# Genetic Regulation of Gelsolin in Lung in Mouse Model and its Potential Broad Spectrum of Biological Functions

Xiaoyun Liu[1,2], Helin Feng[3,4], Jiaqian Zhu[5], Rachel Stein[4,6], Yue Huang[4], Yan Jiao[1,4*] and Xiaodong Zhu[1]

[1]Mudanjiang Medical College, Mudanjiang, HeilongJiang, P.R China
[2]Department of Pharmacology, Zhongshan School of Medicine, Sun Yat-sen University, Guangzhou, Guangdong, P.R China
[3]The Fourth Hospital, Hebei Medical University, Shijiazhuang, Hebei, P.R China
[4]Department of Orthopedic Surgery and BME, Campbell-Clinic, University of Tennessee Health Science Center, Memphis, Tennessee, USA
[5]Department of Biological Science, Rust College, Holly Springs, Mississippi, USA
[6]Department of Medicine, University of Tennessee Health Science Center, Memphis, Tennessee, USA

## Abstract

Considerable studies have been done on the potential biological function of gelsolin and its connection to human immune system, diseases and other disorders. The objective of our study was to identify genetic factors that regulate gelsolin in mouse lung and analyze its function immune system using well defined recombinant inbred strains. For this purpose we chose the BXD recombinant inbred (RI) strains derived from progeny of the C57BL/6J (B6) and DBA/2J (D2) progenitor strains. Whole genome gene expression in lung was used for the eQTL mapping. Bioinformatics tools and genotyping data were used for the candidate gene analysis. Gene network and correlation processes were used to assess the association between gelsolin and biological traits. Data indicated that an eQTL on chromosome 9 covering a genomic area between 21Mb and 30Mb is a major play in regulation of the gelsolin expression level. Analysis of genetic elements within this region revealed that Ncapd3 is the most favorite candidate gene. Its expression level is highly associated to that of gelsolin. The expression level of gelsolin between mouse strains with two genotype of SNP (rs13480109) in a regulatory region of the *Ncapd3* showed a significant difference. Additional association analysis suggest that gelsolin may has a broad spectrum of biological function. The expression level of gelsolin has very high correlation with genes in a variety of biological function. These highly associated genes are mainly for protein binding. The expression of gelsolin is also correlated to multiple known immune phenotypes. These data contribute significantly to our current knowledge on the biological function of gelsolin.

**Keywords:** Candidate genes; eQTL; Gelsolin; Gene network; Immune system; Mouse

## Introduction

Gelsolin is an actin-binding protein that is a key regulator of actin filament assembly and disassembly [1]. Considerable studies have been done on the potential biological function and its connection to human immune system, diseases and other disorders. Animal models such as mouse knockout model have also been created for illustration of molecular pathways and its impact on the diseases. In spite of the tremendous progress, its genetic regulation and molecular pathways is still elusive.

Gelsolin expression was linked to cancer for almost two decades ago. By examined gelsolin expression in 12 cultured non-small cell lung cancer (NSCLC) cell lines, Akita et al. concluded that frequent loss of gelsolin expression may be involved in the development of NSCLCs as a potential molecular target of tobacco-induced carcinogenesis [2]. Late Shieh et al. reported that gelsolin expression appears to be a significant prognostic factor for cancer recurrence in cases of Stage I NSCLC [3]. Most recently, gelsolin was proposed as an important cellular target for cotinine, through which this compound influences on the basic processes involved in neoplastic transformation and metastasis, such as migration and apoptosis [4].

The molecular pathway between lung cancer and gelsolin was also studied. Sagawa reported that gelsolin suppresses tumorigenicity through inhibiting PKC activation in a human lung cancer cell line, PC10 [5]. By Proteomic analysis of a neoplastic mouse lung epithelial cell line whose tumorigenicity has been abrogated by transfection with the gap junction structural gene for connexin 43, *Gja1*, Peebles et al. found that *Gja1* transfection affected the concentrations of four proteins: PDI, alpha-enolase, aldolase A, and gelsolin-like protein [6]. Recently, *Nm23-h1* was found to bind to gelsolin and inactivates its actin-severing capacity to promote tumor cell motility and metastasis [7].

Lack of gelsolin has been linked to inflammation and fibrosis. By investigation of the in vivo function of gelsolin in the transgenic gelsolin-null (Gsn-) mice, Witke et al. found that gelsolin is required for rapid motile responses in cell types involved in stress responses such as hemostasis, inflammation, and wound healing [8]. Also using genetically-modified mice lacking gelsolin expression, Oikonomou concluded that gelsolin expression is necessary for the development of modelled pulmonary inflammation and fibrosis, while the caspase-3-mediated gelsolin fragmentation was shown to be an apoptotic effector mechanism in disease pathogenesis and a marker of lung injury [9].

Most importantly, in addition to its role in known diseases and disorders, gelsolin is important in human health and host defense against several infections [10,11]. For example, García-Expósito and colleagues found that, through its severing of cortical actin and by controlling the amount of actin available for reorganization during HIV-1 Env-mediated viral fusion, entry and infection, gelsolin can constitute a barrier that restricts HIV-1 infection of CD4+

lymphocytes in a pre-fusion step [12]. A study indicated that group B Streptococcus (GBS)-β-haemolysin is solely responsible for gelsolin increase causing, through membrane permeability defects, calcium influx and calpain activation [13]. Recently, Yang et al. found that Plasma gelsolin improves lung host defense against pneumonia by enhancing macrophage NOS3 function. [14].

Mouse models have been widely used in the study of molecular pathways of important genes. In this study, we were able to take advantage of whole genome expression profiles of recombinant inbred (RI) strains derived from C57BL/6J (B6) X DBA/2J (D2), known as BXD, for the identification of gelsolin gene pathways and regulatory locus, and its association to diseases [15]. The combined BXD strain set is the largest mouse RI mapping panel currently available. By analysis of the genetic composition of the QTL we were able to identify a candidate gene and to identify a specific mutation responsible for the effect of the QTL.

