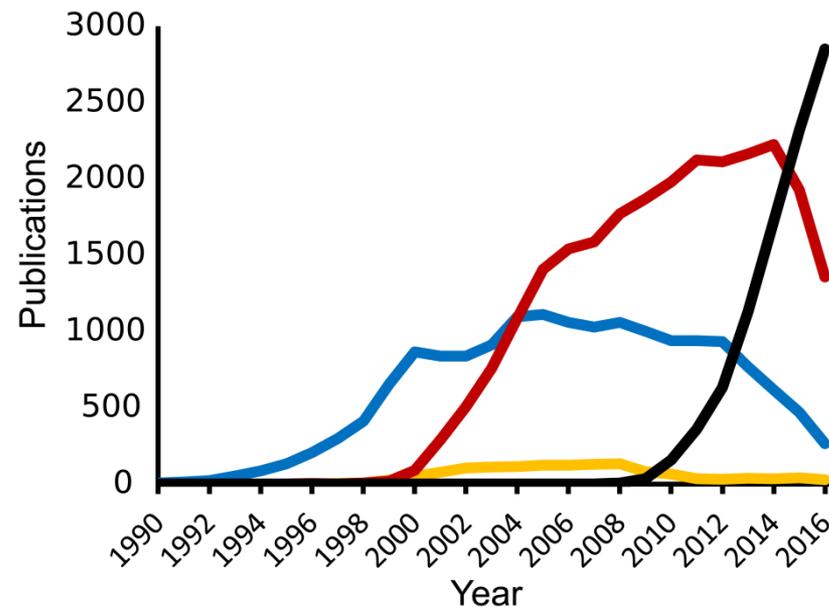


Fig 1. Transcriptomics method use over time. Published papers since 1990, referring to RNA sequencing (black), RNA microarray (red), expressed sequence tag (blue), and serial/cap analysis of gene expression (yellow)[12].

<https://doi.org/10.1371/journal.pcbi.1005457.g001>

Conventional RNA-seq

Análisis individualizado centrado en el hospedante o patógeno después de su separación física (*Westermann et al., 2012. Nature Reviews Microbiology 10: 618–630*)

Dual RNA-seq

Permite también revelar funciones de RNA no codificantes en la interacción hospedante/patógeno (*Westermann et al., 2016. Nature 529: 496–501*)

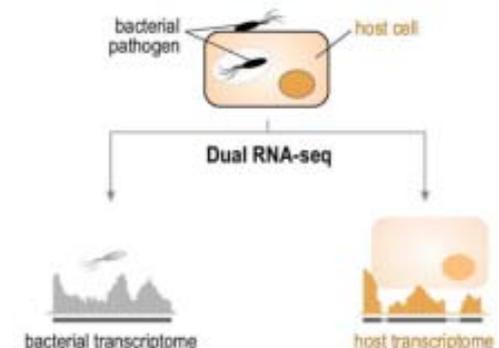


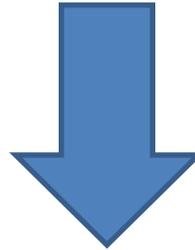
Fig 1. Methods for RNA sequencing of bacterial infections. A. Concept of dual RNA-seq.

Objetivos

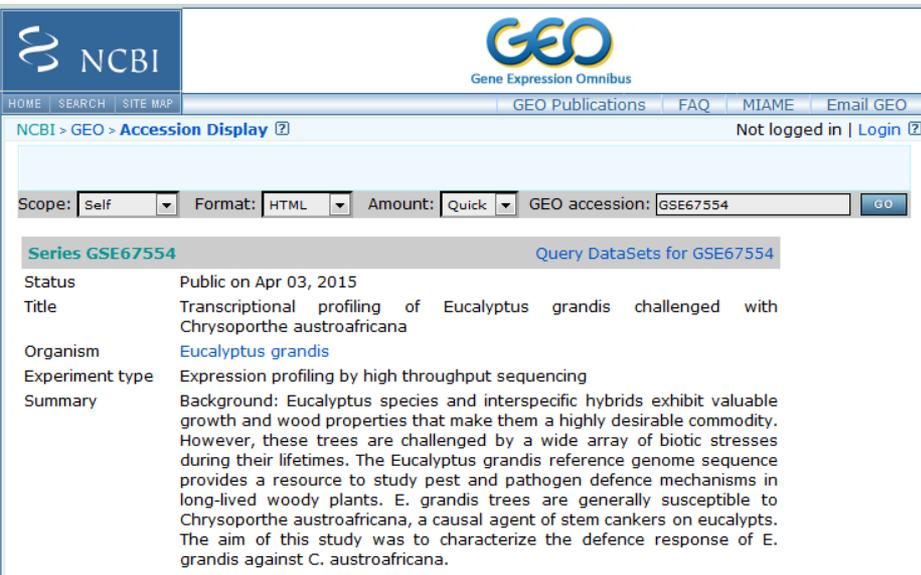
- Desarrollar el pipeline completo para el análisis dual hospedante/patógeno de datos de RNA-seq
- Obtener una serie de scripts que puedan ser implementados en el análisis de datos de RNA-seq dual en otros patosistemas

Material y métodos

COSTE DE SECUENCIACIÓN INABORDABLE



Gene Expression Omnibus (GEO)



NCBI logo and Gene Expression Omnibus logo are visible at the top. The page shows search filters: Scope: Self, Format: HTML, Amount: Quick, and GEO accession: GSE67554. The main content area displays the following information:

Status	Public on Apr 03, 2015
Title	Transcriptional profiling of <i>Eucalyptus grandis</i> challenged with <i>Chrysosporthe austroafricana</i>
Organism	Eucalyptus grandis
Experiment type	Expression profiling by high throughput sequencing
Summary	Background: <i>Eucalyptus</i> species and interspecific hybrids exhibit valuable growth and wood properties that make them a highly desirable commodity. However, these trees are challenged by a wide array of biotic stresses during their lifetimes. The <i>Eucalyptus grandis</i> reference genome sequence provides a resource to study pest and pathogen defence mechanisms in long-lived woody plants. <i>E. grandis</i> trees are generally susceptible to <i>Chrysosporthe austroafricana</i> , a causal agent of stem cankers on eucalypts. The aim of this study was to characterize the defence response of <i>E. grandis</i> against <i>C. austroafricana</i> .

