

Перевод на английский язык <https://vavilov.elpub.ru/jour>

Промоторы генов, кодирующих β -амилазу, альбумин и глобулин пищевых растений в сравнении с непищевыми, характеризуются более низкой аффинностью к ТАТА-связывающему белку: *in silico* анализ

О.В. Вишневский¹, И.В. Чадаева¹, Е.Б. Шарыпова¹, Б.М. Хандаев¹, К.А. Золотарева¹, А.В. Казачек¹,
П.М. Пономаренко¹, Н.Л. Подколотный^{1, 2}, Д.А. Рассказов¹, А.Г. Богомолов¹, О.А. Подколотная¹,
Л.К. Савинкова¹, Е.В. Землянская¹, М.П. Пономаренко¹ ✉

¹ Федеральное исследовательское учреждение Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия

² Институт вычислительной математики и математической геофизики Сибирского отделения Российской академии наук, Новосибирск, Россия
✉ pon@bionet.nsc.ru

Аннотация. Принято считать, что при доместикации пищевых растений отбор шел на урожайность, технологичность переработки в продукты питания, устойчивость к патогенам и стрессовым воздействиям окружающей среды. При этом также могли оцениваться вкусовые качества продуктов питания растительного происхождения и их ценность для здоровья. Однако неясно, проводил ли человек в прошлом наряду с положительным отбором на полезные свойства растений одновременно отбор против таких вредоносных свойств, как способность вызывать аллергические реакции. Этот вопрос становится все более актуальным по мере роста аллергии населения как вызова современной медицине. В связи с этим селекционерами уже ведутся интенсивные исследования по созданию гипоаллергенных форм пищевых растений. В этой работе рассмотрены альбумин, глобулин и β -амилаза мягкой пшеницы *Triticum aestivum* L. (1753), идентифицированные ранее как мишени для атак иммуноглобулинов класса Е человека. Нашей целью было найти на геномном уровне следы отрицательного отбора в прошлом против гипераллергенности трех белков (альбумин, глобулин и β -амилаза) при одомашнивании предковых форм современных пищевых растений. Для этого мы сфокусировали поиск на сайте связывания ТАТА-связывающего белка (ТВР) как локализованном в узком районе [–70; –20] относительно старта транскрипции, консервативном, необходимом для первичной инициации транскрипции и наиболее изученном регуляторном сигнале в геномах эукариот. Ранее нами был создан свободно доступный веб-сервис Plant_SNP_TATA_Z-tester для оценки величин равновесной константы диссоциации (K_D) комплексов ТВР с проксимальными промоторами генов растений по их последовательностям ДНК длиной 90 п.о. В настоящей работе с его помощью проанализированы 363 последовательности ДНК промоторов генов 43 видов растений. Обнаружено, что пищевые растения, в сравнении с непищевыми, характеризуются достоверно более низкой аффинностью ТВР к проксимальным промоторам их генов, гомологичных генам глобулина, альбумина и β -амилазы мягкой пшеницы как пищевых аллергенов ($p < 0.01$, Z-критерий Фишера). Это свидетельствует об отборе при доместикации пищевых растений в прошлом на снижение уровня данных аллергенных белков. Ключевые слова: пищевые аллергены; альбумин; глобулин; β -амилаза; ген; промотор; мягкая пшеница *Triticum aestivum* L. (1753); растения; ТАТА-связывающий белок; ТАТА-бокс; доместикация; отбор; оценки *in silico*.

Для цитирования: Вишневский О.В., Чадаева И.В., Шарыпова Е.Б., Хандаев Б.М., Золотарева К.А., Казачек А.В., Пономаренко П.М., Подколотный Н.Л., Рассказов Д.А., Богомолов А.Г., Подколотная О.А., Савинкова Л.К., Землянская Е.В., Пономаренко М.П. Промоторы генов, кодирующих β -амилазу, альбумин и глобулин пищевых растений в сравнении с непищевыми, характеризуются более низкой аффинностью к ТАТА-связывающему белку: *in silico* анализ. Вавиловский журнал генетики и селекции. 2022;26(8):798-805. DOI 10.18699/VJGB-22-96

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

O.V. Vishnevsky¹, I.V. Chadaeva¹, E.B. Sharypova¹, B.M. Khandaev¹, K.A. Zolotareva¹, A.V. Kazachek¹,
P.M. Ponomarenko¹, N.L. Podkolodny^{1, 2}, D.A. Rasskazov¹, A.G. Bogomolov¹, O.A. Podkolodnaya¹,
L.K. Savinkova¹, E.V. Zemlyanskaya¹, M.P. Ponomarenko¹ ✉

¹ Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia

² Institute of Computational Mathematics and Mathematical Geophysics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia
✉ pon@bionet.nsc.ru

Abstract. It is generally accepted that during the domestication of food plants, 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 simultaneously selected against such detrimental properties as allergenicity. This topic is becoming increasingly relevant as the allergization of the population grows, being a major challenge for modern medicine. That is why intensive research by breeders is already underway for creating hypoallergenic forms of food plants. 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. 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. 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 the corresponding transcription start sites), is the most conserved, necessary for primary transcription initiation, and is the best-studied regulatory genomic signal in eukaryotes. Our previous studies presented our publicly available Web service Plant_SNP_TATA_Z-tester, which makes it possible to estimate the equilibrium dissociation constant (K_D) of TBP complexes with plant proximal promoters (as output data) using 90 bp of their DNA sequences (as input data). In this work, by means of this bioinformatics tool, 363 gene promoter DNA sequences representing 43 plant species were analyzed. 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) ($p < 0.01$, Fisher's Z-test). This evidence suggests that in the past humans carried out selective breeding to reduce the expression of food plant genes encoding these allergenic proteins.

