МЕДИЦИНСКИ УНИВЕРСИТЕТ – СОФИЯ
МЕДИЦИНСКИ ФАКУЛТЕТ
КАТЕДРА МЕДИЦИНСКА ГЕНЕТИКА

Аделина Горанова Йосифова

Асоциативно проучване на кандидат гени и цялостен геномен скрининг при пациенти с биполарно афективно разстройство

А В Т О Р Е Ф Е Р А Т
на дисертация за присъждане
на образователна и научна степен “ДОКТОР”

София, 2011

Дисертационния труд съдържа 156 страници и включва 17 фигури и 22 таблици. Оформен е в 5 глави, изводи, приноси, приложения и литературни източници. Библиографията обхваща 336 източника на английски език, по–голямата част, от които от последните 10 години.

Изследванията са проведени в *Laboratory for International Alliance, RIKEN Center for Genomic Medicine, Йокохама, Япония* и *Катедра медицинска генетика, Медицински факултет, Медицински Университет София, България*

Защитата на дисертационния труд ще се състои на 17.10.2011 г. от 15.00 часа в Аулата на СБАЛАГ „Майчин дом", ул. „Здраве" 2, София, пред научно жури в състав:

1. Проф. д-р Драга Тончева – **научен ръководител**
2. Доц. д-р Стоян Лалчев – **председател на журито**
3. Чл.-кор. Проф. Георги Русев – **рецензент**
4. Проф. Севдалин Георгиев – **рецензент**
5. Чл.-кор. Проф. Иван Иванов

Материалите по защитата са на разположение в Катедрата по медицинска генетика и са публикувани на интернет страницата на МУ – София, www.mu-sofia.bg

МЕДИЦИНСКИ УНИВЕРСИТЕТ – СОФИЯ
МЕДИЦИНСКИ ФАКУЛТЕТ
КАТЕДРА ПО МЕДИЦИНСКА ГЕНЕТИКА

Аделина Горанова Йосифова

Асоциативно проучване на кандидат гени и цялостен геномен скрининг при пациенти с биполарно афективно разстройство

А В Т О Р Е Ф Е Р А Т

на дисертация за присъждане на образователна и научна степен “ДОКТОР”

Професионално направление ”Медицина” 7.1

Научна специалност ”Генетика”

Научни ръководители:

ПРОФ. Д-Р ДРАГА ТОНЧЕВА, ”ДОКТОР”

ПРОФ. Д-Р ЮСУКЕ НАКАМУРА, ”ДОКТОР”

София, 2011

УВОД

Биполарното афективно разстройство, наричано още ”манийно – депресивна болест” е често срещано, тежко психиатрично заболяване, което засяга от 0.5% до 1.5% от общата популация. Заболяването се характеризира с периодично повтарящи се епизоди на мания или хипомания и депресия, което води до нарушения в нормалния ритъм на живот.

Честата повтаряемост на двете фази, ранното начало, продължителната терапия, многократните хоспитализации, честата злоупотреба с различни лекарства, нарушената работоспособност и не на последно място, повишения риск от самоубийство, ако заболяването не бъде третирано, са тежко емоционално и финансово бреме, както за самия болен, така и за неговите роднини и за обществото като цяло. Биполарното афективно разстройство може да има сравнително ранно начало, като първия епизод обикновено настъпва преди 30 годишна възраст. Въпреки това е установено, че заболяването може да се манифестира и в по-късен етап от живота. Установено е, че при 9.3% от пациентите с афективни разстройства манийните епизоди могат да се проявят едва след 60 годишна възраст. Заболяването засяга в еднаква степен и двата пола, но профилът на симптомите може да бъде различен. Съществуват предположения,че жените имат склонност към повече смесени епизоди или епизоди на мания в сравнение с мъжете. Диагностицирането на биполарното афективно разстройство не винаги е лека и безпроблемна задача. Основната причина за това могат да бъдат редица физически състояния, които имат симптоми характерни за тези при биполарното заболяване. Поставянето на диагнозата се затруднява и поради факта, че не съществуват лабораторни тестове за откриване на заболяването. В повечето случаи диагнозата се базира основно на това, което пациента споделя сам, на неговото поведение, на наблюдения на негови близки, приятели и колеги. Съгласно DSM IV (Диагностичен и статистически справочник на умствените заболявания, IV издание) към групата на афективните разстройства се причисляват следните заболявания: биполарно разстройство (манийно-депресивна психоза), циклотимия, дистимия и униполарно разстройство или депресивна психоза. Биполарното афективно разстройство, от своя страна, се подразделя на биполарно разстройство тип I и биполарно разстройство тип II. Биполарното афективно разстройство може да възникне в резултат от комбиниране на генетични, биологични фактори и фактори на околната среда, за изясняването, на които са проведени редица проучвания. Резултатите от *фамилни* изследвания, изследвания на *близнаци* и *осиновени деца* сочат генетичния компонент като основен в етиологията на биполарното разстройство. Големият брой *фамилни изследвания* проведени през

последните години, недвусмислено сочи фамилният характер на заболяването, като се предполага, че родствениците от първа степен на пациенти с биполарно разстройство имат 5 – 10 пъти по-висок риск за развитие на заболяването в сравнение с риска за общата популация. Резултатите от проведените *близначни проучвания* установяват ниво на съвпадение (конкордантност) за заболяването, значително по – високо сред монозиготни близнаци (78%) в сравнение с дизиготните (29%), а установената унаследяемост варира в границите между 59% и 93%, което показва силното генетично присъствие в етиологията на заболяването. Въпреки това, фактът, че наблюдаваната конкордантност дори при монозиготни близнаци е по-малка от 100%, недвусмислено показва въздействието на фактори и на околната среда. Резултатите от *изследвания на осиновени деца* показват, че честотата на афективните заболявания (униполарно, биполарно, циклотимия) е значително по – висока при биологичните родители на биполарни пробанди (31%) в сравнение с приемните родители на такива пациенти (12%) и в сравнение с биологичните родители на здрави осиновени деца (2%).

Проведените през годините многобройни фамилни проучвания, проучвания на близнаци и осиновени деца категорично показват ролята на генетичните фактори в етиологията на заболяването, но не предлагат ясен модел за неговото унаследяване. Ако се вземат предвид данните за рекурентния риск от многобройните генетични изследвания, рязкото намаляване на този риск в посока: монозиготни близнаци → роднини от първа степен → обща популация, по – скоро се обяснява с взаимодействието на няколко гена (епистаза). Това взаимодействие може да включва един ген с голям ефект и няколко с малък или няколко гена със среден ефект. Друга възможност е взаимодействие на голям брой гени, всеки от които е с малък ефект. Въпреки голямата роля на генетичните компоненти, в етиологията на заболяването се включват и фактори на околната среда, поради което се приема, че биполарното разстройство е *мултифакторно заболяване*. Според мултифакторния модел, предразположеността към дадено заболяване е променлива величина с Гаусов тип на разпределение в популацията и когато тази величина премине дадена граница (праг на развитие на болестта), заболяването се изявява. Днес вече са познати генетични механизми, които не следват Менделовите закони, дават сложни модели на унаследяване и за които се предполага, че могат да бъдат свързани с биполарното разстройство. Пример в това отношение са динамичните мутации като експанзията на тринуклеотидни повтори, геномния импринтинг и митохондриалната наследственост.

Откриването на *хромозомни нарушения* при индивиди с умствено заболяване, може директно да насочи към определен хромозомен район, съдържащ кандидат гени,

предразполагащи към развитие на съответното заболяване. При биполарното разстройство, най – честите хромозомни нарушения включват: t(1;11)(q42.1;q14.3); t(9;11)(p24;q23.1); r(18)(18p11pter;18q23qter); t(2;18)(q21q23); inv(18)(p11.3q21.1).

Две са основните групи подходи, които се използват за откриване и определяне на специфичните гени и молекулни варианти в тях, отговорни за дадено заболяване: *анализ за скаченост* (Linkage mapping), който се провежда в големи родословия и *асоциативни проучвания* (Linkage disequilibrium mapping) на ниво популация. При *анализа за скаченост* се проследява ко-сегрегацията между даден фенотип и определен генетичен маркер в поколението. При този метод се изследват големи родословия с няколко засегнати от болестта индивида с помощта на голям брой генетични маркери (микросателити), разпръснати из целия геном. При анализа за скаченост не се изисква познаване патофизиологията на изследваното заболяване и тук се детектират гени с голям ефект на големи разстояния. Успешното прилагане на анализа за скаченост изисква уточняване на болестен статус, модел на унаследяване на заболяването (доминантен, рецесивен), както и честота и пенетрантност на болестния алел (т.е., вероятността даден индивид да е болен, ако се приеме, че е носител на този алел). С помощта на параметричния анализ за скаченост са открити и картирани гените за около 1000 моногенни болести. При комплексните заболявания, каквото е биполарното афективно разстройство, този вид анализ няма такъв успех. Тук моделът на унаследяване на заболяването, пенетрантността и честотата на болестния алел не могат да бъдат определени със сигурност. Ето защо при изследването на тези заболявания по – подходящ е непараметричният анализ за скаченост. Този анализ се основава на сравнение между наблюдаваното и очаквано поделяне на еднакви алели, известно като идентичност по произход (identity by descent, (IBD) между двойки засегнати родственици. Най – широко използваният метод при непараметричния анализ за скаченост е метода на засегнатите сибси (affected sib pair (ASP) method). Засегнатите сибси се генотипират за голям брой генетични маркери (няколко стотин). И двата вида анализи за скаченост: параметричен и непараметричен са широко използвани при генетичните изследвания на биполарното разстройство. В резултат на многобройните изследвания проведени през годините, много райони от човешкия геном са били асоциирани с биполарното афективно разстройство: 4p16, 12q24, 18p11.2, 18q22, 21q22, 22q11–13. Други хромозомни райони, представляващи интерес са: 5p15.33, 6q21, 8q24.22, 10q26.2, 11p15.5, 13q32.1–3, 17q25.3, 20q13.33, Xq26.1. *Асоциативните проучвания* се базират на предположението, че комплексните заболявания се причиняват от голям брой различни гени, всеки от които допринася по малко (има малък ефект) за развитието на заболяването.

Основната цел на асоциативните изследвания е да се открият специфични алели в дадени генетични маркери, които се срещат с по – голяма честота сред болни индивиди в сравнение със здравите. За постигането на тази цел се използват два основни подхода: *директн*а и *индиректна* асоциация. При *директната* асоциация самият маркер съдържа функционален полиморфизъм, който директно повлиява фенотипа, като кодира протеин с различна структура или причинява разлики в експресията на гена. При *индиректната* асоциация алелът не причинява болестта, но е много близко разположен до болестия алел и се намира в неравновесна скаченост (LD) с него, т. е. алелът на даден генетичен маркер и болестният алел се предават заедно от поколение на поколение, защото са толкова близко разположени, че между тях много рядко настъпва рекомбинация при мейозата. За провеждане на асоциативно изследване могат да се използват два типа извадки: *тип 'случаи – контроли'* и *семейства с болен пробанд и неговите родители (триоси)*.

Поради своето изобилие и гъстота в човешкия геном, SNPs са идеални полиморфни маркери при асоциативни изследвания на комплексните заболявания и се използват за установяване на гените, отговорни за дадено заболяване, както и за идентифициране на мутациите в тези гени. Те представляват замени на единични нуклеотиди, които се срещат с честота >1% в поне една популация. Най – интересни за анализиране са кодиращите SNPs, водещи до замяна на разнотипни аминокиселини (замяна на основна с киселинна аминокиселина), тъй като този тип замени водят до структурни и функционални промени на съответния генен продукт. Тези полиморфизми обаче са най-рядко срещаните (~5%) и най-малко хетерозиготни в сравнение с останалите SNPs, което е нормално като се вземе пред вид действието на естествения отбор. Близо 60% от всички кодиращи SNPs, водещи до аминокиселинна замяна, имат честота на редкия алел <5%. Поради тези причини, при традиционния анализ на индивидуални маркери се включват и полиморфизми локализирани в некодиращите райони на гена и дори в извънгенните области, тъй като те могат да са в неравновесна скаченост с все още неизвестен функционален полиморфизъм, който създава предразположение за развитие на дадено заболяване. В последните няколко години при асоциативните изследвания, анализирането на единични, случайно избрани SNPs се комбинира с анализиране на хаплотипи. Най – голямата атрактивност на хаплотипните методи е идеята, че често срещаните хаплотипи обхващат повече от генетичните варианти в един хаплоблок и могат да бъдат тествани като се използват малък брой SNPs, т. нар. 'представителни' SNPs (htSNPs - haplotype tag SNPs). Изборът на кандидат гени за асоциативни проучвания става въз основа на 1) тяхната биологична роля в патофизиологията на заболяването и 2) тяхната локализация в райони, показали скаченост

със заболяването в предишни изследвания. При биполарното разстройство повечето изследвания са фокусирани върху невротрансмитерните системи, които са основна мишена за действие на лекарствата използвани при лечение на заболяването, а именно допамин, серотонин и норадренергичните системи. Сред т. нар. 'традиционни' кандидати, получили най – голямо внимание са гените кодиращи моноамино оксидаза А (*MAOA*), тирозин хидроксилазата (*TH*), катехол –О – метилтрансферазата (*COMT*), и серотониновия транспортер (*5HTT*).

Завършването на човешкия геномен проект (The Human Genome project) и Hap-Map проекта (The HapMap project) и наличието на високо – ефективни технологии, позволяващи едновременното тестване на стотици хиляди SNPs в целия геном, в голям брой проби, направи технически възможно и финансово изгодно провеждането на широко мащабните асоциативни проучвания. При тези проучвания могат да използват и двата типа извадки, характерни за асоциативните изследвания. Понастоящем се използват три подхода при избора на маркери за провеждане на цялостен геномен скрининг: 1) случайна селекция, 2) селекция въз основа на неравновесна скаченост (LD) и 3) избор на функционални SNPs (кодиращи, cSNPs). При селекцията на маркери въз основа на LD се избират минимален брой представителни SNPs (tagSNPs), които обхващат максималния брой варианти, наблюдавани в дадената популация. При случайната селекция LD не се взима под внимание. При третият подход се анализират само кодиращи SNPs, водещи до разнотипна аминокиселинна замяна или всички варианти, разположени в гени, като се предполага, че те са свързани с изследваното заболяване.

До 2008г. проведените широко мащабни асоциативни проучвания при биполарно разстройство са четири. Три от тези изследвания използват Affymetrix, а едно Illumina като генотипиращи платформи. Анализирайки ДНК пулове (комбиниране на много индивидуални ДНК проби) от 1233 случая и 1439 контроли, Baum и колеги (2008) установяват асоциация между биполарното разстройство и SNP (rs1012053, $P = 1.5 \times 10^{-8}$) в гена за диацилглицерол киназа (*DGKH*). Генът е картиран в хромозома 13q14.11 и представлява интерес, тъй като неговият продукт е включен във фосфатидил инозитоловия път, чрез който литият упражнява някои от своите стабилизиращи настроението ефекти.

В друг геномен скан (WTCCC, 2007), при индивидуално генотипиране на 1868 случая и 2938 контроли е идентифициран локус в район с висока неравновесна скаченост в късото рамо на хромозома 16 (16p12) включващ няколко гена - *PALB2*, чиито продукт е необходим за стабилността на ключови ядрени структури, включително на хроматина; *NDUFAB1*, който кодира един от компонентите в митохондриалната дихателна верига и

DCTN5, който взаимодейства с *DISC1* (ген асоцииран с шизофрения и биполарно разстройство).

Sklar и колеги (2008) провеждат широко мащабно асоциативно изследване върху 1461 случая и 2008 контролни индивида и получават статистически значими резултати за два варианта - rs4939921 ($P = 1.7 X 10^{-7}$) в гена за миозин (*MYO5B*) и rs1705236 ($P = 6.1 X 10^{-7}$) в гена за тетраспанин – 8 (*TSPAN8*). Проведеният хаплотипен анализ също показва асоциация. Въпреки това, допълнителното генотипиране на тези и други сигнификантни SNPs в две нови извадки (409 триоса; 365 случая и 351 контроли) не потвърждава първоначалната находка.

Използвайки генотипираща платформа, като тези на Sklar et al и The WTCCC, Ferreira и колеги (2008), анализират 1098 биполарни пациенти и 1267 контроли, като установяват асоциация за SNP (rs7221510, $P = 1.4 X 10^{-6}$) в гена *SKAP1*. Проведеният в последствие мета – анализ, включващ гореспоменатата извадка и данни от WTCCC (2007), Sklar et al (2008) (общо 4387 случая и 6209 контроли) установява силна асоциация на два варианта: rs10994336 ($P = 9.1 X 10^{-9}$) в гена за анкирин – G (*ANK3*) и rs1006737 ($P = 7.0 X 10^{-8}$) в *CACNA1C* гена.

Мета – анализът на данните от цялостните геномни сканове, проведени от Baum et al (2008) и The WTCCC (2007), идентифицира още два допълнителни гена *JAM3* и *SLC39A3* като евентуално асоциирани с биполарното разстройство.

ЦЕЛ И ЗАДАЧИ

Цел:

Да се проучат генетични варианти в известни кандидат гени за асоциираност с биполарното разстройство и да се идентифицират нови полиморфизми, предразполагащи към развитие на заболяването в българска популация.

Задачи:

1.· Да се направи подбор и да се анализират полиморфизми в:

А.· гени, доказали асоциация със заболяването в други популации;

В.· кандидат гени, неизследвани до момента за асоциация в българска популация;

С.· гени, неизследвани за асоциация със заболяването, но с предполагаемо участие в неговата патогенеза, поради нарушена експресия в мозъка при пациенти с биполарно разстройство.

2.· Да се определи алелната и генотипната честота на анализираните полиморфизми в български пациенти и контроли и да се изследва тяхната асоциация с биполарното афективно разстройство.

3.· Да се верифицират резултатите за всички полиморфизми, показали положителна асоциация, посредством анализирането им в допълнителна група от случаи и контроли.

