Allele Mining with the Soybean Allele Catalog Tool

Kristin Bilyeu

Plant Genetics Research Unit

USDA/ARS

University of Missouri

Columbia, MO 65211 USA

kristin.bilyeu@usda.gov

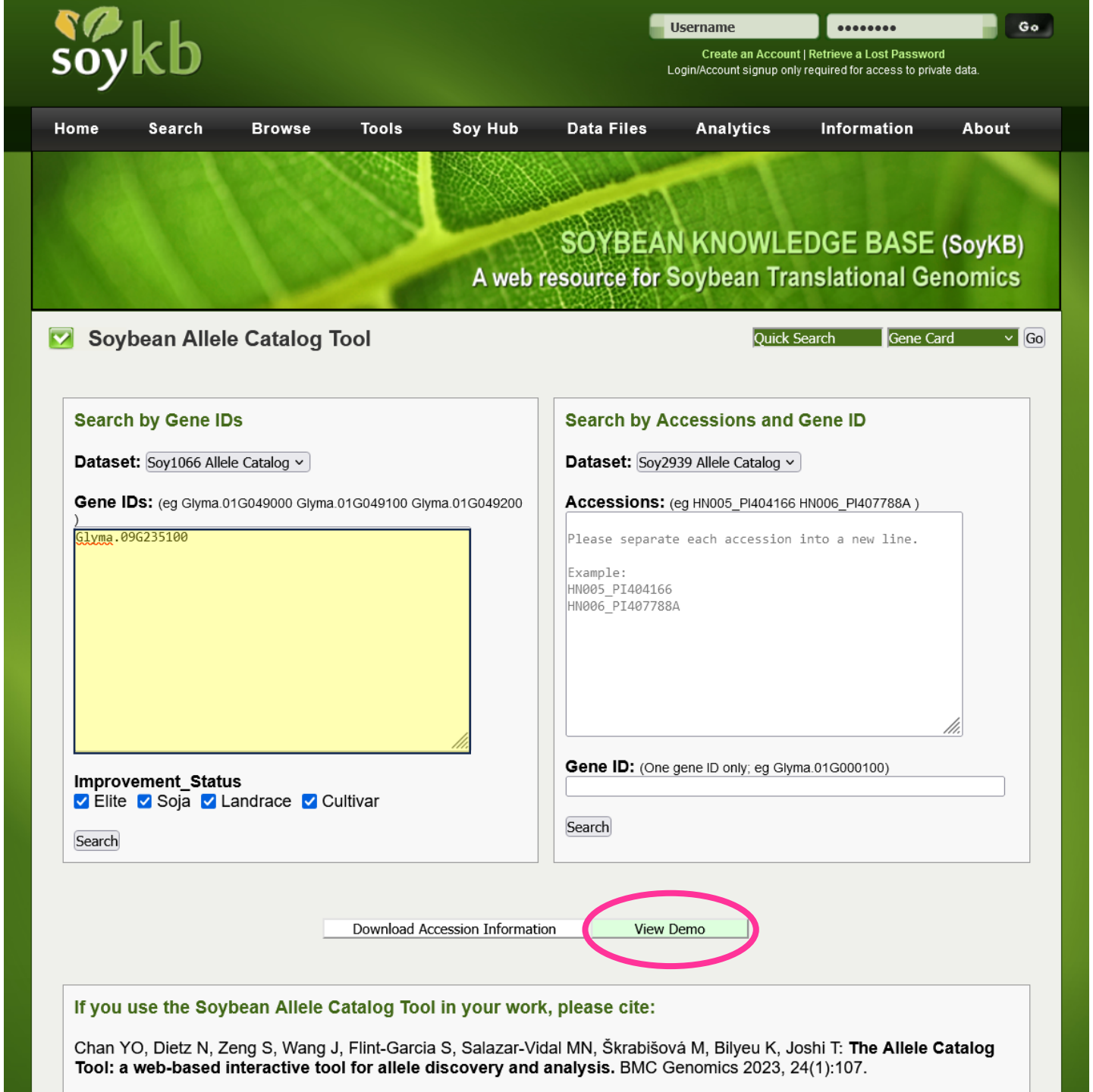
February 14, 2024

Soybean Allele Catalog Tool

Based on whole genome sequence data for large accession panels

soykb.org

Or the Soyhub: https://soykb.org/soyhub.php/

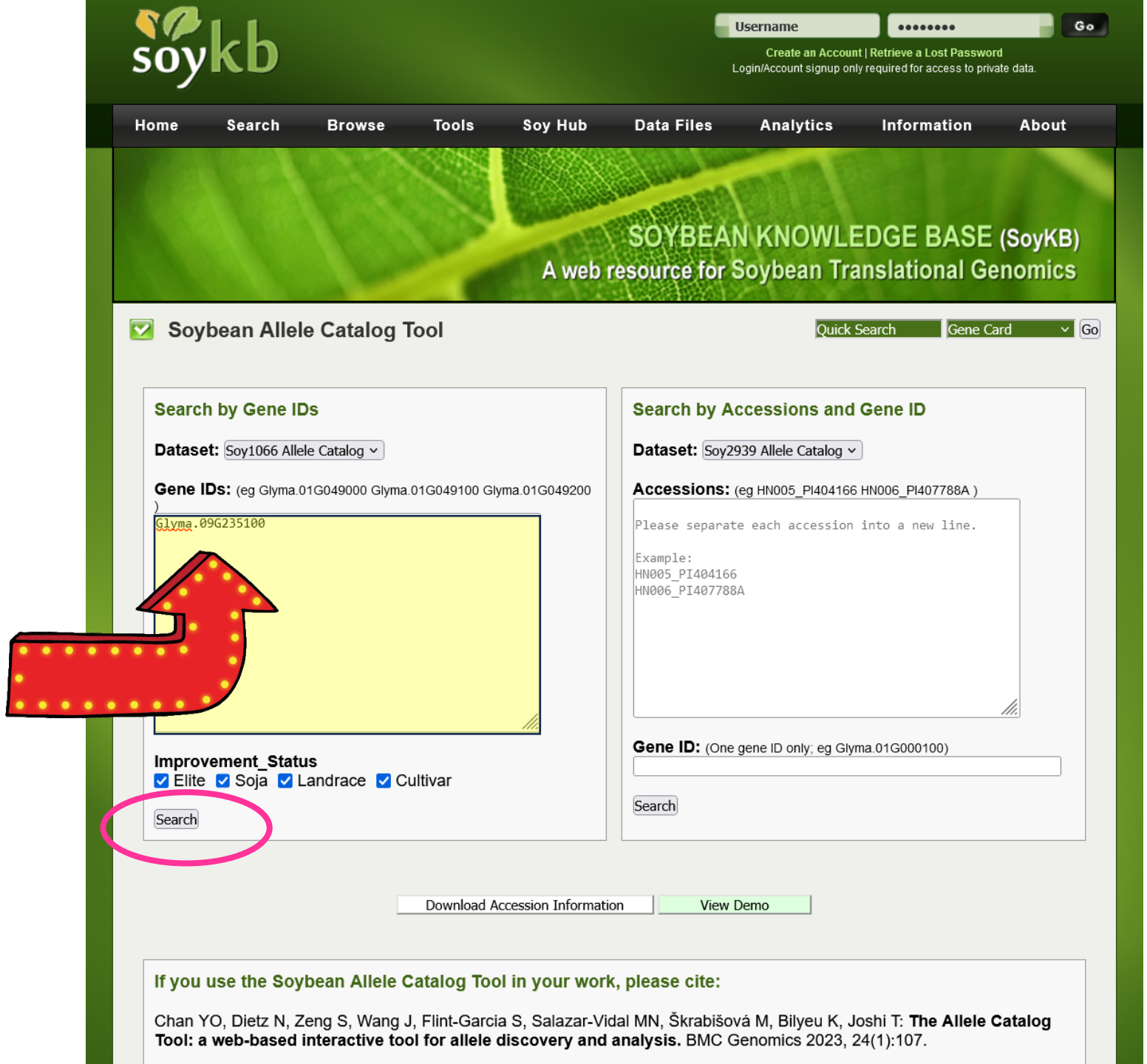


Soybean Allele Catalog Tool

Based on whole genome sequence data for large accession panels

soykb.org

Or the Soyhub: https://soykb.org/soyhub.php/



Results for *R* gene controlling **black** or **brown** seed coat/hilum pigments

Soja	Landrace	Elite	Total	Cultivar	Gene	Chromosome	45758816	45758833	45758856	45759100	45759137	45759165	45760553	45760555
107	258	199	626	93	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref
0	99	94	239	31	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	T|frameshift_variant|R75fs	G|Ref	A|Ref	G|Ref
0	44	46	90	5	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	T|intron_variant&splice_donor_variant	A|Ref	G|Ref
0	37	18	55	9	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	C|W32S	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref
0	26	24	50	2	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	C|frameshift_variant|G63fs	TC|Ref	G|Ref	A|Ref	G|Ref
2	1	0	3	0	Glyma.09G235100	Chr09	A|L19I	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref
1	0	0	1	0	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	C|E166Q
1	0	0	1	0	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	C|E165A	G|Ref
