Published in final edited form as:

Proc IEEE Int Symp Bioinformatics Bioeng. 2007 October ; 2007: 433–439. doi:10.1109/BIBE.2007.4375598.

Identifying Genomic Regulators of Set-Wise Co-Expression

Jung Hoon Woo^{*,†,¶}, Tian Zheng^{‡,¶}, and Ju Han Kim^{*,§,||}

^{*}Seoul National University Biomedical Informatics (SNUBI), Seoul National University College of Medicine, Seoul 110-799, Korea

[†]Macrogen Inc., Seoul, Korea

[‡]Department of Statistics, Columbia University, New York, NY, USA

[§]Human Genome Research Institute, Seoul National University College of Medicine, Seoul 110-799, Korea

Abstract

The genetical genomics approach has been used to study the genetic basis of variation in gene expression, where putative transcriptional regulators of genes are identified via genetic quantitative trait mapping. The genetic regulators identified through such efforts can partially account for an individual gene's natural variation. However, genes in a molecular pathway often exhibit coordinated activities, the patterns and levels of which are also regulated. In an effort to understand these complicated mechanisms, we propose a method that searches for the genomic regulators of set-wise co-expression of related genes, based on current genetical genomics data. Using this method, we studied genomic regulators of 233 biological pathways for a BXD RI data set. For 15 pathways, we obtained significant regulatory loci after controlling for the false discovery rate. The results presented in this paper constitute important evidence of the heritability of mRNA co-expression between individuals. We have shown that, by defining new phenotypes using existing genetical genomics data, evidence on regulation of co-expression can be derived.

I. Introduction

A key objective of biomedical research is to elucidate the complex genetic interactions underlying molecular phenotypic variations [1]. The “genetical genomics approach” is a recent development that offers a new perspective on genetic analysis by integrating large scale genotype data with high throughput genomics data such as microarray gene expression profiles. Specifically, this method treats molecular mRNA transcription levels (gene expression profiles) as quantitative traits and identifies regulatory loci that can explain the genetic variation in these molecular traits. In most current genetical genomics studies, each gene's expression is studied as an individual trait and the genetic factors that explain its variation are studied [2]-[6]. Although widely used, this strategy is insufficient to study the complicated molecular mechanism regulating mRNA transcription due to a problematic assumption that genes are individually expressed [7]. It has been observed that the expressions of genes are correlated and may be controlled by the same regulators. For instance, genes on the same pathway usually have correlated gene expressions. Therefore, it is important for genetical genomics methods to consider and investigate, in the analysis, such dependence structure.

Several approaches have been proposed to incorporate gene set information, and to identify regulators or regulatory conditions for a set of related genes. Not specifically designed for genetical genomics, Mootha et al. (2003) [8] introduced the gene set enrichment analysis (GSEA) test, which associated the rankings of genes with gene set membership information to see whether certain previously defined molecular gene sets were “enriched” among the top genes associated with an outcome of interest. This strategy has been adopted in genetical genomics studies such as those by Lan et al. [7] and Ghazalpour et al. [9]. In both studies, quantitative trait locus (QTL) mapping was first conducted for each transcript (gene expression trait) individually. For genomic regions with excessive numbers of transcripts mapped to them, GSEA [9] or Gene Ontology (GO) [10] enrichment analysis [7] was conducted to identify regulatory pathways or molecular functional categories that were likely to have set-wise regulators within these regions.

Another way of acknowledging the inter-relatedness among genes in a set is to use mega-phenotypes defined directly on the related genes. Clustering and principal component analysis have been popular tools to capture correlation structure for conventional gene expression analysis (e.g., [11], [12]). In the study by Lan et al. [13], the principal components of the multiple transcripts in a set were used as traits to identify the regulators of the main phenotypic variations for a set of genes.

The above two strategies have focused on identifying the co-regulators of a set of genes (e.g., on a pathway) or the regulators of common variations of a gene set. However, regulators of a molecular pathway may affect not only the expression levels of these genes but also the extent to which they are inter-correlated. Firstly, the interaction of two genes may be regulated or affected by a regulator. Secondly, interacting transcripts that are regulated by a common upper cascade regulator may show high correlation under the functioning variant of the regulator, but not under the non-functioning variant. We hypothesize that if a genomic regulator affects the upper cascade of an interacting gene set (e.g. a pathway), the correlation tendencies among the set should vary as genotypic differences of the regulator (Figure 1).

