

ВИЗНАЧЕННЯ МІШЕНЕЙ ДЛЯ ПОШУКУ ПІД ЧАС АНАЛІЗУ НЕКОДУЮЧИХ ДІЛЯНОК ДНК ЛЮДИНИ

Вінницький національний технічний університет

Анотація

Робота присвячена обчислювальному аналізу некодуєчих ділянок (НКД) ДНК людини. Розглянуто ключові мішені для пошуку, зокрема регуляторні елементи (промотори, енхансери) та гени некодуєчих РНК. Описано важливість ідентифікації консервативних мотивів та епігенетичних маркерів (метилування ДНК, модифікації гістонів) для розуміння їхньої функціональної ролі. Пошук таких ділянок буде корисним як для фундаментальної науки, так і для задач прикладної медицини.

Ключові слова: некодуєча ДНК, пошук мотивів, регуляторні елементи, епігенетичні маркери, некодуєчі РНК, біоінформатика.

Abstract

The paper is devoted to computational analysis of non-coding regions (NCR) of human DNA. Key targets for search are considered, in particular regulatory elements (promoters, enhancers) and non-coding RNA genes. The importance of identifying conservative motifs and epigenetic markers (DNA methylation, histone modifications) for understanding their functional role is described. The search for such regions will be useful both for basic science and for problems of applied medicine.

Keywords: non-coding DNA, motif search, regulatory elements, epigenetic markers, non-coding RNA, bioinformatics.

Вступ

Аналіз некодуєчих ділянок ДНК (НКД ДНК) є важливим напрямком досліджень, оскільки ці ділянки, які колись помилково називали «сміттевою ДНК», відіграють критичну роль у регуляції генів та функціонуванні геному. НКД ДНК складають більшу частину геному еукаріотів (у людини до 98–99 %) [1].

Проведення аналізу некодуєчих ділянок ДНК [2] має на меті вирішення таких завдань:

- розкриття регуляторної картини геному, щоб зрозуміти, як і коли вмикаються та вимикаються гени, що кодують білки;
- ідентифікація функціональних елементів, що передбачає виявлення послідовностей, які виконують регуляторні функції (наприклад, енхансери, промотори, ізолятори), або кодують некодуєчі РНК (нкРНК), такі як мікроРНК (міРНК) та довгі некодуєчі РНК (днкРНК);
- розуміння механізмів організації геному, що необхідно для дослідження ролі НКД ДНК у підтримці структури хромосом та стабільності геному (наприклад, вплив на теломери);
- пошук генетичних причин захворювань через виявлення мутацій або варіацій у НКД ДНК, які можуть бути пов'язані з різними хворобами, включно зі спадковими захворюваннями, раком та хронічними хворобами.

Метою цієї роботи є визначення найважливіших інформативних ділянок в некодуєчій частині ДНК людини, які будуть виступати мішенями для пошуку в процесі комп'ютеризованого аналізу послідовності ДНК.

Інформативність некодуєчих ділянок ДНК

Аналіз НКД ДНК дозволяє отримати низку цінних відомостей, що мають застосування як у фундаментальній науці, так і в медицині, зокрема індивідуалізованій.

У галузі генетичної регуляції це може забезпечити ідентифікацію регуляторних послідовностей, визначення ділянок, які контролюють експресію генів (наприклад, мутації в енхансерах можуть призводити до неправильної активації або пригнічення сусідніх генів) [3].

Виявлення базових ділянок для некодуєчих РНК (нкРНК) також є перспективною задачею, адже багато НКД ДНК є джерелом синтезу нкРНК (наприклад, мікроРНК), які відіграють ключову роль у

регуляції експресії генів, впливаючи на стабільність і трансляцію мРНК [4].

У медицині дані, отримані з аналізу НКД ДНК можуть бути використані при діагностиці та профілактиці захворювань. Мутації в некодуючих ділянках можуть бути асоційовані зі схильністю до поширених захворювань (діабет другого типу, серцеві хвороби, хвороба Альцгеймера, астма) та спадковими синдромами. Виявлення цих змін дозволяє проводити ранню профілактику та вживати профілактичних заходів. Аналіз НКД ДНК допомагає скласти індивідуальні рекомендації щодо способу життя, харчування та фізичної активності на основі унікальних генетичних особливостей людини.

