

Identification of metabolic pathways and
candidate genes associated with the process
of foliar senescence in cultivated sunflower
(*Helianthus annuus L.*) by integrating
transcriptomic, metabolomic and phenomics
data

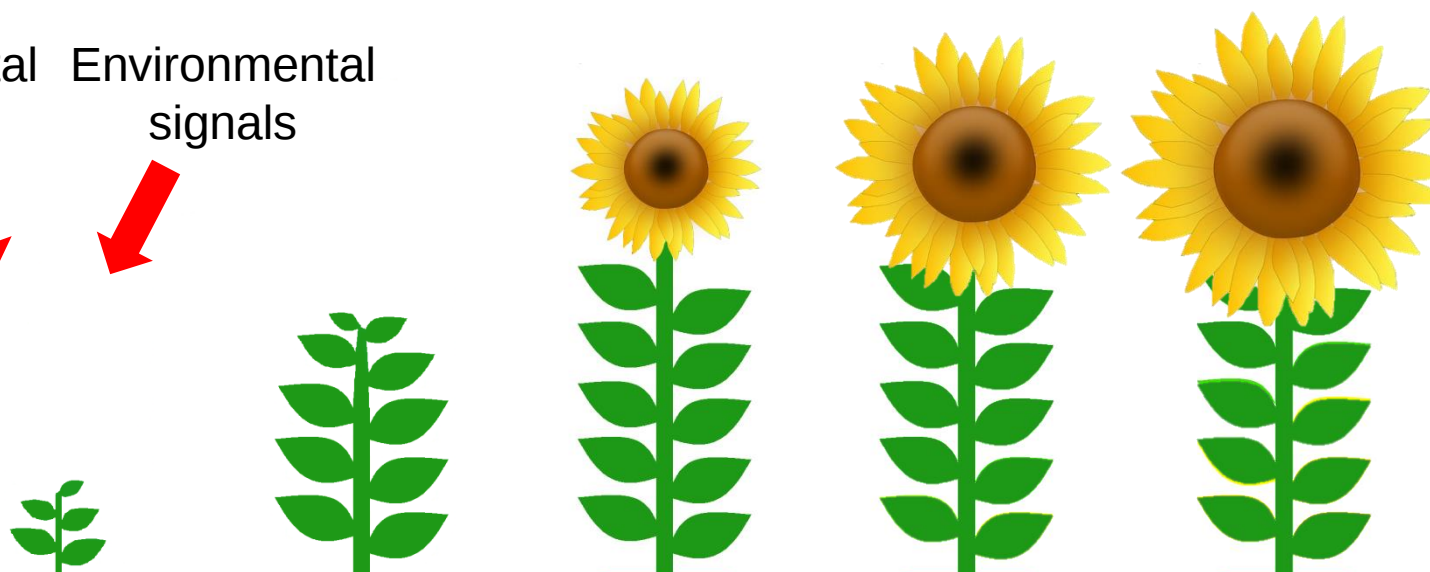
Paula Fernandez
Instituto de Biotecnología/IABIMO
INTA – CNIA
fernandez.pc@inta.gob.ar

Leaf senescence process in sunflower. Why?

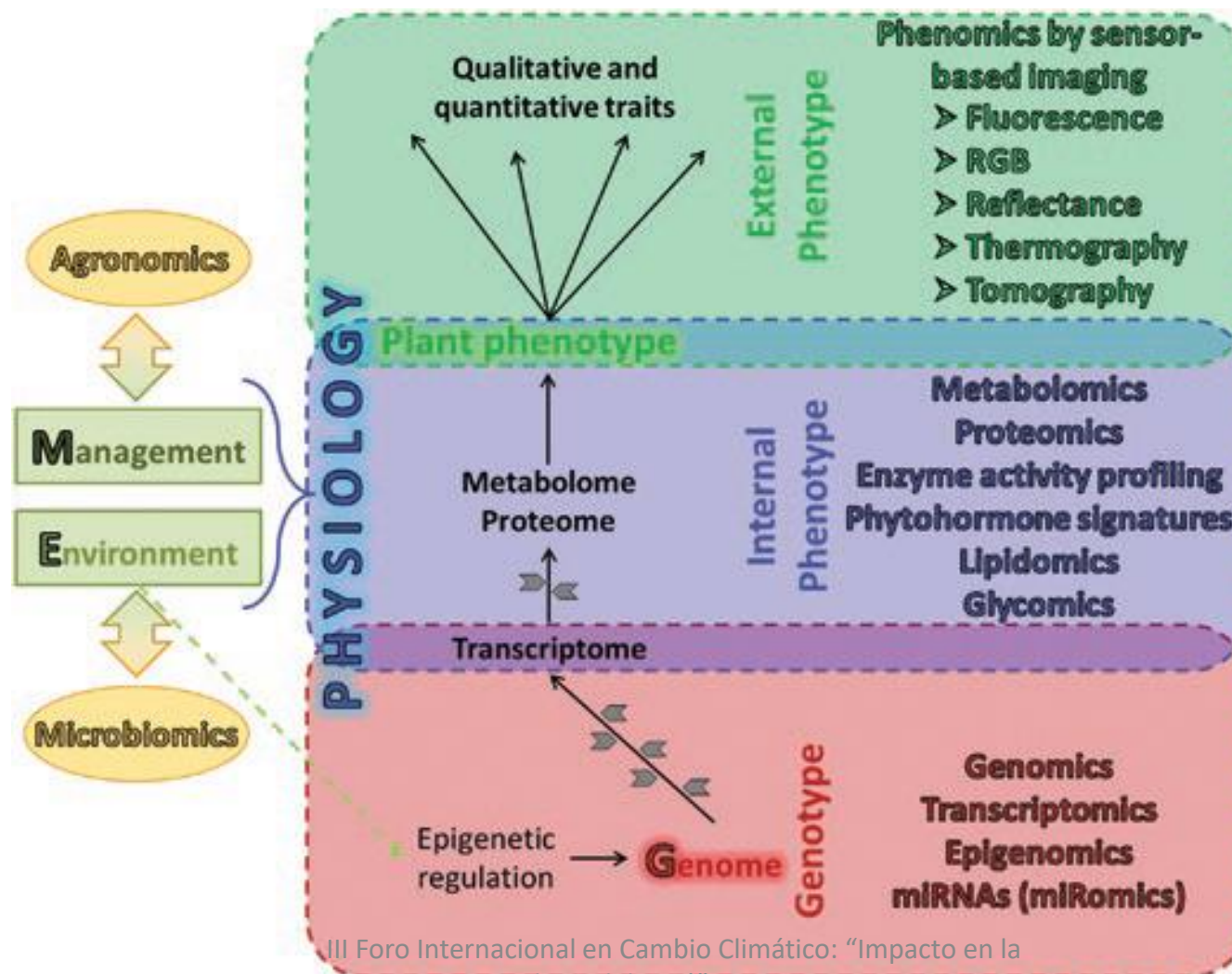
- Real yields $\neq$ Potential yields. (INTA “Brechas” Yield Program)
- In annual crops, and after anthesis, senescence process is induced and the nutrients are relocated into young leaves and grain.
- The intercepted radiation during the grain filling phase plays an important role in crop yield and oil content (*Dosio et al., 2000; Aguirrezábal et al., 2003*). Climate change and levels of radiation

Developmental
signals

Environmental
signals

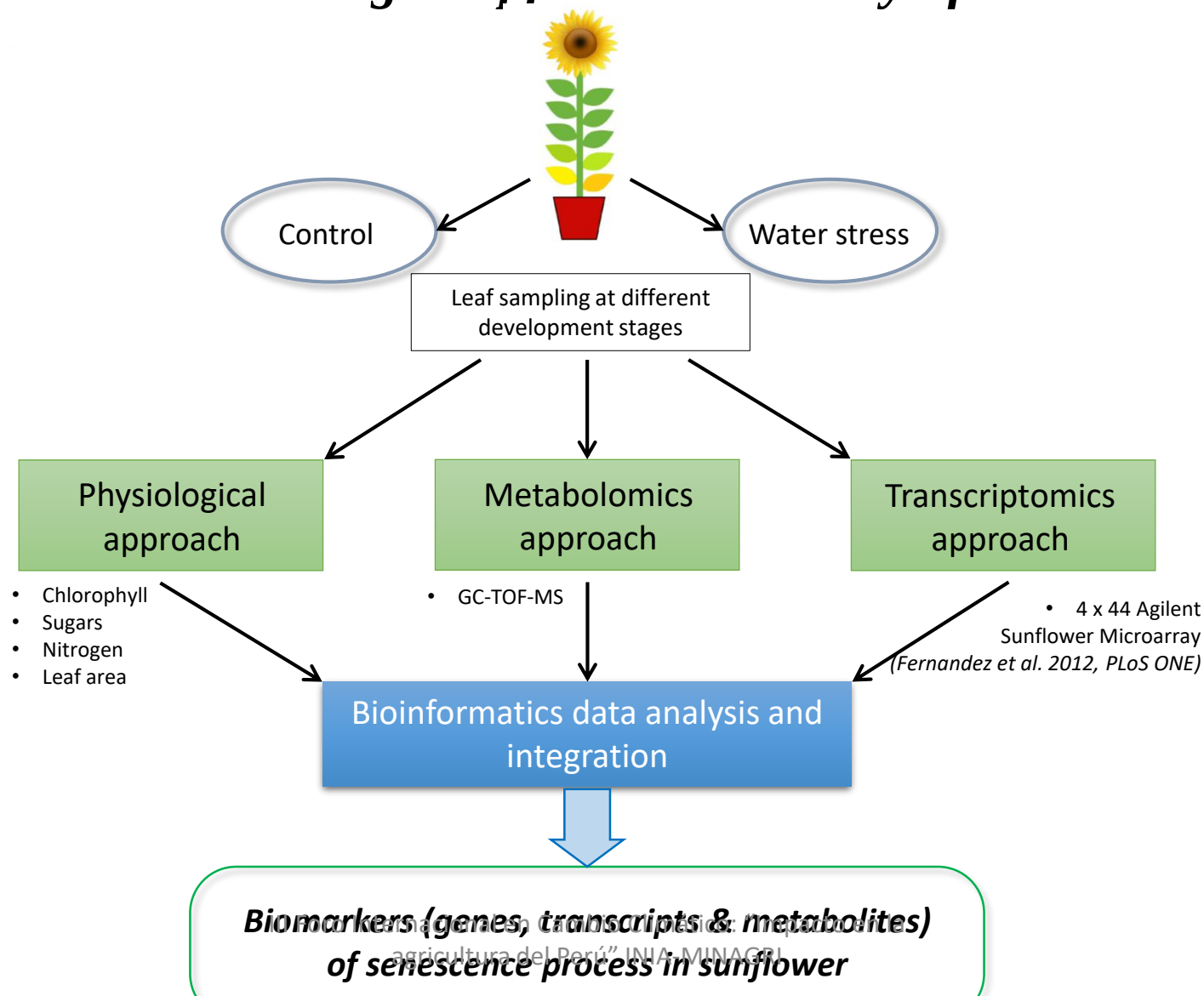


Integration of multi-omics techniques and physiological phenotyping within a holistic phenomis approach



