Genomic Microarray Quality Assurance

Catherine D. Kashork, Lisa G. Shaffer and Kyle S. Sundin
Signature Genomics Laboratories
USA

1. Introduction

The use of microarray technology is revolutionizing the field of clinical cytogenetics. This new technology has transformed the cytogenetics laboratory by the adaptation of techniques that had previously been the province of molecular geneticists. Proficiency with these techniques is now a must for the modern cytogeneticist. This chapter will focus on quality assurance principles associated with microarray analysis for the diagnosis of copy number changes associated with genetic disease.

Microarrays consist of a glass slide or other solid support on which small amounts of DNA (“probes” or “targets”) are deposited and immobilized in an ordered fashion (DeRisi et al., 1996; Schena et al., 1995). Probes vary in size from oligonucleotides manufactured to represent genomic regions of interest (25-85 base pairs [bp] of DNA) to large genomic clones such as bacterial artificial chromosomes (BACs, 80-200 thousand base pairs [kb]). Analysis methodology for microarray-based comparative genomic hybridization (aCGH) is consistent regardless of the probe content. First, DNA is extracted from a test sample (e.g., blood, skin, cells from pregnancy). The patient DNA is labeled with a fluorescent dye, while a DNA from a normal control (reference) sample or pooled control samples is labeled with a different-colored fluorescent dye. The two genomic DNAs, test and reference, are then mixed together and applied to the array. Because the DNAs have been denatured, they are single strands; when applied to the array, the single-strand DNAs hybridize with the arrayed single-strand probes. Using a dual-color scanner, digital images are captured, and the relative fluorescence intensities of the hybridized labeled DNA probes are quantified. The fluorescence ratio of the test and reference hybridization signals is determined at different positions along the genome and provides information on the relative copy number of sequences in the test genome compared to the normal diploid genome, enabling the detection of submicroscopic chromosomal deletions and duplications at an unprecedented level (Beaudet & Belmont, 2008; Shaffer & Bejjani, 2009).

Launching a new assay in the clinical setting requires an effective validation of the assay, clear protocols for use at the bench and clearly defined quality assurance (QA) and quality control (QC) procedures prior to the launch. Every laboratory must develop a strong Quality Management System (QMS) that is coordinated with the defined policies under regulatory bodies, such as CLIA ’88 (Schwartz, 1999), College of American Pathologists and state regulating agencies. These agencies perform rigorous inspections and verify that a diagnostic laboratory follows defined principles to ensure quality patient care and correct diagnosis. This chapter covers many of the QA and QC principles identified and monitored for laboratories offering microarray-based diagnostics.
2. Quality systems with strong monitoring for quality metrics

Regulatory bodies require diagnostic laboratories to build a strong QMS (Deming, 2000). A robust QMS integrates the organization’s processes, policies and procedures for total quality management. In the diagnostic laboratory industry, CAP and other accrediting bodies require defined metrics throughout all phases of testing, including pre-analytical, analytical and post-analytical.

For microarray technology, pre-analytical metrics may include assessment of DNA quality and yield. Each laboratory must define the ideal quality of DNA prior to implementing the assay into clinical testing. In our experience, a gel assessment that indicates clean genomic DNA free of RNase and degradation should lead to quality microarray results. If a DNA specimen has artifacts or appears to have degraded (Fig. 1), the laboratory should inform the client that results may be compromised because of DNA quality or obtain a new specimen from which to perform the analysis. In addition to a visual assessment of the DNA via gel electrophoresis, the laboratory should assess the DNA yield post-extraction. The quantity of DNA required in the analytic labeling phase of microarray analysis determines DNA yield requirements. Spectrophotometric assessment of DNA offers two indicators of quality DNA, including quantity and purity. A 260/280 nm reading indicates quantity, and a 260/230 nm reading indicates purity. These measurements are imperative for the downstream labeling process. Insufficient DNA quantity and quality (purity) will compromise successful microarray analysis. The spectrophotometer measures optical density (OD), which is the physical process of absorbing light. The OD, or absorbance, is calculated as a mathematical quantity. OD readings for pure DNA should measure at 1.8 (Sambrook & Russell, 2001). Our laboratory uses OD measurements from 1.4 to 1.8, although quality of labeling product can be compromised at the lower OD readings. There are many causes of poor yield, including compromised technique during extraction and poor sample quality (e.g., from increased age or exposure).

