



**ПЛОВДИВСКИ УНИВЕРСИТЕТ – ПАИСИЙ
ХИЛЕНДАРСКИ**



БИОЛОГИЧЕСКИ ФАКУЛТЕТ

КАТЕДРА ФИЗИОЛОГИЯ НА РАСТЕНИЯТА И МОЛЕКУЛЯРНА БИОЛОГИЯ

РЪКОВОДИТЕЛ: ПРОФ. Д-Р ВАЛЕНТИНА ТОНЕВА

Гл. ас. Д-р ТИХОМИР ИЛИЕВ ВЪЧЕВ

**СРАВНИТЕЛНИ ГЕНОМНИ, ТРАНСКРИПТОМНИ И
ПРОТЕОМНИ ИЗСЛЕДВАНИЯ ПРИ НЕВРОРАЗВИТИЙНИ
РАЗСТРОЙСТВА**

АВТОРЕФЕРАТ

на дисертационен труд за присъждане на научна степен “ДОКТОР НА НАУКИТЕ”

Област на висше образование: 4, „Природни науки, математика и информатика”

Професионално направление: 4.3. „Биологически науки”

Научна специалност: „Молекулярна биология”

ПЛОВДИВ 2020

Дисертационният труд е обсъден и насочен за защита на заседание на разширен Катедрен съвет на Катедра “Физиология на растенията и молекулярна биология” към Биологически факултет на Пловдивски Университет “Паисий Хилендарски”. Изследванията по дисертационния труд са финансирани в рамките на проекти “Анализ на експресионния профил на небелтък кодиращи – микро РНК гени при шизофрения” № НО-4/2011; „Изследване на серумните нива на микро РНК биомаркерни молекули при деца с аутизъм с цел създаване на платформа за целите на неинвазивната молекулярна диагностика на аутистично болни“ № НО-12/2013; „Изследване на метилационния статус на кандидат гени асоциирани с разстройството от аутистичния спектър с цел изясняване на епигенетичните механизми участващи в патогенезата му“ № НО-09/2014 “Анализ на експресионния профил на небелтък кодиращи – микро РНК гени при шизофрения” 2011.

Публичната защита ще се състои на 26. 02. 2021 г., от 10 часа в 15 аудитория на Биологически факултет, ул. „Тодор Самодумов“, № 2, гр.Пловдив.

Научно жури в състав:

Вътрешни членове:

Проф. д-р Балик Маломиров Джамбазов

Доц. д-р Еленка Стоилова Георгиева

Външни членове:

Акад. д.б.н Иван Георгиев Иванов

Проф. д.б.н Антон Божидаров Тончев

Проф. д.т.н Алберт Иванов Кръстанов

Проф. д.м.н Надежда Петрова Маджирова

Доц. д-р Весела Иванова Чалова-Жекова

Материалите по защитата са на разположение в библиотеката на Пловдивски Университет “Паисий Хилендарски” и са публикувани на интернет страницата на университета.

I.	СЪДЪРЖАНИЕ.....	3
II.	ВЪВЕДЕНИЕ.....	7
III.	ЦЕЛ И ЗАДАЧИ.....	9
IV.	МАТЕРИАЛИ И МЕТОДИ.....	10
IV.1.	Участници в проучването (пациенти с диагноза шизофрения и контролна група от здрави лица).....	10
IV.2.	Участници в проучването (пациенти с диагноза РАС и контролна група от здрави деца).....	11
IV.3.	Участници и клинична оценка (пациенти с диагноза шизофрения и контролна група от здрави лица).....	12
IV.4.	Вземане на венозна кръв.....	12
IV.5.	Засекретяване на пробите.....	12
IV.6.	Екстракция на тотална РНК от периферна кръв с помощта на <i>PAXgene Blood miRNA Kit</i>	12
IV.7.	Екстракция на тотална ДНК от периферна кръв с помощта на QIAmp DNA Blood Mini Kit Spin протокол (Qiagen).....	13
IV.8.	Хоризонтална агарозна гел-електрофореза на РНК.....	13
IV.9.	Количествен анализ на изолираната РНК.....	13
IV.10.	Изготвяне на сборни РНК проби (пулове) и преципитация на РНК образците.....	13
IV.11.	In silico дизайн и анализ на праймерни двойки.....	14
IV.12.	Провеждане на обратна транскрипция на миРНК молекули.....	14
IV.13.	Синтез на комплементарни ДНК вериги.....	14
IV.14.	Предварителна амплификация (преамплификация) на получените миРНК специфични комплементарни ДНК последователности.....	14
IV.15.	Количествен PCR експресионен анализ на миРНК молекули.....	15
IV.16.	Количествен PCR анализ (Quantitative Reverse Transcription Real-Time PCR) на миРНК молекули.....	15
IV.17.	Провеждане на qRT-PCR (Fermentas SYBR Green PCR Kit/Thermo) анализ.....	16
IV.18.	Провеждане на обратна транскрипция на протеин-кодиращи гени.....	16
IV.19.	Провеждане на qRT-PCR (Fermentas SYBR Green PCR Kit (Thermo)).....	17
IV.20.	Анализ на експресионният профил на малки РНК молекули от периферна кръв при РАС (Small RNA sequencing).....	17

IV.21. Анализ на експресионния профил на малки РНК молекули (Small RNA sequencing).....	17
IV.22. Биоинформатичен анализ на миРНК експресионния профил чрез секвениране на малки РНК молекули от периферна кръв при РАС.....	17
IV.23. Анализ на експресионния профил на информационни РНК молекули (RNA Sequencing).....	18
IV. 24. Секвениране на РНК (Транскриптомен анализ - Методика).....	18
IV. 25. Генно онтологичен анализ на диференциално експресирани гени	21
IV.26. KEGG анализ на биологичните пътищата обогатени с диференциално експресирани гени.....	21
IV.27. Количествен протеомен анализ – Isobaric Tag for Relative and Absolute Quantification (ITRAQ) при РАС.....	22
IV.28. Мащабно екзомно секвениране (Whole-Exome Sequencing) за изследване на мутации в белтък-кодиращи последователности в човешкия геном асоциирани с РАС.....	23
IV.29. Идентифициране на мутации при РАС с помощта на секвениране на РНК от следващо поколение (NGS RNA sequencing)...	24
IV.30. Транскриптомно профилиране на белтък-кодиращи гени в периферна кръв при шизофрения (Digital Gene Expression Tag Profiling (DGE)).....	25
IV. 31. Потвърждаване на диференциално експресирани гени от анализ чрез провеждане на количествен Real Time PCR анализ.....	25
V. РЕЗУЛТАТИ.....	25
V.1. Резултати от проведеното мащабно екзомно секвениране (<i>Whole-Exome Sequencing</i>) на белтък-кодиращи последователности в човешкия геном асоциирани със РАС.....	25
V. 2. Биоинформатичен анализ на ITRAQ протеомни данни.....	29
V.3. Сравнителен анализ на диференциално експресирани протеини от проведения количествен протеомен анализ – Isobaric Tag for Relative and Absolute Quantification (ITRAQ).....	30

V.4. Анотация по клъстери от ортоложни групи от протеини (Clusters of Orthologous Groups of Proteins (COG)).....	33
V. 5. Генно онтологичен (GO) анализ на диференциално експресирани протеини.....	33
V.6. Резултати от проведеният анализ на експресионния профил на малки РНК молекули от периферна кръв при РАС.....	34
V.7. Идентифициране на диференциално експресирани анотирани миРНК молекули в анализирани групи чрез секвениране на малки РНК молекули.....	36
V.8. Резултати от проведеният анализ на миРНК експресионния профил чрез секвениране на малки РНК молекули при РАС.....	36
V.9. Индивидуален експресионен анализ на миРНК молекули в серум от пациенти с РАС.....	37
V.10. Резултати от проведените изследвания на кандидат миРНК биомаркерни молекули в периферна кръв при деца с диагноза РАС.....	42
V.11. Идентифициране на биологични пътища свързани с диференциално експресирани миРНК молекули при РАС.....	43
V.12. Идентифициране на диференциално експресирани белтък-кодиращи гени в анализирани групи чрез секвениране на РНК (RNA-Sequencing).....	46
V.13. Генно онтологичен анализ на диференциално експресирани гени при РАС.....	47
V.14. Потвърждаване на диференциално експресирани гени от транскриптомния (RNA-Seq) анализ чрез провеждане на количествен Real Time PCR анализ.....	48
V.15. Резултат от количественият експресионен анализ (RT-qPCR) на гените (AGO2, AGO3, AGO4 и Drosha) при РАС.....	49
V.16. KEGG анализ на биологичните пътищата обогатени с диференциално експресирани гени при РАС.....	50
V.17. Идентифициране на нова митохондриална мутация в гена, кодиращ цитохром с III оксидаза (CO3) при деца с РАС с помощта на секвениране на РНК от следващо поколение (NGS RNA sequencing).....	51

V.18. Резултати от проведеното транскриптомно профилиране на белтък-кодиращи гени в периферна кръв при шизофрения.....	52
V.19. ROC анализ (receiver operating characteristic) за определяне на диагностична точност на изследваните диференциално експресирани гени при шизофрения.....	54
V.20. ROC анализ (receiver operating characteristic) за определяне на диагностична точност на изследваните диференциално експресирани гени при шизофрения.....	54
V.21. ROC анализ (receiver operating characteristic) за определяне на диагностична точност на изследваните диференциално експресирани гени при шизофрения.....	56
VI. ДИСКУСИЯ.....	57
VI.1. Обсъждане на резултатите получени от проведеното мащабно екзомно секвениране на ДНК (<i>Whole-Exome Sequencing</i>) за изследване на белтък-кодиращи последователности в човешкия геном асоциирани със PAC.....	59
VI.2. Обсъждане на резултатите получени от проведените експресионни изследвания на мРНК при PAC.....	60
VI.3. Обсъждане на резултатите от проведения количествен протеомен анализ – Isobaric Tag for Relative and Absolute Quantification (ITRAQ) при PAC.....	63
VI.4. Обсъждане на резултатите получени от проведените транскриптомни (RNA-Seq) изследвания на белтък-кодиращи гени при PAC.....	64
VI.5. Обсъждане на резултатите получени от ген онтологичния анализ на диференциално експресирани гени и KEGG анализ на биологичните пътищата.....	65
VI.6. Обсъждане на резултатите от сравнителният експресионен анализ (DGE-tag profiling) при шизофрения.....	67
VII. ЗАКЛЮЧЕНИЕ.....	67
VIII. ПРИНОСИ.....	69

II. ВЪВЕДЕНИЕ

Стотици милиони хора по света страдат от психични заболявания. Световната Здравна Организация (СЗО) оценява, че 25% от времето, което европейските служители прекарват като отпуск поради болест, е причинено от психични разстройства. Това включва инвалидизиращи заболявания като шизофрения, разстройство от аутистичния спектър (РАС), специфично нарушение на развитието на речта (СНРР) и др, които нанасят огромна тежест: загуба на продуктивност на засегнатите пациенти, роднините и здравната система. В много от случаите тези разстройства и заболявания представляват трудност за диагностициране, като изискват много и различни подходи и протоколи както за диагностика, така и за лечение в хода на заболяването. За да се справи с клиничните нужди, които възникват при диагностицирането на такъв тип заболявания, съвременната медицина е в търсене на специфични диагностични и прогностични биомаркери при заболяванията, които могат коректно да идентифицират заболяването посредством специфични модел на кандидат-биомаркерни молекули.

Благодарение на съвременни напредък в областта на молекулярната генетика и биотехнологии, развитието и прилагането през последните години на нови технологии за секвениране от следващо поколение като транскриптомни и протеомни анализи, позволи откриването на голям брой молекулярни биомаркери, в това число малки РНК молекули, белтъккодиращи информационни РНК молекули, ДНК специфични варианти и протеини и метаболити. Всичко това разшири обхвата от потенциалните биомаркери, включвайки генетично профилиране, транскриптомен и протеомен анализ. Въпреки привидното предположение, че единственият подходящ материал за изследване и валидиране на биомаркери асоциирани с невроразвитийните разстройства е мозъчната тъкан, едно ново направление в търсене на биомаркери при тези заболявания е профилирането на експресията на небелтък кодиращи регулаторни гени като миРНК както и прилагането на транскриптомен и протеомен анализ, които показват висока чувствителност и специфичност в тъкани различни от ЦНС. МиРНК молекулите са малки регулаторни РНК молекули, които показват промяна в експресионния профил при редица заболявания. Ролята на миРНК молекулите и предимството, което те предлагат при асоциирането на специфични експресионни промени със съответно заболяване подчертава тяхната стойност, като потенциални молекулярни биомаркери. Редица изследвания демонстрират ролята на миРНК молекулите и по специално комплексни миРНК експресионен профил, който може да бъде използван за диагностика на заболявания. Специфичният миРНК профил показва висока диагностична стойност и анализът от телесни течности, като периферна

кръв предлага възможност за провеждането на относително неинвазивен диагностичен тест. МиРНК молекулите като потенциални биомаркери притежават обективно измерими биологични характеристики, които могат да бъдат използвани като индикатори за нормални или патологични процеси: миРНК профила служи като високо специфичен маркер за диагностициране, прогнозиране, проследяване на развитието на болестта и предвиждане на терапевтичният отговор. Периферните кръвни клетки в качеството им на динамично рефлектираща система позволяват профилирането на миРНК експресията, което практически прави възможно използването им като потенциален обект за изследване. От тази гледна точка миРНК молекулите се разглеждат като потенциален нов клас биомаркери при редица заболявания, включително разстройства на централната нервна система (ЦНС). Приложението на тези открития в клиничната практика, ни доближава до персонализираната медицина, като подобрява грижите за пациентите, увеличавайки капацитета на диагностичните и терапевтичните процедури.

МиРНК молекулите представляват надеждни биомаркери при невропсихиатрични заболявания поради тяхната значителна стабилност. Тези аспекти на миРНК молекулите позволяват тяхната детекция не само от телесни тъкани, а така също и от телесни флуиди като кръв (както от свободно циркулиращи нуклеинови киселини в кръвната плазма и/или серум, така и от ядрени кръвни клетки). Докато диагностичната стойност на единични биомаркери е ограничена поради липсата на чувствителност и специфичност, общият миРНК профил получен от мултиплексен анализ с използването на платформи за секвениране на малки РНК молекули, характеризиращи се с висока информативност и специфичност, като прави възможно, както идентифицирането на индивидуалният миРНК, така и пълният миРНК експресионен профил с висока степен на точност. В последните години обаче стана ясно, че подходите свързани с изучаването на миРНК молекули са много подходящи при изследване на етиологията и патогенезата на мултифакторните заболявания, към които спадат невроразвитийните разстройства като шизофрения и РАС. В България все още не са провеждани експресионни изследвания на миРНК молекули в областта на молекулярната психиатрия, което допълнително аргументира и провеждането на подобен род изследвания и у нас.

Протеомиката е друг потенциален подход, който може да генерира нови хипотези от съществено значения при патогенезата и диагностиката на заболяванията чрез идентифициране на кандидат белтъчни биомаркери. Протеомните подходи допълват генетичните и геномните изследвания и така да фокусират вниманието върху белтъчните продукти, които са свързани с генетични аномалии, свързани с невродегенеративните и невроразвитийните разстройства.

III. ЦЕЛ И ЗАДАЧИ

Цел

Провеждане на сравнителни геномни, транскриптомни и протеомни изследвания при невроразвитийни разстройства.

Задачи

1. Провеждане на сравнителен експресионен анализ на белтък-кодиращи гени (RNA Sequencing) при РАС.
2. Провеждане на сравнителен експресионен анализ на малки РНК молекули (Small RNA sequencing) при РАС.
3. Провеждане на сравнителен протеомен анализ – Isobaric Tag for Relative and Absolute Quantification (ITRAQ) при РАС.
4. Провеждане на мащабно екзомно секвениране (*Whole Exome Sequencing*) при РАС.
5. Провеждане на сравнителен експресионен анализ на белтък-кодиращи гени (Digital Gene Expression) при шизофрения.
6. Провеждане на сравнителен експресионен (qRT-PCR) анализ на белтък-кодиращи гени при РАС.
7. Провеждане на сравнителен експресионен (qRT-PCR) анализ на миРНК при РАС.
8. Провеждане на сравнителен експресионен (qRT-PCR) анализ на белтък-кодиращи гени при шизофрения.
9. Провеждане на ROC (Receiver Operating Characteristic) анализ на диференциално експресирани белтък-кодиращи гени и миРНК молекули при РАС и шизофрения.

IV. МАТЕРИАЛИ И МЕТОДИ

IV.1. Участници в проучването (пациенти с диагноза шизофрения и контролна група от здрави лица)

Дизайнът на настоящото проучване и формата за информирано съгласие бяха одобрени от комисията по научна етика към Медицинския университет Пловдив. Проведено беше Mini-International Neuropsychiatric Interview и рутинен психиатричен преглед, при който бяха получена анамнестични данни за заболяването. Интервюто беше извършено от сертифициран психиатър, за оценка на диагнозата параноидна шизофрения на базата на критериите заложи в Diagnostic and Statistical Manual of Mental Disorders forth text revised edition (DSM IV TR) (Диагностичен и статистически наръчник на психичните разстройства - четвърто преработено издание) и с цел изключване на каквито и да е други психични разстройства при контролите. При всички пациенти, включени в изследването е поставена консенсусна диагноза - шизофрения съгласно международната система за диагностика DSM-IV (American Psychiatric Association. 1994. Diagnostic and Statistical Manual of mental disorders, 4thed.).

Основната част от пациентите за настоящето проучване е набрана в рамките на проект "Анализ на експресионния профил на небелтък кодиращи – микро РНК гени при шизофрения" № НО-4/2011; „Изследване на серумните нива на микро РНК биомаркерни молекули при деца с аутизъм с цел създаване на платформа за целите на неинвазивната молекулярна диагностика на аутистично болни“ № НО – 12/2013; и съответно проект "Аутизъм и специфично нарушение на развитието на речта – сравнителни геномни изследвания за идентифицирането на обща патогенетична основа " № ДФНИ-Б01/21 към Катедрата Педиатрия и Медицинска генетика, Медицински университет – Пловдив:

Групата анализирани индивиди включени в насоящата експериментална работа обхваща 33 пациента с диагноза шизофрения и група от 25 здрави лица-контроли. Средната възраст на изследваните пациентите е 45,6 год. (SD = 12.565). В контролната група средната възраст е 46,3 год. (SD = 12.673).

Всички участници, включени в настоящото изследване са от български произход. Пациентите с диагноза шизофрения включени в настоящото изследване са пациенти на ДПБ - Пазарджик. Пациентите са набирани от лекуващите ги лекари на базата на цялата достъпна информация за пациента, включително епикризи, картони на пациента.

IV.2. Участници в проучването (пациенти с диагноза РАС и контролна група от здрави деца)

В представеното проучване са включени 40 деца с генерализирано разстройство на развитието, на възраст от 3 до 11 години (средна възраст на извадката е 6,86). За целите на проучването са използвани диагностичните критерии на DSM IV-TR и на пациентите е поставена диагноза съобразно класификацията на Американската Психиатрична Асоциация. За потвърждаване на диагноза РАС при всеки пациент е извършено подробно психиатрично изследване с прилагане на ADI-R, CARS и GARS. След публикуването на DSM 5 диагнозата РАС е препотвърдена съгласно новите критериите. В контролната група са включени 30 здрави деца на възраст от 3 до 11 години (средна възраст на извадката е 6,43), след подписване на информирано съгласие от родителите.

