

آنالیز شبکه برهمکنش پروتئین - پروتئین بر اساس ژن‌های تغییر بیان یافته در مایع مغزی نخاعی برای بیماری مالتیپل اسکلروزیس

ناهد صفری علی قیارلو^۱ (Ph.D student)، مصطفی رضایی طاویرانی^{۱*} (Ph.D)، محمد تقی زاده^۲ (Ph.D)، سید محمد طباطبائی^۳ (Ph.D student)، سعید نمکی^۴ (Ph.D)

۱ - مرکز تحقیقات پروتئومیکس، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

۲ - گروه بیوانفورماتیک، مرکز تحقیقات بیوشیمی و بیوفیزیک (IBB)، دانشگاه تهران، تهران، ایران

۳ - گروه انفورماتیک پزشکی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

۴ - گروه ایمونولوژی، دانشکده پزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران

چکیده

هدف: ام.اس شایع‌ترین بیماری تخریب میلین سیستم عصبی مرکزی است. روش‌های جدید، برای فهم بهتر بیماری‌زایی و معرفی اهداف دارویی جدید ضروری می‌باشند. هدف این مطالعه، استفاده از روش شبکه زیستی به منظور شناسایی نشانگرهای کلیدی جدید درگیر در بیماری‌زایی ام.اس بود.

مواد و روش‌ها: در این مطالعه، پروفایل بیان ژن مایع مغزی نخاعی مربوط به ۲۶ بیمار تازه مبتلا شده به ام.اس و ۱۸ فرد کنترل از پایگاه داده Array Express گرفته و آنالیز شد. ژن‌های متفاوت بیان شده با داده‌های برهمکنش پروتئین-پروتئین تلفیق و زیرشبکه برهمکنش پروتئین-پروتئین ساخته شد. ویژگی‌های توپولوژیکی شبکه حاصل توسط نرم‌افزار cytoscape مورد آنالیز قرار گرفت.

یافته‌ها: با آنالیز پروفایل بیان ژن، ۳۰۶۲ ژن متفاوت بیان شده ($FDR < 0.05$) در مایع مغزی نخاعی بیماران ام.اس در مقایسه با افراد کنترل مشخص شدند که ۱۰۸۰ ژن افزایش بیان و ۱۹۸۱ ژن کاهش بیان داشتند. با ادغام ژن‌های متفاوت بیان شده و داده‌های برهمکنش پروتئین-پروتئین، پروتئین زیرشبکه‌ای با ۱۴۴۰ رأس و ۳۵۰۰ یال ساخته شد. پس از آنالیز ویژگی‌های توپولوژیکی زیرشبکه، پنج ژن جدید که مقادیر بالای پارامترهای مرکزیت را داشتند، به عنوان نشانگرهای کاندید معرفی شدند. این نشانگرها در مسیرهای زیستی نظیر تنظیم دینامیک اسکلت سلولی، چرخه سلولی، اسپلیسوزوم نقش دارند. با بررسی متون، ارتباط نشانگرها در بیماری‌زایی ام.اس تأیید شد. نتیجه‌گیری: بنابراین، آنالیزهای بر پایه شبکه می‌تواند نشانگرهایی را معرفی کند که بتوان آن‌ها را به عنوان اهداف دارویی بالقوه برای ام.اس مورد بررسی بیشتری قرار داد.

واژه‌های کلیدی: ام.اس، پروفایل بیان ژن، نقشه‌های برهمکنش پروتئین-پروتئین، ویژگی‌های توپولوژیکی

مقدمه

ناتوانی‌های نورولوژیکی در بیماران ام.اس می‌گردد [۱، ۲]. ام.اس بیش‌تر در افراد جوان (بین ۲۰-۴۰ سال) و غالباً در خانم‌ها مشاهده می‌شود. شروع مراحل خودایمنی در ام.اس ناشناخته است. نتایج مطالعات پیشنهاد می‌دهد که ام.اس

ام.اس بیماری خودایمنی التهابی ایجادشونده به واسطه T cell می‌باشد که در نتیجه حمله نادرست سیستم ایمنی به پوشش میلین سیستم عصبی مرکزی است. تخریب میلین سبب

درجه رأس (degree)، closeness، betweenness، centrality برای شناسایی رأس‌هایی که از نظر عمل‌کردی در شبکه مهم‌اند، بسیار قابل توجه هستند [۸، ۱۸، ۱۹]. در شبکه برهمکنش پروتئین-پروتئین به هر پروتئین رأس و به برهمکنش بین پروتئین‌ها، یال می‌گویند. رأس‌ها با درجه بالا، "هاب" و رأس‌ها با betweenness بالا، "گلوگاه" نامیده می‌شوند که این رأس‌ها نقش بسیار مهمی در شبکه دارند [۲۰].

در این مطالعه پروفاایل بیان ژن و داده‌های برهمکنش پروتئین-پروتئین با یکدیگر تلفیق شدند تا زیرشبکه برهمکنش برای ژن‌های تغییر بیان‌یافته در مایع مغزی نخاعی بیماران ام‌اس ساخته شود. ویژگی‌های توپولوژیکی زیرشبکه حاصل مورد آنالیز قرار گرفتند و ژن‌هایی که پارامترهای مرکزیت بالایی داشتند، مشخص شدند. پس از بررسی متون، پنج ژن جدید با اندازه‌های مرکزی بالا در زیرشبکه، به عنوان نشانگرهای کاندید معرفی شدند.

