Genotyping of Measles Virus in Clinical Specimens on the Basis of Oligonucleotide Microarray Hybridization Patterns†

Alexander A. Neverov,^{1,6*} Michaela A. Riddell,^{2,4} William J. Moss,² Dmitriy V. Volokhov,¹
Paul A. Rota,³ Luis E. Lowe,³ Doris Chibo,⁴ Sheilagh B. Smit,⁵ Diane E. Griffin,²
Konstantin M. Chumakov,¹ and Vladimir E. Chizhikov¹

Center for Biologics Evaluation and Research, Food and Drug Administration, Rockville,¹ and W. Harry Feinstone Department of Molecular Microbiology and Immunology, Johns Hopkins University Bloomberg School of Public Health, Baltimore,² Maryland; Measles, Mumps, Rubella, and Herpesvirus Branch, Centers for Disease Control and Prevention/WHO Global Measles Reference Laboratory, Atlanta, Georgia³; Victorian Infectious Diseases Reference Laboratory/WHO Western Pacific Region Measles Reference Laboratory, Melbourne, Victoria, Australia⁴; Vaccine-Preventable Virus Infections Unit, National Institute for Communicable Diseases, Johannesburg, Republic of South Africa⁵; and Institute of Molecular Biology, State Research Center of Virology and Biotechnology “Vector,” Koltsovo, Novosibirsk Region, Russian Federation⁶

Received 12 May 2006/Returned for modification 6 June 2006/Accepted 20 July 2006

An oligonucleotide microarray hybridization method for identification of most known measles virus (MV) genotypes was developed. Like the conventional genotyping method, the microarray relied on detecting sequence differences in the 450-nucleotide region coding for the COOH-terminal 150 amino acids of the nucleoprotein (N). This region was amplified using PCR primers binding to all known MV genotypes. The microarray included 71 pairs of oligonucleotide probes (oligoprobes) immobilized on glass slides. Each pair consisted of a genotype-specific oligoprobe, which matched the sequence of only one target genotype, and a control oligoprobe, which contained mismatches at the nucleotide positions unique to this genotype. A pattern recognition algorithm based on cluster analysis of the ratios of hybridization signals from specific and control oligoprobes was used to identify the specific MV genotype. Following the initial validation, the method was used for rapid genotyping of two panels of coded samples. The results of this study showed good sensitivity (90.7%), specificity (100%), and genotype agreement (91.8%) for the new method compared to the results of genotyping conducted using phylogenetic analysis of viral sequences of the C terminus of the N gene. In addition, the microarray demonstrated the ability to identify potential new genotypes of MV based on the similarity of their hybridization patterns with those of known MV genotypes.

Before the introduction of live attenuated vaccines, measles was endemic in all countries and was a leading cause of childhood morbidity and mortality. High vaccination coverage through routine and supplemental vaccination programs has significantly reduced the circulation of measles virus (MV) in many industrialized nations (14); however, the virus remains endemic in many developing countries, leading to 30 million cases and approximately 454,000 deaths annually (36). The goal of the Global Measles Strategic Plan, sponsored by the World Health Organization (WHO) and the United Nations Children's Fund (UNICEF), is to significantly reduce measles mortality in countries where the virus is endemic and to maintain measles-free status in countries that have already interrupted measles transmission (38). Laboratory confirmation of suspected cases is an essential component of measles surveillance. To this end, WHO has developed a global network of measles laboratories. These laboratories perform serologic assays to detect MV-specific immunoglobulin M antibodies and conduct genetic characterization of wild-type MVs isolated from outbreaks and sporadic cases (2, 33). Molecular epide-

miological studies, along with standard case investigation and reporting, provide the necessary tools to monitor MV circulation and gauge the success of vaccination programs (23).

The classification system for wild-type MV is based on the sequences of the 450 nucleotides coding for the 150 amino acids at the C terminus of the N protein (24, 32, 34, 35, 37). The intergenotype diversity within this region of the N gene is greater than 2.5%, and the most divergent MV genotypes differ by as much as 12%. Measles viruses are classified into eight clades (A to H) that are currently subdivided into 23 recognized genotypes (A, B1 to B3, C1, C2, D1 to D10, E, F, G1 to G3, H1, and H2) (24, 37). Sequence analysis of PCR products is currently the most practical, cost-effective, and accurate method for MV genotyping. Other genotyping methods, such as restriction fragment length polymorphism (RFLP) (22, 29), the heteroduplex mobility assay (HMA) (18), refractory mutation analysis (27), genotyping by nucleotide-specific multiplex PCR (19), and real-time PCR (31), have been proposed. Most of these methods have inherent shortcomings and can differentiate only a limited number of MV genotypes. RFLP-based methods depend on the availability of restriction sites suitable for analysis, and the results of HMA are often difficult to interpret and reproducibility is low. Some of these techniques are technically challenging, expensive, and difficult to standardize. Furthermore, these methods are not amenable to high-throughput screening and lack the sensitivity of se-

* Corresponding author. Mailing address: HFM-470, LMD/CBER, FDA, 1401 Rockville Pike, Rockville, MD 20852-1448. Phone: (301) 827-7388. Fax: (301) 827-9531. E-mail: alexander.neverov@fda.hhs.gov.

