

# Sample Selection for Microarray Gene Expression Studies

D. Repsilber<sup>1</sup>, L. Fink<sup>2</sup>, M. Jacobsen<sup>3</sup>, O. Bläsing<sup>4</sup>, A. Ziegler<sup>1</sup>

<sup>1</sup>Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Lübeck, Germany

<sup>2</sup>Institut für Pathologie, Justus-Liebig-Universität Gießen, Gießen, Germany

<sup>3</sup>Max-Planck-Institut für Infektionsbiologie, Berlin, Germany

<sup>4</sup>Max-Planck-Institut für Molekulare Pflanzenphysiologie, Potsdam, Germany

## Summary

**Objectives:** The choice of biomedical samples for microarray gene expression studies is decisive for both validity and interpretability of results. We present a consistent, comprehensive framework to deal with the typical selection problems in microarray studies.

**Methods:** Microarray studies are designed either as case-control studies or as comparisons of parallel groups from cohort studies, since high levels of random variation in the experimental approach thwart absolute measurements of gene expression levels. Validity and results of gene expression studies heavily rely on the appropriate choice of these study groups. Therefore, the so-called principles of comparability, which are well known from both clinical and epidemiological studies, need to be applied to microarray experiments.

**Results:** The principles of comparability are the study-base principle, the principle of deconfounding and the principle of comparable accuracy in measurements. We explain each of these principles, show how they apply to microarray experiments, and illustrate them with examples. The examples are chosen as to represent typical stumbling blocks of microarray experimental design, and to exemplify the benefits of implementing the principles of comparability in the setting of microarray experiments.

**Conclusions:** Microarray studies are closely related to classical study designs and therefore have to obey the same principles of comparability as these. Their validity should not be compromised by selection, confounding or information bias. The so-called study-base principle, calling for comparability and thorough definition of the compared cell populations, is the key principle for the choice of biomedical samples and controls in microarray studies.

## Keywords

Confounding, gene expression profiling, biometry, research design, selection bias, microarray studies

Methods Inf Med 2005; 44: 461–7

## Introduction

How is gene expression altered in tumor tissue as compared to normal tissue? The design of a microarray gene expression study to answer such a typical question seems straightforward. Tissue samples should be taken from tumor tissue and from healthy tissue, RNAs extracted and expression levels compared with microarrays to detect candidates for differential gene expression. However, where should the healthy tissue be sampled? Depending on our choice, the results of this study may or may not answer our initial question. For example, healthy tissue from the direct neighborhood of the tumor is probably the wrong choice since the metabolic or blood flow situation may be altered. Thus, measured differences in gene expression may not reflect the typical changes in tumor tissue as compared to “normal” healthy tissue. On the other hand, samples from remote parts of the tissue may also cause interpretation problems, as gene expression in these remote parts may not be typical for the tissue region where the tumor was found, and regulated differentially for other reasons. The sampling strategy thus directly determines the interpretability of the experiment [1].

This example points to a general design problem, which is well known from clinical and epidemiological studies. It is the problem of choosing appropriate *controls* for comparison with the given *cases* in case-control studies [2, 3]. More generally, originating from biostatistics of epidemiological studies, there exist the so-called *principles of comparability* (for an overview see e.g. [3]), which offer a consistent framework for selection of controls.

Microarray gene expression studies are also generally *comparative* in their design since high levels of random variation in the experimental approach thwart absolute measurements of gene expression levels. Moreover, sampling for microarray gene expression studies most often relies on the choice of “cases” together with an appropriate “control”. In this context, the latter is often called “reference”. Therefore, sampling recommendations for conducting clinical comparisons as derived from epidemiological principles also apply to microarray gene expression studies.

In our contribution, we survey the implementation of design principles from comparisons in epidemiological and clinical studies for designing and conducting microarray gene expression studies and illustrate these with typical examples. The three design principles in focus are (compare [3]):

- 1) the study-base principle, which includes the avoidance of ascertainment bias, and/or selection bias in gene expression studies,
- 2) deconfounding, and
- 3) comparable accuracy for measurements in study subjects, e.g. tissues.

