

DOI: 10.18523/2617-4529.2025.8.55-61

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 та геноміки НАН України», Україна

ОДНОНУКЛЕОТИДНІ ПОЛІМОРФІЗМИ В ПОСЛІДОВНОСТЯХ ГЕНА *PINA* ДЕЯКИХ ДИПЛОЇДНИХ ВИДІВ РОДУ *AEGILOPS*

Види роду *Aegilops* L. є генетичним ресурсом для перенесення нових генів, зокрема нових алелів генів пуроіндолінів, у пшеницю м'яку. Пуроіндоліни *a* і *b* – низькомолекулярні білки, які визначають текстуру ендосперму зерна у *Triticum aestivum* та споріднених видів. Мета роботи – аналіз частот трапляння однонуклеотидних поліморфізмів (SNP) кодуючих послідовностей гена *Pina* диплоїдних видів егілопсів порівняно з референсною послідовністю сорту *T. aestivum* Chinese Spring серед послідовностей, представлених у базі NCBI.

З бази даних NCBI було відібрано послідовності гена *Pina* диплоїдних видів егілопсів: 32 послідовності *Ae. speltoides*, 8 послідовностей *Ae. bicornis*, 5 послідовностей *Ae. sharonensis*, 6 послідовностей *Ae. searsii*, 8 послідовностей *Ae. caudata*, 10 послідовностей *Ae. comosa* та 14 послідовностей *Ae. umbellulata*. Як референсну послідовність використовували послідовність гена пуроіндоліну *a* (алель *Pina-D1a*) DQ363911.1 сорту CS. Послідовності вирівнювали за допомогою програми MEGA 11.

Серед проаналізованих послідовностей гена *Pina* *Ae. speltoides*, *Ae. bicornis*, *Ae. sharonensis*, *Ae. searsii*, *Ae. caudata*, *Ae. comosa*, *Ae. umbellulata* сумарно виявлено SNP у 61 позиції кодуючої послідовності. У різних видів було від 11 до 30 SNP. Види егілопсів охарактеризовано за частотами трапляння нуклеотидних замін. У більшості видів егілопсів переважають несинонімічні заміни. SNP у двох позиціях трапляються у всіх досліджених диплоїдних видів егілопсів, а сім SNP є притаманними лише видам секції *Sitopsis*. Серед усіх проаналізованих видів егілопсів тільки *Ae. searsii* має SNP, що є унікальними для виду і зафіксовані у всіх представлених у базі даних послідовностях.

Ключові слова: пуроіндолін, SNP, *Aegilops*, *Pina*, радикальні амінокислотні заміни, консервативні амінокислотні заміни, синонімічні заміни.

Вступ

Види роду *Aegilops* L. є цінним генетичним ресурсом для перенесення нових генів, що визначають якість зерна та стійкість до абіотичних і біотичних факторів, у пшеницю *Triticum aestivum* L. (геномна формула AABBDD, $2n = 42$) [1-3]. За класифікацією Ван Слагерена (Van Slageren) [4], виділяють 22 види егілопсів. За деякими іншими таксономічними системами егілопси зараховують до роду *Triticum* [5]. Серед них 10 є диплоїдними видами ($2n = 14$). Це *Ae. tauschii* (DD) (синонім *Ae. squarrosa*), *Ae. speltoides* (SS), *Ae. bicornis* (S^bS^b), *Ae. longissima* (S^iS^i), *Ae. sharonensis* ($S^{sh}S^{sh}$),

Ae. searsii (S^sS^s), *Ae. caudata* (CC) (синонім *Ae. markgrafii*, *T. dichasians*), *Ae. comosa* (MM), *Ae. uniaristata* (NN), *Ae. umbellulata* (UU). *Ae. speltoides*, *Ae. bicornis*, *Ae. longissima*, *Ae. sharonensis*, *Ae. searsii* належать до секції *Sitopsis*, *Ae. caudata* – до секції *Cylindropyrum*, *Ae. comosa* та *Ae. uniaristata* – до секції *Comopyrum*, *Ae. umbellulata* – до секції *Aegilops*. *Ae. tauschii*, який є донором геному D пшениці м'якої, належить до секції *Vertebrata* [4,6]. Саме *Ae. tauschii* завдячує *T. aestivum* такою ознакою, як м'якозерність, рівень якої визначається алелями пуроіндолінових генів [7]. Локус *Ha*, що містить тісно зчеплені гени *Pina-D1* та

Pinb-D1, наявний на короткому плечі хромосоми 5D і відсутній на гомеологічних хромосомах 5A і 5B *T. aestivum* та *T. turgidum* [7,8].

