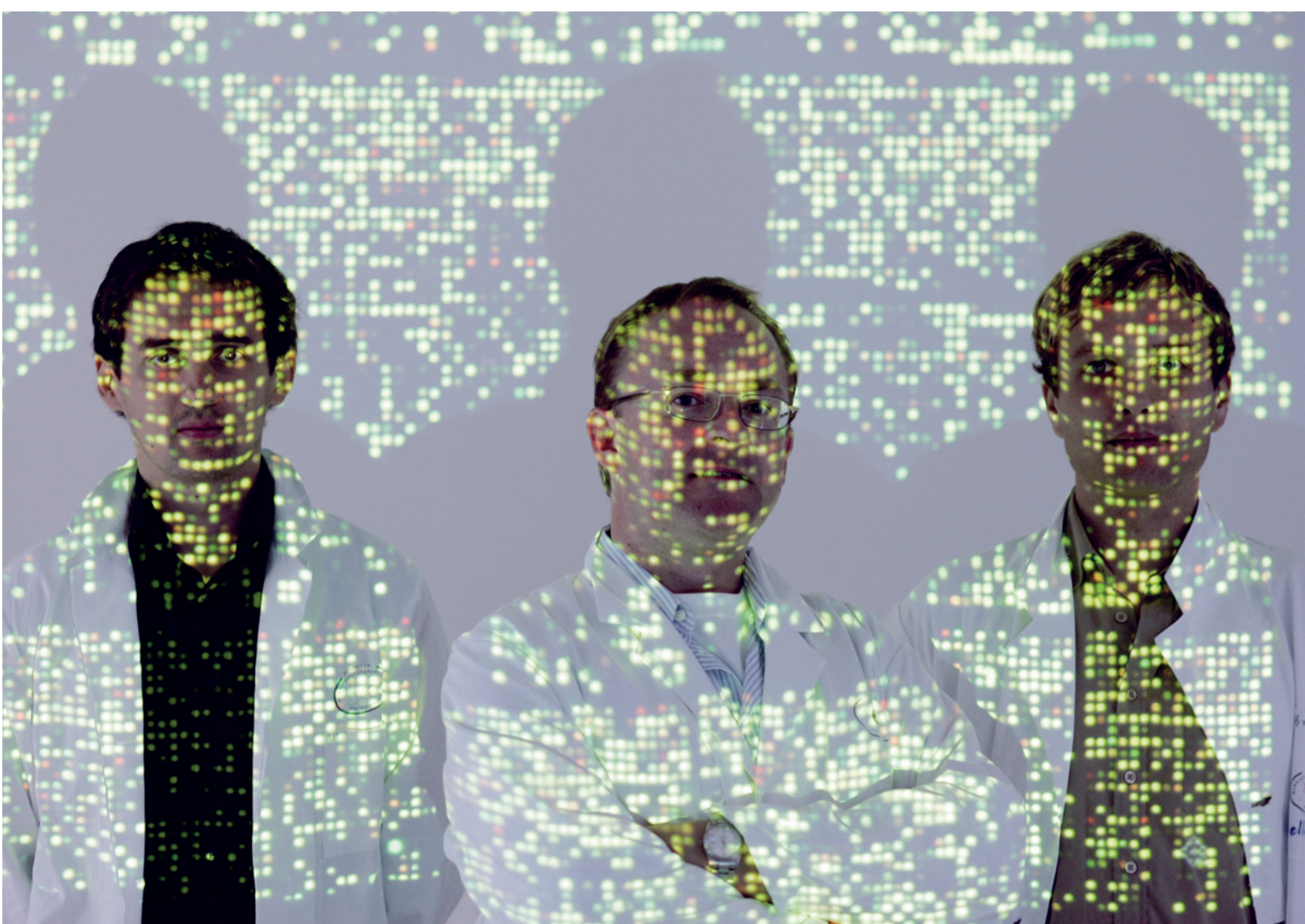
GENOMXPRESS
SONDERAUSGABE

NGFN
Nationales Genomforschungsnetz

Nationales Genomforschungsnetz



Das menschliche Genom – viel komplexer als vermutet

Ressourcen und Anwendungen für die Hochdurchsatz-Funktionsanalyse von Krankheitsgenen

