

DOI: 10.18523/2617-4529.2024.7.27-35

УДК 633.11:575:591.151

Созінова О. І.^{1,2}, Блюм Я. Б.²

¹ Інститут захисту рослин НААН, Україна

² Державна установа «Інститут харчової біотехнології і геноміки НАН України», Україна

АНАЛІЗ ОДНОНУКЛЕОТИДНИХ ПОЛІМОРФІЗМІВ У ПОСЛІДОВНОСТЯХ ГЕНІВ *Pina* ТА *Pinb* ДИПЛОЇДНИХ ПШЕНИЦЬ *Triticum monococcum* І *T. urartu*

Пуруіндоліни (пуруіндолін а і пуруіндолін б) – низькомолекулярні білки, що визначають текстуру ендосперму зерна видів триб *Triticeae* та *Avena*. Метою цієї роботи був аналіз однонуклеотидних поліморфізмів (SNP) кодуєчих послідовностей генів *Pina* та *Pinb* диплоїдних пшениць *Triticum monococcum* (*A^mA^m*) і *T. urartu* (*AA*) з бази даних NCBI порівняно з референсними послідовностями сорту пшениці м'якої *Chinese Spring* (*CS*).

З бази даних NCBI було відібрано 62 послідовності гена пуруіндоліну а *T. monococcum* (*Pina-A^m1*), 22 послідовності *Pina T. urartu* (*Pina-A1*), 32 послідовності гена пуруіндоліну б *T. monococcum* (*Pinb-A^m1*) і 13 послідовностей *Pinb T. urartu* (*Pinb-A1*). Як референсні послідовності використовували послідовності гена пуруіндоліну а (алель *Pina-D1a*) *DQ363911.1* сорту *CS* та гена пуруіндоліну б (алель *Pinb-D1a*) *DQ363913.1* *CS* з бази даних NCBI. Послідовності вирівнювали з використанням програми MEGA 11. Всього в загальній вибірці 84 послідовностей *Pina* диплоїдних пшениць ідентифіковано 34 SNP (13 синонімічних відмінностей, 21 несинонімічних, з яких 15 призводять до радикальних амінокислотних замін, 6 – до консервативних), частина з яких спостерігалась у всіх послідовностях, а частина були поодинокими. Серед 45 послідовностей *Pinb* трапляється 36 SNP, але, на відміну від гена *Pina*, тут переважали синонімічні заміни (22); 7 замін призводили до радикальних амінокислотних замін та ще 7 – до консервативних. Заміни в послідовностях пуруіндолінових генів відносно генів *CS* можна розподілити на фіксовані в обох видів диплоїдних пшениць, ті, що фіксовані в *T. urartu* і за якими є поліморфізм у *T. monococcum*, та специфічні для кожного виду. Виявлено істотні відмінності в частотах трапляння альтернативних нуклеотидів у певних позиціях (81, 318, 322 і 384 *Pina* та 135 і 359 *Pinb*) у дикорослої пшениці *T. monococcum ssp. aegilopoides* і культивованої однозернянки *T. monococcum ssp. monococcum*.

Ключові слова: пуруіндолін, *Triticum monococcum ssp. aegilopoides*, *T. monococcum ssp. monococcum*, *T. urartu*, SNP, радикальні амінокислотні заміни, консервативні амінокислотні заміни.

Вступ

Однією з головних ознак якості зерна пшениці м'якої *Triticum aestivum* L. (геномна формула AABBDD) є текстура ендосперму, за якою сорти поділяють на твердозерні та м'яззерні [1-3]. Рівень твердозерності [4] визначається локусом *Ha* на хромосомі 5D, а саме алелями генів пуруіндоліну а (*Pina-D1*) і пуруіндоліну б (*Pinb-D1*) [1,5]. Пуруіндоліни – основні білки з молекулярною масою 13 кДа, які мають 10 цистеїнових залишків та триптофан-багатий домен [6], пов'язаний

з антимікробними властивостями цих білків [7]. Довжина кодуєчої послідовності пуруіндолінових генів – 450 п.н. (447 + стоп-кодон). Подібно до проламінових генів вони також не містять інтронів [6]. Пуруіндоліни синтезуються з N-термінальним сигнальним пептидом, який відщеплюється при їх переході в ендоплазматичний ретикулум, також ще додатково відщеплюються фрагменти з N- і C-кінців [6]. У локусі *Ha* також міститься ген, що кодує білок GSP-1 (grain softness protein), пов'язаний з м'язозерністю, хоча для

нього впливу на текстуру зерна не показано [8,9]. У *T. aestivum* і *T. turgidum* (AABB) гени пуруіндолінів відсутні на хромосомах 5A і 5B [10] через великі геномні делеції [11], тоді як ген *Gsp-1* там присутній [9]. Гени *Pina-1* та *Pinb-1* є на хромосомі 5A^m пшениці-однозернянки *T. monococcum* (A^mA^m) [10,12], 5A *T. urartu* (AA), 5A *T. timopheevii* (AAA^mA^mGG), проте відсутні в G геномі цього виду [13]. У *T. zhukovskiyi* (AAA^mA^mGG) гени *Pina-1* та *Pinb-1* присутні лише на хромосомі 5A^m [13].