Пуроіндоліни (пуроіндолін а та пуроіндолін б) – низькомолекулярні цистеїн- і триптофан-багаті білки зерна з надродини проламінів; пуроіндолін а складається зі 120 амінокислотних залишків [9]. У складі незрілого білкового продукту пуроіндолінових генів є сигнальний пептид із 19 амінокислотних залишків та невеликі N- та C-кінцеві фрагменти, які відщеплюються з формуванням зрілого білка. У позиції 66–73 у пуроіндоліну а (відносно початку незрілого білка) є триптофан-багатий домен, який ще називають триптофановим мотивом, WRWWKWWK [9]. Триптофан-багатий домен відповідає за антимікробні властивості пуроіндолінів [10]. Гени пуроіндолінів не містять інтронів [9] подібно до генів проламінів (гліадинів та субодинаць глютенінів), та, на відміну від власне проламінів та інших білків надродини проламінів, вони є лише в одичинній копії [11]. Алейний склад пуроіндолінів визначає текстуру зерна, яка, своєю чергою, впливає на помел зерна пшениці та водопоглинальні властивості борошна. Більшу кількість енергії, необхідну для помелу твердозерних зразків зерна пшениці порівняно з м'якозерними, пов'язують із сильнішою адгезією між поверхнями крохмальних гранул і білковим матриксом ендосперму зернівки [7,11,12].

Джерелом різноманітності генів пуроіндолінів для збагачення генофонду пшениці можуть бути різні види егілопсів [1]. Відповідні білки егілопсів також називають пуроіндолінами [13]. Послідовності пуроіндолінових генів різних видів егілопсів представлені в базі даних NCBI (National Center for Biotechnology Information) у результаті досліджень різних авторів [14–20]. Метою нашої роботи був аналіз частот трапляння одонуклеотидних поліморфізмів (SNP) кодуючих послідовностей гена *Pina* диплоїдних видів егілопсів секцій *Sitopsis*, *Cylindropyrum*, *Comorugum* та *Aegilops* порівняно з референсною послідовністю сорту *T. aestivum* Chinese Spring серед послідовностей, представлених у базі NCBI.

Матеріали і методи

З бази даних NCBI було відібрано послідовності гена *Pina* диплоїдних видів егілопсів *Ae. speltoides*, *Ae. bicornis*, *Ae. sharonensis*, *Ae. searsii*, *Ae. caudata*, *Ae. comosa* та *Ae. umbellulata*.

Послідовності *Ae. speltoides*: AY622788.2, AY622789.2, AY622787.2, JX648382.1, JX648381.1, JX648380.1, JX648379.1, JX648378.1, EU268474.1,

FJ898230.1, FJ898229.1, FJ898228.1, FJ898227.1, FJ898226.1, FJ898225.1, FJ898224.1, FJ898223.1, FJ898222.1, FJ898221.1, FJ898220.1, FJ898219.1, EU307588.1, EU307587.1, AJ302097.1, AJ302096.1, DQ269835.1, DQ269834.1, DQ269833.1, DQ269832.1, DQ269831.1, DQ269830.1, DQ269829.1.

Послідовності *Ae. bicornis*: AY622795.2, AY622794.2, FJ898215.1, FJ898207.1, AY608587.1, DQ269844.1, DQ269843.1, DQ269842.1.

Послідовності *Ae. sharonensis*: AY622796.2, EU268476.1, FJ898217.1, FJ898216.1, DQ269841.1.

Послідовності *Ae. searsii*: AY622793.2, AY622792.2, JX648377.1, JX648376.1, DQ269840.1, DQ269839.1.

Послідовності *Ae. caudata*: JX648375.1, JX648374.1, JX648373.1, JX648372.1, JX648371.1, AY608594.2, AY608586.1, DQ269848.1.

