

مقایسه‌ی بیان ژن در سلول‌های فیروبلاست در افراد مبتلا به بیماری هانتینگتون و افراد معمولی

مهسا قربانی

mghorbani@ce.sharif.edu

۱۶ بهمن ۱۳۹۵

باشد [۳]. علائم بیماری هانتینگتون می‌تواند دیرتر یا زودتر در بدن انسان ظهور کند. در صورتی که علائم پیش از ۲۰ سالگی ظهور کند، با نام هانتینگتون نوجوانی (JHD) شناخته می‌شود [۴]. اگرچه هانتینگتون بیماری کشنده‌ای نیست، اما ضعف در بخش‌های مختلف بدن می‌تواند موجب مرگ مبتلایان به بیماری بشود. ذات‌الریه و بیماری‌های قلبی شایع‌ترین دلایل مرگ در بیماران مبتلا به هانتینگتون هستند. احتمال وقوع خفگی، بیماری‌های دستگاه گوارشی (مانند سرطان پانکراس)، مرگ ناشی از ضایعات عروقی سیستم عصبی و خودکشی ناشی از افسردگی در این بیماران بیش‌تر از افراد عادی است [۵، ۶]. شیوه‌ی وراثت در این بیماری به صورت اتوزومی غالب است به این معنی که بیماری خود را در فردی که دارای حداقل یک الل جهش‌یافته است، نشان می‌دهد. اگر یکی از والدین دارای ژن جهش‌یافته باشند، با احتمال ۵۰ درصد، فرزند نیز دارای این ژن خواهد بود، اما اگر فردی ژن را از والد دریافت نکند، امکان انتقال به فرزندانش وجود ندارد [۷].

تا به حال اطلاعات دقیقی از عملکرد پروتئین هانتینگتین در زیست‌شناسی مشخص نشده است. از سوی دیگر، در نوروپاتی‌هایی که تحت تاثیر بیماری‌های از بین‌برنده‌ی نوروپاتی هستند، تفاوت قابل توجهی از نظر ریخت‌شناسی، فیزیولوژی و عملکرد نسبت به نوروپاتی‌های معمولی مشاهده شده است [۸]. بنابراین بررسی تفاوت بیان ژن، میان افراد بیمار و افراد عادی می‌تواند به شناخت بهتر تاثیر پروتئین بر فرآیندهای بدن و زمان شروع و نحوه‌ی پیشرفت بیماری کمک کند. مطالعات مختلفی روی سلول‌های مختلف بدن بیماران، مانند سلول‌های خونی و پوستی صورت گرفته است. در این مطالعه قصد داریم به بررسی تفاوت بیان ژن میان سلول‌های فیروبلاست (که از مهم‌ترین انواع سلول‌های سوماتیک هستند) در بدن بیماران مبتلا به هانتینگتون و افراد عادی بپردازیم.

۲- داده‌ها و روش پیش‌پردازش

در این بخش به توضیح داده‌ها و پیش‌پردازش روی آن‌ها می‌پردازیم.

چکیده: بیماری هانتینگتون بیماری‌ای موروثی و ژنتیکی است که بر اثر تکرار سه‌تایی CAG در ژن هانتینگتون اتفاق می‌افتد. پروتئین هانتینگتین جهش‌یافته باعث تغییر مسیرهای زیستی در بدن فرد مبتلا می‌گردد، به طوری که نوروپاتی‌های بدن فرد به مرور از بین رفته و در نهایت به دلیل اختلال در مسیرهای زیستی، مرگ اتفاق می‌افتد. با توجه به عدم شناخت کافی نسبت به بیماری، تا کنون درمان مشخصی برای آن ارایه نشده است. احتمال آن می‌رود که بیماری علاوه بر سیستم عصبی بدن، بخش‌های دیگری را نیز تحت تاثیر قرار بدهد. به همین جهت در این مقاله با استفاده از داده‌های بیان ژن در سلول فیروبلاست افراد مبتلا و افراد عادی، ابتدا ژن‌هایی را که در بدن بیماران نسبت به افراد سالم تغییرات قابل‌توجهی داشته‌اند، استخراج کرده و سپس اهمیت آن‌ها را مسیرها و فرآیندهای زیستی بدن مورد بررسی قرار می‌دهیم. مشخص می‌شود که ژن‌های استخراج‌شده، در مسیرها و فرآیندهای زیستی متفاوتی از جمله چرخه‌ی سلولی و اصلاح دی‌ان‌ای تاثیر به‌سزایی دارند. همچنین تاثیر این ژن‌ها بر لیزوزوم‌ها و هسته‌ی سلول‌های فیروبلاست نشان داده می‌شود.

