Product Application Focus

genotools SNP MANAGER: A New Software for Automated High-Throughput MALDI-TOF Mass Spectrometry SNP Genotyping

W. Pusch, K.-O. Kraeuter, T. Froehlich[1], Y. Stalgies, and M. Kostrzewa[1]
Bruker Daltonik GmbH, Bremen, and [1]Bruker Saxonia Analytik GmbH, Leipzig, Germany

BioTechniques 30:210-215 (January 2001)

ABSTRACT

Analysis of single nucleotide polymorphisms (SNPs) is a rapidly growing field of research that provides insights into the most common type of differences between individual genomes. The resulting information has a strong impact in the fields of pharmacogenomics, drug development, forensic medicine, and diagnostics of specific disease markers. The technique of matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) has been shown to be a highly suitable tool for the analysis of DNA. It supplies a very versatile method for addressing a high-throughput SNP genotyping approach. Here, we present the Bruker genotools SNP MANAGER, a new software tool suitable for highly automated MALDI-TOF MS SNP genotyping. The genotools SNP MANAGER administers the sample preparation data, calculates masses of allele-specific primer extension products, performs genotyping analysis, and displays the results. In the current study, we have used the genotools SNP MANAGER to perform an automated duplex SNP analysis of two biallelic markers from the promoter of the gene encoding the inflammatory mediator interleukin-6.

INTRODUCTION

Single nucleotide polymorphisms (SNPs) are the most common type of genetic variations between different individuals or between different ethnic groups (22).

Current estimations expect statistically one SNP per 1000 bp genomic DNA (3). Thus, the haploid human genome, consisting of approximately 3 billion bp (13), is expected to contain markedly more than a million SNPs. Considerable efforts have been started in industrial and academic research to identify a large portion of these human SNPs (10,14,20,21). The SNP consortium, which was founded in 1999 by leading pharmaceutical companies and genomic research centers, has already mapped more than 100 000 SNPs (9,15).

SNPs are of particular interest in genetic mapping, association studies, forensic medicine, and as diagnostic disease markers. Moreover, SNPs can cause alterations of gene function by amino acid exchanges in the translated proteins, by the creation of splice site variants, or by change of promoter function. A central point of pharmacogenomic research is the impact of SNPs on amino acid exchanges in drug metabolizing enzymes (1). Therefore, a consideration of the respective genotyping results is necessary for the choice and/or the dosage of an applied drug.

As a result of the vast amount of SNPs and the even higher number of potential patient samples, there is a strong need for a reliable method suitable for an automated high-throughput analysis. Because of its speed, cost effectiveness, and multiplex possibilities, matrix-assisted laser desorption/ionization time of flight mass spectrometry (MALDI-TOF MS) has been shown to be extremely suitable for high-throughput SNP genotyping (6,7,16,18). Furthermore, because of its accuracy, it is a high-quality option for future DNA diagnostics (2,12). In this study, we present a new software tool capable of performing an automated SNP genotyping analysis from a MALDI-TOF MS spectrum in 2–3 s.

MATERIALS AND METHODS

PCR Amplification

Human genomic DNA from anonymous unrelated individuals was screened for interleukin-6 (IL6) promoter polymorphism genotypes. All PCR and primer extension reactions were carried out in a PTC-225™ Tetrad thermal cycler (MJ Research, Waltham, MA, USA). PCR primers for a portion of the IL6 promoter spanning both polymorphic sites were (5′-GAGACGCCTTGAAGTAACTGCAC-3′) and (5′-CCTGGAGGGGAGATAGAGCTTCT-3′).

