

GENOMXPRESS

Informationen aus der deutschen Genomforschung

1.05

März 2005

Springende Erreger und Entzündungen ohne Infektion · Hochdurchsatz gegen den Krebs: Gemeinsam geht's besser! · **MapMan und die Systembiologiedatenbank CSB.DB** · Intraspezifische Introgression – wie die Genetik biochemische Phänotypen beeinflusst · **Koexistenz bei Mais machbar? Ergebnisse des Erprobungsanbaus 2004** · Das Bielefelder GenDB-Annotationsystem · **Reduzierte Genome und das Konzept eines minimalen Gensets** · Firmenportrait BRAIN AG · **Gesichter der Forschung: Portrait Evelyn Möller** Tagungsberichte · **Buchbesprechung** · Science Digest

Hochdurchsatz gegen den Krebs: Gemeinsam geht's besser!

Das Zusammenspiel von Genen und Proteinen wird im Nationalen Genomforschungsnetz erforscht, um neue Ansatzpunkte für die Diagnose und Therapie von Tumorerkrankungen zu finden. Jetzt bündeln die Systematisch-Methodischen Plattformen RNA und Cell ihr Know How im NGFN-Projekt „Funktionelle Netzwerke der Genexpression“ um wichtige Schnittstellen bei der Tumorentstehung zu identifizieren.

Holger Sültmann, Ruth Wellenreuther, Stefan Wiemann, Silke Argo und Annemarie Poustka

Krebs wird durch Störungen der komplexen biologischen Vorgänge in Zellen und Zellverbänden verursacht. Nur die Kenntnis dieser Vorgänge wird es ermöglichen, die besten Angriffspunkte für spezifische Therapien von Tumoren bei gleichzeitiger Minimierung der Nebenwirkungen zu finden.

Auf der Suche nach geeigneten Angriffspunkten hat sich die Analyse von Genexpressionsmustern mit Hilfe von sogenannten „DNA-Chips“ (Microarrays) innerhalb weniger Jahre zu einer der wirkungsvollsten Methoden der systematischen Datengewinnung für funktionelle Studien entwickelt. Veränderungen der Genaktivitäten in Krankheitsgeschehen können mit dieser Technologie genomweit, exakt und schnell untersucht werden. Auf diese Weise werden in der Systematisch-Methodischen Plattform-RNA (SMP-RNA) Experimente durchgeführt, die Aussagen über die Expression nahezu aller menschlichen Gene in einer Tumorprobe liefern. Das Ergebnis solcher genomweiten Expressionsstudien sind meist Listen mit vielen hundert, im Vergleich zur „normalen“ Zelle, differentiell exprimierten Genen.

Die Rolle differentiell exprimierter Gene und ihrer Genprodukten (Proteine) in Tumoren, lässt sich sehr gut in zellulären Systemen untersuchen. Die SMP-Cell hat zu diesem Zweck eine Reihe spezifischer zellulärer Hochdurchsatz-Assays etabliert, die es erlauben, die Auswirkungen der veränderten Genexpression direkt in Zellen und hier insbesondere in Tumor-relevanten Prozessen zu beobachten.

SMP-RNA und SMP-Cell bündeln ihre im Nationalen Genomforschungsnetz (NGFN) etablierten Ressourcen und Technologien in dem gemeinsamen Projekt „Funktionelle Netzwerke der Genexpression“. Zentrales Element dieses Projektes ist das „Effektorgen-Profilung“, in dem die RNA-Interferenz (RNAi) in zellulären Systemen mit der Microarraytechnologie verknüpft wird, um „Effektor“-Gene zu identifizieren und deren komplexe Interaktionen in der Zelle zu modellieren.

Das Ziel ist, hierdurch neue Ansatzpunkte (Targets) für die Tumordiagnose und die gezielte Tumorthherapie zu definieren.

Profile verraten Tumor-Gene

Bereits in der ersten Förderphase des NGFN hat die Genexpressions-Gruppe in der SMP-RNA Microarrays für die Analyse von Nierenzell-, Brustdrüsen- und Gehirntumoren hergestellt und in vielen kooperativen Projekten mit Klinikern genutzt. Aus diesen Projekten liegt eine Fülle von Signaturen der Genexpression vor. So wurden im Nierenzellkarzinom molekulare Muster gefunden, die der histopathologischen Einteilung entsprechen. In bestimmten Mammakarzinomen korreliert die Genaktivität mit der Expression des Östrogenrezeptors und mit morphologischen Charakteristika der Tumoren.