In this paper, we describe an approach to identify the genomic regulators of set-wise co-expression of related genes according to current biological knowledge. To uncover and define co-expression quantitatively, we use the level of correlation between the expression values of two transcripts, in keeping with current practices [11]. More specifically, in this study, we introduce a differential allelic co-expression (DACE) test to identify regulatory association between an SNP marker and a gene set, by searching for allelic association between an SNP and the correlation levels within a gene set under a linear regression model. Using the DACE test, we identified significant putative genomic regulators that appear to affect mRNA co-expression in 15 biological pathways.

II. Materials and Methods

A. Differential allelic co-expression (DACE) test

Identification of differential allelic co-expression is performed through a series of tests, each conducted between a given SNP marker and a given gene set, to test correlation structure differentiation of mRNA transcripts (of the gene set) arising from the genotypes of an SNP.

Assume that expression phenotypes and SNP genotypes are measured on n subjects. Given an SNP, subjects are divided into its G genotypes. To study a set of p transcripts (expression phenotypes), first compute, within each genotype group, the *Pearson correlation coefficients* [14] of the expression levels between all pairs of transcripts in a set. Denote r_{ijg} as the correlation between transcripts i and j within genotype group g . Correlation coefficients are

not normally distributed [15]. For our test, we perform the “Fisher’s z ’ transformation” [15] on the original correlation values as follows:

$$z_{ijg} = \frac{1}{2} \ln \left(\frac{1+r_{ijg}}{1-r_{ijg}} \right),$$

$$i=1, \dots, p-1, \quad j=i+1, \dots, p,$$

$$g=1, \dots, G. \quad (1)$$

To test whether the SNP under study has a significant effect on the levels of correlation among these genes, we adopt the general framework of a linear model:

$$z_{ijg} = \beta_0 + \beta_1 X_g + \varepsilon_{ijg} \quad (2)$$

where z is defined as above and X_g represents genotype variation. For an SNP with alleles A and B, we code X_g 0 for genotype AA, 1 for genotype AB and 2 for genotype BB. Therefore, the t test for $H_0 : \beta_1 = 0$ versus $H_1 : \beta_1 \neq 0$ searches for significant *allelic association* between an SNP and the trait of interest (correlations).

B. Compilation of gene set

We used pathway information for grouping probes in an array to create a gene set. These data were obtained from publicly available pathway resources (BioCarta [16]) for mapping genes to pathways. A total of 276 pathways were present for the 12,488 probes on the Affymetrix U74Av2 array.

C. BXD RI data set

Recombinant inbred strains were generated by repeated inbreeding of F2 mice derived from 2 parental inbred strains, C57BL/6 (B6) and DBA/2 in the case of BXD RI strains. Each of the BXD strains has evolved a unique combination of homozygous parental alleles; the usefulness of the strains in genetic mapping has led to hundreds of experiments to study the genetics of a wide variety of phenotypes [17]. We assembled expression phenotype and genotype data from different repositories concerning the BXD RI strains.

D. Microarray gene expression data

We downloaded DNA microarray data sets which measure transcriptome expression in hematopoietic stem cell (HSC) across 22 BXD RI strains [6], from National Center for Biotechnology Center (NCBI) Gene Expression Omnibus (GEO). Affymetrix U74Av2 arrays were used for HSC transcriptome profiling. For each chip, we computed 12,488 gene expression values for the Affymetrix data using the robust multichip average (RMA) algorithm [18] which uses background adjustment, quantile normalization and summarization.

E. SNP genotype data

The genotype data of the 22 BXD RI strains of mice were downloaded from webQTL [19] on the GeneNetwork website (<http://www.genenetwork.org>). The original BXD genotype data file included a set of 3,795 markers, both SNPs and microsatellites. We chose 3,033 SNP genotypes for our study.

III. Results

We compiled a new BXD RI data set for the DACE test as follows: as the genotype of each RI strain was fixed, experimental mice samples in each data set can be regarded as the same individual if they were from the same RI strain. 22 BXD RI strain samples were assembled with both gene expression data and genotype data. We assumed regulatory association between a genomic locus and gene set when the extent of inter-correlation significantly varied between genotypes. The DACE test described in the Materials and Methods section was used to identify regulatory association in the BXD RI data set.