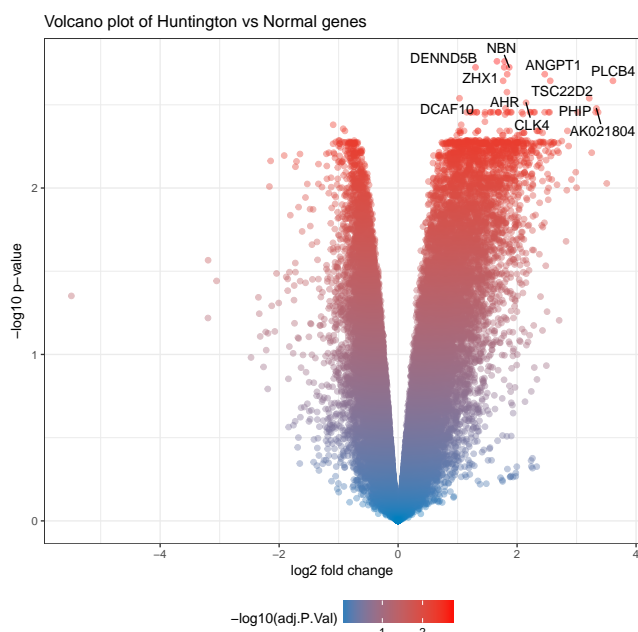
۱- مقدمه

بیماری هانتینگتون (HD) بیماری‌ای نادر و موروثی است که به دلیل اختلال در شبکه‌ی عصبی بدن ایجاد شده و موجب اختلال‌های حرکتی، نقص‌های شناختی و مشکلات روانی می‌شود. بیماری هانتینگتون به دلیل تکرار سه‌تایی CAG در اگزون اول ژن HD روی بازوی کوتاه کروموزوم ۴ ایجاد می‌شود. تعداد تکرارهای توالی ۳ تایی CAG در یک ژن سالم در حدود ۱۰ تا ۲۹ بار است، در حالی‌که این تعداد در یک فرد بیمار در حدود ۳۶ تا ۱۲۱ بار است [۱]. ترجمه‌ی توالی افزایش‌یافته‌ی CAG به پروتئین هانتینگتین (HTT)، باعث ایجاد توالی پلی‌گلوتامین در این پروتئین می‌شود [۲]. هانتینگتین پروتئینی واقع در سیتوزول است که در نوروپاتی‌ها و سلول‌های سوماتیک یافت می‌شود. پروتئین جهش‌یافته‌ی هانتینگتین، باعث تخریب نوروپاتی‌های مغزی به ویژه دو جز از گره‌های قاعده‌ای مغز به نام‌های هسته‌ی پوسته‌ای و هسته‌ی دمدار می‌شود. زمان شروع بیماری به طور میانگین در بین سال‌های ۳۰ تا ۵۰ زندگی فرد مبتلا اتفاق می‌افتد و انتظار می‌رود تا در حدود ۱۵ تا ۲۰ سال بعد ادامه داشته

در این بخش نتایج حاصل از تحلیل داده‌ها و ژن‌های استخراج شده آورده شده است.