† Supplemental material for this article may be found at <http://jcm.asm.org/>.

quence analysis. Data from these various alternative approaches cannot be easily reported to a central database, so results obtained in different laboratories cannot be easily compared.

As the WHO measles laboratory network expands, additional methods for the genetic characterization of MV will be desirable. For example, high-throughput screening techniques may be necessary if the number of specimens increases significantly. Characterization of larger regions of the genome or complete genomes may be required in order to increase the sensitivity of the molecular epidemiologic analysis and to efficiently monitor multiple genetic characteristics of the virus.

DNA microarray technology is an efficient tool for rapid genetic analysis of microorganisms including transcription profiling, resequencing, single-nucleotide polymorphism (SNP) analysis, and genotyping of bacterial and viral pathogens (4–6, 10, 21). Short oligonucleotide probes (oligoprobes) enable discrimination of samples with minor genetic differences. MV genotypes may differ by only a few nucleotides, and these differences are not always conserved even within a specific genotype. Unique signature nucleotide patterns capable of distinguishing all genotypes may not be readily identifiable, and accurate microarray discrimination of closely related genotypes presents a challenge. We propose a novel approach for microarray design and analysis that relies on the recognition of patterns of hybridization signals from a large number of genotype-specific and control oligoprobes. This method has allowed us to correctly identify the genotypes of most tested samples, including a previously unidentified genotype that was not included in the initial microarray design.

MATERIALS AND METHODS

Sample sources. PCR products containing the 450 nucleotides coding for the 150 amino acids at the C terminus of the N protein from different MV strains were prepared from cDNA specimens provided by the Measles, Mumps, Rubella, and Herpesvirus Branch, Centers for Disease Control and Prevention (Atlanta, GA); Department of Molecular Microbiology and Immunology, Johns Hopkins Bloomberg School of Public Health (Baltimore, MD); Victorian Infectious Diseases Reference Laboratory (North Melbourne, Victoria, Australia); Vaccine-Preventable Virus Infections Unit, National Institute for Communicable Diseases (Johannesburg, Republic of South Africa); Public Health Agency of Canada (Winnipeg, Manitoba); and Hospital Ramón y Cajal, Instituto Nacional de la Salud (Madrid, Spain). Detailed information about the samples is given in Table S1 in the supplemental material. Coded panels for assay validation were shipped on dry ice to the Johns Hopkins Bloomberg School of Public Health for testing. Panel 1, collected in Australia, consisted of 55 cDNAs prepared, at the time of collection, from clinical samples (nose/throat swab, oral fluid, urine, serum, or nasopharyngeal aspirate) or viral isolates. Thirty-nine samples contained the following MV genotypes: A (number of samples analyzed [n] = 2), C2 (n = 1), D1 (n = 1), D2 (n = 1), D3 (n = 3), D4 (n = 4), D5 (n = 6), D7 (n = 2), D8 (n = 4), D9 (n = 3), G2 (n = 2), G3 (n = 2), H1 (n = 7), and H2 (n = 1). MV genotypes were assigned after sequence analysis in Australia as previously described (7–9). In addition, panel 1 contained 16 samples of viruses other than MV: human parvovirus B19 (n = 2), respiratory syncytial virus (n = 2), parainfluenza virus type 2 (n = 2), human herpesvirus 6 (n = 2), Epstein-Barr virus (n = 2), varicella-zoster virus (n = 2), cytomegalovirus (n = 2), and picornavirus (rhinovirus untyped) (n = 2). Panel 2, collected in the Southern African region, consisted of 15 samples of MV cDNA (SA samples) prepared and sequenced during the initial laboratory investigation from clinical specimens (urine or serum) or viral isolates and contained MV genotypes B2 (n = 4), D2 (n = 9), and D4 (n = 2). The MV genotypes and nucleotide sequences of all tested samples, previously determined in other studies, were revealed after completion of the assay evaluation and validation.