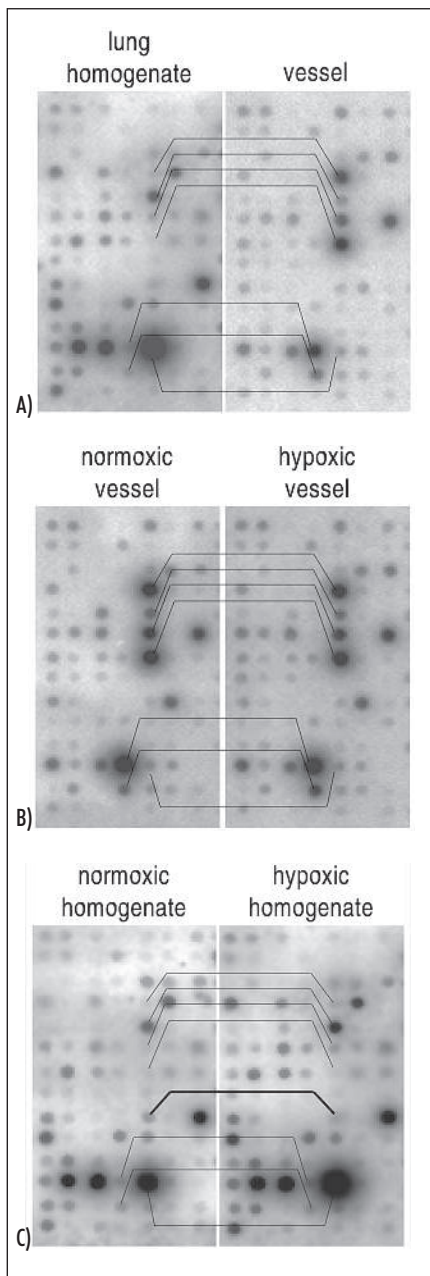
The study-base principle, also termed principle of homogeneity, appears to be the key principle for microarray experiments. It is illustrated in the following section using two typical examples for the problem of selection bias on different levels of biological organization. The principle of deconfounding is exemplified in the following section using two examples. The next section then discusses how gene expression studies should be conducted in order to achieve comparable measurement accuracies in cases and controls. Finally, conclusions are drawn

about applying epidemiological design principles for sample selection to microarray gene expression experiments.

## The Study-base Principle: Homogeneity of Samples and References

Clinical and epidemiological case-control studies aim at revealing possible correlations between certain risk factors on the one hand and the disease of interest on the other. The means of investigation is to seek significantly higher abundances of the risk factors in question in the group of affected individuals as compared to the group of healthy controls. In the context of microarray gene expression studies, affected individuals, e.g. with a specific tumor, are compared with healthy controls with respect to their gene expression levels. In most of these applications, researchers speak of tumor tissue samples and healthy tissue samples. Since the aim is to find those gene expression levels which consistently differ between the sample of interest and the chosen control, significantly differentially expressed genes are the analogue of significant risk factors from epidemiological studies.

Within the epidemiological context, the study-base is defined as the population from which both affected and healthy individuals are taken. Here, the *study-base* principle seeks to avoid a selection bias by thoroughly defining the study-base population as well as the traits distinguishing affected from healthy individuals. The simplest way to satisfy this principle is to draw a population-based random sample of controls from the same source as the cases [3]. However, cases also need to be chosen carefully in order to avoid systematic differences in characteristics between those who are selected for study and those who are not. These potential and important sources of bias directly apply to microarray gene expression studies by demanding that the reason for differential investigated gene expression should be *the only* difference between cells



**Fig. 1** Differential gene expression as measured on macroarrays for hypoxic/normoxic mice. A) Tissue-specific gene expression in homogenate of total lung preparations versus micro-dissected lung vessels. B) Gene expression in dissected vessels under normoxic conditions as compared to hypoxic conditions. Small distinct changes for few genes on a background of otherwise comparable expression levels. C) Gene expression in lung homogenate under normoxic conditions as compared to hypoxic conditions. The bold lines point to a gene which seems to be upregulated under normoxic conditions, which is not the case for the comparison of gene expression in dissected vessels (compare B). Reprinted from *Am J Pathol* 2002; 160: 81-90 with permission from the American Society for Investigative Pathology.

chosen as sample and those chosen as control.

In the following we give two examples of how the *study-base principle* may be violated in typical microarray gene expression studies, surveying the consequences for the interpretability of study results as well as presenting possible solutions for the respective study designs.

## Differential Gene Expression in Lungs from Mice under Hypoxic Conditions

We [4] have studied gene expression in both total lung homogenates and dissected lung arteries from mice under hypoxic conditions as compared to mice under normal  $O_2$  partial pressure. In this case our focus is on the choice of the tissue chosen for comparison and how this choice affects results and their interpretation.

