APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Jan. 2006, p. 207–211 Vol. 72, No. 1
0099-2240/06/$08.00+0 doi:10.1128/AEM.72.1.207–211.2006

Evaluation of Affymetrix Severe Acute Respiratory Syndrome Resequencing GeneChips in Characterization of the Genomes of Two Strains of Coronavirus Infecting Humans

Irshad M. Sulaiman,[1]* Xin Liu,[2] Michael Frace,[1] Nikhat Sulaiman,[1] Melissa Olsen-Rasmussen,[1] Elizabeth Neuhaus,[1] Paul A. Rota,[2] and Robert M. Wohlhueter[1]

Biotechnology Core Facility Branch, Scientific Resources Program,[1] and Respiratory and Enteric Viruses Branch, Division of Viral and Rickettsial Diseases,[2] National Center for Infectious Diseases, Centers for Disease Control and Prevention, Atlanta, Georgia 30333

Received 24 June 2005/Accepted 11 October 2005

Severe acute respiratory syndrome (SARS) was discovered during a recent global outbreak of atypical pneumonia. A number of immunologic and molecular studies of the clinical samples led to the conclusion that a novel coronavirus (SARS-CoV) was associated with the outbreak. Later, a SARS resequencing GeneChip was developed by Affymetrix to characterize the complete genome of SARS-CoV on a single GeneChip. The present study was carried out to evaluate the performance of SARS resequencing GeneChips. Two human SARS-CoV strains (CDC#200301157 and Urbani) were resequenced by the SARS GeneChips. Five overlapping PCR amplicons were generated for each strain and hybridized with these GeneChips. The successfully hybridized GeneChips generated nucleotide sequences of nearly complete genomes for the two SARS-CoV strains with an average call rate of 94.6%. Multiple alignments of nucleotide sequences obtained from SARS GeneChips and conventional sequencing revealed full concordance. Furthermore, the GeneChip-based analysis revealed no additional polymorphic sites. The results of this study suggest that GeneChip-based genome characterization is fast and reproducible. Thus, SARS resequencing GeneChips may be employed as an alternate tool to obtain genome sequences of SARS-CoV strains pathogenic for humans in order to further understand the transmission dynamics of these viruses.

A unique coronavirus associated with severe acute respiratory syndrome (SARS-CoV) was discovered in a global outbreak of atypical pneumonia during 2002 to 2003. The first case of SARS was identified in the Guangdong Province of China, and SARS eventually spread to over 29 countries. The impact of SARS was severe, with 800 deaths among the approximately 8,400 infected individuals. Several studies revealed the SARS agent to be a novel coronavirus with a positive-stranded RNA genome, which had never been detected in humans before. Furthermore, nucleotide sequence characterization of the whole genome confirmed that SARS-CoV had the largest genome of any known RNA virus and was not closely related to any previously described coronaviruses (1, 6–8). Nevertheless, the zoonotic origin of the outbreak has also been suggested (3).

The National Institute of Allergy and Infectious Diseases (NIAID) launched the SARS-CoV array program to develop a rapid diagnostic tool that could improve understanding of the pathogen. As a part of this research program, the SARS resequencing GeneChip was made by Affymetrix (Santa Clara, Calif.), which can interrogate 29,724 bases of SARS-CoV in a single hybridization. In this study, the performance of SARS resequencing GeneChips was assessed by hybridizing two strains of SARS-CoV infecting humans.

* Corresponding author. Mailing address: Biotechnology Core Facility Branch, Scientific Resources Program, National Center for Infectious Diseases, Centers for Disease Control and Prevention, Mailstop G-36, 1600 Clifton Road, Atlanta, GA 30333. Phone: (404) 639-1644. Fax: (404) 639-1331. E-mail: isulaiman@cdc.gov.