Seit 2003 konzentriert sich die Gruppe auf die Prozessierung und Anwendung von globalen Microarrays mit 37.500 menschlichen cDNA-Klonen. Bis Ende 2004 sind ca. 4.000 dieser Arrays prozessiert und in verschiedenen kooperativen Studien mit klinischen Partnern eingesetzt worden. So zeigte sich ein Genexpressionsmuster, mit dem eine Vorhersage der Metastasierung von klarzelligen Nierenzellkarzinomen möglich wird. Eine andere Studie zur differentiellen Genexpression in Prostatakarzinomen bestä-

tigte die aus der Literatur bereits bekannten Marker und assoziierte darüber hinaus eine Reihe weiterer Gene mit dieser Krebsform.

Die klinischen Kooperationen im Bereich der Microarray-Analyse führten – unter Verwendung einer Vielzahl von Gewebeproben von Tumorpatienten – zur Identifizierung von mehreren Tausend differentiell exprimierten Genen. Hieraus haben wir 1.500 differentiell exprimierte Gene für weitere in den folgenden Abschnitten beschriebene Analysen in den SMPs RNA und Cell ausgewählt. Diese Auswahl beruht auf Vergleichen zwischen Tumor- und Normalgewebe (Nieren- und Prostatakarzinom), verschiedenen Tumor-Subtypen (Mammakarzinom oder Gehirntumoren), Tumorstadien (z. B. Metastasierung) und Tumor-Differenzierung (Grading). Einige charakteristische Tumor-relevante Gene sind Teil dieser Auswahl, die Funktionen der meisten enthaltenen Gene im Zusammenhang mit Tumoren sind jedoch nicht bekannt.

Kandidatengene zeigen ihre Funktionen

In der SMP-Cell wurden zelluläre Assays entwickelt, die ein breites Spektrum Tumor-relevanter molekular- und zellbiologischer Prozesse abdecken (Abb. 1): Ein möglicher Einfluss auf die Zellteilung wird mit einem Proliferations-Assay (BrdU-Einbau während der S-Phase des Zellzy-

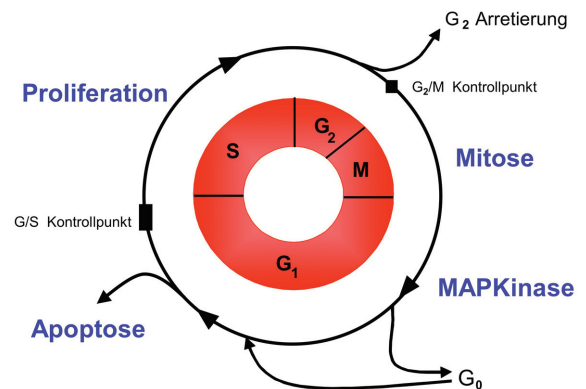


Abb. 1: In der Systematisch-Methodischen Plattform (SMP)-Cell werden am DKFZ Heidelberg Kandidatengene auf ihre Funktionen in verschiedenen Tumor-relevanten Prozessen, beispielsweise im Zellzyklus, getestet.

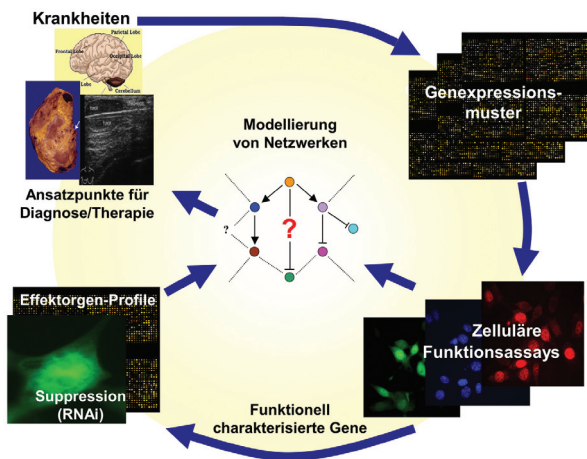


Abb. 2: Das Projekt „Funktionelle Netzwerke der Genexpression“ im Nationalen Genomforschungsnetz zielt auf die Identifizierung bisher unerkannter Geninteraktionen, um diese als potentielle Ansatzpunkte für neue Targets zur Diagnose und Therapie von Tumorerkrankungen nutzbar zu machen.

klus) sowie mit verschiedenen MAPKinase-Assays (aktivierte p44/42, ERK 1/2 und aktivierte p38, PKB/Akt) untersucht. Verringerter programmierter Zelltod dagegen wird in einem Apoptose-Assay (aktivierte Caspase-3) erkennbar. Ein Assay für Kontakt-unabhängiges Wachstum wird u.a. mit Caki1 und Caki2 Nierenzellkarzinom-Zelllinien durchgeführt.

