aJava

Asian Journal of Animal and Veterinary Advances

Academic
Journals Inc.

www.academicjournals.com

Asian Journal of Animal and Veterinary Advances 7 (8): 745-753, 2012
ISSN 1683-9919 / DOI: 10.3923/ajava.2012.745.753

Simulation Analysis of the Selection of Differential Expression Genes Correlated with Phenotype Data in Chicken Microarray Experiments

[1,2]W.Y. Liu and [3]C.J. Zhao

[1]Department of Scientific Research, Fuyang Normal College, Fuyang City 236041, China
[2]Department of Biology Sciences, Fuyang Normal College, Fuyang City 236041, China
[3]Department of Animal Science, China Agricultural University, Beijing City 100094, China

Corresponding Author: W.Y. Liu, Department of Biological Sciences, Fuyang Normal College No. 100 West Qing He Road, Fuyang City 236037, People's Republic of China Tel: +86-558-2596562, +86-18298199628 Fax: +86-558-2596561, +86-558-2593670

ABSTRACT

DNA microarrays can measure the differential expression of hundreds and thousands of genes to identify changes in their expression between pair-wise biological states. Although, many methods are developed to determine the significance of these changes while accounting for the enormous number of genes, there is no uniform or standard method for different microarray experiments to date. Researches are often confused with diverse algorithm issues, especially the changing fold of differential expression genes, in the selection of suitable candidate genes. Here, a simulation was described to analyze the microarray experimental data correlated with corresponding biological phenotypes, aiming at to measure the proper fold changes of differential expression genes with rigorous statistics of p values and acceptable FDRs (false discovery rates) in permutated univariate tests. The result showed that the commonly used fold change method of candidate gene selection was relatively reasonable and the variation scope was suggested as fold changes of 1.5-2.1. The fold changes no less than 1.5 of gene expression would be a good choice of gene selection with stable gene numbers passed the filtering criteria. When the data filtering fold change was rigid, the number of genes passed the filtering criteria turned to decrease sharply. The present analysis provided a good basis for further gene expression studies.

Key words: DNA microarray, differential expression gene, phenotype data, Spearman correlation, Pearson correlation

INTRODUCTION

Recent years have seen an explosion of work on DNA microarray experiments. For biological problems, such data mainly arise from DNA microarrays (both genome-wide gene expression profile and Single Nucleotide Polymorphism (SNP) mutation assays), from which investigators try to classify disease categories, tumor types, response to drugs, gene expression regulation or other categories (Ludwig and Weinstein, 2005). DNA microarray is a kind of high-throughput biotechnology containing millions of cDNA or oligonucleotide probes for the measuring of gene expression and it can be used to detect the expression hundreds and thousands of genes in a hybridization experiment or single arrays. Despite the success and popularity of oligonucleotide arrays as a high-throughput technique for measuring mRNA expression levels, quantitative calibration studies have until now been limited due to unsuitable data. However, a new

high-throughput technology produced by Affymetrix Corporation contains calibration data and generated their unique oligonucleotide arrays permitting the detailed study of intensity dependent gene expression. Affymetrix oligonucleotide arrays (GeneChip®, Affymetrix Corporation) have been applied to many gene expression studies in the past years (Tavazoie *et al.*, 1999; Cho *et al.*, 2001; Hakak *et al.*, 2001). It adopted a set of perfect match and mismatch oligonucleotide probes, usually 11 to 20 pairs, to measure the mRNA concentrations of genes expressed in an array. Besides linear normalization, average difference and signal methods provided by MAS software (Affymetrix Corporation), researchers have now proposed alternative multi-level analysis methods such as feature extraction (Schadt *et al.*, 2001a, b), normalization (Hill *et al.*, 2001), statistical inferences of gene expression changes (Baldi and Long, 2001; Liu *et al.*, 2005; Fox and Dimmic, 2006), expression index computation (Li and Wong, 2001a, b; Irizarry *et al.*, 2003; Lazaridis *et al.*, 2002; Zhou and Abagyan, 2002) and other aspects (Pavelka *et al.*, 2004; Lu, 2004; Weng *et al.*, 2006) in the attempts to improve these issues.