تحلیل داده‌های پردازش نشده

در گام اول از نمودار آتش فشان برای نمایش منهای لگاریتم پی-مقدار در مبنای ۱۰ نسبت به مقدار تغییرات لگاریتم بیان ژن‌ها، استفاده می‌کنیم. در این نمودار ژن‌هایی مهم هستند، که پی-مقدار در آن‌ها کم و تغییرات ژن‌ها در دو دسته زیاد باشد، بنابراین هرچه ژن به قسمت بالا-راست یا بالا-چپ نمودار نزدیک‌تر باشد، اهمیت بیش‌تری می‌یابد. با توجه به



شکل ۲: نمودار آتش فشان ژن‌های اندازه‌گیری شده

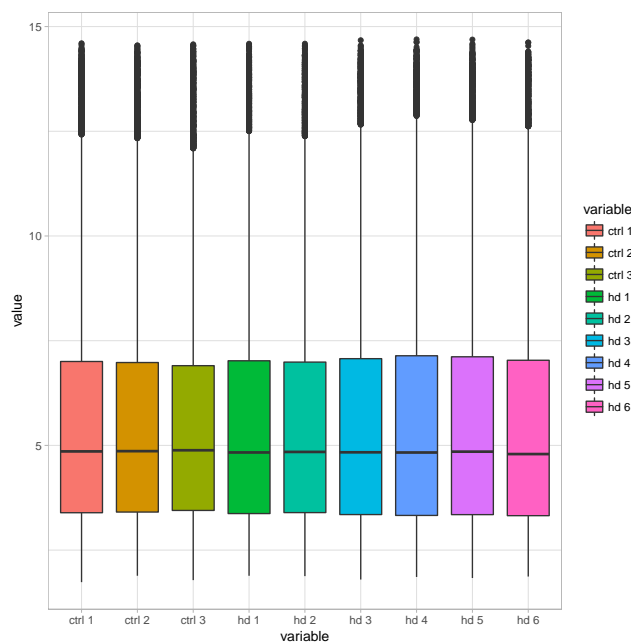
نمودار ۲، می‌توان گفت که مقادیر آستانه‌ی ۰,۰۱ برای پی-مقدار و ۱ برای لگاریتم نسبت تغییرات، تعداد قابل قبولی از ژن‌ها را از میان تمامی ۵۴۶۷۵ ژن جدا می‌کند.

در گام دوم، نمونه‌های مساله را به کمک تحلیل مولفه‌های اصلی، رسم می‌کنیم. شکل ۳ آریانس داده‌ها را در مولفه‌های مختلف نشان می‌دهد. در این شکل مشخص است که بیش‌تر تغییرات در دو مولفه‌ی اول و دوم اتفاق می‌افتد. شکل ۳ ب نیز نشان می‌دهد که مولفه‌ی اصلی اول، به خوبی می‌تواند نمونه‌های بیمار و سالم مساله را از یک‌دیگر جدا کند. در بخش بعدی به تحلیل ژن‌های انتخاب شده در مرحله‌ی پیش‌پردازش می‌پردازیم.

تحلیل ژن‌های استخراج شده

همانطور که پیش‌تر گفته شد ۱۶۲ ژن از میان ژن‌های اندازه‌گیری شده به عنوان ژن‌هایی که تفاوت قابل توجهی در نمونه‌های سالم و بیمار دارند و احتمال شانس بودن اختلاف آن‌ها پایین است، انتخاب می‌شوند. نمودار