4.· Да се проведе широко мащабно асоциативно проучване за идентифициране на нови полиморфизми, чрез цялостно сканиране на генома и анализиране на повече от 550 000 SNPs при български пациенти с биполарно разстройство и контроли.

5.· Да се валидизират и реплицират резултатите от цялостния геномен скрининг, посредством повторно генотипиране на най–силно асоциираните SNPs в допълнителна група от случаи и контроли.

МАТЕРИАЛИ И МЕТОДИ

Клиничен материал

Извадката от пациенти (случаи) включва общо 310 неродствени индивида от български произход, от които: 232 (74.8%) с диагноза биполарно разстройство тип I, 9 пациента (2.9%) с диагноза биполарно разстройство тип II и 69 индивида (22.3%) с диагноза биполарно разстройство (неуточнен тип). Пациентите са набрани от регионални психиатрични болници и диспансери в градовете София, Пловдив, Русе, Благоевград, Раднево в България. Диагнозата е поставена от лекуващите лекари въз основа на цялата налична клинична документация (епикризи от предишни хоспитализации, здравни картони), подробни анамнестични данни и полуструктурирани интервюта (съкратената версия на СКАН (SCAN; Schedule of Clinical Assesment in Neuropsychiatry), съгласно изискванията на DSM-IV критериите. При несъответствие по отношение на диагнозата, с пациента е проведено допълнително детайлно интервю от двама независими лекари-специалисти, с цел постигне на консенсусна диагноза.

Контролната група се състои от 704 неродствени, психически здрави индивида, подбрани от същите региони на България и съответстващи на болните по пол, възраст и етнически произход. Контролите са набрани сред пациенти на общо практикуващи лекари сред медицински персонал и сред кандидати за придобиване на свидетелство за

управление на МПС. Някои демографски характеристики за двете изследвани групи са представени в таблица 1.

Таблица 1: Демографски данни за случаи и контроли

Показател	Контроли		Случаи	
Брой индивиди	704		310	
Пол	*мъже*	*жени*	*мъже*	*жени*
	350 (49.7%)	354 (50.3%)	119 (38.4%)	191 (61.6%)
Средна възраст (стандартно отклонение)*	47 (±16)	50 (±15.2)	39 (±15.3)	37 (±17.2)

* липсват данни за възрастта на 27 пациента (11 мъже и 16 жени)

На всички пациенти и контроли е дадена пълна информация за изследването и те са подписали информирано съгласие за участие. Одобрение за провеждане на изследването е получено от всички местни етични комисии, където са събирани пробите, както и от етичната комисия в института в Йокохама (Yokohama), RIKEN (The Institutes of Physical and Chemical Research), Япония.

Биологичен материал

Значителна част от ДНК пробите (220 пациенти) и всички контроли са набрани в рамките на съвместен проект между Катедрата по медицинска генетика, при Медицинския университет София и Катедрата по невропсихиатрична генетика, при Медицински университет, Кардиф, Великобритания, в периода 2000 – 2006 г., с финансовата подкрепа на Janssen Research Foundation. Останалата част от пробите (90 пациента) са набрани допълнително от Катедра медицинска генетика, при Медицински университет София през 2007 г.

За генетичен анализ е използвана високомолекулна ДНК изолирана от ядрени кръвни клетки. Използвани са 10 ml венозна кръв, взета стерилно в пластмасови епруветки (вакутейнери), съдържащи K_3EDTA като антикоагулант. Кръвта е транспортирана на 4ºC до 24-48h в молекулярно-генетичната лаборатория в Катедрата по медицинска генетика, където веднага е подложена на обработка за изолиране на високомолекулна ДНК посредством фенол – хлороформна екстракция. След изолирането, високомолекулната геномна ДНК е съхранявана в ТЕ буфер при -20ºC в ДНК-банката на Катедрата по медицинска генетика. От всяка проба е взето равно

количество (200 µl); пробите са транспортирани върху сух лед до Institute of Medical Science, Laboratory of Pharmacogenetics, RIKEN, Йокохама, Япония, където са проведени анализите.

Концентрацията и чистотата на всяка ДНК проба са определени със спектрофотометър NanoDrop® ND-1000, V3.3 (NanoDrop Technologies, Inc.), като за целта е използвана 1 µl проба. За определяне чистотата на ДНК е взето съотношението от абсорбцията на ДНК при 260nm и 280nm дължина на вълната. При съотношение A260:A280=~1.8, ДНК се приема за 'чиста' (свободна от примеси).

За количествено оценяване на наличието и степента на деградация на ДНК е проведена електрофореза в 1% агарозен гел, с 1 µl проба. Регистрирането на smear и липсата на ясно ограничена ивица е показател за ДНК-деградация. Концентрациите на двадесет и пет проби, показали ниска стойност на съотношението A260:A280 и наличие на smear са оценени допълнително в институт RIKEN, Yokohama, SNP research center, с PicoGreen®, dsDNA Quantification Kit (Molecular Probes, Eugene, Oregon, USA) съгласно протокола на производителя. След измерване на концентрацията, пробите са разредени с Tris EDTA буфер (pH 8.0) до 50ng/µl за цялостния геномен скрининг и до 5ng/µl за асоциативно проучване на кандидат гени.

Дизайн на експерименталната работа

Експерименталната работа при настоящия научно – изследователски труд обхваща най – общо два етапа. През първият етап, с помощта на базата данни PubMed на NCBI (National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/) са проучени максимален брой литературни източници, описващи молекулярно-генетични изследвания при биполарно разстройство и е изготвен списък от 199 кандидат маркера (единични нуклеотидни полиморфизми) от 65 гена. След направения подбор, за всеки вариант е избран метод за генотипиране, а именно: за 47 полиморфизма е използван метод TaqMan, а останалите 152 маркера са анализирани чрез Invader метод, комбиниран с мултиплексна полимеразна верижна реакция (MPCR, multiplex PCR) с проби и праймери по собствен дизайн.

Всички маркери, независимо от избрания за генотипиране метод, първоначално са анализирани в серия от 94 случая и 184 контроли. След прилагане на точен тест на Фишер, полиморфизмите показали статистически значими разлики в алелните и

генотипните честоти сред случаи и контроли са генотипирани в нова серия от 78 случая и 372 контролни индивида. В последствие, тест на Фишер е проведен отделно върху данните от второто генотипиране, както и върху обобщените данни от двете генотипирания и е извършена корекция за множествени сравнителни процедури.

Вторият етап на научната разработка е посветен на цялостното геномно асоциативно проучване. С помощта на Illumina Bead Array 550K генотипиращ чип първоначално са анализирани 554 496 единични нуклеотидни полиморфизма в 188 български пациенти и 376 здрави индивида (Група 1). След филтриране на получените данни по определени критерии, за по-нататъшен анализ са подбрани общо 497 732 полиморфизма. За валидизиране на резултатите от цялостното геномно сканиране, първите 100 маркера, показали най – значима асоциация със заболяването са генотипирани чрез Inavder метод, комбиниран с мултиплексна полимеразна верижна реакция в същата извадка от 188 случая и 376 контроли (Група 1). С цел потвърждаване на резултатите, стоте маркера са анализирани и в допълнителна серия от 122 случая и 328 контроли (Група 2) със същия метод. Така, общият брой изследвани проби с цел валидизиране и потвърждаване на резултатите от цялостното геномно асоциативно проучване е 310 случая и 704 контроли (Група 1 + Група 2). За статистическа оценка на резултатите от генотипирането в Група 2 е използван точния тест на Фишер, а върху обобщените данни от анализирането на всички проби (Група 1 + Група 2) е приложен тестът на Cochran-Mantel-Haenszel (CMH test). Извършена е и корекция за множествени сравнителни процедури по метода на Бонферони.

Методи за генотипиране на генетични варианти

Метод за генотипиране на полиморфизми чрез алел – специфична хибридизация (TaqMan метод)

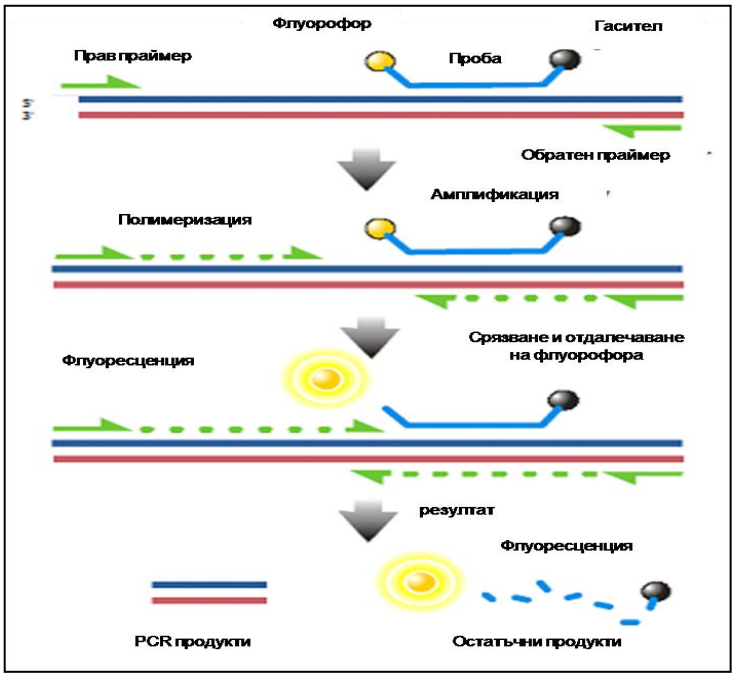
Поради своята висока производителност, специфичност, чувствителност, бързина, точност и финансова рентабилност, TaqMan е предпочитан метод за генотипиране при анализ на сравнително малък брой полиморфизми в голяма извадка. Методът е описан за първи път през 1991 г. от Holland и колеги, а впоследствие Roche Molecular Diagnostics и Applied Biosystems разработват технологията за диагностични и изследователски цели. Понастоящем, компанията Applied Biosystems поддържа и предлага една от най-големите в света колекции на предварително конструирани (pre-

designed), валидизирани и налични проби (установена е минималната алелна честота на полиморфизмите в 2 до 4 популационни групи) за генотипиране на SNPs чрез TaqMan (inventoried probes), както и такива, които се произвеждат и преминават процес на валидизация, само ако бъдат поръчани (made to order probes). При липса на проба за даден полиморфизъм, съществува възможност за създаването й по избор на клиента (custom made probes).

TaqMan пробите използвани за генотипиране на полиморфизми в настоящата научна разработка са предварително конструирани и предоставени от Applied Biosystems, заедно с универсалния TaqMan микс.

Принцип на метода

TaqMan методът комбинира полимеразна верижна реакция и детекция на мутацията в една стъпка, като се използва 5'→3'екзонуклеазната активност на ензима Taq ДНК полимераза. За провеждането на PCR са необходими двойка специфични праймери (прав и обратен) за амплифициране на даден ДНК фрагмент в който се намира нуклеотидната замяна и двойка алел – специфични проби, комплементарни на всеки един от двата алела на изследвания SNP. Всяка проба е двойно белязана, като в 5' края си носи флуорофор (FAM и VIC), а в 3' края си – 'заглушител' на флуоресценцията (minor groove binder, MGB). Когато пробата е интактна, близостта на 'заглушителя' до флуорофора, подтиска неговата флуоресценция, индуцирана от светлинен източник, посредством флуоресцентен резонансен енергиен трансфер (FRET). Процес, при който енергията от възбудена флуорофорна молекула се пренася върху акцепторна молекула - "заглушител" на флуоресцирането. Пробите хибридизират с определена последователност от амплифициращия се ДНК фрагмент, в мястото на нуклеотидната замяна. По време на фазата на удължаване на праймера и синтез на новата верига от Taq ДНК полимеразата, 5'екзонуклеазната активност на ензима разгражда разположената отпред проба, хибридизирала с ДНК матрицата. По този начин флуорофора се отдалечава от 'заглушителя', което води до увеличаване интензитета на флуоресценция. При несъответствие между пробата и таргетната ДНК, поради намалената хибридизация, не се генерира флуоресцентен сигнал (Фигура 1).



Фигура 1: Принцип на TaqMan метода

Методът обхваща следните етапи: 1. изготвяне на реакционна смес; 2. изготвяне на реакционна плака; 3. провеждане на PCR; 4. разчитане и анализ на данните посредством Applied BiosystemsTM 7900HT Real-Time PCR system - ABI Prism® 7900HT Sequence Detection System, която разполага с удобен за работа софтуер.

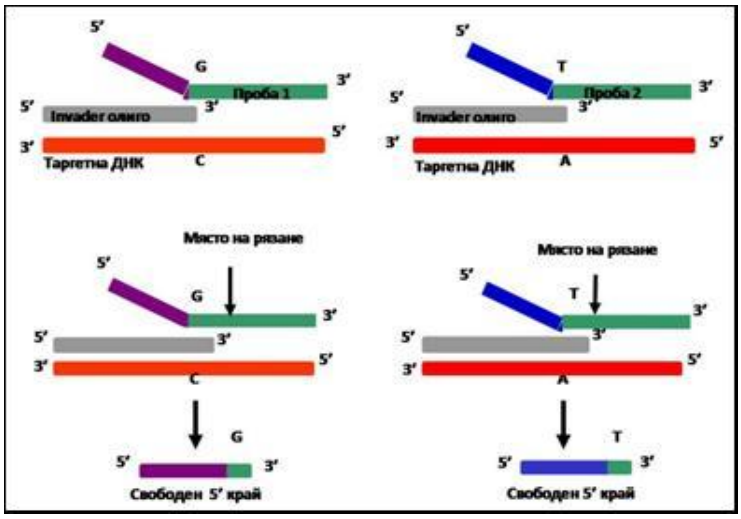
Генотипиране на полиморфизми чрез Invader метод, комбиниран с мултиплексна полимеразна верижна реакция

Принцип на Invader метода

Invader технологията, разработена от Third Wave Technologies (http://www.twt.com) комбинира структурно – специфични ендонуклеазни ензими за рязане и флуоресцентен резонансен енергиен трансфер. Методът протича в две едновременни, циклични, изотермални реакции и разчита на идеалното съвпадение на ензима с неговия субстрат. В първата реакция много прецизно се детектира мястото на нуклеотидната замяна, а втората се използва за амплификация на сигнала и отчитане на резултатите. Всяка Invader реакция включва две таргетни, SNP специфични сигнални проби (съответстващи на ‘дивия’ и на ‘мутантния’ тип алели на изследвания полиморфизъм), Invader олигонуклеотид (проба), две универсални FRET проби (universal fluorescence resonance energy transfer cassettes), белязани с два различни флуорофора (VIC и FAM) и ензим – клеваза (Cleavase enzyme; термостабилна Flap

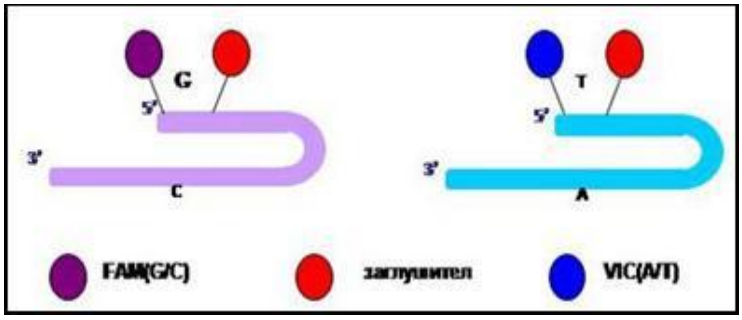
ендонуклеаза (FEN), изолирана от едноклетъчни микроорганизми), която разпознава и реже образуваната при реакцията тройна структура.

В първичната реакция, два олигонуклеотида (проби) хибридизират с едноверижната таргетна ДНК (съдържаща SNP), като образуват специфична тройна структура в мястото на SNP. Тя се разпознава от ензима – клеваза и бива срязана от него, като се освобождава къс ДНК фрагмент, наричан още 5’ рамо (flap) (фигура 2).



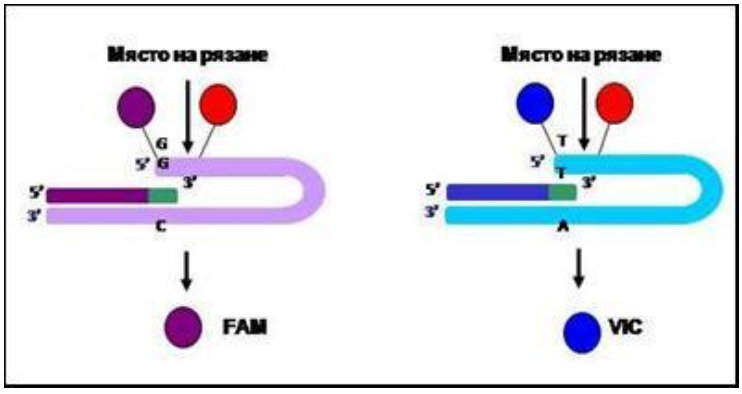
Фигура 2: Първична реакция при Invader

Във вторичната реакция се включват и двете универсални FRET проби, чиято структура е изобразена на фигура 3. На 5’ края си всяка FRET проба носи флуорофор (донорна молекула, VIC и FAM), съответстващ на всеки един от двата алела на анализирания SNP (в случая, FAM за G/C и VIC за A/T). Флуоресценцията на флуорофора първоначално е заглушена от близко разположената до него акцепторна молекула (заглушител).



Фигура 3: Структура на универсалните FRET проби

Освободените при първичната реакция 5’ рамена, участват във вторичната реакция като Invader олиго. Те отново хибридизират с FRET пробите, като образуват специфичната тройна структура. Клевазата отново разпознава образуваната тройна структура и срязва съответната FRET проба. В резултат на това, флуорофора се освобождава от ‘заглушителя’ и се генерира флуоресцентен сигнал (Фигура 4).



Фигура 4: Вторична реакция при Invader

В настоящия дисертационен труд Invader методът е комбиниран с мултиплексна полимеразна верижна реакция (mPCR), която позволява едновременното амплифициране на голям брой различни таргетни ДНК райони в една реакция и генотипирането на повече на брой полиморфизми в големи групи от индивиди.

