

**МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ
ЧЕРНІВЕЦЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ЮРІЯ ФЕДЬКОВИЧА**

Навчально-науковий інститут біології, хімії та біоресурсів

Кафедра молекулярної генетики та біотехнології

Організація мультигенної родини sHSP у *Nicotiana sylvestris*

Кваліфікаційна робота

Рівень вищої освіти – перший (бакалаврський)

Виконав:

студент 4-го курсу, групи 407

спеціальність 162 Біотехнології та
біоінженерія

Піхоцький Микола Миколайович

Керівник:

доктор біологічних наук,

професор **Панчук І.І.**

До захисту допущено

на засіданні кафедри

протокол №15 від 4 червня 2024 р.

Зав. кафедри

проф. Волков Р.А.

Чернівці – 2024

Анотація

Аналіз кодувальних ділянок (CDS) sHSP виявив варіабельність N-кінцевої ділянки та консервативність α -кристалінового домену. Делеції та додаткові амінокислотні послідовності у α -кристаліновому домені були характерні для різних класів sHSP і вказували на високий ступінь варіацій та можливі адаптаційні механізми до теплового стресу. Філогенетичний аналіз показав чітке розділення між цитоплазматичними та органельними sHSP, а також виявив значні мутації та делеції, що впливають на формування окремих гілок на філогенетичному дереві. Це дослідження підкреслює значення консервативних та варіабельних областей sHSP у відповідях на тепловий стрес.

Ключові слова: sHSP, *N. sylvestris*, кодувальна ділянка, α -кристаліновий домен, стрес, біоінформатичний аналіз, філогенія.

Abstract

The analysis of the coding regions (CDS) of sHSPs revealed the variability of the N-terminal region and the conservation of the α -crystallin domain. Deletions and additional amino acid sequences in the α -crystallin domain were characteristic of different classes of sHSPs and indicated a high degree of variation and possible adaptive mechanisms to heat stress. The phylogenetic analysis showed a clear separation between cytoplasmic and organellar sHSPs, and revealed significant mutations and deletions affecting the formation of individual branches on the phylogenetic tree. This study emphasizes the importance of conserved and variable regions of sHSPs in responses to heat stress.

Key words: sHSP, *N. sylvestris*, coding region, α -crystallin domain, stress, bioinformatic analysis, phylogeny.

ЗМІСТ

ВСТУП.....	4
РОЗДІЛ 1. ОГЛЯД ЛІТЕРАТУРИ.....	5
1.1. Загальна характеристика роду <i>Nicotiana</i>	5
1.2. Огляд диплоїдного виду <i>N. sylvestris</i>	9
1.3. Загальні аспекти будови та організації sHSP.....	11
1.4. sHSP у родини Solanaceae за дії стресових умов.....	18
РОЗДІЛ 2. МАТЕРІАЛИ ТА МЕТОДИ.....	20
2.1. Пошук послідовностей.....	20
2.2. Вирівнювання послідовностей.....	22
2.3. Побудова філогенетичного дерева.....	23
РОЗДІЛ 3. РЕЗУЛЬТАТИ ТА ОБГОВОРЕННЯ.....	26
ВИСНОВКИ.....	38
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ.....	39

ВСТУП

Малі білки теплового шоку (sHSP) - це велика група білків розміром від 16-42 кДа, які містять консервативний α -кристаліновий домен, коротку С-кінцеву послідовність та варіабельну N-кінцевої область. sHSP здатні утворювати гранули теплового стресу і відіграють важливу роль у стресостійкості. Вони є найбільш різноманітною групою HSP у рослин.

З огляду на глобальне потепління, актуальним стає вивчення механізмів відповіді рослин на високі температури, важливим компонентом якої є експресія sHSP. Структура та функція sHSP у рослин родини *Solanaceae* все ще залишаються недостатньо вивченими. Зручним модельним об'єктом для таких досліджень є диплоїдний вид *Nicotiana sylvestris*. На сьогодні існує обмежена кількість наукових досліджень, присвячених вивченню структури та функцій sHSP *N. sylvestris*.

Метою даної роботи було провести біоінформатичний аналіз sHSP *Nicotiana sylvestris* class CI-CV, а також sHSP, які мають мітохондріальну, хлоропластну та перексосимальну локалізацію.

Відповідно до мети були поставлені наступні **завдання**:

- 1) провести пошук гомологів sHSP для *Nicotiana sylvestris* у генетичній базі даних.
- 2) здійснити вирівнювання амінокислотних послідовностей кодувальної ділянки sHSP та визначити генетичні варіації sHSP у *Nicotiana sylvestris* та їх можливий вплив на функції та експресію цих білків.
- 3) побудувати філогенетичне дерево, проаналізувати розташування різних класів sHSP на філогенетичному дереві та виявити можливі причини формування окремих гілок.

РОЗДІЛ 1. ОГЛЯД ЛІТЕРАТУРИ

1.1 Загальна характеристика роду *Nicotiana*

Родина Solanaceae (Пасльонові) належить до покритонасінних, яка включає близько 100 родів та 2500 видів рослин. На основі недавньої класифікації вони поділяються на 9 підродин і 14 триб. Ці рослини поширені майже по всьому світу, за винятком Антарктиди, проте найбільше різноманіття видів спостерігається в Центральній та Південній Америці. Деякі ключові світові культури, такі як *Solanum*, *Capsicum* та *Nicotiana*, входять до цієї родини. Вони також відомі своєю різноманітністю біологічно активних речовин, корисних для здоров'я людей та захисту культур від шкідників. Займаючи близько 28 мільйонів гектарів сільськогосподарських угідь. Пасльонові включають не лише культурні рослини, але і дикорослі види, які використовуються у традиційній кухні або як вихідний матеріал для корисних генів у селекції. (Motti, 2021) . До родини Пасльонових належать помідор (*Solanum lycopersicum*), картопля (*Solanum tuberosum*) та баклажан (*Solanum melongena*). (Olmstead et al., 2008).

Одним із найпоширеніших представників родини Solanaceae є рід *Nicotiana*. Роду *Nicotiana* присвячена всесвітня монографія Гудспіда (1954), а взаємовідносини між його видами та секціями вивчалися останнім часом з використанням комбінації молекулярних та морфологічних методів. Нині розпізнано близько 82 видів роду *Nicotiana*, що поширені в Америці та Австралії, один таксон зустрічається в Африці (*N. africana* Merxm.) та ще три таксони в Тихому океані. Представники даного роду є надзвичайно різноманітними за способом життя, формою та розміром листків, морфологією квіток та хромосомним набором. Так види *Nicotiana* представлені як однорічними травами (наприклад, *N. linearis* Phil., *N. plumbaginifolia* Viv.), так і невеликими деревами м'яколистяних порід (наприклад, *N. tomentosa* Ruiz i Pav., *N. wigandioides* Koch та Fintelm). Більшість представників ростуть у вологих або сезонно вологих

середовищах і часто є ефемерами (наприклад, *N. attenuata* S.Wats. у Великому Басейні або види австралійських пустель у секції *Suaveolentes*, (Chase & Christenhusz, 2018a; Knapp, 2020).

Листя зазвичай велике, перетинчасте, вкрите липкими залозистими волосками. Ці особливості надають рослинам характерного запаху та в'язкості. Однак, деякі види мають дуже маленькі листки (наприклад, *N. linearis*). *N. glauca* Graham та деякі її родичі (секція *Noctiflorae*) не мають копієчних залоз і мають гумоподібні, дещо воскоподібні або гладенькі листки. Спеціалізовані та універсальні лускокрилі та інші травоядні комахи живляться листям нікотіани, незважаючи на високий вміст алкалоїдів і токсичність. Тютюн відтягнутий (*N. attenuata*) став модельним організмом для вивчення еволюції захисних реакцій рослин у природних популяціях завдяки дослідженням Болдвіна та його колег цієї рослини в пустелях Великого Басейну в штаті Юта (Knapp, 2020).

Забарвлення квітів *Nicotiana* варіюється від чистого білого до кремового та різних відтінки рожевого або червоного, а деякі види мають зелені або жовті квіти. Загалом, квітки розкриваються або вдень і запилюються денною міллю, метеликами та/або бджолами, або вночі - нічною міллю. Деякі види мають квіти, які розкриваються вночі, а потім залишаються відкритими вдень протягом деякого часу, що забезпечує змішану систему запилення (наприклад, *N.attenuata*, (Kessler, 2010). Запах квіток може бути досить сильним, особливо у видів, що цвітуть вночі, але денна австралійська *N. umbratica* N.T. Burb. має квіти з сильним запахом гвоздики (Chase & Christenhusz, 2018b). Квітковий запах зазвичай не описується в таксономічних дослідженнях, але може виявитися цікавим з еволюційної точки зору (Haverkamp, 2018). Незважаючи на те, що всі квітки нікотинових певною мірою трубчасті, всі вони мають кінцівку (розпростерті пелюстки віночка), яка є більш-менш п'ятикутною, але через відмінності в довжині та ширині трубки віночка вони мають відносно широке розмаїття загальних форм. Багато видів нічних квіток (наприклад,

N. longiflora Cav., *N. repanda* Willd., *N. sylvestris* Speg.) мають дуже довгі, вузькі трубки (близько 1 мм в діаметрі у деяких видів), тоді як інші (наприклад, *N. obtusifolia* M.Martens та Galeotti, *N. rustica*, *N. wigandioides*) мають коротші, ширші трубки (McCarthy et al., 2016; Knapp, 2020).

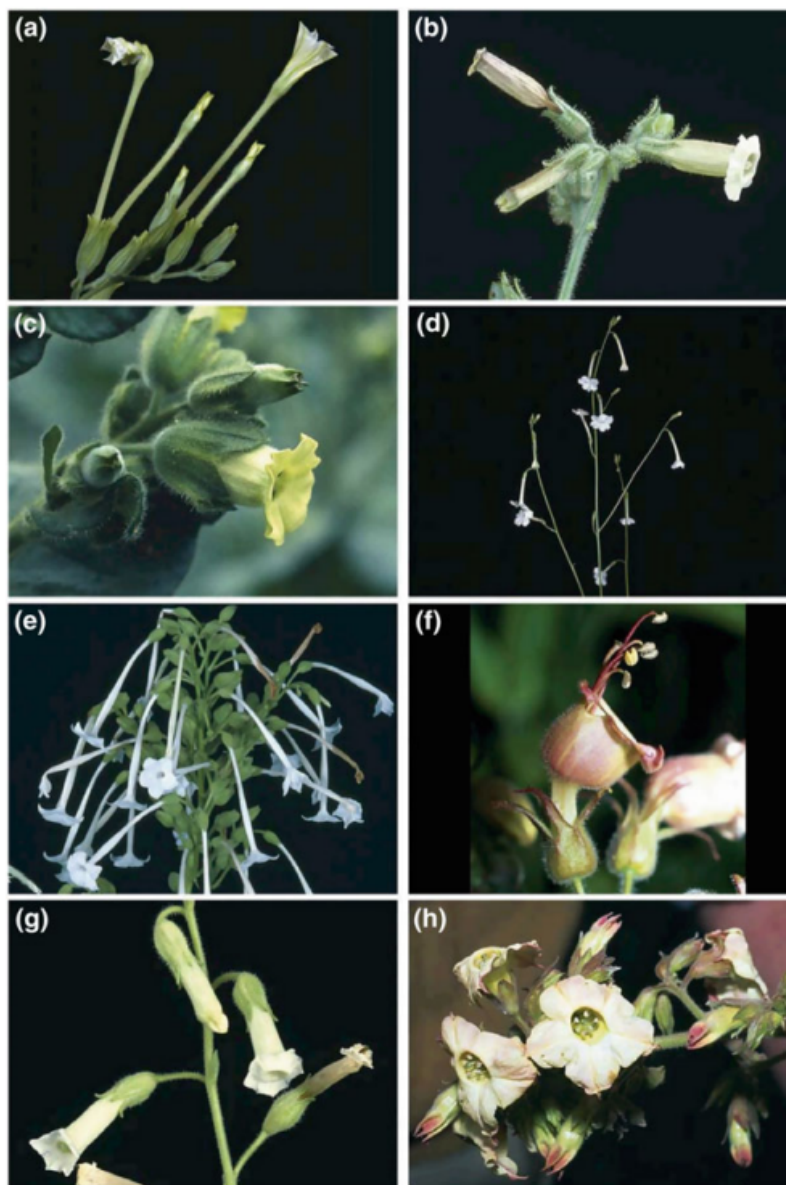


Рис. 1.1.1. Види квіток представників роду *Nicotiana*. А. Секція *Repandae*, *N. nesophila* (cult. RBG Kew); В. Секція *Repandae*, *N. nudicaulis* (cult. RBG Kew); С. Секція *Rusticae*, *N. rustica* (cult. RBG Kew); D. Секція *Suaveolentes*, *N. suaveolens* (cult. RBG Kew); Е. Секція *Sylvestres*, *N. sylvestris* (cult. RBG Kew); F. Секція *Tomentosae*, *N. otophora* (Болівія); G. Секція *Trigonophyllae*, *N. obtusifolia* (cult. RBG Kew); H. Секція *Undulatae*, *N. wigandioides* (Болівія). (Knapp, 2020)

