

Review

Microarray Meta-Analysis and Cross-Platform Normalization: Integrative Genomics for Robust Biomarker Discovery

Christopher J. Walsh ^{1,2}, Pingzhao Hu ³, Jane Batt ^{1,2} and Claudia C. Dos Santos ^{1,2,*}

¹ Keenan and Li Ka Shing Knowledge Institute of Saint Michael's Hospital, Toronto ON M5B 1W8, Canada; E-Mails: chrisj.walsh@mail.utoronto.ca (C.J.W.); batt.jane@utoronto.ca (J.B.)

² Institute of Medical Sciences and Department of Medicine, University of Toronto, Toronto ON M5B 1W8, Canada

³ Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg R3E 0J9, MB, Canada; E-Mail: pingzhao.hu@umanitoba.ca

* Author to whom correspondence should be addressed; E-Mail: dossantosC@smh.ca; Tel.: +1-416-864-8575; Fax: +1-416-864-6013.

Academic Editor: Quan-Zhen Li

Received: 6 July 2015 / Accepted: 17 August 2015 / Published: 21 August 2015

Abstract: The diagnostic and prognostic potential of the vast quantity of publicly-available microarray data has driven the development of methods for integrating the data from different microarray platforms. Cross-platform integration, when appropriately implemented, has been shown to improve reproducibility and robustness of gene signature biomarkers. Microarray platform integration can be conceptually divided into approaches that perform early stage integration (cross-platform normalization) *versus* late stage data integration (meta-analysis). A growing number of statistical methods and associated software for platform integration are available to the user, however an understanding of their comparative performance and potential pitfalls is critical for best implementation. In this review we provide evidence-based, practical guidance to researchers performing cross-platform integration, particularly with an objective to discover biomarkers.

Keywords: microarray platform; meta-analysis; normalization; biomarker

1. Introduction

The discovery of highly-reliable biomarkers from high-dimensional microarray data is an important goal in molecular medicine, with wide-ranging clinical applications. Potential roles for biomarkers include early detection of disease in healthy individuals, disease classification, prognosis, prediction of response to therapy, and as surrogate outcomes in clinical trials [1]. The ideal biomarker is inexpensive, robust, easily interpretable, well-validated, and clinically useful (e.g., improving prognosis or choice of therapy) compared to current standards of practice, meaning that the result is “actionable, leading to patient benefit” [1]. Publicly-available microarray data has vast potential to serve as a source of biomarker discovery as there is an enormous quantity of existing gene expression data [2,3]. At the present time, the Gene Expression Omnibus, a repository of array- and sequence-based expression data, currently contains 1,413,278 samples performed on 14,346 platforms [4]. The most widely known of these platforms include the Affymetrix GeneChips (*in situ* synthesized oligonucleotide microarray) and the Illumina high-density bead arrays [5]. While other types of microarrays exist, such as protein and microRNA [6,7], this review will focus on integration of gene expression data from multiple cDNA microarray platforms as it relates to the discovery of gene signatures that may serve as biomarkers for clinical applications. The integration of multiple data types (e.g., transcriptomic and proteomic data) has been proposed [8], however this is also beyond the scope of our paper.

While microarrays measure the expression of thousands of genes simultaneously, it is expected that only a small subset of the genes will be associated with the clinical or biological outcome of interest. This subset of genes, often termed a “gene signature” or “prognostic signature”, has a collective expression pattern that is unique to the outcome of interest and thus has potential to function as a biomarker [9]. The gene signature is typically composed of far fewer number genes (often less than 100 genes) than that on a microarray chip (often more than 20,000 genes) making it feasible for further study using approaches such as quantitative RT-PCR. Point of Care (POC) devices that rely on transcriptional signatures are progressively gaining momentum as diagnostic tools for routine use in the clinical setting, resulting from their practical and affordable application making this approach highly accessible as cheaper diagnostic kits [10,11].

Biomarkers for the monitoring of disease activity of POC are currently lacking. A number of published gene signatures validated using independent samples have been shown to serve as significant predictors of clinical outcome [12–15]. However, the development of prognostic signatures that are robust and stable (e.g., the same biomarkers are identified in both discovery and validation sets) [16] has proven challenging [17–19]. In Section 3, we will discuss recent examples of promising transcriptomic biomarkers for disease diagnosis and prognosis that have been identified using meta-analysis approaches.

Published prognostic gene signatures derived from internal validation often show little overlap with genes identified by other study groups [15]. Potential causes of small reproducibility include differences in sample collection methods, processing protocols, and microarray platforms, patient heterogeneity, and small sample sizes [12]. Due to the difficulty of acquiring samples, particularly from human tissue and the associated costs, microarray experiments from single-institution patient cohorts are often composed of small sample sizes. Predictive models trained on the gene signatures identified from these smaller-sized individual studies are less robust [15,20]. Michiels *et al.* [21]

re-analyzed data from nine studies predicting cancer prognosis and found an unstable misclassification rate for the gene signature (defined as the 50 genes for which expression was most highly correlated with outcome) using training sets derived using a re-sampling approach, with performance increasing as the size of the training set increases.

Integration of multiple microarray data sets has been advocated to improve gene signature selection [22]. Increasing sample sizes increases the statistical power to obtain a more precise estimate of integration of (differential) gene expression and to assess the heterogeneity of the overall estimate, as well as to reduce the effects of individual study-specific biases [23–26]. Meta-analysis is most commonly applied for the purpose of detecting differentially-expressed (DE) genes [27] which may serve as a candidate gene signature or be used as features in classification models or classifiers to further refine a clinically useful gene signature [28]. Supervised classification techniques (also known as prediction analysis or supervised machine learning) are the most commonly used methods in microarray analysis that lead to identification of clinically-useful biomarkers (*i.e.*, gene signatures providing improved discrimination between two or more patient groups) [27]. Classification methods for gene signature selection are beyond the scope of this article and have been reviewed elsewhere [29].

2. Integrative Transcriptomic Data Analysis

Two fundamental approaches to combine the information of multiple independent microarray studies from different platforms (termed “integrative analysis” [23]) are meta-analysis and cross-platform normalization (also termed “merging”). A conceptual framework by Hamid *et al.* [22] classifies microarray meta-analysis as “late stage” data integration as it combines the final statistic results from different studies, whereas cross-platform normalization integrates data at the “early stage”. Application of these approaches necessitates that all of the included studies are testing the same hypothesis and/or performed under comparable conditions or treatments [2,30,31]. While the degree of similarity that is required between “suitably similar” datasets still remains to be determined, cross-platform integration for the purpose of biomarker discovery is most appropriate using relatively homogenous datasets selected to answer well-defined questions [32]. Early or late stage integration of data can be used regardless of the biological question (e.g., differential expression analysis or class prediction). The overall principle of these two approaches is summarized in Figure 1.

