

КИЇВСЬКИЙ НАЦІОНАЛЬНИЙ УНІВЕРСИТЕТ
ІМЕНІ ТАРАСА ШЕВЧЕНКА
НАВЧАЛЬНО-НАУКОВИЙ ІНСТИТУТ ВИСОКИХ ТЕХНОЛОГІЙ

Завідувач кафедри кафедри молекулярної біотехнології та біоінформатики:

к.б.н., доцент Нипорко О. Ю.

Протокол № ____ засідання кафедри

від “ ____ ” _____ 2023 р.

**ДОСЛІДЖЕННЯ ВПЛИВУ ФІЗИЧНИХ КОНТАКТІВ
ЕНХАНСЕР-ПРОМОТОР НА ЕКСПРЕСІЮ ГЕНІВ**

Випускна кваліфікаційна робота магістра
студента спеціальності 091 Біологія
ОП «Біоінформатика та структурна біологія»
Смялковської Валерії Вікторівни

Науковий керівник від кафедри:
професор кафедри молекулярної біології та біоінформатика
д.б.н. Солдаткін Олексій Петрович

Робота виконана у групі механізмів регуляції геному

Німецького центру дослідження раку

під керівництвом PhD **Angelika Feldmann**

Оцінка захисту роботи

КИЇВ – 2023

АНОТАЦІЯ

Смялковська В. В. Дослідження впливу фізичних контактів енхансер-промотор на експресію генів. Випускна кваліфікаційна робота магістра за спеціальністю 091 Біологія ОП «Біоінформатика та структурна біологія».

У роботі було розроблено модель таргетованого анфолдингу контактів енхансер-промотор, протестовано її функціонування на локусі Sox2 у ембріональних стовбурових клітинах миші. В результаті, модель успішно рекрутується до визначених цілей, проте не впливає на структуру хроматину та транскрипцію.

Також було оцінено вплив деградації когезину на експресію генів протягом чотирьох годин нейрональної диференціації. Встановлено, що відсутність когезину змінює транскрипцію у стовбурових клітинах незначним чином, проте помітно порушує транскрипцію генів, індукованих протягом перших чотирьох годин диференціації.

Ключові слова: регуляція геному, енхансер, промотор, експресія генів, просторова структура хроматину, стовбурові клітини, диференціація клітин.

ЗМІСТ

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ	5
ВСТУП.....	6
РОЗДІЛ 1. Огляд літератури	8
1.1. Компоненти регуляції транскрипції.....	8
1.1.1. Транскрипція у еукаріотах	8
1.1.2. Структура хроматину	9
1.1.3. Транскрипційні фактори	10
1.1.4. Епігенетичні модифікації.....	11
1.1.5. Цис-регуляторні елементи	14
1.2. Енхансери як важливий елемент регуляції транскрипції	15
1.2.1. Історія дослідження та визначення енхансерів.....	15
1.2.2. Патології, спричинені дизрегуляцією енхансерів	17
1.2.3. Механізми функціонування енхансерів.....	18
РОЗДІЛ 2. Матеріали та методи	21
2.1. Аналіз даних	21
2.2. Вектор та гайдові РНК (гРНК).....	21
2.3. Клонування	24
2.3.1. гРНК	24
2.3.2. PARP1	24
2.4. Бактеріальна культура	24
2.5. Культура клітин.....	25

2.5.1. Культура клітин для трансфекції.....	25
2.5.2. Диференціація клітин та деградація когезину	25
2.6. Трансфекція	26
2.7. Імунопреципітація хроматину (ChIP)	26
2.8. Полімеразна ланцюгова реакцій у реальному часі (qPCR).....	27
2.9. Capture-C	29
2.9.1. Підготовка бібліотеки.....	29
2.9.2. Гібридизація та секвенування.....	31
2.10. Підготовка бібліотек для ncRNA-seq	34
2.11. Вестерн блот	34
РОЗДІЛ 3. Результати та обговорення	36
3.1. Модель таргетованого анфолдингу	36
3.1.1. Вибір початкового таргету	36
3.1.2. Тестування таргетованого анфолдингу на локусі Sox2	37
3.2. Модель нетаргетованого анфолдингу	41
ВИСНОВКИ	46
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	47

ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ

mESCs – ембріональні стовбурові клітини миші (mouse embryonic stem cells)

ТФ – транскрипційний фактор

RNAPII – РНК-полімераза II (RNA polymerase II)

ТАД – топологічно асоційований домен

qPCR – полімеразна ланцюгова реакція в реальному часі (quantitative polymerase chain reaction)

RA – ретиноева кислота (retinoic acid)

ВСТУП

Енхансери є ділянками геному, що представляють один з шарів регуляції експресії генів [1]. Вони можуть бути розташовані проксимально або дистально відносно промоторів, знаходитися до або після них, але необхідною умовою для їх функціонування є специфічний фізичний контакт з промотором гена, що регулюється .

Функція енхансерів полягає в активації транскрипції гена, забезпечуючи точне та вчасне включення гена у клітинній активності. Порушення регуляції енхансерів, такі як мутації або порушення комунікації з промотором, можуть приводити до ряду хвороб (енхансеропатій) [2], таких як порушення розвитку та ракові захворювання, що робить енхансери важливим об'єктом для дослідження.

Проте, механізм функціонування енхансерів досі є питанням для дискусії. Наразі, усі науково підтверджені моделі взаємодії енхансера з промотором передбачають фізичний контакт між даними ділянками. Найбільш вивченою моделлю є формування петлі [3], але також існують дані в підтримку таких моделей, як трекінг [4], лінкінг [5] та релокація [6].

На противагу «контактним» моделям взаємодії, було продемонстровано, що під час активації експресії гену *Shh* відбувається дистанціювання промотору від підтверджених енхансерів протягом нейрональної диференціації ембріональних стовбурових клітин миші (mESCs) [7]. На підтримку даної теорії також виступає експеримент з деградації когезину у mESCs [8], основного посередника в моделі формування петлі, за якого рівень експресії більшості генів зазнав незначних змін. Дані дослідження ставлять під питання роль когезину та безпосередніх контактів енхансер-промотор у активації та підтримці експресії генів.

Тому, метою даного дослідження є з'ясування ролі контактів енхансер-промотор та регуляторних елементів, направлених на їх формування, у активації та підтримці транскрипції.

Відповідно до мети були поставлені наступні завдання:

1. Створити гнучку систему для таргетованого розділення підтверджених контактів енхансер-промотор, провести її тестування на моделі mESCs.
2. Дослідити ефект нетаргетованого анфолдингу шляхом деградації когезину у процесі нейрональної диференціації, що може пролити світло на роль когезину у формуванні контактів *de novo*.

РОЗДІЛ І

ОГЛЯД ЛІТЕРАТУРИ

1.1. Компоненти регуляції транскрипції

1.1.1. Транскрипція у еукаріотах

Першим кроком в експресії генів є транскрипція. Це один з механізмів, що лежить в основі центральної догми молекулярної біології [9]. Протягом транскрипції на матриці ДНК відбувається синтез РНК, яка згодом стане матрицею для трансляції білку.

Експресія генів у еукаріотах - це багаторівневий і тісно регульований процес. Першим рівнем складності для клітинного транскрипційного механізму є організація геному в хроматин. Хроматин-асоційовані білки, включаючи гістони та не-гістонові білки, допомагають компактно упакувати великий геном в межах ядра. Динамічна взаємодія різноманітних клітинних факторів, таких як ферменти, що модифікують гістони, АТФ-залежні ремоделери, та гістонові шаперони разом з компонентами хроматину сприяють експресії генів та іншим необхідним клітинним процесам, які залежать від ДНК [10]. Транскрипція починається з послідовного рекрутування загальних факторів транскрипції та РНК-полімерази II (RNAPII) до промотору гена, що відбувається після специфічного для послідовності зв'язування активатора з елементами енхансера [11]. Існує кілька критичних етапів, що призводять до продуктивної транскрипції, включаючи формування функціонального преініціаторного комплексу (PIC), ініціацію транскрипції, продуктивну елонгацію транскрипту та завершення транскрипції на сайтах термінації.

Цикл еукаріотичної транскрипції включає в себе ініціацію, елонгацію та термінацію та регулюється на кількох етапах. Складна регуляторна мережа транскрипції включає в себе просторову структуру хроматину, регульовану

модифікаціями гістонів, ТФ та регуляцію RNAPII, цис-регуляторні елементи та епігенетичні модифікації. Продуктивна транскрипція вимагає вивільнення зупиненої RNAPII з промотора та переходу до елонгації транскрипту. Загальною тенденцією, що виникла, є те, що баланс між позитивними та негативними регуляторними факторами визначає загальну частоту транскрипції [12]. Фактори, пов'язані з термінацією транскрипції, також можуть посередниками в генетичному петлінгу та тим самим визначати ефективність повторної ініціації [13].

1.1.2. Структура хроматину

Геномна ДНК у еукаріотичних клітинах упаковується в хроматин - структуру, яка контролює практично всі ядерні процеси, пов'язані з ДНК: транскрипцію, реплікацію та репарацію ДНК. Упаковка ДНК в хроматин в основному залежить від двох основних типів невеликих позитивно заряджених білків: корових гістонів (H2A, H2B, H3 і H4) та зв'язувального гістону H1 [14, 15].

Базовою одиницею хроматинової нитки є нуклеосомне ядро, яке складається з двох копій кожного з чотирьох гістонових корових білків, що утворює октамер, здатний упаковувати 147 пар основ лівозакрученої суперспіралізованої ДНК [16]. У межах нуклеосомного ядра кожен гістон складається з альфа-спірального С-термінального домену, так званої гістонової складки, що приймає участь у формуванні гістон-гістонових зв'язків у октамері, та вільного N-кінцевого хвоста, відкритого до посттрансляційних модифікацій. Нуклеосомні ядра зв'язані лінкерною ДНК, довжина якої зазвичай варіює між 10 та 90 парами основ, що утворює регулярну структуру, так звані «намистини на нитці». Створюючи ще більш компактну структуру, лінкерний гістон H1 зв'язує вільну ДНК в безпосередній близькості до нуклеосомного ядра, що називають хроматосоною [17]. Хроматосоми створюють другий рівень компактизації ДНК, більш конденсовану 30-нм хроматинову нитку [18]. Корові та лінкерні гістони, здебільшого,

використовують позитивно заряджені залишки амінокислот аргінін та лізин для електростатичної взаємодії з фосфатним остовом ДНК.

Упаковка ДНК є початковим механізмом регуляції експресії, оскільки здатність ТФ та інших білків взаємодіяти з ДНК визначається її доступністю. Нуклеосома є гнучкою структурою, що здатна підвищувати або зменшувати доступність ДНК у відповідь на посттрансляційні модифікації N-термінальних неупорядкованих хвостів, про що йтиме мова нижче.

1.1.3. Транскрипційні фактори

Ще однією деталлю пазлу регуляції транскрипції є транскрипційні фактори (ТФ) – білки, що зв'язують послідовність-специфічні ділянки [19, 20]. ТФ зазвичай мають типово структуровані та еволюційно консервативні ДНК-зв'язуючі домени, що є важливою характеристикою ТФ. Специфічні послідовності ДНК, що зв'язують ТФ, часто класифіковані у мотиви – набір коротких послідовностей ДНК (в середньому 6-12 пар основ), зв'язуваних певним ТФ [21]. Окрім цього, певні ТФ також можуть впливати на структуру хроматину, взаємодіючи з гістонами чи рекрутуючи ремоделери хроматину.

ТФ дуже відрізняються за тим, як вони впливають на транскрипцію при зв'язуванні ДНК. Певні людські ТФ можуть безпосередньо рекрутувати РНК-полімеразу, тоді як інші залучають додаткові фактори, що впливають на конкретні етапи транскрипції [22]. Додатково, ТФ можуть діяти за допомогою стеричних механізмів, які не передбачають специфічної ефекторної функції, наприклад, шляхом блокування сайту зв'язування для інших протеїнів [23]. Однак вважається, що більшість еукаріотичних ТФ діють шляхом рекрутування кофакторів [24]. Коактиватори чи корепресори, задіяні в таких механізмах, початково визначені як медіатори ефекторної активності ТФ, часто є великими мультисубодиничними протеїновими комплексами чи мультидоменними протеїнами, що регулюють

транскрипцію різними шляхами. Зазвичай домени у їх складі залучені у зв'язування хроматину, ремоделінг нуклеосом, та/або ковалентні модифікації гістонів чи інших білків, включно з ТФ та РНК-полімеразою [22]. Загалом, кофактори можуть бути класифіковані у три основні класи: комплекс Медіатор [25], нуклеосомні ремоделери [26] та модифікатори гістонів [27], а також інші скафолди чи адаптори, що зв'язують ТФ та кофактори [28]. Комплекс Медіатор грає центральну роль у фосфорилуванні та активації пре-ініціаторного комплексу на багатьох промоторах, а нуклеосомні ремоделери забезпечують відкритість хроматину, необхідну для активності промоторів та віддалених енгансерів. Класичним прикладом даного механізму є $IFN\beta$ енгансосома. Зв'язування численних ТФ результує в рекрутменті GCN5/KAT2A та CBP/p300 гістонацетилтрансфераз [29]. Привнесені зміни до локального хроматинового середовища рекрутують такі нуклеосомні ремоделери, як SWI/SNF комплекс, щоб створити простір для зв'язування РНК полімерази та ініціації транскрипції.

Важливо зауважити, що функціональність ТФ залежить від контексту, тому поділ ТФ на активатори та репресори є складною задачею.

1.1.4. Епігенетичні модифікації

Епігенетична регуляція транскрипції представляє собою модифікації ДНК, що впливають на експресію генів без зміни послідовності нуклеотидів [30-32]. Більшість відомих епігенетичних модифікацій є оборотними. Вони включають метилювання ДНК, посттрансляційні модифікації гістонів та некодуючі РНК. Ці механізми є важливими для ініціації та підтримки епігенетичного сайленсингу та регулювання профілю експресії генів і є основоположними для ряду клітинних процесів, включаючи клітинну диференціацію, інактивацію Х-хромосоми, ембріогенез тощо.

