



THE UNIVERSITY OF
WESTERN
AUSTRALIA

Using transcriptomics to explain the mechanisms of plant hormone perception

Mark Waters

School of Molecular Sciences

SCIE4001 | 05 March 2024

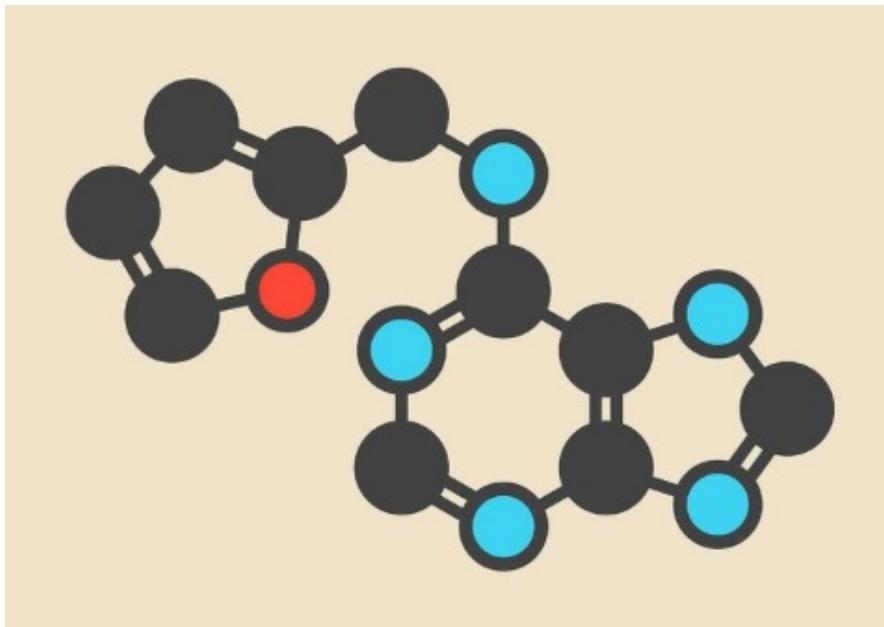


watersmt.org

@watersmt

@SMS_UWA

How do plants co-ordinate growth?



- A plant growth regulator is an organic compound, either natural or synthetic, that modifies or controls one or more specific physiological processes within a plant
- Commonly they are identical to, or mimics of, natural plant hormones that regulate growth
- Act at a distance and at very low concentrations

Manipulation of plant growth by external factors

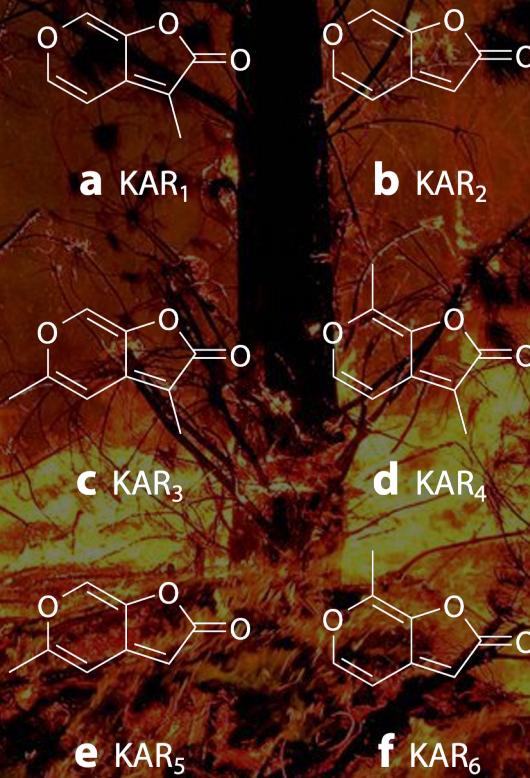
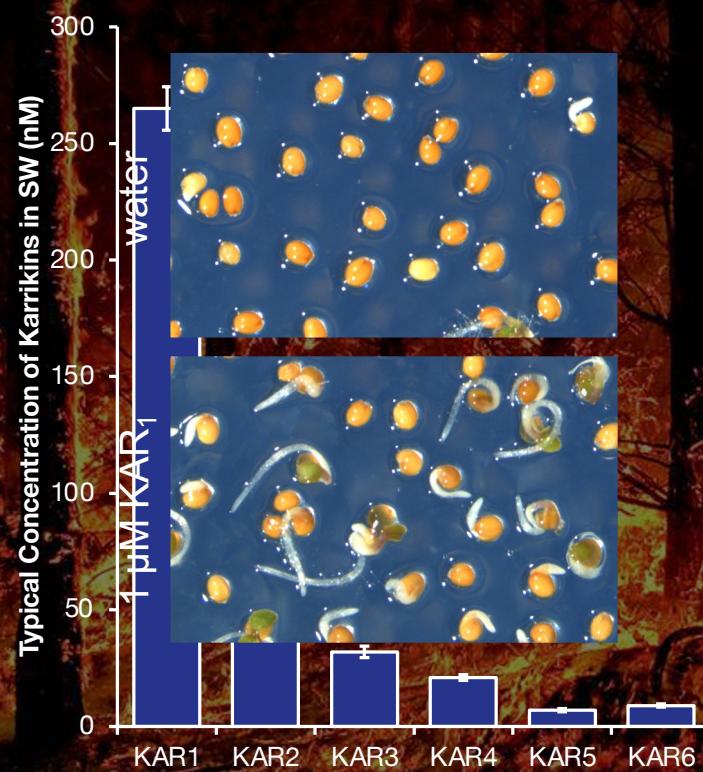
PGRs can be **endogenous** (plant hormones) or **exogenous** (hormone mimics)

Observations from the natural world can inform us about potential PGRs

Bakanae (“foolish seedling”) caused by fungus *Gibberella fujikuroi*



Karrikins are butenolide seed germination stimulants produced by fire



Flematti GR et al. (2004) Science 305: 977; Flematti GR et al. (2009) J. Agric. Food Chem. 57: 9475-9480
Nelson DC et al. (2012) Ann. Rev. Plant Biol. 29 April 2012

Post-fire emergence of yellow tailflower for first time in 20 years



2010



2009

Anthocercis littorea (Solanaceae)

Two projects from my lab:

1. Using transcriptomics to probe **diversity in karrikin receptors** in a pernicious weed, *Brassica tournefortii* (Sun et al. 2020 *Nature Communications* doi:[10.1038/s41467-020-14991-w](https://doi.org/10.1038/s41467-020-14991-w))
2. Using transcriptomics to **identify genes activated or repressed** by butenolide signalling in the bacterium *Bacillus subtilis* (Melville et al. 2024 *Current Biology* doi:[10.1016/j.cub.2023.12.035](https://doi.org/10.1016/j.cub.2023.12.035))

Two projects from my lab:

1. Using transcriptomics to probe **diversity in karrikin receptors** in a pernicious weed, *Brassica tournefortii* (Sun et al. 2020 *Nature Communications* doi:[10.1038/s41467-020-14991-w](https://doi.org/10.1038/s41467-020-14991-w))



ARTICLE

<https://doi.org/10.1038/s41467-020-14991-w>

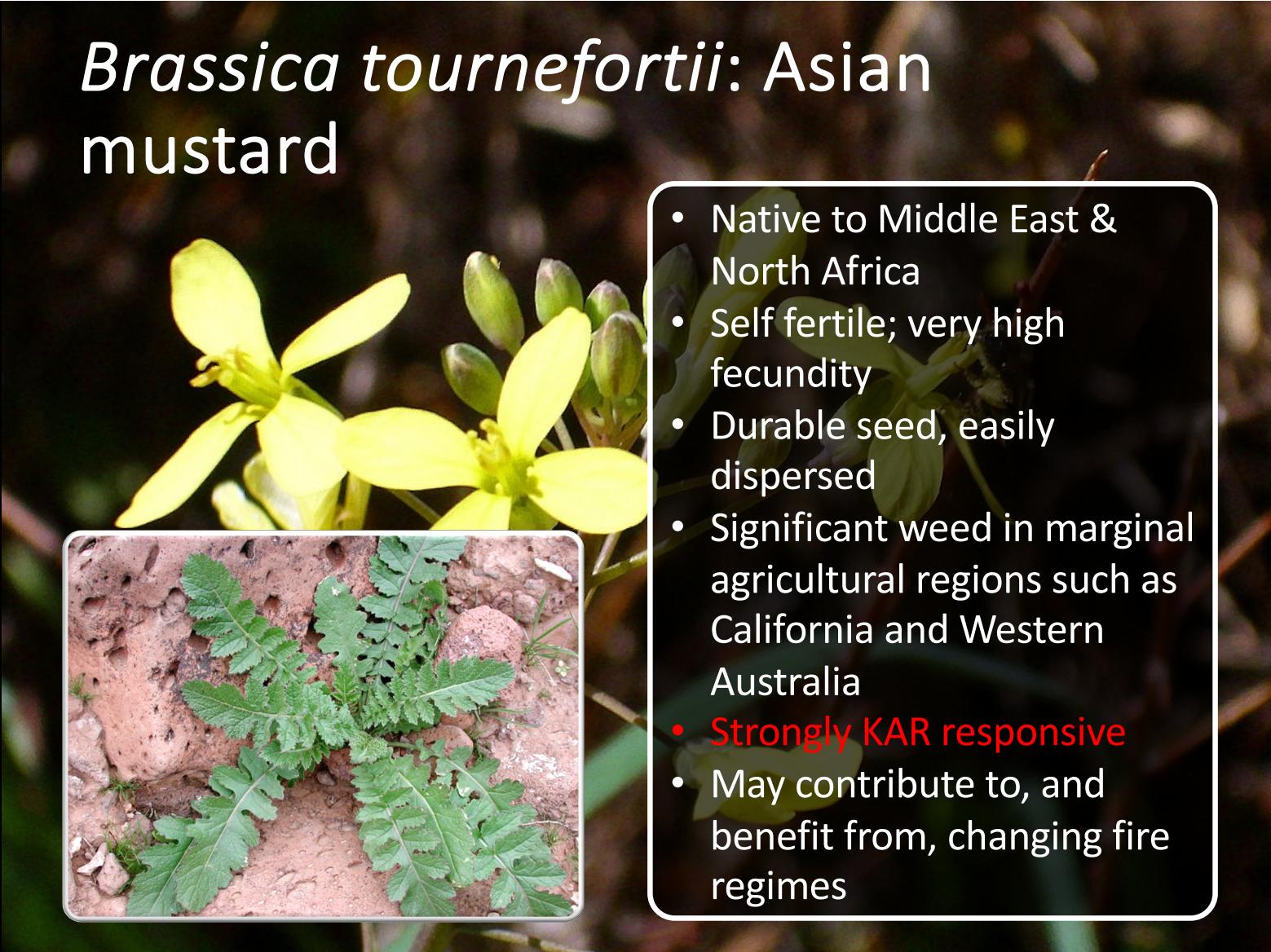
OPEN

Check for updates

Divergent receptor proteins confer responses to different karrikins in two ephemeral weeds

Yueming Kelly Sun^{1,2}, Jiaren Yao¹, Adrian Scaffidi¹, Kim T. Melville¹, Sabrina F. Davies¹, Charles S. Bond¹, Steven M. Smith^{3,4}, Gavin R. Flematti¹ & Mark T. Waters^{1,2}

Brassica tournefortii: Asian mustard

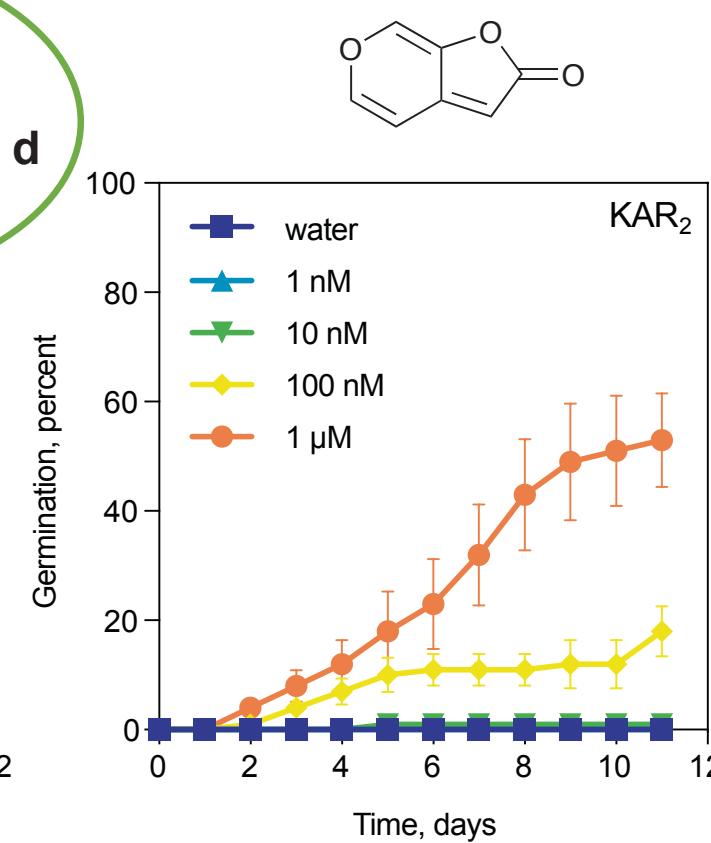
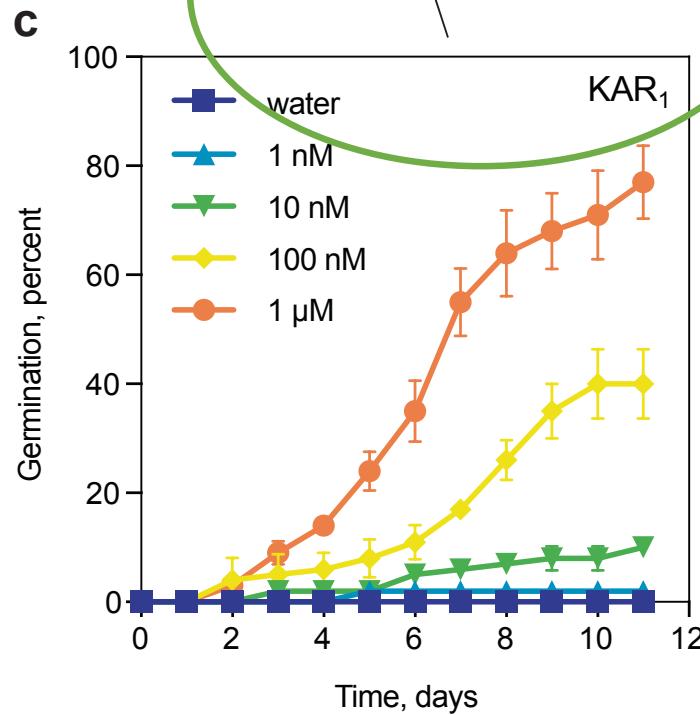