مواد و روش‌ها

تهیه پروفاایل بیان ژن. پروفاایل بیان ژن مربوط به بافت مایع مغزی نخاعی بیماران ام‌اس با شناسه EMTAB-69 از پایگاه داده ArrayExpress دانلود شد. داده‌های EMTAB-69 توسط Brynedal و همکارانش در سال ۲۰۱۰ منتشر شد [۵]. در این مطالعه، ۲۶ بیمار ام‌اس مراجعه‌کننده به بیمارستان دانشگاه Karolinska در سوئد، شرکت داشتند که از آن‌ها ۱۲ بیمار در فاز عود (relapse) و ۱۴ بیمار در فاز فروکش (remission) بودند. بیماران ام‌اس، از یک جمعیت بیمار تازه مبتلا شده به بیماری انتخاب شده و هیچ‌کدام از این بیماران داروی تنظیم‌کننده ایمنی دریافت نکرده بودند. جمعیت کنترل نیز شامل ۱۸ فرد با سابقه بیماری‌های سیستم عصبی غیر التهابی به منظور شناسایی پروفاایل بیان ژن مخصوص ام‌اس، بودند. میانگین سن در بیماران ام‌اس ۳۹/۲ و در گروه کنترل ۴۰/۶ و نسبت جنسیت (زن به مرد) ۱ به ۳ در بیمار و ۱ به ۲ در گروه کنترل گزارش شده است. از همه افراد، نمونه‌های

حاصل ترکیب فاکتورهای ژنتیکی، محیطی و عفونی می‌باشد [۳]. با این وجود، دلیل ام‌اس، شروع غیرمنتظره بیماری، سرعت مختلف پیشرفت بیماری به صورت معما باقی مانده است. بنابراین روش‌های جدیدی برای فهم بهتر مکانیسم بیماری‌زایی و اهداف دارویی جدید برای هر مرحله از بیماری مورد نیاز است [۴].

بررسی پروفاایل بیان ژن اندام هدف یعنی سیستم عصبی مرکزی، می‌تواند تحقیق جامعی از بیان ژن‌های تغییر یافته در مراحل مختلف بیماری را فراهم کند. اما دسترسی به نمونه‌های مغزی و نخاعی بیماران ام‌اس به دلیل محدودیت‌های روش‌های موجود، مشکل می‌باشد. از طرفی، نمونه‌های پس از مرگ به سختی فاز بیماری را معلوم می‌کنند. بنابراین، در بسیاری از مطالعات، سلول‌های مایع مغزی نخاعی به عنوان جانشین اندام هدف در بیماری‌های سیستم عصبی مرکزی مورد استفاده قرار می‌گیرند [۵].

تغییر سطح بیان ژن در پروفاایل بیان ژن ممکن است در نتیجه تغییر بیان، ژن دیگر باشد و در واقع، ممکن است نتیجه مستقیم تغییر فنوتیپ سلولی نباشد. در نتیجه اطلاعات دیگری مورد نیاز است تا این تغییرات را در بستر زیستی (سلولی) قرار دهد [۶]. آنالیزهای بر پایه شبکه برهمکنش پروتئین-پروتئین ارتباطات شناخته شده بین پروتئین‌ها را در بستر بیوشیمیایی، آبشارهای سلولی و شبکه‌های بیومولکولی معرفی می‌کنند [۷-۹]. آنالیز تلفیق پروفاایل بیان ژن و داده‌های برهمکنش پروتئین-پروتئین، ژن‌های تغییر بیان‌یافته را در بستر زیستی قرار می‌دهند [۱۰-۱۲]. شبکه‌های پروتئینی، گروه‌بندی عمل‌کردی ژن‌های برهم‌کنش‌دهنده افزایش/کاهش بیان‌یافته را منعکس می‌کنند. بنابراین، ممکن است نقش زیرمجموعه‌ای از این ژن‌ها، با این آنالیزهای تلفیقی مشخص شود [۶].

در سال‌های اخیر، آنالیزهای توپولوژیکی (ظاهری) به منظور شناسایی پروتئین‌ها با نقش‌های زیستی مهم، روی شبکه برهمکنش پروتئین-پروتئین به صورت گسترده انجام می‌گیرد [۱۳-۱۷]. پارامترهای مرکزیت (Centrality) نظیر

یال‌های متصل به رأس. ۲- Betweenness به صورت زیر محاسبه می‌شود:

$$C_B(v) = \sum_{s \neq t \neq v \in V} \frac{\rho_{st}(v)}{\rho_{st}}$$

در واقع این فرمول نشان‌دهنده نسبت تعداد کوتاه‌ترین طول مسیرهای عبورکننده از رأس v نسبت به کل کوتاه‌ترین طول مسیرها بین رأس‌های s و t است.

۳- Closeness centrality به صورت زیر محاسبه می‌شود:

$$C_C(v) = \frac{1}{\sum_{u \in V} \text{dis}(u, v)}$$

در واقع این فرمول، نشان‌دهنده میزان فاصله یک رأس با بقیه رأس‌ها در شبکه است که هر چه این فاصله کم‌تر باشد، رأس مرکزیت بیشتری در شبکه دارد. بنابراین، اندازه زیاد برای یک رأس بیانگر نزدیک بودن سایر رأس‌ها به رأس مورد نظر است.

برای تعیین ژن‌هایی که درجه رأس بالا دارند، از روش زیر استفاده شد. ابتدا میانگین و انحراف معیار درجه رأس همه رأس‌ها محاسبه و طبق فرمول (میانگین + ۲ برابر انحراف معیار) Cut-off، حد آستانه‌ای تعریف شد [۳۰]. ژن‌هایی که درجه رأسشان از این حد آستانه بالاتر بود به عنوان ژن‌های هاب معرفی شدند. برای معرفی ژن‌های گلوگاه، ژن‌هایی که ۱۰ درصد بالای betweenness را داشتند، در نظر گرفته شدند.