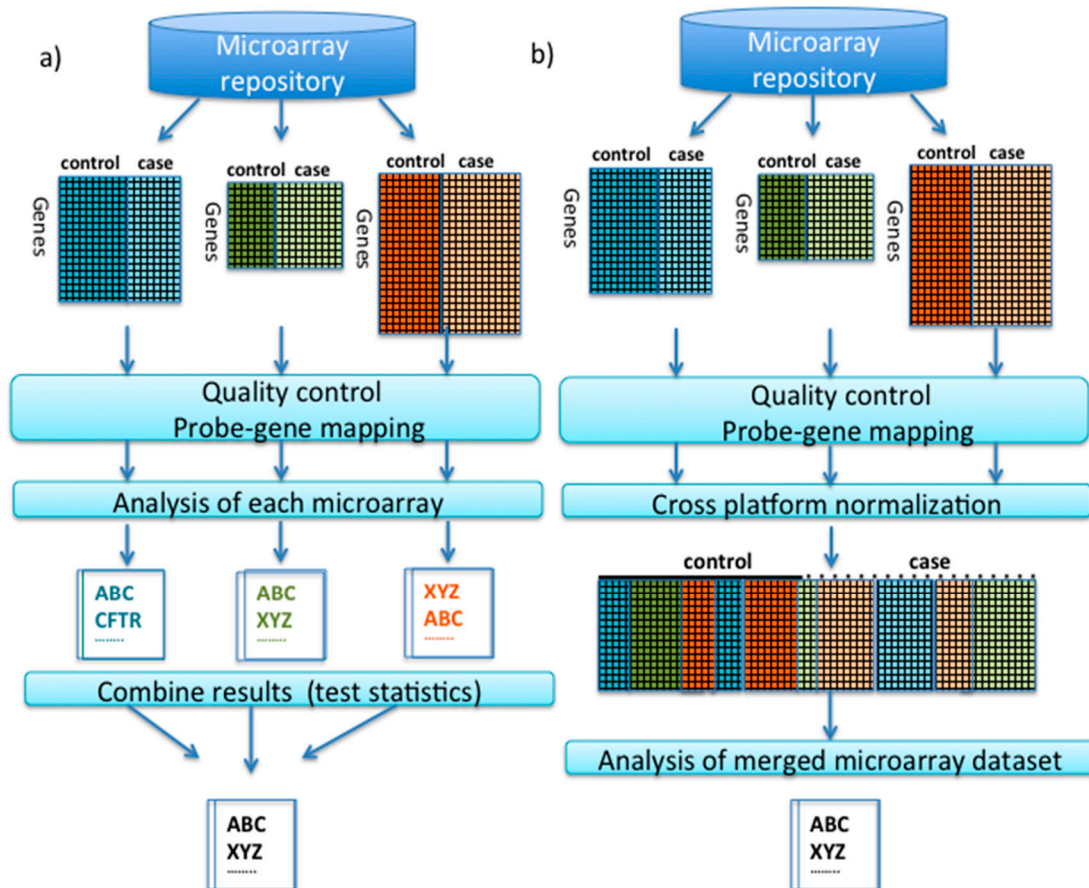


Figure 1. Outline of two microarray integration methods: (a) meta-analysis (“late integration”). Individual case-cohort microarray studies are pre-processed and each study is used to identify ranked gene lists which are then combined in the final step; (b) Cross-platform merging and normalization (“early integration”). After pre-processing of individual studies, a single unified case-cohort dataset is generated (“clustered” into cases and cohorts, indicating removal of batch to batch variation) and in this example, used to identify a ranked gene list.

2.1. Pre-Processing and Quality Control Prior to Integrative Analysis

Ramasamy *et al.* [24] identified key issues and steps for performing a meta-analysis including identifying suitable microarrays, pre-processing and preparing individual datasets, selection of meta-analysis method, and interpretation of results. A systematic review of microarray meta-analysis studies in the literature has found that the criteria to include or exclude microarray studies is mostly subjective and *ad hoc* and remains an open question in the field [27]. Two critical pre-processing steps we will highlight here are (i) removing arrays with poor quality and (ii) determining the relationships between probes and genes. Identifying microarrays of poor quality is essential prior to integrative analysis because inclusion of poor quality studies may reduce statistical power and adversely affect the outcome of meta-analysis [27,33]. There are a number of quality assessment packages available for Bioconductor, including Simpleaffy [34] and affyPLM [35] for Affymetrix. The MetaQC package provides six quality control measurements to identify problematic studies across multiple platforms for further assessment of causes of lower quality to determine their exclusion from meta-analysis [36,37].

Another important pre-processing step is ascertaining which probes represent a given gene within and across the different microarray platforms. The relationship between probes and genes may be determined by mapping probes to the gene using sequence-matched datasets or using gene-level identifiers such as Entrez Gene ID available in the annotations packages in R/Bioconductor [38] to unify the microarray datasets. Sources of high-quality probe re-annotation include alternative chip definition files (CDFs) for Affymetrix [39] and ReMOAT (Re-annotation and Mapping for Oligonucleotide Array Technologies) and its associated annotation packages in R/Bioconductor for Illumina [40]. Only genes that are present across the different platforms being integrated will remain for further analysis, while those absent in one or more platforms will be “lost”, reflecting the tradeoff between increasing sample size and power *versus* decreasing the number of genes analyzed [32]. Co-inertia analysis, a multivariate analysis method that describes the common trends or co-relationships between datasets of two conditions, has been applied to determine the loss of information incurred by reducing the number of genes to the subset common to different platforms [41]. Imputation of gene expression present in some datasets, but not others, to allow these genes to be part of predictive models has been proposed [42].

If multiple probes match a single gene, selecting the probe with the highest interquartile range (IQR) has been recommended [43]. Genes with low mean expression across most studies are typically filtered out prior to meta-analysis. Turnbull *et al.* [32] applied relatively strict filter thresholds for their microarray integration analysis based on a prior study that found genes with low or intermediate expression have poorer inter-platform reproducibility than highly-expressed genes [17,44]. Furthermore, incorporation of a quality measure based on detection *p*-values estimated from Affymetrix arrays into the study-specific test statistics within a meta-analysis of two Affymetrix array studies using an effect sized model produced more biologically meaningful results than an unweighted model [25,45].

2.2. Meta-Analysis

In the meta-analysis approach, each experiment is first analyzed separately and the results of each study are then combined. Meta-analysis methods that combine primary statistics (e.g., *p*-values or effect sizes) require the use of raw gene expression data whereas secondary statistics rely only on ranked lists of genes. Popular methods for meta-analysis mainly combine one of three types of statistics: *p*-value [46], effect size [47], and ranked gene lists (“rank aggregation”) [27,33,48]. Ranked lists of genes produced for each study (e.g., ranked by order of *p*-value for DE of each gene) have been aggregated into a single gene ranking (“consensus”) using a number of methods including the rank product method [48].