Mohawk Dunes, western Arizona

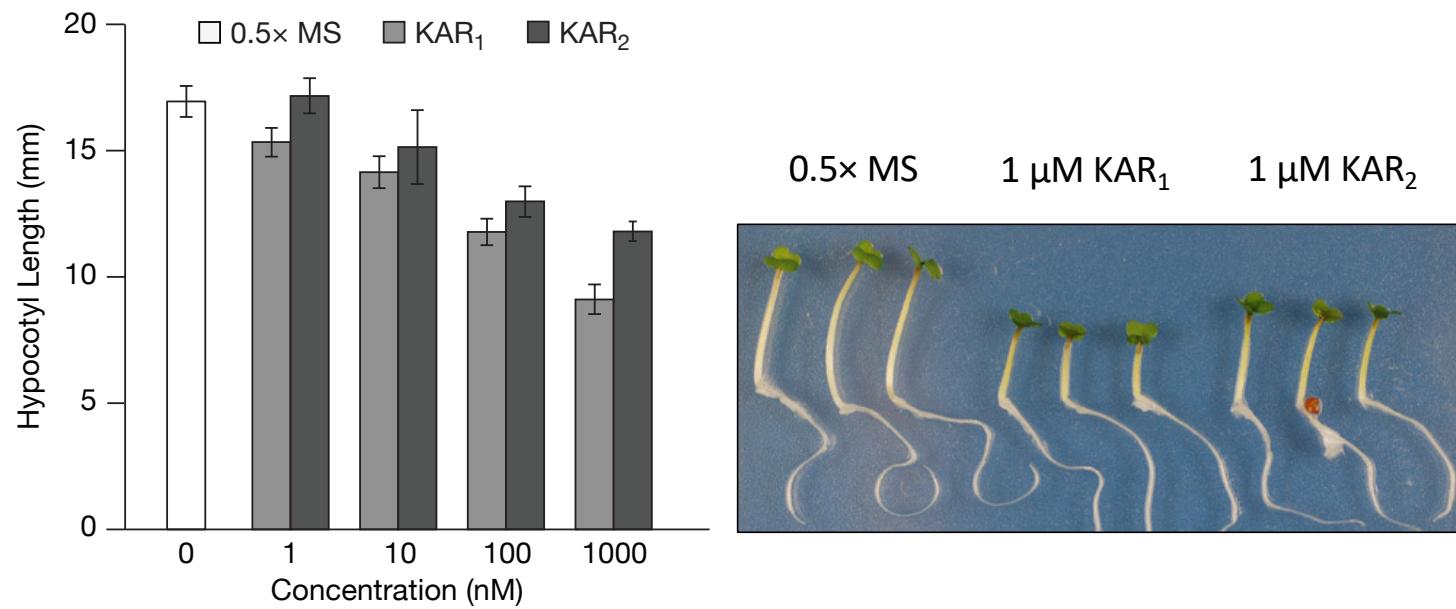


https://www.desertmuseum.org/invaders/invaders_saharamustard.php

B. tournefortii responds preferentially to KAR

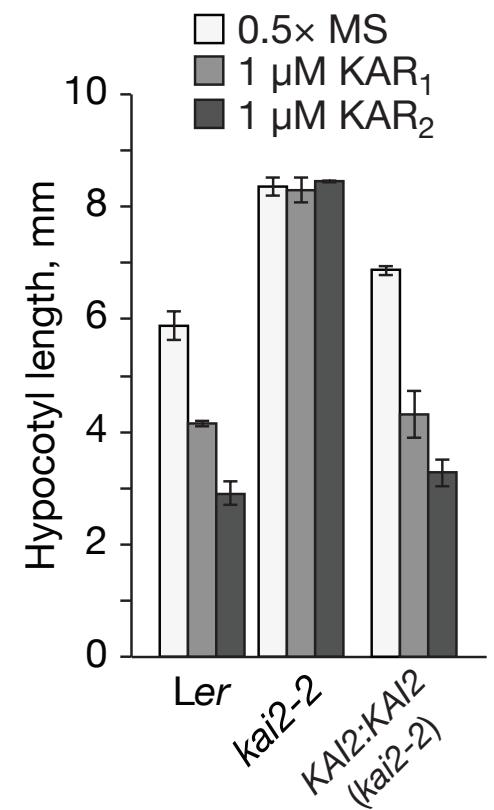
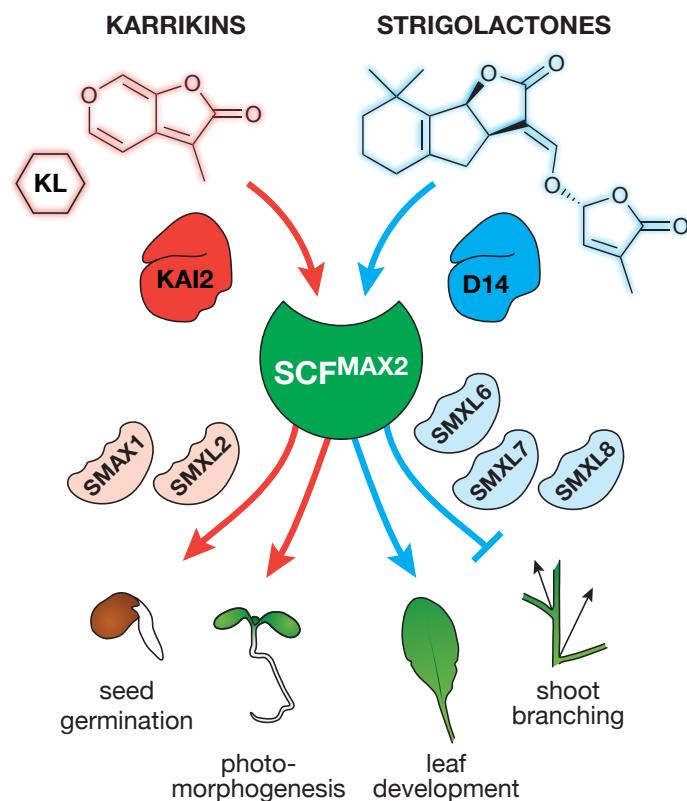


B. tournefortii responds positively and preferentially to KAR₁

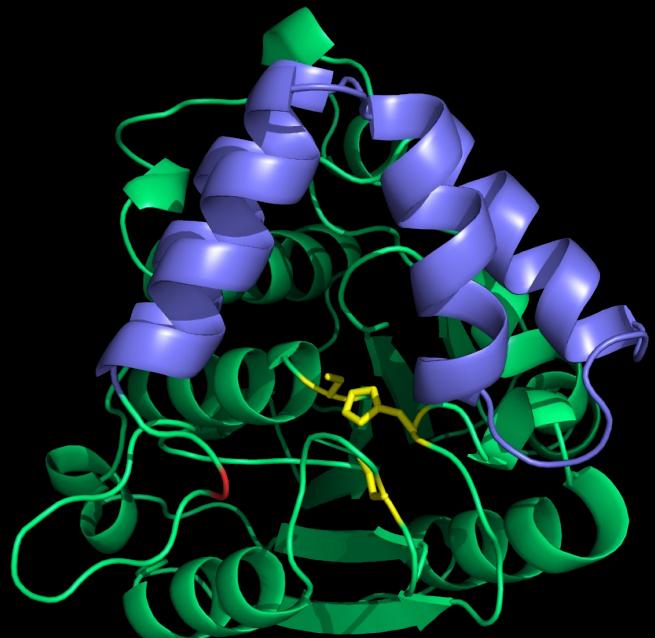


B. tournefortii is approximately 10-fold more sensitive to KAR₁ than to KAR₂
at level of physiological response

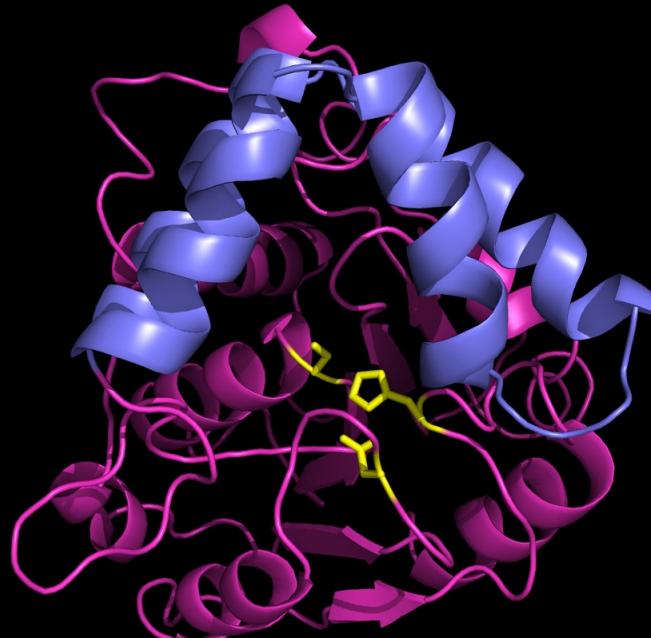
Arabidopsis thaliana responds preferentially to KAR₂ through KAI2



KAI2 and D14 are hydrolases with receptor-like properties



Arabidopsis KAI2
Karrikins
and "KL"



Arabidopsis D14
Strigolactones

Zhao et al (2013) Cell Research

Hypothesis

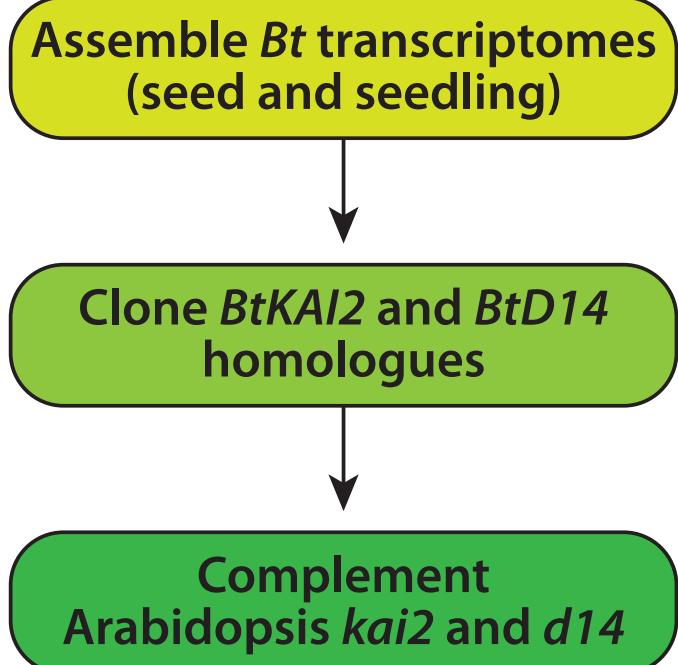
That differences in karrikin response between *Brassica tournefortii* and *Arabidopsis thaliana* are due to the expression and sequence divergence of different KAI2 homologues.

The *Brassica* lineage has undergone genome triplication in recent past, so we might expect up to three KAI2 homologues (*Arabidopsis* only has one).

We could sequence the genome *de novo*, but:

- a transcriptome gives us useful information (expression level, UTRs etc.).
- triplicated genomes are notoriously tricky to assemble (lots of very similar sequences), and it's overkill for our needs.

What factors account for differences in karrikin response between species?

