

دانشگاه علوم بهزیستی و توانبخشی

گروه ژنتیک

پایان نامه کارشناسی ارشد

رشته ژنتیک انسانی

عنوان

شناسایی مسیرهای مولکولی درگیر در بیماری ناتوانی ذهنی در خانواده ای با جهش در ژن CDK9
با استفاده از تکنیک RNA-seq

نگارنده

فاطمه پیمانی فروشانی

اساتید راهنما

دکتر کلثوم اینانلو

دکتر حسین نجم آبادی

استاد مشاور

دکتر کیمیا کهریزی

ماه سال

بهمن ۱۳۹۷

شماره ثبت

۹۵۱۱۶۳۰۰۳

تقدیم

به محضر ارزشمند پدر و مادر عزیزم

بابت تمام زحمات و پشتیبانی آنها در تک تک لحظه‌های زندگی ام ...

تقدیر و تشکر

شکر ایزد یزدان که توفیق اتمام این پایان نامه را نصیب من ساخت.

با تشکر از خانواده عزیزم که حمایت‌هایشان مشوق من در تمام طول زندگی بوده است.

با تشکر از استاد راهنمای نازنینم دکتر اینانلو عزیز که با حمایت‌های همیشگی و سعه صدر از هیچ کمکی بر من در

راستای انجام این پایان نامه دریغ نمودند و در طی این دوران بسیار به من آموختند.

با تشکر از استاد عزیز پروفسور نجم آبادی که بی شک بدون حمایت‌های ایشان اتمام این پایان نامه مقدور نبود.

از دکتر کهریزی عزیز بابت حمایت‌های همه جانبه کمال تشکر را دارم.

با تشکر صمیمانه از کلیه اساتید و کارکنان مرکز تحقیقات ژنتیک که مرا در انجام این پایان نامه یاری ساختند.

با تشکر از کلیه دوستانم که وجود آن‌ها مایه دلگرمی من در طول ایام تحصیل بوده است.

چکیده :

بیماری ناتوانی ذهنی یک اختلال عصبی-تکاملی با هتروژنیتی بالای ژنتیکی و کلینیکی و نرخ بروز ۱-۳٪ است. علی رغم اهمیت بالای این بیماری از نظر کلینیکی، اطلاعات کمی در خصوص پاتوژنیسیته این بیماری در دسترس است. مطالعه ترانسکریپتوم ابزاری کلیدی در بررسی ارتباط بین جهش‌ها و تغییر بیان ژن‌ها در افراد بیمار است. شناسایی ژن‌ها و ایزوفرم‌هایی با بیان متفاوت، تعیین عملکرد واریانت‌های جدید و شناخته شده و ارتباط بین تغییرات ژنتیکی و فنوتیپ تعدادی از کاربردهای بی شمار تکنیک RNA-seq است. در این مطالعه جهت شناسایی مسیرهای مولکولی و زیستی عامل بیماری ناتوانی ذهنی در خانواده‌ای با جهش در ژن تنظیم کننده رونویسی CDK9، ترانسکریپتوم کل خون افراد بیمار و سالم با استفاده از تکنیک RNA-seq مورد بررسی قرار گرفت. در نهایت ۹۷۲ ژن با $q\text{-value} < 0.05$ و $\text{foldchange} > 1.5$ به عنوان ژن‌هایی با بیان متفاوت در بین افراد بیمار و سالم، گزارش شدند (۵۱۳ ژن در افراد بیمار، افزایش بیان و ۴۵۹ ژن کاهش بیان نشان دادند). بر اساس آنالیزهای صورت گرفته مهم‌ترین مسیرهای مولکولی تغییر بیان یافته در افراد بیمار عبارتند از: مسیرهای بیان ژن، اجزای ساختاری ریبوزوم و مولکول‌های متصل شونده به RNA، تغییرات هیستون، پروتئین‌های تیروزین کیناز، پروتئین‌های GTPase انتقال دهنده پیام و اسکلت سلولی اکتینی. این یافته‌ها دیدگاه جدیدی را در خصوص مسیرهای مولکولی عامل بیماری ناتوانی ذهنی فراهم آورد و بر نقش فاکتورهای رونویسی در تنظیم عملکرد نورون‌ها تاکید کرد.

کلمات کلیدی:

ناتوانی ذهنی، ترانسکریپتوم، مسیرهای مولکولی، بیان ژن، RNA-seq، ژن CDK9

فهرست مطالب

فهرست تصاویر..... ج

فهرست جداول..... د

فصل اول کلیات پژوهش

- ۱-۱-۱. مقدمه..... ۱
- ۱-۱-۱. تعریف، طبقه بندی و علل بیماری ناتوانی ذهنی..... ۱-۱-۱
- ۱-۱-۲. شناسایی ژن‌های درگیر در بیماری ناتوانی ذهنی..... ۱-۱-۲
- ۱-۱-۳. نقش ژن *CDK9* در تکامل مغز..... ۱-۱-۳
- ۱-۱-۴. آنالیز ترانسکریپتوم..... ۱-۱-۴
- ۱-۱-۵. مسیرهای مولکولی درگیر در بیماری ناتوانی ذهنی..... ۱-۱-۵
- ۱-۱-۶. شباهت بافت خون و مغز..... ۱-۱-۶
- ۱-۲. بیان مسئله، اهمیت و ضرورت تحقیق..... ۱-۲
- ۱-۳. اهداف، فرضیات و سوالات..... ۱-۳
- ۱-۳-۱. هدف کلی..... ۱-۳-۱
- ۱-۳-۲. اهداف اختصاصی..... ۱-۳-۲
- ۱-۳-۳. اهداف کاربردی..... ۱-۳-۳
- ۱-۳-۴. سوال‌ها و فرضیه‌ها..... ۱-۳-۴

فصل دوم پیشینه پژوهش

- ۱-۲. مطالعات ترانسکریپتوم..... ۱-۲
- ۲-۲. کاربرد پروفایل بیان ژن در بیماری‌های عصبی-تکاملی..... ۲-۲
- ۳-۲. ارتباط بین پروفایل بیانی خون و مغز..... ۳-۲

فصل سوم روش شناسی تحقیق

- ۱-۳. نوع مطالعه..... ۱-۳
- ۲-۳. جامعه و نمونه آماری و روش نمونه گیری..... ۲-۳
- ۳-۳. روش جمع آوری داده‌ها..... ۳-۳
- ۴-۳. متغیرها..... ۴-۳
- ۵-۳. روش اجرا و تجزیه و تحلیل داده‌ها..... ۵-۳
- ۶-۳. ملاحظات اخلاقی..... ۶-۳
- ۷-۳. توالی‌یابی RNA از خون..... ۷-۳

۴۳ RNA استخراج ۱-۷-۳
۴۳ CDNA آماده سازی کتابخانه ۲-۷-۳
۴۵ توالی‌یابی نسل بعد ۳-۷-۳
۴۷ ۸-۳ آنالیز بیوانفورماتیکی داده های حاصل از مرحله توالی‌یابی
۶۳ ۹-۳ تأیید ژن‌هایی با بیان تغییر یافته با استفاده از روش QRT-PCR

فصل چهارم توصیف و تحلیل داده‌ها

۶۵ ۱-۴ نتایج حاصل از توالی‌یابی RNA
۶۵ ۲-۴ نتایج کیفیت سنجی (گزارش‌های حاصل از FASTQC)
۶۸ ۳-۴ نتایج TRIMMOMATIC
۶۸ ۴-۴ نتایج MAPPING
۶۹ ۵-۴ نتایج شمارش خوانش‌ها توسط HTSEQ
۷۰ ۶-۴ نتایج حاصل از DESEQ2
۷۱ ۷-۴ نتایج حاصل از آنالیز عملکردی بر روی لیست ژنی در FUNRICH
۷۳ ۸-۴ نتایج آنالیز ENRICHMENT
۷۵ ۹-۴ نتایج TOPPGENE
۷۶ ۱۰-۴ نتایج حاصل از GENOMATIX
۷۷ ۱۱-۴ نتایج تنظیم کننده‌های بالادست
۷۸ ۱۲-۴ نتایج حاصل از QRT-PCR

فصل پنجم بحث و نتیجه گیری

۸۰ ۱-۵ بحث و نتیجه گیری
۸۰ ژن‌های تغییر بیان یافته و کاندید بیماری ناتوانی ذهنی
۸۱ ۲-۵ G پروتئین‌ها، اسکلت سلولی و ناتوانی ذهنی
۸۴ ۳-۵ پروتئین SHOOTIN1
۸۷ ۴-۵ FGFR2
۸۸ ۵-۵ مسیر RAS/MAPK
۹۰ ۶-۵ ارتباط ریبوزوم و مسیرهای ترجمه با ناتوانی ذهنی
۹۱ ۷-۵ مسیر NFkB
۹۲ ۸-۵ ژن‌های ZINC FINGER
۹۳ ۹-۵ پروتئین‌های تنظیم کننده بالادست ژنهای تغییر بیان یافته
۹۴ پیشنهادات پژوهشی
۹۵ منابع

ضمائم	۱۰۶
ضمیمه شماره ۱: لیست ژن‌های افزایش بیان یافته در بیماران با q-value کمتر از ۰/۰۵	۱۰۶
ضمیمه ۲: لیست ژن‌های کاهش بیان یافته در بیماران با q-value کوچکتر از ۰/۰۵	۱۱۱
ضمیمه شماره ۳: لیست ژن‌هایی با بیان تغییر یافته که جهش در آن‌ها در مطالعات قبلی به عنوان عامل بیماری گزارش شده است	۱۱۵

