



ДЕРЖАВНА СЛУЖБА  
ІНТЕЛЕКТУАЛЬНОЇ  
ВЛАСНОСТІ  
УКРАЇНИ

УКРАЇНА

(19) **UA** (11) **110790** (13) **C2**  
(51) МПК  
**C12Q 1/68** (2006.01)

## (12) ОПИС ДО ПАТЕНТУ НА ВИНАХІД

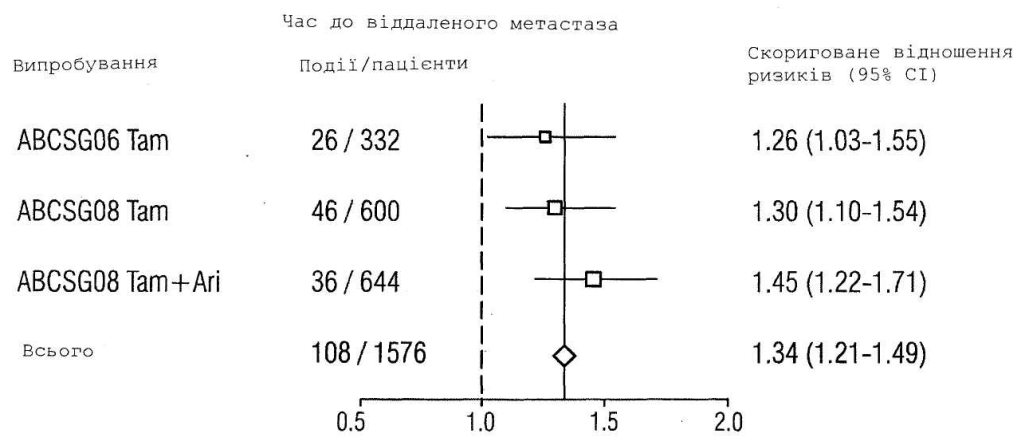
(21) Номер заявки:	<b>а 2012 12456</b>	(72) Винахідник(и):	<b>Дартманн Марайке (DE), Федер Інке Сабін (DE), Германн Матіас (DE), Хенніг Гідо (DE), Вебер Карстен (DE), фон Тьорн Крістіан (DE), Кроненветт Ральф (DE), Петри Крістоф (DE)</b>
(22) Дата подання заявки:	<b>29.03.2011</b>	(73) Власник(и):	<b>СІВІДОН ДІАГНОСТИКС ГМБХ, Nattermannallee 1, 50829 Cologne, Germany (DE)</b>
(24) Дата, з якої є чинними права на винахід:	<b>25.02.2016</b>	(74) Представник:	<b>Маслова Тетяна Михайлівна, реєстр. №61</b>
(31) Номер попередньої заявки відповідно до Паризької конвенції:	<b>10158561.0</b>	(56) Перелік документів, взятих до уваги експертизою:	<b>WO 2006/119593 A1, 16.11.2006 WO 2009/158143 A1, 30.12.2009 WO 2010/003773 A1, 14.01.2010 WO 2008/079269 A2, 03.07.2008</b>
(32) Дата подання попередньої заявки відповідно до Паризької конвенції:	<b>31.03.2010</b>		
(33) Код держави-учасниці Паризької конвенції, до якої подано попередню заявку:	<b>EP</b>		
(41) Публікація відомостей про заявку:	<b>25.12.2012, Бюл.№ 24</b>		
(46) Публікація відомостей про видачу патенту:	<b>25.02.2016, Бюл.№ 4</b>		
(86) Номер та дата подання міжнародної заявки, поданої відповідно до Договору РСТ	<b>PCT/EP2011/054855, 29.03.2011</b>		

## (54) СПОСІБ ПЕРЕДБАЧУВАННЯ РЕЦИДИВУ РАКУ МОЛОЧНОЇ ЗАЛОЗИ ПРИ ЕНДОКРИННОМУ ЛІКУВАННІ

### (57) Реферат:

Винахід належить до способу, набору для прогнозування наслідків захворювання при раку молочної залози, причому вказаний спосіб включає: (а) визначення в зразку пухлини від вказаного пацієнта рівнів експресії РНК наступних 8 генів: UBE2C, BIRC5, DHCR7, STC2, AZGP1, RBBP8, IL6ST і MGP; (b) математичне комбінування величин рівня експресії для генів вказаного набору, причому ці величини визначають в зразку пухлини з одержанням комбінованого показника, де вказаний комбінований показник указує на прогноз у вказаного пацієнта.

UA 110790 C2



Фіг. 1

## Технічна галузь

Даний винахід відноситься до способів, наборів і систем для прогнозування наслідків рака молочної залози. Конкретніше, даний винахід відноситься до прогнозування рака молочної залози на основі вимірювання рівнів експресії маркерних генів в зразках пухлини пацієнтів з раком молочної залози.

Рівень техніки, до якого відноситься винахід

Рак молочної залози є однією з основних причин смертності від раку у жінок в західних країнах. Конкретніше, тільки в США щорічно рак молочної залози забирає життя приблизно 40000 жінок, і його діагностують приблизно у 200000 жінок. Протягом останніх декількох десятиліть ад'ювантна системна терапія привела до значного збільшення виживаності при ранньому раку молочної залози. Цей клінічний досвід привів до узгоджених рекомендацій, що пропонують ад'ювантну системну терапію для переважної більшості пацієнтів з раком молочної залози (EBCAG). При раку молочної залози доступна безліч варіантів лікування, які можна застосовувати на додаток до хірургічного видалення пухлини, що зазвичай проводиться, і подальшого опромінювання ложа пухлини. Три основних і принципово різних стратегії є ендокринне лікування, хіміотерапія і лікування націленою терапією. Передумовою для лікування ендокринними засобами є експресія рецепторів гормонів в пухлинній тканині, тобто або рецептора естрогену, або рецептора прогестерону, або обох з них. Доступні декілька ендокринних засобів з різним способом дії і відмінностями в результаті захворювання при тестуванні у великих групах пацієнтів. Тамоксифен є основою ендокринного лікування впродовж останніх трьох десятиліть. Значні клінічні випробування продемонстрували, що тамоксифен значно знижує ризик рецидиву пухлини. Додатковий варіант лікування заснований на інгібіторах ароматази, які відносяться до нового класу ендокринних лікарських засобів. В протилежність тамоксифену, який є конкурентним інгібітором зв'язування естрогену, інгібітори ароматази блокують продукування самого естрогену, тим самим знижуючи стимул до зростання для позитивних по рецептору естрогену пухлинних кліток. Крім того, у деяких пацієнтів виникає рецидив, не зважаючи на ендокринне лікування, і конкретно для цих пацієнтів можуть бути корисні додаткові терапевтичні лікарські засоби. Було показано, що хіміотерапія антрациклінами, таксанами і іншими засобами є ефективною для придушення рецидиву захворювання як у позитивних по рецептору естрогену пацієнтів, так і у негативних по рецептору естрогену пацієнтів. У дослідженні NSABP-20 порівнювали тамоксифен окремо з тамоксифеном плюс хіміотерапія у негативних по вузлах позитивних по рецептору естрогену пацієнтів і показали, що комбіноване лікування було ефективнішим, ніж лікування тамоксифеном окремо. Проте дослідження IBCSG IX, в якому порівнювали тамоксифен окремо з тамоксифеном плюс хіміотерапія, не показало якої-небудь користі додавання цитотоксичних засобів. Недавно було показано, що введене системне антитіло, що направлене проти антигена HER2/neu на поверхні пухлинних клітин, знижує ризик рецидиву у декілька разів у пацієнтів зі зверхекспресуючими Her2neu пухлинами. Проте більшість, якщо не всі, з різних способів медикаментозного лікування мають численні побічні дії, які можуть значно погіршити якість життя пацієнтів (Shapiro and Recht, 2001; Ganz et al., 2002). Це робить обов'язковим вибір стратегії лікування на основі ретельної оцінки ризику у конкретного пацієнта, щоб уникнути надмірного, а також недостатнього лікування. Оскільки користь хіміотерапії є відносно великою у разі позитивних по HER2/neu пухлин і пухлин, що характеризуються відсутністю експресії HER2/neu і рецепторів естрогену (базальний тип), в порівнянні з негативними по HER2/neu і позитивних по рецепторах естрогену пухлинах (люминальний тип), найбільш важке прийняття рішення про лікування стосується люминальних пухлин, для яких класичні клінічні чинники, такі як ступінь, розмір пухлини або залучення лімфатичних вузлів, не забезпечують чіткої відповіді на питання про використання хіміотерапії. Для задоволення цієї медичної потреби були розроблені новіші молекулярні інструменти, такі як аналіз 21 гена, аналіз геномного показника ступеня і інші аналізи.

Методичні рекомендації по лікуванню зазвичай розробляються відомими в даній області експертами. У Європі методичні рекомендації St Gallen 2009 року рекомендували хіміотерапію у пацієнтів з позитивним по HER2 раком молочної залози, а також у пацієнтів з негативним по HER2 і негативним по ER захворюванням. Неясність про користь хіміотерапії виникає у разі пацієнтів з негативним по HER2 і позитивним по ER захворюванням. Щоб прийняти збалансоване рішення для суб'єкта як найбільш корисний критерій використовують вірогідність рецидиву раку. Клінічні критерії, такі як стан лімфатичних вузлів, ступінь пухлини, розмір пухлини та інші, є корисними, оскільки вони забезпечують інформацію про ризик рецидиву. Пізніше було показано, що мультигенні аналізи забезпечують інформацію краще або додатково до стандартних клінічних чинників ризику. Є загальновизнаним, що маркери проліферації,

очевидно, забезпечують основну прогностичну інформацію. Відомими прикладами цих прогностичних засобів є тест Mammaprint від Agendia, показник рецидиву від Veridex і геномний індекс ступеня, що розроблений в інституті Jules Bordet і ліцензований для Ipsogen. Всі з цих аналізів засновано на визначенні рівнів експресії щонайменше 70 генів і всі з них були розроблені для РНК, не деградованою значно фіксацією формаліном і зануренням в парафін, а виділеною із свіжої тканини (транспортованою в RNALater™). Іншим важливим мультигенним аналізом є тест показника рецидиву від Genomic Health Inc. Цей тест визначає рівень експресії 16 пов'язаних з раком генів і 5 еталонних генів після екстракції РНК з фіксованих у формаліні і занурених в парафін зразків тканини.

Проте сучасні інструменти мають недолік відсутності клінічної надійності і застосовності в найбільш важливій клінічній групі ризику, тобто у тих пацієнтів з раком молочної залози, що мають проміжний ризик рецидиву на основі стандартного клінічного параметра. Таким чином, потрібні кращі інструменти для оптимізації прийняття рішення про лікування на основі прогнозу для пацієнта. Для клінічної користі, що полягає в уникненні хіміотерапії, потрібний тест з високою чутливістю і високою негативною прогностичною силою, щоб не піддати пацієнта недостатньому лікуванню, у якого зрештою розвинеться віддалений метастаз після хірургічної операції. Що стосується потреби, що залишається, в матеріалах і способах, придатних для прийняття клінічних рішень відносно ад'ювантної терапії, даний винахід задовольняє потреби в перспективних способах прогнозування раку молочної залози на основі добре доступних клінічних і експериментальних даних.

#### Визначення

Якщо не визначено інакше, технічні і наукові терміни, використовувані в цьому документі, мають таке ж значення, яке зазвичай має на увазі фахівець в області, до якої відноситься даний винахід.

Термін "рак" не обмежується якою-небудь стадією, ступенем, гістохімічним ступенем, агресивністю або злоякісністю ураженої тканини або клітинного агрегату.

Термін "передбачення наслідків" захворювання, як використовують в рамках винаходу, означає як передбачення наслідків у пацієнта, що піддається даній терапії, так і прогноз у пацієнта, якого не піддають лікуванню. Термін "передбачення наслідків", зокрема, може відноситися до ризику розвитку у пацієнта метастаза, локального рецидиву або загибелі.

Термін "передбачення", як використовують в рамках винаходу, відноситься до індивідуальної оцінки раки або до очікуваної виживаності (OAS, загальна виживаність або DFS, виживання без захворювання) пацієнта, якщо пухлину лікують даною терапією. В протилежність цьому, термін "прогноз" відноситься до індивідуальної оцінки раку, або до очікуваної виживаності (OAS, загальна виживаність або DFS, виживання без захворювання) у пацієнта, якщо пухлина залишається без лікування.

"Результат" в рамках даного винаходу є певним станом, що досягається в ході захворювання. Цей результат захворювання може бути, наприклад, клінічний стан, такий як "рецидив захворювання", "розвиток метастаза", "розвиток вузлового метастаза", "розвиток віддаленого метастаза", "виживання", "загибель", "показник ремісії пухлини", стадія або ступінь захворювання або схожі з ними.

Під "ризиком" мають на увазі число, пов'язане з вірогідністю розвитку у суб'єкта або пацієнта або досягнення суб'єктом або пацієнтом певного результату захворювання. Термін "ризик" в контексті даного винаходу не має на увазі наявності якого-небудь позитивного або негативного значення, що має на увазі, відносно благополуччя пацієнта, а відноситься тільки до можливості або вірогідності виникнення або розвитку даного стану.

Термін "клінічні дані" відноситься до всіх доступних даних і інформації, що стосуються стану здоров'я пацієнта, включаючи, але не обмежуючись ними, вік, стать, масу тіла, менопаузальний/гормональний статус, етіопатологічні дані, дані анамнезу, дані, отримані за допомогою діагностичних способів in vitro, таких як гістопатологія, аналізи сечі або крові, дані, отримані способами візуалізації, такими як рентген, комп'ютерна томографія, MRI, PET, spect, ультразвукове дослідження, електрофізіологічні дані, генетичний аналіз, аналіз експресії генів, оцінка за допомогою біопсії, внутрішньоопераційні дані.

Термін "позитивний по вузлах", "діагностований як позитивний по вузлах", "залучення вузлів" або "залучення лімфатичних вузлів" означає пацієнта, у якого раніше були діагностовані метастази в лімфатичних вузлах. Це охоплює метастаз як в дренажний лімфатичний вузол, близькорозташований лімфатичний вузол, так і у віддалений лімфатичний вузол. Цей попередній діагноз не складає частина способу за винаходом. Замість цього, він є передумовою для вибору пацієнтів, зразки яких можна використовувати для одного з варіантів здійснення даного винаходу. Цей попередній діагноз може бути поставлений будь-яким придатним

способом, відомим в даній області, включаючи, але не обмежуючись ними, видалення лімфатичних вузлів і патологічний аналіз, аналіз за допомогою біопсії, аналіз біомаркерів *in vitro*, вказуючих на метастаз, способи візуалізації (наприклад, комп'ютерна томографія, рентген, магнітно-резонансна томографія, ультразвукове дослідження) і внутрішньоопераційні дані.

5 У контексті даного винаходу "біологічним зразком" є зразок, що отриманий з біологічного організму або знаходився у контакті з біологічним організмом. Прикладами біологічних зразків є: клітини, тканина, рідини організму, змивна рідина, мазання, біоптати, кров, сеча, слина, мокрота, плазма, сироватка, супернатант клітинної культури та інші.

10 "Зразком пухлини" є біологічний зразок, що містить пухлинні клітини, або цілі, або зруйновані. Зразок може бути зразком будь-якої біологічної тканини або рідини. Такі зразки включають, але не обмежуються ними, мокроту, кров, сироватку, плазму, клітини крові (наприклад, лейкоцити), тканину, зразки центральної або тонкоїгольної біопсії, що містять клітини рідини організму, сечу, рідину черевної порожнини і плевральну рідину, спинномозкову рідину, слізну рідину або клітини, виділені з них. Також вони можуть включати зрізи тканин, такі як заморожені або фіксовані зрізи, узяті для гістологічних цілей або піддані мікродисекції клітини або їх позаклітинні частини. Зразок пухлини, що підлягає аналізу, може бути тканинним матеріалом з неопластичного осередку пошкодження, узятим аспіруванням або пункцією, вирізуванням або будь-яким іншим хірургічним способом, що забезпечує отриманий біопсією або резекцією клітинний матеріал. Такий зразок містить пухлинні клітини або фрагменти пухлинних клітин, отримані від пацієнта. Клітини можуть бути знайдені в клітинному "мазку", узятим, наприклад, шляхом аспірування через соски, лаважа проток, тонкоїгольної біопсії або шляхом стимульованого або самовільного виділення з сосків. У іншому варіанті здійснення зразком є рідина організму. Такі рідини включають, наприклад, рідини крові, сироватку, плазму, лімфу, асцитичні рідини, гінекологічні рідини або сечу, але не обмежуючись цими рідинами.

25 "Ген" є набір сегментів нуклеїнової кислоти, який містить інформацію, необхідну для отримання функціонального РНК-продукту. "Продукт гена" є біологічною молекулою, що продукується шляхом транскрипції або експресії гена, наприклад мРНК, кДНК або трансльованого білка.

30 "мРНК" є транскрибованим продуктом гена і цей термін має звичайне значення, що використовується фахівцем в даній області. "Молекула, що походить з мРНК" є молекулою, яка хімічно або ферментативний отримана з мРНК-матриці, такої як кДНК.

35 Термін "рівень експресії" відноситься до певного рівня експресії гена. Це може бути певний рівень експресії гена як абсолютна величина або в порівнянні з еталонним геном (наприклад, геном домашнього господарства), з середнім значенням для двох або більш еталонних генів, або з вирахованою середньою величиною експресії (наприклад, в аналізі з ДНК-чіпом) або з іншим інформативним геном без застосування еталонного зразка. Рівень експресії гена можна вимірювати прямо, наприклад, шляхом отримання сигналу, де сила сигналу корелює з кількістю мРНК-транскриптів цього гена, або він може бути отриманий непрямо на білковому рівні, наприклад, за допомогою способів імуногістохімії, CISH, ELISA або RIA. Рівень експресії також можна отримувати за допомогою конкурентної реакції з еталонним зразком. Величині експресії, яку визначають шляхом вимірювання деякого фізичного параметра в аналізі, наприклад, випускання флуоресценції, можна привласнювати числові значення, яке можна використовувати для подальшої обробки інформації.

