*Proc. Natl. Acad. Sci. USA*
Vol. 96, pp. 10016–10020, August 1999
Biochemistry

# Chip-based genotyping by mass spectrometry

**(DNA chip/single nucleotide polymorphism)**

KAI TANG*, DONG-JING FU†, DOMINIQUE JULIEN, ANDREAS BRAUN, CHARLES R. CANTOR, AND HUBERT KÖSTER

Sequenom Inc., 11555 Sorrento Valley Road, San Diego, CA 92121

*Contributed by Charles R. Cantor, July 12, 1999*

**ABSTRACT** **Silicon chips with immobilized target DNAs were used for accurate genotyping by mass spectrometry. Genomic DNAs were amplified with PCR, and the amplified products were covalently attached to chip wells via *N*-succinimidyl (4-iodoacetyl)aminobenzoate (SIAB) chemistry. Primer annealing, extension, and termination were performed on a 1-μl scale directly in the chip wells in parallel. Diagnostic products thus generated were detected *in situ* by using matrix-assisted laser desorption ionization mass spectrometry. This miniaturized method has the potential for accurate, high-throughput, low-cost identification of genetic variations.**

The advances of the Human Genome Project hold promise that a high-quality sequence of the human genome will be completed by the end of 2003 (1). This will provide unique opportunities for studying genetic variation in humans and its relationship with disease risk. The most common DNA sequence variations, single nucleotide polymorphisms (SNPs), are stable and widely scattered across the genome. Once constructed, a high-density SNP map of several hundred thousand markers will be an indispensable tool for genome-wide association studies to identify genes that contribute to disease risk and individual differences in drug response. To facilitate large-scale SNP identification, new technologies are being developed to replace gel-based resequencing. In one approach, highly redundant, sequence-specific oligonucleotide arrays have been used to hybridize against fluorescently labeled DNA targets. The hybridization patterns are scanned for possible mismatches in sequences (2–5).

We present here a different approach that combines mass spectrometric detection with enzymatic extension of primers hybridized to immobilized DNA target arrays. The advantage of this combination is high specificity and unparalleled accuracy of allele identification. In previous studies we have shown that immobilized DNA strands, either through biotin-streptavidin binding to magnetic or controlled-pore glass beads (6) or through covalent attachment to silicon surfaces (7), survive laser irradiation during matrix-assisted laser desorption ionization (MALDI) mass spectrometry (MS). Only the intact hybridized strands can be detected. Subsequent developments led to solid-phase enzymatic reactions and MALDI-MS detection for DNA sequencing (8, 9) and mutation/polymorphism analyses (10, 11). The primer–oligo base extension (PROBE) assay was designed to target sequence variations at specific sites (10). Because it generates products only a few bases longer than primers, PROBE is ideal for SNP identification using MALDI-MS. Other technologies have been developed to facilitate MALDI-MS in a miniaturized-array format by nano-dispensing of matrix solution and diagnostic products onto a silicon chip (12, 13). In this report, we demonstrate that PCR products can be covalently attached to a silicon chip and PROBE reactions can be performed successfully in the chip wells on a 1-μl scale without the use of conventional sample tubes or microtiter plates. Upon nano-dispensing of matrix into the chip wells, diagnostic products can be detected by MALDI-MS *in situ* such that homozygous as well as heterozygous alleles can be identified accurately. The key issue of SNP identification for high-throughput genomics is the cost per SNP. By transferring a microtiter plate/bead-based assay to a chip-based assay, the sample volume needed for analysis can be reduced from microliters to nanoliters. This kind of scale-down is essential for cost reduction.

## MATERIALS AND METHODS

Oligonucleotides including 5′ thiolated or biotinylated materials were purchased from Operon Technologies (Alameda, CA). Their quality was checked by using MALDI-MS. Individual genomic DNA was extracted from anonymous blood samples according to standard procedures by using a Puregene DNA isolation kit (Gentra, Minneapolis, MN). Silicon wafers ($20 \times 20$ mm$^2$) with 36 wells were used.

