CONF-940723--45

Laser mass spectrometry for DNA fingerprinting for forensic applications

C. H. Chen, K. Tang, N. I. Taranenko, S. L. Allman, and L. Y. Ch'ang

Oak Ridge National Laboratory, Health Sciences Research Division,
Oak Ridge, Tennessee 37831-6378

1. INTRODUCTION

The application of DNA fingerprinting has become very broad in forensic analysis, patient identification, diagnostic medicine, and wildlife poaching, since every individual's DNA structure is identical within all tissues of their body. DNA fingerprinting was initiated by the use of restriction fragment length polymorphisms (RFLP).[1] In 1987, Nakamura et al.[2] found that a variable number of tandem repeats (VNTR) often occurred in the alleles. The probability of different individuals having the same number of tandem repeats in several different alleles is very low. Thus, the identification of VNTR from genomic DNA became a very reliable method for identification of individuals. Take the Huntington gene as an example, there are CAG trinucleotide repeats. For normal people, the number of CAG repeats is usually between 10 and 40. Since people have chromosomes in pairs, the possibility of two individuals having the same VNTR in the Huntington gene is less than one percent, if we assume equal distribution for various repeats. When several allels containing VNTR are analyzed for the number of repeats, the possibility of two individuals being exactly identical becomes very unlikely. Thus, DNA fingerprinting is a reliable tool for forensic analysis. In DNA fingerprinting, knowledge of the sequence of tandem repeats and restriction endornuclease sites can provide the basis for identification.

The major steps[3] for conventional DNA fingerprinting include (1) specimen processing, which usually involves the chemical extraction of genomic DNA, (2) amplification of selected DNA segments by PCR, and (3) gel electrophoresis to do the final DNA analysis. In general, the process of gel electrophoresis involves the radioactive material labeling or dye tagging.[4] It is a very time-consuming process. In this work we propose to use laser desorption mass spectrometry for fast DNA fingerprinting.

The conventional method used for sizing DNA fragments is gel electrophoresis. In this process, an electric field is applied to the gel medium to cause the drift of DNA fragments that typically exist as charged particles. The drift velocity tends to be slower as DNA fragments become larger. However, this process is slow and it takes hours to separate different sizes of DNA fragments. The resolution is not very good due to the limitation caused by the physical size of the electrophoresis instrument, which limits the separation of different sizes of DNA. In addition, other physical and chemical properties of gel, such as viscosity, temperature, and temperature gradients, tend to influence the drift velocity of DNA fragments. Thus, a gel electrophoresis method is not suitable to accurately measure the molecular weight of biomolecules. It can only be used to detect biomolecules that have an electrical charge in the gel solution.

Mass spectrometry has been used for several decades to measure the molecular weight of gas

samples by mass-to-charge ratio. In order to use a mass spectrometer for molecular weight determination, ions must be produced in gas phase and the absolute pressure in the mass spectrometer needs to be less than 10^{-4} Torr to prevent collisions between ions and other gaseous molecules. In general, a mass spectrometer needs to have three key components. They are (1) an ionizer, (2) a mass selector, and (3) an ion detector. Since mass spectrometry is based on the detection of gas phase ions, a means of producing ions is obviously needed.

A mass spectrometer can be used for molecular weight determination only under the condition that the molecular ions are produced in space. However, DNA fragments in general do not have enough vapor pressure for gas phase analysis. In order to use mass spectroscopy for DNA fragments, it is necessary to develop a method which places DNA fragments in space without breaking them up.

For the analysis of materials with no vapor pressure at room temperature, several desorption methods have been used to carry the analytes into space for mass analysis. Major approaches include fast atom bombard (FAB) mass spectrometry, secondary ion mass spectrometry (SIMS), and laser desorption mass spectrometry (LDMS). However, none of these have been very successful for analysis of large biomolecules. Since most biomolecules are fragile, when subjected to heat or bombardment, a fragmentation process almost always occurs. However, in 1987 Hillenkamp and his coworkers[5] discovered that large protein molecular ions can be produced without much fragmentation by laser desorption if these biomolecules are mixed with smaller organic compounds which serve as a matrix for strong absorption of a laser beam. This process is now called matrix-assisted laser desorption (MALD). The typical preparation technique for MALD is to dissolve biomolecular samples in solution, then prepare another solution containing a small organic compound such as nicotinic acid. These two solutions are subsequently mixed and a small amount of solution is placed on a metal place to dry. After the crystallization of the sample, the sample plate is placed in the mass spectrometer for analysis. The molar ratio of matrix to analyte is typically more than 1000 to 1. For a MALD process, matrix materials strongly absorb the laser energy and become vaporized, and large biomolecules are carried out during the fast vaporization process. Large biomolecules can be delivered into space without breakup, which is probably due to minimal direct absorption of laser energy; thus, "soft" desorption can be achieved. However, it has also been found that parent biomolecule ions are also produced during the MALD process in addition to the expected neutral molecules. Thus, these desorbed ions can be directly detected by a mass spectrometer. The process involving ionization and matrix-assisted laser desorption at the same time is abbreviated MALDI (namely, matrix-assisted laser desorption and ionization).