Метилування залишків цитозину у 5'-положенні представляє собою оборотну ковалентну модифікацію ДНК, результатом якої є 5-метил-цитозин. У ссавців, здебільшого, метилуванню піддаються цитозини розташовані в так званих CpG-острівцях, де p позначає фосфодієфірний зв'язок між нуклеотидами [33]. Загалом, метилування ДНК асоційоване з репресією генів [34]. Вважається, що метилування ДНК призводить до транскрипційного сайленсингу шляхом кількох механізмів. Один механізм передбачає специфічну взаємодію ДНК-метилтрансфераз з факторами транскрипції, що призводить до сайт-специфічного метилування в промоторних областях [35]. Це сайт-специфічне метилування згодом відповідає за збирання в цих місцях білків, які розпізнають метильовану ДНК. Потім ці збірки безпосередньо впливають на подальшу роботу механізму транскрипції або викликають зміни в структурі хроматину, що, у свою чергу, впливає на механізм експресії генів [36, 37].

Іншим епігенетичним важелем є модифікації неупорядкованих N-термінальних хвостів гістонів [38]. Існують різні модифікації гістонів, серед яких ацетилювання [38], фосфорилування [39], метилування [40], убіквітилування [41] та сумоїлювання [42]. Ці модифікації зустрічаються як на канонічних гістонах [43] (H2A, H2B, H3 і H4), так і на варіантах (такі як H3.1, H3.3 і H2T.1) [44]. Більшість модифікацій локалізуються на аміно- та карбоксикінцевих гістонових хвостах.

Наприклад, гістонацетильтрансферази (HATs) і гістондеацетильтрансферази (HDACs) додають і та прибирають ацетилювання [45]. Ацетилювання може зменшити позитивний заряд залишків лізину, що потім пригнічує зв'язування між хвостами гістону та негативно зарядженою ДНК, залишаючи ДНК відкритою [46, 47]. Тому ацетилювання гістонів зазвичай розглядають як активну гістонову мітку.

Метилування гістонів, яке зазвичай відбувається в залишках лізину гістону H3 і H4 шляхом додавання метильних груп, є однією з найважливіших посттранскрипційних модифікацій. Воно каталізується гістонметилтрансферазою

(HMT), яка використовує S-аденозилметіонін як субстрат для перенесення метильних груп на залишки лізину гістонів [48]. Залишки лізину гістонів можуть бути моно-, ди- та триметильованими (me1, me2 та me3 відповідно). Загалом H3K4, H3K36 і H3K79 вважаються активними мітками, які локалізовані на активно транскрибованих генних ділянках хроматину. H3K9, H3K27 і H4K20 відомі як репресивні модифікації, які зазвичай асоціюються зі зниженою експресією генів і конденсованим хроматином [49].

Фосфорилування також відбувається на залишках серину та тирозину гістонів [50]. Фосфорилування гістонів бере участь у великій кількості клітинних процесів, включаючи експресію генів, регуляцію клітинного циклу, репарацію ДНК та асиметричний поділ клітин [51, 52]. Сайти метилювання, H3K9 і H3K27, мають залишок серину, який може бути фосфорильований, утворюючи домен «KS». Таке близьке розташування між цими двома сайтами модифікації наводить на гіпотезу про те, що фосфорилування H3 може змінити, принаймні частково, спорідненість білків «читачів» або «письменників» до контактуючих залишків лізину [53]. Цей взаємозв'язок між фосфорилуванням і метилюванням гістонів дозволяє просторово-часово контролювати мітки гістонів, що призводить до складної, але точної регуляції структури хроматину [54].

Малі некодуєчі РНК (нкРНК) зазвичай утворюються з більших молекул попередників РНК шляхом розщеплення ферментами родини рибонуклеаз III (зазвичай Drosha та Dicer) і включають мікроРНК, короткі інтерферуючі РНК, РНК, що взаємодіють з PIWI, та РНК, асоційовані з повторами, на додаток до інших, менш вивчених видів.

В порівнянні з дволанцюговими РНК, що виконують функцію інгібування трансляції (РНК інтерференція), малі ядерні РНК приймають участь в ядерному процесі, відомому як пригнічення експресії генів, керуючи не тільки збірку гетерохроматину та сайленсинг генів в *S. pombe* [55], але й запрограмовану

елімінацію ДНК в *T. Thermophila* [56]. У 2002, 4 наукових групи продемонстрували взаємодію малих РНК, і, ймовірно, направлення, хроматинових ремоделерів до геномних таргетів, використовуючи моделі дріждів [57, 58] або *T. thermophila* [59, 60].

Дані дослідження підкреслюють думку, що ДНК, РНК та гістони, разом з їх модифікаціями, працюють разом, впливаючи на стани хроматину, тим самим регулюючи важливі геномні функції.

1.1.5. Цис-регуляторні елементи

Блок-кодуючі ділянки геному є переважно консервативними між людьми та шимпанзе, тоді як фенотипові відмінності між спорідненими видами забезпечуються кількісними змінами у експресії генів [61]. Дані зміни, здебільшого, є генетично кодованими, що підтверджується високою спадковістю багатьох фенотипових рис у різних видів та людських популяцій, вказуючи на існування іншого коду на основі ДНК, що цим керує [62]. Його часто називають цис-регуляторний код, оскільки експресія генів контролюється регуляторними елементами ДНК, які діють в «цис». Це означає, що такі елементи знаходяться на тому ж самому хромосомному алелі, як і регульований ген. Цис-регуляторний код дуже варіює в залежності від клітинного типу, оскільки фактори транскрипції, що прямо зчитують ДНК, є різними з-поміж різних клітинних типів [21]. Транскрипція відбувається так званими імпульсами, частота та амплітуда яких є об'єктом регуляції, але які є стохастичними за своєю суттю [63]. Цис-регуляторні елементи класифіковані залежно від їх розташування відносно генів (проксимальні до сайту старту транскрипції промотери чи дистальні елементи) та їх впливу на транскрипцію (енхансери, що підвищують рівні транскрипції, сайленсери які зменшують транскрипцію, чи інсулятори що розділяють регуляторні домени).

Цис-регуляторні елементи можуть бути різні за розміром. Наприклад, вони можуть бути досить короткі, як even-skipped енхансер 2 смуги у дрозофіл, що кодується лише 480 парами основ [64]. Багато цис-елементів знаходяться в мегабазах від ділянок, що вони регулюють, та взаємодіють з ними шляхом фізичного контакту чи епістатично, наприклад, локуси Sox9 [65] та Мус [66].

Початково енхансери визначались як короткі послідовності, що автономно регулюють транскрипцію незалежно від розташування відносно промотора [1]. репортерним аналізом [67]. Кожний цис-регуляторний елемент може виступати як набір сайтів зв'язування для різних наборів ТФ відносно клітинного контексту. Тому, один і той самий елемент може бути зв'язаний репресивним ТФ та функціонувати як сайленсер в одному клітинному типі, та бути зв'язаним активуючим ТФ і функціонувати як енхансер в іншому.

Організація сайтів зв'язування транскрипційних факторів в цис-регуляторні елементи дозволяє інтеграцію багаточисленних сигналів для досягнення просторово-часової специфічності [68].

Хоча цис-регуляторні елементи, такі як енхансери, є модульними "блоками" транскрипційної регуляції, що залежать від послідовності та клітинного контексту, кінцевим результатом цис-регуляторного коду є транскрипція генів. В складних багатоклітинних організмах енхансери часто функціонують на відстані десятків чи сотень кілобаз від промоторів, що вони регулюють. Деякі енхансери людей регулюють гени на відстані однієї мегабази від них [65]. Енхансери здатні демонструвати промоторну специфічність, пропускаючи близько розташовані гени для регуляції дистальних [69].

1.2. Енхансери як важливий елемент регуляції транскрипції

1.2.1. Історія дослідження та визначення енхансерів

У 1981 році було вперше продемонстровано некодуючу послідовність ДНК, яка "посилювала" експресію цис-кодованого гена, у спосіб, відмінний від транскрипційної активації, опосередкованої промоторами [1, 70]. Зокрема, у епізомальному репортерному векторі, некодуюча ділянка геному вірусу SV40 збільшувала експресію на відстані, віддаленій від промотора репортерного гена, і не залежала від орієнтації підсилювальної ділянки. З цього експерименту з'явилося оригінальне визначення енхансера, яке широко застосовується і сьогодні: "елемент транскрипційного енхансера може діяти в будь-якій орієнтації в багатьох позиціях, навіть знаходячись нижче від сайту ініціації транскрипції" [1]. Кілька років по тому, використовуючи подібний метод *in vitro*, в локусі IgH був ідентифікований перший ендogenous енхансер, специфічний для клітин ссавців [71-73].

Протягом наступного десятиліття, з ростом інтересу до енхансерів, були сформовані кілька загальних принципів. По-перше, енхансери не містять нуклеосом, про що свідчить підвищена чутливість до ДНКаз I [74, 75], але оточені нуклеосомами зі специфічними, пов'язаними з транскрипцією модифікаціями гістонів [76-78]. По-друге, енхансери містять кластери мотивів зв'язування ТФ [79], і зв'язування ТФ з цими мотивами лежить в основі їхньої енхансерної активності [80, 81]. По-третє, енхансери, ймовірно, розташовуються в безпосередній близькості до своїх промоторів-мішеней шляхом формування просторових петель ДНК [5, 82].

Наше сучасне розуміння полягає в тому, що енхансери зв'язуються специфічними для типу клітин ТФ, асоційовані з ділянками відкритого хроматину і фланкуються гістонами, що несуть модифікації H3K27ac та/або H3K4me1. Вони взаємодіють зі специфічними промоторами в 3D-просторі і можуть бути латентними, праймованими або активними [83]. Енхансери хребетних можуть мати довжину 100-1000 пар основ, а декілька енхансерів можуть існувати в кластері,

утворюючи супер-енхансер [84] - аналогічно до раніше описаних локус-контрольованих ділянок [85]. Енхансери знаходяться переважно в міжгенних та інтронних ділянках, тоді як кілька енхансерів було виявлено в екзонах.

1.2.2. Патології, спричинені дизрегуляцією енхансерів

Дисфункція енхансерів стала центральним механізмом у патогенезі деяких захворювань. Зокрема, дисфункція енхансерів внаслідок точкових мутацій або структурних варіантів є важливим механізмом, що лежить в основі аберантної регуляції генів при раку [86] та менделічних хворобах [87]. Крім того, генетичні варіанти, пов'язані з поширеними захворюваннями, часто зустрічаються в цис-регуляторних елементах, включаючи енхансери [62].

Структурна варіація може спричинити "викрадення енхансерів", якщо активний енхансер транслокується таким чином, що він регулює гени, які не є його початковими мішенями [88]. Це явище було описано в ракових захворюваннях, де енхансери можуть бути переміщені через хромосомні перебудови [88] або мікроампліфікацію геномних ділянок [89], щоб індукувати експресію онкогену. Нещодавній аналіз профілів повногеномного секвенування у понад 1200 геномах ракових клітин виявив сотні генів, для яких зміни експресії спостерігалися в межах 100 кілобаз від структурної варіації, деякі з яких можуть бути пов'язані з викраденням енхансерів [90]. Ці дані свідчать про те, що викрадення енхансерів може бути досить поширеним механізмом при ракових захворюваннях.

Інший шлях, який може спричинювати містартетинг енхансерів - це порушення меж хроматинових доменів, так званих топологічно асоційованих доменів (ТАД). Порушення кордонів ТАД може призвести до взаємодії енхансера з генами за його межами, що було ідентифіковано як механізм захворювання при трьох вроджених розладах, включаючи вади розвитку кінцівок [91].

Загальна кількість консервативних варіантів в енхансерах, що регулюють певний ген, є хорошим предиктором його патогенності [18], що дозволяє припустити, що мутації в енхансерах, ймовірно, сприяють розвитку вроджених захворювань. Ряд мутацій енхансерів призводить до вроджених захворювань в тому числі до деяких неврологічних захворювань, як нещодавно було розглянуто в [92].

1.2.3. Механізми функціонування енхансерів

Як саме енхансер активує транскрипцію гена-мішені - постійне питання в галузі регуляції генів еукаріот. Щоб здійснювати регуляцію, енхансери повинні взаємодіяти з цільовими промоторами. Однак у вищих еукаріотів енхансери фізично відокремлені від промоторів генів-мішеней вздовж геному. Було запропоновано кілька моделей комунікації між енхансером і промотором. Вони включають трекінг, лінкінг, релокацію та петлеутворення.

Модель трекінгу припускає, що RNAPII після зв'язування з енхансером переміщується вздовж ДНК, доки не зв'яжеться з промотором [4, 93]. Якщо припустити, що RNAPII та інші активатори, зв'язані з енхансерами, рухаються вздовж ланцюга ДНК в односпрямованому напрямку, то після досягнення контакту з промотором ланцюг ДНК успішно формує петлеву структуру. На користь цієї моделі свідчать результати імунопреципітації хроматину, які продемонстрували односпрямовані рухи комплексу RNAPII-ТФ-енхансер на ДНК, що рухається до промотора [4]. Іншим доказом є існування некодуючих РНК, в тому числі коротких поліаденільованих РНК, що транскрибуються з ділянки енхансера, які іноді охоплюють область між енхансером і промотором [94-96]. Однак, модель трекінгу не може пояснити, як енхансери та промотори взаємодіють між собою на міжхромосомному рівні [97], що свідчить про те, що взаємодія енхансер-промотор відбувається або за іншим, або за кількома механізмами.

Модель лінкінгу є розвитком спостережень за репресором лямбда, згідно з якою зв'язок між енхансером і промотором відбувається через утворення білкового скаффолду. Одним із запропонованих з'єднувальних факторів у тварин є білок Chip *Drosophila* [98], який, як було запропоновано, олігомеризується від енхансера до промотора. Однак більш недавні дослідження з ортологом хребетних, *Lbd1*, свідчать про те, що він може утворювати цільові петлі шляхом гомодимеризації при зв'язуванні з енхансерами та промоторами [99].

Ще однією моделлю є релокація. Ця модель припускає, що і енхансер, і промотор переміщуються до ядерного компартменту, де вони взаємодіють один з одним. У цій ділянці, яку називають транскрипційною фабрикою [100], високі локальні концентрації RNAPII сприятимуть ефективній транскрипції. В одній клітині спостерігаються сотні таких фабрик, причому, за оцінками, кожна фабрика містить кілька генів [6]. Хоча точна стабільність транскрипційних фабрик у живих клітинах залишається дискусійною [101, 102], модель фабрики RNAPII сумісна з орієнтаційно-незалежною природою енхансерів.

