

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

Завідувач кафедри молекулярної біотехнології
та біоінформатики
доцент Нипорко Олексій Юрійович
Протокол №__ засідання кафедри
від «__»_____2022 р.

Потенційні маркери для раннього виявлення раку голови та шиї
методом рідинної біопсії на основі патернів метилювання ДНК із
застосуванням технологій Oxford Nanopore

Випускна кваліфікаційна робота бакалавра
студентки спеціальності 091 «Біологія»

ОП «Біологія (високі технології)»

Мамойлик Вікторії Олегівни

Науковий керівник:

доцент кафедри молекулярної

біотехнології та біоінформатики

к.б.н. **Нипорко Олексій Юрійович**

Оцінка захисту роботи

Київ – 2022

РЕФЕРАТ

Дипломна робота бакалавра: сторінок 33, рисунків 13, джерел 34, сторінок додатків 3.

Мамоєлик В.О. Потенційні маркери для раннього виявлення раку голови та шиї методом рідинної біопсії на основі порівняння метилювання ДНК з застосуванням технологій Oxford Nanopore. — Випускна кваліфікаційна робота бакалавра за спеціальністю 091 Біологія ОП «Біологія (високі технології)».

У даній роботі було проведено секвенування зразків слини здорових людей та оцінено можливість застосування Oxford Nanopore для детекції метилювання ДНК як біомаркера.

Було виявлено, що технології Oxford Nanopore є ефективним методом для секвенування фрагментів позаклітинної ДНК та являє собою надійну альтернативу секвенуванню Illumina.

Отримані результати можуть бути використані для розкриття повного потенціалу рідинної біопсії як для дослідницьких, так і для клінічних цілей.

Ключові слова: рак голови та шиї, рідинна біопсія, метилювання, циркулююча пухлинна ДНК, позаклітинна ДНК, Oxford Nanopore.

ЗМІСТ

Вступ.....	4
Розділ 1. Огляд літератури	6
1.1. Рідинна біопсія на основі слини при раку голови та шиї... ..	6
1.2. Циркулюючі нуклеїнові кислоти пухлин	7
1.2.1. Циркулююча пухлинна ДНК	9
1.2.2 Позаклітинна пухлинна ДНК	11
1.2.3. Пухлинна РНК	12
1.3. Метилювання цпДНК в клінічній медицині	13
1.4. Моніторинг реакції на лікування	14
1.4.1. Прогнозування метастазів	16
1.5. Методи на основі перетворення бісульфіту	17
1.6. Рання діагностика.....	18
1.7. Точне виявлення циркулюючої пухлинної ДНК за допомогою консенсусного секвенування Nanopore	20
Розділ 2. Методи та матеріали досліджень	24
2.1 Матеріали дослідження	24
2.2 Протокол очищення позаклітинної ДНК	24
2.3 Контроль якості ДНК після очищення.....	25
2.4 Очищення за допомогою гелю	25
2.5 Підготовка бібліотеки та секвенування	27
РОЗДІЛ 3. Результати дослідження.....	29
3.1 Результати секвенування	29
3.2 Метилювання позаклітинної ДНК на хромосомах	30
РОЗДІЛ 4. Обговорення результатів дослідження	33
ВИСНОВКИ.....	35
Список використаних джерел	36
Додатки.....	42

ВСТУП

Рак голови та шиї (РГШ) залишається однією з основних причин смертності у всьому світі, через таку сукупність факторів як пізня стадія діагностики пухлини, локорегіональна агресивність і віддалені метастази. Рак голови та шиї є шостим за поширеністю раком у світі, пов'язаним з етіологічними факторами, такі як вживання тютюну та алкоголю, інфекції ВПЛ та певною мірою генетична схильність. Сучасний стандартизований діагностичний метод для РГШ — тканина біопсія, яка не дозволяє визначити повну динаміку розвитку пухлини. На сьогоднішній день біопсія тканин є найбільш широко використовуваним методом діагностики; однак ця техніка є інвазійною, досить складною, болючою, трудомісткою і потенційно небезпечною для пацієнта. Крім того, внутрішньопухлинна і метастатична гетерогенність залишається не виявленою, що впливає на конкретність, чутливість і точність оцінки.

Завдяки легкості відбору матеріалу, останні дослідження були зосереджені на використанні рідинної біопсії на основі слини. Збір слини є зручним, неінвазивним, безболісним методом та пропонує можливість забору повторюваних проб для моніторингу захворювання в реальному часі. Крім того, дозволяє виявлення, виділення та аналіз пухлин-похідних компонентів, таких як циркулюючі пухлинні нуклеїнові кислоти, позаклітинні везикули, циркулюючі пухлинні клітини та метаболіти. Очікується, що рідинна біопсія продемонструє високий рівень точності з точки зору представлення геному пухлин і мутацій.

Актуальність цієї роботи полягає в тому, що циркуляційні біомаркери мають широкий спектр застосування в клінічних умовах, але ще немає підтверджених даних для їх використання в клінічній практиці для діагностування РГШ. Покращення технології виявлення та секвенування нового покоління вирішили б багато проблем, що дасть можливість широкого спектру застосування рідкої біопсії на основі слини для клінічної практики.

Метою роботи є пошук потенційних маркерів для виявлення РГШ на основі порівняння метилювання ДНК у вибраних генах. Відповідно до мети, ми поставили такі задачі:

1. Ознайомитися з проведеними дослідженнями на цю тему та актуалізувати дані.
2. Провести забір слини здорових та хворих на РГШ людей.
3. Ізольовати ДНК відповідно до протоколу.
4. Підготувати бібліотеки для секвенування Nanopore.
5. Проаналізувати отримані результати.
6. Оцінити можливість застосування Oxford Nanopore для детекції метилювання ДНК як біомаркера для виявлення РГШ.

РОЗДІЛ 1. Огляд літератури

1.1 Рідинна біопсія на основі слини при раку голови та шиї

Рак голови та шиї займає шосте місце серед злоякісних пухлин у світі. Переважаючим гістологічним типом є плоскоклітинний рак, який в основному виникає в порожнині рота, ротоглотці і гортані [1]. Незважаючи на прогрес у діагностиці та терапії, загальне виживання пацієнтів з РГШ залишається низьким протягом чотирьох десятиліть. Звичайні діагностичні стратегії включають фізичне обстеження, методи візуалізації, такі як комп'ютерна томографія (КТ), ультразвукове дослідження (УЗД), магнітно-резонансна томографія (МРТ) та біопсія тканини з подальшим гістопатологічним аналізом [2].

Слина як потенційне джерело для рідкої біопсії пацієнтів з РГШ має ряд переваг порівняно з іншими рідинами організму, оскільки відображає будь-які геномні, епігеномні, протеомні та фізіологічні/патологічні зміни в ротовій порожнині, гортані та глотці; служить неінвазивним, недорогим, простішим і доступнішим інструментом скринінгу; і надає можливість моніторингу пацієнтів з РГШ в режимі реального часу завдяки гнучкості повторного відбору проб і більших обсягів для обстеження без потреби підготовленого медичного персоналу для збору [3].

Незважаючи на потенційну цінність використання біомаркерів, отриманих зі слини, як діагностичного інструменту, його клінічна корисність обмежена через деякі проблеми. Насамперед, слина містить різні непухлинні компоненти, які перешкоджають здатності виявляти біомолекули пухлинного походження. Більше того, відносний внесок різних субкомпонентів у середовище слини ускладнює ідентифікацію специфічних маркерів РГШ [4]. Однак потенційна корисність слини як інструменту рідинної біопсії для діагностики, прогнозу та терапевтичного моніторингу РГШ широко досліджується. У даний час найпоширенішими компонентами для рідинної

біопсії РГШ є позаклітинні нуклеїнові кислоти пухлин (ДНК, мРНК і мікроРНК), позаклітинні везикули, циркулюючі пухлинні клітини і метаболіти слини (Рис. 1.1) [5].

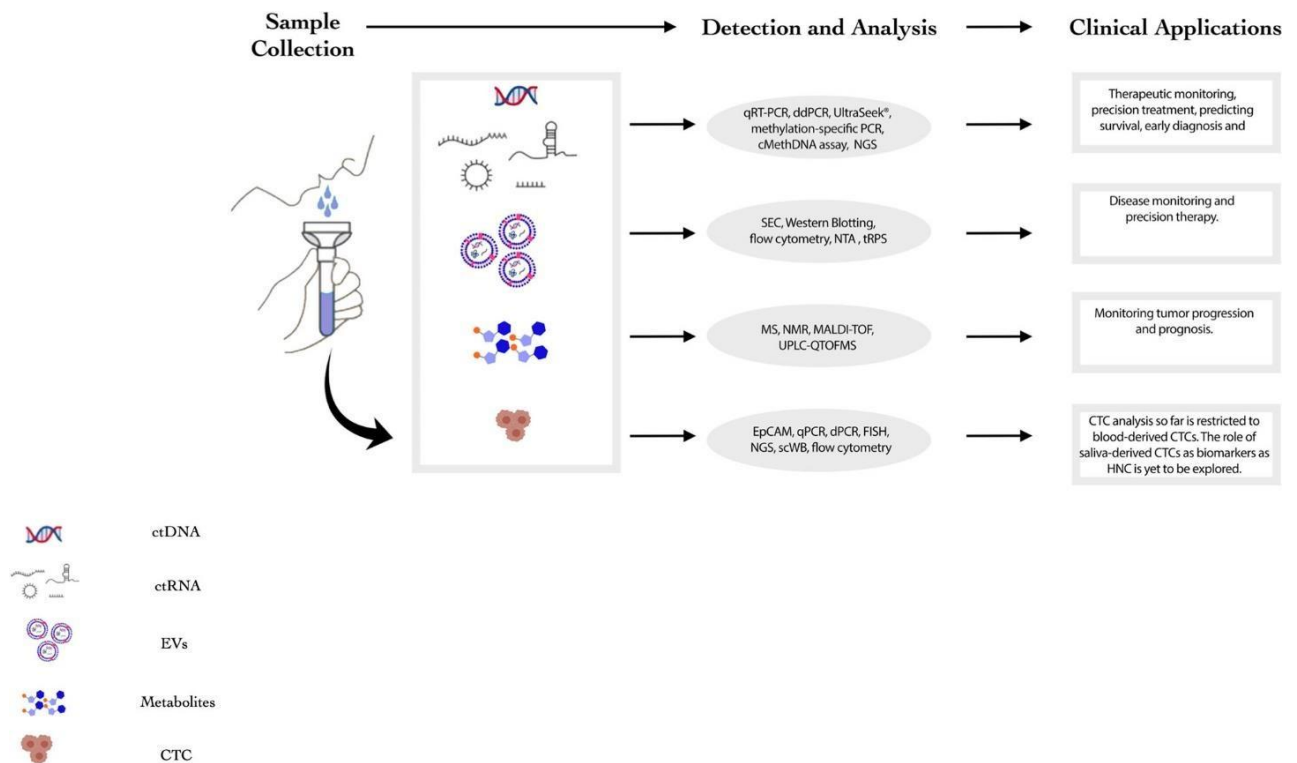


Рис. 1.1. Компоненти слини, які потенційно можуть діяти як біомаркери для РГШ. Методи виявлення та аналізу для кожного компонента наведені на малюнку.