The key advantage to using mass spectrometry for DNA detection is speed. Using gel electrophoresis for DNA separation typically takes a few hours, and overnight if radioactive tagging is used. However, a mass spectrum can be obtained within one second. For example, a time-of-flight mass spectrometer (TOFMS) usually needs less than one millisecond to separate different sizes of DNA; thus speed is improved by six to seven orders of magnitude. In addition, hundreds of samples can be placed in a mass spectrometer for fast analysis if the sample holder is designed to hold multiple targets. For gel electrophoresis, tagging of radioactive materials or organic chromophore is required for detection. Thus, radioactive or chemical wastes are produced. Mass spectrometry promises to eliminate the need for tagging; therefore, no waste will be produced.

Although matrix-assisted laser desorption/ionization has been very successful for proteins, the success for DNA has been limited to only small oligonucleotides due to the fragile chemical structures of DNA. Recently, we found a few new matrices that enable the use of MALDI for detecting large DNA segments. In addition, the more innovative design of a time-of-flight mass spectrometer also improved the detection sensitivity.

2. EXPERIMENTAL

A schematic diagram of the experimental apparatus is shown in Fig. 1.[6-7] A Nd:YAG laser capable of delivering four different wavelengths (i.e., 1064 nm, 532 nm, 355 nm, and 266 nm) was used for laser ablation. Typical laser fluence used in this work was less than 200 mJ/cm^2 to prevent any possible production of plasma or fragmentation. An ion deflector was used to deflect electrons and small ions whenever it was necessary since the large number of desorbed electrons and small ions was often enough to saturate the electron multiplier. The pulse duration of the ion deflector was adjustable, ranging between 4 and 40 μs, and the maximum switching voltage was 5 kV. A conversion box was used to receive ions and emit secondary electrons. A Johnston Laboratories, Inc., electron multiplier was used to detect secondary electrons from an aluminum foil in the conversion box. The conversion box was installed to reduce the probability of saturation of the electron multiplier by scattered laser light, and to prevent the contamination of the first dynode of the Johnston multiplier by the organic and bio-organic samples used. Signals from the multiplier went through a preamplifier and subsequently to a fast digital oscilloscope with a maximum sampling rate of 280 samples/s. TOF waveforms from each laser pulse were transferred to a computer for data storage and retrieval. The maximum resolution (m/Δm) of the TOF spectrometer is only about 400. However, the TOF resolution of heavy DNA segments is somewhat lower. This is probably due to the matrix materials, since the initial velocity distribution of desorbed oligomers is expected to be similar to the velocity distribution of the matrix material. Single-stranded DNA segments used in this work were obtained from Oligos Etc., Inc. Double-stranded DNA segments were obtained by using polymerase chain reaction (PCR) for DNA amplification from selected genomic DNA.

3. RESULTS AND DISCUSSION

A laser-desorption negative ion mass spectrum of mixtures of 3, 4, 5, 8, 11, 15, 20, and 34 nucleotides with 3-methlsalicylic acid and 3-hydroxy-4-methoxybenzaldehyde mixtures excited by a 355 nm laser beam is shown in Fig. 2. Parent ions of each oligomer were observed. However, the observation of parent ions is only limited to small oligonucleotides. Figure 3 shows the spectrum of somewhat larger single-stranded DNA with 3-hydroxypicolinic acid as the matrix. Recently, we found that picolinic acid is a better matrix for MALDI of DNA. A mass spectrum of mixtures of single-stranded DNA of 150 mer and 100 mer is shown in Fig. 4. A MALDI spectrum of double-stranded DNA with 246 base pairs using the mixture of picolinic acid and 3-hydroxypicolinic acid as matrices is shown in Fig. 5. It was found the DNA ions detected in the mass spectrometer always corresponds to the molecular weight of single-stranded DNA. Although the molecular weights of each single-stranded DNA are not the same, the limited resolution prevents the observation of two separated peaks. The DNA segments used represent a partial sequence of a human immunoglobulin V_λ gene. The quantity of DNA in a sample was estimated as 200 femtomole by measuring fluorescence intensity using a fluorometer. Figure 6 shows the mass spectrum of double-stranded

DNA of 500 base pairs in size. The quantity of DNA in the sample was estimated as 100 femtomole. Both single- and double-charged ions were detected. Figure 7 shows the mass spectrum of mixtures of poly-A. The resolution was good enough to resolve the DNA segments with one base difference. However, for DNA larger than 100 nucleotides, the resolution usually is still not good enough to distinguish the oligonucleotide with a single base difference.