При подбор на извадките са съобразени следните критерии за изключване:

1. Епилепсия
2. ДЦП, дори без установена с образна находка мозъчна структурна лезия.
3. Наличие на огнищна неврологична симптоматика (насочва мозъчна структурна лезия)
4. Доказана с образно изследване (ТФЕ, КТ, МРТ) мозъчна структурна лезия, дори без наличие на ДЦП или огнищна неврологична симптоматика.
5. Суспектни анамнестични или клинични данни за хронично соматично, неврологично или друго психиатрично заболяване.
6. Наличието на остро инфекциозно заболяване (Острото инфекциозно заболяване обикновено води до левкоцитоза, т.е. по-голям обем левкоцитна маса, т.е. промяна на количеството миРНК).
7. Фамилна обремененост с наследствени метаболитни заболявания
8. Наличието на фамилност със шизофрения или друга психиатрична патология
9. Отказ за участие в проучването.
10. Наличие на умствено изоставане (за контролната група децата).
11. Билингвисти

На всички родителите, изразили желание децата им да участват в проучването, е предоставено информирано съгласие, одобрено от комисията по етика към МУ-Пловдив, след заявяване на изрично съгласие в писмена форма, пациентите бяха оценени според включващите и изключващите критерии заложи в настоящото изследване. Основната част от пациентите включени в настоящото изследване са набрани в рамките на проект: *“Аутизъм и специфично нарушение на развитието на речта – сравнителни геномни изследвания за идентифицирането на обща патогенетична основа”* № ДФНИ-Б01/21 и проект *„Изследване на серумните нива на микро РНК биомаркерни молекули при деца с аутизъм с цел създаване на платформа за целите на неинвазивната молекулярна диагностика на аутистично болни“* № НО – 12/2013.

IV.3. Участници и клинична оценка (пациенти с диагноза шизофрения и контролна група от здрави лица)

След подписване на удобен от Етичната комисия на МУ-Пловдив формуляр за Информирано Съгласие бяха набрани тридесет пациента с шизофрения, хоспитализирани в Държавна психиатрична болница – Пазарджик, България. Всички подходящи пациенти бяха интервюирани от сертифициран по MINI изследовател за потвърждаване на диагнозата шизофрения по критериите на DSM IV TR. Диагноза беше определяна от двама независими психиатри, използвайки цялата налична информация, включително болничните записи и стандартното психиатрично клинично интервю. Клиничните симптоми на всеки пациент са оценени по скалата за позитивни и негативни симптоми – PANSS и скалата за общо клинично впечатление за тежест - CGI-S от сертифициран психиатър, използвайки SCI - PANSS интервюто. Основният критерий за включване е, че пациента не е приемал лекарствата си поне 2 седмици. След подписването на формуляра за информиране на съгласието, беше събрана контролната група от здрави доброволци, съответна по възраст и пол на пациентите. Участниците в контролната група бяха интервюирани от клиницист за значими симптоми и фамилна анамнеза за роднини до втора степен с психиатрична диагноза/симптоматика по ос I според критериите на DSM IV TR.

IV.4. Вземане на венозна кръв

При вземането на венозна кръв от пациентите включени в настоящото проучване се използва тънка игла – венкат, „бътерфлайка“ и PAXgene RNA Tube моновета.

IV.5. Засекретяване на пробите

След вземането на кръвта пробите бяха засекретени съответно с изискванията на стандартите на етичната комисия, като пробите получаваха съответно буквен код на пациента и пореден номер. Достъп до информацията имаха единствено лицата, включени в провеждането на проучването, а конфиденциалността бе спазена според изискванията на ИС и етичните норми.

IV.6. Екстракция на тотална РНК от периферна кръв с помоща *PAXgene Blood miRNA Kit*

Екстракцията на тотална РНК от периферна кръв беше проведена с помоща на *PAXgene Blood miRNA Kit*. Технологията на *PAXgene Blood miRNA Kit*

представлява добре утвърдена методика за пречистване (екстракция) на тотална РНК от венозна кръв, предварително изолирана в PAX gene RNA Blood моновета.

IV.7. Екстракция на тотална ДНК от периферна кръв с помощта на QIAamp DNA Blood Mini Kit Spin протокол (Qiagen)

Принцип и процедура

Китът QIAamp DNA Blood Mini е предназначен за бързо изолиране и пречистване на средно 6 μg тотална ДНК (геномна, вирусна, митохондриална) от 200 μl тотална човешка кръв и до 50 μg ДНК от 200 μl лимфоцитна фракция, 5×10^6 пречистени лимфоцити или култивирани клетки с нормален хромозомен набор. Процедурата е подходяща за екстракция от тотална кръв, третирана с цитрат, хепарин или EDTA; лимфоцити; плазма; серум и други телесни течности.

IV.8. Хоризонтална агарозна гел-електрофореза на РНК

Качествена оценка за наличието и степента на деградацията на РНК е направена с помощта на хоризонтална агарозна гел-електрофореза в 1% агарозен гел с етидиев бромид. Визуализацията на РНК беше проведена върху еднопроцентна агарозна TBE гел - електрофореза с използването на 0,5X TBE буферна разделяща система (50mM трис, 50mM борна киселина, 1mM EDTA, pH 8.3) при напрежение 4,5 V/cm² и визуализиращ агент етидиев бромид.

IV.9. Количествен анализ на изолираната РНК

Спектрофотометричното определяне на концентрацията на изолираната РНК беше проведено с помощта на Epoch Micro-Volume Spectrophotometer System (BioTec).

IV.10. Изготвяне на сборни РНК проби (пулове) и преципитация на РНК образците

Общите РНК пулове бяха приготвени като еднакво количество РНК от всеки един пациент се поставя в една обща епруветка (пул от шизофренно болни). По идентичен начин, еднакво количество от всяка една контролна РНК проба беше поставено в друга епруветка (пул от контролна група) за получаването съответно на сборна контролна проба.

За успешното транспортиране и съхранение на изолираната РНК пробите бяха подложени на преципитация както следва:

Добавяне на 3 обема абсолютен етанол;

Добавяне на 1/10 обем 3М NaOAc, рН 5.2.

След създаването на двете комплексни РНК проби, образците бяха транспортирани до централата на BGI - Пекин в термостабилни опаковки на сух лед (-78°C) за поддържане на минимална температура, съгласно изискванията на лабораторията за провеждане на транскриптомен анализ на малки РНК както и за провеждане на транскриптомен анализ на белтък кодиращи гени (RNA Seq).

IV.11. In silico дизайн и анализ на праймерни двойки

В настоящата работа беше използван специфичен софтуер за диазайн на праймери, като PrimerExpress® или Primer3 (<http://primer3.ut.ee/>).

IV.12. Провеждане на обратна транскрипция на миРНК молекули

За провеждане на обратна транскрипция на миРНК молекули беше използван кит Maxima® First Strand cDNA Synthesis Kit на “Thermo scientific”.

IV.13. Синтез на комплементарни ДНК вериги

Реакционните смеси, необходими за реакциите на синтез на комплементарни ДНК последователности, могат да бъдат приготвени чрез смесване на агентите поотделно за всяка РНК или в обща смес (master mix), съдържаща всички компоненти за реакцията на обратна транскрипция освен РНК матриците.

IV. 14. Предварителна амплификация (преамплификация) на получените миРНК специфични комплементарни ДНК последователности

Преамплификацията е често необходима стъпка за генериране на

достатъчен брой копия на целеви ДНК молекули, които са необходими за прецизна последваща амплификация. Преамплификация на копии ДНК молекули специфични за анализиранияте миРНК беше проведена с помощта на *reqGOLDTaq DNA Polymerase*.

IV. 15. Количествен PCR експресионен анализ на иРНК молекули

В настоящото изследване е използван вариант със свързващо се към ДНК зеленото флуоресцентно багрило (SYBR Green I) SYBR Green I е интеркалиращ агент, който се свързва към малката бразда на двойно верижна ДНК молекула. Багрилото не се свързва към едноверижни ДНК молекули. В разтвор SYBR Green I флуоресцира слабо, но започва да излъчва силен флуоресцентен сигнал след свързване с двуверижна ДНК.

В изследването беше използван методът на относителния количествен анализ, при който се получават данни за степента на изменение в генната експресия на изследваните елементи в сравнение с референтен ген актин В.

IV. 16. Количествен PCR анализ (Quantitative Reverse Transcription Real-Time PCR) на миРНК молекули

Методът на полимеразна верижна реакция в реално време Real-Time qRT-PCR (Quantitative Reverse Transcription Real-Time PCR) може да бъде използван за количествен, полуколичествен и качествен анализ. Методът позволява надеждна детекция и измерване на генерираните PCR продукти при всеки цикъл от процеса на полимеразна верижна реакция.

При провеждането на настоящото изследване беше използвано багрило SYBR Green, което е от първия тип – свързва се към голямата бразда на двуверижната ДНК. Съответно с увеличаване количеството на ампликона, се засилва флуоресценцията, която се отчита след всеки цикъл. Но багрилата от този тип могат да се свържат към всички двуверижни PCR продукти, включително към неспецифични такива (например праймерни димери).

Количественият PCR анализ беше извършен с помощта на *Maxima Hot Start Taq DNA Polymerase* е *Taq* ДНК-полимераза, която е химически модифицирана, чрез добавяне на термолабилни блокиращи групи. Ензимът е неактивен при

стайна температура, което предотвратява елонгацията на неспецифично свързани праймери или на праймерни димери и осигурява по-висока специфичност на ДНК амплификацията. Ензимът осигурява удобство при подготовката на реакциите на стайна температура. Реакционният буфер е специфично оптимизиран за qPCR анализ с използването на SYBR Green I. Той съдържа както KCl и $(\text{NH}_4)_2\text{SO}_4$, за да се обезпечи високата специфичност на свързването на праймерите. Максимумите на възбуждане и излъчване на SYBR Green I са съответно при 494 nm и 521 nm, които са съвместими с всеки Real-time PCR апарат.

Референтното багрило – ROX Passive reference dye – е включено в реакционната смес на кита, за да служи за вътрешна нормализация на флуоресцентния сигнал на SYBR Green I, когато се използва машина, която може да детектира ROX, напр. Applied Biosystems 7500. ROX позволява корекция на разликите в отделните съдове, получени поради неточности в пипетирането и флуктоации във флуоресценцията. Присъствието на ROX не влияе върху qPCR, когато се използват други системи, напр. iCycler iQ, защото багрилото не се включва в реакцията и има различен емисионен спектър (максимумите на възбуждане/ излъчване са при 580 nm/621 nm), в сравнение със SYBR Green I. В кита не са включени праймери. Праймерите се добавят от експериментатора допълнително в хода на работа. Полимеразните верижни реакции в реално време бяха проведена в апарат 7500 на Applied Biosystems. Съобразно с изискванията да не се допуска излагането на SYBR Green/ROX qPCR Master Mix (2X) на светлина по време на работа, за да се избегне загуба на интензитет на флуоресцентния сигнал.

IV.17. Провеждане на qRT-PCR (Fermentas SYBR Green PCR Kit/Thermo) анализ

Количественият PCR анализ беше извършен с помощта на Maxima Hot Start Taq DNA Polymerase е Taq ДНК-полимераза, която е химически модифицирана, чрез добавяне на термолабилни блокиращи групи.

IV.18. Провеждане на обратна транскрипция на протеин-кодиращи гени

Поради невъзможността РНК да бъде директно използвана за амплификация с полимеразна верижна реакция, за PCR анализа е необходимо да

се синтезира копи-ДНК от РНК чрез обратна транскрипция. Използван е кит Maxima® First Strand cDNA Synthesis Kit на „Thermo scientific“.

IV.19. Провеждане на qRT-PCR (Fermentas SYBR Green PCR Kit (Thermo))

За провеждането на реакциите беше използван PCR-апарат 2720 на Applied Biosystems.

IV. 20. Анализ на експресионният профил на малки РНК молекули от периферна кръв при РАС (Small RNA sequencing)

Създаване на комплексни проби (пул) и преципитация на РНК, подготовка на получените проби за анализ (Small RNA sequencing)

След определяне на концентрациите на всички изолирани РНК образци беше изчислено необходимото количество от всяка проба за получаване на комплексна пул проба, включваща всички проби на деца с поставена диагноза разстройство от аутистичен спектър и съответно контролна група деца. Смесването на пробите бе извършено върху сух лед, с цел минимализиране възможността за деградация на РНК. За успешното транспортиране и съхранение на изолираните РНК образци, пробите са подложени на преципитация.

IV. 21. Анализ на експресионният профил на малки РНК молекули (Small RNA sequencing)

Посредством прилагането на протокол за провеждане на широко-машабното секвениране на малки РНК молекули беше използван напълно потенциала на метода като беше получена информация за количествената експресия на всички малки РНК присъстващи и в анализиранияте проби, в контраст с други методи за анализ при които се анализират основно силно експресираните или тези молекули които показват значителна разлика в експресията между изследваните състояния. Идентифицирани бяха 259 статистически значимо променени миРНК. Анализът включваше и изследването на диференциалният профил на експресия на секвенираните малки РНК по класове който включва: свързаните с повтори малки РНК молекули, картирани по екзонни секвенции, картирани по интронни секвенции, картирани по транспортни РНК и др.

IV. 22. Биоинформатичен анализ на миРНК експресионният профил чрез секвениране на малки РНК молекули от периферна кръв при РАС

Стандартен биоинформатичен анализ - Стандартният биоинформатичен анализ беше проведен на базата на получените данни от профила от малки РНК молекули (експресионния профил).

Детайлен биоинформатика анализ

Диференциален експресионен анализ на миРНК молекулите. 2. Идентифициране на таргетни и РНК молекули на идентифицираните миРНК молекули. 3 . Анотации на таргетни гени GO анотация и KEGG анализ на миРНК. 4 . Анализ на базите в миРНК секвенциите претърпели редактираните (editing).

IV. 23. Анализ на експресионния профил на информационни РНК молекули (RNA Sequencing)

Създаване на комплексни проби (пул) и подготовка на получените проби за транскриптомен анализ (RNA Sequencing)

След определяне на концентрациите на всички изолирани РНК образци беше изчислено необходимото количество от всяка проба за получаване на комплексна пул проба, включваща всички проби на деца с поставена диагноза разстройство от аутистичен спектър и съответно контролна група деца. Смесването на пробите бе извършено върху сух лед, с цел минимализиране възможността за деградация на РНК. За успешното транспортиране и съхранение на изолираните РНК образци, пробите са подложени на преципитация както следва:

Добавяне на 3 обема абсолютен етанол;

Добавяне на 1/10 обем 3M NaOAc, pH 5.2

IV. 24. Секвениране на РНК (Транскриптомен анализ - Методика)

РНК – секвенирането, още известно като "Whole Transcriptome Shotgun Sequencing" е определен за "революционен инструмент за транскриптомни изследвания", методът се базира на използването на високопроизводителни технологии за секвениране на комплементарни ДНК (кДНК) последователности, с цел получаване на информация за секвенцията на проба РНК. РНК – секвенирането бързо се превръща в инструмент за изследването на редица

заболявания болести като рак и др. В продължение на много години, стандартният метод за определяне на последователността на транскрибираните гени включваше улавянето на секвенция от последователността на информационна РНК чрез използването на експресирани маркерни секвенции (expressed sequence tags: ESTs) или пълното секвениране на кДНК последователности, чрез използването на конвенционални технологии за секвениране като секвениране по Сенджър (Sanger seq.). Въпреки това, РНК – секвенирането има редица предимства пред конвенционалните методи за секвениране. РНК – секвенирането дава информация за иРНК като генерира десетки милиони къси фрагменти които многократно повтарят секвенция от определен локус, като тези данни могат да бъдат използвани в измерване на нивото на генна експресия.

Изборът на подходяща платформа за количествен анализ на РНК посредством РНК секвениране е продиктуван от добрата възпроизводимост на резултатите от предлагана платформата на BGI за транскриптомен анализ, както и от ниският лимит на детекция (чувствителността) на предложеният анализ.

Пробите от тотална РНК първоначално се подлагат на третиране с ДНК-аза I с цел разграждане на всяко възможно замърсяване на пробите от тотална РНК с геномна ДНК. В последствие се провежда обогатяване на иРНК молекули с помощта на олиго (dT) магнитни частици. Чрез инкубиране на иРНК молекули в буфер за фрагментиране, иРНК се фрагментират до къси фрагменти с размер от около 200 бази. Непосредствено след фрагментирането се провежда протокол за синтез на едноверижна кДНК верига чрез използване на неспецифични (random) хексамерни праймери, последвано от добавяне на буфер, дНТФ, RNase H и ДНК полимераза I за синтезиране на втората ДНК верига. Получената двойноверижна ДНК се пречиства с помощта на магнитни зърна. След пречистването на двойноверижната ДНК се извършва поправка на краищата и добавяне в 3'-край на единичен нуклеотид A (аденин). Следва добавяне на адаптори за секвениране към получените фрагменти. Фрагментите са обогатени чрез провеждане на PCR амплификация. По време на етапа на качествена оценка (QC), е използван Agilent 2100 Bioanalyzer и ABI StepOne Plus Real-Time PCR система, за количествен и качествен анализ на получените библиотеки. Така получените библиотеки са подложени на секвениране с помощта на Illumina HiSeq™ 2000.

Качествена оценка на секвенирането

Филтриране на данни от секвенирането

Поради наличието на адапторни последователности и/или секвенирани фрагменти (reads) с ниско качество, които присъстват в необработените данни

(row data), се извършва филтриране на данните за получаване на секвенирани фрагменти с високо качество (чисти данни). Процедурата на филтруване включва следните стъпки:

1) Премахване на секвенирани фрагменти съдържащи адапторни последователности.

2) Премахване на секвенирани фрагменти, в които процентът на неизвестни бази (N) е по-голям от 10%.

3) Премахване на секвенирани фрагменти с ниско качество. Ако процентът на бази с ниско качество (база със стойност ≤ 5) е по-голям от 50% в фрагментите, тези фрагменти се определят като такива с ниско качество.

Количествен анализ на генната експресия

Нивото на експресия на всеки анализиран ген се определя от броя на секвенираните фрагменти картирани по секвенцията на специфичният ген към общият брой на специфично картирани фрагменти от анализираната проба.

Изследване на генно покритие

Генното покритие е процентът от гена обхванати от секвенираните фрагменти. Тази стойност се определя като съотношението от броя бази на съответният ген обхванати от специфично картираните фрагменти към общия брой бази на съответният ген.

Изследване на нива на генна експресия

Нивото на генна експресия се изчислява по RPKM (Reads Per kb per Million reads) (Mortazavi, 2008) метода (секвенирани фрагменти на килобаза на милион фрагменти) и следната формула:

$$RPKM = \frac{10^6 C}{NL / 10^3}$$

Където RPKM (A) е нивото на експресия на гена A, C е броя на секвенираните фрагменти, който еднозначно се картират върху гена A, а N е общият брой на секвенираните фрагменти, който еднозначно се картират върху всички гени и L е броя на базите в гена A. Чрез използването на RPKM метода е възможно да се елиминира влиянието на различната дължина на гена и несъответствията, получени от секвенирането при изчисляването на нивото на генната експресия.

Следователно, RPKM стойностите могат да бъдат директно използвани за сравняване на разликата в генната експресия между анализирани проби.

Идентифициране на диференцирано експресирани гени (DEGs)

За идентифицирането на диференцирано експресирани гени е използвана р-стойността като показател за диференциална генната експресия и FDR стойността (False Discovery Rate) за оценка на прага на р-стойността. В настоящата работа е използвано $FDR \leq 0.001$ и абсолютната стойност на съотношението $\log_2 \geq 1$, като праг за преценка на разликата при генната експресия.

V. 25. Генно онтологичен анализ на диференциално експресирани гени

Генната онтология (ГО) е международна стандартизирана класификация за генна функция, която предлага контролирана и динамично-обновяваща се база данни със стриктно дефинирани свойства на гените и техните продукти. ГО обхваща три аспекта: клетъчен компонент, молекулярна функция и биологичен процес. Основна единица на ГО е ГО единицата (елемент). Анализът предоставя всички ГО елементи, които съдържат значими диференциално експресирани гени, сравнени с генома, и филтрира тези, които отговарят на биологични функции.

