

Zentrum für Bioinformatik

Bericht 2010 – 2015

Zentrum für Bioinformatik

Bericht 2010 – 2015

1	Präambel	6
2	Überblick	9
	Historie	10
	Struktur	12
	Forschungsschwerpunkte	14
	Lehre	15
	Impact	16
3	Forschungsgebiete	21
	Bioinformatische Methoden	22
	<i>Analyse von genomischer Variation</i>	22
	<i>Statistisches Lernen in der Bioinformatik</i>	28
	<i>Biomolekulare Interaktionen</i>	34
	<i>Biologische Netzwerke</i>	40
	<i>Epigenetik</i>	48
	Anwendungen der Bioinformatik	54
	<i>Infektionskrankheiten</i>	54
	<i>Krebs</i>	62
	<i>Weitere medizinische Anwendungen</i>	68
	<i>Biotechnologische Anwendungen</i>	74
4	Aktuelles und Zahlen	79
	Aussagen von Absolventen	80
	Nachgefragt bei.....	91
	Aktivitäten des Zentrums	96
	Preise und Auszeichnungen	103
	Kolloquien und Vorträge	108
	Drittmittel in Zahlen	114
	Patente und Erfindungen	115
	Kooperationen	116
	Graduierungen	117
	Rufe auf Professuren	121
5	Lehre	123
	Studienprogramm	124
	Ziele und Inhalte der Bioinformatik-Studiengänge	125
	Berufsfelder für Bioinformatiker	126
	Informationen für Studierende und Bewerber	127
6	Ausgewählte Publikationen	133
7	Wege & Anfahrt	143
8	Impressum	144

1 Präambel

6

Das Zentrum für Bioinformatik Saar (ZBI) betreibt Forschung und Lehre in einem hochmodernen Bereich der Naturwissenschaften und ist eine Vorbildinstitution für interdisziplinäre Forschung auf dem Campus der Universität des Saarlandes (UdS).

In diesem Bericht wollen wir das Zentrum einer breiteren Öffentlichkeit vorstellen. Wir wollen damit allen an der Wissenschaft Interessierten die spannenden Themen und Methoden der Bioinformatik sowie die Arbeiten unseres Zentrums nahebringen. Wir laden Sie, liebe Leser und Leserinnen, in die faszinierende Welt dieses hoch aktuellen Wissenschaftsgebietes ein.

Die Bioinformatik ist eine interdisziplinäre Wissenschaftsdisziplin mit Anteilen der Fächer Biologie und Informatik, aber auch Medizin, Chemie, Pharmazie, Physik und Mathematik. Ziel der Bioinformatik ist der Einsatz von Computern zur Konfiguration von hochtechnologischen Experimenten in den Lebenswissenschaften sowie (vor allem) zur Analyse der dabei generierten umfangreichen und komplexen Datensätze.

Mit dem ersten Entwurf einer Sequenz des menschlichen Genoms, der vor 15 Jahren vorlag, war ein Meilenstein in der Biologie erreicht. Sie wandelte sich von einer in vielen Bereichen vorwiegend beschreibenden qualitativen zu einer vornehmlich quantitativen Wissenschaft. Aus der Genomsequenz auf die biologischen Prozesse im menschlichen Organismus zu schließen, hat sich allerdings als deutlich schwerer herausgestellt, als zunächst angenommen.

Die evolutiv entstandenen lebenden Organismen sind hochkomplex und nicht darauf ausgelegt, leicht verstanden zu werden. Hochdimensionale genetische und molekularbiologische Daten, die ein tiefergehendes Verständnis der komplexen biologischen Prozesse ermöglichen, können nur noch mit Hilfe des Computers analysiert werden. Daher ist die Bioinformatik aus den Lebenswissenschaften und der Biotechnologie heute nicht mehr wegzudenken.

Im Zentrum für Bioinformatik Saar verfolgt ein interdisziplinäres Team von Wissenschaftlern und Wissenschaftlerinnen gemeinsam folgende Ziele:

1. Exzellente Forschung im Bereich der Bioinformatik, sowohl die computergestützte Methodik betreffend als auch die Anwendungen in Biologie, Pharmazie und Medizin – mit direkter Auswirkung auf die Mehrung des Wissens in diesen Bereichen.
2. Publikation der Forschungsergebnisse in hochrangigen internationalen Fachzeitschriften.
3. Entwicklung und Bereitstellung von frei verfügbarer Software zur praxisnahen Anwendung in den Lebenswissenschaften.
4. Translation zentraler Forschungsergebnisse entweder gemeinsam mit bestehenden Firmen oder als Neugründungen, um die Gesundheit von Patienten zu fördern.
5. Ausbildung von hervorragenden jungen Bioinformatikern und Bioinformatikerinnen für Wissenschaft und Service in den Bereichen der Biotechnologie, Pharmazie und Medizin.
6. Schaffung eines Leuchtturms im Bereich der Bioinformatik und eines Vorbilds für fakultätsübergreifende interdisziplinäre Forschung am Standort Saarbrücken.
7. Stärkung der Bioinformatik im nationalen und internationalen Umfeld.

Die Gründung des Zentrums für Bioinformatik Saar jährt sich in diesem Jahr zum 15. Mal. So fallen die Gründung des ZBI im Jahr 2001, und die Ankündigung eines ersten Entwurfs der Sequenz des menschlichen Genoms, in eine Zeit, in der sich in Deutschland ein Bewusstsein für die Wichtigkeit der Bioinformatik gebildet hat. Die Gründung des Zentrums als eines von fünf Bioinformatikzentren in Deutschland war eine von der Deutschen Forschungsgemeinschaft (DFG) unterstützte Reaktion auf diese Entwicklung.

Nach dem Auslaufen der Förderung im Jahr 2007 hat die DFG in einer Abschlussbewertung dem Zentrum die größte Forschungsstärke unter allen deutschen Bioinformatikzentren attestiert. Dank der nachhaltigen Unterstützung durch die Universität des Saarlandes und die Max-Planck-Gesellschaft sowie durch diverse große Drittmittelprojekte und Industriekooperationen, die von Mitgliedern des Zentrums akquiriert wurden, konnte das ZBI seine Forschungsstärke erhalten und ausbauen. Heute weist das Zentrum bei einer Beteiligung von drei Fakultäten der Universität – Fakultät 2 (Medizin), Fakultät 6 (Mathematik/Informatik), und Fakultät 8 (Chemie, Pharmazie, Bio- und Werkstoffwissenschaften) ein ausgewogenes Fächerprofil auf.

Alle drei Fakultäten steuern je einen Lehrstuhl zum Zentrum bei. Diese Lehrstühle werden durch diverse Juniorprofessuren und Nachwuchsgruppen ergänzt. Ferner ist eine Abteilung des Max-Planck-Instituts für Informatik ausschließlich dem Thema Bioinformatik gewidmet. Weitere Partner sind das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI), das Fraunhofer-Institut für Biomedizinische Technik (IBMT) in Sankt Ingbert und das Helmholtz Zentrum für Pharmazeutische Forschung Saarland (HIPS).

Die intensive Arbeit zwischen den am ZBI beteiligten Forschungsgruppen bildet den Puls des Lebens im Zentrum. Besondere Forschungsleistungen hat das Zentrum zum Beispiel in den Bereichen Epigenetik, Krebsdiagnose, Analyse von viraler Wirkstoffresistenz sowie in der Entwicklung von Medikamenten, insbesondere von antibakteriellen Wirkstoffen, vorzuweisen.

Ein wesentlicher Aspekt des Zentrums ist darüber hinaus die Lehre. Das Zentrum bietet grundständige Lehre in dem interdisziplinären Bachelorstudiengang der Bioinformatik sowie einen weiterführenden Masterstudiengang an. In den 15 Jahren seiner Existenz hat das Zentrum 181 Abschlüsse zum Bachelor (70 Frauen) und 156 Abschlüsse zum Master (48 Frauen) verliehen. Der Masterstudiengang wird dabei in englischer Sprache angeboten. Seine Absolventen haben ein hochgradig internationales Profil: 48 Absolventen kommen aus dem Ausland.

Ein Großteil der Forschung im Zentrum wird jedoch von Doktoranden und Postdoktoranden durchgeführt. Seit der Gründung des Zentrums wurden 61 Doktoranden promoviert, 15 von ihnen Frauen. Zwei Drittel aller Absolventen bleiben im universitären (forschungsnahen) Umfeld und etwa ein Drittel wechselt in industrielle Bereiche. Ferner hat das Zentrum 16 Professoren und Professorinnen (siehe Seite 121) hervorgebracht, die an Universitäten im Inland und Ausland lehren und forschen. So hat das ZBI nicht nur durch seine forschersche Arbeit, sondern auch durch die hervorragend ausgebildeten jungen Wissenschaftler und Wissenschaftlerinnen, die es hervor gebracht hat, einen bleibenden Eindruck in der nationalen und internationalen Bioinformatik-Szene hinterlassen.

Dieser Bericht gliedert sich wie folgt: Nach einem kurzen einführenden Abschnitt, der das Zentrum prägnant in Gänze vorstellt, folgt die Präsentation der inhaltlichen Arbeit im Zentrum. Diese ist gegliedert in einen methodischen Teil und einen anwendungsorientierten Teil. Danach folgen Bemerkungen von Absolventen des Zentrums sowie von externen Kollegen und Kooperationspartnern. Eine Präsentation von Aktivitäten zur Öffentlichkeitsarbeit und von besonderen Ereignissen sowie ein tabellarischer Teil, der die Ergebnisse des Zentrums quantitativ erfasst, schließen daran an. Am Ende des Berichts geben wir Informationen für Leser und Leserinnen, die sich für ein Studium oder eine Forschungstätigkeit im Zentrum interessieren.

Wir wünschen Ihnen bei der Lektüre dieses Berichtes viel Freude.



Saarbrücken | März 2016

Prof. Dr. Dr. Thomas Lengauer
Sprecher des Zentrums für Bioinformatik Saar

2 Überblick

2 Überblick



Historie	10
Struktur	12
Forschungsschwerpunkte	14
Lehre	15
Impact	16

Historie

10

Die ersten Aktivitäten im Saarland auf dem Forschungsgebiet der Bioinformatik wurden von Prof. Dr. Kurt Mehlhorn, dem Gründungsdirektor des Max-Planck-Instituts für Informatik, initiiert. Er hatte im Jahr 1989 dem angehenden Doktoranden Hans-Peter Lenhof vorgeschlagen, sich ein spannendes Thema an der Schnittstelle zwischen der Informatik und der Chemie und Biologie zu suchen. Diese Anregung führte 1993 zur ersten Dissertation über bioinformatische Fragestellungen an der Universität des Saarlandes, anschließend Mitte der neunziger Jahre zu den ersten Lehrveranstaltungen in diesem Gebiet und gegen Ende der neunziger Jahre schließlich zur Einrichtung einer Nachwuchsgruppe am Max-Planck-Institut für Informatik.

Der erste Doktorand der Nachwuchsgruppe war Knut Reinert, der später als Postdoktorand in der Abteilung von Eugene Myers bei Celera Genomics in Bethesda, Maryland einen maßgeblichen Beitrag zur Entwicklung der dortigen innovativen Genomassemblierungspipeline im Rahmen der Sequenzierung des menschlichen Genoms geleistet hat.

Der Gründung des Zentrums für Bioinformatik im Jahr 2001 ging der Erfolg des Saarbrücker Konzepts und Projekts (MPII, UdS, DFKI) das „Virtuelle Biolabor“ bei der DFG-Initiative „Bioinformatik“ voraus. In diesem Wettbewerb hat sich das Saarbrücker Projekt mit vier anderen Konzepten unter Konzepten von 31 Regionen und Forschungsverbänden aus ganz Deutschland durchgesetzt. Folgende 15 Arbeitsgruppen des MPII, der UdS und des DFKI waren an diesem Antrag unter Federführung von Hans-Peter Lenhof beteiligt:

Prof. Dr. Rita Bernhardt	Biochemie, UdS
Prof. Dr. Friedrich Giffhorn	Angewandte Mikrobiologie, UdS
Prof. Dr. Rolf Hartmann	Pharmazeutische und Medizinische Chemie, UdS
Prof. Dr. Elmar Heinzle	Technische Biochemie, UdS
Prof. Dr. Jürgen Hüttermann	Biophysik, UdS
Dr. Joachim José	Pharmazeutische und Medizinische Chemie, UdS
Prof. Dr. Claus-Michael Lehr	Biopharmazie und Pharmazeutische Technologie, UdS
PD Dr. Hans-Peter Lenhof	Bioinformatikgruppe der AG1 des MPII
Prof. Dr. Alfred Louis	Angewandte Mathematik, UdS
Prof. Dr. Kurt Mehlhorn	AG1: Algorithmen und Komplexität, MPII
Prof. Dr. Eckart Meese	Humangenetik, UdS
Prof. Dr. Hans-Peter Seidel	AG4: Computergaphik, MPII
Prof. Dr. Raimund Seidel	Informatik, UdS
Prof. Dr. Michael Springborg	Physikalische Chemie, UdS
Prof. Dr. Gerhard Weikum	Datenbanken, UdS
Prof. Dr. Wolfgang Wahlster	Künstliche Intelligenz, DFKI & UdS

Das Zentrum wurde in den folgenden Jahren mit circa 6,2 Mio Euro von der DFG gefördert. Der Erfolg bei der DFG-Initiative „Bioinformatik“ hatte unter anderem sehr positive Effekte auf die Entwicklung der Lebenswissenschaften und vor allem der Pharmazie auf dem Saarbrücker Campus.

Fast zeitgleich mit der Gründung des Zentrums wurden an der UdS im Jahr 2000 der erste Lehrstuhl für Bioinformatik (Hans-Peter Lenhof) und die Geschäftsführung des Zentrums unter der Leitung von Frau Pia Scherer-Geiß eingerichtet. Ferner wurden drei Studiengänge für Bioinformatik geplant und im darauffolgenden Jahr eingerichtet, darunter auch der erste Bachelor- und Masterstudiengang der Universität des Saarlandes.

Ein Meilenstein der Entwicklung des ZBI war die Einrichtung der Bioinformatik-Arbeitsgruppe am Max-Planck-Institut für Informatik (MPII) im Jahr 2001. Dank gemeinsamer Anstrengungen der Max-Planck-Gesellschaft, des MPII, der Universität und der saarländischen Landesregierung gelang es Prof. Dr. Dr. Thomas Lengauer, einen der Gründungsväter der nationalen und internationalen Bioinformatik, als Direktor einer Arbeitsgruppe des MPII zu gewinnen und damit die internationale Sichtbarkeit des Zentrums für Bioinformatik erheblich zu stärken. Der Forschungsschwerpunkt des ZBI war zunächst die Entwicklung von bioinformatischen Methoden für die Pharmazie oder genauer für den Bereich Wirkstoffentwicklung. Durch die Arbeitsgruppe Lengauer wurde das Forschungsspektrum des ZBI um den Schwerpunkt bioinformatische Methoden und Software für biomedizinische Fragestellungen und hierbei insbesondere für Therapieoptimierung (Personalisierte Medizin) bereichert. Seit 2001 hat Thomas Lengauer auch die Funktion des Sprechers des Zentrums inne.

Im Jahr 2002 hat das ZBI die erste europäische Bioinformatik-Konferenz (ECCB) mit 450 Teilnehmern in Saarbrücken organisiert und damit eine neue internationale Konferenzserie initiiert, die heute zu den größten und bedeutendsten Bioinformatik-Konferenzen weltweit zählt.

Der nächste wichtige Ausbauschritt erfolgte im Jahre 2003, als Volkhard Helms den Ruf an die Universität des Saarlandes angenommen hat und in der Naturwissenschaftlich-Technischen Fakultät III (Chemie, Pharmazie, Bio- und Werkstoffwissenschaften) einen Lehrstuhl für „Computational Biology“ aufgebaut hat. Kurz danach übernahm Prof. Helms die Funktion des Studiendekans und hat sich in den folgenden Jahren mit großem Engagement dem Ausbau und der Verbesserung der Bioinformatik-Studiengänge gewidmet.

Unter anderem hat er die Akkreditierung der Bioinformatik-Studiengänge organisiert und geleitet. Gemeinsam mit den Studiengängen der Informatik und Computerlinguistik

wurden der Bioinformatik-Bachelor- und Masterstudiengang als erste Studiengänge der UdS 2005 akkreditiert. Der Lehrstuhl Helms verwendet Methoden des Molecular Modelling, um die Prinzipien von biomolekularen Interaktionen zu entschlüsseln, darüber hinaus erforscht er regulatorische Mechanismen zellulärer Netzwerke.

Die fünf durch die DFG geförderten Kompetenzzentren wurden nach Abschluss der Förderung im Jahr 2007 durch ein internationales Gutachtergremium evaluiert. Die Gutachter haben das ZBI nicht nur sehr positiv bewertet, sondern ihrer Einschätzung nach war es das forschungstärkste Zentrum der fünf geförderten Kompetenzzentren für Bioinformatik.

Ein weiterer Meilenstein in der Geschichte des ZBI war die Fertigstellung des Neubaus und der Umzug in den ZBI-Neubau im Jahr 2009. Die neuen lichtdurchfluteten Räumlichkeiten mit moderner technischer Ausstattung bieten eine sehr angenehme Arbeits- und Lernumgebung und werden daher auch von vielen externen Gruppen gern für Veranstaltungen genutzt.

Das Zentrum für Bioinformatik bildet seit seiner Gründung eine Brücke zwischen der wissenschaftlich herausragenden Saarbrücker Informatik und den Lebenswissenschaften an der UdS. Die Abteilung Lengauer und der Lehrstuhl Lenhof sind in der Naturwissenschaftlich-Technischen Fakultät I (Mathematik und Informatik), der Lehrstuhl Helms in der Naturwissenschaftlich-Technischen Fakultät III (Chemie, Pharmazie, Bio- und Werkstoffwissenschaften) angesiedelt. Im Jahr 2013 wurde ein weiterer Bioinformatik-Lehrstuhl mit Schwerpunkt „Klinische Bioinformatik“ in der Fakultät 2 (Medizin) eingerichtet. Durch die Arbeitsgruppe Keller hat das ZBI unter anderem an Expertise und internationaler Sichtbarkeit im hochaktuellen Forschungsbereich „MicroRNAs“ und hier besonders im Gebiet „MicroRNAs als molekulare Biomarker für neue leistungsfähige diagnostische Verfahren“ gewonnen.

Eine zentrale Rolle beim Aufbau des ZBI haben auch die Nachwuchsgruppenleiter Oliver Kohlbacher, Andreas Hildebrandt, Rainer Böckmann, Dirk Neumann, Mario Albrecht, Alice McHardy, Jan Baumbach, Marcel Schulz, Nico Pfeifer und Olga Kalinina gespielt. Sie haben sowohl in der Forschung als auch in der Lehre wichtige Beiträge geleistet. Besonderer Dank gilt Oliver Kohlbacher und Andreas Hildebrandt, die zudem wesentlich am Aufbau der Bioinformatik-IT-Infrastruktur des ZBI beteiligt waren. Des Weiteren hat Dirk Neumann gemeinsam mit zwei anderen Absolventen des ZBI den ersten Spin-Off (Scientific Consilience GmbH) des ZBI im Jahr 2010 gegründet.

Fünf Nachwuchsgruppenleiter erhielten zudem Rufe auf Bioinformatik-Professuren und haben heute Bioinformatik-Lehrstühle an nationalen und internationalen Universitäten

inne. Insgesamt zeigte sich der Erfolg in Forschung und Ausbildung auch in der Tatsache, dass das ZBI in relativ kurzer Zeit die beträchtliche Zahl von 16 Professoren hervorgebracht hat (siehe Seite 121), so dass an einer Reihe von deutschen und ausländischen Universitäten heute Kollegen in der Bioinformatik tätig sind, die in Saarbrücken ausgebildet wurden.

2014 gelang es dem ZBI trotz starker Konkurrenz von mehreren anderen Universitäten Tobias Marschall als ersten Juniorprofessor für Bioinformatik am Standort Saarbrücken zu gewinnen.

Das ZBI verbindet als wissenschaftliche Brücke die Informatik nicht nur mit den Lebenswissenschaften, sondern auch mit der Pharmazie, der Chemie und der Physik, was sowohl durch gemeinsame Publikationen als auch durch gemeinsame Projekte belegt ist. Das ZBI und seine Bioinformatik-Gruppen sind in viele Exzellenz- und Vorzeigeprojekte der UdS involviert, vom Exzellenzcluster (MMCI) und der Graduiertenschule der Informatik über den Sonderforschungsbereich 1027 (Physikalische Modellierung von Nicht-Gleichgewichtsprozessen in biologischen Systemen) bis hin zu den EU-Projekten P-Medicine, Blueprint und DEEP.

Struktur

12

Das Zentrum für Bioinformatik (ZBI) ist eine wissenschaftliche Einrichtung der Universität des Saarlandes. Die Organe des ZBI umfassen die Zentrumsleitung, den Zentrumsrat und den Studiendekan der Bioinformatik. Die Zentrumsleitung setzt sich aus je einem Vertreter der beteiligten Fakultäten zusammen. Prof. Dr. Dr. Thomas Lengauer, Direktor am Max-Planck-Institut für Informatik in Saarbrücken, ist der Zentrumsprecher. Die administrative Geschäftsführung übernimmt die hauptamtliche Geschäftsführerin Pia Scherer-Geiss. Der Zentrumsrat koordiniert die Organisation der Bioinformatik-Studiengänge und anderer gemeinsamer Aktivitäten. Mitglieder sind alle Professoren für Bioinformatik, die Geschäftsführung sowie Vertreter der drei Fakultäten, der Mitarbeiter und Mitarbeiterinnen und der Studierenden. Das Zentrum für Bioinformatik betreibt ein eigenes Prüfungsamt (geleitet von Karin Jostock) und wirkt federführend bei der Erstellung von Studien- und Prüfungsordnungen mit.

Die inhaltliche Struktur des Zentrums folgt einem Schalenmodell. Den Kern bilden die derzeit acht Bioinformatikgruppen, die jeweils zur Hälfte aus Lehrstühlen und Juniorgruppen bestehen. Die nächste Schale umfasst Gruppen aus den beteiligten Fächern Informatik, Biologie und Medizin, die sowohl an der Bioinformatik-Lehre als auch an disziplinübergreifenden Forschungsprojekten teilnehmen. Die nächste Schale umfasst Gruppen, die das Zentrum vornehmlich in der Lehre unterstützen. In der äußeren Schale befinden sich Forschungsgruppen mit forscherschem Kontakt zum Zentrum.

Seit dem Auslaufen der Initialförderung des Zentrums durch die DFG im Jahre 2007 finanziert sich das ZBI durch die Grundfinanzierung der beteiligten Fakultäten und Institute sowie durch eingeworbene Drittmittel (siehe Seite 108).

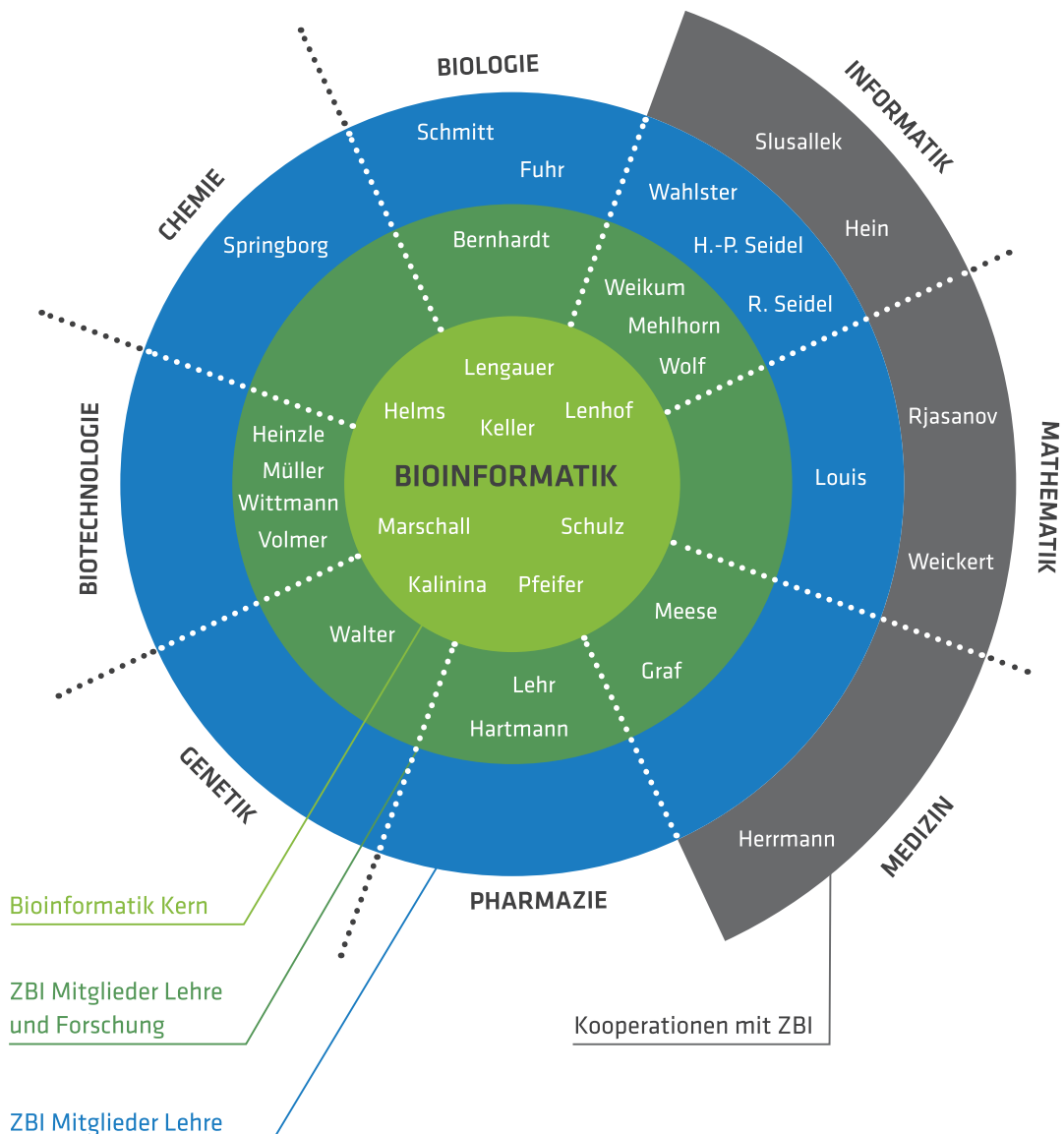


Abb.: Schalenmodell des Zentrums für Bioinformatik

Die im Schalenmodell genannten Partnergruppen des ZBI

Leitung	Forschungsgebiet	
Prof. Dr. Rita Bernhardt	Biochemie	UdS
Prof. Dr. Günter Fuhr	Biomedizinische Technik	IBMT
Prof. Dr. Norbert Graf	Pädiatrische Onkologie und Hämatologie	UdS
Prof. Dr. Rolf Hartmann	Pharmazeutische und Medizinische Chemie	HIPS
Prof. Dr. Matthias Hein	Maschinelles Lernen	UdS
Prof. Dr. Elmar Heinzle	Technische Biochemie	UdS
Prof. Dr. Volkhard Helms	Bioinformatik	UdS
Prof. Dr. Mathias Herrmann	Medizinische Mikrobiologie und Hygiene	UdS
Dr. Olga Kalinina	Structural Bioinformatics of Protein Interactions	MPII
Prof. Dr. Andreas Keller	Klinische Bioinformatik	UdS
Prof. Dr. Claus-Michael Lehr	Biopharmazie und Pharmazeutische Technologie	HIPS
Prof. Dr. Hans-Peter Lenhof	Bioinformatik	UdS
Prof. Dr. Dr. Thomas Lengauer	Bioinformatik und Angewandte Algorithmik	MPII
Prof. Dr. Alfred Louis	Angewandte Mathematik	UdS
Juniorprofessor Dr. Tobias Marschall	Algorithms for Computational Genomics	UdS/MPII
Prof. Dr. Eckart Meese	Humangenetik	UdS
Prof. Dr. Dr. h.c. Kurt Mehlhorn	Algorithmen und Komplexität	MPII
Prof. Dr. Rolf Müller	Pharmazeutische Biotechnologie	HIPS
Dr. Nico Pfeifer	Statistical Learning in Computational Biology	MPII
Prof. Dr. Sergej Rjasanow	Angewandte Mathematik	UdS
Prof. Dr. Manfred Schmitt	Molekular- und Zellbiologie	UdS
Dr. Marcel Schulz	High-throughput Genomics & Systems Biology	MMCI/MPII
Prof. Dr. Hans-Peter Seidel	Computergrafik	MPII
Prof. Dr. Raimund Seidel	Theoretische Informatik	UdS
Prof. Dr.-Ing. Philipp Slusallek	Computer Graphics Lab & Intel Visual Computing Institute	UdS
Prof. Dr. Michael Springborg	Physikalische Chemie	UdS
Prof. Dr. Dietrich A. Volmer	Institut für Bioanalytische Chemie	UdS
Prof. Dr. Dr. h.c. mult. Wolfgang Wahlster	Informatik	DFKI
Prof. Dr. Jörn Walter	Genetik	UdS
Prof. Dr. Joachim Weickert	Mathematische Bildanalyse	UdS
Prof. Dr. Gerhard Weikum	Datenbanken und Informationssysteme	MPII
Prof. Dr. Christoph Wittmann	Systembiotechnologie	UdS
Prof. Dr. Verena Wolf	Modeling and Simulation	UdS

Am ZBI beteiligte Forschungsgruppen

UdS *Universität des Saarlandes*

MPII *Max-Planck-Institut für Informatik*

DFKI *Deutsches Forschungszentrum für Künstliche Intelligenz*

IBMT *Fraunhofer Institut für Biomedizinische Technik*

HIPS *Helmholtz Zentrum für Pharmazeutische Forschung Saarland*

MMCI *Excellence Cluster Multimodal Computing and Interaction*

Forschungsschwerpunkte

14

Da wir jedem Forscher die Möglichkeit bieten, sich mit dem bioinformatischen Thema seiner Wahl auseinander zu setzen, ist das Zentrum für Bioinformatik Saar forschersich breit aufgestellt. Trotz der großen Breite der Forschung hat sich das ZBI ein konkretes wissenschaftliches Profil gegeben, das auf pharmazeutische und medizinische Themen ausgerichtet ist und das es von den anderen deutschen Bioinformatikzentren deutlich abgrenzt. Kurz gesagt, wir wollen durch Bioinformatik zum besseren Verständnis von Krankheiten beitragen, um sie besser und früher zu diagnostizieren und sie zielgerichteter zu behandeln.