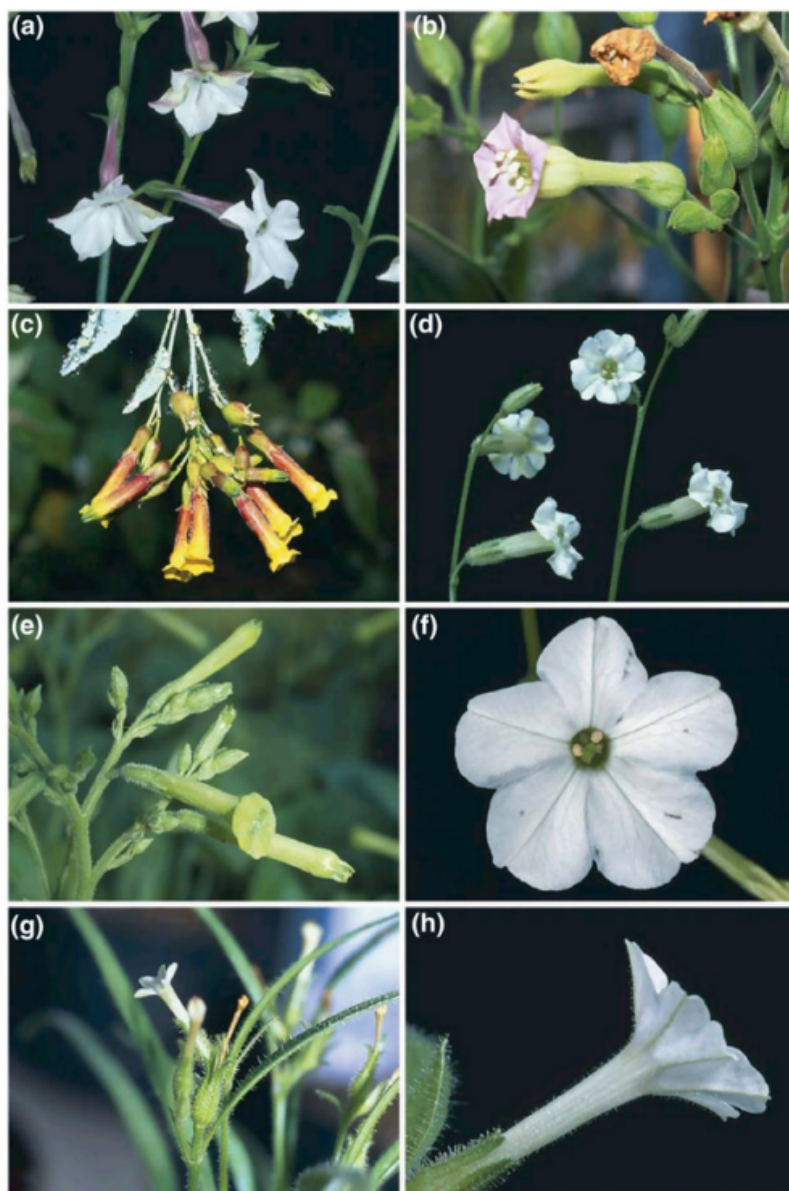


Рис. 1.1.2. Види квіток представників роду *Nicotiana*. А. Секція *Alatae*, *N. bonariensis* (cult. RBG Kew); В. Секція *Nicotiana*, *N. tabacum* (cult. Radboud University, acc. #A04750053); С. Секція *Noctiflorae*, *N. glauca* (Болівія); D. Секція *Noctiflorae*, *N. noctiflora* (cult. RBG Kew); Е. Секція *Paniculatae*, *N. paniculata* (cult. RBG Kew); F. Секція *Petunioides*, *N. attenuata* (cult. RBG Kew); G. Секція *Petunioides*, *N. linearis* (cult. Radboud University, acc. #96450099); H. Секція *Polydicliae*, *N. clevelandii* (cult. RBG Kew). (Knapp, 2020)

Рід *Nicotiana* є багатим джерелом терпеноїдів, які відіграють важливу роль як аттрактанти для деяких комах-запилювачів. У *N. tabacum* у трихомних залозах синтезуються як мембраноїдні, так і лабданоїдні

дитерпеноїди, тоді як *N. sylvestris* виробляє переважно мембраноїдні дитерпеноїди, а *N. tomentosiformis* - переважно лабданоїдні дитерпеноїди (Tissier, 2012).

Хоча за останнє десятиліття було опубліковано кілька геномів видів *Nicotiana*, в тому числі для *N. sylvestris* (Sierro et al., 2013), *N. tomentosiformis* (Sierro et al., 2013) і *N. tabacum* (Sierro et al., 2014; Edwards et al., 2017), ці геноми в основному базуються на збірці даних секвенування другого покоління і тому страждають від значної фрагментації, що призводить лише до часткової прив'язки до хромосом (Sierro et al., 2024).

1.2 Огляд диплоїдного виду *N. sylvestris*

Nicotiana sylvestris - це диплоїдна рослина (S-геном; $2n = 2x = 24$), що походить з Південної Америки регіону Анд в Аргентині та Болівії. Гібридизація із *Nicotiana tomentosiformis* (Т-геном; $2n = 2x = 24$), що відбулась 200,000 років тому, дала початок існуванню *Nicotiana tabacum*, що є алотетраплоїдом ($2n = 4x = 48$) (D'Andrea et al., 2023).

N. sylvestris (тютюн лісовий), зазвичай ще називають квітучим тютюном або білими падаючими зірками, виростає до 1-1,5 м заввишки. Має довгі білі трубчасті квіти, що мають духмяний та характерний аромат, що нагадує ваніль або жасмин. Цей аромат стає особливо сильним ввечері. Період цвітіння починається влітку та закінчується восени з початком холодів. Квіти є привабливими для метеликів та колібрі. Листя – довгасте та прикореневе, завдовжки до 40 см. Зараз, *N. sylvestris* культивується як декоративна рослина. (*Flowering tobacco Nicotiana sylvestris*)



Рис. 1.2.1. Морфологія *N.sylvestris* (*Flowering tobacco Nicotiana sylvestris*)

Як і інші представники роду *Nicotiana*, *N. sylvestris* продукує широкий спектр алкалоїдів, які, як відомо, токсичні для комах і є добре відомим механізмом захисту від травоядних тварин. *N. sylvestris* накопичує майже однакову кількість алкалоїдів у коренях і листі (3,5 мг/г у коренях і 2,1 мг/г у листі), *N. tomentosiformis* накопичує більше алкалоїдів у коренях (8,8 мг/г у коренях і 0,6 мг/г у листі), а *N. tabacum* - у листі (1,3 мг/г у коренях і 12,5 мг/г у листі) . Склад накопичених алкалоїдів варіює між трьома видами, причому *N. tabacum* отримує вигоду від генетичного і регуляторного внеску обох своїх предків (Kaminski et al., 2020). У коренях *N. sylvestris* 87% алкалоїдів становить нікотин, 11% - анатабін і 1,9% - анабазин, тоді як у листі 100% алкалоїдів - нікотин. (Kaminski et al., 2020).

Для *N. sylvestris* притаманне облігатне цвітіння в умовах довгого дня, тоді як для *N. tomentosiformis* – факультативне цвітіння в умовах короткого дня. *N. tabacum* зміг зберегти особливості своїх батьківських рослин та

демонструє нейтральне відношення до світлоперіода. *N. tabacum* поєднує два диплоїдні геноми предків і втратив лише 4-8% загальної ДНК, таким чином залишаючись тісно пов'язаним з обома видами. (Schmidt et al., 2024)

У рослинній біології *N. sylvestris* слугує диплоїдним модельним організмом для вивчення продукції терпеноїдів у залозистих трихомах (Ennajdaoui et al., 2010; Sallaud et al., 2012), конструювання геномів пластид (Maliga & Svab, 2010; Thyssen et al., 2012), функції мітохондрій (Pellny et al., 2008; Priault et al., 2006), стійкості до гербіцидів (Yemets et al., 2008) та стійкості до вірусів рослин (Sekine et al., 2012) (Sierro et al., 2013).

N. sylvestris є сучасним нащадком одного з предків інших поліплоїдних видів *Nicotiana* (наприклад, *N. nudicaulis*, *N. repanda*, *N. nesophila* і *N. stocktonii*) (Knapp et al., 2004). Chase et al. (2003) навіть припустили, що *N. sylvestris* міг бути «предком» для всіх видів *Nicotiana*, оскільки він легко продукує схрещування в межах роду. *N. tomentosiformis* здебільшого досліджували як можливого предка *N. tabacum* (Sierro et al., 2013).

1.3. Загальні аспекти будови та організації sHSP

Білки теплового шоку (HSP) є широко розповсюдженими і висококонсервативними білками, що присутні в усіх організмах і в усіх клітинах. Окремі sHSP, також відомі як шаперони, відіграють важливу роль у згортанні та розгортанні білків, складанні мультибілкових комплексів, транспортуванні та сортуванні білків у правильні субклітинні компартменти, контролі клітинного циклу, передачі сигналу, а також захисті клітин від стресу та апоптозу (Hagymasi et al., 2022).

Залежно від молекулярної маси, HSP зазвичай поділяють на такі підродини: HSP100, HSP90, HSP70, HSP60 та малі HSP (Hassan et al., 2020; Ibrahim et al., 2019; Gong et al., 2020; Wu et al., 2022). Назви HSP у бактерій відрізняються від рослинних HSP, але класифікація залишається незмінною на основі молекулярної маси (Haq et al., 2019).

Оскільки рослини ведуть прикріплений спосіб життя, є вразливими до різноманітних згубних факторів, то, відповідно, мають набагато більше sHSP, ніж інші еукаріоти. Кількість sHSP варіює у різних видів рослин: *Arabidopsis* має 19 sHSP, рис (*Oryza sativa*) - 23, а тополя (*Populus trichocarpa*) - 36. Крім того, що sHSPs дуже численні, вони також дуже різноманітні як за послідовністю, так і за місцем функціонування в клітині. Однією з найбільших відмінностей між рослинними sHSP і sHSP в інших організмах є наявність окремих підродин рослинних sHSP. Тільки у рослин існують еволюційно пов'язані підродини sHSP, які перетинають кілька видових меж. Деякі з цих підродин мають вік >400 мільйонів років, а інші з'явилися нещодавно (Waters, 2012).

Малі білки теплового шоку (sHSP) є найбільш поширеною підгрупою білків теплового шоку (HSP) з молекулярною масою від 16 до 42 кДа (Singh & Grover, 2008; Waters, 2012). У рослинах sHSP діють як важливі молекулярні шаперони, які зв'язують і стабілізують білки на проміжних стадіях згортання, збирання, деградації та транслокації через мембрани (Waters, 2012). Ці білки мають еволюційно дивергентний N-кінець, за яким слідує консервативний α -кристалічний домен і короткий C-кінець (Waters, 2012). На сьогоднішній день ідентифіковано одинадцять (Waters & Vierling, 2020) окремих родин sHSP як у однодольних, так і у дводольних рослин. Чотири з цих родин локалізовані в цитоплазмі, хоча для *N. sylvestris* є характерним наявність п'яťох родин, а інші шість - у клітинних органелах, включаючи ядро, хлоропласти, мітохондрії, ендоплазматичний ретикулум та пероксисоми. Різна внутрішньоклітинна локалізація може вказувати на функціональну дивергенцію між цими сімействами sHSP (Waters & Vierling, 2020; Zhang et al., 2015).

Раніше було виявлено, що більшість sHSP не виявляються у вегетативних тканинах за нормальних умов росту, але можуть бути швидко індуковані тепловим стресом (Waters, 2012). Пізніші дослідження виявили складні патерни експресії sHSP. Деякі sHSP експресуються не тільки під

впливом тепла, але й на певних стадіях розвитку, наприклад, під час ембріогенезу та дозрівання насіння, а також під час розвитку пилку та дозрівання плодів (Volkov et al., 2005; Zhang et al., 2015).

Експресія генів деяких sHSPs була виявлена у всіх рослинних тканинах за різних умов (Zhang et al., 2015). У рису ген *Oshsp26* індукується тепловим стресом: рівень транскрипту був найвищим, коли листя рису обробляли при 42°C протягом 2 годин. Також було виявлено, що ген *Oshsp26* експресується після окислювального стресу навіть за відсутності теплового стресу (Lee et al., 2000). У шовковиці (*Morus bombycis* Koidz.) малий білок теплового шоку (WAP20), локалізований в ендоплазматичному ретикулумі, накопичується під час сезонної холодової акліматизації (Ukaji et al., 2010). Також повідомлялося, що малі білки теплового шоку можуть підвищувати толерантність до токсичності, спричиненої наноматеріалами (Park et al., 2014).