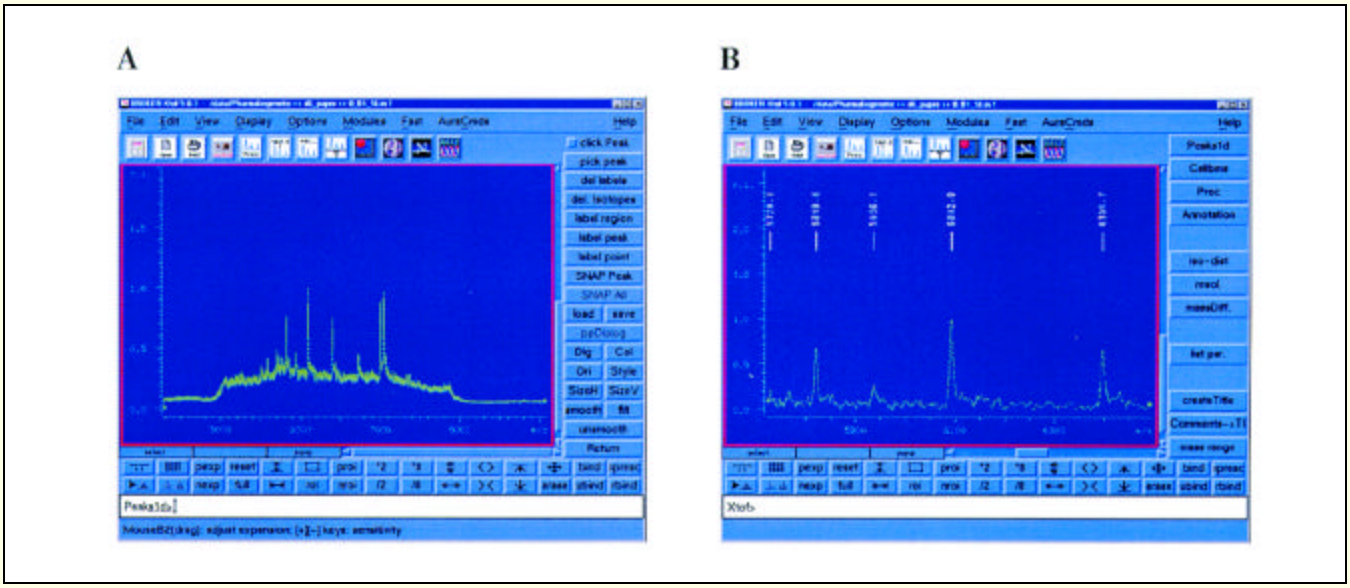


Figure 1. MALDI-TOF MS spectrum corresponding to the target position B1 of the IL6-DX55, IL6-DX1 duplex SNP analysis in Figure 2F. (A) Raw spectrum. Automatic acquisition was performed with the XACQ 4.0 acquisition software in combination with the AUTOXECUTE 5.0 package. The spectrum is displayed in the XMASS/XTOF 5.0 processing software. (B) The same spectrum as in panel A after automatic processing according to the IL6-DX1 method. Internal calibration was performed. Peaks corresponding to the primer and to both extension products have been labeled. Note the high mass accuracy compared with the expected masses (Table 1). Additional peaks (e.g., caused by extension products from other SNP methods in a multiplex analysis) are ignored for the genotyping result.

Table 1. Allele-Specific Primer Extension Reactions for the Performed IL6 Gene Promoter Duplex Analysis

Primer Extension	Parameter	IL6 Gene Promoter (IL6-DX55)	IL6 Gene Promoter (IL6-DX1)
Primer	Sequence Average Mass (Da)	5´-GCAATGTGACGTCCTTTAGCAT-3´ 6725.4	5´-GTTGTAGACTGCCTGGCCA-3´ 5819.8
Extension Product (Allele A)	Sequence Average Mass (Da)	Primer + ddC 6998.5	Primer + ddC 6092.9
Extension Product (Allele B)	Sequence Average Mass (Da)	Primer + ddG 7038.5	Primer + dTddC 6397.1

For the PCRs, 1 U *Taq* DNA polymerase (Qiagen, Hilden, Germany), 200 µM dNTPs (Promega, Mannheim, Germany), and 5 pmol each primer were used together with 30 ng genomic DNA as template. PCR conditions were 2 min at 94°C, followed by 35 cycles of 10 s at 94°C, 15 s at 63°C, 30 s at 72°C, and a final extension time of 2 min at 72°C. To remove excess primers, dNTPs, and salts, the genopure dsDNA purification kit (Bruker Saxonia, Leipzig, Germany) for dsDNA was used. Purified PCR products were recovered in 5 µL elution buffer and directly used for subsequent primer extension.