Although, massive amounts of data are generated and methods are developed to determine whether changes in gene expression are experimentally significant, there is no universal uniform method or justified filtering criterion for different microarray experiment results. Most of the efforts in method development have appropriately focused on what to do with their own datasets. Researches are often confused with diverse algorithms in practice. Many experiments were reported with 2 fold or at least a 1.5 fold changes with an estimated FDR (false discovery rate) (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001) and other experiments were analyzed with lower fold changes since few genes were selected out with 2 fold or 1.5 fold changes sometimes (Tusher *et al.*, 2001; Naef *et al.*, 2002). It seems arbitrary to determine the fold changes of candidate gene selection for further researches to freshmen or laymen. Moreover, most experiments reported only the diverse differential expression genes without correlation with biological phenotypes, especially phenotype data. In the study, a simulation analysis was made in an attempt to resolve the issue for candidate genes selection based on fold changes using their correlation with phenotype data in DNA microarray experiments.

MATERIALS AND METHODS

Data sets: The datasets were retrieved from GEO http://www.ncbi.nlm.nih.gov/geo/or Array Express http://www.ebi.ac.uk/arrayexpress/ (GSE12675) for the chicken embryo gene expression profiling in heart tissues at Hamburger-Hamilton stage 43. The data (GSE12675) were produced with Affymetrix GeneChip® Chicken Genome Arrays which contains 32773 transcripts corresponding to over 28000 chicken genes and 684 transcripts from 17 avian viruses. It included 60 animals and 12 chicken genome arrays in total. The definite animal assembly, as well as the procedures of sample preparation, RNA preparation, microarray hybridization and quantitative PCR experiment, were described in the previous study (Li and Zhao, 2009). In the simulation analysis, the phenotype data were the tender weights of organism muscle tissues (breast muscle and thigh muscle) produced by the corresponding 60 full or half sibs (i.e., mixed family) at the age of eight weeks in the same assembly as described (Li and Zhao, 2009) with proper statistical amendment based on the farm's statistical data. It was aimed at probing into the effect of differential expression genes on muscle traits and assumed the chicken full or half sibs' phenotype data were genetically the same as those animals used in previous studies (Li and Zhao, 2009; Song *et al.*, 2010).

Data preprocessing: The raw datasets (*.CEL files) were re-analyzed by Affymetrix GCOS software (Affymetrix Corporation) for quality issues. Data preprocessing was performed using software packages developed in version 2.5.0 of Bioconductor (Gentleman *et al*., 2004) and R version 2.10.0 (R Development Core Team, 2009). Each Affymetrix dataset was background adjusted, normalized and log2 probe-set intensities calculated using the Robust Multichip Averaging (RMA) algorithm in affy package (Irizarry *et al*., 2003; Gautier *et al*., 2004). When the multiple probe sets target one gene, the probe set with largest variability was kept.

The correlation of differential expression genes with phenotype data was accomplished with a program written in the Fortran language and Bioconductor 2.5.0 (Gentleman *et al*., 2004) for permutated univariate tests of a random variance model (Wright and Simon, 2003). Two correlation algorithms, i.e. Pearson correlation and Spearman correlation, were used in the analysis with three levels of significance of p value of the permutated univariate tests (0.05, 0.01 and 0.001) and corresponding FDRs. Pearson correlation is computed between the measurements from two datasets. The Pearson correlation is commonly used in statistical models, but Spearman correlation is a more robust measure of correlation. The Spearman correlation is defined as a rank correlation like the Pearson correlation except that within each profile, each gene expression measurement is replaced by its rank within that profile and it is those ranks that are used in computing the correlation. The Spearman correlation is less sensitive to outlying observations and therefore tends to be more stable, but it is not clear how to handle assignment of rank if there are some missing expression measurements.

RESULTS AND DISCUSSION

In recent years, gene expression profile studies, including real-time quantitative PCR (quantitative RT-PCR or RT-PCR) and DNA microarray experiments, are increasingly common in domesticated animal researches (Salleh *et al*., 2004; Endrini *et al*., 2007; Chaudhry, 2008; Baatartsogt *et al*., 2009; Cossio-Bayugar *et al*., 2009; Chen *et al*., 2011; Huang *et al*., 2011; Long *et al*., 2011; Baatartsogt *et al*., 2011; Suzuki *et al*., 2011), especially in birds (Mekki *et al*., 2006; Lim *et al*., 2009; Goyal *et al*., 2010). However, few studies reported the comparative selection of suitable differential expression genes or candidate genes for further studies with dynamic fold changes of gene expression. Correlation of differential expression genes with biological phenotype information was also seldom reported and many of the experiments were designed concerning the gene interaction or interactome predicted from the genome-wide correlation (Ge *et al*., 2001). To the best of our knowledge, there were few researches or reports on the justification of candidate gene selection determined by fold changes in diverse microarray experiments. It is for the first time that the selection of differential expression genes based on fold changes had been experimentally investigated in chicken microarray experiments with reasonable and/or acceptable parameters correlated with specific phenotype data in the simulations.