IV. 26. KEGG анализ на биологичните пътищата обогатени с диференциално експресирани гени

Гените обикновено взаимодействат едни с други като така реализират специфичните си биологични функции. KEGG е основната база данни с публичен достъп свързана с данни за биологични пътища и генно взаимодействие. Анализът идентифицира метаболитни и сигнални пътища свързани с диференциално експресирани гени сравнени с целия геном. Изчисленията се извършват по следната формула:

$$P = 1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}$$

Където N е броя на всички анотирани гени в KEGG, n е броя на диференциално експресирани гени в N; M е броя на всички гени свързани със специфичен път, Докато, m е броя на диференциално експресирани гени в M. Списъкът на пътищата установени с този анализ е интерактивен. Изборът на

специфичен биологичен път отвежда към детайлна информация за същият в KEGG базата данни.

IV. 27. Количествен протеомен анализ – *Isobaric Tag for Relative and Absolute Quantification (iTRAQ)* при PAC

Идентифицирането и количественото определяне на протеини от комплексни смеси беше улеснено от въвеждането на масспектрометрични методи (MS) за количествен протеомен анализ. Използването на изобарни маркери за относителна и абсолютната количествена оценка (*Isobaric Tag for Relative & Absolute Quantification (iTRAQ)*) се състои в свързването на амин-специфична, стабилна изотопна маркировка, чрез която може да се постави специфични пептиди на различни биологични проби. Протеомният анализ изисква като първоначална стъпка фрагментиране на белтъчните молекули с трипсин до получаване на малки пептидни фрагменти.

Пептидните фрагменти, генерирани след трипсинизиране са разделени посредством SDS-PAGE гел-електрофореза за улеснение, след което се маркират с iTRAQ специфичен маркер. Реагентът iTRAQ се състои от един репортерна група, и балансираща част и пептид реактивна група която взаимодейства с N-края на пептидите или свободна аминокислотна група на Лизинови остатъци, като дава обща молекулна маса от 145. Репортерната група се използва да маркира уникално всеки пептид, с молекулна маса варираща между 114-117, като по този начин позволява количествено определяне на протеините в отделни проби едновременно. Маркираните проби се смесват до получаване на обща смес. Общата смес съдържаща маркираните проби се пречиства на катионо-обменна колона за премахване на несвързания iTRAQ реагент. (*Strong Cation Exchange Chromatography (SCX)* - Маркираните пептиди се фракционират на катионо-обменна колона, за отстраняване на всички несвързани реагенти). Това улеснява пречистването на пробите преди по-нататъчно финно разделяне и пречистване с помощта на обратно фазова хроматография. Допълнително пречистване на SCX обработените проби е извършено с помощта на течна хроматография с обръщане на фазите където пробата преминава през колона, съдържаща стационарна фаза матрикс, където селективно се адсорбират само някои молекули. елюираните фракции се характеризират с помощта на MS.

От биологична гледна точка, връзката между експресията на протеини и нивото на съответните им информационните РНК транскрипти до голяма степен може да бъде обяснена като резултат на молекулярните събития включващи: ефективност на транслацията, алтернативен сплайсинг, нагъване, асемблиране в комплекси, транспорт и локализация, ковалентни модификация, секреция, и деградация, всички от които процеси засягат нивата на протеини, независимо от

транскрипцията. Значението на тези посттранскрипционни процеси беше разкрито от съвременни данни, които показват, че присъствието на генетична вариация в някои тези посттранскрипционни процеси се асоциира с определени човешки заболявания.

Предимства на iTRAQ метода

1) Тъй като iTRAQ реагент може да маркира всеки един пептид, съдържащ пост-транслационни модификации, така методът може да подобри надеждността и покритието от идентифицирани белтъци. 2) Методът може да подобри надеждността на количественото измерване, тъй като позволява количествено определяне на пептидната маса на протеина. 3) Това е изследователски метод с висок потенциал за откриване на биомаркери. 4) Метода предлага висока точност на количествената оценка. 5) Метода има потенциала да сравнява нивото на протеин в няколко проби едновременно.

IV. 28. Мащабно екзомно секвениране (Whole-Exome Sequencing) за изследване на мутации в белтък-кодиращи последователности в човешкия геном асоциирани с РАС

Поради изключителната генетична комплексност на разстройството и връзката на множество локуси и относително слабата корелация генотип-фенотип, РАС се разглежда като модел на заболяване, при което прилагането на пълното екзомно секвениране е добре аргументирано. На базата на така получените резултати от пълно-екзомно секвениране е развит модел, включващ участието на стотици гени, с потенциал за възникване на *de novo* мутации. По-голямата част от тези мутации могат да увеличат риска, но не и да се превърнат в "причина" за разстройството, което е в подкрепа на вече предложеният полигенен модел за етиопатогенезата и развитието на РАС. Трудоемкият подход за пълното секвениране на всички кодиращи регион има потенциала да се превърне както в клинично значимо за генетичната диагностика изследване, така също и за идентифицирането на варианти в ДНК с потенциално значим патологичен ефект при заболявания с все още неизвестна етиология. Целта на този подход е идентифицирането на функционални промени в генома, които са отговорни за развитието на специфична патология.

Идентифициране на генетични вариантите с помощта на IGV браузър

Идентифицирането на генетични варианти от проведеното екзомно секвениране беше извършено с помощта на IGV браузър, позволяващ проследяването на всеки идентифициран вариант като предоставя детайлна информация относно: 1) геномната локализация на идентифицираният вариант 2)

типа мутация до която води идентифицираният вариант (загуба на стоп кодон, мутации с изместване на рамката на четене, мутации в места за извършване на сплайсинг, синонимни мутации и др.) 3) ген, в който се наблюдава идентифицираният вариант 4) типа на получения вариант в резултат на последващият клиничен ефект: увреждащи (*deleterious*), варианти с неизвестна клинична значимост и благоприятни варианти (*benign*).

IV. 29. Идентифициране на мутации при РАС с помощта на секвениране на РНК от следващо поколение (NGS RNA sequencing)

Проведеният биоинформатичен анализ за идентифициране на мутации при РАС е извършен използвайки предварително получени данни от РНК транскриптомен анализ съдържащи 27,75 млн. секвенирани фрагмента от комплексна проба на децата с диагноза РАС и 28,35 млн. секвенирани фрагмента от комплексна проба представляваща извадка от контролната група деца. Обработката на „суровите“ данни беше проведена, посредством компютърни програми, налични на страницата на достъпния онлайн компютърен клъстер „Galaxy“ (<https://usegalaxy.org>). Проверката на качеството на получените секвенирани фрагменти беше извършена с програмата FastQC.. Програма TopHat 2.0 беше използвана като възможност за бързо и точно картиране на голям обем „сурови“ секвенции върху референтен геном и генериране на алайнмънт файл в “.bam” - формат. За провеждането на анализа бяха използвани настройки по подразбиране на програмата TopHat и актуалната версия на човешкия геном от (*Homo sapiens* hg38 Canonical) като референтен геном. В резултат от проведения алайнмънт повече от 97% от всички секвенирани фрагменти бяха картирани успешно. Анализът на вариантите (Variant calling) беше осъществен посредством FreeBayes детектор на варианти. Детекцията на вариантите беше проведена с избраните настройки, по идентичен начин за данните получени от двете комплексни проби представителни за анализираните извадки. Като следваща стъпка от анализа, получените данни бяха сравнени и филтрирани, за идентифициране само на онези варианти които имат качество на детекция по-голямо от 500 (QUAL > 500, където QUAL стойността е отрицателен десетичен логаритъм на вероятността алтернативният алел да е отчетен погрешно) и присъстват значимо само в комплексната проба на пациентите с РАС, и не се срещат при пробата от здрави деца. Идентифицирани бяха 176 варианта, всички те бяха анотирани с програма SnpEff Variant effect and annotation (*Galaxy Version 4.0.0*). Потенциалният фенотипен ефект на аминокиселинната замяна бе оценен статистически с онлайн достъпната програма PolyPhen-2 (<http://genetics.bwh.harvard.edu/pph2/>), достъпна и през базата данни Ensembl (<http://www.ensembl.org/index.html>).

IV. 30. Транскриптомно профилиране на белтък-кодиращи гени в периферна кръв при шизофрения (Digital Gene Expression Tag Profiling (DGE))

Създаване на комплексни проби (пул) за провеждане на сравнителен експресионен анализ (Digital Gene Expression Tag Profiling (DGE))

След определяне на концентрациите на всички изолирани РНК образци беше изчислено необходимото количество от всяка проба за получаване на комплексна пул проба, включваща всички проби от анализирани групи. Смесването на пробите бе извършено върху сух лед, с цел минимализиране възможността за деградация на РНК. За успешното транспортиране и съхранение на изолираните РНК образци, пробите бяха преципитирани съответно изискванията. Beijing Genomics Institute (BGI) беше използван като сертифициран доставчик на услуги за NGS секвениране.

IV. 31. Потвърждаване на диференциално експресираните гени от анализ чрез провеждане на количествен Real Time PCR анализ

Валидиране на данните от проведените транскриптомен анализ се осъществи с помощта на количествен Real-Time PCR анализ. На индивидуален количествен Real-Time PCR анализ бяха подложени 5 диференциално експресираните гени във всички проби от пациенти с диагноза шизофрения и контролна група от здрави лица. Промяната в генната експресия е анализирана чрез използване на $2^{-\Delta\Delta CT}$ метода.

V. РЕЗУЛТАТИ

V.1. Резултати от проведеното мащабно екзомно секвениране (*Whole Exome Sequencing*) на белтък-кодиращи последователности в човешкия геном асоциирани със РАС

С цел идентифициране на геномни варианти при пациенти с диагноза РАС беше извършено мащабно екзомно секвениране (*Whole Exome Sequencing*) в комплексни проби от извадките обект на изследване, които включват кохорта от деца с диагноза РАС и кохорта от контролна група здрави деца, всяка от която включва 40 деца.

Получените от пълното екзомно секвениране геномни варианти бяха анотирани по отношение на принадлежност към ген и съответно генна функция с помощта на данни от базата данни *Ensembl*. Резултатите от проведените анализ са насочени върху това кои гени и съответните им транскрипционни варианти са

засегнати от идентифицираните генетични промените, и дали тези варианти могат да причинят значителни функционални проблеми.

ДНК секвенирането беше извършено на платформа за секвениране от следващо поколение (NGS) на Illumina в лабораторията на Oxford Gene Technology (OGT). Накратко, получените комплексни образци от геномна ДНК бяха подложени на фрагментиране с използването на ултразвук, лигирани към мултиплексни адаптори (*pairedend adapters*) и амплифицирани с помощта на PCR с използването на праймери за секвениране, белязани със специфични баркодове (индекси). Секвенирането беше проведено на платформа HiSeq 2000, Illumina.

Геномните варианти идентифицирани след провеждането на пълното екзомно секвениране бяха както следва: мутации в ДНК последователността, включващи инерции, делеции, единични нуклеотидни полиморфизма (SNPs) и др. Вариантите които преминаха през дискриминативните филтрите (стъпките описано по-горе) бяха класифицирани като: увреждащи (*deleterious*) мутации (потенциално патогенни варианти), вариантите с неизвестна клинична значимост, или благоприятни варианти (*benign*). Класификацията на идентифицираните варианти е извършена в съответствие с насоките за тълкуване на Американския колеж по Медицински Генетика и геномика *American College of Medical Genetics and Genomics* (ACMG).

Двете основни цели, които бяха поставени при провеждането на настоящото изследване бяха: **1)** да се потвърдят установени геномни варианти характерни за РАС в изследваната кохорта, показващи тенденция и подкрепящи хипотезата за генетична хетерогенност при РАС. **2)** да се идентифицират специфични геномни варианти в анализираната извадка от пациенти с диагноза РАС, които не са установени до момента. По отношение на поставената първа цел в проучването, получените данни подкрепят съществуването на патогенетични мутации в изследваните извадки някои от тях характерни за РАС.

Известните варианти от dbSNP (Release 135) бяха анотирани, в отделни набори от данни, което позволява сравнително лесно идентифициране на нови варианти със сериозни прогнозни клинични последствията. На групите включени в анализа беше оценяван статуса за носителство на генетични варианти (според критериите изброени по горе) чрез секвениране на всеки от кодиращите екзонни региона. Екзоните бяха анализирани на базата на данните получени от мащабното екзомно секвениране. Генерираните данни бяха получен при използвана дълбочина на секвениране от 50X екзомно покритие на всеки специфичен локус, което позволява оценката на наблюдаваните варианти в изследваните групите да бъде извършена чрез оценка на мутационната честота, като функция от броя

секвенирани фрагменти (*reads*), получени от съответният геномен регион. Сред идентифицираните кандидат гени, един от най интересните е ANK3 (*Ankyrin 3*) гена. Редица от широкомащабните проучвания проведени при РАС показват силна асоциация на гена ANK3 със РАС. Наскоро, чрез използването на екзомно секвениране бяха идентифицирани миссенс мутации в ANK3 гена при четири от 67 пациента с диагноза РАС. Предвид факта, че ANK3 гена се асоциира силно с шизофрения и биполарно разстройство, получените данни подкрепят хипотезата за асоциацията между мутации в ANK3 гена и РАС, което е в подкрепа на твърденията за наличието на обща молекулярна основа между и РАС и други невропсихиатрични разстройства.

Обобщеният анализ на получените данни от пълното екзомно ДНК секвениране демонстрира наличие на мутации при РАС, водещи до формирането на нефункционални протеини.

Основните мутации, с потенциал за сериозни биологични последствия могат да бъдат разделени в няколко отделни категории:

1. **Мутации, водещи до поява на терминаращи кодони**, важни поради принципното приемане, че получаването на терминаращ кодон води до формирането на по-малка иРНК и съответно нефункционален белтък.
2. **Мутации, водещи до загуба на терминаращи кодони**, важни поради принципното приемане, че промяната на всеки терминаращ кодон води до формирането на по-голяма иРНК и съответно патологично функционален белтък.
3. **Мутации, водещи до появата на сериозни несинонимни замени**, важни поради принципното приемане, че всяка промяна в аминокиселинната последователност може да има за резултат формирането на патологично функционален или нефункционален белтък.
4. **Мутации, свързани с изместване на рамката на четене (*frameshift mutation*) – делеции или инсерции**, важни поради принципното приемане, че водейки до промяна в аминокиселинната последователност може да имат за резултат формирането на белтък с патологична функция или напълно нефункционален протеин.

В групата от деца с диагноза РАС прави впечатление, че мутациите, **водещи до формирането на стоп кодони** са свързани с гени като MOB3C (*MOB kinase activator 3C*), асоцииран със свързването на метални йони. *Cytochrome P450 Family 4*

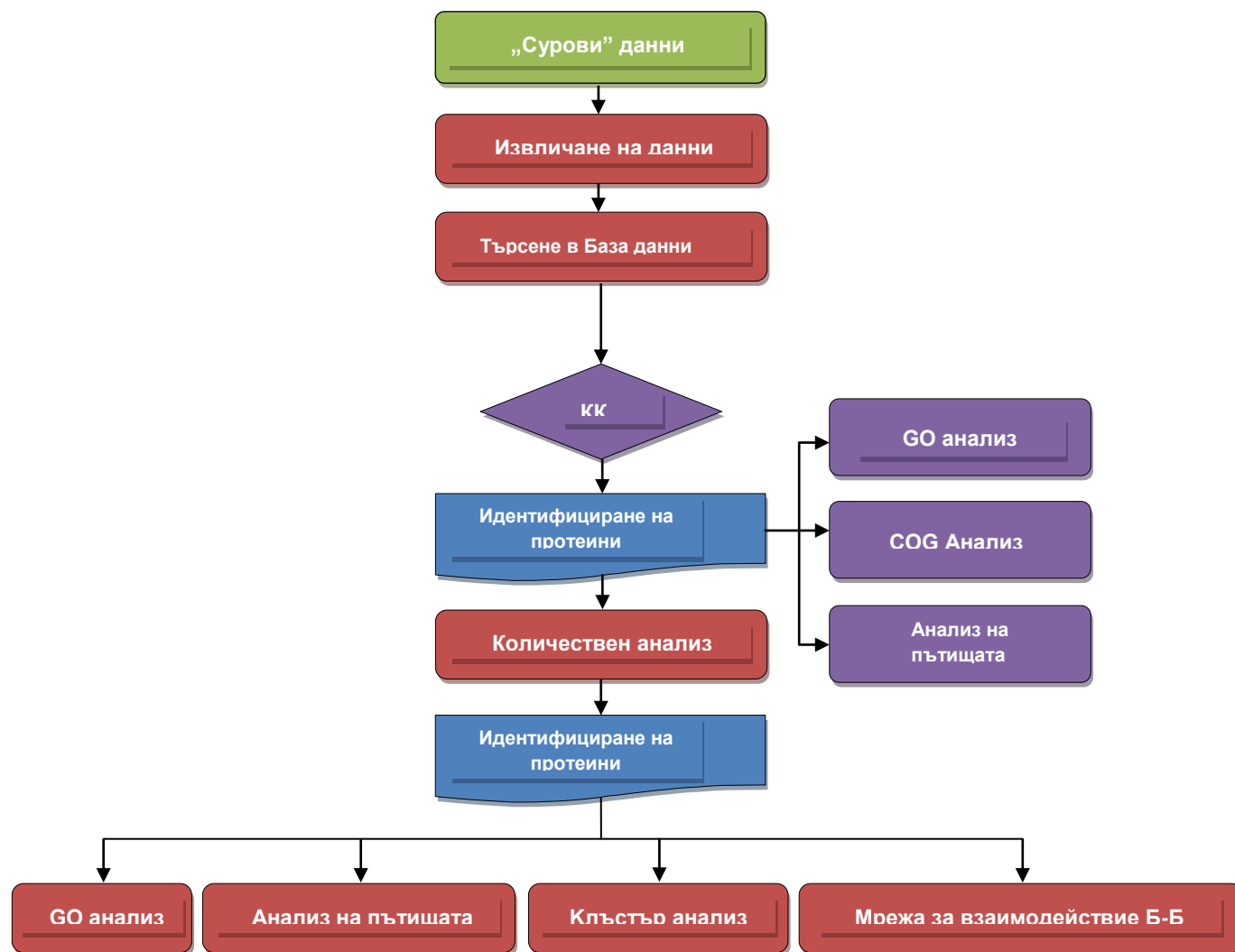
Subfamily B Member 1 (CYP4B1), ген кодиращ член от цитохром P450 суперфамилията от ензими. Протеините на цитохром P450 катализират реакции, свързани с метаболизма на лекарствени препарати и синтеза на холестерол, стероиди и други липиди. *Regenerating islet-derived protein 4* (REG4), вероятно свързан с гастроинтестиналния епител и възпалението, *PDE4DIP*, фосфодиестеразна котва в центрозомата, *RHBG*, транс мембранен протеин, свързан с Rh фамилията, *OR10X1*, G белтък-асоцииран *рецептор*, свързани с олфакториума, *FM06P*, флавин съдържащ монооксигеназен псевдоген, *CAPN8*, асоцииран с калциево-зависима цистеин ендопептидазна активност и невронална цериодна липофусциноза. *TSSC1*, тумор супресиращ фактор, асоцииран със синдрома на Beckwith Wiedemann, тумор на Wilms, рабдомиосарком, адренокортикален карцином и др.

Идентифицираните гени с мутации, водещи до загуба на стоп кодони включват гени като: *Fasciculation and elongation protein zeta 2* (FEZ2), свързан с невроналното свързване, *Cytosolic beta-glucosidase* (GBA3), свързан с хидролиза на гликозиди, *Alcohol dehydrogenase 1C (class I), gamma polypeptide* (ADH1C), алкохол дехидрогеназа, *matrilin 2* (MATN2), свързан с формирането на филаментозни връзки в извънцелуларния матрикс, *N-Acetylneuraminase Synthase* (NANS), свързан с биосинтезата на сиаловата киселина, *Actinin alpha 3* (ACTN3), асоцииран с невроналната свързаност, *Neurofibromin 1* (NF1), свързан с неврофиброматоза тип 1 и синдромен аутизъм.