45 "Еталонний профіль рівнів експресії" в рамках винаходу означає будь-який профіль рівнів експресії, який можна використовувати для порівняння з іншим профілем рівнів експресії. У переважному варіанті здійснення винаходу еталонний профіль рівнів експресії є, наприклад, середнім профілем рівнів експресії, що спостерігається в групі здорових суб'єктів, суб'єктів із захворюванням або суб'єктів із захворюванням, яким проводили конкретний тип терапії, службовців як еталонна група, або суб'єкти з хорошим або поганим результатом.

50 Термін "математичне комбінування рівнів експресії" в рамках даного винаходу означає отримання числової величини з певного рівня експресії гена і застосування алгоритму до однієї або декількох таких числових величин для отримання комбінованої величини або комбінованого показника.

"Алгоритм" є процес, який здійснює деяку послідовність дій для вироблення інформації.

55 "Показник" є числовою величиною, яка отримана математичним комбінуванням рівнів експресії з використанням алгоритму. Також вона може бути отримана з рівнів експресії і іншої інформації, наприклад, клінічних даних. Показник може бути пов'язаний з результатом захворювання пацієнта.

60 "Дискримінантна Функція" є функцією набору перемінних, використовуваних для класифікації об'єкту або події. Таким чином, дискримінантна функція дозволяє класифікацію

пацієнта, зразка або події в категорію або безліч категорій згідно з доступними даними або параметрами від вказаного пацієнта, зразка або події. Така класифікація є стандартним інструментом статистичного аналізу, добре відомим кваліфікованому спеціалісту. Наприклад, пацієнта можна класифікувати як такого, що "має високий ризик" або що "має низький ризик",

5 "має високу вірогідність метастаза" або що "має низьку вірогідність метастаза", що "потребує лікування" або "не потребує лікування" згідно даним, отриманим від вказаного пацієнта, зразка або події. Класифікація не обмежується "високим проти низького", а може здійснюватися на безліч категорій, ступенів або подібних з ними. Класифікацію також слід розуміти в ширшому сенсі як відрізняючий показник, де, наприклад, вищий показник відповідає вищій вірогідності віддаленого метастаза, наприклад, (загальний) ризик віддаленого метастаза. Приклади функцій дискримінантів, які дозволяють класифікацію, включають, але не обмежуються ними, функції, що визначаються методом опорних векторів (SVM), методом к-найближчих сусідів (kNN), (наївні) байєсовські моделі, моделі лінійної регресії або фрагментарно визначені функції, наприклад, такі як при виявленні підгруп, в деревовидній схемі рішень, в логічному аналізі даних (LAD) і тому подібне. У ширшому значенні, прикладами для цілей ілюстрації є величини безперервних показників математичних методів або алгоритмів, таких як коефіцієнти кореляції, проєкції, показники методу опорних векторів, інші способи на основі схожості, їх комбінації і тому подібне.

Термін "методика лікування", "порядок лікування", "режим", а також "режим лікування" відноситься до послідовного або одночасного за часом проведення введення протипухлинних, і/або антивазкулярних, і/або стимулюючих імунну систему засобів, і/або засобів, що забезпечують проліферацію клітин крові, і/або променевої терапії, і/або гіпертермії, і/або гіпертермії для лікування раку. Їх введення можна проводити в ад'ювантному і/або неоад'ювантному порядку. Склад такого "протоколу" може варіювати з погляду дози одного засобу, тимчасових рамок застосування і частоти введення в межах певного терапевтичного діапазону. В даний час проводяться дослідження різних комбінацій різних лікарських засобів і/або фізичних способів, і різних схем.

Термін "цитотоксична хіміотерапія" відноситься до різних методик лікування, що впливають на проліферацію і/або виживання клітин. Лікування може включати введення алкілюючих засобів, антиметаболітів, антрациклінів, алкалоїдів рослин, інгібіторів топоізомерази і інших протипухлинних засобів, включаючи моноклональні антитіла і інгібітори кінази. Зокрема, цитотоксичне лікування може відноситися до лікування таксанами. Таксани є алкалоїдами рослин, які блокують ділення клітин шляхом перешкоди функціонуванню мікротрубочок. Прототипним таксаном є природний продукт паклітаксел, початково відомий як таксол і вперше отриманий з кори тиса тихого.

Доцетаксел є напівсинтетичним аналогом паклітаксела. Таксани підвищують стабільність мікротрубочок, перешкоджаючи розбіжності хромосом в ході анафази.

Термін "ендокринне лікування" або "гормональне лікування" (також іноді зване "антигормональним лікуванням") означає лікування, яке націлене на передачу сигналу гормонами, наприклад, інгібування гормонів, інгібування рецепторів гормонів, використання агоністів або антагоністів рецепторів гормонів, використання рецепторів-сміттярів або сирітських рецепторів, використання похідних гормонів і перешкода продукуванню гормонів. Конкретним прикладом є терапія тамоксифеном, яка модулює передачу сигналу рецептором естрогену, або ароматазне лікування, яке перешкоджає продукуванню стероїдних гормонів.

Тамоксифеном є перорально активний селективний модулятор рецепторів естрогену (SERM), який застосовують для лікування рака молочної залози, і в даний час він є лікарським засобом, що найбільш продається в світі, для цієї мети. Тамоксифен продається під торговими назвами Nolvadex, Istubal і Valodex. Проте цей лікарський засіб, навіть до закінчення дії патенту на нього, був і залишається широко згадуваним за його непатентованим найменуванням "тамоксифен". Тамоксифен і похідні тамоксифену конкурентно зв'язуються з рецепторами естрогену на пухлинах і інших тканинних мішенях, утворюючи ядерний комплекс, який знижує синтез РНК і інгібує ефекти естрогену.

Рецепторами стероїдів є внутріклітинні рецептори (як правило, цитоплазма), які здійснюють передачу сигналу стероїдних гормонів. Їх приклади включають рецептори типу I, зокрема, рецептори статевих гормонів, наприклад, рецептор андрогенів, рецептор естрогену, рецептор прогестерону; рецептор глюкокортикоїдів, рецептор мінералокортикоїдів; і рецептори типу II, наприклад, рецептор вітаміну А, рецептор вітаміну D, рецептор ретиноїдів, рецептор гормону щитовидної залози.

Термін "спосіб на основі гібридизації", як використовують в рамках винаходу, відноситься до способів, що залучають процес об'єднання комплементарних одноланцюжкових нуклеїнових

кислот або нуклеотидних аналогів в одну дволанцюжкову молекулу. Нуклеотиди або нуклеотидні аналоги зв'язуються з комплементарними ним молекулами в нормальних умовах, так що дві абсолютно комплементарних нитки легко зв'язуються одна з одною. У біоаналізах дуже часто використовують мічені одноланцюжкові зонди для пошуку комплементарних послідовностей-мішеней. Якщо такі послідовності існують в зразку, зонди гібридизуються з вказаними послідовностями, які потім можна виявляти завдяки мітці. Інші способи на основі гібридизації включають способи на мікроматрицях і/або біочіпах. У цих способах зонди іммобілізують на твердій фазі, яку потім контактують із зразком. Якщо в зразку існують комплементарні нуклеїнові кислоти, вони гібридизуються із зондами, і, таким чином, їх можна виявити. Ці підходи також відомі як "способи на основі чіпів". Іншим способом на основі гібридизації є ПЦР, яка описана нижче. Коли потрібне визначення рівнів експресії, способи на основі гібридизації, наприклад, можна використовувати для визначення кількості мРНК даного гена.

Олігонуклеотид, здатний специфічно зв'язуватися з послідовностями гена або їх фрагментами, відноситься до олігонуклеотида, який специфічно гібридизується з геном або продуктом гена, таким як мРНК або кДНК гена, або з його фрагментом. Для специфічної детекції гена або продукту гена необов'язково виявляти всю послідовність гена. Фрагмент розміром приблизно 20-150 основ містить достатню специфічну інформацію про послідовність для забезпечення специфічної гібридизації.

Термін "спосіб на основі ПЦР", як використовують в рамках винаходу, відноситься до способів, що включають полімеразну ланцюгову реакцію (ПЦР). Вона є способом експоненціальної ампліфікації нуклеїнових кислот, наприклад ДНК, шляхом ферментативної реплікації *in vitro*. Оскільки ПЦР є технологією *in vitro*, її можна проводити без обмежень форми ДНК, і її можна широко модифікувати для проведення широкого набору генетичних маніпуляцій. Коли потрібно визначити рівні експресії, спосіб на основі ПЦР можна використовувати, наприклад, для детекції присутності даної мРНК шляхом (1) зворотної транскрипції повного набору мРНК (так званого транскриптома) в кДНК за допомогою ферменту зворотної транскриптази, і (2) детекції присутності даної кДНК за допомогою відповідних праймерів. Цей підхід широко відомий як ПЦР із зворотною транскриптазою (ОТ-ПЦР). Більш того, способи на основі ПЦР включають, наприклад, ПЦР в реальному часі, і, зокрема, придатні для аналізу рівнів експресії, кінетичної або кількісної ПЦР (q-ПЦР).

Термін "кількісна ПЦР" (q-ПЦР) відноситься до будь-якого типу способу ПЦР, який дозволяє кількісне визначення матриці в зразку. Кількісна ПЦР в реальному часі включає різні способи виконання або детекції продуктів, як наприклад, спосіб TaqMan або спосіб LightCycler. Наприклад, в способі TaqMan використовується флуорогенний зонд з подвійним міченням. У ПЦР в реальному часі TaqMan вимірюється накопичення продукту через флуорофор на експоненціальних стадіях ПЦР, а не на кінцевій стадії, як в загальноприйнятій ПЦР. Експоненціальне збільшення продукту використовують для визначення порогового циклу, СТ, тобто кількості циклів ПЦР, при якому виявляється значне експоненціальне збільшення флуоресценції, і яке прямо корелює з кількістю копій ДНК-матриці, присутніх в реакції. Умови реакції у високому ступені схожі із загальноприйнятою ПЦР, проте її проводять в термоблоці для аналізу в реальному часі, який дозволяє вимірювання флуоресцентних молекул в пробірках для ПЦР. На відміну від звичайної ПЦР, в ПЦР в реальному часі TaqMan в реакційну суміш додають зонд, тобто одноланцюжковий олігонуклеотид, комплементарний сегменту із 20-60 нуклеотидів в ДНК-матриці і розташований між двома праймерами. З 5'- і 3'-кінцями зонду пов'язані флуоресцентний репортер або флуорофор (наприклад, 6-карбоксифлуоресцеїн, скорочена назва: FAM, або тетрахлорфлуоресцеїн, скорочена назва: TET) і гасильник (наприклад, тетраметилпродамін, скорочена назва: TAMRA, або дигідроциклопірролоіндолтрипептидний "гасильник black hole" скорочена назва: BHQ), відповідно [2]. Близьке розташування флуорофору і гасильника, пов'язаних із зондом, інгібує флуоресценцію флуорофору. В процесі ПЦР, з початком синтезу ДНК, 5'-3' екзонуклеазна активність Taq-полімерази порушує це співвідношення в зонді, який піддався відпалу з матрицею. Деградація зонда вивільняє флуорофор з нього і порушує близьке розташування до гасильника, таким чином, усуваючи ефект гасіння і забезпечуючи флуоресценцію флуорофору. Таким чином, флуоресценція, що виявляється в термоблоці ПЦР в реальному часі, прямо пропорційна вивільненому флуорофору і кількості ДНК-матриці, присутньої в ПЦР.

Під "матрицею" або "впорядкованою системою" мають на увазі розташування областей, що адресуються, або "адресу" на пристрої. Положення можуть бути організовані у вигляді двовірних матриць, тривимірних матриць або інших форматів впорядкованих систем. Кількість положень може знаходитися в діапазоні від декількох до щонайменше сотень тисяч. Найбільш

важливо, кожним розташуванням є повністю незалежна реакційна ділянка. Матриці включають, але не обмежуються ними, матриці нуклеїнових кислот, матриці білків і матриці антитіл. "Матриця нуклеїнових кислот" відноситься до матриці, що містить зонди нуклеїнових кислот, такі як олігонуклеотиди, нуклеотидні аналоги, полінуклеотиди, полімери нуклеотидних аналогів, morpholinos або крупніші частини генів. Нуклеїнова кислота і/або аналог на матриці переважно є односторонніми. Матриці, де зонди є олігонуклеотиди, називають "олігонуклеотидними матрицями" або "олігонуклеотидними чіпами". "Мікроматриця" в справжньому описі також відноситься до "біочіпа" або "біологічного чіпа", матриці з ділянок, що має щільність окремих ділянок щонайменше приблизно 100/см<sup>2</sup>, і переважно щонайменше приблизно 1000/см<sup>2</sup>.

"Пари праймерів" і "зонди" в рамках винаходу мають звичайне значення цього терміну, яке добре відоме фахівцям в області молекулярної біології. У переважному варіанті здійснення винаходу "пари праймерів" і "зонди" означають полінуклеотидні молекули, що мають послідовність, ідентичну, комплементарну, гомологічну областям полінуклеотиду-мішені, або гомологічну комплементарній послідовності областей полінуклеотиду-мішені, який підлягає детекції або кількісному визначенню. Інший варіант здійснення також включає нуклеотидні аналоги для застосування як праймери і/або зонди. Технології зондів, використовуваних для застосування в кінетичній ПЦР або ПЦР в реальному часі, можуть бути, наприклад, системи TaqMan®, що отримуються від Applied Biosystems, зонди для подовження, такі як Scorpion® Primers, подвійні гібридаційні зонди Amplifluor®, що отримуються від Chemicon International, Inc, або малоборознисті ліганди.

"Індивідуально мічені зонди" в рамках винаходу означають молекулярні зонди, що містять полінуклеотид, олігонуклеотид або нуклеотидний аналог і мітку, яка сприяє детекції або кількісному визначенню зонда. Переважними мітками є флуоресцентні молекули, люмінесцентні молекули, радіоактивні молекули, ферментні молекули і/або молекули гасників.

"Розміщені зонди" в рамках винаходу означає набір іммобілізованих зондів, переважно у впорядкованому розташуванні. У переважному варіанті здійснення винаходу окремі "розміщені зонди" можуть бути ідентифіковані по їх відповідному положенню на твердій підкладці, наприклад, на "чіпі".

При використанні відносно послідовності односторонньої нуклеїнової кислоти, термін "по суті гомологічний" відноситься до будь-якого зонду, який може гібризуватися (тобто він є комплементарною послідовністю для) з послідовністю односторонньої нуклеїнової кислоти в умовах низької жорсткості, як описано вище.

Суть винаходу

У загальних рисах, даний винахід відноситься до способу оцінки ризику рецидиву у пацієнта з негативним або позитивним по вузлах, позитивних по рецептору естрогену і негативних по HER2/NEU раком молочної залози, зокрема, у пацієнтів, яким проводять ендокринну терапію, наприклад, при лікуванні тамоксифеном. Статус рецепторів естрогену, як правило, визначають з використанням імуногістохімії, статус HER2/NEU (ERBB2), як правило, визначають з використанням імуногістохімії і флуоресцентної гібридації *in situ*.

Проте для цілей винаходу статус рецепторів естрогену і статус HER2/NEU (ERBB2) можна визначати за допомогою будь-якого придатного способу, наприклад, імуногістохімії, флуоресцентної гібридації *in situ* (FISH) або аналізу експресії РНК.

Даний винахід відноситься до способу передбачення результату рака молочної залози у пацієнта з раком молочної залози, що має позитивну по рецептору естрогену і негативну по HER2 пухлину, причому вказаний спосіб включає:

(а) визначення в зразку пухлини від вказаного пацієнта рівнів експресії РНК щонайменше 2 з наступних 9 генів: UBE2C, BIRC5, RACGAP1, DHCR7, STC2, AZGP1, RBBP8, IL6ST і MGP;

(b) математичне комбінування величин рівня експресії для генів вказаного набору, причому ці величини визначають в зразку пухлини з отриманням комбінованого показника, де вказаний комбінований показник указує на прогноз у вказаного пацієнта. У одному варіанті здійснення вибирають щонайменше 3, 4, 5 або 6 генів.

У наступному варіанті здійснення винаходу спосіб включає:

(а) визначення в зразку пухлини від вказаного пацієнта рівнів експресії РНК наступних 8 генів: UBE2C, RACGAP1, DHCR7, STC2, AZGP1, RBBP8, IL6ST і MGP;

(b) математичне комбінування величин рівня експресії для генів вказаного набору, причому ці величини визначають в зразку пухлини з отриманням комбінованого показника, де вказаний комбінований показник указує на прогноз у вказаного пацієнта.

У наступному варіанті здійснення спосіб за винаходом включає:

(а) визначення в зразку пухлини від вказаного пацієнта рівнів експресії РНК наступних 8 генів: UBE2C, BIRC5, DHCR7, STC2, AZGP1, RBBP8, IL6ST і MGP;



(b) математичне комбінування величин рівня експресії для генів вказаного набору, причому ці величини визначають в зразку пухлини з отриманням комбінованого показника, де вказаний комбінований показник указує на прогноз у вказаного пацієнта.