The simulation analysis was time consuming since all the datasets were Affymetrix GeneChip® Chicken Genome Arrays which contains more than 28000 chicken genes' transcript information. In total, twenty-four simulations were done in over 150 permutated univariate tests. The result was processed through permutated univariate tests of Pearson correlation and Spearman correlation with statistical significance (Table 1). Table 1 showed the summary of the prime result in detail. It was noticed that all the positive result was mainly from the Spearman correlation tests due to Spearman correlation is more robust and stable than Pearson correlation. Simulations using the Pearson correlation tests were almost unavailable and made no use later. As showed in Table 1,

Table 1: No. of significant genes detected with different filtering criteria in simulations (ouly Spearman correlations were stably available)

Fold changes of data filtering criteria	No. of genes passed the filtering criteria	Spearman correlation			Pearson correlation
		p<0.05	p<0.01	p<0.001	p<0.001
2.5	163	06	2	0	0
2.4	207	06	2	0	0
2.3	265	06	2	0	0
2.2	339	06	2	0	0
2.1	461	09	2	0	0
2.0	633	15	2	0	0
1.9	903	26	2	0	0
1.8	1296	35	3	0	0
1.7	1895	72	10	0	0
1.6	2909	115	14	0	0
1.5	4505	176	21	3	0
1.4	7045	268	28	3	0
1.3	11044	439	48	10	0
1.2	17039	681	70	15	0
1.1	25484	1041	102	28	0
1.09	26353	1073	103	30	0
1.08	27143	1104	109	31	0
1.07	27912	1136	114	34	0
1.06	28585	1164	116	34	0
1.05	29182	1184	119	35	0
1.04	29589	1206	121	37	0
1.03	30189	1224	122	37	0
1.02	30082	1223	122	37	0
1.01	30189	1224	122	37	0

Pearson correlation tests with p value less than 0.001 found no differential expression genes. The fold changes ranged from 1.01 to 2.5 with detailed analyses from 1.01 to 1.09 due to tremendous variations in gene numbers (Table 1).

Although, p value of 0.05 or 0.01 is significant in the context of common experiments designed to evaluate small numbers of specimens or genes, microarray experiments designed for over 10,000 genes would identify 100 genes falsely by chance. This was why we gave a crucial consideration and took integrated correlation tests with p<0.001 (Table 1, Fig. 1). Spearman correlation tests with three significance levels revealed that number of genes passed the filtering criteria and correlation tests decreased rapidly before the fold change of 1.5. However, the number of genes passed the filtering criteria had changed more than 1.5 fold and turned into a relatively stable trend. Especially, the Spearman correlation tests with p<0.001 clearly produced no differential expression genes more than 1.5 fold change. These suggested that there would be no chance of mistake in the selection of differential expression genes correlated with the phenotype data. The fold change of 1.5 should be regarded as the lower limit of filtering criteria of candidate genes selection. Moreover, the selection of differential expression genes with more rigorous filtering criteria will result in less candidate genes available and the fold change of 2.1 were suggested as the suitable upper limit in candidate genes' filtering (Table 1, Fig. 1). The variation scope should be 1.5-2.1 fold changes while the 1.8-2.0 fold changes of gene expression could be a better choice of gene selection. When the data filtering fold changes were rigid, the number of genes passed the