Послідовності *Ae. comosa*: JX648370.1, JX648369.1, JX648368.1, JX648367.1, FJ898210.1, FJ898211.1, FJ898209.1, FJ898206.1, DQ269846.1, DQ269845.1.

Послідовності *Ae. umbellulata*: JX648389.1, JX648388.1, JX648387.1, JX648386.1, JX648385.1, JX648384.1, JX648383.1, LC375780.1, LC375779.1, LC375778.1, LC375777.1, LC375776.1, LC375775.1, DQ269847.1.

Референсною послідовністю слугувала послідовність гена пуроіндоліну а (алеель *Pina-D1a*) DQ363911.1 сорту *T. aestivum* Chinese Spring (CS) з бази даних NCBI. Послідовності вирівнювали за допомогою програми MEGA 11 [21]. Аналіз відмінностей від референсної послідовності (SNP або заміни) проводили лише для кодуючих послідовностей. Заміни в нуклеотидних і амінокислотних послідовностях характеризували згідно з класифікацією, наведеною в [22], як синонімічні та несинонімічні. Несинонімічні заміни, своєю чергою, поділяли на радикальні та консервативні. Істотність відмінностей частот аналізували за допомогою критерію χ^2 .

Результати та обговорення

Проведено аналіз відмінностей від референсної послідовності гена пуроіндоліну а сорту пшениці CS у послідовностях гена *Pina* диплоїдних видів егілопсів *Ae. speltoides* (32 послідовності), *Ae. bicornis* (8), *Ae. sharonensis* (5), *Ae. searsii* (6), *Ae. caudata* (8), *Ae. comosa* (10), *Ae. umbellulata* (14). SNP та частоти їх трапляння серед вибірки послідовностей певного виду показано в таблиці 1. SNP траплялись у 11–30 позиціях, залежно від виду, найбільше їх було у *Ae. speltoides* (30) і найменше у *Ae. comosa* (11). У більшості видів егілопсів серед SNP переважну частку становили несинонімічні заміни

(табл. 2). Сумарно для всіх видів це співвідношення становить 75 : 46, що істотно відрізняється від 1 : 1 ($\chi^2 = 6,95$, $P < 0,01$).

У всіх проаналізованих видів егілопсів з високою частотою (77,8–100 %) траплялась синонімічна заміна в позиції 24 (A→C), лише у *Ae. umbellulata* її частота була 43,9 %. Таку заміну також мають усі послідовності *Pina* диплоїдних пшениць *T. monococtum* (A^mA^m) і *T. urartu* (AA) [23]. Ще однією заміною, що траплялась у всіх досліджених видів егілопсів, була заміна в позиції 268 (A→C), яка призводила до заміни метіоніну на лейцин в амінокислотній позиції 90 поблизу триптофан-багатого домена. Частота такої заміни була від 87,5 до 100 % у видів секції Sitopsis, а також 100 % у *Ae. comosa*, проте була значно нижчою у видів *Ae. caudata* та *Ae. umbellulata*. Низка SNP були притаманними лише егілопсам секції Sitopsis: у позиціях 154 (C→T), 209 (A→G), 415 (A→G), 422 (C→G), 424 (G→A), 431 (T→G) та 443 (G→A). Заміни в позиціях 154 та 443 синонімічні, а решта призводять до консервативних (SNP 209, 424) і радикальних (SNP 415, 422, 431) амінокислотних змін.

Заміна в позиції 121 (G→C), що призводить до консервативної заміни валіну на лейцин, є характерною для послідовностей *Ae. speltooides*, *Ae. bicornis*, *Ae. sharonensis* (87,5–100 %) і трапляється у 30 % послідовностей *Ae. comosa*, є рідкісною у *Ae. umbellulata*, але її не виявлено у *Ae. searsii* та *Ae. caudata*. Аналогічну заміну мають також *T. monococtum* і *T. urartu* [23]. У цій самій позиції 50 % послідовностей *Ae. umbellulata* мають іншу заміну – G→A, яка призводить до заміни валіну на ізолейцин. Така сама заміна з низькою частотою наявна серед послідовностей *Ae. speltooides*. Крім згаданих вище SNP, для *Ae. speltooides* також є характерними синонімічні заміни в позиціях 230 і 270. У всіх послідовностях *Ae. sharonensis*, на відміну від інших видів секції Sitopsis, є SNP в позиції 89 (T→C) (результатом є заміна валіну на аланін). Заміна в позиції 139 (T→G) є характерною для *Ae. sharonensis* та *Ae. bicornis*. Усі послідовності *Ae. searsii* мають SNP у позиціях 7, 162, 186 на відміну від інших видів егілопсів.

