

RESEARCH

Open Access

# Unveiling transcription factor regulation and differential co-expression genes in Duchenne muscular dystrophy

Lijun Tian<sup>\*</sup>, Junhua Cao, Xingqiang Deng, Chuanling Zhang, Tong Qian, Xianxiang Song and Baoshan Huang

## Abstract

**Background:** Gene expression analysis is powerful for investigating the underlying mechanisms of Duchenne muscular dystrophy (DMD). Previous studies mainly neglected co-expression or transcription factor (TF) information. Here we integrated TF information into differential co-expression analysis (DCEA) to explore new understandings of DMD pathogenesis.

**Methods:** Using two microarray datasets from Gene Expression Omnibus (GEO) database, we firstly detected differentially expressed genes (DEGs) and pathways enriched with DEGs. Secondly, we constructed differentially regulated networks to integrate the TF-to-target information and the differential co-expression genes.

**Results:** A total of 454 DEGs were detected and both KEGG pathway and ingenuity pathway analysis revealed that pathways enriched with aberrantly regulated genes are mostly involved in the immune response processes. DCEA results generated 610 pairs of DEGs regulated by at least one common TF, including 78 pairs of co-expressed DEGs. A network was constructed to illustrate their relationships and a subnetwork for DMD related molecules was constructed to show genes and TFs that may play important roles in the secondary changes of DMD. Among the DEGs which shared TFs with *DMD*, six genes were co-expressed with *DMD*, including *ATP1A2*, *C1QB*, *MYOF*, *SAT1*, *TRIP10*, and *IFI6*.

**Conclusion:** Our results may provide a new understanding of DMD and contribute potential targets for future therapeutic tests.

**Virtual slides:** The virtual slide(s) for this article can be found here: [http://www.diagnosticpathology.diagnomx.eu/vs/13000\\_2014\\_210](http://www.diagnosticpathology.diagnomx.eu/vs/13000_2014_210)

**Keywords:** Duchenne muscular dystrophy, Differential co-expression analysis, Transcription factor

## Background

Duchenne muscular dystrophy (DMD) is a severe and progressive inherited neuromuscular disease characterized by muscle fiber degeneration and central nervous system disorders that affects 1 in 3600–6000 live male births [1]. This disease is caused by mutations or dysregulation of the X-linked dystrophin gene. Absence of or defects of dystrophin protein result in disruption of the dystrophin-associated protein complex (DAPC), resulting in chronic inflammation and progressive muscle degeneration [2].

Although mutations of the DMD protein have been identified to be primarily responsible for the pathology, comprehensive understanding of the downstream mechanisms due to dystrophin absence is still lacking. Secondary changes in DMD involve calcium homeostasis [3], nitric oxide synthase [4], inflammation [5] and mast cell degranulation [6], suggesting that the pathological process of DMD is highly complicated. Gene expression profile analysis is a powerful strategy to investigate the pathophysiological mechanisms of DMD. Several gene expression profiling studies [7–11] have been performed earlier, providing insights into the pathology of DMD.

\* Correspondence: [tlj1168@163.com](mailto:tlj1168@163.com)

Clinical laboratory, Xuzhou Children's Hospital, No. 18 Sudibei Road, Xuzhou, Jiangsu Province, China

However, most of them focused on expression level changes of individual genes, without considering the differences in gene interconnections. Since the majority of proteins function with other proteins, differential co-expression analysis (DCEA) would provide a better understanding of the mechanism underlying DMD progression. The rationale of DCEA is that changes in gene co-expression patterns between two disease statuses (case and control) provide hints regarding the disrupted regulatory relationships in patients. Therefore, integrate the information of Transcription factors (TFs) into DCEA may detect the regulatory causes of observed co-expression changes since TFs can regulate gene expression through binding the cis-elements in the target genes' promoter regions. TFs are reported to be important in both normal and disease states [12]. Investigation of TF activities in mdx mice has proposed several TFs involved in DMD pathogenesis [13]. Thus, integrative analysis of TFs and target differential co-expression genes may provide better understanding on molecular mechanism of the secondary changes of DMD.

