Experimental Designs for Array Comparative Genomic Hybridization Technology

S.K. McDonnell^a S.M. Riska^a E.W. Klee^a E.C. Thorland^b N.E. Kay^c
S.N. Thibodeau^d A.S. Parker^e J.E. Eckel-Passow^a

^aDivision of Biomedical Statistics and Informatics, ^bMedical Genetics, ^cDivision of Hematology and Internal Medicine, and ^dDepartment of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, Minn., and ^eDepartment of Urology, Mayo Clinic, Jacksonville, Fla., USA

Key Words

Balanced-block design · Balanced incomplete-block design · Dye-swap design · Reference design

Abstract

Array comparative genomic hybridization (aCGH) technology is commonly used to estimate genome-wide copy-number variation and to evaluate associations between copy number and disease. Although aCGH technology is well developed and there are numerous algorithms available for estimating copy number, little attention has been paid to the important issue of the statistical experimental design. Herein, we review classical statistical experimental designs and discuss their relevance to aCGH technology as well as their importance for downstream statistical analyses. Furthermore, we provide experimental design guidance for various study objectives.

Copyright © 2013 S. Karger AG, Basel

An experimental design is critical for effectively facilitating studies proposing to evaluate the association between copy number and disease. Recently, there has been an explosion of methodologies for preprocessing and an-

alyzing genome-wide copy-number data. However, little attention has been paid to the important topic of experimental design. Most studies that utilize array comparative genomic hybridization (aCGH) for copy-number interrogation use a reference design, where target samples are hybridized with one fluorescent dye and a common reference sample is included on every array and labeled with the other fluorescent dye (fig. 1a). The reference design is intuitive as the relative copy number for every target sample is calculated using the same common reference sample co-hybridized on each array and under the same experimental conditions (denoted here after as the on-chip reference). Although intuitive, the reference design allocates half of the available resources to a reference sample that is ultimately not of biological interest and thus is clearly not cost effective.

There are numerous publications in the 2-color gene-expression microarray literature that advocate for the use of more efficient aCGH experimental designs, including the seminal paper by Kerr and Churchill [2007]. In fact, such publications relate microarray experiments with classical field trials and show that balanced-block and incomplete-block designs are more efficient than a reference design. With respect to aCGH technology, Buffart et al. [2008] showed empirically that on-chip reference sam-

ples are not necessary for calculating relative copy number and demonstrated that off-chip reference samples can be utilized. Despite this finding, they did not suggest alternative experimental designs that utilize off-chip reference samples for aCGH analysis. Thus, herein we provide further evidence that an off-chip reference can be used to estimate copy number and provide guidance for choosing alternative experimental designs.

Estimating Copy Number for aCGH

Our focus is on microarray-based platforms for estimating genome-wide copy number. The primary manufacturer of microarray-based platforms is Agilent, which uses a 2-color system. In a 2-color system, 2 samples are separately labeled using fluorescent dyes, the labeled samples are subsequently mixed together, and the mixture is hybridized on an array. A machine then scans the array at 2 distinct wave lengths to obtain separate fluorescent intensity values for each probe on the array, generating sample specific measurements for the 2 samples assayed.

Although the measured intensities are correlated with copy number, absolute copy number is not obtainable due to experimental artifacts in the data. Thus, a baseline sample is required as a comparator when estimating copy number. When using a reference design, the on-chip reference sample is used as the baseline sample for estimating copy number. For example, for the commonly employed reference design, the endpoint of interest is the ratio of the 2 labeled samples for each probe on the microarray, which represents the relative copy number for each probe.

Experimental Design

There are 2 general experimental scenarios when using a 2-color array platform: (i) analysis of paired samples and (ii) analysis of unpaired or independent samples. Paired samples can be used to identify novel copy-number aberrations. For example, analyzing patient-matched tumor and normal tissues eliminates the discovery of common copy-number variants and instead allows for identification of tumor-specific copy-number aberrations. When analyzing paired samples, an efficient design is to hybridize the subject-matched samples to the same array. Although the optimal experimental design of paired samples is rather intuitive, optimal designs for research studies evaluating un-paired samples are rarely

Fig. 1. Experimental designs for 2-color microarrays: reference design (a), dye-swap design (b), balanced-block design (c), and balanced incomplete-block design (d). R denotes a reference sample and A, B and C denote 3 different experimental groups of interest. Numbers denote biological replicates for each experimental group. The tip of each arrow points to the sample labeled with one of the dyes (e.g. Cy3), and the tail of each arrow points to the sample labeled with the other dye (e.g. Cy5).

used. Thus, the focus of this paper is to discuss efficient experimental designs for research projects that propose to analyze unpaired (or independent) samples.