Серед 18 SNP, які трапляються в послідовностях *Pina* *Ae. umbellulata*, високі частоти понад 75 % мають лише три: в позиції 230 (синонімічна заміна), 257 (G→A), призводить до радикальної заміни амінокислот, аргінін→глутамін, та 392 (G→A, заміна аргініну на лізин). Серед п'яти SNP, притаманних лише *Ae. caudata*, з високою частотою (75,0 %) представлені лише дві: в позиції 136 (A→C) і 273 (A→G). Перша заміна є радикальною (аспарагін на гістидин), а друга синонімічною.

Серед усіх проаналізованих видів егілопсів лише *Ae. searsii* має SNP, що є унікальними для виду і зафіксовані у всіх представлених у базі даних послідовностях. Це синонімічні заміни в позиціях 162 і 186, які можна використати для розробки специфічних праймерів для ідентифікації цього виду серед інших видів секції Sitopsis.

Зразки егілопсів можуть бути джерелом нових алелів для створення генотипів пшениці з новими властивостями щодо рівня м'язокзерності та антимікробної дії. Очевидно, що найбільш значні відмінності спостерігатимуться для алельних варіантів із замінами в триптофановому домені або поблизу (амінокислотні позиції 66–73 у незрілому білку пуроіндоліну α). Відмінності в послідовності ділянки, що відповідає триптофановому домену, мають усі види секції Sitopsis у нуклеотидній позиції 209 (A→G), що відповідає амінокислотній позиції 70 і призводить до заміни лізину на аргінін. Практично всі послідовності видів секції Sitopsis, *Ae. comosa* та частина послідовностей *Ae. umbellulata* та *Ae. caudata* мають заміну в позиції 268, яка реалізується в заміні метіоніну на лейцин в амінокислотній позиції 90. Ще одним потенційно важливим SNP є SNP у позиції 257, що призводить до радикальної заміни аргініну на глутамін у позиції 86. Такий SNP є в більшості послідовностей *Ae. umbellulata* та в одиничних послідовностях *Ae. speltooides*, *Ae. bicornis* та *Ae. caudata*. Така сама заміна в позиції 257 наявна у всіх послідовностях диплоїдних пшениць *T. monococtum* і *T. urartu*, які мають екстрем'язокзерну текстуру зерна [23]. Загалом, егілопсам притаманна м'яка структура зерна. Зокрема, Чен та ін. (Chen et al.) [16] аналізували рівень твердозерності зразків *Ae. longissima*, *Ae. sharonensis*, *Ae. bicornis*, а також поліплоїди *Ae. kotschyi* (UUSS), *Ae. triuncialis* (UUCS), *Ae. juvenalis* (DDMMUU) і виявили м'язокзерну текстуру зерна у всіх зразків, крім одного зразка *Ae. sharonensis*. Твердозерні зразки *Ae. umbellulata* також виявили Окада та ін. (Okada et al.) [20,24], що вказує на широку різноманітність пуроіндолінових алелів егілопсів щодо рівня м'язокзерності.

Висновки

Серед проаналізованих послідовностей егілопсів *Ae. speltooides*, *Ae. bicornis*, *Ae. sharonensis*, *Ae. searsii*, *Ae. caudata*, *Ae. comosa*, *Ae. umbellulata* сумарно виявлено SNP у 61 позиції кодуючої послідовності гена *Pina*. У більшості видів переважають несинонімічні заміни. Два SNP трапляються у всіх досліджених видів егілопсів, а сім SNP є притаманними лише видам секції Sitopsis. Унікальні видоспецифічні SNP виявлено лише для *Ae. searsii*.

Роботу виконано в межах НДР 0125U001978.