In order to achieve fast sequencing for DNA fingerprinting by laser mass spectrometry, in our opinion, the following three conditions are required: (1) the size of the DNA fragment to be detected should be larger than 300 nucleotides, (2) the sensitivity needs to be high enough to detect quantities typical of PCR production, usually in the femtomole region, and (3) the resolution of mass spectra needs to be sufficiently high for sequence determination. We have proven that laser desorption mass spectrometry can be used to detect large DNA segments and for good sensitivity. If the resolution can be significantly improved, fast DNA sequencing for DNA fingerprinting by laser desorption mass spectrometry should become a practical reality.

Assuming better resolution can be obtained in the near future, we give this example of mass spectrometry for DNA sequencing: a small single-stranded oligonucleotide with the structure of GAA TTC CGC CGA TAC TGA CGG GCT CA GGA GTC GTC GCC ACC is to be sequenced. Sanger's method will produce four different sets of DNA ladder products with selected terminal nucleotides such as A (adenine), T (thymine), C (cytosine), and G (guanine). These four sets of DNA ladder products are to be mixed with a matrix such as the mixture of PA and 3-HPA, and subsequently placed into the mass spectrometer for detection. Figure 8 is the illustration of the concept for using MALDI-MS for DNA sequencing. The time required for running gel electrophoresis is usually in terms of hours. However, the time for analyzing a DNA sample by mass spectrometry is typically in the range of a few hundred microseconds.

In addition to using sequencing for DNA fingerprinting, the measurements of the size of DNA can also be used for DNA fingerprinting. Figure 9 shows the mass spectrum of a pair of double-stranded DNA, clearly indicating that one set of double-stranded DNA has three base pairs less than the other set of double-stranded DNA. These results indicate that VNTR can also be detected by laser desorption mass spectrometry.

In conclusion, laser desorption mass spectrometry is emerging as a powerful tool for fast DNA fingerprinting. We expect a real demonstration and routine use in the near future.

4. ACKNOWLEDGEMENT

Research was sponsored by the office of Health and Environmental Research, U. S. Department of Energy under contract DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc. Preparation of the manuscript by Darlene Holt is also acknowledged.

5. REFERENCES

1. A. J. Jeffreys, V. Wilson, and S. L. Thein, *Nature* **314**, 67 (1985).
2. Y. Nakamura, M. Leppert, P. O'Connell, R. Wolff, T. Holm, M. Culver, C. Martin,

E. Fujimoto, M. Hoff, E. Kumlin, and R. White, *Science* **235**, 1616 (1987).

3. L. T. Kirby, "DNA Fingerprinting," Stockton Press, 1990.

4. C. H. Chen, M. G. Payne, and K. B. Jacobson, *International Journal of Genome Research* **1**, 25 (1992).

5. M. Karas, D. Backmann, U. Bahr, and F. Hillenkamp, *International Journal of Mass Spectrometry and Ion Processes* **78**, 531 (1987).

6. K. Tang, S. L. Allman, and C. H. Chen, *Rapid Communications in Mass Spectrometry* **6**, 365 (1993).

7. K. Tang, S. L. Allman, R. B. Jones, C. H. Chen, and S. Araghi, *Rapid Communications in Mass Spectrometry* **7**, 63 (1993).