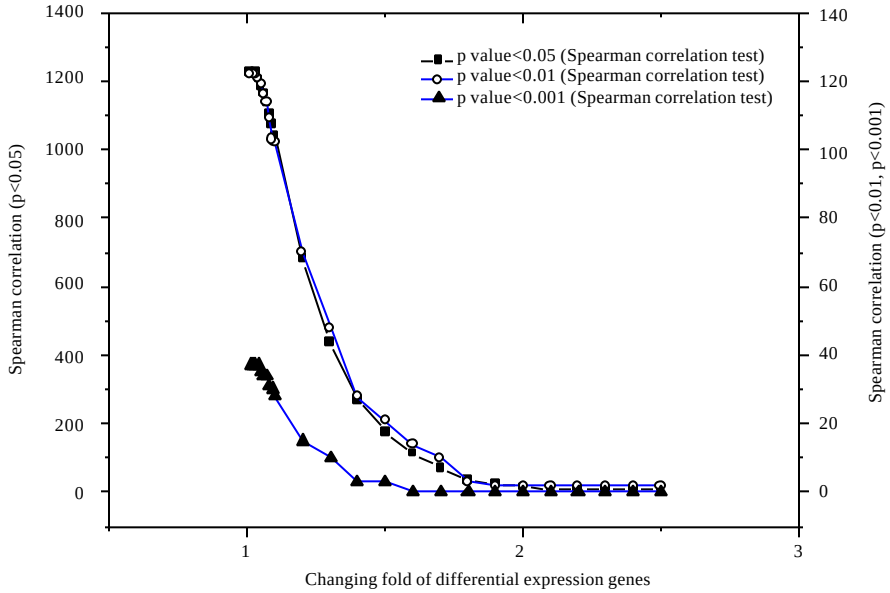

Fig. 1: Plot of the No. of significant genes detected with different filtering criteria in simulations (only Spearman correlations were stably available and showed), The lateral axis displayed the fold changes (i.e. filtering criteria) of differential expression genes and the vertical axes indicated the differential expression gene numbers passed the filtering criteria with Spearman correlations at the significance levels of p-values. The left vertical axis gave the gene numbers passed at the significance level of 0.05 of p-values while the right vertical axis provided the gene numbers passed at the significance levels of 0.01 and 0.001 of p- values

filtering criteria turned to decrease sharply. Figure 1 showed the plot of Spearman correlation test results.

The analyses above gave fine results and data with significant p values and acceptable FDRs in permutated univariate tests. Table 2 showed the first 37 genes sorted by the nominal 0.001 level of p value with Spearman correlation in a permutated univariate test. The significance of the p values and FDRs values were relatively acceptable (Benjamini and Yekutieli, 2001; Tusher *et al*., 2001; Wu, 2008) in permutated tests or multiple tests. Benjamini and Hochberg's method (Benjamini and Hochberg, 1995; Benjamini and Yekutieli, 2001) assumes independent tests and guarantees an upper bound for the FDR by a step-up or step-down procedure applied to the individual p values. Because of the limited number of permutations, the FDR is too granular. For data used in the simulations, the p value for each gene is calculated from permutations of only twelve microarray experiments and the FDR values (0.3-0.7) were relatively acceptable depending on how the p value was defined. Table 2 also indexed the simulation information and parametric p values and FDR values. Another important factor that might impact the FDR values was the selection of phenotype data since the simulation analysis was based on the correlation of differential expression genes with the full or half sibs' organism muscle tender weights. The phenotype data were produced from full or half sibs genetically the same as those chicken in the previous study and the farm's housing environment and management measures was constant and stable in the farm However, the full or half sibs were not the chicken used in the microarray experiments and the time was not same as that of gene expression. Therefore, the experiments' result could only be regarded

Table 2: The first 37 genes sorted by the nominal 0.001 level of p-value in the permutated univariate test with the 1.01 fold change of gene expression