The aim of this study was to identify genetic factors that regulate gelsolin in mouse lung and analyze its function immune system using RI strains derived from BXD. In order to do so, several analytic experiment were conducted. The quantities trait loci for expression of gelsolin was mapped. The potential genes that regulate the expression level of gelsolin was analyzed. The gene network of gelsolin were constructed. Finally the potential function of gelsolin and its gene network in immune system was analyzed.

## Materials and Methods

### Mouse gene expression data sets from recombinant inbred strains

Gene expression data from the mouse lung in recombinant inbred (RI) strains and standard inbred strains was obtained with Affy Mouse Genome 430 2.0 (GPL1261) [16] (http://www.genenetwork.org/webqtl/main.py? FormID=sharinginfo&GN_AccessionId=160). The data set includes the whole gene expression profiles from 61 mouse strains, including 47 RI strains from BXD (derived from C57BL/6J and DBA/2J), two parents, two F1s, and 10 standard inbred strains.

### QTL/interval analysis

QTL mapping was conducted using publically available software on GeneNetwork (http://www.genenetwork.org/webqtl/main.py). One important feature of the GeneNetwork is WebQTL, which is the leading GeneNetwork module, and has been optimized for on-line analysis of traits that are controlled by combinations of allelic variants and environmental factors [15]. A simple graphical user interface enables rapid, intuitive mapping, and analysis of the reconstructed network [16]. For mapping, 5,000 permutation tests were conducted to determine statistical significance. The threshold was computed by evaluating the distribution of highest LRS scores generated by a set of 5,000 random permutations of strain means (http://www.genenetwork.org/glossary.html). The significant threshold by this calculation represents the approximate LRS value that corresponds to a genome-wide p-value of 0.05, or a 5% probability of falsely rejecting the null hypothesis that there is no linkage anywhere in the genome [17]. The suggestive threshold represents the approximate LRS value that corresponds to a genome-wide p-value of 0.63, or a 63% probability of falsely rejecting the null hypothesis that there is no linkage anywhere in the genome [17,18]. After we obtained the initial loci of QTL from WebQTL from GeneNetwork [17], we further examined the genomic regions of QTL of significant loci for potential associated genes [19,20].

### Analysis for candidate genes within QTL

Using the term gelsolin, we identified potential candidate genes for QTL with PGMapper (http://www.genediscovery.org/pgmapper/index.jsp) [21]. Mouse strain-specific single nucleotide polymorphism (SNP) information of potential candidate genes was obtained using NCBI (http://www.ncbi.nlm.nih.gov/projects/SNP/MouseSNP.cgi) and The Jackson Laboratory database (http://www.informatics.jax.org/javawi2/servlet/WIFetch?page=snpQF). The association of genotype and phenotype e.g. the expression level of gelsolin then was examined.
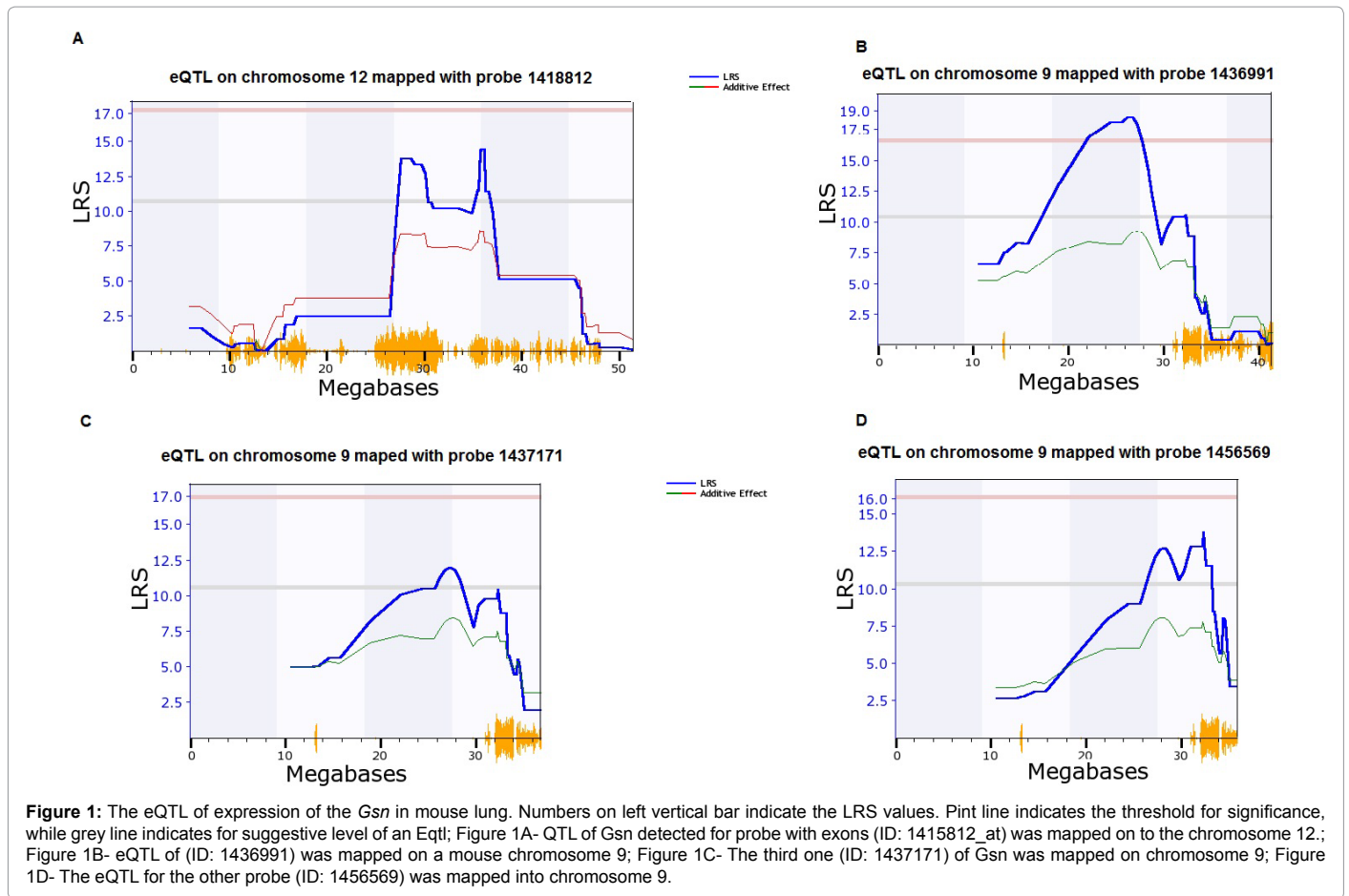
### Gene network analysis methods