فهرست تصاویر

تصویر ۱-۱: شجره نامه خانواده ای مبتلا به ناتوانی ذهنی	۵
تصویر ۱-۲: نقش سوپر کمپلکس طولیل سازی در از سرگیری شروع رونویسی	۶
تصویر ۱-۳: آنالیز سیستماتیک ژن‌های مرتبط با ناتوانی ذهنی	۱۱
تصویر ۱-۴: شبکه اینترکشن پروتئین‌های کد شده توسط ژن‌های شناخته شده برای بیماری ناتوانی ذهنی	۱۲
تصویر ۱-۵: مسیرهای سلولی جهت فعالیت صحیح نورون	۱۳
تصویر ۱-۶: مسیرهای مولکولی فعال در نورون‌ها	۱۴
تصویر ۳-۱: مراحل تهیه کتابخانه CDNA از RNA	۴۵
تصویر ۳-۲: مراحل آنالیز بیوانفورماتیکی داده‌های خام RNA-SEQ	۴۷
تصویر ۳-۳: فرمت فایل FASTQ ایجاد شده توسط دستگاه توالی‌یابی	۴۸
تصویر ۴-۱: گزارش کیفیت ایجاد شده توسط FASTQC	۶۶
تصویر ۴-۲: نمودار کیفیت بازها بر اساس موقعیت آن‌ها در خوانش بر مبنای PHRED	۶۷
تصویر ۴-۳: نمودار PER BASE SEQUENCE CONTENT	۶۷
جدول ۴-۴: تعداد کل شمارش‌ها برای هر نمونه	۶۹
تصویر ۴-۴: ده ژن با بالاترین میزان افزایش و کاهش بیان در بیماران نسبت به کنترل‌ها	۷۰
تصویر ۴-۵: HEATMAP نشان دهنده ژن‌های با بیان متفاوت و میزان تفاوت	۷۱
تصویر ۴-۶: عملکرد مولکولی و مسیرهای زیستی درگیر ژن‌هایی با بیان افزایش یافته	۷۲
تصویر ۴-۷: عملکرد مولکولی و مسیرهای زیستی درگیر ژن‌هایی با کاهش بیان در بیماران	۷۳
تصویر ۴-۸: اجزای سلولی با بیشترین تعداد ژن‌های درگیر	۷۶

تصویر ۴-۹. مسیر مرتبط با تمایز سلول‌های عصبی	۷۷
تصویر ۴-۱۰. مسیر مرتبط با تنظیم متابولیسم RNA	۷۷
تصویر ۴-۱۱. تنظیم کننده‌های بالادست پیشنهاد شده برای ژن‌های تغییر بیان یافته	۷۸
تصویر ۴-۱۲. نتایج تفاوت بیان ژن در نمونه کنترل و بیمار برای دو ژن <i>FGFR2</i> و <i>SHTN1</i>	۷۹
تصویر ۵-۱. مکانیسم فعال شدن رسپتور <i>GPR56</i>	۸۴
تصویر ۵-۲. مولکول‌های تعیین کننده در شکل گیری آکسون و اینترکشن آن‌ها با یکدیگر	۸۶
تصویر ۵-۳. مسیر مولکولی فعال کننده اسکلت سلولی	۸۷

فهرست جداول

جدول ۱-۱. طبقه بندی بیماران ناتوان ذهنی بر اساس میزان بهره هوشی	۳
جدول ۱-۲. لیست ژن‌هایی با بیان متفاوت در افراد مبتلا به اوتیسم در مقایسه با افراد سالم	۳۵
جدول ۲-۲. مسیرهای مولکولی درگیر در بیماری اوتیسم	۳۶
جدول ۱-۳. متغیرهای مطالعه مورد نظر	۴۱
جدول ۲-۳. توالی پرایمرهای طراحی شده برای ژن‌های کاندید	۶۴
جدول ۱-۴. نتایج کیفیت توالی‌یابی دستگاه توالی‌یابی کننده	۶۵
جدول ۲-۴. درصد خوانش‌های حفظ شده پس از اجرای TRIMMOMATIC	۶۸
جدول ۳-۴. آمار نقشه برداری نرم افزار STAR	۶۹
جدول ۴-۶. مسیرهای پیشنهاد شده توسط PANTHER	۷۴
جدول ۴-۷. مسیرهای پیشنهاد شده توسط REACTOME	۷۵

فصل اول کلیات پژوهش

۱-۱. مقدمه

ناتوانی ذهنی^۱ از جمله بیماری‌های رایج عصبی-تکاملی با سن شروع زود هنگام (قبل از ۱۸ سالگی) می‌باشد. این بیماری از جمله بیماری‌ها با هتروژنیتی لکوسی و فنوتیپی بسیار زیاد است و تاکنون مطالعات فراوانی جهت شناسایی ژن‌های درگیر در بیماری انجام گرفته است [۱] [۲].

با توجه به پیشرفت‌های اخیر صورت گرفته در توالی‌یابی DNA، سرعت بالا و هزینه مناسب آن نسبت به روش‌های قبل، یکی از روش‌های متداول برای یافتن ژن‌ها و جهش‌های عامل بیماری، روش توالی‌یابی کل اگزوم^۲ و یا کل ژنوم^۳ با استفاده از توالی‌یابی‌های نسل جدید^۴ می‌باشد [۳]. علی‌رغم شناسایی تعداد زیادی ژن درگیر در بیماری ناتوانی ذهنی، علت مولکولی بیماری در بسیاری از افراد مبتلا ناشناخته است [۴] [۵].