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

For citation: Vishnevsky O.V., Chadaeva I.V., Sharypova E.B., Khandayev B.M., Zolotareva K.A., Kazachek A.V., Ponomarenko P.M., Podkolodny N.L., Rasskazov D.A., Bogomolov A.G., Podkolodnaya O.A., Savinkova L.K., Zemlyanskaya E.V., Ponomarenko M.P. 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

Введение

В настоящее время проблема аллергенности продуктов питания исключительно актуальна, так как наблюдающийся стремительный рост аллергизации населения становится одним из наиболее ключевых вызовов современной медицины (Prescott et al., 2022). В связи с этим современными селекционерами ведется работа в двух направлениях: (1) создание новых гипоаллергенных форм сельскохозяйственных пищевых растений и (2) выявление новых пищевых аллергенов растений и молекулярных механизмов их действия (Hong et al., 2021; Cavazza et al., 2022).

Цель нашего исследования заключалась в поиске на молекулярно-генетическом уровне следов отрицательного отбора против аллергенов в процессе одомашнивания предковых форм современных пищевых растений. Были изучены три пищевых аллергена мягкой пшеницы *Triticum aestivum* L. (1753): β -амилаза, альбумин и глобулин, идентифицированные как мишени атаки иммуноглобулинов класса Е человека (Wang et al., 2021).

Работа выполнялась с использованием созданного нами ранее свободно доступного веб-сервиса Plant_SNP_TATA_Z-tester, который предназначен для оценки величин равновесной константы диссоциации (K_D) комплексов TBP-1 из резуховидки Таля *Arabidopsis thaliana* (L.) Heynh., 1842 (здесь и далее – растительный TBP) с проксимальными промоторами генов растений (Rasskazov et al., 2022). С его помощью проанализированы 363 нуклеотидные последовательности проксимальных промоторов генов 43 видов растений. Обнаружено, что пищевые растения, в сравнении с непищевыми, характеризуются значимо

более низкими оценками аффинности растительного TBP к промоторам генов, гомологичных β -амилазе, альбумину и глобулину мягкой пшеницы как пищевым аллергенам. Это указывает на имевший место искусственный отбор, направленный на снижение уровня экспрессии генов пищевых растений, кодирующих аллергенные белки, который осуществлялся человеком в прошлом по мере одомашнивания таких растений.

Материалы и методы

Нуклеотидные последовательности промоторов растительных генов, проанализированных в работе. Рассмотрены три аллергенных белка мягкой пшеницы *T. aestivum*: β -амилаза, альбумин и глобулин, которые ранее были экспериментально идентифицированы в качестве мишеней для атаки иммуноглобулинов класса Е человека (Wang et al., 2021). Из базы данных GenBank (Benson et al., 2015) были экстрагированы нуклеотидные последовательности проксимальных промоторов длиной 90 п. о., расположенных непосредственно перед стартами транскрипции растительных генов, гомологичных генам β -амилазы, альбумина и глобулина мягкой пшеницы *T. aestivum*. После исключения ДНК-последовательностей промоторов с неопределенными нуклеотидами, w, s, r, y, k, m, b, d, h, v и n (согласно номенклатуре (IUPAC-IUB..., 1970)), остались 363 промоторных последовательности, относящиеся к 43 видам растений. Затем все 43 вида растений были разделены на две непересекающиеся группы: группа I, представленная 235 проксимальными промоторами 28 видов пищевых растений, для каждого из которых удалось найти сведения об их многовековом использо-

Таблица 1. Характеристика 235 нуклеотидных последовательностей проксимальных промоторов генов пищевых растений – гомологов глобулина (*Glo*), альбумина (*Alb*) и β -амилазы (*Bmy*) мягкой пшеницы *T. aestivum*, изученных в этой работе

Виды пищевых растений		Количество промоторов		
№ п/п	Название	<i>Glo</i>	<i>Alb</i>	<i>Bmy</i>
1	Гречиха посевная <i>Fagopyrum esculentum</i> Moench, 1794	1	–	–
2	Гинкго двуплодный <i>Ginkgo biloba</i> L., 1771	1	–	–
3	Вишня Ёсино <i>Prunus yedoensis</i> var. <i>nudiflora</i> Koehne, 1912	1	–	2
4	Кукуруза <i>Zea mays</i> L., 1753	1	–	–
5	Овес посевной <i>Avena sativa</i> L., 1753	2	–	–
6	Восковница красная <i>Morella rubra</i> Siebold & Zucc.	2	2	1
7	Киноа <i>Chenopodium quinoa</i> Willd., 1798	2	–	–
8	Рис посевной <i>Oryza sativa</i> L., 1753	3	–	–
9	Дыня <i>Cucumis melo</i> L., 1753	4	2	6
10	Артишок испанский <i>Cynara cardunculus</i> L.	4	–	–
11	Дуб пробковый <i>Quercus suber</i> L.	4	–	–
12	Виноград <i>Vitis vinifera</i> L.	9	–	9
13	Кофе конголезский <i>Coffea canephora</i> Pierre ex A. Froehner, 1897	1	–	–
14	Перец стручковый <i>Capsicum annuum</i> L., 1753	26	8	27
15	Кунжут <i>Sesamum indicum</i> L.	–	1	–
16	Киви <i>Actinidia rufa</i> Franch. & Sav.	–	1	–
17	Бразильский орех <i>Bertholletia excelsa</i> Humb. & Bonpl.	–	1	–
18	Соя культурная <i>Glycine max</i> (L.) Merr., 1917	–	2	–
19	Горох посевной <i>Pisum sativum</i> L., 1753	–	4	–
20	Перилла кустарниковая <i>Perilla frutescens</i> var. <i>hirtella</i> (Nakai) Makino	–	5	–
21	Миндаль обыкновенный <i>Prunus dulcis</i> (Mill.) D.A. Webb, 1967	–	8	4
22	Мандарин уншиу <i>Citrus unshiu</i> (Tanaka ex Swingle) Marow., 1921	–	15	–
23	Чай <i>Camellia sinensis</i> (L.) Kuntze, 1887	–	–	1
24	Ячмень обыкновенный <i>Hordeum vulgare</i> L. (1753)	–	–	2
25	Гибискус сирийский <i>Hibiscus syriacus</i> L. (1753)	–	–	2
26	Ананас настоящий <i>Ananas comosus</i> (L.) Merr., 1917	–	–	3
27	Олива европейская <i>Olea europaea</i> L., 1753	–	–	4
28	Полынь однолетняя <i>Artemisia annua</i> L.	13	35	16
Количество исследованных видов пищевых растений		15	12	12

