

# Anwendungen der DNA-Array-Technologie in der Laboratoriumsmedizin

## Uses of DNA Microarray Technology in Laboratory Medicine

P. Cullen<sup>1,2</sup>, S. Lorkowski<sup>1</sup>

**Zusammenfassung:** Die kürzlich erfolgte Sequenzierung des menschlichen Genoms hat eine Fülle an neuen Daten generiert. Die in diesen Daten enthaltenen Informationen werden sich schon bald auf die humane Diagnostik im allgemeinen und insbesondere auf die Laboratoriumsmedizin auswirken. Eine der wichtigsten Applikationen, die diese Informationen nachhaltig nutzen wird, ist die Technologie der sogenannten DNA-Mikroarrays, die im Prinzip zweidimensionale Arrangements von DNA-Molekülen darstellen. Diese Technologie ermöglicht eine schnelle und hochgradig parallele Analyse der Genexpression und von Gensequenzen und wird sicherlich bald im großen Umfang für Routineanalysen im diagnostischen Labor eingesetzt werden. Der vorliegende Artikel beschreibt die wichtigsten aktuellen Applikationen dieser Technologie und gibt einen kurzen Ausblick auf zukünftige Entwicklungen.

**Schlüsselwörter:** DNA-Mikroarrays; DNA-Chips; Laboratoriumsmedizin; Diagnostik; humanes Genom.

**Summary:** The recent decoding of the human genome has provided a wealth of new information, which will soon impact human diagnostics in general, and laboratory medicine in particular. One of the most important applications that will use this information is DNA microarray technology. DNA microarrays are two-dimensional arrangements of DNA molecules, which allow rapid and massively parallel analysis of gene expression and sequence. We will soon see routine use of this technology in the field of diagnostics. This article describes the most important current applications of this technology and discusses future developments in this field.

**Keywords:** DNA microarrays; DNA chips; laboratory medicine; diagnostics; human genome.

## Einleitung

Nach der beinahe vollständigen Entschlüsselung des humanen Genoms im letzten Jahr ist die labormedizinische Diagnostik in eine neue Phase eingetreten. In zahlreichen Laboratorien und Unternehmen weltweit werden Werkzeuge entwickelt, um dieses Wissen sowohl in der Forschung als auch in der Entwicklung neuer Diagnostika und Therapeutika umzusetzen. Nach der Phase der „Genomics“, wendet man sich nun den „Transkriptomics“ und den „Proteomics“ zu, die sich mit der Summe aller Transkripte bzw. aller Proteine einer Zelle beschäftigen und im Falle der Proteomics auch versuchen, posttransnationale Modifikation und die Interaktion von Proteinen untereinander oder mit ihren Liganden aufzuklären. Sogar der Begriff von „Metabolomics“, die Summe aller metabolischen Interaktionen innerhalb einer Zelle, macht derzeit die Runde.

Diese Ansätze zeichnen sich durch die folgenden drei Merkmale aus:

1. Massive Parallelität. Es besteht nun die Möglichkeit mit neuen Technologien nicht nur Hunderte, sondern oft Tausende oder sogar Zehntausende von Molekülen gleichzeitig zu analysieren.
2. Ein sehr hoher Bedarf an Datenverarbeitung und Integration mit anderen Datenquellen, die sogenannte Bioinformatik. Oft gewinnen ermittelte Daten erst durch diesen Schritt überhaupt eine biologische Bedeutung.
3. Hoher Durchsatz. Durch Automatisierung und Standardisierung wird versucht, eine möglichst hohe Probenzahl zu analysieren.

In den kommenden Jahren wird durch Anwendung der neuen Technologien unser Wissen über die molekulare Funktionsweise und die Komplexität lebender Systeme in bisher ungeahnter Weise vertieft.

In der Laboratoriumsmedizin werden nach unseren Einschätzungen die ersten Testsysteme dieser neuen technologischen Welle integriert, die aus dem Bereich der DNA-Mikroarray-Technologie, die oft allgemein als „Genchip“-Technologie bezeichnet wird, stammen. Aus diesem Grund werden wir uns in der vorliegenden Übersichtsarbeit auf eine Beschreibung dieser Technologie und ihrer Anwendungen beschränken.

<sup>1</sup>Institut für Arterioskleroseforschung, Domagkstraße 3, 48149 Münster

<sup>2</sup>Ogham Diagnostics GmbH, Mendelstraße 11, 48149 Münster  
Korrespondenzadresse: PD Dr. Paul Cullen, Institut für Arterioskleroseforschung, Domagkstraße 3, D-48149 Münster.  
E-mail: cullen@uni-muenster.de

## Hauptergebnisse der Entschlüsselung des menschlichen Genoms

Um die labormedizinische Anwendung von DNA-Mikroarrays in den richtigen Kontext zu setzen, ist es zunächst nützlich, die Hauptergebnisse der Entschlüsselung des humanen Genoms darzustellen. Das humane Genom wurde im Jahr 2000 von einem öffentlichen internationalen Konsortium, geführt vom amerikanischen Wissenschaftler *Francis Collins*, sowie von dem privaten Unternehmen Celera Genomics, unter der Leitung von *Craig Venter*, entschlüsselt und zeitgleich in den Zeitschriften *Nature* [1] und *Science* [2] im Jahr 2001 veröffentlicht. Die Ergebnisse beider Gruppen stimmen in den wesentlichen Punkten gut überein. Das menschliche Erbgut besteht aus circa 3,2 Milliarden Basen in 24 Molekülen mit einer Gesamtlänge von etwa 1,8 m. Etwa 25 % des Genoms scheint keine Gene zu enthalten, und nahezu weitere 50 % besteht hauptsächlich aus hochrepetitiven Elementen (vom Biologen *Richard Dawkins* als „selfish DNA“ bezeichnet [3]). Etwa 28 % des Genoms kodieren für RNA und enthalten nach derzeitigen Schätzungen zwischen 30.000 und 40.000 Gene. Diese Zahl ist viel niedriger als frühere Schätzungen, die von möglichen 100.000 bis 150.000 Genen ausgingen. Etwa siebenhundert Gene des menschlichen Genoms sind RNA-Gene, die hauptsächlich t-RNA kodieren. Aufgrund von Rekombination, die viel häufiger vorzukommen scheint, als bisher angenommen wurde, geht man aber davon aus, daß die etwa 35.000 Gene für bis zu 250.000 verschiedene Proteine kodieren könnten. Das „Durchschnittsgen“ erstreckt sich auf 27 kB des Genoms und enthält neun Exone über eine Länge von 1,3 kB, die wiederum für 447 Aminosäuren kodieren. Man geht davon aus, daß etwa die Hälfte aller Gene Spleißvarianten aufweist. Es darf jedoch nicht außer Acht gelassen werden, daß die Länge der Gene sehr stark variiert. Das Titin-Gen, das Gen mit der längsten kodierenden Sequenz im Genom, enthält als Beispiel 80.780 Basen in mehr als 170 Exonen und kodiert für ein Protein mit über 30.000 Aminosäuren.