Mangwanda et al. *BMC Genomics* (2015) 16:319
DOI 10.1186/s12864-015-1529-x



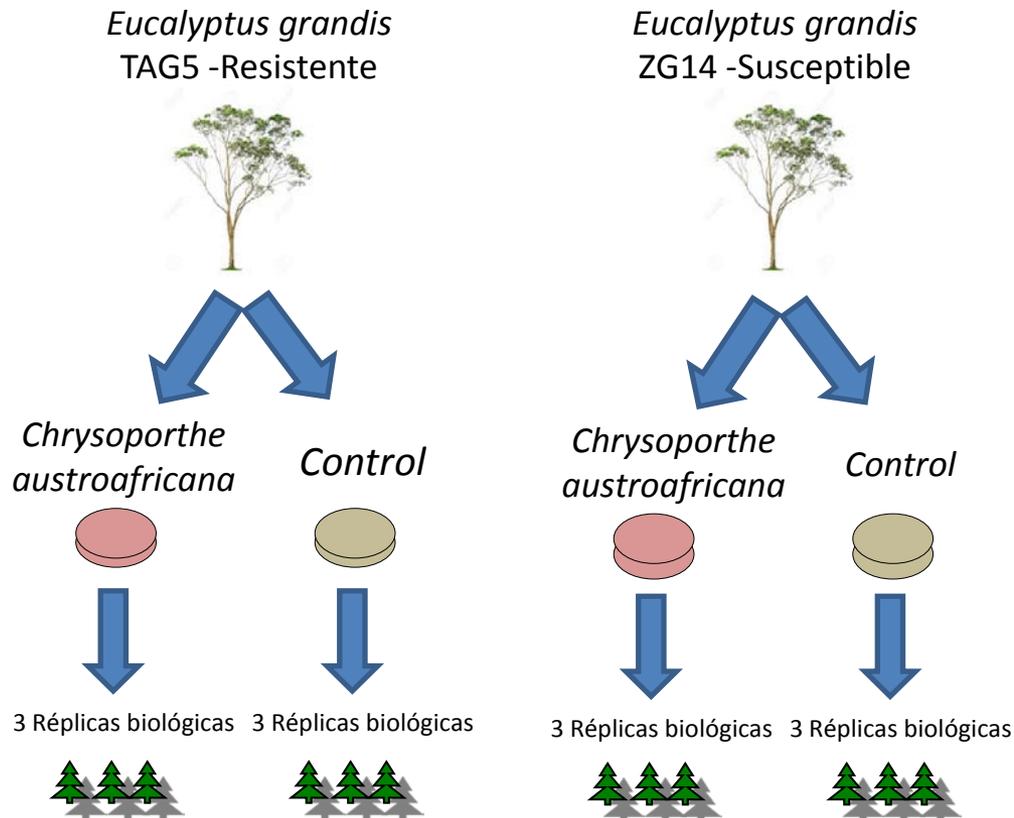
RESEARCH ARTICLE

Open Access

Transcriptome and hormone profiling reveals *Eucalyptus grandis* defence responses against *Chrysosporthe austroafricana*

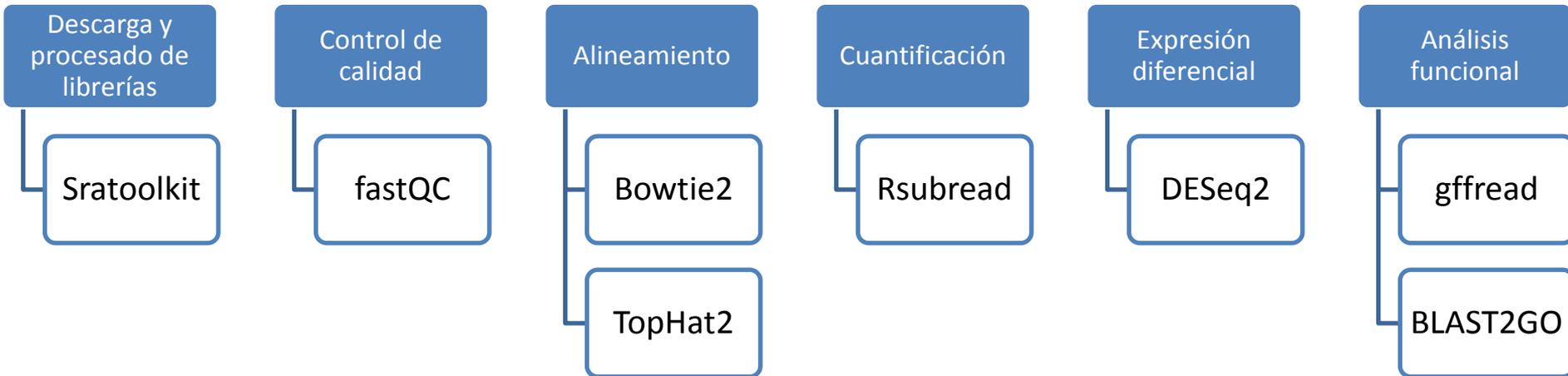
Ronishree Mangwanda, Alexander A Myburg and Sanushka Naidoo*

DISEÑO EXPERIMENTAL



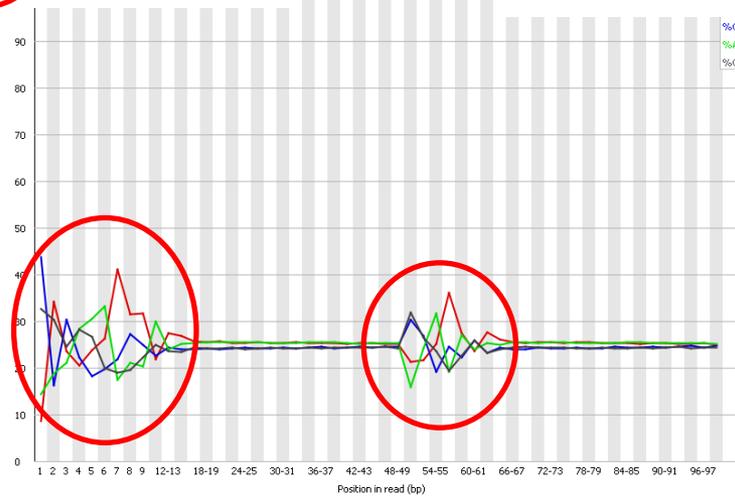
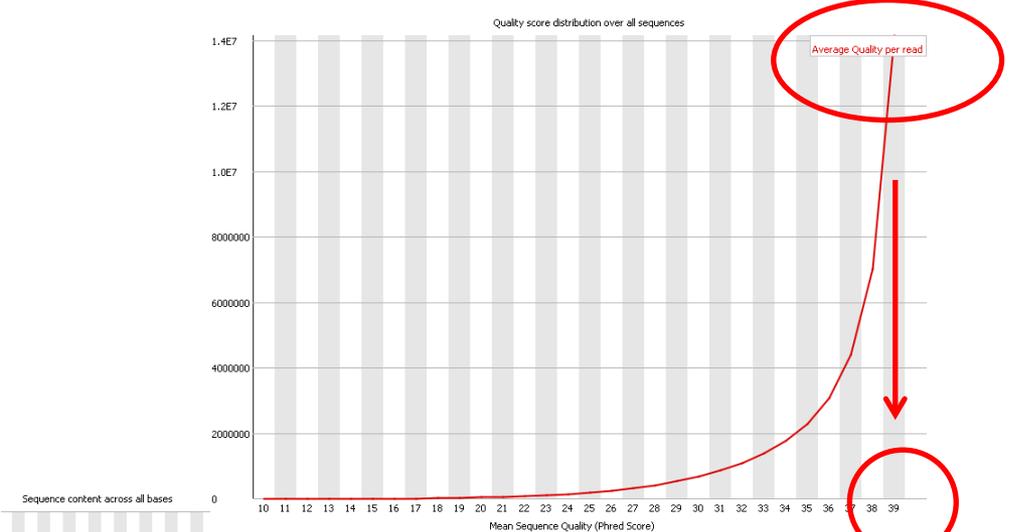
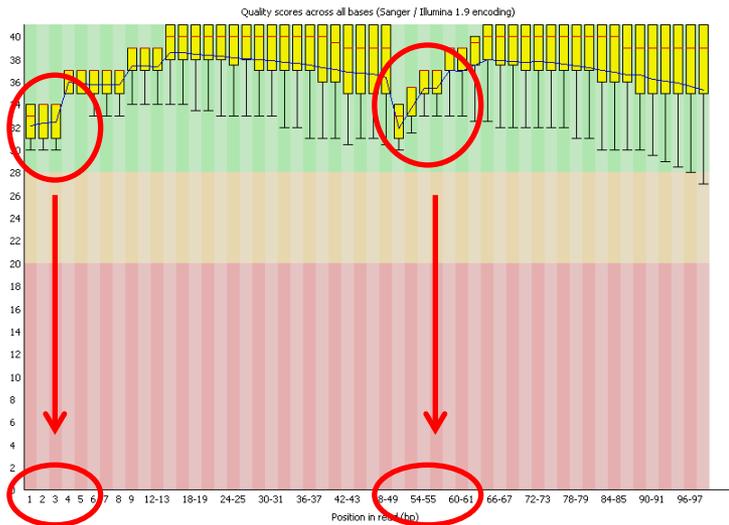
Material y métodos

