

## بررسی ویژگی‌های توپولوژیکی ژن‌های تغییر بیان یافته خون در شبکه برهم کنش پروتئین - پروتئین برای بیماری دیابت نوع ۱

ناهد صفری علی قیارلو<sup>۱</sup> (Ph.D student)، محمد تقی زاده<sup>۲</sup> (Ph.D student)، سید محمد طباطبائی<sup>۳</sup> (Ph.D student)، سوده شهسواری<sup>۴</sup> (Ph.D student)، سعید نمکی<sup>۵</sup> (Ph.D)، سهیلا خداکریم<sup>۶</sup> (Ph.D)، مصطفی رضایی طاویرانی<sup>۱\*</sup> (Ph.D)، علی رشیدی پور<sup>۷</sup> (Ph.D)

۱- مرکز تحقیقات پروتئومیکس، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران  
۲- گروه بیوانفورماتیک، مرکز بیوشیمی و بیوفیزیک (IBB)، دانشگاه تهران، تهران، ایران  
۳- گروه انفورماتیک پزشکی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران  
۴- گروه آمار زیستی، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران  
۵- گروه ایمونولوژی، دانشکده پزشکی، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران  
۶- گروه اپیدمیولوژی، دانشکده بهداشت، دانشگاه علوم پزشکی شهید بهشتی، تهران، ایران  
۷- مرکز تحقیقات و گروه فیزیولوژی، دانشکده پزشکی، دانشگاه علوم پزشکی سمنان، سمنان، ایران

### چکیده

سابقه و هدف: بیماری دیابت نوع ۱ در نتیجه تخریب خودایمنی سلول‌های بتا تولیدکننده انسولین در پانکراس ایجاد می‌شود. برای گسترش درمان بیماری، شناسایی نشانگرهای درمانی کارآمد مورد نیاز است. روش‌های سیستم بیولوژی سبب فهم بهتری از عوامل عملکردی در بیماری می‌گردند. هدف این مطالعه، جستجوی نشانگرهای کاندید در این بیماری با بررسی ویژگی‌های توپولوژیکی زیرشبکه حاصل از تلفیق داده‌های بیان ژن و برهمکنش پروتئین-پروتئین بود.

مواد و روش‌ها: در این مطالعه، پروفایل بیان ژن سلول‌های تک هسته‌ای خون محیطی بیماران تازه مبتلا شده به دیابت از پایگاه داده Gene Expression Omnibus (GEO) دانلود و ژن‌های تغییر بیان یافته مشخص گردیدند. سپس این ژن‌ها روی شبکه برهم کنش پروتئین-پروتئین منطبق شدند. پنج ویژگی توپولوژیکی توسط نرم‌افزار Cytoscape آنالیز شد. در نهایت از سه پارامتر مهم درجه رأس، Betweenness و Closeness centrality برای معرفی نشانگرهای کاندید استفاده گردید.

یافته‌ها: با انجام آنالیز آماری، ۲۴۶۷ ژن تغییر بیان یافته مشخص گردیدند که ۱۰۲۴ افزایش و ۱۴۴۳ کاهش بیان داشتند. با قرار دادن این ژن‌ها در شبکه، زیر شبکه‌ای با ۹۴۹ نود و ۱۷۷۶ یال ساخته شد. با آنالیز توپولوژیکی زیرشبکه، پروتئین‌های با درجه رأس بالا (Hub) و پروتئین‌های با Betweenness بالا (گلوگاه یا Bottleneck) شناخته شدند. ۹ پروتئین Hub-bottleneck که مرکزیت بالا (Closeness centrality) در شبکه داشتند، به عنوان نشانگرهای کاندید معرفی شدند.

نتیجه‌گیری: نشانگرهای کاندید به دست آمده از شبکه با نگاه سیستماتیک، می‌توانند به عنوان نشانگرهای تشخیصی جدید و اهداف دارویی مورد بررسی بیش تر قرار گیرند.

واژه‌های کلیدی: دیابت شیرین نوع ۱، بیان ژن، شبکه برهم کنش پروتئین-پروتئین

دیابت نوع ۱، بیماری خود ایمنی مزمن است که در افراد

مقدمه

است [۱۲،۱۱]. در این راستا، آنالیزهای توپولوژیکی (ظاهری) شبکه برهمکنش پروتئین-پروتئین در شناسایی اهمیت زیستی پروتئین‌ها کاربرد داشته است [۱۳-۱۵]. ویژگی‌های توپولوژیکی مرکزی نظیر درجه رأس، Betweenness، Closeness centrality برای شناسایی پروتئین‌های عمل‌کردی مهم در شبکه مفید هستند. در شبکه برهمکنش پروتئین-پروتئین، نودها با درجه رأس بالا را Hub و نودها با Betweenness بالا را گلوگاه (Bottleneck) می‌نامند که نقش بسیار اساسی در شبکه دارند [۱۶].

در این مطالعه، ژن‌های متفاوت بیان شده در سلول‌های تک هسته‌ای خون محیطی کودکان دیابتی، با آنالیزهای آماری مشخص و با شبکه برهمکنش پروتئین-پروتئین تلفیق شدند. هدف از این مطالعه شناسایی پروتئین‌های مهم درگیر در بیماری بود. آنالیزهای توپولوژیکی انجام شد و پروتئین‌های با ارتباطات زیاد (Hub) و گلوگاه (Bottleneck) که در مرکز شبکه بودند، معرفی شدند.