Figure 1A shows hybridized membrane filters opposing gene expression from total lung homogenate and lung arteries. Apparently, differences in gene expression span almost the whole set of genes probed for this experiment. The figure clearly shows that lung tissue is a *heterogeneous* tissue. Therefore, a gene expression study built upon this tissue material is prone to violate the study-base principle mainly for two reasons: 1) It is unclear whether the composition of such a heterogeneous tissue remains constant for the two conditions under comparison. Any uncontrolled changes, however, would obscure the interpretation of observed changes of RNA levels as changes in gene expression. 2) It is unclear which cell population should be declared responsible for observed changes in gene expression, thus interpretability is thwarted again. As a consequence, it is clearly questionable if the use of total lung homogenates leads to valid, thus interpretable results for the intended comparison of normoxic and hypoxic mice.

In this case, a specific homogeneous tissue element of primary interest should be used instead. Figure 1B shows the results for the comparison of gene expression in dissected lung arteries from normoxic mice as opposed to hypoxic mice. The expression

profiles are extremely homogeneous between the compared conditions. As a result, interpretability of results is achieved, as the observed changes in gene expression nicely match the changes in artery morphology as observed with the microscope [4]. In addition, as illustrated in Figure 1C, comparisons of normoxic with hypoxic conditions lead to different results if either vessel tissues or total lung homogenates are compared.

In conclusion, and again from the epidemiological study point of view, an improvement in the definition of the study base resulted in both avoiding the uncertainty of uncontrollable changes in the composition of a heterogeneous tissue as well as assuring the validity, thus interpretability of the study results.

## Gene Expression during the Cell Cycle

Our second example illustrating the importance of the study-base principle exemplifies not only the necessity to strictly define and choose the populations of cells for which gene expressions should be compared. It also illustrates selection biases introduced by technical problems of the selection method.

Cooper and Shedden [5] review gene expression experiments aimed at identifying genes which are differentially regulated during the cell cycle in yeast. The experimental design in this case intends to compare gene expression levels from different stages of the cell cycle. In this case, it is not the definition of the study base which causes the problems. The cell population of interest is very well defined as yeast cells under certain conditions during cell culturing. It is rather the selection of the cases for comparison that causes serious obstacles.

Each individual single cell within the cell culture population is in another state of the cell cycle. A cross-sectional study would therefore result in assessing a mean expression value for all genes under study and, hence, impede the observation of cell cycle dynamics in regulation of gene expression. Therefore, both a measure and a method for

cell cycle *synchronization* are needed. To be specific, cells under investigation have to be cell-cycle-aligned and are expected to pass through the cell cycle as a unified and coherent cohort.

Firstly, it is difficult to determine whether cells are really synchronized or not. Cooper and Shedden [5] give a detailed list of criteria for determining the status of synchronization in cell cultures. Cells can be viewed as being cell-cycle-aligned and sensible experimental time series become possible only if their main criteria are fulfilled.

The second problem, however, is the decisive one. Which method should be chosen to synchronize cells? There are two groups of methods frequently applied, the so-called *whole-culture methods* and the *selective synchronization methods*. Both seem to arrive at synchronized populations of yeast cells. They differ, however, significantly in their eligibility for a sensible gene expression experiment.

A popular example for a whole-culture method is cell-starving. Starved cells are arrested in a specific phase of the cell cycle. Supplying standard medium to the starved cell culture allows for normal cell growth again, and is thought to be the start of a cell-cycle-aligned culture. The drawbacks are overwhelming, and especially illustrative as a violation of the study-base principle. Firstly, the method itself represents a rather severe intervention, so that “normal” gene regulatory rhythms are not likely to be observed on a background of repair and rescue programs of the arrested cells. Hence, the selection method imposes a severe bias with respect to the observed gene expression levels. Secondly, after having released cells from arrest they are far from being synchronized as defined before. Depending on the stage of the cell cycle where the arrest methods were applied, returning to “normal” rhythms is rather different for different individual cells. This obscures detection of cell-cycle-regulated genes. Again, selection introduces severe biases, and possible results would not be interpretable with respect to the target question of cell-cycle-regulated genes.