معرفی نشانگرهای کاندید. از سه ویژگی توپولوژیکی مهم ژن‌ها در زیرشبکه، برای شناسایی نشانگرهای کاندید استفاده شد. ژن‌ها بر اساس درجه رأس، betweenness و closeness centrality از مقادیر بالا به پایین رتبه‌بندی شدند. سپس، ۲۰ ژن که در هر ستون بالاترین مقادیر را داشتند، مورد بررسی بیش‌تر قرار گرفتند. در این میان پنج ژن جدید که در هر سه ستون مشترک بودند و از طریق بررسی متون، ارتباط معنادارتری با بیماری ام.اس داشتند، به عنوان نشانگرهای کاندید معرفی شدند.

مابع مغزی نخاعی که حدود ۸۰ درصد حاوی لئوسیت‌های T بود، گرفته شده است. نمونه‌ها بلافاصله سانتریفوژ شده‌اند و تا زمان استفاده، در دمای ۷۰- درجه سانتی‌گراد ذخیره شده‌اند. با استفاده از کیت‌های حاوی RNeasy، RNA استخراج شده است. غلظت RNA با استفاده از بیوآنالیزور Agilent 2100 اندازه‌گیری شده است. مقدار RNA در نمونه‌های مایع مغزی نخاعی در دامنه ۰/۸ - ۱۰ نانوگرم و به طور متوسط ۷/۱ نانوگرم گزارش شده است. اهداف نشان‌دار شده به طور جداگانه با ریزارایه‌های U133 plus 2.0 arrays شرکت افی متریکس هیبرید شده‌اند که بیان بیش از ۴۷۰۰۰ رونوشت مربوط به ۲۹۰۰۰ ژن را مورد بررسی قرار می‌دهد.

آنالیز داده‌های بیان ژن. با استفاده از الگوریتم موجود در پکیج affy، داده‌های خام تبدیل به مقادیر بیان ژن شدند [۲۱]. پس از پردازش اولیه، پروفایل بیان ژن حاوی ۵۴۶۷۵ پروب بود که از این میان، پروب‌هایی که با توان کم‌تری تشخیص داده می‌شدند، توسط پکیج genefilter حذف شدند [۲۲]. پس از پردازش اولیه، ۲۷۳۳۶ پروب برای هر نمونه برای آنالیزهای بعدی مورد استفاده قرار گرفتند. برای تعیین ژن‌های تغییر بیان‌یافته، پکیج limma استفاده شد [۲۳]. در این پکیج از برای تعیین ژن‌های تغییر بیان یافته شده است. همه آنالیزهای فوق در محیط برنامه‌نویسی R انجام شد.

استخراج زیرشبکه مربوط به ژن‌های تغییر بیان یافته و آنالیز ویژگی‌های توپولوژیکی. از چهار پایگاه داده اصلی برهمکنش پروتئین-پروتئین، IntAct، MINT، DIP و InnateDB به عنوان منبع داده‌های برهمکنش پروتئین-پروتئین انسانی استفاده شد [۲۴-۲۷]. ژن‌های تغییر بیان یافته در مایع مغزی نخاعی بیماران ام.اس نسبت به گروه کنترل روی شبکه برهمکنش پروتئین-پروتئین قرار گرفتند. این زیرشبکه فقط حاوی ژن‌های تغییر بیان یافته به صورت رأس و ارتباط بین آن‌ها به صورت یال بود. زیرشبکه توسط نرم‌افزار Cytoscape ساخته و نمایش داده شد [۲۸]. سه پارامتر اصلی مرکزیت برای تعیین ژن‌های با عمل‌کرد زیستی مهم مورد استفاده قرار گرفتند [۲۹]. ۱- درجه رأس: اندازه‌گیری تعداد

نتایج

آنالیز پروفایل بیان ژن. از پکیج limma برای آنالیز پروفایل بیان ژن و مقایسه ژن‌های متفاوت بیان شده در مایع مغزی نخاعی در بیماران ام.اس در مقایسه با کنترل استفاده شد. نتایج آنالیز آماری نشان داد که ۳۰۶۲ ژن ($0.05 < FDR$) تفاوت بیان داشتند که از این تعداد ۱۰۸۰ ژن افزایش بیان و ۱۹۸۱ ژن کاهش بیان داشتند.

آنالیز توپولوژیکی زیرشبکه برهمکنش پروتئین- پروتئین در این مطالعه فقط از برهمکنش‌های مستقیم بین ژن‌های متفاوت بیان شده برای ساختن زیرشبکه برهمکنش پروتئین- پروتئین استفاده شد. زیرشبکه مربوط به ژن‌های تغییر بیان یافته در مایع مغزی نخاعی بیماران ام.اس شامل ۱۴۴۰ رأس و ۳۵۰۰ یال بود. ویژگی‌های توپولوژیکی مورد بررسی قرار گرفت تا شبکه زیستی از شبکه تصادفی متمایز گردد. در شبکه‌های زیستی رأس‌ها از قانون $P(k) = k^{-\lambda}$ power law پیروی می‌کنند که $P(k)$ توزیع درجه رأس، k درجه رأس و λ توان درجه می‌باشد. در زیرشبکه برهمکنش پروتئین- پروتئین مایع مغزی نخاعی، $\lambda = 1.94$ بود که نشان‌دهنده این است که رأس‌ها از این قانون پیروی می‌کنند (شکل ۱، بالا- سمت چپ). علاوه بر این، چندین ویژگی توپولوژیکی زیرشبکه نیز مورد بررسی قرار گرفتند. پارامتر betweenness تعداد کوتاه‌ترین مسیرهایی که از یک رأس در شبکه می‌گذرد را اندازه می‌گیرد. پارامتر Closeness centrality یک رأس، نشان‌دهنده میزان فاصله یک رأس با بقیه رأس‌ها در شبکه است که هر چه این فاصله کم‌تر باشد،