یکی از روش‌های مورد استفاده جهت تفسیر واریانت‌های گزارش شده و برقراری ارتباط بین واریانت و فنوتیپ بیماری و یافتن مسیرهای مولکولی درگیر در بیماری استفاده از آنالیز ترانسکریپتوم است [۶] [۷] [۸]. آنالیز ترانسکریپتوم با استفاده از روش RNA-seq، با توجه به پوشش کل RNAهای بیان شده توسط سلول، بررسی دقیق اختلاف بیان ژن در حالت نرمال و بیماری، مشخص کردن ایزوفرم‌های مختلف کد شده توسط ژن و یافتن RNAهای جدید، روش مناسبی برای بررسی مسیرهای مولکولی درگیر در بیماری‌های مختلف از جمله بیماری‌های ذهنی با پیچیدگی‌های بالا در مسیرهای کنترلی و سیگنالینگ می‌باشد.

۱-۱-۱. تعریف، طبقه بندی و علل بیماری ناتوانی ذهنی

طبق تعریف ارائه شده توسط DSM-5 (دستور عمل‌های تشخیصی و آماری اختلالات ذهنی ویرایش پنجم)^۵

¹Intellectual disability

²Whole exome sequencing

³Whole genome sequencing

⁴Next generation sequencing(NGS)

⁵ The Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition

بیماری ناتوانی ذهنی که پیش تر تحت عنوان عقب ماندگی ذهنی^۱ شناخته می‌شد، به صورت محدودیت چشمگیر در عملکرد ذهنی و رفتار اکتسابی که شامل مهارت‌های ادراکی، اجتماعی و عملی است، شناخته می‌گردد. سن بروز بیماری قبل از ۱۸ سالگی و سطح IQ در افراد مبتلا پایین تر از ۷۰ می‌باشد. این اختلال مزمن می‌تواند به صورت تنها (ایزوله) و یا همراه با سایر اختلالات ذهنی از جمله افسردگی، اختلال بیش فعالی/کم‌توجهی، صرع و اختلال طیف اوتیسم و... و یا به صورت بخشی از فنوتیپ یک سندرم بدشکلی^۲ که اندام‌ها و عملکرد آن‌ها را تحت تأثیر قرار می‌دهد، رخ دهد. فهم مطالب جدید و پیچیده، یادگیری مهارت‌های جدید و مدیریت مستقل مسائل از جمله چالش‌های موجود در این بیماری است [۹][۱۰][۱۱][۱۲].

دلیل بروز بیماری در حدود ۴۰ درصد از افراد مبتلا، مبهم و ناشناخته است. پنجاه درصد از موارد بیماری با علت شناخته شده، به دلیل عوامل محیطی از جمله فقر غذایی، ایسکمی مغزی پیش از/ هنگام تولد، عفونت‌های پس از تولد، مصرف الکل در طول دوران بارداری و... رخ می‌دهد. پنجاه درصد باقی مانده به علل ژنتیکی از جمله اختلالات کروموزومی، جهش‌های تک ژنی وابسته به X و اتوزومال، تغییرات تعداد کپی و... رخ می‌دهند [۳][۱۲]. به صورت دقیق تر تقریباً ۲۸ درصد از موارد بیماری با استفاده از فاکتورهای ژنتیکی قابل توجیه هستند [۱۳].

میزان بروز بیماری به دلیل تنوع زیاد عوامل ایجاد کننده آن در مناطق مختلف، متفاوت است ولی به صورت کلی با استناد به تعریف ارائه شده از ناتوانی ذهنی شیوع این بیماری حدود ۱ تا ۳ درصد در جمعیت‌های غربی و دو تا سه برابر در مناطقی با عوامل مستعد کننده‌ای هم‌چون سوء تغذیه، محرومیت‌های فرهنگی و خدمات بهداشتی نامناسب گزارش شده است [۱۰].

میزان بروز این بیماری در ایران حدود ۱۳ نفر به ازای هر ۱۰۰۰ نفر گزارش شده است که این میزان در

¹ Mental retardation

² Malformation

مردان نسبت به زنان و در بزرگسالان نسبت به جوانان بالاتر است. نتایج به دست آمده طبق اطلاعات مرکز آمار ایران، شیوع کمتر این بیماری (۱,۳٪) را در کشور ما نسبت به سایر مناطق جهان نشان می‌دهد که می‌تواند ناشی از روشهای مختلف جمع آوری داده و تعاریف مختلف ارائه شده برای ناتوانی ذهنی باشد [۱۴].