1.2 Циркулюючі нуклеїнові кислоти пухлин

Циркулюючі пухлинні нуклеїнові кислоти (цпНК) — це фрагменти позаклітинної геномної/вірусної ДНК та/або РНК, які викидаються пухлинними клітинами через пасивні процеси, такі як некроз і апоптоз, або за допомогою активних механізмів, таких як спонтанне вивільнення нуклеїнових кислот у систему кровообігу. Ці фрагменти циркулюючої пухлинної ДНК і РНК знаходяться в різних рідинах організму, включаючи слину. Вони відображають генетичну інформацію про пухлину, відображають клональну гетерогенність та еволюцію пухлини [6, 7]. Швидкість, з якою ці циркулюючі

нуклеїнові кислоти вивільняються в кров, залежить від локалізації пухлини, васкуляризації, масштабу, що призводить до варіабельності у пацієнтів. Аналіз цДНК більше покладений на ідентифікацію та націлений на певні мутації, специфічні для пухлини і на розуміння епігенетичного ландшафту, тоді як цРНК акцентує увагу на ідентифікації нових або диференціальних моделей експресії інформаційної РНК (мРНК), мікроРНК (міРНК) як потенційних біомаркер слини. Біомаркери на основі цРНК, ймовірно, дають краще динамічне уявлення про стан клітини та регуляцію порівняно з біомаркерами цДНК.

Виявлення та аналіз цНК є досить складним завданням. На даний момент ПЛР у реальному часі (qRT-PCR), цифрова крапельна ПЛР (ddPCR) і метод ПЛР на основі мас-спектрометрії UltraSeek® (Agena Bioscience) є найбільш широко використовуваними методами, оскільки вони допомагають оптимізувати зразки з низькою концентрацією цНК при РГШ [8]. ddPCR все ще є найбільш переважаючим методом, який демонструє більш високу чутливість, специфічність і здатність до мультиплексування. Крім того, такі методи, як специфічна для метилювання ПЛР, аналіз сMethDNA використовуються для виявлення різниці в моделях метилювання на промоторі цДНК у зразках пацієнтів з РГШ.

Методи, засновані на ПЛР, є кращими, коли існує низька кількість цільових ділянок (≤ 20 мішеней), обмежена кількість введення зразка та обмежена оцінка гетерогенності пухлини або ідентифікація відомих варіантів. Методи секвенування наступного покоління (NGS), такі як CAPP-Seq (персоналізоване профілювання раку шляхом глибокого секвенування), TAm-Seq (глибоке секвенування з тегамі ампліконів), Safe-Seq (система безпечного секвенування) і AmpliSeq, використовуються для виділення та захоплення цНК. Ці методи NGS можуть виявляти як відомі, так і невідомі пухлиноспецифічні мутації та аналізувати моделі диференціальної експресії окремих маркерів або групи маркерів. Цільові методи NGS є менш трудомісткими, призводять до меншої втрати ресурсів і пропонують більш

високий рівень відкриття, таким чином допомагаючи ідентифікувати нові варіанти. Незважаючи на поточні обмеження, ці методи продемонстрували потенціал для виявлення та виділення менших концентрацій цпДНК зі слини, відкриваючи таким чином нові шляхи для клінічного застосування. Завдяки технологічному прогресу вища специфічність і чутливість виявлення цпДНК можуть ефективно розширити їх клінічне застосування. Тим не менш, розробка економічно ефективних аналізів NGS має вирішальне значення для їх широкого клінічного застосування [9].

1.2.1 Циркулююча пухлинна ДНК

Циркулююча пухлинна ДНК (цпДНК) являє собою тривіальну фракцію (<1%) повної пухлинної ДНК, що викидається з пухлинних клітин в кровообіг. Однак вважається, що ця невелика субпопуляція відображає соматичні мутації та геномний ландшафт первинних пухлин, що може бути корисним для ранньої діагностики та прогнозування ризику РГШ. Останнім часом кілька досліджень підкреслюють використання цпДНК, отриманої зі слини, для раннього виявлення раку. Було проведено комплексний аналіз соматичних мутацій (TP53, CDKN2A, NRAS, NOTCH1, PIK3CA, FBXW7 і HRAS) у слині та плазмі 93 зразків пацієнтів з РГШ. Дослідження продемонструвало виявлення цпДНК з частотою 96% незалежно від розміру, стадії та розташування пухлини. Крім того, у більшості пацієнтів із зазначеними соматичними мутаціями спостерігався рецидив після операції. У пацієнтів з РГШ швидкість виявлення цпДНК була вищою в слині порівняно з плазмою, що вказує на те, що цпДНК слини можна використовувати для виявлення РГШ [10].

Аналогічно, мутація p53 в екзоні 4 кодону 63 була виявлена в слині ранньої стадії РГШ. Однак було повідомлено, що підхід цільового виявлення мутацій не зміг продемонструвати значну узгодженість у виявленні мутацій

TP53 від цпДНК, отриманої з пухлини та слини. Було виявлено мікросателітну нестабільність та втрату гетерозиготності певних хромосом у клітинах слизової оболонки порожнини рота у пацієнтів з РГШ та слини ($p < 0,001$) з різною чутливістю та специфічністю на основі розміру зразка та підсайтів вибірки. Більше того, генетичні зміни в генах RMAIP1 і RTPN1 могли відрізнити пацієнтів з РГШ від здорових людей [11].

У сукупності ці дослідження свідчать про те, що оцінка соматичних мутацій цпДНК слини може бути ефективною неінвазивною заміною біопсії тканини для ранньої діагностики, спостереження за захворюванням та прогнозу пацієнтів з РГШ. Однак для клінічного застосування цього підходу знадобляться численні дослідження на основі виявлення мутацій зі стандартизованими протоколами та більшою когортою пацієнтів. Низький вихід цпДНК після очищення від слини також є ключовим обмежуючим фактором. Щоб підвищити ефективність та чутливість цпДНК слини як біомаркера, необхідно розробити конкретні панелі цпДНК, які можуть допомогти виявляти та відстежувати випадки РГШ в режимі реального часу та економічно ефективним способом.

Кілька досліджень продемонстрували значну клінічну кореляцію між гіперметилуванням у промоторній області цпДНК слини у пацієнтів з РГШ. Зокрема, локуси метильованого гена були ідентифіковані як у тканинах, так і в передопераційних зразках слини, і вони могли б служити класифікатором для диференціації між передопераційними та післяопераційними зразками для пацієнтів із РГШ. Аналогічно цьому показано, що виявлення промоторного гіперметилування одного або всіх генів (TIMP-3, CCNA1, DCC, MGMT, MINT-31, DAPK p16) у попередньо обробленій ДНК слини може ефективно передбачити погану виживаність та рецидив пацієнтів з РГШ. Ці результати свідчать про те, що моделі гіперметилування цпДНК мають потенціал передбачити агресивність захворювання, загальну виживаність та терапевтичний моніторинг та нагляд за пацієнтами з РГШ.

1.2.2 Позаклітинна пухлинна ДНК

Позаклітинна ДНК вивільняється з клітин у кровоносну систему по всьому тілу. Вперше його виявили Мандел і Метаїс у 1948 році. Позаклітинна ДНК може бути знайдена в плазмі та інших рідинах організму, таких як спинномозкова рідина, плевральна рідина, сеча, слина та ін. У певних фізіологічних або патологічних захворюваннях, такі як вагітність, трансплантація органів та рак, пов'язані/уражені тканини можуть вивільнити додаткову ДНК у периферичний кровообіг. Таким чином, виявлення позаклітинної ДНК в периферичній крові може виявити аномалії індивідів неінвазивним способом [12].

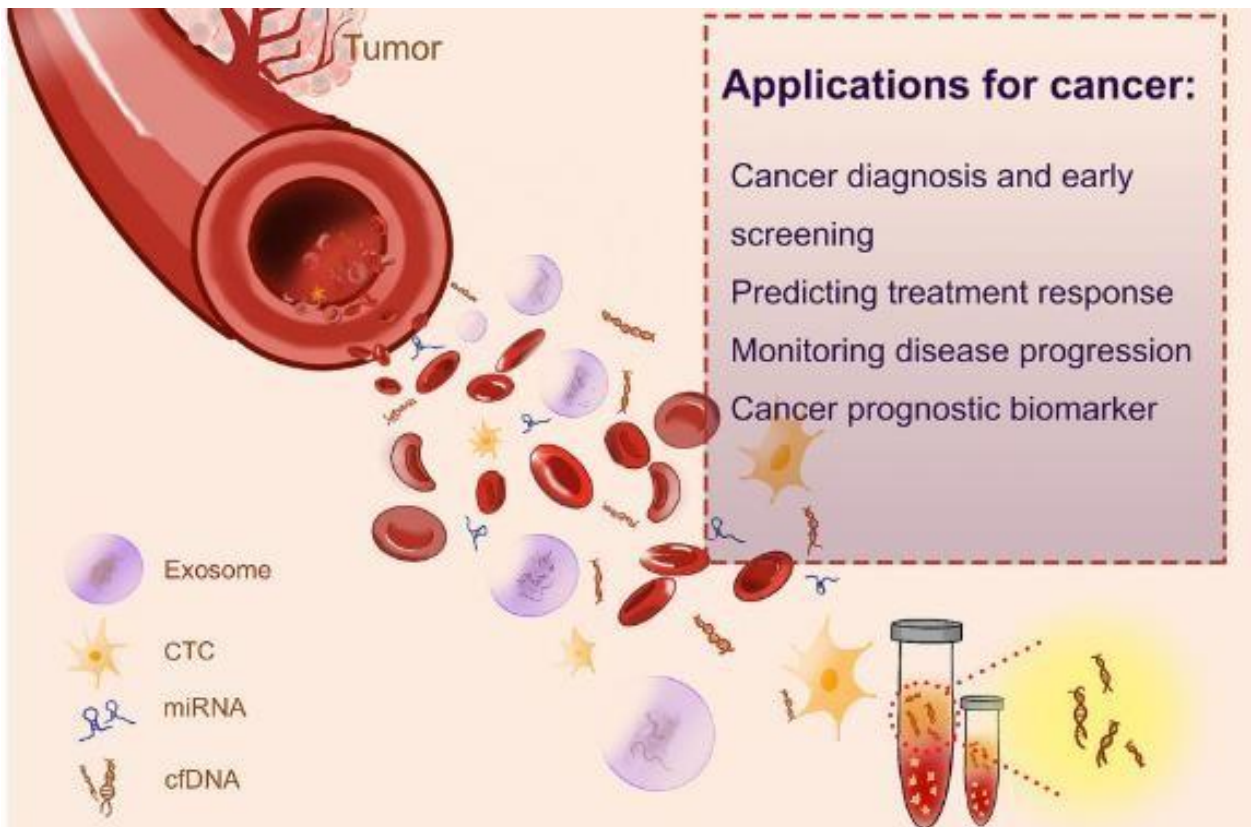


Рис. 1.2. Компоненти пухлин, які вивільняються в кров.

Хворі на рак зазвичай мають високий рівень позаклітинної ДНК у сироватці або плазмі в результаті клітинного некрозу або апоптозу, тому що пухлинні клітини діляться швидше, ніж звичайні клітини, і позаклітинна ДНК вивільняється у високій пропорції.

1.2.3 Циркулююча пухлинна РНК

Аналіз транскриптомних профілів циркулюючих рідин організму є широко досліджуваним методом для раннього виявлення раку, і кілька досліджень показали значний зв'язок транскриптому з прогресуванням захворювання. Кілька досліджень продемонстрували зв'язок мРНК слини з розвитком і виявленням РГШ. Девід Елашофф і його колеги обґрунтували ефективність біомаркерів у більшій когорті пацієнтів (382 пацієнти), припустивши потенційну роль маркерів мРНК слини у виявленні раку. Що стосується показників індивідуальних маркерів у п'яти когортах, збільшення IL8 та SAT було статистично значущим ($p < 0,02$). Перевірка цих біомаркерів у більших когортах пацієнтів показує їх доцільність у розрізненні РГШ від здорових контрольних. Рівні експресії генів слини відповідали пухлинній тканині та пов'язували із загальним виживанням ($p = 0,011$) пацієнтів, демонструючи його величезний потенціал як багатообіцяючого біомаркера та незалежного прогнозу РГШ. ВПЛ-16 також був ідентифікований як основний етіологічний фактор, що відповідає за пухлинний генез РГШ. мРНК ВПЛ-16 показала суттєво змінену експресію в слині пацієнтів з РГШ з одночасним впливом на p16(INK4a), відомий пухлинний супресор, який має життєво важливу роль у регуляції клітинного циклу ($p < 0,05$). Таким чином, картина експресії різних мРНК слини корелює з різними важливими клінічними параметрами, включаючи прогресування пухлини, диференціювання та загальне виживання. Що ще важливіше, експресія мРНК слини зображує незалежний фактор прогнозу для РГШ, що припускає, що мРНК слини може бути потенційним біомаркером для раннього виявлення РГШ і створення прогнозу для пацієнтів з РГШ [13].