Information related to expression level and gene association and network were all analyzed using the Gene Network [16,22]. In case of multiple probes, we first used all the probes in the construction of the gene network. If they are all highly positively correlated. Then one probe (usually the one with the highest expression level) is used for the final construction of the gene network. Graphic connections and associations within the gene network are accomplished with the Metrix and Network Graphs at GeneNetwork. Initial edge lengths were computed by applying an R-to-Z transform to the correlation coefficients and then inverting the results. The graph drawing algorithm found a configuration that minimizes the total stretching of the edges. Curves show Pearson correlation coefficients >0.4 or <-0.4. The graph's canvas is 40.0 by 40.0 cm, and the node labels are drawn with a 16.0 point font, and the edge labels are drawn with a 16.0 point font. A graph based on literature correlation was constructed by selecting the commend of literature in the graphic section. Based on Gene Network, the literature correlation is defined as a measure of the similarity of words used to describe genes. Sets of words that are associated with genes are compared using latent semantic indexing methods. Sets of words associated with genes are extracted from MEDLINE/PubMed abstracts (www.genenetwork.org) [16,22].

## Results

### Transcriptome mapping of expression quantitative trait loci (eQTL) that regulates expression level of gelsolin

We first mapped the Expression Quantitative Trait Loci (eQTL) that regulates the expression level of gelsolin using data of the mouse model, the HZI Lung M430v2 (Apr08) RMA Database [23]. The dataset has seven probes of gelsolin. One from antisense in intron 1 or promoter, one from last four exons and proximal half of 3' UTR, four from 3' UTR and one probe of poor specificity. The eQTL for probe with exons (ID: 1415812_at) was mapped on to the chromosome 12, with maximum LRS of 14.4 which is higher than the suggestion level (10.66) and lower than the significant level (17.22) (Figure 1A). The peak region (area covers the highest LRS score) of the eQTL is located between 25 Mb and 38 Mb, which contains 94 genetic elements.

The eQTL for three probes from 3' UTR was mapped on to chromosome 9. One probe (ID: 1436991) was mapped on a genome region between with a LSR level of 18.1, which is higher than that of significant threshold (thresholds for suggestion level 10.42, significant level 16.64). The peak region is located between 21 Mb and 28 Mb, which contains 82 genomic elements (Figure 1B). The eQTL for the other probe (ID: 1456569) was mapped into chromosome 9 in a similar genomic region, with LRS of 13.8, which is higher than the suggestion level (10.31) and lower than the significant level (16.11) (Figure 1D). The third one (ID: 1437171) has a LRS of approximately 12.0 (threshold for suggestion level 10.63, significant level 17.12) (Figure 1C).

**Figure 1:** The eQTL of expression of the *Gsn* in mouse lung. Numbers on left vertical bar indicate the LRS values. Pint line indicates the threshold for significance, while grey line indicates for suggestive level of an Eqtl; Figure 1A- QTL of Gsn detected for probe with exons (ID: 1415812_at) was mapped on to the chromosome 12.; Figure 1B- eQTL of (ID: 1436991) was mapped on a mouse chromosome 9; Figure 1C- The third one (ID: 1437171) of Gsn was mapped on chromosome 9; Figure 1D- The eQTL for the other probe (ID: 1456569) was mapped into chromosome 9.

The mapping for the fourth probe (ID: 1456312) for the 3' UTR produced multiple suggestive loci on chromosome 2, 6, and 16 (Supplementary Figure S1).

The mapping for the rest two probes did not produce LRS more than the suggestive level at any locus.

## Initial analysis of candidate genes for regulation of expression of gelsolin on mouse chromosome (Chr) 9

Because three probes of gelsolin located the eQTL onto the same location on Chr 9, we further examined the genes within the eQTL region to identify the potential candidate genes. We first examined the association of expression levels of these genes in the eQTL with that of gelsolin. According to the map based on probe #1436991, the peak region of the eQTL is between 21Mb and 28Mb. The other two probes mapped the peak region near 30Mb. To make sure we did not miss the real candidates, we examined the genomic region between 21Mb and 30 Mb. Within this region, there are 103 genetic elements (Table S1) including 69 known genes, (*Acad8    Acp5    Adamts15*

| | | | | | | |
|---|---|---|---|---|---|---|
| *Adamts8* | *Anln* | *Ap1m2* | *Aplp2* | *Atg4d* | *B3gat1* | *Barx2* |
| *Bbs9* | *Bmper* | *Carm1* | *Cnn1* | *Cypt10* | *Cypt4* | *Cypt9* |
| *Dnm2* | *Dock6* | *Dpy19l1* | *Dpy19l2* | *Ecsit* | *Edg8* | *Eepd1* |
| *Elavl3* | *Elof1* | *Epor* | *Glb1l2* | *Glb1l3* | *Gm1110* | *Herpud2* |
| *Hnt* | *Igsf9b* | *Ilf3* | *Jam3* | *Kank2* | *Keap1* | *Kri1* |
| *Ldlr* | *Ncapd3* | *Nfrkb* | *Npsr1* | *Opcml* | *Pde4a* | *Pigyl* |
| *Prdm10* | *Prkcsh* | *Qtrt1* | *Rab3d* | *Rgl3* | *rp9* | *Sept7* |
| *Slc44a2* | *Smarca4* | *Snx19* | *Spata19* | *Spc24* | *St14* | *Tbx20* |

| | | | | |
|---|---|---|---|---|
| *Thyn1* | *Tmed1* | *Tmem205* | *Tmem45b* | *Vps26b* |
| *Yipf2* | *Zbtb44* | *Zfp653* | *Zfp809* | *Zfp810* ). |

For the convenience of identifying genes linked to gelsolin, we used the circular layout model for the building of gene network in this case. We used three probes of gelsolin, 1436991 and 1456569, 1437171, to build the gene network with all 69 genes (Table S2). From the whole genome gene expression profile, we identified a total of 108 probes which represent the 69 genes. Correlation matrix analysis indicated that genes *Dnm2, Pigyl, Smarca4* and Prkcsh are significantly negatively correlated to the expression levels of gelsolin, while genes *Herpud2, Ncapd3, Sept7*, and *Zbtb44* were significantly positively associated to the expression level of gelsolin (Figure 2A).

