

Multipurpose Instantaneous Microarray Detection of Acute Encephalitis Causing Viruses and Their Expression Profiles

Desh Deepak Singh · Amita Jain

Received: 28 March 2012 / Accepted: 14 May 2012 / Published online: 7 June 2012
© Springer Science+Business Media, LLC 2012

Abstract Detection of multiple viruses is important for global analysis of gene or protein content and expression, opening up new prospects in terms of molecular and physiological systems for pathogenic diagnosis. Early diagnosis is crucial for disease treatment and control as it reduces inappropriate use of antiviral therapy and focuses surveillance activity. This requires the ability to detect and accurately diagnose infection at or close to the source/outbreak with minimum delay and the need for specific, accessible point-of-care diagnosis able to distinguish causative viruses and their subtypes. None of the available viral diagnostic assays combine a point-of-care format with the complex capability to identify a large range of human and animal viruses. Microarray detection provides a useful, labor-saving tool for detection of multiple viruses with several advantages, such as convenience and prevention of cross-contamination of polymerase chain reaction (PCR) products, which is of foremost importance in such applications. Recently, real-time PCR assays with the ability to confirm the amplification product and quantitate the target concentration have been developed. Furthermore, nucleotide sequence analysis of amplification products has facilitated epidemiological studies of infectious disease outbreaks and monitoring of treatment outcomes for infections, in particular for viruses that mutate at high frequency. This review discusses applications of microarray technology as a potential new tool for detection and identification of acute encephalitis-causing viruses in human serum, plasma, and cell cultures.

Introduction

Microarray technology is a promising technology used to study the expression level of various genes simultaneously or to genotype multiple regions of a genome in a particular cell type of an organism, at a particular time and conditions. This allows comparison of gene expression between normal and infected cells. The method involves placing thousands of gene sequences in known locations on a glass slide called a gene chip. Based on such easy implementation of a huge number of tests for identification of viral agents affecting humans, these tests are helpful for blood screening; they can also be used as a diagnostic tool, albeit imperfect as they provide only an indirect measure of infection but cannot advise clinicians regarding whether the infection is recent or in progress, or regarding response to therapy [52, 102, 104]. Antibody-based testing can also fail to detect current infection, because it usually takes a few days to weeks for the immune system to raise an antibody response to an infectious agent [56, 79, 125]. In the past, tests have been developed based on direct quantification of the infectious agent in a sample taken from the patient [41]. Such tests detect the presence of nucleic acids (the genetic material) of infectious agents in blood or other samples. The most common method is the polymerase chain reaction (PCR), which can detect 100 copies or more of an infectious agent in a single sample [29, 79]. PCR employs an enzymatic reaction to amplify specific nucleic acid sequences of infectious agents that may be present in the sample [59, 114]. There are several problems with this technique [78]: viral agents can mutate very quickly, and PCR primers may therefore not recognize the infectious agent, producing false or weak results [22]. The subsequent process is based on direct hybridization of the infectious agent's nucleic acid to a synthetic nucleic acid probe [23, 62]. The hybridized

D. D. Singh (✉) · A. Jain
Virology Laboratory, Department of Microbiology,
C. S. M. Medical University, Lucknow, UP 226003, India
e-mail: ddsbms@gmail.com

infectious agent is then detected by extension using non-enzymatic methods. The main disadvantage of this technique is required an additional quantity of the patient's sample (minimum 1 ml blood compared with 0.1 ml for PCR) [51, 70, 100]. Virus culture generally results in good specificity, but not all viruses can be cultured, and technical expertise is required to understand the cytopathic effect (CPE) and to read stained preparations. Also, this method is time consuming and labor intensive due to the long incubation period of some viruses, and it is very difficult to culture a variety of cell types at once [104, 110].

During the last two decades and today, viral diseases remain the main cause of mortality in humans. In recent times, the appearance of infectious diseases has become more serious, as represented by new pathogens such as re-emerging viruses causing acquired immune-deficiency syndrome, acute encephalitis syndrome, Hendra, Nipah, severe acute respiratory syndrome (SARS), and avian influenza [27, 28, 73]. Many factors contribute to the emergence of viral infections, potentially including genetic exchange or mutation, adaptation to new hosts or vectors, rapid transport, trade, migration of people, and changing values or lifestyles [8]. The resulting rapid epidemiological changes in the community may lead to both new and old viruses that can emerge and cause outbreaks at unexpected times and locations. The fight against such emergent viral infections requires the development of a comprehensive strategy.

Using conventional techniques to analyze gene expression, researchers are able to survey a relatively small number of genes at once [24]. Microarray profiling offers many potential advances in terms of diagnostic and therapeutic interventions in human diseases because of its unparalleled ability for high-throughput gene expression analysis. This technology provides powerful tools for the scientific community [15], and scientists are using microarray technology to try to understand fundamental aspects of growth and development as well as to explore the underlying genetic causes of many human diseases [48, 49]. However, the limitations of this technique are related in part to issues regarding the various methodologies and experimental designs, as well as difficulties in the interpretation of results. Despite these limitations, microarray technology has been used efficiently in disease diagnosis.

Acute Encephalitis Syndrome

Encephalitis means inflammation of brain matter. More than 100 different infectious pathogens and several toxins have been identified as causative agents of encephalitis, although in many cases no pathogen can be detected. Accurate etiological diagnosis is required to increase the

usefulness of surveillance of acute encephalitis, especially in view of concerns about new and re-emerging infections. Viruses that infect the central nervous system (CNS) may selectively involve the spinal cord (myelitis), brain stem (e.g., rhomb encephalitis), cerebellum (cerebellitis), or cerebrum (encephalitis). Almost every acute viral CNS infection results in meningeal as well as parenchymal inflammation to varying degrees [12]. Fundamental clinical and laboratory findings are mainly similar despite the different causative agents, consisting of fever and headache in addition to distorted cerebral position, frequently accompanied by seizures and central neurologic abnormality. Cerebrospinal fluid (CSF) is abnormal in >90 % of cases, characteristically showing lymphocytic pleocytosis, slightly elevated protein level, and normal glucose [113]. In unusual diseases, such as West Nile virus (WNV) meningoencephalitis or cytomegalovirus (CMV) radiculomyelitis, polymorphonuclear cells rather than lymphocytes could be the principal cell type, providing useful diagnostic evidence. Nevertheless, in spite of these variations, standard CSF study rarely leads to exact identification of the etiologic agent [11].

A clinician may be faced with a patient of any age, at any time of year, with acute-onset fever and altered mental status including symptoms such as confusion, disorientation, coma, and inability to talk, or new-onset seizures (excluding simple febrile seizures) [61]. Other symptoms may include an increase in irritability, abnormal behavior greater than that seen with usual febrile illness, and altered mental status, enabling initial differentiation of encephalitis from noninfectious causes of brain dysfunction (encephalopathy) [11, 41, 97].

A huge number of tests to detect viral pathogens are available, but they have not been validated for identification of divergent viruses using traditional methods of assay gene expression; researchers can therefore identify only a relatively small number of genes at once [24]. Cell cultures using the traditional tube method can be used for isolation and detection of a wide variety of viruses, including unanticipated agents, mixed-culture antiviral susceptibility testing, serotyping, and epidemiologic studies. They offer increased sensitivity over rapid antigen tests, but require a long incubation period for some viruses as well as acquisition and maintenance of a variety of cell culture types in-house. Shell vials with centrifugation can also be used, but reading of pre-CPE stained preparations is both time consuming and labor intensive. Also, unanticipated agents may be missed when pre-CPE staining targets only one or a few viruses, and isolates from fixed/stained vials are not available. Nonculture antigen detection using Immunofluorescence (IF) monoclonal antibodies (mAbs) takes 40 min per sample, generally offering good sensitivity (which varies with the virus detected) and excellent specificity.

