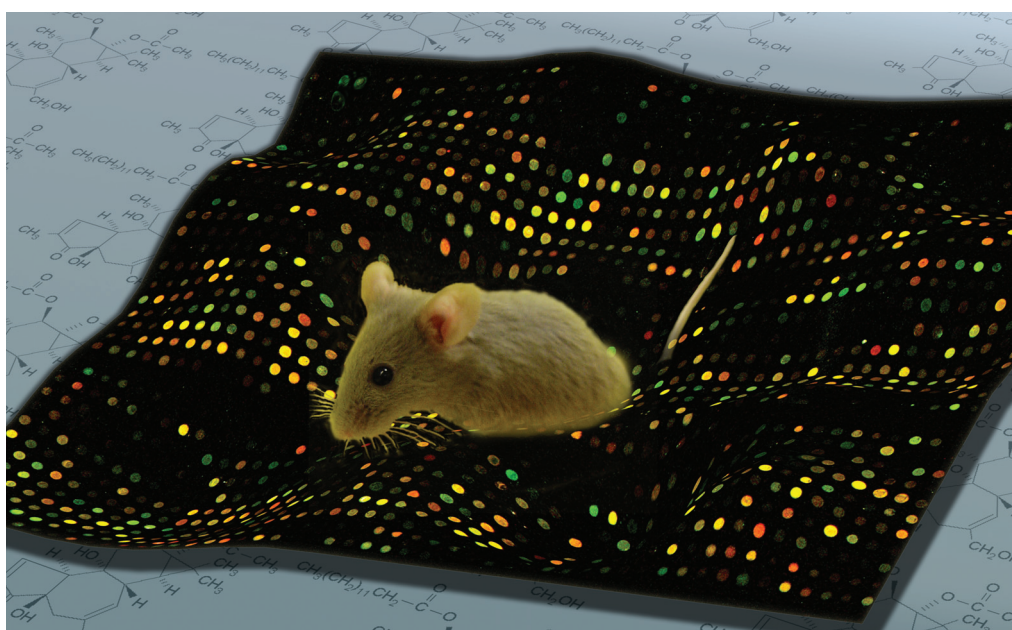


BIO spektrum

ZEITSCHRIFT DER GESELLSCHAFT FÜR BIOCHEMIE UND MOLEKULARBIOLOGIE (GBM),
DER VEREINIGUNG FÜR ALLGEMEINE UND ANGEWANDTE MIKROBIOLOGIE (VAAM),
DER GESELLSCHAFT FÜR GENETIK (GfG), DER GESELLSCHAFT FÜR ENTWICKLUNGSBIOLOGIE
(GfE) UND DER DEUTSCHEN GESELLSCHAFT FÜR EXPERIMENTELLE UND KLINISCHE
PHARMAKOLOGIE UND TOXIKOLOGIE (DGPT)



- DIE ALTERSABHÄNGIGE MAKULADEGENERATION (AMD)
- „ERSTE-KLASSE-TICKET“ FÜR ANTIGENE • DER KERNPoren-KOMPLEX: VON DER STRUKTUR ZUR FUNKTION • tRNA IMPORT IN MITOCHONDRIEN • HORIZONTALER GENTRANSFER IN THERMOPHILEN MIKROORGANISMEN • POLYMER OBERFLÄCHEN FÜR BIOMEDIZINISCHE ANWENDUNGEN • MODIFIZIERTE PROTEINE DURCH KOMBINIERTES SYNTHES UND BIOSYNTHES
- STAND UND PERSPEKTIVEN DER BAKTERIELLEN GENOMFORSCHUNG IN DEUTSCHLAND • **SPECIAL: GENOMICS AND FUNCTION**
- AKTUELLES AUS FORSCHUNG UND TECHNOLOGIE