Among all the genetic elements, only 38 have SNP polymorphism between the two parental strains. (*Edg8*

| | | | | |
|---|---|---|---|---|
| *AB124611* | *Ldlr* | *6530413G14Rik* | *Acp5* | *Zfp809* |
| *Anln* | *E130101E03Rik* | *Herpud2* | *Eepd1* | *Glb1l3* | *Gm1110* |
| *Ncapd3* | *Adamts8* | *5033425B01Rik* | *Ilf3* | *Dock6* | *Jam3* |
| *Zbtb44* | *St14* | *Dnm2* | *Dpy19l1* | *Snx19* | *Bmper* | *Tbx20* |
| *Adamts15* | *Aplp2* | *Nfrkb* | *Npsr1* | *Dpy19l2* | *Barx2* |
| *EG666539* | *Tmem45b* | *Bbs9* | *Opcml* | *Prdm10* |
| *Hnt* | *4930434F21Rik*) |

Among genes that have polymorphism and the genes significantly affect the expression levels of Gsn, Ncapd3 is the only gene that are both correlated to the expression level of Gsn and has a polymorphism. Therefore, Ncapd3 is treated as the most favorite candidate gene. Figure 2B shows that the expression level of Ncapd3 is positively correlated
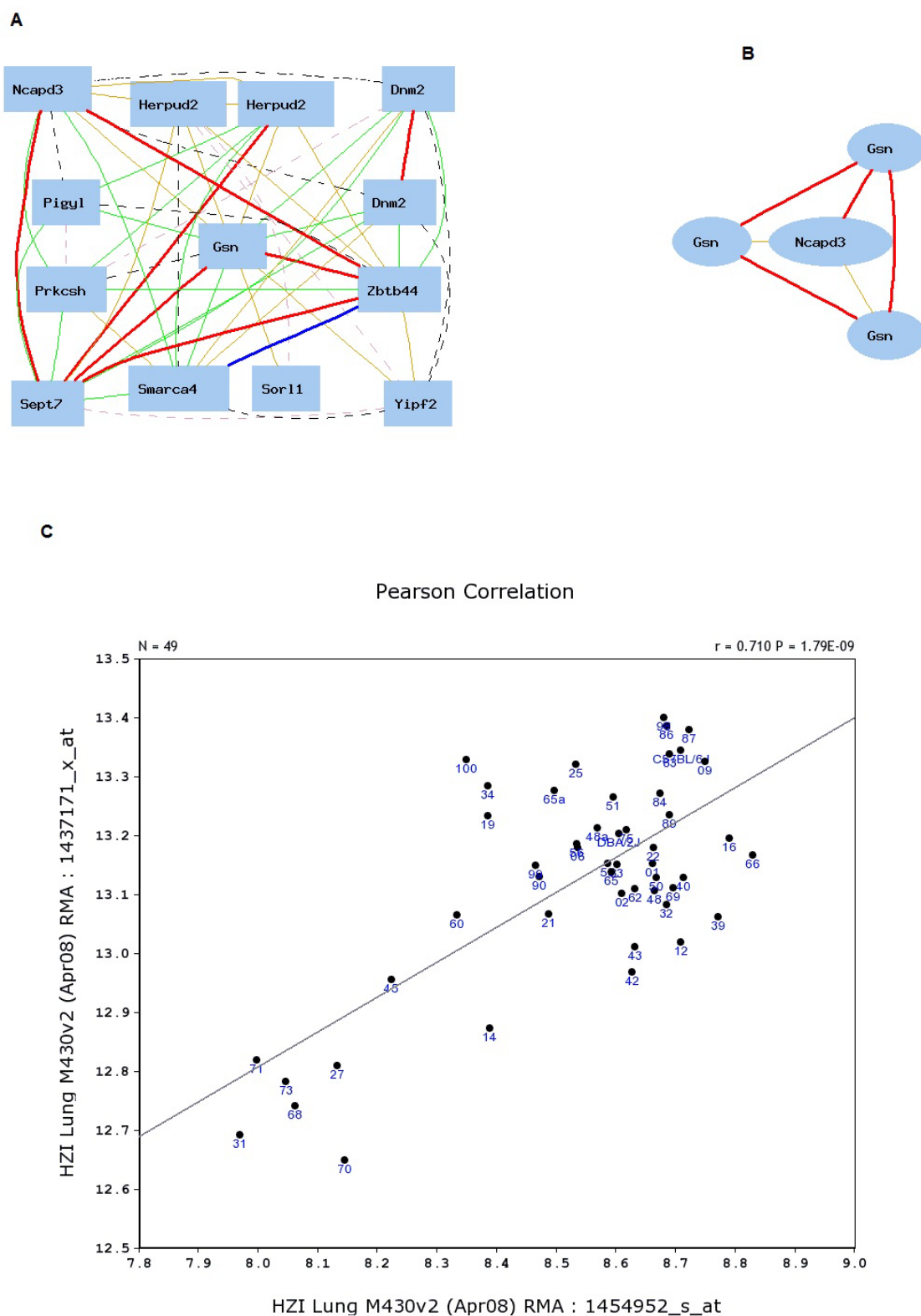
**Figure 2:** Candidate genes and their correlation with gelsolin in expression levels; Figure 2A-Gene network of Gsn and candidate genes for eQTL on chromosome 9. The 13 nodes in the graph below show the selected traits. The 54 edges between the nodes, filtered from the 78 total edges and drawn as curves, show Pearson correlation coefficients greater than 0.4 or less than -0.4; Figure 2B-Genenetwork of three probes of *Gsn* and *Ncapd3*. The 4 nodes in the graph below show the selected traits. The 6 edges between the nodes, filtered from the 6 total edges and drawn as curves, show Pearson correlation coefficients greater than 0.5 or less than -0.5; Figure 2C-The positive correlation between the expression levels of Gsn (Probe: 1437171) and Ncapd3 (Probe: 1454952).

to all three probes at the 3' UTR of the Gsn gene. Figure 2C shows the correlation between the Gsn and Ncapd3 in 57 mouse strains.