CMV antigenemia testing is more sensitive than traditional or shell-vial culture for CMV in blood, but generally not as sensitive as cell culture. Also, it requires expertise for reading and is not useful for all viruses; flavivirus sensitivity is especially poor. Non-IF antigen detection requires 30 min per sample, offering generally good specificity for respiratory syncytial virus (RSV) and influenza A and B viruses. No special technical expertise is required, and results are available very rapidly, enabling application in point-of-care testing. However, it generally offers poor sensitivity compared with cell culture and is currently only available for RSV and influenza A and B viruses. Additional testing of negative samples by cell culture is recommended, the most common method being by PCR, which can detect 100 copies or more of an infectious agent in a single sample [79]. PCR uses an enzymatic reaction to amplify specific nucleic acid sequences from the infectious agent if they are present in the sample [114]. There are several problems with this method [78]. PCR uses specific nucleic acid sequences (primers) from a known sequence of the infectious agent [55, 88]. Therefore, if the infectious agent has not been sequenced, PCR cannot be used. Similarly, if the infectious agent mutates very rapidly, the primers may not recognize the infectious agent and a false-negative test will result [21, 22]. This is a major problem with detection of viruses, which undergo very rapid mutation, especially in response to drug treatment [14, 27]. Since PCR uses an enzymatic reaction, the enzyme can be inhibited by impurities in the patient sample, also leading to false-negative results [7]. In addition, the short sequences of primers are only specific to an infectious agent at definite temperatures, making the test reliant on very strict conditions [117]. The specificity of the primers also makes it difficult to detect more than one agent simultaneously in a single PCR reaction [19, 46]. Multiplex PCR reactions exist, but generally they are not quantitative and can only detect two or a maximum of three agents concurrently [6].

Detection of multiple viruses causing acute encephalitis syndrome (AES) is important for global analysis of gene content and expression, opening prospects for new molecular and physiological systems to control pathogenic diagnosis. Early diagnosis is crucial for disease treatment and control, as it reduces inappropriate use of antiviral therapy and focuses surveillance activity [85, 88]. This requires the ability to detect and accurately diagnose infection at or close to the source/outbreak with minimum delay and the need for specific, accessible point-of-care diagnosis able to distinguish causative viruses and their subtypes. None of the available viral diagnostic assays combines a point-of-care format with the complex capability to identify a large range of viruses causing AES. Biomedical research evolves and advances not only through the compilation of knowledge but also through the

development of new technologies [9]. Microarray detection provides a useful, labor-saving tool for detection of multiple viruses, offering several advantages such as convenience and prevention of cross-contamination. Microarray technology aims to monitor the whole genome on a single chip, providing researchers with a clearer picture of the interactions among thousands of genes simultaneously [25]. This represents a major methodological advance and illustrates how the advent of a new technology can provide powerful tools for research [15].

Microarrays can help answer gradually more multifaceted questions and execute additional complicated experiments. Researchers may be able to infer possible functions of new genes based on similarities in expression profile compared with known genes [32, 115]. Finally, such studies increase the number of accessible gene families, providing a novel guide for synchronized gene expression across gene families as well as for completely new groups of genes. In addition, since discovery of any gene usually interrelates with that of many others, knowledge of how these genes are organized can be improved through such analyses, and exact information of these interrelationships will emerge. Use of microarrays may also accelerate identification of genes involved in the progress of various diseases by enabling scientists to examine a much larger number of genes [104]. This technology will also assist with assessment of gene expression and function at cellular level, illuminating how manifold gene products work mutually to construct substantial chemical responses to both static and changing cellular needs. Scientists may use microarray technology to try to understand fundamental aspects of growth and development as well as to explore the underlying genetic causes of many human diseases [114, 122].

Principles of Microarray Technology

Microarray technology is based on hybridization of two DNA strands formed of complementary nucleic acids linked by tight noncovalent bonds. In the hybridization technique, after washing, nonspecific bonding sequences are removed while only strongly paired strands will remain hybridized. Consequently, fluorescently labeled target sequences designed to bind to a particular probe sequence create a signal which depends on the potency of the hybridization and the number of paired bases, the hybridization conditions (such as temperature), and washing after hybridization. The overall potency of the signal from a spot depends on the amount of test sample bound to the corresponding probes at that location [84]. Hybridized targets can then be detected using one of many reporter-molecule systems. Microarrays use relative quantization in which the intensity of a feature is compared to the intensity of the

same feature under a different condition, and the identity of the feature is known by its position [127].

Types of Microarray

Two types of DNA microarray are widely available for data analysis: complementary DNA (cDNA) arrays and oligonucleotide arrays (Fig. 1).

cDNA Arrays

This type of chip offers a high-density microarray, most often being derived from cDNA (hence the name). Such chips are usually made by robotically spotting onto a large glass surface [76, 77]. Hybridization is carried out using fluorescently labeled messenger RNA (mRNA) corresponding to the cDNA, and hybridized duplexes are identified by the color fluorescence detection method. Thus, they can be used to study gene expression patterns in time and space. If a gene is overexpressed in a particular disease state, then more sample cDNA, as compared with control cDNA, will hybridize to the spot representing that gene. In turn, the red fluorescence of this spot will have greater intensity than its green fluorescence [22]. Once the expression profile of different genes involved in a disease has been distinguished, cDNA derived from a diseased sample taken from any person can be hybridized to

conclude whether the expression profile of the gene from the individual corresponds to the expression profile of a known disease [113].

Oligonucleotide Arrays

In these arrays, the expression strength of a gene is assessed by means of a probe set consisting of 11–20 individual probe pairs. In most recent gene chips, the number of probe pairs has stabilized and is now 10–13 [27, 28, 96]. Every probe pair includes an ideal-match 22–25-mer oligonucleotide probe, which is planned to hybridize exclusively to a unique gene transcript, and a variance probe of the same length, which varies from the ideal match probe by a single base in the core of the sequence [46, 55, 118]. The aim of the variance probe is to calculate unfocused hybridization. Probe position algorithms created by Affymetrix read the signals from each 22–25 oligonucleotide probe set to obtain the particular values as a hybridization blueprint of the 22–25 probes. Each ideal match probe has a corresponding mismatch probe which contain the same 22–25 bases long sequences as the ideal match probe, except for the fact that the middle base (11–13) in the chain is substituted for the complement of the 11–13th base of its consequent ideal match probe. This is meant to give an estimate of non-specific binding, which occurs when m-RNA that not targeted binds to ideal match [60, 94, 95]. In addition it has been exposed to the sign perverse by the ideal match probes not completely unfocused signal. In particular when a transcript is there at far above the ground levels, the label point strength also hybridizes to the variance probe [113, 114]. However, observed expression levels also include variation that is introduced during the process of carrying out the experiment, which could be classified as obscuring variation [57]. Affymetrix has approached the normalization problem by proposing that intensities should be scaled so that each array has the same average value. The distribution of probe intensities is the same across a set of arrays [80, 89, 96]. Propose arametric and non-parametric methods to achieve this. All these approaches depend on the choice of a baseline array. Currently it is up to the researcher to decide on the most significant result for their particular intention for high sensitivity and low variability, or a low false-positive rate [66].

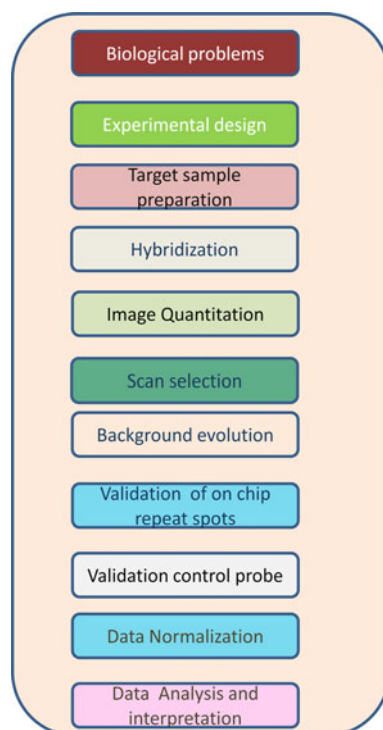


Fig. 1 Microarray assignment workflow

Methods

Briefly, a limited sequence is immobilized on the microarray surface and binds to the target RNA during hybridization. The captured target is labeled with an additional fluorophore-conjugated DNA oligonucleotide (specifically, the label sequence). Positive control spots, in which a

capture sequence hybridizes directly to a complementary label sequence, are included to aid visual analysis. After hybridization and precise washing, the microarray is scanned in a laser-based fluorescence scanner with 5- μm resolution [8].

Sequence Selection and AES Chip Microarray Design

AES virus-specific capture and label sequences may be selected [97]. The possibility of false-positive signals resulting from direct hybridization of label sequences to capture sequences is examined by incubation of the label sequences in the absence of any other nucleic acids at room temperature for 2 h in standard hybridization buffer. Capture sequences found to exhibit cross-reactivity with label sequences are removed from the array layout, along with the corresponding label sequence, and the arrays are reprinted [33]. This process is repeated until the microarray demonstrates no false-positive signals in the absence of viral RNA. The resulting array contains capture sequences and the corresponding label sequences. Each capture sequence is spotted in triplicate, and a single limited sequence with a complementary fluorescence-labeled sequence in solution is used as a positive control on each array. The positive control serves as a direct indication of whether the hybridization conditions are adequate and also as a spatial marker for ease of presentation [40, 53, 60, 67].