Großkinsky et al. (2015b)

Functional Genomics Strategies applied to the study of senescence in sunflower



Commercial Sunflower Hybrid
Putative "stay green"
Advanta Seeds

Moschen et al, 2014

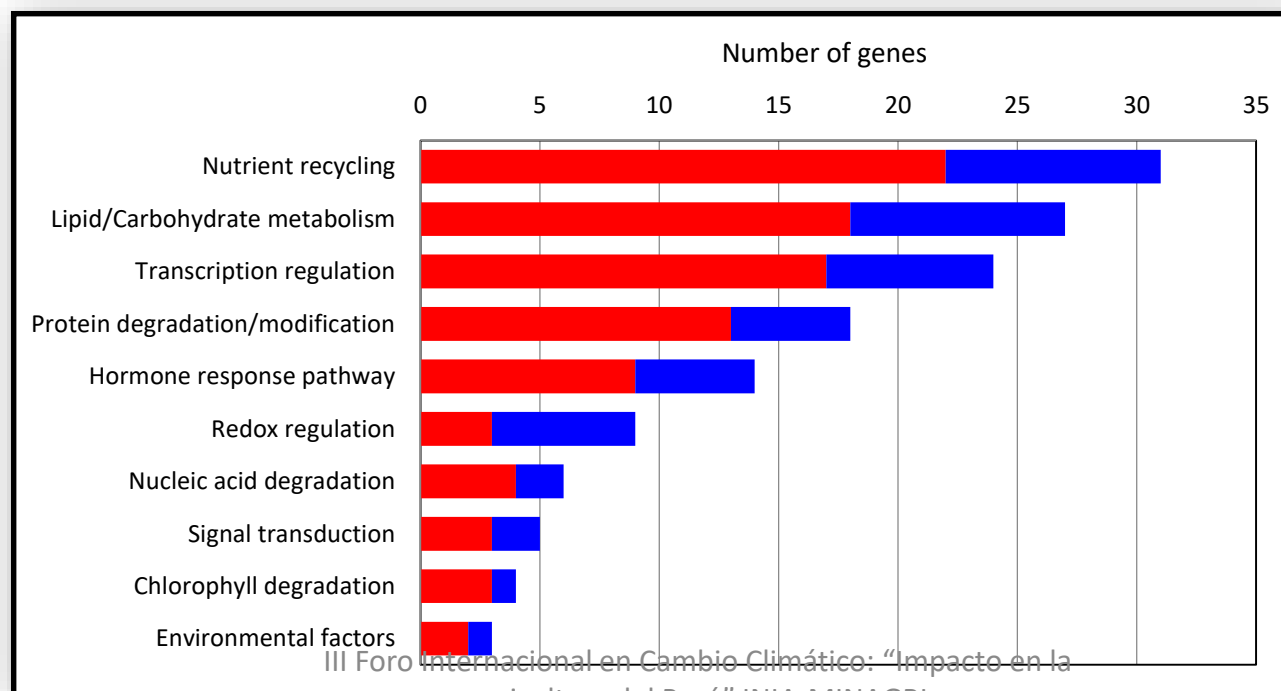
Transcriptomic profiles

Custom sunflower microarray (Agilent platform) $\approx$ 44,000 probes

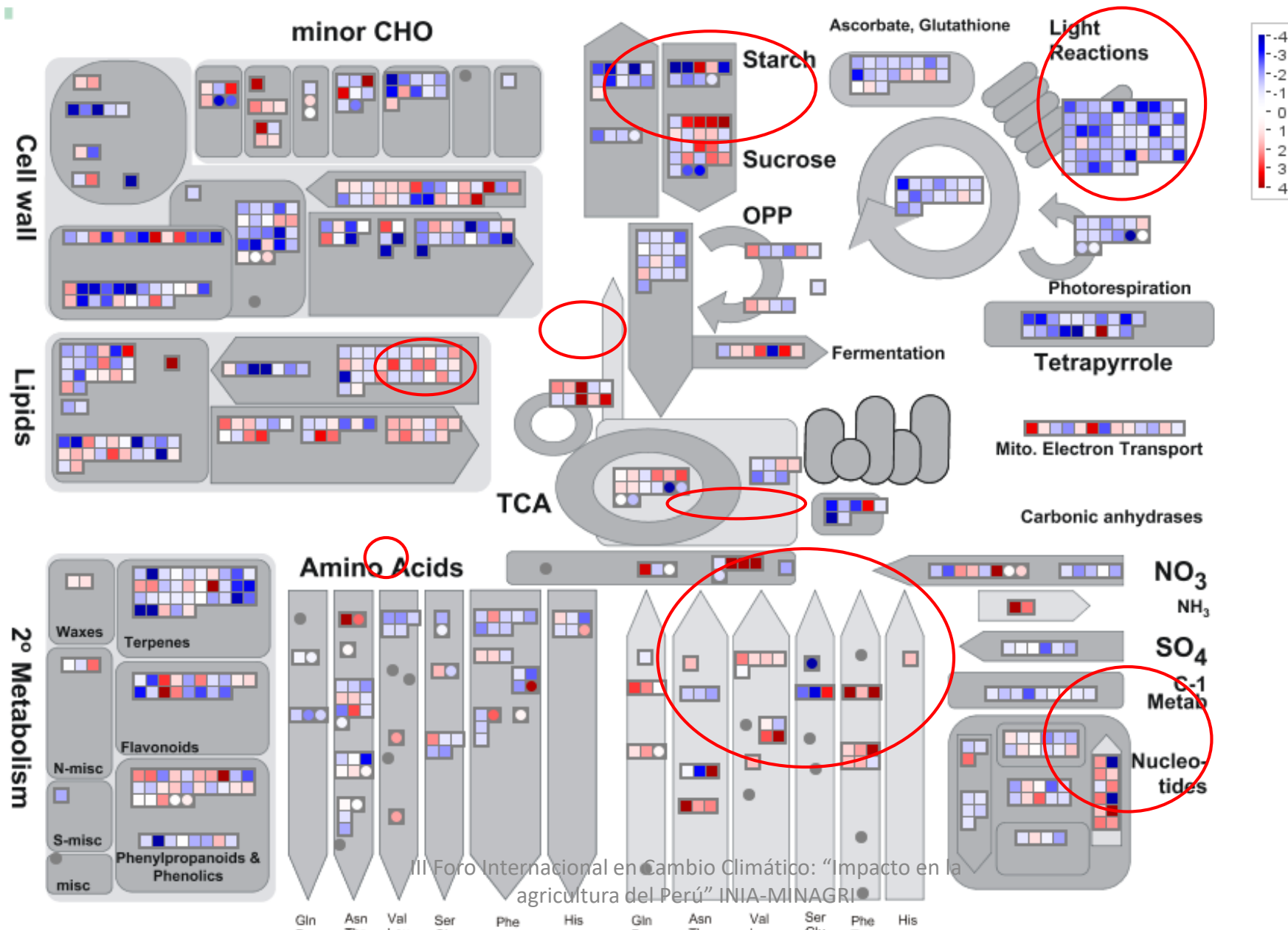
Sunflower Unigene Database (SUR) <http://atgc-sur.inta.gob.ar/> (Fernandez et al. 2012, PLoS ONE)

4,910 unigenes up or down regulated

168 significant Blast matches to Senescence Associated Genes (SAGs) reported in *Arabidopsis thaliana* using the “Leaf Senescence Database” (Liu, X. et al. 2011)
(<http://psd.cbi.pku.edu.cn/>)