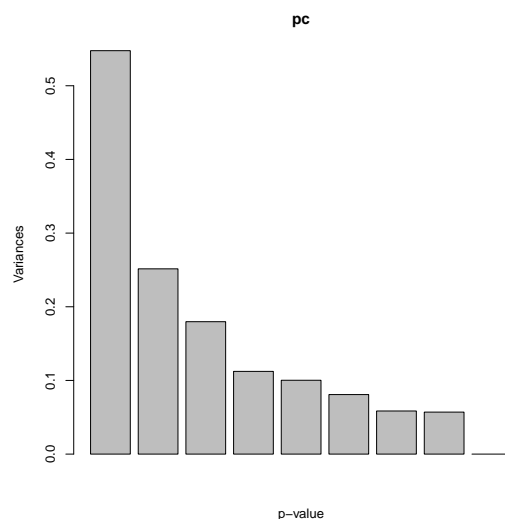
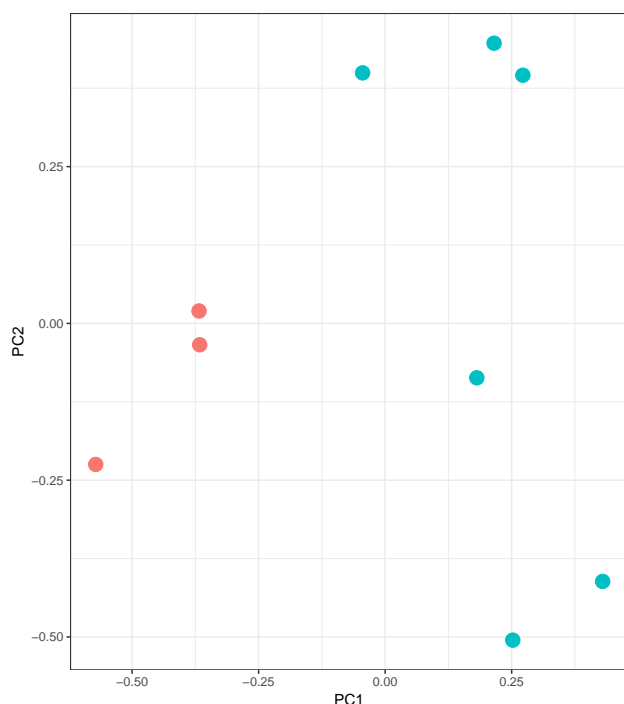
داده‌های این آزمایش به کمک تکنولوژی میکروآرای با استفاده از تراشه‌ی HG-U 133 plus 2.0 از ۳ فرد سالم و ۶ فرد مبتلا به هانتینگتون جمع‌آوری شده است. تمامی افراد مورد آزمایش مرد بوده و سن افراد سالم برابر ۴۸، ۲۸ و ۵۷ و سن افراد بیمار برابر ۳۸، ۶۳، ۵۷، ۴۷ و ۳۸ است. ۵۴۶۷۵ ژن در هر یک از افراد اندازه‌گیری شده است. شکل ۱ توزیع نمونه‌های مساله را نشان می‌دهد.



شکل ۱: توزیع داده‌های بدون پردازش مربوط به ۹ نمونه

روش پیش‌پردازش

همان‌طور که در شکل ۱ نشان داده شده است، چارک‌های داده‌های مساله با هم برابر بوده و داده‌ها اصلاح شده و نرمال‌سازی شده هستند. برای تعیین تفاوت بیان ژن در دو دسته، ابتدا از روش آزمایشی بیز [۹] برای تعیین اختلاف لگاریتم بیان ژن در دو دسته (logFC) و احتمال شانس بودن اختلاف بیان در صورت عدم اختلاف میان آن‌ها در واقعیت (p-value یا همان پی-مقدار) استفاده شده است. سپس ژن‌هایی که احتمال شانس بودن اختلاف بیان در آن‌ها کم‌تر از ۰,۰۱ بوده و قدر مطلق اختلاف لگاریتم بیان میان آن‌ها حداقل ۱ است، انتخاب می‌شوند. تعداد ۸۴۸ ژن در این مرحله انتخاب می‌شوند، سپس به کمک ابزار GSEA ژن‌ها و امتیاز آن‌ها را که به کمک تست جایگشت به دست آمده‌اند، استخراج کرده و ژن‌هایی که مقدار قدر مطلق امتیاز آن‌ها حداقل ۰,۵ است، انتخاب می‌کنیم. تعداد ژن‌های این دسته برابر ۵۵۰ است. در نهایت برای انتخاب ژن‌های نهایی اشتراک میان این دو دسته ژن را در نظر می‌گیریم. تعداد ۲۲۸ ژن در این دو دسته مشترک هستند که ۱۶۲ تا از آن‌ها یکتا هستند. از این ژن‌ها به عنوان ژن‌های اصلی متفاوت میان افراد بیمار و سالم استفاده می‌کنیم. در ادامه نتایج روی داده‌های پردازش نشده و ژن‌های استخراج شده آورده شده است.



