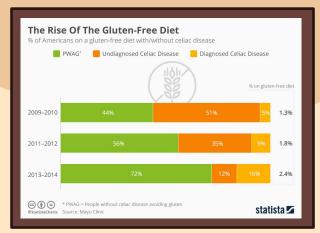
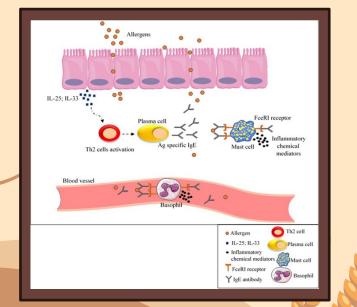


## Background

- Glutenin is a gluten protein that is a staple component in food worldwide across every culture
- Its acts as a key genetic determinant responsible for encoding proteins that contribute to the structural integrity and functionality of gluten
- Celiac disease is a chronic autoimmune disorder that is triggered by the consumption of that affects 1% of the world population, inhibiting them from consuming gluten containing products without consequences
- Proportion of individuals consuming gluten-diets amongst those who are diagnosed, undiagnosed, and not diagnosed with Celiac disease is increasing from 5%, 51%, and 44% respectively in 2009-2010 to 16%, 12%, and 72% in 2013-2014.







## Glutenin Chromosome

- Gliadins are represented by the ω, γ, and α symbols.
- Glutenins are represented by the HMW (High Molecular Weight) and LMW (Low Molecular Weight) labels.

#### Chromosome 1:

Chromosome 1A (red): Has genes Gli-A6, Gli-A1, Gli-A3, Gli-A5, and Glu-A3, and Glu-A1.

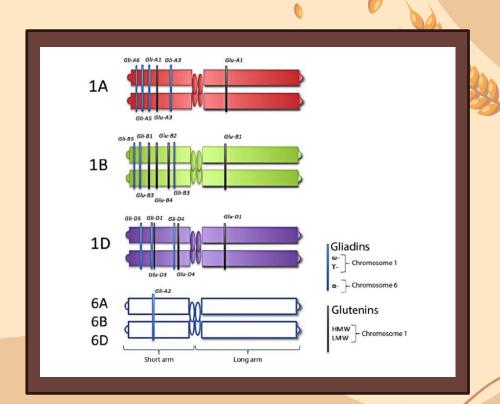
Chromosome 1B (green): Contains genes Glu-B3, Glu-B4, Gli-B5, Glu-B1, Glu-B2, and Gli-B3.

Chromosome 1D (purple): Features genes Gli-D5, Gli-D1, Gli-D4, and Glu-D3, Glu-D4, and Glu-D1.

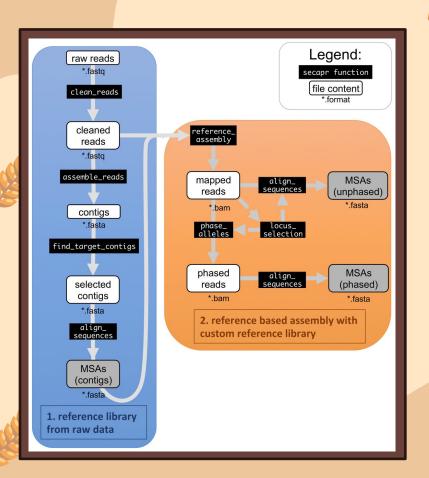
#### Chromosome 6:

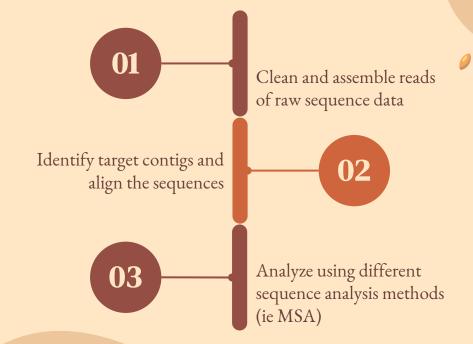
Chromosome 6A (blue): Contains the Gli-A2 gene.

Chromosome 6B and 6D are also depicted, but specific genes are not labeled on them.