Загалом, за винятком *T. turgidum*, пуруіндолінові гени *Pina* та *Pinb* або відповідні індолінові гени (гордоіндолінів у ячменю, секалоіндолінів у жита, авеноіндолінів та вроміндолінів у вівса) [14] присутні у видів злаків з триб Triticeae та Avenae (Poae), для яких характерна м'яка текстура ендосперму: егілопсів, ячменю, жита, вівса, видів роду *Elymus*, *Agropyron* та ін. [10,13,15,16]. Вважається, що предковий пуруіндоліновий ген виник як новий проламіновий ген у спільного предка підродин Pooideae та Oryzoideae (syn. Ehrhartoideae) після їх відділення від Panicoideae та був втрачений у Oryzoideae [15]. Тому пуруіндоліни зараховують до надродини проламінів, яка містить власне проламіни (гліадини і глютеніни), білки-переносники ліпідів, інгібітори екзогенних α -амілаз і трипсину та α -глобуліни [14].

До диплоїдних пшениць належать два види – *T. monococcum* та *T. urartu* [17]. *T. monococcum* має два підвиди: *T. monococcum* ssp. *aegilopoides* – дикоросла пшениця, яка є предковою формою domestикованого (культивованого) підвиду *T. monococcum* ssp. *monococcum*. *T. monococcum* ssp. *aegilopoides* поширений у північно-східному Середземномор'ї та в західній Азії. Інший вид, *T. urartu*, який є донором геному A *T. turgidum* і, відповідно, *T. aestivum* – дикорослий вид, ареал якого збігається з ареалом *T. monococcum* ssp. *aegilopoides* (Вірменія, південно-східна Туреччина, північ Іраку, північно-західна частина Ірану) [17]. Пшениця-однозернянка характеризується як екстрем'якозерна [18,19]. Синтетичні гексаплоїди від міжвидового схрещення між *T. durum* і *T. monococcum* ssp. *aegilopoides* є м'якозерними [20]. Секвенування пуруіндолінових генів диплоїдних пшениць було проведено в багатьох дослідженнях [12,13,21-23], а послідовності цих генів можна знайти в базі даних NCBI (National Center for Biotechnology Information). Метою нашої роботи був аналіз однострункових поліморфізмів (SNP) кодуючих послідовностей генів *Pina* та *Pinb* *T. monococcum* і *T. urartu* з бази даних NCBI порівняно з референсними послідовностями дикого типу сорту пшениці м'якої Chinese Spring.

Матеріали і методи

З бази даних NCBI було відібрано такі послідовності генів *Pina* і *Pinb* диплоїдних пшениць. Послідовності гена *Pina* *T. monococcum*: EU329377.1, EU329376.1, EU329375.1, EU329374.1, EU329371.1, EU329369.1, EU329367.1, EU307592.1, EU307591.1, EU329372.1, EU329368.1, EU329373.1, EU329370.1, EU329366.1, EU329365.1, AY622786.1, EU329362.1, EU329360.1, EU329359.1, EU329350.1, EU329348.1, EU329344.1, EU329347.1, AJ302092.1, EU329364.1, EU329363.1, EU329361.1, EU329358.1, EU329357.1, EU329356.1, EU329355.1, EU329354.1, EU329352.1, EU329353.1, EU329351.1, EU329349.1, EU329346.1, EU329345.1, EU329343.1, EU329342.1, EU329341.1, EU329340.1, EU329339.1, EU329338.1, EU329337.1, EU329336.1, FJ898231.1, AJ302093.1, HQ696585.1, HQ696587.1, HQ696586.1, EU268473.1, EU307585.1, AJ242715.2, AJ249933.1, DQ269824.1, DQ269823.1, DQ269822.1, DQ269821.1, DQ269820.1, DQ269819.1, DQ269825.1. Послідовності гена *Pinb* *T. monococcum*: AY622798.1, AY622797.1, AY622799.1, EU268487.1, EF382940.1, EF382939.1, EF382936.1, EF382935.1, EF382934.1, EF382933.1, EF382932.1, EF382931.1, EF382930.1, EF382929.1, EF382928.1, HQ696594.1, HQ696593.1, HQ696592.1, HQ696591.1, HQ696590.1, FJ898262.1, EU307611.1, EU307610.1, EU307608.1, AJ302102.1, AJ302101.1, DQ269857.1, DQ269856.1, DQ269855.1, DQ269854.1, DQ269853.1, DQ269852.1. Послідовності гена *Pina* *T. urartu*: EU268495.1, EU329391.1, EU329390.1, EU329389.1, EU329387.1, EU329386.1, EU329385.1, EU329384.1, EU329383.1, EU329382.1, EU329381.1, EU329380.1, EU329379.1, EU329378.1, HQ696584.1, EU307590.1, EU307589.1, AJ302094.1, AJ302095.1, DQ269828.1, DQ269827.1, DQ269826.1. Послідовності гена *Pinb* *T. urartu*: EU268479.1, EF382938.1, EF382937.1, HQ696589.1, HQ696588.1, EU307613.1, EU307612.1, AJ302104.1, AJ302103.1, DQ269861.1, DQ269860.1, DQ269859.1, DQ269858.1.

