

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ

Қ.И. Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті

Қ.Тұрысов атындағы Геология және мұнай-газ ісі институты

Химиялық және биохимиялық инженерия кафедрасы

Байжума Жулдыз Багдадқызы,

Уалихан Алина Серікқызы

«Биоинформатикалық құралдарды пайдалана отырып сүт безі қатерлі ісігіндегі негізгі гендердің экспрессиясын бағалау»

ДИПЛОМДЫҚ ЖОБА

6В05101—«Химиялық және биохимиялық инженерия» мамандығы

Алматы 2024

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ ҒЫЛЫМ ЖӘНЕ ЖОҒАРЫ БІЛІМ
МИНИСТРЛІГІ

Қ.И.Сәтбаев атындағы Қазақ ұлттық технологиялық зерттеу университеті
Қ.Тұрысов атындағы Геология және мұнай-газ ісі институты
Химиялық және биохимиялық инженерия кафедрасы

ҚОРҒАУҒА ЖІБЕРІЛДІ
«ХиБИ» кафедрасы кафедра
менгерушісі, PhD докторы
Амитова А.А
«04» 06 2024ж.



ДИПЛОМДЫҚ ЖОБА

Тақырыбы: «Биоинформатикалық құралдарды пайдалана отырып сүт
безі қатерлі ісігіндегі негізгі гендердің экспрессиясын бағалау»

6B05101–«Химиялық және биохимиялық инженерия» мамандығы

Орындаған:

Байжума Жулдыз,
Уалихан Алина

Пікір беруші
Ph.D, биология ғылымдарының
кандидаты

 Асрадина С.Ш.
«07» маусым 2024 ж.

Ғылыми жетекші
Магистр, аға оқытушы


 Ботбаев Д.М.
«06» маусым 2024 ж.

Алматы 2024

ҚАЗАҚСТАН РЕСПУБЛИКАСЫ БІЛІМ ЖӘНЕ ҒЫЛЫМ МИНИСТРЛІГІ

Қ.И.Сәтбаев атындағы Қазақ ұлттық техникалық зерттеу университеті

«Қ.Тұрысов атындағы Геология және мұнай-газ ісі» институты

«Химиялық және биохимиялық инженерия»

кафедрасы 6В05101–«Химиялық және

биохимиялық инженерия»

БЕКІТЕМІН

Кафедра меңгерушісі

«ХиБИ» кафедрасы

PhD доктор

Амитова А.А

2024ж.



**Дипломдық жоба орындауға
ТАПСЫРМА**

Білім алушы Байжума Жұлдыз Бағдадқызы, Уалихан Алина Серікқызы
Тақырыбы «Биоинформатикалық құралдарды пайдалана отырып сүт безі қатерлі ісігіндегі негізгі гендердің экспрессиясын бағалау»

Университет ректорының 2023 жылғы "04" желтоқсан №548 -п/ө
бұйрығымен бекітілген.

Аяқталған жобаны тапсыру мерзімі: 2024 жылғы "7" маусым

Дипломдық жобада қарастырылатын мәселелер тізімі:

а) Әдебиетке шолу;

ә) Тәжірибелік бөлім;

б) Ген экспрессиясын талдау үшін биоинформатикалық құралдарды пайдалану;

в) GEO және TCGA сияқты әртүрлі көздер мен дерекқорлардан алынған деректердің үлкен көлемімен жұмыс істеу;

г) Ісіктерді люминальды А және В, HER2 оң және үш рет теріс сияқты әртүрлі молекулалық кіші түрлерге бөлу;

г) Зерттеу нәтижелерін талдау және талқылау

ж) Пайдаланылған әдебиеттер

Ұсынылатын негізгі әдебиет 43 атаудан тұрады.

Дипломдық жобаны даярлау

КЕСТЕСІ

Бөлім атаулары, дайындалатын сұрақтардың тізімі	Ғылыми жетекшіге, кеңесшілерге өткізу мерзімі	Ескерту
Әдеби шолу	24.11.2023ж	
Әдістер мен жұмысты орындау барысы	22.01.2024-20.03.2024ж	
Алынған нәтижелерді талдау	25.03.2024-08.04.2024ж	
Графикалық бөлім	10.04.2024ж	

Аяқталған дипломдық жобаның және оларға қатысты бөлімдерінің кеңесшілері мен қалып бақылаушының қолтаңбалары

Бөлімдер атауы	Ғылыми жетекші, кеңесшілер (аты-жөні, тегі, ғылыми дәрежесі, атағы)	Қолтаңба қойылған мерзім	Қолы
Әдістер мен жұмысты орындау барысы	Аға оқытушы, магистр Ботбаев Д.М		
Алынған нәтижелерді талдау	Аға оқытушы, магистр Ботбаев Д.М		
Қалып бақылаушы	Аға оқытушы, магистр Ботбаев Д.М		

Ғылыми жетекші



Ботбаев Д.М.

Тапсырманы орындауға алған білім алушы

Байжума Ж.Б.
 Уалихан А.С.

Күні

«20» 01 2024 ж.

АҢДАТПА

«Биоинформатикалық құралдарды пайдалана отырып сүт безі қатерлі ісігіндегі негізгі гендердің экспрессиясын бағалау» атты дипломдық жұмыс қағаз түрде 46 беттен тұрады. Жұмыс кіріспеден, 1 диаграммадан, 1 кестеден, 3 графиктен, 8 суреттен тұрады.

Мақсаты. Биоинформатикалық құралдарды қолдана отырып, сүт безі қатерлі ісігіндегі гендік экспрессияны толықтай зерттеу, бағалау және бүкіләлемдік дерекқорларын пайдалана отырып, сүт безі қатерлі ісігін диагностикалау және туындау себептерін анықтау.

Қойылған мақсатқа жету үшін биоинформатикалық құралдарды пайдалана отырып RNA-Seq project орындалды. Бұл деректерді өңдеу мен талдауға арналған негізгі қадамдармен көрсетіледі: шикі деректерді алу, оларды индекстеу және геноммен сәйкестендіру, дифференциалды ген экспрессиясын талдау жүргізілді.

Нәтижесінде қатерлі ісікке қатысатын негізгі гендер мен молекулалық жолдар қарастырылады. Негізгі ген экспрессиясын талдау нәтижелері ұсынылған. Қорытындыда зерттеу нәтижелері бойынша жасалған негізгі қорытындылар баяндалған.

АННОТАЦИЯ

Дипломная работа «Оценка экспрессии ключевых генов при раке молочной железы с использованием биоинформатических инструментов» состоит из 46 страниц в бумажном обложке. Работа состоит из введения, 1 диаграммы, 1 таблицы, 3 графиков, 8 рисунков.

Цель. Полное исследование экспрессии генов рака молочной железы с использованием биоинформатических инструментов, оценка и диагностика и выявление причин рака молочной железы с использованием всемирной базы данных.

Для достижения поставленной цели был реализован проект RNA-Seq с использованием биоинформатических инструментов. Это выражается в основных шагах для обработки и анализа данных: были получены необработанные данные, их индексация и идентификация с геномом, проведен дифференциальный анализ экспрессии генов.

В результате рассматриваются ключевые гены и молекулярные пути, участвующие в развитии рака. Представлены результаты анализа экспрессии генов. В заключении изложены основные выводы, сделанные по результатам исследования.

ANNOTATION

Thesis "Evaluation of the expression of key genes in breast cancer using bioinformatic tools" consists of 46 paperback pages. The work consists of an introduction, 1 diagram, 1 table, 3 graphs, 8 figures.

Purpose. A complete study of breast cancer gene expression using bioinformatic tools, assessment and diagnosis and identification of the causes of breast cancer using a worldwide database.

To achieve this goal, the RNA-Seq project was implemented using bioinformatic tools. This is expressed in the main steps for data processing and analysis: raw data were obtained, their indexing and identification with the genome, and a differential analysis of gene expression was performed.

As a result, key genes and molecular pathways involved in the development of cancer are considered. The results of the analysis of the expression of the main genes are presented. In conclusion, the main conclusions drawn from the results of the study are outlined.

МАЗМҰНЫ

	КІРІСПЕ	9
	НЕГІЗГІ БӨЛІМ	
1	ӘДЕБИЕТКЕ ШОЛУ	10
1.1	Сүт обыры жайлы статискалық мәліметтер	10
1.2	Сүт безі қатерлі ісігіндегі ген экспрессиясын профильдеу	12
1.3	Биоинформатика құралдары арқылы биомаркерлер таңдау	13
1.4	Сүт безі қатерлі ісігіндегі ген экспрессиясының классификациясы	16
1.5	Биоинформатикалық құралдармен сүт безі қатерлі ісігін диагностикалау	18
1.6	Сүт безіндегі машиналық оқыту тәсілдері	21
1.7	Сүт безі қатерлі ісігін емдеу әдістері	23
1.8	Биоинформатикалық талдау арқылы сүт безі қатерлі ісігіндегі ықтимал маңызды гендер	25
1.9	Биоинформатикадағы сүт безі қатерлі ісігінің жоғары өнімді геномдық тәсілдері	26
2.	Сүт безінің қатерлі ісігі кезінде биоинформатикалық бағдарламалық жасақтамасы	29
2.1	RNA-Seq деректерін талдау құралдары	29
2.2	Микрочип талдау құралдары	33
2.3	Жолдар мен желілерді талдау	34
3	Қолданылған материалдар мен әдістер	35
4	Зерттеу нәтижелері	

Қорытынды

Пайдаланылған әдебиеттер

Кіріспе

Сүт безі қатерлі ісігінің дамуының молекулалық механизмі әлі де жақсы түсінілмеген, бірақ тиімді емдеу үшін оның дамуы мен болжауына байланысты гендерді анықтау қажет. Бұл зерттеуде біз сүт безінің аденокарциномасындағы ықтимал патогенді және дифференциалды экспрессияланған гендерді анықтау үшін қолжетімді деректер жиынының биоинформатикалық талдауын қолдандық.

Сүт безінің қатерлі ісігі молекулалық және гистопатологиялық айырмашылықтарға байланысты әртүрлі түрлерге бөлінеді. Молекулалық классификация алты кіші топты ажыратады, олардың әрқайсысының өзіндік сипаттамалары мен болжамдары бар. Жақында жүргізілген мета-анализ TNRMG-дің төрт кіші түрін анықтады, олар бірегей клиникалық ерекшеліктерге ие және емдеудің жеке тәсілін қажет етеді. Сүт безі қатерлі ісігінің кіші топтары әртүрлі сипаттамаларға ие болғанымен, олар ұқсас белсендірілген немесе басылған гендер мен сигнал беру жолдарын көрсетеді. Бұл гендер мен жолдар сүт безі қатерлі ісігінің қалыптасуы мен дамуында шешуші рөл атқаруы мүмкін. Ген экспрессиясы профилі мен клиникалық деректердің биоақпараттық талдауы онкогенезге және аурудың болжамына байланысты орталық гендерді анықтауға көмектеседі.

Осылайша, біздің зерттеуіміз ауруды диагностикалау мен емдеудің дәл әдістерін қамтамасыз ету үшін сүт безі қатерлі ісігімен байланысты негізгі гендер мен молекулалық жолдарды зерттеуге бағытталған.

1. ӘДЕБИЕТКЕ ШОЛУ

1.1 Сүт безінің қатерлі ісігі таралу жиілігі

Қазіргі уақытта сүт безі қатерлі ісігі бүкіл әлемде қатерлі ісік ауруының негізгі себебі ретінде өкпе обырын басып озды, шамамен 2,3 миллион жаңа жағдай тіркелді, бұл барлық қатерлі ісіктердің 11,7% құрайды [1]. Сүт безінің қатерлі ісігі - бұл молекулалық төрт түрлі молекулалық кіші типтермен сипатталатын гетерогенді ауру. Метастатикалық емес аурудың ерте сатысында сүт безі қатерлі ісігімен ауыратын науқастардың шамамен 70-80% емдеуге болады, ал алыстағы органдардың метастаздары бар дамыған сүт безі қатерлі ісігі қазіргі уақытта қол жетімді емдеу әдістерімен емделмейді. Қазақстанда көптеген елдердегідей сүт безі қатерлі ісігі әйелдердің қатерлі ісіктері құрылымында бірінші орында. Орташа алғанда, елде жыл сайын 4000-ға жуық әйелге сүт безі қатерлі ісігі диагнозы қойылады және 1400 науқас сүт безі қатерлі ісігінен қайтыс болады [2]. Эпидемиологиялық зерттеулер планетаның әртүрлі аймақтарында аурудың гетерогенді таралуын көрсетті. 2022 жылғы мәліметтер бойынша сүт безі қатерлі ісігі 2,3 миллион әйелге диагноз қойылды және бүкіл әлемде бойынша 670 000 өлім тудырды [3]. Мұндай эпидемиологиялық жағдайда сүт безі қатерлі ісігі әрбір оныншы әйел үшін қауіпті диагноз нақты болады. Ісік жасушалары аз уақыт аралығында іргелес сүт безі тініне таралуы мүмкін. Ісік одан әрі өскен сайын көлемді масса немесе тығыздала болады. Инвазивті қатерлі ісік жақын маңдағы лимфа түйіндеріне немесе басқа мүшелерге таралуы мүмкін. Әйелдер ерлерге карағанда қатерлі ісікке жиірек ұшырайды және жиілік негізінен жыныс мүшелері мен сүт бездеріне түседі.

Тарихи тұрғыдан алғанда, сүт безі қатерлі ісігін емдеудің сипаты ауру туралы білім деңгейімен және ол туралы негізгі түсініктермен анықталды: Гиппократ кезінде емделуден толық бас тартудан, орта ғасырлардағы сүт безі қатерлі ісігін шөптермен және жергілікті каутеризациямен символдық емдеуден бастап, XX ғасырдың 50-ші жылдарының ортасында өте кеңейтілген кесу операцияларына дейін [4]. Кеңес Одағының аумағында алғашқы онкология институты 1928 жылы пайда болды. Онкологиялық аурулардың тарихы адамзат өмір бойы белгілі. Әйелдерде сүт безі қатерлі ісігінің жиілігі жасына қарай артып, шамамен 50 жаста шыңына жетеді. Сонымен қатар, менопаузадан кейін әйелдерде эстроген деңгейінің жоғарылауымен сүт безі қатерлі ісігінің даму қаупі артады. Сонымен қатар, қант диабеті мен семіздік сүт безі қатерлі ісігінің даму қаупінің жоғарылауымен байланысты болды. Ауылдық жерлерде тұратын әйелдер қалалық, әсіресе өнеркәсіптік аймақтарға карағанда сүт безінің қатерлі ісігімен аз ауырады [5]. Өнеркәсіптік аудандарда тұратын тұрғындарда сүт безі обырымен аурушылар санында байырғы тұрғындарда жиі байқалады, бұл әйелдердің денсаулығына және сүт безі обыры ауруына өнеркәсіптік фактордың әсерін растайды.

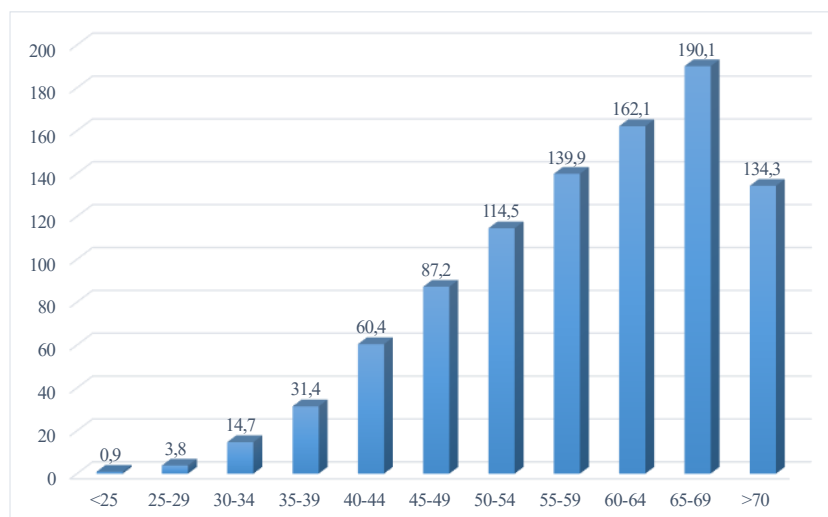
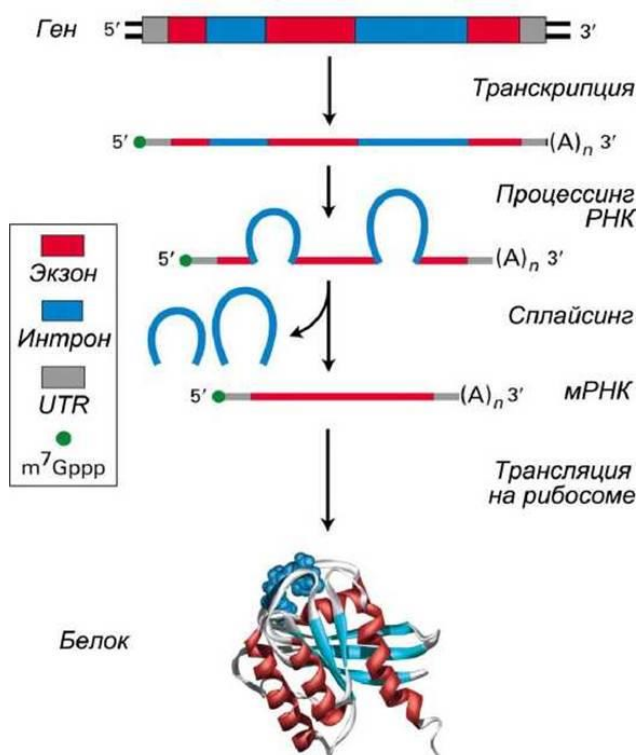


Диаграмма 1. 2012-2021 жылдары Қазақстанда жас топтары бойынша бөліністе 100 000 әйел халыққа шаққандағы сырқаттанушылық көрсеткіші.