رأس مرکزیت بیشتری در شبکه دارد. ضریب خوشه‌بندی، میزان اتصالات همسایه‌های یک رأس را اندازه می‌گیرد. همان‌طور که در شکل ۱ (پایین- سمت چپ) مشاهده می‌شود، هرچه تعداد همسایه‌ها بیشتر شود، تمایل به خوشه‌بندی نیز بیشتر می‌شود. کوتاه‌ترین طول مسیر دلالت بر ارتباطات نزدیک پروتئین‌ها دارد. دامنه کوتاه‌ترین طول مسیر پروتئین‌های زیرشبکه بین ۴-۵ است (شکل ۱، پایین- سمت راست). پارامتر Network centralization، میزان تأثیر حذف برخی از رأس‌های مرکزی در شبکه را اندازه‌گیری می‌کند. پارامتر Network density بیانگر نسبت بالقوه از اتصالات در یک شبکه که بیانگر اتصالات واقعی هستند، می‌باشد. پارامتر Network diameter، نشان‌دهنده اندازه خطی یک شبکه است. جدول ۱، ویژگی‌های توپولوژیکی مهم زیرشبکه ژن‌های متفاوت بیان شده در مایع مغزی نخاعی بیماران ام.اس را نشان می‌دهد.

معرفی نشانگرهای کاندید. بیست ژن که بالاترین مقادیر برای سه پارامتر مرکزی داشتند، به عنوان منبع انتخاب نشانگرها مورد بررسی قرار گرفتند (جدول ۲). سپس پنج ژن *RAC1*، *CDKN1A*، *EZH2*، *SRPK1* و *PPP1CA* که دارای درجه رأس بالاتر (هاب)، betweenness بالاتر (گلوگاه) و closeness centrality بالاتر (کم‌ترین فاصله را با بقیه ژن‌ها) در زیرشبکه بودند، و هم‌چنین در بررسی متون ارتباط معنادارتری با بیماری ام.اس داشتند به عنوان نشانگر شناسایی شدند (جدول ۳).

جدول ۱. ویژگی‌های توپولوژیکی زیرشبکه برهمکنش پروتئین- پروتئین مایع مغزی نخاعی در بیماری ام.اس

Topological parameters	Values
Number of nodes	1440
Number of edges	3500
Clustering coefficient	0.056
Average shortest pathlength	4.4
Network centralization	0.046
Network density	0.003
Network diameter	11

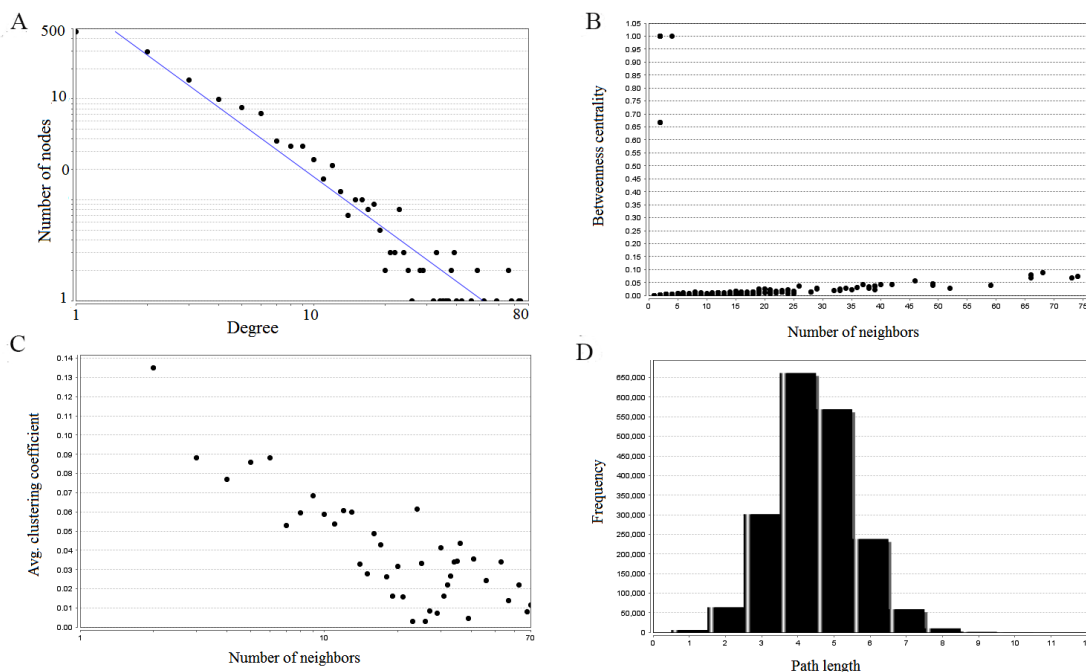
جدول ۲. ژن‌های دارای مقادیر بالای سه پارامتر مرکزیت در زیرشبکه برهمکنش پروتئین-پروتئین مایع مغزی نخاعی در بیماری ام.اس

Node	Degree	Node	Betweenness Centrality	Node	Closeness Centrality
IKBKG	74	FN1	0.089797	FN1	0.337911
GABARAPL1	73	ESR1	0.080572	ESR1	0.333737
FN1	68	IKBKG	0.073661	YWHAG	0.327627
ESR1	66	YWHAG	0.069476	IKBKG	0.327005
YWHAG	66	GABARAPL1	0.067134	CTNNB1	0.323171
GOLGA2	59	PPP1CA	0.056265	GABARAPL1	0.322415
ARRB2	52	CTNNB1	0.04659	MCM7	0.319055
CTNNB1	49	ATXN1	0.04353	EZH2	0.314468
EED	49	RAC1	0.042199	PPP1CA	0.313895
PPP1CA	46	SRPK1	0.041248	SRPK1	0.313824
ATXN1	42	EED	0.041154	EED	0.311624
RAC1	40	GOLGA2	0.038837	YBX1	0.31078
CDKN1A	39	CDKN1A	0.038047	RAC1	0.308071
YBX1	39	MCM7	0.037438	RBBP4	0.307178
MCM7	39	CREB3	0.036152	CDKN1A	0.30629
EZH2	38	EZH2	0.03354	TUBA1A	0.306222
ARRB1	38	SFN	0.032221	SFN	0.304665
SRPK1	37	TNFRSF1A	0.029617	FOS	0.304261
SFN	36	TK1	0.028969	RPS16	0.303993
STK4	35	ARRB1	0.027865	ARRB1	0.303859

