

Genotypisierung

DNA-Analyse mittels Massenspektrometrie – Was wiegt die Individualität?

NICOLAS VOGEL, ANGELA SEEBAHN, CORD-MICHAEL BECKER
INSTITUT FÜR BIOCHEMIE UND MOLEKULARE MEDIZIN, EMIL-FISCHER-ZENTRUM,
FRIEDRICH-ALEXANDER-UNIVERSITÄT ERLANGEN-NÜRNBERG

Subtile Variationen der DNA begründen nicht nur unsere Individualität, sondern können auch zu Krankheiten führen. Die Methode der Massenspektrometrie erlaubt eine individuelle Risikoeinschätzung durch eine schnelle und sensitive Analyse dieser Variationen.

Subtle variations within the DNA not only render us individual, but may also be associated with diseases. Methods of mass spectrometry permit an individual risk assessment by rapid and sensitive analysis of these variations.

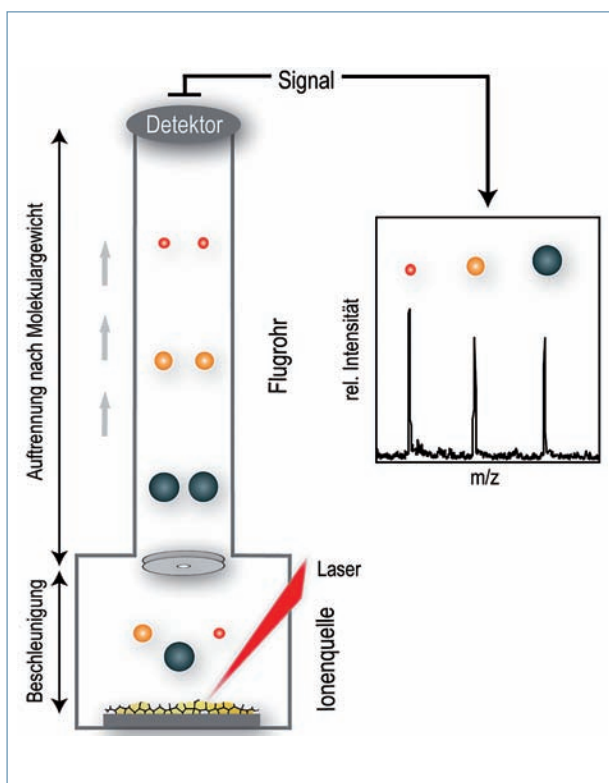
■ Seit dem Jahr 2003 gilt die Sequenz des humanen Genoms als vollständig entschlüsselt. Seither sind neue Fragestellungen in den Mittelpunkt der Forschung getreten. Darunter die Untersuchung von Sequenzvariationen der DNA, die Einfluss auf die Trans-

kriptionsrate, das Spleißen oder die Regulation verschiedener Gene haben und zu pathologischen Veränderungen führen können. Ferner bestimmen diese Variationen des menschlichen Erbguts im Zusammenspiel mit Umweltfaktoren individuelle Unter-

schiede, in denen die Einzigartigkeit jedes Menschen begründet liegt. Die Matrix-assistierte Laserdesorption/Ionisation-Flugzeit-Massenspektrometrie (MALDI-TOF-MS) hat sich als ein geeignetes Werkzeug zur effektiven und schnellen Untersuchung solcher Variationen herausgestellt (**Abb. 1**). So können heute Einzelnukleotidpolymorphismen (SNPs), DNA-Methylierung oder Mikrosatelliteninstabilität (MSI) mit ihrer Hilfe nachgewiesen werden. Nach Einführung dieser ursprünglich für die Untersuchung von Peptiden und Proteinen entwickelten Methode [1] wurde sie für die Nukleinsäureanalyse modifiziert und zur Anwendung gebracht. MALDI-TOF-MS genießt heute eine große Akzeptanz und breite Anwendung in der Analyse dieser Molekülklasse. Die Hauptanwendungen basieren auf einer spezifischen Verlängerung eines kurzen Oligonukleotids und werden zur Genotypisierung einzelner Allele und zur Bestimmung der Mikrosatellitenlänge oder des Cytosin-Methylierungsstatus verwendet (**Abb. 2**). Insgesamt betrachtet verfügen Labs mit einer MALDI-TOF-MS-Ausstattung über eine Plattform zur breiten Analyse von Makromolekülen verschiedenster Art.