0	1	0	1	0	Glyma.09G235100	Chr09	C|Ref	G|frameshift_variant|Q25fs	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref

Download (Accession Counts) Download (All Accessions)

Download Accession Information Download All (Accession Counts) Download All (All Accessions)

Reference missense frameshift splice

Each row of the results is an allele with gene and variant position details

Gene and variant position information

Soja	Landrace	Elite	Total	Cultivar	Gene	Chromosome	45758816	45758833	45758856	45759100	45759137	45759165	45760553	45760555
107	258	199	626	93	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref
0	99	94	239	31	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	T|frameshift_variant|R75fs	G|Ref	A|Ref	G|Ref
0	44	46	90	5	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	T|intron_variant&splice_donor_variant	A|Ref	G|Ref
0	37	18	55	9	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	C|W32S	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref
0	26	24	50	2	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	C|frameshift_variant|G63fs	TC|Ref	G|Ref	A|Ref	G|Ref
2	1	0	3	0	Glyma.09G235100	Chr09	A|L19I	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref
1	0	0	1	0	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	C|E166Q
1	0	0	1	0	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	C|E165A	G|Ref
0	1	0	1	0	Glyma.09G235100	Chr09	C|Ref	G|frameshift_variant|Q25fs	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref

Download (Accession Counts) Download (All Accessions)