Модель петлеутворення припускає, що фактори, зв'язані з двома різними сайтами, фізично взаємодіють один з одним, що призводить до екструзії проміжної ДНК. Перші докази на користь петлеутворення були отримані на прикладі *Escherichia coli* понад 30 років тому [103]. Оскільки зв'язування ТФ є динамічним, формування петлі також відбувається динамічно. Для прикладу, когезин, що має кільцеву структуру, захоплює ДНК у просвіт своєї кільцеподібної структури, утворюючи невелику хроматинову петлю. Ця спочатку невелика хроматинова петля надалі екструдується когезином, доки когезин не зіткнеться з CTCF-зв'язаними ділянками у протилежній орієнтації, так званими кордонами ТАД. Накопичені дані свідчать на користь даної моделі. По-перше, хроматинові петлі і ТАД різко послаблюються при деградації хроматин-зв'язаного когезину [104, 105]. По-друге, кількість і час перебування зв'язаного з хроматином

когезину безпосередньо визначають розмір хроматинової петлі: клітини з дефіцитом фактора розвантаження когезину WAPL демонструють хроматинові петлі збільшеного розміру [105], тоді як інактивація комплексу завантаження когезину NIPBL-MAU2 призводить до вкорочення хроматинових петель [106]. По-третє, експериментальна інверсія орієнтації обраних CTCF-зв'язуючих сайтів призводить до порушення хромосомних петель, що в деяких випадках супроводжується зміною експресії генів [107, 108].

Можливо, також існують безконтактні механізми дії енхансерів. Наприклад, експресія гену *Shh* в мозку миші, що розвивається, контролюється кількома енхансерами *Shh*, але в клітинах-попередниках нейронів під час експресії *Shh* не було виявлено жодних контактів між промотором *Shh* і цими енхансерами [7]. З іншої сторони, деградація когезину призводила до незнаних порушень у експресії більшості генів [8]. Ці дослідження вказують на те, що енхансерний контроль експресії генів може включати механізми, які не передбачають прямого контакту з промотором.

РОЗДІЛ 2

МАТЕРІАЛИ ТА МЕТОДИ

2.1. Аналіз даних

Аналіз та візуалізація даних qPCR проводився у Microsoft Office Excel за методом $2^{-\Delta\Delta C_t}$.

Контроль якості прочитань був виконаний у програмі FastQC. Для Capture-C Fastq-файли було вирівняно за геномом mm10 і фільтровано за допомогою HiCUP [109] і Bowtie 2 [110]. Потім парні bam-файли було оброблено за допомогою пакета R Bioconductor Chicago [111].

Для ncRNA-seq риди було вирівняно за геномами в mm10 та dm6 за допомогою STAR (Dobin et al., 2013). Аналіз диференціальної експресії генів проведено за допомогою пакета R DESeq2 [112].

2.2. Вектор та гайдові РНК (гРНК)

У дослідженні було використано вектор рх330-dCas9 довжиною 8511 пар основ, в якому кодовані наступні основні компоненти (рис. 2.1):

- dCas9, мутант, що позбавлений нуклеазної активності, який містить 3xFLAG-тег для подальшої детекції розташування експресованого dCas9;
- AmpR – ген резистентності до ампіциліну 100мг/мл, що застосовується для селекції трансформованих бактерій;

Плазміда рKLV2 (рис. 2.2) була використана як носій гРНК та селективного маркера – гену резистентності до пуроміцину. Це полегшило клонування векторів та дало можливість відбирати для аналізу тільки трансфіковані клітини.

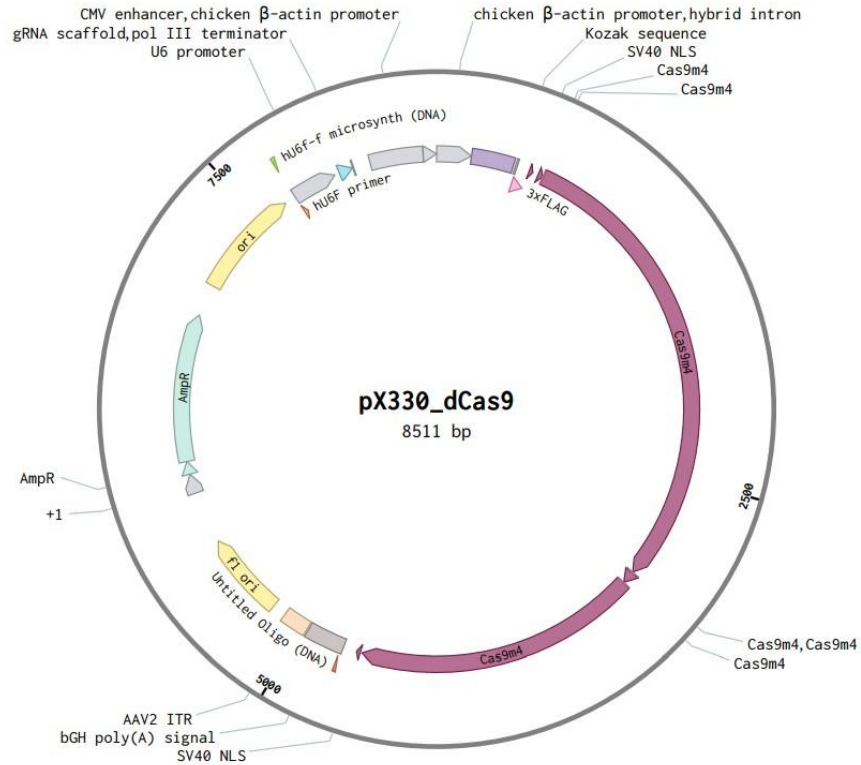


Рис. 2.1. Схема плазмідного конструктору pX330-dCas9.

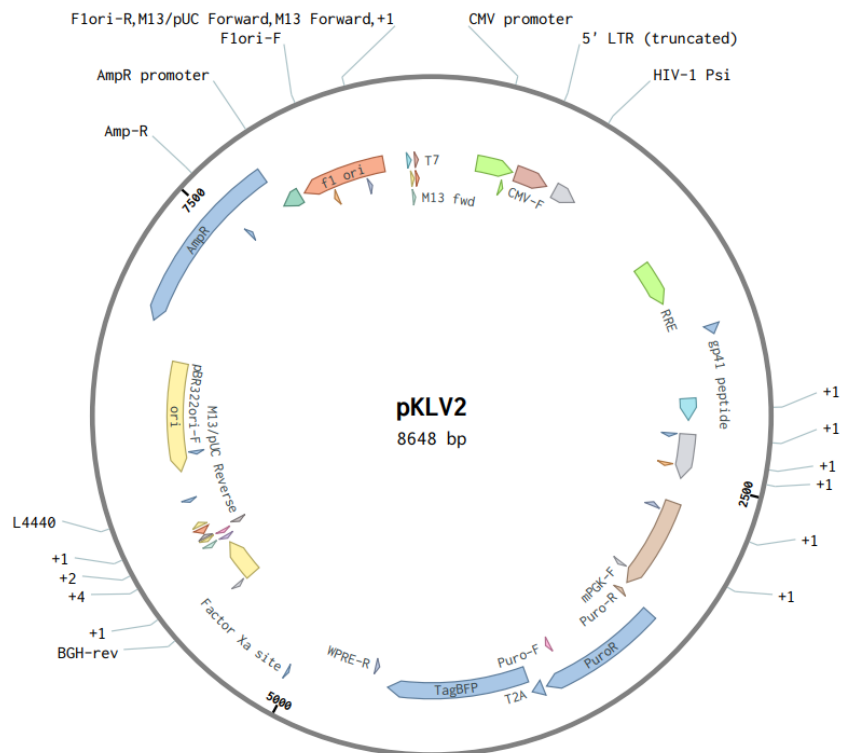


Рис. 2.2. Схема плазмідного конструктору pKLV2.

Олігонуклеотиди, що використовувалися для кодування гРНК (таблиця 2.1), були вироблені на замовлення компанією Invitrogen Thermo Fisher Scientific.

Також було використано ген PARP1, довжина якого ~3 кілобази, синтезований у векторі компанією Twist Biosciences.

Для аналізу та збору векторів було використано веб-сервіс Benchling (<https://benchling.com/>).

Таблиця 2.1 Послідовності гРНК, застосованих в дослідженні

Назва олігонуклеотиду	Послідовність олігонуклеотиду
SOX2_prom3_F	CACCGGACAGCCCCATCACATGGA
SOX2_prom3_R	AAACTCCATGTGATGGGGGCTGTCC
SOX2_prom5_F	CACCGGGCGCTCGCTGATTGGCCGC
SOX2_prom5_R	AAACGCGGCCAATCAGCGAGCGCCC
vs45_Sox2Enh_sgRNA_f1	CACCGCTCATCTGGCTCGGAAAGCT
vs46_Sox2Enh_sgRNA_r1	AAACAGCTTTCCGAGCCAGATGAGC
vs47_Sox2Enh_sgRNA_f2	AAACAGCTTTCCGAGCCAGATGAGC
vs48_Sox2Enh_sgRNA_r2	CACCGTGGGCCAGTTGTAAACCCCC
vs49_Sox2Enh_sgRNA_f3	AAACGGGGGTTTACAACCTGGCCCAC
vs50_Sox2Enh_sgRNA_r3	CACCGAGACCTTGTTTACGTGTTAG
vs51_Sox2Enh_sgRNA_f4	AAACCTAACACGTAAACAAGGTCTC
vs52_Sox2Enh_sgRNA_r4	CACCGATTCCACTAACACGTAAACA
vs53_Sox2Enh_sgRNA_f5	AAACTGTTTACGTGTTAGTGGAATC
vs54_Sox2Enh_sgRNA_r5	CACCGACATATCAATGACTAGGCG
vs55_Sox2Enh_sgRNA_f6	AAACCGCCTAGTCATTGATATGTC
vs56_Sox2Enh_sgRNA_r6	CACCGCGAGGCCTTAGCTACGAAAC

2.3. Клонування

2.3.1. гРНК

Для підготовки кожної гРНК 100мкМ прямого та зворотнього олігонуклеотидів відпалювались за 95°C 5 хв, далі температура знижувалась на 5 °C за хвилину до 25 °C.

Суміш рестрикції лігації містить: 1мкл вектору (100нг/мкл), 2 мкл олігонуклеотидного дуплексу, 2мкл Tango буферу (Thermo Fisher Scientific), 1 мкл 10мМ DTT, 1 мкл 10мМ АТР, 1 мкл BbsI (NEB), 0.5мкл лігази T7 (NEB),

11.5мкл води.

Суміш було інкубовано 15 циклів за наступною схемою:

- 37°C протягом 5 хвилин
- 21°C протягом 5 хвилин

2.3.2. PARP1

Щоб мати можливість використовувати ген PARP1, його необхідно вирізати з вектору. Для цього було поставлено реакцію рестрикції, яка містила по 20 одиниць ензимів EcoRI-HF та SpeI-HF (NEB), 10x rCutSmartBuffer (NEB), 500нг вектору. Реакція інкубувалась за температури 37 °C протягом ночі.

Реакційну суміш було розділено електрофорезом в агарозному гелі, фракцію, що відповідала 3 кілобазам, вирізано та виділено з гелю з використанням ZymoClean DNA Recovery Kit (Zymo Research).

Вектор рх330 перед вставкою PARP1 було лінеаризовано таким самим чином.

Реакція лігації фрагменту PARP1 з лінеаризованим вектором включає в себе фрагмент та вектор у співвідношенні 3:1, 20000 одиниць лігази T4 (NEB), 10x T4 DNA Ligase Buffer (NEB). Реакцію було інкубовано за температури 16 °C протягом ночі, після чого ензим було деактивовано за 65 °C протягом 10 хвилин.

2.4. Бактеріальна культура

Ампліфікація плазмиди проводилась з використанням культури *E. coli* XL10-Gold® Ultracompetent Cells (Agilent). 25мкл бактерій було трансформовано 300 нг вектору. Тепловий шок було проведено за температури 42°C протягом 45 сек. Далі бактерії інкубувалися з додаванням 250мкл середовища S.O.C. (Invitrogen) 1 год за 37 °C для стабілізації.

25мкл бактеріальної суспензії було посіяно на агар з додаванням ампіциліну 1000:1 та інкубовано одну ніч при 37 °C. Наступного дня відібрані колонії були посіяні у 5 мл LB з додаванням ампіциліну 1000:1 та інкубовано одну ніч при 37 °C. Для виділення та очищення вектору було використано ZymoPURE Plasmid Miniprep Kit (Zymo Research). Контроль якості отриманих конструктів було проведено за допомогою секвенування за Сенгером (виконане компанією Microsynth, Гейдельберг).

2.5. Культура клітин

2.5.1. Культура клітин для трансфекції

Для експерименту було використано mESCs E14. Для культивування використовується DMEM з додаванням 15% фетальної телячої сироватки (Gibco), 1% 200мМ L-глутаміну (Gibco), 1% незамінних аміноксилот (Gibco), 1% пеніцилін-стрептоміцину (Gibco), 0,1% бетамеркаптоетанолу (Gibco) та 1:125 LIF (фактор інгібування лейкемії, власного виробництва).

Клітини були посіяні на чашки Петрі з 0.1% желатиною підложкою для імітації міжклітинного матриксу. Клітини піддавалися трипсинізації кожні два дні з використанням TrypLE (Gibco).

2.5.2. Диференціація клітин та деградація когезину

Для даного експерименту було використано Scc1-AID модифіковані mESCs [8]. Середовище для диференціації клітин та деградації когезину містить у собі усі

компоненти, окрім LIF. У середовище для диференціації було додано 1:100 розчин ретиноєвої кислоти у DMSO у пропорції 1:1000. Для деградації когезину у середовище було додано 500 μ M ауксину. Клітини були інкубовані 2 години за присутності когезину, та 4 години за присутності ретиноєвої кислоти.

2.6. Трансфекція

Транз'єнтна трансфекція проводилась за ліпід-опосередкованим механізмом з використанням Lipofectamine 3000 (Invitrogen). Для таргетування промотора гену Sox2 клітини були трансфіковані двома плазмідами разом, для енансеру – шістьма. Загальна кількість ДНК для кожної трансфекції складала 2.5мкг, об'єм трансфекційної суміші складав 200мкл.

