

ІДЕНТИФІКАЦІЯ ТА АНАЛІЗ ОРГАНІЗАЦІЇ ГЕНІВ *sHSP* У *NICOTIANA ATTENUATA* ТА *SOLANUM LYCOPERSICUM* (SOLANACEAE)

Л.В. КОЗУБ, Ю.О. ТИНКЕВИЧ, Р.А. ВОЛКОВ, І.І. ПАНЧУК

Чернівецький національний університет імені Юрія Федьковича,
вул. Коцюбинського, 2, 58002, Чернівці, Україна

E-mail: i.panchuk@chnu.edu.ua

Автор для кореспонденції – І.І. Панчук, e-mail: i.panchuk@chnu.edu.ua

Низькомолекулярні білки теплового шоку (*sHSP*) відіграють ключову роль у захисті рослин від абіотичних стресів, а також беруть участь у процесах росту й розвитку. У цій роботі за допомогою біоінформатичних підходів проведено комплексний аналіз мультигенної родини *sHsp* у представників родини Пасльонові (*Solanaceae*), яка об'єднує численні харчові, лікарські та декоративні види. У результаті скринінгу геномів двох представників цієї родини *Nicotiana attenuata* (коштовий тютюн), *Solanum lycopersicum* (томат) та модельної рослини *Arabidopsis thaliana* виявлено, відповідно, 52, 45 і 34 ймовірні гени *sHsp* та визначено їхнє хромосомне розташування. Всі гени тютюну, п'ять генів томату та один ген арабідопсису ідентифіковані вперше. Передбачувані *sHSP* належать до 19 класів. Вони відрізняються за розміром структурних доменів і містять сигнали локалізації в різних клітинних компартментах. Більшість класів сформувалися ще до дивергенції філогенетичних ліній Суперрозидів та Суперастеридів і збереглися впродовж подальшої еволюції дводольних, тоді як окремі класи виявилися специфічними лише для *A. thaliana* або представників родини *Solanaceae*. Для кількох *sHsp* характерна кластерна організація на хромосомах, що виникла внаслідок тандемних дуплікацій окремих генів протягом еволюції. У досліджених геномах ідентифіковано 15 *sHsp*, які кодують від двох до чотирьох альтернативних транскриптів. Аналіз промоторних ділянок показав наявність численних сайтів впізнання транскрипційних факторів, залучених до регуляції стресових відповідей, дії світла, фітогормонів, а також контролю росту й розвитку.

Ключові слова: абіотичний стрес, біоінформатичний аналіз, будова промотора, молекулярна геноміка, мультигенні родини, α -кристаліновий домен (ACD), *sHSP*/*HSP20*, *Arabidopsis*, *Nicotiana*, *Solanum*.

Вступ. Рослини ведуть прикріплений спосіб життя і постійно зазнають впливу несприятливих чинників довкілля. Абіотичні стресори —

такі як підвищена температура, посуха, засолення ґрунту, надмірне освітлення та високі концентрації іонів важких металів — можуть викликати значне порушення фізіологічних процесів, зниження життєздатності, росту й продуктивності (Hassan et al., 2020; Yadav et al., 2020; Buzduga et al., 2022; Saharan et al., 2022).

Для нейтралізації негативних наслідків стресу рослини активують чисельні захисні механізми, які включають зміни експресії багатьох генів, що призводить до синтезу стрес-асоційованих метаболітів та спеціалізованих захисних білків. Серед них важливими є білки теплового шоку (Heat Shock Proteins — HSP), які виконують роль молекулярних шаперонів, стабілізуючи інші білки та запобігаючи їх денатурації або сприяючи правильному фолдингу. HSP класифікують відповідно до молекулярної маси, амінокислотної послідовності, клітинної локалізації та функцій. За молекулярною масою, HSP зазвичай групуються на такі підродини: HSP100, HSP90, HSP70, HSP60, HSP40 та HSP20/*sHSP*. У вищих рослин найбільш чисельною та різноманітною групою є низькомолекулярні HSP20 (small HSP або *sHSP*), які діють незалежно від АТФ і є ключовими у запобіганні агрегації частково денатурованих білків (Feder and Hofmann, 1999; Haslbeck et al., 2005; Haslbeck and Vierling, 2015).

Спільною структурною рисою *sHSP* є наявність консервативного α -кристалінового домену (ACD), розташованого в центральній частині білкової молекули. Цей домен розміром близько 90 АК зазвичай має структуру β -сендвіча, утворену з двох антипаралельних β -листів, які забезпечують здатність до олігомеризації та взаємодії з іншими білками. Крім ACD до складу *sHSP* входять відносно довга варіабельна N-кінцева та коротка C-кінцева ділянки. N-кінцева область, яка відрізняється

у різних класів sHSP, відповідає за специфічну взаємодію із певними субстратами, тоді як С-кінцева ділянка часто містить IXI/V-мотив, який стабілізує взаємодію між мономерами. Функціональна ефективність sHSP обумовлена здатністю до утворення стабільних олігомерних структур та взаємодії з пошкодженими білками (Van Montfort et al., 2001; Haslbeck et al., 2005).

У рослин sHSP представлені великими мультигенними/мультипротейновими родинами, члени яких локалізуються в різних клітинних компартментах: цитоплазмі, ядрах, хлоропластах, мітохондріях, ендоплазматичному ретикулумі та пероксисомах (Scharf et al., 2001; Waters and Vierling, 2020). Така різноманітна локалізація свідчить про їх функціональну спеціалізацію у забезпеченні клітинного гомеостазу під час стресу.

Окрім теплового, sHSP захищають рослини від дії холодового та осмотичного стресів та надмірного освітлення, покращуючи роботу фотосинтетичного апарату шляхом стабілізації клітинних мембран, що призводить до зниження продукції АФК (Jiang et al., 2009; Ruibal et al., 2013; Wu et al., 2022). Також ці білки беруть участь у захисті рослин від патогенів, оскільки задіяні у координації імунної відповіді (Maimbo et al., 2007). Крім того, за нестресових умов sHSP активуються під час мітозу, ембріогенезу, при розвитку пилку та дозрівання плодів та насіння (Wehmeyer et al., 1996; Volkov et al., 2005; Ji et al., 2019; González-Gordo et al., 2023).

Характерний для sHSP ACD виявлений також у деяких інших білках, наприклад у IDM-білках (Increased DNA Methylation proteins), які беруть участь у регуляції епігенетичних процесів, зокрема в контролі метилування ДНК (Grimanelli and Ingouff, 2020).

В еволюції вищих рослин відбувалось збільшення кількості sHSP та їх диверсифікація (Waters and Vierling, 2020). У Покритонасінних збільшення мультигенної родини sHSP пов'язане із повногеномними дуплікаціями, які мали місце в еволюції різних таксономічних груп, а також із дуплікаціями окремих генів sHSP (Krsticevic et al., 2016; Zhang et al., 2023; Huang et al., 2025). Збільшення кількості копій генів sHSP є передумовою для їх регуляторної (диференційна експресія дуплікованих копій у відповідь на стресові сигнали, у різних тканинах та/або на

різних стадіях онтогенезу) або функціональної (різна субстратна специфічність) диверсифікації, наслідком чого є підвищена пластичність стресової відповіді та еволюційної адаптованості рослин (Waters and Vierling, 2020).

З огляду на глобальні кліматичні зміни, розуміння механізмів стресового захисту рослин набуває все більшої актуальності. Водночас, значний прогрес у розшифруванні повних геномів еукаріот, якого було досягнуто за останнє десятиліття, відкрив принципово нові можливості у дослідженнях молекулярного різноманіття мультигенних родин. Зокрема, поліморфізм генів *sHsp* на сьогодні досліджено у представників багатьох родин Покритонасінних (Krsticevic et al., 2016; Ji et al., 2019; Chen et al., 2021; Zhang et al., 2023; Huang et al., 2025).

Значний інтерес становить вивчення sHSP у представників родини Пасльонові (Solanaeae), до якої належать чисельні харчові, лікарські та декоративні рослини світового значення (Hunziker, 2001). На сьогодні мультигенна родина *sHsp* охарактеризована для таких важливих видів, як картопля (*Solanum tuberosum* L. – Zhao et al., 2018), томат (*S. lycopersicum* L. – Yu et al., 2016), баклажан (*S. melongena* L. – Gong et al., 2021) та паприка (*Capsicum annuum* L. – González-Gordo et al., 2023), які є відносно спорідненими між собою (Komarova et al., 2008; Tynkevich et al., 2022; Deanna et al., 2025). Проте, різноманіття sHSP все ще залишається недостатньо вивченим у представників таксономічно більш віддаленого роду Тютюн (*Nicotiana*), хоча синтез HSP у рослин після впливу підвищених температур було вперше описано саме у *N. tabacum* (Barnett et al., 1980). Попередньо для *N. tabacum* нами було ідентифіковано 11 генів sHSP та оцінено їх експресію під час розвитку пилку та за дії теплового та холодового стресів (Volkov et al., 2005). У нашій новій роботі ми застосували біоінформатичні методи для ідентифікації та аналізу структурної організації, хромосомного розташування та еволюції генів sHSP у геномі іншого представника цього роду, *N. attenuata*. Також у геномах *S. lycopersicum* та *Arabidopsis thaliana* ми ідентифікували декілька нових генів *sHsp*, які не було виявлено у попередніх дослідженнях (Yu et al., 2016; Scharf et al., 2001; Siddique et al., 2008).

Матеріали і методи. Виявлення генів *sHsp* у геномах досліджуваних видів здійснювали з використанням біоінформатичних підходів. Початковий пошук амінокислотних послідовностей *sHsp* проводили в базі даних GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) із застосуванням алгоритму BLASTp (Boratyn et al., 2013). Для *N. attenuata* було використано два варіанти зборки геному: (1) анотована референтна версія NIATTr2 (GCF_001879085.1) від 15.11.2016 та (2) новіша неанотована версія Niat.v3 (GCA_030864195.1) від 25.08.2023. Для *S. lycopersicum* було використана остання анотована референтна версія SLM_r2.1 (GCF_036512215.1) від 03.09.2024, а для *A. thaliana* – TAIR10.1 (GCA_000001735.2) від 15.03.2018.

Як запити для пошуку були застосовані описані раніше амінокислотні послідовності *sHSP* *S. lycopersicum*, *S. tuberosum* та *A. thaliana* (Scharf et al., 2001; Siddique et al., 2008; Yu et al., 2016; Zhao et al., 2018). Для подальшого аналізу відбирали послідовності, які показували високу ідентичність, містили типові ACD та мали молекулярну масу не більше 42 кДа, що вважається характерною ознакою *sHSP* (Haslbeck and Vierling, 2015; Hassan et al., 2020; Hagymasi et al., 2022).

Передбачення субклітинної локалізації *sHSP* здійснювали інструментами WoLF PSORT (<https://wolfpsort.hgc.jp>) та PISP (<https://bioinfo.usu.edu/Plant-mSubP/>). Для підтвердження використовували пошук N-термінальних сигнальних послідовностей за допомогою TargetP 2.0 (<https://services.healthtech.dtu.dk/service.php?TargetP-2.0>).

Оцінка подібності *sHSP* та побудова дендрограми. Вирівнювання амінокислотних послідовностей проводили методом G-INS-i на сервері MAFFT (<https://mafft.cbrc.jp/alignment/software/>) (Katoh et al., 2017). Філогенетичну дендрограму будували методом Maximum Likelihood (ML) з використанням плагіна PhyML для Geneious Prime 2023.2.1 (Guindon and Gascuel, 2003). Для статистичної підтримки гілок дерев використовували метод aLRT Chi2 (Anisimova and Gascuel, 2006). Отриману дендрограму експортували у форматі Newick та візуалізували за допомогою онлайн-інструменту Interactive Tree of Life (iTOL) v6 (<https://itol.embl.de>) (Letunic and Bork, 2021).

Аналіз екзон-інтронної структури відібраних генів проводили за допомогою онлайн-ресурсу Gene Structure Display Server (GSDS) 2.0 (<http://gsds.gao-lab.org>) (Hu et al., 2014). Геномні та кодувальні послідовності відповідних генів завантажували у форматі FASTA.

Аналіз хромосомної локалізації *sHsp*. GFF-файли для геномів *N. attenuata* і *S. lycopersicum* були створені на основі отриманих нами координат *sHsp* генів. Для визначення хромосомного розташування *sHsp* у *A. thaliana* та *S. lycopersicum* використовували доступні у базі GenBank анотації до повних геномних збірок, зокрема координати локалізації генів на конкретних хромосомах.

Для *N. attenuata* використовували інший підхід, оскільки у анотованій версії NIATTr2 геному станом на серпень 2025 р. наведено хромосомну локалізацію лише 23 з 52 генів *sHsp*, тоді як решта 29 генів розташовані лише у нерозміщених ДНК-скефолдах (табл. 1). Відповідно, ми здійснили пошук *sHsp* на послідовностях окремих хромосом у більш новій неанотованій збірці геному Niat.v3 за допомогою локального вирівнювання у програмі UGENE v46.0 (<http://ugene.net>) (Okonechnikov et al., 2012). Це дозволило визначили хромосомах локалізацію 49 генів *sHsp* *N. attenuata*. Для решти трьох генів виявлено, що вони знаходяться у ДНК-скефолдах (табл. 1). Візуалізацію розташування *sHsp* на хромосомах здійснювали у програмі TBtools.