MATERIALS AND METHODS

Virus RNA and RT-PCR. SARS-CoV RNA was extracted from infected Vero cells by the guanidinium acid-phenol method as described before (8). Five overlapping fragments (5.0 kb, 6.5 kb, 6.8 kb, 6.5 kb, and 7.3 kb) were generated for the CDC#200301157 and Urbani genomes. The Urbani genome was amplified using a long reverse transcription-PCR (RT-PCR) protocol as described before (8). However, the CDC#200301157 genome was amplified by the SuperScript one-step RT-PCR kit for long templates (Invitrogen, Carlsbad, Calif.). The cDNA products were analyzed by agarose gel electrophoresis and visualized after ethidium bromide staining.

Pooling and quantitation of PCR products. The cDNA products were purified by the QIAquick cleanup kit (QIAGEN, Valencia, Calif.). The concentration of these amplified products was measured at A_{260} in a NanoDrop-1000 spectrophotometer (Rockland, Del.). Equimolar amounts of these amplicons covering the entire SARS genome were pooled prior to fragmentation and labeling to achieve maximum sequence information in a single GeneChip hybridization.

Fragmentation and labeling of PCR products. Pooled cDNA amplicons were fragmented using the GeneChip fragmentation reagent (Affymetrix) in a thermocycler (preheated at 37°C with a single cycle of 37°C for 15 min, 95°C for 15 min, and 4°C hold) and visualized on a 4 to 20% Novex Tris-borate-EDTA gel (Invitrogen) to ensure that an optimal range of fragment sizes (20 to 200 bp) was achieved. Subsequently, the fragmented samples were end labeled using the GeneChip labeling reagent (Affymetrix) while incubating at 37°C for 2 h, followed by inactivation at 95°C for 15 min. Samples were then cooled on ice and stored at −20°C until hybridization was achieved.

Hybridization and staining of the amplified product. The SARS resequencing GeneChips were equilibrated at room temperature for 15 min before hybridization. Prehybridization was accomplished by filling each GeneChip with 200 μl of prehybridization buffer (10 mM Tris, pH 7.8, 0.01% Tween 20) followed by placing it in a GeneChip hybridization oven (Affymetrix) set at 45°C and rotating at 60 rpm for 15 min. To 60 μl of fragmented and labeled cDNA samples was added 160 μl of freshly prepared hybridization cocktail (3 M tetra methyl ammonium chloride, 10 mM Tris, pH 7.8, 0.01% Tween 20, 500 μg/ml acetylated bovine serum albumin [BSA], 100 μg/ml herring sperm DNA, with 0.26 μg of fragmented and labeled 7.5-kb DNA serving as positive hybridization control), and this mixture was denatured by placing the tubes at 95°C for 5 min and

TABLE 1. Call rate in SARS resequencing GeneChips hybridized with two strains of SARS-CoV infecting humans at GDAS default homozygote model settings

Strain and GeneChip	Call rate (%)	No. of discordant calls	Accuracy (%)
CDC#200301157[a]			
SARS3[b]	93.6	0	100
SARS4	90.7	0	100
SARS7	96.5	0	100
SARS8[b]	96.4	0	100
SARS11[b]	95.8	0	100
Urbani[a]			
SARS6	90.2	0	100
SARS9	96.5	0	100
SARS10[b]	96.3	0	100
SARS12[b]	95.7	0	100

[a] The results are based on 28,720 bases for each SARS GeneChip. The last 1,004 bases were excluded from the data analysis.
[b] GeneChips were rehybridized with freshly prepared hybridization cocktail using the same target cDNA long-range RT-PCR.

equilibrated at 45°C for 5 min. The prehybridization buffer was removed and replaced with equilibrated hybridization solution and placed in a GeneChip hybridization oven rotating at 60 rpm for 16 h at 45°C. After completion of hybridization, the hybridization solution was removed and the GeneChips were completely filled with 200 μl of nonstringent buffer (6× sodium phosphate buffer [SSPE; 1× SSPE is 0.18 M NaCl, 10 mM NaH_2PO_4, and 1 mM EDTA, pH 7.7, 0.01% Tween 20]).