Індукція експресії sHSP під час розвитку пилку була досліджена Volkov et al. (2005) у *N. tabacum*. Дана індукція пов'язана з потенційною можливістю стрес-індукованого ембріогенезу мікроспор. За нормальної температури більшість досліджуваних генів, що кодують цитозольні sHSP, не транскрибуються у рослинних тканинах, проте всі гени експресуються під час розвитку пилку. Низькі рівні мРНК sHSP-1A і -1B були виявлені на ранній неклітинній стадії, але зростають у чотири-сім разів у зрілому пилку. В одноклітинному пилку дев'ять інших генів мають підвищену регуляцію, тоді як у двоклітинному їхня регуляція знижувалась. Три з цих генів проявляють стадіально-специфічну експресію. За допомогою Вестерн-аналізу було показано, що цитозольні sHSP class CI та CII експресуються на всіх стадіях розвитку пилку. Відповідно, можна припустити, що певні гени sHSP можуть відігравати специфічну роль на ранніх, а інші - на пізніх стадіях розвитку пилку.

sHSP також експресуються під час ембріогенезу рослин. Збільшення кількості білків-шаперонів може відігравати важливу роль у соматичному

ембріогенезі, полегшуючи стрес пов'язаний під час переходу від соматичного до ембріонального стану. Так у *Pinus pinaster* було виявлено, що два sHSP гіперекспресуються в сім'ядольних соматичних ембріонах (дві ізоформи HSP 18.2), а один в сім'ядольних зиготичних ембріонах (class CII HSP17.6). У *Phoenix dactylifera* пов'язані зі стресом білки родини sHSP (17,6 і 70 кДа) були виявлені лише в зиготних ембріонах. Присутність sHSP 17.6 лише в зиготичних ембріонах робить його потенційним маркером зрілості зиготних ембріонів у *P. pinaster* та *P. dactylifera*. (Tchorbadjieva, 2015).

A. thaliana, є зручним модельним організмом, для якого наявна велика кількість досліджень різноманітності та експресії генів sHSP. Використовуючи аналіз доступних послідовностей повного геному ангіоспермів, дослідниками було ідентифіковано 11 підродин sHSP (таблиця 1.3). Вони включають шість підродин, які локалізовані в цитоплазмі/ядрі (CI-CVI), і п'ять підродин sHSP, які локалізуються в органелах. Підродини органел включають одну під родину, яка локалізована в ендоплазматичному ретикулумі (ER), іншу, яка локалізована в пероксисомі (PX), одну під родину, яка локалізована в хлоропласті (CP), і дві під родини, які локалізовані в мітохондріях (MTI і MTII). Десять з цих підродин присутні як в однодольних, так і в дводольних (CP, MTI, MTII, ER, PX, CI, CII, CIII, CIV і CV), а одна зустрічається тільки в однодольних (CVI) (Waters, 2012).

Важливо зазначити, що не всі sHSP належать до цих 11 консервативних підродин. Ці унікальні sHSP можуть представляти новіші генні дуплікації і можуть бути новими підродинами (Waters et al., 2008; Bondino et al ., 2011).

Таблиця 1.3. Патерни експресії та клітинна локалізація членів підродини sHSP *A. thaliana*. Інформація наведена на основі опублікованих даних (Siddique et al., 2008; Waters et al., 2008a).

	Subfamily	Abbreviation	<i>A. thaliana</i> sHSP	Stress-induced expression	Developmental expression
1	Cytoplasmic/nuclear I	CI	Hsp17.4CI	H, Os, S, D, Ox, UV-B, W	Pollen, seeds
			Hsp17.6ACI	H, Os, S, D, Ox, UV-B, W	Pollen, seeds
			Hsp17.6BCI	H, Os, S, D, Ox, UV-B, W	Pollen, seeds
			Hsp17.6CI	H, Os, S, D, Ox, UV-B, W	Pollen, seeds
2	Cytoplasmic/nuclear II	CII	Hsp17.6CII	H, Os, S, Ox, UV-B, W	
			Hsp17.7CII	H, Ox, S, D, Ox, UV-B	None
3	Cytoplasmic/nuclear III	CIII	Hsp17.4CIII	Constitutive, H	Apex, flowers, pollen, siliques, seeds
4	Cytoplasmic/nuclear IV	CIV	Hsp15.4CIV	Constitutive, H, W	Apex, flowers, siliques
5	Cytoplasmic/nuclear V	CV	Hsp21.7CV	Constitutive	
6	Cytoplasmic/nuclear VI	CVI	Hsp18.5CVI	Constitutive, H	Flowers, pollen, siliques, seeds
7	Chloroplast	CP	Hsp21CP	H, Os, S	None
8	Mitochondrial I	MTI	Hsp23.5MTI	H, Ox, UV-B	
			Hsp23.6MTI	H, Ox, UV-B	
9	Mitochondrial II	MTII	Hsp23.5MTII	H, Os, S	Seeds
10	Endoplasmic reticulum	ER	Hsp22ER	H, Os	
11	Peroxisome	PX	Hsp15.7PX	H, Os, S, Ox, B	Seeds

Хоча у всіх sHSP присутні декілька консервативних амінокислотних залишків, найбільш значущими є консервативні структурні ділянки, що знаходяться перед початком С-кінцевої області. Область HSP20, підродини sHSP *A. thaliana* містить дев'ять висококонсервативних β -листів (рис. 1.3.1), знайдених у всіх sHSP. Ця ділянка є важливою для олігомеризації білків. (Basha et al. 2012)

At21CP	1	MAST---LSFAA-SALCSPLAPSPS-VSSKSATPFVSFPRKIPSRIRAQDQRENSIDVVQQGQQKGN--	63
At23.6MT	1	MASA---LALK-RLSSSIAPRS-----RSVLRPAVSSRIFNTNAVRSYDDDGENG	47
At26.5MTII	1	MALARLARLNQ-QKLSPSLMGQSCERGLVGNRHNPMKLNRFMATAG--EQEDKMNTEVSVSEKKSP--	65
At22ER	1	MMK-----HLLSIFFIGALLGNIKTSEGS-----	25
At15.7PX	1	M-----	1
At17.6AI	1	MS-----	2
At17.6II	1	MD-----	2
At17.4BIII	1	MSAVA-----	5
At15.4CIV	1	M-----	1
At21.7CV	1	MTSSSGSLKLEIHTDDKT-----PGKWSVPLGDDVFRRLS-----	36
At18.5CVI	1	MSMIP-----	5
At21CP	64	QGSSVEKRPQRLTMDVSPFGLLDPLSF-----MRTMRQMLDTMDRMFEDTMPVSG---RNRG	118
At23.6MT	48	DGVDLYRRSVP-RRRGDFFSDFVDFPSP-----TRSVSQVLNLMQFMENPLLSAT-----R	98
At26.5MTII	66	RQNFPRRRGRKSLWRNTDDHGYFTPTLN-----GNTLIQATENMNRIFDNFN-----	112
At22ER	26	-----LSSALETTPGSLSDLWDRFPDPFKILERIPLGLER-----	62
At157PX	2	-----ADRGIFLYPFR-----R	13
At17.6AI	3	-----LIPSFPGNNRR-----INNIFDPFSLDVWDPFKELQFP-----S	37
At17.6II	3	-----LGRFPI-----ISILEDMLEVPEDHNNKTRNNPSRVYM	36
At17.4BIII	6	-----INHEFGLPEAEIKLILPISRSGE---SNES	33
At15.4CIV	2	-----DFQTIQVMPWEYVLAS-----	17
At21.7CV	37	-GGGGSEKAVF-----GEGSLSPFLFGKY-----FDPSDAFLWEFEAEVLL-----	78
At18.5CVI	6	-ISNRRLS-----PGDRIWEFFEL-----MNTFLDFPSPALFLSHHFPSLSR----EIFP	51
At21CP	119	GSVSEIRAPWDIKEEHEIKMRFDMPLGSKEDVKISVEDN--VLVIKGEQKKEDSDDD-----SWS	177
At23.6MT	99	GMGASGARRGWDIKEKDDALYLRIEMPGLSREDVKLALQD--TLVIRGEGKNEEDGG-----EEG	157
At26.5MTII	113	---VNPFLQMGQVKEQDDCYKLYREVPLTKEDVKITVNDG--ILTIKGDHKAEEKGS---PEEDEYWS	174
At22ER	63	DTSVALSPARVDWKETAEGHEIMLDIPGLKKDEVKIEVEEN-GVLRVSGERKKREEKK---GDQWHRV	126
At15.7PX	14	FQWRSRSTALIDWMESNNSHIFKINVPGYNKEDIKVQIEEG-NVLSIRGEGIKEEKKENLVHVAEREAF	82
At17.6AI	38	SSSAIANARVDWKETAEAHVFKADLPGMKKEEVKVEIEDD-SVLKISGERHVEKEEK---QDTWHRV	101
At176II	37	RDAKAMAATPADVIEHPNAYAFVVDMPGIKGDEIKVQVEND-NVLVVSGERQRENKEN---EGVKYVRM	101
At17.4BIII	34	RGRGSSNNIPIDILESPKEYIFYLDIPGISKSDIQVTVEE-RTLVIKSNKGRKRDDDESEEGSKYIRLE	102
At15.4CIV	18	QSLNNYQENHVRWSQSPDSHTFSDVPLGLRKEEIKVEIEDS-IYLIIRTEATPMSPP-----	73
At21.7CV	79	ASLRSLGQCRVDWSQTDQAYVLKSDIPVVGKNNVQVYVDINGRVMELISGQWNSKKAA---TNSDWRS	144
At18.5CVI	52	QTSSSTVNTQLNWTETPTAHVFKAYLPGVDQDEIVAFVDEE-GYLQICTGDN-----	97
		β2 β3 β4 β5 β6	
At21CP	178	GRSVSSYGTRELQLPDNC-EKDKIKAEKLN---GVLFITIPKTKVER-----KVIDVQIQ---	227
At23.6MT	158	ESGNRRFTSRIGLPDKIYKIDIKAEKLN---GVLKVVIPIKMKQERN-----DVRQIEIN---	210
At26.5MTII	175	SKSYGYNTSLSLPDDA-KVEDIKAEKLN---GVLNLVIPRTEKPKK-----NVQEISVE---	225
At22ER	127	ERSYGKFWRQFKLPDNC-DMESVKAKLEN---GVLTLNLTLSPEKVK-----GPRVNVIAAED	182
At15.7PX	83	SGGGSEFLRRIELPENV-KVDQVKAYVEN---GVLTVVVPKDTSSKSS-----KVRNVNITSKL-	137
At17.6AI	102	ERSGGGFSRKFRLPENV-KMDQVKASMEN---GVLTVTVPKVETNKKK-----AQVKSIDISG---	155
At17.6II	102	ERRMGKFMKFKLPENA-DLDKISAVCHD---GVLKVTVQKLPPPEPK-----KPKTIQVQVA---	155
At17.4BIII	103	RRLAQNLVKKFRLPEDA-DMASVTAKYQE---GVLTVVIKKLPPQPP-----KPKTVQIAVS---	155
At15.4CIV	74	DQPLKTFKRKFRLPESI-DMIGISAGYED---GVLTVIVPKRIMTRR-----LIDPSDVPESLQ---	128
At21.7CV	145	RWWEHGYVRRLELPEDA-DAKYSEAFLSNDDYSFLEIRIPKINSKNKF-----	192
At18.5CVI	98	KFMSRFKLPNNA-----LTDQVTAMMED---EFLVVVFVEKDASSSPQLPEIEENNRNVVEITGDDD	162
		β7 β8 β9 β10	

Рисунок 1.3.1 Вирівнювання sHSPs *A. thaliana*. По одному представнику від кожної з 11 підродин sHSP *A. thaliana* було вирівняно за допомогою програми структурного вирівнювання Promals3D. Кожна передбачувана α-спіраль позначена червоним кольором, а β-спіралі - синім. Консервативні β-плікативні листки, знайдені в домені HSP20, пронумеровані відповідно до *Triticum aestivum* Hsp16.9. У N-кінцевій ділянці (залишки 1-140) немає консервативних амінокислот або мотивів. (Waters, 2012)

Незважаючи на значну консервативність регіону HSP20 у рослин, в ньому все ж існують деякі цікаві відмінності між підродинами. У двох підродинах рослин (CIV та CVI) відсутній важливий β6-листок. Це

прикметно, оскільки було показано, що $\beta 6$ лист важливий для утворення димерів та олігомеризації. Незважаючи на відсутність $\beta 6$ -листа, було показано, що AtHsp18.5 CVI може утворювати димери. Siddique et al. (2008) також продемонстрували, що AHsp18.5 (CVI) може функціонувати як шаперон.

Дослідження структурних та функціональних відмінностей між підродинами CI та CII показало, що обидві підродини CI і CII утворюють додекамери і обидві мають шаперонну активність. Однак білки CII ефективніше, ніж білки CI, запобігають незворотній агрегації (Basha et al., 2010). CII sHSPs зберігають свою додекомерну структуру за вищих температур, ніж білки CI. Цікавим є те, що хоча білки CI і CII знаходяться в цитозолі, вони не утворюють гетероолігомерів (Basha et al., 2010). Однак, коли розглядаються дослідження інших видів, роль білків CII у захисті білків стає менш зрозумілою. У дослідженні протопластів томатів Tripp et al. (2009) створили нокдаун генів CI та CII і продемонстрували, що CI sHSPs мають чітку термозахисну роль, а CII sHSPs - ні.