**PCR Amplification.** The primers used, including types of modification, are listed in Table 1. To generate thiolated PCR products, 50 μl of PCRs were performed containing 30~120 ng of genomic DNA, 2 units of AmpliTaq Gold (Applied Biosystems), 10 pmol forward primer, 10 pmol reverse primer, 0.2 mM dNTPs, 2 mM $MgCl_2$, and 1× enzyme buffer. For PCRs with universal primers, 10 pmol universal primers and 1 pmol universal-tagged gene-specific primers were used. The template was denatured at 94°C for 2 min, followed by 12 touchdown cycles [94°C, 20 sec; 56°C (0.5°C decrease for each cycle until 50°C), 20 sec; 72°C, 30 sec], 35 regular cycles (94°C, 20 sec; 50°C, 20 sec; 72°C, 30 sec), and a final incubation at 72°C for another 2.5 min. The PCR products were purified by using a High Pure PCR product purification kit (Boehringer Mannheim) and eluted with 1× TE buffer (10 mM Tris/1 mM EDTA, pH 8.0).

**Generation of Single-Stranded PCR Products.** A suspension of 1.5 mg of streptavidin-coated magnetic beads (Dynal, Great Neck, NY) in 150 μl of 5× binding and washing buffer (5 M $NH_4Cl$/50 mM Tris·HCl, pH 8.0) was added to 400~500 μl of PCR products. After incubation on a shaker at 37°C for 30 min, the supernatant was removed, and the beads were washed once with 10 mM Tris·HCl, pH 8.0 and then resuspended in 40 μl of freshly diluted 0.1 M NaOH (from 1.000 M stock). The mixture was incubated at room temperature for 5 min, and the supernatant was collected. The beads were washed again with 10 μl of 0.1 M NaOH, and the supernatant was collected and combined with the supernatant from the first wash. The solution then was neutralized by adding 50 μl of 0.1 M HCl

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

PNAS is available online at www.pnas.org.

Abbreviations: SNP, single nucleotide polymorphism; HPA, human platelet alloantigen(s); MALDI-MS, matrix-assisted laser desorption ionization mass spectrometry; PROBE, the primer–oligo base extension.

*To whom reprint requests should be addressed. E-mail: ktang@sequenom.com.

†Present address: Eli Lilly and Company, Lilly Corporate Center, Drop Code 0444, Indianapolis, IN 46285.

Table 1. PCR primers

| Antigen | Forward primers | Reverse primers | Product length, bp |
|---|---|---|---|
| HPA-1 | 5′ss-HPA-1F (5′ ss-TTCTGATTGCTGGACTTCTC) | 5′bio-HPA-1R (5′bio-GGAGTGCAATCCTCTGGG) | 224 |
| HPA-2 (ref. 16) | 5′bio-HPA-2F (5′bio-GCCTGCTCTCACCGTCCT) | 5′ss-HPA-2R (5′-ss-AGTCAAGTTGTTGTTAGCCA) | 179 |
| HPA-4 | 5′bio-US-1 (5′-bio-GCC AGG GTT TTC CCA GTC AC)<br>US-1-HPA-4F (5′-GCC AGG GTT TTC CCA GTC ACC GGC AGG TGG AGG ATT AC) | 5′ss-US-2 (5′ss-CGA ATT CGA GCT CGG TAC CC)<br>US-2-HPA-4R (5′-CGA ATT CGA GCT CGG TAC CCA CTT ACT CAT AGC AGG GGT TTT) | 270 |

bio, biotin; 5′ss, 5′-HO-$(CH_2)_6$-S-S-$(CH_2)_6$)-.

(freshly diluted from 1.000 M HCl). The above 100-μl solution was precipitated by adding 10 μl of 3 M $NH_4OAc$ (pH 4.5), 1 μl of 10 ng/μl glycogen (Sigma), and 220 μl of ethanol and incubation at 4°C for more than 1 hr. The pellet DNA was collected by centrifuging at 13,000 × *g* for 12 min followed by washing once with 70% ethanol and dried.