تشخیص رسمی بیماران ناتوان ذهنی از طریق انجام تست IQ صورت می‌گیرد و طبقه بندی این بیماران بر مبنای میزان بهره هوشی به صورت زیر است [۱۱][۱۲] (جدول ۱-۱):

جدول ۱-۱. طبقه بندی بیماران ناتوان ذهنی بر اساس میزان بهره هوشی

عمیق (خیلی شدید)	شدید	متوسط	خفیف	شدت بیماری
کمتر از ۱۹	۲۰-۳۵	۳۶-۵۱	۵۲-۶۹	میزان بهره هوشی (IQ)

۱-۱-۲. شناسایی ژن‌های درگیر در بیماری ناتوانی ذهنی

مغز اندامی بسیار پیچیده و متشکل از هزاران نوع سلول مرتبط با یکدیگر است. در زمان تکامل و هم چنین در طی فعالیت‌های روزانه پروتئین‌های بی‌شماری باید به میزان متناسب، در زمان و مکان صحیح فعال گردند. قابل تصور است که رخداد جهش و حذف و جابه‌جایی در هر یک از ژن‌های کدکننده این پروتئین‌ها، تکامل مغز و عملکردهای ذهنی را با اختلال مواجه می‌کند [۱۲].

تشخیص ژنتیکی بیماران مبتلا به ناتوانی ذهنی با توجه به اطلاعات زیادی که در خصوص پیش بینی وضعیت بیماری، پیچیدگی‌های احتمالی و طبقه بندی آن، انواع گزینه‌های درمانی و هم چنین الگوی توارث بیماری جهت مدیریت سایر اعضا در اختیار ما قرار می‌دهد، از اهمیت بالایی برخوردار است [۱۲].

هتروژنیتی ژنتیکی بالای این بیماری، تحقیقات ژنتیکی را به سمت بررسی‌های غیر جهت دار گسترده ژنومی^۱ شامل تکنیک میکروآرای^۲ و اخیراً توالی‌یابی کل اگزوم^۳، کل ژنوم^۴ و یا پنل‌های گسترده ژنومی^۵ با استفاده از تکنیک توالی‌یابی نسل جدید^۶، سوق داده است [۱۲]. ظهور تکنیک‌های پربازده توالی‌یابی کل اگزوم و توالی‌یابی کل ژنوم انقلابی در زمینه شناسایی ژن‌های عامل بیماری ناتوانی ذهنی ایجاد کرده است [۱۰]. استفاده از تکنولوژی‌های جدید همزمان با افزایش دانش زیستی ما، بازده تشخیصی و مزایای نتایج تست‌های ژنتیکی برای بیماران و خانواده‌های آن‌ها را به طرز شگفتی افزایش داده است [۱۲].

در یکی از مطالعات صورت گرفته توسط مرکز تحقیقات ژنتیک دانشگاه علوم بهزیستی و توانبخشی بر روی خانواده‌ای با ازدواج خویشاوندی با سه فرزند مبتلا به ناتوانی ذهنی، واریانت بدمعنی هموزیگوس $c.280C>T$ (p.R94C) در ژن *CDK9* با استفاده از روش توالی‌یابی کل اگزوم به عنوان جهش عامل بیماری در بیماران مورد مطالعه گزارش شد (تصویر ۱-۱) [۱].