Таким чином, закономірності збереження N-кінцевої послідовності в межах підродини, але не між підродинами, узгоджуються з даними, отриманими на окремих sHSP, що свідчать про біохімічні та біологічні відмінності між підродинами. Подальші дослідження N-кінцевої ділянки і того, як відмінності в послідовності тут відображають функціональні зміни, матимуть вирішальне значення для розуміння взаємозв'язку між послідовністю і функцією в sHSPs (Waters, 2012).

Варіабельні властивості N-кінцевої ділянки (NTR) α -кристалінового домену, найкраще ілюструє структура пшеничного HSP16.9 (*Triticum aestivum*; TaHSP16.9). Цей еукаріотичний sHSP є додекамером, що складається із двох кілець, де кожне кільце є тримером з трьох альфа-кристалінових опосередкованих димерів. Таким чином, структура цього еукаріотичного sHSP ілюструє, що ледь помітна різниця в

локальному оточенні може мати потужний вплив на структуру NTR. (Heirbaut et al., 2015)

1.4. sHSP у родини Solanaceae за дії стресових умов.

Zafar et al. (2020) досліджували функцію sHSP 17.6 за різних абіотичних стресів у *B. napus*. Ген sHSP був виділений з *N. tabacum* та перенесений у *B. napus*. Філогенетичний аналіз та прогнозування субклітинної локалізації показали, що NtHSP17.6 належить білків class C1. Піддаючи рослини тепловому стресу, рослини *B. napus* з надмірною експресією NtHSP17.6 демонстрували вищий рівень експресії, вміст води, проліну, розчиненого цукру та хлорофілу, ніж рослини дикого типу. Подібні результати були отримані за дії сольового та посушливого стресів. Можна зробити висновок, що NtHSP17.6 відіграє значну роль у стійкості до високих температур та посухи, а також має деякий незначний вплив при сольовому стресі.

Надмірна експресія гену PtsHSP 17.2 (class CII), що був отриманий з *Pinellia ternata* та перенесений в *N. tabacum*, помітно підвищує терmostійкість у трансгенних рослин. У порівнянні з рослинами дикого типу, трансгенний тютюн демонстрував менші зміни вмісту хлорофілу, та вмісту малонового діальдегіду під впливом теплового стресу. Також зростала активність антиоксидантних ферментів та вмісту проліну, збільшувалась здатність утримувати воду та антиоксидантна здатність. Крім того, рівні експресії генів, що кодують антиоксидантні ферменти, були більш сильно індуковані тепловим стресом у трансгенних ліній, ніж у дикого типу. (Tian et al., 2022)

Підвищену стійкість до теплового та окислювального стресу показав хлоропластний ген малого білка теплового шоку LeHSP 21, що був конститутивно експресований у рослинах тютюну (*N. tabacum* L. cv Wisconsin 38) з помідору (*Lycopersicon esculentum* cv PKM-1). Крім того, коли розсаду піддавали високотемпературній обробці, активність

антиоксидантних ферментів і вміст проліну були значно вищими у трансгенних рослин, ніж у рослин дикого типу. (Zhang et al., 2015)

Liu et al. (2019) вивчали механізми захисту *Citrullus lanatus* та *N. benthamiana* до вірусу зеленої крапчастої мозаїки огірків (CGMMV). Так було зафіксовано, що цитоплазматичні sHSP (NbsHSP та ClsHSP) підвищується за початкової інфекції CGMMV у рослинах *N. benthamiana* та *C. lanatus*, але не в системно інфікованих листках. Надмірна експресія sHSP пригнічувала накопичення вірусної РНК і сповільнювала розвиток хвороби, тоді як пригнічення sHSP не мало помітного впливу на інфекцію CGMMV. Таким чином, дослідники припустили, що ідентифікований ними sHSP може бути одним із факторів, що модулюють захисні механізми хазяїна у відповідь на інфекцію CGMMV.

Підвищену стійкість до холодового стресу було зафіксовано Li et al. (2018) у sHSP, що локалізований у ендоплазматичному ретикулумі (ER-sHSP), CaHSP22.5. Дослідниками було виділено цей ген з солодкого перцю (*Capsicum annuum* L.) та перенесено у рослини тютюну (*N. tabacum* L.).

За дії холодового стресу у трансгенних ліній, у порівнянні з нетрансформованими, зріс індекс продуктивності та максимальна ефективність фотохімічної реакції PSII (F_v/F_m), а також знизилось накопичення H_2O_2 та супероксидних радикалів ($O_2^{\cdot-}$). Це може свідчити, що накопичення CaHSP22.5 підвищувало фотохімічну активність та стійкість до окислення. Конститутивна експресія CaHSP22.5 знижувала експресію інших молекулярних шаперонів ендоплазматичного ретикулуму, що означає, що конститутивна експресія ER-sHSP послаблює стрес ендоплазматичного ретикулуму, спричинений низькими температурами. Вчені припускають, що CaHSP22.5 стабілізує розгорнуті білки як шаперон і підвищує активність ферментів, що поглинають активні форми кисню, щоб уникнути окислювального пошкодження під час холодового стресу.

РОЗДІЛ 2. МАТЕРІАЛИ ТА МЕТОДИ

2.1. Пошук послідовностей

Матеріалом роботи слугували амінокислотні послідовності sHSP *N. sylvestris*. Для пошуку амінокислотних та нуклеотидних послідовностей ми використовували базу даних GenBank.

GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) – це загальнодоступна база даних послідовностей для 420 000 описаних видів і супровідної бібліографічної та біологічної анотації. GenBank створюється і поширюється Національним центром біотехнологічної інформації (NCBI), підрозділом Національної медичної бібліотеки (NLM), розташованим на території Національного інституту здоров'я США (NIH) в Бетесді, штат Меріленд, США.

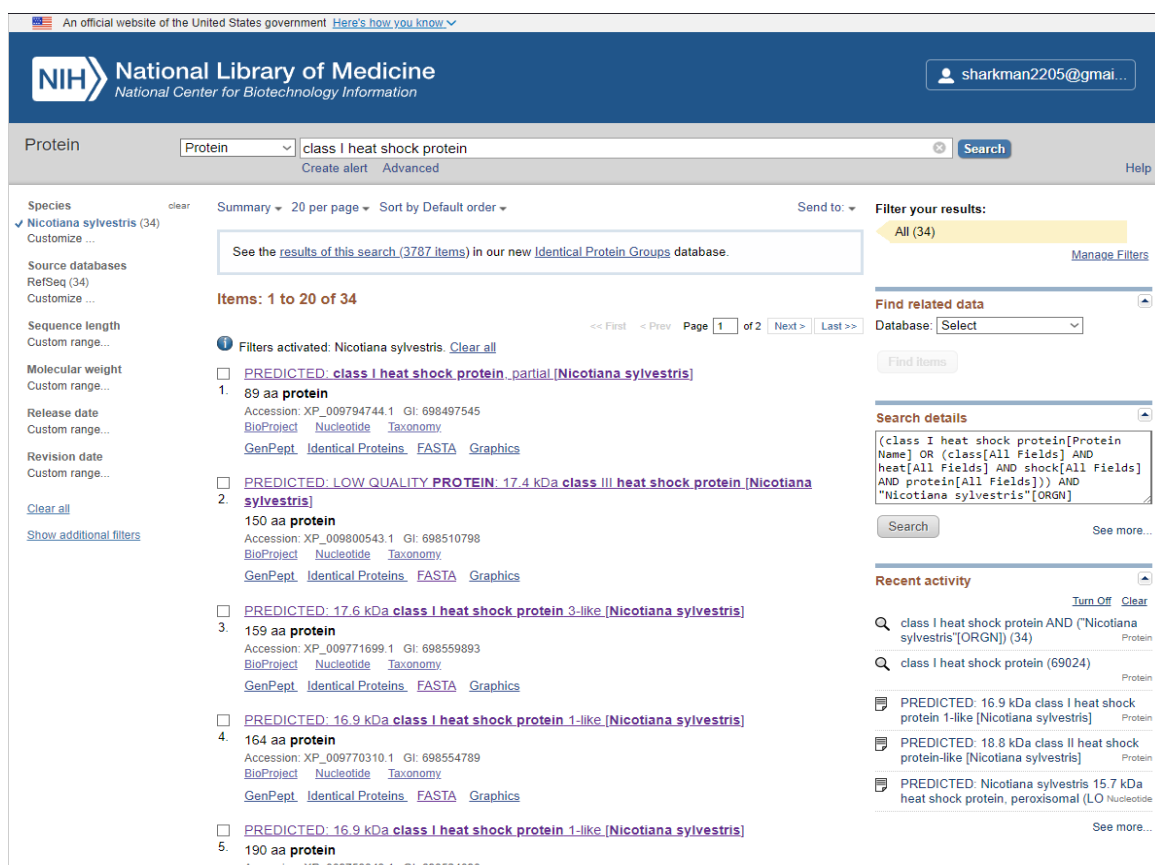


Рис. 2.1.1. Пошукове вікно GenBank

Щоденний обмін з Європейським нуклеотидним архівом (ENA) та Японським банком даних ДНК (DDBJ) забезпечує глобальне покриття даних. Більшість документів GenBank надсилаються за допомогою

BankIt, порталу подання NCBI або інструменту tbl2asn і надходять від окремих лабораторій і групових повідомлень із великомасштабних проектів секвенування, включаючи проекти повного генома (WGS) і проекти відбору проб навколишнього середовища. GenBank доступний через базу даних нуклеотидів NCBI, яка містить посилання на супутню інформацію, таку як таксономія, геноми, білкові послідовності та структури, а також біомедичну журнальну літературу в PubMed (Sayers et al., 2018).

За допомогою ПЗ BLAST ми змогли розширити пошук необхідних нам послідовностей. BLAST (Basic Local Alignment Search Tool (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) - один з найпопулярніших інструментів для пошуку та вирівнювання послідовностей.

NIH National Library of Medicine
National Center for Biotechnology Information

sharkman2205@gmail...

BLAST® » blastp suite » results for RID-6G120V1R016

Home Recent Results Saved Strategies Help

[Edit Search](#) Save Search Search Summary

How to read this report? BLAST Help Videos Back to Traditional Results Page

Job Title XP_009770310:PREDICTED: 16.9 kDa class I heat...

RID 6G120V1R016 Search expires on 06-12 16:01 pm [Download All](#)

Program BLASTP [Citation](#)

Database nr [See details](#)

Query ID XP_009770310.1

Description PREDICTED: 16.9 kDa class I heat shock protein 1-like [Nic ...

Molecule type amino acid

Query Length 164

Other reports [Distance tree of results](#) [Multiple alignment](#) [MSA viewer](#)

Filter Results

Organism only top 20 will appear ☐ exclude

Type common name, binomial, taxid or group name

[Add organism](#)

Percent Identity to **E value** to **Query Coverage** to

[Filter](#) [Reset](#)

Descriptions Graphic Summary Alignments Taxonomy

Sequences producing significant alignments Download Select columns Show 100

☒ select all 100 sequences selected

[GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#) [MSA Viewer](#)

	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per Ident	Acc. Len	Accession
☒	PREDICTED: 16.9 kDa class I heat shock protein 1-like [Nicotiana sylvestris]	Nicotiana sylvestris	332	332	100%	2e-114	100.00%	164	XP_009770310.1
☒	PREDICTED: 16.9 kDa class I heat shock protein 1-like [Nicotiana attenuata]	Nicotiana attenuata	317	317	100%	2e-108	95.73%	172	XP_019227102.1
☒	16.9 kDa class I heat shock protein 1-like [Nicotiana tomentosiformis]	Nicotiana tomentosiformis	269	269	99%	1e-89	92.64%	172	XP_009613293.1
☒	PREDICTED: 16.9 kDa class I heat shock protein 1-like [Nicotiana tabacum]	Nicotiana tabacum	269	269	99%	2e-89	92.64%	172	XP_016515155.1
☒	16.9 kDa class I heat shock protein 1-like [Nicotiana tomentosiformis]	Nicotiana tomentosiformis	225	225	100%	3e-72	72.12%	162	XP_009599116.1
☒	hypothetical protein KY284_021141 [Solanum tuberosum]	Solanum tuberosum	215	215	100%	2e-68	67.07%	161	KAH0680056.1

Рис. 2.1.2. Пошукове вікно BLAST

BLAST приймає нуклеотидну або білкову послідовність як вхідні дані і шукає її в базі даних нуклеотидних або білкових послідовностей. Також дане ПЗ здатне перекладати нуклеотидні послідовності за

необхідності. Таким чином, BLAST може здійснювати пошук за нуклеотидним запитом у базі даних білків або за білковим запитом у базі даних нуклеотидів. BLAST використовує евристики для прискорення пошуку. BLAST також надає статистику, яка оцінює ймовірність випадкового збігу. Веб-сайт BLAST видає результати з посиланнями на інші ресурси NCBI. (Boratyn et al., 2013).