2.7. Імунопреципітація хроматину (ChIP)

Для ChIP було зібрано 10^7 клітин. Подвійне зшивання хроматину для подальшого ChIP проводилась з використанням 8мкл 500 μ M розчину DSG у DMSO. Після 45 хвилин інкубації за кімнатної температури було додано 125мкл 16% формальдегіду (Thermo Fisher Scientific) та інкубовано протягом 15 хвилин. Далі, реакція була погашена 750мкл розчину 1M гліцину та інкубовано 5 хвилин за кімнатної температури.

Після цього було проведено ізоляцію ядер послідовними промивками розчинами LB1 (25x інгібітори протеаз, 100% гліцерол, 1M HEPES-KOH (pH 7.9), 5M NaCl, 0.5M EDTA, 10% NP40, 10% Triton-X100), LB2 (25x інгібітори протеаз, 1M Tris-HCl pH 8.0, 5M NaCl, 0.5M EDTA, 1M EGTA), LB3 (25x інгібітори протеаз, 1M Tris-HCl pH 8.0, 5M NaCl, 0.5M EDTA, 1M EGTA, 10% Nlaurylsarcosine, 10% sodium deoxycholate).

Після очищення зразки були соніковані 25 циклів (Bioruptor Pico, 30 сек вкл/вимк.) та змішані з 10% Triton-X100 LB3.

Для імунопреципітації хроматину були використані моноклональні antiFLAG антитіла, продуковані у миші (Merck), 2мкг на кожну реакцію. Інкубація зразків з антитілами відбувалася 16 годин за 4°C. Перед додаванням антитіл було відібрано 2% об'єму зразка як вхідний зразок.

Далі, преципітовані ділянки хроматину разом з антитілами були очищені за допомогою надпарамагнітних сферичні полімерних частинок Protein A Dyna Beads (Invitrogen), які були блоковані 20мг/мл BSA та 10мг/мл тРНК дріжджів з метою уникнення неспецифічних зв'язувань.

В подальшому, зв'язаний мікрогранулами матеріал був послідовно промитий низькосолевым буфером (10% SDS, Triton X100, 0.5M EDTA, 1M Tris-HCl pH8, 5M NaCl), високосолевым буфером (10% SDS, Triton X100, 0.5M EDTA, 1M Tris-HCl pH 8, 5M NaCl), LiCl буфером (1M LiCl, 10% NP-40, sodium deoxycholate, 0.5M EDTA, 1M Tris-HCl pH8). Елюція преципітату була виконана в буфері для елюції, що містив 10% SDS, 1M NaHCO₃ та H₂O.

Після цього, зразки інкубувалися 16год за 65°C з RNase A (20мг/мл, Qiagen) та 1 год при 55 °C з протеїназою K (Invitrogen). Зразки були очищені з використанням Zymo ChIP DNA Clean and Concentrator Kit (Zymo Research).

Зразки були проаналізовані методом полімеразної ланцюгової реакції.

2.8. Полімеразна ланцюгова реакцій у реальному часі (qPCR)

Умови ПЛР реакції є універсальними для ChIP та RT-qPCR, температура плавлення кожного праймеру складає 60°C

ПЛР реакція включала в себе 5 мкл ДНК, 0.375мкл прямого праймеру (10мМ), 0.375мкл зворотнього праймеру (10мМ), 7.5 мкл PowerUp SYBR Green Master Mix (Applied Biosystems) та 1.75 мкл H₂O. Використані праймери перераховані у таблиці 2.2, умови ПЛР вказані у таблиці 2.3.

Таблиця 2.2. Послідовності праймерів для ПЛР

Назва праймеру	Послідовність праймеру
Int3_qRT_FP	ATGCCCCTCAGCTATCACAC
Int3_qRT_RP	GGACAGACATCTGCCAAGGT
H3K27ac_ES_21252_FP	TGAACAGGGTTGGGAGGGTA
H3K27ac_ES_21252_RP	AATGGGGATGCCACTCACTG
H3K27ac_notES_32091_FP	GGACCAGGTGTAAAGAGGCC
H3K27ac_notES_32091_RP	CCCTACACCGTGCCAGAAAT
Sox2P_SMsgFlanking_F5	GCCGCCGGAAACCCATTTA
Sox2P_SMsgFlanking_R5	GCAAACACTCTCTTCTCTGCC
Hoxb3_F	CCAACTCCACCCTCACAAA
Hoxb3_R	GCCACCACCACAACCTTCT
Hoxa7_F	AAGCCAGTTTCCGCATCTAC
Hoxa7_R	CTTCTCCAGTTCCAGCGTCT
Rad23b_F	TAATTGCAGCCCTGAGAGCC
Rad23b_R	TAGTTGCTGTTCGTGGTTGCT
Ikbkap_F	GCTGGCGGTATCTATGCTGT
Ikbkap_R	TCTGCCGAAAGACTGTCACC
Nanog_F	CAAAGGATGAAGTGCAAGCGG
Nanog_R	GGTGCTGAGGCCTTCTGAAT
Sox2_F	GAACGCCTTCATGGTATGGT
Sox2_R	TTGCTGATCTCCGAGTTGTG

Таблиця 2.3 - Умови ПЛР

Температур а, °С	Час, сек	Кількість циклів
50	120	1
95	120	1

95	3	45
60	30	
Крива плавлення		
95	15	1
60	60	1
95	15	1

ПЛР проводилась на ампліфікаторі StepOnePlus (Applied Biosystems).

2.9. Capture-C

2.9.1. Підготовка бібліотеки

Для бібліотеки було зібрано 10^7 клітин у 9.3мл клітинного середовища. Для фіксації хроматину клітини було інкубовано з 1.25мл 16% формальдегіду 10 хвилин за кімнатної температури. Реакцію було погашено з додаванням 1.5м 1М гліцину, клітинний осад промито PBS та піддано лізису протягом 20 хв за температури 4°C з використанням буферу для лізису, що містив 10mM Tris pH 8, 10 mM NaCl, 0.2% NP-40, 25x інгібітори протеаз. Після цього осад було ресуспендовано у 1мл лізисного буферу та розподілено навпіл (орієнтовно $5 \cdot 10^6$ клітин в аліквоті), зразки були швидко заморожені на сухому льоді та зберігались за -80°C до подальшої обробки.

Згодом $5 \cdot 10^6$ клітин було розморожено на льоді, центрифуговано, осад був розчинений у 250мкл 1x DpnII buffer (Thermo Fisher Scientific). До зразків було додано 10% SDS (фінальна концентрація 0.28%) та інкубовано 1 годину при 37°C. Потім було додано 20% Triton X-100 (фінальна концентрація 1.67%) та інкубовано додаткову годину. Після цього було додано 10мкл рекомбінантного DpnII (500

одиниць) та інкубовано при 37°C. Після ~5 годин було додано додаткові 10мкл, після чого ще 10мкл наступного дня. Наступним кроком була лігація, за якої до кожного експериментального зразка було додано 8мкл лігази T4 (240 одиниць, Thermo Fischer Scientific) та інкубовано протягом 22 годин за температури 16°C. Слідуючи за цим, зразки були очищені фенол-хлороформом. Зразки були соніковані 13 циклів (Bioruptor Pico, 30 сек вкл/вимк). Далі зразки об'ємом 200мкл були двічі очищені Ampure XP SPRI beads (Beckman Coulter), додано 0.85x бідсів перший раз, та 0.4x другий, елюція зразку в 86мкл Elution buffer (0.1x TE, Zymo Research). Дана процедура дозволяє здійснити відбір фрагментів довжиною ~200 пар основ.

Після сонікації було проведено відновлення кінців ДНК, для чого до 3мкг ДНК було додано 10мкл ER Reaction buffer та 5мкл ER Enzyme Mix (NEB) та інкубовано 30 хвилин за 20°C. Зразок був очищений 1.8x Ampure XP SPRI beads. Далі було проведено dA-tailing, для чого до попереднього міксу додавалися 5мкл dA Reaction Buffer та 3мкл Klenow (NEB). Реакція була інкубована при 37°C 30 хвилин. Було повторено очистку бідсами, як описано вище. Надалі зразки були адаптор ліговані при 20°C 45 хвилин з додаванням 10мкл Quick Ligation buffer, 15мкл NEB Adaptor, 5мкл T4 ДНК лігази. Надалі додано 3мкл USER Enzyme Mix та інкубовано 15 хвилин за 37°C. Була повторена очистка бідсами, описана вище. Після того до послідовностей ДНК було додано індекси у процесі ПЛР-ампліфікації з використанням полімерази Herculase II (Agilent). Умови ПЛР описані в таблиці 2.3. За цим слідувала подвійна очистка з 1.8x Ampure XP SPRI beads.

Для оцінки якості зразків, концентрації було виміряно спектрофотометричним методом на Qubit (Thermo Fisher Scientific), а розмір фрагментів оцінено методом капілярного електрофорезу на Bioanalyzer 2000 (Agilent).

2.9.2. Гібридизація та секвенування

Для гібридизації були використані пул біотинільованих проб, таргетованих на промотори, що представляють інтерес в даному дослідженні (табл. 2.4), в якому кожна проба 2.9nM.

1.5мкг кожної з раніше підготованих бібліотек були змішані разом перед початком процедури. Для блокування повторів в ДНК було додано 5мкг на бібліотеку COT DNA Mouse (Invitrogen). Зразки були висушені у вакуумній центрифугі за температури 60°C, після чого розчинені у гібридизаційному буфері та змішані з реагентами Nimblegen KAPA HyperCapture Reagent kit (Roche). До пулу зразків було додано 4.5мкл пулу гібридизаційних проб на бібліотеку. Умови гібридизації включають 95°C перші 5 хвилин, а потім 55°C. Зразки піддавалися гібридизації двічі, перший раз протягом 72 годин, другий – протягом 48.

Після цього проводиться очистка матеріалу від не гібридизованих фрагментів з використанням Nimblegen KAPA HyperCapture Bead Kit (Roche).

Молярність бібліотеки кількісно визначено за допомогою qPCR з використанням стандартів KAPA Illumina ДНК-стандартів (Roche) та секвеновано на платформі Illumina NextSeq 500.

Таблиця 2.3. Умови ПЛР для додавання індексів

Температура	Час, сек	Кількість циклів
98°C	30	1
98°C	10	6
65°C	30	
72°C	30	
72°C	500	1

af1173_ Stat3_CC_1	GATCTCCACCTGAGTGGCAGTGGCTCTGGTGGCCCG AACGTCCCTAGCCTCCGGGAGCAGTCACAGGAGG
af1174_ Stat3_CC_2	CCGAGGGCCGCCTGGCCTCTCCTAGCCGACGCGGTC AGGCGACACCGAGCGGGGTTGTACCTCAGCGATC
af1163_ Esrrb.2_CC_ 1	GATCCAGACCCGGGCTGGCATGTTTTGCCAGTATTTG TTGTGTGACCTGGAGAAGGACAATTCTGCACAC
af1164_ Esrrb.2_CC_ 2	TCTGGCGCTCCGCGTCCCATCGCGCTTCGGTGGAACT CGGAGCGTTCGTTGTGCGGTCCTGGAAATGATC
af1167_ Nanog_CC_1	GATCAGAGGATGCCCCCTAAGCTTCCCTCCCTCCCA GTCTGGGTCACCTTACAGCTTCTTTTGCATTAC
af1168_ Nanog_CC_2	GGCTGCGGCTCACTTCCTTCTGACTTCTTGATAATTT TGCATTAGACATTTAACTCTTCTTTCTATGATC
af2328_ Lifr_CC_1	GATCCGTTTCTTCTCTGCACAAATAAGCCCAACACGC CTCAGAGGCACGACTAGATGGCTCCAGGATGTT
af2329_ Lifr_CC_2	AGGCAAGCAAAGTTAGGCTTCCTAATGGATGGCTTT CCGGATGTGTTGTCTAAAGCACAAAGCTGAGATC
af2330_ Tfcp2l1_CC_ 1	GATCGGGGTCCGAGCGCTGACTAACCTCAGGCCTCC AGAGAAGGGCCCTCCGGGAGGCCCTAGGGAATG
af2331_ Tfcp2l1_CC_ 2	TTTTTCTTTTAAAAAATAACTCCAAACCATTGAAAAT GGTAATCAAACCTAGGGACATGAGACTGCAGATC

af2332_ Dppa5a_CC_ 1	GATCAAGCTTATCCACCACGCACGGCCCACAGCTCC AGG TTCAGGAAGTTTTAGTACCTGTGTGGGCGGT
af2333_ Dppa5a_CC_ 2	CGCTGGAAAGTTCAAAGACTTCTAAAAGACTATAGC TTTTCCCATCCCGAAAGAATCAGATACTGGGATC

2.10. Підготовка бібліотек для ncRNA-seq

10 мільйонів мишачих mESCs було змішано з 4 мільйонами клітин дрозофіли SG4. Для ізолювання клітинних ядер, зразки були послідовно промиті буферами HS (50mM KCl, 10mM MgSO₄·7H₂O, 5mM HEPES, 0.05% NP-40), двічі RSB (10mM NaCl, 10mM Tris pH 7.4, 3mM MgCl₂) та центрифуговані після кожного кроку. Остаточний осад був розчинений у PBS.

Тотальну РНК було виділено за допомогою набору RNeasy Mini Kit (QIAGEN) відповідно до протоколу з подальшою обробкою за допомогою набору TURBODNA-free Kit (Thermo Fischer Scientific). Якість РНК оцінено за допомогою набору 2100 Bioanalyzer RNA 6000 Pico (Agilent).

Для підготовки бібліотек кожен зразок був спочатку очищений від рРНК за допомогою набору NEBNext rRNA Depletion kit (NEB). Подальша обробка була виконана за допомогою набору NEBNext Ultra II Directional RNA Library Prep Kit for Illumina (NEB).

Молярність бібліотеки кількісно визначено за допомогою qPCR з використанням стандартів КАРА Illumina ДНК-стандартів (Roche) та секвеновано на платформі Illumina NextSeq 500.

2.11. Вестерн блот

До клітинного осаду було додано 100мкл буферу RIPA (20 mM Tris pH 7.5, 150 mM NaCl, 1% Nonidet P-40, 0.5% sodium deoxycholate, 1 mM EDTA, 0.1% SDS, інгібітори протеаз cOmplete) та інкубовано протягом 30 хвилин при 4°C. Концентрацію було виміряно методом Бредфорда.