PIPELINE



Resultados y discusión

CONTROL DE CALIDAD



Resultados y discusión

ALINEAMIENTO

Muestra	% Reads mapeadas frente al genoma del hospedante (<i>E. grandis</i>)	% Reads mapeadas frente al genoma del patógeno (<i>C. austroafricana</i>)
TAG5 Control BR1	88,1 %	0,1 %
TAG5 Control BR2	88,1 %	0,8 %
TAG5 Control BR3	88,1 %	0,1 %
TAG5 Infected BR1	88,7 %	1,0 %
TAG5 Infected BR2	87,7 %	1,5 %
TAG5 Infected BR3	87,2 %	0,9 %
ZG14 Control BR1	55,1 %	0,1 %
ZG14 Control BR2	86,8 %	0,1 %
ZG14 Control BR3	88,4 %	0,1 %
ZG14 Infected BR1	88,1 %	1,4 %
ZG14 Infected BR2	88,2 %	1,4 %
ZG14 Infected BR3	87,9 %	1,5 %

88.1%

0.3%

87.9%

1.0%

76.8%

0.1%

88.1%

1.4%

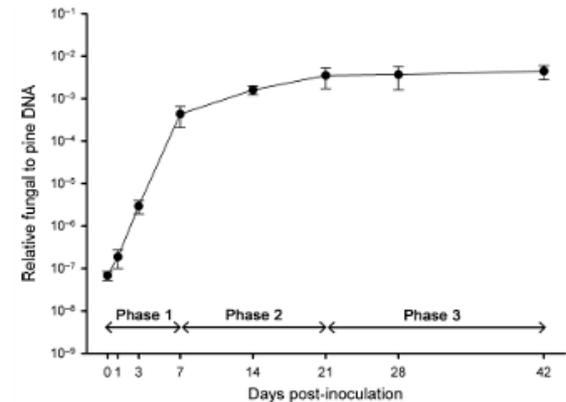
Resultados y discusión



Research

Spatial and temporal dynamics of the colonization of *Pinus radiata* by *Fusarium circinatum*, of conidiophora development in the pith and of traumatic resin duct formation

Noemí Martín-Rodríguez¹, Santiago Espinel², Joseba Sanchez-Zabala¹, Amaia Ortiz², Carmen González-Murua¹ and Miren K. Duñabeitia¹



Molecular Plant Pathology

MOLECULAR PLANT PATHOLOGY (2016) 17(2), 210–224

DOI: 10.1111/mpp.12273

Genome-wide gene expression dynamics of the fungal pathogen *Dothistroma septosporum* throughout its infection cycle of the gymnosperm host *Pinus radiata*

ROSIE E. BRADSHAW^{1,*}, YANAN GUO¹, ANDRE D. SIM¹, M. SHAHJAHAN KABIR¹, PRANAV CHETTRI¹, IBRAHIM K. OZTURK¹, LUKAS HUNZIKER¹, REBECCA J. GANLEY² AND MURRAY P. COX¹

	Total reads	Number of mapped fungal reads*	Fungal reads percentage of total reads†
Fungal mycelium <i>in vitro</i>	86 392 832	60 533 198	100
Early <i>in planta</i>	683 221 605	512 742	0.1
Mid <i>in planta</i>	1 033 815 851	871 668	0.5
Late <i>in planta</i>	17 146 358	1 883 653	17.1
Total <i>in planta</i>	1 734 183 814	3 268 063	

*Mapped to *D. septosporum* coding sequences.

†Normalized based on the percentage of mapped reads in the *in vitro* samples (see Experimental procedures).

TABLE 2 | Flagstat and FastQC RNA-seq mapping statistics of *Eucalyptus nitens* reads to the v1.1 *Eucalyptus grandis* and *Phytophthora cinnamomi* var *cinnamomi* genomes.

Sample name	Total reads mapped to <i>E. grandis</i>	Properly paired to <i>E. grandis</i> (%) ^a	Singletons mapped to <i>E. grandis</i> (%) ^b	% GC content	Expressed genes in host	Average FPKM in host	Total reads mapped to pathogen	% reads mapped to pathogen
Control 1	37444809	66.44	10.24	50	29024	493598	30756	0.08
Control 2	36111678	68.02	10.16	50	29250	492972	29274	0.08
Control 3	37060251	68.76	8.90	49	29135	467006	29116	0.08
Inoculated 1	37234371	66.13	11.65	49	29429	471923	444935	1.19
Inoculated 2	36622434	67.30	12.48	49	29407	497171	552202	1.51
Inoculated 3	36022978	68.19	10.18	49	29576	473466	196534	0.55

Dual RNA-Sequencing of *Eucalyptus nitens* during *Phytophthora cinnamomi* Challenge Reveals Pathogen and Host Factors Influencing Compatibility

Febè E. Meyer^{1†}, Louise S. Shuey^{1†}, Sitha Naidoo¹, Thandekile Mami¹, Dave K. Berger², Alexander A. Myburg¹, Noëlani van den Berg¹ and Sanushka Naidoo^{1*}

OPEN ACCESS

ORIGINAL RESEARCH
published: 02 March 2016
doi: 10.3389/fpls.2016.00191



Resultados y discusión

CUANTIFICACIÓN - HOSPEDANTE

Muestra	Genes identificados TFM	Genes identificados Mangwanda et al., 2015	% Genes identificados respecto a Mangwanda et al., 2015	% Genes identificados respecto a genoma referencia*
TAG5 Control BR1	26250	28448	92,3%	72,2%
TAG5 Control BR2	27690	28122	98,5%	76,1%
TAG5 Control BR3	26119	28239	92,5%	71,8%
TAG5 Infected BR1	26357	28487	92,5%	72,5%
TAG5 Infected BR2	26355	28688	91,9%	72,5%
TAG5 Infected BR3	25855	29829	86,7%	71,1%
ZG14 Control BR1	25318	27823	91,0%	69,6%
ZG14 Control BR2	25787	28043	92,0%	70,9%
ZG14 Control BR3	25788	28053	91,9%	70,9%
ZG14 Infected BR1	25389	27714	91,6%	69,8%
ZG14 Infected BR2	25718	27889	92,2%	70,7%
ZG14 Infected BR3	25579	27729	92.2%	70.3%

8 %

* Myburg et al. 2014. The genome of *Eucalyptus grandis*. Nature 510:356–62

CUANTIFICACIÓN – PATÓGENO

No se pudo llevar a cabo la cuantificación en el patógeno al no disponer de un archivo de anotación.