Ein wichtiges Ergebnis der Sequenzierung des humanen Genoms ist der hohe Grad an Ähnlichkeit zwischen verschiedenen Spezies. Die Homologie zwischen Menschen und Schimpansen beträgt zum Beispiel über 99 %, während die Homologien auf der Proteinebene zum Nematoden *Caenorhabditis elegans* 46 % und zur Fruchtfliege *Drosophila melanogaster* 61 % betragen. Sogar zu einzelligen Hefen besteht eine 41 %ige Homologie auf der Proteinebene. Auch untereinander sind sich die Menschen erstaunlich ähnlich: zwischen allen Mitgliedern der Spezies *Homo sapiens* besteht nämlich eine Homologie von über 99,9 % (siehe dazu Abbildungen 1 und 2). Es wird vermutet, daß über 90 % der Diversität zwischen Menschen aufgrund von etwa 60.000 Einzelnukleotidpolymorphismen (auf Englisch *single nucleotide polymorphisms* oder SNPs) in den kodierenden Regionen des Genoms zustande kommt.

Für beide Sequenzierungsgruppen lag ein großer Teil der Arbeit darin, den verschiedenen DNA-Sequenzen eine Funktion bzw. sie einem Gen zuzuweisen. Durch diese Prozedur, die als Annotation bezeichnet wird, wurden in bisher nicht-sequenzierten DNA-Regionen putative Gene identifiziert. Obwohl beide Konsortien mehrere unabhängige Methoden verwendeten, stellte die Suche nach Homologien zu bekannten Genen in Menschen und anderen Spezies für beide Gruppen das wichtigste Werkzeug dar. Durch diese Prozedur gelang es den Gruppen, etwa 58 % der putativen Proteine, die durch das Genom kodiert werden, eine Funktion zuzuweisen. Die größte Gruppe stellte die Transkriptionsfaktoren dar, gefolgt von Rezeptoren, Komponenten des Zytoskeletts und Protoonkogenen.

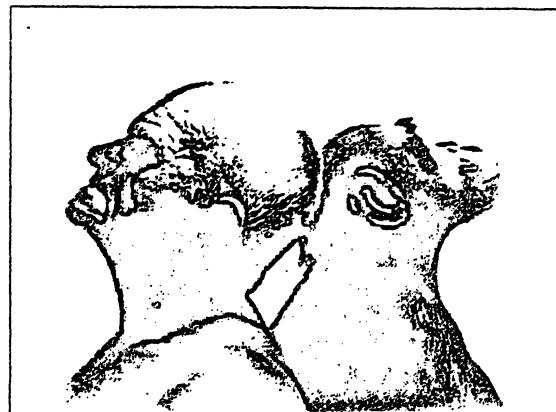


Abbildung 1 Die Homologie der Genome des Menschen und des Schimpansen beträgt mehr als 99 %. Obwohl beide Spezies genetisch fast identisch sind, ist ihr Phänotyp und ihr Habitus doch sehr verschieden. Die Ursachen für diese unterschiedliche Ausprägung sind bislang nicht eindeutig erklärt. Nach Meinung einiger Wissenschaftler, wie zum Beispiel Prof. Svante Pääbo, ist die unterschiedlich starke Genexpression ein Grund hierfür. Aus diesem Grund untersucht der Leipziger Wissenschaftler auch die Genexpression in den Gehirnen von Menschen und Affen. Nach seiner Meinung ist die stärkere Expressionsaktivität zahlreicher Gene (im Mittel etwa dreimal mehr) auch möglicherweise für die größere Intelligenz des Menschen verantwortlich.



Abbildung 2 Obwohl die genetische Übereinstimmung zwischen Menschen noch kleiner ist als die zum Affen (sie ist größer als 99,9 %), unterscheiden sich einzelne Menschen (im Bild Gina Lollobrigida und Rowan Atkinson) in ihrem Aussehen und ihrem Verhalten sehr stark voneinander.

nen. Weitere wichtigen Gruppen waren die Kinasen, Matrixproteine, Immunglobuline, Motor-Proteine und Proteine, die eine Rolle spielen bei der interzellulären Adhäsion [1, 2].

Vor dem Hintergrund der sehr hohen Homologie zwischen Menschen und anderen Spezies ist es von Interesse, daß die Proteine, die evolutionär gesprochen bei den Menschen „neu“ auftreten, eine Rolle in der Entwicklung des zentralen Nervensystems, in der Steuerung der Organogenese und in den Gerinnungs- und Immunsystemen zu spielen scheinen [1, 2].

Natürlich werden die Daten, die hier angegeben werden, mit großer Wahrscheinlichkeit im Laufe der nächsten Jahre revidiert werden. Insbesondere können wir eine Präzisierung der Genzahl erwarten. Nichtsdestotrotz können wir davon ausgehen, daß die Kartierung des menschlichen Erbguts weitestgehend abgeschlossen ist und daß die Grunddaten somit als verbindlich angenommen werden können.

