

ОЦІНКА ГЕНОМНОЇ ЕВОЛЮЦІЇ РОДИНИ ГЕНІВ ТУБУЛІНУ ДЛЯ ГЕНОТИПУВАННЯ ВИДІВ *CAMELINA*

Р.Я. БЛЮМ^{1,*}, А.М. РАБОКОНЬ¹, В.Г. САХАРОВА¹, Д.Б. РАХМЕТОВ², Я.В. ПІРКО¹, Я.Б. БЛЮМ¹

¹ Інститут харчової біотехнології та геноміки НАН України, Україна, 04123, Київ, вул. Байди-Вишневецького, 2А

² Національний ботанічний сад ім. М.М. Гришка НАН України, Україна, 01014, Київ, вул. Садово-Ботанічна, 1

E-mail: blume.rostislav@gmail.com

Автор для кореспонденції – Р.Я. Блюм, e-mail: blume.rostislav@gmail.com

Тубулін відіграє ключову роль у функціонуванні систем цитоскелету, котрі регулюють такі фундаментальні процеси як поділ та ріст клітини. Коректна ідентифікація ізотипів та визначення ортології генів тубуліну у рослин є нетривіальною задачею, котра вимагає залучення комплексу біоформатичних підходів. В ході проведеного дослідження було здійснено повногеномний пошук та ідентифікацію генів тубуліну у диплоїдних представників роду *Camelina*, зокрема у видів *C. neglecta*, *C. laxa*, *C. hispida*, що дозволило ідентифікувати повні набори генів α -, β - та γ -тубулінів, а також їх псевдогенів. Філогенетичний аналіз та серія повногеномних порівнянь дозволили встановити ортологію генів тубуліну, визначити ізотипову приналежність кодованих тубулінів та відслідкувати еволюційні зміни в наборах генів тубуліну в ході дивергенції видів та виникнення алогексаплоїдного *C. sativa*. Генотипування зразків різних представників *Camelina* за допомогою ТВР-, сТВР- та γ ТВР-маркерів дозволило ефективно диференціювати види на основі оцінки поліморфізму інтронних ділянок генів β - та γ -тубуліну. Одержані результати створюють міцне підґрунтя для подальших досліджень ізотипового та функціонального різноманіття тубулінів у Хрестоцвітних та інших груп квіткових рослин, а також сприяють розробленню та впровадженню нових високоефективних систем молекулярних маркерів для ДНК-баркодингу та маркер-опосередкованої селекції, в тому числі й таких перспективних олійних культур як *C. sativa*.

Ключові слова: *Camelina*, дикі родичі, тубулін, геномна еволюція, поліплоїдія, ДНК баркодинг, поліморфізм генів тубуліну.

Вступ. Поділ клітин є невід’ємною частиною росту та розмноження живих організмів. В еукаріотичних організмів цей процес значною мірою залежить від функціонування цитоскелетних систем, зокрема правильної організації мікротрубочок під час поділу клітини та формування веретена поділу (Lowe et al., 2001; Wickstead and Gull, 2011). Базовою структурною складовою мікротрубочок є білки тубуліни, котрі можуть бути представлені різними підродинами, зокрема α -, β - та γ -, або менш поширеними ϵ -, δ - чи ζ -тубулінами. Еволюційні причини та механізми появи такого різноманіття тубуліну у різних груп живих організмів наразі залишаються дискусійним питанням. У вищих рослин найбільш поширеними є α - та β -тубуліни (понад 80 % генів цих білків), а також γ -тубуліни, які є ключовою складовою центрів організації мікротрубочок (Demchuk and Blume, 2005; Breviario et al., 2013; Gavazzi et al., 2017). Оскільки вищі рослини в ході еволюції зазнали значної кількості подій повногеномної дуплікації, кількість та варіативність ізотипового набору тубулінів в них є вищою, ніж, наприклад, у тварин (Findeisen et al., 2014).

Кількість генів тубуліну у рослин різниться залежно від виду і в окремих випадках може сягати декількох десятків генів, особливо у складних поліплоїдних видів (Findeisen et al., 2014; Rao et al., 2016; Pydiura et al., 2019; Chen et al., 2021; Blume et al., 2024). Достеменно невідомо, чому надмірна кількість копій генів деяких ізотипів тубуліну не елімінується в процесі еволюції, а, навпаки, зберігає надзвичай-

но високу консервативність кодуєчих послідовностей. Нещодавно на прикладі *Arabidopsis thaliana* нами було показано, що поява основних класів ізотипів β -тубуліну пов'язана з відносно нещодавніми процесами повногеномної дуплікації, котрі передували дивергенції Хрестоцвітих, в той час як різноманіття генів α -тубуліну зумовлене локальними або транспозиційними дуплікаціями (Yemets et al., 2024). Також було показано, що у складного алогексаплоїдного виду *Camelina sativa* значна частина гомеологічних генів тубуліну збереглась у потроєнному стані, ймовірно, завдяки субфункціоналізації їх експресії (Jost et al., 2004; Blume et al., 2024). Вивчення механізмів субфункціоналізації на прикладі генів тубуліну у складних поліплоїдних видів є важливим, оскільки може дозволити поглибити розуміння балансування набору генів після кратного збільшення геному (Silflow et al., 1987; Gasanov et al., 2021).

Рід рижію (*Camelina*) привернув увагу великої кількості дослідників через високу перспективність рижію посівного (*C. sativa*) як альтернативної олійної культури для отримання рідких біопалив (Vollmann and Eynck, 2015; Berti et al., 2016). Оптимальний жирнокислотний склад олії насіння, короткий вегетаційний цикл та близькість цього виду до *A. thaliana* зумовили значний інтерес до використання рижію у геномних, біотехнологічних та селекційних дослідженнях (Faure and Terfer, 2016; Zanetti et al., 2021; Blume et al., 2022; Ghidoli et al., 2023). Більше того, рід *Camelina* часто розглядається як модельний для вивчення еволюції складних поліплоїдних видів (Mandáková et al., 2019; Blume et al., 2023). Зокрема, було показано, що лише три види цього роду є диплоїдними: *Camelina neglecta* (N⁶), *Camelina laxa* (L⁶) та *Camelina hispida* (H⁷). Усі інші представники цього роду є поліплоїдними нащадками з різними комбінаціями підгеномів, зокрема: *Camelina intermedia* nom. provis. (N⁶N⁷), *Camelina rumelica* (N⁶H⁷), *Camelina microcarpa* тип 1 (N⁶N⁷H⁷) та тип 2 (N⁶N⁷N⁶), *C. sativa* (N⁶N⁷H⁷) та *Camelina alyssum* (N⁶N⁷H⁷) (Žerdoner Čalasan et al., 2019; Mandáková et al., 2019; Chaudhary et al., 2020; Brock et al., 2022a,b; Mandáková and Lysak, 2022; Blume et al., 2023). На сьогодні наявні дані повногеном-

ного сиквенування диплоїдних видів рижію (*C. neglecta*, *C. laxa* та *C. hispida*), а також алогексаплоїдного рижію посівного (*C. sativa*), що дозволяє аналізувати походження родин генів у широкому еволюційному контексті (Kagale et al., 2014; Martin et al., 2022; Chaudhary et al., 2023; Wang et al., 2024).

Важливість вивчення родини тубуліну у різних видів обумовлюється також можливістю використання некодуєчих послідовностей цих генів, зокрема інтронів, у якості молекулярних маркерів (Breviario et al., 2013; Rabokon, 2021; Braglia et al., 2023). Існуючий на сьогодні метод оцінки поліморфізму довжини інтронів генів тубуліну (TBP – Tubulin based polymorphism) набув широкого використання в якості підходу для генотипування та ДНК-баркодингу різних груп рослин (Blume et al., 2020; Lykholat et al., 2022; Rabokon et al., 2023). За допомогою даного підходу вдається як диференціювати окремі види на основі їх молекулярно-генетичних профілів, так і розрізняти окремі лінії та сорти рослин, оцінюючи їх генетичні відмінності, що може бути успішно використано у селекції (Rabokon, 2021; Braglia et al., 2023). Нещодавно також було показано можливість ефективного генотипування представників *Camelina* з різною плоїдністю за допомогою TBP-методу (Galasso et al., 2015; Sakharova et al., 2023, 2025). Оцінка генетичного різноманіття як окультурених, так і диких представників *Camelina* набуває актуальності через необхідність інтрогресії диких алелів до селекційних ліній *C. sativa*, через недостатню поліморфність рижію посівного, котра, в свою чергу, значно обмежує можливості селекційного удосконалення цієї олійної культури (Manca et al., 2012; Blume et al., 2020, 2023).

Саме тому метою даної роботи було здійснення повногеномного пошуку та еволюційного аналізу родини генів тубуліну диплоїдних видів *C. neglecta*, *C. laxa*, *C. hispida* var. *hispida* та *C. hispida* var. *grandiflora*, спираючись на актуальні дані повногеномного сиквенування, а також оцінка еволюційної консервативності генів тубуліну в межах роду *Camelina* та демонстрація ефективності використання TBP-маркерів для генотипування цих видів.

Матеріали та методи. Пошук та ідентифікація генів. Для дослідження використовували

гени тубуліну, ідентифіковані у геномах чотирьох таксонів роду *Camelina*, зокрема, *C. neglecta*, *C. laxa*, *C. hispida* var. *hispida* та var. *grandiflora* (Martin et al., 2022; Chaudhary et al., 2023; Wang et al., 2024). Дослідження було проведено спираючись на найбільш актуальні доступні збірки геномів зазначених видів у базах даних NCBI Genomes (<https://www.ncbi.nlm.nih.gov/home/genomes/>) та NDGC Genome Warehouse (GWH) (<https://ngdc.cncb.ac.cn/gwh/>) (табл. 1). Геноми зазначених видів *Camelina* були анотовані *ab initio* за допомогою програмного забезпечення AUGUSTUS (Hoff and Stanke, 2013).

Гени ідентифікували за допомогою бібліотеки руHMMER (<https://pyhmmmer.readthedocs.io/en/stable/>) шляхом пошуку в масиві послідовностей трансльованих пептидів таких амінокислотних послідовностей, які б містили ННМ-профілі (з бази даних Pfam), характерні для тубулінів (Mistry et al., 2021; Larralde and Zeller, 2023). ННМ-пошук здійснювали з використанням домену PF00091.30 (Tubulin/FtsZ family, GTPase domain). Оскільки даний домен є спільним як для усіх тубулінів, так і для FtsZ білків, нами також додатково було проведено зворотній пошук консервативних доменів у послідовностей, ідентифікованих в результаті ННМ-пошуку. Ідентифікацію консервативних доменів, характерних для α -, β - та γ -тубулінів, здійснювали за допомогою InterPro (<https://www.ebi.ac.uk/interpro/>) (Payson-Lafosse et al., 2022), спираючись на інформацію з бази даних CDD (Wang et al., 2022). Гени усіх тубулінів нетипової довжини були переанотовані WebScipio (<https://www.webscipio.org/search>) (Natje et al., 2011) з метою виправлення можливих помилок анотації та/

або встановлення можливої причини псевдогенізації. При такій ручній переанотації як референтні використовували послідовності тубулінів арабідопсису (*Arabidopsis thaliana*). Дані про екзон-інтронну структуру генів тубуліну отримували, спираючись на AUGUSTUS-анотацію з урахуванням корекцій, зроблених шляхом WebScipio-анотації генів тубуліну в геномах зазначених видів *Camelina*. Додатково структура генів була верифікована за допомогою Gene Structure Display Server (<http://gsds.cbi.pku.edu.cn>) (Hu, 2015).

Реконструкція філогенії генів тубуліну. Трансльовані амінокислотні послідовності ідентифікованих генів α -, β - та γ -тубулінів були використані при реконструкції філогенії тубулінів у Хрестоцвітих. Також, у якості референтних були використані послідовності тубулінів *A. thaliana* та *A. lyrata*, а також ідентифіковані та описані раніше послідовності тубулінів *C. sativa* (Blume et al., 2024). Трансльовані амінокислотні послідовності були вирівняні за допомогою алгоритму MUSCLE (Edgar, 2004). У подальшому вирівняні послідовності були відредаговані з метою видалення переривчастих ділянок (гепів) та надмірно гетерогенних ділянок, котрі негативно впливають на точність реконструкції філогенії.

Філогенетичні дерева були побудовані з використанням веб-версії програмного забезпечення IQ-TREE (<http://iqtree.cibiv.univie.ac.at/>) (Nguyen et al., 2015; Trifinopoulos et al., 2016) за допомогою методу максимальної правдоподібності (ML - Maximum Likelihood) з підбором оптимальної моделі заміни за допомогою ModelFinder та з бустреп-підтримкою в 1000 ітерацій, обрахованою за допомогою UFBoot

Таблиця 1. Використані у дослідженні збірки геномів *Camelina*

Вид	ID збірки	Назва збірки	Джерело
<i>C. hispida</i> var. <i>hispida</i>	GCA_023657505.1*	ASM2365750v1	NCBI
<i>C. hispida</i> var. <i>grandiflora</i>	GCA_023864115.1	ASM2386411v1	NCBI
<i>C. laxa</i>	GCA_024034495.1*	ASM2403449v1	NCBI
<i>C. neglecta</i>	GCA_023864065.1*	ASM2386406v1	NCBI
<i>C. neglecta</i>	GCA_029034625.1	Cneglecta_1CN	NCBI
<i>C. neglecta</i>	GWHCBFX00000000	—	GWH

Примітка. * Збірка має статус референтної для виду.

(Kalyaanamoorthy et al., 2017; Hoang et al., 2018). Для набору трансльованих амінокислотних послідовностей α -тубуліну найоптимальнішою моделлю була визначена JTTDCMut+I, для β -тубулінів JTTDCMut+I+G4, а для γ -тубулінів – JTT. Отримані філогенетичні дерева були візуалізовані за допомогою програмного забезпечення iTOL v7 (<https://itol.embl.de/>) (Letunic and Bork, 2024). Отримані філогенетичні дерева були вкорінені за середньою точкою (midpoint). Класи в межах підродин α - та β -тубулінів були визначені на основі результатів проведеного філогенетичного аналізу, спираючись на усталену класифікацію (Oakley et al., 2007; Radchuk, 2008; Rao et al., 2016; Blume et al., 2024). Номенклатура генів (їх приналежність до ізотипів) переважно була також визначена на основі реконструйованої філогенії тубулінів. У випадку тубулінів з високою подібністю амінокислотних послідовностей застосовувався інший підхід, описаний нижче. Назви ідентифікованим генам тубулінів та їх класів присвоювали, спираючись на раніше запропоновану номенклатуру (Rao et al., 2016; Blume et al., 2024).