## مواد و روش‌ها

تهیه پروفایل بیان ژن. داده‌های خام بیان ژن میکرواری با شناسه GSE9006 از پایگاه داده Gene Expression Omnibus (GEO) (<http://www.ncbi.nlm.nih.gov/geo/>) دانلود شدند. داده‌های GSE9006 توسط Ellen C همکارانش در سال ۲۰۰۷ منتشر شد [۶]. نمونه‌های خونی از ۴۳ بیمار تازه مبتلا شده به دیابت نوع ۱ (۲۸ بیمار با نژاد Caucasian) و ۷ بیمار با نژاد (Hispanic) ۳ بیمار با نژاد (African-American) و ۴ بیمار با نژاد مختلط و ۲۴ نمونه کنترل (۱۱ بیمار با نژاد Caucasian)، ۷ بیمار با نژاد (Hispanic) و ۶ بیمار با نژاد مختلط تهیه شده است. میانگین سن ۹/۵ در بیمار و ۱۰/۵ در سالم و نسبت جنسیت (زن به مرد) ۲۵/۱۸ در بیمار و ۱۴/۱۰ گزارش شده است. نمونه‌های خون در لوله‌های حاوی EDTA جمع‌آوری شده است. سلول‌های تک هسته‌ای خون محیطی در بافر حاوی بتا مرکاپتواتانول تجزیه و در ۸۰- درجه سانتی‌گراد نگهداری

مستعد ژنتیکی، توسط عوامل محیطی تسریع می‌گردد. سلول‌های ایمنی بدن به سلول‌های بتا در جزایر لانگرهانس پانکراس حمله کرده و به آن‌ها آسیب می‌رسانند. در نتیجه تولید انسولین در این سلول‌ها کاهش و در نهایت حذف می‌شود [۱،۲]. آنتی‌بادی‌های تولیدشونده بر علیه اجزای سلول‌های بتا، در سرم افراد پیش‌دیابتی نشان‌دهنده فرایند خودایمنی پیش‌رونده در طی دوره پیش‌کلینیکی بدون علائم بیماری است [۳]. بیماری زمانی از نظر کلینیکی آشکار می‌شود که حدود ۹۰ درصد سلول‌های بتا تخریب شده‌اند [۴].

بیش‌تر اطلاعات ما در مورد مکانیسم‌های تخریب انتخابی سلول‌های بتا، از مدل‌های حیوانی به دست می‌آیند [۵]. اگرچه این روش‌ها ضرورتاً نمی‌توانند موقعیت انسانی را منعکس کنند و سؤال‌های زیادی بی‌پاسخ مانده است. پروفایل بیان کل ژنوم توسط میکرواری یک تکنولوژی بسیار کارآمد در زمینه مطالعات ژنتیکی بیماری‌ها است. اخیراً این تکنیک برای شناسایی ژن‌ها و مسیرهای دخیل در دیابت نوع ۱ با استفاده از سلول‌های تک هسته‌ای خون محیطی (PBMC) مورد استفاده قرار گرفته است [۷،۶]. در این مطالعات این‌طور فرض شده است که سلول‌های ایمنی مهاجم به سلول‌های بتا پانکراس در تعادل با سلول‌های ایمنی در حال گردش هستند و ممکن است در سلول‌های تک هسته‌ای خون محیطی یافت شوند [۹،۸]. علاوه بر این، بی‌نظمی‌های متابولیکی مرتبط با دیابت نیز، بالقوه سلول‌های بدن را تحت تأثیر قرار می‌دهند که این تغییرات ممکن است در پروفایل بیان ژن سلول‌های تک هسته‌ای خون محیطی، منعکس گردند [۶].

آنالیزهای بر اساس شبکه برهمکنش پروتئین-پروتئین برای شناسایی ارتباطات بین مولکول‌های پروتئینی استفاده می‌شود، که از این طریق می‌توان دیدگاه کلی از این ارتباطات در بستر بیوشیمیایی، انتقال سیگنال و شبکه‌های بیومولکولی به دست آورد [۱۰]. در سال‌های اخیر تلفیق داده‌های بیان ژن در مقیاس بالا با شبکه برهمکنش پروتئین-پروتئین، روش نوید بخشی برای به دست آوردن بستر زیستی معنی‌دار از ارتباطات عمل‌کردی ژن‌های متفاوت بیان شده، فراهم کرده

دانلود و با یکدیگر ادغام شدند [۲۱-۲۳]. سپس با استفاده از نرم افزار Cytoscape [۲۴]، ژن‌هایی که تفاوت بیان داشتند روی شبکه برهمکنش پروتئین-پروتئین منطبق و زیرشبکه حاصل استخراج شد. پنج ویژگی توپولوژیکی، درجه رأس (closeness centrality, betweenness, degree)، کوتاه‌ترین طول مسیر (Average shortest pathlength) و ضریب خوشه‌بندی (clustering coefficient) توسط نرم‌افزار Cytoscape آنالیز شدند. دو ویژگی توپولوژیکی درجه رأس و betweenness به منظور شناسایی پروتئین‌های مهم مورد بررسی بیش‌تر قرار گرفتند. درجه رأس (Ki) نشان‌دهنده تعداد اتصالات یک نود  $i$  می‌باشد. لازم به توضیح است که در شبکه برهمکنش پروتئین-پروتئین، نود اشاره به یک پروتئین و یال اشاره به اتصال بین دو پروتئین دارد. Betweenness به صورت زیر محاسبه می‌شود:

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

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

برای تعیین پروتئین‌هایی که درجه رأس بالا دارند، از روش زیر استفاده شد. ابتدا میانگین و انحراف معیار درجه رأس همه پروتئین‌ها محاسبه و طبق فرمول Cut-off  $(M+2*S)$ ، حد آستانه‌ای تعریف شد. پروتئین‌هایی که درجه رأسشان از این حد آستانه بالاتر بود به عنوان پروتئین‌های hub معرفی شدند. برای معرفی پروتئین‌های گلوگاه (bottleneck)، پروتئین‌هایی که ۱۰ درصد بالای betweenness را داشتند، در نظر گرفته شدند.