Идентифицирани също бяха гени с мутации, водещи до появата на несинонимни замени в групата на пациентите с PAC, като *Spermatogenesis associated 3* (SPATA3), асоцииран със сперматогенезата, *Solute Carrier Family 22 Member 1* (SLC22A1), транс мембранен белтък, свързан с допаминовия транспорт, *AarF Domain Containing Kinase 5* (ADCK5), вероятно с тирозинкиназна активност, *NLR Family CARD Domain Containing 5* (NLRC5), ген играещ роля в цитокиновия отговор и антивирусния имунитет чрез инхибиране на NF-каппа-B активирането и отрицателната регулация на интерферон тип I сигналните пътища, *Apolipoprotein B receptor* (APOBR), *NUDT11*, спадащ към фосфохидролазна фамилия.

Установени бяха също 537 мутации в гени, водещи до изместване на рамката на четене между които са гени като: *Cub Domain-Containing Protein 2* (CDCP2), *Caspase 9* (CASP9), от групата на каспазите, *Hornerin* (HRNR), *Ubiquitin-like-conjugating enzyme ATG3* (ATG3), убиквитин-подобен конюгиращ ензим, *NOP16*, кодиращ нуклеоларен протеин и *Fibroblast Growth Factor Receptor 4* (FGFR4-фибробластен растежен фактор 4).

V. 2. Биоинформатичен анализ на ITRAQ протеомни данни



Фигура 1. Общ план на процедурите за провеждане на биоинформатичен анализ. На фигурата е посочена информация за основните процедури при провеждане на биоинформатичния анализ. 1) идентифициране на пикове от „сурови“ данни, получаване на списък от идентифицирани пикове. Изграждане на база данни с информация за идентифицирани пептиди и протеини. Най-накрая, сравняване на относителното количество на протеини между различните проби, получавайки информация за важни кандидат протеини. Също така анализът позволява комбиниране на протеомни с транскриптомни данни, за получаване на комбиниран протеомен и транскриптомен анализ.

Сравнителен анализ на диференциално експресирани протеини между анализирани групи (РАС vs. контролна група) с понижена експресия

От получените данни от iTRAQ-LC-MS/MS анализа бяха идентифицирани общо 351 протеина. В сравнение с контролната група, бяха установени общо 60 диференциално експресирани протеини, включително 24 с повишена експресия и 36 с понижена експресия.



Фигура 2. Анализ на диференциално експресирани протеини. По оста – **x**; изследвани проби; По оста – **y** брой на диференциално експресирани протеини. Колоната в зелено показва съответно броя на протеини (36) с понижена експресия.

Установените протеини с понижена експресия включват:

SERPINE2 Glia-derived nexin (GDN)

Серинов протеазен инхибитор с активност към тромбин и трипсин. SERPINE2 е фактор участващ в удължаването на невоните чрез инхибиране на тромбина. Невронното удължаване и разклоняване са ключови клетъчни събития по време на развитието на мозъка, тъй като те са в основата на формирането на правилно изграждане на невронните мрежи.

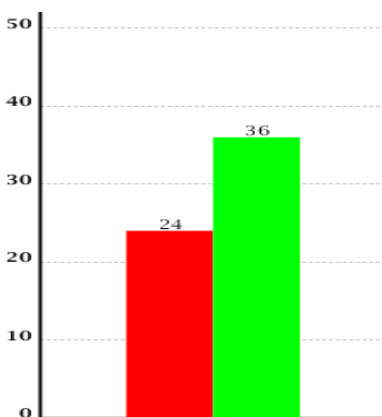
Интеллектин-1

Имуноглобулин J

Ангиогенин

Фибулин 5

Сравнителен анализ на диференциално експресираните протеини между анализираниите групи (РАС vs. контролна група) с повишена експресия

[illegible]

Фигура 3. Анализ на диференциално експресирани протеини. По оста – x; изследвани проби; По оста – y брой на диференциално експресирани протеини. В червено са посочени броя на протеините (24) показали повишена експресия.

Аполипопротеин С-II

Аполипопротеин С-II е компонент на хиломикроните, липопротеини с много ниска плътност (VLDL), и липопротеини с ниска плътност (LDL) и липопротеини с висока плътност (HDL) в плазмата. Играе важна роля в метаболизма на липопротеините като активатор на липопротеиновата липаза.

Аполипопротеин С4 (АРОС4)

Аполипопротеин С4 е липид-свързващ протеин, принадлежащ към семейството на аполипопротеиновите гени. Смята се, че протеинът играе роля в метаболизма на липидите.

Аполипопротеин F

Този протеин образува комплекси с липопротеини и може да участва в транспорт и /или естерификацията на холестерола.

Комплемент С4-А протеин и Комплемент С1q субкомпонент субединица А

Неензимен компонент на С3 и С5 конвертазата, от съществено значение за извършването на класическия път на комплемента. С1q се свързва с проензимите С1r и С1s, до получаването на С1, първият компонент на серумната система на комплемента.

Периостин

Секреторен протеин на екстрацелуларният матрикс, който функционира в развитието и регенерацията на тъканите, включително зарастване на рани и процеса на камерно ремоделиране след миокарден инфаркт. Протеин се свързва с интегрини, за да поддържа адхезията и миграцията на епителните клетки.

Миозин-9

Клетъчният миозин, играе роля в цитокинезата, клетъчната форма и специализираните функции като секреция. Миозина участва в движението на клетките (клетъчна подвижност); поддържане на формата на клетките и процеса на цитокинеза.

Пластин-2

Актин-свързващ протеин. Играе роля в активирането на Т-клетките. Модулира експресията на клетъчните повърхностни протеини IL2RA/CD25 и CD69.

V.4. Анотация по клъстери от ортоложни групи от протеини (Clusters of Orthologous Groups of Proteins (COG))

С цел извличане на максимална информация от бързо натрупващите се данни от получените протеомни данни, възниква необходимостта всички протеини да бъдат класифицирани в съответствие с техните хомоложни взаимоотношения.

От COG анотацията внимание привличат определени функционалните класове в които се наблюдава висок брой диференциално експресираните протеини: Обмен на протеин и посттранслационни модификация, чаперони (O), с ангажирани 98 протеина, Общи функции (единствено предсказани) (R), 73 протеина и Цитоскелет (Z), 28 протеина. Докато при останалите категории броят на диференциално експресираните протеини е значително по малък Енергиен метаболизъм (C), 14 протеина, Метаболизъм и транспорт на аминокиселини (E), 14 протеина, Метаболизъм и транспорт на въглехидрати (G), 16 протеина, Транслация, структура на рибозома и биогенеза (J), 10 протеина.

V. 5. Генно онтологичен (GO) анализ на диференциално експресирани протеини

Генната онтология (GO) е международна стандартизирана класификация за генна функция, която предлага контролирана и динамично-обновяваща се база данни със стриктно дефинирани свойства на гените и техните продукти. GO обхваща три аспекта: клетъчен компонент, молекулярна функция и биологичен процес. Основна единица на GO е GO единицата (елемент). Всеки GO елемент принадлежи към тип онтология. Анализът предоставя всички GO елементи, които съдържат значими диференциално експресирани гени, сравнени с генома, и филтрира тези, които отговарят на биологични функции. Анализът първоначално картира диференциално експресираните гени в GO-елементите в базата данни, изчислява броя на гените във всеки GO-елемент като след това чрез геометрично изследване, открива GO елементи с значимо съдържание на диференциално експресирани гени сравнено с генома.

В съответният генно онтологичния аспект **клетъчен компонент** впечатление прави концентрацията на диференциално експресирани гени в елементи като екстрацелуларни региони 10.40% и клетъчна мембрана 8.80%, докато други остават по слабо застъпени. В генно онтологичния аспект **молекулярна функция** впечатление прави концентрацията на диференциално експресирани протеини в елементи като свързване 49.30%, в който елемент са средоточени приблизително 50% от диференциално експресираните протеини в генно онтологичния аспект, Други елементи с концентрация на диференциално експресирани протеини това са елементите: рецепторна активност 24.23%, каталитична активност 24.23%, и регулация на ензимната активност 7.34%. В генно онтологичния аспект **биологичен процес** концентрацията на диференциално експресирани протеини се наблюдава в елементи като метаболитни процеси 8.35%, регулация на биологичните процеси 7,53%, организация на клетъчните процеси и биогенеза 4,35%, клетъчни процеси 9,71% и др.

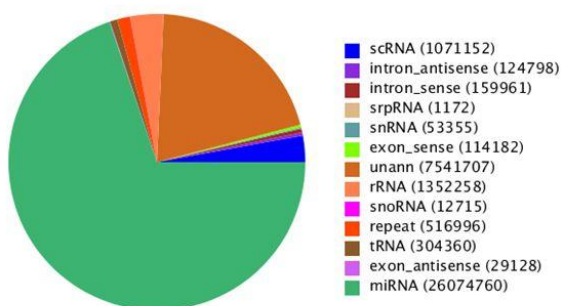
Фигура 22. Кръгова диаграма демонстрираща анализа на генната онтология. Кръговата диаграма показва разпределението на съответните GO единици (елементи). С различни цветове са обозначени различни елементи на GO анализа, докато площта на кръговата диаграма показва процента на GO елемента от общия брой протеини.

V.6. Резултати от проведеният анализ на експресионният профил на малки РНК молекули от периферна кръв при РАС

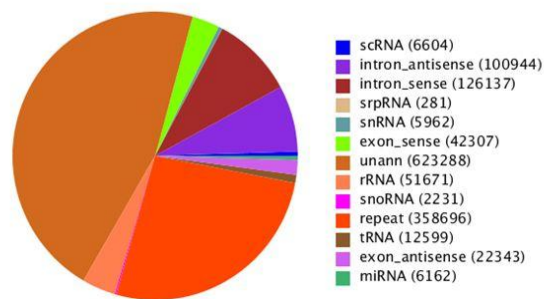
Анализа на резултатите от широко-мащабният експресионен анализ позволи детекцията на специфични промени в генната експресия при РАС като спомага за функционалното разграничаване на ключови регулаторни пътища и мрежи.

Стандартен биоинформатичен анализ за профилиране от малките РНК молекули по референтният геном.

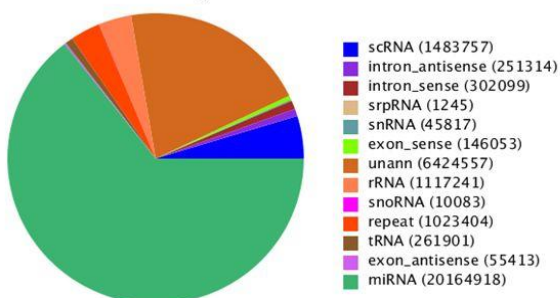
Представяне на данни в кръгова диаграма за аотирани малкиРНК при СТ-общ брой



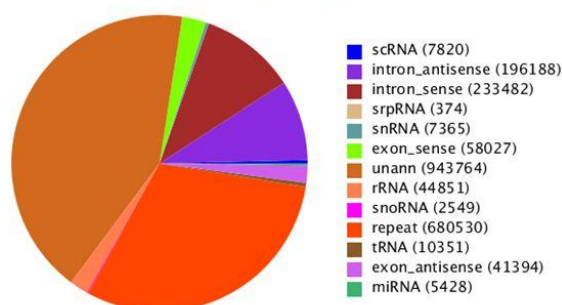
Представяне на данни в кръгова диаграма за аотирани малкиРНК при СТ-уникални



Представяне на данни в кръгова диаграма за аотирани малкиРНК при АТ-общ брой



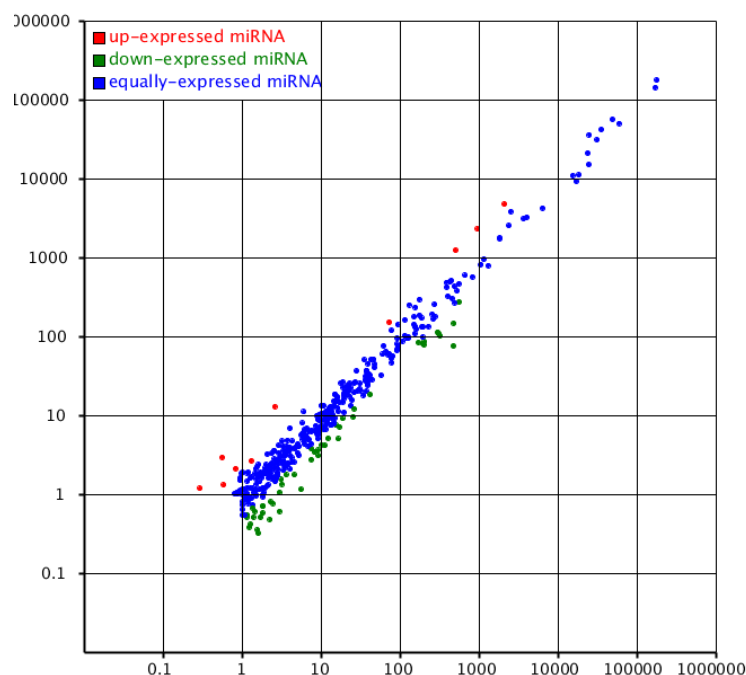
Представяне на данни в кръгова диаграма за аотирани малкиРНК при АТ-уникални



Фигура 4. Обобщена кръгова диаграма на аотираните малките РНК молекули по референтният геном (общ брой и уникални).

V.7. Идентифициране на диференциално експресирани анотирани миРНК молекули в анализирани групи чрез секвениране на малки РНК молекули

Диференциално експресирани миРНК молекули при РАС



Фигура 5. Скатър плот диаграма отразяваща експресията на миРНК молекули. На диаграмата е посочено сравнението между (РАС и Контролна група) Отбелязани са позиции показващи промени между анализирани групи с червено и зелено (съответно с повишена (Up regulated) и понижена (Down regulated) експресия) докато транскриптите които не показват диференциална експресия са показани в синьо.

V. 8. Резултати от проведеният анализ на миРНК експресионният профил чрез секвениране на малки РНК молекули при РАС

Изследването на профила от малки РНК молекули (експресионния профил) при РАС в сравнение с контролна група деца в периферна кръв показва наличието на специфични промени характерни за разстройството. Всичко това ясно демонстрира потенциала на метода и използването на периферна кръв за целите на идентифицирането на потенциални биомаркери при РАС. След проведения широкомащабен експресионен анализ бяха идентифицирани значителен брой миРНК молекули със статистически значима диференциална

експресия между анализирани групи както следва: let-7i-3p, miRНК-103a-2-5p, miRНК-106b-5p, miRНК-1249, miRНК-1307-5p, miRНК-134-5p, miRНК-139-5p, miRНК-142-3p, miRНК-145-5p, miRНК-15a-5p, miRНК-18b-3p, miRНК-193b-3p, miRНК-197-5p, miRНК-20b-3p, miRНК-210-5p, miRНК-29c-5p, miRНК-301a-3p, miRНК-3064-5p, miRНК-3620-3p, miRНК-365a-3p, miRНК-378d, miRНК-487b-3p, miRНК-3909, miRНК-424-5p, miRНК-4707-3p, miRНК-500a-5p, miRНК-500b-5p, miRНК-5010-3p, miRНК-501-5p, miRНК-5187-5p, miRНК-550a-5p, miRНК-5690, miRНК-584-3p, miRНК-589-3p, miRНК-619-5p, miRНК-664a-3p, miRНК-664b-3p, miRНК-671-3p, miRНК-6799-3p, miRНК-6850-5p, miRНК-96-5p, miRНК-183-5p, miRНК-199a-5p, miRНК-4734, miRНК-6849-3p, miRНК-223-5p, miRНК-3135a, miRНК-328-3p, miRНК-3674, miRНК-3687, miRНК-4489, miRНК-504-5p, miRНК-576-5p, miRНК-8052 и miRНК-937-3p).

Настоящото проучването представя доказателство, че периферната кръв е подходяща и лесно достъпна тъкан (материал) за изучаване на динамиката в експресията на кандидат биомаркери при РАС.

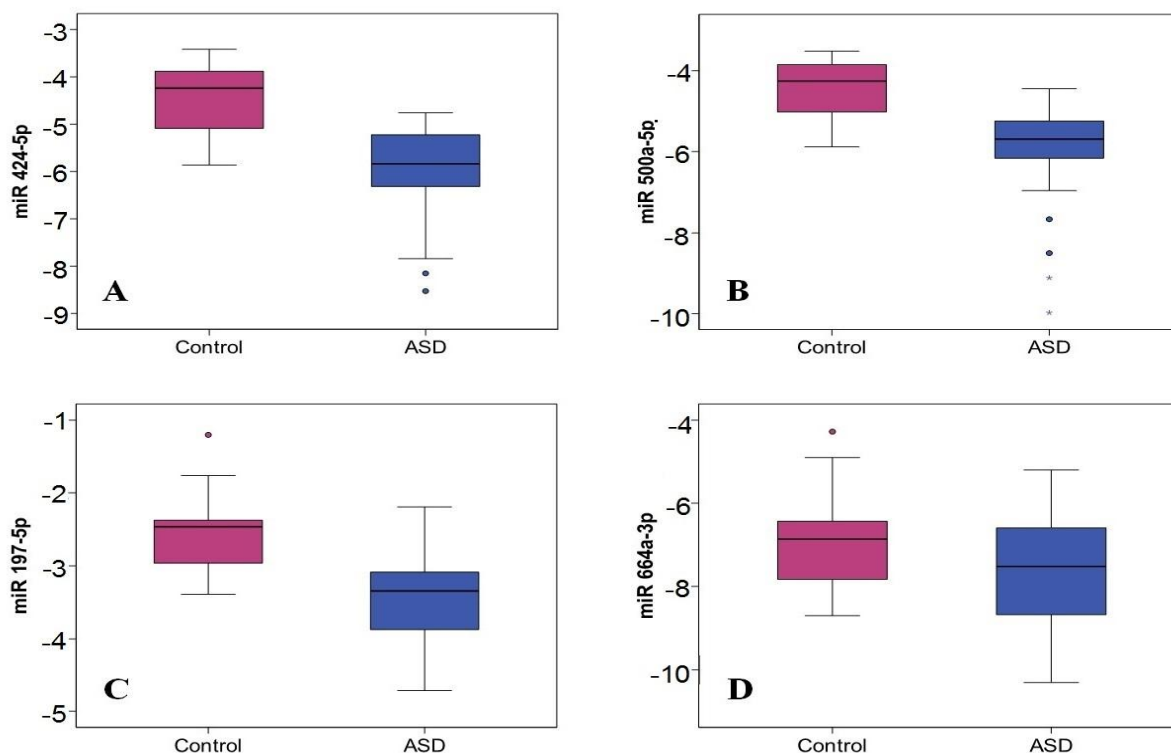
V. 9. Индивидуален експресионен анализ на miRНК молекули в серум от пациенти с РАС

За изследване на експресията на индивидуални miRНК молекули в серумни образци при РАС, беше използван miRНК специфичен stem-loop qRT-PCR анализ. Относителните нива на експресия бяха определени използвайки, Cel-miR-39 като екзогенна контрола, използвана за изчисляване на промените в miRНК експресията с помощта на $2^{-\Delta\Delta C_t}$ метод.