вании людьми в качестве продуктов питания (табл. 1), и группа II, представленная 128 проксимальными промоторами непищевых растений (остальные 15 видов) (табл. 2).

Анализ нуклеотидных последовательностей проксимальных промоторов растений. С помощью созданного нами ранее веб-сервиса Plant_SNP_TATA_Z-tester (Rasskazov et al., 2022) по нуклеотидной последовательности каждого промотора, охарактеризованного в табл. 1 и 2, рассчитывали в единицах моли на литр (М) величины K_D комплексов растительного ТВР с этим промотором.

Расчеты выполняли в соответствии с разработанной нами ранее моделью трехэтапного связывания ТВР с промотором: (i) ТВР скользит вдоль двойной спирали ДНК промотора (Coleman, Pugh, 1995) ↔ (ii) ТВР останавлива-

ется на потенциальном сайте ТВР-связывания (Berg, von Hippel, 1987; Bucher, 1990) ↔ (iii) комплекс ТВР/промотор стабилизируется за счет изгиба оси двойной спирали ДНК под прямым углом (Flatters, Lavery, 1998). Впоследствии такое трехэтапное связывание экспериментально наблюдали *in vitro* R.F. Delgadillo с коллегами (2009).

Статистический анализ. С использованием стандартного пакета программ Statistica (Statsoft™, США) усредняли полученные с помощью Plant_SNP_TATA_Z-tester (Rasskazov et al., 2022) оценки K_D комплексов растительного ТВР к промоторам генов β -амилазы, альбумина и глобулина для пищевых и непищевых растений по отдельности. На этой основе оценивали статистическую значимость различия между пищевыми и непищевыми растениями согласно Z-критерию Фишера.

Таблица 2. Характеристика 128 нуклеотидных последовательностей проксимальных промоторов генов непищевых растений – гомологов глобулина (*Glo*), альбумина (*Alb*) и β-амилазы (*Bmy*) мягкой пшеницы *T. aestivum*, изученных в этой работе

Виды непищевых растений		Количество промоторов		
№ п/п	Название	<i>Glo</i>	<i>Alb</i>	<i>Bmy</i>
1	Маклея сердцевидная <i>Macleaya cordata</i> (Willd.) R. Br.	1	–	–
2	Стрига желтая <i>Striga asiatica</i> (L.) Kuntze	1	1	1
3	Генлиси <i>Genlisea aurea</i> A.St. Hil. (1833)	1	–	2
4	Теосинте флоридский <i>Zea luxurians</i> (Durieu & Asch.) R.M. Bird, 1978	1	–	–
5	Ярутка <i>Microthlaspi erraticum</i> (Jord.) T. Ali & Thines, 2016	2	1	2
6	Трипсакум ежевидный <i>Tripsacum dactyloides</i> (L.) L., 1759	2	–	–
7	Теосинте балсасский <i>Zea mays</i> subsp. <i>parviglumis</i> Ittis & Doebley, 1980	3	–	–
8	Резуховидка Таля <i>Arabidopsis thaliana</i> (L.) Heynh., 1842	4	5	3
9	Дихантелий <i>Dichantheium oligosanthos</i> (Schult.) Gould	6	–	6
10	Водяная лилия <i>Nymphaea thermarum</i> Eb. Fisch., 1988	8	1	9
11	Василистник анемоновидный <i>Thalictrum thalictroides</i> (L.) A.J. Eames & B. Boivin	9	15	6
12	Томат чилийский <i>Solanum chilense</i> (Dunal) Reiche	15	13	6
13	Стрига египетская <i>Striga hermonthica</i> (Delile) Benth.	–	1	–
14	Вшивосемянник японский <i>Phtheirospermum japonicum</i> (Thunb.) Kanitz	–	–	1
15	Китайская роза <i>Rosa chinensis</i> Jacq., 1768	–	–	2
Количество видов непищевых растений		12	7	10

Результаты

Глобулин

Результаты оценки *in silico* K_D комплексов растительного ТВР с 74 проксимальными промоторами генов глобулинов из 15 видов пищевых растений в сравнении с 53 такими промоторами из 12 видов непищевых растений, полученные с использованием Plant_SNP_TATA_Z-tester (Rasskazov et al., 2022), представлены в табл. 3. Согласно этой таблице, у пищевых растений оценки K_D комплексов растительного ТВР с промоторами этих генов варьировали от 1.67 ± 0.12 до 6.75 ± 5.23 нМ при среднем значении, равном 2.97 ± 0.21 нМ, а у непищевых – от 1.25 ± 0.06 до 3.33 ± 0.23 нМ при среднем 2.15 ± 0.08 нМ.