Se trata de dos organismos (hospedante/patógeno) muy dispares. Mientras el hospedante (640 Mb) tiene un genoma entorno a 10 veces más grande que el patógeno (~45 Mb), el número de genes es tan sólo entorno a 3 veces mayor (36376 vs 13484). Ante esta circunstancia una extrapolación a lo que sucedería con el patógeno en función de los resultados obtenidos en el hospedante no sería muy fiable.

Resultados y discusión

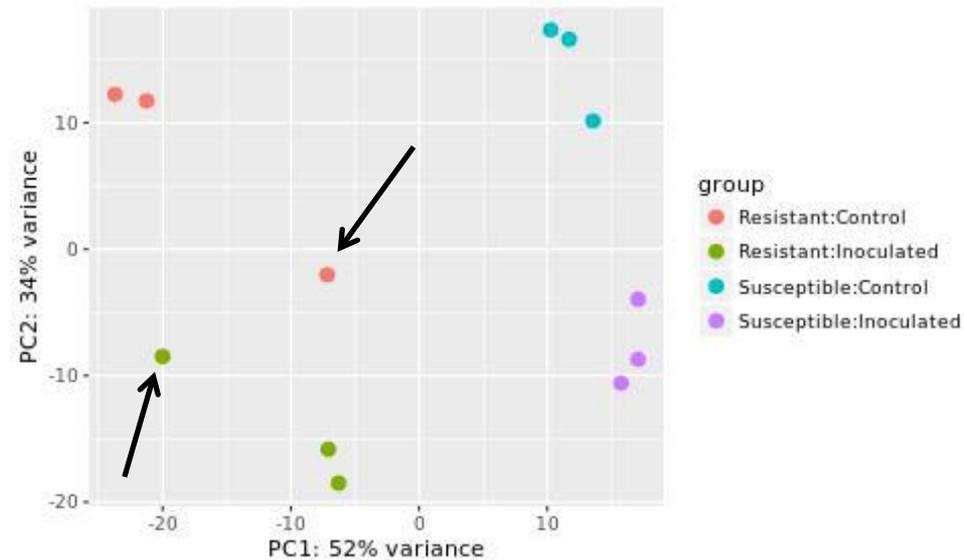
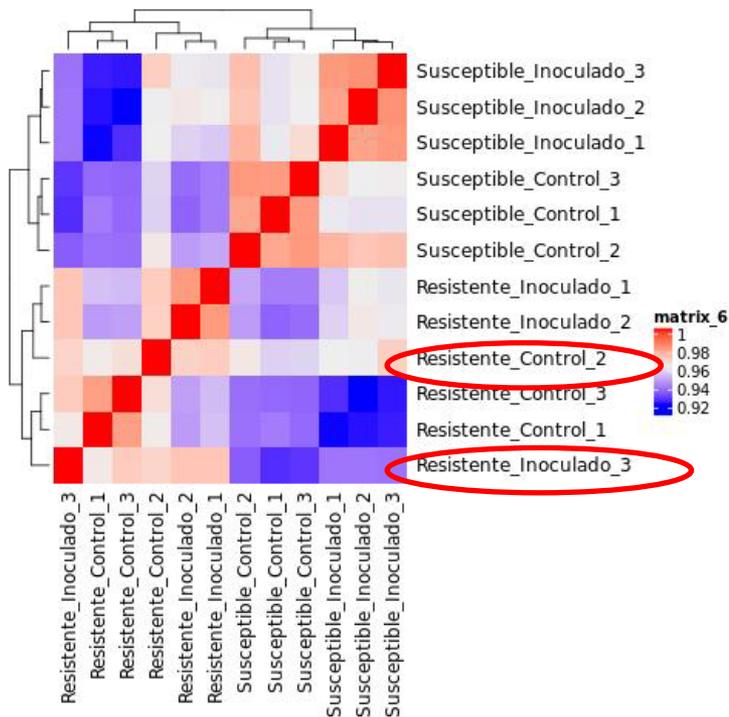
EXPRESIÓN DIFERENCIAL

Tratamiento	TFM		Mangwanda et al., 2015	
	Genes sobre-regulados	Genes infra-regulados	Genes sobre-regulados	Genes infra-regulados
Interacción Resistencia*Inoculación	92	128	-	-
TAG5 (resistente) Inoculado vs Control	1317	1112	856	682
ZG14 (susceptible) Inoculado vs Control	2020	1346	749	746
TAG5 Inoculado vs. ZG14 Inoculado	1181	1490	-	-