С цел идентифицирането на специфичните експресионни профили на серумни miRНК молекули, беше използван qRT-PCR анализ за проверка на потенциални кандидат биомаркери в сборни проби от групите от пациенти с диагноза РАС в сравнение със здрави контроли. В последното проучване бяха анализирани общо 42 miRНК, от които 29 бяха с понижено ниво на експресия при пациенти с РАС, 11 от тях показаха повишена експресия, а 2 от изследваните 42 miRНК молекули не показаха промени в експресионния профил. Създаването на сборни проби, обаче може значително да намали чувствителността на подхода и да предостави информация само за средната стойност за всички проби от представителната извадка. Поради това е необходимо потвърждение, което включва оценка на експресията на индивидуални проби от отделни пациенти включени в изследването. Четирите от miRНК молекулите, показали най-големи промени между изследваните сборни серумни проби, а именно miRНК-500a-5p, miRНК-197-5p, miRНК-424-5p и miRНК-664a-3p, бяха селектирани за

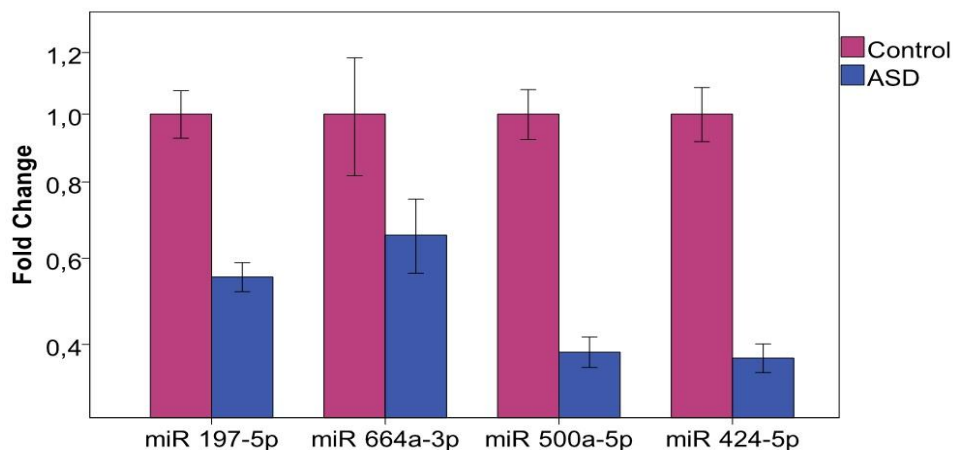
индивидуален анализ. Експресията на миРНК-500а-5р, миРНК-197-5р, миРНК-424-5р и миРНК-664а-3р беше анализирана при 38 деца, диагностицирани с РАС и 28 клинично здрави деца с помощта на количествен PCR в реално време (qRT-PCR).

Както може да бъде видяно на **Фигура 6**, тенденция към промяна в експресионния профил свързана с РАС, се наблюдава при всички изследвани миРНК молекули. Средната промяна в стойностите на Ct за миРНК-197-5р и тази на екзогенната контрола Cel-miR-39, използвана за нормализиране (делта Ct) беше около 2.83 цикъла по-ниско в проби от пациенти с РАС, в сравнение със съответните здрави контроли (**Фигура 6С**). В същото време средната промяна в Ct стойността за миРНК-664а-3р е 2.83 цикъла по-ниска (**Фигура 6D**); 2.16 цикъла по-ниска стойност за миРНК-424-5р (**Фигура 6А**) и съответно 1.63 цикъла по-висока за миРНК-500а-5р (**Фигура 6В**). Приложеният статистически анализ използвайки непараметричен тест на Kolmogorov-Smirnov демонстрира, че наблюдаваната понижена експресията е статистически значима за всички изследвани миРНК молекули, със съответните *p*-стойности, както следва: *p* < 0,001 за миРНК-424-5р; *p* < 0,001 за миРНК-197-5р; *p* < 0,001 за миРНК-500а-5р и *p* < 0,001 за миРНК-664а-3р. Идентичен експериментален подход беше приложен и за миРНК-365а-3р, която беше една от миРНК молекулите с повишени нива на експресия в едно от предходните изследвания извършено със сборни проби [Kichukova et al., 2017]. Резултатите от проведените анализ на миРНК-365а-3р не възпроизведеха първоначалните данни поради което миРНК-365а-3р не беше подложена на допълнителни анализи.



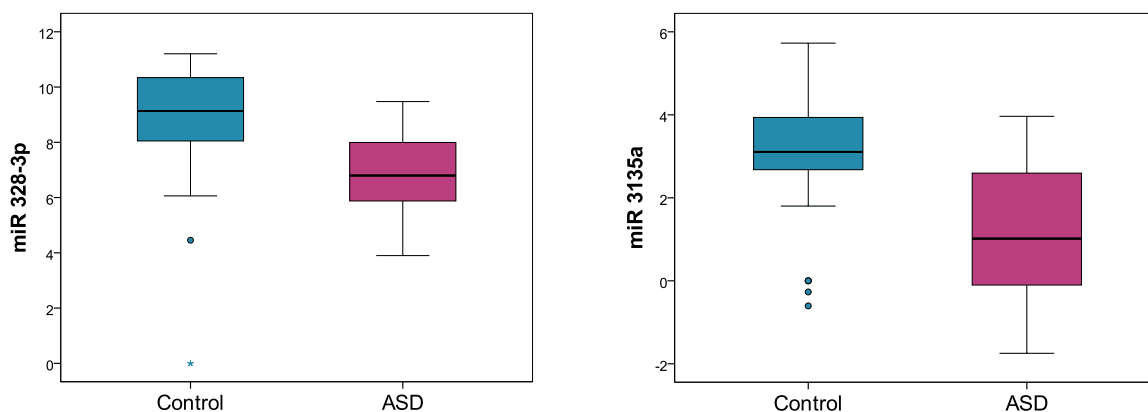
Фигура 6. Разлика в нивото на експресия на миРНК-424-5р, миРНК-500а-5р, миРНК-197-5р и миРНК-664а-3р в серумни образци при пациенти с РАС спрямо контролната група от здрави деца. Статистическата значимост е изчислена чрез непараметричен Kolmogorov–Smirnov тест. Нивата на експресия бяха нормализирани спрямо екзогенна контрола Cel-miR-39. Бокс плот диаграмите визуализиращи експресията на съответните миРНК молекули, показват понижена експресия на всички изследвани миРНК молекули при деца с диагноза РАС. Стойностите показали отклонения (outliers) са показани като отделни точки.

От четирите изследвани миРНК молекули, миРНК-197-5р и миРНК-664а-3р се характеризират с най-висока стойност на понижена експресия при пациенти с РАС, съответно със 19,5% и 21,1% от техните съответни стойности при контролната група (**Фигура 7**). От своя страна, миРНК-197-5р е около два пъти (приблизително 46%) по слабо експресирана при деца с диагноза РАС, в сравнение със здрави деца, докато за миРНК-664а-3р този ефект е по-слабо изразен, приблизително 60%. Изследвания експресионен модел се оказва доста стабилен при различните индивиди, както се вижда от малката вариабилност на данните (грешка в стандартното отклонение). По-широк диапазон на относителното ниво на експресия се наблюдава само в случай на миРНК-664а-3р.

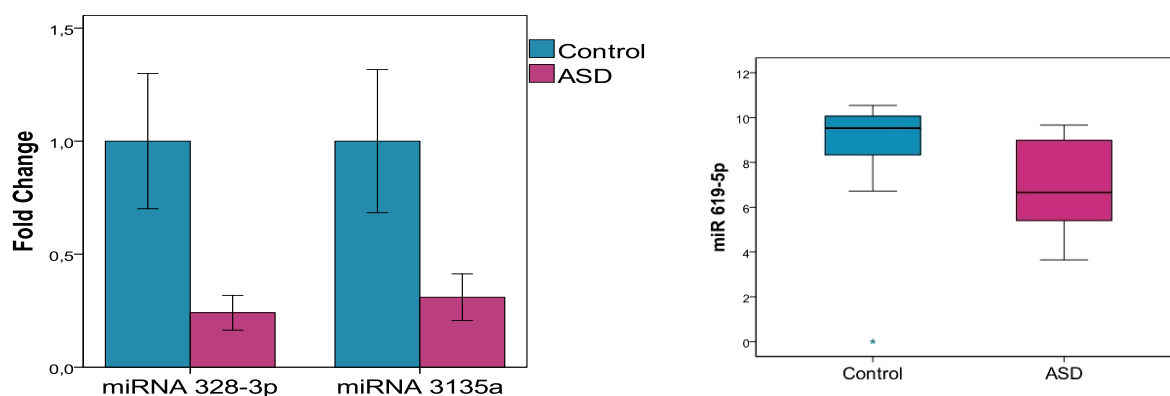


Фигура 7. Разлика в нивото на експресия на на четири от изследваните миРНК молекули между РАС и контролна група от здрави деца. Данните са представени като пъти промяна (*Fold change*) на средната стойност на $2^{-\Delta\Delta C_t}$ за всяка миРНК, след като е извършена нормализация с екзогенна (*spike-in*) Cel-miR-39 контрола.

В други изследвания на серумни образци при РАС, на експресионен анализ бяха подложени други три миРНК молекули, показали диференциална експресия от NGS данни, миРНК-3135a, миРНК-328-3p и миРНК-619-5p. Експресионният анализ на серумните миРНК-619-5p, миРНК-3135a и миРНК-328-3p, беше извършен с помощта на миРНК специфичен stem-loop qRT-PCR метод. Получените данни демонстрират промени в относителните нива на експресия на миРНК-3135a и миРНК-328-3p, които са значително по-ниски при пациенти с РАС, в сравнение с извадката от контролната група от здрави деца. (**Фигура 8**). Относителните серумни нива на тези миРНК молекули могат да диференцират РАС от здравите контролни пациенти (**Фигура 9**).



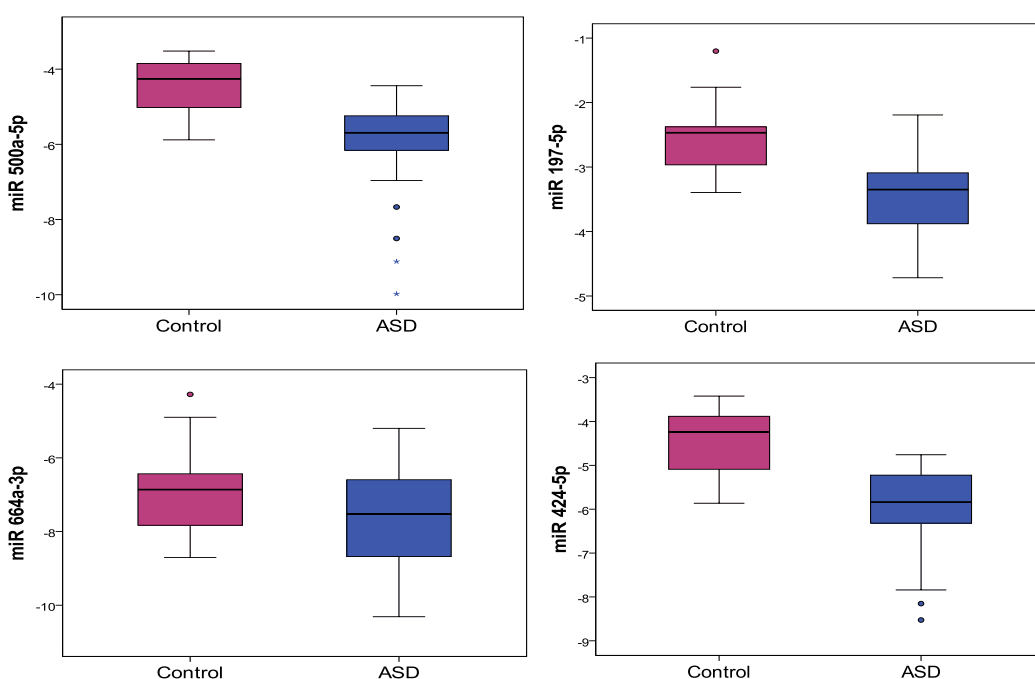
Фигура 8. Диференциална експресия на серумни миРНК молекули при пациенти с РАС в сравнение с контролна група от здрави деца. Представен е резултат от проведен количествен RT-PCR анализ на нивата на миРНК-3135a и миРНК 328-3p. Циркулиращите серумни миРНК молекули са идентифицирани с помоща на миРНК специфичен stem-loop qRT-PCR анализ.



Фигура 9. 9А. Разлика в нивото на експресия на анализирани серумни миРНК молекули (миРНК-3135a и миРНК 328-3p) между групите от пациенти с РАС и контролна група от здрави деца. Данните са представени като пъти разлика на средната стойност на $2^{-\Delta\Delta Ct}$ за всяка миРНК, след като е извършена нормализация с екзогенна (*spike-in*) контрола Cel-miR-39. **9В.** Диференциална експресия на миРНК-619-5p в серум при пациенти с РАС в сравнение с контролна група от здрави деца. Нивото на експресия е нормализирано спрямо екзогенна контрола Cel-miR-39.

V. 10. Резултати от проведените изследвания на кандидат миРНК биомаркерни молекули в периферна кръв при деца с диагноза РАС

За изследване на експресията на кандидат миРНК молекулите в образци от периферна кръв при РАС, също беше използван миРНК специфичен stem-loop qRT-PCR анализ. Относителните нива на експресия бяха определени използвайки U6 като ендогенна контрола, използвана за изчисляване на промените в миРНК експресията с помощта на $2^{-\Delta\Delta C_t}$ метод. Статистически значима промяна в експресията в периферна кръв беше установена за миРНК-424-5р и миРНК-500a-5р (Фигура 10).



Фигура 10. Разлика в нивото на експресия на миРНК-424-5р, миРНК-500a-5р, миРНК-197-5р и миРНК-664a-3р в образци от периферна кръв при пациенти с РАС спрямо контролната група от здрави деца. Бокс плот диаграмите визуализиращи експресията на съответните миРНК молекули, показват понижена експресия на миРНК-424-5-р и миРНК-500a-5р при деца с диагноза РАС.

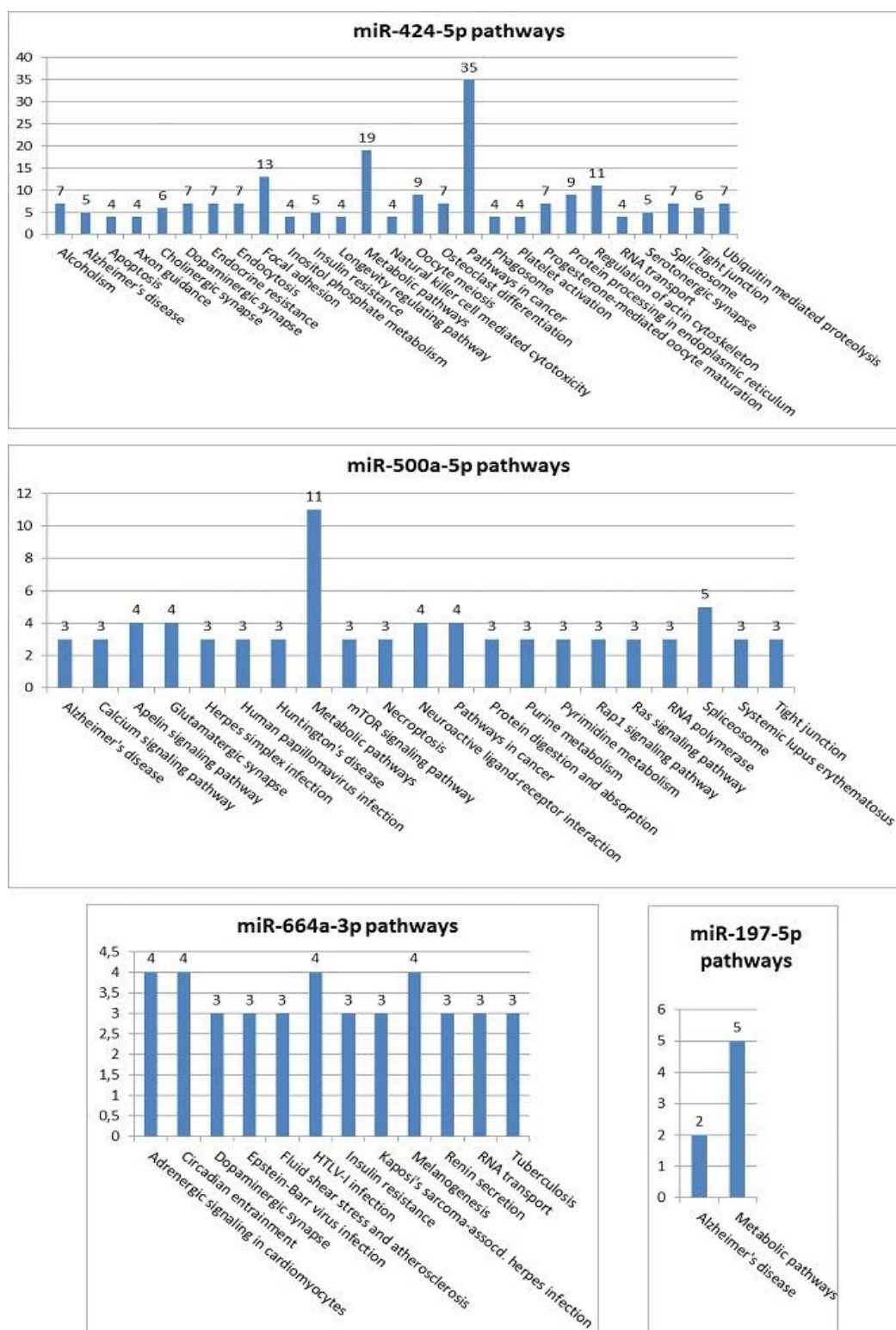
V.11. Идентифициране на биологични пътища свързани с диференциално експресирани миРНК молекули при РАС

Измежду всички 162 гена, обозначени като уникални валидирани таргетни гени на диференциално експресираните миРНК-3135a и миРНК-328-3p в базата данни miRWalk, 71 гена вземат участие в биологични пътища, описани в базата данни Kyoto Encyclopedia of Genes and Genomes (KEGG). Един от целевите гени, получени от проведенят анализ е амилоид бета прекурсорния протеин (APP), ген с участие в синаптичните пътища. Този ген кодира мембранен протеин, който се подлага на протеолитична обработка. В серотонинергичните синапси разтворимият APP алфа фрагмент взаимодейства с протеин на транслацията, стимулирайки невропротекцията. Друго ген, (SLC8A1), участва в специфичен път на екстероцептивна трансдукция. Кодираният от гена SLC8A1 протеин играе роля на $\text{Na}^+/\text{Ca}^{2+}$ (K^+) антипортер. При обонятелните неврони той участва в процеса на мембранна реполяризация като елиминира последствията от предходен потенциален ефект. Други шест целеви гени показват участие в невродегенеративни заболявания, като болестта на Алцхаймер, болестта на Хънтингтън или болестта на Паркинсон. Два гена: APP и BACE1, които се асоциират с патогенезата на болестта на Алцхаймер. Амилоидният прекурсорен протеин, кодиран от APP, се разцепва от бета-секретаза (кодирана от BACE1), което води до получаването на отделен пептид, отговорен за образуването на амилоидна плака. Два белтъка (DNAL1 и POLR2I) са описани като участници в патогенезата на болестта на Хънтингтън по индиректен начин, докато взаимодействат или са повлияни от продукта на гена *huntingtin* (HTT). В резултат на което възникват неспецифични промени в организацията на цитоскелета или в процеса на генната експресия. Два други гена, чиито белтъчни продукти участват в митохондриалната респираторна верига NADH: убихинон оксидоредуктазната субединица A1 (NDUFA1) и NADH: убихинон оксидоредуктазна субединица V3 (NDUFV3), са асоциирани по рано не само при болестта на Алцхаймер, но също така и при болестта на Хънтингтън и Паркинсон. Взаимовръзката между болестта на Алцхаймер и намалената експресия на гените свързани с енергийния метаболизъм е добре установена. Въпреки това, точните механизми, при които NDUFA1 и NDUFV3 гените допринасят за специфично състояние на невровъзпроизводство и тяхната регулаторна роля, все още предстои да бъде изяснена.