Das experimentelle Design der Assays wurde so gewählt, dass sowohl aktivierende als auch inhibierende Effekte in den transfizierten Zellen gemessen werden können. Weitere Assays werden in der SMP-Cell im Rahmen von NGFN-2 entwickelt und eingesetzt. In einer ersten Studie, in der über 100 verschiedene Proteine auf ihren Effekt in der DNA-Replikation untersucht wurden, fielen 16 Proteine als Modulatoren des Zellzyklus auf. In Kombination mit den Microarray-Ergebnissen konnte zudem eine unmittelbare Tumorrelevanz festgestellt werden.

Einsatz neuer Technologien

Die rasche Fortentwicklung der RNA-Interferenz (RNAi) in den letzten Jahren hin zu einem hochwertigen Werkzeug für die funktionelle Genomforschung wird genutzt, um mit Hilfe von RNAi-Screens in den zellulären Assays die in Tumoren differentiell exprimierten Gene herauszufiltern, die einen unmittelbaren Einfluss auf die jeweils untersuchten Prozesse haben. Die Funktionsweise der RNAi wurde ausführlich in dem Artikel über die SMP RNAi (GenomXPress 4/04) beschrieben. Um ein umfassendes Bild der Genfunktionen zu erhalten, wurde allerdings ein Ansatz gewählt, in dem die Reprimierung der Expression spezifischer Gene – mittels RNAi – durch die gerichtete Überexpression ihrer Proteine – mittels Expressions-Plasmiden – komplementiert wird. Interessanterweise bewirken ein Zuviel und ein Zuwenig der jeweils untersuchten Proteine häufig keine gegensätzlichen Resultate; vielmehr

weisen die Ergebnisse der in NGFN-1 durchgeführten Analysen auf wesentlich komplexere Zusammenhänge wie Konzentrationsabhängigkeiten hin. Vor einem Screening in den Assays wird zunächst die endogene Expression der Kandidatengene mittels quantitativer RT-PCR validiert. Hierfür werden die gleichen Primer/Sonden-Sets verwendet, mit denen im weiteren Verlauf des Projektes der RNAi-induzierte „Knock-down“ verifiziert wird.

Das Ergebnis jedes Assays werden Gene sein, die den hier getesteten Tumor-relevanten Prozess durch ihr Fehlen verändern und so als Prozess-relevant identifiziert werden können. Die Gene werden zusätzlich durch den komplementären Ansatz, die gerichtete Überexpression in den selben Assays, kontrolliert. Hierfür werden vollständige cDNAs mit GFP- und YFP-Tags versehen und in die Zellen transfiziert.

Auf dem Pfad der Interaktionen

Zentrales Element des gemeinsamen Projektes ist die Aufdeckung komplexer Netzwerke der Genregulation. Als Basis der Analysen werden wir Gene verwenden, die signifikante Auffälligkeit in zellulären Assays zeigen, krankheitsrelevante funktionelle Domänen aufweisen sowie eine Homologie zu bekannten, an Tumor-relevanten Prozessen beteiligten Genen, haben. Die Wirkungen und Interaktionen dieser Gene im zellulären Kontext werden wir genomweit mit einer Methode untersuchen, die wir als „Effektorgen-Profilung“ (Abb. 2) bezeichnen. Wissen wir zum Beispiel über ein bestimmtes Gen aus den zellulären Assays, dass es in einem bestimmten Prozess einen Effekt hat, erhebt sich die Frage, auf wie viele und welche der menschlichen Gene dieses Gen wiederum wirkt. Für das Effektorgen-Profilung kombinieren wir dabei die Methode der RNAi mit dem Expression Profiling mit Hilfe von Microarrays. Wir transfizieren siRNA gegen das zu

untersuchende Gen in Zellen geeigneter Zellkultursysteme, um seine endogene Expression zu unterdrücken und um hieraus den Effekt des fehlenden Genproduktes auf weitere Gene im Zellkontext zu erkennen. Wir isolieren die RNA aus diesen Zellen und hybridisieren sie auf die humanen globalen Microarrays (37.500 Gene und ESTs), die uns erlauben, ein genomweites Expressionsprofil zu erstellen und damit die Mehrzahl der Effektorgene zu erfassen.

