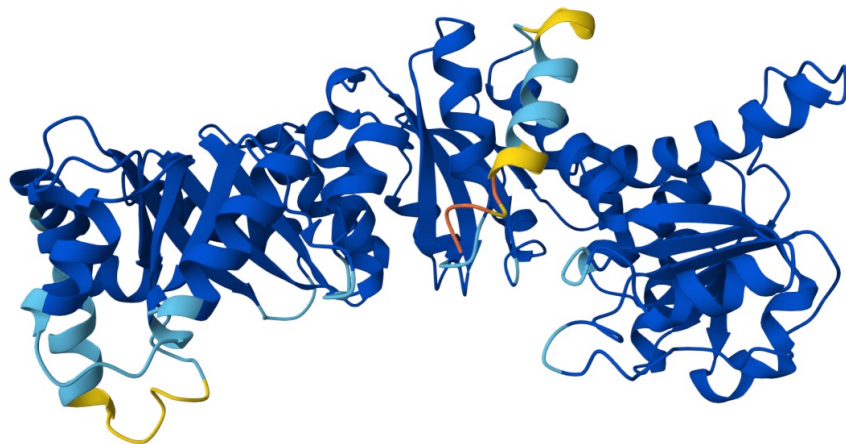


Поиск бифункциональных 3-дегидрохинатадегидратаз/ шикиматдегидрогеназ *Cornus sericea* L. на основе *de novo* секвенирования транскриптома

Евгения Никонорова, Татьяна Кроль, Дмитрий Балеев

ФГБНУ «Всероссийский научно-исследовательский
институт лекарственных и ароматических растений»,
Москва, Россия



Новосибирск, 2025



Гидролизуемые танины *Cornus sericea* L.

- *C. sericea* L. (свида шелковистая/кизил блестящий) накапливает значительное количество ГТ со сложной химической структурой
- В листьях *C. sericea* L. содержатся мономерные (изомеры гемина D, теллимаграндин I и II), димерные (изомеры камптотина A, корнусиина A, D или камптотина B, изомеры корнусиина E, G) и тримерные эллагитаннины (изомер корнусиина F, C и трапанина A)
- Нет данных о геноме и транскриптоме *C. sericea* L. – невозможно определить какие гены и кодируемые ими ферменты ответственны за синтез
- Механизм синтеза гидролизуемых танинов *C. sericea* L. остается малоизученным

Цель и задачи исследования

Цель исследования – изучить механизм регуляции синтеза гидролизуемых танинов на примере *Cornus sericea* L.



Провести секвенирование транскриптома *Cornus sericea* L. методами NGS и ONT секвенирования