A number of methods have been developed to test the statistical significance of results based on combining *p*-values from each study including Fisher’s method, Stouffer’s method, minP, and maxP. Fisher’s method sums log-transformed *p*-values, whereas Stouffer’s method sums inverse-normal-transformed *p*-values, to combine statistical significance across studies. The minP method takes the minimum *p*-value from combined studies, whereas the maxP method takes the maximum of the combined *p*-values. Rhodes *et al.* [49] published one of the first papers to combine

p -values from individual studies of DE gene expression using Fisher's method which found improved statistical significance using the combed analysis compared to individual studies.

Combined effect size to generate an estimate of the overall effect size and its confidence interval is frequently used in meta-analysis of clinical research data. Choi *et al.* [47] described one of the first methods to combine effect sizes using a random-effects modeling approach for combining datasets from individual studies of two groups to form an overall estimate of the weighted effect size. The effect size was measured by the standardized mean difference obtained by dividing the difference in the average gene expression between the treatment and control groups by a pooled estimate of standard deviation. The effect size was used to measure the magnitude of treatment effect in each study and a random effects model was used to incorporate inter-study variability.

Meta-analysis methods have been categorized based on the hypothesis settings that gene biomarkers are differentially expressed "in all studies" (HS_A), "in the majority of studies" (HS_r), or in "one or more studies" (HS_B) [33,50]. In Fisher's, Stouffer's, and minP method, an extremely small p -value in one study likely meets criteria for statistical significance; thus, it detects DE in "in one or more studies" (HS_B), whereas the maxP or rank product method tends to detect gene biomarkers DE in "all studies" (HS_A).

The choice of the statistical meta-analysis method is selected based on the biological purpose of the analysis. A gene serving as a biomarker from a meta-analysis is expected to show concordant biological effects across all or most experiments for a given condition derived from relatively homogenous sources (e.g., up-regulation of a gene predicting risk of lung cancer detection from lung epithelium biopsied from a cohort of smokers *versus* healthy non-smokers) [51]. While detecting biomarkers DE in all studies seems an ideal goal, it can be too stringent when the number of samples is large, increasing the heterogeneity of experimental, platform, or biological samples [50]. Meta-analysis methods detecting DE in the majority of samples (HS_r) are generally recommended as they provide robustness and detection of relevant signals across the majority of samples [33]. Song and Tseng [52] proposed a robust order statistic, r th ordered p -value (rOP), which tests the alternative hypothesis that there are significant p -values in at least a given percentage of studies. This method detects biomarkers DE in the majority of studies (e.g., >70% of studies) based on a user-specific threshold of studies.

2.2.1. Comparison of Meta-Analysis Methods

Several comparative studies systematically comparing meta-analysis methods for microarray data have been previously published [33,53,54]. Chang *et al.* [33] benchmarked the performance of six p -value combination methods (Fisher, Stouffer, adaptively weighted Fisher, minP, maxP, and rOP), two combined effect size methods (fixed effects and random effects) and four combined ranks methods (RankProd, RankSum, product of ranks, and sum of ranks). The 12 meta-analysis methods were categorized into three hypothesis settings (candidate markers DE in "all" [HS_A], "most" [HS_r], or "one or more" [HS_B] studies) based on their strengths for detecting DE genes. They then applied four statistical criteria to the assessment of each meta-analysis method: (1) detection capability (the number of DE genes detected); (2) biological association (degree of association between DE list with predefined genes from pathways related to the disease), stability (randomly splitting the data and comparing results of the two-meta-analyses) and robustness (effect of including an outlying irrelevant study to the meta-analysis).

Among the methods based on HSA setting, the maxP performed the worst based on their four criteria and the investigators recommend that it be avoided. Rank product method had improved performance but weaker detection capability. The two methods that tended to detect DE in the majority of samples were the Random Effect Model (REM) and the r th order p -value (rOP). rOP outperformed REM based on stronger biological association and detection capabilities, but this was achieved at the expense of diminished stability and robustness.

It is important to note that differentially-expressed genes determined by combing p -values or ranks obtained by two-sided hypothesis testing may result in genes with discordant DE across two-class outcomes which can be difficult to interpret [27]. Wang *et al.* [37] have proposed one-sided correction of p -values to guarantee identification of DE genes with concordant DE direction.

2.2.2. Association of Meta-Analysis Method to Outcome Variable

The objective and type of outcome types (e.g., two-class, multi-class, survival) [24] will govern the choice of both the test statistic (t -statistic, F -statistic, log-rank statistic) and the meta-analysis method (combing p -values, effect sizes, or ranks). Methods combing effect sizes (standardized mean differences or odds ratios) are appropriate for combining two-class outcomes. Meta-analysis of expression studies with continuous outcomes (e.g., using regression or correlation coefficients) and survival outcomes (based on log-rank statistics) have typically been performed using combined p -values [50,55] and can be performed using the MetaDE package [37]. To capture concordant expression patterns for multi-class outcomes, Lu *et al.* [52] have applied multi-class correlation (min-MCC) because the F -statistic has been found to frequently fail to capture concordant patterns of gene expression.

2.3. Cross-Platform Normalization

Cross-platform normalization (also termed “data merging” [23]) considers all data from experiments across different microarray platforms as a single data set from the same experiment. Direct integration of data sets performed on different microarray platforms may introduce undesirable batch effects due to systematic multiplicative biases [23,32,56]. The level of difficulty present to combine multiple datasets has been termed “dataset complexity” [53]. For example, integrating different Affymetrix platforms is less complex to analyze by meta-analysis or cross platform normalization than datasets performed across very different platforms. Studies using low complexity datasets, mainly from the Affymetrix platform, have directly merged the studies to construct a gene signature [41,57–59].

Cross-platform transformation and normalization methods have been developed with an aim to remove the artifactual differences between data from different microarray platforms while preserving the underlying biological differences between conditions. This step is essential, as non-biological differences (“batch effects”) in the gene signature discovery data can obscure real biological differences found between clinical groups.

Early attempts at cross-platform merging applied straightforward transformation methods of location and scale (mean and variance) to process the gene expression data from different studies. Batch mean centering [56] is a simple transformative method that standardizes the expression of each gene to have the same center (mean expression). Probe sets can be further transformed to have the same variance or distributions on different platforms [60,61]. While these methods are relatively easy and

intuitive, the batch mean centering method has been shown to have only marginal improvement compared to uncorrected data for cross-platform integration of Illumina and Affymetrix data [32]. Probability of expression (POE), a model-based transformation that is estimated based on a method that adopts an underlying mixture distribution that transforms each data value into range $[-1,1]$ has been used for cross-platform merging based on a unified scale as an alternative to using gene-specific summaries [62,63]. While this transformation has been applied for identifying meta-signatures, it has been found to be difficult to compare to other normalization methods [26].

