

УДК 577.245

ПЕРВИННИЙ АНАЛІЗ РЕЗУЛЬТАТІВ ВСЕГЕНОМНОГО ПОШУКУ ГЕНІВ ВІДПОВІДІ НА ДІЮ ІНТЕРФЕРОНУ АЛЬФА

О. О. ДРАГУЩЕНКО, Б. Т. ТОКОВЕНКО, М. Ю. ОБОЛЕНСЬКА

Інститут молекулярної біології і генетики НАН України, Київ;
e-mail: olena.imbig@gmail.com

У роботі проведено первинний аналіз потенційних генів відповіді на дію ІФНа, які було виявлено у процесі пошуку їх в геномі щура за допомогою розробленої в нашій лабораторії програми COTRASIF. З цією метою застосовано веб-інструменти *FatiGO* і *GOTM*, з використанням системи ІНОР проведено пошук у літературі, гени класифіковані за концептуальною схемою «Онтології генів» (ОГ, англ. «Gene Ontology») та ранжировані за першочерговістю їх для експериментальної перевірки.

Базуючись на результатах проведеного аналізу, серед 162 генів імовірної первинної відповіді на дію ІФНа виокремили групу із 61 гена, для яких існують експериментальні підтвердження відповіді їх на дію ІФНа; для інших генів з нашого переліку (101) таких підтверджень не знайдено, тобто вони «невідомі» як такі, що регульовані інтерфероном. На основі функціонального аналізу «невідомих» генів за схемою «Онтології генів» виявлено, що ця група є збагаченою на гени функціональної категорії «компонент синапсу». Ця категорія представлена трьома генами *Rattus norvegicus*: *PRKCA-binding protein*, *NMDAR-L*, *GABA_A – receptor subunit alpha-2*. Серед 162 генів імовірної первинної відповіді на дію ІФНа виявлено збагачену категорію «іммунна відповідь», що містить ген *Mb11*, який також належить до групи «невідомих». Ці гени обрано першими на черзі для подальшої експериментальної перевірки регуляції їх інтерфероном.

Ключові слова: інтерферон альфа, ймовірні гени відповіді, класифікація генів відповіді, «Онтологія генів», *FatiGO*, *GOTM*.

Інтерферон альфа (ІФНа) відомий як цитокін, що регулює значну кількість процесів в організмі людини. Найбільш широко вивченими є антипроліферативна, антивірусна та імуномодельююча активність ІФНа [1, 2]. Інші регуляторні властивості ІФНа вивчено мало, хоча використання ІФНа у клінічній практиці, неоднозначність відповіді пацієнтів на цей препарат та його побічні ефекти вимагають досконаліших знань. Одним із підходів для розширення наших уявлень щодо функціональної активності ІФНа може бути визначення в геномі будь-якого організму потенційних генів первинної відповіді на дію ІФНа. До останніх відносять гени, які мають у своїх промоторах ділянку ISRE (interferon stimulated response element), консервативний регуляторний елемент інтерферон-стимульованих генів (interferon stimulated genes, ISGs). Ділянка ISRE є сайтом зв'язування транскрипційного фактора ISGF3 (interferon stimulated gene factor 3), який утворюється в цитоплазмі внаслідок взаємодії ІФНа зі своїм рецептором і потрапляє до ядра.

У нашій лабораторії раніше було розроблено інструмент COTRASIF (conservation-aided transcription-factor-binding site finder) для всеге-

номного пошуку еволюційно-консервативних сайтів зв'язування транскрипційних факторів у промоторах генів еукаріот [3]. Серед усіх 17 725 протейнкодуєчих генів *Rattus norvegicus* було виявлено 162 гени, у промоторі яких є сайт ISRE та які можуть бути потенційними генами первинної відповіді на дію ІФНа (повні списки генів доступні онлайн за адресою <http://biomed.org.ua/COTRASIF/supplement.html>).

Метою нашої роботи було провести аналіз цих генів за даними літератури, виявити ті, для яких існують відомості щодо їхньої відповіді на дію ІФНа, та «невідомі» гени, для яких такі відомості відсутні, а також провести класифікацію «невідомих» генів за схемою «Онтології генів» (ОГ, «Gene Ontology») та визначити першочергові гени для експериментальної перевірки.

На основі функціонального аналізу раніше «невідомих» генів за схемою ОГ виявлено, що ця група є збагаченою на гени функціональної категорії *компонент синапсу*. Також виділено низку генів *Rattus norvegicus* (*Mb11*, *PRKCA-binding protein*, *NMDAR-L*, *GABA_A – receptor subunit alpha-2*) для подальшої першочергової експериментальної перевірки регуляції їх інтерфероном.

