The *Aegilops tauschii* pangenome: A tool for harnessing wild wheat genetic diversity

Andrea Gonzalez
PhD student
The Wulff Lab
King Abdullah University of Science and Technology
Thuwal, Saudi Arabia

The Open Wild Wheat Consortium

Illustration by Robyn Palescandolo

Unlocking the Diversity of Wild Wheat

We are an international consortium of researchers with strategic interests in Tricticeae breeding, genomics, genetics, pathology, and germplasm collection, preservation, and distribution.

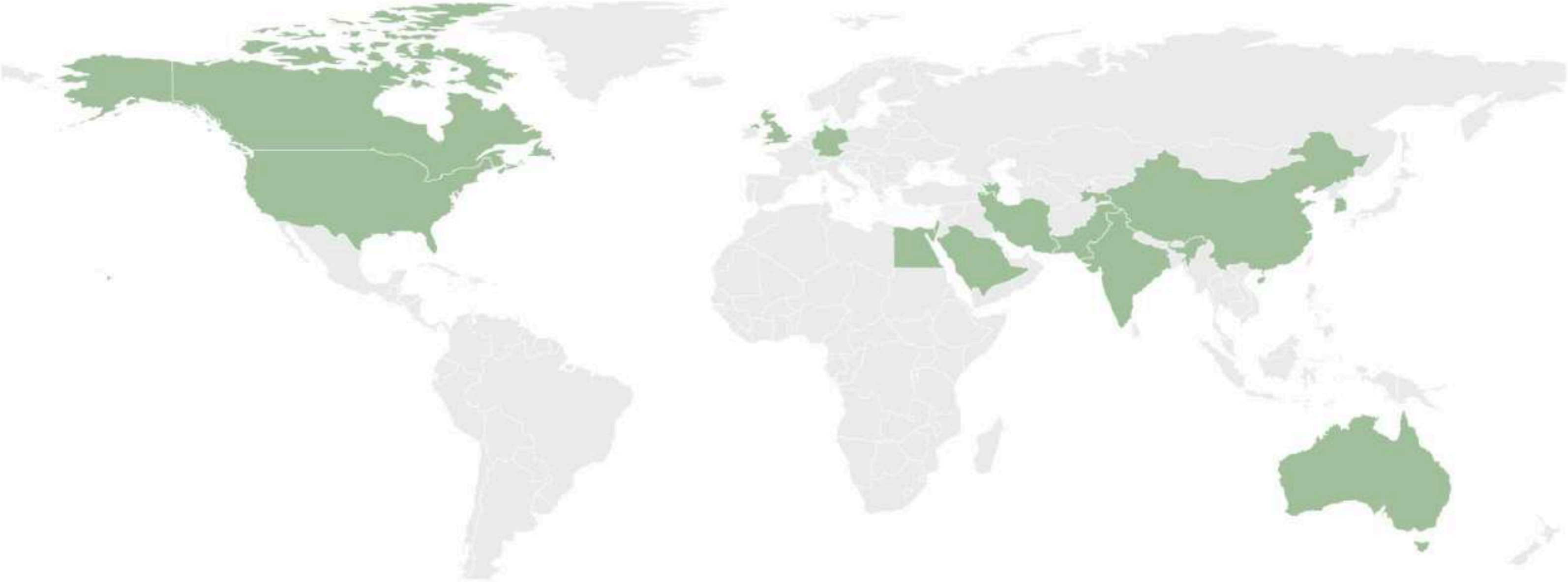
OPEN WILD WHEAT

www.openwildwheat.org

The Open Wild Wheat Consortium

Cross-Continental Collaboration

The Open Wild Wheat Consortium is represented by 15 countries across America, Europe, Asia, Australia, and Africa.



OPEN WILD WHEAT

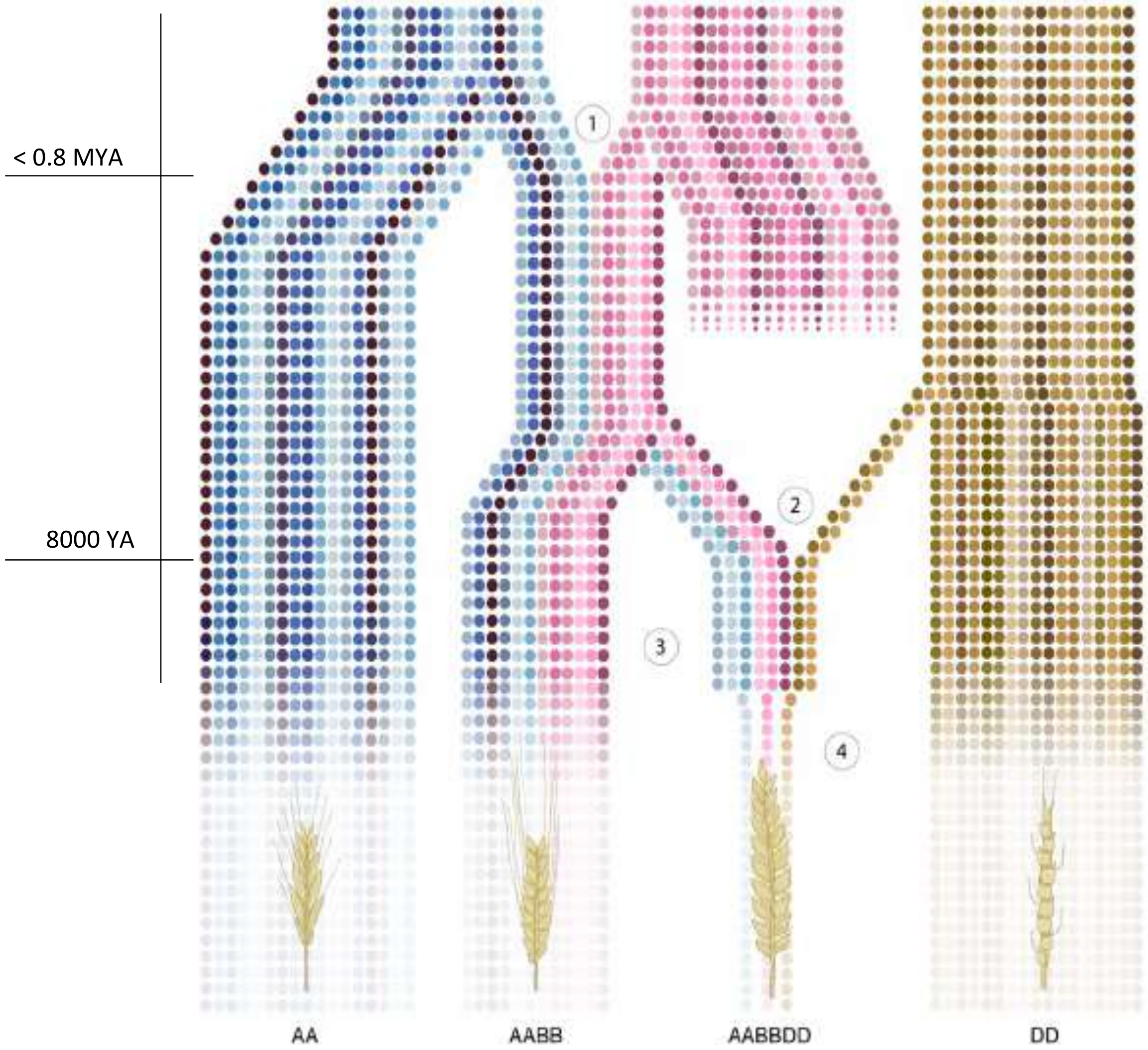
www.openwildwheat.org

1. Agricultural Research Center, Egypt
2. Research Institute of Forests and Rangelands, Iran
3. Tel Aviv University, Israel
4. King Abdullah University of Science and Technology, Saudi Arabia
5. Quaid-i-Azam University, Pakistan
6. University of Minnesota, USA
7. Sichuan Agricultural University, Chengdu, China
8. Julius Kühn-Institut, Germany
9. National Academy of Sciences of Tajikistan, Tajikistan
10. University of British Columbia, Canada
11. University of California Davis, USA
12. Leibniz Institute of Plant Genetics and Crop Plant Research (IPK) Gatersleben, Germany
13. Wheat Genetics Resource Center, Kansas State University, USA
14. National Institute of Agricultural Science, Republic of Korea
15. Institute of Crop Science, Chinese Academy of Agricultural Sciences, China
16. University of Sheffield, United Kingdom
17. Genetic Resources Institute, Ministry of Science and Education, Azerbaijan
18. John Innes Centre, United Kingdom
19. Punjab Agricultural University, India
20. University of Sydney, Australia
21. Institute of Molecular Biology and Biotechnologies, Azerbaijan
22. Kansas State University, USA
23. USDA-Agricultural Research Service, USA
24. University of Nottingham, United Kingdom
25. University of Maryland, USA
26. South Dakota State University, USA

Aegilops tauschii as a source of genetic diversity for bread wheat improvement

Aegilops tauschii
Wild relative of hexaploid wheat

- D-genome donor
- Diploid
- Genome size 4.36 Gb (Wheat = 16 Gb)



Hafeez et al. 2021

Aegilops tauschii as a source of genetic diversity for bread wheat improvement

Aegilops tauschii
Wild relative of hexaploid wheat

- D-genome donor
- Diploid
- Genome size 4.36 Gb (Wheat = 16 Gb)

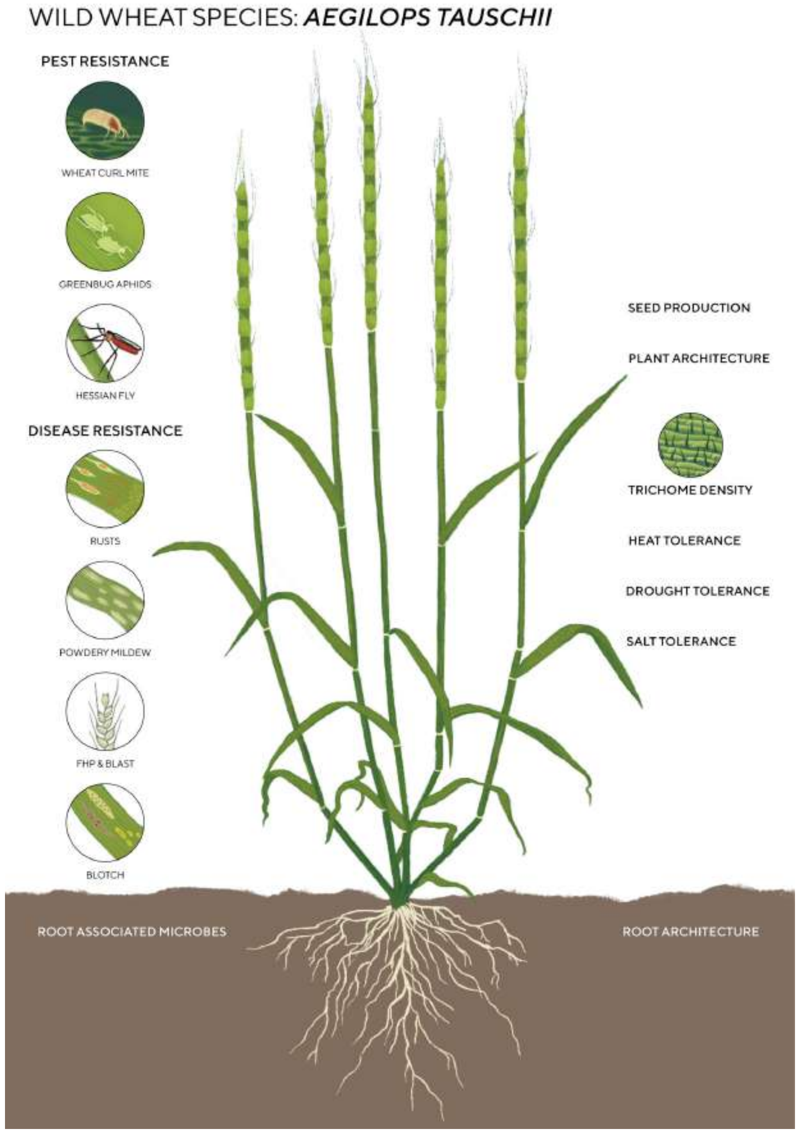
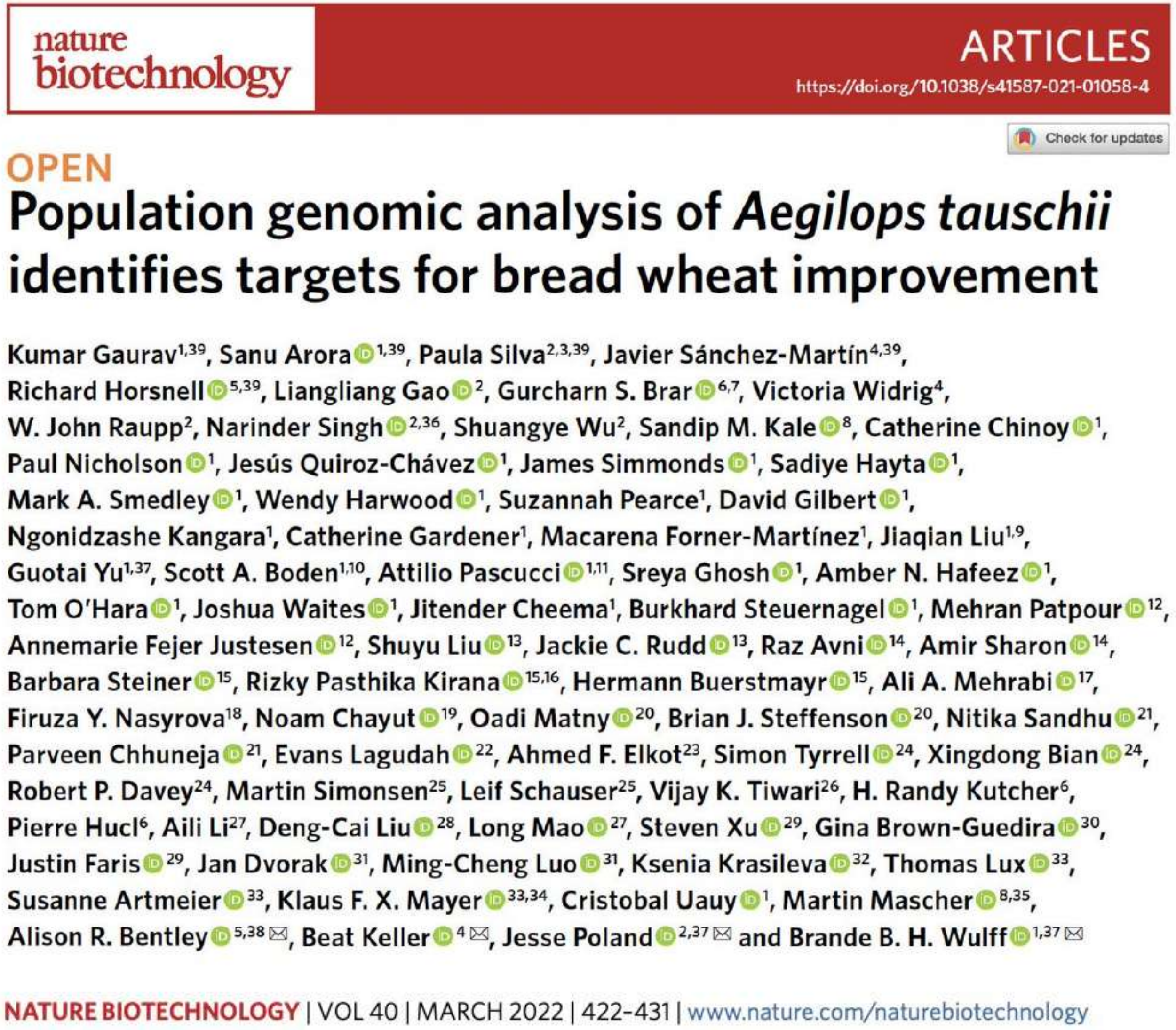


