

## ІДЕНТИФІКАЦІЯ ТА БІОІНФОРМАТИЧНИЙ АНАЛІЗ ГЕНІВ *sHSP* *ORYZA SATIVA*

Л.В. КОЗУБ, І.І. ПАНЧУК

Кафедра молекулярної генетики та біотехнології,  
Чернівецький національний університет імені Юрія Федьковича  
вул. Коцюбинського, 2, м. Чернівці, Україна, 58012  
e-mail: [i.panchuk@chnu.edu.ua](mailto:i.panchuk@chnu.edu.ua)

Рослини введуть прикріпленій спосіб життя та постійно зазнають впливу негативних факторів навколишнього середовища, таких як посуха, підвищення температури та висока засоленість ґрунтів, що значно знижують їхню врожайність. Для протидії цим умовам рослини активують захисні механізми, одним із ключових елементів яких є низькомолекулярні білки теплового стресу (*sHSP*). Ці білки діють як молекулярні шаперони, захищаючи клітини від пошкодження за стресових умов. Незважаючи на їх важливу роль, механізми дії *sHSP* у рису (*Oryza sativa*), однієї з основних продовольчих культур світу, досі вивчені недостатньо.

У дослідженні застосовано біоінформатичні методи для ідентифікації, класифікації та аналізу *sHSP* генів у геномі *O. sativa*. Амінокислотні послідовності вирівнювали методом *G-INS-I* на сервері *MAFFT*. Філогенетичний аналіз проводили методом *Maximum likelihood* за допомогою плагіна *PhyML* для *Geneious Prime 2023.2.1*. Статистична підтримка гілок була розрахована за методом *aLRT Chi2*. Екзон-інтронну структуру було згенеровано на *Gene Structure Display Server 2.0*.

В геномі *O. sativa* виявлено 35 генів *sHSP*, амінокислотні послідовності яких класифіковано на вісім структурних класів. Аналіз показав значну варіабельність N- та C-кінцевих ділянок *sHSP* при збереженні консервативності  $\alpha$ -кристалінового домену. Пластидні, мітохондріальні, ендоплазматичні та пероксисомальні білки формують окремі гілки на філогенетичному дереві, що свідчить про їх еволюційну дивергенцію. Також визначено білок, що не увійшов до відомих класів. Побудова екзон-інтронної структури підтвердила наявність інтронів у частини генів, що може свідчити про їхню роль у регуляції відповіді на тепловий стрес.

Отримані результати дозволяють краще зрозуміти роль *sHSP* у механізмах стійкості рослин до абіотичних стресів та є основою для подальших досліджень у напрямку створення високопродуктивних сортів рису.

**Ключові слова:** абіотичний стрес, біоінформатичний аналіз, генетичний поліморфізм, молекулярна геноміка, молекулярна еволюція, мультигенна родина,  $\alpha$ -кристаліновий домен, *sHSP*, *O. sativa*.

**Вступ.** Рослини постійно зазнають стресових впливів довкілля, таких як підвищення температури, посуха, висока концентрація іонів важких металів, висока засоленість ґрунту на оброблюваних і зрошуваних землях тощо, що призводить до пригнічення росту та розвитку, а також до зниження продуктивності сільськогосподарських культур (Yadav et al., 2020; Buzduga et al., 2022; Hassan et al., 2022).

Спільною рисою різних форм абіотичного стресу є збільшення продукції у клітині активних форм кисню (АФК), які викликають окислювальний стрес. Зокрема, АФК пошкоджують ліпіди мембран, білки та ДНК (Scharf et al., 1998; Rusnak et al., 2013; Ji et al., 2019). Це викликає індукцію відповіді на стрес, яка полягає у зміні експресії багатьох генів, активності захисних білків та концентрацій стрес-асоційованих метаболітів (Lee et al., 2000; Panchuk

et al., 2002; Pyrizhok et al., 2009; Doliba et al., 2011; Buzduga et al., 2014; Poudel et al., 2020).

Особливістю високотемпературного (теплого) стресу є денатурація білків, що призводить до порушення багатьох клітинних функцій (Volkov et al., 2005). Тому серед захисних білків теплового шоку (Heat Shock Proteins – HSP), які синтезуються у відповідь на зростання температури, головне місце займають чисельні білки-шаперони, які здатні протидіяти денатурації інших білків (Zhang et al., 2015; Waters & Vierling, 2020; Nagymasi et al., 2022).

Низькомолекулярні білки теплового шоку (small HSP - *sHSP*) мають меншу молекулярну масу, ніж інші HSP, зазвичай у діапазоні 12–42 кДа (Haslbeck & Vierling, 2015). Однією з їхніх відмінних рис є наявність консервативного  $\alpha$ -кристалінового домену, варіабельної N-кінцевої та короткої C-кінцевої ділянок. Ці білки відіграють важливу роль у стабілізації структури

інших білків під час стресу, зв'язуючись із частково денатурованими білками для утворення стабільних комплексів та запобігаючи їх необоротній агрегації АТФ-незалежним шляхом (Reddy et al., 2016; Hibshman et al., 2023).

У рослин sHSP представлені мультигенними/мультипротеїновими родинами, які нараховують десятки членів. Індивідуальні sHSP мають різну клітинну локалізацію, завдяки чому вони забезпечують захист від теплового стресу у різних органелах (Scharf et al., 2001; Waters, 2012; Khan et al., 2021).

Хоча захисна функція sHSP доведена для цілого ряду видів (Sun et al., 2012), уявлення про їх молекулярне різноманіття все ще залишаються фрагментарними навіть для найважливіших сільськогосподарських рослин. Зокрема, недостатньо охарактеризованими залишаються sHSP рису (*Oryza sativa*), хоча цей вид є однією з найважливіших продовольчих культур світового значення, яку вирощують переважно в тропічних і субтропічних зонах. Зависока температура негативно впливає на фертильність, розвиток пилку та врожайність зерна рису (Chandel et al., 2012; Chen et al., 2014; Wang et al., 2015).

*Oryza sativa* є однодольною диплоїдною рослиною ( $2n=24$ ), яка має маленький геном (Salse et al., 2008). Завдяки цьому рис став першим представником однодольних, геном якого було повністю розшифровано (Sakai et al., 2013). Тим не менш, різноманіття sHSP рису і досі залишається недослідженим. Тому метою даної роботи стала ідентифікація генів sHSP цієї рослини із застосуванням біоінформатичних підходів, аналіз їхньої структури та еволюції.

**Матеріали та методи.** Для пошуку послідовностей генів sHSP у базі даних GenBank ми використали алгоритм BLAST (Boratyn et al., 2013).

Амінокислотні послідовності вирівнювали методом G-INS-I на сервері MAFFT (Katoh et al., 2019). Філогенетичний аналіз проводили методом Maximum likelihood за допомогою плагіна PhyML для Geneious Prime 2023.2.1 (Guindon & Gascuel, 2003). Статистична підтримка гілок була розрахована за методом aLRT Chi2 (Anisimova & Gascuel, 2006). Розраховане дерево експортували у форматі Newick та візуалізували з використанням онлайн інструменту iTOL v6 - Interactive tree of life (Letunic & Bork, 2021). Екзон-інтронну структуру було згенеровано на Gene Structure Display Server 2.0 (Hu et al., 2014)

**Результати та їх обговорення.** У базі даних GenBank в геномі *O. sativa* було ідентифікованого 35 генів, які імовірно кодують sHSP. Для них характерна варіативність довжини амінокислотних послідовностей, що знаходиться

у діапазоні від 148 (OsHSP15.9-Per) до 251 (OsHSP27.5-P) АК, а передбачувана молекулярна маса коливається від 15.9 до 27.5 kDa (табл. 1).