Повногеномне порівняння та аналіз синтенії. З метою аналізу еволюційного походження ідентифікованих генів тубуліну нами також було проведено повногеномне порівняння генів аналізованих видів за допомогою алгоритму синтенії MCScan, інтегрованого у бібліотеку JCVI (<https://github.com/tanghaibao/jcvi>) (Tang et al., 2024). Для повногеномного аналізу використовували кодуєчі послідовності генів досліджуваних геномів (табл. 1). Результати аналізу також були візуалізовані за допомогою JCVI. Ортологію цільових генів підтверджували в результаті окремого аналізу, котрий передбачав оцінку колінеарності геномного оточення (фланкуючих генів) цільового гена (Cheng et al., 2012). Для цього оцінювали подібність геномного оточення гена з досліджуваного геному (один з диплоїдних видів *Camelina*) з його гомологами (синтелогами) з референтного геному (*C. sativa*). Враховуючи значну еволюційну близькість видів, при порівнянні була використана однакова кількість генів з геномного оточення як для досліджуваного виду, так і для референтного – по 50 генів з 5'- та 3'-кінця. Коефіцієнт син-

тенної ортології обраховували як $(m+n)/100$, де m – кількість генів у 5'-сторони від аналізованого гена, що утворили пари з геномним оточенням гена референтного виду, n – з 3'-кінця відповідно, а 100 – загальна кількість генів з геномного оточення у досліджуваного виду. Таким чином було обраховано коефіцієнти синтенної ортології для усіх синтенних пар, утворених досліджуваним геном з генами референтного виду. Парою синтенних ортологів вважалась така пара, для якої значення коефіцієнту було найвищим. За результатами такого аналізу було прояснено ортологію та уточнено ізотипову приналежність генів тубуліну у диплоїдних видів *Camelina*.

Рослинний матеріал та генотипування за допомогою TBP-маркерів. Рослинний матеріал, використаний у дослідженні, було отримано від USDA National Plant Germplasm System (USDA-NPGS) (США) та від Національного ботанічного саду ім. М.М. Гришка НАН України (НБС). Повний перелік генотипів різних видів *Camelina* наведено у табл. 2. Геномна ДНК була виділена з проростків насіння зразків (приблизно по 10 рослин для кожного генотипу) за допомогою методу ЦТАБ (цетил-триметиламонія бромід) (Sambrook and David, 2001). Якість отриманої ДНК та її концентрацію визначали за допомогою електрофорезу в 1,5%-вому агарозному гелі та за допомогою спектрофотометру NanoDrop™ Lite (Thermo Scientific, USA). Одержані зразки ДНК зберігали за температури -20°C .

Для полімеразної ланцюгової реакції (ПЛР) використовували по 50 нг ДНК з кожної проби. Кожну ПЛР проводили щонайменше у двократній повторності (з використанням негативного контролю), що дозволяло підтвердити специфічність отриманих продуктів ампліфікації. ПЛР проводили в мікропробірках об'ємом 200 мкл за допомогою ампліфікатора Thermal Cycler 2720 («Applied Biosystems», США). Реакційна суміш загальним об'ємом 10 мкл містила п'ятикратний ПЛР-буфер з сульфатом амонію, 2,5 ммоль MgCl_2 , 50 нг рослинної ДНК, 1 мкМ кожного з праймерів, 0,2 мМ кожного дНТФ, 0,5 од. Таq полімерази («Fermentas», Литва).

Аналіз поліморфізму довжини 1-го та 2-го інтронів β -тубуліну проводили відповідно до

методики, описаної раніше (Braglia et al., 2023), оскільки інтрони генів β -тубуліну були продемонстровані як надійні маркери для *Camelina* (Galasso et al., 2015; Blume et al., 2020). Для ПЛР використовували пари вироджених праймерів, що фланкують 1-й (TBP-аналіз) та 2-й (сTBP-аналіз) інтрони β -тубуліну, а саме для TBP-аналізу (5'→3') – F: AACTGGGC-A-ARGGNCAYTAYAC; R: ACCATRCAYTCRT-CDGCRTTYTC, а для сTBP-аналізу (5'→3') – F: GARAAYGCHGAYGARTGYATG; R: CRAA-VCCBACCATGAARAARTG (Braglia et al., 2023). Також були використані вироджені праймери до найбільш консервативної ділянки γ -тубуліну, котра фланкує 2-й інтрон (у *A. thaliana*), для проведення γ TBP-аналізу (Pirko et al., 2018a): F: AYGTVTTTATTTACCARG-CKGA-3'; R: 5'-GAGTTGTARGGYTGGACR-AC-3'.

Ампліфікацію проводили відповідно до наступного протоколу: початкова денатурація (94 °C) – 3 хв, 40 циклів ампліфікації (денатурація 94 °C – 30 с, відпал праймерів 55 °C – 40 с, елонгація 72 °C – 1,5 хв), завершальна елонгація 72 °C – 8 хв, 15 °C – утримання (Braglia et al., 2023). Продукти ампліфікації розділяли за

допомогою електрофорезу в 6%-ному неденатуруючому поліакриламідному гелі в 1 × TBE-буфері (тривалість розділення фрагментів 2–3 год) (Sambrook and David, 2001) з подальшим фарбуванням нітратом срібла (Benbouza et al., 2006). Довжину фрагментів визначали за допомогою маркера довжини ДНК (O'Gene Ruler™ 100bp Plus DNA Ladder; «Thermo Fisher Scientific», США) (Braglia et al., 2023).

Аналіз та статистична обробка даних генотипування. Цифрові фотографії гелів аналізували за допомогою програми Gel Analyzer (gelanalyzer.software.informer.com/1.0/). Наявність амплікону певного розміру оцінювали за бінарною системою: 1 – наявний, 0 – відсутній. Отримані дані молекулярних маркерів були використані для подальшого філогенетичного аналізу. Оптимальні моделі для побудови дерева за допомогою методу максимальної правдоподібності (Maximum Likelihood, ML) були визначені за допомогою ModelFinder (Kalyanamoorthy et al., 2017). Для набору використаних маркерів оптимальною моделлю була GTR2+FO. Філогенетичний аналіз (ML) проводили за допомогою веб-версії інструменту IQ-TREE (<http://iqtree.cibiv.univie.ac.at>) (Ngu-

Таблиця 2. Перелік видів та генотипів *Camelina* використаних у дослідженні

Вид <i>Camelina</i>	ID зразка	Місце походження	Колекція
<i>C. microcarpa</i> тип 1*	PI633191	Монтана, США	USDA-NPGS
<i>C. microcarpa</i> тип 1*	Ames 31219	Грузія	USDA-NPGS
<i>C. microcarpa</i> тип 2*	PI650136	Німеччина	USDA-NPGS
<i>C. microcarpa</i> тип 2*	PI633186	Угорщина	USDA-NPGS
<i>C. neglecta</i>	PI650135**	Франція	USDA-NPGS
<i>C. microcarpa</i> (перевизначено як <i>C. rumelica</i>)	PI650134	Іспанія	USDA-NPGS
<i>C. microcarpa</i> тип 2*	PI633187	Польща	USDA-NPGS
<i>C. rumelica</i>	PI650138	Іран	USDA-NPGS
<i>C. hispida</i> var. <i>grandiflora</i>	PI650133**	Туреччина	USDA-NPGS
<i>C. alyssum</i>	PI650132	Німеччина	USDA-NPGS
<i>C. laxa</i>	Ames 32852	Вірменія	USDA-NPGS
<i>C. sativa</i> (сорт Suneson)	—	США	НБС
<i>C. sativa</i> (сорт Клондайк)	03097001	Україна	НБС
<i>C. sativa</i> (сорт Міраж)	94097001	Україна	НБС

Примітка.* Дані про цитотипову приналежність зразків отримано з попередніх досліджень (Chaudhary et al., 2020; Brock et al., 2022a,b). ** Ці зразки були використані для повногеномного секвенування у попередніх дослідженнях (Martin et al., 2022; Chaudhary et al., 2023; Wang et al., 2024)

yen et al., 2015; Trifinopoulos et al., 2016) з бутстреп-підтримкою 1000, використовуючи UFBoot для ультрашвидкого бутстрепа (Hoang et al., 2018). Отримані філогенетичні дерева візуалізували за допомогою веб-версії iTOL v7 (<https://itol.embl.de>) (Letunic and Bork, 2024).

Для визначення популяційної структури та генетичного різноманіття було використано програму STRUCTURE 2.2.3 (Hubisz et al., 2009). Для аналізу були встановлені такі параметри: 10^5 ітерацій випалювання (burn-in) та 2×10^5 ітерацій для обрахунку Марківських ланцюгів Монте-Карло (Markov Chain Monte Carlo, MCMC). Передбачувана кількість популяцій (K) варіювала від 1 до 10. Найімовірніше значення K було визначено за допомогою StructureSelector (<https://lmm.ac.cn/StructureSelector/>) (Raj et al., 2014; Li and Liu, 2018). Цей підхід базується на використанні методу ΔK , який є спеціальним статистичним критерієм, що оцінює швидкість зміни логарифмічної ймовірності даних при різній кількості кластерів, відповідно до опису Evanno et al. (2005). Для виявлення структури популяцій було також використано додатковий метод відповідно до опису Puechmaille et al. (2016). Для візуалізації результатів STRUCTURE використовували програмне забезпечення Clumpak, інтегроване в StructureSelector (Kopelman et al., 2015). Аналіз основних координат та відповідна діаграма PCA була обчислена у R за допомогою бібліотеки adegenet та візуалізована з використанням бібліотеки ggplot2.

Результати досліджень та їх обговорення.

За результатами проведеного повногеномного пошуку нами було виявлено, що диплоїдні види *Camelina*, зокрема *C. neglecta*, *C. laxa*, *C. hispida* var. *hispida* та *C. hispida* var. *grandiflora*, мають від 5 до 6 генів α -тубуліну (табл. 3). Номенклатуру генів визначали на основі приналежності передбачуваних білкових продуктів ідентифікованих генів до конкретних ізотипів та їх подібності до ідентифікованих раніше генів *C. sativa* (Blume et al., 2024). Розмір усіх виявлених генів α -тубуліну варіював за рахунок зміни довжини некодуючих ділянок (інтронів), в той час як кодуєча частина цих генів вирізнялась значною консервативністю. Кількість інтронів варіювала залежно від класу α -тубулінів. Зокрема, гени ізотипів TUA1,

TUA3, TUA5 вирізнялись наявністю 4-х інтронів, а гени ізотипів TUA2, TUA4, TUA6 мали по 3 інтрони в усіх випадках. Подібна клас-специфічна відмінність у екзон-інтронній структурі генів α -тубуліну спостерігається також і у ряду інших видів (Rao et al., 2016; Blume et al., 2024).

Слід також зазначити, що у *C. hispida* var. *hispida* та *C. hispida* var. *grandiflora* був відсутнім ізотип TUA1. Найвірогідніше, що цей ген був втрачений в ході численних крупномасштабних перебудов геному, яких зазнав вид *C. hispida* протягом еволюційної дивергенції (Mandáková et al., 2019). Це також може пояснити відсутність гену ізотипу TUA1 у третьому С-підгеномі *C. sativa* (Blume et al., 2024). Слід також зазначити, що у збірці GCA_029034625.1 геному *C. neglecta* також був втрачений ген *CnTUA1*. Однак, втрата даного гену з високою вірогідністю є артефактом збірки, оскільки ген *CnTUA1* наявний у двох інших збірках (GCA_023864065.1, GWHCBFX000000000). На підтримку цього також свідчить той факт, що усі три збірки геному *C. neglecta* були отримані в результаті секвенування одного й того ж генотипу – PI650135 (Martin et al., 2022; Chaudhary et al., 2023; Wang et al., 2024).

Використовуючи трансльовані амінокислотні послідовності, кодовані ідентифікованими генами, нами було реконструйовано філогенію α -тубулінів (рис. 1). Послідовності описаних раніше α -тубулінів *C. sativa*, *A. thaliana*, *A. lyrata* були використані як референтні (Blume et al., 2024). При філогенетичній реконструкції було використано 60 амінокислотних послідовностей, для яких 82,52 % сайтів були неваріативними, що підтверджує високу консервативність α -тубулінів, принаймні, в межах аналізованих родів *Camelina* та *Arabidopsis*. За результатами філогенетичного аналізу було виявлено наявність двох класів ізотипів α -тубуліну, при чому кожен з аналізованих видів характеризувався наявністю генів α -тубуліну кожного з класів. Ідентифіковані α -тубуліни *C. neglecta*, *C. laxa*, *C. hispida* var. *hispida* та *C. hispida* var. *grandiflora* групувались відповідно до їх ортологічної приналежності, зокрема, часто поділяючи спільні клади з білками *C. sativa*.