The most commonly employed experimental design for aCGH studies using unpaired samples is the *reference design* (fig. 1a). As discussed above, the reference design utilizes the same reference sample on every array. This reference sample is not of biological interest; however, the probe intensity for a sample of interest is measured relative to the intensity of the same probe on the same array for the on-chip reference sample. With respect to resources, to process n biological samples of interest, the reference design requires n arrays.

Buffart et al. [2008] showed that an on-chip reference sample is not necessary for estimating relative copy number. In fact, they compared the use of an on-chip reference with the use of an off-chip reference that was processed on the same day as well as with an off-chip reference that was processed on a different day. The study showed that the use of an off-chip reference processed on the same day resulted in copy-number estimates with similar precision as an on-chip reference. It further showed that the use of an off-chip reference that was processed on a different day decreased the precision almost by half. Although Buffart et al. [2008] demonstrate that

an off-chip reference can be used to estimate copy number if the samples are processed at the same time point, they do not provide alternative designs for investigators to consider. Herein, we present data pertaining to alternative experimental designs that utilize an off-chip reference to estimate copy number that are more efficient than a reference design.

Recommendations with respect to experimental designs for gene-expression and immunoaffinity-protein arrays are widely available in the literature [e.g. Simon et al., 2002; Eckel-Passow et al., 2005; Kerr and Churchill, 2007], and we propose that the same recommendations hold for aCGH studies. The reference and dye-swap designs (fig. 1a, b) both use an on-chip reference and thus are intuitive for addressing the dye biases inherent in 2-color arrays and estimating relative intensity values produced by 2-color arrays. The main disadvantage of the reference design is that half of the hybridizations (and thus, half of the cost of the study) are used for a reference sample that is ultimately not of biological interest. Additionally, it has been demonstrated that probe-specific dye biases exist and these biases cannot be accounted for when using a reference design [Churchill, 2002; Kerr and Churchill, 2007]. Probe-specific dye biases are defined as probes that exhibit an increased intensity when labeled with one dye or the other, regardless of the sample, and cannot be detected nor corrected for using a reference design resulting in biased estimates of relative copy number. To eliminate artifacts due to probe-specific dye biases, a *dye-swap design* has been previously proposed (fig. 1b). In a dye-swap design, each sample is run in duplicate, once labeled with the Cy3 dye and then again on a separate array with the Cy5 dye. Although the dye-swap design corrects for probe-specific dye bias, it requires that every sample be hybridized twice and, therefore, doubles the cost of a study in comparison to the reference design. Here, we propose that block designs, which are more efficient with respect to both cost and variability, are an appealing alternative in comparison to the reference and dye-swap designs.

Two-color aCGH arrays can be considered experimental blocks with block size of 2 (i.e. 2 samples can be hybridized to each array). In statistics, a *blocking factor* is a variable that is not of primary interest, yet it contributes unwanted variability in the endpoint under investigation. In aCGH, arrays should be considered a blocking factor and any experimental design should account for this. When there are 2 groups of interest, e.g. prostate cancer versus healthy controls, both groups can appear on each array. This is referred to as a *balanced-block design*

(fig. 1c). If there are more than 2 groups of interest, not every group can appear simultaneously on each array resulting in a *balanced incomplete-block design* (fig. 1d). The design is still balanced since for each group half the samples are labeled with the Cy3 fluorescent dye and likewise half are labeled with the Cy5 fluorescent dye. With respect to estimating copy number, the reference and dye-swap designs use an on-chip reference while the balanced-block and incomplete balanced-block designs would necessitate an off-chip reference.

