

2. Структура и организација на растителен геном

Вовед

Секоја карактеристика на еден организам е кодирана во информацијата која се содржи во DNA-молекулите, односно во геномот. Техниките на секвенционирање на DNA овозможиле добивање на комплетни геномски секвенци кај различни микроорганизми и виши еукариотски организми, вклучувајќи голем број растителни видови.

Повеќето гени кај организмите се наоѓаат во јадрените DNA-молекули. Исто така, пластидите и митохондриите кај растенијата содржат сопствени DNA-молекули за кодирање на генски производи кои се неопходни за функцијата на клеточните органели и репродукцијата на растенијата. Така на пр., некои особини кои се кодираат од гените на клеточните органели имаат комерцијална примена кај некои растителни култури, како што е цитоплазматската машка стерилност кај сирак (*Sorghum* sp.) и пченка (*Zea mays*). Особините кои потекнуваат од гените на клеточните органели можат да се препознаат преку нивното наследување од едниот родител (унипарентално наследување). Кај повеќето растителни видови, митохондриите и хлоропластите се наследуваат само преку наследниот материјал од мајката-растение.

Во педесеттите години на минатиот век било откриено дека одредени делови од генетскиот материјал можат да се движат од едно на друго место во геномот. Подоцна било потврдено дека овие мобилни генетски елементи претставуваат универзални компоненти на геномите. Според тоа, овие мобилни генетски елементи имаат поголемо влијание врз промените на геномот во споредба со точкестите мутации и на таков начин имаат значаен придонес во еволуциските промени кај растенијата.

2.1 Структура на растителен геном

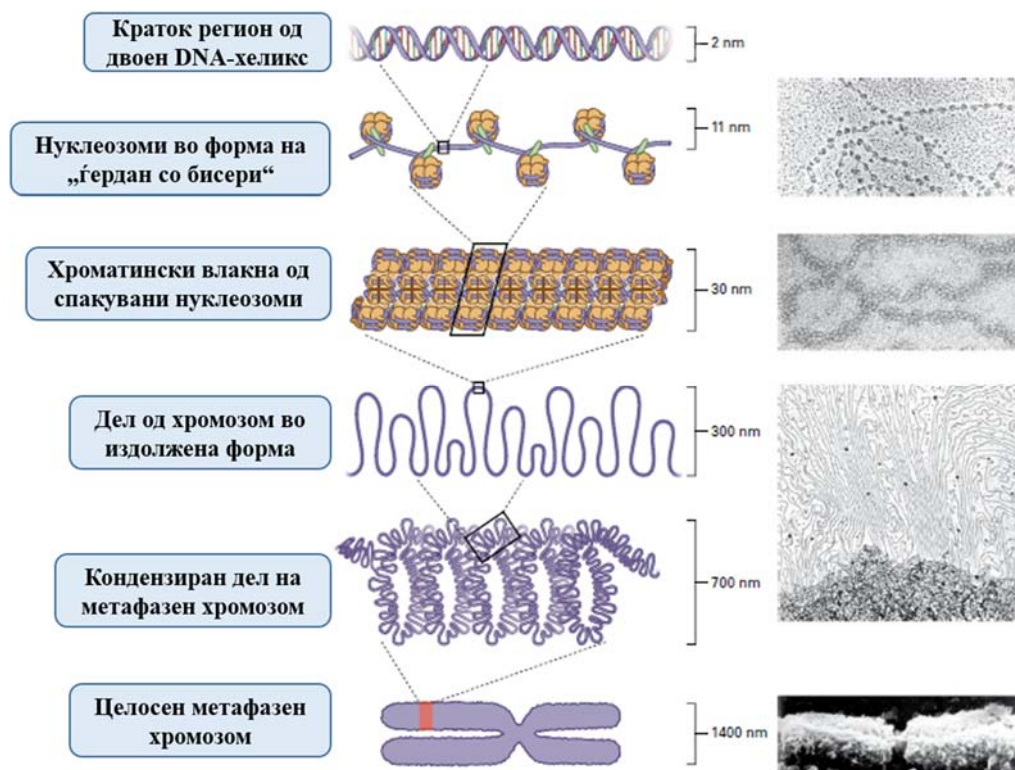
Првото комплетно секвенционирање на растителниот геном било објавено во 2000 година кај *Arabidopsis thaliana*. Од тогаш, податоците за структурата и содржината на растителните геноми биле добиени со примена на голем број техники кои биле развиени во текот на 20^{-тиот} век.

2.1.1 Микроскопијата при висока резолуција открива детали за структурата на хромозомите

Хромозомите за прв пат биле визуелизирани на почетокот на 20^{-тиот} век и се сметале како носители на наследниот материјал. Подоцна било откриено дека хромозомите кај вишите еукариотски организми содржат двојноверижни DNA-молекули кои се организирани во компактни хромозомски структури. Овие хромозомски структури за време на метафаза од митозата имаат должина од неколку микрометри. Зголемувањето на резолуцијата на микроскопските техники овозможило визуелизација на особините на хромозомите до степен на DNA-секвенца (Сл. 2.1).

Примената на светлосната микроскопија овозможува визуелизација на основните особини на хромозомите во метафаза од митозата: центромери кои се одговорни за сегрегација на сестринските хроматиди при митоза, хромозомски краци и теломери (Сл. 2.1). Исто така, хромозомите можат да се визуелизираат при понизок степен на кондензација за време на

пахитен од мејоза. Во оваа фаза, хромозомите се користат како супстрати за флуоресцентна *in situ* хибридизација (FISH) за да се одреди локацијата на одредени типови на DNA-секвенци во хромозомите. При овој степен на резолуција може да се направи дистинкција помеѓу хетерохроматинот и еухроматинот како хетерохроматински јазли. Во границите на резолуцијата на светлосната микроскопија, DNA-нишките можат да се визуелизираат и да се примени FISH-методот или да се дигестираат со рестриктивни ензими за да се добие оптичка мапа.

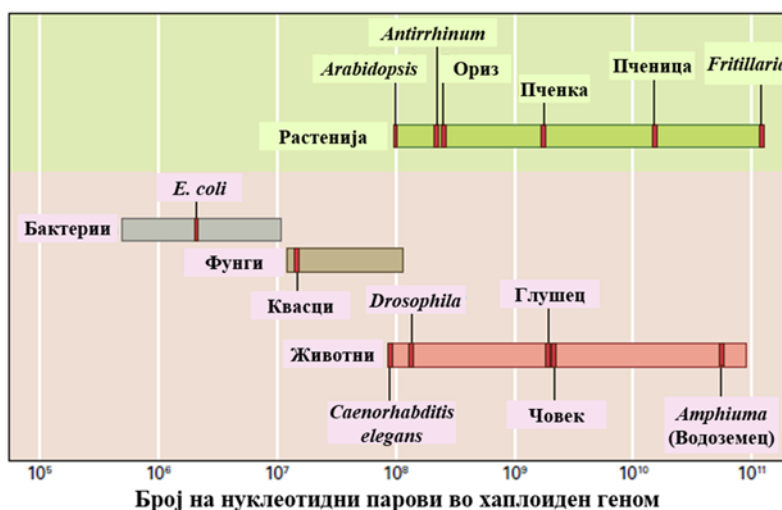


Слика 2.1 Организација на DNA и протеини во хроматин при различни степени на кондензација. На левиот панел се претставени шематски прикази на DNA, намотана DNA околу нуклеозомите, спакувани нуклеозоми, издолжени јамки од спакувани нуклеозоми, кондензирана DNA во метафазни хромозоми и цели метафазни хромозоми. Сликите на десниот панел се скенинг-електронски микрографии со различен степен на резолуција (Buchanan и сор., 2015).

При највисок степен на резолуција можат да се визуелизираат цели хромозоми преку скенинг-електронска микроскопија. Всушност, хромозомите можат парцијално да се дезинтегрираат со цел да се открие нивниот нуклеопротеински матрикс за кој се прикрепуваат DNA-веригите. Исто така, хромозомите можат да се визуелизираат како „ѓердан со бисери“, при што се открива начинот на кој хистоните учествуваат во пакување на DNA во хроматин (Сл. 2.1). Основниот степен на хроматинската структура е **нуклеозом**, кој претставува кондензација од прв степен за да се разликува од натамошните повисоки степени на кондензација. Всушност, нуклеозомот се состои од двојно намотана DNA (166 bp) околу глобуларен октамер од хистонски протеини, кои сочинуваат два тетрамери, а секој тетрамер е составен од хистоните H2A, H2B, H3 и H4.

2.1.2 Кај скриеносемените растенија постои значајна варијација во големината на геномот

Големината на јадрениот геном помеѓу организмите со хаплоидна содржина на DNA (**C-вредност**) варира од 10^7 до 10^{11} базни парови (bp). Хуманиот геном има интермедиерна вредност од овој опсег која изнесува 3×10^9 bp. Според поранешните претпоставки, комплексноста на организмот е во корелација со големината на геномот. Така на пр., хуманата популација има поголем геном од повеќето инсекти, додека пак, инсектите имаат поголем геном во споредба со фунгите. Сепак, оваа корелација не е универзална бидејќи некои водоземци имаат речиси 50 пати поголем геном во споредба луѓето, како и 'рскавичните риби кои имаат поголем геном од коскените риби. Растителните геноми се карактеризираат со широка варијација во големината. Така на пр., најмалиот геном припаѓа на *Arabidopsis thaliana*, додека пак, еден од најголемите геноми се наоѓа кај *Fritillaria assyriaca* од фамилијата Liliaceae (Сл. 2.2).



Слика 2.2 Големината на хаплоиден геном во базни парови (C-вредности) кај различни организми. Повеќето еукариотски организми имаат хаплоиден геном со големината на DNA од 10^7 до 10^{11} bp (Buchanan и сор., 2015).

Во средината на 20-тиот век бил воведен терминот **парадокс на C-вредноста** за да се опише недостатокот на директна корелација помеѓу големината на геномот и комплексноста на организмите. Секвенционирањето на голем број геноми кај организмите овозможило откривање на особените на секвенците кај различни класи на DNA кои имаат придонес во овој парадокс. Варијацијата во големината на геномот кај растенијата може да се должи на една или повеќе серии на геномска дупликација, експанзија на фамилиите на (ретро)транспозони, како и на сегашната полиплоидија.

Секвенционирањето на DNA претставува ефикасна технологија за прецизно пресметување на големината на растителните геноми. Сепак, кај голем број растителни и анимални геноми воопшто не се достапни податоци за DNA-секвенци. Во овие случаи, пресметувањето на содржината на DNA во изолирано јадро кое е обоено со флуоресцентна боја претставува релативно прецизна техника за одредување на големината на геномот.

2.2 Организација на растителен геном

2.2.1 Јадрените геноми содржат единични копии на секвенци и репетитивни DNA-секвенци

Пред воведување на технологијата на DNA-секвенционирање, составот на геномските секвенци бил проучуван со Cot-анализа. Всушност, оваа анализа се применува за одредување на содржината на репетитивни DNA-секвенци во примероци. Принципот на Cot-анализата се базира на кинетика на реасоцијација на DNA, која вклучува денатурација на DNA-примерок, а потоа се овозможува негово реанилирање, при што се следи преостаната фракција на едноверижна DNA во функција од времето. Терминот „Cot“ произлегува од графичкиот приказ на Cot-анализата кој ја претставува фракцијата на едноверижна DNA во функција од концентрацијата на DNA при нулта време (c_0) и време на мерење (t). Ова претставува реакција од втор ред, при која едноверижните молекули задолжително се ренатурираат со нивните комплементарни молекули. Како резултат на зголемување на геномот, бројот на копии на секоја секвенца на грам DNA се намалува, при што ренатурацијата се одвива подолго време.

Со примена на овој метод, геномот може да се подели на три компоненти: високо репетитивна DNA, интермедиерно репетитивна DNA и единични копии на DNA. Иако постојат голем број на достапни DNA-секвенци, овие широки категории на геномски компоненти сè уште се користат за одредување на структурата на растителниот геном. Сепак, границите помеѓу секоја од овие компоненти на геномот не се целосно разјаснети. Единичните копии на DNA (single-copy DNA) се присутни само во една копија во хаплоидниот геном и се сметаат како низок број на копии на DNA (low-copy DNA). Оваа класа на DNA-секвенци ги содржат речиси сите кодирачки генски секвенци во геномот. Кај повеќето растителни геноми, половина од гените припаѓаат на генски фамилии со различна големина. Во зависност од критериумот за сличноста во генските секвенци (>80% идентичност во нуклеотидите), сродните гени имаат способност да хибридизираат и не можат да се класифицираат како единична копија.

Високо репетитивните DNA-секвенци можат лесно да се дефинираат и карактеристичен пример се репетитивните тандем-секвенци, кои се присутни во центромерите на растителните хромозоми. Некои високо репетитивни секвенци имаат значајно различен состав на бази во споредба со геномот и се нарекуваат сателитска DNA, бидејќи овие секвенци биле сепарирани со примена на градиент на густина со цезиум хлорид. Кај повеќето животни и квасци, сателитската DNA е богата со базите AT, додека пак, сателитската DNA кај растенијата има висока содржина на базите GC. Сателитската DNA примарно е асоцирана со центромерот или со теломерите кај растенијата и најчесто има хетерохроматична структура. Според тоа, сателитската DNA има кондензирана структура и останува неактивна во интерфазата од клеточната делба.

Интермедиерните репетитивни DNA-секвенци се присутни од неколку десетици до десетина илјади копии во геномот. Постојат неколку илјади фамилии на интермедиерни репетитивни DNA-секвенци, каде што претставниците од една фамилија имаат многу слични секвенци кои не се идентични помеѓу себе. Притоа, претставниците од една фамилија можат вкрстено да се хибридизираат помеѓу себе бидејќи се карактеризираат со повеќе од 80% идентичност во нуклеотидните секвенци. Во рамките на оваа широка категорија на репетитивни DNA-секвенци се вбројуваат: транспозони, ретротранспозони,

структурни RNA молекули, како што се рибозомални, транспортни и други RNA-форми, како и некои големи протеин-кодирачки генски фамилии. Голем број од овие фамилии на репетитивни секвенци се јавуваат како тандемски повторувања, како што е примерот со рибозомалните RNA молекули во нуклеоларниот организирачки регион (nucleolar organizer region-NOR) и 5S RNA-кластерите. Некои примери од големи протеин-кодирачки генски фамилии кои се јавуваат како тандемски повторувања се гените за хистони, како и гените за резистентност кон заболувања.

2.2.2 Генетските и физичките мапи овозможуваат комплементарен преглед на организацијата на геномот

Генетските мапи претставуваат организирани листи на дистинктивни особини (маркери) по должината на секој хромозом, како и релативното растојание помеѓу нив (одредено преку фреквенцијата на рекомбинација помеѓу нив за време на мејозата). Овие маркери можат да се базираат на DNA-секвенци или фенотипски особини (пр. висина на растение, форма на лист, резистентност кон заболувања или други стрес фактори).

При споредување на генетските мапи со секвенционирани геноми, постои висок степен на конзистентност помеѓу редоследот на маркерите кои се базирани на DNA-секвенци во генетската мапа и појавувањето на овие маркери во самиот геном. Во некои проекти на секвенционирање на геноми, генетските мапи можат да го олеснат одредувањето на редоследот и ориентацијата на DNA секвенците. Синхронизирањето на генетската мапа со соодветна геномска секвенца овозможува идентификација на гените-кандидати од генетската мапа за одреден фенотип од интерес (пр. принос на семе или состав на масло). Исто така, голем број на методи можат да се користат за да се идентификуваат гените-кандидати кои го контролираат фенотипот од интерес.

За разлика од генетската мапа која ги дефинира хромозомите во однос на рекомбинацијата помеѓу локусите за да се добие релативното генетско растојание помеѓу нив (генетски простор), **физичката мапа** претставува организиран приказ на голем број големи DNA-фрагменти (100 kb) кои се добиени со деградација на геномот. Често пати, главен предуслов за секвенционирање на растителните геноми е подготовката на физичка мапа. При физичкото мапирање со висока резолуција се користат клонирани секвенци на DNA со цел да се добијат мапи од соседни, препокривачки DNA-фрагменти, кои можат да опфатат многу долги растојанија. Овој регион од препокривачки DNA-фрагменти се нарекува **контиг (contig)**, бидејќи се формира од соседни генски клонови.

Методологијата за подготовка на физичка мапа опфаќа конструкција на библиотека од бактериски вештачки хромозом (bacterial artificial chromosome-BAC). Овие BAC се вектори за клонирање на плазмиди, кои го користат изворот на репликација на F-факторите и на таков начин се пропагираат во *E. coli* како единични плазмиди или плазмиди со многу низок број на копии. Нискиот број на копии и примената на клетки-домаќин кои се дефициентни во рекомбинација имаат значаен придонес кон стабилноста на DNA-молекулите. Повеќето BAC-библиотеки имаат инсерции од 100 до 150 kb геномска DNA. Потоа се користи техниката која е наречена фингерпринтинг „fingerprinting“ за да се одреди начинот според кој овие случајно клонирани фрагменти се прекриваат. Според оваа техника, DNA-молекулот од секој BAC се дигестира со еден или повеќе рестрикциски ензими и се одредува големината на добиените фрагменти. Притоа, секој BAC продуцира карактеристичен сет на фрагменти, при што два BAC-фрагменти се препокриваат (нивните

DNA-инсерти произлегуваат од ист регион на геномот) ако имаат одреден број на бандови со иста големина.

Компарацијата на големините на бандовите се одвива на агарозен гел, иако може да се користи и нова техника позната како фингерпринтинг со висока информација за содржината (high information content fingerprinting). При оваа техника се користи комбинација на повеќе рестрикциски ензими и означување со четири бои, што е проследено со капиларно DNA секвенционирање за да се добијат поголем број на фрагменти со прецизна големина, како и мапа со повисока резолуција. Овој метод на одредување на препокривањето на фрагменти е заменет со сет од техники за означување на секвенците за секој BAC, како и користење на мултиплекс секвенционирање (геномско профилирање).

Оптичкото мапирање претставува натамошна обработка на физичкото мапирање. Притоа, изолирани DNA-молекули (од неколку стотици килобазни парови во должина) се имобилизираат на стаклена површина и се дигестираат со рестрикциски ензими. Независно од методот што се користи за означување на секој BAC или DNA-секвенца, контизите на препокривачките BAC или DNA-секвенци се асемблираат врз база на заедничките ознаки.