На рис. 1 показано сравнение среднеарифметических оценок K_D комплексов растительных ТВР к промоторам генов, кодирующих глобулины, для двух контрастных выборок (пищевых и непищевых растений) с помощью Z-критерия Фишера. Различие между сравниваемыми выборками было достоверным с $Z = 3.59$, с уровнем статистической значимости $p < 0.001$.

Альбумин

В табл. 4 приведены полученные с помощью веб-сервиса Plant_SNP_TATA_Z-tester (Rasskazov et al., 2022) оценки K_D комплексов растительного ТВР с 84 промоторами генов альбуминов из 12 видов пищевых растений и с 37 промоторами из 7 видов непищевых растений. Как можно видеть из этой таблицы, оценки K_D комплексов ТВР/промотор для этих генов в случае пищевых растений наблюдались от 1.65 ± 0.12 до 4.49 ± 1.39 нМ (в среднем

3.10 ± 0.22 нМ), тогда как у непищевых растений они были от 1.65 ± 0.05 до 2.70 ± 0.22 нМ (в среднем 2.18 ± 0.10 нМ).

Сравнение исследуемых выборок пищевых и непищевых растений с помощью Z-критерия Фишера представлено на рис. 2: пищевые растения достоверно отличаются от непищевых ($Z = 3.85$, $p < 0.001$).

β-Амилаза

В табл. 5 приведены результаты оценки K_D комплексов растительного ТВР с 77 проксимальными промоторами

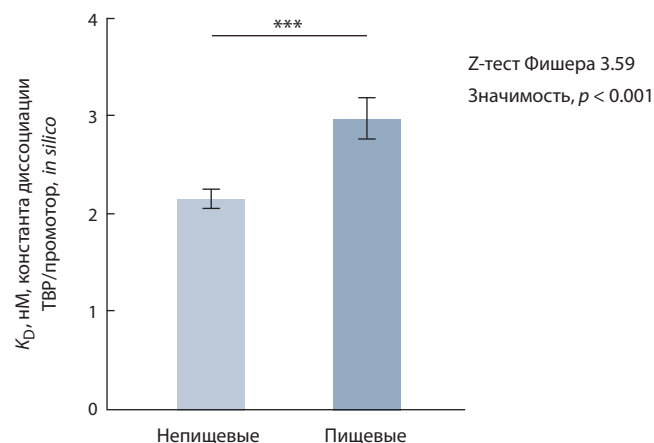


Рис. 1. Достоверное различие исследованных выборок пищевых и непищевых растений по оценке *in silico* K_D комплексов растительного ТВР к промоторам их генов, кодирующих глобулины.

Здесь и на рис. 2: *** достоверность $p < 0.001$ согласно Z-критерию Фишера.

Таблица 3. Среднеарифметические оценки величин равновесной константы диссоциации (K_D) комплексов ТВР с проксимальными промоторами генов растительных глобулинов

№ п/п	Вид растения	N	$K_D, M_0 \pm \Delta, \text{нМ}$
Пищевые растения			
1	Кофе конголезский	1	2.17 ± 0.13
2	Гречиха посевная	1	2.04 ± 0.14
3	Гинкго двулопастный	1	1.67 ± 0.12
4	Вишня Ёсино	1	1.76 ± 0.12
5	Кукуруза	1	2.30 ± 0.16
6	Овес посевной	2	2.57 ± 0.15
7	Восковница красная	2	6.75 ± 5.23
8	Киноа	2	2.47 ± 0.12
9	Рис посевной	3	2.84 ± 0.38
10	Дыня	4	2.66 ± 0.30
11	Артишок испанский	4	3.22 ± 0.37
12	Дуб пробковый	4	4.84 ± 1.41
13	Виноград	9	2.67 ± 0.35
14	Полынь однолетняя	13	2.51 ± 0.25
15	Перец стручковый	26	3.01 ± 0.37
Итого		74	2.97 ± 0.21
Непищевые растения			
1	Макля сердцевидная	1	1.25 ± 0.06
2	Стрига желтая	1	3.33 ± 0.23
3	Генлиси	1	2.70 ± 0.19
4	Теосинте флоридский	1	1.96 ± 0.14
5	Ярутка	2	2.25 ± 0.83
6	Трипсакум ежевидный	2	2.19 ± 0.07
7	Теосинте балсасский	3	2.12 ± 0.11
8	Резуховидка Таля	4	1.90 ± 0.13
9	Дихантелия	6	2.27 ± 0.15
10	Водяная лилия	8	2.60 ± 0.39
11	Василистник анемоновидный	9	2.01 ± 0.16
12	Томат чилийский	15	1.97 ± 0.11
Итого		53	2.15 ± 0.08

Примечание. Здесь и в табл. 4 и 5: N – количество промоторов; M_0 – среднеарифметическая оценка; Δ – стандартная ошибка.

