

Soybean Breeding: Where Digital Phenotyping and Genomics Meet in the Field

Johann Vollmann

BOKU University
Dept. of Agricultural Sciences
Institute of Crop Breeding and Genomics
Konrad Lorenz Str. 24
3430 Tulln an der Donau / Austria

johann.vollmann@boku.ac.at

Legume Generation

RISE-SOY

Contents



- **Soybean breeding**
- Adaptation to Europe
- Stem termination
- Digital phenotyping
- How does genomics support soybean breeding ?
- Conclusions





























Austrian soy products (examples)



Soybean seeds of different cultivars









Legume Generation



Soybean root nodules

Basics of soybean breeding and genetics

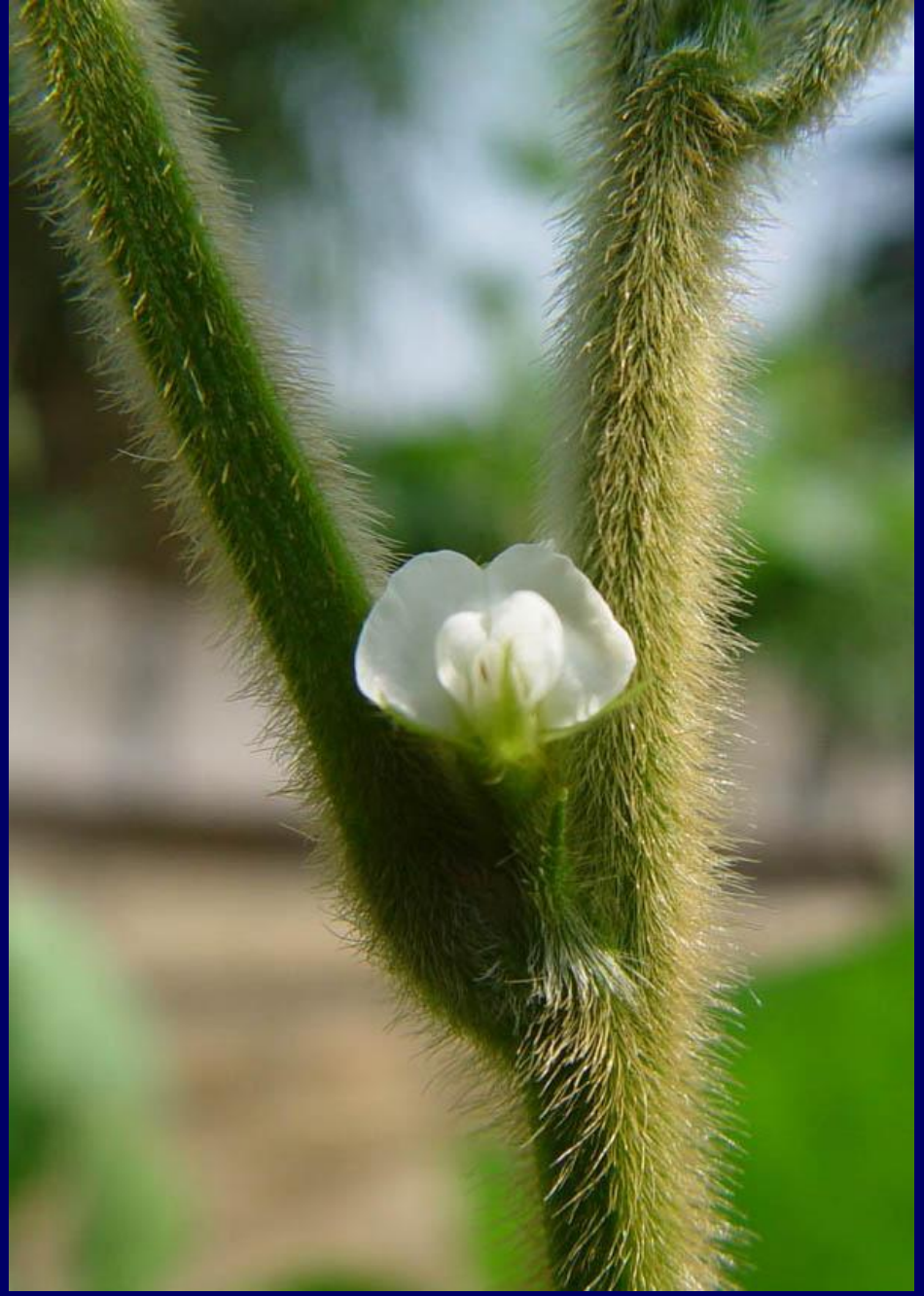
Soybean: ***Glycine max*** (L.) Merr., Fabaceae
Domestication from wild *G. soja* 3500 years ago in China
Reproductive biology: **self-pollination** (>99%)
Linkage groups (chromosomes): **2n = 40**
Genome sequenced: 1.115 Mb genome size (1.1 gigabases)
46.430 protein coding genes, genome is paleopolyploid
(i.e. 75% genes in multiple copies) (Schmutz et al. 2010, Nature 463:178-183)
From genome sequence data: Over 33.000 SSR markers available for marker-assisted selection (Song et al. 2010, Crop Sci. 50:1950-1960)

Commercial breeding: SNP-panels for genomic selection, SSR markers etc.

Markers



purple / grey



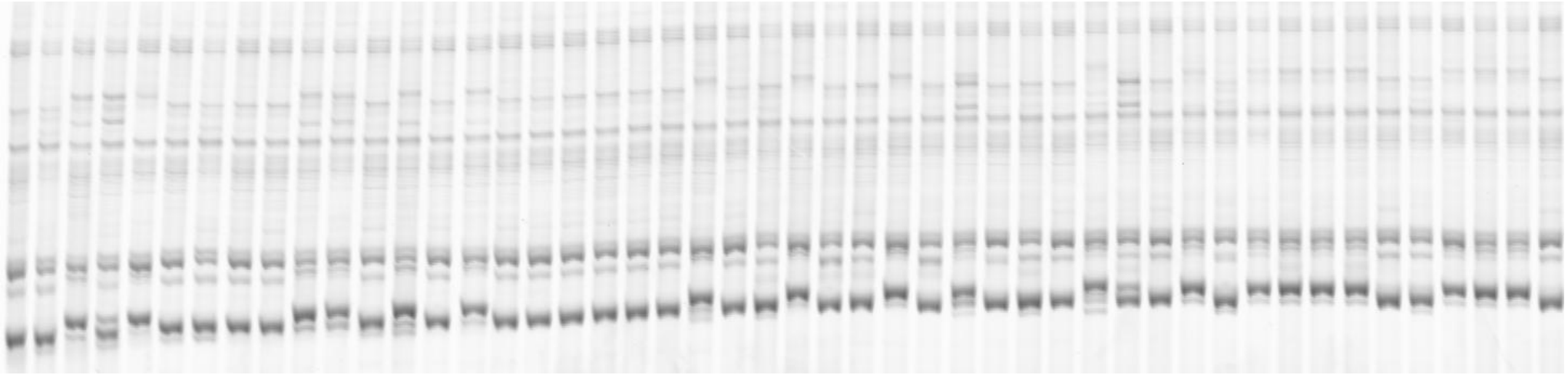
white / tawny

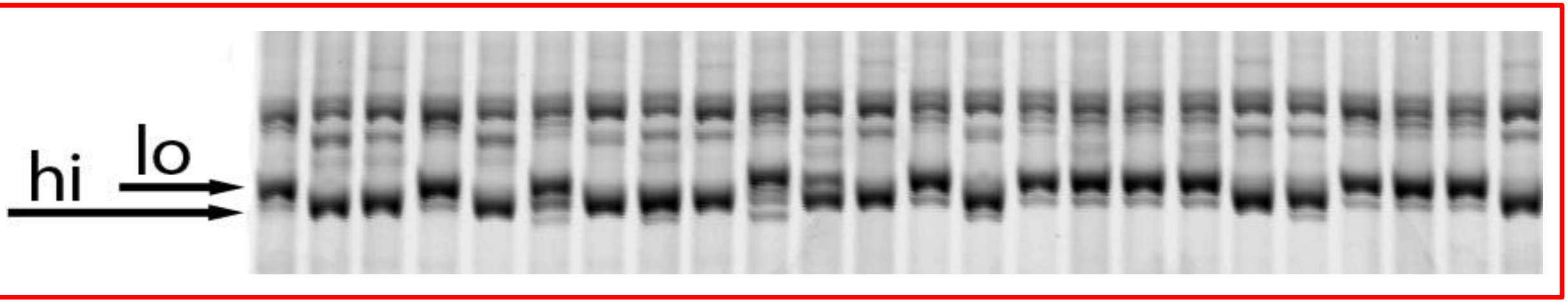


F_2-segregation in flower color

flower color (*W1/w1*) visible on hypocotyl – pleiotropic effect

Markers





SSR marker SacK149 associated with either low (lo) or high (hi) **cadmium** (Cd) accumulation in soybean seed

Contents



- Soybean breeding
- **Adaptation to Europe**
- Stem termination
- Digital phenotyping
- How does genomics support soybean breeding ?
- Conclusions

Soybean variety classification

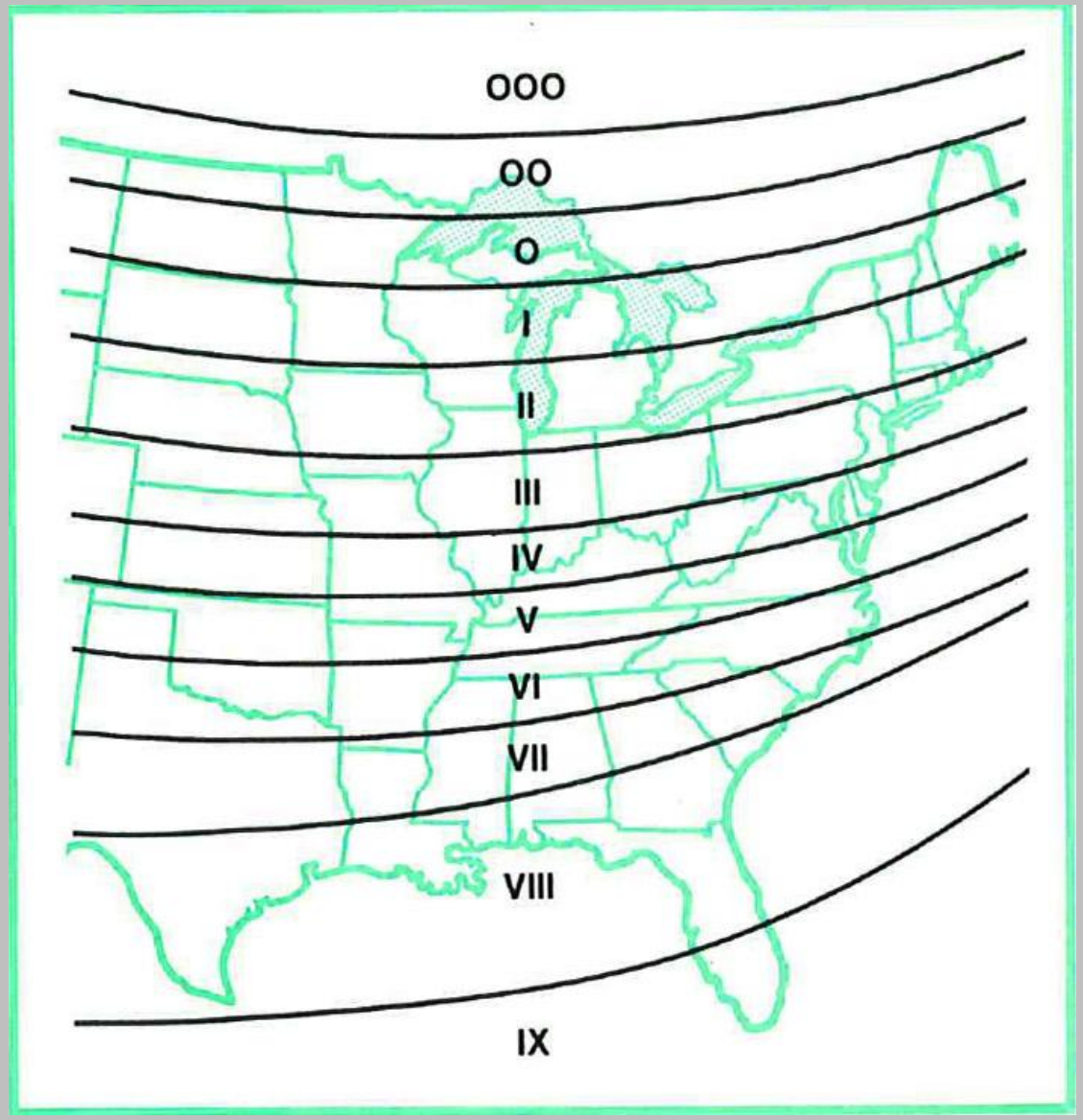
Maturity groups:

0000: very early
000: early
00: mid early
0: mid late

I - X: later MGs not suitable for central Europe

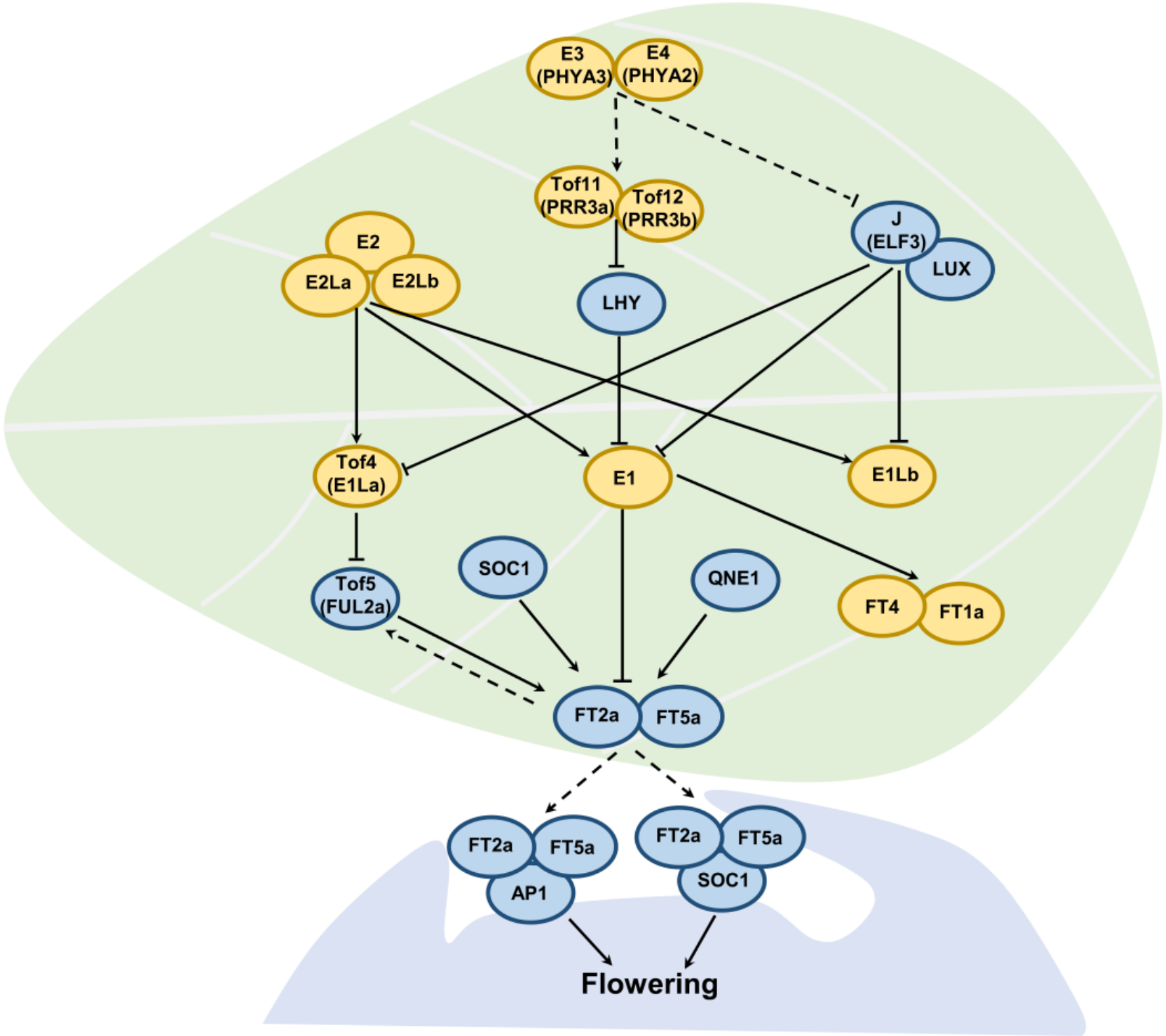
Maturity groups: Specific reaction to daylenght / photoperiod controlling the flowering and maturity process. Due to maturity groups, each soybean variety is adapted to a specific geographic latitude.