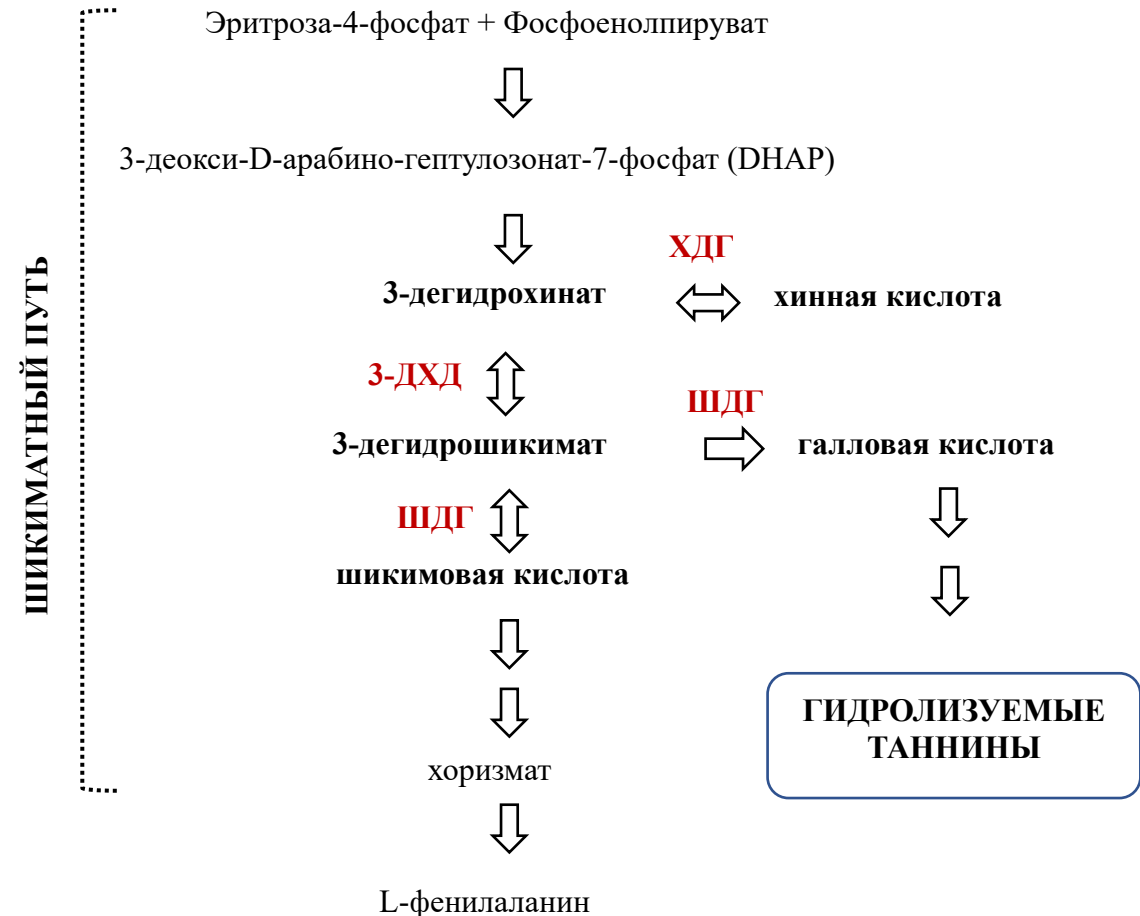


Выполнить *de novo* сборку транскриптома и провести функциональную аннотацию



Идентифицировать бифункциональный фермент 3-дегидрохинатадегидратаза/шикиматдегидрогеназа *Cornus sericea* L.

Бифункциональная 3-дегидрохинатадегидратаза / шикиматдегидрогеназа – ключевой фермент синтеза гидролизуемых танинов



Материалы и методы



Сбор растительного материала

Образцы молодых и зрелых листьев, бутонов, соцветий и незрелых плодов *C. sericea* были собраны в Ботаническом саду ФГБНУ ВИЛАР (г. Москва) в термос с жидким азотом и хранились при -80 до начала исследования



Выделение РНК

Тотальную РНК извлекали с помощью набора RNeasy Mini Kit и обрабатывали турбо-ДНКазой для удаления геномной ДНК



Контроль качества

Концентрацию РНК определяли с использованием Qubit RNA HS Assay Kit для обеспечения достаточного количества материала для секвенирования. Целостность РНК оценивалась с использованием RNA 6000 Pico Kit for Bioanalyzer

Секвенирование нового поколения (NGS)

Подготовка библиотеки проводилась с использованием набора TruSeq Stranded mRNA (Illumina). Секвенирование выполнялось на платформе NovaSeq 6000, 2 x 101 п.н. (Illumina Technologies, Сан-Диего, США). Демультиплексирование коротких ридов и обрезку адаптеров проводили с помощью Illumina bcl2fastq 2.20 и Skewer 0.2.2. Качество анализировали с помощью FastQC.

Нанопоровое секвенирование (ONT)