У іншому варіанті здійснення винаходу BIRC5 може бути замінений UBE2C, або TOP2A, або RACGAP1, або AURKA, або NEK2, або E2F8, або PCNA, або CYBRD1, або DCN, або ADRA2A, або SQLE, або CXCL12, або EPHX2, або ASPH, або PRSS16, або EGFR, або CCND1, або TRIM29, або DHCR7, або PIP, або TFAP2B, або WNT5A, або APOD, або PTPRT, за умови, що після заміни вибирають 8 різних генів; і

UBE2C може бути замінений BIRC5, або RACGAP1, або TOP2A, або AURKA, або NEK2, або E2F8, або PCNA, або CYBRD1, або ADRA2A, або DCN, або SQLE, або CCND1, або ASPH, або CXCL12, або PIP, або PRSS16, або EGFR, або DHCR7, або EPHX2, або TRIM29 за умови, що після заміни вибирають 8 різних генів; і

DHCR7 може бути замінений AURKA, BIRC5, UBE2C або будь-яким іншим геном, який може замінювати BIRC5 або UBE2C, за умови, що після заміни вибирають 8 різних генів; і

STC2 може бути замінений INPP4B, або IL6ST, або SEC14L2, або MAPT, або CHPT1, або ABAT, або SCUBE2, або ESR1, або RBBP8, або PGR, або PTPRT, або HSPA2, або PTGER3, за умови, що після заміни вибирають 8 різних генів; і

AZGP1 може бути замінений PIP, або EPHX2, або PLAT, або SEC14L2, або SCUBE2, або PGR за умови, що після заміни вибирають 8 різних генів; і

RBBP8 може бути замінений CELSR2, або PGR, або STC2, або ABAT, або IL6ST за умови, що після заміни вибирають 8 різних генів; і

IL6ST може бути замінений INPP4B, або STC2, або MAPT, або SCUBE2, або ABAT, або PGR, або SEC14L2, або ESR1, або GJA1, або MGP, або EPHX2, або RBBP8, або PTPRT, або PLAT за умови, що після заміни вибирають 8 різних генів; і

MGP може бути замінений APOD, або IL6ST, або EGFR за умови, що після заміни вибирають 8 різних генів.

Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де вказаний комбінований показник указує на користь цитотоксичної хіміотерапії.

Використання способу за винаходом до проведення у пацієнта ендокринної терапії дозволяє передбачення ефективності ендокринної терапії.

У таблиці 2 нижче показано, чи указує надекспресія кожного з представлених вище маркерних генів на хороший результат або на поганий результат у пацієнта, якому проводять ендокринну терапію. Таким чином, кваліфікований фахівець може сконструювати математичну комбінацію, тобто алгоритм, враховуючи ефект даних генів. Наприклад, підсумовування або зважене підсумовування генів, надекспресія яких указує на хороший результат, забезпечує алгоритм, де високий показник ризику указує на хороший результат. Надійність алгоритму можна досліджувати шляхом аналізу зразків пухлин пацієнтів з історією хвороби, де, наприклад, показник для пацієнтів з хорошим результатом і для пацієнтів з поганим результатом можна визначати окремо і порівнювати. Кваліфікованому фахівцеві, біостатистику, відоме застосування інших математичних методів, таких як дискримінантні функції, для отримання оптимізованих алгоритмів. Алгоритми можна оптимізувати, наприклад, для чутливості або специфічності. Алгоритми можна адаптувати для конкретної аналітичної платформи, використовуюваної для вимірювання експресії маркерних генів, такий як кількісна ПЦР.

Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де вказана ендокринна терапія включає тамоксифен або інгібітор ароматази.

Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де передбачають ризик розвитку рецидиву.

Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де вказаний рівень експресії визначають як рівень небілкової експресії.

Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де вказаний рівень експресії визначають як рівень експресії РНК.

Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де вказаний рівень експресії визначають з допомогою щонайменше одного з:

спосіб на основі ПЦР

спосіб на основі мікроматриці, і

спосіб на основі гібридизації.

Згідно одному аспекту винаходу передбачений спосіб, як описано вищим, де вказане визначення рівнів експресії проводять у фіксованому у формаліні зануреному в парафін зразку пухлини або в свіжозамороженому зразку пухлини.

Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де рівень експресії вказаного щонайменше одного маркерного гена визначають як патерн експресії відносно щонайменше одного еталонного гена, або щодо обчисленої середньої величини експресії.

5 Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де вказана стадія математичного комбінування включає стадію застосування алгоритму до величин, характерних для рівня експресії даного гена.

Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де вказаним алгоритмом є лінійне комбінування вказаних величин, характерних для рівня експресії даного гена.

10 Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де величину, характерну для рівня експресії даного гена, множать на коефіцієнт.

Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де для вказаного комбінованого показника визначають одне, два або більш порогових значень і його виділяють в групи високого і низького ризику, високого, проміжного і низького ризику, або більшу кількість груп ризику шляхом застосування порогового значення до комбінованого показника.

15 Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де високий комбінований показник указує на користь агресивнішої терапії, наприклад, цитотоксичної хіміотерапії. Кваліфікованому фахівцеві зрозуміло, що "високий показник" в цьому відношенні відноситься до еталонної величини або порогової величини. Крім того, кваліфікованому фахівцеві зрозуміло, що, залежно від конкретного алгоритму, використовуюваного для отримання комбінованого показника, також "низький" показник нижче за межу або еталонну величину може указувати на користь агресивнішої терапії, наприклад, цитотоксичної хіміотерапії. Це відбувається у тому випадку, коли гени, що мають позитивну кореляцію з високим ризиком метастазів, враховуються в алгоритмі з позитивним коефіцієнтом, так що загальний високий показник указує на високу експресію генів, що мають позитивну кореляцію з високим ризиком.

25 Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де інформація, що стосується статусу вузлів у пацієнта, обробляється на стадії математичного комбінування величин рівня експресії генів з отриманням комбінованого показника.

Згідно одному аспекту винаходу передбачений спосіб, як описано вище, де вказана інформація, що стосується статусу вузлів, є числовою величиною  $>0$ , якщо вказаний статус вузлів є негативним, і вказана інформація є числовою величиною  $>0$ , якщо статус вузлів є позитивним або невідомим. У ілюстративних варіантах здійснення винаходу негативному статусу вузлів привласнюють величину 0, невідомому статусу вузлів привласнюють величину 0,5 і позитивному статусу вузлів привласнюють величину 1. Можна вибирати інші величини, так щоб вони відображали різну вагу статусу вузлів в алгоритмі.

30 Крім того, винахід відноситься до набору для виконання способу, як описано вище, причому вказаний набір містить набір олігонуклеотидів, здатних специфічно зв'язувати послідовності, або зв'язуватися з послідовностями, фрагментів генів в комбінації генів, де

40 (i) вказана комбінація містить щонайменше 8 генів: UBE2C, BIRC5, DHCR7, STC2, AZGP1, RBBP8, IL6ST і MGP; або

(ii) вказана комбінація містить щонайменше 10 генів: BIRC5, AURKA, PVALB, NMU, STC2, RBBP8, PTGER3, CXCL12, CDH1 і PIP; або

(iii) вказана комбінація містить щонайменше 9 генів BIRC5, DHCR7, RACGAP1, PVALB, STC2, IL6ST, PTGER3, CXCL12 і ABAT; або

45 (iv) вказана комбінація містить щонайменше 9 генів: DHCR7, RACGAP1, NMU, AZGP1, RBBP8, IL6ST і MGP;

Крім того, винахід відноситься до застосування набору для здійснення способу за будь-яким з пп.1-17 формули винаходу, причому вказаний набір містить набір олігонуклеотидів, здатних специфічно зв'язувати послідовності, або зв'язуватися з послідовностями, фрагментів генів в комбінації генів, де

50 (i) вказана комбінація містить щонайменше 8 генів: UBE2C, BIRC5, DHCR7, STC2, AZGP1, RBBP8, IL6ST і MGP; або

(ii) вказана комбінація містить щонайменше 10 генів BIRC5, AURKA, PVALB, NMU, STC2, RBBP8, PTGER3, CXCL12, CDH1 і PIP; або

55 (iii) вказана комбінація містить щонайменше 9 генів: BIRC5, DHCR7, RACGAP1, PVALB, STC2, IL6ST, PTGER3, CXCL12 і ABAT; або

(iv) вказана комбінація містить щонайменше 9 генів: DHCR7, RACGAP1, NMU, AZGP1, RBBP8, IL6ST і MGP; 19. Комп'ютерний програмний продукт, здатний обробляти величини, характерні для рівня експресії генів AKR1C3, MAP4 і SPP1, шляхом математичного

комбінування вказаних величин з отриманням комбінованого показника, де вказаний комбінований показник указує на користь цитотоксичної хіміотерапії у вказаного пацієнта.

Крім того, винахід відноситься до комп'ютерного програмного продукту, здатного обробляти величини, характерні для рівнів експресії комбінації генів, шляхом математичного комбінування вказаних величин з отриманням комбінованого показника, де вказаний комбінований показник указує на ефективність або користь ендокринної терапії вказаного пацієнта, згідно з описаними вище способами.

Вказаний комп'ютерний програмний продукт може зберігатися на носіїві даних або він може бути реалізований в діагностичній системі, здатній видавати величини, характерні для рівня експресії даного гена, такий як система ПЦР в реальному часі.

Якщо комп'ютерний програмний продукт зберігається на носіїві даних або виконується на комп'ютері, працівник, що управляє, може вводити величини експресії, отримані для рівня експресії відповідних генів. Потім комп'ютерний програмний продукт може застосовувати алгоритм для отримання комбінованого показника, вказуючого на користь цитотоксичної хіміотерапії для даного пацієнта.

Способи за даним винаходом мають перевагу забезпечення надійного передбачення результату захворювання на основі використання тільки невеликої кількості генів. Було виявлено, що способи за даним винаходом особливо придатні для аналізу відповіді на ендокринне лікування, наприклад, тамоксифеном, у пацієнтів з пухлинами, класифікованими як позитивні по ESR1 і негативні по ERBB2.

Докладний опис винаходу

Винахід пояснено за допомогою ілюстративних варіантів здійснення і фігур, що додаються:

На фиг.1 представлений графік типу "форрест-плот" для скоректованого відношення ризиків з 95 % довірчим інтервалом для показника T5 в комбінованій групі, а також в окремих ланках лікування досліджень ABCSG06 і 08, з використанням віддаленого метастаза як результат.

На фиг.2 представлений аналіз Каплана-Мейера пацієнтів ER+, HER-, N0-3 з комбінованих груп ABCSG06 і 08, стратифікованих як високий або низький ризик згідно величині показника T5.

У даному описі описані унікальні комбінації маркерних генів, які можна комбінувати в алгоритм для представленого в даному описі нового передбачуваного тесту. Технічно, спосіб за винаходом можна застосовувати на практиці з використанням двох технологій: 1.) Виділення загальної РНК зі свіжої або фіксованої пухлинної тканини і 2.) Кінетична ОТ-ПЦР виділених нуклеїнових кислот. Альтернативно передбачається вимірювання рівнів експресії з використанням альтернативних технологій, наприклад, за допомогою мікроматриці, або шляхом вимірювання на білковому рівні.

Способи за винаходом засновані на кількісному визначенні молекул РНК, виділених з пухлини, для отримання величин експресії і подальшого біоінформативного аналізу вказаних певних величин експресії. Молекули РНК можна виділяти з будь-якого типу зразка пухлини, наприклад, біоптатів, мазків, підданого резекції матеріалу пухлини, свіжомороженої тканини пухлини або із зануреної в парафін і фіксованої у формаліні пухлинної тканини. По-перше, визначають рівні РНК генів, що кодують конкретні комбінації генів UBE2C, BIRC5, DHCR7, RACGAP1, AURKA, PVALB, NMU, STC2, AZGP1, RBBP8, IL6ST, MGP, PTGER3, CXCL12, ABAT, CDH1 і PIP, або їх конкретні комбінації, як вказано. Виходячи з цих величин експресії, обчислюють прогностичний показник шляхом математичного комбінування, наприклад, згідно формулам T5, T1, T4 або T5b (див. нижче). Висока величина показника указує на високий ризик розвитку віддаленого метастаза, низька величина показника указує на низький ризик віддаленого метастаза. Отже, високий показник також указує на те, що пацієнт є пацієнтом з високим ризиком, для якого буде корисна агресивніша терапія, наприклад, цитотоксична хіміотерапія.

Представлені приклади засновані на ідентифікації прогностичних генів з використанням пухлин пацієнтів з гомогенним лікуванням в ад'ювантному режимі тамоксифеном. Більш того, ідентифікація відповідних генів була обмежена пухлинами, класифікованими як позитивні по ESR1 і негативні по ERBB2, виходячи з рівнів експресії РНК. Крім того, для розробки алгоритму розглядали гени, що дозволяють виділяти пухлини проміжного ризику, ступені 2. Нарешті, проводили перенесення платформи з матриць Affymetrix HG U133a на кількісну ПЦР в реальному часі, а також перемикання типу зразка зі свіжомороженої тканини на тканину FFPE для забезпечення надійного виконання алгоритму, незалежно від платформи і типу тканини. В результаті визначення рівня експресії молекули РНК з первинної пухлини і подальший комплексний і багатофакторний аналіз, як описано вище, забезпечують кращий спосіб передбачення вірогідності рецидиву захворювання у пацієнтів, у яких діагностований негативний або позитивний по лімфатичних вузлах ранній рак молочної залози, при лікуванні

тільки тамоксифеном в ад'ювантному режимі. Таким чином, тест заснований на меншій кількості генів, ніж конкуруючі тести, але забезпечує кращу інформацію з погляду високої чутливості і негативної передбачуваної величини, зокрема, для пухлин, що вважаються такими, які проявляють проміжний ризик рецидиву, виходячи із стандартних клінічних чинників.

5 Тотальну РНК екстрагували способом Siemens, заснованим на гранулах з кремнезема і який є повністю автоматизованим способом виділення РНК, з одного цілісного 10-мкм зрізу тканини FFPE на роботі для роботи з рідинами Hamilton MICROLAB STARlet (17). Робот, буфери і хімічні реагенти були частиною системи Siemens VERSANT® kPCR Molecular System (Siemens Healthcare Diagnostics, Tarrytown, NY; не є комерційно доступною в США). У короткому викладі, 10 150 мкл буфера FFPE (буфер FFPE, реагент для дослідження, Siemens Healthcare Diagnostics) додавали до кожного зрізу і інкубували протягом 30 хвилин при 80 °C при струшуванні для плавлення парафіну. Після охолодження додавали протеазу K і інкубували протягом 30 хвилин при 65 °C. Після лізису залишковий дебрис тканині видаляли з рідини для лізису за допомогою стадії інкубації протягом 15 хвилин при 65 °C з 40 мкл покритих кремнеземом гранул оксиду заліза. Гранули зі зв'язаним з поверхнею дебрисом відокремлювали за допомогою магніта і лізати переносили в стандартний 2-мл планшет з глибокими лунками (96 лунок). У них 15 тотальну РНК і ДНК пов'язувала з 40 мкл невикористаних гранул і інкубувала при кімнатній температурі. Хаотропні умови отримували шляхом додавання 600 мкл буфера для лізису. Потім, гранули піддавали магнітному розділенню і супернатанти видаляли. Після цього зв'язані з 20 поверхнею нуклеїнові кислоти промивали три рази з подальшим застосуванням намагнічення, аспірування і видалення супернатантів. Після цього нуклеїнові кислоти елюювали шляхом інкубації гранул з 100 мкл буфера для елювання протягом 10 хвилин при 70 °C при струшуванні. Нарешті, гранули відокремлювали і супернатант інкубували з 12 мкл суміші ДНКази I (2 мкл ДНКази I (без РНКази); 10 мкл 10x буфера ДНКази I; Ambion/Applied Biosystems, Darmstadt, Німеччина) для видалення контамінучої ДНК. Після інкубації протягом 30 хвилин при 25 37 °C, розчин, що не містить ДНК, тотальною РНК розділяли на аліквоти і зберігали при -80 °C, або безпосередньо використовували для аналізу експресії мРНК за допомогою кінетичної ПЦР із зворотною транскрипцією (RTkPCR). Всі зразки аналізували за допомогою одностадійної От-к-ПЦР відносно експресії аж до трьох еталонних генів (RPL37A, CALM2, OAZ1) і аж до 16 генів-мішеней в ABI PRISM® 7900HT (Applied Biosystems, Darmstadt, Німеччина). Систему для одностадійної кількісної ОТ-ПЦР Superscript® III Platinum® з ROX (6-карбокси-X-родамін) (Invitrogen, Karlsruhe, Німеччина) використовували згідно інструкціям виготівника. Відповідні зонди і праймери представлені в таблиці 1. Умови ПЦР були наступними: 30 хвилин при 50 °C, 2 35 хвилини при 95 °C, а потім 40 циклів по 15 секунд при 95 °C і 30 секундах при 60 °C. Всі аналізи ПЦР проводили у трьох екземплярах. Як замісний маркер для виходу РНК використовували порогову (Ct) величину циклів для гена домашнього господарства RPL37A, як описано в літературі (17). Відповідні рівні експресії генів-мішеней обчислювали способом дельта-Ct з використанням формули:

$$20-(Ct(\text{мішень}) - \text{середнє значення } (Ct(\text{еталонні гени}))).$$

40 Перенесення платформи з матриць Affymetrix HG U133a (свіжозаморожена тканина) на кількісну ПЦР в реальному часі обчислювали таким чином. Проводили вимірювання для матеріалу від 158 пацієнтів з використанням обох платформ з отриманням парних вибірок. З даних ПЦР обчислювали величини дельта-Ct. З даних Affymetrix обчислювали Log2-вираження шляхом застосування нижньої межі (що встановлює всі величини нижче за нижню межу як 45 нижню межу), а потім обчислюючи логарифм з основою 2. Застосування нижньої межі знижує ефект збільшеного відносного шумового сигналу при вимірюванні для генів/зразків з низькою експресією; використовували нижню межу 20, також придатна нижня межа від 0,1 до 200. Для кожного заміряного способом ПЦР гена був вибраний набір зондів HG-U133a шляхом максимізації коефіцієнта кореляції Пірсона між величиною дельта-Ct (від ПЦР) і log2-вираженням (від Affymetrix). Також придатні інші заходи кореляції, наприклад, коефіцієнт кореляції Спірмана. В більшості випадків найкращий корелюючий набір зондів належав передбачуваному гену, для решти випадків ПЦР-ген видаляли для подальшої обробки. Гени, що демонструють погану кореляцію між платформами, також видаляли, де використовували поріг коефіцієнта кореляції Пірсона, рівний 0,7 (також придатні величини від 0,5 до 0,8). 55 Перенесення платформи завершували обчисленням тих, що не піддавалися контролю z-переходів для обох платформ і комбінуванням їх; потім єдину величину Пцр-дельта-ст трансформували в шкалу Affymetrix за допомогою наступних стадій: (i) застосувати афінну лінійну трансформацію, де коефіцієнти визначають за допомогою z-трансформації даних ПЦР, (ii) застосувати зворотну афінну лінійну трансформацію, де коефіцієнти визначали за 60 допомогою z-трансформів даних Affymetrix, (iii) інвертувати log2, тобто обчислити експоненціал

по відношенню до основи 2. Альтернативами двократним z-трансформаціям є лінійна регресія або регресія вищого порядку, стійка регресія, або основні компонентні способи, які також є придатними.