Основни стъпки при провеждане на mPCR-Invader:

1. Дизайн на праймери за провеждане на мултиплексен PCR

Всички праймери използвани за провеждане на мултиплексната полимеразна верижна реакция са конструирани по собствен дизайн, като са използвани ДНК последователности с размер около 1 kb, за всеки полиморфизъм от базата данни BioMart (Ontario Institute for Cancer Research (OICR) и European Bioinformatics Institute (EBI; http://www.biomart.org/index.html). Получените ДНК последователности са експортирани във вид на текстов файл с който може да се работи в програмата Sequencher (Gene Codes Corporation, http://genecodes.com/). Непосредствено преди дизайна на праймерите всяка таргетна ДНК последователност е проверена за наличие на повторени ДНК елементи посредством опцията Repeat Masker на интернет страницата на Institute for Systems Biology (http://www.repeatmasker.org/). За дизайн на

праймерите е използван олигокалкулатора на SIGMA Genosys (http://www.genosys.jp/adt/SGJ_NN_3rd.html) – фирмата производител.

2. Валидизиране на праймерите

За всяка двойка праймери задължително са проверени условията на работа (DMSO(+) и DMSO(-)) и способността им да инициират амплификация, посредством Test – PCR. При липса на амплификация и при двата вида условия, двойката праймери се тества посредством Touch – down PCR. По–голяма част от използваните в настоящата работа двойки праймери работят в среда без DMSO.

3. Провеждане на мултиплекс PCR

Едновременно се амплифицират 20 таргетни ДНК последователности или при неуспех се провежда Singleplex PCR (амплифициране на една таргетна ДНК последователност).

4. Хоризонтална гел електрофореза

След приключване на реакцията, PCR продуктите се подлагат на хоризонтална нисковолтова гел електрофореза в 1.5% агарозен гел. PCR реакцията се смята за успешно протекла при наличие на един или няколко широки бенда (до около 500 бази).

5. Дизайн на алел специфични и Invader проби за провеждане на Inavder

За дизайн на проби при Invader е необходимо да се знаят 40 – 50 бази от последователността, непосредствено около SNP. При дизайна на пробите е използвана отново програмата Sequencher® 4.14, и същите последователности, върху които са конструирани праймерите за мултиплексния PCR. За определяне температурата на топене (Tm) при синтез на пробите е използван специално създаден за целта онлайн адрес на институт RIKEN (http://www.riken.go.jp/lab-ww/help2/members/kagawa/tmnumber.html). Всички проби за провеждане на Invader са изработени и доставени от Sigma Genosys® Japan.

6. Провеждане на Invader

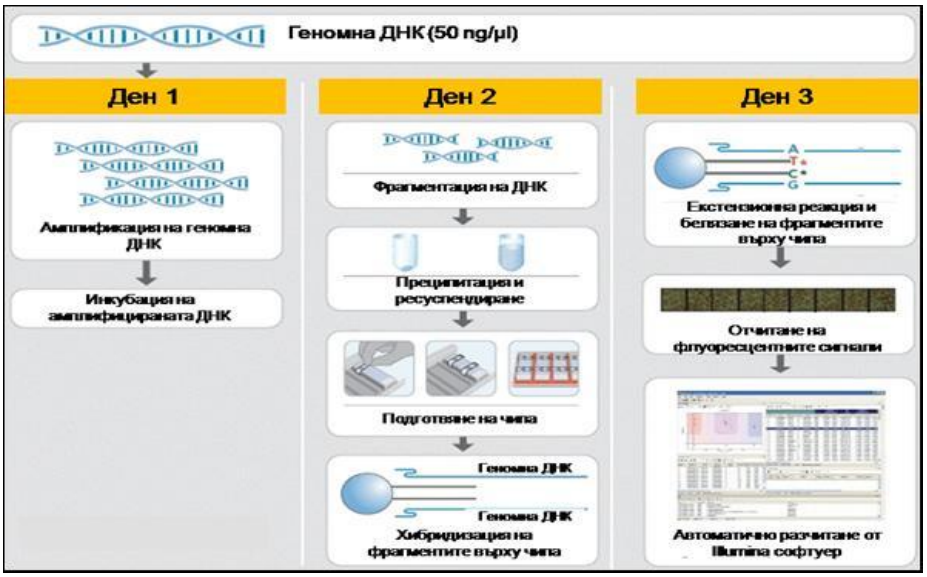
С помощта на роботизирания автоматичен пипетор Beckman Multimek, 2µl от разредения MPCR продукт се разпределят в 384 ямкови плаки. Плаките се оставят да изсъхнат на стайна температура за една нощ. На другия ден се опаковат и съхраняват на 4ºC до провеждане на Invader реакцията. Предварително се приготвя микса от Invader и алелни проби за всеки SNP, изготвя се и реакционния микс. С помощта на

мултиканална пипета, към 384 – ямковата плака, съдържаща изсъхналия вече PCR продукт се добавят по 3 µl реакционен микс. Разчитането на резултатите след Invader става по същия начин, както и след провеждане на TaqMan анализ.

Цялостно геномно асоциативно проучване

Цялостният геномен скрининг беше проведен в лабораторията за развитие на методите за генотипиране (Laboratory for Genotyping Development) към RIKEN институт в Йокохама Япония, върху 550K генотипиращия чип на Illumina (Illumina HumanHap550v3 Genotyping BeadChip), съгласно протокола на производителя за технологията Infinium II.

Методът Infinium II за цялостно геномно генотипиране, разработен от Illumina и комбиниран с технологията BeadArray, позволява едновременното сканиране на стотици хиляди полиморфизми в рамките на един геном. Най – общо, методът обхваща четири етапа. Първият етап включва изотермална амплификация на геномната ДНК и последващо ензим контролирано фрагментиране, което след това не изисква провеждане на гел електрофореза за разделяне на специфични по размер фрагменти. След алкохолна преципитация и ресуспендиране, пробата се нанася върху чипа, което осигурява хибридизация на амплифицираните фрагменти с локус – специфичните 50 – мерни олигонуклеотиди, ковалентно свързани към едно от мънистата (едно мънисто съдържа проби за всеки един от двата алела на даден SNP). Следва ензим-медиирано алел специфично удължаване на олигонуклеотидите върху мънистото (base extension) и флуоресцентно белязане на пробите. Флуоресцентните сигнали се детектират от специален четец (Illumina BeadArray Reader), а обработката на получените данни – идентификация на SNP и неговите алели, се извършва от специално разработен софтуер на Illumina (GenomeStudio Data Analysis Software) (Фигура 5). Статистическата оценка на данните от цялостния геномен скан е извършена с помощта на точния тест на Фишер.



Фигура 5: Принцип на метода Infinium II, комбиниран с BeadArray технологията на Illumina. Методът обхваща четири основни стъпки: 1.цялостно геномно амплифициране и фрагментиране на ДНК; 2.Хибридизация на амплифицираните фрагменти върху чипа; 3.Алел – специфична екстензионна реакция; 4. Белязане на продуктите с флуорохроми (адаптирано от Infinium® II Assay Workflow протокол; © 2006 Illumina, Inc)

Цялостното геномно асоциативно проучване беше проведено върху 550K генотипиращия чип на Illumina (Illumina HumanHap 550K chip), като успешно бяха генотипирани 554 496 единични нуклеотидни полиморфизма в серия от 188 случая с биполарно разстройство и 376 контролни индивида, всички от български произход. След изключване на полиморфизмите локализирани в половите хромозоми и на вариантите в митохондриалната ДНК, получените данни бяха подложени на качествен контрол.

Филтрирането беше извършено въз основа на изчислените от софтуера на Illumina данни чрез филтърната функция на Microsoft Excel.

Валидизиране и реплициране на резултатите от цялостното геномно асоциативно проучване

След филтриране на данните общо 497 732 варианта от всички успешно генотипирани полиморфизми отговориха на гореспоментатите критерии и бяха включени в последващия анализ. Тези полиморфизми бяха подредени по нарастване на най-ниската от трите стойности на p (въз основа на един от трите модела на теста на Фишер) за

всеки SNP, като първите 100 маркера, показали най – силна асоциация със заболяването бяха използвани за индивидуално генотипиране.

С цел валидизиране и реплициране на резултатите от цялостния геномен скан, индивидуалното генотипиране на избраните 100 полиморфизма беше проведено в два етапа. През първия етап стоте кандидат маркера бяха генотипирани в същата серия от случаи и контроли, използвана при цялостното геномно асоциативно проучване. През втори етап, който обхваща реплицирането на резултатите, маркерите бяха анализирани в допълнителна серия от случаи и контроли.

Статистически методи

1. Анализ на равновесието по закона на Харди-Вайнберг (Hardy-Weinberg equilibrium)

Съгласно Закона на Харди-Вайнберг, в идеална популация, което означава достатъчно голяма популация, без действие на отбор спрямо разглеждания ген, без възникване на нови алели на гена в резултат на мутации и без асортативно кръстосване по този ген, съотношението между алелните и генотипните честоти в даден локус остава непроменено с течение на поколенията. В един локус с два алела, доминантен A и рецесивен a и генотипове AA (хомозиготи по домонантния алел), Aa (хетерозиготи) и aa (хомозиготи по рецесивния алел), сумата от алелните и генотипните честоти е единица, или

$$(p+q)2=p2+2pq+q2=1,$$

където p и q са алелните честоти, съответно на доминантния и рецесивния алели.

Равновесните генотипни честоти на хомозиготите по двата алела се изразяват чрез квадрата на съответния алел, а генотипната честота на хетерозиготите – чрез произведението на алелните честоти умножени по две.

Проверка за наличие на равновесие по закона на Харди-Вайнберг е извършена при изследването на полиморфизми в областта на кандидат гени сред двете изследвани групи с помощта на програмата HaploView 4.0 (Broad Institute of MIT & Harvard), като за ниво на значимост е прието $P < 0.001$. Маркерите, показали значително отклонение от принципа на Харди-Вайнберг са изключени от анализ.

2. Точен тест на Фишер (Fisher's exact test)

Точният тест на Фишер принадлежи към групата на непараметричните статистически методи и се прилага при анализ на две независими извадки, независимо от техните характеристики (размер, пропорционалност). Тестът на Фишер е изключително подходящ в случаите, когато използваната извадка е с малък размер, т.е., когато е възможно очакваната честота да е под 10 при една степен на свобода. Наименованието на теста – 'точен', показва възможността значимостта на отклонението от нулевата хипотеза да се изчисли точно, а не на базата на приближения спрямо теоретична извадка.

При използване на точния тест на Фишер може да се установи статистически значима асоциация между две номинални променливи, когато анализираните обекти се класифицират по два различни показателя. По този начин се формира двумерна таблица с размерност 2X2 (2X2 кростаблица) (Таблица 2).

Таблица 2: 2X2 кростаблица при точен тест на Фишер

	Случаи	**Контроли**	**Общо**
Алел 1	a	b	**r1**
Алел 2	c	d	**r2**
Общо	**c1**	**c2**	**t**

r1 – общ брой индивиди в първи ред, носители на алел 1 (a+b); r2 – общ брой индивиди във втори ред, носители на алел 2 (c+d); c1 – общ брой индивиди в първа колона, случаи (a+c); c2– общ брой индивиди във втора колона, контроли (b+d); t – общ брой анализирани индивиди (a+b+c+d). Адаптирано от http://statpages.org/ctab2x2.html

Съгласно нулевата хипотеза (H0) между две групи не съществува статистически значима асоциация, а за проверка достоверността на хипотезата е прието ниво на значимост P = 0.05. Според алтернативната хипотеза (HA), между двете изследвани групи в извадката съществува статистически значима, неслучайна разлика. Тази хипотеза може да бъде левостранна, когато между променливите се очаква да има негативна асоциа-ция, дясностранна – очаква се позитивна асоциация или двустранна – алтернативата не е предварително известна.

Тестът на Фишер бива едностранен (one – tailed) или двустранен (two – tailed). При едностранния тест предварително се знае посоката на разликата между параметъра и оценката, а вероятността (P) се изчислява като сума от стойността на P във всички

възможни случаи на хипергеометрично разпределение, при които разпределението на данните в клетките със същата насока на разликата е същото като емпирично полученото или още по-крайно. При двустранният тест, вероятността (P) е сума от стойността на P във всички възможни случаи на същото или по-крайно разпределение на данните, но и в двете посоки на стохастично разпределение, независимо от посоката на връзката. Стойност на P ≤ 0.05 показва, че вероятността полученото разпределение да е случайно е по-малка или равна на 5%.

В настоящата научна разработка е приложен двустранния точен тест на Фишер съгласно три модела: на алелните честоти (allele – frequency mode), на доминантно унаследяване (dominant – inheritance mode) и на рецесивно унаследяване (recessive – inheritance mode), като по този начин за всеки изследван полиморфизъм са получени три възможни зависимости, както по отношение на алелните честоти, така и по отно-шение на генотипните честоти на хомозиготите по единия или другия алтернативен алел. Тестът на Фишер и използван за установяване на статистически значима асоциа-ция както при асоциативното проучване на кандидат гени, така и след цялостното ге-номно асоциативно проучване, а също и при валидизирането и реплицирането на ре-зултатите след цялостния геномен скан. Изчисленията са направени с помощта на Microsoft Excel и програма Perl в DOS.

3. Корекция при множествени изследвания

Множественият сравнителен анализ (корекция при множествени изследвания; Correction for multiple testing) е процедура, с помощта на която се коригира стойността на величината P (P – value), получена при провеждане на голям брой аналогични статистически тествания. Прилагането на тази процедура има за цел да намали вероят-ността от възникване на грешка от I род (α грешка, получаване на фалшиво позитивен резултат), изразяваща се в погрешно отхвърляне на нулевата хипотеза. Прието ниво на значимост от 0.05 означава, че вероятността за възникване на грешка от I род е 5%, т.е. вероятността да сме получили статистически значим резултат случайно е 5%.

Една от най – често прилаганите процедури за множествено сравняване, но и много консервативна е корекцията на Бонферони (Bonferroni correction), която има за цел да намали риска за възникване на грешки от I род при многократно използване на даден статистически анализ. Съгласно тази корекция, при проведени n на брой незави-сими изследвания, корекцията на нивото на значимост за целия обем на експеримента (Ptotal), се изчислява като всяко отделно изследване се разгледа при ниво на значимост равно на 1/nта

от нивото, което трябва да се приеме при едно проведено изследване и се изразява по следния начин:

$$Ptotal=P/n,$$

където P е нивото на значимост в едно изследване, прието за 0.05, а n е броят на проведените изследвания. Така получената стойност представлява коригираното ниво на значимост, съгласно броя направени изследвания, като всички тествания при които p продължава да има стойност по – ниска от зададената се интерпретират като данни за наличие на истинска асоциация между даден вариант и заболяването.

Въпреки че прилагането на корекцията на Бонферони намалява риска за грешка от I род, поради своята консервативност, тя увеличава риска за възникване на грешка от II род (β грешка, получаване на фалшиво отрицателен резултат), изразяваща се в нарас-тване на вероятността за отхвърляне на истински асоциации.

В настоящия научен труд, корекцията на Бонферони е приложена както при асо-циативното изследване на кандидат гени, така и след получаване на данните от цялост-ното геномно асоциативно проучване и филтрирането на резултатите.

4. Други статистически методи

Неравновесната скаченост между изследваните маркери (D'/r^2) е изчислена с помощта на програмата HaploView 4.0 (Broad Institute of MIT & Harvard), а LD – блоковите структури са конструирани по метода "solid spine of LD" (в програмата HaploView), съгласно който първият и последният маркер в хаплотипния блок трябва да са в силна неравновесна скаченост с междинните SNPs, като те самите не е необхо-димо да са в неравновесна скаченост. Хаплотипният анализ беше извършен с помощта на програмата SNPALyze версия 3.1 (Dynacom, Chiba, Japan, http://www2.dynacom.co.jp), а също така и с програмата HaploView 4.0.

За оценяване размера на ефекта (effect size, сила на връзката между две промен-ливи при статистически анализ) е анализирано съотношението на шансовете (Odds Ratio, OR). OR е относителна мярка за риска или вероятността дадено събитие да се случи сред дадена група обекти, когато е налице даден фактор, в сравнение с група обекти, при които факторът липсва. Оценка на 95% доверителен интервал (95%CI) и съотно-шението на шансовете е направена на базата на заключенията за рисков алел за всеки полиморфизъм поотделно.

Статистическите изчисления след цялостното геномно асоциативно проучване са извършени с помощта на програмата R в лабораторията по статистически анализи към институт RIKEN (Laboratory of Statistical Analysis, RIKEN Center for Genomic Medicine, Tokyo, Japan) от Проф. Н. Каматани и А. Такахаши, съвместно с Д-р М. Кубо от лабораторията по развитие на методите за генотипиране в същия институт (Laboratory for Genotyping Development, RIKEN Center for Genomic Medicine, Yokohama, Japan). Програмата R е използвана и за построяване на квантилна вероятностна диаграма (Q – Q plot), с цел установяване наличието на популационна стратификация сред изследваните групи. За оценка на статистически значимите разлики от цялостния геномен скан е построена диаграма тип Manhattan (Manhattan plot). За целта е използвана програмата HaploView 4.0 и данните за хромозомната локализация и трансформираната в отрицателен десетичен логаритъм стойност на p за всеки изследван SNP.

За установяване на асоциация след индивидуалното генотипиране на стоте кандидат маркера, върху обобщените данни (валидизиране и реплициране на резултатите) от цялостния геномен скан е приложен тестът на Cochran-Mantel-Haenszel (CMH test statistics), който отчита евентуалното наличие на популационна стратификация сред използваната извадка. Тестът е приложен със съдействието на Д-р Й. Каматани от лабораторията по молекулярна медицна към института по медицински науки в Токио (Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, Japan).