Special: Genomics and Function

Anwendungen der DNA-Chiptechnologie in der Krebsforschung

Björn Fritz, Felix Kokocinski, Jörg Schlingemann und Meinhard Hahn

Abt. Molekulare Genetik, DKFZ, Heidelberg

Tumore sind nach heutigem Wissen auf Akkumulation genetischer Veränderungen in einzelnen somatischer Zellen zurückzuführen, wobei spezifische Schlüsselgene betroffen sind. Deren Produkte erfüllen essentielle Funktionen in Kontroll- und Regulationsmechanismen, die in gesunden Zellen Wachstums- und Teilungsaktivität auf die Bedürfnisse des Körpers abstimmen. Entgehen Zellen dieser Kontrolle, so kommt es zu unkontrollierter klonaler Proliferation und letztlich zur Tumorentwicklung. Durch Mikroevolutionsprozesse reichern sich in der DNA u. a. Punktmutationen, Alleldeletionen von Tumorsuppressorgenen sowie Amplifikationen von (Proto)-Onkogenen in Tumorzellen und deren Vorläufern an. Viele Gene werden zudem fehlreguliert und zeigen veränderte mRNA- und Protein-Expressionsmuster. Bereits heute werden Erkenntnisse der Molekularen Medizin über solche genetischen Veränderungen für Diagnose, Prognose und Therapie genutzt. Wesentliche Grundlage hierfür sind leistungsfähige und zuverlässige Detektionsverfahren wie sie mit der DNA-Chiptechnologie^[1-4] zur Verfügung stehen.

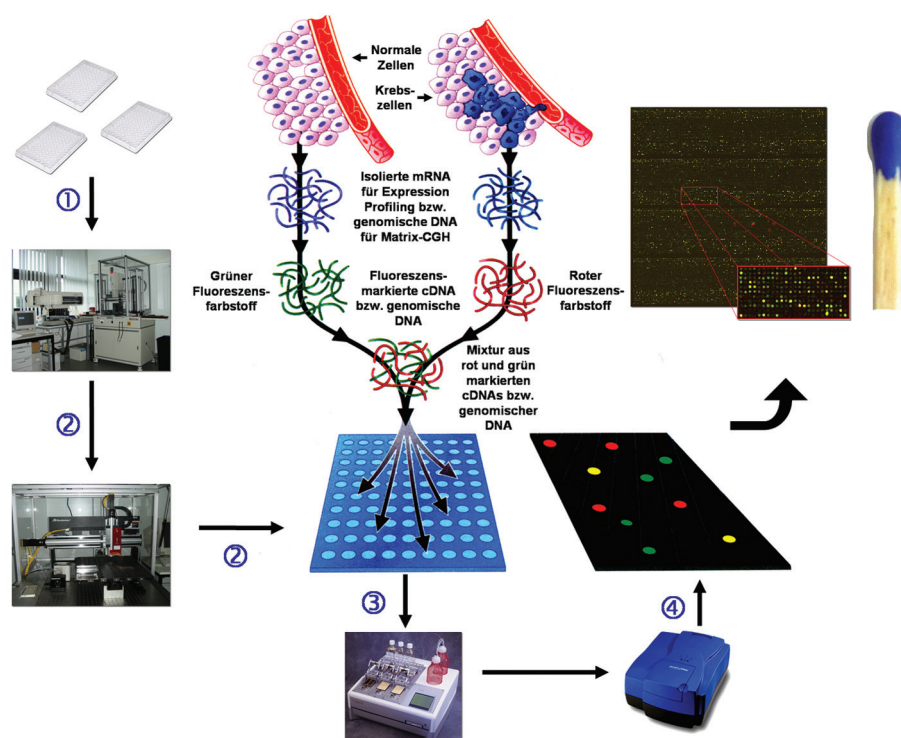


Abb. 1: Workflow-Diagramm für Expression Profiling- sowie Matrix-CGH-Experimente. 1: Herstellung der Zielsequenzen. Ausgehend von cDNA-Klonen (Expression Profiling) bzw. BAC-Klonen (Matrix-CGH) werden die Zielsequenzen unter Einsatz eines Pipettierroboters PCR-amplifiziert. 2: DNA-Chip-Herstellung: Ein Microarray druckt im Kontaktverfahren jeweils 0,2 nl der verschiedenen PCR-Produkte auf speziell beschichtete Glasobjektträger. 3: Hybridisierung: Die komparative Hybridisierung der fluoreszenzmarkierten Proben-DNA sowie anschließende Waschung der Microarrays erfolgt vollautomatisch. 4: Detektion: mittels eines Laserscanners werden die Fluoreszenzsignale ausgelesen und der weiteren Auswertung zugänglich gemacht. Teile der Abbildung mit freundlicher Genehmigung von Aventis Pharma Deutschland GmbH.