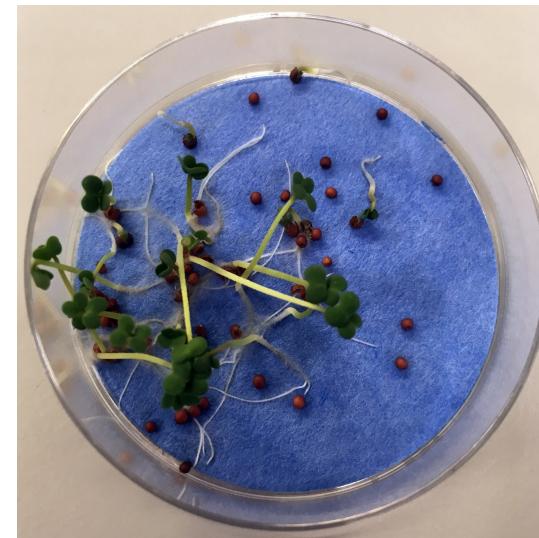


Generating transcriptomes

Seeds 24 h imbibed



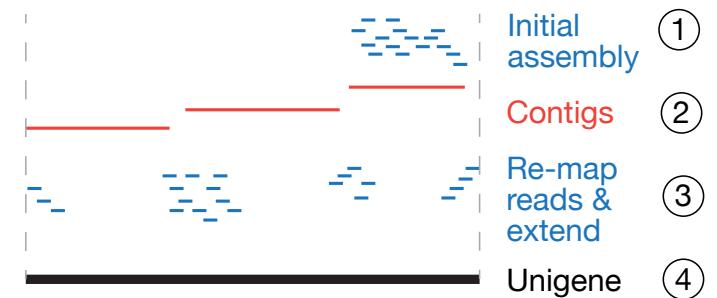
Seedlings 4 day old



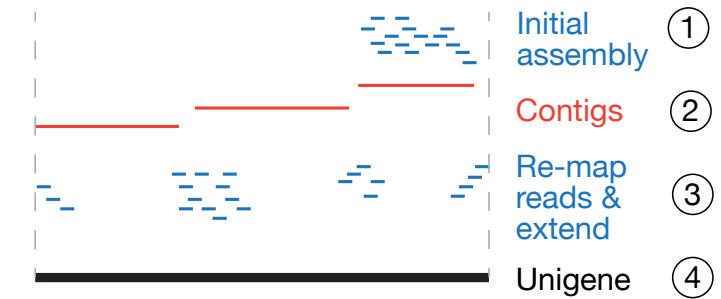
Generating transcriptomes



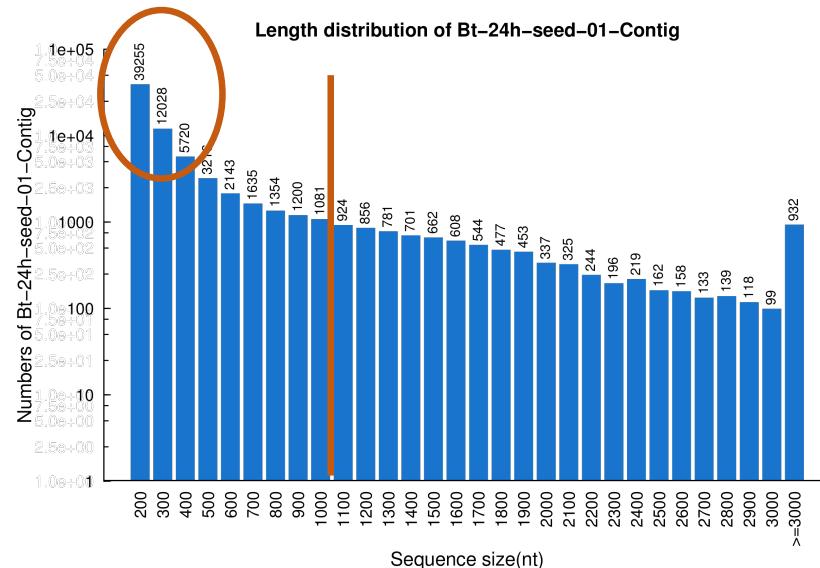
1. Isolate mRNA (polyA+ selection)
2. Generate cDNA libraries (fragmentation, RT)
3. Paired-end sequencing on Illumina HiSeq2000
4. Reads cleaned (removal of adapters, ambiguous/low quality reads): ~35 million reads, 99.9% >Q20
5. Assembled into contigs (mean length ~430 bp) and then longer sequences (“Unigenes”; mean length ~1000 bp) using Trinity



From contigs to unigenes

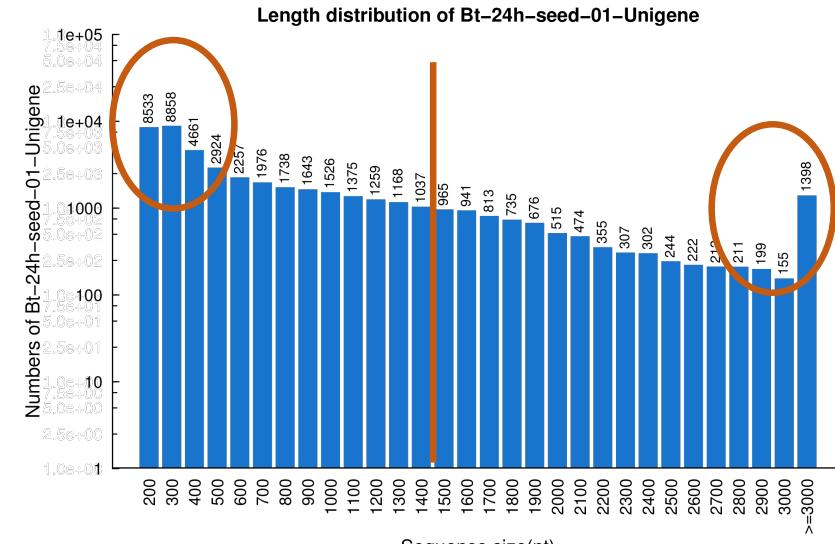


Contigs (chunks of contiguous reads)



N50: 1012 bp, mean: 445 bp

“Unigenes” (~= transcripts)



N50: 1434 bp, mean: 816 bp

Comparison with assembly statistics from *Arabidopsis thaliana* (Oxford Nanopore)

Stage	Arabid. seedling	Arabid. buds	Brassica seeds	Brassica seedling
Average length	1003	924	774	816
Median length	851	764	—	—
N50 length	1188	1141	1348	1434
Max read length	24734	24080	—	—

Note different sequencing technologies used (Illumina vs ONT; ONT produces much longer reads that are more likely to be complete transcripts, so good test of whether unigenes are indeed complete transcripts)

Arabidopsis data from:

New insights into *Arabidopsis* transcriptome complexity revealed by direct sequencing of native RNAs

Shoudong Zhang, Runsheng Li, Li Zhang, Shengjie Chen, Min Xie, Liu Yang, Yiji Xia, Christine H Foyer, Zhongying Zhao, and Hon-Ming Lam

[Nucleic Acids Res.](https://doi.org/10.1093/nar/gkaa588) 2020 48: 7700–7711. [10.1093/nar/gkaa588](https://doi.org/10.1093/nar/gkaa588)

Finding KAI2 homologues in *B. tournefortii*

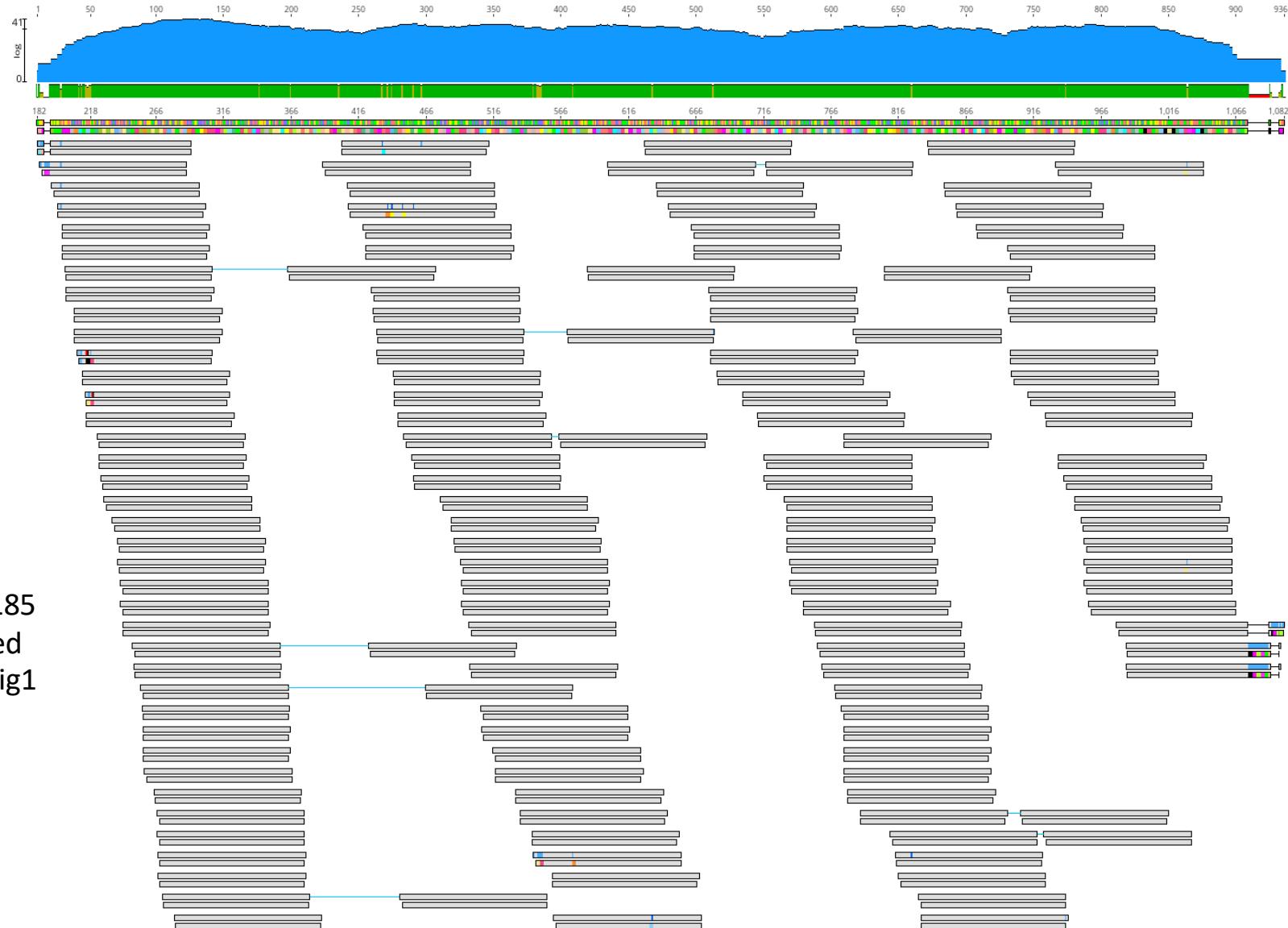
1. Unigenes from both libraries were combined and aligned by BLASTx to protein databases to derive 45,553 predicted coding region sequences (CDS)
2. Make custom BLASTn database with CDS and search for hits with query sequence (i.e. Arabidopsis KAI2)
3. Found 3 unique contigs => three putative homologues with different apparent expression levels:

Library	Total reads	Assembled reads	Contig 1	Contig 2	Contig 3
Seed	34,953,924	1838	185	1407	246
Seedlings	35,040,596	3630	913	2715	2

What are the differences between these sequences?

Are they “real” transcripts or imaginary?