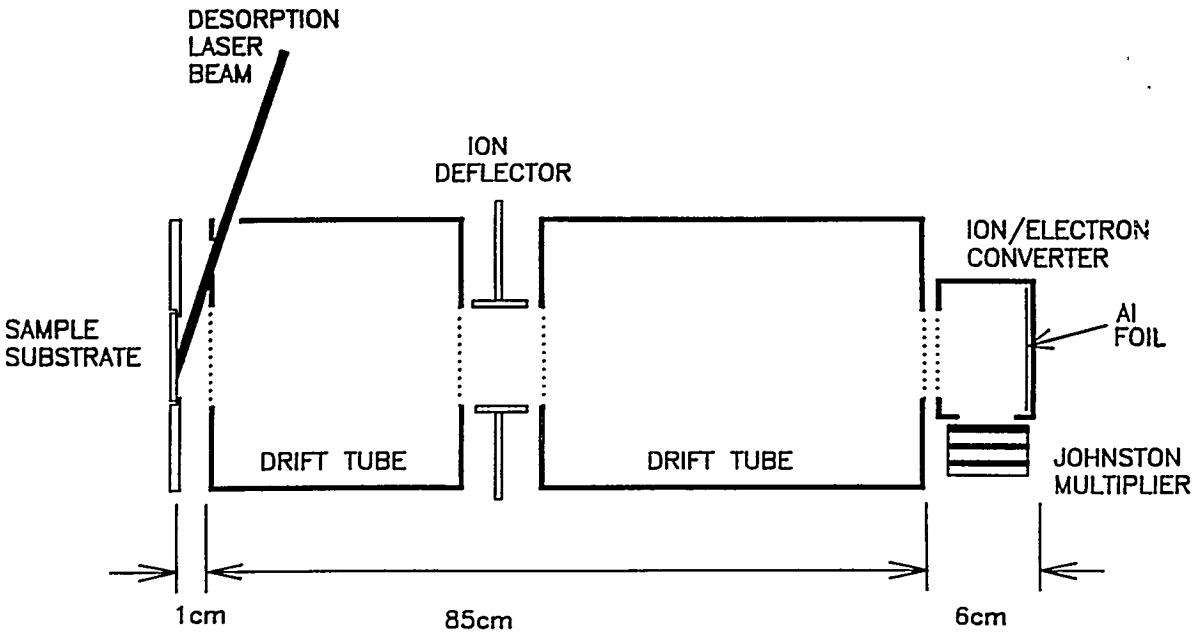

Figure 1. Schematic diagram of the time-of-flight mass spectrometer.

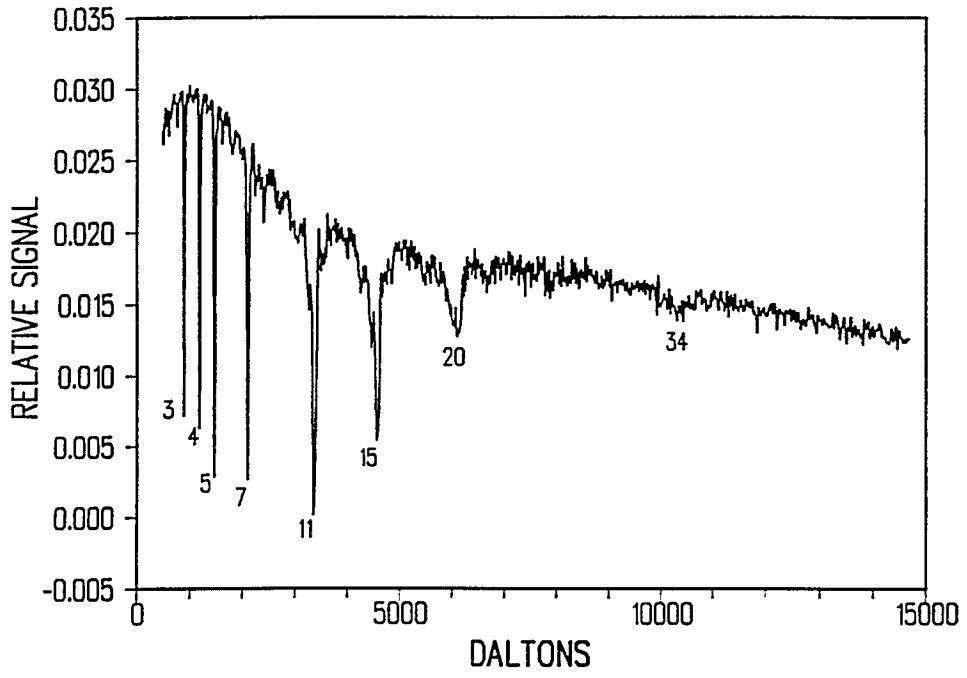

Figure 2. Negative ion mass spectrum obtained from mixtures of oligonucleotides which include 3-mer (5'-AGT-3'), 4-mer (5'-AGTC-3'), 5-mer (5'-AGTCC-3'), 7-mer (5'-AGTCCTG-3'), 11-mer (5'-AGTCCTGAAGT-3'), 15-mer (5'-AGTCCTGAAGTCCTG-3'), 20-mer (5'-AGTCCTGAAGTCCTGAAGTC-3'), and 34-mer (5'-AGTCCTGAAGTCCTGAAGTCAGTCCTGAAGTCCT-3'). Each analyte loaded is about 20 pmol and the quantity laser desorbed is estimated as ~350 fmol. A mixture of 3-methylsalicylic acid and 3-hydroxy-4-methoxybenzaldehyde with a ratio of 2 to 1 was used as matrix material. The mole ratio of matrix to analyte is 15,000:1. The laser wavelength is 355 nm with laser fluence at 160 mJ/cm^2.