## Aufbau und Funktionsprinzip von DNA-Mikroarrays

Das Funktionsprinzip eines DNA-Chips ähnelt herkömmlichen Hybridisierungstechniken der Molekularbiologie wie den Northern- oder Southern-Blot-Analysen. Diese Verfahren nutzen die Eigenschaft von Nukleinsäuren aus, miteinander sequenzspezifisch zu hybridisieren. Unter Hybridisierung wird die nichtkovalente Bindung zweier zueinander komplementäre Nukleinsäurestränge verstanden, die auf der Ausbildung von Wasserstoffbrückenbindungen zwischen den heterozyklischen Basen der Nukleinsäuremoleküle beruht. Aufgrund der hohen Spezifität dieser Watson-Crick-Basenpaarung kommt es nur bei einer perfekten Paarung der Basen (Guanin mit Cytosin und Adenin mit Thymin oder Uracil) zu einer Bindung, die ausreichend stark ist, um stabil zu sein. Schon eine einzelne Basenmißpaarung der Nukleinsäuremoleküle beeinträchtigt unter geeigneten Bedingungen die Bindungsaffinität beträchtlich oder unterbindet diese vollständig.

Auf einem DNA-Chip werden Nukleinsäuren bekannter Sequenz, die als Sonden bezeichnet werden, auf einem geeigneten Substrat (Träger) in einem bekannten Muster immobilisiert. Die zu untersuchende Nukleinsäure wird markiert und mit den Nukleinsäuren auf dem Chip hybridisiert. Eine Hybridisierung erfolgt aufgrund der spezifischen Watson-Crick-Basenpaarung nur zwischen exakt komplementären Nukleinsäuremolekülen. Darüber hinaus ist die Intensität des gemessenen Signals zur Menge an hybridisierter Probe proportional, so daß quantitative Aussagen über einzelne an die Nukleinsäuren des Chips hybridisierten Nukleinsäurespezies möglich sind.

Für die Herstellung von DNA-Mikroarrays können entweder Nylonmembranen oder auch Glas oder Silizium als Träger verwendet werden. Als Sonden werden Plasmide oder PCR-Produkte mit 500 bis 5.000 Basen eingesetzt, die mit einer Dichte von 80 bis 100

Spots/cm<sup>2</sup> auf Membranen mit Ausmaßen einiger Quadratzentimeter aufgetragen werden (sogenannte *low density chips*). Eine Weiterentwicklung stellen jedoch Arrays höherer Spottichten (*high density chips*) mit mehr als 10.000 Spots/cm<sup>2</sup> dar. Diese werden allgemein als DNA-Chips oder Genchips bezeichnet [6].

## Anwendung von DNA-Mikroarrays

Grundsätzlich gibt es drei große Anwendungsbereiche für DNA-Mikroarrays:

1. die Analyse der Genexpression,
2. die Detektion von Polymorphismen mit sogenannten SNP-Chips (wird „snip chips“ ausgesprochen) und
3. die Resequenzierung von Genen.

### Analyse der Genexpression

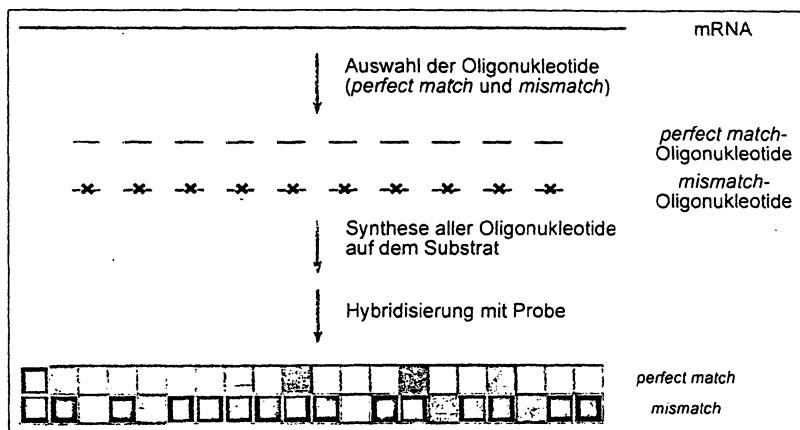
#### *Oligonukleotid-Arrays*

Die amerikanische Firma Affymetrix hat sich auf die Herstellung von DNA-Mikroarrays für die Untersuchung der Genexpression spezialisiert. Solche Arrays werden mittels der Technik der Photolithographie, die direkt aus der Halbleiterindustrie stammt, hergestellt. Mittels dieser Technik kann eine extrem hohe Anzahl an Oligonukleotiden *in situ* auf der Trägeroberfläche hergestellt werden. Der derzeit erhältliche Chipsatz enthält mehr als zwei Millionen Sonden (inklusive Kontrollsonden für z. B. zugesetzte Standards, die als *spikes* bezeichnet werden, und für sogenannte *mismatch*-Sonden) in Anordnungen, die *Features* genannt werden, auf insgesamt fünf einzelnen Chips. Somit werden etwa 12.000 Gene und rund 50.000 *expressed sequence tag clusters* (EST-Cluster) durch den Chipsatz abgedeckt. ESTs sind Teilsequenzen, die aus cDNA-Genbanken identifiziert wurden, die aber bis jetzt keiner eindeutigen Gensequenz zugeordnet wurden.

Pro zellulärer RNA werden auf den Chips von Affymetrix bis zu 20 Oligonukleotide, die etwa 20 Basen lang sind, verstreut auf der Chip-Oberfläche aufgebracht. Diese Oligonukleotide werden anhand ihrer chemischen Eigenschaften so ausgewählt, daß ihre Hybridisierungseigenschaften aufeinander abgestimmt sind. Sie sind zu bis zu 20 verschiedenen Sequenzabschnitten der RNA, vornehmlich aus dem nichtkodierenden 3'-Bereich, exakt komplementär und werden daher als *perfect-match*-Oligonukleotide bezeichnet (Abb. 3).