Download Accession Information Download All (Accession Counts) Download All (All Accessions)

Reference missense frameshift splice

A frequency table has improvement status with most frequent (Total) on top

Frequency table

Soja	Landrace	Elite	Total	Cultivar	Gene	Chromosome	45758816	45758833	45758856	45759100	45759137	45759165	45760553	45760555
107	258	199	626	93	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref
0	99	94	239	31	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	T|frameshift_variant|R75fs	G|Ref	A|Ref	G|Ref
0	44	46	90	5	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	T|intron_variant&splice_donor_variant	A|Ref	G|Ref
0	37	18	55	9	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	C|W32S	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref
0	26	24	50	2	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	C|frameshift_variant|G63fs	TC|Ref	G|Ref	A|Ref	G|Ref
2	1	0	3	0	Glyma.09G235100	Chr09	A|L19I	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref
1	0	0	1	0	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	C|E166Q
1	0	0	1	0	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	C|E165A	G|Ref
0	1	0	1	0	Glyma.09G235100	Chr09	C|Ref	G|frameshift_variant|Q25fs	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref

Download (Accession Counts) Download (All Accessions)

Download Accession Information Download All (Accession Counts) Download All (All Accessions)

Reference missense frameshift splice

Mutations in the *R* gene cause normally **black** pigments in the seed coat/hilum to be **brown**