РЕЗУЛТАТИ

Резултати от асоциативното проучване на кандидат гени

В първият етап от нашето проучване, кандидат полиморфизмите бяха подбрани от публично достъпните бази данни на NCBI и HapMap (www.ncbi.nlm.nih.gov; www.hapmap.org/index.html.en) въз основа на съществуващи данни за асоциация със заболяването. Общо 199 маркера бяха селектирани за генотипиране в този първи етап на проучването, като повечето от тях са ‘представителни’ SNPs (tagSNPs) във фаза II на HapMap проекта с $r2>0.8$ и честота на редкия алел 0.1 ($MAF\geq0.1$) в Европейска популация (CEU population), с изключение на тези, асоциирани със заболяването в други популации.

След изготвянето на списъка с избраните кандидат полиморфизми, за всеки един от тях беше определен метода за генотипиране. За 47 маркера беше избран методът TaqMan с предварително конструирани, преминали процес на валидизация и поддържани в наличност проби от Applied Biosystems, а останалите 152 полиморфизма бяха анализирани

чрез Invader метод, комбиниран с мултиплексна полимеразна верижна реакция, с проби и праймери, конструирани по собствен дизайн.

Маркерите, за които беше избран методът TaqMan, бяха генотипирани успешно, като един от тях, rs1058573 (*5HTR2A*) се оказа мономорфен за нашата популация и беше изключен от изследването. От останалите 152 маркера, генотипирани чрез метода Invader, резултат беше получен за 144 от тях. За 8 полиморфизма генотипирането се оказа неуспешно и разграничаване на отделните клъстери от флуоресцентни сигнали не беше постигнато, дори и след използването на два отделни набора от Invader проби. От 144 успешно генотипирани SNPs, след анализ на разпределението на алелните и генотипни честоти, 9 показаха отклонение от принципа на Харди–Вайнберг за равновесието в контролната група индивиди, поради което бяха изключени от анализ. Два SNPs, rs872123 (*GAD1*) и rs16927997 (*GRP78*/*MIF2*) се оказаха мономорфни за изследваната извадка и също не бяха включени в последващия анализ.

Всички маркери бяха генотипирани в серия от 94 случая и 184 контроли, като валиден резултат беше получен за 179 полиморфизма от 65 гена. За установяване на асоциация с биполарното разстройство, върху получените данни от успешно генотипираните маркери, беше приложен двустранния точен тест на Фишер.

В този първи етап на изследването, при ниво на статистическа значимост, P=0.05, общо 17 SNPs показаха предполагаема асоциация със заболяването (P стойности по–малки от 0.05), поради което бяха подложени на допълнителен анализ в допълнителна извадка от 78 случая и 372 контроли. Генотипирането на тези полиморфизми във втората група от случаи и контроли беше извършено с пробите в наличност и със същия метод, с който е бил анализиран съответния маркер в първия етап на изследването. Статистическият анализ беше проведен с точен тест на Фишер отделно върху резултатите от първи и втори скрининг, а също и върху обобщените данни от генотипирането на всички проби. Подробна информация относно генотипирането на избраните 17 полиморфизма е посочена в таблица 3.

Таблица 3: Обобщени резултати от генотипирането на 17 SNPs в 172 случая с биполарно разстройство и 556 контроли

Локус	Ген	db SNP		Случаи (*n*=172)			Контроли (*n*=556)			Честота, случаи			Честота, контроли			Риск алел	P-стойност		
				11	12	22	11	12	22	11	12	22	11	12	22		1vs2	vs11	vs22
4p16.1	*GPR78*	rs9799720	1ви	2	26	66	0	36	148	0.02	0.28	0.70	0.00	0.20	0.80	1	0.038	0.11	0.070
			2ри	1	15	62	2	66	300	0.01	0.19	0.79	0.01	0.18	0.82	1	0.66	1.00	0.75
			Общо	3	41	128	2	102	448	0.02	0.24	0.74	0.00	0.19	0.81	1	**0.035**	0.090	0.065
6p21.3	*GRM4*	rs1565366	1ви	5	35	54	3	53	128	0.05	0.37	0.57	0.02	0.29	0.70	1	0.029	0.12	0.047
			2ри	5	14	59	10	105	254	0.06	0.18	0.76	0.03	0.28	0.69	2	0.72	0.16	0.28
			Общо	10	49	113	13	158	382	0.06	0.28	0.66	0.02	0.29	0.69	1	0.17	**0.029**	0.45
6p21.3	*GRM4*	rs3798529	1ви	50	40	4	131	48	5	0.53	0.43	0.04	0.71	0.26	0.03	2	0.0063	0.0035	0.72
			2ри	56	19	3	255	104	12	0.72	0.24	0.04	0.69	0.28	0.03	1	0.73	0.69	1.00
			Общо	106	59	7	386	152	17	0.62	0.34	0.04	0.70	0.27	0.03	2	0.064	0.062	0.62
7q36.1	*HTR5A*	rs1800883	1ви	8	37	49	39	69	73	0.09	0.39	0.52	0.22	0.38	0.40	2	0.0050	0.0066	0.073
			2ри	8	21	48	60	152	160	0.10	0.27	0.62	0.16	0.41	0.43	2	**0.0038**	0.23	**0.0025**
			Общо	16	58	97	99	221	233	0.09	0.34	0.57	0.18	0.40	0.42	2	**0.000097**	**0.0081**	**0.00085**
8p22-p21	*DPYSL2*	rs11863	1ви	21	44	28	68	84	32	0.23	0.47	0.30	0.37	0.46	0.17	2	0.0028	0.020	0.020
			2ри	19	42	17	103	181	88	0.24	0.54	0.22	0.28	0.49	0.24	2	0.93	0.58	0.77
			Общо	40	86	45	171	265	120	0.23	0.50	0.26	0.31	0.48	0.22	2	0.055	0.068	0.21
8p22-p21	*DPYSL2*	rs17666	1ви	7	36	49	34	83	67	0.08	0.39	0.53	0.18	0.45	0.36	2	0.0019	0.019	0.0096
			2ри	11	28	39	56	160	156	0.14	0.36	0.50	0.15	0.43	0.42	2	0.31	0.86	0.21
			Общо	18	64	88	90	243	223	0.11	0.38	0.52	0.16	0.44	0.40	2	**0.0045**	0.084	**0.0079**
9q34.3	*GRIN1*	rs11146020	1ви	79	7	1	146	34	1	0.91	0.08	0.01	0.81	0.19	0.01	1	0.068	0.049	1.00
			2ри	65	13	0	301	62	5	0.83	0.17	0.00	0.82	0.17	0.01	1	0.65	0.87	0.59
			общо	144	20	1	447	96	6	0.87	0.12	0.01	0.81	0.17	0.01	1	0.082	0.10	0.70
11p13	*BDNF*	rs6265	1ви	68	25	1	107	64	12	0.72	0.27	0.01	0.58	0.35	0.07	1	0.0079	0.026	0.067
			2ри	52	25	1	221	131	19	0.67	0.32	0.01	0.60	0.35	0.05	1	0.14	0.25	0.22

			Общо	120	50	2	328	195	31	0.70	0.29	0.01	0.59	0.35	0.06	1	**0.0033**	**0.015**	**0.019**
12p13	*GNB3*	rs5443	1ви	41	37	16	74	93	16	0.44	0.39	0.17	0.40	0.51	0.09	2	0.57	0.70	0.048
			2ри	31	32	15	180	159	30	0.40	0.41	0.19	0.49	0.43	0.08	2	**0.017**	0.17	**0.0050**
			общо	72	69	31	254	252	46	0.42	0.40	0.18	0.46	0.46	0.08	2	**0.018**	0.38	**0.00051**
12q21.1	*TPH2*	rs1007023	1ви	2	34	58	4	44	136	0.02	0.36	0.62	0.02	0.24	0.74	1	0.069	1.00	0.039
			2ри	1	20	57	7	89	275	0.01	0.26	0.73	0.02	0.24	0.74	1	1.00	1.00	0.89
			Общо	3	54	115	11	133	411	0.02	0.31	0.67	0.02	0.24	0.74	1	0.12	1.00	0.079
12q22-q23	*DUSP6*	rs2279574	1ви	14	54	26	57	93	33	0.15	0.57	0.28	0.31	0.51	0.18	2	0.0041	0.0035	0.088
			2ри	14	41	23	84	186	98	0.18	0.53	0.29	0.23	0.51	0.27	2	0.43	0.37	0.67
			Общо	28	95	49	141	279	131	0.16	0.55	0.28	0.26	0.51	0.24	2	**0.026**	**0.013**	0.23
12q24.2-q24.31	*NOS1*	rs532967	1ви	2	17	75	2	57	123	0.02	0.18	0.80	0.01	0.31	0.68	2	0.10	0.61	0.035
			2ри	6	32	40	6	111	255	0.08	0.41	0.51	0.02	0.30	0.69	1	**0.00084**	**0.0087**	**0.0041**
			Общо	8	49	115	8	168	378	0.05	0.28	0.67	0.01	0.30	0.68	1	0.33	**0.019**	0.78
15q11.2-q12	*GABRA5*	rs140681	1ви	0	14	77	3	44	136	0.00	0.15	0.85	0.02	0.24	0.74	2	0.048	0.55	0.064
			2ри	1	18	58	1	74	297	0.01	0.23	0.75	0.00	0.20	0.80	1	0.32	0.31	0.44
			Общо	1	32	135	4	118	433	0.01	0.19	0.80	0.01	0.21	0.78	2	0.55	1.00	0.52
18p11.21	*MPPE1*	rs3974590	1ви	35	44	15	106	64	14	0.37	0.47	0.16	0.58	0.35	0.08	2	0.00059	0.0015	0.038
			2ри	37	37	4	178	149	45	0.47	0.47	0.05	0.48	0.40	0.12	1	0.45	1.00	0.11
			Общо	72	81	19	284	213	59	0.42	0.47	0.11	0.51	0.38	0.11	2	0.10	**0.036**	0.89
22q11.2	*DGCR8*	rs2073778	1ви	65	27	2	152	29	3	0.69	0.29	0.02	0.83	0.16	0.02	2	0.019	0.014	1.00
			2ри	62	15	1	278	81	8	0.79	0.19	0.01	0.76	0.22	0.02	1	0.51	0.56	0.71
			Общо	127	42	3	430	110	11	0.74	0.24	0.02	0.78	0.20	0.02	2	0.35	0.25	1.00
22q13.1	*SYNGR1*	rs909685	1ви	16	35	43	14	77	93	0.17	0.37	0.46	0.08	0.42	0.51	1	0.10	0.024	0.53
			2ри	8	30	39	55	155	159	0.10	0.39	0.51	0.15	0.42	0.43	2	0.16	0.37	0.26
			Общо	24	65	82	69	232	252	0.14	0.38	0.48	0.12	0.42	0.46	2	0.90	0.60	0.60
22q13.1	*SYNGR1*	rs715505	1ви	55	29	9	110	66	6	0.59	0.31	0.10	0.60	0.36	0.03	2	0.33	0.90	0.046
			2ри	46	25	6	207	135	27	0.60	0.32	0.08	0.56	0.37	0.07	1	0.69	0.61	1.00

Общо	101	54	15	317	201	33	0.59	0.32	0.09	0.58	0.36	0.06	2	0.89	0.72	0.22

$1^{ви}$ - първи етап на изследването, използвани 94 случая и 184 контроли; $2^{ри}$ –втори етап, използвани 78 случая и 372 контроли; Общо– комбиниран анализ на 172 случая и 556 контроли.

В обобщение, в първия етап на асоциативното проучване успешно бяха генотипирани 179 полиморфизма от 65 кандидат гена в 94 случая и 184 контролни индивида. Статистическата обработка на данните с двустранен точен тест на Фишер разкри възможната асоциация на 17 маркера с биполарното разстройство. Във втория етап на изследването тези полиморфизми бяха подложени на допълнителен молекулярно – генетичен анализ в 78 случая и 372 контроли. Статистическият анализ на данните само от втория скрининг показа възможна асоциация на заболяването с три полиморфизма, а именно: rs1800883 (*HTR5A*) с $P=2.5X10^{-3}$, rs5443 (*GNB3*) с $P=5.00X10^{-3}$ и rs532967 (*NOS1*) с $P=8.40X10^{-4}$.

Комбинираният анализ на данните от генотипирането във всички проби (172 случая и 556 контроли), установи възможна асоциация на биполарното разстройство с общо девет маркера: rs9799720 (*GPR78*) с $P=3.50X10^{-2}$, rs1565366 (*GRM4*) с $P=2.90X10^{-2}$, rs1800883 (*HTR5A*) с $P=9.70X10^{-5}$, rs17666 (*DPYSL2*) с $P=4.50X10^{-3}$, rs6265 (*BDNF*) с $P=3.30X10^{-3}$, rs5443 (*GNB3*) с $P=5.10X^{-4}$, rs2279574 (*DUSP6*) с $P=1.30X10^{-2}$, rs532967 (*NOS1*) с $P=1.90X10^{-2}$, rs3974590 (*MPPE1*) с $P=3.60X10^{-2}$.

Най–силна статистически значима асоциация показа вариант rs1800883 (C→G замяна), разположен в промоторния участък на гена, кодиращ рецептора за серотонин (*HTR5A*, 5-hydroxytryptamine (serotonin) receptor 5A). Генотипиранто на този полиморфизъм се оказа неуспешно за 1 индивид от групата на случаите (0.6%) и за трима от контролната група (0.5%), което е индикация, че установената асоциация не е резултат от грешка при генотипирането. Рисковият генотип (GG) беше наблюдаван с много по–голяма честота сред пациентите - 56.7%, в сравнение с 42.1% от индивидите в контролната група (P=0.00085, odds ratio (OR)=1.80; 95% CI, 1.27–2.54). Рисковият алел (G) беше наблюдаван много по–често в групата на случаите с алелна честота от 73.7%, в сравнение с 62.1% честота при контролите (P=0.000097).

Слаба асоциация беше наблюдавана и за SNP rs6265, водещ до амино – киселинна замяна (Val66Met), разположен в екзон 2 на гена кодиращ невротрофния фактор в мозъка (BDNF, brain-derived neurotrophic factor). Рисковият алел (C) беше наблюдаван с по–голяма честота в групата на случаите, в сравнение с контролната група индивиди (P=0.0033).

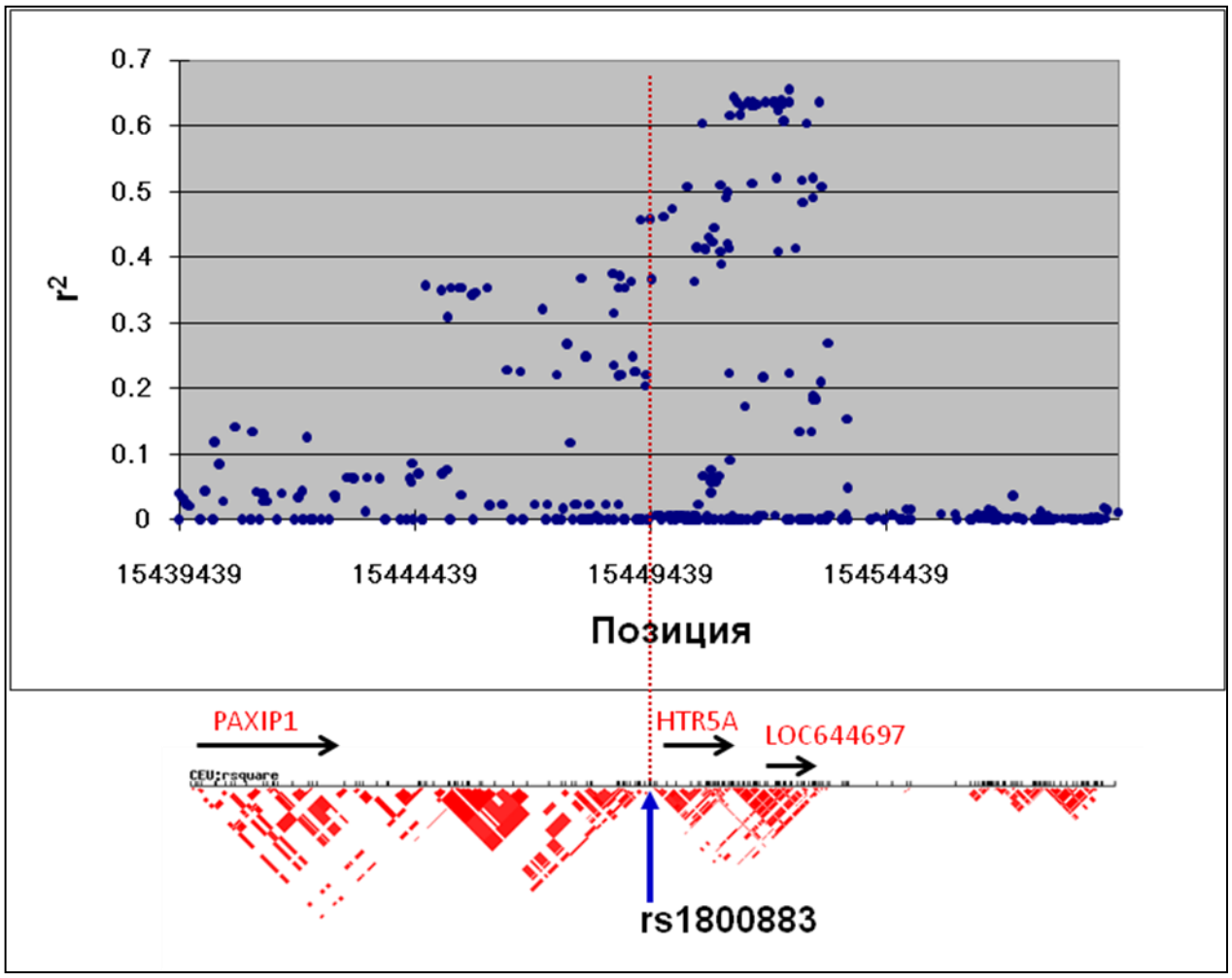
Слаба асоциация показа и вариант rs5443, локализиран в екзон 10 на гена кодиращ β субединицата на гуанин нуклеотид – свързващия протеин (GNB3, guanine nucleotide binding protein (G protein), beta polypeptide 3). 18% от случаите са хомозиготни носители на рисковия генотип (TT), в сравнение с 8% при контролите (P=0.00051).