Sample amplification, transcription, and hybridization. Samples of cDNA/DNA for the assay validation were initially amplified using previously de-

scribed primers (MV 60 and MV 63) and cycling conditions (26). Second-round amplification of DNA samples was conducted using two newly designed PCR primers, Measles_F (GCTATGCCATGGGAGTAGGAGTGAACCTTG) and Measles_R_T7 (TAATACGACTCACTATAGGGCGGCCTCTCGACCTAGTCTAGAAG). The latter contained the bacteriophage T7 promoter (underlined in the sequence) for facilitation of later *in vitro* transcription. The 50- μ l reaction mixture contained 2.5 U of HotStarTaq DNA polymerase, 1 $\times$ reaction buffer supplemented with 2.5 mM MgCl₂ (QIAGEN, Chatsworth, CA), 0.4 μ M each primer, 0.2 mM each deoxynucleoside triphosphate, and 2 μ l (~100 ng) of template DNA. PCR was performed using GeneAmp PCR system 9700 or 2720 (Applied Biosystems, Foster City, CA) with an initial 15-min activation at 95°C, followed by 40 cycles consisting of 40 s of incubation at 94°C, 40 s at 57°C, 1 min at 72°C, and a final 10-min extension at 72°C. Further *in vitro* transcription, RNA labeling, and hybridization were performed as previously described (30), except that second-round PCR products from the validation study were not purified prior to single-stranded RNA (ssRNA) synthesis.

Design of oligonucleotide probes. The genetic variability within the 450 nucleotides coding for the 150 amino acids at the C terminus of the N protein (nucleotides 1233 to 1682) was analyzed by aligning all of the sequences available in GenBank (more than 700). Preliminary analysis of these sequences allowed us to reduce the data set to 405 unique sequences. The final data set contained 15 sequences for genotype A, 2 for B1, 2 for B2, 18 for B3.1 (similar to New York.USA/94), 46 for B3.2 (similar to Ibadan.NIE/97/1), 29 for C1, 36 for C2, 12 for D1, 6 for D2, 22 for D3, 41 for D4, 29 for D5.1 (similar to Palau.BLA/93), 24 for D5.2 (similar to Bangkok.THA/93/1), 38 for D6, 7 for D7.1 (similar to Victoria.AUS/16.85), 6 for D7.2 (similar to Illinois.USA/50.99), 11 for D8, 4 for D9, 6 for E, 2 for F, 2 for G1, 3 for G2, 4 for G3, 33 for H1, and 7 for H2. Phylogenetic analysis of samples was conducted using Mega (version 3.1) software (20).

A set of oligoprobes specific to different MV genotypes was designed using custom Oligoscan (version 2.12) software. While most genotypes could be unambiguously identified by unique genotype-specific oligoprobes, some did not contain unique sequences suitable for probe design. The identification of these genotypes was based on recognition of unique combinations of signals from probes common to two or more genotypes. Due to the low level of genetic divergence between MV genotypes, the majority (46.5%) of the genotype-specific microarray probes differed from the sequences of other MV genotypes by only one nucleotide. Among other probes, 42.5% contained two mismatches, 9.6% contained three, and 1.4% contained four. To increase the reliability of MV genotyping, all genotype-specific oligoprobes were complemented with control oligoprobes in which the unique bases (located near the center of each probe) were replaced with the nucleotide most commonly observed in other genotypes. Therefore, the ratios of genotype-specific oligoprobes to control oligoprobes were used for analysis rather than the hybridization intensities from genotype-specific oligoprobes. The sequences and characteristics of all oligoprobes used in the study can be found in Table S2 in the supplemental material. The oligonucleotides had melting temperatures of ~45°C and were synthesized by Operon Biotechnologies (Huntsville, AL).

Microarray design, manufacture, and hybridization conditions. The microarray consisted of 145 oligoprobes (71 pairs and 3 controls) spotted three times each onto the surfaces of CodeLink activated slides according to the protocols described previously (30). During microarray fabrication, each spotting mixture contained, besides the MV oligonucleotide, a quality control (QC) oligonucleotide with an arbitrary sequence unrelated to measles virus RNA (molar ratio, 20:1, respectively). Before hybridization, a Cy3-labeled ssRNA MV sample was mixed with a Cy5-labeled anti-QC oligonucleotide (with a sequence complementary to that of the QC oligonucleotide probe present in each spot.) The final hybridization mixture contained 1 to 2 μ M Cy3-labeled ssRNA sample and 0.2 μ M Cy5-labeled anti-QC oligonucleotide (complementary to the QC probe immobilized on the chip) in 1 $\times$ MICROMAX hybridization buffer III (Perkin-Elmer, Boston, MA). Hybridization was conducted at 50°C for 60 min, followed by the standard washing procedure described previously (30).

Scanning and data analysis. The ScanArray 5000 microarray analysis system (Perkin-Elmer, Boston, MA) with 632-nm (for Cy5) and 543-nm (for Cy3) lasers was used. The Cy3 fluorescent signal and local background were measured for each microarray element and analyzed by ScanArray Express software (Perkin-Elmer, Boston, MA). Cy5 images were used for assessing the appearance of each spot, the spot morphology, and the uniformity of hybridization conditions among all surfaces of the microarray. The validation study was performed using an Axon 4200AL scanner and GenePix Pro software (version 6.0; Molecular Devices, Sunnyvale, CA).