(آ) نسبت واریانس داده‌ها در مولفه‌های اصلی. مولفه‌ها از چپ به راست (ب) نمودار نمایش بیان ژن در افراد مبتلا به هانتینگتون و افراد سالم مولفه‌های اصلی اول و دوم. مولفه‌های اول تا ۱۹م را نشان می‌دهند. هر نقطه نمایش دهنده‌ی یک نمونه در مساله است.

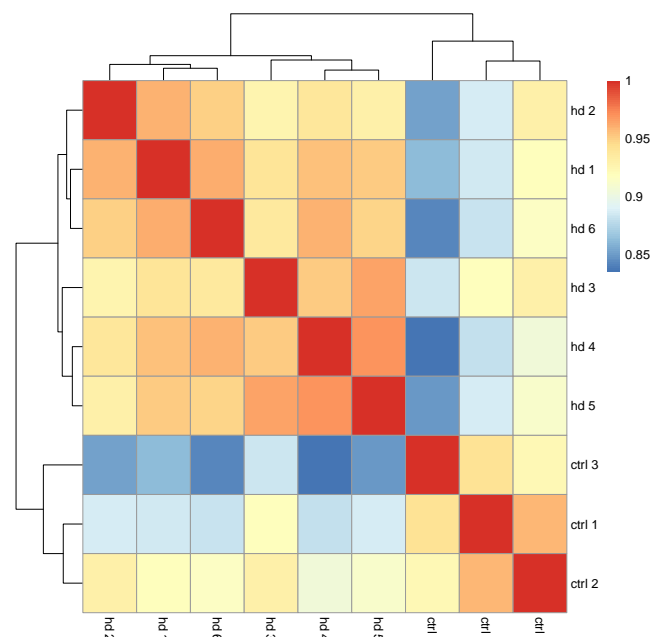
شکل ۳: نمودار نمایش تغییرات ژن‌ها در مولفه‌های اصلی و جدایی نمونه‌ها بر اساس مولفه‌های اصلی اول و دوم

است. بخش‌های زرد در شکل ۵ نشان‌دهنده‌ی اهمیت ژن‌ها در مسیرهای مشخص شده است. همان‌طور که در شکل ۵ دیده می‌شود، یکی از مسیرهای مشخص شده توسط ژن‌ها مسیر اصلاح دی‌ان‌ای است، از سوی دیگر می‌دانیم که بیش‌تر مبتلایان به هانتینگتون به صورت ارثی دارای این بیماری هستند و دی‌ان‌ای در تمام سلول‌های آنان دارای تکرار بیش‌ازحد سه‌تایی CAG است، بنابراین می‌توان گفت که ژن‌های تولید شده، در راستای اصلاح دی‌ان‌ای تولید می‌شوند. مسیر مهم دیگری که در شکل دیده می‌شود، چرخه‌ی سلولی است. این موضوع نشان می‌دهد که ژن‌های متفاوت در افراد بیمار می‌تواند بر چرخه‌ی سلولی اثر بگذارد و احتمال با همین اثر موجب مرگ سلول بشود.