جدول ۳. لیست ژن‌های کاندید در زیرشبکه برهمکنش پروتئین-پروتئین مایع مغزی نخاعی در بیماری ام.اس

Symbol	Full name	Fold change	direction
RAC1	Ras-related C3 botulinum toxin substrate 1	-1	-1
CDKN1A	Cyclin-dependent kinase inhibitor 1 A	-1	-1
EZH2	enhancer of zeste 2 polycomb repressive complex 2 subunit	1	1
SRPK1	SRSF protein kinase 1	1	1
PPP1CA	protein phosphatase 1 catalytic subunit alpha	1	1

۱- به معنای کاهش بیان و ۱ به معنای افزایش بیان است.



شکل ۱. ویژگی‌های توپولوژیکی زیر شبکه برهمکنش پروتئین-پروتئین. A. نمودار توزیع درجه رأس. B. نمودار betweenness. C. نمودار میانگین ضریب خوشه بندی. D. نمودار میانگین کوتاه ترین طول مسیر.

بحث و نتیجه گیری

اگرچه مطالعات ژنتیکی زیادی در مورد بیماری زایی ام.اس انجام گرفته است، اما هنوز تعدادی از ژن‌های درگیر در بیماری ناشناخته باقی مانده‌اند [۳۲،۳۱]. آنالیز داده‌ها بر پایه شبکه زیستی ابزارهای بسیار کاربردی برای تعیین ژنتیک بیماری‌های پیچیده می‌باشند که ژن‌ها و فرایندهای غیرقابل تشخیص توسط مطالعات ژنتیکی را معرفی می‌کنند [۳۴،۳۳]. در این مطالعه ژن‌های تغییر بیان یافته در مایع مغزی نخاعی بیماران ام.اس در مقایسه با کنترل، با داده‌های برهمکنش پروتئین-پروتئین تلفیق شدند. ویژگی‌های توپولوژیکی زیرشبکه مربوطه مورد بررسی قرار گرفت و ژن‌هایی که پارامترهای مرکزیت بالاتری (درجه رأس، betweenness و closeness centrality) در زیرشبکه داشتند، به عنوان نشانگرهای کاندید معرفی شدند.

در سال‌های اخیر، شبکه‌های برهمکنش پروتئین-پروتئین به طور قابل توجهی برای معرفی ژن‌های هدف در بیماری‌ها مورد استفاده قرار می‌گیرند. علاوه بر این، مطالعه بر اساس شبکه، دیدگاهی به سمت شناخت مکانیسم‌های بیماری زایی [۳۵،۱۱]، تعیین اهداف دارویی [۳۷،۳۶] را فراهم می‌کند. آنالیز ویژگی‌های توپولوژیکی پروتئین‌ها برای شناخت ساختار شبکه برهمکنش‌ها بسیار کمک‌کننده است؛ به طوری که اهمیت عمل‌کردی پروتئین‌ها از طریق آنالیزهای پارامترهای مرکزی مشخص می‌شود [۱۸]. بنابراین، آنالیز شبکه‌های برهمکنش پروتئین-پروتئین می‌تواند در تفسیر ژن‌های ناشناخته مؤثر باشد که این امر، درک مکانیسم‌های بیماری را تسهیل می‌کند و کمک‌کننده برای اولیت‌بندی ژن‌های درگیر در بیماری است [۳۸].

در این مطالعه، پنج ژن با افزایش یا کاهش بیان معنی‌دار در مایع مغزی نخاعی بیماران ام.اس که درجه رأس بالا (هاب)، betweenness بالا (گلوگاه) و closeness centrality بالا (کم‌ترین فاصله با بقیه رأس‌ها)، در زیرشبکه برهمکنش پروتئین-پروتئین داشتند، از طریق بررسی متون به عنوان نشانگرهای کاندید معرفی شدند. اولین نشانگر کاندید، ژن