Spearman correlation coefficient	p-value	FDR	Probe set of GeneChip®
0.996	0.0000031	0.4708371	GgaAffx.7110.1.S1_at
0.993	0.0000081	0.4708371	Gga.10656.1.S1_at
0.992	0.0001053	0.4708371	GgaAffx.24437.2.S1_s_at
0.99	0.0001603	0.4708371	Gga.5875.1.S1_a_at
0.988	0.0002168	0.4708371	Gga.10670.1.S1_s_at
-0.986	0.0003027	0.4708371	GgaAffx.13150.1.S1_s_at
0.985	0.0003164	0.4708371	Gga.14257.1.S1_at
-0.985	0.0003185	0.4708371	Gga.17981.1.S1_at
-0.985	0.0003353	0.4708371	GgaAffx.10499.1.S1_at
0.985	0.0003375	0.4708371	Gga.1168.1.S1_at
0.985	0.0003392	0.4708371	Gga.9077.1.S1_at
0.985	0.0003555	0.4708371	Gga.5520.1.S1_at
-0.984	0.0003684	0.4708371	GgaAffx.444.1.S1_at
-0.983	0.0004146	0.4708371	Gga.4562.1.S1_s_at
-0.982	0.0004655	0.4708371	Gga.13065.1.S1_at
0.981	0.0005228	0.4708371	Gga.19209.1.S1_at
-0.981	0.0005476	0.4708371	GgaAffx.20308.1.S1_at
-0.979	0.0006303	0.5934176	GgaAffx.1650.1.S1_s_at
0.979	0.0006346	0.5934176	Gga.10998.1.S1_s_at
0.979	0.0006452	0.5934176	Gga.4715.2.S1_a_at
-0.979	0.0006473	0.5934176	Gga.12684.1.S1_s_at
-0.979	0.0006554	0.5934176	Gga.18834.1.S1_at
0.978	0.0006898	0.5934176	Gga.4194.1.S1_at
-0.978	0.0007029	0.5934176	GgaAffx.26312.1.S1_s_at
0.978	0.0007342	0.5934176	Gga.16560.2.S1_s_at
0.977	0.0007582	0.5934176	Gga.10515.3.S1_a_at
-0.977	0.0007871	0.5934176	Gga.4864.1.S1_at
-0.977	0.0007909	0.5934176	Gga.7691.1.S1_at
0.976	0.0008326	0.5934176	Gga.1909.1.S1_at
0.976	0.0008526	0.5934176	Gga.10905.1.S1_at
0.976	0.0008850	0.5934176	GgaAffx.10124.2.S1_s_at
0.976	0.0008884	0.5934176	Gga.5786.1.S1_at
-0.975	0.0009538	0.5934176	GgaAffx.23629.1.S1_s_at
0.975	0.0009543	0.5934176	GgaAffx.5641.1.S1_s_at
0.975	0.0009604	0.5934176	Gga.6534.1.A1_at
0.974	0.0009663	0.5934176	Gga.8241.1.S1_at
-0.974	0.0009939	0.5934176	Gga.14825.1.S1_at

Positive coefficients showed the genes were significantly up-regulated and negative coefficients indicated the genes were significantly download-regulated

as "simulations" and it remains to be confirmed by further studies in a more realistic experiment. The resulted data of the simulation analysis are available through e-mail upon demand.

CONCLUSION

So far, many studies appear as taking advantage of novel advanced biotechnologies such as DNA microarray to select suitable candidate genes and dissect the complex of gene regulation network. In this study, for the first time, simulation analysis was made to find out differential

expression genes correlated with corresponding biological phenotypes with rigorous statistics of p values ($p<0.05$, 0.01, 0.001) and acceptable FDR values with stable correlation analyses in the permutated univariate tests. Generally speaking, the result showed that the traditional fold change method of candidate gene selection was relatively reasonable. The variation scope should be 1.5-2.1 fold changes and the 1.8-2.0 fold changes of gene expression should be the better choice of gene selection with stable gene numbers passed the filtering criteria. The simulation gave the first step to resolve the issue for candidate genes selection by means of observing fold changes and correlations with corresponding phenotype data. Potentially, this might offer a good reference and further connection to study gene regulatory network kinetics, as well as bridging the gap between the microarray experiments and individual biological phenotypes.

ACKNOWLEDGMENTS

We are grateful to the anonymous reviewers for their constructive comments and suggestions. This work was supported by Chinese Research Funds to LWY (2011PTFY04ZD, KJ2012A216).

REFERENCES

Baatartsogt, O., K. Choi, P.K. Mandal, H.K. Lim and G.H. Li *et al.*, 2009. Gene expression profile of synovial cells in experimental post-traumatic arthritis of knee in swine. Biotechnology, 8: 70-77.

Baatartsogt, O., P.K. Mandal, H.K. Lim, C.H. Lee, J.H. Lee and K. Choi, 2011. Differential expression of immuno-inflammatory genes in synovial cells from knee after inducing post-traumatic arthritis in swine. Asian J. Anim. Vet. Adv., 6: 333-343.

Baldi, P. and A.D. Long, 2001. A Bayesian framework for the analysis of microarray expression data: Regularized t-test and statistical inferences of gene changes. Bioinformatics, 17: 509-519.

Benjamini, Y. and D. Yekutieli, 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Stat., 29: 1165-1188.

Benjamini, Y. and Y. Hochberg, 1995. Controlling the false discovery rate: A practical approach to multiple testing. J. R. Stat. Soc. Ser. B. Stat. Methodol., 57: 289-300.