در گام بعدی به بررسی اهمیت ژن‌های مشخص شده در فرآیندهای زیستی بدن می‌پردازیم. برای این کار از سایت Enrichr [۱۱، ۱۲] استفاده می‌کنیم. نتایج استخراج شده در جدول ۱ خلاصه شده است. همان‌طور که در جدول دیده می‌شود، بیماری هانتینگتون باعث تغییر در عملکرد لیزوزوم و وزیکول در سلول می‌شود. نوروپاتی تحت‌تأثیر در بیماری هانتینگتون، پروتئین هانتینگتین جهش‌یافته را در هسته و سیتوپلاسم سلول ذخیره می‌کنند. نقص در لیزوزوم‌ها در سلول‌های فیبروبلاست به عنوان یکی از نشانه‌های بیماری هانتینگتون شناخته می‌شود [۱۵، ۱۶]. ذخیره‌ی هانتینگتین به نحوی موجب مرگ سلول می‌شود.

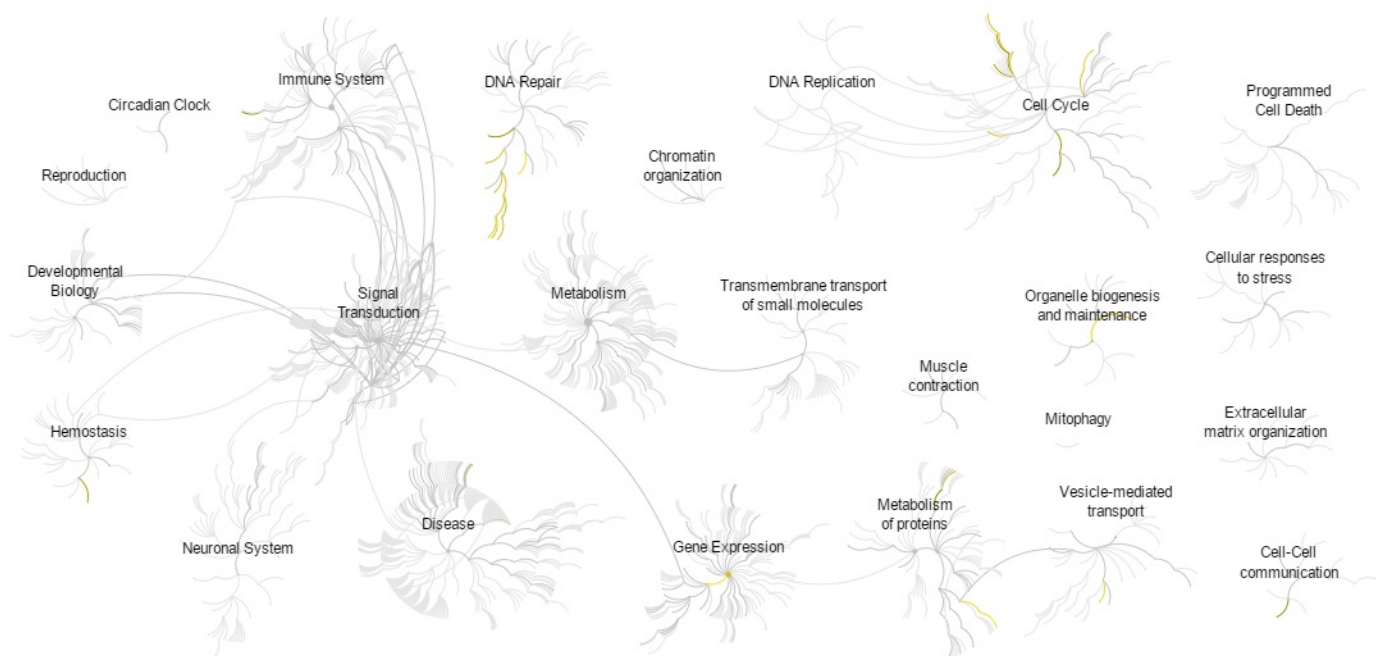
در نهایت با بررسی هستی‌شناسی فنوتیپ (Phenotype Ontology) به نتیجه‌ی گزارش‌شده در جدول ۲ می‌رسیم. به طور کلی نتایج به‌دست‌آمده در جدول بسیار نزدیک به علایم شناخته‌شده در سلول‌های سوماتیکی است که تحت تأثیر بیماری هانتینگتون قرار گرفته‌اند. این

شباهت داده‌های مساله بر اساس ژن‌های انتخابی در شکل ۴ آورده شده است. در این شکل مشخص است که شباهت میان افراد سالم و نیز شباهت میان افراد بیمار کاملاً آن‌ها را از یک‌دیگر جدا ساخته است.