Transcriptomics and metabolomics profiles



Moschen et al. 2016

Transcription Factors Analysis

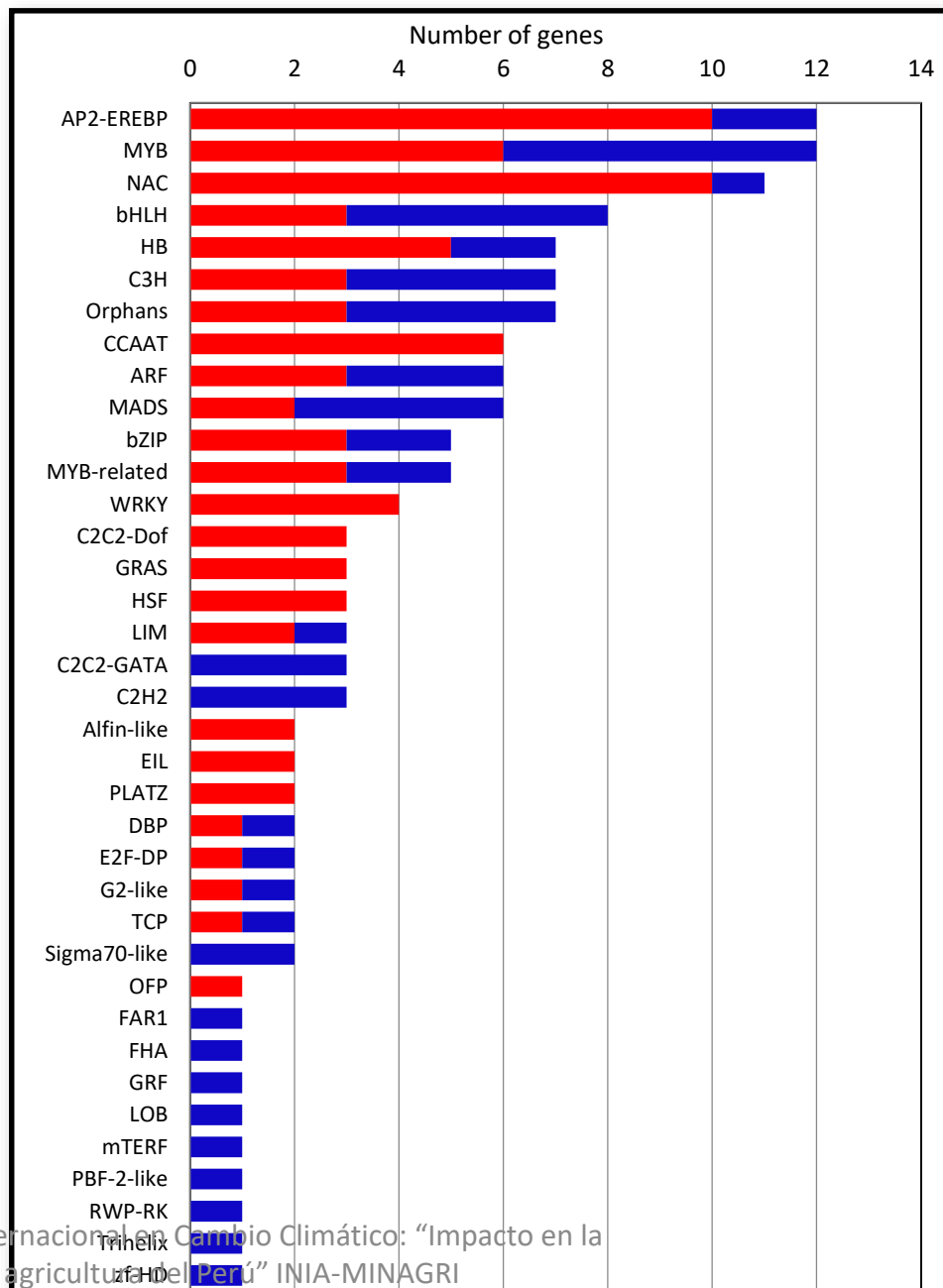
- **Plant transcription factor database**
(Perez-Rodriguez P. et al 2009)
(<http://plntfdb.bio.uni-potsdam.de/v3.0/>)

≈23000 TF sequences:

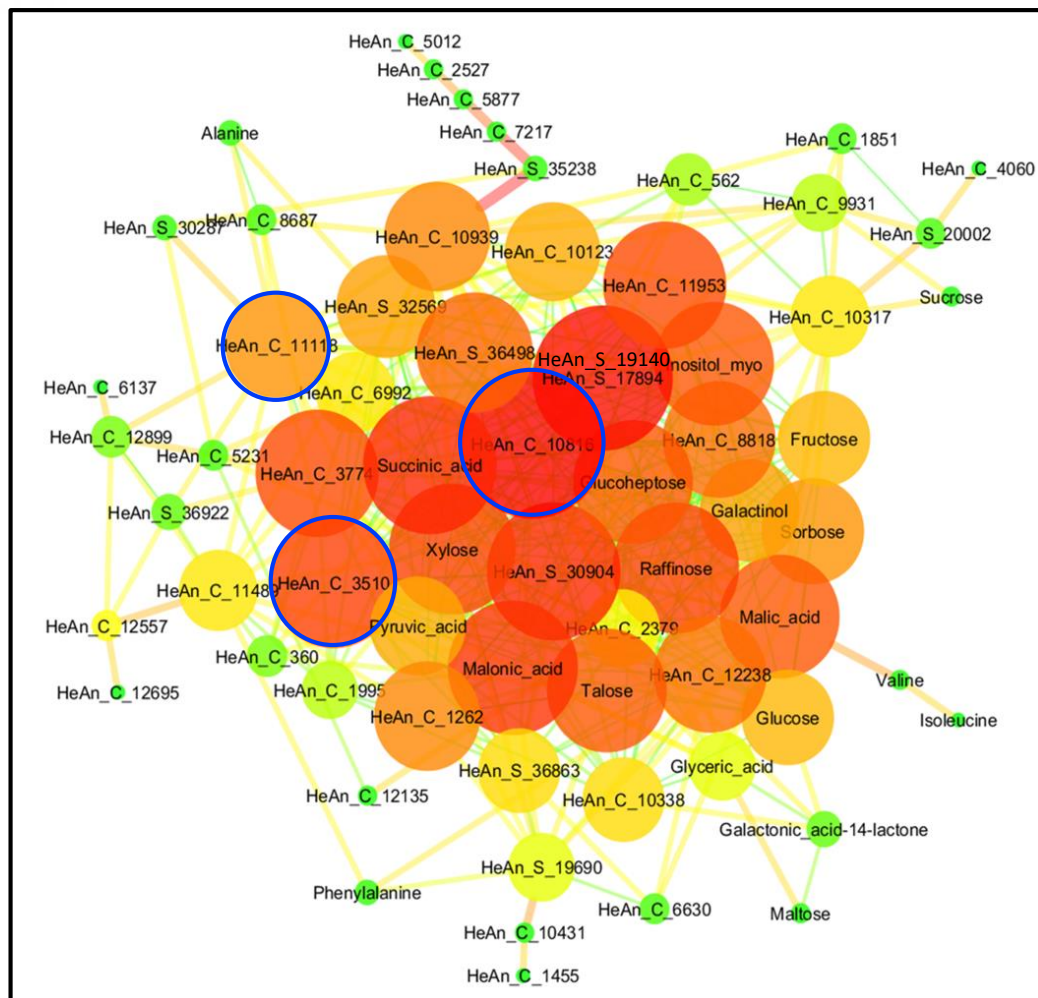
- *Arabidopsis lyrata*
- *Arabidopsis thaliana*
- *Oriza sativa*
- *Populus trichocarpa*
- *Vitis vinifera*
- *Zea mays*

**680 significant Blast matches
unigenes identified in sunflower
database**

140 TFs up or down regulated

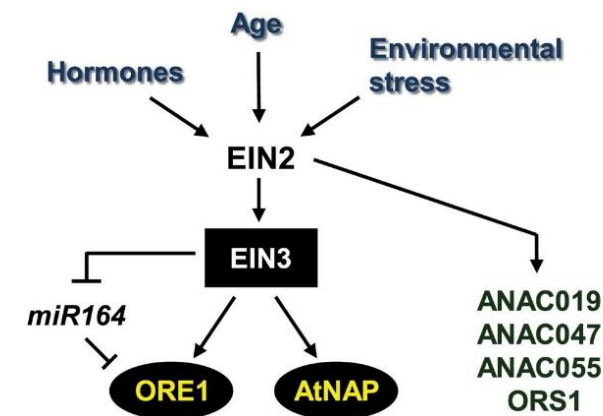


Transcription Factors and metabolites integration (WGCNA)



Moschen et al. 2016

HeAn	<i>H. annuus</i>	T1	T2	T3	Arabidopsis	
HeAn_S_19140	<i>HaNAC01</i>	0	0,48	0,86	AT5G39610	<i>(ANAC092/ORE1)</i>
HeAn_C_3510	<i>HaNAC03</i>	0	5,48	6,09	AT4G27410	<i>(ANAC072)</i>
HeAn_C_10816	<i>Ha-NAC05</i>	0	5,13	5,50	AT1G52890	<i>(ANAC019)</i>
HeAn_S_40317	<i>Ha-NAC06</i>	0	4,33	4,92	AT3G15500	<i>(ANAC055)</i>
HeAn_C_11118	<i>Ha-NAC07</i>	0	3,91	5,23	AT1G69490	<i>(NAC029/AtNAP)</i>
HeAn_C_8526	<i>Ha-NAC08</i>	0	2,28	2,62	AT1G01720	<i>(ANAC002/ATAF1)</i>
HeAn_S_17088	<i>Ha-NAC09</i>	0	1,23	3,30	AT3G15510	
HeAn_S_22690	<i>Ha-NAC11</i>	0	1,15	1,38	AT4G35580	
HeAn_S_17779	<i>Ha-NAC12</i>	0	1,30	1,52	AT4G28530	



Kim et al. 2014, J Exp Bot

INTA Breeding Sunflower Population

Selection of senescence contrasting sunflower genotypes

135 sunflower genotypes were phenotyped:

- Growing cycle.
- Anthesis time.
- Number of leaves.
- Plant size
- Evolution of total/dry leaf.
- Green leaf at anthesis time
- SPAD

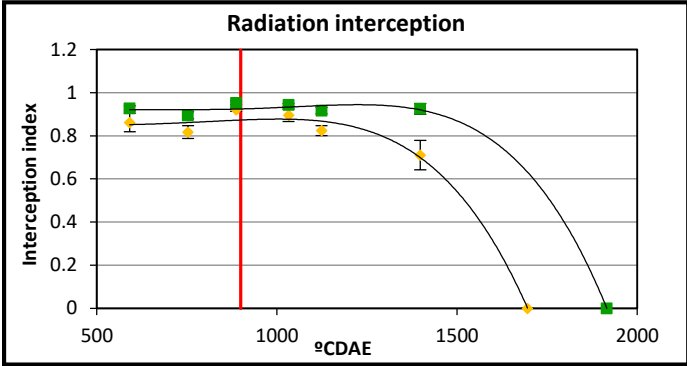
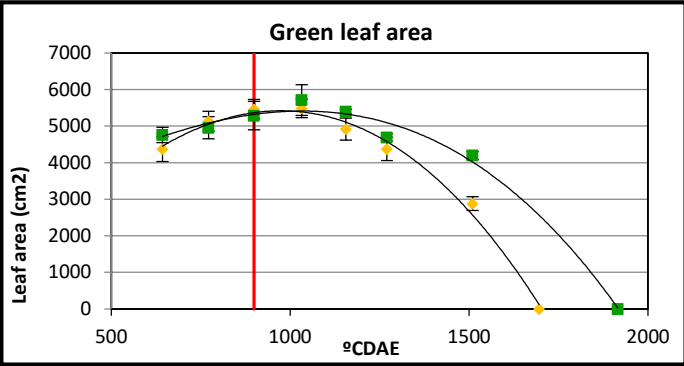
Pre-selection:

10 genotypes (5 contrasts)

Early senescence genotype	Delayed senescence genotype
c977-b	C154B
R453	B481-6
2091	2021
B473-2	C829B
C818	B10



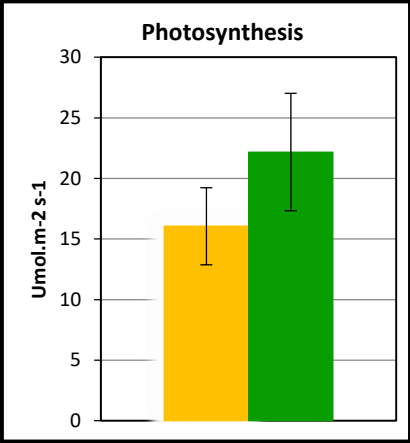
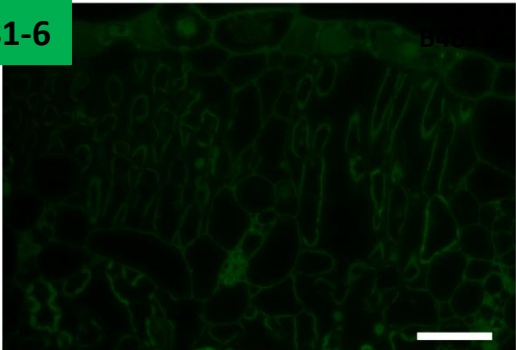
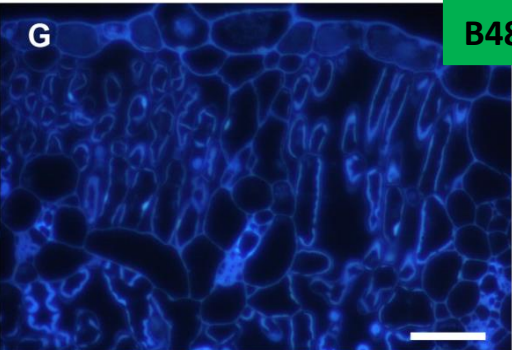
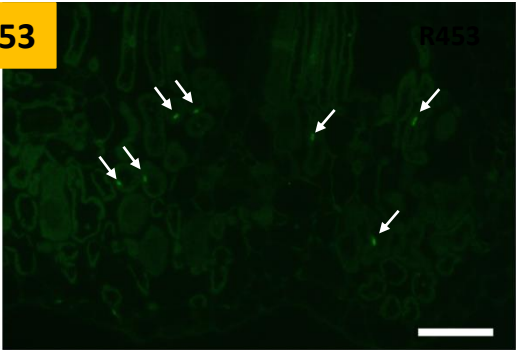
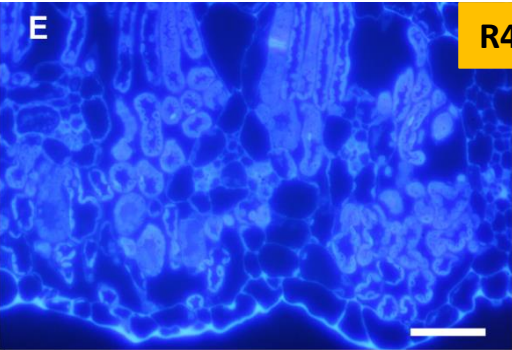
Early senescence genotype	Delayed senescence genotype
R453	B481-6



Programed cell death (PCD) detection
10 days after flowering

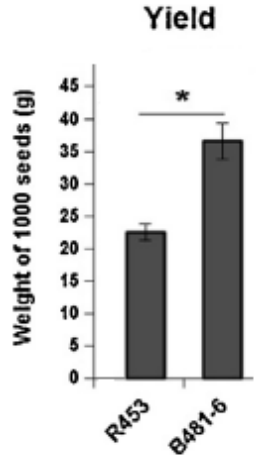
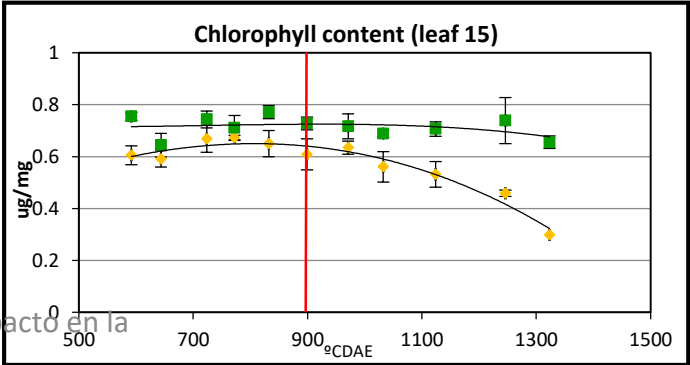
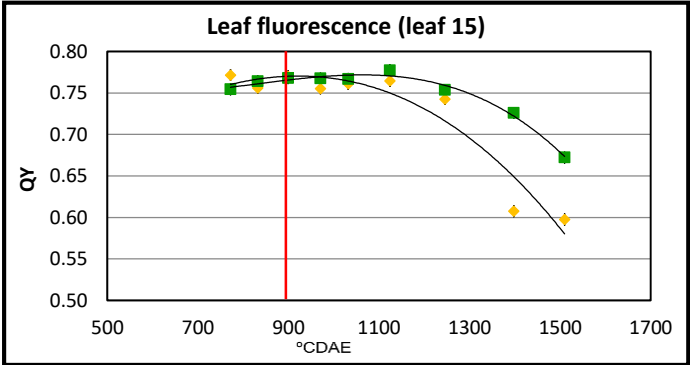
DAPI

TUNEL



10 days after flowering

Lopez-Gialdi et al. 2016
Moschen et al, 2017
Marino et al 2017



Early senescence
genotype

R453



10 DAA

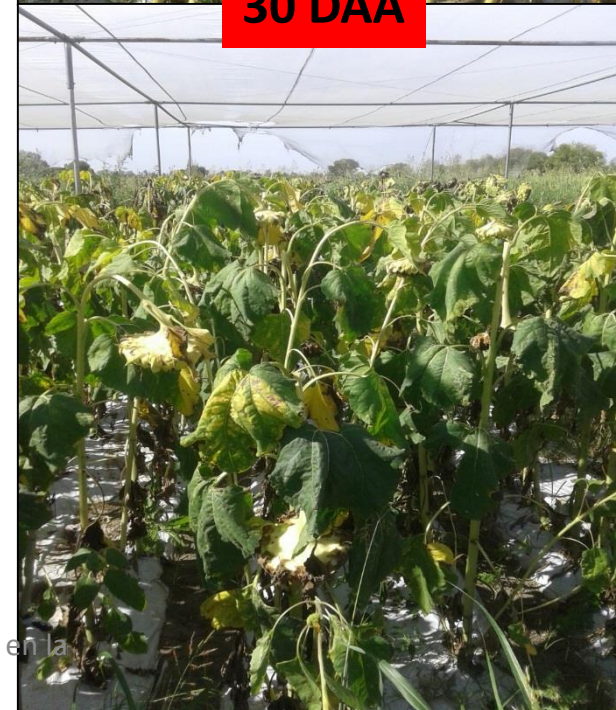


Delayed senescence
genotype

B481-6



30 DAA



LETTER

OPEN

doi:10.1038/nature22380

The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution

Hélène Badouin^{1*}, Jérôme Gouzy^{1*}, Christopher J. Grassa^{1,2*}, Florent Murat³, S. Evan Staton², Ludovic Cottret¹, Christine Lelandais-Brière^{4,5}, Gregory L. Owens², Sébastien Carrère¹, Baptiste Mayjonade¹, Ludovic Legrand¹, Navdeep Gill², Nolan C. Kane^{2,6}, John E. Bowers⁷, Sarel Hubner^{2,8,9}, Arnaud Bellec¹⁰, Aurélie Bérard¹⁰, Hélène Bergès¹⁰, Nicolas Blanchet¹, Marie-Claude Boniface¹, Dominique Brunel¹¹, Olivier Catrice¹, Nadia Chaidir^{2,12}, Clotilde Claude¹³, Cécile Donnadieu¹⁴, Thomas Faraut¹⁵, Ghislain Fievet¹, Nicolas Helmstetter¹⁰, Matthew King^{2,16}, Steven J. Knapp¹⁷, Zhao Lai^{18,19}, Marie-Christine Le Paslier¹¹, Yannick Lippi¹, Lolita Lorenzon¹, Jennifer R. Mandel²⁰, Gwenola Marage¹, Gwenaelle Marchand¹, Elodie Marquand¹¹, Emmanuelle Bret-Mestries²¹, Evan Morien², Savithri Nambesan²², Thuy Nguyen^{2,23}, Prune Pegot-Espagnet¹, Nicolas Pouilly¹, Frances Raftis², Erika Sallet¹, Thomas Schiex²⁴, Justine Thomas¹, Céline Vandecasteele¹⁴, Didier Varès¹, Felicity Year³, Sonia Vautrin¹⁰, Martin Crespi^{4,5}, Brigitte Mangin¹, John M. Burke⁷, Jérôme Salse¹, Stéphane Muñoz³, Patrick Vincourt³, Loren H. Rieseberg^{2,18} & Nicolas B. Langlade³

The domesticated sunflower, *Helianthus annuus* L., is a global oil crop that has promise for climate change adaptation, because it can maintain stable yields across a wide variety of environmental conditions, including drought¹. Even greater resilience is achievable through the mining of resistance alleles from compatible wild sunflower relatives^{2,3}, including numerous extremophile species⁴. Here we report a high-quality reference for the sunflower genome (3.6 gigabases), together with extensive transcriptomic data from vegetative and floral organs. The genome mostly consists of highly similar, related sequences⁵ and required single-molecule real-time sequencing technologies for successful assembly. Genome analyses enabled the reconstruction of the evolutionary history of the Asterids, further establishing the existence of a whole-genome triplication at the base of the Asterids II clade⁶ and a sunflower-specific whole-genome duplication around 29 million years ago⁷. An integrative approach combining quantitative genetics, expression and diversity data permitted development of comprehensive gene networks for two major breeding traits, flowering time and oil metabolism, and revealed new candidate genes in these networks. We found that the genomic architecture of flowering time has been shaped by the most recent whole-genome duplication, which suggests that ancient paralogues can remain in the same regulatory networks for dozens of millions of years. This genome represents a cornerstone for future research programs aiming to exploit genetic diversity to improve biotic and abiotic stress resistance and oil production, while also considering agricultural constraints and human nutritional needs^{8,9}.