Матеріали і методи

Пошук експериментальних доказів регуляції генів інтерферонами в джерелах літератури

Оскільки ми проводимо саме всегеномний пошук генів первинної відповіді на ІФН α , а також враховуючи, що система інтерферону достатньо добре досліджена, то наявність експериментальних відомостей для частини генів з нашого переліку щодо здатності їх відповідати на дію ІФН α є прямим підтвердженням того, що інструмент COTRASIF працює «правильно». У свою чергу, відсутність таких відомостей для інших генів з нашого переліку розкриває широкі можливості для експериментального виявлення «нових» генів первинної відповіді на дію ІФН α . Як наслідок, результат аналізу джерел літератури спрямовує роботу експериментатора.

Отже, важливість пошуку літератури в цьому разі безсумнівна, а його раціональна організація суттєво прискорює сам пошук, що дозволяє швидше досягти бажаного результату. Для цього серед існуючих на даний момент програм, біоінформативних проектів, електронних бібліотек та різноманітних пошукових систем були обрані мінімальні за кількістю, які забезпечували б максимально якісний пошук літератури відповідної тематики.

Структура проведеного нами пошуку мала наступний вигляд (рис. 1):

Кожен ген має власний унікальний ідентифікатор у базі даних анотацій геномів Ensembl. Анотація гена, крім фактичної інформації про ген, містить також посилання на зовнішні ресурси, зокрема на UniGene. Проект UniGene групує транскрипти за ознакою подібності послідовності. Групування проводиться між багатьма геномами одночасно. Таким чи-

ном, кожен запис у базі даних UniGene – це видонезалежна інформація про ген (подібність на рівні амінокислотної послідовності, рівні експресії гена, геномного розташування і т.д.). Крім наведеної інформації UniGene містить також перелік найчастіше вживаних назв гена. Це дозволяє за ідентифікатором UniGene з'ясувати, як у літературі найчастіше називають потрібний нам ген у різних видів. Далі за допомогою пошукової системи ІНОР (Information Hyperlinked Over Proteins, <http://www.ihop-net.org/UniPub/iHOP/>) та за назвою гена проводили пошук джерел літератури, в яких одночасно згадується одна із назв-ідентифікаторів досліджуваного гена та одна із назв інтерферону (наприклад, IFN або interferon). Подальший перегляд контексту одночасного вживання двох термінів дозволяв визначити, чи йде мова про експериментальне підтвердження регуляції експресії досліджуваного гена інтерфероном. Пошукова система ІНОР здійснює автоматичний лексичний аналіз тез публікацій, які викладені в базі даних «Pubmed». Під час лексичного аналізу кожна біологічна одиниця максимально однозначно ідентифікується, та встановлюються зв'язки між різними біологічними одиницями (генами, протеїнами тощо). Тобто ІНОР формує «мережу» генів та протеїнів на основі наукової літератури і дозволяє пов'язати гени та протеїни із фенотипами, патологіями та функціями. Така «мережа» є легким способом візуалізації зв'язків та доступу до більш ніж 10 000 000 рефератів у «Pubmed» [4].

У тому разі, якщо пошук безпосередньо через ІНОР не давав позитивного результату, виконували розширений пошуковий запит, який можна сформувати за допомогою ІНОР для пошуку в «Pubmed» та «Google Scholar» за ключовими словами «назва гена/продукту» + «інтерферон» (причому «назва гена» – склад-

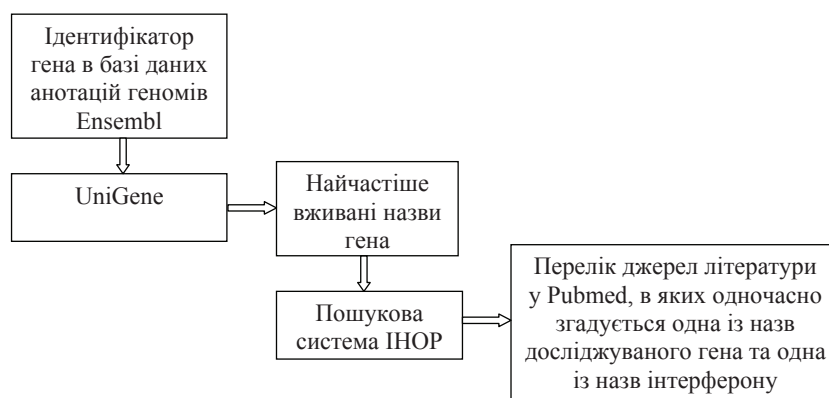


Рис. 1. Схема пошуку в літературі «відомих» генів відповіді на дію ІФН α

ний вираз, що містить всі відомі ІНОР назви досліджуваного гена).