Клас I, яких включає α -тубуліни ізотипів TUA2, TUA4 та TUA6, був виділений в окре-

му кладу з високим рівнем бутстреп-підтримки (понад 80 %). Слід зазначити, що α -тубуліни класу I вирізняються високою подібністю послідовностей. Характерною особливістю генів α -тубулінів класу I є консервативна структура

генів, яка передбачає наявність трьох інтронів. Висока подібність генів класу I добре узгоджується з їх еволюційною історією, оскільки ізотипове різноманіття цієї гілки α -тубулінів у Хрестоцвітих найімовірніше могло виникнути

Таблиця 3. Перелік ідентифікованих генів α -тубуліну, ідентифікованих у геномах диплоїдних видів *Camelina*

Назва гена	Локалізація, ID хромосоми та координати у п.н.	Орієнтація	Кількість інтронів
<i>C. hispida</i> var. <i>grandiflora</i>			
<i>ChvgTUA2</i>	CM043447.1:33431111..33433062	+	3
<i>ChvgTUA6</i>	CM043448.1:44001789..44003607	—	3
<i>ChvgTUA4</i>	CM043449.1:12635199..12637131	+	4
<i>ChvgTUA3</i>	CM043453.1:12494485..12496213	—	4
<i>ChvgTUA5</i>	CM043453.1:12506319..12508047	+	4
<i>C. hispida</i> var. <i>hispida</i>			
<i>ChvhTUA2</i>	CP094631.1:33431111..33433062	+	3
<i>ChvhTUA6</i>	CP094632.1:44001789..44003607	—	3
<i>ChvhTUA4</i>	CP094633.1:12635199..12637131	+	3
<i>ChvhTUA3</i>	CP094637.1:12494485..12496213	—	4
<i>ChvhTUA5</i>	CP094637.1:12506319..12508047	+	4
<i>C. laxa</i>			
<i>CITUA3</i>	CM043881.1:8105807..8107530	—	4
<i>CITUA5</i>	CM043881.1:8110875..8112589	+	4
<i>CITUA6</i>	CM043883.1:1421015..1422827	+	3
<i>CITUA2</i>	CM043884.1:4743129..4745086	—	3
<i>CITUA4</i>	CM043885.1:4637940..4639853	+	3
<i>CITUA1</i>	JALLKC010000234.1:40781..42804	—	4
<i>C. neglecta</i> ¹			
<i>CnTUA2</i>	CM043455.1:25646663..25648636	+	3
<i>CnTUA1</i>	CM043456.1:12766171..12768216	+	4
<i>CnTUA6</i>	CM043456.1:31367289..31369129	—	3
<i>CnTUA4</i>	CM043457.1:9301264..9303203	+	3
<i>CnTUA3</i>	CM043460.1:8661640..8663357	—	4
<i>CnTUA5</i>	CM043460.1:8666835..8668552	+	4
<i>C. neglecta</i> ²			
<i>CnTUA2</i>	CM054955.1:24373567..24375540	+	3
<i>CnTUA6</i>	CM054956.1:31136218..31138058	—	3
<i>CnTUA4</i>	CM054957.1:9279935..9281874	+	3
<i>CnTUA3</i>	CM054959.1:8630872..8632589	—	4
<i>CnTUA5</i>	CM054959.1:8636070..8637787	+	4
<i>C. neglecta</i> ³			
<i>CnTUA6</i>	GWHCBBFX00000002:1491583..1493423	+	4
<i>CnTUA1</i>	GWHCBBFX00000002:20484244..20486289	—	3
<i>CnTUA2</i>	GWHCBBFX00000003:4929458..4931431	—	3
<i>CnTUA3</i>	GWHCBBFX00000005:18678862..18680579	—	4
<i>CnTUA5</i>	GWHCBBFX00000005:18684061..18685779	+	4
<i>CnTUA4</i>	GWHCBBFX00000006:9273481..9275420	+	3

Примітка. Збірка: ¹GCA_023864065.1 (референтна); ²GCA_029034625.1; ³GWHCBBFX000000000

ISSN 0564-3783. Цитологія і генетика. 2025. Т. 59. № 5

Таблиця 4. Перелік генів β -тубуліну, ідентифікованих у геномах диплоїдних видів *Camelina*

Назва гена	Локалізація, ID хромосоми та координати у п.н.	Орієнтація	Кількість інтронів
<i>C. hispida</i> var. <i>grandiflora</i>			
<i>ChvgTUB5</i>	CM043447.1:9941593..9943997	—	2
<i>ChvgTUB1</i>	CM043448.1:38003405..38005530	—	2
<i>ChvgTUB</i>	CM043450.1:9174266..9176243	+	2
<i>ChvgTUB9</i>	CM043451.1:15624579..15626144	—	2
<i>ChvgTUB3-1</i>	CM043452.1:31840640..31842552	+	2
<i>ChvgTUB3-2</i>	JALGBY010000009.1:2786695..2788607*	—	2
<i>ChvgTUB8</i>	CM043453.1:17070739..17072282	+	2
<i>ChvgTUB4</i>	CM043453.1:31079413..31081020	—	2
<i>ChvgTUB7</i>	CM043450.1:16735114..16737075	—	2
<i>C. hispida</i> var. <i>hispida</i>			
<i>ChvhTUB5</i>	CP094631.1:9941593..9943997	—	2
<i>ChvhTUB1</i>	CP094632.1:38003405..38005530	—	2
<i>ChvhTUB</i>	CP094634.1:9174266..9176243	+	2
<i>ChvhTUB9</i>	CP094635.1:15624579..15626144	—	2
<i>ChvhTUB3</i>	CP094636.1:31840640..31842552	+	2
<i>ChvhTUB8</i>	CP094637.1:17070739..17072282	+	2
<i>ChvhTUB4</i>	CP094637.1:31079413..31081020	—	2
<i>ChvhTUB7</i>	CP094634.1:16735114..16737075	—	2
<i>C. laxa</i>			
<i>CITUB9</i>	CM043880.1:14670222..14671788	—	2
<i>CITUB6</i>	CM043881.1:4606900..4608526	—	2
<i>CITUB8</i>	CM043881.1:10245936..10247467	+	2
<i>CITUB4</i>	CM043881.1:32297304..32298905	—	2
<i>CITUB</i>	CM043882.1:5933487..5935547	+	2
<i>CITUB2</i>	CM043882.1:14551795..14553758	+	2
<i>CITUB1</i>	CM043883.1:3799963..3802120	—	2
<i>CITUB7</i>	CM043883.1:26296966..26298786	—	2
<i>CITUB5</i>	CM043884.1:23051194..23053622	+	2
<i>CITUB3</i>	CM043885.1:26967372..26969068	+	2
<i>C. neglecta</i> ¹			
<i>CnTUB5</i>	CM043455.1:9306194..9308594	—	2
<i>CnTUB7</i>	CM043456.1:6496434..6498463	—	2
<i>CnTUB1</i>	CM043456.1:24501821..24503950	—	2
<i>CnTUB2</i>	CM043458.1:13701263..13703283	—	2
<i>CnTUB</i>	CM043458.1:21661464..21663454	—	2
<i>CnTUB9</i>	CM043459.1:10810963..10812550	—	2
<i>CnTUB4</i>	CM043459.1:32510300..32511919	—	2
<i>CnTUB3</i>	CM043459.1:45772269..45774258	+	2
<i>CnTUB6</i>	CM043460.1:4761951..4763602	—	2
<i>CnTUB8</i>	CM043460.1:11284398..11285919	+	2
<i>C. neglecta</i> ²			
<i>CnTUB5</i>	CM054955.1:8373808..8376208	—	2
<i>CnTUB7</i>	CM054956.1:6441144..6443173	—	2
<i>CnTUB1</i>	CM054956.1:24267579..24269708	—	2
<i>CnTUB2</i>	CM054958.1:13266463..13268483	—	2

Назва гена	Локалізація, ID хромосоми та координати у п.н.	Орієнтація	Кількість інтронів
<i>C. neglecta</i> ²			
<i>CnTUB</i>	CM054958.1:21252543..21254534	—	2
<i>CnTUB6</i>	CM054959.1:4729574..4731225	—	2
<i>CnTUB8</i>	CM054959.1:11254727..11256248	+	2
<i>CnTUB9</i>	CM054960.1:10833643..10835230	—	2
<i>CnTUB4</i>	CM054960.1:32268480..32270103	—	2
<i>CnTUB3</i>	CM054960.1:45509157..45511146	+	2
<i>C. neglecta</i> ³			
<i>CnTUB9</i>	GWHCBFX00000001:11007930..11009517	—	2
<i>CnTUB4</i>	GWHCBFX00000001:32877656..32879279	—	2
<i>CnTUB3</i>	GWHCBFX00000001:46273698..46275687	+	2
<i>CnTUB1</i>	GWHCBFX00000002:8360801..8362930	+	2
<i>CnTUB7</i>	GWHCBFX00000002:26762328..26764357	+	2
<i>CnTUB5</i>	GWHCBFX00000003:21086390..21088790	+	2
<i>CnTUB</i>	GWHCBFX00000004:6799075..6801065	+	2
<i>CnTUB2</i>	GWHCBFX00000004:14764837..14766857	+	2
<i>CnTUB8</i>	GWHCBFX00000005:16059610..16061131	—	2
<i>CnTUB6</i>	GWHCBFX00000005:22585655..22587306	+	2

Примітка. Збірка: ¹GCA_023864065.1 (референтна); ²GCA_029034625.1; ³GWHCBFX00000000. * Потенційно артефакт збірки геному.

C. laxa було виявлено по десять генів β-тубуліну, які відображають повне ізотипове різноманіття цих білків, відоме для представників Хрестоцвітих (Blume et al., 2024). Локалізація відповідних генів на хромосомах дещо відрізняється у різних релізах геномів *C. neglecta*, що, однак, може бути пов'язано, в першу чергу, з особливостями процедури процесингу сиквенсів при збиранні геномів.

На відміну від інших видів, представники *C. hispida* мали дещо меншу кількість генів β-тубуліну. Зокрема, у обох таксонів var. *hispida* та var. *grandiflora* було виявлено втрату функціональних генів ізотипу TUB2 та TUB6. При цьому у геномі *C. hispida* var. *grandiflora* було виявлено ознаки можливої дуплікації гена ізотипу TUB3, представленого генами *ChvgTUB3-1* та *ChvgTUB3-2*. Однак, слід зазначити, що дуплікований ген *ChvgTUB3-2* локалізується на нерозміщеному скефолді JALGBY010000009.1 та мав повністю ідентичний сиквенс до *ChvgTUB3-1*. Враховуючи це, можна припустити, що ген *ChvgTUB3-2* потенційно може бути артефактом збірки геному та копією регіону,

котрий включає *ChvgTUB3-1*. Однак, підтвердження цього припущення вимагає подальших досліджень, які залучали б більшу кількість даних повногеномного секвенування *C. hispida* var. *grandiflora*.

Окрім ідентифікації генів повноцінних генів β-тубуліну нами також були ідентифіковані псевдогени, виявлені у геномах *C. hispida* var. *grandiflora* та *C. hispida* var. *hispida* (табл. 5). Більшість ідентифікованих псевдогенів характеризувалися наявністю передчасних стоп-кодів, мутацій, зсуву рамки читування, внаслідок чого могла спостерігатись фрагментація екзонів або втрата окремих окремих кодуючих ділянок. Найімовірніше, такі псевдогени не кодують функціонального білкового продукту через частковість сиквенсів β-тубулінів. Наприклад, псевдогени *ChvgTUB2p-1*, *ChvgTUB2p-2*, *ChvgTUB4p-1*, *ChvgTUB4p-2* (та їх ортологи *ChvhTUB2p-1*, *ChvhTUB2p-2*, *ChvhTUB4p-1*, *ChvhTUB4p-2*) виникли в результаті дуплікації функціональних генів та, імовірно, були псевдогенізовані через відсутність еволюційної переваги у більшій копійності генів цих ізотипів.

Слід також відзначити, що *ChvgTUB2p-1*, *ChvgTUB2p-2* (та їх ортологи *ChvhTUB2p-1*, *ChvhTUB2p-2*) виникли як тандемні дуплікації. При цьому жодного функціонального гена ізо типу TUB2 не залишилось у геномах *C. hispida* var. *grandiflora* та *C. hispida* var. *hispida*, однак еволюційно близькі гени ізо типу TUB3 збереглись у цих видів. Аналогічно, у обох таксонів *C. hispida* були втрачені гени ізо типів TUB1 та TUB6. Оскільки набори псевдогенів у var. *grandiflora* та var. *hispida* є ідентичними (табл. 5), слід припустити, що втрата функціональних варіантів цих генів відбулась до еволюційної дивергенції внутрішньовидових таксонів *C. hispida*. Цікавим також є той факт, що третій (C) підгеном *C. sativa*, котрий, імовірно, походить від *C. hispida* характеризується також високою кількістю псевдогенів β -тубуліну (Blume et al., 2024).

Спираючись на трансльовані амінокислотні послідовності ідентифікованих β -тубулінів нами було реконструйовано філогенію цієї підродини білків (рис. 2). Послідовності описаних раніше β -тубулінів *C. sativa*, *A. thaliana*, *A. lyrata* були використані у якості референтних (Blume et al., 2024). При філогенетичній

реконструкції було використано 102 амінокислотні послідовності, у яких 84,69 % сайтів були консервативними. Проведений філогенетичний аналіз дозволив виявити наявність чотирьох окремих класів ізо топів. Різноманіття β -тубулінів у аналізованих видів були представлено ізо типами майже усіх класів, окрім *C. hispida*, геноми якого не мали генів, характерних для класу I. В той же час *C. hispida* var. *hispida* та var. *grandiflora* зазнали втрати функціональних генів ізо типу TUB6 через їх псевдогенізацію (*ChvgTUB6p* та *ChvhTUB6p*). Клас II, який включає ізо типи TUB2, TUB3, TUB7 та TUB8, був найбільш численною групою β -тубулінів, широко представленою у всіх аналізованих видів, навіть враховуючи втрату генів ізо типу TUB2 у *hispida* var. *hispida* та var. *grandiflora*. Клас III був представлений білками TUB4, TUB9 та TUB (TUB10), а клас IV – ізо типами TUB1 та TUB5 (рис. 2).