Genetics: *E*-genes for daylength sensitivity

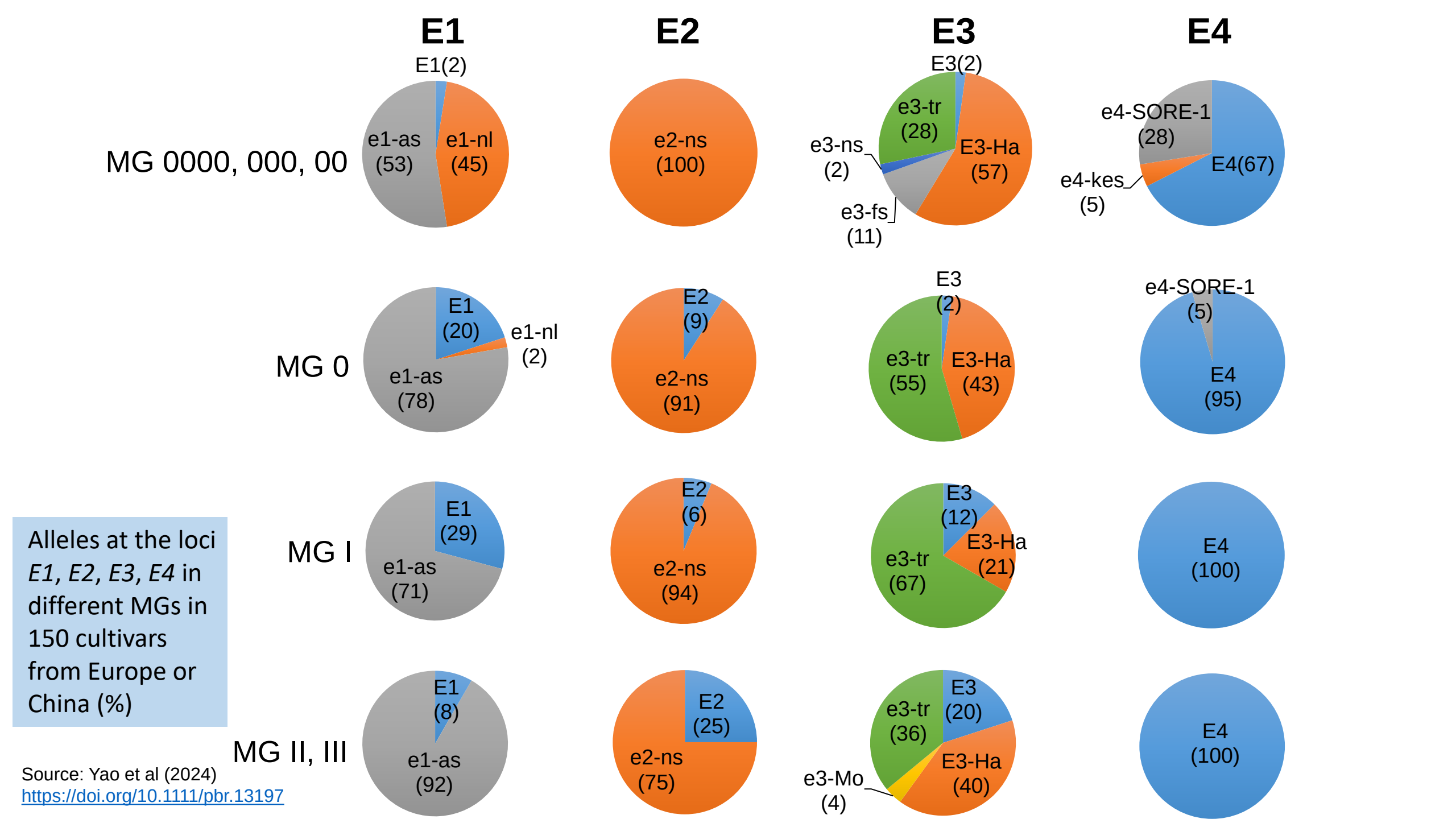


Source: Scott & Aldrich, 1983

Fig. 1 Photoperiodic flowering regulatory mechanisms in soybean. Yellow oval represents flowering repressors; blue oval indicates flowering promoters. The solid and dotted lines represent direct and indirect regulation, respectively. Arrow means activation, while T-shape symbol means suppression

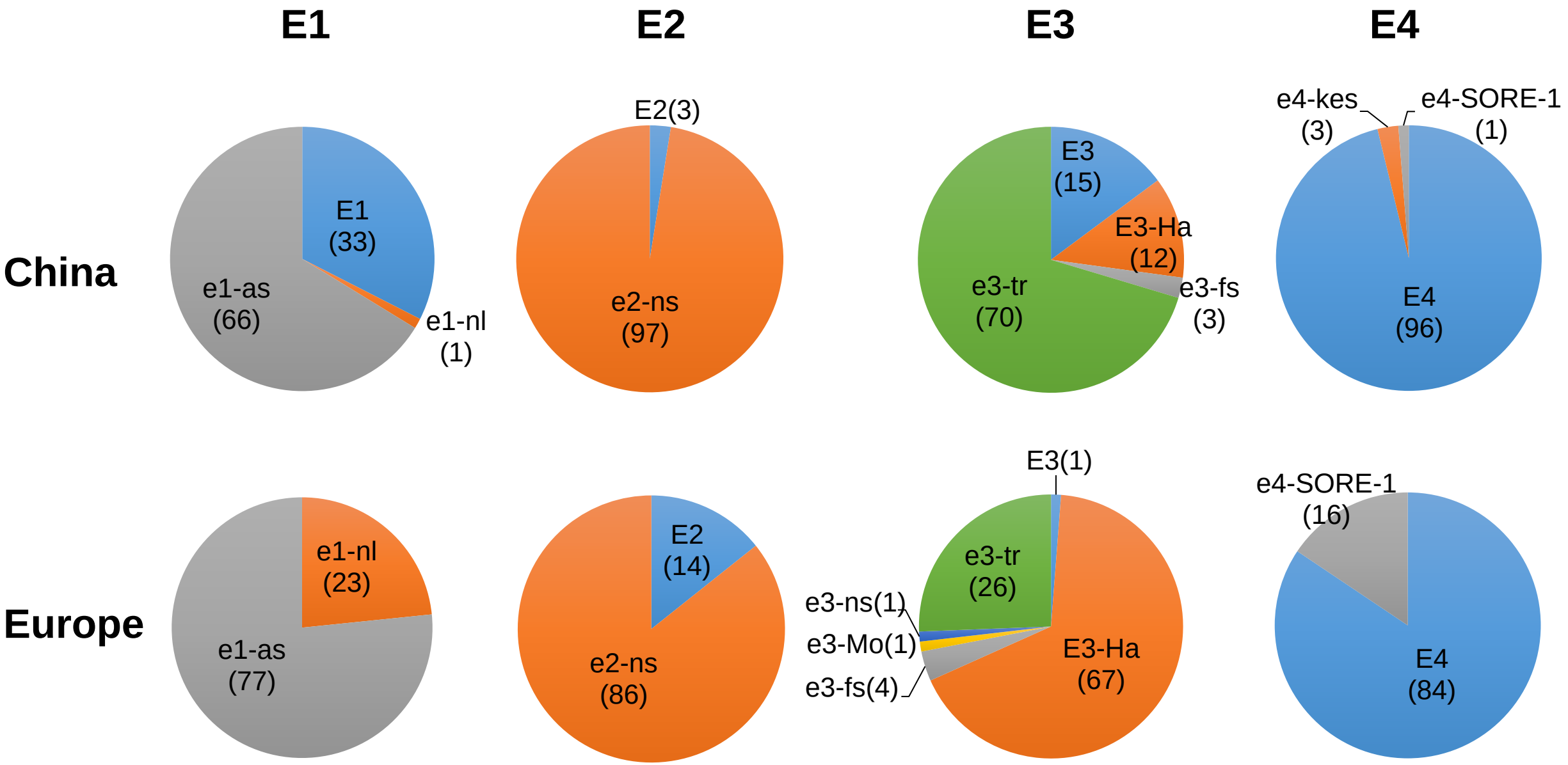


Source:
Hou et al, 2023,
Molec Breeding 43:36



Alleles at the loci *E1, E2, E3, E4* in different MGs in 150 cultivars from Europe or China (%)

Source: Yao et al (2024)
https://doi.org/10.1111/pbr.13197



Distribution of alleles at the loci *E1*, *E2*, *E3*, and *E4* in 156 Chinese and European elite cultivars across the two regions of origin (% cultivars per pie chart section in parentheses)

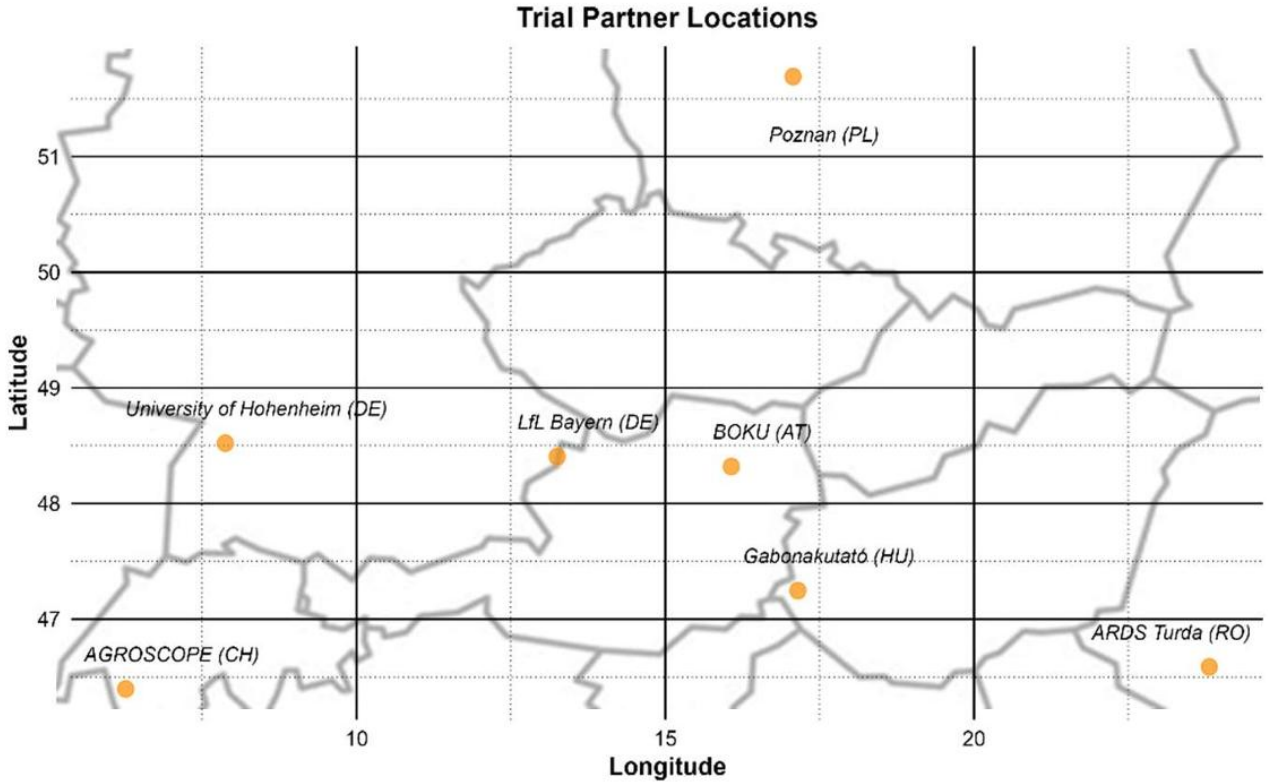
Source: Yao et al (2024)
https://doi.org/10.1111/pbr.13197

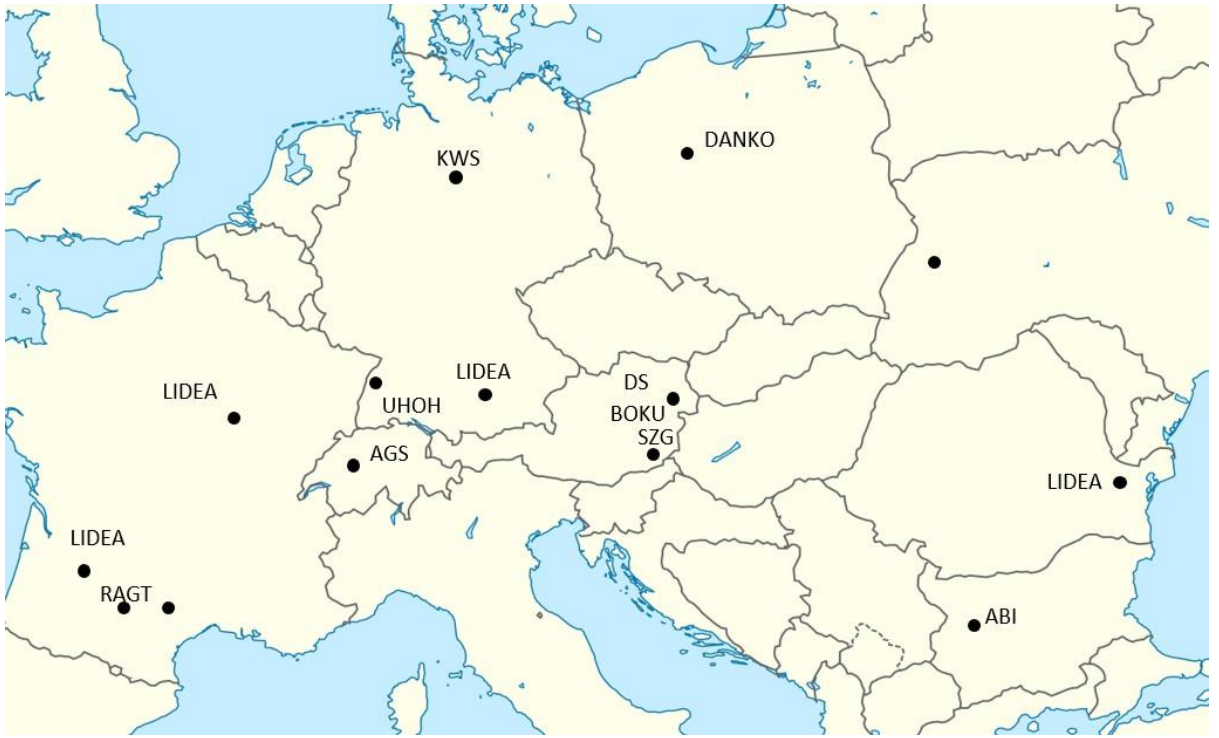
Breeding for yield and adaptation





Differences in maturity between genotypes and experiments
(Legume Generation, Tulln, Austria, 2024)