У фундаментальній та системній біології дослідження НКД ДНК допомагає зрозуміти, як організовані та еволюціонують геноми різних видів організмів, що сприяє розумінню еволюції геному.

Некодуючі РНК, що продукуються НКД ДНК, можуть регулювати «внутрішній годинник» організму, що важливо для розуміння циркадних ритмів та їхньої кореляції з хворобами, що є предметом вивчення хронобіології.

Деякі ділянки НКД ДНК, які містять повторювані послідовності, використовуються для встановлення батьківства та інших видів генетичного споріднення.

Мішені для пошуку при аналізі НКД ДНК

Аналізі НКД ДНК часто зводиться до пошуку різноманітних патернів (послідовностей) та епігенетичних маркерів, які вказують на їхню функціональну значущість у регуляції генів і підтримці структури геному. Основні значимі патерни та маркери, можна розділити на дві великі групи: патерни на рівні послідовності ДНК та епігенетичні маркери.

Патерни на рівні послідовності ДНК (функціональні елементи)

Це специфічні послідовності нуклеотидів у НКД ДНК, які слугують "місцями посадки" для білків або є матрицями для синтезу РНК. Серед них виділяють три групи:

1. Регуляторні послідовності (цис-елементи), які включають промотори (послідовності, розташовані безпосередньо перед геном, які слугують місцем початку транскрипції РНК), енхансери (підсилювачі) та сайленсери (пригнічувачі), ізолятори (послідовності, які відокремлюють регуляторні елементи одного гена від сусідніх, запобігаючи небажаній взаємодії).

Енхансери та сайленсери – це віддалені ділянки, які при згинанні нитки ДНК, контактують із промотором гена, посилюючи або пригнічуючи його активність. Їх шукають за певними короткими консенсусними послідовностями (TFBS – ділянки зв'язування факторів транскрипції [6]).

2. Ділянки, що кодують некодуючі РНК (нкРНК) [7]: основа для синтезу мікроРНК, або міРНК (послідовностей, які можуть утворювати шпилькоподібну структуру – прекурсор для міРНК, які потім регулюють експресію генів) та довгі некодуючі РНК, або днкРНК (довгі транскрибовані ділянки, які не кодують білок і регулюють генну активність іншими способами, наприклад, беруть участь у хроматиновому ремоделюванні).

3. Повторювані послідовності (репетитивні елементи), до яких належать мікросателіти та теломерні послідовності [8]. Мікросателіти (Short Tandem Repeats, STRs) – це короткі тандемні повтори нуклеотидів (наприклад, CACACA...). Вони використовуються як важливі генетичні маркери для встановлення спорідненості особин та в популяційній генетиці. Теломерні послідовності – це спеціалізовані повтори на кінцях хромосом, що підтримують їхню стабільність (наприклад, мотив TTAGGG у людини).

Епігенетичні маркери (модифікації)

Епігенетичні маркери – це хімічні модифікації ДНК або асоційованих із нею білків (гістонів), які не змінюють саму нуклеотидну послідовність, але впливають на активність НКД ДНК [9]. Вони є ключовими маркерами активності регуляторних елементів.

1. Пошук ділянок метилування ДНК – додавання метильних груп до цитозину (зазвичай у послідовності CpG). Високий рівень метилування в промоторних областях, як правило, є маркером пригніченого (неактивного) гена. Низький рівень метилування або його відсутність часто асоціюється з активними регуляторними елементами (наприклад, активними промоторами).

2. Модифікації гістонів – білків, навколо яких упаковується ДНК, утворюючи хроматин. Модифікації "хвостів" гістонів (ацетилювання, метилування, фосфорилування) слугують маркерами стану хроматину. Маркери активного хроматину – це ацетилювання гістонів (H3K27ac, H3K9ac) або певне метилування (H3K4me1, H3K4me3), які часто вказують на активні енхансери та промотори.

Маркери репресованого хроматину, наприклад, метилування H3K27me3 чи H3K9me3, вказують на інактивовані або пригнічені ділянки геному.