شکل ۴: نمودار شباهت نمونه‌ها بر اساس ژن‌های انتخاب شده

در مسیریابی زیستی ژن‌های به‌دست‌آمده، از منبع [۱۰] استفاده شده



شکل ۵: نمودار اهمیت ژن‌های انتخاب‌شده در فعالیت‌های زیستی بدن [۱۰]

جدول ۱: نتیجه‌ی فرآیندهای زیستی حاصل از ژن‌های استخراج‌شده از بیماران هانتینگتون نسبت به افراد عادی

| Process | Ajusted P-value | Reference |
|--|-----------------|-----------|
| Protein localization to lysosome | 0.01670 | [13, 14] |
| Protein localization to vacuole | 0.01670 | [13, 14] |
| Protein targeting to vacuole | 0.01670 | [15] |
| Protein targeting to lysosome | 0.01670 | [14] |
| Establishment of protein localization to vacuole | 0.01670 | [14] |

جدول ۲: نتیجه‌ی هستی‌شناسی فنوتیپ در انسان با توجه به ژن‌های شناخته‌شده

| Name | Ajusted P-value | Reference |
|-----------------------|-----------------|-----------|
| Neurodegeneration | 0.1422 | [17] |
| Autism | 0.1897 | [18] |
| Heterogeneous | 0.1897 | [19] |
| Aganglionic megacolon | 0.1897 | |
| Macrotia | 0.1897 | |

انتخاب ژن‌های مورد مقایسه برای دو دسته از نمونه‌ها بررسی شده است. در گام بعد، مسیرهای زیستی‌ای که تحت تاثیر ژن‌ها قرار گرفته‌اند، فرآیندهای زیستی مرتبط و هستی‌شناسی فنوتیپ آن‌ها مورد بررسی قرار گرفت. نتایج نشان می‌دهد که سلول‌های فیروبلست، سلول‌های مهمی در بیماری هانتینگتون بوده و تغییر ژن‌ها در آن‌ها قابل تشخیص است. در گام‌های آینده می‌توان به بررسی تعداد بیش‌تری از ژن‌ها پرداخته و تاثیر ژن‌ها را بر عملکرد سلولی مورد بررسی قرار داد.

موضوع در مقاله‌های پیشین نیز بررسی شده است. نکته‌ی جالب توجه در جدول وجود بیماری اوتیسم در این جدول است، زیرا هردو از بیماری‌های نورولوژی هستند.

۴- نتیجه‌گیری

در این مقاله به بررسی تغییر مقادیر ژن‌ها در سلول‌های فیروبلست در بیماران مبتلا به هانتینگتون و افراد سالم پرداختیم. ابتدا نحوه‌ی

