Integrative analysis of multiple gene expression profiles applied to liver cancer study

Jung Kyoon Choi[a,b], Jong Young Choi[c], Dae Ghon Kim[d], Dong Wook Choi[e], Bu Yeo Kim[e], Kee Ho Lee[e], Young Il Yeom[f], Hyang Sook Yoo[f], Ook Joon Yoo[b], Sangsoo Kim[a,*]

[a] *National Genome Information Center, Korea Research Institute of Bioscience and Biotechnology, Yuseong-gu, Eoeun-dong 52, Daejeon 305-333, Republic of Korea*
[b] *Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Daejeon, Republic of Korea*
[c] *Kangnam St. Mary's Hospital, The Catholic University of Korea, Seoul, Republic of Korea*
[d] *Division of Gastroenterology and Hepatology, Chonbuk National University Medical School and Hospital, Chonju, Chonbuk, Republic of Korea*
[e] *Laboratory of Molecular Oncology, Korea Cancer Center Hospital, Seoul, Republic of Korea*
[f] *Korea Research Institute of Bioscience and Biotechnology, Daejeon, Republic of Korea*

Received 30 January 2004; revised 5 March 2004; accepted 8 March 2004

First published online 9 April 2004

Edited by Lukas Huber

Abstract A statistical method for combining multiple microarray studies has been previously developed by the authors. Here, we present the application of the method to our hepatocellular carcinoma (HCC) data and report new findings on gene expression changes accompanying HCC. From the cross-verification result of our studies and that of published studies, we found that single microarray analysis might lead to false findings. To avoid those pitfalls of single-set analyses, we employed our effect size method to integrate multiple datasets. Of 9982 genes analyzed, 477 significant genes were identified with a false discovery rate of 10%. Gene ontology (GO) terms associated with these genes were explored to validate our method in the biological context with respect to HCC. Furthermore, it was demonstrated that the data integration process increases the sensitivity of analysis and allows small but consistent expression changes to be detected. These integration-driven discoveries contained meaningful and interesting genes not reported in previous expression profiling studies, such as growth hormone receptor, erythropoietin receptor, tissue factor pathway inhibitor-2, etc. Our findings support the use of meta-analysis for a variety of microarray data beyond the scope of this specific application.
© 2004 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.

Keywords: Microarray; Meta-analysis; Hepatocellular carcinoma; Liver cancer

1. Introduction

Recently, many microarray results have been derived from cancer studies with an increasing interest in the classification of tumors and prediction of clinical outcome. We present here a meta-analysis result of expression profiles of hepatocellular carcinoma (HCC), which is among the five leading causes of cancer death in the world. The causes of HCC are better understood than those of other cancers in human. Major risk factors for HCC are chronic hepatitis resulting from hepatitis B virus (HBV) or hepatitis C virus. Nevertheless, the molecular pathogenesis of HCC, including changes in gene expression and aberrations in gene structure as well, is not well understood, leaving prognosis of HCC very poor.

We collected four independent cDNA microarray datasets generated with a common objective to identify differentially expressed genes in HBV-positive HCC samples. Despite the fact that they had been created in relatively controlled experimental conditions, we failed to directly merge the data even after normalization of each dataset. A clustering result showed that the tissue samples tended to cluster accordingly more to their experimental origin (i.e., which dataset they were derived from) rather than to their biological origin (i.e., whether they are malignant or normal). As was pointed out by Brazma et al. [1], a wide variety of experimental variables might have caused this problem. These include microarray platforms and formats, reference samples, experimental procedures, laboratory conditions, parameters on image processing, and so on.

One way to draw a conclusion from these heterogeneous datasets is to combine their analysis results using meta-analysis methods [2–5]. There are three major approaches to meta-analysis: vote counting, combining significance levels, and combining effect sizes. Vote counting represents the result of analysis as either a positive or negative vote to a hypothesis. Suppose we have eight microarray datasets with a common study design. The change in the expression level of a gene caused by a given treatment might be significant in five datasets and not significant in three datasets, yielding five positive and three negative votes. While intuitive, this is a rather crude approach, and does not take into account the level of significance of an individual result. By combining significance levels, this limitation can be overcome. For the example in the above, the P values can be combined across the eight datasets to estimate the overall P value. Recently, this method was successfully applied to the meta-analysis of four prostate cancer microarray datasets making new findings [6,7]. The most recent development in meta-analysis introduced the novel concept of combining effect sizes. It is a procedure that permits

* Corresponding author. Fax: +82-42-879-8519.
E-mail address: sskimb@kribb.re.kr (S. Kim).

Abbreviations: HCC, hepatocellular carcinoma; HBV, hepatitis B virus; GO, gene ontology; FEM, fixed effects model; REM, random effects model

0014-5793/$22.00 © 2004 Published by Elsevier B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/j.febslet.2004.03.081

the combination of effect size estimates to obtain an overall estimate of the average effect size. Effect size is a name given to a family of indices that measure the magnitude of treatment effect given in a study. In [8], a methodology for using effect size models for microarray data was established and shown to be able to manage interstudy variation effectively.

In the analysis of the multi-center HCC datasets, three main questions were asked:

1. Do the independent studies produce similar results? Does a typical single-set analysis with small sample size produce reproducible results? To answer this question, we performed a cross-verification by comparing individual significant gene lists. We also compared five published HCC studies that had been performed with laboratory scale sample size [9–13].
2. Does the meta-analysis of microarray analyses work from theory to practice? The question could be addressed by investigating gene ontology (GO) terms associated with the dysregulated genes identified by meta-analysis. This would also provide an overview of their biological implications with regard to HCC.
3. What are the advantages of meta-analysis versus single analyses? One of our expectations was that increased sample size would increase statistical power. By comparing the meta-analysis result with the individual results, we could specify the advantage of meta-analysis from the biological as well as from the statistical perspective. Furthermore, here we show that meta-analysis is essential for making reliable results from microarrays.

2. Materials and methods

2.1. Microarrays and sequence verification

Human cDNA clone set (UniGEM^{TM}) was obtained from Incyte Genomics (Fremont, CA, USA). The identities of more than 10 000 individual clones were verified by single-pass sequencing. A total of 7844 sequences were generated after passing PHRED base calling and quality control processes. About 5461 known genes and 994 ESTs, including replicated ones, were confirmed in the set by annotating with UniGene (http://www.ncbi.nlm.nih.gov/UniGene).