The fluorescence intensities obtained from genotype-specific oligoprobes were divided by signals obtained from the respective control oligoprobes. If both the

FIG. 1. Hierarchical cluster analysis of hybridization patterns of viral samples. The color-coded data matrix shows ratios of fluorescence (specific to the respective control oligoprobes) arranged in both dimensions by hierarchical clustering. The last two rows show hybridization data from oligoprobes common for all MV strains. Individual hybridization profiles for each sample are shown in columns (labeled at the top of each column), and hybridization data from each oligoprobe are presented in rows, labeled on the right. Upper and lower color palettes show data matrix scales for specific/control oligoprobe ratios and normalized signals from broadly specific oligoprobes, respectively. The tree on the left shows clustering based on the similarity of hybridization oligoprobe specificities. The tree at the top shows clusters obtained by comparison of hybridization patterns of samples with different MVs. Genotypes are given at the top. Genotype B2 samples SA1, SA2, SA3, and SA15 were erroneously identified as belonging to genotype B3 (asterisked). VIDRL samples are from the Victorian Infectious Diseases Reference Laboratory.

specific and control oligoprobes yielded weak hybridization signals (less than four times the local background), the ratios were ignored and considered to be zero.

The distance between two hybridization patterns was estimated based on the Pearson correlation coefficient (16) between two sets of hybridization data expressed as ratios (see "Design of oligonucleotide probes" above). The Pearson correlation coefficient was calculated using the formula

$$P_{XY} = \frac{\sum_i^N (X_i - \bar{X}) \cdot (Y_i - \bar{Y})}{\sqrt{\sum_i^N (X_i - \bar{X})^2 \cdot \sum_i^N (Y_i - \bar{Y})^2}}$$

where X_i and Y_i are ratios calculated for each of N oligoprobes, and $\bar{X}$ and $\bar{Y}$ are the respective averages for all probes in hybridization patterns for two samples. The distance between two hybridization patterns was postulated to be $-\log(P_{XY})$, and the complete matrix of all pairwise distances was used to construct the dendrogram showing relationships between the hybridization patterns. The dendrogram was built by the topological optimization method previously described by Chumakov and Iushmanov (11, 15). All distance calculations and tree construction were done automatically by using custom Oligoscan software.

For illustration purposes, after completion of the study, all hybridization pattern data were organized in the form of a rectangular table and analyzed by Cluster (version 3.0) software (12). Independent clustering of columns and rows of a data set were performed to classify different samples according to their hybridization profiles (column clustering) and to outline pairs of oligonucleotides associated with each group of MV strains (row clustering). The same agglomerative algorithm was used to perform both hierarchical clusterings. The distance matrix was calculated using a metric based on Pearson's correlation coefficient (see "Scanning and data analysis" above). The average linkage model for measuring distances between items was implemented in the Cluster program. Clustering results were visualized using Java TreeView (version 1.0.12) software (http://sourceforge.net/project/showfiles.php?group_id=84593). Exact binomial 95% confidence intervals (95% CI) for sensitivity and specificity were calculated using STATA software (version 7.0; StataCorp, College Station, TX).

RESULTS

Development, optimization, and preliminary evaluation of the MV microarray. The development and optimization of the microarray assay included the design of oligoprobe pairs, testing of their discriminatory efficiency, and optimization of protocols for PCR, hybridization sample preparation, chip fabrication, microarray hybridization, and image analysis. Optimization was conducted using a panel of samples obtained from the Centers for Disease Control and Prevention, representing the WHO reference strains of MV genotypes A, B1, B2, B3.1, B3.2, C1, C2, D3, D4, D5.1, D5.2, D6, D7.1, D7.2, D9, G1, G2, G3, H1, and H2. Even though most of the genotype- and group-specific oligoprobes demonstrated good specificity, relatively high cross-hybridization (more than four times above the local background) was observed for oligoprobes specific to genotypes in clade C. Therefore, these oligoprobes were redesigned. After optimization, the microarray assay correctly identified each genotype from the reference sample collection. All the results of the microarray genotyping were 100% concordant with those of sequence-based genotyping.

Preliminary evaluation of the microarray was performed using 63 DNA samples, which included 20 previously tested reference strains (REF samples) (see Table S1 in the supplemental material); 41 coded MV DNA samples, also obtained from the Centers for Disease Control and Prevention (CDC samples); and 2 samples of F and E genotypes, provided by R. Fernández-Muñoz (Hospital Ramón y Cajal, Instituto Nacio-

nal de la Salud, Madrid, Spain) and G. Tipples (Public Health Agency of Canada, Winnipeg, Manitoba). The results of microarray analysis are shown in Fig. 1.