SNP Detection Reaction

The purified PCR product (5 µL) was mixed with the extension mixture containing 1 U Thermo Sequenase™ (Amersham Pharmacia Biotech, Freiburg, Germany) and polymorphism detection primers each of 15 pmol in 1× reaction buffer (Amersham Pharmacia Biotech). Chain terminator 2′,3′-dideoxynucleotide triphosphates ddCTP and ddGTP (each 200 µM; Roche Molecular Biochemicals, Mannheim, Germany) were used together with regular dTTP nucleotide (200 µM; Promega). The polymorphism detection primers were IL6-DX1: (5′-GTTGTAGACTGCCTGGCCA-3′) for G/A transition at relative position -594 and IL6-DX55: (5′-GCAATGTGACTGCCTTTAGCAT-3′) for G/C transversion at relative position -172. Cycling conditions were 2 min at 94°C, followed by 35 cycles of 10 s at 94°C, 15 s at 55°C, 15 s at 72°C, and a final extension time of 2 min at 72°C. Excess salts and nucleotides were removed using the genopure oligo DNA purification system (Bruker Saxonia). Purified extension products and residual primers were eluted in 10 µL elution buffer and subsequently subjected to MALDI-TOF MS.

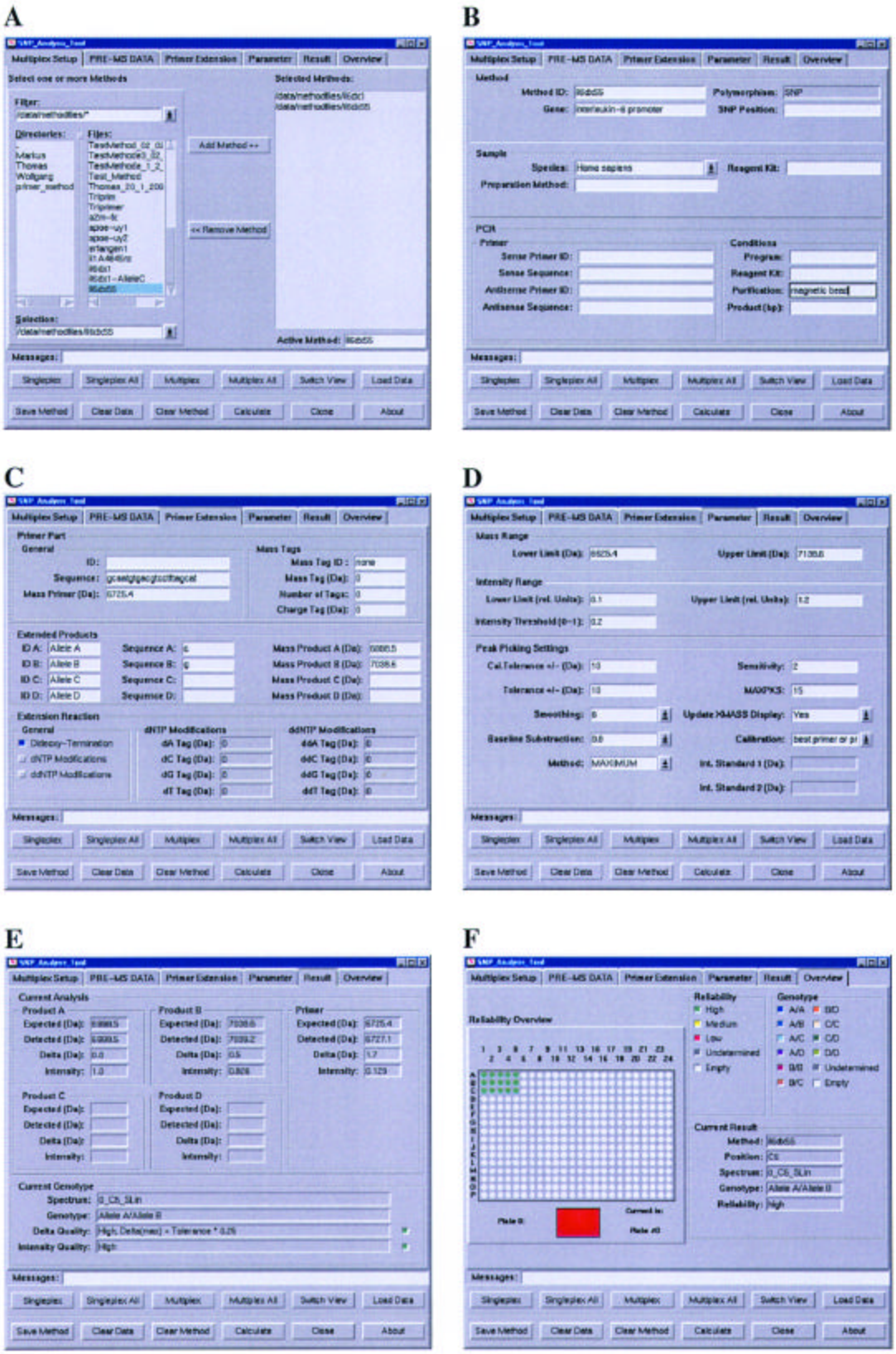
Mass Spectrometry

One microliter of matrix solution (3-hydroxypicolinic acid, 6 mg/mL and dibasic ammonium citrate, 1.25 mg/mL in water) was pipetted onto an Anchor Chip™ (Bruker Daltonik, Bremen, Germany) sample target (anchor size: 400 µm) and allowed to dry at room temperature. Subsequently, 1 µL analyte DNA was added to the matrix spots and dried at room temperature. The target was introduced into the source region of a BIFLEX III mass spectrometer (Bruker Daltonik) equipped with a SCOUT MTP operated in linear mode with 19 and 16.3 kV on the conversion dynode and the sample target, respectively. Theoretical average molecular masses were calculated according to their atomic compositions. Experimental values, all detected with external calibration, are reported as neutral deprotonated forms. The instrument was calibrated using ssDNA standards between 4 and 8 kDa.

RESTRICTION FRAGMENT-LENGTH POLYMORPHISM ANALYSIS