**Reducing the 5′ Disulfide DNA to a Terminal Free Thiol.** The purified PCR products (from 250 μl of original crude PCR, approximately 25 pmol, single-stranded or double-stranded) were dissolved in 3.8 μl of $ddH_2O$, to which 0.22 μl of 20× TCEP buffer (1 M $KH_2PO_4$/20 mM EDTA, pH 4.5) and 0.44 μl of 100 mM TCEP [5.74 mg tris(2-carboxyethyl)phosphine (Pierce) freshly dissolved in 200 μl of $ddH_2O$] were added. The final 4.46 μl of solution was incubated in an Eppendorf tube for 5 hr at 37°C.

**Amination of Si Chips.** Silicon chips (Accelerator Technology, Bryan, TX) were washed in ethanol, dropped into a vial containing 3 ml of solution of 3-aminopropyltriethoxysilane (Aldrich) in anhydrous toluene (3-APTES/toluene, 1:3 in volume). After 3 hr, the chips were washed with toluene followed by DMSO.

***N*-Succinimidyl (4-Iodoacetyl)aminobenzoate (SIAB) Functionalization of Chips.** The aminated chips then were treated as described (7).

**Coupling Thiol-DNA to SIAB Chip.** To the 4.46 μl of reduced thio-DNA, 1 μl of 1 M $KH_2PO_4$, pH 8, and 4.54 μl of $H_2O$ were added. Then 1∼1.5 μl of the mixture was applied to each well of the 6 × 6 chip and kept overnight at room temperature. Finally, the chip was washed with $H_2O$ and treated with 5× SSC/50% formamide solution followed by washing with water.

**Generating Single-Stranded DNA Template for PROBE Reactions.** After double-stranded DNA templates were immobilized on the chip, the noncovalently attached strand was denatured by immersing the chip in 0.1 M NaOH and incubating at room temperature for 5 min. The chip then was washed once with 0.1 M NaOH, three times with water, and twice with 10 mM Tris·HCl.

**PROBE Reactions on the Chip.** In each well containing covalently attached, single-stranded DNA templates, 1∼1.5 μl of reaction mix was added containing 5∼6.75 pmol primer, 2 units of thermosequenase (Amersham Pharmacia), 50 μM dNTPs, 50 μM ddNTPs (see Table 2), and 1× thermosequenase buffer. The chip then was placed in an *in situ* PCR instrument (MJ Research, Cambridge, MA). The following program was used: 80°C for 1 min, 50°C for 1 min, followed by ramping from 50°C to 72°C at the rate of 0.1°C/sec, and then 72°C for 2 min, 50°C for 1 min, and 25°C for 5 min. After the reaction, the chip was washed three times with 10 mM Tris·HCl, pH 8.0, and three times with 70 mM ammonium citrate and then left to dry.

**Matrix Dispensing.** Matrix stock solution was made by dissolving 69.5 mg of 3-hydroxypicolinic acid (Aldrich) and 11.3 mg of ammonium citrate in 700 μl of 50% acetonitrile. It was diluted by adding 1.5 vol of pure water before use. A piezo-electric pipette (Microdrop, Norderstedt, Germany) was used to dispense matrix (10 drops or ≈3 nl) in each well.

**MALDI-MS.** A linear time-of-flight mass spectrometer with delayed extraction (Voyager DE; PerSeptive Biosystems, Framingham, MA) was used. The target holder was modified to hold the 20 × 20-$mm^2$ chip. A nitrogen laser firing at 3 Hz was used. The target and middle grids were kept at +18.2 kV for 200 nsec after each laser shot. Then, the target potential was raised to +20 kV. The ion guide in the flight tube was kept at −2 V. The final spectra were smoothed by a 19-point Savisky–Golay average.