In this study, using microarray gene expression data collected from the Gene Expression Omnibus (GEO) database, we carried out an integrative bioinformatics analysis. Firstly, differentially expressed genes (DEGs) were detected and pathways enriched with DEGs were acquired. Secondly, differentially regulated link networks were constructed to integrate the TF-to-target information and the differential co-expression genes. Our results may provide new understanding of DMD and new targets for future therapeutic tests.

## Methods

### Microarray data

We used two datasets (GSE6011 and GSE3307) from the GEO (<http://www.ncbi.nlm.nih.gov/geo/>) database. For further combined analysis, we only selected samples generated with the platform GPL96: [HG-U133A] Affymetrix Human Genome U133A Array.

### Detection of differentially expressed genes (DEGs)

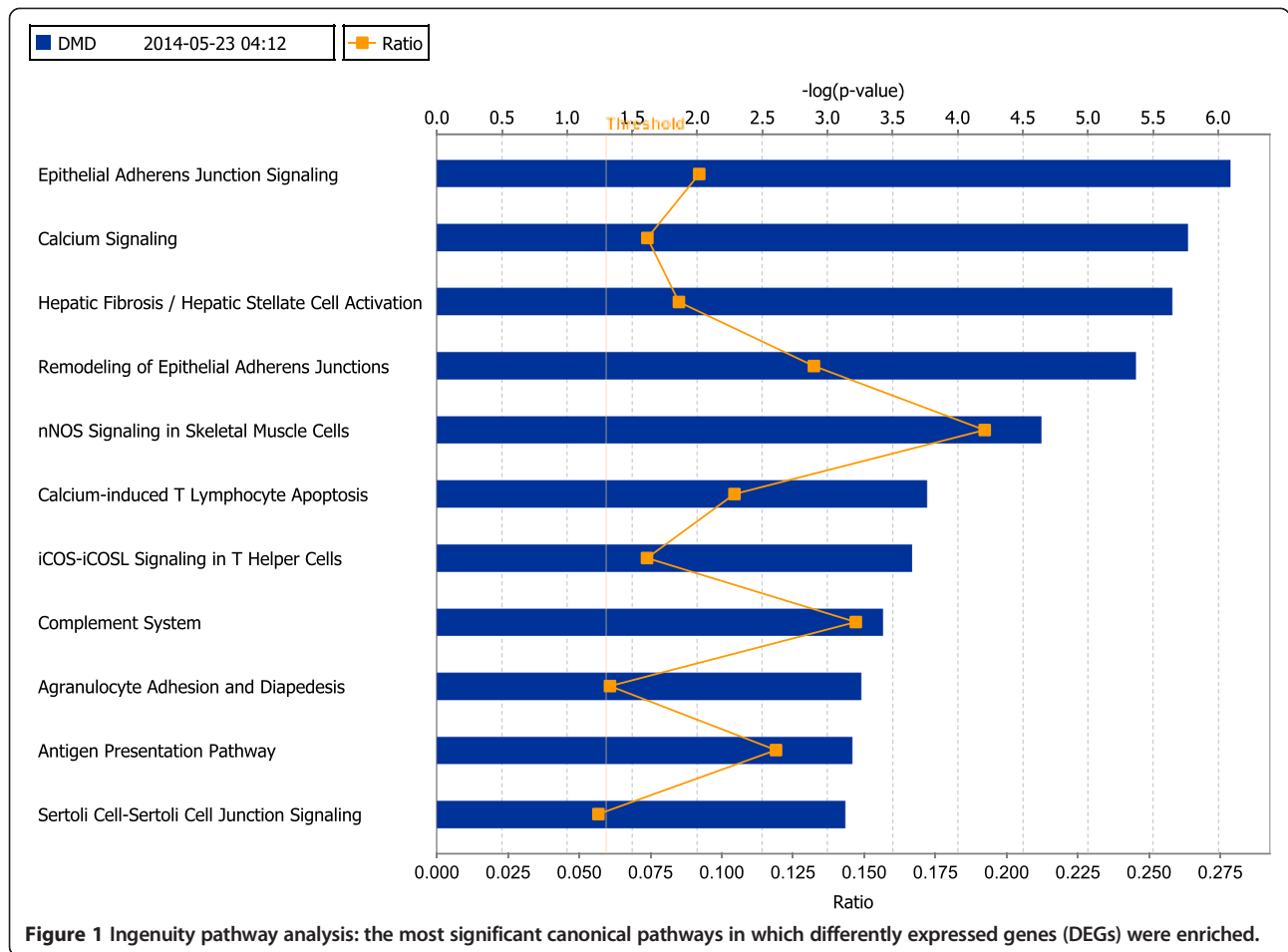
Raw data including simple omnibus format in text (SOFT) family files and CEL files of all samples were downloaded. The CEL files contained the expression information for each probe. Robust Multiarray Analysis (RMA) [14] was used for raw intensity values normalization: firstly, background noise and processing artifacts were neutralized by model-based background correction; secondly, expression values were set to a common scale by quantile normalization; thirdly, expression value of each probe was generated by an iterative median polishing procedure. The resulting log<sub>2</sub>-transformed RMA expression value was then used for detecting DEGs. Statistical *t* test was used to identify DEGs and the Benjamini and Hochberg procedure [15] were carried out to correct the multiple testing problems. The threshold of significant DEGs was set as  $P < 0.01$ . Up- or down-regulation of the DEGs were determined according to the fold-change. All of the above procedures were performed using the R software (v3.0.3) with BioConductor and limma packages (3.12.1) and libraries [16].