## **Bioinformatics Methodologies**

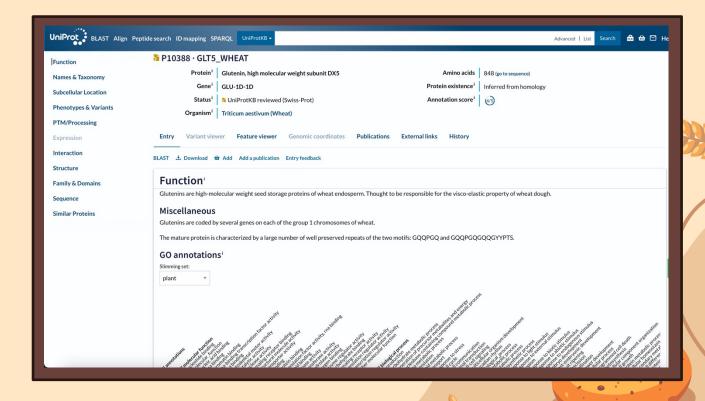






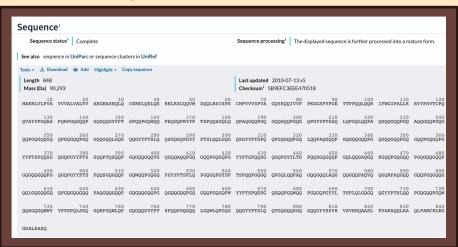
# **Sequence Data**

 Glutenin showcases a variety of similar genes and models it derives from

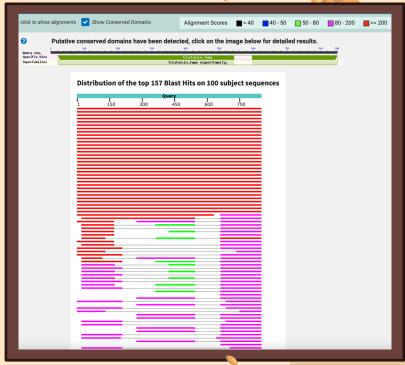


## **Sequence Data**

#### Amino Acid Sequence



## Nucleotide Alignment Scores

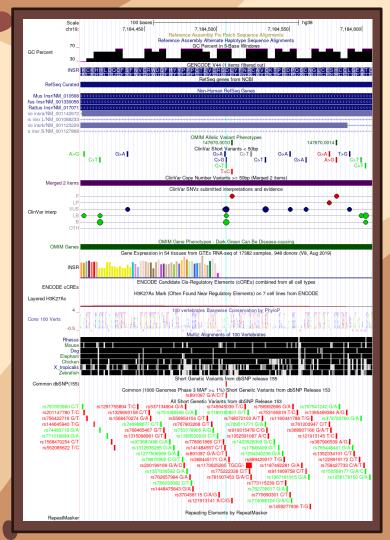




#### **Sequence Data**

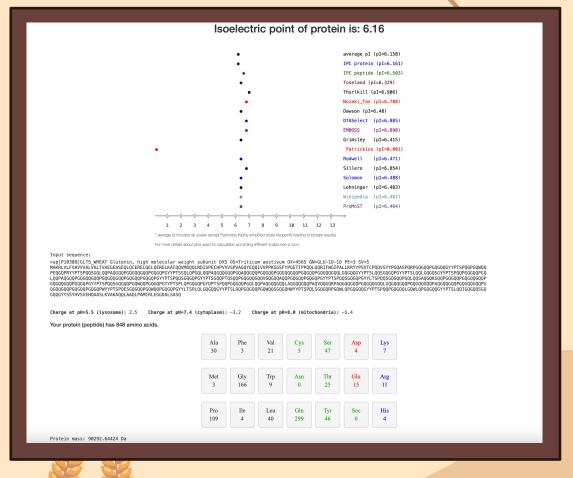
- 1. The genomic region spans from about 7,184,450 bases to 7,184,550 bases, representing a 100 base-pair segment.
- 2. There are several genetic variants depicted, with both ClinVar and dbSNP references.
- 3. OMIM Allelic Variants and Phenotypes: Displays phenotypic information associated with certain alleles with distinction between pathogenic (red) and non-pathogenic (green) variants.
- 4. Gene Expression: Data about gene expression in 50+ tissues from GTEx is shown.
- 5. Histone Marks: There's a layered H3K27Ac mark, which often indicates active regulatory regions.
- 6. Conservation: Visualization of conservation scores across many vertebrates, including several mammalian species like elephants, mouse, and dogs.