## RESULTS AND DISCUSSIONS

Human platelet alloantigens (HPA-1, HPA-2, and HPA-4) were chosen as models for genotyping because they are excellent examples with which to demonstrate the feasibility of highly accurate, chip-based clinical diagnostics. Those alloantigens are involved in diseases such as neonatal alloimmune thrombocytopenia, posttransfusion purpura, and refractoriness to platelet transfusions in which platelet-specific antibod-

Table 2. PROBE primers and products

| Reaction mix | Alleles (frequencies*) | Product sequences | Expected $[M+H]^+$, Da |
|---|---|---|---|
| HPA-1P ddA/ddC/dG/ddT | Primer | 5′-GGTCACAGCGAGGTGAGCCC | 6,169.0 |
| | 33L (81.2%) | 5′-GGTCACAGCGAGGTGAGCCCA | 6,466.2 |
| | 33P (18.8%) | 5′-GGTCACAGCGAGGTGAGCCCGGA | 7,124.6 |
| HPA-2P ddA/ddC/dG/dT | Primer | 5′-CCTGCCCCCAGGGCTCCTGA | 6,016.0 |
| | 163T (91.4%) | 5′-CCTGCCCCCAGGGCTCCTGAC | 6,289.2 |
| | 163M (8.6%) | 5′-CCTGCCCCCAGGGCTCCTGATGC | 6,922.6 |
| HPA-4P ddA/ddC/dG/ddT | Primer | 5′-GCTGGCCACCCAGATGC | 5,157.4 |
| | 143R (>99.5%) | 5′-GCTGGCCACCCAGATGCGA | 5,783.8 |
| | 143Q (<0.5%) | 5′-GCTGGCCACCCAGATGCA | 5,454.6 |

*Allelic frequencies according to ref. 16.

ies may be involved (14–16). DNA fragments containing the polymorphic sites were amplified by the PCR. The primers used and the corresponding product lengths are listed in Table 1. The specific polymorphisms then were interrogated by PROBE reactions (10). Diagnostic products that were generated, which were only a few bases longer than corresponding PROBE primers, were analyzed by MALDI-MS. The PROBE primers used and their allele-specific products with expected mass values are listed in Table 2.

To transform currently established microtiter plate/bead-based assays (9) to a chip-based format, we chose a silicon chip with 36 wells (in a 6 × 6 array) as our immobilization and enzymatic reaction surface as well as the surface for subsequent MALDI-MS. The experimental protocol is shown schematically in Fig. 1. We prepared PCR products (with one primer thiolated) off-line and then attached these covalently onto individual wells of the silicon chip via a modification of established silicon dioxide derivatization chemistry (7). The immobilized double-stranded PCR products then were denatured, leaving the covalently attached single-stranded DNA in the chip wells for further reaction. Alternatively, the nonthiolated PCR primer was biotinylated, and the PCR products were denatured before attachment to the silicon surface. The biotinylated strand was collected by streptavidin-coated magnetic beads, and the thiolated strand in the supernatant was used for covalent attachment.

In each chip well containing immobilized single-stranded DNA template, a mixture of selected dNTPs and ddNTPs, PROBE primer was added followed by thermosequenase, and the chip was placed in a chamber made from a microscope slide and seal film (MJ Research). The chamber was positioned in an *in situ* PCR machine (MJ Research), and the PROBE reaction was carried out. After removing enzyme and buffer salts by washing, matrix solution was loaded onto the chip well by using a piezo-electric pipette. The chip then was transferred into the mass spectrometer and analyzed directly by MALDI-MS.

Genotyping of codon 163 (T163M) in HPA-2 and codon 33 (L33P) in HPA-1 was studied in further detail. Results from two different individuals with homozygous 163T and heterozygous 163T/M in HPA-2, respectively, are shown in Fig. 2. The allele frequency for 163M is relatively low; for instance, it is 8.6% in the German population (16). In all cases, the PROBE primers were not fully extended, contrary to most PROBE reactions in solution. This may be caused by nonspecific binding of enzyme to the oxidized silicon surface or surface poisoning of enzyme activity. However, the unextended primer can serve as an internal standard for accurate mass calibration of the allele-specific peaks. Generally, the mass accuracy of diagnostic products measured on the chips is better than 0.02%.