### Functional analysis

Differentially expressed probes were annotated based on the SOFT files. All genes were then mapped to the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways (<http://www.genome.jp/kegg/>) database. The hyper geometric distribution test was used to identify pathways significantly enriched with DEGs. In addition to KEGG pathway analysis, DEGs were also uploaded into the Ingenuity Pathway Analysis (IPA) software (Ingenuity® Systems, <http://www.ingenuity.com>) and was associated with the canonical pathways, molecular and cellular functions, diseases and disorders in the Ingenuity Knowledge Base. Fisher's exact test was implemented to assess the significance of the associations between DEGs and canonical pathways, functions, or diseases. For canonical pathways, a ratio was also computed between the number of DEGs and the total number of molecules in the pathway.

**Table 1 KEGG pathways enriched with differentially expressed genes**

ID	Pathway description	Pathway class	P-value
hsa05322	Systemic lupus erythematosus	Immune diseases	7.40E-05
hsa05416	Viral myocarditis	Cardiovascular Diseases	1.60E-04
hsa04514	Cell adhesion molecules (CAMs)	Signaling Molecules and Interaction	4.00E-03
hsa05310	Asthma	Immune Diseases	1.00E-02
hsa04510	Focal adhesion	Cell Communication	1.00E-02
hsa05130	Pathogenic Escherichia coli infection	Infectious Diseases	1.10E-02
hsa04672	Intestinal immune network for IgA production	Immune system	1.90E-02
hsa05330	Allograft rejection	Immune Diseases	1.70E-02
hsa04612	Antigen processing and presentation	Immune System	2.20E-02
hsa05332	Graft-versus-host disease	Immune Diseases	2.10E-02



### DCEA and TF-to-target analysis

The DEGs were subjected to DCEA and TF-to-target analysis using the DCGL package (v2.0) [17] in the R software. Human TF-to-target library was incorporated into the package to identify differentially regulated genes and links. A total of 215 human TFs and 16,863 targets were included in the package. Differentially co-expressed genes (DCGs) were firstly identified and a network was constructed with DCGs and their shared TFs. Since dysregulation of *DMD* gene is the primary cause of the disease, we also constructed a subnetwork surrounding the predefined gene *DMD*.

### Results

With the two downloaded datasets (GSE6011 and GSE3307), expression profile of 27 DMD patients (22 from GSE6011 and 5 from GSE3307) and 14 healthy controls were included in this study. Dystrophin protein was absent in all patients [9,10]. A total of 454 genes were detected to be DEGs, including 284 up regulated genes and 170 down regulated ones. KEGG pathway analysis revealed that aberrantly regulated genes were enriched in 10 pathways (Table 1). Most of these pathways (7/10) are involved in the

inflammatory/immune response process, such as the immune system, immune disease and infectious disease. In addition, a DMD related pathway, the viral myocarditis pathway, was included. The other two pathways are Cell adhesion molecules (CAMs) and Focal adhesion.