Matrix-CGH als Werkzeug der klinischen Diagnostik

► 1992 wurde die Technik der „Comparative Genomic Hybridisation“ (CGH) eingeführt, die erstmals die genomweite Analyse der DNA-Kopienzahl als Funktion der chromosomalen Lokalisation erlaubte. CGH beantwortete also die Frage: An welcher Stelle des Genoms gibt es Verluste oder Zugewinne. Im vergangenen Jahrzehnt trug die CGH grundlegend zu unserem heutigen Wissen über rekurrente Veränderungen in

Tumoren bei und legte die Basis für die Identifizierung einer Reihe von Onkogenen und neuer Tumorklassifikationsschemata. Dass die chromosomale CGH nie den Eingang als klinisches Routineverfahren fand, lag hauptsächlich an geringer Auflösung (beschränkt auf etwa 3–10 Mbp) und fehlender Automatisierbarkeit. Die Auswertung erfordert zudem erfahrene Zytogenetiker. Im Zuge der DNA-Chiptechnologie (Abb. 1) wurden die für die CGH eingesetzten Metaphase-Chromosomen durch genomische Fragmente definierter chromosomaler Lo-

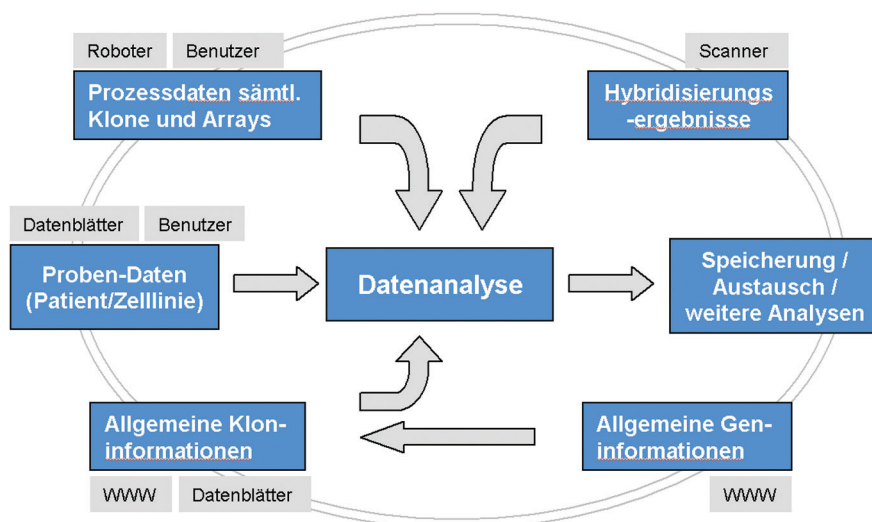


Abb. 2: Datenquellen und Datenvernetzung bei Microarray-Experimenten.

on vergesellschaftet sind, ist zu erwarten, dass derartige maßgeschneiderte Chips ein wertvolles Instrument in der klinische Diagnostik sein werden.

Erstellen von Expressionsprofilen mit cDNA-Microarrays

Zur Untersuchung der Änderung von Transkriptmustern stehen cDNA- und Oligonukleotid-Microarrays zur Verfügung, die die parallele Expressionsanalyse mehrerer Tausend Gene erlauben^[1, 2]. Für cDNA-Arrays werden Sonden für interessante Gene mittels PCR aus einer geeigneten Klonsammlung gewonnen. Nach Aufreinigung und Kontrolle der PCR-Produkte werden diese durch einen Roboter auf beschichtete Glasoberflächen aufgebracht und immobilisiert. Sodann werden Gesamt-RNA oder mRNA