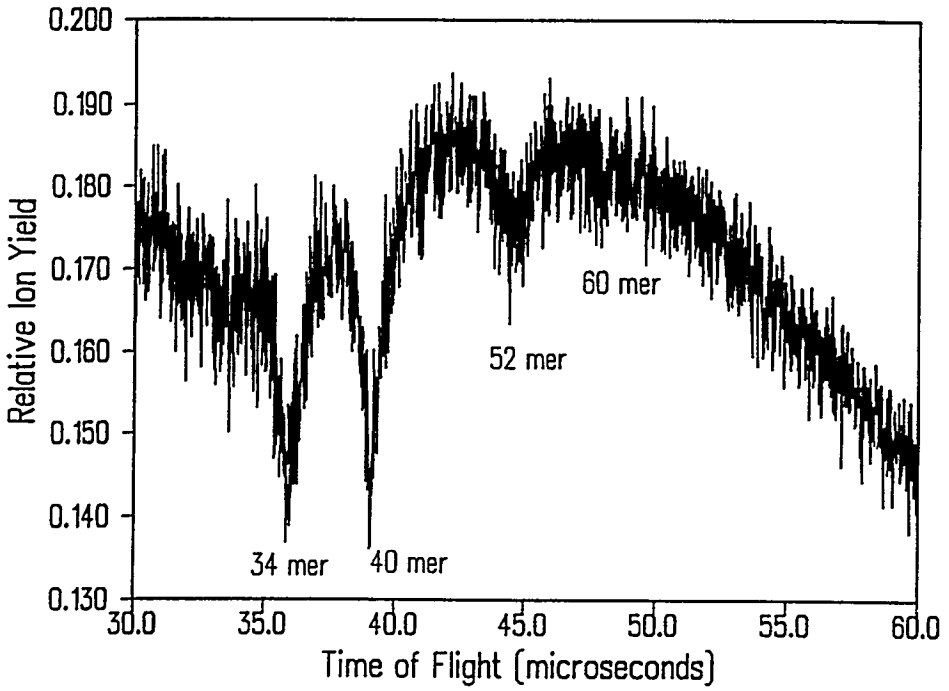

Figure 3. Positive ion spectrum of oligomer mixtures of 34-mer, 40-mer, 52-mer, and 60-mer with 3-hydroxypicolinic acid as the matrix. The sequences of these oligomers are the same as those described in Fig. 2 and Fig. 4 except for the 40-mer which has the sequence 5'-AGTCCTGAAGTCCTGAAGTCAGTCCTGAAGTCCTGAAGTC-3'). The sample had 30 pmol of each oligomer and 1.2 μmol of 3'-hydroxypicolinic acid. The laser fluence was 150 mJ/cm^2 with wavelength at 355 nm.

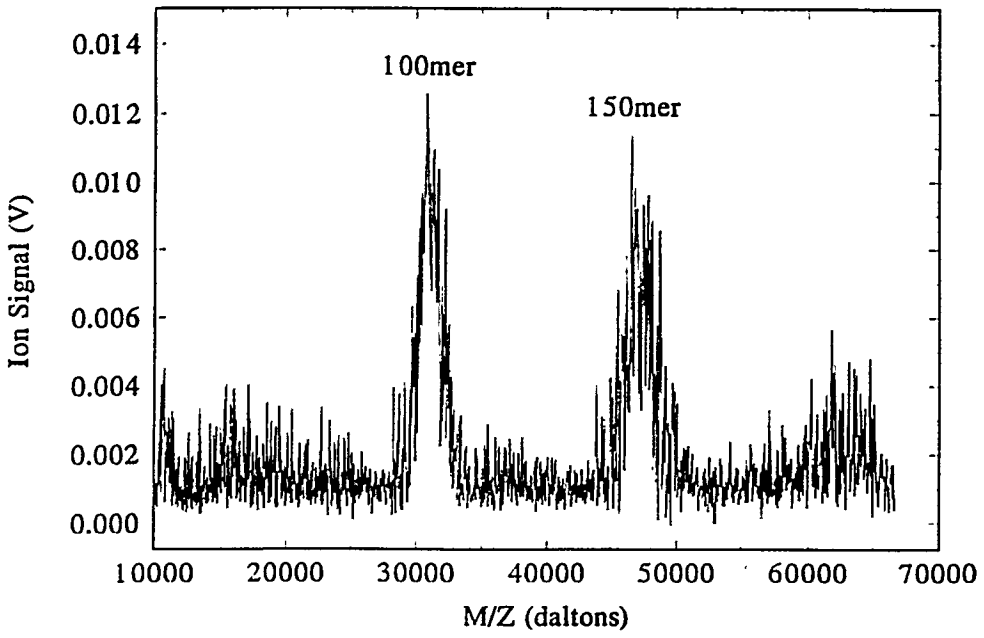

Figure 4. Negative ion mass spectrum of synthetic single-stranded DNA of 100 and 150 nucleotides.