2.2.3 Секвенционирањето на растителниот геном овозможува преглед на геномската структура со висока резолуција

Секвенционирањето на јадрените геноми кај растенијата започнало во 1997 година. Видот *Arabidopsis thaliana* бил прифатен како модел за растителна геномика поради две есенцијални особини: мал геном (≈ 120 Mb) и мал број на репетитивни DNA-секвенци. Во текот на следната деценија бил забележан значаен напредок на техниките на секвенционирање, асемблирање на секвенци и геномска биоинформатика. Овој напредок во технологиите овозможил секвенционирање на растителни видови со комплексни геноми, како што е пченицата (*Triticum* sp.), (Сл. 2.2). Во следниот текст се прикажани принципите на различни технологии што се користат за добивање на секвенци на целокупниот геном, како и предностите и недостатоците на различните методи.

Првите проекти за секвенционирање на геномот се базирале на BAC-методот, при што била конструирана физичка мапа од BAC-секвенци кои биле индивидуално секвенционирани. Индивидуалните BAC-секвенци биле одвоено асемблирани, а потоа споени врз база на преклопувањето на секвенци со помош на физички и генетски мапи за да се добијат комплетни хромозоми (познати како псевдомолекули). Со оглед на тоа што овој пристап вклучува интензивна лабораториска работа и е неопходно секвенционирање на илјадници BAC-секвенци, натамошните истражувања биле насочени кон развој на методи за „шотган на цел геном“ (whole genome shotgun-WGS).

Овој WGS-метод опфаќа фрагментација на целиот геном и секвенционирање на DNA во еден чекор и започнува интензивно да се користи поради зголемување на софистицираноста на алгоритмите за асемблирање на секвенци и моќта на компјутерските техники. Главен недостаток на WGS-методот претставува ко-асемблирањето на секвенци со висок степен на сличност кои потекнуваат од различни делови на геномот. Овој недостаток се јавува и кај технологиите на DNA-секвенционирање со високи перформанси кои се карактеризираат со висок степен на грешка и читање на кратки секвенци.

Исто така, проектите за секвенционирање на геномот опфаќаат секвенционирање на краевите на различни фрагменти со дефинирана големина со цел да се добијат т.н. секвенци на „соседен пар“ (“mate pair”), кои се користат за идентификација на DNA-репетитивните

секвенци. Во повеќето случаи, геномите со ниска содржина на репетитивни DNA-секвенци не можат целосно да се асемблираат и на таков начин значаен дел од геномот претставува колекција на непрочитани секвенци. Генетските и физичките мапи кои иницијално биле користени за валидација на секвенционирањето на геномот врз база на ВАС-системот, денес имаат важна улога во редоследот и ориентацијата на асемблираните секвенци од геномот.

2.2.4 Геномската анотација ги открива биолошките функции кои се кодирани во DNA-секвенците

Геномската анотација е процес на идентификација на локацијата на гените, како и одредување на структурата (егзони, интрони и други особини) и функцијата на кодирачките региони во геномот. Структурните особини на гените се одредуваат со примена на два метода. Првиот метод се базира на достапните секвенци на mRNA (секвенционирани како cDNA) и протеините (одредени од нуклеотидните секвенци или од директно секвенционирање на протеини). Различни алгоритми се користат за подредување на препостоечките секвенци кон нов геном. Голем број од овие алгоритми можат да препознаат конзервирани сплајсинг места (GT...AG) за да се дефинираат границите помеѓу егзонот и интронот. Прецизните генски структури се дефинираат преку cDNA-молекули кои целосно ја опфаќаат протеин-кодирачката секвенца на одреден ген. Подредувањето на нуклеотидните и протеинските секвенци од други видови има информативен карактер, меѓутоа станува помалку значајно бидејќи се зголемува еволуиската дистанца помеѓу новиот геном и достапните секвенци. Исто така, целокупното подредување на два слични генома може да има информативен карактер, бидејќи конзервираните секвенци помеѓу двата генома најверојатно имаат функционално значење и кодираат гени или регулаторни секвенци. Вториот метод за дефинирање на структурата на гените е задолжително компјутерска техника која се базира на разликите во составот на базите и фреквенцијата на соседните нуклеотиди (ди-, три-, тетра-нуклеотиди), како и веројатноста за нивна транзиција помеѓу егзоните, интроните и интергенските секвенци. Често пати, овие софистицирани програми вклучуваат т.н. скриени Markov-и модели (hidden Markov models). Овие модели се имплементираат на мали и дефинирани структури кои биле идентификувани во новиот геном (или сличен геном) со користење на постоечки секвенци. Овие два метода се користат за идентификација на структурите на гени низ целиот нов геном, односно во регионите за кои што постојат евидентни податоци од нуклеотидните и протеинските секвенци, како и во регионите за кои не постојат никакви податоци. Сепак, структурите на гените кои се потврдени исклучиво со примена на компјутерските методи не се сметаат за сигурни податоци во растителната геномика.

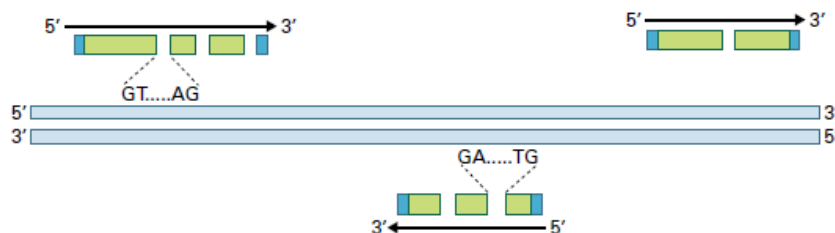
Предвидувањето на генската структура овозможува дефинирање на протеин-кодирачката секвенца за секој ген. Идентификацијата на функцијата на овие гени се базира на нивното споредување со другите гени во databазите, како што се GenBank или SwissProt. Во овој случај, споредувањето се врши на ниво на аминокиселинскиот состав, бидејќи протеинските секвенци се поконзервирани помеѓу видовите во споредба со нуклеотидните секвенци. Потоа, функцијата на споредување на databазата се пренесува на новиот протеин. Овој метод може да има потенцијално погрешна анотација, бидејќи само кај мал број протеини од databазите е утврдена нивната функција преку експериментални докази. Овој феномен на „пропагација на минлива грешка“ (transitive error propagation) ќе се намали во иднина, бидејќи експериментално се потврдуваат функциите на голем број протеини.

2.2.5 Секвенционирањето на DNA има значаен удел во одредувањето на организацијата на геномот

Организацијата на геномот од *Arabidopsis* претставува ефикасен модел и појдовна точка за одредување на структурата на голем број растителни геноми. Проектот за секвенционирање на геномот кај *Arabidopsis* овозможил добивање на 105 Mb од DNA-секвенца со несеквенциониран дел околу центромерите кој изнесува околу 15 Mb од целиот геном со вкупна големина од 120 Mb. Натамошните анализи со точна цитометрија покажале дека содржината на несеквенционирана DNA е поголема и се однесува за геном со вкупна големина од околу 150 Mb. Сепак, овие истражувања немаат значајно влијание врз основните особини на геномот кај повеќето други растителни геноми. На сличен начин како и кај повеќето растенија и животни, малата големина на геномот кај *Arabidopsis* е во корелација со неговиот краток животен циклус, односно растенијата со големи геноми имаат тенденција за бавен репродуктивен циклус. Денес, главните особини на хромозомите можат да се опишат на ниво на DNA-секвенци (Сл. 2.1).

Повеќето гени се локализирани во еухроматинските хромозомски краци

Хромозомските краци кои се протегаат од теломерите до перицентромерните региони во близина на центромерите опфаќаат секвенци од околу 100 Mb и се карактеризираат со униформна содржина на гени. Геномот на *Arabidopsis* кодира околу 30 000 протеин-кодирачки гени, како и голем број гени за транспортни RNA-молекули (tRNA) и различни класи на некоидачки RNA-гени со регулаторни функции. Овој број постојано се зголемува како резултат на откривањето на нови мали гени кои можат да бидат протеин-кодирачки гени или регулаторни гени. Во просек, протеин-кодирачкиот ген кај *Arabidopsis* содржи пет егзони и опфаќа околу 2 kb од геномската секвенца од старт до стоп кодонот. Регулаторните секвенци на генот можат да се наоѓаат спротиводно, низводно или во рамките на овој протеин-кодирачки регион. Сепак, повеќето промотори се локализирани спротиводно од транскрипциското стартно место.



Слика 2.3 Гените (протеин-кодирачки секвенци) се наоѓаат во едната или другата DNA-верига, а поретко во истиот регион од двете вериги. Со оглед на тоа што транскрипцијата секогаш се одвива во правец 5'→3', гените од спротивните вериги се транскрибираат во спротивен правец. Транскрипцијата може да биде конвергентна или дивергентна. На сликата се прикажани ориентациите на граничните места на спојување на двете вериги. Зелените полиња го претставуваат протеин-кодирачкиот регион на транскриптите, додека пак, сините полиња ги означуваат 5' и 3' непреводливиите региони. Во ретки случаи, во ист регион на DNA можат да се наоѓаат два гена на спротивни вериги, при што едниот ген се наоѓа во интронот на другиот ген (Buchanan и сор., 2015).

Гените се одвоени со DNA-секвенца од околу 2,5 kb (вклучувајќи регулаторни региони), при што густината на гени изнесува еден ген на секои 4,5 kb-секвенци. Ова е пример за највисока густина на гени кај растенијата и може да се спореди со геномите кај *Brachypodium* (Poaceae), *Physcomitrella* (Bryophyta) и едноклеточната зелена алга *Chlamydomonas*. Со неколку исклучоци, само едната верига од било кој регион на двојниот хеликс кодира протеин. Според тоа, протеин-кодирачките гени можат да се наоѓаат на едната или другата верига и можат да се транскрибираат во било кој правец (Сл. 2.3).

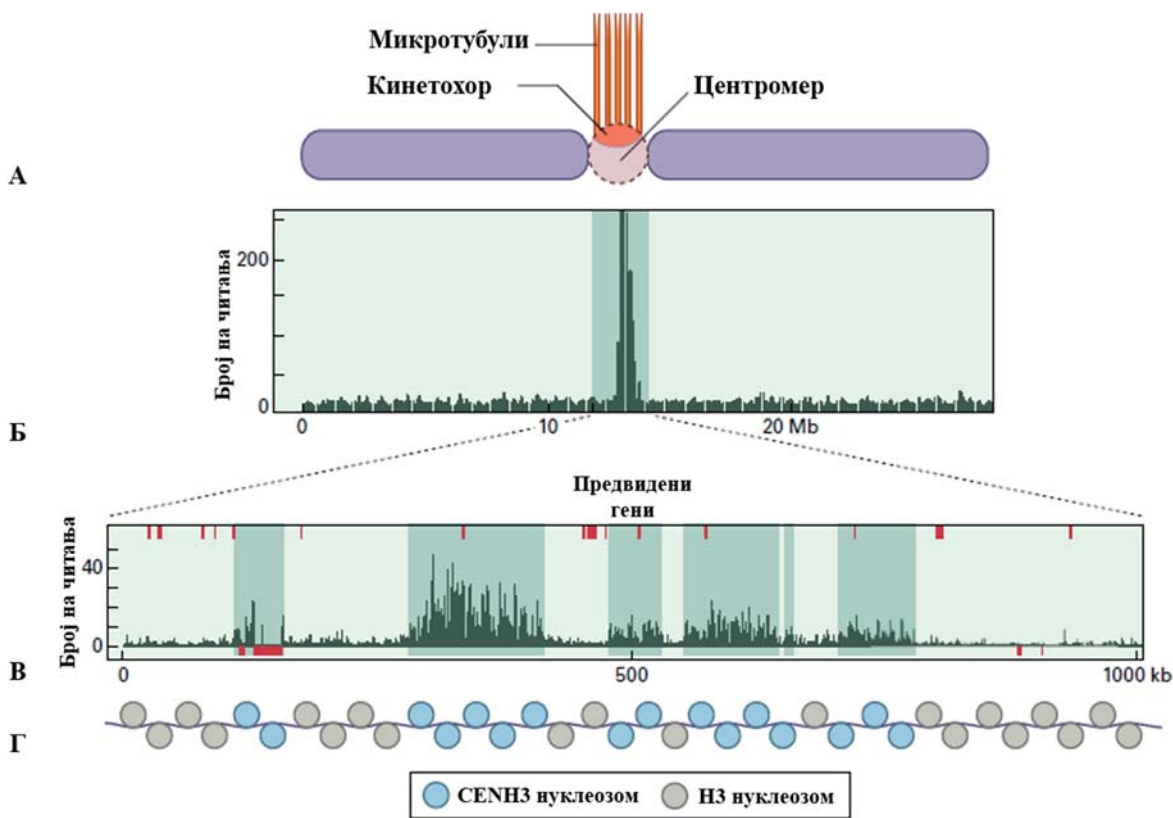
Теломерите ги заштитуваат краевите на хромозомите

Хромозомските краци завршуваат со теломери, кои претставуваат специјализирани структури за заштита на хромозомските краеве. Исто така, теломерите овозможуваат нивна прецизна репликација и оневозможуваат скратување на хромозомите за време на последователните циклуси на DNA-синтеза. Кај речиси сите растителни видови е идентификувана и секвенционирана консензус репетитивна секвенца (TTTAGGG) во теломерите. Теломерите кај *Arabidopsis* имаат должина од околу 2-3 kb.

Кај еукариотите е идентификуван специјализиран ензим наречен **теломераза** кој овозможува одржување на едноверижен регион по репликацијата на DNA. На таков начин, теломеразата учествува во заштита на хромозомот од скратување при секоја клеточна делба. Теломеразата претставува реверзна транскриптаза со интегриран RNA-молекул во тродимензионалната структура на ензимот. Всушност, RNA-молекулот функционира како урнек за нова едноверижна теломерна секвенца. Исто така, теломерите имаат функција во одржување на „нелепливи“ краеве на хромозомите. Цитогенетските анализи покажуваат дека краевите на оштетените (испрекинати) хромозоми се многу лепливи и можат да се спојат со било кој друг достапен DNA-фрагмент. Важно е да се напомене дека теломерите имаат улога во организација на хромозомите во јадрото преку прикрепување на хромозомите за фиброзната јадрена ламина од внатрешната страна на јадрената мембрана.

Центромерите се одговорни за прецизна сегрегација на хромозомите во новодобиените клетки за време на клеточната делба

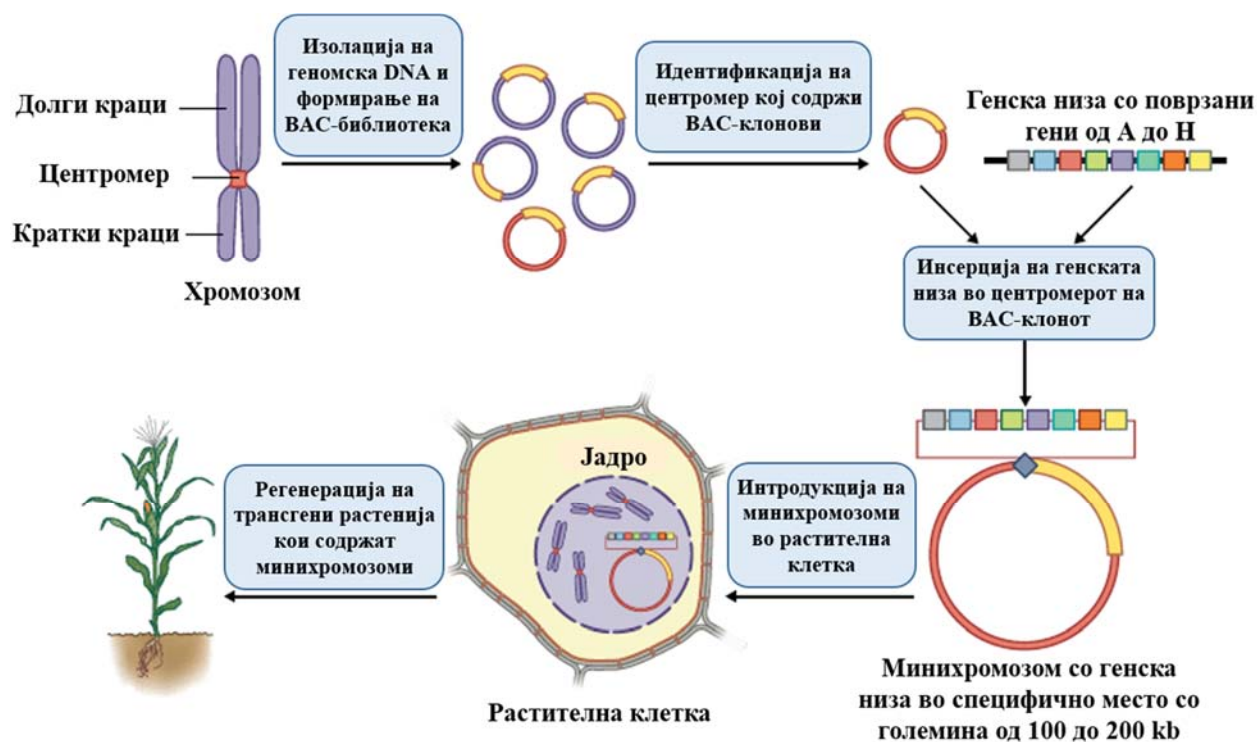
Центромерите на хромозомите (Сл. 2.4) се хетерохроматински (високо кондензирани) констрикции на хроматинот, кои се прикрепуваат за делбеното вретено за да се олесни сепарацијата на реплицираните хроматиди во митоза и мејоза. Квасците *Saccharomyces cerevisiae* и *Schizosaccharomyces pombe* имаат најмали центромери со должина од неколку десетина до стотина нуклеотиди. Сепак, центромерите кај повеќето растенија не се варијабилни и имаат поголема и покомплексна структура. Аналогно на алфа сателитските репетитивни секвенци кај центромерите од хуманиот геном, растителните центромери содржат тандемски низи од една или повеќе DNA-секвенци. Овие репетитивни секвенци се специфични за секој растителен вид и имаат должина помеѓу 150 и 200 bp (приближна должина на DNA која го обвиткува нуклеозомот). Исто така, центромерите кај растенијата се богати со ретротранспозонски секвенци. Центромерите можат лесно да се препознаат со помош на протеините кои се врзуваат за DNA од овој регион, особено хистонот CENH3. Кај квасците се идентификувани околу 50 протеини кои конститутивно или привремено се асоцирани со центромерите. Голем број хомолози на овие протеини (CENP-C, CENP-H-1, CCAN, ASK, CBF5 и Bub-слични протеини) се идентификувани кај растенијата.