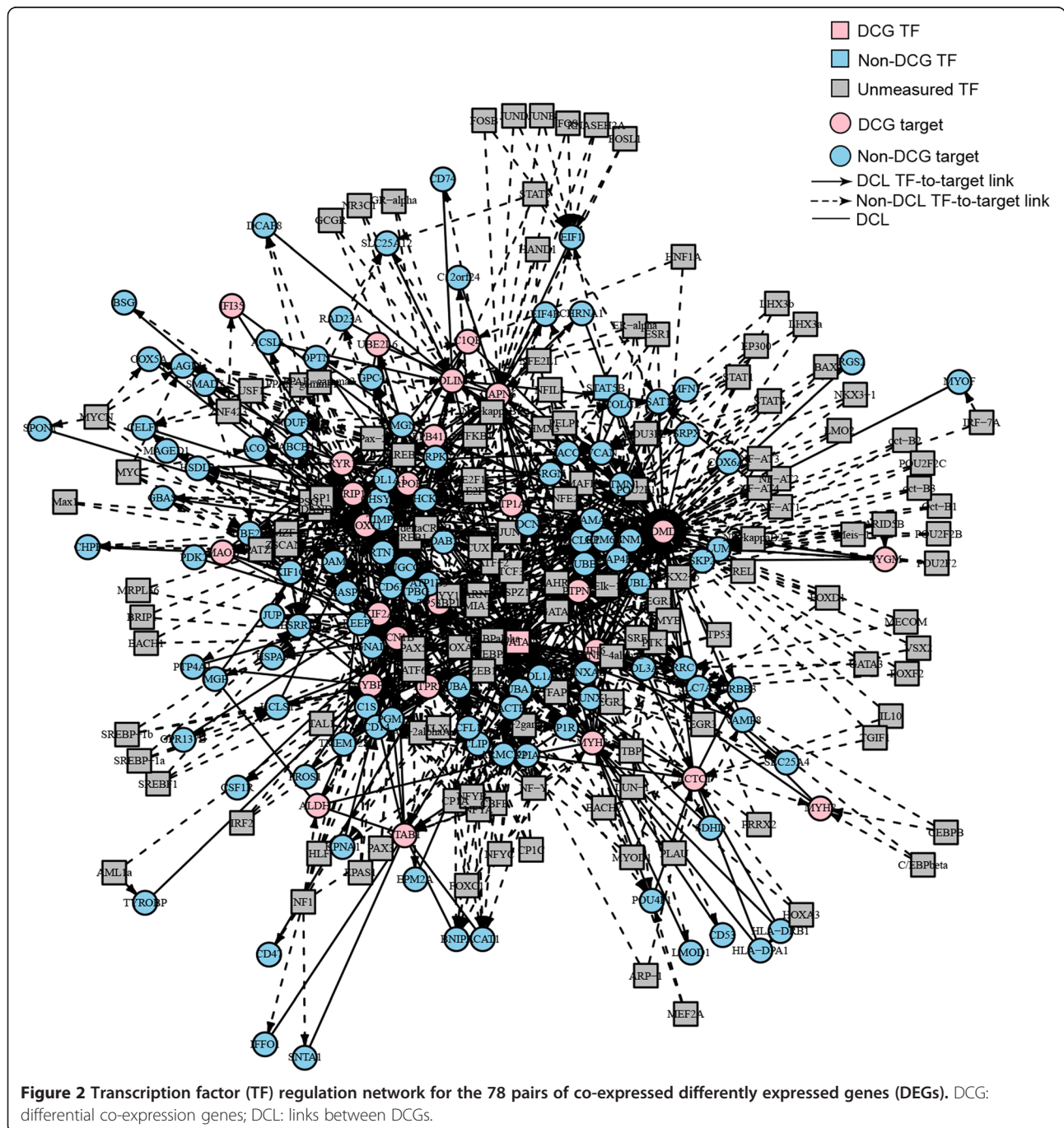
**Table 2** Ingenuity pathway analysis: diseases and functions related to differentially expressed genes