Аналіз цис-регуляторних елементів у промоторних ділянках генів *sHsp*. Для ідентифікації ймовірних цис-регуляторних елементів у промоторних регіонах генів родини *sHsp* було використано комбінований підхід. Пошук цис-елементів в послідовностях довжиною 2000 нп перед старт-кодоном кожного гена проводився за допомогою онлайн-сервісу PlantCARE (<https://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). Це дозволило виявити функціонально охарактеризовані елементи, пов'язані із відповіддю на освітлення, рослинні гормони, біотичний і абіотичний стрес, а також елементи, пов'язані із ростом і розвитком. Додатково за допомогою інструменту FIMO з пакету MEME Suite (<https://meme-suite.org/meme/tools/fimo>) було проведено пошук промоторних елементів HSE (heat shock elements), які є сайтами зв'я-

Таблиця 1. Характеристика генів *sHsp* *Nicotiana attenuata*, *Solanum lycopersicum* та *Arabidopsis thaliana*

Назва гена		Локус
Нова	Стара	
<i>NatHsp16.0-CI</i>		109230573
<i>NatHsp17.2-CI</i>		109239375
<i>NatHsp17.4A-CI</i>		109239381
<i>NatHsp17.4B-CI</i>		109239378
<i>NatHsp17.4C-CI</i>		109239380
<i>NatHsp17.5-CI</i>		109239377
<i>NatHsp17.9A-CI</i>		109229327
<i>NatHsp17.9B-CI</i>		109211377
<i>NatHsp17.9C-CI</i>		109242695
<i>NatHsp17.9D-CI</i>		109242698
<i>NatHsp18.0A-CI</i>		109211376
<i>NatHsp18.0B-CI</i>		109219819
<i>NatHsp18.8-CI</i>		109239374
<i>NatHsp20.6-CI</i>		109229430
<i>NatHsp22.2-CI</i>		109214667
<i>SlyHsp7.8-CI</i>		NA
<i>SlyHsp9.1-CI</i>		NA
<i>SlyHsp17.5A-CI</i>	<i>SlHsp17.6A</i>	101264936
<i>SlyHsp17.5B-CI</i>	<i>SlHsp17.6B</i>	101266525
<i>SlyHsp17.5C-CI</i>	<i>SlHsp17.6C</i>	543848
<i>SlyHsp17.6A-CI</i>	<i>SlHsp17.7A</i>	101266231
<i>SlyHsp17.6B-CI</i>	<i>SlHsp15.2</i>	101268307
<i>SlyHsp17.7-CI</i>		138349016
<i>SlyHsp24.4-CI</i>	<i>SlHsp24.5</i>	101263239
<i>SlyHsp25.4-CI</i>	<i>SlHsp17.7B</i>	101268020
<i>SlyHsp27.0-CI</i>	<i>SlHsp27.1</i>	104644406
<i>AthHsp17.3-CI</i>	<i>AtHSP17.4-CI</i>	823768
<i>AthHsp17.5A-CI</i>		841789
<i>AthHsp17.5B-CI</i>	<i>AtHSP17.6A-CI</i>	842280
<i>AthHsp17.4-CI</i>	<i>AtHSP17.6B-CI</i>	817499
<i>AthHsp17.7-CI</i>	<i>AtHSP17.8-CI</i>	837252
<i>AthHsp18.0-CI</i>	<i>AtHSP18.1-CI</i>	836093
<i>AthHsp18.4-CI(CIV)</i>	<i>AtHSP18.5-CIV</i>	816448
<i>NatHsp15.2-CII</i>		109220085
<i>NatHsp17.4A-CII</i>		109217313
<i>NatHsp17.4B-CII</i>		109236242
<i>NatHsp17.5A-CII</i>		109214022
<i>NatHsp17.5B-CII</i>		109236218
<i>NatHsp17.5C-CII</i>		109220084
<i>NatHsp18.1-CII</i>		109236241
<i>NatHsp21.1-CII</i>		109235601
<i>NatHsp30.6-CII</i>		109236008
<i>SlyHsp17.2-CII</i>	<i>SlHsp17.3B</i>	543572
<i>SlyHsp17.5-CII</i>	<i>SlHsp17.6D</i>	544282
<i>AthHsp17.5-CII</i>	<i>AtHSP17.6-CII</i>	831075
<i>AthHsp17.6-CII</i>	<i>AtHSP17.7-CII</i>	831076
<i>NatHsp16.8-CIII</i>		109223973
<i>SlyHsp16.0-CIII</i>	<i>SlHsp16.1A</i>	101248279
<i>AthHsp17.2-CIII</i>	<i>AtHSP17.4-CIII</i>	841843
<i>SlyHsp15.6-CV</i>	<i>SlHsp15.7</i>	101256444

Геномна збірка 1			Геномна збірка 2		
Ch/Sc	№ доступу у GenBank	Розташування, нп	Ch/Sc	№ доступу у GenBank	Розташування, нп
Ch7	NC_031995.1	87909685-87910536	Ch2	CM061563.1	2177229-2177654
Sc	NW_017670436.1	165132-65876	Ch1	CM061562.1	444011622-444012366
Sc	NW_017670436.1	278655-279449	Ch1	CM061562.1	443886612-443887406
Sc	NW_017670436.1	154798-55554	Ch1	CM061562.1	444021942-444022698
Sc	NW_017670436.1	255828-256629	Ch1	CM061562.1	443910684-443911485
Sc	NW_017670436.1	140380-141184	Ch1	CM061562.1	444036316-444037120
Ch7	NC_031995.1	76785566-76786954	Ch1	CM061562.1	240029578-240030966
Sc	NW_017672021.1	145033-145866	Ch1	CM061562.1	150073446-150074279
Sc	NW_017670725.1	425054-426015	Ch4	CM061565.1	116896013-116896974
Sc	NW_017670725.1	416016-416911	Ch4	CM061565.1	116886978-116887873
Sc	NW_017672021.1	139691-140628	Ch1	CM061562.1	150078679-150079616
Sc	NW_017675222.1	55324-56284	Ch6	CM061567.1	33089497-33090457
Sc	NW_017670436.1	138463-139146	Ch1	CM061562.1	444038354-444039037
Ch7	NC_031995.1	87987727-87989487	Ch2	CM061563.1	2089350-2091110
Sc	NW_017672840.1	132682-133595	Ch8	CM061569.1	38562615-38563528
Ch9	NC_090808.1	50526163- 50526366			
Ch7	NC_090806.1	60921512-60922395			
Ch6	NC_090805.1	50066124-50066933			
Ch6	NC_090805.1	50060408-50061143			
Ch6	NC_090805.1	50061689-50062492			
Ch6	NC_090805.1	50050440-50051228			
Ch9	NC_090808.1	7986871-7987796			
Ch6	NC_090805.1	50054746-50055492			
Ch9	NC_090808.1	5489933-5491472			
Ch9	NC_090808.1	7999557-8000441			
Ch10	NC_015447.3	65576220-65577277			
Ch3	NC_003074.8	16984118-16984867			
Ch1	NC_003070.9	19980299-19981199			
Ch1	NC_003070.9	22031343-22032250			
Ch2	NC_003071.7	12633065-12634120			
Ch1	NC_003070.9	2274926-2275783			
Ch5	NC_003076.8	24062323-24063300			
2	NC_003071.7	8369568-8370675			
Sc	NW_017675447.1	31200-32073	Ch9	CM061570.1	66325011-66325884
Sc	NW_017673793.1	78890-79701	Ch9	CM061570.1	66384854-66385665
Ch12	NC_032000.1	40457523-40458464	Ch9	CM061570.1	66520515-66521456
Sc	NW_017672664.1	71486-72851	Ch7	CM061568.1	137460151-137461516
Ch12	NC_032000.1	32064487-32065309	Ch9	CM061570.1	66237679-66238501
Sc	NW_017675447.1	25469-26205	Ch9	CM061570.1	66319280-66320016
Ch12	NC_032000.1	40451274..40452294	Ch9	CM061570.1	66526833-66527853
Ch12	NC_032000.1	40448784-40449690	Ch9	CM061570.1	66535217-66536123
Ch12	NC_032000.1	32105241-32113853	Ch9	CM061570.1	66197081-66197545
Ch8	NC_090807.1	52329516-52330178			
Ch8	NC_090807.1	52525545-52526281			
Ch5	NC_003076.8	3882231-3883111			
Ch5	NC_003076.8	3883957-3884838			
Ch3	NC_031991.1	66126649-66128171	Ch4	CM061565.1	188234086-188234280
Ch3	NC_090802.1	67226665-67227517			
Ch1	NC_003070.9	20179357-20180566			
Ch2	NC_090801.1	60829219-60830177			

Назва гена		Локус
Нова	Стара	
<i>AthHsp15.3-CV</i>	<i>AtHSP15.4-CV</i>	828276
<i>NatHsp21.0-CVI</i>		109209891
<i>SlyHsp21.4-CVI</i>	<i>SlHsp21.6B</i>	101260554
<i>AthHsp21.5-CVI</i>	<i>AtHSP21.7-CVI</i>	835555
<i>NatHsp18.4-CVII</i>		109222355
<i>NatHsp19.2-CVII</i>		109208449
<i>SlyHsp17.8-CVII</i>	<i>SlHsp17.9</i>	101261225
<i>NatHsp24.4-CVIII</i>		109206036
<i>SlyHsp23.6-CVIII</i>	<i>SlHsp23.7</i>	101246615
<i>SlyHsp23.7-CVIII</i>		101252073
<i>NatHsp21.6A-ER</i>		109234957
<i>NatHsp21.6B-ER</i>		109242134
<i>NatHsp22.4-ER</i>		109214575
<i>SlyHsp19.1-ER</i>	<i>SlHsp21.5C</i>	544205
<i>SlyHsp21.4-ER</i>	<i>SlHsp21.5A</i>	101256272
<i>SlyHsp21.5-ER</i>	<i>SlHsp21.6A</i>	101254946
<i>AthHsp21.9-ER</i>	<i>AtHSP22.0-ER</i>	826616
<i>NatHsp18.2-MPI</i>		109216819
<i>NatHsp21.4-MPI</i>		109216814
<i>NatHsp24.3-MPI</i>		109216817
<i>SlyHsp18.3-MPI</i>	<i>SlHsp18.2</i>	101254323
<i>SlyHsp21.4-MPI</i>	<i>SlHsp21.5B</i>	101250041
<i>SlyHsp23.7-MPI</i>	<i>SlHsp23.8B</i>	543507
<i>AthHsp14.6-MPI</i>	<i>AtHSP14.7-CVII</i>	834810
<i>AthHsp21.6-MPI</i>		834809
<i>AthHsp23.3-MPI</i>	<i>AtHSP23.5-MI</i>	835218
<i>AthHsp23.5-MPI</i>	<i>AtHSP23.6-MI</i>	828623
<i>AthHsp25.2-MPI</i>	<i>AthAcd25.4</i>	827353
<i>AthHsp27.6-MPI</i>	<i>AthAcd26.0</i>	814831
<i>AthHsp28.8-MPI</i>	<i>AthAcd28.9</i>	834809
<i>NatHsp25.0-MPII</i>		109224675
<i>NatHsp26.6-MPII</i>		109235413
<i>SlyHsp26.2-MPII</i>	<i>SlHsp9.0</i>	101261304
<i>AthHsp26.6-MPII</i>	<i>AtHSP26.5-MII</i>	841687
<i>NatHsp21.9-NPI</i>		109224761
<i>NatHsp16.7-NPI</i>		109224759
<i>SlyHsp16.6-NPI</i>		101266717
<i>SlyHsp21.8-NPI</i>		101267233
<i>AthHsp15.4-NPI</i>	<i>AthAcd15.5</i>	843977
<i>AthHsp22.2-NPI</i>	<i>AthAcd22.2</i>	841924
<i>AthHsp39.3-NPI</i>	<i>AthAcd39.0</i>	841923
<i>NatHsp29.8-NPII</i>		109210066
<i>SlyHsp26.4-NPII</i>	<i>SlHsp26.5</i>	101259555
<i>AthHsp31.1-NPII</i>	<i>AthAcd31.2</i>	837158
<i>NatHsp25.5-PI</i>		109233228
<i>NatHsp26.6-PI</i>		109227913
<i>SlyHsp26.0-PI</i>	<i>SlHsp25.7B</i>	543846
<i>SlyHsp26.1-PI</i>	<i>SlHsp26.2</i>	544024
<i>AthHsp25.2-PI</i>	<i>AtHSP25.3-P</i>	828881
<i>AthHsp16.8-PII</i>	<i>AthAcd16.9</i>	827140