Отримані послідовності геному *O. sativa* аналізували та порівнювали між собою, а також з раніше охарактеризованими генами sHSP (Scharf et al., 2001) дводольної модельної рослини - арабідопсису (*Arabidopsis thaliana*), яка належить до родини Brassicaceae (рис. 1). Аналіз показав, що порівнювані амінокислотні послідовності містять три основні частини: мінливу N-кінцеву область, консервативний  $\alpha$ -кристаліновий домен (ACD) та C-кінцеву ділянку, що є характерною особливістю для структурної будови sHSP.

Для того, щоб охарактеризувати молекулярну еволюцію sHSP, було побудовано дендрограму подібності на основі вирівнювання амінокислотних послідовностей. На ній з високою статистичною підтримкою виділяються декілька клад (рис. 2).

На наступному етапі аналізу було визначено приналежність генів/білків до певного структурного класу. Для цього ми звертали увагу на пептидні сигнали клітинного локалізації/транспорту, які присутні у N-кінцевій області та C-кінцевій ділянці білкових послідовностей, а також на розташування та збереження консервативних мотивів у ACD, додатково враховуючи анотацію цих послідовностей, які наводяться у базі даних GenBank. Також при аналізі враховували розташування послідовностей sHSP рису на дендрограмі відносно sHSP арабідопсису, для яких структурна класифікація була запропонована раніше (Scharf et al., 2001; Siddique et al., 2008).

Філогенетичний аналіз показав, що мультипротеїнову родину sHSP *O. sativa* можна розділити на 8 структурних класів (табл. 1). Білки трьох класів мають цитоплазматичну та/або ядерну локалізацію, тоді як інші розташовані у мітохондріях, пластидах, ендоплазматичному ретикулумі та пероксисомах. Крім sHSP, для яких визначено приналежність до певного класу, ідентифіковано білок, який не увійшов у жоден із цих класів.

Показано, що довжина N-кінцевої ділянки цитоплазматичних sHSP лежить у межах від 42 до 71 АК. Для пластидних sHSP цей показник становить від 80 до 141 АК, тоді як для мітохондріальних білків – від 113 до 148 залишків. Ендоплазматичні sHSP містять від 23 до 77 АК у N-кінцевій ділянці, пероксисомальний білок – лише 31 АК, а некласифікований OsHSP22.1 – 94 АК. C-кінцева ділянка є значно коротшою і містить від 15 до 32 АК у цитоплазматичних, 13-33 у пластидних, 14-15 у мітохондріальних, 34-61 ендоплазматичних та 18 у пероксисомального sHSP. У некласифікованого

Таблиця 1.

Характеристика нуклеотидних та амінокислотних послідовностей sHSP *O. sativa*

Table 1.

Characteristics of nucleotide and amino acid sequences of *O. sativa* sHSP

№	Назва гена/білка	Довжина CDS		Молекулярна маса, kDa	Номер нуклеотидної послідовності	Номер амінокислотної послідовності	Анотація гена у GenBank
		нп	АК				
<b>Клас CI</b>							
1	OsHSP16.8A-CI	450	149	16.8	NM_001401476.1	NP_001388405.1	<i>O. sativa</i> Japonica Group 16.9 kDa class I HSP 3-like
2	OsHSP16.8B-CI	453	150	16.8	NM_001401477.1	NP_001388406.1	<i>O. sativa</i> Japonica Group 16.9 kDa class I HSP 1-like
3	OsHSP16.8C-CI	453	150	16.8	NM_001401478.1	NP_001388407.1	<i>O. sativa</i> Japonica Group 16.9 kDa class I HSP 2-like
4	OsHSP17.8A-CI	501	166	17.8	NM_001401475.1	NP_001388404.1	<i>O. sativa</i> Japonica Group 17.9 kDa HSP 2-like
5	OsHSP18.0A-CI	483	160	18.0	KT304314.1	ALS88372.1	<i>O. sativa</i> Indica Group cultivar Pathum Thani 1 small HSP gene, complete cds
6	OsHSP18.0B-CI	486	161	18.0	NM_001417479.1	NP_001404408.1	<i>O. sativa</i> Japonica Group 18.1 kDa class I HSP -like
7	OsHSP17.5-CI	480	159	17.5	NM_001417480.1	NP_001404409.1	<i>O. sativa</i> Japonica Group 17.7 kDa class I HSP -like
8	OsHSP17.2-CI	465	154	17.2	NM_001402128.1	NP_001389057.1	<i>O. sativa</i> Japonica Group 17.4 kDa class I HSP -like
9	OsHSP18.1-CI	486	161	18.1	KT304316.1	ANW06227.1	<i>O. sativa</i> cultivar Nam Roo class I Hsp18 gene
10	OsHSP17.8B-CI	486	161	17.8	NM_001402127.1	NP_001389056.1	<i>O. sativa</i> Japonica Group 17.9 kDa class I HSP -like
11	OsHSP17.7-CI	480	159	17.7	U83671.1	AAC78394.1	<i>O. sativa</i> low molecular mass HSP17.7 gene, complete cds
12	OsHSP17.4-CI	465	154	17.4	EU846992.1	ACJ54896.1	<i>O. sativa</i> Japonica Group clone KCS121E02 17.4 kDa HSP mRNA
13	OsHSP16.5-CI	453	150	16.5	NM_001401474.1	NP_001388403.1	<i>O. sativa</i> Japonica Group 16.6 kDa HSP-like
14	OsHSP17.6-CI	495	164	17.6	NM_001416846.1	NP_001403775.1	<i>O. sativa</i> Japonica Group 17.8 kDa HSP-like (LOC4330496), mRNA
15	OsHSP18.7A-CI	534	177	18.7	NM_001409259.1	NP_001396188.1	<i>O. sativa</i> Japonica Group 18.9 kDa HSP (LOC4328171), mRNA
16	OsHSP18.7B-CI	522	173	18.7	NM_001403103.1	NP_001390032.1	<i>O. sativa</i> Japonica Group 18.8 kDa class V HSP-like (LOC4343386), mRNA
<b>Клас CII</b>							
17	OsHSP18.0-CII	501	166	18.0	DQ180746.1	ABA29610.1	<i>O. sativa</i> (japonica cultivar-group) class II small HSP mRNA
18	OsHSP17.9-CII	501	166	17.9	NM_001406238.1	NP_001393167.1	<i>O. sativa</i> Japonica Group 18.0 kDa class II HSP -like