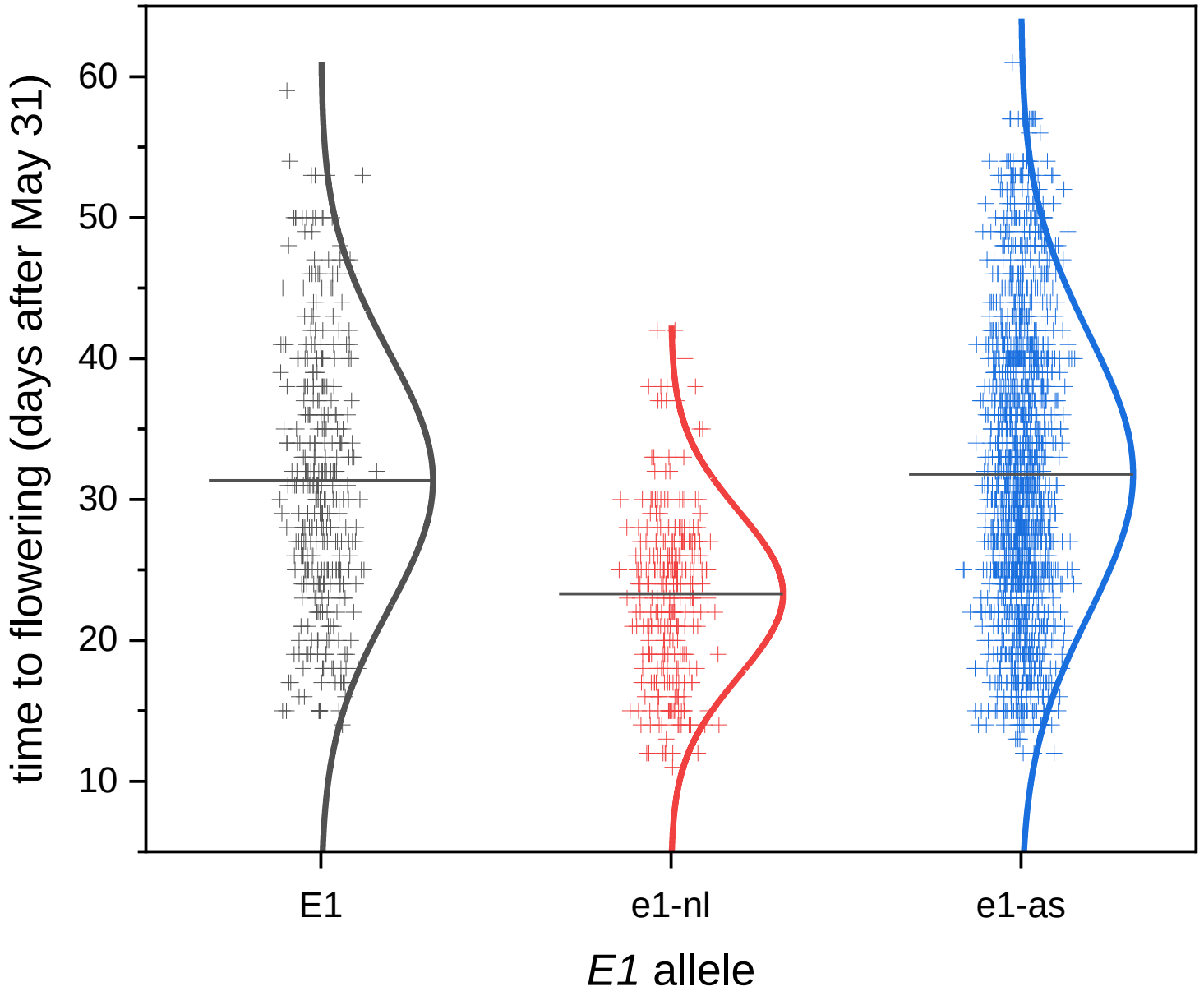


Project **HABERLANDT**

Project **LEGUME GENERATION**

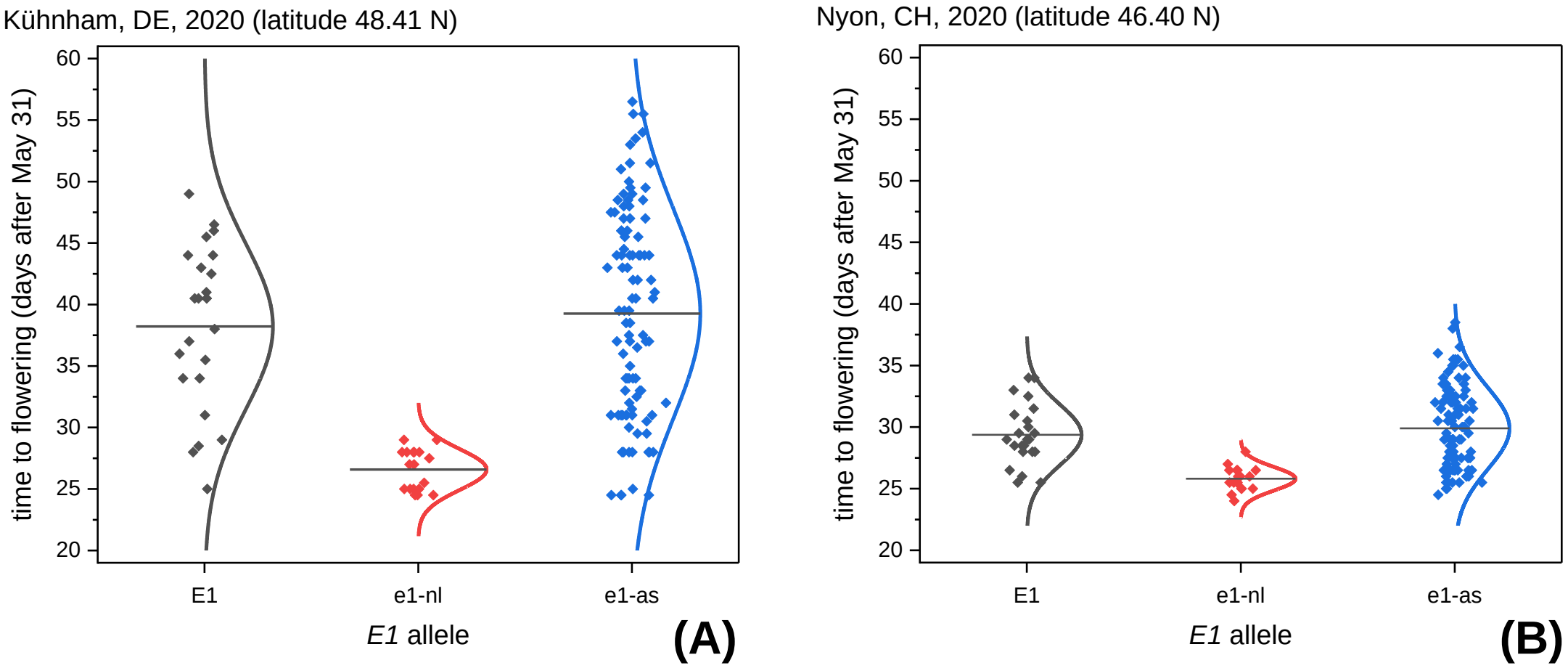
Legume Generation

Funded by the European Union



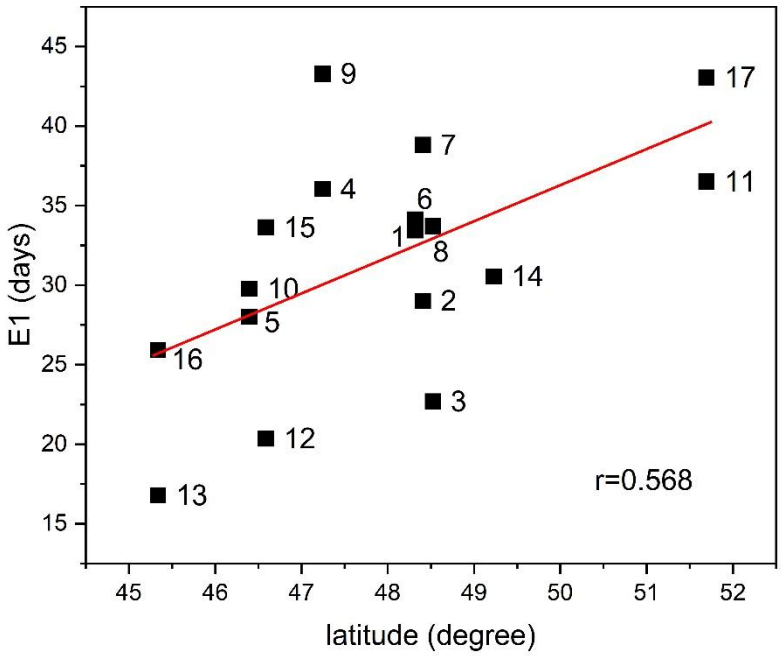
Time to flowering across all cultivars and 17 environments as affected by the *E1* allele

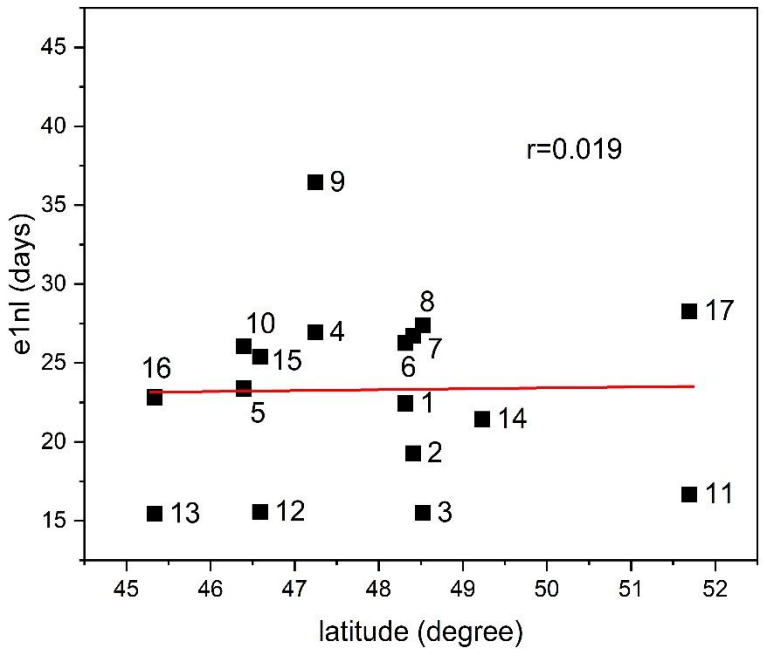
Source: Yao et al (2024)
https://doi.org/10.1111/pbr.13197

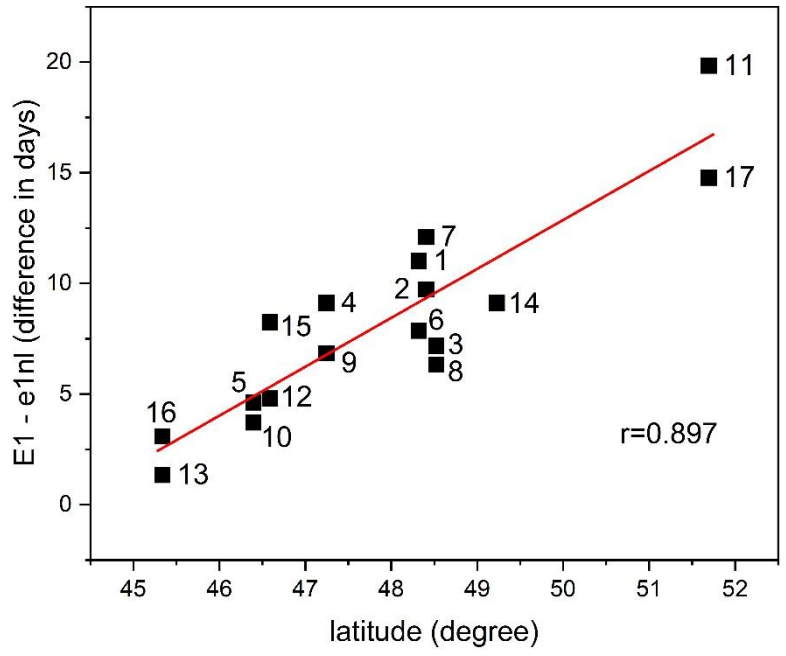


Distribution of time to flowering across all cultivars as affected by the E1 allele at a higher (A) and a lower (B) latitude environment