Слика 2.4 Структура на центромерот од хромозомот 8 кај ориз (*Oryza sativa*). (А) Структура на хромозомот со хромозомски краци, центромер, кинетохор и микротубули. На сликата е прикажана само една сестринска хроматида. (Б) Локализација на CENH3 (центромер-специфична варијанта на хистон 3) во центромерниот регион во споредба со CENH3 низ целиот хромозом. Затемнетото поле го прикажува супресираниот регион за crossing over на хромозомот. (В) Зголемен преглед на CENH3-врзувачкиот домен. Предвидените гени (црвени полиња) главно се локализирани надвор од CENH3-врзувачките домени (сиви полиња), додека пак, репетитивните секвенци во центромерот со големина од 155 bp (CentO) се прикажани како црвени полиња под мапираните читања. (Г) Шематски приказ на нуклеозомот од центромерот на хромозомот 8 кај ориз. Сивите кругови претставуваат нуклеозоми со каноничен хистон H3, додека пак, сините кругови се нуклеозоми кои го содржат хистонот CENH3 (Buchanan и сор., 2015).

Големината на центромерите значајно варира помеѓу видовите (од 50 kb до повеќе Mb), како и помеѓу хромозомите од еден вид. Покрај ограничените познавања за функционалните елементи на растителните центромери, DNA-секвенците и асоцираните протеини, денес се врши дизајнирање на артифицијални хромозоми кои имаат способност за одржлива автономна репликација во друг организам. На таков начин се овозможува пренесување на важни метаболички особини во друга сорта или вид со цел да се подобрат карактеристиките на растенијата (Сл. 2.5).

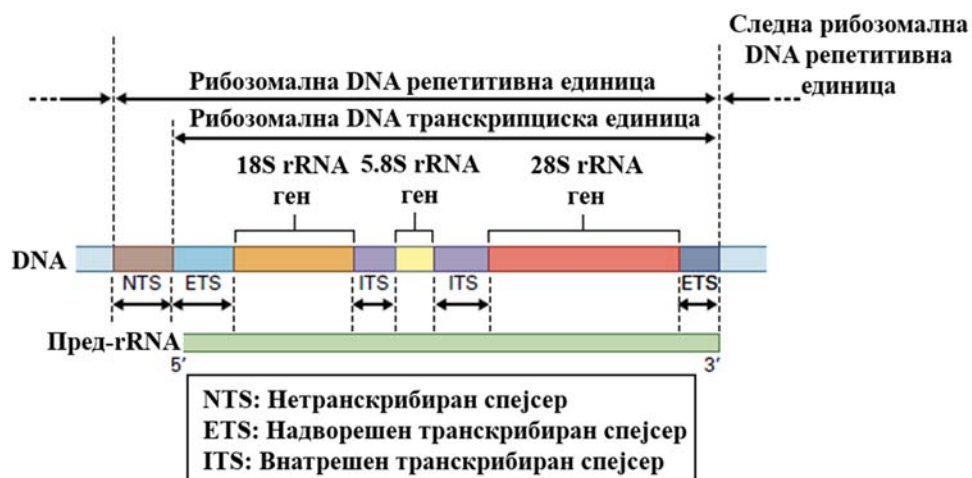
Перицентромерните региони кои ги ограничуваат центромерите од двете страни на хромозомот имаат должина од повеќе мегабази и се карактеризираат со ниска густина на гени, висока содржина на транспозони, особено ретротранспозони. Во овие перицентромерни региони е супресирана генетската рекомбинација.



Слика 2.5 Артифицијални минихромозоми кои носат сет од гени можат да се користат за генетски инженеринг кај растенијата (Buchanan и сор., 2015).

Гените кои кодираат RNA-компоненти на рибозомите се локализирани во мал број на геномски региони

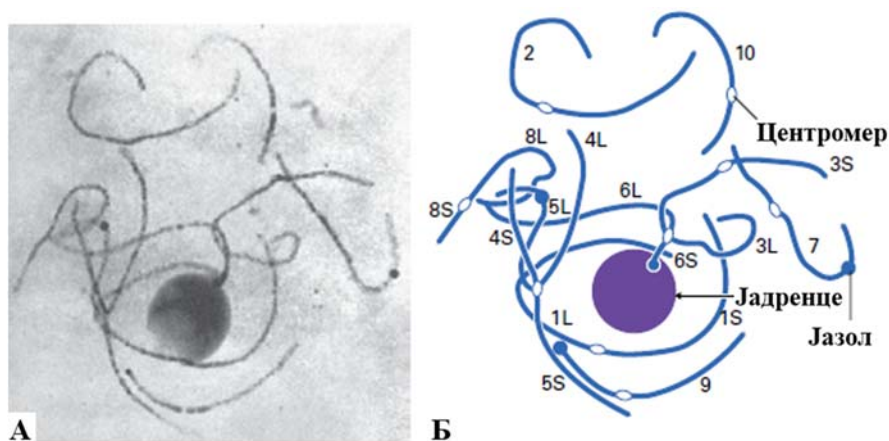
Нуклеоларните организирачки региони (Nucleolar organizer regions-NOR) се региони на хромозомот каде што се одвиваат процесите на транскрипција, биогенеза и асемблирање на рибозомалните RNA-молекули. Високата содржина на рибозомалната RNA (rRNA) е неопходна за одржување на структурата на рибозомите и синтезата на протеините во активните клетки. Според тоа, практично е невозможно една или неколку копии од rRNA-гени да ги задоволат потребите на клетките за транслација на протеини. Нуклеоларните организирачки региони се состојат од голем број тандемски репетитивни секвенци со должина од околу 10 kb кои ги кодираат 28S, 18S и 5.8S RNA-компонентите на рибозомите (Сл. 2.6). Кодирачките секвенци на овие три гени се високо конзервирани помеѓу видовите. Сепак, интергенскиот транскрибиран спејсер (intergenic transcribed spacer-ITS) е многу варијабилан регион и најчесто се користи за испитување на поврзаноста помеѓу видовите. Секој растителен вид содржи само еден или два нуклеоларни организирачки региони и нивниот број и локализацијата во хромозомите се одредуваат со примена на FISH-методот. Другата RNA-компонента на рибозомот (5S RNA) се кодира во тандемски низи, меѓутоа, овие кодирачки секвенци се наоѓаат на други локации во хромозомот. Кај *Arabidopsis*, главниот 5S RNA генски-кластер е детектиран во центромерните региони на хромозомите 3, 4 и 5.



Слика 2.6 Еукариотска рибозомална DNA репетитивна единица. Повеќето рибозомални гени кај растенијата се наоѓаат во рамки на репетитивна единица со големина од 7,8 до 185 kb. Репетитивната единица е составена од високо конзервирани rRNA-гени кои се одвоени со долги спејсер DNA-секвенци кои не се транскрибираат (NTS, ETS и ITS). Рибозомите содржат четири различни RNA-молекули со големина од 18S, 5.8S, 28S и 5S кои учествуваат во нивната структура. Гените за првите три RNA молекули се наоѓаат на рибозомалната DNA репетитивна единица, додека пак, 5S-генот има друга локација во геномот (Buchanan и сор., 2015).

Кондезираниот хроматин може да се визуелизира како хетерохроматински јазли

Хетерохроматинските јазли се препознаваат како специфични региони со високо кондензиран хроматин кои се одвоени од центромерите. Често пати, овие структури можат да се забележат кај растенијата со големи геноми, како што е пченката (Сл. 2.7). Единичниот хетерохроматински јазол кај *Arabidopsis* е локализиран на хромозомот 4 и доминантно е составен од ретротранспозони.



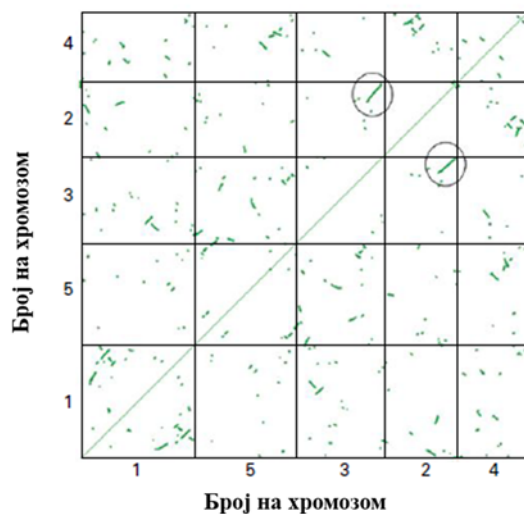
Слика 2.7 Хромозомите кај пченка (*Zea mays*) се карактеристичен пример за хетерохроматински јазли (Buchanan и сор., 2015). (А) Микрографија од микроспороцитно јадро кај пченка во стадиум на пахитен на која се прикажани 10 хромозоми и јадренце. (Б) Шематски приказ на идентификуваните хромозоми кај пченка. Хомологните хромозоми во стадиум на пахитен се поставуваат еден наспроти друг и се доволно кондензирани за да се овозможи идентификација на поединечните хромозоми според нивните структурни особини (позиција на центромер и хетерохроматински јазли).

2.2.6 Големите геноми имаат структурна комплексност и „генски острови“ кои се распоредени помеѓу големите региони на репетитивните секвенци

Сите гореспоменати структурни особини на хромозомите се карактеристични за голем број мали растителни геноми (<1 Gb). Од друга страна пак, геномот кај пченка (*Zea mays*), како и другите големи геноми кај Poaceae (треви/житни растенија) имаат поинаков тип на организација на геномот. Геномот на пченка се состои од околу 80% репетитивни DNA секвенци, а најголем дел од овие секвенци кодираат ретротранспозони. За разлика од повеќето растителни геноми кои се организирани во еухроматински хромозомски краци со висока густина на гени и хетерохроматински центромери со ниска густина на гени, геномот кај пченката се состои од т.н. „генски острови“ кои се распоредени помеѓу ретротранспозоните. Овие „генски острови“ се мали региони (околу 50 kb) кои содржат од еден до седум гени (во просек четири гени) и се наоѓаат на растојание кое е неколку пати поголемо од нивната должина. Овој модел на организација на геномот е карактеристичен и за другите претставници од Poaceae со големи и високо-репетитивни геноми, како што е пченицата (*Triticum* sp.) и јачменот (*Hordeum vulgare*). Сепак, не е доволно познато дали овој модел на организација на геномот е карактеристичен за другите монокотиледони растенија кои не припаѓаат на фамилијата Poaceae или дикотиледоните растенија со големи геноми, како што се зелената салата (*Lactuca sativa*), сончогледот (*Helianthus annuus*) или кромидот (*Allium cepa*).

2.2.7 Дупликациите на геномот за време на еволуцијата имаат главна улога во структурата на растителниот геном

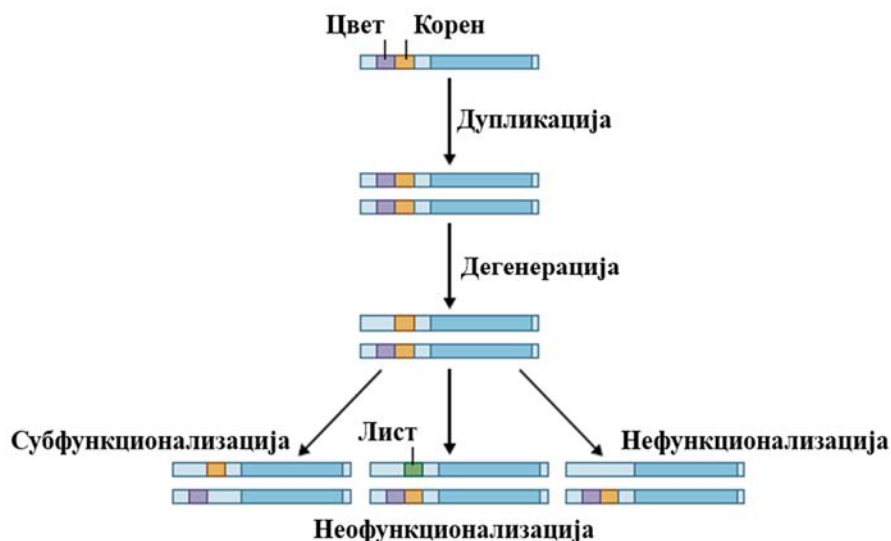
Секвенците на растителниот геном обезбедуваат јасни податоци за постоењето на интензивна дупликација на растителниот геном која се одвивала пред повеќе од 100 милиони години. Овие случувања за време на еволуцијата можат да се прикажат во форма на dot plot-анализа (Сл. 2.8), при која што координатите на двете оски претставуваат позиција на геномска особина, а секоја точка претставува регион на сличност помеѓу два дела од геномот.



Слика 2.8 Dot plot-статистичка анализа на геномот од *Arabidopsis thaliana* покажува настани на дупликација во текот на еволуцијата (Buchanan и сор., 2015). Хромозомите се наредени на координатите според големината (од најголем кон најмал хромозом). Секоја точка претставува кластер на слични гени кои се најдени на две различни места во геномот. Заокружените линии претставуваат регион од долниот дел на хромозомот 2 кој е дуплиран во долниот дел на хромозомот 3. На сликата се забележува симетрија околу дијагоналната линија која ја означува самоидентичноста.

Често пати, dot plot анализата се применува кај структурните особини на протеините, бидејќи протеинските секвенци се поконзервирани во рамките на долг еволуциски период во споредба со DNA-секвенците. На таков начин се овозможува детекција и визуелизација на раните еволуциски настани на дупликација. При dot plot-анализа на одреден геном, централната дијагонала ја претставува позицијата на истата конзервирана особина на двете координати, додека пак, серијата на точки кои формираат дијагонали покрај оваа централна линија претставуваат доказ за сегментални дупликации (DNA-сегменти со речиси идентични секвенци).

Времето на одвивање на дупликација на геномот може да се одреди од дивергенцијата на две секвенци. Ова се пресметува како синонимен степен на супституција во протеин-кодирачките секвенци, кој се калибрира преку референтни секвенци со фосилни податоци. Всушност, синонимните супституции се однесуваат на промените во третата база на кодоните кои не влијаат врз аминокиселинската секвенца на протеинот. Еволуциската историја на растителните DNA-секвенци претставува комплексна задача поради акумулацијата на голем број податоци и примената на софистицирани методи на анализа. Процесот на дупликација на геномот може да има влијание врз губењето на гени и промените во функцијата на гените. Термините субфункционализација и неофункционализација се однесуваат на дупликација на поединечни гени и не опфаќаат дупликација на цел геном. При субфункционализација, две генски секвенци кои потекнуваат од оригинална единична генска копија добиваат различна функција (начин на експресија и биохемиската активност) во споредба со оригиналниот ген. Неофункционализацијата претставува процес кога едниот дуплиран ген превзема нова функција, која претходно не била поврзана со неговиот родителски ген (Сл. 2.9).

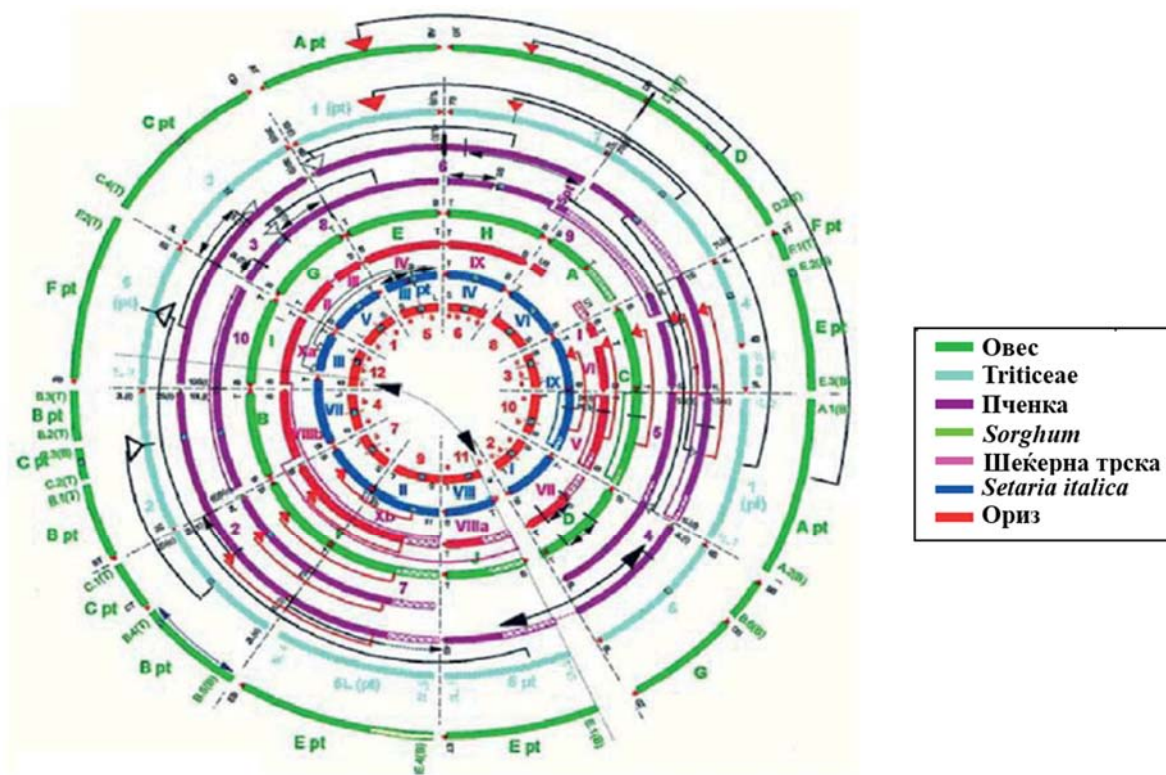


Слика 2.9 Субфункционализација и неофункционализација. Единичен ген содржи промоторни елементи кои ја регулираат експресијата во цветот и коренот. По одвивање на процесите на дупликација и дегенерација, постојат три можни резултати. При субфункционализација, секој ген ги задржува промоторните елементи и станува специјализиран за експресија во само еден од двата растителни органа. При неофункционализација, еден од промоторните елементи мутира и ја регулира експресијата во листовите. При нефункционализација, едната копија на генот повеќе не се експресира и овој ген се нарекува псевдоген (Buchanan и сор., 2015).

Овие промени можат да се јават во регулаторниот или структурниот (протеин-кодирачки) регион на генот. Откако ќе се изврши дупликацијата, еден од дуплираните гени може да дивергира и еволуира без да ја одржи својата оригинална функција, што може да биде од есенцијално значење за преживување на организмот.

2.2.8 Сродните растителни видови имаат конзервирана организација на единични копии на гени

Истражувањата за генетско мапирање и геномско секвенционирање покажале дека содржината и редоследот на гените по должината на големи сегменти од хромозомите се конзервирани помеѓу сродните видови на растенија и животни. Оваа колинеарност на генетските локуси или гени се нарекува **синтенија**. Првиот впечатлив пример за синтенија кај голем број растителни видови бил прикажан кај фамилијата Poaceae (Сл. 2.10). Сепак, секвенционирањето на повеќе геноми показало дека овој тип на зависност може да се забележи и кај други фамилии на растенија, како Solanaceae (домат, компир, пиперка, модар патлиџан) и Leguminosae (*Medicago*, *Lotus*, грав, соја, грашок).