Illustration by Emma Waller. OpenWildWheat.org

Population genomics of *Aegilops tauschii*

OPEN
WILD
WHEAT

Phase I
2017-2021



nature biotechnology

ARTICLES
https://doi.org/10.1038/s41587-021-01058-4

Check for updates

OPEN

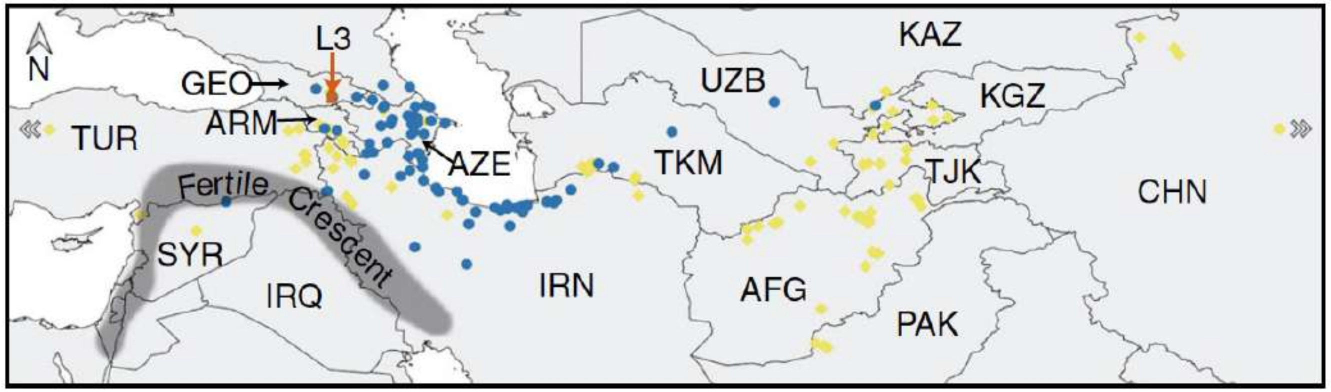
Population genomic analysis of *Aegilops tauschii* identifies targets for bread wheat improvement

Kumar Gaurav[1,39], Sanu Arora[1,39], Paula Silva[2,3,39], Javier Sánchez-Martín[4,39], Richard Horsnell[5,39], Liangliang Gao[2], Gurcharn S. Brar[6,7], Victoria Widrig[4], W. John Raupp[2], Narinder Singh[2,36], Shuangye Wu[2], Sandip M. Kale[8], Catherine Chinoy[1], Paul Nicholson[1], Jesús Quiroz-Chávez[1], James Simmonds[1], Sadiye Hayta[1], Mark A. Smedley[1], Wendy Harwood[1], Suzannah Pearce[1], David Gilbert[1], Ngonidzashe Kangara[1], Catherine Gardener[1], Macarena Forner-Martínez[1], Jiaqian Liu[1,9], Guotai Yu[1,37], Scott A. Boden[1,10], Attilio Pascucci[1,11], Sreya Ghosh[1], Amber N. Hafeez[1], Tom O'Hara[1], Joshua Waites[1], Jitender Cheema[1], Burkhard Steuernagel[1], Mehran Patpour[12], Annemarie Fejer Justesen[12], Shuyu Liu[13], Jackie C. Rudd[13], Raz Avni[14], Amir Sharon[14], Barbara Steiner[15], Rizky Pasthika Kirana[15,16], Hermann Buerstmayr[15], Ali A. Mehrabi[17], Firuza Y. Nasyrova[18], Noam Chayut[19], Oadi Matny[20], Brian J. Steffenson[20], Nitika Sandhu[21], Parveen Chhuneja[21], Evans Lagudah[22], Ahmed F. Elkot[23], Simon Tyrrell[24], Xingdong Bian[24], Robert P. Davey[24], Martin Simonsen[25], Leif Schauser[25], Vijay K. Tiwari[26], H. Randy Kutcher[6], Pierre Hucl[6], Aili Li[27], Deng-Cai Liu[28], Long Mao[27], Steven Xu[29], Gina Brown-Guedira[30], Justin Faris[29], Jan Dvorak[31], Ming-Cheng Luo[31], Ksenia Krasileva[32], Thomas Lux[33], Susanne Artmeier[33], Klaus F. X. Mayer[33,34], Cristobal Uauy[1], Martin Mascher[8,35], Alison R. Bentley[5,38] ✉, Beat Keller[4] ✉, Jesse Poland[2,37] ✉ and Brande B. H. Wulff[1,37] ✉

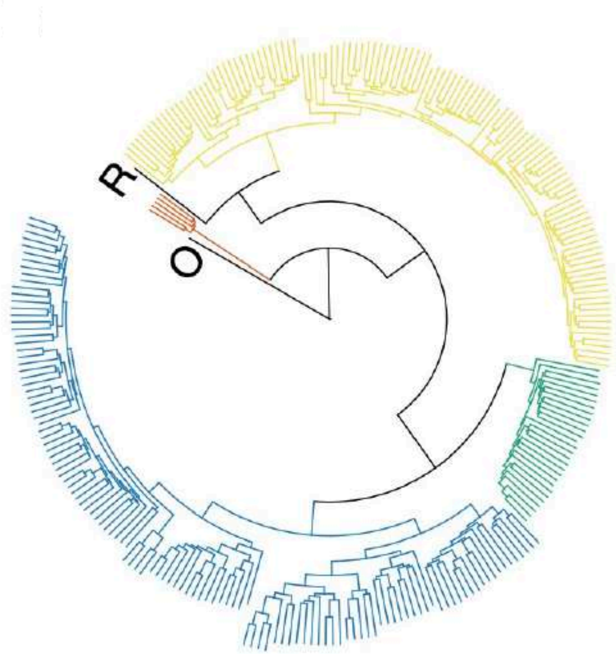
NATURE BIOTECHNOLOGY | VOL 40 | MARCH 2022 | 422–431 | www.nature.com/naturebiotechnology

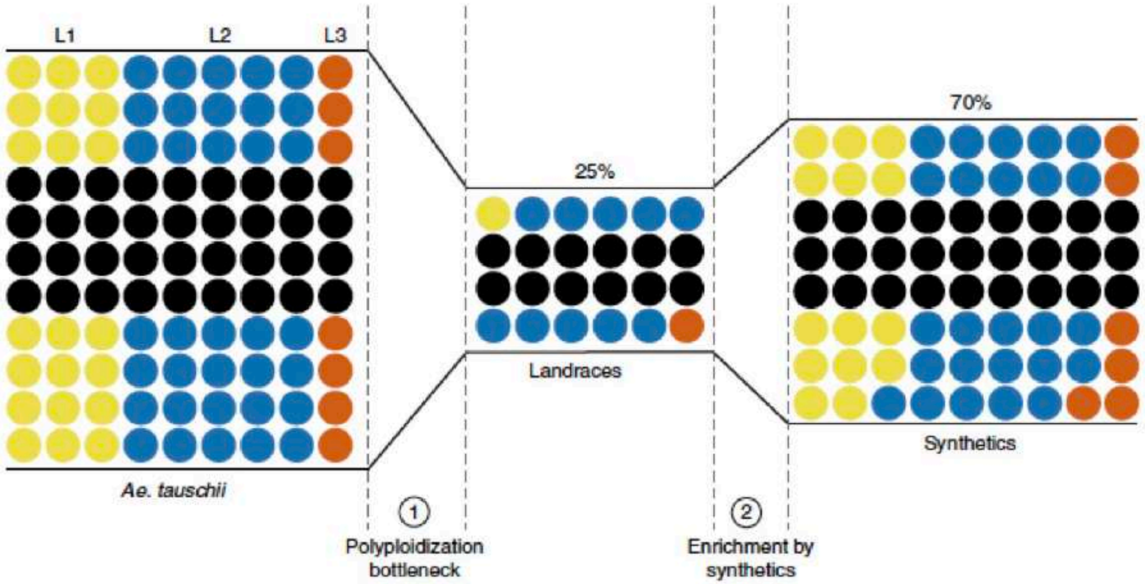
Population genomics of *Aegilops tauschii*

🟩 Wheat
🟨 L1
🟦 L2
🟧 L3



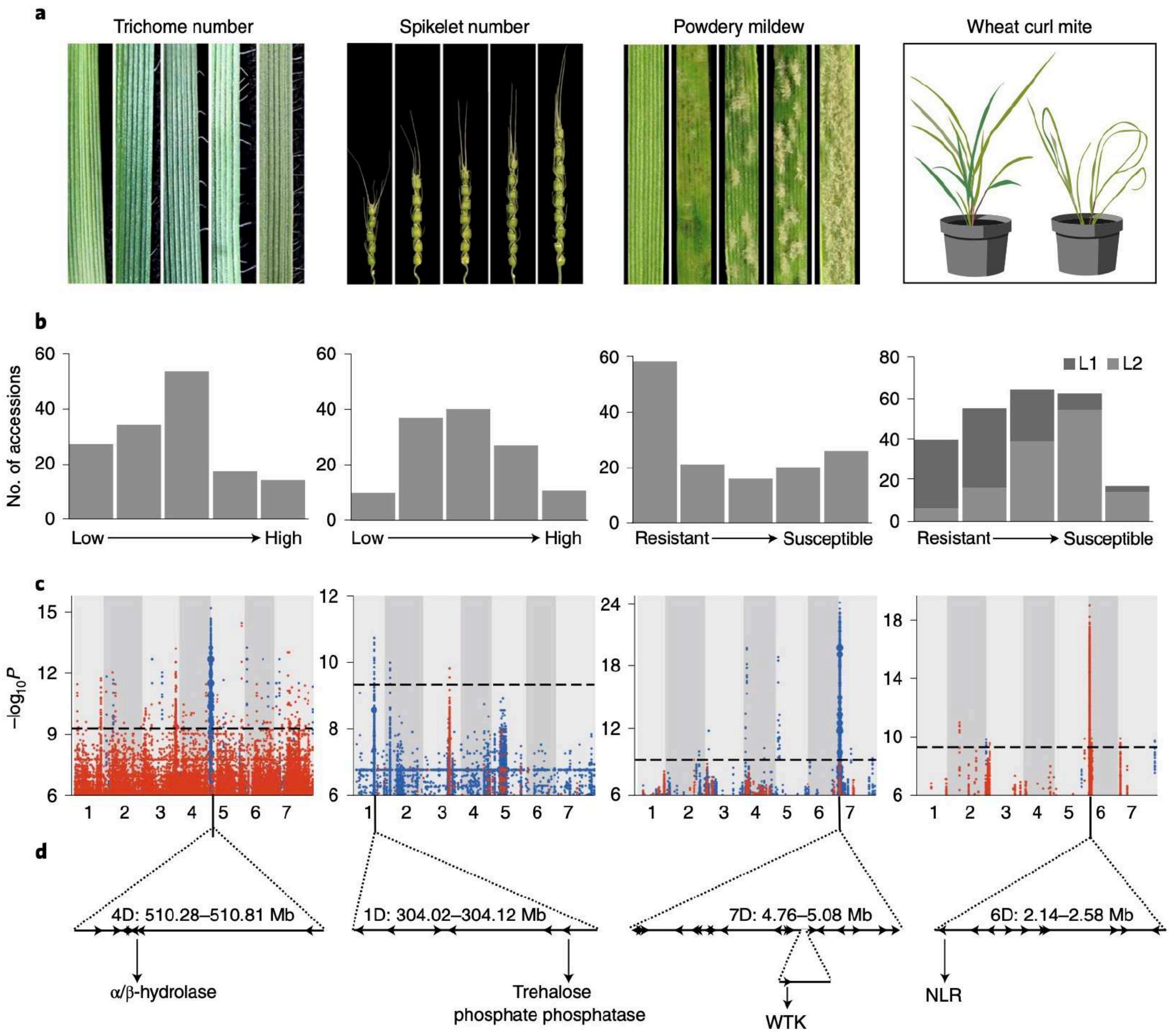
Sequenced diversity panel of 242 non-redundant *Ae. tauschii* accessions





Gaurav et al. Nature Biotechnology. 2022

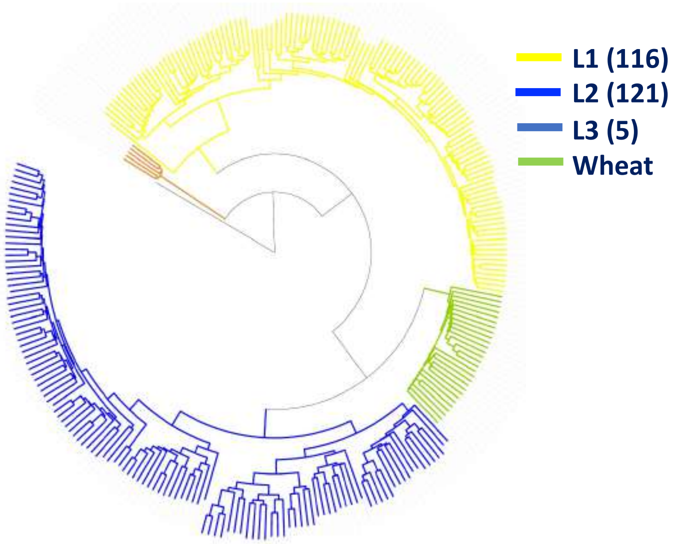
Dissecting traits of interest using the resequenced and phenotyped panel