Якщо для певного гена було знайдено відповідне літературне підтвердження його регуляції інтерфероном, то такий ген вносили до переліку «відомих», в іншому ж разі ген вносили до переліку «невідомих».

Функціональний аналіз обох груп генів проводили за категоріями «Онтології генів», використовуючи для цього два веб-інструменти – FatiGO і GOTM.

Функціональний аналіз генів за допомогою концептуальної схеми «Онтологія генів»(ОГ)

Для функціонального аналізу генів користувалися концептуальною схемою ОГ, що формалізує ознаки генів та їхніх продуктів для багатьох організмів. ОГ являє собою не звичайну базу даних генних послідовностей і каталог генних продуктів, а описує поведінку генних продуктів у контексті функціонування клітини як цілісної системи. На даний момент в базі даних ОГ анотовано гени та їхні продукти організмів людини, миші, щура, корови, курки, дрозофіли, арабідопсиса, черв'яка та дріжджів. Всі гени кожного з цих організмів належать до низки певних категорій ОГ, і за кожним геном закріплюється список категорій, до яких він належить. Доступ до таких списків забезпечується значною кількістю існуючих на даний момент веб-інструментів (ОВО-Edit, Ontology Evolution Explorer, Gene Ontology Explorer, Generic GO Term Mapper та ін., повний перелік та описання веб-інструментів є за адресою <http://www.geneontology.org/GO.tools.shtml>).

Концептуальна схема ОГ представлена у вигляді деревоподібної структурованої мережі категорій, що описують властивості генів та їхніх продуктів. Існують три основні незалежні категорії ОГ: **клітинний компонент** (компаратмент клітини чи зовнішньоклітинного оточення, де функціонує продукт гена), **молекулярна функція** (елементарна участь генних продуктів у реакціях на молекулярному рівні, таких як зв'язування або каталіз) та **біологічний процес** (послідовності молекулярних реакцій, що мають певний початок і кінець та відбуваються в ході функціонування інтегрованих живих одиниць: клітин, тканин, органів та організмів). Ці три основні категорії розділяються на вужчі та більш функціонально орієнтовані.

Кожна категорія ОГ має своє ім'я (слово або словосполучення), власний номер, джерело цитування. Будова ОГ – *спрямований ациклічний граф*, де кожна категорія має певні

відносини з однією чи багатьма іншими категоріями. Словник ОГ сконструйований без врахування видоспецифічності – він містить у собі категорії для прокариот та еукаріот, однак багатоклітинних організмів, тобто є універсальним.

Для функціональної характеристики генів, одержаних внаслідок проведення певних широкомасштабних експериментів (наприклад, з мікромасивами генів), виявляють наявність збагачених категорій ОГ, тобто таких категорій, в яких гени зустрічаються з вірогідністю, що перевищує випадкову, і згруповані за певною функціональною ознакою. Наявність збагаченої категорії у вибірці генів свідчить про групування генів за певним критерієм.

Збагачені категорії ОГ можна виявити за допомогою веб-інструментів FatiGO, GOTM, BayGO та ін. Візуальним результатом такого розподілу є *спрямований ациклічний граф*, що відображає ієрархію категорій, до яких належить той чи інший ген. Один ген може одночасно належати до декількох категорій різних рівнів ієрархічного підпорядкування. В роботі використано два веб-інструменти – FatiGO та GOTM.

FatiGO – це веб-інструмент, за допомогою якого порівнюють дві групи генів і визначають ті гени, розподіл яких в цих групах відрізняється із статистичною значущістю. Основою цього функціонального аналізу є застосування критерію Фішера для спряжених таблиць розмірності 2 x 2. Для побудови спряжених таблиць обираються дві групи генів, які складають цільову та загальну вибірки. У нашому випадку цільова вибірка – це гени первинної відповіді на дію ІФНа, а загальна вибірка – це всі протеїнкодуєчі гени геному щура. Шляхом порівняння двох вибірок ми визначаємо, на гени яких категорій ОГ збагачена цільова вибірка. Гени у двох групах характеризуються за своєю приналежністю/неприналежністю до певної категорії ОГ. Результати аналізу коригуються для множинної вибірки з метою отримання коригованого *P-value*.

Кінцевий результат подається у текстовому форматі або HTML, а також може бути представлений у вигляді деревоподібної мережі категорій ОГ, пов'язаної як зі списком генів, так і з числом генів, анотованих певною категорією ОГ [5].

GOTM (Gene Ontology Tree Machine) – це ще один веб-інструмент, за допомогою якого можна провести статистичний аналіз результатів пошуку збагачених категорій ОГ серед генів певної вибірки відносно загальної групи.