Each well of the 6 × 6 silicon chip held about 1 μl of liquid and had a surface area of 6.25 $mm^2$/well. From previous experiments (7) we assumed that maximum loading per well of PCR product would be about 1.56 pmol. Our standard 50-μl PCR usually generates approximately 5 pmol of amplified products. Half of the original crude PCR product (25 μl or 2.5 pmol) was used in each well, which would provide roughly a 1.6-fold excess of DNA for covalent attachment. The amount of PROBE primer used in each well was 5–6.75 pmol, which is about 2- to 3-fold excess over the maximum template loading. However, the hybridization efficiency of PROBE primer to the covalently attached template was expected to be no more than 40% based on previous results (7). Therefore,

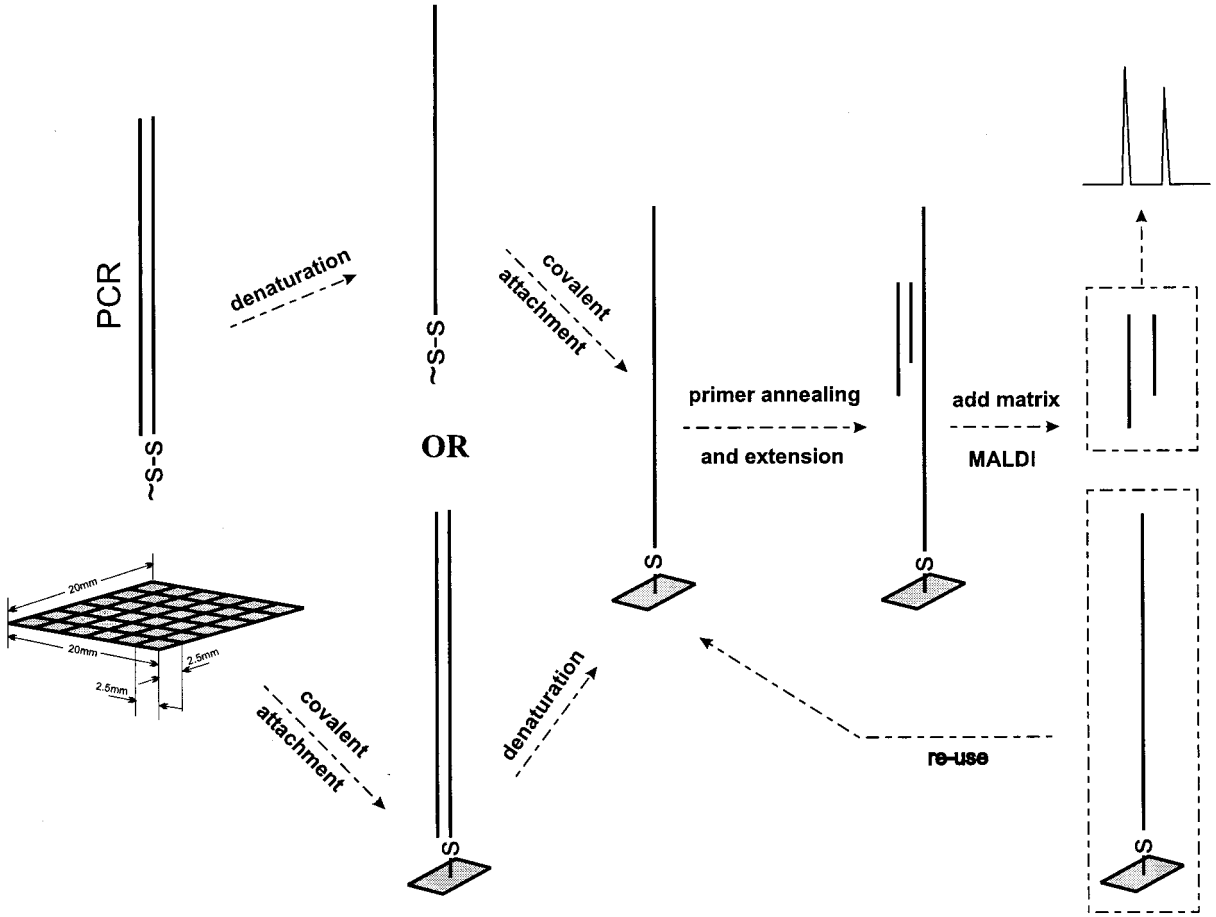


FIG. 1. Schematic overview of PCR product immobilization, PROBE reaction, and MALDI-MS on a chip.