Name	P-value	# Molecules
<b>Diseases and Disorders</b>		
Neurological disease	3.98E-15–1.20E-02	133
Skeletal and muscular disorders	2.66E-13–1.07E-02	118
Cardiovascular disease	5.89E-13–9.41E-03	75
Cancer	9.06E-11–1.14E-02	293
Psychological disorders	3.24E-10–1.13E-02	92
<b>Molecular and Cellular functions</b>		
Cellular growth and proliferation	1.38E-14–1.12E-02	136
Cell death and survival	1.24E-13–1.20E-02	128
Cellular development	8.68E-08–1.12E-02	106
Cellular movement	1.15E-07–1.20E-02	70
Cellular assembly and organization	6.30E-06–9.41E-03	73

The IPA results of canonical pathways showed that the DEGs were enriched for epithelial adherens junction signaling, calcium signaling, nNOS signaling in skeletal muscle cells and some additional pathways (Figure 1). Consistent with the KEGG pathway analysis, aberrantly regulated genes were enriched in immune system pathways, such as calcium-induced T lymphocyte apoptosis, iCOS-iCOSL signaling in T helper cells, complement system, and antigen presentation pathway (Figure 1). Diseases and Disorders analysis (Table 2) revealed that

“Neurological Disease” ( $P = 3.98E-15-1.20E-02$ ) was the top disease affected by these DEGs followed by “Skeletal and Muscular Disorders” ( $P = 2.66E-13-1.07E-02$ ). “Cellular Growth and Proliferation” ( $P = 1.38E-14-1.12E-02$ ) was the top biological function mediated by these DEGs (Table 2).

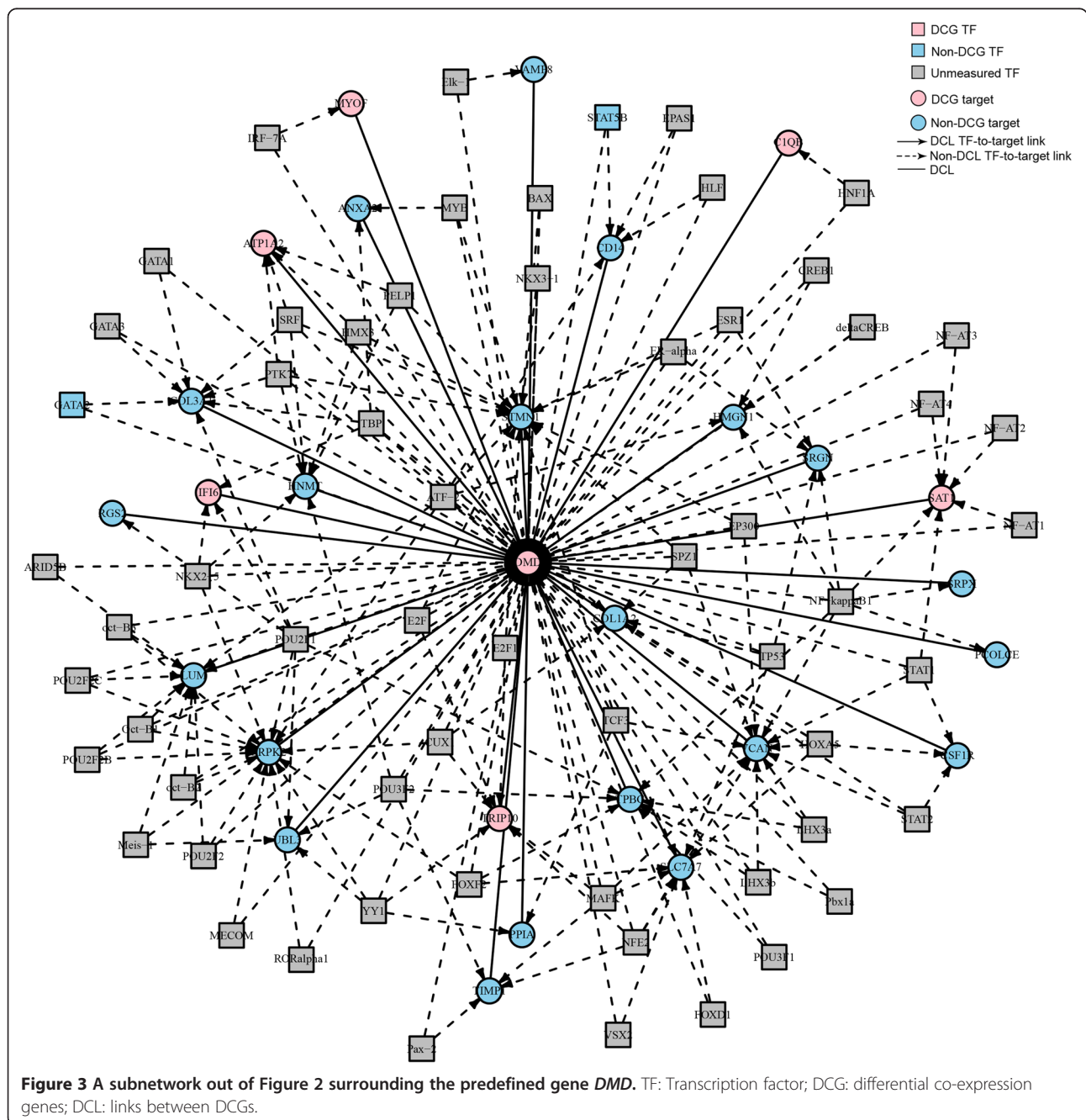
DCEA results generated 610 pairs of DEGs regulated by at least one common TF. Among these pairs, 78 pairs DEGs were identified as differentially co-expressed. The expression values of a total of 25 pairs of DEGs were positively correlated with each other. A network was