След прилагане корекцията на Бонферони, само вариант rs1800883 в *HTR5A* гена остана статистически значимо асоцииран с биполарното разстройство, с коригирана P стойност от 0.017. Отчитайки вероятността за възникване на грешка от II род (получаване на фалшиво отрицателен резултат) при прилагане на корекцията на Бонферони, евентуалната асоциация на останалите полиморфизми със заболяването не може да бъде напълно изключена.

Резултати от анализа на връзката между варианти в *HTR5A* и биполарното разстройство. Фино картиране и хаплотипен анализ

За да разберем дали rs1800883 не е маркер за друг, все още неанализиран вариант, който се намира в неравновесна скаченост с него и е включен в етиологията на биполарното разстройство, предприехме фино картиране на региона около маркерния SNP.

Използвайки базата данни на HapMap за европейска популация (CEU), проучихме LD – блоковата структура в района на гена *HTR5A* и оценихме стойностите на r^2 (като мярка за LD) между rs1800883 и други представителни SNPs (tSNPs) в района на и около гена. Както се вижда от фигура 6, маркерният полиморфизъм (rs1800883) е локализиран между два LD блока, и показва по–силна неравновесна скаченост спрямо маркерите в десния LD блок ($r^2 \geq 0.5$), в сравнение с тези в левия.



Фигура 6: Диаграма на LD – блоковата структура на *HTR5A*, съгласно базата данни на HapMap за европейска популация (долната част на фигурата) и разпределение на стойностите на r^2 (по ордината) между rs1800883 и други представителни полиморфизми в района на *HTR5A*, според позицията им. Всяка точка от диаграмата е отделен полиморфизъм (база данни HapMap, 2007)

С помощта на базата данни на HapMap и програмата HaploView, от двата LD – блока за анализ бяха подбрани общо 12 представителни полиморфизма. Генотипирането на маркерите беше извършено посредством метода Invader, комбиниран с мултиплексна полимеразна верижна реакция, първоначално в 94 слуая и 184 контроли, с праймери и проби по собствен дизайн. Статистическата обработка на данните след това първо генотипиране, разкри статистически значима разлика за четири полиморфизма, а именно: rs4960681 (*LOC644697*) с минимално $P=2.0X10^{-2}$ (OR=1.83 и 95% CI от 1.11-3.03), rs2581841 (*HTR5A*) с минимално $P=1.6X10^{-2}$ (OR=1.88 и 95% CI от 1.14-3.10), rs7807281 (*LOC644697*) с минимално $P=2.6X10^{-3}$ (OR=1.95 и

95% CI от 1.16-3.26), rs10254608 (*LOC644697*), с минимално P=2.4X10^{-3} (OR=2.22 и 95% CI от 1.29-3.80).

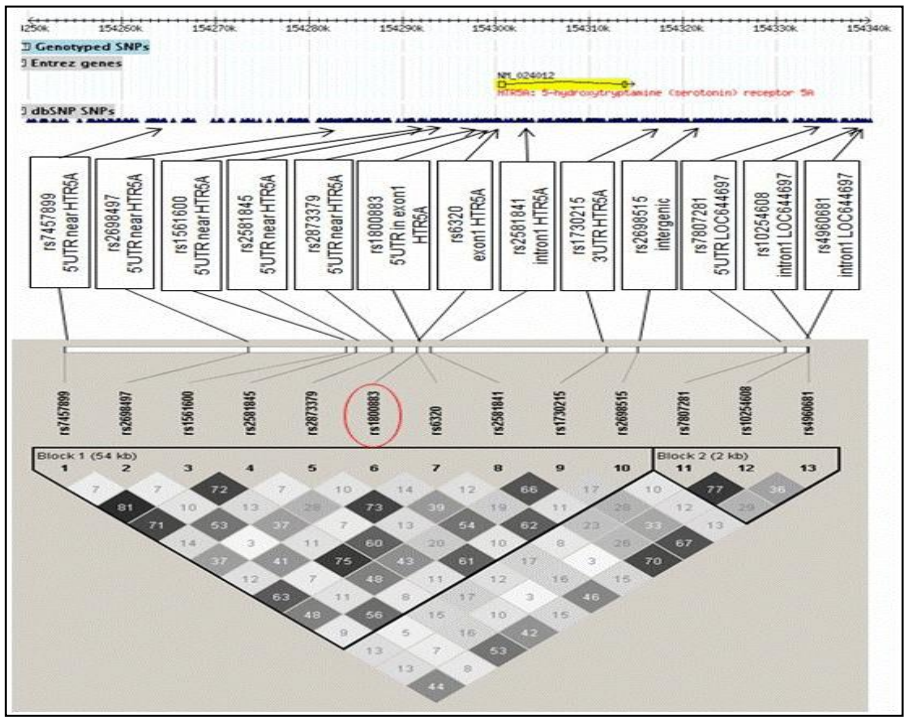
Тези четири полиморфизма бяха подложени на допълнителен анализ в останалите случаи и контроли. Двустранен точен тест на Фишер беше приложен, както върху данните от второто генотипиране, така и върху обобщените данни от двете генотипирания. Като резултат, два от четирите полиморфизма показаха статистически значима разлика, с P стойност близка до тази на rs1800883. Нито един от тези два полиморфизма обаче, не достигна статистическата значимост на маркерния SNP. В таблица 4 са посочени данни относно генотипирането на двата полиморфизма във всички проби, статистическата им значимост след Фишер тест, както и данни относно генотипирането на rs1800883 във всички налични проби.

Таблица 4: Обобщени данни от генотипирането на два представителни полиморфизма, в района на *HTR5A*, показали стойност на P близка до тази на маркерния SNP. Посочени са и данните от анализирането на rs1800883 във всички проби.

Ген	SNP	Случаи, генотип (*n*=172)			Контроли, генотип (*n*=556)			Случаи, честота			Контроли, честота			Рисков алел	P – стойност			OR	95% CI
		11	12	22	11	12	22	11	12	22	11	12	22		pa	p11	p22		
HTR5A	**rs1800883**	**16**	**58**	**97**	**99**	**221**	**233**	**0.09**	**0.34**	**0.57**	**0.18**	**0.40**	**0.42**	**2**	**$9.70X10^{-5}$**	**$8.10X10^{-3}$**	**$8.50X10^{-4}$**	**1.80**	**(1.27-2.54)**
HTR5A	rs2581841	17	56	99	59	259	237	0.10	0.33	0.58	0.11	0.47	0.43	2	$8.10X10^{-3}$	$8.90X10^{-1}$	$6.50X10^{-4}$	1.82	(1.29-2.57)
LOC644697	rs10254608	44	88	40	211	257	86	0.26	0.51	0.23	0.38	0.46	0.16	2	$9.00X10^{-4}$	$3.40X10^{-3}$	$2.20X10^{-2}$	1.79	(1.22-2.62)

Означенията са както в предната таблица

За да установим степента на неравновесна скаченост между изследваните маркери за нашата популация, използвахме данните от генотипирането на 12те представителни полиморфизма, заедно с маркерния SNP и с помощта на програма HaploView, проследихме структурата на LD в района на гена *HTR5A*. Както се вижда на фигура 7, анализираните 13 маркера принадлежат на два LD блока (блок 1 с големна 54 kb и блок 2 с големина 2 kb), в които първият и последният маркер се намират в силна неравновесна скаченост с междинните полиморфизми, като те самите не е необходимо да са в неравновесна скаченост по между си (т.нар.'solid spine of LD' approach в програмата HaploView). Степента на неравновесна скаченост е представена чрез стойностите на r^2. Черните фигури отразяват $r^2=1$ (цифрите не се изписват във фигурата), т.е. между съответните маркери съществува пълна неравновесна скаченост, сивите - $0< r^2<1$, като стойността е означена в съответната фигура, а белите отразяват $r^2=0$, т.е. липса на неравновесна скаченост между съответните маркери (случайно рекомбиниране на маркерите).



Фигура 7: Диаграма на LD структурата в района на гена *HTR5A* и локализация на анализираните представителни полиморфизми, според базата данни на NCBI (горната част на фигурата). За построяване на LD блоковите структури (очертаните с черен контур триъгълни фигури) е използван подходът 'solid spine of LD ap-proach' в програмата HaploView 4.0.

Хаплотипният анализ беше проведен с трите маркера показали статистически значима асоциация (rs1800883, rs2581841 и rs10254608) с програмата SNPALyze 3.1. Таблица 5 демонстрира честотите на установените осем хаплотипа при случаите и контролите.

Таблица 5: Хаплотипен анализ с 3 маркера в *HTR5A* гена при 172 случая и 556 контроли

N	Хаплотип 123*	Честота		P стойност*
		Случаи	Контроли	Odds Ratio (95% доверителен интервал)
X1	G-T-A	0.46	0.36	0.00045; 1.6 (1.22 - 1.99)
X2	C-C-C	0.19	0.28	0.00041; 0.6 (0.43 - 0.78)
X3	G-T-C	0.21	0.22	незначим
X4	C-T-C	0.05	0.06	незначим
X5	G-C-C	0.06	0.05	незначим
X6	C-T-A	0.01	0.03	незначим
X7	C-C-A	0.01	0.01	незначим
X8	G-C-A	0.00	0.00	незначим

1 - rs1800883; 2 - rs2581841; 3 - rs10254608;
* P стойностите са базирани на точен тест на Фишер.

Най–често срещаният хаплотип в нашата популация (X1) е установен при 46% от пациентите в сравнение с 36% при контролната група и включва рисковите алели на трите маркера. Въпреки това, този хаплотип показа по–слаба асоциация със заболяването в сравнение с асоциацията на индивидуалния маркер, rs1800883.

Резултати от широко мащабното асоциативно проучване

Цялостния геномен скрининг върху биполарното разстройство беше проведен в два етапа. В първият етап на изследването с помощта на 550K генотипиращия чип на Illumina успешно бяха анализирани повече от 500 000 полиморфизма за 188 пациенти със заболяването и 376 контролни индивида. Във втория етап с цел валидизиране на резултатите от цялостния геномен скан, 100 маркера бяха генотипирани в същата серия от случаи и контроли, а реплицирането беше проведено в допълнителна серия от проби (122 случая и 328 контроли).

За провеждане на цялостния геномен скрининг, ДНК пробите, предварително доведени до крайна концентрация 50ng/µl, бяха транспортирани до лабораторията за развитие на генотипиращите методи в института RIKEN в Йокохама, Япония (Laboratory for Genotyping Development, RIKEN Center for Genomic Medicine, Yokohama,

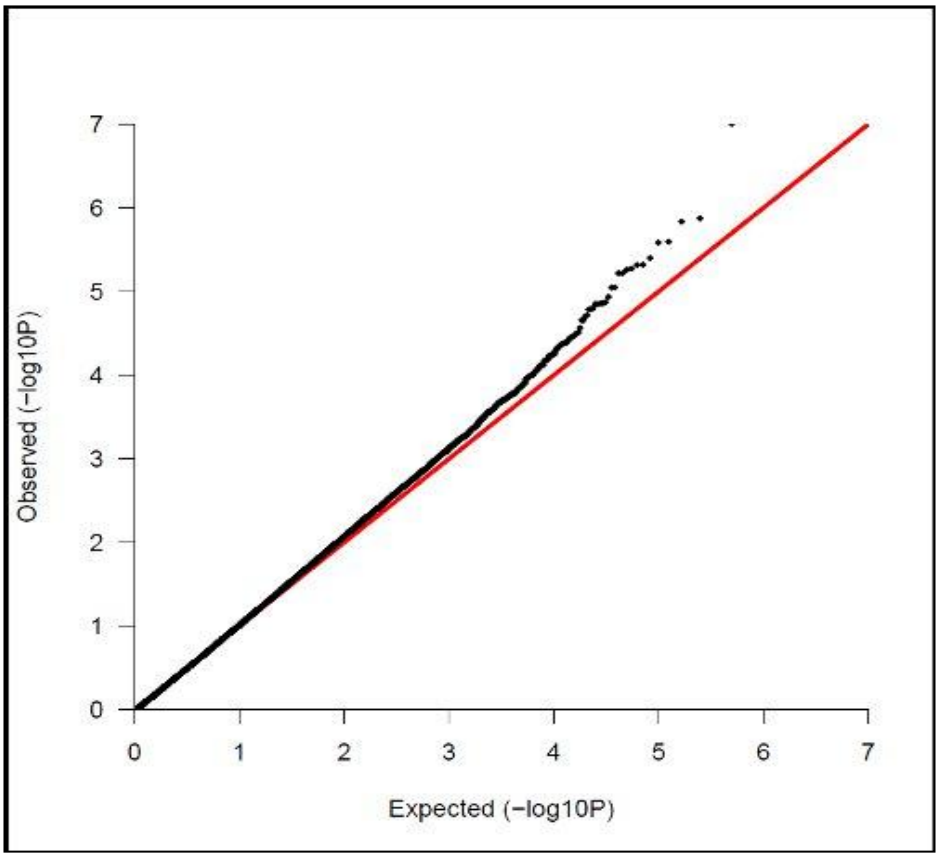
Japan), където генотипирането беше извършено съгласно протокола на фирмата производител (Illumina, Inc. San Diego, California, USA). Получените след скрининга данни бяха подложени на статистически анализ в лабораторията за статистически анализ към същия институт (Laboratory for Statistical Analysis, RIKEN Center for Genomic Medicine, Yokohama, Japan), като за всеки анализиран маркер бяха установени: броя успешно генотипирани индивиди, алелните и генотипни честоти, неговата значимост, изразена чрез стойността P с двустранен точен тест на Фишер за трите модела на унаследяване, рисковия алел, качеството на получения сигнал върху чипа сред случаи, контроли и общо, честотата на по–рядко срещания алел (MAF) за случаи, контроли и общо, стойността на P по отношение на закона на Харди–Вайнберг за равновесието, общо и поотделно (за случаи и контроли), съотношението на шансовете (OR). Разпределението на стойността на P за всичките 554 496 анализирани полиморфизма, след тази първа обработка на данните, е посочено в таблица 6

Таблица 6: Разпределение на стойностите на P след цялостния геномен скан и провеждане на двустранен точен тест на Фишер, съгласно трите модела на унаследяване

P value	Брой полиморфизми		
	vs11	vs22	1vs2
≥0.01	550 020	549 746	548 604
<0.01	4031	4276	5232
<0.001	393	409	577
<0.0001	49	56	69
<0.00001	3	8	13
<0.000001	0	1	1
Общо	**554 496**	**554 496**	**554 496**

Получените данни бяха подложени на филтриране по определени критерии, приети и използвани от лабораторията за развитие на генотипиращите методи в института RIKEN в Япония. В резултат на проведения качествен контрол, общо 497 732 маркера, които показаха качество на получения флуоресцентен сигнал (SNP call rate) ≥0.99, сред случаи и контроли, честота на редкия алел (MAF) > 0 и за двете изследвани групи и стойност на P по отношение на принципа на Hardy-Weinberg за равновесието ≥0.000001, в контролната група, бяха подложени на последващ анализ, с цел установяване на асоциация с биполарното разстройство.

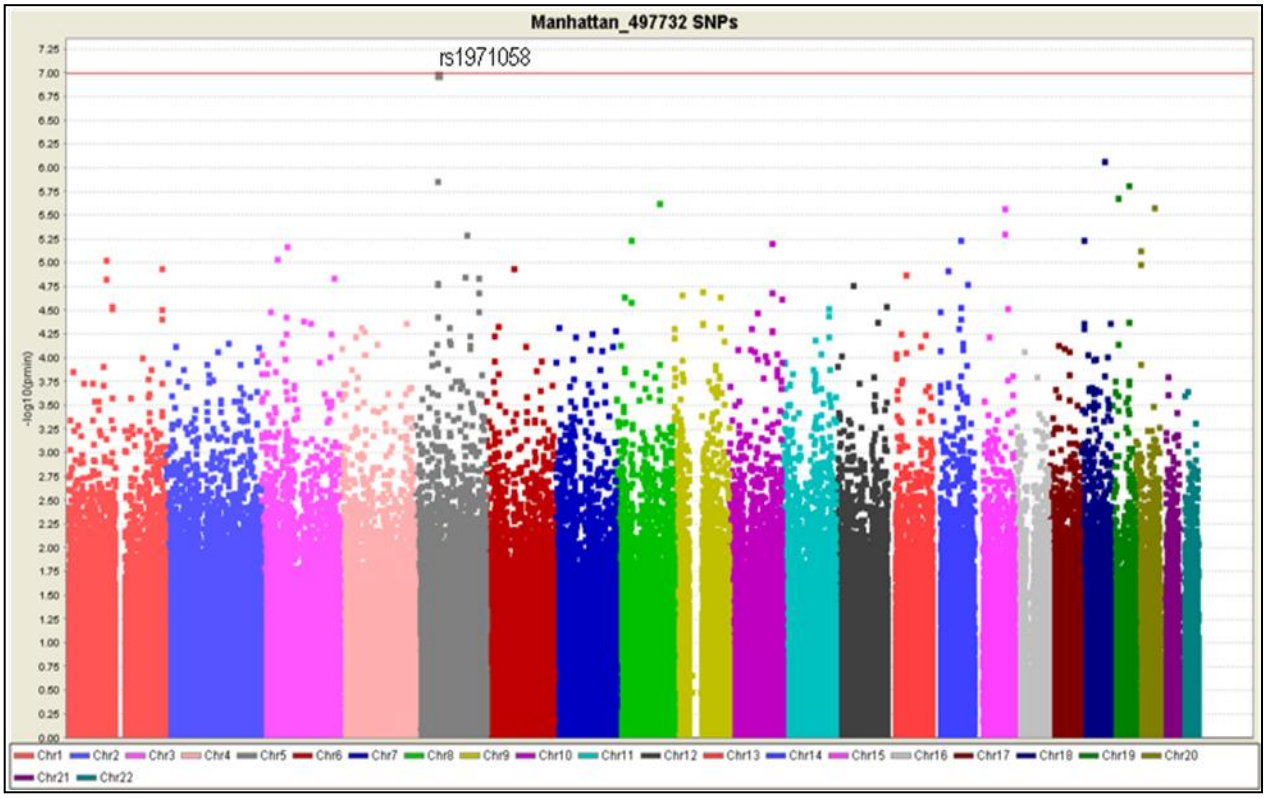
Данните за стойностите на P за 497 732 полиморфизма, изчислени по модела на алелните честоти (pa) с тест на Фишер, бяха използвани за построяване на нормална квантилна диаграма (Q – Q plot), изобразена на фигура 8. Както е видно от представената диаграма, при сравняване разпределенията на -log10 от наблюдаваните стойности на p (ординатa) и очакваните стойности на p (абсциса) при вярна нулева хипотеза, не беше установено значително отклонение от нормалното разпределение и наличие на популационна стратификация сред използваната извадка. Квантилната диаграма е построена с помощта на статистическа програма R, свободно достъпна в Internet.