Over the past decade, a number of more complex cross-platform normalization methods have been published and their performance has been compared in several studies [2,32]. Four cross-platform normalization methods found to be generally effective in a comparative review by Rudy and Palafer [2] are the Empirical Bayes (EB) method, known as Combat [64]; the Cross-Platform Normalization (XPN) method [26], Distance Weighted Discrimination (DWD) [65], and the Gene Quantiles (GQ) method developed as part of the WebArrayDB service [66]. Of these four programs, the authors favour DWD and XPN, while the comparative analysis of cross-platform normalization methods on clinical datasets by Turnbull *et al.* [32] favoured Combat and XPN. We will discuss the results of these comparative analyses in more detail in the following Section 2.3.1.

The Distance Weighted Discrimination method, like Support Vector Machines (SVM), is a margin-based classification method that was developed to improve performance over the latter method. Essentially, SVM finds a hyperplane that separates the two classes (*i.e.*, each systematic bias) to maximize the minimum distance of all the data on the hyperplane (the margin). However SVM has data pile-up problems along the margin which have been improved by modifying the margin to maximize the sum of the inverse distance in DWD [67]. DWD adjusts the microarray data by projecting the different batches onto the hyperplane, finding the batch mean and then subtracting out the plane multiplied by this mean.

Combat, an empirical Bayes method, estimates parameters that represent the batch effects by pooling information across genes in each batch to shrink the batch effect parameter toward the overall mean of the batch effect estimates across genes [64]. The data are then transformed to remove the effects of the different batch effect parameters across platforms. Combat is performed using either a parametric prior method or a non-parametric method based on the prior distributions of the estimated parameters [68].

Unlike the gene-wise linear approaches of DWD and Combat, the cross-platform normalization (XPN) method developed by Shablin *et al.* [26] seeks to borrow information across genes and samples via linked row and column clusters in a two-step procedure. First, K-means clustering is used to find blocks of similar genes and samples across the platforms. This approach is robust to the number of row (K) and column (L) clusters. Then, within each block the data is normalized between platforms within this block. The normalized values obtained over multiple clustering performed over repeated runs is then averaged to better capture the data structure.

2.3.1. Comparison of Cross-Normalization Methods

A comparative analysis of cross-platform normalization methods by Rudy and Valafar [2] found the DWD classification method to provide effective batch adjustment for microarray data [67] and to be the most robust to variation in treatment group sizes between the platforms with the least loss of treatment information (lower underdetection), while XPN showed the greatest inter-platform concordance [2]. Turnbull *et al.* [32] also found that XPN had the highest inter-platform concordance. However, they found that DWD removed not only the platform specific systematic bias, but also relevant biological variability between samples (reduced inter-sample variance), while Combat and XPN preserved this biological signal (slightly increased inter-sample variance) while appropriately correcting the platform specific bias (reduced inter-platform variance). The drawback of DWD to “over-normalize” by removing all systematic expression differences between two datasets, including the relevant biological variability has been cautioned by other authors, prompting development of newer methods [60]. Although Combat and XPN have been found to perform well in previous analyses, the user must be cautious when applying this method to datasets that are unbalanced (e.g., different subtypes within each of the batches) as these methods will not be able to distinguish batch effects from biologically-relevant signals [42].

One limitation of some existing cross-platform normalizing methods is that they can only be applied to two batches at a time. While cross-platform normalization steps can be chained together, the effect of these multiple normalization steps or which chaining method is still unclear [60].

2.3.2. Software and Websites Implementing Microarray Meta-Analysis and Cross-Platform Merging/Normalization

Software, including packages in R/Bioconductor and websites allowing users to implement microarray meta-analysis and cross-platform merging and normalization methods are listed in Table 1. Different experiments from multiple different arrays can be directly merged from the CEL files simultaneously using several packages implemented in R [69] including inSilico Merging [70], the CONOR [2], and virtualArray [71]. The inSilico Merging package implements XPN, DWD, and Combat, and the package CONOR additionally implements the GQ method. The virtualArray package allows cross-platform normalization using empirical Bayes methods (default) or the user may select one quantile discretization, normal discretization normalization, gene quantile normalization, median rank scores, quantile normalization, or mean centering [71]. This batch effect removal step can be supervised allowing the user to specify samples into groups based on platform as well as other attributes (e.g., cell type). Before the combined expression data undergoes cross-platform normalization, the data must be transformed to a common scale (e.g., log₂) and resolution (e.g., 12, 14, 16, or 20 bit) [71]. As with meta-analysis, low expression and low variance genes are typically filtered out.

Table 1. List of software and websites for performing microarray meta-analysis.

Microarray Meta-Analysis (Command Line Packages)		
Software Name	Language	Features
metaDE (metaOmics)	R	Implements 12 major meta-analysis methods [37]
MAMA	R	Implements combined effect size, combined <i>p</i> -values, combined ranks
metaMA	R	Implements combined moderated effect size, combined <i>p</i> -values
metaGEM	R	Implements combined effect size, combined <i>p</i> -values, vote counting [24]
metahdep	R	Effect size estimates particularly when hierarchical dependence is present
GeneMeta	R	Implements combined effect size [47]
OrderedList	R	Combine ranks with or without expression data
RankProd	R	Implements Product of Ranks method
RankAggreg	R	Aggregation of ordered lists based on the ranks using several different algorithms
Automated web applications for microarray meta-analysis/normalization		
Software Name	Features and URL	
INMEX	Meta-analysis. Support for 45 microarray platforms for human, mouse rat. Combines <i>p</i> -values, effect sizes, rank order, others http://www.inmex.ca/INMEX/	
Network Analyst	Meta-analysis. Combines <i>p</i> -values, effect sizes, rank order. Significantly altered genes are then presented within the context of protein-protein interaction networks. http://www.networkanalyst.ca/NetworkAnalyst/faces/home.xhtml	
A-MADMAN	Affymetrix platform normalization using quantile distribution transformation http://compgen.bio.unipd.it/bioinfo/amadman/	
MAAMD	Affymetrix meta-analysis http://www.biokepler.org/use_cases/maamd-workflow-standardize-meta-analyses-affymetrix-microarray-data	
Microarray cross-platform merging/normalization (command line packages)		
Software Name	Language	Features
mergeMaid	R	Implements Probability of Expression transformation (POE) [62]
metaArray	R	Implements POE [62]
CONOR	R	Implements XPN, Empirical Bayes (EB), Quantile normalization (QN), Quantile discretization (QD), others [2]
VirtualArray	R	Implements EB, QN, QD, others [71]
inSilico Merging Automated	R	Implements XPN, EB, DWD, others [23]
Microarray Data Analysis v2.13	R	Implements. Allows analysis of Illumina, Affymetrix and Agilent.
XPN	R	Implements Cross Platform Normalization [26]
DWD	JAVA, R MATLAB	Implements Distance Weighted Discrimination method [65]
Combat	R	Implements Empirical Bayes methods [64]
PLIDA	MATLAB	Normalizes an arbitrary number of platforms [60]
metAnalyzeAll	R	Elastic net classifier [42]

