International Journal of Engineering, Science and Mathematics
Vol. 7 Special Issue 4(3), April 2018
ISSN: 2320-0294 Impact Factor: 6.765
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Identifying causal genes for Amyotrophic Lateral Sclerosis (ALS) by Meta Analysis of Gene Expression Data

Uttiyo Silroy*
Ankur Bhowal**

Abstract

Amyotrophic lateral sclerosis (ALS) is an idiopathic, fatal neurodegenerative disease characterized by progressive muscular paralysis reflecting degeneration of upper andlowers motor neurones in the primary motor cortex, corticospinal tracts, brainstemand spinal cord. Analysing multiple microarray data from multiple studies available in public database such as GEO (Gene Expression Omnibus) database is more effective and will give hint to more detailed gene-gene interaction patterns that is particularly more helpful in Identifying novel genes as disease biomarker or identifying regulatory pathways for further in depth functional study. Present study attemps to identify genes associated with ALS by meta analysis of published gene expression data.

Keywords:

Amyotrophic lateral sclerosis;
Meta analysis;
Gene Expression Omnibus;
GEO2R;
STRING.

Author correspondence:

Uttiyo Silroy,
M.Sc. Student,
Vidyasagar College,
Email: usilroy@gmail.com

1. Introduction

Amyotrophic lateral sclerosis (ALS) is an idiopathic, fatal neurodegenerative disease characterized by progressive muscular paralysis reflecting degeneration of upper and lower motor neurons in the primary motor cortex, corticospinal tracts, brainstem and spinal cord. It is a genetic disorder combining multiple hereditary and epigenetic factors which has made it nearly impossible to identify genetic biomarkers for the disease. Due to its complex nature a meta-analysis combining multiple different expression studies is done to understand the altered biological and molecular functions in ALS. Analyzing multiple microarray data from multiple studies available in public database such as GEO (Gene Expression Omnibus) database is particularly more effective and will give hint to more detailed gene-gene interaction patterns that is particularly more helpful in Identifying novel genes as disease biomarker or identifying regulatory pathways for further in depth functional study. However, there may be disagreement among different study results of different microarray experiments, Meta-analysis focuses on contrasting and combining results from different studies in the hope of identifying patterns among study results of the disease. The target of the study is to find out differentially expressed genes based on data from the GEO database to from a coexpressed gene list followed by Gene Ontology analysis to find out deregulated biological pathways and deregulated molecular function. A significant downregulation in Muscular proliferation and growth coupled with an increase in ATP synthesis by upregulation of Mitochondrial ETC pathways is observed along with significant reduction of protein binding associated gene expression.

*Uttiyo Silroy, Post Graduate Student in the Department of Zoology, Vidyasagar College, Kolkata.
**Assistant Professor, Department of Zoology, Vidyasagar College, Kolkata.

2. Research Method

A. Extraction of genes for analysis

This meta-analysis was based on the studies of Bakay et al.,(2005); Shtilbans et al., (2011) and Bernardini et al., (2013). Gene expression data was acquired as GEO datasets and recognized as GSE3307, GSE41414 and GSE26276 respectively. The composite dataset from these three studies consisted of 87022 genes from three studies, and of which 12219 up-regulated genes in were ALS and 12938 genes were downregulated. We extracted the top significant differentially expressed (≥1.5 and ≤ 1.5 fold up regulated and down regulated) genes. Overall from the first study of Bakay et al., (2005) we have selected 18 ALS and 26 control. From the second study of Shtilbans et al., (2011) we have selected 7 ALS and 7 control. And from the study of Bernardini et al., (2013) we have selected 3 ALS and 3 Control. The patient expression were analyzed in the Web based GEO2R software (based on t statistics for pair wise comparison). Results are obtained as a table of genes ordered by significance.

B. Extraction of Common genes in between studies and GO analysis.

From the narrowed down list of over expressed and down regulated genes, overlapping genes between the studies were found using various MS Excel functions to find more important candidate genes in ALS pathway. Genes found to be common between studies were used for Biological Function GO analysis. The Gene Ontology (GO) project (http://www.geneontology.org/) provides structured, controlled vocabularies and classifications that we have used to classify the differentially expressed genes into various biological function categories using GO enrichment analysis. The p-value < 0.01 was considered as significant level.

C. Protein-Protein Interaction network construction

Protein-Protein Interaction network construction was done using STRING, (Search Tool for the Retrieval of Interacting Genes/Proteins, http://string-db.org) for the preparation of Protein-Protein Interaction network. In the present study we have selected the following active interaction sources: 1. Text mining, 2. Experiments, 3. Databases, 4. Co-expression, 5. Neighbourhood, 6. Gene Fusion and 7. Co-occurrence. The minimum required interaction score was: 0.150 for constructing the interaction network. Then a Functional enrichments analysis was done for identifying enriched biological pathways from KEGG (Kyoto Encyclopaedia of Genes and Genomes) pathways.