Фигура 8: Нормална квантилна диаграма изобразяваща разпределението на отрицателните десетични логаритми на стойностите на P за 497 732 SNPs, успешно преминали качествения контрол. Y (observed (-log10P)) – отрицателен десетичен логаритъм от наблюдаваните стойности на P; X (expected (-log10P)) – отрицателен десетичен логаритъм от очакваните стойности на P при вярна нулева хипотеза (червената диагонална линия).

Данните от минималната получена стойност на P за всеки полиморфизъм (най–ниската стойност на P за всеки маркер, получена след изчисляване на стойностите на P

според трите модела с точен тест на Фишер) за 497 732 маркера преминали успешно качествения контрол, бяха използвани за построяване на диаграма тип Manhattan (Manhattan plot). По оста X на получената диаграма (Фигура 9) се вижда разпределението на 497 732 полиморфизма генотипирани в цялостното геномно асоциативно изследване, според локализацията им в 22-те двойки автозоми, като всяка хромозома е изобразена в различен цвят, а всяка точка е полиморфизъм. По оста Y са изобразени стойностите на отрицателния десетичен логаритъм от Pmin. След прилагане на корекцията на Бонферони за множествени тествания, коригираното ниво на статистическа значимост от $1X10^{-7}$ (0.05/497 732) беше преминато само от един маркер, rs1971058 (интергенен, между *LOC728103* и *LOC100130514*) с минимална стойност на $P=9.86X10^{-8}$.



Фигура 9: Диаграма тип Manhattan, отразяваща отрицателен десетичен логаритъм от разпределението на минималните P стойности за 497 732 SNPs в цялостния геномен скрининг. Всяка хромозома е изобразена в различен цвят. Червената линия изобразява коригираното ниво на статистическа значимост ($-\log10(1X10^{-7})=7$), базирано на корекцията на Бонферони за множествени тествания.

Преминалите успешно качествен контрол 497 732 маркера бяха подредени в нарастващ ред в зависимост от тяхната минимална стойност на P (след точен тест на Фишер), като първите 100 маркера показали най–ниска стойност бяха избрани за генотипиране във всички проби с цел потвърждаване и реплициране на резултатите от

цялостния геномен скан. За валидизиране на резултатите, 100те маркера бяха генотипирани в същата серия от случаи и контроли (188 случая и 376 контроли – група 1), използвана при цялостното геномно сканиране, а реплицирането беше проведено в допълнителна, независима серия от 122 случая и 328 контроли (група 2). И при двете проведени генотипирания, като генотипираща платформа беше използван Invader метод, комбиниран с мултиплексна полимеразна верижна реакция, с проби и праймери по собствен дизайн.

Статистическият анализ на данните от второто генотипиране (реплициране на резултатите) разкри статистически значима разлика с P стойност < 0.05 за следните осем полиморфизма: rs6859725 (*LOC728103*) с Pmin=0.05, rs10875032 (*PTBP2*) с Pmin=0.03, rs6673531 (*PTBP2*) с Pmin=0.04, rs11958003 (*LOC728103*) с Pmin=0.0047, rs1862019 с Pmin=0.012, rs875069 с Pmin= 0.032, rs4387777 (*USP34*) с Pmin=0.0035 и rs1472034 (*LOC728103*) с Pmin=0.037. След изключване на 16 маркера, показали пълна неравновесна скаченост с останалите (r^2>0.8), и прилагане на корекцията на Bonferroni за множествени тествания, нито един от тези осем маркера не достигна зададеното коригирано ниво на статистическа значимост в репликационното проучване, P<0.0006 (0.05/84).

Най – силна асоциация от цялостния геномен скрининг показа вариант rs1971058 (минимална P стойност 9.86 X 10^{-8}), локализиран в 5q11.2 хромозомен локус. Съгласно базата данни на NCBI (версия 36.3), в този район са картирани два гена, *LOC728103* (18 kb от SNP) и *LOC100130514* (35 kb от SNP), който е подобен на гена за фосфоглицерат мутаза 1 (*PGAM1*), експресиран в много тъкани, включително и в мозъка. Въпреки това, силната асоциация на този SNP не беше потвърдена при генотипирането му в допълнителната извадка от случаи и контроли (група 2) (P=0.18, MAF=0.50).

За установяване на крайна асоциация от цялостния геномен скан и репликационното проучване, върху данните от двете колекции беше приложен тестът на Cochran-Mantel-Haenszel. Въпреки че, повторното генотипиране на 100те полиморфизма не разкри силна статистически значима асоциация на някой от тези варианти със заболяването, комбинираният анализ на резултатите в 310 случая и 704 контроли, установи тенденция към асоциация за 10 варианта с P стойност <10 x 10^{-5}, след теста на Cochran-Mantel-Haenszel. Тези варианти, с изключение на един, който е

междугенен, са локализирани в некодиращите участъци на осем гена и са представени в таблица 7. Маркерите са подредени в зависимост от тяхната P стойност след теста на Mantel-Haenszel. За всеки маркер са посочени неговата позиция в базови двойки (според базата данни на NCBI), съответната хромозома, алелите му, най–близкия ген (до 100 бд), разпределението сред случаи и контроли, най–ниската стойност на P съгласно теста на Фишер след цялостния геномен скан и репликационното проучване, както и крайната асоциация (в P стойност) след прилагане теста на Cochran-Mantel-Haenszel.

Таблица 7: Генотипни характеристики и данни за 10 SNPs, показали тенденция за асоциация с биполарното разстройство в пълната извадка от случаи и контроли.

SNP ID (NCBI)	Х3	Позиция (бд)	Алели[a]	Функция	Близкостоящ ген	Група	Случаи[b]		Контроли[c]		P_{min} ЦГАП	P_{min} Реплициране	Mantel-Haenszel P
							11/12/22[d]	MAF	11/12/22[d]	MAF	OR (95%CI)	OR (95%CI)	OR (95%CI)
rs8099939	19	47212948	G>T	Интронен	*GRIK5*	ЦГАП	55/94/37	0.45	129/184/52	0.39	$1.46X10^{-6}$	0.097	$2.12X10^{-6}$
						РП	33/59/29	0.48	96/159/66	0.45	2.53 (1.65-3.88)	1.45 (0.90-2.34)	1.95 (1.43-2.67)
rs6122972	20	48808649	A>G	3' UTR	*PARD6B*	ЦГАП	158/27/1	0.08	245/106/14	0.18	$2.54X10^{-6}$	0.078	$3.11X10^{-6}$
						РП	92/25/2	0.12	225/86/13	0.17	2.76 (1.75-4.35)	1.50 (0.92-2.44)	2.02 (1.46-2.80)
rs2289700	15	77011738	G>A	Интронен	*CTSH*	ЦГАП	133/48/5	0.16	318/46/1	0.07	$2.60X10^{-6}$	0.171	$9.14X10^{-6}$
						РП	94/26/2	0.12	266/49/5	0.09	2.70 (1.74-4.19)	1.47 (0.88-2.45)	2.13 (1.53-2.95)
rs11129950	3	42357775	A>G	5' UTR	*CCK*	ЦГАП	159/25/2	0.08	248/109/8	0.17	$8.83X10^{-6}$	0.098	$2.18X10^{-5}$
						РП	97/25/0	0.10	235/77/8	0.15	2.78 (1.75-4.40)	1.40 (0.85-2.32)	1.99 (1.42-2.79)
rs16970287	15	77015312	A>G	Интронен	*CTSH*	ЦГАП	136/46/4	0.15	322/42/1	0.06	$4.77X10^{-6}$	0.257	$2.70X10^{-5}$
						РП	97/23/2	0.11	270/45/5	0.09	2.75 (1.75-4.33)	1.39 (0.82-2.36)	2.14 (1.52-3.00)
rs8109263	19	20632051	G>T	3' UTR	*ZNF626*	ЦГАП	168/16/2	0.05	265/94/6	0.15	$1.99X10^{-6}$	0.229	$4.08X10^{-5}$
						РП	102/16/3	0.09	253/59/8	0.12	3.52 (2.07-6.00)	1.42 (0.82-2.47)	2.35 (1.60-3.44)
rs10507408	13	32986583	A>G	Интергенен	-	ЦГАП	175/11/0	0.03	297/66/2	0.10	$5.39X10^{-5}$	0.102	$4.88X10^{-5}$
						РП	111/11/0	0.05	276/46/2	0.08	3.64 (1.89-7.00)	1.76 (0.89-3.46)	2.58 (1.61-4.15)
rs3123209	10	133079181	G>A	5' UTR	*TCERG1L*	ЦГАП	57/101/28	0.42	116/164/85	0.46	$2.29X10^{-5}$	0.101	$7.56X10^{-5}$
						РП	34/58/18	0.43	80/168/76	0.49	2.63 (1.67-4.15)	1.57 (0.89-2.75)	2.06 (1.45-2.94)
rs10491291	5	132944258	C>T	Интронен	*FSTL4*	ЦГАП	120/58/8	0.20	296/66/3	0.10	$4.82X10^{-6}$	0.471	$9.53X10^{-5}$

rs10875032	1	96970919	G>T	Интронен	*PTBP2*	РП	86/34/2	0.16	236/77/5	0.14	2.36 (1.59-3.51)	1.21 (0.76-1.91)	1.72 (1.28-2.31)
						ЦГАП	63/101/22	0.39	104/178/83	0.47	$8.99X10^{-6}$	0.030	$9.67X10^{-5}$
						РП	34/73/15	0.42	95/155/68	0.46	2.97 (1.81-4.88)	1.94 (1.07-3.52)	2.76 (1.88-4.05)

ХЗ, хромозома; бд, базови двойки; ЦГАП, Цялостно Геномно Асоциативно Проучване; РП, Репликационно проучване; MAF, Честота на по-рядко срещания алел; P_{min} ЦГАП, минимална стойност на P след тест на Фишер в цялостния геномен скан; P_{min} Реплициране, минимална стойност на P след тест на Фишер при реплициране на резултатите; Mantel-Haenszel P, стойност на P след прилагане на теста на Mantel-Haenszel върху данните от двете генотипирания (310 случая и 704 контроли); OR (95%CI), Съотношение на шансовете (95% доверителен интервал)

a: По-често срещан (Major, главен) алел > по-рядко срещан (Minor) алел

b: Цялостно геномно асоциативно проучване, *n*=188; Реплициране на резултатите, *n*=122

c: Цялостно геномно асоциативно проучване, *n*=376; Реплициране на резултатите, *n*=328

d: По-често срещаният алел сред случаите и контролите е дефиниран като алел 1

Статистическата обработка на данните от двете генотипирания (във всички налични проби) с тест на Cochran-Mantel-Haenszel, разкри аналогична стойност на P за следните пет полиморфизма в сравнение с цялостното геномно асоциативно проучване:

rs8099939 (PMH=2.12X10^{-6}) от гена *GRIK5*

rs6122972 (PMH=3.11X10^{-6}) от гена *PARD6B*

rs2289700 (PMH=9.14X10^{-6}) от гена *CTSH*

rs10507408 (PMH=4.88X10^{-5}) – интергенен

rs3123209 (PMH=7.56X10^{-5}) от гена *TCERG1L*

Освен rs2289700, за още един вариант от гена *CTSH* (cathepsin H), rs16970287 беше наблюдавана слаба асоциация със стойност на P от 2.70X10^{-5} (OR=2.14, 95%CI=1.52-3.00) след теста на Cochran-Mantel-Haenszel. С помощта на програмата Haploview (v 4.2) и въз основа на нашите генотипни данни за тези два полиморфизма установихме, че те се намират в неравновесна скаченост (r^2=0.9). Въпреки това, нито един маркер не показа стойност на P по – ниска от установената в цялостния геномен скан.

Използвайки базата данни за генетична асоциация на Националните институти по здравеопазване (NIH), проверихме дали някой от нашите предполагаеми гени, определени от топ 100 SNPs, не е бил вече включен в етиологията на биполарното разстройство или друго психиатрично заболяване при предишни асоциативни проучвания. Нашето търсене установи пет гена, изследвани за подобна асоциация. Това са генът за холецистокинин (*CCK*), картиран в късото рамо на хромозома 3p22-p21.3, гените за GABA-A рецепторите, alpha 1 и 6 (*GABRA1* и *GABRA6*), картирани в район 5q34-q35, генът за ацетилхолиновият рецептор, субединица alpha 9 (*CHRNA9*), локализиран в 4p14 района и генът за глутаматния рецептор 5 (*GRIK5*), картиран в 19q13.2 района. Общо 6 полиморфизма локализирани близо до или в тези гени бяха сред топ 100 вариантите след цялостното геномно асоциативно проучване. По литературни данни, само два от гореспоментатите 6 полиморфизма са били изследвани в две предишни независими проучвания. Подробна информация относно генотипирането на rs11129950 от гена CCK не беше открита в статията на Cratacos и колеги (2009) Вторият полиморфизъм, rs8099939 (*GRIK5*) е анализиран от Shibata и

колеги (2006) за асоциация с шизофрения в японска група от случаи и контроли, но позитивни резултати не са установени.

Допълнителна проверка в базата данни на NCBI (http://www.ncbi.nlm.nih.gov/) установи, че всички тези гени са разположени близо до или в райони, показали положителна асоциация с биполарното разстройство и шизофренията при предишни проучвания чрез анализ за скаченост. Генът *CCK*, чиито продукт, холецистокинин е повсеместно разпространен в ЦНС е широко изследван за редица невропсихиатрични състояния, като паническо разстройство, шизофрения, униполарна депресия, биполарно разстройство, въпреки че не всички проведени изследвания съобщават позитивни резултати. Интензивното изследване на този ген е продиктувано от многобройните функции, които той изпълнява, а именно регулиране на количеството храна, процеса на обучение, безпокойство/паника и други неврологични състояния. Най – близкият вариант до този ген, rs11129950 показа слаба асоциация след теста на Mantel-Haenszel със стойност на P от 2.18 x 10-5, OR=2.18. За възможна асоциация на гена *GABRA1* (кодиращ ГАМК рецептора, α1) с биполарното разстройство съобщават Horiuchi и колеги през 2004, изследвайки японска група от случаи/контроли и потвърждават своите резултати в друга независима извадка от европейски произход. В допълнение, една година по – късно, Petryshen и колеги получават положителни данни за асоциация на GABA A генния клъстер (включващ гените *GABRA1* и *GABRA6*) с шизофрения. Shi и колеги, 2007 изследват за асоциация с биполарното разстройство гена *CHRNA9*, като използват две отделни групи – 474 индивида от 152 семейства от Националния институт по психично здраве (National Institute of Mental Health, NIMH) и 83 индивида от 22 фамилии от Clinical Neurogenetics (CNG) pedigrees, но не получават позитивни резултати. Генът *GRIK5* е изследван за асоциация с шизофрения от Shibata и колеги въз основа на хипотезата за нарушена функция на глутаматергичната система в етиологията на заболяването. Въпреки че в Японска популация варианти асоциирани със заболяването не са идентифицирани, генът е един от обещаващите кандидат гени с възможна роля в етиологията на биполарното разстройство, тъй като глутаматните рецептори са важни за структурната и функционалната пластичност на синапса, включително процесите на обучение и запаметяване. Въпреки че след теста на Mantel-Haenszel гореспоменатите гени показаха слаба асоциация с биполарното разстройство в нашето проучване, тяхната локализация близо до, или в райони показали скаченост със

заболяването в предишни проучвания ги прави добри кандидати, с възможна роля в етиологията на биполарното разстройство. В допълнение, сравнихме резултатите от цялостното геномно асоциативно проучване със списък от полиморфизми и кандидат гени, за които в литературата съществуват положителни данни за асоциация с биполарното разстройство. Тъй като няколко скорошни цялостни геномни асоциативни проучвания и мета – анализи върху биполарно разстроство асоциират със заболяването полиморфизми и гени в 3p22-p21.3 хромозомен район, решихме да проследим генотипния статус на тези варианти в нашата популация. Както се вижда от таблица 8 нито един от тези варианти, локализирани в 3p22-p21.3 хромозомен район и застъпени в използваната от нас генотипираща платформа не показа асоциация със заболяването в нашата извадка.

Таблица 8: Резултати от цялостното геномно асоциативно проучване в българска популация за полиморфизми, асоциирани с биполарното разстройство в предишни проучвания.