Слід також зазначити, що представники TUB (TUB10) не виділялись в окрему підгрупу в межах класу III, як це було показано у попередніх філогенетичних реконструкціях (Lykholat et al., 2022; Blume et al., 2024). Наразі походження цього ізо типу є недостатньо вивченим

Таблиця 5. Перелік псевдогенів β -тубуліну, ідентифікованих у геномах диплоїдних видів *Camelina*

Назва гена	Локалізація, ID хромосоми та координати у п.н.	Орієнтація	Причина псевдогенізації
<i>ChvgTUB1p</i>	CM043447.1:9938047..9940256	—	Передчасний стоп-кодон у 3-му екзоні
<i>ChvgTUB2p-1</i>	CM043448.1:19854890..19856741	—	Передчасний стоп-кодон у 2-му екзоні
<i>ChvgTUB2p-2</i>	CM043448.1:20101858..20107966	—	Два зсуви рамки зчитування, утворення додаткових екзонів (розрив 2-го екзону)
<i>ChvgTUB4p-1</i>	CM043449.1:6528442..6529554	—	Передчасний стоп-кодон, численні делеції у фланкуючих регіонах
<i>ChvgTUB4p-2</i>	CM043452.1:30801704..30802801	—	Передчасний стоп-кодон, численні делеції у фланкуючих регіонах
<i>ChvgTUB6p</i>	CM043453.1:5712870..5714449	—	Зсув рамки зчитування у 3-му екзоні
<i>ChvhTUB1p</i>	CP094631.1:9938047..9940256	—	Передчасний стоп-кодон у 3-му екзоні
<i>ChvhTUB2p-1</i>	CP094632.1:19854890..19856741	—	Передчасний стоп-кодон у 2-му екзоні
<i>ChvhTUB2p-2</i>	CP094632.1:20101858..20107966	—	Два зсуви рамки зчитування, утворення додаткових екзонів (розрив 2-го екзону)
<i>ChvhTUB4p-1</i>	CP094633.1:6528442..6529554	—	Передчасний стоп-кодон, численні делеції у фланкуючих регіонах
<i>ChvhTUB4p-2</i>	CP094636.1:30801704..30802801	—	Передчасний стоп-кодон, численні делеції у фланкуючих регіонах
<i>ChvhTUB6p</i>	CP094637.1:5712870..5714449	—	Зсув рамки зчитування у 3-му екзоні

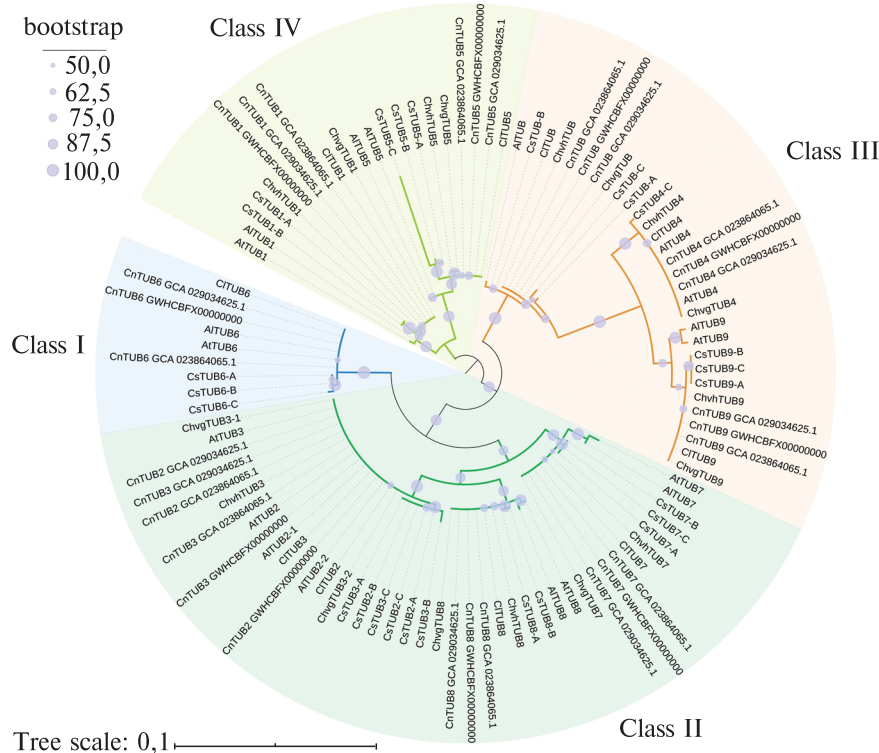


Рис. 2. Філогенетичне дерево β-тубулінів диплоїдних представників *Camelina* та *C. sativa*, *A. thaliana*, *A. lyrata*, побудоване за допомогою методу ML з бутстреп-підтримкою у 1000 ітерацій. Філогенетичне дерево вкорінено за середньою точкою (midpoint)

та потребує подальших досліджень. Цікавим є те, що даний ізотип β-тубуліну втрачений у модельному організмі *A. thaliana*, але зберігся у його близького родича *A. lyrata*. Втрата гена ізотипу TUB (TUB10) може бути пов'язана з тим, що *A. thaliana* в ході еволюції зазнав значної реорганізації геному, котра супроводжувалась втратою окремих його ділянок (Hu et al., 2011; Jiao et al., 2020).

Аналогічно, дискусійним питанням залишається і кількість класів ізотипів β-тубуліну. Значна частина філогенетичних реконструкцій еволюції β-тубулінів у квіткових рослин демонструє наявність лише 4-х класів ізотипів (Rao et al., 2016; Gavazzi et al., 2017; Blume et al., 2024). Однак у ряді робіт вказується на наявність додаткового п'ятого (під)класу β-тубулінів, який включає або еволюційно віддалені ізотипи, або високоспеціалізовані ізотипи класу I (Radchuk, 2008; Oakley et al., 2007; Lykholat et al., 2022). Часто ця група включає окре-

мі ізотипи β-тубулінів представників *Populus*, *Salix*, *Medicago*, *Prunus*, *Eucalyptus*, тощо (Oakley et al., 2007; Lykholat et al., 2022).

Здійснений повногеномний пошук дозволив встановити наявність від двох до трьох генів γ-тубуліну у кожного з диплоїдних представників *Camelina* (табл. 6). Усі ідентифіковані гени γ-тубуліну мали високонсервативну структуру, котра складалась з десяти екзонів та дев'яти інтронів. Ідентифіковані гени кодували γ-тубуліни двох ізотипів, котрі наявні й у інших видів Хрестоцвітих (Blume et al., 2024). У двох таксонів *C. hispida* було виявлено по три гени γ-тубулінів, зокрема було наявно по дві копії генів ізотипу TUG1 (*ChvgTUG1-1*, *ChvgTUG1-2* та *ChvhTUG1-1*, *ChvhTUG1-2*). Наявність двох копій гена на доволі невеликій відстані вказує на видоспецифічну сегментальну (можливо тандемну) дуплікацію регіону, котрий містить в собі ці гени. При цьому у геномах *C. neglecta* та *C. laxa* було виявлено по одному гену

кожного з ізотипів: *CITUG1*, *CITUG2*, *CnTUG1* та *CnTUG2*. Наявність лише двох генів γ -тубуліну у *C. neglecta* підтверджується у всіх трьох геномних збірках.

Також нами було проведено філогенетичний аналіз, що дозволило реконструювати еволюційні взаємозв'язки ідентифікованих γ -тубулінів диплоїдних представників *Camelina* з гомологічними білками інших видів (рис. 3). При реконструкції отриманого філогенетичного дерева послідовності описаних раніше γ -тубулінів *C. sativa*, *A. thaliana*, *A. lyrata* були використані як референтні. Результати проведеного аналізу 24-х послідовностей γ -тубуліну різних ізотипів дозволили виявити надзвичайно високий рівень консервативності амінокислотних послідовностей цих білків. Так, кількість консервативних неваріабельних сайтів становила 96,41 % у проаналізованих γ -тубулінів, що добре узгоджується з даними попередніх досліджень (Blume et al., 2024).

Реконструйоване філогенетичне дерево чітко демонструє розділення γ -тубулінів на два різні ізотипи. Наразі для γ -тубуліну не було виявлено поділу на класи за ізотипами через обмежене різноманіття цього білка, його високу консервативність та обмежену кількість генів у більшості видів квіткових рослин (Find-eisen et al., 2014). При цьому, варто зазначити, що розділення γ -тубулінів на окремі ізотипи характерне лише для обмеженої кількості видів рослин, зокрема лише для частини Хрестоцвітних, саме триби Camelineae (Blume et al., 2024). Слід також зазначити, що хоча такі види, як *C. neglecta* та *C. laxa*, мали по одному гену γ -тубуліну кожного з ізотипів, представники *C. hispida* мали по три відповідні гени у геномах. Білки ChvgTUG1-1, ChvgTUG1-2 та ChvhTUG1-1, ChvhTUG1-2 групувались в окремий кластер, що підтверджує їх спільне походження в результаті дуплікації. Цікавим є той факт, що, незважаючи на ідентичні кодо-

Таблиця 6. Перелік генів γ -тубуліну, ідентифікованих у геномах диплоїдних видів *Camelina*

Назва гена	Локалізація, ID хромосоми та координати у п.н	Орієнтація	Кількість інтронів
<i>C. hispida</i> var. <i>grandiflora</i>			
<i>ChvgTUG1-1</i>	CM043450.1:42322655..42325059	—	9
<i>ChvgTUG1-2</i>	CM043450.1:42473649..42476115	—	9
<i>ChvgTUG2</i>	CM043453.1:2256433..2258756	+	9
<i>C. hispida</i> var. <i>hispida</i>			
<i>ChvhTUG1-1</i>	CP094634.1:42322655..42325059	—	9
<i>ChvhTUG1-2</i>	CP094634.1:42473649..42476115	—	9
<i>ChvhTUG2</i>	CP094637.1:2256433..2258756	+	9
<i>C. laxa</i>			
<i>CITUG2</i>	CM043881.1:1971649..1973993	+	9
<i>CITUG1</i>	CM043883.1:30400860..30403422	—	9
<i>C. neglecta</i> ¹			
<i>CnTUG1</i>	CM043456.1:1073671..1076257	+	9
<i>CnTUG2</i>	CM043460.1:1998555..2000961	+	9
<i>C. neglecta</i> ²			
<i>CnTUG1</i>	CM054956.1:1015691..1018277	+	9
<i>CnTUG2</i>	CM054959.1:1964919..1967325	+	9
<i>C. neglecta</i> ³			
<i>CnTUG1</i>	GWHCBFX000000002:32194783..32197369	—	9
<i>CnTUG2</i>	GWHCBFX000000005:25350036..25352442	—	9

Примітка. Збірка: ¹GCA_023864065.1 (референтна); ²GCA_029034625.1; ³GWHCBFX000000000.

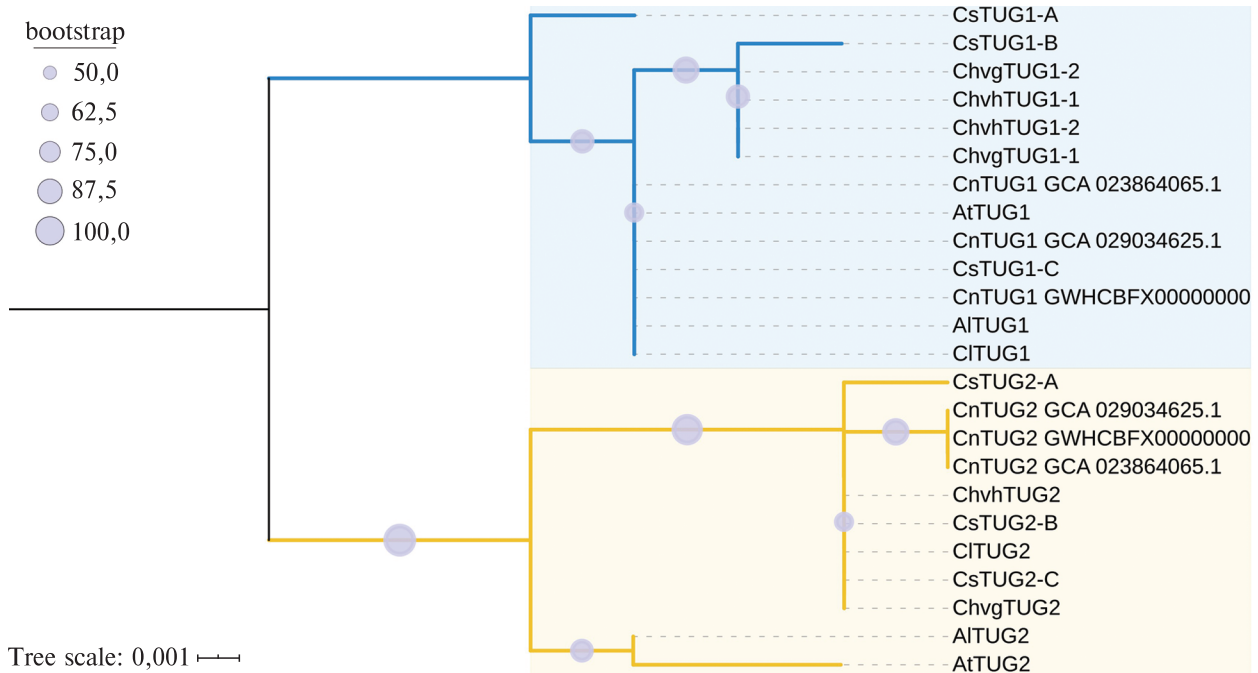


Рис. 3. Філогенетичне дерево γ -тубулінів диплоїдних представників *Camelina* та *C. sativa*, *A. thaliana*, *A. lyrata*, побудоване за допомогою методу ML з бутстреп-підтримкою у 1000 ітерацій. Філогенетичне дерево вкорінено за середньою точкою (midpoint)

вані амінокислотні послідовності цих генів у *C. hispida* var. *hispida* та var. *grandiflora*, усі дупліковані копії TUG1 збереглися в ході еволюції. Причини такої консервації збільшеної кількості генів або їх субфункціоналізації потребують подальшого більш детального дослідження.

З метою подальшої оцінки геномного контексту та еволюційної консервативності ідентифікованих генів тубуліну нами було здійснено порівняння диплоїдних геномів *C. neglecta* (N⁶) та *C. hispida* (H⁷) з підгеномним алогексплоїдного виду *C. sativa* (N⁶N⁷H⁷), для якого *C. neglecta* та *C. hispida* є предковими видами (рис. 4). Синтенія α -тубулінів предкових диплоїдних видів та алогексаплоїдного виду-нащадка чітко простежується на рис. 4, а, що підкреслює високий рівень консервативності цих генів. Більшість генів α -тубуліну була напряму успадкована алогексаплоїдним видом *C. sativa*, окрім псевдогенізованого *CsTUA5p-A* у першому підгеномі (Blume et al., 2024). При цьому, навіть для псевдогенізованого *CsTUA5p-A* були виявлені синтенні поєднання з локуса-

ми ортологічних функціональних генів α -тубуліну у геномах предкових видів. Подібний патерн збереження набору генів α -тубуліну навіть після поліплоїдизації є доволі цікавим та може слугувати додатковим доказом ортологічності досліджуваних генів різних ізотипів α -тубуліну у видів *Camelina*. Найвірогідніше, збереження великої кількості копій α -тубулінів у геномі *C. sativa* може вказувати на те, що ці гени зазнали субфункціоналізації шляхом диференціації патернів експресії (Blume et al., 2024).