Бұл зерттеуде 2012 жылдан 2021 жылға дейін Қазақстанда сүт безі обырының эпидемиологиялық сипаттамасы ұсынылған. Зерттеу он жыл ішінде сүт безі қатерлі ісігінің эпидемиологиялық сипаттамаларында айтарлықтай өзгерістерді көрсетті, бұл сүт безі қатерлі ісігінің алдын алу шараларының тиімділігін және елдегі соңғы жылдары сүт безі қатерлі ісігін емдеудің жаңа тәсілдерін енгізуді көрсетеді. Алынған нәтижелер соңғы онжылдықта Қазақстанның әйелдер халқының орташа жылдық өсу қарқынымен (1,33 қарсы 3,92) салыстырғанда сүт безі обырының таралуының орташа жылдық өсу қарқынының жоғарырақ екенін көрсетеді. Бұл тенденция бүкіләлемдік мәліметтерге сәйкес және оны сүт безі қатерлі ісігінің скринингтік бағдарламасының тиімділігімен түсіндіруге болады. Қазақстанда сүт безі обырымен сырқаттанушылықтың стандартталған салыстырмалы көрсеткіші (44,4) Батыс (46,6) және Шығыс (43,3) Азиядағы көрсеткіштермен бір деңгейде, бірақ Оңтүстік және Орталық Азияға (26,2) қарағанда жоғары. Сонымен қатар, зерттеу кезеңінде өлім-жітімнің төмендеу тенденциясы байқалды. Дегенмен, сүт безі қатерлі ісігінен болатын өлім-жітім басқа жас топтарымен салыстырғанда 65-69 және ≥ 70 жас топтарында айтарлықтай жоғары болды, ал 70 жастан асқан әйелдер арасында стандартталған өлім деңгейі ең жоғары болды. Бұл деректер 70 жастан асқан жас тобында сүт безі қатерлі ісігінен болатын жаһандық өлім-жітімнің статистикалық маңызды өсуін көрсетті. Басқа зерттеулер көрсеткендей, көптеген елдер мен аймақтарда сүт безі қатерлі ісігінің стандартталған салыстырмалы көрсеткіші 70 жастан асқан халық арасында ең жоғары болды және 2035 жылға дейін барлық жас топтарында өседі. Зерттеу нәтижелері 2012 жылдан бері салыстырмалы түрде тұрақты эпидемиологиялық жағдаймен салыстырғанда 2020 жылы сүт безі қатерлі ісігінен болатын ауру мен өлім-жітімнің айтарлықтай төмендегенін көрсетті.

1.2 Сүт обыры кезіндегі ген экспрессиясын профильдеу

Ген экспрессиясы - бұл ДНҚ гендерінен алынған ақпарат ақуыздар немесе рибонуклеин қышқылдары (РНҚ) сияқты функционалды өнімдерге айналатын процесс. Бұл процесс жасушалардағы өсу, даму, дифференциация және қоршаған ортаға жауап беру сияқты көптеген биологиялық процестерді реттеуде шешуші рөл атқарады. Гендік экспрессия транскрипциядан басталады, бұл процесс ДНҚ-дан алынған ақпарат мРНҚ молекулаларына көшіріледі. Содан кейін мРНҚ трансляция процесінде ақуыздарды синтездеу үшін қолданылады немесе егер ол соматикалық емес жасуша болса, жасушада басқа функцияларды орындай алады. Бұл биологиялық процестерді зерттеуге, ауруларды анықтауға, диагностика мен емдеудің жаңа әдістерін жасауға және әртүрлі физиологиялық және патологиялық жағдайлардың негізінде жатқан механизмдерді түсінуге көмектеседі. Микрочиптердегі ген экспрессиясының профилі және жаңа буын секвенциясы қазіргі уақытта сүт безі қатерлі ісігінің пайда болуы мен таралуы негізінде жатқан молекулалық механизмдердің толық себебін ашу үшін пайдаланылатын геномдық деректердің үлкен көлемін алуға мүмкіндік береді. Бұл биоинформатика сүт безі обырын скринингтен алынған кешенді деректерді талдау мен түсіндіруде басым рөлге ие болады.



Сурет 1. Ген экспрессиясының схемасы

Қазіргі уақытта маммографиялық скрининг пен халықтың жалпы қартаюына байланысты сүт безі қатерлі ісігінің жиілігі артып келеді.

PR - прогестеронға арналған рецептор.

ER-эстрогенге рецептор.

Ki-67-онкогенді ақуыз.

HER-2-адамның эпидермиялық өсу факторының екінші типті рецепторы.

Ісіктің пролиферативті белсенділік деңгейі Ki-67 ақуызының экспрессиясымен бағаланады. HER2 ақуызының экспрессиясының жоғарылауы ісіктің агрессивті сипатын көрсетеді, емдеу тиімділігін төмендетеді және қайталану қаупін арттырады. Иммуногистохимиялық талдау негізінде сүт безі қатерлі ісігінің 4 кіші түрі ажыратылады, олардың әрқайсысының клиникалық көрінісі мен емдеу тактикасында өзіндік ерекшеліктері бар. Сүт безі қатерлі ісігінің люминальды түрлері HER-2 оң және үш есе теріс қатерлі ісіктерімен салыстырғанда аз агрессивті ағыммен және қолайлы болжаммен сипатталады. Жоғары агрессивтілікпен, метастаздардың ерте таралуымен және қолайсыз болжаммен сипатталатын үштік теріс қатерлі ісікті атап өту өте маңызды. Биоинформатика биология мен ақпараттық технологиялардың қиылысында ерекше орын алады, биологиялық зерттеулерден алынған деректерді сақтау, тарату және талдау құралдарын ұсынады. Ол ген экспрессиясының жиынтық дерекқоры (GEO) сияқты ашық дерекқорларда қол жетімді мРНҚ экспрессия профильдері сияқты ауқымды деректерді басқару қажеттілігінен туындады. Аппараттық және бағдарламалық қамтамасыз етуді дамыту осы деректерді талдау процесін айтарлықтай жақсартуға мүмкіндік берді [6]. Тар мағынада, биоинформатика - бұл биология мен информатиканың тоғысындағы пән, оның міндеті биологиялық ақпаратты тиімді басқару және талдау әдістерін әзірлеу болып табылады [7]. Бұл биологиялық деректерді талдауға қатысты есептерді шешудің алгоритмдері мен статистикалық әдістерін әзірлеуді қамтиды. 1980 жылдары биоинформатиканың пайда болуымен ғылымның бұл саласы тез дамыды, әсіресе 2000 жылдардың басында адам геномын картаға түсіру жобасы аяқталғаннан кейін [8]. ДНҚ микрочиптері және протеомдық эксперименттер сияқты ген экспрессиясын талдаудағы технологиялық жетістіктер биоинформатиканың маңыздылығын айтарлықтай арттырды [9]. Бүгінгі таңда биоинформатикалық әдістер биологиялық және клиникалық зерттеулерде эксперименттік тәсілдермен қатар қолданылады [10].

1.2.1 Биоинформатика құралдары арқылы биомаркерлер таңдау

Биомаркерлер-бұл организмдегі процестерді көрсететін және биологиялық үлгілерде өлшенетін немесе анықталатын негізгі көрсеткіштер. Олардың медицинадағы маңыздылығын бағаламауға болмайды, өйткені олар диагностикада, ауруларды бақылауда және жаңа емдеу әдістерін әзірлеуде шешуші рөл атқарады. Биомаркерлер гендер, РНҚ, пептидтер, ақуыздар,

метаболиттер және тіпті микроРНК сияқты әртүрлі сипаттамаларды көрсете алады. Биоақпараттық тәсілдер күрделі биологиялық деректерді талдауға және ықтимал маңызды гендерді, молекулаларды немесе сигнал беру жолдарын анықтауға мүмкіндік беретін сүт безі қатерлі ісігінің биомаркерлерін анықтауда маңызды рөл атқарады. Міне, осы саладағы бірнеше негізгі тәсілдер:

Ген экспрессиясын талдау: зерттеушілер ісік пен қалыпты сүт безі тіндері арасындағы ген экспрессиясының профильдерін салыстыру үшін ген экспрессиясы микрочиптерін немесе РНК секвенциясын пайдаланады. Бұл әлеуетті биомаркерлер ретінде қызмет ете алатын дифференциалды экспрессияланған гендерді анықтауға мүмкіндік береді.

МикроРНК талдауы: микроРНК-ген экспрессиясын реттейтін шағын РНК молекулалары. РМЖ-дегі микроРНК профильдерін зерттеу сүт безі қатерлі ісігімен байланысты және диагностика немесе болжау үшін биомаркер ретінде қызмет ете алатын микроРНК-ны анықтауға көмектеседі.

Биоақпараттық жолдарды талдау: KEGG немесе Reactome сияқты биологиялық жол дерекқорларын пайдалану зерттеушілерге сүт безі қатерлі ісігінде өзгерген сигналдық жолдарды анықтауға мүмкіндік береді. Бұл ісіктің дамуы мен прогрессиясының артындағы молекулалық механизмдерді түсінуге және емдеудің ықтимал мақсаттарын анықтауға көмектеседі.

Биомаркерлер мен қолтаңбаларды талдау: биоақпараттық әдістерді сүт безі қатерлі ісігінің диагностикасы, болжауы және бақылауы үшін пайдалануға болатын қолтаңбаларды немесе биомаркерлер жиынтығын жасау үшін пайдалануға болады. Бұған машиналық оқыту мен статистикалық талдаудың әртүрлі әдістері кіреді.

Интегративті деректерді талдау: геномика, протеомика, транскриптомика және клиникалық деректер сияқты әртүрлі көздерден алынған деректерді біріктіру сүт безі қатерлі ісігімен байланысты биологиялық процестер туралы толық түсінік алуға және жаңа биомаркерлерді анықтауға мүмкіндік береді.

Биоинформатика тәсілдері сүт безі қатерлі ісігінің биомаркерлерін анықтау үшін өте маңызды, бұл айтарлықтай клиникалық нәтижелерге әкелуі мүмкін. CBioPortal, RCGAToolbox Және Genomic Data Commons деректер порталы сияқты құралдар генотиптерді фенотиптермен байланыстыруға көмектеседі, зерттеу нәтижелерін нақты клиникалық қолданбаларға аударады. Машиналық оқыту және тереңдетіп оқыту omics деректерін талдау арқылы болжамды, диагностикалық және болжамды биомаркерлерді анықтау үшін де қолданылады². Биоинформатика талдауын эксперименттік тексерумен

біріктіретін интегративті тәсіл сүт безі қатерлі ісігінің перспективалы биомаркерлерін ашудың кілті болып табылады.

CA 15-3: бұл ісік маркерін пациенттің сүт безі қатерлі ісігін емдеуге реакциясын бағалау үшін қолдануға болады. CA 15-3 деңгейінің жоғарылауы аурудың дамуын көрсетуі мүмкін.

CA 125: сүт безі зақымданған кезде анықталуы мүмкін басқа ісік биомаркері. Ол басқа қатерлі ісіктермен де байланысты болуы мүмкін, бірақ оның деңгейі науқастың жағдайын бақылау үшін пайдалы болуы мүмкін.

MammaPrint: бұл тест сүт безі қатерлі ісігінің қайталану және алыс метастаз қаупін анықтайды.

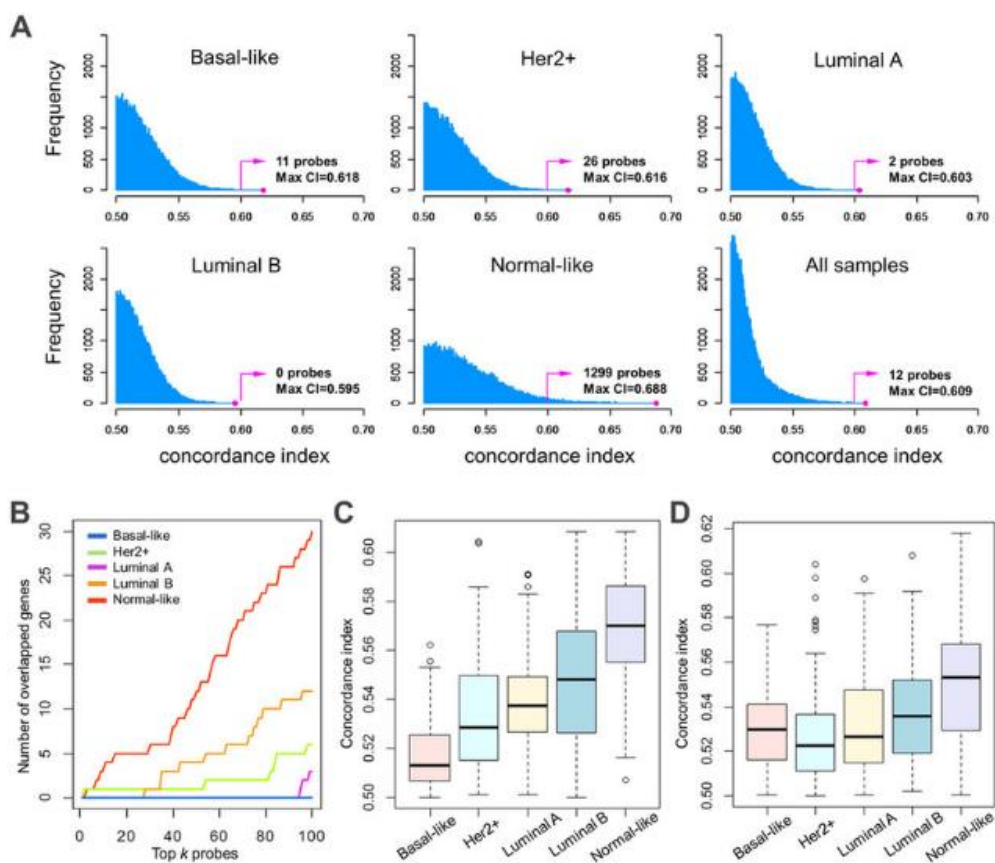
Биоинформатика саласындағы зерттеулер сүт безі қатерлі ісігінің бірқатар ықтимал емдік мақсаттарын, соның ішінде DLGAP5, AURKA және UBE2C сияқты гендерді анықтады [11]. Осылайша, биоинформатика күрделі биологиялық деректерді талдау және түсіндіру үшін құралдар мен әдістерді қамтамасыз ететін заманауи биологиялық зерттеулердің маңызды саласы болып табылады. Біріктірілген геномика арқылы сүт безі қатерлі ісігіндегі он маңызды генді анықтады және сүт безі обырын емдеуге арналған 65 ықтимал шағын молекулалы қосылыстарды болжады [12]. Биоинформатика геномдық деңгейде ауруларды зерттеуге жақындауға және ауру механизмдері туралы жаңа гипотезаларды тұжырымдауға мүмкіндік беретін ген экспрессиясының толық профильдерін талдау үшін қолданылады. Бұл ісік жасушаларының түзілуіне, сақталуына және таралуына қатысатын сигнал беру жолдарын зерттеуді қамтиды. Сонымен қатар, биоинформатикалық талдау микроРНК, ДНК көшірмелерінің саны туралы деректер, бір нуклеотидті полиморфизмдер (SNPS), ДНК тізбегі және ДНК метилденуі туралы ақпаратты қоса алғанда, деректердің кең ауқымына қолданылады. Соңғы онжылдықта сүт безі қатерлі ісігінің молекулалық құрылымы мен ісік гетерогенділігін түсінуде революциялық секіріс болды. Доминантты генетикалық факторлардың аз саны ісіктің дамуына және көшірме санының өзгеруіне ықпал ететін сирек мутациялармен синергетикалық түрде әрекет етеді деген тұжырым бар. Осы уақытқа дейін 1000 сүт безі ісіктерінің үлкен параллель тізбегін ескере отырып, қозғаушы гендердің көпшілігі анықталған. TNRMF кезінде қайталанатын мутациялар және мақсатты жолдардың айқын жетіспеушілігі байқалады [13]. Хромосомалық тұрақсыздықтан туындаған бұл қатерлі ісік үшін геномдық тұрақсыздық механизмдеріне әсер ету қажет. Сүт безі қатерлі ісігінің әртүрлі кіші түрлеріндегі көптеген ауытқулардың функционалдық маңыздылығы мен гендік тәуелділігі әлі анықталған жоқ. Бұл негізгі молекулалық жолдар мен өзара байланысты түйіндерді шешу үшін жан-жақты функционалды скринингті (мысалы, CRISPR-Cas9 өңдеу) тереңірек талдаумен және деректерді біріктірумен біріктіруді қамтиды [14]. Қазіргі уақытта сүт безі

қатерлі ісігі туралы генетикалық ақпарат өте бай және терең болғанымен, оны дәл медицинаға және әдеттегі клиникалық тәжірибеге көшіру қиын міндет болып қала береді. Дегенмен, геномдық және экспрессивті белгілерді клиникалық және патологиялық белгілермен үнемі нақтылау және біріктіру сүт безі қатерлі ісігінің негізгі механизмдері туралы түсінігімізді айтарлықтай кеңейтіп, сайып келгенде, емдеудің прогрессиясы мен деэскалациясын басқаруға арналған биомаркер құралдарының жетілдірілуіне әкелуі керек. Жекелендірілген емдеудің жаңа тәсілдері мультимиксті машиналық оқытуды қолданудан да пайда көруі керек және соңғы зерттеулер емдеуге жауаптарды болжайды. Бұл талдау тұқым қуалайтын ауруларды тудыруы мүмкін геномдық өзгерістерді анықтау, аурулардың дамуына жауап беретін гендерді анықтау және аурудың фенотипін қалпына келтіруге бағытталған генетикалық терапияны әзірлеу үшін медицинада қолданылады. Кең деректер жиынтығын өңдеу және талдау үшін арнайы мәліметтер базасы мен шолғыш құралдары құрылды. Осылайша, биоинформатика аурулардың молекулалық негіздерін терең түсіну және диагностика мен емдеудің жаңа тәсілдерін әзірлеу құралдары мен әдістемелерін ұсына отырып, қазіргі заманғы ауруларды зерттеуде шешуші рөл атқарады.