2.2. Microarray experiments and datasets

HCC and adjacent normal samples were obtained with informed consent from patients at three different hospitals. All the HCC samples were HBV-positive. Each hospital's Clinical Research Review Board has approved its own study. The human cDNA clone set was spotted using a commercial arrayer. Two different versions differing in terms of spot-locations were distributed to the hospitals for four independent hybridizations of HCC and normal liver samples. Sample preparation, microarray hybridizations, and fluorescence signal acquisitions were carried out independently at each institution according to similar but not identical experimental protocols and laboratory conditions. One experiment used a different reference sample from the three other experiments (see Table 1 for detailed information).

The raw data generated from these experiments were collected and subjected to a normalization procedure after being passed a null value and a variation filter. A lowess smoothing of the scatter plot between $\log_2 R/G$ and $\log_2 \sqrt{R \times G}$ was proposed recently [14]. We applied the so-called 'within-print-tip-group' normalization method using the R package available at http://www.stat.berkeley.edu/users/terry/zarray/Html.

2.3. Effect size method

The statistical method has been detailed in [8] and is described only briefly here. Standardizing the mean difference between the two groups gives $d = (\bar{X}_t - \bar{X}_n)/S_p$, where $\bar{X}_t$ and $\bar{X}_n$ represent the means of the tumor and normal group, respectively, and S_p is the pooled standard deviation. For each gene, we obtained the effect size d_i for studies $i = 1, 2, \ldots, k$.

With a well-established model for combining effect sizes, we easily estimate the average effect size μ. The general model was given by $d_i = \mu + \delta_i + \varepsilon_i$, where $\delta_i \sim N(0, \tau^2)$ and $\varepsilon_i \sim N(0, s_i^2)$. Between-study variance τ^2 represents the variability between studies, while within-study variance s_i^2 represents the sampling error conditioned on the ith study. We let $\tau^2 = 0$ if we can safely assume that effect sizes d_is are not originally different from one another. This can be addressed by testing for the homogeneity of study effects. We used a widely used test for homogeneity based on the statistic $Q = \sum w_i(d_i - \bar{d})^2$, where $w_i = s_i^{-2}$ and $\bar{d} = (\sum w_i d_i)/\sum w_i$ [15]. Model selection led to an overall estimate of μ given by weighted average of the observed effect sizes as $\hat{\mu} = (\sum w_i d_i)/\sum w_i$, where $w_i = (s_i^2 + \tau^2)^{-1}$, if s_i^2 and τ^2 are known. The estimation of s_i^2 and τ^2 was given in [8] with Bayesian approach.

Standard normal score for the average effect size (z_{avg}) was computed by estimating the standard error of the average effect size. Standard normal score for each effect size (z_i) was computed as the ratio of d_i to s_i.

3. Results

3.1. Data collection and preliminary examinations

Four independent datasets (D1, D2, D3 and D4) were created from multi-center microarray studies on HCC (Table 1).

Table 1
Dataset information

Dataset ID	Number of tumor samples	Number of normal samples	Chip type	Sample labeling (Cy5:Cy3)	Statistic	Author[a] or experimenter
P1	9	9	Membrane	Radioactivity	Paired-sample fold change	Xu et al. [11]
P2	6	8	Affymetrix	SAPE	Two-sample fold change	Tackels-Horne et al. [10]
P3	20	20	cDNA	Normal:tumor	Paired-sample fold change	Okabe et al. [9]
P4	12	12	cDNA	Tumor:normal	Paired-sample fold change	Li et al. [12]
P5	8	8	cDNA	Tumor:normal	Paired-sample fold change	Chung et al. [13]
D1	16	16	cDNA (Ver. 1)	Sample:normal liver	Two-sample effect size	Local hospital A
D2	23	23	cDNA (Ver. 1)	Sample:placenta[b]	Two-sample effect size	Local hospital B
D3	29	5	cDNA (Ver. 1)	Sample:placenta	Two-sample effect size	Local hospital C
D4	12	9	cDNA (Ver. 2[c])	Placenta:sample	Two-sample effect size	Local hospital C

[a] Refs. [9–13].
[b] The human placenta RNA samples were from the same batch.
[c] Ver. 2 chips have different spot locations from Ver. 1 chips. They were printed using the same arrayer.

Differential expressions were measured as effect sizes for each dataset. The effect size of a gene that has missing values in greater than 50% of the observations in a particular dataset was considered as missing for that dataset. Two gene sets, G1 and G2, were defined to contain genes or ESTs which generated all four effect sizes (6610 elements) and more than two effect sizes (9982 elements), respectively. G1, the subset of G2, was used as default throughout this study.

On the other hand, we attempted to find commonly reported genes among five published studies, P1, P2, P3, P4 and P5 (Table 1). However, to our astonishment, no up-regulated gene and only two down-regulated genes (cytochrome P450IIC, polypeptide 9 and hepatocyte growth factor (HGF) activator) were found. To quantify consistency between studies, cross-verifications were performed as follows: first, significant genes from each dataset were listed; second, in each list, we counted the number of genes which appeared also in at least one of the other lists; third, a verification score was computed as the ratio of the number of verified genes to the total number of the listed genes. The procedure was applied between P1–P5 and between D1–D4, separately. For D1–D4, a z score was assigned for each gene and genes with $z > 2.5$ were identified as significant genes (the same threshold as used for our meta-analysis). As shown in Table 2, the verification scores were about 9–20% for P1–P5 and 14–22% for D1–D4, which means that about 80% of the genes called significant in a single study were not confirmed by at least one of the other studies. This inconsistency is most likely to be caused from different experimental conditions and limited scopes of study populations. This finding explicitly presents the artifacts of single-set analyses and emphasizes the importance of integrating multiple studies.

3.2. Model selection and the identification of significant genes

To explore variability between the studies, the heterogeneity of the effect sizes was measured by the statistic Q, which is distributed as a χ^2 distribution under the hypothesis of homogeneity [15]. A quantile plot for the genes of G1 was shown in Fig. 1(a) in [8]. The observed Q values perfectly fitted the χ^2_3 distribution, implying the homogeneity of the effect sizes. In other words, the effect sizes can be assumed to be sampled from the common normal population, and the differences are likely to be due to sampling error alone. Therefore, we could consider the between-study variance τ^2 to be zero.