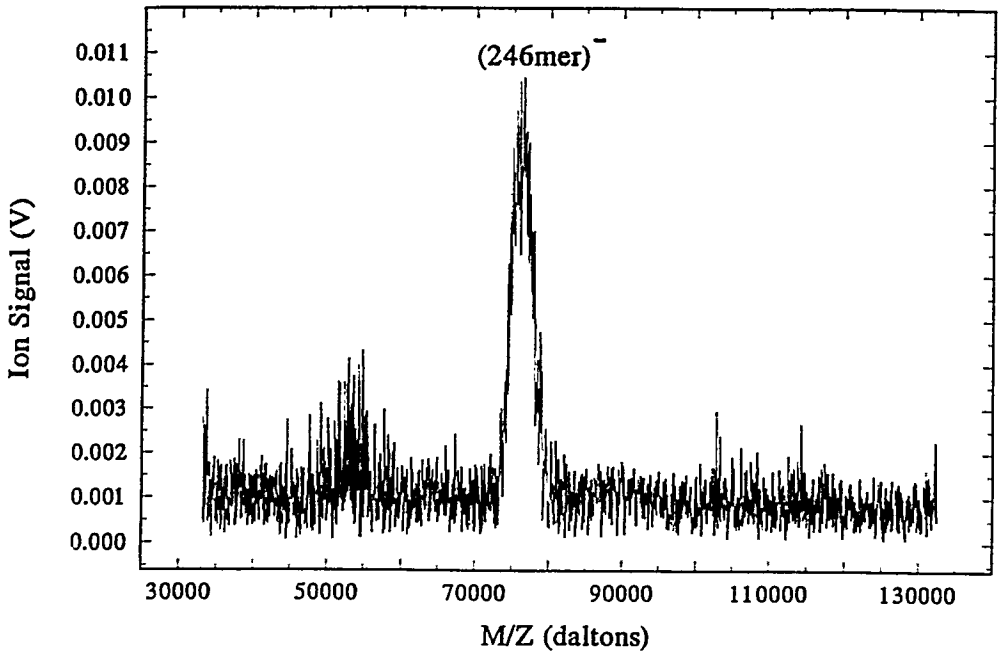

Figure 5. Negative ion mass spectrum of a 246-base-pair DNA amplified by a PCR process.

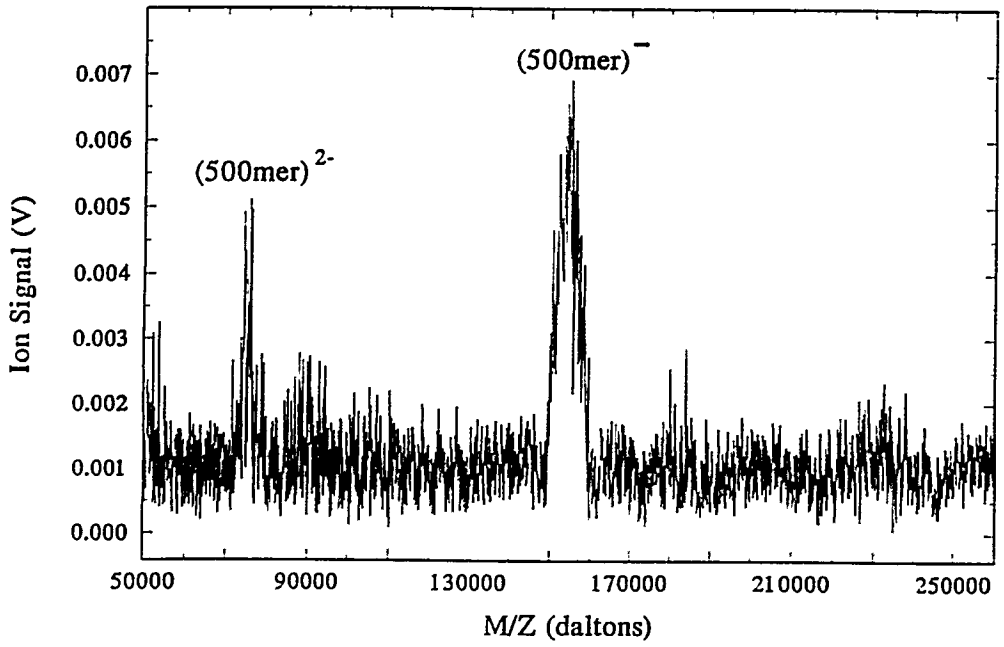

Figure 6. Mass spectrum of a 500-base-pair DNA amplified from bacteriophage lambda genome.