Microarray Slide Preparation

The substrate used for these studies comprises an Omni-Grid microarray spotter with solid core pins with 550 μm pitch between spots.

Samples

Viral samples can be purified from whole blood, plasma, serum, throat swabs, cerebral spinal fluid, virus-infected supernatants, and other cell-free body fluids [101].

Chip Processing

Figure 1 shows a schematic of the dispensing protocol. The details of each processing step are described below.

Nucleic Acid Extraction

Nucleic acids may be extracted from clinical samples by using a nucleic acid purification kit, omitting RNase digestion, or manually by triazole methods according to the manufacturer's recommended protocols. RNA is bound to an advanced silica gel membrane under optimal buffering conditions [39, 97]. A simple two-step washing protocol

ensures that PCR inhibitors such as proteins or divalent cations are completely removed, leaving high-quality RNA to be eluted in Milli-Q water [12]. RNA purification processes are generally performed by TRIzol or using a kit method for degradation of the omnipresent RNases and DNA contamination from genomic DNA in the source material [79]. Purification from a viral source carries with it the possible additional challenge of low or varying viral titers. In addition to looking for a kit that can handle low titers as well as overcome the other traditional challenges, users must find a method that is easy to use and that ensures that the typical concentration of extracted RNA meets the requirements of downstream applications [61].

Primer Design, Probes, and Arrays to Confirm the Identity of Viral Strains

The most critical aspect of successful PCR is primer design. All things being equal, a poorly designed primer can result in a PCR reaction that will not work [47, 48]. The primer sequence determines numerous parameters such as the length of the product, its melting temperature, and ultimately the yield [4, 42]. A poorly designed primer can result in little or no product due to nonspecific amplification and/or primer-dimer formation, which can become competitive enough to suppress product formation [13, 19, 111]. This subsection provides rules that should be considered when designing PCR primers. More comprehensive coverage of this subject can be found elsewhere [23, 26]. Several variables must be taken into account when designing PCR primers [126, 127]. Among the most critical are: primer length, melting temperature (T_m), specificity, complementary primer sequences, G/C content and polypurine (A, G) stretches, and 3'-end sequence; each of these critical elements is discussed in turn [121, 127, 129].

Practically all viral strains carry a unique DNA sequence that differentiates it from other strains, and this sequence can be used for probe design [1, 35]. The probe DNA binds specifically to the target gene corresponding to the viral strain prepared from a clinical sample [31, 53, 120]. Using a variety of such probes, various pathogens in a clinical sample can be determined in a single trial. The DNA probe representing the selected target should be designed considering various aspects such as probe length, GC content, molar concentrations, self-hybridization possibilities, and a limit on the number of single-nucleotide repeats [60, 63]. The melting temperature, secondary structure, and binding position in the target DNA are factors that can affect the signal intensity, specificity, and sensitivity [66, 71]. Typical parameters for such DNA probes include the following: minimum length of 35 and maximum up to 40 bases, melting temperature minimum (T_{min}) of 70 $^{\circ}\text{C}$ and

maximum (T_{\max}) of 75 °C, and GC content of 45–50 % [86, 92, 93].

Exceptional to the limits of DNA-DNA hybridization models, determining the array equal to the most favorable DNA-DNA duplex on a microarray is sturdy [69, 116]. Computationally, the best possible arrangement among two DNA sequences should be clearly described in terms of a generalized, accurate distance algorithm [19, 94, 123]. Enormously set, the accurate impassiveness between two sequences correspond to the whole number of insertions, deletions, and substitutions that are needed to transform one sequence into the other. From the standpoint of DNA cross-hybridization, a substitution corresponds to a mismatched pair of nucleotides whereas insertions/deletions correspond to gaps in the duplex DNA-DNA [23, 26]. The lower the number of mismatches and gaps in the alignment, the smaller the edit distance. On the other hand, exact area does not make available enough information with regards to the effectiveness of hybridization [46, 127].

Oligonucleotide probes, are usually much shorter (9–24-mer) and are often customized to integrate an amine or thiol linker that allow covalent attachment of the oligonucleotide to a covered glass face [71, 86, 120, 121]. The alterations of probes put in substantial expense to array construction [86]. For illustration, unchanged oligonucleotide probes are able to be balanced in alkaline buffer (pH 12) and put down instinctively on to acid-washed slides; they stick on to the slide face via hydrogen bonds and electrostatic attraction and are then accessible to form duplexes with corresponding strands of target DNA [93]. This addition scheme is vigorous transversely a broad range of temperature (4–95 °C) pH (1–10), and ionic buffers (e.g., 0–4 M NaCl) [116]. The compassion of recognition can be improved if acid-washed slides are coated with epoxy-silane before probe deposition [19, 112]. Underprivileged eminence slides have rough float up and may well auto-fluorescence, thereby producing background signal that interferes with spot finding and quantification [23, 124]. Auto-fluorescence can be predominantly challenging at what time signal strength is short, since is the case with expression arrays. To keep away from these harms, premium ready slides are available [19, 66].

RNA Quantification

The RNA concentration can be used to determine the amount of sample lost while cleaning using a kit. Transcribed viral nucleic acid may be purified by using a kit or manually, and quantified by measurement of the optical absorbance at 260 nm [50, 61, 97]. The concentration of RNA in the crude transcription product is calculated beforehand [128, 129].

Many post-PCR applications require removal of unincorporated primers, primer-dimers, and other reaction components from the PCR product. Traditional purification methods such as ethanol precipitation are difficult to automate. The clean-up system can be automated to purify 96 or 384 reactions in less time. PCR clean-up is especially amenable to automation because no vacuum or centrifugation steps are required, in contrast to many filter-based methods [5, 11, 26, 31].

Sample Amplification and Labeling

Reverse-transcription (RT) reactions may be prepared using a reasonable quantity (40–200 ng) of total RNA or random hexamers/gene-specific primers (Fig. 2). The mixture is heated and immediately cooled on ice before addition of dithiothreitol, dNTPs/aa-dUTP (a mixture of dGTP, dATP, dCTP, dTTP, and aminoallyl-dUTP), Superscript III reverse transcriptase, and first-strand buffer. The mixture can then be incubated, followed by the application of specific reaction conditions [3, 20, 24].

cDNA from cells under two different conditions is extracted and labeled with two different fluorescent labels, for example, a green dye (cyanine 3) for cells for condition 1 and a red dye (cyanine 5) for cells at condition 2 (to be more accurate, the labeling is typically done by synthesizing single-strand DNAs that are complementary to the extracted mRNA using an enzyme called reverse

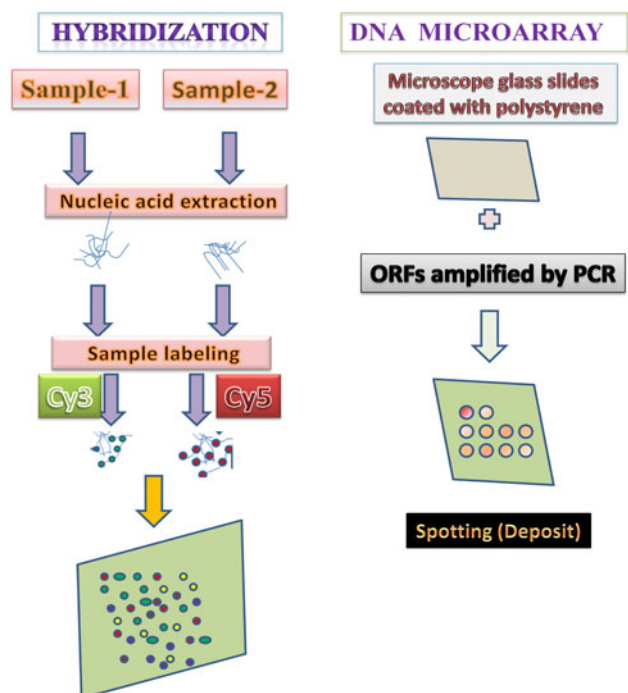


Fig. 2 Principle of microarray assay for gene expression. *ORFs* open reading frames

transcriptase) [34, 41, 48]. Both extracts are washed over the microarray. Labeled gene products from the extracts hybridize to their complementary sequences at the spots due to their preferential binding; complementary single-strand nucleic acid sequences tend to attract each other, and the longer the complementary parts, the stronger the attraction [61, 75, 79].