## Correlation between genotype of *Ncapd3* and *Gsn* expression level in mouse lung

We next examined the SNPs of Ncapd3 among the RI strains in the BXD mice. In the genome map of GeneNetwork, Ncapd3 starts at 26837759 bs and ends at 26902900 bs. There is a SNP (rs 13480109) in the up-stream of the gene, located at 25697555 bs. The SNP is located in

a non-gene region. According to MGI (http://www.informatics.jax.org/javawi2/servlet/WIFetch?page=snpQF), the SNP has a polymorphism between C57BL/6J and DBA/2J with A/G changes. We then collected the data on the expression levels of RI strains. The average expression level of Gsn in 22 RI strains with A genotype is 12.632 while the 23 RI strains with G genotypes is 12.842. T test indicated there is a significant difference between these two groups, with a P values of 0.0001 (Figure 3).

## Gene network of gelsolin-related genes in lung in mouse model

Using data on the expression level of Probe 1436991 and whole genome expression profiles in the RI strains, we identified the top 100 probes of genes with expression levels most correlated to that of gelsolin from the mouse data of HZI Lung M430v2 (Apr08) RMA Database (Table S3 and Figure S2). The top 100 correlations ranked by the Genetic Correlation (Spearman's rho) showed three distinguishable characterizations. The first one is that the levels of correlations of expression level of gesoline to these genes are very high. The rho values of the first 100 genes with gesoline are all at the significant level, with the rho value of 0.764 as the minimum. The second characterization is that all of the top 100 genes are positively correlated to the expression of gelsoline. The third characterization is that their molecular function are mostly on the protein binding (Supplementary Figure S2). Using the WEB-based GEne SeT AnaLysis Toolkit [24], we examined the biological function of these top 100 genes. Surprisingly, 62 of these 100 top genes are involved in molecular binding, including 32 for heterocyclic compound binding, 32 for organic cyclic compound binding, and 43 for protein binding. There are 33 genes involved in macromolecular complex (Figure 4). None of these genes involves in



**Figure 3:** The expression levels of gelsolin in mouse strains with different genotypes. Numbers on left vertical bar indicate the relative expression levels of gelsolin. A="A" polymorphism of SNP (rs 13480109); G="G" polymorphism of SNP. The expression level of mouse strains with "G" genotype are significantly higher than that of the A genotype.
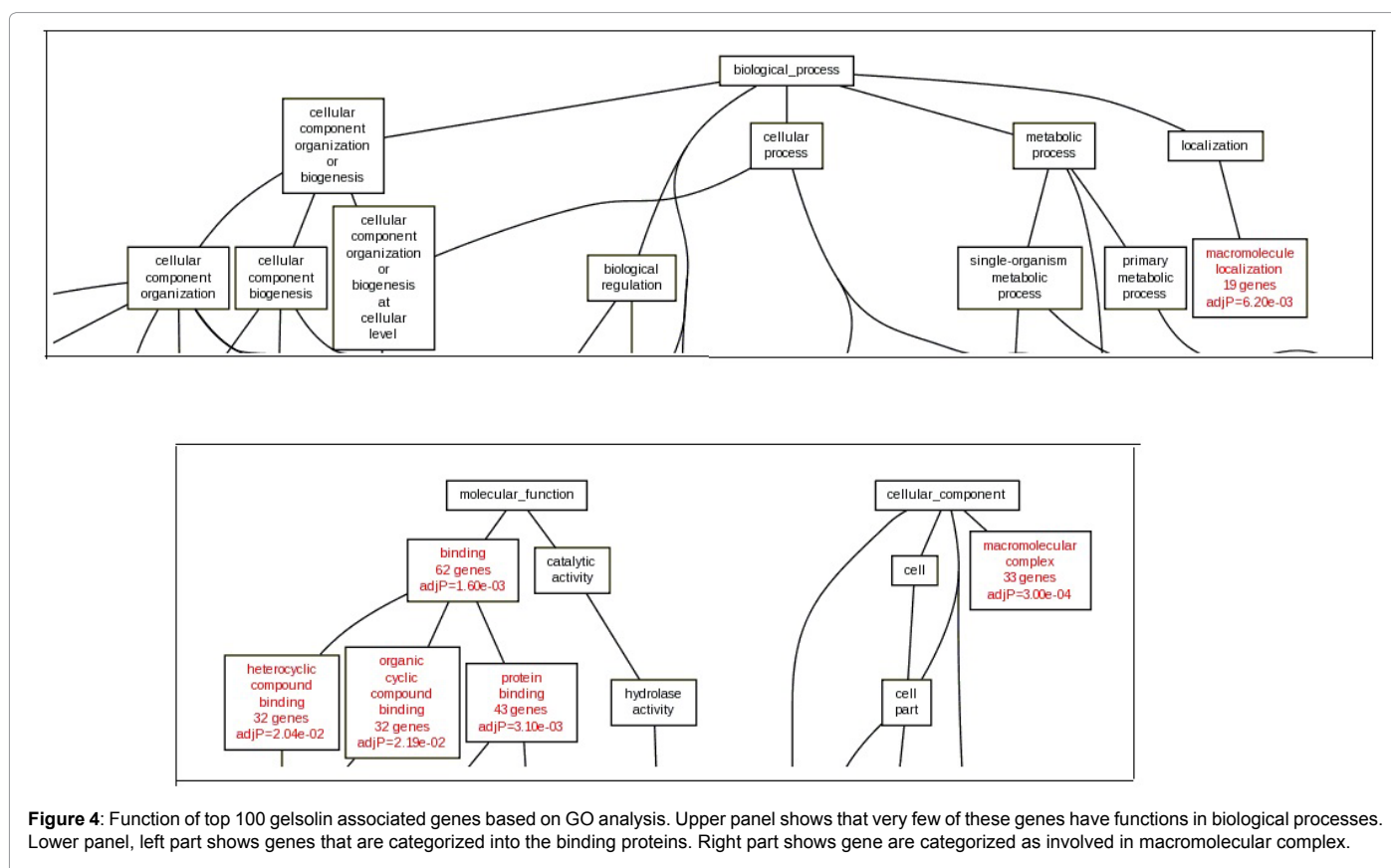