Only samples belonging to genotypes D6 and D8 demonstrated clear-cut hybridization patterns. Other genotypes were expected to have various degrees of cross-hybridization with oligoprobes of different specificity. However, the ambiguities in interpretation were overcome by using cluster analysis of the microarray hybridization patterns, allowing us to place samples with similar hybridization patterns into separate groups and identify the genotype on the basis of the reference sample that belonged to each group. In total, we were able to identify 20 distinct groups (see the tree at the top of Fig. 1). Dendrograms built from the microarray data and phylogenetic trees built from nucleotide sequence analysis (Fig. 2) were similar and formed the same number of branches, with identical samples composing each group. MV genotypes were identified correctly for 19 of 20 groups. The only unidentified sample, which had a pattern that did not match any known MV genotype cluster, was CDC 41. The pattern for this sample was unique and contained signals only from the probes universal to all MV strains. No significant hybridization of this sample with other genotype- or group-specific oligoprobes was observed. After the sample code was broken, the sample was found to belong to a newly established genotype (D10) recently identified in Uganda (24). The oligoprobes for this genotype had not been included in the microarray because genotype D10 sequence data were not available at the time when the microarray was developed. Thus, inclusion in the microarray of oligoprobes for regions of the N gene common to all known MV genotypes enabled detection of previously unknown genotypes. Discovery of samples with unusual hybridization patterns should lead to their sequencing and genetic analysis and to the design of oligoprobes for the new genotype, followed by incorporation of these probes into updated microchips.

Analysis of the hybridization patterns allowed us to unambiguously distinguish closely related genetic groups (e.g., samples C1 REF and C2 REF). Established genetic clusters could be additionally subdivided into separate subgroups. Thus, genotype D7 contained two distinct subgroups composed of samples D7.1 REF and D7.2 REF and of samples CDC 17 and CDC 18. Genotype D5 also contained two separate groups, represented by sample CDC 13 and the reference strains. Similarly, the cluster of genotype B3 samples showed two genetically distant groups, which have been described previously (17). Therefore, the cross-reactivity of microarray oligoprobes did not interfere with MV genotyping but instead provided additional and valuable information for discriminating between closely related samples within a genotype.

To assess the reproducibility of the microarray assay, several MV samples were independently analyzed using different lots of microarray slides prepared at different times (Fig. 1, samples B2 REF and B2 REF R, CDC 13 and CDC 13 R, C1 REF and C1 REF R, CDC 19 and CDC 19 R, CDC 20 and CDC 20 R, CDC 21 and CDC 21 R, and G1 REF and G1 REF R). In all cases, the hybridization patterns were almost identical. Samples belonging to the same MV genotype, with identical N gene sequences, always produced identical hybridization patterns (Fig. 1, samples CDC 11 and CDC 12 of genotype D4, samples CDC 19, CDC 20, and CDC 21 of genotype D8, samples G3

REF and CDC 30 of genotype G3, and samples CDC 33, CDC 34, and CDC 35 of genotype H1).

To demonstrate that the microarray assay was specific, we tested samples of other *Paramyxovirinae*, mumps virus and Nipah virus. The presence of PCR products for the Nipah virus samples was unexpected and was probably caused by nonspecific primer binding at the high template concentration. However, the RNA transcribed from those amplicons did not hybridize with any MV-specific oligoprobes of the microarray.

Validation of MV microarray genotyping. In addition, two different blinded panels of samples were evaluated at the Johns Hopkins Bloomberg School of Public Health using microarray slides made at the FDA laboratory. Of the 55 samples in panel 1, 39 represented 14 different MV genotypes, including genotypes D1 and D2, which were not previously tested during assay evaluation (see Materials and Methods). The remaining 16 samples of the panel represented other viruses also causing fever and rash. MV was detected and subsequently genotyped for 35 samples (sensitivity, 89.7% [95% CI, 75.8 to 97.1%]). The results of genotyping were 100% concordant with the genotypes previously identified by sequence analysis. All other samples from panel 1 were negative by both PCR and microarray hybridization (100% specificity). Panel 2 contained 15 samples of different MV genotypes. The microarray detected MV in 14 samples (sensitivity, 93.3% [95% CI, 68.1 to 99.8%]). It is not clear why the last sample failed to be amplified by PCR. The specificity for this panel could not be assessed because all samples were positive for MV. The genotypes of 10 of 14 samples were correctly identified by microarray analysis (71.4% genotype agreement). The reason why we failed to identify four samples of genotype B2 was a greater degree of genetic drift between these recently isolated MV strains (28) and the reference strain for genotype B2, isolated in 1983. Due to the 1.7 to 2.0% nucleotide difference between the recent wild-type strains and the reference sequence, the microarray mistakenly identified these samples as genotype B3.2. Sequence analysis of these genotype B2 isolates showed that correct genotype assignment by microarray was indeed possible but would require designing new oligoprobes. We intend to include these oligoprobes in newer versions of the MV microarray. Thus, the combined sensitivity of the microarray based on these two panels was 90.7% (95% CI, 79.7 to 96.9%). Because the reduction in sensitivity was caused exclusively by the failure to amplify MV RNA from samples that may have been compromised by inadequate international shipping conditions, we expect that sensitivity would be increased by testing samples that are shipped and stored under appropriate conditions.