## نتایج

تعیین ژن‌های متفاوت بیان شده، در این مطالعه پروفایل بیان ژن سلول‌های تک هسته‌ای خون محیطی حاصل از میکروآرای بیمارانی دیابت نوع ۱ مورد آنالیز قرار گرفت. علاوه بر تعیین ژن‌هایی که تفاوت بیان دارند، تعیین افزایش یا کاهش بیان آن‌ها نیز برای تفسیر زیستی حائز اهمیت است. نتایج

شده است. RNA با استفاده از کیت‌های حاوی RNeasy استخراج شده است. غلظت RNA با استفاده از بیوآنالیزور Agilent ۲۱۰۰ اندازه‌گیری شده است. cDNA دورشته‌ای حاوی توالی پروموتور T7-dT [۲۴] با استفاده از cDNA synthesis kit GeneChip one-cycle، از ۲-۵ میکروگرم RNA تولید شده است. این cDNA تولیدی به عنوان الگو برای رونویسی *in vitro* با برجسب‌های بیوتین با استفاده از کیت labeling GeneChip IVT مورد استفاده قرار گرفته است. cRNA بیواتینیل‌شده با استفاده از محلول‌های پاک‌سازی تخلیص شده‌اند و با ریزارایه‌های U133A و U133B شرکت افی متریکس هیبرید شده‌اند. ریزارایه‌ها با استفاده از اسکنر laser confocal اسکن شده‌اند. نمونه‌هایی که کیفیت cRNA ضعیف داشتند، از آنالیز حذف شده‌اند.

انتخاب ژن‌های متفاوت بیان شده در دو گروه. داده‌های خام بیان mRNA که توسط ریزارایه‌های (HG\_U133A and B (HG\_U133B) شرکت افی متریکس تهیه شده بودند، با استفاده از الگوریتم (RMA) تصحیح زمینه و نرمال شدند [۱۷]. از پکیج "affy" که از وبسایت (<https://www.bioconductor.org>) تهیه شده بود، برای این منظور استفاده گردید [۱۸]. ژن‌هایی که تفاوت بیان داشتند، با تست آماری ANOVA [۱۹] با  $p\text{-value} < 0.05$  تعیین شدند که در این تست اثر فاکتورهای سن و جنس در آنالیز ANOVA لحاظ شدند. برای تعیین Fold change از روش رگرسیون خطی استفاده شد که در آن متغیرهای دوتایی (بیمار و سالم) به عنوان متغیرهای وابسته و تفاوت بیان ژن، سن و جنس به عنوان متغیرهای مستقل در نظر گرفته شدند [۲۰]. جهت ضرایب رگرسیون نشان‌دهنده افزایش یا کاهش بیان ژن در محاسبات Fold change است. همه آنالیزهای فوق در محیط برنامه‌نویسی R انجام شد.

ادغام ژن‌های متفاوت بیان شده با داده‌های برهمکنش پروتئین-پروتئین و ساخت و آنالیز توپولوژیکی زیرشبکه مربوطه، برای تهیه شبکه برهمکنش پروتئین-پروتئین جامع، داده‌های برهمکنش از سه پایگاه داده IntAct، MINT و DIP

مشاهده است. همان‌طور که در شکل ۱ مشاهده می‌شود، بیش‌تر پروتئین‌ها ضرایب خوشه‌بندی ایشان در فواصل (۰/۲) - ۰) پراکنده شده‌اند که نشان‌دهنده تمایل نسبتاً پایین پروتئین‌ها به خوشه‌بندی است. کوتاه‌ترین طول مسیر دلالت بر ارتباطات نزدیک پروتئین‌ها دارد. دامنه کوتاه‌ترین طول مسیر پروتئین‌های زیرشبکه بین ۵-۳ است.

پروتئین‌های Hub و Bottleneck طبق روش ذکر شده در قسمت روش‌ها، تعیین شدند که به ترتیب ۳۲ و ۴۸ پروتئین به عنوان Hub و Bottleneck معرفی شدند. خلاصه نتایج در جدول ۲ ذکر شده است. هم‌چنین در شکل شماره ۲، نمایی کلی از زیرشبکه حاصل و پروتئین‌های Hub و Bottleneck قابل مشاهده است.

معرفی پروتئین‌های کاندید بر اساس ویژگی‌های توپولوژیکی. برای شناسایی پروتئین‌های کاندید از سه ویژگی توپولوژیکی مهم پروتئین‌ها در شبکه استفاده شد. ابتدا پروتئین‌هایی که هم Hub و هم Bottleneck (Hub-bottleneck) بودند، مشخص شده‌اند. سپس این پروتئین‌ها بر اساس پارامتر Closeness centrality به صورت کاهش‌ی از بالا به پایین مرتب شدند. سپس ۱۰ درصد بالای پروتئین‌های واجد شرایط انتخاب شدند. در مرحله بعد عمل‌کرد پروتئین‌های کاندید در مقالات جستجو گردید و با توجه به سطح بیان ژن‌های مربوطه به صورت افزایشی یا کاهش‌ی، ارتباط کاندیدها با بیماری دیابت نوع ۱ مورد بحث قرار گرفت. جدول ۳، پروتئین‌های کاندید و میزان Closeness centrality آن‌ها را نشان می‌دهد.

جدول ۲. تعیین حد آستانه برای hub، تعداد hubs و bottlenecks در

زیرشبکه برهمکنش پروتئین-پروتئین

Parameters	Values
Mean (M)	3.7
Standard Deviation (SD)	6.07
Cut-off (M+2*S)	15.7
Number of hubs	32
Number of bottlenecks	48

آنالیز آماری ANOVA نشان داد که ۲۴۶۷ ژن تفاوت بیان داشتند ( $p < 0.05$ ) که ۱۰۲۴ ژن افزایش بیان و ۱۴۴۳ ژن کاهش بیان داشتند.