Інша частина цпРНК — це довгі некодуючі РНК (lncRNA), які мають довжину приблизно більше 200 нуклеотидів і не транслюються в білок. Враховуючи їх неминучу роль у прогресуванні пухлини та метастазуванні, сигнатури lncRNA, отримані із слини, були досліджені як ймовірні біомаркери

для моніторингу прогресування захворювання РГШ. У пілотному дослідженні повідомлялося про вимірювані рівню lncRNA HOTAIR і MALAT у слині пацієнтів з РГШ. Крім того, підвищені рівні були пов'язані з вузловими метастазами, що з'ясували його потенціал як прогностичного маркеру.

1.3 Метилування цДНК в клінічній медицині

Метилування ДНК є однією з найкраще вивчених епігенетичних модифікацій, які відбуваються, коли метильна група додається до позиції вуглецю-5 цитозину за допомогою ДНК-метилтрансфераз, відбувається переважно в динуклеотидах CpG. Все більше дослідів показують, що аномальне метилування ДНК сприяє пухлиногенезу та прогресуванню пухлини, головним чином через глобальне гіпометилування, вогнищеве гіперметилування в багатьох геномних областях (переважно в CpG-сайтах) і прямий мутагенез на метилованих цитозинах. У людини близько 28 мільйонів CpG у геномі, 60–80% якого, як правило, метильовані. На відміну від сильно індивідуалізованого і гетерогенного характеру генних мутацій, пухлинні клітини одного типу мають тенденцію до прояву подібних змін метилування ДНК. Крім того, моделі метилування позаклітинної ДНК узгоджуються з клітинами або тканинами, з яких вони походять, що означає, що виявлення аберацій специфічного для пухлини метилування ДНК у плазмі пацієнта може бути підходом до розробки аналізу крові для ранньої діагностики та прогнозу раку [14].

Існує кілька потенційних переваг оцінки метилування ДНК порівняно з іншими маркерами. По-перше, метилування ДНК відбувається на ранніх стадіях пухлинного генезу і може бути тканинним і специфічним для кожного типу раку. По-друге, метилування ДНК при раку має велику кількість геномних ділянок, що дозволяє використовувати декілька динуклеотидів CpG для виявлення. Іншою ключовою перевагою є те, що моделі метилування

часто асоціюються з специфічними видами раку і використовуються для детекції тканин походження раку.

Важлива віха в області аналізу метилювання цпДНК була схвалена FDA, це був Epi proColon тест у 2016 році [15], скринінговий тест на колоректальний рак, який базується на виявленні метилювання промоторної області SEPT9. Накопичені докази також підтвердили цінність метилювання цпДНК як біомаркерів для скринінгу, ранньої діагностики, моніторингу відповіді на терапію та прогнозу раку.

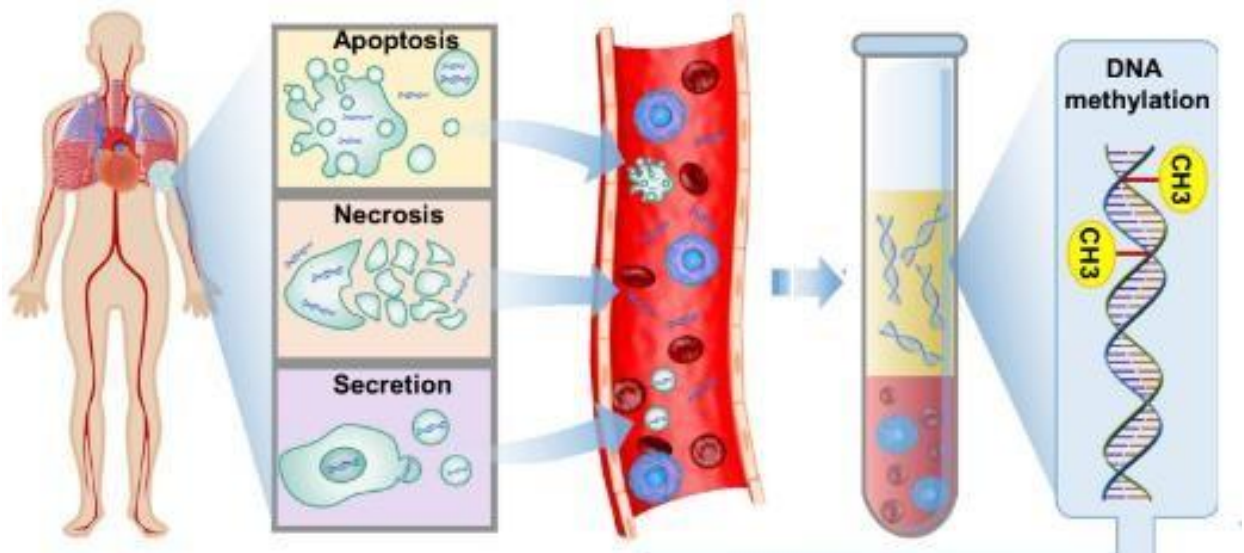


Рис. 1.3. Шляхи потрапляння компонентів пухлин в кров та приклад метилювання ДНК.

1.4 Моніторинг реакції на лікування

цпДНК можна використовувати для моніторингу відповіді на лікування раку. У порівнянні, цпДНК пропонує діагностичну перевагу для моніторингу реакції на лікування в режимі реального часу [16]. Недавнє дослідження повідомило про ранній стрибок в плазмі рівня цпДНК (збільшення мутованої ДНК BRAF) при меланомі у пацієнтів із імунотерапією з перенесенням Т-клітин, що відображає транс-загибель пухлинних клітин [17]. Інше

дослідження виявило зниження рівня цпДНК (мутації EGFR) при раку легенів у пацієнтів після терапії інгібіторами тирозинкінази (ІТК) [18].

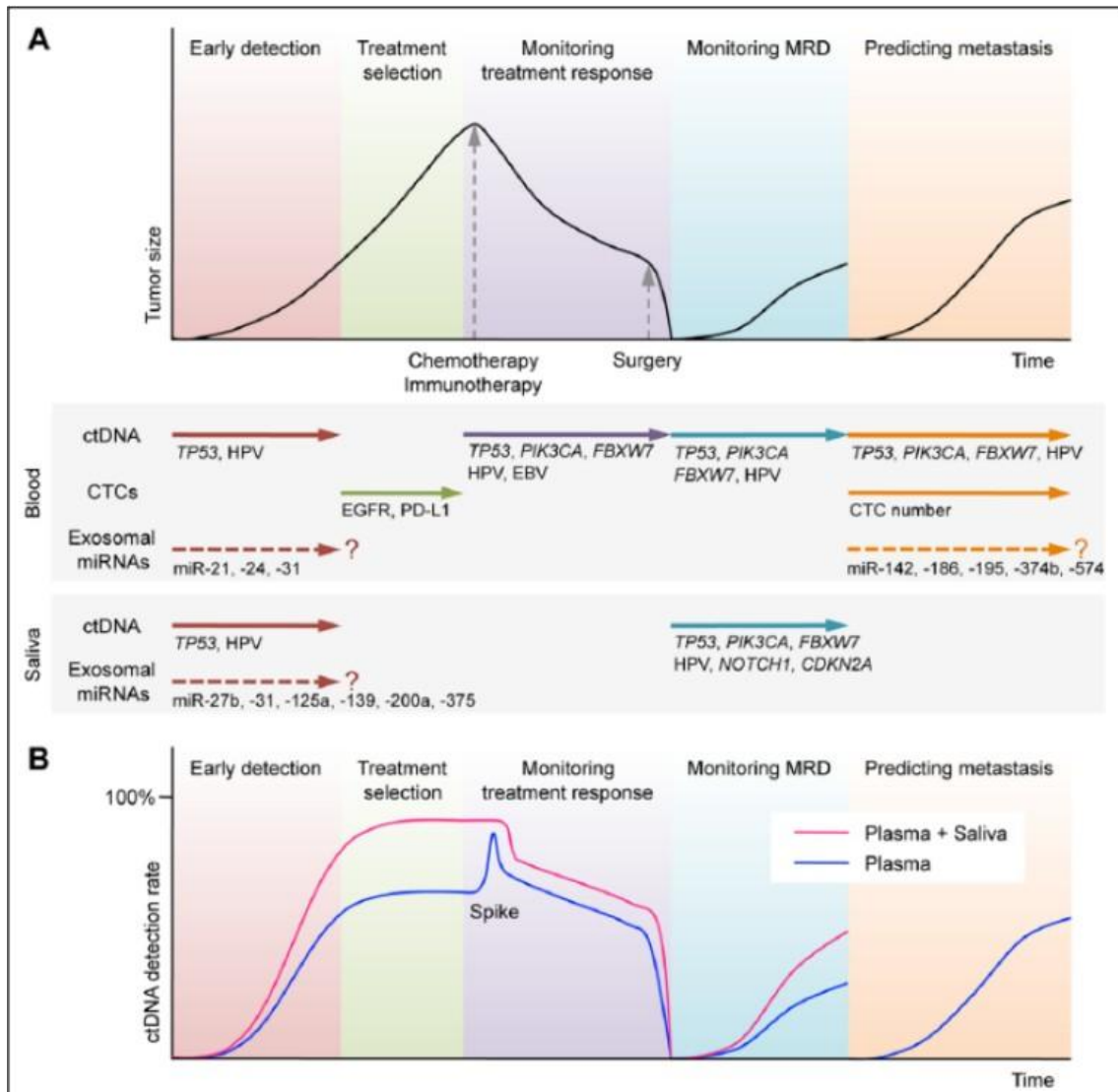


Рис. 1.4. Потенційні клінічні застосування циркулюючих біомаркерів у лікуванні РГШ. (А) Схематичний час лікування захворювання та розмір пухлини у пацієнтів з РГШ, які проходять хіміотерапію (або імунотерапію) та хірургічні операції. Аналіз плазми цпДНК дозволяє раннє виявлення та моніторинг відповіді на лікування, моніторинг мінімальної залишкової хвороби та прогнозування метастазування. Аналіз циркулюючих пухлинних клітин може допомогти у виборі цільової терапії. (В) Використання цпДНК плазми та слини в поєднанні дозволяє вищу ймовірність виявлення раку, ніж використання лише цпДНК плазми. Сплеск рівня цпДНК відображає загибель пухлинних клітин при системній терапії.

Дослідження кліренсу, що вивчає період напіввиведення з плазми крові цпДНК у пацієнтів з карциною носоглотки продемонструвало, що середній період напіввиведення під час хіміотерапії становив 3,99 днів (діапазон, 1,85–28,29). Інше дослідження повідомляє, що період напіввиведення цпДНК з плазми (APC, KRAS, TP53 і PIK3CA) при колоректальному раку був через 114 хвилин після операції, підтверджуючи, що цпДНК є ідеальним біомаркером для швидкого моніторингу зміни розміру пухлини через її швидку динаміку [19].