Washing and staining were carried out in a GeneChip FS-450 fluidics station (Affymetrix). Staining was done twice with a solution containing 6× SSPE, 0.01% Tween 20, 2 mg/ml acetylated BSA, and 10 μg/ml streptavidin–R-phycoerythrin conjugate (SAPE). One additional cycle was completed with an antibody-wash mixture (6× SSPE, 0.01% Tween 20, 2 mg/ml acetylated BSA, 3 μg/ml biotinylated antistreptavidin, and 100 μg/ml goat immunoglobulin G) to remove excess SAPE. Finally, the hybridized, washed, and stained GeneChips were scanned on a GeneChip scanner (Affymetrix). The GeneChip operating software version 3.0.2 (GCOS) program (Affymetrix) was used to operate both the fluidics station and scanner.

SARS GeneChips and data analysis. The SARS resequencing GeneChip consists of eight unique 25-mer probes per base position, which is varied at the central position to incorporate each possible nucleotide (A, C, G, or T) to detect both known as well as novel single-nucleotide polymorphisms (SNPs). The published sequence (AY274119) from the Canada Genome Science Centre was tiled as the reference sequence, and additional variants were identified by tiling three other published sequences (AY278741, AY278491, and AY278554). The data were analyzed by GeneChip DNA analysis software (GDAS), which employs the ABACUS algorithm of Cutler et al. (4). ABACUS is an automated statistical system that provides quality scores to individual genotypes and determines whether the site is polymorphic or not; in the GDAS implementation, it can analyze the Affymetrix GeneChip hybridization data and provides automatic calls. It can be applied in experiments with diploid or haploid target sequences. The algorithm of this program provides specific models for the presence or absence of various genotypes in the samples. There are 5 genotype models (A, C, G, T, and no call) for haploid data sets and 11 genotype models (A, C, G, T, AC, AG, AT, CG, CT, GT, and no call) for diploid data sets (4). The data were also examined by RATools, which is an implementation of the ABACUS algorithm and provides a rigorous framework for the analysis of resequencing GeneChips (4). The open source of RATools is publicly available at http://www.dpgp.org/. Multiple alignments of the SARS-CoV nucleotide sequences were carried out with the DNAsp (9) and CLUSTALX (10) programs.

RESULTS

A total of nine Affymetrix SARS GeneChips were hybridized to resequence the two strains of SARS-CoV infecting humans (four with strain Urbani and five with strain CDC#200301157; Table 1). These strains had been sequenced before by us to high redundancy by capillary-based dideoxy sequencing on an ABI 3730 sequencer. For approximately 60 kb of the two SARS genomes generated by capillary sequencing, the average consensus phred quality value was 90 with 9.1-fold average redundancy (data not shown). Nucleotide sequences obtained by GeneChips (present study) were compared with previously generated sequences. The PCR amplification of the last 1,004 bases for both strains was inconsistent. Therefore, this portion of the genome was excluded from the data analysis while evaluating the performance of SARS resequencing GeneChip hybridization results in the study.

A small amount of purified RT-PCR products (11.6 μl with a 100-ng/μl concentration) was required to hybridize each GeneChip. However, during hybridization some of the GeneChips either completely failed to hybridize due to improper fragmentation, labeling, or washing or in some cases only partially hybridized, with a large number of uncalled bases being evident (data not shown). For these GeneChips, attempts were made to rehybridize with the same target cDNA previously used in hybridization experiments. Rehybridization of all five SARS GeneChips with fresh target was successful (Table 1).

The base calls were determined with GDAS and RATools for each of the hybridized SARS GeneChips. These two programs are implementations of ABACUS (4, 15). Since SARS is an RNA virus, the homozygote model was selected for the GDAS program at default settings to analyze the resequencing data. While using RATools for the base call analysis, the total threshold was set at 30 and the strand threshold was set at −2 (4, 15). The default parameters values for both GDAS and RATools were roughly similar, which provided similar levels of base calling performance (4, 15). A base was scored as an "uncalled base" (N) when the ABACUS algorithm could not make a confident call for the same position on replicates. The "discordant calls" were those that varied between GeneChips and the data obtained from conventional sequencing. The percentage of nucleotide sequence with high confidence across the genome was the "call rate," and the "accuracy" was the total number of correct calls excluding the uncalled bases (Table 1).