DISCUSSION

Genotyping of MV isolates is an important component of measles surveillance because it provides a means to track the transmission pathways of the virus (2, 3, 23, 25). Genetic characterization of viral isolates is the only means to distinguish between a vaccine reaction and disease caused by wild-type virus. Molecular characterization is conducted at specialized global laboratories operating under WHO auspices as well as at some regional and national laboratories (13). Nucleotide sequencing of PCR products currently provides the most ac-

FIG. 2. Results of phylogenetic analysis of sequences of viral samples. A tree was built by the neighbor-joining method, using Mega (version 3.1) software (20), on the basis of the number of nucleotide substitutions. MV genotypes (boldfaced) are arranged outside the circle, and the strains belonging to each genotype are given inside the circle. Recently isolated genotype B2 strains (28) are boldfaced and italicized.

curate and cost-effective method to identify MV genotypes. However, alternative high-throughput techniques that target multiple regions of the MV genome are desirable.

The microarray technique described here is amenable to high-throughput implementation and enables simultaneous multilocus analysis of the pathogen genotype as well as accurate analysis of mixed pathogen populations. It should be noted that microarray technology has emerged recently and is still in the developmental stage. It certainly has the potential to be significantly improved and simplified and to become more robust and efficient, particularly for the purposes of pathogen detection and identification. For example, the replacement of

fluorescent dyes (e.g., Cy5 and Cy3) with nanogold particles (1) will allow users to significantly reduce the cost of analysis and to use inexpensive imagers with digital cameras instead of sophisticated microarray scanners. Therefore, microarray technology has unique features that can make it very attractive and a valuable instrument for rapid detection and genotyping of different viral and bacterial pathogens.

Regardless of the platform, all microarray methods for genotyping of microorganisms are based on the use of relatively short (15- to 40-mer) oligoprobes that hybridize only with sequences from a particular subset of species. Although several computer programs are currently available for the au-

tomated design of unique microarray oligoprobes, the design of oligoprobes continues to be a challenge, particularly for closely related species that differ by one or a few point mutations. Possible cross-hybridization with unrelated oligoprobes also decreases the utility of this approach. The use of routine methods for the design of oligoprobes for phylogenetically informative parts of the N gene did not produce a sufficient number of unique genotype-specific oligoprobes. To overcome this obstacle, we used multiple group-specific oligonucleotides in addition to genotype-specific oligonucleotides. Thus, MV genotyping relied not only on the comparison of hybridization data from genotype-specific oligoprobes but also on analysis of complex hybridization patterns, which included signals from multiple oligoprobes specific to overlapping groups of MVs. Pairs of oligonucleotides that included specific and control probes were used to detect single-base substitutions.

To identify the genotype of an unknown MV sample, we compared its hybridization pattern with patterns obtained from reference strains. The Pearson correlation coefficient was used to evaluate the difference between hybridization patterns. Computing all possible pairwise comparisons between microarray results for clinical and reference samples resulted in a distance matrix used to construct a dendrogram showing the relatedness of samples. This technique, routinely used in microarray-based methods of transcription analysis, has not previously been used for genotyping. We found that MV genotypes identified by cluster analysis of microarray hybridization patterns were consistent with genotypes determined by sequencing. Therefore, microarray hybridization could be an alternative method for rapid, high-throughput genotyping of new clinical isolates, including those with novel genotypes. The procedures used in microarray studies are not yet fully optimized for routine use and are still relatively expensive. Despite these disadvantages, recent breakthroughs in microfluidics, silicon chips, electronics, and nanotechnology promise the creation of small, inexpensive, and user-friendly devices, combining different laboratory instruments into single, self-contained units that will dramatically reduce individual assay costs.

The main goal of our study was to evaluate the feasibility and usefulness of a pattern recognition approach, widely used in transcription profiling studies, for viral genotyping, particularly for discrimination among viruses with closely related sequences. This approach eliminates the need to design strictly genotype specific oligoprobes for all known genotypes and allows the use of oligoprobes that bind more than one genotype to produce interpretable genotyping information.