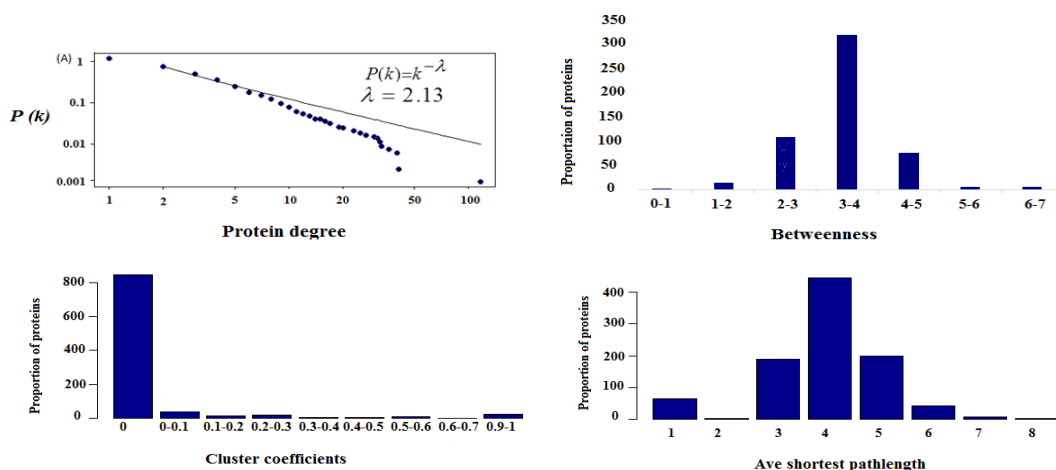
آنالیز ویژگی‌های توپولوژیکی زیرشبکه برهمکنش پروتئین-پروتئین. ژن‌های متفاوت بیان شده روی شبکه برهمکنش پروتئین-پروتئین منطبق شدند که پس از حذف پروتئین‌هایی که برهمکنش با سایر پروتئین‌ها در شبکه نداشتند، زیرشبکه حاصل شامل ۹۴۹ نود و ۱۷۷۶ یال بود. پنج ویژگی توپولوژیکی محاسبه و نتایج در جدول ۱ آورده شده‌اند.

جدول ۱. ویژگی‌های توپولوژیکی زیرشبکه برهمکنش پروتئین-پروتئین

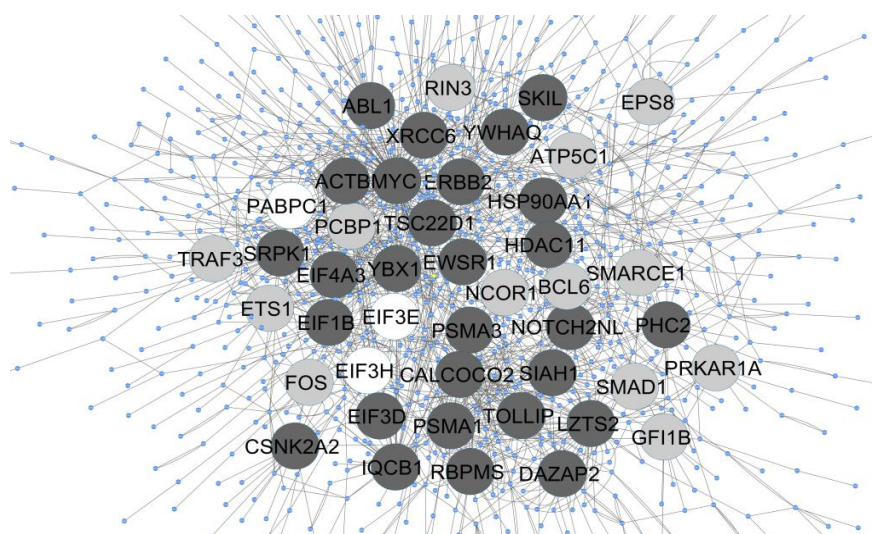
Topological parameters	Values
Number of nodes	949
Number of edges	1776
Average degree	3.75
Average betweenness	0.011
Average closeness centrality	0.268
Average shortest pathlength	4.40
Average clustering coefficient	0.048

در شبکه‌های برهمکنش پروتئین-پروتئین، درجه رأس پروتئین‌ها از قانون Power-law ( $P(k) = k^{-\gamma}$ ) پیروی می‌کنند. طبق این قانون، همان‌طور که در شکل ۱ سمت چپ مشاهده می‌شود، بیش‌تر پروتئین‌ها درجه رأس پایین و تعداد کم‌تری درجه رأس بالا دارند. پارامتر Betweenness تعداد کوتاه‌ترین مسیرهایی که از یک نود در شبکه می‌گذرد را اندازه می‌گیرد. این اندازه‌گیری ممکن است میزان سیگنال‌هایی که از یک نود در شبکه می‌گذرد را منعکس کند.

Closeness centrality یک نود، نشان‌دهنده میزان فاصله یک نود با بقیه نودها در شبکه است که هر چه این فاصله کم‌تر باشد، نود مرکزیت بیش‌تری در شبکه دارد. ضریب خوشه‌بندی، میزان اتصالات همسایه‌های یک نود را اندازه می‌گیرد. ضریب خوشه‌بندی بالای یک نود، نشان‌دهنده چگالی بالایی از ارتباطات شبکه‌اش است. ضرایب خوشه‌بندی در فواصل ۰/۱ اندازه‌گیری شده‌اند و نسبت حضور پروتئین‌ها در فواصل به صورت نمودار ستونی قابل



شکل ۱. ویژگی‌های توپولوژیکی زیرشبکه برهمکنش پروتئین-پروتئین. a. درجه رأس b. Betweenness c. ضریب خوشه بندی d. کوتاهترین طول مسیر



شکل ۲. شبکه برهمکنش پروتئین-پروتئین. پروتئین‌های Hub، Hub-bottleneck و Bottleneck به ترتیب به رنگ‌های سفید، خاکستری تیره و خاکستری روشن دیده می‌شوند.