PCR conditions for restriction fragment-length polymorphism (RFLP) analysis were as described above. PCR primers were (5′-ATGCCAAAGTGCTGAGTCACTA-3′) and (5′-GCAGAATGAGCCTCAGACATCTC-3′) for G/C transversion at position -172 and (5′-CTTGAAGTAACTG CACGAAATTTGA-3′) and (5′-TTCCTCTGACTCCATCG CAGC-3′) for G/A transition at position -594. Products were purified using the genopure dsDNA purification kit. The PCR product spanning the -594 G/A transition was digested with *Fok*I (1 U for 1 h at (37°C; New England Biolabs, Frankfurt,

Figure 2. genotools SNP MANAGER software. (A) Multiplex setup. All previously designed SNP methods are available for the setup of the next genotyping analysis. Each method corresponds to one biallelic marker. Multiplex analyses are set up by combining different methods. (B) Pre-MS data. Optionally, data concerning the DNA preparation and the PCR conditions can be assigned to each method. (C) Primer extension. The masses of the primer and the extension products are derived from the respective sequences. A primer charge tag can be specified for the GOOD Assay, a purification-free SNP genotyping method by MALDI-TOF MS that has been recently described (19). Moreover, a further primer mass tag is possible. Mass modifications at the dNTPs and ddNTPs used in a primer extension experiment (5) can be specified for each type of nucleotide. (D) Parameter. Allows access to critical XMASS/XTOF settings for the peak picking, calibration, baseline correction, and smoothing. (E) Results. Shows the detailed results for the current spectrum with the active SNP method. The expected masses as well as the detected masses and the corresponding peak intensities are displayed. In the current analysis, the peaks corresponding to the extension primer and to both alleles have been found. The resulting genotype is heterozygous AB. The analysis quality is accepted, as all product peaks have an intensity higher than the user-defined threshold and a low mass deviation. Accordingly, the analysis is labeled with a green color code. (F) Overview. Gives a graphical overview of the results from a complete 384-well target. All 15 samples, which have been analyzed with the IL6-DX1 SNP method, have been accepted.


A
B
C
D
E
F

Germany), and the PCR product spanning the -172 G/C transition was digested with *Nla*III (1 U for 1 h at 37°C; New England Biolabs). Restriction fragments were analyzed on a 2% agarose gel.