Although each SARS GeneChip was capable of resequencing 29,724 bp of the complete genome of SARS-CoV, the present study includes the analysis of the results for 28,720 bp per GeneChip, excluding the last 1,004 bases of the SARS genome. Nine of the successfully hybridized SARS GeneChips revealed 94.6% bases of good quality calls (244,522 bp out of a total possible of 249,691 bp). The data generated by GDAS were similar to the RATools-based data. The distribution of quality scores across these base calls is shown in Fig. 1. The remainder of uncalled bases (Ns) was perhaps due to failure of long amplicons generated by RT-PCR. The repeatability and accuracy of ABACUS calls were evaluated by employing replicate experiments of independent amplifications of a SARS strain followed by hybridization with Affymetrix SARS GeneChips (Table 1).

At the homozygote default settings, when the hybridized SARS GeneChip was analyzed individually or subjected to batch analysis or when all nine GeneChips were analyzed collectively, the number of uncalled bases (Ns) did not change or decrease. Within the successful hybridized region of GeneChips that produced good quality calls, the number of uncalled

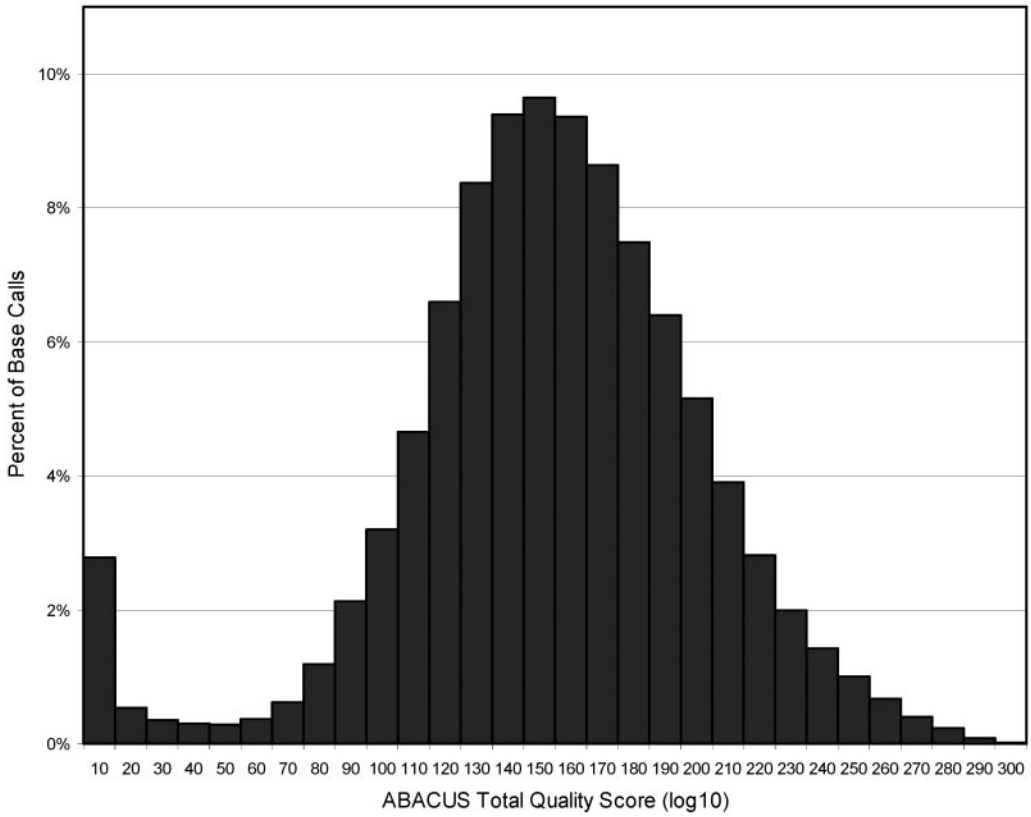

FIG. 1. ABACUS quality scores for base calls in SARS-CoV. A quality score measures the difference in $\log_{10}$ units between the likelihood level for the best base call model and that for the second best model (4, 15). Among all the bases, 94.6% possess quality scores that exceeded the threshold used in the study.