Таргетните иРНК молекули на диференциално експресираните миРНКи в серум при РАС; миРНК-424-5p, миРНК-500a-5p, миРНК-197-5p и миРНК-664-3p бяха получени от базата данни miRWalk. Връзката на валидираните таргетни

иРНК молекули със съответните идентифицирани биологични пътища е представена на **Фигура 11**. Анализ на информацията в базата данни MirWalk идентифицира 671 уникални валидирани прицелни гени за всичките четири диференциално експресирани миРНК (35 от които са обект на регулация на от повече от една миРНК молекула). Установено е, че 287 от тези валидирани таргетни гени участват в 478 различни биологични пътища, описани в базата данни на KEGG. Повечето от проучените миРНК молекули са свързани с биологични процеси, като сигнална трансдукция, метаболизъм и ракови заболявания. Показано е, че 19 валидирани таргетни гена участват в синаптични пътища (в холинергични, допаминергични, ГАВАергични или глутаматергични синапси). Сред тях са четири отделни гена, които кодират G-протеинови субединици и по този начин играят важна роля във вътреклетъчната сигнална интеграция: GNAL, GNAQ, GNB1 и GNG12. Допълнителни три гена (GRIN2B, GABRG2 и GABARAPL1) кодират субединици на рецептори за невроактивни лиганди или рецептор-свързани протеини. И накрая, други два таргетни гена участват в обратното усвояване на глутамат: SLC1A2 и SLC1A1.

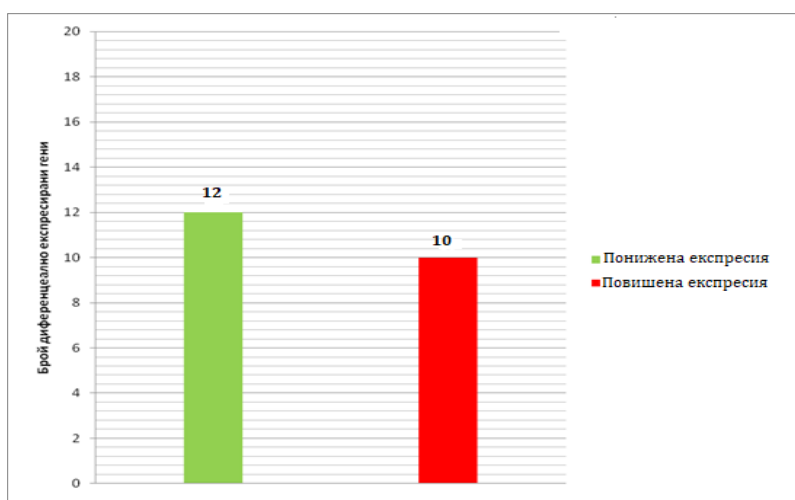
Въпреки, че миРНК-664-3р има множество валидирани таргетни гени, функционалният им профил е разпределен в голям набор от биологични процеси, които не са пряко свързани с ЦНС и с известни ключови аспекти от поведението при РАС. (**Фигура 11**) До известна степен подобна е ситуацията и с миРНК-197-5р, с тази разлика, че панелът на таргетните иРНКи е по-малък. Освен това миРНК-197-5р действително има 2 попадения в категория на болестта на Алцхаймер. Докато прицелните гени на миРНК-424-5р и миРНК-500a-5р имат много по-изразено участие в специфични за нервната тъкан биологични пътища, включително на таргетни гени, засягащи различни видове синапси, насочване на аксони, взаимодействия на невроактивни лиганди и рецептори и някои патологии. (**Фигура 11**)



Фигура 11. Обогащение на биологичните пътища съдържащи таргетни гени на диференциално експресирани миРНК: миРНК-424-5p; миРНК-500a-5p; миРНК-664-3p и миРНК-197-5p.

V.12. Идентифициране на диференциално експресирани белтък-кодиращи гени в анализирани групи чрез секвениране на РНК (RNA-Sequencing)

За идентифицирането на диференциално експресирани гени беше използвана p -стойността като показател за диференциална генната експресия, FDR стойността (False Discovery Rate), като метод за определяне на прага на p -стойността и абсолютната стойност на съотношението $\log_2 \geq 1$, като праг за преценка на разликата в генната експресия. Всички анализирани в транскриптомният анализ информационни РНК молекули (транскрипти) в пул пробата на децата с PAC са съпоставени със същите на контролната групата. Двадесет и два РНК транскрипта показаха диференциална експресия със статистическа значимост $p < 0,001$, $FDR \leq 0.001$ и $\log_2 \geq 1$.



Фигура 12. Гени показващи променена експресия със статистическа. Двадесет и два РНК транскрипта показаха диференциална експресия със статистическа значимост $p < 0,001$, $FDR \leq 0.001$ и $\log_2 \geq 1$. При десет гена се установи повишено ниво на експресия, а 12 гена показаха понижени нива на експресия.

V.13. Генно онтологичен анализ на диференциално експресирани гени при РАС

За категоризиране на установените диференциално експресирани гени, беше използван генно онтологичен анализ, като международна стандартизирана класификация за генна функция, която предлага контролирана и динамично-обновяваща се база данни със стриктно дефинирани свойства на гените и техните продукти.

Функционална класификация на диференциално експесираните гени в генно онтологичния аспект клетъчен компонент.

Три от анализираните белтъл-кодиращи гена са асоциирани с функцията в определена част на клетката: FLNC с понижена експресия и TPX2, и DPM3 с повишена експресия. Един ген показва връзка с екстрацелуларния регион - SERPINE1, с понижена експресия. Четири от гените имат участие в макромолекулни белтъчни комплекси: HLA-DQA2, TPX2 и DPM3 с повишена експресия; в рибонуклеопротеинов комплекс - IGF2BP1 с понижена експресия. Свързани с клетъчната мембрана са два от гените: TM4SF1 с понижена експресия и DPM3 с повишена експресия. С органелите са свързани три гена: цитоскелета - FLNC с понижена експресия и TPX2 с повишена експресия; ендоплазмения ретикулум - DPM3 с повишена експресия.

Функционална класификация на диференциално експесираните гени в генно онтологичния аспект молекулярна функция.

Молекулната функция на шест от диференциално експесираните гени се свързва със - “binding” – процеса на свързване (селективното, нековалентно, често стехиометрично, взаимодействието на молекулата с едно или повече специфични места на друга молекула) - FOSL1 свързване с ДНК и активност на транскрипционен фактор; IGF2BP1 свързване с РНК; FLNC и TPX2 свързване с протеини на цитоскелета; CRB2 - показва рецепторна активност. Седем гена с каталитична молекулярна функция: - регулаторна активност - SERPINE1, IGF2BP1 и RAP1GAP; хидролазна активност - SERPINE1, ANKRD22 и LTF; лигазна активност TRIM17; трансферазна активност - DPM3. Два гена имат молекулярна функция като ензимни регулатори - SERPINE1 (инхибитор на ензимна активност) и RAP1GAP (регулира активността на малка GTP-аза). Молекулната

функция на един ген - FOSL1, като транскрипционен фактор. Други два гена имат молекулярна функция като структурна молекула (действието на молекула, която допринася за структурната цялост на комплекс в или извън клетката) - FLNC и TPX2. SLC8A2, ген свързан с трансмембрана транспортна активност. Функционална класификация на диференциално експресирани гени в генно онтологичния аспект биологични процеси

Два от гените се свързват с процеси индуциращи апоптоза - IGF2BP1 и MAGEA4 и двата с намалена експресия. В процеси свързани с биологична регулация са SERPINE1, FOSL1 с намалена експресия и RAP1GAP и CRB2 с повишена. FLNC участва в организация на клетъчни компоненти и има намалена експресия. Девет гена се асоциират с участие в клетъчни процеси – SLC8A2, RAP1GAP, CRB2, DPM3 и TPX2 с повишена експресия и FOSL1, MAGEA4, FLNC и IGF2BP1 с понижена експресия. В развитийни процеси участие имат четири гена - MAGEA4, FLNC, и IGF2BP1 с понижена експресия и CRB2 с повишена. Ген онтологичния анализ показва връзка на два гена с процеси на имунната система - CRB2 и HLA-DQA-2 с повишена експресия. Връзка с GO термина локализация (всеки процес, при който клетка, вещество, или клетъчна единица, като протеинов комплекс или органел, се транспортира до или поддържа на определено място) показаха гените - SLC8A2, CRB2 и LTF с повишена експресия и IGF2BP1 с понижена. В метаболитни процеси вземат участие - RAP1GAP, CRB2, ANKRD22, DPM3, и LTF, които са с повишена експресия и FOSL1, SERPINE1, IGF2BP1 и TRIM17 с понижена. IGF2BP1 и OTOF показаха връзка с процеси характерни за многоклетъчните организми. В процеси отговор на стимули връзка има IGF2BP1 гена, който показва намалена експресия.

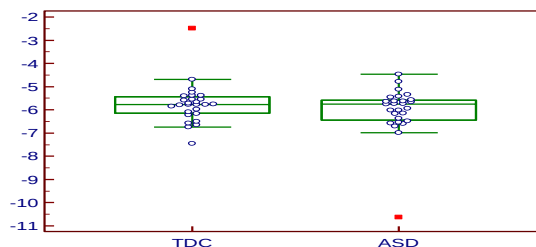
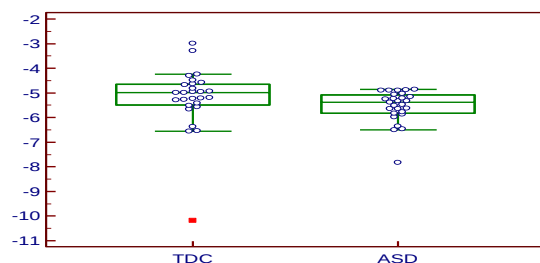
V.14. Потвърждаване на диференциално експресирани гени от транскриптомния (RNA-Seq) анализ чрез провеждане на количествен Real Time PCR анализ

Валидиране на данните от проведения транскриптомен анализ се осъществи с помощта на количествен Real-Time PCR анализ. На индивидуален количествен Real-Time PCR анализ бяха подложени 10-те диференциално експресирани гени с най добри показатели (FDR и *p*-стойност) във всички проби от децата с диагноза

РАС и контролна група от здрави деца. Промяната в генната експресия е анализирана чрез използване на $2^{-\Delta\Delta CT}$ метод.

V.15. Резултат от количественият експресионен анализ (RT-qPCR) на гените (AGO2, AGO3, AGO4 и Drosha) при РАС

Интереса в настоящата работа също беше насочен към провеждане на сравнителен експресионен анализ на гени участващи в процеса на миРНК-

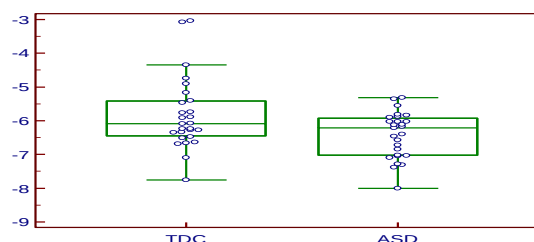


био
гене
за
при
дец
а с

РАС.

Фигура 13. Бокс плот диаграма на **Ago2** гена при пациенти с диагноза РАС и контролна група деца. Проведеният статистическият анализ показва понижена експресия на **Ago2** гена при РАС, но без

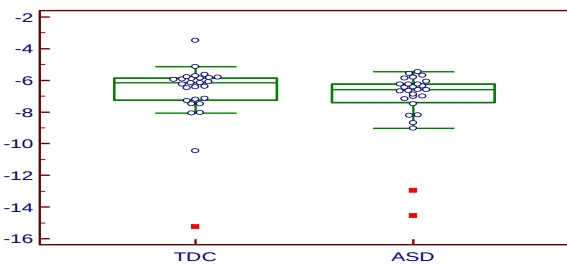
статистическа значима промяна. Бокс плот диаграма на **Ago3** гена при пациенти с диагноза РАС и контролна група деца. Проведеният статистическият анализ показва понижена експресия на **Ago3** гена при РАС но без статистическа значима промяна.



Фигура 14. Бокс плот диаграма на **Drosha** гена при пациенти с диагноза РАС и контролна група деца. Проведеният статистическият анализ показва понижена експресия на **Drosha** гена при РАС, със статистическа значима промяна. Бокс плот диаграма на **Ago4** гена при пациенти с диагноза РАС и контролна група деца. Проведеният статистическият анализ показва повишена експресия на **Ago4** гена но без статистическа значима промяна.

V.16. KEGG анализ на биологичните пътищата обогатени с диференциално експресирани гени при РАС

Киото-енциклопедия за гени и геноми (KEGG) е основната база данни с публичен достъп свързана с данни за биологични пътища и генно взаимодействие. Анализът на биологичните пътища с участието на диференциално експресирани гени при РАС идентифицира метаболитни и сигнални пътища свързани с диференциално експресирани гени. В **Таблица 1** са представени биологичните пътища със статистическа значимост, в които вземат участие установените от настоящото изследване диференциално експресирани гени.

		Диференциално експресирани гени (17)		P-value	ID на пътя
		3 (17.65%)	0.002537809	ko04020	
2	Amphetamine addiction	2 (11.76%)	0.007751596	ko05031	
3	Leishmaniasis	2 (11.76%)	0.007983234	ko05140	
4	GABAergic synapse	2 (11.76%)	0.008100223	ko04727	
5	Retrograde endocannabinoid signaling	2 (11.76%)	0.008575949	ko04723	
6	MAPK signaling pathway	3 (17.65%)	0.009102987	ko04010	

7	Arrhythmogenic right ventricular cardiomyopathy (ARVC)	2 (11.76%)	0.01196498	ko05412
8	Cholinergic synapse	2 (11.76%)	0.01253102	ko04725
9	GnRH signaling pathway	2 (11.76%)	0.01506537	ko04912
10	Glutamatergic synapse	2 (11.76%)	0.01553389	ko04724
11	Serotonergic synapse	2 (11.76%)	0.01813789	ko04726
12	Dopaminergic synapse	2 (11.76%)	0.01864557	ko04728
13	Steroid biosynthesis	1 (5.88%)	0.02947291	ko00100
14	Wnt signaling pathway	2 (11.76%)	0.0308568	ko04310
15	Leukocyte transendothelial migration	2 (11.76%)	0.03652306	ko04670
6	Hypertrophic cardiomyopathy (HCM)	2 (11.76%)	0.04232472	ko05410
17	Dilated cardiomyopathy	2 (11.76%)	0.04449439	ko05414
18	Asthma	1 (5.88%)	0.04538188	ko05310
19	Influenza A	2 (11.76%)	0.04547207	ko05164
20	Carbohydrate digestion and absorption	2 (11.76%)	0.04895749	ko04973
21	Endocrine and other factor-regulated calcium reabsorption	2 (11.76%)	0.04921018	ko04961

Таблица 1. Биологични пътища с участието на диференциално експресираните гени при РАС.

V.17. Идентифициране на нова митохондриална мутация в гена, кодиращ цитохром с III оксидаза (CO3) при деца с РАС с помощта на секвениране на РНК от следващо поколение (NGS RNA sequencing)

Благодарение на съвременният напредък в молекулярната генетика, развитието и прилагането през последните години на нови технологии за

секвениране от следващо поколение на ДНК и РНК позволи откриването на голям брой единични нуклеотидни полиморфизми (SNP), в периферна кръв чрез използването на транскриптомни данни, които се характеризират с висока чувствителност и специфичност, като функция от броя секвенирани фрагменти (reads) получени от съответният геномен регион. Аналитичният подход представен в настоящата работа е приложим за идентифициране на ДНК варианти при заболявания и разстройства за които се наблюдава хетерогенност в определени локуси в ДНК. Подобен ярък пример представляват и мутациите в митохондриалния геном, поради характерната хетероплазмията на митохондриалната ДНК. От всички анализирани варианти като най-сериозен кандидат с най-високо качество на детекция (QUAL = 13461.3) се оказва локусът chrM:9214. Идентифицираната мутация води до замяна на аминокиселината хистидин с аргинин на трета позиция в молекулата на митохондриално-кодирания белтък цитохром с оксидаза субединица III.

В настоящото изследване беше идентифицирана точкова мутация на позиция 9214 в митохондриалния геном при деца с РАС, като използваме метода на пълното транскриптомно секвениране. Посочената мутация: m.9214A>G присъства преимуществено в комплексната проба на децата с РАС (при 22% от всички картирани фрагменти), докато в пробата от здравите деца почти напълно отсъства. Проведеният анализ *in silico* на прогнозния фенотипен ефект определи изследваната нуклеотидна замяна като потенциално увреждаща. Направените наблюдения ни дават основания да предположим евентуално участие на изследваната мутация и последствията от нея в механизмите на етиопатогеназа в определени случаи на РАС.

V.18. ROC анализ (receiver operating characteristic) за определяне на диагностична точност на изследваните диференциално експресирани серумни миРНК при РАС

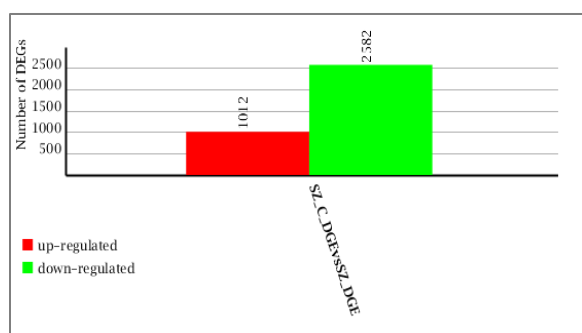
С цел оценка на потенциала на анализирани миРНК молекули в серумни образци като потенциални биомаркери при РАС, беше извършен ROC (receiver operating characteristic) анализ който имаше за цел да определи диагностичната чувствителност и специфичност на всяка от изследваните миРНК молекули. За да се оцени характеристиката на диференциално експресирани миРНК-500a-5p, миРНК-197-5p и миРНК-664-3p и миРНК-424-5p като потенциални серумни биомаркери при РАС, беше извършен ROC анализ за определяне на диагностичната точност на изследваните миРНК. В допълнение, точността на прогнозиране на миРНК панела, базиран на всички изследвани миРНКи беше оценена чрез изготвяне на комбинирана (адитивна) ROC-крива. Когато обект на изследване е индивидуалното определяне на диагностичната чувствителност за

всяка серумна миРНК, резултатите са както следва: 86,1% за миРНК-197-5р; 77,8% за миРНК-500а-5р; 25% за миРНК-664-3р и 88,9% за миРНК-424-5р. От своя страна индивидуалният анализ на специфичността на изследваните кандидат биомаркери показва специфичност от 78,6% за миРНК-197-5р, 92,9% за миРНК-500а-5р, 100% за миРНК-664-3р и съответно 75% за миРНК-424-5р. Получените параметри на AUC за миРНК-197-5р, миРНК-500а-5р и миРНК-424-5р бяха съответно 0.825, 0.796 и 0.756, докато за миРНК-664-3р AUC стойността е определена под 0.7. Според така получените резултати, миРНК-197-5р, миРНК-500а-5р и миРНК-424-5р могат да бъдат определени като добри диагностични биомаркери при РАС, с диагностична чувствителност (т.е. процент на истински положителни) близо или над 80%, и съответната висока специфичност (т.е. процент на истински отрицателни) и високи AUC стойности (т.е. точност). МиРНК-664-3р се характеризира с по-ниска диагностична ефективност, тъй като има много по-малка вероятност за откриване в проби от пациенти с РАС като показва само 25% чувствителност и AUC <0.7. Въпреки това, експресионният профил на тази миРНК в серум може да се използва за да се изключи възможността за развитие на РАС при здрави индивиди, поради демонстрираната от тази миРНК изключително високата специфичност (100%). Данните за експресионният профил на четирите серумни миРНК, описани в изследването, бяха едновременно взети под внимание за провеждане на комбиниран (адитивен) ROC анализ. Интегрирането на резултатите за миРНК-197-5р и миРНК-500а-5р повиши стойността на AUC до 0,975 и показва изключително висока стойност на диагностична чувствителност на комбинираните класификатори, миРНК-197-5р и миРНК-500а-5р (адитивен ROC анализ) от 97,4% със съответна специфичност от 92,6%. Оценката на данните за чувствителност и специфичност на получения диагностичен панел бяха впечатляващи, тъй като и при двете изследвани миРНК молекули стойностите надвишаваха 90%. Експресионните нива на миРНК-424-5р и миРНК-664-3р не показаха промени, които биха могли да ги отнесат към този панел поради което те бяха изключени от него. Вzeti заедно всички тези резултати демонстрират, че когато се разглеждат самостоятелно, три от избраните класификатори, а именно миРНК-197-5р, миРНК-500а-5р и миРНК-424-5р, могат да дискриминират случаите на деца с диагноза РАС от здрави контроли с достатъчно висока точност. Докато обаче, най-висока точност за диагностициране на пациенти с РАС се постига чрез панела, който е изграден въз основа на данни за експресионните нива на най-добрите класификатори установени от проведеното изследване, миРНК-197-5р и миРНК-500а-5р. В друго проведено изследване на серумни образци при РАС, на експресионен анализ бяха подложени други две миРНК молекули (миРНК-3135а и миРНК-328-3р), показали диференциална експресия от NGS данни, Проведеният ROC анализ на експресионните данни показва добри диагностични характеристики на изследваните предиктори.