Hybridization

Samples for determination probably include DNA from a number of species, at diverse concentrations depending on the relative species abundances. A microarray protocol for quantitative evaluation of species diversity has not yet been developed. However, a universal significance of microarray hybridization in area of genomic transcription levels take on vigorous hybridization of two differentially labeled samples of cDNA to the same microarray slide [82, 114]. The main principle behind microarray technology is hybridization between two DNA strands, i.e., the property of complementary nucleic acid sequences to specifically pair with each other by forming hydrogen bonds between complementary nucleotide base pairs [79]. A greater number of complementary base pairs in a nucleotide sequence means tighter noncovalent bonding between the two strands. After washing off nonspecific bonding sequences, only strongly paired strands will remain hybridized [82, 83]. So, fluorescently labeled target sequences that bind to a probe sequence generate a signal that depends on the strength of the hybridization as determined by the number of paired bases, hybridization conditions (such as temperature), and washing after hybridization. The total strength of the signal from a spot (feature) depends on the amount of target sample bound to the probes present at that spot [114, 115]. Microarrays use relative quantization in which the intensity of a feature is compared with the intensity of the same feature under a different condition, while the identity of the feature is determined by its position [43]. An alternative to microarrays is serial analysis of gene expression, where the transcriptome is sequenced, allowing absolute measurements [54]. This allows quantitative discrimination between transcription levels in two samples when the relative variation is greater than about twofold [2, 20]. Using the same technique, it should be possible to quantitate >2-fold differences in species abundance between two samples, allowing rapid and sensitive examination of differences in species abundance [8, 10, 22].

Microarray Imaging

Microarray images are acquired by a laser scanner that executes a regional scan of the slide and creates a digital map or image for each dye, showing the fluorescent

intensities for each pixel [24, 45]. For a particular microarray examination, the scanner generates two 16-bit tagged image file (TIFF) layouts, one for each fluorescent dye [57, 125]. Dissimilar dyes absorb and emit light at different wavelengths [115, 116]. To quantify the amount of each of the two fluorescent dyes at each spot, the scanner applies light excitation at the different wavelengths and measures at the different emission wavelengths [44, 48]. The dyes used usually are Cy3 and Cy5, with emission in the ranges of 510–550 nm and 630–660 nm, respectively [51, 54]. These dyes enable measurement of the amount of sample bound to a spot based on the level of fluorescence emitted when excited by the laser. If the RNA from the sample in condition 1 is abundant, the spot will be green, whereas if the RNA from the sample in condition 2 is abundant, the spot will be red. If both are equal, the spot will be yellow, while if neither is present it will not fluoresce and will appear black [58, 75]. Thus, from the fluorescence intensities and colors for each spot, the relative expression levels of the genes in both samples can be estimated.

A number of conditions [e.g., scan rate, laser power, photomultiplier tube (PMT) voltage] can be adjusted by the user at the time of scanning [24, 80, 103]. New, higher-power lasers provide additional photons for excitation and generate higher signal and noise [58, 75]. Higher PMT voltage results in greater amplification of photons to electrons and generates more detector noise and signal [81, 129].

It is preferable to employ higher laser fluence rather than higher PMT voltage, as this stimulates more signal photons rather than producing more signal per photon [87, 127]. However, elevated laser power can destroy hybridized samples through photobleaching, and depending on the number of scans to be performed on each sample, the laser power needs to be adjusted accordingly [63, 64].

Image quantification limitations (e.g., adaptive, fixed sphere, spot distance) should be carefully assessed and determined for each project as a whole, also depending on the array design, slide type, and spot morphology [8, 89, 106]. It should be noted that the image quantification method should be identical for all slides in a project, whereas the image acquisition parameters, for instance, laser power and/or photomultiplier parameters, can be optimized from slide to slide [9, 10].

Commonly designed for each slide a secure alignment of laser power and PMT is chosen though scanning [23]. The choice of these two parameters is finally determined so that almost all expression on the chip can be captured [37, 90]. However, it has been observed that not all genes spotted onto a chip can be measured accurately for a single scanner setting [44], as there may be genes with expression of 50,000 or more to genes with expression as low as 200 or even less [78]. The choice of these two parameters is finally determined so that almost all expression on the chip

can be captured [105]. However, it has been observed that not all the genes spotted on the chip can be measured accurately for a single scanner setting [107]. There are genes with expression ranging from 50,000 or more to genes with expression as low as 200 or even less [115]. Such a wide range of expression is impossible to capture accurately in a single scan with a fixed setting [116]. However, a single scan with particular PMT and laser settings is certainly suitable for most, but certainly not all, of the intensity range [8–10]. Thus, there is a need to capture various ranges of gene expression values and then combine the information from all scans before further analysis is carried out [37, 44, 78, 107].

Data Analysis

The probes represent (partial) genomic sequences of a gene, positioned or fixed onto a glass slide, whereas the aim is the surveys of gene expression, in a highly parallel and comprehensive manner [16, 18, 128]. Probe variants also provide discrete, dissimilar nucleotide sequences corresponding to the same gene [10]. Signal potency is commonly sequence dependent. For this reason, averaging of the indicated intensities is not appropriate, and probe variants should be investigated separately in the final data evaluation steps. Probe copies exhibit similar response for various instances on the chip [23]. In theory, they should exhibit the same expression, being included for signal amplification. Confidence in the reliability of gene investigation is based on authentication research [17, 25]. Probe replicates are mainly developed based on two important approaches, as described below [74].

Our preferred option to specify a sufficient number of biological samples per condition is to determine the core value for each replicated set of genetic material probes on a chip and use this as the “true value” for that gene probe [66]. The median provides a robust measure and its determination inherently ignores outlier values within a replicate set, in contrast to the arithmetic mean [30, 58].

An alternative is to consider all replica values in the considered dataset, e.g., to apply analysis of variance (ANOVA)-type methods afterwards [53, 54]. Investigational inaccuracy is then pooled with genetic variation. Microarray data may require further processing aimed at reducing the dimensionality of the data to aid comprehension and more focused analysis for each gene [123, 124]. Other methods permit analysis of data consisting of a low number of biological or technical replicates [24, 37].

Data Normalization

Normalization is an important process in the investigation of DNA microarrays when evaluating data from diverse

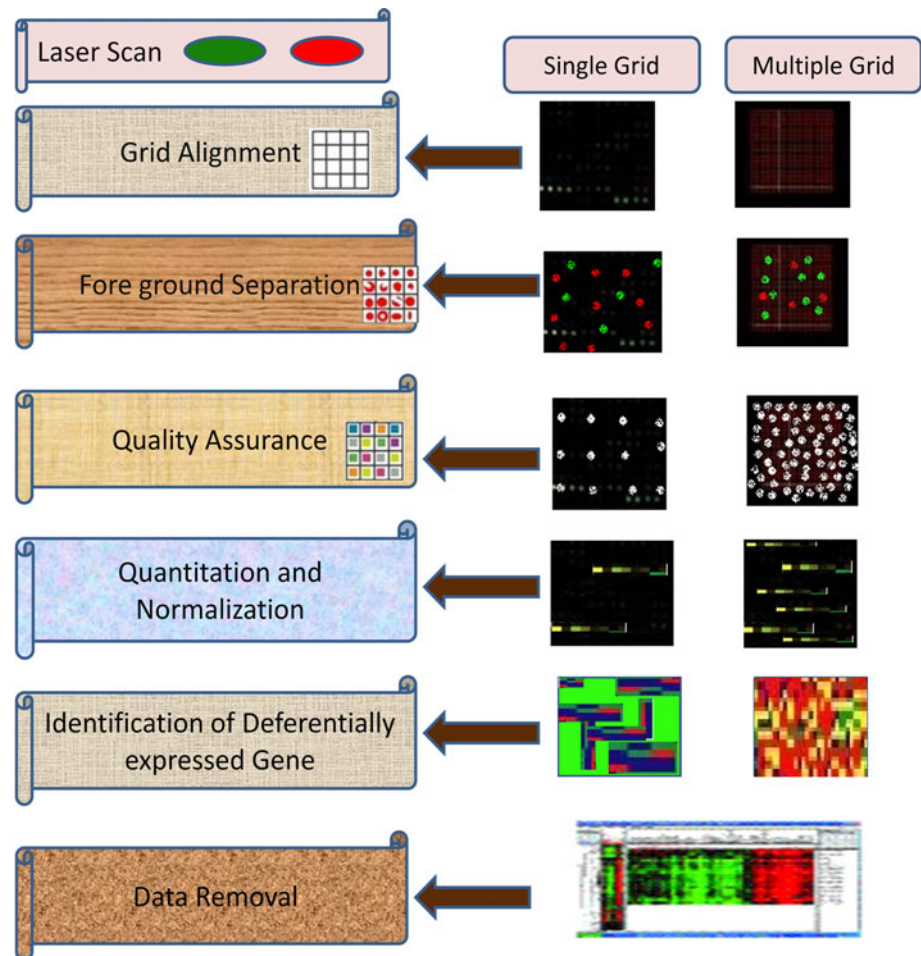
arrays or color controls [36, 38]. Analysis of microarrays can be scientifically influenced by various effects such as nucleic acid extraction, cDNA preparation, sample labeling, assimilation, imaging, and spot detection. In addition, there are effects that are exceptional to individual arrays, such as special effects of various probes, spotting effects, region effects, and pin effects. Normalization endeavors to compensate for such effects through use of internal controls [51–53].