**Figure 4:** Function of top 100 gelsolin associated genes based on GO analysis. Upper panel shows that very few of these genes have functions in biological processes. Lower panel, left part shows genes that are categorized into the binding proteins. Right part shows gene are categorized as involved in macromolecular complex.
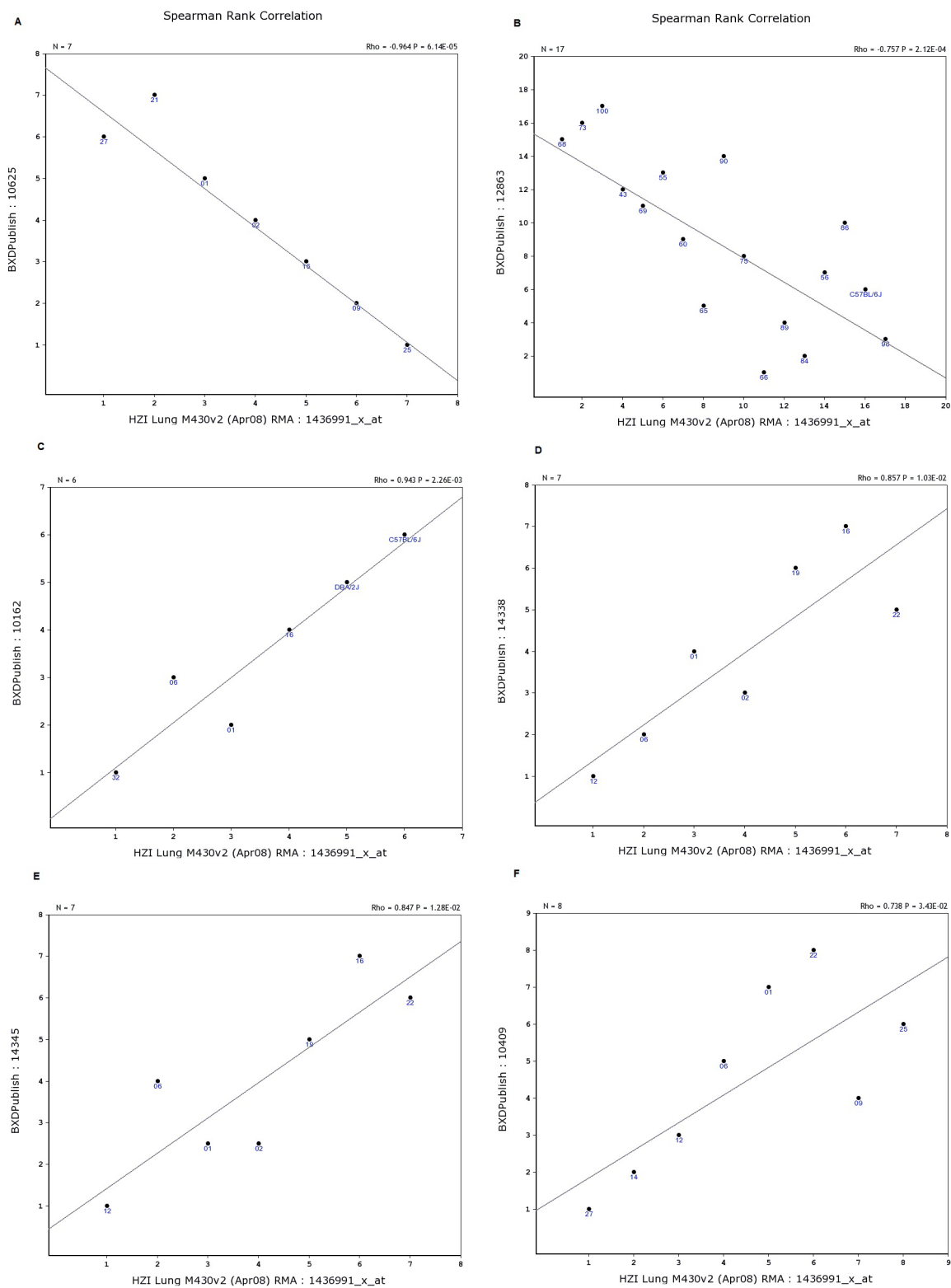
**Figure 5:** Correlation between the expression level of gelsolin and immune function; Figure 5A-The expression of gelsolin is highly negatively correlated to the Protein kinase C (PKC) activity for cortex particulate fraction; Figure 5B-The expression of gelsolin is negatively correlated to the heart rate during echocardiography of 17-week old females; Figure 5C-The expression of gelsolin is negatively correlated to the low-density lipoprotein (LDL) cholesterol levels; Figure 5C-The expression of gelsolin positively correlated to G-CSF (granulocyte colony stimulating factor) (log); Figure 5E-The expression of gelsolin is positively correlated to TNF-a (log) at 72 hour after inoculation with $1 \times 10^6$ cfu Candida albicans (ATCC 10231) IV via retroorbital sinus immediately after light onset; Figure 5F-The expression of gelsolin is positively correlated to Antigenic activity of irradiated BXD spleen cells for Thy-1+CD3+CD4+CD8- T-cell clone TGVH 9.

cellular component organization and cellular process. These unique features of gelsolin associated genes provide fundamental information on the potential biological function of gelsolin. It may function by partnering with or binding to other genes/protein, thus, a much broad regulatory role may be played by gelsolin.