SNP_ID (NCBI)	Х3	Позиция	Ген	Случаи, *n*=188		Контроли, *n*=376		ЦГАП P_{min}^{*} OR (95%CI)	Източник
				11/12/22	MAF	11/12/22	MAF		
rs1042779	3	52796051	*ITIH1*	62/96/28	0.41	115/177/73	0.44	0.164 1.41 (0.88-2.27)	Scott et al., 2009
rs2289247	3	52702297	*GNL3*	38/93/55	0.45	87/179/99	0.48	0.372 1.22 (0.79-1.87)	Scott et al., 2009
rs11177	3	52696345	*GNL3*	31/95/60	0.42	81/182/102	0.47	0.125 1.43 (0.90-2.25)	Scott et al., 2009
rs3617	3	52808845	*ITIH3*	52/92/42	0.47	103/181/80	0.47	0.898 1.04 (0.68-1.58)	Scott et al., 2009
rs2251219	3	52559827	*PBRM1*	56/96/33	0.44	102/178/82	0.47	0.223 1.35 (0.86-2.11)	McMahon et al., 2010
rs1010554	3	52517959	*STAB1*	63/89/34	0.42	132/174/59	0.40	0.517 1.16 (0.73-1.84)	Baum et al., 2008
rs7620081	3	51475631	*VPRBP*	160/26/0	0.07	318/44/3	0.07	0.328 inf	Baum et al., 2008
rs9833750	3	51171463	*DOCK3*	159/27/0	0.07	319/43/3	0.07	0.328 inf	Baum et al., 2008
rs9869826	3	50973856	*DOCK3*	159/27/0	0.07	320/42/3	0.07	0.328 inf	Baum et al., 2008
rs6265	11	27636492	*BDNF*	2/54/130	0.156	21/123/221	0.226	0.007 5.62 (1.44-21.79)	Muller DJ et al., 2006; Okada T et al., 2006; Kremeyer B et al., 2006; Sklar P et al., 2002; Geller B et al., 2004;
rs1012053	13	41551437	*DGKH*	128/45/13	0.191	247/106/11	0.176	0.045 2.38 (1.07-5.33)	Baum et al., 2008
rs9315897	13	41636982	*DGKH*	134/44/8	0.16	227/123/15	0.21	0.023 1.57 (1.07-2.30)	Baum et al., 2008
rs420259	16	23541527	*PALB2*	104/60/22	0.28	176/159/30	0.30	0.105 1.36 (0.96-1.94)	WTCCC, 2007
rs4939921	18	45716326	*MYO5B*	157/26/3	0.086	292/68/4	0.10	0.246 1.34 (0.83-2.14)	Sklar, 2008

ЦГАП, цялостно геномно асоциативно проучване; CI, доверителен интервал,(confidence interval); Х3,хромозома; MAF, честота на редкия алел, (Minor Allele Frequency);

OR, съотношение на шансовете, (odds ratio; SNP, единична нуклеотидна замяна, (single nucleotide polymorphism); inf, безкрайност, (infinity)
* Двустранен точен тест на Фишер. Представени са минималните стойности на P

В допълнение, малко вероятно е вариантът rs11129950 (*CCK*), който ние идентифицирахме в нашето проучване да се намира в неравновесна скачесност (LD) с някой от по–рано идентифицираните полиморфизми в хромозомен район 3p22-p21.3, тъй като най–близкия до нашия маркер SNP, rs9869826 от гена *DOCK3* (експресиран изключително в ЦНС, а неговият продукт е известен като модификатор на клетъчната адхезия или пресинилин – свързващ протеин) е разположен на 8.68Mbp разстояние от нашия маркер (съгласно базите данни на NCBI и HapMap). Тенденция към асоциация беше установена за гена *DGKH* (diacylglycerol kinase, eta). Генът е картиран в дългото рамо на хромозома 13q14, в район показал скаченост с биполарното разстройство и шизофренията, а неговият продукт играе важна роля в литий – чувствителния фосфатидил инозитолов сигнален път. Най–силно асоциирания вариант от този ген, rs1012053 идентифициран през 2008 г. от Baum и колеги в тяхното проучване, показа слаба асоциация в нашата извадка с минимална стойност на P от 4.50×10^{-2} след цялостното геномно асоциативно проучване, но въпреки това предразполагащият алел, установен от Baum et al е A, а в нашето проучване – G. Според базата данни на HapMap за европейска популация, rs1012053 не е в LD ($r^2 = 0.03$) с друг вариант от същия ген, rs9315897, с минимално $P=2.30 \times 10^{-2}$ (в нашата извадка), също идентифициран от Baum и колеги. Полиморфизъм rs420259, локализиран в гена *PALB2* (partner and localizer of *BRCA2* gene) идентифициран от WTCCC като асоцииран с болестта, не показа никаква асоциация в нашата извадка. Полиморфизъм rs1485171, локализиран в гена за глутамания рецептор, подтип 7, също споменат от WTCCC в тяхното изследване не беше включен в използваната от нас генотипираща платформа. Вариантът rs4939921 в гена за миозин (*MYO5B*) асоцииран със заболяването в изследването на Sklar и колеги, през 2008 не показа никаква асоциация в изследваната от нас група. Не на последно място, асоциацията с гена *BDNF* също би трябвало да бъде спомената. Полиморфизмът rs6265 от този ген, за когото данни за асоциация със заболяването съществуват от предишни проучвания в други популации, показа слаба асоциация в нашата извадка след цялост-ното геномно асоциативно изследване (минимална стойност на $P = 6.92 \times 10^{-3}$). Въпреки това, rs6265 загуби значимостта си след прилагане на корекцията за множествени тествания.

ОБСЪЖДАНЕ

Обсъждане на резултатите от асоциативното проучване на кандидат гени

Асоциативното проучване на кандидат гени и полиморфизми в тях беше проведено в първия етап от нашето проучване, за да бъдат идентифицирани генетични варианти, асоциирани с биполарното афективно разстройство в българска популация. За целта, подробно бяха проучени няколко стотин източници, съобщаващи за молекулярно генетични изследвания при заболяването. При избора на кандидат гени, вниманието ни основно беше насочено към тези от тях, за които положителна асоциация беше установена в минимум едно проучване; гени, неизследвани до момента за асоциация със заболяването в българска популация; гени, неизследвани за асоциация със заболяването, но за които съществуват данни за нарушена експресия в мозъка при пациенти с биполарно разстройство. С помощта на базите данни на NCBI и HapMap и въз основа на съществуващи данни за асоциация със заболяването, за генотипиране в първият етап на нашето проучване бяха избрани общо 199 маркера, от 65 гена, повечето от които 'представителни' SNPs (tagSNPs) във фаза II на HapMap проекта с r^2>0.8 и честота на редкия алел 0.1 (MAF≥0.1) в европейска популация (CEU population), с изключение на тези, асоциирани със заболяването в други популации. За генотипиране на избраните полиморфизми в този първи етап от проучването бяха използвани 2 метода. 47 полиморфизма от общо 199те избрани, бяха успешно генотипирани чрез метода TaqMan с проби на Applied Biosystems, а за останалите 152 маркера беше използвана Invader технологията, комбинирана с мултиплексна полимеразна верижна реакция, с проби и праймери по собствен дизайн. Методът TaqMan, при който алелната дискриминация се постига посредством хибридизация предлага сравнително бърз и лесен начин за установяване на мутационния статус на ДНК в една стъпка. Тъй като при разграничаването на алелите тук не се включват ензими, ключова роля има дизайнът на пробата. Всички проби, използвани за генотипиране на SNPs чрез метода TaqMan, са предоставени от компанията Applied Biosystems, която понастоящем предлага една от най-големите в света колекции на валидизирани, предварително конструирани (pre-designed), налични проби за генотипиране на SNPs, които използват универсални PCR условия и не изискват допълнителна оптимизация на протокола. Въпреки че се отличава с изключителна специфичност, чувствителност, бързина и точност, TaqMan изисква алел – специфични, флуоресцентно белязани олигонуклеотиди (проби) за всеки анализиран маркер, което увеличава цената за един анализ, изисква сравнително голямо количество геномна ДНК

(при анализиране на голям брой полиморфизми) и не на последно място, позволява едновременното анализиране на един полиморфизъм в една реакция, което значително намалява производителността на метода. Поради тези причини, като алтернативен метод за останалите 152 маркера беше избрана Invader технологията, кобминирана с мултиплексна полимеразна верижна реакция, а пробите и праймерите бяха конструирани по собствен дизайн.

Независимо от избрания метод за генотипиране, всички полиморфизми първоначално бяха анализирани в група от 94 случая и 184 контроли. Генотипирането се оказа успешно за 179 маркера от 65 гена, от които 17 показаха предполагаема асоциация със заболяването с P стойност по–малка от 0.05. За потвърждаване истинността на получените резултати, сигнификантните маркери бяха генотипирни в допълнителна, независима извадка от 78 случая и 372 контролни индивида. Статистическият анализ на данните само от второто генотипиране, показа възможна асоциация на заболяването с три полиморфизма, а комбинирания анализ на резултатите от двете генотипирания разкри предполагаема асоциация на биполарното разстройство с общо девет маркера. След коригиране на стойността на P посредством корекцията на Бонферони за множествени сравнения, само един полиморфизъм, rs1800883 от гена HTR5A остана статистически значимо асоцииран с биполарното разстройство (P=0.000097, коригирана стойност P=0.017).

Според базите данни на NCBI и HapMap (http://www.ncbi.nlm.nih.gov/SNP/; www.hapmap.org/index.html.en) rs1800883 (-19G/C) е локализиран в промоторния участък на гена кодиращ серотониновия рецептор подтип 5A (*HTR5A*, 5-hydroxytryptamine (serotonin) receptor 5A). По литературни данни към момента, полиморфизмът е изследван за асоциация с биполарното разстройство в няколко проучвания. Анализирайки 112 биполарни пациенти, 75 униполарни пациенти и 187 контролни индивида от бри-тански произход, Birkett и колеги, 2000 установяват асоциация между G алела на SNP и биполарното разстройство с P стойност 0.025 и OR (Odds Ratio) 0.56, индикиращо евентуалният протективен ефект на този алел в изследваната популация. Година по – късно, Arias и колеги анализират същия този вариант при 88 биполарни пациента и 164 контроли от испански произход, но не откриват асоциация между SNP и заболяването. Въпреки че в настоящата работа е установена статистически значима асоциация между rs1800883 и биполарното

разстройство, G алелът на този SNP, определен от Birkett и колеги (2000) като протективен, се оказва рисков за нашата популация. Когато изследваният вариант е асоцииран директно със заболяването, един и същи алел на маркера би трябвало да покаже асоциация при проучвания в други популации. При положение, че изследваният полиморфизъм се намира в неравновесна скаченост с варианта причинител, то тогава може да се очаква различните алели на този маркер да покажат асоциация със заболяването при различни популации дори когато каузалния вариант е представен. В допълнение, в нашата контролна група индивиди, честотата на G алела (0.62) е сходна с тази при германски (0.63) и испански контроли (0.62), докато при британски контроли тя е доста по–висока (0.71). Имайки предвид произхода на изследваните проби и факта, че полиморфизмът е изследван само в европейска популация, разлики по отношение на алелната честота между британските контроли и българските, испанските и германските контроли не би трябвало да се очакват и е много вероятно да се дължат на наличие на популационна стратификация в британската контролна извадка.

Получените от нас резултати осигуряват допълнителни доказателства за значението на гена *HTR5A* за предразположеността към биполарно афективно разстройство, но не могат да обяснят връзката между идентифицирания SNP и болестта. Добре известно е, че двете фази на заболяването (мания и депресия) се отличават с различен серотонинов профил (излишък и недостиг на серотонин, респективно), като тази трансформация е силно повлияна от фактори на околната среда. Тъй като районът около SNP е богат на CpG повтори, той би могъл да бъде добър кандидат за епигенетичнен контрол чрез метилиране, въпреки че до момента по литературни данни, функционални проучвания, касаещи този полиморфизъм не са провеждани.

Човешкият *5-HT5A* рецепторен ген се състои от два екзона, разделени от дълъг интронен участък. Целият кодиращ район на гена е подробно проучен при здрави индивиди за наличие на генетични варианти, като първоначално са установени 2 полиморфизма в 5’ нетраслируемия участък на гена (единия от които rs1800883) и други два в кодиращия регион. Генът е експресиран основно в мозъчната кора (по–специално слоеве II-III и V-VI на неокортекса), хипокампуса (gyrus dentatus и клетъчно пирамидалния слой на зони CA1 и CA3) и малкия мозък, което разпространение най–вероятно е индикатор за това, че продукта на този ген е включен във висши мозъчни и

лимбични функции. Поради тази причина, *HTR5A* е смятан за един от добрите кандидат гени с евентуална роля в етиологията на психиатричните заболявания. Той кодира протеин от 357 аминокиселини и е картиран в дългото рамо на хромозома 7q36.1, в район показал предполагаема скаченост с биполарното афективно разстройство (с ранно начало) в няколко проучвания.

Серотонергичната система е широко застъпена в ЦНС повлиявайки почти всяка сфера от физиологията на бозайниците: от сърдечносъдовата регулация, дишането, храносмилателната система, усещането за болка и терморегулация до по–централно контролирани функции, които включват, поддържане на циркадни ритми, апетит, агресия, сензоримоторна активност, сексуално поведение, настроение, познавателни способности, учение и памет. Предположението, че нарушения в серотонергичната система са свързани с етиологията на депресивните разстройства, датира от 1955 г., когато Pletscher и колеги установяват, че резерпинът (антихипертензивно лекарство, което подтиска депресията) намалява серотонина в мозъка. През 1958 г. установяването на антидепресивните свойства на имипрамина (imipramine) и ипрониазида (iproniazid), впоследствие свързано с действието им съответно като инхибитор на обратното захващане на серотонина и инхибитор на моноаминооксидазата, допълнително потвърждава ролята на серотонина в патофизиологията на депресивните разстройства. Тези открития допринасят за формулирането на т.нар. индоламинна хипотеза за депресията. Според нея склонността към депресия е свързана с ниска серотонергична активност, която може да се обясни с: 1.намалено освобождаване на серотонин; 2. по – малко серотонинови рецептори или 3. нарушена серотонин рецепторно – медиирана сигнална трансдукция. През 1974 г. Prange и колеги разширяват индоламинната хипотеза, като включват и биполарното афективно разстройство, изказвайки предположението, че двете фази на заболяването – манийна и депресивна се характеризират с дефицит в серотонергичната невротрансмисия. По отношение на афективните разстройства, серотониновата система привлича голям интерес поради факта, че антидепресантите, които инхибират обратното захващане на серотонина или активират неговите рецептори, са ефикасни при лечението на болестта.

Към днешна дата, почти всички гени кодиращи различните подтипове серотонинови рецептори са анализирани в голям брой асоциативни проучвания. В настоящата научна разработка освен *HTR5A* гена, са изследвани полиморфизми в още

три гена, кодиращи серотонинови рецептори подтип 2A (*HTR2A*), 3B (*HTR3B*) и 4 (*HTR4*).

От молекулярно генетична гледна точка, сред многобройните ‘серотонергични кандидат гени’, серотониновия 2A рецепторен ген, локализиран в дългото рамо на хромозома 13q14-q21, е най–широко изследвания, вероятно поради факта, че подпомага благоприятния терапефтичен ефект на инхибиторите на обратното захващане на серотонина. Към момента по литературни данни, пет *HTR2A* полиморфизма са тествани в асоциативни проучвания, но положителни резултати са получени само за два от тях - rs1058573 (-1438A/G;), локализиран в промоторния участък на гена и rs6313 (102T/C; Ser34Ser), в екзон 1. И двата полиморфизма бяха включени в асоциативното проучване на кандидат гени и бяха изследвани за асоциация със заболяването в българската извадка. Единият, rs1058573 се оказа мономорфен за нашата популация, поради което беше изключен от анализ, а rs6313 не показа статистически значима асоциация със заболяването (Pmin=0.071).

Два полиморфизма от гена *HTR3B*, водещи до аминокиселинна замяна - rs1176744 (A/C Tyr129Ser) в екзон 5 и rs17116138 (G/A Val183Ile) в екзон 6 бяха подложени на анализ за асоциация с биполарното разстройство в нашата извадка, но нито един от тях не показа статистически значима стойност на P (Pmin= 0.50 и 1.00, респективно).

Генът *HTR4* също беше застъпен с два полиморфизма в асоциативното проучване: rs2278392 и rs1432919. Поради наличие на силна неравновесна скаченост между тях, на анализ за установяване на асоциация с биполарното афективно разстройство в българската извадка беше подложен само един от двата маркера - rs2278392. Статистически значима разлика в разпределението на алелните и генотипни честоти между случаи и контроли не беше установена (Pmin= 0.30).

Въпреки че след прилагане на стриктната корекция на Бонферони, само вариант rs1800883 от гена *HTR5A* остана статистически значимо асоцииран със заболяването, вероятната асоциация на още няколко полиморфизма не може да бъде изцяло изключена и те би трябвало да бъдат споменати.

Разлика в разпределението на алелните и генотипни честоти между случаи и контроли за вариант rs6265 (*BDNF*) беше установена и в българската извадка.

Рисковият за нашата популация алел C, беше наблюдаван по–често при случаи в сравнение с контролната група индивиди (P=0.0033). Заедно с rs6265, още два полиморфизма от гена - rs2049045 и rs4923463 бяха включени в асоциативното проучване на кандидат гени, но нито един от двата варианта не запази статистическата си значимост след обработката на резултатите от двете генотипирания.

Общо четири полиморфизма от гена *DPYSL2* (DRP-2) бяха включени в проведеното от нас асоциативно проучване на кандидат гени и бяха анализирани в първия етап на изследването в 94 случая и 184 контроли. Два от четирите полиморфизма, rs11863 и rs17666 показаха стойност на P по – ниска от 0.05 (Pmin=0.0028 и 0.0019, респективно) и бяха генотипирани във втория етап на проучването в нова извадка от случаи и контроли. След комбинирането на резултатите от първия и втория скрининг, само rs17666 показа асоциация със заболяването, като рисковият за нашата популация алел T беше установен в 70.3% от случаите срещу 62% при контролните индивиди (P=0.0045).