The statistical analysis begins with the scanning file itself. Various parameters for the distribution of the pixels in a particular spot are given like mean, mode, median, standard deviation and the main part correct reserved to elucidate the potency assessment of a précised spot in both channels [10, 109]. The scanned files provide the position values for the center intensity of both channels and their background [20]. The background noise measures the intensity of the mRNA for the slide even if no material was spotted [23]. Using all this information go to the next step, those spots with disgusting excellence that should not be used for advance analysis [34, 37]. The different incorporation properties of the dyes and their different physical characteristics make this the most important source of systematic error in two-color microarrays [38, 42]. The difference in overall intensity between different arrays can be due to real biological variation from one condition to another or just experimental noise [84]. The different expression level of a particular gene in a particular array can also be due to biological variability of the gene or to some noise [37]. If the overall intensity of the hybridized samples is different, this may also be due to some experimental error or to real biological activity [48]. This factor is therefore important in the choice of the within-array normalization method. Besides these factors, it must also be considered that some part of the probe will attach to the slide even when there is no spotted material, thereby contributing to the foreground intensity [44]. However, a reliable estimator for background intensity has not yet been provided. The data should be normalized sequentially to eliminate all nonbiological variation introduced by the investigational procedure and to facilitate assessment of the intensity values contained across slides [24].

Analysis

A brief outline of the initial data analysis of the absolute expression levels within an experiment is shown in Fig. 3. For microarray projects that are designed to study defined gene pathways and interactions, a maximum of annotation and statistical reliability is required [40, 65]. We suggest that the minimum result set for each gene should include: fold-changes of mean expression level per condition and *P* values from significance testing [24].

Fig. 3 Schematic flow diagram of bioinformatics in microarray development



Data Visualization

Data points for biological chip replicates are usually the mean values of each gene for a given condition. Limited subsets of interesting genes can obviously be plotted by means of simple vertical bar charts of $\log(\text{ratio})$ or absolute values [53]. This can be complemented by custom, project-dependent graphs, often integrating annotation data with gene expression results. Expression profiles of genes within one condition or of each gene across a number of conditions can be subjected to cluster analysis [99].

Condition Means and Confidence Intervals

These parameters are required to present the expression level of a gene, and enable better interpretation of fold-changes or variation. In most dual-dye experiments, confidence intervals can be calculated for the fold-change itself [48, 51, 53].

Calculation of 95 % confidence intervals (CI) for means can be done in various ways. If there are sufficient numbers of observations for each condition, the common formula

based on a t -distribution can be used [91]. Broadly speaking, the values making up both groups of observations are randomly reassigned to each group a large number of times, with the desired statistic (mean, CI95, t , etc. [97, 98]) being calculated for each such similar run [107]. The accumulated set of newly generated statistics is then used to “estimate” the corresponding parameter for the original data [125].

Significance Tests

A simple two-sample t test or Welch t test is often the first tool of choice for statistical inference. Adjustment for multiple testing changes the obtained P values but not the order of sorted significance values [108]. If two conditions can be assumed to be dependent (e.g., cell lines) then paired t tests can increase the statistical power [119]. Nonparametric testing (e.g., using the Mann–Whitney U test) is an alternative with less power that nonetheless works better under the assumption that the underlying distributions are nonsymmetrical. However, for the very small numbers of observations (i.e., 5–7) typically available in microarray

studies, the resulting P values can be less useful as a filtering tool [72]. Statistical power can generally be improved by employing bootstrap versions of significance tests. For most microarray studies, P values resulting from significance testing must be interpreted with care in cases with few biological replicates per group [8, 10, 98].

ANOVA-type methods [53] are somewhat more involved, being appropriate where there is more than one experimental factor under investigation (e.g., treatment and dose, or biological replicates and hybridization replicates) [84]. It is important to note that the expression of individual genes of interest is usually backed up by verification using other techniques such as RT-PCR, in situ hybridization, and Northern blotting [31, 48, 66].

Explorative Methods

Explorative methods can be used to identify genes or samples with similar expression profiles, indicating coregulation or sample type, respectively [84, 91]. If coregulation or time effects are of interest, (graphical) principal component analysis can be used to assess the number of clusters that may be contained within the data, which can then serve as the input parameter for the number of expected clusters in a K -means or self-organizing map (SOM) clustering approach [127]. Because of the nature of explorative methods, we recommend using several combinations of algorithms and distance measures (SOM and hierarchical, both with Pearson correlation and Euclidean distance as a minimum) to highlight different features in the data [68].

Conclusions

Available techniques to screen a broad range of viruses are intrinsically biased and thereby constrained to detection of a restricted number of candidate viruses. To overcome this difficulty, an approach to widen viral recognition based on a combination of viral genomics and long-oligonucleotide microarray technology is required. To accomplish this objective, extremely conserved nucleotide sequences within a viral family can be selected for presentation on the microarray. By using these most conserved sequences, it is hoped to maximize the possibility of detecting all members of each viral family, as well as unsequenced, unknown, or newly evolved family members. A secondary, but corresponding, ambition is to take advantage of the high resolution of microarray hybridization to distinguish among viral subtypes, which is a complex and difficult task with conventional methods.

References

1. Afrakhte M, Schultheiss TM (2004) Construction and analysis of a subtracted library and microarray of cDNAs expressed specifically in chicken heart progenitor cells. *Dev Dynam* 230: 290–298
2. Alizadeh A, Eisen M, Davis RE, Ma C, Sabet H, Tran T, Powell JI, Yang L (1999) The lymphochip: a specialized cDNA microarray for the genomic-scale analysis of gene expression in normal and malignant lymphocytes. *Cold Spring Harb Symp Quant Biol* 64:71–78
3. Alizadeh AA, Eisen MB, Davis R, Ma EC, Lossos IS, Rosenwald A, Boldrick JC, Sabet H et al (2000) Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. *Nature* 403:503–511
4. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–410
5. Anonymous (March 31, 2006) Posting date. Biological diagnostics manufacturing: request for information. U.S. Department of Health and Human Services Solicitation 2006-Q-08478. <http://www.fbo.gov/spg/HHS/CDCP/PGOA/2006%2DQ%2D08478/SynopsisR.html>. Accessed 2 Dec 2011
6. Bellau-Pujol S, Vabret A, Legrand L, Dina J, Gouarin S, Petitjean-Lecherbonnier J, Pozzetto B, Ginevra C, Freymuth F (2005) Development of three multiplex RT-PCR assays for the detection of 12 respiratory RNA viruses. *J Virol Methods* 126: 53–63
7. Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J R Stat Soc B* 57:289–300
8. Berthet N, Leclercq I, Dublineau A, Shigematsu S, Burguiere AM et al (2010) High-density resequencing DNA microarrays in public health emergencies. *Nat Biotechnol* 28:25–27
9. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC et al (2001) Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. *Nat Genet* 29:365–371
10. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach Ansorge W et al (2001) Minimum information about a microarray experiment (MIAME)—toward standards for microarray data. *Nat Genet* 29:365–371
11. Briese T, Palacios G, Kokoris M, Jabado O, Liu Z, Renwick N, Kapoor V, Casas I, Pozo F, Limberger R, Perez-Brena P, Ju J, Lipkin WI (2005) Diagnostic system for rapid and sensitive differential detection of pathogens. *Emerg Infect Dis* 11: 310–313
12. Briese T, Paweska JT, McMullan LK, Hutchison SK, Street C, Palacios G, Khristova ML, Weyer J, Swanepoel R, Egholm M, Nichol ST, Lipkin WI (2009) Genetic detection and characterization of Lujo virus, a new hemorrhagic fever-associated arenavirus from southern Africa. *PLoS Pathog* 5:e1000455
13. Brownie J, Shawcross S, Theaker J, Whitcombe D, Ferrie R, Newton C, Little S (1997) The elimination of primer-dimer accumulation in PCR. *Nucleic Acids Res* 25:3235–3241
14. Brust S, Duttman H, Feldner J, Gurtler L, Thorstenson R, Simon F (2000) Shortening of the diagnostic window with a new combined HIV p24 antigen and anti-HIV-1/2/O screening test. *J Virol Methods* 90:153–165
15. Bryant PA, Venter D, Robins-Browne R, Curtis N (2004) Chips with everything: DNA microarrays in infectious diseases. *Lancet Infect Dis* 4:100–111
16. Buckley MJ (2000) Spot user's guide. CSIRO Mathematical and Information Sciences, Sydney, Australia. <http://www.cmis.csiro.au/iap/Spot/spotmanual.htm>. Accessed 5 Jan 2012