RAC1 است که کاهش بیان یافته است. این ژن یک G (GTPase) پروتئین سیگنالی کوچک و عضو خانواده Rho می‌باشد. این ژن، نقش تنظیمی در دینامیک اسکلت سلولی اکتین و میکروتوبول دارد که این فرایند برای رشد و پایداری اکسون، دندریت و انعطاف‌پذیری ساختاری ستون فقرات، لازم می‌باشد. علاوه بر این، این ژن تنظیم‌کننده مهم NADPH-dependent membrane oxidase (NOX) است که این آنزیم نقش مهمی در تولید گونه‌های واکنش اکسیژن دارد. در نتیجه، RAC1 نقش مهمی در پاسخ التهابی و سمیت عصبی (neurotoxicity) میانجی‌شونده توسط سلول‌های میکروگلیا در سیستم عصبی دارد [۳۹]. در این راستا، نتایج یک مطالعه ارتباط اختلال تنظیم ژن RAC1 در بیماری زایی Amyotrophic Lateral Sclerosis (ALS) که یکی از بیماری‌های سیستم عصبی است، را نشان می‌دهد [۳۹]. بنابراین، می‌توان این ژن را مورد بررسی بیشتری قرار داد. دومین نشانگر کاندید، ژن CDKN1A است که کاهش بیان یافته است. این ژن، تنظیم‌کننده فرایند چرخه سلولی در مرحله G1 است. شواهد نشان می‌دهد که CDKN1A به عنوان بازدارنده چرخه سلولی، فرایند خودایمنی را مهار می‌کند [۴۰]. نتایج مطالعه‌ای نشان داده است که این ژن، تنظیم‌کننده مخصوص هومئوستاز لنفوسیت‌های T خاطره/فعال است [۴۱]. بنابراین می‌توان این ژن را مورد مطالعه بیشتری قرار داد. سومین نشانگر کاندید، ژن EZH2 است که افزایش بیان یافته است. این ژن، سه گروه متیل را به اسید آمینه لیزین ۲۷ از هیستون H3 در پروموتور ژن هدف اضافه می‌کند. مطالعه Li و همکاران نشان داد که فعالیت ژن EZH2 باید به صورت مداوم در نورون‌ها مهار گردد تا نورون‌ها به صورت مجدد وارد فرایند چرخه سلولی نشوند. بنابراین، افزایش بیان این ژن می‌تواند آغازگر فرایندهایی شود که در نهایت منجر به تخریب سلول‌های سیستم عصبی مرکزی شود [۴۲]. این نقش می‌تواند دلیل کاندید بودن این ژن باشد. چهارمین نشانگر کاندید، ژن SRPK1 است که افزایش بیان یافته است. این ژن در مراحل اولیه آپوپتوز فعال می‌شود. نقش‌های زیستی پیشنهادی برای

برهمکنش‌های موجود بین ژن‌های متفاوت بیان شده در بیماری‌ها است.

تشکر و قدردانی

این مقاله برگرفته از پایان‌نامه دوره دکترای تخصصی (Ph.D) خانم ناهید صفری علی قیارلو می‌باشد. از تمام کسانی که در انجام این تحقیق کمک کردند، تشکر و قدردانی می‌شود.

منابع

- [1] Noseworthy JH, Lucchinetti C, Rodriguez M, Weinshenker BG. Multiple sclerosis. *N Engl J Med* 2000; 343: 938-952.
- [2] Compston A, Coles A. Multiple sclerosis. *Lancet* (London, England) 2002; 359: 1221-1231.
- [3] Compston A, Coles A. Multiple sclerosis. *Lancet* (London, England) 2008; 372: 1502-1517.
- [4] Achiron A, Gurevich M, Friedman N, Kaminski N, Mandel M. Blood transcriptional signatures of multiple sclerosis: unique gene expression of disease activity. *Annal Neurol* 2004; 55: 410-417.
- [5] Brynedal B, Khademi M, Wallström E, Hillert J, Olsson T, Duvefelt K. Gene expression profiling in multiple sclerosis: a disease of the central nervous system, but with relapses triggered in the periphery? *Neurobiol Dis* 2010; 37: 613-621.
- [6] Wachi S, Yoneda K, Wu R. Interactome-transcriptome analysis reveals the high centrality of genes differentially expressed in lung cancer tissues. *Bioinformatics* 2005; 21: 4205-4208.
- [7] Tavirani PhD MR. Meningioma protein-protein interaction network. *Arch Iran Med* 2014; 17: 262.
- [8] Zamanian-Azodi M, Rezaei-Tavirani M, Rahmati-Rad S, Hasanzadeh H, Tavirani MR, Seyyedi SS. Protein-protein interaction network could reveal the relationship between the breast and colon cancer. *Gastroenterol Hepatol Bed Bench* 2015; 8: 215-224.
- [9] Safaei A, Tavirani MR, Oskouei AA, Azodi MZ, Mohebbi SR, Nikzamir AR. Protein-protein interaction network analysis of cirrhosis liver disease. *Gastroenterol Hepatol Bed Bench* 2016; 9: 114.
- [10] Li M, Wu X, Wang J, Pan Y. Towards the identification of protein complexes and functional modules by integrating PPI network and gene expression data. *BMC Bioinformatics* 2012; 13: 109.
- [11] Safari-Alighiarloo N, Taghizadeh M, Rezaei-Tavirani M, Goliaei B, Peyvandi AA. Protein-protein interaction networks (PPI) and complex diseases. *Gastroenterol Hepatol Bed Bench* 2014; 7: 17-31.
- [12] Safari Alighiarloo N, Taghizadeh M, Tabatabaei SM, Shahsavari S, Namaki S, Khodakarim S, Rezaei Tavirani M. Identification of new key genes for type 1 diabetes through construction and analysis of the protein protein interaction networks based on blood and pancreatic islet transcriptomes. *J Diabetes* 2016; 9: 764-777.
- [13] Huang KC, Yang KC, Lin H, Tsun-Hui TT, Lee WK, Lee SA, Kao CY. Analysis of schizophrenia and hepatocellular carcinoma genetic network with corresponding modularity and pathways: novel insights to the immune system. *BMC Genomics* 2013; 14: S10.