Einzelnukleotidpolymorphismen (SNPs) – *A single base can change your life*

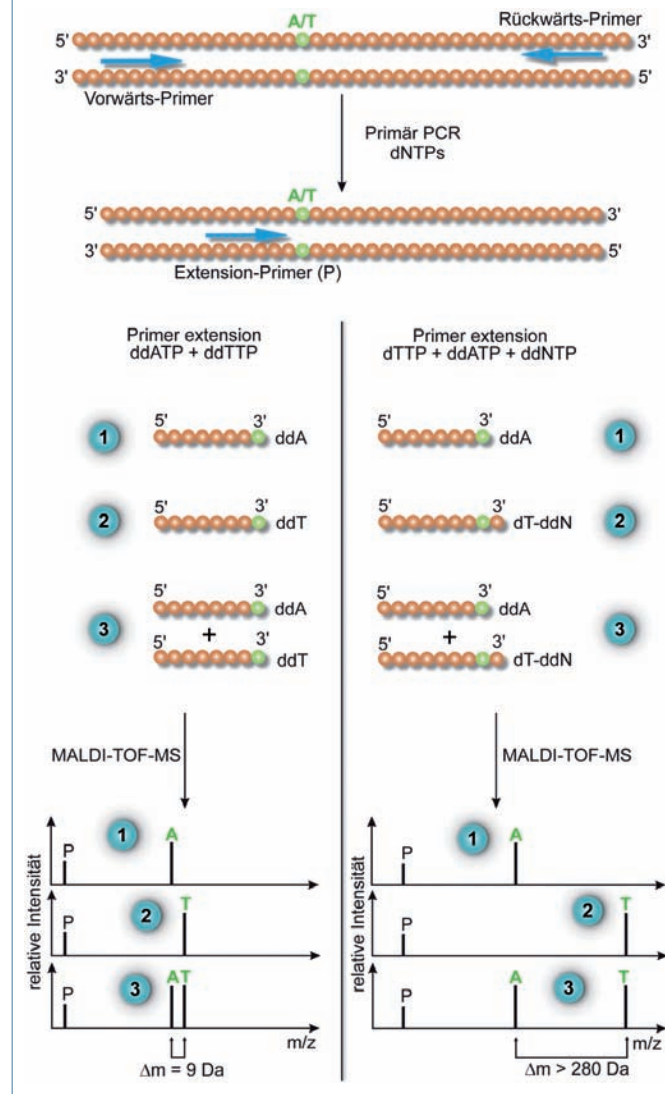
Unter SNP versteht man die natürliche genetische Variabilität einzelner Nucleobasen innerhalb eines Genoms. Durchschnittlich findet man solche variablen Positionen im menschlichen Genom in einem Abstand von 500 bis 1.000 Basenpaaren. Etwa 90 Prozent aller genetischen Varianten des Menschen werden durch sie bedingt. Hierbei unterscheidet man codierende und nicht-codierende SNPs, wobei codierende SNPs innerhalb von transkribierten Genabschnitten liegen und je nach Position die Aminosäuresequenz des resultierenden Proteins verändern können. SNPs in nicht-codierenden Bereichen können sich auf die Transkription eines Gens durch Zerstörung oder Erzeugung von regu-



◀ **Abb. 1:** Schematische Darstellung eines MALDI-TOF-Massenspektrometers. Die Analyten werden, eingebettet in eine Matrix auf dem Probenträger, durch Laserbeschuss verdampft und dabei ionisiert (geladen). Die gebildeten Analyt-Ionen werden in einem elektrischen Feld in eine feldfreie Driftstrecke, das Flugrohr, beschleunigt. Hier findet die Trennung der Molekül-Ionen nach ihrem Molekulargewicht statt. Ein Detektor am Ende des Flugrohrs registriert auftreffende Ionen und gibt ein für den Computer lesbares Signal aus. Durch eine geeignete Software können die Signale visualisiert und diskreten Molekulargewichten (genauer: Masse-zu-Ladungs-Verhältnissen, m/z) zugeordnet werden.

latorischen Elementen wie Bindungsstellen für Transkriptionsfaktoren auswirken. Damit können SNPs allgemein die individuelle Anfälligkeit für eine Krankheit oder deren Ausprägung maßgeblich beeinflussen. So ist z. B. ein bestimmtes Allel im Gen des Blutgerinnungsfaktors V (Faktor-V-Leiden, G1691A) für ein erhöhtes Thromboserisiko verantwortlich [2]. Die Kenntnis über den jeweiligen Genotyp erlaubt eine frühzeitige, vorbeugende Behandlung des betroffenen Patienten oder die Prognose eines Krankheitsverlaufs. Daher kann ein Screening von bekannten SNPs im Zuge von Korrelationsstudien wertvolle Informationen über die Kopplung von allelischen Varianten mit der untersuchten Krankheit liefern. Ferner dient die SNP-Analyse der Suche nach potenziellen Markern für klinische Diagnosen.