Gaurav et al. Nature Biotechnology. 2022

Resource limitations of Open Wild Wheat Phase I

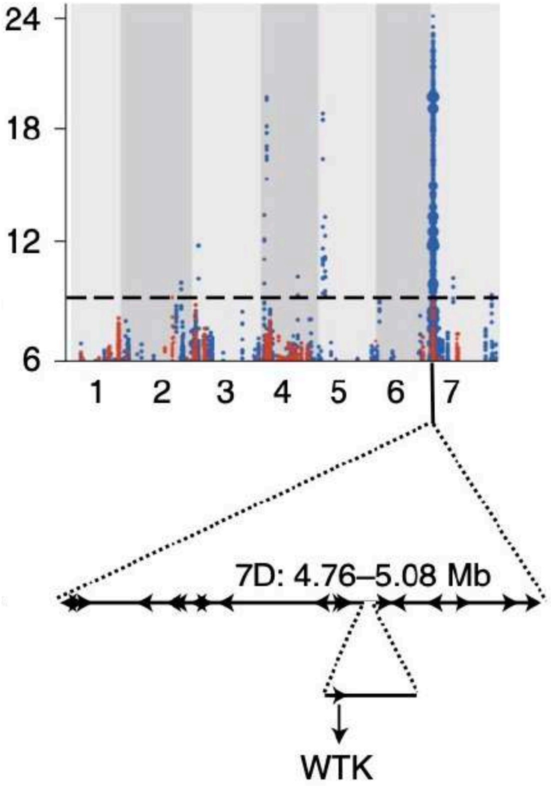
Detecting rare alleles



Adapted from Gaurav et al., 2022

- ✓ **Increase diversity panel size**

High-quality genome assemblies



Gaurav et al. Nature Biotechnology. 2022

- ✓ **Generate the pangenome**

The expanded *Aegilops tauschii* diversity panel

OPEN
WILD
WHEAT

Phase II

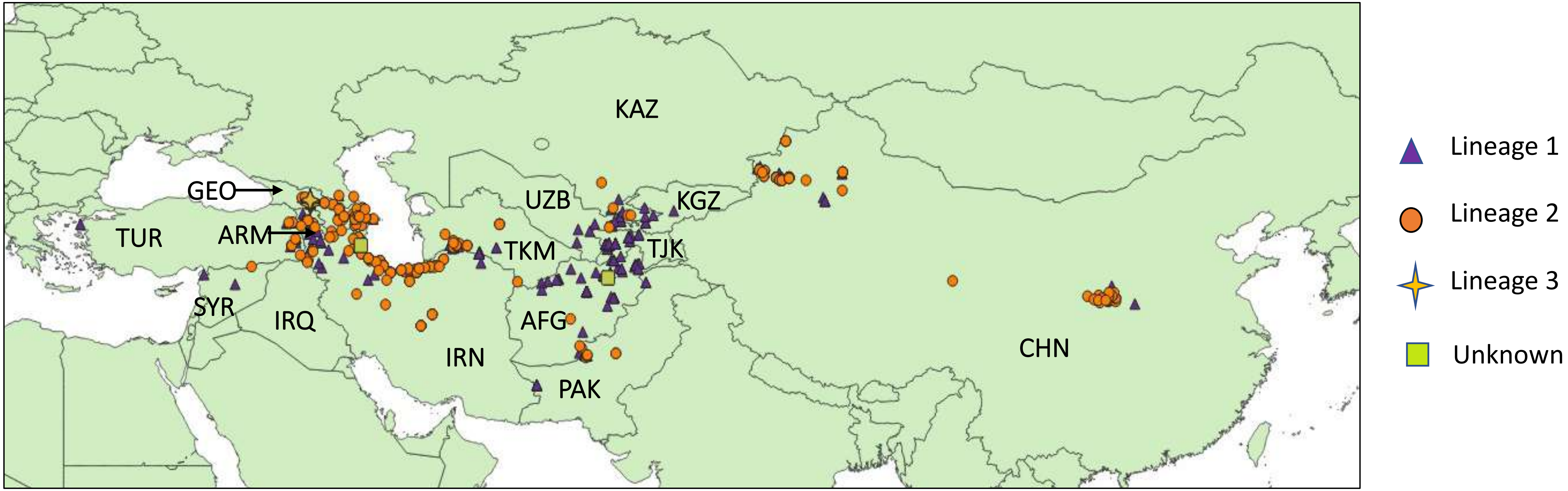
Ae. tauschii panel

675 accessions

OWWC collection (GRU John Innes Centre) Dr. Noam Chayut	242 accessions
Historic collection and SHW D-genome donors (GRU John Innes Centre) Dr. Noam Chayut	84 accessions
Tajikistan collection (Tajikistan National Academy of Sciences) Prof. Firuza Nasyrova	84 accessions
WGRC collection (WGRC, Kansas State University) Dr. Jon Raupp	178 accessions
Azerbaijan collection (Azerbaijan National Academy of Sciences) Prof. Samira Rustamova	20 accessions
PGRC collection (Plant Genetic Resources of Canada) Dr. Dallas Kessler	34 accessions
Pakistan collection (Quaid-i-Azam University) Prof. Awais Rasheed	33 accessions

Future additions: Iranian collection (Ilam University) Prof. Ali Mehrabi >50 accessions

Geographical distribution of the *Aegilops tauschii* diversity panel



The *Aegilops tauschii* diversity panel

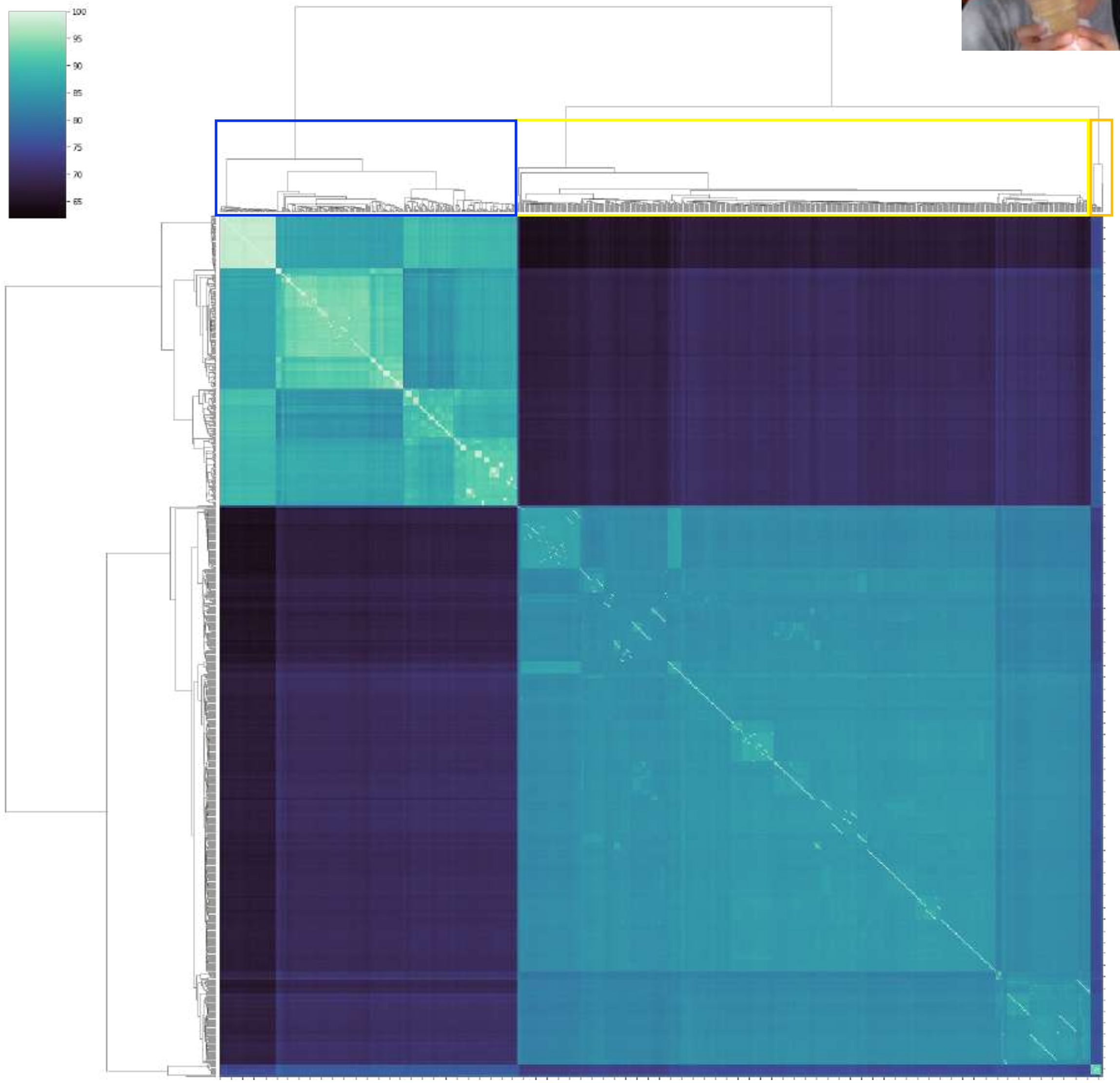
✓ 621/675 accessions -> Illumina WGS sequencing PE150

~7.5x coverage

\+ Public sequencing data for 324 accessions

(Zhou et al. 2021. *Nature Plants*; Zhao et al. 2023. *Nature Plants*)

Lineage	Total accessions	Non-redundant accessions	Genetic redundancy
1	667	352	**47%**
2	270	153	**43%**
3	8	6	**25%**
	945	511	

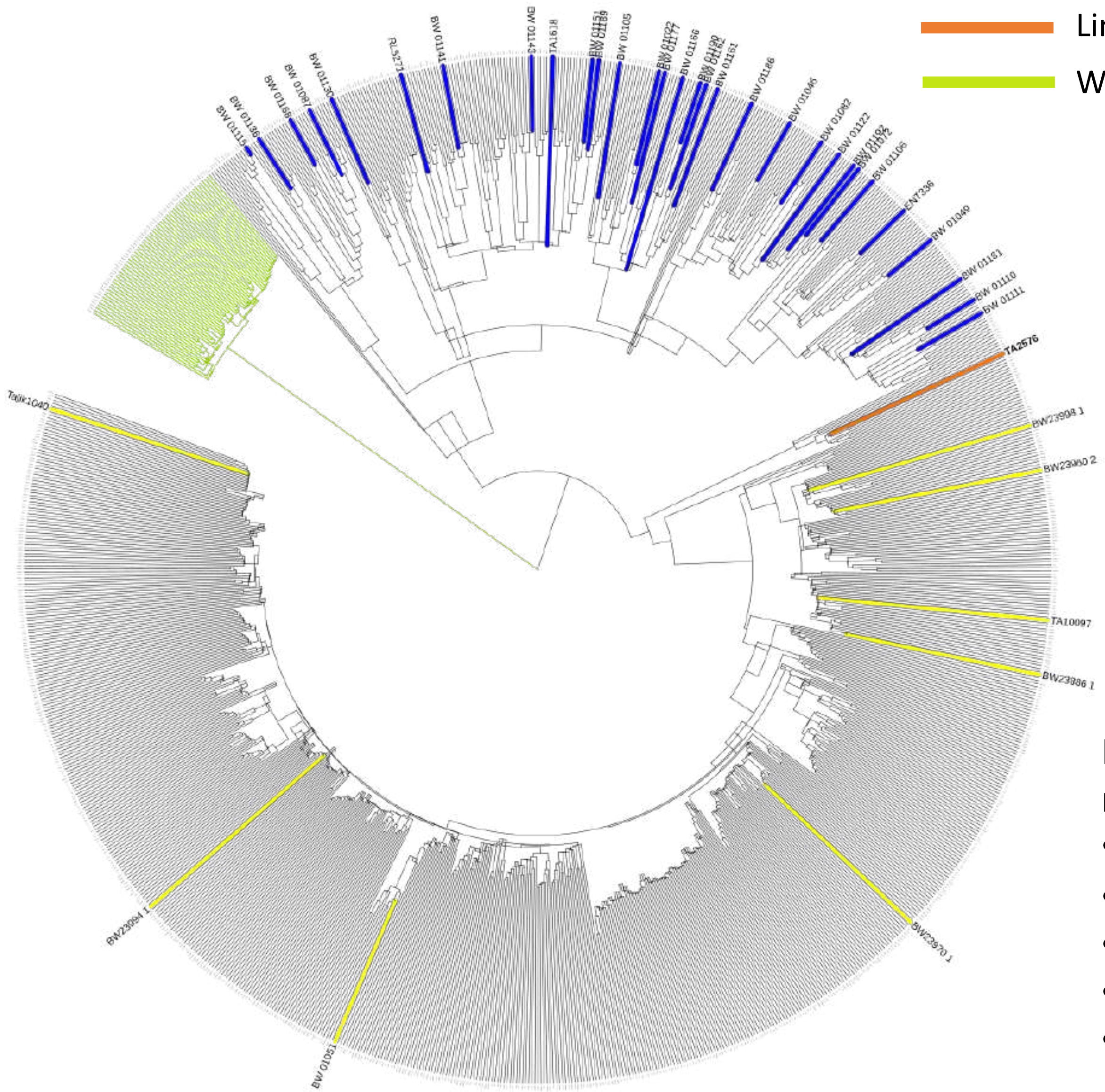


The pangenome of *Aegilops tauschii*

44 accessions

Strategic traits of interest
Disease and pest resistance
Starch
Yield
Salinity tolerance
D-genome donors of SHW
Heat tolerance
Nitrogen use efficiency
Tiller number
Iron content
Agromorphological traits

- 🟨 Lineage 1
- 🟦 Lineage 2
- 🟧 Lineage 3
- 🟩 Wheat D-genome

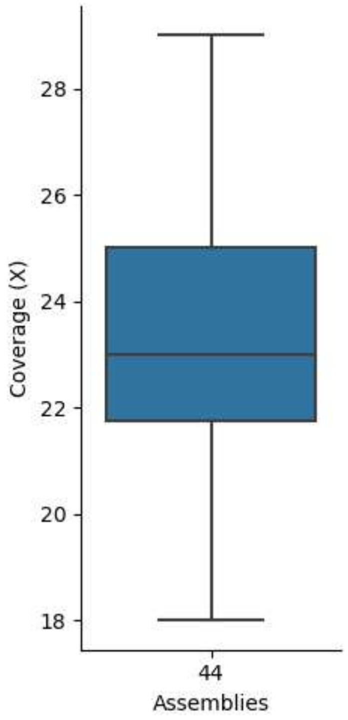


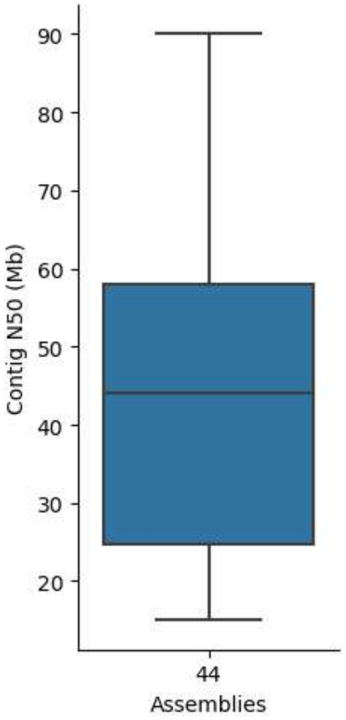
Pangenome accessions not shown:

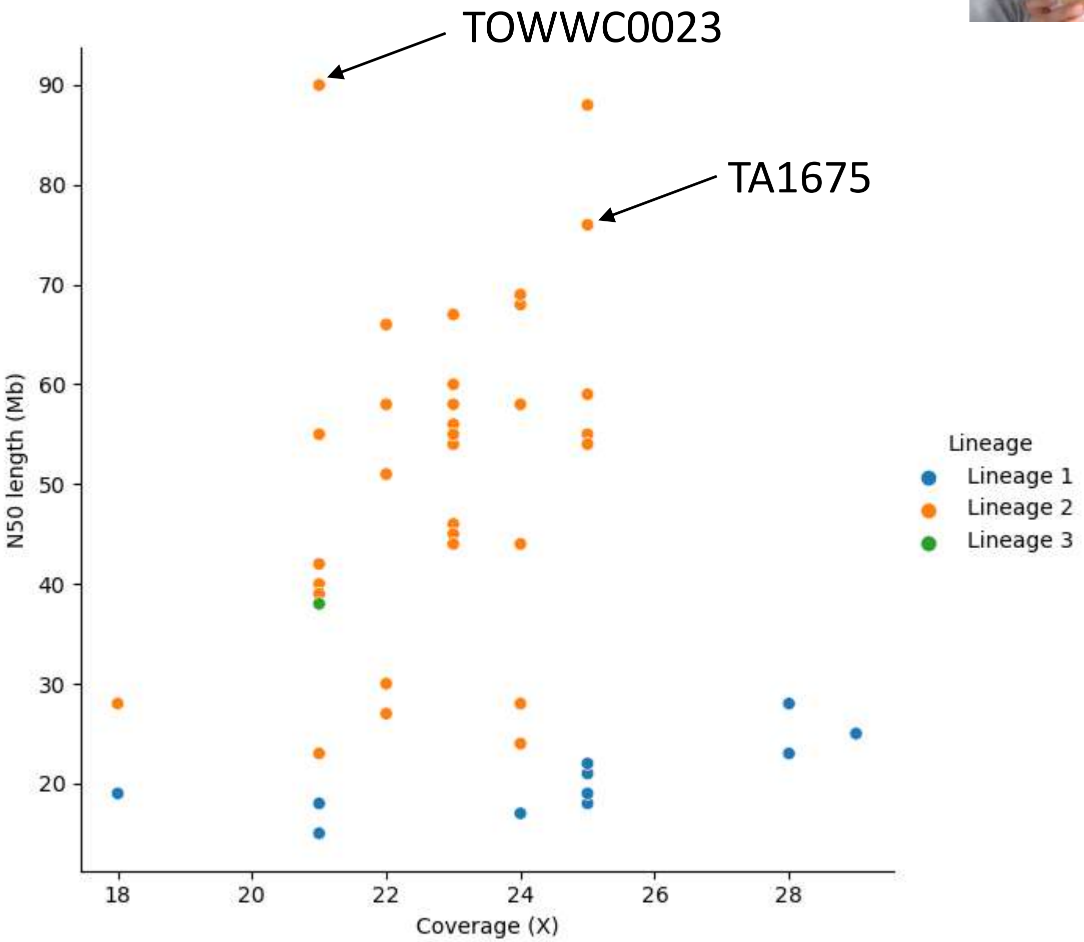
- PI690713 (L2)
- BW_01001 (L2)
- P99-95.1-1 (L1)
- BW_26072 (L1)
- TA10187 (L1)

Contig-level assemblies

- ✓ Primary assemblies generated from PacBIO HiFi reads

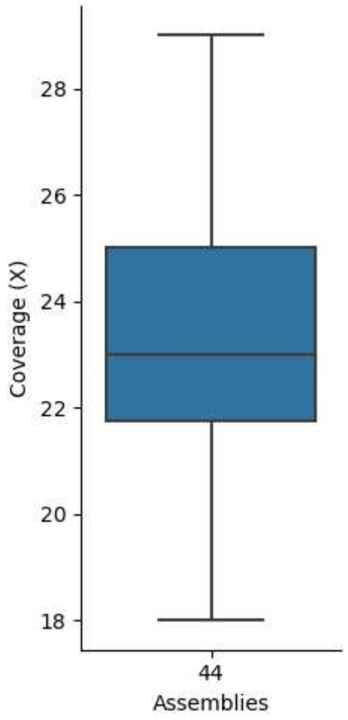


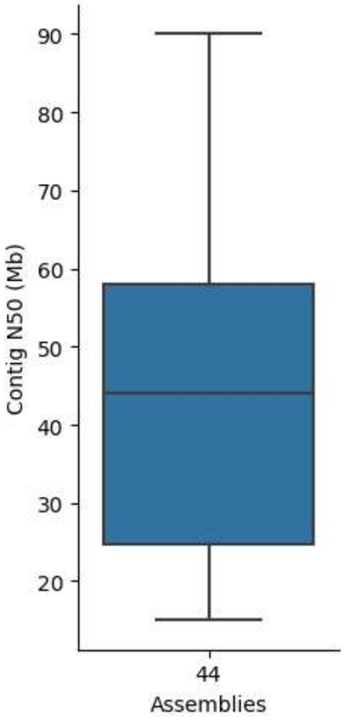


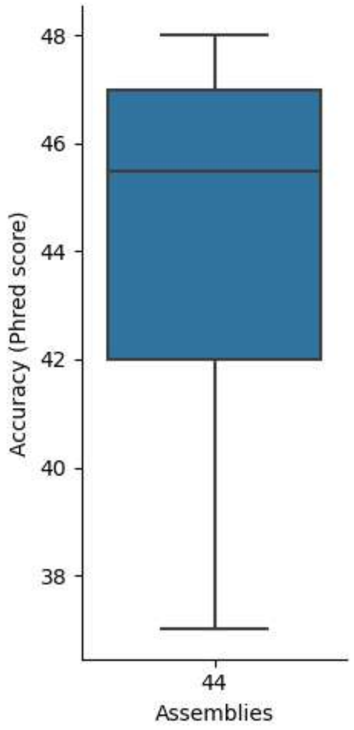


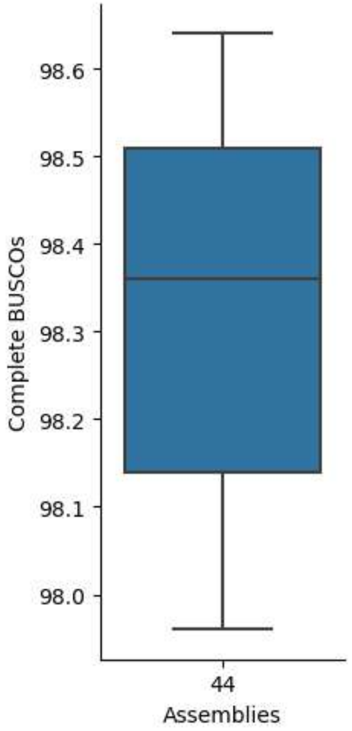
Contig-level assemblies

- ✓ Primary assemblies generated from PacBIO HiFi reads









Chromosome-level assemblies per lineage

Gold Level

- ✓ High-quality assemblies for anchoring pangenome to a lineage-level reference
- ✓ Provide high resolution assemblies for structural variant calling

Assembly	**TA10171_L1**	**TA1675_L2**	**TA2576_L3**	**AL8/78***
Superscaffolds	7	7	7	7
Total assembled length (bp)	4,151,983,908	4,159,914,615	4,245,074,256	4,218,179,011
Unplaced scaffolds	924	884	3,045	109,190
Unplaced assembled length (bp)	45,447,308	53,352,240	121,041,196	123,803,079
Unplaced scaffold N50 (bp)	49,187	58,980	36,307	27,953

*NCBI RefSeq assembly: AL8/78 v5.0 (RefSeq GCF_002575655.2)

Chromosome-level assemblies per lineage

Gold Level

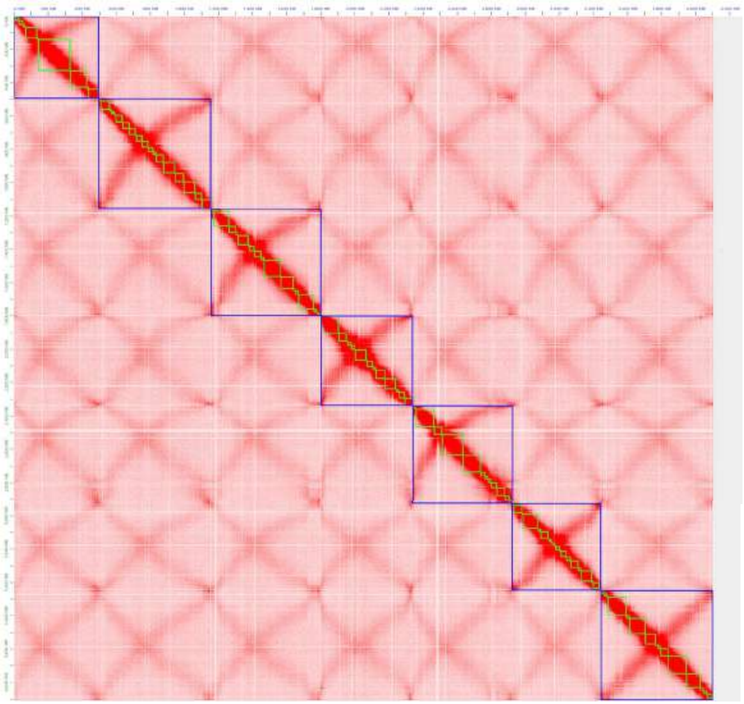
- ✓ PacBIO CCS (HiFi) 67-97x coverage + Illumina HiC 300 M reads
- ✓ HiFi + HiC phased assemblies
- ✓ HiC scaffolding to the chromosome-scale

Assembly	**TA10171_L1**	**TA1675_L2**	**TA2576_L3**	**AL8/78***
Superscaffolds	7	7	7	7
Total assembled length (bp)	4,151,983,908	4,159,914,615	4,245,074,256	4,218,179,011
Unplaced scaffolds	924	884	3,045	109,190
Unplaced assembled length (bp)	45,447,308	53,352,240	121,041,196	123,803,079
Unplaced scaffold N50 (bp)	49,187	58,980	36,307	27,953

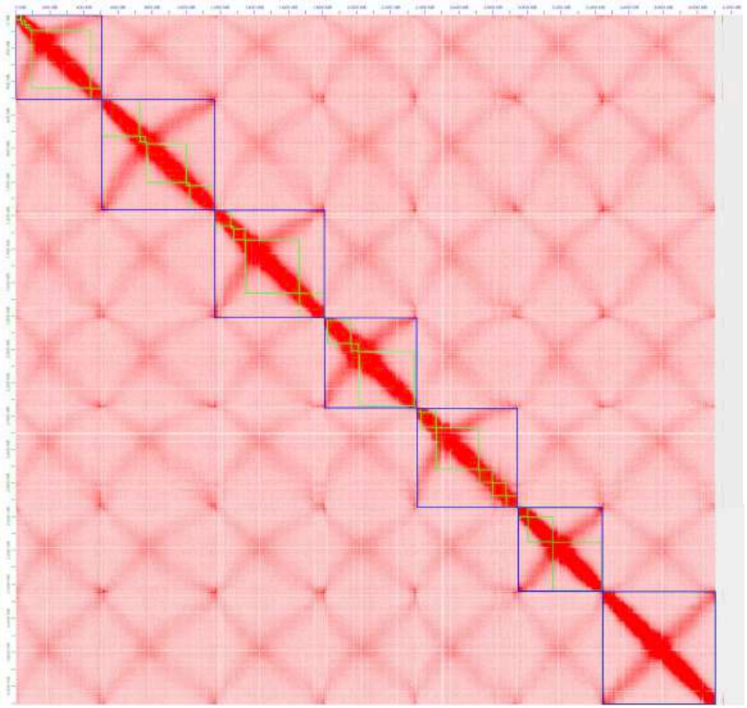
*NCBI RefSeq assembly: AL8/78 v5.0 (RefSeq GCF_002575655.2)

Chromosome-level assemblies per lineage: HiC scaffolding

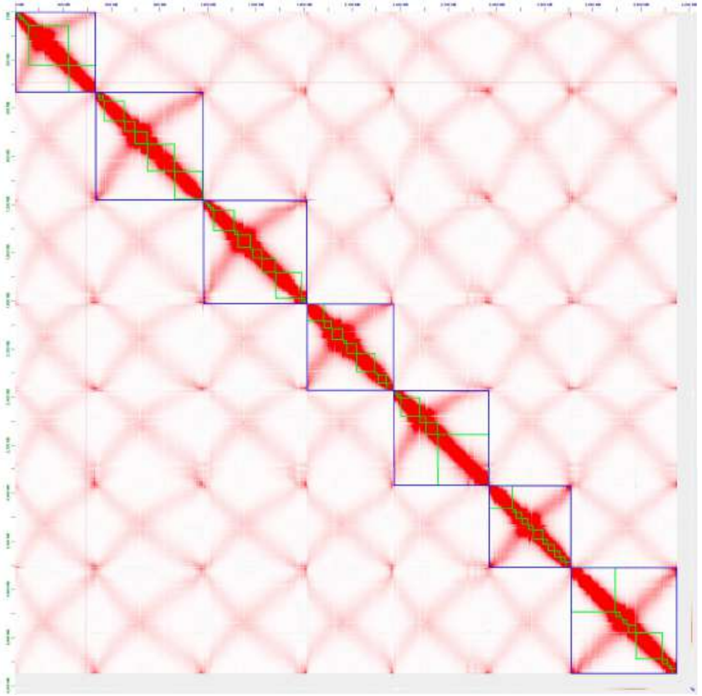
L1: TA10171
77x coverage
Contig N50 = 53 Mb



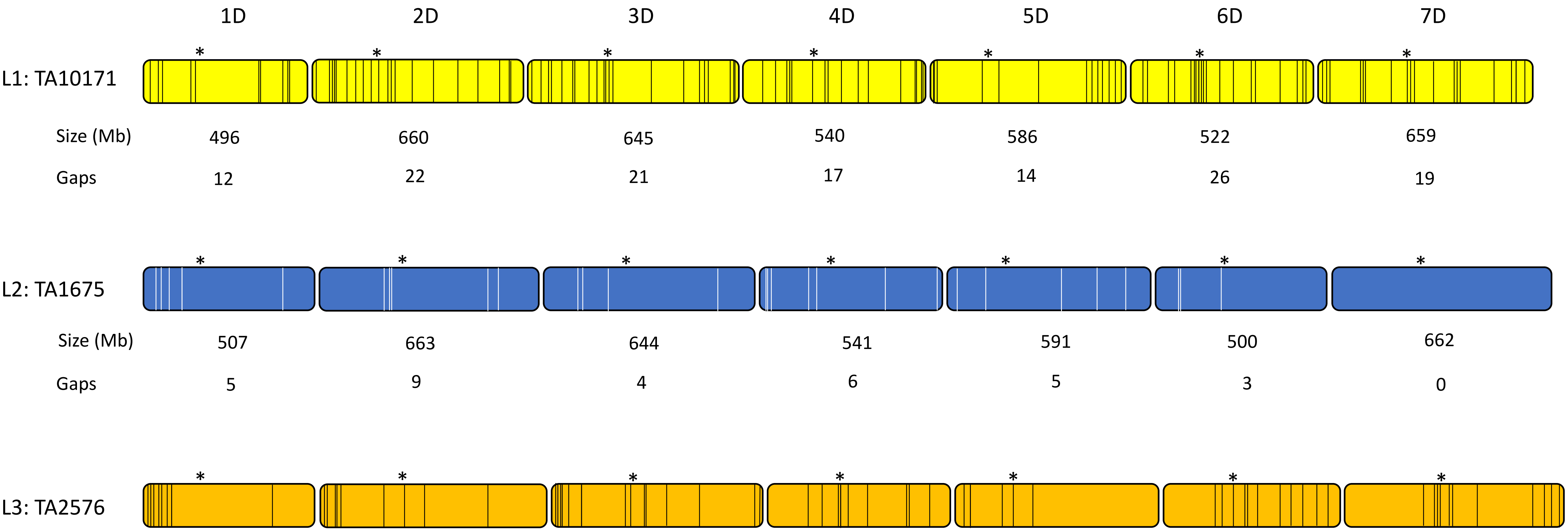
L2: TA1675
97x coverage
Contig N50 = 221 Mb