ژن‌های (1 and 2) SRPK، درگیری در آبشارهای سلولی پیش‌برنده آپوتوز، پردازش متناوب (alternative splicing) mRNA، پایداری RNA و احتمالاً تولید اتو آنتی‌بادی‌هایی که بر علیه فاکتورهای پردازشگر عمل می‌کنند، می‌باشد [۴۳]. بنابراین با توجه به ماهیت خودایمنی بیماری ام.اس، می‌توان این ژن را مورد بررسی بیش‌تری قرار داد. آخرین نشانگر کاندید، ژن PPP1CA (PP1) است که افزایش بیان یافته است. این ژن برای تقسیم سلولی ضروری است به طوری‌که مهار چرخه سلولی و/یا آپوتوز را تسهیل می‌کند [۴۴]. فعالیت ژن‌های PP1 و PP2A در لنفوسیت‌های B و T موش‌های مستعد به خودایمنی مشاهده شده است [۴۵]. با توجه به این نقش، می‌توان این ژن را نیز مورد مطالعه بیش‌تری قرار داد. به طور کلی، این مطالعه نشان داد که آنالیز داده‌های بیان ژن بر اساس شبکه برهمکنش پروتئین-پروتئین می‌تواند دیدگاه بیش‌تری را برای تعیین ژن‌های درگیر و هم‌چنین بیماری‌زایی در بیماری ام.اس در مرحله post-genomic را فراهم کند. به طور خلاصه، با ادغام ژن‌های متفاوت بیان شده در مایع مغزی نخاعی بیماران ام.اس با داده‌های برهمکنش پروتئین-پروتئین، شبکه مربوطه استخراج شد. پارامترهای اندازه‌گیری‌کننده مرکزیت ژن‌ها در شبکه که نشان‌دهنده اهمیت این ژن‌ها می‌باشند، محاسبه شدند. پنج نشانگر کاندید که درجه رأس بالا (هاب)، betweenness بالا (گلوگاه) و closeness centrality بالا (کم‌ترین فاصله با بقیه رأس‌ها در شبکه) را دارا بودند، معرفی شدند. این ژن‌های کاندید معرفی شده با رویکرد سیستم بیولوژی، می‌توانند مورد تأیید بیش‌تر آزمایشگاهی قرار بگیرند و تحقیقات در زمینه بیماری ام.اس را افزایش دهند. اگر چه ادغام داده‌های بیان ژن و برهمکنش پروتئین-پروتئین، ابزار بسیار مؤثری برای اولویت‌بندی ژن‌های درگیر در بیماری محسوب می‌شود، اما به دلیل کامل نبودن پایگاه داده‌های ثبت‌کننده برهمکنش‌های پروتئین-پروتئین، فقدان بسیاری از برهمکنش‌های پروتئینی محتمل است و این امر محدودکننده مطالعه برای شناسایی همه

- [30] Ray M, Ruan J, Zhang W. Variations in the transcriptome of Alzheimer's disease reveal molecular networks involved in cardiovascular diseases. *Genome Biol* 2008; 9: R148.
- [31] Baranzini SE, Galwey NW, Wang J, Khankhanian P, Lindberg R, Pelletier D, et al. Pathway and network-based analysis of genome-wide association studies in multiple sclerosis. *Hum Mol Genet* 2009; 18: 2078-2090.
- [32] Zhang F, Shi Y, Wang L, Sriram S. Role of HDAC3 on p53 expression and apoptosis in T cells of patients with multiple sclerosis. *PLoS One* 2011; 6: e16795.
- [33] Sharma A, Gulbahce N, Pevzner SJ, Menche J, Ladenvall C, Folkersen L, et al. Network-based analysis of genome wide association data provides novel candidate genes for lipid and lipoprotein traits. *Mol Cell Proteomics* 2013; 12: 3398-3408.
- [34] Yu FD, Yang SY, Li YY, Hu W. Co-expression network with protein-protein interaction and transcription regulation in malaria parasite *Plasmodium falciparum*. *Gene* 2013; 518: 7-16.
- [35] Feldman I, Rzhetsky A, Vitkup D. Network properties of genes harboring inherited disease mutations. *Proc Natl Acad Sci U S A* 2008; 105: 4323-4328.
- [36] Zhu M, Gao L, Li X, Liu Z, Xu C, Yan Y, et al. The analysis of the drug-targets based on the topological properties in the human protein-protein interaction network. *J Drug Target* 2009; 17: 524-532.
- [37] Zali H, Zamanian-Azodi M, Tavirani MR, Baghban AA-z. Protein Drug Targets of *Lavandula angustifolia* on treatment of Rat Alzheimer's Disease. *Iran J Pharm Res* 2015; 14: 291-302.
- [38] Stelzl U, Worm U, Lalowski M, Haenig C, Brembeck FH, Goehler H, et al. A human protein-protein interaction network: a resource for annotating the proteome. *Cell* 2005; 122: 957-968.
- [39] D'Ambrosi N, Rossi S, Gerbino V, Cozzolino M. Rac1 at the crossroad of actin dynamics and neuroinflammation in Amyotrophic Lateral Sclerosis. *Front Cell Neurosci* 2014; 8: 279.
- [40] Trakala M, Arias CF, García MI, Moreno-Ortiz MC, Tsilingiri K, Fernández PJ, et al. Regulation of macrophage activation and septic shock susceptibility via p21 (WAF1/CIP1). *Eur J Immunol* 2009; 39: 810-819.
- [41] Arias CF, Ballesteros-Tato A, García MI, Martín-Caballero J, Flores JM, Martínez-A C, Balomenos D. p21CIP1/WAF1 controls proliferation of activated/memory T cells and affects homeostasis and memory T cell responses. *J Immunol* 2007; 178: 2296-2306.
- [42] Li J, Jiang D. The role of epigenomics in the neurodegeneration of ataxia-telangiectasia. *Epigenomics* 2015; 7: 137-141.
- [43] Kamachi M, Le TM, Kim SJ, Geiger ME, Anderson P, Utz PJ. Human autoimmune sera as molecular probes for the identification of an autoantigen kinase signaling pathway. *J Exp Med* 2002; 196: 1213-1226.
- [44] Flores-Delgado G, Liu CW, Sposto R, Berndt N. A limited screen for protein interactions reveals new roles for protein phosphatase 1 in cell cycle control and apoptosis. *J Proteome Res* 2007; 6: 1165-1175.
- [45] Zhu T, Matsuzawa S-i, Mizuno Y, Kikuchi K. Alterations in activities of protein phosphatases PP1 and PP2A in and B lymphocytes of autoimmune MRL/MpJ-lpr/lpr mice. *J Biochem* 1993; 114: 50-54.
- [14] Lee SA, Tsao TT, Yang KC, Lin H, Kuo YL, Hsu CH, et al. Construction and analysis of the protein-protein interaction networks for schizophrenia, bipolar disorder, and major depression. *BMC Bioinformatics* 2011; 12: S20.
- [15] Safari-Alighiarloo N, Taghizadeh M, Tabatabaei M, Shahsavari S, Namaki S, Khodakarim S, et al. Topological analysis of blood differentially expressed genes in protein-protein interaction network in type 1 diabetes. *Koomesh* 2016; 18: 86-94.
- [16] Rezaei Tavirani M, Vafaei R, Mansouri V, Zamanian Azodi M. Protein-protein interaction network analysis of major depression disorder via proteomic approach from cerebrospinal fluid sample. *Koomesh J* 2017; 19: 577-583.
- [17] Zamanian Azodi M, Rezaei Tavirani M, Peyvandi M, Safaei A, Rezaei-Tavirani M, Mahboubi M, et al. Protein-protein interaction network analysis of human fibroblast cells treated with ethanol. *Koomesh J* 2016; 18: 35-41.
- [18] Hindumathi V, Kranthi T, Rao S, Manimaran P. The prediction of candidate genes for cervix related cancer through gene ontology and graph theoretical approach. *Mol Biosyst* 2014; 10: 1450-1460.
- [19] Safari-Alighiarloo N, Rezaei-Tavirani M, Taghizadeh M, Tabatabaei SM, Namaki S. Network-based analysis of differentially expressed genes in cerebrospinal fluid (CSF) and blood reveals new candidate genes for multiple sclerosis. *Peer J* 2016; 4: e2775.
- [20] Yu H, Kim PM, Sprecher E, Trifonov V, Gerstein M. The importance of bottlenecks in protein networks: correlation with gene essentiality and expression dynamics. *PLoS Comput Biol* 2007; 3: e59.
- [21] Gautier L, Cope L, Bolstad BM, Irizarry RA. affy—analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 2004; 20: 307-315.
- [22] Gentleman R, Carey V, Huber W, Hahne F. Gensefilter: Methods for filtering genes from microarray experiments. *R Package Version* 2011; 1.
- [23] Smyth GK. Limma: linear models for microarray data. *Bioinformatics and computational biology solutions using R and Bioconductor*: Springer; 2005. p. 397-420.
- [24] Kerrien S, Aranda B, Breuza L, Bridge A, Broackes-Carter F, Chen C, et al. The IntAct molecular interaction database in 2012. *Nucleic Acids Res* 2012; 40: D841-846.
- [25] Ceol A, Aryamontri AC, Licata L, Peluso D, Briganti L, Perfetto L, et al. MINT, the molecular interaction database: 2009 update. *Nucleic Acids Res* 2010; 38: D532-539.
- [26] Xenarios I, Salwinski L, Duan XJ, Higney P, Kim SM, Eisenberg D. DIP, the database of interacting proteins: a research tool for studying cellular networks of protein interactions. *Nucleic Acids Res* 2002; 30: 303-305.
- [27] Lynn DJ, Winsor GL, Chan C, Richard N, Laird MR, Barsky A, et al. InnateDB: facilitating systems-level analyses of the mammalian innate immune response. *Mol Syst Biol* 2008; 4: 218.
- [28] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; 13: 2498-2504.
- [29] Zhang A. Protein interaction networks: computational analysis: Cambridge University Press; 2009.