To estimate the statistical significances of the average effect sizes without a normal approximation, we applied the recently described analytical method called SAM [16]. The procedure also includes adjusting for multiple testing. G2 was used in this procedure so that the genes with some missing results were included. For each dataset, column-wise permutation was performed to produce a null effect size for each gene. Permutations between datasets were not allowed to occur. These null effect sizes from the four datasets were combined under the assumption that τ^2 is zero. In this way, the z score of the null average effect size, z^*_{avg}, was obtained for each gene. The genes with $|z_{\text{avg}}| > 2.5$ from the original data were chosen. This criterion identified 150 genes significantly up-regulated and 327 genes significantly down-regulated. For each permutation, the number of the genes with $|z^*_{\text{avg}}| > 2.5$ was counted. 300 different permuted datasets generated an average of 48.5 such genes. The number implies an estimated number of falsely significant discoveries. The ratio over the number of significant genes in the original data, i.e., 477, yields an estimated false discovery rate of 10%. The gene names and expression pattern of 393 known genes among 477 significant genes can be visualized at http://centi.kribb.re.kr/MMA. Complete datasets also can be downloaded from the site.

3.3. Comparison of meta-analysis with individual analyses

Of the 477 significant genes, 365 genes were found also in G1. Using these 365 genes, we compared the meta-analysis result with the individual analyses. Significant genes in each dataset were selected using the same z threshold ($z_{\text{th}} = 2.5$) as that used in the meta-analysis. These gene lists are the ones that were used in our cross-verification described above. For each of the gene lists, we counted the number of genes appearing in the list of the 365 genes. For the meta-analysis result, we counted the number of genes appearing in at least one of the four lists. Verification scores were generally higher than those in cross-verification (Table 2), which implies that the meta-analysis produced results in reasonable agreement with each individual analysis.

Notably, in contrast to the 215 genes present in at least one of the individual lists, the other 150 genes were not found in any of the lists. We denoted these genes as integration-driven discoveries. As a general statement, among total discoveries satisfying $|z_{\text{avg}}| \geqslant z_{\text{th}}$, those with $|z_i| < z_{\text{th}}$ for all studies $i = 1, \ldots, k$ are called integration-driven discoveries with regard to the z_{th}.

They can be considered as being detected by the integration process. The plot in Fig. 1 shows the relationship between the number of integration-driven discoveries and the number of total discoveries for up- or down-regulated genes. Of the most significant 500 genes in the meta-analysis, approximately up to 100 up- and 100 down-regulated genes could be detected by

Table 2
Verifications of significant gene lists from different studies

Verification	Published gene lists					Gene lists in this study				
	P1	P2	P3	P4	P5	D1	D2	D3	D4	Meta-analysis
No verification[a]	68	80	176	66	42	165	188	125	62	365
Cross-verification[b]	14	15	18	6	5	23	38	28	12	NA
(% score)	**(20.6)**	**(18.8)**	**(10.2)**	**(9.1)**	**(11.9)**	**(13.9)**	**(20.2)**	**(22.4)**	**(19.4)**	NA
Meta verification[c]	NA	NA	NA	NA	NA	62	130	59	16	215
(% score)	NA	NA	NA	NA	NA	**(37.6)**	**(69.1)**	**(47.2)**	**(25.8)**	**(58.9)**

Verification scores are shown in boldface.
[a] Total number of genes present in each gene list.
[b] Number of genes appeared in at least one of other gene lists (between P1–P5 and between D1–D4).
[c] Number of genes appearing in the meta-analysis result (for D1–D4) or number of genes appearing in at least one of D1–D4 (for meta-analysis).

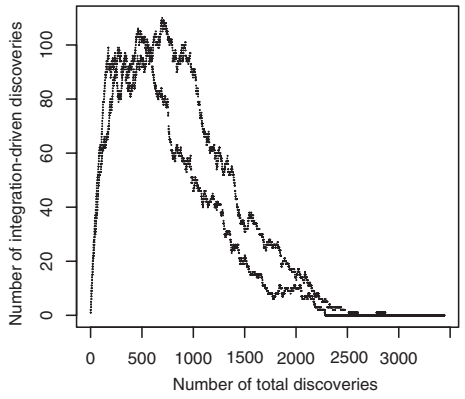

Fig. 1. Number of integration-driven discoveries versus number of total discoveries. Up- and down-regulated genes were treated separately. Total discoveries are genes called significant in the meta-analysis result with a given z threshold. Integration-driven discoveries are genes that are not significant in any of individual results but that are significant in the combined result.

none of the individual analyses using the same statistical significance level.

3.4. Biological implications of identified genes

Biological aspects of 477 dysregulated genes were explored using GO terms. We discovered GO terms with high frequency of dysregulation. The odds of dysregulation was given by the ratio of the number of dysregulated genes to the number of the other genes (in G2). Odds ratio (OR) was computed as the ratio of the odds of dysregulation of the GO term to the odds of dysregulation of the category to which the GO term belongs (i.e., biological process, molecular function, or cellular component). For example, of 110 genes associated with lipid metabolism, we found 17 dysregulated genes. Since lipid metabolism belongs to biological process, OR is computed as $(17/(110-17))/(146/(2145-146))$. GO terms with more than five dysregulated genes and OR $\geqslant 1.5$ are listed in Table 3.

Remarkably, inspection of GO terms revealed important liver-specific signatures that represent main liver functions (e.g., lipid metabolism, amino acid metabolism, carbohydrate metabolism, hormone metabolism, biosynthesis, circulation, complement component, etc.). About 90% of the corresponding dysregulated genes were down-regulated except for ribosomal proteins and eukaryotic translation initiation factors. The term "intracellular signaling cascade" in Table 3 includes MAPKKK, JNK, and JAK/STAT cascades, G-protein signaling coupled to cAMP, and RAS protein signal transduction. "Receptor signaling protein" in molecular function is closely linked to "intracellular signaling cascade" sharing several genes (e.g., small GTPase regulatory/interacting proteins are connected with RAS protein signal transduction). "Chemotaxis" contains chemokines and chemokine receptors.