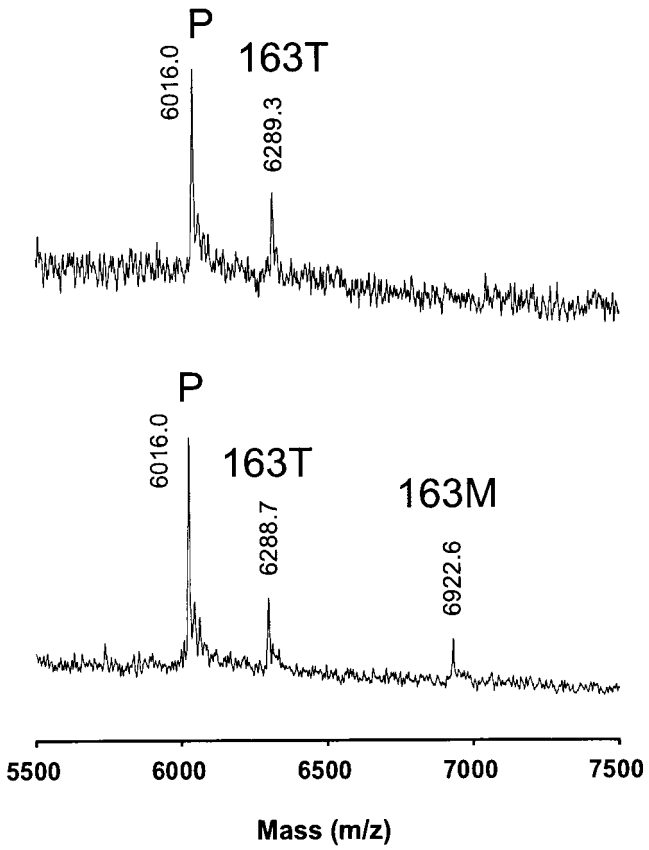


FIG. 2. Mass spectra for genotyping of HPA-2 showing a homozygous sample (*Upper*) and a heterozygous sample (*Lower*). The primer was extended by one base for the 163T allele and three bases for the 163M allele. The unextended primer (P) was present in both cases.

the surface density of PROBE products, including unextended primer, was expected to be no more than 100 fmol/mm$^2$. When matrix solution (pH = 4.0) is added to the surface, about 10% of the hybridized strands are expected to be denatured (17) and cocrystallized with matrix molecules. A piezo-electric pipette was used for matrix dispensing. The small droplet size (~3 nl) contributed to fast solvent evaporation (~10 sec), forming homogeneously distributed small crystals across the area (d $\approx$ 300 $\mu$m) (Fig. 3). Therefore, the amount of single-stranded DNA in the matrix crystals of each droplet spot was estimated to be 0.7 fmol [10 fmol/mm$^2$ $\times$ $\pi$(0.15 mm)$^2$]. In

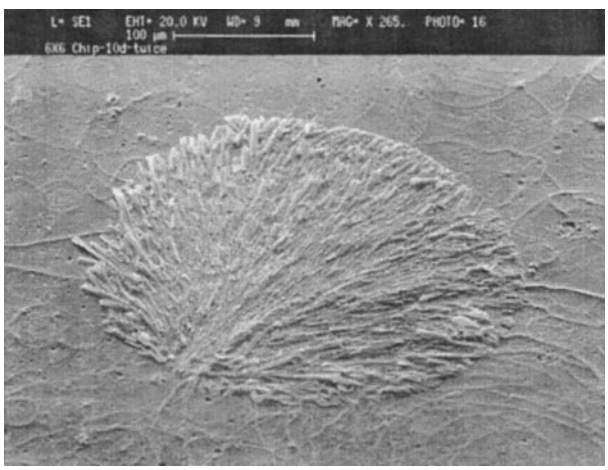


FIG. 3. Scanning electron microscopy of one recrystallized matrix spot formed by dispensing 3 nl of matrix twice.