generated to illustrate the relationship between the differentially co-expressed genes and their regulatory TFs (Figure 2). Considering the primary causative effect of the aberrant expression of *DMD*, we construct a subnetwork with *DMD* and its differentially co-expressed genes (Figure 3) to illustrate genes and TFs that may play important roles in the secondary changes of DMD patients. As shown in Figure 3, among the DEGs which shared TFs with *DMD*, six genes were co-expressed with *DMD*, including *ATP1A2*, *C1QB*, *MYOF*, *SAT1*, *TRIP10*, and *IFI6*.

### Discussion

The pathophysiology of DMD is highly complex, involving the dysregulation of many downstream cascades. Gene expression profiling is powerful for investigating the secondary changes of DMD patients. However, previous studies mainly focused on individual gene expression changes, without considering the gene co-expression pattern or the TF information. In this study, with combined two DMD microarray datasets, we implemented DCEA and constructed a TF regulated network with the hope to provide new understanding of the pathogenesis.



Both KEGG pathway and IPA canonical pathway analysis of the DEGs revealed that pathways enriched with aberrantly regulated genes are mostly involved in the immune response processes. This may be due to the infiltration of immune cells into the muscles and the elevated levels of various inflammatory cytokines [18-20]. In addition, the calcium signaling and nNOS signaling in skeletal muscle cells pathway were found to be enriched with deregulated genes, which is consistent with previous findings [3,4].

DCEA results generated 610 pairs of DEGs regulated by at least one common TF, including 78 pairs of co-expressed DEGs. As shown in Figure 2, many TFs are involved in the regulation of these DEGs. Since *DMD* is the causative gene, a subnetwork was constructed to illustrate important genes and TFs that may play important roles in the secondary changes of DMD patients (Figure 3). Among the DEGs which shared TFs with *DMD*, six genes were co-expressed with *DMD*, including *ATPIA2*, *C1QB*, *MYOF*, *SAT1*, *TRIP10*, and *IFI6*. Among them, *ATPIA2*, *MYOF*, and *SAT1* have previously been reported to be involved in muscular dystrophy [21-23]. The relationship between the three left genes and DMD is still known. However, it is worth further investigation. Take *IFI6* for example, this gene was highly expressed in muscle and bone and it is reported to inhibit cytochrome c release from mitochondria by regulating the  $Ca^{2+}$  channel which consequently attenuate apoptosis [24]. Alterations in intracellular  $Ca^{2+}$  homeostasis are involved in the alteration of apoptosis since a sustained increase in cytosolic  $Ca^{2+}$  concentration accompanies with apoptosis in cells [25]. Previous study [26] have demonstrated the high influx of extracellular calcium through a dystrophin-deficient membrane, which may lead to subsequent muscle necrosis or apoptosis along with the inflammation response. Whether this gene contributes to the pathogenesis through its regulation of the  $Ca^{2+}$  channel needs further investigation. Among the TFs detected in the array, several TFs of *DMD* are dysregulated, including *GATA2* and *STAT5B*. Further studies are needed to investigate their involvement in the disease and the mechanisms of other unmeasured TFs (Figure 3).