	Soja	Landrace	Elite	Total	Cultivar	Gene	Chromosome	45758816	45758833	45758856	45759100	45759137	45759165	45760553	45760555	
☐	107	258	199	626	93	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref	***R***
☐	0	99	94	239	31	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	T|frameshift_variant|R75fs	G|Ref	A|Ref	G|Ref	*r*
☐	0	44	46	90	5	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	T|intron_variant&splice_donor_variant	A|Ref	G|Ref	*r*
☐	0	37	18	55	9	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	C|W32S	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref	*r*
☐	0	26	24	50	2	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	C|frameshift_variant|G63fs	TC|Ref	G|Ref	A|Ref	G|Ref	*r*
☐	2	1	0	3	0	Glyma.09G235100	Chr09	A|L19I	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref	☐
☐	1	0	0	1	0	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	C|E166Q	☐
☐	1	0	0	1	0	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	C|E165A	G|Ref	☐
☐	0	1	0	1	0	Glyma.09G235100	Chr09	C|Ref	G|frameshift_variant|Q25fs	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref	☐

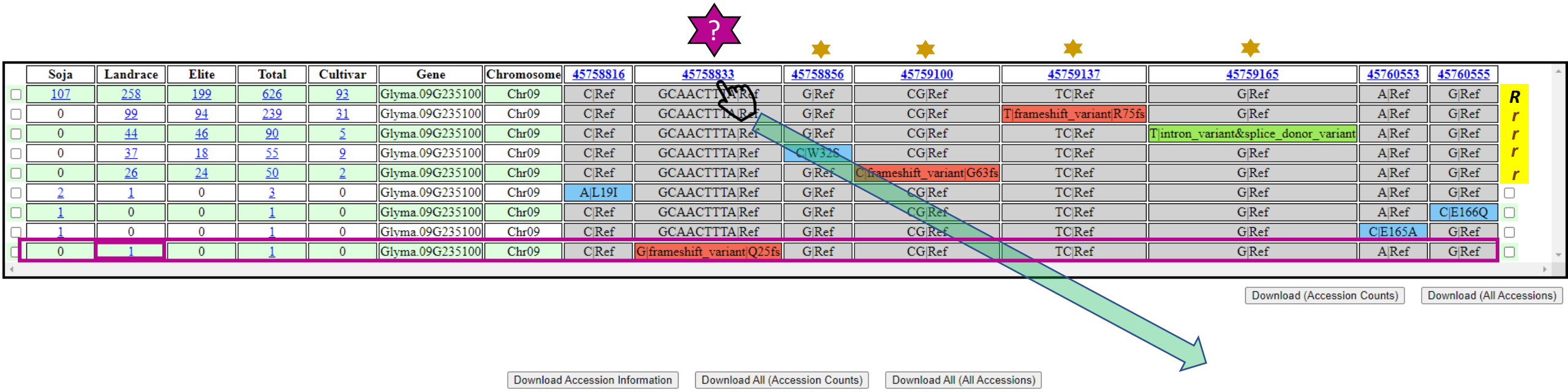
Download (Accession Counts) Download (All Accessions)

Download Accession Information Download All (Accession Counts) Download All (All Accessions)

Reference missense frameshift splice

Gillman, J., Tetlow, A., Lee, J.-D., Shannon, J.G., and Bilyeu, K., *Loss-of-function mutations affecting a specific Glycine max R2R3 MYB transcription factor result in brown hilum and brown seed coats. BMC Plant Biology, 2011.* ***11(1): p. 155.***

Allele mining-is 09:4578833 variant a new *r* allele?



	Soja	Landrace	Elite	Total	Cultivar	Gene	Chromosome	45758816	45758833	45758856	45759100	45759137	45759165	45760553	45760555	
☐	107	258	199	626	93	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref	R
☐	0	99	94	239	31	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	T|frameshift_variant|R75fs	G|Ref	A|Ref	G|Ref	r
☐	0	44	46	90	5	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	T|intron_variant&splice_donor_variant	A|Ref	G|Ref	r
☐	0	37	18	55	9	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	C|W32S	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref	r
☐	0	26	24	50	2	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	C|frameshift_variant|G63fs	TC|Ref	G|Ref	A|Ref	G|Ref	r
☐	2	1	0	3	0	Glyma.09G235100	Chr09	A|L19I	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref	☐
☐	1	0	0	1	0	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	C|E166Q	☐
☐	1	0	0	1	0	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	C|E165A	G|Ref	☐
☐	0	1	0	1	0	Glyma.09G235100	Chr09	C|Ref	G|frameshift_variant|Q25fs	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref	☐