ACKNOWLEDGMENTS

We thank David Asher of the Center for Biologics Evaluation and Research, Food and Drug Administration, Rockville, MD, for review of the manuscript and Rafael Fernández-Muñoz of the Hospital Ramón y Cajal, Instituto Nacional de la Salud, Madrid, Spain, and Graham Tipples of the Public Health Agency of Canada, Winnipeg, Manitoba, for generously providing samples for genotyping.

This work was partially supported by grants from the DHHS Biotechnology Engagement program (BTEP 16, to K.M.C.), the National Health and Medical Research Council of Australia (grant 282418, to M.A.R.), and the Bill and Melinda Gates Foundation (grant 3522, to D.E.G.) and by grants from the Elizabeth Glaser Pediatric AIDS Foundation (51331-28-PG) and the Thrasher Research Fund (02818-9) to W.J.M.

REFERENCES

- Bao, Y. P., M. Huber, T. F. Wei, S. S. Marla, J. J. Storhoff, and U. R. Muller. 2005. SNP identification in unamplified human genomic DNA with gold nanoparticle probes. *Nucleic Acids Res.* **33**:e15.
- Bellini, W. J., and R. F. Helfand. 2003. The challenges and strategies for laboratory diagnosis of measles in an international setting. *J. Infect. Dis.* **187**(Suppl. 1):S283–S290.
- Bellini, W. J., and P. A. Rota. 1998. Genetic diversity of wild-type measles viruses: implications for global measles elimination programs. *Emerg. Infect. Dis.* **4**:29–35.
- Boriskin, Y. S., P. S. Rice, R. A. Stabler, J. Hinds, H. Al-Ghusein, K. Vass, and P. D. Butcher. 2004. DNA microarrays for virus detection in cases of central nervous system infection. *J. Clin. Microbiol.* **42**:5811–5818.
- Butcher, L. M., E. Meaburn, L. Liu, C. Fernandes, L. Hill, A. Al-Chalabi, R. Plomin, L. Schalkwyk, and I. W. Craig. 2004. Genotyping pooled DNA on microarrays: a systematic genome screen of thousands of SNPs in large samples to detect QTLs for complex traits. *Behav. Genet.* **34**:549–555.
- Cherkasova, E., M. Laassri, V. Chizhikov, E. Korotkova, E. Dragunsky, V. I. Agol, and K. Chumakov. 2003. Microarray analysis of evolution of RNA viruses: evidence of circulation of virulent highly divergent vaccine-derived polioviruses. *Proc. Natl. Acad. Sci. USA* **100**:9398–9403.
- Chibo, D., C. J. Birch, P. A. Rota, and M. G. Catton. 2000. Molecular characterization of measles viruses isolated in Victoria, Australia, between 1973 and 1998. *J. Gen. Virol.* **81**:2511–2518.
- Chibo, D., M. Riddell, M. Catton, and C. Birch. 2002. Novel measles virus genotype, East Timor and Australia. *Emerg. Infect. Dis.* **8**:735–737.
- Chibo, D., M. Riddell, M. Catton, M. Lyon, G. Lum, and C. Birch. 2003. Studies of measles viruses circulating in Australia between 1999 and 2001 reveals a new genotype. *Virus Res.* **91**:213–221.
- Chizhikov, V., A. Rasooly, K. Chumakov, and D. D. Levy. 2001. Microarray analysis of microbial virulence factors. *Appl. Environ. Microbiol.* **67**:3258–3263.
- Chumakov, K. M., and S. V. Iushmanov. 1988. A principle of maximum topological similarity in molecular systematics. *Mol. Gen. Mikrobiol. Virusol.* **1988**:3–9. (In Russian.)
- Eisen, M. B., P. T. Spellman, P. O. Brown, and D. Botstein. 1998. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. USA* **95**:14863–14868.
- Featherstone, D., D. Brown, and R. Sanders. 2003. Development of the Global Measles Laboratory Network. *J. Infect. Dis.* **187**(Suppl. 1):S264–S269.
- Hersh, B. S., G. Tambini, A. C. Nogueira, P. Carrasco, and C. A. de Quadros. 2000. Review of regional measles surveillance data in the Americas, 1996–99. *Lancet* **355**:1943–1948.
- Iushmanov, S. V., and K. M. Chumakov. 1988. Algorithms for constructing phylogenetic trees of maximum topological similarity. *Mol. Gen. Mikrobiol. Virusol.* **1988**:9–15. (In Russian.)
- Kenney, J. F., and E. S. Keeping. 1962. *Mathematics of statistics*, 3rd ed., p. 252–285. Van Nostrand, Princeton, N.J.
- Kouomou, D. W., E. Nerrienet, J. Mfoupouendoun, G. Tene, H. Whittle, and T. F. Wild. 2002. Measles virus strains circulating in Central and West Africa: geographical distribution of two B3 genotypes. *J. Med. Virol.* **68**:433–440.
- Kreis, S., and T. Whistler. 1997. Rapid identification of measles virus strains by the heteroduplex mobility assay. *Virus Res.* **47**:197–203.
- Kremer, J. R., F. Fack, C. M. Olinger, M. N. Mulders, and C. P. Muller. 2004. Measles virus genotyping by nucleotide-specific multiplex PCR. *J. Clin. Microbiol.* **42**:3017–3022.
- Kumar, S., K. Tamura, and M. Nei. 2004. MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform.* **5**:150–163.
- Laassri, M., V. Chizhikov, M. Mikheev, S. Shchelkunov, and K. Chumakov. 2003. Detection and discrimination of orthopoxviruses using microarrays of immobilized oligonucleotides. *J. Virol. Methods* **112**:67–78.
- Mori, T. 1994. A simple method for genetic differentiation of the AIK-C vaccine strain from wild strains of measles virus. *Biologicals* **22**:179–185.
- Mulders, M. N., A. T. Truong, and C. P. Muller. 2001. Monitoring of measles elimination using molecular epidemiology. *Vaccine* **19**:2245–2249.
- Muwonge, A., M. Nanyunja, P. A. Rota, J. Bwogi, L. Lowe, S. L. Liffick, W. J. Bellini, and S. Sylvester. 2005. New measles genotype, Uganda. *Emerg. Infect. Dis.* **11**:1522–1526.
- Rota, J. S., J. L. Heath, P. A. Rota, G. E. King, M. L. Celma, J. Carabana, R. Fernandez-Munoz, D. Brown, L. Jin, and W. J. Bellini. 1996. Molecular epidemiology of measles virus: identification of pathways of transmission and implications for measles elimination. *J. Infect. Dis.* **173**:32–37.
- Rota, P. A., S. L. Liffick, J. S. Rota, R. S. Katz, S. Redd, M. Papania, and W. J. Bellini. 2002. Molecular epidemiology of measles viruses in the United States, 1997–2001. *Emerg. Infect. Dis.* **8**:902–908.
- Samuel, D., S. Beard, H. Yang, N. Saunders, and L. Jin. 2003. Genotyping of measles and mumps virus strains using amplification refractory mutation system analysis combined with enzyme immunoassay: a simple method for outbreak investigations. *J. Med. Virol.* **69**:279–285.