kalisation ersetzt. Diese Technologie wurde in Anlehnung an die herkömmliche CGH als Matrix- oder Array-CGH bezeichnet (Abb. 1)^[3, 4]. Arrays können vollautomatisiert prozessiert und analysiert werden. Die Auflösung der Matrix-CGH wird durch die Größe und den genomischen Abstand der DNA-Elemente bestimmt, die auf dem Chip aufgebracht werden. Als Fragmente kommen Bakterielle Artifizielle Chromosomen (BAC-Klone) in Betracht, die etwa 100–200 kbp genomischer DNA repräsentieren. Inzwischen konnten unsere sowie eine Reihe anderer Arbeitsgruppen zeigen, dass die Matrix-CGH monoallelische Verluste und Zugewinne mit einer Sicherheit nachweist, die der klinischen Diagnostik genügt. Es gibt Bestrebungen, in naher Zukunft Matrix-CGH-Arrays zu entwickeln, die das komplette menschliche Genom lückenlos repräsentieren, wozu etwa 30.000 BAC-Klone notwendig sind. Diese Gesamtgenom-Chips stellen ein ideales Werkzeug dar, um nach bisher unidentifizierten Mikroaberrationen in Tumorgenomen zu suchen. Bestimmte, in Tumoren häufig veränderte Regionen wurden zur Definition der minimal veränderten Region bereits erfolgreich durch Matrix-CGH eingegrenzt.

Ein weiterer Vorteil der Matrix-CGH liegt in der Marker-orientierten Analyse. So wurde in unserem Labor ein CGH-Chip entwickelt, der den genomischen Status mehrerer hundert Onkogene, Tumorsuppressorgene und suspekter genomischer Loci abfragt. Diese Marker-orientierte Analyse führte zur Entwicklung krankheitsspezifischer Arrays, die erfolgreich zur Tumorklassifikation eingesetzt wurden. Da bestimmte zytogenetische Veränderungen nicht nur mit bestimmten Tumorentitäten sondern auch mit verschiedenen Stadien der Tumorprogressi-