Assessment of quality should be implemented throughout all phases of testing including the analytic phase. For aCGH, analytical metrics may include, but are not limited to, spectrophotometric assessment of the labeling product and the identification of labeling efficiency, which has an impact on results. Microarray analysis usually requires a dye incorporation using a random priming method. The dyes are tagged to a promoter

Fig. 1. Gel electrophoresis for the assessment of DNA degradation. Lane 1 has the molecular mass standard. Lanes 2 & 5 show high molecular weight samples that do not exhibit any signs of degradation. The two samples in lanes 3 & 4 show lower molecular weight DNA below the main high molecular weight bands in the other lanes. Degraded DNA typically leads to compromised array results.
molecule, and the promoter molecule is incorporated into the genomic DNA. As with pre-analytic assessment of DNA, the quantity assessment of the labeling product is performed by assessing the 260/280 nm readings from the spectrophotometer. Laboratories should define post-labeling quantity requirements that indicate labeling efficiency. In addition to labeling efficiency as a quality indicator, visual assessment of the set-up of labeling product on the array should be considered. In the event that there are air bubbles or non-complete contact of array product to the hybridization area (Fig. 2), the quality of the microarray result may be compromised and should be documented.

Fig. 2. Microarray after hybridization with an air bubble. The air bubble creates an area of incomplete contact of array product to the hybridization area, which compromises the array.

Post-analytic assessments may include average standard deviation (SD), intensity and background values. The SD value is the standard deviation of the normalized log₂ intensity ratios for autosomal regions (excluding large copy number imbalances) and provides a measure of quality for aCGH experiments (Vermeesch et al., 2005). The SD value provides a quantitative metric that is relative to the overall noise on an array. As the overall noise of an aCGH experiment increases, so does the SD value. Our laboratory has established SD values
that indicate whether an aCGH experiment is optimal, suboptimal, or failed. In our laboratory, the average SD is used daily to monitor the collective SD values for all patients. Daily monitoring of the average SD value allows for establishment of a system to monitor the average SD over time. If a shift in the average SD value is observed, the laboratory processes can be evaluated to determine the potential cause and potentially prevent a system-wide failure. Each laboratory will need to define and validate a quality metric to measure the quality of the array and implement a system to track the performance of the metric.

In addition to monitoring the SD value of the array, the signal intensity of the two fluorescent dyes relative to the background can be monitored. These two values can be tracked independently or together by monitoring the signal-to-noise ratio (SNR). The SNR is the signal intensity divided by the background noise. Low signal intensity or high background noise will result in a low SNR value. A low SNR is an indicator for poor-quality array data (Basarsky et al., 2000). Low signal intensities can result from several factors, including poor fluorescent dye incorporation in labeling, inadequate denaturation of the probe, inadequate quantity of the probe, and suboptimal hybridization. Several factors can result in high background noise, including labeling reaction impurities, drying of the array during hybridization or during the post-hybridization washes, and incorrect Cot-1 to probe ratio. Constant monitoring of these metrics allows the laboratory staff to anticipate potential system failures leading to failed or inaccurate findings.

3. Verification of array results

In addition to the microarray assessments, there are other post-analytical assessments of quality that validate the microarray findings and lead to a complete, quality result used by the clinician for the diagnosis of the patient. Methods for confirmation of array results may include fluorescence in situ hybridization (FISH), multiplex ligation-dependent probe amplification (MLPA), quantitative PCR (Q-PCR) and other PCR techniques. FISH is an established technique that is used to identify numerical and structural chromosome abnormalities by using fluorescently labeled DNA probes to detect the presence or absence of the DNA in the interphase nucleus or in metaphase, the stage of active cell division when the chromosomes are visibly condensed and can be observed in a microscope (Kashork et al., 2010). FISH commonly uses unique-sequence BAC probes; depending on the specific probe used, the resolution of metaphase FISH is ~80-200 kb (Shaffer et al., 2001). In our experience, deletions are easy to visualize by FISH. However, tandem duplications represent a challenge to any laboratory using confirmatory FISH because the duplicated material is not of sufficient distance from the original genomic location to generate a distinct fluorescent signal to allow detection by interphase or metaphase analysis. In some cases the intensity of the signal may be twice as intense on the duplicated homolog compared to the normal homolog, but this is not always the case.