Послідовності праймерів і зондів були наступними:

Таблица 1

Последовательности праймеров и зондов для соответствующих генов

ген	зонд	Seq ID	прямой праймер	Seq ID	обратный праймер	Seq ID
ABAT	TCGCCCTAAGAGGCTCTTCCTC	1	GGCAACTTGAGGTCTGACTTTTG	2	GGTCAGCTCACAAGTGGTGTGA	3
ADRA2A	TTGTCCCTTTCCCCCTCCGTGC	4	CCCCAAGAGCTGTTAGGTATCAA	5	TCAATGACATGATCTCAACCAGAA	6
APOD	CATCAGCTCTCAACTCTGGTTTAACA	7	ACTCACTAATGGAAAACGGAAAGATC	8	TCACCTTCGATTGTATTCACAGTT	9
ASPH	TGGGAGGAAGGCAAGGTGCTCATC	10	TGTGCCAACAGACCAAGAC	11	TCGTGCTCAAAGGAGTCATCA	12
AURKA	CCGTCAGCTGTGCTAGGCAT	13	AATCTGGAGGCAAGGTTTCA	14	TCTGGATTGGCTCCTGTGAA	15
BIRC5	AGCCAGATGACGACCCATAGAGGAACA	16	CCCAGTGTTTCTCTGCTCAAG	17	CAACCGGACGAATGCTTTTT	18
CELSR2	ACTGACTTTCCCTTCTGGAGCAGGTGGC	19	TCCAGTGATGTATTCAGAGCTTGT	20	TGCCACAGCCTCTTTTTCT	21
CHPT1	CCAGGCCACCGAAGAGGCAC	22	CGCTCGTCTCATCTCCTACT	23	CCCAGTGCACATAAAGGTATGTC	24
CXCL12	CCACAGCAGGGTTTCAGGTTC	25	GCCACTACCCCTCTCTGAA	26	TCACCTTGCCAACAGTCTCTGAT	27
CYBRD1	AGGGCATCGCCATCATCGTC	28	GTCACCGGCTTCGTCTTCA	29	CAGGTCCACGGCAGTCTGT	30
DCN	TCCTTTTCAGCAACCCGGTCCA	31	AAGGCTCTCTATTCTGGGTGTGA	32	TGGATGGCTGTATCTCCAGTA	33
DHCR7	TGAGCGCCACCCCTCTCGA	34	GGGCTCTGCTTCCCGATT	35	AGTCATAGGGCAAGCAGAAAATTC	36
E2F8	CAGGATACCTAATCCCTCTCACGCAG	37	AAATGTCTCCGCAACCTTGTTT	38	CTGCCCCAGGGATGAG	39
EPHX2	TGAAGCGGGAGGACTTTTGTAAA	40	CGATGAGAGTGTTTTATCCATGCA	41	GCTGAGGCTGGGCTCTTCT	42
ESR1	ATGCCCTTTTGCCGATGCA	43	GCCAAATTTGTTTGTATGGATTAA	44	GACAAAACCGAGTCACATCAGTAATAG	45
GJA1	TGCACAGCCTTTTGATTTCCCGAT	46	CGGGAAGCACCCTCTCTAATC	47	TTTATGTCAGCAGCTAGTTTTTT	48
HSPA2	CAAGTCAGCAAAACACGCAAAA	49	CATGCACGAATAATCAAAAATGC	50	ACATTATTCGAGGTTTCTCTTTAATGC	51
IL6ST	CAAGCTCCACCTTCCAAGGACCT	52	CCCTGAATCCATAAAGGCATACC	53	CAGCTTCGTTTTTCCCTACTTTTT	54
INFP4B	TCCGAGCGCTGGATTGTCATGAG	55	GCACCAGTTACACAAGGACTTCTT	56	TCTCTATGCGGCATCCTTCTC	57
MAPT	AGACTATTTGACACTGCCGCT	58	GTGGCTCAAAGGATAATATCAAAAC	59	ACCTTGCTCAGGTCAACTGGTT	60
MGP	CCTTCATATCCCTCAGCAGAGATGG	61	CGTTTCATTAACAGGAGAAATGCAA	62	ATTGAGCTCGTGGACAGGCTTA	63
NEK2	TCCTGAACAAATGAATCGCATGTCCTACAA	64	ATTGTTGGCACACCTTATTACATGT	65	AAGCAGCCCAATGACCAGATa	66
PCNA	AAATACTAAAATGCGCCGCAATGA	67	GGGCGTGAACCTCACCAGTA	68	CTTCGGCCCTTAGTGTAATGATATC	69
PGR	TTGATAGAAACGCTGTGAGCTCGA	70	AGCTCATCAAGGCAATTTGGTTT	71	ACAAGATCATGCAAGTTATCAAGAAGTT	72
PIP	TGCATGGTGGTTAAACTTTACCTCA	73	TGCTTGCAGTTCAAAACAGAAATTG	74	CACCTTGTAGAGGGATGCTGCTA	75
PLAT	CAGAAAGTGGCCATGCCACCCTG	76	TGGGAAGACATGAATGCACACTA	77	GGAGGTTGGGCTTTAGCTGAA	78
PRSS16	CACTGCGGTCACCCACACCA	79	CTGAGGAGCACAGAACCTCAACT	80	CGAACTCGGTACATGTCTGATACAA	81
PTGER3	TCGGTCTGCTGCTCTCCGCTCC	82	CTGATTGAAGATCATTTTCAACATCA	83	GACGCCATTACAGTTATGG	84
PTPR	TTGGCTTCTGGACACCTTCACA	85	GAGTTGTGGGCTCTACCATTTGC	86	GAGCGGGAACCTTGGGATAG	87
RACGAP1	ACTGAGAACTCCACCCGGCGCA	88	TCGCCAATCTGGATAAATTTGA	89	GAATGTGCGGAATCTGTTTGAG	90
RBBP8	ACCGATTCCGCTACATTCACCCCAAC	91	AGAAATTTGGCTTCTCTGCTCAAG	92	AAAACCAACTTCCAAAAATCTCT	93
SCUBE2	CTAGAGGGTTCCAGGTCCCATACGTGACATA	94	TGTGGATTGAGTTCAAGTCCCAATG	95	CCATCTCGAACTATGTCTTCAATGAGT	96
SEC14L2	TGGGAGGATGCAACGCGTG	97	AGGTCTTACTAAGCAGTCCCATCTCT	98	CGACCGGACCTGAACTC	99
SQLE	TATGCTCTCCCAAAAGAAACACCTCG	100	GCAAGCTTCTCTCTCTCTCA	101	CCTTTAGCAGTTTTTCTCCATAGTTTTATATC	102
TFAP2B	CAACACCACCACTAACAGGCACACGTC	103	GGCATGGACAAGATGTTCTTGA	104	CCTCCTTGTGCGCAGTTTTACT	105
TOP2A	CAGATCAGGACCAAGATGGTTCCACAT	106	CATTGAAGACGCTTCGTTATGG	107	CCAGTTGTGATGGATAAAATTAATCAG	108
TRIM29	TGCTGTCTCACTACCGCCATTCTACG	109	TGGAAATCTGGCAAGCAGACT	110	CAATCCCGTTGCCTTTGTTG	111
UBE2C	TGAACACACATGCTCCGAGCTCTG	112	CTTCTAGGAGAACCCCAACATTGATAGT	113	GTTTCTTGAGGTAATCTTAAAAAGCT	114
WNT5A	TATTCACATCCCTCAGTTGAGTGAATTG	115	CTGTGGCTCTTAATTTATGCATAATG	116	TTAGTGCTTTTTGCTTTCAAGATCTT	117
STC2	TCTCACCTTGACCTCAGCCCAAG	118	ACATTTGACAAATTTCCCTTAGGATT	119	CCAGGACGCAGCTTTACCAA	120
AZGP1	CACCAGCCACCAGGCCCCAG	121	TCCTGGACCGCAAGATC	122	TAGGCCAGGCATTCAGTTTC	123
CALM2	TCGCTCTCGGAAACCGGTAGC	124	GAGCGAGCTGAGTGGTTGTG	125	AGTCAGTTGGTCAGCCATGCT	126
CDH1	CCTGCCAATCCGATGAAATTTGAAAT	127	TGAGTGTCCCCCGGTATCTTC	128	TCAGCCGCTTTCAGATTTTCA	129
NMU	ACCCTGCTGACCTTCTTCCATTCCGT	130	AGAAATTTGGCTTCTCTGCTCAAG	131	AAAACCAACTTCCAAAAATCTCT	132
OAZ1	TGCTTCCACAAGAACCGCGAGGA	133	CGAGCCGACCATGTCTTCAT	134	AAGCCCAAAAAGCTGAAGTT	135
PVALB	AAGTTCTTCCAAATGGTCCGCC	136	CCGACTCCTTCGACCACAA	137	CATCATCCGCACTCTTTTCTTC	138
RPL37A	TGGTGGCGGTGCCTGGA	139	TGTGGTTCTGCATGAAGACA	140	GTGACAGCGGAAGTGGTATTGTAC	141

У таблиці 2 нижче представлені гени, використані в способах за винаходом і в конкретних варіантах здійснення T5, T1, T4 і T5b. Також в таблиці 2 показано, чи указує надекспресія даного гена на хороший або поганий результат при терапії тамоксифеном. У таблиці 2 приведена функція гена, компартментна локалізація в клітині і клітинні процеси, в які він залучений.

Таблиця 2

Перелік генів для алгоритмів T5, T1, T4 і T5b

Ген	Назва	Висока експресія	Функція	Компонент	Процес
UBE2C	убіквітин-кон'югуючий фермент E2C	Поганий результат	зв'язування АТР	цитозоль	ділення клітин
BIRC5	бакуловірусний IAP-повтор, що містить, 5	Поганий результат	зв'язування GTP-ази Ran	цитозоль	клітинний цикл
DHCR7	7-дегідро-холестерин-редуктаза	Поганий результат	активність 7-дегідро-холестерин-редуктази	мембрана ендоплазматичної мережі	регуляція проліферації кліток
RACGAP1	Активуючий GTP-азу Rac білок 1	Поганий результат	активність активатора GTP-ази	цитоплазма	клітинний цикл
AURKA	кіназа aurora A	Поганий результат	зв'язування АТР	центросома	мітотичний клітинний цикл
PVALB	парвальбумін	Поганий результат	зв'язування іонів кальцію		
NMU	нейромедин U	Поганий результат	зв'язування рецепторів	позаклітинна область	передача сигналу
STC2	станніокальцин 2	Хороший результат	активність гормону	позаклітинна область	пов'язана з рецептором клітинної поверхні передача сигналу
AZGP1	альфа-2-глікопротеїн 1	Хороший результат	активність транс-мембранного переносника білків	позаклітинна область	негативна регуляція проліферації клітин
RBBP8	ретинобластомний зв'язуючий білок 8	Хороший результат	зв'язування білків	ядро	контрольна точка клітинного циклу
IL6ST	передавач сигналу інтерлейкіну 6	Хороший результат	рецепторна активність	позаклітинна область	передача сигналу
MGP	матриксний білок Gla	Хороший результат	структурний компонент позаклітинного матриксу	позаклітинна область	дифференціювання клітин
PTGER3	рецептор 3 простагландину E	Хороший результат	активність ліганд-залежного рецептора	плазматична мембрана	передача сигналу
CXCL12	ліганд 12 хемокіна (мотив C-XC)	Хороший результат	активність хемокіна	позаклітинна область	передача сигналу
ABAT	4-амінобутират-амінотрансфераза	Хороший результат	активність трансферази	мітохондрії	катаболічний процес гамма-аміно-масляної кислоти
CDH1	кадгерин 1	Хороший результат	зв'язування молекул клітинної адгезії	плазматична мембрана	гомофільная адгезія кліток

Таблиця 2

Перелік генів для алгоритмів T5, T1, T4 і T5b

Ген	Назва	Висока експресія	Функція	Компонент	Процес
PIP	пролактин-індукуючий білок	Хороший результат	зв'язування актину	позаклітинна область	
CALM2					Еталонний ген
OAZ1					Еталонний ген
RPL37A					Еталонний ген

У таблиці 3 нижче представлені комбінації генів, використані для кожного алгоритму.

Таблиця 3

Комбинация генов для соответствующих алгоритмов

ген	Algo T1	Algo T4	Algo T5	Algo T5b
UBE2C			X	
BIRC5	X	X	X	
DHCR7		X	X	X
RACGAP1		X		X
AURKA	X			
PVALB	X	X		
NMU	X			X
STC2	X	X	X	
AZGP1			X	X
RBBP8	X		X	X
IL6ST		X	X	X
MGP			X	X
PTGER3	X	X		
CXCL12	X	X		
ABAT		X		
CDH1	X			
PIP	X			

5

У таблиці 4 нижче представлений ID набору зондів Affy і картування ID моделі TaqMan для маркерних генів за даним винаходом.

Таблиця 4

Обозначение гена ID набора зондов Affy и картирование ID модели TaqMan

ген	ID модели	ID набора зондов
UBE2C	R65	202954 at
BIRC5	SC089	202095 s at
DHCR7	CAGMC334	201791 s at
RACGAP1	R125-2	222077 s at
AURKA	CAGMC336	204092 s at
PVALB	CAGMC339	205336 at
NMU	CAGMC331	206023 at
STC2	R52	203438 at
AZGP1	CAGMC372	209309 at
RBBP8	CAGMC347	203344 s at
IL6ST	CAGMC312	212196 at
MGP	CAGMC383	202291 s at
PTGER3	CAGMC315	213933 at
CXCL12	CAGMC342	209687 at
ABAT	CAGMC338	209460 at
CDH1	CAGMC335	201131 s at

У таблиці 5 нижче представлені повні назви, ID гена Entrez, номер доступу в gene bank і хромосомна область маркерних генів за даним винаходом.

Таблиця 5

Офіційне позначення	Офіційна повна назва	ID гена Entrez	Номер доступу	Розташування
UBE2C	убіквітин-кон'югуючий фермент E2C	11065	U73379	20q13.12
BIRC5	бакуловірусний IAP-повтор, що містить, 5	332	U75285	17q25
DHCR7	7-дегідро-холестерин-редуктаза	1717	AF034544	11q13.4
STC2	станніокальцин 2	8614	AB012664	5q35.2
RBBP8	ретинобластомний зв'язуючий білок 8	5932	AF043431	18q11.2
IL6ST	передавач сигналу інтерлейкіну 6	3572	M57230	5q11
MGP	матриксний білок Gla	4256	M58549	12p12.3
AZGP1	альфа-2-глікопротеїн 1, цинк-зв'язаний	563	BC005306	11q22.1
RACGAP1	активуючий GTP-азу Rac білок 1	29127	NM_013277	12q13
AURKA	кіназа аутогана A	6790	BC001280	20q13
PVALB	парвальбумін	5816	NM_002854	22q13.1
NMU	нейромедін U	10874	X76029	4q12
PTGER3	рецептор 3 простагландину E (підтип Ep3)	5733	X83863	1p31.2
CXCL12	ліганд 12 хемокіна (мотив C-XC) (стромальний клітинний чинник 1)	6387	L36033	10q11.1
ABAT	4-амінобутират-амінотрансфераза	18	L32961	16p13.2
CDH1	кадгерин 1, тип 1, E-кадгерин (епітеліальний)	999	L08599	16q22.1
PIP	пролактин-індукуючий білок	5304	NMM_002652	7q32-qter



Приклад алгоритму T5:

Алгоритм T5 є групою з чотирьох членів, де кожен член є лінійною комбінацією з двох генів. Математичні формули для T5 представлені нижче; позначення є такими ж, як і для T1. T5 можна обчислювати тільки з даних про експресію генів.

член 1 ризику= $0,434039 [0,301...0,567]*(0,939*BIRC5-3,831)$

$-0,491845 [-0,714...-0,270]*(0,707*RBBP8-0,934)$

член 2 ризику= $0,488785 [0,302...0,675]*(0,794*UBE2C-1,416)$

$-0,374702 [-0,570...-0,179]*(0,814*IL6ST-5,034)$

член 3 ризику= $-0,39169 [-0,541...-0,242]*(0,674*AZGP1-0,777)$

$+ 0,44229 [0,256...0,628]*(0,891*DHCR7-4,378)$

член 4 ризику= $-0,377752 [-0,543...-0,212]*(0,485*MGP+4,330)$

$-0,177669 [-0,267...-0,088]*(0,826*STC2-3,630)$

ризик=член 1 ризику+член 2 ризику+член 3 ризику+член 4 ризику

Коефіцієнти зліва на кожному рядку обчислювали як коефіцієнти регресії пропорційних ризиків COX, числа в квадратних дужках визначають 95 % довірчих інтервалів для цих коефіцієнтів. Іншими словами, замість множення значення виразу  $(0,939*BIRC5-3,831)$  на  $0,434039$ , воно може бути помножене на будь-який коефіцієнт між  $0,301$  і  $0,567$  і, проте, даватиме передбачуваний результат з 95 % довірчим інтервалом. Значення виразів в круглих дужках справа в кожному рядку позначають перенесення платформи з ПЦР на Affymetrix: змінні PVALB, CDH1 ... означають засновану на ПЦР експресію, нормалізовану до еталонних генів (величини дельта-Ct), повне значення виразу в круглих дужках відповідає логарифму (основа 2) величин експресії на мікроматриці Affymetrix для відповідних наборів зондів.