Слика 2.10 Синтенија помеѓу различни геноми на треви (Buchanan и сор., 2015). Компаративна консензус-мапа кај треви. На сликата се прикажани податоци од голем број истражувања: овес-пченица-пченка-ориз; пченица-ориз; пченица-ориз; пченка-пченица; пченка-Sorghum-шеќерна трска и *Setaria italica*-ориз. Стрелките означуваат инверзии и транспозиции за да се опишат денешните хромозоми. На сликата се прикажани локациите на теломерите (Δ) и центромерите (□). Шпартаните површини покажуваат хромозомски региони за кои постојат мал број на компаративни податоци. Номенклатурата на хромозомите ги опишува нивните краци (кратки/долги или горе/доле).

Кај еволутивно оддалечените видови, синтенијата примарно се констатира според редоследот на ортологните протеински секвенци, а поретко преку компарација на нуклеотидните секвенци. Должината на овие синтенски секвенци се намалува кај еволутивно оддалечените видови. Сепак, кратки секвенци на синтенија можат да се идентификуваат и кај видови од различни фамилии кои дивергирале пред повеќе од 50 милиони години (микросинтенија).

2.3 Транспозонски елементи

Транспозонските елементи се DNA-секвенци кои се движат или преместуваат од едно на друго место од геномот. Овие мобилни DNA-елементи опфаќаат значаен дел од јадрениот геном и за време на транспозицијата можат со себе да ја носат генетската информација. На таков начин, транспозоните имаат особено значајна улога при организација на геномот. Транспозонските елементи од различни организми (*Drosophila*, *Saccharomyces cerevisiae* и *Zea mays*) покажуваат значајна конзервација во организацијата, секвенцата и начинот на транспозиција. Транспозоните се класифицираат според организација на секвенците и начинот на транспозиција (Табела 2.1).

2.3.1 Ретротранспозоните се мобилни елементи кои се движат преку RNA-интермедиери

Во зависност од начинот на транспозиција, основните типови на транспозонски елементи се поделени во две класи. Класа I елементите се познати како **ретротранспозони** и овие транспозони се пренесуваат преку RNA-интермедиер. Така на пр., LTR (long terminal repeats)-ретротранспозоните поседуваат долги терминални репетитивни секвенци кои кодираат GAG протеин (капсиден протеин вклучен во продукција на партикули слични на вируси), POL протеин (реверзна транскриптаза), како и ENV (протеин на обвивка), кои заедно се фузионираат како полипротеин. Потеклото на ретротранспозоните наликува на вирусите, бидејќи механизмот на транспозиција преку интермедиери се одвива на идентичен начин како при репликација на ретровирусот во геномот на клетката-домаќин (Сл. 2.11А).

Класа II елементите се **DNA-транспозони** кои се движат низ геномот преку DNA-интермедиер и често пати после нивната ексцизија оставаат изменета DNA-секвенца (footprint) на оригиналното место од каде што потекнуваат (Сл. 2.11Б).

Во однос на квантитетот или бројот на копии, ретротранспозоните значајно ги надминуваат DNA-транспозоните кај растителните геноми. Кај растенијата со големи геноми, како што се житните растенија и некои други видови (*Iris* или *Lilium*), ретротранспозоните можат да опфаќаат од 50% до 90% од геномот. Класичен пример за тоа како транспозоните и нивната експанзија учествуваат во големината на геномот се т.н. група на инсертирани или „вгнездени“ транспозони (nested transposons) во геномот од пченка (Сл. 2.12).

Геномското секвенционирање кај голем број примероци од пченка покажува значаен диверзитет во нивниот геном. Всушност, голем број примероци/сорти од пченка имаат иста генетска мапа (редослед на гени) со минимални разлики во рамките на генските секвенци, а варијациите се констатирани во секвенците помеѓу гените (Сл. 2.13А), како и во содржината на гени (број на копии на гени). Исто така, овој тип на варијација во геномот е забележан кај други растителни видови кои се карактеризираат со транспозонска активност

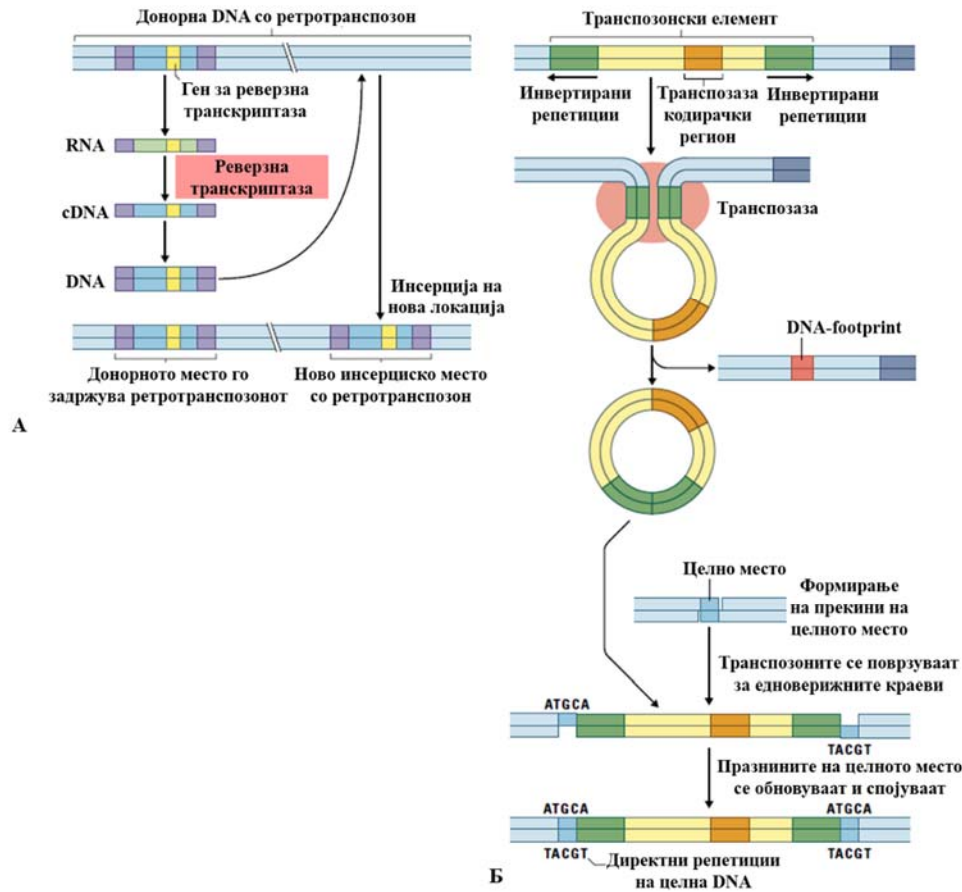
за време на последните неколку милиони години. Како резултат на оваа транспозонска активност, кај сите претставници од еден ист вид не била евидентирана иста содржина на генетски материјал. Овие податоци биле основа за воспоставување на концептите на основен геном (core genome) и пангеном (pan genome) кај одреден вид, кои претходно биле предложени за бактериските видови. Пангеномот го опфаќа вкупниот сет на гени кај сите претставници на еден вид. Всушност, пангеномот е претставен со основен геном кој се однесува на сите гени кои се заеднички за сите претставници на видот, како и со дополнителен геном кој се однесува на гените кои се присутни само кај некои претставници на видот (Сл. 2.13В).

Табела 2.1 Транспозонските елементи се класифицираат според механизмот на транспозиција и кодирачките генски секвенци (Buchanan и сор., 2015).

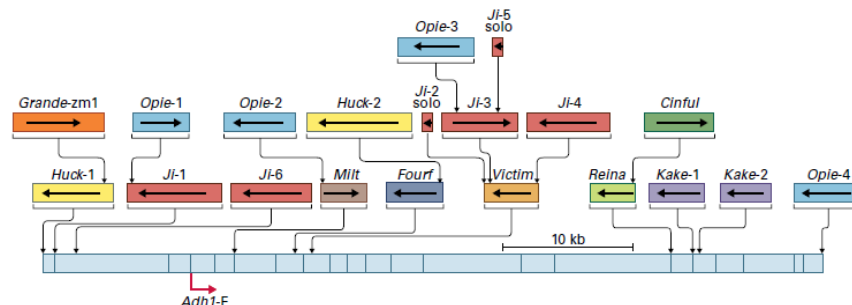
Класа	Структурни карактеристики на ниво на суперфамилија	TSD*
Класа I (ретротранспозони)		
LTR	<i>Copia</i> → GAG AP INT RT RH →	4-6
	<i>Retrovirus</i> → GAG AP RT RH INT ENV →	4-6
DIRS	<i>DIRS</i> ↔ GAG AP RT RH YR ↔	0
	<i>Ngaro</i> → GAG AP RT RH YR → >>>	0
PLE	<i>Penelope</i> ↔ RT EN ↔	Варијабилно
LINE	<i>R2</i> — RT EN —	Варијабилно
	<i>L1</i> — ORF1 — APE RT —	Варијабилно
SINE	<i>tRNA</i> — — —	Варијабилно
Класа II (транспозони)-субкласа 1		
TIR	<i>Tc1-Mariner</i> ↔ Tase* ↔	TA
	<i>CACTA</i> ↔ Tase* — ORF2 ↔	2-3
Crypton	<i>Crypton</i> — YR —	0
Класа II (транспозони)-субкласа 2		
Helitron	<i>Helitron</i> — RPA — // Y2 HEL —	0
Maverick	<i>Maverick</i> — C-INT — ATP — // CYP — POL B —	6

*TSD: Целни места на транспозиција.

Структурни особини	Протеин кодирачки домени
→ Долги терминални репетиции	AP Аспартат протеиназа
↔ Терминални инвертирани репетиции	APE Апурииска ендонуклеаза
— Кодирачки регион	ATP Пакувачка АТР-аза
— Некодирачки регион	C-INT C-интеграза
— Дијагностичка особина во некодирачки регион	CYP Цистеин протеаза
— Регион кој содржи една или повеќе дополнителни рамки на читање (ORF)	EN Ендонуклеаза
	ENV Протени на обвивка
	GAG Капсиден протени
	HEL Хеликаза
	INT Интеграза
	ORF Отворена рамка на читање со непозната функција
	POL B DNA-полимераза B
	RH RNаза H
	RPA Репликациски протени A (само во пластиди)
	RT Реверзна транскриптаза
	Tase Транспозаза (* со DDE мотив)
	YR Тирозин рекомбиназа
	Y2 YR со YY мотив

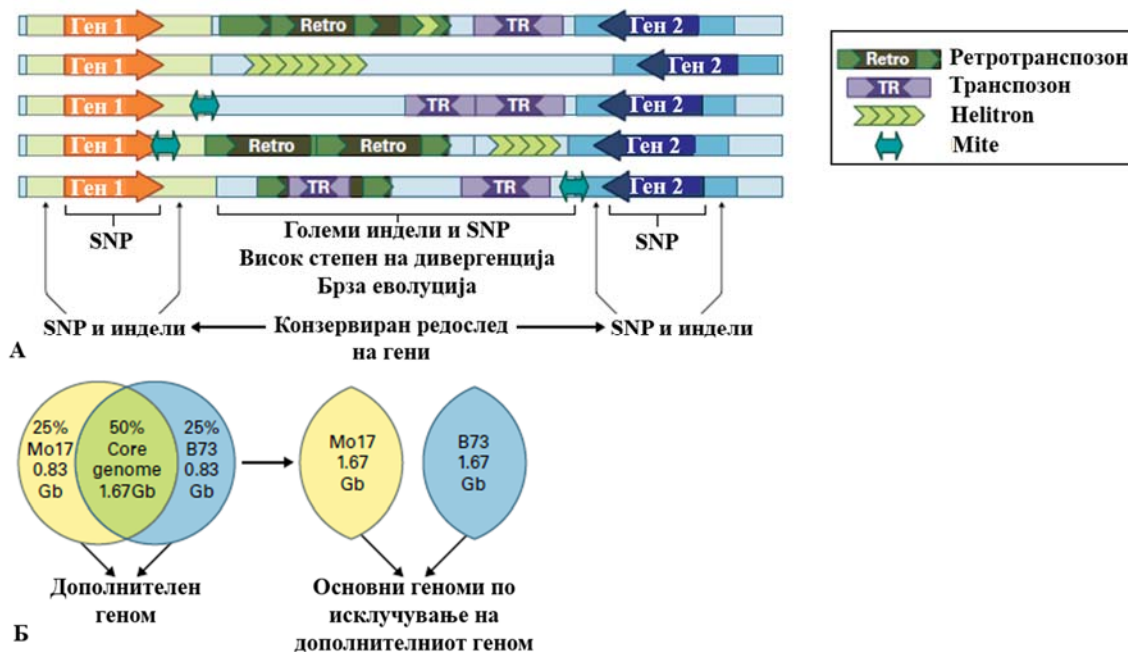


Слика 2.11 (А) Механизам на транспозиција на ретротранспозони. Елементот кој подлежи на транспозиција се копира (транскрибира) во интермедиерен RNA-молекул. Овој RNA-молекул преминува во cDNA со реверзна транскрипција, а потоа се инсертира во нова локација. (Б) Транспозиција на класа II транспозонски елементи. Овие транспозонски елементи се ограничени од двете страни со кратки инвертирани репетитивни секвенци. Според овој модел на транспозиција, ензимот транспозаза ги препознава овие секвенци и формира стебло/јамка структура. Потоа, оваа јамка се отсекува од тој регион на геномот. Се прави прекин на ново инсерциско место. Откако транспозонот ќе се инсертира, формираните едноверижни региони од пресекот на целното место се обновуваат, при што се добива дупликација на целното место (Buchanan и сор., 2015).



Слика 2.12 Повеќекратни настани на транспозиција во геномот на пченка. Репетитивни DNA-секвенци во регионот околу генот *Adh1* кај геномот од пченка. Во регионот околу генот *Adh1* од геномот на пченка доминираат ретротранспозони. Класите на секвенциските мотиви се именувани (пр. Opie за сино) и се поставени на мапата според позициите каде се јавуваат инсерциите (Buchanan и сор., 2015).

Голем број региони од геномот класифицирани како ретротранспозон-слични врз база на нивните секвенци не се сметаат како функционални ретротранспозони поради промените во секвенците кои ја прекинуваат нивната способност за кодирање на протеини. Некои од овие DNA секвенци можат сè уште да се движат под дејство на интактни ретротранспозони и нивните промотери можат да влијаат врз експресијата на нормални гени во близина на местото на нивна инсерција во геномот.

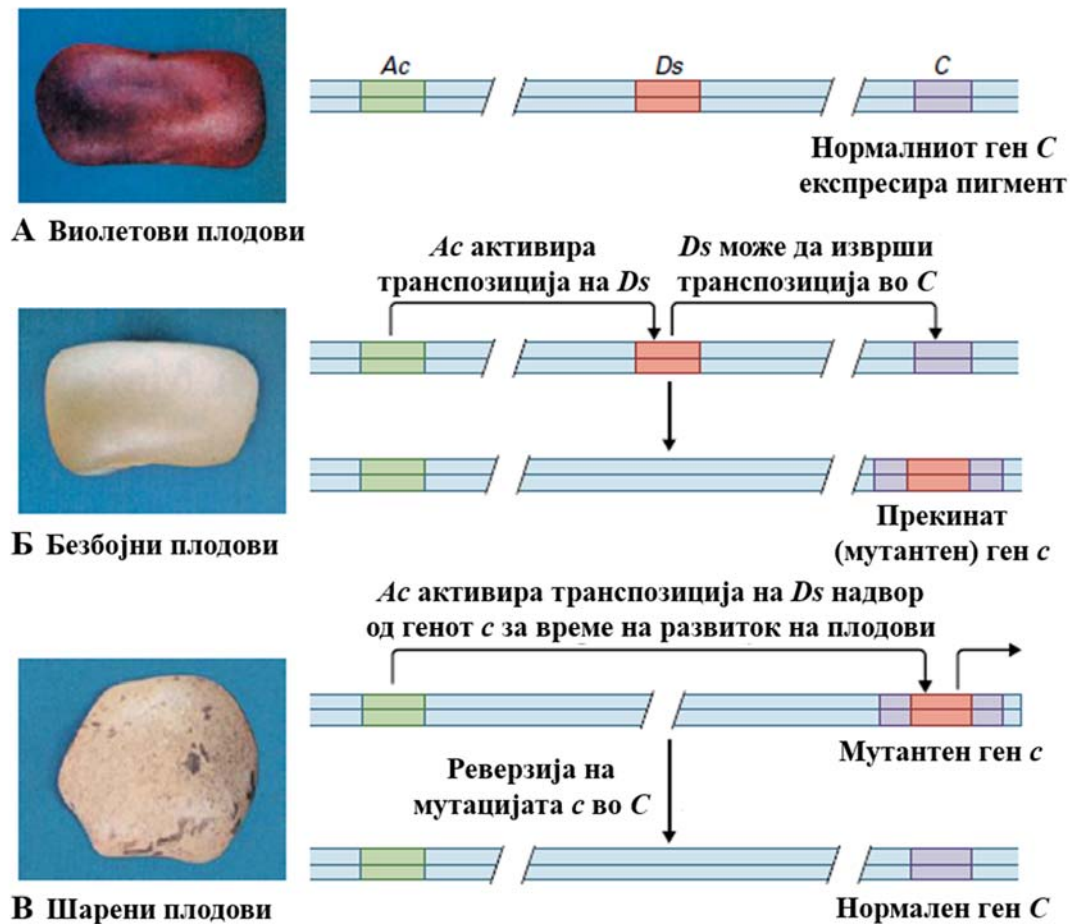


Слика 2.13 Диверзитет кај различни сорти на пченка и концептот на пангеном (Buchanan и сор., 2015). (A) Варијација на секвенци во краток сегмент од геномот на пченка. На сликата се прикажани пет различни алели од краток сегмент кој опфаќа два гена. (Б) Концепт на пангеном. Околу $\frac{2}{3}$ од геномот на пченката е заеднички кај сортите Mo17 и B73, додека пак, останатите $\frac{1}{3}$ од геномот е уникатен за секоја сорта. SNP: Еднуклеотидни полиморфизми; Индели: инсерции/делеции на бази.

Друга класа I транспозони се долги дисперзирани елементи (LINE-long interspersed elements) и кратки дисперзирани елементи (SINE-short interspersed elements). Вкупната должина на LINE изнесува од 4 до 9 kb и за нивната транспозиција е неопходна активност на ендонуклеаза, реверзна транскриптаза и врзување за нуклеински киселини. Од друга страна пак, SINE се мали (80-500 bp) неавтономни елементи кои го користат ензимскиот систем на LINE за транспозиција.