Пошук збагачених категорій ОГ базується на гіпергеометричному тесті. Для цього веб-інструментом GOTM розраховуються три параметри: O, E та R, де O (observed) – кількість генів нашої вибірки, що належать до певної категорії ОГ; E (expected) – очікувана кількість генів з нашої вибірки, що мають належати до тієї ж категорії ОГ. Розрахунки очікуваної кількості генів (E) в нашій цільовій вибірці базуються на твердженні, що в цій вибірці міститься такий самий відсоток генів, що належать до певної категорії ОГ, як і в загальній вибірці, тобто в усьому геномі. R (ratio) – ступінь збагаченості для даної категорії ОГ в нашій вибірці – розраховується як відношення отриманої кількості генів у певній категорії ОГ до очікуваної ($R=O/E$). Якщо параметр $R > 1$, то є підстави вважати категорію ОГ в нашій вибірці збагаченою. У цьому разі додатково розраховується значення P – вірогідність збагачення для даної категорії ОГ. Категорія ОГ вважається збагаченою, якщо $P < 0,01$ [6].

Крім того, програмне забезпечення GOTM дозволяє побудувати дерево онтології генів (GOTree) – деревоподібну структуру ієрархічно підпорядкованих збагачених та незбагачених категорій ОГ, представлених у нашій вибірці генів. Таким чином, візуалізується зв'язок між даними категоріями ОГ, тобто ми отримуємо наочний результат аналізу.

Результати та обговорення

Первинний аналіз результатів всегеномного пошуку генів відповіді на дію ІФНа

Внаслідок пошуку експериментальних доказів регуляції 162 генів інтерфероном було виявлено, що для 61 з них є такі підтвердження в літературі, а для інших 101 таких підтверджень не знайдено. Ми провели функціональний аналіз за категоріями ОГ отриманих трьох груп генів, що містять 162, 61 та 101 ген, застосувавши два веб-інструменти – FatiGO і GOTM. Ці інструменти використовують різні значення P -value для визначення вірогідності збагачення категорій ОГ: в GOTM $P < 0,01$, а

в FatiGO – P -value, кориговане до множинної вибірки.

Функціональний аналіз імовірних генів-мішеней ІФНа за категоріями ОГ за допомогою веб-інструмента FatiGO

Аналіз загальної групи генів. Серед загальної групи (162 гени) веб-інструментом FatiGO виявлено лише одну вірогідно збагачену категорію ОГ – «іммунна відповідь», що ієрархічно підпорядковується категорії **біологічний процес**. До цієї категорії належать 9 генів. Для восьми з них відома регуляція їх інтерферонами I типу, а для гена *Mbl1* (Mannose-binding protein A, ENSRNOG00000011706) опублікованих експериментальних даних немає. *Mbl1* анотований у GOTM як ген вродженої імунної відповіді, має дві ділянки ISRE. Білок, що є продуктом цього гена, кальційзалежно активує систему комплементу.

Одержана нами єдина вірогідно збагачена категорія ОГ «іммунна відповідь» відповідає очікуванням, оскільки найкраще охарактеризовані саме імуностимулювальні ефекти ІФНа.

Аналіз групи генів, для яких не знайдено в літературі підтверджень щодо регуляції їх інтерфероном. Аналізуючи гени із переліку «невідомих», який налічує 101 ген, за допомогою інструмента FatiGO виявлено єдину збагачену категорію **компонент синапсу**, що ієрархічно підпорядковується категорії ОГ **клітинний компонент**. До цієї категорії належать 3 гени (таблиця).

Продукти зазначених вище генів анотовані як складові частини синапсу, що знаходяться в місцях контакту між двома нейронами або між нейроном та м'язовим волокном чи гліальною клітиною.

PRKCA-binding protein (PICK1) – присутній у збуджуваних синапсах протеїн, який має PDZ-домен, і через нього зв'язується з протеїнкіназою C альфа, функціонує як якорний протеїн, що специфічно спрямовує PRKCA до мітохондрій [7, 8]. Також він функціонує як адаптерний протеїн: забезпечує зв'язування та субклітинну локалізацію низки мембранних протеїнів, взаємодіє із глутаматним рецепто-

Гени категорії клітинний компонент

№	Ідентифікатор гена	Назва гена
1	ENSRNOG00000011507	PRKCA-binding protein (Protein kinase C-alpha binding protein)
2	ENSRNOG00000005723	N-methyl-D-aspartate (NMDA) receptor subunit 3A precursor
3	ENSRNOG00000002349	Gamma-aminobutyric acid (GABA _A) receptor subunit alpha-2

ром, переносниками моноамінів крізь плазматичну мембрану, потенціалнезалежними натрієвими каналами, а також може націлювати PRKCA до цих мембранних протеїнів і, таким чином, регулювати функціонування їх [9]. За будовою – може бути мономером та гомодимером, залучений до процесу утворення рецепторних кластерів.