Геномна збірка 1			Геномна збірка 2		
Ch/Sc	№ доступу у GenBank	Розташування, нп	Ch/Sc	№ доступу у GenBank	Розташування, нп
Ch4	NC_003075.7	11603481-11604637			
Sc	NW_017671728.1	141093-142667	Ch2	CM061563.1	67369101-67370675
Ch7	NC_090806.1	68766565-68768564			
Ch5	NC_003076.8	22203717-22205152			
Sc	NW_017688760.1	942-2438	Sc	JAPQWU010000013.1	27796003-27797499
Sc	NW_017671463.1	219968-221296	Sc	JAPQWU010000015.1	12571685-12572796
Ch4	NC_090803.1	61012042-61013791			
Sc	NW_017671119.1	271106-272300	Ch4	CM061565.1	21259311-21260374
Ch3	NC_090802.1	60219585-60221122			
Ch3	NC_090802.1	60215399-60216387			
Ch11	NC_031999.1	6767199-6768231	Ch6	CM061567.1	39703792-39704824
Sc	NW_017670673.1	69850-70909	Ch3	CM061567.1	147443451-147444510
Sc	NW_017672805.1	58673-59509	Ch4	CM061565.1	177514339-177515175
Ch11	NC_090810.1	11700552-11701450			
Ch3	NC_090802.1	60758453-60759296	Ch3		60758453-60759296
Ch1	NC_090800.1	87611452-87612417	Ch1		87611452-87612417
Ch4	NC_003075.7	6370338-6371372			
Sc	NW_017673598.1	74342-75635	Ch9	CM061570.1	44749889-44751182
Sc	NW_017673598.1	59721-61074	Ch9	CM061570.1	44764450-44765803
Sc	NW_017673598.1	16497-18158	Ch9	CM061570.1	44810880-44812541
Ch8	NC_090807.1	64628352-64630100			
Ch8	NC_090807.1	64625806-64627261			
Ch8	NC_090807.1	64622980-64624109			
Ch5	NC_003076.8	19299717-19300503			
Ch5	NC_003076.8	19297614-19299174			
Ch5	NC_003076.8	20891163-20892416			
Ch4	NC_003075.7	12916926-12918110			
Ch4	NC_003075.7	9313602-9317262			
Ch2	NC_003071.7	882645-884164			
Ch5	AB025628.1	19297614-19299174			
Ch3	NC_031991.1	59500541-59502650	Ch4	CM061565.1	3775052-3777160
Ch12	NC_032000.1	21892775-21894700	Ch7	CM061568.1	169170395-169172320
Ch12	NC_090811.1	60221319-60223746			
Ch1	NC_003070.9	19574538-19575866			
Ch4	NC_031992.1	31098-34663	Ch11	CM061572.1	28976817-28978717
Ch4	NC_031992.1	21184-24009	Ch11	CM061572.1	28987635-28988561
Ch4	NC_090803.1	68183435-68186592			
Ch4	NC_090803.1	68190691-68194430			
Ch1	NC_003070.9	28682084-28683785			
Ch1	NC_003070.9	20454389-20456220			
Ch1	NC_003070.9	20452593-20454266			
Sc	NW_017671761.1	110800-113708	Ch2	CM061563.1	91156544-91159451
Ch7	NC_090806.1	66144079-66146023			
Ch1	NC_003070.9	1967087..1969656			
Ch10	NC_031998.1	30395436-30397965	Ch3	CM061567.1	35735355-35737883
Ch1	NC_031989.1	65988880-65990547	Ch4	CM061565.1	48412157-48413824
Ch5	NC_090804.1	8420977-8423572			
Ch3	NC_090802.1	49105500-49106445			
Ch4	NC_003075.7	13818825-13819971			
Ch4	NC_003075.7	8510678-8511928			

Назва гена		Локус
Нова	Стара	
<i>AthHsp22.0-PII</i>	<i>AthAcd22.1</i>	821824
<i>NatHsp15.6-Po</i>		109206223
<i>SlyHsp15.9-Po</i>	<i>SlHsp16.1B</i>	101268271
<i>AthHsp15.6-Po</i>	<i>AtHSP15.7-Per</i>	833746
<i>SlyHsp11.9-MCI</i>		NA
<i>SlyHsp15.8-MCI</i>		NA
<i>SlyHsp17.3-MCI</i>		NA
<i>SlyHsp26.0-MCI</i>	<i>SlHsp14.5</i>	109119557
<i>SlyHsp15.5-MCII</i>	<i>SlHsp15.6</i>	101256536
<i>NatHsp38.1-MCIII</i>		109229692
<i>SlyHsp27.7-MCIII</i>	<i>SlHsp27.2</i>	104644273
<i>SlyHsp36.8-MCIII</i>	<i>SlHsp37.0</i>	101244182
<i>AthHsp28.6-MCIII</i>	<i>AthAcd28.7</i>	844011
<i>AthHsp41.2-MCIII</i>	<i>AthAcd41.3</i>	830370
<i>NatHsp17.8-MCIV</i>		109241832
<i>NatHsp17.3-MCIV</i>		109241846
<i>NatHsp18.8-MCIV</i>		109210789
<i>NatHsp26.8-MCIV</i>		109233836
<i>NatHsp27.4-MCIV</i>		109239196
<i>NatHsp28.1A-MCIV</i>		109239261
<i>NatHsp28.1B-MCIV</i>		109226300
<i>NatHsp40.5-MCIV</i>		109219931
<i>SlyHsp15.5-MCIV</i>	<i>SlHsp15.5</i>	101247576
<i>SlyHsp23.7-MCIV</i>	<i>SlHsp23.8A</i>	101249849
<i>SlyHsp24.5-MCIV</i>	у нас	101250148
<i>SlyHsp25.6-MCIV</i>	<i>SlHsp25.7A</i>	101250437
<i>SlyHsp26.8-MCIV</i>	<i>SlHsp26.8</i>	101262846
<i>SlyHsp27.4-MCIV</i>	<i>SlHsp27.5</i>	101258622
<i>AthHsp20.2-MCIV</i>	<i>AthAcd16.6</i>	841882
<i>AthHsp25.0-MCIV</i>	<i>AthAcd25.1</i>	817256
<i>AthHsp28.0-MCIV</i>	<i>AthAcd28.1</i>	832222

Примітка. Ch – хромосома, NA – не анотовано, Sc – скефолд.

зування транскрипційних факторів із родини HSF (heat shock factors). Для цього використовували об'єднаний набір PWM-матриць для 23 HSE-мотивів *Arabidopsis thaliana* із бази даних JASPAR (https://jaspar.elixir.no/search?q=&collection=CORE&tax_group=plants). Результати обох пошуків були інтегровані та візуалізовані у вигляді теплової карти у програмі TBtools.

Результати та обговорення. Ідентифікація генів *sHsp* та характеристика кодованих ними білків. Для ідентифікації генів *sHsp* *N. attenuata*, *S. lycopersicum* та *A. thaliana* було проведено скринінг геномів, які наявні у базі даних GenBank. Мультигенні родини *sHsp* арабідоп-

сису та томату були описані раніше (Scharf et al., 2001; Siddique et al., 2008; Yu et al., 2016). Проте, беручи до уваги, що з того часу було здійснено нові зборки геномів *A. thaliana* та *S. lycopersicum*, ми провели пошук генів *sHsp* у геномах цих видів з використанням останніх доступних варіантів зборки.

Загалом для *N. attenuata*, *S. lycopersicum* та *A. thaliana* нами було ідентифіковано, відповідно, 52, 45 та 34 гени, які кодують білки, що містять ACD та мають молекулярну масу від 7,8 до 41,2 кДа (табл. 1 та 2). Для *S. lycopersicum* та *A. thaliana* було виявлено всі гени *sHsp*, описані раніше (Yu et al., 2016; Scharf et al.,

Геномна збірка 1			Геномна збірка 2		
Ch/Sc	№ доступу у GenBank	Розташування, нп	Ch/Sc	№ доступу у GenBank	Розташування, нп
Ch3	NC_003074.8	7977604-7978635	Ch1		343283170-343285370
Sc	NW_017671144.1	35040-37240			
Ch4	NC_090803.1	5385313-5386777			
Ch5	NC_003076.8	14968758-14969591			
Ch12	NC_090811.1	67053622-67054314			
Ch1	NC_090800.1	30716117-30720431			
Ch1	NC_090800.1	26884566-26888873			
Ch1	NC_090800.1	28761025-28769720			
Ch2	NC_090801.1	70737945-70738601			
Ch7	NC_031995.1	109310992-109314217			
Ch12	NC_090811.1	64706359-64707265	Sc	JAPQWU010000015.1	14249903-14251611
Ch4	NC_090803.1	60094682-60096139			
Ch1	NC_003070.9	28813454-28814378			
Ch5	NC_003076.8	1427117-1428514			
Ch1	NC_031989.1	99397287-99400200			
Ch1	NC_031989.1	99440245-99441951			
Sc	NW_017671896.1	139341-140255			
Ch10	NC_031998.1	11709039-11710535			
Sc	NW_017670421.1	505545-507171			
Ch1	NC_031989.1	104370393-104372344	Ch10	CM061571.1	7759961-7762546
Ch1	NC_031989.1	50361092-50363062	Ch10	CM061571.1	7802993-7804368
Ch1	NC_031989.1	24820824-24824250	Ch8	CM061569.1	56294703-56295617
Ch10	NC_090809.1	58928640-58929663	Ch3	CM061564.1	32395570-32397066
Ch1	NC_090800.1	3616000-3617694	Ch6	CM061567.1	6057102-6058728
Ch1	NC_090800.1	3613417-3614625	Ch3	CM061567.1	232943189-232945140
Ch1	NC_090800.1	3610188-3611570	Ch3	CM061567.1	233570022-233571991
Ch9	NC_090808.1	1256737-1258130	Ch3	CM061567.1	228062878-228066307
Ch11	NC_090810.1	60433217-60434946			
Ch1	NC_003070.9	20307572-20308846			
Ch2	NC_003071.7	11598305-11599322			
Ch5	NC_003076.8	7122800-7124001			

2001; Siddique et al., 2008), а також ідентифіковано, відповідно, п'ять (*SlyHsp17.7-CI*, *SlyHsp23.7-CVIII*, *SlyHsp16.6-NPI*, *SlyHsp21.8-NPI*, *SlyHsp24.5-MCIV*) та один (*AthHsp21.6-MPI*) нові гени. Довжина передбачуваних *sHSP* варіює від 67 до 366 АК, а розрахована ізоелектрична точка (pI) становить від 4.28 до 9.97.

Для ідентифікованих нами білків було визначено їх можливу локалізацію у клітині. Встановлено, що вони можуть транспортуватись у різні компартменти, а саме — у цитозоль, ядро, ендоплазматичний ретикулум, апарат Гольджі, вакуоль, пероксисоми, мітохондрії та пластиди. Для багатьох білків передбачається локалізація у двох або більше компартментах (табл. 2).

Філогенетична подібність *sHSP*. Із використанням амінокислотних послідовностей ACD було побудовано ML-дендрограму (рис. 1), яка відображає подібність між *sHSP N. attenuata*, *S. lycopersicum* та *A. thaliana*, а також *S. tuberosum* (Zhao et al., 2018). Аналіз дендрограми свідчить, що всі білки можна поділити на 19 класів, яким були присвоєні назви з урахуванням їх клітинної локалізації. При цьому було збережено традиційні назви класів CI-CVI (cytoplasm/nucleus), ER (endoplasmatic reticulum), MPI, MPII (mitochondria, plastids), PI (plastids) та Po (peroxisomes), запропоновані для *sHSP A. thaliana* (Scharf et al., 2001; Siddique et al., 2008). Білок *AthHSP18.5-CI(IV)* із вкороченим ACD,

Таблиця 2. Характеристика білків sHSP *Nicotiana attenuata*, *Solanum lycopersicum* та *Arabidopsis thaliana*

Назва гена	Кількість екзонів	Розмір білка, АК	М. м., кДа	pI	Розмір, АК			Клітинна локалізація	
					N	ACD	C	WoLF	PISP
<i>NatHsp16.0-CI</i>	1	141	16	5,62	32	94	15	Ц	Ц
<i>NatHsp17.2-CI</i>	1	150	17,2	6,43	45	90	15	Ц	Ц
<i>NatHsp17.4A-CI</i>	1	153	17,4	5,56	48	90	15	Ц	Ц
<i>NatHsp17.4B-CI</i>	1	153	17,4	6,43	48	90	15	Ц	Ц
<i>NatHsp17.4C-CI</i>	1	153	17,4	6,43	46	92	15	Ц	Я
<i>NatHsp17.5-CI</i>	1	153	17,5	6,43	48	90	15	Ц	Ц
<i>NatHsp17.9A-CI</i>	2	159	17,9	5,38	54	90	15	Ц	Ц
<i>NatHsp17.9B-CI</i>	1	159	17,9	5,94	54	90	15	Ц	Ц
<i>NatHsp17.9C-CI</i>	1	159	17,9	5,95	54	90	15	Ц	Ц
<i>NatHsp17.9D-CI</i>	1	159	17,9	7,41	54	90	15	Ц	Я
<i>NatHsp18.0A-CI</i>	1	159	18	5,84	54	90	15	Ц	Я
<i>NatHsp18.0B-CI</i>	1	159	18	5,61	54	90	15	Ц	Ц
<i>NatHsp18.8-CI</i>	1	163	18,8	7,18	58	90	15	Ц	Я
<i>NatHsp20.6-CI</i>	2	176	20,6	7,84	70	90	16	Ц	Я
<i>NatHsp22.2-CI</i>	1	195	22,2	9,97	92	89	14	М	Пл
<i>SlyHsp7.8-CI</i>	1	67	7,8	4,69	11	55	1	Ц	Я
<i>SlyHsp9.1-CI</i>	2	78	9,1	5,17	18	59	1	Ц	Я
<i>SlyHsp17.5A-CI</i>	1	154	17,5	5,57	49	90	15	Ц	Ц
<i>SlyHsp17.5B-CI</i>	1	154	17,5	5,91	49	90	15	Ц	Я
<i>SlyHsp17.5C-CI</i>	1	154	17,5	5,93	59	90	15	Ц	Я
<i>SlyHsp17.6A-CI</i>	1	154	17,6	5,93	49	90	15	Ц	Я
<i>SlyHsp17.6B-CI</i>	1	156	17,6	8,32	50	90	16	Ц	Ц
<i>SlyHsp17.7-CI</i>	1	154	17,5	5,91	49	90	15	Ц	Ц
<i>SlyHsp24.4-CI</i>	2	208	24,4	7,52	104	88	16	Ц	Я
<i>SlyHsp25.4-CI</i>	1	222	25,4	6,53	116	90	16	Ц	Я
<i>SlyHsp27.0-CI</i>	1	234	27	9,49	129	90	15	М	Пл
<i>AthHsp17.3-CI</i>	1	156	17,3	5,15	51	90	15	Ц	Я
<i>AthHsp17.5A-CI</i>	1	157	17,5	5,32	52	90	15	Ц	Ц
<i>AthHsp17.5B-CI</i>	1	155	17,5	7,39	48	90	17	Ц	Ц
<i>AthHsp17.4-CI</i>	1	153	17,4	6,64	48	90	15	Ц	Ц
<i>AthHsp17.7-CI</i>	1	157	17,7	6,18	50	90	17	Ц	Ц
<i>AthHsp18.0-CI</i>	1	161	18	7,31	54	90	17	Ц	Ц
<i>AthHsp18.4-CI(CIV)</i>	1	162	18,4	4,28	62	72	28	Ц, Я	Я
<i>NatHsp15.2-CII</i>	1	138	15,2	9,82	31	90	17	Ц	Ц
<i>NatHsp17.4A-CII</i>	1	157	17,4	6,4	50	88	19	Ц	Ц
<i>NatHsp17.4B-CII</i>	1	157	17,4	6,4	50	88	19	Ц	Ц
<i>NatHsp17.5A-CII</i>	2	157	17,5	6,77	50	90	17	Ц	Ц
<i>NatHsp17.5B-CII</i>	1	157	17,5	6,39	50	90	17	Ц	Ц
<i>NatHsp17.5C-CII</i>	1	157	17,5	5,83	50	90	17	Ц	Ц
<i>NatHsp18.1-CII</i>	1	160	18,1	7,04	50	93	17	Ц	Ц
<i>NatHsp21.1-CII</i>	1	192	21,1	5,34	52	90	50	Ц	Я
<i>NatHsp30.6-CII</i>	2	271	30,6	8,82	164	90	17	Я	Я
<i>SlyHsp17.2-CII</i>	1	155	17,2	7,24	48	90	17	Ц	Ц
<i>SlyHsp17.5-CII</i>	1	158	17,5	5,57	51	90	17	Ц	Ц
<i>AthHsp17.5-CII</i>	1	155	17,5	6,63	47	91	17	Ц	Ц
<i>AthHsp17.6-CII</i>	1	156	17,6	5,56	48	91	17	Ц	Ц
<i>NatHsp16.8-CIII</i>	3	150	16,8	7,84	41	94	15	ПМ	Я
<i>SlyHsp16.0-CIII</i>	2	144	16	8,24	35	94	15	Я, ПМ	Я