19	OsHSP18.9-CII	528	175	18.9	NM_001409862.1	NP_001396791.1	<i>O. sativa</i> Japonica Group 19.0 kDa class II HSP
<b>Клас CII</b>							
20	OsHSP18.5- CIII	519	172	18.5	NM_001416979.1	NP_001403908.1	<i>O. sativa</i> Japonica Group 18.6 kDa class III HSP -like (LOC4330933), mRNA
<b>Клас P</b>							
21	OsHSP26.6-P	720	239	26.6	AB020973.1	BAA78385.1	<i>O. sativa</i> mRNA for HSP 26, complete cds
22	OsHSP22.1-P	723	240	22.1	NM_001402123.1	NP_001389052.1	<i>O. sativa</i> Japonica Group 26.7 kDa HSP, chloroplastic-like
23	OsHSP21.4-P	603	200	21.4	XM_015755840.3	XP_015611326.1	<i>O. sativa</i> Japonica Group small HSP, chloroplastic
24	OsHSP27.5-P	756	251	27.5	XM_015782094.3	XP_015637580.1	<i>O. sativa</i> Japonica Group small HSP, chloroplastic
25	OsHSP21.0-P	594	197	21.0	XM_015757781.3	XP_015613267.1	<i>O. sativa</i> Japonica Group small HSP, chloroplastic
<b>Клас M</b>							
26	OsHSP23.6A-M	663	220	23.6	NM_001401986.1	NP_001388915.1	<i>O. sativa</i> Japonica Group 24.1 kDa heat shock protein, mitochondrial-like
27	OsHSP23.6B-M	666	221	23.6	NM_001401985.1	NP_001388914.1	<i>O. sativa</i> Japonica Group 24.1 kDa HSP, mitochondrial-like
28	OsHSP23.5-M	660	219	23.5	XM_015772047.3	XP_015627533.1	<i>O. sativa</i> Japonica Group 23.6 kDa HSP, mitochondrial
29	OsHSP26.2-M	747	248	26.2	NM_001421446.1	NP_001408375.1	<i>O. sativa</i> Japonica Group 26.2 kDa HSP, mitochondrial
30	OsHSP24.1-M	663	220	24.1	GU120342.1	ADR66976.1	<i>O. sativa</i> Japonica Group clone KCG005G03 22 kDa HSP mRNA, complete cds
<b>Клас ER</b>							
31	OsHSP19.1-ER	621	208	19.1	NM_001423181.1	NP_001410110.1	<i>O. sativa</i> Japonica Group 21.9 kDa HSP -like (LOC4350180), mRNA
32	OsHSP19.3-ER	552	183	19.3	XM_015757600.3	XP_015613086.1	<i>O. sativa</i> Japonica Group 16.6 kDa HSP -like (LOC9268009), mRNA
33	OsHSP20.7-ER	648	215	20.7	NM_001419362.1	NP_001406291.1	<i>O. sativa</i> Japonica Group 23.2 kDa HSP -like (LOC4335956), mRNA
<b>Клас Per</b>							
34	OsHSP15.9-Per	441	148	15.9	NM_001421521.1	NP_001408450.1	<i>O. sativa</i> Japonica Group 16.0 kDa HSP, peroxisomal-like
<b>Некласифікований</b>							
35	OsHSP22.1	612	203	22.1	NM_001420689.1	NP_001407618.1	<i>O. sativa</i> Japonica Group 22.3 kDa class VI HSP -like

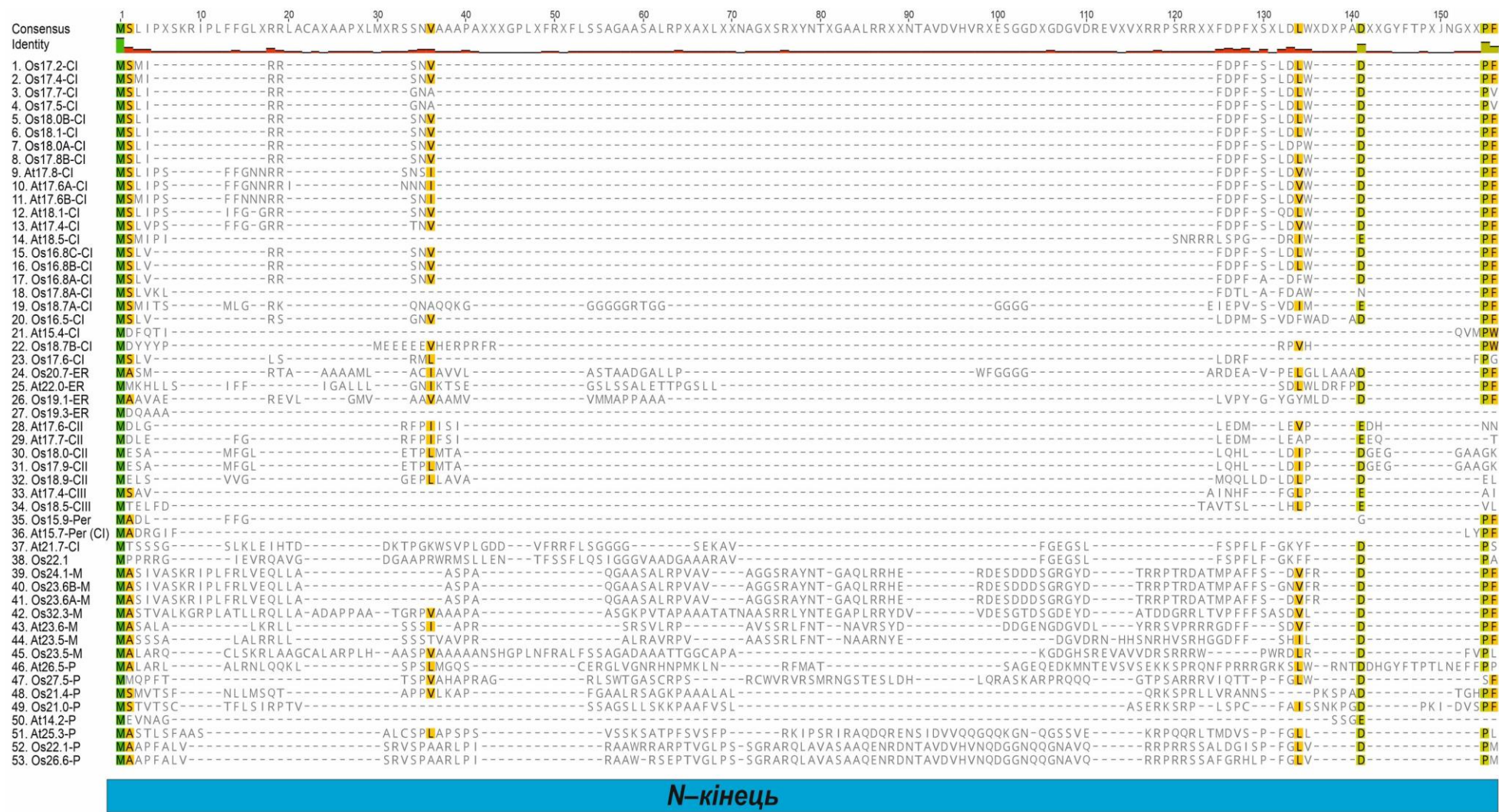


Рис. 1. (початок). Порівняння амінокислотних послідовностей sHSP для *Oryza sativa* (Os), *Arabidopsis thaliana* (At). Стрілками вказано позиції потенційних вторинних структурних елементів у межах ACD.

Fig. 1. Amino acid sequence comparison of sHSP of *Oryza sativa* (Os), *Arabidopsis thaliana* (At). Arrows indicate the positions of potential secondary structural elements within the ACD.