L3: TA2576
67x coverage
Contig N50 = 117 Mb



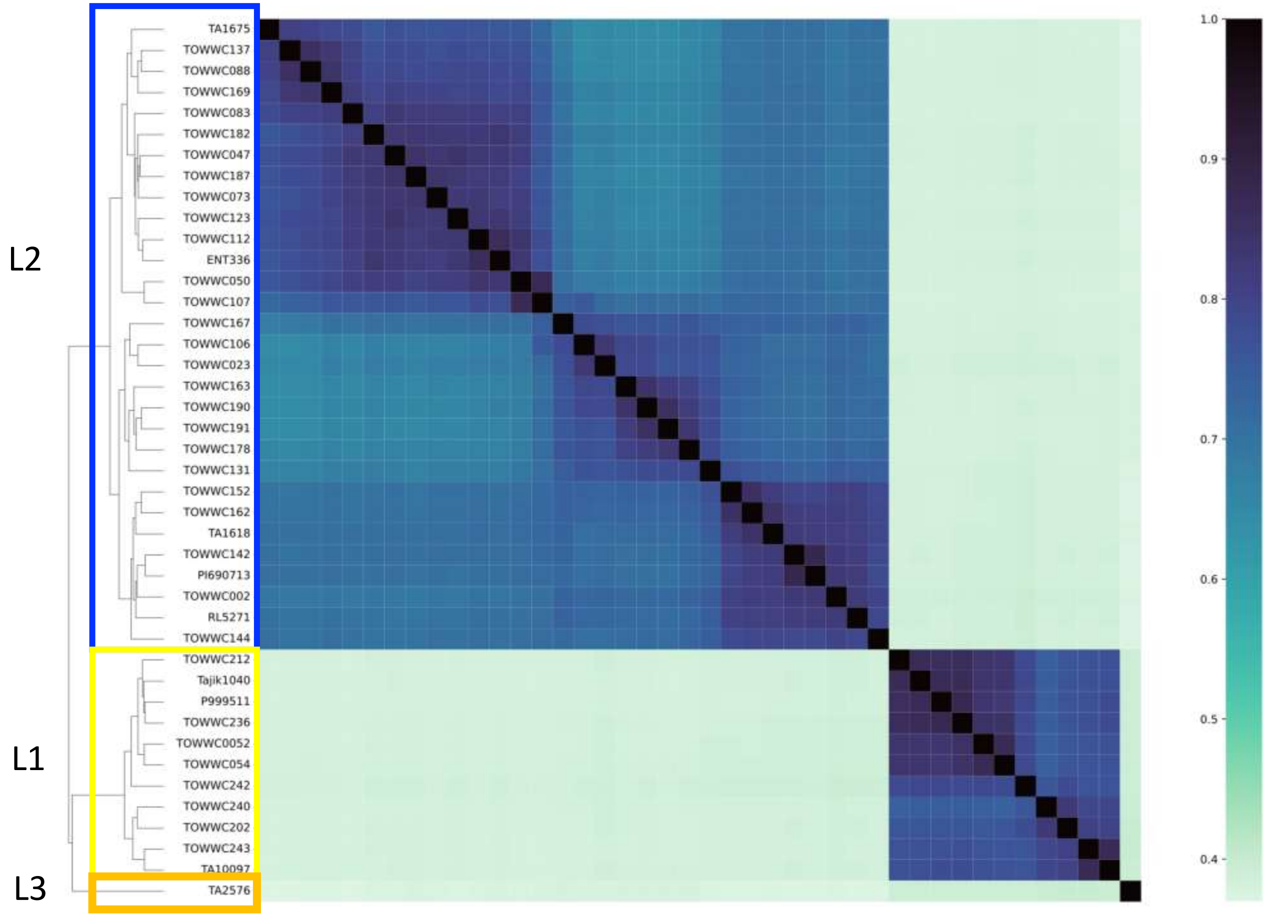
Inter-lineage variation in number and location of gaps

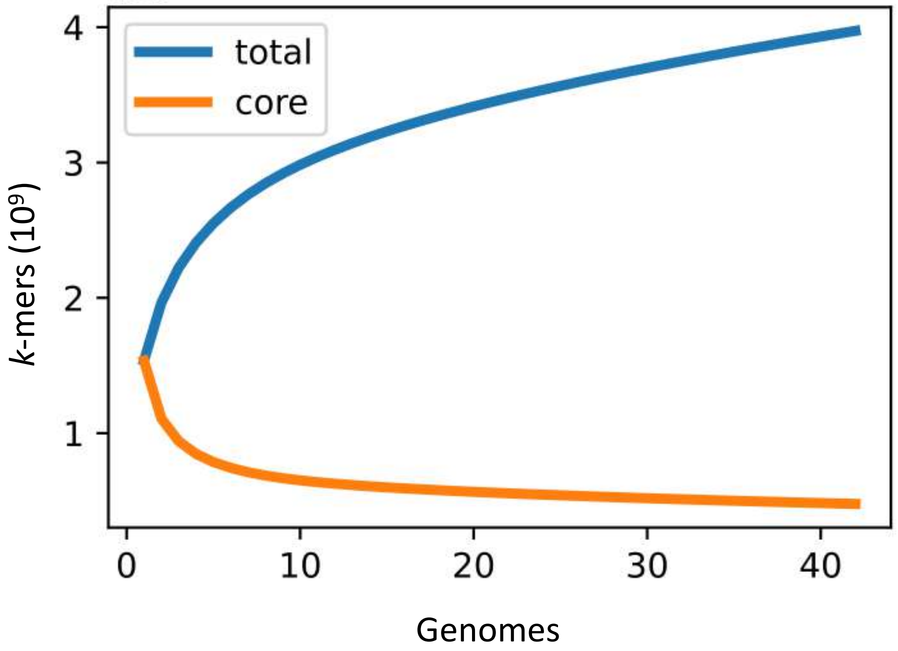


Lineage		1D	2D	3D	4D	5D	6D	7D
L1: TA10171	Size (Mb)	496	660	645	540	586	522	659
	Gaps	12	22	21	17	14	26	19
L2: TA1675	Size (Mb)	507	663	644	541	591	500	662
	Gaps	5	9	4	6	5	3	0
L3: TA2576	Size (Mb)	501	674	646	540	594	511	659
	Gaps	10	11	15	11	7	13	12

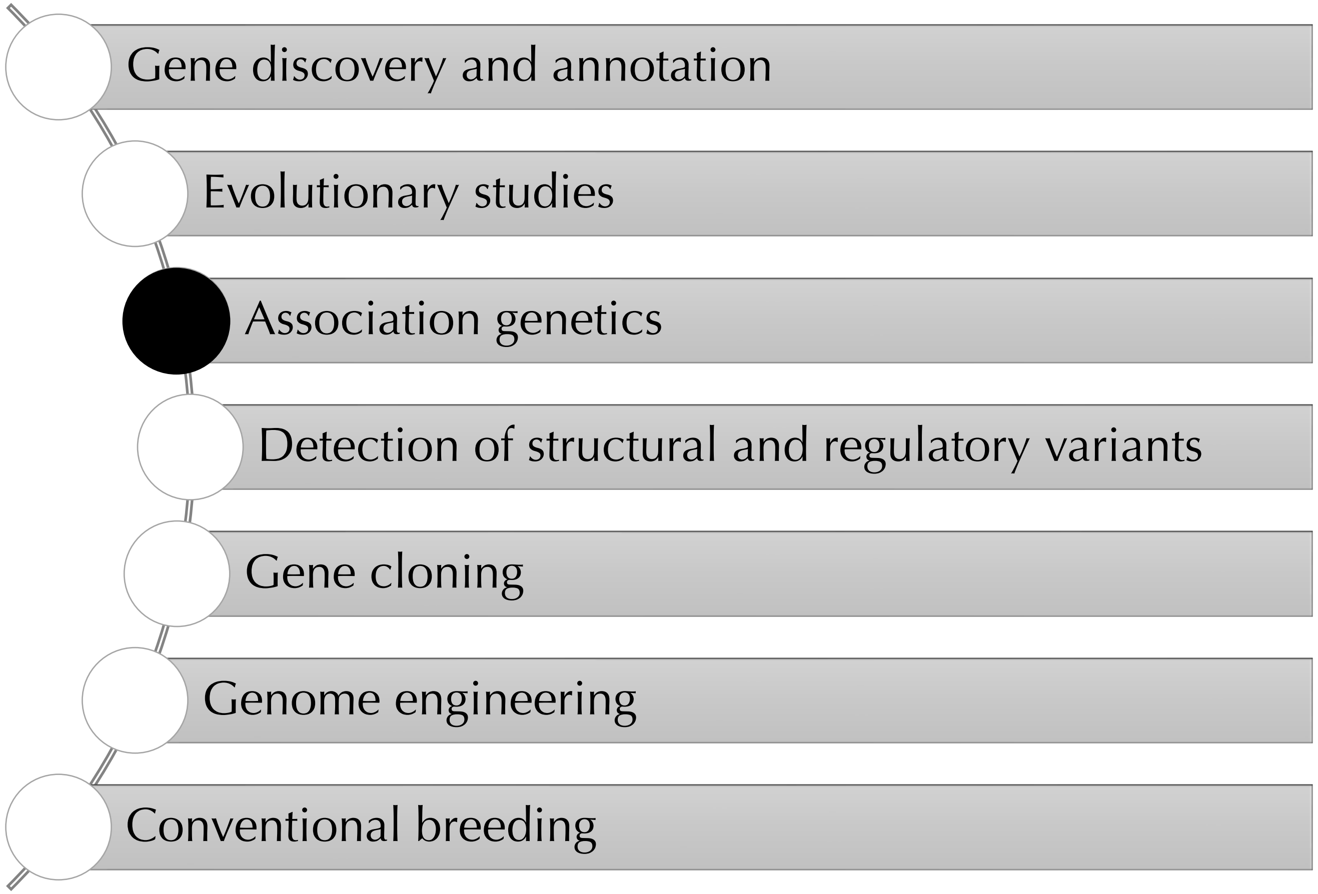
* = centromeric region

The pangenome of *Aegilops tauschii*





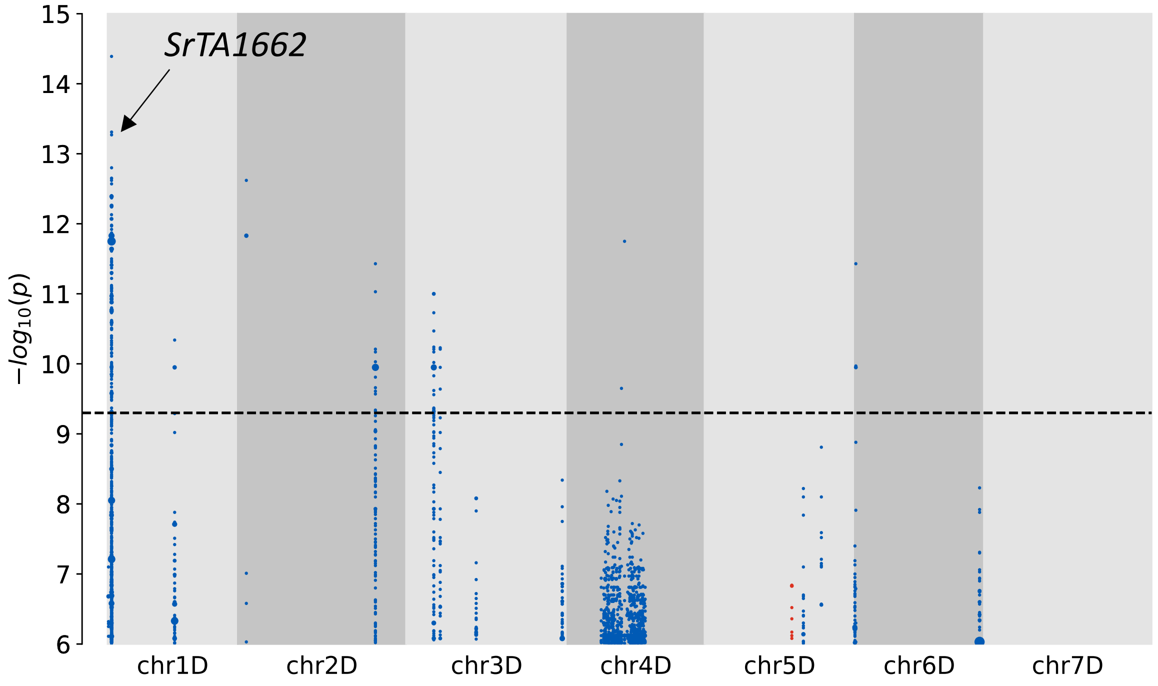
Applications of the *Ae. tauschii* diversity panel and pangenome



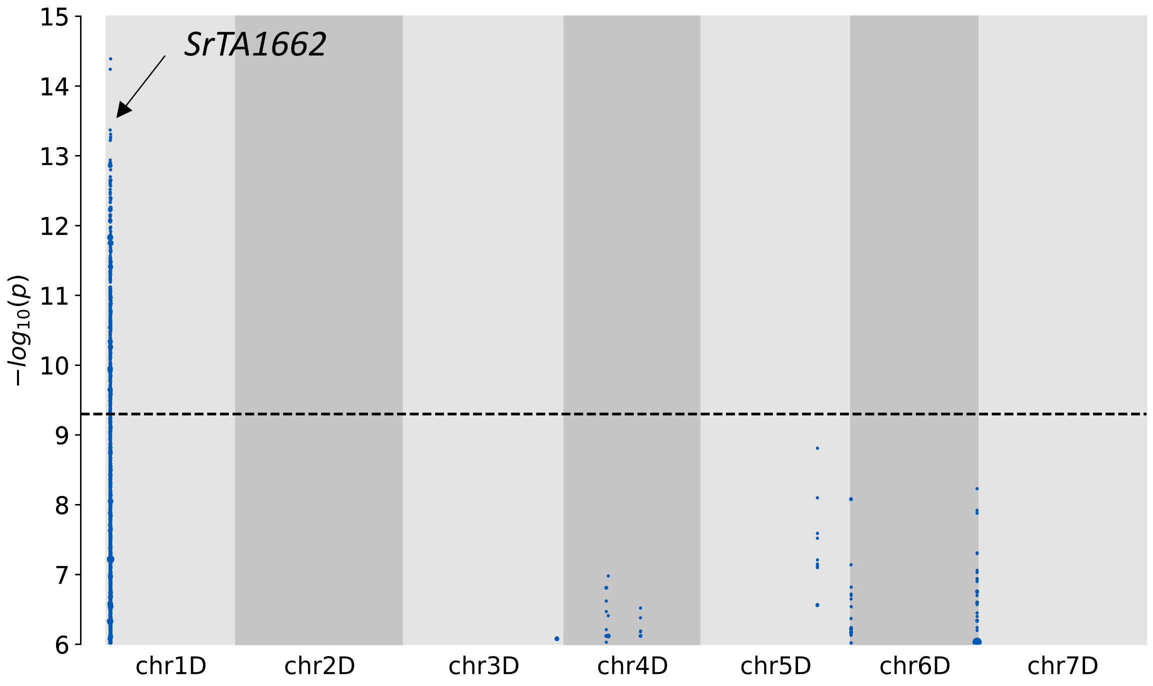
- ○ Gene discovery and annotation
- ○ Evolutionary studies
- ● Association genetics
- ○ Detection of structural and regulatory variants
- ○ Gene cloning
- ○ Genome engineering
- ○ Conventional breeding