2.4. Comparison of Meta-Analysis vs. Cross-Platform Normalization

Directly-merged microarray data (or applying cross-platform normalization) has been argued to have better performance than meta-analysis for the identification of robust biomarkers on the premise that “deriving separate statistics and then averaging is often less powerful than directly computing statistics from aggregated data” [57]. In a comparative study, Taminau *et al.* [23] found significantly more differentially-expressed genes using cross-platform normalization than meta-analysis. An additional advantage of cross-platform normalization is that it allows prediction models applied to a subset of studies to be applied across additional studies from other platforms [27]. While cross-platform normalization has been applied in multiple studies [72–74], it has less frequently been used in the literature compared to meta-analysis [2]. A recent comprehensive systematic literature review of studies applying microarray integration methods found that only 27% of the studies directly merged microarray data and this subset of studies were mostly performed on the same platform [27].

One major limitation of existing cross-platform normalization is that they require that every treatment group or sample type be represented on each platform to allow differentiation of treatment effects from platform effects. Furthermore, cross-platform normalization methods do not guarantee elimination of laboratory or batch effects across experiments and Rung and Brazma [3] have argued that microarray meta-analysis provides better control of between-laboratory heterogeneity, which can be estimated using Cochrane’s Q statistic and be correspondingly adjusted.

3. Promising Transcriptomic Biomarkers Identified Using Meta-Analysis Approaches

Sweeney *et al.* [75] recently identified a transcriptomic signature to improve discrimination of patients with sepsis (infection) from those with sterile inflammation using blood samples. Their work analyzed publicly-available gene expression datasets from 22 independent cohorts (composed of 2903 microarrays in total) and applied a meta-analysis strategy implementing both effect size and *p*-values of differential gene expression. The investigators identified 82 genes differentially expressed between sepsis and inflammation and then performed a greedy forward search to determine which combination of these 82 genes produced the best improvement of area under the curve (AUC) in their discovery datasets. This resulted in an 11-gene transcriptional signature that was applied to 15 independent validation cohorts and was found to improve discrimination of patients with infection from those with sterile inflammation compared to use of clinical data alone. This gene signature requires further validation using prospective cohorts, however its excellent discriminatory power in both the discovery and validation cohorts suggests that it is likely to become a useful clinical assay in the future.

Santiago and Potashkin [76] implemented a transcriptomic and network-based meta-analysis in NetworkAnalyst (Table 1) to identify potential key hub genes in the blood of patients with Parkinson’s disease (PD). Their analysis identified hepatocyte nuclear factor 4 alpha (*HNF4A*) and polypyrimidine tract binding protein 1 (*PTBPI*), as the most significant up- and down-regulated genes in blood samples from PD patients. The relative abundance of *HNF4A* mRNA was found to correlate with disease severity in PD and the results were validated using samples obtained from two independent clinical trials. The abundance of *HNF4A* and *PTBPI* mRNAs significantly decreased and increased,

respectively, in PD patients during a 3-year follow-up period suggesting that these biomarkers may be useful for monitoring disease-modifying therapies for PD.

4. Confounding Adjustment

In the previously discussed cross-platform normalization approaches (Section 2.3), the major batch effect (the platform) is clearly identified (“supervised”), in distinction to other confounding adjustment methods such as using surrogate variable analysis (SVA) which detect “latent” (unknown) variables such as experimental variability or patient subgroups (e.g., breast cancer subtypes). It is important to account for possibly confounding (e.g., age or sex) or possibly predictive variables (e.g., smoking history) in addition to gene expression for building a gene signature. Additional categorical and continuous variables can be easily included along with the gene expression data using regression methods such as the elastic net penalty to fit a generalized linear model (GLM) [42]. These models can also be readily adapted for different outcomes such as categorical, continuous, and survival times. Cho *et al.* [77] developed a software program (rbsurv) to detect survival-associated genes based on the partial likelihood of the Cox model that allows adjustment of for risk factors in survival modeling.

Modelling confounding factors with variable selection in meta-analysis has recently been shown to improve robustness and sensitivity of DE gene detection [43] and inter-study concordance [78]. Chikina *et al.* [78] produced corrected differential expression lists using surrogate variables calculated using a modified version of SVA with improved inter-study agreement over uncorrected analysis. A two-class meta-analysis by Wang *et al.* [43] applied a random intercept model to account for confounding covariates in each single study analysis and combined *p*-values of the candidate biomarker list from each study using Fisher’s and maxP methods. Statistical approaches to allow synthesis of regression slopes in meta-analysis have been described [79] and applied to meta-analysis [51].