As the only major crop domesticated in North America, with its sun-like inflorescence that inspired artists, the sunflower is both a social icon and a major research focus for scientists. In evolutionary biology, the *Helianthus* genus is a long-time model for hybrid speciation and adaptive introgression¹⁰. In plant science, the sunflower is a model

for understanding solar tracking¹¹ and inflorescence development¹². Despite this large interest, assembling its genome has been extremely difficult as it mainly consists of long and highly similar repeats. This complexity has challenged leading-edge assembly protocols for close to a decade¹³.

To finally overcome this challenge, we generated a 102× sequencing coverage of the genome of the inbred line XHQ using 407 single-molecule real-time (SMRT) cells on the PacBio RS II platform. Production of 32 million very long reads allowed us to generate a genome assembly that captures 3 gigabases (Gb) (80% of the estimated genome size) in 13,957 sequence contigs. Four high-density genetic maps were combined with a sequence-based physical map to build the sequences of the 17 pseudo-chromosomes that anchor 97% of the gene content (Fig. 1 and Supplementary Note 1.1–1.6). This compares favourably to an assembly of another sunflower genotype (HA412-HO; Supplementary Note 1.7), based on second-generation sequencing data, in which 2 Gb of sequence are placed in 816,854 contigs and 31,392 scaffolds. The sunflower genome encodes 52,232 inferred protein-coding genes and 5,803 spliced long non-coding RNAs (lncRNAs, Supplementary Note 2.1). To build the first small-RNA-mediated regulatory network for the sunflower, we identified 123 microRNA (miRNA) genes that we classified into 43 families (Supplementary Data 1), including 16 novel families. Sixty-three lncRNAs and 1,020 miRNAs are predicted to be miRNA targets, including 71 loci that probably produce secondary phased short-interfering RNAs (siRNAs, Supplementary Note 2.2).

More than three quarters of the sunflower genome consisted of long terminal repeat retrotransposons (LTR-RTs), of which 59% belong to the *Gypsy* evolutionary lineage. Sunflower LTR-RT lineages are predominantly young and exhibit minimal sequence divergence owing to significant expansion in the past one million years⁵. This pattern contrasts with that of DNA transposons, where the greatest density of

GENOME SEQUENCING

Illuminating the sunflower genome

A high-quality sunflower genome provides insight into Asterid genome evolution. Moreover, integrative analyses based on quantitative genetics, expression and diversity data uncover the gene networks and candidate genes for oil metabolism and flowering time, two important agronomic traits for sunflowers.

Sébastien Renaut

The common sunflower *Helianthus annuus*, well known for its large composite flowers, is cultivated around the world as an important source of seeds and oil. The sunflower and its wild relatives have also been used as a model system to study fundamental questions in evolutionary biology, such as the appearance of new species. Yet despite its economic, cultural and scientific importance, efforts by researchers to decipher this complex and large genome had remained largely unfruitful until the recent publication of Badouin and colleagues' work¹, which reports the first high-quality genome assembly of the common sunflower.

The key to circumventing some of the inherent complexity of this genome was to sequence very long stretches of DNA using the technology developed by Pacific Biosciences². Once these random DNA sequences were assembled into large sections, the researchers used genetic maps to order the pieces of this enormous puzzle back together (Fig. 1). In the end, the results are impressive and yielded a nearly complete high-quality reference genome

measuring 3.6 gigabases long (that is 3.6 billion A, C, T or G nucleotides), spread amongst 17 chromosomes. For the sake of comparison, this is slightly larger than the human genome with its 3.1 gigabases and 23 chromosomes. The real difficulty with the sunflower, however, is the fact that more than three-quarters of the total genome's length is comprised of long repeated sections of DNA, called long terminal repeat retrotransposons. These stretches of DNA all look very much alike, making it challenging to know which pieces belong where.

To better understand the evolutionary history of the sunflower's genome, Badouin *et al.* compared it to several closely (lettuce, artichoke) and more distantly (coffee, grape) related species. They were able to confirm a previously known large increase of the genome's size, which tripled its length. This whole-genome triplication, common to many flowering plants (the eudicots), dates back 122–164 million years. In addition, a second round of genome triplication took place 38–50 million years ago in the common ancestor of the sunflower, artichoke and lettuce³. Finally,

the sunflower genus underwent another whole-genome duplication nearly 29 million years ago. These multiple rounds of genome replication, which are followed by complex reorganization of the chromosomes (over time, pieces of chromosomes tend to fall off, be deleted and/or move around), are largely responsible for the complex and repetitive nature of the sunflower genome. Whether having a large genome where many genes are present in multiple copies has allowed the sunflower and its wild relatives to easily and quickly adapt to a wide variety of environmental conditions, including extreme ones, remains to be explicitly tested.

During more modern times, wild sunflowers were one of the few crops domesticated in North America⁴. Nearly 4,000 years ago, Native Americans started selecting for plants with large seeds containing a greater quantity of nutritive oil. Following this first step of domestication, modern breeders then improved these locally adapted cultivars, for example by selecting for plants that flower early and produce seeds over a short growing season. This allowed sunflower cultivation to spread

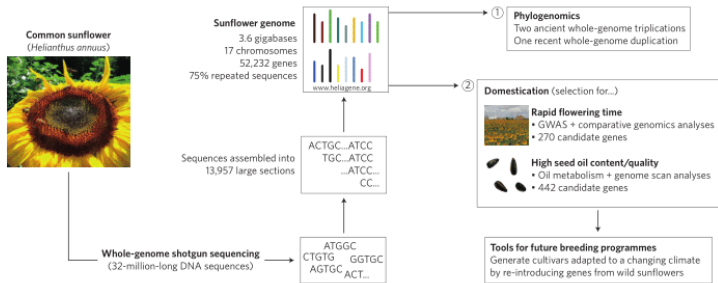
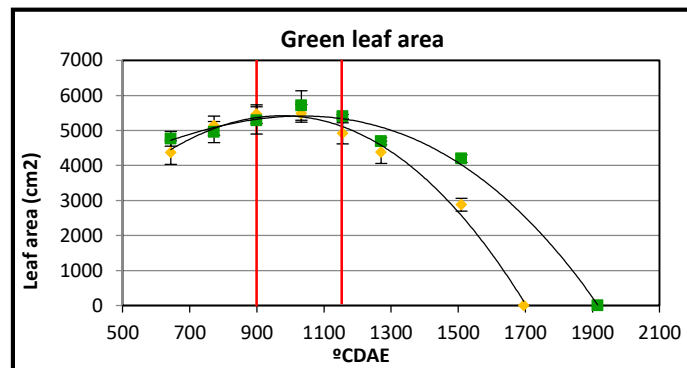


Figure 1 | Summary of the experimental design used to construct the sunflower genome, from whole-genome shotgun sequencing to sequence assembly and the high-quality genome. Phylogenomic, domestication and selection analyses were then performed to shine light on the sunflower's ancient and modern evolutionary history and identify the genetic basis of key agricultural traits. GWAS, genome-wide association study.

RNAseq analysis of sunflower leaf senescence: Illumina HiSeq

Early senescence genotype	R453
Delayed senescence genotype	B481-6

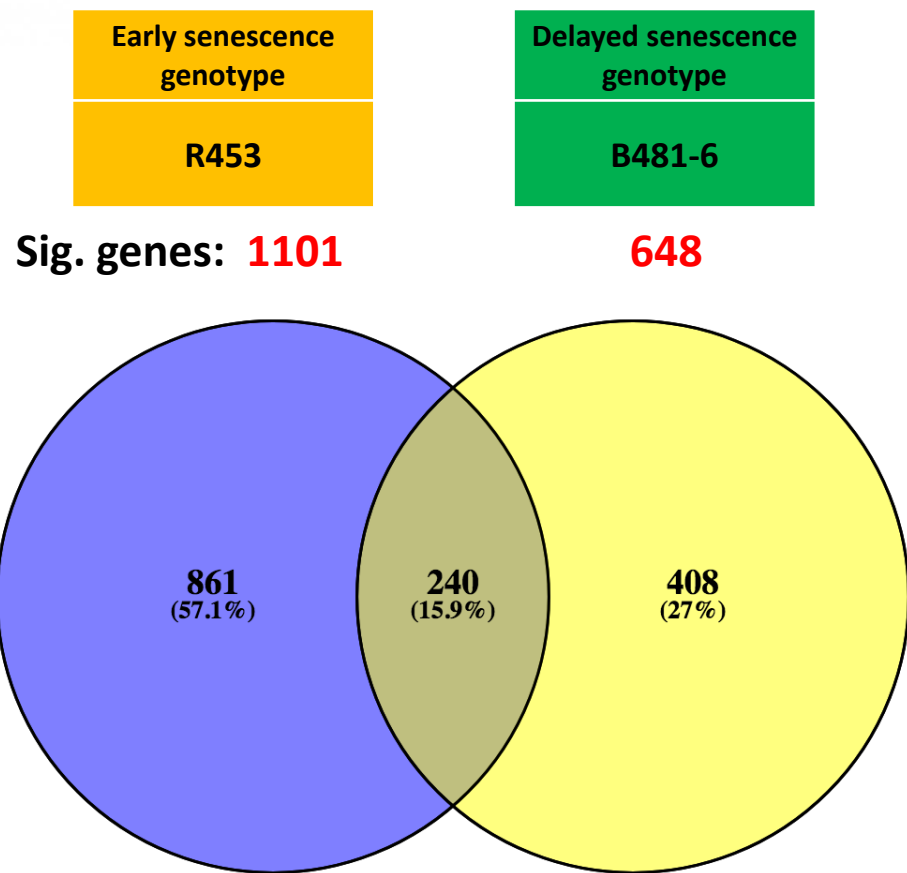


- Leaf 15
- 2 Genotypes
- Anthesis and Post-anthesis-12DDA
- 3 Biological replicates (pool of 3 plants)