جدول ۳. Closeness centrality مربوط به پروتئین‌های Hub - bottleneck

Symbol	Full name	Closeness centrality
MYC	Myc proto-oncogene protein	0.358
YBX1	Y-box binding protein 1	0.317
XRCC6	DNA non-homologous end-joining repair	0.310
PSMA1	Proteasome subunit alpha 1	0.306
CALCOCO2	Calcium binding and coiled-coil domain 2	0.301
PSMA3	Proteasome subunit alpha 3	0.298
YWHAQ	14-3-3 protein theta	0.298
ACTB	Actin, beta	0.296
ERBB2	Receptor tyrosine-protein kinase erbB-2	0.294

تاکنون درمان قطعی برای آن نیست [۲۵]. برای پیشبرد اهداف درمان برای این بیماری، شناسایی مسیرهای درگیر در بیماری و نشانگرهای درمانی مورد نیاز هستند. روش‌های سیستم

## بحث و نتیجه‌گیری

دیابت نوع ۱ به دلیل تخریب سلول‌های بتا ترشح‌کننده انسولین در پانکراس به واسطه سلول‌های ایمنی رخ می‌دهد و

است [۳۳]. موش ترانسژنیک که در آن ژن MYC در سلول‌های بتا پانکراس افزایش بیان یافته است، سبب ایجاد دیابت نوزادی می‌گردد که در آن هایپرپلازی جزیره، افزایش آپوپتوز و کاهش بیان ژن انسولین، از نشانه‌های بیماری هستند [۳۴]. دومین نشانگر، پروتئین YBX1 است. سطح بیان ژن YBX1 که در پردازش متناوب (alternative splicing mRNA) نقش دارد، کاهش یافته است. این پروتئین نقش مهمی در پایداری mRNA سیتوکین‌ها دارد که در بسیاری از بیماری‌های خودایمن نقش دارند. نتایج مطالعات در موش که تعادل پایداری mRNA سیتوکین‌ها تغییر یافته است، پیشنهاد می‌کند که عدم تعادل بین پایداری و تخریب mRNA سیتوکین‌های التهابی می‌تواند نشان‌دهنده مکانیسم پایه‌ای منجرشونده به بیماری‌های خودایمنی گردد [۳۵]. علاوه بر این شواهد نشان می‌دهد، ژن YBX1 به عنوان تنظیم‌کننده ژن PTP1B است که این ژن، تنظیم‌کننده مهمی برای مسیر سیگنالینگ به واسطه انسولین و سیتوکین می‌باشد [۳۶]. بنابراین با توجه به نقش این ژن در پایداری سیتوکین‌های التهابی و تاثیر روی ژن درگیر در مسیر سیگنالینگ وابسته به انسولین، می‌توان این ژن را به عنوان نشانگر کاندید معرفی کرد. سومین نشانگر، پروتئین XRCC6 است. سطح بیان ژن XRCC6 که در DNA non-homologous end joining (NHEJ) نقش دارد، در این مطالعه کاهش یافته است. مطالعات Khanna و همکارانش نشان داد که نقص پروتئینی XRCC6 نه تنها باعث کاهش ظرفیت تعمیر دو رشته DNA می‌گردد بلکه سبب نقص ایمنی ناشی از ظرفیت نو ترکیبی ناقص در گیرنده‌های سلول‌های ایمنی می‌گردد [۳۷]. نقص گیرنده‌های سلول‌های ایمنی به ویژه T cells نقش اساسی در بیماری‌های خودایمن از جمله دیابت نوع ۱ دارد [۳۸] که این نقش می‌تواند دلیل کاندیدی بودن ژن XRCC6 باشد. زیرواحدهای ۱ و ۳ پروتئین‌های PSMA1/PSMA3 سبب تجزیه پپتیدها به واسطه ATP/ubiquitin می‌گردند. سیستم پروتئوزوم یوبی کوئیتین نقش زیستی قابل توجهی در پردازش آنتی‌ژن و پاسخ ایمنی

بیولوژی با فراهم آوردن میزان زیادی اطلاعات سبب نگاه سیستمیک و کلی به مکانیسم‌های مولکولی پیچیده بیماری‌ها و شناسایی عناصر عمل‌کردی می‌گردند. برای این نیاز، پروفایل بیان کل ژنوم تهیه شده توسط روش‌های مانند میکرواری یا RNA-sequencing اطلاعات بسیار مفیدی را در سطح ژن فراهم می‌کنند که تلفیق این اطلاعات با داده‌های برهمکنش پروتئین-پروتئین در شناسایی مسیرهای درگیر در بیماری نقش بسزایی دارند. در این مطالعه، پروفایل بیان ژن در بیماران تازه مبتلا شده به دیابت از پایگاه داده تهیه و آنالیز گردید. پس از تعیین ژن‌های متفاوت بیان شده، این ژن‌ها روی شبکه برهمکنش پروتئین-پروتئین منطبق شدند و زیرشبکه مربوطه استخراج شد. با آنالیز ویژگی‌های توپولوژیکی زیرشبکه، ۹ پروتئین کاندید معرفی و ارتباط آن‌ها با بیماری توسط مقالات مورد تأیید قرار گرفت.

شبکه‌های برهمکنش پروتئین-پروتئین دیدگاهی به سمت شناخت مکانیسم‌های بیماری‌زایی [۲۷،۲۶]، تعیین اهداف دارویی [۲۹،۲۸] و کشف نشانگرهای جدید [۳۰] بر اساس شبکه را فراهم می‌کنند. آنالیز ویژگی‌های توپولوژیکی پروتئین‌ها برای شناخت ساختار شبکه برهمکنش‌ها بسیار کمک‌کننده است؛ به طوری که می‌توان منجر به شناخت فرایندهای مولکولی درگیر در بیماری شود. به عنوان مثال، پروتئین‌های hub و گلوگاه (bottleneck) برای شناسایی ژن‌های بیماری‌زا اهداف دارویی جدید هدف‌گیری می‌شوند [۳۲،۳۱].