Overall, the experimental design should depend on the objective of the study as well as on the sources of variability. The reference design assumes that array effects are large, and thus, the optimal approach for estimating copy number is to hybridize a reference sample to every array and subsequently, to use the on-chip reference sample to estimate copy number. We evaluated this assumption by quantifying assignable causes for the observed variability using an in-house study where 52 chronic lymphocytic leukemia (CLL) patients were analyzed on the Agilent Human Genome CGH 1×1M Microarray, which has nearly one million distinct copy-number probes [Kay et al., 2010]. Of the 52 samples, only 50 were analyzed herein: 1 sample did not have adequate DNA and 1 sample had a low derivative log ratio spread (DLRS <0.35). A reference design was utilized for the CLL case study; the reference sample consisted of a pool of 9 female lymphoblastoid cell lines. A variance component model was used to estimate variability due to array and subject and the results are presented in table 1. For the CLL case study, the subject-to-subject variance of 0.067 is nearly 2 times larger than the array-to-array variance. As a second case study, a variance component model was run on 199 male clinical samples that were processed in the Mayo Clinic Cytogenetics Laboratory within a 2-week time span on the Agilent Human Genome CGH 4×180k Microarray. A reference design was utilized for all samples; the reference sample consisted of a single male sample. As shown in table 1, the subject-to-subject variance is again nearly 2 times larger than the array-to-array variance. Thus, both case studies evaluated herein suggest that it is more important to spend resources on additional independent biological samples versus running a reference sample on every array in an attempt to reduce array variability. Furthermore, both case studies corroborate the results provided by Buffart et al. [2008] and suggest that alternative experimental designs, such as a balanced-block or incomplete-block design, should be considered since such designs allow almost twice as many samples to be analyzed at the same cost as a reference design.

Table 1. Variance components model

Source	CLL case study (n = 50) Agilent 1×1M			Clinical case study (n = 199) Agilent 4×180k		
	σ (SE)	% variation	p value	σ (SE)	% variation	p value
Array	0.034 (0.015)	9.7	0.013	0.024 (0.005)	2.8	<0.0001
Dye	0.044 (0.064)	12.5	0.2463	0.001 (0.002)	0.1	0.2809
Subject	0.067 (0.013)	19.0	<0.0001	0.045 (0.004)	5.3	<0.0001
Error	0.207 (3.3×10^{-5})	58.8	<0.0001	0.781 (0.0001)	91.8	<0.0001

A variance components model was used to determine sources of observed variability in aCGH data. The variance attributed to error denotes the variability across probes. Variance components were estimated for 2 case studies that utilized 2 different Agilent whole-genome CGH microarrays; base-two logarithm transformed intensity data were analyzed. For the CLL case study, vari-

ance components were estimated using 800,000 randomly chosen probes. σ = Estimated variance associated with the corresponding source; SE = standard error associated with the estimated variance; % variation = percent of total variance attributed to the corresponding source; p value = result of testing if the variance component equals zero.

As mentioned above, in order to utilize a balance-block or balanced incomplete-block design, an off-chip reference needs to be used to estimate copy number. Thus, we empirically evaluated the use of an off-chip reference for estimating copy number. We estimated the bias and variability associated with averaging the probe-level intensity across a set of reference samples and using this average-reference array as the baseline for estimating copy number in comparison to using an on-chip reference as the baseline for estimating copy number. Using the CLL case study, which utilized a reference design, we constructed average-reference arrays of 50 different sample sizes by averaging each probe intensity across $k = 1, 2, \dots$ up to all 50 reference samples. Subsequently, we evaluated the use of each average-reference array as the baseline for estimating copy number for each of the 50 CLL samples. For $k = 1$, each of the 50 reference samples were evaluated in turn. When averaging $k = 2, 3, \dots, 49$ reference samples, 200 bootstrap datasets were obtained for each size k , and copy number estimates were estimated for each bootstrap dataset using the average reference array constructed with k randomly selected reference samples as the baseline sample. Lastly, all $k = 50$ reference samples were averaged to create an average reference sample, which was used as the baseline for estimating copy number for each of the 50 CLL samples. We evaluated both bias (fig. 2a) and variance (fig. 2b) of the estimated copy numbers associated with using an average of k reference arrays as the baseline for estimating copy number. We compared the bias and variance with that obtained using the on-chip reference (denoted as $n = 0$ in fig. 2). Bias was calculated for each probe as the difference from a value of 2. The median ab-

solute deviation of the copy-number estimates across all probes for a given subject was used as a robust measure of variability. Figure 2a demonstrates the law of large numbers and shows that as we averaged across more samples, the average bias approaches the true expected average, which is zero in the current scenario. Similarly, figure 2b also demonstrates the law of large numbers. For the CLL case study, figure 2b demonstrates that the average of 10 or more reference samples (i.e. 20% of the total number of samples) results in more stable estimated copy numbers than using the on-chip reference as the baseline. In summary, the CLL case study we report here suggests that only 20% of the resources need to be used on a reference sample, and thus, the remaining 80% of resources can be used on the case samples of primary interest.