Назва гена	Кількість екзонів	Розмір білка, АК	М. м., кДа	pI	Розмір, АК			Клітинна локалізація	
					N	ACD	C	WoLF	PISP
<i>AthHsp17.2-CIII</i>	3	155	17,2	8,31	44	95	16	Я, Пл	Пл
<i>SlyHsp15.6-CV</i>	2	137	15,6	4,84	31	80	26	Ц	Пл
<i>AthHsp15.3-CV</i>	2	134	15,3	5,1	28	82	24	Я	Я
<i>NatHsp21.0-CVI</i>	2	184	21	5,05	88	88	8	Пл	Я
<i>SlyHsp21.4-CVI</i>	2	188	21,4	5,73	86	93	9	Ц	Ц
<i>AthHsp21.5-CVI</i>	2	192	21,5	5,44	89	96	7	Пер	Ц
<i>NatHsp18.4-CVII</i>	2	164	18,4	5,26	58	90	16	Ц	Я
<i>NatHsp19.2-CVII</i>	2	172	19,2	5,23	66	90	16	Ц	Ц
<i>SlyHsp17.8-CVIII</i>	2	163	17,8	5,48	58	89	16	Ц	Ц
<i>NatHsp24.4-CVIII</i>	3	216	24,3	4,94	28	79	109	Ц, Г	Я
		217	24,4	4,89	29	79	109	Ц, Г	Я
<i>SlyHsp23.6-CVIII</i>	2	209	23,6	4,89	28	79	102	Ц	Ц
<i>SlyHsp23.7-CVIII</i>	2	209	23,7	5,07	22	79	108	Г, Ц	Ц
<i>NatHsp21.6A-ER</i>	1	192	21,6	5,57	69	90	33	ПМ, В	ЕР
<i>NatHsp21.6B-ER</i>	1	190	21,6	6,08	69	90	31	ПМ	ЕР
<i>NatHsp22.4-ER</i>	1	198	22,4	5,42	75	92	31	В	Пл
<i>SlyHsp19.1-ER</i>	2	190	19,1	5,84	69	92	29	ПМ, В	Пл
<i>SlyHsp21.4-ER</i>	1	188	21,4	7,4	68	90	20	Пл, В	Пл
<i>SlyHsp21.5-ER</i>	1	189	21,5	8,27	69	91	29	В, ПМ	Пл
<i>AthHsp21.9-ER</i>	1	195	21,9	5,61	74	90	31	Пл	Пл
<i>NatHsp18.2-MPI</i>	2	169	18,2	5,56	81	82	6	М	М
<i>NatHsp21.4-MPI</i>	2	192	21,4	7,15	94	82	16	Пл	М
<i>NatHsp24.3-MPI</i>	2	216	24,3	5,91	117	84	15	Ц, М	Пл
<i>SlyHsp18.3-MPI</i>	2	170	18,3	4,98	82	82	6	М, Пл	М
<i>SlyHsp21.4-MPI</i>	2	196	21,4	8,24	97	82	17	М, Пл	М
<i>SlyHsp23.7-MPI</i>	2	210	23,7	6,82	111	84	15	М, Пл	Пл
<i>AthHsp14.6-MPI</i>	2	131	14,6	8,41	28	82	17	Пл, Ц	Пл
<i>AthHsp21.6-MPI</i>	3	193	21,6	9,24	37	84	72	Пл	Г
<i>AthHsp23.3-MPI</i>	2	210	23,3	9,01	109	86	15	М, Пл	Пл
<i>AthHsp23.5-MPI</i>	2	210	23,5	6,84	107	88	15	Пл, М	М, Пл
<i>AthHsp25.2-MPI</i>	6	167	18,2	7,21	139	28	0	Ц	Пл
	6	216	23,3	8,08	139	76	1	Я, Ц	Пл
	6	232	25,2	8,65	147	84	1	Пл, Я	Пл
<i>AthHsp27.6-MPI</i>	6	247	26,7	8,99	162	84	1	Я, Ц	Пл
	6	255	27,6	8,79	170	84	1	Я, Пл	Пл
<i>AthHsp28.8-MPI</i>	3	193	21,6	9,24	37	84	72	Пл	Г
	3	264	28,8	6,18	37	84	143	Пл	Ц
<i>NatHsp25.0-MPII</i>	3	222	25	9,52	126	95	1	Пл, М	Пл
<i>NatHsp26.6-MPII</i>	2	234	26,6	8,62	130	90	14	Пл, М	Пл
<i>SlyHsp26.2-MPII</i>	2	231	26,2	8,89	127	90	14	Пл, Я	Пл
<i>AthHsp26.6-MPII</i>	2	225	25,6	7,92	120	91	14	Пл, Я	Пл
		232	26,4	7,28	127	91	14	Пл	Пл
<i>NatHsp21.9-NPI</i>	4	148	16,2	8,69	46	99	3	Пл, Я	Пл
		156	17,1	8,68	54	99	3	Пл, Я	Пл
		200	21,9	8,31	98	99	3	Я	Пл
		200	21,9	8,31	98	99	3	Я	Пл
<i>NatHsp16.7-NPI</i>	3	154	16,7	7,72	34	97	23	Пл	Пл
<i>SlyHsp16.6-NPI</i>	3	153	16,6	7,15	33	97	23	Я, Пл, Ц	Пл

Назва гена	Кількість екзонів	Розмір білка, АК	М. м., кДа	pI	Розмір, АК			Клітинна локалізація	
					N	ACD	C	WoLF	PISP
<i>SlyHsp21.8-NPI</i>	5	200	21,8	9,24	98	99	3	Я, Ц	Пл
<i>AthHsp15.4-NPI</i>	3	143	15,4	7,71	37	98	8	Пл	Пл
<i>AthHsp22.2-NPI</i>	4	206	22,2	4,8	93	99	5	Я	Пл
<i>AthHsp39.3-NPI</i>	3	268	30,3	5,52	243	25	0	Ц, Я	Я
	3	349	39,3	6,28	239	99	7	Я	Я
<i>NatHsp29.8-NPII</i>	7	258	28,4	9,1	162	95	1	Пл	Я
	7	268	29,8	9,1	162	104	2	Пл	Я
<i>SlyHsp26.4-NPII</i>	6	238	26,4	9,35	142	95	1	Я, Пл	Я
<i>AthHsp31.1-NPII</i>	6	285	31,1	9,34	188	96	1	Пл	Пл
<i>NatHsp25.5-PI</i>	2	221	25,5	8,84	120	88	13	Пл	Пл
<i>NatHsp26.6-PI</i>	2	242	26,6	7,12	141	88	13	Пл	Пл
<i>SlyHsp26.0-PI</i>	3	221	25,6	9,32	119	89	13	Пл	Пл
	3	225	26	9,32	123	89	13	Пл	Пл
<i>SlyHsp26.1-PI</i>	2	235	26,1	8,1	133	89	13	Ц	Пл
<i>AthHsp25.2-PI</i>	2	227	25,2	8,38	127	87	13	Пл	Пл
<i>AthHsp16.8-PPI</i>	1	152	16,8	5,59	49	90	13	Пл, Ц	Пл
<i>AthHsp22.0-PPI</i>	1	198	22	5,72	80	91	27	Пл, М	Пл
<i>NatHsp15.6-Po</i>	2	143	15,6	8,88	28	97	18	Пер	Пл
<i>SlyHsp15.9-Po</i>	2	145	15,9	7,45	28	99	18	Пер	Ц
<i>AthHsp15.6-Po</i>	1	137	15,6	8,36	24	95	18	Пер	Пер
<i>SlyHsp11.9-MCI</i>	4	107	11,9	6,25	15	87	5	Ц	Пл
<i>SlyHsp15.8-MCI</i>	5	142	15,8	5,1	50	87	5	РМ	КМ
<i>SlyHsp17.3-MCI</i>	5	153	17,3	8,99	61	87	5	Пл	Пл
<i>SlyHsp26.0-MCI</i>	7	233	26	4,89	141	87	5	Ц, Я	ПМ
<i>SlyHsp15.5-MCII</i>	2	136	15,5	7,76	25	95	16	Пл, ПМ	Я
<i>NatHsp38.1-MCIII</i>	4	302	34,6	5,93	33	94	175	Ц	Ц
	4	313	35,8	6,16	44	94	175	Г, Пл	Я
	4	333	38,1	6,16	64	94	175	ПМ, В	ЕР
<i>SlyHsp27.7-MCIII</i>	2	245	27,7	8,48	33	99	113	Ц	Ц
<i>SlyHsp36.8-MCIII</i>	2	325	36,8	6	31	94	200	Ц	Ц
<i>AthHsp28.6-MCIII</i>	2	244	28,6	4,98	33	94	117	Ц	ЕР
<i>AthHsp41.2-MCIII</i>	2	366	41,2	8,83	23	80	263	Г, Ц	Пл
<i>NatHsp17.8-MCIV</i>	3	138	15,2	9,82	3	70	65	Пл	Пл
	3	160	17,8	9,06	17	78	65	Пл, ПМ	Ц
<i>NatHsp17.3-MCIV</i>	2	154	17,3	9,04	16	80	58	Ц	Ц
<i>NatHsp18.8-MCIV</i>	2	166	18,8	9,19	17	80	69	Ц, Г	Пл
<i>NatHsp26.8-MCIV</i>	2	240	26,8	9,39	22	79	139	Г, Я	Я
<i>NatHsp27.4-MCIV</i>	2	247	27,4	6,18	19	79	149	Г, Ц	Я
<i>NatHsp28.1A-MCIV</i>	2	254	28,1	5,9	22	79	143	Г, Пл	Я
<i>NatHsp28.1B-MCIV</i>	2	254	28,1	5,31	22	79	143	Г, Пл	Я
<i>NatHsp40.5-MCIV</i>	3	366	40,5	8,34	19	78	229	Г, РМ	Я
<i>SlyHsp15.5-MCIV</i>	2	139	15,5	9,16	13	79	47	Ц	Пл
<i>SlyHsp23.7-MCIV</i>	2	213	23,7	9,44	25	78	120	Ц, Г	Я
<i>SlyHsp24.5-MCIV</i>	2	222	24,5	8,72	16	80	126	Г	Я
<i>SlyHsp25.6-MCIV</i>	2	232	25,6	6,06	19	79	134	Г, В	Я
<i>SlyHsp26.8-MCIV</i>	2	236	26,8	4,96	19	79	138	Г, Ц	Я
<i>SlyHsp27.4-MCIV</i>	2	247	27,4	7,27	19	79	149	Г, В	Я
<i>AthHsp20.2-MCIV</i>	2	183	20,2	9,43	20	80	83	Ц	Ц

Назва гена	Кількість екзонів	Розмір білка, АК	М. м., кДа	pI	Розмір, АК			Клітинна локалізація	
					N	ACD	C	WoLF	PISP
<i>AthHsp25.0-MCIV</i>	2	224	25	9,31	21	79	124	Ц, Г	Я
<i>AthHsp28.0-MCIV</i>	2	249	28	9,6	17	79	153	Г, В	Я

Примітка. М. м. – молекулярна маса; ACD, N та C – розміри ACD, N та C; В – вакуоль, Г – апарат Гольджі, ЕР – ендоплазматичний ретикулум, КМ – клітинна мембрана, М – мітохондрії, Пер – пероксисома, Пл – пластиди, ПМ – позаклітинний матрикс, Ц – цитозоль, Я – ядро.