Натомість, синтенія генів β -тубуліну була значно складнішою (рис. 4, б), оскільки в процесі балансування набору генів після поліплоїдизації та виникнення алогексаплоїдного геному *C. sativa* значна кількість копій генів β -тубуліну була втрачена. Слід також зазначити, що часто гени, котрі представляють різні ізоטיפи β -тубуліну одного й того ж класу (напр. TUB1 та TUB5), могли утворювати синтенні поєднання. Це пов'язано з тим, що різні ізоטיפи тубуліну в межах окремих класів (TUB1 та TUB5 для класу IV, а також TUB4 та TUB9 для

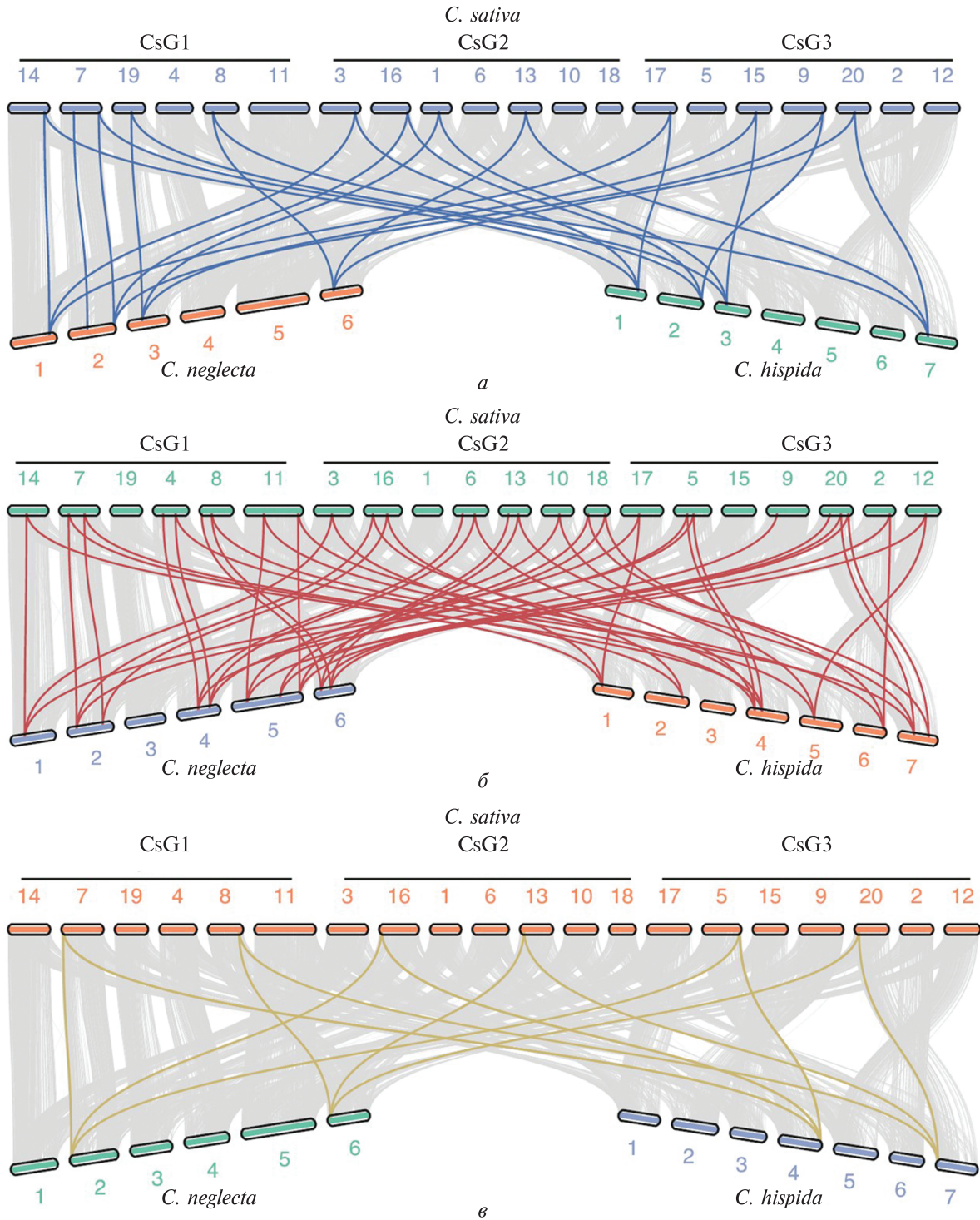


Рис. 4. Синтенія між підгеномами алогексаплоїдного виду *C. sativa* ($N^6N^7H^7$) та геномами його предкових видів *C. neglecta* (N^6) та *C. hispida* var. *hispida* (H^7): (а) – синім кольором виділені поєднання між синтелогічними генами α -тубулінів, (б) – червоним – синтелогои β -тубулінів, (в) – жовтим – синтелогои γ -тубулінів. CsG1 – перший (N^6) підгеном *C. sativa*, CsG2 – другий (N^7) підгеном *C. sativa*, CsG3 – третій (H^7) підгеном *C. sativa*

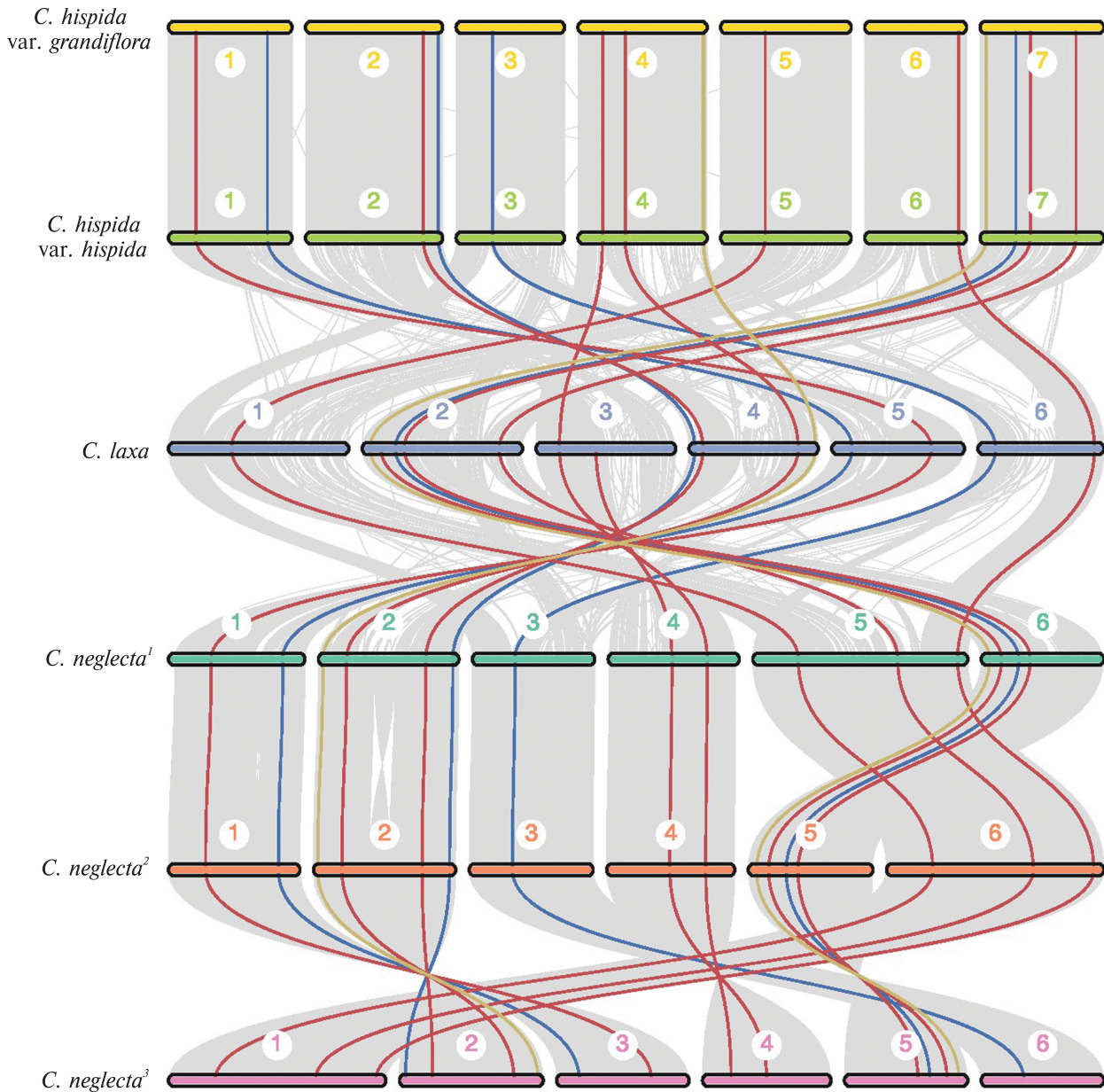


Рис. 5. Синтенія між підгеномами алогексаплоїдного виду *C. sativa* ($N^6N^7H^7$) та геномами його предкових видів *C. neglecta* (N^6) та *C. hispida* (H^7): (а) – синім кольором виділені поєднання між синтелогічними генами α -тубулінів, (б) – червоним – синтелогі β -тубулінів, (в) – жовтим – синтелогі γ -тубулінів. CsG1 – перший (N^6) підгеном *C. sativa*, CsG2 – другий (N^7) підгеном *C. sativa*, CsG3 – третій (H^7) підгеном *C. sativa*. Цифри біля видової назви *C. neglecta* вказує на конкретну збірку генома, зокрема: 1 – GCA_023864065.1 (референтна); 2 – GCA_029034625.1; 3 – GWHCBFX00000000

класу III) найімовірніше виникли в результаті древніх повногеномних дуплікацій (Yemets et al., 2024). Найбільш цікавим є той факт, що у третьому підгеномі *C. sativa* наявні функціо-

нальні копії генів *CsTUB6-C* та *CsTUB2-C*, втрачені (псевдогенізовані) у *C. hispida* var. *hispida*, який є предковим для даного підгеному (Martin et al., 2022; Blume et al., 2023, 2024). Даний

факт свідчить на користь того, що втрата генів ізотипів TUB2 та TUB6 у представників *C. hispida* відбулась вже після того, як даний вид взяв участь у поліплоїдизації та утворенні *C. sativa*. Більше того, оскільки два цих гени втрачені у обох таксонів — var. *hispida* та var. *grandiflora*, — можна припустити, що псевдогенізація відбулась до дивергенції та появи var. *hispida* та var. *grandiflora*. Це дозволяє припустити, що в утворенні виду *C. sativa* брала участь предкова форма *C. hispida*, а не окрема форма var. *hispida* або var. *grandiflora*, як наразі вважається (Martin et al., 2022; Blume et al., 2023).

Синтенія γ -тубулінів виявилась доволі простою, оскільки зазвичай синтелогічні поєднання були виявлені між ортологами певного ізо типу (рис. 4, в). При цьому геном *C. sativa* зберігає усі шість копій генів γ -тубуліну, успадкованих від предкових видів. Варто зазначити, що третій підгеном *C. sativa* має лише одну копію *CsTUG1-C*, в той час як форми *C. hispida* мають ці гени у дуплікованому стані (*ChvgTUG1-1*, *ChvgTUG1-2* та *ChvhTUG1-1*, *ChvhTUG1-2*). Слід очікувати, що дана дуплікація у *C. hispida* виникла вже після того як даний вид взяв участь в утворенні *C. sativa*, оскільки алогексаплоїдний геном рижію посівного не має ознак втрати додаткової копії TUG1 та/або її псевдогена.

Для того, щоб оцінити консервативність локусів тубулінів у різних диплоїдних видів *Camelina*, нами було проведено аналіз синтенії геномів *C. neglecta*, *C. laxa*, *C. hispida* var. *hispida* та *C. hispida* var. *grandiflora* (рис. 5). Геноми *C. hispida* var. *hispida* та *C. hispida* var. *grandiflora* вказували максимальний рівень структурної подібності та мали ідентичні набори генів тубуліну (не враховуючи *ChvgTUB3-2*, локалізованого на послідовності JALGBY010 000009.1, яка не показана рис. 5), оскільки ці таксони є внутрішньовидовими та максимально спорідненими з досліджуваних. Слід також зазначити, що локалізація генів тубуліну сильно відрізнялась у *C. hispida*, *C. laxa*, *C. neglecta*. Вважається, що види *C. hispida* та *C. laxa* зазнали множинних незалежних геномних перебудов в ході їх еволюції, якщо порівнювати з реконструйованим предковим геномом представників *Camelina* (Mandáková et al., 2019; Martin et al., 2022). При цьому вва-

жається, що структура геному *C. neglecta* є найбільш близькою до предкового генома *Camelina* (Mandáková et al., 2019). Жодних відмінностей у структурах геномів різних збірок *C. neglecta* виявлено не було (хромосоми мали різну нумерацію), а ідентифіковані гени тубуліну розташовувались у гомологічних локусах.

Виявлені патерни консервації набору генів тубуліну та варіативності їх послідовностей дозволяють використовувати окремі ділянки їх генів (зокрема більш варіабельних регіонів — інтронів) як молекулярні маркери. Спираючись на дані щодо консервативності генів тубуліну у проаналізованих видів *Camelina*, нами було проведено оцінку поліморфізму довжини інтронів β - та γ -тубулінів (рис. 6). Вибір саме генів β - та γ -тубулінів для використання як маркерів обґрунтовується їх висококонсервативною екзон-інтронною структурою, котра є більш сталою, ніж у α -тубулінів (Breviario et al., 2013; Pirko et al., 2018a, b). Більше того, раніше нами було показано, що інтрони α -тубулінів у *C. sativa* є доволі великими за розмірами (можуть сягати майже 4 тис. п.н.), через що використання їх у якості маркерів є доволі складним, менш надійним та менш інформативним (Blume et al., 2020, 2024). Слід також зазначити, що проаналізовані в даній роботі зразки були також використані іншими дослідниками для сиквенування геному. Зокрема, на основі генотипу PI650133 (*C. hispida* var. *grandiflora*) була створена збірка GCA_023864115.1, а генотип PI650135 (*C. neglecta*) був використаний при створенні усіх трьох наразі існуючих збірок геному для цього виду, а саме, GCA_023864065.1, GCA_029034625.1, GWHCBFX00000000 (Martin et al., 2022; Chaudhary et al., 2023; Wang et al., 2024).

Результати проведеного TBP- та cTBP-аналізу (рис. 6, а, б) дозволили виявити чіткі відмінності між генотипами різних видів, спираючись на отримані молекулярно-генетичні профілі. Так, в першу чергу кількість ампліконів добре узгоджується з плоїдністю генома того чи іншого виду: найменшою кількістю ампліконів вирізнялись диплоїдні види *C. hispida* var. *hispida*, *C. laxa*, *C. neglecta*, в той час, як найбільшу кількість ампліконів було виявлено у гексаплоїдних видів *C. microcarpa* та *C. sativa*.