The gene names and quality scores of integration-driven discoveries of $z_{th} = 2.5$ were given at http://centi.kribb.re.kr/MMA. We selected 70 interesting genes from the perspective of oncogenesis (see Table 4). The list contains previously reported genes or their related genes (e.g., the same family or a different subunit) including RPL35, PSMD4, EIF3S7, EIF3S8, GSTA2, NNMT, UGT2B7, PON3, HPD, FGA, C6, C3, ALB, and CYP7B1, and genes expected to be dysregulated in terms of liver functions, including TPI1, TRIP6, DGUOK, DUT, GBE1, FDFT1, ACAT1, ACAA1, ARG1, SERPIN1C, ALOX15, SCP2, SLC25A20, GSTA3, TPO, SORD, ODC1, and FUT1.

The list also contains interesting genes not reported in previous expression profiling studies. Some of these genes are worthy of further investigation in association with HCC. We illustrate three down-regulated genes: growth hormone receptor (GHR), tissue factor pathway inhibitor-2 (TFPI-2), and regucalcin or senescence marker protein 30 and two up-regulated genes: membrane cofactor protein (MCP or CD46) and erythropoietin receptor (EPOR).

Growth harmone receptor has eight different 5′-untranslated region variants (V1–V8) and the exclusive expression of V1 variant in adult liver has been reported. Marked or complete suppression of V1 in HCC as compared to paired normal liver has also been found [17]. According to Fernandez et al. [18] and Ji et al. [19], growth hormone (GH) administration to rat hepatoma cells desensitizes the JAK2/STAT5 pathway, possibly because of the down-regulation of GHR. These findings may explain why GH fails to stimulate growth in hepatoma

Table 3
Biological aspects of the genes dysregulated in HCC

Gene ontology[a]	Dysregulated[b]	Total[c]	OR[d]
Biological process	146	2145	1.0
Lipid metabolism	17	110	2.5
(Steroid metabolism)	(9)	(30)	(5.9)
Amino acid and derivative metabolism	7	37	3.2
Carbohydrate metabolism	7	58	1.9
Hormone metabolism	5	11	11.4
Biosynthesis	16	137	1.8
Circulation	8	34	4.2
Metal ion transport	5	38	2.1
Chemotaxis	5	41	1.9
Cell adhesion	9	90	1.5
Apoptosis	9	92	1.5
Intracellular signaling cascade	13	99	2.1
(Small GTPase mediatedsignal transduction)	(5)	(19)	(4.9)
Molecular function	155	2161	1.0
Complement component	5	9	16.2
Enzyme activator	6	39	2.4
Enzyme inhibitor	5	47	1.5
Serin-type peptidase	5	39	1.9
Transferase, transferring hexosyl groups	6	24	4.3
Oxidoreductase	11	106	1.5
Receptor signaling protein	9	88	1.5
(Small GTPaseregulatory/interacting protein)	(5)	(35)	(2.2)
Cellular component	118	1729	1.0
Extracellular	24	205	1.8
(Extracellular space)	(13)	(115)	(1.7)
(Extracellular matrix)	(6)	(47)	(2.0)
Ribonucleoprotein complex	5	40	2.0
Cytosol	9	78	1.8

[a] Gene ontology terms with more than five dysregulated genes and OR $\geqslant 1.5$.
[b] Number of genes associated with the GO term among 477 dysregulated genes.
[c] Number of genes associated with the GO term among 9982 genes in G2.
[d] Odds ratio given by the ratio of the odds of dysregulation of the GO term to the odds of dysregulation in the overall category (e.g., for lipid metabolism, OR = $(17/(110-17))/(146/(2145-146))$).

cells as it does in other cancer cells [20–23]. Therefore, transcriptional regulation of GHR V1 in HCC is worth the investigation in association with GH-mediated JAK2/STAT5 signaling.

Tissue factor pathway inhibit-2 inhibits several extracellular matrix degrading serine proteases and may play a role in tumor invasion and metastasis. The regulatory role of TFPI-2 in the invasiveness of human lung, prostate and brain cancer cell was suggested based on its mRNA level [24–26]. It was also shown that the expression of TFPI-2 is correlated inversely with the progression of human gliomas [27]. Since invasion is a characteristic feature of HCC, the role of TFPI-2 in the HGF-induced invasion of HCC cells [28] is worth the investigation.

Regucalcin is preferentially expressed in liver. Inagaki and Yamaguchi [29] suggested that regucalcin plays a suppressive role in the enhancement of various protein kinase activities associated with the proliferation of the cloned rat hepatoma cells. Over-expressing regucalcin in rat hepatoma cell was shown to result in the significant suppression of cell proliferation and DNA synthesis [30]. Therefore, HCC may deactivate one of tumor-suppressing factors by decreasing the transcript level of regucalcin.

Membrane cofactor protein is one of the complement regulatory proteins and is widely distributed in human organs. Kinugasa et al. [31], using Western blot analysis, reported that the protein level of MCP in HCC was significantly higher than that in both liver cirrhosis and chronic hepatitis. Our result suggests that MCP is regulated at the mRNA level during tumor progression. HCC may acquire increased MCP expression and escape from tumor-specific complement-mediated cytotoxicity.

Because autonomous erythropoietin (EPO) expression can mediate the autocrine growth of EPOR-bearing erythrocytes, the expressions of EPO and EPOR by tumors of other tissues may also stimulate cancer. Recent studies have reported the expression of EPOR in various human cancer cells and the possible contribution of EPO–EPOR pathway to the promotion of cancer [32–34]. EPO is produced mainly in kidney and, to a lesser extent, in liver. Autocrine role for EPO signaling in promoting the growth of renal cancers has been suggested [35]. The increase of EPOR in HCC is also notable in that increased serum EPO level often characterizes HCC.

4. Discussion

In our previous study, an effect size model was employed to analyze microarray data. Here, the effect size method was used to discover significant genes and to confirm the results by interstudy validation from our multi-center microarray studies on HCC. There are two distinct approaches in effect size method: a fixed effects model (FEM) and a random effects model (REM). For a FEM, we assume that the studies under examination share the common true effect size μ, and that the differences of the actual effect sizes are due to sampling errors alone. In a REM, the differences are due not only to sampling errors, but also to factors other than chance, such as measurement errors and inherent differences between studies. Therefore, each study has its own true effect size, which can be considered as being sampled from a super-population of possible studies with the variance of τ^2.