Sunflower genome

- **90935** protein coding genes (*Sunflower Genome Consortium Meeting 2015*)
 - ✓ **59817** with Full Length Best Hits (spanning 60% of the length of the *A. thaliana* | SwissProt | Unitprot_plant protein) OR « EST/RNAseq assembly » support.
 - ✓ 39050 with EST support over 80% of the mRNA.
 - ✓ 13568 gene models correspond to full length *A. thaliana* proteins.
- **New Sunflower genome 2017** (*Badouin, Helene et al. 2017, Nature*)

Functional analysis Post-anthesis vs. anthesis time (Mapman annotation)



R453 and B481-6: (240 genes)

- All the genes showed the same expression pattern.
- Most of the genes presented higher or lower expression level in the early senescence genotype (R453).

Early senescence genotype R453: (861 genes)

- Higher expression levels of genes related to transporter family, cell wall and lipid degradation, protein degradation, redox homeostasis.
- Lower expression levels of cell cycle, RNA transcription, secondary metabolism and genes related auxin metabolism and synthesis of different macromolecules.

delayed senescence genotype B481-6: (408 genes)

- Higher expression levels of genes related to heat shock proteins, signalling light, signalling receptor kinases, biotic and abiotic stress response.
- Lower expression levels of genes related to transport, some genes related to lipid metabolism.

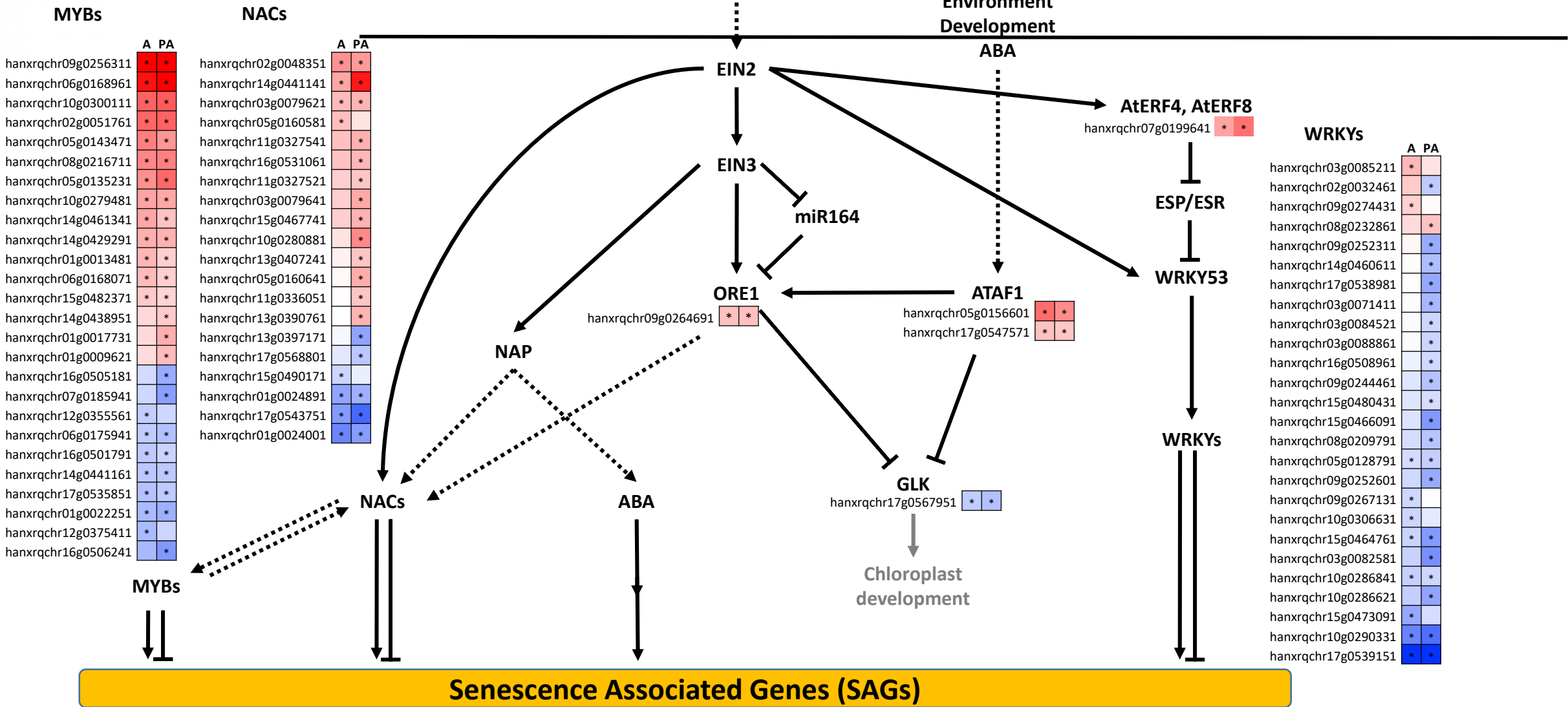
Padj<0.05 and Log2FC>1 o <-1



PERÚ

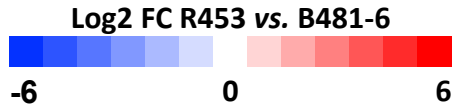
Ministerio de Agricultura y Riego

Transcription factor analysis R453 vs. B481-6: anthesis (A) and Post-anthesis (PA)



Early senescence genotype	R453
Delayed senescence genotype	B481-6

III Foro Internacional en Cambio Climático: “Impacto en la agricultura del Perú” INIA-MINAGRI





PERÚ

Ministerio
de Agricultura y Riego

Transcription factors and QTLs associated to leaf senescence

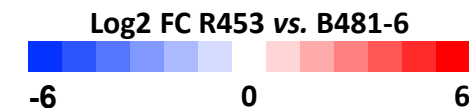
Collaboration Dr. Nicolas Langlade, LIPM-INRA Toulouse

Sunrise (France Sunflower Breeding Program)



- NAC, MYB and WRKY TFs were associated to QTLs for leaf senescence and with RNAseq data of Argentina genotypes.

QTL	SNP Distance	Nº env	Locus Tag	TF_Family	log2FC_A	log2FC_PA	Description
LES03.126	240844	2	hanxrqchr03g0079621	NAC	*	*	Putative NAC domain
LES03.126	60158	2	hanxrqchr03g0079641	NAC		*	Probable NAC domain containing protein 19 (HaNAC05 - transgenic line)
LES05.209	254237	4	hanxrqchr05g0160581	NAC	*		Probable NAC (No Apical Meristem) domain transcriptional regulator superfamily protein
LES05.209	458893	4	hanxrqchr05g0161341	NAC			Putative nascent polypeptide-associated complex NAC domain
LES05.209	44268	4	hanxrqchr05g0160711	WRKY			Putative WRKY domain
LES05.209	22316	4	hanxrqchr05g0162521	NAC			Putative NAC domain
LES05.209	156624	4	hanxrqchr05g0160641	NAC		*	Probable NAC domain containing protein 1
LES06.14	1162035	4	hanxrqchr06g0168961	MYB	*	*	Putative myb/SANT-like domain; Harbinger transposase-derived nuclease domain
LES06.14	144607	4	hanxrqchr06g0168071	MYB	*	*	Probable myb-like transcription factor family protein
LES06.14	170572	4	hanxrqchr06g0168101	WRKY			Putative WRKY domain
LES06.14	845073	4	hanxrqchr06g0170361	WRKY			Probable WRKY family transcription factor



- 4 Contrasting INTA sunflower genotypes
 - R456 y B451-6
 - B473 y C817
- Genotypes SUNRISE Project



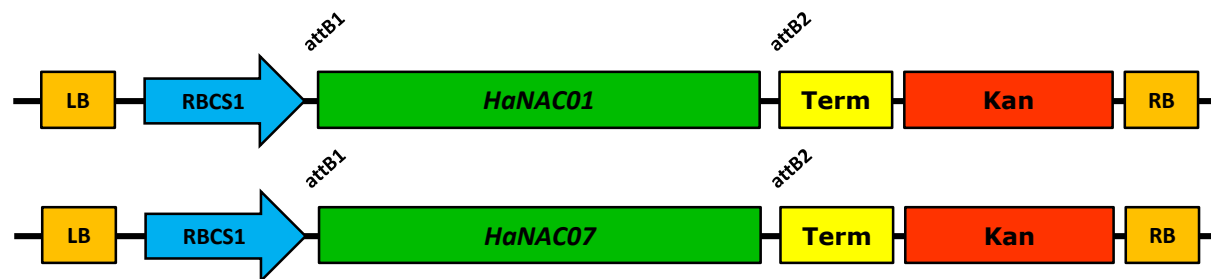
Measurements

- Number of senescent leaves
- Rank of youngest senescent leaf
- Insertion height of petiole of youngest senescent leaf
- Total leaf number
- Ratio of senescent leaf/total leaf number
- Plant height
- Stem diameter above cotyledons
- Leaf area at anthesis
- Dualex: Indices of Anthocyanins, Flavonols, Chlorophyll and Nitrogen Status

Collaboration with Dr. Nicolas Langlade (INRA)

Functional characterization of sunflower NAC transcription factors associated to leaf senescence in Arabidopsis

Transcription factors were isolated from sunflower, cloned and overexpressed in Arabidopsis.