Як референсну послідовність використовували послідовність гена пуруіндоліну a (алель *Pina-D1a*) DQ363911.1 сорту Chinese Spring (CS) та гена пуруіндоліну b (алель *Pinb-D1a*) DQ363913.1 CS з бази даних NCBI. Послідовності вирівнювали з використанням програми MEGA 11 [24]. Аналізували SNP (відмінності від референсної послідовності, які для зручності ще називали «заміни») лише в кодуючих послідовностях. Для характеристики нуклеотидних і амінокислотних заміни використовували класифікацію, наведену в [25]. Для аналізу відмінностей частот використовували критерій χ^2 .

Результати і обговорення

З бази даних NCBI було відібрано 62 послідовності гена пуруіндоліну a *T. monococcum*

(*Pina-A^m1*) і 22 послідовності *T. urartu* (*Pina-A1*). Вирівнювання послідовностей дало змогу охарактеризувати відмінності в певних позиціях нуклеотидів (SNP) порівняно з референсною послідовністю гена сорту CS (алель *Pina-D1a*, який називають алелем «дикого типу») (табл. 1). Всього в загальній вибірці 84 послідовностей було ідентифіковано 34 SNP (13 синонімічних відмінностей, 21 несинонімічних, з яких 15 призводять до радикальних амінокислотних замін, 6 – до консервативних), частина з яких спостерігалась у всіх послідовностях, а частина були поодинокими. Всі послідовності *Pina T. monococcum* і *T. urartu* мали відмінності у 5 позиціях: 24 (A→C), 70 (G→A), 121 (G→C), 230 (A→T), 257 (G→A). Заміна в позиції 24 є синонімічною, несинонімічні заміни в позиціях 70 і 121 є консервативними (Val→Ile та Val→Leu, відповідно), тоді як заміни в позиціях 230 та 257 призводять до радикальної заміни амінокислот (Gln→Leu та Arg→Gln, відповідно). За позиціями 318 і 322 заміни T→C і G→A, з яких друга призводить до радикальної заміни Gly→Ser, присутні у всіх послідовностях *T. urartu* та в близько половини послідовностей *T. monococcum*. Ще три синонімічні заміни в позиціях 81, 249 і 384 трапляються тільки в майже половини зразків *T. monococcum*. Вказані тут позиції збігаються з SNP, ідентифікованими в дослідженнях Guzman et al. [23] під час аналізу 21 зразка *T. monococcum* та 7 зразків *T. urartu*.

Порівняння частот трапляння альтернативних нуклеотидів у позиціях 81, 318, 322 і 384 показало істотні відмінності в дикій і культивованій однозернянок (табл. 2). Заміна в позиції 81 частіше трапляється в підвиді *aegilopoides*, тоді як заміни в позиціях 318, 322 і 384 більш характерні для культивованого підвиду. Однак лише одна з цих замін (322) призводить до амінокислотної заміни.

Крім замін, що є фіксованими або трапляються з високою частотою, серед послідовностей гена *Pina* є низка рідкісних замін, які містяться в певних позиціях лише в одній послідовності гена *Pina* (табл. 1). Переважну більшість таких SNP ідентифіковано Chen et al. [22]. У *T. monococcum* виявлено 15 рідкісних замін, із них 5 синонімічних і 10 несинонімічних, що призводять до 8 радикальних і 2 консервативних амінокислотних замін. У *T. urartu* трапляються 8 поодиноких замін, серед яких 3 синонімічні, 4 призводять до радикальних амінокислотних замін, 1 – до консервативної. Співвідношення різних видів рідкісних замін є подібним у двох видів диплоїдних пшениць ($\chi^2 = 0,1$). Таке саме співвідношення зберігається і в результаті аналізу загальної кількості нуклеотидних замін у послідовностях (10 синонімічних, 11 радикальних, 5 консервативних у *T. monococcum*, 5, 8 і 3,

відповідно, у *T. urartu*). Низька частота таких замін вказує на відносно недавнє виникнення таких мутацій. Зразки з радикальними амінокислотними замінами є потенційним джерелом пуринолінів з новими властивостями.