Ефективність алгоритму T5 тестували у пацієнтів, яких лікували тамоксифеном або анастрозолом з не більше ніж 3 позитивними лімфатичними вузлами і ER+, HER2-пухлинами, які брали участь в рандомізованих клінічних випробуваннях ABCSG06 (n=332) або ABCSG08 (n=1244). Як представлено на фіг.1, регресійний аналіз Cox показує, що показник T5 має значущу асоціацію з розвитком віддаленого метастаза у всіх протестованих групах.

Аналіз Каплана-Мейера проводили після класифікації пацієнтів об'єднаних груп ABCSG з використанням заданої межі для показника T5.

Пацієнти з низьким ризиком розвитку віддаленого метастаза мали показник T5? -9,3, тоді як пацієнти з високим ризиком розвитку віддаленого метастаза мали показник T5 вище-9,3. Як представлено на фіг.2, спостерігають високо значуще розділення для обох груп ризику.

Важливо, що показник T5 оцінювали і порівнювали проти "Adjuvant! Online", інтерактивного інструменту, що допомагає вибрати терапію на основі введених клінічних параметрів, таких як розмір пухлини, ступінь пухлини і статус вузлів. Коли показник T5 тестували за допомогою регресії Cox з двома змінними проти показника ризику рецидива в Adjuvant! Online, обидва показники зберігали значущу асоціацію з розвитком віддалених метастазів. Регресія Cox з двома змінними з використанням дихотомізованих даних, які стратифікували згідно T5 (межа -9,3), відповідно, відносно Adjuvant! Online (межа 8), знов давала високо значущі і незалежні асоціації з часом до метастаза як клінічний результат.

Таблиця 6

Регресія Cox з двома змінними для T5 і Adjuvant! Online

Змінна	Відношення ризиків	95 % CI*	P
Adjuvant! Online	2,36	1,58-3,54	<0,0001
Сигнатура експресії генів (група ризику)	2,62	1,71-4,01	<0,0001
Adjuvant! Online (показник)	1,04	1,02-1,06	<0,0001
Сигнатура експресії генів (група ризику)	1,35	1,21-1,49	<0,0001

де HR=відношення ризиків, 95 %CI=95 % довірчий інтервал, p=величина P.

Ілюстративні криві Каплана-Мейера представлені на Фіг. 1, де Високий = група високого ризику, Низький = група низького ризику, згідно заданій межі.

Висока величина показника T5 указує на збільшений ризик виникнення віддаленого метастаза в даний період часу.

Було показано, що це має місце для пацієнтів, яких лікували тамоксифеном, а також для пацієнтів, яких лікували інгібіторами ароматази.

## Приклад алгоритму T1:

Алгоритм T1 є групою з трьох членів, де кожен член є лінійною комбінацією з аж до чотирьох змінних. Як правило, змінні можуть бути експресією генів або клінічними змінними. У T1 єдиною не пов'язаною з генами змінною є статус вузлів, що кодується 0, якщо пацієнт є негативним по вузлах, і 1, якщо пацієнт є позитивним по лімфатичних вузлах. Математичні формули для T1

представлені нижче.

$$\begin{aligned} \text{член 1 ризику} &= +0,193935 [0,108...0,280]*(0,792*PVALB-2,189) \\ &-0,240252 [-0,400...-0,080]*(0,859*CDH1-2,900) \\ &-0,270069 [-0,385...-0,155]*(0,821*STC2-3,529) \\ &+ 1,2053 [0,534...1,877]*\text{статус вузлів} \\ \text{член 2 ризику} &= -0,25051 [-0,437...-0,064]*(0,558*CXCL12+0,324) \\ &-0,421992 [-0,687...-0,157]*(0,715*RBBP8-1,063) \\ &+0,148497 [0,029...0,268]*(1,823*NMU-12,563) \\ &+0,293563 [0,108...0,479]*(0,989*BIRC5-4,536) \\ \text{член 3 ризику} &= +0,308391 [0,074...0,543]*(0,812*AURKA-2,656) \\ &-0,225358 [-0,395...-0,055]*(0,637*PTGER3+0,492) \\ &-0,116312 [-0,202...-0,031]*(0,724*PIP+0,985) \end{aligned}$$

Ризик = + член 1 ризику + член 2 ризику + член 3 ризику

Коефіцієнти зліва на кожному рядку обчислювали як коефіцієнти регресії пропорційних ризиків COX, числа в квадратних дужках позначають 95 % довірчих інтервалів для цих коефіцієнтів. Значення виразів в круглих дужках справа в кожному рядку позначають перенесення платформи з ПЦР на Affymetrix: змінні PVALB, CDH1 ... позначають засновану на ПЦР експресію, нормалізовану до еталонних генів (величини дельта-Ct), повне значення виразу в круглих дужках відповідає логарифму (основа 2) величин експресії на мікроматриці Affymetrix для відповідних наборів зондів.

## Приклад алгоритму T4:

Алгоритм T4 є лінійною комбінацією мотивів. Верхні 10 генів декількох аналізів наборів даних Affymetrix і даних ПЦР об'єднували в мотиви. Гени, що не належать до кластера, використовували як мотиви одного гена. Коефіцієнти регресії пропорційних ризиків COX були знайдені в багатофакторному аналізі.

Як правило, мотиви можуть бути експресією одного гена або середньою експресією корелюючих генів. Математичні формули для T4 представлені нижче.

$$\begin{aligned} \text{prolif} &= ((0,84 [0,697...0,977]*RACGAP1-2,174)+(0,85 [0,713...0,988] \\ &*DHCR7-3,808)+(0,94 [0,786...1,089]*BIRC5-3,734))/3 \\ \text{мотив2} &= ((0,83 [0,693...0,96]*IL6ST-5,295)+(1,11 [0,930...1,288]* \\ &ABAT-7,019)+(0,84 [0,701...0,972]*STC2-3,857))/3 \\ \text{ptger3} &= (PTGER3*0,57 [0,475...0,659]+1,436) \\ \text{cxcl12} &= (CXCL12*0,53 [0,446...0,618]+0,847) \\ \text{pvalb} &= (PVALB*0,67 [0,558...0,774]-0,466) \end{aligned}$$

Коефіцієнти і відхилення для кожного гена означають перенесення платформи з ПЦР на Affymetrix: Змінні RACGAP1, DHCR7 ... позначають засновану на ПЦР експресію, нормалізовану до CALM2 і PPIA, повне значення виразу в круглих дужках відповідає логарифму (основа 2) величин експресії на мікроматриці Affymetrix для відповідних наборів зондів.

Числа в квадратних дужках означають 95 % довірчих інтервалів для цих коефіцієнтів. Оскільки алгоритм мав ще кращі характеристики в комбінації з клінічною змінною, додавали статус вузлів. У T4 статус вузлів кодується як 0, якщо пацієнт є негативним по лімфатичних вузлах, і 1, якщо пацієнт є позитивним по лімфатичних вузлах. При цьому алгоритм T4 є:

$$\begin{aligned} \text{ризик} &= -0,32 [-0,510...-0,137]*\text{мотив 2} \\ &+0,65 [0,411...0,886]*\text{prolif} \\ &-0,24 [-0,398...-0,08]*\text{ptger3} \\ &-0,05 [-0,225...0,131]*\text{cxcl12} \\ &+0,09 [0,019...0,154]*\text{pvalb} \\ &+\text{статус вузлів} \end{aligned}$$

Коефіцієнти ризику обчислювали як коефіцієнти регресії пропорційних ризиків COX, числа в квадратних дужках означають 95 % довірчий інтервал для цих коефіцієнтів.

Алгоритм T5b є групою з двох членів, де кожен член є лінійною комбінацією з чотирьох генів. Математичні формули для T5b представлені нижче, позначення є такими ж, як і для T1 і T5. У T5b не пов'язана з генами змінна є статусом вузлів, що кодується 0, якщо пацієнт є негативним по лімфатичних вузлах, і 1, якщо пацієнт є позитивним по лімфатичних вузлах, і 0,5, якщо статус лімфатичних вузлів невідомий. T5b визначають як:

член 1 ризик= $0,359536 [0,153...0,566]*(0,891*DHCR7-4,378)$   
 $-0,288119 [-0,463...-0,113]*(0,485*MGP+4,330)$   
 $+0,257341 [0,112...0,403]*(1,118*NMU-5,128)$   
 $-0,337663 [-0,499...-0,176]*(0,674*AZGP1-0,777)$

5 член 2 ризик= $-0,374940 [-0,611...-0,139]*(0,707*RBBP8-0,934)$   
 $-0,387371 [-0,597...-0,178]*(0,814*IL6ST-5,034)$   
 $+0,800745 [0,551...1,051]*(0,860*RACGAP1-2,518)$   
 $+0,770650 [0,323...1,219]*\text{статус вузлів}$

ризик = член 1 ризику + член 2 ризику

10 Кваліфікованому фахівцеві зрозуміло, що ці алгоритми відображають конкретні приклади, і що на основі інформації, що стосується асоціації експресії генів з результатом, як приведено в таблиці 2, можна встановлювати альтернативні алгоритми з використанням звичайної кваліфікації.

Спрощення алгоритму з використанням підгруп генів

15 "Прикладом алгоритму T5" є груповий провісник, що складається з 4 членів по 2 з тих, що представляють інтерес гена в кожному. Кожен член є незалежним і самостійним провісником віддаленого рецидиву, кожен додатковий член вносить внесок до передбачуваної надійності і потужності алгоритму для прогнозу часу до метастаза, часу до загибелі або вірогідності виживання для пацієнта з раком молочної залози. Рівняння нижче демонструє "приклад алгоритму T5"; для простоти прочитання кількість розрядів після десяткового знаку скорочена до 2; діапазон в квадратних дужках указує на оцінений діапазон коефіцієнтів (середнє значення +/- 3 стандартних відхилення).

Алгоритм T5:

25  $+0,41 [0,21...0,61]*BIRC5-0,33 [-0,57...-0,09]*RBBP8$   
 $+0,38 [0,15...0,61]*UBE2C-0,30 [-0,55...-0,06]*IL6ST$   
 $-0,28 [-0,43...-0,12]*AZGP1+0,42 [0,16...0,68]*DHCR7$   
 $-0,18 [-0,31...-0,06]*MGP-0,13 [-0,25...-0,02]*STC2$

індекси з: досліджувана група=0,724

30 Назви генів в алгоритмі позначають різницю в експресії мРНК гена в порівнянні з одним або декількома генами домашнього господарства, як описано вище.

При аналізі групи, що відрізняється від досліджуваної групи (234 зразки пухлини) несподівано було відкрито, що деякі спрощення "початкового алгоритму T5" все ще забезпечували діагностичну ефективність, що не є значно нижчою щодо початкового алгоритму T5. Найбільш простим спрощенням було зменшення групового провісника тільки до одного члена. Приклади ефективності "груп з одного члена" представлені нижче:

тільки 1 член:

$+0,41 [0,21...0,61]*BIRC5-0,33 [-0,57...-0,09]*RBBP8$

індекси з: досліджувана група=0,653, незалежна група=0,681

тільки 2 член:

40  $+0,38 [0,15...0,61]*UBE2C-0,30 [-0,55...-0,06]*IL6ST$

індекси з: досліджувана група =0,664, незалежна група =0,696

тільки 3 член:

$-0,28 [-0,43...-0,12]*AZGP1+0,42 [0,16...0,68]*DHCR7$

індекси з: досліджувана група=0,666, незалежна група=0,601

45 тільки 4 член:

$-0,18 [-0,31...-0,06]*MGP-0,13 [-0,25...-0,02]*STC2$

індекси з: досліджувана група=0,668, незалежна група=0,593

Ефективність груп з одного гена, як представлено на незалежній вибірці з 234 зразків, значно понижена в порівнянні з ефективністю повного алгоритму. Проте, використання групи, що складається з меншого числа членів, забезпечує простішу, менш дорогу оцінку ризику рецидиву раки молочної залози, або смерті від раки молочної залози, яка може бути прийнятною для певної діагностичної мети.

Послідовне комбінування більш за одне, але менше чотирьох членів в новий груповий передбачуваний алгоритм, часто призводить до невеликого, але значущого збільшення діагностичної ефективності в порівнянні з групою з одного члена. Було несподівано, що існували виражені поліпшення для деяких комбінацій групових членів, тоді як інші комбінації не забезпечували майже ніякого поліпшення. Спочатку, гіпотеза полягала в тому, що комбінація членів, відповідних схожим біологічним мотивам, що відбивалося використовуваними генами, призводила до меншого поліпшення, ніж комбінація членів, що відображають явно біологічні мотиви, що відрізняються. Проте, справа була не в цьому. Не можна було ідентифікувати

правила для прогнозу комбінації деяких генів для отримання алгоритму, що володіє більшою прогностичною потужністю, ніж інша комбінація генів. Перспективні комбінації можна було вибрати тільки на основі експериментальних даних.

Ідентифіковані комбінації комбінованих групових членів для отримання спрощених, проте, потужних алгоритмів, представлені нижче.

5 тільки 1 і 2 члени:

+0,41 [0,21...0,61]\*BIRC5-0,33 [-0,57...-0,09]\*RBBP8

+0,38 [0,15...0,61]\*UBE2C-0,30 [-0,55...-0,06]\*IL6ST

індекси з: досліджувана група =0,675, незалежна група=0,712

10 тільки 1 і 3 члени:

+0,41 [0,21...0,61]\*BIRC5-0,33 [-0,57...-0,09]\*RBBP8

-0,28 [-0,43...-0,12]\*AZGP1+0,42 [0,16...0,68]\*DHCR7

індекси з: досліджувана група=0,697, незалежна група=0,688

тільки 1 і 4 члени:

15 +0,41 [0,21...0,61]\*BIRC5-0,33 [-0,57...-0,09]\*RBBP8

-0,18 [-0,31...-0,06]\*MGP-0,13 [-0,25...-0,02]\*STC2

індекси з: досліджувана група =0,705, незалежна група=0,679

тільки 2 і 3 члени:

20 +0,38 [0,15...0,61]\*UBE2C-0,30 [-0,55...-0,06]\*IL6ST

-0,28 [-0,43...-0,12]\*AZGP1+0,42 [0,16...0,68]\*DHCR7

індекси з: досліджувана група=0,698, незалежна група=0,670

тільки 1, 2 і 3 члени:

25 +0,41 [0,21...0,61]\*BIRC5-0,33 [-0,57...-0,09]\*RBBP8

+0,38 [0,15...0,61]\*UBE2C-0,30 [-0,55...-0,06]\*IL6ST

-0,28 [-0,43...-0,12]\*AZGP1+0,42 [0,16...0,68]\*DHCR7

індекси з: досліджувана група=0,701, незалежна група=0,715

Виключення не повних членів групи, а окремого гена або генів з різних членів групи також можливо, але це вимагає перепідготовки всього алгоритму. Проте, його також можна було переважно проводити. Ефективність спрощених алгоритмів, отриманих шляхом виключення цілих членів або окремих генів здебільшого є ідентичною.

Варіанти алгоритмів із заміною генів

Описані алгоритми, такі як "приклад алгоритму T5", вище, також можна модифікувати шляхом заміни одного або декількох генів одним або декількома іншими генами. Метою таких модифікацій є заміна генів, які важко виміряти на конкретній платформі, геном, який простіше аналізувати на цій платформі. Тоді як таке перенесення необов'язково може приводити до покращеної ефективності в порівнянні з початковим алгоритмом, він може забезпечувати ключ для впровадження прогностичного алгоритму в конкретну діагностичну платформу. Як правило, заміну одного гена іншим геном при збереженні діагностичної потужності прогностичного алгоритму можна якнайкраще здійснювати шляхом заміни одного гена коекспресуючим геном з високою кореляцією (показаною, наприклад, за допомогою коефіцієнта кореляції Пірсона). Проте, необхідно враховувати, що експресія мРНК двох генів високо корелює на одній платформі, може виявитися абсолютно незалежною між ними при оцінці на іншій платформі. Таким чином, така здавалося б легка заміна при зведенні до експериментальної практики може приводити до невтішно поганих результатів, а також до несподівано значних результатів, що завжди залежить від непередбачуваної обліку використовуваної платформи. Шляхом повторення цього процесу можна замінити декілька генів.

Ефективність такого підходу може бути продемонстрована шляхом оцінки передбачуваної ефективності показника алгоритму T5 і його варіанту на групах для підтвердження. У представленій нижче таблиці представлений індекс с відносно кінцевого результату у вигляді віддаленого рецидиву в двох групах для підтвердження.

Таблиця 7

Варіант	Дослідження для підтвердження А	Дослідження для підтвердження В
початковий алгоритм T5	індекс с=0,718	індекс с=0,686
виключення BIRC5 (установка експресії на деяку константу)	індекс с=0,672	індекс с=0,643
заміна BIRC5 на UBE2C (без корекції коефіцієнта)	індекс с=0,707	індекс с=0,678

Можна бачити, що виключення одного з генів T5, тут представлене, наприклад, для BIRC5 значно знижує ефективність передбачування. Заміна його іншим геном призводить приблизно до такої ж ефективності.

Кращим способом заміни гена є перепідготовка алгоритму. Оскільки T5 складається з чотирьох незалежних членів групи, необхідно перепідготувати тільки член, який містить замінений ген. Представлені нижче рівняння демонструють заміни генів алгоритму T5, представлені вище, підготовлені в групі з 234 пацієнтів з раком молочної залози. Нижче представлений тільки один член, для обчислення індексу с решта членів використовувала незмінними в порівнянні з початковим алгоритмом T5. У діапазоні в квадратних дужках приведений оцінений діапазон коефіцієнтів: середнє значення +/- 3 стандартних відхилень.