bases varied from 261 to 1,108, with an average of 502.0 ± 157.2 per GeneChip (Fig. 2). Further, the numbers of uncalled bases were almost the same in GeneChips hybridized with the two SARS strains. Nevertheless, there was no improvement in call rate with increase in the number of replicates. The uncalled bases of some of the SARS GeneChips were manually verified by comparing the probe intensity (including the A, C, G, and T bars of both forward and reverse strands) of the

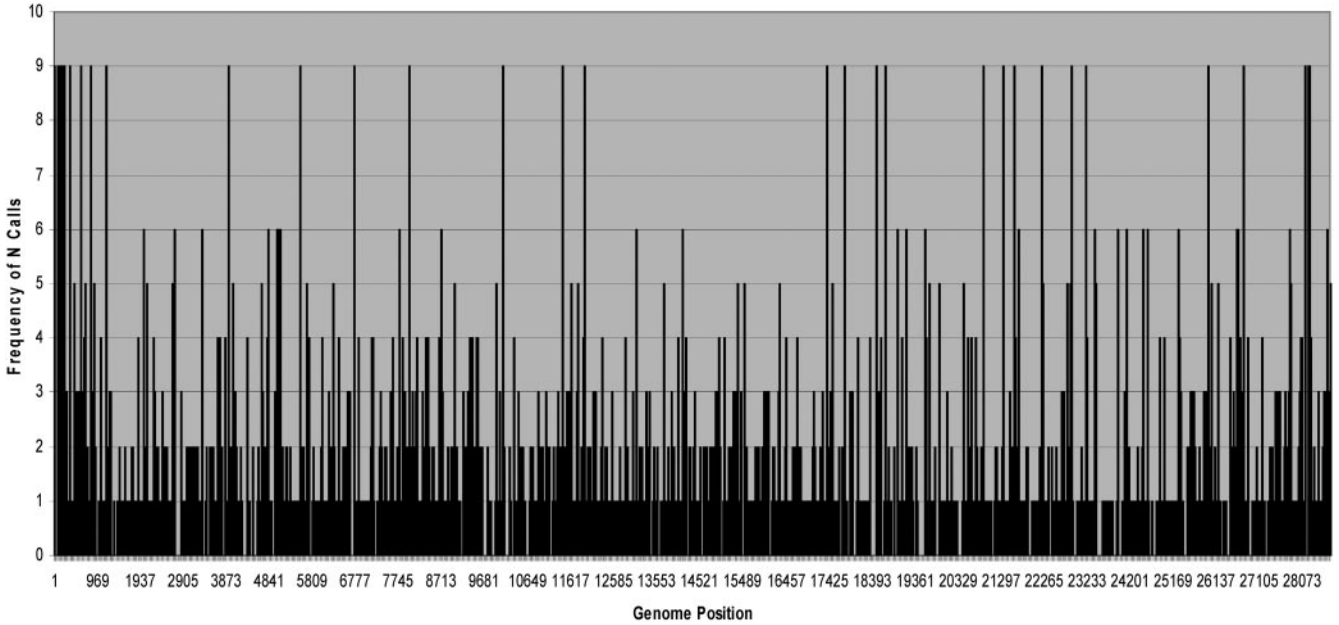

FIG. 2. Distribution and frequency of uncalled bases across the nine SARS-CoV genomes resequenced.

TABLE 2. Genetic diversity across the five human SARS-CoV genomes

Position	Residue at position given[c]				
	AY274119 (Canada)	AY278741 (Urbani)[a]	AY714217 (CDC)[b]	AY278491 (HKU)	AY278554 (CHU)
2601	T	T	T	C	T
7746	G	G	G	G	T
7919	C	T	C	C	C
7930	G	G	G	A	G
8387	G	G	G	C	G
8417	G	G	G	C	G
9404	T	T	T	T	C
9479	T	T	T	T	C
13494	G	G	G	A	G
13495	T	T	T	G	T
16622	C	T	C	C	C
17564	T	T	T	T	G
17846	C	C	C	C	T
18065	G	G	G	A	G
18974	A	A	G	A	A
19064	A	G	A	A	G
21721	G	G	G	G	A
22222	T	T	T	T	C
23220	G	T	T	T	T
23445	A	A	G	A	A
24872	T	C	T	T	T
25298	A	G	G	G	G
25569	T	T	T	A	T
26600	C	C	C	T	C
26857	T	C	T	T	T
27827	T	T	T	T	C