Для аналізу замін у гені *Pinb* проведено вирівнювання 32 послідовностей *T. monococcum* (*Pinb-A^m1*) і 13 послідовностей *T. urartu* (*Pinb-A1*). Серед них позицій із замінами було 36, але, на відміну від гена *Pina*, тут переважали синонімічні заміни (22); замін, які приводили до радикальних і консервативних амінокислотних замін, було по 7 кожного типу (табл. 3). У всіх проаналізованих послідовностях *Pinb* диплоїдних пшениць обох видів було зафіксовано заміни у 22 позиціях: 96, 98, 120, 125, 171, 201, 228, 237, 259, 285, 294, 306, 307, 310, 323, 339, 342, 357, 370, 411, 423, 426. Синонімічна заміна в позиції 24 (T→C) присутня у всіх послідовностях *T. urartu* та в половини послідовностей *T. monococcum*. В однозернянок спостерігається поліморфізм у позиціях 135, 243, 359, де достатньо висока частота послідовностей із заміною, якої немає у *T. urartu*. У свою чергу, лише в *T. urartu* поліморфними є позиції 210, 290 та зафіксована заміна в позиції 75, і вони є синонімічними. Заміни в позиціях 135 і 359 частіше трапляються в дикого підвиду, ніж у культивованого (табл. 2). Якщо перша заміна є синонімічною, то друга є консервативною (Val→Ala). Переважну кількість замін у *Pinb* описано Guzman et al. [23] у процесі аналізу послідовностей 28 зразків диплоїдних пшениць. Додатково серед наявних у базі NCBI послідовностей є рідкісні заміни в позиціях 10 (T→A), 196 і 410 (G→A). Загалом, за співвідношенням різних типів замін у гені *Pinb* обидва види диплоїдних пшениць практично не відрізняються.

Отже, заміни в послідовностях пуринолінових генів можна розподілити на фіксовані в обох видів, фіксовані в *T. urartu* і ті, за якими є поліморфізм у *T. monococcum*, та специфічні для кожного виду. Крім того, для обох видів співвідношення трьох типів замін відрізняється для генів *Pina* та *Pinb* ($\chi^2 = 13,6$, $P < 0,01$). За розрахунками Guzman et al. [23], нуклеотидна різноманітність *Pina* є значно нижчою, ніж очікуване значення за тестом на нейтральність, на відміну від *Pinb*, що, на думку цих авторів, свідчить про різні селективні обмеження в еволюції цих генів. Massa and Moris [21] серед сайтів гена *Pina*, що підлягали позитивному добору, виділили сайти 121 (V→L) і 230 (Q→L) і також не виявили дії позитивного добору на ген *Pinb*. Такі відмінності автори пояснюють можливою більш високою антимікробною активністю пуриноліну а, чий триптофановий домен містить більше залишків

Таблиця 1

Відмінності в послідовності гена пуруіндоліну а *Pina T. monococcum (T.mon)* і *T. urartu (Tur)* від референсної послідовності сорту Chinese Spring (CS) та відмінності в амінокислотній послідовності

N	Позиція в гені <i>Pina</i>	CS	<i>T.mon/ Tur</i>	Позиція в кодоні	Кількість зразків <i>T.mon</i> із заміною	Частота у <i>T.mon</i>
1	11	t	c	2	1	0,016
2	24	a	c	3	62	1,000
3	56	c	a	2	1	0,016
4	70	g	a	1	62	1,000
5	81	t	c	3	28	0,452
6	121	g	c	1	62	1,000
7	158	t	g	2		
8	161	a	g	2		
9	164	g	a	2	1	0,016
10	182	a	g	2	1	0,016
11	213	g	c	3	1	0,016
12	222	a	c	3	1	0,016
13	230	a	t	2	62	1,000
14	249	t	c	3	25	0,403
15	253	a	g	1	1	0,016
16	257	g	a	2	62	1,000
17	273	a	g	3	1	0,016
18	283	c	t	1	1	0,016
19	318	t	c	3	33	0,532
20	322	g	a	1	37	0,597
21	348	t	c	3		
22	358	a	g	1		
23	362	t	c	2		
24	372	a	g	3	1	0,016
25	380	a	c	2		
26	384	g	a	3	32	0,516
27	392	g	a	2	3	0,048
28	409	c	t	1	1	0,016
29	410	c	t	2	1	0,016
30	415	a	g	1		
31	416	a	t	2	1	0,016
32	418	a	g	1	1	0,016
33	421	c	t	1	1	0,016
34	441	c	a	3		

триптофану. Водночас результати досліджень Kim et al. [26] свідчать про кооперативну дію пуруіндоліну а і b для запобігання розпаду полярних ліпідів у процесі дозрівання зерна. Також було показано взаємодію пуруіндолінів із власне проламінами [27,28], зокрема вплив пуруіндоліну а на агрегацію проламінів у процесі розвитку зерна [29], що, у свою чергу, може визначати різне адаптивне значення певних замін, і як результат – пуруіндолінових алелів.