Choosing a suitable model for given studies upon the effect size method is quite challenging. In this study, the QQ plot led to the conclusion that the heterogeneity is within normal sampling error supporting a FEM. However, because this is merely based on the measure of the magnitude of the heterogeneity, we needed to investigate the real cause of the heterogeneity. This would be possible by using replicated spots having non-zero effect sizes. Fortunately, our clone set had 12 such lactotransferrin (LTF) spots. When applying one-way repeated measures ANOVA to the effect sizes of the data, we found that the variability between the experiments is larger than expected from the variability within the experiments (data not shown). Although we could not extend the conclusion of the LTF case to every gene without global evidence, this finding supports the idea of a REM and presents the potential problem of estimating the heterogeneity merely based on the Q statistic.

When we are not sure which model is suitable, Bayesian would offer unified modeling approach and allow us to overcome the debate over the appropriateness of a FEM and a REM. Another approach to this issue is to select a model a priori according to the type of inference to be made. FEMs are appropriate for the inferences that extend only to the studies included in meta-analysis, whereas REMs facilitate the inferences that generalize beyond the given studies.

During the process of combining multiple microarray datasets, we found that the data integration increased sensitivity for the detection of small expression changes, which the individual analyses were unable to detect. The genes found this way appeared to be meaningful from the perspective of liver functions and tumorigenesis of HCC. They also included the genes that had not been reported in previous microarray studies, to the best of our knowledge. Furthermore, it appeared that several of these new genes play important roles in tumor progression, based on the findings of previous HCC studies.

Statistical aspect of these integration-driven discoveries was discussed in [8]. They occur when we put together weak but consistent effects. They are false negatives by individual analyses. We believe that collecting separate datasets brings the same effect of increasing sample size, and that this increased sample size in turn decreases the chance of false negatives. Another important aspect of integration-driven discoveries concerns homogeneity between individual study effects. This fact is supported by the quite large P values of the Q statistics in Table 4 and ensures that the separate results are in good agreement. It should also be emphasized that this interstudy validation does more than simply increasing sample size, because it synthesizes studies performed under various experimental conditions and study populations. Thus, it enables us to overcome the limitation of single microarray studies. The risk of making conclusions based on a single-set analysis was made apparent by our cross-verification analysis.

From these findings, we propose that the integration of more than one independent dataset should be encouraged to be widely used in a variety of microarray applications. By enhancing both the reliability and sensitivity of analysis, meta-analysis would provide an opportunity of unearthing new target genes that are beyond the reach of conventional single-set analyses. Moreover, public archives will facilitate a variety of applications of microarray meta-analyses on larger scales.