در این مطالعه ۹ پروتئین که ویژگی‌های توپولوژیکی مهمی در شبکه داشتند، به عنوان نشانگر کاندید معرفی شدند. شرح مختصری از عمل‌کرد این پروتئین‌ها و ارتباط آن‌ها با بیماری با جستجو در مقالات فراهم شده است. به عنوان اولین نشانگر، پروتئین MYC، فسفوپروتئین هسته‌ای است که نقش مهمی در چرخه سلولی، پیشرفت آپوپتوز و ترانسفرماسیون سلول دارد. سطح بیان این ژن در این مطالعه افزایش یافته است. این پروتئین پروتئین‌های انکوژن به دلیل نقش داشتن در مسیر آپوپتوز، در بسیاری از بیماری‌های خودایمن گزارش شده

استرسی که توسط سیتوکین التهابی TNF شروع می‌شود، ErbB2/ErbB3 را فعال می‌کند که منجر به سیگنال انسولین ناقص از طریق فسفریلاسیون سرین/ترئونین IRS-1 و IRS-2 با تحریک مسیر تنظیم‌شونده توسط PI3K و مقاومت به انسولین می‌شود. مقاومت به انسولین وضعیت پاتولوژیکی عمومی منجرشونده به دیابت است [۴۶]. بنابراین این ژن به عنوان آخرین نشانگر کاندید معرفی شد.

در نتیجه، این مطالعه نشان می‌دهد که شبکه زیستی به عنوان یک ابزار مؤثر برای شناسایی ژن‌های کاندید بیماری‌های چندعاملی نظیر دیابت نوع ۱ در نظر گرفته می‌شود و ادغام داده‌های بیان ژن و برهمکنش پروتئین-پروتئین برای دستیابی به این منظور ضروری می‌باشد. به طور خلاصه، با تلفیق ژن‌های متفاوت بیان شده در دیابت نوع ۱ با داده‌های برهمکنش پروتئین-پروتئین زیر شبکه برهمکنش‌های پروتئینی ساخته شد. پارامترهای اندازه‌گیری‌کننده مرکزیت پروتئین‌ها در شبکه (درجه رأس، Betweenness و Closeness centrality) که نشان‌دهنده اهمیت این پروتئین‌ها در شبکه می‌باشند، محاسبه شدند. در نتیجه، ژن‌های کاندید برای بیماری دیابت نوع ۱ معرفی شدند که می‌توانند به عنوان نشانگرهای تشخیصی و یا اهداف دارویی مورد تأیید بیش‌تر آزمایشگاهی قرار بگیرند و تحقیقات در زمینه بیماری دیابت نوع ۱ را افزایش دهند.

#### محدودیت این مطالعه:

این مطالعه اهمیت ادغام داده‌های "omics" را برای فهم بهتر مکانیسم‌های مولکولی بیماری‌های چندعامله نظیر دیابت نوع ۱ را نشان می‌دهد 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 برهمکنش‌های موجود بین ژن‌های متفاوت بیان شده در این بیماری است.

دارد که بالقوه می‌تواند در بسیاری از بیماری‌های ایمنی نقش داشته باشد [۳۹]. پردازش ناقص آنتی‌ژن یکی از عوامل اساسی در بیماری‌های خودایمن از جمله دیابت نوع ۱ است [۴۰]. بنابراین این دو ژن نیز به عنوان چهارمین و پنجمین نشانگرهای کاندید معرفی می‌گردند. پروتئین CALCOCO2 به عنوان گیرنده ubiquitin-coated باکتری، نقش مهمی در سیستم ایمنی ذاتی به واسطه ماکروفاژها دارد. بیان این ژن در این مطالعه کاهش یافته است. پروتئین CALCOCO2 به عنوان گیرنده اتوفازی شناخته شده است. اهمیت اتوفازی برای حفظ حجم و عمل‌کرد سلول‌های بتا پانکراس نقش دارد. اتوفازی ناقص با دیابت و پانکراتیت حاد ارتباط دارد [۴۲، ۴۱] که می‌تواند دلیل بر کاندید بودن این ژن به عنوان ششمین نشانگر باشد. هفتمین نشانگر کاندید، YWHAQ، پروتئین آداپتور است که در تنظیم مسیرهای سیگنالینگ زیادی نقش دارد. در این مطالعه سطح بیان این ژن کاهش یافته است. به دلیل نقشی که در مسیر آپوپتوز دارد، می‌تواند به عنوان مانع یا میانجی آپوپتوز بین هسته یا میتوکندری و سیتوپلاسم باشد [۴۳]. با توجه به نقش آپوپتوز در بیماری‌های خودایمنی، این ژن می‌تواند در ایجاد بیماری دیابت نوع ۱ نقش مهمی داشته باشد [۴۴]. هشتمین نشانگر، پروتئین اکتین (ACTB) در حرکت، ساختار و یکپارچگی سلول نقش دارد. در این مطالعه، این ژن کاهش بیان داشته است. تغییرات تعداد کپی در جایگاهی در کروموزوم ۷ که حاوی ژن اکتین می‌شود، ممکن است موجب عدم تنظیم عمل‌کرد اکتین سیتواسکلتی می‌گردد که برای عمل‌کرد T cell و B cell مورد نیاز است. در نتیجه این موضوع ممکن است حساسیت به بیماری‌های خودایمنی نظیر لوپوس را افزایش دهد [۴۵] که می‌تواند به عنوان نشانگر کاندید در دیابت نوع ۱ به عنوان بیماری خودایمن نیز محسوب شود. نهمین نشانگر کاندید، پروتئین ERBB2 عضوی از خانواده گیرنده فاکتور رشد اپیدرمی است که در بسیاری از مسیرهای سیگنالینگ نقش دارد. سطح بیان این ژن در این مطالعه افزایش یافته است. شواهدی وجود دارد که نشان می‌دهد که در شرایط

## تشکر و قدردانی

این مقاله برگرفته از پایان نامه دوره دکترای تخصصی (PhD) خانم ناهید صفری علی قیارلو میباشد. از تمام کسانی که در انجام این تحقیق ما را کمک کردند تشکر و قدردانی می‌شود.