2.3.2 DNA транспозоните кај пченка се откриени како елементи кои се одговорни за прекини на хромозомите и варијација во бојата

Класа II елементите (Табела 2.1) биле за прв пат откриени како Ac/Ds транспозонски систем кај пченката (Сл. 2.14). Овој Ac/Ds-систем се состои од два типа на елементи: **автономен елемент** (Ac; активатор) и **неавтономен елемент** (Ds; дисоцијатор).



Слика 2.14 Ефекти на транспозицијата врз обојувањето на плодите од пченка (Buchanan и сор., 2015). Barbara McClintock била првиот научник која го открила постоењето на транспозонски елементи во геномот. Генетските анализи на фенотиповите од плодови на пченка покажале дека гените или сегменти од гените можат да „скокаат“ од една на друга позиција во геномот. На сликата е прикажана прогресијата на настани поврзани со транспозиција на гени: (А) Генот *C* кодира пигмент кај пченката кој е одговорен за виолетова боја на плодите, што значи дека кај виолетовите плодови е активен генот *C*. (Б) Елементот *Ac* кодира транспозаза која посредува во транспозиција на *Ds*-елементот. При инсерција на транспозонскиот елемент *Ds* во генот *C* настанува инактивација на генот, што предизвикува добивање на безбоен мутантен фенотип означен како *c*. (В) Ако настане транспозиција или „скокање“ на *Ds* надвор од мутантниот ген *c* за време на развојот на плодите, тогаш се обновува активноста на генот *C* и се забележуваат виолетови дамки низ безбојната површина на плодите.

Автономниот *Ac*-елемент претставува 4,6-kb DNA-секвенца која ги кодира сите неопходни производи за негова транспозиција. Централниот регион кодира 3,5-kb транскрипт на транспозаза која има способност за ексцизија на *Ac*-елементот и сите претставници од фамилијата *Ds*. Инвертираните репетитивни секвенци со големина од 11 bp од секој крај на *Ac*-елементот имаат улога во препознавање на транспозаза ензимот. При инсерција на *Ac*-елементот во нова позиција од геномот настанува додавање на директни репетитивни секвенци (8 bp) од геномот-домаќин со цел да се опфати овој елемент од двете страни. Регионот од овие директни репетитивни секвенци кои остануваат на истата позиција во геномот по извршената транспозиција на елементот се нарекува **footprint** на

транспозицијата (Сл. 2.11B). Краевите на инвертираните репетиции и формираното footprint-место на инсерциите/ексцизиите се карактеристични за фамилиите на DNA-транспозонските елементи, како што е *Ac/Ds*-системот. Сепак, големината и секвенците на овие региони претставуваат специфични карактеристики на секоја фамилија на DNA-транспозонските елементи.

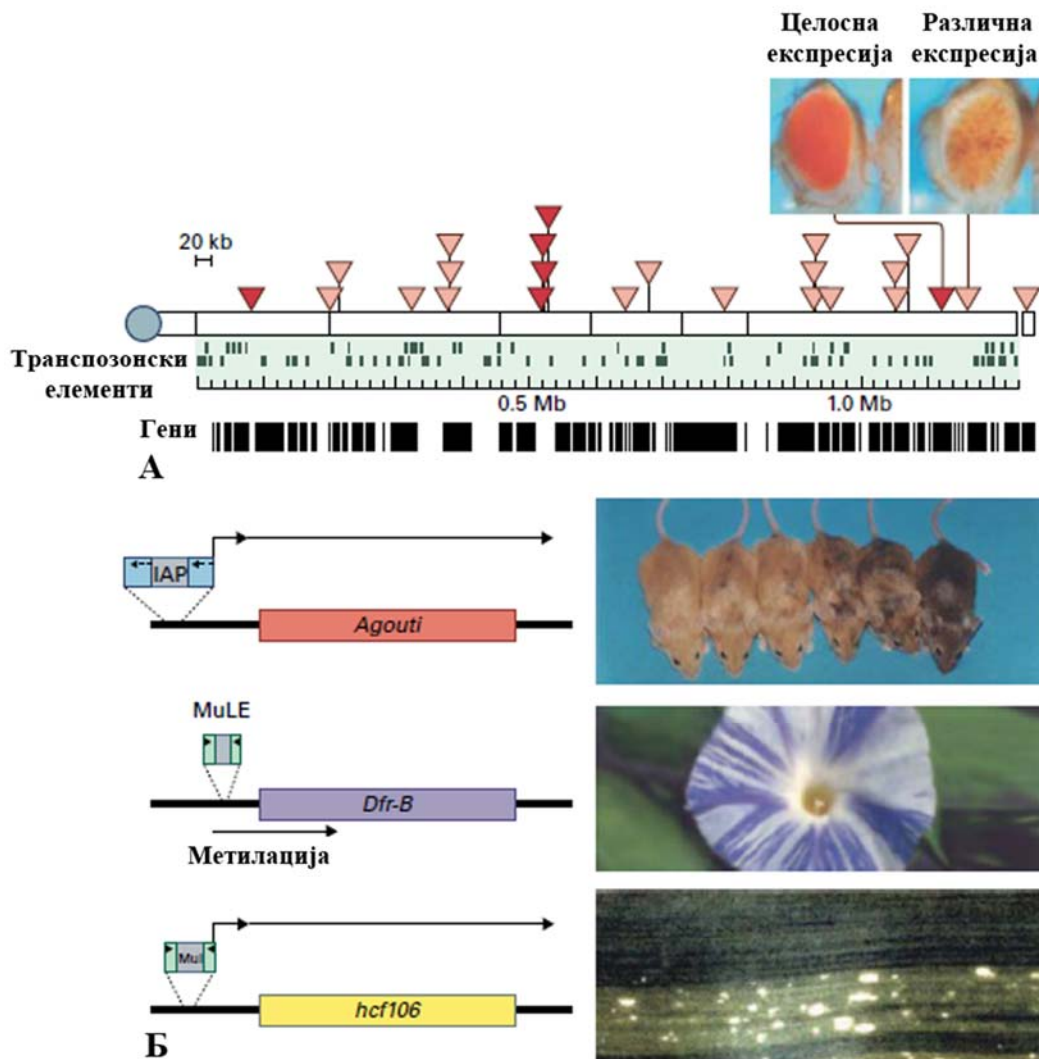
Неавтономните *Ds*-елементи претставуваат фамилија на сродни DNA-секвенци кои се препознаваат од ензимот транспозаза што се кодира од *Ac*-елементот. Овие неавтономни елементи не кодираат сопствена транспозаза и нивната транспозиција зависи од присуството на *Ac*-елементот и неговите генски производи. Кај растенијата се откриени седум различни суперфамилии на класа II транспозони и неколку фамилии од овие транспозони се потврдени кај сите растителни геноми. Минијатурните инвертирано-репетитивни транспозонски елементи (MITES-miniature inverted-repeat transposable elements) се кратки, неавтономни DNA-елементи кои се широко распространети кај растителните геноми кои се зависни од транспозази од други елементи за нивната транспозиција. Врз основа на структурните карактеристики и секвенците, повеќето MITE се поделени во две главни групи: *Tourist* и *Stowaway* транспозони.

2.3.3 Транспозонските елементи можат да функционираат кај различни растителни видови

Транспозонските елементи од еден растителен вид можат да бидат активни и кај други видови. Оваа технологија е позната како **транспозонско обележување** и се користи за нарушувања на функцијата на гените и добивање на фенотипски мутанти. Така на пр., транспозонот TnT од тутун бил интродуциран во геномот од *Medicago truncatula* преку *Agrobacterium*-посредувана генетска трансформација. Притоа, линиите на растенија означени со транспозон можат да подлежат на скрининг на фенотипот од интерес, а локацијата на случајно инсертираните места може да се идентификува преку секвенционирање. На таков начин, технологијата на транспозонско обележување овозможува идентификација на растителните линии кои ги содржат инсерциите во гените од интерес.

2.3.4 Транспозициите можат да предизвикаат мутации и промени во генската експресија

Повеќето растителни геноми содржат мал број на интактни и потенцијално активни транспозони, како и голем број на дефективни елементи. Според тоа, транспозицијата може да предизвика мутации. За разлика од класа I елементите, класа II елементите индуцираат нестабилни мутации. По извршената ексцизија на транспозонот, промените на целното место зависат од неговата локација во регулаторните или кодирачките региони на гените и дали настанува обновување на нарушената рамка на читање. Повеќето форми на разнобојност на цветовите или семињата кај пченка се јавуваат поради одвивање на процесите на транспозиција (Сл. 2.14). Исто така, транспозоните можат да влијаат врз експресијата на соседните гени преку континуирана транскрипција на гени преку стоп кодони (read-through transcription) и ширење на епигенетско стивнување (epigenetic silencing) од транспозонските елементи на соседните региони од геномот (Сл. 2.15).



Слика 2.15 Транспозонските елементи можат да имаат различни ефекти врз гените во близина. (А) Позиционен ефект на варијација кај *Drosophila melanogaster*. Експресијата на трансген кој е одговорен за црвена боја на очите зависи од неговата близина до транспозонските елементи. Црвените триаголници ги покажуваат инсерциските места на трансгенот кои се одговорни за црвена боја на очи кај мутанти со бела боја. Од друга страна пак, инсерцијата на трансгените на местата означени со розови триаголници предизвикува шаренолика боја на очи поради модулирање на нивната експресија од страна на транспозонските елементи во близина. (Б) Три примери на индуцирана варијација под дејство на инсерција на транспозон во близина на протеин-кодирачки ген. Кај стаорци, ретротранспозонот IAP продуцира транскрипт кој продолжува да се чита во генот *agouti* за боја на крзно. Содржината на *agouti*-транскриптот и бојата на крзно зависат од епигенетскиот статус на ретротранспозонот и овие особини можат да се наследуваат во потомството. Кај цветовите од *Ipomoea* sp., метилацијата на DNA кај неавтономен MuLE-транспозон може да се прошири до промоторот на генот за дихидрофлавонол редуктаза (*Dfr-B*), при што се добиваат обоени линии на венечните ливчиња. Кај пченката (*Zea mays*), активноста на една фамилија на транспозонски елементи регулира две епигени. Кај ткивата со активна Mutator транспозон фамилија се јавуваат мутантни фенотипови кај двете епигени, при што се формираат единични делови од ткивото со светло зелена боја (висока флуоресценција на хлорофил, *hcf106* мутант) и некротично шарено ткиво (*les28* мутант) на листовите. Иницијацијата на транскриптите *hcf106* настанува под дејство на неавтономен (Mu1) транспозон во услови кога не е активен автономниот транспозон во геномот (Buchanan и сор., 2015).

2.4 Генска експресија

Развитокот на растенијата и нивната способност да одговорат на промените во животната средина значајно зависат од експресијата на информацијата која се содржи во геномот. За овој процес на **генска експресија** е неопходна транскрипција на гени во RNA и транслација на овие RNA-молекули во генски производи, односно протеини. Во ова поглавје е опишан молекуларниот механизам на генска експресија и регулаторните механизми кои ја контролираат активноста на гените кај растенијата.

2.4.1 RNA-полимеразите имаат различни функции при транскрипција на јадрени гени

Ензимите **RNA-полимерази** синтетизираат различни типови на RNA-молекули во растителните клетки. Пет различни RNA-полимерази посредуваат во транскрипција на гените од јадрениот геном на растенијата, а дополнителни RNA-полимерази се одговорни за транскрипција на геномот од клеточните органели. Јадрените RNA-полимерази (означени од RNA I до RNA V) имаат различни функции и синтетизираат различни типови на RNA-молекули (**Табела 2.2**).

Табела 2.2 Јадрени RNA-полимерази кај растенијата.

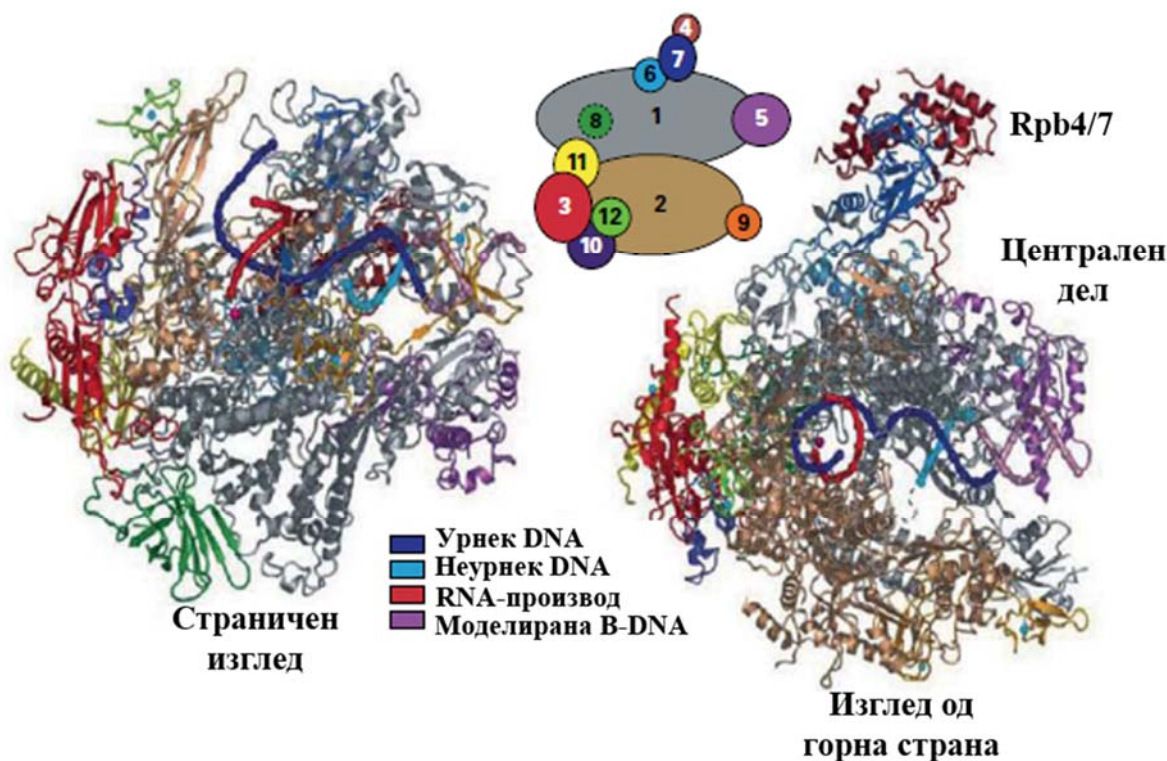
Тип на RNA-полимераза	Производи
RNA-полимераза I	5.8S, 18S и 28S rRNA
RNA-полимераза II	mRNA, miRNA, snRNA, ta-siRNA
RNA-полимераза III	5S rRNA, tRNA, snRNA
RNA-полимераза IV	siRNA
RNA-полимераза V	Некодирачки RNA-молекули

RNA-полимераза I има високо специфична улога при генската експресија бидејќи транскрибира само прекурсор на 45S рибозомална RNA (rRNA), кој подоцна се процесира во 5.8S, 18S и 28S rRNA-молекули. Транскрипцијата со посредство на RNA-полимераза I опфаќа до 50% од синтезата на RNA-молекули во активните клетки.

RNA-полимераза II е одговорна за транскрипција на протеин-кодирачките гени и синтетизира информациска RNA (mRNA). Исто така, оваа полимераза транскрибира микроRNA (miRNA) гени и синтетизира мали јадрени RNA-молекули (small nuclear RNA-snRNA), како и *trans*-дејствувачки мали интерферирачки RNA-молекули (*trans*-acting small interfering RNA-ta-siRNA). Поради есенцијалната улога во експресија на гени, RNA-полимераза II била предмет на детални проучувања (**Сл. 2.16**).

RNA-полимераза III транскрибира гени кои кодираат транспортна RNA (tRNA), 5S rRNA, како и различни други мали RNA-молекули.

RNA-полимераза IV и **V** биле неодамна откриени и за разлика од RNA-полимеразите I, II и III кои се присутни кај сите еукариоти, овие полимерази се специфични за растенијата (RNA-полимераза V е присутна само кај скриеносемените растенија). Исто така, овие RNA-полимерази не се есенцијални за вијабилноста на растенијата. Функцијата на RNA-полимераза IV е синтеза на мали интерферирачки RNA-молекули (small interfering RNA-siRNA), додека пак, RNA-полимераза V транскрибира некодирачки RNA-молекули кои се вклучени во siRNA-посредувано генско стивнување.

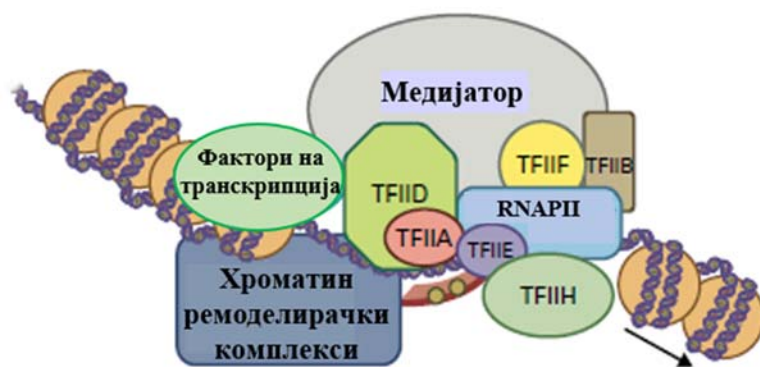


Слика 2.16 Кристална структура на елонгациониот комплекс од RNA-полимераза II кај *Saccharomyces cerevisiae*. Субединиците на RNA-полимераза II се обоени според легендата на сликата (Buchanan и сор., 2015).

Според структурните карактеристики, RNA-полимеразите се разликуваат во сензитивноста кон дејството на инхибиторите на транскрипција. Така на пр., RNA-полимераза II е високо сензитивна кон α -аманитин, кој претставува цикличен октапептид присутен кај родот *Amanita* и претставува главен извор на интоксикација на луѓето со печурки. Од друга страна пак, RNA-полимераза I не е сензитивна кон α -аманитин, додека пак, RNA-полимераза III има интермедиерна сензитивност кон дејството на овој инхибитор. Ензимот RNA-полимераза III може селективно да се инхибира под дејство на тагетитоксин, кој претставува бактериски фитотоксин. Имајќи ги предвид овие разлики во сензитивноста, α -аманитин и тагетитоксин како специфични инхибитори на RNA-полимерази можат да се користат при дистинкција на транскриптите кои се добиени со активноста на различни RNA-полимерази.