Член 1 T5:

Початковий член 1:

$+0,41 [0,21 \dots 0,61] * BIRC5 - 0,33 [-0,57 \dots -0,09] * RBBP8$

індекси з: досліджувана група = 0,724, незалежна група = 0,705

заміна BIRC5 на TOP2A в членові 1:

$+0,47 [0,24 \dots 0,69] * TOP2A - 0,34 [-0,58 \dots -0,10] * RBBP8$

з індекси з: досліджувана група = 0,734, незалежна група = 0,694

заміна BIRC5 на RACGAP1 в членові 1:

$+0,69 [0,37 \dots 1,00] * RACGAP1 - 0,33 [-0,57 \dots -0,09] * RBBP8$

індекси з: досліджувана група = 0,736, незалежна група = 0,743

заміна RBBP8 на CELSR2 в членові 1:

$+0,38 [0,19 \dots 0,57] * BIRC5 - 0,18 [-0,41 \dots 0,05] * CELSR2$

індекси з: досліджувана група = 0,726, незалежна група = 0,680

заміна RBBP8 на PGR в членові 1:

$+0,35 [0,15 \dots 0,54] * BIRC5 - 0,09 [-0,23 \dots 0,05] * PGR$

індекси з: досліджувана група = 0,727, незалежна група = 0,731

Член 2 T5:

Початковий член 2:

$+0,38 [0,15 \dots 0,61] * UBE2C - 0,30 [-0,55 \dots -0,06] * IL6ST$

індекси з: досліджувана група = 0,724, незалежна група = 0,725

заміна UBE2C на RACGAP1 в членові 2:

$+0,65 [0,33 \dots 0,96] * RACGAP1 - 0,38 [-0,62 \dots -0,13] * IL6ST$

індекси з: досліджувана група = 0,735, незалежна група = 0,718

заміна UBE2C на TOP2A в членові 2:

$+0,42 [0,20 \dots 0,65] * TOP2A - 0,38 [-0,62 \dots -0,13] * IL6ST$

індекси з: досліджувана група = 0,734, незалежна група = 0,700

заміна IL6ST на INPP4B в членові 2:

$+0,40 [0,17 \dots 0,62] * UBE2C - 0,25 [-0,55 \dots 0,05] * INPP4B$

індекси з: досліджувана група = 0,725, незалежна група = 0,686

заміна IL6ST на MAPT в членові 2:

$+0,45 [0,22 \dots 0,69] * UBE2C - 0,14 [-0,28 \dots 0,01] * MAPT$

індекси з: досліджувана група = 0,727, незалежна група = 0,711

Член 3 T5:

Початковий член 3:

$-0,28 [-0,43 \dots -0,12] * AZGP1 + 0,42 [0,16 \dots 0,68] * DHCR7$

індекси з: досліджувана група = 0,724, незалежна група = 0,705

заміна AZGP1 на PIP в членові 3:

$-0,10 [-0,18 \dots -0,02] * PIP + 0,43 [0,16 \dots 0,70] * DHCR7$

індекси з: досліджувана група = 0,725, незалежна група = 0,692

заміна AZGP1 на EPHX2 в членові 3:

$-0,23 [-0,43 \dots -0,02] * EPHX2 + 0,37 [0,10 \dots 0,64] * DHCR7$

індекси з: досліджувана група = 0,719, незалежна група = 0,698

заміна AZGP1 на PLAT в членові 3:

$-0,23 [-0,40 \dots -0,06] * PLAT + 0,43 [0,18 \dots 0,68] * DHCR7$

індекси з: досліджувана група = 0,712, незалежна група = 0,715

заміна DHCR7 на AURKA в членові 3:

$-0,23 [-0,39 \dots -0,06] * AZGP1 + 0,34 [0,10 \dots 0,58] * AURKA$

індекси з: досліджувана група = 0,716, незалежна група = 0,733

Член 4 T5:

Початковий член 4:

- 0,18 [-0,31...-0,06]\*MGP-0,13 [-0,25...-0,02]\*STC2  
індекси з: досліджувана група =0,724, незалежна група=0,705  
заміна MGP на APOD в членові 4:  
-0,16 [-0,30...-0,03]\*APOD-0,14 [-0,26...-0,03]\*STC2
- 5 індекси з: досліджувана група =0,717, незалежна група=0,679  
заміна MGP на EGFR в членові 4:  
-0,21 [-0,37...-0,05]\*EGFR-0,14 [-0,26...-0,03]\*STC2  
індекси з: досліджувана група =0,715, незалежна група=0,708  
заміна STC2 на INPP4B в членові 4:
- 10 -0,18 [-0,30...-0,05]\*MGP-0,22 [-0,53...0,08]\*INPP4B  
індекси з: досліджувана група =0,719, незалежна група=0,693  
заміна STC2 на SEC14L2 в членові 4:  
-0,18 [-0,31...-0,06]\*MGP-0,27 [-0,49...-0,06]\*SEC14L2  
індекси з: досліджувана група =0,718, незалежна група =0,681
- 15 Можна бачити, що заміна окремих генів, експериментально ідентифікованих для кількісного визначення за допомогою кінетичної ПЦР, зазвичай впливає на передбачувану ефективність алгоритму T5, оціненого за допомогою індексу  $s$ , тільки незначно.
- У наступній таблиці (таблиця 8) представлена потенційна заміна генів-кандидатів для генів в алгоритмі T5. Кожен ген-кандидат представлений в одному елементі таблиці: за назвою гена слідує поміщений в дужки абсолютний коефіцієнт кореляції Пірсона для експресії початкового гена в алгоритмі T5 і кандидата для заміни, і ID набору зондів HG-U133A.
- 20

Таблиця 8

BIRC5	RBBP8	UBE2C	IL6ST	AZGP1	DHCR7	MGP	STC2
UBE2C (0.775), 202954_at	CELSR2 (0.548), 204029_at	BIRC5 (0.775), 202095_s_at	INPP4B (0.477), 205376_at	PIP (0.530), 206509_at	AURKA (0.345), 204092_s_at	APOD (0.368), 201525_at	INPP4B (0.500), 205376_at
TOP2A (0.757), 201292_at	PGR (0.392), 208305_at	RACGAP1 (0.756), 222077_s_at	STC2 (0.450), 203438_at	EPHX2 (0.369), 209368_at	BIRC5 (0.323), 202095_s_at	IL6ST (0.327), 212196_at	IL6ST (0.450), 212196_at
RACGAP1 (0.704), 222077_s_at	STC2 (0.361), 203438_at	TOP2A (0.753), 201292_at	MAPT (0.440), 206401_s_at	PLAT (0.366), 201860_s_at	UBE2C (0.315), 202954_at	EGFR (0.308), 201983_s_at	SEC14L2 (0.417), 204541_at
AURKA (0.681), 204092_s_at	ABAT (0.317), 209459_s_at	AURKA (0.694), 204092_s_at	SCUBE2 (0.418), 219197_s_at	SEC14L2 (0.351), 204541_at			MAPT (0.414), 206401_s_at
NEK2 (0.680), 204026_s_at	IL6ST (0.311), 212196_at	NEK2 (0.684), 204026_s_at	ABAT (0.389), 209459_s_at	SCUBE2 (0.331), 219197_s_at			CHPT1 (0.410), 221675_s_at
E2F8 (0.640), 219990_at		E2F8 (0.652), 219990_at	PGR (0.377), 208305_at	PGR (0.302), 208305_at			ABAT (0.409), 209459_s_at
PCNA (0.544), 201202_at		PCNA (0.589), 201202_at	SEC14L2 (0.356), 204541_at				SCUBE2 (0.406), 219197_s_at
CYBRD1 (0.462), 217889_s_at		CYBRD1 (0.486), 217889_s_at	ESR1 (0.353), 205225_at				ESR1 (0.394), 205225_at
DCN (0.439), 209335_at		ADRA2A (0.391), 209869_at	GJA1 (0.335), 201667_at				RBBP8 (0.361), 203344_s_at
ADRA2A (0.416), 209869_at		DCN (0.384), 209335_at	MGP (0.327), 202291_s_at				PGR (0.347), 208305_at
SQLE (0.415), 209218_at		SQLE (0.369), 209218_at	EPHX2 (0.313), 209368_at				PTPRT (0.343), 205948_at
CXCL12 (0.388), 209687_at		CCND1 (0.347), 208712_at	RBBP8 (0.311), 203344_s_at				HSPA2 (0.317), 211538_s_at
EPHX2 (0.362), 209368_at		ASPH (0.344), 210896_s_at	PTPRT (0.303), 205948_at				PTGER3 (0.314), 210832_x_at
ASPH (0.352), 210896_s_at		CXCL12 (0.342), 209687_at	PLAT (0.301), 201860_s_at				
PRSS16 (0.352), 208165_s_at		PIP (0.328), 206509_at					
EGFR (0.346), 201983_s_at		PRSS16 (0.326), 208165_s_at					
CCND1 (0.331), 208712_at		EGFR (0.320), 201983_s_at					
TRIM29 (0.325), 202504_at		DHCR7 (0.315), 201791_s_at					
DHCR7 (0.323), 201791_s_at		EPHX2 (0.315), 209368_at					
PIP (0.308), 206509_at		TRIM29 (0.311), 202504_at					
TFAP2B (0.306), 214451_at							
WNT5A (0.303), 205990_s_at							
APOD (0.301), 201525_at							
PTPRT (0.301), 205948_at							

У представленій нижче таблиці (таблиця 9) приведені послідовності праймерів і зондів для Q-ОТ-ПЦР, використані для представленої вище таблиці.

Таблица 9

Ген	зонд	прямой праймер	обратный праймер
ABAT	TCGCCCTAAGAGGCTCTTCCTC	GGCAACTTGAGGTCTGACTTTTG	GGTCAGCTCACAAGTGGTGTGA
ADRA2A	TTGTCTTTTCCCCCTCCGTGC	CCCCAAGAGCTGTTAGGTATCAA	TCAATGACATGATCTCAACCAGAA
APOD	CATCAGCTCTCAACTCTGGTTTAACA	ACTCACTAATGGAAAACGGAAAGATC	TCACCTTCGATTTGATTACAGTT
ASPH	TGGGAGGAAGGCAAGGTGCTCATC	TGTGCCAACGAGACCAAGAC	TCGTGCTCAAAGGAGTCATCA
AURKA	CCGTACGCTGTGCTAGGCAT	AATCTGGAGGCAAGGTTTGA	TCTGGATTTGCTCTCTGTGAA
BIRC5	AGCCAGATGACGACCCCATAGAGGAACA	CCCAGTGTCTTCTCTGCTTCAAG	CAACCGGACGAATGCTTTTT
CCND1			
CELSR2	ACTGACTTTCTTCTGGAGCAGGTGGC	TCCAAGCATGTATTCCAGACTTGT	TGCCCCACAGCTCTTTTTCT
CHPT1	CCACGGCCACCGAAGAGGCAC	CGCTCGTGCTCATCTCTACT	CCCAGTGACATATAAAGGTATGTC
CXCL12	CCACAGCAGGGTTTCAGGTTCC	GCCACTACCCCTCTGAA	TCACCTTGCCAAAGTCTCTGAT
CYBRD1	AGGGCATCGCCATCATCGTC	GTCAACGGCTTCGTCTTCA	CAGGTCCACGGCAGTCTGT
DCN	CTTTTTCAGCAACCCGGTCCA	AAGGCTCTCTTATTCGGGTGTGA	TGGATGGCTGATCTCCAGTA
DHCR7	TGAGCGCCACCCCTCTCGA	GGGCTCTGTTCCCGATT	AGTCATAGGGCAAGCAGAAAATTC
E2F8	CAGGATACCTAATCCCTCTCACGCAG	AAATGTCTCCGCAACCTTGTTT	CTGCCCCCAGGGATGAG
EGFR			
EPHX2	TGAAGCGGGAGGACTTTTTGTAAA	CGATGAGAGTGTATTATCCATGCA	GCTGAGGCTGGGCTCTTCT
ESR1	ATGCCCTTTTGCCGATGCA	GCCAAATTGTGTTTGTATGGATTAA	GACAAAACCGAGTCACATCAGTAATAG
GJA1	TGCACAGCCTTTTGATTTCCCGAT	CGGGAAGCACCATCTCTAACTC	TTTATGTCCAGCAGCTAGTTTTT
HSPA2	CAAGTCAGCAAAACGCAAAA	CATGCACGAATAATCAAAATGC	ACATTATTCAGGTTTCTCTTAATGC
IL6ST	CAAGCTCCACCTTCCAAAGGACCT	CCCTGAATCCATAAAGGCATACC	CAGCTTCGTTTTTCCCTACTTTTT
INPP4B	TCCGAGCGCTGGATTGCATGAG	GCACCAGTTACACAAGGACTTCTTT	TCTCTATGCGGCATCCTTCTC
MAPT	AGACTATTGCACACTGCCGCT	GTGGCTCAAAGGATAATATCAACAC	ACCTTGCTCAGGTCAACTGGTT
MGP	CCTTCATATCCCTCAGCAGAGATGG	CCTTCATTAAACAGGAGAAATGCAA	ATTGAGCTCGTGGACAGGCTTA
NEK2	TCCTGAACAAATGAATCGCATGCTACAA	ATTGTGTGCACACCTTATTACATGT	AAGCAGCCCAATGACCAGATa
PCNA	AAATACTAAAATGCGCCGGCAATGA	GGGCGTGAACCTCACCAGTA	CTTCGCCCTTAGTGTATGATATC
PGR	TTGATAGAAACGCTGTGAGCTCGA	AGCTCATCAAGGCAATTGTTT	ACAAGATCATGCAAGTTATCAAGAAGTT
PIP	TGCATGGTGGTTAAACTTACCTCA	TGCTGCAGTTCAACAGAAATTG	CACCTTGTAGAGGATGCTCTCA
PLAT	CAGAAAGTGGCCATGCCACCTG	TGGGAAGACATGAATGCACACTA	GGAGGTTGGGCTTTAGTGAA
PRSS16	CACTGCCGGTCACCCACACCA	CTGAGGAGCACAGAACCTCAACT	CGAACTCGGTACATGTCTGATACAA
PTGER3	TCGGTCTGCTGGTCTCCGCTCC	CTGATTGAAGATCATTTCAACATCA	GACGGCCATTAGCTTATGG
PTPRT	TTGGCTTCTGGACACCTCACA	GAGTGTGGCTCTACCAATTGC	GAGCGGGAACCTTGGGATAG
RACGAP1	ACTGAGAATCTCCACCCGGCGCA	TCGCCAACTGGATAAATTGGA	GAATGTGCGGAATCTGTTTGAG
RBBP8	ACCGATTCCGCTACATTCACCCAAC	AGAAATTGGCTTCTGCTCAAG	AAAACCAACTTCCAAAAATCTCT
SCUBE2	CTAGAGGGTTCAGGTCCATACGTGACATA	TGTGGATTAGTTCAAGTCCAATG	CCATCTCGAACTATGTCTTCAATGAT
SEC14L2	TGGGAGGCATGCAACGCGTG	AGGTCTTACTAAGCAGTCCCATCTCT	CGACCGGCACCTGAACTC
SQLE	TATGCGTCTCCAAAAGAAGAACCTCG	GCAAGCTTCTTCTCTCTCA	CCTTTAGCAGTTTTCTCCATAGTTTTATATC
STC2	TCTCACCTTGACCTCAGCCAAG	ACATTGACAAATTTCCCTTAGGATT	CCAGGACGCAGCTTACCAA
TFAP2B	CAACACCACCACTAACAGGCACAGTC	GGCATGGACAAGATGTTCTTGA	CCTCCTGTGCGCAGTTTACT
TOP2A	CAGATCAGGACCAAGATGGTCCACAT	CATTGAAGACGCTTCTGTATGG	CCAGTTGTGATGGATAAAATTAATCAG
TRIM29	TGCTGTCTCACTACCGGCCATTCTACG	TGGAAATCTGGCAAGCAGACT	CAATCCCGTTGCCCTTGTG
UBE2C	TGAACACACATGCTGCCGAGCTCTG	CTTCTAGGAGAACCACATTGATAGT	GTTTCTTGACGGTACTCTTAAAGCT
WNT5A	TATTACATCCCTCAGTTGACAGTGAATTG	CTGTGGCTCTAATTTATTGCATAATG	TTAGTGCTTTTTGCTTCAAGATCTT

Друга альтернатива для неконтрольованого вибору можливих генів-кандидатів для заміни заснована тільки на даних Affymetrix. Це має ту перевагу, що його можна проводити тільки на основі вже опублікованих даних (наприклад, з [www.ncbi.nlm.nih.gov/geo/](http://www.ncbi.nlm.nih.gov/geo/)). У представленій нижче таблиці (таблиця 10) приводяться кандидати для заміни наборів зондів HG-U133a, використаних в алгоритмах T1-T5. Це засновано на підготовчих даних цих алгоритмів. Заголовок стовпця містить назву генів і ID набору зондів напівжирним шрифтом. Потім приведені 10 якнайкраще корелюючих наборів зондів, де кожен елемент таблиці містить ID набору зондів, коефіцієнт кореляції в дужках і назву гена.