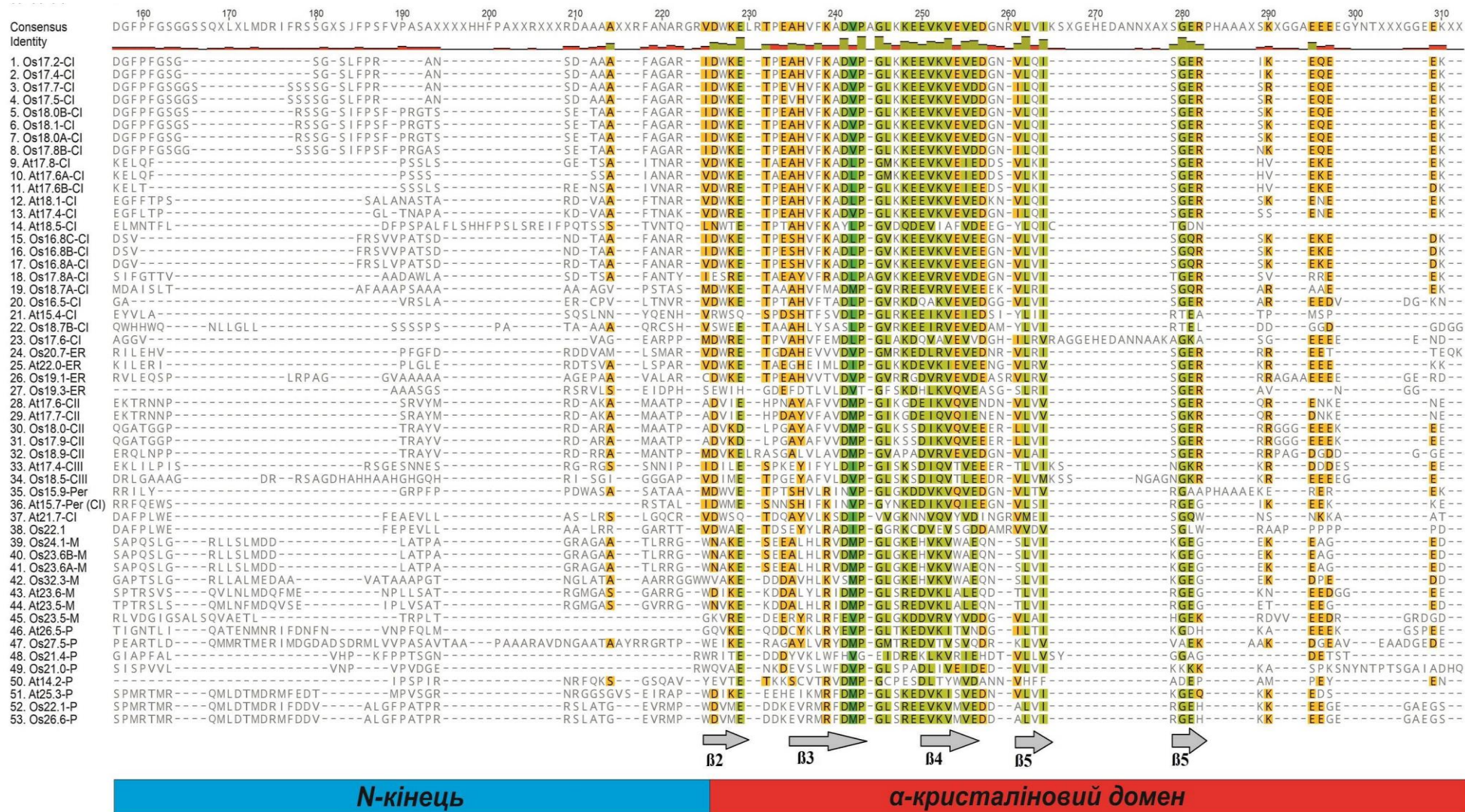


Рис. 1. (продовження).