3. Results and Analysis

The results displays tables that lists **significant shared GO terms (or parents of GO terms)** used to describe the set of genes that was found to be commonly downregulated or upregulated in the pair wise studies.

Table: 1 Partial list showing GO enrichment from total 12938 downregulated genes and from 12219 upregulated genes (pool of 3 studies).

GO biological function category enriched by downregulated genes in ALS condition	NO. OF GENES	GO biological function category enriched by upregulated genes in ALS condition	NO. OF GENES
ERK1 and ERK2 cascade (GO:0070371)	5	Mitochondrial electron transport, cytochrome c to oxygen(GO:0006123)	6
Skeletal muscle contraction (GO:0003009)	5	Mitochondrial ATP synthesis coupled electron transport (GO:0042775)	6
Labyrinthine layer development (GO:0010830)	6	ATP synthesis coupled electron transport (GO:0042773)	6
Regulation of myotube differentiation (GO:0010830)	7	Oxidative phosphorylation (GO:0006119)	6
Regulation of striated muscle cell differentiation (GO:0051153)	7	Respiratory electron transport chain (GO:0022904)	6
striated muscle cell proliferation (GO:0014855)	5	Hydrogen ion transmembrane transport (GO:1902600)	6
cardiac muscle cell proliferation (GO:0060038)	4	NADH regeneration (GO:0006735)	5
muscle cell proliferation (GO:0033002)	5	striated muscle contraction (GO:0006941)	9
cardiac muscle tissue growth (GO:0055017)	5	actin-mediated cell contraction (GO:0070252)	8
regulation of cellular protein localization (GO:1903827)	27	purine ribonucleoside triphosphate metabolic process (GO:0009205)	8
positive regulation of apoptotic process (GO:0043065)	25	Skeletal muscle contraction (GO:0003009)	5
positive regulation of cell development (GO:0010720)	20	Actin mediated cell contraction (GO:0070252)	8
enzyme linked receptor protein signalling pathway (GO:0007167)	25	Actin filament based movement (GO:0030048)	8
actin filament-based process (GO:0030029)	14	Proton transport (GO:0015992)	7
Regulation of cellular amide metabolic process (GO:0034248)	14	Carbohydrate catabolic process (GO:0016052)	9

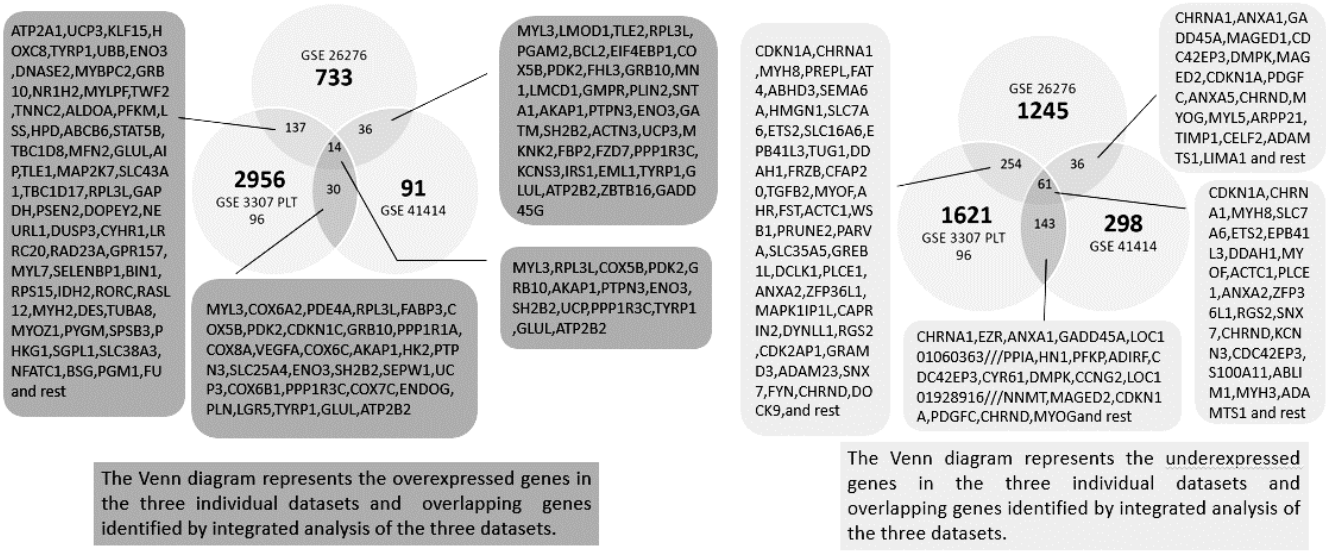

Figure 1. Diagram showing the number of common upregulated and down regulated genes between the studies: GSE3307, GSE41414 and GSE26276. The numbers in bold depicts the total significant gene in the study and the numbers under the overlapped region indicates the genes that are common.