Competitive Binding

One concern with using experimental designs such as the balanced-block or incomplete-block design has to do with competitive binding. That is, there are concerns that co-hybridizing 2 case samples of interest that have the same chromosomal abnormalities will result in biased estimates of copy number. Buffart et al. [2008] co-hybridized their BT474 cell line and showed that the relative copy numbers of the *ERBB2* and *CYP24A1* genes were not suppressed and thus concluded that comparative hybridization is not competitive. To further evaluate competitive binding, we performed a more in-depth experiment. Specifically, we co-hybridized technical replicates for samples that had known regions of homozygous dele-

Fig. 2. Using the CLL case study, which utilized a reference design, we empirically estimated bias (a) and variability (b) associated with averaging the probe-level intensity across a set of k reference samples and using this average-reference array as the baseline for estimating copy number for each of the 50 CLL subjects. The x-axis denotes k , the number of reference samples that were averaged

to create the average-reference baseline array; $k = 0$ denotes the use of the on-chip reference as the baseline sample, which is considered the gold standard. The boxplots denote the distribution across 200 bootstrap samples. n = Number of bootstrap datasets generated.

Table 2. Competitive binding self-self experiment

Copy number	Sample description	Cytoband	Physical location
2	normal male reference		
0	<i>NPHP1</i> homozygous deletion	2q13	110,219,766–110,322,026
1	unbalanced translocation ¹	6	152,739,166–170,763,014
3	unbalanced translocation ¹	12	100,682–2,982,747
4	PW tandem triplication	15q11q13	20,249,886–26,829,558

¹ A single subject with both one copy gain (chr12) and one copy loss (chr6).

tions, heterozygous deletions, heterozygous duplications, and homozygous duplications as well as a male reference sample (table 2). Utilizing an off-chip average baseline approach, we averaged the technical replicates for the male reference sample and used it as the baseline for calculating copy number for each of the co-hybridized sam-

ples. The off-chip reference baseline results were compared with those obtained using the on-chip reference from a reference design (previously generated data). In all instances, the use of an off-chip baseline identified the known copy-number regions (data not shown).

Fig. 3. Experimental designs for aCGH research studies that include a reference sample(s) for estimating copy number but spend fewer resources on the reference sample in comparison to the experimental groups of interest. Example designs are provided for cohort studies (**a**), studies comparing 2 experimental groups (**b**) and studies comparing 3 experimental groups (**c**). For further information, see figure 1.

Experimental Design Examples

Here, we provide examples of experimental designs that can be used for aCGH research-based studies. For each, the 3 basic principles of experimental design are adhered to: replication, randomization and blocking [Montgomery, 1997]. Biological *replication* is essential for assuring that results are real and reproducible to the broader population. *Randomization* refers to randomly sampling subjects from the population and is important for obtaining valid statistical tests. Randomization can be further utilized to reduce bias in aCGH studies by randomizing samples to arrays. Lastly, as discussed previously, in 2 color aCGH studies, arrays should be considered a *blocking* factor. Measurements made within a block should be more similar than measurements across blocks, and thus, if possible, it is important to include samples from each group of interest within each block. Additionally, if an aCGH study requires processing and running samples over a span of several days or weeks, then day/week should be considered another blocking factor. For aCGH studies, we advocate that *balance* is a 4th principle of experimental design that should be adhered to. Specifically, we and others [Kerr and Churchill, 2007] suggest that each group of interest should be balanced with respect to dyes, i.e. half of the group should be labeled with the Cy3 fluorescent dye and likewise half of the group should be labeled with the Cy5 fluorescent dye. The importance of replication, randomization and blocking in gene expression microarray studies is nicely discussed in detail by Churchill [2002].

Example 1: Cohort Study

Consider a study where the objective is to evaluate the association between copy number and clinical outcome

for a particular disease of interest (e.g. survival for brain cancer patients) and that a baseline reference sample(s) is desired in order to estimate copy number for each case subject of interest. Furthermore, assume that there are financial resources to run 10 arrays and that sample processing will take 2 days to complete. For the *cohort study* described here, we suggest co-hybridizing 2 case subjects together on every array and additionally including 2 arrays that denote the reference sample(s). Thus, 20% of the array resources are spent on the reference sample. Figure 3a displays the design to use for processing samples for each of the 2 days that are necessary to complete this study. Within each block (day is the blocking factor in the current scenario), the reference samples can be averaged and subsequently the average can be used as the baseline for estimating copy number for each disease subject within the corresponding block. Based on our empirical studies, we suggest a minimum of 2 off-chip references be averaged together and used as the baseline for estimating copy number. We further advise that the reference samples be balanced with respect to dye. Even so, if one would like to utilize a larger number of reference samples, a balanced-block design (fig. 1c) could be considered where half the resources are spent on the case samples of interest and the other half on the reference sample. The balanced-block design still offers twice as many case subjects to be studied than the reference design, having a dramatic effect on the statistical power to detect associations.