High-quality reference genome assemblies improve association genetics

Associated *k*-mers for resistance to *Puccinia graminis* f. sp. *tritici* race QTHJC in accession TA1662



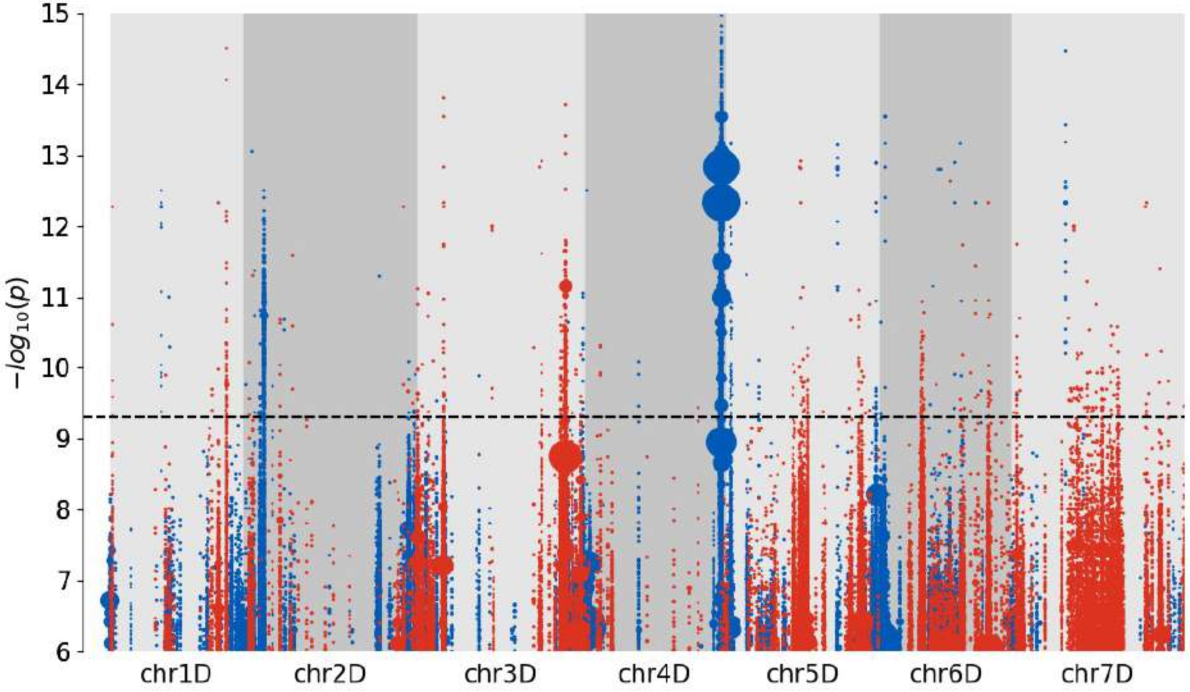
Low quality assembly (contig N50 of 196 kb)



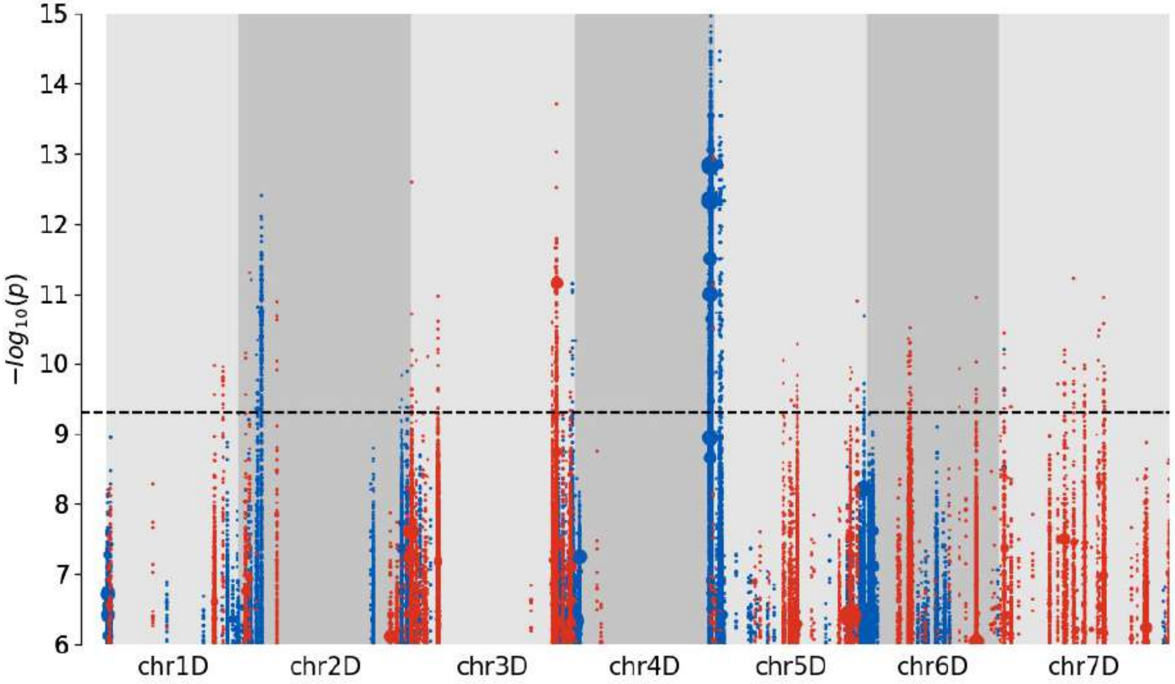
High-quality assembly (contig N50 of 58.21 Mb)

High-quality reference genome assemblies improve association genetics

Associated *k*-mers for trichome number in accession TA1662



Low quality assembly (contig N50 of 196 kb)



High-quality assembly (contig N50 of 58.21 Mb)

Applications of the *Ae. tauschii* diversity panel and pangenome

- ● Gene discovery and annotation
- ○ Evolutionary studies
- ○ Association genetics
- ○ Detection of structural and regulatory variants
- ● Gene cloning
- ○ Genome engineering
- ○ Conventional breeding

Unraveling new disease resistance genes

Theor Appl Genet (2013) 126:1179–1188
DOI 10.1007/s00122-013-2045-5

ORIGINAL PAPER

Simultaneous transfer, introgression, and genomic localization of genes for resistance to stem rust race TTKSK (Ug99) from *Aegilops tauschii* to wheat

Eric L. Olson · Matthew N. Rouse ·
Michael O. Pumphrey · Robert L. Bowden ·
Bikram S. Gill · Jesse A. Poland

Unraveling new disease resistance genes

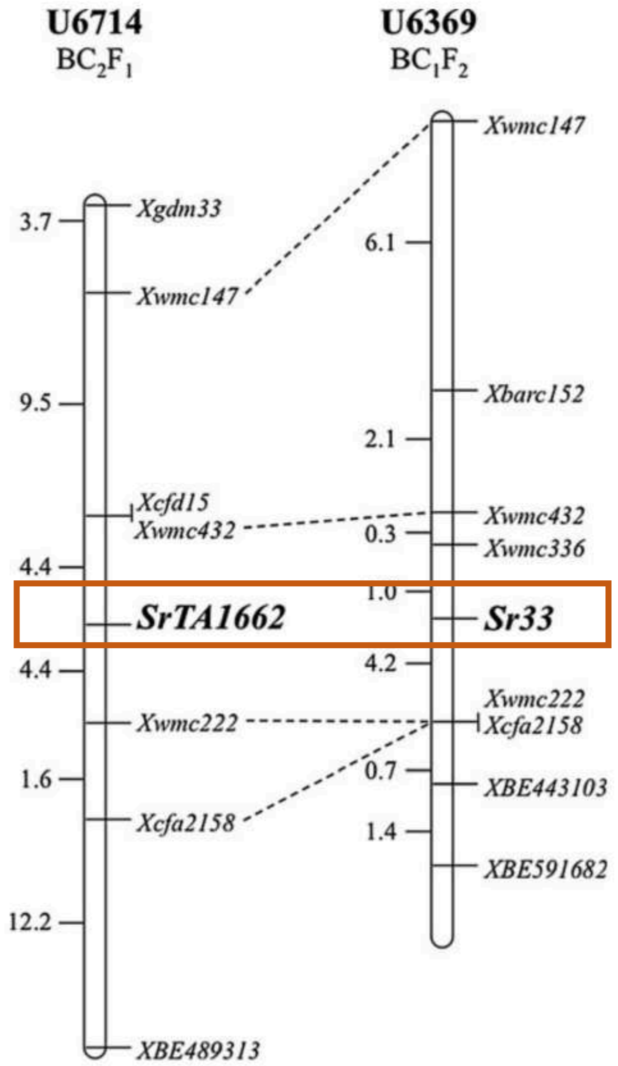


Fig. 1 Genetic linkage maps showing the position of *SrTA1662* and *Sr33* from PI 603225 on 1DS


nature biotechnology

ARTICLES

https://doi.org/10.1038/s41587-021-01058-4

Check for updates

OPEN

Population genomic analysis of *Aegilops tauschii* identifies targets for bread wheat improvement

SrTA1662 does not confer resistance to stem rust isolate Ug99!

Unraveling new disease resistance genes

Predicted NLR genes

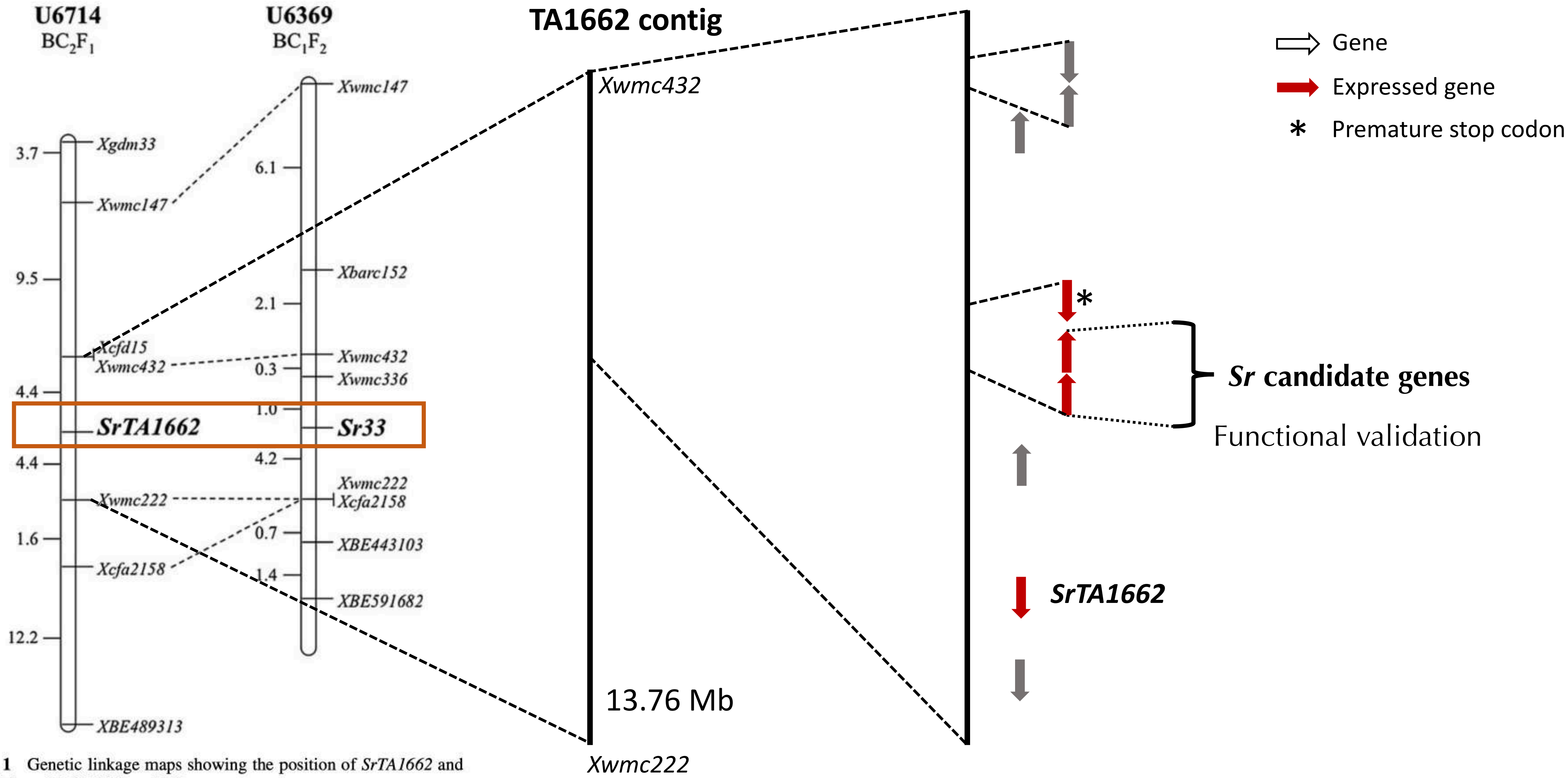


Fig. 1 Genetic linkage maps showing the position of *SrTA1662* and *Sr33* from PI 603225 on 1DS

SrTA1662 and *Sr33*: alleles or paralogs?