Download (Accession Counts) Download (All Accessions)

Download Accession Information Download All (Accession Counts) Download All (All Accessions)

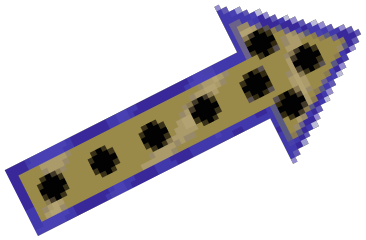
Reference missense frameshift splice

A new *r* frameshift allele for brown color

Soja	Landrace	Elite	Total	Cultivar	Gene	Chromosome	45758816	45758833	45758856	45759100	45759137	45759165	45760553	45760555	
107	258	199	626	93	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref	R
0	99	94	239	31	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	T|frameshift_variant|R75fs	G|Ref	A|Ref	G|Ref	r
0	44	46	90	5	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	T|intron_variant&splice_donor_variant	A|Ref	G|Ref	r
0	37	18	55	9	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	C|W32S	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref	r
0	26	24	50	2	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	C|frameshift_variant|G63fs	TC|Ref	G|Ref	A|Ref	G|Ref	r
2	1	0	3	0	Glyma.09G235100	Chr09	A|L19I	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref	
1	0	0	1	0	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	C|E166Q	
1	0	0	1	0	Glyma.09G235100	Chr09	C|Ref	GCAACTTTA|Ref	G|Ref	CG|Ref	TC|Ref	G|Ref	C|E165A	G|Ref	
0	1	0	1	0	Glyma.09G235100	Chr09	C|Ref	G|frameshift_variant|Q25fs	G|Ref	CG|Ref	TC|Ref	G|Ref	A|Ref	G|Ref	r

Download (Accession Counts) Download (All Accessions)

Chromosome	Position	Accession	SoyKB_Accession	GRIN_Accession	Improvement_Status	Classification	Genotype	Functional_Effect	Imputation	HILUMCOLOR	SCOATCOLOR
Chr09	45758833	SRR1533319	PI_567258	PI567258	Landrace	Other	G	frameshift_variant		Br	Br