In addition, although microarray analysis can detect DNA copy number changes, it does not identify the provenance of the abnormality; seemingly identical array results may be caused by distinct molecular mechanisms. Complete understanding of the rearrangement so that accurate genetic counseling can be provided requires visualization of the rearrangement, which can be accomplished with FISH. For example, a copy-number gain identified by microarray analysis (Fig. 3) may be a duplication, an insertion, a marker chromosome or an unbalanced translocation (Fig. 4).
Fig. 3. (A-B) Microarray plots from two subjects showing identical single-copy gains of 17 oligonucleotide probes from the terminal long arm of chromosome 3, approximately 189 kb in size (chr3:199,067,024-199,255,755, hg18 coordinates). Probes are ordered on the x-axis according to physical mapping positions, with the most proximal 3q29 probes to the left and the most distal 3q29 probes to the right. Values along the y-axis represent log2 ratios of patient:control signal intensities. Results are visualized using Genoglyphix (Signature Genomics, Spokane, WA).

One disadvantage of FISH visualization of microarray results, particularly when using high-density oligonucleotide arrays, is that high-density oligonucleotide arrays can detect abnormalities well below the size of the smallest FISH probes, which are 100-350 kb in size. However, PCR-based methodologies such as MLPA and Q-PCR can visualize small gains and losses. MLPA targets the region of interest with two oligonucleotide probes, one probe containing a forward primer sequence and the second probe containing the reverse primer sequence (Schouten et al., 2002). The oligonucleotide probes are allowed to hybridize to the DNA followed by a ligation step. If the two probes are adjacent to each other the ligation will combine the two probes into a single probe with a fluorescently tagged forward primer on one end and reverse primer on the other end. The probe is amplified by PCR and only the ligated probe is amplified. The amplified product is dependent on the number of target sites present in the DNA. The forward primer is fluorescently labeled, which allows a comparison of the ratio of the fluorescent intensity between reference sample and the test subject to determine the relative quantity of the probe.

Q-PCR amplifies and simultaneously quantifies the relative amount of DNA when compared against a reference. Two Q-PCR methods have been developed. The first method uses fluorescent dyes that intercalate nonspecifically with the double-stranded DNA which produces fluorescent signals relative the quantity of DNA present (VanGuilder et al., 2008). The ratio is compared against a normal reference to confirm the relative quantity of the sample to the control. The second method uses a fluorescently labeled probe that is targeted to the region of interest. The fluorescently labeled probe has a fluorescent reporter and a
quencher to hide the fluorescent signal until the region is amplified (Udvardi et al., 2008). During each round of the PCR process, the exonuclease activity of the polymerase releases the fluorescent reporter, unquenching the signal and allowing detection.

Additional PCR-based methods such as polymorphic microsatellite analysis have also been used as a confirmatory assay for aCGH. Although these molecular assays can confirm a copy number gain or loss, they cannot reveal the chromosomal rearrangement or mechanism giving rise to the copy number variant (CNV). Each laboratory must determine the appropriate confirmatory assay to meet its needs.

Fig. 4. (A) FISH visualization of the gain shown in Fig. 3A revealed an unbalanced translocation of the 3q29 material to 1p. BAC clone RP11-23M2 from 3q29 is labeled in red, and BAC clones RP11-9A9 from 3q11.2 and RP11-438F14 from 1q44 are labeled in green as controls. The presence of one red signal on one of the chromosome 1 homologues indicates translocation of 3q29 onto 1p (arrow). (B) FISH of the gain shown in Fig. 3B. BAC clone RP11-159K3 from 3q29 is labeled in red, and chromosome 3 centromere probe D3Z1 is labeled in green as a control. The presence of two red signals on metaphase FISH rules out an unbalanced translocation, while the additional finding of three red signals on interphase FISH (inset) suggests a duplication. The patient shown in Fig. 3A may have inherited the unbalanced translocation from a parent with a balanced translocation, whereas the duplication in the patient shown in Fig. 3B may have arisen de novo, although parental testing is necessary to confirm the inheritance.