Die SNP-Genotypisierung ist eine der am weitesten verbreiteten Anwendungen für MALDI-TOF-MS innerhalb der Genomforschung. Basierend auf der Methode der Primer-Verlängerung (*primer extension*), der allelspezifischen Verlängerung eines Oligonukleotids, kann die vorliegende Variante eines bekannten SNPs bestimmt werden. Zu diesem Zweck wird die genomische DNA im Bereich der zu untersuchenden Base (typischerweise 300 bis 400 Basenpaare) mittels PCR vervielfältigt und die resultierende Doppelstrang-DNA gereinigt. Der zweite Schritt der Analyse beinhaltet die allelspezifische *primer extension*-Reaktion. Die Sequenz des zu verlängernden Oligonukleotids wird so gewählt, dass sein 3'-Ende exakt eine Base



◀ **Abb. 2:** Schematischer Ablauf der MALDI-TOF-MS-basierten DNA-Analyse zur Detektion von genetischen Heterogenitäten (SNP, MSI, Methylierung). Ausgehend von isolierter genomischer DNA wird der zu untersuchende DNA-Bereich mittels PCR vervielfältigt. Anschließend wird ein Oligonukleotid allelspezifisch verlängert (*primer extension*), idealerweise um eine unterschiedliche Anzahl an Nucleotiden. Die Extensionsprodukte werden dann mittels MALDI-TOF-MS bezüglich ihrer molekularen Massen analysiert und damit der Genotyp (homo- bzw. heterozygot) bestimmt. Weitere Erklärungen im Text.

vor dem zu untersuchenden SNP zum Liegen kommt. Für die folgende Verlängerung kommen zwei unterschiedliche Strategien zum Einsatz. Beim GOOD- [3], dem PinPoint- [4]

und dem GenoSNP-Assay [5] werden der Reaktion nur Didesoxynucleotide (ddNTPs) zugesetzt, was zu einem Elongationsstopp bereits nach Einbau einer einzelnen Base

führt. Je nach Massenzunahme des Oligonukleotids in der MALDI-TOF-MS lässt sich das eingebaute Nukleotid und damit die allelische Variante des SNPs bestimmen. Ein wesentlicher Nachteil dieser Strategie ist die nur geringfügige Massenzunahme durch die Verlängerung des Oligonukleotids um eine Base (**Abb. 2**, linkes Schema). Zieht man die Möglichkeit in Betracht, dass durch die Anlagerung von niedermolekularen Substanzen wie Natrium- oder Kalium-Ionen (M+23 Da bzw. M+39 Da) die Extensionsprodukte in ihren Massen zusätzlich verändert werden können, so führt dies unweigerlich zu einer erschwerten Auswertung und möglicherweise zur Überlagerung von informativen Signalen. Eine Alternative dazu besteht in der zweiten Strategie, die im PROBE- [6], VSET- [7] oder MassEXTEND-Assay [8] zum Einsatz kommt. Hierbei wird der *primer extension*-Reaktion eine Mischung aus Desoxynukleotiden (dNTPs) und ddNTPs zugesetzt, was – je nach Wahl der Nukleotide – zur Verlängerung um eine oder mehrere Basen führt. Dadurch rücken informative Massensignale im MALDI-TOF-Massenspektrum weiter auseinander, und die einwandfreie Zuordnung der Allelvarianten wird erleichtert (**Abb. 2**, rechtes Schema).

Im Multiplex-Verfahren können mehrere verschiedene SNPs in nur einem Reaktionsansatz analysiert werden. Mittlerweile ist es möglich, auf diese Weise bis zu 30 Polymorphismen parallel zu bestimmen. Ermöglicht wurde dieser hohe Grad an Komplexität durch die Entwicklung eines Algorithmus zur automatisierten Auswahl von miteinander kompatiblen Oligonukleotiden und gleichzeitiger BLAST-Suche zur Gewährleistung ihrer Spezifität [9].