3. Ділянки ДНК, чутливі до ДНКаз I – це ті ділянки, що активно використовуються (наприклад, активні промотори чи енхансери), є відкритими і менш щільно упакованими. Це робить їх чутливими до розщеплення ферментом ДНКазою I. Виявлення таких "відкритих" ділянок є потужним маркером функціонально активної НКД ДНК.

4. Просторова організація геному може бути опосередковано встановлена шляхом пошуку патернів взаємодії між віддаленими НКД ДНК (наприклад, між енхансером та промотором), які свідчать про те, що ці елементи співпрацюють для регуляції експресії гена.

Висновки

Аналіз некодуючих ділянок ДНК людини націлений на виявлення функціональних елементів геному, які не пов'язані безпосередньо із синтезом білків. Ключовими мішенями є регуляторні послідовності, такі як промотори та енхансери, що контролюють активність генів. Ці ділянки ідентифікують шляхом пошуку специфічних коротких мотивів, які слугують місцями зв'язування для факторів транскрипції. Іншою важливою мішенню є гени, що є основою для синтезу некодуючих РНК (наприклад, мікроРНК), які виконують власні регуляторні функції. Окрім самої послідовності ДНК, критичний пошук зосереджений на епігенетичних маркерах, таких як метилування ДНК та модифікації гістонів, оскільки вони вказують на активний чи пригнічений стан цих регуляторних елементів. Зрештою, пошук також включає ідентифікацію генетичних варіантів у цих функціональних ділянках для розуміння їхнього зв'язку зі схильністю до захворювань, що може бути використано в індивідуалізованій медицині.

СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

1. Ludwig M. Z. Functional evolution of noncoding DNA. *Current opinion in genetics & development*. 2002. No. 12 (6). P. 634–639.
2. Shabalina S. A., Spiridonov N. A. The mammalian transcriptome and the function of non-coding DNA sequences. *Genome biology*. 2004. No. 5 (4). Art. 105. 8 p.
3. Shanmugam A., Nagarajan A., Pramanayagam S. Non-coding DNA-a brief review. *Journal of Applied Biology and Biotechnology*. 2017. No. 5. P. 42–47.
4. López-Urrutia E., Montes L. P. B., de Guevara Cervantes D. L., Pérez-Plasencia C., Campos-Parra A. D. Crosstalk between long non-coding RNAs, micro-RNAs and mRNAs: deciphering molecular mechanisms of master regulators in cancer. *Frontiers in oncology*. 2019. No. 9. Art. 669. 8 p.
5. Elliott K., Larsson E. Non-coding driver mutations in human cancer. *Nature Reviews Cancer*. 2021. Vol. 21 (8). P. 500–509.
6. Yang L., Zhou T., Dror I., Mathelier A., Wasserman W. W., Gordân R., Rohs R. TFBSshape: a motif database for DNA shape features of transcription factor binding sites. *Nucleic acids research*. 2014. Vol. 42 (D1). P. D148–D155.
7. Paraskevopoulou M. D., Hatzigeorgiou A. G. Analyzing miRNA–lncRNA interactions. *Long non-coding RNAs: methods and protocols*. New York, NY : Springer New York, 2016. P. 271–286.
8. Shapiro J. A., von Sternberg R. Why repetitive DNA is essential to genome function. *Biological Reviews*. 2005. No. 80 (2). P. 227–250.
9. Al Aboud N. M., Tupper C., Jialal I. Genetics, Epigenetic Mechanism. *StatPearls*. StatPearls Publishing, Treasure Island (FL), 2025. URL: <https://europepmc.org/article/nbk/nbk532999>
10. Richa R., Sinha R. P. Hydroxymethylation of DNA: an epigenetic marker. *EXCLI journal*. 2014. Vol. 13. P. 592–610.

Каплунський Олександр Андрійович – аспірант гр 163-23а, кафедра біомедичної інженерії та оптико-електронних систем, Вінницький національний технічний університет, м. Вінниця, email: okaplunskyi@vntu.edu.ua

Kaplunkkyi Oleksandr A. – postgraduate student of gr 163-23a, department of biomedical engineering and optical-electronic systems, Vinnytsia National Technical University, Vinnytsia. email: okaplunskyi@vntu.edu.ua