1.4.1 Прогнозування метастазів

Нинішнє використання цпДНК засноване на доказах того, що вона має спільні мутаційні профілі з первинними або вторинними пухлинами. Це означає, що ми можемо отримати опис метастазування без необхідності інвазійної біопсії тканин. На підтримку цього, численні дослідження повідомляють про цпДНК як високочутливого біомаркера для метастазів, що відображає пухлинне навантаження та гетерогенетичність при різних видах раку, включаючи РГШ [20].

1.5 Методи на основі перетворення бісульфіту

Основна ідея методу перетворення бісульфіту (WGBS) полягає в тому, що бісульфітні солі дезамінують неметильовані залишки цитозину в урацил, тоді як метиловані цитозини залишаються незмінними під час перетворення [21]. Після ампліфікації за допомогою ПЛР двома наборами ланцюг-специфічних праймерів, метилювання перетвореної ДНК можна було визначити шляхом подальшого секвенування. Шляхом інтеграції

високопродуктивних підходів, можна охопити тисячі сайтів метилювання за одне тестування, що дозволяє виявити біомаркер *de novo* на геномному рівні.

WGBS є найбільш повною та інформативною технологією профілювання метилювання ДНК, яка безпосередньо створює бібліотеку та виконує секвенування після перетворення бісульфіту, забезпечуючи метилювання для всіх цитозинів, включаючи області з низькою щільністю CpG і не-CpG-сайти (CpA, CpT і CpC) [22].

WGBS можна виконати з ~30 нг ДНК і в деяких випадках лише з 125bp, що робить його основним методом для аналізу позаклітинної ДНК. RRBS (секвенування бісульфіту зі зменшеним представленням) використовує розщеплення MspI шляхом секвенування наступного покоління для виявлення статусу метилювання збагачених CpG областей [23]. Він коштує дешевше, ніж WGBS, але має відносно низьке охоплення всього геному і залежить від цілісності ДНК і методів сортування фрагментів. На жаль, RRBS вимагає багато високоякісної ДНК, і тому її важко застосувати до позаклітинної ДНК. MCTA-seq використовує специфічні праймери з тандемною послідовністю CpG CGGCGG для ампліфікації гіперметилюваних сайтів CpG. Завдяки багаторазовій ампліфікації в процесі цього методу можна виявити гіперметилювані сайти CpG зразки ктДНК всього 7,5 пг [24]. Хоча MCTA-seq переважно націлений на багаті CGGCGG CpG сайтів і, таким чином, може пропустити інші потенційні біомаркери, цей метод все ще має прийнятну чутливість. У порівнянні з RRBS, аналіз MCTA-seq може зменшити фон метилювання та це простіше і рентабельніше: оскільки будівництво бібліотеки займає всього кілька годин і відносно потрібна низька глибина секвенування. Піросеквенування є найбільш часто використовуваним кількісним методом і надає в реальному часі кількісні дані про статус метилювання кількох CpG всередині амплікону [25].

1.6 Рання діагностика

Рання діагностика раку важлива для покращення прогнозу лікування пацієнтів, але ефективні підходи ранньої діагностики на основі крові досі відсутні. Останні дослідження щодо застосування маркерів метилювання цпДНК переважно зосереджено на ранній діагностиці раку. Дослідники використовували повне бісульфітне секвенування для скринінгу та підтвердження метилювання цпДНК для діагностики та прогнозу гепатоцелюлярної карциноми (ГЦР). Прогностичні моделі прогнозування з використанням біомаркерів селективного метилювання досягнули підтвердженої чутливості 90,5% і специфічності 83,2% у розрізненні між пацієнтами з ГЦК та здоровими особами контролю [26]. В іншому дослідженні також використовували повне бісульфітне секвенування для профілювання списку вибраних CpG-сайтів у цпДНК для виявлення раку. Вони використовували новий алгоритм, щоб визначити оцінку метилювання для кожного зразка і оцінку метилювання. За допомогою цього можна було точно визначити наявність раку в 83,8% випадків зі 100% специфічністю.

Ключове обмеження цього дослідження полягало в тому, що в ньому брали участь пацієнти з поширеними формами раку, а не пацієнти на ранній стадії ракового захворювання. Недавнє дослідження застосувало cfMeDIP-seq для ДНК плазми без бісульфітів у всьому геномі профілювання метилювання. Цей метод збагатив CpG, потенційно більш інформативні фрагменти, таким чином підвищується економічна ефективність. Використовуючи аналіз на основі машинного навчання, цей метод може відрізнити пацієнтів із сімома різними типами раку від здорових контрольних пацієнтів із високою чутливістю на ранніх стадіях. Зовсім нещодавно в дослідженні оцінювалась ефективність повного метилювання аналізу цпДНК для виявлення та локалізації понад 50 типів раку на всіх стадіях. З повної панелі метилювання, що охоплює 103 456 окремих регіонів (17,2 Мб) і 1 116 720 цитозин-гуанін-динуклеотиди (CpGs), вони розробили класифікатор пан-раку, за допомогою

якого рак був виявлений на всіх стадіях (чутливість I–III стадій: 43,9%; чутливість I–IV стадії: 54,9%) при специфічності >99%. Крім того, цей підхід може також локалізувати тканину походження з точністю >90%, що важливо для спрямування подальшого спостереження. Статистичні результати цього звіту можуть містити деякі оманливі методологічні проблеми. Усі ці дослідження включали використання специфічних для типу раку, поряд із типоспецифічними для тканин, цпДНК біомаркери метилювання для діагностики раку. Крім того, застосування складної статистики та біоінформатичної стратегії в цих дослідженнях можуть ефективно підвищити чутливість і специфічність [27].

Крім того, специфічні для тканин і пухлин біомаркери 5hmC також можуть мати потенційну цінність при діагностиці раку. Вважається, що 5hmC відіграє роль у багатьох видах раку і демонструє тканинно-специфічний характер. Останні дослідження продемонстрували потенційне використання 5hmC в цпДНК як біомаркерів для прогнозування типу і стадій пухлини. Нещодавнє дослідження виявило підтверджено надійні асоційовані з раком сигнатури 5hmC у цпДНК, які були характерними для специфічних видів раку. Насправді виявлення 5hmC технічно складніше, оскільки потрібна висока глибина секвенування через низький рівень 5hmC в клітинах крові [28].

Подальша оцінка буде необхідною у більшій кількості пацієнтів з раком на ранніх стадіях і, зрештою, безсимптомно хворих особин, щоб краще оцінити потенційну користь метилювання цпДНК. Подібні до питань що зустрічається під час скринінгу пухлин, метилювання цпДНК також стикається з проблемою розрізнення цпДНК та геномною ДНК у ранній діагностиці пухлин. Причому кожен аналіз розроблений різними дослідницькими колективами має своє визначення позитивних чи негативних результатів, як а також власні методи отримання цпДНК, її аналізу та вимірювання; тому майбутні дослідження необхідно зосередити на створенні стандартизованого перехресного порівняння для забезпечення контролю якості і дозволити неупереджене порівняння між аналізами.

1.7 Точне виявлення циркулюючої пухлинної ДНК за допомогою консенсусного секвенування Nanopore

Наявність соматичних мутацій у молекулах безклітинної ДНК зазвичай використовується для приблизного визначення вмісту цпДНК. Однак виявлення цпДНК є складним завданням, оскільки неракові клітини також викидають позаклітинну ДНК в кров. Частка пухлинних молекул у крові, як правило, набагато нижча за 5 %, а фракції – до 0,1 %. Тому діагностичний аналіз цпДНК повинен бути швидким і дешевим, а також високочутливим. цпДНК можна виявити з хорошою чутливістю за допомогою цифрової крапельної ПЛР (ddPCR), але ця методика вимагає досить тривалого часу, оскільки вона зазвичай може досліджувати лише один локус на аналіз, а варіанти повинні бути відомі апіорі. В якості альтернативи використовуються підходи до секвенування наступного покоління, але вони вимагають високо оптимізованих робочих процесів лабораторії, що є складним завданням у невеликих лікарнях. Крім того, підходи NGS часто вимагають об'єднання кількох вибірок, щоб стати економічно ефективними, що збільшує час виконання [29].

Технологія Oxford Nanopore Technology (ONT), яка нещодавно з'явилася як потужна платформа секвенування, має переваги з точки зору швидкості, економічної ефективності та гнучкості (розподілене секвенування замість централізованого секвенування). Таким чином, секвенування ONT може бути дуже актуальним для швидкого та невідкладного клінічного тестування рідинної біопсії [30]. Однак існують два важливі обмеження для секвенування ONT, які перешкоджають його використанню в клінічних умовах. По-перше, поточні протоколи оптимізовані для довгих молекул ДНК. Найкоротший фрагмент, секвенований на цій платформі на сьогоднішній день, становить ~425 bp, що набагато довше, ніж середні 145 bp позаклітинної ДНК [31]. По-друге, рівень базальної помилки становить ~5–10%, що занадто високо для надійного виявлення позаклітинної ДНК. Кілька досліджень показали, що

багаторазове читання однієї і тієї ж молекули може знизити частоту помилок секвенування. Однак деякі з цих методів можуть виявити лише фракції цпДНК >5%, тоді як інші покладаються на самоциркуляризацію, що неможливе для коротких молекул цпДНК [32].

Було представлено нову техніку під назвою CyclomicsSeq, яка використовує циркуляризацію та конкатемеризацію коротких молекул ДНК та оптимізовану основну послідовність ДНК у поєднанні з секвенуванням ONT. Як доказ концепції, розроблено специфічний для TP53 протокол CyclomicsSeq і спеціальний конвеєр програмного забезпечення для визначення мутаційного навантаження серії зразків цпДНК, отриманих з рідких біопсій пацієнтів з негативними плоскоклітинними клітинами РГШ на вірус папіломи людини. TP53 є найбільш часто мутованим геном-супресором пухлини при раку людини, і тому служить широко застосовною мішенню для моніторингу раку на основі рідинної біопсії. Існує відносно небагато мутацій гарячих точок, що робить цей ген особливо придатним для підходів на основі NGS. Застосування до ВПЛ-негативного РГШ мотивується тим фактом, що п'ятирічна виживаність відносно низька, а значні переваги лікування можна отримати при ранній діагностиці рецидивуючої хвороби та/або відповіді на лікування. Більше того, диференціація залишкової або рецидивуючої пухлини та радіаційних ефектів часто є важкою під час оцінки відповіді або у разі підозри на рецидив, навіть за допомогою сучасних методів візуалізації. Приблизно 90% пацієнтів з ВПЛ-негативним РГШ мають соматичну мутацію в TP53 [33]. Ці мутації TP53 виникають на ранніх стадіях пухлинного генезу РГШ і як такі присутні (практично) у всіх пухлинних клітинах, включаючи субклони, які метастазують. З цієї причини виявлення мутованих молекул цпДНК TP53 в рідких біопсіях вважається ідеальним біомаркером для РГШ.

Було продемонстровано, що CyclomicsSeq призводить до високоточних консенсусних послідовностей, придатних для виявлення мутацій при роздільній здатності однієї молекули. Тестування поздовжньої рідинної біопсії за допомогою CyclomicsSeq правильно визначає наявність та відсутність

вмісту цпДНК, що може бути інформативним для лікування пацієнтів з РГШ. CyclomicsSeq можна застосувати до однієї або кількох геномних областей вибору, в принципі, таким чином представляючи новий тест рідинної біопсії, який є релевантним для діагностичного моніторингу будь-якої солідної пухлини, для якої цпДНК є відповідним біомаркером [34].