SNP Genotyping

SNP genotyping was performed automatically using the genotools SNP MANAGER software. For online analysis, SNP genotyping was automatically performed during the acquisition of subsequent data from a 384-well target. Offline analysis was automatically performed using a set of previously acquired MALDI-TOF MS spectra.

RESULTS

We have analyzed DNA samples from 15 patients regarding two different biallelic markers located in the IL6 gene promoter (4,17,23). All genotypes derived from the automated MALDI-TOF MS analysis were controlled by an independent RFLP analysis. IL6-DX55 is a G/C transversion at position -172 relative to the transcription start site (4,17) generating a novel *Nla*III restriction site. Similarly, a novel *Fok*I restriction site results from a G/A transition at position -594 (IL6-DX1). MALDI-TOF MS analysis of these two markers (IL6-DX1 and IL6-DX55) yielded identical genotypes for each analyzed DNA sample (Table 2). This may indicate a linkage of the two polymorphic sites. All determined genotypes were in 100% agreement with an alternative established RFLP control analysis (data not shown). As a prerequisite for the analysis, short allele-specific DNA fragments, suitable for the MALDI-TOF process, were generated. This was done as described in Materials and Methods.

In a first step, the DNA region of interest was amplified by a conventional PCR. Subsequently, the resulting PCR fragment was used as the template in a primer extension reaction, using a primer located upstream from the polymorphic site.

In the primer extension reactions, allele-specific ssDNA fragments were generated for the respective SNPs (Table 1). The termination conditions were arranged to generate DNA fragments of unique size for all possible reaction products, thus allowing simultaneous MALDI-TOF analysis of both SNPs in a duplex assay. The spectra were automatically acquired using the XACQ 4.0 acquisition software in combination with the AUTOXECUTE 5.0 automation package (Bruker Daltonik).

Subsequently, these spectra were used for an automated SNP genotyping analysis using the XMASS/XTOF 5.0 processing software in combination with the new genotools SNP MANAGER module.

A typical raw spectrum containing peaks from both primer extensions (IL6-DX1, IL6-DX55) is shown in Figure 1A. Figure 1B shows the same spectrum after analysis with the genotools SNP MANAGER. The spectrum was internally calibrated. The expected primer peak (5819.7 Da) and the allele-specific product peaks (6092.9 Da, 6396.7 Da) are labeled.

Figure 2 shows the graphical user interface of the genotools SNP MANAGER. Detailed results of the analysis are shown in the Result section of the genotools SNP MANAGER (Figure 2E). A result overview of all spectra acquired from a complete 384-well plate target is displayed in the Overview section (Figure 2F). As the default, the reliability of the results is shown in a green/yellow/red color code for each target position. The reliability is ranked by the intensity of the detected product peaks and by the difference between the detected and the expected masses. Optionally, the overview can be switched to a display of the resulting genotypes.

Analysis results of insufficient quality (red or yellow quality code) can easily be screened by a mouse click on the desired target postion. The corresponding spectrum is loaded in XMASS, and the analysis data for the active method are loaded in the genotools SNP MANAGER. Insufficient quality may result from poorly converted extension primer or from

Table 2. SNP Genotyping Results from Automated MALDI-TOF MS Analysis

Sample	SNP	Genotype genotools SNP MANAGER
A1	IL6-DX1	A/A
	IL6-DX55	A/A
A2	IL6-DX1	B/B
	IL6-DX55	B/B
A3	IL6-DX1	A/A
	IL6-DX55	A/A
A4	IL6-DX1	A/B
	IL6-DX55	A/B
A5	IL6-DX1	A/A
	IL6-DX55	A/A
B1	IL6-DX1	A/B
	IL6-DX55	A/B
B2	IL6-DX1	A/B
	IL6-DX55	A/B
B3	IL6-DX1	A/B
	IL6-DX55	A/B
B4	IL6-DX1	A/A
	IL6-DX55	A/A
B5	IL6-DX1	A/A
	IL6-DX55	A/A
C1	IL6-DX1	A/B
	IL6-DX55	A/B
C2	IL6-DX1	B/B
	IL6-DX55	B/B
C3	IL6-DX1	A/B
	IL6-DX55	A/B
C4	IL6-DX1	A/B
	IL6-DX55	A/B
C5	IL6-DX1	A/B
	IL6-DX55	A/B

additional peaks caused by Na^+ or K^+ ion adducts.