Відмінності в послідовності гена *Pina* диплоїдних видів егілопсів від референсної послідовності

Позиція в гені	CS	SNP	Позиція в АК послідовності	АК CS	Заміна АК	Характер заміни*	<i>Ae. speltoides</i>
1	A	T	1	M	L	C	
7	G	A	3	A	T	R	0,031
10	C	T	4	L	F	C	
23	G	C				S	0,250
24	A	C				S	0,969
27	G	C				S	
33	T	A				S	
36	G	C				S	
41	C	T	14	A	V	C	
57	G	A				S	
66	C	A	22	S	R	R	
66	C	T				S	
67	G	A	23	E	K	R	0,063
73	G	C				S	
78	C	T				S	0,125
82	T	A	28	Y	N	R	
89	T	C	30	V	A	C	0,031
96	C	T				S	0,031
97	G	A	33	G	R	R	0,031
100	G	T	34	G	C	R	
104	G	C	35	G	A	C	
105	T	A	36			S	0,031
106	G	C	36	A	P	C	0,313
121	G	C	41	V	L	C	0,969
121	G	A	41	V	I	C	0,031
130	A	G	44	K	E	R	
136	A	C	46	N	H	R	
138	T	A	46			S	
139	T	G	47	S	A	R	0,031
140	C	A	47	S	A	R	0,031
141	T	G	47	S	L	R	0,031
154	C	T				S	0,969
162	T	C				S	
186	C	T				S	
189	G	A				S	
201	T	C				S	
209	A	G	70	K	R	C	0,969
230	A	T				S	0,969
231	A	G				S	
249	T	C				S	
256	C	T				S	
257	G	A	86	R	Q	R	0,031
268	A	C	90	M	L	C	0,969
270	G	A				S	0,938
273	A	G				S	
314	G	T	105	G	V	C	
316	G	A	106	D	N	R	0,031
320	T	A	107	L	H	R	
333	C	T				S	
351	G	A				S	
382	C	A	128	L	M	C	
390	C	T				S	0,031
392	G	A	131	R	K	C	0,031
399	C	T				S	
400	C	G	134	Q	E	R	
415	A	G	139	N	D	R	0,969
422	C	G	141	P	R	R	0,969
424	G	A	142	G	S	C	0,969
429	T	C				S	0,125
431	T	G	144	I	S	R	0,969
443	G	A				S	0,844

Таблиця 1

сорту Chinese Spring (CS) та відмінності в амінокислотній (АК) послідовності

<i>Ae. sharonensis</i>	<i>Ae. searsii</i>	<i>Ae. bicornis</i>	<i>Ae. umbellulata</i>	<i>Ae. comosa</i>	<i>Ae. caudata</i>
			0,143		
	1,000				0,500
	0,333				
1,000	1,000	0,875	0,429	1,000	0,875
				0,300	
				0,300	
			0,143	0,600	0,125
			0,143		
		0,125			
				0,600	0,125
					0,250
0,200					
					0,125
1,000			0,143	1,000	0,125
			0,286		
	0,333				
1,000		0,875	0,071	0,300	
			0,500		
		0,25			
					0,750
0,400					
1,000		0,875			
			0,286		
			0,500		
1,000	1,000	0,875			
	1,000				
	1,000				
0,400					
				0,300	
1,000	1,000	0,875			
			0,786		
	0,333				
0,200					
			0,286		
		0,125	0,857		0,125
1,000	1,000	0,875	0,143	1,000	0,125
					0,750
			0,071		
0,400					
		0,25	0,071		
0,200					
0,200					
				0,100	0,125
			0,857		
			0,071	0,100	
1,000	1,000	0,875			
1,000	1,000	0,875			
1,000	1,000	0,875			
0,200					
1,000	1,000	0,875			
1,000	0,833	0,875			

Таблиця 2

**Кількість синонімічних (S), консервативних (C)
і радикальних (R) замін у кодуючій послідовності гена *Pina* диплоїдних видів егілопсів**

Вид	Кількість замін			
	C	R	S	Разом
<i>Ae. speltooides</i>	8	11	11	30
<i>Ae. sharonensis</i>	5	5	10	20
<i>Ae. searsii</i>	5	4	6	15
<i>Ae. bicornis</i>	4	6	5	15
<i>Ae. umbellulata</i>	8	6	4	18
<i>Ae. comosa</i>	4	1	6	11
<i>Ae. caudata</i>	3	5	4	12