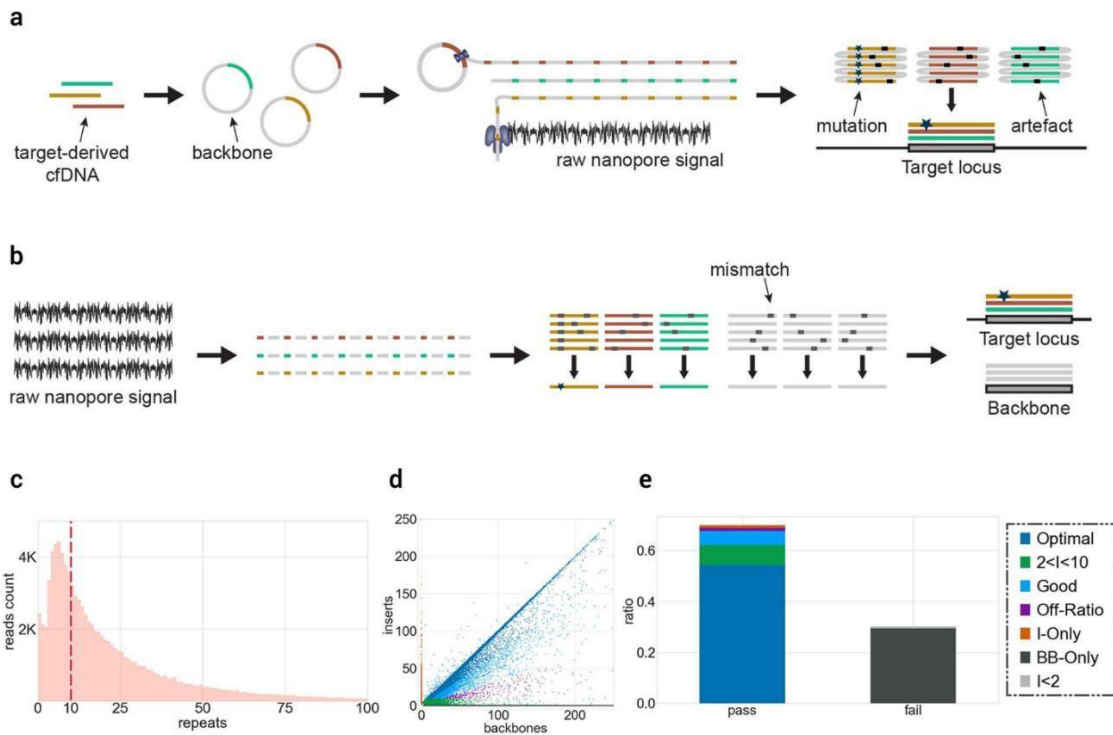


Рис. 1.5. а Експериментальна установка CyclomicsSeq. Ампліфікована ПЛР ("цільове") цпДНК циркуляризується за допомогою оптимізованого ДНК. Ампліфікація Rolling Circle генерує довгу молекулу ДНК з чергуванням послідовностей вставки та скелета, яка секвенується за допомогою секвенування ONT. Консенсусний виклик послідовності ДНК дозволяє розрізняти мутації та артефакти секвенування. б Схематичний огляд біоінформаційного трубопроводу. с Розподіл копій вставки в залежності від кількості читань для репрезентативного циклу CyclomicsSeq (#CY_SM_PC_HN_0002_001_000). d Співвідношення вставки та магістралі для зчитування CyclomicsSeq для репрезентативного запуску CyclomicsSeq (#CY_SM_PC_HN_0002_001_000). Кожне зчитування представлено точкою даних (точкою). Кольори, зазначені в легенді, представляють різні категорії, до яких може належати читання. Оптимальне зчитування CyclomicsSeq є

результатом співвідношення один до одного копій вставки та магістралі та містить щонайменше 10 повторів (синій). Інші категорії включають: читання з меншою кількістю повторів (зелений), зчитування без хребта (помаранчевий), читання без цільової вставки (сірий). Зчитування з коефіцієнтами ВВ:І від 0,35 до 3 визначаються як «Добре» (Блакитний), тоді як інші класифікуються як «Відношення поза межами» (фіолетовий). е Співвідношення даних секвенування, згрупованих за типом читання, для репрезентативного циклу CyclomicsSeq (#CY_SM_PC_HN_0002_001_000). У цьому випадку понад 60% даних було використано для формування консенсусних зчитування. Залишок даних було відкинуто, оскільки вони містили послідовності лише магістралі.

РОЗДІЛ 2. Матеріали та методи дослідження

2.1 Матеріал дослідження

Матеріалом дослідження були зразки слини хворих на РГШ та здорових людей. Зразки слини пацієнтів були надані Fakultní nemocnice Hradec Králové (Чехія), за умов інформованої згоди, висновок етичної комісії у Додатку 1.

2.2 Протокол очищення позаклітинної ДНК

Валідація протоколу відбувалася на основі QIAamp® Circulating Nucleic Acid Handbook.

До протейнінази К додаємо зразок слини. Додаємо буфер ACL та буфер ATL, закриваємо і швидко перемішуємо протягом 30 с. Інкубуємо 60°C при 30 хв. Додаємо буфер ACB до лізату, закриваємо кришку та ретельно перемішуємо. Інкубуємо суміш лізат-буфер ACB протягом 5 хв на льоду.

Вставляємо колонку QIAamp Mini в VacConnector на QIAvac 24 Plus. Вставляємо розширювач 20 мл у відкриту колонку QIAamp Mini. Обережно наносимо лізат у подовжувач пробірки колонки QIAamp Mini. Вмикаємо вакуумний насос. Коли всі лізати протягнуто через колонки повністю вимкаємо вакуумний насос і знижуємо тиск до 0 мбар.

Наносимо буфер ACW1 на колонку QIAamp Mini. Залишаємо кришку колонки відкритою та вмикаємо вакуумний насос. Після того, як весь буфер ACW1 пройде через колонку QIAamp Mini, вимикаємо вакуумний насос і зменшуємо тиск до 0 мбар.

Наносимо буфер ACW2 на колонку QIAamp Mini. Залишаємо кришку колонки відкритою та вмикаємо вакуумний насос. Після того, як весь буфер ACW2 пройде через колонку QIAamp Mini, вимикаємо вакуумний насос і скидаємо тиск до 0 мбар.

Наносимо етанол (96–100%) на колонку QIAamp Mini. Залишаємо кришку відкритою і вмикаємо вакуумний насос. Після того, як весь етанол був витягнутий через колонку QIAamp Mini, вимикаємо вакуумний насос і скидаємо тиск до 0 мбар.

Закриваємо кришку колонки QIAamp Mini, виймаємо її з вакуумного колектора і відкидаємо VacConnector. Поміщаємо колонку QIAamp Mini в чисту пробірку і центрифугуємо на повній швидкості (20 000 x g; 14 000 об/хв) 3 хв.

Поміщаємо колонку QIAamp Mini в нову пробірку для збору, відкриваємо кришку та інкубуємо при 56°C протягом 10 хв до повного висихання мембрани.

Поміщаємо колонку QIAamp Mini в чисту пробірку для елюювання об'ємом 1,5 мл. Обережно наносимо 20–150 мкл буфера AVE в центр QIAamp. Закриваємо кришку і інкубуємо при кімнатній температурі протягом 3 хв.

Центрифугуємо на повній швидкості (20 000 x g; 14 000 об/хв) протягом 1 хв для елюювання нуклеїнових кислот.

2.3 Контроль якості ДНК після очищення

Концентрація розраховується за допомогою Qubit™ dsDNA HS. Розмір фрагментів досліджуємо за допомогою The Agilent TapeStation system. Маркери для порівняння у Додатку 3.

2.4 Очищення за допомогою гелю

Додатковий етап очищення був проведений за допомогою гелелектрофорезу TBE 6%.

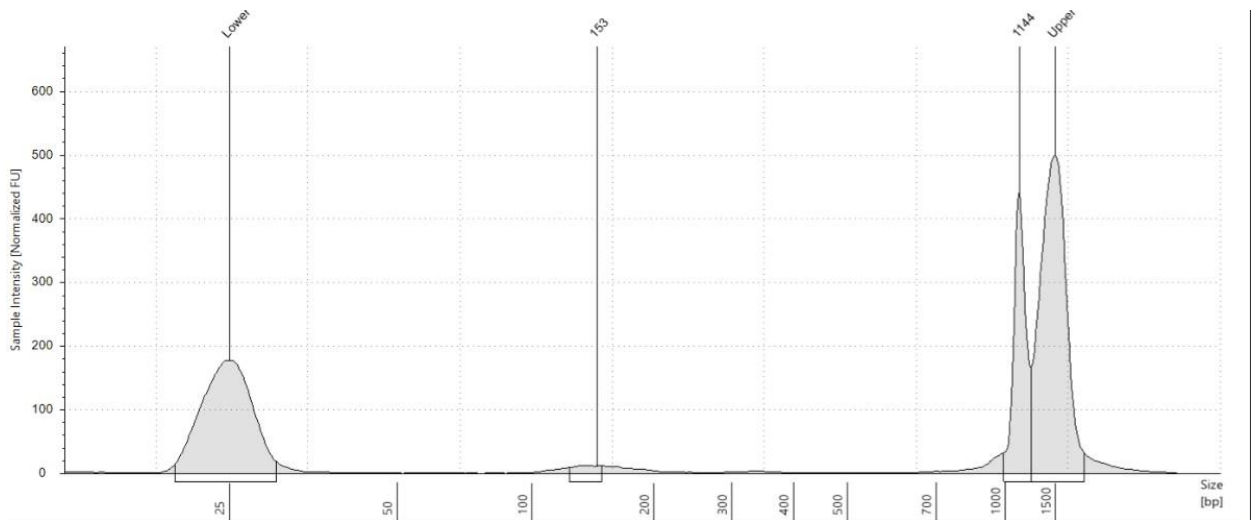


Рис. 2.1. Результати The Agilent TapeStation: бачимо пік в 153 пари і піки більше 1000 пар. Згідно зі статтями позаклітинна ДНК знаходиться в межах 150-200 пар.



Рис. 2.2. Результати гел-електрофорезу.



Рис. 2.3. Гель-електрофорез без фрагментів довжиною 200 пар.

2.5 Підготовка бібліотеки та секвенування

Для бібліотеки використовували стандартний протокол Oxford Nanopore.

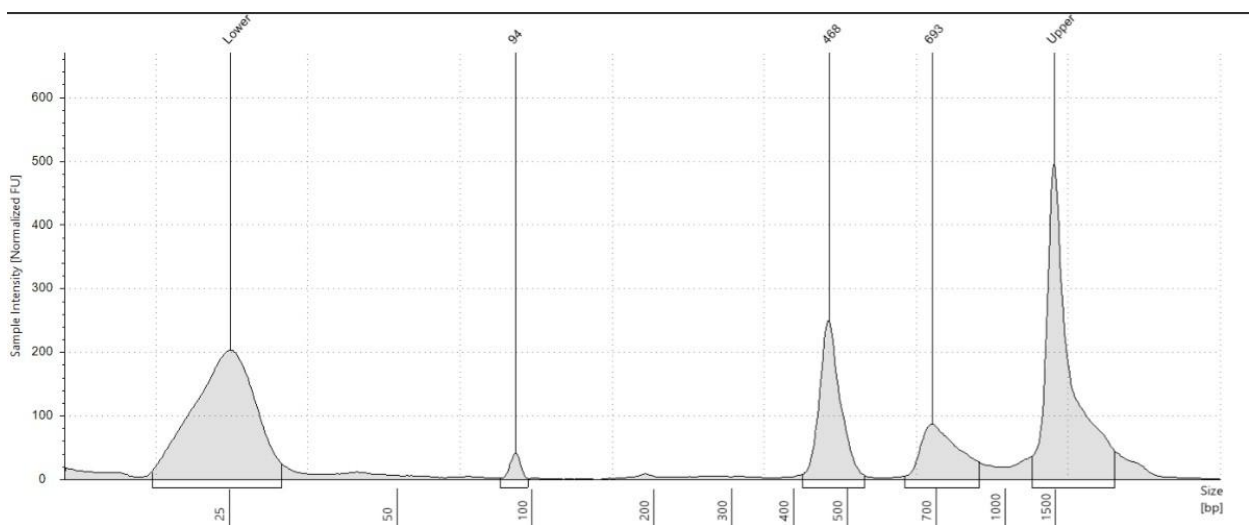


Рис.2.4. Результати The Agilent TapeStation після підготовки бібліотеки.