Source: Yao et al (2024)
https://doi.org/10.1111/pbr.13197

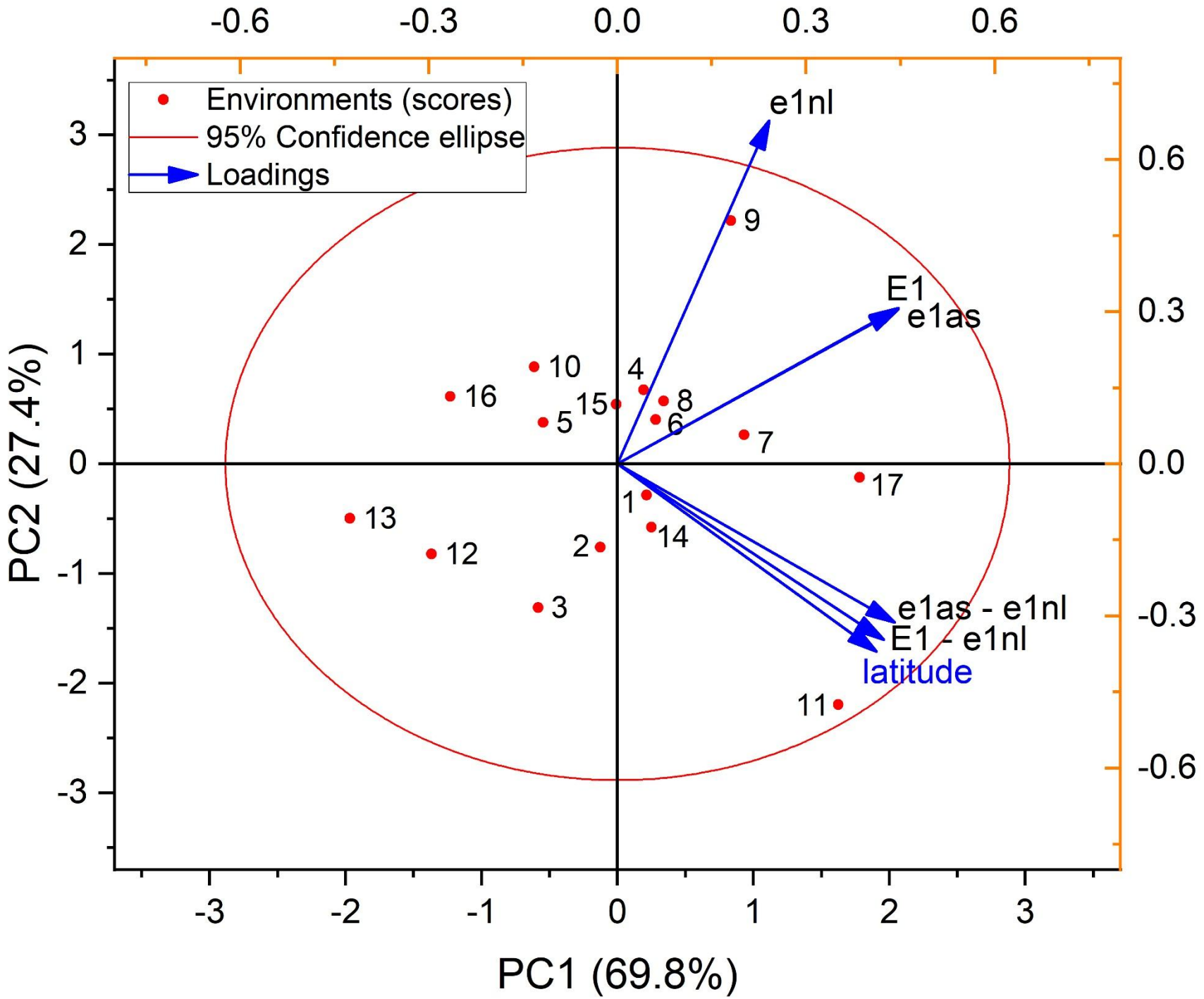






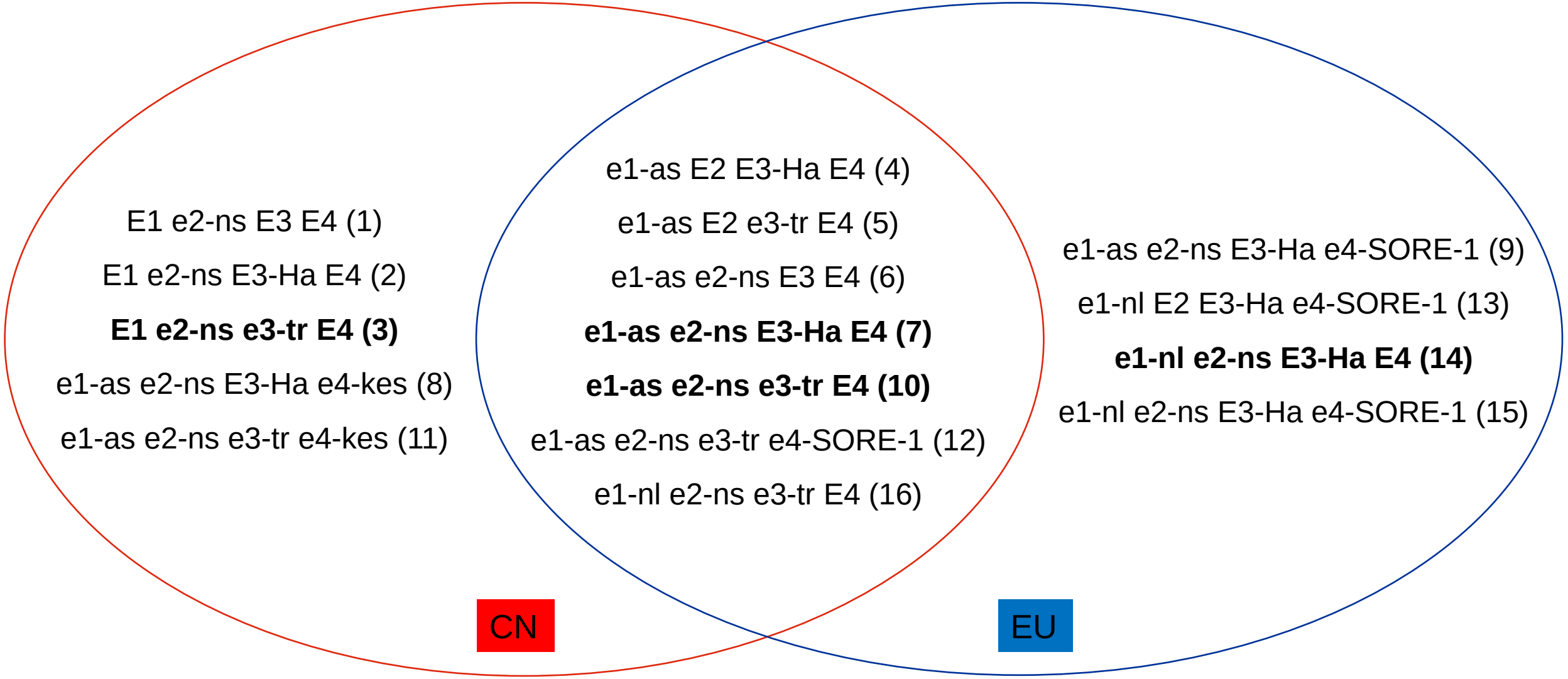
Correlation between latitude and time to flowering for *E1* (A) and *e1-nl* (B) alleles, and for the difference between *E1* and *e1-nl* alleles (C) over 17 environments

Source: Yao et al (2024)
https://doi.org/10.1111/pbr.13197



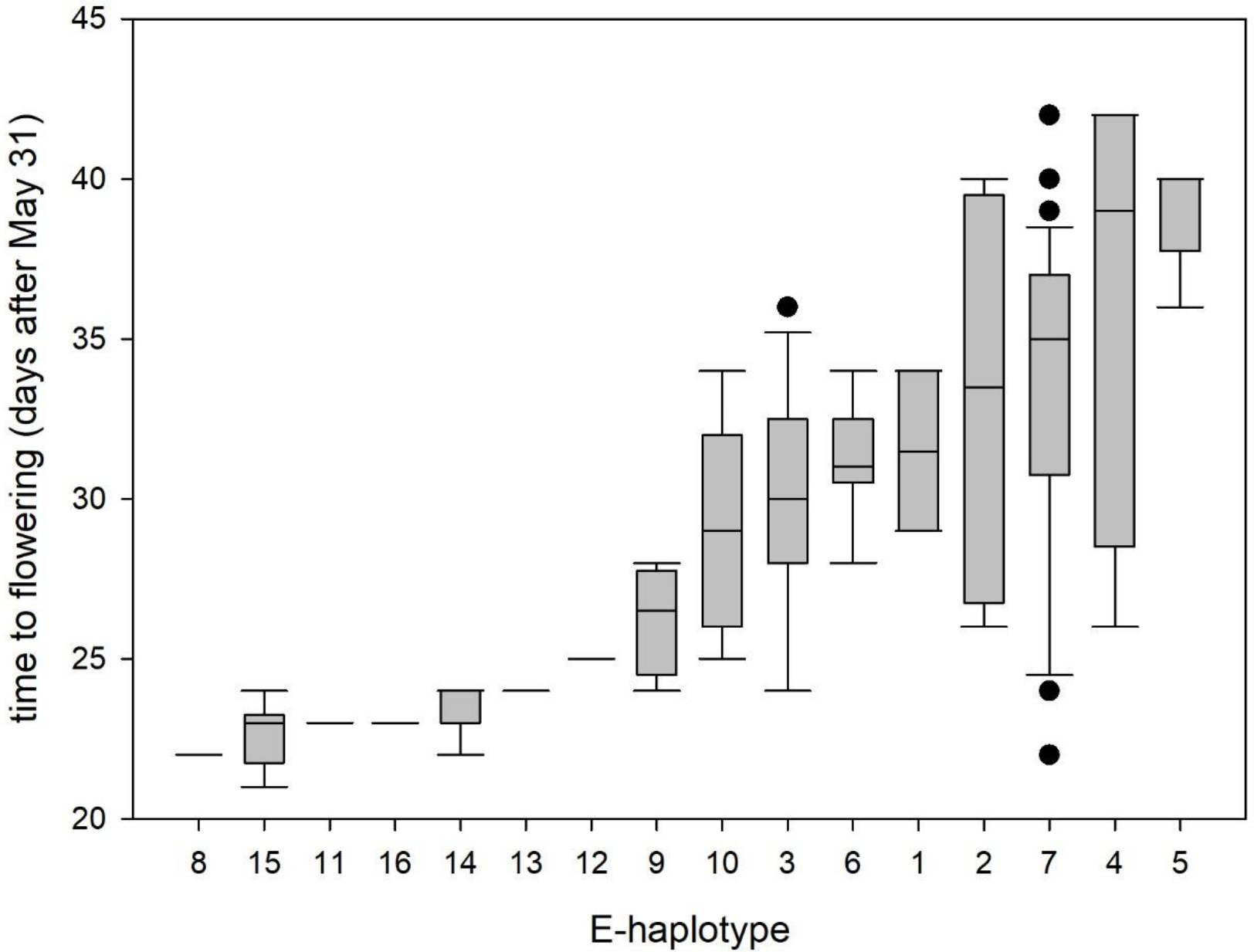
PCA bi-plot of *E1* alleles and latitude for time to flowering in 17 environments

Source: Yao et al (2024)
https://doi.org/10.1111/pbr.13197



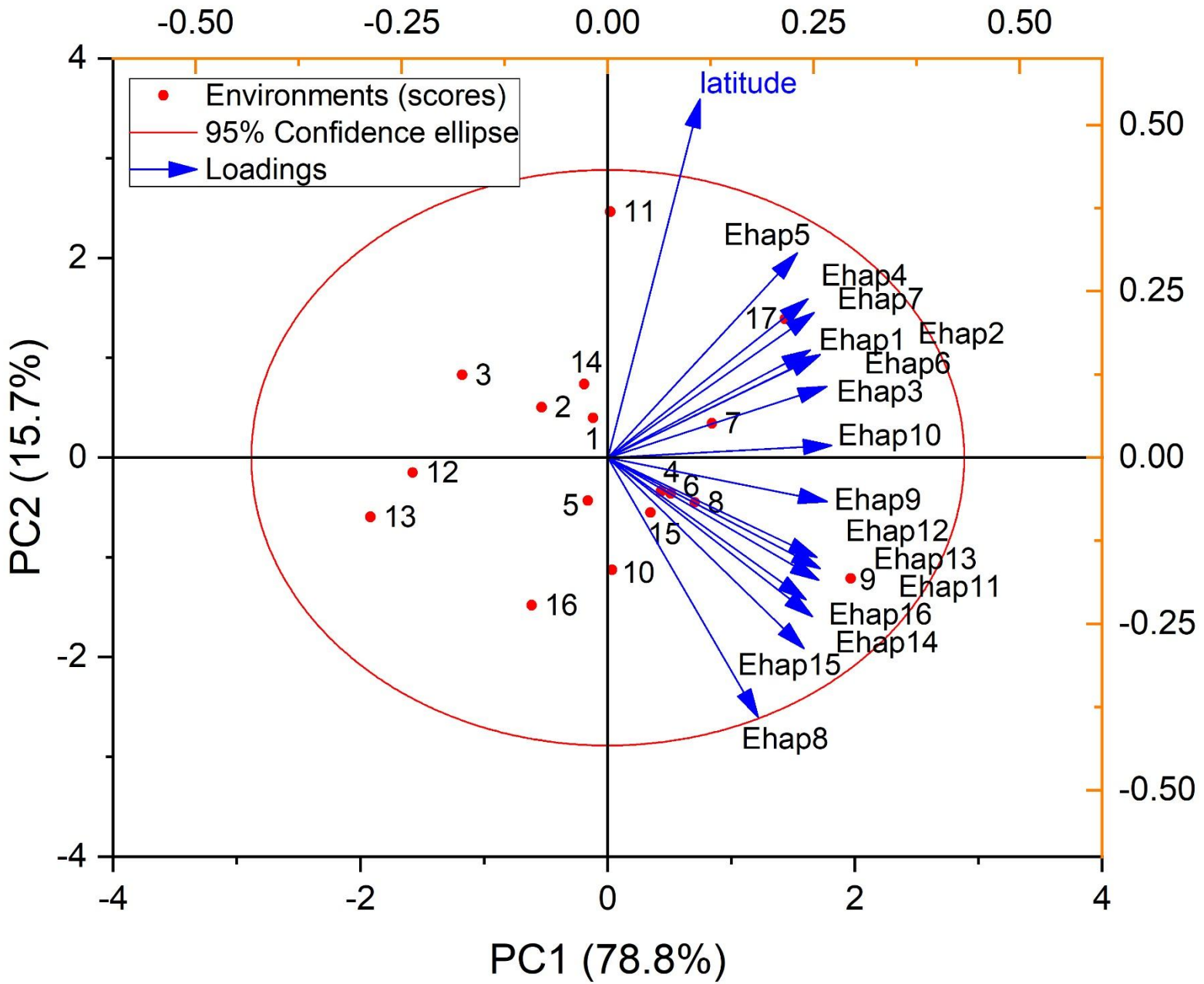
Distribution of *E* gene alleles and *E*-haplotypes in Chinese and European cultivars (haplotype no. in parenthesis; most frequent haplotypes in bold)

Source: Yao et al (2024)
https://doi.org/10.1111/pbr.13197



Time to flowering of soybean cultivars as affected by the *E*-haplotype (genotype mean data across all 17 environments).

Source: Yao et al (2024)
https://doi.org/10.1111/pbr.13197



PCA bi-plot of 16 *E*-haplotypes and latitude for time to flowering in 17 environments

Source: Yao et al (2024)
https://doi.org/10.1111/pbr.13197



Example of missing adaptation

Tulln, 2 Oct 2023

Contents



- Soybean breeding
- Adaptation to Europe
- **Stem termination**
- Digital phenotyping
- How does genomics support soybean breeding ?
- Conclusions

Yield formation

Plant architecture - flower initiation - pod formation - grain yield

Dt2: semi-determinate stem growth (shoot apical meristem generates flowers terminating stem growth)



dt2: indeterminate stem growth (shoot apical meristem remains vegetative continuously)



dt2 vs. *Dt2*
stem determination

Dt2 (“semi-determined” stem growth, shortes, pods on terminal raceme, better pod-set, shorter indernodes, stronger stem, less lodging) is dominant over dt2 (“in-determined” stem growth)




Symptoms
of drought


Severe drought

Dt2 Is a Gain-of-Function MADS-Domain Factor Gene That Specifies Semideterminacy in Soybean[C][W]

Jieqing Ping,[a,1] Yunfeng Liu,[a,1] Lianjun Sun,[a,1] Meixia Zhao,[a] Yinghui Li,[b] Maoyun She,[a] Yi Sui,[a,2] Feng Lin,[a] Xiaodong Liu,[a] Zongxiang Tang,[a] Hanh Nguyen,[c] Zhixi Tian,[a,3] Lijuan Qiu,[b] Randall L. Nelson,[d] Thomas E. Clemente,[c] James E. Specht,[c] and Jianxin Ma[a,4]

[a] Department of Agronomy, Purdue University, West Lafayette, Indiana 47907
[b] Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081, China
[c] Department of Agronomy and Horticulture/Center for Plant Science Innovation, University of Nebraska, Lincoln, Nebraska 68583
[d] Soybean/Maize Germplasm, Pathology, and Genetics Research Unit, U.S. Department of Agriculture–Agricultural Research Service, Department of Crop Sciences, University of Illinois, Urbana, Illinois 61801

Plant Cell, 2014, 26, 2831-2842.