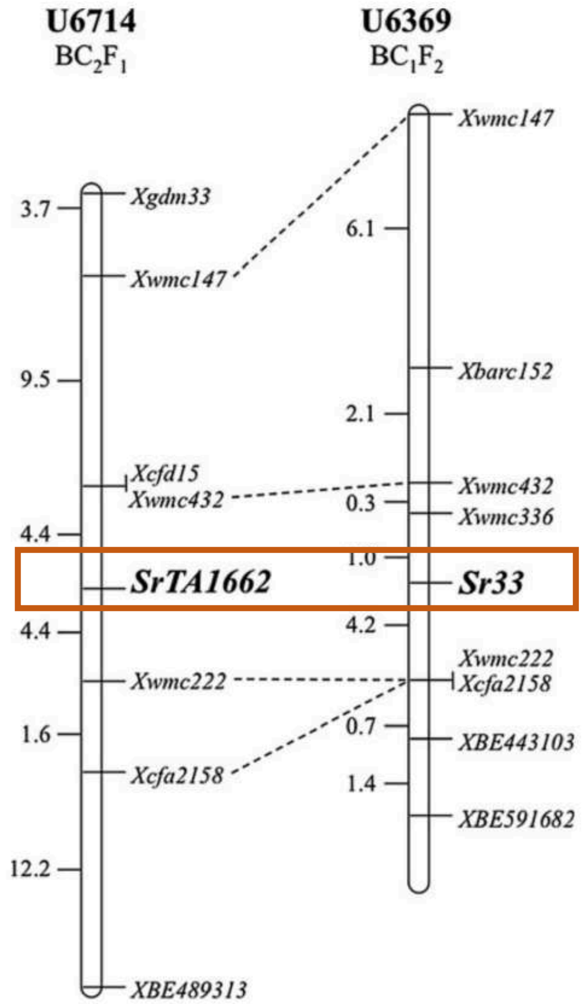
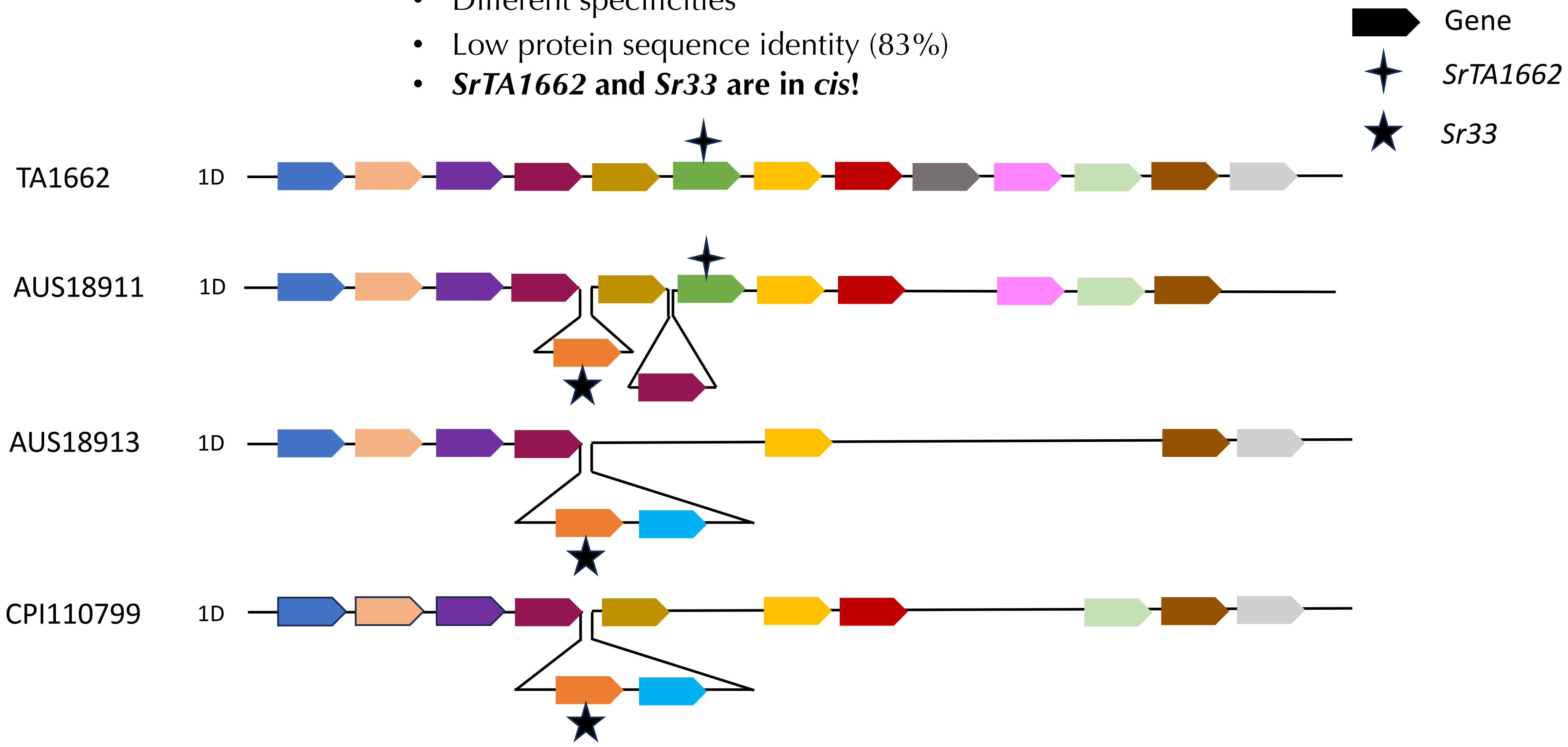


Fig. 1 Genetic linkage maps showing the position of *SrTA1662* and *Sr33* from PI 603225 on 1DS

Olson et al. *Theor Appl Genet*. 2013

SrTA1662 and *Sr33:* alleles or paralogs?

- Different specificities
- Low protein sequence identity (83%)
- ***SrTA1662* and *Sr33* are in *cis*!**



Applications of the *Ae. tauschii* diversity panel and pangenome

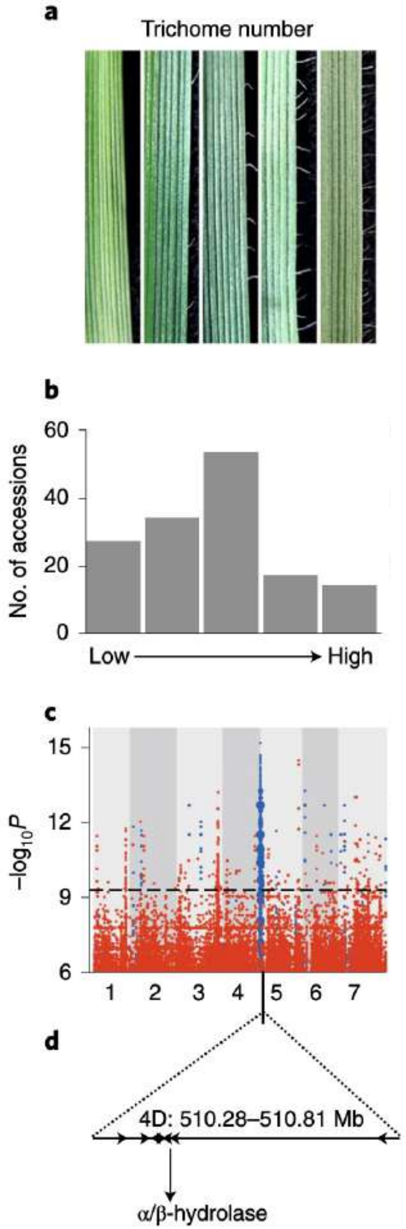
- ● Gene discovery and annotation
- ○ Evolutionary studies
- ● Association genetics
- ○ Detection of structural and regulatory variants
- ● Gene cloning
- ○ Genome engineering
- ○ Conventional breeding

Dissecting leaf trichome regulation

Phenotypic trait or marker used to characterize wheat:

- ✓ Easy observation
- ✓ High heritability
- ✓ Linked to pest resistance and abiotic stress