1.3 Сүт безі қатерлі ісігіндегі ген экспрессиясының классификациясы

Биоинформатика генетикалық ақауларды анықтау үшін үш миллиард базалық жұпты іздеу қолға алынды. Сүт безінің қатерлі ісігі енді бір ғана ауру ретінде қарастырылмайды, керісінше, ол гетерогенді, молекулалық, гистопатологиялық және клиникалық деңгейде әртүрлі кіші түрлерден тұрады, әртүрлі болжамдық және емдік әсерлері бар. Алинен және оның әріптестері қалыпты сүт безі тінін құрайтын жасушалардың әрбір түріндегі ген экспрессиясы мен SAGE (ген экспрессиясының үздіксіз талдауы) жан-жақты профилін жүргізді және жасуша бетінің маркерлері мен магниттік моншақтарды пайдаланатын *in situ* және инвазивті сүт безі карциномаларын құрайтын әрбір жасуша түріндегі ген экспрессиясының жан-жақты профильдерін сипаттады [15]. Адам ісіктерінің гетерогенділігін түсіну және жеке геномиканы қолдана отырып ісіктерді диагностикалау мен емдеудің жаңа стратегияларын жасау үшін осы аурумен байланысты механизмдерді нақтылауды қажет етеді. Сүт безі обырын ерте анықтау пациенттердің өмір сүруіне айтарлықтай әсер ететіні кеңінен мойындалған. Ісік геномының экспрессия үлгілерінің ісік биологиясын қалай көрсететінін бақылау өте маңызды, өйткені салыстырмалы геномдық будандастыру және транскрипциялық профильдеу сияқты жоғары өнімді геномдық массивтерге негізделген әдістер ген экспрессиясының үлгілерін клиникалық нәтижелермен байланыстыру арқылы ісік прогрессиясы кезінде арнайы белсендірілген немесе инактивацияланған гендер мен жолдарды анықтау үшін пайдалануға

болады. Ген экспрессиясының профилі сүт безі қатерлі ісігін әртүрлі кіші топтарға жаңа жолмен жіктеуге мүмкіндік берді. Ген экспрессиясының профилі сүт безі обырын бес биологиялық әртүрлі ішкі кіші түрге бөлуге мүмкіндік берді: люминальды А, люминальды В, HER2-оң (HER21), базальды және қалыпты. А және В люминальды кіші типтері ER-оң, ал В люминальды кіші типтері салыстырмалы түрде нашар нәтижемен байланысты. HER21-және базальды сүт безі қатерлі ісігі нашар нәтиже береді. Ісік үлгілеріндегі ген экспрессиясын талдау әдетте тұтас ісік гомогенаттарында жүргізіледі және осылайша ісіктегі барлық жасуша түрлерінің экспрессиясын көрсететін үлгі болып табылады. Басқа зерттеуде гендік онтологиялық биологиялық процесте анықталған гендер жиынтығын картаға түсіру арқылы оң эстроген рецепторлары (ER+) немесе теріс G ақуызы ісіктің метастаздық потенциалымен байланысты ER+ немесе ER-мен байланысты болды [16].



Сурет 2. Сүт безі қатерлі ісігінің бес кіші түрінің гендік экспрессия профилдері әртүрлі болжамдық күшті көрсетеді.

Люминальды А: сүт безі қатерлі ісігінің бұл кіші түрі эстрогендік рецепторларға (ER+) және мүмкін прогестерон рецепторларына (PR±) оң реакциямен сипатталады, HER2 теріс (HER-2 теріс). Сонымен қатар, Ki-67 пролиферация индексі әдетте төмен (<16%). Бұл кіші түр әдетте қолайлы болжамға ие және гормондық терапияға жақсы жауап береді. Ісік дамыған кезде пролиферативті белсенділік 19% - дан аз болады. Дұрыс тағайындалған емдеу кезінде асқинулар мен қайталану қаупі барынша азайтылады.

Люминальды В: бұл кіші типте эстрогендік рецепторларға (ER+) және/немесе прогестерон рецепторларына (PR=) реакция бар, бірақ Ki-67 пролиферация индексі жоғары (>16%). HER2 теріс (her-2 теріс) немесе оң (HER-2 оң) болуы мүмкін. Люминальды в әдетте люминальды А-ға қарағанда қатерлі ісіктің агрессивті ағымымен байланысты.

HER2-оң. HER2-жасушалардың өсуі мен бөлінуін реттеуге қатысатын ақуыз. HER2-оң сүт безі қатерлі ісігінде бұл ақуыздың шамадан тыс реттелуі орын алады, бұл қатерлі ісік жасушаларының өсуі мен көбею сигналдарының жоғарылауына әкеледі. Бұл әдетте қатерлі ісіктің агрессивті түрімен және оның таралу жылдамдығымен байланысты. HER2-оң қатерлі ісік (HER2-positive breast cancer) сүт безі қатерлі ісігінің кіші тобы рак клеткаларының бетінде HER2 рецепторының (2 типті эпидермиялық өсу факторының рецепторы) жоғарылауымен сипатталады.

Сүт безі қатерлі ісігінің базальды кіші түрі, сонымен қатар үштік-теріс сүт безі қатерлі ісігі. Сүт безінің базальды қатерлі ісігі әдетте эстроген (ER-), прогестерон (PR-) және HER2 (HER2 -) рецепторларына теріс әсер етеді. Бұл қатерлі ісіктің бұл түрі гормондық терапияға немесе HER2 тежеуге бағытталған дәрілерге жауап бермейтінін білдіреді. Базальды сүт безі қатерлі ісігі көбінесе агрессивті ағыммен және басқа кіші түрлермен салыстырғанда ісіктің тез өсуімен сипатталады. Бұл метастаздың жоғары қаупімен және аз қолайлы болжаммен байланысты болуы мүмкін.

1.4 Биоинформатикалық құралдармен сүт безі қатерлі ісігін диагностикалау

Сүт безі қатерлі ісігімен ауыратын науқастарды емдеудің әртүрлі әдістері бар, соның ішінде ең көп қолданылатындар: хирургия, сәулелік терапия, эндокриндік терапия, химиотерапия және мақсатты терапия. Алайда, сүт безі қатерлі ісігімен ауыратын науқастардың 40% ісіктердің қайталануы байқалады, ал олардың 60-70% алыстағы метастаздар байқалуда. [17] Биоинформатиканың дамуымен, одан да басқа геномика мен транскриптомика сияқты үлкен көлемдегі деректердің пайда болуымен көптеген биомолекулалар арасындағы қалыптасқан байланысты зерттеу мақсатында осы деректерді талдау үшін биоинформатика мен информатика әдістерін қолдану, аурудың механизмін нақтылау және терапевтік мақсаттарды болжау үшін маңызды зерттеу әдісі болды. Сүт безі қатерлі ісігін диагностикалаудың дәстүрлі әдістерімен салыстырғанда, молекулалық маркерлер жоғары ерекшелік, динамикалық бақылаудың қарапайымдылығы және т.б. сияқты ерекше артықшылықтарға ие болуда. [18]

Көбінесе, DAVID онлайн талдау веб-сайты арқылы дифференциалды экспрессия гендерін (DEG) және байытылған GO және KEGG

дифференциалды гендерін анықтау үшін сүт безі қатерлі ісігінің гендік экспрессиясының (GEO) жан-жақты дерекқорлары қолданылады. Дерекқорлар дифференциалды гендердің қызметін және олардың аурумен байланысты сигнал беру жолдарындағы рөлін талдау сүт безі қатерлі ісігінің жаңа диагностикалық биомаркерлерін анықтау, сондай-ақ бастапқы реттеуші бұзылулар мен сүт безі қатерлі ісігінің негізгі механизмдерін анықтау бойынша идеяларды ұсынады. Ықтималды сүт безі обырын бағалау үшін қажетті сынақтардың саны әдетте пациент немесе оның дәрігері маммографияға скрининг жасау, қалыптан тыс кальцинацияға зертханалық зерттеулер жүргізу және өзін-өзі тексеру кезінде кеудеде ісік немесе түйін тапқан кезден басталады. Көбінесе адам кеудесінің қызарғанын немесе ісінгенін, сондай-ақ қолтық астындағы түзілімдерді немесе түйіндерді байқауы мүмкін.

Ұлттық Биологиялық Ғылымдар Орталығының (GEO) дерекқоры ашық қолжетімді деректер платформасы болып табылады. Бұл микрочиптерді және жоғары өнімді реттілікті пайдалана отырып, ген экспрессиясын профильдеуге арналған тарихтағы ең үлкен деректер жинағын қамтиды. Gse29431 деректер жинағын алу үшін GEO дерекқорына жарияланған сүт безі қатерлі ісігінің гендік чип деректер жинағын енгізу үшін "Сүт безі обырын" іздеу сұрауы ретінде пайдаланылады. Деректер жинағы U133 plus 2.0 Адам Геномының Үлкен Аффиметриялық платформасын пайдаланады және сүт безі қатерлі ісігінің 54 жағдайын және сүт безі қатерлі ісігі бар 12 науқаста қалыпты сүт безі тінін қамтиды. Сүт безі қатерлі ісігінің 54 үлгісінің түрлері 1-кестеде келтірілген.

Салауатты бақылау	12 жағдай	
Сүт безі қатерлі ісігі	Люминальды А	11 жағдай
	Люминальды В	28 жағдай
	HER2 шамадан тыс экспрессиясы	13 жағдай
	Базальды тәрізді	2 жағдай

1-кесте. Сүт безі обырын талдауға арналған үлгілердің түрлері

TSCGA онкологиялық науқастардың түсінігін тереңдету және қатерлі ісіктің алдын алу, диагностикалау және емдеу әдістерін жақсарту үшін жоғары өнімді геномдық талдау технологиясын пайдаланады. GTEx тіндерге тән ген экспрессиясын және реттелетін дерекқорларды жасау үшін адамның көптеген тіндері мен мүшелерінен алынған үлгілердің транскрипциялық реттілігі мен генотиптік талдауын әзірледі. Тексеру үшін сүт безі қатерлі ісігінің TSCGA және GTEx деректерін және қалыпты сүт безі үлгілерін жүктеп алу керек.

Мәселен, қытайлық ғалымдар Юэхун Ку мен Чанчун Ней өз мақалаларында биоинформатикалық талдауды қолдана отырып, сүт безі

қатерлі ісігінің диагностикалық және болжамды биомаркерлеріне скрининг жүргізді. Нәтижесінде, сүт безі обыры бар емделушілерден алынған қатерлі ісік үлгілерін және қалыпты сүт безі үлгілерін дифференциалды талдау арқылы барлығы 965 дифференциалды ген алынды, олардың 833 - і төмендеді және 132-сі көтерілді. Дифференциалды экспрессияланған гендер ангиогенез, плазмалық мембрана және интегриннің байланысуы, PPAR сигнал беру жолдары, адипоциттердегі липолиздің реттелуі және глицеролипидтердің метаболизмі сияқты әртүрлі GO ішкі жиындары үшін байытылды. Ең маңызды айырмашылықтары бар бес ген CA4, PLIN4, GPD1, TUSC5 және S100B болды және сүт безі қатерлі ісігінің тіндеріндегі осы бес геннің экспрессия деңгейі қалыпты сүт безі тіндеріне қарағанда төмен болды. Ең айқын дифференциалды экспрессиясы бар бес геннің *genia* онлайн талдауында біз S100B генінің пациенттердің болжамымен айтарлықтай байланысы бар екенін анықтадық. S100B генінің экспрессиясы неғұрлым жоғары болса, пациент үшін болжам соғұрлым жақсы болады [19]. Дегенмен, CA4, PLIN4, GPD1 және TUSC5 гендерінің экспрессия деңгейлері пациенттердің болжамымен айтарлықтай байланысты емес. TCGA және GTEx дерекқорларындағы осы бес геннің экспрессия деңгейі сүт безі қатерлі ісігінде төмендеді және бұл статистикалық маңызды болды. Осы зерттеудегі ген экспрессиясының кең дерекқорынан алынған микрочиптер арқылы сүт безі қатерлі ісігінің деректерін талдау арқылы анықталған ең айтарлықтай ерекшеленетін бес CA4, PLIN4, GPD1, TUSC5 және S100B гендері сүт безі қатерлі ісігінің тіндерінде басылды, бұл қатерлі ісіктің ықтимал диагностикалық биомаркері болуы мүмкін сүт безі. Сонымен қатар, дифференциалды экспрессияланған 100B гендері сүт безі қатерлі ісігін болжау үшін әлеуетті биомаркер болуы мүмкін.

Бүгінгі таңда жаңа дамыған технологиялардың арқасында қатерлі ісік ауруының ықтималдығын білу үшін диагностика мен тест жүргізудің көптеген жолдары бар. 2016 жылы ASCO геномдық сынақтарды қолдануды мақұлдады, олар:

Dx™ онкотипі. Бұл сынақ ER-оң немесе PR-оң, HER2-теріс сүт безі қатерлі ісігі бар, лимфа түйіндеріне таралмаған адамдарға арналған. Бұл сондай-ақ кейбір жағдайларда қатерлі ісік 1-3 лимфа түйіндеріне тараған, мысалы, менопаузадан аман қалған әйелдерде мүмкін. Бұл тест пациенттер мен олардың дәрігерлеріне гормондық терапияға химиотерапияны қосу туралы шешім қабылдауға көмектеседі. Бұл тестте қатерлі ісікпен байланысты 16 ген және «рецидив көрсеткішін» есептеу үшін 5 эталондық ген қарастырылады, егер науқас 5 жыл бойы гормондық терапияны қабылдаған жағдайда, диагноз қойылғаннан кейін 10 жыл бойы сүт безінен немесе өңірлік лимфа түйіндерінен тыс қатерді бағалайды. Қайталану көрсеткіші химиотерапияны қолдану жөніндегі ұсынымдарды басшылыққа алу үшін пайдаланылады. Мысалы, Американдық клиникалық онкология қоғамы

(ASCO) қайталану деңгейі 26 немесе одан жоғары рецидив көрсеткіші бар адамдарға кейіннен гормондық терапия мен химиотерапия ұсынуды кеңес береді.