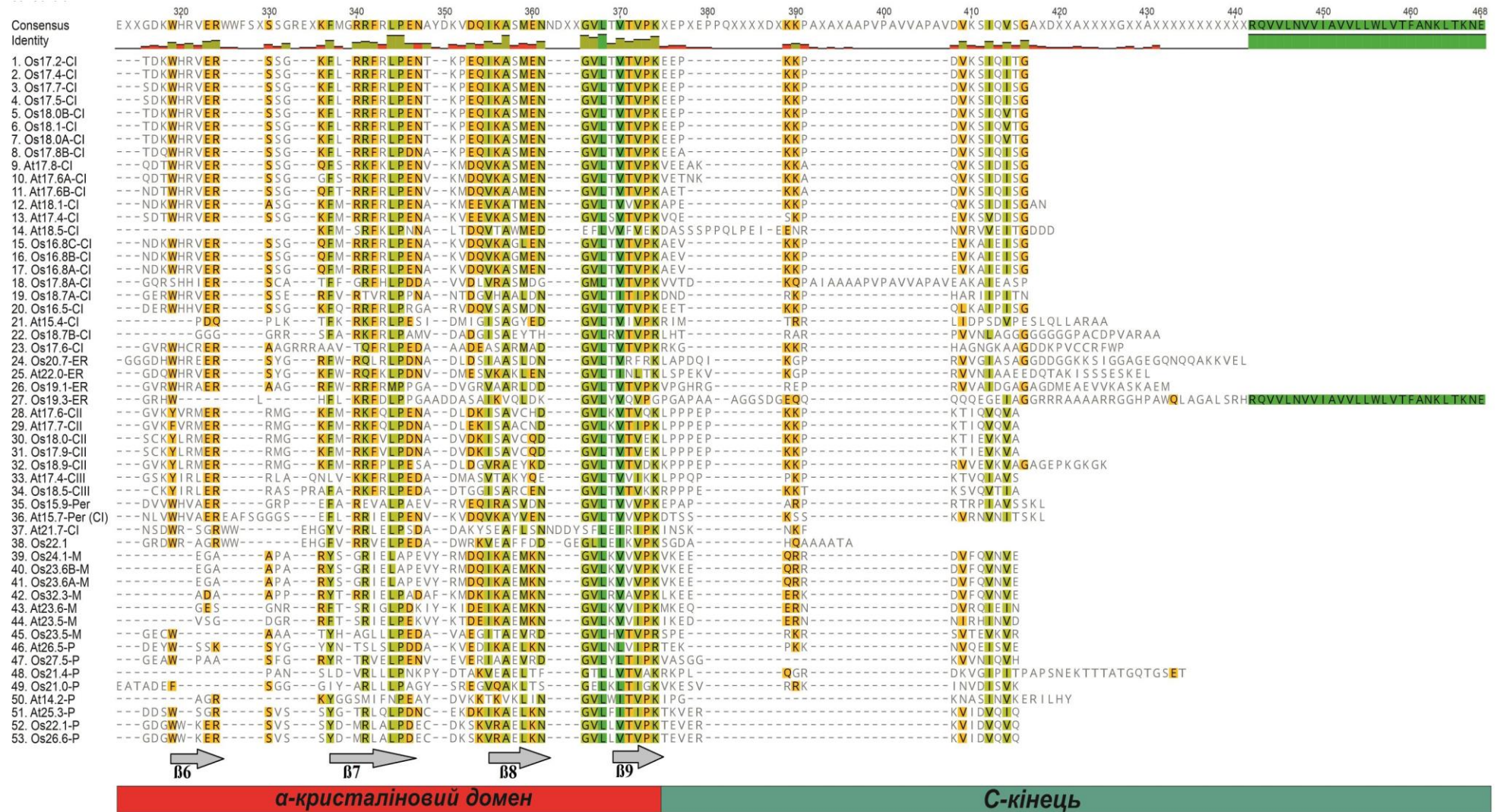
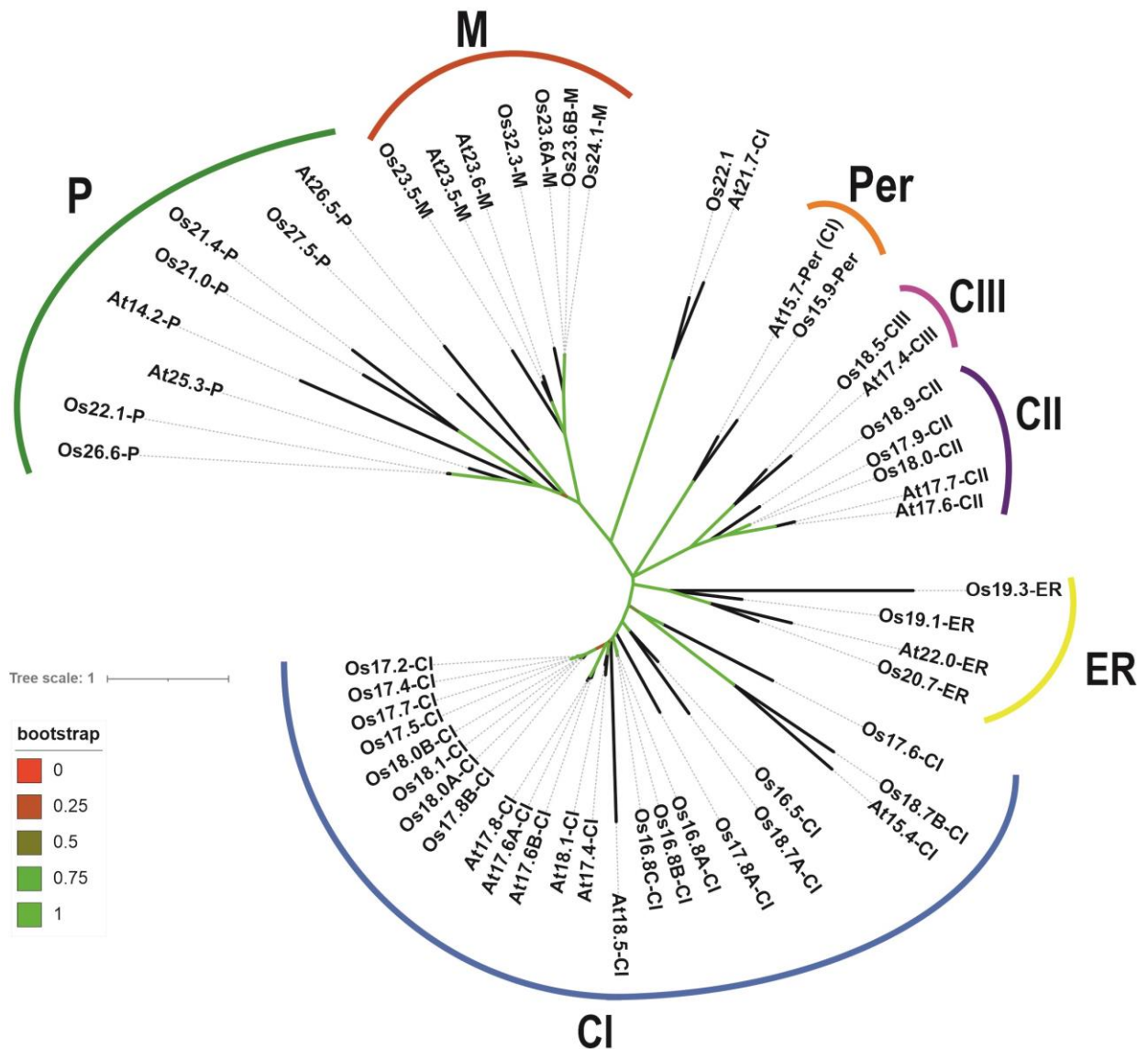


Рис. 1. (закінчення).



**Рис. 2. ML-дендрограма подібності амінокислотних послідовностей sHSP *Oryza sativa*.**  
**Fig. 2. ML-dendrogram reflecting amino acid sequence similarity of *Oryza sativa* sHSP.**

білка OsHSP22.1 ця ділянка складає 18 АК. Незважаючи на різні варіації у N та C ділянках, альфа-кристаліновий домен всіх sHSP містить близько 90 амінокислотних залишків (рис. 1).

До класу CI (cytosolic I) належать 16 sHSP *O. sativa*, з яких вісім білків (Os17.2-CI, Os17.4-CI, Os17.5-CI, Os17.7-CI, Os17.8B-CI, Os18.0A-CI, Os18.0B-CI, Os18.1-CI) є високо подібними між собою і утворюють окрему підгрупу на дендрограмі. Ці білки характеризуються наявністю специфічної амінокислотної послідовності GFPFGSG у N-кінці. Сестринську групу на дендрограмі формують шість послідовностей *A. thaliana* (AtHSP17.4-CI (CAB90950), AtHSP17.6A-CI (AAD39328), AtHSP17.6B-CI (AAC95188), AtHSP17.8-CI (AAF79569), AtHSP18.1-CI (BAB09509), AtHSP18.5-CI (ACC16461) –рис. 2).

В межах класу CI на дендрограмі присутня ще одна підгрупа, яка містить білки рису та арабідопсису (OsHSP17.6-CI, OsHSP18.7B-CI, AtHSP15.4-CI). Всі решта sHSP рису не мають гомологів у геномі арабідопсису і представлені окремими гілками на рис. 2. Виходячи з їх розташування над дендрограмі, можна висунути припущення, що гени цих білків існували у стародавньому предковому геномі ще до дивергенції одно- та дводольних рослин. Пізніше вони були втрачені у арабідопсису, але збереглися у геномі рису.

До класу CII належать три білки (OsHSP17.9-CII, OsHSP18.0-CII, OsHSP18.9-CII), які демонструють високий рівень подібності між собою. Особливістю послідовності OsHSP18.9-CII є наявність 6-амінокислотної делеції – GEGGAA в N-кінцевій ділянці, що відрізняє її від OsHSP17.9-