The alternative group of methods – selective synchronization – does better for both points of criticism regarding the whole

culture methods. Here, cells of a specific phase are selected for example by their specific size or DNA content. The selection itself is not disturbing the cells in their physiology as severe as the arrest methods. Moreover, cell cycle phase-selected cells could be shown to exhibit coherent cell cycles for at least three rounds.

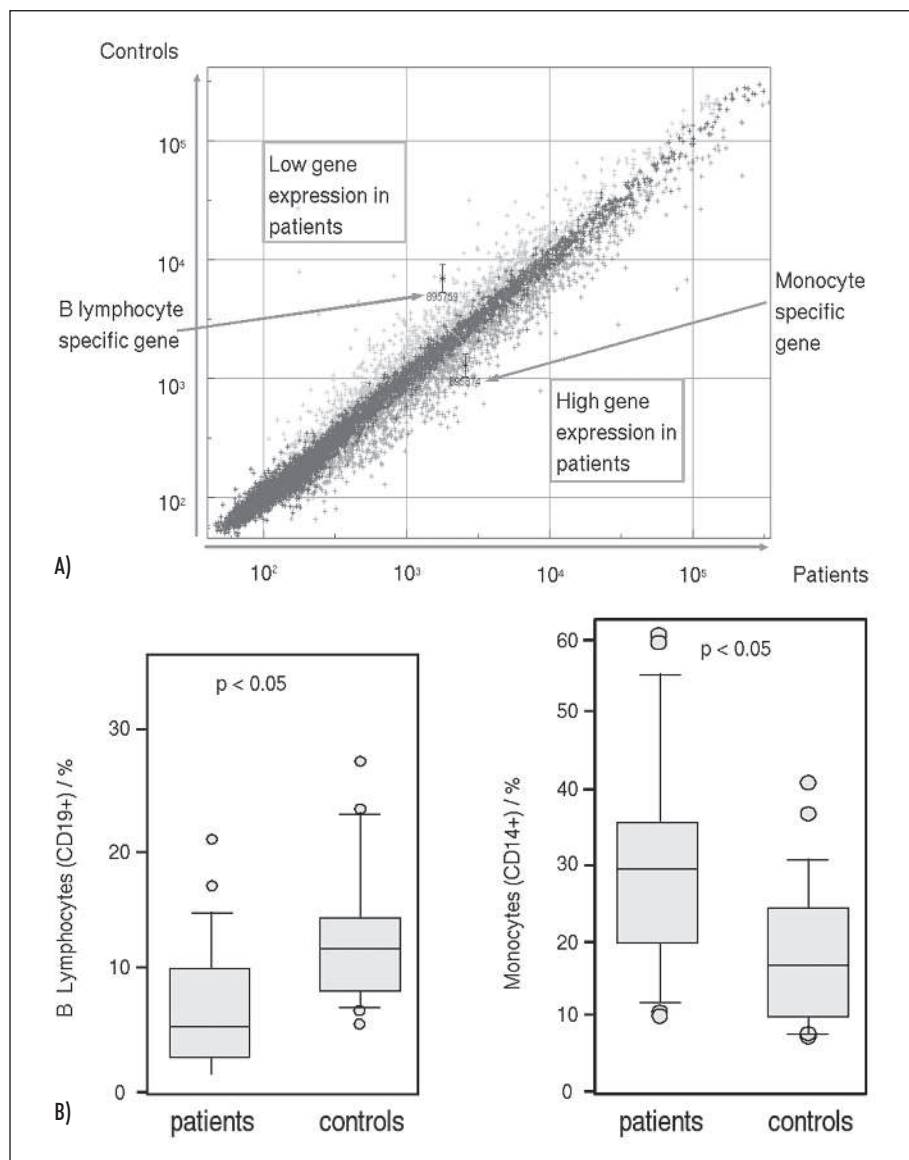
In summary, this example illustrates the typical problems of comparability. Here, the study population is well-defined. However, the selection of cases and controls, i.e. non-cases, poses severe problems. Part of the solution of the dilemma is sticking to a thorough definition of the cases. Here, a cell-cycle-aligned population of cells in culture is better defined, which helps in avoiding selection biases.