МаммаПринт™. Бұл сынақ 3 немесе одан аз лимфа түйіндеріне таралған сүт безінің ER-оң немесе PR-оң, HER2-теріс обыры бар 50 жастан асқан адамдарға арналған. Бұл сынақ сүт безі қатерлі ісігінің ерте сатыдағы қайталану қаупін бағалау үшін 70 ген туралы ақпаратты пайдаланады. Қатерлі ісіктің қайталану қаупі жоғары адамдар үшін бұл сынақ пациенттер мен олардың дәрігерлеріне гормондық терапия алдында химиотерапия жүргізу керектігі туралы шешім қабылдауға көмектеседі. Бұл тест қатерлі ісіктің қайталану қаупі төмен адамдарға, 50 жастан кіші адамдарға немесе 3 лимфа түйінінен артық қатерлі ісігі бар адамдарға ұсынылмайды. [20]

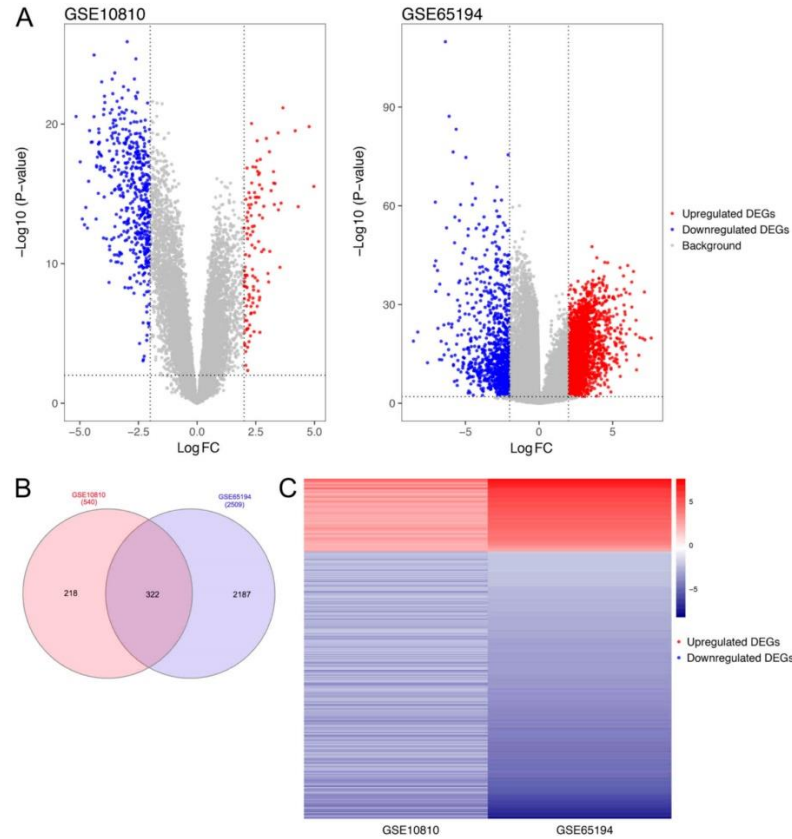
ЭндоБолжам. Бұл сынақ 3 немесе одан аз лимфа түйіндеріне таралған және менопауза пайда болған сүт безінің ER-оң, HER2-теріс обыры бар адамдарға арналған. Бұл тест диагноз қойылғаннан кейін 10 жыл ішінде қатерлі ісіктің қайта даму қаупін бағалау үшін 12 гендегі ақпаратты пайдаланады. Бұл тест пациенттер мен олардың дәрігерлеріне химиотерапияны және операциядан кейін гормондық терапияны жүргізу туралы шешім қабылдауға көмектеседі.

Prosigna™ (PAM50). Бұл тест менопаузадан аман қалған және лимфа түйіндеріне таралмаған ER-оң, HER2-теріс сүт безі обыры бар адамдарға арналған. Бұл тест диагноз қойылғаннан кейін 10 жыл ішінде қатерлі ісіктің қайта даму қаупін бағалау үшін 50 ген туралы ақпаратты пайдаланады. Бұл тест пациенттер мен олардың дәрігерлеріне химиотерапияны және операциядан кейін гормондық терапияны жүргізу туралы шешім қабылдауға көмектеседі. [21]

1.5 Сүт безіндегі машиналық оқыту тәсілдері

RNA-Seq транскрипт деңгейіндегі экспрессияны талдау үшін ақпарат береді және микрочип технологияларымен салыстырғанда айқас будандастыруға және өлшенетін экспрессия деңгейлерінің шектеулі диапазонына байланысты шектеулерді жеңеді. Сүт безі қатерлі ісігінде гендердің күшеюі кейбір хромосомалық жерлерде қайталанатын түрде жүреді (мысалы 1q, 8p12, 8q24, 11q13, 12p13, 12q13, 17q21-q23, 20q13), бұл ісіктің өсуі кезінде кейбір онкогендердің жоғары жиілікте белсендірілуін көрсетеді [22]. Ісіктің өсуі кезінде жоғары жиілікті күшейту геннің экспрессиясын физиологиялық тұрғыдан тыс тудырады, вариацияның қалыпты деңгейлері конституциялық күшейту механизмдері болып табылады, сондықтан ісіктің дамуы үшін маңызды. Қатерлі ісік жасушаларында онкогенді күшейтудің маңыздылығы канцерогендік массивтерді қолдану арқылы ісік жасушаларының

экспрессиялық профилін анықтау нәтижесінде пайда болды. Аномальды ген экспрессиясы сүт безі қатерлі ісігінің пайда болуы мен дамуында маңызды рөл атқарады. Соңғы жылдары сүт безі қатерлі ісігін емдеу әдістерін жетілдіре отырып, сүт безі қатерлі ісігімен ауыратын науқастардың болжамы мен өмір сапасы жақсарды.



Сурет 3. Әрбір деректер жиынындағы дифференциалды экспрессияланған ген және екі ортақ гендік экспрессияның жиынтық дерекқоры жиынымен ортақ ортақ ген экспрессиясы.

(А) Әрбір гендік экспрессияның жиынтық дерекқорындағы деректер жиынындағы дифференциалды экспрессияланған гендер графигі. Қызыл нүктелер ісік үлгілерінде айтарлықтай үлкейген гендерді білдіреді. Көк нүктелер ісік үлгілерінде айтарлықтай төмендеген гендерді білдіреді. Нүктелі тік сызықтар сүзгінің маңыздылық шегін көрсетеді.

(В) Екі деректер жиынтығымен ортақ гендік экспрессияның жиынтық дерекқоры.

(С) бірдей ген экспрессиясының үлгісі бар екі деректер жиынындағы жалпы DEG ген экспрессиясының жылу картасы. Қызыл сызықтар ісік үлгілерінде айтарлықтай үлкейген гендерді білдіреді. Көк сызықтар ісік үлгілерінде айтарлықтай төмендеген гендерді білдіреді. DEG, экспрессиясы өзгертілген гендер; GEO, ген экспрессиясы туралы геоақпараттық мәліметтер базасы; FC, өзгеру коэффициенті.

Маңызды емдеу әдістері, әсіресе иммундық бақылау нүктелерін бұғаттау терапиясы жатады. Иммунотерапиямен байланысты мақсаттар ісік иммунотерапиясының дамуына ықпал ететін негіз болып табылады. FAM83H-AS1 аномальды геніне дейін биоинформатикалық талдау арқылы сүт безі қатерлі ісігінің кестесін тапты. FAM83H-AS1 әр түрлі аномальды экспрессия бауыр қатерлі ісігі, сүт безі қатерлі ісігі, жатыр мойны обыры, қуық қатерлі ісігі, өңеш қатерлі ісігі, өкпе ісігі гендерінің сипаттамасы ретінде көрінеді [23]. Әдебиеттегі зерттеулер протеинкиназа R (PKR) және фосфатидилинозитол-4-киназа 2-альфа (PI4K2A) гендерінің экспрессиясы сүт безі қатерлі ісігімен ауыратын науқастарды болжау үшін маңызды екенін көрсетті, ал PI4K2A/PKR бағытталған лизосомалық кешен HER3-оң сүт безі қатерлі ісігі жасушаларының пролиферациясын тежеудің маңызды механизмі болуы мүмкін [24]. PI4K2A/PKR-ге бағытталған лизосомалық кешен қатерлі ісіктің осы түрін емдеудің тиімді әдісі бола алады. Зерттеуде [25] Fattyacidtransportprotein1 (FATP1) сүт безі қатерлі ісігі жасушалары мен қатерлі ісік емес жасушалар арасындағы май қышқылдарын тасымалдауда шешуші рөл атқаратыны анықталды, ал FATP1 белсенділігінің тежелуі ісік жасушаларының май қышқылдарын алу қабілетін тежейді және олардың өміршеңдігінің төмендеуіне әкеледі, сондықтан FATP1 сүт безі қатерлі ісігін емдеудің мақсаты болуы мүмкін. Осылайша, FATP1 сүт безі қатерлі ісігін емдеудің мақсаты болуы мүмкін. Авторлар қабаттасатын гендердің көпшілігі сүт безінің канцерогенезінде маңызды рөл атқаратын фокальды адгезияға және жасушадан тыс матрицаның рецептормен әрекеттесуіне қатысты екенін анықтады. Әртүрлі зерттеулер арқылы ген қолтаңбасы туралы деректер неғұрлым көп болса, ген экспрессиясының эпигенетикалық реттелуін түсіну және түзету араласуы соғұрлым жақсы болады.

1.6 Сүт безі қатерлі ісігін емдеу әдістері

Сүт безі қатерлі ісігі қазіргі заманғы медицинаның қатерлі ісік ауруын зерттеудегі үлкен жетістіктеріне қарамастан, бүкіл әлемдегі әйелдердің денсаулығына негізгі қауіптердің бірі болып қала береді [26]. Зерттеулер көрсеткендей, сүт безі қатерлі ісігі 45 жасқа дейінгі әйелдерде қатерлі ісік өлімінің негізгі себебі болып табылады [27]. Сүт безі обырын емдеудің дәстүрлі әдістеріне хирургиялық кесу, радиотерапия, химиотерапия, гормондық терапия және молекулалық мақсатты терапия жатады [28]. Сүт безі обырын ерте анықтаған кезде радио және химиотерапиямен біріктірілген хирургиялық кесу пациенттердің өмір сүруін айтарлықтай арттыруы мүмкін [29]. Алайда, аурудың алғашқы белгілері айқын емес және оларды елемей оңай болғандықтан, пациенттердің көпшілігі анықталғанға дейін кеш сатысында, ал рак клеткалары басқа органдарға метастаз берген, сондықтан хирургия, химиотерапия және гормондық терапияның болжамы қолайсыз және молекулалық мақсатты терапия, олардың мақсатты және

цитотоксикалығының арқасында сүт безі обырын емдеу саласында зерттеудің ыстық нүктесіне айналды. Молекулалық мақсатты терапияның төмен цитоуыттылығы, жанама әсерлерінің аздығы және қайталану жиілігінің төмендігі сияқты артықшылықтары болса да, сүт безі қатерлі ісігінің патогенезі күрделі және аурудың әртүрлі молекулалық кіші түрлерінің мақсаттары бірдей емес, сондықтан қолайлы мақсатты анықтау ауруды мақсатты емдеудің маңызды шарты болып табылады [30]. Бұл ауруды емдеуге қолайлы мақсаттарды табу үшін ғалымдар көптеген зерттеулер жүргізді. In vitro эксперименттер жоғары спецификалық цитоуыттылыққа бағытталған шағын интерференциялық РНҚ (siRNA) адамның эпидермиялық өсу факторының (HER3) 3 рецепторымен байланысатынын және HER3-оң сүт безі қатерлі ісігі жасушаларының көбеюін тиімді тежейтінін көрсетті [31]. Сүт безі обырын емдеуді екі компонентке бөлуге болады - бұл жергілікті немесе жергілікті, хирургиялық және сәулелік терапияны қамтитын экспозиция және химиотерапия және эндокриндік әсерлер арқылы жүзеге асырылатын жүйелік.

Хирургиялық емдеу:

1. Үлкен кеуде бұлшықетін сақтай отырып - радикалды мастэктомия. Бір блок субклавиялық, аксиларлы, субкапулярлы аймақтардың және кіші кеуде бұлшықетінің талшықтары мен лимфа түйіндері бар сүт безін алып тастайды.
2. Қолтық асты лимфа түйіндерін алып тастайтын - мастэктомия. Бір блок қолтық асты және вектораралық аймақтың лимфа түйіндерін сүт безімен алып тастайды. Екі кеуде бұлшықеті де сақталады.
3. Мастэктомия- үлкен кеуде бұлшықетінің фассиясымен сүт безін алып тастау. Операцияның бұл түрі паллиативті ретінде қолданылады.
4. Сүт безінің радикалды резекциясы - субклавиялық, аксиларлы және субкапулярлы аймақтардың лимфа түйіндері бар бір блокта сүт безі секторын алып тастау. Ісіктің ішкі локализациясында операция екі тіліктен жасалады - аксиларлы аймақтағы екінші кесу.

Сәулелік терапия бастапқы сүт безі обырын емдеудің дербес әдісі ретінде салыстырмалы түрде сирек қолданылады. Егер операция жасалуы мүмкін болса, сәулелік әсер хирургиялық араласумен бірге орындалады. Операцияға дейінгі сәулелік терапияны, операциядан кейінгі немесе екеуінің комбинациясын қалған рак клеткаларын жою үшін қолданылады. Сәулелік терапия өздігінен немесе дәрілік терапиямен бірге ісікті операциялық күйге келтіру мақсатында сүт безі қатерлі ісігінің жергілікті кең таралған түрлерін емдеудің бірінші кезеңінде маңызды болып қала береді. Операцияға қарсы көрсетілімдер болған кезде сәулелік терапия консервативті емдеу тұрғысынан қолданылады.

Гормондық терапия сүт безі қатерлі ісігінің бастапқы кең таралған және жалпыланған түрлерінде жүргізілетін кешенді емдеудің міндетті құрамдас

бөлігі болып табылады. Ол рак клеткаларында гормоналды рецепторлар болған кезде қолданылады. Рак клеткаларының өсуін ынталандыратын гормондардың әсерін блоктауға бағытталған. Гормондық терапияның жетістіктері антиэстрогендер, ароматаза ингибиторлары мен анактиваторлары, прогестиндер сияқты жүйелі әсер ететін препараттарды қолданумен байланысты.

Химиотерапия - бүкіл денеде рак клеткаларын жою үшін қолданылады. Ісікті азайту үшін операциядан бұрын (неoadьювантты химиотерапия) немесе қалған рак клеткаларын жою үшін операциядан кейін (адьювантты химиотерапия) қолдануға болады. Сүт безі қатерлі ісігінде жиі қолданылатын химиотерапиялық препараттардың ішінде әсер ету механизмі өте әртүрлі, бірақ көбейетін ісік жасушасын өлтіретін: таксотера, паклитаксел, доксорубицин-адриамицин, фармарубицин, навельбин, кселода, 5-фторурацил, циклофосфамид. Бұл препараттар әртүрлі комбинацияда және дәйектілікте сүт безі қатерлі ісігінің адьювантты терапиясында сәтті қолданылады. Оксфорд университетінде жүргізілген халықаралық талдау сүт безі қатерлі ісігінің қайталану қаупін 50% -70% - ға азайтатын және сол арқылы сүт безі қатерлі ісігімен ауыратын ондаған мың әйелдердің өмірін сақтайтын адьювантты химиотерапияның жоғары тиімділігін растады.

1.7 Биоинформатикалық талдау арқылы сүт безі қатерлі ісігіндегі ықтимал маңызды гендер

Сүт безі қатерлі ісігін зерттеуде биоинформатика оның дамуы мен прогрессиясында рөл атқаратын гендер мен биологиялық жолдарды анықтаудың негізгі құралы ретінде әрекет етеді. Пациенттердің геномдық деректерін және гендік экспрессия деректерін талдау жаңа терапевтік стратегиялар үшін маңызды болуы мүмкін молекулалық мақсаттарды анықтауға мүмкіндік береді. Сонымен қатар, биоинформатикалық әдістер ER-оң немесе ER-теріс қатерлі ісік сияқты сүт безі қатерлі ісігінің әртүрлі фенотиптерінің механизмдерін түсінуге көмектесуі арқылы емдеуді таңдауда маңызды болуы мүмкін. Көптеген қатерлі ісіктер сияқты, сүт безі қатерлі ісігі молекулалық сипаттамаларының, гистопатологиялық түрінің және клиникалық нәтижесінің айырмашылығына байланысты әртүрлі түрлерге жіктеледі. Базаль тәрізді және клаудин-төмен эстроген рецепторларының (ER), прогестерон рецепторларының (PR) және HER2 экспрессиясының болмауымен сипатталатын кіші типтер үш рет теріс сүт безі қатерлі ісігінің (TNRMF) түріне жатады және аурудың ұзақ мерзімді қайталану ықтималдығы жоғары және висцеральды метастаздардың жоғары жиілігі бар [32]. Жақында TNRMF жағдайларының үлкен когортасы бар мета-талдау TNRMJ-ді кем дегенде төрт кіші түрге бөлді: базальды иммуноактивті (BLIA), базальды иммуносупрессияланған (BLIS), люминальды андрогендік рецептор (LAR) және мезенхималық (MES) ісік [33]. Бұл кіші классификацияны мРНҚ, микроРНҚ, ДНҚ және эпигенетикалық профильдерді талдау арқылы қатерлі

ісік геномының атласы (TCGA) бағдарламасы одан әрі қолдайды [34]. Молекулалық биологиялық зерттеулердің деректеріне сүйене отырып, сүт безі қатерлі ісігінің пайда болуы қалыпты эпителийден *in situ* түтік карциномасына (DCIS) дейін дамып, содан кейін инвазивті сүт безі қатерлі ісігіне (IBC) және метастаздық карциномаға дейін жетеді [35]. *In situ* түтік карциномасы инвазивті сүт безі қатерлі ісікпен тікелей байланыста болып табылды және оның өзіне тән патологиялық өзгерістері бар [36]. Сонымен қатар, DCIS локализацияланған сәулелік терапия мен эндокриндік терапиядан кейін қайталанудың жоғары жылдамдығына ие. Дегенмен, сүт безі қатерлі ісігін молекулалық патологиясы туралы түсінігіміз әлі де шектеулі. Сүт безі қатерлі ісігін клиникалық емдеу нұсқалары туралы көптеген пікірталастар бар, дегенмен жақында DCIS емдеу стратегияларын стандарттау жүргізілді. Осылайша, сүт безі қатерлі ісігінің дамуы молекулалық механизмдерін және сүт безі қатерлі ісігінің ерте сатысын жіктеуді зерттеу үшін көбірек зерттеулер жүргізу қажет. Айта кету керек, жеті ген ASPN, DNAPTP3, FN1, C2orf40, CNN1, CRYAB және KRT5 сияқты барлық кезеңдерде дәйекті өзгерістерге ие, бұл осы гендердің сүт безі қатерлі ісігінің дамуындағы "жетекші" рөлін көрсетеді. Атап айтқанда, DNAPTP3-ті анықталды, ол адамның сүт безі қатерлі ісігінің дамуының ерте және кеш кезеңдерінде реттелетін жаңа ген. Биоинформатикалық талдау арқылы сүт безі қатерлі ісігіндегі ықтимал маңызды гендер мен негізгі жолдарды анықтау гендер мен ақуыздардың өзара әрекеттесу деректерін талдаудың түрлі әдістерін қамтитын кешенді тәсілді қолдануды мәңзейді. Бұл тәсіл сүт безі қатерлі ісігінің дамуы мен дамуында шешуші рөл атқаратын гендер мен жолдарды анықтауға мүмкіндік береді. Сүт безі қатерлі ісігінің әрбір кіші тобының өзіндік емдік схемасы және нақты болжамы бар. Дегенмен, жинақталған деректер сүт безі қатерлі ісігінің бұл кіші топтарында ұқсас белсендірілген немесе басылған гендер және ортақ сигнал беру жолдары бар деген гипотезаны қолдайды. Толық геномдық молекулалық профильдеу ісік генезисі мен прогрессиясындағы молекулалық өзгерістерді анықтай алады және негізгі гендерді анықтаудың жоғары тиімді әдісі ретінде көрсетті. Зерттеуде биоинформатикалық талдау арқылы сүт безі қатерлі ісігінің онкогенезі мен болжамындағы ықтимал маңызды гендер мен негізгі жолдарды зерттеуге тырысты. Нәтижесінде кейбір хаб гендерінің онкогенезбен және сүт безі қатерлі ісігінің болжамымен байланысты екенін көрсетеді.