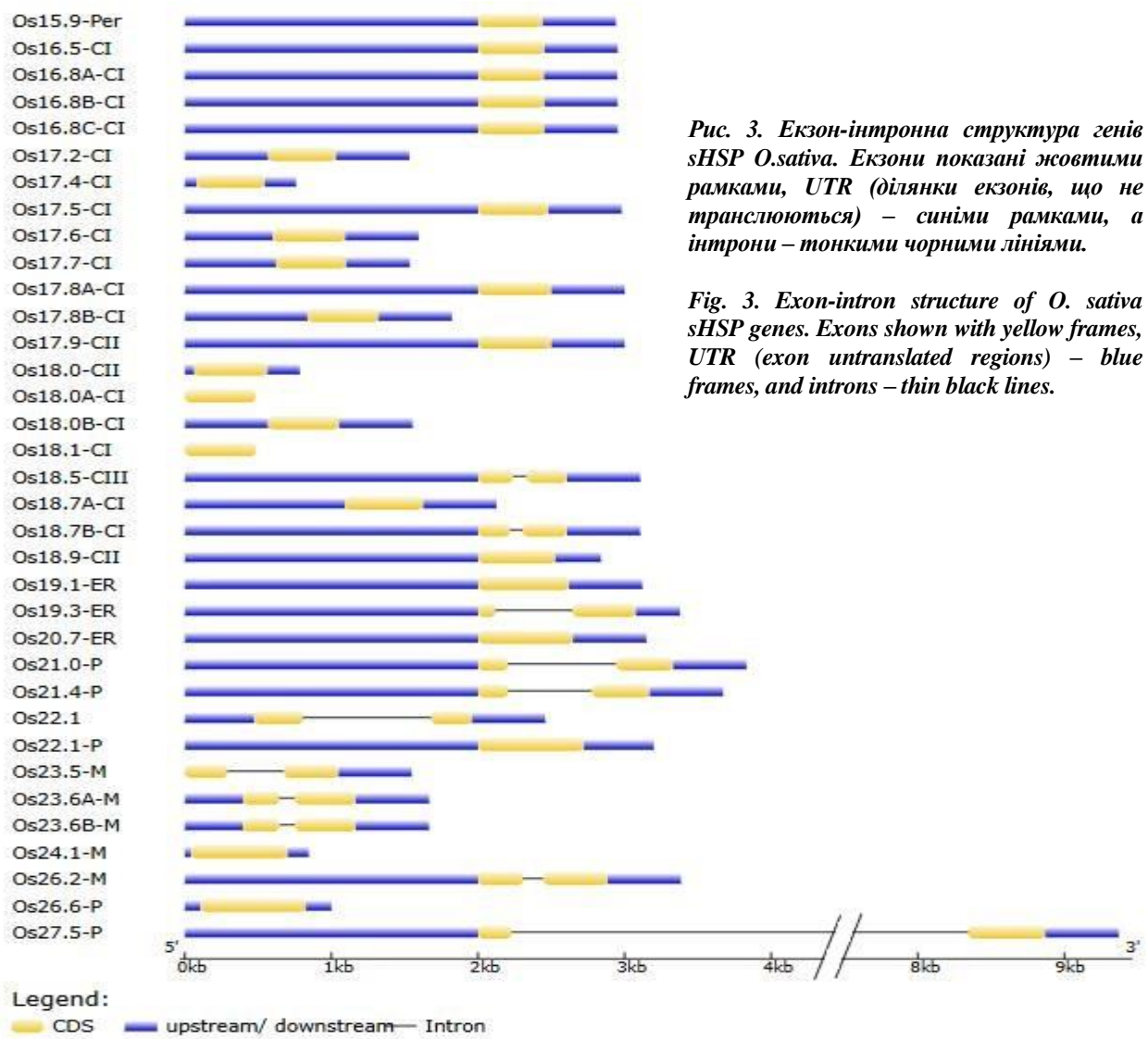


Рис. 3. Екзон-інтронна структура генів sHSP *O.sativa*. Екзони показані жовтими рамками, UTR (ділянки екзонів, що не транслюються) – синіми рамками, а інтрони – тонкими чорними лініями.

Fig. 3. Exon-intron structure of *O. sativa* sHSP genes. Exons shown with yellow frames, UTR (exon untranslated regions) – blue frames, and introns – thin black lines.

СІ та OsHSP18.0-СІ. Крім того, в ACD OsHSP18.9-СІ між β2 та β3 листами наявна вставка двох амінокислот – LR, а також одноамінокислотні заміни.

Клас СІІ представлений у рисі лише одним білком (OsHSP18.5-СІІ), який попри структурну схожість із білками класу СІ, має відмінності в ACD. Білки цих двох класів відрізняються від цитоплазматичних sHSP класу СІ та формують спільну гілку на дендрограмі, що вказує на їх еволюційну дивергенцію від спільного предкового білка. Також ці гени мають високий ступінь подібності із відповідними білками арабідопсису класу СІ та СІІ (AtHSP17.6-СІІ (CAB87675), AtHSP17.7-СІІ (CAB87676), AtHSP17.4-СІІ (AAD25777)).

У геномі *O. sativa* також ідентифіковано ділянки, які кодують п'ять пластидних, п'ять мітохондріальних, три ендоплазматичні та один пероксисомальний sHSP. Пластидні та мітохондріальні білки формують окремі гілки на дендрограмі.

Пластидні білки (OsHSP21.0-Р, OsHSP21.4-Р, OsHSP22.1-Р, OsHSP26.6-Р, OsHSP27.5-Р) характеризуються значною варіабельністю в N-кінцевій ділянці. Крім того, в ACD між β4 та β5 листами послідовність OsHSP21.4-Р містить додатковий амінокислотний залишок треоніну, тоді як у OsHSP27.5-Р в області β5 листа спостерігається вставка залишку аланіну. Делеції в β6 області характерні для Os21.0-Р та Os21.4-Р. Для цього класу також знайдено гомологічні білки у арабідопсису (AtHSP14.2-Р (BAB09086), AtHSP25.3-Р (CAB38279)).

Мітохондріальні білки (OsHSP23.5-М, OsHSP23.6А-М, OsHSP23.6В-М, OsHSP24.1-М, OsHSP26.2-М) мають ряд особливостей. Зокрема, N-кінцева ділянка OsHSP26.2-М, OsHSP23.5-М містить унікальні амінокислотні залишки, які відрізняються від інших білків цієї групи. В ACD цих білків виявлена часткова втрата амінокислотної послідовності для β6 листа. Встановлено високий рівень подібності між OsHSP23.6А-М, OsHSP23.6В-М, OsHSP24.1-М,

OsHSP32.3-M та білками *A. thaliana* AtHSP23.5-M (BAB09755), AtHSP23.6-M (CAA67022).

Ендоплазматичні sHSP (OsHSP19.1-ER, OsHSP19.3-ER, OsHSP20.7-ER) демонструють загальну консервативність в ACD. Проте, у Os19.3-ER в ділянці β6 листа наявна делеція п'яти амінокислот – HRXER Ці білки утворюють на дендрограмі спільну гілку із AtHSP22.0-ER (CAB39778).

Єдиний пероксисомальний sHSP рису (OsHSP15.9-Per) має коротку N-кінцеву область та на C-кінці містить характерний сигнал локалізації для пероксисомальних білків - SKL. Цікаво, що цей білок рису показав високу подібність із sHSP арабідопсису, AtHSP15.7-CI (BAB08313), який початково був описаний як цитоплазматичний (Scharf et. al., 2001), але пізніше в базі даних GenBank переанотований як пероксисомальний.

Окрім згаданих вище, було виявлено білок OsHSP22.1, який не належать до жодного з відомих класів. Він відрізняється додатковим двоамінокислотним залишком GE між β8 та β9 листами, який відсутній у решті білків та формує спільну гілку із AtHSP21.7-CI (BAB09345). Цей білок арабідопсису був раніше анований як CI-like (Scharf et. al., 2001).