Stefan Wiemann, Ingo Schupp, Stephanie Bechtel, Dorit Arlt, Özgür Sahin, Meher Majety, Ruth Wellenreuther, Heiko Rosenfelder, Jitao Zhang, Christian Schmidt, Alexander Mehrle, Florian Hahne, Annemarie Poustka, Rainer Pepperkok, Philippe Bastiaens, Helmut Blöcker, Michael Böcher, Karl Köhrer, Günther Michel, André Bahr, Jürgen Lauber, Birgit Ottenwälder, Dagmar Heubner, Werner Mewes, Andreas Hörlein

Bis vor wenigen Jahren ist man davon ausgegangen, dass Gene wie auf einer Perlenschnur aufgereiht auf unseren Chromosomen vorliegen und dass jedes Gen für ein spezifisches Produkt – also entweder ein Protein oder ein RNA-Molekül – kodiert. Von diesen und vielen weiteren Vorstellungen musste man sich aufgrund neuer Erkenntnisse verabschieden. Unser Genom ist wesentlich komplexer als lange vermutet wurde, und diese Komplexität setzt sich auch in der Regulierung der Gene und ihren Aktivitäten fort. Im Hinblick auf die relativ geringe Zahl von Genen (ca. 30 000) in jeder unserer Zellen – mit der wir uns übrigens kaum von dem winzigen Fadenwurm *C. elegans* unterscheiden – wurde inzwischen klar, dass der Regulierung von Genen und deren Funktionen eine zentrale Bedeutung für die Erhaltung von Gesundheit oder die Entwicklung von Krankheiten zukommt. Die Wissenschaftler im Verbundprojekt SMP Cell haben es sich daher zur Aufgabe gemacht, sowohl auf molekularer als auch auf zellulärer Ebene die Funktionen von Genen und Genprodukten zu erforschen sowie deren Fehlfunktionen im Krankheitsgeschehen zu entschlüsseln. Die Experimente werden überwiegend im Hochdurchsatz durchgeführt. Über die Ziele und die Struktur der SMP Cell wird ausführlich auf der Webseite www.smp-cell.de informiert. Anhand von drei Schwerpunkten soll hier der Beitrag der SMP Cell zur krankheitsorientierten Genomforschung im NGFN exemplarisch verdeutlicht werden.

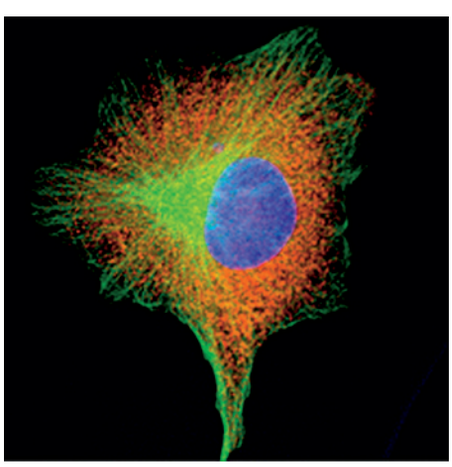


Abb.1: Eine Zelle im Fluoreszenzmikroskop, bei der einzelne Bestandteile unterschiedlich gefärbt wurden, um in einem Funktionsassay die hervorgerufenen Veränderungen besser erkennen zu können.

Auf Umwegen zu den Genen

Die meisten Gene des Menschen liegen auf den Chromosomen nicht in zusammenhängenden Blöcken oder „Sätzen" vor, sondern sind von scheinbar unwichtigen Abschnitten unterbrochen. Dies macht es zum einen schwierig, die Gene überhaupt zu entdecken, und zum anderen ist es unmöglich, diese Gene im Labor direkt für die Herstellung von Proteinen oder anderen Genprodukten zu verwenden. Stattdessen muss ein Umweg über sogenannte cDNAs gegangen werden. Solche cDNAs werden aus einem Zwischenprodukt der Zellen, der mRNA, hergestellt. Sie können in Bakterien vermehrt und anschließend in vielen verschiedenen Experimenten weiter verwendet werden. Aus der Sequenz, also der Abfolge der Basenbausteine, kann abgeleitet werden, welches Protein von der jeweiligen cDNA kodiert wird, und dieses Protein kann schließlich auch gentechnisch produziert werden. Einige dieser sogenannten rekombinanten Proteine werden inzwischen routinemäßig bei der Arzneimittelherstellung eingesetzt (z. B. menschliches Insulin), da die Funktion dieser Proteine sowie deren veränderte Aktivität oder Fehlfunktion in Krankheiten (in dem Beispiel der Insulinabhängigen Diabetes Typ 1) bekannt sind.