Висновки

У результаті аналізу послідовностей пуруіндолінових генів *T. monococcum* і *T. urartu* з бази даних NCBI загалом ідентифіковано 34 SNP

у гені *Pina* і 36 у гені *Pinb*, частина з яких були рідкісними, та охарактеризовано частоту їх трапляння в обох видів. Встановлено відмінності в частотах трапляння альтернативних нуклеотидів у певних позиціях (81, 318, 322 і 384 *Pina* та 135 і 359 *Pinb*) у дикої і культивованої однозернянок. Гени *Pina* та *Pinb* відрізняються за співвідношенням синонімічних, консервативних і радикальних замін. Отже, виявлені зразки диплоїдних пшениць із несинонімічними замінами в пуруіндолінових генах можуть слугувати джерелом білків із новими властивостями для розширення генофонду пшениці м'якої щодо ознак м'якозерності та антимікробних властивостей пуруіндолінів.

Продовження табл. 1

N	Кількість зразків <i>T.ur</i> із заміною	Частота у <i>T.ur</i>	Позиція в амінокислотній послідовності	CS	<i>T.mon/T.ur</i>	Характер заміни*
1			4	L	P	рад
2	22	1,000	8	G	G	
3			19	A	E	рад
4	22	1,000	24	V	I	конс
5			27	S	S	
6	22	1,000	41	V	L	конс
7	1	0,045	53	L	R	рад
8	1	0,045	54	D	G	рад
9			55	R	Q	рад
10			61	D	G	рад
11			71	W	C	рад
12			74	G	G	
13	22	1,000	77	Q	L	рад
14			83	C	C	
15			85	S	G	рад
16	22	1,000	86	R	Q	рад
17			91	P	P	
18			95	R	K	конс
19	22	1,000	106	D	D	
20	22	1,000	108	G	S	рад
21	1	0,045	116	D	D	
22	1	0,045	120	K	E	рад
23	1	0,045	121	V	A	конс
24			124	E	E	
25	1	0,045	127	N	T	рад
26			128	L	L	
27			131	R	K	конс
28			137	P	F	рад
29			137	P	P	
30	1	0,045	139	N	I	рад
31			139	N	I	рад
32			140	I	V	конс
33			141	P	P	
34	1	0,045	147	Y	Y	

* рад – радикальна; конс – консервативна.

Таблиця 2

Відмінності в частоті трапляння нуклеотидів у певних позиціях гена *Pina* і *Pinb* у дикої і культурної однозернянок

Ген	SNP	subsp. <i>aegilopoides</i>	subsp. <i>monococcum</i>	χ^2	P
<i>Pina</i>	81T	6	25	7,7	0,006
	81C	13	11		
	318T	13	11	7,7	0,006
	318C	6	26		
	322G	13	9	10,2	0,001
	322A	6	28		
	384G	15	11	12,23	0,0005
384A	4	26			
<i>Pinb</i>	135G	4	12	10,4	0,001
	135A	9	1		
	359T	2	9	7,7	0,006
	359C	11	4		

Таблиця 3

Відмінності в послідовності гена пуроіндоліну b *Pinb* *T. monocossum* (*T.mon*) і *T. urartu* (*T.ur*) від референсної послідовності сорту Chinese Spring (CS) та відмінності в амінокислотній послідовності