[a] Data based on four SARS GeneChips hybridized.
[b] Data based on five SARS GeneChips hybridized.
[c] HKU, Hong Kong SARS sequence from the University of Hong Kong; CHU, Hong Kong sequence from the Chinese University of Hong Kong.

respective cell, and at none of the positions examined was the genetic polymorphism discernible (data not shown). Furthermore, no discordant calls were evident among the nine hybridized SARS GeneChips. Nevertheless, the number of discordant calls increased when the default homozygote model resequencing algorithm settings of GDAS were changed to more permissive base-calling algorithm settings (data not shown). Consequently, it is important to note that across the nine SARS-hybridized GeneChips, 100% accuracy was recorded (i.e., there was less than 1 error in a total of 244,522 bp called). Furthermore, these high-confidence nondiscordant calls covered 90.2 to 96.5% of the genomes.

Multiple alignments on the nine nearly complete genome sequences derived for CDC#200301157 and Urbani strains (this study) were aligned with the five published (AY274119, AY278741, AY714217, AY278491, and AY278554) human coronavirus sequences. The analysis revealed a distinctly conserved pattern across the genomes of human SARS coronaviruses. However, 26 SNPs were evident across the five SARS-CoV strains. The CDC#200301157 and Urbani strains had 4- and 7-point mutations, respectively, compared with the reference sequence from the Canada Genome Science Centre. Strain CDC#200301157 differed from the Urbani strain at seven positions (Table 2). The analysis also confirmed that the sequences generated by GeneChips for the two strains were identical to those previously generated by conventional sequencing.

DISCUSSION

The GeneChip-based hybridization assay is an emerging technology that has been used to study gene expression, SNP detection, mapping, genetic linkage, and genetic polymorphism in various organisms. The earlier versions of GeneChips were developed with a 35- by 35-μm feature size. Such GeneChips containing 25-mer probes complementary to the human mitochondrial genome (16.2 kb of sense and antisense DNAs, 30.5 kb total) were used to identify SNPs in human genomes (2, 12). Subsequently, GeneChips with higher density and smaller feature size (20 by 24 μm) were developed for resequencing and SNP detection that could screen 30 kb sense DNA and 30 kb antisense DNA on each GeneChip (13). The data derived from GeneChips were validated by sequencing the fragments on ABI sequencers and were found to be comparable (5). More recently, a high-throughput platform of similar capacity was developed by NimbleGen (Madison, Wis.) that has the capacity to interrogate the complete genome of SARS-CoV (14). The SARS resequencing GeneChips used in this study had about 60-kb sense and antisense sequences complementary to SARS-CoV with a 20- by 25-μm feature size.

The high call rates obtained for SARS resequencing GeneChips (this study) were comparable to those reported for *Bacillus anthracis* resequencing GeneChips (15). However, the present study also illustrates a vexing problem with GeneChip resequencing, namely, the propensity for large numbers of uncalled bases (though not necessarily wrong calls) by the GDAS software at the Affymetrix resequencing platform. While resequencing some SARS isolates at the NimbleGen platform, which also employed the ABACUS algorithm for data analysis, the uncalled bases were also obvious (14). It was suggested that probes with low G/C contents hybridized weakly and in turn produced insufficient signals for base calling (14). Lower base calling has been reported in the regions of Affymetrix GeneChips with elevated frequencies of purines, which may reflect synthesis differences in the chemistries used by Nimblegen and Affymetrix (4).