який раніше відносили до класу CIV, за нашими даними належить до класу CI. Крім того, на дендрограмі візуалізуються дев'ять класів білків, які не було описано раніше – CVII, CVIII, NPI, NPII (nucleus, plastids), MCI-IV (multiple compartments) та PII. В кількох випадках передбачувана локалізація білків виявилась іншою, ніж вважалось попередньо. Особливо це стосується білків класів CIII та ER, які окрім цитоплазми/ядра та ендоплазматичного ретикулуму вірогідно транспортуються у пластиди та вакуолі (табл. 2).

Білки, які належать до різних класів відрізняються між собою за розмірами ACD та N- та C-кінцевих ділянок. Розмір N-кінцевої ділянки знаходиться у межах від 3 (*NatHsp17.8-MCIV*) до 243 (*AthHsp39.3-NPI*) АК, C-кінцевої – від 0 (*AthHsp39.3-NPI*) до 263 (*AthHsp41.2-MCIII*) АК, а ACD від 25 (*AthHsp39.3-NPI*) до 104 (*NatHsp29.8-NPII*) АК – (табл. 2). При цьому, білки, які належать до одного класу, переважно подібні між собою за розмірами N- та C-кінцевих ділянок.

Більшість класів sHSP на дендрограмі охоплюють білки всіх чотирьох досліджуваних видів, що вказує на існування цих класів до дивергенції філогенетичних ліній Суперрозидів (Brassicaceae: *Arabidopsis*) та Суперастеридів (Solanaceae: *Nicotiana*, *Solanum*) (Angiosperm Phylogeny Group, 2016) та їх збереження упродовж подальшої еволюції дводольних. Проте, є декілька класів, які специфічні лише для *A. thaliana* (PII) або представників родини Solanaceae (CVII, CVIII, MCI, MCII).

Найбільш численними у всіх досліджених видів виявились sHSP класу CI, що є характерним і для багатьох інших рослин. Проте,

особливістю *N. attenuata* виявилась наявність значної кількості (9) sHSP класу CII порівняно з *S. lycopersicum* (2) та *S. tuberosum* (2).

Аналіз структури транскриптів показав, що всі гени класу PII та більшість генів класів CI, CII та ER не містять інтронів, тоді як гени інших класів як правило складаються із кількох екзонів (рис. 2). Зокрема, по сім екзонів знайдені у генах *NatHsp29.8-NPII* та *SlyHsp26.0-MCI*. По п'ять чи шість екзонів виявлено у багатьох генів класів MPI, NPI, NPII та MCI.

Також встановлено, що внаслідок альтернативного сплайсингу деякі гени можуть кодувати ізоформи білка, що відрізняються між собою розміром N- (*NatHsp24.4-CVIII*, *AthHsp27.6-MPI* *AthHsp26.6-MPII*, *NatHsp21.9-NPI* *Sly26.0-PI*, *NatHsp38.1-MCIII*) та C- (*AthHsp28.8-MPI*) кінцевої ділянок, ACD (*NatHsp29.8-NPII*) або кількох сусідніх ділянок одночасно (*AthHsp25.2-MPI* *AthHsp39.3-NPI*).

У геномах досліджених видів виявлено 15 sHsp, які кодують від двох до чотирьох альтернативних варіантів РНК (табл. 1, рис. 2). У чотирьох випадках ці альтернативні транскрипти кодують лише один варіант білка, а в 11 випадках – дві або три ізоформи, які можуть мати різну клітинну локалізацію. Особливо часто альтернативні транскрипти зустрічаються у класах NPI та MPI.

Хромосомна локалізація генів sHsp була визначена для *N. attenuata* та *S. lycopersicum*. За результатами картування встановлено, що у *N. attenuata* гени sHsp присутні на 10 з 12 хромосом, крім хромосом 5 та 12 (табл. 1, рис. 3). Найбільша кількість генів – 11, 10, 7 та 6 – виявлена на хромосомах 9, 1, 4 та 3, відповідно. На хромосомах 1–4 та 9–11 по декілька

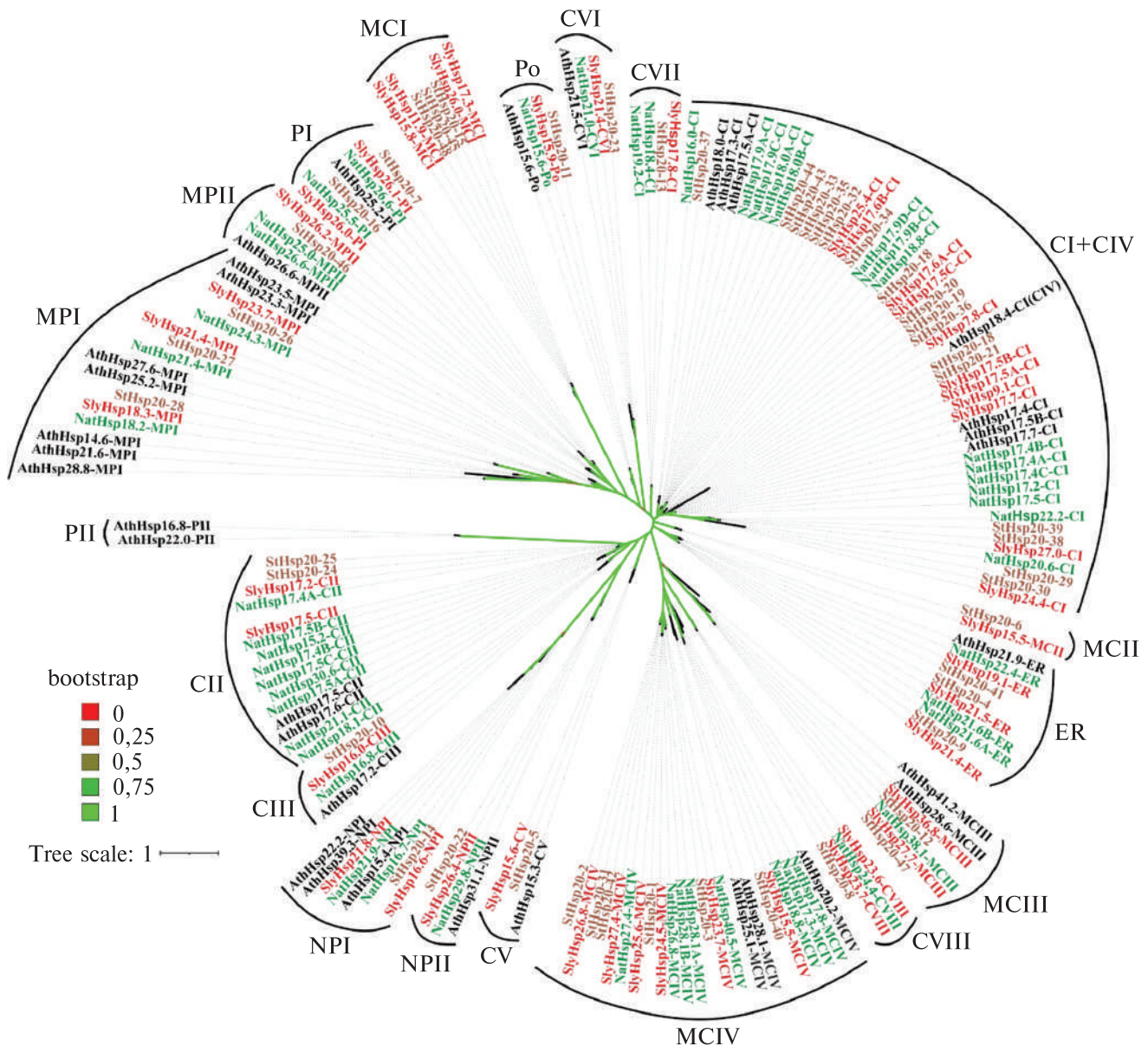


Рис. 1. Філогенетична дендрограма, отримана методом Maximum likelihood при порівнянні послідовностей ACD генів *sHsp* *Nicotiana attenuata*, *Solanum lycopersicum*, *Solanum tuberosum* та *Arabidopsis thaliana*. Колір гілок на дереві відповідає значенням bootstrap підтримки відповідних груп

високоподібних генів розташовані у безпосередній близькості один до одного, утворюючи кластери. Найбільший кластер розміром 336 kb знаходиться на хромосомі 9 та охоплює вісім з дев'яти генів *sHsp* класу CII, які виявлено у геномі *N. attenuata*. Ще один ген цього класу, *NatHsp17.5A-CII*, знаходиться на хромосомі 7.

Другий за кількістю генів кластер, розташований на кінці хромосомі 1, містить шість з

15 *sHsp* класу CI. Цей кластер складається з двох частин, які мають розмір 23.3 та 26.0 kb і містять два гени (*NatHsp17.4A-CI*, *NatHsp17.4C-CI*) у антисенс-орієнтації та чотири гени (*NatHsp17.2-CI*, *NatHsp17.4B-CI*, *NatHsp17.5-CI*, *NatHsp18.8A-CI*) у сенс-орієнтації, відповідно. Дві частини кластера розділені ділянкою хромосоми, яка має розмір 101.7 kb та не містить *sHsp* генів. Найпростішим поясненням

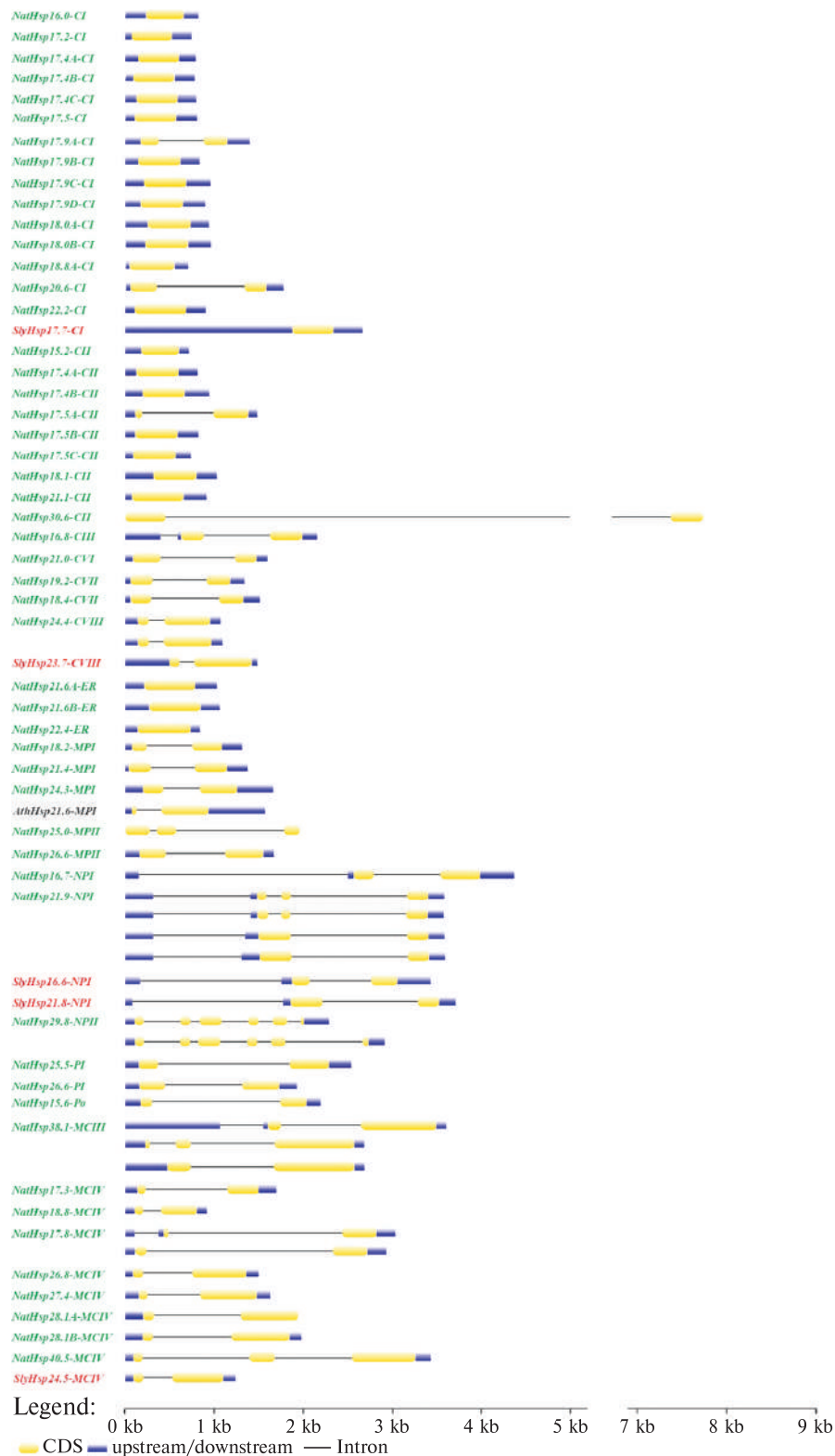


Рис. 2. Організація екзонів/інтронів ідентифікованих у цій роботі генів *sHsp* *Nicotiana attenuata*, *Solanum lycopersicum* та *Arabidopsis thaliana*. Жовті прямокутники представляють екзони, а чорні лінії — інтрони. Нетрансльовані ділянки на 5' та 3' кінцях транскриптів позначені синіми прямокутниками

такої будови кластеру є те, що попередньо всі шість генів виникли внаслідок кількох тандемних дуплікацій спільної предкової послідовності та утворили єдиний кластер. Пізніше внаслідок інверсії ділянки хромосоми два гени змінили орієнтацію, а кластер розділювався на дві частини. Решта дев'ять генів класу СІ знаходяться на хромосомах 1, 2, 4, 6 та 8 і частково представлені дуплікованими копіями.