## منابع

- [15] Zali H, Rezaei-Tavirani M. Meningioma protein-protein interaction Network. Arch Iran Med 2014; 17: 262.
- [16] 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.
- [17] Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, Speed TP. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics 2003; 4: 249-264.
- [18] Gautier L, Cope L, Bolstad BM, Irizarry RA. affy—analysis of affymetrix geneChip data at the probe level. Bioinformatics 2004; 20: 307-315.
- [19] Rutherford A. Introducing ANOVA and ANCOVA: a GLM approach: Sage; 2001.
- [20] Pedhazur EJ. Multiple regression in behavioral research: Explanation and prediction. 1997.
- [21] Kerrien S, Aranda B, Breuza L, Bridge A, Broackes-Carter F, Chen C, Duesbury M, Dumousseau M, Feuermann M, Hinz U. The IntAct molecular interaction database in 2012. Nucleic Acids Res 2012; 40: D841-846.
- [22] Ceol A, Aryamontri AC, Licata L, Peluso D, Briganti L, Perfetto L, Castagnoli L, Cesareni G. MINT, the molecular interaction database: 2009 update. Nucleic Acids Res 2010; 38: D532-539.
- [23] Xenarios I, Salwinski L, Duan XJ, Higney P, Kim S-M, 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.
- [24] 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.
- [25] Bergholdt R, Brorsson C, Palleja A, Berchtold LA, Fløyel T, Bang-Berthelsen CH, et al. Identification of novel type 1 diabetes candidate genes by integrating genome-wide association data, protein-protein interactions, and human pancreatic islet gene expression. Diabetes 2012; 61: 954-962.
- [26] 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.
- [27] 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.
- [28] 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.
- [29] Zali H, Zamanian-Azodi M, Rezaei-Tavirani M, Akbarzadeh-Baghban A. Protein drug targets of *lavandula angustifolia* on treatment of Rat Alzheimer's Disease. Iran J Pharm Res 2015; 14: 291.
- [30] Taylor IW, Linding R, Warde-Farley D, Liu Y, Pesquita C, Faria D, et al. Dynamic modularity in protein interaction networks predicts breast cancer outcome. Nature Biotechnol 2009; 27: 199-204.
- [31] Jeong H, Mason SP, Barabási A-L, Oltvai ZN. Lethality and centrality in protein networks. Nature 2001; 411: 41-42.
- [32] Hase T, Tanaka H, Suzuki Y, Nakagawa S, Kitano H. Structure of protein interaction networks and their implications on drug design. PLoS Comput Biol 2009; 5: e1000550.
- [33] Boumpas D, Tsokos G, Mann D, Eleftheriades E, Harris C, Mark G. Increased proto-oncogene expression in peripheral blood lymphocytes from patients with systemic
- [1] Atkinson MA, Eisenbarth GS. Type 1 diabetes: new perspectives on disease pathogenesis and treatment. Lancet 2001; 358: 221-229.
- [2] Van Belle TL, Coppieters KT, Von Herrath MG. Type 1 diabetes: etiology, immunology, and therapeutic strategies. Physiol Rev 2011; 91: 79-118.
- [3] Bach JF. Insulin-dependent diabetes mellitus as an autoimmune disease. Endocr Rev 1994; 15: 516-542.
- [4] Tanaka Y, Asakawa T, Asagiri K, Akiyoshi K, Hikida S, Mizote H. Nesidioblastosis treated successfully by 85% pancreatectomy. Kurume Med J 2004; 51: 99-103.
- [5] Atkinson MA, Leiter EH. The NOD mouse model of type 1 diabetes: as good as it gets? Nature Med 1999; 5: 601-604.
- [6] Kaizer EC, Glaser CL, Chaussabel D, Banchereau J, Pascual V, White PC. Gene expression in peripheral blood mononuclear cells from children with diabetes. J Clin Endocr Metab 2007; 92: 3705-3711.
- [7] Wang X, Jia S, Geoffrey R, Alemzadeh R, Ghosh S, Hessner MJ. Identification of a molecular signature in human type 1 diabetes mellitus using serum and functional genomics. J Immunol 2008; 180: 1929-1937.
- [8] Kent SC, Chen Y, Clemmings SM, Viglietta V, Kenyon NS, Ricordi C, et al. Loss of IL-4 secretion from human type 1a diabetic pancreatic draining lymph node NKT cells. J Immunol 2005; 175: 4458-4464.
- [9] Ott PA, Berner BR, Herzog BA, Guerkov R, Yonkers NL, Durinovic-Bello I, et al. CD28 costimulation enhances the sensitivity of the ELISPOT assay for detection of antigen-specific memory effector CD4 and CD8 cell populations in human diseases. J Immunol Methods 2004; 285: 223-235.
- [10] Wu J, Vallenius T, Ovaska K, Westermarck J, Mäkelä TP, Hautaniemi S. Integrated network analysis platform for protein-protein interactions. Nature Methods 2009; 6: 75-77.
- [11] 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.
- [12] Bapat SA, Krishnan A, Ghanate AD, Kusumbe AP, Kalra RS. Gene expression: protein interaction systems network modeling identifies transformation-associated molecules and pathways in ovarian cancer. Cancer Res 2010; 70: 4809-4819.
- [13] 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.
- [14] Zamanian-Azodi M, Rezaei-Tavirani M, Rahmati-Rad S, Hasanzadeh H, Rezaei-Tavirani M, Seyyedi SS. Protein-Protein Interaction Network could reveal the relationship between the breast and colon cancer. Gastroenterol Hepatol Bed Bench 2015; 8: 215-224.