35мкг білку для кожного зразка було розділено у 10% поліакриламідному гелі. Далі, розділені білкові фракції були перенесені на нітроцелюлозну мембрану з використанням Trans-Blot Turbo Transfer System (Bio-Rad) відповідно до рекомендацій виробника. Після блокування мембрани 5% розчином сухого молока у PBS-T, її було розділено навпіл. Верхня частина інкубувалася з антитілами проти RAD21 (ABCAM ab154769, 1:1000), нижня – проти TBP (ABCAM ab220788, 1:1000) протягом ночі. Наступного дня обидві мембрани інкубувалися з вторинними кролиними антитілами (1:5000) протягом 40 хвилин. Блот був проявлений з використанням BIORAD Clarity Western ECL Substrate (пероксид та люмінол змішані 1:1). Візуалізація проводилась на BIORAD ChemiDoc Imaging System з витримкою 3 хвилини.

РОЗДІЛ ІІІ

РЕЗУЛЬТАТИ ТА ОБГОВОРЕННЯ

3.1. Модель таргетованого анфолдингу

3.1.1. Вибір початкового таргету

Як демонструють дослідження, механізми функціонування енхансерів можуть бути різними, на що впливають різноманітні фактори, наприклад, локалізація енхансера та промотора, структура хроматину. Так як ми припускаємо, що механізми енхансер-промоторної взаємодії можуть бути локус-специфічними, досить важливо враховувати контекст в дослідженні даних механізмів. Саме тому ми ставимо за мету розробку таргетованої системи.

Одним з основних критеріїв для вибору гену є його транскрипційна активність. Оскільки ми тестуємо модель на mESCs, які містять не так багато активних генів через відсутність клітинної специфікації, увагу було звернено на гени плюрипотентності. Промотор гену Sox2 має виражену взаємодію з супер-енхансером, що був експериментально підтверджений [113].

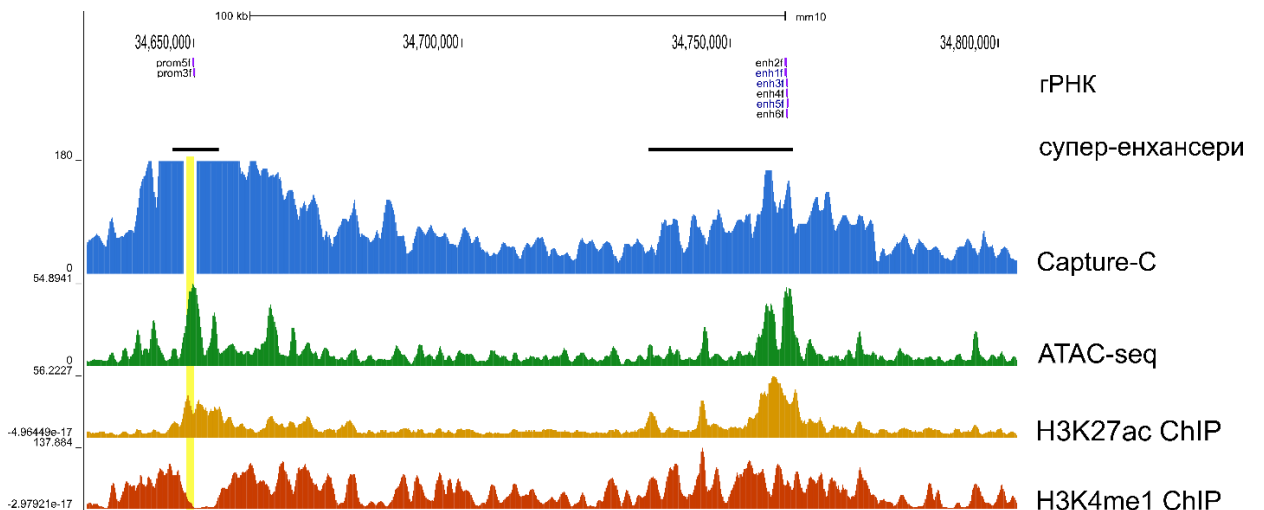


Рис. 3.1. Знімок геномного браузеру для гену Sox2 та його суперенхансера, а також локалізація гРНК використаних в дослідженні. Жовта лінія позначає viewpoint для Capture-C. Дані про супер-енхансери взяті з [84].

Також дана ідея підтримується результатами аналізу геномного браузеру UCSC для гену *Sox2*, розташованого на третій хромосомі (рис. 3.1). Capture-C даного регіону демонструє активні контакти промотора з енансером інтересу. Гістонова мітка H3K27Ac є міткою активних енансерів, і зі знімку ми можемо бачити, що обраний регіон містить супер-енансер, що також підтверджується модифікацією H3K4me1. ATAC-seq, метод для пошуку відкритого хроматину, також демонструє, що дана ділянка є активною та відкритою для взаємодій.

3.1.2. Тестування таргетованого анфолдингу на локусі *Sox2*

Для досягнення нашої мети ми використали вектор px330-dCas9, в якому розташування сайтів рестрикції відкривало можливість для додавання інших протеїнів. Також, в ході експерименту було застосовано методику зворотної трансфекції через її зручність, гнучкість та швидкість отримання результату. Загалом, було протестовано здатність до анфолдингу для dCas9, а також для комплексу dCas9-PARP1, в якому dCas9 забезпечував таргетну доставку до локусів інтересу, а PARP1 грав у анфолдингу як механічна та каталітична перепона, оскільки даний протеїн [114] може полі(АТФ)-рибозилювати білки, зокрема гістони, що приводить до релаксації хроматину [115]. Також, було продемонстровано здатність віддаляти таргетований енансер від промотору у локусі *Shh* [7].

Експериментом було передбачено наступні зразки: негативний контроль (neg ctrl), що не був трансфікований; клітини, трансфіковані плазмідом з dCas9, таргетованою на промотор (prom) чи енансер (enh) гену *Sox2*; клітини, трансфіковані плазмідом з dCas9-PARP1, таргетовані так само на промтор чи

енхансер цільового гену (prom-PARP та enh-PARP). Експеримент був відтворений у двох біологічних реплікатах.

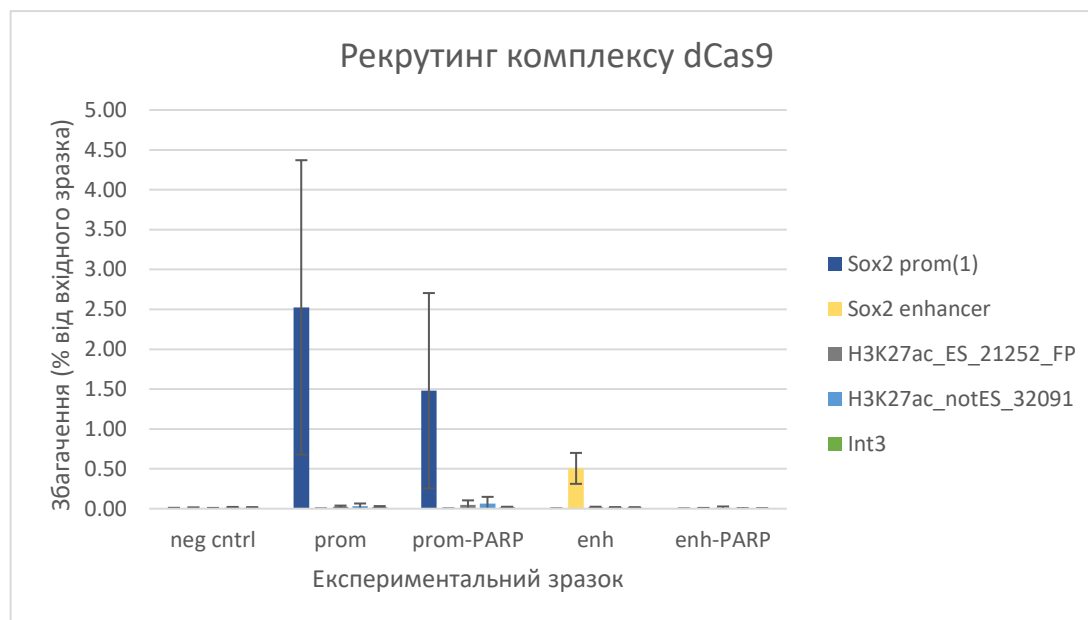


Рис. 3.1. Результати ChIP-qPCR

Як можна побачити з результатів ChIP-ПЛР на рис. 3.1, рівень рекрутменту обох комплексів dCas9 та dCas9-PARP1 до цільових таргетів, представлених праймерами Sox2-prom та Sox2-enh є значно вищими за фонові сигнали, в даному випадку представлені парами праймерів H3K27ac_ES, H3K27ac_notES та Int3. Це свідчить про те, що, в загальній мірі, нам вдалося досягти таргетованого рекрутменту обох комплексів. Однак, рівень рекрутменту зменшується за наявності PARP1, а у випадку enh-PARP і зовсім зникає. Дане явище потребує подальших досліджень, але початковим припущенням може бути смерть клітин до моменту збору, спричинена порушеннями в регуляторній мережі експресії.

Щоб оцінити дистанцію між промотором та енхасером цільового гену у всіх експериментальних зразках, було проведено Capture-C секвенування, що

демонструє контакти ДНК, які представляють інтерес, відібрані сетом проб, описаних вище. З результатів, представлених на рис. 3.2 для гену *Sox2*, можна побачити, що змін у паттернах контактів не відбулося в обох реплікатах, в порівнянні з негативним контролем.

Для оцінки змін у експресії генів було проведено ЗТ-ПЛР для генів, що активно експресуються у mESCs (*Sox2* та *Nanog*), неактивних клітин-специфічних генів (клітин-специфічні гени *Нохb3* та *Ноха7*), та стабільно експресованих генів (*Rad23b* та *Ikbkap*). Розрахунок змін проводився за допомогою методу $2^{-\Delta\Delta C_t}$ з нормалізацією відносно гену *Rad23b*. Як можемо бачити на рис. 3.3, рівень експресії гену *Sox2* знижується у трансфікованих зразках, при тому незалежно від присутності PARP1, тоді як експресія *Nanog* не зазнає істотних змін. Також варто звернути увагу на профілі експресії генів *Нохb3* та *Ноха7*.

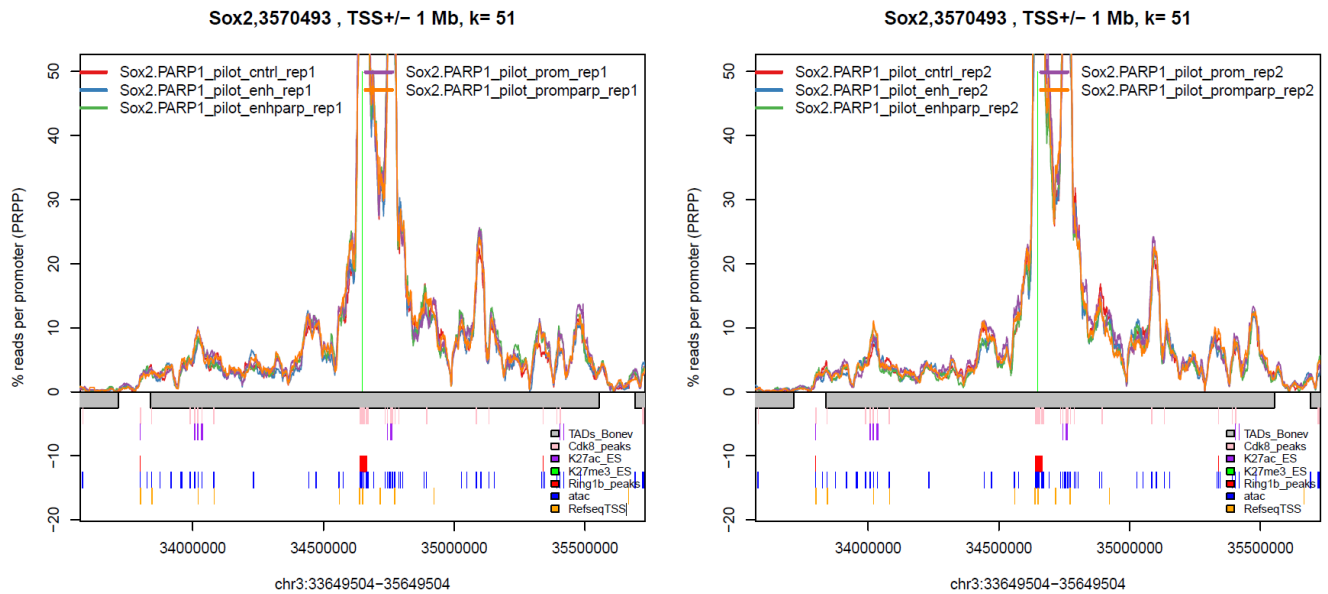


Рис. 3.2. Мапа геномних контактів viewpoint (зелена пряма) з +/- 1 мегабаза від сайту старту транскрипції гену *Sox2* для обох реплікатів

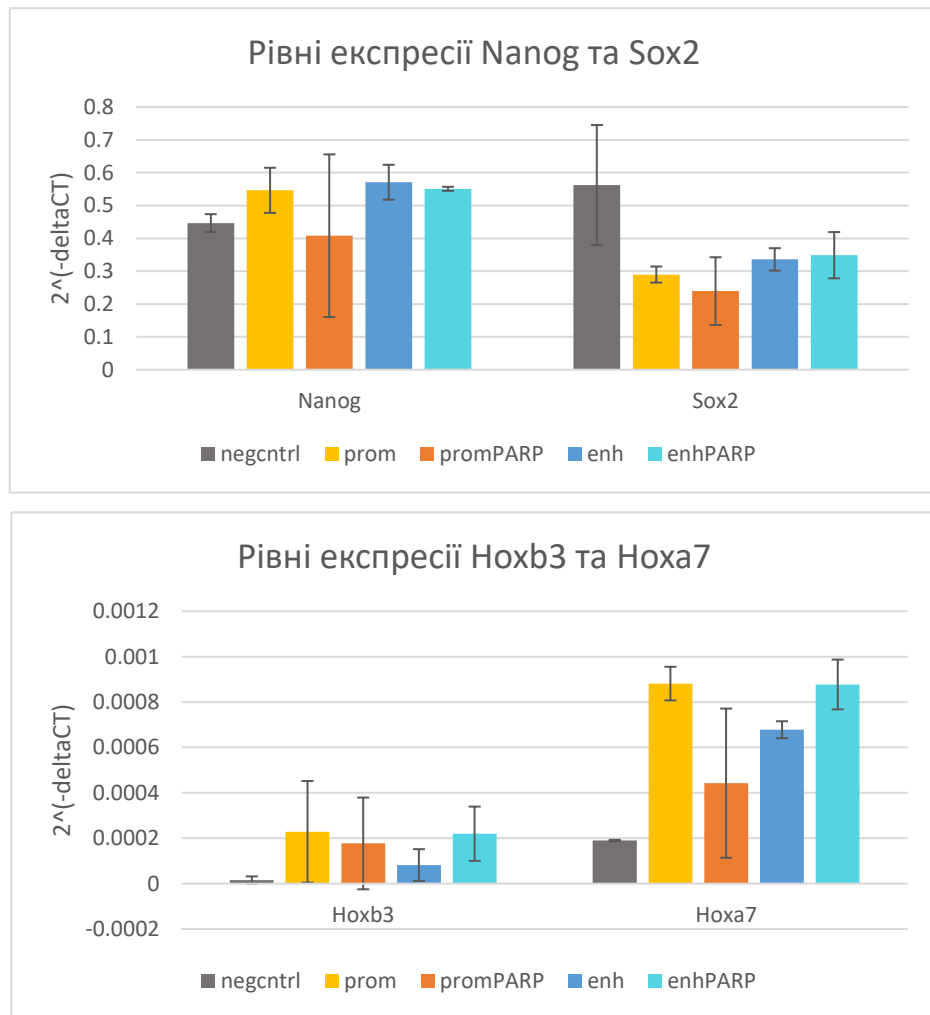


Рис. 3.3. Рівні експресії генів Nanog, Sox2, Hoxb3 та Hoxa7, нормалізовані відносно експресії гену Rad23b.