5. Conclusions

Gene signature discovery for prognostic and diagnostic purposes is improved with knowledgeable selection and appropriate application of integration methods on microarray data performed on multiple platforms. While no consensus for the best implementation of cross-platform integration is currently available, previous benchmarking and comparative analyses have established the strengths and limitations of many of the existing methods. The recent evidence suggesting improved performance of cross-platform normalization methods over meta-analysis may lead to an increasing proportion of studies in the literature implementing the former method. Further refinement of existing methods and development of new methods for cross-platform normalization and classification to exploit the vast quantity of microarray data currently available are expected. As elimination of platform specific bias becomes well-established with these methods, future studies addressing the performance of prognostic signature discovery in light of the existing *biological* heterogeneity will become a central focus.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Pepe, M.S.; Feng, Z. Improving biomarker identification with better designs and reporting. *Clin. Chem.* **2011**, *57*, 1093–1095.
2. Rudy, J.; Valafar, F. Empirical comparison of cross-platform normalization methods for gene expression data. *BMC Bioinform.* **2011**, *12*, doi:10.1186/1471-2105-12-467.
3. Rung, J.; Brazma, A. Reuse of public genome-wide gene expression data. *Nat. Rev. Genet.* **2013**, *14*, 89–99.
4. Gene Expression Omnibus. Secondary. Available online: <http://www.ncbi.nlm.nih.gov/geo/> (accessed on 5 May 2015).
5. Miller, M.B.; Tang, Y.W. Basic concepts of microarrays and potential applications in clinical microbiology. *Clin. Microbiol. Rev.* **2009**, *22*, 611–633.
6. Liu, C.G.; Calin, G.A.; Volinia, S.; Croce, C.M. MicroRNA expression profiling using microarrays. *Nat. Protoc.* **2008**, *3*, 563–578.
7. Hall, D.A.; Ptacek, J.; Snyder, M. Protein microarray technology. *Mech. Ageing Dev.* **2007**, *128*, 161–167.
8. Wu, S.; Xu, Y.; Feng, Z.; Yang, X.J.; Wang, X.G.; Gao, X. Multiple-platform data integration method with application to combined analysis of microarray and proteomic data. *BMC Bioinform.* **2012**, *13*, doi:10.1186/1471-2105-13-320.
9. Chang, C.; Wang, J.; Zhao, C.; Fostel, J.; Tong, W.; Bushel, P.R.; Deng, Y.; Pusztai, L.; Symmans, W.F.; Shi, T. Maximizing biomarker discovery by minimizing gene signatures. *BMC Genom.* **2011**, *12*, doi:10.1186/1471-2164-12-S5-S6.
10. McCollum, E.D.; Preidis, G.A.; Maliwichi, M.; Fostel, J.; Tong, W.; Bushel, P.R.; Deng, Y.; Pusztai, L.; Symmans, W.F.; Shi, T. Clinical *versus* rapid molecular HIV diagnosis in hospitalized African infants: A randomized controlled trial simulating point-of-care infant testing. *J. Acquir. Immune. Defic. Syndr.* **2014**, *66*, e23–e30.
11. Park, S.; Zhang, Y.; Lin, S.; Wang, T.H.; Yang, S. Advances in microfluidic PCR for point-of-care infectious disease diagnostics. *Biotechnol. Adv.* **2011**, *29*, 830–839.
12. Director's Challenge Consortium for the Molecular Classification of Lung A; Shedden, K.; Taylor, J.M.; Enkemann, S.A.; Tsao, M.S.; Yeatman, T.J.; Gerald, W.L.; Eschrich, S.; Jurisica, I.; Giordano, T.J.; *et al.* Gene expression-based survival prediction in lung adenocarcinoma: A multi-site, blinded validation study. *Nat. Med.* **2008**, *14*, 822–827.
13. Van Laar, R.; Flinchum, R.; Brown, N.; Ramsey, J.; Riccitelli, S.; Heuck, C.; Barlogie, B.; Shaughnessy, J.D., Jr. Translating a gene expression signature for multiple myeloma prognosis into a robust high-throughput assay for clinical use. *BMC Med. Genom.* **2014**, *7*, doi:10.1186/1755-8794-7-25.
14. Gesthalter, Y.B.; Vick, J.; Steiling, K.; Spira, A. Translating the transcriptome into tools for the early detection and prevention of lung cancer. *Thorax* **2015**, *70*, 476–481.
15. Shen, R.; Chinnaiyan, A.M.; Ghosh, D. Pathway analysis reveals functional convergence of gene expression profiles in breast cancer. *BMC Med. Genom.* **2008**, *1*, doi:10.1186/1755-8794-1-28.

16. Shi, L.; Campbell, G.; Jones, W.D.; Campagne, F.; Wen, Z.; Walker, S.J.; Su, Z.; Chu, T.M.; Goodsaid, F.M.; Pusztai, L.; *et al.* The MicroArray Quality Control (MAQC)-II study of common practices for the development and validation of microarray-based predictive models. *Nat. Biotechnol.* **2010**, *28*, 827–838.
17. Simon, R. Genomic biomarkers in predictive medicine: An interim analysis. *EMBO Mol. Med.* **2011**, *3*, 429–435.
18. Diamandis, E.P. Cancer biomarkers: can we turn recent failures into success? *J. Natl. Cancer Inst.* **2010**, *102*, 1462–1467.
19. Baker, S.G. Improving the biomarker pipeline to develop and evaluate cancer screening tests. *J. Natl. Cancer Inst.* **2009**, *101*, 1116–1169.
20. Cruz, J.; Wishart, D. Applications of machine learning in cancer prediction and prognosis. *Cancer Inf.* **2006**, *2*, 59–77.
21. Michiels, S.; Koscielny, S.; Hill, C. Prediction of cancer outcome with microarrays: A multiple random validation strategy. *Lancet* **2005**, *365*, 488–492.
22. Hamid, J.S.; Hu, P.; Roslin, N.M.; Ling, V.; Greenwood, C.T.; Beyene, J. Data integration in genetics and genomics: Methods and challenges. *Hum. Genom. Proteom.* **2009**, *2009*, doi:10.4061/2009/869093.
23. Taminau, J.; Lazar, C.; Meganck, S.; Nowé, A. Comparison of merging and meta-analysis as alternative approaches for integrative gene expression analysis. *ISRN Bioinform.* **2014**, *2014*, doi:10.1155/2014/345106.
24. Ramasamy, A.; Mondry, A.; Holmes, C.C.; Altman, D.G. Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med.* **2008**, *5*, e184.
25. Hu, P.; Greenwood, C.M.; Beyene, J. Integrative analysis of multiple gene expression profiles with quality-adjusted effect size models. *BMC Bioinform.* **2005**, *6*, doi:10.1186/1471-2105-6-128.
26. Shabalín, A.A.; Tjelmeland, H.; Fan, C.; Perou, C.M.; Nobel, A.B. Merging two gene-expression studies via cross-platform normalization. *Bioinformatics* **2008**, *24*, 1154–1160.
27. Tseng, G.C.; Ghosh, D.; Feingold, E. Comprehensive literature review and statistical considerations for microarray meta-analysis. *Nucleic Acids Res.* **2012**, *40*, 3785–3799.
28. Hu, P.; Wang, X.; Haitsma, J.J.; Furmli, S.; Masoom, H.; Liu, M.; Imai, Y.; Slutsky, A.S.; Beyene, J.; Greenwood, C.M.; *et al.* Microarray meta-analysis identifies acute lung injury biomarkers in donor lungs that predict development of primary graft failure in recipients. *PLoS ONE* **2012**, *7*, e45506.
29. Perez-Diez, A.; Morgun, A.; Shulzhenko, N. Microarrays for cancer diagnosis and classification. *Adv. Exp. Med. Biol.* **2007**, *593*, 74–85.
30. Xia, J.; Gill, E.E.; Hancock, R.E. NetworkAnalyst for statistical, visual and network-based meta-analysis of gene expression data. *Nat. Protoc.* **2015**, *10*, 823–844.
31. Kitchen, R.R.; Sabine, V.S.; Simen, A.A.; Dixon, J.M.; Bartlett, J.M.; Sims, A.H. Relative impact of key sources of systematic noise in Affymetrix and Illumina gene-expression microarray experiments. *BMC Genom.* **2011**, *12*, doi:10.1186/1471-2164-12-589.