генов β-амилазы из 12 видов пищевых растений, а также оценки для 38 промоторов из 10 видов непищевых растений согласно расчетам веб-сервиса Plant_SNP_TATA_Z-tester (Rasskazov et al., 2022). Эта таблица характеризует пищевые растения диапазоном K_D от 1.30 ± 0.09 до 8.77 ± 7.36 нМ при среднеарифметической оценке 2.85 ± 0.21 нМ, тогда как у непищевых растений диапазон K_D был от 1.66 ± 0.32 до 6.75 ± 5.23 нМ при среднем

Таблица 4. Среднеарифметические оценки K_D комплексов ТВР с проксимальными промоторами генов растительных альбуминов

№ п/п	Вид растения	N	$K_D, M_0 \pm \Delta, \text{нМ}$
Пищевые растения			
1	Кунжут	1	2.28 ± 0.16
2	Киви	1	1.77 ± 0.12
3	Бразильский орех	1	3.04 ± 0.15
4	Восковница красная	2	1.96 ± 0.14
5	Дыня	2	2.04 ± 0.14
6	Соя культурная	2	1.65 ± 0.12
7	Горох посевной	4	4.49 ± 1.39
8	Перилла кустарниковая	5	1.98 ± 0.40
9	Миндаль обыкновенный	8	3.74 ± 0.66
10	Перец стручковый	8	3.00 ± 0.59
11	Мандарин уншиу	15	3.51 ± 0.61
12	Полынь однолетняя	35	3.07 ± 0.35
Итого		84	3.10 ± 0.22
Непищевые растения			
1	Водяная лилия	1	2.70 ± 0.22
2	Ярутка	1	2.00 ± 0.14
3	Стрига египетская	1	1.65 ± 0.12
4	Стрига желтая	1	2.19 ± 0.15
5	Резуховидка Таля	5	2.03 ± 0.21
6	Томат чилийский	13	2.33 ± 0.20
7	Василистник анемоновидный	15	2.11 ± 0.16
Итого		37	2.18 ± 0.10

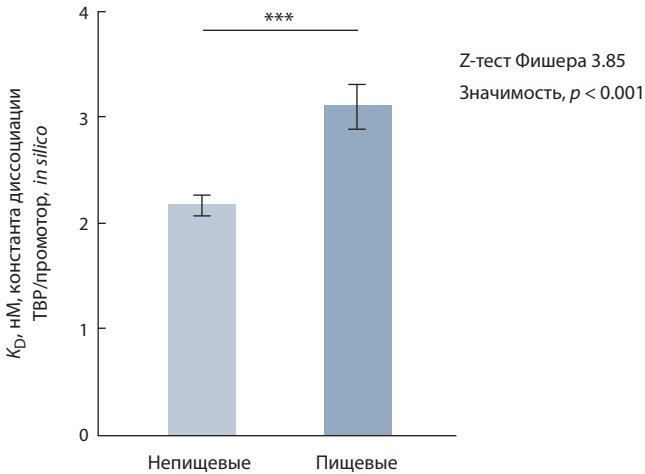


Рис. 2. Достоверное различие исследованных выборок пищевых и непищевых растений по оценке *in silico* K_D комплексов ТВР с проксимальными промоторами генов растительных ТВР к промоторам их генов, кодирующих альбумины.

Таблица 5. Среднеарифметические оценки K_D комплексов TBP с проксимальными промоторами генов растительных β -амилаз

№ п/п	Вид растения	N	$K_D, M_0 \pm \Delta, \text{нМ}$
Пищевые растения			
1	Чай	1	4.59 ± 0.28
2	Восковница красная	1	2.21 ± 0.13
3	Ячмень обыкновенный	2	1.30 ± 0.09
4	Гибискус сирийский	2	3.58 ± 1.86
5	Вишня Ёсина	2	3.19 ± 1.68
6	Ананас настоящий	3	8.77 ± 7.36
7	Миндаль обыкновенный	4	6.56 ± 1.63
8	Олива европейская	4	5.24 ± 0.93
9	Дыня	6	4.79 ± 0.96
10	Виноград	9	3.97 ± 0.73
11	Полынь однолетняя	16	4.29 ± 0.77
12	Перец стручковый	27	2.59 ± 0.23
Итого		77	3.89 ± 0.32
Непищевые растения			
1	Стрига желтая	1	3.50 ± 0.25
2	Вшивосемянник японский	1	3.43 ± 0.21
3	Генлиси	2	1.38 ± 0.82
4	Ярутка	2	1.93 ± 0.37
5	Китайская роза	2	1.79 ± 0.23
6	Резуховидка Таля	3	1.66 ± 0.32
7	Дихантелия	6	3.26 ± 0.43
8	Василистник анемоновидный	6	2.91 ± 0.43
9	Томат чилийский	6	2.89 ± 0.47
10	Водяная лилия	9	3.30 ± 0.62
Итого		38	2.85 ± 0.21

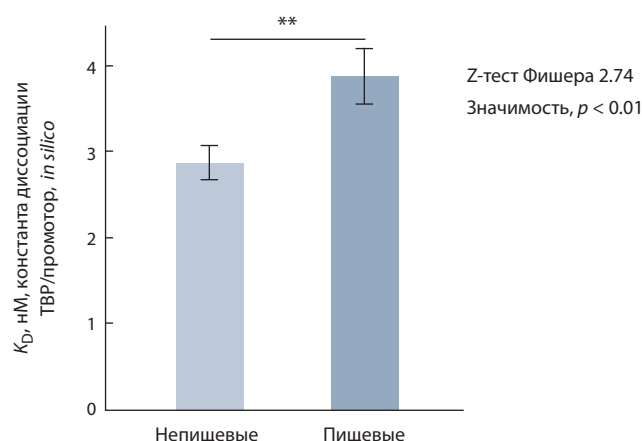


Рис. 3. Достоверное различие исследованных выборок пищевых и непищевых растений по оценке *in silico* K_D комплексов растительного TBP к промоторам их генов, кодирующих β -амилазы.