Coverage
Identity



Amino acid sequences of all three homologues after confirmation by RT-PCR, cloning and sequencing

CONSENS

Consensus 1 10 20 30 40 50 60 70
MGGVVEEAHNVKVIGSGNQXTIVLGHGFGTDQS VWKHLVPHLVDDYRIVLYDNMGAGTTNPXYFDFDRYSTLE

AtKAI2

AtKAI2 M-GVVEEAHNVKVIGSG-EATIVLGHGFGTDQS VWKHLVPHLVDDYRIVLYDNMGAGTTNPDYFDFDRYSNLE

BtKAI2a

BtKAI2a MGGVVEEAHNVKVIGTGTQATIVLGHGFGT DQS VWKHLVPHLL EDYRIVLYDNMGAGTTNPDYFDFDRYSTLE

BtKAI2b MGGVVEEAHNVKVIGSGNQGTIVLGHGFGT DQS VWKHLVPHLVDDYRIVLYDNMGAGTTNPPEYFDFEIRYSTLE

BtKAI2c MGGVVEEAHNVKVIGSGNKGTIVLGHGFGT DQS VWRHLVPHLVDDYRIVLYDNMGAGTTNPPEYFDFDRYSTLE

80 90 100 110 120 130 140
GFSFDLIAILEDLQIESCIFVGHSVSAMGVVLASLNRPDLFSKIVMISASPRYVNDVDYQGGFEQXDLNQLFE

AtKAI2 GYSFDLIAILEDLKIESCIFVGHSVSAMI GVVLASLNRPDLFSKIVMISASPRYVNDVDYQGGFEQEDLNQLFE



alpha T1

BtKAI2a GFSFDLIAILEDLQIDSCIFVGHSVSAMGVLLASLNRPDLFSKIVMISASPRYVNDVDYQGGFEQDDLNQLFE

BtKAI2b GFSFDLIAILEDLQIESCIFVGHSLSAMGVVLASLNRPDLFSKIVMISASPRYVNDVDYQGGFEQDDLNQLFE

BtKAI2c GFAFDLIAILEDLQIESCIFVGHSRSAMGVVLASLNRPPELFSKLVMVSASPRFVNEDDYEGGFDQEDLKQLFE



150 160 170 180 190 200 210
AMRSNYKAWCLGFAPLAVGDDLD SVAVQEFSRTL FNMRPDIALSVAQTIFQS

AtKAI2 A IRSNYKAWCLGFAPLAVGDDMDSIAVQEFSRTL FNMRPDIALSVGQTIFQS



alpha T2



Dave's SMXL s...



alpha T4

BtKAI2a AMRSNYKAWCLGFAPLAVGDDLE SVAVQEFSRTL FNMRPDIALSVAQTIFQS

BtKAI2b AMRSNYKAWCLGFAPLAVGDDLD SVAVQEFSRTL FNMRPDIALSVAQTIFQS

BtKAI2c AMRSNYKAWCLGFAPLVVGGDLD SIAVQEFSRTL FNMRPDIALSMAQTIFS



220 230 240 250 260 270 272
LAVPVAVSEYLHTNLGCESVVEVXXSDGHLPQLSSPDSVIPVLLRHIRNDIAV

AtKAI2 LAVPVVVSEYLHANLGCE SVVEVIPS



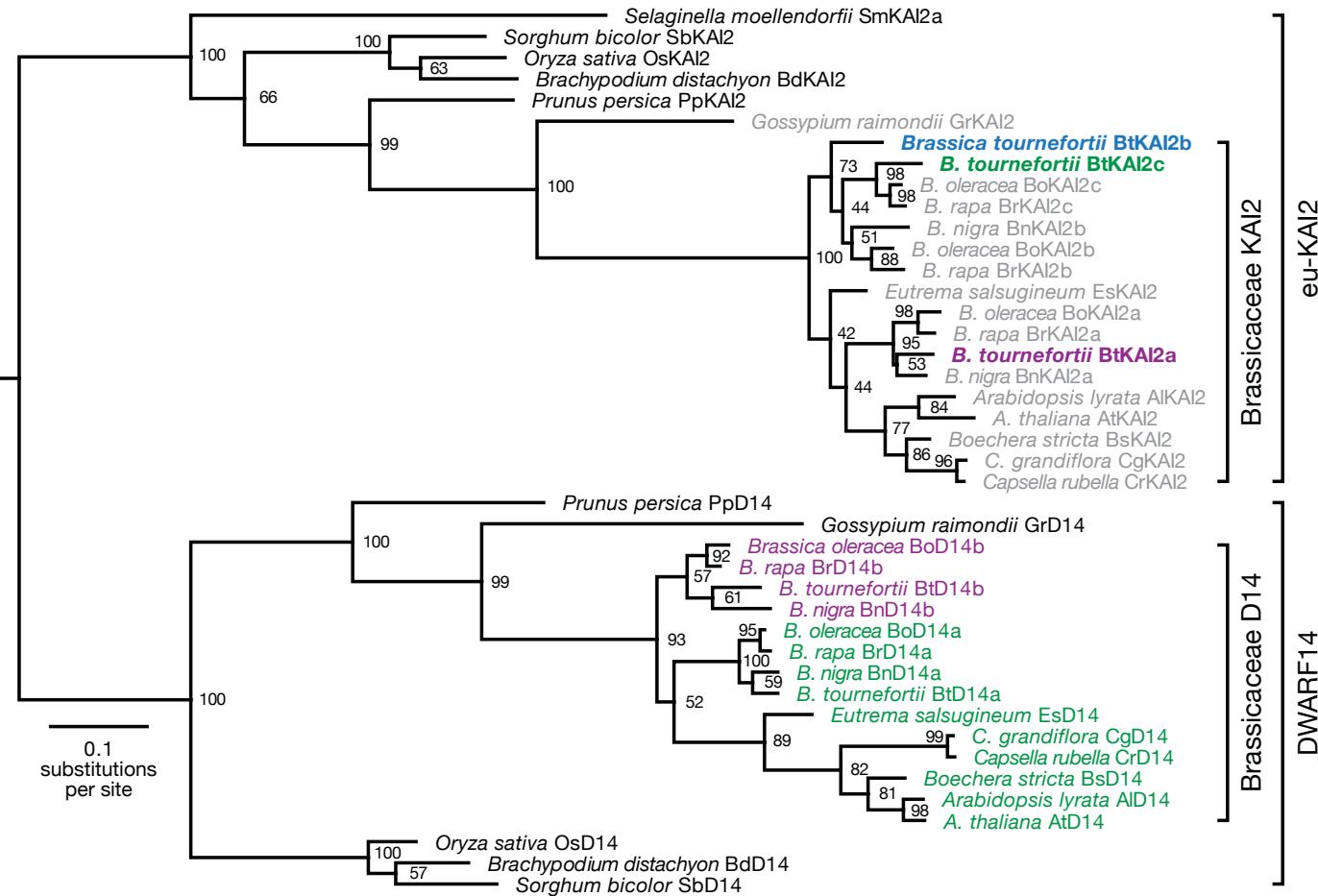
BtKAI2a LAVPVAVSEYLNTNLGCESVVEVIPS

BtKAI2b LAVPVAVSEYLHKNLGCESVVEVIMS

BtKAI2c LAVPVAVSEYLHTNLGSES VVEVIMSS



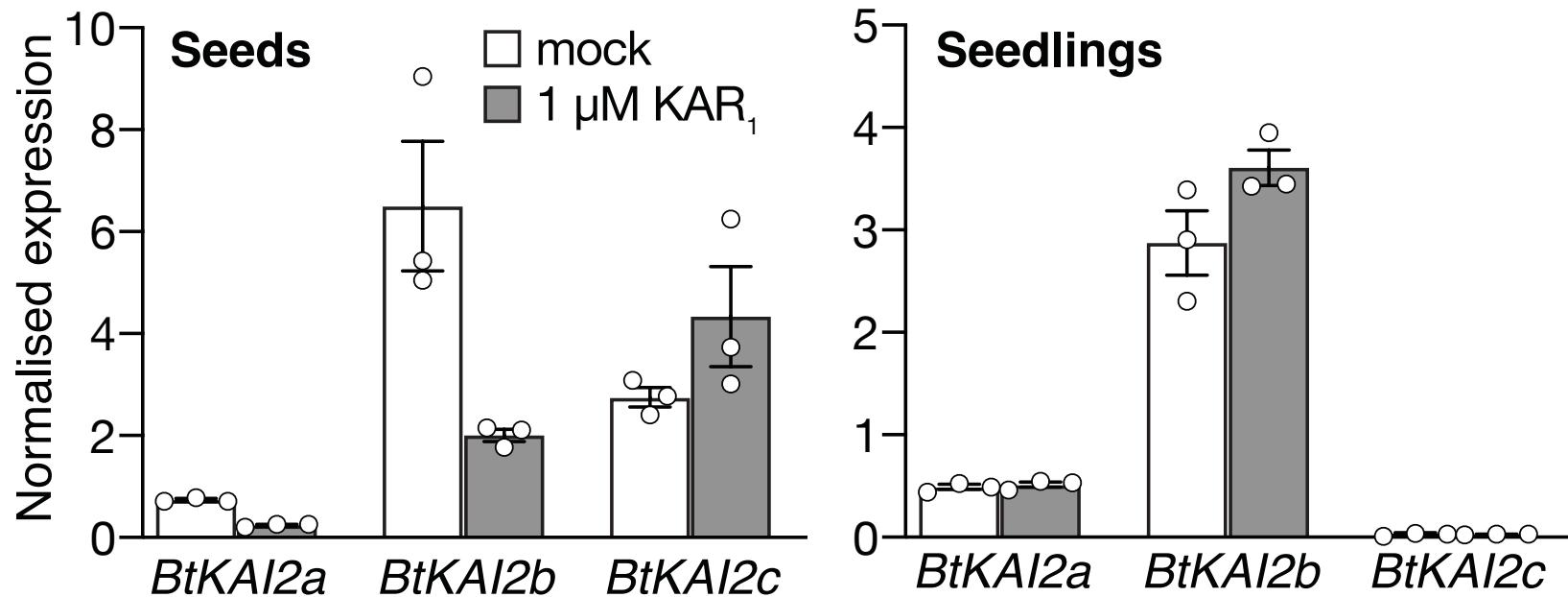
Catalytic residue
(highly conserved
functional site)



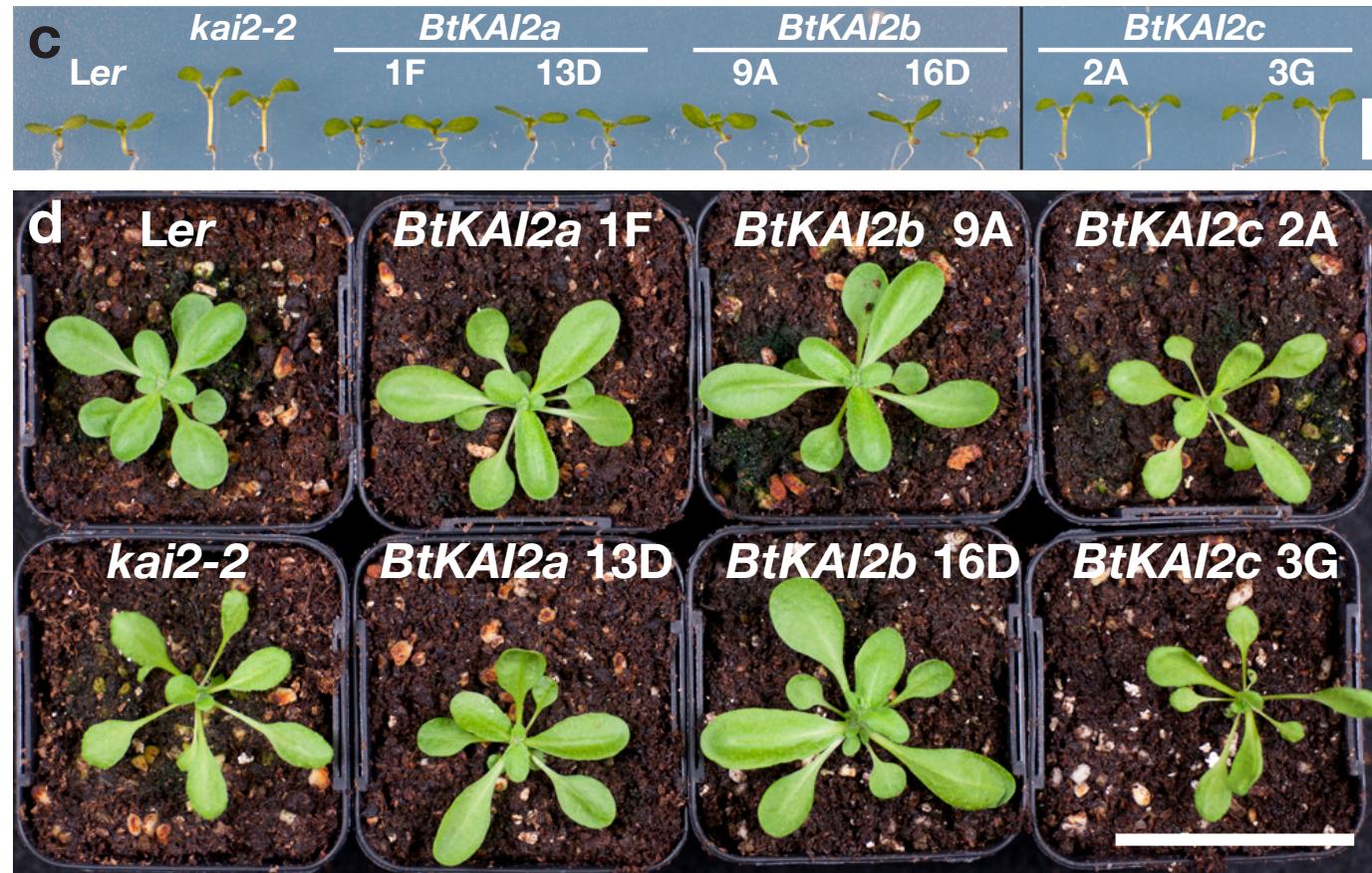
Differential homeostatic expression

Library	Total reads	Assembled reads	Contig 1 (<i>BtKAI2a</i>)	Contig 2 (<i>BtKAI2b</i>)	Contig 3 (<i>BtKAI2c</i>)
Seed	34,953,924	1838	185	1407	246
Seedlings	35,040,596	3630	913	2715	2