¹ Unbiased genome-wide approaches

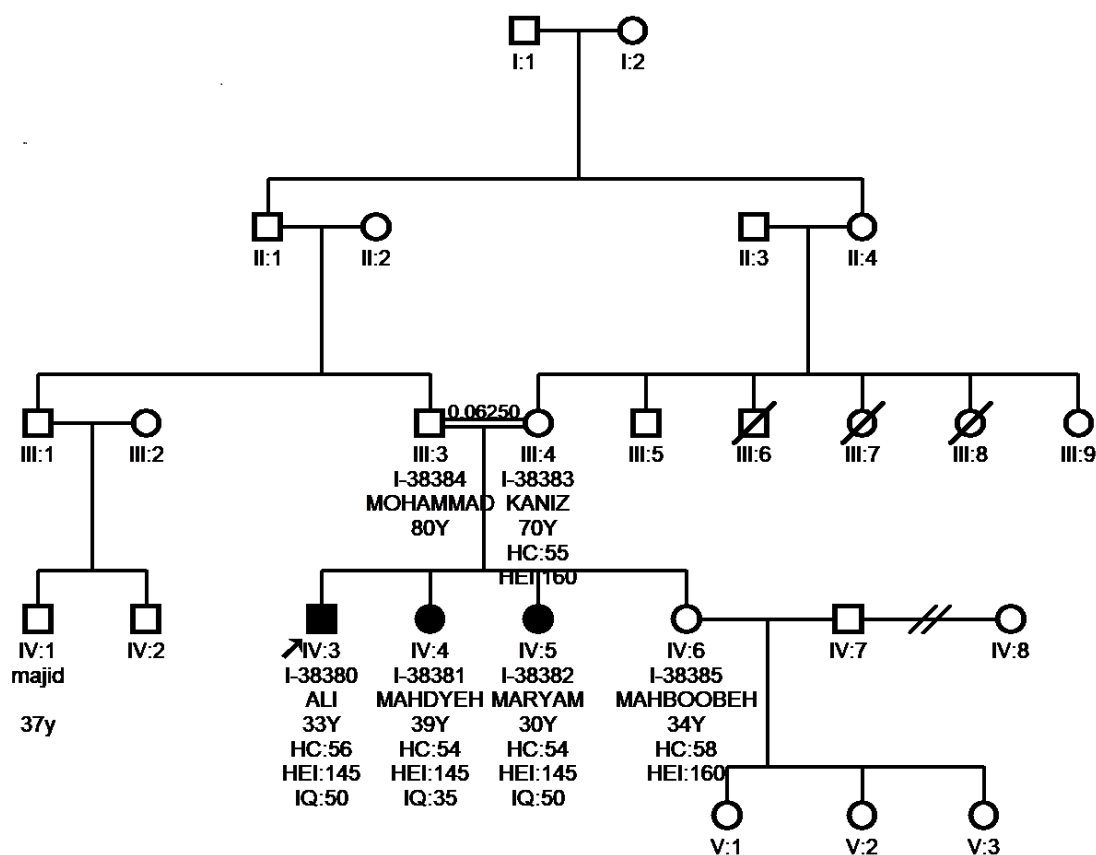
² Microarrays

³ Whole Exome Sequencing (WES)

⁴ Whole Genome Sequencing (WGS)

⁵ extensive gene panels

⁶ Next generation sequencing (NGS)



تصویر ۱-۱: شجره نامه خانواده ای مبتلا به ناتوانی ذهنی با ازدواج خویشاوندی که ژن عامل بیماری در این خانواده *CDK9* است.

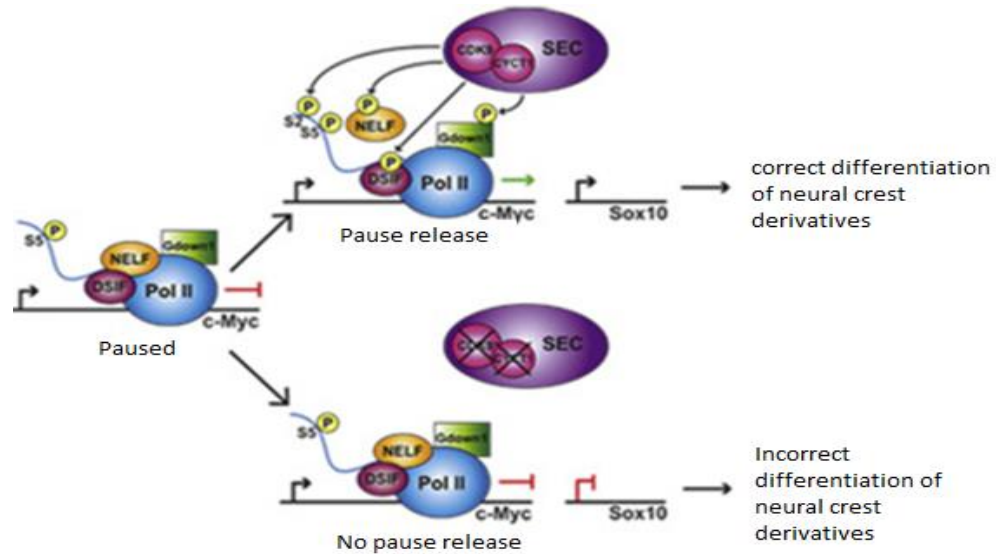
۱-۱-۳. نقش ژن *CDK9* در تکامل مغز:

تکامل صحیح جنین از نظر ساختاری و ذهنی در طول دوره نه ماهه بارداری، مستلزم کنترل دقیق و صحیح بیان ژن‌ها از جنبه مکانی و زمانی است. یکی از سطوح تنظیم بیان ژن، کنترل رونویسی ژن در مرحله طویل سازی^۱ است. RNA پلیمراز II^۲ پس از شروع رونویسی و سنتز حدود ۵۰ نوکلئوتید از mRNA، در ژن‌هایی که کنترل بیان آن‌ها در سطح طویل سازی رونویسی رخ می‌دهد، از حرکت باز می‌ایستد؛ شروع مجدد حرکت پلیمراز و از سرگیری پلیمره کردن نوکلئوتیدها، مستلزم فسفریلاسیون دم انتهایی کربوکسی

^۱elongation

^۲RNA polymerase II

پلیمراز و فاکتورهای توقف توسط سوپرکمپلکس طولی سازی^۱ است. یکی از اجزای سوپر کمپلکس طولی سازی پروتئین هتروداایمر P-TEFb (متشکل از CDK9+cyclinT1) است، که در آن CDK9^۲ وظیفه فسفریله کردن انتهای سرین دم RNA پلیمراز II و فاکتورهای توقف NELF، Gdown1، DSIF را بر عهده دارد (تصویر ۱-۲)[۱۵]. ژن *CDK9* به عنوان کیناز فعال کننده آنزیم RNA پلیمراز II یکی از اجزای اصلی سوپر کمپلکس طولی سازی را تشکیل می‌دهد[۱۶].