28. **Smit, S. B., D. Hardie, and C. T. Tiemessen.** 2005. Measles virus genotype B2 is not inactive: evidence of continued circulation in Africa. *J. Med. Virol.* **77**:550–557.
29. **Takahashi, M., T. Nakayama, Y. Kashiwagi, T. Takami, S. Sonoda, T. Yamanaka, H. Ochiai, T. Ihara, and T. Tajima.** 2000. Single genotype of measles virus is dominant whereas several genotypes of mumps virus are co-circulating. *J. Med. Virol.* **62**:278–285.
30. **Wade, M. M., D. Volokhov, M. Peredelchuk, V. Chizhikov, and Y. Zhang.** 2004. Accurate mapping of mutations of pyrazinamide-resistant *Mycobacterium tuberculosis* strains with a scanning-frame oligonucleotide microarray. *Diagn. Microbiol. Infect. Dis.* **49**:89–97.
31. **Waku-Koumou, D., A. Alla, B. Blanquier, D. Jeantet, H. Caidi, A. Rguig, F. Freymuth, and F. T. Wild.** 2006. Genotyping measles virus by real-time amplification refractory mutation system PCR represents a rapid approach for measles outbreak investigations. *J. Clin. Microbiol.* **44**:487–494.
32. **World Health Organization.** 1998. Expanded Programme on Immunization (EPI). Standardization of the nomenclature for describing the genetic characteristics of wild-type measles viruses. *Wkly. Epidemiol. Rec.* **73**:265–269.
33. **World Health Organization.** 2005. Global measles and rubella laboratory network—update. *Wkly. Epidemiol. Rec.* **80**:384–388.
34. **World Health Organization.** 2001. Nomenclature for describing the genetic characteristics of wild-type measles viruses (update). *Wkly. Epidemiol. Rec.* **76**:249–251.
35. **World Health Organization.** 2001. Nomenclature for describing the genetic characteristics of wild-type measles viruses (update). Part I. *Wkly. Epidemiol. Rec.* **76**:242–247.
36. **World Health Organization.** 2006. Progress in reducing global measles deaths: 1999–2004. *Wkly. Epidemiol. Rec.* **81**:90–94.
37. **World Health Organization.** 2003. Update of the nomenclature for describing the genetic characteristics of wild-type measles viruses: new genotypes and reference strains. *Wkly. Epidemiol. Rec.* **78**:229–232.
38. **World Health Organization.** 2002. WHO-UNICEF joint statement on strategies to reduce measles mortality worldwide. *Wkly. Epidemiol. Rec.* **77**:224–228.