References

- Kilian B, Mammen K, Millet E, Sharma R, Graner A, Salamini F, Hammer K, Özkan H. *Aegilops*. In: Kole C, editor. Wild Crop Relatives: Genomic and Breeding Resources. Berlin Heidelberg: Springer-Verlag; 2011. p. 1-76. doi:10.1007/978-3-642-14228-4
- Ceoloni C, Kuzmanović L, Ruggeri R, Rossini F, Forte P, Cuccurullo A, et al. Harnessing genetic diversity of wild gene pools to enhance wheat crop production and sustainability: Challenges and opportunities. *Diversity*. 2017 Dec 1;9(4):55. doi:10.3390/d9040055
- Kishii M. An update of recent use of *Aegilops* species in wheat breeding. *Front Plant Sci*. 2019 May 9;10. doi:10.3389/fpls.2019.00585
- Van Slageren MW. Wild wheats: a monograph of *Aegilops* L. and *Amblyopyrum* (Jaub. & Spach) Eig (Poaceae). Wageningen Agricultural University Papers; 1994. 512 p.
- Kimber G, Feldman M. Wild Wheat, an Introduction. Columbia: College of Agriculture University of Missouri Special Report 353; 1987. 146 p.
- Feldman M, Levy AA. Wheat Evolution and Domestication. Cham: Springer International Publishing; 2023. doi:10.1007/978-3-031-30175-9
- Pauly A, Pareyt B, Fierens E, Delcour JA. Wheat (*Triticum aestivum* L. and *T. turgidum* L. ssp. *durum*) kernel hardness: II. Implications for end-product quality and role of puroindolines therein. *Comp Rev Food Sci Food Saf*. 2013 Jun 12;12(4):427-38. doi:10.1111/1541-4337.12018
- Turnbull KM, Turner M, Mukai Y, Yamamoto M, Morell MK, Appels R, Rahman S. The organization of genes tightly linked to the *Ha* locus in *Aegilops tauschii*, the D-genome donor to wheat. *Genome*. 2003 Apr 1;46(2):330-8. doi:10.1139/g02-124
- 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 Apr;25(1):43-57. doi:10.1007/bf00024197
- Morris CF. The antimicrobial properties of the puroindolines, a review. *J Microbiol Biotechnol*. 2019 May 27;35(6). doi:10.1007/s11274-019-2655-4
- Morris CF. Puroindolines: the molecular genetic basis of wheat grain hardness. *Plant Mol Biol*. 2002 Jan 1;48(5/6):633-47. doi:10.1023/a:1014837431178
- Shewry P (2022) Wheat grain proteins: past, present and future. *Cereal Chem*. 100(1):9-22. doi:10.1002/cche.10585
- Morris CF, Luna J, Caffè-Treml M. The vromindolines of cv. *Hayden* oat (*Avena sativa* L.) – A review of the Poaceae and Triticeae indolines and a suggested system for harmonization of nomenclature. *J Sci*. 2021 Jan;97:103135. doi:10.1016/j.jcs.2020.103135
- 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*. 2007 Nov 16;146(1):200-12. doi:10.1104/pp.107.108852
- Lillemo M, Cosimo Simeone M, Morris CF. Analysis of puroindoline a and b sequences from *Triticum aestivum* cv. 'Penawawa' and related diploid taxa. *Euphytica*. 2002; 126(3):321-31. doi:10.1023/A:1019908325078
- Chen M, Wilkinson M, Tosi P, He G, Shewry P. Novel puroindoline and grain softness protein alleles in *Aegilops* species with the C, D, S, M and U genomes. *Theor Appl Genet*. 2005 Aug 17;111(6):1159-66. doi:10.1007/s00122-005-0047-7
- 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 Jul 28;63(4):526-36. doi:10.1007/s00239-005-0292-z
- 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*. 2007 Nov 16;146(1):200-12. doi:10.1104/pp.107.108852
- Cuesta S, Guzmán C, Alvarez JB. Allelic diversity and molecular characterization of puroindoline genes in five diploid species of the *Aegilops* genus. *J Exp Bot*. 2013 Sep 21; 64(16):5133-43. doi:10.1093/jxb/ert299
- Okada M, Ikeda TM, Yoshida K, Takumi S. Effect of the U genome on grain hardness in nascent synthetic hexaploids derived from interspecific hybrids between durum wheat and *Aegilops umbellulata*. *J Cereal Sci*. 2018 Sep;83:153-61. doi:10.1016/j.jcs.2018.08.011
- Tamura K, Stecher G, Kumar S. MEGA11: Molecular evolutionary genetics analysis Version 11. *Mol Biol Evol*. 2021 Apr 23;38(7):3022-7. doi:10.1093/molbev/msab120
- Zhang J. Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. *J Mol Evol*. 2000 Jan;50(1):56-68. doi:10.1007/s002399910007
- 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 Breed*. 2011 Jun 18;30(1):69-78. doi:10.1007/s11032-011-9599-1
- Okada M, Michikawa A, Yoshida K, Nagaki K, Ikeda TM, Takumi S. Phenotypic effects of the U-genome variation in nascent synthetic hexaploids derived from interspecific crosses between durum wheat and its diploid relative *Aegilops umbellulata*. *PLoS ONE*. 2020 Apr 2;15(4):e0231129. doi:10.1371/journal.pone.0231129