## Deconfounding Samples

Confounding [6] may occur in gene expression studies in a similar way as in epidemiological studies. For example, gene expression levels might have an effect on the outcome. Differences in gene expression levels between samples may, however, be distorted by other factors that influence the outcome. In principle, there are two approaches to overcome confounding. The quantitative impact of the confounding factor on the study results cannot exactly be measured. In this case, complete assurance of control of confounding can be achieved by eliminating the variability in the confounding factor over cases and controls [3]. The more convincing approach, however, is to correct for confounding effects in the analyses if the confounder can be measured. In the following we give an example for each of these possible cases.

## Differential Gene Expression in Blood as a Heterogeneous Tissue

The following example again deals with the problem of measuring gene expression in heterogeneous tissues. However, focusing on a well-defined subpopulation by microdissection, as for the example above, dealing with dissected lung arteries from mice,



**Fig. 2** Gene expression in tuberculosis patients (x-axis) as compared to healthy controls (y-axis) in peripheral blood mononuclear cells (PBMCs). Among genes that appear to be down-regulated (light-grey, above line of identity) in patients are B-cell specific genes, among genes appearing to be upregulated (dark-grey, below line of identity) are monocyte specific genes. One specific gene in each group is marked by a whisker-plot. However, B-cells are decreased in their population in patients and monocytes increased, such that it remains unclear if results point to regulation of gene expression in PBMCs or to regulation of PBMC population composition [unpublished data].

is obscured this time, and, hence, another work-around is necessary.

We [unpublished data] have searched for candidate genes differentially expressed in peripheral blood mononuclear cells (PBMCs) in tuberculosis patients as compared to healthy contacts. Tissue heterogeneity in this case is reflected as different ratios of the specific cell types in the assessed population of mononuclear cells under in-

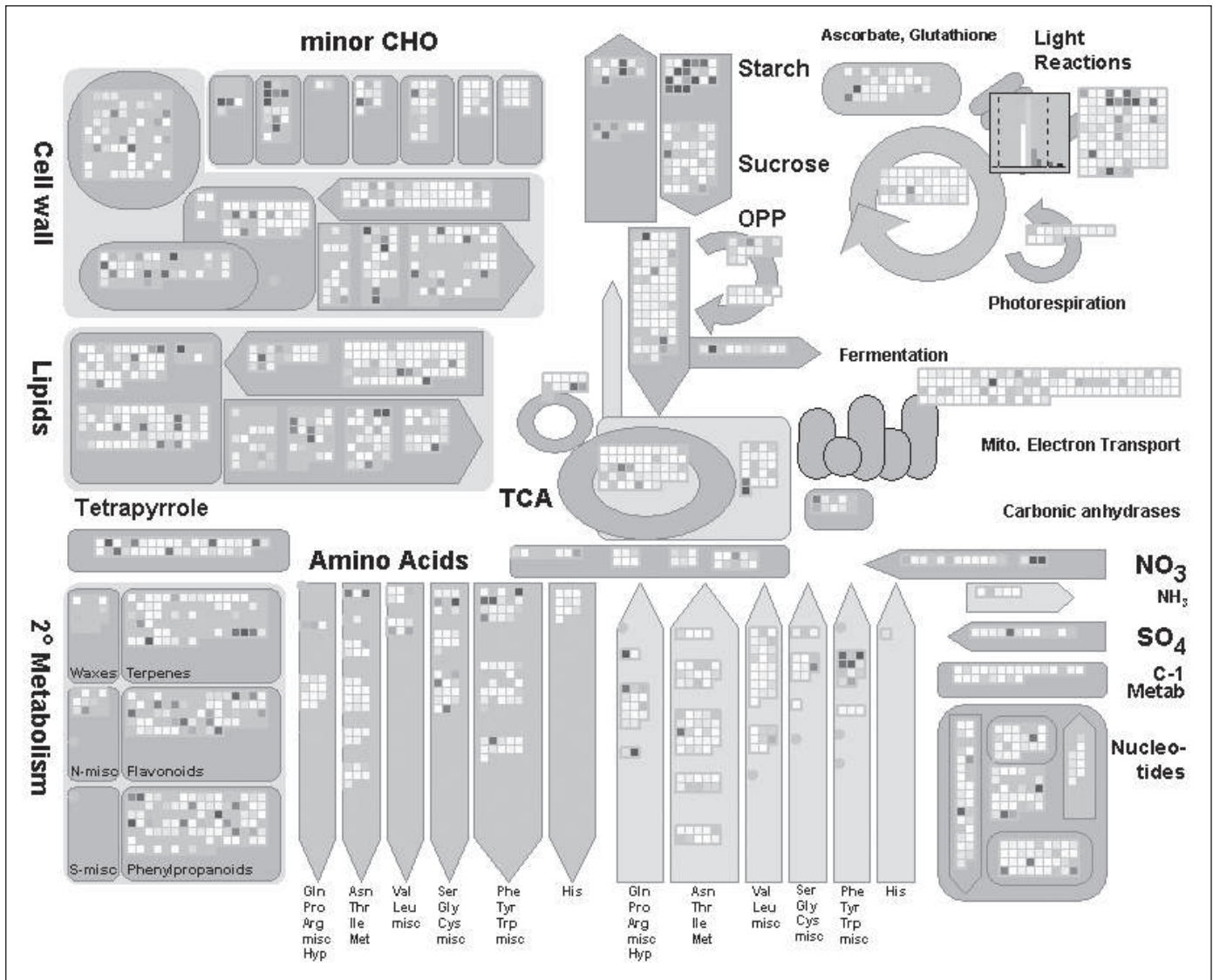
vestigation. Figure 2 illustrates the confounding effect. For all genes investigated, Figure 2A shows the expression levels in both tuberculosis patients (x-axis) and healthy contacts (y-axis) as mean values over the nine patients from each study group. Genes clearly above the line of identity would be thought to be down-regulated in patients as compared to the healthy controls. Conversely, those genes