NMDA-рецептори (NMDARs) – іонотропні рецептори до глутамату (збуджуючого нейромедіатора), які селективно зв'язують N-метил-D-аспартат (NMDA). Вони локалізовані на постсинаптичній мембрані нейронів, складаються із двох субодиниць – NR1 та NR2, функціонують у тетрамерній формі та характеризуються високою проникністю для іонів Ca^{2+} [10–12]. Вхідний струм іонів Ca^{2+} активує низку біохімічних каскадів, що опосередковують процеси синаптогенезу та синаптичної пластичності, тобто зміни трансмембранного потенціалу у відповідь на активацію постсинаптичних рецепторів. Це вважається основним механізмом, за допомогою якого реалізується феномен пам'яті та навчання. Так, відомо, що активація NMDA-рецепторів підвищує активність піроглутаміл пептидази II (екзопептидази) в гіпокампі щурів, тобто в ділянці мозку, що відповідає за короткотривалу пам'ять [13]. NMDA-рецептори функціонують як гетеромультимерні комплекси, до складу яких входять субодиниця NR1 та, щонайменш, одна з чотирьох (A–D) субодиниць NR2. Коекспресія додаткової субодиниці NR3A модулює активність NMDA-рецептора. Гетеромерний рецептор, що містить NR3A (NR1/NR2/NR3A), відрізняється за своїми властивостями, а саме: має знижену проникність для Ca^{2+} та низьку чутливість до блокування магнієм (Mg^{2+}) [14].

Під час розвитку організму збуджуючі синапси в центральній нервовій системі підлягають ремоделюванню. Одним з таких прикладів є переключення у постнатальний період з форми NMDA-рецептора, що розвивається, на «дорослу» форму рецептора. На ранніх етапах розвитку організму глутаматергічні синапси вже містять NMDA-рецептори, і протягом певного періоду, що залежить від активності нервової системи, змінюється субодиничний склад цих рецепторів, а також їхні кінетичні властивості та міжсинаптична взаємодія. NR3A утворюється тільки протягом короткого періоду постнатального розвитку, після чого відбувається вилучення субодиниці NR3A з поверхні мембрани шляхом ендоцитозу, і синаптична мережа стає сталою [15]. Таким чи-

ном, модулюючи активність NMDA-рецептора, ІФНа може бути залучений до функціонування збуджуючих синапсів, регуляції процесу росту аксонів та синаптогенезу.

GABA_A-рецептор (Gamma-aminobutyric acid receptor) – лігандзалежний хлорний канал, рецептор до γ -аміномасляної кислоти (ГАМК) – інгібіторного нейромедіатора. Після зв'язування ГАМК з рецептором іонний канал відкривається, і вхідний струм іонів Cl^- зумовлює гіперполяризацію постсинаптичної мембрани нейрона, знижуючи його здатність до збудження. GABA_A-рецептор має сайти зв'язування низки психоактивних речовин, таких як бензодіазепін, пікротоксин, барбітурати, етанол, нейрональні стероїди та ін., які посилюють дію природного ліганду, ГАМК, і, таким чином, зумовлюють седативний ефект [16].

GABA_A-рецептор є гетеропентамером. У більшості випадків до нього належать субодиниці 2 α , 2 β та 1 γ ($\alpha_2\beta_2\gamma$), з яких α і β є обов'язковими компонентами, а замість γ -субодиниці можуть бути ϵ , π , δ або θ . Залежно від субодиничного складу, GABA_A-рецептори виявляють різні властивості. Сайт зв'язування ГАМК локалізований між субодиницями α і β , а сайт зв'язування такого анксиолітика як бензодіазепін – між α і γ -субодиницями; зв'язування бензодіазепіну вимагає наявності амінокислотного залишку гістидину, який є в субодиницях α_1 , α_2 , α_3 , і α_5 . Через це бензодіазепін виявляє нижчу афінність до рецепторів, що містять α_4 та α_6 , оскільки в них гістидин заміщений на аргінін [17]. Отже, GABA_A-рецептори, що містять субодиницю α_2 , проявляють більший афінитет до деяких анксиолітиків.

Оскільки раніше не було відомостей про участь інтерферону в регуляції процесів у нервовій системі, то ці дані були дещо неочікуваними. Проте у разі лікування інтерфероном хворих на гепатит та деякі форми раку спостерігаються побічні ефекти, серед яких – головні болі, депресія та ін. [18], природа яких невідома. Одержані нами результати дають змогу експериментально перевірити нові, раніше невідомі функції ІФНа.