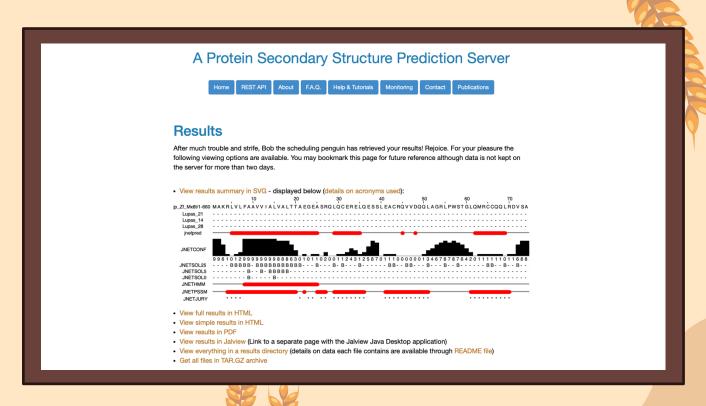
#### **Isoelectric Point**

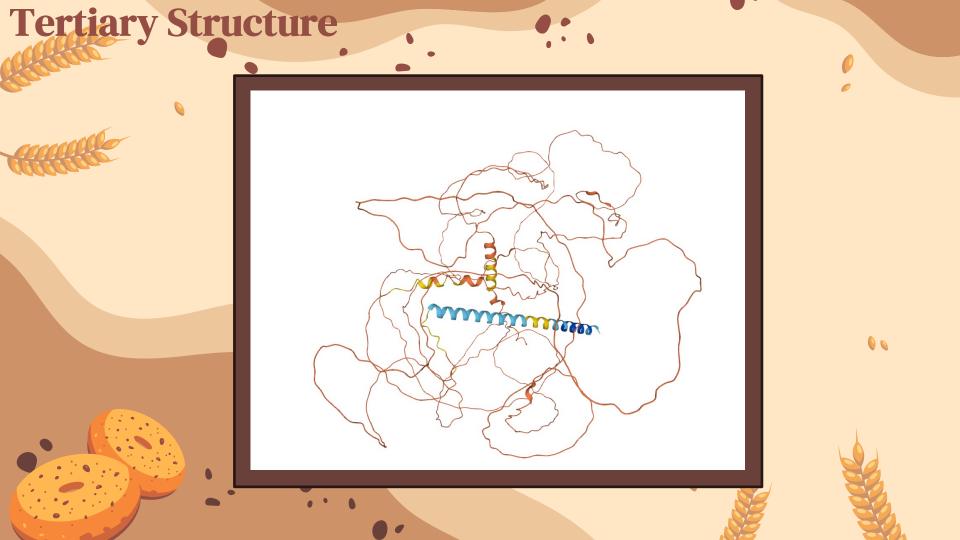
- 0
- Isoelectric point of 6.16
- Depending on environment, the net charge will change



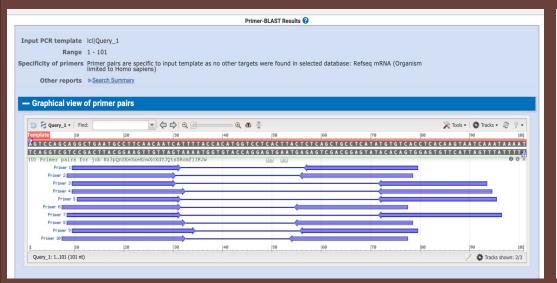
## **Secondary Protein Structure**

- Contains Alpha Helices
- Determines some functionality and components of 3D structure