** – Достоверность $p < 0.01$ согласно Z-критерию Фишера.

$3.89 \pm 0.32 \text{ нМ}$. Сравнение проанализированных выборок пищевых и непищевых растений с использованием Z-критерия Фишера, согласно которому они статистически достоверно отличаются друг от друга при $Z = 2.74$ с уровнем статистической значимости $p < 0.01$, продемонстрировано на рис. 3.

Обсуждение

Общеизвестно, что в процессе стихийной доместикации предковых форм современных пищевых растений отбор шел, прежде всего, на их хозяйственно ценные признаки, такие как «урожайность», «устойчивость к патогенам и стрессовым воздействиям окружающей среды», а также «технологичность переработки в конечные продукты питания». В процессе одомашнивания также оценивались вкусовые качества продуктов питания и их значение для здоровья.

Остается неясным, проводился ли наряду с положительным отбором на полезные свойства сельскохозяйственных растений одновременно отбор против их вредоносных свойств, к числу которых можно отнести способность приготовленных из них пищевых продуктов вызывать аллергические реакции у человека. Для ответа на этот вопрос мы сконцентрировали внимание на поиске молекулярно-генетических маркеров селекции, ассоциированных со структурно-функциональной организацией проксимальных промоторов растительных генов.

Проанализированы растительные гены – гомологи трех генов мягкой пшеницы *T. aestivum*, кодирующих пищевые аллергены – β -амилазу, альбумин и глобулин, идентифицированные как мишени для атаки иммуноглобулинов класса E человека (Wang et al., 2021). Рассмотрены 363 гена-гомолога, относящихся к 28 и 15 видам пищевых и непищевых растений соответственно. Для каждого гена-гомолога с использованием веб-сервиса Plant_SNP_TATA_Z-tester (Rasskazov et al., 2022) была рассчитана величина K_D комплекса растительного TBP с проксимальным промотором этого гена.

Интерес к белку TBP и сайту его связывания в проксимальном промоторе (каноническая форма: TATA-бокс) обусловлен тем, что они играют ключевую роль в инициации транскрипции генов эукариот. Экспериментально установлено (Coleman, Pugh, 1995), что TBP скользит вдоль двойной спирали ДНК благодаря неспецифической аффинности между ними, $K_D \sim 10^{-5} \text{ М}$ (Hahn et al., 1989). Затем TBP останавливается на сайте TBP-связывания в силу их взаимного молекулярного распознавания (Berg, von Hippel, 1987; Bucher, 1990) из-за большей специфической аффинности TBP к этому сайту, $K_D \sim 10^{-9} \text{ М}$ (Hahn et al., 1989). Далее под действием TBP происходят плавление двойной спирали ДНК в сайте TBP-связывания и излом ее оси под прямым углом, что стабилизирует комплекс TBP/промотор (Flatters, Lavery, 1998). Сформировавшийся комплекс TBP/промотор принято считать обязательным «якорем» на ДНК, который необходим для связывания РНК полимеразы II (Muller et al., 2001; Martianov et al., 2002; Choukrallah et al., 2012; Rhee, Pugh, 2012) в качестве ключевого шага сборки преиницирующего комплекса транскрипции (Auble, 2009), ответственного за базальную транскрипцию (Fire et al., 1984). Вследствие ключевой

значимости ТАТА-боксов мутации, локализованные в проксимальных промоторах, оказывают выраженное влияние на уровень экспрессии генов (Савинкова и др., 2009).

Молекулярный механизм связывания ТВР с промоторами генов эукариот за три последовательных шага впервые был предложен в работе (Пономаренко и др., 2008) и позднее подтвержден экспериментально (Delgadillo et al., 2009). На основе этого механизма разработана биоинформатическая модель для расчета изменения K_D комплексов ТВР с проксимальными промоторами генов эукариот в случае полиморфизма сайтов ТВР-связывания в этих промоторах (Пономаренко и др., 2009). Результаты расчетов с использованием этой модели были подтверждены в независимых экспериментах *ex vivo* на культурах клеток, трансфицированных плазмидой pGL4.10 (Promega, США), несущей нормальный или мутантный вариант исследуемого промотора перед репортерным геном люциферазы (Ponomarenko et al., 2017), а также *in vitro* в реальном времени (Arkova et al., 2017) на спектрометре остановленной струи SX.20 (Applied Photophysics, Великобритания), в равновесных (Savinkova et al., 2013) и неравновесных (Drachkova et al., 2014) условиях гелетардации. В итоге столь разносторонней верификации этой биоинформатической модели на ее основе был создан веб-сервис Plant SNP_TATA_Z-tester (Rasskazov et al., 2022), использованный в данной работе для оценки K_D комплексов растительного ТВР с проксимальными промоторами генов пищевых и непищевых растений.

Осуществленный нами анализ показал, что пищевые растения, в сравнении с непищевыми, характеризуются значимо более низкими значениями аффинности ТВР к промоторам генов, гомологичных β -амилазе, альбумину и глобулину мягкой пшеницы как пищевым аллергенам ($p < 0.01$, Z-критерий Фишера). Рассматривая полученные результаты, примем во внимание тот экспериментально установленный факт, что уровень экспрессии генов эукариот повышается при увеличении аффинности ТВР к промоторам данных генов (Mogno et al., 2010). Это позволяет нам интерпретировать выявленное у пищевых растений, в сравнении с непищевыми, снижение аффинности ТВР к промоторам генов, гомологичных генам β -амилазы, альбумина и глобулина мягкой пшеницы как пищевых аллергенов. Это может свидетельствовать о проведенном человеком в прошлом отборе пищевых растений на пониженное содержание этих аллергенных белков.