32. Turnbull, A.K.; Kitchen, R.R.; Larionov, A.A.; Renshaw, L.; Dixon, J.M.; Sims, A.H. Direct integration of intensity-level data from Affymetrix and Illumina microarrays improves statistical power for robust reanalysis. *BMC Med. Genom.* **2012**, *5*, doi:10.1186/1755-8794-5-35.
33. Chang, L.C.; Lin, H.M.; Sibille, E.; Tseng, G.C. Meta-analysis methods for combining multiple expression profiles: Comparisons, statistical characterization and an application guideline. *BMC Bioinform.* **2013**, *14*, doi:10.1186/1471-2105-14-368.
34. Wilson, C.L.; Miller, C.J. Simpleaffy: A BioConductor package for Affymetrix Quality Control and data analysis. *Bioinformatics* **2005**, *21*, 3683–3685.
35. Bolstad, B. affyPLM: Model Based QC Assessment of Affymetrix GeneChips. Available online: <http://www.cse.unsw.edu.au/~mike/myrlibrary.old/affyPLM/doc/QualityAssess.pdf> (accessed on 16 April 2015).
36. Kang, D.D.; Sibille, E.; Kaminski, N.; Tseng, G.C. MetaQC: Objective quality control and inclusion/exclusion criteria for genomic meta-analysis. *Nucleic Acids Res.* **2012**, *40*, doi:10.1093/nar/gkr1071.
37. Wang, X.; Kang, D.D.; Shen, K.; Song C; Lu, S.; Chang, L.C.; Liao, S.G.; Huo, Z.; Tang, S.; Ding, Y.; *et al.* An R package suite for microarray meta-analysis in quality control, differentially expressed gene analysis and pathway enrichment detection. *Bioinformatics* **2012**, *28*, 2534–2536.
38. Gentleman, R.C.; Carey, V.J.; Bates, D.M.; Bolstad, B; Dettling, M.; Dudoit, S.; Ellis, B.; Gautier, L.; Ge, Y.; Gentry, J.; *et al.* Bioconductor: Open software development for computational biology and bioinformatics. *Genome Biol.* **2004**, *5*, doi:10.1186/gb-2004-5-10-r80.
39. Dai, M.; Wang, P.; Boyd, A.D.; Kostov, G.; Athey, B.; Jones, E.G.; Bunney, W.E.; Myers, R.M.; Speed, T.P.; Akil, H.; *et al.* Evolving gene/transcript definitions significantly alter the interpretation of GeneChip data. *Nucleic Acids Res.* **2005**, *33*, doi:10.1093/nar/gni179.
40. Barbosa-Morais, N.L.; Dunning, M.J.; Samarajiwa, S.A.; Darot, J.F.; Ritchie, M.E.; Lynch, A.G.; Tavaré, S. A re-annotation pipeline for Illumina BeadArrays: Improving the interpretation of gene expression data. *Nucleic Acids Res.* **2010**, *38*, doi:10.1093/nar/gkp942.
41. Konstantinopoulos, P.A.; Cannistra, S.A.; Fountzilas, H.; Culhane, A.; Pillay, K.; Rueda, B.; Cramer, D.; Seiden, M.; Birrer, M.; Coukos, G.; Zhang, L.; *et al.* Integrated analysis of multiple microarray datasets identifies a reproducible survival predictor in ovarian cancer. *PLoS ONE* **2011**, *6*, e18202, doi:10.1371/journal.pone.0018202
42. Hughey, J.J.; Butte, A.J. Robust meta-analysis of gene expression using the elastic net. *Nucleic Acids Res.* **2015**, doi:10.1093/nar/gkv229.
43. Wang, X.; Lin, Y.; Song, C.; Culhane A; Pillay K; Rueda, B.; Cramer, D.; Seiden, M.; Birrer, M.; Coukos, G.; *et al.* Detecting disease-associated genes with confounding variable adjustment and the impact on genomic meta-analysis: With application to major depressive disorder. *BMC Bioinform.* **2012**, *13*, doi:10.1186/1471-2105-13-52.
44. Sabine, V.S.; Sims, A.H.; Macaskill, E.J.; Renshaw, L.; Thomas, J.S.; Dixon, J.M.; Bartlett, J.M. Gene expression profiling of response to mTOR inhibitor everolimus in pre-operatively treated post-menopausal women with oestrogen receptor-positive breast cancer. *Breast Cancer Res. Treat.* **2010**, *122*, 419–428.
45. Hu, P.; Beyene, J.; Greenwood, C.M. Tests for differential gene expression using weights in oligonucleotide microarray experiments. *BMC Genom.* **2006**, *7*, doi:10.1186/1471-2164-7-33.