Для генів рису, які кодують sHSP було також проаналізовано екзон-інтронну структуру (рис. 3). Аналіз показав, що 12 із 35 досліджуваних генів містять по 1 інтрону. Серед усіх генів найдовший інтрон був виявлений у гені OsHSP27.5-P, а найкоротший – у генах OsHSP18.5-СП, OsHSP18.7В, OsHSP23.6А-М, OsHSP23.6В-М та OsHSP26.2-М. У генах, які кодують білки класів СІ та Per інтрони відсутні. Можлива роль інтронів у регуляції відповіді на тепловий стрес залишається недослідженою.

**Висновки.** За результатами біоінформатичного аналізу у геномі *O. sativa* присутні 35 генів, які кодують sHSP. Ідентифіковані білки мають типову для sHSP структуру. Ці білки класифікуються на кілька основних груп залежно від наявності специфічних амінокислотних послідовностей на їхніх N- та C-кінцях та варіацій у α-кристаліновому домені. Усі досліджувані білки містять консервативний α-кристаліновий домен, що вказує на важливу роль цієї структури в їх функціонуванні як молекулярних шаперонів.

**Подяки.** Автори висловлюють щире подяку професору Р.А. Волкову за участь в обговоренні отриманих результатів та ст. наук. співр. Ю.О. Тинкевичу за корисні поради при проведенні біоінформатичного скринінгу.

## Список літератури / References:

1. Anisimova, M., & Gascuel, O. (2006). Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic Biology*, 55(4), 539–552. <https://doi.org/10.1080/1063515060075545>
2. Boratyn, G. M., Camacho, C., Cooper, P. S., Coulouris, G., Fong, A., Ma, N., Madden, T. L., Matten, W. T., McGinnis, S. D., Merezuk, Y., Raytselis, Y., Sayers, E. W., Tao, T., Ye, J., & Zaretskaya, I. (2013). BLAST: a more efficient report with usability improvements. *Nucleic Acids Research*, 41(W1),W29–W33. <https://doi.org/10.1093/nar/gkt282>
3. Buzduga, I. M., Salamon, I., Volkov, R. A., & Panchuk, I. I. (2022). Rapid accumulation of cadmium and antioxidative response in tobacco leaves. *The Open Agriculture Journal*, 16(1). <https://doi.org/10.2174/18743315-v16-e2206271>
4. Buzduga, I.M., Volkov, R.A., & Panchuk I.I. (2014). Heat stress affects lipid peroxidation and activity of ascorbate peroxidase in *Nicotiana tabacum* [Vplyv teplovoho stresu na peroksidne okyslennia lipidiv ta activnist askorbat peroksidasy u *Nicotiana tabacum*]. *Physiologia Rastenij i Genetika*. 46 (2), 151-157. [In Ukrainian].
5. Chandel, G., Dubey, M., & Meena, R. (2012). Differential expression of heat shock proteins and heat stress transcription factor genes in rice exposed to different levels of heat stress. *Journal of Plant Biochemistry and Biotechnology*, 22(3), 277–285. <https://doi.org/10.1007/s13562-012-0156-8>
6. Chen, X., Lin, S., Liu, Q., Huang, J., Zhang, W., Lin, J., Wang, Y., Ke, Y., & He, H. (2014). Expression and interaction of small heat shock proteins (sHsps) in rice in response to heat stress. *Biochimica Et Biophysica Acta (BBA) - Proteins and Proteomics*, 1844(4),818–828. <https://doi.org/10.1016/j.bbapap.2014.02.010>
7. Doliba, I. M., Volkov, R. A., & Panchuk, I. I. (2011). Activity of catalase and ascorbate peroxidase in Cat2 knock-out mutant of *Arabidopsis thaliana* upon cadmium stress. *Visn. Ukr. Tov. Genet. Sel*, 9(2), 200–209.
8. Guindon, S., & Gascuel, O. (2003). A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, 52(5), 696–704. <https://doi.org/10.1080/10635150390235520>
9. Hagymasi, A. T., Dempsey, J. P., & Srivastava, P. K. (2022). Heat-Shock proteins. *Current Protocols*, 2(11). <https://doi.org/10.1002/cpz1.592>
10. Haslbeck, M., & Vierling, E. (2015). A first line of stress defense: small heat shock proteins and their function in protein homeostasis. *Journal of Molecular Biology*, 427(7), 1537–1548. <https://doi.org/10.1016/j.jmb.2015.02.002>
11. Hassan, M. U., Chattha, M. U., Khan, I., Chattha, M. B., Barbanti, L., Aamer, M., Iqbal, M. M., Nawaz, M., Mahmood, A., Ali, A., & Aslam, M. T. (2020). Heat stress in cultivated plants: nature, impact, mechanisms, and mitigation strategies – a review. *Plant Biosystems*, 155(2), 211–234. <https://doi.org/10.1080/11263504.2020.1727987>
12. Hibshman, J. D., Carra, S., & Goldstein, B. (2023). Tardigrade small heat shock proteins can limit