Для аналізу результатів секвенування використовували R скрипт,
Додаток 2.

РОЗДІЛ 3. Результати дослідження

3.1. Результати секвенування

Позаклітинна ДНК, як правило, досить коротка (~180), тому не ідеальна для секвенування ONT. Однак, якщо секвенувати безпосередньо (без ампліфікації ПЛР), інформація про метилювання зберігається у файлах fastq, і вже є деякі публікації, які використовують ONT для секвенування цих біопроб.

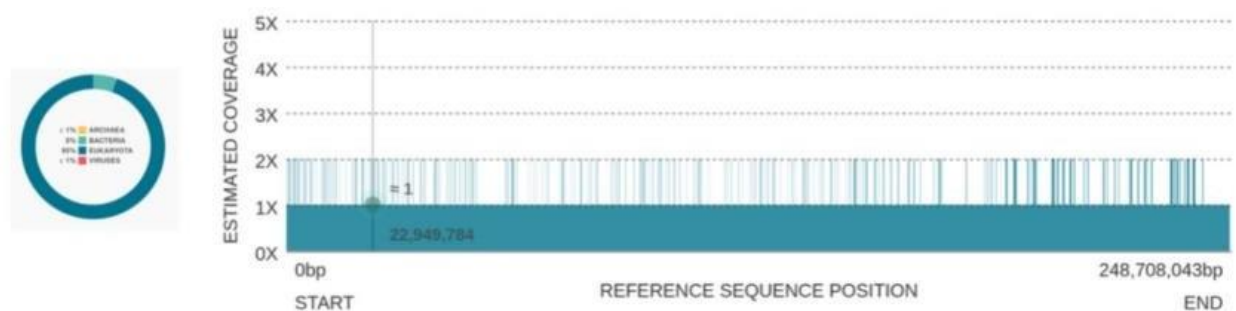


Рис. 3.1. Покриття секвенування.

У наших руках одна проточна комірка (з одним лише зразком – без мультиплексування) генерувала прибіл. 4 млн читань, середня довжина зчитування ~600 bp, прибіл. 2,7Gb були секвеновані (але прибіл. 50% були невдалими зчитуваннями). По-перше, ми використали файли Epi2Me та fastq (які були згенеровані в режимі реального часу під час виконання секвенування БЕЗ конфігурації для збереження інформації про метилювання), яка є дуже інтуїтивною програмою від ONT для аналізу даних, просто щоб підтвердити, що те, що ми секвенували, є ДНК людини. Справді, >95% читань відповідають людині.

Коли ми потім викликали метилювання, ми втратили значну частину зчитування (вихід від nanopolish):

- total reads: 1183314
- unparseable: 0
- qc fail: 888

- could not calibrate: 75943
- no alignment: 648600
- bad fast5: 0

Нарешті, ми отримали дані про метилювання від приблизно 1М окремих CpG-сайтів по всьому геному. Більшість із цих сайтів були охоплені одним читанням:

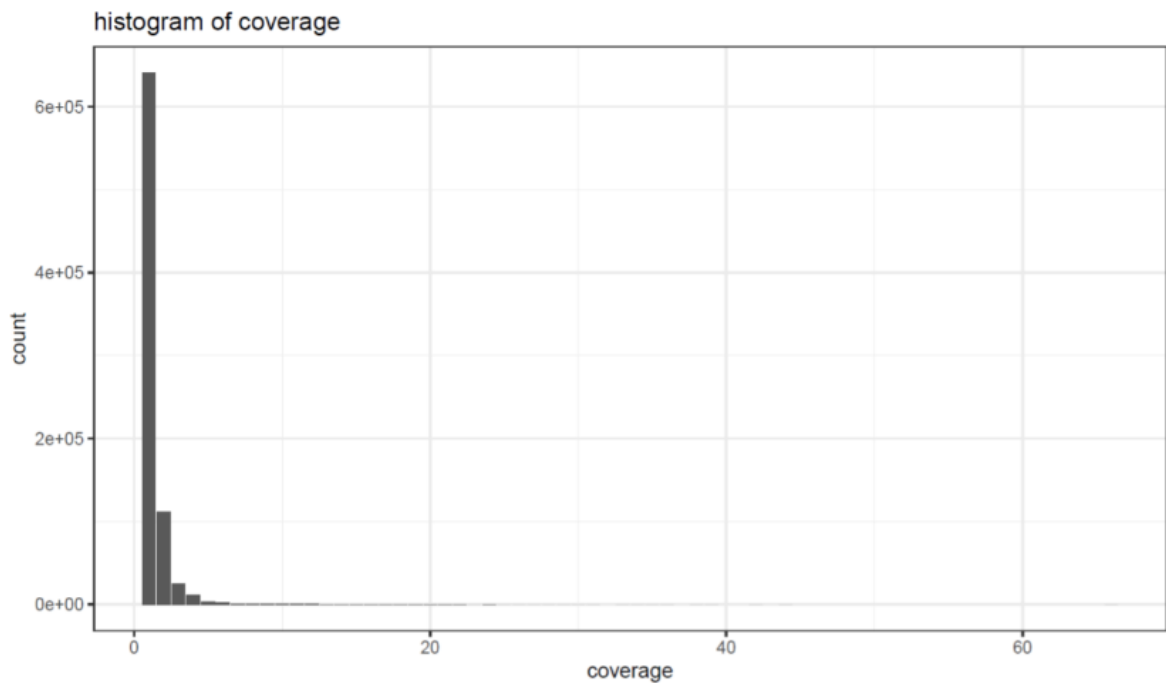


Рис. 3.2. Гістограма покриття секвенування.

Згідно з кількома статтями, які використовують ONT для вивчення метилювання позаклітинної ДНК, навіть з таким низьким охопленням, можна досить правдиво зобразити метилом, оскільки існує дуже висока кореляція між оцінкою ONT та іншими методами, що забезпечують більше охоплення та менше помилок (наприклад, бісульфітне секвенування всього геному за допомогою Illumina). Тому не дивно з'ясувати, що більшість сайтів мали

значення метилювання 1 або 0 (оскільки зазвичай було лише одне вимірювання на сайт).

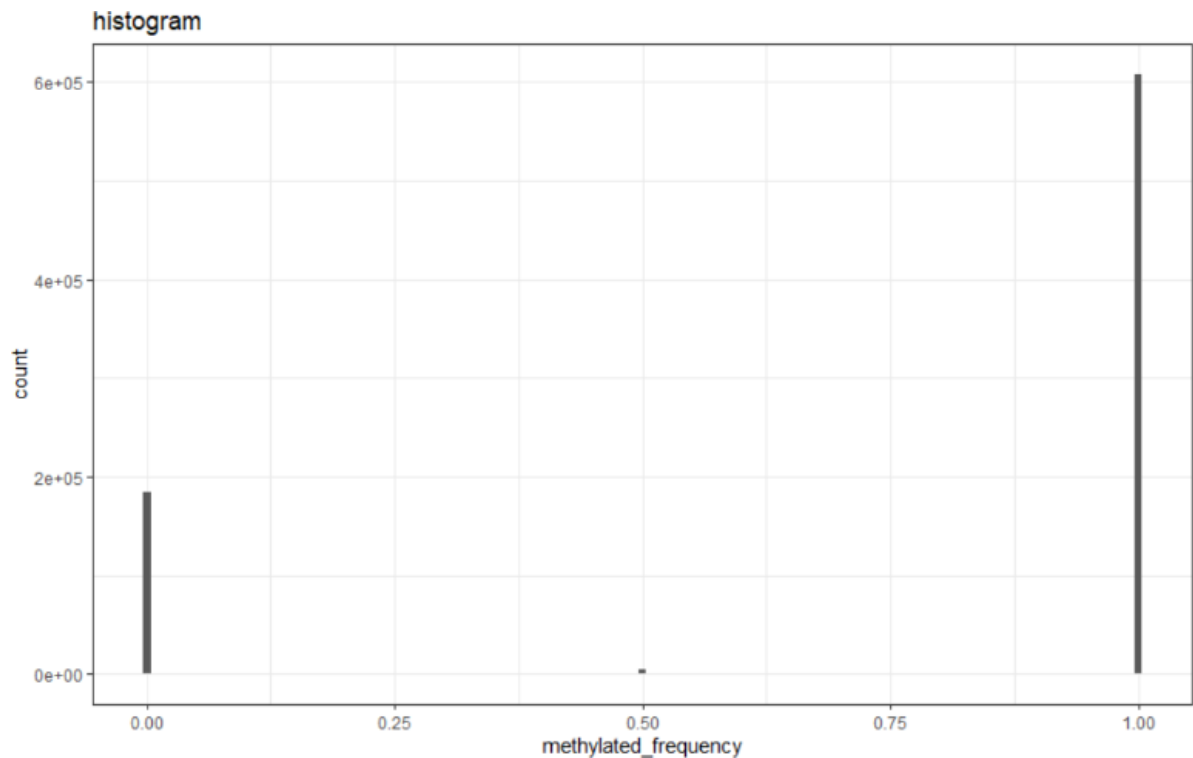


Рис. 3.3. Гістограма частоти метилювання.

3.2 Метилювання позаклітинної ДНК на хромосомах

Ми побудували на графіку значення метилювання окремих сайтів CpG на хромосому і помітили, що середнє метилювання хромосом було високим і наближалось до середнього метилювання по всьому геному (прибл. 0,76, червона пунктирна лінія на графіку). Це узгоджується з інформацією, знайденою в літературі, де середнє метилювання у здорової людини є досить високим. На графіку нижче окремі вимірювання зі значенням метилювання 0, 1 і 0,5 створюють лінію, оскільки ці значення метилювання часто зустрічаються в наших даних. Одним із підходів до вирішення цього було б створення фрейму даних, де середнє метилювання розраховувалося б на 10 Мб вікон геному, що ми хочемо навчитися робити в майбутньому.

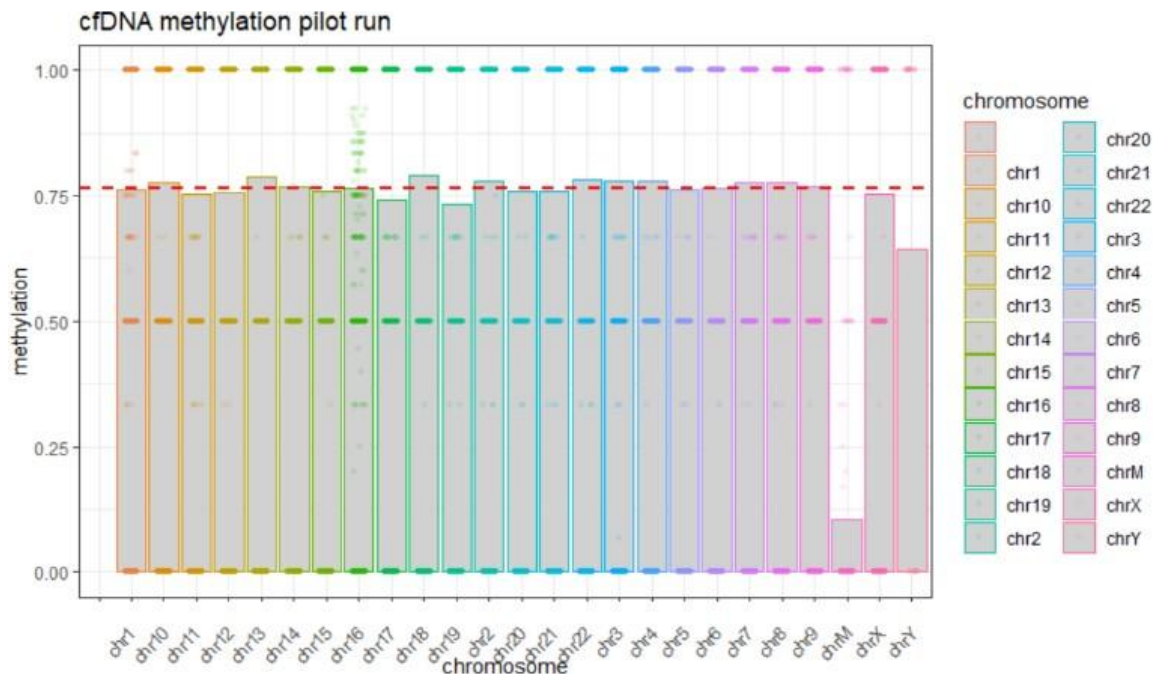


Рис. 3.4. Метилування позаклітинної ДНК на хромосомах.