The ultimate cure for DMD will lie in the stable introduction of a functional dystrophin gene into the muscles of DMD patients, however, when gene therapy or transplantation of stem cells/ muscle precursor cells will be clinically available is unpredictable. In the interim, reducing secondary features of the pathologic progression of dystrophin deficiency could improve the quality and length of life for DMD patients. Dysregulation of the genes which were co-expressed with DMD may be caused by the malfunction of DMD. Therapeutic strategies aimed to compensate for the dysregulation of these genes may help

to reduce the secondary features of DMD. Therefore, genes co-expressed with DMD we identified here may be considered as therapeutic targets in further investigations to treat the secondary effects. In addition, TFs that regulated these genes and DMD may also serve as potential therapeutic markers in future studies.

## Conclusions

In summary, with combined microarray data from the GEO database, we conducted DECA and acquire the TF information of co-expressed DEGs to capture pathogenic characteristics of DMD. Our results may provide a new understanding of DMD and further contribute potential targets for future therapeutic tests.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

LT conceived of the study and drafted the manuscript. JC, XD and CZ performed the statistical analysis. TQ, XS and BH revised the manuscript. All authors read and approved the final manuscript.

## Acknowledgement

This study is supported by the Planning program of Xuzhou Science and Technology Bureau in Jiangsu, China (Y3014).

Received: 6 June 2014 Accepted: 12 October 2014

Published online: 23 October 2014

## References

- Bushby K, Finkel R, Birnkrant DJ, Case LE, Clemens PR, Cripe L, Kaul A, Kinnett K, McDonald C, Pandya S, Poysky J, Shapiro F, Tomezska J, Constantin C: **Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management.** *Lancet Neurol* 2010, **9**:77-93.
- Hoffman EP, Brown RH, Kunkel LM: **Dystrophin: the protein product of the Duchenne muscular dystrophy locus.** *Cell* 1987, **51**:919-928.
- Head SI: **Branched fibres in old dystrophic mdx muscle are associated with mechanical weakening of the sarcolemma, abnormal  $Ca^{2+}$  transients and a breakdown of  $Ca^{2+}$  homeostasis during fatigue.** *Exp Physiol* 2010, **95**:641-656.
- Altamirano F, Lopez JR, Henriquez C, Molinski T, Allen PD, Jaimovich E: **Increased resting intracellular calcium modulates NF-kappaB-dependent inducible nitric-oxide synthase gene expression in dystrophic mdx skeletal myotubes.** *J Biol Chem* 2012, **287**:20876-20887.
- Monici MC, Aguenouz M, Mazzeo A, Messina C, Vita G: **Activation of nuclear factor-kappaB in inflammatory myopathies and Duchenne muscular dystrophy.** *Neurology* 2003, **60**:993-997.
- Gorospe JR, Tharp MD, Hinckley J, Kornegay JN, Hoffman EP: **A role for mast cells in the progression of Duchenne muscular dystrophy? Correlations in dystrophin-deficient humans, dogs, and mice.** *J Neurol Sci* 1994, **122**:44-56.
- Chen YW, Nagaraju K, Bakay M, McIntyre O, Rawat R, Shi R, Hoffman EP: **Early onset of inflammation and later involvement of TGFbeta in Duchenne muscular dystrophy.** *Neurology* 2005, **65**:826-834.
- Haslett JN, Sanoudou D, Kho AT, Bennett RR, Greenberg SA, Kohane IS, Beggs AH, Kunkel LM: **Gene expression comparison of biopsies from Duchenne muscular dystrophy (DMD) and normal skeletal muscle.** *Proc Natl Acad Sci U S A* 2002, **99**:15000-15005.
- Pescatori M, Broccolini A, Minetti C, Bertini E, Bruno C, D'Amico A, Bernardini C, Mirabella M, Silvestri G, Giglio V, Modoni A, Pedemonte M, Tasca G, Galluzzi G, Mercuri E, Tonali PA, Ricci E: **Gene expression profiling in the early phases of DMD: a constant molecular signature characterizes DMD muscle from early postnatal life throughout disease progression.** *FASEB J* 2007, **21**:1210-1226.