Die obige Abbildung illustriert dieses Forschungsprofil. Unsere Beiträge zur Diagnose von Krankheiten zielen darauf ab, aus Omics-Daten von Patienten Informationen zu extrahieren, die auf die Art ihrer Krankheit (z. B. Subtyp eines bösartigen Tumors) schließen lassen. Unter Omics-Daten verstehen wir die gesamte Zelle abdeckende molekulare Daten betreffend das Genom (Genomics, Epigenomics), die in der Zelle abgelesenen Gene (Transcriptomics), die in der Zelle vorhandenen Proteine (Proteomics) und Metabolite (Metabolomics) sowie die molekularen Wechselwirkungen (Interactomics). Die Differentialdiagnose einer Krankheit ist im Allgemeinen eine wesentliche Vorbedingung für eine effektive Therapie. Hier geht es aber nicht so sehr darum, Mechanismen des Krankheitsprozesses zu verstehen, sondern vielmehr um eine genaue Bestimmung des Krankheitstyps. Diese ist oft auch ohne ein genaues Verständnis der Krankheitsprozesse möglich. Der Bereich „Zielmoleküle“ repräsentiert unsere Beiträge zur biomedizinischen Grundlagenforschung und speziell zur Aufklärung der Mechanismen von Krankheitsprozessen. Wir bedienen uns dazu wiederum der Omics-Daten.

Hier steht die Analyse von molekularen Netzen und deren Fehlfunktionen bei Krankheiten im Vordergrund. Auf der Grundlage eines solchen Verständnisses der molekularen Basis der Krankheit können dann auch Zielmoleküle identifiziert werden – Proteine, an die noch zu entwickelnde oder bereits existierende Wirkstoffmoleküle binden sollen, um den Krankheitsprozess im Sinne des Patienten zu moderieren.

Der Bereich Design repräsentiert den pharmazeutisch ausgerichteten Teil unserer Forschung. Hier gilt es, für ein identifiziertes Zielmolekül aus der übergroßen Vielzahl von möglichen chemischen Verbindungen erste Kandidaten zu selektieren, die vielversprechend für den weiteren Wirkstoffentwurf sind. Die Datenbasis für diese Forschung bilden große Datenbanken von pharmazeutischen Verbindungen. In diesem Bereich profitieren wir von den Kooperationen mit dem Fachbereich Pharmazie und dem auf unserem Campus beheimateten Helmholtz Institut für Pharmazeutische Forschung Saarland (HIPS).

Der Bereich Optimierung steht für die Verfeinerung des Wirkstoffs hinsichtlich pharmazeutischer Eigenschaften wie Spezifität, Bioverfügbarkeit und Verträglichkeit. Der Bereich Therapie schließlich steht für die Optimierung von Kombinationen verschiedener Medikamente auf die individuellen Bedürfnisse des Patienten. Dieser Bereich ist bei uns besonders stark ausgeprägt und stellt die translationale Komponente der Forschung des ZBI dar.

Neben diesen Schwerpunkten ist das ZBI aber auch noch in anderen Bereichen aktiv. Ein weiteres Forschungsfeld ist zum Beispiel der Bereich der Biotechnologie. Hier werden Mikroorganismen für die Produktion von gewünschten chemischen Verbindungen modifiziert oder für die Optimierung von Nahrungsmitteln hinsichtlich agrar-ökonomischer Eigenschaften eingesetzt.

Lehre

Die Universität des Saarlandes bietet seit dem Wintersemester 2001/2002 bzw. dem Wintersemester 2003/2004 einen sechs-semesterigen Bachelorstudiengang und einen vier-semesterigen forschungsorientierten Masterstudiengang Bioinformatik an. Beide Studiengänge werden von dem Zentrum für Bioinformatik organisiert. Das Bachelorstudium vermittelt ein breites Spektrum an Fachwissen und die für den Einstieg in die berufliche Praxis notwendigen Grundlagen. Das Masterstudium lehrt, aufbauend auf einem ersten berufsqualifizierenden Abschluss, tiefer gehendes Fachwissen, das auf eine anspruchsvolle Tätigkeit in der Industrie oder eine Promotion vorbereitet.

An der Lehre in den sehr interdisziplinären Bioinformatik-Studiengängen sind Lehrende dreier Fakultäten und mehrerer An-Institute beteiligt. Das Fächerangebot umfasst Mathematik, Informatik, Bioinformatik, Chemie, Biologie und Pharmazie. Für die eigentliche Bioinformatik-Lehre sind der Lehrstuhl Professor Helms (Biologie), der Lehrstuhl Professor Lenhof (Informatik), der Lehrstuhl Professor Keller (Medizin), Juniorprofessor Marschall (Informatik) sowie die Abteilung von Professor Lengauer am MPI für Informatik zuständig. An der Lehre im Kernbereich Bioinformatik sind derzeit weiterhin Prof. Verena Wolf (Informatik) und Dr. Marcel Schulz (Informatik) beteiligt.

Der Bachelorstudiengang wird in deutscher Sprache unterrichtet. Unsere Studierenden besitzen im Allgemeinen sowohl ein starkes Interesse an Informatik und Mathematik als auch ein ausgeprägtes Interesse an den Lebenswissenschaften. Etwa die Hälfte der Studienanfänger stammt aus dem Saarland, die andere Hälfte meist aus dem restlichen Bundesgebiet. Die Studierenden können ihr individuelles Studienprogramm gemäß einer eher informatischen oder einer eher biologischen Vertiefung gewichten. Der größte Teil der Absolventen unseres Bachelorstudiengangs bleibt danach in Saarbrücken und setzt sein Studium im Masterstudiengang fort. Ein kleiner Teil der Studierenden wechselt z.B. in den Masterstudiengang für Biotechnologie an der UdS.



Der Masterstudiengang wird in englischer Sprache unterrichtet und wendet sich sowohl an die Absolventen unseres Bachelorstudiengangs als auch an Absolventen von vergleichbaren Studiengängen im In- und Ausland. Unser Masterstudiengang genießt eine hohe internationale Reputation und wird deshalb stark nachgefragt, insbesondere unter Bewerbern aus dem asiatischen Raum. In den vergangenen Jahren war die Anzahl der aufgenommenen externen Bewerber vergleichbar mit denen, die den Bachelorstudiengang am ZBI abgeschlossen haben. Die Gruppe der Studierenden hat daher eine sehr internationale Prägung.

Die Absolventen des Masterstudiengangs sind national und international begehrte Kandidaten für Promotionsstellen an führenden Forschungsinstituten und Universitäten. Etwa zwei Drittel unserer Absolventen setzt daher ihre Ausbildung mit einem Promotionsvorhaben fort. Der andere Teil wechselt direkt in die Industrie oder in den Dienstleistungssektor, sowohl im Saarland als auch im restlichen Bundesgebiet.

Impact

16

Das Zentrum für Bioinformatik Saar hat wichtige wissenschaftliche Impulse auf dem Campus der Universität des Saarlandes sowie in der deutschen und internationalen Bioinformatik gesetzt. Um auch weiterhin erfolgreich zu bleiben, orientiert sich die Arbeit im ZBI stets an sieben Zielen, die in der Präambel für das Zentrum formuliert wurden.

Exzellente Forschung

Das Zentrum kann hervorragende Forschungsergebnisse in einer Vielzahl von Forschungsbereichen mit Bioinformatikanteil vorweisen. Die Forschungsschwerpunkte im Zentrum für Bioinformatik schließen ein:

Epigenetik

In diesem für die modernen Biowissenschaften zentralen Gebiet hat sich das Zentrum einen Schwerpunkt gesetzt. Das große Deutsche Epigenom Programm (DEEP), das von 2012 bis 2017 vom BMBF gefördert wird, leitet Prof. Jörn Walter, Inhaber des Lehrstuhls für Genetik an der Universität des Saarlandes. Prof. Thomas Lengauer vom Max-Planck-Institut für Informatik koordiniert die Bioinformatik-Aktivitäten dieses Projektes. Darüber hinaus haben sich im Umfeld von DEEP umfangreiche interdisziplinäre Forschungsarbeiten entwickelt, die weitere Lehrstühle einschließen (z.B. Prof. Helms, Prof. Wolf, Prof. Hein). In dem Projekt sind hochrangige Publikationen und wichtige Software entstanden, die weltweit von einer Vielzahl von Wissenschaftlern genutzt werden.

Biomarker für die Diagnose von Krankheiten

Die ZBI-Arbeitsgruppen Prof. Keller, Prof. Meese und Prof. Lenhof haben für eine Reihe von Erkrankungen, darunter viele Krebsarten, neue Biomarker identifiziert. Die Biomarker für Krebs setzen sich aus verschiedenen, in Tumoren auftretenden Antigenen bzw. aus verschiedenen kurzen RNA-Molekülen (MicroRNAs) zusammen. Da sich diese zusammengesetzten Biomarker, die man auch als molekulare Signaturen bezeichnet, im Blut von Patienten nachweisen lassen, können mit ihnen minimal- bzw. nicht-invasive diagnostische Tests entwickelt werden. Die diagnostischen Signaturen sind in mehr als 50 wissenschaftlichen Veröffentlichungen beschrieben und bilden die Grundlage mehrerer Patentanmeldungen bzw. Patente. Aufbauend auf diesen Ergebnissen sollen die Biomarker auf ihren Einsatz in der Klinik getestet werden.

Bakterielle Resistenz – Neue Wirkstoffe

Durch die Zusammenführung der Expertisen aus Bioinformatik, Biochemie, Pharmazie und Chemie konnten Arbeitsgruppen des Zentrums vielversprechende antibakterielle Naturstoffe isolieren und deren molekulare Ziele identifizieren. Hervorzuheben sind Arbeiten an einem völlig neuartigen Gyraseinhibitor und translationale Arbeiten zur Entwicklung von Griselimycinen als Wirkstoffe gegen Tuberkulose. Letztere zeigen hervorragende Aktivitäten in Tiermodellen und weisen durch ihren neuen Wirkmechanismus der DNA-Syntheseblockade keine Kreuzresistenz mit bekannten Antibiotika auf.

Bakterielle Resistenz – Resistenztests

Aufbauend auf der Erfahrung mit Biomarker-Panels arbeitet das ZBI daran, die Diagnose- und Therapieentscheidung bei Patienten mit bakteriellen Infektionskrankheiten zu verbessern. Ziel ist es, klassische kultur-basierte Resistenztests, wie sie heute gängig sind, durch genetische Testung zu ergänzen. Entsprechende Tests haben das Potenzial schneller zu relevanten Therapieentscheidungen zu führen und die Behandlung von Patienten dadurch zu verbessern.

Biotechnologie

Das Zentrum hat international führende Konzepte zu Design und Entwicklung mikrobieller Zellfabriken vorangetrieben, die die Basis für eine zukunftsweisende nachhaltige industrielle Produktion bilden. Neben hochrangigen Publikationen unterstreichen eine Reihe internationaler Patente und gemeinsam mit der chemischen und pharmazeutischen Industrie entwickelte Materialien, wie biobasierte Kunststoffe oder Plattformchemikalien wie Bio-Bernsteinsäure, den hohen Anwendungsgrad der hier einfließenden bioinformatischen Ansätze.

Biomolekulare Interaktionen

Dieses Thema bildet einen Schwerpunkt am ZBI, zu dem zahlreiche Arbeitsgruppen beitragen. Die Forschungen umfassen z.B. Analysen zum Bindeverhalten von Molekülen. Hier wurde mit einer neuartigen Analyse transient – also zwischenzeitlich – auftretender Bindetaschen eine neue Forschungsrichtung initiiert. Ferner gibt es Untersuchungen zur Bindung von Wirkstoffen an Proteinen, die für Medizin und Biotechnologie relevant sind und zur Modellierung großer aus Proteinen bestehender Komplexe. Auf höherer Ebene wurden Interaktions-Netzwerke mit zahlreichen beteiligten Molekülen analysiert, z.T. im Rahmen der Grundlagenforschung, z.T. vor dem Hintergrund medizinischer Fragestellungen.

Veröffentlichungen

Im Berichtszeitraum (2010–2015) hat das Zentrum für Bioinformatik über 500 Publikationen in internationalen Zeitschriften und begutachteten Konferenzen publiziert. Beispiele für hochrangige Zeitschriften sind sowohl interdisziplinär ausgerichtete Journale, die an ein breites Publikum adressiert sind (Zeitschriftenfamilie Nature, Science, Cell und PLOS, Proceedings of the National Academy of Sciences (PNAS)), als auch die führenden Bioinformatikzeitschriften (PLOS Computational Biology, Bioinformatics und BMC Bioinformatics) sowie führende Journals, die auf Genetik spezialisiert sind (Nucleic Acids Research, Genome Biology und Genome Research). Der translationale Aspekt der Forschung zeigt sich durch Veröffentlichungen in anwendungsnahen klinisch orientierten Journals, wie z. B. Gastroenterology oder Cancer Research. Eine reine Darstellung in führenden Zeitschriften ist jedoch für eine Bewertung des wissenschaftlichen Ertrages des Zentrums allein nicht ausreichend.

Viele hochwertige Publikationen erscheinen daher auch in spezifischen Fachzeitschriften. Einen Überblick über die wesentlichen Publikationen des Zentrums im Berichtszeitraum gibt die Publikationsliste auf Seite 133 ff.

Software

Am Zentrum für Bioinformatik sind vielfältige und viel genutzte Bioinformatik-Softwaresysteme entstanden. Führend sind zu nennen:

Der über das Internet frei verfügbare geno2pheno Server zur viralen Resistenzanalyse, der weltweit zur Behandlung von Patienten mit HIV, Hepatitis B und Hepatitis C Infektionen genutzt wird.

Die von internationalen Wissenschaftlern in der Epigenetik vielfach eingesetzten Softwaresysteme Rnbeads (Analyse von DNA Methylierungsdaten), BiQAnalyzer (Qualitätskontrolle von DNA Methylierungsdaten) und EpiExplorer (Durchforstung genomischer und epigenomischer Daten).

Der über das Internet frei verfügbare Server GeneTrail, der von vielen Gruppen weltweit für die Analyse (Gene Set Enrichment Analysis) von deregulierten biologischen Prozessen und Kategorien bei verschiedensten Erkrankungen eingesetzt wurde und wird.

Die C++-Softwarebibliothek BALL (Biochemical Algorithms Library), die in der Strukturbiologie zur Modellierung von Proteinen und anderen Makromolekülen und beim Wirkstoffdesign eingesetzt wird und die mit mehreren Softwarepreisen ausgezeichnet wurde.

FunSimMat: eine umfassende Datenbank von funktionellen Ähnlichkeiten für Proteine und Proteinfamilien.

Translation

In dem Bereich der viralen Resistenzanalyse hat das Zentrum durch die Arbeit in der Abteilung von Prof. Lengauer am Max-Planck-Institut für Informatik weltweit führende Forschungsergebnisse mit hohem translationalem Anteil vorzuweisen. Neben hochrangigen Veröffentlichungen ist in diesem Bereich auch Software entstanden, die direkt und breitflächig bei der Behandlung von HIV-Patienten Einsatz findet (siehe Seite 54).

Ausbildung

Gut ausgebildete Absolventen sind eines der wertvollsten Ergebnisse des Zentrums für Bioinformatik. In den 15 Jahren seiner Existenz hat das Zentrum 181 Abschlüsse zum Bachelor (70 Frauen) und 156 Abschlüsse zum Master (48 Frauen) verliehen. 143 der Bachelor-Absolventen haben in Saarbrücken auch ein Masterstudium angeschlossen. Da der Masterstudiengang auf Englisch angeboten wird, haben seine Absolventen ein hochgradig internationales Profil: 48 kommen aus dem Ausland.

Seit der Gründung des Zentrums wurden 61 Doktoranden promoviert, 15 von ihnen Frauen. Die Promovierten des Zentrums gehen etwa zu einem Drittel in die Industrie und arbeiten zu etwa zwei Drittel in universitätsnahen Bereichen. Ferner hat das Zentrum zwölf Professoren und vier Professorinnen hervorgebracht, die an Universitäten im Inland und Ausland lehren und forschen.

Bedeutung für die Region

Mit der Arbeit im Zentrum für Bioinformatik wurden auf dem Campus der Universität des Saarlandes neue Türen aufgestoßen. Einen Studiengang mit derart hohem interdisziplinären Anteil hat es hier bis dato nicht gegeben, er war auch in der nationalen Universitätslandschaft ein Novum. In der Tat genießt die partnerschaftliche Zusammenarbeit zwischen verschiedenen Fakultäten, wie sie im Zentrum für Bioinformatik Saar zum Wohle der Studierenden stattfindet, auch im Ausland hohes Ansehen und dient als Blaupause für ähnliche Unternehmungen. Unsere Bachelor- und Masterstudiengänge waren auch die ersten, die an der Universität des Saarlandes akkreditiert wurden. Mit seiner Arbeit und der Vorbildfunktion, die es für die Universität und darüber hinaus bietet, hat das Zentrum für Bioinformatik daher einen wertvollen Beitrag zur Überwindung des monolithischen Selbstverständnisses vieler Fächer geleistet. Dies ist umso wichtiger vor dem Hintergrund, dass sich die immer komplexer werdenden Innovationen heute meist an den Grenzen zwischen den Forschungsdisziplinen bewegen.

Nationale und Internationale Bedeutung

Über die Region hinaus hat das Zentrum wichtige Beiträge zur Entwicklung der nationalen und internationalen Bioinformatik geleistet. Im Folgenden werden die bedeutendsten aufgelistet:

Prof. Lengauer ist Gründungsmitglied und Vizepräsident der International Society for Computational Biology, die unter anderem die größte jährliche Bioinformatiktagung weltweit organisiert.

Im Jahre 2002 gründete das Zentrum in Gestalt von Prof. Lengauer und Prof. Lenhof die größte jährliche europäische Bioinformatik Tagungsreihe European Conference on Computational Biology (ECCB). Die erste Tagung wurde in Saarbrücken abgehalten und zog direkt 450 Teilnehmer an. Heute liegt die Anzahl der Tagungsteilnehmer zwischen 900 und 1200.

Prof. Walter koordiniert das vom BMBF geförderte DEEP Projekt (Deutsches Epigenom Programm), das für die Entwicklung der Epigenetik in Deutschland prägend ist.

Prof. Lengauer war von 2006 bis 2015 Obmann der Sektion Informationswissenschaften der Deutschen Wissenschaftsakademie Leopoldina. Diese Sektion beheimatet auch die Bioinformatik. Derzeit ist Prof. Lengauer Mitglied des Präsidiums der Leopoldina. Ferner sind Prof. Lengauer und Prof. Müller Mitglieder der Deutschen Akademie für Technikwissenschaften acatech.

Seit dem Jahr 2001 organisiert Prof. Helms zusammen mit Prof. Helmut Grubmüller (Max-Planck-Institut für Biophysikalische Chemie, Göttingen) den jährlichen „*Workshop on Computer Simulation and Theory of Macromolecules*“ mit jeweils etwa 120 Teilnehmern. Diese Konferenz wirkt strukturbildend für dieses Forschungsgebiet in Deutschland.

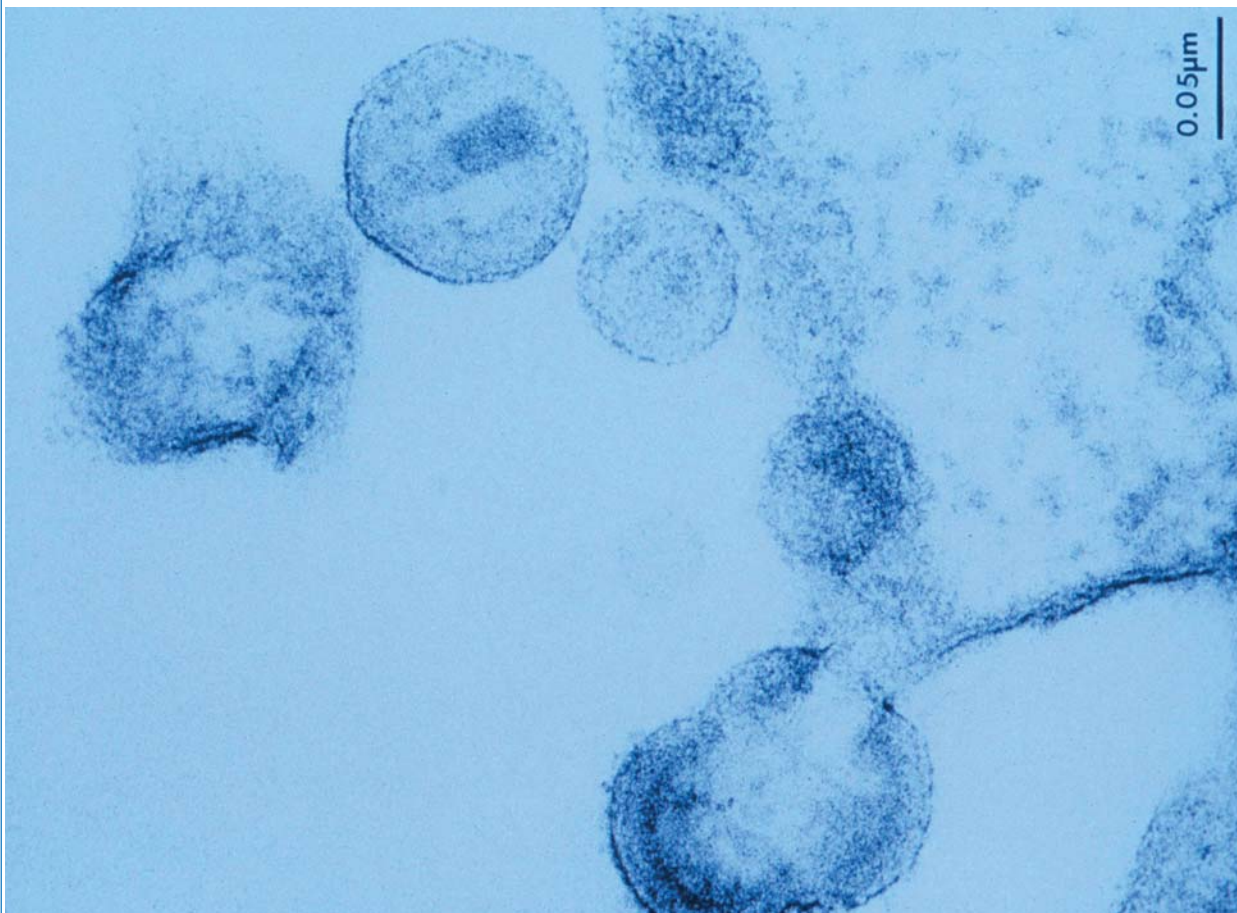
Mitglieder des Zentrums haben bedeutende Wissenschaftspreise erhalten, so etwa den Phoenix-Pharmazie Wissenschaftspreis (R. Müller, 2010, 2014), die Konrad Zuse Medaille (T. Lengauer, 2003), den Karl-Heinz Beckurts Preis (T. Lengauer, 2003), den Preis für Innovatoren unter 35 des Magazins Technology Review (V. Wolf, 2013), den AIDS Forschungspreis der Heinz-Ansmann Stiftung (T. Lengauer, 2010), den Dechema-Preis der Max Buchner-Forschungstiftung (R. Müller, 2010), das Bundesverdienstkreuz am Bande (N. Graf, 2011), den Grant Fertility Innovation (A. Keller, 2015) und den Hector Wissenschaftspreis (T. Lengauer, 2015). C.-M. Lehr erhielt Forschungspreise des Bundesministeriums für Ernährung, Landwirtschaft und Verbraucherschutz (2011) und des Landes Rheinland-Pfalz (2011). Ferner ging der hoch angesehene Preis des European Research Council (ERC Grants) an C. Bock (2016).

Eine ausführliche Präsentation der ab dem Jahr 2010 an ZBI Mitglieder überreichten Preise findet sich von Seite 103 bis 107.

Zusammenfassend hat das Zentrum für Bioinformatik Saar seit seiner Gründung eine Tür in die interdisziplinäre Forschung aufgestoßen. Durch seine erfolgreiche Arbeit wurden wichtige Impulse in der Region und darüber hinaus gesetzt. Absolventen des Zentrums bevölkern heute sowohl die regionale, als auch die internationale Bioinformatik-Szene und prägen so den Fortschritt in diesem Gebiet mit. Das Zentrum für Bioinformatik hat so das Ansehen des Saarlandes in der weltweiten Wissenschaftsszene signifikant gestärkt. Unser Anspruch ist es, aufbauend auf diesen Ergebnissen und Erfolgen, auch in der Zukunft weitere Beiträge für das Land zu leisten und so entscheidend an der Stärkung der Biotechnologie und des Gesundheitswesens im Saarland mitzuwirken.

3 Forschungsgebiete

3 Forschungsgebiete



Mikroskopisches Bild von HI-Viren, die von einer infizierten Zelle abknospen (Foto: Professor Schneeweis)

Bioinformatische Methoden	22
Analyse von genomischer Variation	22
Statistisches Lernen in der Bioinformatik	28
Biomolekulare Interaktionen	34
Biologische Netzwerke	40
Epigenetik	48
Anwendungen der Bioinformatik	54
Infektionskrankheiten	54
Krebs	62
Weitere medizinische Anwendungen	68
Biotechnologische Anwendungen	74

Analyse von genomischer Variation

Jede unserer Körperzellen trägt eine vollständige Kopie unserer DNA in sich und damit unser vollständiges Erbgut, das auch als Genom bezeichnet wird. Das Genom ist in Chromosomen organisiert: Wir erben jeweils 23 Chromosomen von unserer Mutter und unserem Vater. Ein Chromosom besteht aus einem sehr langen DNA-Molekül, welches durch spezielle Proteine stabilisiert und räumlich angeordnet wird. Die DNA kodiert genetische Informationen als Abfolge von vier verschiedenen Basen: Adenin (A), Cytosin (C), Guanin (G) und Thymin (T). In diesem Sinne kann ein Chromosom vereinfacht als eine Sequenz der Buchstaben A, C, G und T betrachtet werden. Die Genomforschung beschäftigt sich demnach mit der Frage, wie unsere Genomsequenz individuelle Merkmale beeinflusst, also welche kausalen Zusammenhänge zwischen genetischer Variation und der Ausprägung verschiedener Merkmale bestehen. Viele verschiedene Eigenschaften können dabei von Interesse sein: Körpergröße, die Anfälligkeit für eine bestimmte Krankheit oder die Intoleranz gegenüber einem bestimmten Arzneimittel.