Table 4
Known genes selected from integration-driven discoveries

Gene description	D1 (SE)	D2 (SE)	D3 (SE)	D4 (SE)	Avg (SE)	z_{avg}	$P(Q)$
30 up-regulated genes							
Ribosomal protein L35 (**RPL35**)	0.885 (0.370)	0.628 (0.309)	0.799 (0.494)	0.702 (0.682)	0.742 (0.204)	3.635	0.95992
Triosephosphate isomerase 1 (**TPI1**)	0.810 (0.374)	0.546 (0.304)	0.534 (0.489)	0.740 (0.499)	0.646 (0.195)	3.306	0.9427
Thyroid hormone receptor interactor 6 (**TRIP6**)	0.705 (0.376)	0.537 (0.304)	0.426 (0.487)	1.089 (0.539)	0.639 (0.198)	3.233	0.79383
H2A histone family, member Z	0.804 (0.374)	0.385 (0.298)	0.613 (0.490)	0.855 (0.460)	0.611 (0.191)	3.194	0.77173
Proteasome 26S subunit (**PSMD4**)	0.777 (0.367)	0.448 (0.299)	0.571 (0.489)	0.720 (0.470)	0.601 (0.191)	3.143	0.90536
Deoxyguanosine kinase (**DGUOK**)	0.477 (0.466)	0.615 (0.305)	0.565 (0.490)	1.239 (0.593)	0.657 (0.212)	3.106	0.76051
MGC11287 similar to ribosomal protein S6 kinase	0.725 (0.405)	0.575 (0.304)	0.619 (0.491)	0.548 (0.508)	0.615 (0.200)	3.069	0.99076
Transmembrane 9 superfamily member 1	0.934 (0.443)	0.404 (0.301)	0.547 (0.489)	0.825 (0.475)	0.613 (0.201)	3.049	0.74676
Eukaryotic translation initiation factor 3, subunit 7 (**EIF3S7**)	0.744 (0.398)	0.736 (0.305)	0.133 (0.487)	0.596 (0.624)	0.616 (0.205)	3.01	0.74289
DUTP pyrophosphatase (**DUT**)	0.817 (0.417)	0.512 (0.300)	0.148 (0.485)	1.018 (0.512)	0.598 (0.200)	2.985	0.59423
Eukaryotic translation initiation factor 3, subunit 8 (**EIF3S8**)	0.921 (0.372)	0.308 (0.297)	0.571 (0.489)	0.640 (0.476)	0.565 (0.192)	2.948	0.63818
Corticotropin releasing hormone receptor 1	0.687 (0.460)	0.516 (0.300)	0.214 (0.489)	1.620 (0.653)	0.612 (0.211)	2.893	0.36504
Erythropoietin receptor	0.810 (0.467)	0.696 (0.307)	0.358 (0.486)	0.285 (0.502)	0.588 (0.207)	2.842	0.81646
Glucan (1,4-alpha-), branching enzyme 1 (**GBE1**)	0.637 (0.381)	0.361 (0.301)	0.888 (0.498)	0.477 (0.471)	0.533 (0.194)	2.743	0.8195
Kinesin family member 3C	0.532 (0.366)	0.379 (0.298)	0.830 (0.495)	0.988 (0.709)	0.548 (0.201)	2.731	0.79239
Peroxisome biogenesis factor 13	0.523 (0.378)	0.342 (0.307)	1.046 (0.501)	0.576 (0.493)	0.538 (0.197)	2.728	0.69566
Serine/threonine kinase 25 (oxidant stress response kinase1)	0.912 (0.378)	0.513 (0.300)	0.021 (0.484)	0.517 (0.600)	0.541 (0.199)	2.717	0.54621
NADH dehydrogenase (ubiquinone) 1 alpha	0.638 (0.362)	0.401 (0.298)	0.638 (0.490)	0.481 (0.463)	0.515 (0.190)	2.711	0.95439
Farnesyl-diphosphate farnesyltransferase 1 (**FDFT1**)	0.508 (0.407)	0.353 (0.301)	0.985 (0.499)	0.665 (0.506)	0.540 (0.200)	2.705	0.74079
Thioredoxin	0.445 (0.358)	0.364 (0.297)	0.915 (0.497)	0.651 (0.476)	0.514 (0.190)	2.698	0.79505
Transcription elongation factor B (SIII), polypeptide	0.631 (0.388)	0.255 (0.296)	0.468 (0.488)	1.141 (0.491)	0.523 (0.195)	2.687	0.47697
Heterogeneous nuclear ribonucleoprotein A1	0.531 (0.425)	0.595 (0.312)	0.124 (0.486)	0.802 (0.474)	0.537 (0.202)	2.658	0.78477
APEX nuclease (multifunctional DNA repair enzyme)	0.670 (0.429)	0.601 (0.315)	0.381 (0.498)	0.332 (0.487)	0.531 (0.205)	2.588	0.93784
Membrane cofactor protein	0.035 (0.486)	0.714 (0.304)	0.693 (0.500)	0.437 (0.524)	0.539 (0.210)	2.568	0.67366
MutS (*E. coli*) homolog 6	0.385 (0.375)	0.744 (0.305)	0.233 (0.485)	0.331 (0.468)	0.496 (0.194)	2.563	0.76065
Capping protein (actin filament), gelsolin-like	0.467 (0.370)	0.638 (0.306)	0.556 (0.489)	0.144 (0.476)	0.496 (0.194)	2.557	0.85334
Translocating chain-associating membrane protein (TRAM)	0.446 (0.376)	0.693 (0.304)	0.015 (0.484)	0.652 (0.549)	0.506 (0.198)	2.554	0.68194
DEAD/H box polypeptide 9 (RNA helicase A)	0.652 (0.376)	0.374 (0.297)	0.621 (0.490)	0.448 (0.593)	0.501 (0.198)	2.523	0.93803
Myosin, heavy polypeptide 3	0.168 (0.438)	0.549 (0.304)	0.561 (0.492)	1.052 (0.593)	0.527 (0.208)	2.529	0.6901
Kinesin family member 3C	0.448 (0.480)	0.574 (0.301)	0.554 (0.506)	0.423 (0.509)	0.522 (0.208)	2.511	0.99238
40 down-regulated genes							
Multiple endocrine neoplasia I	−0.717 (0.383)	−0.764 (0.316)	−1.230 (0.513)	−1.061 (0.470)	−0.875 (0.199)	−4.388	0.81896
Glutathione *S*-transferase A2 (**GSTA2**)	−0.826 (0.368)	−0.791 (0.317)	−0.591 (0.490)	−0.970 (0.466)	−0.800 (0.196)	−4.091	0.95599
Nicotinamide *N*-methyltransferase (**NNMT**)	−0.622 (0.368)	−0.687 (0.303)	−1.259 (0.508)	−0.285 (0.573)	−0.707 (0.199)	−3.549	0.61856
Core-binding factor, runt domain (MTG16)	−0.855 (0.445)	−0.555 (0.311)	−1.177 (0.531)	−0.837 (0.636)	−0.761 (0.216)	−3.522	0.774
UDP glycosyltransferase 2 family, B7 (**UGT2B7**)	−0.849 (0.369)	−0.453 (0.299)	−0.570 (0.495)	−1.334 (0.572)	−0.690 (0.197)	−3.496	0.54335
Growth hormone receptor	−0.609 (0.374)	−0.356 (0.297)	−1.191 (0.505)	−1.191 (0.503)	−0.674 (0.195)	−3.461	0.35076

Table 4 (*continued*)