## Potential regulatory roles of Gsn for autoimmune system and broad spectrum of biological functions

Previously, function of gelsolin has been mainly linked to the plasma gelsolin protein; we examined the potential link between expression level of gelsolin in lung and disorders. By examine the association of gelsolin and phenotypic data, potential regulatory role of gelsolin for several immune and other systems have been found.

First of all, the expression of gelsolin (Probe 1436991) is highly negatively correlated to the Protein kinase C (PKC) (GeneNetwork ID: 2400904) activity for cortex particulate fraction (Figure 5A) [25], which agrees with the report that gelsolin suppresses tumorigenicity through inhibiting PKC activation in a human lung cancer cell line [5].

Secondly, it is negatively correlated to the heart rate during echocardiography of 17-week old females (Figure 5B) [26]. Surprisingly, it is negatively correlated to the low-density lipoprotein (LDL) cholesterol levels (Figure 5C) [27].

In addition, expression of gelsolin is positively correlated to several factors in infectious and immune system. It is positively correlated to G-CSF (granulocyte colony stimulating factor) (log) (Figure 5D) (GeneNetwork ID: 14338) and TNF-a (log) (Figure 5E) at 72 hour after inoculation with $1 \times 10^6$ cfu Candida albicans (ATCC 10231) IV via retroorbital sinus immediately after light onset [28]. It is positively correlated to Antigenic activity of irradiated BXD spleen cells for Thy-1+CD3+CD4+CD8- T-cell clone TGVH 9 (Figure 5F) [29].

## Discussion

Our data demonstrated that *Ncapd3* is very likely a major regulator for the expression level of gelsolin in lung. The eQTL on chromosome 9 is confirmed by the mapping on to the same location with three probes of gelsolin. The likelihood of the candidacy of *Ncapd3* for the eQTL is supported by the fact that its expression level is highly association with the expression of gelsolin and it has a polymorphism between the two parental strains of the mapping RI strains. The expression levels of gelsolin of RI strains of two genotypes based on the SNP polymorphism at the regulatory region of the *Ncapd3* showed a significant difference. *Ncapd3* has been reported as a non-structural maintenance of chromosomes (SMC) condensin II subunits [30]. gelsolin is known as an actin-binding protein. The high association between gelsolin and many other genes suggests that gelsolin may has a broad spectrum in regularity of gene expression. The function of *Ncapd3* in maintenance and regulation of chromosome density may directly connect to the expression level of gelsolin. This novel finding needs further study, especially in consideration of how the regulation related to the cancer and inflammation.

The other significant finding from our study is the great potential of broad spectrum of biological function of gelsolin. The expression of gelsolin is very highly correlated to many genes. All of these correlations positive. Gelsolin has also been linked to variety of biological phenotypes. Previously, gelsolin's function has been known in actin filaments [1], inflammation [8,9] and cancer diseases [2-7]. Our finding suggests that gelsolin may have a much broad function such as in metabolism and infections as well. Considering the fact that

gelsolin distributes in many organs and tissues, it is most likely many more biological functions related to gelsolin will be discovered in the near future.

The GO analysis provides a clue on how gelsolin function in multiple biological processes. The fact that 62 of the 100 top gelsolin associated genes are involved in molecular binding indicating that binding to other molecules is the major function of the gelsolin. Through regulation of these binding molecules, gelsolin regulates the biological function of multiple biological processes. The additional 33 genes that are involved in macromolecular complex is the additional emphases that binding in the macromolecular complex in part of binding function of these binding genes. It is clear that gelsolin in general does not involves in the regulation of genes in cellular component organization and cellular process. These information clarify the future research direction on the study of biological function of gelsolin.

The expression of gelsolin is mainly regulated by the eQTL on chromosome 9. There are other loci, such as the eQTLs on chromosome 12, 2, 6, and 16 which may also influence the expression level of gelsolin. Although the expression level of gelsolin may be regulated by multiple factors, the influence by these loci is relatively small and need to be confirmed by additional studies. Gelsolin expresses in multiple tissues at different life stages [1,10,31,32]. The molecular mechanism in lung may not be necessary the same in other tissues. Thus, the mechanism in lung may be the same, similar or totally different from other tissues. Considerable amount of studies may be needed in order to understand the regulatory mechanism of gelsolin in human tissues.

Previous studies have been mainly focused on the expression of plasma secreted protein of gelsolin [1,8,10,31,32]. Our study reveals the molecular mechanism of gelsolin at transcript level of whole tissue of lung. Therefore, our findings may be different from other studies in term of the biological and disease association. Nevertheless, our study suggests that the transcriptional level of gelsolin in lung is an important subject for the understanding of its role in the biological function and disease.

How this novel data may apply to the human population will be important task in the future.

### Authors' contributions

Drs. Liu, Feng, Jiao and Zhu-X had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Study concept and design: Jiao, Zhu-X,

Data collection and analysis: Liu, Feng, Zhu-J, Huang, Stein

Data interpretation of data: All authors.

Drafting of the manuscript: Liu, Feng, Zhu-J, Jiao.

Critical revision of the manuscript for important intellectual content: All authors.

Literature search: Liu, Feng, Stein, Jiao

Obtained funding: Jiao, Zhu-X.

## References

1. Yin HL, Zaner KS, Stossel TP (1980) Ca2+ control of actin gelation. Interaction of gelsolin with actin filaments and regulation of actin gelation. J Biol Chem. 255: 9494-9500.

2. Dosaka-Akita H, Hommura F, Fujita H, Kinoshita I, Nishi M, et al. (1998) Frequent loss of gelsolin expression in non-small cell lung cancers of heavy smokers. Cancer Res. 58: 322-327.

3. Shieh DB, Godleski J, Herndon JE, Azuma T, Mercer H, et al. (1999) Cell motility as a prognostic factor in Stage I nonsmall cell lung carcinoma: the role of gelsolin expression. Cancer. 85: 47-57.

4. Nowak JM, Klimaszewska-Wisniewska A, Izdebska M, Gagat M, Grzanka A (2015) gelsolin is a potential cellular target for cotinine to regulate the migration and apoptosis of A549 and T24 cancer cells. Tissue Cell 47: 105-114.