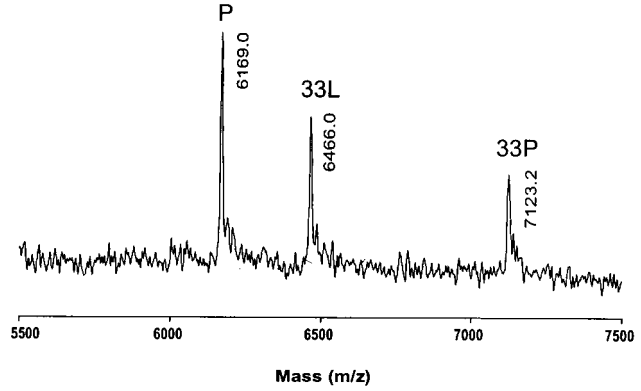


FIG. 4. Mass spectrum for genotyping of HPA-1 showing a heterozygous sample 33L/33P. A recycled HPA-1 chip was used to perform the PROBE reaction and MALDI-MS.

MALDI-MS, the laser was tightly focused (d $\approx$ 100 $\mu$m). Searching for good spot was still necessary to obtain a good spectrum, indicating uneven distribution of DNA in the matrix crystals.

Covalent attachment of target DNA templates on functionalized silicon chips not only facilitates enzymatic reactions and MALDI-MS *in situ*, but it also provides DNA templates that can be reused. After MALDI, the matrix on the chip easily could be rinsed off. Remaining PROBE primers and extension products could be denatured by incubating the chip in 0.1 M NaOH. After washing the chip surface, a new set of primers can be used for PROBE reactions targeting other polymorphic sites. As an example, a recycled chip was used for HPA-1 genotyping. Fig. 4 shows that a heterozygous sample, 33L/33P, was detected unambiguously. Using the recycled chip did not at all compromise the quality of the mass spectrum. By recycling, some significant cost reductions may be achieved.

Another method for cost reduction was to use universal primers for PCR amplification of DNA templates. An example is shown in Table 1 for amplification of HPA-4. Universal-tagged, gene-specific primers US-1-HPA-4F and US-2-HPA-4R were synthesized with predesigned sequences (such as the sequence of the universal sequencing primer) added to 5′ ends of the gene-specific primer sequences. PCRs were carried out at relatively low ratios of these universal-tagged, gene-specific primers (usually 1/10 of the universal primers) to the universal primers US-1 and US-2. In this format, only the

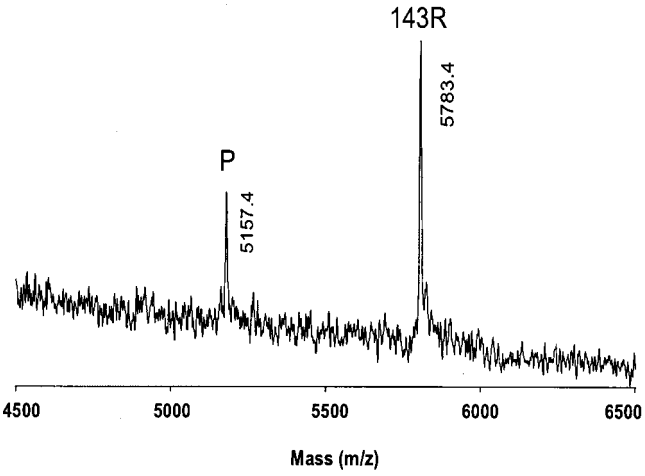


FIG. 5. Mass spectrum for genotyping of HPA-4 showing a homozygous sample 143R. The template DNA was amplified by using universal primers listed in Table 1 before covalent attachment onto the chip surface.

universal primers need to be modified (such as thiolation or biotinylation) for attachment to the chip surface. As a result, the usage of costly gene-specific primers was reduced significantly. Genotyping by using such amplified templates produced results of the same quality as conventional PCR templates. An example is shown in Fig. 5.