Für die meisten Proteine des Menschen ist die Funktion hingegen nicht geklärt. Viele der Gene sind noch nicht als cDNAs vorhanden und können daher auch nicht effizient auf ihre Funktion untersucht werden. Wissenschaftler der SMP Cell bilden das einzige europäische Konsortium, das in großem Maßstab menschliche Gene in Form von cDNAs kloniert, also eine Genbibliothek erstellt, aus der Proteine künstlich hergestellt und auf ihre Funktion untersucht werden können (siehe auch den Abschnitt: Zelluläre Funktionsassays). Mehrere tausend cDNAs wurden inzwischen von diesem Konsortium produziert und werden anderen Projekten des NGFN zur Verfügung gestellt. Mit dieser Ressource bringt sich die SMP Cell auch in ein internationales Projekt ein, in dem alle weltweit führenden Genom-Projekte gemeinsam an einer vollständigen „Bibliothek" der menschlichen Gene arbeiten, um diese für die Wissenschaft und Industrie als Basis für die systematische Funktionsanalyse verfügbar zu machen. Die Bibliothek dieser ORFeome Collaboration (www.orfeomecollaboration.org) ist bereits jetzt die größte frei zugängliche Ressource menschlicher Gene weltweit und bildet damit eine entscheidende Voraussetzung zur systematischen zellulären Genfunktionsanalyse. Bis zum Frühjahr 2008 sollen die meisten menschlichen Gene in diese Bibliothek aufgenommen werden, wozu die SMP Cell einen wichtigen Beitrag liefert.

Kompliziertes Zusammenspiel

Die erstaunlich geringe Zahl an Genen lässt vermuten, dass das Zusammenspiel von Genen und Genprodukten komplizierter sein muss als ursprünglich angenommen wurde. Deshalb müssen alle Aktivitäten innerhalb unserer Zellen streng kontrolliert werden, um den Ausbruch von Krankheiten zu vermeiden. Das richtige Verhältnis zwischen Wachstum und Zelltod ist z. B. entscheidend für die Vermeidung von Krebs, der meist durch Störungen gerade dieser zentralen Prozesse in einer oder wenigen unserer vielen Milliarden (genauer: ca. 10^{14}) Zellen eingeleitet wird.

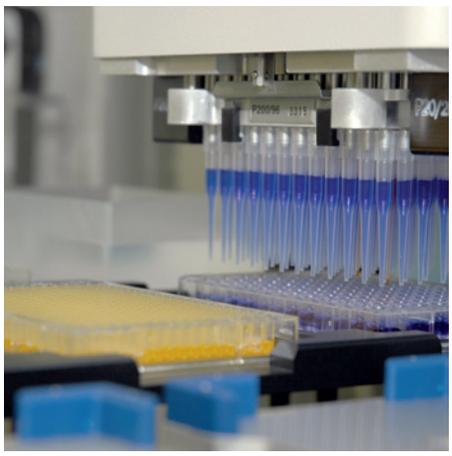


Abb.2: Detail eines Roboters für Hochdurchsatz-Funktionsassays.

Bis vor kurzem wurde davon ausgegangen, dass Proteine in linearen Prozessen arbeiten, vergleichbar einer Produktionsstraße mit einem Anfangs- und einem Endpunkt. Dieser Blick auf unsere Proteine und deren Wirkungen hat sich jedoch gewandelt. Proteine arbeiten und funktionieren in Netzwerken, bei denen Einflüsse an verschiedenen Stellen unterschiedliche Wirkungen hervorrufen können. Insbesondere für das Verständnis der molekularen Ursachen von Krankheiten ist es daher wichtig, die zugrundeliegenden Proteinnetzwerke zu analysieren.

In der SMP Cell arbeiten Wissenschaftler verschiedener Fachrichtungen zusammen, um in Funktionsassays zu untersuchen, wie Zellen auf Störungen normaler Abläufe reagieren. Daraus können die Forscher schlussfolgern, welche Funktion die untersuchten Proteine haben. In diesen Funktionsassays werden einzelne Proteine in Zellkulturen entweder über das normale Maß hinaus produziert (Überexpression der Proteine mithilfe der Reagenzien aus der oben beschriebenen Genbibliothek), oder die in diesen Kulturen natürlich vorkommende Menge an Protein wird vermindert. Für letzteren Ansatz wird eine Technologie eingesetzt, für deren Entdeckung Craig Mello und Andrew Fire den Nobelpreis erhielten: Die sogenannte RNA Interferenz. Bei dieser Technologie werden einzelne Gene gezielt ausgeschaltet, wodurch die Herstellung der kodierten Proteine vermindert wird. Diese beiden experimentellen Ansätze sind komplementär, in Funktionsassays kann folglich sowohl ein Zuviel als auch ein Zuwenig an Proteinen erreicht und im Anschluss die Konsequenz für die Zellen untersucht werden.