N	Позиція в гені <i>Pinb</i>	CS	<i>T.mon</i> / <i>T.ur</i>	Позиція в кодоні	Кількість зразків <i>T.mon</i> із заміною	Частота у <i>T.mon</i>
1	10	T	A	1	1	0,031
2	24	T	C	3	16	0,500
3	75	C	T	3		
4	85	A	G	1	3	0,094
5	96	C	T	3	32	1,000
6	98	G	C	2	32	1,000
7	120	T	C	3	32	1,000
8	125	A	T	2	32	1,000
9	135	G	A	3	12	0,375
10	159	C	T	3	3	0,094
11	171	A	G	3	32	1,000
12	196	G	A	1	1	0,031
13	201	C	T	3	32	1,000
14	210	A	G	3		
15	228	C	T	3	32	1,000
16	237	T	C	3	32	1,000
17	243	T	C	3	25	0,781
18	259	A	C	1	32	1,000
19	285	A	G	3	32	1,000
20	290	G	T	2		
21	294	T	C	3	32	1,000
22	306	G	A	3	32	1,000
23	307	C	G	1	32	1,000
24	310	G	A	1	32	1,000
25	323	G	A	2	32	1,000
26	339	G	C	3	32	1,000
27	342	C	A	3	32	1,000
28	357	G	T	3	32	1,000
29	359	T	C	2	18	0,563
30	370	C	A	1	32	1,000
31	410	G	A	2	1	0,031
32	411	C	A	3	32	1,000
33	414	C	T	3	1	0,031
34	423	G	A	3	32	1,000
35	424	T	C	1	3	0,094
36	426	C	A	3	32	1,000

Продовження табл. 3

N	Кількість зразків <i>T.ur</i> із заміною	Частота у <i>T.ur</i>	Позиція в амінокислотній послідовності	CS	<i>T.mon/T.ur</i>	Характер заміни*
1			4	L	I	конс
2	13	1,000	8	A	A	
3	13	1,000	25	G	G	
4			29	N	D	конс
5	13	1,000	32	G	G	
6	13	1,000	33	G	A	конс
7	13	1,000	40	C	C	
8	13	1,000	42	Q	L	рад
9			45	P	P	
10			53	Y	Y	
11	13	1,000	57	R	R	
12			66	V	V	
13	13	1,000	67	T	T	
14	5	0,385	70	T	T	
15	13	1,000	76	G	G	
16	13	1,000	79	H	H	
17			81	V	V	
18	13	1,000	87	K	Q	рад
19	13	1,000	95	Q	Q	
20	4	0,308	97	R	R	
21	13	1,000	98	C	C	
22	13	1,000	102	R	R	
23	13	1,000	103	R	G	рад
24	13	1,000	104	V	M	рад
25	13	1,000	108	R	K	конс
26	13	1,000	113	L	F	рад
27	13	1,000	114	G	G	
28	13	1,000	119	E	D	конс
29			120	V	A	конс
30	13	1,000	124	L	I	конс
31			137	G	E	рад
32	13	1,000	137	G	G	
33			138	A	A	
34	13	1,000	141	K	K	
35			142	F	L	рад
36	13	1,000	142	F	F	

* рад – радикальна; конс – консервативна.