## **Primer Design**



Primer pair 1	l .								
	Sequence (5'->3')	Template strand	Length	Start	Stop		GC%	Self complementarity	Self 3' complementarity
Forward primer	GGCTGAATGCCTTCAACAATCA	Plus	22	10	31	59.77	45.45	8.00	2.00
Reverse primer Product length	GTGACACATATGAGGCAGCTGAG 70	Minus	23	79	57	61.29	52.17	8.00	3.00
Primer pair 2	2								
	Sequence (5'->3')	Template strand	Length	Start	Stop		GC%	Self complementarity	Self 3' complementarity
Forward primer	AGGCTGAATGCCTTCAACAATC	Plus	22	9	30	59.51	45.45	8.00	0.00
Reverse primer Product length	TGACACATATGAGGCAGCTGAGA 70	Minus	23	78	56	61.45	47.83	8.00	1.00
Primer pair 3	3								
	Sequence (5'->3')	Template strand				Tm	GC%	Self complementarity	Self 3' complementarity
Forward primer	GGCTGAATGCCTTCAACAATC	Plus	21		30	58.12	47.62	8.00	0.00
Reverse primer Product length	TGATTACTTGTGAGGTGACACA 84	Minus	22	93	72	57.32	40.91	5.00	3.00
Primer pair 4	1								
	Sequence (5'->3')	Template strand	Length	Start	Stop		GC%	Self complementarity	Self 3' complementarity
Forward primer	GGCTGAATGCCTTCAACAATCAT	Plus	23	10	32	60.12	43.48	8.00	3.00
Reverse primer Product length	TTGATTACTTGTGAGGTGACACA 85	Minus	23	94	72	57.97	39.13	5.00	3.00
Primer pair 5	5								
	Sequence (5'->3')	Template strand	Length				GC%	Self complementarity	Self 3' complementarity
Forward primer	GCTGAATGCCTTCAACAATCA	Plus	21	11	31	57.42		8.00	2.00
Reverse primer Product length	TTTGATTACTTGTGAGGTGACACA 85	Minus	24	95	72	58.57	37.50	5.00	3.00
Primer pair 6	5								
	Sequence (5'->3')	Template strand	Length				GC%	Self complementarity	Self 3' complementarity
Forward primer	CAGGCTGAATGCCTTCAACAATCA	Plus	24	8	31	62.12		8.00	2.00
Reverse primer Product length	GACACATATGAGGCAGCTGAGAG 70	Minus	23	77	55	60.80	52.17	6.00	1.00
Primer pair 7	7								
	Sequence (5'->3')	Template strand	Length				GC%	Self complementarity	Self 3' complementarit
Forward primer	AGGCTGAATGCCTTCAACAATCA	Plus	23	9	31	61.07			2.00
Reverse primer	ATTTGATTACTTGTGAGGTGACACA 88	Minus	25	96	72	58.94	36.00	5.00	3.00







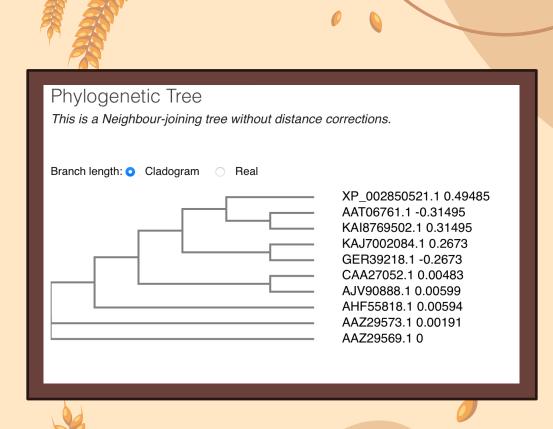
# Multiple Protein Sequence Alignment

- Utilized to determine regions of sequences which have been conserved across different genes
- Conserved regions showcase survival importance
- Those with mutations in conserved regions either get wiped out or become exceptional amongst the population
- Also contains semi-conserved and nonconserved regions which have less importance.