In der SMP Cell wurden solche Funktionsassays für den Hochdurchsatz etabliert. Insbesondere krebsrelevante Prozesse, wie z. B. das Wachstum der Zellen und die Verbreitung von Krebszellen während der Metastasen-Bildung, werden mithilfe dieser Hochdurchsatztechnologien untersucht. Ziel ist es, neue Ansätze für die Diagnose und Therapie von Krebserkrankungen – vor allem von Brustkrebs – zu entdecken. Weitere Funktionsassays wurden entwickelt, um die Reifung von neuronalen Zellen zu untersuchen, und um Proteine zu identifizieren, die den Cholesterinstoffwechsel beeinflussen. In diesen Ansätzen werden die Ergebnisse mit bekannten Informationen verknüpft und dadurch neue Erkenntnisse gewonnen und veröffentlicht. Die identifizierten Gene befinden sich derzeit in einem als Validierung bezeichneten Prozess: Mithilfe von Patientenproben soll untermauert werden, dass diese Gene ursächlich mit den jeweiligen Krankheiten verknüpft sind. Außerdem wird überprüft, ob diese Gene als diagnostische Marker und eventuell als Ansatzpunkte für neue Therapien verwendet werden können.

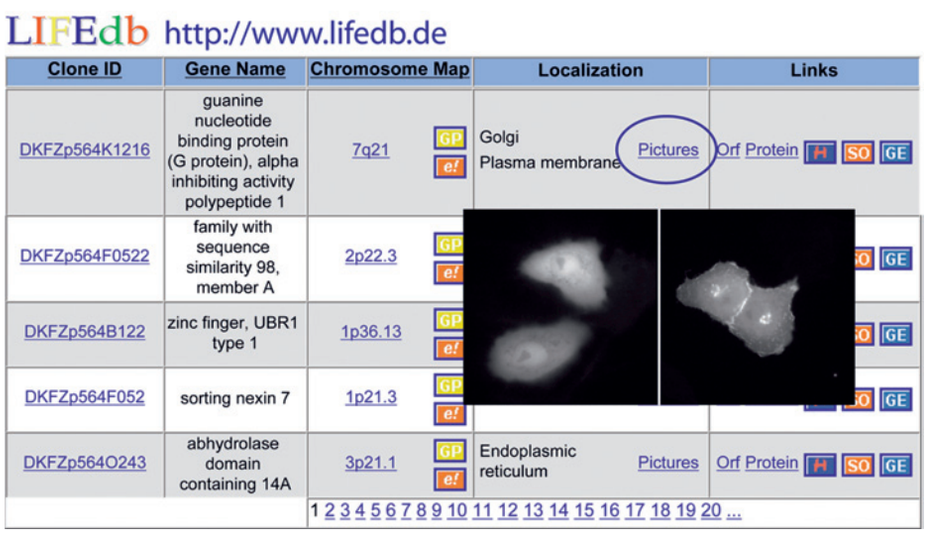
Hochdurchsatz: Daten, Daten, Daten ...

Schon der Begriff „Hochdurchsatz" zeigt an, dass in diesem Forschungszweig eine große Menge an Daten gesammelt wird. Dies ist zum einen auf die große Zahl unserer Gene und Proteine zurückzuführen, zum anderen auf die Vielzahl von experimentellen Ansätzen, mit denen die Funktion und Krankheitsrelevanz dieser Moleküle entdeckt werden soll. Die Experimente sind sehr aufwändig und werden zum großen Teil von Robotern durchgeführt. Wissenschaftler der SMP Cell entwickeln Software zur Steuerung dieser Roboter sowie zur Verfolgung einzelner Proben über die verschiedenen Prozesse hinweg: ein sogenanntes Labor-Informations-Management-System (LIMS).