Genetische Veränderungen, die Ursache bestimmter Merkmale sind, befinden sich meistens in den Genen. Dies sind DNA-Regionen, die als Blaupause für Proteine dienen – oder sie liegen in Regionen, die steuern, wie viel von einem gewissen Protein produziert wird. Aus den in der DNA kodierten Genen, die man sich wie oben erwähnt als Baupläne vorstellen kann, werden in zwei Schritten Proteine hergestellt. Als erstes werden Kopien dieser DNA-Abschnitte erstellt, die dann als RNA-Moleküle vorliegen. Dieser Vorgang heißt Transkription. Danach werden aus den RNA-Molekülen Proteine produziert, die eine Vielzahl von Funktionen ausüben. Somit wirken sich Änderungen in der DNA über den Umweg der RNA auf die hergestellten Proteine aus. Diese Veränderungen können beträchtliche Auswirkungen auf die Ausprägung von Merkmalen haben und zum Beispiel die Ursache einer genetischen Krankheit sein.

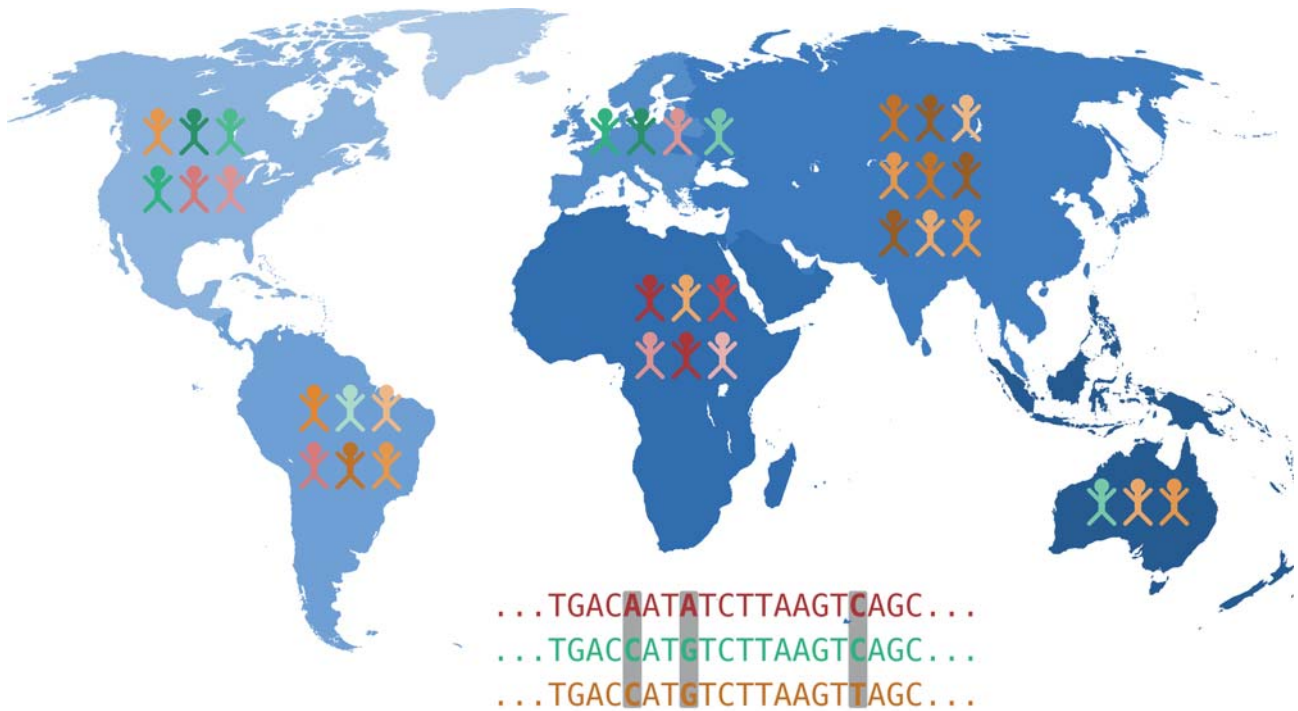
Aber wie können Unterschiede in der DNA oder RNA gemessen werden? In beiden Fällen können hochmoderne Sequenzierverfahren (engl. next generation sequencing, NGS), die ständig weiterentwickelt werden, Verwendung finden. Diese Technologien haben in den letzten Jahren eine enorme Entwicklung durchlaufen, wobei sich der Sequenzierdurchsatz um ein Vielfaches gesteigert hat und die Kosten der Sequenzierung deutlich gesunken sind. Mit neuesten Methoden ist es möglich, die komplette DNA eines Menschen zu einem Preis um 1.000 Euro zu sequenzieren. Die verwendeten Sequenziergeräte lesen in kurzer Zeit Millionen von kurzen DNA- oder RNA-Fragmenten. Das heißt, man erhält nicht den ganzen DNA-Strang, sondern nur sehr viel kleinere Teilstücke, die sogenannten Reads. Die Analyse der erhaltenen Daten ist somit vergleichbar mit der Lösung eines hochkomplexen Puzzles: Man versucht die korrekte DNA-Sequenz aus den kurzen Reads zu rekonstruieren.

Obwohl dies bereits ausgefeilte Methoden erfordert, werden wir häufig mit noch komplexeren Situationen konfrontiert. Das hängt damit zusammen, dass man es oft nicht mit einem einzelnen Genom, sondern mit einer Mischung von Genomen zu tun hat. Denn wie eingangs erwähnt erben wir je einen Satz Chromosomen von Mutter und Vater. Diese sind sich zwar ähnlich, aufgrund der genetischen Unterschiede zwischen den Eltern jedoch nicht identisch. Jedes von der Sequenziermaschine gelesene Fragment stammt also entweder von einer Chromosomenkopie vom Vater oder aber von der Mutter. Bildlich gesprochen haben wir es nun also mit zwei Puzzles zu tun, deren Teile gemischt sind und die fast identische Motive haben.

Eine noch kompliziertere Mischung von Genomen liegt bei Tumorproben vor. Krebszellen entwickeln sich aus normalen Körperzellen, häufen aber im Laufe der Zeit viele genetische Mutationen an. Das Genom einer Krebszelle kann sich daher deutlich vom Originalgenom des erkrankten Menschen unterscheiden. Damit nicht genug: auch innerhalb eines Tumors herrscht Heterogenität und die Genome der einzelnen Krebszellen unterscheiden sich teilweise drastisch. Da beim Sequenzieren in der Regel eine größere Menge Zellen verwendet werden muss, hat man es mit einer komplexen Mischung vieler Krebsgenome zu tun, die es im Computer zu entwirren gilt.

Auch Viren haben ein Genom, das je nach Virustyp als DNA oder als RNA vorliegt. Ist ein Mensch zum Beispiel mit HIV infiziert, so befinden sich im Körper eine Vielzahl von Viruspartikeln, die alle eine Kopie des Virusgenoms in sich tragen. Im Laufe einer langen Infektion mutieren die Viren im Körper und bilden eine Population mit beträchtlicher genetischer Vielfalt. Sequenziert man nun die Viruspartikel aus einer Blutprobe, steht man – ähnlich wie oben für Krebs beschrieben – einer komplexen Mischung an Genomen gegenüber.

Die Analyse der bei der Sequenzierung entstehenden riesigen Datenmengen stellt Wissenschaftler vor Herausforderungen im Bereich der Informatik, Mathematik und Statistik. Forscher am Zentrum für Bioinformatik stellen sich diesen Herausforderungen und haben wesentlich zur Weiterentwicklung moderner Analysemethoden von DNA- sowie RNA-Sequenzierdaten beigetragen. Nachfolgend werden vier wichtige Beispiele erläutert.



Während die Genome aller Menschen zu über 99% identisch sind, zeigen sich jedoch an bestimmten Positionen Unterschiede. Diese genomische Variabilität ist wichtig für die Weiterentwicklung der Spezies Homo sapiens. Einige dieser genomischen Varianten erhöhen aber das Risiko für bestimmte Erkrankungen.

Analyse von genomischer Variation

Genomische Variationen zwischen Menschen	24
Detektion von Fusionsgenen in Krebszellen	25
Gemischte Genome	26
Suche nach viralem Ursprung von Tumoren	27

Genomische Variationen zwischen Menschen

24

Genome von zwei Personen können sich auf viele Arten unterscheiden, wie in Abbildung 1 dargestellt ist. Derzeit konzentrieren sich die meisten genetischen Studien auf die relativ leicht zu identifizierenden Punktmutationen, die auch als SNPs (von engl. Single Nucleotide Polymorphism) bezeichnet werden. Man spricht von einer Punktmutation, wenn ein Buchstabe (Nukleotid) an einer bestimmten Position im Genom gegen einen anderen Buchstaben ausgetauscht wurde. Sich auf Punktmutationen zu beschränken, ist aber oft nicht ausreichend, denn sie erklären nur einen Teil der genetischen Diversität.

Daher werden nun auch verstärkt größere genomische Veränderungen analysiert, die sogenannten strukturellen Variationen wie Deletionen (Löschungen) oder Duplikationen von ganzen DNA-Segmenten. Solche strukturellen Varianten treten häufig auf. Um das gesamte Spektrum der menschlichen genetischen Vielfalt zu erfassen, ist ihr Studium unverzichtbar. Eine besondere Bedeutung kommt ihnen auch in Genomen von Krebszellen zu. Denn Genome von Krebszellen sind oft besonders instabil und damit anfällig für strukturelle Veränderungen. Solche strukturellen Veränderungen können große Auswirkungen auf die Eigenschaften von Krebszellen und ganzen Tumoren haben. Das Studium von strukturellen Variationen in gesundem sowie in Tumorgewebe basiert auf neuen Sequenzieretechnologien in Kombination mit neuen algorithmischen Methoden, die wir aktiv entwickeln.

Im „Genome of the Netherlands“-Projekt wurden die kompletten Genome von Mutter, Vater und einem Kind aus 250 niederländischen Familien sequenziert [1]. Tobias Marschall ist seit seiner Zeit als Postdoktorand am CWI Amsterdam an diesem Projekt beteiligt und arbeitet als Juniorprofessor am ZBI weiter an der Analyse dieses umfangreichen Datensatzes, insbesondere im Hinblick auf strukturelle Variationen. Durch das Lesen der Genome von Eltern und Kindern können wir bestimmen, welche genetischen Varianten in den Kindern vorhanden sind, die nicht von den Eltern geerbt wurden, sondern das Ergebnis von neuen Mutationen sind. Die Häufigkeit, mit der solche neuen Mutationen auftreten, ist dabei von großem Interesse, weil sie ein wichtiger Parameter in Modellen der menschlichen Evolution ist.

Über die Grundlagenforschung hinaus ist das Studium struktureller genetischer Varianten potenziell wichtig für die personalisierte Medizin. Bisher wurden Assoziationsstudien, die Verbindungen zwischen dem Genom einer Person und klinisch relevanten Merkmalen suchen, meist beschränkt auf Punktmutationen. Jüngste Forschungen mit Beteiligung des ZBI lieferten deutliche Hinweise, dass die Berücksichtigung von strukturellen genetischen Varianten in solchen Studien zu neuen, medizinisch relevanten Erkenntnissen führt. Daher arbeiten wir aktiv an statistischen Methoden, um Genotypen von strukturellen Varianten aus SNP-Daten vorherzusagen. Dies versetzt klassische, SNP-basierte Assoziationsstudien in die Lage, ohne Mehraufwand Kandidaten für kausale strukturelle Varianten aufzuspüren, also für solche Varianten, die ursächlich für eine klinische Eigenschaft sind.

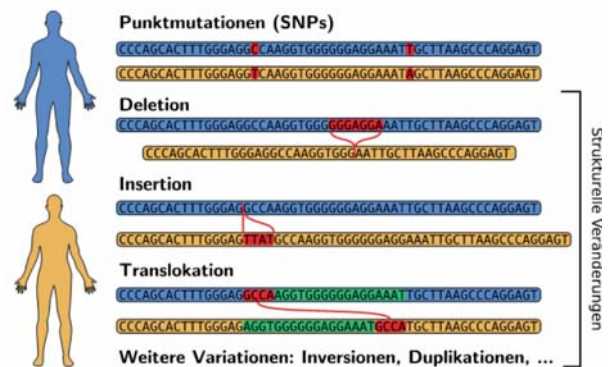


Abbildung 1 // Genetische Unterschiede zwischen Menschen

Referenzen

- [1] *The Genome of the Netherlands Consortium. Whole-genome sequence variation, population structure and demographic history of the Dutch population. Nature Genetics, 46, 818-825, 2014*



Juniorprofessor Dr. Tobias Marschall
 Algorithms for Computational Genomics
 Universität des Saarlandes /
 Max-Planck-Institut für Informatik
 Telefon +49 681 302 70880
 marschall@cs.uni-saarland.de

Detektion von Fusionsgenen in Krebszellen

Gefördert durch „Cluster of Excellence: Multimodal Computing and Interaction“

In Krebsgenomen können viele DNA-Veränderungen gefunden werden, von denen einige für das Krebswachstum verantwortlich sind. Beispiele hierfür sind Fusionsgene, die durch Translokationen von DNA-Bereichen entstehen. Eine solche Verlagerung eines Genomabschnitts führt dazu, dass zwei Gene, die in der Regel auf unterschiedlichen Chromosomen liegen, benachbarte DNA-Regionen im Krebs-Genom werden (siehe Abbildung 2). Dieses neue Fusionsgen kann entweder die Funktionen der beiden Ursprungsgene erben oder eine völlig neue Funktion entwickeln und damit zum Beispiel die normale Funktionsweise eines der ursprünglichen Gene beeinträchtigen.

Beispielsweise ist bei Patienten mit chronischer myeloischer Leukämie in einigen Fällen das Bcr-Abl-Fusionsgen zu finden, welches eine hyperaktive Form des Abl-Gens von gesunden Individuen darstellt. Diese verstärkte Aktivität des Fusionsgens ermöglicht es dem Krebs, schneller zu wachsen. Weil dieses Fusionsgen aber nur in Krebszellen vorkommt, war es möglich, dieses Gen als Ziel für Medikamente einzusetzen und gezielt Wirkstoffe dagegen zu entwickeln (siehe Abbildung 2). Im Laufe der Jahre wurden mehrere Medikamente entwickelt, die das erzeugte Protein des Fusionsgens hemmen und damit erlauben, diese spezifische Form des Krebses erfolgreich zu behandeln.

Mit Hilfe von neuartigen Sequenziermethoden hat man viele unterschiedliche Krebsarten untersuchen können und dabei festgestellt, dass Fusionsgene weit verbreitet sind. Die Entwicklung von Methoden, die der Identifikation von Fusionsgenen aus Sequenzierdaten dienen, ist deshalb zu einem wichtigen Forschungsgebiet geworden. Damit können neue Therapieansätze für die verschiedenen Krebsarten gefunden werden.

In Krebszellen sind viele DNA-Regionen im Gegensatz zu gesunden Zellen verändert. Um neue Fusionsgene zu finden ist der Vergleich mit bekannten Genomsequenzen in Datenbanken oft nicht hilfreich. Ein alternatives und erfolgreiches Konzept für die Erkennung von Fusionsgenen ist aber die de-novo-Assemblierung von exprimierter RNA aus Sequenzierdaten von Krebszellen. Bei der de-novo-Assemblierung wird nicht die bekannte Datenbanksequenz des menschlichen Genoms verwendet, was die Verfahren wesentlich komplizierter macht. Eine der ersten Methoden zur de-novo-Assemblierung von großen RNA-Datensätzen hat Marcel Schulz in seiner Zeit als Doktorand am MPI für molekulare Genetik in Berlin in Zusammenarbeit mit Wissenschaftlern vom European Bioinformatics Institute (EMBL-EBI) entwickelt [1].

Wie erwähnt gleicht die Aufgabe einem Puzzle: Man versucht die kompletten RNA-Sequenzen durch das Aneinanderfügen von vielen Millionen Einzelteilen, also kurzen Teilsequenzen, zu bestimmen, mit dem Unterschied, dass man bei diesem Textpuzzle kein „Bild“ als Vorlage hat, das einem helfen könnte zu bestimmen, welche Teile wie zusammengehören. Die Gruppe von Dr. Schulz hat am ZBI das ursprüngliche Verfahren entscheidend verbessert und es verlässlicher gemacht. Mit Hilfe eines neuen Algorithmus können auch extrem große RNA-Datensätze, wie sie mit neuesten Technologien erzeugt werden, effizient analysiert werden.

Dadurch wird die Entdeckung von Fusionsgenen selbst in komplizierten Krebsgenomen ermöglicht, was mit herkömmlichen Methoden kaum oder gar nicht möglich ist. Unsere Software wurde daher kürzlich schon verwendet, um bisher unbekannte Fusionsgene in Lungenkrebszellen nachzuweisen.

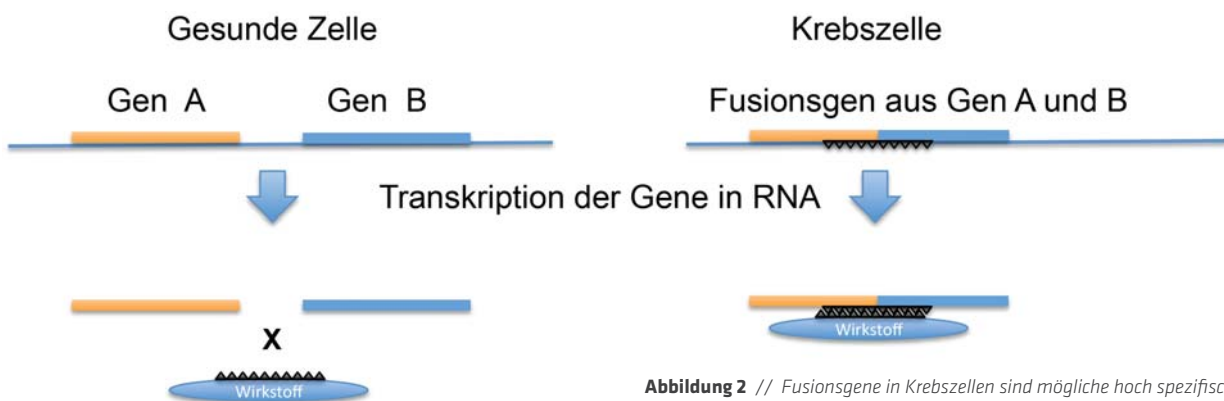


Abbildung 2 // Fusionsgene in Krebszellen sind mögliche hoch spezifische Angriffspunkte für eine effektive medikamentöse Behandlung, da sie in normalen Zellen nicht vorkommen.

Referenzen

- [1] M. H. Schulz, D. R. Zerbino, M. Vingron and E. Birney. **Oases: robust de novo RNA-seq assembly across the dynamic range of expression.** *Bioinformatics* 28 (8): 1086-92. 2012



Dr. Marcel Schulz
 High-throughput Genomics & Systems Biology
 Universität des Saarlandes /
 Max-Planck-Institut für Informatik
 Telefon +49 681 9325 3115
 mschulz@mmci.uni-saarland.de

Gemischte Genome

26

Es gibt etliche Szenarien, in denen wir Proben analysieren, die eine Mischung von verschiedenen Genomen beinhalten. Ein Beispiel hierfür sind diploide Organismen wie der Mensch, die je einen Satz Chromosomen von Mutter und Vater erben. Jede Zelle enthält somit zwei sehr ähnliche Varianten der Chromosome, die gemeinsam sequenziert werden.

Bei der Sequenzierung von Viruspartikeln aus Patientenblut und von Tumorproben hat man es darüber hinaus durch fortwährende Mutation mit teils sehr heterogenen Mischungen von Genomen zu tun. Generell soll jeweils aus heterogenen Sequenzierdaten auf die vorliegende Mischung von Sequenzen in der jeweiligen Probe zurückgeschlossen werden. Dies erfordert neue leistungsfähigere Verfahren, die wir am ZBI entwickeln.

Wie kann man nun Fragmente, die von derselben Kopie eines Chromosoms – auch Haplotyp genannt – stammen, erkennen und zusammensetzen? Zunächst benötigt man eine Sequenzier-technologie, die Textfragmente (Reads) hinreichender Länge generiert. Denn zur Rekonstruktion von Haplotypen benötigen wir Reads, die zwei oder mehr Stellen aufweisen, an denen sich die jeweiligen Haplotypen unterscheiden. Sequenziergeräte der dritten Generation, wie sie zurzeit auf den Markt kommen, haben hier deutliche Vorteile gegenüber den weit verbreiteten Geräten der zweiten Generation. Keines dieser Geräte arbeitet jedoch fehlerfrei und besonders solche mit längeren Reads liefern oft Daten mit vielen Fehlern.

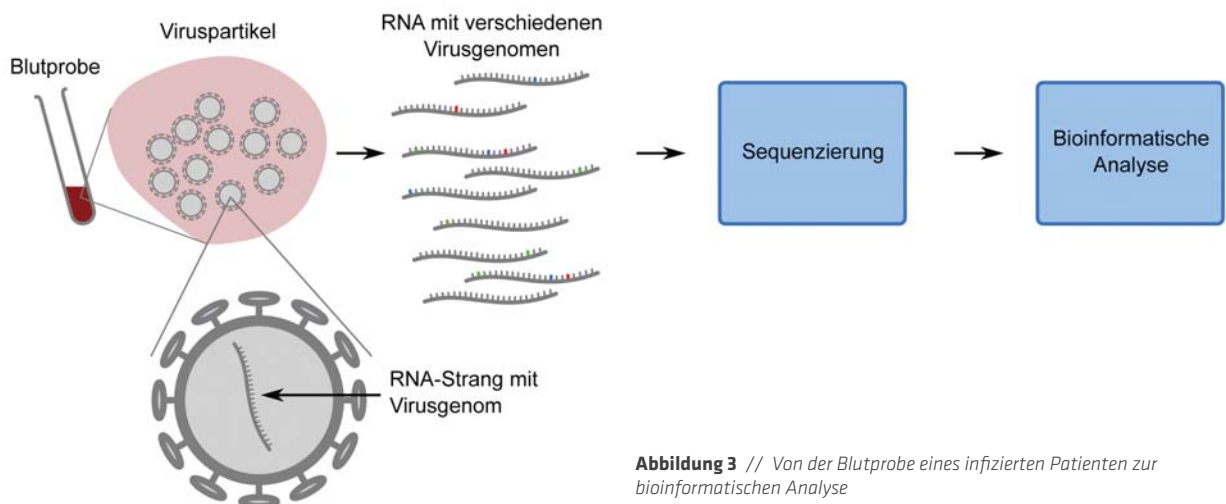


Abbildung 3 // Von der Blutprobe eines infizierten Patienten zur bioinformatischen Analyse

Referenzen

- [1] M. Patterson, T. Marschall, N. Pisanti, L. van Iersel, L. Stougje, G. W. Klau und A. Schönhuth. **WhatsHap: Weighted Haplotype Assembly for Future-Generation Sequencing Reads.** *J Comput Biol* 22 (6): 498-509. 2015



Juniorprofessor Dr. Tobias Marschall
Algorithms for Computational Genomics
Universität des Saarlandes /
Max-Planck-Institut für Informatik
 Telefon +49 681 302 70880
 marschall@cs.uni-saarland.de

Als Bioinformatiker entwickeln wir statistische Modelle, die Sequenzierfehler erkennen, ohne dabei die echten Unterschiede zwischen zwei Haplotypen zu übersehen. Die nächste Herausforderung, der wir uns stellen, besteht im Entwurf effizienter Algorithmen zur Haplotyp-Rekonstruktion, denn oftmals fallen die resultierenden informatischen Probleme in die Klasse der sogenannten NP-harten Probleme.

Das sind Probleme, für die es keine Lösungsverfahren gibt, die in allen Fällen schnell zu einer Lösung kommen. Bereits seit seiner Zeit als Postdoktorand am CWI Amsterdam arbeitet Tobias Marschall an Methoden, die dennoch in der Lage sind, die in der Praxis auftretenden Fälle hinreichend schnell zu lösen [1]. Seine Arbeitsgruppe am ZBI befasst sich auch aktuell mit der Entwicklung entsprechender Algorithmen. Aber welchen Nutzen hat die Rekonstruktion von Haplotypen? Wir wollen dies am Beispiel der HIV-Infektion erläutern. Das HI-Virus hat eine sehr hohe Mutationsrate. Das führt dazu, dass die Viruspartikel in einem Patienten im Laufe der Zeit eine immer größere genetische Vielfalt entwickeln. Der Wirt beherbergt also eine ganze Viruspopulation (siehe Abbildung 3). Zur Behandlung ist die Frage entscheidend, welche Resistenzen gegen die zur Verfügung stehenden Medikamente in dieser Viruspopulation vorhanden sind. Um diese Frage so genau wie möglich beantworten zu können, ist die Kenntnis aller vorhandenen Haplotypen und deren Häufigkeiten notwendig.

Die vielfältigen Anwendungen – von Viruspopulationen über menschliche Genome bis hin zu Tumorerogenität – verdeutlichen, welche Bandbreite an praktischen Problemen sich bisweilen mit sehr ähnlichen Techniken lösen lässt und unterstreicht damit auch die Bedeutung der Forschung an neuen Methoden, wie sie am ZBI stattfindet.

Suche nach viralem Ursprung von Tumoren

Gefördert durch das Oncogene Projekt des Bundesministeriums für Gesundheit und Forschung (BMBF)

Virale und bakterielle Infektionen sind der Grund für die Entstehung von bis zu 20% der bösartigen Tumoren, einschließlich diverser Gebärmutter-, Leber- und Magentumore. Meistens geht die Infektion auf Krebs erzeugende (onkogene) Viren, wie das menschliche Papillomavirus (HPV), Epstein-Barr-Virus (EBV), Hepatitis-B-Virus (HBV) und Hepatitis-C-Virus (HCV) zurück. Epidemiologische Daten legen aber nahe, dass diese Liste noch unvollständig ist und in Zukunft weitere onkogene Viren gefunden werden. Darum hat die Charakterisierung des onkogenen Potenzials von Viren wichtige Konsequenzen für die Vorbeugung, Diagnose und Therapie von bösartigen Tumoren.

Krebs erzeugende Viren sind besonders schwer nachzuweisen.

