

Promoters of genes encoding β -amylase, albumin, and globulin in food plants have weaker affinity for TATA-binding protein as compared to non-food plants: *in silico* analysis

Podkolodnaya O.^{1*}, Chadaeva I.¹, Zolotareva K.¹, Podkolodnyy N.^{1,3}, Khandaev B.^{1,2}, Rasskazov D.¹, Bogomolov A.^{1,2}, Kazachek A.^{1,2}, Vishnevsky O.^{1,2}, Sharypova E.¹, Ponomarenko P.¹, Savinkova L.¹, Zemlyanskaya E.^{1,2}, Ponomarenko M.¹

¹ Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia

² Novosibirsk State University, Novosibirsk, Russia

³ Institute of Computational Mathematics and Mathematical Geophysics, SB RAS, Novosibirsk, Russia

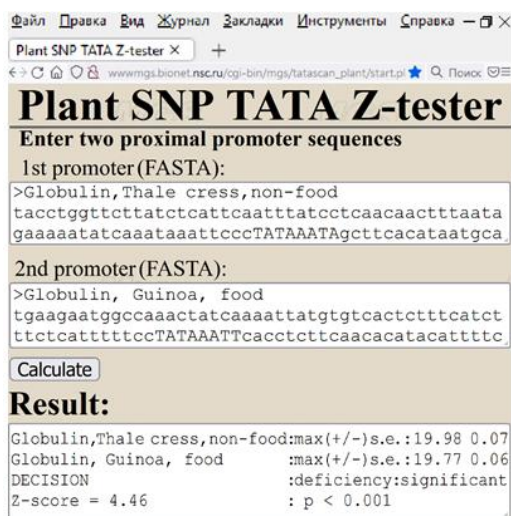
* opodkol@bionet.nsc.ru

Key words: food; allergen; albumin; globulin; β -amylase; gene; promoter; TATA-binding protein; TATA box; common wheat; *Triticum aestivum* L. (1753); plants; spontaneous artificial selection

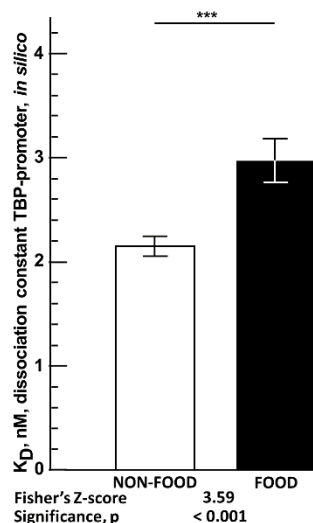
Motivation and Aim: It is generally accepted that during the domestication of food plants, the selection was focused on their productivity, the ease of their technological processing into food, and resistance to pathogens and environmental stressors. Besides, the palatability of plant foods and their health benefits could also be subjected to selection by humans in the past. Nonetheless, it is unclear whether in antiquity, aside from positive selection for beneficial properties of plants, humans also simultaneously selected against such detrimental properties as allergenicity. This topic is becoming increasingly relevant as the allergization of population grows, being a major challenge for modern medicine. That is why, breeders are constantly trying to create hypoallergenic forms of food plants. **Methods and Algorithms:** Accordingly, in this paper, albumin, globulin, and β -amylase of common wheat *Triticum aestivum* L. (1753) are analyzed, which have been identified earlier as targets for attacks by human class E immunoglobulins [1]. At the genomic level, we wanted to find signs of past negative selection against the allergenicity of these three proteins (albumin, globulin, and β -amylase) during the domestication of ancestral forms of modern food plants. Using our public web service Plant_SNP_TATA_Z-tester [2], here we focused the search on the TATA-binding protein (TBP)-binding site because it is located within a narrow region (between positions -70 and -20 relative to corresponding transcription start sites), is the most conserved, necessary for primary transcription initiation, and is the best-studied regulatory genomic signal in eukaryotes. By this way, we have estimated the equilibrium dissociation constant (K_D) of TBP complexes with plant proximal promoters (as its output data) using 90 bp of their DNA sequences (as its input data) for 235 gene promoters representing 28 plant species, as described in Table 1 and illustrated in Fig. 1a.

Table 1. Characteristics of 235 nucleotide sequences of proximal promoters of food plant genes homologous to the studied globulin (*Glo*), albumin (*Alb*), and β -amylase (*Bmy*) gens from common wheat *Triticum aestivum* L. (1753)