Gene description	D1 (SE)	D2 (SE)	D3 (SE)	D4 (SE)	Avg (SE)	z_{avg}	$P(Q)$
Paraoxonase 3 (**PON3**)	−0.460 (0.383)	−0.762 (0.305)	−0.843 (0.502)	−0.441 (0.465)	−0.639 (0.196)	−3.265	0.86687
Acetyl-coenzyme A acetyltransferase 1 (**ACAT1**)	−0.539 (0.399)	−0.636 (0.309)	−0.381 (0.494)	−1.344 (0.557)	−0.662 (0.204)	−3.246	0.58791
4-Hydroxyphenylpyruvate dioxygenase (**HPD**)	−0.390 (0.382)	−0.544 (0.322)	−1.103 (0.504)	−1.053 (0.552)	−0.663 (0.206)	−3.226	0.59179
Arginase, liver (**ARG1**)	−0.969 (0.415)	−0.652 (0.303)	−0.379 (0.489)	−0.415 (0.454)	−0.635 (0.197)	−3.221	0.76255
Fibrinogen, A alpha (**FGA**)	−0.718 (0.377)	−0.719 (0.308)	−0.644 (0.501)	−0.250 (0.458)	−0.623 (0.195)	−3.197	0.84366
Complement component 6 (**C6**)	−0.821 (0.426)	−0.433 (0.302)	−1.321 (0.597)	−0.567 (0.465)	−0.652 (0.204)	−3.19	0.57828
Antithrombin III (**SERPINC1**)	−0.467 (0.383)	−0.608 (0.302)	−1.081 (0.503)	−0.390 (0.445)	−0.601 (0.193)	−3.11	0.73923
Neural cell adhesion molecule 1	−0.671 (0.462)	−0.668 (0.306)	−0.815 (0.508)	−0.336 (0.489)	−0.634 (0.207)	−3.066	0.91506
Tissue factor pathway inhibitor 2	−0.562 (0.411)	−0.547 (0.300)	−0.628 (0.490)	−0.816 (0.494)	−0.608 (0.199)	−3.055	0.97225
Albumin (**ALB**)	−0.712 (0.365)	−0.464 (0.299)	−0.675 (0.491)	−0.475 (0.447)	−0.565 (0.189)	−2.98	0.94691
Complement component 3 (**C3**)	−0.783 (0.367)	−0.486 (0.299)	−0.763 (0.498)	−0.245 (0.485)	−0.572 (0.193)	−2.964	0.7974
Microsomal triglyceride transfer protein	−0.923 (0.398)	−0.377 (0.301)	−0.213 (0.488)	−1.103 (0.528)	−0.590 (0.199)	−2.962	0.43322
Arachidonate 15-lipoxygenase (**ALOX15**)	−0.596 (0.380)	−0.664 (0.306)	−0.926 (0.497)	0.014 (0.504)	−0.583 (0.198)	−2.948	0.58218
Transcription factor (TFIIB)	−0.435 (0.376)	−0.491 (0.306)	−0.720 (0.492)	−0.979 (0.521)	−0.583 (0.198)	−2.945	0.82523
Sterol carrier protein X (**SCP2**)	−0.283 (0.373)	−0.753 (0.309)	−0.110 (0.484)	−1.043 (0.469)	−0.572 (0.194)	−2.944	0.41405
Lymphocyte antigen 6 complex, locus E	0.065 (0.372)	−0.731 (0.308)	−1.031 (0.500)	−0.957 (0.525)	−0.584 (0.198)	−2.943	0.20521
Carnitine/acylcarnitine translocase (**SLC25A20**)	−1.091 (0.457)	−0.374 (0.312)	−0.207 (0.494)	−1.132 (0.530)	−0.614 (0.210)	−2.928	0.3452
Glutathione *S*-transferase A3 (**GSTA3**)	−0.823 (0.374)	−0.552 (0.300)	−0.508 (0.488)	−0.235 (0.451)	−0.559 (0.191)	−2.923	0.79564
Acetyl-coenzyme A acyltransferase 1 (**ACAA1**)	−0.187 (0.366)	−0.707 (0.330)	−0.812 (0.494)	−0.715 (0.470)	−0.572 (0.199)	−2.874	0.65876
Thyroid peroxidase (**TPO**)	−0.294 (0.462)	−0.341 (0.327)	−1.056 (0.509)	−1.300 (0.530)	−0.619 (0.216)	−2.867	0.30696
Sorbitol dehydrogenase (**SORD**)	−0.680 (0.405)	−0.512 (0.300)	−0.331 (0.490)	−0.714 (0.470)	−0.558 (0.196)	−2.841	0.93197
Carboxypeptidase B2 (plasma)	−0.314 (0.356)	−0.601 (0.301)	−0.773 (0.493)	−0.537 (0.457)	−0.534 (0.190)	−2.814	0.88123
Jagged 1 hematopoiesis	−1.023 (0.455)	−0.733 (0.305)	0.045 (0.492)	−0.220 (0.476)	−0.564 (0.203)	−2.772	0.33685
Angiogenin, ribonuclease (ANG)	−0.730 (0.365)	−0.291 (0.296)	−1.043 (0.500)	−0.283 (0.443)	−0.515 (0.189)	−2.722	0.51106
Regucalcin (senescence marker protein-30)	−0.906 (0.397)	−0.316 (0.297)	−0.285 (0.490)	−0.793 (0.492)	−0.530 (0.196)	−2.705	0.58168
cdc-like kinase 2 (CLK2)	−1.271 (0.520)	−0.395 (0.316)	−0.417 (0.491)	−0.551 (0.502)	−0.576 (0.214)	−2.693	0.52682
Claudin 1 (senescence-associated epithelial membrane protein)	−0.783 (0.386)	−0.267 (0.300)	−1.074 (0.501)	−0.241 (0.466)	−0.515 (0.194)	−2.648	0.4299
Chaperonin containing TCP1, subunit 4 (delta)	−0.502 (0.359)	−0.335 (0.300)	−0.774 (0.493)	−0.890 (0.583)	−0.518 (0.196)	−2.636	0.78911
Integrin, beta 1 (fibronectin receptor)	−0.784 (0.367)	−0.327 (0.297)	−0.928 (0.497)	−0.110 (0.441)	−0.496 (0.189)	−2.621	0.48275
Ornithine decarboxylase 1 (**ODC1**)	−0.587 (0.367)	−0.584 (0.301)	−0.826 (0.494)	0.172 (0.484)	−0.501 (0.193)	−2.595	0.47615
Cathepsin O	−0.342 (0.388)	−0.632 (0.309)	−0.554 (0.490)	−0.435 (0.470)	−0.510 (0.197)	−2.592	0.94509
Cytochrome P450, subfamily VIIB polypeptide 1 (**CYP7B1**)	−0.495 (0.365)	−0.596 (0.301)	−0.558 (0.490)	−0.177 (0.457)	−0.490 (0.191)	−2.567	0.89381
Vanin 1	−0.585 (0.361)	−0.232 (0.326)	−0.953 (0.498)	−0.539 (0.483)	−0.505 (0.198)	−2.546	0.66803
Fucosyltransferase 1 (**FUT1**)	−0.312 (0.356)	−0.638 (0.302)	−0.556 (0.489)	−0.382 (0.575)	−0.496 (0.196)	−2.534	0.90936

Of integration-driven discoveries with $z_{th} = 2.5$, 30 up-regulated and 40 down-regulated genes were selected to be displayed. D1–D4, effect size for D1–D4; S.E., standard error; Avg, average effect size calculated based on FEM; z_{avg}, z score of average effect size; $P(Q)$, P value of Q statistic. Larger $P(Q)$ indicates that the effect sizes are more homogeneous. Genes mentioned in text are indicated by bold type.

Our work established the methodology of using an effect size model for microarray meta-analysis and demonstrated its effectiveness with an application to HCC studies. It was shown that the meta-analysis could offer more exact and extended clues into liver carcinogenesis. Here, we reported upon genes newly discovered using our method and some of the genes that call for further investigation. We believe that these genes identified by the integration of expression profiles will helpfully expand our knowledge of the mechanism of HCC progression.

Acknowledgements: This work was supported by Grant No. FG-5-01 of the 21C Frontier Functional Human Genome Project from the Ministry of Science & Technology of Korea. The authors thank Drs. U. Yu and Y. Hahn for their critical review of the manuscript.