As described in the Materials and Methods section, we configured 276 gene sets using pathway information. Due to computational concerns, we restricted ourselves to pathways with fewer than 30 gene members. As a result, the test was performed using 233 pathways and 3,033 SNP markers. The false discovery rate (FDR) [20] was controlled at < 0.001 . 15 pathways showed significant association with at least one SNP locus. Five of these, some of which play critical roles in hematopoietic stem cell (HSC) function, are listed in Table I. A complete list of the association found between the 15 pathways and 132 loci is provided in Supplement 1 [21].

For comparison, we performed the widely used single gene tests on the same data set, in which the association between gene and locus detected by the DACE test cannot be detected. Therefore, the set-wise co-expression regulators identified using our approach are different from (and possibly complementary to) the gene set (or GO) enriched regulatory regions studied by Lan et al. [7] and Ghazalpour et al. [9]. For example, using the DACE test, the *role of Tob in the T-cell activation pathway* showed significant association with SNP markers within the region 84.03-88.47 cM on chromosome 6. On the other hand, none of the 17 genes comprising this pathway showed association with the same SNPs in single gene tests. These results indicate that none of the 17 transcripts have differentiated expression levels across the genotype groups of this locus (Figure 2-b). Rather, the extent and patterns of correlation among the transcripts are significantly different between the two genotype groups (Figure 2-c).

In this paper, we detected SNPs that were associated with the difference in the correlation structure among genes in a pathway. To elucidate the biological meaning for the significant association between genomic loci identified and the correlation relation among the genes in a specific pathway, we examined their positional relation on the genome and in the regulatory network. We hypothesize, as one possible biological interpretation, that if a certain genomic locus regulates the expression of genes in the upper cascade of a specific pathway, this locus might show significant association with the pathway in the DACE test. Therefore, we examined, in the regulatory network of known biological interactions, the positional relation between the significant genomic regulatory loci for a specific pathway identified by the DACE test and the genes in this pathway. More specifically, for a pathway with a significant genomic regulatory locus returned by the DACE test, we located the members of this pathway on the genome and identified those that were adjacent to the identified regulatory locus. Here, we empirically defined *adjacent genes* as those located within 10 Mbp from the physical location of the identified regulatory locus. Among the 15 significant pathways with significant loci identified, 7 pathways have at least one gene member that are adjacent to the regulatory genomic loci returned by the DACE test. The results are provided in Supplement 2 [22].

In most cases, the adjacent genes are located in the upper cascade of their pathways. In the case of the *Cyclin E Destruction pathway*, the “chr2. 149122813 - 159760664” locus was detected as a putative genomic regulator of the pathway by the DACE test (p-value

$<2.59\text{E-}06$). The E2F1 gene, a member of the *Cyclin E Destruction pathway*, has been physically located within the locus of “chr2. 149122813 - 159760664,” next to the identified regulatory locus (Figure 3). By single gene association test, the E2F1 gene and the regulatory locus did not show significant association after correcting for multiple comparison (minimum p-value = 0.0639).

For the DACE test, we used the ordinary simple linear regression to identify differentiation of correlation structure. However, the significance evaluation using simple linear regression might potentially be problematic since the distribution of statistics may depart from normality. To evaluate this issue, we used permutation procedure to obtain empirical p-values. For each permutation, we randomly reassigned the genotypes of a given SNP to the 22 samples and repeated the DACE tests, measuring the association between the correlation structure of the mRNA transcripts and the shuffled genotypes of the SNP. The procedure was repeated 1,000 times to get empirical p-values of the original test results. We performed the permutation test for the 15 previously identified significant pathways and false discovery rate (FDR) was controlled at < 0.05 . Not all association originally identified by the DACE test was significant in the permutation procedure, but strong signals were still captured by the permutation test (Table II).