References

1. Pauly A, Pareyt B, Fierens E, Delcour JA. Wheat (*Triticum aestivum* L. and *T. turgidum* L. ssp. *durum*) kernel hardness: I. Current view on the role of puroindolines and polar lipids. *Comprehensive Reviews in Food Science and Food Safety*. 2013;12(4):413-26. doi:10.1111/1541-4337.12019
2. Shewry P. Wheat grain proteins: Past, present, and future. *Cereal Chemistry*. 2023;100(1):9-22. doi:10.1002/cche.10585
3. Rai A, Han SS. Recent developments on the contribution of glutenin and puroindoline proteins to improve wheat grain quality. *Cereal Chemistry*. 2023;100(1). doi:10.1002/cche.10607
4. Morris CF. Puroindolines: The molecular genetic basis of wheat grain hardness. *Plant Molecular Biology*. 2002;48(5-6):633-47. doi:10.1023/a:1014837431178
5. Turnbull KM, Turner M, Mukai Y, Yamamoto M, Morell MK, Rahman S, et al. The organization of genes tightly linked to the *Ha* locus in *Aegilops tauschii*, the D-genome donor to wheat. *Genome*. 2003;46(2):330-338. doi:10.1139/g02-124
6. Gautier MF, Aleman ME, Guirao A, Marion D, Joudrier P. *Triticum aestivum* puroindolines, two basic cysteine-rich seed proteins: cDNA sequence analysis and developmental gene expression. *Plant Mol Biol*. 1994;25(1):43-57. doi:10.1007/BF00024197
7. Morris CF. The antimicrobial properties of the puroindolines, a review. *World Journal of Microbiology and Biotechnology*. 2019;35(6):86. doi:10.1007/s11274-019-2655-4
8. Elmorjani K, Geneix N, Dalgalarondo M, Branlard G, Marion D. Wheat grain softness protein (Gsp1) is a puroindoline-like protein that displays a specific post-translational maturation and does not interact with lipids. *Journal of Cereal Science*. 2013;58(1):117-22. doi:10.1016/j.jcs.2013.03.012
9. Morris CF, Geng H, Beecher BS, Ma D. A review of the occurrence of *Grain softness protein-1* genes in wheat (*Triticum aestivum* L.). *Plant Molecular Biology*. 2013;83(6):507-521. doi:10.1007/s11103-013-0110-8
10. Gautier MF, Cosson P, Guirao A, Alary R, Joudrier P. Puroindoline genes are highly conserved in diploid ancestor wheats and related species but absent in tetraploid *Triticum* species. *Plant Science*. 2000;153(1):81-91. doi:10.1016/S0168-9452(99)00258-7
11. Chantret N, Salse J, Sabot F, Rahman S, Bellec A, Laubin B, et al. Molecular basis of evolutionary events that shaped the hardness locus in diploid and polyploid wheat species (*Triticum* and *Aegilops*). *Plant Cell*. 2005;17(4):1033-45. doi:10.1105/tpc.104.029181
12. Tranquilli G, Lijavetzky D, Muzzi G, Dubcovsky J. Genetic and physical characterization of grain texture-related loci in diploid wheat. *Molecular and General Genetics*. 1999;262(4-5):846-50. doi:10.1007/s004380051149
13. Li W, Huang L, Gill BS. Recurrent deletions of puroindoline genes at the grain hardness locus in four independent lineages of polyploid wheat. *Plant Physiol*. 2008 Jan;146(1):200-12. doi:10.1104/pp.107.108852
14. Morris CF, Luna J, Caffè-Treml M. The Vromindolines of cv. Hayden oat (*Avena sativa* L.) – A review of the Poeae and Triticeae indolines and a suggested system for harmonization of nomenclature. *Journal of Cereal Science*. 2021;97:103135. doi:10.1016/j.jcs.2020.103135
15. Charles M, Tang H, Belcram H, Paterson A, Gornicki P, Chalhou B. Sixty million years in evolution of soft grain trait in grasses: Emergence of the softness locus in the common ancestor of poeidae and Ehrhartoideae, after their divergence from Panicoideae. *Molecular Biology and Evolution*. 2009;26(7):1651-61. doi:10.1093/molbev/msp076
16. Wilkinson MD, King R, Grimaldi R. Sequence diversity and identification of novel puroindoline and grain softness protein alleles in *Elymus*, *Agropyron* and related species. *Diversity*. 2018;10(4):114. doi:10.3390/d10040114
17. Feldman M, Levy AA. *Triticum* L. In: *Wheat Evolution and Domestication*. Springer, Cham; 2023. doi:10.1007/978-3-031-30175-9_10
18. Taddei F, Gazza L, Conti S, Muccilli V, Foti S, Pogna NE. Starch-bound 2S proteins and kernel texture in einkorn, *Triticum monococcum* ssp. *monococcum*. *Theoretical and Applied Genetics*. 2009;119(7):1205-12. doi:10.1007/s00122-009-1121-3
19. Kulathunga J, Reuhs BL, Zwinger S, Simsek S. Comparative study on kernel quality and chemical composition of ancient and modern wheat species: Einkorn, emmer, spelt and hard red spring wheat. *Foods*. 2021;10(4):761. doi:10.3390/foods10040761
20. Michikawa A, Okada M, Ikeda TM, Nagaki K, Yoshida K, Takumi S. Phenotypic effects of A^m genomes in nascent synthetic hexaploids derived from interspecific crosses between durum and wild einkorn wheat. *PLoS ONE*. 2023;18(4):e0284408. doi:10.1371/journal.pone.0284408
21. Massa AN, Morris CF. Molecular evolution of the puroindoline-a, puroindoline-b, and grain softness protein-1 genes in the tribe Triticeae. *J Mol Evol*. 2006;63(4):526-536. doi:10.1007/s00239-005-0292-z
22. Chen Q, Qi PF, Wei YM, Wang JR, Zheng YL. Molecular characterization of the pina gene in einkorn wheat. *Biochem Genet*. 2009;47(5-6):384-96. doi:10.1007/s10528-009-9239-1
23. Guzmán C, Caballero L, Martín MA, Alvarez JB. Molecular characterization and diversity of the *Pina* and *Pinb* genes in cultivated and wild diploid wheat. *Mol Breeding*. 2012;30:69-78. doi:10.1007/s11032-011-9599-1
24. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis version 11. *Molecular Biology and Evolution* 2021;38:3022-27. doi:10.1093/molbev/msab120
25. Zhang J. Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. *Journal of Molecular Evolution*. 2000;50(1):56-68. doi:10.1007/s002399910007
26. Kim KH, Feiz L, Martin JM, Giroux MJ. Puroindolines are associated with decreased polar lipid breakdown during wheat seed development. *Journal of Cereal Science*. 2012;56(2):142-46. doi:10.1016/j.jcs.2012.02.011
27. Lesage VS, Bouchet B, Rhazi L, Elmorjani K, Branlard G, Marion D. New insight into puroindoline function inferred from their subcellular localization in developing hard and soft near-isogenic endosperm and their relationship with polymer size of storage proteins. *Journal of Cereal Science*. 2011;53:231-38. doi:10.1016/j.jcs.2011.01.002
28. Quayson ET, Marti A, Morris CF, Marengo M, Bonomi F, Seetharaman K, et al. Structural consequences of the interaction of puroindolines with gluten proteins. *Food Chemistry*. 2018;253:255-61. doi:10.1016/j.foodchem.2018.01.146
29. Geneix N, Dalgalarondo M, Tassy C, Nadaud I, Barret P, Bakan B, et al. Relationships between puroindoline A-prolamins interactions and wheat grain hardness. *PLoS ONE*. 2020;15(9):e0225293. doi:10.1371/journal.pone.0225293