Диагностичната чувствителност на миРНК-3135а и миРНК-328-3р беше определена съответно на 76,3 %, и 78,9 %, със съответна специфичност оценена на 88,9% и 88,9% и площ под ROC кривата както следва за миРНК-3135а ($AUC = 0,828$) и миРНК-328-3р ($AUC = 0,858$). Проведеният комбинирана ROC анализ (LOGREGR) на диференциално експресирани миРНК молекули (миРНК-3135а и миРНК-328-3р) показва диагностична чувствителност на комбинираните класификатори от 78,9%, със съответна специфичност от 88,9%. От получените данни става ясно, че комплексната характеристика на няколко миРНК молекули като потенциални биомаркери при РАС показва по добри диагностични характеристики в сравнение с използването на единични такива. Допълнително други две кандидат миРНК бяха подложени на ROC анализ, миРНК-619-5р и миРНК-365а-3р. Проведеният ROC анализ на миРНК-619-5р показва добра диагностична чувствителност от 100 %, но със съответно ниска специфичност от 55,6%, докато анализа на миРНК-365а-3р не показва задоволителни резултати.

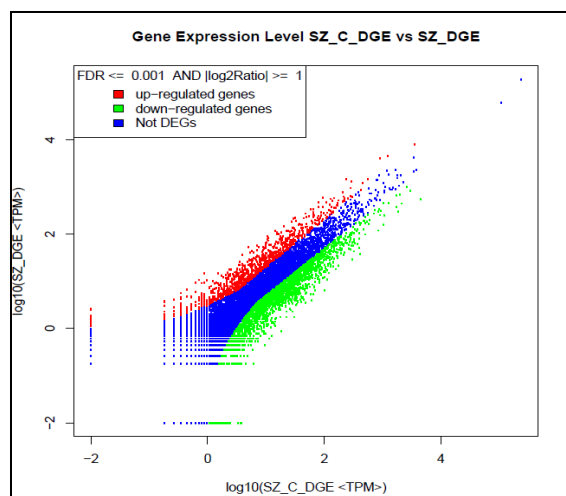
V.19. ROC анализ (receiver operating characteristic) за определяне на диагностична точност на изследваните миРНК биомаркери в образци от периферна кръв при РАС

За оценка на потенциала на изследваните миРНК молекули като потенциални биомаркери в алтернативен образец (периферна кръв) при РАС, беше извършен ROC (receiver operating characteristic) анализ който имаше за цел да определи диагностичната чувствителност и специфичност на всяка от изследваните миРНК-424-5р, миРНК-500а-5р, миРНК-197-5р и миРНК-664а-3р молекули.



V.20. Резултати от проведеното транскриптомно профилиране на белтък-кодиращи гени в периферна кръв при шизофрения

След проведения широкомащабен експресионен анализ (DGE) бяха идентифицирани значителен брой диференциално експресирани гени при шизофрения. Гените показали понижена експресия бяха 1012, докато тези показали повишена бяха значително повече 2582 гена (Фигура 15 и Фигура 16).



Фигура 15. Експресионен профил на диференцирано експресирани гени между сборни проби от пациенти с диагноза шизофрения и контролна група лица.

Фигура 16. Диференциално експресирани гени при пациенти с диагноза шизофрения и контролна група лица. Диференциално експресираните гени са представени в червено (с повишена експресия) и зелено (с понижена експресия). Гените, които не демонстрират промени в експресията, са посочени в синьо; DEG, диференцирано експресирани гени; TPM; тагове на милион; FDR, степен на фалшиво откриване.

Допаминов (DA) D4 рецептор (DRD4)

Допаминовият (DA) D4 рецептор (DRD4) играе роля в медирането на допаминергичната активност.. Допаминовият рецептор е отговорен за невронната сигнализация в мезолимбичната система на мозъка, зона на мозъка, която регулира емоцията и комплексното поведение. Активира се от допамин, но също

и от епинефрин и норепинефрин, и от множество синтетични агонисти и лекарства.

Fasciculation and elongation protein ζ -1 (FEZ1)

Fasciculation and elongation protein ζ -1 (FEZ1) е един от първите идентифицирани свързващи партньори на протеина кодиран от DISC1 гена, ген на чувствителност за основни психични разстройства, включително шизофрения в дрождени two-hybrid системи на библиотеки получени от мозъка на възрастен човек.

Dicer 1, рибонуклеаза III (DICER1)

Генът DICER1 кодира протеин, който играе роля в регулирането на активността (експресията) на други гени. Чрез тази си роля за регулиране на експресията на гените, Dicer участва в редица процеси, включително клетъчен растеж и деление (пролиферация) и съзряване на клетките, за поемане на специализирани функции (диференциация).

ASTN2 астротактин 2

ASTN2 кодира протеин, който се експресира в мозъка с роля при миграцията на неврони, въз основа на функционални изследвания на базата на астротактин 1 гена при хора и мишки. Делеция в тази област е асоциирана с шизофрения.

CASP3 каспаза 3, свързана с апоптоза цистеин пептидаза

Този ген кодира протеин, който е член на семейството протеаза (каспаза) на цистеин-аспарагинова киселина.

CDT1 chromatin licensing and DNA replication factor 1

Протеинът, кодиран от този ген, участва в образуването на предрепликационния комплекс, необходим за репликацията на ДНК. Кодираният протеин може да свързва геминин, който предотвратява репликацията и може да функционира, за да попречи на този протеин да инициира репликация на неподходящо място. Фосфорилирането на този протеин чрез циклин А-зависими кинази води до разграждане на протеина.

ANK3 ankyrin 3, node of Ranvier (ankyrin G)

Анкирините са семейство от протеини, за които се смята, че свързват интегралните протеини на мембраната към цитоскелета и играят ключова роля в

дейности като клетъчна подвижност, активиране, пролиферация, контакт и поддържане на специализирани мембранни домени.

V.21. ROC анализ (receiver operating characteristic) за определяне на диагностична точност на изследваните диференциално експресираните гени при шизофрения

За да се оцени характеристиката на диференциално експресираните гени, DICER, FEZ1, DRD4, GRIN2 като потенциални биомаркери при шизофрения, беше извършен ROC анализ за определяне на диагностичната точност на изследваните белтък-кодиращи гени. Диагностичната чувствителност на DICER, FEZ1, DRD4 и GRIN2D гените беше определена съответно на 72,4%, 55,2%, 86,2% и 48,3%, със съответна специфичност оценена на 83,3%, 87,5%, 91,7% и 83,3%.

VI. ДИСКУСИЯ

Представената в настоящата дисертация научна разработка е първото по рода си комплексно научно изследване обхващащо сравнителни геномни транскриптомни и протеомни изследвания, проведени върху клинично подбрана група от български пациенти с диагноза РАС в сравнение с контролна група от здрави лица. Едно от най-големите предизвикателства при провеждането на геномни, транскриптомни и протеомни изследвания при невроразвитийните разстройства е събирането на проби от първичната прицелна тъкан, т.е. мозък. Поради невъзможността за извършване на изследванията на мозък *in vivo*, за тази цел, се провеждат основно ретроспективни проучвания с *post-mortem* мозъчни образци. Причината за избора на този вид материал е обоснована и се изразява в това, че заболявания като шизофрения и РАС, засягат мозъчната функция и най-аргументирано би било евентуалните транскриптомни и протеомни изследвания да бъдат насочени именно към засегнатите и ангажирани със разстройствата тъкани. Подобен подход би имал голяма стойност при идентифицирането на етиопатогенетичните фактори в развитието на разстройствата/заболяванията, които към момента все още са с ненапълно изяснена етиология.

Към днешна дата, редица кандидат гени са изследвани в търсенето на евентуална връзка със заболяванията. Изборът на кандидат гени е основно базиран на данни получени от анализи на скаченост/цитогенетични данни или наличието на специфични констатации на клинично равнище при пациенти с РАС и шизофрения. Един от основните методи за идентифициране на кандидат гени при

невроразвитийните разстройства е директно секвениране на екзонни секвенции при пациенти, и сравнението им с контролна група лица. В допълнение, изучаването на посттранскрипционната регулация при невроразвитийните разстройства шизофрения чрез анализ на некодиращи РНК допълнително (например, миРНК молекули) може да се осигури необходимата връзка между податливостта/чувствителността към разстройството и включените в процеса потенциални епигенетични фактори. Редица наблюдения могат да бъдат използвани в подкрепа на допускането, че миРНК са молекули, които би било рационално да бъдат изследвани и евентуално обвързани с патогенезата на невроразвитийните разстройства като РАС и шизофрения. Някои миРНК и други малки РНК показват коекспресия в периферна кръв и мозък, както има и миРНК представени в мозъчни тъкани с по висока експресия но също експресирани и в периферна кръв. Настоящото проучване имаше за цел да охарактеризира глобалният профил на миРНК експресията в периферна кръв, с фокус върху идентифицирането на специфични миРНК молекули с характеристики на потенциални биомаркери при невроразвитийните разстройства (РАС и шизофрения).

В настоящата експериментална работа детайлно беше изследван общият експресионен профил на миРНК молекули в периферна кръв от пациенти с РАС. Чрез използването на секвениране на малки РНК молекули, беше установена промяна в експресията на голям брой миРНК молекули. Използваният в настоящата работа подход предлага възможност да се направи стъпка в разбирането на ролята на миРНК молекулите и свързаните с тях биологични процеси в етиопатогенезата на РАС. Към момента редица проучвания демонстрират, че нарушенията в невроналната пластичност и функция при невроразвитийните разстройства могат да бъдат свързани с нарушения в миРНК медираната пост-транскрипционна регулация. Използването на аутопсирана мозъчна тъкан има няколко ограничения, които включват наличността на материал за изследване, периодът на потенциалното забавяне между смърт и пробовземане, при които могат да възникнат бързи клетъчни и молекулни промени. В много случаи липсата на история на прилаганата медикаментозна терапия и трудността за разграничаване на последиците в резултат на заболяването от други съпътстващи фактори представлява допълнително ограничение. Поради наличието на тези комплициращи фактори, нов акцент в изследванията на биомаркери при шизофрения и РАС е подходът базиран на изследване на молекулярния профил на биомолекули от периферни тъкани и по специално миРНК молекули в периферна кръв, в качеството и на динамична и обективно променяща се система, отразяваща физиологичното състояние. В представеното проучване беше използван миРНК специфичен qRT-

PCR анализ, с цел идентифициране на миРНК, молекули с евентуална роля в патогенезата на РАС. Въпреки факта, че ние не знаем дали наблюдаваната дисрегулация на миРНК молекулите също е характерна и за ЦНС на изследваните пациенти, по-голямата част от анализиранияте миРНК показват експресия в мозък съгласно литературни данни. Промените в посттранскрипционните регулаторни механизми в периферна кръв наблюдавани при пациенти с шизофрения имат не само отношение към патофизиологията на заболяването, но също така хвърлят светлина върху един подходящ обект за идентифициране и валидиране на потенциални биомаркери също така и при РАС. Друг аспект в насоките при изследване на връзките между дисрегулацията на миРНК молекулите и невроразвитийните разстройства са мутационните промени, т.е., генетичната връзка, която може да има за резултат промяна в транскрипцията и процесинга на миРНК молекули. МиРНК гените в по-голямата си част се транскрибират от промоторни секвенции специфични за РНК полимераза II, така те също подлежат на регулация от транскрипционни фактори и структурата на хроматина по подобен на информационните РНК молекули начин. МиРНК молекулите се транскрибират обикновено под формата на некодиращи РНК молекули или са транскрибирани от интронни секвенции на протеин-кодиращи информационни РНК молекули. Като такива, степента на транскрипция на миРНК молекулите би се влияла, както по отношение на преобладаващата регулаторната среда, така и според специфичният епигенетичен профил, което е характерно и за всеки белтък-кодиращ ген. Транскрипция на миРНК молекулите също зависи от целостта на консервативните последователности (мотиви) в техните регулаторни (промоторни и енхсерни) елементи. След транскрипция, процеингът до получаването на прекурсор и зряла миРНК молекула също ще е функция от целостта на последователности в първичният транскрипт, която направлява ензимната система за биогенезата на зрели миРНК молекули. С всичко изброено по горе, изследването на детайлните механизми на регулация на миРНК експресията остава обещаваща перспектива в изследванията в областта на невроразвитийните разстройства.

VI. 1. Обсъждане на резултатите получени от проведеното мащабно екзомно секвениране на ДНК (*Whole-Exome Sequencing*) за изследване на белтък-кодиращи последователности в човешкия геном асоциирани със РАС

Представеният анализ може да бъде разглеждан като предварително пилотно изследване което би изисквало разширяване на извадката както и по следващо индивидуално валидиране на установените кандидат генетични варианти при РАС. През последните години се наблюдава значително обогатяване

на информацията относно спектъра и честотата на генетичните изменения в генома, свързани с РАС. Независимо обаче от постигнатият напредък, голяма част от генетичния принос към етиопатогенезата на заболяването остава неизяснен. В настоящото проучване, чрез метода на екзомно секвениране от ново поколение е установено присъствие на значителен брой генетични варианти в генома при пациенти с диагноза РАС в сравнение с клинично здрави деца. Високата информативност на използвания в настоящия дисертационен труд алгоритъм за оценка на генетични варианти, би могъл да послужи за детайлизиране на разбиранията ни за етиологията и патогенезата на заболяването, както и в бъдещ план за подобряване на диагностиката на РАС.

В тази насока, бъдещите изследвания в областта биха могли да се фокусират по отношение на:

- Провеждане на индивидуален анализ на *de novo* идентифицираните, патологични варианти при деца с РАС.
- Разширяване обхвата на изследваните групи от пациенти с РАС и на клинично здрави деца от българска популация.
- Провеждане на комплексни функционални анализи на *de novo* идентифицираните, патологични варианти и за вариантите с неизвестно клинично значение в генома на пациенти с РАС.

VI. 2. Обсъждане на резултатите получени от проведените експресионни изследвания на миРНК при РАС

В настоящото изследване част от представените резултати обхващат проведени експресионни изследвания на серумни миРНК молекули при индивидуален анализ на пациенти с диагноза РАС в сравнение с контролна група от здрави деца. В проведено предварително изследване на експресионен анализ бяха подложени всички миРНК молекули показали диференциална експресия от проведеният РНК транскриптомен анализ на миРНК молекули (Small RNA Seq).

В проведените предварителни експресионни изследвания върху комплексни (пул) проби от анализирани групи, 42 миРНК молекули показаха обещаващи резултати като кандидат биомаркери за първоначален анализ. От тях бяха идентифицирани общо 29 миРНК с понижена експресия (миРНК-589-3р, миРНК-6849-3р, миРНК-3135а, миРНК-15а-5р, миРНК-328-3р, миРНК-183-5р, миРНК-3674, миРНК-96-5р, миРНК-3687, миРНК-6799, миРНК-3р, миРНК-587-3р, миРНК-504-5р, миРНК-576-5р, миРНК-486-3р, миРНК-let-7i-3р, миРНК-29с-5р, миРНК-301а-3р, миРНК-3064-5р, миРНК-145-5р, миРНК-424-5р, миРНК-193b-3р, миРНК-487b-3р, миРНК-197-5р, миРНК-500а-5р, миРНК-664b-3р, миРНК-20b-3р, миРНК-671-3р и миРНК-199а-5р) 11 показаха повишена експресия (миРНК-4489,

миРНК-8052, миРНК-106b-5p, миРНК-1423p, миРНК-3620-3p, миРНК-365a-3p, миРНК-664a-3p, миРНК-374b-5p, миРНК-18b-3p, миРНК-619-5p и миРНК-210-5p), докато две не показаха промени в експресията при деца с РАС в сравнение с контролна група деца от общата популация.

Получените резултати в резюме могат да бъдат обобщени както следва. За една от тях, миРНК-365a-3p, не беше установена статистически значима разлика в експресията между пациентите с РАС и контролна група от здрави деца. За останалите четири: миРНК-500a-5p, миРНК-197-5p, миРНК-424-5p и миРНК-664a-3p, е установена понижена експресия, първоначално наблюдавана в сборни проби при РАС, а в последствие потвърдено и при проведеният индивидуален експресионен анализ на пациенти с диагноза РАС, включени в настоящото изследване. Трябва да се отбележи факта, че миРНК-424-5p и миРНК-500a-5p медираната регулация на генната експресия изглежда има значително въздействие върху процеси в ЦНС, тъй като много от техните валидирани иРНК вземат активно участие в процесите, протичащи в мозъка. Обратно, не се установява подобна видима връзка между миРНК-197-5p, миРНК-664a-3p и ЦНС, което може частично да се обясни с по-малкия брой засегнати иРНК молекули и по-слабото им участие в по-специфични KEGG пътища.

МиРНК-500a-5p показва участие при различни злокачествени заболявания и по-специфично карцином на черния дроб. През 2009 г. миРНК-500a-5p е предложена като диагностичен маркер за хепатоцелуларен карцином (ХЦК) след направените заключения относно повишената експресия в ракови клетъчни линии, тъкан от ХЦК и серум. Особен интерес привлича също и факта, че серумните нива на миРНК-500a-5p се връща към норма при пациенти с ХЦК след извършена хирургическа интервенция за отстраняване на карцинома [Yamamoto et al., 2009]. Няколко други изследвания демонстрират абнормни нива на миРНК-500a-5p и връзката и с туморната патология. Пример за това е, че тази миРНК е една от десетте миРНК молекули, свързани с общата преживяемост в случаи на недребноклетъчен рак на белия дроб (NSCLC). По подобен начин, повишената експресия на миРНК-500a-5p се свързва и с повишена клетъчна пролиферация и лоша прогноза, както при рак на простатата така и при рак на стомаха.

В представеното в дисертацията изследване, се потвърждава промененият профил на експресия на миРНК-500a-5p в серум от пациенти с РАС, но за разлика от ХИВ-асоциираните неврокогнитивни разстройства, в този случай миРНК-500a-5p показва понижена експресия.