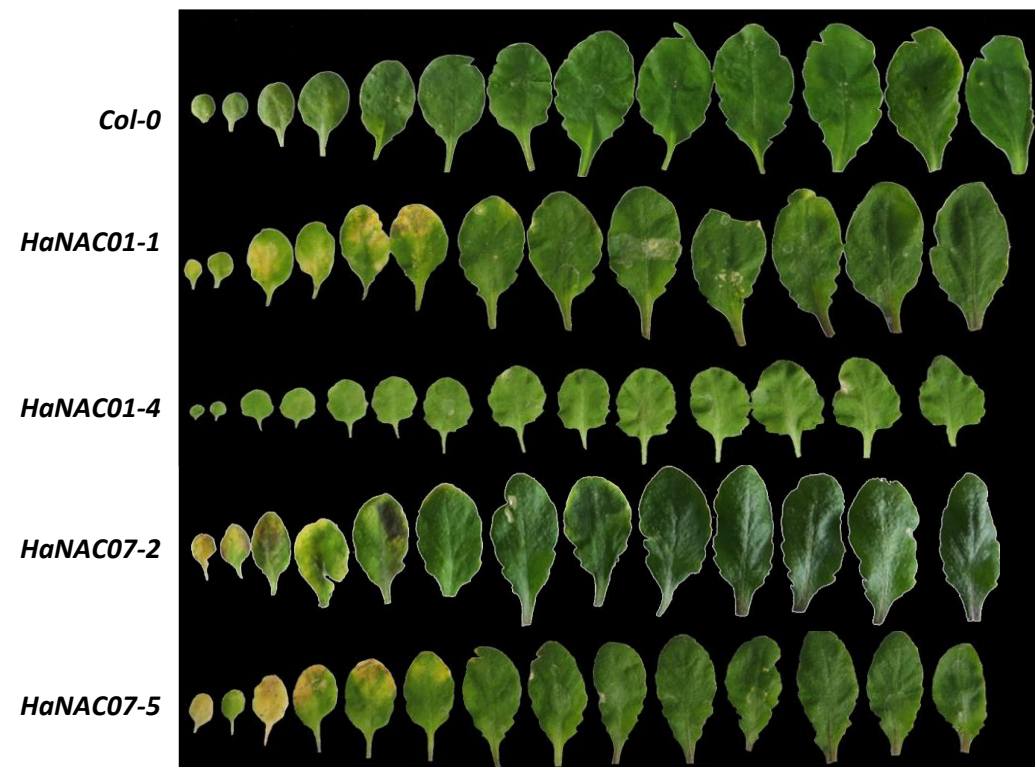
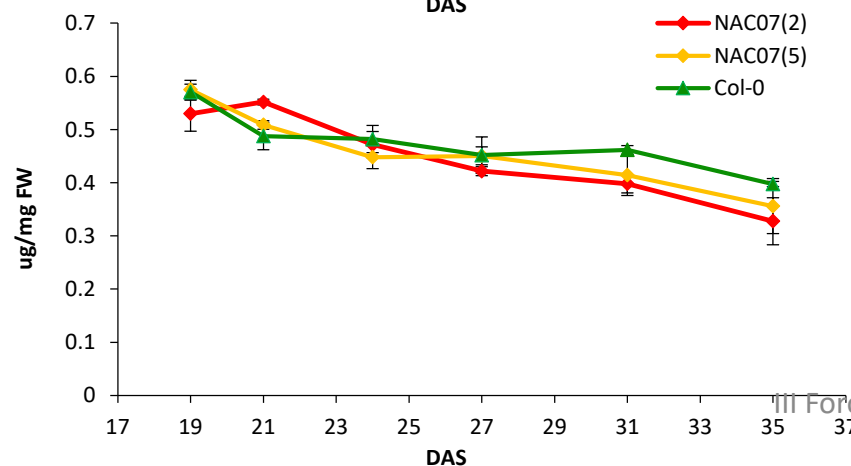
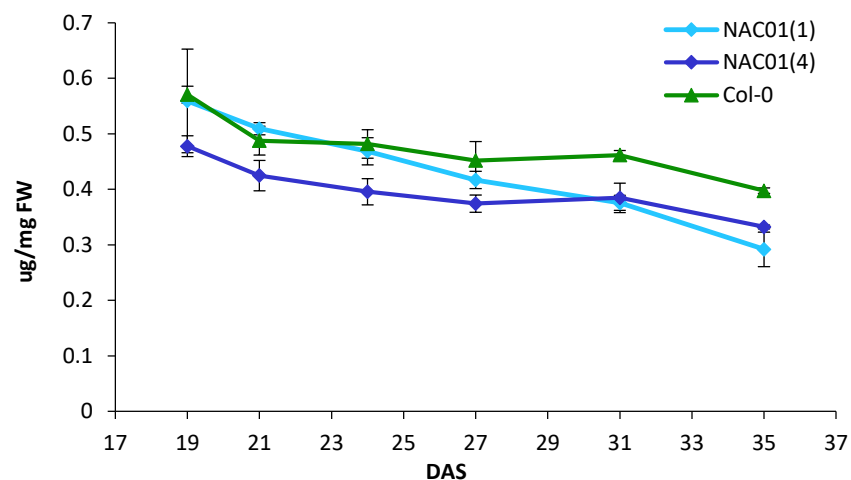


Independent transgenic lines for each transcription factor were performed.



Functional characterization of sunflower NAC transcription factors associated to leaf senescence in Arabidopsis

Chlorophyll content



Johanna Diaz, degree thesis, UNSAM

Phenotyping assays in Phenopsis platform (LEPSE – INRA Montpellier)

Dra. Paula Fernández External fellowship for Researchers CONICET.

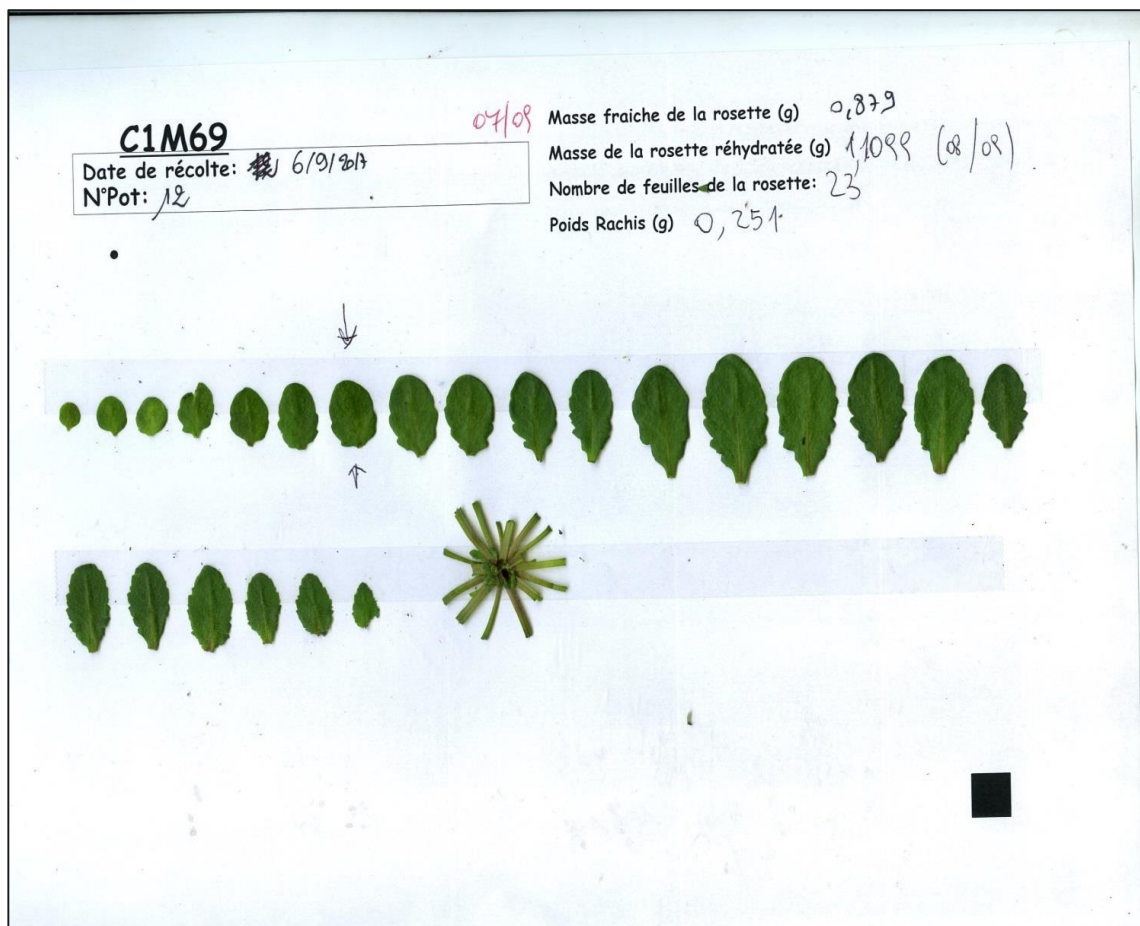
Collaboration with Dra. C .Vazquez and Dr .Denis Vile & Francoise Tardieu (Labintex-INTA / LEPSE-INRA Montpellier)

- Eight (8) *A. thaliana* transgenic lines + Col 0
- Control and drought stress treatment

Line name	Sunflower gene	Arabidopsis ortholog	Transgene expression level
HaNAC01-1	HaNAC01	ORE1 (ANAC100)	+
HaNAC01-4			+++
HaNAC03-3	HaNAC03	ANAC072	+
HaNAC03-5			+++
HaNAC05-1	HaNAC05	ANAC019	+++
HaNAC05-5			+
HaNAC07-2	HaNAC07	ANAC029/ATNAP	+++
HaNAC07-5			+