Кластер на хромосомі 3 включає чотири з восьми генів класу МСІV, з яких три гени (*NatHsp26.8B-MCIV*, *NatHsp28.1A-MCIV*, *NatHsp40.5-MCIV*) розташовані у сенс-, а один (*NatHsp28.1B-MCIV*) — у антисенс-орієнтації. Інші гени цього класу розміщені на хромосомах 6, 8 та 10. Решта шість *sHsp*-кластерів у геномі *N. attenuata* містять лише по 2 або 3 (хромосома 9) гени різних класів.

У геномі *S. lycopersicum* 45 передбачуваних генів *sHsp* знайдено на всіх 12 хромосомах (рис. 3). На хромосомі 1 розташовані 6 генів, а на хромосомах 3, 4, 6, 8 та 9 — по 5 генів. На хромосомі 6 всі п'ять генів належать до класу СІ та утворюють кластер розміром 16,5 kb, в межах якого розташований лише один ген, який кодує білок, що не належить до мультигенної родини *sHsp*. Кластери на інших хромосомах утворені трьома або двома генами.

У попередніх дослідженнях для *S. lycopersicum* було виявлено 42 *sHsp* гени, для 41 з яких визначено хромосомну локалізацію (Yu et al., 2016). Оскільки два гени (*SlHsp39.4* та *SlHsp49.3*), які виникли внаслідок тандемної дуплікації, кодують білки великої молекулярної маси, їх було виключено із нашого подальшого аналізу. Решта 39 генів були виявлені нами на тих самих хромосомах, що і в роботі Yu et al. (2016), але їх розташування у декількох випадках суттєво відрізняється від попереднього. Також з використанням нової геномної збірки SLM_r2.1 ми визначили розташування п'яти нових, ідентифікованих нами генів, а також гена *SlyHsp11.9-MCI*, локалізація якого була невідома.

Аналіз отриманих нами результатів показує, що кластерна організація є характерною рисою генів *sHsp* у геномах представників родини Solanaceae. Раніше про кластерну організацію повідомлялось і для рослин, які належать до різних таксономічних груп, але найчастіше

число генів в кластері не перевищує чотирьох (Krsticevic et al., 2016; Chen et al., 2021; Zhang et al., 2023; Huang et al., 2025), тоді як у геномі *N. attenuata* нами виявлено кластер, що містить вісім *sHsp* класу СІІ.

Біоінформатичний аналіз промоторних ділянок генів *sHsp*. Для виявлення сигнальних шляхів, які можуть брати участь у регуляції експресії *sHsp* ми проаналізували промоторні ділянки 57 генів, які було вперше ідентифіковано нами в цій роботі (рис. 4). Було виявлено імовірні регуляторні цис-елементи, які можна розділити на чотири групи, задіяні у (1) відповідь на тепловий шок (Heat Shock Elements, HSE) та інші види стресу, (2) відповідь на дію світла, (3) передачу сигналів фітогормонів та (4) регуляцію росту і розвитку.

HSE виявились найпоширенішим елементом: вони зустрічаються у проаналізованих промоторних ділянках у кількості від одного у *SlyHsp23.7-CVIII* до 14 у *NatHsp18.0A-CI* та *NatHsp25.5-PI*. У промоторі гена *NatHsp18.4-CVII* HSE відсутні. HSE є сайтом впізнавання транскрипційних факторів теплового шоку HSF, які активуються за умов високотемпературного стресу та індукують експресію *sHSP* та інших захисних білків (Scharf et al., 1998; Scharf et al., 2001; Panchuk et al., 2002; Volkov et al., 2006). Отже, отримані дані вказують, що майже всі досліджувані гени можуть бути задіяні у відповідь на тепловий стрес. Раніше активація експресії 11 генів *sHsp* була продемонстрована нами за дії підвищеної температури на *N. tabacum* (Volkov et al., 2005).

Також у досліджених промоторах знайдено сайти впізнавання для інших транскрипційних факторів, зокрема MYB, MYC, WRKY та ін., які у рослин беруть участь у відповіді на абіотичний (посуха, низькі температури, гіпоксія) та біотичний стреси, механічні пошкодження, а також у гормональному сигналінгу, регуляції вторинного метаболізму, спокою та проростання насіння, росту та розвитку (Dombrecht et al., 2007; Seo et al., 2009; Rushton et al., 2011; Kazan and Manners, 2013; Cao et al., 2013). Цікаво, що промотор одного із досліджуваних генів — *NatHsp17.2-CI* — помітно виділяється серед всіх решта наявністю 14 сайтів ARE, які задіяні у регуляції відповіді на гіпоксію (Olive et al., 1991).

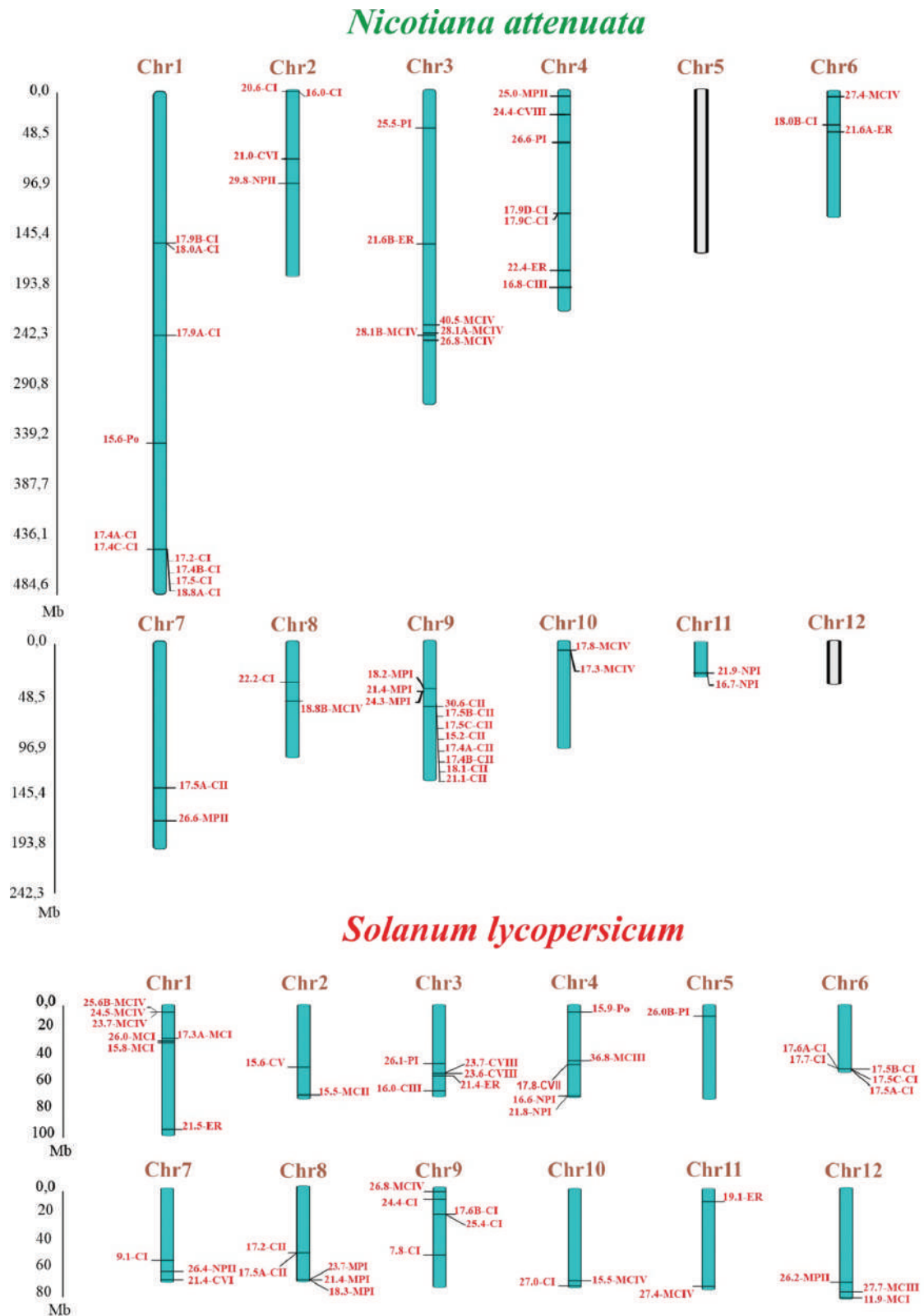


Рис. 3. Розташування генів *sHsp* на хромосомах *Nicotiana attenuata* та *Solanum lycopersicum*. Справа від хромосом наведено назви генів, які розташовані у сенс-, а зліва – у антисенс орієнтації

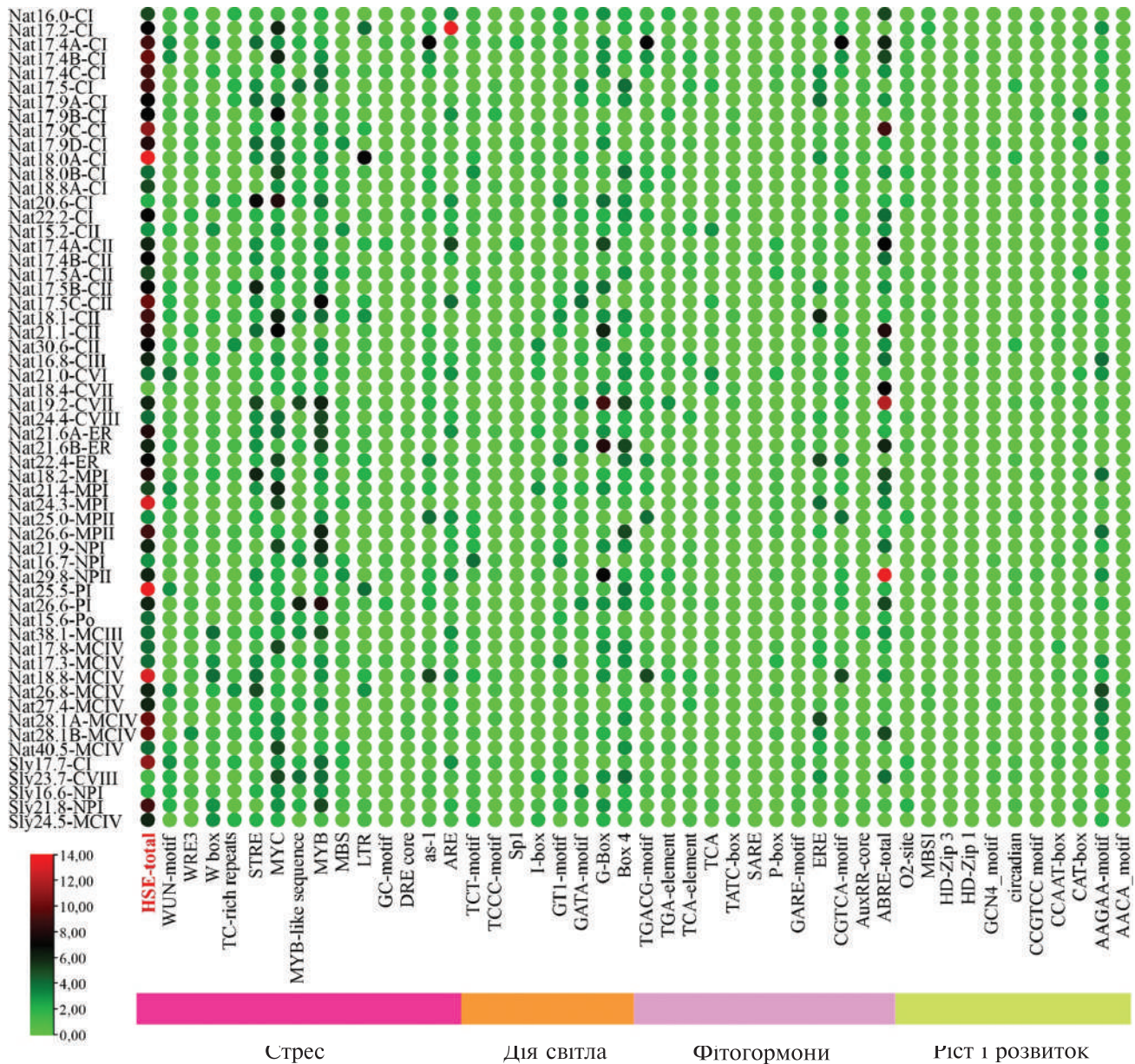


Рис. 4. Цис-елементи у промоторних ділянках генів *sHsp Nicotiana attenuata* та *Solanum lycopersicum*

До другої групи належать такі промоторні елементи, як G- та I-бокси, GATA-, GT1- та TCT-мотиви, що є сайтами впізнавання для транскрипційних факторів bZIP, MYB, GATA та GT1. Ці елементи беруть участь у регуляції експресії генів у відповідь на зміни освітлення та різні стреси, такі як високі та низькі температури, посуха та патогени, а також задіяні у регуляції росту та розвитку.

Третю групу складають промоторні елементи, які задіяні у передачі сигналів фіто-

гормонів, як от абсцизової кислоти (ABRE-елементи), ауксинів (AuxRR-core та TCA-елемент), гіберелінів (GARE-мотив, P- та TATC-бокси), етилену (ERE-елемент), метил-жасмонату (TGACG/CGTCA-мотиви) та саліцилату (SARE- та TCA-елементи) (Narusaka et al., 2003; Shariatipour and Heidari, 2018). Особливо часто у промоторах знайдено ділянки, які пов'язані із дією абсцизової кислоти (гени *NatHsp29.8-NPII*, *NatHsp19.2-CVII*, *NatHsp17.9C-CI*, *NatHsp21.1-CII*, *NatHsp17.4A-CI*, *NatHsp17.4A-CII*), ме-

тил-жасмонату (*NatHsp17.4A-CI*, *NatHsp18.8B-CI*, *NatHsp25.0-MPII*) та етилену (*NatHsp18.1-CII*, *NatHsp22.4-ER*, *NatHsp28.1A-MCIV*).