Das System der Firma Affymetrix hat den Vorteil, daß der Anteil an nicht spezifischen Hybridisierungen am Gesamtsignal ermittelt werden kann. Dazu wird für jedes *perfect-match*-Oligonukleotid ein weiteres, sogenanntes *mismatch*-Oligonukleotid auf den Chip aufgebracht. Das *mismatch*-Oligonukleotid unterscheidet sich von dem *perfect-match*-Oligonukleotid dadurch, daß in der Mitte des Oligonukleotids eine einzelne Base ausgetauscht ist. Unter idealen Bedingungen findet zwischen der Probe und dem *mismatch*-Oligonukleotid keine Hybridisierung statt. Signale, die dennoch

**Abbildung 3** Expressionsanalyse mit dem *perfect-match/mismatch*-System eines Oligonukleotid-Arrays. Aus einer mRNA werden mehrere Bereiche ausgewählt, die in Form von Oligonukleotiden auf dem Chip erzeugt werden (*perfect-match*-Sonde). Zusätzlich werden *mismatch*-Oligonukleotide synthetisiert, die sich nur, dahingehend von den *perfect-match*-Oligonukleotiden unterscheiden, daß in der Mitte des Oligonukleotids eine einzelne Base ausgetauscht wird. Durch den Vergleich der Signalintensitäten zwischen *perfect-match*- und *mismatch*-Proben kann entschieden werden, ob die Hybridisierung an die *perfect-match*-Sequenz spezifisch ist oder nicht. Feld 1 zeigt ein Beispiel, wo markierte RNA weder an die *perfect-match*- noch an die *mismatch*-Oligonukleotide hybridisiert hat. Weil kein Unterschied in der Intensität zwischen *perfect*- und *mismatch*-Signal vorhanden ist, wird dieses Feld nicht ausgewertet. Feld 3 kann ebenfalls nicht ausgewertet werden, da sowohl die die *perfect-match*- wie auch die *mismatch*-Signale gleich stark sind. Diese heißt, das die Probe unspezifisch hybridisiert hat. Alle anderen Felder werden ausgewertet, da hier spezifische Hybridisierungen vorliegen.



an Stellen mit *mismatch*-Oligonukleotiden auftreten, müssen daher von unspezifisch hybridisierten Nukleinsäuren hervorgerufen werden. Durch einen Vergleich der Signalintensitäten zwischen der *perfect-match*- und *mismatch*-Probe kann entschieden werden, ob die Hybridisierung an der *perfect-match*-Sequenz spezifisch ist oder nicht. Anhand dieses Vergleichs können mit einem geeigneten Algorithmus für die *perfect-match*-Oligonukleotide korrigierte Intensitäten berechnet werden, die dann mit den Intensitäten eines zweiten Chips, der mit einer weiteren Probe hybridisiert wurde, verglichen werden.

#### cDNA-Arrays

Andere Unternehmen, zum Beispiel die amerikanische Firma Incyte, haben sich auf die Herstellung von Arrays, die cDNAs als Sonden enthalten, spezialisiert. Diese Sonden werden in der Regel durch eine Polymerase-Kettenreaktion (PCR) aus klonierten cDNA-Abschnitten hergestellt.

Um den Expressionszustand eines Gewebes oder einer Zellkultur zu messen wird die mRNA der Probe zunächst zu cDNA mittels reverser Transkription umgeschrieben, die in der Regel den Expressionszustand des zu untersuchenden zellulären Systems wiederspiegelt. Die cDNA wird oftmals mit Fluoreszenz-Farbstoffen während der reversen Transkription markiert und anschließend mit dem DNA-Array hybridisiert.

Die Mengen an hybridisierten Nukleinsäuren können anhand der Intensitäten der detektierten Signale ermittelt werden, da in bestimmten Grenzen eine lineare Proportionalität zwischen Signalintensität und der Menge an gebundener Probe besteht.

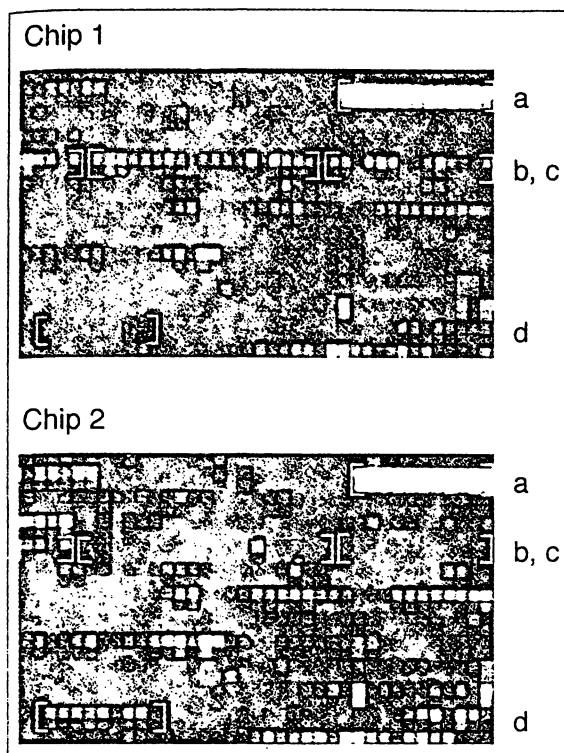
Für die meisten Anwendungen ist jedoch die erhältliche cDNA-Menge zu gering. Ursache hierfür ist zum einen die niedrige Effizienz der reversen Transkription. Zum anderen steht bei zahlreichen Versuchsansätzen (Biopsien, Blutproben etc.) nur sehr wenig Ausgangsmaterial zur Verfügung. Dies macht für viele Anwendungen eine Vermehrung (Amplifikation) der cDNA obligat.