Credit: Eye of Science/Science Photo Library

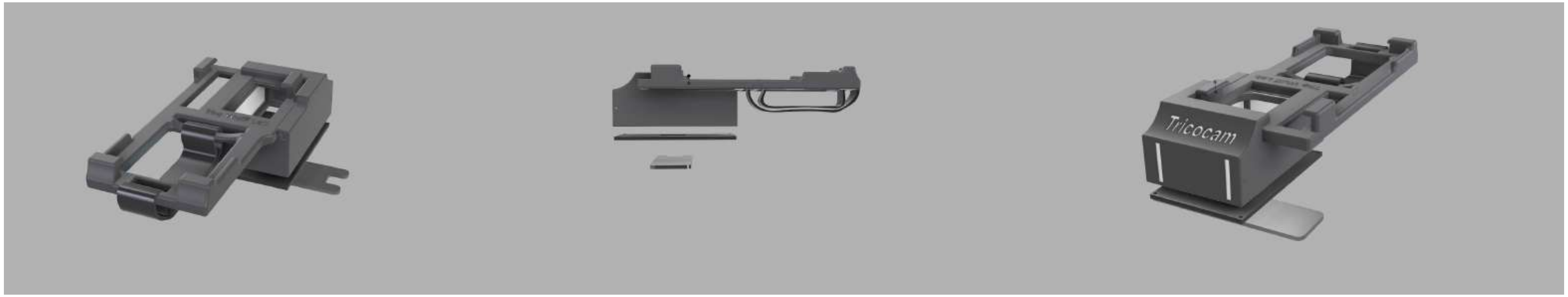


Gaurav et al. Nature Biotechnology. 2022

The trichome phenotyping problem

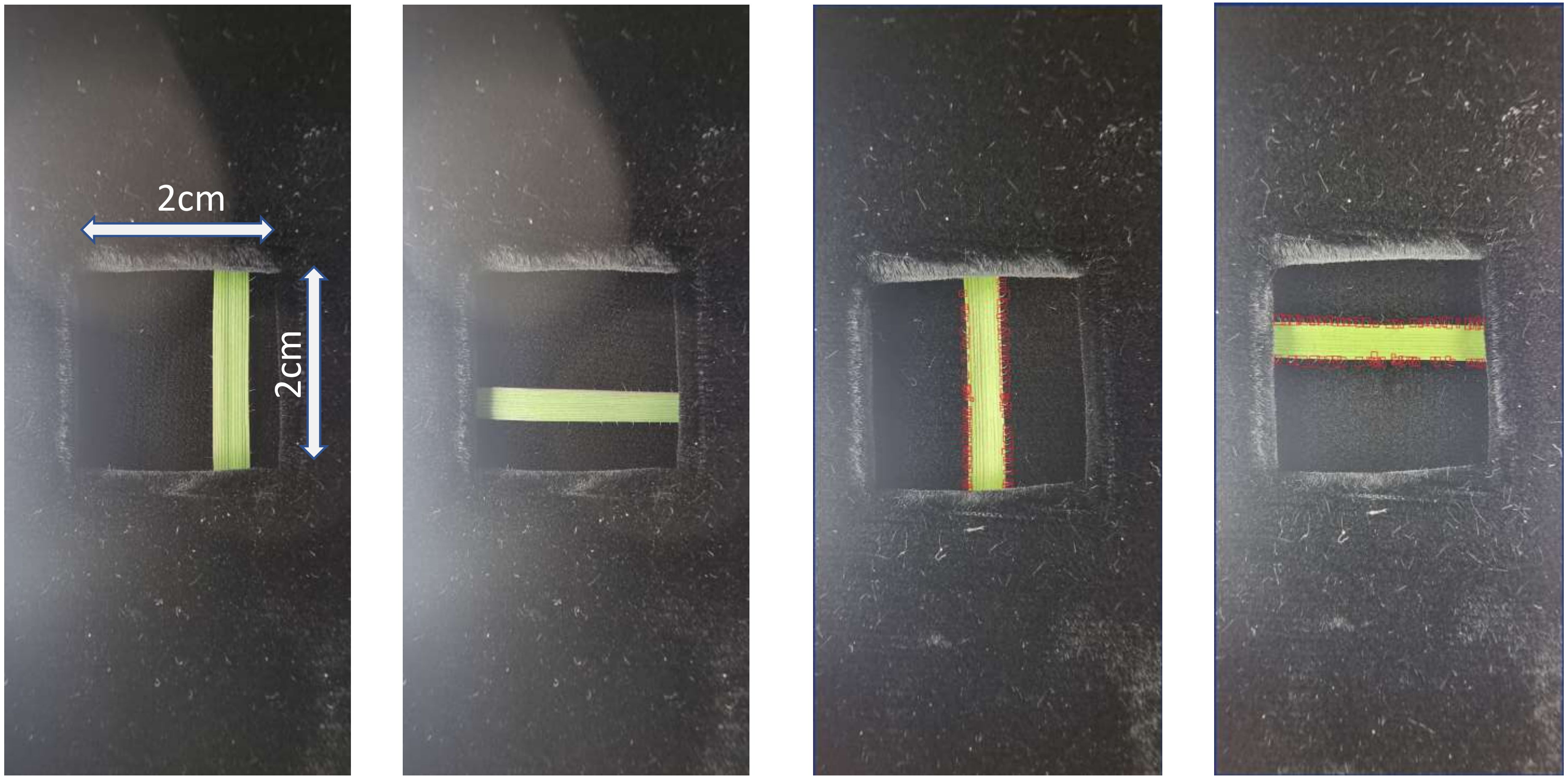
The Tricocam

3D design and optimization

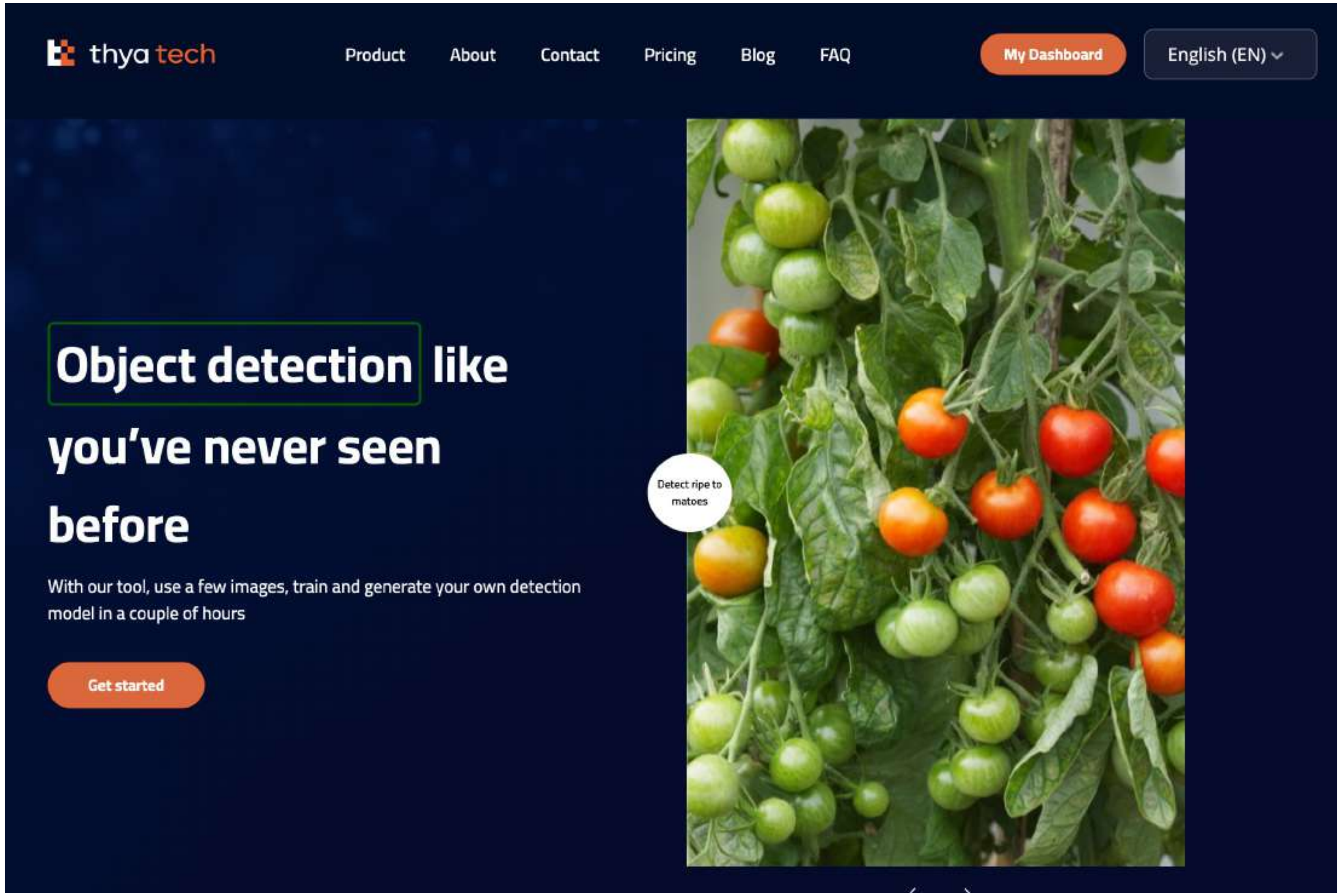


“Final” product

Trichome counting automation process – Thya Technology



Trichome counting automation process – Thya Technology

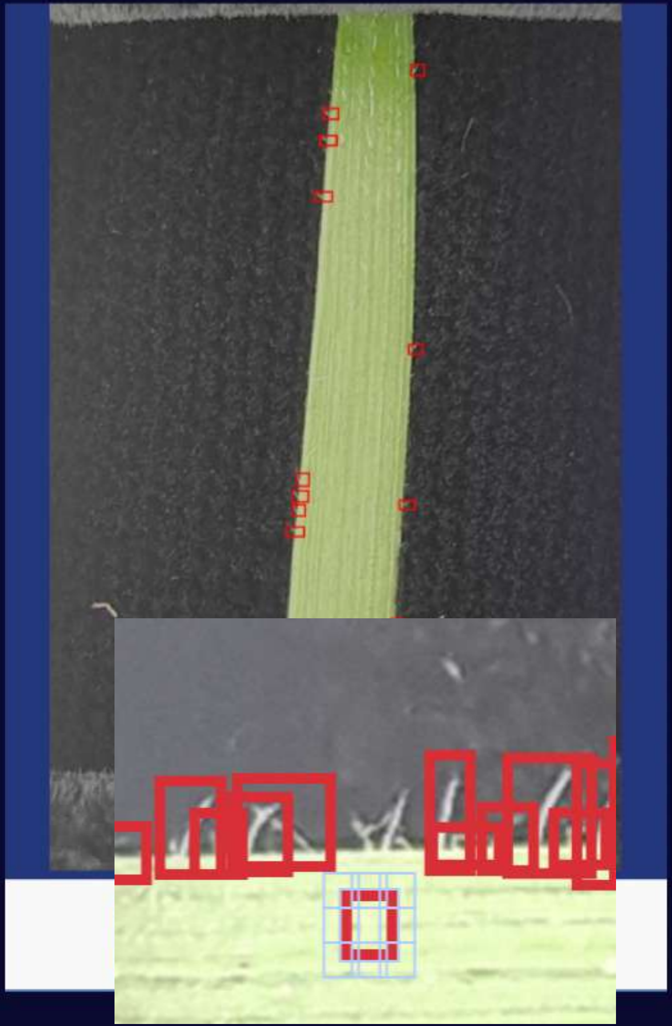


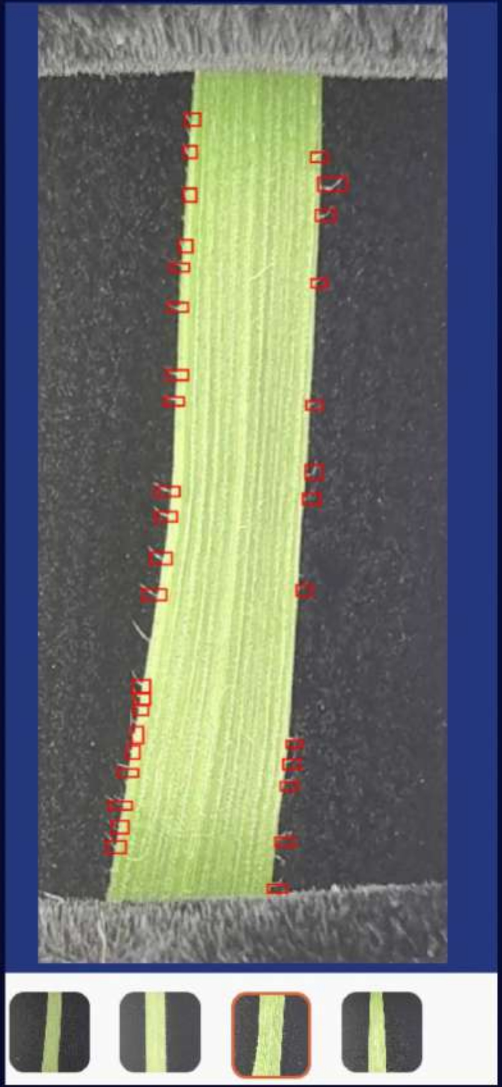
Object detection like you've never seen before

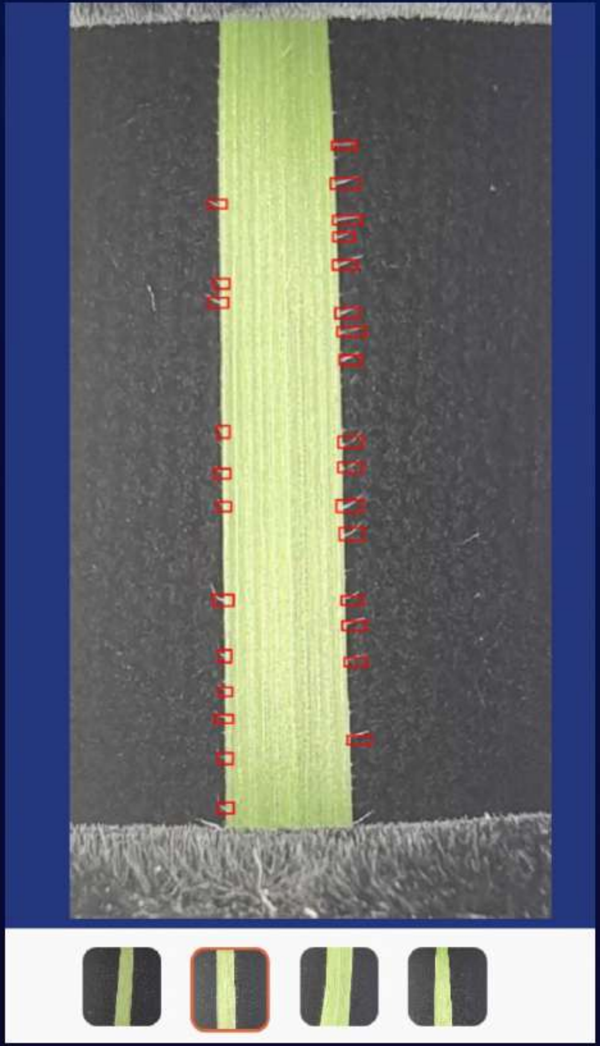
With our tool, use a few images, train and generate your own detection model in a couple of hours

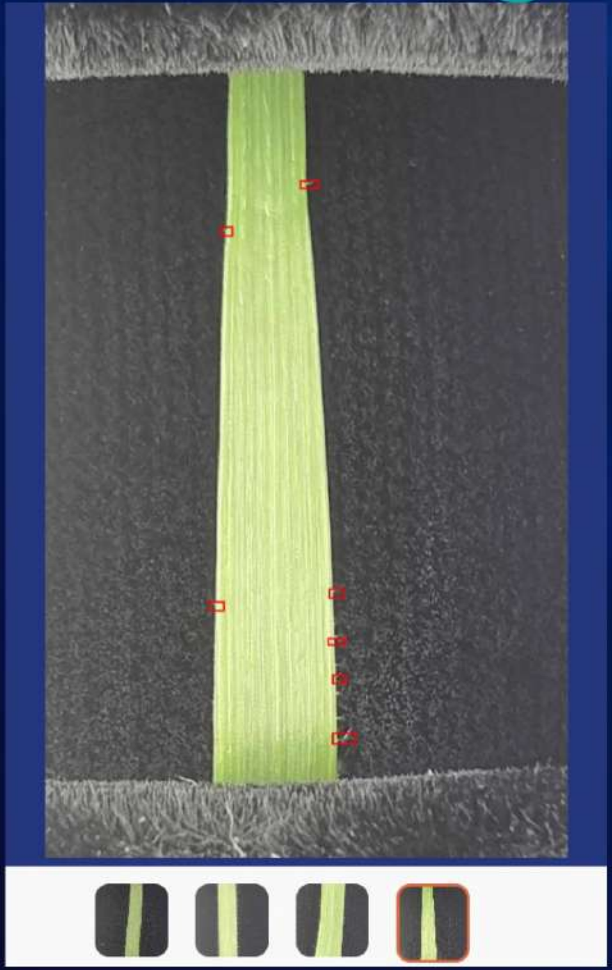
Trichome detection - Results

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

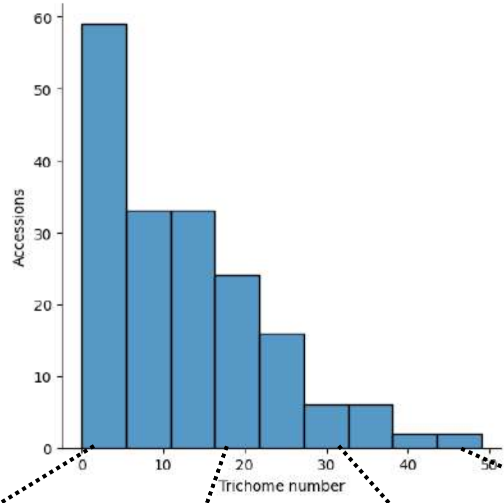


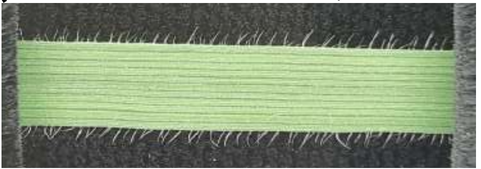


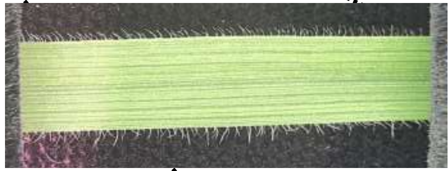




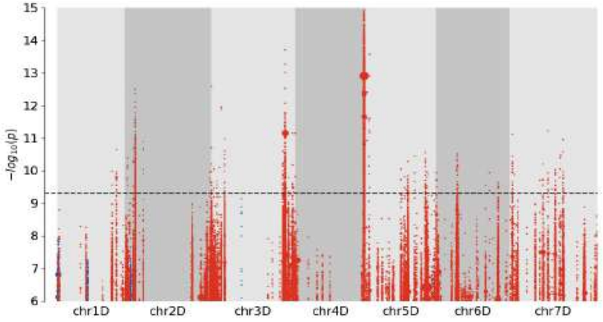
Dissecting leaf trichome regulation in *Ae. tauschii*



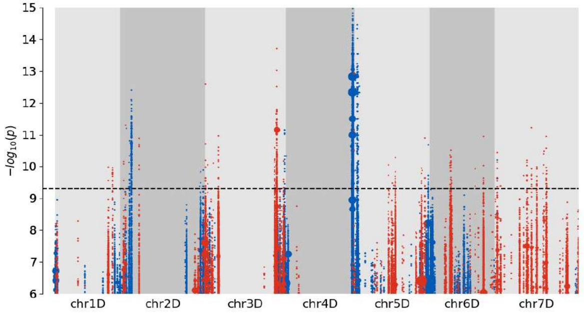




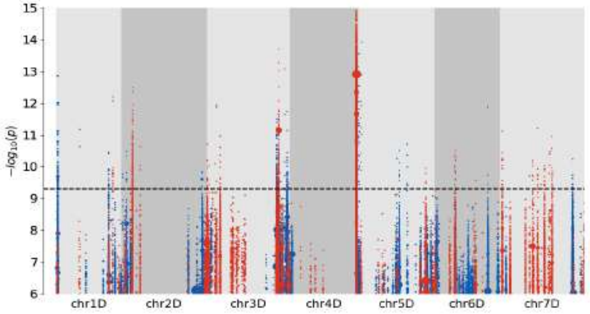
TOWWC0137



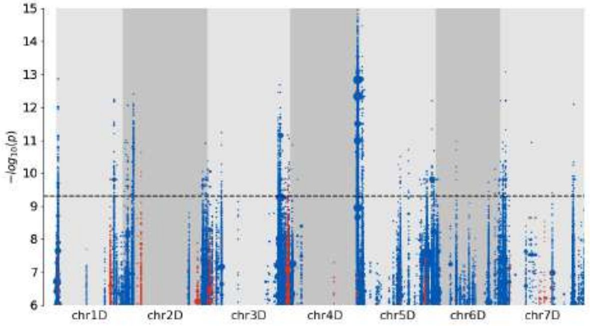
TOWWC0112



TOWWC0178



TOWWC0107



Take-home messages

- *Aegilops tauschii* diversity panel of 511 non-redundant sequenced accessions
- *Ae. tauschii* pangenome with 44 high-quality genome assemblies
- Chromosome-scale reference assemblies for each lineage to anchor the pangenome
- Resource for enabling high-resolution genetic association studies, gene cloning, detection of structural variation, genome engineering, etc.

Acknowledgments

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

The Wulff Lab (KAUST)

- Brande Wulff
- Catherine Gardener
- Guotai Yu
- Jing Lu
- Radim Svačina
- Ana Belen Perera
- Francisco Ayala
- Renjie Chen
- Amer Alrudayan
- Karthick Gajendiran
- Luby Wang
- Konstanze Laquai

Chia-Yi Hu

Acknowledgments

- Simon Krattinger
- Jesse Poland
- Emile Cavalet-Giorsa
- Michael Abrouk
- Hanin Ahmed
- Naveenkumar Athiyannan
- Luis Rivera
- Salim Bougouffa
- Adil Salhi
- Xin Gao

جامعة الملك عبدالله
للعلوم والتقنية
King Abdullah University of
Science and Technology

Center for Desert Agriculture (KAUST)

CDA
CENTER FOR DESERT
AGRICULTURE

Thya Technology

- Justine Braguy

John Innes Centre (JIC)

- Burkhard Steuernagel
- Cristobal Uauy
- Jesus Quiroz-Chavez

The Open Wild Wheat Consortium

OPEN
WILD
WHEAT

Tajikistan National Academy of Sciences

- Firuza Nasyrova

Azerbaijan National Academy of Sciences

- Samira Rustamova

Quaid-i-Azam University

- Awais Rasheed

Germplasm Resource Unit, JIC

- Noam Chayut

Wheat Genetics Resource Center, KSU

- Jon Raupp

Plant Gene Resource of Canada (PGRC)

University of the Free State

- Willem Boshoff

The Wulff Lab

www.wulfflab.org

جامعة الملك عبدالله للعلوم والتقنية
King Abdullah University of Science and Technology

CDA Center for Desert Agriculture

https://cda.kaust.edu.sa/

OPEN WILD WHEAT

www.openwildwheat.org
contact_owwc@openwildwheat.org

Thank you

Illustration by Robyn Palescandolo