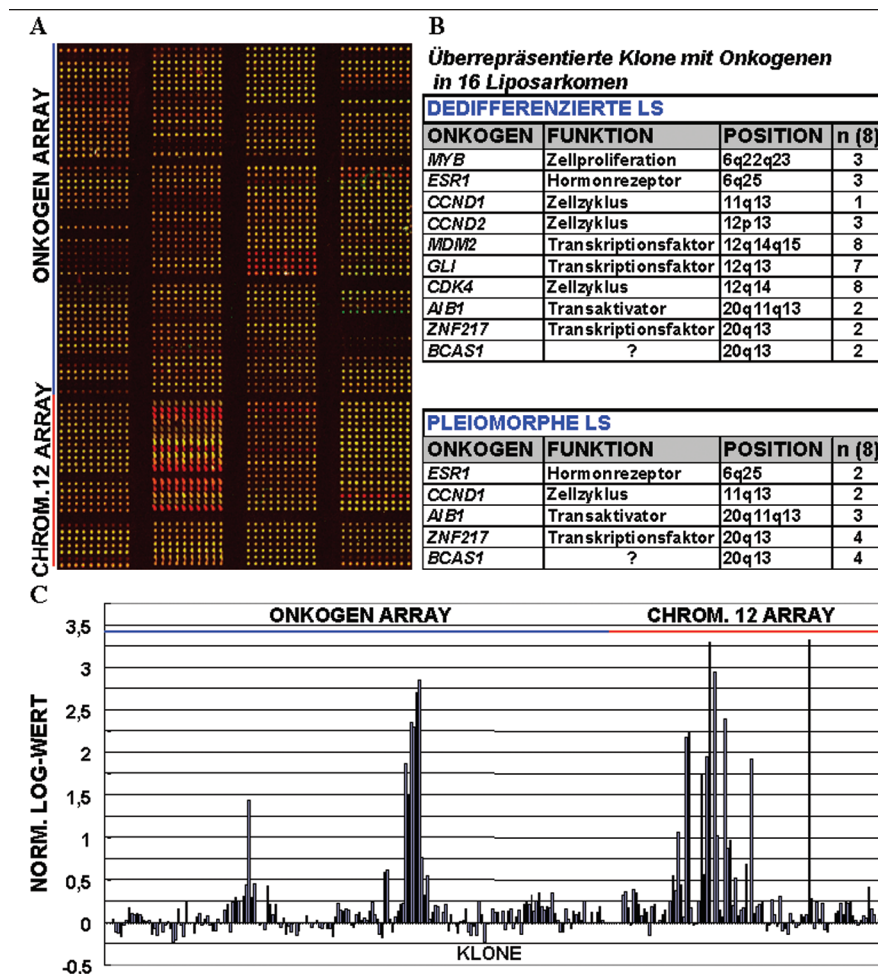


Abb. 3: Matrix-CGH-Analyse. A, Falschfarbenbild der Hybridisierung des Liposarkomfalles DL10 (Tumormarkierung, Cy3, Grün; Kontrollmarkierung, Cy5, Rot) gegen ein Matrix-CGH-Microarray, bestehend aus einem Onkogen-Array mit 188 Klonen und einem für Chromosom 12q spezifischen Subarray mit 112 Klonen. B, Kopienzahlveränderungen von Onkogenen in 8 DL- und 8 PL-Fällen. C, Genomisches Profil: normalisierte Kopienzahlveränderungen (Ordinate, log₂) der genomischen DNA der Probe DL10 und normaler Referenz-DNA. Die BAC-Klone im Onkogen- und 12q-Subarray sind entsprechend ihrer genomischen Abfolge von 1p nach Xq und von 12p cen nach 12q ter geordnet.

aus einer Test- sowie einer Referenz-Probe über eine reverse Transkription mit unterschiedlich fluoreszierenden Nukleotidderivaten separat markiert. Die beiden cDNA-Proben werden vereinigt und unter stringenten Bedingungen auf den Microarray hybridisiert. Wie auch bei der Matrix-CGH werden die nach Anregung mit zwei getrennten Lasersystemen emittierten Fluoreszenzsignale in einem Microarray-Scanner über ein CCD-System sowie Photomultiplier-Röhren detektiert und der weiteren computergestützten Analyse zugänglich gemacht (Abb. 1).

Datenmanagement und Datenanalyse

Hochdurchsatzmethoden wie die DNA-Chiptechnologie erfordern aufgrund der Vernetzung der enormen Datenmengen aus dem komplexen Herstellungsprozess der Arrays (Abb. 1) und dem eigentlichen Experiment maßgeschneiderte bioinformatische Lösungen (Abb. 2).

Prozessdaten. *Labordaten-Systeme* (Laboratory Information Management Systems) werden zur Inventarisierung und Erfassung sämtlicher Produktionsdaten, zur Proben-Rückverfolgung und Fehleranalyse, sowie zur Steuerung des aktuellen Produktions-Protokolls eingesetzt^[5].

Hybridisierungsdaten. Sämtliche Intensitätswerte aus den hybridisierten DNA-Punkten müssen gefiltert (Abzug des Hintergrundsignals) und normalisiert werden, um alle Experiment-spezifischen Einflüsse zu entfernen. Eine mögliche Methode dafür ist die Berücksichtigung von Kontrollwerten, die bekanntermaßen unverändert oder in definierter Weise verändert sind und zu denen sämtliche Werte ins Verhältnis gesetzt werden. Alternativ besteht die Möglichkeit, alle Werte ins Verhältnis zum Gesamt-Mittelwert oder -Medianwert zu setzen, da die Wertemenge üblicherweise normalverteilt ist.

Nach dieser Daten-Vorverarbeitung werden unterschiedliche Methoden aus dem Bereich des *Data Mining* angewendet, deren Ziel es ist, sowohl die Komplexität des Datensatzes zu reduzieren als auch Muster in den Datensätzen zu finden, die biologisch-medizinische Rückschlüsse oder eine Klassifizierung der untersuchten Fälle gestatten. Das beliebteste Verfahren zur Mustererkennung ist das *Clustering*, bei dem Profile entsprechend ihrer Ähnlichkeit gruppiert und mit Hilfe von Farbmuster-Diagrammen und Dendrogrammen visualisiert werden. Methoden zur Klassifizierung „erlernen“ oft anhand bereits klassifizierter Tumorentitäten die Einteilung unbekannter Tumorproben. Erwähnt seien hier Methoden wie *Neuronale Netze*, *self-organizing maps (SOM)*, *de-*