Diese Vermehrung kann durch eine Amplifikation der cDNA mittels einer Polymerase-Kettenreaktion erfolgen, die mit wenigen Syntheszyklen durchgeführt wird. Auf diese Weise wird vermieden, daß unterschiedliche Reaktionskinetiken und -verläufe der verschiedenen Nukleinsäuren während der Amplifikation das Verhältnis einzelner Nukleinsäuren zueinander verfälschen.

Ein anderes Verfahren, das ohne eine PCR-Amplifikation auskommt, ist die *in vitro*-Transkription der cDNA. Dazu wird während der Reversen Transkription ein Promotor für eine RNA-Polymerase an den cDNA-Strang angefügt, der die Transkription der cDNA in cRNA (komplementäre RNA) ermöglicht. Diese cRNA, die mittels einer RNA-Polymerase linear amplifiziert wird, wird mit dem Array hybridisiert.

#### Bestimmung von Expressionsunterschieden

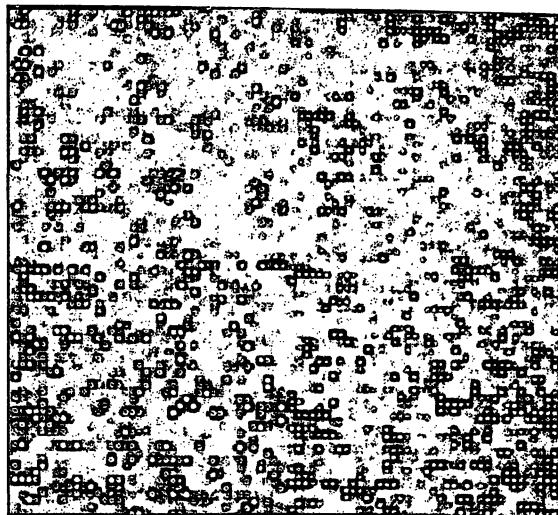
Zur Bestimmung der Expressionsunterschiede zwischen zwei Proben gibt es prinzipiell zwei Verfahren: Beim ersten Verfahren werden zwei identische Chips verwendet, die jeweils mit den markierten Nukleinsäuren zweier verschiedener Proben hybridisiert werden. Die Signalintensitäten der einzelnen Gene werden mittels interner Standards oder statistischer Verfahren nor-



**Abbildung 4** Ausschnitt aus einem Fluoreszenzbild zweier Oligonukleotid-Chips, die mit cRNA-Proben hybridisiert wurden, die aus der mRNA von zwei unterschiedlich kultivierten humanen Makrophagen-Populationen gewonnen wurden. Die markierten Bereiche zeigen Oligonukleotide (*perfect match* und *mismatch*), die mRNAs repräsentieren, die in den beiden Kulturen unterschiedlich stark (b, c, d) bzw. gleich stark (a) exprimiert werden.

malisiert. Hierfür können zum einen Gene verwendet werden, die in Zellen gleichen Typs ein identisches Expressionsniveau aufweisen. In der Praxis werden häufig die mRNAs der Gene für  $\beta$ -Actin, Ribosome und Glycerinaldehyd-3-Phosphat-Dehydrogenase als interne Standards (sogenannte *housekeeping*-Gene) verwendet [5]. Zum anderen kann bei Arrays hoher Dichte die Gesamtintensität aller Signale gemittelt und als Basis einer Normalisierung verwendet werden. Unterschiede zwischen den normalisierten Signalintensitäten deuten auf Expressionsunterschiede hin (Abb. 4).

Bei den Oligonukleotidarrays sind die Hybridisierungseigenschaften der Sonden aufeinander besser abgestimmt als bei den cDNA-Arrays (siehe unten). Nach entsprechender Korrektur der Signale der *mismatch*-Oligonukleotide werden die gemittelten Signalintensitäten eines DNA-Arrays mit den Intensitäten eines zweiten Chips, der mit einer weiteren Probe hybridisiert wurde, verglichen. Da dieser Vergleich anhand aller Oligonukleotidpaare erfolgt, hängt die Entscheidung, ob unterschiedliche Mengen einer RNA in den beiden Proben vorliegen, im Gegensatz zu anderen



**Abbildung 5** Prinzip der Expressionsanalyse mit zwei Proben auf einem Chip. Derartige Systeme arbeiten mit zwei verschiedenen Farbstoffen zur Markierung der Proben (z. B. mit den Fluoreszenzfarbstoffen Cy3 und Cy5). Dies ermöglicht eine simultane Hybridisierung der Proben auf einem Chip. Abgebildet ist ein Chip aus dem Labor von Dr. Patrick O. Brown, Universität Stanford, der 6.116 Hefe-Gene und zahlreiche Kontrollen enthält. mRNAs, die in einer der beiden Zellpopulationen stärker exprimiert wurden als in der anderen, sind entweder grün oder rot dargestellt, während die gelbe Mischfarbe einer Spots Zustand kommt, wenn die Transkripte in beiden Zellpopulationen gleich stark exprimiert wurden.

DNA-Chips von der Auswertung aller Signale ab (Abbildungen 3 und 4).

Bei einem zweiten Verfahren, das mit nur einem Chip für zwei zu untersuchende Proben auskommt, werden zwei verschiedene Markierungen (z. B. die Fluoreszenzfarbstoffe Cy3, 488 nm Anregung, und Cy5, 594 nm Anregung) für die beiden Proben verwendet. Die 1:1 gemischten Proben werden dann auf einem Chip hybridisiert. Gene, die im gleichen Maße repräsentiert sind, weisen eine Mischfarbe der beiden Markierungsfarben auf, während Gene, die unterschiedlich exprimiert werden, die Farbe der entsprechenden Markierung aufweisen (Abb. 5).