17. Call DR, Bakko MK, Krug MJ, Roberts MC (2003) Identifying antimicrobial resistance genes with DNA microarrays. *Antimicrob Agents Chemother* 47:3290–3295
18. Call DR, Borucki MK, Loge FJ (2003) Detection of bacterial pathogens in environmental samples using DNA microarrays. *J Microbiol Methods* 53:235–243
19. Callow MJ, Dudoit S, Gong EL, Speed TP, Rubin EM (2000) Microarray detection and characterization of Plum pox virus: serological methods. *EPPO Bull* 36:254–261
20. Candotti D, Richetin A, Cant B, Temple J, Sims C, Reeves I, Barbara JAJ, Allain JP (2003) Evaluation of a transcription-mediated amplification-based HCV and HIV-1 RNA duplex assay for screening individual blood donations: a comparison with a minipool testing system. *Transfusion* 43:215–225
21. Chen SH, Lin CY, Cho CS, Lo CZ, Hsiung CA (2003) Primer design assistant (PDA): a web-based primer design tool. *Nucleic Acids Res* 31:3751–3754
22. Chiu CY, Greninger AL, Chen EC, Haggerty TD, Parsonnet J, Delwart E, Derisi JL, Ganem D (2010) Cultivation and serological characterization of a human Theiler's-like cardiovirus associated with diarrheal disease. *J Virol* 84:4407–4414
23. Chizhikov V, Rasooly A, Chumakov K, Levy DD (2001) Microarray analysis of microbial virulence factors. *Appl Environ Microbiol* 67:3258–3263
24. Chizhikov V, Wagner M, Ivshina A, Hoshino Y, Kapikian AZ, Chumakov K (2002) Detection and genotyping of human group rotaviruses by oligonucleotide microarray hybridization. *J Clin Microbiol* 40:2398–2407
25. Chuaqui RF, Bonner RF, Best CJM, Gillespie JW, Flaig MJ, Hewitt SM, Phillips JL, Krizman DB, Tangrea MA, Ahram M et al (2001) Post-analysis follow-up and validation of microarray experiments. *Nat Genet* 32:509–514
26. Churchill G (2002) Fundamentals of experimental design for cDNA microarrays. *Nat Genetics Suppl* 32:490–495
27. Coiras MT, Aguilar JC, Garcia ML, Casas I, Perez-Brena P (2004) Simultaneous detection of fourteen respiratory viruses in clinical specimens by two multiplex reverse transcription nested-PCR assays. *J Med Virol* 72:484–495
28. Coiras MT, Lopez-Huertas MR, Lopez-Campos G, Aguilar JC, Perez-Brena P (2005) Oligonucleotide array for simultaneous detection of respiratory viruses using a reverse-line blot hybridization assay. *J Med Virol* 76:256–264
29. Coiras MT, Lopez-Huertas G, Lopez-Campos JC, Aguilar P, Perez B (2005) Oligonucleotide array for simultaneous detection of respiratory viruses using a reverse-line blot hybridization assay. *J Med Virol* 76:256–264
30. Colantuoni C, Henry G, Zeger S, Pevsner J (2002) Local means normalization of microarray element signal intensities across an array surface: quality control and correction of spatially systematic artifacts. *Biotechniques* 32:1316–1320
31. Conejero-Goldberg C, Wang E, Yi C, Goldberg TE, Jones-Brando L, Marincola FM, Webster MJ, Torrey EF (2005) Infectious pathogen detection arrays: viral detection in cell lines and postmortem brain tissue. *Biotechniques* 39:741–751
32. Corless CE, Guiver M, Borrow R, Edwards-Jones V, Fox AJ, Kaczmarek EB (2001) Simultaneous detection of *Neisseria meningitidis*, *Haemophilus influenzae*, and *Streptococcus pneumoniae* in suspected cases of meningitis and septicemia using real-time PCR. *J Clin Microbiol* 39:1553–1558
33. Dacheux L, Berthet N, Dissard G, Holmes EC, Delmas O et al (2010) Application of broad-spectrum resequencing microarray for genotyping rhabdoviruses. *J Virol* 84:9557–9574
34. Davignon L, Walter EA, Mueller KM, Barrozo CP, Stenger DA, Lin B (2005) Use of resequencing oligonucleotide microarrays for identification of *Streptococcus pyogenes* and associated antibiotic resistance determinants. *J Clin Microbiol* 43:5690–5695
35. Dong F, Allawi HT, Anderson T, Neri BP, Lyamichev VI (2001) Secondary structure prediction and structure-specific sequence analysis of single-stranded DNA. *Nucleic Acids Res* 29:3248–3257
36. Dudoit S, Yang YH, Callow MJ, Speed TP (2000) Statistical methods for identifying differentially expressed genes in replicated cDNA microarray experiments. *Science* 283:83–87
37. Dunbar SA (2006) Applications of Luminex xMAP technology for rapid, high-throughput multiplexed nucleic acid detection. *Clin Chim Acta* 363:71–82
38. Eisen MB, Brown PO (1999) DNA arrays for analysis of gene expression. *Methods Enzymol* 303:179–205
39. Epstein CB, Hale W IV, Butow RA (2001) Numerical methods for handling uncertainty in microarray data: an example analyzing perturbed mitochondrial function in yeast. *Methods Cell Biol* 65:439–452
40. Feezor R, Baker HV, Mindrinos M, Hayden D, Tannahill CL, Brownstein BH, Fay A, MacMillan S, Laramie J, Xiao W, Moldawer LL, Cobb JP, Laudanski K, Miller-Graziano CL, Maier RV, Schoenfeld D, Davis RW, Tompkins RG (2004) Whole blood and leukocyte RNA isolation for gene expression analyses. *Physiol Genomics* 19:247–254
41. Fejzo M, Slamon DJ (2001) Frozen tumor tissue microarray technology for analysis of tumor RNA, DNA, and proteins. *Am J Pathol* 159:1645–1650
42. Field LA, Jordan RM, Hadix JA, Dunn MA, Shriver CD, Ellsworth RE, Ellsworth DL (2007) Functional identity of genes detectable in expression profiling assays following globin mRNA reduction of peripheral blood samples. *Clin Biochem* 40:499–502
43. Gentit P (2006) Detection of plum pox virus: biological methods. *EPPO Bull* 36:251–253
44. Ghedin E, Pumfery A, de la Fuente C, Miller Yao K, Lacoste N, Quackenbush V, Jacobson JS, Kashanchi F (2004) Use of a multi-virus array for the study of human pathogens: gene expression studies and ChIP-chip analysis. *Retrovirology* 1:10
45. Glasbey CA, Ghazal P (2003) Combinatorial image analysis of DNA microarray features. *Bioinformatics* 19:194–203
46. Goodman N (2002) Biological data becomes computer literate: new advances in bioinformatics. *Curr Opin Biotechnol* 13:68–71
47. Gorelenkov V, Antipov A, Lejnine S, Daraselia N, Yuryev (2001) A set of novel tools for PCR primer design. *Biotechniques* 31:1326–1330
48. Grace MB, McLeland CB, Gagliardi SJ, Smith JM, Jackson WE III, Blakely WF (2003) Development and assessment of a quantitative reverse transcription-PCR assay for simultaneous measurement of four amplicons. *Clin Chem* 49:1467–1475
49. Greenberg SA (2001) DNA microarray gene expression analysis technology and its application to neurological disorders. *Neurology* 57:755–761
50. Haan CA, Stadler K, Godeke GJ, Bosch BJ, Rottier PJ (2004) Cleavage inhibition of the murine coronavirus spike protein by a furinlike enzyme affects cell-cell but not virus-cell fusion. *J Virol* 78(11):6048–6054
51. Hardegger D, Nadal D, Bossart W, Altwegg M, Dutly F (2000) Rapid detection of *Mycoplasma pneumoniae* in clinical samples by real-time PCR. *J Microbiol Methods* 41:45–51
52. Hegde P, Qi R, Abernathy K, Gay C, Dharap S, Gaspard R, Hughes JE, Snesrud E, Lee N, Quackenbush J (2000) A concise guide to cDNA microarray analysis. *Biotechniques* 29:548–562
53. Ji H, Masse N, Tyler S, Liang B, Li Y, Merks H, Graham M, Sandstrom P, Brooks J (2010) HIV drug resistance surveillance using pooled pyrosequencing. *PLoS ONE* 5:e9263