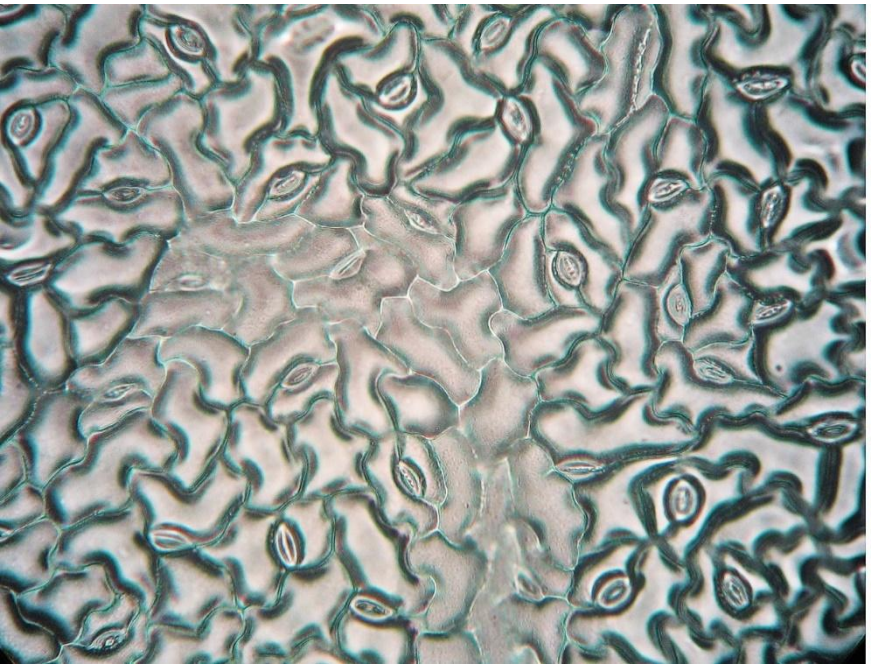
CellPress
PARTNER JOURNAL

Molecular Plant
Research Article

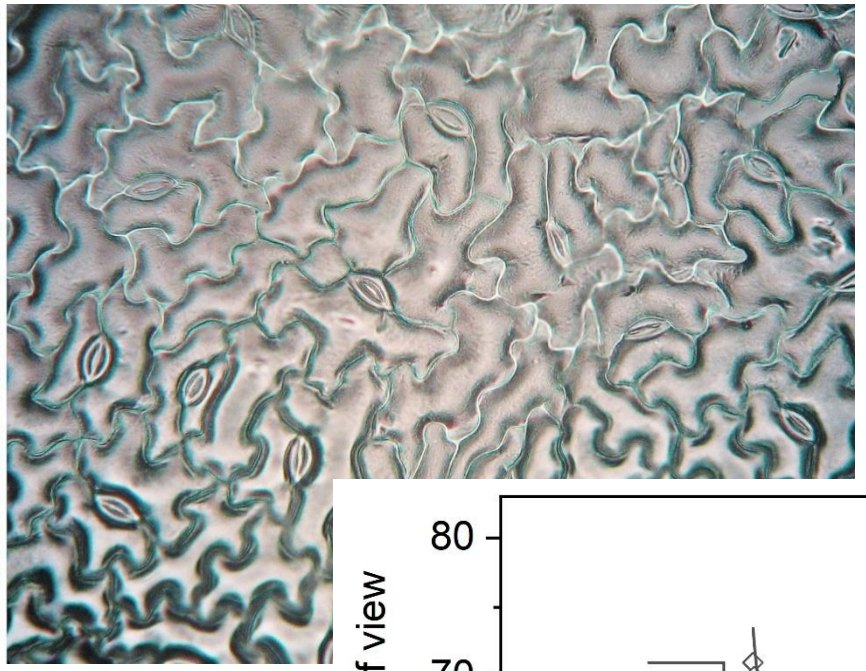
A Post-domestication Mutation, *Dt2*, Triggers Systemic Modification of Divergent and Convergent Pathways Modulating Multiple Agronomic Traits in Soybean

Dajian Zhang[1,2,6], Xutong Wang[1,6], Shuo Li[1,3,6], Chaofan Wang[2], Michael J. Gosney[4], Michael V. Mickelbart[4,5] and Jianxin Ma[1,5,*]

Mol. Plant 2019, 12, 1366-1386.



Dt2



dt2

Stomata density counts

(stomata per 0.142 mm^2)

Data from S. Galla (2021)

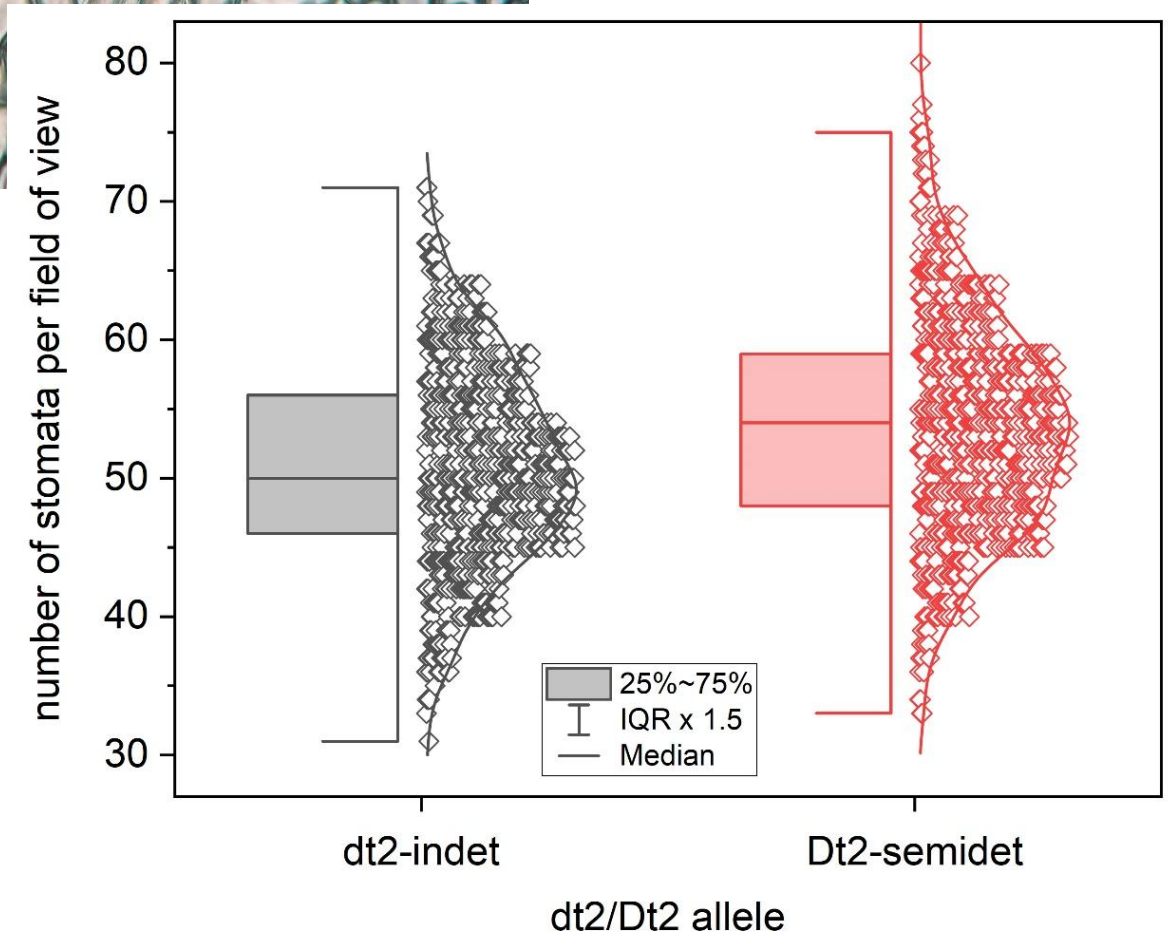


Table: Soybean genotypes sorted according to seed total mean $\delta^{13}C$

No.	Name	Breeder	2024 mean	2025 mean	Total mean seed $\delta^{13}C$	Tukey multiple means comparison
42	UHOH-58	University of Hohenheim	-26,47	-26,59	-26,53	a
37	UHOH-53	University of Hohenheim	-26,32	-26,79	-26,55	a
43	UHOH-59	University of Hohenheim	-26,35	-27,16	-26,75	a b
38	UHOH-54	University of Hohenheim	-26,40	-27,11	-26,75	a b
41	UHOH-57	University of Hohenheim	-26,54	-26,98	-26,76	a b c
13	ES-INSPECTOR	LIDEA	-26,64	-26,91	-26,77	a b c d
30	Xonia	Saatzucht Gleisdorf	-26,50	-27,13	-26,81	a b c d e
39	UHOH-55	University of Hohenheim	-26,15	-27,62	-26,88	a b c d e f
20	RGT SEFORA	RAGT	-26,66	-27,14	-26,90	a b c d e f
1	CH22177 Galice	Agroscope	-26,87	-27,00	-26,93	a b c d e f
29	GL 201117	Saatzucht Gleisdorf	-26,48	-27,42	-26,94	a b c d e f
9	CH22910	Agroscope	-26,74	-27,26	-27,00	a b c d e f g
33	GL 221715	Saatzucht Gleisdorf	-26,86	-27,22	-27,03	a b c d e f g h
36	UHOH-52	University of Hohenheim	-26,74	-27,39	-27,06	a b c d e f g h i
35	UHOH-51	University of Hohenheim	-27,16	-26,97	-27,06	a b c d e f g h i
26	GL Pia (GL 1816014)	Saatzucht Gleisdorf	-26,86	-27,31	-27,08	a b c d e f g h i j
27	GL 201212	Saatzucht Gleisdorf	-26,76	-27,43	-27,09	a b c d e f g h i j
12	ES-DIRECTOR	LIDEA	-26,87	-27,34	-27,10	a b c d e f g h i j
10	CH90166	Agroscope	-26,92	-27,30	-27,11	a b c d e f g h i j
4	CH22681 Prolix	Agroscope	-26,92	-27,37	-27,14	b c d e f g h i j
24	SHIVA	RAGT	-26,97	-27,49	-27,23	b c d e f g h i j k
17	ES-PROFESSOR	LIDEA	-27,12	-27,35	-27,24	b c d e f g h i j k
14	ES-LIBERATOR	LIDEA	-26,73	-27,81	-27,27	b c d e f g h i j k
3	CH22554 Simpol	Agroscope	-26,77	-27,77	-27,27	b c d e f g h i j k l
2	CH22508 Paprika	Agroscope	-27,20	-27,37	-27,28	b c d e f g h i j k l
22	SUZA	RAGT	-27,12	-27,50	-27,31	b c d e f g h i j k l m
19	RGT STUMPA	RAGT	-27,38	-27,30	-27,34	c d e f g h i j k l m
21	RGT SAKUSA	RAGT	-27,45	-27,27	-27,36	d e f g h i j k l m
31	GL 221712	Saatzucht Gleisdorf	-27,13	-27,63	-27,38	e f g h i j k l m
32	GL 221623	Saatzucht Gleisdorf	-27,23	-27,66	-27,44	f g h i j k l m
40	UHOH-56	University of Hohenheim	-27,31	-27,61	-27,46	f g h i j k l m
8	CH90139	Agroscope	-27,26	-27,67	-27,47	f g h i j k l m
34	GL Tilda	Saatzucht Gleisdorf	-26,97	-28,10	-27,53	g h i j k l m
28	GL 201811	Saatzucht Gleisdorf	-27,31	-27,79	-27,55	g h i j k l m
18	ES-VISITOR	LIDEA	-27,34	-27,77	-27,55	g h i j k l m
16	ES-MENTOR	LIDEA	-27,56	-27,56	-27,56	g h i j k l m
25	SIBELLA	RAGT	-27,58	-27,59	-27,58	h i j k l m
44	UHOH-60	University of Hohenheim	-27,72	-27,54	-27,62	i j k l m
7	CH90132 Benno	Agroscope	-27,67	-27,58	-27,62	j k l m
23	SQUADRA	RAGT	-27,63	-27,68	-27,65	j k l m
6	CH90093 Soramax	Agroscope	-27,55	-27,91	-27,73	k l m n
11	LID-CONSTRUCTOR	LIDEA	-27,62	-27,88	-27,74	k l m n
15	ES-LOUXOR	LIDEA	-28,10	-27,55	-27,82	l m n
5	CH90075 Talisse	Agroscope	-27,99	-27,78	-27,88	m n
45	CH22172 Obelix	Agroscope	-28,28	-28,22	-28,25	n

Legume Generation

Carbon isotope discrimination (^{12}C *vs* ^{13}C) for measuring drought sensitivity / water use efficiency (n=45, set Y2, LegGen, Dt2/dt2)

Contents



- Soybean breeding
- Adaptation to Europe
- Stem termination
- **Digital phenotyping**
- How does genomics support soybean breeding ?
- Conclusions



?