b

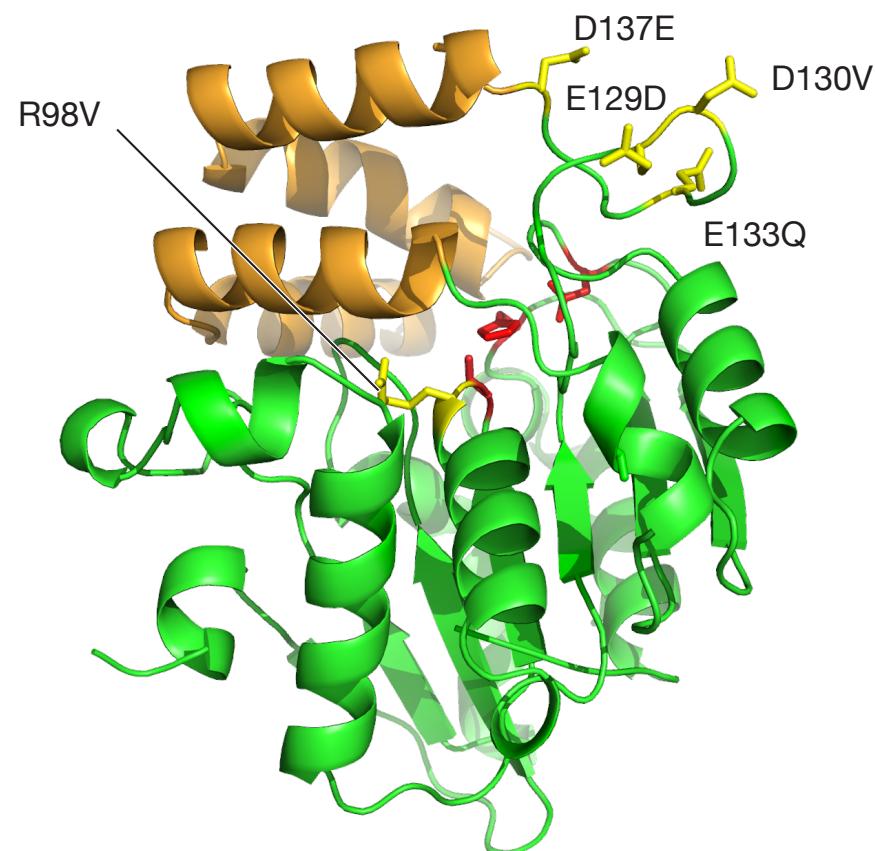


Complementation of *Arabidopsis kai2-2*

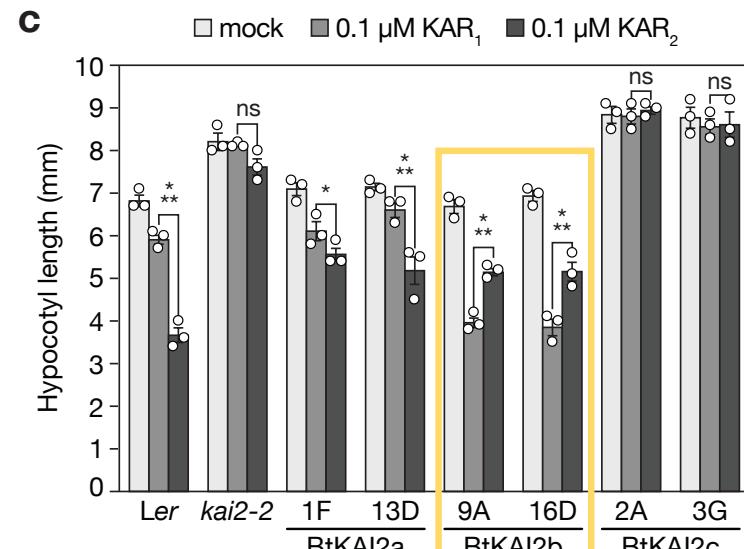


BtKAI2c carries a crucial mutation R98V

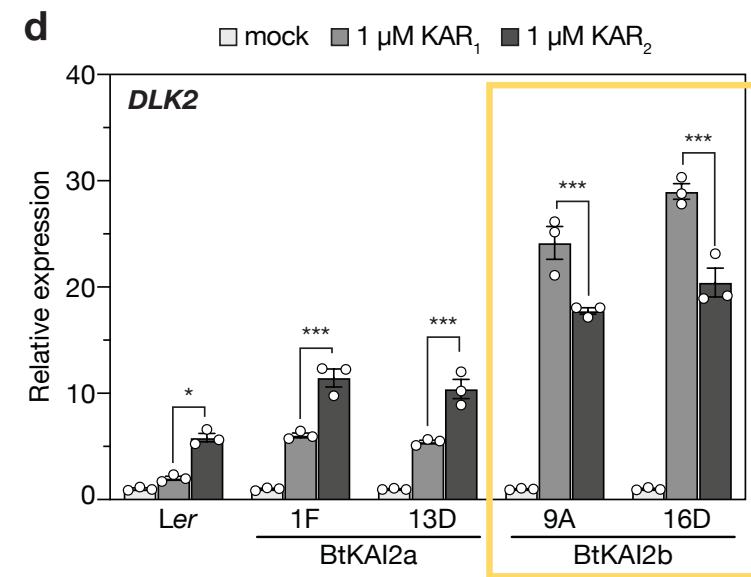
	Ser95	Ser97
	E S C I F V G H S X S A M X G X L A S L N	
AIKAI2	E S C I F V G H S V S A M I G I L A S L N	
AtKAI2	E S C I F V G H S V S A M I G V L A S L N	
BnKAI2a	E S C I F V G H S V S A M V G L L A S L N	
BoKAI2a	E S C I F V G H S V S A M V G L L A S L N	
EsKAI2	E S C I F V G H S V S A M V G V L A S L N	
BrKAI2a	E S C I F V G H S V S A M V S L L A S L N	
BtKAI2a	D S C I F V G H S V S A M V G L L A S L N	
CgKAI2	E S C I F V G H S V S A M V G V L A S L N	
CrKAI2	E S C I F V G H S V S A M V G V L A S L N	
BsKAI2	E S C I F V G H S V S A M V G V L A S L N	
BnKAI2b	E S C I F V G H S V S A M I G V L A S L N	
BoKAI2b	E S C I F V G H S L S G M V G V L A S L N	
BrKAI2b	E S C I F V G H S V S A M I G V L A S L N	
BtKAI2b	E S C I F V G H S L S A M V G V L A S L N	
BoKAI2c	E S C I F V G H S L S A M V G V L A S L N	
BrKAI2c	E S C I F V G H S L S A M V G V L A S L N	
BtKAI2c	E S C I F V G H S R S A M V G V L A S L N	
AtD14	Q N C A Y V G H S V S A M I G I I A S I R	
OsD14	P R C A F V G H S V S A M I G I L A S I R	