Аналіз групи генів, для яких знайдено в літературі підтвердження щодо регуляції їх ІФНа. Аналізуючи гени із переліку «**ВІДОМИХ**» (61 ген) за допомогою інструмента FatIGO, виявлено три збагачені категорії ОГ: *імунна відповідь, відповідь на віруси та ГТР-азна активність*, перші дві з яких ієрархічно підпорядковуються категорії **біологічний процес**, а третя – категорії **молекулярна функція**.

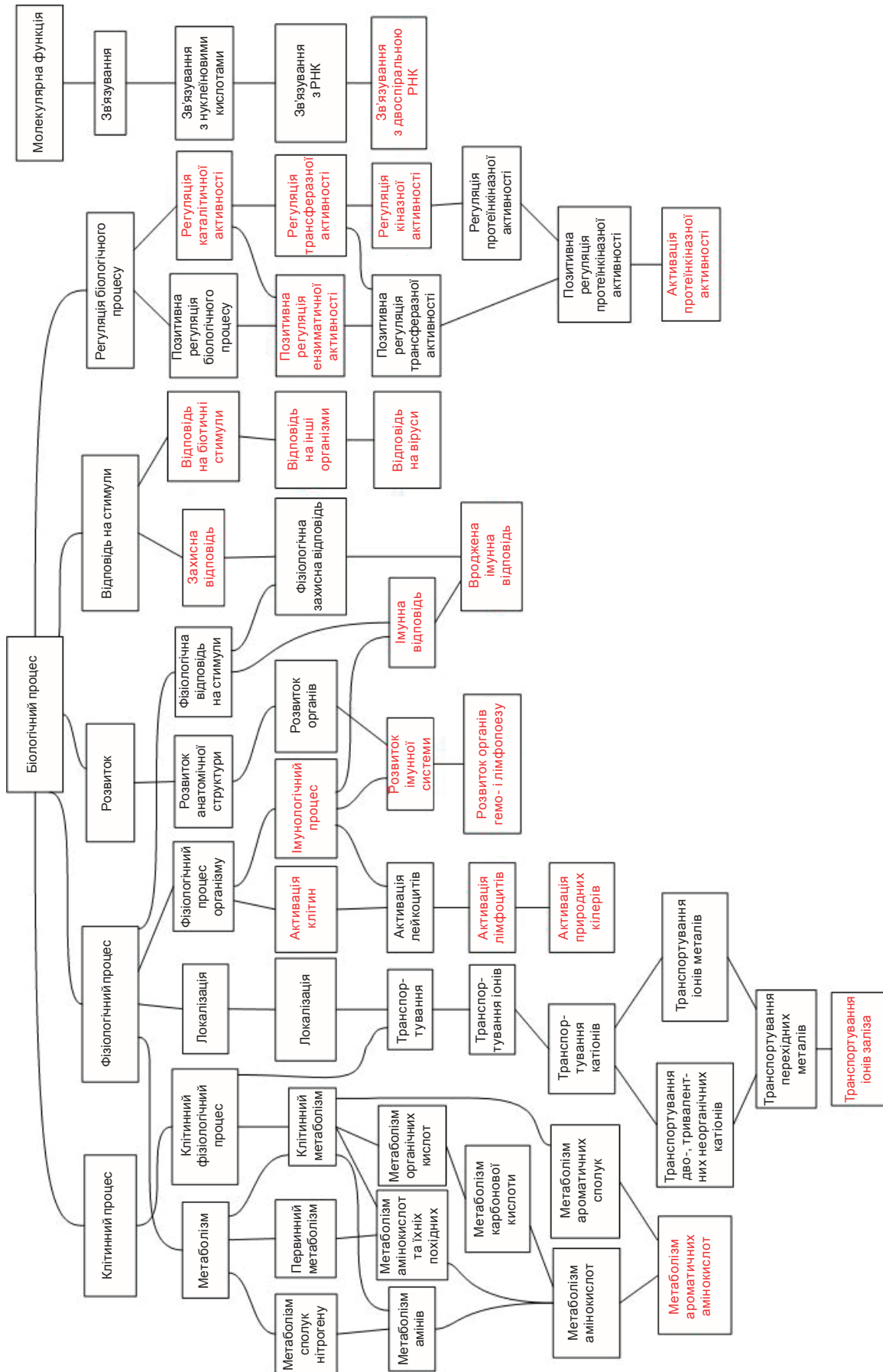


Рис. 2. Розподіл загальної групи генів імовірної відповіді на дію ІФНа за категоріями ОГ веб-інструментом GOTM (червоним кольором позначені збагачені категорії, чорним — незбагачені)

Функціональний аналіз імовірних генів-мішеней ІФНа відповідно до категорій ОГ за допомогою веб-інструмента GOTM

Серед загальної групи генів (162) інструмент GOTM виявив 20 вірогідно збагачених категорій ОГ, які ієрархічно підпорядковуються категоріям **біологічний процес** (19 категорій) та **молекулярна функція** (1 категорія) (рис. 2).