Zudem kann ein „whole-genome Screen", also die Analyse aller menschlichen Gene, abhängig vom jeweiligen Experiment mehrere Terabyte an Daten liefern. Diese müssen zunächst gefiltert werden, um Signale vom Rauschen zu trennen. Im Anschluss müssen diese Daten mit statistischen Methoden analysiert werden, um positive von negativen Signalen zu unterscheiden. In der SMP Cell arbeiten Informatiker, Bioinformatiker und Biologen gemeinsam daran, LIMS Systeme und statistische Algorithmen zu erstellen und anzuwenden. Spezifische Benutzeroberflächen machen die Daten schließlich derart sichtbar, dass sie mit der Vielzahl bereits im Internet vorhandener Daten zu denselben Molekülen verknüpft und so die relevanten Informationen schnell erkannt werden können.

Was ist ein Gen?

Mit der Entschlüsselung des genetischen Codes, des sogenannten Buch des Lebens, erhoffte man sich umfassende Erkenntnisse über die Ursachen von Gesundheit und Krankheit. Tatsächlich ist das menschliche Genom inzwischen fast vollständig sequenziert, die Abfolge der einzelnen Bausteine also bekannt. Anstatt aber die erwarteten Antworten zu erhalten, haben sich neue Fragen und Erkenntnisse ergeben, die eine unerwartete Komplexität und



LIFEdb http://www.lifedb.de

Clone ID	Gene Name	Chromosome Map	Localization	Links
DKFZp564K1216	guanine nucleotide binding protein (G protein), alpha inhibiting activity polypeptide 1	7q21 GP e!	Golgi Plasma membrane Pictures	Orf Protein H SO GE
DKFZp564F0522	family with sequence similarity 98, member A	2p22.3 GP e!		GE
DKFZp564B122	zinc finger, UBR1 type 1	1p36.13 GP e!		GE
DKFZp564F052	sorting nexin 7	1p21.3 GP e!		GE
DKFZp564O243	abhydrolase domain containing 14A	3p21.1 GP e!	Endoplasmic reticulum Pictures	Orf Protein H SO GE

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 ...

Abb.3: Die LIFEdb Datenbank dient der Verknüpfung experimenteller Daten aus der SMP Cell mit allgemein zugänglichen Informationen und zur Veröffentlichung neuer Erkenntnisse im Internet.

Dynamik unseres Genoms erkennen lassen. Im Internet-Lexikon Wikipedia wird der Begriff Gen noch definiert als „ein Abschnitt auf der Desoxyribonukleinsäure (DNA), der die Grundinformationen zur Herstellung einer biologisch aktiven Ribonukleinsäure (RNA) enthält" (August 2007). Die Erkenntnisse aus dem Verbundprojekt SMP Cell haben aber gezeigt, dass diese wie auch alle anderen bisher gemachten Definitionen für den Begriff Gen so nicht stimmen.

Aus einem Gen können nämlich auch verschiedene RNAs hergestellt werden, die völlig unterschiedliche Produkte und Funktionen haben. Außerdem werden ständig neue Gene und sogar ganze Genklassen entdeckt. Sowohl in der Grundlagenforschung als auch in der krankheitsrelevanten Forschung ist daher noch viel Arbeit nötig, um die Zusammenhänge und das Funktionieren unserer Erbinformation sowie deren Zusammenspiel mit Umweltfaktoren verstehen zu können. Die Projekte der SMP Cell sind unmittelbar und auf mehreren Ebenen in diese Forschung involviert. Sie tragen damit wesentlich zur Erforschung der Lebensgrundlagen und Krankheitsursachen bei, wodurch Voraussetzungen geschaffen werden, um neue Strategien zur Vorbeugung und Behandlung dieser Krankheiten zu entwickeln. Mit seinen Ressourcen, Technologien und krankheitsrelevanten Applikationen ist das Verbundprojekt SMP Cell auch für die zukünftige Forschung aufgestellt und wird weiterhin auf internationalem Niveau wichtige Beiträge zum Verständnis von Krankheiten liefern.

Kontakt

Dr. Stefan Wiemann
Abteilung Molekulare Genomanalyse
Deutsches Krebsforschungszentrum (DKFZ)
in der Helmholtz-Gemeinschaft
Im Neuenheimer Feld 580, 69120 Heidelberg
E-Mail: s.wiemann@dkfz.de

gefördert durch:

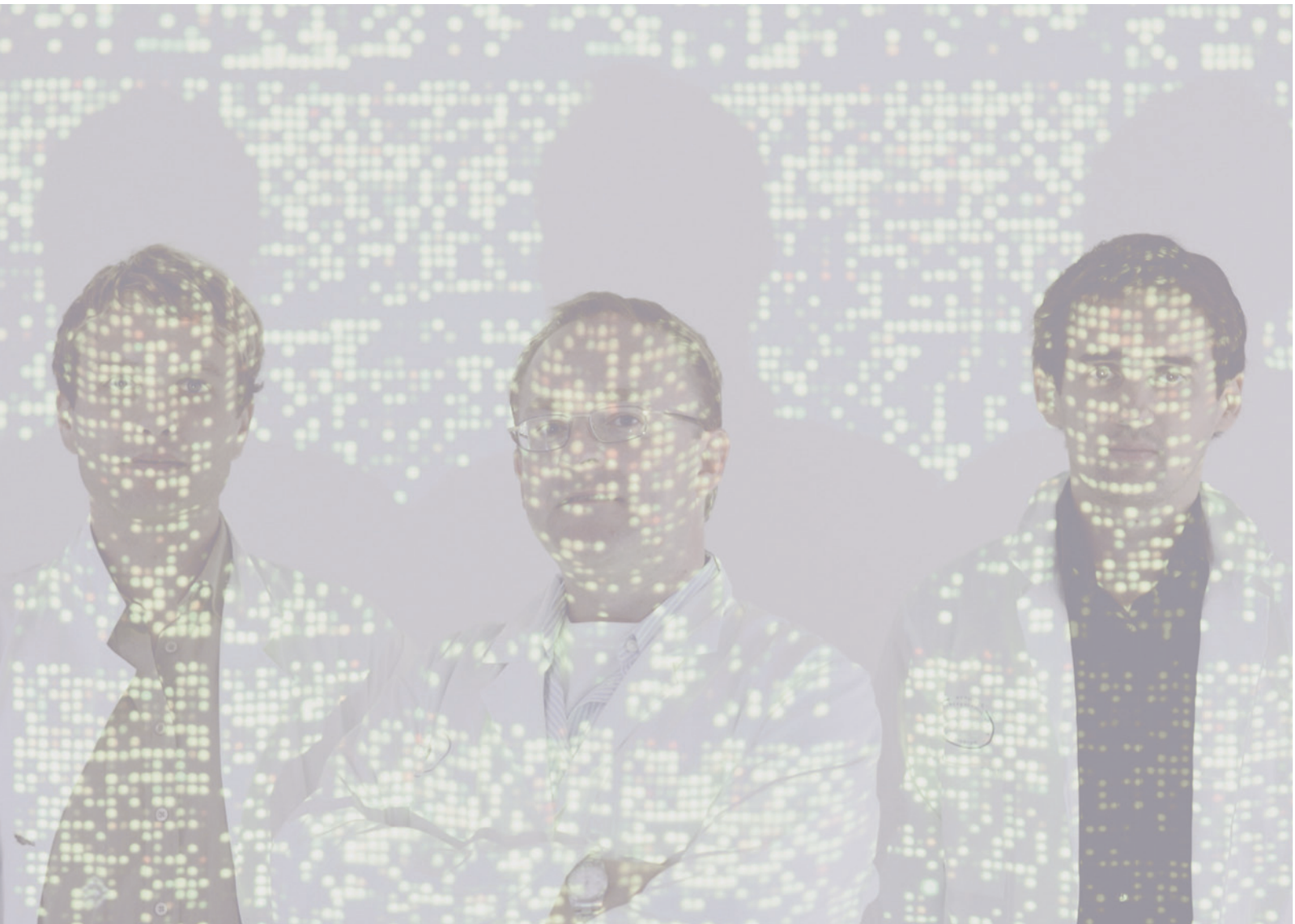
Bundesministerium
für Bildung
und Forschung

NGFN
Nationales
Genomforschungsnetz

Impressum
GenomXPress Sonderausgabe · Dezember 2007

Herausgeber
Das Projektkomitee des
Nationalen Genomforschungsnetzes (NGFN)

Redaktion
Dr. Susanne Dell · Helga Frankenstein
Dr. Olaf Krüger · Britta Paul · Andrea Suhr
Projektmanagement NGFN
Heinrich-Konen-Straße 1 · 53227 Bonn
Tel 0228-3821-0

Redaktionelle Unterstützung
Der Inhalt von namentlich gezeichneten Artikeln
liegt in Verantwortung des jeweiligen Autors.

ISSN 1617-562X
Dieser Newsletter wird aus Mitteln des BMBF gefördert.
Layout & Satz: Dirk Biermann, biermann@potsdam.de
Druck: GS Druck und Medien GmbH Potsdam