Resultados y discusión

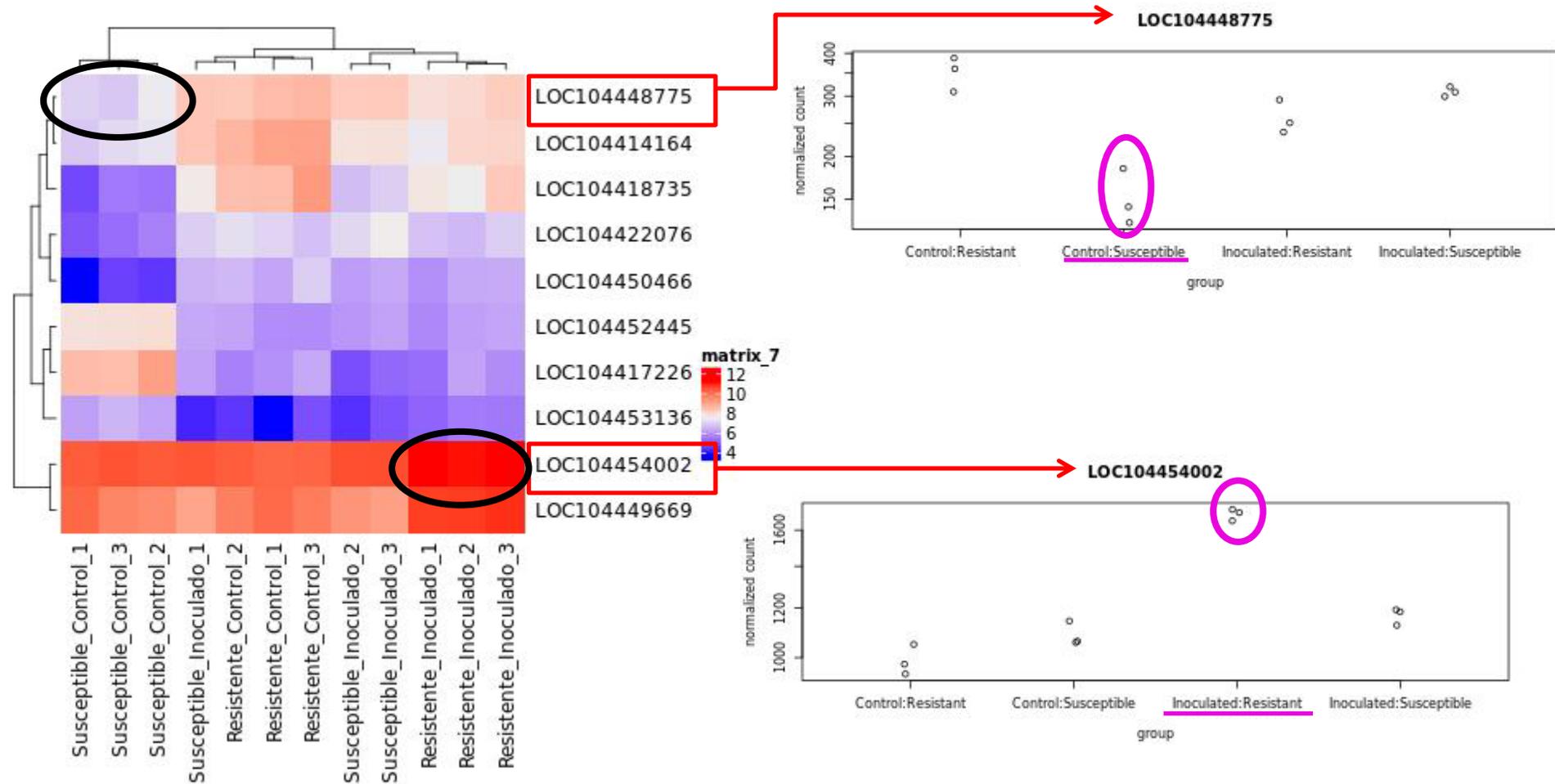
INTERACCIÓN RESISTENCIA:INOCULACIÓN

Correlación de Pearson de conteos ($\log_{10}(\text{conteos}(\text{dds})+1)$)



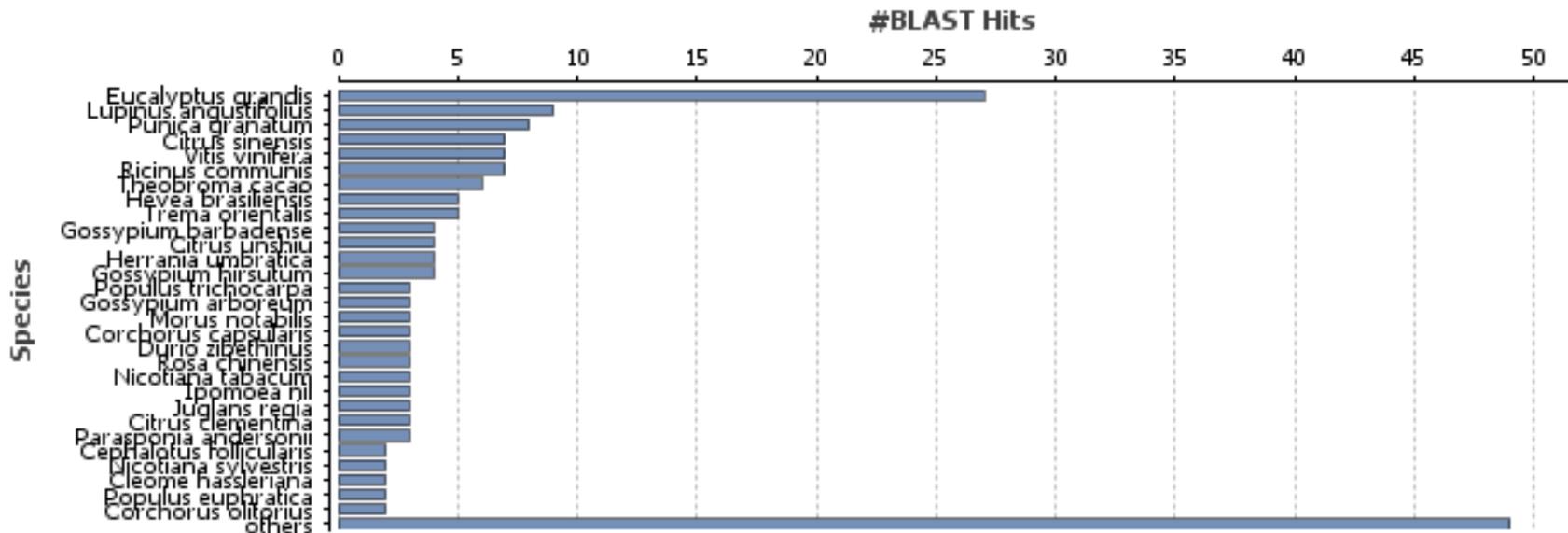
Resultados y discusión

TOP 10 DE GENES DIFERENCIALMENTE EXPRESADOS



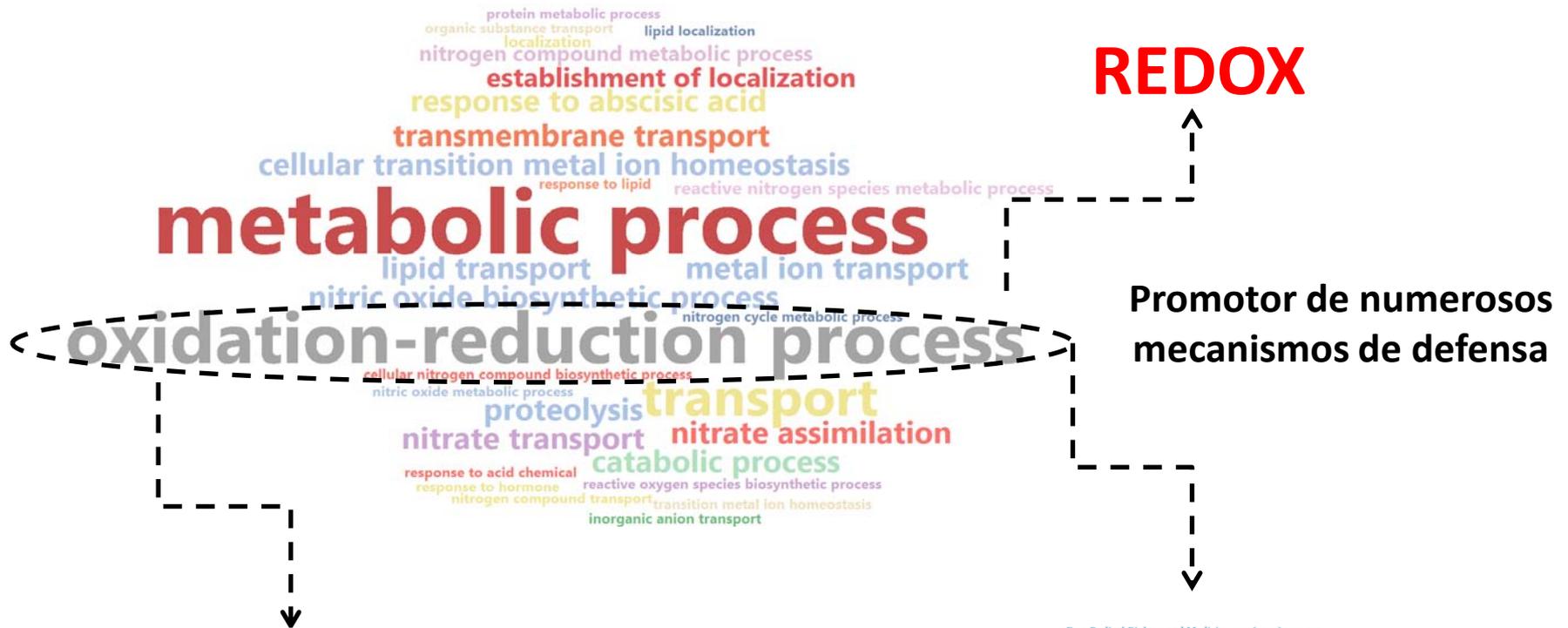
Resultados y discusión

ANÁLISIS DE ENRIQUECIMIENTO FUNCIONAL



Resultados y discusión

ANÁLISIS DE ENRIQUECIMIENTO FUNCIONAL- PROCESO BIOLÓGICO



ANTIOXIDANTS & REDOX SIGNALING
Volume 21, Number 9, 2014
© Mary Ann Liebert, Inc.
DOI: 10.1089/ars.2013.5679