- [12] M. V. Kuleshov, M. R. Jones, A. D. Rouillard, N. F. Fernandez, Q. Duan, Z. Wang, S. Koplev, S. L. Jenkins, K. M. Jagodnik, A. Lachmann *et al.*, "Enrichr: a comprehensive gene set enrichment analysis web server 2016 update," *Nucleic acids research*, p. gkw377, 2016.
- [13] K. B. Kegel, M. Kim, E. Sapp, C. McIntyre, J. G. Castaño, N. Aronin, and M. DiFiglia, "Huntingtin expression stimulates endosomal-lysosomal activity, endosome tubulation, and autophagy," *Journal of Neuroscience*, vol. 20, no. 19, pp. 7268–7278, 2000.
- [14] E. Sapp, C. Schwarz, K. Chase, P. Bhidé, A. Young, J. Penney, J. Vonsattel, N. Aronin, and M. DiFiglia, "Huntingtin localization in brains of normal and huntington's disease patients," *Annals of neurology*, vol. 42, no. 4, pp. 604–612, 1997.
- [15] J. Velier, M. Kim, C. Schwarz, T. W. Kim, E. Sapp, K. Chase, N. Aronin, and M. DiFiglia, "Wild-type and mutant huntingtins function in vesicle trafficking in the secretory and endocytic pathways," *Experimental neurology*, vol. 152, no. 1, pp. 34–40, 1998.
- [16] C. Erie, M. Sacino, L. Houle, M. L. Lu, and J. Wei, "Altered lysosomal positioning affects lysosomal functions in a cellular model of huntington's disease," *European Journal of Neuroscience*, vol. 42, no. 3, pp. 1941–1951, 2015.
- [17] J. M. Gil and A. C. Rego, "Mechanisms of neurodegeneration in huntington's disease," *European Journal of Neuroscience*, vol. 27, no. 11, pp. 2803–2820, 2008.
- [18] R. E. Frye and D. A. Rossignol, "Mitochondrial dysfunction can connect the diverse medical symptoms associated with autism spectrum disorders," 2011.
- [19] D. S. Magazi, V. Bonev, M. MoagiB, Z. Iqbal, M. Dlodla, C. Van der Meyden, and A. Krause, "Huntington's disease: genetic heterogeneity in black african patients," *SAMJ: South African Medical Journal*, vol. 98, no. 3, pp. 200–203, 2008.
- [1] F. O. Walker, "Huntington's disease," *The Lancet*, vol. 369, no. 9557, pp. 218–228, 2007.
- [2] Y. P. Goldberg, H. Telenius, and M. R. Hayden, "The molecular genetics of huntington's disease," *Current opinion in neurology*, vol. 7, no. 4, pp. 325–332, 1994.
- [3] P. Harper, "Huntington's disease: a historical background," *OXFORD MONOGRAPHS ON MEDICAL GENETICS*, vol. 45, no. 1, pp. 3–27, 2002.
- [4] O. Quarrell, K. L. O'Donovan, O. Bandmann, and M. Strong, "The prevalence of juvenile huntington's disease: a review of the literature and meta-analysis," *PLOS Currents Huntington Disease*, 2012.
- [5] A.-W. Heemskerk and R. A. Roos, "Aspiration pneumonia and death in huntington's disease," *PLoS currents*, vol. 4, 2011.
- [6] M. Abildtrup and M. Shattock, "Cardiac dysautonomia in huntington's disease," *Journal of Huntington's disease*, vol. 2, no. 3, pp. 251–261, 2013.
- [7] E. Marchina, S. Misasi, A. Bozzato, S. Ferraboli, C. Agosti, L. Rozzini, G. Borsani, S. Barlati, and A. Padovani, "Gene expression profile in fibroblasts of huntington's disease patients and controls," *Journal of the neurological sciences*, vol. 337, no. 1, pp. 42–46, 2014.
- [8] A. Hodges, A. D. Strand, A. K. Aragaki, A. Kuhn, T. Sengstag, G. Hughes, L. A. Elliston, C. Hartog, D. R. Goldstein, D. Thu *et al.*, "Regional and cellular gene expression changes in human huntington's disease brain," *Human molecular genetics*, vol. 15, no. 6, pp. 965–977, 2006.
- [9] G. K. Smyth *et al.*, "Linear models and empirical bayes methods for assessing differential expression in microarray experiments," *Stat Appl Genet Mol Biol*, vol. 3, no. 1, p. 3, 2004.
- [10] "Reactome," <http://www.reactome.org>, accessed: 2017-02-03.
- [11] E. Y. Chen, C. M. Tan, Y. Kou, Q. Duan, Z. Wang, G. V. Meirelles, N. R. Clark, and A. Ma'ayan, "Enrichr: interactive and collaborative html5 gene list enrichment analysis tool," *BMC bioinformatics*, vol. 14, no. 1, p. 128, 2013.