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

SINGLE NUCLEOTIDE POLYMORPHISMS IN *Pina* GENE SEQUENCES OF SOME DIPLOID SPECIES OF THE GENUS *Aegilops*

Abstract

Species of the genus *Aegilops* L. are a genetic resource for transferring new genes, in particular new alleles of puroindoline genes, into common wheat. Puroindolines a and b are low molecular weight proteins that determine the grain endosperm texture in *Triticum aestivum* and related species. The aim of the study was to analyze frequencies of single nucleotide polymorphisms (SNPs) in the coding sequences of the *Pina* gene in diploid *Aegilops* species in comparison with the reference sequence of the *T. aestivum* variety Chinese Spring (CS) among sequences available in the NCBI database.

Sequences of the *Pina* gene of diploid *Aegilops* species were selected from the NCBI database, comprising 32 sequences of *Ae. speltoides*, 8 sequences of *Ae. bicornis*, 5 sequences of *Ae. sharonensis*, 6 sequences of *Ae. searsii*, 8 sequences of *Ae. caudata*, 10 sequences of *Ae. comosa*, and 14 sequences of *Ae. umbellulata*. The sequence of the puroindoline a gene (the *Pina-D1a* allele) DQ363911.1 of CS was used as a reference. The sequences were aligned using the MEGA11 software.

Among the sequences of the *Pina* gene of *Ae. speltoides*, *Ae. bicornis*, *Ae. sharonensis*, *Ae. searsii*, *Ae. caudata*, *Ae. comosa*, and *Ae. umbellulata*, a total of 61 SNPs were detected in the coding sequence. Different species showed between 11 and 30 SNPs. The *Aegilops* species were characterized with respect to the frequencies of nucleotide substitutions. In the majority of *Aegilops* species, nonsynonymous substitutions predominate. SNPs at two positions occur in all the diploid *Aegilops* species, and seven SNPs are characteristic of only the species from the section Sitopsis. Among the *Aegilops* species, only in *Ae. searsii* there are species-specific SNPs that are fixed in all its sequences presented in the database.

Keywords: puroindoline, SNP, *Aegilops*, *Pina*, radical amino acid substitutions, conserved amino acid substitutions, synonymous substitutions.

ARTICLE HISTORY. Submitted 9 May 2025. Accepted 26 May 2025. Published 18 August 2025

Відомості про авторів Authors Information

Созінова Оксана Ігорівна – молодший науковий співробітник Інституту захисту рослин НААН, Київ, Україна; аспірантка Державної установи «Інститут харчової біотехнології та геноміки НАН України», Київ, Україна

Sozinova Oksana – Junior Researcher of Institute of Plant Protection NAAS, Kyiv, Ukraine; PhD student of Institute of Food Biotechnology and Genomics of the National Academy of Sciences of Ukraine, Kyiv, Ukraine

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

sozinovaoksana1@gmail.com

Блюм Ярослав Борисович – академік НАН України, професор, доктор біологічних наук, директор Державної установи «Інститут харчової біотехнології та геноміки НАН України», Київ, Україна

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

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

cellbio@cellbio.freenet.viaduk.net