Table 2. Tabulation of Enriched GO Categories from the common under expressed genes (>1.5 fold) extracted from the three studies (GSE3307, GSE41414 & GSE26276).

Enriched KEGG pathway between GSE3307, GSE41414 and GSE26276 upregulated genes	Pathway ID – (01230)Biosynthesis of Amino Acid
	Pathway ID - (05012) Parkinson's disease
	Pathway ID - (00190) Oxidative phosphorylation
	Pathway ID - (05016) Huntington's disease
Enriched KEGG pathway between GSE3307, GSE41414 and GSE26276 downregulated genes	Pathway ID - (04510) Focal adhesion
	Pathway ID - (04151) PI3K-Akt signaling pathway
	Pathway ID – (05206) MicroRNAs in cancer
	Pathway ID – (04810) Regulation of actin cytoskeleton
	Pathway ID - (05410) Hypertrophic cardiomyopathy

4. Conclusion

A significant downregulation in Muscular proliferation and growth coupled with an increase in ATP synthesis by upregulation of Mitochondrial ETC pathways is observed. Protein binding ability suffered by a significant reduction of associated gene expression. Simple meta-analysis of similar datasets from multiple studies can identify genetic pathway and causal genes of a disease that are not indicated in a single dataset. In the present study, meta analysis with ALS skeletal muscle sample dataset indicate the important genes and pathways that can add new perspective to the understanding of molecular mechanism of disease progression.

a. Focal Adhesion Pathway: Includes downregulated genes like IGF1, ITGAV, LAMA2, MYL5

Skeletal muscle has a remarkable ability to respond to different physical stresses. Loading muscle through exercise, either anaerobic or aerobic, can lead to increases in muscle size and function while, conversely, the absence of muscle loading stimulates rapid decreases in size and function. A principal mediator of this load-induced change is focal adhesion kinase (FAK), a downstream non-receptor tyrosine kinase that translates the cytoskeletal stress and strain signals transmitted across the cytoplasmic membrane by integrins to activate multiple anti-apoptotic and cell growth pathways. Changes in FAK expression and phosphorylation have been found to correlate to specific developmental states in myoblast differentiation, muscle fiber formation and muscle size in response to loading and unloading. With the capability to regulate costamere formation, hypertrophy and glucose metabolism, FAK is a molecule with diverse functions that are important in regulating muscle cell health.A significant decrease in expression of Focal Adhesion Kinase is seen in patients with ALS, this may be one of the causes for the occurance of muscle dystrophy in ALS.

b. Muscle System Process: Includes Upregulated genes like ALDOA, TNNC2, ATP2A1, MYL3, TNNI2, MYLPF, MYH2, MYH7, TCAP, MYBPC2, MYLPF, MYL7
An increase in genes of muscle system processes and glucose metabolic process is observed. Probable causes can be to offset the load set by degenerated or atrophied muscles.

c. Protein Binding : Includes Downregulated genes like: MYH8, MAPK1, ANXA2, ENO1, MAP2K1, SMARCD3,EZR,ACTC1,PRKAR1A,NCOA1,CDKN1A,SMARCA4,RUVBL2,CD44,NPM1,HDAC4,MYH 3, RAN
d. Biosynthesis of Amino Acids: includes Upregulated genes like : IDH2, GOT2, ALDOA, ACO2, PFKM, ENO3, TPI1, GLUL

e. Actin Binding: includes Downregulated genes like: CORO1C, MYH3, MYH8

f. Oxidative Phosphorylation includes Upregulated genes like: COX5B, COX7C, COX6B1, COX8A, COX6C, COX6A2

4. References

1. Bakay M, Wang Z, Melcon G, Schiltz L, Xuan J, Zhao P, Sartorelli V, Seo J, Pegoraro E, Angelini C, Shneiderman B, Escolar D, Chen YW, Winokur ST, Pachman LM, Fan C, Mandler R, Nevo Y, Gordon E, Zhu Y, Dong Y, Wang Y, Hoffman EP. Brain. Nuclear envelope dystrophies show a transcriptional fingerprint suggesting disruption of Rb-MyoD pathways in muscle regeneration. 2006 Apr;129(Pt 4):996-1013. Epub 2006 Feb 14.

2. Bernardini C, Censi F, Lattanzi W, Barba M, Calcagnini G, Giuliani A, Tasca G, Sabatelli M, Ricci E, Michetti F. Mitochondrial network genes in the skeletal muscle of amyotrophic lateral sclerosis patients. PLoS One. 2013;8(2):e57739.

3. Shtilbans A, Choi SG, Fowkes ME, Khitrov G, Shahbazi M, Ting J, Zhang W, Sun Y, Sealfon SC, Lange DJ. Differential gene expression in patients with amyotrophic lateral sclerosis. Amyotroph Lateral Scler. 2011 Jul;12(4):250-6.