4. Environmental controls

Over time, specific environmental controls have been developed and implemented within the laboratory to ensure quality microarray diagnostics. For aCGH, ozone is an important environmental factor to control. Ozone is a common pollutant found in the lower atmosphere and is the primary component of smog. Ozone is formed when nitric oxides and volatile organic compounds (VOCs) react in the presence of sunlight (US Environmental Protection Agency [EPA], http://www.epa.gov/air/ozonepollution, last accessed May 9, 2008). Nitric oxides and VOCs are emitted by motor vehicle exhaust, industrial emissions, gasoline vapors, chemical solvents, and natural sources. Consequently, ozone levels are higher in urban and industrial areas, especially during the summer months.
National ozone standards have been established by the EPA to protect public health. The established standard peak ozone level set by the EPA is 80 ppb, which is based on the annual fourth maximum 8-hour average (EPA, http://www.epa.gov/air/ozonepollution, last accessed May 9, 2008). The EPA has also established an air quality index system for monitoring the daily pollution levels. The “good” air quality range is 0 to 60 ppb. The fluorescent dyes commonly used in aCGH are sensitive to ozone levels as low as 5 to 10 ppb (Branham et al., 2007; Fare et al., 2003). Thus, ozone levels considered normal for environmental standards are well above those ranges, demonstrating sensitivity of the dyes.

Ozone has been shown to strongly affect dyes that are commonly used in aCGH, including cyanine 5 (Cy5) and Alexa dyes (Alexa Fluor 647 and Alexa Fluor 5) and, to a lesser extent, cyanine 3 (Cy3) and the Alexa equivalents (Alexa Fluor 555 and Alexa Fluor 3) (Branham et al., 2007; Byerly et al., 2009; Fare et al., 2003). Several studies have identified the post-hybridization washes as the most sensitive period for exposure to ozone (Branham et al., 2007; Byerly et al., 2009; Fare et al., 2003). These studies have demonstrated the difficulty with which laboratories identify the source of ozone-related problems, especially considering the extremely low levels of ozone (5 to 10 ppb) that cause these problems, the duration of exposure (as little as 10 to 30 seconds), and the seasonal emergence of ozone itself. The effects of ozone must be addressed when aCGH is performed, particularly in a clinical diagnostic setting, where it is critical to have consistent high quality and reproducible results. Failure to protect the fluorescent dyes from ozone during the post-hybridization washes will result in considerable negative impact on the array data. The implementation of quality control measures such as ozone reduction and monitoring to ensure high-quality aCGH results is mandatory for any aCGH laboratory. There are many commercially available enclosures and scrubbers designed to protect the dyes during the post-hybridizations washes. Some laboratories have gone as far as developing ozone-free rooms where post-hybridization washes and the subsequent scanning and analysis are performed. The latter is most desirable but may not always be feasible.

In addition to ozone degradation, the dyes are also photosensitive and often must be used in a reduced-light environment. Systems should be implemented to prevent photobleaching of the fluorescent dyes. To mitigate against the effect of photobleaching, the dyes should be protected from the light whenever possible. This can be done by using indirect lighting in the work area, using amber tubes, covering the samples with tin foil or placing the samples in areas with little or no light when not being directly handled.

5. Normalization

Normalization, which aims to separate biologically relevant signal from experimental artifacts, is a crucial step of microarray analysis (Neuvial et al., 2006). Each laboratory must identify a system for normalization. Most microarray vendors offer software with built-in normalization methods optimized for their own platforms. Laboratories can use a normalization package that is developed by the microarray vendor or can develop their own package. One normalization method that is used in the laboratory is the locally weighted polynomial regression (LOESS) (Cleveland, 1979). This normalization applies a spatial correction to correct position-dependent non-uniformity of signals across the array.
Another normalization system used in the laboratory is the Qspline fit normalization (Workman et al., 2002). This normalization compensates for the inherent differences in signal intensities between the two dyes. There are many normalization methodologies available; each laboratory will have to define its method of choice.

6. Automation

One of the key elements for any clinical assay is reproducibility. By replacing manual processes with automation, a laboratory can substantially improve the consistency and reproducibility of its daily operations. In addition, automation can increase throughput, which is often an advantage for a growing laboratory and helps to reduce the dependency on staffing levels. The laboratory protocols that have been successfully automated include DNA isolation, labeling and hybridization, washing and analysis. However, automation can present several challenges. Because it is based on a plate format, if the input materials are flawed or a technical issue occurs (e.g., labeling master mix or an automation failure) it will impact the entire plate, which may consist of 48 or more patients. The consequences of a failure of this magnitude are substantial in terms of cost and the potential loss of the sample.