Въз основа на положителните резултати получени от предходни изследвания, генът *GNB3* (guanine nucleotide binding protein (G protein), beta polypeptide 3) беше определен като подходящ кандидат с възможна роля в етиологията на биполарното разстройство и беше включен в асоциативното проучване на кандидат гени. Избраният полиморфизъм, rs5443 беше генотипиран успешно първоначално в група от 94 случая и 184 контролни индивида чрез метода TaqMan. Въпреки че след статистическата обработка на данните от първия скрининг полиморфизмът показа гранична стойост на P=0.048 (при зададено ниво на значимост P=0.05), той беше проследен във втори скрининг в допълнителна, независима извадка от 78 случая и 372 контроли. Рисковият алел T беше наблюдаван по–често в групата на случаите (39.7%) в сравнение с контролната група индивиди (29.7%) със стойност на P (съгласно модела на алелните честоти) от 0.017. Полиморфизмът запази статистически значима асоциация при същия рисков алел (T) и след комбиниране на данните от двата скрининга (172 случая и 556 контроли), като рисковият генотип (TT) беше установен при 18% от случаите в сравнение с 8% при контролните индивиди (P=0.00051). Въпреки установената асоциация, след прилагане на корекцията на Бонферони за множествени тествания, rs5443 загуби статистическата си значимост.

В заключение, получените от нас резултати след асоциативното проучване на кандидат гени дават основание да се предположи, че генът *HTR5A* има важна, но най–вероятно не самостоятелна роля в етиологията на биполарното разстройство в българската популация. Въпреки че, асоциацията на гени като *BDNF, HTR3B, 5HTR2A, HTR4, TPH2* не беше потвърдена, известна тенденция бе наблюдавана за някои от тях, поради което тяхната евентуална роля в патогенезата на болестта не може да бъде отхвърлена напълно. Най–вероятно, малкият размер на използваната извадка или наличието на фалшиво позитивни резултати в предходни проучвания са причина за липсата на асоциация при тях. Поради тази причина, допълнителни проучвания, използващи по–голям брой проби както в нашата, така и в други популации заслужават внимание.

Резултатите от проведеното изследване биха могли да бъдат от полза при планиране на бъдещи анализи на кандидат гени, а също така биха били добър източник на информация при провеждане на мета-анализи върху отделни кандидат гени и варианти в тях.

Обсъждане на резултатите от цялостното геномно асоциативно проучване

С цел идентифициране на нови гени с евентуална роля в етиологията на биполарното разстройство в българска популация, втория етап на настоящата научна разработка е посветен на цялостното геномно асоциативно проучване.

С помощта на HumanHap550K генотипиращият чип на Illumina в съчетание с технологията Infinium II, общо 554 496 представителни единични нуклеотидни полиморфизма бяха успешно анализирани в група от 188 случая с биполарно разстройство и 376 контролни индивида. След прилагане на определени критерии за контрол на качеството, общо 497 732 полиморфизма бяха подбрани за по–нататъшен анализ. За валидизиране на резултатите от цялостното геномно сканиране, първите 100 полиморфизма, показали най–значима асоциация със заболяването бяха генотипирани чрез Inavder метод, комбиниран с мултиплексна полимеразна верижна реакция в същата извадка от случаи и контроли. Във втория етап на цялостното геномно асоциативно проучване, с цел потвърждаане на резултатите, стоте маркера бяха анализирани в допълнителна серия от 122 случая и 328 контроли. И при двете генотипирания, като генотипираща платформа отново беше използван методът Invader,

комбиниран с мултиплексна полимеразна верижна реакция, с проби и праймери по собствен дизайн.

Въпреки че, проведеното от нас цялостно геномно асоциативно проучване не разкри убедителен кандидат ген и нито един от топ100 полиморфизмите не достигна коригираното ниво на значимост след прилагане корекцията на Бонферони ($P < 0.00060$), има няколко варианта, чиято възможна роля в етиологията на заболяването не може да бъде напълно изключена и те заслужават внимание.

Както вече беше споменато, генът кодиращ глутаматния рецептор, подтип 5 (*GRIK5*) първоначално е изследван за асоциация с шизофрения, но генетични варианти асоциирани със заболяването не са установени за японска популация. Генът е картиран в дългото рамо на хромозома 19, включва 19 екзона и има дължина 67.48 kb. Широко експресиран е в имунната, нервната, репродуктивната и дихателната системи. Генът кодира протеин, който принадлежи към фамилията на глутамат – зависимите йонни канали. Глутаматът действа като основен екситаторен невротрансмитер в централна нервна система посредством активирането на лиганд – зависими йонни канали и G – протеин свързани мембранни рецептори. Заедно със субединици, кодирани от други гени във фамилията, продуктът на този ген формира функционални хетеромерни каинат – свързващи йонни канали. В допълнение, *GRIK5* е един от 306те кандидат гена, анализирани от Gratacos и колеги (2009) в техния опит да идентифицират нови гени за няколко психиатрични заболявания, сред които и биполарното разстройство в Европейска популация. Силна асоциация с биполарно разстройство е наблюдавана за друг полиморфизъм, rs10414815, локализиран приблизително на 66kbp от нашия маркер (база данни на HapMap). Този вариант не беше включен в нашата генотипираща платформа и не е в LD ($r^2=0.09$) с rs8099939, според базата данни на HapMap.

Вторият вариант, rs6122972, показал тенденция към асоциация със заболяването в нашата извадка със стойност на P от $3.11X10^{-6}$ и съотношение на шансовете, OR=2.02, след теста на Mantel-Haenszel е локализиран близо до гена *PARD6B* (par-6 partitioning defective 6 homolog beta), който е картиран в дългото рамо на хромозома 20, в регион показал скаченост с шизофрения в предходни проучвания. Генът съдържа 3 екзона, има дължина 22.20 kb, а неговият продукт участва в процесите на асиметрично клетъчно

делене и клетъчна поляризация и има предполагаема роля в аксоногенезата и регулацията на клетъчната миграция.

Интронен вариант rs2289700 в гена за катепсин (*CTSH*) също показа тенденция за асоциация със заболяването след комбинирането на резултатите от цялостния геномен скан и репликативното проучване с P стойност от $9.14X10^{-6}$ и съотношение на шансовете, OR=2.13, след теста на Mantel-Haenszel. Генът е картиран върху дългото рамо на хромозома 15 в район показал предполагаема скаченост, както с биполарното заболяване така и с други психиатрични заболявания, съдържа 14 екзона и има дължина 23.33 kb. Протеинът кодиран от този ген е лизозомна цистеин протеиназа, която участва в разграждането на лизозомните протеини. Увеличена експресия на гена се свързва с малигнената прогресия на простатни тумори. До момента, по литературни данни генът не е изследван за асоциация с биполарното заболяване.

Полиморфизмът rs6265 ($P=6.92X10^{-3}$) от гена *BDNF* (мозъчен невротрофен фактор) също би трябвало да бъде коментиран, въпреки че наблюдаваната след цялостния геномен скан асоциация не запази значимост след прилагане на корекцията на Бонферони. Генът е картиран върху късото рамо на хромозома 11p13, в район показал предполагаема скаченост с биполарното заболяване в предходни проучвания. Полиморфизмът rs6265 (G/A, Val66Met) е бил обект на множество асоциативни проучвания през годините, като по – голяма част от тях установяват асоциация с биполарното заболяване за Val алела на SNP (същата асоциация беше установена и при асоциативното проучване на кандидат гени, проведено в първия етап на нашето изследване). Въпреки че, след цялостния геномен скан, Val алела на полиморфизма беше наблюдаван отново по–често в групата на случаите в сравнение с групата на контролните индивиди със стойност на P от 0.007, rs6265 загуби статистическата си значимост след прилагане на корекцията за множествени тествания. Въпреки този резултат, имайки предвид функцията и локализацията на гена *BDNF*, както и големия брой проучвания проведени през годините, по–голяма част от които съобщават позитивни данни за асоциация със заболяването, генът за мозъчния невротрофен фактор остава добър кандидат ген с възможна роля в етиологията на биполарното разстройство.

В резултат на проведените до момента цялостни геномни асоциативни проучвания и последващи мета – анализи на редица често срещани заболявания, днес са идентифицирани и потвърдени голям брой чести полиморфизми, имащи отношение към етиологията на тези състояния, като едновременно с това става ясно, че всеки един от идентифицираните варианти има малък принос за повишаване риска от дадено заболяване (средно OR=1.36) [336]. Тези изследвания също така ясно показват, че групи съдържащи няколко хиляди пациента трябва да бъдат анализирани с цел идентифициране на полиморфизми с малък ефект (какъвто най–вероятно е случаят и с психиатричните заболявания). В този контекст, основен недостатък на проведеното от нас цялостно геномно сканиране е малкият размер на използваната извадка (188 случая и 376 контроли), което налага проследяването на получените от нас резултати в други проучвания.

Проведените до момента асоциативни проучвания на кандидат гени и цялостни геномни сканове върху биполарно разстройство идентифицират голям брой гени и полиморфизми в тях като асоциирани със заболяването. Въпреки че нашето цялостно геномно асоциативно проучване не установи силен кандидат ген и не успяхме убедително да потвърдим ролята в етиологията на заболяването за Българска популация на полиморфизми, асоциирани с биполарното разстройство в предишни цялостни геномни проучвания, слаба асоциация беше наблюдавана за гена *DGKH*. Трудно е да бъдат направени финални заключения за други известни кандидат гени, като *PALB2, GRM7* и *MYO5B*, тъй като анализираните от тях полиморфизми не показаха асоциация с биполарното заболяване в нашата извадка, което не ги отхвърля като възможни кандидати.

ИЗВОДИ

1. Установено е, че 65 гена от известните голям брой кандидат гени, включени в базата данни на NCBI и HapMap проекта са подходящи за асоциативно проучване на биполарната психоза в пациенти от българска популация чрез методите TaqMan и Invader.
2. В първия етап на изследването успешно са определени алелните и генотипни честоти на 179 полиморфизма от 65 гена и е доказано, че 17 от тях са асоциирани със

заболяването (P < 0.05). Три от тях останаха позитивни при репликативното генотипиране.

3. Установена е статистически значима асоциация на девет SNPs с биполарното разстройство при комбинирания анализ на резултатите от двата етапа на асоциативното проучване (172 случая и 556 контроли).

4. Вариантът rs1800883 в гена за серотониновия рецептор (*HTR5A*) показа най–силна статистически значима асоциация (P =0.000097) и остана асоцииран със заболяването и след прилагане на корекцията на Бонферони за множествени тествания (P = 0.017).

5. В резултат на финото картиране в рамките на LD-региона на rs1800883 (*HTR5A*) са определени алелните и генотипните честоти на 12 представителни полиморфизма (tSNPs) и е установена статистически значима асоциация между 2 от тях и биполарното разстройство. Нито един от тези варианти не достигна значимостта на rs1800883.

6. Доказано е с хаплотипен анализ, че трите най–значими маркера в района на *HTR5A* генът са положително асоциирани със заболяването (P=0.00045), но асоциацията не е по–силна от тази на индивидуалния маркер rs1800883.

7. След първия етап на цялостния геномен скан най–значима асоциация с биполарното разстройство с P стойност от $9.86X10^{-8}$ е установена за rs1971058 (*LOC728103*/*LOC100130514*) от общо успешно генотипирани 594 496 SNPs с 550K чип на Illumina. Тази асоциация не е запазена при потвърждаването на резултатите в допълнителни проби

8. Повторното генотипиране на топ 100 SNPs, в отделна извадка от случаи/контроли не потвърждава силната асоциация на нито един от тези варианти с биполарното разстройство.

9. При комбинираният анализ на резултатите от двете извадки, десет маркера (rs8099939, rs6122972, rs2289700, rs11129950, rs16970287, rs8109263, rs10507408, rs3123209, rs10491291, rs10875032) показаха тенденция за асоциация със заболяването ($P < 10 \times 10^{-5}$).

ПРИНОСИ

А. Научно – приложни приноси

1. За първи път сред група от български пациенти с биполарно афективно разстройство и незасегнати контроли е проведено асоциативно проучване върху голям брой полиморфизми от кандидат гени, при които е установена асоциация със заболяването в други популации.

2. Данните от проведеното асоциативно проучване на кандидат гени послужиха за провеждането на мета-анализ върху варианти в гена *MTHFR*.

3. Разширена е ДНК банката на Катедрата по медицинска генетика, Медицински университет София, с 90 проби от пациенти с биполарно заболяване, както и със съпътстваща клинична информация.

4. За първи път в българска популация е проведено цялостно геномно асоциативно проучване върху биполарното разстройство в резултат, на което е създадена база данни с алелните и генотипните честоти на повече от 500,000 SNPs на индивиди с български произход.

5. Идентифицирани са и нови хромозомни региони и гени, които могат да съдържат варианти с възможна роля в етиологията на заболяването.

Б. Приноси с потвърдителен характер

6. Получени са допълнителни доказателства за значението на варианти в серотониновите гени за развитието на биполарното разстройство, подкрепящи серотониновата хипотеза за етиологията на болестта.

7. Получени са допълнителни данни в подкрепа на асоциацията с биполарно разстройство в българска популация за гени, асоциирани със заболяването в предходни проучвания.

ПУБЛИКАЦИИ ВЪВ ВРЪЗКА С ДИСЕРТАЦИОННИЯ ТРУД

Yosifova A, Mushiroda T, Stoianov D, Vazharova R, Dimova I, Karachanak S, Zaharieva I, Milanova V, Madjirova N, Gerdjikov I, Tolev T, Velkova S, Kirov G, Owen MJ, O'Donovan MC, Toncheva D, Nakamura Y: Case-control association study of 65 candidate genes revealed a possible association of a SNP of HTR5A to be a factor susceptible to bipolar disease in Bulgarian population. J Affect Disord. 2009 Sep;117(1-2):87-97 (IF, 3.76)

Adelina Yosifova, Taisei Mushiroda, Michiaki Kubo, Atsushi Takahashi, Yoichiro Kamatani, Naoyuki Kamatani, Drozdstoi Stoianov, Radoslava Vazharova, Sena Karachanak, Irina Zaharieva, Ivanka Dimova, Savina Hadjidekova, Vihra Milanova, Nadejda Madjirova, Ivan Gerdjikov, Todor Tolev, Nadejda Poryazova, George Kirov, Michael J Owen, Michael C O'Donovan, Draga Toncheva, Yusuke Nakamura. Genome wide association study on bipolar disorder in the Bulgarian population. Genes, Brain and Behavior, [Epub ahead of print] (IF, 3.79)

Peerbooms OL, van Os J, Drukker M, Kenis G, Hoogveld L; **MTHFR in Psychiatry Group**, de Hert M, Delespaul P, van Winkel R, Rutten BP. Meta-analysis of MTHFR gene variants in schizophrenia, bipolar disorder and unipolar depressive disorder: Evidence for a common genetic vulnerability? Brain Behav Immun. 2010. [Epub ahead of print]

MTHFR in Psychiatry Group:…… **Adelina Yosifova**

А. Йосифова, Е. Бечева. Генетика на афективните разстройства в Медицинска генетика в постгеномната ера. Геномна медицина. Под редакцията на Проф. Драга Тончева; София, 2010; 608-611

Е. Бечева, **А.Йосифова**, Д. Тончева. Подходи за идентифициране на патологични генетични варианти при мултифакторните заболявания в Медицинска генетика в постгеномната ера. Геномна медицина. Под редакцията на Проф. Драга Тончева; София, 2010; 565-570

ЦИТИРАНИЯ НА ПУБЛИКАЦИИТЕ ВЪВ ВРЪЗКА С ДИСЕРТАЦИОННИЯ ТРУД

1. Z Zeng, T Wang, T Li, Y Li, P Chen, Q Zhao, J Liu, J Li, G Feng, L He and Y Shi. Common SNPs and haplotypes in DGKH are associated with bipolar disorder and schizophrenia in the Chinese Han population. Molecular Psychiatry 16, 473-475 (May 2011)

2. Piletz, John E.; Zhang, Xiaotongb; Ranade, Rajdeepa; Liu, Chunyub. Database of genetic studies of bipolar disorder. Psychiatric Genetics: April 2011; 21, 2; 57–68

3. Kebir O & Krebs MO. Diethylstilbestrol and risk of psychiatric disorders: A critical review and new insights. World J Biol Psychiatry. 2011 Mar 23. [Epub ahead of print]

УЧАСТИЯ В НАУЧНИ ФОРУМИ

A. Yosifova, T. Mushiroda, M. Kubo, A. Takahashi, Y. Kamatani, D. Stoianov, R. Vazharova, S. Karachanak, I. Zaharieva, V. Milanova, N. Madjirova, I. Gerdjikov, T. Tolev, S. Velkova, G. Kirov, M. J. Owen, M. C O'Donovan, D. Toncheva, Y. Nakamura. Genome wide association study on Bipolar disorder. Bulgarian-Japanese Symposium "Genomics and Proteomics in Personalized Medicine", 19–20 March 2009, Sofia, Bulgaria

A. Yosifova, T. Mushiroda, M. Kubo, A. Takahashi, Y. Kamatani, D. Stoianov, R. Vazharova, I. Dimova, S. Karachanak, I. Zaharieva, S. Hadjidekova, V. Milanova, N. Madjirova, I. Gerdjikov, T. Tolev, N. Poryazova, G. Kirov, M. Owen, M. O'Donovan, D. Toncheva, Y. Nakamura. Case – control association study of candidate genes and genome – wide association study in Bulgarian patients with bipolar affective disorder. European Human Genetics Conference, June 12 - 15, 2010, Gothenburg, Sweden.

A.Yosifova, E. Betcheva, T. Mushiroda, M. Kubo, A. Takahashi, Y. Kamatani, N. Kamatani, D. Stoianov, R. Vazharova, I. Dimova, S. Karachanak, I. Zaharieva, S. Hadjidekova, V. Milanova, N. Madjirova, I. Gerdjikov, T. Tolev, N. Poryazova, G. Kirov, M. Owen, M. O'Donovan, D. Toncheva, Y. Nakamura. Differential analysis of molecular profiles between schizophrenia and bipolar affective disorder after genome-wide association study in Bulgarian population. European Human Genetics Conference, May 28 - 31, 2011, Amsterdam, The Netherlands