References

[1] Brazma, A., Hingamp, P., Quackenbush, J., Sherlock, G., Spellman, P., Stoeckert, C., Aach, J., Ansorge, W., Ball, C.A. and Causton, H.C., et al. (2001) Nat. Genet. 29, 365–371.
[2] Hedges, L.V. and Olkin, I. (1985) Statistical Methods for Meta-Analysis. Academic Press, New York.
[3] Rosenthal, R. (1994) in: The Handbook of Research Synthesis (Cooper, H.M. and Hedges, L.V., Eds.), pp. 231–244, Russell Sage Foundation, New York.
[4] Cooper, H.M. (1998) Synthesizing Research: A Guide for Literature Reviews. Sage, Newbury Park.
[5] DerSimonian, R. and Laird, N.M. (1986) Control. Clin. Trials 7, 177–188.
[6] Rhodes, D.R., Barrette, T.R., Rubin, M.A., Ghosh, D. and Chinnaiyan, A.M. (2002) Cancer Res. 62, 4427–4433.
[7] Xin, W., Rhodes, D.R., Ingold, C., Chinnaiyan, A.M. and Rubin, M.A. (2003) Am. J. Pathol. 23, 255–261.
[8] Choi, J.K., Yu, U., Kim, S. and Yoo, O.J. (2003) Bioinformatics 19 (Suppl. 1), i84–i90.
[9] Okabe, H., Satoh, S., Kato, T., Kitahara, O., Yanagawa, R., Yamaoka, Y., Tsunoda, T., Furukawa, Y. and Nakamura, Y. (2001) Cancer Res. 61, 2129–2137.
[10] Tackels-Horne, D., Goodman, M.D., Williams, A.J., Wilson, D.J., Eskandari, T., Vogt, L.M., Boland, J.F., Scherf, U. and Vockley, J.G. (2001) Cancer 92, 395–405.
[11] Xu, L., Hui, L., Wang, S., Gong, J., Jin, Y., Wang, Y., Ji, Y., Wu, X., Han, Z. and Hu, G. (2001) Cancer Res. 61, 3176–3181.
[12] Li, Y., Li, Y., Tang, R., Xu, H., Qiu, M., Chen, Q., Chen, J., Fu, Z., Ying, K., Xie, Y. and Mao, Y. (2002) J. Cancer Res. Clin. Oncol. 128, 369–379.
[13] Chung, E.J., Sung, Y.K., Farooq, M., Kim, Y., Im, S., Tak, W.Y., Hwang, Y.J., Kim, Y.I., Han, H.S. and Kim, J.C., et al. (2002) Mol. Cells 14, 382–387.
[14] Yang, Y.H., Dudoit, S., Luu, P., Lin, D.M., Peng, V., Ngai, J. and Speed, T.R. (2002) Nucleic Acids Res. 30, e15.
[15] Cochran, B.G. (1954) Biometrics 10, 101–129.
[16] Tusher, V.G., Tibshirani, R. and Chu, G. (2001) Proc. Natl. Acad. Sci. USA 98, 5116–5121.
[17] Zogopoulos, G., Albrecht, S., Pietsch, T., Alpert, L., von Schweinitz, D., Lefebvre, Y. and Goodyer, C.G. (1996) Cancer Res. 56, 2949–2953.
[18] Fernandez, L., Flores-Morales, A., Lahuna, O., Sliva, D., Norstedt, G., Haldosen, L.A., Mode, A. and Gustafsson, J.A. (1998) Endocrinology 139, 1815–1824.
[19] Ji, S., Frank, S.J. and Messina, J.L. (2002) J. Biol. Chem. 277, 28384–28393.
[20] Kolle, S., Sinowatz, F., Boie, G., Temmim-Baker, L. and Lincoln, D. (1999) Int. J. Oncol. 14, 911–916.
[21] Lincoln, D.T., Sinowatz, F., Kolle, S., Takahashi, H., Parsons, P. and Waters, M. (1999) Anticancer Res. 19, 1919–1931.
[22] Temmim, L., Kolle, S., Baker, H. and Sinowatz, F. (2000) Oncol. Rep. 7, 757–760.
[23] Gebre-Medhin, M., Kindblom, L.G., Wennbo, H., Tornell, J. and Meis-Kindblom, J.M. (2001) Am. J. Pathol. 158, 1217–1222.
[24] Lakka, S.S., Konduri, S.D., Mohanam, S., Nicolson, G.L. and Rao, J.S. (2000) Clin. Exp. Metastasis 18, 239–244.
[25] Konduri, S.D., Tasiou, A., Chandrasekar, N. and Rao, J.S. (2001) Int. J. Oncol. 18, 127–131.
[26] Konduri, S.D., Rao, C.N., Chandrasekar, N., Tasiou, A., Mohanam, S., Kin, Y., Lakka, S.S., Dinh, D., Olivero, W.C. and Gujrati, M., et al. (2001) Oncogene 20, 6938–6945.
[27] Rao, C.N., Lakka, S.S., Kin, Y., Konduri, S.D., Fuller, G.N., Mohanam, S. and Rao, J.S. (2001) Clin. Cancer Res. 7, 570–576.
[28] Neaud, V., Faouzi, S., Guirouilh, J., Le Bail, B., Balabaud, C., Bioulac-Sage, P. and Rosenbaum, J. (1997) Hepatology 26, 1458–1466.
[29] Inagaki, S. and Yamaguchi, M. (2001) J. Cell. Biochem. 81, 12–18.
[30] Misawa, H., Inagaki, S. and Yamaguchi, M. (2001) J. Cell. Biochem. 84, 143–149.
[31] Kinugasa, N., Higashi, T., Nouso, K., Nakatsukasa, H., Kobayashi, Y., Ishizaki, M., Toshikuni, N., Yoshida, K., Uematsu, S. and Tsuji, T. (1999) Br. J. Cancer 80, 1820–1825.
[32] Acs, G., Acs, P., Beckwith, S.M., Pitts, R.L., Clements, E., Wong, K. and Verma, A. (2001) Cancer Res. 61, 3561–3565.
[33] Arcasoy, M.O., Jiang, X. and Haroon, Z.A. (2003) Biochem. Biophys. Res. Commun. 307, 999–1007.
[34] Yasuda, Y., Fujita, Y., Matsuo, T., Koinuma, S., Hara, S., Tazaki, A., Onozaki, M., Hashimoto, M., Musha, T. and Ogawa, K., et al. (2003) Carcinogenesis 24, 1021–1029.
[35] Westenfelder, C. and Baranowski, R.L. (2000) Kidney Int. 58, 647–657.