Mikrosatelliten-Instabilität (MSI) – *Could you please repeat this*

MSI repräsentiert eine weitere wichtige genetische Variation, die unter anderem auch mit Krebsentstehung in Verbindung gebracht wird. Mikrosatelliten sind kurze repetitive Sequenzen, die über das gesamte Genom verteilt vorkommen. Dabei gibt es deutliche interindividuelle Unterschiede bezüglich der Anzahl an Wiederholungen innerhalb eines Mikrosatelliten. Dies ermöglicht ihre Nutzung als Marker für diverse Anwendungen in der Humangenetik, z. B. für Verwandtschaftsanalysen. Gesunde Individuen zeigen bereits einen gewissen Grad an Längenvariation von Mikrosatelliten, da diese Bereiche *hotspots* für eine fehlerhafte Replikation darstellen

und für die DNA-Reparaturmechanismen in exakter Länge schwer aufrechtzuerhalten sind. Eine erhöhte Abweichung von der „normalen“ Länge eines Mikrosatelliten deutet auf Defekte im DNA-Reparatursystem hin und wird als Mikrosatelliten-Instabilität bezeichnet (MSI). MSI innerhalb codierender Mikrosatelliten (cMS) kann zur Pathogenese von Krebserkrankungen mit DNA-Reparatur-Defizienz beitragen [10]. Im Besonderen sind die Mikrosatelliten betroffen, die mit der erblichen Form des kolorektalen Karzinoms (*hereditary non-polyposis colorectal cancer*, HNPCC) assoziiert sind, einer in den westlichen Industrieländern häufig auftretenden Krebserkrankung. Die Genotypisierung dieser MSI ist daher eine gängige Methode, um die molekulare Diagnose von HNPCC zu unterstützen.

Die klassischen Techniken zur Detektion von MSI umfassen DNA-Sequenzierung oder Gelelektrophorese, zwei Methoden die sehr zeitaufwendig sind und nur wenige Informationen zu quantitativen Unterschieden liefern. Durch die Anwendung von MALDI-TOF-MS ist die Visualisierung der Längenvariabilität von Mikrosatelliten in einem einzigen Experiment möglich. Bis jetzt wurden zwei Methoden zur Bewertung von MSI entwickelt und eingesetzt. Zum einen die aus der SNP-Genotypisierung bekannte *primer extension*-Reaktion, zum anderen die *in vitro*-Transkription. Erstere basiert auf der spezifischen Verlängerung eines Oligonukleotids, das strangaufwärts vor dem Mikrosatellit bzw. teilweise überlappend bindet. Durch Wahl einer geeigneten Kombination von dNTPs und ddNTPs wird das Oligonukleotid entsprechend der Länge des Mikrosatelliten bis zur ersten sich nicht wiederholenden Base verlängert. Anhand seiner resultierenden Masse im MALDI-TOF-MS kann die Zahl der Basenwiederholungen bestimmt werden. Diese Strategie wurde bereits erfolgreich beim Genotypisieren von fünf genomischen MSI angewandt, die mit neoplastischen Läsionen assoziiert sind [11]. Für die Analyse von längeren Mikrosatelliten ist diese Methode allerdings aufgrund von Massenbeschränkungen der MALDI-TOF-MS nicht geeignet. Eine Alternative ist die *in vitro*-Transkription, wobei informative RNA-Fragmente generiert werden [12]. Für diesen Ansatz wird die zu untersuchende Sequenz in RNA transkribiert und daraufhin durch molekulare Scheren am 3'- und 5'-Ende des Mikrosatelliten geschnitten. Diese Fragmente können dann wiederum durch MAL-

DI-TOF-MS analysiert und die Länge des Fragments bestimmt werden.

DNA-Methylierung – *the Silence of the Genes*

DNA-Methylierung stellt einen wichtigen epigenetischen Faktor für die Genregulation dar. Diese Modifizierung findet nahezu ausschließlich an Cytosinen innerhalb von CpG-Dinukleotiden statt. CpG-reiche Regionen, sogenannte CpG-Inseln, liegen überdurchschnittlich oft in bzw. in der Nähe von Promotorregionen. Im Allgemeinen vermitteln stark methylierte CpG-Inseln das Abschalten von Genabschnitten, demethylierte Abschnitte unterliegen dagegen erhöhter Transkription. Wie bereits SNPs oder MSI können auch Veränderungen im Methylierungsmuster an einer Vielzahl erblicher Erkrankungen beteiligt sein. Verschiedene Methoden wurden bisher entwickelt, um die Methylierung bestimmter DNA-Bereiche zu untersuchen [13]. Da bei einer PCR sämtliche Methylgruppen verloren gehen, wird ein zusätzlicher Arbeitsschritt benötigt, der den Methylierungsstatus vor der Vervielfältigung „einfriert“. Dabei wird die DNA zuvor mit Bisulfit behandelt, wobei nicht-methylierte Cytosine in Uracile umgewandelt werden. Cytosine, die zuvor methyliert vorlagen, sind vor der Bisulfit-Umwandlung geschützt und bleiben unverändert. Während der Vervielfältigung durch die PCR wird gegenüber von jedem neuen Uracil ein Adenin, gegenüber von den unveränderten Cytosinen ein Guanin eingebaut. Durch die Erzeugung dieser künstlichen SNPs ist es nun möglich, den ursprünglich vorhandenen Methylierungsstatus eines DNA-Abschnitts mit den gängigen Methoden der MALDI-TOF-MS-SNP-Analyse zu untersuchen. ■