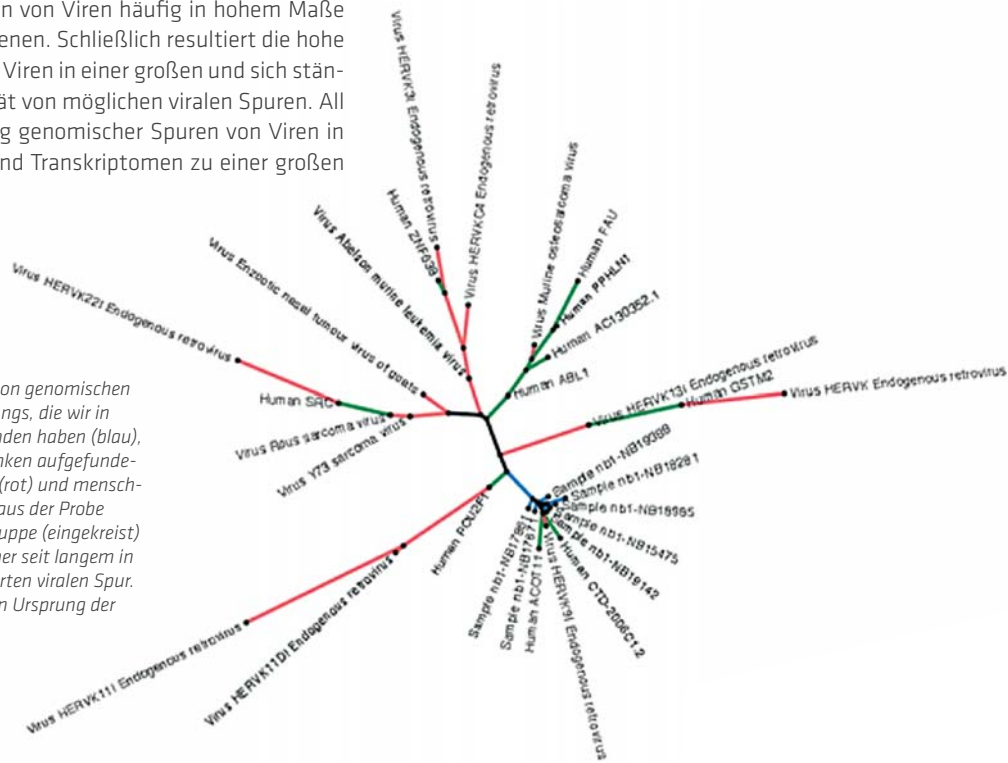
Die jüngsten Fortschritte in der Sequenzierungstechnologie haben die systematische Durchforstung von Genomen und Transkriptomen (also der Gesamtheit der in der Zelle abgelesenen Gene) bösartiger Tumoren möglich gemacht. Wir suchen mit dieser Technologie nach genomischen Spuren von Viren. Aufgrund von Störfaktoren, wie einer geringen Vervielfältigungsrate des Virus und einem zügigen Verlust des genetischen Materials des Virus in der Zelle nach der Tumorentstehung, sind solche genomischen Spuren meistens nur sehr schwach und schwer zu detektieren. Darüber hinaus ähneln genomische Spuren von Viren häufig in hohem Maße normalen menschlichen Genen. Schließlich resultiert die hohe evolutionäre Dynamik von Viren in einer großen und sich ständig verändernden Diversität von möglichen viralen Spuren. All dies macht die Auffindung genomischer Spuren von Viren in menschlichen Genomen und Transkriptomen zu einer großen Herausforderung.

Eine neue Methode zum Auffinden schwacher genomischer Spuren von Viren.

Wir haben uns dieser Herausforderung gestellt und Transkriptome von über 500 Neuroblastom-Tumoren sowie eine Reihe von Virus-positiven Kontrollproben auf das menschliche Genom und über 4.000 bekannte virale Genome abgebildet.

Damit konnten wir sowohl bekannte als auch unbekannte Viren aufspüren. Auf der Basis von theoretischen Schätzungen, der Analyse von positiven Kontrollen sowie Vergleichen mit anderen Methoden konnten wir zudem die hohe Sensitivität unserer Methode nachweisen. Trotzdem förderte eine detaillierte Analyse aller aufgefundenen viralen Gene in Neuroblastom-Proben keine für die Krebsentstehung relevanten Spuren zu Tage (siehe Abbildung 4). Daher legt unsere Analyse nahe, dass die Beteiligung viraler Kofaktoren bei der Entstehung von Neuroblastomen unwahrscheinlich ist [1]. Ein solcher Nachweis der Abwesenheit von viralen genomischen Spuren ist sehr schwierig zu erreichen und ein gewisses Novum in diesem Gebiet. Wir konnten auf der Basis unserer Datenanalyse somit eine Vermutung bestätigen, die bei der Subtypisierung von Neuroblastom-Tumoren klinisch hoch relevant ist. Wir stellen unsere Methode als Open-Source-Software für andere Forscher zur Anwendung und als Basis für Weiterentwicklungen zur Verfügung.

Abbildung 4 // Stammbaum von genomischen Spuren möglichen viralen Ursprungs, die wir in einer Neuroblastom-Probe gefunden haben (blau), gemeinsam mit in den Datenbanken aufgefundenen ähnlichen Spuren von Viren (rot) und menschlichen Genen (grün). Die Spuren aus der Probe sammeln sich in einer kleinen Gruppe (eingekreist) aus menschlichen Genen und einer seit langem in das menschliche Genom integrierten viralen Spur. Das deutet auf den menschlichen Ursprung der Spuren in der Probe hin.



Referenzen

- [1] S. E. Schelhorn, M. Fischer, L. Tolosi, J. Altmüller, P. Nurnberg, H. Pfister, T. Lengauer and F. Berthold. **Sensitive detection of viral transcripts in human tumor transcriptomes.** *PLoS Comput Biol* 9 (10): e1003228. 2013



Prof. Dr. Dr. Thomas Lengauer

Abt. Bioinformatik und Angewandte Algorithmik
 Max-Planck-Institut für Informatik
 Telefon +49 681 9325 3000
 lengauer@mpi-inf.mpg.de

Statistisches Lernen in der Bioinformatik

In vielen Bereichen des Lebens ist es in den letzten Jahren zu einem exponentiellen Wachstum von Daten gekommen. Dies liegt zum einen an immer leistungsfähigeren Computern, die in der Lage sind, große Datenmengen zu verarbeiten und zu speichern, zum anderen aber auch an neuen Messmethoden in vielen Bereichen der Wissenschaft, den sogenannten Hochdurchsatzverfahren. Dies sieht man zum Beispiel in der Physik an Messungen, die am CERN durchgeführt werden, aber auch in der Biologie gibt es immer rasanter ansteigende Datenmengen (z. B. Genom-, Proteom- und Epigenomdaten). Diese Datenflut macht es unmöglich, grundlegende Zusammenhänge in den Daten durch eine reine Betrachtung zu erschließen. So möchten wir bei Genomsequenzen zum Beispiel wissen, an welcher Stelle sich die Gene befinden. Kann man durch zusätzliche Experimente herausfinden, wo sich Gene befinden, so kann man mit Hilfe von statistischen Lernverfahren generelle Sequenzeigenschaften von Genen lernen, die dann in einem Vorhersagemodell verwendet werden können. Dies macht es möglich, für weitere Genomsequenzen die Positionen der Gene vorherzusagen. Diese Vorgehensweise ist insbesondere dann interessant, wenn die zusätzlichen Experimente zeitaufwendig und teuer sind, was für viele Bereiche der Biologie zutrifft.

Die Methoden, die in diesem Szenario eingesetzt werden, gehören zu der Gruppe der statistischen Lernverfahren. Sie werden aber auch in anderen Bereichen angewendet und weiterentwickelt, wie zum Beispiel zum Herausfiltern von Spam Emails, zum Vorhersagen der passendsten Werbung oder zur Vorhersage von Aktientrends. Beim statistischen Lernen gibt es zwei wichtige Bereiche. Der eine umfasst das überwachte Lernen. Hier werden Daten mit bestimmten Labels verwendet um zu lernen, welche Eigenschaften der Daten dazu geeignet sind, diese Labels vorherzusagen. Labels können in dem Fall einfache ja/nein Entscheidungen sein, wie auf Seite 30 beschrieben wird, aber auch komplexe Formen annehmen, wie auf Seite 32 zur Metagenomik beschrieben, wo das Label einem ganzen Pfad in einem Abstammungsbaum entspricht.

Ein anderer wichtiger Bereich des statistischen Lernens ist das unüberwachte Lernen. Hierbei gibt es keine Labels, aber Daten werden aufgrund von Ähnlichkeiten gruppiert, um zum Beispiel interessante Untergruppen zu finden. Dies wird im Beitrag auf Seite 31 thematisiert.

In der Bioinformatik ist das statistische Lernen in allen wichtigen Bereichen vertreten. So werden beispielsweise statistische Lernverfahren entwickelt, um den Zusammenhang zwischen Mutationen und bestimmten Krankheiten aufzudecken.

Auch am Zentrum für Bioinformatik nimmt das statistische Lernen einen wichtigen Platz ein. So werden zum Beispiel in der Abteilung von Prof. Lengauer schon seit mehr als einem Jahrzehnt neue Verfahren entwickelt, um die Resistenz

des HI-Virus gegen bestimmte antivirale Medikamente vorherzusagen. Die Server, die diese Methoden frei zur Verfügung stellen, werden weltweit eingesetzt.

Da auch in diesem Bereich immer wieder neue Messverfahren zur Verfügung stehen, sind auch immer wieder neue Methoden erforderlich, um alle Eigenschaften der Daten bestmöglich abzubilden. Die Tiefensequenzierung ist eines dieser neuen Messverfahren. Hiermit kann man virale Populationen viel genauer untersuchen, als dies mit herkömmlichen Sequenziermethoden möglich war. Wie man diese zusätzlichen Informationen nutzen kann, um damit die Resistenzvorhersage zu verbessern, ist auf Seite 30 dargestellt.

Ein weiterer wichtiger Bereich der Biologie ist die Untersuchung von evolutionären Beziehungen zwischen verschiedenen Spezies. Hierbei werden Stammbäume, die auch phylogenetische Bäume genannt werden, bestimmt. Dort gibt die Entfernung im Baum an, wie weit Organismen evolutionär voneinander entfernt sind. Am ZBI wurden Methoden entwickelt, die das Lernen dieser phylogenetischen Bäume signifikant verbessern (siehe Seite 33).

Daneben stellt auch die Metagenomik ein elementares Teilgebiet der Bioinformatik dar. Hier werden komplexe Bakteriengemeinschaften sequenziert, um Aussagen über die Häufigkeit und potentiell auch den Einfluss bestimmter Bakterien auf die jeweilige Umgebung zu untersuchen. Die Proben können hierbei zum Beispiel aus dem Darm von Patienten, aber auch aus dem New Yorker U-Bahnsystem stammen. Ein wichtiger Schritt in diesen Analysen ist die Vorhersage, welche Teilsequenz von welchem Organismus stammt. Hier haben Forscher am Zentrum für Bioinformatik ein Vorhersagemodell entwickelt, das in der Lage ist, sehr effizient und mit hoher Performanz vorherzusagen, in welchem Bereich des phylogenetischen Baums sich der Organismus befindet, von dem diese Sequenz stammt. Nähere Informationen dazu finden sich auf Seite 32.

Ebenfalls von hoher Bedeutung ist der Bereich der Krebsdatenanalyse. Hier ist es interessant, bestimmte Untergruppen von Patienten mit Methoden des unüberwachten Lernens zu finden, aber auch herauszufinden, welche Veränderungen in der Regulation der Proteine einen Einfluss auf den Schweregrad der Krebserkrankung haben. In diesem Bereich hat das ZBI neue Methoden entwickelt, die es zum Beispiel ermöglichen, Messungen verschiedener biologischer Eigenschaften zu vereinen, um eine einfach verständliche Repräsentation der Daten zu erzeugen und darauf basierende Gruppierungen zu finden (siehe Seite 31). Zudem haben wir uns auch mit Fällen beschäftigt, in denen die Messungen für verschiedene Patienten nicht ganz vergleichbar sind und Lernverfahren entwickelt, die diesen Effekt ausgleichen können.



Statistisches Lernen in der Bioinformatik

Vorhersage des viralen Tropismus bei HIV	30
Integrative Krebsdatenanalyse	31
Taxonomische Einteilung metagenomischer Daten	32
Phylogenomik unter Einbeziehung paraloger Gene	33

Vorhersage des viralen Tropismus bei HIV

30

35 Millionen Menschen waren laut Weltgesundheitsorganisation (WHO) im Jahr 2013 weltweit mit HIV infiziert. Etwa 1,5 Millionen Menschen sind im gleichen Jahr an AIDS gestorben – darunter auch 190.000 Kinder. Die Zahl der HIV-Neuinfektionen lag 2013 laut WHO dabei weltweit etwa bei 2,1 Millionen. Um diesem Problem entgegenzuwirken, sind präventive Maßnahmen unumgänglich. Da eine HIV-Infektion bisher nicht heilbar ist, konzentriert sich ein Großteil der HIV-Forschung darauf, die Symptome der infizierten Patienten mit Hilfe von antiretroviralen Medikamenten zu lindern und deren Leben zu verlängern. Eine Klasse von antiretroviralen Medikamenten gegen HIV ist die der so genannten Eintrittshemmer.

Bevor ein HIV-Eintrittshemmer verschrieben werden kann, muss festgestellt werden, ob er für die Virenpopulation in dem jeweiligen Patienten geeignet ist. HI-Viren können eins von mehreren zellulären Oberflächenproteinen, den sogenannten Corezeptoren, zum Zelleintritt nutzen. Von klinischer Relevanz sind hier der CCR5-Corezeptor und der CXCR4-Corezeptor (siehe Abbildung 5). Da es bisher lediglich Eintrittshemmer gibt, die den CCR5-Corezeptor blockieren, können nur Patienten mit Eintrittshemmern behandelt werden, die hauptsächlich Viren beherbergen, die nur mit Hilfe dieses Corezeptors in die Zelle gelangen können. Viren, die den CXCR4-Corezeptor benutzen können, wären ansonsten in der Lage mit Hilfe dieses Corezeptors in die Zellen zu gelangen, was zu einem Wechsel in der Virenpopulation führen würde.

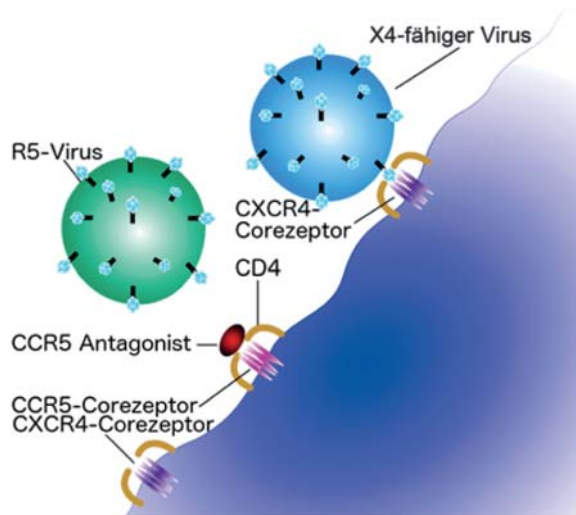


Abbildung 5 // Corezeptornutzung von HI-Viren
Diese Graphik zeigt verschiedene Arten von HI-Viren. R5-Viren (grün) können lediglich den CCR5-Corezeptor für den Zelleintritt verwenden, wohingegen X4-fähige Viren (blau) auch den CXCR4-Corezeptor verwenden können.

Dies möchte man vermeiden, da diese Art von HI-Viren mit schnellerem Fortschreiten der Infektion zum AIDS-Stadium assoziiert ist. Für eine personalisierte Behandlungsempfehlung ist es also wichtig herauszufinden, wie die Virenpopulation des Patienten aussieht.

Um dies festzustellen, verwenden wir den Ansatz, der auf Seite 56 näher beschrieben ist. Hierfür ist jeweils nicht die gesamte genetische Sequenz der HI-Viren notwendig, sondern nur die genetische Sequenz der so genannten V3-Schleife – eines Teils des Proteins von HIV, das in der Virenhülle steckt und beim Zelleintritt an den Corezeptor bindet. Wir haben ein neues Verfahren entwickelt, das mit Hilfe von neuen Sequenzierverfahren generierten komplexen Daten, die Vorhersage der Corezeptor-Nutzung und damit auch die Entscheidung verbessert, ob der Eintrittshemmer verschrieben werden kann [1]. Das Verfahren ist aber prinzipiell auch für andere Szenarien anwendbar. So wäre zum Beispiel auch eine Anwendung in der Analyse von Krebsdaten denkbar. Wir konnten so auch ein statistisches Modell ableiten, welches das Wissen, das man mit den neuen hochaufgelösten Daten in Bezug auf die Vorhersage der Corezeptornutzung erworben hat, auf die Daten anwendet, die mit der ursprünglichen Sequenziermethode gewonnen wurden. Dies ist insbesondere wichtig, da viele Kliniken noch keinen Zugang zu den neuen Sequenzierern haben. Zusätzlich konnten wir wichtige Positionen in der V3-Schleife finden, die aussagekräftig für die jeweilige Corezeptornutzung sind (siehe Abbildung 6). Ein roter Buchstabe bedeutet hierbei, dass der Patient das Medikament eher nicht nehmen sollte, ein grüner steht für Unbedenklichkeit. Dies kann Hinweise darauf geben, wie die Interaktion zwischen V3-Schleife und Corezeptor abläuft, und vielleicht weitere Angriffspunkte gegen HIV aufzeigen.

CTRPSNNTRRRIHIGPGRFYAT~~EE~~IIGEMRQAHC

Abbildung 6 // Ergebnisvisualisierung der Vorhersagemethode
Zusätzlich zu der verbesserten Vorhersagemethode wurde auch eine neuartige Visualisierung entwickelt, die zeigt, welche Aminosäuren an welcher Position einen besonderen Einfluss auf die Vorhersage hatten. Die Sequenz ist ein Beispiel für eine V3-Schleifensequenz eines Virus, das nicht mit dem Eintrittshemmer behandelt werden sollte.

Referenzen

- [1] N. Pfeifer and T. Lengauer. **Improving HIV coreceptor usage prediction in the clinic using hints from next-generation sequencing data.** *Bioinformatics* 28 (18): i589-i595. 2012



Dr. Nico Pfeifer

Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik
Telefon +49 681 9325 3019
nico.pfeifer@mpi-inf.mpg.de

Integrative Krebsdatenanalyse

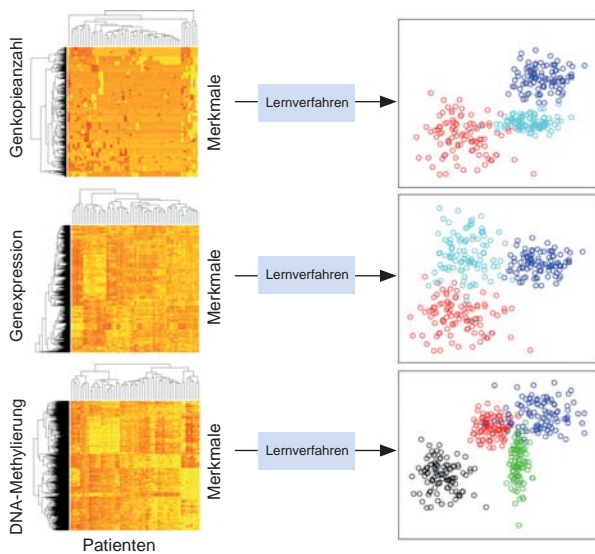


Abbildung 7 // Herkömmliche Ansätze analysieren die einzelnen Datentypen separat.

Viele Krebserkrankungen sind trotz intensiver Forschung schwierig zu heilen und verlaufen in vielen Fällen tödlich. Ein Grund hierfür ist die Heterogenität der Krankheit, d.h. dass ähnliche Symptome sehr unterschiedliche Gründe haben können. Eine personalisierte Therapie, die nicht nur das Gewebe berücksichtigt, in dem der Tumor auftritt, sondern auch speziell auf den Patienten abgestimmt ist, bietet daher erhöhte Chancen auf einen Behandlungserfolg. Die Entscheidung, welche Therapie ein Patient erhält, basiert oft auf sogenannten Krebssubtypen, d.h. ein Krebstyp wird in verschiedene Untergruppen unterteilt. Diese Untergruppen unterscheiden sich bezüglich bestimmter klinischer Merkmale, wie z. B., ob eine Therapie oder ein Medikament anschlägt. Der Patient wird dann einem dieser Subtypen zugeordnet und entsprechend behandelt. Je genauer der Subtyp die Eigenheiten des Tumors widerspiegelt, desto besser kann die Therapie darauf abgestimmt werden. Etablierte Subtypen nutzen aber meist nur einzelne Informationsquellen, wie z. B. die Genexpressionsdaten des Patienten. Tatsächlich findet man in Tumoren aber Veränderungen auf diversen unterschiedlichen Ebenen der Zelle. Diese Veränderungen beeinflussen sich oft gegenseitig, jedoch enthält keine der Ebenen alleine die gesamte Information über den Tumor. Daher wird für Krebspatienten oft eine große Menge unterschiedlicher molekularer Daten gemessen, u. a. auf genetischer sowie epigenetischer Ebene. Ein Ziel der Bioinformatik in diesem Kontext ist die Entwicklung von Methoden, die in der Lage sind, diese verschiedenen Datentypen sinnvoll miteinander zu kombinieren.

Referenzen

- [1] N. K. Speicher and N. Pfeifer. **Integrating different data types by regularized unsupervised multiple kernel learning with application to cancer subtype discovery.** *Bioinformatics* 31 (12): i268-75. 2015
- [2] A. Jalali and N. Pfeifer. **Interpretable per case weighted ensemble method for cancer associations.** *Lecture Notes in Bioinformatics: Proceedings of WABI 2014*. 2014

Ein Ansatz zur Integration dieser Daten ist rMKL-DR (regularized Multiple Kernel Learning for Dimensionality Reduction). Mittels dieser Methode werden in einem adaptiven Prozess Gewichtungen für die unterschiedlichen Datentypen gelernt, sodass ein Datentyp, der viel Information enthält, ein größeres Gewicht bekommt und daher mehr Einfluss auf das Ergebnis hat, als ein Datentyp mit weniger Information. Die integrierten Datenpunkte, in unserem Fall die Tumorproben von Patienten, werden dann in einem zweidimensionalen Raum visualisiert und weiter analysiert (vgl. Abbildungen 7 und 8). Die Distanzen zwischen den Datenpunkten reflektieren deren Ähnlichkeiten basierend auf allen verwendeten Datentypen. Im Fall von Krebspatienten können diese Ähnlichkeiten zur Identifizierung von integrierten Subtypen genutzt werden, mit deren Hilfe über eine mögliche Therapie entschieden werden kann [1].

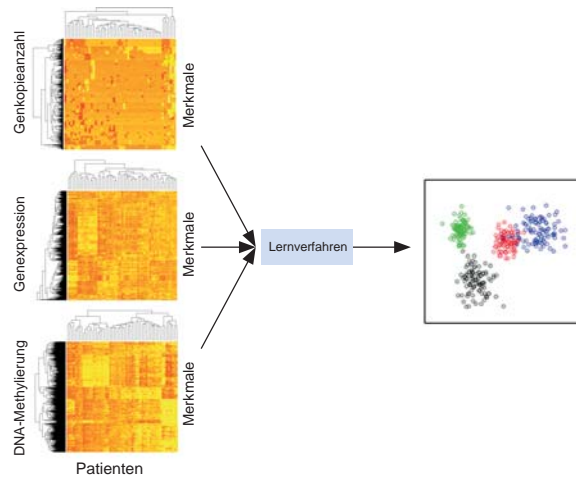


Abbildung 8 // Die integrative Analyse von mehreren Datentypen kann Informationen auffinden, die von verschiedenen Datenquellen unterstützt werden.

Konventionelle statistische Methoden sind im Allgemeinen nicht in der Lage, die molekulare Grundlage für den Tumor zu identifizieren. Darüber hinaus ist die Übertragung der gelernten Ergebnisse auf neue Patienten oft schwierig. Aus diesem Grund haben wir eine Methode entwickelt, die aus den gegebenen Daten mögliche Grundlagen für die Erkrankung extrahiert und zusätzlich einen Hinweis darauf gibt, welche Erklärung gemäß den beobachteten Mustern in den Patientendaten am wahrscheinlichsten ist. Für eine neue Probe wird vorhergesagt, welche der möglichen Erklärungen am zutreffendsten ist [2]. Diese kann Onkologen dabei helfen, die genaue Ursache des Tumors zu untersuchen sowie eine geeignete Therapie zu finden.



Nora Speicher

**Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik**
Telefon +49 681 9325-3027
nora@mpi-inf.mpg.de

Taxonomische Einteilung metagenomischer Daten

32

Bakterien bevölkern fast jedes Milieu dieser Welt, wozu auch der menschliche Körper gehört. Sie beeinflussen damit direkt oder indirekt unsere Gesundheit, die Landwirtschaft sowie die Industrie. Die Diversität der Bakterien in einem Habitat zu analysieren und ihre Funktion zu verstehen ist daher von fundamentalem Interesse. Nur ein ganz geringer Anteil der in der Natur vorkommenden Bakterien kann isoliert und im Labor kultiviert werden. Stattdessen bilden Bakterien allgemein komplexe Gemeinschaften, die fast wie ein Organismus funktionieren. Deswegen sind Studien dieser Bakteriengemeinschaften bis zur Einführung der Metagenomik-Techniken, die es erst seit kurzem gibt, sehr begrenzt gewesen.

Die Metagenomik erlaubt es Bakteriengemeinschaften zu untersuchen, ohne sie im Labor kultivieren zu müssen. Hierbei wird die DNA der Bakterien direkt aus der untersuchten Umgebung extrahiert und dann mit Tiefensequenzierung analysiert. Die DNA-Sequenzen können verwendet werden, um phylogenetische (Stammbaum-), ökologische und funktionale Analysen durchzuführen, wie zum Beispiel die Analyse der Diversität der Bakteriengemeinschaft oder die Untersuchung des enzymatischen Repertoires der verschiedenen Mitglieder der Gemeinschaft.

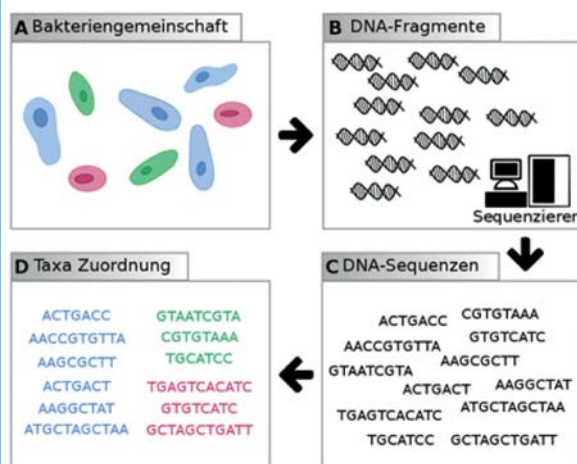


Abbildung 9 // Schematische Darstellung eines Metagenomik-Experiments bis zur Taxa-Zuordnung. (A) Eine Bakteriengemeinschaft, die untersucht werden soll. (B) Die genetische Information der gesamten Gemeinschaft wird in Form von DNA-Fragmenten gesammelt. (C) Kurze DNA-Sequenzen werden gemessen, (D) die mit bioinformatischen Methoden bearbeitet werden, um sie den einzelnen Bakterienarten zuzuordnen (Farbkodierung).

Da alle DNA-Fragmente zusammen extrahiert und sequenziert werden, ist die Zuordnung der Fragmente zu den jeweiligen Bakterien nicht direkt verfügbar. Ein wesentlicher Schritt in der metagenomischen Sequenzanalyse ist daher die Zuordnung der DNA-Fragmente zu den verschiedenen Bakteriengruppen, sogenannten Taxa. Abbildung 9 zeigt eine schematische Darstellung eines Metagenomik-Experiments bis zur Taxa-Zuordnung. Wir haben die neue Methode „PhyloPythiaS“ [1] für diese Taxa-Zuordnung entwickelt, die die Häufigkeit von kurzen Teilsequenzen, sogenannten Oligomeren, verwendet, um die Ähnlichkeit zwischen DNA-Sequenzen zu bewerten. Diese Oligomer-Häufigkeiten sind hochspezifisch für die jeweiligen Organismen. Zudem zeigen ähnliche Organismen auch ähnliche Oligomer-Häufigkeiten. Zusätzlich zu diesen Signaturen nutzt PhyloPythiaS Informationen über die bekannten taxonomischen Beziehungen der Organismen. Die Methode kann verwendet werden, um neue Sequenzen ihren entsprechenden Taxa zuzuordnen.

Die von PhyloPythiaS bestimmten Taxa sind hochspezifisch, enthalten also wenige Fehler verglichen mit anderen Methoden in diesem Bereich. Zudem kommt die Methode mit relativ kurzen Referenzsequenzen aus (ca. 10.000 Basenpaare). Bakterielle Genome sind normalerweise mehrere Millionen Basenpaare lang. Da ein Großteil der bakteriellen Genome noch nicht vollständig sequenziert ist, stehen diese nicht zur Lösung des Taxa-Zuordnungsproblems zur Verfügung, PhyloPythiaS benötigt nur kurze Sequenzen, wie sie direkt aus den metagenomischen Proben entnommen werden können. Dies ist ein entscheidender Vorteil gegenüber anderen Methoden, die längere Sequenzen benötigen. Um unsere Methode anderen Forschern zur Verfügung zu stellen, haben wir einen Webserver entwickelt, der es erlaubt, Sequenzen hochzuladen und eine Taxa-Zuordnung von PhyloPythiaS zu erhalten.