Дані гени активно транскрибуються за нейрональної диференціації, та мають бути неактивними в mESCs, що відслідковується у негативному контролі. Незважаючи на те, що рівень експресії все ще залишається низьким, він є значно вищим за контроль у експериментальних зразках, особливо для гену Hoxa7. Можливим поясненням може бути наші спостереження за негативним контролем, що був оброблений селективним маркером пуроміцином, у порівнянні з контролем без обробки. На рис. 3.4. продемонстровано, що патерни експресії в контролі, обробленому пуроміцином, збігаються з представленими для експериментальних зразків, а саме знижена експресія Sox2 та завищена експресія

Ноха7. Це може свідчити про сторонній вплив пуроміцину, що потребує дослідження.

Отже, в ході даного експерименту нам вдалося досягти таргетованого рекрутменту dCas9 як з PARP1, так і без нього. Проте, ряд питань залишається відкритими. По-перше, зниження рівня рекрутменту до енхансерної ділянки, і, навіть, його зникнення за присутності PARP1 потребує більш детального вивчення. По-друге, незважаючи на рекрутмент, змін у контактах енхансер-промотор не було зафіксовано. Нам відомо, що дія PARP1 може мати локус-специфічний характер [116], тому даний комплекс також має бути протестовано на інших генах. Також, транскрипційні ефекти, що потенційно можуть бути спричинені пуроміцином, потребують подальших експериментів.

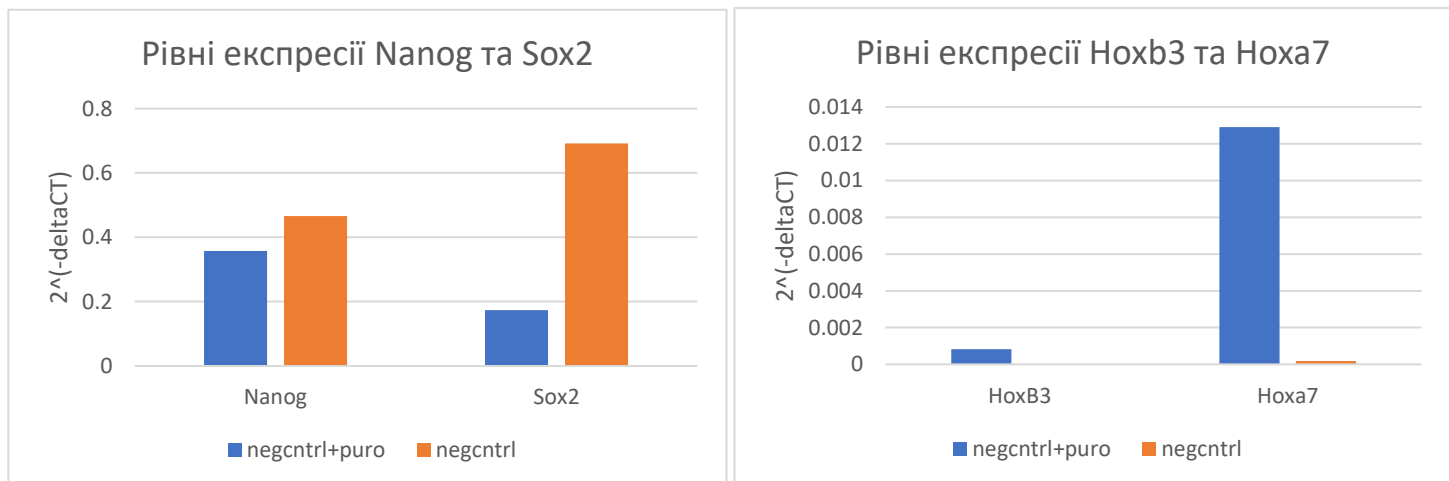


Рис. 3.4. Порівняння рівнів експресії для генів Nanog, Sox2, Hoxb3 та Hoxa7 поміж контролем з доданим селективним марекром і без.

3.2. Модель нетаргетованого анфолдингу

Як було зазначено вище, когезин є важливим посередником у контакті енхансер-промотор. Даний білок відповідальний за екструзію петлі вздовж ТАД, що значно полегшує комунікацію між енхансером та промотором, які розташовані в межах ТАД. У mESCs було продемонстровано, що за деградації

субодиниці когезину Scc1 протягом 6 годин, продовжує відбуватись велика частина контактів енхансер-промотор, а експресія генів порушена лише незначною мірою – експресія лише кількох сотень генів зазнала суттєвого зниження [8], що також підтверджується іншим дослідженням [117].

Нами було вирішено з'ясувати роль когезину у процесі формуванні нових контактів та налагодження транскрипції клітин-специфічних в під час диференціації клітин. З цією метою було застосовано раніше опубліковані mESCs з Scc1-дегрон ауксин-індукованою системою (Scc1-AID) [8]. Додавання ауксину до даних клітин активує процес деградації субодиниці когезину Scc1. Також, було проведено нейрональну диференціацію, індуковану ретиноевою кислотою (RA) [118]. Клітини були інкубовані 2 години з ауксином, та 4 години з RA. Експеримент було проведено у трьох біологічних реплікатах.

Після 2-годинної обробки ауксином було проведено вестерн блот для перевірки деградації субодиниці Scc1. На рис. 3.5 продемонстровано, що білок успішно деградує за даний проміжок часу. ТАТА-зв'язуючий білок виступає в якості контролю навантаження.

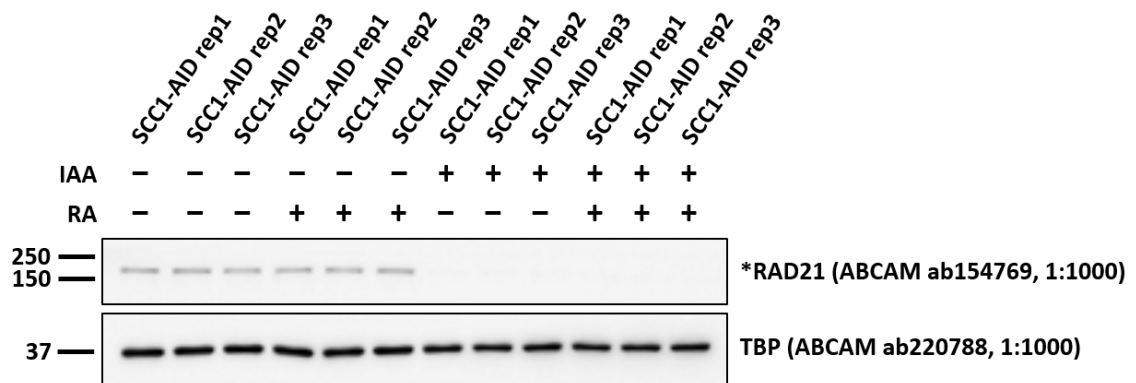


Рис. 3.5. Блотограма білкових екстрактів Scc1-AID клітин демонструє деградацію при обробці ауксином. IAA – ауксин, RA – ретиноева кислота, TBP – ТАТА-зв'язуючий білок.

Значна частина генів змогла досягти ефективної експресії генів навіть за деградації когезину (рис. 3.6). Це може бути пояснене тим, що регуляторна мережа даних генів вже може бути встановлена. Тоді як ситуація для індукованих генів відрізняється. Більшість індукованих протягом диференціації генів, серед яких і ген *Greb11*, не здатні стабілізувати експресію без когезину в тій самій мірі, що з ним. Також, якщо співставити даний результат з мапою Capture-C для гену *Greb11* в ході перших чотирьох годин RA-індукованої диференціації (рис. 3.7, неопубліковані дані лабораторії), то можна помітити, що промотор даного формує нові контакти протягом перших чотирьох годин з моменту додавання RA. Дане спостереження підтримує теорію, що когезин є більш важливим для формування контактів *de novo*, ніж для підтримки вже встановлених регуляторних програм [119].

На противагу цій групі генів, ген *Noxal* майже не втрачає свого рівня експресії за відсутності когезину. Можливо це пояснюється відсутністю нових контактів у перші чотири години диференціації (рис. 3.7).

Підводячи підсумки, можна припустити, що когезин відіграє більш значну роль у створенні контактів *de novo*, ніж у підтримці вже встановлених. В найближчій перспективі, буде зроблено Capture-C для встановлення змін у структурі хроматину.

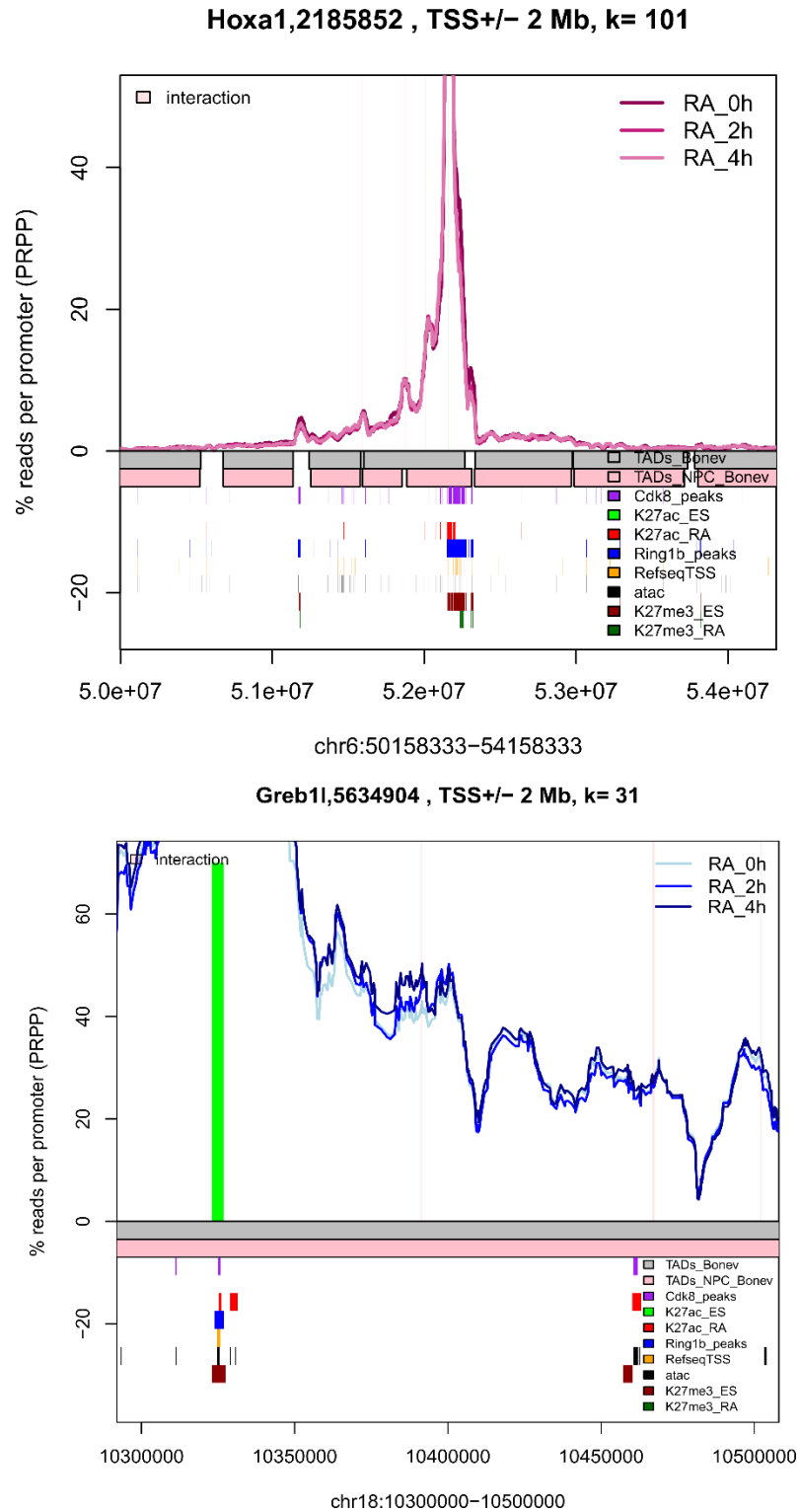


Рис. 3.7. Capture-C для генів *Hoxa1* та *Greb1* протягом перших чотирьох годин RA-індукованої диференціації (неопубліковані дані лабораторії).