Wir erwarten, dass unter den Effektorgenen neben funktionell nicht charakterisierten auch bekannte, gut charakterisierte Gene zu finden sein werden, die aber bislang nicht mit biologischen Prozessen in Tumoren in Verbindung gebracht wurden. Diese Gene sind die Anknüpfungspunkte an bekannte Signalwege und damit die Schnittstellen für die Erweiterung des Verständnisses der Netzwerke. Weitere Gene unbekannter Funktion werden möglicherweise Bindeglieder zwischen bisher getrennten Signalwegen sein und unser Verständnis der intra- und interzellulären Signalgebung erweitern.

Einblicke in ein komplexes Zusammenspiel

Weder die zellulären Assays, noch die Expressionsprofile allein werden uns präzise Aussagen über das komplexe Zusammenspiel von Genen und Proteinen ermöglichen. Erst detaillierte Analysen von Kandidatengenen ordnen die Proteine in einzelne Signalwege ein und machen übergeordnete Netzwerke sichtbar. Effektorgene, die in unabhängigen Profilierungen verschiedener Tumor-relevanter Gene des gleichen Prozesses identifiziert werden, gehören möglicherweise zum selben regulatorischen Netzwerk und können einander beeinflussen. Diese Einflüsse sollen validiert werden, indem die Expression der ausgewählten Gene in siRNA-behandelten Zellen mit qRT-PCR gemessen wird. Jene Proteine, für die eine direkte Interaktion vermutet wird, sollen auf ihre Ko-Lokalisation untersucht und mit Hilfe der Massenspektrometrie direkt verifiziert werden. Daten aus anderen Projekten der SMP-Cell (z.B. Proteinlokalisierung, Annotation) werden wesentliche Beiträge zur Stützung der Aussagen leisten.

Modellierung biologischer Netzwerke

Die umfassende statistische Analyse der Hochdurchsatzdaten ist ein essentieller Bestandteil des Projektes. Überprüfung und Sicherung der Datenqualität stehen an erster Stelle, sie werden durch die in die SMP-RNA und SMP-Cell integrierte Bioinformatik gewährleistet.

Die Entwicklung neuer Analysetools und die Datenanalyse der Hochdurchsatzexperimente sind eng mit den Nasslaborarbeiten verzahnt. Methoden, die hier entwickelt und bereits etabliert wurden, sind zur Normalisierung von Microarray-Daten sowie zur Klassifizierung von Tumoren und Genen geeignet. Das Software-Tool "arrayMagic" war ein weiterer Schwerpunkt der Entwicklung. Es stellt nun eine automatisierte Prozessierungspipeline für große Microarray-Datensätze (bis zu mehreren hundert Arrays) dar und bietet umfangreiche Optionen zu Qualitätskontrolle, Normalisierung und Datenkombination.

Neben den Analysetools für Microarraydaten haben wir eine statistische Analyse für die zellulären Assays entwickelt. Die vorhandenen

Methoden werden durch explorative Analysen (z. B. Clustering) sowie vielfältige Möglichkeiten der Daten-Visualisierung ergänzt. Alle entwickelten Tools wurden als Softwarepakete in die frei verfügbare Analyseplattform R [Bioconductor; (www.bioconductor.org)] implementiert.

Hier laufen die Fäden zusammen

Die Modellierung biologischer Netzwerke der Genregulation (Abb. 2) ist der integrierende und schließende Schritt des Projektes. Hier laufen die Fäden zusammen aus der Identifizierung in Tumoren differentiell exprimierter Gene mittels Expressionsprofilen, ihrer in zellulären Funktionsassays gefundenen Relevanz für Tumorprozesse,

dem Effektorgen-Profilierung auf Basis prozess-relevanter Gene, den Detailexperimenten mit priorisierten Effektorgenen und letztlich jenen Interaktionen, die in der Forschungsgemeinschaft bereits bekannt sind. Die *in silico* Modellierung ist jedoch mehr, als lediglich der Versuch zur Integration der Datenflut. Computersimulierte Modellvorstellungen sollen es uns ermöglichen, die Diversität möglicher Interaktionen systematisch durchzuspielen. Experimente, die virtuell nur Sekunden einnehmen, würden im 'Nasslabor' viele Monate Forschungsarbeit fordern.