Metagenomische Studien haben bereits eine Vielzahl an neuen Organismen und Bakteriengemeinschaften aufgedeckt, zusammen mit ihren biomedizinisch und industriell relevanten Eigenschaften. Trotzdem wird erst langsam klar, welches Potential die Metagenomik hat, unser Verständnis von der mikrobiellen Welt zu revolutionieren. Bioinformatische Methoden sind essentiell, um das gesamte Potential von metagenomischen Studien zu nutzen, und PhyloPythiaS leistet einen wichtigen Beitrag zu diesem Ziel.

Prof. Dr. Alice McHardy



**Helmholtz Zentrum für Infektionsforschung
Technische Universität Braunschweig**

Telefon +49 531 6181 1430
alice.mcHardy@helmholtz-hzi.de

Dr. Kaustubh Patil



**MIT Sloan School of Management
Cambridge, Massachusetts**

Telefon +1 617 253-6399
kpatil@mit.edu

Referenzen

- [1] K. R. Patil, P. Haider, P. B. Pope, P. J. Turnbaugh, M. Morrison, T. Scheffer and A. C. McHardy. **Taxonomic metagenome sequence assignment with structured output models.** *Nat Methods* 8 (3): 191-2. 2011

Phylogenomik unter Einbeziehung paraloger Gene

Einer Studie amerikanischer Wissenschaftler aus dem Jahr 2011 zufolge gibt es auf der Erde schätzungsweise knapp neun Millionen Arten. In welchem verwandtschaftlichen Verhältnis diese stehen, untersuchen Evolutionsbiologen etwa anhand der Gene. Diese werden von Generation zu Generation an die Nachkommen weitergegeben. Manche von ihnen werden dabei immer wieder dupliziert, andere mutieren oder gehen verloren – ein Mechanismus, der mit dafür sorgt, dass stets neue Arten entstehen. Die große Herausforderung für die Wissenschaftler besteht darin, dass sie für die Rekonstruktion der gesamten evolutionären Geschichte der Arten nur auf Gene von lebenden Arten und von einigen wenigen ausgestorbenen Spezies, wie die des Neandertalers, zurückgreifen können.

Ein zentrales Element für die Rekonstruktion ist das Wissen um die verwandtschaftlichen Beziehungen der entsprechenden Gene bzw. Proteine. Wenn Gene einen gemeinsamen Genvorfahren haben (homologe Gene), können folgende Ereignisse aufgetreten sein: Es kam zu einer Duplikation – man bezeichnet die beiden Kopien dann als paraloge Gene, einer Artenbildung (orthologe Gene) oder zu einem horizontalen Gen-Transfer, d.h. zu einer Übertragung von Genen außerhalb der geschlechtlichen Fortpflanzung und über Artgrenzen hinweg (xenologe Gene).

Interessanterweise gibt es Methoden, die es ermöglichen anhand der Sequenzinformation der Gene abzuschätzen, ob sie ortholog sind oder nicht. Bislang stehen beim Erstellen von Stammbäumen nur orthologe Gene im Fokus. Sie sind als ähnliche DNA-Abschnitte in verschiedenen Arten vorhanden und gehen auf einen gemeinsamen Gen-Vorfahren zurück, aus dem neue Arten entstanden sind. Die evolutionäre Geschichte dieser orthologen Gene wird dann als die evolutionäre Geschichte der zugrunde liegenden Arten interpretiert. Diese Art der Rekonstruktion hat allerdings zur Folge, dass die Information von paralogen und xenologen Genen ignoriert wird bzw. als störendes Rauschen bewertet wird, welches es aber eigentlich zu vermeiden gilt. Dadurch wird die Menge möglicher Gene, aus denen wertvolle Information über die evolutionäre Geschichte gewonnen werden kann, deutlich eingeschränkt.

Das Team um Marc Hellmuth und seinen Leipziger Kollegen Nicolas Wieseke hat eine Methode [1] entwickelt, die beim Aufspüren von Verwandtschaftsverhältnissen zwischen Arten erstmals auch weitere genetische Informationen nutzt. Hierbei werden die Informationen von orthologen und paralogen Genen als ein Graph interpretiert, dessen Knoten die Gene und dessen Kanten die Verwandtschaftsverhältnisse der Gene widerspiegeln (Abbildung 10).

Aus dem bereinigten Graphen können nun viele kleine lokale Informationen (sogenannte Triple) des zugrunde liegenden Artenbaums erhalten werden. Die Gesamtheit dieser Triple wird dann genutzt, um die vollständige evolutionäre Geschichte der Arten zu rekonstruieren. Für die Realisierung dieses Vorhabens müssen drei kombinatorisch sehr komplexe Probleme gelöst werden. Den Wissenschaftlern ist es gelungen, Methoden zu entwickeln, die es erlauben, mittels sogenannter ganzzahliger linearer Programmierung diese Probleme exakt zu lösen.

Mit der neuen Methode kann somit eine wesentlich größere Menge an Genen verwendet und daher deutlich mehr Informationen zur präziseren Berechnung der Evolution der Arten benutzt werden. Es wurde gezeigt, dass mit diesem Algorithmus genaue Artenbäume berechnet werden können, sogar dann, wenn es zusätzliche evolutionäre Ereignisse wie einen horizontalen Gen-Transfer gab. Auch andere Wissenschaftler, wie Anthropologen oder Evolutionsforscher, könnten die Technologie künftig nutzen, um genauere Verwandtschaftsverhältnisse aufzuspüren. Neben Hellmuth und Wieseke, beide Erstautoren der Studie, waren an dieser Arbeit auch weitere Kollegen aus Leipzig und Saarbrücken sowie Forscher aus Marburg beteiligt.

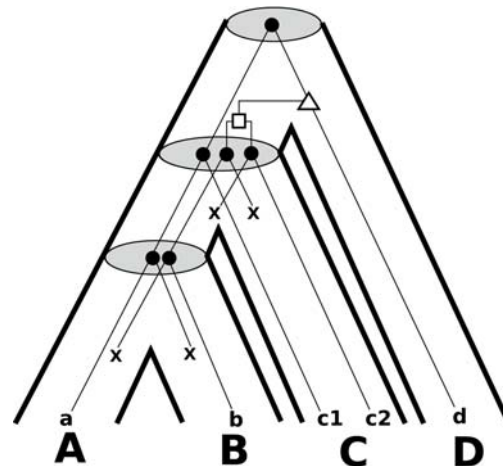


Abbildung 10 // Abgebildet ist ein Genbaum (mit beobachtbaren Genen a, b, c1, c2, d), der entlang eines Artenbaumes (mit Arten A, B, C, D) evolviert. Der Genbaum legt die verwandtschaftlichen Verhältnisse zweier Gene mittels des Event-Labelings eindeutig fest: Xenologe via HGT-Event (Dreieck), Paraloge via Duplikations-Event (Viereck), Orthologs via Speciation-Event (Kreis).



Juniorprofessor Dr. Marc Hellmuth

Institut für Mathematik und Informatik
Ernst Moritz Arndt Universität Greifswald
Telefon +49 3834 86 4636
mhellmuth@mailbox.org



Prof. Dr. Hans-Peter Lenhof

Zentrum für Bioinformatik
Universität des Saarlandes
Telefon +49 681 302 64701
lenhof@bioinf.uni-sb.de

Referenzen

- [1] M. Hellmuth, N. Wieseke, M. Lechner, H. P. Lenhof, M. Middendorf and P. F. Stadler. **Phylogenomics with paralogs**. *Proc Natl Acad Sci U S A* 112 (7): 2058-63. 2015

Biomolekulare Interaktionen

Biomolekulare Interaktionen sind die Grundlage aller physikalischen Prozesse in lebenden Organismen. In der Tat basieren nahezu alle biochemischen Reaktionen in lebenden Zellen auf molekularen Interaktionen zwischen Proteinen, Nucleinsäuren, Zuckern, Lipiden (Fetten) und niedermolekularen Liganden. Das Verständnis dieser Interaktionen ist eine unabdingbare Grundlage, um die Entwicklung von Organismen und Krankheiten, beziehungsweise die Anfälligkeit für Krankheitskeime nachvollziehen zu können. Mithilfe der Bioinformatik kann man detaillierte Strukturmodelle von molekularen Komplexen erzeugen, alternative Positionen von Wirkstoffen in der Bindungstasche eines Proteins energetisch bewerten, Bindungsprozesse modellieren und darüber hinaus einen Einblick in die theoretischen Grundlagen solcher Interaktionen liefern.

Die Forschung im Bereich der strukturellen Bioinformatik und molekularen Modellierung beruht auf der Information über die dreidimensionalen atomaren Strukturen von Biomolekülen. Solche Strukturen werden durch experimentelle Verfahren, wie die Röntgenkristallographie, Kernspinresonanzspektroskopie oder Elektronenmikroskopie, bestimmt. Manche dieser Methoden liefern statische Momentaufnahmen der Moleküle, wogegen andere Ensembles von mehreren Konformationen erzeugen. In der Realität sind die Moleküle aber in ständiger thermischer und Brownscher Bewegung, so dass zum einen ihre Konformation ständig zwischen verschiedenen Formen oszilliert und die Moleküle zudem in der dicht gepackten Umgebung der Zelle ständig aneinanderstoßen. Wir benutzen daher experimentell aufgelöste Strukturen, um diese interne Konformationsdynamik sowie die stattfindenden biophysikalischen und biochemischen Prozesse zu modellieren.

Eine ideale Methode, um dynamische biophysikalische Prozesse in atomarer Auflösung zu modellieren, sind die Moleküldynamiksimulationen. In solchen Simulationen wird jedes Atom eines Biomoleküls als Einheit aufgefasst, die durch chemische Bindungen mit einigen anderen Atomen des Systems verbunden ist. Unter dem Einfluss seiner Wechselwirkungen mit allen anderen Atomen des Systems kann seine Bewegung durch die Newtonsche Bewegungsgleichung beschrieben werden. Auf diese Weise können die konformationelle Dynamik eines Proteins sowie dessen Bindungs-Reaktionen und Komplexe mit anderen Biomolekülen modelliert werden, wobei man detaillierte Informationen über die Bewegung jedes einzelnen Atoms erhält. Somit gewinnt man einen detaillierten Einblick in alle im System ablaufenden Prozesse, sowohl mit atomarer räumlicher und hoher zeitlicher Auflösung als auch auf der Ebene von Einzelmolekülen. So etwas ist heutzutage mit experimentellen Verfahren noch nicht möglich.

Alle biomolekularen Interaktionen beruhen auf physikalischen Gesetzen. Ein molekulares System kann demzufolge unterschiedliche Zustände mit jeweils unterschiedlicher assoziierter Energie annehmen. So kann man die Energie eines Proteins vor und nach der Bindung eines Wirkstoffmoleküls vergleichen. Auf der Grundlage von molekularen Simulationen entwickeln wir theoretische Verfahren, um zu erklären, wie sich die Energie in diesen Fällen ändert und welchen Zustand das molekulare System letztendlich annehmen wird. Unsere Forschung in diesem Bereich hat viele praktische Anwendungen, wie die Optimierung von bindenden Molekülen als Basis für die Entwicklung von neuen Medikamenten, oder die Vorhersage einer Resistenz gegenüber Wirkstoffen, die infolge von Mutationen in dem Ziel-Protein auftreten kann.

Um Krankheiten zu behandeln, entwickelt man oft niedermolekulare Wirkstoffe, die dann an ein bestimmtes Protein binden, dessen Aktivität inhibieren, und so die biologischen Prozesse modulieren, die in einer von einem externen Keim, wie einem Virus oder Bakterium, infizierten Zelle stattfinden. Diese Bindung muss wirksam und spezifisch sein, sodass die anderen Prozesse, die das normale Verhalten der Zelle aufrechterhalten, nicht beeinflusst werden. Oft sind solche Wirkstoffe gegen die Proteine eines Pathogens gerichtet, denn diese kommen in einer gesunden menschlichen Zelle nicht vor. In manchen Pathogenen können aber durch Adaptationsprozesse Mutationen in den Proteinen entstehen, so dass sie weiterhin ihre für die Infektion notwendigen Funktionen ausüben können, aber nicht mehr von dem Liganden erkannt werden. Viren sind besonders effizient darin, auf diese Weise Resistenzen zu entwickeln. Wir verwenden daher Moleküldynamiksimulationen, um die Bindungsprozesse zwischen dem Ligand und dem resistenten oder dem nicht resistenten viralen Protein zu beobachten. Somit können wir die Bindungsaffinität und damit auch die Anwendbarkeit von Medikamenten in bestimmten Situationen vorhersagen, ohne dies experimentell überprüfen zu müssen.

Wirkstoffe binden an Proteine in spezifischen Bindungstaschen. Manche solcher Taschen, z. B. die aktiven Zentren von Enzymen, sind groß und durch ihre geometrische Form in der experimentell bestimmten Proteinstruktur leicht erkennbar. Andere Taschen sind kleiner und werden nur vorübergehend gebildet, wenn der passende Ligand gerade darin bindet. Wir konnten zeigen, dass geeignete Algorithmen durch die Bestimmung von Hohlräumen in den Zielproteinen solche Taschen in den Momentaufnahmen einer Moleküldynamiksimulation aufspüren können. Diese Taschen können anschließend für die Optimierung von Wirkstoffen, die an diese Proteine binden und die Proteinaktivität während einer Krankheit modulieren, verwendet werden. Um ihre biologischen Funktionen



auszuüben, bilden mehr als die Hälfte aller Proteine in der Zelle Komplexe mit anderen Proteinen. Diese Komplexe können ebenfalls entweder stabil sein oder nur zeitweilig gebildet werden. Stabile Komplexe können als biologische Fabriken angesehen werden, die bestimmte, oft aus mehreren Schritten bestehende, chemische Reaktionen ausführen.

In diesem Fall kann die Effizienz des Gesamtprozesses verbessert werden, indem man die Zwischenprodukte der Reaktionen eng beieinander hält. Zeitweilig gebildete Komplexe sind oft an Prozessen der Energieumwandlung oder der Fortleitung von chemischen Signalen beteiligt. Dabei binden die Proteine nur kurz aneinander und eine molekulare Nachricht, z. B. eine chemische Modifikation, wird übertragen. Beide Arten von Komplexen kann man bioinformatisch mit Moleküldynamiksimulationen oder mit kombinatorischen Verfahren untersuchen.

Biomolekulare Interaktionen

Energielandschaften der Proteindynamik	36
Strukturelle Aspekte der Resistenz von HIV	37
Rekonstruktion von Proteinkomplexen	38
Identifikation von Bindestellen in Proteinen	39

Energielandschaften der Proteindynamik

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG)

36

Biomoleküle, wie beispielsweise Proteine, sind mikroskopisch kleine Maschinen, die zusammenarbeiten, um das Leben eines Organismus aufrechtzuerhalten. Unser Hauptinteresse ist es daher, theoretische Modelle zu entwickeln, um die Beziehung zwischen den Strukturen und den Bewegungen von Biomolekülen sowie ihre Rolle in wichtigen biomolekularen Interaktionen, wie zum Beispiel der Interaktion zwischen Proteinen und Medikamenten, zu erklären. Gemittelte atomare dreidimensionale Strukturen von Biomolekülen können z. B. durch die Kristallstrukturanalyse gewonnen werden. Sie stellen eingefrorene Momentaufnahmen von Biomolekülen dar, die eigentlich in ständiger molekularer Bewegung sind. Die hohe Rechengeschwindigkeit moderner Computer gibt uns aber darüber hinaus auch die Möglichkeit, ausgehend von den Protein-Kristallstrukturen die atomaren Bewegungen von Biomolekülen zu simulieren und dadurch den Mechanismus ihrer Funktionsweise zu verstehen.

Zusammenlagerungen mehrerer Proteine zu spezifischen Proteinkomplexen sind essentielle elementare Schritte in praktisch allen zellulären Prozesse. In einer Zelle werden etwa 30% des Zellvolumens von Proteinen eingenommen. Da alle Proteine in steter Bewegung sind („Brownsche Molekularbewegung“), kollidieren daher ständig irgendwelche Proteine in der Zelle miteinander. In der großen Mehrzahl der Fälle sind diese Wechselwirkungen nur von sehr kurzer Dauer. Manchmal führen solche Kollisionen allerdings zur Ausbildung spezifischer Proteinkontakte. Eine wichtige grundlegende Fragestellung ist daher, wie sich spezifische Proteinkontakte von unspezifischen Proteinkontakten unterscheiden. Computersimulationen, wie Moleküldynamik-Simulationen (MD), sind ideal geeignet, um solche Fragestellungen in hohem Detail zu beantworten.

Erstmalig haben wir daher in den letzten Jahren die Assoziation von gut wasserlöslichen, „hydrophilen“ Proteinen mit Hilfe von MD-Simulationen systematisch untersucht. Überraschenderweise beobachteten wir, dass sich die Bestandteile von spezifischen Komplexen bei ihrer Assoziation spontan in die korrekte Orientierung drehen. Man kann sich dies wie ein Raumschiff vorstellen, das auf dem Mond landen soll, und dabei automatisch seine Füße der Mondoberfläche zuwendet. Eine letzte, störende Rolle könnte dabei dem Wasser zwischen den beiden Bindungsschnittstellen der Proteine zufallen, welches den Proteinen quasi „im Weg steht“. Es zeigte sich jedoch, dass diese Wassermoleküle entweder bei der Bindung von hydrophoben Bindungsschnittstellen in einen gasähnlichen Zustand übergehen und somit kein Hin-

dernis darstellen oder bei der Bindung von elektrostatisch polaren Bindungsstellen eine günstige Orientierung einnehmen, welche deren Assoziation begünstigt [1]. Wir zeigten daher, dass solche Assoziationsprozesse im Allgemeinen barrierefrei sind. Wenn die Bindungspartner allerdings in verdrehten Orientierungen aneinanderstoßen, sind die involvierten Kontaktflächen kleiner und die Attraktion der beiden Proteine schwächer. Dieses Modell erklärt, weshalb spezifische Kontakte längere Lebensdauern als unspezifische Proteinkontakte haben. Zukünftig möchten wir diese Untersuchungen auf größere Proteinkomplexe erweitern, die aus mehr als zwei Bindungspartner bestehen, sowie die Rolle von Proteinmutationen berücksichtigen, mit denen die Interaktion der Proteine modifiziert werden kann. Experimentell ermittelbare thermodynamische Größen, wie zum Beispiel Enthalpie, Entropie und freie Enthalpie, dienen dazu, biologische Prozesse, wie die Bindung von Medikament und Protein, sowie deren Beeinflussung durch Veränderungen in der Aminosäuresequenz (Mutation) zu quantifizieren. Die freie Enthalpie ist die Summe der molekularen Interaktionen, die durch die Enthalpie beschrieben werden, und der molekularen Unordnung, die durch Entropie beschrieben wird. Das Verständnis der Mechanismen, in denen Bewegungen der Proteine Änderungen in der freien Enthalpie auslösen, ist entscheidend für viele Bereiche, wie zum Beispiel die Medikamentenentwicklung. Bei biomolekularen Prozessen beobachtet man oft das Phänomen, dass Änderungen eines molekularen Systems zu Änderungen in Enthalpie und Entropie führen, die sich zum großen Teil aufheben. Das heißt, dass die einzelnen Komponenten Enthalpie und Entropie recht großen Änderungen unterworfen sind, während die daraus resultierende freie Enthalpie sich nur vergleichsweise wenig ändert. Wir haben eine allgemeine Theorie entwickelt, die dieses Phänomen erklärt [2]. Sie greift auf die Informationstheorie zurück. Die Einfachheit und Allgemeinheit der Konzepte aus der Informationstheorie ermöglichen uns eine unmittelbare Interpretation der physikalischen Zusammenhänge. Die Anwendung der Informationstheorie in molekularen Simulationen ist ein vielversprechender Ansatz, um bioinformatische Methoden zur Berechnung von Veränderungen in der freien Enthalpie zu verbessern. Eine wichtige Rolle spielen diese Methoden zum Beispiel bei der Vorhersage von Resistenzen gegen bestimmte Medikamente, die durch Mutationen in der Aminosäuresequenz von Proteinen ausgelöst werden.

Referenzen

- [1] M. Ahmad, W. Gu, T. Geyer and V. Helms. **Adhesive water networks facilitate binding of protein interfaces**. *Nat Commun* 2: 261. 2011
- [2] M. Ahmad, V. Helms, T. Lengauer and O. V. Kalinina. **Enthalpy-Entropy Compensation upon Molecular Conformational Changes**. *J Chem Theory Comput* 11 (4): 1410-8. 2015



Dr. Mazen Ahmad

**Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik**

Telefon +49 681 9325-3005
mahmad@mpi-inf.mpg.de



Prof. Dr. Volkhard Helms

**Zentrum für Bioinformatik
Universität des Saarlandes**

Telefon +49 681 302 70701
volkhard.helms@bioinformatik.uni-saarland.de

Strukturelle Aspekte der Resistenz von HIV

Das Humane Immundefizienz-Virus (HIV) verursacht die Krankheit AIDS, die unheilbar ist und für die es auch keinen Impfstoff gibt. In vielen Fällen kann die Vervielfältigung des Virus jedoch durch antiretrovirale Medikamente unterdrückt werden, was die Symptome lindert und den Krankheitsverlauf bremst. Eines der relevantesten Angriffsziele für die antiretrovirale Therapie ist die HIV-Protease. Hierbei handelt es sich um ein virales Protein, das für die Virusreifung, d.h. den Prozess, bei dem das Viruspartikel vollständig infektiös wird, erforderlich ist. Durch die hohe Mutationsrate von HIV entstehen Viruspartikel, die neue Eigenschaften haben und somit resistent gegen bestimmte antivirale Therapien sind. Das kann dazu führen, dass man die medikamentöse Therapie ändern muss. Obwohl sich die meisten wichtigen Mutationsstellen der HIV-Protease an der Stelle im Protein befinden, an der das Medikament (Inhibitor) bindet, existieren auch Mutationen, die weit von dieser Bindestelle entfernt sind. Wenn man sich mit viraler Resistenz beschäftigt, ist es äußerst wichtig zu verstehen, wie diese Mutationen die Bindung der Medikamente beeinflussen.

Experimentell bestimmte dreidimensionale Strukturen des Protease-Inhibitor-Komplexes, gewonnen durch die Kristallstrukturanalyse, stellen statische Information über die Konformationen dieser Moleküle und ihrer Interaktionen zur Verfügung. Die Molekulardynamik, eine Simulationstechnik für Moleküle, kann darüber hinaus verwendet werden, um einen Überblick über relevante Konformationen und ihre Dynamik mit der Kristallstruktur als Ausgangspunkt zu erhalten. Nutzt man diesen Ansatz zusammen mit Methoden der statistischen Mechanik für die thermodynamische Analyse, ist es möglich, Unterschiede in der Bindungsenergie zwischen Inhibitor und Protein zu berechnen. Die freie Energie beschreibt den Zustand eines Systems von Molekülen und bestimmt auch die Wahrscheinlichkeit, mit der eine Konformation auftritt. Daher spielt diese Größe eine wichtige Rolle in molekularen Systemen.

Resistenzmutation L76V

Die Mutation in der Protease-Region des HIV-Genoms, die durch den Austausch der Aminosäure Leucin zu Valin an Position 76 (L76V) gekennzeichnet ist, ist eine wichtige Resistenzmutation. Sie macht das Enzym resistent gegen die Inhibition durch manche Medikamente, aber auf der anderen Seite wiederum anfällig für andere Medikamente. Die Aminosäure liegt dabei weit von der Bindestelle des Proteins entfernt. Wir

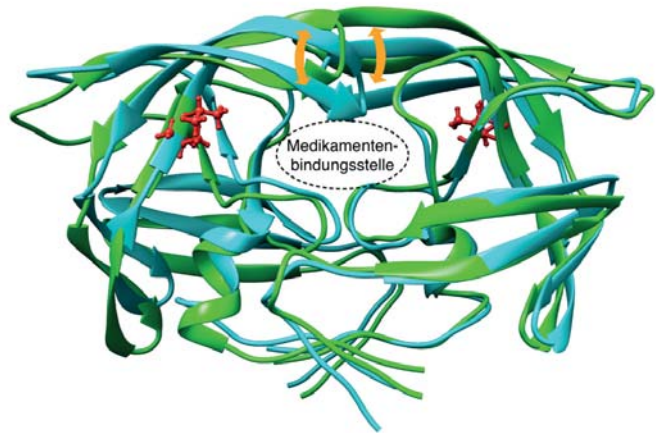


Abbildung 11 // Beispiel für Konformationsunterschiede zwischen der Wildtyp-Protease (grün) und mutierter Protease (blau). Mutierte Aminosäuren sind rot dargestellt. Die orangefarbenen Pfeile deuten die Bewegungsrichtung des Moleküls beim Bindevorgang an.

nutzten daher molekulardynamische Simulationen von Protease-Sequenzen aus Patientenproben, die entweder Leucin oder Valin an Stelle 76 enthalten und testen unterschiedlichen Inhibitoren. Unterschiede in der freien Bindungsenergie zwischen der Wildtyp-Protease (76L) und der mutierten Protease (76V) bei einer Bindung an den Inhibitor wurden dann mittels thermodynamischer Berechnungen bestimmt, um die Stabilität beider Komplexe vergleichen zu können. Der Vergleich von Wildtyp und mutiertem Protein-Inhibitor-Komplex zeigt Unterschiede in der Affinität des Inhibitors zu verschiedenen Proteinvarianten, die das Verständnis dieser speziellen Resistenzbildung fördern. Simulationen der Protease ohne bindende Medikamente offenbarten darüber hinaus Unterschiede zwischen den Konformationen des Wildtyp-Proteins und des mutierten Proteins (Abbildung 11). Weil die ungebundene Protease mit ihren natürlichen Substraten interagiert, gibt die Analyse auch einen Einblick in die allgemeine Evolution dieses Proteins. Die Analyse zeigt darüber hinaus den Effekt von Mutationen auf, die weit entfernt von der Protein-Ligandenschnittstelle sind, und gibt damit Anstöße für die Entwicklung von Medikamenten.

Referenzen

- [1] T. Bastys, N. T. Doncheva, H. Walter, R. Kaiser, M. Albrecht and O. V. Kalinina. **Molecular Mechanisms of HIV-1 Protease Resistance and Resensitization towards Lopinavir and Saquinavir upon L76V Mutation**. *International Workshop on Antiviral Drug Resistance: Meeting the Global Challenge*. 2014



Dr. Olga Kalinina

Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik
 Telefon +49 681 9325 3004
 kalinina@mpi-inf.mpg.de



Tomas Bastys

Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik
 Telefon +49 681 9325 3029
 tbastys@mpi-inf.mpg.de

Rekonstruktion von Proteinkomplexen

Proteine sind molekulare Maschinen, die viele Prozesse in lebenden Organismen umsetzen. Dafür müssen die Proteine oft große Komplexe formen, die aus einer bestimmten Anzahl von gleichen oder unterschiedlichen Proteinen bestehen. Ungefähr die Hälfte aller Proteine in der Zelle ist dabei ein Bestandteil großer stabiler Proteinkomplexe. Auf der anderen Seite sind die kurzlebenden Paarinteraktionen der Proteine ein wichtiger Bestandteil von bioenergetischen Prozessen und Signalwegen. Die Proteineinheiten finden und binden sich in der Natur aufgrund von gegenseitigen chemischen Wechselwirkungen. Diese Wechselwirkungen sind allerdings nicht einfach zu modellieren, und oft fehlen auch die experimentellen Daten dazu. Wir haben daher eine kombinatorische Methode zur Modellierung von Proteinkomplexen entwickelt, die 3D-MOSAIC genannt wird.

Man kann einen Proteinkomplex als ein Puzzle-Spiel betrachten: Jedes Element passt nur dann gut zu den anderen Teilen, wenn es im Gesamtbild korrekt platziert wird. Entsprechend interagieren in einem Proteinkomplex die Proteineinheiten am besten mit ihren Nachbarn, wenn sie korrekt platziert sind. In 3D-MOSAIC wird ein Proteinkomplex schrittweise rekonstruiert, indem neue Einheiten an den wachsenden Komplex angefügt werden. Dabei werden Proteineinheiten bevorzugt, die mit möglichst vielen anderen Proteinen im Komplex interagieren.