54. Kerr K, Churchill G (2001) Statistical design and the analysis of gene expression microarray data. *Genet Res* 77:123–128
55. Kim H, Zhao B, Snesrud EC, Haas BJ, Town CD, Quackenbush J (2002) Use of RNA and genomic DNA references for inferred comparisons in DNA microarray analyses. *Biotechniques* 33: 924–930
56. Lai MM, Brayton PR, Armen RC, Patton CD, Pugh C, Stohlman SA (1981) Mouse hepatitis virus A59: mRNA structure and genetic localization of the sequence divergence from hepatotropic strain MHV-3. *J Virol* 39:823–834
57. Landraud L, Brisse S (2010) Enterobacteriaceae. In: Cohen J, Powderly WG (eds) *Infectious diseases*, 3rd edn. Mosby Elsevier, London, pp 1690–1703
58. Leong WF, Tan H, Ooi C, Koh EE, Chow DR (2005) Microarray and real-time RT-PCR analyses of differential human gene expression patterns induced by severe acute respiratory syndrome (SARS) coronavirus infection of Vero cells. *Microbes Infect* 7:248–259
59. Leski TA, Lin B, Malanoski AP, Wang Z, Long NC (2009) Testing and validation of high density resequencing microarray for broad range biothreat agents detection. *PLoS ONE* 4:e6569
60. Li J, Chen S, Evans DH (2001) Typing and subtyping influenza virus using DNA microarrays and multiplex reverse transcriptase PCR. *J Clin Microbiol* 39:696–704
61. Li X, He Z, Zhou J (2005) Selection of optimal oligonucleotide probes for microarrays using multiple criteria, global alignment and parameter estimation. *Nucleic Acids Res* 33:6114–6123
62. Lin B, Blaney KM, Malanoski AP, Ligler AG, Schnur JM, Metzgar D, Russell KL, Stenger DA (2006) Rapid testing for over 20 respiratory pathogens simultaneously in clinical samples using resequencing arrays. Naval Research Laboratory, Washington, DC
63. Lin BZ, Wang GJ, Vora JA, Thornton JM, Schnur DC, Thach KM, Blaney AG, Ligler AP, Malanoski, Santiago J, Walter EA, Agan BK, Metzgar D, Seto D, Daum LT, Kruzlock R, Rowley RK, Hanson EH, Tibbetts C, Stenger DA (2006) Broad-spectrum respiratory tract pathogen identification using resequencing DNA microarrays. *Genome Res* 16:527–535
64. Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ (1999) High density synthetic oligonucleotide arrays. *Nat Genet* 21:20–24
65. Lönnstedt I, Speed TP (2002) Replicated microarray data. *Stat Sinica* 12:31–46
66. Madhi SA, Klugman KP (2004) A role for *Streptococcus pneumoniae* in virus-associated pneumonia. *Nat Med* 10:811–813
67. Malanoski AP, Lin B, Wang Z, Schnur JM, Stenger DA (2006) Automated identification of multiple micro-organisms from resequencing DNA microarrays. *Nucleic Acids Res* 34: 5300–5311
68. Martinez MJ, Aragon AD, Rodriguez AL, Weber JA, Timlin MB, Sinclair DM, Haaland DM, Werner-Washburne M (2003) Identification and removal of contaminating fluorescence from commercial and in-house printed DNA microarrays. *Nucleic Acids Res*. 31:e18
69. McDonough EA, Barrozo CP, Russell KL, Metzgar D (2005) A multiplex PCR for detection of *Mycoplasma pneumoniae*, *Chlamydia pneumoniae*, *Legionella pneumophila*, and *Bordetella pertussis* in clinical specimens. *Mol Cell Probes* 19:314–322
70. Mehlmann M, Townsend MB, Stears RL, Kuchta RD, Rowlen KL (2005) Optimization of fragmentation conditions for microarray analysis of viral RNA. *Anal Biochem* 347:316–323
71. Mehlmann M, Dawson ED, Townsend MB, Smagala JA, Moore CM, Smith CB, Cox NJ, Kuchta RD, Rowlen KL (2006) Robust sequence selection method used to develop the FluChip diagnostic microarray for influenza virus. *J Clin Microbiol* 44:2857–2862
72. Merino EJ, Wilkinson KA, Coughlan JL, Weeks KM (2005) RNA structure analysis at single nucleotide resolution by selective 2-hydroxyl acylation and primer extension (SHAPE). *J Am Chem Soc* 127:4223–4231
73. Metzgar D, Myers CA, Russell KL, Faix D, Blair PJ et al (2010) Single assay for simultaneous detection and differential identification of human and avian influenza virus types, subtypes, and emergent variants. *PLoS ONE* 5:e8995
74. MGED—Microarray gene expression data society home. <http://www.mged.org>. Accessed 6 Jan 2012
75. Mills J, Roth KA, Cagan RL, Gordon JI (2001) DNA microarrays and beyond: completing the journey from tissue to cell. *Nat Cell Biol* 3(E):175–178
76. Molling P, Jacobsson S, Backman A, Olcen P (2002) Direct and rapid identification and genogrouping of meningococci and *porA* amplification by LightCycler PCR. *J Clin Microbiol* 40: 4531–4535
77. Molling P, Jacobsson S, Backman A, Olcen P (2002) Direct and rapid identification and genogrouping of meningococci and *porA* amplification by LightCycler PCR. *J Clin Microbiol* 40: 4531–4535
78. Muller RA, Ditzen K, Hille M, Stichling R, Ehrlich T, Illmer G, Ehninger J (2008) Detection of herpesvirus and adenovirus co-infections with diagnostic DNA-microarrays. *J Virol Methods* 155:161–166
79. Murphy D (2002) Gene expression studies using microarrays: principles, problems, and prospects. *Adv Physiol Educ* 26: 256–270
80. Natarajan K, Shepard LA, Chodosh J (2002) The use of DNA array technology in studies of ocular viral pathogenesis. *DNA Cell Biol* 21:483–490
81. Pellois X, Zhou O, Srivannavit T, Zhou E, Gulari X (2002) Gao, individually addressable parallel peptide synthesis on microchips. *Nat Biotechnol* 20:922–926
82. Puppe W, Weigl JA, Aron G, Grondahl B, Schmitt HJ, Niesters HG, Groen J (2004) Evaluation of a multiplex reverse transcriptase PCR ELISA for the detection of nine respiratory tract pathogens. *J Clin Virol* 30:165–174
83. Quackenbush J (2002) Microarray data normalization and transformation. *Nat Genet* 32:496–501
84. Raj A, Oudenaarden A (2009) Single-molecule approaches to stochastic gene expression. *Ann Rev Biophys* 38:255–270
85. Raychaudhuri S, Stuart JM, Altman RB (2000) Principal components analysis to summarize microarray experiments: application to sporulation time series. *Pac Symp Biocomput* 455–466
86. Relman DA (1999) The search for unrecognized pathogens. *Science* 284:1308–1310
87. Rimour S, Hill D, Milton C, Peyret P (2005) Go arrays: highly dynamic and efficient microarray probe design. *Bioinformatics* 21:1094–1103
88. Robb JA, Bond CW (1979) Pathogenic murine coronaviruses. I. Characterization of biological behavior in vitro and virus-specific intracellular RNA of strongly neurotropic JHMV and weakly neurotropic A59V viruses. *Virology* 94:352–370
89. Roepman P, Wessels LF, Kettelarij N, Kemmeren P, Miles AJ, Lijnzaad P, Tilanus MG, Koole R, Hordijk GJ, van der Vliet PC, Reinders MJ, Slootweg PJ, Holstege FC (2005) An expression profile for diagnosis of lymph node metastases from primary head and neck squamous cell carcinomas. *Nat Genet* 37(2): 182–186
90. Ross DT, Scherf U, Eisen MB, Perou CM, Rees C, Spellman P, Iyer V, Jeffrey SS et al (2000) Systematic variation in gene expression patterns in human cancer cell lines. *Nat Genet* 24:227–235
91. Rubina AY, Pan'kov SV, Dementieva EI, Pen'kov DN, Butygin AV, Vasiliskov VA, Chudinov AV, Mikheikin AL et al (2004)