In the model system used for DNA genotyping of HPA, the DNA array chip coupled with *in situ* MALDI analysis unambiguously determined the specific alleles of the individual DNAs tested. The size of the chip wells can be miniaturized further because each matrix spot ($d \approx 300\ \mu m$) was usually enough for obtaining a satisfactory MALDI mass spectrum. Thus, it should be possible to develop even more miniaturized reaction chambers to reduce reaction volume by up to 100-fold. This will reduce further the cost of genotyping. Such miniaturization also will enable higher-density DNA chips that will minimize sample stage movement in mass spectrometers and increase the speed of signal acquisition for high-throughput analysis. All of these future steps seem practical, and they ensure that the genotyping cost per SNP soon will be significantly lower.


1. Collins, F. S., Patrinos, A., Jordan, E., Chakravarti, A., Gesteland, R., Walters, L., and the members of the Department of Energy and National Institutes of Health planning groups (1998) *Science* **282,** 682–689.
2. Chee, M., Yang, R., Hubbell, E., Berno, A., Huang, X. C., Stern, D., Winkler, J., Lockhart, D. J., Morris, M. S. & Fodor, S. P. A. (1996) *Science* **274,** 610–614.
3. Kozal, M. J., Shah, N., Shen, N., Yang, R., Fucini, R., Merigan, T. C., Richman, D. D., Morris, D., Hubbell, E., Chee, M., *et al.* (1996) *Nat. Med.* **2,** 753–759.
4. Hacia, J. G., Brody, L. C., Chee, M., Fodor, S. P. A. & Collins, F. S. (1996) *Nat. Genet.* **14,** 441–447.
5. Wang, D. G., Fan, J., Siao, C., Berno, A., Young, P., Sapolsky, R., Ghandour, G., Perkins, N., Winchester, E., Spenser, J., *et al.* (1998) *Science* **280,** 1077–1082.
6. Tang, K., Fu, D., Kötter, S., Cotter, R. J., Cantor, C. R. & Köster, H. (1995) *Nucleic Acids Res.* **23,** 3126–3131.
7. O'Donnell, M. J., Tang, K., Köster, H., Smith, C. L. & Cantor, C. R. (1997) *Anal. Chem.* **69,** 2438–2443.
8. Köster, H., Tang, K., Fu, D., Braun, A., van den Boom, D., Smith, C. L., Cotter, R. J. & Cantor, C. R. (1996) *Nat. Biotechnol.* **14,** 1123–1128.
9. Fu, D., Tang, K., Braun, A., Reuter, D., Darnhofer-Demar, B., Little, D. P., O'Donnell, M. J., Cantor, C. R. & Köster, H. (1998) *Nat. Biotechnol.* **16,** 381–384.
10. Braun, A., Little, D. P. & Köster, H. (1997) *Clin. Chem.* **43,** 1151–1158.
11. Braun, A., Little, D. P., Reuter, D., Müller-Mysok, B. & Köster, H. (1997) *Genomics* **46,** 18–23.
12. Little, D. P., Cornish, T. J., O'Donnell, M. J., Braun, A., Cotter, R. J. & Köster, H. (1997) *Anal. Chem.* **69,** 4540–4546.
13. Little, D. P., Braun, A., O'Donnell, M. J. & Köster, H. (1997) *Nat. Med.* **3,** 1413–1416.
14. Kunicki, T. J. & Newman, P. J. (1992) *Blood* **80,** 1386–1404.
15. Skogen, B., Bellissimo, D. B., Hessner, M. J., Santoso, S., Aster, R. H., Newman, P. J. & McFarland, J. G. (1994) *Transfusion* **34,** 955–960.
16. Legler, T. J., Köhler, M., Mayr, W. R., Panzer, S., Ohto, H. & Fischer, G. F. (1996) *Transfusion* **36,** 426–431.
17. Tang, K., Allman, S. L., Chen, C. H., Chang, L. Y. & Schell, M. (1994) *Rapid Commun. Mass Spectrom.* **8,** 183–186.