Wir haben 3D-MOSAIC getestet, indem wir das Programm experimentell aufgelöste Proteinkomplexstrukturen mit mehr als sechs Untereinheiten rekonstruieren ließen. Die Erfolgsrate lag dabei bei 81,8%. Zu den bemerkenswerten Beispielen gehört die erfolgreiche Rekonstruktion des 20S Proteasom-Komplexes, des Transmembran-Komplexes des Cytochroms BC1 und auch die von Kapsiden des Satelliten-Panicum-Mosaik-Virus und des Satelliten-Tabak-Nekrose-Virus (siehe Abbildung 12).

Die Expression von Genen wird molekular durch eine logische Verschaltung regulatorischer Signale mit verschiedenen Faktoren kontrolliert. Eine besondere Rolle kommt hier der Bildung von Proteinkomplexen zwischen verschiedenen Transkriptionsfaktoren und anderen regulatorischen Proteinen zu. Interessanterweise sind die Bindungsmotive für solche Komplexe zwischen Maus und Mensch viel besser konserviert, als die Bindungsmotive für einzelne Transkriptionsfaktoren.

Zur Vorhersage solcher Komplexe haben wir daher den kombinatorischen Algorithmus DACO entwickelt [2]. Da Proteine meist aus mehreren strukturellen Einheiten, den soge-

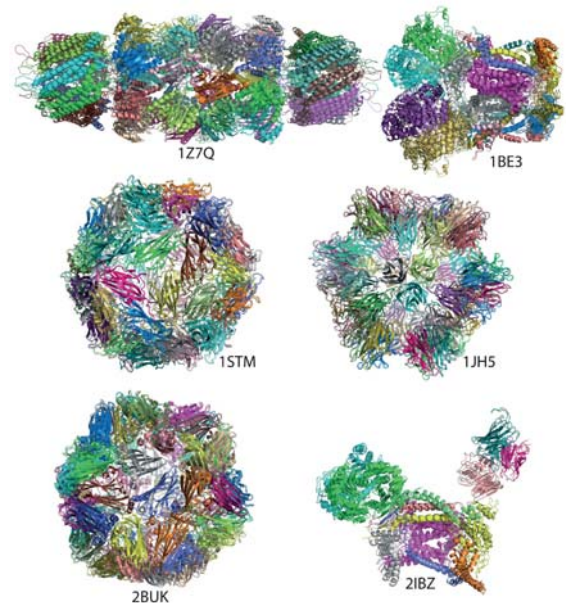


Abbildung 12 // Erfolgreich rekonstruierte Proteinkomplexe, über die entsprechenden Strukturen der Referenzkomplexe gelegt.

nannten Proteindomänen, bestehen, verwendet DACO genomweite, experimentelle Daten über Proteininteraktionen eines Organismus sowie bekannte Interaktionen zwischen Proteindomänen. Wir nehmen an, dass jede Proteindomäne nur eine Interaktion mit einem Bindungspartner eingehen kann. Durch diese Darstellung ist DACO erstmalig in der Lage, eine kombinatorische Vielfalt an alternativ möglichen Proteinkomplexen zu modellieren.

Für die Bäckerhefe *Saccharomyces cerevisiae*, die häufig als Modellsystem für Proteinkomplexe dient, lieferte DACO im Vergleich mit existierenden Algorithmen eine viele größere und zugleich viel genauere Anzahl an Proteinkomplexen, die mehrere Transkriptionsfaktoren enthalten. Die durch solche Komplexe regulierten Zielgene haben dadurch meist wesentlich ähnlichere zelluläre Funktionen als die von einzelnen Transkriptionsfaktoren regulierten Zielgene. DACO liefert somit einen wichtigen Beitrag zur Aufklärung von genregulatorischen Mechanismen in zellulären Systemen. Wir wenden DACO nun auch dazu an, Änderungen in den kombinatorischen Regulationsprozessen zwischen verschiedenen Entwicklungsstadien von menschlichen Zellen sowie bei der Entstehung von Krebs zu entschlüsseln.

Referenzen

- [1] M. Dietzen, O. V. Kalinina, K. Taškova, B. Kneissl, A. K. Hildebrandt, E. Jaenicke, H. Decker, T. Lengauer and A. Hildebrandt. **Large oligomeric complex structures can be computationally assembled by efficiently combining docked interfaces.** *Proteins* 83 (10): 1887-99. 2015
- [2] T. Will and V. Helms. **Identifying transcription factor complexes and their roles.** *Bioinformatics* 30 (17): i415-21. 2014



Dr. Olga Kalinina

**Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik**

Telefon +49 681 9325 3004
kalinina@mpi-inf.mpg.de



Prof. Dr. Volkhard Helms

**Zentrum für Bioinformatik
Universität des Saarlandes**

Telefon +49 681 302 70701
volkhard.helms@bioinformatik.uni-saarland.de

Identifikation von Bindestellen in Proteinen

Viele wichtige zelluläre Prozesse beruhen auf der Wechselwirkung von Proteinen mit kleinen chemischen Molekülen. Dies können Ausgangsstoffe, hemmende Moleküle von enzymatischen Reaktionen oder sogenannte allosterische Effektoren sein. Hemmende Moleküle von enzymatischen Reaktionen binden in vorgeformten Bindungstaschen am aktiven Zentrum, allosterische Effektoren binden in vorgeformten Bindungstaschen an anderen Positionen auf den Proteinoberflächen und beeinflussen auf diese Weise die Funktion des Enzyms negativ. Eine zentrale Rolle spielen dabei stets Bindungstaschen von geeigneter Größe und Beschaffenheit.

Mit Hilfe von geeigneten Computeralgorithmen können solche Bindungstaschen für kleine Ligandmoleküle auf den Proteinoberflächen charakterisiert werden. Dieser Schritt spielt daher seit Jahren eine wichtige Rolle in der struktur-basierten Wirkstoffforschung. Im Idealfall würde eine solche Bindungstasche das Ligandmolekül vollständig aufnehmen können. Dies ist jedoch oft nicht der Fall. So fanden wir in einer Studie heraus, dass in 25-50% aller Fälle der Ligand mit mehr als einer gefundenen Tasche überlappt, wobei die unterschiedlichen Zahlen aus der Verwendung unterschiedlicher Algorithmen resultieren.

Wir stellten uns daher die Frage, ob man nicht in solchen Fällen durch das Zusammenfassen benachbarter Taschen günstigere, größere Taschen bilden könnte, die dann die entsprechenden Liganden voll abdecken. Wir zeigten, dass sowohl durch hierarchisches Clustering als auch durch eine auf maximalem Fluss basierende Methode aus den Ausgangstaschen solche fusionierte Taschen konstruiert werden können. Durch Parameteroptimierung erzielten wir eine optimale Balance zwischen maximaler Abdeckung der Liganden und einem möglichst geringen zusätzlichen unbesetzten Volumen der fusionierten Taschen.

Im Gegensatz zu Enzymen, die ihre natürlichen Liganden in tiefen, vorgeformten Taschen am aktiven Zentrum binden, sind die Bindungsschnittstellen von Protein-Protein-Komplexen üblicherweise sowohl im ungebundenen Zustand als auch im gebundenen Komplex ziemlich plan. Dies bewirkt, dass wenige Konturen existieren, an die ein Ligand spezifisch binden könnte. Durch die Röntgenkristallographie konnte aber gezeigt werden, dass kleine Moleküle manchmal die Öffnung geeigneter Bindungstaschen an solchen Bindungsschnittstellen induzieren können und durch ihre Bindung die Bildung von Proteinkomplexen kompetitiv verhindern können. Wenn man diesen Effekt durch einen Computer-Algorithmus nachvollziehen könnte, würde dies sehr interessante Möglichkeiten für die Entwicklung der Inhibitoren von Protein-Protein-Komplexbildung eröffnen.

In diesem Zusammenhang beobachteten wir in Molekulardynamiksimulationen von wassergelösten Proteinen im ungebundenen Zustand, dass sich fast überall auf ihren Oberflächen Taschen öffnen und schließen, die man mit Algorithmen finden kann [1]. Diese zeitweilig auftretenden Taschen werden vermutlich durch die enge Kopplung der Proteindynamik mit der Dynamik des Lösungsmittels induziert und existieren meist nur für Bruchteile von Nanosekunden (Abbildung 13). Für mehrere wichtige, am Zelltod beteiligte Proteine, die unter anderem an das als „Wächter der Zelle“ bekannte p53-Protein binden, konnten wir zeigen, dass diese zeitweiligen Taschen günstige Abmessungen und physikochemische Parameter aufweisen, um darin bekannte kleine Inhibitormoleküle zu binden. Diese Arbeiten öffnen einen neuen Zugang zur Entwicklung von Inhibitoren von Protein-Protein-Interaktionen, der nun von einer Reihe von Forschergruppen im akademischen Sektor und in der Industrie weiterverfolgt wird.

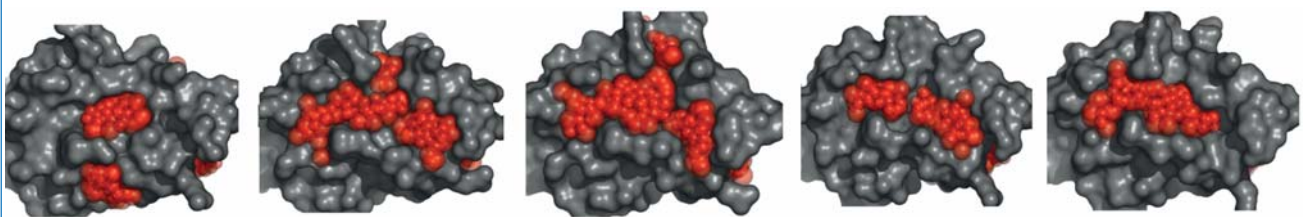


Abbildung 13 // Proteinoberflächen ähneln Flüssigkeiten. Als Beispiel ist hier die Oberfläche des Proteins MDM2 in einer Molekulardynamiksimulation gezeigt. Die 5 Schnappschüsse des Proteins sind jeweils 100 Pikosekunden voneinander getrennt (eine Pikosekunde ist ein Millionstel einer Millionstel-Sekunde). Benachbarte rote Kugeln kennzeichnen zeitweilige Taschen auf der Proteinoberfläche.

Referenzen

- [1] S. Eyrich, J. L. Medina-Franco and V. Helms. **Transient pockets on XIAP-BIR2: towards the characterization of putative binding sites of small-molecule XIAP inhibitors.** *Journal of Molecular Modeling*, Vol. 18, p. 2031-2042. 2012



Prof. Dr. Volkhard Helms

**Zentrum für Bioinformatik
Universität des Saarlandes**

Telefon +49 681 302 70701

volkhard.helms@bioinformatik.uni-saarland.de

Biologische Netzwerke

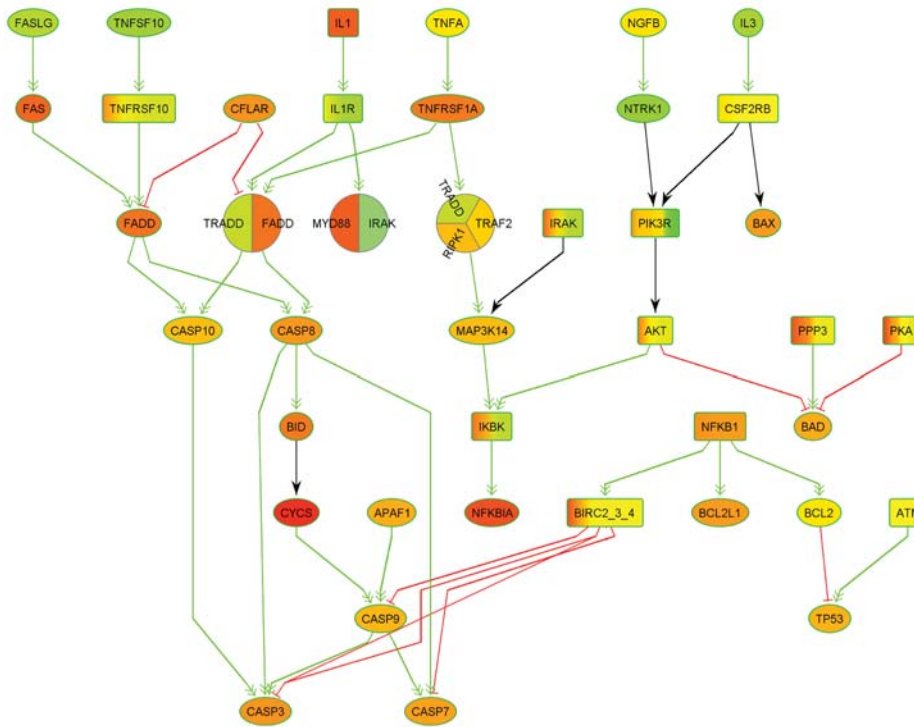
Der Stoffwechsel aller Lebewesen basiert auf komplexen biochemischen Interaktionen und Reaktionen hunderttausender von Molekülen, darunter Proteine, DNA- und RNA-Moleküle, Fette, und Zuckermoleküle. Eine zentrale Rolle als Informationsträger spielen die Genome und hier vor allem gewisse Abschnitte in den Genomen, die sogenannten Gene, welche die Information für die Synthese von Proteinen und nicht-Protein-kodierenden RNA-Molekülen speichern. Dank der rasanten wissenschaftlichen Fortschritte in der Biotechnologie und der Bioinformatik hat sich in den letzten Jahrzehnten unser Wissen über Stoffwechselprozesse so vermehrt, dass allein die Aufbereitung und Speicherung dieses Wissens eine große Herausforderung darstellt: Sie ist nur noch mittels leistungsfähiger bioinformatischer Datenbanken zu bewältigen. Die entsprechenden Daten bilden auch die Grundlage für Simulationen der Stoffwechselprozesse, welche nicht nur neue Erkenntnisse über natürliche biologische Prozesse, sondern auch über die Ursachen von Krankheiten und über pathogene Mechanismen liefern. Diese Erkenntnisse können wiederum zu neuen diagnostischen und therapeutischen Ansätzen führen. Bei entsprechenden Simulationen modelliert man die Stoffwechselprozesse meist mittels Graphen, die auch als biologische Netzwerke bezeichnet werden. Hierbei repräsentieren die Knoten der Netzwerke die beteiligten Moleküle und die Kanten die biochemischen Interaktionen zwischen den Molekülen.

Bei genetisch bedingten Erkrankungen verursachen die vorliegenden genetischen Aberrationen Aktivitätsänderungen in Netzwerken, die zu Störungen in der molekularen Informationsübertragung führen. Die modifizierten Komponenten bezeichnet man auch als deregulierte Prozesse oder Netzwerke. Ein Schwerpunkt der Forschung des ZBI ist die Entwicklung neuer Verfahren für die Detektion deregulierter Prozesse und Subnetzwerke. Da für viele biologische Prozesse zwar die beteiligten Moleküle, aber (noch) nicht die Interaktionen (Topologie des Netzwerks) zwischen den Molekülen bekannt sind, entwickelt die Bioinformatik statistische Verfahren zur Detektion deregulierter Prozesse, welche die Topologie nicht benötigen. Obwohl diese Ansätze mittlerweile nicht nur auf Gene angewendet werden, bezeichnet man sie aufgrund ihres ersten Anwendungsgebietes heute immer noch als Gene-Set-Analyse-Methoden (GSA). Das ZBI hat international beachtete wissenschaftliche Beiträge zur Entwicklung von GSA-Methoden und GSA-Software geleistet (siehe Seite 42). Unter anderem bietet das ZBI einen Webservice namens GeneTrail an, der von vielen Gruppen weltweit genutzt wird und dessen Einsatzgebiet den gesamten Bereich der Lebenswissenschaften umfasst. Die aktuelle Version GeneTrail2 ist zurzeit der leistungsfähigste Webservice für GSA-Analysen weltweit mit dem umfangreichsten Angebot an Methoden.

Eine der wichtigsten Klassen von biologischen Netzwerken sind genregulatorische Netzwerke und Signalkaskaden, die modellhaft den molekularen Informationsfluss in einer Zelle oder auch zwischen Zellen aufzeigen. Auch im Bereich der Entwicklung bioinformatischer Methoden für die Simulation und Analyse von regulatorischen Netzwerken haben Gruppen des ZBI international beachtete wissenschaftliche Beiträge geleistet (siehe Seite 43). Unter anderem erhielt Frau Prof. Verena Wolf für die Entwicklung höchst innovativer Simulationsverfahren im Dezember 2013 beim Nachwuchswettbewerb des Magazins *Technology Review* die Auszeichnung „Innovatoren unter 35“. Diese neuen Verfahren erlauben eine enorme Beschleunigung der Simulationen, so dass realistische Modelle simuliert werden können, die bisher kaum oder gar nicht simuliert werden konnten.

Eine weitere wichtige Klasse von biologischen Netzwerken sind Protein-Protein-Interaktionsnetzwerke (PPIs), deren Knoten Proteine und deren Kanten Wechselwirkungen zwischen den Proteinen repräsentieren. Eine besonders wichtige Klasse von Proteinen sind Transkriptionsfaktoren. Sie regulieren die Produktion von Proteinen. Transkriptionsfaktoren bilden eine große Zahl von Protein-Komplexen, wobei jeder Komplex die Transkription einer anderen Gruppe von Genen steuert. Die Arbeitsgruppe Helms forscht hier an der Schnittstelle zwischen PPI- und genregulatorischen Netzwerken und hat einen sehr innovativen Ansatz entwickelt, um zu bestimmen, welche Komplexe von Transkriptionsfaktoren in vorgegebenen Zelltypen die Transkription von bestimmten Genen oder Gruppen von Genen regulieren (siehe Seite 44). Der Ansatz wurde mit einem „Best Paper Award“ der europäischen Bioinformatikkonferenz ECCB 2014 ausgezeichnet.

Die dritte große Klasse von biologischen Netzwerken sind metabolische Netzwerke. Hier repräsentieren die Knoten kleinere Moleküle, sogenannte Metabolite, und die gerichteten Kanten Enzyme oder Familien von Enzymen, welche die Umsetzung der Metaboliten katalysieren. Am Institut für Systembiotechnologie der UdS erforschen die Wissenschaftler das komplexe Innenleben von Mikroorganismen. In den letzten fünfzehn Jahren haben sie an der Universität des Saarlandes, und zuvor an der Universität Münster und der TU Braunschweig, international führend, systembiologische Methoden entwickelt, mit denen sie die zellulären Vorgänge in Mikroorganismen erfassen und den „molekularen Verkehr“ durch das komplexe Stoffwechselnetzwerk der Zellen verfolgen können. Ergänzend nutzen die Forscher metabolische Netzwerkmodelle, um Voraussagen über ihre Stoffwechselleistungen zu treffen (siehe Seite 45). Hierbei ist das Ziel, mit Hilfe mathematischer und bioinformatischer Methoden ein grundlegendes Verständnis der zugrundeliegenden Stoffwechselfvorgänge zu gewinnen und die richtigen Stellschrauben zur gezielten Veränderung der Stoffwechselleistung zu finden.



Ausschnitt aus einem regulatorischen Netzwerk

Biologische Netzwerke

Analyse biologischer Prozesse	42
Simulation stochastischer Reaktionsnetzwerke	43
Wenn Dynamik wichtig wird: Regulatorische Netzwerke für Zeitreihen	44
Identifizierung deregulierter Signalkaskaden	45
Funktionelle Einheiten in Proteinnetzwerken	46
Metabolische Netze von Mikroorganismen	47

Analyse biologischer Prozesse

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) (Schwerpunktprogramm 1335: Scalable Visual Analytics)

42

Vielen biologischen Prozessen konnte man bereits die Gene oder Proteine zuordnen, die an diesen Prozessen beteiligt sind. Jedoch ist oft noch nicht bekannt, wie die beteiligten Gene genau miteinander interagieren, d.h. die Topologie des entsprechenden biologischen Netzwerkes ist noch nicht oder nur teilweise erschlossen. Unter dem Oberbegriff „biologische Kategorien“ fasst man alle biologischen Prozesse mit bekannter oder unbekannter Topologie zusammen. Bioinformatische Verfahren zur Bestimmung deregulierter Prozesse, die auf alle bekannten biologischen Kategorien angewendet werden können, fasst man unter dem Begriff „Gene Set Analysis“ (GSA) zusammen. Ein Schwerpunkt der Forschung des ZBLs war und ist die Entwicklung entsprechender neuer leistungsfähiger GSA-Methoden und -Software. Unter anderem hat das Zentrum Webserver für die Analyse der Anreicherung von Genen (GeneTrail) und MicroRNAs (MirTrail, MIEAA) entwickelt und implementiert, die von vielen Gruppen weltweit genutzt werden. Die Abbildung 14 zeigt eine Auswertung der Nutzerstatistik (links weltweit und rechts Nordamerika) von GeneTrail für die Jahre 2013 – 2015.

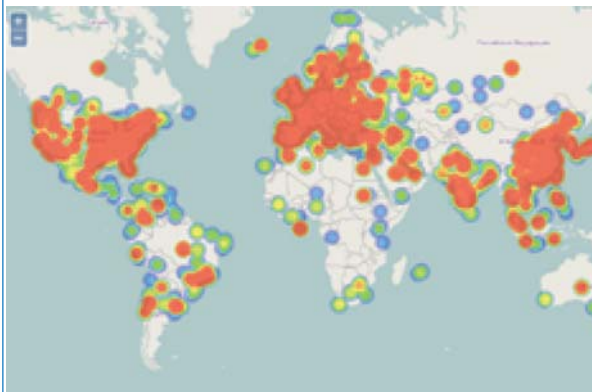
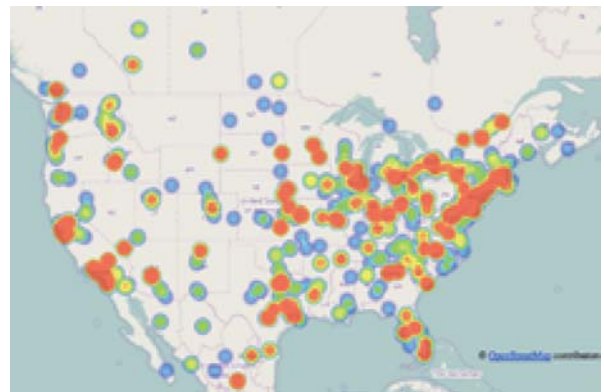


Abbildung 14 // Nutzerstatistik: rot = häufige Nutzung des GeneTrail-Servers, blau = seltene Nutzung des GeneTrail-Servers.

Insgesamt wurden seit der Veröffentlichung mehr als 45.000 Datenanalysen mit GeneTrail durchgeführt. Ein Blick auf die Veröffentlichungen, die GeneTrail referenzieren, zeigt, dass das Einsatzgebiet den gesamten Bereich der Lebenswissenschaften von der Medizin über die Genetik bis hin zur Pflanzenbiologie überdeckt. Betrachtet man nur die Veröffentlichungen im Bereich der Medizin, so stellt man fest, dass GeneTrail zur Untersuchung und Aufklärung von pathogenen Mechanismen vieler bekannter Krankheiten, insbesondere einer Reihe von Krebsarten, eingesetzt wurde [1, 2].

Basierend auf diesen positiven Erfahrungen haben wir die Funktionalität von GeneTrail weiterentwickelt und an neue Einsatzgebiete angepasst. GeneTrail2, das seit 2015 verfügbar ist, bietet in allen Bereichen noch reichhaltigere Funktionalität und gehört zurzeit zu den leistungsfähigsten Software-Paketen für GSA-Studien weltweit. Ein Alleinstellungsmerkmal ist dabei die integrierte Analyse von Genom-, Transkriptom-, Metabolom- und Proteom-Daten. Verglichen mit anderen Programmen bietet GeneTrail2 eine umfangreichere Funktionalität hinsichtlich statistischer Verfahren an und erlaubt es Nutzern, allein für die Spezies Mensch über 46.000 Kategorien aus mehr als 30 Datenbanken (darunter KEGG, Reactome, GO, WikiPathways, DrugBank, Pfam, mirWalk, mirDB etc.) zu analysieren.



Referenzen

- [1] A. Keller, P. Leidinger, A. Bauer, A. Elsharawy, et al. **Toward the blood-borne miRNome of human diseases.** *Nat Methods* 8 (10): 841-3. 2011
- [2] D. Li, C. Beisswenger, C. Herr, J. Hellberg, G. Han, T. Zakharkina, M. Voss, R. Wiewrodt, R. M. Bohle, M. D. Menger, R. M. Schmid, D. Stockel, H. P. Lenhof and R. Bals. **Myeloid cell RelA/p65 promotes lung cancer proliferation through Wnt/beta-catenin signaling in murine and human tumor cells.** *Oncogene* 33 (10): 1239-48. 2014



Prof. Dr. Hans-Peter Lenhof

**Zentrum für Bioinformatik
Universität des Saarlandes**
Telefon +49 681 302 64701
lenhof@bioinf.uni-sb.de



Prof. Dr. Andreas Keller

**Klinische Bioinformatik
Universität des Saarlandes**
Telefon +49 681 302 68611
andreas.keller@ccb.uni-saarland.de

Simulation stochastischer Reaktionsnetzwerke

Gefördert durch „Cluster of Excellence: Multimodal Computing and Interaction“

In diesem Projekt geht es darum, komplexe zelluläre Prozesse durch Computersimulationen darzustellen. Solche Simulationen können teure und zeitaufwändige Laborexperimente ersetzen oder auch im Zusammenspiel mit Laborexperimenten Aussagen über solche Teile eines Systems machen, die im Labor nicht direkt beobachtbar bzw. messbar sind. Denn mit neuartigen Algorithmen können nun bislang extrem zeitaufwendige oder unmögliche Simulationen effizient durchgeführt werden. Ein besonderer Schwerpunkt liegt dabei auf der Simulation genregulatorischer Netzwerke, bei denen das Zusammenspiel zwischen Genen und Proteinen untersucht wird. Die Erforschung der funktionalen Zusammenhänge und Mechanismen in solchen Netzwerken liefert einen wichtigen Beitrag zum Verstehen und zur Entwicklung von erfolgreichen Therapien für Krankheiten, wie zum Beispiel Krebs.

Im Gegensatz zu früheren Modellen verfolgt die Gruppe um Prof. Wolf einen sogenannten hybriden Ansatz, bei dem nur solche Teile des Modells sehr detailliert beschrieben werden, die einen wesentlichen Einfluss auf die gesamte Dynamik haben, während bei anderen Teilen des Modells gröber abstrahiert wird. Zum Beispiel wird der aktuelle Zustand eines Gens (aktiv oder inaktiv) durch eine Wahrscheinlichkeitsverteilung beschrieben, während bei Proteinkonzentrationen nur der Mittelwert über die Zeit integriert wird. Dies erlaubt eine enorme Beschleunigung der Simulationsverfahren und ermöglicht somit die Simulation von realistischen Modellen, die bisher kaum oder gar nicht simuliert werden konnten.

Weiterhin hat die Gruppe in den letzten Monaten eine Methode entwickelt, die es erlaubt, zu einem beliebigen Zeitpunkt der Simulation den detaillierten Systemzustand wieder zu rekonstruieren. Dies ist eine wichtige Weiterentwicklung der Methode im Hinblick auf die Kalibrierung der Modelle hinsichtlich der experimentellen Beobachtungen.

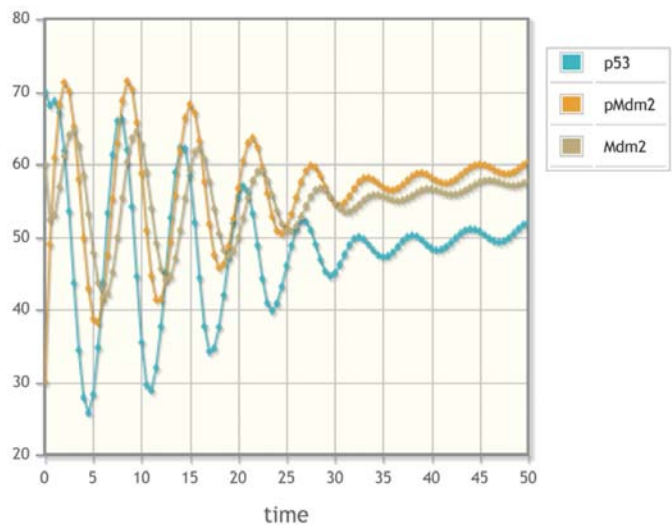


Abbildung 15 // Zeitabhängige Konzentrationsverläufe von drei verschiedenen Proteinen.