Die beschriebene Vorgehensweise wird in der Regel für Arrays verwendet, die PCR-Produkte oder Plasmide als Sonden auf dem Träger verwenden. Der große Vorteil dieser Arrays besteht in ihrer einfachen Herstellung. Nachteilig macht sich jedoch bemerkbar, daß ein solches aus vielen großen Nukleinsäuren bestehendes Gemisch sehr inhomogene Hybridisierungskinetiken aufweist. Daher sind die Hybridisierungsbedingungen (Temperatur, pH, Salzkonzentrationen) für einzelne Moleküle oft nicht optimal. Dies hat zur Folge, daß häufig unspezifische oder unvollständige Hybridisierungen auftreten, die die Ergebnisse verfälschen können. Ferner sind die Nukleinsäuren der Bibliothek ungeordnet auf dem Substrat aufgebracht und daher für die Nukleinsäuren der zu hybridisierenden Probe teilweise aus sterischen Gründen unzugänglich.

### **Detektion von Polymorphismen und Resequenzierung mittels Genchips**

Eine weitere interessante Anwendung der Oligonukleotid-Chips stellt die Resequenzierung mittels Hybridisierung dar. Dieses Verfahren ermöglicht die Identifizierung von Mutationen und Polymorphismen einzelner Basen in bekannten Gensequenzen. Dazu werden den Genabschnitten entsprechende Oligonukleotide auf dem DNA-Chip aufgetragen. Die Nukleinsäuren einer zu untersuchenden Probe werden amplifiziert und mit den Oligonukleotiden des Chips hybridisiert. Anhand des resultierenden Hybridisierungsmusters kann die Sequenz des zu untersuchenden Fragments bestimmt werden, da die Sequenzen der komplementären Oligonukleotide auf dem Chip bekannt sind.

Das Sequenzieren mittels DNA-Chips ist jedoch nicht unproblematisch, da die zu untersuchende mutierte Sequenz hinsichtlich der Hybridisierungsbedingungen nicht optimiert werden kann. Dies hat zur Folge, daß es zu Fehlern bei der Hybridisierung der Probe mit den Oligonukleotiden auf dem Chip kommen kann. Gerade im diagnostischen Sektor ist jedoch eine eindeutige Bestimmung aller Mutationen notwendig. Daher wurden Resequenzierungsschips entwickelt, bei denen jedes Oligonukleotid an jeder Position in der Sequenz variiert wird. So werden Oligonukleotide mit allen denkbaren einfachen Basenaustauschen aufgetragen. Diese dienen ähnlich wie die *mismatch*-Proben der Expressionschips als Kontrolle für nicht spezifische Hybridisierungen bzw. zur Detektion heterozygoter Proben. Nachteile dieser Sequenzierungsstrategie sind die derzeit noch relativ hohen Fehlerraten im Vergleich zu klassischen Sequenzierverfahren und die fehlende Möglichkeit, multiple Mutationen oder komplexe Sequenzabweichungen wie Insertionen oder Deletionen in einem Fragment identifizieren zu können.

Prinzipiell ist mit dem Sequenzieren durch Hybridisierung auch die komplette Sequenzierung unbekannter DNA-Sequenzen anhand ihres Hybridisierungsmusters möglich. Aufgrund der großen Zahl der dazu benötigten Oligonukleotide resultiert eine sehr schwierige Datenanalyse der großen Datens Mengen, die immens aufwendig ist und sehr für Fehler anfällig sein kann. Daher ist ein einfaches Sequenzieren mit herkömmlichen Methoden bislang effizienter. Dank der sich rasch entwickelnden Möglichkeiten der Bioinformatik könnte dieses Verfahren jedoch bald konkurrenzfähig zu konventionellen Verfahren werden. Daher arbeiten bereits heute Firmen an der Entwicklung von DNA-Chips, die zur universellen Sequenzanalyse unbekannter DNA-Proben geeignet sind.

### **Herstellungsmethoden von DNA-Mikroarrays**

#### **Chip-basierte Systeme**

Zur Herstellung von DNA-Chips stehen prinzipiell zwei Verfahren zur Verfügung, die sich in der Art der Immobilisierung der Nukleinsäuren bzw. in der Art der

Oligonukleotid-Herstellung auf der Substratoberfläche unterscheiden. Im ersten Verfahren werden Nukleinsäuren (PCR-Fragmente, Oligonukleotide etc.) *ex situ* hergestellt und erst dann auf einem Substrat mit unterschiedlichen Techniken aufgebracht. Bei dem zweiten Verfahren werden Oligonukleotide *in situ* auf der Oberfläche des Substrats erzeugt. Zu den ersten Methoden gehören neben dem oben beschriebenen photolithographischen Verfahren das *contact tip deposition printing*, das *micro contact printing* ( $\mu$ CP), die *micro fluidics networks* ( $\mu$ FN) und das *electro capture*. Photolithographische Verfahren verwenden ausschließlich *in situ* auf einem Substrat erzeugte Oligonukleotide. Das *piezoelectric printing* und das *micro wet printing* ( $\mu$ WP) sind prinzipiell geeignet, sowohl vorgefertigte Sonden auf einer Oberfläche aufzubringen, als auch Oligonukleotide direkt auf einem Träger zu erzeugen. Die Details dieser Methoden haben wir an anderer Stelle ausführlich beschrieben [4].

#### **Suspensionsarrays**

Eine Sonderstellung unter den derzeit vorhandenen Systemen nehmen die sogenannten Suspensionsarrays ein, die zum Beispiel von Karl Lackner *et al.* verwendet [6] und von der Firma Luminex (Austin, Texas, USA) als FlowMetrix™-System entwickelt wurden. Dieses Verfahren verwendet eine Suspension mit Mikropartikeln mit Durchmessern von bis zu 5,5  $\mu$ m, die aus Polystyrol und Methacrylat bestehen und in verschiedenen Intensitäten mit Fluoreszenzfarbstoffen markiert sind. Durch die Verwendung von Mikropartikeln unterschiedlicher Größen und Farben entstehen diskrete Partikelpopulationen, die in Durchflußzytometern identifiziert werden können. Die mögliche Anzahl diskreter Populationen ist von den verfügbaren Farbstoffen und Techniken zur Markierung der Beads sowie von der Zahl unterscheidbarer Farben im Durchflußzytometer abhängig. Momentan sind mit zwei Farben 64 verschiedene Partikel herstellbar. Durch eine weitere Farbe kann die Zahl der verfügbaren Populationen auf über 500 erhöht werden.