46. Rhodes, D.R.; Yu, J.; Shanker, K.; Varambally, R.; Ghosh, D.; Barrette, T.; Pandey, A.; Chinnaiyan, A.M. Large-scale meta-analysis of cancer microarray data identifies common transcriptional profiles of neoplastic transformation and progression. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 9309–9314.
47. Choi, J.K.; Yu, U.; Kim, S.; Yoo, O.J. Combining multiple microarray studies and modeling interstudy variation. *Bioinformatics* **2003**, *19*, i84–i90.
48. Hong, F.; Breitling, R.; McEntee, C.W.; Wittner, B.S.; Nemhauser, J.L.; Chory, J. RankProd: A bioconductor package for detecting differentially expressed genes in meta-analysis. *Bioinformatics* **2006**, *22*, 2825–2827.
49. Rhodes, D.R.; Barrette, T.; Rubin, M.A.; Ghosh, D.; Chinnaiyan, A.M.; *et al.* Meta-analysis of microarrays: Interstudy validation of gene expression profiles reveals pathway dysregulation in prostate cancer. *Cancer Res.* **2002**, *62*, 4427–4433
50. Song, C.; Tseng, G.C. Hypothesis setting and order statistic for robust genomic meta-analysis. *Ann. Appl. Stat.* **2014**, *8*, 777–800.
51. Ma, S.; Huang, J. Regularized gene selection in cancer microarray meta-analysis. *BMC Bioinform.* **2009**, *10*, doi:10.1186/1471-2105-10-1.
52. Lu, S.; Li, J.; Song, C; Tseng, G.C. Biomarker detection in the integration of multiple multi-class genomic studies. *Bioinformatics* **2010**, *26*, 333–340.
53. Campaign, A.; Yang, Y.H. Comparison study of microarray meta-analysis methods. *BMC Bioinform.* **2010**, *11*, doi:10.1186/1471-2105-11-408.
54. Hong, F.; Breitling, R.A comparison of meta-analysis methods for detecting differentially expressed genes in microarray experiments. *Bioinformatics* **2008**, *24*, 374–382.
55. Li, Z.; Herold, T.; He, C.; Valk, P.J.; Chen, P.; Jurinovic, V.; Mansmann, U.; Radmacher, M.D.; Maharry, K.S.; Sun, M.; *et al.* Identification of a 24-gene prognostic signature that improves the European LeukemiaNet risk classification of acute myeloid leukemia: An international collaborative study. *J. Clin. Oncol.* **2013**, *31*, 1172–1181.
56. Sims, A.H.; Smethurst, G.J.; Hey, Y.; Okoniewski, M.J.; Pepper, S.D.; Howell, A.; Miller, C.J.; Clarke, R.B. The removal of multiplicative, systematic bias allows integration of breast cancer gene expression datasets—Improving meta-analysis and prediction of prognosis. *BMC Med. Genom.* **2008**, *1*, doi:10.1186/1755-8794-1-42.
57. Xu, L.; Tan, A.C.; Winslow, R.L.; Geman, D. Merging microarray data from separate breast cancer studies provides a robust prognostic test. *BMC Bioinformatics* **2008**, *9*, doi:10.1186/1471-2105-9-125.
58. Liu, C.C.; Hu, J.; Kalakrishnan, M.; Huang, H.; Zhou, X.J. Integrative disease classification based on cross-platform microarray data. *BMC Bioinform.* **2009**, *10*, doi:10.1186/1471-2105-10-S1-S25.
59. Lee, Y.; Scheck, A.C.; Cloughesy, T.F.; Lai, A.; Dong, J.; Farooqi, H.K.; Liau, L.M.; Horvath, S.; Mischel, P.S.; Nelson, S.F. Gene expression analysis of glioblastomas identifies the major molecular basis for the prognostic benefit of younger age. *BMC Med. Genom.* **2008**, *1*, doi:10.1186/1755-8794-1-52.
60. Deshwar, A.G.; Morris, Q. PLIDA: Cross-platform gene expression normalization using perturbed topic models. *Bioinformatics* **2014**, *30*, 956–961.

61. Jiang, H.; Deng, Y.; Chen, H.S.; Tao, L.; Sha, Q.; Chen, J.; Tsai, C.J.; Zhang, S. Joint analysis of two microarray gene-expression data sets to select lung adenocarcinoma marker genes. *BMC Bioinform.* **2004**, *5*, doi:10.1186/1471-2105-5-81.
62. Shen, R.; Ghosh, D.; Chinnaiyan, A.M. Prognostic meta-signature of breast cancer developed by two-stage mixture modeling of microarray data. *BMC Genom.* **2004**, *5*, doi:10.1186/1471-2164-5-94.
63. Parmagiani, G.; Garret-Mayer, E.S.; Anbazhagan, R.; Gabrielson, E. A cross-study comparison of gene expression studies for the molecular classification of lung cancer. *Clin. Cancer Res.* **2004**, *10*, 2922–2927.
64. Johnson, W.E.; Li, C.; Rabinovic, A. Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* **2007**, *8*, 118–127.
65. Huang, H.; Lu, X.; Liu, Y.; Haaland, P.; Marron, J.S. R/DWD: Distance-weighted discrimination for classification, visualization and batch adjustment. *Bioinformatics* **2012**, *28*, 1182–1183.
66. WebArray, D.B. Secondary. Available online: <http://www.webarraydb.org/webarray/index.html>. (accessed on 12 May 2015).
67. Benito, M.; Parker, J.; Du, Q.; Wu, J.; Xiang, D.; Perou, C.M.; Marron, J.S. Adjustment of systematic microarray data biases. *Bioinformatics* **2003**, *20*, 105–114.
68. Chen, C.; Grennan, K.; Badner, J.; Zhang, D.; Gershon, E.; Jin, L.; Liu, C. Removing batch effects in analysis of expression microarray data: An evaluation of six batch adjustment methods. *PLoS ONE* **2011**, *6*, e17238.
69. R, Core Team. *A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2012.
70. Taminau, J.; Meganck, S.; Lazar, C.; Steenhoff, D.; Coletta, A.; Molter, C.; Duque, R.; de Schaetzen, V.; Weiss Solís, D.Y.; Bersini, H.; *et al.* Unlocking the potential of publicly available microarray data using inSilicoDb and inSilicoMerging R/Bioconductor packages. *BMC Bioinform.* **2012**, *13*, doi:10.1186/1471-2105-13-335.
71. Heider, A.; Alt, R. virtualArray: A R/bioconductor package to merge raw data from different microarray platforms. *BMC Bioinform.* **2013**, *14*, 75, doi:10.1186/1471-2105-14-75.
72. Warnat, P.; Eils, R.; Brors, B. Cross-platform analysis of cancer microarray data improves gene expression based classification of phenotypes. *BMC Bioinform.* **2005**, *6*, doi:10.1186/1471-2105-6-265.
73. Fielden, M.R.; Nie, A.; McMillian, M.; Yi, Y.; Morrison, C.; Yang, P.; Sun, Z.; Szoke, J.; Gerald, W.L.; Watson, M.; *et al.* Interlaboratory evaluation of genomic signatures for predicting carcinogenicity in the rat. *Toxicol. Sci.* **2008**, *103*, 28–34.
74. Lu, Y.; Lemon, W.; Liu, P.Y.; Yi, Y.; Morrison, C.; Yang, P.; Sun, Z.; Szoke, J.; Gerald, W.L.; Watson M; *et al.* A gene expression signature predicts survival of patients with stage I non-small cell lung cancer. *PLoS Med.* **2006**, *3*, e467.
75. Sweeney, T.E.; Shidham, A.; Wong, H.R.; Khatri, P. A comprehensive time-course-based multicohort analysis of sepsis and sterile inflammation reveals a robust diagnostic gene set. *Sci. Transl. Med.* **2015**, *7*, doi:10.1126/scitranslmed.aaa5993.
76. Santiago, J.A.; Potashkin, J.A. Network-based metaanalysis identifies HNF4A and PTBP1 as longitudinally dynamic biomarkers for Parkinson’s disease. *Proc. Natl. Acad. Sci USA* **2015**, *112*, 2257–2262.

77. Cho, H.; Yu, A.; Kim, S.; Kang, J.; Hong, S.M. Robust likelihood-based survival modeling with microarray data. *J. Stat. Softw.* **2009**, *29*, 1–16
78. Chikina, M.D.; Sealfon, S.C. Increasing consistency of disease biomarker prediction across datasets. *PLoS ONE* **2014**, *9*, e91272.
79. Becker, B.J.; Wu, M.-J. The Synthesis of Regression Slopes in Meta-Analysis. *Stat. Sci.* **2007**, *22*, 414–429.

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).