Die Kalibrierung von Modellen ist von enormer Wichtigkeit im Bereich der Systembiologie, da mathematische Modelle, die reale Systeme beschreiben, häufig auf starken Vereinfachungen der realen Dynamik basieren und mit Hilfe von Computersimulationen verbessert bzw. korrigiert werden müssen. Dies geschieht, indem unter den vielen simulierten Modellen eines ausgewählt wird, das den im Labor gemachten Beobachtungen am besten entspricht. Neben der Bestimmung unbekannter Parameter können aber auch Modelle, die verschiedenen Hypothesen über das System entsprechen, im Hinblick auf experimentelle Beobachtungen verglichen werden.

Zurzeit wird in der Arbeitsgruppe von Prof. Wolf ein Softwaretool entwickelt, das eine anwenderfreundliche Benutzung dieser neuartigen Simulationsverfahren erlaubt (siehe Abbildung 15).

Für ihre Forschungsarbeit zur Simulation stochastischer Reaktionsnetzwerke [1] erhielt Frau Prof. Wolf im Dezember 2013 beim Nachwuchswettbewerb des Magazins Technology Review die Auszeichnung „Innovatoren unter 35“.

Referenzen

- [1] Mikeev, L. and Wolf, V. **Parameter Estimation for Stochastic Hybrid Models of Biochemical Reaction Networks**. *Proceedings of the 15th International Conference on Hybrid Systems: Computation and Control (HSCC'12)*. 2012



Prof. Dr. Verena Wolf

Modeling and Simulation
Universität des Saarlandes
 Telefon +49 681 302 5586
 wolf@cs.uni-saarland.de

Wenn Dynamik wichtig wird: Regulatorische Netzwerke für Zeitreihen

Die idiopathische Lungenfibrose (IPF) ist eine schwer zu diagnostizierende Krankheit, für die es derzeit keine Heilung gibt und deren Ursache man noch nicht verstanden hat. Bei einer Lungenfibrose bildet sich vermehrt vernarbtes Gewebe in der Lunge, was die Lungenfunktion immer stärker beeinträchtigt, bis eine Lungentransplantation notwendig wird, um das Überleben des Patienten sicherzustellen. Wie es zu solch komplexen Krankheiten wie IPF kommt, stellt die heutige Medizin immer noch vor große Rätsel. Man vermutet jedoch, dass bei Patienten mit IPF Mechanismen aktiviert werden, die bei gesunden Menschen nur während der Lungenentwicklung vor und kurz nach der Geburt stattfinden.

Um diese Mechanismen zu verstehen, hat Marcel Schulz als Postdoc in Pittsburgh zusammen mit Wissenschaftlern von der Carnegie Mellon University und der University of Pittsburgh eine neue Methode entwickelt, die den Zeitverlauf von regulatorischen Prozessen, wie den der Lungenentwicklung, als ein dynamisches Netzwerk darstellen kann [1]. Dabei befinden sich in dem Netzwerk verschiedene Typen von Regulatoren, zum Beispiel kleine nicht-Protein-kodierende RNAs und Proteine, die DNA binden. Wir haben diese neuartige Methode auf Mausdaten angewandt, weil bei Mäusen die finale Lungenentwicklung erst nach der Geburt abläuft.

Mit Hilfe unserer neuentwickelten Methode haben wir ein komplexes, dynamisches Netzwerk gebaut, welches in Abbildung 16 dargestellt ist. Es beschreibt die regulatorischen Prozesse der Lungenentwicklung über die Zeit mit noch nie dagewesener Genauigkeit. Interessanterweise fanden wir heraus, dass 40 % der kleinen nicht-Protein-kodierenden RNAs in diesem Netzwerk der Maudaten auch in Patienten mit IPF dereguliert sind. Diese Übereinstimmung ist ein weiteres Indiz dafür, dass die Lungenentwicklung und IPF viele Gemeinsamkeiten haben. Die neuen Erkenntnisse, die mit unserem Netzwerk gewonnen wurden, können daher jetzt eingesetzt werden, um die Mechanismen der Lungenentwicklung besser zu verstehen und damit auch der Heilung von IPF-Patienten vielleicht einen Schritt näher zu kommen.

Am ZBI erweitert die Arbeitsgruppe von Marcel Schulz das Anwendungsspektrum der Methode. In Zusammenarbeit mit Prof. Alexandra Kiemer von der UdS analysieren wir die Entstehung von Leberkarzinomen in Mäusen. Dabei berücksichtigen wir neben nicht-Protein-kodierenden RNAs auch RNA-bindende Proteine bei der Konstruktion des dynamischen regulatorischen Netzwerkes.

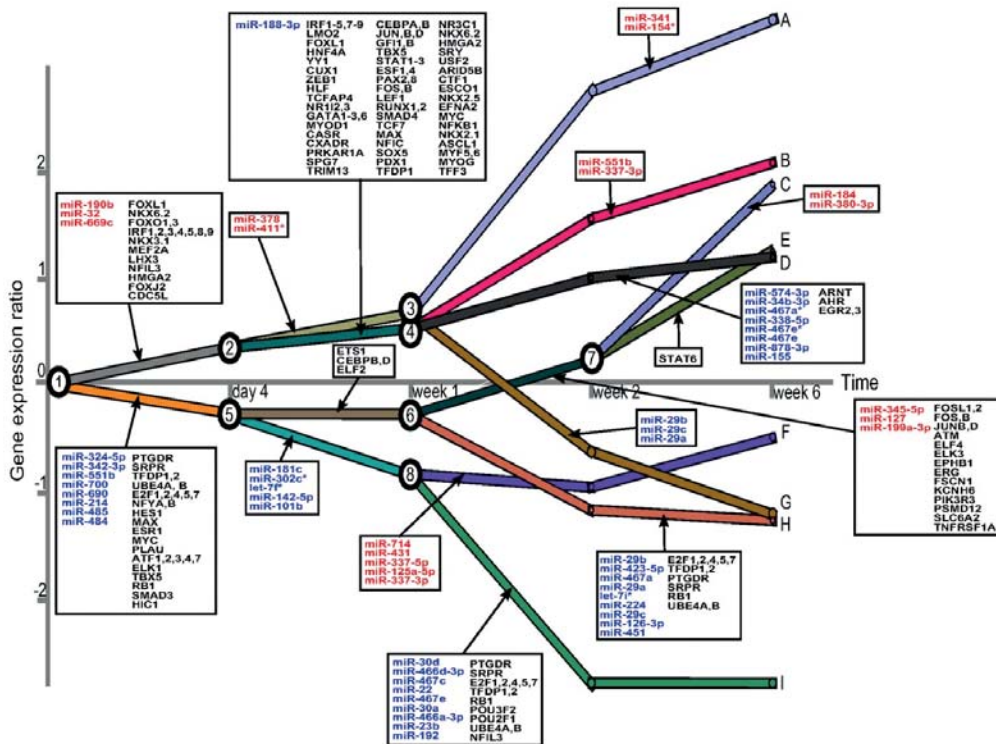


Abbildung 16 // Dieses dynamische Netzwerk der Lungenentwicklung stellt den Einfluss verschiedener regulatorischer Elemente über einen Zeitraum von 6 Wochen dar. Jede farbige Linie stellt eine Gruppe von Genen dar. Die Namen von DNA-bindenden Proteinen und nicht-Protein-kodierender RNA sind in den weißen Boxen dargestellt. Jede Box ist einem bestimmten Zeitpunkt zugeordnet. Die zeitliche Einordnung der regulatorischen Ereignisse erlaubt die bessere experimentelle Validierung.

Referenzen

[1] M. H. Schulz, K. V. Pandit, C. L. Lino Cardenas, N. Ambalavanan, N. Kaminski and Z. Bar-Joseph. **Reconstructing dynamic microRNA-regulated interaction networks.** *Proc Natl Acad Sci U S A* 110 (39): 15686-91. 2013



Dr. Marcel Schulz
 High-throughput Genomics & Systems Biology
 Universität des Saarlandes /
 Max-Planck-Institut für Informatik
 Telefon +49 681 9325 3115
 mschulz@mimi.uni-saarland.de

Identifizierung deregulierter Signalkaskaden

Förderung durch die Deutsche Forschungsgemeinschaft (DFG) (Schwerpunktprogramm 1335: Scalable Visual Analytics)

Bei vielen Krankheiten führen Mutationen im Erbgut zu Störungen des Informationsflusses in den betroffenen Zellen. Diese deregulierten Signalkaskaden verursachen im nächsten Schritt Änderungen in den Aktivitäten nachfolgender biologischer Prozesse, welche schließlich zu beobachtbaren Symptomen führen. Deregulierte Signalkaskaden spielen zum Beispiel bei Krebs eine zentrale Rolle. Die Krebsforschung hat in den letzten Jahrzehnten aufgedeckt, dass eine Fülle von Signalkaskaden bei der Entstehung und der Progression von Krebs beteiligt sein kann und dass aufgrund der Heterogenität von Tumoren selbst in Tumoren vom gleichen Subtyp viele unterschiedliche Signalkaskaden involviert sein können. Die Bestimmung der deregulierten Signalkaskaden in Tumoren vertieft aber nicht nur unser Verständnis der pathogenen Mechanismen, sondern sie hilft auch bei der genetischen und molekularen Typisierung und Klassifikation der Tumore, die wiederum eine individuelle Optimierung von Therapien (personalisierte Medizin) erleichtert.

Arbeitsgruppen des ZBI haben verschiedene effiziente und effektive Verfahren zur Detektion von deregulierten Pfaden und Teilgraphen in regulatorischen Netzwerken entwickelt und publiziert. Ein Beispiel für ein sehr leistungsfähiges Verfahren ist NetworkTrail [1], das mittels eines neuen Optimierungsalgorithmus den am stärksten deregulierten Teilgraphen in sehr großen regulatorischen Netzwerken bestimmen kann.

Darüber hinaus ist NetworkTrail eines der ersten Verfahren, das es gleichzeitig ermöglicht, die Gene und Proteine zu identifizieren, welche die Störungen in den Netzwerken verursachen. Die Anwendungen eines solchen Ansatzes reichen von der Aufklärung pathogener Prozesse bis hin zur Optimierung von Tumortherapien (siehe Abbildung 17). NetworkTrail wurde unter anderem eingesetzt, um die Effekte von Brustkrebsauslösenden Mutationen auf Epithelzellen der Brustdrüsen zu untersuchen. Hierbei konnte gezeigt werden, dass diese Mutationen zellulären Stress verursachen, der wiederum den Zelltod (die Apoptose) hemmt und somit die Überlebensfähigkeit von Zellen mit genetischen Aberrationen erhöht, was letztendlich die Entstehung von Brustkrebs begünstigt.

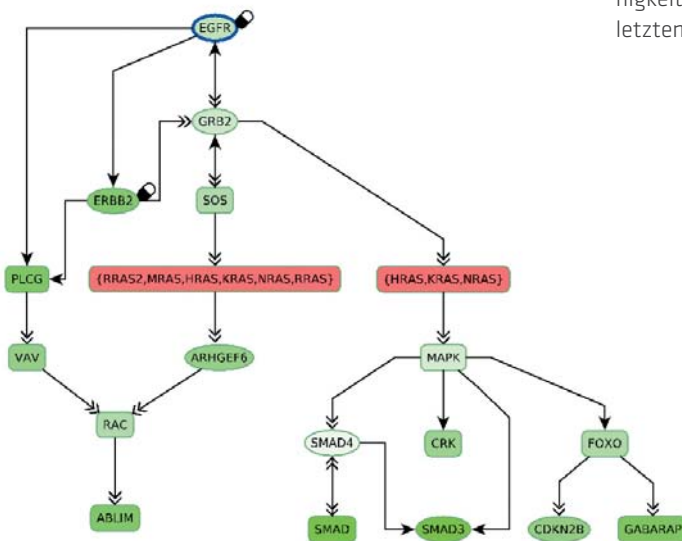


Abbildung 17 // Die „American Cancer Society“ empfiehlt im Falle von Adenokarzinomen des Darms fünf Medikamente als gerichtete Wirkstoffe, darunter zwei Antikörper (Cetuximab und Panitumumab) gegen das Zielmolekül EGFR. Eine Analyse des von EGFR beeinflussten Subnetzwerks zeigt in dem hier betrachteten Fall aber sehr deutlich, dass das Netzwerk im Tumor weniger aktiv ist als in Normalgewebe (grün bedeutet deaktiviert), was auf einen Funktionsverlust von EGFR durch Mutationen zurückzuführen sein könnte. Eine Therapie mit diesen beiden Medikamenten wäre in diesem Fall also unwirksam und würde schlimmstenfalls üble Nebenwirkungen verursachen. Die verstärkte Aktivität der RAS-Familie (rot) wird durch andere Gene induziert, welche im obigen Bild nicht gezeigt werden.



Prof. Dr. Hans-Peter Lenhof

Zentrum für Bioinformatik
Universität des Saarlandes
Telefon +49 681 302 64701
lenhof@bioinf.uni-sb.de



Prof. Dr. Andreas Keller

Klinische Bioinformatik
Universität des Saarlandes
Telefon +49 681 302 68611
andreas.keller@ccb.uni-saarland.de

Referenzen

- [1] C. Backes, A. Rurainski, G. W. Klau, O. Müller, D. Stockel, et al. **An integer linear programming approach for finding deregulated subgraphs in regulatory networks.** *Nucleic Acids Res* 40 (5): e43. 2012

Funktionelle Einheiten in Proteinnetzwerken

Förderung durch die Deutsche Forschungsgemeinschaft (DFG), Sonderforschungsbereich 1027

46

Das Erbgut menschlicher Zellen enthält etwa 20.000 Gene, die als Basis zur Synthese von Proteinen dienen sowie viele weitere Abschnitte, die zwar in RNA umgeschrieben werden können, aber nicht in Proteine übersetzt werden. Zu jedem Zeitpunkt werden in einzelnen Zellen des Menschen etwa 5.000 bis 8.000 verschiedene Proteine hergestellt. Ihre Gesamtheit bestimmt dann den Zustand der Zelle sowie deren Fähigkeit, bestimmte biochemische Reaktionen durchführen zu können, auf externe Signale zu reagieren oder auch einfach, um was für eine Art von Zelle es sich handelt. Es gibt nämlich etwa 200 verschiedene Typen von Körperzellen beim Menschen. Von großem Interesse für die moderne Biologie und Medizin ist es daher herauszufinden, wie die einzelnen zellulären Komponenten miteinander verschaltet sind und welche Bestandteile die Kontrolle haben.

Die Arbeitsgruppe Helms konstruiert mittels experimenteller Daten über den Ableseprozess (die sogenannte Transkription) einzelner Gene genregulatorische Netzwerke. Sie bestimmen, welche Transkriptionsfaktoren die Kontrolle über das Ablesen von welchen Genen – den so genannten Zielgenen des Transkriptionsfaktors – ausüben [1]. Ein besonderer Fokus der Arbeitsgruppe ist es herauszufinden, welche Faktoren die Entwicklung von Stammzellen (Zellen, die noch nicht ausdifferenziert sind) in verschiedene differenzierte Zelltypen bestimmen. Im Blickpunkt der Forschung stehen hier Transkriptionsfaktoren, die diese Prozesse kontrollieren bzw. steuern. Obwohl einzelne Interaktionen zwischen Biomolekülen stochastisch stattfinden, ist das „große Ganze“ im Körper zum Glück sehr stark reguliert. Da es im Menschen aber nur ein paar hundert Transkriptionsfaktoren gibt, kann die Zuordnung von Zielgenen und Transkriptionsfaktoren nicht eins zu eins erfolgen. Stattdessen wird das Ablesen eines Einzelgens meist durch Komplexe aus mehreren Transkriptionsfaktoren bestimmt. Ein besonderes Interesse der Arbeitsgruppe ist daher die Identifizierung der kombinatorischen Vielfalt solcher Multi-Proteinkomplexe. Auf diesem Gebiet treffen die Bereiche der Proteinstruktur und der Systembiologie aufeinander.

Die Arbeitsgruppe entwickelt zudem Techniken, um gewebespezifische Protein-Interaktionsnetzwerke für einzelne Zelltypen zu konstruieren [2]. Diese Erkenntnisse könnten dazu dienen, einen besseren Einblick in die detaillierten Verschaltungen dieser Netzwerke zu gewinnen. Ebenso könnten solche Proteinkomplexe mögliche Zielmoleküle für die Wirkstoffentwicklung sein. Ein besonders leicht zugängliches Organ ist dabei das menschliche Blut. Aus den vornehmlich im Knochenmark befindlichen Blut-Stammzellen bilden sich durch fein-regulierte Kontrollprozesse die verschiedenen Zelltypen unseres Blutes. Ein anderer wichtiger Bereich ist die Analyse von Krebswachstum.

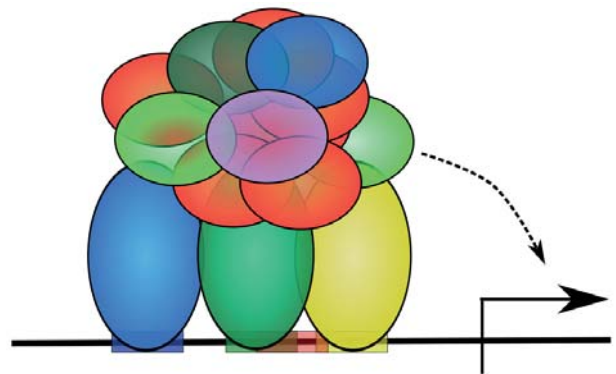


Abbildung 18 // Proteinkomplexe bestehend aus mehreren Transkriptionsfaktoren und rekrutierten Kofaktoren steuern die eukaryotische Transkription.

Referenzen

- [1] M. Hamed, C. Spaniol, M. Nazarieh and V. Helms. **TFmiR: A web server for constructing and analysing disease-specific transcription factor and miRNA co-regulatory networks.** *Nucleic Acids Res* 43 (W1): W283-8. 2015
- [2] T. Will and V. Helms. **PPIXpress: construction of condition-specific protein interaction networks based on transcript expression.** *Bioinformatics*, Vol. 32, P. 571-578, 2016



Prof. Dr. Volkhard Helms

**Zentrum für Bioinformatik
Universität des Saarlandes**

Telefon +49 681 302 70701

volkhard.helms@bioinformatik.uni-saarland.de



Thorsten Will

**Zentrum für Bioinformatik
Graduiertenschule für Informatik**

Telefon +49 681 302 70713

thorsten.will@bioinformatik.uni-saarland.de

Metabolische Netze von Mikroorganismen

Förderung im Rahmen des DFG-Projektes „Identification of host-adapted metabolic functions important for *Yersinia pseudotuberculosis* virulence“ im Schwerpunktprogramm 1316 der DFG „Host-Pathogen-Interaktionen“ des BMBF-Verbundprojektes „Natur-Präklinik“ sowie des BMBF-VIP-Projektes „Bio2Nylon“

47

Bakterien sind winzige Alleskönner. Sie halten extreme Bedingungen aus, können im ewigen Polareis und in kochendem Wasser überleben und in wenigen Minuten alles herstellen, was für eine vollständige Zelle an Bestandteilen nötig ist. Oft brauchen sie nur ein bisschen Salz und Zucker zum glücklich sein. Auch wenn sie nur ein bisschen Salz und Zucker zum glücklich sein. Auch wenn sie nur wenige Mikrometer groß sind, macht sie ihre einzigartige Stoffwechselvielfalt für die biotechnologische Forschung sehr interessant. Mikroorganismen enthalten tausende verschiedene Moleküle, die sie für ihre ausgeklügelten Synthesewege verwenden (siehe Abbildung 19). Am Institut für Systembiotechnologie der Universität des Saarlandes erforschen die Wissenschaftler daher das komplexe Innenleben von Mikroorganismen. In den letzten fünfzehn Jahren haben sie an der UdS, und zuvor an der Universität Münster und der TU Braunschweig, international führend, systembiologische Methoden entwickelt, mit denen sie die zellulären Vorgänge erfassen können.

So lässt sich durch eine Verfütterung von Nährstoffen, die an ausgewählten Atompositionen spezifische Isotope tragen, und durch eine anschließende modellgestützte Auswertung des Einbaus der Isotope in die zahlreichen Metabolite der Zellen, der „molekulare Verkehr“ durch das komplexe Stoffwechselnetzwerk der Zellen verfolgen. Ergänzend nutzen die Forscher metabolische Netzwerkmodelle, um auf Basis des genomischen Repertoires der Zelle Voraussagen über Ihre Stoffwechsellösungen zu treffen. Das Ziel der dabei erzeugten umfangreichen Datensätze, die mit Hilfe mathematischer und bioinformatischer Methoden generiert werden, ist ein grundlegendes Verständnis der zugrunde liegenden Stoffwechsellösungen. Daneben versucht man, die richtigen Stellschrauben zur gezielten Veränderung der Stoffwechsellösung mit Hilfe der synthetischen Biologie zu entdecken. So lassen sich später unerwünschte biochemische Reaktionen durch Entfernen der entsprechenden Genabschnitte im Erbgut der Zelle eliminieren oder gewünschte Aktivitäten durch den Einbau entsprechend aktivierender Gensequenzen verstärken.



Abbildung 19 // Selbst kleinste Mikroorganismen besitzen ein komplexes Netzwerk biochemischer Reaktionen als Grundlage der zellulären Vorgänge. Die technologischen Spezialitäten des Teams von Prof. Wittmann (Institut für Systembiotechnologie, Universität des Saarlandes) sind isotopebasierte metabolische Stoffflussanalysen sowie computerbasierte Simulationen von Stoffwechselwegen und metabolischen Netzwerken. Diese erlauben die Entschlüsselung hunderter gleichzeitig ablaufender Stoffwechsellösungen. Dies ist eine zentrale Grundlage, um wichtige Regulationsmechanismen zu identifizieren und genetische Ziele für das Design gewünschter Phänotypen vorherzusagen, zum Beispiel für den Einsatz von Bakterien, Hefen, Pilzen und sogar ganzen Pflanzen als Zellfabriken für die Biotechnologie.

Referenzen

- [1] J. Becker and C. Wittmann. **Advanced biotechnology: metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products.** *Angew Chem Int Ed Engl* 54 (11): 3328-50. 2015.



Prof. Dr. Christoph Wittmann

**Institut für Systembiologie
Universität des Saarlandes**

Telefon +49 681 302 71971

christoph.wittmann@uni-saarland.de

Epigenetik

Die Epigenetik ist ein Forschungsgebiet, das heute im Zentrum der weltweiten Aufmerksamkeit liegt. Der Grund hierfür ist, dass nach der Sequenzierung des menschlichen Genoms vor zwölf Jahren schnell klar wurde, dass wesentliche Aspekte der Zellregulation nicht direkt aus der genomischen DNA abzulesen sind – denn diese ist ja in allen Zellen gleich. Dagegen unterscheiden sich verschiedene Zellen erheblich voneinander – etwa Zellen in verschiedenen Geweben sowie kranke von gesunden Zellen. Ferner beeinflussen zentrale biologische Prozesse, wie die Antwort auf Stress und das Altern, den Zellzustand umfassend. Diese Faktoren der Zellregulation gehen auf die dynamische Organisation des Genoms im Zellkern zurück, die wiederum auf chemischen Modifikationen der genomischen DNA und des sie umgebenden Molekülgerüsts aus Proteinen beruht.

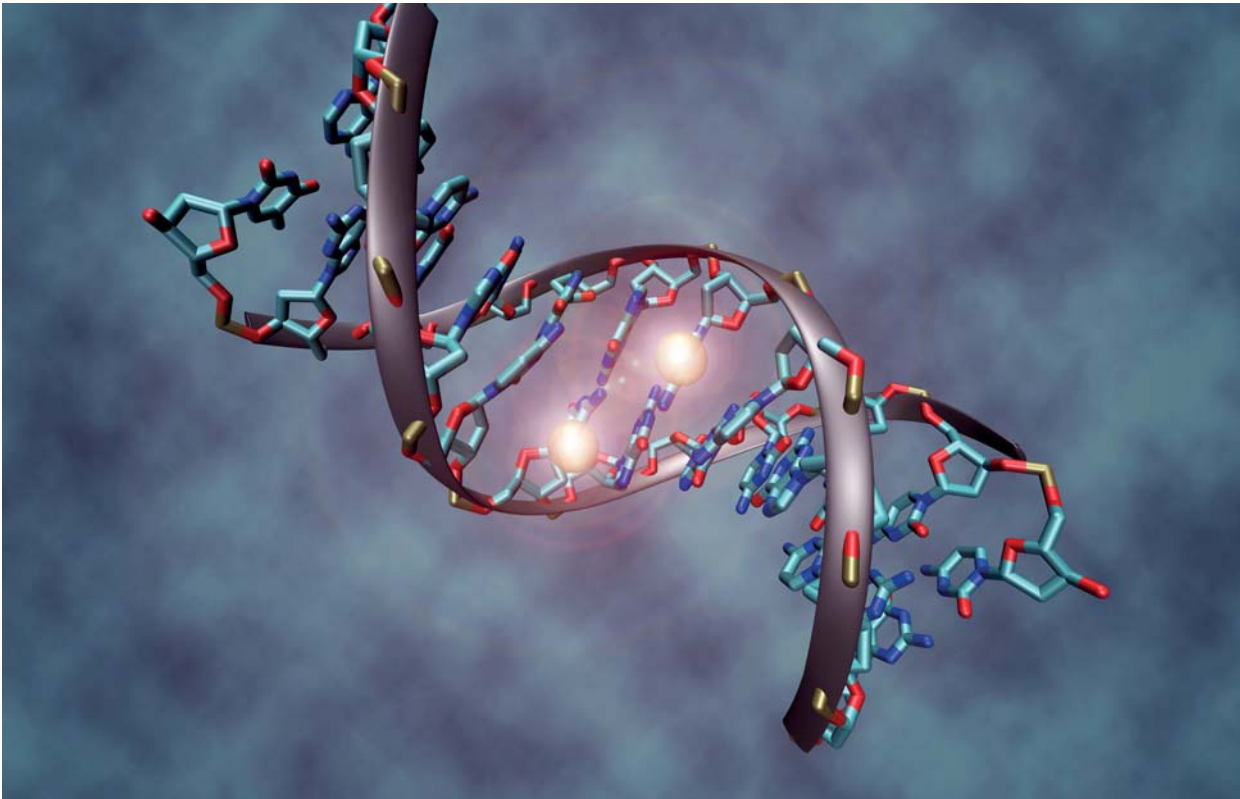
Die Epigenetik wurde in den 1980er Jahren zunächst als die Wissenschaft der Vererbung solcher Eigenschaften definiert, die sich nicht durch die in DNA codierte Information manifestieren. Es stellte sich dann jedoch schnell heraus, dass solche Vererbungsvorgänge, wie auch die Zellregulation allgemein, auf die komplexe Art gestützt sind, in der das Genom im Zellkern verpackt wird. An dieser Verpackung sind spezielle Proteine beteiligt, die zusammen mit der genomischen DNA das sogenannte Chromatin bilden. Das Chromatin erfüllt zwei Aufgaben. Zum einen muss die genomische DNA, die in ihrer gesamten Länge 2 m misst, in dem Zellkern mit einem Radius von einem tausendstel Millimeter untergebracht werden. Zum anderen sollen die Teile der genomischen DNA, in denen Information codiert ist, die nicht von der Zelle gebraucht wird, vor dem Ablesen geschützt werden. Da die genomische Information, die die Zelle benötigt, von dem Gewebe und dem Zellzustand abhängt, ist das Chromatin hoch dynamisch. Seine Struktur wird durch vielfältige chemische Modifikationen der DNA selbst und der sie umgebenden Proteine gesteuert. Die Gesamtheit dieser Modifikationen bezeichnet man als das Epigenom.