1.7.2 Биоинформатикадағы сүт безі қатерлі ісігінің жоғары өнімді геномдық тәсілдері

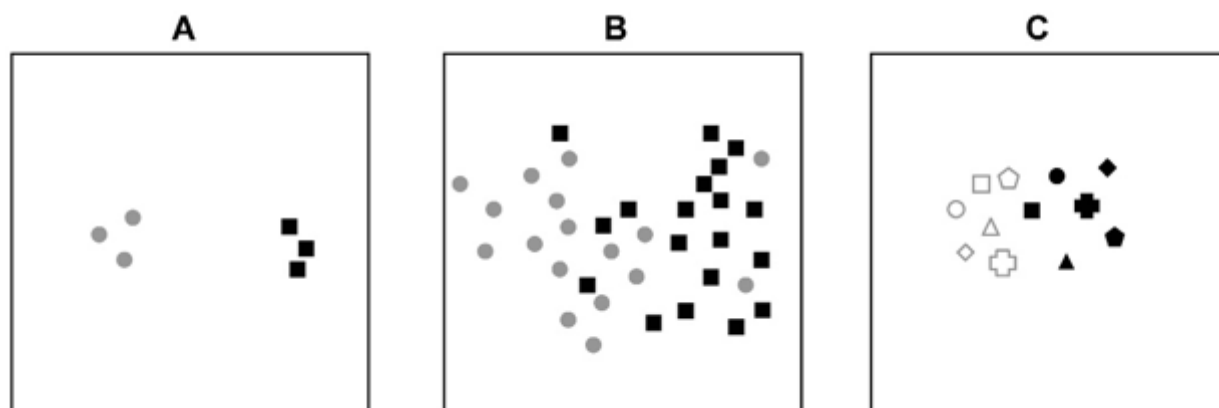
Сүт безі қатерлі ісігінің мамандары жоғары өнімді геномдық тәсілдердің әлеуетін тез игерді. [37] Бұл әдістердің тартымдылығы мындаған ДНҚ тізбегіндегі, мРНҚ транскрипттеріндегі, пептидтердегі немесе метаболиттердегі вариацияларды бір уақытта өлшеу мүмкіндігін ескере отырып айқын көрінеді. [38] Жиі қолданылатын клиникалық-патологиялық сипаттамаларға сүйене отырып (мысалы, ісік мөлшері, лимфа түйіндерінің

катысуы, кіші түрі, дәрежесі және эстроген рецепторларының (ER) экспрессиясы) сүт безі қатерлі ісігінің өте гетерогенді ауру екені анықталған. Молекулалық профильдеу мұны растады және оның негізінде жатқан күрделілікті көрсетті, сондықтан әртүрлі ісіктердің әртүрлі емдеу әдістеріне жауап беруі қазіргі таңда таңқаларлық емес [39]

Пациенттер тобындағы бір транскрипт немесе ақуыз деңгейін өлшеуден мыңдаған гендердің немесе ақуыздардың деңгейін бір уақытта өлшеуге көшу бірнеше тестілеу тұжырымдамалары мен жалған анықтау жиілігін жақсырақ түсіну қажеттілігін тудырады. Гипотезаны растау немесе жоққа шығарудың дәстүрлі тәсілдерінде бір геннің немесе ақуыздың деңгейі өлшенеді. Жоғары өнімді әдістердің тартымды жақтарының бірі - олар гипотезаларға емес, деректерге негізделген, сондықтан олар алдын ала біліммен шектелмейді. Бұл әдістер берілген гипотезаны дәлелдеу немесе жоққа шығару үшін тиімді пайдаланылуы мүмкін болса да, нақты құндылық жаңаларын жасау болып табылады. Үлгілер оның санына қарағанда әлдеқайда көп белгілердің (гендер, транскриптер, бір нуклеотидті полиморфизмдер, пептидтер және т.б.) болуының салдары нақты дифференциалданған белгілердің көпшілігі нақты биологиялық айырмашылықтардан гөрі кездейсоқ болуы мүмкін. Деректер жиынтығының өлшемділігі мен бірнеше рет тестілеу мәселелерін егжей-тегжейлі қарау үшін оқырмандар мысалы ретінде, Clarke et al шолуын тексере алады. Ең көп өзгертін гендердің гендік қолтаңбалары арасындағы вариация когортты таңдаудағы (биологиялық айнымалылар) және эксперименттік ауытқулардағы (техникалық айнымалылар) айырмашылықтарға байланысты болуы мүмкін. [40]

Қосымша тізбектерді будандастырудың негізгі принциптері ген экспрессиясының барлық тәсілдері үшін бірдей болғанымен, микрочип платформаларын әзірлеу мен өндіруде кейбір түбегейлі айырмашылықтар бар. Ерте микрочиптер әдетте жеке зертханаларда клондалған кДНҚ ПТР талдау өнімдерінен немесе слайдтарға басылған синтетикалық олигонуклеотидтерден жасалған. [41] Технология тез дамып, кеңейіп, көптеген басқа айнымалыларды, соның ішінде геномдық ДНҚ мутацияларын және көшірмелер санын, метилдену мен микроРНҚ-ны, ақуыз антиденелерін немесе тіндер мен жасушалардың лизаттарын анықтауға мүмкіндік берді. Коммерциялық микрочиптерге қол жеткізуге бірнеше компаниялар, соның ішінде Affymetrix (Санта-Клара, Калифорния, АҚШ), Agilent Technologies (Санта-Клара, Калифорния, АҚШ) және Illumina (Сан-Диего, Калифорния, АҚШ) көмектесті, бұл белгілі бір платформаны пайдалана отырып, зерттеу нәтижелерін салыстыру мен сәйкестікті біршама жақсартты. Тәсілдердегі айқын айырмашылықтарға және алынған гендік қолтаңбаларға қатысты консенсустың жоқтығына қарамастан, жоғарыда сипатталған барлық үш жалпы тәсілдің болжамды болжау қабілеті ұқсас. Сынақ деректер жинағы

арқылы бірнеше қолтаңбаны бағалау жеке пациенттер үшін болжанған нәтижелердің жоғары сәйкестігін көрсетті.



Сурет-4. Үлгілер арасындағы вариация деңгейі кіші топтарды ажырататын елеулі айырмашылықпен экспрессияланатын гендерді сәйкестендіру үшін қажетті санды айқындайды. Графиктердің мысалдары екі өлшемді негізгі компоненттерді талдау немесе көп өлшемді масштабтау арқылы үлгі транскриптомдарының гипотетикалық жалпы ұқсастығын білдіреді. (А) сұр шеңбер мен қара шаршының көшірмелері бір-біріне тығыз, бірақ бір-бірінен айқын ерекшеленеді. ((В) сұр шеңберлер қара шаршылардан онша айқын емес, сондықтан (А) сияқты сенімділік деңгейімен тұрақты айырмашылықтарды анықтау үшін көп сандар қажет. (С) салыстырылған үлгілер, мысалы, өлшеулерге дейін (толтырылған таңбалар) және кейін (ашық таңбалар), белгілі бір емдеуге немесе процедураға байланысты экспрессияның жалпы өзгерістерін анықтауды жеңілдетеді.

Сүт безінің ісіктерін әртүрлі емдеу әдістеріне болжамды немесе сезімталдықты болжау үшін көптеген гистопатология және молекулалық патология әдістері арқылы оқшаулауға болады. [42] Сүт безі қатерлі ісігі саласында ген экспрессиясының микрочиптерін талдаудың үш негізгі әдісі бар. Біріншісі ол - бақыланбайтын талдау әдісі, онда ісіктер "ішкі" гендер жиынтығы бойынша кіші топтарға топтастырылады, бұл ісіктердің ішіндегі емес, ісіктер арасындағы ген экспрессиясындағы айырмашылықтарды көрсетеді. Люминальды және базальды кіші типтер арасындағы ең жарқын молекулалық айырмашылықтар бірнеше рет анықталды және әртүрлі технологиялар мен платформалар арқылы расталды. [43] Сондай-ақ «молекулалық апокриндік» ісіктерді анықтау және одан әрі ER-теріс ісіктерді кем дегенде бес түрлі кіші түрге бөлу болды. Анықталған молекулалық кіші типтер клиникалық нәтижелердегі айтарлықтай айырмашылықтармен байланысты, олар емдеудің әртүрлі тәсілдеріне жақсы жауап береді. Қазіргі уақытта ER/прогестерон рецепторларының (PR)/HER2 антиандрогендік терапиясының екінші кезеңі жүргізілуде. Ал екінші екі әдіске тоқталсақ, бұл

жеке клиникалық бақылау деректеріне немесе ER күйі, дәрежесі және пролиферациясы сияқты ісіктің биологиялық сипаттамаларына негізделген бақыланатын тәсілдерді пайдаланады. Амстердам тобының 70 генінің қолтаңбасы (сDNA массивтері) мен Роттердам тобының 76 генінің қолтаңбасы (Affymetrix олигонуклеотидтік массивтері) арасында сәйкестіктің болмауы (үш ген) бақылау деректеріне негізделген геномдық тәсілдердің сенімсіз екендігінің дәлелі ретінде айтылып өткен. Бірақ логикалық тұрғыдан алғанда, бақыланбайтын тәсілдермен көрсетілген гетерогенділік әртүрлі үлгілер топтарын пайдалана отырып, екі шағын зерттеу нәтижелерінің қайталануын жоққа шығарады. Қолтаңбалар арасындағы сәйкессіздікті қосу критерийлеріндегі (лимфа түйіндерінің жағдайы, ісік диаметрі, адьювантты емдеу және т.б.), платформадағы (кДНҚ немесе олигонуклеотид массивтері немесе сандық кері транскрипциялық ПТР (qRT-ПТР)) және әртүрлі қолданылатын деректерді талдау әдістеріндегі айырмашылықтарды зерттеу кезінде түсіндіруге болады.

1.8 Сүт безінің қатерлі ісігі кезінде биоинформатикалық бағдарламалық жасақтамасы

1.8.1 RNA-Seq деректерін талдау құралдары

Бұл құралдар RNA-Seq деректерін өңдеуге және талдауға арналған, олар гендердің өзгерісі, баламалы оқиғалар және басқалары туралы мағлұмат береді. Транскриптерді анықтау және ген экспрессиясын сандық бағалауда РНҚ-ның геном мен протеом арасындағы негізгі делдал ретіндегі рөлі ашылғаннан бері молекулалық биологиядағы негізгі әрекеттерге айналды. РНҚ секвенциясының артықшылығы мынада: бұл екі аспект - анықтау және сандық бағалау бір жоғары өнімді секвенирлеу талдауында біріктірілуі мүмкін. RNA-Seq-ті кеңінен енгізу геномика қауымдастығынан әлдеқайда асып түсті және өмір туралы ғылымдар саласындағы зерттеушілер қауымдастығы қолданатын құралдар жинағының стандартты бөлігі болды. Келесі ұрпақ секвенирлеу деректеріне (NGS) негізделген RNA-seq талдауы соңғы уақытта бүкіл транскриптом деңгейінде ген экспрессиясын талдаудың іс жүзінде стандартына айналды. Бұл талдау көбінесе жасушалар мен тіндердегі молекулалық оқиғалар туралы механикалық гипотезаларды құру үшін өте маңызды. Мысал ретінде физиологиялық тітіркендіргіштерге жасушалық реакциялар, эксперименттік бұзылулардың белгілі бір гендер мен жолдарға әсері және аурулардағы гендік реттеудің бұзылуы болуы мүмкін. РНҚ секвенциясы қызығушылық тудыратын тіндерден немесе жасушалардан жалпы РНҚ-ны оқшаулауды, содан кейін ДНҚ кітапханаларын құруды және осы кітапханаларды келесі ұрпақ секвенирлеу құралымен ретке келтіруді қамтиды.

STAR:

Жылдам RNA-seq жұптастырғышы, ол сплайстық жүйкелерді өңдей алады және геномға оқуларды картаға түсіруге жиі қолданылады. Жоғары өнімді секвенирлеудің үлкен оқу жиынтығын анықтамалық геномға сәйкестендіру РНҚ секвенирлеу деректерін талдаудың негізгі қадамдарының бірі болып табылады. STAR бағдарламалық жасақтамасы бұл тапсырманы жоғары дәлдікпен және жылдамдықпен орындайды. STAR бағдарламалық пакеті анықтамалық геном бойынша РНҚ секвенирлеу оқуларының жоғары дәлдікті және ультра жылдам туралануын қамтамасыз етеді. Аннотацияланған және жаңа сплайс қосылыстарын анықтаудан басқа, STAR химерикалық және сақиналы РНҚ сияқты күрделі РНҚ тізбегін анықтауға қабілетті. STAR кез-келген ұзындықтағы біріктірілген тізбектерді қателіктердің орташа деңгейімен теңестіре алады, бұл жаңа секвенирлеу технологияларының масштабталуын қамтамасыз етеді. STAR көптеген кейінгі талдаулар үшін пайдаланылуы мүмкін шығыс файлдарын жасайды, мысалы, транскрипт/ген экспрессиясын сандық бағалау, дифференциалды ген экспрессиясы, жаңа изоформаларды қайта құру, сигналдарды бейнелеу және т.б.

TopHat және Cufflink

RNA-seq деректерін талдауда, әсіресе транскрипт жинағы мен өзгерістерді талдауда алғашқы танымал құралдар, бірақ олар HISAT және StringTie сияқты басқа құралдармен алмастырылды. TopHat және Cufflinks-жоғары өнімді мРНҚ секвенциясы (RNA-seq) деректеріндегі гендерді анықтауға және экспрессияны жан-жақты талдауға арналған тегін, ашық бастапқы бағдарламалық құрал. Олар бірге биологтарға белгілі гендердің жаңа гендері мен жаңа сплайс нұсқаларын анықтауға, сондай-ақ екі немесе одан да көп жағдайларда ген мен транскрипт экспрессиясын салыстыруға мүмкіндік береді.

HISAT2

Тиімді және жылдам бағдарлау құралы, ол бүкіл геномды және транскриптомды бағдарлауға арналған. HISAT2-адамның жалпы популяциясымен (сонымен қатар бір анықтамалық геноммен) жаңа буын секвенирлеу оқуларын (тұтас геномды, транскриптомды және экзомды секвенирлеу деректері) сәйкестендіруге арналған жылдам және сезімтал туралау бағдарламасы. GCSA (графикке арналған BWT кеңейтімі) негізінде біз FM индексінің (GFM) графигін әзірледік және жүзеге асырдық-бұл өзіндік тәсіл және алғашқы енгізу. Жалпы популяцияны білдіретін бір жаһандық GFM индексіні пайдаланудан басқа, HISAT2 бүкіл геномды қамтитын шағын GFM индекстерінің үлкен жиынтығын пайдаланады (әрбір индекс адам популяциясын қамту үшін қажетті 55 000 индексті құрайтын 56 Кбп геномдық

аймақты білдіреді). Бірнеше туралау стратегияларымен біріктірілген бұл шағын индекстер (жергілікті индекстер деп аталады) реттілік оқуларын тиімді теңестіруге мүмкіндік береді. Бұл жаңа индекстеу схемасы Hierarchical graph FM index (HGFM) деп аталды. Біз hisat 2-ді HISAT және Bowtie2 іске асырулары негізінде жасадық. HISAT2 Sam форматындағы туралауды шығарады, бұл SAM қолданатын көптеген басқа құралдармен (мысалы, SAMtools, GATK) өзара әрекеттесуге мүмкіндік береді. HISAT2 GPLv3 лицензиясы бойынша таратылады және Linux, Mac OS X және Windows астында пәрмен жолынан іске қосылады.