Source: Wikipedia (2020 Aug 5) - https://en.wikipedia.org/wiki/Infrared



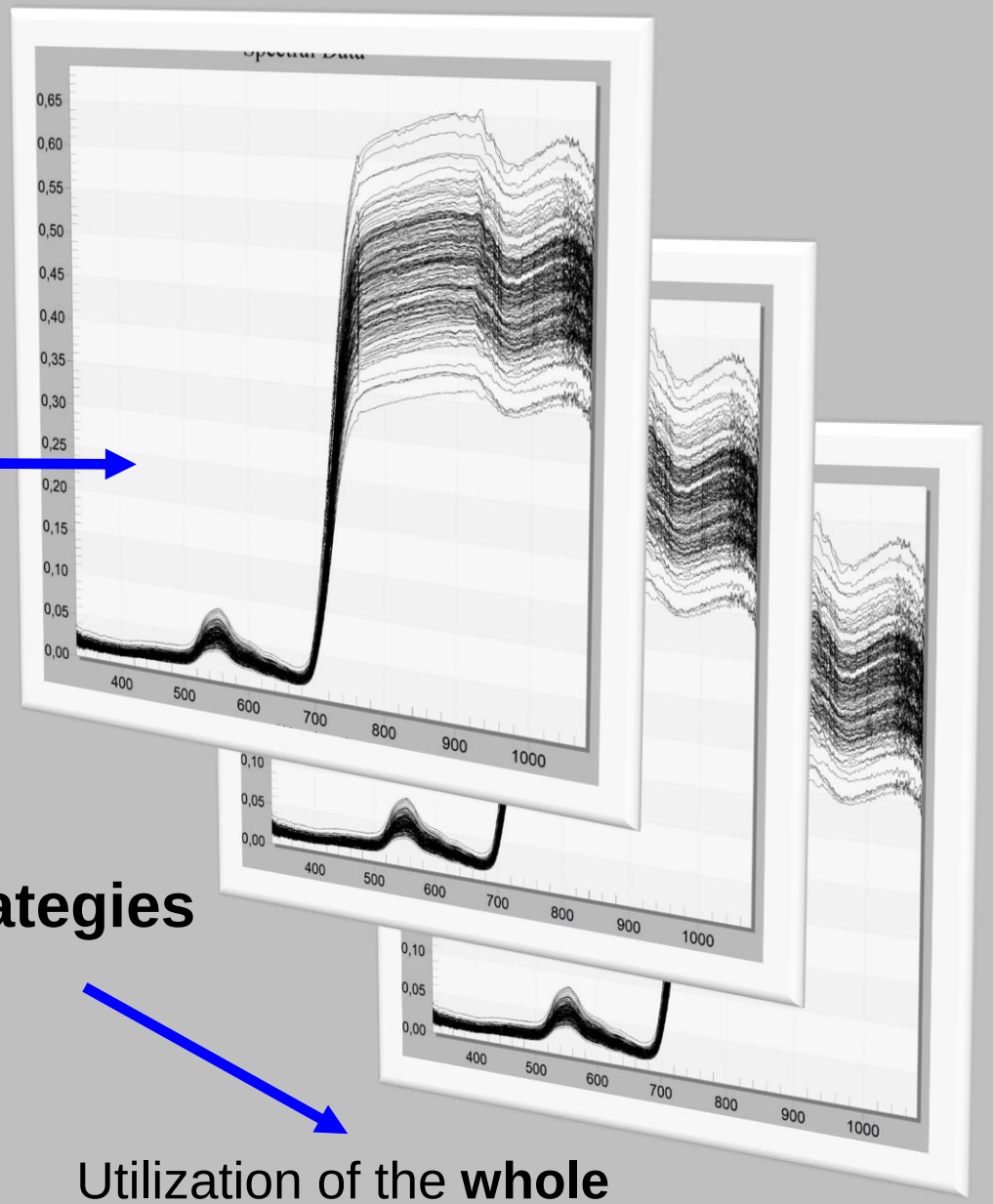


Source: Wikipedia (2020 Aug 5) - https://en.wikipedia.org/wiki/Infrared



Hyperspectral reflectance

ASD handheld spectroradiometer (λ 325-1075 nm)



Phenotypic data

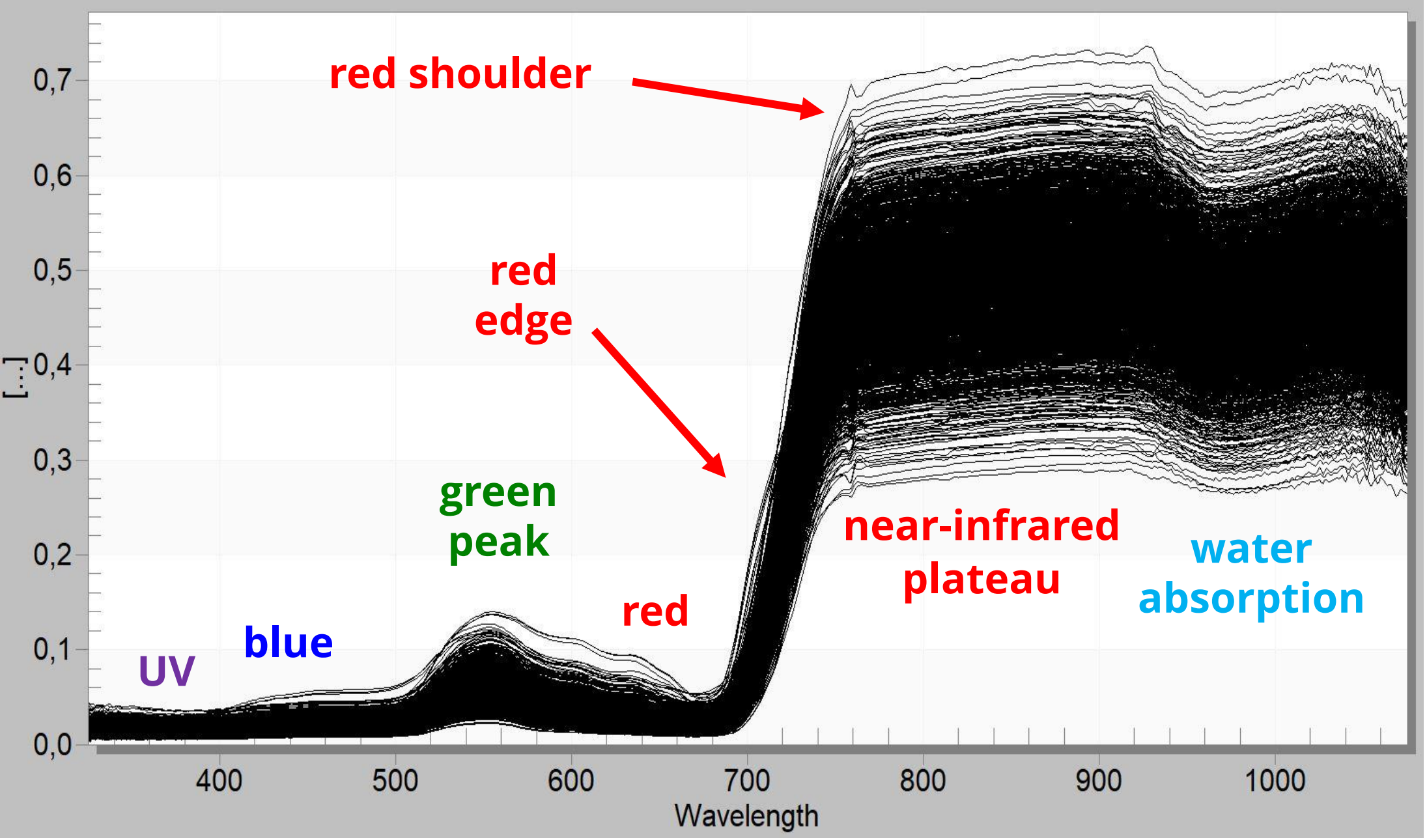
(for model development)

three strategies

Correlation between individual wavelengths and plant characters

Calculation of **spectral reflectance indices** (SRIs) for predicting traits

Utilization of the **whole spectrum** for developing **predictive regression models**



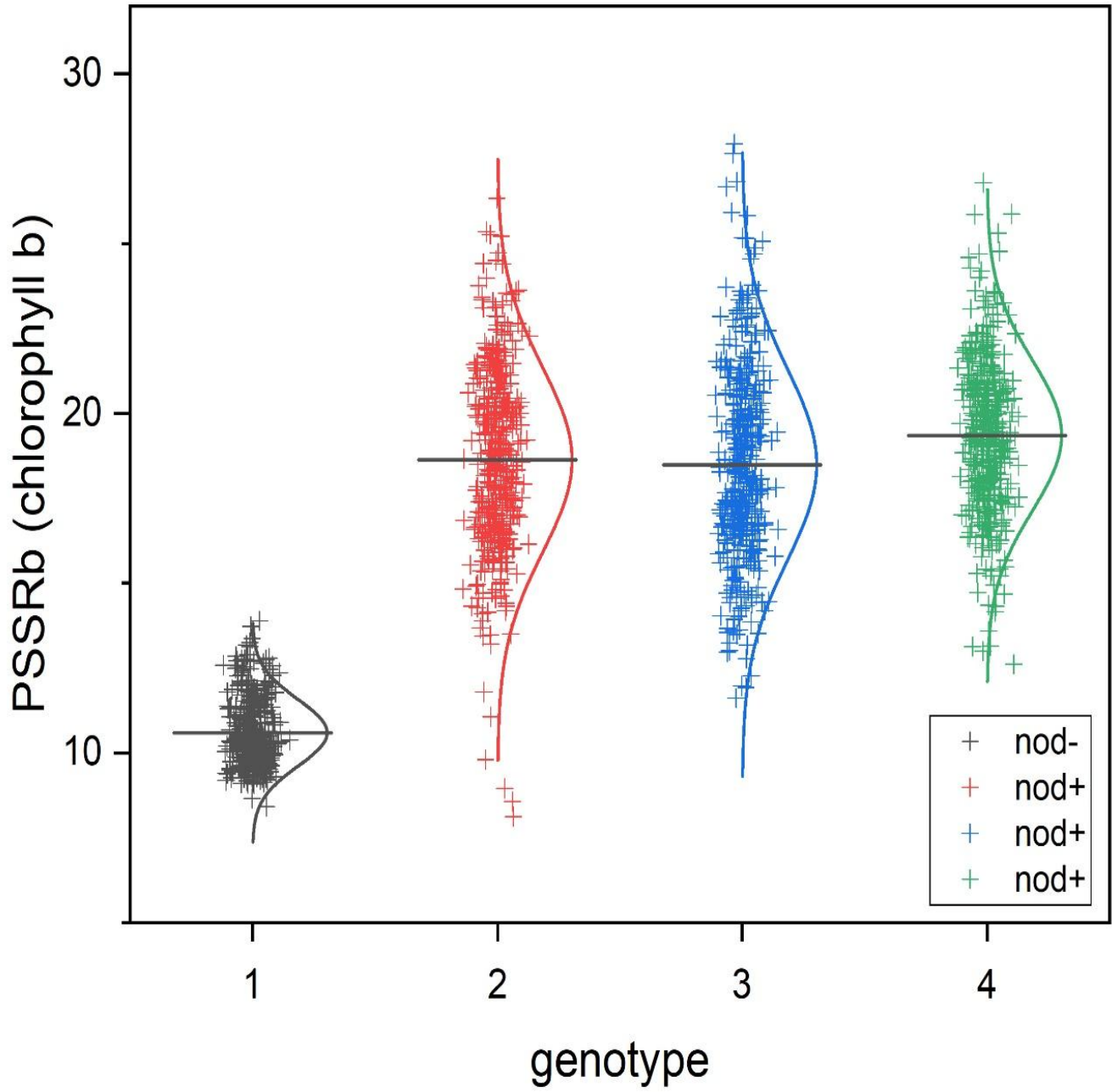
ASD HandHeld 2 reflectance spectra (325-1075 nm) from soybean leaves

Spectral reflectance indices (SRI)

NDVI	NDVI, normalized difference vegetation index (biomass, leaf area)
SIPI	SIPI, structure insensitive pigment index (carotenoid / chlorophyll a ratio)
PSSRa	PSSRa, pigment specific simple ratio (chlorophyll a)
PSSRb	PSSRb (chlorophyll b)
CRI	CRI 1 , carotenoids reflectance index (carotenoid content)
ARI	ARI 1, anthocyanin reflectance index (anthocyanin content)
PSSRc	Pigment-specific simple ratio (carotenoid content)
NRI	Nitrogen reflectance index
RVSI	Red-edge vegetation stress index
PSRI	Plant senescence reflectance index
CI	CI read edge; Canopy chlorophyll and nitrogen content, red edge chlorophyll index
PRI570	Light use efficiency, drought indicator, PRI=phytochemical reflectance index; plant water status
DCNI	Double-peak canopy nitrogen index
GI	Greenness index
GNDVI	Green normalized difference vegetation index
VOG1	Vogelmann index 1
VOG2	Vogelmann index 2
VOG3	Vogelmann index 3
R705	Blue nitrogen index (leaf nitrogen content) / Tian et al 2010
R780/740	Total N uptake
MSR705_445	Modified simple ratio 705/445
REIP	Red edge inflection point
WI	WI, reflectance water index (water content)
NWI-1	Norm. water index 1
NWI-2	Norm. water index 2
NWI-3	Norm. water index 3
NWI-4	Norm. water index 4
NWI-5	Norm. water index 5
WI_1	Water index 1 (Christenson et al. 2016, soybean)
WI_2	Water index 2 (Christenson et al. 2016, soybean)
WI_3	Water index 3 (Christenson et al. 2016, soybean)
RNDVI_1	red NDVI 1
RNDVI_2	red NDVI 2
RNDVI_3	red NDVI 3
RENDVI_1	red edge NDVI 1
RENDVI_2	red edge NDVI 2
RENDVI_3	red edge NDVI 3
WI / NDVI	WI : NDVI
WBI	Water band index (1/WI)
MA1_N	R5 NDVI normalized diff vegetation index 638_674
MA1_R	R5 RVI Ratio vegetation index 638_674
MB1_N	R5 NDVI normalized diff vegetation index 634_678
MB1_R	R5 RVI Ratio vegetation index 634_678

Indices calculated from soybean hyperspectral reflectance data

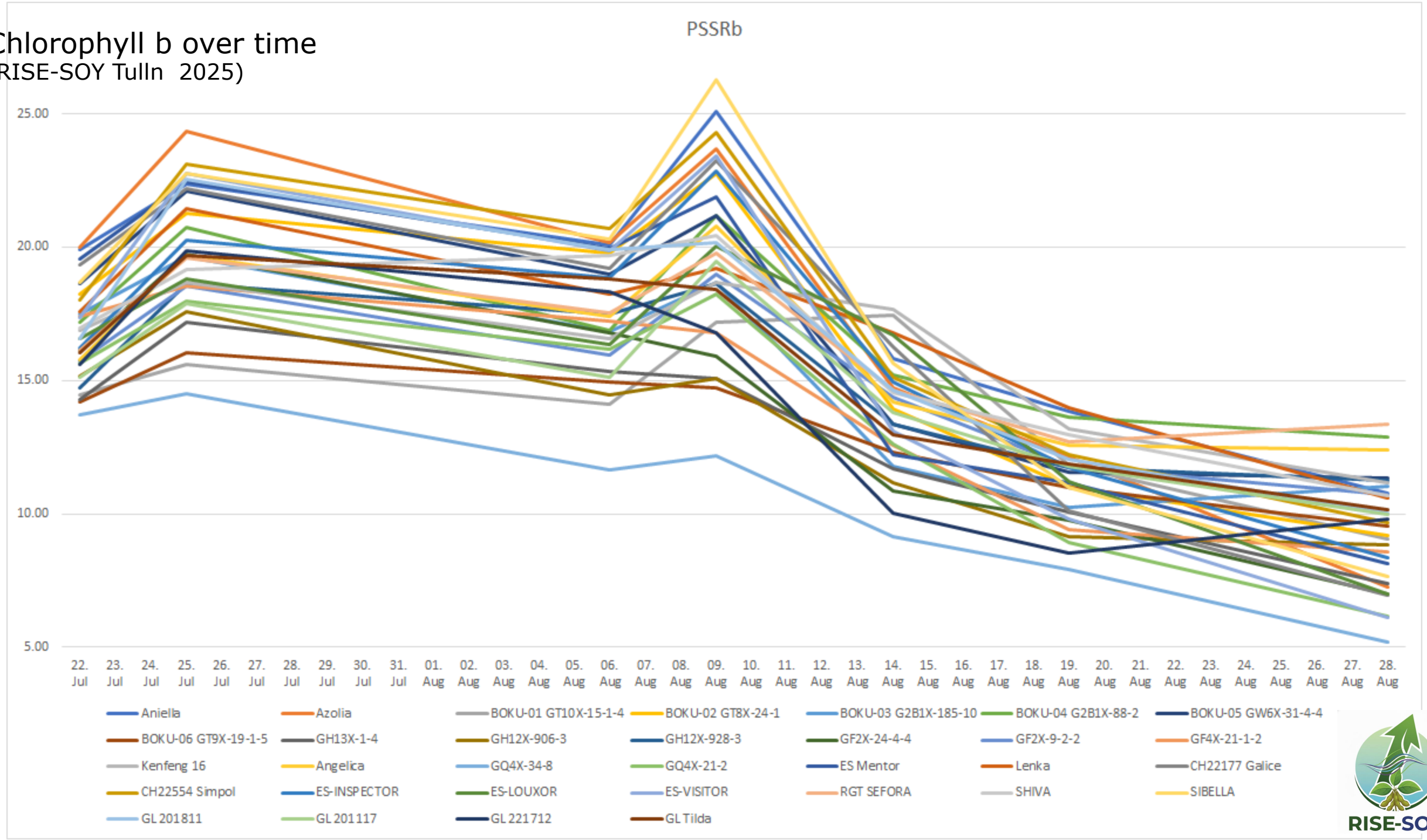




The chlorophyll b index illustrates differences between nodulating and non-nodulating soybean genotypes.