Серед 61 гена, які входять до групи «**відомих**», інструментом GOTM було виявлено 36 вірогідно збагачених категорій ОГ, що ієрархічно підпорядковуються категоріям **біологічний процес** (35 категорій) та **молекулярна функція** (1 категорія).

Аналіз групи «**невідомих**» генів інструментом GOTM не виявив жодної вірогідно збагаченої категорії ОГ.

162 гени нашого переліку містять у промоторі ISRE-елементи, і всі вони є передбаченими програмою COTRASIF генами первинної відповіді на дію ІФНа. У цій роботі проведено пошук експериментальних підтверджень нашого передбачення, і серед тих генів, що не мають такого підтвердження, обрано гені-кандидати для першочергової експериментальної перевірки.

Для цього гени класифікували за категоріями ОГ за допомогою двох веб-інструментів – GOTM і FatiGO. Кожен з них шукає вірогідно збагачені категорії та візуалізує результати пошуку, що забезпечує наочне сприйняття цього результату. Але GOTM має переваги у графічному представленні, він будує ациклічний граф, відображаючи розташування збагачених категорій ОГ відносно усіх проміжних, ієрархічно вищих категорій незалежно від їхньої збагаченості. Граф FatiGO являє собою ієрархічну приналежність збагаченої категорії безпосередньо до головної категорії ОГ, вилучаючи з поля зору усі проміжні, що не є збагаченими.

Крім того, результат розподілення вибірки генів цими двома інструментами дещо відрізняється у зв'язку з різними значеннями *P*-value, які використовуються для визначення вірогідності збагачення тієї чи іншої категорії ОГ в нашій вибірці. В цьому відношенні інструмент FatiGO має перевагу перед GOTM, оскільки в ньому *P*-value коригується до множинної вибірки, через це FatiGO обирає меншу кількість збагачених категорій, проте ставить суворіші вимоги до їхнього збагачення.

Поєднання отриманої інформації з даними мікромасив-експериментів (microarray) можна використати для реконструкції мережі

генної регуляції, що активується інтерфероном.

Таким чином, з проведеного нами аналізу результатів всегеномного пошуку генів відповіді на дію інтерферону альфа, можна дійти висновку:

1. Серед 162 генів імовірної первинної відповіді на дію інтерферону альфа виокремили групу із 61 гена, для яких існують експериментальні підтвердження відповіді їх на дію інтерферону, та 101 ген, для яких таких підтверджень не знайдено.

2. Група потенційних генів первинної відповіді на інтерферон альфа є збагаченою на гени функціональної категорії *компонент сигналу* концептуальної схеми «Онтологія генів». Передбачається, що інтерферон може брати участь у регуляції діяльності нервової системи.

3. Серед знайдених раніше невідомих потенційних генів-мішеней інтерферону альфа виявлено низку генів (*Mb11*, *PRKCA-binding protein*, *NMDAR-L*, *GABA_A – receptor subunit alpha-2*), які обрано для першочергової експериментальної перевірки регуляції їх інтерфероном.

ПЕРВИЧНИЙ АНАЛІЗ РЕЗУЛЬТАТОВ ВСЕГЕНОМНОГО ПОИСКА ГЕНОВ ОТВЕТА НА ДЕЙСТВИЕ ИНТЕРФЕРОНА АЛЬФА

Е. О. Драгущенко, Б. Т. Токовенко, М. Ю. Оболенская

Институт молекулярной биологии и генетики НАН Украины, Киев;
e-mail: olena.imbig@gmail.com

В работе проведен первичный анализ потенциальных генов ответа на действие ИФНа, которые были найдены в процессе их поиска в геноме крысы с помощью разработанной в нашей лаборатории программы COTRASIF. С этой целью были применены веб-инструменты FatiGO и GOTM, проведен поиск в литературе с использованием системы ИНОР, гены классифицированы согласно концептуальной схеме «Онтологии генов» (ОГ) и ранжированы по их первоочередности для экспериментальной проверки.

Основываясь на результатах проведенного анализа, среди 162 генов предполагаемого первичного ответа на действие ИФНа выделили группу генов (61 ген), для которых существуют экспериментальные подтверждения их ответа на действие ИФНа, а для оставшихся генов нашего списка (101) таких подтвержде-

ний не найдено (группа «неизвестных» генов). На основе функционального анализа ранее «неизвестных» генов по схеме ОГ обнаружено, что данная группа является обогащенной на гены функциональной категории «компонент синапса». Эта категория представлена тремя генами *Rattus norvegicus*: *PRKCA-binding protein*, *NMDAR-L* и *GABA_A – receptor subunit alpha-2*. Среди общей группы генов (162) обнаружено обогащенную категорию «иммунный ответ», содержащую ген *Mbll*, который также входит в число «неизвестных» генов. Перечисленные гены выбраны в качестве первых кандидатов для экспериментальной проверки их регуляции интерфероном.