HTSeq

Бұл Python пакеті бағдарлау құралдарымен бірге жұмыс істейді және мүмкіндіктерде оқуларды санайды. HTSeq-бұл RNA секвенирлеуінен алынған мәліметтер негізінде геномдағы гендерге немесе басқа функционалды элементтерге теңестірілген оқулар санын санау үшін қолданылатын жоғары өнімді деректерді талдау құралы. Ол биоинформатикада ген экспрессиясын талдау және әртүрлі жағдайларда немесе үлгілерде транскрипция деңгейін бағалау үшін жиі қолданылады. HTSeq оқуды теңестіру файлдарымен (мысалы, SAM/BAM форматында) және геномдық Аннотация файлдарымен (мысалы, GTF/GFF форматтарында) жұмыс істейді және RNA реттілігіне негізделген ген экспрессиясына сапалық және сандық талдау жасауға мүмкіндік береді.

DESeq2

R статистикалық пакеті, ол RNA-seq деректерінен санау мәліметтерін талдайды және өзгерістерді тексереді. DESeq2 бағдарламасы RNA реттілігі (RNA-seq) деректеріне негізделген гендердің дифференциалды экспрессиясын талдауға арналған. Оның жұмыс принципі эксперименттік деректердің бірегей сипаттамаларын ескеретін және жалпы оқылымдағы өзгергіштік және ішкі өзгергіштік факторлары сияқты жағымсыз әсерлерді жоятын статистикалық әдістерге негізделген. DESeq2 бағдарламасының негізгі кезендері:

1. Деректерді дайындау: DESeq2 эксперимент үлгілеріндегі әрбір ген үшін оқу деректерін дайындаудан басталады. Бұған оқуларды геномға туралау, әрбір генге түсетін оқулар санын санау және үлгілер арасындағы оқулардың жалпы санындағы айырмашылықтарды есепке алу үшін деректерді қалыпқа келтіру кіреді.
2. Дифференциалды экспрессияны бағалау: деректерді дайындағаннан кейін DESeq2 әртүрлі үлгілер немесе үлгілер топтары арасындағы гендердің дифференциалды экспрессиясын бағалау үшін статистикалық модельдерді пайдаланады. Бұл p мәндері сияқты әртүрлі статистикалық көрсеткіштерді есептеуді және шарттар арасындағы экспрессияда

кандай гендер айтарлықтай ерекшеленетінін анықтау үшін бірнеше салыстыруды түзетуді қамтиды.

3. Нәтижелерді түсіндіру: талдаудан кейін DESeq2 шарттар арасындағы экспрессияда айтарлықтай ерекшеленетін гендердің тізімдерін, сондай-ақ олармен байланысты статистикалық көрсеткіштерді ұсынады. Бұл нәтижелерді ген экспрессиясындағы айырмашылықтардың биологиялық маңыздылығын түсіну үшін биологиялық мәліметтер базасы мен талдаулар арқылы одан әрі түсіндіруге болады.

edgeR

DESeq2 сияқты, ол өзгерістерді талдауда қолданылады, әсіресе аз үлгілері бар үшін тұрақты. EdgeR бағдарламасы RNA реттілігі (RNA-seq) деректеріне негізделген гендердің дифференциалды экспрессиясын талдауға арналған. Оның жұмыс принципі әртүрлі үлгілер немесе үлгілер топтары арасындағы ген экспрессиясындағы айырмашылықтардың статистикалық маңыздылығын анықтау үшін дисперсияны модельдеуге және статистикалық тестілеуге негізделген. EdgeR бағдарламасының негізгі кезеңдері:

1. Деректерді дайындау: edgeR бағдарламасы эксперимент үлгілеріндегі әрбір ген үшін оқу деректерін дайындаудан басталады. Бұл оқуларды геномға теңестіруді, әрбір генге түсетін оқулар санын санауды және үлгілер арасындағы оқулардың жалпы санындағы айырмашылықтарды есепке алу үшін деректерді қалыпқа келтіруді қамтиды.
2. Дисперсияны модельдеу: EdgeR әртүрлі үлгілердегі өзгергіштігін ескеру үшін ген экспрессиясының дисперсиясын модельдейді. Бұл гетероскедастиканы, сондай-ақ деректердің басқа ерекшеліктерін ескеруге мүмкіндік береді, бұл талдауды дәлірек және сенімді етеді.
3. Дифференциалды экспрессияны бағалау: дисперсияны модельдеуден кейін edgeR үлгілер немесе үлгілер топтары арасындағы ген экспрессиясының айырмашылықтарының статистикалық маңыздылығын бағалау үшін теріс биномдық үлестіру сияқты статистикалық сынақтарды пайдаланады. Бұл p мәндерін есептеуді және шарттар арасындағы экспрессияда қандай гендер айтарлықтай ерекшеленетінін анықтау үшін бірнеше салыстыруды түзетуді қамтиды.
4. Нәтижелерді түсіндіру: талдаудан кейін edgeR шарттар арасындағы экспрессияда айтарлықтай ерекшеленетін гендердің тізімдерін, сондай-ақ олармен байланысты статистикалық көрсеткіштерді ұсынады. Бұл нәтижелерді ген экспрессиясындағы айырмашылықтардың биологиялық маңыздылығын түсіну үшін биологиялық мәліметтер базасы мен талдаулар арқылы одан әрі түсіндіруге болады.

limma (voom):

Бастапқыда микрочиптер үшін жасалған, бірақ RNA-seq үшін бейімделген, ол күрделі статистикалық модельдермен өзгерістерді талдауда пайдалы. Voom (variance modeling at the observational level) әдісін қолдана отырып, Limma (Microarray data үшін Linear Models) бағдарламасы микрочип деректері мен RNA секвенциясы (RNA-seq) негізінде гендердің дифференциалды экспрессиясын талдауға арналған. Voom әдісімен Limma-әртүрлі биологиялық жағдайлар мен процестерге байланысты гендерді анықтау үшін биомедициналық зерттеулерде кеңінен қолданылатын дифференциалды ген экспрессиясын талдаудың қуатты құралы.

1.8.2 Микрочип талдау құралдары

DNA микрочиптер деректерін талдауға арналған, ол RNA-seq сияқты танымал болса да, кейбір қолданыстар үшін әлі де қолданылады. Микрочиптермен алынған деректерді өңдеуге, талдауға және түсіндіруге арналған бағдарламалық жасақтама мен әдістер (microarrays). Бұл құралдар гендік экспрессияны, генетикалық вариацияны және биология мен медицинаның басқа аспектілерін зерттеуде шешуші рөл атқарады.

Affymetrix Expression Console

Affymetrix микрочиптер деректерін нормалау және сапа бақылауына арналған құралдар ұсынады. Бұл Affymetrix микрочиптері арқылы алынған деректерді талдауға және өңдеуге арналған бағдарламалық құрал. Бұл зерттеушілерге GeneChip сияқты Affymetrix микрочиптері арқылы алынған ген экспрессиясының деректерін жүктеуге, қалыпқа келтіруге және талдауға мүмкіндік береді. Пайдаланушылар деректерді сапалы бағалауды, стандарттауды, қалыпқа келтіруді және дифференциалды ген экспрессиясын талдауды жүргізе алады, бұл әртүрлі жағдайларда немесе үлгілерде экспрессиясы өзгертін гендерді анықтауға көмектеседі.

GeneSpring

Микрочиптер деректерін визуализациялау және талдау құралдарын ұсынады, оның ішінде жолдарды талдау бар. Бұл гендік экспрессия деректерін және жоғары өнімді секвенирлеудің басқа түрлерін талдауға және визуализациялауға арналған бағдарламалық құрал. GeneSpring дифференциалды ген экспрессиясын талдау, кластерлік талдау, функционалдық аннотация және нәтижелерді визуализациялау құралдарын ұсынады. Бұл құрал биология мен медицинаны зерттеу үшін, сондай-ақ Биотехнология мен фармацевтика саласындағы зерттеулер үшін кеңінен қолданылады.

GEO2R

Пайдаланушыларға GEO дерекқорындағы үлгілер арасында гендік өріс профильдерін салыстыруға және талдауға мүмкіндік беретін интерактивті веб-құрал. Бұл gene Expression omnibus (GEO) дерекқоры ұсынатын веб-құрал, ол зерттеушілерге GEO-да қол жетімді микрочип деректеріне негізделген гендердің дифференциалды экспрессиясын жылдам талдауға мүмкіндік береді. Пайдаланушылар шикі ген экспрессиясының деректерін немесе алдын ала өңделген деректерді жүктей алады, содан кейін олар үлгі топтарын салыстыру және осы топтар арасындағы ген экспрессиясының статистикалық маңызды айырмашылықтарын анықтау арқылы талдау жасай алады.

Бұл құралдар зерттеушілерге микрочиптер немесе басқа әдістер арқылы алынған гендік экспрессия деректерін талдауға және түсіндіруге көмектеседі. Олар зерттелетін үлгілерде болатын биологиялық процестерге сапалы талдау мен қорытынды жасау үшін әртүрлі функциялар мен мүмкіндіктер береді.

1.8.3 Жолдар мен желілерді талдау

Биоинформатикада жолдар мен желілерді талдау гендер, ақуыздар, метаболиттер және жол деректері ретінде ұсынылған басқа молекулалық компоненттер сияқты биологиялық объектілер арасындағы күрделі қатынастар мен өзара әрекеттесулерді зерттеу үшін қолданылады. Бұл әдіс желілік ұйым деңгейінде биологиялық процестер мен патологияларды түсінуде шешуші рөл атқарады. Гендік өрістер деректеріндегі әрекеттесулер мен жолдарды талдау және көрсету құралдары

Gene Set Enrichment Analysis (GSEA)

Екі биологиялық жағдай арасында алдын ала анықталған гендер топтарының статистикалық тұрғыдан маңызды, үйлесімді өзгерістер көрсететінін анықтайтын есептеу әдісі. бұл әртүрлі биологиялық күйлер немесе жағдайлар арасындағы ген экспрессиясындағы статистикалық маңызды айырмашылықтарды анықтау үшін қолданылатын гендік деректерді талдау әдісі. Ол биологиялық процестерге, жолдарға немесе функцияларға бірлесіп қатысатын гендік жиынтықтар немесе гендік желілер деп аталатын гендік топтарды анықтауға мүмкіндік береді. GSEA жеке гендердің орнына гендер тобының экспрессиясындағы өзгерістерді анықтауға мүмкіндік береді, бұл биологиялық процестердегі кеңірек өзгерістерді анықтауға мүмкіндік береді.

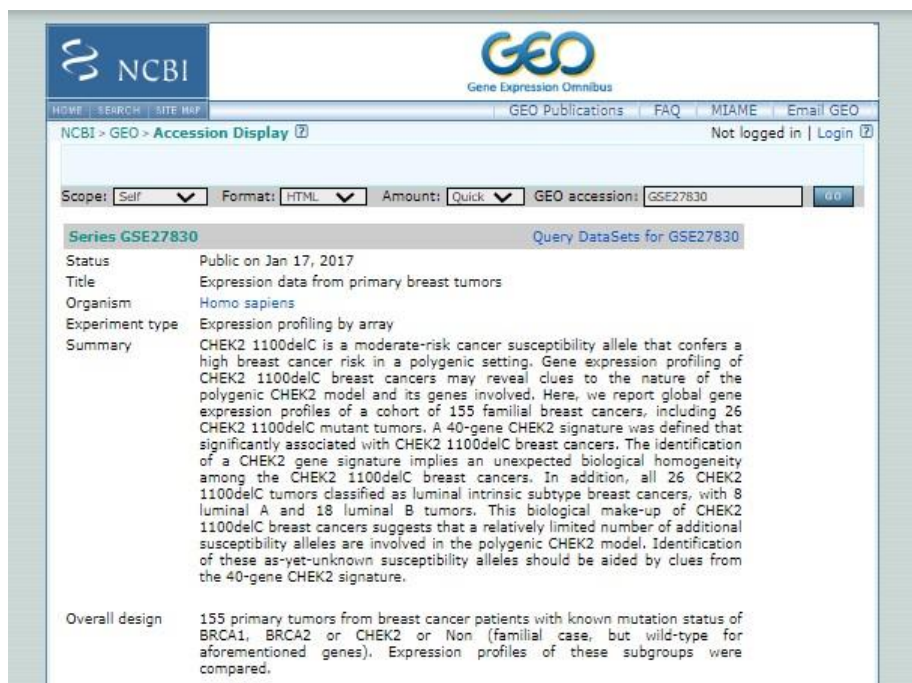
Ingenuity Pathway Analysis (IPA)

Ingenuity Pathway Analysis (IPA) - зерттеушілерге биологиялық жолдар мен биомолекулалар арасындағы өзара әрекеттесу туралы білімді пайдалана отырып, гендік экспрессиялар, мутациялар, протеомдық деректер және т.б.

сияқты биологиялық деректерді зерттеуге және түсіндіруге мүмкіндік беретін биоинформатика саласындағы деректерді талдау бағдарламалық құралы. IPAD белгілі бір биологиялық процестермен немесе патологиялармен байланысты биологиялық маңызды жолдарды, гендер мен метаболиттер желілерін анықтауға, сондай-ақ дәрі-дәрмектер мен препараттардың әсер ету механизмдерін болжауға мүмкіндік береді. Эксперименталды деректер жиынтықтарына немесе қызығушылық тудыратын гендерге ең қатысты биологиялық механизмдер, жолдар және функциялар туралы түсініктер береді.

2. Қолданылған материалдар мен әдістер

Жұмыстың көп бөлігі әдебиет шолуға бағыталған болатын. Зерттеу барысында 155 отбасылық сүт безі қатерлі ісігінің, соның ішінде CHEK2 мутациясы бар 26 ісіктің жаһандық ген экспрессиясының профилдері алынды. Осы жұмыстың барысында біз сүт безі қатерлі ісігінің пайда болу механизмін, емдеу шараларын және молекулалық тұрғыдан жете түсіндік. Дипломдық жұмыстың келесі бөлігінде түрлі биоинформатикалық құралдарды пайдалана отырып RNA-Seq project орындалды. Жұмыстың көп бөлігі NCBI сайтында GEO базасын қолданылып тәжірибе жүзеге асты.



Сурет 5. GEO базасынан келесі образецтің толық гендегі экспрессия профайлын таңдап алдық.

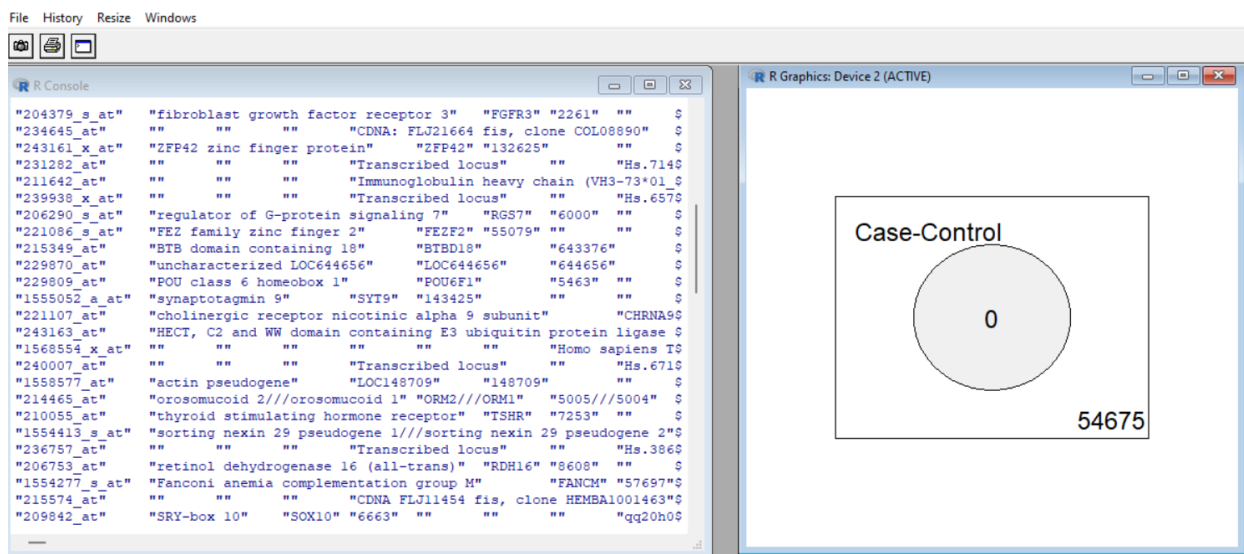
Platforms (1) [GPL570 \[HG-U133_Plus_2\]](#) Affymetrix Human Genome U133 Plus 2.0 Array

Samples (155) [GSM687017](#) primary breast cancer, sample_1028
[More...](#) [GSM687018](#) primary breast cancer, sample_1159
[GSM687019](#) primary breast cancer, sample_1181

This SubSeries is part of SuperSeries:

[GSE54219](#) Molecular genomic and transcriptomic profiling of familial breast cancer.

Сурет 6. 155 таңдап алынған үлгілер



Сурет 7. Жұмыс барысында R 4.2.2 және RStudio бағдарламалық жабдықталары қолданылды.