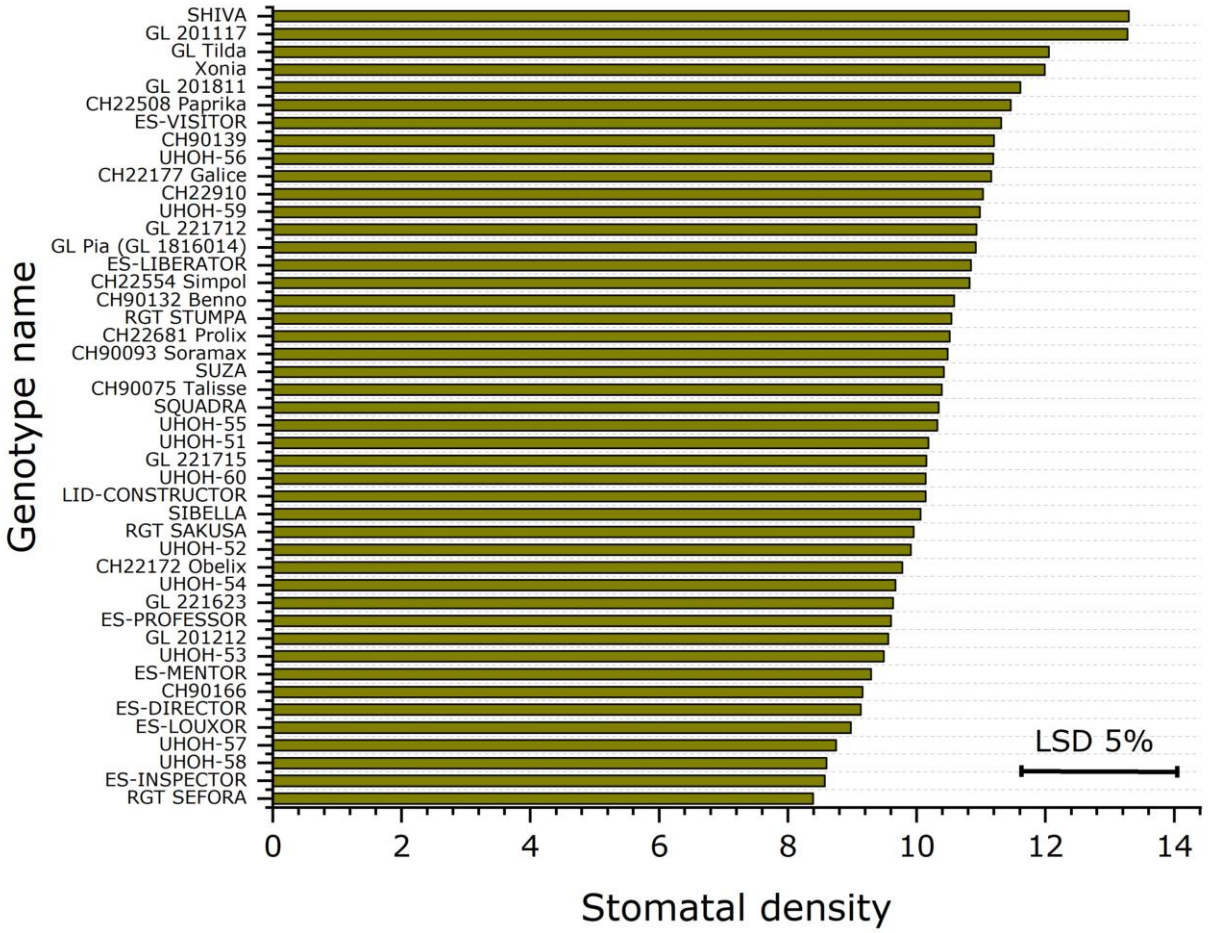
Source: Vollmann et al (2022), https://doi.org/10.1016/j.compag.2022.107169

Chlorophyll b over time
(RISE-SOY Tulln 2025)



RISE-SOY

Drought stress

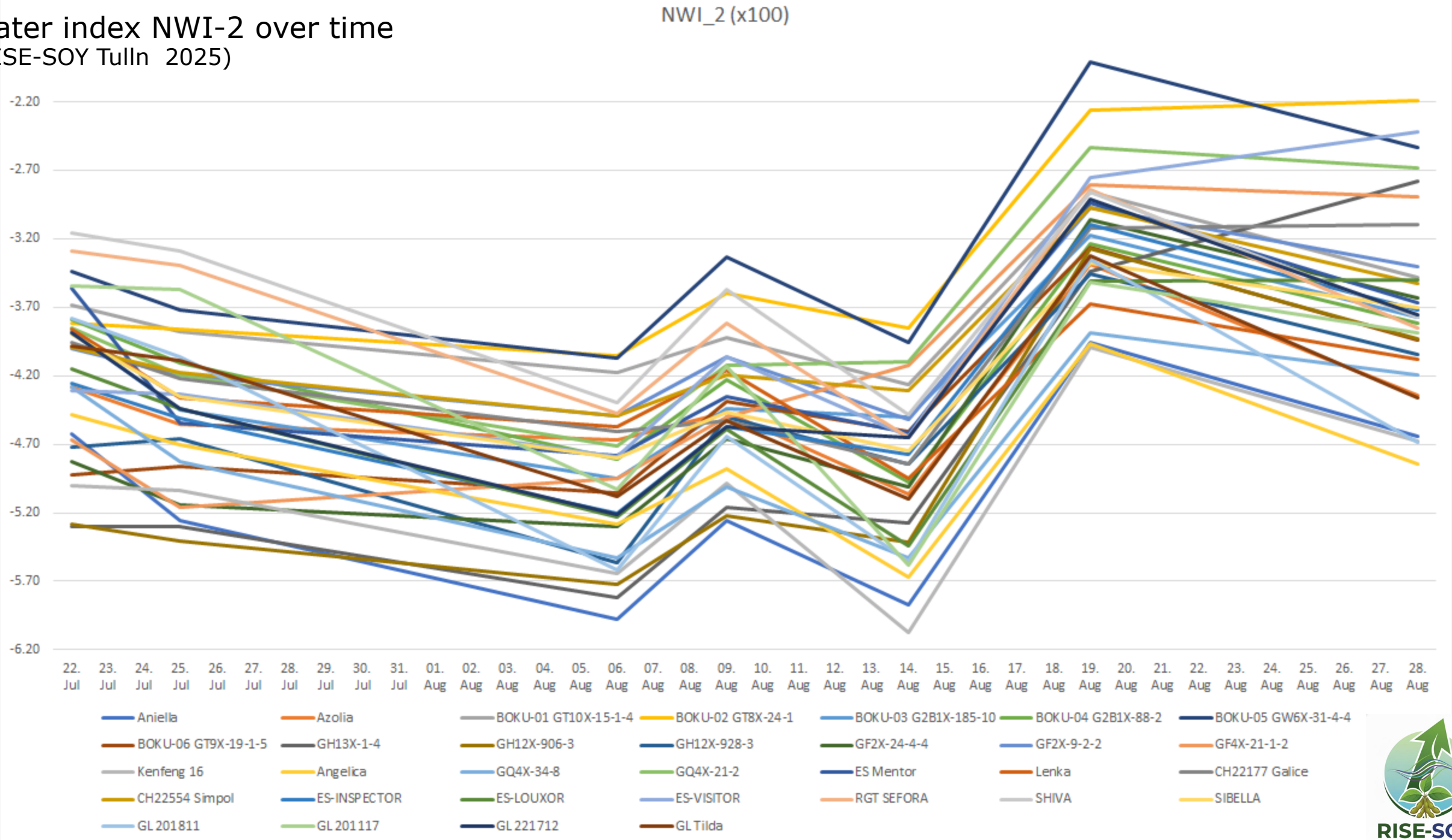


Ranking of genotypes of the Y2 (MG 00) set according to leaf stomatal density (number of stomata per 0.04 mm^2)

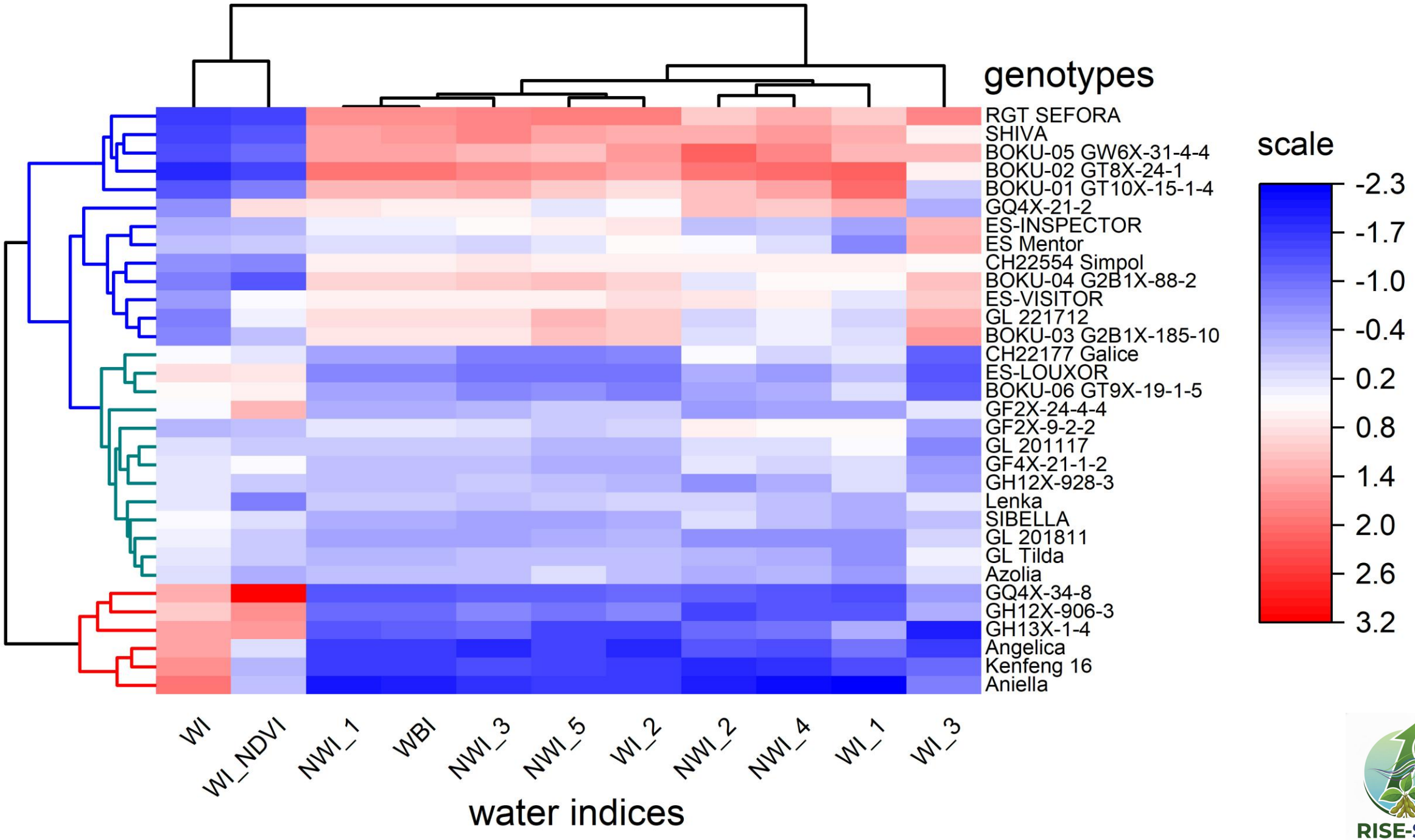


Variation in wilting level between Y2 cultivars during high temperature periods (38°C, Tulln/AT, Aug 2026)

Water index NWI-2 over time
(RISE-SOY Tulln 2025)



RISE-SOY


genotypes
RGT SEFORA
SHIVA
BOKU-05 GW6X-31-4-4
BOKU-02 GT8X-24-1
BOKU-01 GT10X-15-1-4
GQ4X-21-2
ES-INSPECTOR
ES Mentor
CH22554 Simpol
BOKU-04 G2B1X-88-2
ES-VISITOR
GL 221712
BOKU-03 G2B1X-185-10
CH22177 Galice
ES-LOUXOR
BOKU-06 GT9X-19-1-5
GF2X-24-4-4
GF2X-9-2-2
GL 201117
GF4X-21-1-2
GH12X-928-3
Lenka
SIBELLA
GL 201811
GL Tilda
Azolia
GQ4X-34-8
GH12X-906-3
GH13X-1-4
Angelica
Kenfeng 16
Aniella
scale
-2.3
-1.7
-1.0
-0.4
0.2
0.8
1.4
2.0
2.6
3.2
WI
WI_NDVI
NWI_1
WBI
NWI_3
NWI_5
WI_2
NWI_2
NWI_4
WI_1
WI_3
water indices

Legume Generation

Drought stress adaptation

Grouping of genotypes of the Y2 (MG 00) set according to nine **spectral water indices** describing differences in water status of genotypes (Tulln, Austria, 2024)