7. Multiplexing of microarray platforms

As laboratories become more accustomed to using microarrays, the demand for the assay may increase. In addition to automation, multiplexed array formats can help the laboratory satisfy the increased demand. Multiplexed array formats allow for the simultaneous hybridization of 2 to 12 or more samples depending on the probe coverage of the array and the array manufacturer. The multiplex design (Fig. 5) has many positive features including decreased costs and higher throughput. As this technology continues to advance, higher multiplexed formats are likely to be developed. Although increased throughput has its advantages, it also creates challenges for any quality system. The laboratory must ensure that there is no cross contamination between each sub-array, which would affect patient samples. Some manufacturers have included QA/QC features into the development of multiplexed arrays, such as tracking controls that can be spiked into the experiments that identify unique positions on the array such that each position of the array can have a unique tracking control. Unique tracking controls are added to each sample before they are introduced to the array, which allows the laboratory to monitor each sub-array for cross contamination. In the event of cross contamination or leaking between the sub-arrays, the laboratory can determine the sample involved in the cross-contamination event based on the unique tracking controls involved.

8. Validation

As laboratories begin to adopt microarray technology within their facility, they should identify the validation criteria which they are responsible for meeting. The American College of Medical Genetics (ACMG; Shaffer et al., 2007) and other US state guidelines have been developed to ensure laboratories have thoroughly tested and reviewed the capacity and expectation of the assay prior to clinical release. These validations include testing known abnormal specimens to verify the expected outcomes. Section E13.2 of the ACMG
9. Control samples

When defining the control specimen of choice, a laboratory can choose to use same-sex or sex-mismatched controls from the patient. Same-sex controls offer detection of autosomal gains and losses, and complex sex chromosome abnormalities are more easily visualized. Sex-mismatched controls offer the same detection of autosomal gains and losses but can be
more challenging when assessing sex chromosome changes. However, sex-mismatched controls offer the laboratory an internal assessment of hybridization success because of the expected deviations when comparing a male against a female. The deviations that are inherent to sex-mismatched controls are a result of the copy number variation of the X/Y chromosome ratio. When a female genome is compared to a male genome, there is an apparent gain of chromosome X (two copies in the female against the single copy in males) and a loss of chromosome Y (Fig. 6).

Fig. 6. Microarray plots from three different gender pairings: male/male, female/male, and female/female. The probes are ordered on the x-axis according to physical mapping positions starting with chromosome 1 on the left and the X and Y chromosomes on the right. Values along the y-axis represent log₂ ratios of patient:control signal intensities. (A) A same-sex male/male comparison showing identical dosage at the X and Y regions. (B) An opposite-sex female/male comparison showing a gain of the X chromosome (two copies in the female versus a single copy in the male) and a loss of Y (no copies in the female versus one copy in the male). (C) A same-sex female/female comparison showing identical dosage at the X region and no hybridization at the Y region. Results are visualized using Genoglyphix (Signature Genomics, Spokane, WA).

In addition to determining the ideal sex of the control, the laboratory must decide if a single control or pooled controls will be used in the facility. Often, these decisions are made based on available controls. Some institutions have used a consistent male and female control for easy identification of known CNVs and for monitoring the performance of arrays. However, in the absence of consistent candidates for controls, laboratories can create or purchase a pooled DNA control. All variations of these controls must be assessed with any change in the pool so that they will not significantly impact reporting. In addition, CNVs present in the control DNA can be used as a positive indicator of assay performance when using same-sex controls. However, pooled control samples will have diluted CNVs, which may not be apparent on the microarray results or may appear as mosaics or background noise.
10. Conclusion
This chapter highlights many of the quality assurance principles that impact a laboratory setting up or using aCGH. This is not an exhaustive set of challenges to implementation as there may be lab-, region-, environment- and vendor-specific variations. Each laboratory should perform initial quality verification at the time of test development. Post development, a thorough validation must be performed, which may uncover variation that should be controlled prior to launching the clinical assay. A laboratory's role in developing new assays should include an established, documented and maintained quality system that ensures that the test conforms to specified requirements and ultimately leads to accurate results.

11. Acknowledgments
We thank Aaron Theisen (Signature Genomics, Spokane, WA) for his critical review of our manuscript.

12. References


Rapid advances have been made in the last decade in the quality control procedures and techniques, most of the existing books try to cover specific techniques with all of their details. The aim of this book is to demonstrate quality control processes in a variety of areas, ranging from pharmaceutical and medical fields to construction engineering and data quality. A wide range of techniques and procedures have been covered.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