За разлика от изследванията проведени с миРНК-500a-5p и обема от налични данни, проучванията, които са свързани конкретно с миРНК-197-5p, са значително по-ограничени. Забележително е, че в две от проучванията се разкрива ролята на миРНК-197-5p при неврологични заболявания. Първата е посветена на изследване върху болестта на Алцхаймер с късното начало (late-onset Alzheimer's

disease (LOAD)) и представя данни, показващи повишена експресия на миРНК-197-5p в засегнатия от заболяването главен мозък. В други изследвания миРНК-197-5p показва значително понижена експресия в плазмени екзосоми при пациенти с мезиална темпорално-дялова *епилепсия* и хипокампадна склероза и следователно е една от предложените потенциални терапевтични цели и биомаркери за диагностициране на състоянието. Очевидно, участието на миРНК-424-5p в контрола на скоростта на клетъчното деление в различни тъкани може също да играе важна роля в развитието на патологията наблюдавана при РАС. Наскоро, Wu и сътрд. изказват предположението, че еволюционната роля на някои миРНК в мозъка на примати се свързва с инхибиране на прекомерната клетъчна пролиферация, феномен, наблюдаван също и при деца с диагноза РАС. Следователно, може да се заключи, че понижената експресия на миРНК-424-5p при пациенти с РАС може да отразява недостатъчната способност за ограничаване на ранният постнатален мозъчен растеж и развитие.

Най-значимото изследване на миРНК-664a-3p по отношение на неврологични проблеми е посветено на болестта на Алцхаймер и голямо депресивно разстройство. В посоченото изследване авторите извършват систематичен преглед на дисрегулираните миРНК молекули и при двете състояния и откриват, че само седем от изследваните миРНК са общи за тях, една от които е миРНК-664a-3p. И въпреки че, в представените в дисертацията изследвания, понижената експресия на тази миРНК при РАС е статистически значима, то тя показва най-ниска промяна в нивото на експресия и най-слаба прогностична стойност.

Множество гени, обект на регулация от изследвани в предходни проучвания миРНК молекули показват участие в неврологични заболявания и психични разстройства. В три отделни проучвания се открива повишаване на експресията на GNB1 протеина в префронталния кортекс при шизофрения. Други автори описват понижаване на нивата на транскрипция на гуанин нуклеотид-свързващ (G протеин) алфа q полипептид (*gnaq*) при мишки, подложени на хроничен стрес. Мутациите в друг ген, гуанин нуклеотид-свързващ протеин (G протеин) (*gnal*), ген, показващ експресия основно в мозъка, са сред основните причини за първична торсионна дистония и краниоцервикална дистония. Промяна в експресията на *gnal* се съобщава също и при шизофрения. Един от двата установени миРНК таргетни гена, кодиращ глутаматни транспортери, *slc1a1*, е широко изследван в контекста на епилепсия, OCD, MS, болест на Алцхаймер и шизофрения. Значителна свръхекспресия на SLC1A1 протеина е установена в *post-mortem* образци при пациенти с епилепсия, а също така и при плъхове с индуцирана от пилокарпин епилепсия. Значително повишаване на експресията на протеина SLC1A1 при пациенти с епилепсия също е докладвана и от други изследователи. Въпреки това, в други проучвания се съобщава също и за понижена експресия на

slc1a1 протеина. Освен това, повишената експресия на *slc1a1* гена на ниво иРНК и протеин също е демонстрирана при пациенти с диагноза шизофрения. Също така все повече доказателства предполагат участието на гена SLC1A1 и при развитието на множествена склероза. Но все още, предстои да се намери категорично доказателство за ролята на гена при развитието на множествената склероза.

В заключение, представените в настоящата дисертация данни демонстрират, че изследваните миРНК молекули: миРНК-500a-5p, миРНК-197-5p, миРНК-424-5p и миРНК-664a-3p, показват понижени нива в серум на деца с диагноза РАС, както и, че тези миРНК молекули притежават потенциала да бъдат използвани като лесно достъпни и обективно измерими биомаркери при това разстройство. Диагностичната сила на молекулярния тест, основан на миРНК-500a-5p и миРНК-197-5p, ще бъде най-висок, ако тези миРНК се оценяват *en block*.

VI. 3. Обсъждане на резултатите от проведеният количествен протеомен анализ – Isobaric Tag for Relative and Absolute Quantification (ITRAQ) при РАС

Използвайки двуизмерно iTRAQ базирано-LC-MS/MS профилиране, бяха идентифицирани редица диференциално експресирани протеини в серум от пациенти с диагноза РАС в сравнение с нормално развиващи се деца от контролна група. Използвайки iTRAQ метода за количествен протеомен анализ, в представеното изследване бяха идентифицирани 60 протеина, от които 24 протеина с повишена експресия и 36 протеина с понижена. В обобщение, протеомният подход, използван в представеното изследване, спомогна за идентифицирането на редица диференциално експресирани пептиди, които съответстват на известни протеини. Получените данни отразяват различия в имунните молекули, включително такива от системата на комплемента, които биха могли да повлияят косвено върху развиващия се мозък при РАС. Измежду диференциално експресираните протеини от анализа са протеини които; участват в клъстер от протеини с роля в регулацията на Notch сигналния път, протеини участващи във вроденият имунен отговор, (активацията на комплемента), включително участващи в каскадата на комплемента, комплемент C4-A протеин и комплемент C1q субкомпонент субединица А. Сглобяването на нуклеозомите и тяхното позициониране, направляването на аксоните, протеини участващи в метаболизма на холестерола и транспорта на липопротеини като аполипопротеин С-II, аполипопротеин C4 (APOC4) и аполипопротеин F, както и протеини участващи в активирането на Т клетките на, имунния отговор, холестеролният ефлукс, както и в процеса на положителна регулация на биосинтезата на мастни киселини, показват значителна диференциална

регулация. Аполипопротеин С-II играе важна роля в метаболизма на липопротеините като активатор на липопротеиновата липаза. Аполипопротеин С4 (APOC4) е липид-свързващ протеин, принадлежащ към семейството на аполипопротеиновите гени. Смята се, че протеинът играе роля в метаболизма на липидите. Аполипопротеин F, този протеин образува комплекси с липопротеини и може да участва в транспорт и/или естерификацията на холестерола. Като цяло проведените пилотни протеомни изследвания предполагат, че поне при някои от децата с РАС включени в изследваната извадка се наблюдава дисрегулация на метаболизма на холестерола и подкрепят настоящите теории, че РАС може да бъде свързано с повишен отговор на оксидативен стрес. В редица изследвания е демонстрирано, че факторите на комплемента C1q, C3, C4, C5, C6, C9, CR1, C1qRr, CR3, CR4, фактор H както и други фактори показват повишена експресия в мозъка при хора и опитни животни с инфекция, исхемия и инсулт, болест на Алцхаймер, хорея на Хънтингтън и много други заболявания на ЦНС. Reelin, който трябва да се секретира в извънклетъчният матрикс, за да прояви биологичния си ефект, разгражда фибронектин и следователно може да регулира премахването на апоптотични клетки свързано с фибронектин-C1q в нервната система. В контекста на наблюдаваната промяна на протеините на комплемента в серум на лица с РАС в представеното изследване, интересно е да се отбележи проучването на Vargas и съав. които представят доказателства за възпалителни процеси в мозъка на пациенти с РАС.

VI. 4. Обсъждане на резултатите получени от проведените транскриптомни (RNA-Seq) изследвания на белтък-кодиращи гени при РАС

В настоящия дисертационен труд, чрез прилагане на набор от техники, са определени диференциално експресирани протеин-кодиращи гени в периферна кръв от пациенти с диагноза разстройство от аутистичния спектър (РАС). Представеното в настоящата дисертация изследване е първото по рода си комплексно изследване, проведено върху подбрана група от български пациенти. До момента в литературата са докладвани ограничен брой пълни транскриптомни анализи при деца с РАС.

В представената научна разработка се разглеждат данни, получени от цялостен транскриптомен анализ, разкриващ разлики в експресията на протеин-кодиращи гени при деца с РАС, което предполага, че някои генетични аспекти на разстройството са споделени от повечето от децата, включени в изследваната извадка..

Получените резултати разкриват, че процеси като синаптична пластичност, смущения в ориентирането на аксоните, невронното преживяване, диференциацията и възпалението имат роля в патофизиологията на РАС.

CD177 (CD177 молекула) е ген, кодиращ протеина - гликозил-фосфатидилинозитол (GPI) - свързан повърхностен гликопротеин, който играе роля при активиране на неутрофилите. Променената експресия на гена може би води до променена миграция на левкоцити в ЦНС. Има достатъчно данни за ключовата роля на имунитета при развитието на РАС.

CRB2 е протеин кодиращ ген. Данните получени от GO анотации асоциират протеина със свързване на калциеви йони. Смята се, че протеина може да има участие в морфогенезата на клетките (<http://www.genecards.org/>).

DPM3 е протеин кодиращ ген, който кодира субединица на долихол-фосфат манозилтрансфераза и действа като стабилизатор при формиране на комплекса. Протеина участва в биосинтеза на предшествениците на N-гликаните. N-гликопротеините имат ключова роля по време на ембрионалното развитие, диференциация и поддръжката на клетъчните функции (<http://www.genecards.org/>).

IGF2BP1 (инсулин-подобен растежен фактор 2, иРНК свързващ протеин 1) е ген, кодиращ протеин. Този ген кодира иРНК-свързващ протеин, член на семейството на инсулин-подобен растежен фактор 2. В представените данни генът също показва намалена експресия, което предполага участието му в патогенезата на РАС.

LTF (Lactotransferrin)

Няколко автори съобщават значително повишена експресия на LTF в мозъка на пациенти с болестта на Алцхаймер. Открит обаче остава въпросът дали повишената експресия на LTF е следствие на развитието на РАС или е опит организма да отговори на патологични имунни процеси в мозъка и има протективна функция?

MAGEA4 (Melanoma Antigen Family A4) е протеин кодиращ ген. Смята се, че протеина има отношение в процеса на ембрионалното развитие (<http://www.genecards.org/>).

OTOF (Otoferrlin) е ген, кодиращ мулти C2 - домен протеин, участващ в свързаното на Ca^{2+} , сливането на везикули с плазмената мембрана и в контрола на освобождаване на невротрансмитери при синапсите. Полученият резултат за

понижена експресия в представеното проучване може да бъде разгледан като потенциален механизъм на нарушаване на баланса при синапсите, участващ в патогенезата на РАС.

VI. 5. Обсъждане на резултатите получени от ген онтологичния анализ на диференциално експресираните гени и KEGG анализ на биологичните пътищата

Резултатите от проведеният ген-онтологичен анализ показват биологични връзки и общи пътища между диференциално експресираните гени участващи в структурни компоненти на клетката, апоптозата, развитийни, метаболитни и регулаторни процеси, процеси участващи в имунитета и междуклетъчната сигнализацията.

Получените данни демонстрират: 1) изключителната хетерогенност в патогенезата на това разстройство; 2) наличието на доказателства за мултифакторния модел при РАС; 3) връзката на имунната система с РАС; 4) отношение към това разстройство на гени участващи в междуклетъчната сигнализация; 5) дисрегулацията на основни клетъчни процеси може да допринесе за развитие на разстройството. Проведеният анализ идентифицира метаболитни и сигнални пътища като ендоканабиноидна сигнализация, калциева сигнализация и др. в които участват диференциално експресираните гени, получени в настоящото изследване.

Калциев сигнален път

Мутации косвено, осигуряващи завишени нива на калций в клетката чрез повишен инфлукс на Ca^{2+} или чрез потискане механизмите на обратна връзка, са открити при пациенти с РАС. В тази връзка е и проучването на Palmieri и сътр., които докладват повишена концентрация на калций в неокорикална мозъчна тъкан (*post-mortem*) при шест пациента с РАС в сравнение с контролна група.

Ендоканабиноидна сигнализация

Ендоканабиноидите са от ключово значение за модулиране на синаптичната функция. Чрез активиране на канабиноидните рецептори, експресирани в ЦНС, тези липидни посредници могат да регулират невронните функции и поведението. Ретроградната сигнализация е основния начин, чрез който ендоканабиноидите медиират краткосрочни и дългосрочни форми на пластичност при възбуждащите и инхибиторните синапси. Делеции и точкови мутации в гена кодиращ невролигин 3 (NLG 3) се свързват с РАС. Интерес предизвиква и друго изследване при мишки с модел на Fra X синдромен РАС (FMR1-/Y), което показва, че делецията на FMR1 гена води до по-голям

ендоканабиноидно-медиран отговор на GABA-ергични синапси в дорзалния стриатум и CA1 регион на хипокампуса.

Допаминаергичен синапс

Допаминът (DA) е важен невротрансмитер в мозъка, където контролира различни функции като двигателна активност, ендокринна регулация, учене, памет, мотивация и възнаграждение.

Wnt сигналния път

Предвид широкия спектър от процеси, в които участва Wnt/ β -катениновия път, нарушената му регулация може да има произволен брой негативни ефекти върху невронното развитие и по този начин да допринесе за патогенезата на невроразвитийни разстройства, включително PAC.

VI. 6. Обсъждане на резултатите от сравнителният експресионен анализ (DGE-tag profiling) при шизофрения

Въпреки идентифицирането на множество кандидат гени при шизофрения, етиопатогенезата на шизофренията остава неизвестна. Използването на DGE анализ за идентифициране на диференциално експресирани гени притежава потенциала да добави ново измерение към разбирането ни за регулацията на транскрипцията при шизофрения. Периферната кръв е идеална заместваща тъкан, тъй като има потенциала да отразява настъпилите физиологични промени чрез промяна на нивата на генна експресия. Промяната в експресията на миРНК молекули може да се обясни отчасти с нарушение на регулацията на гени свързани с биогенезата на миРНК като DICER1, ген, който бе установен като диференцирано експесиран в представеното проучване. Освен това е възможно промяната в експресията на DICER1 гена да допринесе за промени в експресията на протеин-кодиращи гени чрез модулиране на процесинга на миРНК молекулите.

VII. ЗАКЛЮЧЕНИЕ

Изследванията в областта на миРНК биологията все още са в начален стадий. Предизвикателствата в областта изискват решаването на редица технически проблеми като установяване на стандартизирани процедури за пробоподготовка, усъвършенстване на комбинирани методологии за откриване и количествено идентифициране на миРНК молекули и интерпретация на получените резултати. С напредъка на новите техники и последващите изследвания, миРНК несъмнено се оказват обещаващи при диагностиката на нарушения на ЦНС както и за оценката на свързаната с тях терапия.

Ролята на миРНК при психиатричните разстройства е интензивно развиваща се област в която се интегрират непрекъснато подобряващи се подходи за изследване. В допълнение, включването на мета-анализи на миРНК, обхващащи генетични вариации, експресия и биологична функция, ще предоставят допълнително ценна информация за потенциалната роля на миРНК молекулите при РАС което би могло да подпомогне диагностиката на РАС и психиатричните разстройства и хвърли допълнителна светлина върху етиопатогенезата на разстройствата. В заключение, биомаркерите базирани на миРНК молекули могат да бъдат много полезни при разграничаване на различни подтипове психиатрични разстройства. Транскриптомните данни, базирани на RNA-Seq подхода, могат да предоставят ценна информация за диференциална генна експресия при РАС. Съвременните разбирания относно транскриптомното профилиране, базирано на RNA-Seq подхода, определят обещаващо място на технологията като спомагаща за поставянето на прецизна медицинска диагностика в близко бъдеще. В заключение, приложена технологията за секвениране от ново поколение "Whole-Exome Sequencing" за анализ на генетичната вариабилност при РАС разкри полезността си при идентифицирането на генетични варианти при РАС. Получените резултати също са в подкрепа на твърдението, че в близко бъдеще диагностиката на базата на цялостното екзомно секвениране ще бъде ефективен първичен диагностичен метод при пациенти с диагноза РАС. Използването на DGE технологията за идентифициране на диференциално експресирани белтък-кодиращи гени при пациенти с диагноза шизофрения добавя ново измерение към разбирането ни за генната функция при шизофрения.

Научни публикации свързани с дисертационният труд

1. **Circulating miRNAs as a novel class of potential diagnostic biomarkers in neuropsychiatric disorders.** Tatyana M. Kichukova, Nikolay T. Popov, Hristo Y. Ivanov, Tihomir I. Vachev *Folia Medica* 2016; 57(3&4): 159-172.
2. **Autism Spectrum Disorder – A Complex Genetic Disorder.** Hristo Y. Ivanov, Vili K. Stoyanova, Nikolay T. Popov, Tihomir I. Vachev *Folia Medica* 57(1) (2015).
3. **Profiling of circulating serum microRNAs in children with autism spectrum disorder using stem-loop qRT-PCR assay.** Tatyana M. Kichukova, Nikolay T. Popov, Ivan S. Ivanov, Tihomir I. Vachev *Folia Medica* 2017; 59(1).
4. **Identification of a novel mitochondrial mutation in the cytochrome c oxidase III gene in children with autistic spectrum disorders using next generation RNA-sequencing.** Danail Minchev, Nikolay Popov, Veselin Petrov, Ivan Minkov, Tihomir Vachev, *Comptesrendus de l'Académiebulgare des Science*. 2019 IN PRESS

5. **Epigenetic aspects in schizophrenia etiology and pathogenesis.** Nikolay T. Popov, Vili K. Stoyanova, Nadezhda P. Madzhirova, Tihomir I. Vachev *Folia Medica* 2012; 54(2): 12-16.

6. **Investigation of fasciculation and elongation protein zeta-1 (FEZ1) gene expression changes in schizophrenia.** Tihomir Iliev Vachev, Vili Krasteva Stoyanova, Ivan Minkov, Nikolay Todorov Popov *BJMG* 18 (1), 2015 | 31-38.

7. **Investigation of circulating serum microRNAs miR-328-3p and miR-3135a expression as a promising novel biomarkers for ASD.** Nikolay Popov, Danail Minchev, Ivan S. Ivanov, Vili Stoyanova, Tihomir Vachev *BJMG* 21 (2) 2018.

8. **Blood-Based Gene Expression in children with Autism spectrum disorder.** Hristo Y. Ivanov¹, Vili K. Stoyanova, Nikolay T. Popov, Tihomir I. Vachev, *Biodiscovery* 2015; 17: 2; DOI: 10.7750.

9. **Comparative expression analysis of miR-619-5p in serum and PBMCs as a promising candidate biomarker for autism spectrum disorder.** Nikolay T. Popov, Veselin D. Petrov, Danail S. Minchev, Ivan N. Minkov, Tihomir I. Vachev *Comptesrendus de l'Académiebulgare des Science*. 2019 IN PRESS.

VIII. ПРИНОСИ

Приноси с научно-фундаментален характер

Идентифицирани са кандидат миРНК молекули и белтък-кодиращи гени с потенциални характеристики на биомаркери при РАС.

Идентифицирани са кандидат белтък-кодиращи гени с потенциални характеристики на биомаркери при шизофрения.

Приноси с научно-методологичен характер

Успешно е приложена технологията за секвениране от ново поколение "Whole-Exome Sequencing" за анализ на генетичната вариабилност в кодиращите области на гените при пациенти с РАС.

Успешно е приложена технологията за секвениране от ново поколение "RNA-Sequencing" за идентифициране на диференциално експресирани белтък-кодиращи гени при пациенти с диагноза РАС.

Успешно е приложена технологията за секвениране от ново поколение "Small RNA-Sequencing" за анализ на диференциалният миРНК профил при пациенти с диагноза РАС.

Успешно е приложена технологията за провеждане на сравнителен протеомен анализ „*Isobaric Tag for Relative and Absolute Quantification (ITRAQ)*“ за анализ на диференциалният протеомен профил при пациенти с РАС.

Успешно е приложена технологията за секвениране “Digital Gene Expression Tag Profiling (DGE)” за идентифициране на диференциално експресирани белтък-кодиращи гени при пациенти с диагноза шизофрения.

Приноси с приложен характер:

Създадена е ДНК банка от пациенти с диагноза РАС и шизофрения, както и съответните им кохорти от здрави лица предоставящи възможност за по-мащабни бъдещи проучвания.