Ключевые слова: интерферон альфа, потенциальные гены ответа, классификация генов ответа, «Онтология генов», FatiGO, GOTM.

INITIAL ANALYSIS OF THE RESULTS OF THE GENOME-WIDE SEARCH FOR THE GENES OF RESPONSE TO INTERFERON ALPHA

O. O. Dragushchenko, B. T. Tokovenko, M. Yu. Obolenskaya

Institute of Molecular Biology and Genetics,
National Academy of Sciences of Ukraine, Kyiv;
e-mail: olena.imbig@gmail.com

Summary

The paper is focused on the primary analysis of the potential genes of primary response to IFN α , predicted by the program COTRASIF elaborated in our laboratory. Two web-instruments FatiGO and GOTM were applied for this purpose; the literature search was carried out using IHOP; genes were classified according to the conceptual diagram «Gene Ontology» (GO) and were ranked according to their first priority for the experimental validation.

On the basis of conducted analysis 162 genes of potential primary response to IFN α were subdivided into two groups – 61 genes, for which the experimental data of their responsibility to IFN α were found and 101 genes for which such data are absent. We have performed the functional analysis of the «unknown» genes according to GO. The enriched category «Synapse part» encountering three genes of *Rattus norvegicus*: *PRKCA-binding of protein*, *NMDAR-L* and *GABA_A – receptor subunit alpha-2* were revealed. Among 162 genes the en-

riched category «Immune response» was detected with *Mbll* gene, which is reckoned among «unknown» genes. The four designated genes are chosen as the candidates for the prior validation.

Key words: interferon alpha, genes of potential response, classification of genes of response, «Gene Ontology», FatiGO, GOTM.

1. *Randall R. E., Goodbourn S.* // *J. Gen. Virol.* – 2008. – **89**, N 1. – P. 1–47.
2. *Samuel Ch. E.* // *Clin. Microbiol. Rev.* – 2001. – **14**, N 4. – P. 778–809.
3. *Tokovenko B., Golda R., Protas O. et al.* // *Nucleic Acids Res.* – 2009. – **37**, N 7. – P. 49.
4. *Hoffmann R., Valencia A.* // *Nat. Genet.* – 2004. – **36**, N 7. – P. 664.
5. *Al-Shahrour F, Diaz-Uriarte R, Dopazo J.* // *Bioinformatics.* – 2004. – **20**, N 4. – P. 578–580.
6. *Zhang B., Schmoyer D., Kirov S. et al.* // *BMC Bioinformatics.* – 2004. – **5**. – P. 16.
7. *Wang W. L., Yeh S. F., Chang Y. I. et al.* // *J. Biol. Chem.* – 2003. – **278**, N 39. – P. 37705–37712.
8. *Pan L., Wu H., Shen C. et al.* // *EMBO J.* – 2007. – **26**, N 21. – P. 4576–4587.
9. *Dev K. K., Nishimune A., Henley J. M. et al.* // *Neuropharmacology.* – 1999. – **38**, N 5. – P. 635–644.
10. *Chen P., Wyllie D.* // *Br. J. Pharmacol.* – 2006. – **147**, N 8. – P. 839–853.
11. *Furukawa H., Singh S. K., Mancusso R. et al.* // *Nature.* – 2005. – **438**, N 7065. – P. 185–192.
12. *Wrighton D. C., Baker E. J., Chen P. E. et al.* // *J. Physiol.* – 2008. – **586**, N 1. – P. 211–225.
13. *Rodríguez-Molina V., Vargas M. Á. et al.* // *Neurosci. Lett.* – 2009. – **449**, Issue 3. – P. 211–214.
14. *Perez-Otano I., Schulteis C. T., Contractor A. et al.* // *J. Neurosci.* – 2001. – **21**, N 4. – P. 1228–1237.
15. *Pérez-Otaco I., Luján R., Tavalin S. J. et al.* // *Nat. Neurosci.* – 2006. – **9**, N 5. – P. 611–621.
16. *Johnston G.* // *Pharmacol. Ther.* – 1996. – **69**, N 3. – P. 173–198.
17. *Wafford K. A., Macaulay A. J., Fradley R. et al.* // *Biochem. Soc. Trans.* – 2004. – **32**, N 3. – P. 553–556.
18. *Lindsay K. L., Davis G. L., Schiff E. R. et al.* // *Hepatology.* – 1996. – **24**. – P. 1034–1040.

Отримано 01.12.2009