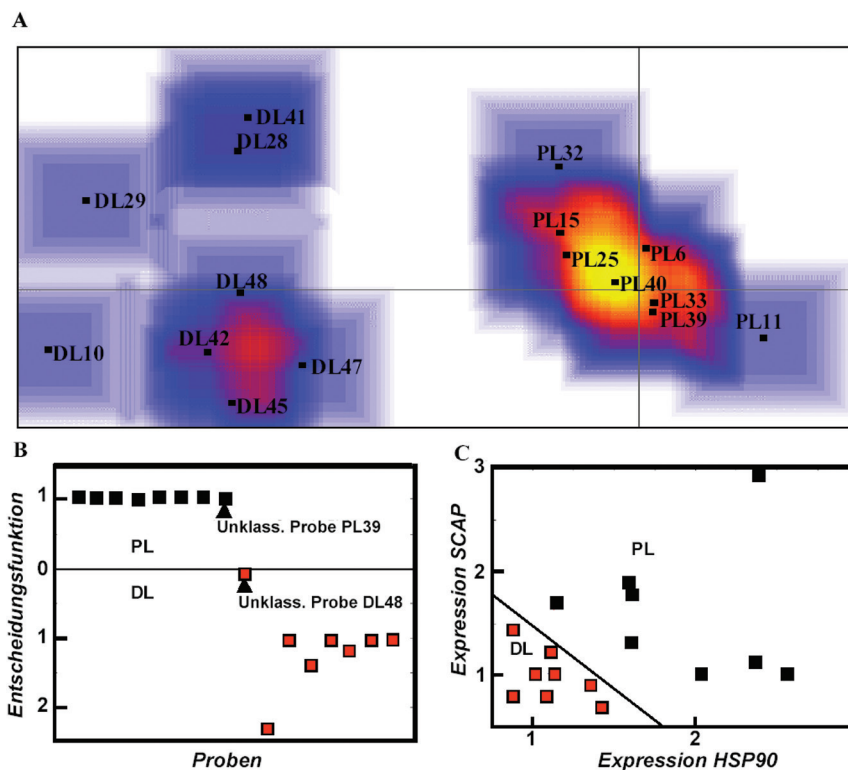


Abb. 4: A, Clusteranalyse von je 8 PL- und 8 DL-Fällen anhand deren Matrix-CGH-Daten und einer Hauptkomponentenanalyse. DL- und PL-Tumore werden eindeutig getrennt, wie aus der unterschiedlichen Ebenenposition zu ersehen ist. B, Klassifizierung der zwei unbekannt Proben PL39 und DL48 auf der Basis eines etablierten Matrix-CGH Klassifizierungsalgorithmus. C, Klassifizierung von PL- und DL-Fällen anhand der Genexpressionsmuster von HSP90 und SCAP. unklass., unklassifiziert.

cision trees, *Hauptkomponenten-Analysen (PCA)* und *support vector machines (SVM)* zur Auswertung von Microarray-Experimenten.

Annotationsdaten. Grundlegend für die Interpretation der Ergebnisse sind umfassende Kenntnisse der Zielsequenzen auf den Arrays. Informationen über kodierte Gene, Sequenzen, chromosomale Lokalisierungen, Funktionshinweise werden mittels Computerprogrammen aus unterschiedlichen Quellen vernetzt und analysiert.

Anwendungsbeispiel: Liposarkome

Sowohl die Charakterisierung der chromosomalen Kopienzahlveränderungen als auch die Genexpressionsanalyse geben Aufschluss über Pathomechanismen, die in der Ätiologie und Entwicklung eines Tumors involviert sind. Darüber hinaus haben sie das Potential, neue diagnostische und prognostische Markermoleküle zu identifizieren. Im vorliegenden Beispiel wurden Matrix-CGH (Abb. 3 A,B,C) und cDNA-Microarrayanalyse zum molekularen Screening von zwei Subtypen von Fettgewebstumoren (DL: differenzierte Liposarkome; PL: pleomorphe Liposarkome) eingesetzt. Zur Verifizierung der Expressionsdaten wurde die quantitative RT-PCR benutzt. Die Kombination