Analysis of protein-protein interactions network based on differentially expressed genes in cerebrospinal fluid for multiple sclerosis

Nahid Safari-Alighiarloo (Ph.D student)¹, Mostafa Rezaei-Tavirani (Ph.D)^{*1}, Mohammad Taghizadeh (Ph.D)², Seyyed Mohammad Tabatabaei (Ph.D student)³, Saeed Namaki (Ph.D)⁴

1 - Proteomics Research Center, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

2 - Bioinformatics Department, Institute of Biochemistry and Biophysics, Tehran University, Tehran, Iran

3 - Medical Informatics Department, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

4- Department of Immunology, Faculty of Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran

(Received: 25 Oct 2016; Accepted: 28 Oct 2017)

Introduction: Multiple sclerosis (MS) is the most common central nervous system (CNS) demyelinating disease. New methods are substantive in order to have a better understanding its pathogenesis and also identifying new therapeutic targets for the disease. The aim of this study was to use network biology approach to identify potential key markers involved in MS pathogenesis.

Materials and Methods: In this study, gene expression profile of cerebrospinal fluid (CSF) samples from 26 newly diagnosed MS patients and 18 controls, which obtained from Array Express Database, were analyzed. Differentially expressed genes (DEG) were integrated with protein-protein interaction (PPI) data to construct PPI sub network. The sub network underwent further topological analysis by Cytoscape software.

Results: Gene expression profile analysis unraveled 3062 differential expressed genes (FDR < 0.05) in CSF of MS patients compared to control in which 1080 genes were up regulated and 1981 genes were down regulated. Integrating of DEGs with PPI data lead to construction the sub network with 1440 nodes and 3500 edges. After topological analysis of the network, five new genes with high centrality measures identified as candidate markers. These markers involved in biological processes such as the regulation of cytoskeletal dynamics, cell cycle and spliceosome some. By literature survey, it has been confirmed their potential contributions in MS pathogenesis.

Conclusion: Therefore, network-based analysis could identify new markers which can be further explored as potential therapeutic targets for MS.

Keywords: Multiple Sclerosis , Transcriptome, Protein Interaction Maps, Topological Features.

* Corresponding author. Tel: +98 21-22714248
tavirany@yahoo.com