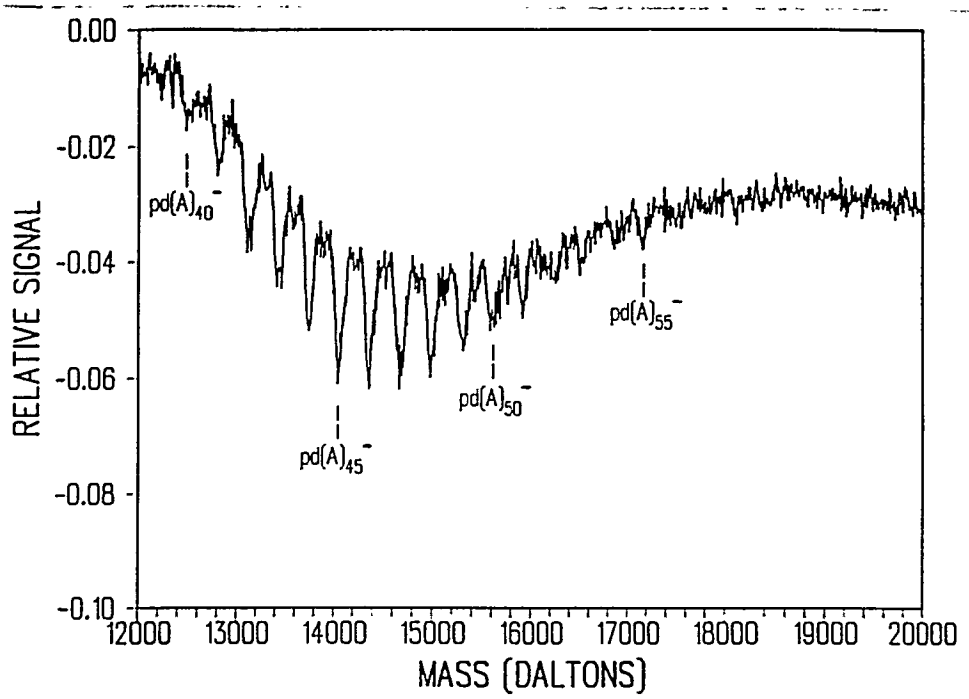

Figure 7. Negative ion spectrum of mixtures of $pd(A)_{40}$, $pd(A)_{41}$, $pd(A)_{42}$, $pd(A)_{43}$, $pd(A)_{44}$, $pd(A)_{45}$, $pd(A)_{46}$, $pd(A)_{47}$, $pd(A)_{48}$, $pd(A)_{49}$, $pd(A)_{50}$, $pd(A)_{51}$, $pd(A)_{52}$, $pd(A)_{53}$, $pd(A)_{54}$, $pd(A)_{55}$, $pd(A)_{56}$, $pd(A)_{57}$, $pd(A)_{58}$, $pd(A)_{59}$, and $pd(A)_{60}$. The sample had about 5 pmol of each oligonucleotide and 0.6 μmol of 3-HPA as the matrix. Approximately 20 polymer beads with ammonium acetate were also put into the sample. The laser fluence was 45 mJ/cm^2.