5. Sagawa N, Fujita H, Banno Y, Nozawa Y, Katoh H, et al. (2003) gelsolin suppresses tumorigenicity through inhibiting PKC activation in a human lung cancer cell line, PC10. Br J Cancer 88: 606-612.

6. Peebles KA, Duncan MW, Ruch RJ, Malkinson AM (2003) Proteomic analysis of a neoplastic mouse lung epithelial cell line whose tumorigenicity has been abrogated by transfection with the gap junction structural gene for connexin 43, Gja1. Carcinogenesis 24: 651-657.

7. Marino N, Marshall JC, Collins JW, Zhou M, Qian Y, et al. (2013) Nm23-h1 binds to gelsolin and inactivates its actin-severing capacity to promote tumor cell motility and metastasis. Cancer Res 73: 5949-5962.

8. Witke W, Sharpe AH, Hartwig JH, Azuma T, Stossel TP, et al. (1995) Hemostatic, inflammatory, and fibroblast responses are blunted in mice lacking gelsolin. Cell 81: 41-51.

9. Oikonomou N, Thanasopoulou A, Tzouvelekis A, Harokopos V, Paparountas T, et al. (2009) gelsolin expression is necessary for the development of modelled pulmonary inflammation and fibrosis. Thorax 64: 467-475.

10. Peddada N, Sagar A, Ashish, Garg R (2012) Plasma gelsolin: a general prognostic marker of health. Med Hypotheses 78: 203-210.

11. Li GH, Arora PD, Chen Y, McCulloch CA, Liu P (2012) Multifunctional roles of gelsolin in health and diseases. Med Res Rev 32: 999-1025.

12. García-Expósito L, Ziglio S, Barroso-González J, de Armas-Rillo L, Valera MS, et al. (2013) gelsolin activity controls efficient early HIV-1 infection. Retrovirology 10: 39.

13. Fettucciari K, Ponsini P, Palumbo C, Rosati E, Mannucci R, et al. (2015) Macrophage induced gelsolin in response to Group B Streptococcus (GBS) infection. Cell Microbiol 17: 79-104.

14. Yang Z, Chiou TT, Stossel TP, Kobzik L (2015) Plasma gelsolin improves lung host defense against pneumonia by enhancing macrophage NOS3 function. Am J Physiol Lung Cell Mol Physiol 309: L11-16.

15. Peirce JL, Lu L, Gu J, Silver LM, Williams RW (2004) A new set of BXD recombinant inbred lines from advanced intercross populations in mice. BMC Genet 5: 7.

16. Chesler EJ, Lu L, Shou S, Qu Y, Gu J, et al. (2005) Genetic dissection of gene expression reveals polygenic and pleiotropic networks modulating brain structure and function. Nature Genetics 37: 233-242.

17. Wang J, Williams RW, Manly KF (2003) WebQTL: web-based complex trait analysis. Neuroinformatics 1: 299-308.

18. Calabrese G, Bennett BJ, Orozco L, Kang HM, Eskin E, et al. (2012) Systems genetic analysis of osteoblast-lineage cells. PLoS Genet 8: e1003150.

19. Wang L, Lu W, Zhang L, Huang Y, Scheib R, et al. (2014) Trps1 differentially modulates the bone mineral density between male and female mice and its polymorphism associates with BMD differently between women and men. PLoS One 9: e84485.

20. Wang L, Huang Y, Jiao Y, Chen H, Cao Y, et al. (2013) Expression of alcoholism-relevant genes in the liver are differently correlated to different parts of the brain. Gene Expr 16: 31-38.

21. Xiong Q, Qiu Y, Gu W (2008) PGMapper: a web-based tool linking phenotype to genes. Bioinformatics 24: 1011-1013.

22. Wang L, Jiao Y, Sun S, Jarrett HW, Sun D, et al. (2015) Gene network of a phosphoglycerate mutase in muscle wasting in mice. Cell Biol Int 39: 666-677.

23. Alberts R, Lu L, Williams RW, Schughart K (2011) Genome-wide analysis of the mouse lung transcriptome reveals novel molecular gene interaction networks and cell-specific expression signatures. Respir Res 12: 61.

24. Wang J, Duncan D, Shi Z, Zhang B (2013) WEB-based GEne SeT AnaLysis Toolkit (WebGestalt): update 2013. Nucleic Acids Res 41: W77-83.

25. Wehner JM, Sleight S, Upchurch M (1990) Hippocampal protein kinase C activity is reduced in poor spatial learners. Brain Res 523: 181-187.

26. Andreux PA, Williams EG, Koutnikova H, Houtkooper RH, Champy MF, et al. (2012) Systems genetics of metabolism: the use of the BXD murine reference panel for multiscalar integration of traits. Cell 150: 1287-1299.

27. Colinayo VV, Qiao JH, Wang X, Krass KL, Schadt E, et al. (2003) Genetic loci for diet-induced atherosclerotic lesions and plasma lipids in mice. Mamm Genome 14: 464-471.

28. Toth LA, Trammell RA, Williams RW (2014) Mapping complex traits using families of recombinant inbred strains: an overview and example of mapping susceptibility to Candida albicans induced illness phenotypes. Pathog Dis 71: 234-48.

29. Miconnet I, Huchet R, Bonardelle D, Motta R, Canon C, et al. (1990) Graft-versus-host mortality induced by noncytolytic CD4+ T cell clones specific for non-H-2 antigens. J Immunol 145: 2123-2131.

30. Liu W, Tanasa B, Tyurina OV, Zhou TY, Gassmann R, et al. (2010) PHF8 mediates histone H4 lysine 20 demethylation events involved in cell cycle progression. Nature 466: 508-512.

31. Chaponnier C, Kocher O, Gabbiani G (1990) Modulation of gelsolin content in rat aortic smooth muscle cells during development, experimental intimal thickening and culture. An immunohistochemical and biochemical study. Eur J Biochem 190: 559-565.

32. Ehrlich HP, Allison GM, Page MJ, Kolton WA, Graham M (2000) Increased gelsolin expression and retarded collagen lattice contraction with smooth muscle cells from Crohn's diseased intestine. J Cell Physiol 182: 303-309.