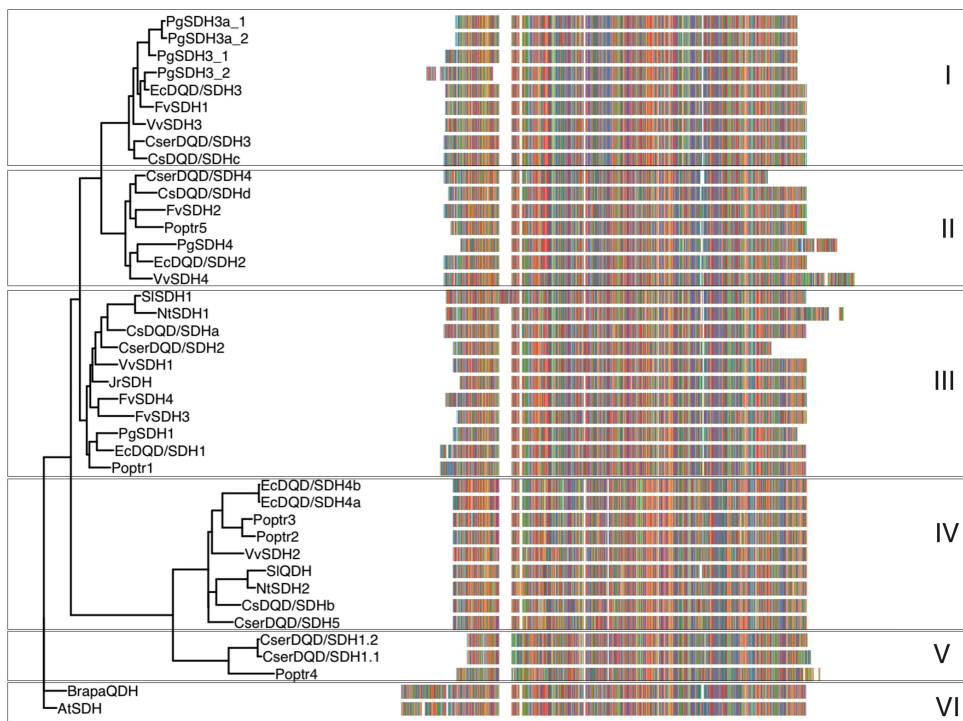
РНК была преобразована в кДНК с использованием набора для синтеза кДНК Mint с 20-25 циклами амплификации с модификациями: 1) синтез кДНК праймируется олигонуклеотидами, содержащими не только часть dT и адаптерную часть, но и индивидуальный штрихкод, специфичный для каждого образца; 2) оптимальное количество циклов амплификации выбирается на основе результатов количественной ПЦР с отжигом праймеров на 3' и 5' концах. Амплифицированную кДНК (150 нг) использовали в качестве входных данных для подготовки библиотеки с использованием стандартного протокола для геномной ДНК с набором SQK-LSK114. Обработка ONT прочтений выполнялась с помощью Guppy 6.5.7 и NTproc. PolyA удаляли с помощью cutadapt v.5.0. Качество анализировали с помощью NanoPlot.

Анализ данных

Сборка транскриптома выполнялась с помощью trans2express. Для оценки качества сборки использовали rnaQUAST и BUSCO. Функциональная аннотация проводилась с помощью blastx с e-value 1e-10 и баз данных Uniprot и RefSeq. Анализы путей KEGG выполняли на сервере KAAS, GO – при помощи eggNOG-mapper. Множественное выравнивание выполняли с помощью Clustal Omega в программе Uniprot. Для предсказания локализации изоформ в клетке онлайн платформ TargetP 2.0, PredSL, iPSORT Prediction, WoLF PSORT, DeepLoc-2.0. Построение филогенетического дерева с помощью iqtree в режиме ModelFinder (Q.plant+I+G4 модель) и 100 итераций.