O. Sozinova^{1,2}, Ya. Blume²

¹ Institute of Plant Protection NAAS, Ukraine

² Institute of Food Biotechnology and Genomics of the National Academy of Sciences of Ukraine, Ukraine

ANALYSIS OF SINGLE NUCLEOTIDE POLYMORPHISMS IN SEQUENCES OF *Pina* AND *Pinb* GENES OF THE DIPLOID WHEATS *Triticum monococcum* AND *T. urartu*

Abstract

Puroindolines (puroindoline a and puroindoline b) are low molecular weight proteins that determine the endosperm texture of grain in the tribes Triticeae and Avenae. The aim of our study was to analyse single nucleotide polymorphisms (SNPs) of the coding sequences of the *Pina* and *Pinb* genes of the diploid wheats *Triticum monococcum* ($A^m A^m$) and *T. urartu* (AA) from the NCBI database in comparison with the reference sequences of the common wheat variety Chinese Spring (CS).

For this, 62 sequences of the *T. monococcum* puroindoline a gene (*Pina-A^m1*), 22 sequences of *T. urartu* *Pina* (*Pina-A1*), 32 sequences of the *T. monococcum* puroindoline b gene (*Pinb-A^m1*) and 13 sequences of *T. urartu* *Pinb* (*Pinb-A1*) were retrieved from the NCBI database. The sequences of the puroindoline a gene DQ363911.1 of the variety CS (the *Pina-D1a* allele) and the puroindoline b gene DQ363913.1 of CS (the *Pinb-D1a* allele) from the NCBI database were used as the reference sequences. The sequences were aligned using MEGA 11. In total, 34 SNPs (13 synonymous and 21 nonsynonymous differences, of which 15 result in radical amino acid substitutions and 6 in conservative ones) were identified in the total sample of 84 diploid wheat *Pina* sequences, some of which were observed in all sequences and some were rare. Among the 45 *Pinb* sequences, there were 36 SNPs, but, unlike the *Pina* gene, synonymous substitutions prevailed (22); 7 substitutions led to radical amino acid substitutions and 7 to conservative ones. Substitutions in the sequences of the puroindoline genes relative to the CS genes can be divided into those fixed in both diploid wheat species, those fixed in *T. urartu* and polymorphic in *T. monococcum*, and species-specific ones. Significant differences in the frequencies of alternative nucleotides at certain positions (81, 318, 322 and 384 of *Pina* and 135 and 359 of *Pinb*) were found between the wild einkorn wheat *T. monococcum* ssp. *aegilopoides* and the cultivated wheat *T. monococcum* ssp. *monococcum*.

Keywords: puroindoline, *Triticum monococcum* ssp. *aegilopoides*, *T. monococcum* ssp. *monococcum*, *T. urartu*, SNPs, radical amino acid substitutions, conservative amino acid substitutions.

Матеріал надійшов 03.06.2024

Відомості про авторів Notes about authors

Созінова Оксана Ігорівна – Інститут захисту рослин НААН, Державна установа «Інститут харчової біотехнології і геноміки НАН України», Україна

Sozinova Oksana – Institute of Plant Protection NAAS, Institute of Food Biotechnology and Genomics of the National Academy of Sciences of Ukraine, Ukraine

<https://orcid.org/0000-0002-0981-3433>

e-mail: sozinovaoksana1@gmail.com

Блюм Ярослав Борисович – академік НАН України, професор, доктор біологічних наук, Державна установа «Інститут харчової біотехнології і геноміки НАН України», Україна

Blume Yaroslav – Academician of the National Academy of Sciences of Ukraine, Professor, Doctor of Biological Sciences, Institute of Food Biotechnology and Genomics of the National Academy of Sciences of Ukraine, Ukraine

<https://orcid.org/0000-0001-7078-7548>

e-mail: cellbio@cellbio.freenet.viaduk.net