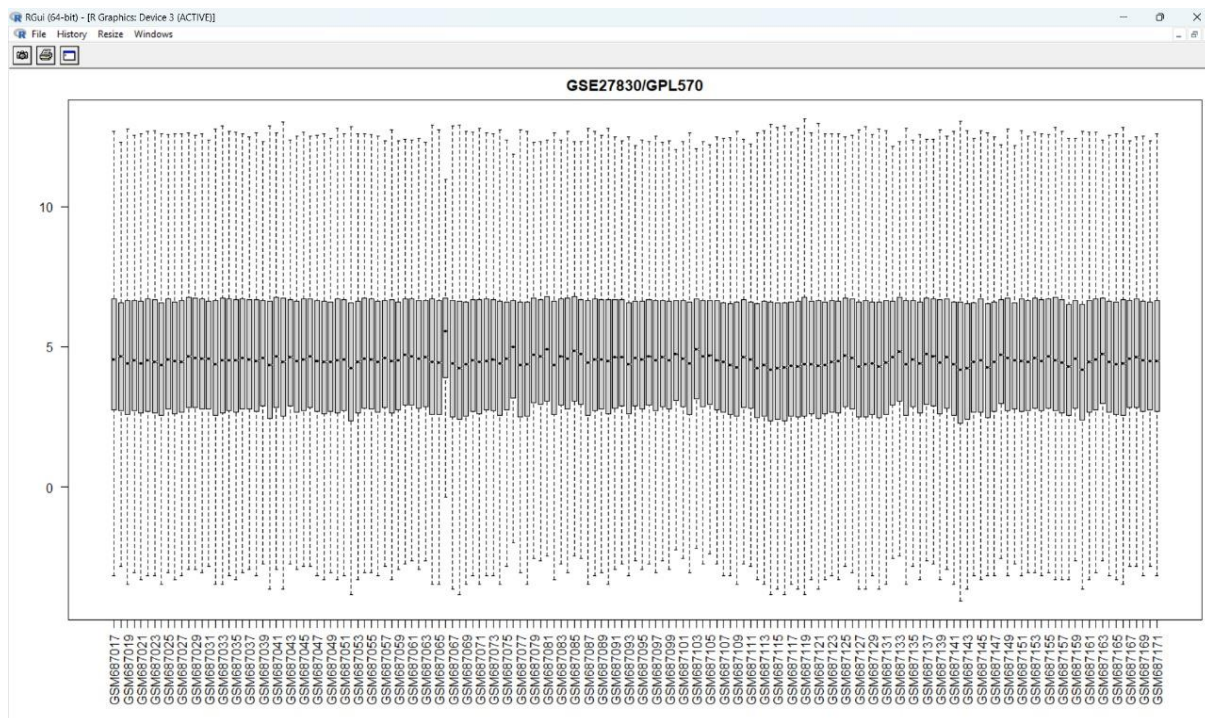
Бұл R сценарийі бастапқы сүт безі ісіктерінің гендік экспрессиясы туралы деректерді қамтитын GSE27830 деректер жинағындағы ген экспрессиясы деректерін талдау және визуализациялау үшін пайдаланылды. Талдау барысында қажетті пакеттер: GEOquery, limma және uthar жүктелді. Логарифмдік трансформацияны қолдана отырып, ген экспрессиясының деректерін түрлендірілді.

Ген экспрессиясы мәндерінің таралу тығыздығының графигін құрылды.

Өрнектің орташа мәні мен дисперсия арасындағы трендті талдау үшін орташа-дисперсиялық тәуелділіктің графигін құру.

UMAP алгоритмін қолдана отырып, екі өлшемді кеңістікте деректерді визуализациялау үшін UMAP графигін құру (multidimensional scaling).

3. Зерттеу нәтижелері



Сурет 8. "Box-and-whisker" графигі — мәндердің таралуын көрнекі түрде көрсетуге және шығарындыларды анықтауға мүмкіндік беретін статистикалық деректерді визуализациялау.

Бұл сурет деректердің графикалық бейнесі, атап айтқанда микрочиптегі ген экспрессиясының жылу картасы. Мұндай бейнелеу биоинформатика мен биологияда әртүрлі жағдайларда немесе үлгілерде мыңдаған гендердің экспрессия деңгейлерін көрсету үшін қолданылады. Бұнда бағандар әртүрлі уақыт нүктелерін, ал жолдар жеке гендерді білдіреді. Әр ұяшықтағы түс қарқындылығы белгілі бір жағдайда геннің экспрессия деңгейін көрсетеді, қараңғы түстер әдетте жоғары экспрессияны көрсетеді. Мұндай жылу карталары биологиялық процестерді түсінуге және белгілі бір жағдайларда белсенді гендерді анықтауға пайдалы болуы мүмкін. Олар әсіресе қатерлі ісік зерттеулерінде пайдалы, мұнда гендік белсенділікті анықтау аурудың механизмдері мен ықтимал емдеу әдістері туралы жаңа түсінуімізге көмектесті. Мұнда біздің тәжірибемізде алынған 155 геннің экспрессия деңгейі көрсетілді. СНЕК2 1100delc генінің қолтаңбасын анықтау сүт безі қатерлі ісігінің арасында күтпеген биологиялық гетерогенділікті көрсетеді. Бұл СНЕК2 1100delc мутациясының болуы сүт безі қатерлі ісігінің көрінісіне әртүрлі дәрежеде немесе оның молекулалық ерекшеліктеріне әсер етуі мүмкін дегенді білдіреді.

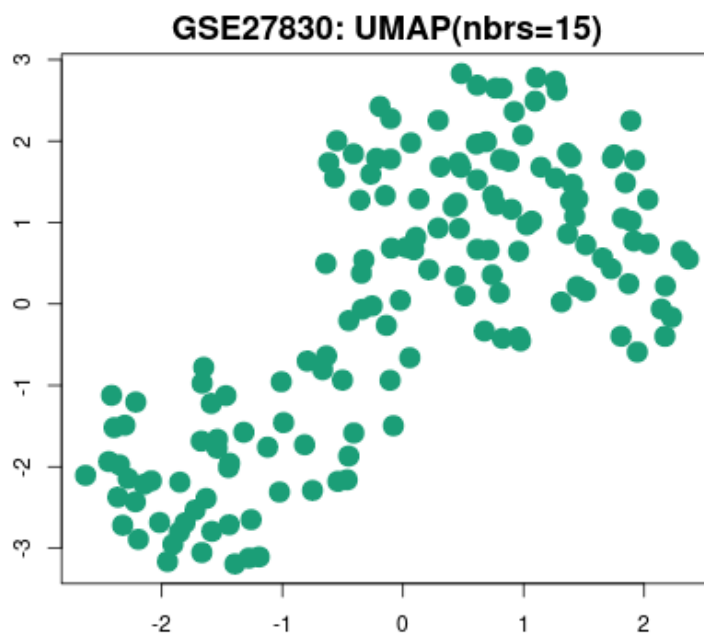


График 1. Деректердің өлшемін азайту және оларды екі өлшемді кеңістікте визуализациялау үшін қолданылады, бұл мәліметтер құрылымы мен олардың кластерлерін жақсы түсінуге мүмкіндік береді. X-осі: шамамен -2.5-тен 2.5-ке дейінгі мәндері бар. Y осі: шамамен -3.5-тен 3.5-ке дейінгі мәндерге ие.

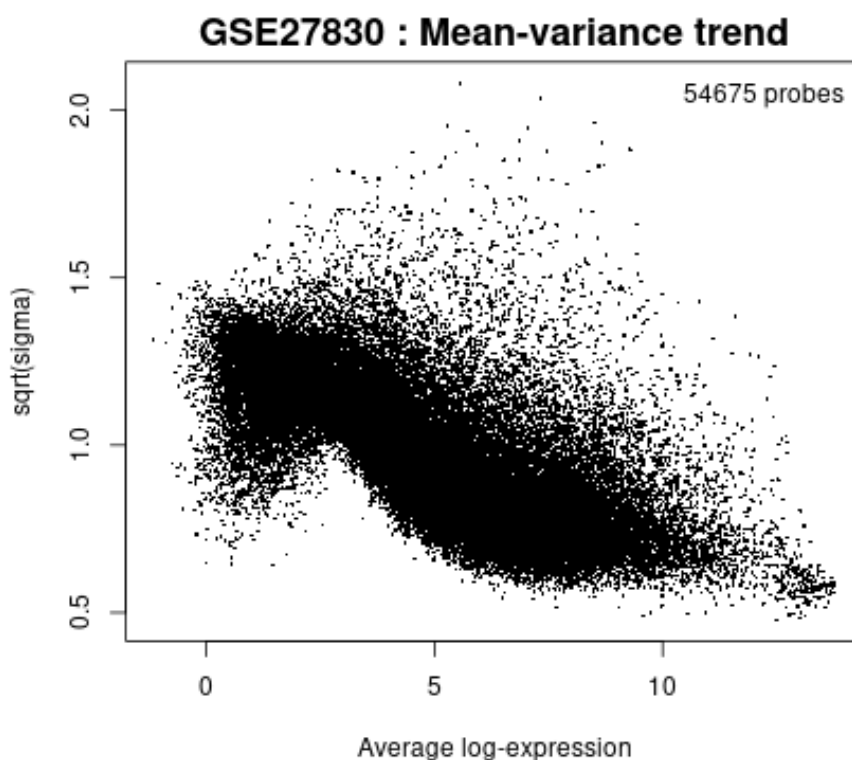


График 2. Бұл график сызықтық модельге сәйкес келгеннен кейін өрнек деректерінің өзгеруіне орташа мәннің тәуелділігін тексеру үшін пайдаланылады. Бұл деректерде елеулі айырмашылықтардың бар-жоғын анықтауға көмектесті. Бұл диаграмма орташа өзгеру тенденциясын есепке алу

арқылы бағалауға көмектеседі. Орташа мәннің өзгеру тенденциясы күшті болған кезде сынақ нәтижелерінің дәлдігін арттырады.

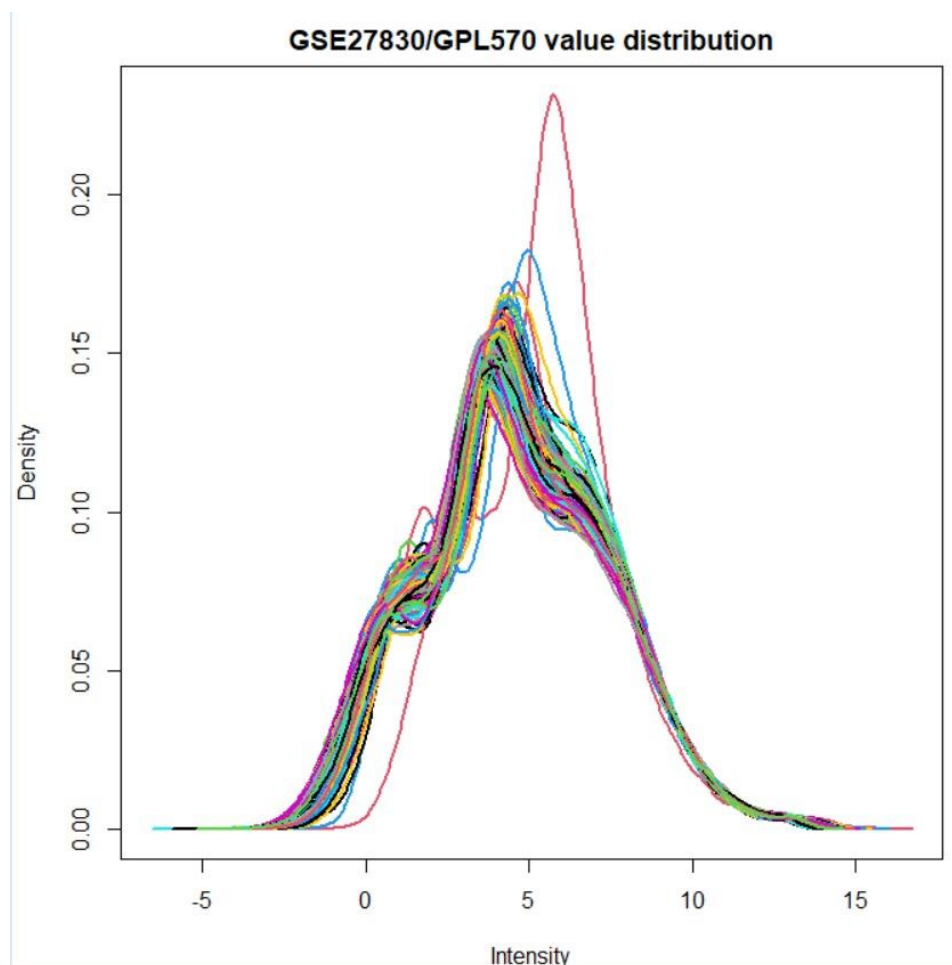


График 3. Бұл диаграмма таңдалған үлгілердің мәндерінің таралуын қарау үшін қолданылады. Тығыздық графигі деректер жиынында мәндердің қалай бөлінетінін түсіну үшін ген экспрессиясының қарқындылық мәндерінің таралуын визуализациялау үшін қолданылады. Графикте қарқындылық мәндерінің көпшілігі 0-ден 10-ға дейінгі диапазонда шоғырланғанын көруге болады, шыңы шамамен 5, бұл зерттелетін үлгілердегі экспрессия мәндерінің қалыпты таралуын көрсетуі мүмкін.

ҚОРЫТЫНДЫ

Бұл зерттеуде біз жалпыға қолжетімді деректер жиынындағы ген экспрессиясының профилі мен клиникалық сипаттамаларын биоинформатикалық талдау арқылы сүт безі қатерлі ісігінің онкогенезі мен болжамындағы ықтимал маңызды гендер мен негізгі жолдарды зерттеуге тырыстық.

Нәтижелер кейбір орталық гендердің сүт безі қатерлі ісігінің онкогенезімен және болжамымен байланысты екенін көрсетті. Жұмыс барысында деректер қорынан 155 отбасылық бастапқы сүт безі қатерлі ісігінің үлгілері үшін ген экспрессиясының профильдерін қарастырдық. Жұмыс нәтижесінде 47- BRCA1 үлгісі, 6 - BRCA2 мутантты үлгілері, 26-СНЕК2 үлгілері және осы үш гендегі мутациясыз 76 үлгілер қаралды. Осы кіші топтардың экспрессиялық профильдерін салыстыру жүргізілді.

Талдау нәтижесінде сүйек пен миға метастаз беру ықтималдығымен байланысты гендік қолтаңбалар анықталды. Бұл қолтаңбалар сүт безі қатерлі ісігімен ауыратын науқастарда метастатикалық қауіпті болжау үшін пайдалы болуы мүмкін. Біз анықталған бірнеше негізгі нәтижелер:

1. СНЕК2 қолтаңбасы: СНЕК2 1100delC мутациясы бар науқастарда сүт безі қатерлі ісігімен айтарлықтай байланысты гендік қолтаңба табылды. Бұл мутация болған кезде ісіктердің бірегей биологиялық ерекшеліктерінің болуын көрсетеді.
2. Люминальды кіші түрлері: СНЕК2 1100delC мутациясы бар барлық ісіктер сүт безі қатерлі ісігінің люминальды кіші түрлері ретінде жіктеледі. Бұл осы мутациясы бар ісіктердің молекулалық сипаттамалары ұқсас екенін көрсетеді.
3. Сезімталдық аллельдерінің шектеулі саны: нәтижелер СНЕК2 полигендік моделіне сүт безі қатерлі ісігінің дамуына әсер ететін генетикалық факторлардың шектеулі саны қатысатынын көрсетеді. Бұл қатерлі ісіктің осы түріне байланысты қосымша генетикалық факторларды іздеуге көмектеседі.

Сонымен, соңғы уақытта онтогенезге қатысты жинақталған мәліметтерге қарамастан, жаңа мақсаттарды белсенді іздеу жалғасуда, оған препараттарды әзірлеу сүт безі обырын емдеудің тиімділігін едәуір арттыруға мүмкіндік береді.