Елементи промотора, які належать до четвертої групи, зокрема AACA_motif, AAGAA_motif, CAT-box, CCAAT-box, CCGTCC_motif, circadian, GCN4_motif, HD-Zip 1, HD-Zip 3, MBSI, O2-site, асоційовані із широким спектром процесів – ростом та розвитком, диференціацією клітин, контролем добових ритмів тощо. Серед цих елементів у досліджених промоторах найбільш широко представлений мотив AAGAA, пов'язаний із розвитком вторинної ксилеми (Ain-Ali et al., 2021; Maqsood et al., 2022). Наявність CAT-боксу і CCGTCC-мотиву у промоторах свідчить про потенційну експресію генів *sHSP* у меристемах. Це вказує, що підсилена експресія *sHSP* може бути необхідна для забезпечення фолдингу білків, які активно синтезуються під час активного росту та розвитку, а не лише як відповідь на стрес. Регуляція експресії *sHSP* на певних етапах онтогенезу (розвиток коренів, цвітіння) може бути також пов'язана із присутністю у промоторах CCAAT-боксів, які є сайтами впізнавання транскрипційних факторів з родини NF-Y (Laloum et al., 2013; Brambilla and Fornara, 2017).

Наявність GCN4_motif та O2-site у промоторах вказує на можливу роль деяких *sHSP* у розвитку насіння, що імовірно важливо для стабілізації та укладання запасних білків (Wu et al., 1998; Fauteux et al., 2009; Contreras et al., 2025). Крім того, до розвитку насіння (контроль синтезу крохмалю у ендоспермі) мають відношення деякі транскрипційні фактори MYB, які зв'язуються із сайтом MBSI. Ці білки також беруть участь у регуляції стресових реакцій та синтезі флавоноїдів (Xiao et al., 2017).

Присутність елементів circadian свідчать про зміни експресії *sHsp* протягом доби. Це узгоджується з даними, що деякі *sHSP* експресуються ритмічно, навіть за відсутності стресу, підтримуючи клітинний гомеостаз (Li et al., 2019). Цис-елементи HD-Zip 1 і HD-Zip 3 є мішенями специфічних для рослин транскрипційних факторів HD-Zip, які відіграють важливу роль у розвитку рослин та відповіді на стрес. Зокрема, білки групи HD-Zip 1 беруть участь у реакціях, пов'язаних з абіотичним стресом, абсцизовою кислотою, синім світлом, деетіо-

ляцією та ембріогенезом. Білки HD-Zip 3 контролюють ембріогенез, полярність листя, ініціацію латеральних органів та функцію меристеми (Elhiti & Stasolla, 2009; Li et al., 2022).

Отже, результати нашого біоінформатичного аналізу демонструють, що регуляторна архітектура *sHsp*-промоторів включає комплекс структур, які дозволяють гнучко реагувати не тільки на тепловий стрес або інші абіотичні подразники, а й активно інтегруватися в мережу регуляції нормального росту, розвитку, циркадної ритміки, а також формування репродуктивних структур. Така комбінація елементів забезпечує багатофакторну регуляцію та підкреслює еволюційну адаптацію *sHSP* до виконання широкого спектру функцій у різних фізіологічних умовах.

Висновки. Мутагенна родина *sHsp* у *Nicotiana attenuata* та *Solanum lycopersicum* включає 52 та 45 генів, для яких характерна кластерна організація, що виникла внаслідок тандемних дуплікацій окремих генів протягом еволюції. Кодовані цими генами білки належать до 19 класів, які відрізняються між собою розміром структурних ділянок та містять сигнали локалізації у різних клітинних компартментах. Промоторні ділянки *sHsp* містять сайти впізнавання багатьох транскрипційних факторів, які задіяні у відповідь на різні форми стресу, дію світла та фітогормонів, регуляцію росту і розвитку.

Дотримання етичних стандартів. Ця стаття не містить будь-яких досліджень за участю людей і хребетних тварин в якості об'єктів дослідження.

Конфлікт інтересів. Автори заявляють про відсутність конфлікту інтересів.

Фінансування. Дослідження проводились за фінансової підтримки Міністерства освіти і науки України (тема № 0121U111109).

IDENTIFICATION AND ANALYSIS
OF THE ORGANIZATION OF *sHsp* GENES
IN *NICOTIANA ATTENUATA* AND *SOLANUM*
LYCOPERSICUM (SOLANACEAE)

L.V. Kozub, Y.O. Tynkevich,
R.A. Volkov, I.I. Panchuk

Yuriy Fedkovych Chernivtsi National University,
2 Kotsiubynsky str., 58002 Chernivtsi, Ukraine
E-mail: i.panchuk@chnu.edu.ua

Low-molecular-weight heat shock proteins (sHSPs) play

a key role in protecting plants from abiotic stress and are also involved in growth and development processes. In this study, using bioinformatics approaches, a comprehensive analysis of the multigenic *sHsp* family in members of the Solanaceae family, which includes numerous food, medicinal, and ornamental plant species, was performed. As a result of a genomic screening for two members of this family, *Nicotiana attenuata* (coyote tobacco), *Solanum lycopersicum* (tomato), and the model plant *Arabidopsis thaliana*, 52, 45, and 34 putative *sHsp* genes were identified, respectively, and their chromosomal location was determined. All tobacco genes, five tomato genes, and one *Arabidopsis* gene were identified for the first time. The putative sHSPs belong to 19 classes. They differ in the size of their structural domains and contain targeting signals for various cellular compartments. Most classes formed before the divergence of the phylogenetic lineages of Superrosids and Superasterids and were conserved throughout the further evolution of dicots, while some classes proved to be specific to *A. thaliana* or members of the Solanaceae family. Several *sHsps* are characterized by a clustered organization on the chromosomes, which arose through tandem duplications of individual genes during evolution. In the genomes examined, 15 *sHsps* were identified, encoding two to four alternative transcripts. Analysis of the promoter regions revealed the presence of numerous recognition sites for transcription factors involved in the regulation of stress responses, light effects, phytohormones, and the control of growth and development.

СПИСОК ЛІТЕРАТУРИ

- Ain-Ali, Q., Mushtaq, N., Amir, R., Gul, A., Tahir, M., and Munir, F., Genome-wide promoter analysis, homology modeling and protein interaction network of Dehydration Responsive Element Binding (DREB) gene family in *Solanum tuberosum*, *PLoS ONE*, 2021, vol. 16, no. 12, pp. e0261215. <https://doi.org/10.1371/journal.pone.0261215>.
- Anisimova, M., and Gascuel, O., Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative, *Systematic Biology*, 2006, vol. 55, no. 4, pp. 539–552. <https://doi.org/10.1080/10635150600755453>
- Barnett, T., Altschuler, M., McDaniel, C.N., and Mascarenhas, J.P., Heat-shock-induced proteins in plant cells, *Develop Genet.*, 1980, vol. 1, pp. 331–340. <https://doi.org/10.1002/dvg.1020010406>.
- 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., and Zaretskaya, I., BLAST: a more efficient report with usability improvements, *Nucleic Acids Research*, 2013, vol. 41, no. W1, pp. W29–W33. <https://doi.org/10.1093/nar/gkt282>
- Brambilla, V., and Fornara, F., Y flowering? Regulation and activity of Constans and CCT-domain proteins in *Arabidopsis* and crop species, *Biochimica et Biophysica Acta (BBA) – Gene Regulatory Mechanisms*, 2017, vol. 1860, no. 5, pp. 655–660. <https://doi.org/10.1016/j.bbagr.2016.10.009>
- Buzduga, I.M., Salamon, I., Volkov, R.A., and Panchuk, I.I., Rapid accumulation of cadmium and antioxidative response in tobacco leaves, *The Open Agriculture Journal*, 2022, vol. 16, no. 1. <https://doi.org/10.2174/18743315-v16-e2206271>.
- Cao, Z., Zhang, S., Wang, R., Zhang, R., and Hao, Y., Genome wide analysis of the Apple MYB transcription factor family allows the identification of MDO-MYB121 gene conferring abiotic stress tolerance in plants, *PLoS ONE*, 2013, vol. 8, no. 7, pp. e69955. <https://doi.org/10.1371/journal.pone.0069955>
- Chen, X., Wang, Z., Tang, R., Wang, L., Chen, C., and Ren, Z., Genome-wide identification and expression analysis of Hsf and Hsp gene families in cucumber (*Cucumis sativus* L.). *Plant Growth Regulation*, 2021, vol. 95, no. 2, pp. 223–239. <https://doi.org/10.21203/rs.3.rs-189743/v1>
- Contreras, E., Alonso, R., Pastor-Mora, E., Ceballos, M.G., Vicente-Carbajosa, J., and Iglesias-Fernández, R., Regulation of *HvASN1* expression by bZIP transcription factors during barley embryo development and germination, *Planta*, 2025, vol. 262, no. 1. <https://doi.org/10.1007/s00425-025-04730-0>
- Deanna, R., Barboza, G.E., Bohs, L., Dodsworth, S., Gagnon, E., Giacomini, L.L., Knapp, S., Orejuela, A., Pocza, P., Särkinen, T., Smith, S.D., and Olmstead R.G., A new phylogeny and phylogenetic classification for Solanaceae, *bioRxiv*, 2025. <https://doi.org/10.1101/2025.07.10.663745>
- Dombrecht, B., Xue, G.P., Sprague, S.J., Kirkegaard, J.A., Ross, J.J., Reid, J.B., Fitt, G.P., Sewelam, N., Schenk, P.M., Manners, J.M., and Kazan, K., MYC2 differentially modulates diverse jasmonate-dependent functions in *Arabidopsis*, *The Plant Cell*, 2007, vol. 19, no. 7, pp. 2225–2245. <https://doi.org/10.1105/tpc.106.048017>
- Elhiti, M., and Stasolla, C., Structure and function of homodomain-leucine zipper (HD-Zip) proteins, *Plant Signaling and Behavior*, 2009, vol. 4, no. 2, pp. 86–88. <https://doi.org/10.4161/psb.4.2.7692>
- Fauteux, F., and Strömvik, M.V., Seed storage protein gene promoters contain conserved DNA motifs in Brassicaceae, Fabaceae and Poaceae, *BMC Plant Biology*, 2009, vol. 9, no. 1, pp. 126. <https://doi.org/10.1186/1471-2229-9-126>
- Feder, M.E., and Hofmann, G.E., Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology, *Annual Review of Physiology*, 1999, vol. 61, no. 1, pp. 243–282.