Завантажувались та зберігались послідовності у форматі FASTA (.fasta) для наступного опрацювання.

2.2. Вирівнювання послідовностей

Вирівнювання отриманих послідовностей проводилось за допомогою програмного забезпечення MAFFT (множинне вирівнювання з використанням швидкого перетворення Фур'є). Дане ПЗ може використовуватись для вирівнювання як амінокислотних, так і нуклеотидних послідовностей. Вирівнювання проводилось методом L-INS-I («по кінцях»). Це ймовірно, найточніший метод, що рекомендується для <200 послідовностей. Його робота полягає у ітеративному методі уточнення з використанням локальної інформації про попарне вирівнювання (Nakamura et al., 2018)

Advanced settings

Strategy:

☐ Auto (FFT-NS-1, FFT-NS-2, FFT-NS-i or L-INS-i; depends on data size) [Updated](#)

Progressive methods

☐ FFT-NS-1 (Very fast; recommended for >2,000 sequences; progressive method)

☐ FFT-NS-2 (Fast; progressive method)

☐ G-INS-1 (Slow; progressive method with an accurate guide tree)

Iterative refinement methods

☐ FFT-NS-i (Slow; iterative refinement method)

☐ E-INS-i (Very slow; recommended for <200 sequences with multiple conserved domains and long gaps; 2 iterative cycles only) [Help](#) [Updated](#) (2015/Jun)

☒ L-INS-i (Very slow; recommended for <200 sequences with one conserved domain and long gaps; 2 iterative cycles only) [Help](#)

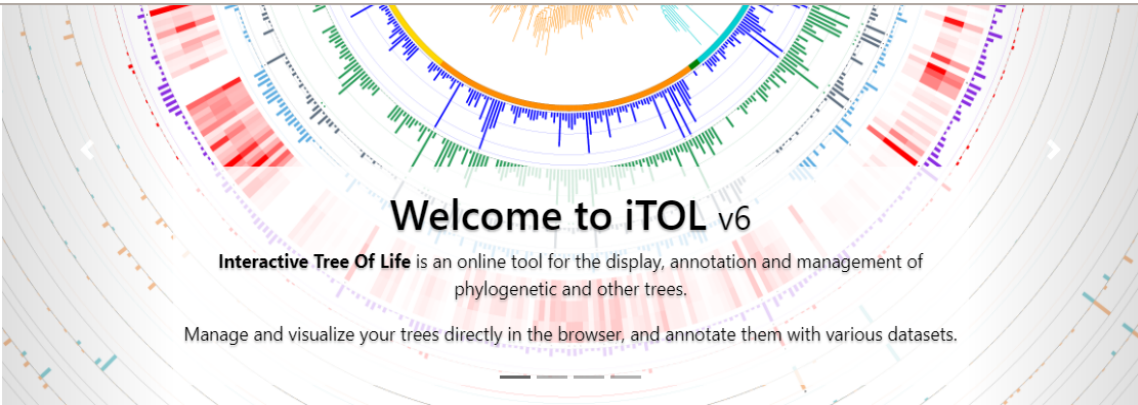
☐ G-INS-i (Very slow; recommended for <200 sequences with global homology; 2 iterative cycles only) [Help](#)

☐ Q-INS-i (Extremely slow; secondary structure of RNA is considered; recommended for a global alignment of highly divergent ncRNAs with <200 sequences × <1,000 nucleotides; the number of iterative cycles is restricted to two, 2016/May) [Help](#)

Рис. 2.2.1. Вибір методу вирівнювання у MAFFT

2.3. Побудова філогенетичного дерева

Філогенетичний аналіз проводили методом Maximum likelihood за допомогою плагіна PhyML для Geneious Prime 2021.0.3.2. Статистична підтримка гілок була розрахована за допомогою тесту aLRT Chi. Отримане дерево було експортоване у форматі Newick та анотоване за допомогою онлайн інструменту iTOL v6 - Interactive Tree Of Life. iTOL (The Interactive Tree Of Life (<https://itol.embl.de>)) - це онлайн-інструмент, доступний за допомогою будь-якого сучасного веб-браузера. Механізм відображення дерева реалізований на чистому JavaScript і використовує елемент HTML5 Canvas для візуалізації. Більшість обчислень і функцій відображення виконуються клієнтським веб-браузером, що дозволяє здійснювати тонкий інтерактивний контроль над різними параметрами відображення. Дане ПЗ підтримує різні формати відображення дерев, такі як філограми або кладограми, вкорінені або невикорінені, прямокутні або радіальні.



Welcome to iTOL v6

Interactive Tree Of Life is an online tool for the display, annotation and management of phylogenetic and other trees.

Manage and visualize your trees directly in the browser, and annotate them with various datasets.

Note: See the details on [iTOL access modes and subscriptions](#)

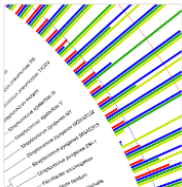
Current changelog: [version 6.9](#)



	Tree upload	Settings	Leaves	Datasets
1	✓	✓	111	4
2	✓	✓	4110	4
3	✓	✓	111	4
4	✓	✓	4110	4
5	✓	✓	111	4
6	✓	✓	4110	4
7	✓	✓	111	4
8	✓	✓	4110	4
9	✓	✓	111	4
10	✓	✓	4110	4

Manage

Organize your trees into workspaces and projects, and access them from any browser. Simply drag and drop multiple tree files onto a project to upload them all at once.



Annotate

19 dataset types. Full control over branch colors, widths and styles. Individually adjustable label fonts, sizes and styles. Check our [gallery of user created trees](#).



Export

Create high quality tree figures for your publications. Direct What-You-See-Is-What-You-Get export of what is displayed on the screen. Export into various vector or bitmap formats.

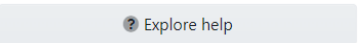
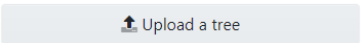
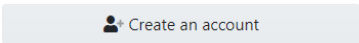


Рис. 2.3.1. Інтерфейс iTOL

Базові функції редагування дозволяють користувачам інтерактивно видаляти або переміщувати окремі вузли або цілі клади. Клади також можна обрізати або згортати, вручну або автоматично, на основі різних параметрів (наприклад, пов'язаних значень бутстрапу або середньої відстані між гілками). Окрім цього, наявні інші різноманітні варіанти маніпуляцій з отриманим деревом, такі як вкорінення вручну у будь-якому вузлі, використовуючи метод вкорінення в середній точці, або автоматично, сортування листя різними способами (Letunic & Bork, 2021).

Побудова філогенетичного дерева проводилась методом Neighbor-joining («ближнього сусіда»). Вперше цей метод був представлений у статті 1987 року і використовується дотепер. Даний метод використовує матрицю відстаней для побудови дерева, визначаючи, які листки є "сусідами" (тобто нащадками одного і того ж внутрішнього батьківського вузла) за допомогою ітеративного процесу кластеризації. Отримане дерево показує мінімальну кількість еволюції, необхідну для пояснення відмінностей між об'єктами, що робить його методом мінімальної еволюції.

Перевагами використання NJ методу є:

- Простота та зрозумілість;
- Він швидкий і недорогий у порівнянні з іншими популярними методами. Особливо методи максимальної правдоподібності та Байєса набагато повільніші;
- Він працює. Об'єднання сусідів виявилось топологічно точним і іноді перевершує більш складні методи, такі як метод максимальної правдоподібності та байєсівський висновок.

Недоліками ж є:

- Втрата даних. Коли ви стискаєте вирівнювання послідовностей або інші дані у відстані, ви здійснюєте зменшення даних. Це не обов'язково погано (методи впорядкування, такі як PCA, також роблять це), але вам слід пам'ятати про це.

- Отримання лише одного можливого дерева. Інші методи, такі як метод максимальної правдоподібності та байєсівський висновок, дають кілька різних дерев, тобто еволюційних гіпотез, які можуть бути корисними для деяких аналізів.
- Приєднання сусідів іноді може призвести до від'ємної довжини гілок. Зауважте, що це не впливає на топологію дерева, лише на довжину гілок (Harrision, 2022).

РОЗДІЛ 3. РЕЗУЛЬТАТИ ТА ОБГОВОРЕННЯ

Нами було проаналізовано 25 амінокислотних послідовностей sHSP родин class CI-CV, а також мітохондріальних (Mit), хлоропластних (Chl) та пероксисомальних (Per) у *N. sylvestris*.

Таблиця 2.1. Амінокислотні та нуклеотидні послідовності sHSP у *N. sylvestris* з GenBank

№	Розмір sHSP	Номер амінокислотної послідовності	Довжина CDS (а.п.)	Номер нуклеотидної послідовності	Довжина CDS (н.п.)
Class CI					
1	16.9	XP_009770310.1	164	XM_009772008.1	495
2		XP_009795364.1	164	XM_009797062.1	495
3		XP_009758848.1	190	XM_009760546.1	573
4	17.6	XP_009771699.1	159	XM_009773397.1	480
5		XP_009762901.1	153	XM_009764599.1	462
6		XP_009761476.1	153	XM_009763174.1	462
7	17.8	XP_009793624.1	177	XM_009795322.1	534
8		XP_009780157.1	159	XM_009781855.1	480
9		XP_009770368.1	154	XM_009772066.1	465
10	18.1	XP_009767864.1	195	XM_009769562.1	588
11		XP_009772746.1	104	XM_009774444.1	315

Class CII					
12	17.1	XP_009765474.1	195	XM_009767172.1	588
13		XP_009779651.1	159	XM_009781349.1	480
14	17.3	XP_009794705.1	157	XM_009796403.1	474
15		XP_009758838.1	157	XM_009760536.1	474
16	18.8	XP_009765475.1	161	XM_009767173.1	486
Class CIII					
17	17.4	XP_009800543.1	150	XM_009802241.1	453
Class CIV					
18	22.0	XP_009763486.1	192	XM_009765184.1	579
Class CV					
19	15.4	XP_009762376.1	137	XM_009764074.1	414
Chloroplastic					
20	-	XP_009791785.1	241	XM_009793483.1	726
21	-	XP_009763042	216	XM_009764740.1	651
22	-	XP_009779611	217	XM_009781309.1	654
Mitochondrial					
23	22.0	XP_009763041.1	192	XM_009764739.1	579
24	26.5	XP_009787010.1	230	XM_009788708.1	693

Peroxisomal					
25	15.7	XP_009779705.1	143	XM_009781403.1	432

Вирівнювання кодувальної ділянки (CDS) для 25 амінокислотних послідовностей показало, що N-кінцева ділянка sHSP є варіабельною, що може бути наслідком динамічних змін або недостатньої стабільності N-кінцевих областей. В свою чергу α -кристаліновий домен зберігає свою консервативність, С-кінець є коротким та містить характерний мотив, який бере участь у гомоолігомеризації при дії теплового стресу.

Область N-кінця цитоплазматичних sHSP коливається від 30 до 90 амінокислотних залишків. Максимальна амінокислотна послідовність спостерігається для класу CI, складаючи 90 амінокислотних залишків, тоді як для класу CV ця послідовність найменша і становить 31 амінокислотний залишок. Щодо sHSP хлоропластів, N-термінальна область містить від 120 до 145 амінокислотних залишків, для мітохондріальних - від 104 до 130 амінокислотних залишків, а для пероксисомальних - лише 28 амінокислотних залишків. Альфа-кристаліновий домен складається з приблизно 90 амінокислотних залишків. У свою чергу, С-кінцева область є короткою, з орієнтовно 14-33 амінокислотними залишками для класу CI, 13-15 залишками для хлоропластів sHSP, від 8 до 14 залишками для мітохондріальних та лише 18 залишками для пероксисомальних (Рис. 3.1).