- Hydrogel drop microchips with immobilized DNA: properties and methods for large-scale production. *Anal Biochem* 325: 92–106
92. Sanchez A, Kondev J (2008) Transcriptional control of noise in gene expression. *Proc Natl Acad Sci USA* 105:5081–5086
 93. Scherf U, Ross DT, Waltham M, Smith LH, Lee JK, Tanabe L, Kohn KW, Reinhold WC et al (2000) A gene expression database for the molecular pharmacology of cancer. *Nat Genet* 24:236–244
 94. Shinde D, Lai Y, Sun F, Arnheim N (2003) Taq DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)_n and (A/T)_n microsatellites. *Nucleic Acids Res* 31:974–980
 95. Sickinger E, Stieler M, Kaufman B, Kapprell H-P, West D, Sandridge A, Devare S, Schochetman G, Hunt JC, Daghfal D, the AxSYM Clinical Study Group (2004) Multicenter evaluation of a new, automated enzyme-linked immunoassay for detection of human immunodeficiency virus-specific antibodies and antigen. *J Clin Microbiol* 42:21–29
 96. Small J, Call DR, Brockman FJ, Straub TM, Chandler DP (2001) Direct detection of 16S rRNA in soil extracts by using oligonucleotide microarrays. *Appl Environ Microbiol* 67:4708–4716
 97. Smith DJ, Lapedes AS, de Jong JC, Bestebroer TM, Stafford GFP, Brun M (2007) Three methods for optimization of cross laboratory and cross-platform microarray expression data. *Nucleic Acids Res* 35:e72
 98. Sobek J, Bartscherer K, Jacob A, Hoheisel JD, Angenendt P (2006) Microarray technology as a universal tool for high-throughput analysis of biological systems. *Comb Chem High Throughput Screen* 9:365–380
 99. Stabenau A, McVicker G, Melsopp C, Proctor G, Clamp M, Birney E (2004) The Ensembl core software libraries. *Genome Res* 14:929–933
 100. Stafford P, Brun M (2007) Three methods for optimization of crosslaboratory and cross- platform microarray expression data. *Nucleic Acids Res* 35(10):e72. *Statistica Sinica* 12:111–140
 101. Sterrenburg E, Turk R, Boer JM, van GB, Ommen den JT (2002) A common reference for cDNA microarray hybridizations. *Nucleic Acids Res* 30:e116
 102. Striebel HM, Birch-Hirschfeld E, Egerer R, Földes-Papp Z (2003) Virus diagnostics on microarrays. *Curr Pharm Biotechnol* 4:401–415
 103. Tan SL, Ganji G, Paepfer B, Proll S, Katze MG (2007) Systems biology and the host response to viral infection. *Nat Biotechnol* 25:1383–1389
 104. Tang P, Chiu C (2010) Metagenomics for the discovery of novel human viruses. *Futur Microbiol* 5:177–189
 105. Tang BS, Chan KH, Cheng VC, Woo PC, Lau SK, Lam CC, Chan TL, Wu AK, Hung IF, Leung SY, Yuen KY (2005) Comparative host gene transcription by microarray analysis early after infection of the Huh7 cell line by severe acute respiratory syndrome coronavirus and human coronavirus 229E. *J Virol* 79(10):6180–6193
 106. Tembe W, Zavaljevski N, Bode E, Chase C, Geyer J, Wasieloski L, Benson G, Reifman J (2007) Oligonucleotide fingerprint identification for microarray-based pathogen diagnostic assays. *Bioinformatics* 23:5–13
 107. Trost E, Hackle H, Maurer M, Trajanoski Z (2003) Java editor for biological pathways. *Bioinformatics* 19:786–787
 108. Tseng GC, Oh MK, Rohlin L, Liao JC, Wong WH (2001) Issues in cDNA microarray analysis: quality filtering, channel normalization, models of variations and assessment of gene effects. *Nucleic Acids Res* 29:2549–2557
 109. Tusher VG, Tibshirani R, Chu G (2001) Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 98:5116–5121
 110. Van de Peppel J, Kemmeren P, Van B H, Radonjic M, Van L, Holstege D (2003) Monitoring global messenger RNA changes in externally controlled microarray experiments. *EMBO Rep* 4(4):387–393
 111. Varga A, James D (2005) Detection and differentiation of plum pox virus using real-time multiplex PCR with SYBR green and melting curve analysis: a rapid method for strain typing. *J Virol Methods* 123:213–220
 112. Vet JAM, Majithia AR, Marras SAE, Tyagi S, Dube, Poiesz SB, Krame JR (1999) Multiplex detection of four pathogenic retroviruses using molecular beacons. *Proc Natl Acad Sci USA* 96:6394–6399
 113. Victoria JG, Kapoor A, Li L, Blinkova O, Slikas B, Wang C, Naeem A, Zaidi S, Delwart E (2009) Metagenomic analyses of viruses in stool samples from children with acute flaccid paralysis. *J Virol* 83:4642–4651
 114. Vinje J, Koopmans MP (2000) Simultaneous detection and genotyping of “Norwalk-like” viruses by oligonucleotide array in a reverse line blot hybridization format. *J Clin Microbiol* 38:2595–2601
 115. Wagner E, Garcia Ramirez JJ, Stingley SW, Aguilar SA, Buehler L, Devi-Rao GB, Ghazal P (2002) Practical approaches to long oligonucleotide-based DNA microarray: lessons from herpesvirus. *Prog Nucleic Acid Res* 71:445–491
 116. Wang D, Coscoy L, Zylberberg M, Avila PC, Boushey HA, Ganem D, De Risi L (2002) Microarray-based detection and genotyping of viral pathogens. *Proc Natl Acad Sci USA* 99:15687–15692
 117. Wang D, Urisman A, Liu YT, Springer M, Ksiazek TG, Erdman DD, Mardis ER, Hickenbotham M, Magrini V, Eldred J et al (2003) Viral discovery and sequence recovery using DNA microarrays. *PLoS Biol* 1(2):E2
 118. Wang HY, Malek RL, Kwitek AE, Greene AS, Luu TV, Behbahani B, Frank B, Quackenbush J, Lee NH (2003) Assessing unmodified 70-mer oligonucleotide probe performance on glass-slide microarrays. *Genome Biol* 4:R5
 119. Wang Z, Daum LT, Vora GJ, Metzgar D, Walter EA, Canas LC, Malanoski AP, Lin B, Stenger DA (2006) Identifying influenza viruses with resequencing microarrays. *Emerg Infect Dis* 12:638–646
 120. Weil MR, Macatee T, Garner HR (2002) Toward a universal standard: comparing two methods for standardizing spotted microarray data. *Biotechniques* 32:1310–1314
 121. Whelen AC, Persing DH (1996) The role of nucleic acid amplification and detection in the clinical microbiology laboratory. *Annu Rev Microbiol* 50:349–373
 122. Wilson KH, Wilson WJ, Radosevich JL, DeSantis TZ, Viswanathan VS, Kuczmarski TA, Andersen GL (2002) High-density microarray of small-subunit ribosomal DNA probes. *Appl Environ Microbiol* 68:2535–2541
 123. Wittberger D, Berens C, Hammann C, Westhof E, Schroeder R (2000) Evaluation of uranyl photocleavage as a probe to monitor ion binding and flexibility in RNAs. *J Mol Biol* 300:339–352
 124. Woo PC, Lau SK, Chu CM, Chan KH, Tsoi HW, Huang Y, Wong BH, Poon RW, Cai JJ, Luk WK, Poon LL, Wong SS, Guan Y, Peiris JS, Yuen KY (2005) Characterization and complete genome sequence of a novel coronavirus, coronavirus HKU1, from patients with pneumonia. *J Virol* 79:884–895
 125. Yang YH, Dudoit S, Luu P, Lin DM, Peng V, Ngai J (2002) Speed TP: normalization for cDNA microarray data: a robust

- composite method addressing single and multiple slide systematic variation. *Nucleic Acids Res* 30(4):e15
126. Yang IV, Chen E, Hasseman JP, Liang W, Frank BC, Wang S, Sharov V, Saeed AI et al (2002) Within the fold: assessing differential expression measures and reproducibility in microarray assays. *Genome Biol* 3:62.1–62.13
127. Yee HY, Speed TP (2002) Design issues for cDNA microarray experiments. *Nat Rev Genet* 3:579–588
128. Zammateo N, Hamels S, De Longueville F, Alexandre I, Gala JL, Brasseur F, Remacle J (2002) New chips for molecular biology and diagnostics. *Biotechnol Annu Rev* 8:85–101
129. Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 31:3406–3415