FORUM REVIEW ARTICLE

Redox Regulation in Plant Immune Function

Debra E. Frederickson Matika and Gary J. Loake



Contents lists available at ScienceDirect

Free Radical Biology and Medicine

journal homepage: www.elsevier.com/locate/freeradbiomed

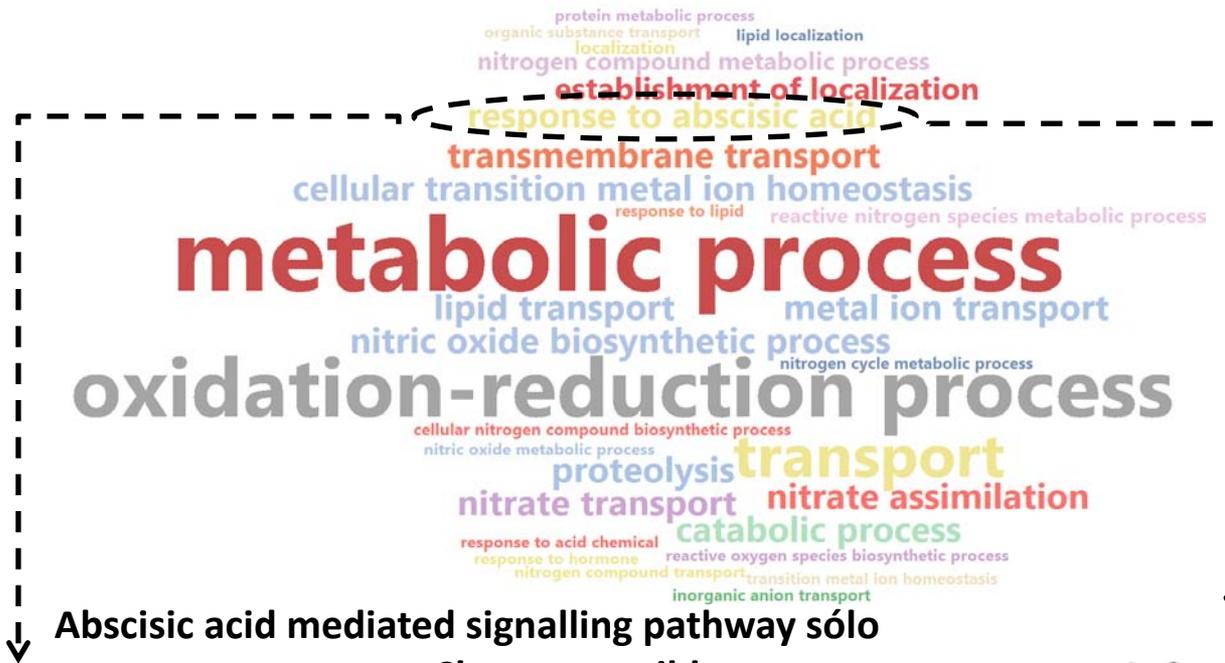


Priming plant resistance by activation of redox-sensitive genes

Carmen González-Bosch

Resultados y discusión

ANÁLISIS DE ENRIQUECIMIENTO FUNCIONAL- PROCESO BIOLÓGICO



Ácido abcísico relacionado con la inducción de defensas

Mangwanda et al. *BMC Genomics* (2015) 16:319
DOI 10.1186/s12864-015-1529-x



RESEARCH ARTICLE

Open Access

Transcriptome and hormone profiling reveals *Eucalyptus grandis* defence responses against *Chrysosporthe austroafricana*

Ronishree Mangwanda, Alexander A Myburg and Sanushka Naidoo*

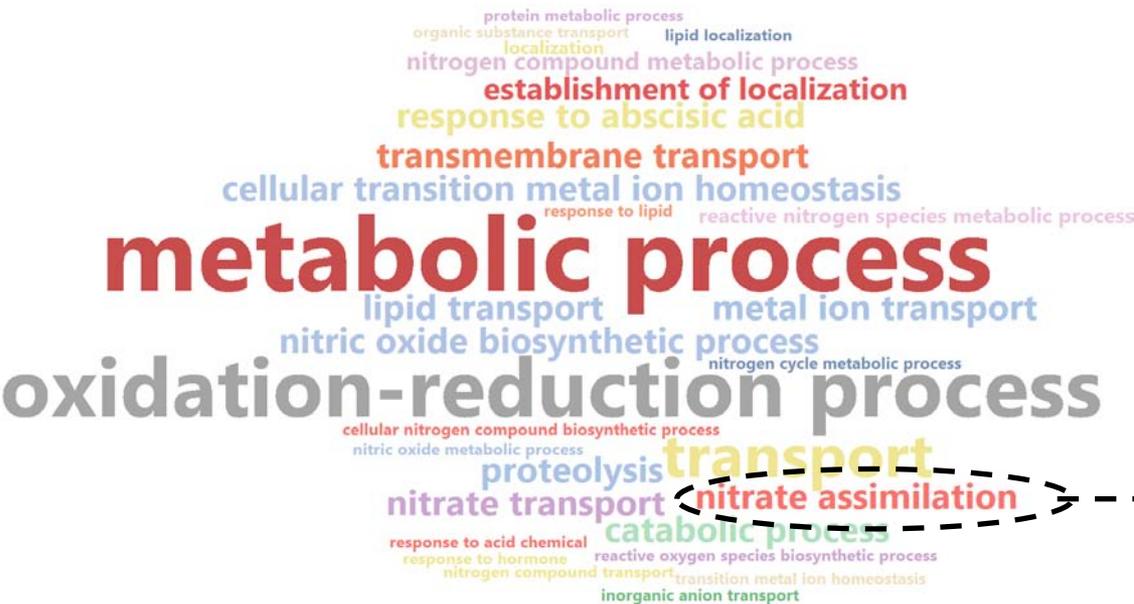
A Central Role of Salicylic Acid in Plant Disease Resistance

Terrence P. Delaney,* Scott Uknes,* Bernard Vernooij, Leslie Friedrich, Kris Weymann, David Negrotto, Thomas Gaffney, Manuela Gut-Rella, Helmut Kessmann, Eric Ward, John Ryals†