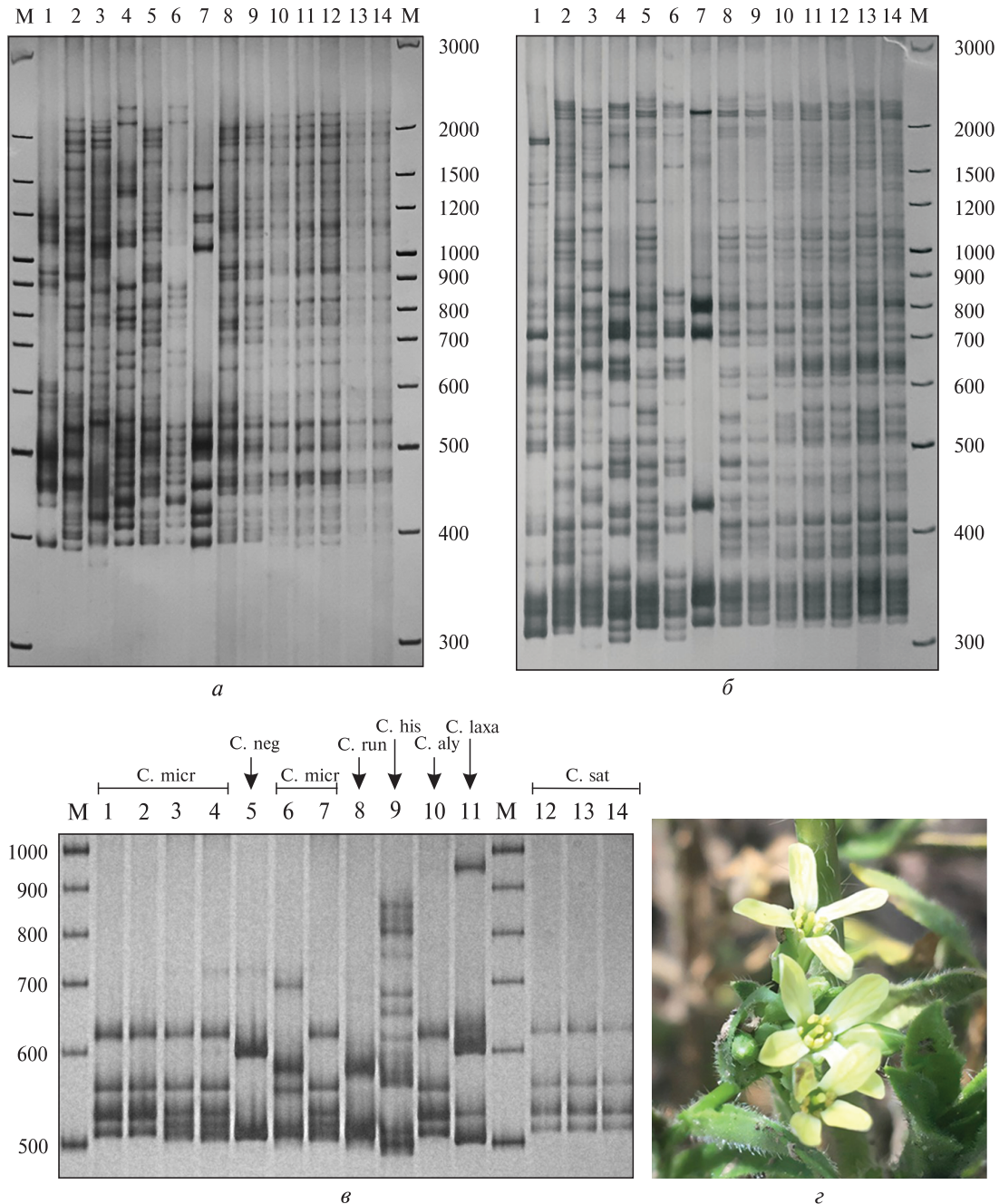


Рис. 6. Молекулярно-генетичні профілі, що відображають поліморфізм довжини інтронів генів β - та γ -тубуліну у різних представників роду *Camelinea*: (а) – аналіз поліморфізму довжини 1-го інтронів β -тубулінів (ТВР); (б) – аналіз поліморфізму довжини 2-го інтронів β -тубулінів (сТВР); (в) – аналіз поліморфізму довжини 2-го інтронів γ -тубулінів (γ ТВР); (г) – зовнішній вигляд квіток *C. rumelica* генотипу PI650134, культивованого в НБС ім. М.М. Гришка протягом 2020 року. М – маркер молекулярної маси, (а, б): 1 – *C. laxa* (Ames 32852); 2 – *C. alyssum* (PI650132); 3 – *C. hispida* var. *grandiflora* (PI650133); 4 – *C. rumelica* (PI650138); 5 – *C. microcarpa* тип 2 (PI650136), 6 – *C. rumelica* (PI650134); 7 – *C. neglecta* (PI650135); 8–9 – *C. microcarpa* тип 2 (PI633186, PI633187); 10–11 – *C. microcarpa* тип 1 (PI633191, Ames 31219); 12–14 – *C. sativa* (сорті Suneson, Міраж, Клондайк); (в) – нумерація зразків відповідно до табл. 2

Отримані нами результати підтверджуються та узгоджуються з іншими дослідженнями, проведеними за допомогою ТВР/сТВР маркерних систем (рис. 6, а, б). Так, у роботі Galasso et al. (2015) ці зразки з різною плоїдністю також характеризуються відмінними ДНК профілями. Більше того, раніше нами було показано, що за допомогою даного підходу можна як розрізнити варіанти гексаплоїдного *C. microcarpa* з різною композицією геному, так і аналізувати популяційну структуру даного виду в окремих регіонах (Sakharova et al., 2023, 2025). Також, варто відзначити, що ДНК-профілі *C. alyssum* і *C. sativa* виявилися схожими з *C. microcarpa*, що узгоджується з усталеними уявленнями про спільне еволюційне походження цих видів. У зразків тетраплоїдного виду *C. rumelica* виявлено меншу кількість фрагментів. Для *C. hispida* характерна більша кількість фрагментів, що може пояснюватися високою поліморфністю даного виду, через що більша кількість локусів β -тубулінів може перебувати у гетерозиготному стані, що виявляється у більшій кількості отриманих ампліконів.

Також нами було проведено γ ТВР-аналіз досліджуваних генотипів рижію з використанням вироджених праймерів, націлених на 2-й інтрон генів γ -тубуліну обох ізотипів (рис. 6, в). Результати γ ТВР-аналізу показали, що з 5 вивчених зразків *C. microcarpa* різного походження, 2 зразки, які належать до *C. microcarpa* тип 1, мають профіль подібний до *C. sativa*. Аналогічно до результатів ТВР/сТВР-аналізу *C. alyssum*, *C. sativa* та *C. microcarpa* мали подібні профілі з великою кількістю ампліконів. Диплоїдні види *C. laxa* та *C. neglecta* вирізнялись малою кількістю ампліконів (2–4), в той час як *C. hispida* мав 9 ампліконів. Більша кількість ампліфікованих фрагментів у *C. hispida* var. *grandiflora* (PI650133) може бути пов'язана з як з більшою кількістю генів γ -тубуліну, так і з потенційно вищим рівнем гетерозиготності даного виду. Найбільш цікавим є той факт, що обидва генотипи тетраплоїдного виду *C. rumelica* характеризуються наявністю різної кількості ампліконів. Зокрема, генотип PI650138 мав два амплікони, а PI650134 – чотири, два з яких співпадали за довжиною з ампліконами першо-

го генотипу. Найбільш вірогідно, що у випадку генотипу PI650138 можна спостерігати перекидання ампліконів, однак даний аспект питання поліморфізму та консервативності інтронів γ -тубуліну потребує подальшого вивчення. Також, слід враховувати, що вид *C. rumelica* є також генетично гетерогенним та у деяких дослідженнях утворює поліфілетичну групу при філогенетичній реконструкції походження видів *Camelina* (Brock et al., 2022a).

Таким чином, в результаті проведеного ТВР/сТВР-аналізу досліджені зразки чітко розподіляються за плоїдністю, при цьому отримані амплікони відповідають очікуванам довжинам ампліконів з локусів виявлених генів β -ту-буліну. Крім того, найбільш поліморфні амплікони можуть бути використані для баркодингу різних генотипів *C. microcarpa*. На основі отриманих молекулярно-генетичних профілів для різних генотипів *Camelina* нами було проведено ряд аналізів, котрі свідчать про здатність ТВР-, сТВР- та γ ТВР-маркерів (загалом дані 146 маркерів) ефективно диференціювати різні види *Camelina* (рис. 7). Зокрема, дані маркерного аналізу були використані для реконструкції філогенетичних співвідношень проаналізованих генотипів *Camelina* за допомогою методу майбільшої правдоподібності (рис. 7, а). Отримане філогенетичне дерево демонструє розподіл генотипів певного таксона (виду) в окремі власні кластери.

Зокрема, представники *C. sativa* були виділені в окремий кластер з високою бустреп підтримкою (понад 74 %). Представники *C. microcarpa* тип 1 були розташовані у сестринській кладі по відношенню до представників *C. sativa*, що добре відображає їх еволюційне співвідношення. Вважається, що *C. microcarpa* тип 1 є прями предком одомашненого виду *C. sativa*, обидва таксони мають однакову плоїдність, хромосомне число та подібну структуру геному ($N^6N^7H^7$) (Chaudhary et al., 2020; Brock et al., 2022b). У той же час, генотипи *C. microcarpa* тип 2 були розташовані в окремій кладі, котра є сестринською по відношенню до групи *C. sativa*-*C. microcarpa* тип 1. Наразі вважається, що представники *C. microcarpa* тип 2 мають відмінне хромосомне число та іншу структуру генома ($N^6N^7 N^6$), оскільки дана таксономічна група могла виникнути в ре-

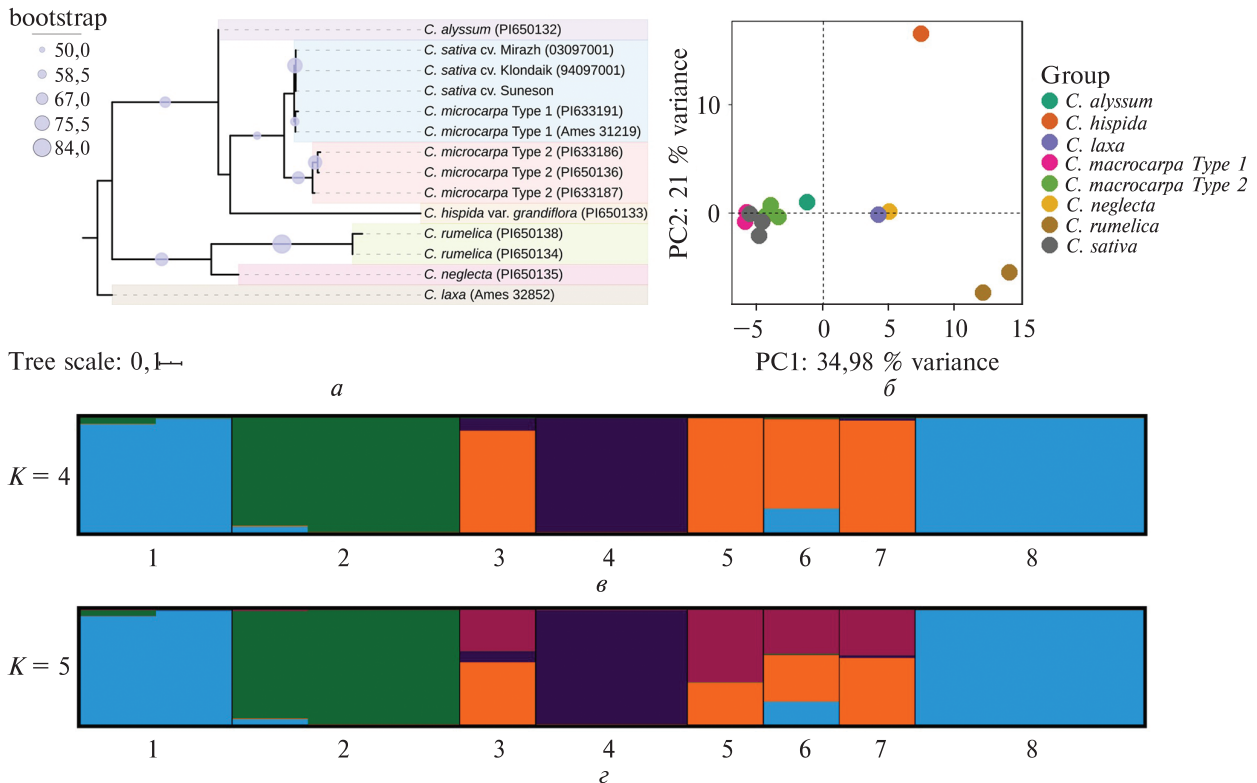


Рис. 7. Диференціація видів *Camelina* на основі генотипування та ДНК-баркодингу за допомогою ТВР-, сТВР- та γТВР-маркерів: (а) — філогенетичне дерево, що відображає розподіл та взаємозв'язки аналізованих генотипів *Camelina*, дерево вкорінене за генотипом *C. laxa* (Ames 32852); (б) — діаграма PCA (аналізу основних координат), що відображає групування та розподіл генотипів різних видів *Camelina*; (в) — результати STRUCTURE-аналізу з розподілом генотипів різних видів *Camelina* за кластерами (популяціями) при $K=4$; (г) — результати STRUCTURE-аналізу з розподілом генотипів різних видів *Camelina* за кластерами (популяціями) при $K=5$. Групи генотипів у STRUCTURE-аналізі (в, г) відповідали таким позначкам: 1 — *C. microcarpa* тип 1; 2 — *C. microcarpa* тип 2; 3 — *C. neglecta*; 4 — *C. rumelica*; 5 — *C. hispida* var. *grandiflora*; 6 — *C. alyssum*; 7 — *C. laxa*; 8 — *C. sativa* (відповідно до переліку генотипів у табл. 2)

зультаті незалежних подій алополіплоїдизації, потенційно без участі *C. hispida* (Chaudhary et al., 2020; Mandáková and Lysak, 2022; Blume et al., 2023).