- [40] Fu Y, Yan G, Shi L, Faustman D. Antigen processing and autoimmunity: evaluation of mRNA abundance and function of HLA-linked genes. *Ann N Y Acad Sci* 1998; 842: 138-155.
- [41] Ebato C, Uchida T, Arakawa M, Komatsu M, Ueno T, Komiya K, et al. Autophagy is important in islet homeostasis and compensatory increase of beta cell mass in response to high-fat diet. *Cell Metab* 2008; 8: 325-332.
- [42] Fujitani Y, Kawamori R, Watada H. The role of autophagy in pancreatic  $\beta$ -cell and diabetes. *Autophagy* 2009; 5: 280-282.
- [43] Li Z, Qiao Z, Zheng W, Ma W. Network Cluster Analysis of Protein-Protein Interaction Network-Identified Biomarker for Type 2 Diabetes. *Diabetes Technol Ther* 2015; 17: 475-481.
- [44] Eguchi K. Apoptosis in autoimmune diseases. *Intern Med* 2001; 40: 275-284.
- [45] Yap T, Adib N, Coman D. Microduplication with autoimmune inflammatory arthritis. *J Mol Genet Med* 2015; 9: 1747-0862.
- [46] Hemi R, Paz K, Wertheim N, Karasik A, Zick Y, Kanety H. Transactivation of ErbB2 and ErbB3 by tumor necrosis factor- $\alpha$  and anisomycin leads to impaired insulin signaling through serine/threonine phosphorylation of IRS proteins. *J Biol Chemistry* 2002; 277: 8961-8969.
- lupus erythematosus and other autoimmune diseases. *Arthritis Rheum* 1986; 29: 755-760.
- [34] Laybutt DR, Weir GC, Kaneto H, Lebet J, Palmiter RD, Sharma A, Bonner-Weir S. Overexpression of c-Myc in  $\beta$ -cells of transgenic mice causes proliferation and apoptosis, downregulation of insulin gene expression, and diabetes. *Diabetes* 2002; 51: 1793-1804.
- [35] Seko Y, Cole S, Kasprzak W, Shapiro BA, Ragheb JA. The role of cytokine mRNA stability in the pathogenesis of autoimmune disease. *Autoimmun Rev* 2006; 5: 299-305.
- [36] Fukada T, Tonks NK. Identification of YB-1 as a regulator of PTP1B expression: implications for regulation of insulin and cytokine signaling. *EMBO J* 2003; 22: 479-493.
- [37] Khanna KK, Jackson SP. DNA double-strand breaks: signaling, repair and the cancer connection. *Nature Genet* 2001; 27: 247-254.
- [38] Zhou J, Kong C, Chen X. The short review on the studies of T cells receptors relate to type 1 diabetes. *Diabetes Metabol Syndr* 2014; 8: 252-254.
- [39] Sjakste T, Paramonova N, Rumba-Rozenfelde I, Trapina I, Sugoka O, Sjakste N. Juvenile idiopathic arthritis subtype-and sex-specific associations with genetic variants in the PSMA6/PSMC6/PSMA3 gene cluster. *Pediatr Neonatol* 2014; 55: 393-403.

# Topological analysis of blood differentially expressed genes in protein-protein interaction network in type 1 diabetes

Nahid Safari-Alighiarloo (Ph.D student)<sup>1</sup>, Mohammad Taghizadeh (Ph.D student)<sup>2</sup>, Seyyed Mohammad Tabatabaei (Ph.D student)<sup>3</sup>, Soodeh Shahsavari (Ph.D student)<sup>4</sup>, Saeed Namaki (Ph.D)<sup>5</sup>, Soheila Khodakarim (Ph.D)<sup>6</sup>, Mostafa Rezaei-Tavirani (Ph.D)<sup>\*1</sup>, Ali Rashidy-pour (Ph.D)<sup>7</sup>

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

*2- Bioinformatics Dept., 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- Biostatistics Department, Faculty of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran*

*5- Dept. of Immunology, Faculty of Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran*

*6-Department of Epidemiology, School Of Public Health, Shahid Beheshti University of Medical Sciences, Tehran, Iran*

*7 – Physiology Research Center and Dept. of Physiology, Faculty of Medicin, Semnan University of Medical Sciences, Semnan, Iran*

(Received: 4 Dec 2015; Accepted: 21 May 2016)

**Introduction:** Type 1 diabetes (T1D) results from autoimmune destruction of insulin-producing beta cells in pancreatic islets of Langerhans. To develop efficient treatments for T1D, it is required to identify suitable therapeutic markers. Systems biology offers approaches to better understanding of functional elements in the disease. Our aim was to investigate larger number of candidate markers in T1D by topological analysis of constructed PPI based on gene expression.

**Materials and Methods:** In this study, gene expression profile of peripheral blood mononuclear cells from newly diagnosed type 1 diabetic children was prepared from Gene Expression Omnibus and analyzed to get differentially expressed genes. Then, these genes were mapped to PPIs data to construct related subnetwork. Five topological features were calculated by Cytoscape software. Finally, degree, betweenness and closeness centrality features were utilized to identify candidate markers.

**Results:** 2467 differentially expression genes were obtained by statistical analyzing of gene expression profile in which 1024 were upregulated and 1443 were downregulated. After mapping these genes on PPI network, there was constructed subnetwork with 949 nodes and 1776 edges. By topological analysis of the subnetwork, we determined high degree nodes (hub) and high betweenness nodes (bottleneck). Then, 9 hub-bottleneck proteins that were more central (high closeness centrality) in the subnetwork were identified and introduced as candidate markers.

**Conclusion:** The obtained markers from network via systemic view can be considered as new diagnostic markers and potential therapeutic targets for T1D.

**Keywords:** Type 1 Diabetes, Gene Expression, Protein-Protein Interaction Network

---

\* Corresponding author. Tel: ++98 21 22714248

tavirany@yahoo.com