Science; Nov 18, 1994; 266, 5188; pg. 1247

Resultados y discusión

ANÁLISIS DE ENRIQUECIMIENTO FUNCIONAL- PROCESO BIOLÓGICO



Asimilación de nitrato asociada
a los procesos de infección
hospedante/patógeno



MINIREVIEW

Nutrition acquisition strategies during fungal infection of plants

Hege H. Divon^{1,2} & Robert Fluhr³

¹Department of Plant and Environmental Sciences, Norwegian University of Life Sciences, Ås, Norway; ²Norwegian Institute for Agricultural and Environmental Research - Bioforsk, Plant Health and Plant Protection Division, Ås, Norway; and ³Department of Plant Science, Weizmann Institute of Science, Rehovot, Israel

Resultados y discusión

ANÁLISIS DE ENRIQUECIMIENTO FUNCIONAL- FUNCIONES MOLECULARES

nitrite reductase activity
nitrate reductase activity
hydroxylase activity
oxidoreductase activity, acting on other nitrogenous compounds as donors
oxidoreductase activity, acting on other nitrogenous compounds as donors, iron-sulfur protein as acceptor
oxidoreductase activity, acting on other nitrogenous compounds as donors, with NAD or NADP as acceptor

heterocyclic compound binding

organic cyclic compound binding

cofactor binding

nitrate reductase (NADPH) activity RNA binding

transition metal ion binding

catalytic activity

prenylcysteine methylesterase activity coenzyme binding

metal ion binding heme binding

molybdenum ion binding iron-sulfur cluster binding binding

tetrapyrrole binding ferredoxin-nitrite reductase activity

molybdopterin cofactor binding

cation binding catalytic activity, acting on a protein

peptidase activity

metal cluster binding

Un incremento de la actividad catalítica es un indicador de la defensa química de la planta

frontiers in
PLANT SCIENCE

ORIGINAL RESEARCH ARTICLE
published: 20 February 2015
doi: 10.3389/fpls.2015.00079

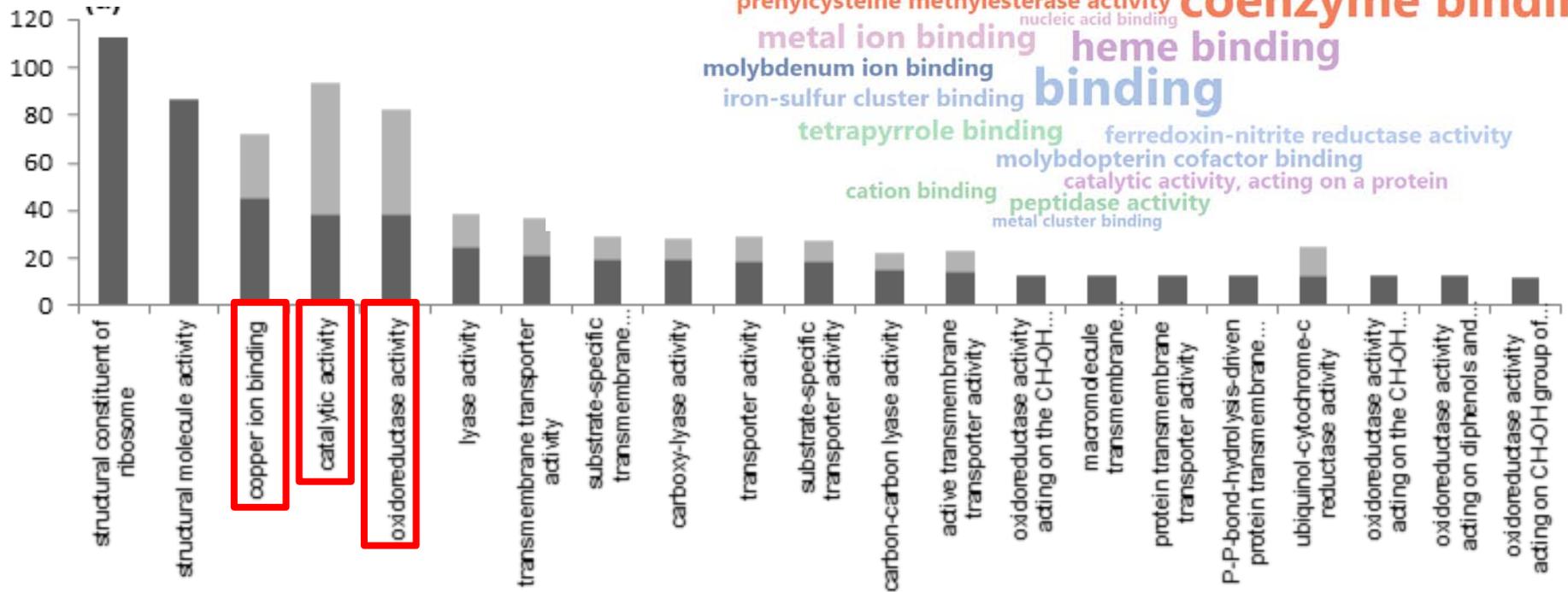


Arabidopsis flower specific defense gene expression patterns affect resistance to pathogens

Luisa Ederli^{1†}, Adam Dawe^{2†}, Stefania Pasqualini¹, Mara Quaglia³, Liming Xiong² and Chris Gehring^{2*}

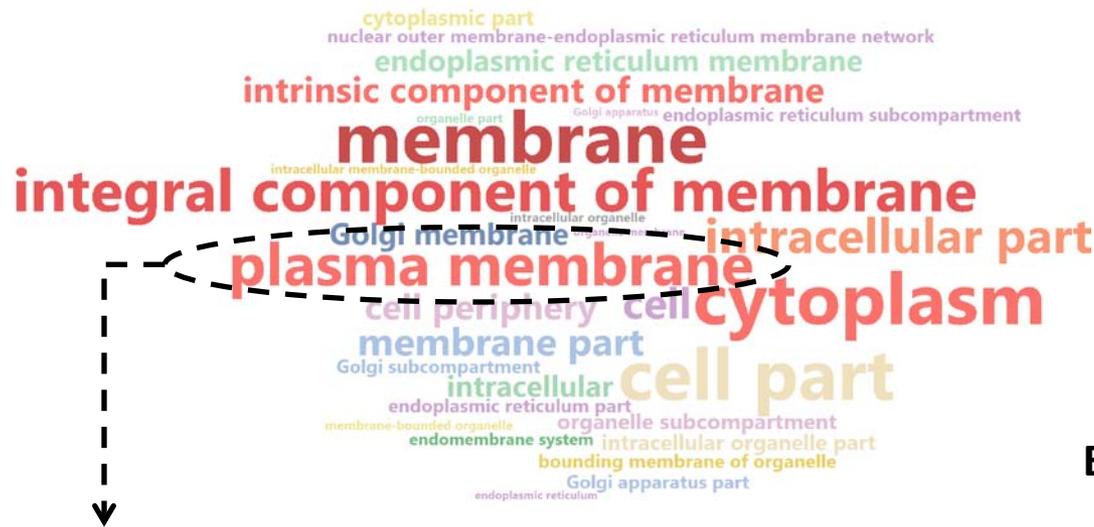
Resultados y discusión

Mangwanda et al., 2015



Resultados y discusión

ANÁLISIS DE ENRIQUECIMIENTO FUNCIONAL- COMPOSICIÓN CELULAR



Es conocido que numerosos genes de resistencia están relacionados con “plasma membrane protein”