ВИСНОВКИ

В ході даної роботи було створено систему для таргетованого розділення промоторів з їх підтвердженими енансерами на основі зворотної трансфекції з використанням вектору, що містив dCas9 та PARP1. Модель було протестовано на mESCs, наразі лише на єдиному локусі – гені плюрипотентності Sox2. У результаті, нам вдалося досягти ефективного рекрутингу до відібраного енансера та промотора гену Sox2, згідно з результатами ChIP-qPCR. Проте, очікуваного ефекту на структуру хроматину та експресію генів не було досягнуто. Як було продемонстровано раніше, PARP1 діє у локус-специфічному манері, тому одним з наступних кроків буде тестування системи на інших генах.

Також, було протестовано нетаргетований анфолдинг протягом нейронального диференціювання клітин шляхом ауксин-індукованої деградації когезину. Дані ncRNA-seq демонструють, що відсутність когезину у mESCs впливає на експресію незначним чином. Проте, це має більший вплив на нейрональні попередники, які не були здатні відновити рівні експресії більшості клітин-специфічних генів, що не є активними у mESCs. Можна припустити, що когезин грає важливішу роль у формуванні нових контактів, ніж у підтримці вже існуючих.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Banerji J, Rusconi S, Schaffner W. Expression of a β -globin gene is enhanced by remote SV40 DNA sequences. *Cell*. 1981;27(2):299-308. doi: 10.1016/0092-8674(81)90413-X.
2. Claringbould A, Zaugg JB. Enhancers in disease: molecular basis and emerging treatment strategies. *Trends in Molecular Medicine*. 2021;27(11):1060-73. doi: 10.1016/j.molmed.2021.07.012.
3. Berg OG, Winter RB, Von Hippel PH. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry*. 1981;20(24):6929-48.
4. Hatzis P, Talianidis I. Dynamics of Enhancer-Promoter Communication during Differentiation-Induced Gene Activation. *Molecular Cell*. 2002;10(6):1467-77. doi: [https://doi.org/10.1016/S1097-2765\(02\)00786-4](https://doi.org/10.1016/S1097-2765(02)00786-4).
5. Ptashne M. Gene regulation by proteins acting nearby and at a distance. *Nature*. 1986;322(6081):697-701.
6. Papantonis A, Cook PR. Transcription Factories: Genome Organization and Gene Regulation. *Chemical Reviews*. 2013;113(11):8683-705. doi: 10.1021/cr300513p.
7. Benabdallah NS, Williamson I, Illingworth RS, Kane L, Boyle S, Sengupta D, et al. Decreased Enhancer-Promoter Proximity Accompanying Enhancer Activation. *Molecular Cell*. 2019;76(3):473-84.e7. doi: 10.1016/j.molcel.2019.07.038.
8. Rhodes JDP, Feldmann A, Hernández-Rodríguez B, Díaz N, Brown JM, Fursova NA, et al. Cohesin Disrupts Polycomb-Dependent Chromosome Interactions in Embryonic Stem Cells. *Cell Reports*. 2020;30(3):820-35.e10. doi: 10.1016/j.celrep.2019.12.057.
9. Crick F. Central Dogma of Molecular Biology. *Nature*. 1970;227(5258):561-3. doi: 10.1038/227561a0.
10. Tyler JK. Chromatin assembly. *European Journal of Biochemistry*. 2002;269(9):2268-74. doi: <https://doi.org/10.1046/j.1432-1033.2002.02890.x>.
11. Thomas MC, Chiang C-M. The general transcription machinery and general cofactors. *Critical reviews in biochemistry and molecular biology*. 2006;41(3):105-78.
12. Li C, Cesbron F, Oehler M, Brunner M, Höfer T. Frequency Modulation of Transcriptional Bursting Enables Sensitive and Rapid Gene Regulation. *Cell Systems*. 2018;6(4):409-23.e11. doi: <https://doi.org/10.1016/j.cels.2018.01.012>.
13. Hampsey M, Singh BN, Ansari A, Lainé J-P, Krishnamurthy S. Control of eukaryotic gene expression: Gene loops and transcriptional memory. *Advances in Enzyme Regulation*. 2011;51(1):118-25. doi: <https://doi.org/10.1016/j.advenzreg.2010.10.001>.

14. Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature*. 1997;389(6648):251-60. doi: 10.1038/38444.
15. Kornberg RD, Thomas JO. Chromatin Structure: Oligomers of the Histones. *Science*. 1974;184(4139):865-8. doi: doi:10.1126/science.184.4139.865.
16. Arents G, Burlingame RW, Wang BC, Love WE, Moudrianakis EN. The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proceedings of the National Academy of Sciences*. 1991;88(22):10148-52. doi: doi:10.1073/pnas.88.22.10148.
17. Simpson RT. Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. *Biochemistry*. 1978;17(25):5524-31. doi: 10.1021/bi00618a030.
18. Widom J, Klug A. Structure of the 3000Å chromatin filament: X-ray diffraction from oriented samples. *Cell*. 1985;43(1):207-13. doi: [https://doi.org/10.1016/0092-8674\(85\)90025-X](https://doi.org/10.1016/0092-8674(85)90025-X).
19. Fulton DL, Sundararajan S, Badis G, Hughes TR, Wasserman WW, Roach JC, et al. TFCat: the curated catalog of mouse and human transcription factors. *Genome Biology*. 2009;10(3):R29. doi: 10.1186/gb-2009-10-3-r29.
20. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM. A census of human transcription factors: function, expression and evolution. *Nature Reviews Genetics*. 2009;10(4):252-63. doi: 10.1038/nrg2538.
21. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, et al. The Human Transcription Factors. *Cell*. 2018;172(4):650-65. doi: <https://doi.org/10.1016/j.cell.2018.01.029>.
22. Fietze S, Farnham PJ. Transcription Factor Effector Domains. In: Hughes TR, editor. *A Handbook of Transcription Factors*. Dordrecht: Springer Netherlands; 2011. p. 261-77.
23. Akerblom IE, Slater EP, Beato M, Baxter JD, Mellon PL. Negative Regulation by Glucocorticoids Through Interference with a cAMP Responsive Enhancer. *Science*. 1988;241(4863):350-3. doi: doi:10.1126/science.2838908.
24. Reiter F, Wienerroither S, Stark A. Combinatorial function of transcription factors and cofactors. *Current Opinion in Genetics & Development*. 2017;43:73-81. doi: <https://doi.org/10.1016/j.gde.2016.12.007>.
25. Richter WF, Nayak S, Iwasa J, Taatjes DJ. The Mediator complex as a master regulator of transcription by RNA polymerase II. *Nature Reviews Molecular Cell Biology*. 2022;23(11):732-49. doi: 10.1038/s41580-022-00498-3.
26. Clapier CR, Iwasa J, Cairns BR, Peterson CL. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nature Reviews Molecular Cell Biology*. 2017;18(7):407-22. doi: 10.1038/nrm.2017.26.

27. Morgan MAJ, Shilatifard A. Reevaluating the roles of histone-modifying enzymes and their associated chromatin modifications in transcriptional regulation. *Nature Genetics*. 2020;52(12):1271-81. doi: 10.1038/s41588-020-00736-4.
28. Matkar S, Thiel A, Hua X. Menin: a scaffold protein that controls gene expression and cell signaling. *Trends in Biochemical Sciences*. 2013;38(8):394-402. doi: <https://doi.org/10.1016/j.tibs.2013.05.005>.
29. Panne D. The enhanceosome. *Current Opinion in Structural Biology*. 2008;18(2):236-42. doi: <https://doi.org/10.1016/j.sbi.2007.12.002>.
30. Waddington CH. The Epigenotype. *International Journal of Epidemiology*. 2011;41(1):10-3. doi: 10.1093/ije/dyr184.
31. Waddington CH. Canalization of Development and Genetic Assimilation of Acquired Characters. *Nature*. 1959;183(4676):1654-5. doi: 10.1038/1831654a0.
32. Waddington CH. An introduction to modern genetics. *An introduction to modern genetics*. 1939.
33. Razin A, Cedar H. DNA methylation and gene expression. *Microbiological Reviews*. 1991;55(3):451-8. doi: doi:10.1128/mr.55.3.451-458.1991.
34. Weber M, Hellmann I, Stadler MB, Ramos L, Pääbo S, Rebhan M, et al. Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nature Genetics*. 2007;39(4):457-66. doi: 10.1038/ng1990.
35. Hervouet E, Vallette FM, Cartron P-F. Dnmt3/transcription factor interactions as crucial players in targeted DNA methylation. *Epigenetics*. 2009;4(7):487-99. doi: 10.4161/epi.4.7.9883.
36. Kass SU, Pruss D, Wolffe AP. How does DNA methylation repress transcription? *Trends in Genetics*. 1997;13(11):444-9. doi: [https://doi.org/10.1016/S0168-9525\(97\)01268-7](https://doi.org/10.1016/S0168-9525(97)01268-7).
37. Wade PA. Methyl CpG binding proteins: coupling chromatin architecture to gene regulation. *Oncogene*. 2001;20(24):3166-73. doi: 10.1038/sj.onc.1204340.
38. Allfrey VG, Faulkner R, Mirsky AE. ACETYLATION AND METHYLATION OF HISTONES AND THEIR POSSIBLE ROLE IN THE REGULATION OF RNA SYNTHESIS^{*}. *Proceedings of the National Academy of Sciences*. 1964;51(5):786-94. doi: doi:10.1073/pnas.51.5.786.
39. Cohen P. The structure and regulation of protein phosphatases. *Annual review of biochemistry*. 1989;58(1):453-508.
40. Byvoet P, Shepherd G, Hardin J, Noland B. The distribution and turnover of labeled methyl groups in histone fractions of cultured mammalian cells. *Archives of biochemistry and biophysics*. 1972;148(2):558-67.
41. Ng HH, Xu R-M, Zhang Y, Struhl K. Ubiquitination of histone H2B by Rad6 is required for efficient Dot1-mediated methylation of histone H3 lysine 79. *Journal of Biological Chemistry*. 2002;277(38):34655-7.

42. Shiio Y, Eisenman RN. Histone sumoylation is associated with transcriptional repression. *Proceedings of the National Academy of Sciences*. 2003;100(23):13225-30. doi: doi:10.1073/pnas.1735528100.
43. Placek BJ, Gloss LM. The N-Terminal Tails of the H2A–H2B Histones Affect Dimer Structure and Stability. *Biochemistry*. 2002;41(50):14960-8. doi: 10.1021/bi026283k.
44. Hake SB, Garcia BA, Kauer M, Baker SP, Shabanowitz J, Hunt DF, et al. Serine 31 phosphorylation of histone variant H3.3 is specific to regions bordering centromeres in metaphase chromosomes. *Proceedings of the National Academy of Sciences*. 2005;102(18):6344-9. doi: doi:10.1073/pnas.0502413102.
45. Hassig CA, Schreiber SL. Nuclear histone acetylases and deacetylases and transcriptional regulation: HATs off to HDACs. *Current Opinion in Chemical Biology*. 1997;1(3):300-8. doi: [https://doi.org/10.1016/S1367-5931\(97\)80066-X](https://doi.org/10.1016/S1367-5931(97)80066-X).
46. Clayton AL, Hebbes TR, Thorne AW, Crane-Robinson C. Histone acetylation and gene induction in human cells. *FEBS letters*. 1993;336(1):23-6.
47. Pogo B, Allfrey V, Mirsky A. RNA synthesis and histone acetylation during the course of gene activation in lymphocytes. *Proceedings of the National Academy of Sciences*. 1966;55(4):805-12.
48. Murray K. The occurrence of ϵ -N-methyl lysine in histones. *Biochemistry*. 1964;3(1):10-5.
49. Vakoc CR, Sachdeva MM, Wang H, Blobel GA. Profile of Histone Lysine Methylation across Transcribed Mammalian Chromatin. *Molecular and Cellular Biology*. 2006;26(24):9185-95. doi: 10.1128/MCB.01529-06.
50. North JA, Simon M, Ferdinand MB, Shoffner MA, Picking JW, Howard CJ, et al. Histone H3 phosphorylation near the nucleosome dyad alters chromatin structure. *Nucleic Acids Research*. 2014;42(8):4922-33. doi: 10.1093/nar/gku150.
51. Sauvé DM, Anderson HJ, Ray JM, James WM, Roberge M. Phosphorylation-induced Rearrangement of the Histone H3 NH₂-terminal Domain during Mitotic Chromosome Condensation. *Journal of Cell Biology*. 1999;145(2):225-35. doi: 10.1083/jcb.145.2.225.
52. Barratt MJ, Hazzalin CA, Cano E, Mahadevan LC. Mitogen-stimulated phosphorylation of histone H3 is targeted to a small hyperacetylation-sensitive fraction. *Proceedings of the National Academy of Sciences*. 1994;91(11):4781-5. doi: doi:10.1073/pnas.91.11.4781.
53. Sawicka A, Seiser C. Sensing core histone phosphorylation—A matter of perfect timing. *Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms*. 2014;1839(8):711-8.
54. Fischle W, Tseng BS, Dormann HL, Ueberheide BM, Garcia BA, Shabanowitz J, et al. Regulation of HP1–chromatin binding by histone H3 methylation and phosphorylation. *Nature*. 2005;438(7071):1116-22. doi: 10.1038/nature04219.