BtKAI2b confers preference for KAR₁

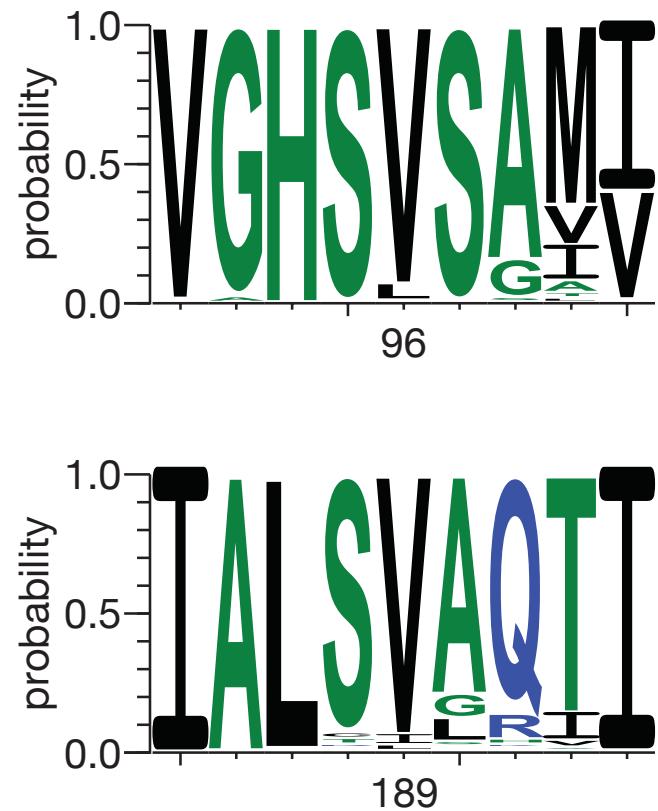


Seedling hypocotyls

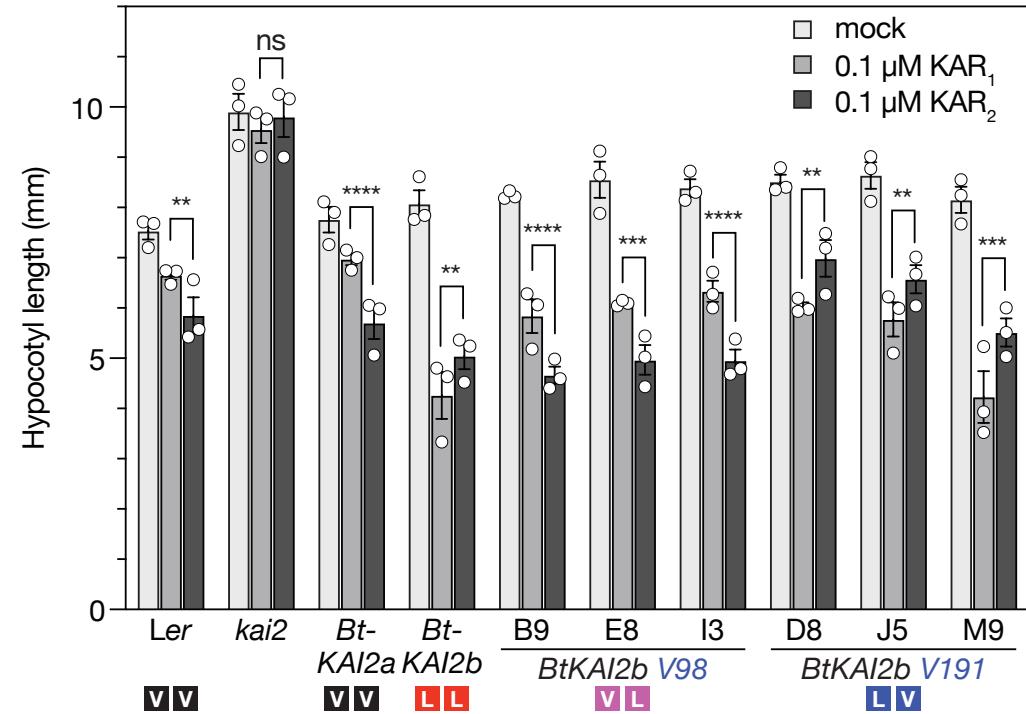


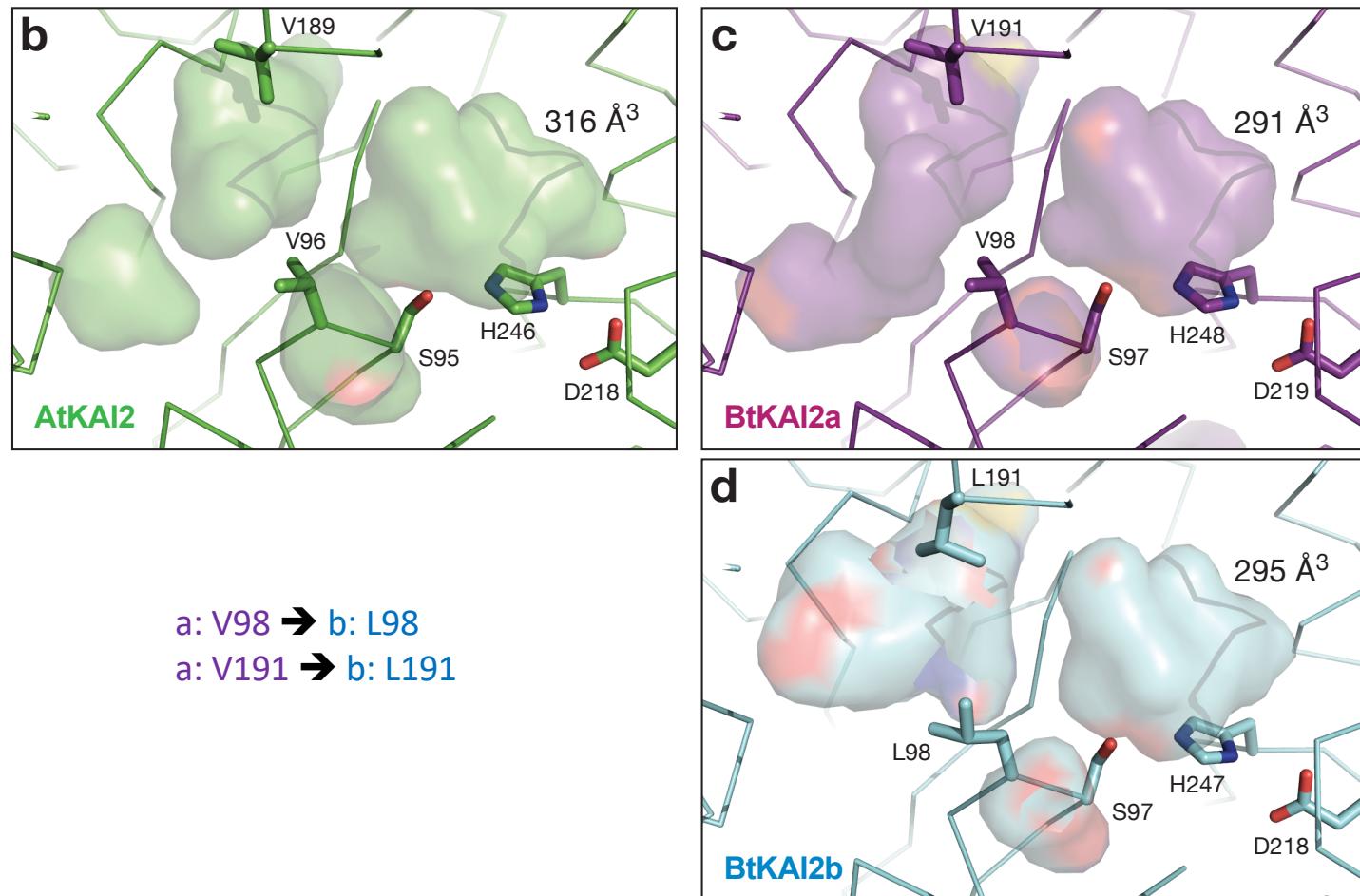
KAI2-dependent transcript *DLK2*

Position 96(98) is key for KAR₁ vs KAR₂ preference



c





Project 2

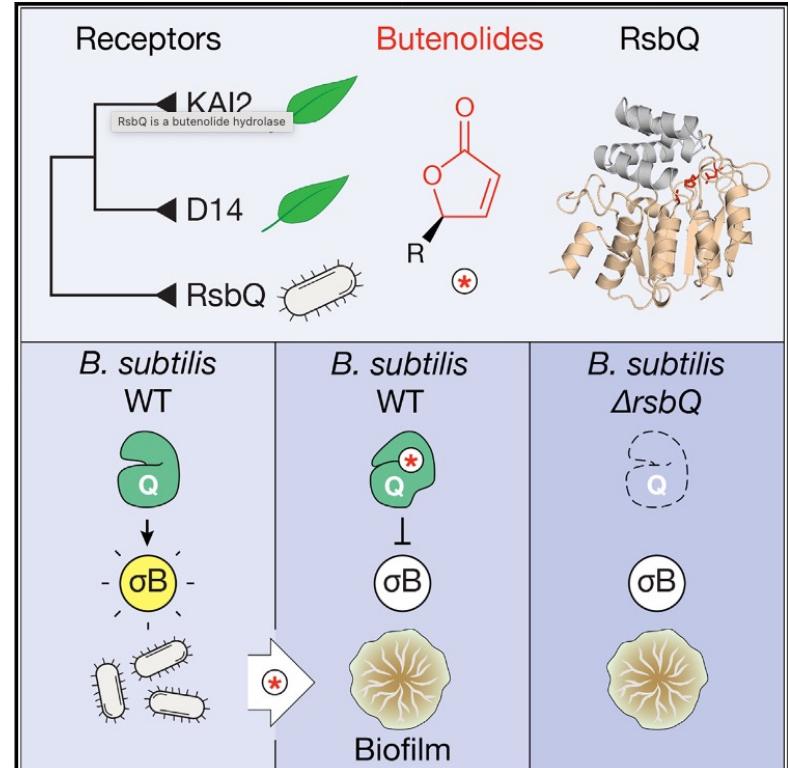
Using transcriptomics to **identify genes activated or repressed** by butenolide signalling in the bacterium *Bacillus subtilis* (Melville et al. 2024 *Current Biology* doi:[10.1016/j.cub.2023.12.035](https://doi.org/10.1016/j.cub.2023.12.035))

Current Biology

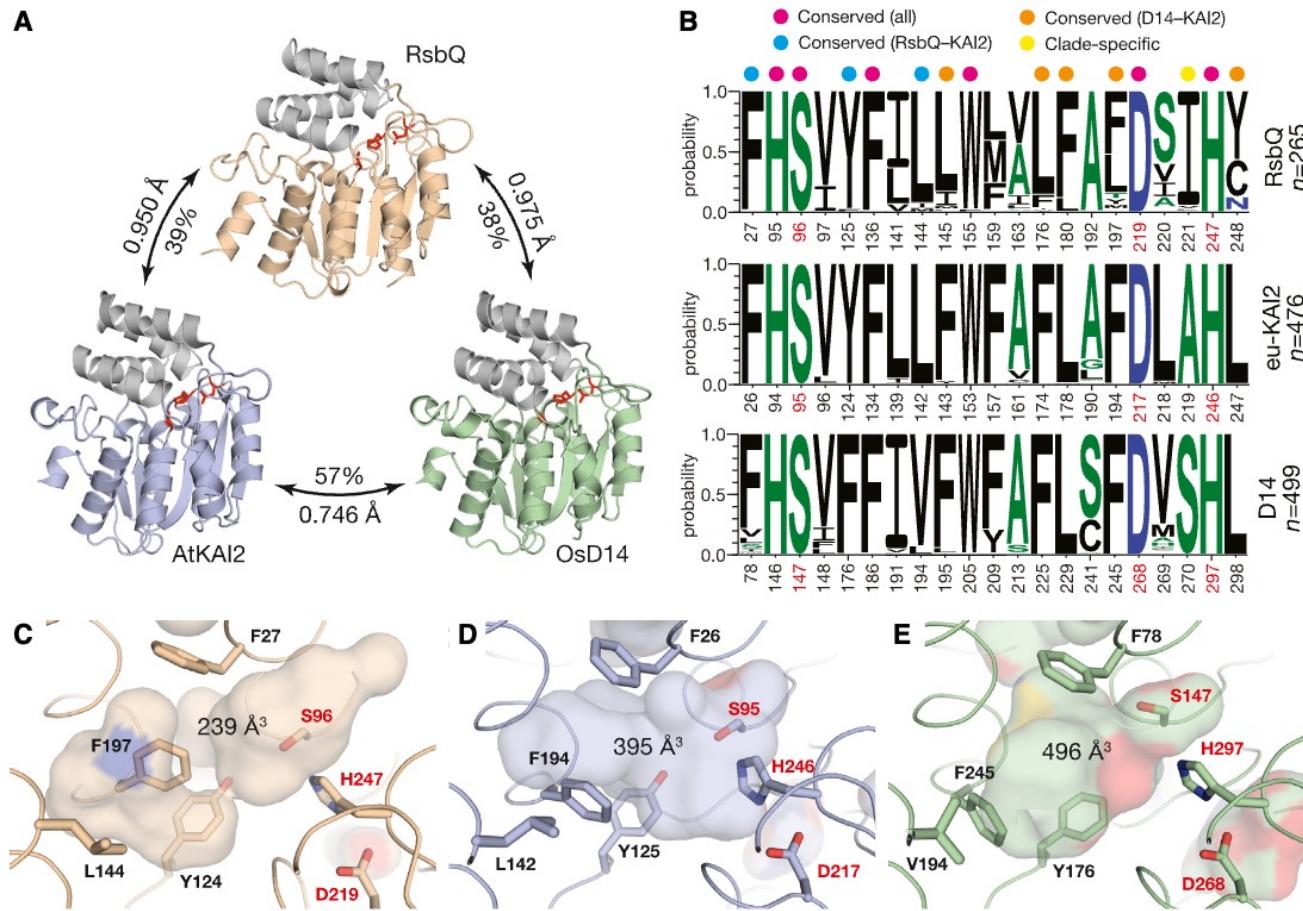
Report

Perception of butenolides by *Bacillus subtilis* via the α/β hydrolase RsbQ

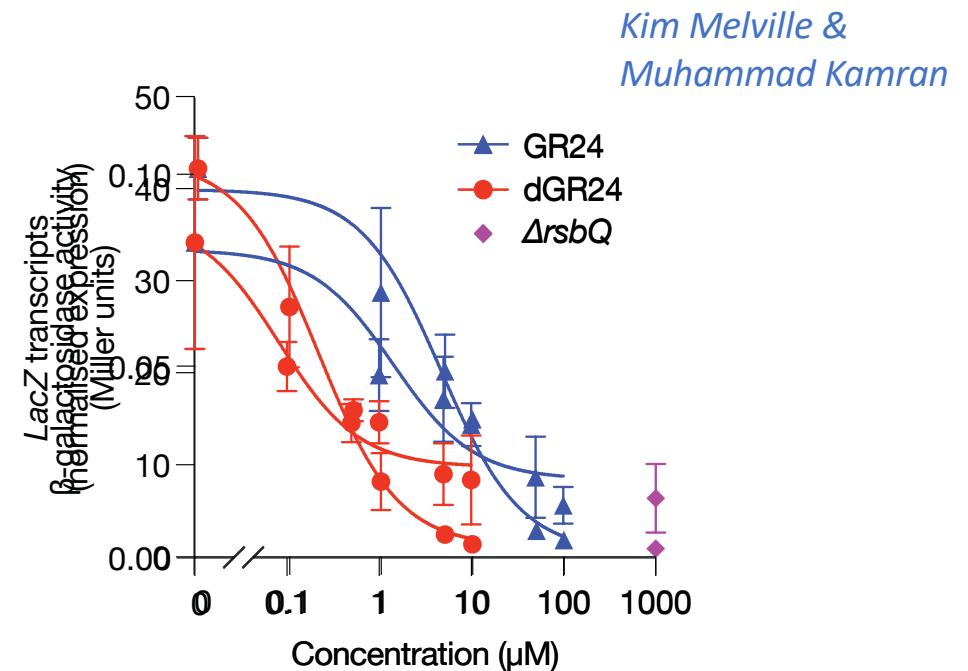
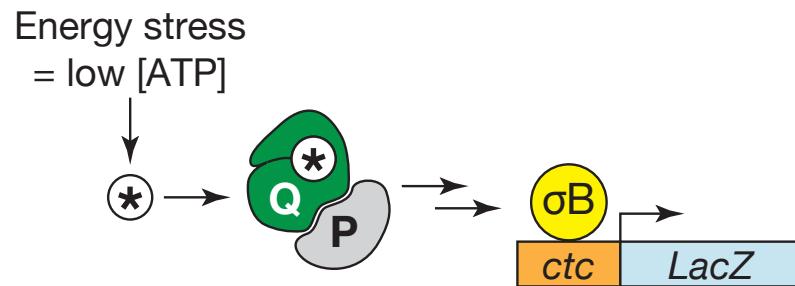
Kim T. Melville,^{1,4} Muhammad Kamran,^{1,4} Jiaren Yao,¹ Marianne Costa,¹ Madeleine Holland,¹ Nicolas L. Taylor,^{1,2} Georg Fritz,¹ Gavin R. Flematti,¹ and Mark T. Waters^{1,3,5,*}



RsbQ is a bacterial homologue of KAI2 & D14 from plants



What happens when you treat *B. subtilis* with butenolides?

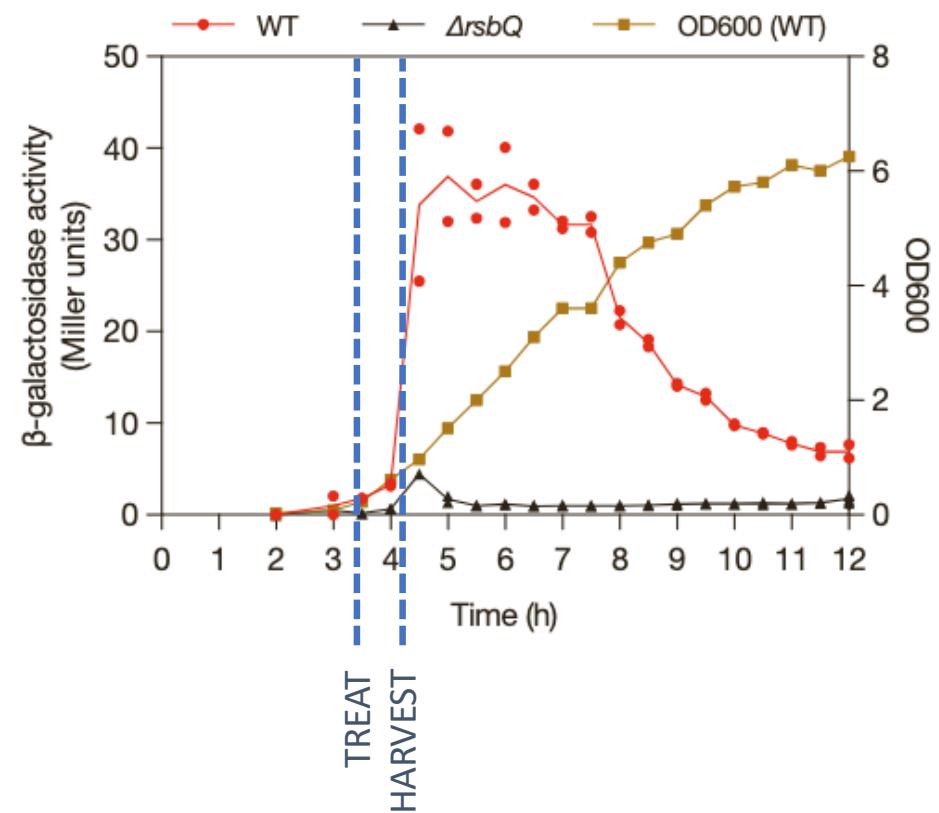


⇒ dGR24 is more potent than GR24, but both INHIBIT RsbQ activity!