Ausblick

Die Modellierung von Netzwerken der Genexpression wird in dem gemeinsamen Projekt der SMPs RNA und Cell im NGFN ein umfassendes Bild komplexer biologischer Vorgänge entstehen lassen. Die Analyse bisher unerkannter Geninteraktionen wird es uns ermöglichen, Schnittstellen zu Tumor-relevanten Prozessen zu identifizieren und somit neue Ansatzpunkte für Diagnose und Therapie von Tumorerkrankungen liefern.

Kontakt

PD Dr. Holger Sültmann

PD Dr. Stefan Wiemann

Deutsches Krebsforschungszentrum – DKFZ
Abt. Molekulare Genomanalyse, TP 3
Heidelberg

E-Mail: h.sueltmann@dkfz.de

Literatur

- Arit, D. et al., 2005. Functional profiling discovers 16 cancer relevant cell cycle modulators. *Eingereicht*.
- Bunes, A. et al., 2005. arrayMagic: two-colour cDNA microarray quality control and preprocessing. *Bioinformatics* 21(4): 554-556.
- Gentleman, R.C. et al. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 5: R80. Epub 2004 Sep 2015.
- Sültmann, H. and A. Poustka. 2004. Recent advances in transcription profiling in human cancer. *Curr. Opin. Mol. Therap.* 6(6), 593-399.
- Sültmann, H., A. von Heydebreck, W. Huber, R. Kuner, A. Buneß, M. Vogt, B. Gunawan, M. Vingron, L. Füzesi, and A. Poustka. 2005. Gene expression in kidney cancer is associated with cytogenetic abnormalities, metastasis formation, and patient survival. *Clinical Cancer Research* 11(2), 646-655.
- Wiemann, S., D.H. Arit, W. Huber, R. Wellenreuther, S. Schleege, A. Mehrle, S. Bechtel, M. Saueremann, U. Korf, R. Pepperkok, H. Sültmann, and A. Poustka. 2004. From ORFeome to biology: a functional genomics pipeline. *Genome Res* 14: 2136-2144.

Glossar

Arrays: Vorlage, welche benutzt wird, um biologisches Material wie z. B. Nukleinsäuren, Peptide oder Proteine in hoher Dichte aufzutragen und zu binden. Arrays bestehen aus oder beinhalten eine feste Trägersubstanz die es erlaubt, große Mengen an Proben in geordneten Abständen aufzubringen, so dass diese für parallele Experimente genutzt werden können.

Assay: Testsystem, hier für die Untersuchung von zellulären Vorgängen wie z.B. des Zellwachstums und des programmierten Zelltodes (Apoptose)

Expression: Auch Genexpression genannt. Die Umsetzung der genetischen Information von der DNA über RNA in Proteine. Man unterscheidet das Kopieren der DNA in RNA (Transkription) und das Übersetzen des genetischen Codes an den Ribosomen in fertige Proteine (Translation).

PCR: Methode, mit der DNA-Abschnitte im Reagenzgefäß vervielfacht werden (Amplifikation). Zu den speziellen Formen der PCR gehört die rT-PCR, in deren Verlauf zunächst aus der Zelle isolierte mRNA (s.u.) mit Hilfe eines Enzyms, der reversen Transkriptase, in cDNA umgeschrieben wird, die dann als Ausgangspunkt für die DNA-Amplifikation dient.

RNA: Einzelsträngige Nukleinsäuremoleküle. Es gibt verschiedene Typen von RNA in einer Zelle, eine der wichtigsten ist die Boten- oder messenger RNA (mRNA), die Vorlage zur Proteinsynthese. Doppelstrang-RNA ist typisch für bestimmte RNA-Viren. Seit langem ist bekannt, dass Doppelstrang-RNA das menschliche Immunsystem zur Bildung bestimmter Moleküle anregt und antivirale Mechanismen, die die Entstehung von Proteinen unspezifisch hemmen auslöst.

RNA-Interferenz-Technik (RNAi) Eingesetzte doppelsträngige RNA-Stücke, auch siRNA genannt (small interfering RNA), führen zur Zerstörung von mRNA, die zu einem der beiden Stränge komplementär ist. Die siRNA, die dabei eine optimale Länge von 21 Nukleotiden besitzen sollte, aktiviert unter Komplexbildung spezifische Proteine der Zelle, die dann die mRNA an entsprechender Stelle spalten. Die Genaktivität wird damit vollständig eingestellt.

Tags: Markierungen die an einen Zellbaustein angeftet werden um ihn in der Zelle sichtbar zu machen. Beispiele sind das grün oder gelb fluoreszierende Protein (GFP und YFP, Y für yellow).