تصویر ۱-۲. نقش سوپر کمپلکس طولی سازی در از سرگیری شروع رونویسی ژن‌هایی که بیان آن‌ها متوقف شده است[۱۵].

توقف RNA پلیمراز II در نزدیکی ۵' ژن و دقیقاً کمی پایین دست نقطه شروع رونویسی، منجر به کنترل زمان رونویسی ژن‌هایی که بیان سریع آن‌ها جهت پاسخ به سیگنال‌های تکاملی ضروری است، می‌شود. به همین دلیل ژن‌هایی که در مرحله‌ی طولی سازی رونویسی (توقف پلیمراز) کنترل می‌گردند، ژن‌های تنظیم

¹ SEC:super elongation complex

²Cyclin dependant kinase

کننده اصلی^۱ و یا ژن‌های پاسخ دهنده سریع هستند که بیان آن‌ها باعث راه اندازی بیان ژن‌های پایین دست می‌گردد[۱۵].

جذب سوپر کمپلکس طویل سازی به RNA پلیمراز نه تنها در القای سریع بیان ژن‌های متوقف شده بلکه در هماهنگ و متوازن کردن بیان همزمان ژن‌ها، در تکامل نیز نقش دارد.

مطالعات در زوفیلا و سلول‌های جنینی نشان داده است که مرحله طویل سازی رونویسی تنها بیان ۱۰ تا ۳۰ درصد از ژن‌های ضروری کنترل کننده تکامل و نه کل ژنوم را تنظیم می‌کند[۱۵].

پروتئین‌های CDK9 و Cyclin-T1، اجزای کمپلکس P-TEFb، جهت تنظیم تمایز سلول‌های ستیغ - عصبی و تمایز نورون‌ها نیز حیاتی هستند[۱۵][۱۷]. هم‌چنین CDK9 به همراه Cyclin-K در حفظ

تمامیت ژنوم در پاسخ به استرس همانندسازی نیز نقش دارد[۱۸]. هم‌چنین ژن CDK9 در ۴ خانواده مبتلا

به سندرم مشابه CHARGE^۲ به عنوان ژن کاندید عامل بیماری گزارش شده است[۱۹،۲۰].

۱-۱-۴. آنالیز ترانسکریپتوم:

علی‌رغم کاربرد فراوان روش‌های توالی‌یابی نسل بعد، تعداد زیاد واریانتهای گزارش شده، هم‌پوشانی اندک بین واریانتهای یافت شده در مطالعات مختلف، نرخ پایین تمایز بین واریانتهای بیماری‌زا و غیر بیماری‌زا (۲۰٪-۵۰٪) و عدم توانایی در برقراری ارتباط دقیق بین واریانت و فنوتیپ بیماری، استفاده از روش‌های

توالی‌یابی ژنوم و اگزوم را با محدودیت رو به رو کرده است[۷][۲۱].

با توجه به اینکه حدود ۸۸ درصد از واریانتهای ژنومی در ناحیه غیرکدکنندهی بین ژنی و اینترونی قرار گرفته‌اند، بسیاری از این تغییرات تأثیری بر روی ساختار پروتئین نداشته و فقط میزان بیان آن را تغییر

^۱ Master regulator

^۲ CHARGE-like syndrome

می‌دهند[۶]. عمده‌ی تنظیمات ژنتیکی در انسان نیز در سطح بیان ژن رخ می‌دهد و هدف از مطالعات بیان

ژن در سطح ترانسکریپتوم، بررسی بیان و یا عدم بیان یک ژن و میزان بیان آن است[۲۲].

آنالیز ترانسکریپتوم در ابتدا با تمرکز بر روی RNA کد کننده پروتئین و بر مبنای تکنیک هیبریداسیون

(ریزآرایه^۱) انجام می‌گرفت و اخیراً بر مبنای تکنیک NGS و توالی‌یابی کل RNAهای سلول (RNA-seq)

صورت می‌پذیرد[۲۳].

اساس تکنیک microarray، هیبرید شدن ترانسکریپتوم با الیگونوکلوئوتیدهایی است که بر مبنای دانش

قبلی ما طراحی شده، بنابراین این روش فاقد پویایی لازم در شناسایی ژن‌ها و ایزوفرم‌های جدید، رونوشت-

های فیوز شده به یکدیگر و ویرایش‌های صورت گرفته بر روی RNA است[۶][۲۳].