Слід зазначити, що при цьому зразок *C. hispida* був розташований у базальній гілці відносно усіх гексаплоїдних рижіїв, окрім *C. alyssum* (рис. 7, а). Наразі статус *C. alyssum* як окремої таксономічної одиниці залишається дискусійним, однак у нашому дослідженні генотип даного виду виказував ознаки значної генетичної віддаленості від основної маси гексаплоїдних представників *Camelina*. Цікавим є також те, що диплоїдний генотип *C. neglecta* був розташований як сестринська гілка до кластеру генотипів *C. rumelica*. На сьогодні

вважається, що вид *C. rumelica* виник шляхом міжвидової *C. neglecta* та *C. hispida*. Близькість *C. neglecta* та *C. rumelica* добре відображається на реконструйованому дереві (рис. 7, а). Варто відмітити, що реконструкція філогенії видів *Camelina* з різною плоїдністю є нетривіальною задачею, оскільки поліплоїдні таксони цього роду в основному представляють різні комбінації *C. neglecta*- та *C. hispida*-подібних геномів у різних пропорціях. Оскільки філогенетичне дерево є лише двовимірною проекцією ієрархічних співвідношень між видами, воно, на жаль, не в змозі точно відобразити складність та переплетеність еволюційних подій, які мали місце в межах роду *Camelina* (Blume et al., 2023). Враховуючи цей

факт, нами було проведено також ряд додаткових аналізів.

З метою подальшого аналізу патернів генетичної варіації досліджуваних генотипів різних видів *Camelina* нами було здійснено аналіз основних компонент (PCA), спираючись на дані проведеного маркерного аналізу (рис. 7, б). Отримана PCA діаграма відображає утворення декількох груп генотипів різних видів. Зокрема, було виявлено групування усіх гексаплоїдних видів у окремий кластер, який складався з *C. microcarpa* тип 1, *C. microcarpa* тип 2, *C. sativa* та більш віддаленого *C. alyssum*. Цікаво, що диплоїдні *C. neglecta* та *C. laxa* були локалізовані поряд на діаграмі, в той час як генотип *C. hispida* var. *grandiflora* був найбільш віддаленим від усіх представників роду. Також, було відмічено групування зразків *C. rumelica* в окремий кластер.

Результати подальшого аналізу кластеризації (популяційний STRUCTURE-аналіз) дозволили виявити достовірне утворення чотирьох або п'яти популяційних кластерів, якщо оцінювати за методиками Puechmaillie et al. (2016) та Evanno et al. (2005), відповідно (рис. 7, в, г). У обох випадках спостерігалось формування окремих популяційних кластерів з утворенням групи *C. sativa*-*C. microcarpa* тип 1 та групи генотипів *C. microcarpa* тип 2. Також, в обох випадках спостерігалось формування окремого популяційного кластеру для генотипів *C. rumelica*. Обидва варіанти реконструкції популяційних кластерів показали об'єднання усіх диплоїдних генотипів (*C. laxa*, *C. neglecta* та *C. hispida*) у спільну групу (рис. 7, в, г). При цьому, *C. neglecta* виявляв наявність домішки (вказує на спільне генетичне походження) від кластеру *C. rumelica*, що добре корелює з групуванням цих таксонів при філогенетичній реконструкції, описаній вище (рис. 7, а). Слід зазначити, що в обох випадках генотип *C. alyssum* був віднесений до популяційного кластеру диплоїдних видів, хоча і зі значною генетичною домішкою від кластеру *C. sativa*-*C. microcarpa* тип 1 (рис. 7, в, г). Збільшення кількості ідентифікованих кластерів лише привносила частку генетичної домішки від нового п'ятого кластеру до групи диплоїдів та *C. alyssum*, що, однак, не змінювало сутності біологічної інтерпрета-

ції отриманих результатів STRUCTURE-аналізу (рис. 7, г).

З огляду на отримані результати можна зробити підсумок, що комбінація використаних для генотипування TBP-, cTBP- та γTBP-маркерів разом з даними повногеномного аналізу генів тубуліну дозволяє більш точно інтерпретувати дані молекулярно-генетичного аналізу та ефективно диференціювати більшість видів *Camelina*. Однак, питання походження та генетичної відмінності *C. alyssum* від інших гексаплоїдних представників *Camelina* вимагатиме подальших досліджень цього виду. Слід також зазначити, що запропонована комбінація маркерів для ДНК-баркодингу видів *Camelina* дозволяє диференціювати навіть такі споріднені таксономічні групи, як *C. sativa* та *C. microcarpa* тип 1. При цьому слід зважати на правильну біоінформатичну обробку отриманих даних, оскільки ці дві таксономічні одиниці є подібними і хоча добре розрізняються при філогенетичному аналізі, вони можуть бути не диференційовані шляхом кластерного популяційного STRUCTURE-аналізу.

Висока консервативність генів тубуліну різних підродин у квіткових рослин робить можливим широке використання елементів послідовностей цих генів в якості маркерних послідовностей для вирішення ширшого спектру молекулярно-генетичних задач, в тому числі і для генотипування (Yemets et al., 2008; Rabokon, 2021; Braglia et al., 2023). У попередніх роботах нами вже було продемонстровано можливість ефективного використання комбінації TBP-, cTBP- та SSR-маркерів як для оцінки генетичного різноманіття сортів та селекційних ліній *C. sativa*, так і для аналізу популяційної структури *C. microcarpa* (Blume et al., 2020; Sakharova et al., 2025). Значною перевагою використання TBP та cTBP маркерних систем є можливість одночасного аналізу великої кількості незчеплених локусів генів β-тубуліну. Це дозволяє в ході однієї ПЛР отримувати значні обсяги інформації про генетичних поліморфізм. З іншого боку, використання γTBP-маркерів дозволяє аналізувати інтронні ділянки більш консервативних генів γ-тубуліну, що добре підходить для ДНК-баркодингу окремих видів або таксономічних одиниць нижчого порядку, як це було показано вище.

Можливість отримання значної кількості генетичної інформації за допомогою ПЛР аналізу стає критичною за умов обмеженої доступності генетичного матеріалу, або пошкодженості його ДНК (Rabokon, 2021; Vog et al., 2022; Sakharova et al., 2025). Раніше нами також було показано можливість успішного використання методу оцінки поліморфізму довжини інtronів генів як для ДНК-баркодингу, так і для аналізу гібридного походження селекційного матеріалу, зокрема, на прикладі *Brassica rapa* (Rabokon et al., 2023) та представників роду *Prunus* (Lykholat et al., 2022). У майбутньому запропонована система молекулярних маркерів може бути успішно використана для аналізу видового та сортового складу селекційного матеріалу, в тому числі й для видів *Camelina*. Оцінка генетичного різноманіття пулу генотипів за допомогою запропонованих методів генотипування (або в комбінації з іншими маркерними системами) може дозволити значно пришвидшити селекційний процес та подальше отримання високопродуктивних генотипів того чи іншого виду рослин.

Дотримання етичних стандартів. Ця стаття не містить будь-яких досліджень з використанням людей і тварин як об'єктів.

Конфлікт інтересів. Автори заявляють про відсутність конфлікту інтересів.

Фінансування. Дослідження проведені в рамках проєкту для дослідницьких груп молодих вчених НАН України «Генетичне різноманіття та популяційно генетична структура рижію дрібноплідного в українській частині його центру походження» (2022–2023 pp.) (Державний реєстраційний № 0122U002196), а також в рамках проєкту «Unraveling the diversity of polyploid wild *Camelina* germplasm for biofuel crop improvement» програми EURIZON, яка фінансується в рамках Horizon 2020 за грантовою угодою № 871072 (Грантова угода УНТЦ — EU#3044) (2024–2025 pp.).

ASSESSING GENOMIC EVOLUTION OF TUBULIN GENE FAMILY FOR *CAMELINA* SPECIES GENOTYPING

R.Y. Blume, A.M. Rabokon,
V.H. Sakharova, D.B. Rakhmetov,
Y.V. Pirko, Y.B. Blume

Institute of Food Biotechnology and Genomics NAS of Ukraine,
Baidy-Vyshnevetskooho str., 2a; Kyiv, 04123, Ukraine
M.M. Gryshko National Botanical Garden NAS of Ukraine, Ukraine,
Sadovo-Botanichna str., 1, 01014, Kyiv, Ukraine
E-mail: blume.rostislav@gmail.com

Tubulins play a key role in the functioning of cytoskeletal systems that regulate such fundamental processes as cell division and growth. Correct identification of isotypes and determination of the orthology of tubulin genes in plants is a non-trivial task that requires the involvement of a complex of bioinformatics approaches. In the present study, a genome-wide search and identification of tubulin genes was carried out in diploid representatives of the genus *Camelina*, in particular in the *C. neglecta*, *C. laxa*, *C. hispida* species, which allowed us to identify complete sets of α -, β -, and γ -tubulin genes, as well as their pseudogenes. Phylogenetic analysis and a series of genome-wide comparisons allowed us to establish the orthology of the tubulin genes, determine isotype identity of the encoded tubulins, and trace evolutionary changes in tubulin gene sets during species divergence and the emergence of allohexaploid *C. sativa* species. Genotyping of the accessions of different *Camelina* species using TBP-, cTBP-, and γ TBP-markers allowed effective differentiation of species based on the assessment of polymorphisms of intronic regions of the β - and γ -tubulin genes. The obtained results lay a strong groundwork for further studies of the isotype and functional diversity of tubulins in Cruciferae and other groups of flowering plants, and will also contribute to the development and implementation of new highly efficient molecular marker systems for DNA-barcoding and marker-assisted breeding of plant species, including such promising oilseed crops as *C. sativa*.

СПИСОК ЛІТЕРАТУРИ

- Benbouza, H., Jean-Marie, J., and Jean-Pierre, B., Optimization of a reliable, fast, cheap and sensitive silver staining method to detect SSR markers in polyacrylamide gels, *Biotechnol. Agron. Soc. Environ.*, 2006, vol. 10, no. 2, pp. 77–81
- Berti, M., Gesch, R., Eynck, C., et al., *Camelina* uses, genetics, genomics, production, and management, *Ind. Crop. Prod.*, 2016, vol. 94, pp. 690–710. <https://doi.org/10.1016/j.indcrop.2016.09.034>
- Blume, R.Y., Kalendar, R.N., Guo, L., et al., Overcoming genetic paucity of *Camelina sativa*: possibilities for interspecific hybridization conditioned by the genus evolution pathway, *Front. Plant Sci.*, 2023, vol. 14, pp. 1259431. <https://doi.org/10.3389/fpls.2023.1259431>
- Blume, R.Y., Rabokon, A.M., Pydiura, M., et al.,

- Genome-wide identification and evolution of the tubulin gene family in *Camelina sativa*, *BMC Genomics*, 2024, vol. 25, pp. 599. <https://doi.org/10.1186/s12864-024-10503-y>
- Blume, R.Y., Rabokon, A.N., Postovitova, A.S., et al., Evaluating diversity and breeding perspectives of Ukrainian spring camelina genotypes, *Cytol. Genet.*, 2020, vol. 54, no. 5, pp. 420–436. <https://doi.org/10.3103/S0095452720050084>
- Blume, R.Y., Rakhmetov, D.B., and Blume, Y.B., Evaluation of Ukrainian *Camelina sativa* germplasm productivity and analysis of its amenability for efficient biodiesel production, *Ind. Crop. Prod.*, 2022, vol. 187B, pp. 115477. <https://doi.org/10.1016/j.indcrop.2022.115477>
- Braglia, L., Gavazzi, F., Gianm, S., et al., Tubulin-based polymorphism (TBP) in plant genotyping, *Plant Genotyping: Methods in Molecular Biology*, 2023, vol. 2638, pp. 387–401. https://doi.org/10.1007/978-1-0716-3024-2_28
- Breviario, D., Gianni, S., and Morello, L., Multiple tubulins: evolutionary aspects and biological implications, *Plant J.*, 2013, vol. 75, pp. 202–218. <https://doi.org/10.1111/tpj.12243>
- Brock, J.R., Mandáková, T., McKain, M., et al., Chloroplast phylogenomics in *Camelina* (*Brassicaceae*) reveals multiple origins of polyploid species and the maternal lineage of *C. sativa*, *Hortic Res.*, 2022a, vol. 9, pp. uhab050. <https://doi.org/10.1093/hortre/uhab050>
- Brock, J.R., Ritchey, M.M., Olsen, K.M., et al., Molecular and archaeological evidence on the geographical origin of domestication for *Camelina sativa*, *Am J Bot.*, 2022b, vol. 109, no. 7, pp. 1177–1190. <https://doi.org/10.1002/ajb2.16027>
- Chaudhary, R., Koh, C.S., Kagale, S., et al., Assessing diversity in the *Camelina* genus provides insights into the genome structure of *Camelina sativa*, *G3: Genes, Genomes, Genetics*, 2020, vol. 10, no. 4, pp. 1297–1308. <https://doi.org/10.1534/g3.119.400957>
- Chaudhary, R., Koh, C.S., Perumal, S., et al., Sequencing of *Camelina neglecta*, a diploid progenitor of the hexaploid oilseed *Camelina sativa*, *Plant Biotechnol. J.*, 2023, vol. 21, pp. 521–535. <https://doi.org/10.1111/pbi.13968>
- Chen, B., Zhao, J., Fu, G., et al., Identification and expression analysis of *Tubulin* gene family in upland cotton, *J. Cotton Res.*, 2021, vol. 4, pp. 20. <https://doi.org/10.1186/s42397-021-00097-1>
- Cheng, F., Wu, J., Fang, L., et al., Biased gene fractionation and dominant gene expression among the subgenomes of *Brassica rapa*, *PLoS One*, 2012, vol. 7, no. 5, pp. e36442. <https://doi.org/10.1371/journal.pone.0036442>
- Demchuk, O.N., and Blume, Ya.B., Construction of phylogenetic tree of plant tubulins basing on the homology of their protein sequences, *Cytol. Genet.*, 2005, vol. 39, no. 1, pp. 3–9.
- Edgar, R.C., MUSCLE: multiple sequence alignment with high accuracy and high throughput, *Nucleic Acid Res.*, 2004, vol. 32, no. 5, pp. 1792–1797. <https://doi.org/10.1093/nar/gkh340>
- Evanno, G., Regnaut, S., and Goudet, J., Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study, *Mol. Ecol.*, 2005, vol. 14, no. 8, pp. 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Faure, J.D., and Tepfer, M., Camelina, a Swiss knife for plant lipid biotechnology, *OCL*, 2016, vol. 23, no. 5, pp. D503. <https://doi.org/10.1051/ocl/2016023>
- Findeisen, P., Muhlhause, S., Dempewolf, S., et al., Six subgroups and extensive recent duplications characterize the evolution of the eukaryotic tubulin protein family, *Genome Biol. Evol.*, 2014, vol. 6, no. 9, pp. 2274–2288. <https://doi.org/10.1093/gbe/evl187>
- Galasso, I., Manca, A., Braglia, L., et al., Genomic fingerprinting of *Camelina* species using cTBP as molecular marker, *Am. J. Plant Sci.*, 2015, vol. 6, pp. 1184–1200. <https://doi.org/10.4236/ajps.2015.68122>
- Gavazzi, F., Pigna, G., Braglia, L., et al., Evolutionary characterization and transcript profiling of β -tubulin genes in flax (*Linum usitatissimum* L.) during plant development, *BMC Plant Biol.*, 2017, vol. 17, pp. 237. <https://doi.org/10.1186/s12870-017-1186-0>
- Ghidoli, M., Ponzoni, E., Araniti, F., et al., Genetic improvement of *Camelina sativa* (L.) Crantz: opportunities and challenges, *Plants*, 2023, vol. 12, pp. 570. <https://doi.org/10.3390/plants12030570>
- Hatje, K., Keller, O., Hammesfahr, B., et al., Cross-species protein sequence and gene structure prediction with fine-tuned Webscipio 2.0 and Scipio, *BMC Res. Notes.*, 2011, vol. 4, pp. 265. <https://doi.org/10.1186/1756-0500-4-265>
- Hoang, D.T., Chernomor, O., von Haeseler, A., et al., UFBoot2: Improving the ultrafast bootstrap approximation, *Mol. Biol. Evol.*, 2018, vol. 35, no. 2, pp. 518–522. <https://doi.org/10.1093/molbev/msx281>
- Hoff, K.J., and Stanke, M., WebAUGUSTUS – a web service for training AUGUSTUS and predicting genes in eukaryotes, *Nucl. Acids Res.*, 2013, vol. 41, no. W1, pp. W123–W128. <https://doi.org/10.1093/nar/gkt418>
- Hu, B., GSDS 2.0: an upgraded gene feature visualization server, *Bioinformatics*, 2015, vol. 31, no. 8, pp. 1296–1297. <https://doi.org/10.1093/bioinformatics/btu817>
- Hu, T., Pattyn, P., Bakker, E., et al., The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change, *Nat. Genet.*, 2011, vol. 43, pp. 476–481. <https://doi.org/10.1038/ng.807>