An die Carboxylgruppen des Methacrylates der markierten Partikel können Oligonukleotide oder Antikörper kovalent über Aminogruppen gebunden werden. Ferner können die Mikropartikel mit Streptavidin gekoppelt werden. Die Bindung der Substanzen erfolgt über eine Biotin-Markierung. Für Expressionsanalysen von Nukleinsäuren werden Oligonukleotide (für Immunoassays Antikörper) derart gekoppelt, daß jeder Partikelpopulation ein spezifisches Oligonukleotid (bzw. ein definierter Antikörper) zugeordnet werden kann. Die Oberfläche eines Partikels bietet ausreichend Platz für eine kovalente Bindung von etwa  $1-2 \times 10^6$  Molekülen. Aufgrund der geringen Größe und Dichte verbleiben diese Partikel mehrere Stunden in Suspension.

Die Analyse von Nukleinsäuren ist mit diesem Verfahren auf zwei Arten möglich, die beide auf der Hybridisierung von Nukleinsäuren einer Probe in einer Bead-Suspension basieren. Beim direkten Hybridisie-

rungssassay erfolgt die Hybridisierung, wie bei den herkömmlichen DNA-Arrays, direkt zwischen einem markierten Amplikon einer Probe und den Oligonukleotiden der Partikel. Der kompetitive Hybridisierungssassay besteht aus zwei getrennten Maßansätzen. Zunächst werden die Beads mit einem fluoreszenzmarkierten Reporter-Oligonukleotid hybridisiert. Diese Bindung wird im Durchflußzytometer als durchschnittliche Fluoreszenz detektiert. In einem zweiten Ansatz wird das Reporter-Oligonukleotid zunächst mit einem Kompetitor (der mit der Nukleinsäure in der Probe identisch ist) inkubiert. Dabei werden Reportermoleküle an den Kompetitor gebunden, die dann nicht mehr zur Hybridisierung an die Mikrobeads zur Verfügung stehen. Es resultiert eine Abnahme der mittleren Fluoreszenz, die nur abhängig von der vorhandenen Kompetitor-Menge ist.

Die Messung der Fluoreszenzintensität und die anschließende Auswertung erfolgt in Durchflußzytometern mit mehreren Fluoreszenzkanälen. Daraus resultiert eine hohe Flexibilität in der Anwendbarkeit der Suspensionsarrays. Mit nur einem Meßgerät sind neben Expressions- und Mutationsanalysen auch nahezu alle gängigen Immunoassays möglich, da neben den Oligonukleotiden auch Antigene, Antikörper, Rezeptoren, Peptide und Enzymsubstrate an die Partikel gebunden werden können.

## Bedeutung der DNA-Mikroarray-Technologie für die Laboratoriumsmedizin

Nach unserer Einschätzung wird die Mikroarray-Technologie in nächster Zukunft Eingang in die routinemäßige Labormedizin finden. Wir gehen davon aus, daß innerhalb von zwei Jahren zahlreiche Mikroarray-basierte Systeme für die Detektion häufiger mutierter Allele auf dem Markt sein werden. Solche Mikroarrays, die unter anderem auch von der Firma Ogham in Münster entwickelt werden, werden es erlauben, für relativ häufige genetisch bedingte Erkrankungen in manchen Fällen eine aufwendige Gensequenzierung zu vermeiden. Zunächst werden einfache Arrays mit niedriger bis mittlerer Spot-Dichte verfügbar sein. Die Probenvorbereitung wird mittels externer Multiplex-PCR erfolgen mit anschließender Hybridisierung und Detektion. Die Detektion in den ersten Systemen wird wahrscheinlich mittels Fluoreszenz-Farbstoffen durchgeführt werden.

Weitere technologische Entwicklungen werden jedoch dieses Bild rasch ändern. Durch Fortschritte in der Mikrosystemtechnik wird es möglich sein, miniaturisierte PCR-Maschinen zu konstruieren, die eine Integration der DNA-Vermehrung, der DNA-Markierung und der Array-Hybridisierung in einem einzigen Modul erlauben – sogenannte „Lab-on-a-chip“-Systeme. Erste Prototypen solcher Systeme sind bereits verfügbar. Solche Module, die als Einwegartikel verfügbar sein werden, werden einen hohen Grad an Automatisierung, eine erhebliche Kostensenkung und einen erhöhten Probendurchsatz erlauben. Mittelfristig wer-

den auch Fluoreszenzfarbstoffe durch andere, preiswerte Detektionssysteme, zum Beispiel mittels Silberfärbung, abgelöst werden [7].

Mit einiger Verzögerung werden unserer Einschätzung nach auch DNA-Arrays der zweiten Generation zur Messung der Genexpression Bedeutung in der Laboratoriumsmedizin erlangen. Derzeit läuft eine Vielzahl von Studien weltweit, um das Expressionsmuster verschiedener Erkrankungen, insbesondere maligner Entartungen, mittels DNA-Mikroarrays mit hoher Dichte zu charakterisieren. Somit werden Merkmale der Genexpression, die mit Variablen wie Prognose und Ansprechen auf Therapie sowie mit verschiedenen Tumor-Subtypen und Tumorstadien korrelieren, identifiziert. Nach Erstellung eines sogenannten RNA-Fingerabdrucks, der oft weniger als einhundert cDNAs enthält, können einfache DNA-Expressionschips der zweiten Generation in großer Stückzahl hergestellt und verwendet werden. Solche DNA-Arrays werden unter anderem Verwendung finden bei der Klassifizierung von Tumoren, bei der Abschätzung der Prognose maligner und anderer Erkrankungen und bei der Auswahl der Therapie.