#### Results for job clustalo-I20231111-201904-0010-74064165-p1m Alignments Result Summary Guide Tree Phylogenetic Tree Results Viewers Submission Details Download Alignment File | Show Colors CLUSTAL 0(1.2.4) multiple sequence alignment KAJ7002084.1 --MITPGPTTVQFQR------GGERCRDQGRIPPRFDQRG MAKRLVLFVAVVVALVALTVAEGEASEQLQCERELQELQERELKACQQV-CAA27052.1 53 AJV90888.1 MAKRLVLFVAVVVALVALTVAEGEASEQLQCERELQELQERELKACQQV 53 AAT06761.1 MAKRLVLFVAVVVALVALTVAEGEASEQLQCERELQELQERELKACQQV-53 AHF55818.1 MAKRLVLFAAVVVALVALTAAEGEASGQLQCEREL---QERELEACRQV------VDQE-50 AAZ29573.1 MAKRLVLFAAVVVALVAHTAAEGEASGQLQCERE----LEACRQV-AAZ29569.1 MAKRLVLFAAVVVALVALTAAEGEASGOLOCERE-----LEACROV-KAI8769502.1 GER39218.1 XP\_002850521.1 KAJ7002084.1 -----LRDISPECHPVVVSPVAGQYEQQIVV-PKGGSFYPGETTPPQQLQQRIFWG-104 CAA27052.1 AJV90888.1 -----LRDISPECHPVVVSPVAGQYEQQIVVPPKGGSLYPGETTPPQQLQQRIFWG-AAT06761.1 -----LRDISPECHPVVVSPVAGOYEOOIVVPPKGGSFYPGETTPPOOLOORIFWG--AHF55818.1 -----LRDASPECHPIAVSPVAROYEOOTVVPPKGGSFYPGETTPPOOLOORIFWG-AAZ29573.1 ----LRDASPECRPVAVSPVARQYEQQTVVPPKGGSFYPGETTPPQQLQQRIFWG--AAZ29569.1 -----LRDASPECRPVAVSPVAROYEOOTVVPPKGGSFYPGETTPPOOLOORIFWG KAI8769502.1 GER39218.1 XP\_002850521.1 NAWLHDTPPAI----PSPLDPGAGNPNVNGHGNG---TAFANPPPPPA-T-KAJ7002084.1 PMNNRDYVP----GGRNMYPG00GNHDPP-----G00GYN-110 CAA27052.1 PALLKRYYPSVTSPQQVSYYPGQASPQRPGQGQQPGQGQQGGQGGQGYYPTSPQQP-160 AJV90888.1 PALLKRYYPSVTSPQQVSFYPGQASPQRPGQGQQPGQGQQGQQGQQGYYPTSPQQP----161 AAT06761.1 ${\tt PALLKRYYPSVTSPQQVSYYPGQASPQRPGRGQQPGQGQQGQGQGYYPTSSQQPTQSQ}$ AHF55818.1 PTLLRRYYPSVTSPQQGSYYPGQASPQRPGQGQQPGQGQQPGQGQQ-AAZ29573.1 PTLLRRYYPSVTSPR0GSYYPG0AFP0RPG0G00PG0G0----0 AAZ29569.1 PTLLRRYYPSVTSPROGSYYPGOAFPORPGOGOOPGOGO-KAI8769502.1 GER39218.1 XP 002850521.1 KAJ7002084.1 110 CAA27052.1 200 AJV90888.1 -GOWOOPEOGOPGYYPTSPOOPGOLOOPAOGOOPGOGOOGO 201 AAT06761.1 OPGOGOOGOOVGOGOOAOOPGOGOOPEOGOPGYYPTSPOOPGOLOOPAOGOOPGOGOOGO 225 AHF55818.1 -PGQGQQGQQPGQGQQPGQGQGYYPTSPQQPGQGQQPGQGQ-189 AAZ29573.1 -PGQGQQSQQPGQGQHPGQGQQGYYPTSPQQPGQGQQPGQGQ 178 AAZ29569.1 -PGQGQQSQQPGQGQHPGQGQQGYYPTSPQQPGQGQQPGQGQ 178 KAI8769502.

# **Phylogenetic Tree**

- Showcases evolutionary relationships between multiple Glutenin sequences
- Determines relative divergence and common ancestor between sequences
- Identifies regions of evolutionary importance



## **Conclusion**

- In conclusion, the extensive research on glutenin and specifically its role in Celiac disease and other gluten-related disorders, indicates that it is a growing dilemma for both medical science and the food industry.
- Current research suggests that glutenin plays a critical role in triggering the autoimmune response in Celiac disease and its can be addressed in one of two fashions, therapeutic treatment to overcome the immune reaction or create a tolerable substitute for the protein [1,2,6].
- We have a detailed mechanistic understanding of the biochemical pathways that are relevant to glutenin and enhances our understanding of its interaction with the human immune system [3,5].
- This knowledge is not only vital for medical research but also for the food industry in developing safe and enjoyable gluten-free products a growing demanded market given by the rise in gluten-free diets [10].
- Additionally, genetic insights into glutenin are invaluable for developing gluten-free wheat varieties and understanding genetic predispositions to gluten-related disorders [12,13].
- The future of glutenin research is supported by the advancements in bioinformatics tools holds promise for both health-related applications and agricultural and food science innovations [14].