ПАЙДАЛАНЫЛҒАН ӘДЕБИЕТТЕР ТІЗІМІ

- [1] Karen S Johnson & Emily F Conant & Mary Scott Soo, MD (2020). Molecular Subtypes of Breast Cancer: A Review for Breast Radiologists.
- [2] Heer E., Harper A., Escandor N., Sung H., McCormack V., Fidler-Benaoudia M.M. Global burden and trends in premenopausal and postmenopausal breast cancer: a population-based study. *Lancet Global Health*. 2020;8(8):e 1027–e1037. [PubMed] [Google Scholar]
- [3] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021. [PubMed] [Google Scholar]
- [4] Breasted JH, editor. *The Edwin Smith Surgical papyrus*. Chicago, Illinois: The University Chicago Press; 1930. special edition, 1984. [Google Scholar]
- [5] Gabrielle LeBlanc, Inkoo Lee, Henry Carretta, Yi Luo, Debajyoti Sinha, George Rust. Rural-Urban Differences in Breast Cancer Stage at Diagnosis. 2022 doi: 10.1089/whr.2021.0082
- [6] Maryam Gholizadeh, S. A. P. , Reza Pasandideh. Proteomics and Bioinformatics Approaches for Breast Cancer Researches. *International Journal of Agriculture and Crop Sciences* 5, 1863-1868 (2013).
- [7] Allinen, M. et al. Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell* 6, 17-32 (2004).
- [8] Узунис, К. Становление биоинформатики: историческая перспектива, краткий обзор и будущие тенденции. в *Биоинформатика в раке и терапии рака* (ред. Гордон, Г. Дж.) 1-11 (Humana Press, 2009).
- [9] Ouzounis, C. A. Rise and demise of bioinformatics? Promise and progress. *PLoS Comput Biol* 8, e1002487 (2012)
- [10] Daisuke Kihara, Y. D. Y. , Troy Hawkins. Bioinformatics resources for cancer research with an emphasis on gene function and structure prediction tools. *Cancer Informatics* 2, 25-35 (2006).
- [11] Adams M. D. , K. J. M. , Gocayne J. D. , Dubnick M. , Polymeropoulos M. H, Xiao H. , Merril C. R. , Wu A. , Olde B. , Moreno R. F. , et al. Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* 252, 1651-1656 (1991)

[12] Velculescu V. E. , Z. L. , Vogelstein B. , Kinzler K. W. Serial analysis of gene expression. *Science* 270, 484-487 (1995)

[13] Sydney Brenner, M. J. , John Bridgham, George Golda, David H. Lloyd, Davida Johnson, Shujun Luo, Sarah McCurdy, Michael Foy, Mark Ewan, Rithy Roth, Dave George, Sam Eletr, Glenn Albrecht, Eric Vermaas, Steven R. Williams, Keith Moon, Timothy Burcham, Michael Pallas, Robert B. DuBridge, James Kirchner, Karen Fearon, Jen-i Mao, Kevin Corcoran. Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nature Biotechnology* 18, 630-634 (2000).

[14] Lon R. Cardon, G. R. A. Using haplotype blocks to map human complex trait loci. *Trends in Genetics* 19, 135-140 (2003).

[15] Andrew G. Clark, R. N. , James Signorovitch, Tara C. Matise, Stephen Glanowski, Jeremy Heil, Emily S. Winn-Deen, Arthur L. Holden, Eric Lai. Linkage Disequilibrium and Inference of Ancestral Recombination in 538 Single-Nucleotide Polymorphism Clusters across the Human Genome. *American Journal of Human Genetics* 73, 285-300 (2003).

[16] K.L. Britt et al. Key steps for effective breast cancer prevention *Nat. Rev. Cancer* (2020)

[17] R.L. Siegel et al. *Cancer statistics, 2020 CA: a Cancer J. Clin.* (2020)

[18] A. Gucalp et al. Male breast cancer: a disease distinct from female breast cancer *Breast Cancer Res. Treat.* (2019)

[19] 耿瑞杰, 姚琳, 黄欣欣, 禹顺英, 苑成梅, 洪武, 吕钦谕, 王庆中, 易正辉, 方贻儒. 基于加权基因共表达网络分析识别抑郁症的差异表达基因模块[J]. *上海交通大学学报(医学版)*, 2021, 41(6): 724-731.

[20] 马燕如, 季林华, 童天颖, 严宇青, 沈超琴, 张昕雨, 曹颖颖, 洪洁, 陈豪燕. 基于单细胞RNA测序的结直肠癌预后预测模型的建立和验证[J]. *上海交通大学学报(医学版)*, 2021, 41(2): 159-165.

[21] Allinen, M. et al. Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell* 6, 17-32 (2004).

[22] Sjöblom T, J. S. , Wood LD, Parsons DW, Lin J, Barber TD, Mandelker D, Leary RJ, Ptak J, Silliman N, Szabo S, Buckhaults P, Farrell C, Meeh P, Markowitz SD, Willis J, Dawson D, Willson JK, Gazdar AF, Hartigan J, Wu L, Liu C, Parmigiani G, Park BH, Bachman KE, Papadopoulos N, Vogelstein B, Kinzler KW, Velculescu VE. The consensus coding sequences of human breast and colorectal cancers. *Science* 314, 268-274 (2006).

- [23] Sorlie, T. et al. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. *Proc Natl Acad Sci U S A* 98, 10869-74 (2001)
- [24] Christos Sotiriou, S. -Y. N. , Lisa M. McShane, Edward L. Korn, Philip M. Long, Amir Jazaeri, Philippe Martiat, Steve B. Fox, Adrian L. Harris, Edison T. Liu. Breast cancer classification and prognosis based on gene expression profiles from a population-based study. *Proc Natl Acad Sci U S A* 100, 10393-8 (2003).
- [25] Jack X Yu, A. M. S. , Yi Zhang, John WM Martens, Marcel Smid, Jan GM Klijn, Yixin Wang, John A Foekens. Pathway analysis of gene signatures predicting metastasis of node-negative primary breast cancer. *BMC Cancer* 7, 182 (2007).
- [26] Yao, J. et al. Combined cDNA array comparative genomic hybridization and serial analysis of gene expression analysis of breast tumor progression. *Cancer Res* 66, 4065-78 (2006).
- [27] Frank Courjal, M. C. , Joelle Simony-Lafontaine, Genevieve Louason, Paul Speiser, Robert Zeillinger, Carmen Rodriguez, and Charles Theillet. Mapping of DNA Amplifications at 15 Chromosomal Localizations in 1875 Breast Tumors: Definition of Phenotypic Groups. *cancer research* 57, 4360-4367 (1997).
- [28] HAN C, FU Y, ZENG N, et al. LncRNA FAM83H-AS1 promotes triple-negative breast cancer progression by regulating the miR-136-5p
- [29] PATAERA, OZPOLATB, SHAOR, et al. Therapeutic targeting of the PI4K2A/PKR lysosome network is critical form is folded protein clearance and survival in cancer cells [J].*Oncogene*, 2020, 39(4),801-813.
- [30] MENDESC, LOPES-COELHO, RAMOSC, et al. Unraveling FATP1, regulated by ER- β , as at argeted breast cancer innovative therapy [J].*Scientific Reports*,2019,9(1):14107
- [31] Bray, F., et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*, 2018.
- [32] DeSantis, C.E., et al., Breast cancer statistics, 2019. *CA Cancer J Clin*, 2019.
- [33] Telang, N., Putative cancer-initiating stem cells in cell culture models for molecular subtypes of clinical breast cancer. *Oncol Lett*, 2015.
- [34] Zhang, J., et al., MicroRNA-138 modulates metastasis and EMT in breast cancer cells by targeting vimentin. *Biomed Pharmacother*, 2016.
- [35] Shimomura, A., et al., Novel combination of serum microRNA for detecting breast cancer in the early stage. *Cancer Sci*, 2016.

- [36] Yen, M.C., et al., S100B expression in breast cancer as a predictive marker for cancer metastasis. *Int J Oncol*, 2018.
- [37] Tian, Z., et al., An immune-related prognostic signature for predicting breast cancer recurrence. *Cancer Med*, 2020.
- [38] N.Lynn, M., et al., Role of Patient and Disease Factors in Adjuvant Systemic Therapy Decision Making for Early-Stage, Operable Breast Cancer: Update of the ASCO Endorsement of the Cancer Care Ontario Guideline. *J Clin Oncol*, 2019.
- [39] Fabrice, N., et al., Use of Biomarkers to Guide Decisions on Adjuvant Systemic Therapy for Women With Early-Stage Invasive Breast Cancer: ASCO Clinical Practice Guideline Update—Integration of Results From TAILORx. . *J Clin Oncol*, 2019.
- [40] Nofisat, K., et al., Biomarkers for Adjuvant Endocrine and Chemotherapy in Early-Stage Breast Cancer. *American Society of Clinical Oncology*, 2022.
- [41] [1] OLIVARES-URBANO M A, GRIÑÁN-LISON C, ZURITA M, etal. Matrix metalloproteases and TIMPs as prognostic biomarkers in breast cancer patients treated with radiotherapy: A pilot study [J]. *Journal of Cellular and Molecular Medicine*, 2020, 24(1):139-148.
- [42] ANASTASIADI Z, LIANOS G D, IGNATIADOU E, etal. Breast cancer in young women: anover view[J]. *Updates in Surgery*, 2017, 69(3):313-317.
- [43] DIABY V, TAWK R, SANOGO V, etal. A review of systematic reviews of the cost-effectiveness of hormone therapy, chemotherapy and targeted therapy for breas tancer [J]. *Breast Cancer Researchand Treatment*, 2015,151(1):27-40.

ХиБИ кафедрасының студенттерінің Байжума Ж. Б., Уалихан А.С.

**ДИПЛОМДЫҚ ЖОБАСЫНА
ҒЫЛЫМИ ЖЕТЕКШІНІҢ ПІКІРІ**

Мамандық: 6B05101– «Химиялық және биохимиялық инженерия»

Тақырып: «Биоинформатикалық құралдарды пайдалана отырып сүт безі қатерлі ісігіндегі негізгі гендердің экспрессиясын бағалау»

Дипломдық жұмыстың негізгі ескертулері және студенттің сипаттамасы

Байжума Жұлдыз және Уалихан Алинаның дипломдық жұмысы бүгінгі күннің маңызды мәселелерінің бірі – «Биоинформатикалық құралдарды пайдалана отырып сүт безі қатерлі ісігіндегі негізгі гендердің экспрессиясын бағалау» тақырыбын зерттеуге арналған. Қазіргі уақытта сүт безі қатерлі ісігі бүкіл әлемдегі әйелдердің денсаулығына айтарлықтай зиян келтіруде, бұл оны диагностикалау мен емдеудің жаңа әдістерін терең зерттеуді және қолдануды талап етеді.

Байжума Жұлдыз және Уалихан Алина биоинформатикалық әдістерді пайдалана отырып, сүт безі қатерлі ісігіндегі негізгі гендердің экспрессиясын талдау бойынша эксперимент жүргізді. Эксперименттік бөлімді орындау барысында RNA-Seq арқылы алынған геномдық және транскриптомдық деректер зерттеліп, арнайы бағдарламалық қамтамасыз етуді пайдалана отырып талданды. Зерттеу мәліметтерді өңдеу және визуализациялау, сондай-ақ алынған нәтижелердің биологиялық интерпретациясын қамтыды. Қойылған мақсаттар мен міндеттер зерттеу тақырыбына толық сәйкес келеді.

Дипломдық жұмыс қорытындысы бойынша студент теориялық және практикалық материалдармен жұмыс істеудің жақсы дағдыларын көрсетті. Жұмыс жазу барысында күнтізбелік кестені сақтап, орындаушылық және тәртіптілік, жауапкершілік және қызығушылық сияқты қасиеттерін көрсетті.

Дипломдық жұмысты бағалау

Байжума Жұлдыздың және Уалихан Алинаның дипломдық жұмысы теориялық және практикалық деңгейде жақсы орындалған, осындай жұмыстарға қойылатын талаптарға сәйкес келеді, қорғауға ұсынылады және сәтті қорғаған жағдайда жоғары бағаға (98%) лайық.

Ғылыми жетекші

Ботбаев Д.М. магистр.

«06» 06 2024 ж.



Дипломдық жобаға

РЕЦЕНЗИЯ

Байжума Жұлдыз Багдадқызы, Уалихан Алина Серікқызы
6B05101–«Химиялық және биохимиялық инженерия» мамандығы

Тақырыбы: «Биоинформатикалық құралдарды пайдалана отырып сүт безі қатерлі ісігіндегі негізгі гендердің экспрессиясын бағалау»

Орындалды:

- а) графикалық бөлім 10 парақ;
- б) түсініктеме 44 бет.

ЖОБАҒА ЕСКЕРТУ

Байжума Жұлдыз және Уалихан Алина студенттерінің дипломдық жұмысы биоинформатикалық құралдарды пайдалана отырып, сүт безі обырындағы негізгі гендердің экспрессиясын бағалауға арналған қызықты зерттеу болып табылады. Авторлар ген экспрессиясының деректерін талдау әдістерін және оларды онкологиялық зерттеулер контекстінде қолдануды мұқият қарастырған.

Мақалада көтерілген маңызды сәттерге ген экспрессиясын талдаудың заманауи әдістеріне шолу, сүт безі қатерлі ісігінің биологиялық ерекшеліктерін сипаттау және оның дамуы мен прогрессиясындағы негізгі гендердің рөлін қарастыру кіреді. Студенттер осы саладағы алдыңғы зерттеулерге сыни талдау жасады және биоинформатикалық құралдарды қолдана отырып, деректерді талдаудың өзіндік тәсілін ұсынады.

Дипломдық жұмыста сүт безі қатерлі ісігімен байланысты қолтаңба гендерін анықтау және оның болжамын болжау үшін машиналық оқыту әдістері мен статистикалық талдау әдістерін қолдануға ерекше назар аударылады. Авторлар осы саладағы болашақ зерттеулердің перспективалық бағыттарын ұсынады, соның ішінде сүт безі қатерлі ісігінің дамуының молекулалық механизмдерін терең зерттеу және диагностика мен емдеудің жаңа әдістерін әзірлеу қамтылған.

Байжума Жулдыз және Уалихан Алина студенттерінің жалпы қорытындысы акпараттық және мазмұнды болып табылады. Олар зерттеу тақырыбы бойынша әдебиеттерге кең шөлу жасады, өз жұмыстарын жақсы құрылымдады және нақты зерттеу нәтижелерін ұсынды. Жұмыс сүт безі қатерлі ісігін зерттеуге құнды үлес болып табылады және авторлардың аналитикалық дағдылары мен кәсібилігінің жоғары деңгейін көрсетеді.

Жұмыстың бағасы

Жалпы, Байжума Жулдыз және Уалихан Алина рецензияланған дипломдық жұмысы қойылатын мемлекеттік стандарттың талаптарына толық сәйкес келеді, қорғауға ұсынылады және сәтті қорғалған жағдайда өте жақсы (98%) бағаға лайық.

Рецензент

Биология ғылымдарының кандидаты, доцент

Әл-Фараби атындағы ҚазҰУ биология және

биотехнология факультетінің

биотехнология кафедрасының профессоры

Асрандина Салтанат Шынтаевна





Метаданные

Название

Биоинформатикалық құралдарды пайдалана отырып сүт безі қатерлі ісігіндегі негізгі гендердің бағалау

Автор

Байжума Жұлдыз, Уалихан Алина

Научный руководитель / Эксперт

Даурен Ботбаев

Подразделение

ИГИНГД

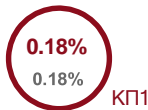
Тревога

В этом разделе вы найдете информацию, касающуюся текстовых искажений. Эти искажения в тексте могут говорить о ВОЗМОЖНЫХ манипуляциях в тексте. Искажения в тексте могут носить преднамеренный характер, но чаще, характер технических ошибок при конвертации документа и его сохранении, поэтому мы рекомендуем вам подходить к анализу этого модуля со всей долей ответственности. В случае возникновения вопросов, просим обращаться в нашу службу поддержки.

Замена букв		3
Интервалы		0
Микропробелы		0
Белые знаки		0
Парафразы (SmartMarks)		2

Объем найденных подоби

КП-ия определяют, какой процент текста по отношению к общему объему текста был найден в различных источниках.. Обратите внимание!Высокие значения коэффициентов не означают плагиат. Отчет должен быть проанализирован экспертом.

**25**

Длина фразы для коэффициента подобия 2

**15186**

Количество слов

**75909**

Количество символов

Подобия по списку источников

Ниже представлен список источников. В этом списке представлены источники из различных баз данных. Цвет текста означает в каком источнике он был найден. Эти источники и значения Коэффициента Подобия не отражают прямого плагиата. Необходимо открыть каждый источник и проанализировать содержание и правильность оформления источника.

10 самых длинных фраз

Цвет текста

ПОРЯДКОВЫЙ НОМЕР	НАЗВАНИЕ И АДРЕС ИСТОЧНИКА URL (НАЗВАНИЕ БАЗЫ)	КОЛИЧЕСТВО ИДЕНТИЧНЫХ СЛОВ (ФРАГМЕНТОВ)	
1	https://stud.kz/referat/show/115800	27	0.18 %

из базы данных RefBooks (0.00 %)



ПОРЯДКОВЫЙ НОМЕР	НАЗВАНИЕ	КОЛИЧЕСТВО ИДЕНТИЧНЫХ СЛОВ (ФРАГМЕНТОВ)
------------------	----------	---


из домашней базы данных (0.00 %)



ПОРЯДКОВЫЙ НОМЕР	НАЗВАНИЕ	КОЛИЧЕСТВО ИДЕНТИЧНЫХ СЛОВ (ФРАГМЕНТОВ)
------------------	----------	---

из программы обмена базами данных (0.00 %)



ПОРЯДКОВЫЙ НОМЕР	НАЗВАНИЕ	КОЛИЧЕСТВО ИДЕНТИЧНЫХ СЛОВ (ФРАГМЕНТОВ)
из интернета (0.18 %)		
ПОРЯДКОВЫЙ НОМЕР	ИСТОЧНИК URL	КОЛИЧЕСТВО ИДЕНТИЧНЫХ СЛОВ (ФРАГМЕНТОВ)
1	https://stud.kz/referat/show/115800	27 (1) 0.18 %

Список принятых фрагментов (нет принятых фрагментов)

ПОРЯДКОВЫЙ НОМЕР	СОДЕРЖАНИЕ	КОЛИЧЕСТВО ИДЕНТИЧНЫХ СЛОВ (ФРАГМЕНТОВ)
------------------	------------	---