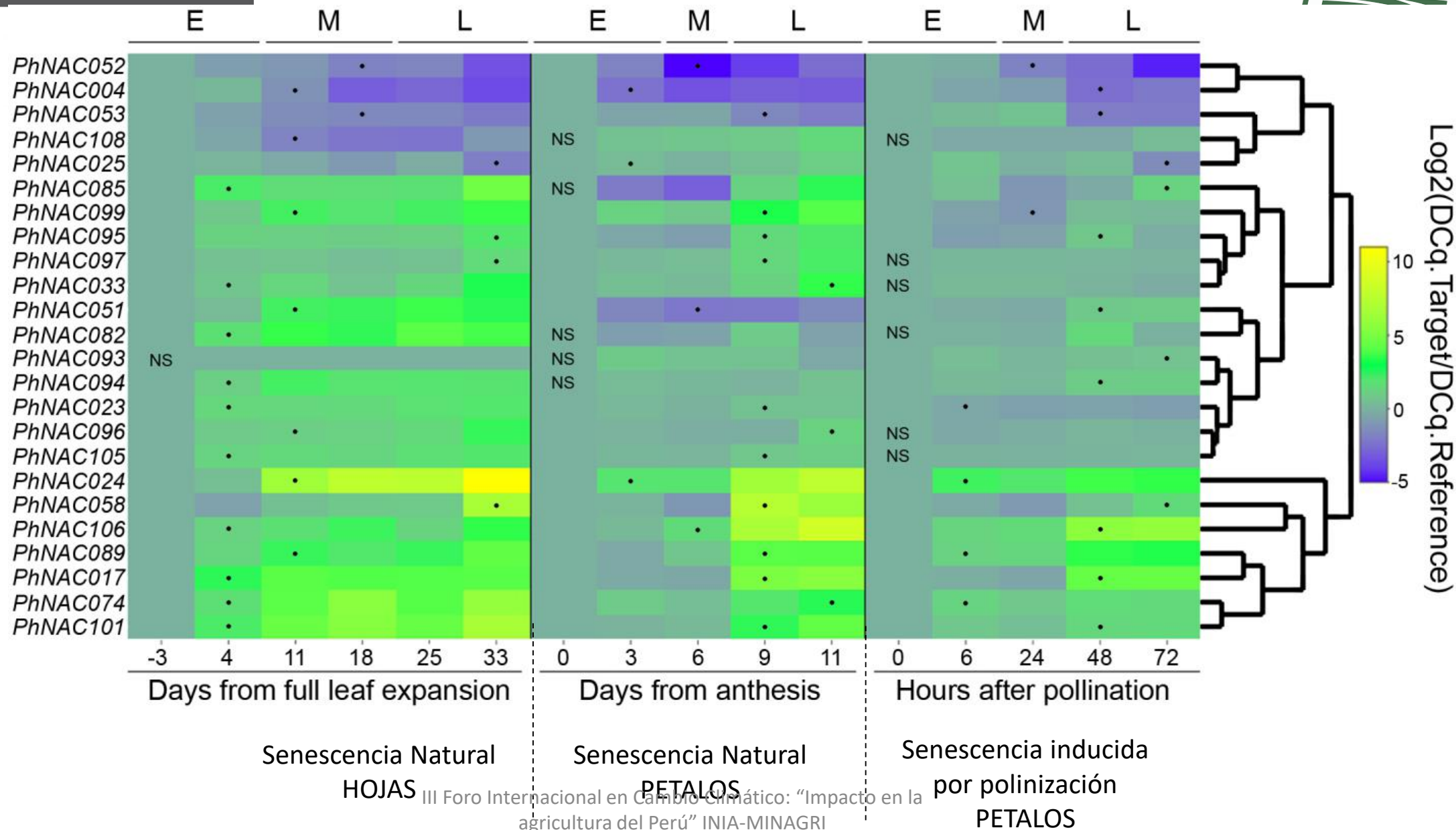
Pots/plants dissection

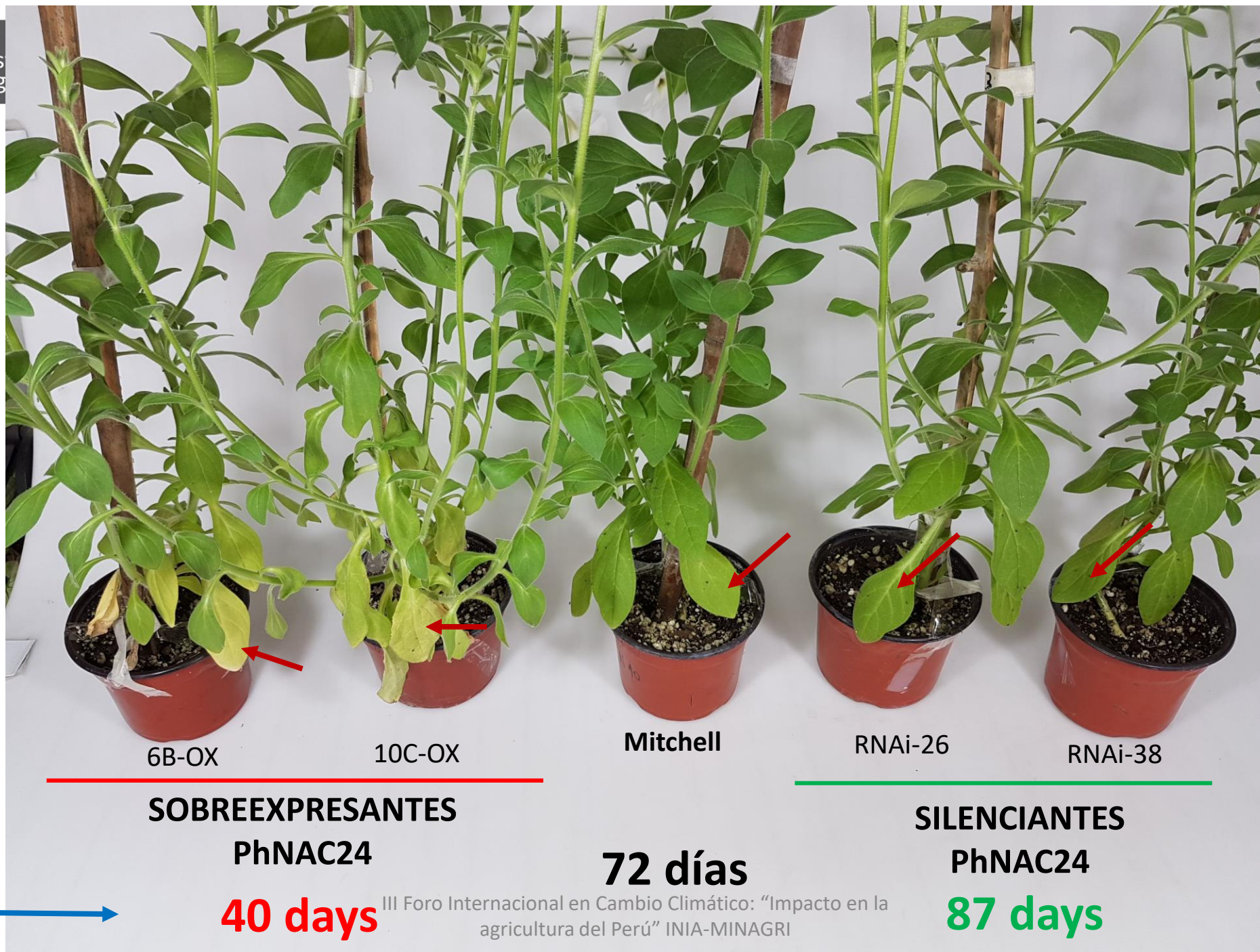


Leaf and flower senescence in *Petunia. hybrida*

Study the triggering of natural **foliar and floral senescence** in *Petunia hybrida* by identifying transcription factors involved in this process through a transcriptional and functional analysis of candidate genes

RNaseq data NAC candidate genes





Sunflower Genomics

Researchers

- Paula Fernandez
- Verónica Lia
- Maximo Rivarola
- Andrea Puebla
- Norma Paniego
- Ruth Heinz
- Esteban Hopp

Postdoc

- Sofia Bengoa Luoni
- Carla Filippi

PhD students

- Sergio Gonzalez
- Mónica Fass
- Juan Montecchia
- Salvador Nicosia

Students

- Johanna Diaz
- Johanna Marino

INTA collaborators

Sunflower pathology group (EEA Balcarce)

- Alberto Escande, Facundo Quiróz and Carla Maringolo

Sunflower breeding group (EEAManfredi)

- Daniel Alvarez and María Valeria Moreno

Sunflower physiology group (EEA Balcarce)

- Guillermo Dosio and Luis Aguirrezabal

Crops Ecophysiology under Abiotic Stress

- Martin Acreche (EEA Cerrillos – Salta)

INTA LABINTEX

- Cecilia Vazquez Rovere
(INTA LABINTEX – LEPSE INRA Montpellier, France)

External collaborators

- Julio Di Rienzo (Universidad Nacional Córdoba)
- Sergio Lew (UBA-FI)
- Gerardo Cervigni (CEFOBI/CONICET)
- Sara Maldonado (UBA-FCEN)
- Bernardo Clavijo and Janet Higgins (EI, UK)
- Giusi Zaina and Michele Morgante UdiUn, IT
- Joaquin Dopazo, Ana Conesa and Paco Garcia (CIPF, España)
- Alisdair Fernie (Max Planck Institute)
- Nicolas Langlade (LIPM-INRA, Toulouse, France)
- Denis Vile (LEPSE INRA Montpellier, France)





PERÚ

Ministerio
de Agricultura y Riego



Funding



AGENCIA
NACIONAL DE PROMOCION
CIENTIFICA Y TECNOLOGICA



INRA
SCIENCE & IMPACT



DAAD

Deutscher Akademischer Austausch Dienst
Servicio Alemán de Intercambio Académico



III Foro Internacional en Cambio Climático: “Impacto en la
agricultura del Perú” INIA-MINAGRI

Thank you very much!

III Foro Internacional en Cambio Climático: "Impacto en la
agricultura del Perú" INIA-MINAGRI