IV. Discussion and Conclusion

It is well known that complex functions of living cells are performed through the concerted efforts of many genes [23]. Because of this modular nature of gene regulatory networks, it is more biologically relevant to identify regulators of gene sets rather than individual genes. Previous single gene analyses assumed that all genes were individually expressed and looked for regulatory factors for each individual gene, an insufficient approach. In contrast, we developed a set-wise approach (the DACE test) to reveal genomic loci involved in the regulation of co-expression, or of the inter-correlation of mRNA expressions, in gene sets such as those involved in pathways. The DACE test treats a gene set as an analytical unit and examines factors associated with correlated activities within a given pathway. This relationship, between a genomic locus and a gene set, cannot be detected by single gene tests. Having configured our gene set using biological knowledge rather than through an optimizing approach to generate a random set, our results are more biologically relevant. Using biological knowledge of pathways also narrows the computational scope of such studies, leading to higher power and better efficiency.

In our analysis, the genomic regulators identified using the BXD RI data set were not associated with the actual expression levels of the genes in the corresponding pathways, as they were not identified as significant in single gene tests. Instead, they regulated the level of correlations between genes and were only detectable when the set of genes was studied as a unit. (See Figure 2 for an example using the *role of Tob in T-cell activation pathway*). This clearly demonstrates the utility of our approach for biological discovery, i.e., studying the *regulation* of interactive activities within pathways, which cannot be achieved with traditional genetical genomics approaches. It should also be noted that our approach is different from studying gene-gene regulation within a pathway, which focuses on the interactive activities of individual gene pairs genes within a pathway.

A biological pathway is defined as a series of molecular interactions and reactions. If there are subtle changes in the expression level of a few genes located in the upper cascade of a pathway, they may alter the overall co-expression patterns of the pathway. This hypothesis is plausible to explain our DACE results. We found, for 7 of the 15 significant pathways, at least one upper cascade gene that were adjacent to the identified regulatory loci of the pathways. In the case of the *Cyclin E destruction pathway*, the E2F1 gene is high in the

upper cascade of the pathway (Figure 4) and the gene was located right next to the significant regulatory locus of the pathway (Figure 3).

As for the 8 pathways that did not show gene member(s) next to the loci of identified regulators, one possible explanation is that our expression data only covered 8500 genes. We may have missed some upper cascade genes of the 8 pathways near by the regulatory locus just because their expressions were not measured by this platform. Or it is possible that there are other pathway members which have not been discovered yet. Because we only checked the regional relationships based on established knowledge, our results still require further validation to establish true regulatory relation. We believe that our results provide good starting points for pathway augmentation if there is undiscovered gene member nearby an identified regulatory loci of a pathway.

While there have been several studies concerned with co-expression or co-regulation of genes under external stimuli, these studies have yet to take into consideration natural variations, such as inherited genetic variation. Our procedure, designed to measure the influence of genetic variation on the co-regulation of gene expression at a genome-wide scale, offers important evidence of heritability of mRNA co-expression between individuals.

Acknowledgments

This study was supported by a grant of the Korea Health 21 R&D Project, Ministry of Health & Welfare, Republic of Korea (0405-BC02-0604-0004), and in part by NIH grant R01 GM070789.