2.4.2 Транскрипциските фактори и RNA-полимеразите го формираат основниот систем на транскрипција

Процесот на транскрипција може да се подели во три последователни фази: (i) иницијација на транскрипција, (ii) елонгација која овозможува формирање на RNA-производ преку активност на RNA-полимераза и (iii) терминација на транскрипција.



Слика 2.17 Шематски приказ на преиницијациониот комплекс на RNA-полимераза II (RNAPII) и неговата интеракција со транскрипциски фактори (TF) и хроматин ремоделирачки комплекси на DNA. Светлокафените топкиња ги претставуваат нуклеозомите (Buchanan и сор., 2015).

Иницијацијата е важен чекор во контрола на транскрипцијата кој зависи од формирањето на **преиницијациониот комплекс на транскрипција** (Сл. 2.17). Овој сложен протеински комплекс е составен од RNA-полимераза и асоцирани протеини (Табела 2.3), кои ја модулираат активноста на RNA-полимераза и го активираат ензимот до генските промотори. Така на пр., ензимот RNA-полимераза II не може да се врзе за DNA бидејќи е неопходно присуство на неколку **транскрипциски фактори** (TFIIA, B, D, E, F и H). Притоа, TFIID е мултипротеински комплекс кој го содржи клучниот сет од протеини на преиницијациониот комплекс на транскрипција.

Табела 2.3 Компоненти на преиницијациониот комплекс на RNA-полимераза кај *Arabidopsis*.*

Фактори	Број на субединици	Функција
TFIIA	3	Стабилизација на интеракцијата помеѓу TBP и TATA
TFIIB	1	Селекција на иницијационо место на транскрипција
TFIID	TBP и 12-15 TAF	Препознавање на промоторни елементи
TFIIE	2	Формирање на иницијациониот комплекс
TFIIF	2	Активирање на TFIIE и TFIIH
TFIIH	10	Ослободување на RNAPII од промоторот и елонгација на транскрипција
RNAPII	12	Иницијација, елонгација и терминација на транскрипција
Медијатор	~34	Пренесување на информации од DNA-асоцирани транскрипциски фактори на RNAPII

*TBP: TATA-врзувачки протеин; TAF: TBP-асоциран фактор; RNAPII: RNA-полимераза II; TF: Транскрипциски фактори

Клучна компонента на TFIID е **TATA-врзувачкиот протеин (TBP)** кој препознава **TATA секвенца** (Сл. 2.17). Имено, TATA претставува мотив од DNA-секвенца која е локализирана во промоторите на голем број протеин-кодирани гени околу позиција -30 (30 нуклеотиди спротиводно од местото на иницијација на транскрипција што кореспондира на првиот нуклеотид од RNA). Всушност, TBP-посредуваното врзување на TFIID-комплексот за промоторот е проследено со активација на дополнителни транскрипциски фактори и RNA-полимераза II во преиницијациониот комплекс, при што настанува иницијација на процесот на транскрипција.

2.4.3 Експресијата на растителните гени се менува за време на развитокот на растенијата или под дејство на сигнали од животната средина

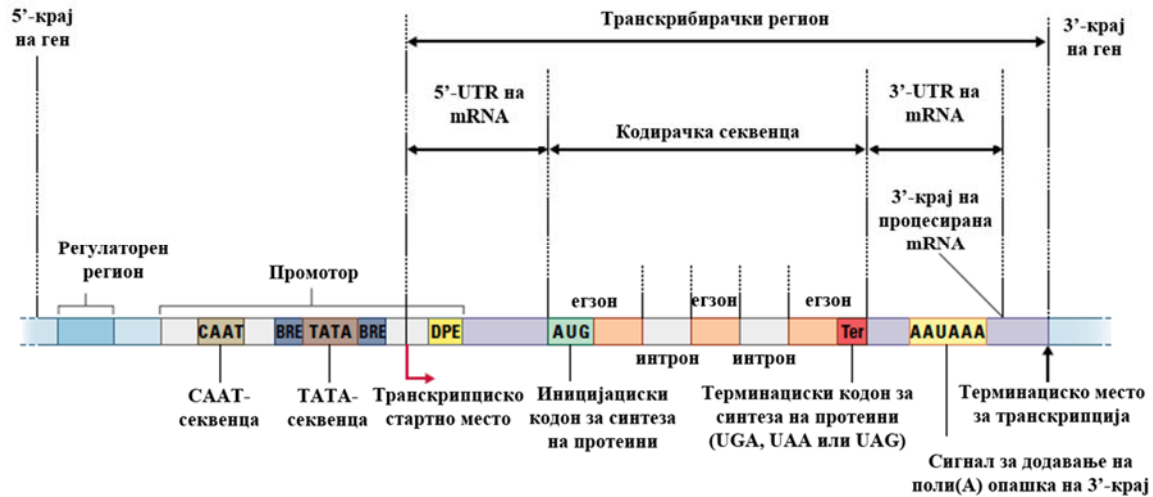
Некои растителни гени конститутивно се експресираат и транскрибираат во сите клетки на растителниот организам. Овие гени се нарекуваат „**домаќински гени**“ (**housekeeping genes**), бидејќи кодираат протеини кои се вклучени во основните клеточни функции, како што е метаболизмот на јаглехидрати или биосинтезата на протеини. Сепак, сите гени од јадрениот геном не се експресираат постојано, бидејќи голем број гени се транскрибираат во одреден стадиум од развитокот на растенијата или во специфични клетки или ткива. Исто така, експресијата на некои други гени зависи од сигналите или стрес факторите од животната средина. Терминот **диференцијална генска експресија** се однесува на промените во транскрипциската активност на гените за време на развитокот или како одговор кон сигналите од животната средина.

Значењето на диференцијалната генска експресија при растот и развитокот на растенијата може да се потврди преку проучување на механизмите на формирање на различни типови на клетки при процесот на **клеточна диференцијација**. Познато е дека преку процесот на клеточна диференцијација не настанува губење на генетската информација. Оваа појава за прв пат била потврдена кај растенијата преку изведување на класични експерименти за **тотипотентност** кај морков и тутун. Всушност, тотипотентноста претставува способност на ткивни експлантати или изолирани клетки од растение-донатор за регенерација во цело растение кое е идентично со донорните клетки. Денес, овие класични експерименти се изведуваат со изолирање на ембрионски стем-клетки од глушец, луѓе и други животни, како и репрограмирање на диференцирани адултни мамалиски клетки преку трансфекција со соодветен сет на транскрипциски регулатори. Според тоа, клеточната диференцијација за време на развитокот на еден организам претставува оркестрирана активација и репресија на соодветен сет од гени во различни органи и ткива. Активацијата на одредени гени при специфични стадиуми на развиток или во специфични растителни органи е строго контролиран процес кој се одвива под дејство на регулаторни региони од геномот кои се познати како генски промотори.

2.4.4 Промоторите се составени од регулаторни DNA-секвенци

Промоторот содржи различни елементи на DNA-секвенци кои активираат протеински фактори за регулирање на транскрипцијата на гени. Секоја класа на RNA-полимерази се врзува за соодветните промотори со цел да се иницира процесот на транскрипција. Промоторите се состојат од два типа на регулаторни елементи: **основни елементи** и **cis-елементи**. Основните елементи се неопходни за врзувањето на RNA-полимераза во позиција од 50 bp спротиводно или низводно од иницијацијското место на транскрипција. Овие елементи се релативно конзервирани помеѓу различни гени и заедно формираат **централен промотор (core promoter)**. Карактеристичен пример за основен елемент е ТАТА-секвенцата, а други примери за основни елементи во централните промотори се **TFIIБ-елементи на препознавање (TFIIБ recognition elements-BRE)** и **низводен промоторски елемент (downstream promoter element-DPE)**. Всушност, TFIIБ-елементите на препознавање (BRE) се локализирани спротиводно или низводно од ТАТА-секвенците (Сл. 2.18) и се препознаваат од страна на основниот транскрипциски фактор TFIIБ. Низводниот промоторски елемент (DPE) е локализиран околу 30 bp низводно од

иницијациското место на транскрипција и овој елемент е врзан за субединицата на TFIID-комплексот. Сите типови на основни елементи не се присутни во даден промотор, односно составот на основните промотори покажува значаен степен на варијабилност. Сепак, основните елементи кај растенијата не се целосно дефинирани поради ограничените податоци за секвенци кои ги одредуваат иницијациските места на транскрипција.



Слика 2.18 Структура и организација на гени кај еукариотски организми (Buchanan и сор., 2015). Генот е поделен на неколку функционални единици. Транскрибирачкиот регион претставува урнек за синтеза на mRNA, кој потоа се процесира и преведува во протеин како генски производ. Транскрибирачкиот регион е составен од некодирачки секвенци кои го делат регионот на кодирачки секвенци (егзони) и некодирачки секвенци (интрони). Транскрибирачкиот регион е ограничен од двете страни со некодирачки секвенци кои имаат улога во регулација на генската експресија. Повеќето елементи на регулаторните секвенци се наоѓаат на 5'-крајот од генот. Првите 1000 базни парови од 5'-крајот припаѓаат на генскиот промотор бидејќи содржи секвенци за иницијација на транскрипцијата. Во првите 50 базни парови од стартното транскрипциско место се наоѓа високо конзервиран регулаторен елемент кој е познат како ТАТА-секвенца. Всушност, ТАТА-секвенцата го координира активирањето на RNAPII за генската секвенца. UTR: Непреводлив регион.

За разлика од основните елементи кои се карактеризираат со одреден степен на конзервираност помеѓу различни гени, *cis*-елементите покажуваат значајна варијација помеѓу гените. Терминот „*Cis*“ (од латински што значи „од иста страна“) се однесува на фактот дека овие регулаторни елементи се локализирани на ист DNA-молекул како и гените што се под нивна контрола. *Cis*-елементите се одговорни за активација, репресија или модулација на генската експресија. За овие елементи се врзуваат **trans-дејствувачки фактори (trans-acting factors)**, кои претставуваат протеини што се кодираат на друго место („*trans*“) во геномот. *Trans*-дејствувачките фактори можат да стапат во директна интеракција со основниот транскрипциски систем (компонентите од TFIID-комплексот на транскрипциски фактори) или во индиректна интеракција со **медијатор**, кој претставува голем протеински комплекс кој дејствува како транскрипциски корегулатор (Сл. 2.17). На таков начин, *trans*-дејствувачките фактори влијаат врз степенот на одвивање на транскрипцијата под дејство на соодветни RNA-полимерази. Всушност, *trans*-дејствувачките фактори кои се врзуваат за *cis*-елементите најчесто се транскрипциски фактори и се нарекуваат **специфични транскрипциски фактори** за да се разликуваат од основните транскрипциски фактори.

Повеќето *cis*-елементи кај растенијата се наоѓаат во регион со големина од 1 до 2 kb, кој е локализиран спротиводно од иницијацијата место на транскрипција (Сл. 2.18). Сепак, овие елементи можат да се најдат на поодалечени растојанија спротиводно од промоторот или низводно од местото каде се одвива терминација на транскрипцијата. Често пати, *cis*-елементите се кратки DNA-секвенци (<10 bp) кои можат значајно да варираат во нуклеотидниот состав (Табела 2.4).

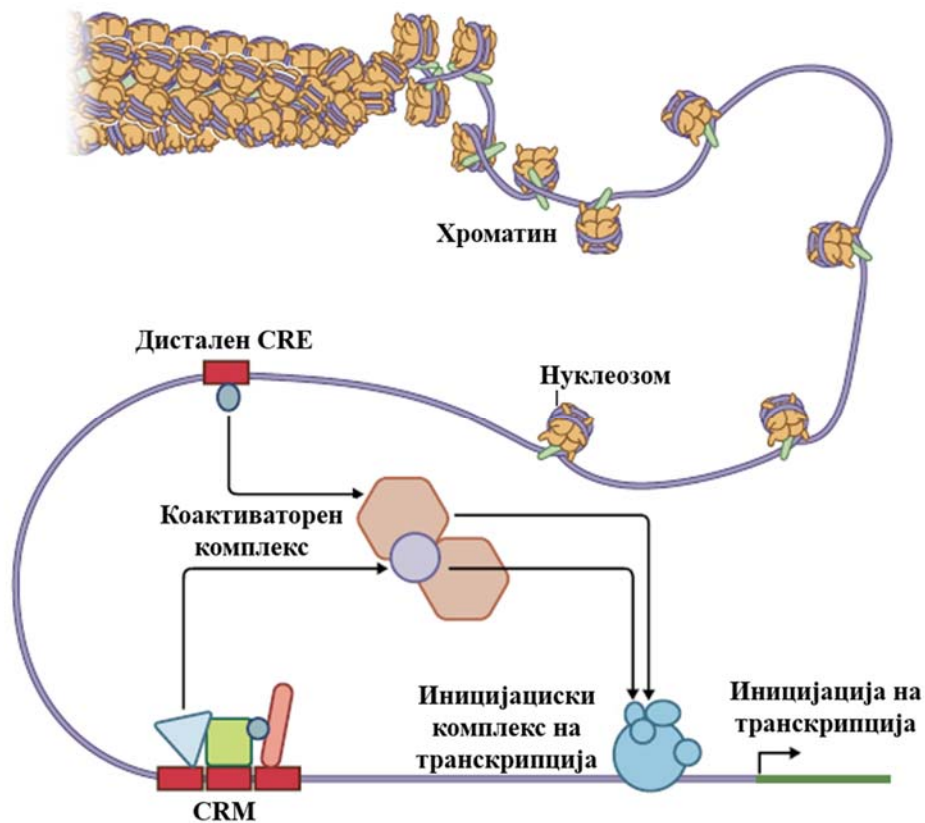
Табела 2.4 Примери на мотиви на DNA-секвенци за кои се врзуваат транскрипциски фактори од *Arabidopsis* и нивна функција во растот и развитокот на растенијата.*

Име на мотив	Секвенца	Врзувачки транскрипциски фактори	Функција
ABA-респонсивен елемент (ABRE)	CACGTGGC	ABRE-врзувачки фактори (Bzip протеини)	Одговор на ABA
Ауксин-респонсивен елемент (ARE)	TGTCTC	Фактори за одговор кон ауксини	Одговор на ауксини
CArG-секвенца	CC(A/T) ₆ GG	MADS-домен протеини	Развиток на цвет
Evening елемент	AAAATATCT	MYB протеини	Регулација преку циркадијален ритам
G-секвенца	CACGTC	Bzip и bHLH протеини	Одговор на светлина
GATA-промоторен мотив	(A/T)GATA(G/A)	GATA-транскрипциски фактори	Различни функции

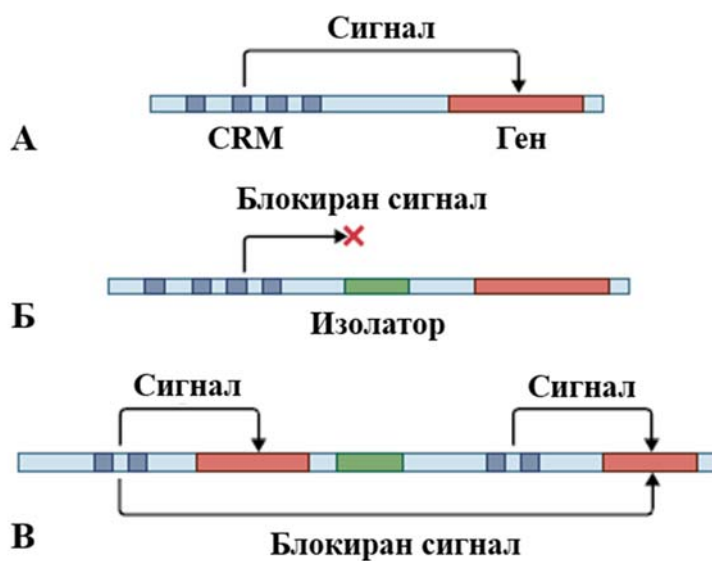
* ABA: Абсцицинска киселина; CArG: C-A/T-богат-G домен; bHLH: хеликс-јамка-хеликс мотив.

Промоторот може да содржи неколку различни *cis*-елементи, кои можат да бидат застапени со повеќе од една копија. Често пати, овие елементи меѓусебно се групираат во кластер и формираат ***cis*-регулаторни модули (*cis*-regulatory modules)**. Групирањето на овие модули во непосредна близина обезбедува кооперативно дејство на неколку (различни или идентични) транскрипциски фактори при регулација на генската експресија (Сл. 2.19). Во зависност од стимулативниот или инхибиторниот ефект врз генската експресија, *cis*-регулаторните модули се класифицираат како **засилувачи (*enhancers*)** или **стивнувачи (*silencers*)**. Засилувачите и стивнувачите претставуваат регулаторни елементи кои се оддалечени од транскрипцискиот регион на генот и се разликуваат од **проксималните промоторни елементи**, кои претставуваат регулаторни елементи локализирани во близина на иницијацијата место на транскрипција. Исто така, *cis*-регулаторните модули можат да функционираат како **изолатори (*insulators*)** за да ги блокираат сигналите од засилувачите или стивнувачите. На таков начин, изолаторите го спречуваат влијанието на засилувачите и стивнувачите врз експресијата на гените (Сл. 2.20).

cis-Елементите можат да посредуваат во експресијата на специфични гени за време на развитокот на растенијата или при одговорот кон различни внатрешни или надворешни сигнали. Таков пример е ауксин-респонсивниот елемент (auxin response element-ARE) со секвенца 5'-TGTCTC-3', кој се наоѓа во промоторите на гените што се контролираат од страна на фитохормонот ауксин. На таков начин, присуството на специфични *cis*-елементи во промоторот го одредува местото и времето на експресија на гените кај растенијата. Според тоа, капацитетот на клетките за диференцијална генска експресија е генетски предодредена од геномот во форма на ***cis*-регулаторен код**.



Слика 2.19 Пример за функцијата на *cis*-регулаторните модули и засилувачи. *cis*-Регулаторниот модул (CRM) во промоторот на генот се врзува со неколку транскрипциски фактори. Овој комплекс на транскрипциски фактори заедно со дистален *cis*-регулаторен елемент (CRE) како засилувач стапува во интеракција со коактиваторен комплекс за да се иницира транскрипцијата (Buchanan и сор., 2015).