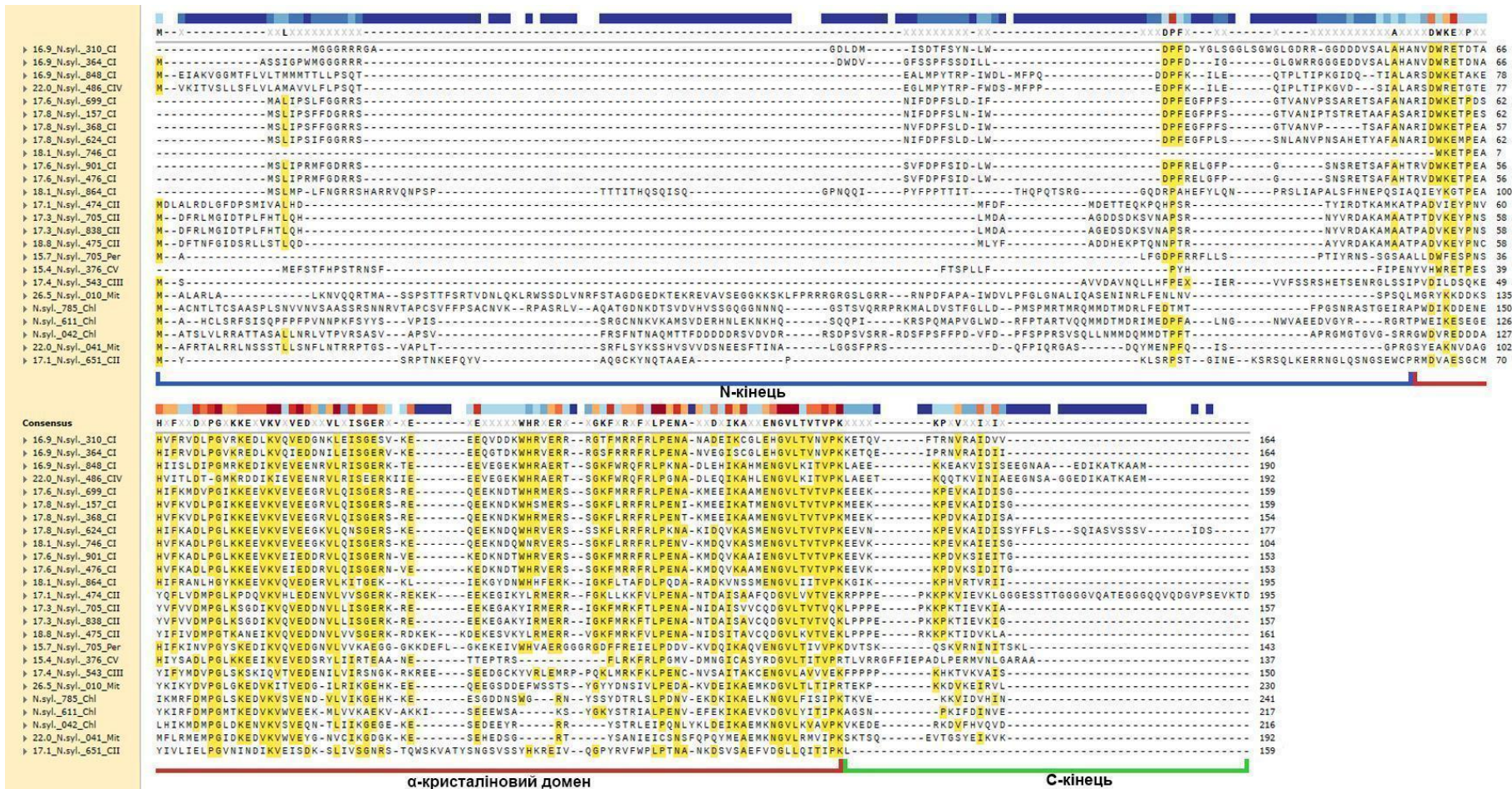


Рисунок 3.1. Результати вирівнювання CDS послідовностей sHSP для *N. sylvestris*

Отримані результати ми порівняли з результатами будови кодувальної ділянки sHSP в *Lycopersicum esculentum* опублікованими Scharf et al.(2001). Так, для α -кристалінового домену LpHSP17.7-CI у мотиві $\beta 2$ характерна амінокислотна послідовність - IDWKE, $\beta 3$ – AHVFKADL, $\beta 4$ – VKVEVE, $\beta 5$ – VLQISG, $\beta 6$ – WQRVER, $\beta 7$ – FMRRFRL, $\beta 8$ – KASME та $\beta 9$ – LTVTVPK, бета-листи зберігають свою консервативність у межах родини. N-кінцева ділянка є довгою та варіабельною, тоді як С-кінець – короткий з наявним мотивом $\beta 10$ (VKS). У дослідженнях Ji et al. (2019) описано будову ACD HSP20 у винограді (*Vitis vinifera*). Лише мотив $\beta 2$ - DWKETPEAHVFKADLPGLKKE, який містить 21 амінокислотний залишок має подібність із мотивами ACD $\beta 2$, $\beta 3$ *N. sylvestris*, всі інші мотиви відрізняються. Також для *Vitis vinifera* характерний довший ACD, ніж для *N. sylvestris*.

В α -кристаліновому домені sHSP (рис. 3.2) class CI на початку домену для послідовностей 16.9_310_CI та 16.9_364_CI наявна делеція в одну амінокислоту, а для послідовності 18.1_746_CI характерна трьох амінокислотна делеція. Для послідовності 17.6_699_CI притаманна делеція VPK перед початком С-кінця. Для послідовностей 18.1_864_CI та 16.9_848_CI відзначається значна кількість мутацій. У кінці α -кристалінового домену присутня характерна послідовність амінокислотних залишків MENGVLTVPK. У послідовностях sHSP class CII (17.1_474, 17.1_651, 17.3_705, 17.3_838, 18.8_475) притаманна двох амінокислотна делеція на початку α -кристалінового домену. При цьому для 17.1_474, 17.3_705 та 17.3_838 є додаткова делеція двох останніх амінокислот. Послідовність 18.8_475 має додатковий амінокислотний набір наприкінці домену LPPPERKKPKTIDV. Делеція першої амінокислоти та наявність додаткової амінокислотної послідовності FPPPPKHKTVM є характерним для sHSP class CIII - 17.4_543. Делеція у 8 положенні наявна для CIV_486. Для послідовності 15.4_376_CV наявна делеція першої амінокислоти. Хлоропластні sHSP (785_Ch1, 611_Ch1 та 042_Ch1) мають

двох амінокислотну делецію на початку, крім того для 785_ChI наявна додаткова амінокислотна послідовність TKVEKKVIDVHIN наприкінці домену. Пероксисомальна послідовність 15.7_705 має одно амінокислотну делецію на початку та додаткову амінокислотну послідовність GG, якої немає в інших класах sHSP. Для мітохондріальних послідовностей 22.0_041 є характерною делеція перших одинадцяти, 26.5_010 - делеція перших чотирьох амінокислот. 26.5_010_Mit характеризується додатковою амінокислотою послідовністю (TEKPKK) наприкінці домену.

Бета-листи, які формуються при просторовому укладанні білкової молекули містять окремі амінокислотні заміни, проте в цілому зберігають консервативні мотиви.

На основі вирівнювання амінокислотних послідовностей sHSP було побудоване філогенетичне дерево (рис.3.3). Послідовності 16.9_310_C1 та 16.9_364_C1 формують окрему гілку від послідовностей sHSP class CI, що може бути пов'язане з присутністю делеції на початку α -кристалінового домену. Також 18.1_864_C1 та 16.9_848_C1 не увійшли у гілку sHSP class CI, що може свідчити про наявність мутацій у даних послідовностей. sHSP class CII сформували окрему гілку, за винятком 17.1_651_CII, що також може свідчити про високу ступінь відмінності від інших послідовностей свого класу. Цікавим є те, що 16.9_848_C1 увійшла в одну гілку разом з 22.0_486_CIV, що могло б свідчити про високу ступінь спорідненості між даними послідовностями. Мітохондріальні, хлоропластні та пероксисомальні послідовності формують окрему гілку від цитоплазматичних sHSP

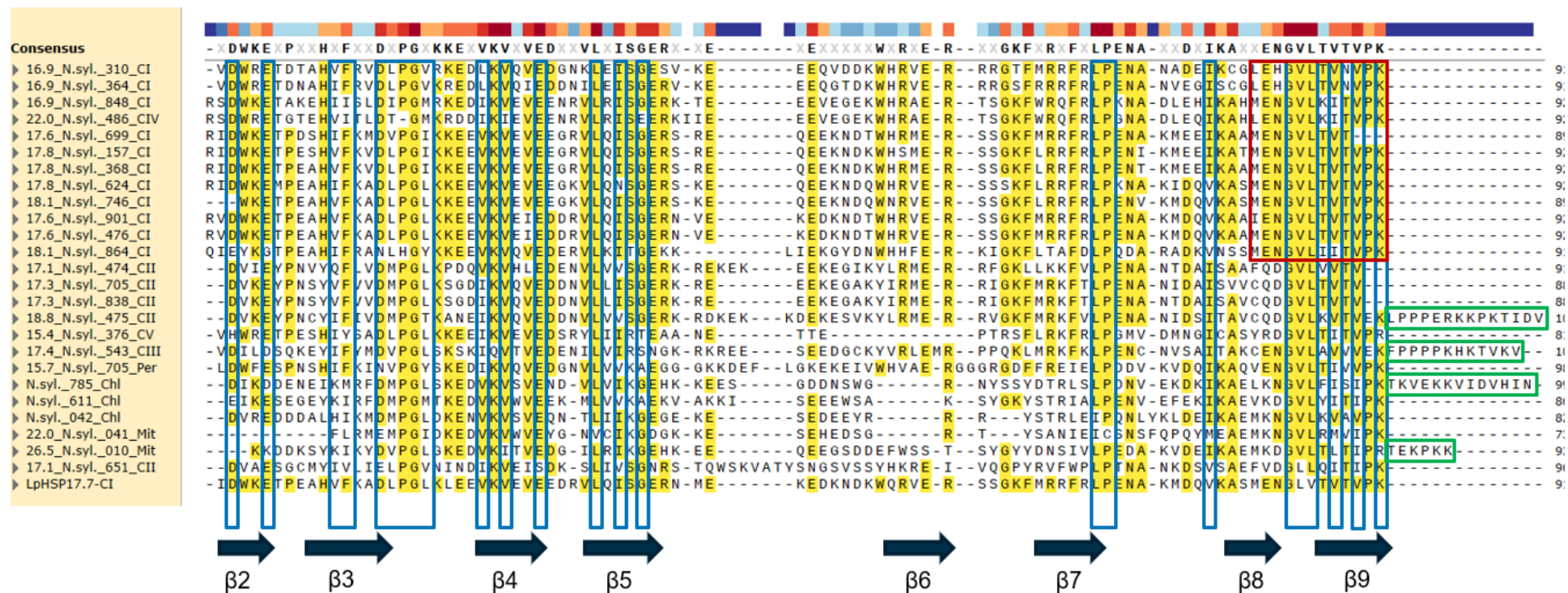


Рисунок 3.2 Амінокислотне вирівнювання ACD послідовностей sHSP класу CI, CII, CIII, CIV, CV, мітохондріальних, хлоропластних та перексисомальних у *N. sylvestris*

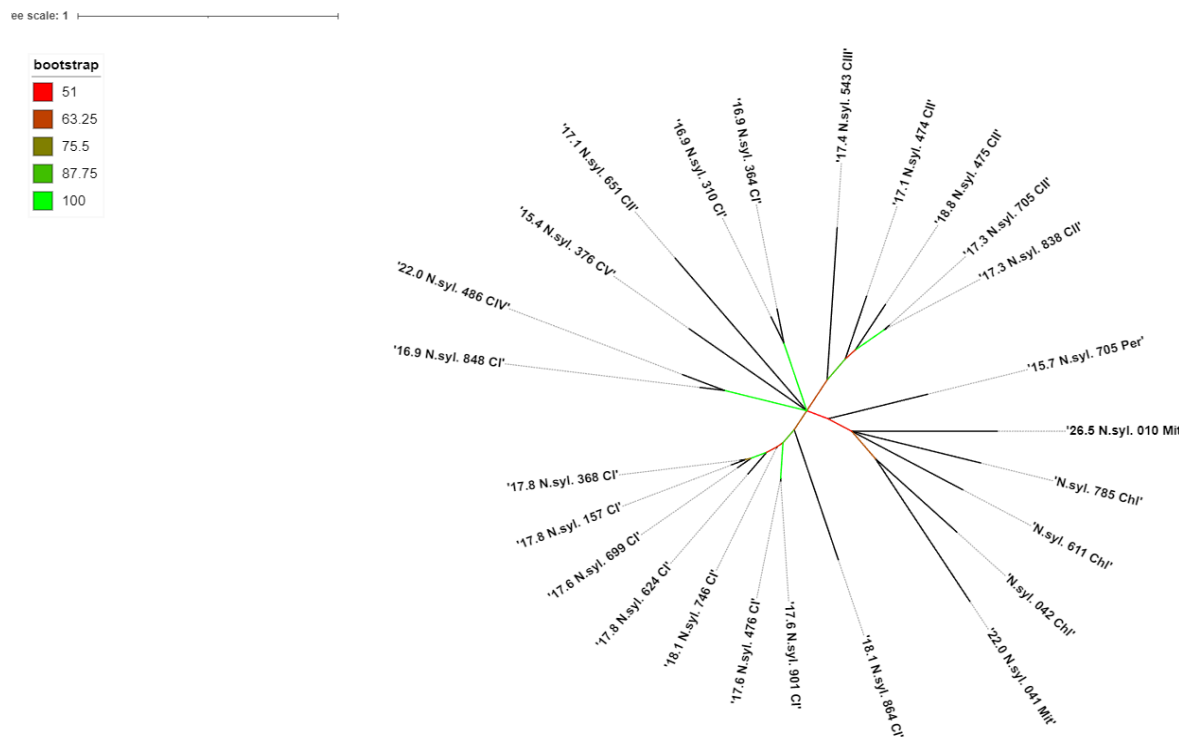


Рисунок 3.3. Філогенетичне дерево на основі вирівнювання послідовностей CDS sHSP у *N. sylvestris*

Використовуючи отримані дані, нами було побудовано таблицю (таблиця 3.1) з відсотками подібності. Так подібність між послідовностями class CI складала від 45,4% (16.9_848 та 18.1_864) до 99,4% (17.6_901 та 17.6_476). Для class CII складала від 38,7% (17.1_474 та 17.1_651) до 98,5%. Для хлороластних sHSP найменшим відсотком подібності є 46% (785_ChI та 042_ChI), а найбільшим - 48,8 (042_ChI та 611_ChI). Подібність мітохондріальних sHSP складає 39%.