Цікаво, що ми отримали кілька читань, що охоплюють мітохондріальну ДНК (chrM на графіку), і там метилування було набагато кращим у порівнянні з ядерною ДНК. Здається, це узгоджується з даними з літератури, де кілька досліджень показують, що метилування ДНК значно нижче в мтДНК, ймовірно, через те, що мітохондрії можуть не мати ферментного механізму для виконання цієї модифікації своєї ДНК.

РОЗДІЛ 4. Обговорення результатів дослідження

Рідинна біопсія проклала спосіб ранньої діагностики, прогнозу рецидивів, моніторингу, а також скринінгу груп високого ризику. Аналіз метилювання позаклітинної ДНК швидко з'явився як багатообіцяючий інструмент з багатьма потенційними можливостями клінічного застосування в області онкології, однак, все ще існує багато проблем і обмежень. Специфічні аналізи та демонстрація ефективності біомаркерів у цільових груп населення – це дві ключові перешкоди, які необхідно подолати.

Ми показали, що застосування технологій секвенування Oxford Nanopore виявило метилювання позаклітинної ДНК здорових пацієнтів при зчитуванні довгих фрагментів. Декілька джерел вказують на те, що отримані нами результати є достовірними. Наступний крок це секвенування та дослідження метилювання позаклітинної ДНК пацієнтів хворих на РГШ. Проте ми вже змогли вказати, що для дослідження метилювання ДНК хворих на РГШ пацієнтів методом рідинної біопсії можна використовувати технології Oxford Nanopore.

Також цей метод, на відміну від Illumina, дозволяє не використовувати ПЛР та Bisulfite Conversion-Based метод, що негативно впливає на ДНК, пошкоджуючі її.

Ми продемонстрували, що середнє метилювання хромосом було високим і наближалось до середнього метилювання по всьому геному (прибл. 0,76). Це узгоджується з інформацією, знайденою в літературі, де середнє метилювання у здорової людини є досить високим.

Таким чином, секвенування ONT може бути дуже актуальним для швидкого та невідкладного клінічного тестування рідинної біопсії. Сучасні дослідження пропонують ряд покращень, такі як CyclomicsSeq, що призводить до високоточних консенсусних послідовностей, придатних для виявлення мутацій при роздільній здатності однієї молекули. Тестування поздовжньої рідинної біопсії за допомогою CyclomicsSeq правильно визначає наявність та

відсутність вмісту позаклітинної ДНК, що може бути інформативним для лікування пацієнтів з РГШ. CyclomicsSeq можна застосувати до однієї або кількох геномних областей вибору, в принципі, таким чином представляючи новий тест рідинної біопсії, який є релевантним для діагностичного моніторингу будь-якої солідної пухлини, для якої позаклітинна ДНК є відповідним біомаркером.

У наших планах продовжити секвенування позаклітинної ДНК пацієнтів, хворих на РГШ з застосуванням Oxford Nanopore.

ВИСНОВКИ

Технології Oxford Nanopore є ефективним методом для секвенування фрагментів позаклітинної ДНК. Хоч цей пілотний запуск не був оптимальним з точки зору згенерованих читань, отримані дані щодо метилювання ДНК здорових пацієнтів відповідають референсам з інших секвенувань.

Середнє метилювання хромосом було високим і наближалось до середнього метилювання по всьому геному приблизно 0,76. ONT для вивчення метилювання позаклітинної ДНК, навіть з таким низьким охопленням, може досить правдиво зобразити метилом, оскільки існує дуже висока кореляція між оцінкою ONT та іншими методами, що забезпечують більше охоплення та менше помилок (наприклад, секвенування бісульфіту всього геному за допомогою Illumina). Більшість сайтів мали значення метилювання 1 або 0 (оскільки зазвичай було лише одне вимірювання на сайт). Тому Oxford Nanopore являє собою надійну альтернативу секвенуванню Illumina.

Наявність протоколів для аналізу позаклітинної ДНК на основі Nanopore зробить цей аналіз нарешті доступним для використання повного потенціалу рідинної біопсії як для дослідницьких, так і для клінічних цілей.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

- [1] Ferlay J, Colombet M, Soerjomataram I, Mathers C, Parkin DM, Piñeros M, Znaor A, Bray F. Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *Int J Cancer*. 2019 Apr 15;144(8):1941-1953. doi: 10.1002/ijc.31937. Epub 2018 Dec 6. PMID: 30350310.
- [2] Yates LR, Gerstung M, Knappskog S, Desmedt C, Gundem G, Van Loo P, Aas T, Alexandrov LB, Larsimont D, Davies H, Li Y, Ju YS, Ramakrishna M, Haugland HK, Lilleng PK, Nik-Zainal S, McLaren S, Butler A, Martin S, Glodzik D, Menzies A, Raine K, Hinton J, Jones D, Mudie LJ, Jiang B, Vincent D, Greene-Colozzi A, Adnet PY, Fatima A, Maetens M, Ignatiadis M, Stratton MR, Sotiriou C, Richardson AL, Lønning PE, Wedge DC, Campbell PJ. Subclonal diversification of primary breast cancer revealed by multiregion sequencing. *Nat Med*. 2015 Jul;21(7):751-9. doi: 10.1038/nm.3886. Epub 2015 Jun 22. PMID: 26099045; PMCID: PMC4500826.
- [3] Kaczor-Urbanowicz KE, Martin Carreras-Presas C, Aro K, Tu M, Garcia-Godoy F, Wong DT. Saliva diagnostics - Current views and directions. *Exp Biol Med (Maywood)*. 2017 Mar;242(5):459-472. doi: 10.1177/1535370216681550. Epub 2016 Dec 8. PMID: 27903834; PMCID: PMC5367650.
- [4] Miller SM. Saliva testing--a nontraditional diagnostic tool. *Clin Lab Sci*. 1994 Jan-Feb;7(1):39-44. PMID: 10172036.
- [5] Patel A, Patel S, Patel P, Tanavde V. Saliva Based Liquid Biopsies in Head and Neck Cancer: How Far Are We From the Clinic? *Front Oncol*. 2022 Mar 21;12:828434. doi: 10.3389/fonc.2022.828434. PMID: 35387114; PMCID: PMC8977527.
- [6] Schwarzenbach H, Hoon DS, Pantel K. Cell-free nucleic acids as biomarkers in cancer patients. *Nat Rev Cancer*. 2011 Jun;11(6):426-37. doi: 10.1038/nrc3066. Epub 2011 May 12. PMID: 21562580.

- [7] Stroun M, Lyautey J, Lederrey C, Olson-Sand A, Anker P. About the possible origin and mechanism of circulating DNA apoptosis and active DNA release. *Clin Chim Acta*. 2001 Nov;313(1-2):139-42. doi: 10.1016/s0009-8981(01)00665-9. PMID: 11694251.
- [8] Dressman D, Yan H, Traverso G, Kinzler KW, Vogelstein B. Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations. *Proc Natl Acad Sci U S A*. 2003 Jul 22;100(15):8817-22. doi: 10.1073/pnas.1133470100. Epub 2003 Jul 11. PMID: 12857956; PMCID: PMC166396.
- [9] Diaz LA Jr, Bardelli A. Liquid biopsies: genotyping circulating tumor DNA. *J Clin Oncol*. 2014 Feb 20;32(6):579-86. doi: 10.1200/JCO.2012.45.2011. Epub 2014 Jan 21. PMID: 24449238; PMCID: PMC4820760.
- [10] Wang Y, Springer S, Mulvey CL, Silliman N, Schaefer J, Sausen M, James N, Rettig EM, Guo T, Pickering CR, Bishop JA, Chung CH, Califano JA, Eisele DW, Fakhry C, Gourin CG, Ha PK, Kang H, Kiess A, Koch WM, Myers JN, Quon H, Richmon JD, Sidransky D, Tufano RP, Westra WH, Bettegowda C, Diaz LA Jr, Papadopoulos N, Kinzler KW, Vogelstein B, Agrawal N. Detection of somatic mutations and HPV in the saliva and plasma of patients with head and neck squamous cell carcinomas. *Sci Transl Med*. 2015 Jun 24;7(293):293ra104. doi: 10.1126/scitranslmed.aaa8507. PMID: 26109104; PMCID: PMC4587492.
- [11] El-Naggar AK, Mao L, Staerckel G, Coombes MM, Tucker SL, Luna MA, Clayman GL, Lippman S, Goepfert H. Genetic heterogeneity in saliva from patients with oral squamous carcinomas: implications in molecular diagnosis and screening. *J Mol Diagn*. 2001 Nov;3(4):164-70. doi: 10.1016/S1525-1578(10)60668-X. PMID: 11687600; PMCID: PMC1906964.
- [12] Luo H, Wei W, Ye Z, Zheng J, Xu RH. Liquid Biopsy of Methylation Biomarkers in Cell-Free DNA. *Trends Mol Med*. 2021 May;27(5):482-500. doi: 10.1016/j.molmed.2020.12.011. Epub 2021 Jan 23. PMID: 33500194.

- [13] Li Y, St John MA, Zhou X, Kim Y, Sinha U, Jordan RC, Eisele D, Abemayor E, Elashoff D, Park NH, Wong DT. Salivary transcriptome diagnostics for oral cancer detection. *Clin Cancer Res.* 2004 Dec 15;10(24):8442-50. doi: 10.1158/1078-0432.CCR-04-1167. PMID: 15623624.
- [14] Nian J, Sun X, Ming S, Yan C, Ma Y, Feng Y, Yang L, Yu M, Zhang G, Wang X. Diagnostic Accuracy of Methylated SEPT9 for Blood-based Colorectal Cancer Detection: A Systematic Review and Meta-Analysis. *Clin Transl Gastroenterol.* 2017 Jan 19;8(1):e216. doi: 10.1038/ctg.2016.66. PMID: 28102859; PMCID: PMC5288600.
- [15] Lamb YN, Dhillon S. Epi proColon® 2.0 CE: A Blood-Based Screening Test for Colorectal Cancer. *Mol Diagn Ther.* 2017 Apr;21(2):225-232. doi: 10.1007/s40291-017-0259-y. PMID: 28155091.
- [16] Reinert T, Schøler LV, Thomsen R, Tobiasen H, Vang S, Nordentoft I, Lamy P, Kannerup AS, Mortensen FV, Stribolt K, Hamilton-Dutoit S, Nielsen HJ, Laurberg S, Pallisgaard N, Pedersen JS, Ørntoft TF, Andersen CL. Analysis of circulating tumour DNA to monitor disease burden following colorectal cancer surgery. *Gut.* 2016 Apr;65(4):625-34. doi: 10.1136/gutjnl-2014-308859. Epub 2015 Feb 4. PMID: 25654990.
- [17] Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, Thornton K, Agrawal N, Sokoll L, Szabo SA, Kinzler KW, Vogelstein B, Diaz LA Jr. Circulating mutant DNA to assess tumor dynamics. *Nat Med.* 2008 Sep;14(9):985-90. doi: 10.1038/nm.1789. Epub 2007 Jul 31. PMID: 18670422; PMCID: PMC2820391.
- [18] Hamana K, Uzawa K, Ogawara K, Shiiba M, Bukawa H, Yokoe H, Tanzawa H. Monitoring of circulating tumour-associated DNA as a prognostic tool for oral squamous cell carcinoma. *Br J Cancer.* 2005 Jun 20;92(12):2181-4. doi: 10.1038/sj.bjc.6602635. PMID: 15928666; PMCID: PMC2361808.
- [19] Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang Y, Agrawal N, Bartlett BR, Wang H, Luber B, Alani RM, Antonarakis ES, Azad NS, Bardelli A, Brem H,