Chaudhry, M.A., 2008. Induction of gene expression alterations by culture medium from trypsinized cells. J. Boil. Sci., 8: 81-87.

Chen, Y., S.L. Yang, W.D. Deng and H.M. Mao, 2011. Differential gene expression in different tissues of black-bone sheep and normal sheep. Asian J. Anim. Vet. Adv., 5: 379-384.

Cho, R.J., M. Huang, M.J. Campbell, H. Dong and L. Steinmetz *et al.*, 2001. Transcriptional regulation and function during the human cell cycle. Nat. Genet., 27: 48-54.

Cossio-Bayugar, R., E. Miranda-Miranda, D. Portilla-Salgado and J. Osorio-Miranda, 2009. Quantitative PCR detection of cholinesterase and carboxylesterase expression levels in acaricide resistant *Rhipicephalus (Boophilus) microplus*. J. Entomol., 6: 117-123.

Endrini, S., A. Rahmat, P. Ismail and Y.H. Taufiq-Yap, 2007. Comparing of the cytotoxicity properties and mechanism of *Lawsonia inermis* and *Strobilanthes crispus* extract against several cancer cell lines. J. Med. Sci., 7: 1098-1102.

Fox, R.J. and M.W. Dimmic, 2006. A two-sample Bayesian t-test for microarray data. BMC Bioinformatics, Vol. 7. 10.1186/1471-2105-7-126.

Gautier, L., L. Cope, B.M. Bolstad and R.A. Irizarry, 2004. Affy-analysis of *Affymetrix GeneChip* data at the probe level. Bioinformatics, 20: 307-315.

Ge, H., Z. Liu, G.M. Church and M. Vidal, 2001. Correlation between transcriptome and interactome mapping data from *Saccharomyces cerevisiae*. Nat. Genet., 29: 482-486.

Gentleman, R.C., V.J. Carey, D.M. Bates, B. Bolstad and M. Dettling *et al.*, 2004. Bioconductor: Open software development for computational biology and bioinformatics. Genome Biol., Vol. 5, No. 10. 10.1186/gb-2004-5-10-r80.

Goyal, G., V. Upmanyu, S.K. Singh, S.K. Shukla, S. Mehra, V. Kumar and D. Sharma, 2010. Differential expression of IL-6 and IGF-II in guinea fowl and chicken. Int. J. Poult. Sci., 9: 756-760.

Hakak, Y., J.R. Walker, C. Li, W.H. Wong and K.L. Davis *et al.*, 2001. Genome-wide expression analysis reveals dysregulation of myelination-related genes in chronic schizophrenia. Proc. Natl. Acad. Sci. USA., 98: 4746-4751.

Hill, A.A., E.L. Brown, M.Z. Whitley, G. Tucker-Kellogg, G.P. Hunter and D.K. Slonim, 2001. Evaluation of normalization procedures for oligonucleotide array data based on spiked cRNA controls. Genome Biol., Vol. 2, No. 12. 10.1186/gb-2001-2-12-research0055.

Huang, J., L. Liu, H. Wang, C. Zhang and Z. Ju *et al.*, 2011. Variants and gene expression of the *TLR2* gene and susceptibility to mastitis in cattle. Asian J. Anim. Vet. Adv., 6: 51-61.

Irizarry, R.A., B. Hobbs, F. Collin, Y.D. Beazer-Barclay, K.J. Antonellis, U. Scherf and T.P. Speed, 2003. Exploration, normalization and summaries of high density oligonucleotide array probe level data. Biostatistics, 4: 249-264.

Lazaridis, E.N., D. Sinibaldi, G. Bloom, S. Mane and R. Jove, 2002. A simple method to improve probe set estimates from oligonucleotide arrays. Math. Biosci., 176: 53-58.

Li, C. and W.H. Wong, 2001a. Model-based analysis of oligonucleotide arrays: Expression index computation and outlier detection. Proc. Natl. Acad. Sci. USA., 98: 31-36.

Li, C. and W.H. Wong, 2001b. Model-based analysis of oligonucleotide arrays: Model validation, design issues and standard error application. Genome. Biol., Vol. 2. No. 8.

Li, M. and C. Zhao, 2009. Study on Tibetan chicken embryonic adaptability to chronic hypoxia by revealing differential gene expression in heart tissue. Sci. China. C. Life. Sci., 52: 284-295.