which are clearly represented below the line of identity would be thought to be up-regulated in patients as compared to controls.

However, we also *measured* the composition of several immune cell populations which are part of the assessed PBMCs. Figure 2B displays the proportions of B-cells and monocytes in patients and controls, respectively, as measured in *fluorescence aided cell sorting* (FACS) analysis. The proportion of B-cells is clearly decreased in patients, while the proportion of monocytes is increased. Therefore, it is unclear if the observed difference in mRNA levels between patients and controls is due to the shift of B-cell and monocyte proportion or to a regulation of gene expression on the single cell level.

We have solved this problem by adjusting for cell proportions in the gene expression analyses. In detail, we have applied a variance analytic approach and modeled the cell proportion, as well as interaction effects. Upon use of this method, we have observed a difference in gene expression in the monocyte-specific gene, while no difference has been detected in the B-cell-specific gene. We were able to confirm our findings by protein expression analyses (results not shown).

In summary, we have shown in this example that tissue heterogeneity may cause confounding in microarray gene expression experiments. We want to stress that it may be extremely complicated to detect tissue heterogeneity. And even if it has been observed, it may be difficult – e.g. technically challenging as in the example – to adjust for. Moreover, it is likely to be of importance in many gene expression studies of multicellular organisms – especially in humans, where the low amount of RNA available might lead to sample heterogeneity. Approaches to subdue this serious confounding are object of current investigations. Also, deconfounding is only as good as the inherent statistical model in which the confounding parameters are located. For a general discussion of the philosophy of modeling within these contexts we want to refer the reader to reviews on these topics [7, 8].



**Fig. 3** MapMan overview about diurnally expressed genes (dark-grey) in the central metabolism of the dicot plant *Arabidopsis thaliana* grown in a 12 h light/12 h dark regime. Shading becomes less with decreasing cycling behavior. Expression values from experimental

time points at day and night were converted to calculate a relative amplitude. The data for the experiment [20] were extracted from a public plant microarray database [21]. For details about the MapMan software see [9].

## Time as Confounder in Gene Expression Studies

In the last example, it was possible to adjust for confounding. In this section, however, we illustrate that confounding cannot be adequately adjusted for in all instances. Many gene expression experiments are either cross-sectional or case-control so that gene expression levels are measured only at a single time point. Gene expression levels may, however, change over time. For many gene regulatory systems it is already known

that they exhibit rhythms on several time scales. To illustrate, we want to point to expression dynamics measured for *Arabidopsis* on a scale of circadian rhythms [9]. Figure 3 gives an overview of which genes of *Arabidopsis* metabolism are regulated according to a circadian rhythm. Roughly a third of all surveyed genes are differentially expressed.

In this case, the confounding effect cannot be independently measured as in the case of compositions of immune cell populations discussed before, where FACS analyses were available. Hence, we are left with

the possibility to eliminate the variability in the confounding factor over cases and controls [3]. In practice, this may be achieved by sampling the plant material at a fixed time of the day.