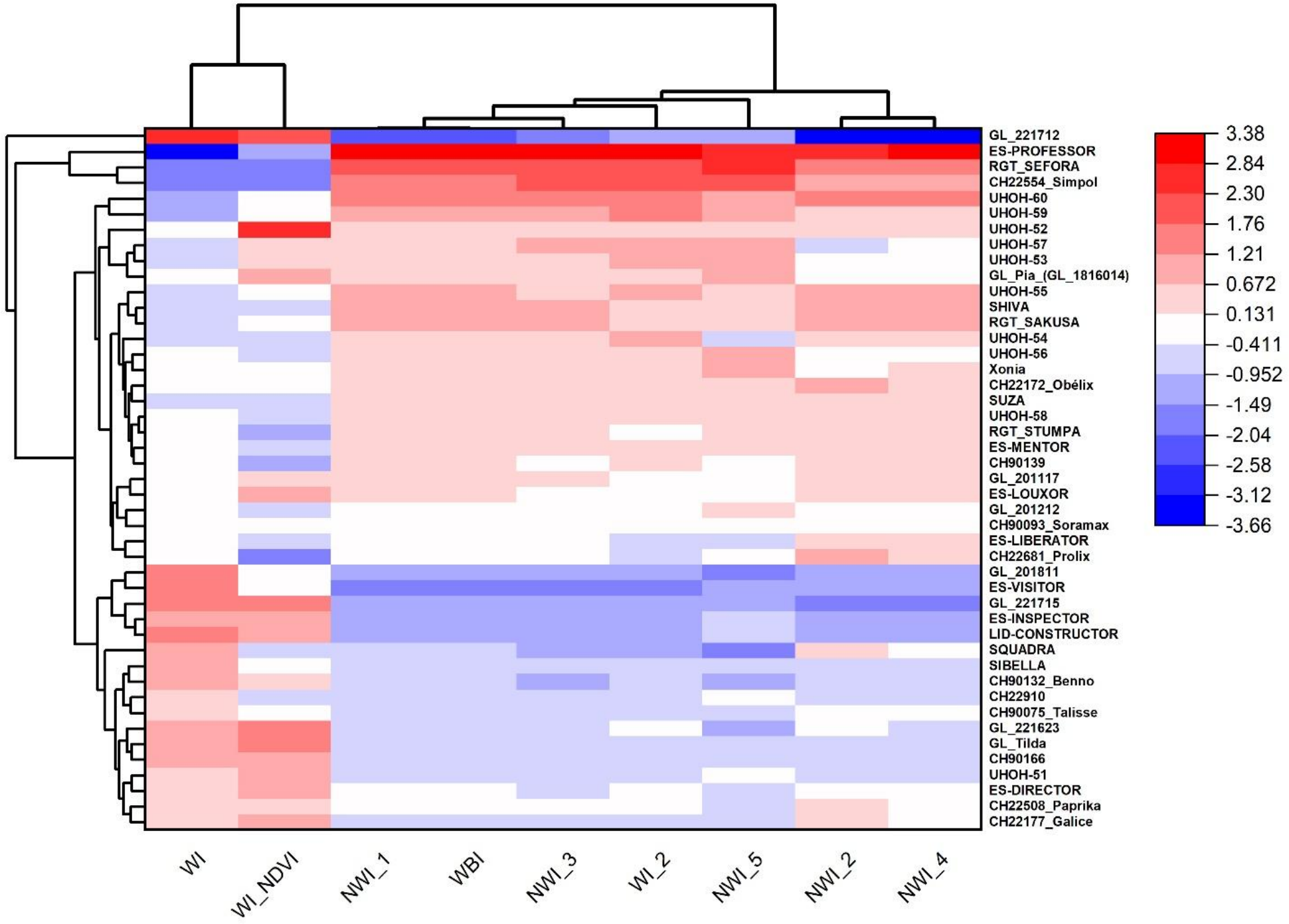


Table. *Coefficients of correlation between selected spectral reflectance indices (determined on July 27) and agronomic or seed quality traits*

INDEX	ttf	ttm	SPAD	lsize	podse	plh	yield	tsw	prot	oil	sucr	proty	stom
NDVI	-0.057	-0.255	-0.031	0.092	0.472**	-0.511**	0.106	-0.076	-0.024	-0.067	-0.221	0.105	0.109
NRI	0.294	0.007	-0.257	0.393**	-0.008	-0.045	-0.147	0.112	-0.011	0.011	0.048	-0.161	0.134
CI	-0.336*	-0.279	0.120	-0.175	0.489**	-0.459**	0.278	-0.153	-0.039	-0.083	-0.214	0.283	0.106
REIP	-0.396**	-0.204	0.138	-0.309*	0.363*	-0.327*	0.357*	-0.299*	-0.071	-0.028	-0.193	0.359*	0.023
PSSRa	-0.125	-0.265	-0.006	0.036	0.460**	-0.496**	0.151	-0.088	-0.067	-0.083	-0.191	0.143	0.107
PSSRb	-0.120	-0.255	0.024	0.032	0.486**	-0.510**	0.170	-0.102	-0.048	-0.104	-0.229	0.167	0.094
MA1_R	-0.023	0.024	-0.052	0.012	-0.268	0.192	-0.202	0.072	-0.125	0.162	0.356*	-0.236	0.092
WI	-0.428**	-0.149	0.029	-0.410**	0.125	-0.038	0.346*	-0.082	-0.012	-0.071	0.125	0.353*	0.165
WI_1	0.109	-0.036	-0.050	0.069	0.124	-0.364*	-0.436**	0.069	-0.198	0.164	-0.027	-0.478**	-0.269
WI_2	0.371*	0.184	-0.044	0.407**	-0.241	0.237	-0.180	0.109	0.023	0.116	-0.104	-0.184	-0.158
WI_3	0.306*	0.186	-0.050	0.385**	-0.334*	0.440**	0.082	0.067	0.138	0.050	-0.150	0.103	-0.076
NWI_1	0.426**	0.152	-0.040	0.407**	-0.126	0.042	-0.339*	0.089	0.013	0.073	-0.128	-0.346*	-0.162
NWI_2	0.355*	0.012	-0.051	0.256	0.072	-0.275	-0.414**	-0.019	-0.087	0.105	-0.198	-0.438**	-0.249
NWI_3	0.447**	0.195	-0.049	0.454**	-0.163	0.109	-0.296*	0.098	0.043	0.047	-0.097	-0.295*	-0.132
NWI_4	0.385**	0.049	-0.065	0.323*	0.015	-0.191	-0.423**	0.048	-0.051	0.098	-0.163	-0.442**	-0.217
NWI_5	0.467**	0.189	-0.060	0.519**	-0.209	0.187	-0.232	0.147	0.112	-0.009	-0.095	-0.217	-0.103
	ttf	ttm	SPAD	lsize	podse	plh	yield	tsw	prot	oil	sucr	proty	stom

ttf: time to flowering; ttm: time to maturity; SPAD: SPAD value; lsize: leaf size; podse: pod set score; plh: plant height; yield: grain yield; tsw: 1000-seed weight; prot: seed protein content; oil: oil content; sucr: sucrose content; proty: protein yield.

Legume Generation

Contents



- Soybean breeding
- Adaptation to Europe
- Stem termination
- Digital phenotyping
- **How does genomics support soybean breeding ?**
- Conclusions

LegGen Soy 5K Trait-specific Marker Panel

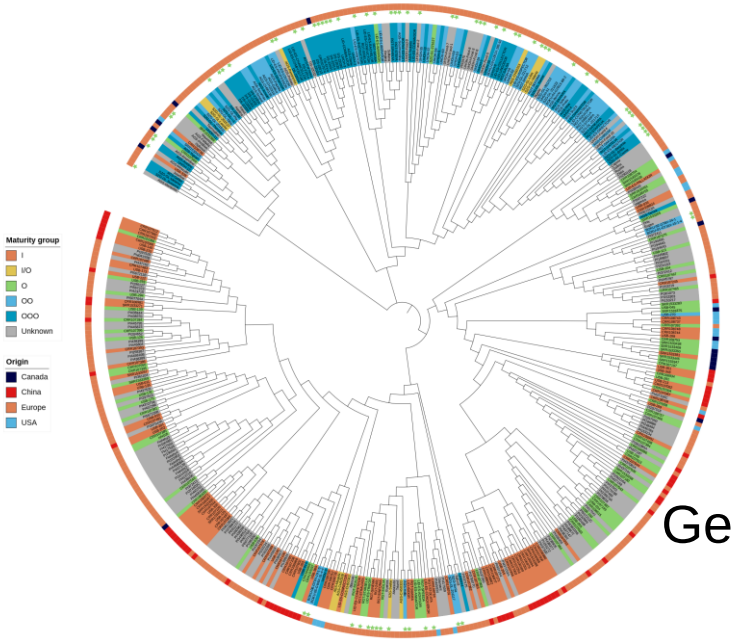
Legume Generation

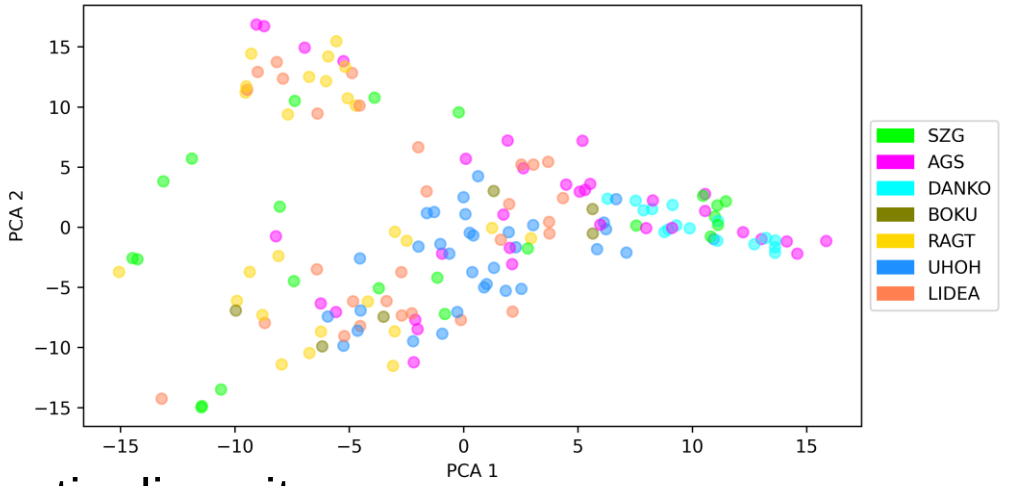
Over 5000 markers including

355 trait-specific and breeder‘s markers (causative mutations, adopted markers)

Other diagnostic and compatibility markers

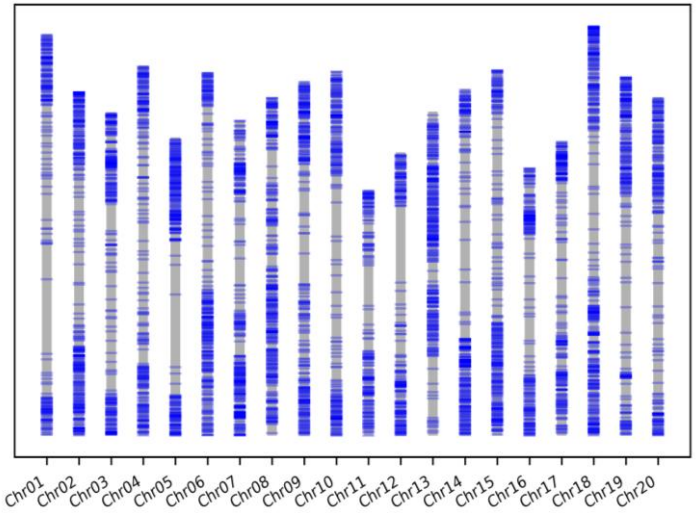
Platform: Flex-Seq™ platform of LGC Biosearch Technologies, Inc.



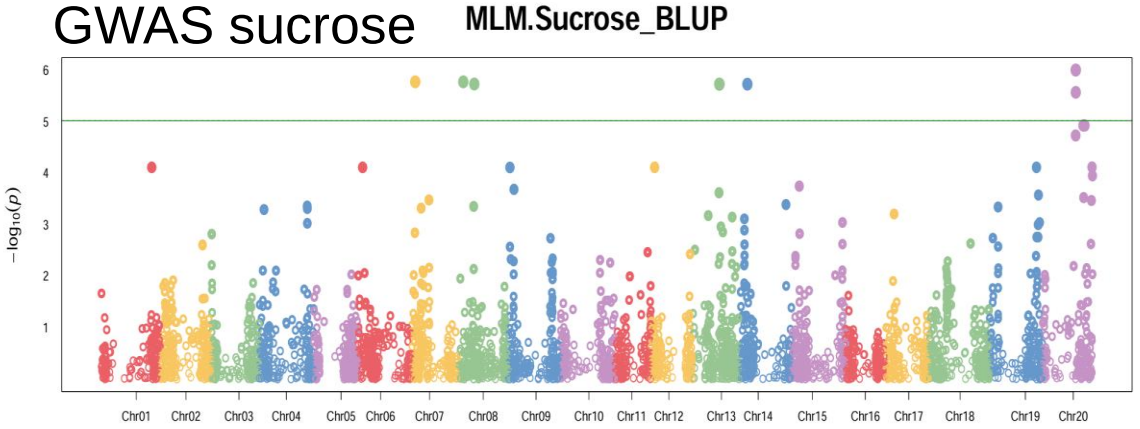


Genetic diversity

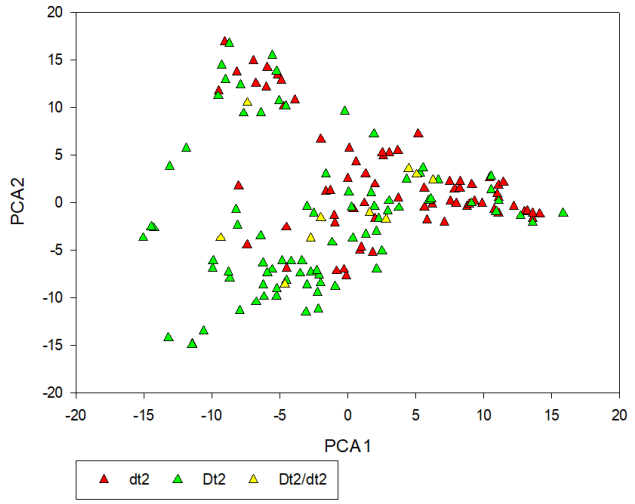
Coverage



GWAS sucrose



Dt2 distribution



Sample applications of LegGen Soy 5K panel:
Courtesy of Mária Škrabišová and Jana Biová, Legume Generation / Palacký University Olomouc

Contents



- Soybean breeding
- Adaptation to Europe
- Stem termination
- Digital phenotyping
- How does genomics support soybean breeding ?
- **Conclusions**



Summary / conclusions

- **Phenological adaptation** to latitude is essential for performance at target environments
- ***Dt2* stem termination:** Ambiguous role of semi-determinacy
- Huge potential of **digital phenotyping**
- **Genomics is supporting** the breeding process, urgent need for implementation research





Acknowledgements

Martin Pachner
Xindong Yao
Leopold Rittler
Benedikt Piller
Matthias Fidesser
Ahmad M. Manschadi
Claudia Jonak
Jonas Pilstl
Mária Škrabišová
Jana Biová
Babak Malekian
Rebecca Nowotny-Hood
Jana Seiter

HABERLANDT Project

Legume Generation

Legume Generation (Boosting innovation in breeding for the next generation of legume crops for Europe) has received funding from the European Union's Horizon Europe research and innovation programme under grant agreement No.101081329. It also receives support from the governments of the United Kingdom, Switzerland and New Zealand.

RISE-SOY

WISSENSCHAFT · FORSCHUNG NIEDERÖSTERREICH

FTI-STRATEGIE NIEDERÖSTERREICH 2021–2027

GESELLSCHAFT FÜR FORSCHUNGS FÖRDERUNG NIEDERÖSTERREICH

FUNDED AS PART OF THE RTI-STRATEGY LOWER AUSTRA 2027