55. Martienssen R, Moazed D. RNAi and Heterochromatin Assembly. *Cold Spring Harbor Perspectives in Biology*. 2015;7(8). doi: 10.1101/cshperspect.a019323.
56. Chalker DL, Meyer E, Mochizuki K. Epigenetics of Ciliates. *Cold Spring Harbor Perspectives in Biology*. 2013;5(12). doi: 10.1101/cshperspect.a017764.
57. Hall IM, Shankaranarayana GD, Noma K-i, Ayoub N, Cohen A, Grewal SIS. Establishment and Maintenance of a Heterochromatin Domain. *Science*. 2002;297(5590):2232-7. doi: doi:10.1126/science.1076466.
58. Volpe TA, Kidner C, Hall IM, Teng G, Grewal SIS, Martienssen RA. Regulation of Heterochromatic Silencing and Histone H3 Lysine-9 Methylation by RNAi. *Science*. 2002;297(5588):1833-7. doi: doi:10.1126/science.1074973.
59. Taverna SD, Coyne RS, Allis CD. Methylation of Histone H3 at Lysine 9 Targets Programmed DNA Elimination in *Tetrahymena*. *Cell*. 2002;110(6):701-11. doi: 10.1016/S0092-8674(02)00941-8.
60. Mochizuki K, Fine NA, Fujisawa T, Gorovsky MA. Analysis of a piwi-Related Gene Implicates Small RNAs in Genome Rearrangement in *Tetrahymena*. *Cell*. 2002;110(6):689-99. doi: [https://doi.org/10.1016/S0092-8674\(02\)00909-1](https://doi.org/10.1016/S0092-8674(02)00909-1).
61. King M-C, Wilson AC. Evolution at Two Levels in Humans and Chimpanzees. *Science*. 1975;188(4184):107-16. doi: doi:10.1126/science.1090005.
62. Maurano MT, Humbert R, Rynes E, Thurman RE, Haugen E, Wang H, et al. Systematic Localization of Common Disease-Associated Variation in Regulatory DNA. *Science*. 2012;337(6099):1190-5. doi: doi:10.1126/science.1222794.
63. Rodriguez J, Larson DR. Transcription in Living Cells: Molecular Mechanisms of Bursting. *Annual Review of Biochemistry*. 2020;89(1):189-212. doi: 10.1146/annurev-biochem-011520-105250.
64. Small S, Blair A, Levine M. Regulation of even-skipped stripe 2 in the *Drosophila* embryo. *The EMBO Journal*. 1992;11(11):4047-57. doi: <https://doi.org/10.1002/j.1460-2075.1992.tb05498.x>.
65. Long HK, Osterwalder M, Welsh IC, Hansen K, Davies JOJ, Liu YE, et al. Loss of Extreme Long-Range Enhancers in Human Neural Crest Drives a Craniofacial Disorder. *Cell Stem Cell*. 2020;27(5):765-83.e14. doi: <https://doi.org/10.1016/j.stem.2020.09.001>.
66. Lin X, Liu Y, Liu S, Zhu X, Wu L, Zhu Y, et al. Nested epistasis enhancer networks for robust genome regulation. *Science*. 2022;377(6610):1077-85. doi: doi:10.1126/science.abk3512.
67. Patwardhan RP, Hiatt JB, Witten DM, Kim MJ, Smith RP, May D, et al. Massively parallel functional dissection of mammalian enhancers in vivo. *Nature Biotechnology*. 2012;30(3):265-70. doi: 10.1038/nbt.2136.
68. Buecker C, Wysocka J. Enhancers as information integration hubs in development: lessons from genomics. *Trends in Genetics*. 2012;28(6):276-84. doi: <https://doi.org/10.1016/j.tig.2012.02.008>.

69. Lettice LA, Horikoshi T, Heaney SJH, van Baren MJ, van der Linde HC, Breedveld GJ, et al. Disruption of a long-range cis-acting regulator for *Shh* causes preaxial polydactyly. *Proceedings of the National Academy of Sciences*. 2002;99(11):7548-53. doi: 10.1073/pnas.112212199.
70. Moreau P, Hen R, Wasylyk B, Everett R, Gaub M, Chambon P. The SV40 72 base repair repeat has a striking effect on gene expression both in SV40 and other chimeric recombinants. *Nucleic acids research*. 1981;9(22):6047-68.
71. Mercola M, Wang X-F, Olsen J, Calame K. Transcriptional enhancer elements in the mouse immunoglobulin heavy chain locus. *Science*. 1983;221(4611):663-5.
72. Banerji J, Olson L, Schaffner W. A lymphocyte-specific cellular enhancer is located downstream of the joining region in immunoglobulin heavy chain genes. *Cell*. 1983;33(3):729-40.
73. Gillies SD, Morrison SL, Oi VT, Tonegawa S. A tissue-specific transcription enhancer element is located in the major intron of a rearranged immunoglobulin heavy chain gene. *Cell*. 1983;33(3):717-28.
74. Tuan D, Solomon W, Li Q, London IM. The "beta-like-globin" gene domain in human erythroid cells. *Proceedings of the National Academy of Sciences*. 1985;82(19):6384-8.
75. Gross DS, Garrard WT. Nuclease hypersensitive sites in chromatin. *Annual review of biochemistry*. 1988;57(1):159-97.
76. Hebbes T, Clayton A, Thorne A, Crane-Robinson C. Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. *The EMBO journal*. 1994;13(8):1823-30.
77. Hebbes TR, Thorne AW, Crane-Robinson C. A direct link between core histone acetylation and transcriptionally active chromatin. *The EMBO journal*. 1988;7(5):1395-402.
78. Lee DY, Hayes JJ, Pruss D, Wolffe AP. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. *Cell*. 1993;72(1):73-84.
79. Serfling E, Jasin M, Schaffner W. Enhancers and eukaryotic gene transcription. *Trends in Genetics*. 1985;1:224-30.
80. Ikuta T, Kan YW. In vivo protein-DNA interactions at the beta-globin gene locus. *Proceedings of the National Academy of Sciences*. 1991;88(22):10188-92.
81. Forsberg M, Westin G. Enhancer activation by a single type of transcription factor shows cell type dependence. *The EMBO Journal*. 1991;10(9):2543-51.
82. Müller H-P, Sogo J, Schaffner W. An enhancer stimulates transcription in trans when attached to the promoter via a protein bridge. *Cell*. 1989;58(4):767-77.
83. Spitz F, Furlong EE. Transcription factors: from enhancer binding to developmental control. *Nature reviews genetics*. 2012;13(9):613-26.
84. Whyte Warren A, Orlando David A, Hnisz D, Abraham Brian J, Lin Charles Y, Kagey Michael H, et al. Master Transcription Factors and Mediator Establish Super-

- Enhancers at Key Cell Identity Genes. *Cell*. 2013;153(2):307-19. doi: <https://doi.org/10.1016/j.cell.2013.03.035>.
85. Pott S, Lieb JD. What are super-enhancers? *Nature Genetics*. 2015;47(1):8-12. doi: 10.1038/ng.3167.
86. Sur I, Taipale J. The role of enhancers in cancer. *Nature Reviews Cancer*. 2016;16(8):483-93.
87. Kioussis D, Vanin E, deLange T, Flavell RA, Grosveld FG. β -Globin gene inactivation by DNA translocation in $\gamma\beta$ -thalassaemi. *Nature*. 1983;306(5944):662-6. doi: 10.1038/306662a0.
88. Northcott PA, Lee C, Zichner T, Stütz AM, Erkek S, Kawauchi D, et al. Enhancer hijacking activates GFI1 family oncogenes in medulloblastoma. *Nature*. 2014;511(7510):428-34. doi: 10.1038/nature13379.
89. Abraham BJ, Hnisz D, Weintraub AS, Kwiatkowski N, Li CH, Li Z, et al. Small genomic insertions form enhancers that misregulate oncogenes. *Nature Communications*. 2017;8. doi: 10.1038/ncomms14385.
90. Zhang Y, Li X, Liang D, Xu F, Liang S, Zhu X, et al. Heavy Chain Deposition Disease: Clinicopathologic Characteristics of a Chinese Case Series. *American Journal of Kidney Diseases*. 2020;75(5):736-43. doi: <https://doi.org/10.1053/j.ajkd.2019.08.013>.
91. Lupiáñez Darío G, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, et al. Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. *Cell*. 2015;161(5):1012-25. doi: <https://doi.org/10.1016/j.cell.2015.04.004>.
92. Carullo NVN, Day JJ. Genomic Enhancers in Brain Health and Disease. *Genes*. 2019;10(1):43.
93. Wang Q, Carroll JS, Brown M. Spatial and temporal recruitment of androgen receptor and its coactivators involves chromosomal looping and polymerase tracking. *Molecular cell*. 2005;19(5):631-42.
94. Kong S, Bohl D, Li C, Tuan D. Transcription of the HS2 enhancer toward a cis-linked gene is independent of the orientation, position, and distance of the enhancer relative to the gene. *Molecular and cellular biology*. 1997;17(7):3955-65.
95. Tchurikov NA, Kretova OV, Moiseeva ED, Sosin DV. Evidence for RNA synthesis in the intergenic region between enhancer and promoter and its inhibition by insulators in *Drosophila melanogaster*. *Nucleic acids research*. 2009;37(1):111-22.
96. Zhu X, Ling J, Zhang L, Pi W, Wu M, Tuan D. A facilitated tracking and transcription mechanism of long-range enhancer function. *Nucleic acids research*. 2007;35(16):5532-44.
97. Lomvardas S, Barnea G, Pisapia DJ, Mendelsohn M, Kirkland J, Axel R. Interchromosomal interactions and olfactory receptor choice. *Cell*. 2006;126(2):403-13.

98. Morcillo P, Rosen C, Baylies MK, Dorsett D. Chip, a widely expressed chromosomal protein required for segmentation and activity of a remote wing margin enhancer in *Drosophila*. *Genes & Development*. 1997;11(20):2729-40.
99. Deng W, Lee J, Wang H, Miller J, Reik A, Gregory PD, et al. Controlling long-range genomic interactions at a native locus by targeted tethering of a looping factor. *Cell*. 2012;149(6):1233-44.
100. Wansink DG, Schul W, van der Kraan I, van Steensel B, van Driel R, de Jong L. Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *J Cell Biol*. 1993;122(2):283-93. doi: 10.1083/jcb.122.2.283.
101. Cisse II, Izeddin I, Causse SZ, Boudarene L, Senecal A, Muresan L, et al. Real-time dynamics of RNA polymerase II clustering in live human cells. *Science (New York, NY)*. 2013;341(6146):664-7. doi: 10.1126/science.1239053.
102. Ghamari A, van de Corput MPC, Thongjuea S, van Cappellen WA, van IJcken W, van Haren J, et al. In vivo live imaging of RNA polymerase II transcription factories in primary cells. *Genes & Development*. 2013;27(7):767-77. doi: 10.1101/gad.216200.113.
103. Dunn TM, Hahn S, Ogden S, Schleif RF. An operator at-280 base pairs that is required for repression of araBAD operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. *Proceedings of the National Academy of Sciences*. 1984;81(16):5017-20.
104. Rao SSP, Huang S-C, Glenn St Hilaire B, Engreitz JM, Perez EM, Kieffer-Kwon K-R, et al. Cohesin Loss Eliminates All Loop Domains. *Cell*. 2017;171(2):305-20.e24. doi: 10.1016/j.cell.2017.09.026.
105. Wutz G, Várnai C, Nagasaka K, Cisneros DA, Stocsits RR, Tang W, et al. Topologically associating domains and chromatin loops depend on cohesin and are regulated by CTCF, WAPL, and PDS5 proteins. *The EMBO Journal*. 2017;36(24):3573-99. doi: <https://doi.org/10.15252/emboj.201798004>.
106. Haarhuis JH, van der Weide RH, Blomen VA, Yáñez-Cuna JO, Amendola M, van Ruiten MS, et al. The cohesin release factor WAPL restricts chromatin loop extension. *Cell*. 2017;169(4):693-707. e14.
107. de Wit E, Vos ES, Holwerda SJ, Valdes-Quezada C, Verstegen MJ, Teunissen H, et al. CTCF binding polarity determines chromatin looping. *Molecular cell*. 2015;60(4):676-84.
108. Sanborn AL, Rao SS, Huang S-C, Durand NC, Huntley MH, Jewett AI, et al. Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. *Proceedings of the National Academy of Sciences*. 2015;112(47):E6456-E65.
109. Wingett S, Ewels P, Furlan-Magaril M, Nagano T, Schoenfelder S, Fraser P, et al. HiCUP: pipeline for mapping and processing Hi-C data [version 1; peer review: 2

- approved, 1 approved with reservations]. *F1000Research*. 2015;4(1310). doi: 10.12688/f1000research.7334.1.
110. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nature Methods*. 2012;9(4):357-9. doi: 10.1038/nmeth.1923.
111. Cairns J, Freire-Pritchett P, Wingett SW, Várnai C, Dimond A, Plagnol V, et al. CHiCAGO: robust detection of DNA looping interactions in Capture Hi-C data. *Genome Biology*. 2016;17(1):127. doi: 10.1186/s13059-016-0992-2.
112. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. 2014;15(12):550. doi: 10.1186/s13059-014-0550-8.
113. Zhou HY, Katsman Y, Dhaliwal NK, Davidson S, Macpherson NN, Sakthidevi M, et al. A Sox2 distal enhancer cluster regulates embryonic stem cell differentiation potential. *Genes & Development*. 2014;28(24):2699-711. doi: 10.1101/gad.248526.114.
114. Huletsky A, De Murcia G, Muller S, Hengartner M, Menard L, Lamarre D, et al. The effect of poly (ADP-ribosyl) ation on native and H1-depleted chromatin: a role of poly (ADP-ribosyl) ation on core nucleosome structure. *Journal of Biological Chemistry*. 1989;264(15):8878-86.
115. Poirier GG, De Murcia G, Jongstra-Bilen J, Niedergang C, Mandel P. Poly (ADP-ribosyl) ation of polynucleosomes causes relaxation of chromatin structure. *Proceedings of the National Academy of Sciences*. 1982;79(11):3423-7.
116. Ciccarone F, Zampieri M, Caiafa P. PARP1 orchestrates epigenetic events setting up chromatin domains. *Seminars in Cell & Developmental Biology*. 2017;63:123-34. doi: <https://doi.org/10.1016/j.semcdb.2016.11.010>.
117. Hsieh T-HS, Cattoglio C, Slobodyanyuk E, Hansen AS, Darzacq X, Tjian R. Enhancer–promoter interactions and transcription are largely maintained upon acute loss of CTCF, cohesin, WAPL or YY1. *Nature Genetics*. 2022;54(12):1919-32. doi: 10.1038/s41588-022-01223-8.
118. Abranches E, Silva M, Pradier L, Schulz H, Hummel O, Henrique D, et al. Neural Differentiation of Embryonic Stem Cells In Vitro: A Road Map to Neurogenesis in the Embryo. *PLOS ONE*. 2009;4(7):e6286. doi: 10.1371/journal.pone.0006286.
119. Calderon L, Weiss FD, Beagan JA, Oliveira MS, Georgieva R, Wang Y-F, et al. Cohesin-dependence of neuronal gene expression relates to chromatin loop length. *eLife*. 2022;11:e76539. doi: 10.7554/eLife.76539.