Minor differences were noticed in the cDNA sequences obtained by Affymetrix GeneChip and ABI cycle sequencing for human immunodeficiency virus type 1 samples. The Affymetrix GeneChips failed to detect the length polymorphism as the Affymetrix methodology utilizes the hybridization technique containing a defined array of oligonucleotides (11). Ambiguity in the results was explained by the Affymetrix GeneChip design being based on the sequence of a human immunodeficiency virus type 1 strain that lacked length polymorphism (11). Thus, the Affymetrix GeneChips can do resequencing but not de novo sequencing, and if a particular variation is not tiled as a reference, the additional insertion or deletion cannot be detected after hybridization, which should be confirmed by the conventional sequencing.

Regardless of the uncalled bases, it is evident that the genetic variations in two SARS-CoV genomes characterized by SARS resequencing GeneChips are valid and are indicative of low genetic diversity across the genomes (Table 2). Multiple alignments on sequences of both SARS strains generated by Affymetrix SARS resequencing GeneChips (this study) and ABI sequencer (previous study) were identical. The SARS GeneChips should be useful in tracking genomic changes of

the pathogens over time and geography. These resequencing GeneChips were also found to be advantageous over the conventional sequencing in which a larger amount of genomic material is required for both amplification and sequencing. Thus, in future, this tool may be used more frequently to characterize clinical samples when the genetic material (DNA or RNA) is a limiting factor.

In conclusion, the SARS resequencing GeneChip can be utilized to obtain genomic sequences of SARS-CoV. The GeneChip-based genome characterization was rapid and sensitive. Nucleotide sequences generated by these GeneChips were reproducible and precise over 94% of the genome. The SARS resequencing GeneChips were successful in detection of all known SNPs in the SARS-CoV isolates characterized. Resequencing by GeneChip hybridization should help in understanding the transmission dynamics and epidemiology of SARS-CoV strains isolated from humans in different geographic locations and time points. This new technique may be generalized as an effective tool for resequencing human-pathogenic genomes of public health importance in the future.