Таблиця 10

UBE2C	BIRC5	DHCR7	RACGAP1	AURKA	PVALB	NMU	STC2
202954_at	202095_s_at	201791_s_at	222077_s_at	204092_s_at	205336_at	206023_at	203438_at
210052_s_at (0.82) TPX2	202954_at (0.82) UBE2C	201790_s_at (0.66) DHCR7	218039_at (0.79) NUSAP1	208079_s_at (0.89) STK6	208683_at (-0.33) CAPN2	205347_s_at (0.45) TMSL8	203439_s_at (0.88) STC2
202095_s_at (0.82) BIRC5	218039_at (0.81) NUSAP1	202218_s_at (0.48) FADS2	214710_s_at (0.78) CCNB1	202954_at (0.80) UBE2C	219682_s_at (0.30) TBX3	203764_at (0.45) DLG7	212496_s_at (0.52) JMJD2B
218009_s_at (0.82) PRC1	218009_s_at (0.79) PRC1	202580_x_at (0.47) FOXM1	203764_at (0.77) DLG7	210052_s_at (0.77) TPX2	218704_at (0.30) FLJ20315	203554_x_at (0.44) PTTG1	219440_at (0.52) RAI2
203554_x_at (0.82) PTTG1	202705_at (0.78) CCNB2	208944_at (-0.46) TGFBR2	204026_s_at (0.77) ZWINT	202095_s_at (0.77) BIRC5		204962_s_at (0.44) CENPA	215867_x_at (0.51) CA12
208079_s_at (0.46) UBE2C	204962_s_at (0.78) CENPA	202954_at (0.46) UBE2C	218009_s_at (0.76) PRC1	203554_x_at (0.76) PTTG1		204825_at (0.43) MELK	214164_x_at (0.50) SEC14L2
202705_at (0.81) CCNB2	203554_x_at (0.78) PTTG1	209541_at (-0.45) IGF1	204641_at (0.76) NEK2	218009_s_at (0.75) PRC1		209714_s_at (0.41) CDKN3	204541_at (0.50) SEC14L2
218039_at (0.81) NUSAP1	208079_s_at (0.77) STK6	201059_at (0.45) CTIN	204444_at (0.75) KIF11	201292_at (0.73) TOP2A		219918_s_at (0.41) ASPM	203963_at (0.49) FLNB
202870_s_at (0.80) CDC20	210052_s_at (0.77) TPX2	200795_at (-0.45) SPARCL1	202705_at (0.75) CCNB2	214710_s_at (0.73) CCNB1		207828_s_at (0.41) CENPF	212495_at (0.50) JMJD2B
204092_s_at (0.77) STK6	202580_x_at (0.77) FOXM1	218009_s_at (0.45) PRC1	203362_s_at (0.75) MAD2L1	204962_s_at (0.73) CENPA		202705_at (0.41) CCNB2	208614_s_at (0.49) FLNB
209408_at (0.80) KIF2C	204092_s_at (0.77) STK6	218542_at (0.45) C10orf3	202954_at (0.75) UBE2C	218039_at (0.73) NUSAP1		219787_s_at (0.40) ECT2	213933_at (0.49) PTGER3

AZGP1	RBBP8	IL6ST	MGP	PTGER3	CXCL12	ABAT	CDH1
209309_at	203344_s_at	212196_at	202291_s_at	213933_at	209687_at	209460_at	201131_s_at
217014_s_at (0.92) AZGP1	36499_at (0.49) CELSR2	212195_at (0.85) IL6ST	201288_at (0.46) ARHGDIB	210375_at (0.74) PTGER3	204955_at (0.81) SRPX	209459_s_at (0.92) ABAT	201130_s_at (0.57) CDH1
206509_at (0.52) PIP	204029_at (0.45) CELSR2	204864_s_at (0.75) IL6ST	219768_at (0.42) VTCN1	210831_s_at (0.74) PTGER3	209335_at (0.81) DCN	206527_at (0.63) ABAT	221597_s_at (0.40) HSPC171
						213392_at (0.54)	
204541_at (0.46) SEC14L2	208305_at (0.45) PGR	211000_s_at (0.68) IL6ST	202849_x_at (-0.41) GRK6	210374_x_at (0.73) PTGER3	211896_s_at (0.81) DCN	MGC35048	203350_at (0.38) AP1G1
200670_at (0.45) XBP1	205380_at (0.43) PDZK1	214077_x_at (0.61) MEIS4	205382_s_at (0.40) DF	210832_x_at (0.73) PTGER3	201893_x_at (0.81) DCN	221666_s_at (0.49) PYCARD	209163_at (0.36) CYB561
209368_at (0.45) EPHX2	203303_at (0.41) TCTE1L	204863_s_at (0.58) IL6ST	200099_s_at (0.39) RPS3A	210834_s_at (0.55) PTGER3	203666_at (0.80) CXCL12	218016_s_at (0.48) POLR3E	210239_at (0.35) IRX5
218627_at (-0.43) FLJ11259	205280_at (0.38) GLRB	202089_s_at (0.57) SLC39A6	221591_s_at (-0.37) FAM64A	210833_at (0.55) PTGER3	211813_x_at (0.80) DCN	214440_at (0.46) NAT1	200942_s_at (0.34) HSBP1
202286_s_at (0.43) TACSTD2	205279_s_at (0.38) GLRB	210735_s_at (0.56) CA12	214629_x_at (0.37) RTN4	203438_at (0.49) STC2	208747_s_at (0.79) C1S	204981_at (0.45) SLC22A18	209157_at (0.34) DNAJA2
213832_at (0.42) ---	203685_at (0.38) BCL2	200648_s_at (0.52) GLUL	200748_s_at (0.37) FTH1	203439_s_at (0.46) STC2	203131_at (0.78) PDGFRA	212195_at (0.45) IL6ST	210715_s_at (0.33) SPINT2
204288_s_at (0.41) SORBS2	203304_at (-0.38) BAMBI	214552_s_at (0.52) RABEP1	209408_at (-0.37) KIF2C	212195_at (0.41) IL6ST	202994_s_at (0.78) FBLN1	204497_at (0.45) ADCY9	203219_s_at (0.33) APRT
			218726_at (-0.36)				
202376_at (0.41) SERPINA3	205862_at (0.36) GREB1	219197_s_at (0.51) SCUBE2	DKFZp762E1312	217764_s_at (0.40) RAB31	208944_at (0.78) TGFBR2	215867_x_at (0.45) CA12	218074_at (0.33) FAM96B

Після вибору гена або набору зондів необхідно визначити математичне картування між величинами експресії гена для заміни і величинами експресії нового гена. Існує декілька альтернатив, які обговорюються в даному описі на основі прикладу "величин дельта-Ct для заміни BIRC5 на RACGAP1". У підготовчих даних сумісний розподіл для експресії схожий з фігурою 3.

Коефіцієнт кореляції Пірсона дорівнює 0,73.

Одним підходом є створення функції картування від RACGAP1 до BIRC5 за допомогою регресії. Лінійна регресія є першим вибором і результатом в даному прикладі

$$BIRC5 = 1,22 \cdot RACGAP1 - 2,85.$$

З використанням цього рівняння можна легко замінити змінну BIRC5, наприклад, в алгоритмі T5 правою частиною. У інших прикладах може бути достатньо стійкої регресії, поліноміальній регресії або одновимірних нелінійних пре-перетворень.

Регресійний метод припускає вимірювання шумового сигналу на BIRC5, але без шумового сигналу на RACGAP1. Таким чином, картування не є симетричним відносно взаємозамінюваності двох змінних. Підхід симетричного картування заснований на двох одновимірних z-перетворюваннях.

$$z = (BIRC5 - \text{середнє значення}(BIRC5)) / \text{std}(BIRC5) \text{ і}$$

$$z = (RACGAP1 - \text{середнє значення}(RACGAP1)) / \text{std}(RACGAP1)$$

$$z = (BIRC5 - 8,09) / 1,29 = (RACGAP1 - 8,95) / 0,77$$

BIRC5=1,67\*RACGAP1+-6,89

Знов, в інших прикладах можуть бути достатніми інші перетворення: нормалізація по медіані і/або середньому абсолютному відхиленню, нелінійні картування або інше.

ПЕРЕЛІК ПОСЛІДОВНОСТЕЙ

```

<110> Sividon Diagnostics GmbH

<120> СПОСІБ ПЕРЕДБАЧУВАННЯ РЕЦИДИВУ РАКА МОЛОЧНОЇ ЗАЛОЗИ
      ПРИ ЕНДОКРИННОМУ ЛІКУВАННІ

<130> Sividon

<160> 141

<170> PatentIn version 3.3

<210> 1
<211> 22
<212> ДНК
<213> Homo sapiens

<400> 1
tcgccctaag aggtctcttc tc                22

<210> 2
<211> 23
<212> ДНК
<213> Homo sapiens

<400> 2
ggcaacttga ggtctgactt ttg                23

<210> 3
<211> 22
<212> ДНК
<213> Homo sapiens

<400> 3
ggtcagctca caagtgggtg ga                22

<210> 4
<211> 22
<212> ДНК
<213> Homo sapiens

<400> 4
ttgtcctttc cccctccgt gc                22

<210> 5
<211> 23
<212> ДНК
<213> Homo sapiens

<400> 5
ccccaagagc tgtaggtat caa                23

<210> 6
<211> 24
<212> ДНК
<213> Homo sapiens

```

<400> 6  
tcaatgacat gatctcaacc agaa 24

<210> 7  
<211> 27  
<212> ДНК  
<213> Homo sapiens

<400> 7  
catcagctct caactcctgg ttaaca 27

<210> 8  
<211> 26  
<212> ДНК  
<213> Homo sapiens

<400> 8  
actcactaat ggaaaacgga aagatc 26

<210> 9  
<211> 24  
<212> ДНК  
<213> Homo sapiens

<400> 9  
tcaccttcga ttgattcac agtt 24

<210> 10  
<211> 24  
<212> ДНК  
<213> Homo sapiens

<400> 10  
tgaggaggaag gcaagggtgct catc 24

<210> 11  
<211> 20  
<212> ДНК  
<213> Homo sapiens

<400> 11  
tgtgccaacg agaccaagac 20

<210> 12  
<211> 21  
<212> ДНК  
<213> Homo sapiens

<400> 12  
tcgtgctcaa aggagtcac a 21

<210> 13  
<211> 21  
<212> ДНК  
<213> Homo sapiens

<400> 13

ccgtcagcct gtgctaggca t	21
<210> 14	
<211> 20	
<212> ДНК	
<213> Homo sapiens	
<400> 14	
aatctggagg caaggttcga	20
<210> 15	
<211> 21	
<212> ДНК	
<213> Homo sapiens	
<400> 15	
tctggatttg cctcctgtga a	21
<210> 16	
<211> 28	
<212> ДНК	
<213> Homo sapiens	
<400> 16	
agccagatga cgaccccata gaggaaca	28
<210> 17	
<211> 23	
<212> ДНК	
<213> Homo sapiens	
<400> 17	
cccagtgttt cttctgcttc aag	23
<210> 18	
<211> 20	
<212> ДНК	
<213> Homo sapiens	
<400> 18	
caaccggacg aatgcttttt	20
<210> 19	
<211> 27	
<212> ДНК	
<213> Homo sapiens	
<400> 19	
actgactttc cttctggagc aggtggc	27
<210> 20	
<211> 24	
<212> ДНК	
<213> Homo sapiens	
<400> 20	
tccaagcatg tattccagac ttgt	24

<210>	21	
<211>	20	
<212>	ДНК	
<213>	Homo sapiens	
<400>	21	
	tgccacagc ctcttttct	20
<210>	22	
<211>	21	
<212>	ДНК	
<213>	Homo sapiens	
<400>	22	
	ccacggccac cgaagaggca c	21
<210>	23	
<211>	21	
<212>	ДНК	
<213>	Homo sapiens	
<400>	23	
	cgctcgtgct catctcctac t	21
<210>	24	
<211>	24	
<212>	ДНК	
<213>	Homo sapiens	
<400>	24	
	cccagtgcac ataaaaggta tgtc	24
<210>	25	
<211>	22	
<212>	ДНК	
<213>	Homo sapiens	
<400>	25	
	ccacagcagg gtttcagggt cc	22
<210>	26	
<211>	19	
<212>	ДНК	
<213>	Homo sapiens	
<400>	26	
	gccactacc cctcctgaa	19
<210>	27	
<211>	22	
<212>	ДНК	
<213>	Homo sapiens	
<400>	27	
	tcaccttgcc aacagttctg at	22

<210>	28	
<211>	20	
<212>	ДНК	
<213>	Homo sapiens	
<400>	28	
	agggcacgc catcatcgtc	20
<210>	29	
<211>	19	
<212>	ДНК	
<213>	Homo sapiens	
<400>	29	
	gtcacccggt tcgtcttca	19
<210>	30	
<211>	19	
<212>	ДНК	
<213>	Homo sapiens	
<400>	30	
	caggccacg gcagtctgt	19
<210>	31	
<211>	21	
<212>	ДНК	
<213>	Homo sapiens	
<400>	31	
	tcttttcagc aaccgggtcc a	21
<210>	32	
<211>	22	
<212>	ДНК	
<213>	Homo sapiens	
<400>	32	
	aaggcttctt attcgggtgt ga	22
<210>	33	
<211>	22	
<212>	ДНК	
<213>	Homo sapiens	
<400>	33	
	tggatggctg tatctcccag ta	22
<210>	34	
<211>	19	
<212>	ДНК	
<213>	Homo sapiens	
<400>	34	
	tgagcgcca ccctctcga	19

<210>	35	
<211>	18	
<212>	ДНК	
<213>	Homo sapiens	
<400>	35	
	gggctctgct tcccgatt	18
<210>	36	
<211>	24	
<212>	ДНК	
<213>	Homo sapiens	
<400>	36	
	agtcataaggg caagcagaaa attc	24
<210>	37	
<211>	26	
<212>	ДНК	
<213>	Homo sapiens	
<400>	37	
	caggatacct aatccctctc acgcag	26
<210>	38	
<211>	22	
<212>	ДНК	
<213>	Homo sapiens	
<400>	38	
	aaatgtctcc gcaaccttgt tc	22
<210>	39	
<211>	17	
<212>	ДНК	
<213>	Homo sapiens	
<400>	39	
	ctgccccag ggatgag	17
<210>	40	
<211>	24	
<212>	ДНК	
<213>	Homo sapiens	
<400>	40	
	tgaagcggga ggactttttg taaa	24
<210>	41	
<211>	24	
<212>	ДНК	
<213>	Homo sapiens	
<400>	41	
	cgatgagagt gttttatcca tgca	24
<210>	42	

```

<211> 19
<212> ДНК
<213> Homo sapiens

<400> 42
gctgaggctg ggctcttct 19

<210> 43
<211> 19
<212> ДНК
<213> Homo sapiens

<400> 43
atgccctttt gccgatgca 19

<210> 44
<211> 24
<212> ДНК
<213> Homo sapiens

<400> 44
gccaaattgt gtttgatgga ttaa 24

<210> 45
<211> 27
<212> ДНК
<213> Homo sapiens

<400> 45
gacaaaaccg agtcacatca gtaatag 27

<210> 46
<211> 25
<212> ДНК
<213> Homo sapiens

<400> 46
tgcacagcct tttgatttcc ccgat 25

<210> 47
<211> 22
<212> ДНК
<213> Homo sapiens

<400> 47
cggggaagcac catctctaac tc 22

<210> 48
<211> 24
<212> ДНК
<213> Homo sapiens

<400> 48
ttcatgtcca gcagctagtt tttt 24

<210> 49
<211> 21

```



```

<212> ДНК
<213> Homo sapiens

<400> 49
caagtcagca aacacgcaaa a 21

<210> 50
<211> 24
<212> ДНК
<213> Homo sapiens

<400> 50
catgcacgaa ctaatcaaaa atgc 24

<210> 51
<211> 27
<212> ДНК
<213> Homo sapiens

<400> 51
acattattcg aggtttctct ttaatgc 27

<210> 52
<211> 24
<212> ДНК
<213> Homo sapiens

<400> 52
caagctccac cttccaaagg acct 24

<210> 53
<211> 23
<212> ДНК
<213> Homo sapiens

<400> 53
ccctgaatcc ataaaggcat acc 23

<210> 54
<211> 24
<212> ДНК
<213> Homo sapiens

<400> 54
cagcttcggt tttccctact tttt 24

<210> 55
<211> 22
<212> ДНК
<213> Homo sapiens

<400> 55
tccgagcgct ggattgcatg ag 22

<210> 56
<211> 25
<212> ДНК

```

<213> Homo sapiens	
<400> 56	
gcaccagtta cacaaggact tcttt	25
<210> 57	
<211> 21	
<212> ДНК	
<213> Homo sapiens	
<400> 57	
tctctatgcg gcaccccttct c	21
<210> 58	
<211> 23	
<212> ДНК	
<213> Homo sapiens	
<400> 58	
agactatttg cacactgccg cct	23
<210> 59	
<211> 26	
<212> ДНК	
<213> Homo sapiens	
<400> 59	
gtggctcaaa ggataatatc aaacac	26
<210> 60	
<211> 22	
<212> ДНК	
<213> Homo sapiens	
<400> 60	
accttgctca ggtcaactgg tt	22
<210> 61	
<211> 26	
<212> ДНК	
<213> Homo sapiens	
<400> 61	
ccttcatatc ccctcagcag agatgg	26
<210> 62	
<211> 24	
<212> ДНК	
<213> Homo sapiens	
<400> 62	
ccttcattaa caggagaaat gcaa	24
<210> 63	
<211> 22	
<212> ДНК	
<213> Homo sapiens	

<400> 63	
attgagctcg tggacaggct ta	22
<210> 64	
<211> 30	
<212> ДНК	
<213> Homo sapiens	
<400> 64	
tcctgaacaa atgaatcgca tgctctacaa	30
<210> 65	
<211> 26	
<212> ДНК	
<213> Homo sapiens	
<400> 65	
atttggtggc acaccttatt acatgt	26
<210> 66	
<211> 21	
<212> ДНК	
<213> Homo sapiens	
<400> 66	
aagcagccca atgaccagat a	21
<210> 67	
<211> 25	
<212> ДНК	
<213> Homo sapiens	
<400> 67	
aaatactaaa atgcgccggc aatga	25
<210> 68	
<211> 20	
<212> ДНК	
<213> Homo sapiens	
<400> 68	
ggcgtgaac ctcaccagta	20
<210> 69	
<211> 25	
<212> ДНК	
<213> Homo sapiens	
<400> 69	
cttcggccct tagtgtaatg atatc	25
<210> 70	
<211> 24	
<212> ДНК	
<213> Homo sapiens	