Заключение

В настоящей работе у 28 пищевых и 15 непищевых видов растений впервые единообразно проанализированы *in silico* 363 последовательности ДНК длиной 90 п. о. проксимальных промоторов генов, гомологичных генам мягкой пшеницы, кодирующим пищевые аллергены: глобулин, альбумин и β -амилазу (Wang et al., 2021). В результате обнаружены пониженные оценки *in silico* аффинности ТВР к промоторам исследуемых генов пищевых растений в сравнении с непищевыми. Это может свидетельствовать в пользу имевшего место искусственного отбора пищевых растений, направленного на пониженный уровень экспрессии генов, кодирующих аллергенные белки, который осуществлялся человеком в прошлом.

Список литературы / References

- Пономаренко П.М., Пономаренко М.П., Драчкова И.А., Лысова М.В., Аршинова Т.В., Савинкова Л.К., Колчанов Н.А. Прогноз изменения аффинности ТАТА-связывающего белка к ТАТА-боксам в результате полиморфизмов ТАТА-боксов промоторов генов человека. *Молекуляр. биология*. 2009;43(3):512-520. [Ponomarenko P.M., Ponomarenko M.P., Drachkova I.A., Lysova M.V., Arshinova T.V., Savinkova L.K., Kolchanov N.A. Prediction of the affinity of the TATA-binding protein to TATA boxes with single nucleotide polymorphisms. *Mol. Biol. (Moscow)*. 2009; 43(3):472-479. DOI 10.1134/S0026893309030157.]
- Пономаренко П.М., Савинкова Л.К., Драчкова И.А., Лысова М.В., Аршинова Т.В., Пономаренко М.П., Колчанов Н.А. Пошаговая модель связывания ТВР/ТАТА-бокс позволяет предсказать наследственное заболевание человека по точечному полиморфизму. *Докл. Академии наук*. 2008;419(6):828-832. [Ponomarenko P., Savinkova L., Drachkova I., Lysova M., Arshinova T., Ponomarenko M., Kolchanov N. A step-by-step model of TBP/TATA box binding allows predicting human hereditary diseases by single nucleotide polymorphism. *Dokl. Biochem. Biophys.* 2008; 419:88-92. DOI 10.1134/S1607672908020117.]
- Савинкова Л.К., Пономаренко М.П., Пономаренко П.М., Драчкова И.А., Лысова М.В., Аршинова Т.В., Колчанов Н.А. Полиморфизмы ТАТА-боксов промоторов генов человека и ассоциированные с ними наследственные патологии. *Биохимия*. 2009; 74(2):149-163. [Savinkova L.K., Ponomarenko M.P., Ponomarenko P.M., Drachkova I.A., Lysova M.V., Arshinova T.V., Kolchanov N.A. TATA box polymorphisms in human gene promoters and associated hereditary pathologies. *Biochemistry (Moscow)*. 2009;74(2):117-129. DOI 10.1134/s0006297909020011.]
- Arkova O., Kuznetsov N., Fedorova O., Savinkova L. A real-time study of the interaction of TBP with a TATA box-containing duplex identical to an ancestral or minor allele of human gene *LEP* or *TPI*. *J. Biomol. Struct. Dyn.* 2017;35(14):3070-3081. DOI 10.1080/07391102.2016.1241190.
- Auble D.T. The dynamic personality of TATA-binding protein. *Trends Biochem. Sci.* 2009;34(2):49-52. DOI 10.1016/j.tibs.2008.10.008.
- Benson D.A., Clark K., Karsch-Mizrachi I., Lipman D.J., Ostell J., Sayers E.W. GenBank. *Nucleic Acids Res.* 2015;43(Database issue):D30-D35. DOI 10.1093/nar/gku1216.
- Berg O.G., von Hippel P.H. Selection of DNA binding sites by regulatory proteins. Statistical-mechanical theory and application to operators and promoters. *J. Mol. Biol.* 1987;193(4):723-750. DOI 10.1016/0022-2836(87)90354-8.
- Bucher P. Weight matrix descriptions of four eukaryotic RNA polymerase II promoter elements derived from 502 unrelated promoter sequences. *J. Mol. Biol.* 1990;212(4):563-578. DOI 10.1016/0022-2836(90)90223-9.
- Cavazza A., Mattarozzi M., Franzoni A., Careri M. A spotlight on analytical prospects in food allergens: From emerging allergens and novel foods to bioplastics and plant-based sustainable food contact materials. *Food Chem.* 2022;388:132951. DOI 10.1016/j.foodchem.2022.132951.
- Choukrallah M.A., Kobi D., Martianov I., Pijnappel W.W., Mischerikow N., Ye T., Heck A.J., Timmers H.T., Davidson I. Interconversion between active and inactive TATA-binding protein transcription complexes in the mouse genome. *Nucleic Acids Res.* 2012;40(4):1446-1459. DOI 10.1093/nar/gkr802.
- Coleman R.A., Pugh B.F. Evidence for functional binding and stable sliding of the TATA binding protein on nonspecific DNA. *J. Biol. Chem.* 1995;270(23):13850-13859. DOI 10.1074/jbc.270.23.13850.
- Delgadillo R.F., Whittington J.E., Parkhurst L.K., Parkhurst L.J. The TATA-binding protein core domain in solution variably bends TATA sequences via a three-step binding mechanism. *Biochemistry*. 2009; 48(8):1801-1809. DOI 10.1021/bi8018724.
- Drachkova I., Savinkova L., Arshinova T., Ponomarenko M., Peltek S., Kolchanov N. The mechanism by which TATA-box polymorphisms