Найбільший відсоток подібності між представниками class CI та CII складає 66,5% (18.1_746_CI та 17.3_838_CII), тоді як найменший - 32% (18.1_864_CI та 17.1_474_CII). Між CI та CIII найбільший відсоток 64,6 (18.1_746_CI та 17.4_543_CIII), найменший - 44,8% (18.1_864_CI та 17.4_543_CIII). Між CI та CIV найбільшим є 84,8% (16.9_848_CI та 22.0_486_CIV), а найменшим - 42,7% (18.1_864_CI та 22.0_486_CIV). Між CI та CV найбільшим є 68,9% (18.1_746_CI та 15.4_376_CV), найменшим - 47,6% (18.1_864_CI та 15.4_376_CV). Між CI та хлоропластними найбільший відсоток це 43,9 (17.6_699_CI / 17.6_476_CI та 042_ChI), найменший - 29,9% (17.8_624_CI та 785_ChI). Найбільшим відсотком між CI та мітохондріальними є 46,6 (18.1_746_CI та 22.0_041_Mit), найменшим - 29% (17.8_624_CI та 26.5_010_Mit). Між CI та пероксисомальною найбільшим є 70,4% (18.1_746_CI та 15.7_705_Per), найменшим є 50,3% (18.1_864_CI та 15.7_705_Per).

Між class CII та CIII найбільший відсоток становить 64,3 (17.3_705_CII / 17.3_838_CII / 18.8_475_CII та 17.4_543_CIII), найменший - 52,7% (17.1_651_CII та 17.4_543_CIII). Між class CII та CIV найбільший є 53,4% (17.3_705_CII / 17.3_838_CII та 22.0_486_CIV), найменшим - 38,1% (17.1_651_CII та 22.0_486_CIV). Між CII та CV найбільшим відсотком є 56,1 (17.3_838_CII та 15.4_376_CV), тоді як найменшим - 44,2% (17.1_474_CII / 17.1_651_CII та 15.4_376_CV). Між CII та хлоропластними найбільший відсоток це 42,4 (17.3_838_CII та 042_ChI), найменший - 25% (17.1_474_CII та 785_ChI).

Між СІІ та мітохондріальними найбільший - 43,3% (17.3_838_СІІ та 22.0_041_Mit), найменший - 23,5% (17.1_474_СІІ та 26.5_010_Mit). Між СІІ та пероксисомальною найбільший - 60,4% (17.3_838_СІІ та 15.7_705_Per), а найменший - 50% (17.1_474_СІІ та 15.7_705_Per).

Між class СІІ та СІV відсоток подібності становить 46,6% (17.4_543_СІІ та 22.0_486_СІV). Між СІІ та СV - 53,4% (17.4_543_СІІ та 15.4_376_CV). Між СІІ та хлоропластними найбільшим є 39,6% (17.4_543_СІІ та 611_ChI), найменшим - (17.4_543_СІІ та 785_ChI). Між СІІ та мітохондріальними найбільший відсоток це 38,1% (17.4_543_СІІ та 22.0_041_Mit), а найменшим - 32,9% (17.4_543_СІІ та 26.5_010_Mit). Між СІІ та пероксисомальною відсоток становить 64,3.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25		
	83,2	54,3	69,8	70,7	71,0	64,0	70,1	70,7	67,1	53,0	42,4	47,6	57,0	57,0	53,7	56,1	52,7	57,0	32,9	39,9	39,9	34,5	34,5	60,4	1	16.9_310_CI
		57,0	71,3	72,6	72,9	65,2	70,7	72,3	66,2	54,0	46,0	48,2	60,4	60,4	57,6	57,3	55,8	58,8	36,0	42,1	40,5	39,6	32,3	59,8	2	16.9_364_CI
			63,7	60,1	60,4	61,3	63,7	63,4	59,5	45,4	45,4	40,2	54,9	55,2	52,1	49,1	84,8	52,4	31,4	38,1	39,6	36,9	29,6	53,0	3	16.9_848_CI
				85,4	86,0	84,5	93,9	93,9	77,4	59,1	48,2	47,6	60,1	60,4	58,2	58,5	60,1	61,0	34,1	43,9	42,7	39,9	36,3	64,3	4	17.6_699_CI
					99,4	79,0	83,8	85,4	79,9	59,8	50,0	48,8	62,8	63,1	60,1	57,0	58,2	63,1	35,7	43,6	41,5	39,9	36,3	64,0	5	17.6_901_CI
						79,6	84,5	86,0	79,9	60,1	50,0	48,8	62,8	63,1	60,1	57,0	58,2	63,1	35,7	43,9	41,5	40,2	36,6	64,0	6	17.6_476_CI
							83,2	83,8	74,4	55,2	46,3	43,6	53,4	53,4	51,2	51,2	57,6	56,7	29,9	36,9	36,6	33,8	29,0	59,1	7	17.8_624_CI
								94,8	78,0	59,5	47,9	47,3	60,1	60,4	57,6	57,9	60,1	61,3	34,5	42,1	43,0	39,9	35,7	64,0	8	17.8_157_CI
									79,6	59,8	47,6	47,3	59,8	60,1	57,6	57,3	60,4	62,8	33,8	43,0	43,0	41,5	37,2	63,1	9	17.8_368_CI
										56,7	54,0	54,6	66,2	66,5	63,7	64,6	56,7	68,9	36,3	43,0	42,4	46,6	42,4	70,4	10	18.1_746_CI
											32,0	42,1	44,5	44,5	43,6	44,8	42,7	47,6	33,8	40,9	39,9	38,7	32,9	50,3	11	18.1_864_CI
												38,7	69,5	71,0	70,4	53,0	45,4	44,2	25,0	31,4	30,8	30,5	23,5	50,0	12	17.1_474_CII
													48,8	49,1	49,4	52,7	38,1	44,2	28,7	31,4	37,2	35,7	29,6	53,7	13	17.1_651_CII
														98,5	84,5	64,3	53,4	55,5	36,9	42,1	40,9	43,0	34,8	60,1	14	17.3_705_CII
															84,1	64,3	53,4	56,1	36,9	42,4	40,9	43,3	35,4	60,4	15	17.3_838_CII
																64,3	50,6	53,4	34,1	39,9	40,5	40,5	32,3	60,1	16	18.8_475_CII
																	46,6	53,4	32,0	37,5	39,6	38,1	32,9	64,3	17	17.4_543_CIII
																		51,2	29,6	36,0	37,8	34,1	26,8	51,5	18	22.0_486_CIV
																			31,7	39,3	38,4	40,9	33,8	58,5	19	15.4_376_CV
																				46,0	47,0	40,9	43,0	34,1	20	785_ChI
																					48,8	52,7	39,6	38,1	21	042_ChI
																						48,5	38,4	41,2	22	611_ChI
																							39,0	42,1	23	22.0_041_Mit
																								35,1	24	26.5_010_Mit
																									25	15.7_705_Per

Таблиця 3.1. Відсоток подібності послідовностей CDS sHSP у *N. sylvestris*

Між послідовностями CIV та CV подібність у 51,2%. Між CIV та хлоропластними найбільший - 37,8% (22.0_486_CIV та 611_ChI), найменший - 29,6% (22.0_486_CIV та 785_ChI). Між CIV та мітохондріальними найбільший відсоток становить 34,1 (22.0_486_CIV та 22.0_041_Mit), найменший - 26,8% (22.0_486_CIV та 26.5_010_Mit). Між CIV та пероксисомальною відсоток складає 51,5.

Між CV та хлоропластними найбільший відсоток це 39,3 (15.4_376_CV та 042_ChI), найменший - 31,7% (15.4_376_CV та 785_ChI). Між CV та мітохондріальними найбільший - 40,9% (15.4_376_CV та 22.0_041_Mit). Між CV та пероксисомальною відсоток складає 58,5.

Найбільший відсоток подібності між хлоропластними та мітохондріальними послідовностями складає 52,7 (042_ChI та 22.0_041_Mit), тоді як найменшим - 38,4% (611_ChI та 26.5_010_Mit). Між хлоропластними та пероксисомальною найбільший це 41,2% (042_ChI та 15.7_705_Per), а найменший - 34,1% (785_ChI та 15.7_705_Per).

Між мітохондріальними та пероксисомальною найбільший відсоток становить 42,1 (22.0_041_Mit та 15.7_705_Per), найменший - 35,1% (26.5_010_Mit та 15.7_705_Per).

Найбільшим відсотком подібності є 99,4% для 17.6_901_CI та 17.6_476_CI, тоді як найменшим - 23,5% для 17.1_474_CII та 26.5_010_Mit.

ВИСНОВКИ

1. N-кінцева ділянка sHSP є варіабельною та коливається від 30 до 90 амінокислотних залишків. С-кінець є коротким та містить характерний мотив.
2. Кінець α -кристалінового домену має характерні послідовності для різних класів sHSP, для класу CI MENGVLTVTVPK, для CII - LPPPERKKPKTIDV, для CIII - FPPPPKHKTVKV, для 785_ChI - TKVEKKVIDVHIN, для 26.5_010_Mit - TEKPKK.
3. 16.9_848_CI увійшла в одну гілку разом з 22.0_486_CIV, що свідчить про високу ступінь спорідненості між даними послідовностями.
4. Мітохондріальні, хлоропластні та пероксисомальні послідовності формують окрему гілку від цитоплазматичних sHSP.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- 1) Basha, E., Jones, C., Wysocki, V., & Vierling, E. (2010). Mechanistic Differences between Two Conserved Classes of Small Heat Shock Proteins Found in the Plant Cytosol. *Journal of Biological Chemistry/the Journal of Biological Chemistry*, 285(15), 11489–11497. <https://doi.org/10.1074/jbc.m109.074088>
- 2) Basha, E., O'Neill, H., & Vierling, E. (2012). Small heat shock proteins and α -crystallins: dynamic proteins with flexible functions. *Trends in Biochemical Sciences*, 37(3), 106–117. <https://doi.org/10.1016/j.tibs.2011.11.005>
- 3) Bondino, H. G., Valle, E. M., & Have, A. T. (2011). Evolution and functional diversification of the small heat shock protein/ α -crystallin family in higher plants. *Planta*, 235(6), 1299–1313. <https://doi.org/10.1007/s00425-011-1575-9>
- 4) 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>
- 5) Chase, M. W. (2003). Molecular Systematics, GISH and the origin of hybrid taxa in *Nicotiana* (Solanaceae). *Annals of Botany*, 92(1), 107–127. <https://doi.org/10.1093/aob/mcg087>.
- 6) Chase, M. W., & Christenhusz, M. J. M. (2018). *Nicotiana Karijini*. *Curtis's Botanical Magazine*, 35(3), 228–236. <https://doi.org/10.1111/curt.12242>
- 7) Chase, M. W., & Christenhusz, M. J. M. (2018b). *Nicotiana Umbratica*. *Curtis's Botanical Magazine*, 35(3), 278–285. <https://doi.org/10.1111/curt.12248>
- 8) D'Andrea, L., Sierro, N., Ouadi, S., Hasing, T., Rinaldi, E., Ivanov, N. V., & Bombarely, A. (2023). Polyploid *Nicotiana* section *Suaveolentes*