References

1. Cheung VG, Spielman RS. The genetics of variation in gene expression. *Nature Genetics* 2002;32:522–525. [PubMed: 12454648]
2. Jansen RC, Nap JP. Genetical genomics: the added value from segregation. *Trends in Genetics* 2001;17:388–391. [PubMed: 11418218]
3. Brem RB, Yvert G, Clinton R, Kruglyak L. Genetic dissection of transcriptional regulation in budding yeast. *Science* 2002;296:752–5. [PubMed: 11923494]
4. Schadt EE, Monks SA, Drake TA, Lusis AJ, Che N, Colinayo V, Ruff TG, Milligan SB, Lamb JR, Cavet G, Linsley PS, Mao M, Stoughton RB, Friend SH. Genetics of gene expression surveyed in maize, mouse and man. *Nature* 2003;422:297–302. [PubMed: 12646919]
5. Morley M, Molony CM, Weber TM, Devlin JL, Ewens KG, Spielman RS, Cheung VG. Genetic analysis of genome-wide variation in human gene expression. *Nature* 2004;430:743–747. [PubMed: 15269782]
6. Bystrykh L, Weersing E, Dontje B, Sutton S, Pletcher MT, Wiltshire T, Su AI, Vellenga E, Wang JT, Manly KF, Lu L, Chesler EJ, Alberts R, Jansen RC, Williams RW, Cooke MP, de Haan G. Uncovering regulatory pathways that affect hematopoietic stem cell function using ‘genetical genomics’. *Nature Genetics* 2005;37:225–232. [PubMed: 15711547]
7. Lan H, Chen M, Flowers JB, Yandell BS, Stapleton DS, Mata CM, Mui ET, Flowers MT, Schueler KL, Manly KF, Williams RW, Kendziorski C, Attie AD. Combined expression trait correlations and expression quantitative trait locus mapping. *PLoS Genet* 2(1):e6. [PubMed: 16424919]
8. Mootha VK, Lindgren CM, Eriksson KF, Subramanian A, Sihag S, Lehar J, Puigserver P, Carlsson E, Ridderstrale M, Laurila E, Houstis N, Daly MJ, Patterson N, Mesirov JP, Golub TR, Tamayo P, Spiegelman B, Lander ES, Hirschhorn JN, Altshuler D, Groop LC. Pgc-1 alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics* 2003;34:267–273. [PubMed: 12808457]
9. Ghazalpour A, Doss S, Sheth SS, Ingram-Drake LA, Schadt EE, Lusis AJ, Drake TA. Genomic analysis of metabolic pathway gene expression in mice. *Genome Biology* 2005;6:R59. [PubMed: 15998448]

10. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. the gene ontology consortium. *Nat Genet* 2000;25:25–9. [PubMed: 10802651]
11. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America* 1998;95:14863–14868. [PubMed: 9843981]
12. West M, Blanchette C, Dressman H, Huang E, Ishida S, Spang R, Zuzan H, Olson J, A J, Marks JR, Nevins JR. Predicting the clinical status of human breast cancer by using gene expression profiles. *Proc Natl Acad Sci U S A* 2001;98:11462–7. [PubMed: 11562467]
13. Lan H, Stoehr JP, Nadler ST, Schueler KL, Yandell BS, Attie AD. Dimension reduction for mapping mrna abundance as quantitative traits. *Genetics* 2003;164:1607–1614. [PubMed: 12930764]
14. Pearson K. mathematical contributions to the theory of evolution viii on the correlation of characters not quantitatively measurable. *Proceedings of the Royal Society of London* 1900;66:241–244.
15. Fisher RA. On the ‘probable error’ of a coefficient of correlation deduced from a small sample. *Metron* 1921;1:1–32.
16. Biocarta pathways database. <http://www.biocarta.com/genes/all-Pathways.asp>
17. Chesler EJ, Lu L, Shou SM, Qu YH, Gu J, Wang JT, Hsu HC, Mountz JD, Baldwin NE, Langston MA, Threadgill DW, Manly KF, Williams RW. Complex trait analysis of gene expression uncovers polygenic and pleiotropic networks that modulate nervous system function. *Nature Genetics* 2005;37:233–242. [PubMed: 15711545]
18. Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of affymetrix genechip probe level data. *Nucleic Acids Research* 2003;31:e15.
19. Chesler EJ, Lu L, Wang JT, Williams RW, Manly KF. Webqtl: Rapid exploratory analysis of gene expression and genetic networks for brain and behavior. *Nature Neuroscience* 2004;7:485–486.
20. Benjamini Y, Hochberg Y. Controlling the false discovery rate - a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B-Methodological* 1995;57:289–300.
21. Woo, JH.; Zheng, T.; Kim, JH. Supplement table 1: Gene sets with significantly associated regions as indicated by the dace test. <http://www.snubi.org/publication/dace/>
22. Woo, JH.; Zheng, T.; Kim, JH. Supplement table 2: Significant pathways identified with gene member(s) adjacent to the regulatory genomic loci identified by dace. <http://www.snubi.org/publication/dace/>
23. Segal E, Shapira M, Regev A, Pe’er D, Botstein D, Koller D, Friedman N. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nature Genetics* 2003;34:166–176. [PubMed: 12740579]

Fig. 1.

A hypothetical pathway with an upper cascade regulator. Gene A is in the upper cascade of a four-gene pathway. When one has the variant *red* in the promoter region of gene A on the genome, transcription factor (TF)- α binds and leads to the expression of gene A, which, in turn, regulates genes B, C and D and their interactions. When the variant *blue* is present in the promoter of A, TF- β binds and gene A is suppressed. Genes B, C and D will show different interacting patterns due to other regulation they receive.