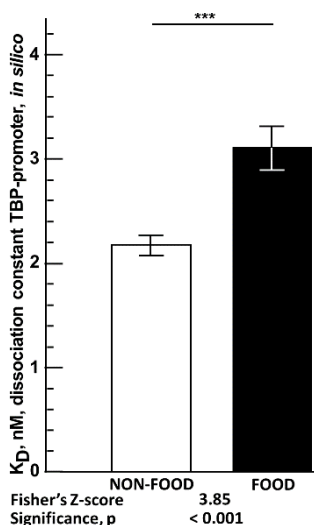
Food plant species		Number of promoters		
#	Name	<i>Glo</i>	<i>Alb</i>	<i>Bmy</i>
1	buckwheat <i>Fagopyrum esculentum</i> Moench, 1794	1		
2	maidenhair tree <i>Ginkgo biloba</i> L., 1771	1		
3	Yoshino cherry <i>Prunus yedoensis</i> var. <i>nudiflora</i> Koehne, 1912	1		2
4	maize (<i>Zea mays</i> L., 1753	1		
5	oat <i>Avena sativa</i> L., 1753	2		
6	waxberry <i>Morella rubra</i> Siebold & Zucc.	2	2	1
7	quinoa <i>Chenopodium quinoa</i> Willd., 1798	2		
8	rice <i>Oryza sativa</i> L., 1753	3		
9	melon <i>Cucumis melo</i> L., 1753	4	2	6
10	cardoon <i>Cynara cardunculus</i> L.	4		
11	cork oak <i>Quercus suber</i> L.	4		
12	wine grape <i>Vitis vinifera</i> L.	9		9
13	Congolese coffee <i>Coffea canephora</i> Pierre ex A.Froehner, 1897	1		
14	pepper <i>Capsicum annum</i> L., 1753	26	8	
15	sesame <i>Sesamum indicum</i> L.		1	
16	kiwifruit nashi-kazura <i>Actinidia rufa</i> Franch. & Sav		1	27
17	Brazil nut <i>Bertholletia excelsa</i> Humb. & Bonpl.		1	
18	soybean <i>Glycine max</i> (L.) Merr., 1917		2	
19	pea <i>Pisum sativum</i> L., 1753		4	
20	perilla <i>Perilla frutescens</i> var. <i>hirtella</i> (Nakai) Makino		5	
21	almond <i>Prunus dulcis</i> (Mill.) D.A.Webb, 1967		8	4
22	mandarin unshiu <i>Citrus unshiu</i> (Tanaka ex Swingle) Marcow., 1921		15	
23	tea <i>Camellia sinensis</i> (L.) Kuntze, 1887			1
24	barley <i>Hordeum vulgare</i> L. (1753)			2
25	hibiscus <i>Hibiscus syriacus</i> L. (1753)			2
26	pineapple <i>Ananas comosus</i> (L.) Merr., 1917			3
27	olive <i>Olea europaea</i> L., 1753			4
28	sweet wormwood <i>Artemisia annua</i> L.	13	35	16
TOTAL number of food plant species		15	12	12



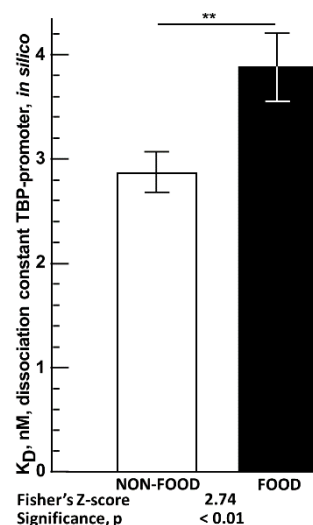
(a)



(b)



(c)



(d)

Fig. 1. The statistically significant difference between the studied food plants and non-food plants in the *in silico* estimates of K_D for complexes of plant TBP with 90 bp proximal promoters of their genes encoding globulins (b), albumins (c) and β -amylases (d), as calculated by our publicly available Web service *Plant_SNP_TATA_Z-tester* [2] (a). Legend: Bar height, the arithmetic mean of the equilibrium dissociation constant (K_D) for complexes between plant TBP and the 90 bp proximal promoters of the indicated plant genes; error bars, standard error of the mean (SEN); asterisks, statistical significance according to Fisher's Z-test, namely: ** for $p < 0.01$ and *** for $p < 0.001$

Results: It was found that compared with non-food plants, food plants are characterized by significantly weaker affinity of TBP for proximal promoters of their genes homologous to the genes of common-wheat globulin, albumin, and β -amylase (food allergens), as presented in Fig. 1b, 1c and 1d: $p < 0.01$ according to Fisher's Z-test [3].

Conclusion: This suggests that in the past, humans to reduce the expression of food plant genes encoding these allergenic proteins carried out spontaneous artificial selection.

Funding: The study is supported by the Government Budget Project FWNR-2022-0020.

References

1. Wang Y., Weng J., Zhu C., Ai R., Zhou J., Wang C., Chen Q., Fu L. Allergenicity assessment and allergen profile analysis of different Chinese wheat cultivars. *World Allergy Organ J.* 2021;14(7):100559. doi 10.1016/j.waojou.2021.100559
2. Rasskazov D., Chadaeva I., Sharypova E. et al. Plant_SNP_TATA_Z-Tester: A Web service that unequivocally estimates the impact of proximal promoter mutations on plant gene expression. *Int J Mol Sci.* 2022;23(15):8684. doi 10.3390/ijms23158684
3. Vishnevsky O.V., Chadaeva I.V., Sharypova E.B. et al. Promoters of genes encoding β -amylase, albumin and globulin in food plants have weaker affinity for TATA-binding protein as compared to non-food plants: in silico analysis. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov Journal of Genetics and Breeding.* 2022;26(8):798-805. doi 10.18699/VJGB-22-96