### RESOURCES

- 1. "What Is Celiac Disease?" Celiac Disease Foundation, celiac.org/about-celiac-disease/what-is-celiac-disease/. Accessed 4 Dec. 2023.
- 2. Sapone, A., Bai, J.C., Ciacci, C. *et al.* Spectrum of gluten-related disorders: consensus on new nomenclature and classification. *BMC Med* 10, 13 (2012). <a href="https://doi.org/10.1186/1741-7015-10-13">https://doi.org/10.1186/1741-7015-10-13</a>
- 3. T. W. J. M. Van Herpen, M. Riley, C. Sparks, H. D. Jones, C. Gritsch, E. H. Dekking, R. J. Hamer, D. Bosch, E. M. J. Salentijn, M. J. M. Smulders, P. R. Shewry, L. J. W. J. Gilissen, Detailed Analysis of the Expression of an Alpha-gliadin Promoter and the Deposition of Alpha-gliadin Protein During Wheat Grain Development, *Annals of Botany*, Volume 102, Issue 3, September 2008, Pages 331–342, https://doi.org/10.1093/aob/mcn114
- 4. Sollid LM, Jabri B. Triggers and drivers of autoimmunity: lessons from coeliac disease. Nat Rev Immunol. 2013 Apr;13(4):294-302. doi: 10.1038/nri3407. Epub 2013 Mar 15. PMID: 23493116; PMCID: PMC3818716.
- 5. Wieser H. Chemistry of gluten proteins. Food Microbiol. 2007 Apr;24(2):115-9. doi: 10.1016/j.fm.2006.07.004. Epub 2006 Sep 7. PMID: 17008153.
- Ludvigsson JF, Leffler DA, Bai JC, Biagi F, Fasano A, Green PH, Hadjivassiliou M, Kaukinen K, Kelly CP, Leonard JN, Lundin KE, Murray JA, Sanders DS, Walker MM, Zingone F, Ciacci C. The Oslo definitions for coeliac disease and related terms. Gut. 2013 Jan;62(1):43-52. doi: 10.1136/qutjnl-2011-301346. Epub 2012 Feb 16. PMID: 22345659; PMCID: PMC3440559.
- 7. Catassi C, Elli L, Bonaz B, Bouma G, Carroccio A, Castillejo G, Cellier C, Cristofori F, de Magistris L, Dolinsek J, Dieterich W, Francavilla R, Hadjivassiliou M, Holtmeier W, Körner U, Leffler DA, Lundin KE, Mazzarella G, Mulder CJ, Pellegrini N, Rostami K, Sanders D, Skodje GI, Schuppan D, Ullrich R, Volta U, Williams M, Zevallos VF, Zopf Y, Fasano A. Diagnosis of Non-Celiac Gluten Sensitivity (NCGS): The Salerno Experts' Criteria. Nutrients. 2015 Jun 18;7(6):4966-77. doi: 10.3390/nu7064966. PMID: 26096570; PMCID: PMC4488826.

### RESOURCES

- 8. Wieser, Herbert and Peter Koehler. "The Biochemical Basis of Celiac Disease." Cereal Chemistry85 (2008): 1-13.
- 9. Belitz, H.D., Grosch, W. and Schieberle, P. (2009) Food Chemistry. 4th Edition, Springer-Verlag, Berlin, 1070 p.
- 10. McCarthy, Niall, and Felix Richter. "Infographic: The Rise of the Gluten-Free Diet." *Statista Daily Data*, 18 Jan. 2017, www.statista.com/chart/7639/the-rise-of-the-gluten-free-diet/.
- 11. Asri N, Rostami-Nejad M, Anderson RP, Rostami K. The Gluten Gene: Unlocking the Understanding of Gluten Sensitivity and Intolerance. Appl Clin Genet. 2021 Feb 11:14:37-50. doi: 10.2147/TACG.S276596. PMID: 33603437; PMCID: PMC7886246.
- 12. Li Y, Fu J, Shen Q, Yang D. High-Molecular-Weight Glutenin Subunits: Genetics, Structures, and Relation to End Use Qualities. Int J Mol Sci. 2020 Dec 26;22(1):184. doi: 10.3390/ijms22010184. PMID: 33375389; PMCID: PMC7795185.
- 13. Delorean, E., Gao, L., Lopez, J.F.C. *et al.* High molecular weight glutenin gene diversity in *Aegilops tauschii* demonstrates unique origin of superior wheat quality. *Commun Biol* 4, 1242 (2021). <a href="https://doi.org/10.1038/s42003-021-02563-7">https://doi.org/10.1038/s42003-021-02563-7</a>
- 14. Andermann T, Cano Á, Zizka A, Bacon C, Antonelli A. 2018. SECAPR—a bioinformatics pipeline for the rapid and user-friendly processing of targeted enriched Illumina sequences, from raw reads to alignments. *PeerJ* 6:e5175 <a href="https://doi.org/10.7717/peerj.5175">https://doi.org/10.7717/peerj.5175</a>
- 15. Chromosomal Locaton of Glutenin and Gliadin Loci in Hexaploid Wheat ..., www.researchgate.net/figure/Chromosomal-locaton-of-glutenin-and-gliadin-loci-in-hexaploid-wheat-The-high-molecular\_fig1\_273772041. Accessed 5 Dec. 2023.