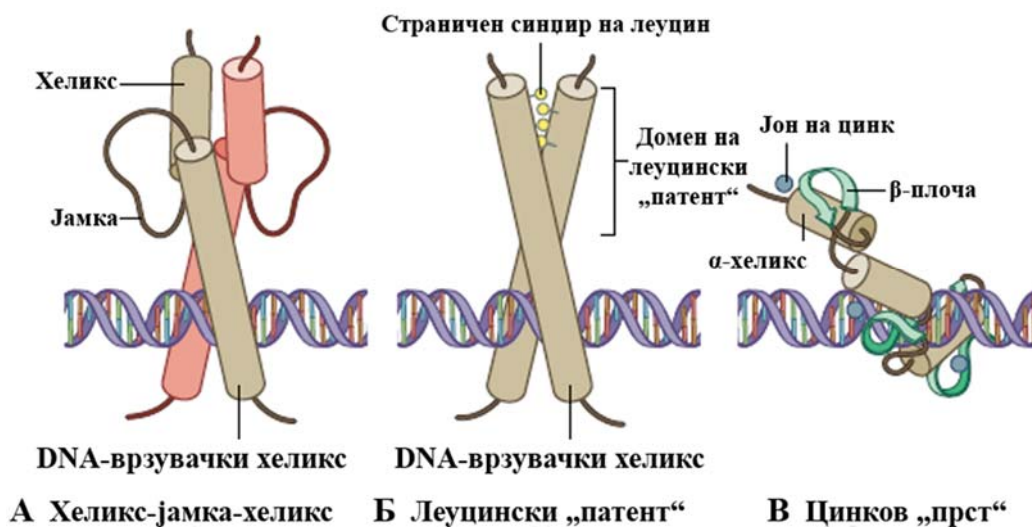
Слика 2.20 Функција на генетски изолатори (Buchanan и сор., 2015). Ефектот на *cis*-регулаторниот модул (CRM) врз генот (А) може да се блокира во присуство на изолятор (Б). Исто така, изоляторите можат да ги ограничат ефектите на CRM врз специфичен ген и на таков начин ја спречуваат активноста на CRM врз соседните гени (В).

2.4.5 Различни фамилии на транскрипциски фактори ја контролираат генската експресија

Познато е дека транскрипциските фактори се врзуваат за *cis*-елементите за да се стимулира или инхибира генската експресија преку интеракција со компонентите од транскрипцискиот систем. Притоа, транскрипциските фактори се врзуваат за специфична DNA-секвенца, што значи дека одреден транскрипциски фактор врши таргетирање на специфичен *cis*-елемент. Така на пр., претставниците од фамилијата на транскрипциски фактори наречени ауксин-респонсивни фактори (auxin response factors-ARF) се врзуваат за ауксин-респонсивниот елемент (ARE).

Транскрипциските фактори се составени од најмалку два домени со различни функции. Имено, **DNA-врзувачкиот домен (DNA-binding domain)** посредува во препознавањето и врзувањето за *cis*-елементот, додека пак, **трансактивиращкиот домен (transactivating domain)** стапува во интеракција со транскрипцискиот систем и ја одредува активноста на протеинот. Исто така, транскрипциските фактори содржат **секвенца за локализирање на јадрото (nuclear localization sequence-NLS)**, која им овозможува на синтетизираните протеини во цитоплазмата да навлезат во јадрото.

Постојат неколку различни типови на DNA-врзувачки домени кои се разликуваат според структурата и начинот на интеракција со DNA-молекулите (Сл. 2.21). Така на пр., транскрипциските фактори означени како **цинков „прст“ (zinc finger)** се карактеризираат со домен составен од две антипаралелни β -плочи и еден α -хеликс кој се инсертира во главното вдлабнување на двојноверижниот DNA-хеликс. Овој домен се формира преку поврзување на јонот на цинк со четири аминокиселини (две хистидински и две цистеински аминокиселински остатоци) и на таков начин се добива структура која наликува на прст.



Слика 2.21 Три главни типа на DNA-врзувачки домени (Buchanan и сор., 2015). (А) Хеликс-јамка-хеликс мотив. (Б) леуцински „патент“. (С) Zinc finger домен. Секој модел прикажува DNA-протеински комплекс.

Често пати, транскрипциските фактори содржат специфични структурни мотиви, како што е доменот на **базен леуцински „патент“ (basic leucine zipper-bZIP)**. Протеините со овој домен формираат димери, кои се врзуваат за големиот завој од DNA-молекулот. Секој bZIP-мономер содржи α -хеликс со леуцински остаток на секоја седма аминокиселина.

Леуцинските остатоци од двете α -хеликални структури меѓусебно се поврзуваат и на таков начин се овозможува „затворање“ („zip“) на α -хеликсите и формирање на централно преплетени спирали. Овие преплетени спирали ја одржуваат структурата на димерот и го позиционираат базниот регион во местото на главната вдлабнатина од DNA. За разлика од DNA-врзувачките домени, трансактивиращките домени кај транскрипциските фактори не се јасно дефинирани.

Транскрипциските фактори можат да се поделат на репресори и активатори во зависност од нивниот ефект врз генската експресија. Некои транскрипциски фактори имаат бифункционална функција на репресори и активатори. Која од овие две функции има доминантна улога во контролата на одреден ген може да зависи од присуството на дополнителни транскрипциски фактори на генскиот промотор. Исто така, посттранслациските модификации на бифункционалниот транскрипциски фактор можат да предизвикаат инактивација на еден од неговите трансактивиращки домени.

Транскрипциските фактори имаат значајна улога во регулацијата на растот и развитокот на растенијата. Во геномот на растенијата се идентификувани голем број гени кои кодираат транскрипциски фактори. Така на пр., геномот кај *Arabidopsis* содржи околу 2000 гени кои кодираат транскрипциски фактори. Овој број на гени кореспондира на 6% од сите генски секвенци кај *Arabidopsis*, што претставува значајно зголемена вредност во споредба со другите модел организми. Идентификуваните транскрипциски фактори кај вишите растенија можат да се групираат во 50 различни фамилии во зависност од сличностите во секвенците. Имињата на овие фамилии можат да се однесуваат на заедничките структурни особини на транскрипциските фактори. Така на пр., **базниот хеликс-јамка-хеликс протеини (basic helix-loop-helix proteins-bHLH)** содржат DNA-врзувачки домен кој е составен од две α -хеликални структури поврзани со интермедиерна јамка (Сл. 2.21). Исто така, транскрипциските фактори се именуваат според нивната физиолошка функција кај растенијата, како што е примерот со ауксин-респонсивниот фактор (ARF). Некои транскрипциски фактори се именуваат според основниот претставник на фамилијата. Така на пр., **AP2-слични транскрипциски фактори (AP2-like transcription factors)** ја претставуваат најголемата фамилија на транскрипциски фактори кај растенијата. Овие транскрипциски фактори го добиле името според APETALA2 (AP2) протеинот, кој има клучна регулаторна улога во развитокот на цветови кај растенијата.

Голем број фамилии на транскрипциски фактори кај растенијата (Табела 9.5) можат да се присутни и кај другите еукариоти, како што се bZIP-, MADS- и MYB-протеините.

Табела 9.5 Вкупен број и специфичност на фамилии на транскрипциски фактори кај *Arabidopsis* и ориз.

Фамилија	<i>Arabidopsis</i>	Ориз (<i>Oryza sativa</i>)	Специфичност за растенија
AP2/ERF	147	161	Не*
bZIP	74	94	Не
GRAS	32	57	Да
MADS	108	77	Не
MYB	198	182	Не
NAC	100	149	Да
WRKY	71	109	Не*

* Доминантно се идентификувани кај растенија.

Бројот на претставниците во фамилиите на транскрипциски фактори кај растенијата може да биде значајно зголемен или намален во споредба со животните или фунгите. Така на пр., фамилијата на **транскрипциски фактори со MADS домен (MADS domain transcription factors)** опфаќа повеќе од 100 претставници кај *Arabidopsis*, а само два претставника од оваа фамилија се потврдени кај *Drosophila*. Од друга страна пак, геномот кај *Drosophila* кодира значајно зголемен број на транскрипциски фактори со цинк „прст“-домени во споредба со *Arabidopsis*. Покрај еволутивно конзервираните фамилии на транскрипциски фактори, растенијата содржат голем број транскрипциски фактори кои не се идентификувани кај другите царства на организми, како што е ауксин-респонсивниот фактор (ARF).

Големиот број на транскрипциски фактори кај растенијата укажува на висок степен на комплексност во регулацијата на генската експресија. Оваа комплексност во регулацијата на генската експресија дополнително се зголемува поради фактот дека голем број транскрипциски фактори дејствуваат како димери (bZIP-протеини). Покрај **хомодимери** (парови на идентични протеини), голем број на транскрипциски фактори формираат **хетеродимери** (парови на различни транскрипциски регулатори). Единичен транскрипциски фактор може да формира хетеродимери со голем број на различни протеини, што овозможува значајно зголемување на бројот на можни комплекси на транскрипциски фактори. Исто така, активноста на транскрипциските фактори може да се регулира преку голем број посттранслационски модификации, кои се јавуваат како одговор на активација на специфични патишта на сигнална трансдукција. Така на пр., фосфорилацијата на транскрипциски фактор под дејство на специфична протеинска киназа може да ја измени DNA-врзувачката активност или да се препознае како сигнал за негова деградација. Активноста на другите транскрипциски фактори може да се контролира преку промените во нивната интрацелуларна локализација. Така на пр., задржувањето на транскрипцискиот фактор во цитоплазмата резултира со негова неактивност, додека пак, индукцијата на импортоот во јадрото предизвикува негова активација.

Клучен проблем во регулација на генската експресија е идентификацијата на гените кои се под контрола на транскрипциски фактор, како и *cis*-елементите за кои се врзува овој фактор. Во таа насока се развиени бројни експериментални методи за да се лоцираат врзувачките места на транскрипциските фактори во геномот. Особено софистицирана и ефикасна техника за идентификација на **целни гени** на транскрипциски фактор е методот на **хроматинска имунопреципитација (ChIP)**. Според овие анализи, транскрипциските фактори кај *Arabidopsis* можат да се врзуваат за стотина до илјадници региони по должината на геномот. Сепак, сите процеси на врзување на транскрипциските фактори не предизвикуваат промени во генската експресија, бидејќи овие места се локализирани надвор од функционалните *cis*-регулаторни модули.

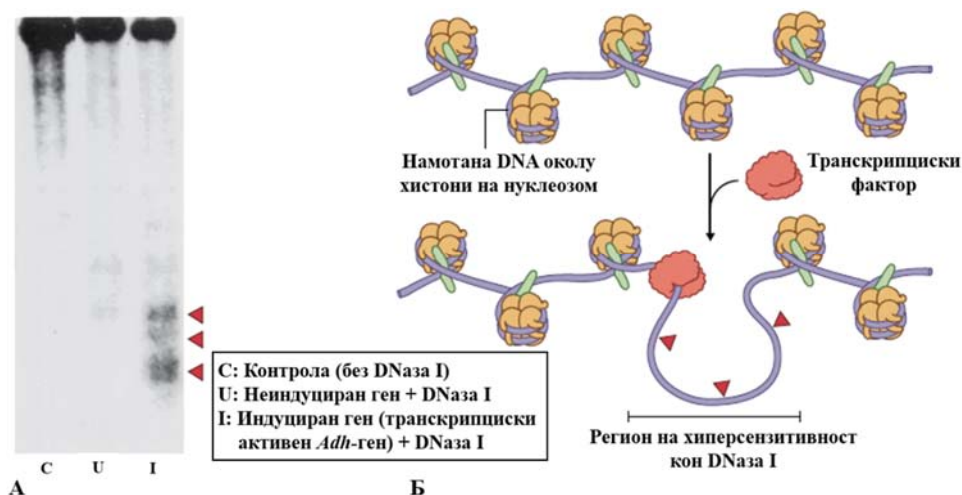
2.5 Хроматин и епигенетска регулација на генската експресија

2.5.1 Јадрената DNA е спакувана во хроматин како динамична структура

Јадрените геноми кај растенијата имаат големи димензии и се карактеризираат со висока содржина на DNA. Сепак, јадрената DNA се кондензира и се редуцира нејзината должина со цел ефикасно да се вклопи во јадрото кое има дијаметар од неколку микрометри. Оваа ефикасна стратегија на пакување на јадрената DNA се овозможува преку формирање на **хроматин** како резултат на асоцијација на DNA-молекулите со различни типови на

протеини. Најважни и најзастапени DNA-асоцирани протеини се хистоните. Хистоните се релативно мали протеини кои се карактеризираат со висок позитивен полнеж поради високата содржина на базни аминокиселини. На таков начин, позитивниот полнеж на хистоните има значајна улога во одржувањето на цврста интеракција помеѓу овие протеини и DNA-молекулите со негативен полнеж. Всушност, главната функција на хистоните е формирањето на структурата на нуклеозоми (Сл. 2.1).

Високиот степен на кондензација на јадрената DNA во хроматинот значајно ја ограничува генската експресија, бидејќи транскрипциските единици и промоторните секвенци не се достапни за RNA-полимеразите и транскрипциските фактори. Студиите за корелација помеѓу генската експресија и структурата на хроматинот покажуваат дека активните гени се повеќе подложни на дигестија од страна на DNаза во споредба со неактивните гени (Сл. 2.22). Всушност, хроматинот кој ги опкружува транскрипциски активните гени е помалку кондензиран во споредба целокупниот геномски хроматин и поради тоа овие гени се повеќе достапни за дигестија со нуклеази. Според тоа, DNA-регионите кои се хиперсензитивни кон дејството на DNаза I се ослободени од структурата на нуклеозомот, што овозможува лесна достапност на транскрипциските фактори до промоторните *cis*-елементи. Овие региони со низок степен на кондензација на хроматинот се нарекуваат **еухроматин**, додека пак, високо кондензираните региони се означуваат како **хетерохроматин**.

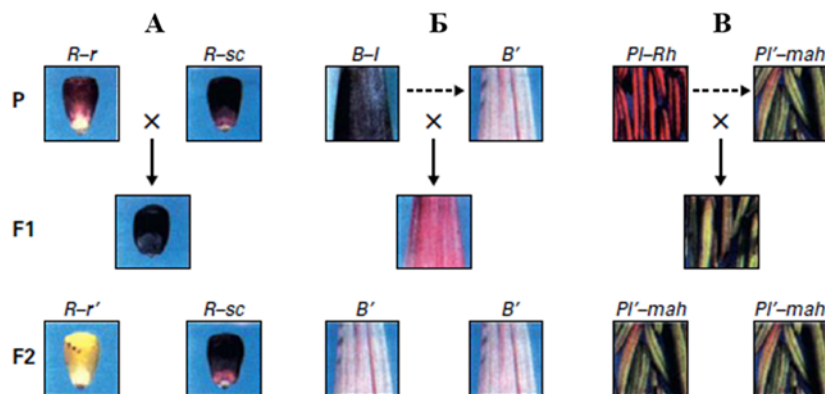


Слика 2.22 Хиперсензитивни места кон DNаза I кај промоторот на генот за алкохол дехидрогеназа 1 (ADH1) кај пченка. При транскрипциска активација на генот, регионот на промоторот станува помалку кондензиран со цел да се врзат транскрипциските фактори. Оваа намалена кондензираност на хроматинот околу промоторот може да се визуелизира како зголемување на достапноста за дигестија со DNаза I. Ензимот DNаза I претставува ендонуклеаза со низок степен на специфичност кон DNA-секвенците. Регионот на DNA со кондензиран хроматин не е достапен за дејството на DNаза I, меѓутоа, намалувањето на кондензацијата на хроматинот на генскиот промотор (се овозможува пристап до транскрипцискиот систем) овозможува достапност за DNаза I. Според тоа, хиперсензитивноста на DNA-регионите кон дејството на DNаза I претставува индикатор за хроматин со низок степен на кондензација, што претставува главен предуслов за транскрипциска активација на одреден ген. (А) Генскиот промотор на ADH1 е помалку достапен за DNаза I кај клетки со „исклучен ген“ (Неиндуциран ген: U) во споредба со клетки кои поседуваат транскрипциски активен ген (индуциран ген: I). Транскрипциски активните промотори кај индуцираните клетки покажуваат повеќе региони на хиперсензитивност кон DNаза I (црвени триаголници) во споредба со клетките каде не се врши транскрипција на генот (контрола: C). (Б) Хиперсензитивниот регион кон DNаза I не поседува структура на нуклеозом, при што се овозможува врзување на еден или повеќе транскрипциски фактори.

Постоењето на еухроматински и хетерохроматински DNA-региони покажува дека хроматинот е динамична структура која подлежи на значајни структурни модификации. Структурата на хроматинот може да се модифицира преку активноста на две различни класи на протеини: **хистонски хаперони** и **АТР-зависни ензими за ремоделирање на хроматин**. Во класата на хистонски хаперони се вбројува факторот за асемблирање на хроматин (chromatin assembly factor 1-CAF-1), кој учествува во асемблирање на хистоните за време на репликацијата и ремоделирањето на DNA. Исто така, овие протеини го спречуваат погрешното асемблирање и агрегацијата на хистоните. Ензимите за хроматинското ремоделирање ја изменуваат позицијата на нуклеозомите во хроматинот со цел да ја редуцираат или зголемат достапноста на другите протеини до DNA-молекулите. Овие ензими содржат DNA-зависна АТРаза која претставува систем за снабдување на клетката со енергија. Класични примери на протеини за хроматинско ремоделирање се производите на гените за промена на начинот на размножување (SWI) и способноста за ферментирање на сахароза (SNF) кај квасците, кои меѓусебно дејствуваат во ист комплекс. Кај растителните клетки се идентификувани хомолози на гените SWI/SNF, кои се вклучени во регулација на различни процеси на развојот кај растенијата.

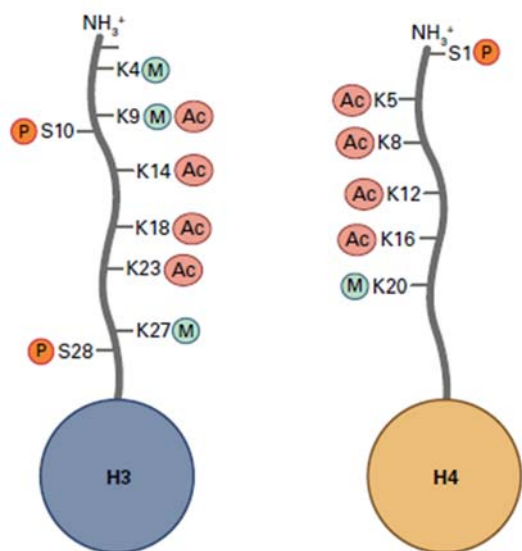
2.5.2 Генската експресија може да се контролира преку епигенетски механизми

Зависноста помеѓу генотипот и фенотипот на еден организам покажува дека сите наследни особини кои го одредуваат моделот и времето на генската експресија се кодирани во примарната DNA-секвенца. Во последните неколку децении се евидентирани наследни промени кои не вклучуваат промени во примарната DNA-секвенца на организмите. Овие промени се познати како **епигенетски промени**.