What happens to gene expression in *B. subtilis* treated with butenolides?

Experimental design:

- Two treatments ($\pm 10 \mu\text{M}$ dGR24)
- Three genotypes (WT, $\Delta rsbQ$, $rsbQ^{S96A}$) – why?
- Three biological replicates – why?
- Total = 18 samples

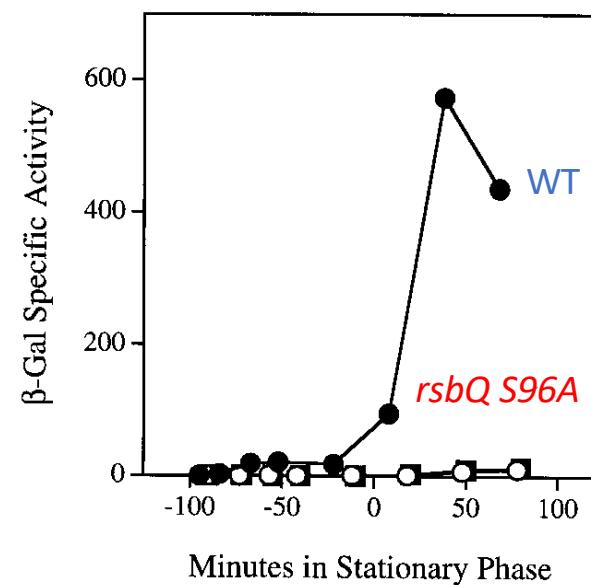
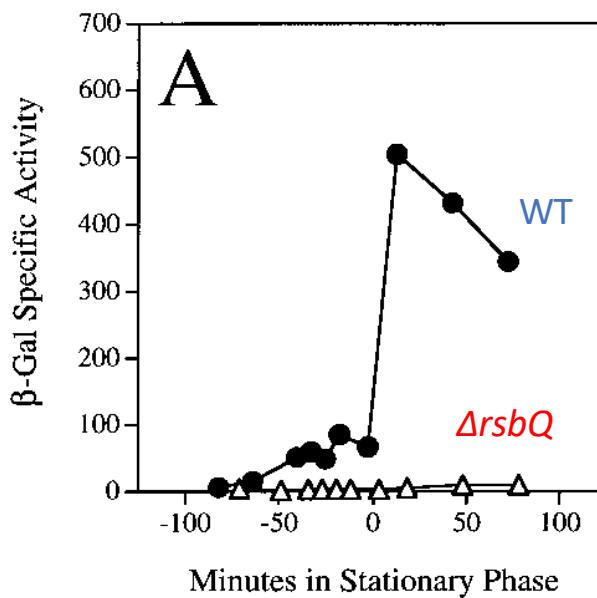


Why use two *rsbQ* mutants?

$\Delta rsbQ$ is a knockout mutant (complete loss of protein)

$rsbQ^{S96A}$ is a point mutant (missense mutant) targeting active site Ser

They are phenotypically the same – or are they?



a) **Redundancy:** If genes are dependent on RsbQ function, they should be misregulated in BOTH mutants = more powerful conclusions

b) **Information about protein function:** if one mutation is phenotypically more severe than the other

Experimental details

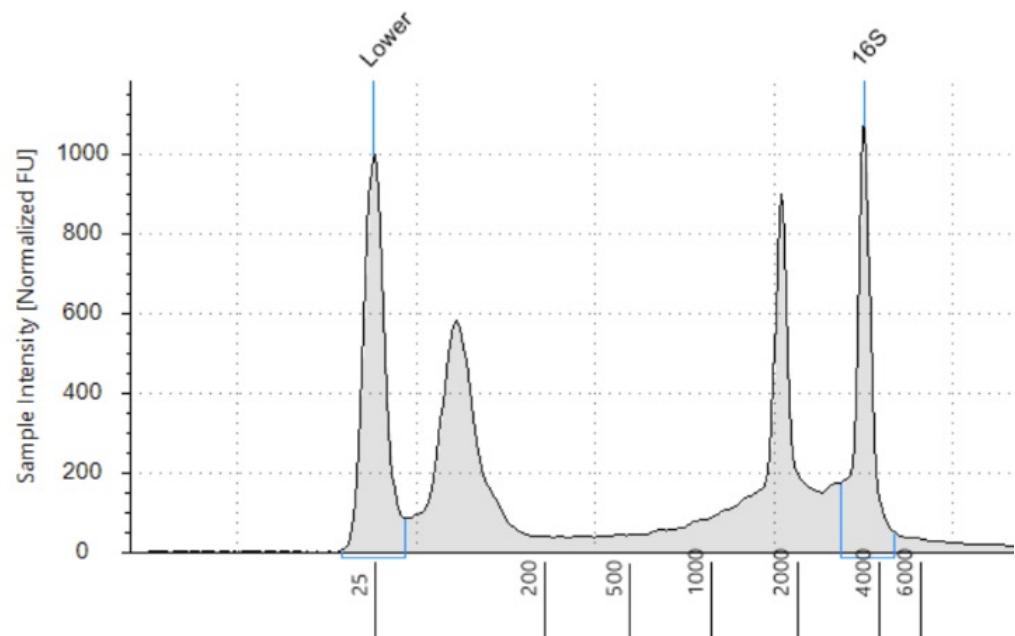
DNase treatment – why?

RNA integrity assessment – RIN
> 7 – why?

Ribodepletion workflow –
removal of rRNA – why?

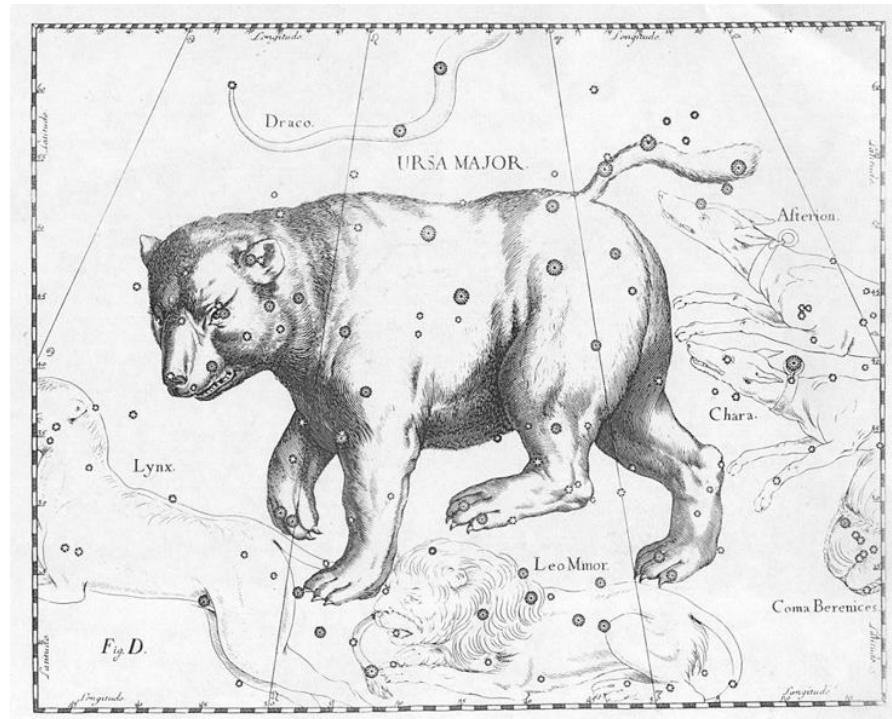
Paired-end libraries – why?

~10m reads per sample



RNAseq workflow

Read counts using **kallisto** (fast; uses *pseudoalignment* “which focuses only on identifying the transcripts from *which the reads could have originated* and does not try to pinpoint exactly how the sequences of the reads and transcripts align”; [link](#))



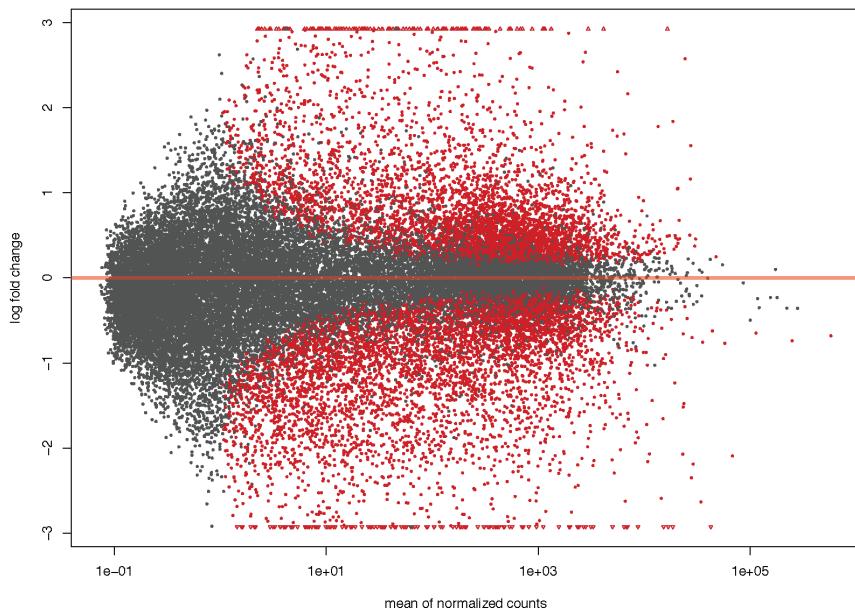
RNAseq workflow

Differentially expressed gene
(DEG) identification using DEseq2

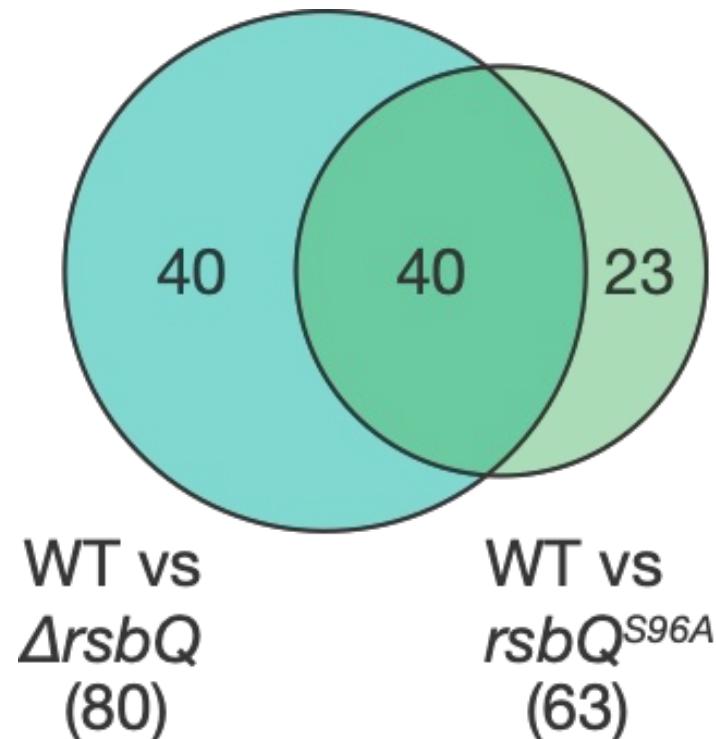
Gene selection criteria:

FDR-adjusted P-value <0.05

$\log_2 FC >1$ or < 1



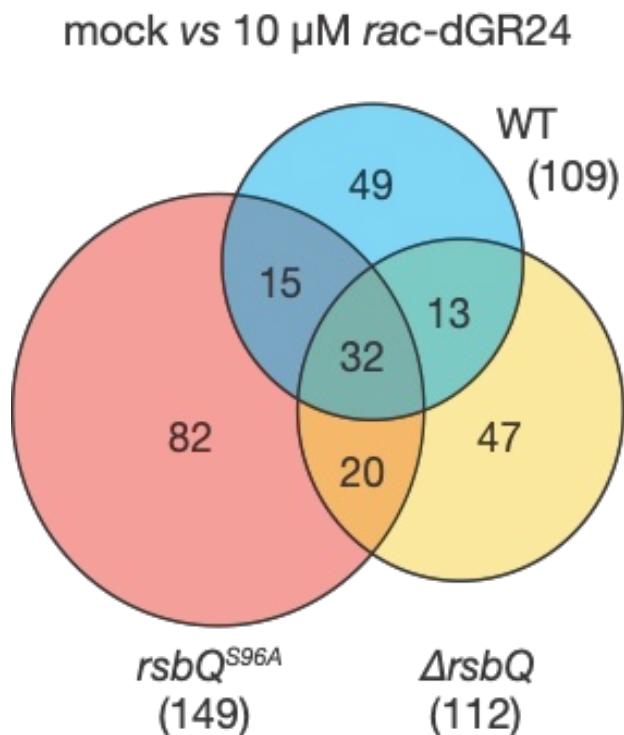
Some results



Just comparing UNTREATED bacteria:

- Only a subset of genes are differentially expressed, relative to WT in BOTH mutants
- That is, 80 genes were mis-expressed in $\Delta rsbQ$, and 63 in $rsbQ^{S96A}$, but only 40 were common

Some results

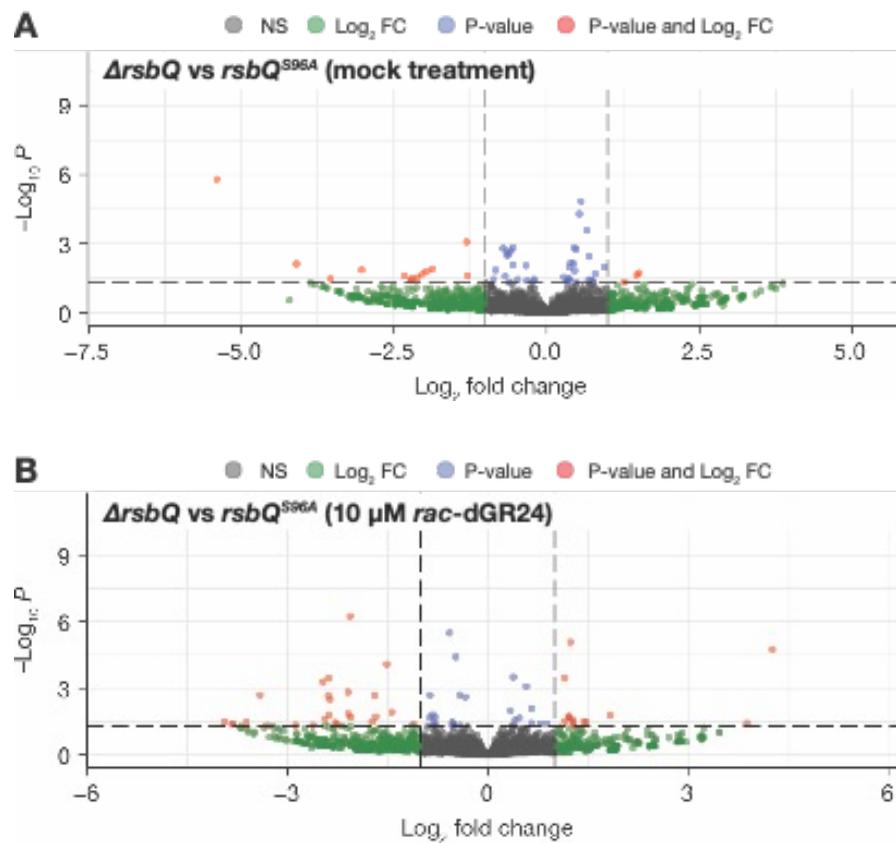


Comparing all three genotypes when treated with dGR24:

Mutants still show a transcriptional response, but different from each other

- How do we identify genes that are likely RsbQ-dependent?
- How do we identify genes that are likely RsbQ-independent?

Visualising the data with volcano plots



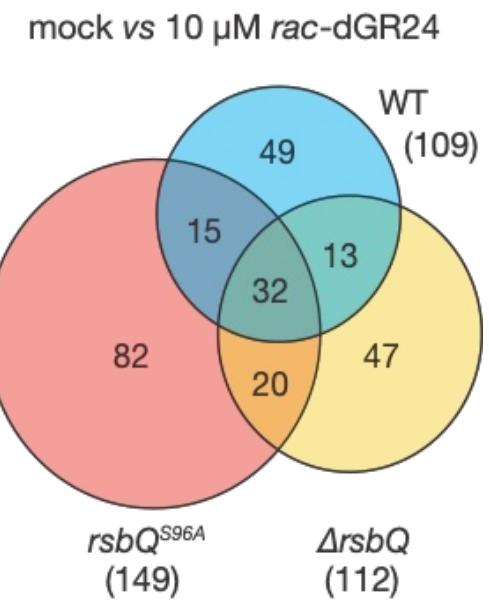
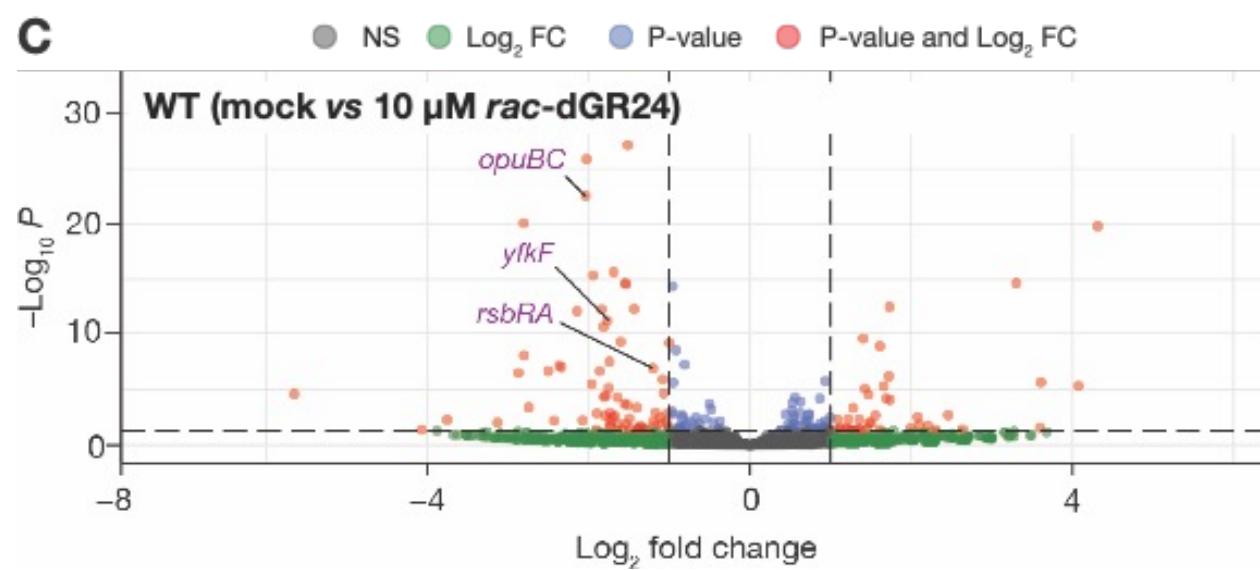
These two plots compare the two mutants

Red dots indicate genes that are differentially expressed

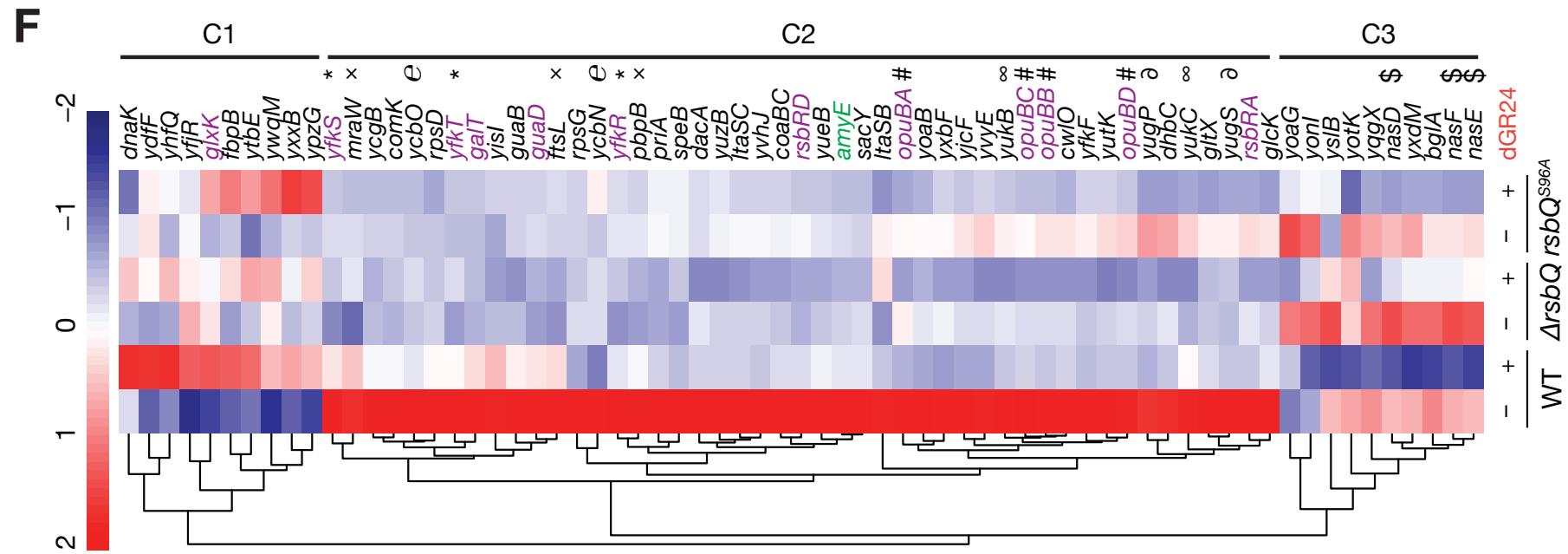
Evidence that mutants are not identical

More DEGs under treatment (B) and no treatment (A) suggesting that the two mutants differ in their response and possibly one mutant retains some function

The blue circle as a volcano plot



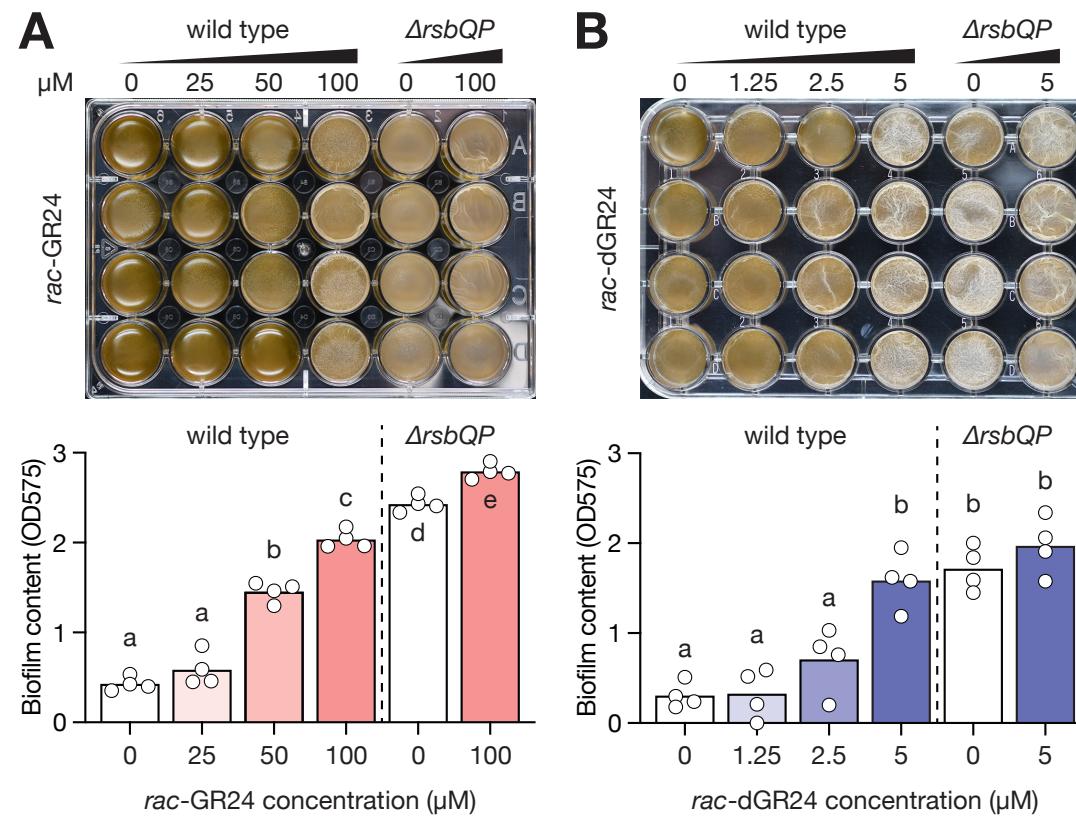
Hierarchical clustering to identify core RsbQ- dependent genes



Conclusion: looking at C2, which is enriched for sigB-dependent transcripts (purple), treatment with dGR24 phenocopies (recapitulates/is similar to) loss-of-function RsbQ mutations

Therefore: at a transcriptome scale, treatment with dGR24 inhibits RsbQ function

Back to biology: butenolides promote biofilm formation via RsbQ



Summary

Discussed two examples of transcriptomics:

- One for gene discovery in non-model organism
- Another for identifying downstream effectors of a signaling pathway in bacteria

Both have a focus on a biological question used for hypothesis testing

Other applications/further analysis could include:

- Measuring/quantifying drug responses;
- Identifying alternatively-spliced transcripts
- Single cell transcriptomics – assaying cell-to-cell heterogeneity in gene expression