Literatur

- [1] Karas M, Hillenkamp F (1988) Laser desorption/ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* 60:2299–2301
- [2] Humeny A, Bonk T, Berkholz A et al. (2001) Genotyping of thrombotic risk factors by MALDI-TOF mass spectrometry. *Clin Biochem* 34:531–536
- [3] Sauer S, Lechner D, Berlin K et al. (2000) Full flexibility genotyping of single nucleotide polymorphisms by the GOOD assay. *Nucleic Acids Res* 28:E100
- [4] Haff LA, Smirnov IP (1997) Single-nucleotide polymorphism identification assays using a thermostable DNA polymerase and delayed extraction MALDI-TOF mass spectrometry. *Genome Res* 7:378–388
- [5] Wenzel T, Elssner T, Fahr K et al. (2003) Genosnip: SNP genotyping by MALDI-TOF MS using photocleavable oligonucleotides. *Nucleosides Nucleotides Nucleic Acids* 22:1579–1581
- [6] Braun A, Little DP, Koster H (1997) Detecting CFTR gene mutations by using primer oligo base extension and mass spectrometry. *Clin Chem* 43:1151–1158
- [7] Sun X, Ding H, Hung K et al. (2000) A new MALDI-TOF based mini-sequencing assay for genotyping of SNPs. *Nucleic Acids Res* 28:E68

- [8] Buetow KH, Edmonson M, MacDonald R et al. (2001) High-throughput development and characterization of a genome-wide collection of gene-based single nucleotide polymorphism markers by chip-based matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *Proc Natl Acad Sci USA* 98:581–584
- [9] Rachlin J, Ding C, Cantor C et al. (2005) MuPlex: multi-objective multiplex PCR assay design. *Nucleic Acids Res* 33:W544–W547
- [10] Thibodeau SN, Bren G, Schaid D (1993) Microsatellite instability in cancer of the proximal colon. *Science* 260:816–819
- [11] Bonk T, Humeny A, Gebert J et al. (2003) Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry-based detection of microsatellite instabilities in coding DNA sequences: a novel approach to identify DNA-mismatch repair-deficient cancer cells. *Clin Chem* 49:552–561
- [12] Sasayama T, Kato M, Aburatani H et al. (2006) Simultaneous genotyping of indels and SNPs by mass spectroscopy. *J Am Soc Mass Spectrom* 17:3–8
- [13] Fraga MF, Esteller M (2002) DNA methylation: a profile of methods and applications. *Biotechniques* 33:632, 634, 636–649

Korrespondenzadresse:

Prof. Dr. Cord-Michael Becker
 Dr. Nicolas Vogel
 Institut für Biochemie und Molekulare Medizin
 Emil-Fischer-Zentrum
 Friedrich-Alexander-Universität Erlangen-Nürnberg
 Fahrstraße 17
 D-91054 Erlangen
 Tel.: 09131-8524190 (Becker), -8524196 (Vogel)
 Fax: 09131-8522485
 cmb@biochem.uni-erlangen.de
 nico.vogel@biochem.uni-erlangen.de

AUTOREN



Angela Seebahn

Jahrgang 1982. 2001–2005 Ingenieurstudiengang der Biopharmazeutischen Technologie an der Fachhochschule Gießen. 2010 Promotion am Lehrstuhl für Biochemie und Molekulare Medizin der Med. Fakultät, Universität Erlangen; dort seit 2009 Betreuung der Massenspektrometrie.



Nicolas Vogel

Jahrgang 1977. 1998–2004 Biologiestudium an der Friedrich-Alexander-Universität Erlangen-Nürnberg. 2009 Promotion am Lehrstuhl für Biochemie und Molekulare Medizin der Med. Fakultät, Universität Erlangen; dort seit 2009 Postdoktorand.



Cord-Michael Becker

Jahrgang 1957. 1976–1984 Medizinstudium an der Universität Kiel. 1985 Promotion. 1990 Habilitation für das Fach Molekulare Neuropharmakologie/Neurobiologie. 1994–1995 Hermann-und-Lilly-Schilling-Professur für Med. Forschung, Universität Heidelberg. Seit 1995 Lehrstuhl für Biochemie und Molekulare Medizin, Institut für Biochemie der Med. Fakultät, Universität Erlangen.