Слика 2.23 Парамутацијата опишува интеракција помеѓу адели, каде што експресијата на една адела кај хетерозигот се изменува со присуство на друга адела. Парамутација била испитувана кај генски локуси кои ја контролираат пигментацијата кај *Zea mays*. На сликата се прикажани примери на парамутација кои влијаат врз обојување на семе (*r*-експресија), семенова обвивка (*b*-експресија) и антера (*pl*-експресија). Фенотиповите на родителските (P) линии, F1 и F2-генераиите се прикажани во последователни редици. (A) Активноста на парамутираната адела *R-r* се редуцира кај потомството по експозиција на парамутагена адела *R-sc* кај F1-хетерозиготот. Аделите *R-r* и *R-sc* се структурно различни помеѓу себе и аделата *R-sc* има висок степен на експресија. (Б) Ниската експресија кај парамутагената *B'*-состојба може спонтано да потекнува од високо експресираните парамутирани адели *B-I* (испрекината стрелка). Аделите *B-I* можат да се сменат во *B'* кога се експонираат на *B'* кај F1-хетерозиготот. (В) Ниската експресија кај парамутагената *Pl'-mah* (*Pl'*) состојба може да потекнува од високо експресираните парамутирани *Pl-Rh* (*Pl*)-алели (испрекината стрелка). Аделите *Pl-Rh* се менуваат во *Pl'-mah* кога се експонираат на *Pl'-mah* кај F1-хетерозиготот (Buchanan и сор., 2015).

Карактеристични примери на епигенетски феномени се **парамутација** (Сл. 2.23) и **косупресија**. Сепак, епигенетските механизми имаат клучна улога во голем број клеточни и развојни процеси, како и во регулацијата на генската експресија. Епигенетските механизми се базираат на различни типови на ковалентни модификации на хроматинот, кои се означени како **епигенетски промени**. Овие промени вклучуваат метилација на DNA и различни посттранслациски модификации на хистоните кои најчесто се јавуваат на аминокиселинскиот крај на овие протеини (Сл. 2.24) и карбоксилниот крај на овие протеини.



Слика 2.24 Ковалентни модификации на хистоните H3 и H4 (Buchanan и сор., 2015). Кратенките „M“, „Ac“ и „P“ означуваат метилација, ацетилација и фосфорилирање, соодветно. На сликата се прикажани модифицираните аминокиселински остатоци (означени со кратенки со една буква) и нивните позиции во однос на аминокиселинскиот крај на протеините.

Метилацијата на DNA опфаќа метилација на цитозинските бази во јадрените DNA-молекули под дејство на ензимите **DNA-метилтрансферази** (Сл. 2.25). Метилацијата на цитозинските бази кај животните главно се одвива кај CG-динуклеотидите. Кај растенијата, целни места на дејствување на DNA-метилтрансферазите се цитозините во секвенците CHG и CHH (H = C, T или A). Метилацијата на DNA претставува ефикасен модел за проучување на епигенетските промени, бидејќи постојат механизми кои обезбедуваат митотско и мејотско наследување на специфични модели на DNA-метилација. Новореплицираните DNA-молекули кои содржат хемиметилирани места претставуваат специфичен супстрат за активноста на метилтрансфераза, при што се обезбедува одржување на постоечкиот модел на метилација во двете ќеркини хромозоми.

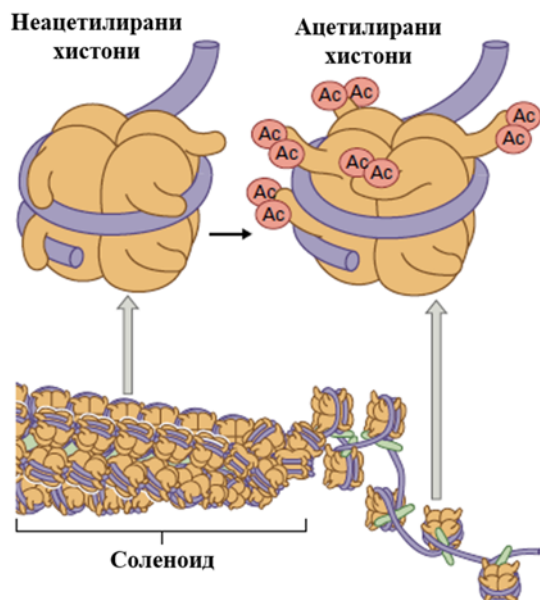


Слика 2.25 Метилација на цитозин. Ензимот метилтрансфераза врши трансфер на метилна група од S-аденозил метионин (SAM) на цитозин, при што се формира 5-метилцитозин (Buchanan и сор., 2015).

Кај растенијата се карактеризирани три различни патишта на DNA-метилација. METHYLTRANSFERASE 1 (MET1) претставува хомолог на мамалискиот ензим DNMT1 и главно одржува метилација на CG-места, додека пак, CHROMOMETHYLASE3 (CMT3) е специфичен ензим кај растенијата кој одржува метилација на CHG-секвенци. Третиот пат на DNA-метилација ги вклучува ензимите DOMAINS REARRANGED METHYLASES 1 и 2 (DRM1/2), кои одржуваат метилација на цитозин во сите контексти на секвенци.

Метилацијата на цитозин во DNA-молекулите се проучува со примена на **бисулфитно секвенционирање**. Оваа техника опфаќа третман на геномска DNA со натриум бисулфит, кој врши конвертирање на цитозин во урацил (метираните цитозини не се конвертираат). Секвенционирањето на третираната DNA овозможува дистинкција помеѓу метилираните и неметилираните цитозини. Со примена на оваа техника е потврдено дека 20% од геномот на *Arabidopsis* е во метилирана состојба и степенот на DNA-метилација зависи од големината на геномот и содржината на репетитивни секвенци. Најголема фракција на метилирана DNA е констатирана во рамките на транспозони и други репетитивни секвенци. Метилацијата на DNA има важна улога во стивнување на активноста на транспозоните, бидејќи мутанти кои не поседуваат активност на метилација се карактеризираат со транскрипциска активација на транспозони. Ефектот на DNA-метилација врз генската експресија зависи од локацијата на метилираните цитозини. Притоа, метилацијата на промоторот се поврзува со намалена транскрипција, додека пак, метилацијата во кодирачкиот регион на генот е во корелација со висок степен на експресија.

Посттранслационската модификација на хистоните опфаќа метилација, ацетилација и фосфорилација (Сл. 2.24). **Ацетилацијата на хистоните** се одвива под дејство на ензимите **хистон ацетил трансферази (НАТ)**, кои вршат трансфер на ацетилна група од ацетил CoA на ϵ -амино групата од различни остатоци на лизин на хистоните H3 и H4. Ацетилираните хистони ја регулираат транскрипцијата преку интеракција со различни транскрипциски фактори и хроматин ремоделирачки комплекси кои содржат **бромодомен** (мотив од 110 аминокиселини). Ацетилацијата на хистоните активира експресија на гени преку редуцирање на кондензацијата на хроматинот (Сл. 2.26). Сепак, ацетилацијата на хистоните е реверзибилен процес кој се одвива под дејство на ензимите **хистон деацетилази (HDAC)**.

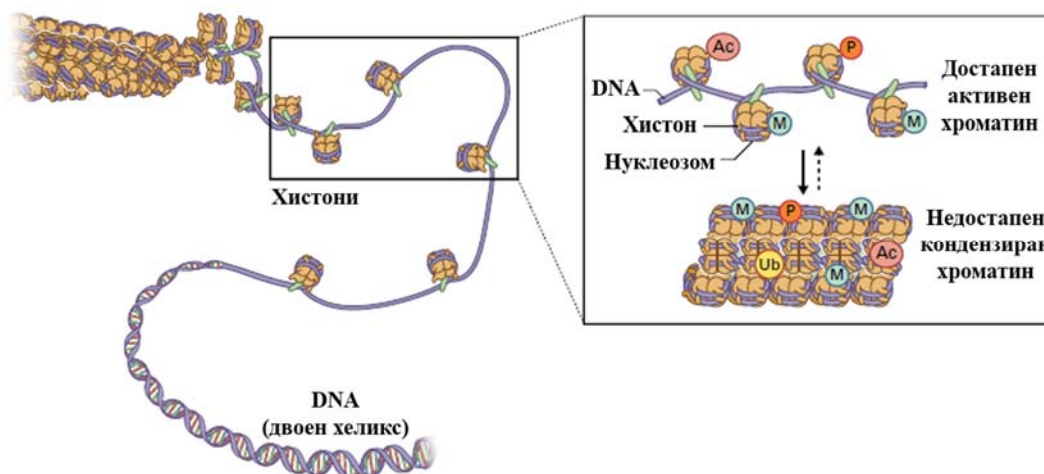


Слика 2.26 Хистоните можат да се модифицираат под дејство на ензимите ацетилтрансферази и деацетилази, кои влијаат врз степенот на кондензација на хроматинот. Хистоните од нуклеозомот се модифицираат на аминокрајот под дејство на ацетилтрансферази. Ацетилната група на хистонот формира стерична бариера за околните структури и на таков начин се формира локален регион на декондензација. Декондензацијата на хроматинот може да ја олесни достапноста на транскрипциските фактори за околната секвенца. Ензимите хистон деацетилази вршат отстранување на ацетилните групи и повторно се формира цврста интеракција помеѓу аминокрајот на хистоните и DNA-молекулите (Buchanan и сораб., 2015).

Метилацијата на хистоните доминантно се одвива на лизинските остатоци од хистоните H3 и H4 под дејство на ензимите **хистон метил трансферзи (HMT)**, кои користат *S*-аденозил метионин како донор на метилни групи. За разлика од ацетилација на хистоните каде се врши адиција на единични ацетилни групи, метилацијата на хистоните може да се одвива со адиција на моно-, ди- или три метилни групи. Ефектите на метилација на хистоните врз генската експресија зависат од аминокиселинските остатоци во хистонските продолжетоци. Во тој контекст, метилацијата на одредени аминокиселински остатоци ја зголемуваат експресијата на гените, додека пак, метилацијата на други аминокиселини има спротивен ефект. Така на пр., три-метилацијата на лизин 27 од хистонот H3 под дејство на специфичен протеински комплекс (**polycomb repressive complex 2-PRC2**) предизвикува репресија на гените. Од друга страна пак, три-метилацијата на лизин 4 од хистонот H3 преку **trithorax група на протеини (trxG)** се спротивставува на активноста на PRC2 и предизвикува активирање на гените. Исто така, степенот на метилација на аминокиселинските остатоци од хистоните има значајно влијание врз генската експресија. Така на пр., ди-метилацијата на лизин 9 од хистонот H3 има репресивен ефект, додека пак, три-метилацијата на истата аминокиселина предизвикува активна транскрипција на гени.

2.5.3 Епигенетските механизми се базираат на интеракција со транскрипциските фактори

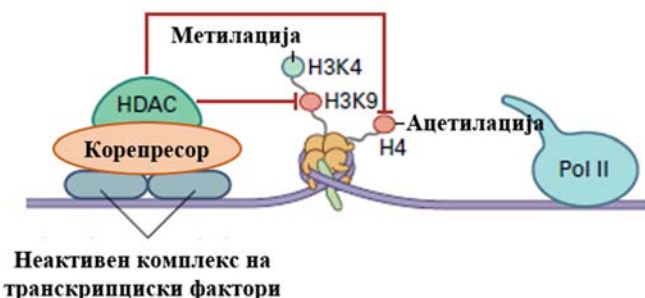
Познато е дека различни епигенетски промени можат да имаат спротивни ефекти врз транскрипцијата на гените. Исто така, различни епигенетски промени се одвиваат со посредство на специјализирани протеини, кои можат да ја активираат или инхибираат експресијата на гените преку модифицирање на хроматинската структура. Овој процес на епигенетски промени е основа за поставување на **хипотезата на хистонски код** или **епигенетски код**. Постооењето на голем број епигенетски промени овозможува добивање на неограничен број на комбинации во геномот. Сепак, одредени модели на епигенетски промени кои преовладуваат во геномот се познати како **епигенетски состојби** (Сл. 2.27).



Слика 2.27 Состојба на хроматин. Специфични епигенетски промени во геномскиот регион можат да предизвикаат кондензација на хроматинот или релаксирана конфигурација на хроматин. Овие промени во структурата на хроматинот ја спречуваат или олеснуваат транскрипцијата на гените (Buchanan и сор., 2015).

Постојат голем број на податоци за поврзаноста на епигенетските механизми со активноста на транскрипциските фактори. Истражувањата покажуваат дека ензимите кои учествуваат во индукција или бришење на епигенетските промени (пр. ацетилази и деацетилази) ги користат информациите од транскрипциските фактори за модификација на хроматинската структура. На молекуларно ниво е потврдено дека ензимите за модификација на хроматинот вршат таргетирање на генските секвенци преку директна или индиректна интеракција со транскрипциските фактори. Карактеристичен пример за индиректен механизам вклучува **транскрипциски корепресори** (Сл. 2.28). Овие протеини се врзуваат за одредени транскрипциски фактори и мобилизираат хистон деацетилази за генските промотори, при што настанува отстранување на ацетилните групи од хистоните и промени во хроматинската структура.

Основна состојба



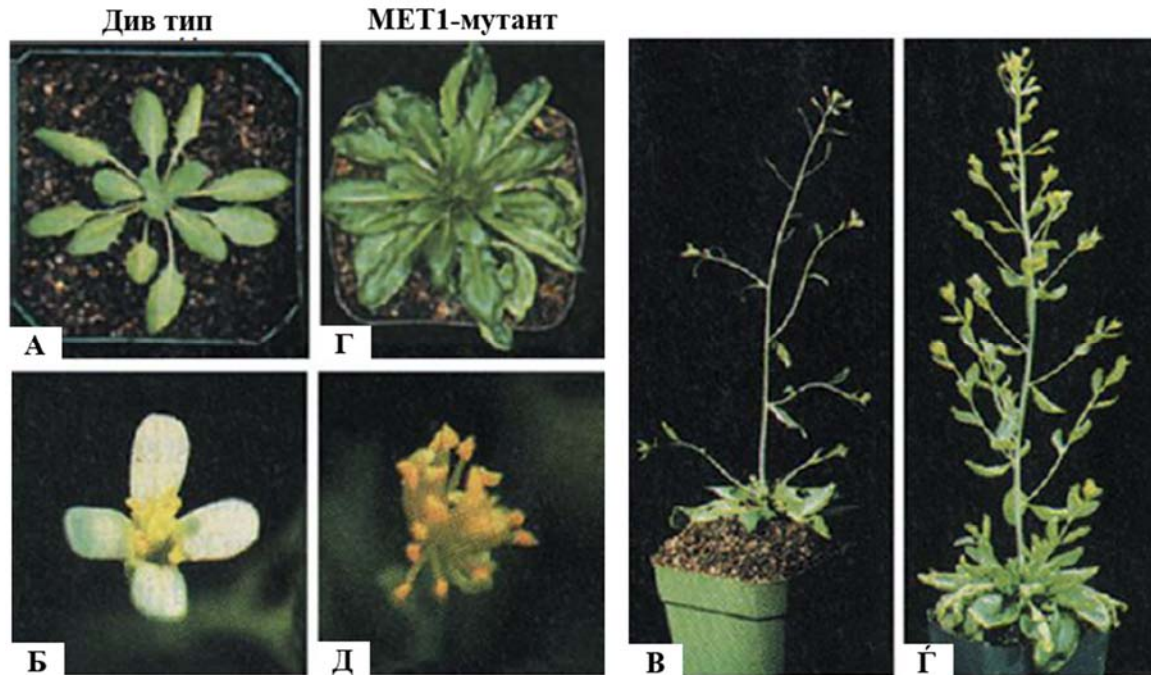
Слика 2.28 Механизам на дејство на транскрипциски корепресор. Врзувањето на корепресорот за комплексот на транскрипциски фактори предизвикува мобилизирање на хистон деацетилаза (HDAC) за промоторот на целиот ген. Потоа, HDAC катализира отстранување на ацетилни групи од специфични позиции на хистонот H3. Оваа промена во хроматинската структура ја инактивира транскрипцијата (Buchanan и сор., 2015).

2.5.4 Епигенетските механизми го контролираат растот и развитокот на растенијата

Епигенетските механизми контролираат голем број на процеси кои се поврзани со развитокот на растенијата. Карактеристичен пример за епигенетски механизам кој влијае врз развитокот на растенијата е регулацијата на времето на цветање кај *Arabidopsis*. Кај зимските едногодишни растенија (растенија кои 'ртат во есен, го преживуваат зимскиот период и цветаат во пролет), генот *FLOWERING LOCUS C (FLC)* ја инхибира иницијацијата на цветањето. Инхибицијата на цветањето се отстранува за време на зимскиот период, бидејќи растенијата се експонирани на пролонгиран период со ниска температура преку процесот означен како **вернализација**. Основниот молекуларен механизам зависи од активноста на polycomb репресивен комплекс 2 (PRC2), кој го репресира *FLC*-генот преку три-метилација на лизин 27 од хистонот H3. Откако ќе се исклучи експресијата на *FLC*-генот, сигналите од животната средина (фотопериод) овозможуваат иницијација на цветањето.

Исто така, polycomb групата (PcG) на протеини се вклучени во репресија на развитокот на растенијата преку активноста врз гените за развој на семе и ембрион. Губењето на активноста на PcG предизвикува намалување на експресијата на овие гени во вегетативните ткива на растенијата, што може да индуцира **соматска ембриогенеза** (формирање на ембрионот од соматски клетки без оплодување). Друг пример за влијание на епигенетските механизми во контролата на растот на растенијата е улогата на DNA-метилација во феноменот познат како **геномски импринтинг**. Овој термин претставува невообичаен начин на наследување, бидејќи активноста на алелите зависи од тоа дали се пренесени во потомството од растението мајка или татко. Всушност, импринтингот се базира на

разликите во моделите на метилација на цитозинот помеѓу геномите на таткото или мајката, кои настануваат под дејство на ензимот DNA-деметирирачка гликозилаза (DEMETETER). Овој ензим примарно се експресира во женскиот гаметофит и на таков начин се редуцира степенот на метилација во геномот на мајката.



Слика 2.29 Влијание на активноста на DNA-метилтрансфераза врз развитокот на растенија (Buchanan и сор., 2015). Компарација на фенотипови на див тип (А–В) и мутантна линија MET1 (Г–Ѓ) од *Arabidopsis thaliana*. Губењето на оперативноста на генот MET1 за метилација на цитозин е поврзано со абнормален фенотип за време на развитокот на вегетативните органи (А, Г), цветовите (Б, Д) и соцветијата (В, Ѓ).

Исто така, DNA-метилацијата може да има директен ефект врз развитокот на растенијата. Така на пр., нарушувањето на активноста на MET1 предизвикува редуцирана метилација во геномот, при што се добива фенотип со доцно цветање, како и абнормален развиток на изданоци и цветови (Сл. 2.29).