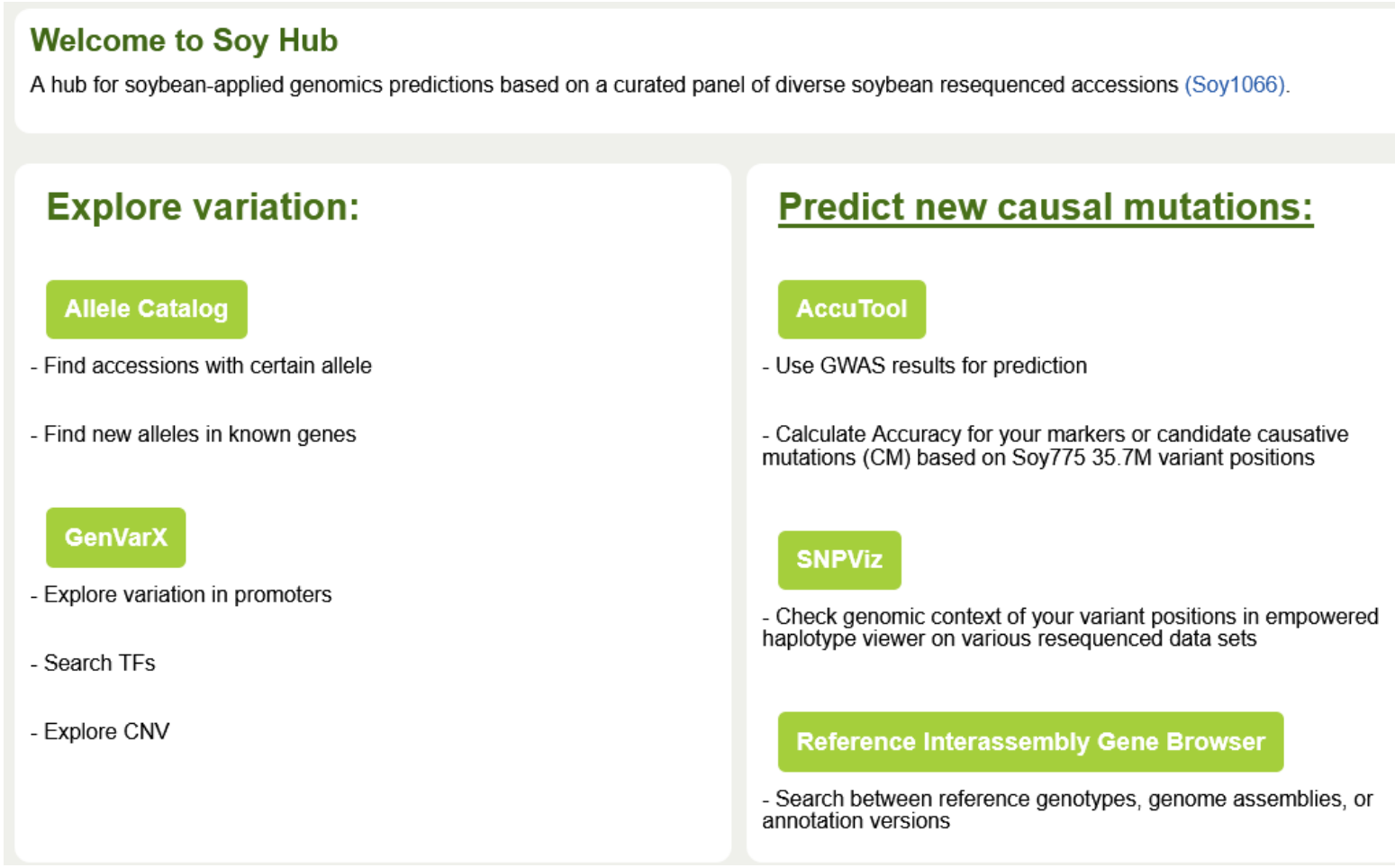


Annual updates to allele catalog data panel

kristin.bilyeu@usda.gov

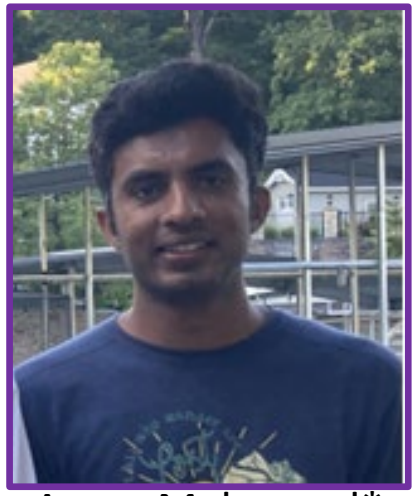
soykb.org

- Contact me about adding your data to the Soybean Allele Catalog Tool (fall deadline?)
- Private data? Pipelines available on Github (https://github.com/yenon118)
- Send feedback about tool or let us know if you want to collaborate!



Meet the team and funders

Dr. Mária Škrabišová
Palacký University Olomouc



Anser Mahmood*

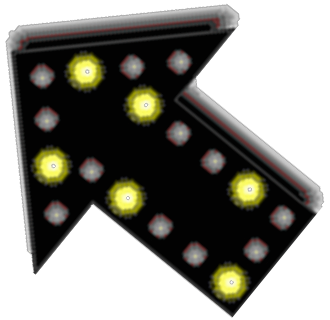
Dr. Nick Dietz (2021)

Yen On Chan*

Dr. Trupti Joshi

Jana Biová*

Dr. Nathan Grant



USDA ARS

UNITED SOYBEAN BOARD™

MU

MISSOURI SOYBEANS
brought to you by Missouri soybean farmers and their checkoff

Allele Mining with the Soybean Allele Catalog Tool

Thanks for your attention!

kristin.bilyeu@usda.gov

February 14, 2024