Cameron JL, Lee CC, Fecher LA, Gallia GL, Gibbs P, Le D, Giuntoli RL, Goggins M, Hogarty MD, Holdhoff M, Hong SM, Jiao Y, Juhl HH, Kim JJ, Siravegna G, Laheru DA, Lauricella C, Lim M, Lipson EJ, Marie SK, Netto GJ, Oliner KS, Olivi A, Olsson L, Riggins GJ, Sartore-Bianchi A, Schmidt K, Shih IM, Oba-Shinjo SM, Siena S, Theodorescu D, Tie J, Harkins TT, Veronese S, Wang TL, Weingart JD, Wolfgang CL, Wood LD, Xing D, Hruban RH, Wu J, Allen PJ, Schmidt CM, Choti MA, Velculescu VE, Kinzler KW, Vogelstein B, Papadopoulos N, Diaz LA Jr. Detection of circulating tumor DNA in early- and late-stage human malignancies. *Sci Transl Med.* 2014 Feb 19;6(224):224ra24. doi: 10.1126/scitranslmed.3007094. PMID: 24553385; PMCID: PMC4017867.

[20] Dawson SJ, Tsui DW, Murtaza M, Biggs H, Rueda OM, Chin SF, Dunning MJ, Gale D, Forsheo T, Mahler-Araujo B, Rajan S, Humphray S, Becq J, Halsall D, Wallis M, Bentley D, Caldas C, Rosenfeld N. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N Engl J Med.* 2013 Mar 28;368(13):1199-209. doi: 10.1056/NEJMoa1213261. Epub 2013 Mar 13. PMID: 23484797.

[21] Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL. A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A.* 1992 Mar 1;89(5):1827-31. doi: 10.1073/pnas.89.5.1827. PMID: 1542678; PMCID: PMC48546.

[22] Huang J, Wang L. Cell-Free DNA Methylation Profiling Analysis-Technologies and Bioinformatics. *Cancers (Basel).* 2019 Nov 6;11(11):1741. doi: 10.3390/cancers11111741. PMID: 31698791; PMCID: PMC6896050.

[23] Lissa D, Robles AI. Methylation analyses in liquid biopsy. *Transl Lung Cancer Res.* 2016 Oct;5(5):492-504. doi: 10.21037/tlcr.2016.10.03. PMID: 27826530; PMCID: PMC5099516.

[24] Wen L, Li J, Guo H, Liu X, Zheng S, Zhang D, Zhu W, Qu J, Guo L, Du D, Jin X, Zhang Y, Gao Y, Shen J, Ge H, Tang F, Huang Y, Peng J. Genome-scale

detection of hypermethylated CpG islands in circulating cell-free DNA of hepatocellular carcinoma patients. *Cell Res.* 2015 Nov;25(11):1250-64. doi: 10.1038/cr.2015.126. Epub 2015 Oct 30. Erratum in: *Cell Res.* 2015 Dec;25(12):1376. PMID: 26516143; PMCID: PMC4650428.

[25] Kristensen LS, Hansen JW, Kristensen SS, Tholstrup D, Harsløf LB, Pedersen OB, De Nully Brown P, Grønbæk K. Aberrant methylation of cell-free circulating DNA in plasma predicts poor outcome in diffuse large B cell lymphoma. *Clin Epigenetics.* 2016 Sep 7;8(1):95. doi: 10.1186/s13148-016-0261-y. PMID: 27610206; PMCID: PMC5015248.

[26] Sethi S, Benninger MS, Lu M, Havard S, Worsham MJ. Noninvasive molecular detection of head and neck squamous cell carcinoma: an exploratory analysis. *Diagn Mol Pathol.* 2009 Jun;18(2):81-7. doi: 10.1097/PDM.0b013e3181804b82. PMID: 19430297; PMCID: PMC2693294.

[27] Shen SY, Burgener JM, Bratman SV, De Carvalho DD. Preparation of cfMeDIP-seq libraries for methylome profiling of plasma cell-free DNA. *Nat Protoc.* 2019 Oct;14(10):2749-2780. doi: 10.1038/s41596-019-0202-2. Epub 2019 Aug 30. PMID: 31471598.

[28] Han D, Lu X, Shih AH, Nie J, You Q, Xu MM, Melnick AM, Levine RL, He C. A Highly Sensitive and Robust Method for Genome-wide 5hmC Profiling of Rare Cell Populations. *Mol Cell.* 2016 Aug 18;63(4):711-719. doi: 10.1016/j.molcel.2016.06.028. Epub 2016 Jul 28. PMID: 27477909; PMCID: PMC4992443.

[29] Bhatti H, Lu Z, Liu Q. Nanopore Detection of Cancer Biomarkers: A Challenge to Science. *Technol Cancer Res Treat.* 2022 Jan-Dec;21:15330338221076669. doi: 10.1177/15330338221076669. PMID: 35229683; PMCID: PMC8891933.

[30] Deamer D, Akeson M, Branton D. Three decades of nanopore sequencing. *Nat Biotechnol.* 2016 May 6;34(5):518-24. doi: 10.1038/nbt.3423. PMID: 27153285; PMCID: PMC6733523.

- [31] Wei S, Weiss ZR, Williams Z. Rapid Multiplex Small DNA Sequencing on the MinION Nanopore Sequencing Platform. *G3 (Bethesda)*. 2018 May 4;8(5):1649-1657. doi: 10.1534/g3.118.200087. PMID: 29540443; PMCID: PMC5940156.
- [32] Wilson BD, Eisenstein M, Soh HT. High-Fidelity Nanopore Sequencing of Ultra-Short DNA Targets. *Anal Chem*. 2019 May 21;91(10):6783-6789. doi: 10.1021/acs.analchem.9b00856. Epub 2019 May 10. PMID: 31038923; PMCID: PMC6533607.
- [33] Baugh EH, Ke H, Levine AJ, Bonneau RA, Chan CS. Why are there hotspot mutations in the TP53 gene in human cancers? *Cell Death Differ*. 2018 Jan;25(1):154-160. doi: 10.1038/cdd.2017.180. Epub 2017 Nov 3. PMID: 29099487; PMCID: PMC5729536.
- [34] Marcozzi A, Jager M, Elferink M, Straver R, van Ginkel JH, Peltenburg B, Chen LT, Renkens I, van Kuik J, Terhaard C, de Bree R, Devriese LA, Willems SM, Kloosterman WP, de Ridder J. Accurate detection of circulating tumor DNA using nanopore consensus sequencing. *NPJ Genom Med*. 2021 Dec 9;6(1):106. doi: 10.1038/s41525-021-00272-y. PMID: 34887408; PMCID: PMC8660781.

Додаток 1. Висновок етичної комісії

Etická komise, Fakultní nemocnice Hradec Králové, Sokolská 581, 500 05 Hradec Králové
Ethics Committee, University Hospital Hradec Kralove, Sokolská 581, 500 05 Hradec Králové

STANOVISKO ETICKÉ KOMISE Opinion of the Ethics Committee

Vážená paní
PharmDr. Natálie Birknerová
Ústav klinické biochemie a diagnostiky
Fakultní nemocnice Hradec Králové

Číslo jednací/Reference number: 202106 P05

Název studie/Full Title of study: Epigenetická analýza tekutých biopsií jako neinvazivní metoda pro stanovení biomarkerů nádorů hlavy a krku

Datum doručení žádosti/Date of submission of the Application Form: 19May2021

Místo hodnocení/ Jméno zkoušejícího / Trial Site / Name of Investigator:
> PharmDr. Natálie Birknerová, Ústav klinické biochemie a diagnostiky, Fakultní nemocnice Hradec Králové

Seznam hodnocených dokumentů/List of all submitted documents:
> Žádost o posouzení EK FNHK

Vyjádření EK/ Ethics Committee's opinion:
EK vydává / EC issues

Souhlasné stanovisko/Favourable opinion
Nesouhlasné stanovisko/Unfavourable opinion

Fakultní nemocnice Hradec Králové
Etická komise
Sokolská 581
500 05 Hradec Králové
e-mail: etikom@fnhk.cz



Datum/Date: 20May2021

MUDr. Jiří Vortel, předseda EK
Signature of Chairperson of the EC

Додаток 2. Скрипт для R

```

# load in a report table
setwd("C:/Users/verom/Desktop/cfDNA/NB october 2021/OneDrive_1_01-12-2021")
myTab2 <- read.table("methylation_frequency2.tsv.txt", sep="\t",
header=TRUE)
dim(myTab2)
mean(myTab2$methylated_frequency)
sd(myTab2$methylated_frequency)
sum(myTab2$called_sites)
names(myTab2)
class(myTab2$methylated_frequency)
boxplot(myTab2$methylated_frequency)

quantile(myTab2$methylated_frequency, probs = c(0, 0.25, 0.5, 0.75, 1))

boxplot(myTab2$methylated_frequency, main="cfDNA methylation pilot run",
ylab="methylation", ylim=c(0, 1.5), las=1)

class(myTab2$chromosome)

library(dplyr)
sample<-filter(myTab2, chromosome=='chr1' & chromosome=='chr2' &
chromosome=='chr3' & chromosome=='chr4' & chromosome=='chr5')

boxplot(myTab2$methylated_frequency ~ myTab2$chromosome, main="cfDNA
methylation pilot run", ylab="methylation", xlab="chromosome", ylim=c(0,
1.5), las=1)

library(ggplot2)
##ggplot of methylation values
ggplot(myTab2, aes(y=methylated_frequency, x=chromosome, col=chromosome))
+
  geom_point(position=position_jitter(h=0, w=0.2), alpha = 0.1, size = 1)
+
  ylab("methylation") +
  xlab("chromosome") +
  ggtitle("cfDNA methylation pilot run") +
  theme_bw() + theme(axis.text.x = element_text(angle = 45, vjust = 0.5,
hjust=1)) +
  geom_hline(yintercept=0.766, linetype="dashed", color = "red", size=1)

##histogram of values
ggplot(myTab2, aes(methylated_frequency)) +
  geom_bar() +
  ggtitle("histogram") +
  theme_bw()

##histogram of coverage
ggplot(myTab2, aes(called_sites)) +
  geom_bar() +
  ggtitle("histogram of coverage") +
  xlab("coverage") +
  theme_bw()
ggsave("coverage.pdf")

```

```
ggplot(myTab2, aes(chromosome, methylated_frequency, col=chromosome)) +  
  geom_bar(stat = "summary", fun = "mean", fill="grey82") +  
  geom_point(position=position_jitter (h=0, w=0.2), alpha = 0.1, size = 1)  
+  
  ylab("methylation") +  
  xlab("") +  
  ggtitle("cfDNA methylation pilot run") +  
  theme(plot.title = element_text(vjust = 0.5, hjust = 1)) +  
  theme_bw() + theme(axis.text.x = element_text(angle = 45, vjust = 0.5,  
hjust=1)) +  
  geom_hline(yintercept=0.766, linetype="dashed", color = "red", size=1)  
  
ggsave("graph1.pdf")
```

Додаток 3. Маркери для The Agilent TapeStation