dieser Ansätze führte zur Identifizierung neuer Kandidatengene in der Pathogenese dieser Tumoren. Eine Clusteranalyse zeigte, dass die beiden Subtypen der Fettgewebstumoren, basierend auf dem genomischen Profil, eindeutig zu unterscheiden sind (Fig 4 A). Eine Reihe von sogenannten „Class predictors“ wurde ermittelt, Gene deren Expression bzw. chromosomale Veränderung klassenspezifisch ist. Im Gegensatz zum vielfach angewendeten Einsatz von Cluster- und Klassifikationsanalyse mit Hilfe von Genexpressionsdaten war zum Zeitpunkt dieser Studie das Potential der Matrix-CGH als Klassifizierungsmethode nicht untersucht. Die erfolgreiche Performance des Klassifikationsalgorithmus konnte anhand von zwei Blindproben gezeigt werden, die mit einem Konfidenzintervall > 95% korrekt klassifiziert werden konnten (Abb. 4 B). Die Expressionsanalyse identifizierte eine Reihe von Onkogenen und Tumorsuppressorgenen (wie die *Cyclin abhängige Kinase 4*, *Cyclin D1* und *Apoptose Aktivator HRK*), die gegenüber dem Referenzgewebe stark differentiell exprimiert waren und zeigte in vielen Fällen einen direkten Zusammenhang zwischen zytogenetisch zugewonnenem Material und Überexpression. Darüber hinaus konnte die cDNA-Expressionsanalyse einen

Satz von wenigen Diskriminatoren ermitteln, die eine eindeutige Trennung der Fettgewebstumoren in die beiden Subtypen erlaubte (Abb. 4 C).

Matrix-CGH und Expression Profiling zur globalen Genomcharakterisierung

Beide Microarray-Verfahren ergänzen sich vortrefflich und liefern zusammen eine umfassendere Charakterisierung der einem spezifischen Tumor zugrunde liegenden genetischen Veränderungen. Die Matrix-CGH erfasst Veränderungen der genomischen Kopienzahl von Genen wie Onko- und Tumorsuppressorgenen. Da die hierfür untersuchte genomische DNA wesentlich stabiler ist als die im Expression Profiling analysierte mRNA, sind sogar Paraffin-fixierte Archivmaterialien der Untersuchung zugänglich. Demgegenüber lassen sich mittels Expression Profiling auch Veränderungen der Genexpression aufspüren, die nicht direkt auf Änderungen der Genkopienanzahl zurückzuführen sind, wie sie durch Punktmutatio-

nen in Promotorbereichen oder veränderte Regulationsmechanismen bedingt sein können. Jedoch müssen für diese Untersuchungen nichtdegradierte mRNA-Pools zur Verfügung stehen. In diesem Fall lassen sich für Tumorproben umfassende Informationen über die zugrunde liegenden genetischen Störungen gewinnen.

Beide Techniken werden deshalb bereits heute in der Tumordiagnostik vielfach gemeinsam eingesetzt, wobei ihr Potential bei weitem noch nicht ausgeschöpft ist. Ihr Einsatz im Tumordiagnostikbereich wird in der nahen Zukunft beträchtlich zunehmen.

Literatur

- [1] **Schena, M., Shalon, D., Davis, R. W. und Brown, P. O.** (1995): Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270:467-470
- [2] **DeRisi, J., Penland, L., Brown, P.O., Bittner, M. L., Meltzer, P. S., Ray, M., Chen, Y., Su, Y. A. und Trent, J. M.** (1996): Use of a cDNA microarray to analyse gene expression patterns in human cancer. *Nat. Genet.* 14:457-460

[3] **Lichter, P., Joos, S., Bentz, M. und Lampel, S.** (2000): Comparative genomic hybridization: uses and limitations. *Semin. Hematol.* 37:348-357

[4] **Fritz, B., Schubert, F., Wrobel, G., Schwaenen, C., Wessendorf, S., Nessling, M., Korz, C., Rieker, R. J., Montgomery, K., Kucherlapati, R., Mechtersheimer, G., Eils, R., Joos, S. und Lichter, P.** (2002): Microarray-based copy number and expression profiling in dedifferentiated and pleomorphic liposarcoma. *Cancer Res.* 62:2993-2998

[5] **Kokocinski, F., Wrobel, G., Hahn, M. und Lichter, P.** (2003): QuickLims: Facilitating the data management for DNA-microarray fabrication. *Bioinformatics* 19 (im Druck)

Korrespondenzadresse:

Dr. Meinhard Hahn
Abt. Molekulare Genetik (B060)
Deutsches Krebsforschungszentrum
Im Neuenheimer Feld 280
D-69120 Heidelberg
Tel.: 06221 42 4677
Fax: 06221 42 4639
m.hahn@dkfz.de