REFERENCES

1. **Centers for Disease Control and Prevention.** 2003. Outbreak of severe acute respiratory syndrome—worldwide. Morb. Mortal. Wkly. Rep. **52:**226–228.
2. **Chee, M., R. Yang, E. Hubbell, A. Berno, X. C. Huang, D. Stern, J. Winkler, D. J. Lockhart, M. S. Morris, and S. P. Fodor.** 1996. Accessing genetic information with high-density DNA arrays. Science **274:**610–614.
3. **Chinese SARS Molecular Epidemiology Consortium.** 2004. Molecular evolution of the SARS coronavirus during the course of the SARS epidemic in China. Science **303:**1666–1669.
4. **Cutler, D. J., M. E. Zwick, M. M. Carrasquillo, C. T. Yohn, K. P. Tobin, C. Kashuk, D. J. Mathews, N. A. Shah, E. E. Eichler, J. A. Warrington, and A. Chakravarti.** 2001. High-throughput variation detection and genotyping using microarrays. Genome Res. **11:**1913–1925.
5. **Fan, J. B., D. Gehl, L. Hsie, N. Shen, K. Lindblad-Toh, J. P. Laviolette, E. Robinson, R. Lipshutz, D. Wang, T. J. Hudson, and D. Labuda.** 2002. Assessing DNA sequence variations in human ESTs in a phylogenetic context using high-density oligonucleotide arrays. Genomics **80:**351–360.
6. **Ksiazek, T. G., D. Erdman, C. S. Goldsmith, S. R. Zaki, T. Peret, S. Emery, S. Tong, C. Urbani, J. A. Comer, W. Lim, P. E. Rollin, S. F. Dowell, A. E. Ling, C. D. Humphrey, W. J. Shieh, J. Guarner, C. D. Paddock, P. Rota, B. Fields, J. DeRisi, J. Y. Yang, N. Cox, J. M. Hughes, J. W. LeDuc, W. J. Bellini, L. J. Anderson et al.** 2003. A novel coronavirus associated with severe acute respiratory syndrome. N. Engl. J. Med. **348:**1953–1966.
7. **Marra, M. A., S. J. Jones, C. R. Astell, R. A. Holt, A. Brooks-Wilson, Y. S. Butterfield, J. Khattra, J. K. Asano, S. A. Barber, S. Y. Chan, A. Cloutier, S. M. Coughlin, D. Freeman, N. Girn, O. L. Griffith, S. R. Leach, M. Mayo, H. McDonald, S. B. Montgomery, P. K. Pandoh, A. S. Petrescu, A. G. Robertson, J. E. Schein, A. Siddiqui, D. E. Smailus, J. M. Stott, G. S. Yang, F. Plummer, A. Andonov, H. Artsob, N. Bastien, K. Bernard, T. F. Booth, D. Bowness, M. Czub, M. Drebot, L. Fernando, R. Flick, M. Garbutt, M. Gray, A. Grolla, S. Jones, H. Feldmann, A. Meyers, A. Kabani, Y. Li, S. Normand, U. Stroher, G. A. Tipples, S. Tyler, R. Vogrig, D. Ward, B. Watson, R. C. Brunham, M. Krajden, M. Petric, D. M. Skowronski, C. Upton, and R. L. Roper.** 2003. The genome sequence of the SARS-associated coronavirus. Science **300:**1399–1404.
8. **Rota, P. A., M. S. Oberste, S. S. Monroe, W. A. Nix, R. Campagnoli, J. P. Icenogle, S. Penaranda, B. Bankamp, K. Maher, M. H. Chen, S. Tong, A. Tamin, L. Lowe, M. Frace, J. L. DeRisi, Q. Chen, D. Wang, D. D. Erdman, T. C. Peret, C. Burns, T. G. Ksiazek, P. E. Rollin, A. Sanchez, S. Liffick, B. Holloway, J. Limor, K. McCaustland, M. Olsen-Rasmussen, R. Fouchier, S. Gunther, A. D. Osterhaus, C. Drosten, M. A. Pallansch, L. J. Anderson, and W. J. Bellini.** 2003. Characterization of a novel coronavirus associated with severe acute respiratory syndrome. Science **300:**1394–1399.
9. **Rozas, J., and R. Rozas.** 1995. DnaSP, DNA sequence polymorphism: an interactive program for estimating population genetics parameters from DNA sequence data. Comput. Appl. Biosci. **11:**621–625.
10. **Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins.** 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **24:**4876–4882.
11. **Vahey, M., M. E. Nau, S. Barrick, J. D. Cooley, R. Sawyer, A. A. Sleeker, P. Vickerman, S. Bloor, B. Larder, N. L. Michael, and S. A. Wegner.** 1999. Performance of the Affymetrix GeneChip HIV PRT 440 platform for antiretroviral drug resistance genotyping of human immunodeficiency virus type 1 clades and viral isolates with length polymorphisms. J. Clin. Microbiol. **37:**2533–2537.
12. **Wang, D. G., J. B. Fan, C. J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins, E. Winchester, J. Spencer, L. Kruglyak, L. Stein, L. Hsie, T. Topaloglou, E. Hubbell, E. Robinson, M. Mittmann, M. S. Morris, N. Shen, D. Kilburn, J. Rioux, C. Nusbaum, S. Rozen, T. J. Hudson, and E. S. Lander.** 1998. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science **80:**1077–1082.
13. **Warrington, J. A., N. A. Shah, X. Chen, M. Janis, C. Liu, S. Kondapalli, V. Reyes, M. P. Savage, Z. Zhang, R. Watts, M. DeGuzman, A. Berno, J. Snyder, and J. Baid.** 2002. New developments in high-throughput resequencing and variation detection using high density microarrays. Hum. Mutat. **19:**402–409.
14. **Wong, C. W., T. J. Albert, V. B. Vega, J. E. Norton, D. J. Cutler, T. A. Richmond, L. W. Stanton, E. T. Liu, and L. D. Miller.** 2004. Tracking the evolution of the SARS coronavirus using high-throughput, high-density resequencing arrays. Genome Res. **14:**398–405.
15. **Zwick, M. E., F. Mcafee, D. J. Cutler, T. D. Read, J. Ravel, G. R. Bowman, D. R. Galloway, and A. Mateczun.** 17 December 2004, posting date. Microarray-based resequencing of multiple *Bacillus anthracis* isolates. Genome Biol. **6:**R10. [Online.] http://genomebiology.com.