<400> 70	
ttgatagaaa cgctgtgagc tcga	24
<210> 71	
<211> 22	
<212> ДНК	
<213> Homo sapiens	
<400> 71	
agctcatcaa ggcaattggt tt	22
<210> 72	
<211> 28	
<212> ДНК	
<213> Homo sapiens	
<400> 72	
acaagatcat gcaagttatc aagaagtt	28
<210> 73	
<211> 25	
<212> ДНК	
<213> Homo sapiens	
<400> 73	
tgcatggtgg ttaaaactta cctca	25
<210> 74	
<211> 23	
<212> ДНК	
<213> Homo sapiens	
<400> 74	
tgcttgtagt tcaaacagaa ttg	23
<210> 75	
<211> 23	
<212> ДНК	
<213> Homo sapiens	
<400> 75	
caccttgtag agggatgctg cta	23
<210> 76	
<211> 23	
<212> ДНК	
<213> Homo sapiens	
<400> 76	
cagaaagtgg ccatgccacc ctg	23
<210> 77	
<211> 23	
<212> ДНК	
<213> Homo sapiens	
<400> 77	

tggaagaca tgaatgcaca cta	23
<210> 78	
<211> 21	
<212> ДНК	
<213> Homo sapiens	
<400> 78	
ggaggttggg ctttagctga a	21
<210> 79	
<211> 21	
<212> ДНК	
<213> Homo sapiens	
<400> 79	
cactgccggt caccacacc a	21
<210> 80	
<211> 23	
<212> ДНК	
<213> Homo sapiens	
<400> 80	
ctgaggagca cagaacctca act	23
<210> 81	
<211> 25	
<212> ДНК	
<213> Homo sapiens	
<400> 81	
cgaactcggg acatgtctga taca	25
<210> 82	
<211> 22	
<212> ДНК	
<213> Homo sapiens	
<400> 82	
tcggtctgct ggtctccgct cc	22
<210> 83	
<211> 26	
<212> ДНК	
<213> Homo sapiens	
<400> 83	
ctgattgaag atcattttca acatca	26
<210> 84	
<211> 20	
<212> ДНК	
<213> Homo sapiens	
<400> 84	
gacggccatt cagcttatgg	20

<210>	85	
<211>	22	
<212>	ДНК	
<213>	Homo sapiens	
<400>	85	
	ttggcttctg gacacctca ca	22
<210>	86	
<211>	22	
<212>	ДНК	
<213>	Homo sapiens	
<400>	86	
	gagttgtggc ctctaccatt gc	22
<210>	87	
<211>	20	
<212>	ДНК	
<213>	Homo sapiens	
<400>	87	
	gagcgggaac cttgggatag	20
<210>	88	
<211>	23	
<212>	ДНК	
<213>	Homo sapiens	
<400>	88	
	actgagaatc tccaccggc gca	23
<210>	89	
<211>	21	
<212>	ДНК	
<213>	Homo sapiens	
<400>	89	
	tcgccaactg gataaattgg a	21
<210>	90	
<211>	22	
<212>	ДНК	
<213>	Homo sapiens	
<400>	90	
	gaatgtgagg aatctgtttg ag	22
<210>	91	
<211>	26	
<212>	ДНК	
<213>	Homo sapiens	
<400>	91	
	accgattccg ctacattcca cccaac	26

<210> 92  
 <211> 22  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 92  
 agaaattggc ttctgtctca ag 22  
  
 <210> 93  
 <211> 25  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 93  
 aaaaccaact tcccaaaaat tctct 25  
  
 <210> 94  
 <211> 31  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 94  
 ctagagggtt ccagggtcca tacgtgacat a 31  
  
 <210> 95  
 <211> 24  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 95  
 tgtggattca gttcaagtcc aatg 24  
  
 <210> 96  
 <211> 27  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 96  
 ccatctcgaa ctatgtcttc aatgagt 27  
  
 <210> 97  
 <211> 20  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 97  
 tgggaggcat gcaacgcgtg 20  
  
 <210> 98  
 <211> 26  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 98  
 aggtcttact aagcagtccc atctct 26

```

<210> 99
<211> 18
<212> ДНК
<213> Homo sapiens

<400> 99
cgaccggcac ctgaactc 18

<210> 100
<211> 29
<212> ДНК
<213> Homo sapiens

<400> 100
tatgcgtctc ccaaaagaag aacacctcg 29

<210> 101
<211> 21
<212> ДНК
<213> Homo sapiens

<400> 101
gcaagcttcc ttcttccttc a 21

<210> 102
<211> 31
<212> ДНК
<213> Homo sapiens

<400> 102
ccttttagcag ttttctccat agttttatat c 31

<210> 103
<211> 27
<212> ДНК
<213> Homo sapiens

<400> 103
caacaccacc actaacaggc acacgtc 27

<210> 104
<211> 22
<212> ДНК
<213> Homo sapiens

<400> 104
ggcatggaca agatgttctt ga 22

<210> 105
<211> 22
<212> ДНК
<213> Homo sapiens

<400> 105
cctccttgtc gccagtttta ct 22

<210> 106

```



<211> 28  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 106  
 cagatcagga ccaagatggt tcccacat 28  
  
 <210> 107  
 <211> 22  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 107  
 cattgaagac gcttcgttat gg 22  
  
 <210> 108  
 <211> 27  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 108  
 ccagttgtga tggataaaat taatcag 27  
  
 <210> 109  
 <211> 27  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 109  
 tgctgtctca ctaccggcca ttctacg 27  
  
 <210> 110  
 <211> 21  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 110  
 tggaaatctg gcaagcagac t 21  
  
 <210> 111  
 <211> 20  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 111  
 caatcccgtt gcctttgttg 20  
  
 <210> 112  
 <211> 25  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 112  
 tgaacacaca tgctgccgag ctctg 25  
  
 <210> 113  
 <211> 27

<212> ДНК  
 <213> Homo sapiens  
  
 <400> 113  
 cttctaggag aaccaacat tgatagt 27  
  
 <210> 114  
 <211> 27  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 114  
 gtttcttgca ggtacttctt aaaagct 27  
  
 <210> 115  
 <211> 30  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 115  
 tattcacatc ccctcagtg cagtgaattg 30  
  
 <210> 116  
 <211> 27  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 116  
 ctgtggctct taatttattg cataatg 27  
  
 <210> 117  
 <211> 26  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 117  
 ttagtgcttt ttgctttcaa gatctt 26  
  
 <210> 118  
 <211> 23  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 118  
 tctcaccttg accctcagcc aag 23  
  
 <210> 119  
 <211> 26  
 <212> ДНК  
 <213> Homo sapiens  
  
 <400> 119  
 acatttgaca aatttcctt aggatt 26  
  
 <210> 120  
 <211> 20  
 <212> ДНК

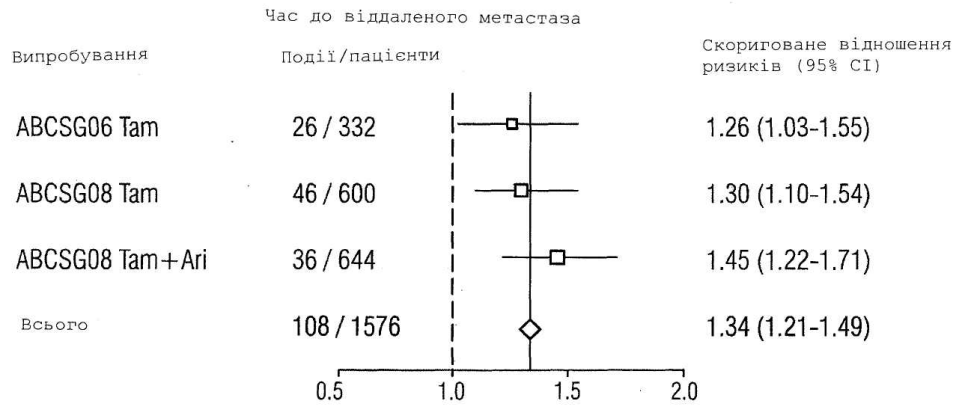
<213> Homo sapiens	
<400> 120	
ccaggacgca gctttaccaa	20
<210> 121	
<211> 20	
<212> ДНК	
<213> Homo sapiens	
<400> 121	
caccagccac caggccccag	20
<210> 122	
<211> 18	
<212> ДНК	
<213> Homo sapiens	
<400> 122	
tcttgaccg gcaagatc	18
<210> 123	
<211> 21	
<212> ДНК	
<213> Homo sapiens	
<400> 123	
taggccaggc acttcagttt c	21
<210> 124	
<211> 22	
<212> ДНК	
<213> Homo sapiens	
<400> 124	
tcgcgtctcg gaaaccgga gc	22
<210> 125	
<211> 20	
<212> ДНК	
<213> Homo sapiens	
<400> 125	
gagcgagctg agtggttg	20
<210> 126	
<211> 21	
<212> ДНК	
<213> Homo sapiens	
<400> 126	
agtcagttgg tcagccatgc t	21
<210> 127	
<211> 27	
<212> ДНК	
<213> Homo sapiens	

<400> 127	
cctgccaatc ccgatgaaat tggaaat	27
<210> 128	
<211> 21	
<212> ДНК	
<213> Homo sapiens	
<400> 128	
tgagtgtccc ccggtatctt c	21
<210> 129	
<211> 21	
<212> ДНК	
<213> Homo sapiens	
<400> 129	
tcagccgctt tcagattttc a	21
<210> 130	
<211> 26	
<212> ДНК	
<213> Homo sapiens	
<400> 130	
accctgctga cctttctcca ttccgt	26
<210> 131	
<211> 22	
<212> ДНК	
<213> Homo sapiens	
<400> 131	
agaaattggc ttctgtctca ag	22
<210> 132	
<211> 25	
<212> ДНК	
<213> Homo sapiens	
<400> 132	
aaaaccaact tcccaaaaat tctct	25
<210> 133	
<211> 23	
<212> ДНК	
<213> Homo sapiens	
<400> 133	
tgcttcaca agaaccgcga gga	23
<210> 134	
<211> 20	
<212> ДНК	
<213> Homo sapiens	

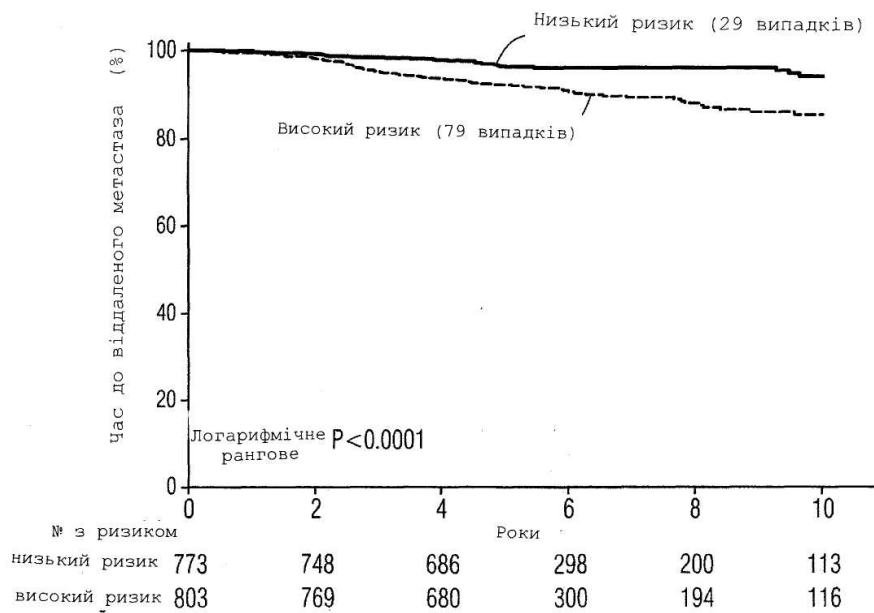
<400> 134	
cgagccgacc atgtcttcat	20
<210> 135	
<211> 21	
<212> ДНК	
<213> Homo sapiens	
<400> 135	
aagcccaaaa agctgaaggt t	21
<210> 136	
<211> 22	
<212> ДНК	
<213> Homo sapiens	
<400> 136	
aagtctctcc aaatggtcgg cc	22
<210> 137	
<211> 19	
<212> ДНК	
<213> Homo sapiens	
<400> 137	
ccgactcctt cgaccacaa	19
<210> 138	
<211> 23	
<212> ДНК	
<213> Homo sapiens	
<400> 138	
catcatcgc actctttttc ttc	23
<210> 139	
<211> 18	
<212> ДНК	
<213> Homo sapiens	
<400> 139	
tggtggtggcgc tgcttgga	18
<210> 140	
<211> 21	
<212> ДНК	
<213> Homo sapiens	
<400> 140	
tgtggttcct gcatgaagac a	21
<210> 141	
<211> 24	
<212> ДНК	
<213> Homo sapiens	
<400> 141	
gtgacagcgg aagtgggtatt gtac	24

## ФОРМУЛА ВИНАХОДУ

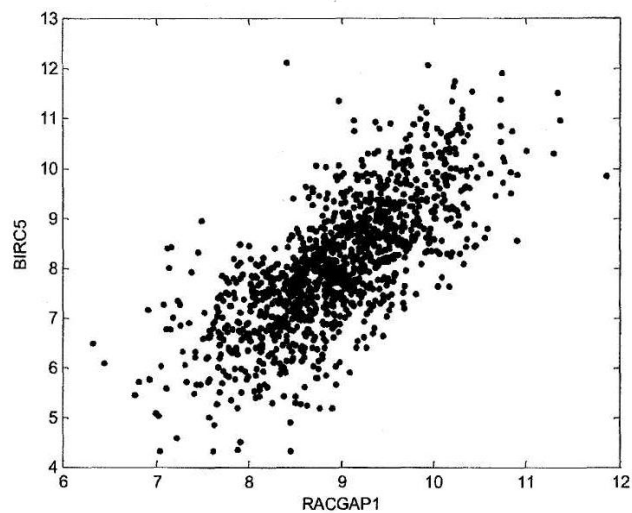
1. Спосіб передбачування наслідків раку молочної залози в позитивній по рецептору естрогена і негативній по HER2 пухлині у пацієнта з раком молочної залози, причому вказаний спосіб включає:
  - (a) визначення в зразку пухлини від вказаного пацієнта рівнів експресії РНК наступних 8 генів: UBE2C, BIRC5, DHCR7, STC2, AZGP1, RBBP8, IL6ST і MGP;
  - (b) математичне комбінування величин рівня експресії для генів вказаного набору, причому ці величини визначають в зразку пухлини з отриманням комбінованого показника, де вказаний комбінований показник указує на прогноз у вказаного пацієнта.
2. Спосіб за п. 1, де вказаному пацієнту проведена ендокринна терапія або йому передбачають проводити ендокринну терапію.
3. Спосіб за п. 2, де вказана ендокринна терапія включає тамоксифен або інгібітор ароматази.
4. Спосіб за будь-яким з пп. 1-3, де передбачають ризик розвитку рецидиву раку молочної залози або пов'язаної з ризиком смерті.
5. Спосіб за будь-яким з пп. 1-4, де вказаний рівень експресії визначають як рівень експресії матричної РНК.
6. Спосіб за п. 5, де вказаний рівень експресії визначають за допомогою щонайменше одного зі способів: спосіб на основі ПЛР, спосіб на основі мікроматриці і спосіб на основі гібридизації.
7. Спосіб за будь-яким з попередніх пунктів, де вказане визначення рівнів експресії проводять у фіксованому у формаліні зануреному в парафін зразку пухлини або в свіжозамороженому зразку пухлини.
8. Спосіб за будь-яким з вказаних вище пунктів, де рівень експресії щонайменше одного маркерного гена визначають як профіль експресії відносно щонайменше одного еталонного гена або щодо обчисленої середньої величини експресії.
9. Спосіб за будь-яким з попередніх пунктів, де вказана стадія математичного комбінування включає стадію застосування алгоритму до величин, що характеризують рівні експресії даних генів, де вказаний алгоритм є лінійною комбінацією вказаних величин, характерних для рівнів експресії даних генів.
10. Спосіб за п. 9, де значення характерної величини рівня експресії даного гена множать на коефіцієнт.
11. Спосіб за будь-яким з попередніх пунктів, де один, два або більше порогових значень визначають для вказаного комбінованого показника, які відокремлюють в групі високого і низького ризику, високого, проміжного і низького ризику або на більшу кількість груп ризику шляхом застосування порогу до комбінованого показника, причому високий комбінований показник указує на користь цитотоксичної хіміотерапії.
12. Спосіб за будь-яким з попередніх пунктів, де інформація відносно статусу вузлів у пацієнта переробляється на стадії математичного комбінування величин рівня експресії для генів з одержанням комбінованого показника.
13. Спосіб за п. 11 або 12, де вказана інформація, що стосується статусу вузлів, є числовою величиною, якщо вказаний статус є негативним, і вказана інформація є відмінною числовою величиною, якщо вказаний статус вузлів є позитивним, і вказана інформація є відмінною або ідентичною числовою величиною, якщо вказаний статус вузлів є невідомим.
14. Застосування набору для виконання способу за будь-яким з пп. 1-13, причому вказаний набір містить набір олігонуклеотидів, здатних специфічно зв'язувати послідовності або зв'язуватися з послідовностями фрагментів генів в комбінації генів, де вказана комбінація містить 8 генів: UBE2C, BIRC5, DHCR7, STC2, AZGP1, RBBP8, IL6ST і MGP.



Фіг. 1



Фіг. 2



Фіг. 3

---

Комп'ютерна верстка О. Рябко

---

Державна служба інтелектуальної власності України, вул. Василя Липківського, 45, м. Київ, МСП, 03680, Україна

---

ДП "Український інститут інтелектуальної власності", вул. Глазунова, 1, м. Київ – 42, 01601