- desiccation-induced protein aggregation. *Communications Biology*, 6(1). <https://doi.org/10.1038/s42003-023-04512-y>
13. Hu, B., Jin, J., Guo, A., Zhang, H., Luo, J., & Gao, G. (2014). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*, 31(8), 1296–1297. <https://doi.org/10.1093/bioinformatics/btu817>
  14. Ji, X., Yu, Y., Ni, P., Zhang, G., & Guo, D. (2019). Genome-wide identification of small heat-shock protein (HSP20) gene family in grape and expression profile during berry development. *BMC Plant Biology*, 19(1). <https://doi.org/10.1186/s12870-019-2031-4>
  15. Katoh, K., Rozewicki, J., & Yamada, K. D. (2017). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Briefings in Bioinformatics*, 20(4), 1160–1166. <https://doi.org/10.1093/bib/bbx108>
  16. Khan, S., Jabeen, R., Deeba, F., Waheed, U., Khanum, P., & Iqbal, N. (2021). Heat shock proteins: classification, functions and expressions in plants during environmental stresses. *Journal of Bioresource Management*, 8(2), 9. <https://doi.org/10.35691/JBM.1202.0183>
  17. Lee, B., Won, S., Lee, H., Miyao, M., Chung, W., Kim, I., & Jo, J. (2000). Expression of the chloroplast-localized small heat shock protein by oxidative stress in rice. *Gene*, 245(2), 283–290. [https://doi.org/10.1016/s0378-1119\(00\)00043-3](https://doi.org/10.1016/s0378-1119(00)00043-3)
  18. Letunic, I., & Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research*, 49(W1), W293–W296. <https://doi.org/10.1093/nar/gkab301>
  19. Panchuk, I. I., Volkov, R. A., & Schöffl, F. (2002). Heat stress-and heat shock transcription factor-dependent expression and activity of ascorbate peroxidase in *Arabidopsis*. *Plant physiology*, 129(2), 838–853. <https://doi.org/10.1104/pp.001362>
  20. Poudel, P. B., & Poudel, M. R. (2020). Heat stress effects and tolerance in wheat: A review. *J. Biol. Today's World*, 9(3), 1–6. <https://doi.org/10.35248/2322-3308.20.09.217>
  21. Pyrizhok, R. Yu., Volkov, R. A., & Panchuk I. I. (2009). Peroxidase activity in maize seedlings upon heat stress [Actyvnyist peroxidasy prorostkiv kukurudzy v umovakh teplovoho stresu] *Physiologia i biochimia kulturnykh rastenij*. 41 (1): 44–49. [In Ukrainian].
  22. Reddy, P. S., Chakradhar, T., Reddy, R. A., Nitnavare, R. B., Mahanty, S., & Reddy, M. K. (2016). Role of heat shock proteins in improving heat stress tolerance in crop plants. In *Heat shock proteins*, 283–307. [https://doi.org/10.1007/978-3-319-46340-7\\_14](https://doi.org/10.1007/978-3-319-46340-7_14)
  23. Rusnak, T. O., Doliba, I. M., Volkov, R. A., & Panchuk I. I. (2013). Guaiacol peroxidase activity in Cat2 knock-out mutant of *Arabidopsis thaliana* upon heat stress treatment. [Actyvnyist guaiacol peroxidasy u Cat2 knock-out mutantiv *Arabidopsis thaliana* za dii teplovoho stresu]. *Physiologia i biochimia kulturnykh rastenij*. 45 (3), 246–253. [In Ukrainian].
  24. Sakai, H., Lee, S. S., Tanaka, T., Numa, H., Kim, J., Kawahara, Y., Wakimoto, H., Yang, C., Iwamoto, M., Abe, T., Yamada, Y., Muto, A., Inokuchi, H., Ikemura, T., Matsumoto, T., Sasaki, T., & Itoh, T. (2013). Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. *Plant and Cell Physiology*, 54(2), e6. <https://doi.org/10.1093/pcp/pcs183>
  25. Salse, J., Bolot, S., Throude, M., Jouffe, V., Piegu, B., Quraishi, U. M., Calcagno, T., Cooke, R., Delseny, M., & Feuillet, C. (2008). Identification and characterization of shared duplications between rice and wheat provide new insight into grass genome evolution. *The Plant Cell*, 20(1), 11–24. <https://doi.org/10.1105/tpc.107.056309>
  26. Scharf, K. D., Höhfeld, I., & Nover, L. (1998). Heat stress response and heat stress transcription factors. *Journal of biosciences*, 23, 313–329. <https://doi.org/10.1007/bf02936124>
  27. Scharf, K., Siddique, M., & Vierling, E. (2001). The expanding family of *Arabidopsis thaliana* small heat stress proteins and a new family of proteins containing  $\alpha$ -crystallin domains (Acid proteins). *Cell Stress & Chaperones*, 6(3), 225. [https://doi.org/10.1379/1466-1268\(2001\)006](https://doi.org/10.1379/1466-1268(2001)006)
  28. Siddique, M., Gernhard, S., Von Koskull-Döring, P., Vierling, E., & Scharf, K. (2008). The plant sHSP superfamily: five new members in *Arabidopsis thaliana* with unexpected properties. *Cell Stress and Chaperones*, 13(2), 183–197. <https://doi.org/10.1007/s12192-008-0032-6>
  29. Sun, L., Liu, Y., Kong, X., Zhang, D., Pan, J., Zhou, Y., Wang, L., Li, D., & Yang, X. (2012). ZmHSP16.9, a cytosolic class I small heat shock protein in maize (*Zea mays*), confers heat tolerance in transgenic tobacco. *Plant Cell Reports*, 31(8), 1473–1484. <https://doi.org/10.1007/s00299-012-1262-8>
  30. Volkov, R. A., Panchuk, I. I., & Schöffl, F. (2005). Small heat shock proteins are differentially regulated during pollen development and following heat stress in tobacco. *Plant Molecular Biology*, 57, 487–502. <https://doi.org/10.1007/s11103-005-0339-y>
  31. Wang, A., Yu, X., Mao, Y., Liu, Y., Liu, G., Liu, Y., & Niu, X. (2015). Overexpression of a small heat-shock-protein gene enhances tolerance to abiotic stresses in rice. *Plant Breeding*, 134(4), 384–393. <https://doi.org/10.1111/pbr.12289>
  32. Waters, E. R. (2012). The evolution, function, structure, and expression of the plant sHSPs. *Journal of Experimental Botany*, 64(2), 391–403. <https://doi.org/10.1093/jxb/ers355>
  33. Waters, E. R., & Vierling, E. (2020). Plant small heat shock proteins – evolutionary and functional diversity. *New Phytologist*, 227(1), 24–37. <https://doi.org/10.1111/nph.16536>
  34. Yadav, S., Modi, P., Dave, A., Vijapura, A., Patel, D., & Patel, M. (2020). Effect of abiotic stress on crops. In *IntechOpen eBooks*. <https://doi.org/10.5772/intechopen.88434>
  35. Zhang, J., Chen, H., Wang, H., Li, B., Yi, Y., Kong, F., Liu, J., & Zhang, H. (2015). Constitutive expression of a tomato small heat shock protein gene LEHSP21 improves tolerance to high-temperature stress by enhancing antioxidation capacity in tobacco. *Plant Molecular Biology Reporter*, 34(2), 399–409. <https://doi.org/10.1007/s11105-015-0925-3>

# IDENTIFICATION AND BIOINFORMATICS ANALYSIS OF sHSP GENES *ORYZA SATIVA*

**L. V. Kozub, I. I. Panchuk**

*Department of molecular genetics and biotechnology,  
Yuriy Fedkovych Chernivtsi National University  
2 Kotsiubynskoho str. 2, 58012 Chernivtsi, Ukraine*

*Plants have a sessile lifestyle and are constantly exposed to adverse environmental factors such as drought, high temperature, and high soil salinity, which significantly reduce their yield. To counteract these conditions, plants activate defense mechanisms, one of the key elements of which are small heat shock proteins (sHSP). These proteins act as molecular chaperones, protecting cells from damage under stress conditions. Despite their important role, the mechanisms of action of sHSP in rice (*Oryza sativa*), one of the world's main food crops, remain poorly understood.*

*In this study, bioinformatics methods were used to identify, classify, and analyze sHSP genes in the *O. sativa* genome. Amino acid sequences were aligned using the G-INS-I method on the MAFFT server. Phylogenetic analysis was performed using the Maximum likelihood method with the PhyML plugin for Geneious Prime 2023.2.1. Statistical support for branches was calculated using the aLRT Chi2 method. Exon-intron structure was generated on the Gene Structure Display Server 2.0.*

*35 sHSP genes were identified in the *O. sativa* genome, and their amino acid sequences were classified into eight structural classes. The analysis showed significant variability of the N- and C-terminal regions of sHSP with the conservation of the  $\alpha$ -crystalline domain. Plastidic, mitochondrial, endoplasmic and peroxisomal proteins form separate branches on the phylogenetic tree, indicating their evolutionary divergence. A protein that did not fall into known classes was also identified. The construction of the exon-intron structure confirmed the presence of introns in some genes, which may indicate their role in regulating the response to heat stress.*

*The obtained results allow us to better understand the role of sHSP in the mechanisms of plant resistance to abiotic stresses and provide a basis for further research aimed at creating high-yielding rice varieties.*

*Keywords: abiotic stress, bioinformatic analysis, genetic polymorphism, molecular genomics, molecular evolution, multigene family,  $\alpha$ -crystallin domain, sHSP, *O. sativa*.*

*Отримано редколегією 02.11.2024 р.*

## **ORCID ID**

Людмила Козуб: <https://orcid.org/0000-0002-2675-5896>

Ірина Панчук: <https://orcid.org/0000-0002-2837-4480>