Результаты

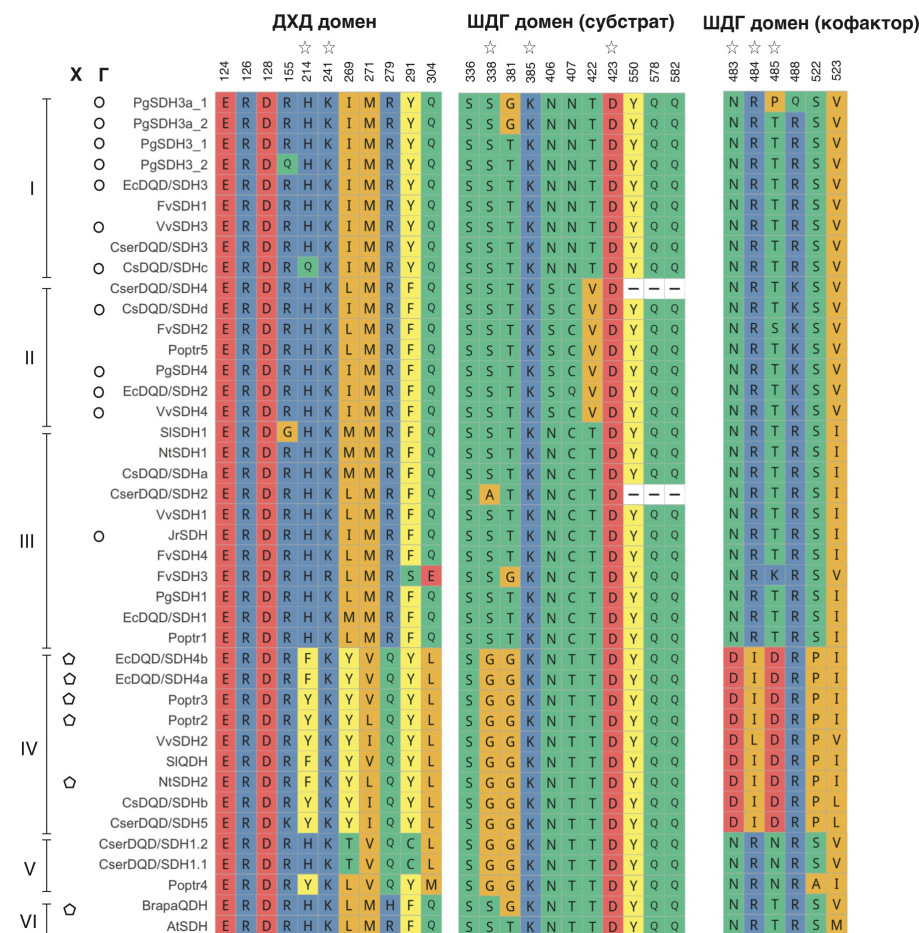
- Окончательная сборка транскриптома состояла из 29 124 транскриптов общей длиной 48 139 967 п.н. Самый большой транскрипт длиной 27 663 п.н.
- По данным BUSCO, сборка охватила 88.4% ортологических генов из набора данных эвдикот, 90.4% генов растений и 90.6% эукариотических генов, что характеризует ее как достаточно полную.
- При анализе GO данных только 19 098 транскриптов (65,6%) были аннотированы. С помощью KEGG Mapper были реконструированы карты путей, из которых 3741 относились к метаболизму



Идентифицировано пять потенциальных изоформ бифункциональной 3-дегидрохинатадегидратазы / шикиматдегидрогеназы

- Диапазон длин аминокислотных последовательностей 500-532 АК
- Предсказанная молекулярная масса 54.3-57.8 кДа
- Наибольшая вероятность пластидной локализации – для изоформ 3 и 4
- Проведен сравнительный филогенетический анализ и анализ ключевых аминокислот изоформ 3-ДХД/ШДГ *C. sericea* L.

Изоформы 3-ДХД/ШДГ *C. sericea* L. структурно отличаются друг от друга, что может предопределять функциональные различия: наличие «классической» ШДГ активности у изоформы 3 и ХДГ активности с отсутствием ДХД активности у изоформы 5. Изоформы 3 и 4 возможно обладают галлат-образующей активностью. Изоформа 1 может проявлять сочетанную активность



Заключение

Данное исследование является первым шагом к пониманию молекулярных механизмов синтеза гидролизуемых танинов в *C. sericea* и может послужить основой для дальнейших исследований биосинтеза вторичных метаболитов в растениях. Идентификация пяти изоформ 3-ДХД/ШДГ открывает новые возможности для изучения их специфических функций и регуляции синтеза гидролизуемых танинов у *C. sericea*.

Автор для переписки: Евгения Никонорова (Eugenia Nikonorova)

к.м.н., в.н.с. лаборатории геномики и биохимии
лекарственных растений ФГБНУ ВИЛАР

e-mail: gatiatulinaer@gmail.com
+7 (922) 855 72 78

Мы благодарим Алексея Пенина (Aleksey Penin) за предоставленные праймеры для проведения ONT секвенирования, а также Марию Логачеву (Maria Logacheva) за помощь в планировании эксперимента

ONT секвенирование проведено с использованием научного оборудования ЦКП «Геномика, протеомика, метаболомика» ФГБУ ФНКЦ ФХМ им. Ю.М. Лопухина ФМБА России