- Hubisz, M.J., Falush, D., Stephens, M., and Pritchard, J.K., Inferring weak population structure with the assistance of sample group information, *Mol. Ecol. Resour.*, 2009, vol. 9, no. 5, pp. 1322–1332. <https://doi.org/10.1111/j.1755-0998.2009.02591.x>
- Jiao, W.B., and Schneeberger, K., Chromosome-level assemblies of multiple *Arabidopsis* genomes reveal hotspots of rearrangements with altered evolutionary dynamics, *Nat. Commun.*, 2020, vol. 11, pp. 989. <https://doi.org/10.1038/s41467-020-14779-y>
- Jost, W., Baur, A., Nick, P., et al., A large plant beta-tubulin family with minimal C-terminal variation but differences in expression, *Gene.*, 2004, vol. 340, no. 1, pp. 151–160. <https://doi.org/10.1016/j.gene.2004.06.009>
- Kagale, S., Koh, C., Nixon, J., et al., The emerging bio-fuel crop *Camelina sativa* retains a highly undifferentiated hexaploid genome structure, *Nat. Commun.*, 2014, vol. 5, pp. 3706. <https://doi.org/10.1038/ncomms4706>
- Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., et al., ModelFinder: Fast model selection for accurate phylogenetic estimates, *Nat. Method.*, 2017, vol. 14, pp. 587–589. <https://doi.org/10.1038/nmeth.4285>
- Kopelman, N.M., Mayzel, J., Jakobsson, M., et al., CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across *K*, *Mol. Ecol. Resour.*, 2015, vol. 15, pp. 1179–1191. <https://doi.org/10.1111/1755-0998.12387>
- Larralde, M., and Zeller G., PyHMMER: a Python library binding to HMMER for efficient sequence analysis, *Bioinformatics*, 2023, vol. 39, no. 5, pp. btad214. <https://doi.org/10.1093/bioinformatics/btad214>
- Letunic, I., and Bork, P., Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool, *Nucleic Acid. Res.*, 2024, vol. 52, no. W1, pp. W78–W82. <https://doi.org/10.1093/nar/gkae268>
- Li, Y.-L., and Liu, J.-X., StructureSelector: A web-based software to select and visualize the optimal number of clusters using multiple methods, *Mol. Ecol. Resour.*, 2018, vol. 18, pp. 176–177. <https://doi.org/10.1111/1755-0998.12719>
- Lowe, J., Li, H., Downing, K.H., and Nogales, E., Refined structure of $\alpha\beta$ -tubulin at 3.5 resolution, *J. Mol. Biol.*, 2001, vol. 313, pp. 1045–1057. <https://doi.org/10.1006/jmbi.2001.5077>
- Lykholat, Y.V., Rabokon, A.M., Blume, R.Ya., et al., Characterization of β -tubulin genes in *Prunus persica* and *Prunus dulcis* for fingerprinting of their interspecific hybrids, *Cytol. Genet.*, 2022, vol. 56, no. 6, pp. 481–493. <https://doi.org/10.3103/S009545272206007X>
- Manca, A., Pecchia, P., Mapelli, S., et al., Evaluation of genetic diversity in a *Camelina sativa* (L.) Crantz collection using microsatellite markers and biochemical traits, *Genet. Resour. Crop Evol.*, 2012, vol. 60, pp. 1223–1226. <https://doi.org/10.1007/s10722-012-9913-8>
- Mandáková, T., and Lysak, M.A., The identification of the missing maternal genome of the allohexaploid camelina (*Camelina sativa*), *Plant J.*, 2022, vol. 112, pp. 622–629. <https://doi.org/10.1111/tpj.15931>
- Mandáková, T., Pouch, M., Brock, J.R., et al., Origin and evolution of diploid and allopolyploid *Camelina* genomes were accompanied by chromosome shattering, *Plant Cell*, 2019, vol. 31, no. 11, pp. 2596–2612. <https://doi.org/10.1105/tpc.19.00366>
- Martin, S.L., Lujan-Toro, B., James, T., et al., Insights from the genomes of 4 diploid *Camelina* spp., *G3: Genes, Genomes, Genetics*, 2022, vol. 12, no. 12, pp. jkac182. <https://doi.org/10.1093/g3journal/jkac182>
- Mistry, J., Chuguransky, S., Williams, L., et al., Pfam: The protein families database in 2021, *Nucleic Acid. Res.*, 2021, vol. 49, no. D1, pp. D412–D419. <https://doi.org/10.1093/nar/gkaa913>
- Nguyen, L.-T., Schmidt, H.A., von Haeseler, A., and Minh, B.Q., IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies, *Mol. Biol. Evol.*, 2015, vol. 32, pp. 268–274. <https://doi.org/10.1093/molbev/msu300>
- Oakley, R., Wang, Y., Ramakrishna, W., et al., Differential expansion and expression of alpha- and beta-tubulin gene families in *Populus*, *Plant Physiol.*, 2007, vol. 145, no. 3, pp. 961–973. <https://doi.org/10.1104/pp.107.107086>
- Paysan-Lafosse, T., Blum, M., Chuguransky, S., et al., InterPro in 2022, *Nucleic Acid. Res.*, 2023, vol. 51, no. D1, pp. D418–D427. <https://doi.org/10.1093/nar/gkac993>
- Pirko, Ya.V., Buy, D.D., Postovoitova, A.S., et al., New ILP method based on γ -tubulin genes intron length polymorphism, *Rep. Natl. Acad. Sci. Ukraine*, 2018a, no. 12, pp. 1025–1030
- Pirko, Ya.V., Postovoitova, A.S., Rabokon, A.M., et al., Study of intron length polymorphism of the α -tubulin genes as a method of analysis of the genetic differentiation in plants, *Ukr. Bot. J.*, 2018b, vol. 75, no. 6, pp. 576–584
- Puechmaille, S.J., The program STRUCTURE does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem, *Mol. Ecol. Res.*, 2016, vol. 16, pp. 608–627. <https://doi.org/10.1111/1755-0998.12512>
- Pydiura, N., Pirko, Y., Galinousky, D., et al., Genome-wide identification, phylogenetic classification, and exon-intron structure characterization of the tubulin and actin genes in flax (*Linum usitatissimum*), *Cell*

- Biol. Int.*, 2019, vol. 43, no. 9, pp. 1010–1019. <https://doi.org/10.1002/cbin.11001>
- Rabokon, A.M., Blume, R.Y., Sakharova, V.G., et al., Genotyping of interspecific *Brassica rapa* hybrids implying β -tubulin gene intron length polymorphism (TBP/cTBP) assessment, *Cytol. Genet.*, 2023, vol. 57, no. 6, pp. 538–549. <https://doi.org/10.3103/S0095452723060075>
- Rabokon, A.M., Intron length polymorphism of tubulin genes as an effective tool for genetic plant differentiation, *Rep. Natl. Acad. Sci. Ukraine*, 2021, no. 10, pp. 30–35. <https://doi.org/10.15407/vsn2021.10.030>
- Radchuk, V., The transcriptome of the tubulin gene family in plants, in *The Plant Cytoskeleton: A Key Tool for Agro-Biotechnology*, Blume, Y.B., Baird, W.V., Yemets, A.I., and Breviario, D., Eds., New York: Springer, 2008, pp. 219–241. https://doi.org/10.1007/978-1-4020-8843-8_11
- Raj, A., Stephens, M., and Pritchard, J.K., fast STRUCTURE: Variational inference of population structure in large SNP data sets, *Genetics*, 2014, vol. 197, no. 2, pp. 573–589. <https://doi.org/10.1534/genetics.114.164350>
- Rao, G., Zeng, Y., He, C., and Zhang, J., Characterization and putative post-translational regulation of α - and β -tubulin gene families in *Salix arbutifolia*, *Sci Rep.*, 2016, vol. 6, pp. 19258. <https://doi.org/10.1038/srep19258>
- Sakharova, V.H., Blume, R.Y., Rabokon, A.N., et al., Efficiency of genetic diversity assessment of little-pod false flax (*Camelina microcarpa* Andr. ex DC.) in Ukraine using SSR- and TBP-marker systems, *Rep. Natl. Acad. Sci. Ukraine*, 2023, vol. 4, pp. 85–94. <https://doi.org/10.15407/dopovidi2023.04.093>
- Sakharova, V.H., Blume, R.Y., Rabokon, A.M., et al., Exploring the genetic diversity and population structure of little-pod false flax (*Camelina microcarpa*: *Brassicaceae*) in Ukraine, *Ukr. Bot. J.*, 2025, vol. 82, no. 2, pp. 144–161. <https://doi.org/10.15407/ukrbotj82.02.144>
- Sambrook, J., and David, W.R. *Molecular Cloning: A Laboratory Manual*. Vol. 2, Cold Spring Harbor, 2001
- Silflow, C.D., Oppenheimer, D.G., Kopozak, S.D., et al., Plant tubulin genes structure and differential expression during development, *Dev. Genet.*, 1987, vol. 8, pp. 435–460. <https://doi.org/10.1002/dvg.1020080511>
- Tang, H., Krishnakumar, V., Zeng, X., et al., JCVI: A versatile toolkit for comparative genomics analysis, *Imeta*, 2024, vol. 3, no. 4, pp. e211. <https://doi.org/10.1002/imt2.211>
- Trifinopoulos, J., Nguyen, L.T., von Haeseler, A., and Minh, B.Q., W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis, *Nucleic Acid. Res.*, 2016, vol. 44, no. W1, pp. W232–W235. <https://doi.org/10.1093/nar/gkw256>
- Vollmann, J., and Eynck, C., Camelina as a sustainable oilseed crop: Contributions of plant breeding and genetic engineering, *Biotechnol. J.*, 2015, vol. 10, pp. 525–535. <https://doi.org/10.1002/biot.201400200>
- Wang, J., Chitsaz, F., Derbyshire, M.K., et al., The conserved domain database in 2023, *Nucleic Acid. Res.*, 2023, vol. 51, no. D1, pp. D384–D388. <https://doi.org/10.1093/nar/gkac1096>
- Wang, S., Blume, R.Y., Zhou, Z.-W., et al., Chromosome-level assembly and analysis of *Camelina neglecta* – a novel diploid model for camelina biotechnology research, *Biotechnol. Biofuels Bioproducts*, 2024, vol. 17, pp. 17. <https://doi.org/10.1186/s13068-024-02466-9>
- Wickstead, B., and Gull, K., The evolution of the cytoskeleton, *J. Cell Biol.*, 2011, vol. 194, no. 4, pp. 513–525. <https://doi.org/10.1083/jcb.201102065>
- Yemets, A., Radchuk, V., Bayer, O., et al., Development of transformation vectors based upon a modified plant α -tubulin gene as the selectable marker, *Cell Biol. Int.*, 2008, vol. 32, no. 5, pp. 566–570. <https://doi.org/10.1016/j.cellbi.2007.11.012>
- Yemets, A., Shadrina, R., Blume, R., et al., Autophagy formation, microtubule disorientation, and alteration of ATG8 and tubulin gene expression under simulated microgravity in *Arabidopsis thaliana*, *npj Microgravity*, 2024, vol. 10, pp. 31. <https://doi.org/10.1038/s41526-024-00381-9>
- Zanetti, F., Alberghini, B., Jeromela, A.M., et al., Camelina, an ancient oilseed crop actively contributing to the rural renaissance in Europe. A review, *Agron. Sustain. Dev.*, 2021, vol. 41, pp. 2. <https://doi.org/10.1007/s13593-020-00663-y>
- Žerdoner Čalasan, A., Seregin, A.P., Hurka, H., et al., The Eurasian steppe belt in time and space: Phylogeny and historical biogeography of the false flax (*Camelina* Crantz, *Camelineae*, *Brassicaceae*), *Flora*, 2019, vol. 260, pp. 151477. <https://doi.org/10.1016/j.flora.2019.151477>

Надійшла в редакцію 05.06.25
Після доопрацювання 10.07.25
Прийнята до друку 18.09.2025