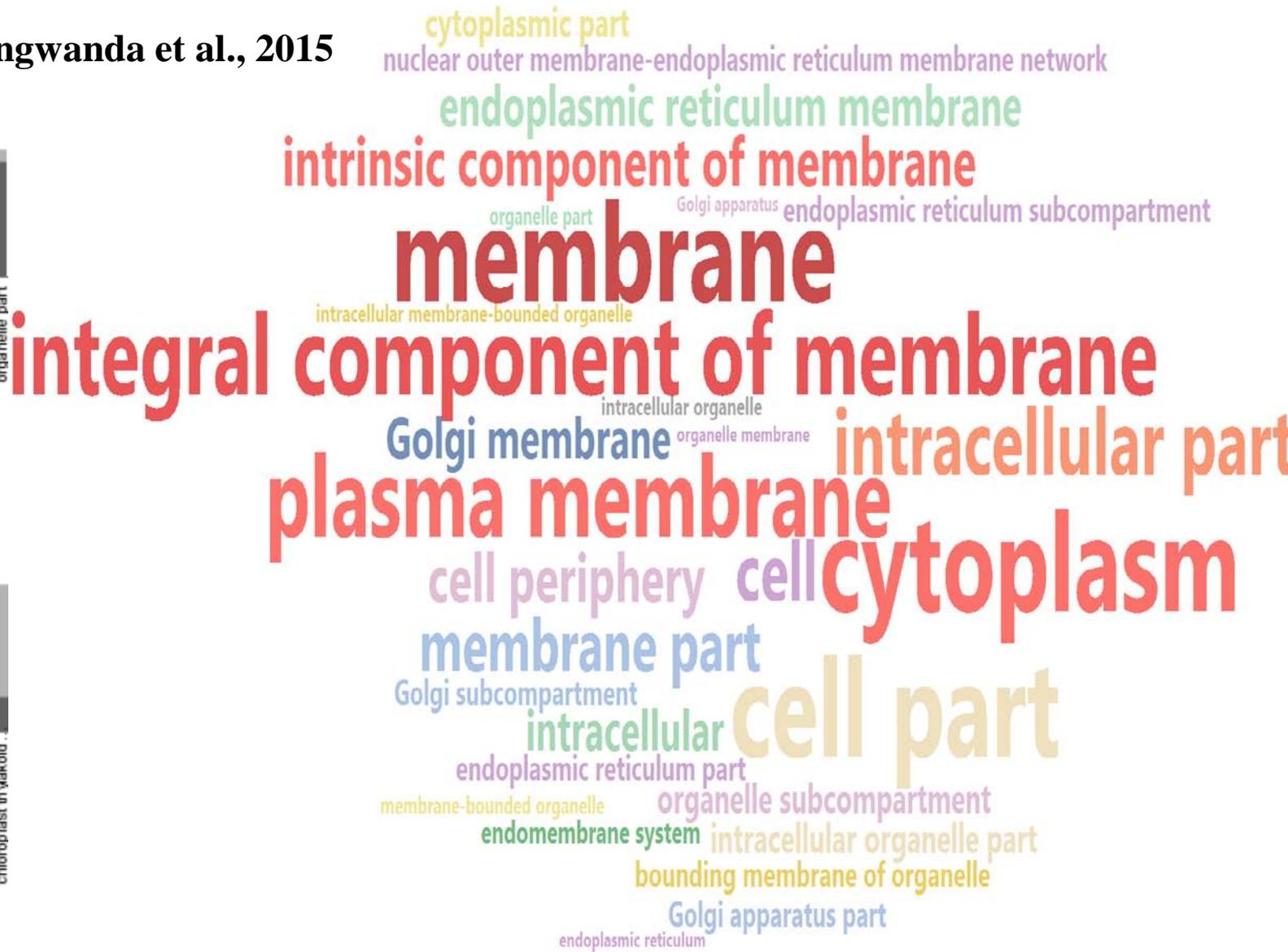
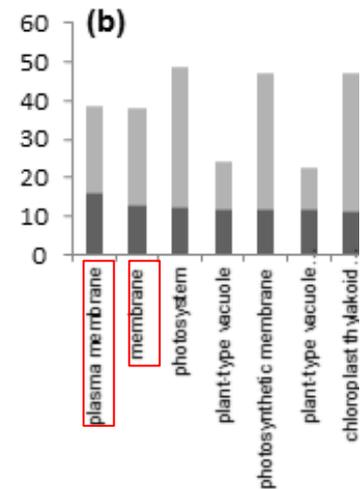
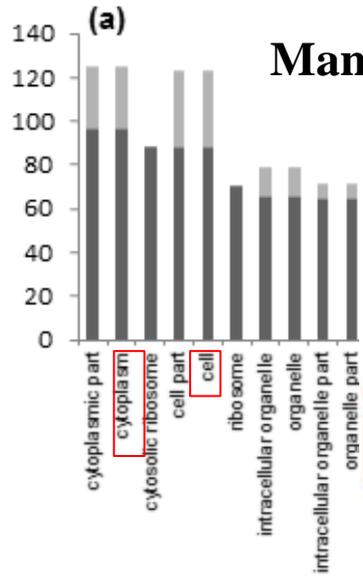
Proc. Natl. Acad. Sci. USA
Vol. 95, pp. 15849-15854, December 1998
Plant Biology

The *Arabidopsis thaliana* RPM1 disease resistance gene product is a peripheral plasma membrane protein that is degraded coincident with the hypersensitive response

DOUGLAS C. BOYES^{*2}, JAESEUNG NAM^{*}, AND JEFFERY L. DANGL^{*3}§

Resultados y discusión

Mangwanda et al., 2015



Conclusiones

El RNA-seq dual ha permitido capturar RNA del hospedante (*E. grandis*) y del patógeno (*C. austroafricana*) en una proporción del 85-90% y 1-1,5%, respectivamente

En el presente patosistema una profundidad de secuenciación de 4 millones fue suficiente para identificar aproximadamente el 70% de genes del hospedante

El porcentaje de reads mapeadas frente al genoma del patógeno es muy bajo por efecto del ratio biomasa hospedante/patógeno en la material vegetal

El análisis funcional de los 10 genes más diferencialmente expresados permitió identificar procesos biológicos, funciones moleculares y componentes celulares previamente vinculados con mecanismos de defensa de las plantas

Entre los procesos biológicos alterados por efecto de la resistencia del hospedante y la presencia del patógeno se encuentra el proceso redox implicado en la producción ROS, la cual es conocida por ser una de las primeras respuestas celulares de la planta tras el reconocimiento del patógeno

Mientras que la actividad catalítica, indicador de la defensa química de la planta y la membrana plasmática, la cual está relacionada con genes de resistencia, fueron identificadas como la principal función molecular y componente celular alterada en el presente estudio

Conclusiones



Este TFM ha establecido un pipeline y desarrollado un script para el estudio dual hospedante/patógeno mediante RNA-seq

El pipeline y scripts desarrollados pueden ser aplicados a cualquier otro patosistema vegetal

La capacidad computacional requerida para un estudio de RNA-seq con muestras cuya profundidad es de 40 millones de reads fue superior a la disponible actualmente en mi institución. Si bien, una submuestra de 4 millones de reads fue suficiente para el desarrollo del pipeline

El presente TFM me ha permitido conocer los repositorios de libre acceso disponibles y familiarizarme con distintos software que han sido usados para el desarrollo y optimización de un pipeline completo para el RNA-seq dual de un patosistema modelo