## Comparable Accuracies in Measurements for Sample and Reference

How can the third *principle of comparability* be explained and assured in typical

designs for microarray gene expression studies? The principle of achieving comparable accuracies for measurements of gene expression in cases and controls seeks to reduce the information bias, also termed observational bias. This bias is involved in measurement errors which are differential in cases and controls. In epidemiological studies, this occurs e.g. by different quality (accuracy) of information between comparison groups.

The two possible scenarios are that, firstly, the measurement error can be removed in the analysis or, secondly, that no correction is possible in the analysis. In the latter case, choice of samples and experimental design are the only means to encounter the problem. In microarray gene expression experiments we typically struggle with both types of information bias. In the following we discuss the dye bias and two approaches to deal with it by applying a so-called *normalization* [10-12] or an *experimental design approach* [12], respectively.

## Eliminating Dye-bias in Two-color Microarray Studies – Normalization or Dye-swap?

For two-color microarray gene expression experiments, it is known that the two fluorescent dyes used to differentially label sample and control are different both in their labeling efficiency as well as in their biophysical properties influencing the hybridization reaction [12, 13]. Therefore, the *principle of comparable accuracies* [3] is severely violated.

On the one hand, it may be overcome by normalization strategies which are discussed in this special issue by Ittrich as well as by Boes [14, 15], or by estimating the bias effects by ANOVA modeling as discussed in the contribution of Bretz et al. [16]. Normalization assumes that the biasing effect can be estimated from the experimental data, given a model for this effect. However, the underlying assumptions for modeling the bias effect may interfere with the reality of the experiment in question. This may occur if many low expressed genes are indeed differentially expressed such that normaliza-

tion procedures overcorrect these effects and make them non-detectable. Thus again, as in the situation for deconfounding, also regarding normalization, the “correction” is only as good as the underlying model.

On the other hand, assuming that no correction is possible in the analysis, the comparable accuracy principle calls for nondifferential measurement errors. In the case of a two-color microarray experiment this can be achieved by repeating all arrays with the dyes, i.e. colors, reversed for cases and controls. For technical details see the contribution of Repsilber and Ziegler in this special issue [17]. This way, each case and control is measured with either dye, assuring nondifferential measurement errors. Another approach to reduce dye-bias in experiments involving more than one individual (human, animal, plant) is the so-called balanced labeling design where the dyes for labeling are assigned to both groups under comparison in equal proportions [12].

In summary, the comparable accuracy principle calls for design modification if the underlying bias cannot be corrected for. This decision, however, is often unclear and represents a compromise between use of resources and guarantee of unbiased results.

## Summary: Applying the Principles of Comparability to Microarray Gene Expression Studies

In this paper, we have shown that the *principles of comparability* as advised for epidemiological studies provide a comprehensive framework also for recommending principles for the choice of biomedical material in microarray gene expression studies. Our contribution is intended for both sides; the statistician as well as the wet-lab researcher. Necessarily, our language is simple from the statistician’s point of view and at the same time introductory for the wet-lab researcher. Conversely, the wet-lab examples have to be introductory for the statistician, and are, hence, everyday-knowledge for the biologist or medical re-

searcher. Thus, the deliberate, mutual introductory character of our contribution intends to improve the evidently needed better communication just between the wet-lab researcher and the statistician. Without doubt, a deepening and consolidation is necessary from both sides – once the starting point is made, which our contribution aims to establish.

Most importantly, the *study-base principle*, as applied in the context of microarray gene expression experiments, confronts the researcher with the need to thoroughly define the research question. This especially includes appropriate descriptions of which cell populations should be compared under which conditions. Generally, it is advisable to compare different conditions for the same cell type, as gene expression experiments very sensitively reflect any changes in conditions. Thus, tissue heterogeneity represents a major problem for all microarray gene expression experiments involving multicellular organisms [1, 18, 19]. Solutions lie in either dissection of homogeneous parts of the tissue of interest, or in correcting for independently observed tissue heterogeneity during analysis. The latter can be classified as an approach to satisfy the *deconfounding principle*. Here, especially for microarray analysis, it is important to be aware of the dynamic of gene expression levels which may exhibit circadian rhythms on numerous time scales. The third principle, the *comparable accuracy principle*, is not as important for the question of which biomedical material should be selected for in a microarray gene expression experiment. However, it has implications for experimental design and analysis approaches.