- associated with human hereditary diseases influence interactions with the TATA-binding protein. *Hum. Mutat.* 2014;35(5):601-608. DOI 10.1002/humu.22535.
- Fire A., Samuels M., Sharp P.A. Interactions between RNA polymerase II, factors, and template leading to accurate transcription. *J. Biol. Chem.* 1984;259(4):2509-2516. DOI 10.1016/S0021-9258(17)43382-5.
- Flatters D., Lavery R. Sequence-dependent dynamics of TATA-Box binding sites. *Biophys. J.* 1998;75(1):372-381. DOI 10.1016/S0006-3495(98)77521-6.
- Hahn S., Buratowski S., Sharp P.A., Guarente L. Yeast TATA-binding protein TFIID binds to TATA elements with both consensus and nonconsensus DNA sequences. *Proc. Natl. Acad. Sci. USA.* 1989; 86(15):5718-5722. DOI 10.1073/pnas.86.15.5718.
- Hong L., Pan M., Xie X., Liu K., Yang J., Wang S., Wang S. Aptamer-based fluorescent biosensor for the rapid and sensitive detection of allergens in food matrices. *Foods.* 2021;10(11):2598. DOI 10.3390/foods10112598.
- IUPAC-IUB Commission on Biochemical Nomenclature (CBN). Abbreviations and symbols for nucleic acids, polynucleotides and their constituents. Recommendations 1970. *Biochem. J.* 1970;120(3): 449-454. DOI 10.1042/bj1200449.
- Martianov I., Viville S., Davidson I. RNA polymerase II transcription in murine cells lacking the TATA binding protein. *Science.* 2002; 298(5595):1036-1039. DOI 10.1126/science.1076327.
- Mogno I., Vallania F., Mitra R.D., Cohen B.A. TATA is a modular component of synthetic promoters. *Genome Res.* 2010;20(10):1391-1397. DOI 10.1101/gr.106732.110.
- Muller F., Lakatos L., Dantonel J., Strahle U., Tora L. TBP is not universally required for zygotic RNA polymerase II transcription in zebrafish. *Curr. Biol.* 2001;11(4):282-287. DOI 10.1016/S0960-9822(01)00076-8.
- Ponomarenko P., Chadaeva I., Rasskazov D.A., Sharypova E., Kashina E.V., Drachkova I., Zhechev D., Ponomarenko M.P., Savinkova L.K., Kolchanov N. Candidate SNP markers of familial and sporadic Alzheimer's diseases are predicted by a significant change in the affinity of TATA-binding protein for human gene promoters. *Front. Aging Neurosci.* 2017;9:231. DOI 10.3389/fnagi.2017.00231.
- Prescott S.L., Logan A.C., Bristow J., Rozzi R., Moodie R., Redvers N., Haahela T., Warber S., Poland B., Hancock T., Berman B. Exiting the anthropocene: achieving personal and planetary health in the 21st century. *Allergy.* 2022;77(12):3498-3512. DOI 10.1111/all.15419.
- Rasskazov D., Chadaeva I., Sharypova E., Zolotareva K., Khandaev B., Ponomarenko P., Podkolodny N., Tverdokhlebo N., Vishnevsky O., Bogomolov A., Podkolodnaya O., Savinkova L., Zemlyanskaya E., Golubyatnikov V., Kolchanov N., Ponomarenko M. 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.
- Rhee H., Pugh B. Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature.* 2012;483(7389):295-301. DOI 10.1038/nature10799.
- Savinkova L., Drachkova I., Arshinova T., Ponomarenko P., Ponomarenko M., Kolchanov N. An experimental verification of the predicted effects of promoter TATA-box polymorphisms associated with human diseases on interactions between the TATA boxes and TATA-binding protein. *PLoS One.* 2013;8(2):e54626. DOI 10.1371/journal.pone.0054626.
- 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.

ORCID ID

O.V. Vishnevsky orcid.org/0000-0002-0347-2252
I.V. Chadaeva orcid.org/0000-0002-2724-5441
E.B. Sharypova orcid.org/0000-0002-1467-9312
P.M. Ponomarenko orcid.org/0000-0003-2715-9612
N.L. Podkolodny orcid.org/0000-0001-9132-7997

D.A. Rasskazov orcid.org/0000-0003-4795-0954
A.G. Bogomolov orcid.org/0000-0003-4359-6089
O.A. Podkolodnaya orcid.org/0000-0003-3247-0114
L.K. Savinkova orcid.org/0000-0003-4543-4104
E.V. Zemlyanskaya orcid.org/0000-0002-4304-1190
M.P. Ponomarenko orcid.org/0000-0003-1663-318X

Благодарности. Работа выполнена при поддержке Российского научного фонда, грант № 20-14-00140. Авторы выражают благодарность Центру коллективного пользования «Биоинформатика» за использование вычислительных ресурсов в рамках бюджетного проекта Правительства РФ FWNR-2022-0020 и Федеральной научно-технической программы развития генетических технологий России.

Конфликт интересов. Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 16.09.2022. После доработки 29.11.2022. Принята к публикации 30.11.2022.