Example 2: Comparing 2 Groups

Consider a study where the objective is to compare copy numbers across 2 biologically interesting groups. For example, the objective might be to compare ER-positive versus ER-negative breast cancer and assume that a baseline reference sample(s) is desired to estimate copy

number for each case subject. Again, we assume that there are sufficient financial resources to run 10 arrays and that sample processing will take 2 days to complete. For the case-case study described here, we suggest using an *augmented balanced-block design*. Figure 3b displays an augmented balanced-block design that could be used for processing samples for each of the 2 days that are necessary to complete this study. This design is essentially a classical balanced-block block design; however, 2 arrays that denote the reference sample(s) are included for the sole purposes of estimating copy number. As was the case in example 1, 20% of the array resources are spent on the reference sample. And, within each block (day is the blocking factor in the current scenario) the reference samples can be averaged, and subsequently, the average can be used as the baseline for estimating copy number for each disease subject within the corresponding block. If one would like to utilize a larger number of reference samples, an incomplete balanced-block design (fig. 1d) could be considered where two-thirds of the resources are spent on the case samples of interest and one-third on the reference sample. The incomplete balanced-block design still offers twice as many case subjects to be studied than the reference design, which has a significant effect on the statistical power to detect differences between the 2 case groups of interest.

Example 3: Comparing More than 2 Groups

For a study comparing more than 2 groups, we would similarly advocate the use of an *augmented balanced incomplete-block design*. Figure 3c displays an augmented balanced incomplete-block design that could be used for processing samples for each of the 2 days that are necessary to complete this study. This design is essentially a classical balanced incomplete-block design; however, 2 arrays that denote the reference sample(s) are included for the sole purposes of estimating copy number. Again, within each block (day is the blocking factor in the current scenario) the reference samples can be averaged, and subsequently, the average can be used as the baseline for estimating copy number for each disease subject within the corresponding block.

Discussion

Even though Buffart et al. [2008] published their empirical results 4 years ago suggesting that an off-chip reference sample is sufficient for estimating copy number, the reference and dye-swap designs are still the primary

Table 3. Variance estimates for an in-house custom-designed CGH microarray (Agilent) targeting a region on chromosome 22

Source	All subjects (n = 60)		
	σ (SE)	% variation	p value
Array	0.127 (0.024)	12.1	<0.0001
Dye	0.008 (0.011)	0.008	0.2409
Subject	0.002 (0.0004)	0.002	<0.0001
Error	0.914 (0.0004)	87.0	<0.0001

experimental designs utilized for aCGH studies. From a research perspective, the balanced-block and incomplete-block designs allow investigators to nearly double their sample size while using the same number of arrays as required for a reference design. This doubling of sample size has an enormous effect on statistical power, which is of utmost importance in genome-wide research studies exploring the association between copy number and disease.

Using 2 different human Agilent whole-genome CGH microarrays (Agilent Human Genome CGH 1×1M Microarray and Agilent Human Genome CGH 4×180k Microarray), we showed empirically that subject-to-subject variance was almost twice as large as array-to-array and dye-to-dye variance. In addition to these 2 whole-genome Agilent microarrays, we also evaluated a microarray that was manufactured by Agilent; however, it was designed in-house to target a 12-Mb region on chromosome 22 (contained 105,070 total probes, including 1,456 control probes from other chromosomes). Interestingly, the results varied dramatically from the 2 whole-genome microarrays we studied. For the custom microarray, the array-to-array variability was more than 10 times larger than both subject-to-subject and dye-to-dye variance (table 3). We hypothesize that the array-to-array variability was much larger in the custom microarray because the probes were chosen to achieve the highest density coverage possible for chromosome 22 and thus included probes that were less reliable. Based on these empirical results, for custom aCGH platforms where the array-to-array variability is expected to be large, a reference design is likely more appropriate. Additionally, for custom microarrays, it might be beneficial to consider running additional sources of replicates in order to identify the largest sources of variability.