- originated by hybridization of two ancestral *Nicotiana* clades. *Frontiers in Plant Science*, 14. <https://doi.org/10.3389/fpls.2023.999887>
- 9) Edwards, K. D., Fernandez-Pozo, N., Drake-Stowe, K., Humphry, M., Evans, A. D., Bombarely, A., Allen, F., Hurst, R., White, B., Kernodle, S. P., Bromley, J. R., Sanchez-Tamburrino, J. P., Lewis, R. S., & Mueller, L. A. (2017). A reference genome for *Nicotiana tabacum* enables map-based cloning of homeologous loci implicated in nitrogen utilization efficiency. *BMC Genomics*, 18(1). <https://doi.org/10.1186/s12864-017-3791-6>
 - 10) Ennajdaoui, H., Vachon, G., Giacalone, C., Besse, I., Sallaud, C., Herzog, M., & Tissier, A. (2010). Trichome specific expression of the tobacco (*Nicotiana sylvestris*) cembratrien-ol synthase genes is controlled by both activating and repressing cis-regions. *Plant Molecular Biology*, 73(6), 673–685. <https://doi.org/10.1007/s11103-010-9648-x>
 - 11) *Flowering tobacco Nicotiana sylvestris*. (2019, April 15). FineGardening
<https://www.finegardening.com/plant/flowering-tobacco-nicotiana-sylvestris>
 - 12) Gong, Z., Xiong, L., Shi, H., Yang, S., Herrera-Estrella, L. R., Xu, G., Chao, D., Li, J., Wang, P., Qin, F., Li, J., Ding, Y., Shi, Y., Wang, Y., Yang, Y., Guo, Y., & Zhu, J. (2020). Plant abiotic stress response and nutrient use efficiency. *Science China. Life Sciences/Science China. Life Sciences*, 63(5), 635–674. <https://doi.org/10.1007/s11427-020-1683-x>
 - 13) Hagymasi, A. T., Dempsey, J. P., & Srivastava, P. K. (2022). Heat-Shock proteins. *Current Protocols*, 2(11). <https://doi.org/10.1002/cpz1.592>
 - 14) Haq, N. U., Khan, N., Ali, N., Khattak, N., Gai, N., Zhang, N., Wei, N., & Gong, N. (2019). Heat shock proteins: dynamic biomolecules to counter plant biotic and abiotic stresses. *International Journal of Molecular Sciences*, 20(21), 5321. <https://doi.org/10.3390/ijms20215321>

- 15) Harrision, A. (2022). Bioinformatics by hand: Neighbor-joining trees. *Tender Is The Byte*.
<https://www.tenderisthebyte.com/blog/2022/08/31/neighbor-joining-trees/>
- 16) 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>
- 17) Haverkamp, A., Hansson, B. S., Baldwin, I. T., Knaden, M., & Yon, F. (2018). Floral trait variations among wild tobacco populations influence the foraging behavior of hawkmoth pollinators. *Frontiers in Ecology and Evolution*, 6. <https://doi.org/10.3389/fevo.2018.00019>
- 18) Heirbaut, M., Strelkov, S. V., & Weeks, S. D. (2015). Everything but the ACD, Functional Conservation of the Non-conserved terminal regions in SHSPs. In *Heat shock proteins* (pp. 197–227).
https://doi.org/10.1007/978-3-319-16077-1_8
- 19) Ibrahim, I. M., Abdelmalek, D. H., & Elfiky, A. A. (2019). GRP78: A cell's response to stress. *Life Sciences*, 226, 156–163.
<https://doi.org/10.1016/j.lfs.2019.04.022>
- 20) 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>
- 21) Kaminski, K. P., Bovet, L., Laparra, H., Lang, G., De Palo, D., Sierro, N., Goepfert, S., & Ivanov, N. V. (2020). Alkaloid chemophenetics and transcriptomics of the Nicotiana genus. *Phytochemistry*, 177, 112424.
<https://doi.org/10.1016/j.phytochem.2020.112424>

- 22) Kessler, D., Diezel, C., & Baldwin, I. T. (2010). Changing pollinators as a means of escaping herbivores. *CB/Current Biology*, 20(3), 237–242. <https://doi.org/10.1016/j.cub.2009.11.071>
- 23) Knapp, S. (2020). Biodiversity of *Nicotiana* (Solanaceae). In *Compendium of plant genomes* (pp. 21–41). https://doi.org/10.1007/978-3-030-29493-9_2
- 24) Knapp, S., Chase, M. W., & Clarkson, J. J. (2004). Nomenclatural changes and a new sectional classification in *Nicotiana* (Solanaceae). *Taxon*, 53(1), 73–82. <https://doi.org/10.2307/4135490>
- 25) 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)
- 26) 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>
- 27) Li, M., Ji, L., Jia, Z., Yang, X., Meng, Q., & Guo, S. (2018). Constitutive expression of CaHSP22.5 enhances chilling tolerance in transgenic tobacco by promoting the activity of antioxidative enzymes. *Functional Plant Biology*, 45(5), 575. <https://doi.org/10.1071/fp17226>
- 28) Liu, S., Liu, L., Aranda, M. A., Peng, B., & Gu, Q. (2019). Expression and Localization Patterns of a Small Heat Shock Protein that Interacts with the Helicase Domain of Cucumber Green Mottle Mosaic Virus. *Phytopathology*, 109(9), 1648–1657. <https://doi.org/10.1094/phyto-11-18-0436-r>
- 29) Maliga, P., & Svab, Z. (2010). Engineering the Plastid Genome of *Nicotiana sylvestris*, a Diploid Model Species for Plastid Genetics. In *Methods in molecular biology* (pp. 37–50). https://doi.org/10.1007/978-1-61737-957-4_2

- 30) McCarthy, E. W., Chase, M. W., Knapp, S., Litt, A., Leitch, A. R., & Comber, S. C. L. (2016). Transgressive phenotypes and generalist pollination in the floral evolution of *Nicotiana* polyploids. *Nature Plants*, 2(9). <https://doi.org/10.1038/nplants.2016.119>
- 31) Motti, R. (2021). The Solanaceae family: botanical features and diversity. *In Compendium of plant genomes* (pp. 1–9). https://doi.org/10.1007/978-3-030-30343-3_1
- 32) Nakamura, T., Yamada, K. D., Tomii, K., & Katoh, K. (2018). Parallelization of MAFFT for large-scale multiple sequence alignments. *Bioinformatics*, 34(14), 2490–2492. <https://doi.org/10.1093/bioinformatics/bty121>
- 33) Olmstead, R. G., Bohs, L., Migid, H. A., Santiago-Valentin, E., Garcia, V. F., & Collier, S. M. (2008). A molecular phylogeny of the Solanaceae. *Taxon*, 57(4), 1159–1181. <https://doi.org/10.1002/tax.574010>
- 34) Park, H., Ko, E., & Ahn, Y. (2014). Small heat shock proteins can confer tolerance to nanomaterial-induced toxicity. *HortScience*, 49(8), 1116–1121. <https://doi.org/10.21273/hortsci.49.8.1116>
- 35) Pellny, T. K., Van Aken, O., Dutilleul, C., Wolff, T., Groten, K., Bor, M., De Paepe, R., Reyss, A., Van Breusegem, F., Noctor, G., & Foyer, C. H. (2008). Mitochondrial respiratory pathways modulate nitrate sensing and nitrogen-dependent regulation of plant architecture in *Nicotiana sylvestris*. *Plant Journal*, 54(6), 976–992. <https://doi.org/10.1111/j.1365-3113x.2008.03472.x>
- 36) Sallaud, C., Giacalone, C., Töpfer, R., Goepfert, S., Bakaher, N., Rösti, S., & Tissier, A. (2012). Characterization of two genes for the biosynthesis of the labdane diterpene Z-abienol in tobacco (*Nicotiana tabacum*) glandular trichomes. *Plant Journal*, 72(1), 1–17. <https://doi.org/10.1111/j.1365-3113x.2012.05068.x>

- 37) Sayers, E. W., Cavanaugh, M., Clark, K., Ostell, J., Pruitt, K. D., & Karsch-Mizrachi, I. (2018). GenBank. *Nucleic Acids Research*, 47(D1), D94–D99. <https://doi.org/10.1093/nar/gky989>
- 38) Schmidt, F. J., Grundmann, L., Lahme, M., Seidemann, M., Schwarze, A., Lichtenauer, S., Twyman, R. M., Prüfer, D., & Noll, G. A. (2024). COL2-dependent photoperiodic floral induction in *Nicotiana sylvestris* seems to be lost in the *N. sylvestris* × *N. tomentosiformis* hybrid *N. tabacum*. *Frontiers in Plant Science*, 14. <https://doi.org/10.3389/fpls.2023.1249879>
- 39) Sekine, K., Tomita, R., Takeuchi, S., Atsumi, G., Saitoh, H., Mizumoto, H., Kiba, A., Yamaoka, N., Nishiguchi, M., Hikichi, Y., & Kobayashi, K. (2012). Functional differentiation in the Leucine-Rich repeat domains of closely related plant Virus-Resistance proteins that recognize common AVR proteins. *Molecular Plant-microbe Interactions*, 25(9), 1219–1229. <https://doi.org/10.1094/mpmi-11-11-0289>
- 40) 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 & Chaperones*, 13(2), 183–197. <https://doi.org/10.1007/s12192-008-0032-6>
- 41) Sierro, N., Auberson, M., Dulize, R., & Ivanov, N. V. (2024). Chromosome-level genome assemblies of *Nicotiana tabacum*, *Nicotiana sylvestris*, and *Nicotiana tomentosiformis*. *Scientific Data*, 11(1). <https://doi.org/10.1038/s41597-024-02965-2>
- 42) Sierro, N., Battey, J. N., Ouadi, S., Bakaher, N., Bovet, L., Willig, A., Goepfert, S., Peitsch, M. C., & Ivanov, N. V. (2014). The tobacco genome sequence and its comparison with those of tomato and potato. *Nature Communications*, 5(1). <https://doi.org/10.1038/ncomms4833>
- 43) Sierro, N., Battey, J. N., Ouadi, S., Bovet, L., Goepfert, S., Bakaher, N., Peitsch, M. C., & Ivanov, N. V. (2013). Reference genomes and

- transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis*. *Genome Biology*, 14(6). <https://doi.org/10.1186/gb-2013-14-6-r60>
- 44) Singh, A., & Grover, A. (2008). Genetic engineering for heat tolerance in plants. *Physiology and Molecular Biology of Plants/Physiology and Molecular Biology of Plants*, 14(1–2), 155–166. <https://doi.org/10.1007/s12298-008-0014-2>
 - 45) Tchorbadjieva, M. I. (2015). Advances in proteomics of somatic embryogenesis. In *Springer eBooks* (pp. 67–90). https://doi.org/10.1007/978-81-322-2683-3_5
 - 46) Thyssen, G., Svab, Z., & Maliga, P. (2012). Exceptional inheritance of plastids via pollen in *Nicotiana sylvestris* with no detectable paternal mitochondrial DNA in the progeny. *Plant Journal*, 72(1), 84–88. <https://doi.org/10.1111/j.1365-313x.2012.05057.x>
 - 47) Tian, C., Zhang, Z., Huang, Y., Xu, J., Liu, Z., Xiang, Z., Zhao, F., Xue, J., Xue, T., & Duan, Y. (2022). Functional characterization of the *Pinellia ternata* cytoplasmic class II small heat shock protein gene PtsHSP17.2 via promoter analysis and overexpression in tobacco. *Plant Physiology and Biochemistry*, 177, 1–9. <https://doi.org/10.1016/j.plaphy.2022.02.017>
 - 48) Tissier, A. (2012). Trichome specific expression: promoters and their applications. *Transgenic plants-advances and limitations*, 17, 353-378.
 - 49) Tripp, J., Mishra, S. K., & Scharf, K. (2009). Functional dissection of the cytosolic chaperone network in tomato mesophyll protoplasts. *Plant, Cell & Environment/Plant, Cell and Environment*, 32(2), 123–133. <https://doi.org/10.1111/j.1365-3040.2008.01902.x>
 - 50) Ukaji, N., Kuwabara, C., Kanno, Y., Seo, M., Takezawa, D., Arakawa, K., & Fujikawa, S. (2010). Endoplasmic reticulum-localized small heat shock protein that accumulates in mulberry tree (*Morus bombycis* Koidz.) during seasonal cold acclimation is responsive to abscisic acid. *Tree Physiology*, 30(4), 502–513. <https://doi.org/10.1093/treephys/tpp125>

- 51) 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(4), 487–502. <https://doi.org/10.1007/s11103-005-0339-y>
- 52) 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>
- 53) Waters, E. R., Aebermann, B. D., & Sanders-Reed, Z. (2008). Comparative analysis of the small heat shock proteins in three angiosperm genomes identifies new subfamilies and reveals diverse evolutionary patterns. *Cell Stress & Chaperones*, 13(2), 127–142. <https://doi.org/10.1007/s12192-008-0023-7>
- 54) 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>
- 55) Wu, J., Gao, T., Hu, J., Zhao, L., Yu, C., & Ma, F. (2022). Research advances in function and regulation mechanisms of plant small heat shock proteins (sHSPs) under environmental stresses. *Science of the Total Environment*, 825, 154054. <https://doi.org/10.1016/j.scitotenv.2022.154054>
- 56) Yemets, A., Stelmakh, O., & Blume, Y. B. (2008). Effects of the herbicide isopropyl-N-phenyl carbamate on microtubules and MTOCs in lines of *Nicotiana sylvestris* resistant and sensitive to its action. *Cell Biology International*, 32(6), 623–629. <https://doi.org/10.1016/j.cellbi.2008.01.012>
- 57) Zafar, S., YanKun, Y., KeMing, Z., WeiJie, W., & XiaoLi, T. (2020). Overexpression of *nicotiana tabacum* HSP17.6 enhances abiotic stress tolerance in *brassica napus*. *International Journal of Agriculture and Biology*, 23(1), 164–170. <https://www.cabdirect.org/cabdirect/abstract/20203268831>

- 58) 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>