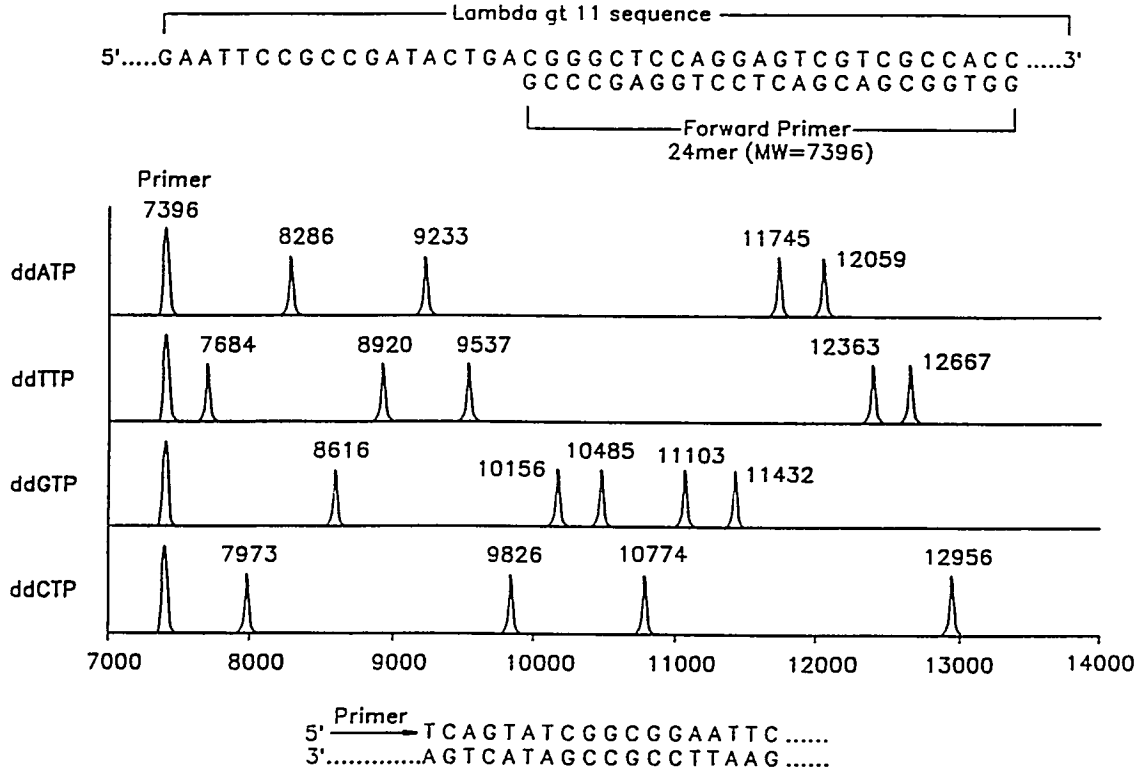

Figure 8. Conceptual mass spectra for DNA sequencing.

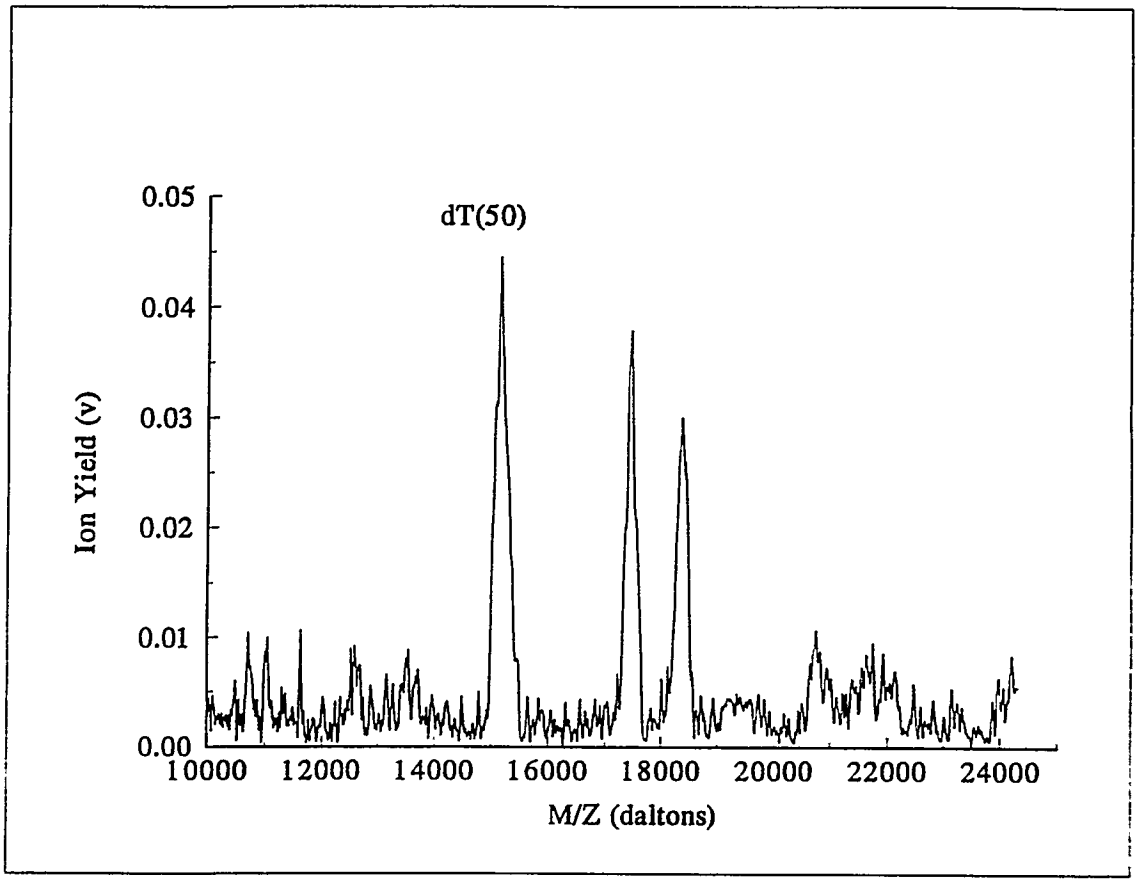

Figure 9. Mass spectrum for a pair of double-stranded DNA. Two peaks show one for 60 base pairs, the other for 57 base pairs. Although double-stranded DNA segments were in the sample, the observed ions correspond to the molecular weight of single-stranded DNA. $d(T)_{50}$ was used for calibration of the molecular weights of other DNA peaks. This result shows that fast DNA fingerprinting by mass spectrometry should be feasible.