The on-chip reference design is currently the standard for clinical CNV genetic testing [Miller et al., 2010]. Based on our empirical results utilizing a set of 199 clinical sam-

ples in which the subject-to-subject variance was estimated to be 2 times larger than the array-to-array variance, it is worth considering whether an off-chip reference may be useful in the clinical setting. In this case, a clinical lab could run one array with a reference sample labeled with both Cy3 and Cy5 and use the average of the 2 samples as the reference for all clinical samples. Of course, further testing will be necessary before use of an off-chip reference is accepted into clinical practice.

For SNP-based platforms (e.g. Illumina and Affymetrix platforms), where a single sample is represented on each array, off-chip reference samples are utilized for estimating copy number. For example, the intensity measures across samples are averaged at each probe, and this average array is used as the baseline for estimating copy number. Similarly, for SNP-based platforms, the HapMap samples can be averaged and used as the baseline for estimating copy number. However, McMullan et al. [2009] showed that locally run reference samples provide more precise estimates of copy number in comparison to reference samples that were run at a different institution. They further discuss that reference samples should be run in parallel to the case samples under study; however, they do not discuss how many reference samples are necessary. The experimental designs discussed herein – the balanced-block and balanced incomplete-block designs – require a similar approach to estimating copy number, namely the use of an off-chip reference.

As is well documented in the 2-color gene expression array literature, performing statistical association analyses on relative fold change values (analogous to relative copy-number values) is not optimal. Specifically, there is a wealth of literature stating that in order to account for the variability resulting from the various experimental artifacts associated with a microarray experiment, it is advised to model the individual intensity values at each probe using analysis of variance (ANOVA) models [Kerr et al., 2000; Wolfinger et al., 2001]. One could argue that the same holds true for copy-number association analyses. That is, instead of performing analyses on the relative copy-number values, one should perform analyses on the probe intensity values after appropriate normalization to account for experimental artifacts (e.g. array and dye effects). The analysis of copy-number data is outside the scope of the current review; however, the use of appropriate experimental designs – such as the balanced-block and balanced incomplete-block designs discussed herein – allow for the analysis of probe-intensity values directly.

Acknowledgement

This work was funded in part by the Fraternal Order of Eagles (J.E.E.-P. and S.N.T.) and NIH R01-CA95241 (N.E.K.).

References

- Buffart TE, Israeli D, Tijssen M, Vosse SJ, Msić A, et al: Across array comparative genomic hybridization: a strategy to reduce reference channel hybridizations. *Genes Chromosomes Cancer* 47:994–1004 (2008).
- Churchill GA: Fundamentals of experimental design for cDNA microarrays. *Nat Genet* 32 Suppl:490–495 (2002).
- Eckel-Passow JE, Hoering A, Therneau TM, Ghorbail I: Experimental design and analysis of antibody microarrays: applying methods from cDNA arrays. *Cancer Res* 65:2985–2989 (2005).
- Kay NE, Eckel-Passow JE, Braggio E, VanWier E, Shanafelt TD, et al: Progressive but previously untreated CLL patients with greater array CGH complexity exhibit a less durable response to chemoimmunotherapy. *Cancer Genet Cytogenet* 203:161–168 (2010).
- Kerr MK, Churchill GA: Statistical design and the analysis of gene expression microarray data. *Genet Res* 89:509–514 (2007).
- Kerr MK, Martin M, Churchill GA: Analysis of variance for gene expression microarray data. *J Comput Biol* 7:819–837 (2000).
- McMullan DJ, Bonin M, Hehir-Kwa JY, de Vries BBA, Dufke A, et al: Molecular karyotyping of patients with unexplained mental retardation by SNP arrays: a multicenter study. *Hum Mutat* 30:1082–1092 (2009).
- Miller DT, Adam MP, Aradhya S, Biesecker LG, Brothman AR, et al: Consensus statement: chromosomal microarray is a first-tier clinical diagnostic test for individuals with developmental disabilities or congenital anomalies. *Am J Hum Genet* 86:749–764 (2010).
- Montgomery DC: *Design and Analysis of Experiments*, ed 4 (John Wiley & Sons, New York 1997).
- Simon R, Radmacher MD, Dobbin K: Design of studies using DNA microarrays. *Genet Epidemiol* 23:21–36 (2002).
- Wolfinger RD, Gibson G, Wolfinger ED, Bennett L, Hamadeh H, et al: Assessing gene significance from cDNA microarray expression data via mixed models. *J Comput Biol* 8:625–637 (2001).