10. Chen YW, Zhao P, Borup R, Hoffman EP: **Expression profiling in the muscular dystrophies: identification of novel aspects of molecular pathophysiology.** *J Cell Biol* 2000, **151**:1321–1336.
11. Wong B, Gilbert DL, Walker WL, Liao IH, Lit L, Stamova B, Jickling G, Apperson M, Sharp FR: **Gene expression in blood of subjects with Duchenne muscular dystrophy.** *Neurogenetics* 2009, **10**:117–125.
12. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM: **A census of human transcription factors: function, expression and evolution.** *Nat Rev Genet* 2009, **10**:252–263.
13. Dogra C, Srivastava DS, Kumar A: **Protein-DNA array-based identification of transcription factor activities differentially regulated in skeletal muscle of normal and dystrophin-deficient mdx mice.** *Mol Cell Biochem* 2008, **312**:17–24.
14. 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.
15. Benjamini Y, Hochberg Y: **Controlling the false discovery rate: a practical and powerful approach to multiple testing.** *J R Stat Soc Ser B Methodol* 1995, **57**:289–300.
16. Smyth GK, Michaud J, Scott HS: **Use of within-array replicate spots for assessing differential expression in microarray experiments.** *Bioinformatics* 2005, **21**:2067–2075.
17. Yang J, Yu H, Liu BH, Zhao Z, Liu L, Ma LX, Li YX, Li YY: **DCGL v2.0: an R package for unveiling differential regulation from differential co-expression.** *PLoS One* 2013, **8**:e79729.
18. Spencer MJ, Walsh CM, Dorshkind KA, Rodriguez EM, Tidball JG: **Myonuclear apoptosis in dystrophic mdx muscle occurs by perforin-mediated cytotoxicity.** *J Clin Invest* 1997, **99**:2745–2751.
19. McDouall RM, Dunn MJ, Dubowitz V: **Nature of the mononuclear infiltrate and the mechanism of muscle damage in juvenile dermatomyositis and Duchenne muscular dystrophy.** *J Neurol Sci* 1990, **99**:199–217.
20. Cai B, Spencer MJ, Nakamura G, Tseng-Ong L, Tidball JG: **Eosinophilia of dystrophin-deficient muscle is promoted by perforin-mediated cytotoxicity by T cell effectors.** *Am J Pathol* 2000, **156**:1789–1796.
21. An HB, Zheng HC, Zhang L, Ma L, Liu ZY: **Partial least squares based identification of Duchenne muscular dystrophy specific genes.** *J Zhejiang Univ Sci B* 2013, **14**:973–982.
22. Burr AR, Millay DP, Goonasekera SA, Park KH, Sargent MA, Collins J, Altamirano F, Philipson KD, Allen PD, Ma J, Lopez JR, Molkentin JD: **Na<sup>+</sup> dysregulation coupled with Ca<sup>2+</sup> entry through NCX1 promotes muscular dystrophy in mice.** *Mol Cell Biol* 2014, **34**:1991–2002.
23. Davis DB, Delmonte AJ, Ly CT, McNally EM: **Myoferlin, a candidate gene and potential modifier of muscular dystrophy.** *Hum Mol Genet* 2000, **9**:217–226.
24. Cheriya V, Glaser KB, Waring JF, Baz R, Hussein MA, Borden EC: **G1P3, an IFN-induced survival factor, antagonizes TRAIL-induced apoptosis in human myeloma cells.** *J Clin Invest* 2007, **117**:3107–3117.
25. Cerella C, D'Alessio M, De Nicola M, Magrini A, Bergamaschi A, Ghibelli L: **Cytosolic and endoplasmic reticulum Ca<sup>2+</sup> concentrations determine the extent and the morphological type of apoptosis, respectively.** *Ann N Y Acad Sci* 2003, **1010**:74–77.
26. Blake DJ, Weir A, Newey SE, Davies KE: **Function and genetics of dystrophin and dystrophin-related proteins in muscle.** *Physiol Rev* 2002, **82**:291–329.

**Submit your next manuscript to BioMed Central and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