Fig. 2.

Genome-wide DACE test results for the *role of Tob in the T-cell activation pathway*. The DACE test was conducted between individual SNPs on the genome and the gene set constructed from pathway information. As some genes are represented by more than one probe set, they are indicated more than one times in this figure. (a) Genome-wide distribution of negative log10 of p-values from the DACE test for the *role of Tob in T-cell activation pathway*. The horizontal red line is our threshold for evidence of significant regulatory association ($FDR < 0.001$). Eleven SNP markers within the region 84.03-88.47 cM in chromosome 6 are above the threshold. P-values of the markers in a relatively large region were identical because the region was in complete linkage disequilibrium (LD). (b) Heatmap for the gene expression matrix. There is no significant change of expression level between two groups when each gene is considered individually. Therefore, none of the 17 genes showed association with the region in single gene tests. (c) Heatmaps for the correlation matrices. Hierarchical clustering was done on the genes to show clearer patterns. The two heatmaps show significant change of cofelation tendencies in the *role of Tob in the T-cell activation pathway* between different groups defined by the genotypes at the identified significant locus.

Fig. 3.
Physical map of the genomic regulator identified for the *Cyclin E Destruction* pathway.

Fig. 4.
The Cyclin E destruction pathway.

TABLE I

Gene sets with significantly associated regions as indicated by the DACE test.

Gene set (= pathway)	Chr	cM	p-value	FDR
E2F1 Destruction	2	93.84 - 101.21	2.39 E-07	1.21 E-04
	5	0 - 5.07	1.49 E-06	3.77 E-04
Cyclin E Destruction	2	91.77 - 101.80	3.62 E-07	1.83 E-04
Role of Tob in T-cell activation	6	84.03 - 88.47	2.95 E-06	8.14 E-04
IL22 Soluble Receptor Signaling	7	39.61 - 41.94	1.47 E-06	4.94 E-04
Regulation of hematopoiesis by cytokines	6	112.66 - 112.67	7.88 E-06	8.84 E-04
	7	80.62- 123.23	1.21 E-07	1.49 E-04
	12	36.64 - 37.82	9.22 E-06	8.84 E-04
	13	34.71 - 34.72	1.97 E-07	1.49 E-04
	13	47.34 - 48.77	5.28 E-06	8.84 E-04
	13	66.95 - 67.24	2.92 E-06	8.84 E-04
	14	28.33 - 28.62	7.84 E-06	8.84 E-04

TABLE II

Significant association identified by permutation procedure.

Gene set (= pathway)	Chr	cM	Perm p	FDR
Cytokines and Inflammatory Response	6	85.48-86.38	0.0	0.0
The Co-Stimulatory Signal During T-cell Activation	3	96.78-97.06	0.0	0.0
Role of Tob in T-cell activation	6	84.61	0.0	0.0
The role of FYVE-finger proteins in vesicle transport	6	91.14 - 96.24	0.0	0.0
	12	57.72	0.0	0.0
Regulation of hematopoiesis by cytokines	7	80.62	0.0	0.0
IL22 Soluble Receptor Signaling Pathway	7	40.51	0.0	0.0
Cyclin E Destruction Pathwa	2	93.84-105.82	0.0	0.0
	5	0.28	0.0	0.0
	18	102.66-102.67	0.0	0.0
E2F1 Destruction Pathway	2	93.85-104.67	0.0	0.0
	5	0.28-2.79	0.0	0.0
Visceral Fat Deposits and the Metabolic Syndrome	5	66.21-66.49	0.0	0.0
	7	50.97-55.97	0.0	0.0
Regulation of p27 Phosphorylation during Cell Cycle Progression	2	86.43-104.65	0.0	0.0
	5	0.29-10.84	0.0	0.0
	9	15.96	0.0	0.0
Angiotensin-converting enzyme 2 regulates heart function	6	64.75-72.38	0.0	0.0
CBL mediated ligand-induced downregulation of EGF receptors	6	91.14-92.72	0.0	0.0
	10	87.31-89.36	0.0	0.0
	15	18.39-23.87	0.0	0.0