The detailed results are stored together with the spectrum. Additionally, a summary of the results is stored in a result table. The result table is in the ASCII format and can easily be further processed using standard spreadsheet programs or laboratory information management systems (LIMS) systems.

DISCUSSION

For a high-throughput SNP analysis, MALDI-TOF MS is a method that has shown great promise (6,8,11). It is accurate, fast, cost effective, and suitable for automation. Moreover, MALDI-TOF MS SNP genotyping is based on the detection of the mass of the allele-specific products. Accordingly, it is not susceptible to secondary structure effects, unlike detection based on hybridization of fluorescent probes. We have developed a software package that automates not only the acquisition process but also the genotyping analysis.

The genotools SNP MANAGER software *(i)* administers the data concerning the sample preparation and PCR conditions, *(ii)* facilitates the setup of the primer extension reactions by calculating the molecular masses of the extension primer and the expected products from the input sequences (Figure 2C), *(iii)* defines critical data processing settings for the peak picking and the calibration, *(iv)* performs the genotyping analysis, *(v)* displays the detailed results with an evaluation of its reliability for the current spectrum, *(vi)* displays a graphical overview concerning the results of a complete 384-well target plate for the specified method, and *(vii)* stores the detailed results for each analysis and an overview table for all spectra of a 384-well target.

The genotools SNP MANAGER software is suitable for the use with current BIFLEX/REFLEX III MALDI-TOF mass spectrometers (Bruker Daltonik). It requires BIFLEX/REFLEX standard software for the acquisition (XACQ 4.0, AUTOXECUTE 5.0) and the data processing (XMASS/XTOF 5.0). The genotools SNP MANAGER allows automatic SNP genotyping of samples from complete 384-well targets. Analysis can be performed offline with previously generated spectra or online during data acquisition. Simultaneous SNP genotyping for multiple biallelic markers (multiplex analysis) is supported. Depending on the utilized hardware, a single offline SNP genotyping event takes 2–3 s. Accordingly, the spectra from a complete 384-well target can be analyzed in 13–20 min, thus allowing several thousand offline analyses per day. Moreover, the data acquisition of new spectra is also automated and can be controlled with the AUTOXECUTE software.

Depending on the settings, a new spectrum can be recorded every 7–10 s. During data acquisition, the system is capable of performing at least a duplex or triplex online analysis before the next set of data is acquired. For higher multiplex analysis, the data acquisition may overtake the postprocessing. In this case, the acquisition will be uncoupled from the postprocessing steps (batch mode). Conclusively, a BIFLEX/REFLEX MALDI-TOF mass spectrometer equipped with the AUTOXECUTE automation software, in combination with the genotools SNP MANAGER analysis software, supplies a reliable, highly automated system for high-throughput SNP analysis.