- <https://doi.org/10.1146/annurev.physiol.61.1.243>
Gong, C., Pang, Q., Li, Z., Li, Z., Chen, R., Sun, G., and Sun, B., Genome-wide identification and characterization of *Hsf* and *Hsp* gene families and gene expression analysis under heat stress in eggplant (*Solanum melongena* L.), *Horticulturae*, 2021, vol. 7, pp. 149. <https://doi.org/10.3390/horticulturae7060149>
- González-Gordo, S., Palma, J.M., and Corpas, F.J., Small Heat Shock Protein (*sHSP*) gene family from sweet pepper (*Capsicum annuum* L.) fruits: Involvement in ripening and modulation by nitric oxide (NO), *Plants*, 2023, vol. 12, pp. 389. <https://doi.org/10.3390/plants12020389>
- Grimanelli, D., and Ingouff, M., DNA methylation readers in plants. *J. Mol. Biol.*, 2020, vol. 432, no. 6, pp. 1706–1717. <https://doi.org/10.1016/j.jmb.2019.12.043>
- Guindon, S., and Gascuel, O., A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood, *Systematic Biology*, 2003, vol. 52, no. 5, pp. 696–704. <https://doi.org/10.1080/10635150390235520>
- Hagymasi, A.T., Dempsey, J.P., and Srivastava, P.K., Heat-shock proteins, *Current Protocols*, 2022, vol. 2, no. 11. <https://doi.org/10.1002/cpz1.592>
- Haslbeck, M., and Vierling, E., A first line of stress defense: small heat shock proteins and their function in protein homeostasis, *Journal of Molecular Biology*, 2015, vol. 427, no. 7, pp. 1537–1548. <https://doi.org/10.1016/j.jmb.2015.02.002>
- Haslbeck, M., Franzmann, T., Weinfurter, D., and Buchner, J., Some like it hot: the structure and function of small heat-shock proteins, *Nature Structural and Molecular Biology*, 2005, vol. 12, no. 10, pp. 842–846. <https://doi.org/10.1038/nsmb993>
- Hassan, M.U., Chattha, M.U., Khan, I., Chattha, M.B., Barbanti, L., Aamer, M., Iqbal, M.M., Nawaz, M., Mahmood, A., Ali, A., and Aslam, M.T., Heat stress in cultivated plants: nature, impact, mechanisms, and mitigation strategies – a review, *Plant Biosystems*, 2020, vol. 155, no. 2, pp. 211–234. <https://doi.org/10.1080/11263504.2020.1727987>
- Hu, B., Jin, J., Guo, A., Zhang, H., Luo, J., and Gao, G., GSDS 2.0: An upgraded gene feature visualization server, *Bioinformatics*, 2014, vol. 31, no. 8, pp. 1296–1297. <https://doi.org/10.1093/bioinformatics/btu817>
- Huang, X., Liu, Y., Yu, X., Cai, Y., Hou, L., Zhang, J., and Yang, H., Genome-wide identification and evolution-profiling analysis of heat shock protein gene family in Poaceae Barnhart, *International Journal of Molecular Sciences*, 2025, vol. 26, no. 9, pp. 4269. <https://doi.org/10.3390/ijms26094269>
- Hunziker, A.T., *The Genera of Solanaceae*. A.R.G. Gantner Verlag K.G., Ruggell, Liechtenstein, 2001. ISBN 3-904144-77-4
- Ji, X., Yu, Y., Ni, P., Zhang, G., and Guo, D., Genome-wide identification of small heat-shock protein (*HSP20*) gene family in grape and expression profile during berry development, *BMC Plant Biology*, 2019, vol. 19, no. 1. <https://doi.org/10.1186/s12870-019-2031-4>
- Jiang, C., Xu, J., Zhang, H., Zhang, X., Shi, J., Li, M., and Ming, F., A cytosolic class I small heat shock protein, RcHSP17.8, of *Rosa chinensis* confers resistance to a variety of stresses to *Escherichia coli*, yeast and *Arabidopsis thaliana*, *Plant Cell and Environment*, 2009, vol. 32, no. 8, pp. 1046–1059. <https://doi.org/10.1111/j.1365-3040.2009.01987.x>
- Katoh, K., Rozewicki, J., and Yamada, K.D., MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization, *Briefings in Bioinformatics*, 2017, vol. 20, no. 4, pp. 1160–1166. <https://doi.org/10.1093/bib/bbx108>
- Kazan, K., and Manners, J.M., MYC2: The master in action, *Molecular Plant*, 2013, vol. 6, no. 3, pp. 686–703. <https://doi.org/10.1093/mp/sss128>
- Komarova, N.Y., Grimm, G.W., Hemleben, V., and Volkov, R.A., Molecular evolution of 35S rDNA and taxonomic status of *Lycopersicon* within *Solanum* sect. *Petota*, *Plant Systematics and Evolution*, 2008, vol. 276, no. 1, pp. 59–71. <https://doi.org/10.1007/s00606-008-0091-2>
- Krsticevic, F.J., Arce, D.P., Ezpeleta, J., and Tapia, E., Tandem duplication events in the expansion of the small heat shock protein gene family in *Solanum lycopersicum* (cv. Heinz 1706), *G3 Genes Genomes Genetics*, 2016, vol. 6, no. 10, pp. 3027–3034. <https://doi.org/10.1534/g3.116.032045>
- Laloum, T., De Mita, S., Gamas, P., Baudin, M., and Niebel, A., CCAAT-box binding transcription factors in plants: Y so many?, *Trends in Plant Science*, 2013, vol. 18, no. 3, pp. 157–166. <https://doi.org/10.1016/j.tplants.2012.07.004>
- Letunic, I., and Bork, P., Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation, *Nucleic Acids Research*, 2021, vol. 49, no. W1. <https://doi.org/10.1093/nar/gkab301>
- Li, B., Gao, Z., Liu, X., Sun, D., and Tang, W., Transcriptional profiling reveals a time-of-day-specific role of REVEILLE 4/8 in regulating the first wave of heat shock-induced gene expression in *Arabidopsis*, *The Plant Cell*, 2019, vol. 31, no. 10, pp. 2353–2369. <https://doi.org/10.1105/tpc.19.00519>
- Li, Y., Yang, Z., Zhang, Y., Guo, J., Liu, L., Wang, C., Wang, B., and Han, G., The roles of HD-ZIP proteins in plant abiotic stress tolerance, *Frontiers in Plant Science*, 2022, vol. 13. <https://doi.org/10.3389/fpls.2022.1027071>
- Maimbo, M., Ohnishi, K., Hikichi, Y., Yoshioka, H.,

- and Kiba, A., Induction of a small heat shock protein and its functional roles in *Nicotiana* plants in the defense response against *Ralstonia solanacearum*, *Plant Physiology*, 2007, vol. 145, no. 4, pp. 1588–1599. <https://doi.org/10.1104/pp.107.105353>
- Maqsood, H., Munir, F., Amir, R., and Gul, A., Genome-wide identification, comprehensive characterization of transcription factors, cis-regulatory elements, protein homology, and protein interaction network of DREB gene family in *Solanum lycopersicum*, *Frontiers in Plant Science*, 2022, vol. 13. <https://doi.org/10.3389/fpls.2022.1031679>
- Narusaka, Y., Nakashima, K., Shinwari, Z.K., Sakuma, Y., Furihata, T., Abe, H., Narusaka, M., Shinozaki, K., and Yamaguchi-Shinozaki, K., Interaction between two cis-acting elements, ABRE and DRE, in ABA-dependent expression of *Arabidopsis* rd29A gene in response to dehydration and high-salinity stresses, *The Plant Journal*, 2003, vol. 34, no. 2, pp. 137–148. <https://doi.org/10.1046/j.1365-3113x.2003.01708.x>
- Okonechnikov, K., Golosova, O., and Fursov, M., Unipro UGENE: a unified bioinformatics toolkit, *Bioinformatics*, 2012, vol. 28, no. 8, pp. 1166–1167. <https://doi.org/10.1093/bioinformatics/bts091>
- Olive, M.R., Peacock, W., and Dennis, E.S., The anaerobic responsive element contains two GC-rich sequences essential for binding a nuclear protein and hypoxic activation of the maize *Adh1* promoter, *Nucleic Acids Research*, 1991, vol. 19, no. 25, pp. 7053–7060. <https://doi.org/10.1093/nar/19.25.7053>
- Panchuk, I.I., Volkov, R.A., and Schöffl, F., Heat stress-and heat shock transcription factor-dependent expression and activity of ascorbate peroxidase in *Arabidopsis*, *Plant physiology*, 2002, vol. 129, no. 2, pp. 838–853. <https://doi.org/10.1104/pp.001362>
- Ruibal, C., Castro, A., Carballo, V., Szabados, L., and Vidal, S., Recovery from heat, salt and osmotic stress in *Physcomitrella patens* requires a functional small heat shock protein PpHsp16.4, *BMC Plant Biology*, 2013, vol. 13, no. 1. <https://doi.org/10.1186/1471-2229-13-174>
- Rushton, D.L., Tripathi, P., Rabara, R.C., Lin, J., Ringler, P., Boken, A.K., Langum, T.J., Smidt, L., Boomsma, D.D., Emme, N.J., Chen, X., Finer, J.J., Shen, Q.J., and Rushton, P.J., WRKY transcription factors: key components in abscisic acid signaling, *Plant Biotechnology Journal*, 2011, vol. 10, no. 1, pp. 2–11. <https://doi.org/10.1111/j.1467-7652.2011.00634.x>
- Saharan, B.S., Brar, B., Duhan, J.S., Kumar, R., Marwaha, S., Rajput, V.D., and Minkina, T., Molecular and physiological mechanisms to mitigate abiotic stress conditions in plants, *Life*, 2022, vol. 12, no. 10, pp. 1634. <https://doi.org/10.3390/life12101634>
- Scharf, K.D., Höfeld, I., and Nover, L., Heat stress response and heat stress transcription factors, *Journal of Biosciences*, 1998, vol. 23, pp. 313–329. <https://doi.org/10.1007/bf02936124>
- Scharf, K., Siddique, M., and Vierling, E., The expanding family of *Arabidopsis thaliana* small heat stress proteins and a new family of proteins containing α -crystallin domains (Acid proteins), *Cell Stress & Chaperones*, 2001, vol. 6, no. 3, pp. 225. [https://doi.org/10.1379/1466-1268\(2001\)006](https://doi.org/10.1379/1466-1268(2001)006)
- Seo, P.J., Xiang, F., Qiao, M., Park, J., Lee, Y.N., Kim, S., Lee, Y., Park, W.J., and Park, C., The MYB96 transcription factor mediates abscisic acid signaling during drought stress response in *Arabidopsis*, *Plant physiology*, 2009, vol. 151, no. 1, pp. 275–289. <https://doi.org/10.1104/pp.109.144220>
- Shariatipour, N., and Heidari, B., Investigation of drought and salinity tolerance related genes and their regulatory mechanisms in *Arabidopsis* (*Arabidopsis thaliana*), *The Open Bioinformatics Journal*, 2018, vol. 11, no. 1, pp. 12–28. <https://doi.org/10.2174/1875036201811010012>
- Siddique, M., Gernhard, S., Von Koskull-Döring, P., Vierling, E., and Scharf, K., The plant sHSP superfamily: five new members in *Arabidopsis thaliana* with unexpected properties, *Cell Stress and Chaperones*, 2008, vol. 13, no. 2, pp. 183–197. <https://doi.org/10.1007/s12192-008-0032-6>
- The Angiosperm Phylogeny Group, Chase M.W., Christenhusz M.J.M., Fay M.F., Byng J.W., Judd W.S., Soltis D.E., Mabberley D.J., Sennikov A.N., Soltis P.S., and Stevens P.F., An update of the Angiosperm Phylogeny Group classification for the orders and families of flowering plants: APG IV, *Botanical Journal of the Linnean Society*, 2016, vol. 181, no. 1, pp. 1–20. <https://doi.org/10.1111/boj.12385>
- Tynkevich, Y.O., Shelyfist, A.Y., Kozub, L.V., Hemleben, V., Panchuk, I.I., and Volkov, R.A., 5S ribosomal DNA of genus *Solanum*: molecular organization, evolution, and taxonomy, *Frontiers in plant science*, 2022, vol. 13, pp. 852406. <https://doi.org/10.3389/fpls.2022.852406>
- Van Montfort, R.L.M., Basha, E., Friedrich, K.L., Slingsby, C., and Vierling, E., Crystal structure and assembly of a eukaryotic small heat shock protein. *Nature Structural & Molecular Biology*, 2001, vol. 8, no. 12, pp. 1025–1030. <https://doi.org/10.1038/nsb722>
- Volkov, R.A., Panchuk, I.I., and Schöffl, F., Small heat shock proteins are differentially regulated during pollen development and following heat stress in tobacco, *Plant Molecular Biology*, 2005, vol. 57, pp. 487–502. <https://doi.org/10.1007/s11103-005-0339-y>
- Volkov, R.A., Panchuk, I.I., Mullineaux, P.M., and

- Schöffl, F., Heat stress-induced H₂O₂ is required for effective expression of heat shock genes in *Arabidopsis*, *Plant Molecular Biology*, 2006, vol. 61, no. 4–5, pp. 733–746. <https://doi.org/10.1007/s11103-006-0045-4>
- Waters, E.R., and Vierling, E., Plant small heat shock proteins – evolutionary and functional diversity, *New Phytologist*, 2020, vol. 227, no. 1, pp. 24–37. <https://doi.org/10.1111/nph.16536>
- Wehmeyer, N., Hernandez, L.D., Finkelstein, R.R., and Vierling, E., Synthesis of small heat-shock proteins is part of the developmental program of late seed maturation, *Plant Physiology*, 1996, vol. 112, no. 2, pp. 747–757. <https://doi.org/10.1104/pp.112.2.747>
- Wu, C., Suzuki, A., Washida, H., and Takaiwa, F., The GCN4 motif in a rice glutelin gene is essential for endosperm-specific gene expression and is activated by Opaque-2 in transgenic rice plants, *The Plant Journal*, 1998, vol. 14, no. 6, pp. 673–683. <https://doi.org/10.1046/j.1365-3113x.1998.00167.x>
- Wu, J., Gao, T., Hu, J., Zhao, L., Yu, C., and Ma, F., Research advances in function and regulation mechanisms of plant small heat shock proteins (sHSPs) under environmental stresses, *The Science of the Total Environment*, 2022, vol. 825, pp. 154054. <https://doi.org/10.1016/j.scitotenv.2022.154054>
- Xiao, Q., Wang, Y., Du, J., Li, H., Wei, B., Wang, Y., Li, Y., Yu, G., Liu, H., Zhang, J., Liu, Y., Hu, Y., and Huang, Y., ZmMYB14 is an important transcription factor involved in the regulation of the activity of the ZmBT1 promoter in starch biosynthesis in maize, *FEBS Journal*, 2017, vol. 284, no. 18, pp. 3079–3099. <https://doi.org/10.1111/febs.14179>
- Yadav, S., Modi, P., Dave, A., Vijapura, A., Patel, D., and Patel, M., Effect of abiotic stress on crops, *In IntechOpen eBooks*, 2020. <https://doi.org/10.5772/intechopen.88434>
- Yu, J., Cheng, Y., Feng, K., Ruan, M., Ye, Q., Wang, R., Li, Z., Zhou, G., Yao, Z., Yang, Y., and Wan, H., Genome-wide identification and expression profiling of tomato HSP20 gene family in response to biotic and abiotic stresses, *Frontiers in Plant Science*, 2016, vol. 7. <https://doi.org/10.3389/fpls.2016.01215>
- Zhang, C., Zhang, Y., Su, Z., Shen, Z., Song, H., Cai, Z., and Yu, M., Integrated analysis of HSP20 genes in the developing flesh of peach: identification, expression profiling, and subcellular localization, *BMC Plant Biology*, 2023, vol. 23, no. 1, pp. 663. <https://doi.org/10.1186/s12870-023-04621-0>
- Zhao, P., Wang, D., Wang, R., Kong, N., Zhang, C., Yang, C., Wu, W., Ma, H., and Chen, Q., Genome-wide analysis of the potato *Hsp20* gene family: identification, genomic organization and expression profiles in response to heat stress, *BMC Genomics*, 2018, vol. 19, no. 1. <https://doi.org/10.1186/s12864-018-4443-1>

Надійшла в редакцію 07.08.2025
Після доопрацювання 17.09.2025
Прийнята до друку 18.01.2026