Ein weiteres Gebiet, welches derzeit große Beachtung findet, ist die Verwendung von DNA-Arrays, die alle SNPs im Genom, die gesamte kodierende Region des Genoms oder sogar selbst das gesamte Genom enthalten, um Korrelationen zwischen dem Vorhandensein von Polymorphismen und das Ansprechen auf Medikamenten festzustellen. Durch dieses Feld, das als Pharmakogenomics bezeichnet wird, erhofft man sich eine weitgehende Personalisierung der pharmazeutischen Therapie, die nicht nur die Ansprechbarkeit des Patienten auf die Therapie, sondern auch die Wahrscheinlichkeit von Nebenwirkungen berücksichtigt. Unseres Erachtens wird es einige Jahre dauern, bis solche Ansätze Routineverwendung in der Laboratoriumsmedizin finden. Dennoch halten wir es nicht für ausgeschlossen, daß in der Zukunft solche Systeme die heutige Medikamentenüberwachung und Toxikologie ergänzen werden.

## Voraussetzungen für den Einsatz von DNA-Chips in der Laboratoriumsmedizin

Grundsätzlich gibt es drei Punkte, deren Etablierung eine notwendige Voraussetzung für den Einsatz von DNA-Arrays, insbesondere von DNA-Chips für die Genexpressionsanalyse, in der Laboratoriumsmedizin ist:

1. Standardisierung der Probenaufarbeitung. Für alle Messungen mit DNA-Chips (dies gilt insbesondere für Genexpressionsanalysen) ist eine sorgfältig standardisierte Entnahme und Aufarbeitung der Proben die wichtigste Voraussetzung. Bereits kleinste Abweichungen von der Standardprozedur können die Expression zahlreicher Gene beeinflussen und müssen daher vermieden werden, um eine korrekte Diagnose zu ermöglichen [8].

2. Ermittlung von Referenzwert. Ähnlich wie dies in den letzten Jahrzehnten bereits für zahlreiche Blutwertparameter geschehen ist, müssen auch für relevante Analysen der Genexpression und von Polymorphismen Referenzwerte ermittelt werden, indem umfangreiche standardisierte Analysen in einem als gesund definierten Teil der Bevölkerung durchgeführt werden. Durch entsprechend ausgelegte Studien müssen die Zusammenhänge zwischen differenzieller Genexpression bzw. genomischen Polymorphismen und einem Krankheitsbild oder einer Prädisposition gezeigt werden.
3. Kalibrierung und Quantifizierung. Damit Analysen mit DNA-Chips sinnvoll in der Laboratoriumsmedizin eingesetzt werden können, ist die Entwicklung standardisierter Kalibrierungsverfahren notwendig, die dann, zum Beispiel im Falle der Genexpressionsanalysen, eine Quantifizierung derselben ermöglichen.

## Ausblick

Wir können davon ausgehen, daß wir erst am Anfang der Revolution, die durch die Entschlüsselung des menschlichen Genoms ausgelöst wurde, stehen. In den kommenden Jahren wird die Analyse der Proteinexpression, die Untersuchung von posttranslationalen Modifikationen, sowie die Untersuchung von Interaktionen zwischen Proteinen immer mehr in den Mittelpunkt rücken. Wie Eingangs erwähnt dürften nach den Genomics und Transkriptomics die Proteomics die nächste große Welle der diagnostischen Möglichkeiten in Gang setzen. Hier wird die Array-Technologie zweifelsohne eine wichtige Rolle spielen, obwohl die Herstellung von sogenannten „Protein-Chips“ eine technische Herausforderung darstellt, die die Herstellung von DNA-Arrays um mindestens eine Größenordnung übersteigt. Nichtsdestotrotz sind auch hier bereits erste Ergebnisse publiziert worden [9, 10]. Wahrscheinlich

werden durch die Kombination der Array-Technologie mit hochempfindlichen Detektions- und Meßverfahren wie der *matrix-assisted laser desorption/ionization time of flight* (MALDI-TOF) ganz neue Möglichkeiten der Diagnostik im medizinischen Labor eröffnet. Erste Ansätze zur kombinierten Verwendung von Massenspektrometrie und Array-Technologie sind bereits in der DNA-Sequenzanalyse und in der Proteinanalytik zu finden. So verwendet beispielsweise die Firma Sequenom (Hamburg) diese Technik, um eine Identifikation von *single nucleotide polymorphisms* und *short tandem repeats* (SNP- und STR-Analysen) mit hoher Genauigkeit und Sensitivität zu ermöglichen.

## Literatur

1. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature* 2001;409(6822):860-921.
2. Venter JC, Adams MD, Myers EW, et al. The sequence of the human genome. *Science* 2001;291(5507):1304-51.
3. Dawkins R. *The selfish gene*. Oxford: Oxford University Press, 1989.
4. Lorkowski S, Lorkowski G, Cullen P. Biochips – Das Labor in der Streichholzschachtel. *Chemie in unserer Zeit* 2000;34(6): 356-73.
5. Warrington JA, Nair A, Mahadevappa M, Tsyganskaya M. Comparison of human adult and fetal expression and identification of 535 housekeeping/maintenance genes. *Physiological Genomics* 2000;2(3):143-7.
6. Lackner KJ, Kilwinski J, Langmann T, Aslanidis C, Schmitz G. Multiplex DNA- und RNA-Analyse an fluoreszenteren Microbeads als Alternative zum DNA-Array. *Medizinische Genetik* 1999;11: 16-17.
7. Taton TA, Mirkin CA, Letsinger RL. Scanometric DNA array detection with nanoparticle probes. *Science* 2000;289(5485): 1757-60.
8. Schuchhardt J, Beule D, Malik A, Wolski E, Eickhoff H, Lehrach H, Herzel H. Normalization strategies for cDNA microarrays. *Nucleic Acids Research* 2000;28(10):E47.
9. MacBeath G, Schreiber SL. Printing proteins as microarrays for high-throughput function determination. *Science* 2000;289(5485): 1760-3.
10. MacBeath G. Proteomics comes to the surface. *Nature Biotechnology* 2001;19(9):828-9.