بررسی تغییرات بیان RNAهای غیرکدکننده، با توجه به رونویسی آن‌ها توسط بخش بزرگی از ژنوم (در

مقایسه با رونویسی RNAهای کد کننده توسط تنها ۲ درصد از ژنوم) و نقش آن‌ها در تنظیم بیان ژن‌ها،

کنترل پروسه‌های اپی ژنتیکی، مسیرهای مولکولی و تکاملی و... حائز اهمیت می‌باشد[۲۴]. RNAهای غیر

کد کننده بلند^۲ نقش گسترده‌ای در تکامل مغز و پیچیدگی‌های مسیرهای مولکولی دخیل در آن دارند. هم

چنین بررسی ایزوفرم‌های مختلف با توجه به وقوع پیرایش متناوب^۳ در بیش از ۹۰ درصد از ژن‌ها دارای

اهمیت بالایی است. توالی‌یابی RNA^۴ با استفاده از توالی‌یابی نسل بعدی، مسیر را برای بررسی اصولی کل

ترانسکریپتوم (شامل فاکتورهای ژنتیکی و اپی ژنتیکی) و شبکه‌های تنظیمی فراهم آورده است[۲۴].

پروفایل مقایسه‌ای بیان ژن، ابزاری مفید است در شناسایی ژن‌هایی با بیان متفاوت در ارتباط با بیماری‌های

مختلف از جمله اختلالات ذهنی[۲۵].

¹ microarray

² Long non-coding RNA(lncRNA)

³Alternative splicing

⁴ RNA Sequencing

مزایای RNA-seq شامل: توانایی شناسایی تمام انواع RNA های سلول اعم از شناخته شده یا ناشناخته، توانایی تفکیک ایزوفرم‌های مختلف ژن، بررسی‌های بیان مختص آلی^۱، اندازه‌گیری دقیق میزان رونویسی ژن‌هایی با بیان اندک، تخمین نسبتا دقیق از بیان به صورت sense/antisense، آنالیز نقطه آغاز رونویسی، شناسایی ویرایش‌های^۲ صورت گرفته بر روی RNA و... می‌باشد. همچنین تکنیک توالی‌یابی، میزان بیان ژن را با شمارش خوانش‌های^۳ تولید شده برای یک ناحیه خاص اندازه‌گیری می‌کند که در مقایسه با تخمین میزان بیان ژن بر اساس قدرت سیگنال نوری تولید شده در اثر هیبریداسیون در روش microarray. دقت بالاتری دارد [۶].

با استفاده از روش RNA-Seq امکان بررسی مستقیم تغییرات رخ داده در توالی و میزان RNA وجود دارد که در تفسیر اطلاعات ژنتیکی حاصل از توالی‌یابی DNA کمک‌کننده است [۲۶].

۱-۱-۵. مسیرهای مولکولی درگیر در بیماری ناتوانی ذهنی

درصد قابل توجهی از بیماری ناتوان ذهنی در اثر جهش در یک ژن رخ می‌دهد. تاکنون بالغ بر ۱۴۱۶ ژن عامل بیماری ناتوانی ذهنی گزارش شده است؛ که از این تعداد ۸۰۲ ژن در فرم اتوزوم مغلوب بیماری، ۵۲۵ ژن در فرم غالب بیماری و ۱۳۲ ژن در فرم وابسته به X ناتوانی ذهنی گزارش شده است [۲۷]. پروتئین‌های کد شده توسط ژن‌های شناخته شده‌ی این بیماری از طریق برهم‌کنش مستقیم و یا برهم‌کنش‌های پیچیده و غیرمستقیم در یک و یا چند مسیر مولکولی مشترک فعالیت می‌کنند و همگرایی این مسیرهای مولکولی منجر به بروز ویژگی‌های فنوتیپی مشخص و رایج در بیماران می‌شود (تصویر ۱-۳ و ۱-۴). با توجه به هتروژنی بالای بیماری ناتوانی ذهنی، درمان این بیماران در سطح ژنی غیرممکن است و مسیرهای مولکولی مشترک اختلال یافته در بیماران می‌بایست هدف درمان قرار گیرند. ژن‌های مرتبط با ناتوانی ذهنی با اختلال

¹Allele specific expression

²RNA editing

³ read

در پروسه‌های سلولی نظیر شکل‌گیری نورون‌ها، مهاجرت نورونی، عملکرد سیناپسی، مسیرهای سیگنالینگ سلولی و ایمنی ذاتی منجر به بروز بیماری می‌شوند. در اغلب موارد، فرم مغلوب بیماری ناتوانی ذهنی در اثر جهش در ژن‌هایی که در رونویسی، ترجمه، تقسیم سلولی، تخریب پروتئین و متابولیسم انرژی فعالیت می‌کنند، ایجاد می‌شود [۴][۱][۱۲].

همانطور که در تصویر ۱-۳ مشخص است مسیرهای مرتبط با تکامل سیستم عصبی، متابولیسم RNA، انتقال دهنده‌ها، رونویسی و متابولیسم دارای بیشترین سهم از بین ژن‌هایی است که تاکنون در ارتباط با بیماری ناتوانی ذهنی شناخته شده است. هم‌چنین ژن‌های مرتبط با مسیرهای Hedgehog، رسپتور گلوتامات، پراکسی‌زوم، گلیکوزیلاسیون و مژک‌ها^۱ بیشترین میزان غنی‌سازی ژن^۲ را نسبت به پس‌زمینه ژنتیکی نشان داده‌اند [۴].

¹ cilia

² Enrichment