Lim, H.K., K. Choi, P.K. Mandal, O. Baatartsogt, C.H. Lee, J.H. Lee and H.B. Kim, 2009. Transcriptional profiling of spleen lymphocyte in fowl typhoid of broilers. Asian J. Anim. Vet. Adv., 4: 66-75.

Liu, X., A. Krishnan and A. Mondry, 2005. An Entropy-based gene selection method for cancer classification using microarray data. BMC Bioinformatics, Vol. 6. 10.1186/1471-2105-6-76.

Long, X.Y., Y.X. Liu, H. Rocheleau, T. Ouellet and G.Y. Chen, 2011. Identification and validation of internal control genes for gene expression in wheat leaves infected by strip rust. Int. J. Plant Breed. Genet., 5: 255-267.

Lu, C., 2004. Improving the scaling normalization for high-density oligonucleotide GeneChip expression microarrays. BMC Bioinformatics, Vol. 5. 10.1186/1471-2105-5-103.

Ludwig, J.A. and J.N. Weinstein, 2005. Biomarkers in cancer staging, prognosis and treatment selection. Nat. Rev. Cancer, 5: 845-856.

Mekki, D.M., B. Liang-yong, W. Jin-yu, Y. Yan, Y. Yabo and H.H. Musa, 2006. Avian E2A expression level variation associated with phenotypic variation of two Chinese indigenous duck breeds (*Anas platyrhynchos*). Int. J. Poult. Sci., 5: 1023-1028.

Naef, F., C.R. Hacker, N. Patil and M. Magnasco, 2002. Empirical characterization of the expression ratio noise structure in high-density oligonucleotide arrays. Genome Biol., Vol. 3. 10.1186/gb-2002-3-4-research0018.

Pavelka, N., M. Pelizzola, C. Vizzardelli, M. Capozzoli, A. Splendiani, F. Granucci and P. Ricciardi-Castagnoli, 2004. A power law global error model for the identification of differentially expressed genes in microarray data. BMC Bioinformatics, Vol. 5. 10.1186/1471-2105-5-203.

R Development Core Team, 2009. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Salleh, M.N., P. Ismail, Y.H. Taufiq-Yap and P. Carmichael, 2004. Identification of *Estrogen-responsive* genes in $p53^{+/-}$ knockout and isogenic wild-type parent strain mice by CDNA macroarray analysis. J. Med. Sci., 4: 312-317.

Schadt, E.E., C. Li, B. Ellis and W.H. Wong, 2001a. Feature extraction and normalization algorithms for high-density oligonucleotide gene expression array data. J. Cell. Biochem., 84: 120-125.

Schadt, E.E., C. Li, C. Su and W.H. Wong, 2001b. Analyzing high-density oligonucleotide gene expression array data. J. Cell. Biochem., 80: 192-202.

Song, M., W. Han, H. Bao, C. Liu, C. Wu and C. Zhao, 2010. Physiological adaptability of tibet chicken embryo liver to hypoxia and its protein profile analysis. Asian J. Anim. Vet. Adv., 5: 547-556.

Suzuki, H., R. Ohtsuka and M. Takeda, 2011. Differences in gene expression profiles among the proximal, middle and distal peyer's patches in the mouse small intestine. Res. J. Immunol., 4: 19-24.

Tavazoie, S., J.D. Hughes, M.J. Campbell, R.J. Cho and G.M. Church, 1999. Systematic determination of genetic network architecture. Nat. Genet., 22: 281-285.

Tusher, V.G., R. Tibshirani and G. Chu, 2001. Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. USA., 98: 5116-5121.

Weng, L., H.Y. Dai, Y.H. Zhan, Y.D. He, S.B. Stepaniants and D.E. Bassett, 2006. Rosetta error model for gene expression analysis. Bioinformatics, 22: 1111-1121.

Wright, G.W. and R.M. Simon, 2003. A random variance model for detection of differential gene expression in small microarray experiments. Bioinformatics, 19: 2448-2455.

Wu, W.B., 2008. On false discovery control under dependence. Ann. Stat., 36: 364-380.

Zhou, Y. and R. Abagyan, 2002. Match-Only Integral Distribution (MOID) algorithm for high-density oligonucleotide array analysis. BMC Bioinformatics, Vol. 3. 10.1186/1471-2105-3-3.