In summary, the choice of biomedical material is important at the design stage since many biases can be introduced. Results from gene expression experiments can be well interpreted and valid conclusions may be drawn only if the fundamental principles for designing epidemiological and clinical epidemiological studies are applied.

### Acknowledgments

This work was supported in part by the German Ministry of Education and Research within the National Genome Research Net (NGFN), Grants: NIE-S02T21, PGE-S26T11.

## References

1. Bakay M, Chen Y-W, Borup R, Zhao P, Nagaraju K, Hoffman E. Sources of variability and effect of experimental approach on expression profiling data interpretation. *BMC Bioinform* 2002; 3: 4.
2. Schulz K, Grimes DA. Case-control studies: research in reverse. *Lancet* 2002; 359: 431-4.
3. Wacholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies. I. Principles. *Am J Epidemiol* 1992; 135 (9): 1019-28.
4. Fink L., Kohlhoff S, Stein MM, Hänze J, Weissmann N, Rose F, Akkayagil E, Manz D, Griminger F, Seeger W, Bohle RM. cDNA Array Hybridization after Laser-Assisted Microdissection from Nonneoplastic Tissue. *Am J Pathol* 2002; 160 (1): 81-90.
5. Cooper S, Shedden K. Microarray analysis of gene expression during the cell cycle. *Cell Chrom* 2003; 2: 1.
6. Weinberg CR. Towards a clearer definition of confounding. *Am J Epidemiol* 1993; 137: 1-8.
7. Breiman L. Statistical Modeling: The Two Cultures. *Stat Sci* 2001; 16 (3): 199-231.
8. Hilborn R, Mangel M. *The Ecological Detective: Confronting Models with Data*. Princeton: University Press; 1997.
9. Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller L A, Rhee S Y, Stitt M. *Plant J* 2004; 37: 914-39.
10. Quackenbush J. Computational Analysis of Microarray data. *Nat Rev Genet* 2001; 2: 418-26.
11. Quackenbush J. Microarray data normalization and transformation. *Nat Genet Suppl* 2001; 32: 496-501
12. Simon RM, Korn EL, McShane LM, Radmacher MD, Wright GW, Zhao Y. *Design and Analysis of DNA Microarray Investigations*. Heidelberg: Springer; 2003.
13. Schena M. *Microarray Analysis*. Heidelberg: Springer; 2003.
14. Boes T, Neuhäuser M. Normalization for Affymetrix GeneChips. *Methods Inf Med* 2005; 44: 414-7 (this issue).
15. Itrich C. Normalization for Two-channel Microarray Data. *Methods Inf Med* 2005; 44: 418-22 (this issue).
16. Bretz F, Landgrebe J, Brunner E. Design and Analysis of Two-color Microarray Experiments Using Linear Models. *Meth Inf Med* 2005; 44: 423-30 (this issue).
17. Reipsilber D, Fink L, Jacobsen M, Bläsing O, Ziegler A. Sample Selection for Microarray Studies. *Methods Inf Med* 2005; 44: 461-7 (this issue).
18. The Tumor Analysis Best Practices Working Group. Expression profiling – best practices for data generation and interpretation in clinical trials. *Nat Rev Genet* 2004; 5 (3): 229-37.
19. King H C, Sinha A A. Gene Expression Profile Analysis by DNA Microarrays – Promise and Pitfalls. *JAMA* 2001; 286: 2280-8.
20. Thorneycroft D, Chapple A, Dunstan H, Hylton C, Zeeman SC, Smith AM. Diurnal Changes in the Transcriptome Encoding Enzymes of Starch Metabolism Provide Evidence for Both Transcriptional and Posttranscriptional Regulation of Starch Metabolism in Arabidopsis Leaves. *Plant Physiol* 2004; 136: 2687-99.
21. Craigon DJ, James N, Okyere J, Higgins J, Jotham J, May S. NASCArrays: A repository for Microarray Data generated by NASC's Transcriptomics Service. *Nucleic Acids Res* 2004; 32: D575-D577. (<http://affymetrix.arabidopsis.info/>)

### Correspondence to:

Prof. Dr. Andreas Ziegler  
 Institut für Medizinische Biometrie und Statistik  
 Universitätsklinikum Schleswig-Holstein, Campus Lübeck  
 Ratzeburger Allee 160, Haus 4  
 23538 Lübeck, Germany  
 E-mail: [ziegler@imbs.uni-luebeck.de](mailto:ziegler@imbs.uni-luebeck.de)