REFERENCES

1. **Aynocioglu, A.S., J. Brockmoller, S. Bauer, C. Sachse, P. Guzelbey, Z. Ongen, M. Nacak, and I. Roots.** 1999. Frequency of cytochrome P450 CYP2C9 variants in a Turkish population and functional relevance for phenytoin. Br. J. Clin. Pharmacol. *48*:409-415.
2. **v.d. Boom, D., C. Jurinke, S. Higgins, T. Becker, and H. Köster.** 1998. Mass spectrometric DNA diagnostics. Nucleosides Nucleotides *17*:2157-2164.
3. **Brookes, A.J.** 1999. The essence of SNPs. Gene *234*:177-186.
4. **Faulds, G., D. Fishman, P. Woo, and S. Humphries.** 1998. Novel *Fok*I polymorphism in the 5′ flanking region of the human IL-6 gene. GenBank Accession No. AF048692.
5. **Fei, Z., T. Ono, and L.M. Smith.** 1998. MALDI-TOF mass spectrometric typing of single nucleotide polymorphisms with mass-tagged ddNTPs. Nucleic Acids Res. *26*:2827-2828.
6. **Griffin, T.J., J. Hall, J.R. Prudent, and L.M. Smith.** 1999. Direct genetic analysis by matrix-assisted laser desorption/ionization mass spectrometry. Proc. Natl. Acad. Sci. USA *96*:6301-6306.
7. **Griffin, T.J. and L.M. Smith.** 2000. Single-nucleotide polymorphism analysis by MALDI-TOF mass spectrometry. Trends Biotechnol. *18*:77-84.
8. **Gut, I.G. and S. Beck.** 1995. A procedure for selective DNA alkylation and detection by mass spectrometry. Nucleic Acids Res. *23*:1367-1373.
9. **Hodgson, J.** 1999. Analysts, firms pour cold water on SNP consortium. Nat. Biotechnol. *17*:526.
10. **Kwok, P.Y. and Z. Gu.** 1999. Single nucleotide polymorphism libraries: why and how are we building them? Mol. Med. Today *5*:538-543.
11. **Li, J., J.M. Butler, Y. Tan, H. Lin, S. Royer, L. Ohler, T.A. Shaler, J.M. Hunter, D.J. Pollart, J.A. Monforte, and C.H. Becker.** 1999. Single nucleotide polymorphism determination using primer extension and time-of-flight mass spectrometry. Electrophoresis *20*:1258-1265.
12. **Little, D.P., A. Braun, B. Darnhofer-Demar, A. Frilling, Y. Li, R.T. McIver, Jr., and H. Köster.** 1997. Detection of RET proto-oncogene codon 634 mutations using mass spectrometry. J. Mol. Med. *75*:745-750.
13. **Lodish, H., D. Baltimore, A. Berk, S.L. Zipursky, P. Matsudeira, and J. Darnell.** 1996. Rekombinante DNA technologie, p. 234. *In* Molekulare Zellbiologie, 2nd ed. Walter de Gruyter, Berlin.
14. **Marth, G.T., I. Korf, M.D. Yandell, R.T. Yeh, Z. Gu, H. Zakeri, N.O. Stitziel, L. Hillier, P.Y. Kwok, and W.R. Gish.** 1999. A general approach to single-nucleotide polymorphism discovery. Nat. Genet. *23*:452-456.
15. **Masoon, E.** 1999. As consortium plans free SNP map of human genome. Nature *398*:545-546.
16. **Nordhoff, E., F. Kirpekar, and P. Roepsdorff.** 1996. Mass spectrometry of nucleic acids. Mass Spectrometry Rev. *15*:67-138.
17. **Olomolaiye, O.O., N.A.P. Wood, and J.L. Bidwell.** 1997. A novel *Nla*III polymorphism in the human IL-6 promoter. GenBank Accession No. AF005485.
18. **Ross, P., L. Hall, I. Smirnov, and L. Haff.** 1998. High level multiplex genotyping by MALDI-TOF mass spectrometry. Nat. Biotechnol. *16*:1347-1351.
19. **Sauer, S., D. Lechner, K. Berlin, H. Lehrach, J.L. Escary, N. Fox, and I.G. Gut.** 2000. A novel procedure for efficient genotyping of single nucleotide polymorphisms. Nucleic Acids Res. *28*:e13.
20. **Sherry, S.T., M. Ward, and K. Sirotkin.** 2000. Use of molecular variation in the NCBI dbSNP database. Hum. Mutat. *15*:68-75.
21. **Smigielski, E.M., K. Sirotkin, M. Ward, and S.T. Sherry.** 2000. dbSNP: a database of single nucleotide polymorphisms. Nucleic Acids Res. *28*:352-355.
22. **Wang, D.G., J.B. Fan, C.J. Siao, A. Berno, P. Young, R. Sapolsky, G. Ghandour, N. Perkins et al.** 1998. Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science *280*:1077-1082.
23. **Yasukawa, K., T. Hirano, Y. Watanabe, K. Muratani, T. Matsuda, S. Nakai, and T. Kishimoto.** 1987. Structure and expression of the human B cell stimulatory factor-2 (BSF-2/IL-6) gene. EMBO J. *6*:2939-2945.

Address correspondence to Dr. Wolfgang Pusch, Bruker Daltonik GmbH, Fahrenheitstrasse 4, 28359 Bremen, Germany. e-mail: wpu@bdal.de