Mit in den letzten zwei Jahrzehnten entwickelten molekularbiologischen Techniken können diese Modifikationen (epi-) genomweit vermessen werden. Damit ist das Epigenom kartierbar geworden. Während ein Mensch nur ein Genom hat, hat er im Gegensatz dazu viele Epigenome. Im Prinzip unterscheiden sich die Epigenome in allen der etwa 200 Gewebetypen des Organismus. Darüber hinaus unterscheiden sie sich in jedem Zelltyp in gesunden und krankhaft veränderten Zellen. Die Kenntnis des Epigenoms gilt als Voraussetzung für das Verständnis der Prozesse in dem betreffenden Zelltyp. Aus diesem Grund wurde vor einigen Jahren das International Human Epigenome Consortium (IHEC) gegründet – ein weltweiter Zusammenschluss von Wissenschaftlern mit dem Ziel, in den nächsten Jahren mindestens 1.000 Epigenome von menschlichen Zellen zu kartieren. Das Zentrum für Bioinformatik ist an zwei Stellen an IHEC beteiligt.

Zum einen sind zwei seiner Arbeitsgruppen Partner im BLUEPRINT Projekt der EU. Dieses Projekt mit einer EU Fördersumme von 30 Mio Euro, der größten einem Einzelprojekt im Bereich der Biologie verliehenen Summe, hat zum Ziel, 100 Epigenome von Zellen der Blutlinie (hämatopoetische Zellen), sowohl kranke (hauptsächlich bösartig veränderte) als auch gesunde Zellen, zu kartieren. Zum anderen ist das ZBI Partner im Deutschen Epigenom Programm (DEEP), dem vom BMBF mit rund 20 Mio Euro finanzierten deutschen Beitrag zu IHEC. In DEEP werden vor allem Stoffwechsel- und Immunkrankheiten untersucht und weitere 70 Epigenome kartiert. Das Zentrum für Bioinformatik stellt mit Prof. Jörn Walter den Koordinator von DEEP. Prof. Thomas Lengauer koordiniert den Bioinformatikbereich des Projektes.

Neben der Kartierungsarbeit selbst treibt das Zentrum für Bioinformatik die experimentelle Technologie zur Bestimmung des Chromatinzustandes voran und entwickelt Methoden zur biologischen Interpretation der anfallenden Daten.

Dazu gehört das Auffinden von epigenetischen Mustern, die charakteristisch für bestimmte Krankheitstypen und -stadien sind. Zu diesem Zweck werden im Zentrum für Bioinformatik seit Jahren Softwaresysteme entwickelt, die weltweit eingesetzt werden. Neben der Entwicklung von Software ist das Zentrum für Bioinformatik auch an Validierungsstudien beteiligt, in denen die Software eingesetzt wird, um aus epigenetischen Daten neue biologische Erkenntnisse zu gewinnen. So konnten in einer im Jahre 2012 veröffentlichten Studie die Ursprungsgewebe von zahlreichen verschiedenen bösartigen Tumoren identifiziert werden, die im Patienten nur noch in gewebefremden Metastasen zu finden waren. Eine solche Gewebsbestimmung ist zentral für eine effektive Tumorthherapie.



Methylierungsstellen in DNA-Doppelhelix

Epigenetik

Integrative Methoden zur Analyse epigenomischer Daten	50
Software für die Analyse von DNA Modifikationsdaten.	51
Zwillingsstudien: epigenetische Studien in Reinform	52
Integrative epigenetische Analyse von Genaktivität	53

Integrative Methoden zur Analyse epigenomischer Daten

Förderung durch das Bundesministerium für Bildung und Forschung (BMBF) (Deutsches Epigenomprogramm) und EU (BLUEPRINT Projekt)

50

Unsere DNA kodiert das Aussehen und Verhalten aller Zellen unseres Körpers. Es gibt jedoch auch eine Vielzahl zusätzlicher Schichten der Regulation, die das Schicksal der Zelle bestimmen. Die Struktur der Packung der DNA im Zellkern beeinflusst, ob und wie in der DNA codierte Gene abgelesen werden können und ist damit für die äußere Erscheinung und Funktion der Zelle, dem sogenannten Phänotyp, verantwortlich. Die Packungsstruktur wird durch chemische Modifikationen der DNA selbst (DNA Methylierung) oder des molekularen Gerüsts, z. B. durch Modifikation der an die DNA bindenden Histonproteine, beeinflusst. Die Analyse dieser epigenetischen Träger molekularer Information wird durch Softwares ermöglicht, die am ZBI entwickelt werden: EpiExplorer basiert auf einer Technologie, die für die Suche nach textueller Information im Internet entwickelt wurde und stellt ein Werkzeug dar, mit dem genomische und epigenomische Merkmale schnell und intuitiv erkannt, dargestellt und interaktiv erkundet werden können. Bei RnBeads handelt es sich um ein leicht zu bedienendes Softwarepaket für die umfassende Analyse experimentell ermittelter DNA Methylierungsdaten.

EpiExplorer – ein Werkzeug für die effiziente Suche und den Vergleich genomischer und epigenomischer Merkmale

Moderne Labortechniken ermöglichen eine umfassende „Kartierung“ der epigenetischen Modifikationen zu erschwinglichen Kosten. Nationale und internationale Forschungskonsortien wie etwa das International Human Epigenome Consortium (IHEC), das von der EU geförderte BLUEPRINT Konsortium und das vom BMBF geförderte Deutsche Epigenom Programm (DEEP) haben es sich zur Aufgabe gemacht diese Karten für verschiedenste Zelltypen und Individuen zu erstellen. Große Herausforderungen der Bioinformatik bestehen aber weiterhin in der integrativen Analyse der sich ergebenden, äußerst vielschichtigen Datenflut und darin, das Verständnis des Zusammenspiels genomischer und epigenomischer Faktoren zu ermöglichen. Mithilfe von EpiExplorer können Wissenschaftler schnell und intuitiv große, heterogene Datensätze durchforsten und wissenschaftliche Hypothesen entwickeln. Dies geschieht genomweit in Echtzeit über das Internet.

Das Tool integriert eine Vielzahl (epi-)genomischer Karten und ermöglicht die Identifikation besonders interessanter genomischer Regionen. Unter der Haube basiert EpiExplorer auf einem vielseitigen Indizierungsverfahren, das in Sekundenschnelle umfassende Analysen und Visualisierungen ermöglicht. Des Weiteren können Vergleichsansichten für verschiedene Datensätze erstellt werden und Benutzer können eigene Datensätze, die genomische Abschnitte von speziellem Interesse genauer beschreiben, in ihre Analysen mit einbeziehen [1].

Umfassende Analyse von DNA Methylierungsdaten mithilfe von RnBeads

Die Methylierung der DNA ist ein wichtiger Bestandteil der epigenetischen Regulation einer Zelle und die derzeit wahrscheinlich am intensivsten erforschte epigenetische Modifikation. Sie spielt eine wichtige Rolle in der Embryonalentwicklung und ist an der Entstehung komplexer Krankheiten wie beispielsweise Krebs beteiligt. Derzeit gibt es eine Reihe experimenteller Verfahren, um genomweit DNA Methylierungsmuster zu kartieren. In Zusammenarbeit zwischen den Arbeitsgruppen Walter und Lengauer wurde z. B. das RnBeads Softwarepaket entwickelt, mit dessen Hilfe sowohl experimentell als auch bioinformatisch orientierte Wissenschaftler die Daten aus Laborexperimenten umfassend analysieren können: Verzerrungen der Daten durch experimentelle Einflüsse können identifiziert und korrigiert sowie Muster in der Methylierung quantifiziert, visualisiert und verglichen werden.

Beispielsweise ist es möglich, genomische Abschnitte zu identifizieren, deren Methylierungs-Fingerabdrücke sich signifikant zwischen Normal- und Krebsgewebe unterscheiden und diese mit biologischen Merkmalen zu annotieren. Identifizierte Regionen können dann in den EpiExplorer exportiert und dort mit anderen (epi-)genomischen Karten verglichen werden. Obwohl eine umfassende Standardanalyse in RnBeads keine Programmiererfahrung voraussetzt, ermöglicht der modulare Aufbau des Softwarepakets auf den Studienzweck zugeschnittene, individuelle Analysen. Standardmäßig werden umfassende und interpretierbare Resultate in HTML Dokumenten ausgegeben. Somit ist es leicht, die Ergebnisse von Analysen zu verwalten, zu vergleichen und mit Kollegen oder Kollaborationspartnern auszutauschen [2].

Referenzen

- [1] K. Halachev, H. Bast, F. Albrecht, T. Lengauer and C. Bock. **EpiExplorer: live exploration and global analysis of large epigenomic datasets.** *Genome Biol* 13 (10): R96. 2012
- [2] Y. Assenov, F. Muller, P. Lutsik, J. Walter, T. Lengauer and C. Bock. **Comprehensive analysis of DNA methylation data with RnBeads.** *Nat Methods* 11 (11): 1138-40. 2014



Fabian Müller

Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik

Telefon + 49 681 9325 3009
fmuller@mpi-inf.mpg.de

Software für die Analyse von DNA Modifikationsdaten

Förderung durch das Bundesministerium für Bildung und Forschung (BMBF) und EU (BLUEPRINT Projekt)

Eine wichtige Aufgabe bei der Kartierung des Epigenoms einer Zelle ist die experimentelle Messung und computerbasierte Auswertung der sogenannten DNA Methylierungsmuster. Methylierung ist eine einfache chemische Modifikation des Cytosins – einer der vier Basenbausteine der DNA –, die in unterschiedlichen Regionen des Genoms unterschiedlich häufig auftritt. Wie jedes Experiment im Labor ist auch die Bestimmung der DNA Methylierung nicht frei von Messfehlern und Hintergrundrauschen, was eine detaillierte Bewertung der Messwerte durch Experten erforderlich macht. Hierbei ist es von enormer Bedeutung, dass Fehlerkorrekturen und qualitätssteigernde Anpassungen der Daten durch die Experten manuell und wenig fehleranfällig durchgeführt werden können. Dafür ist die gezielte Betrachtung eines einzelnen Locus oder mehrerer Loci, also begrenzter Bereiche der DNA, von erwiesener oder zumindest vermuteter Relevanz bezüglich der allgemeinen Qualität des Datensatzes notwendig. Sie lässt Rückschlüsse über die Verwertbarkeit der Daten für weitere Analysen zu. Des Weiteren ist eine locuspezifische Analyse beim Vergleich von verschiedenen Zellzuständen, beispielsweise gesund und krank, häufig für Experten aufschlussreicher als eine (epi-)genomweite algorithmische Zusammenfassung der Daten, da abnormale Veränderungen der DNA Methylierungsmuster in frühen Entwicklungsstadien von Krankheiten auf diesem Wege einfacher zu erkennen sind. Fehler in den Methylierungsmustern der DNA haben einen starken Einfluss auf das Entstehen und Fortschreiten diverser Krankheitsbilder, darunter sehr schwerwiegende Krankheiten, wie Krebs oder Alzheimer. Allerdings ermöglicht die moderne Forschung nicht nur ein frühzeitiges Erkennen von Krankheiten anhand von Abweichungen in der DNA Methylierung, man kann auch verbesserte Diagnosen stellen und so z. B. eine individuelle Einschätzung über den Erfolg einer Chemotherapie für bestimmte Krebsarten liefern. Um diesen Fortschritt auch und gerade im medizinischen Bereich zu gewährleisten, bedarf es zeitgemäßer und einfach zu bedienender Software, die die Analyse von Methylierungsdaten auch für Nicht-Bioinformatiker erlaubt.

Schritt halten mit dem Stand der Forschung

In enger Zusammenarbeit mit der Gruppe Genetik/Epigenetik von Prof. Walter entwickeln wir Softwares, die genau diesen Aufgabenbereich abdecken (siehe auch EpiExplorer und RnBeads – integrative Methoden zur Analyse epigenomischer Daten). BiQ Analyzer ist ein von uns entwickeltes und im Jahr 2005 veröffentlichtes Programm zur Visualisierung, Qualitätskontrolle und lokusspezifischen, d.h. zielgerichteten Analyse von DNA Methylierungsdaten, das sich schnell als ein Standardwerkzeug etablierte. Der technische Fortschritt, sprich die immer größere und schnellere Datenproduktion im Labor, erforderte wenige Jahre später eine Anpassung unserer Software auf die sogenannten Hochdurchsatzverfahren [1]. Wichtig hierbei war die Erhaltung der einfachen Bedienbarkeit über dieselbe grafische Oberfläche, um Nutzern eine Umgewöhnung zu ersparen. Denn nicht nur im technologischen Bereich, sondern auch in der Biologie erfordern Neuentdeckungen eine Aktualisierung der Programme. So wurden kürzlich chemische Varianten der DNA Methylierung entdeckt, deren Bedeutung immer intensiver erforscht wird. Im Falle der sogenannten Hydroxymethylierung verdichten sich beispielsweise die Hinweise, dass es sich hierbei um einen wichtigen Faktor im Prozess der Demethylierung, also der gezielten Entfernung von Methylgruppen im Genom, handelt – ein Vorgang, für den bis dato kein passendes Erklärungsmodell zur Verfügung stand. BiQ Analyzer HiMod [2] unterstützt die Analyse von regulären DNA Methylierungsdaten und diesen neu entdeckten Varianten. Ein zusätzliches Problem ist hierbei, dass Standardexperimente keine Unterscheidung zwischen Methylierung und Hydroxymethylierung zulassen. Speziell angepasste Laborprotokolle sind erforderlich, um das Auftreten der Hydroxymethylierung in BiQ Analyzer HiMod über eine vergleichende Datenanalyse festzustellen. Die kontinuierliche Pflege und Weiterentwicklung von Bioinformatiksoftware ist daher essenziell für die Unterstützung von Forschern weltweit, nicht nur, um die Biologie zu verstehen, sondern auch, um neue Ansätze in der Therapie von komplexen Krankheiten zu entdecken.

51

Referenzen

- [1] P. Lutsik, L. Feuerbach, J. Arand, T. Lengauer, J. Walter and C. Bock. **BiQ Analyzer HT: locus-specific analysis of DNA methylation by high-throughput bisulfite sequencing.** *Nucleic Acids Res* 39 (Web Server issue): W551-6. 2011
- [2] D. Becker, P. Lutsik, P. Ebert, C. Bock, T. Lengauer and J. Walter. **BiQ Analyzer HiMod: an interactive software tool for high-throughput locus-specific analysis of 5-methylcytosine and its oxidized derivatives.** *Nucleic Acids Res* 42 (Web Server issue): W501-7. 2014



Peter Ebert

**Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik**

Telefon +49 681 9325 3008
pebert@mpi-inf.mpg.de



Pavlo Lutsik

**Genetik / Epigenetik
Universität des Saarlandes**

Telefon +49 681 302 2881
p.lutsik@mx.uni-saarland.de

Zwillingsstudien: epigenetische Studien in Reinform

Förderung durch die Humboldt-Stiftung und DFG (Schwerpunktprogramm Nr. 1029)

Im Menschen gibt es zwei wesentliche Faktoren die eine einfache Interpretation epigenetischer Daten erschweren. Zum einen kann man die „betroffenen“, d.h. physiologisch veränderten Zellen (z.B. Neuronen, Muskelzellen, etc.) aus Probanden aus ethischen Gründen nicht gewinnen oder gar in Reinform isolieren. Für vergleichende Studien stehen dann meist „nur“ Gemische aus Speichel, Blut- oder Mundschleimhautzellen mit unbekannter Zusammensetzung zur Verfügung. Zum anderen ist die epigenetisch zu untersuchende Merkmalsausprägung im Menschen auch durch die individuelle genetische Grundkonstitution des Einzelnen mitbestimmt. Die genetische Variabilität überdeckt so häufig epigenetische Effekte. Eines unserer Forschungsziele ist es daher neue experimentelle und bioinformatische Verfahren zu etablieren, die solche störenden Faktoren und Einflüsse besser berücksichtigen.

Studien an eineiigen Zwillingen bieten eine exzellente Möglichkeit, einen der Faktoren – die genetische Variabilität – auszuschließen. Ein Beispiel solch einer Studie ist die Analyse, die wir an 12 eineiigen Zwillingspaaren durchgeführt haben. Jedes Zwillingsspaar unterschied sich in einem messbaren körperlichen oder psychisch testbaren Merkmal, wie zum Beispiel ihrem Gewicht bei der Geburt. Epidemiologische Untersuchungen deuteten an, dass diese Geburtsgewichtsunterschiede auf verschiedene epigenetische Programme im Stoffwechsel der Zwillinge während der gemeinsam erfahrenen Schwangerschaft zurückzuführen sind. Es wurde angenommen, dass diese individuell unterschiedlichen epigenetischen Programme der Zwillinge bis ins Erwachsenenalter erhalten bleiben und ein erhöhtes Risiko für eine Ausbildung von Herz-Kreislauf- und Stoffwechselerkrankungen im Erwachsenenalter darstellen. Ziel unserer Studie war es, solche langfristigen epigenetischen Prägungen im Genom der erwachsenen Zwillinge aufzuspüren. Als Quelle für unsere Analysen diente uns DNA aus Speichelproben. Für die Studie nutzten wir neue moderne hochauflösende Methoden der Epigenom-Analytik, mit der wir eine große Anzahl epigenetischer Markierungen im gesamten Genom erfassen konnten. Die epigenetischen Daten wurden dann statistischen Lern- und Korrekturverfahren unterzogen, um zunächst Messfehler zu korrigieren.

Mit Hilfe bekannter Vergleichsdaten, die nicht auf Zellmixturen basieren, sondern nach Zelltyp getrennt sind, konnten wir dann zusätzlich auch die unterschiedliche Zusammensetzung der Speichelproben rechnerisch angleichen.

Diese erstmalig vorgenommene Vorgehensweise zeigte, dass die Epigenome der Zwillingsspaare viel ähnlicher sind als zuvor vermutet und keine nachhaltigen Prägungen zu finden waren [1]. Die experimentelle und bioinformatische Herangehensweise unserer Studie setzte Maßstäbe für andere, vergleichende epigenetische Studien. Analoge Korrektur-Methoden gehören mittlerweile zum Standard vergleichender epigenetischer Studien.

Modellierung epigenetischer Vererbung

Epigenetische Muster können über lange Zeit und viele Zellteilungen hinweg stabil „vererbt“ werden. Sie können sich aber auch in bestimmten Entwicklungsphasen des Menschen durch natürliche und künstliche Prozesse stark verändern. Im Verlauf der normalen Entwicklung werden Muster z.T. in großem Umfang gelöscht und wieder neu und anders gesetzt, um so neue Entwicklungsprogramme zu starten. Insbesondere in Krebs-Zellen findet man vergleichbare massive epigenetische Musterveränderungen. Es ist daher sehr bedeutsam, die molekularen Ursachen dieser epigenetischen Umformungen zu verstehen, mit dem Ziel, neue Therapieansätze zu entwickeln.

Bereits zu Beginn des Lebens, d.h. während der ersten Zellteilungen nach der Befruchtung, beobachtet man natürlich vorkommende umfassende Umformungen von epigenetischen Mustern. So werden zunächst erste epigenetisch festgelegte Alleskönner-Stammzellen gebildet, die dann wieder in Zellen mit spezielleren Aufgaben differenziert werden. Im Zuge dieser Zell-Programm-Änderungen werden epigenetische Muster zunächst umfassend gelöscht und dann wieder neu gesetzt. Unsere Forschung widmet sich der Frage, welche Mechanismen das erste Löschen, die Neugestaltung und dann die perfekte Vererbung (Weitergabe) epigenetischer Informationen steuern. Wir untersuchen dabei exemplarisch die Vererbung der DNA Methylierung. Mit speziellen Verfahren bestimmen wir in einzelnen DNA Molekülen die Muster der An- oder Abwesenheit von Methylgruppen in der DNA mit Hilfe neuer tiefer Sequenzierungsmethoden. Auf der Basis der dabei generierten umfangreichen Daten werden in enger Zusammenarbeit mit Frau Prof. Verena Wolf epigenetische Prozessabläufe mit Hilfe von Markov-Ketten modelliert, die den Erhalt bzw. den Umbau (Verlust oder Neugewinn) solcher Markierungen vermutlich steuern [2]. Diese Analysen liefern uns ganz neue Einblicke in epigenetische Prozessabläufe, die mit großer Wahrscheinlichkeit auch bei der Entstehung von komplexen Krankheiten, wie z. B. bei Krebs, eine Rolle spielen.

Referenzen

- [1] Souren et al. **Adult monozygotic twins discordant for intra-uterine growth have indistinguishable genome-wide DNA methylation profiles.** *Genome Biol.* (2013); 14:R44 doi:10.1186/gb-2013-14-5-r44.
- [2] J. Arand, D. Spieler, T. Karius, M. R. Branco, D. Meilinger, A. Meissner, T. Jenuwein, G. Xu, H. Leonhardt, V. Wolf and J. Walter. **In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases.** *PLoS Genet* 8 (6): e1002750. 2012



Prof. Dr. Jörn Walter

Genetik / Epigenetik
 Universität des Saarlandes
 +49 681 302 2425
 j.walter@mx.uni-saarland.de

Integrative epigenetische Analyse von Genaktivität

Gefördert durch „Cluster of Excellence: Multimodal Computing and Interaction“

Seit Jahrzehnten sind Forscher auf der ganzen Welt dabei zu entschlüsseln, wie unser Genom reguliert und gesteuert wird, um zu verstehen, worin sich die verschiedenen Zellen unseres Körpers unterscheiden, z.B. Herzzellen von Leberzellen. Nur wenn wir gesunde Varianten aller existierenden Zelltypen ausreichend verstehen, können wir identifizieren, was den Unterschied zu einem kranken Zelltyp eines Gewebes ausmacht.

Mit der Sequenzierung des menschlichen Genoms kam die Erkenntnis, dass die Epigenetik die entscheidende Rolle bei der Differenzierung von Zellen spielt. Sie ist somit ein wichtiger Baustein für das Verständnis der Regulation von Genen. Mit dem Vorhandensein von hunderten von epigenetischen Datensätzen, die im Rahmen großer, weltweiter Forschungs-kooperationen von gesunden Zellen verschiedenster Gewebe gemessen wurden, wird es der Forschung ermöglicht, dieses Verständnis entscheidend verbessern zu können.

In diesen Datensätzen werden verschiedene Komponenten des epigenetischen Systems vermessen: z. B. chemische Modifikationen der Histonproteine, um die die DNA gewickelt wird, um so das Chromatin zu bilden. Darüber hinaus messen wir die Positionierungen der aus dieser Wicklung resultierenden knotenförmigen Nukleosomen entlang des langen fadenförmigen DNA Moleküls, die Methylierung des DNA Moleküls selbst, die Aktivität von kleinen regulatorischen RNA Molekülen, die Aktivität von Genen, sowie die Aktivität von vielen an die DNA-bindenden Proteinen. Jede dieser Komponenten wird mit Hilfe einer anderen Technologie vermessen, die ihren eigenen Schwankungen unterliegt und deren Entwicklung meist noch nicht vollständig abgeschlossen ist. Oft wird dabei auch nicht die direkte Aktivität eines epigenetischen Regulators gemessen, z. B. welches DNA-bindende Protein an welcher Stelle in einem Chromosom bindet, was die Analysen zusätzlich erschwert.

Aufgrund der Komplexität dieser unterschiedlichen Datenquellen und der zu Grunde liegenden Biologie, mussten frühere Studien diese Komponenten isoliert betrachten, z. B. sich auf die DNA-Methylierung beschränken. Obwohl jede dieser Komponenten für sich betrachtet eine entscheidende Rolle spielen kann, so ist es doch unabdingbar, alle Komponenten in einen Zusammenhang zu setzen, wie es in der Zelle auch passiert, um das große Ganze zu verstehen.

Ein wichtiges Ziel ist daher die Erstellung von realistischen Modellen für die Regulation von Genen, welche verschiedene epigenetische Komponenten miteinander verknüpfen, Modelle, die erklären können, wie sich Herzzellen von Leberzellen unterscheiden – Modelle, die dazu benutzt werden können, zu beurteilen, welche Komponenten in diesem komplexen System für das Entstehen einer Krankheit verantwortlich sind – Modelle, die von vielen Wissenschaftlern weltweit benutzt werden können, um das Verständnis von Genregulation voranzutreiben.

Doch die Integration der verschiedenen epigenetischen Datensätze in solchen Modellen ist aus den bereits erwähnten Gründen eine große Herausforderung. Wir benutzen in unseren Ansätzen neueste Methoden aus dem Bereich des maschinellen Lernens, um solche Modelle zu entwerfen und sind dadurch in der Lage, Messfehler und überflüssige Informationen zu reduzieren. Unser Ziel dabei ist es, die Aktivität von Genen so genau wie möglich vorherzusagen. In ersten Studien ist es uns gelungen, solche Modelle von unterschiedlichen Leberzelltypen zu bauen. Damit haben wir den Grundstein dafür gelegt, das Verfahren auf hunderte verschiedene Zelltypen auszuweiten und dadurch das Verständnis von epigenetischer Regulation zu verbessern.

Die Bioinformatik als Schnittstelle zwischen den verschiedenen Fachrichtungen ist unabdingbar dafür, diesen komplexen integrativen Ansatz erfolgreich zu bewältigen. Dabei spielt die Bioinformatik am ZBI in Saarbrücken mit ihren Beiträgen zu weltweiten Kooperationen und hochmoderner Forschung eine entscheidende Rolle.



Dr. Marcel Schulz

High-throughput Genomics & Systems Biology

Universität des Saarlandes /

Max-Planck-Institut für Informatik

Telefon +49 681 9325 3115

mschulz@mmci.uni-saarland.de

Infektionskrankheiten

Infektionskrankheiten sind für die medizinische Forschung eine ständige Herausforderung. Der Grund liegt darin, dass sich in der Natur seit frühester Zeit ein harter Wettbewerb zwischen den verschiedenen Spezies entwickelt hat. Dieser Wettbewerb ist auch heute noch in vollem Gange. Uns Menschen wird er besonders durch die Aktivität von krankheitserzeugenden Erregern bewusst. Der Erreger sucht seinen Vorteil im Infizieren oder Besiedeln des menschlichen Wirts. Der infizierte menschliche Organismus wehrt sich gegen den Erreger mithilfe seines Immunsystems und, in neueren Zeiten, unterstützt durch eine medikamentöse Therapie. Der Erreger wiederum reagiert darauf durch eine Anpassung an den Wirt. Diese Anpassung findet nicht vornehmlich im einzelnen Erreger statt, sondern vielmehr durch eine evolutive Veränderung des Genoms des Erregers in den folgenden Generationen. Sie bringen damit Eigenschaften hervor, die zukünftig für das Überleben im Wirt günstig sind. Durch die ständige Veränderung des Erregers müssen dann immer wieder neue therapeutische Strategien entwickelt werden – es findet ein Wettrennen zwischen dem Erreger und uns Menschen als Wirten statt.

Aufgrund ihrer hohen Relevanz für die menschliche Gesundheit aber auch aufgrund des beispielhaften evolutiven Szenarios, das ihnen zu Grunde liegt, stellen Infektionskrankheiten einen reichen Fundus an hoch interessanten und relevanten Forschungsproblemen dar. Das Zentrum für Bioinformatik hat dieses Gebiet zu einem seiner Schwerpunkte gewählt. Wir untersuchen Infektionskrankheiten aus einer Vielzahl von Perspektiven. Dabei steht allerdings das evolutive Wettrennen zwischen Wirt und Erreger im Vordergrund.

Im Zentrum untersuchen wir Infektionskrankheiten mit besonders hoher evolutiver Dynamik. Die höchste Dynamik hat dabei HIV vorzuweisen, das eine hohe Mutationsrate besitzt und sich daher rasant im Körper eines jeden Patienten weiter entwickeln kann. Im Zentrum wird HIV auf verschiedenen Ebenen beforscht. Zum einen untersuchen wir die molekulare Grundlage für das Krankheitsgeschehen, die aus einem Netzwerk zwischen Interaktionen von Molekülen (meistens Proteinen) des Virus und seines Wirts besteht. Diese Forschung ist ausgesprochen schwierig, da die evolutiven Veränderungen im Virus höchst subtil sind und an der Grenze dessen liegen, was heute mit dem Computer modelliert werden kann (siehe Seite 37). Deshalb können wir auf der Basis solcher Erkenntnisse bisher noch keine für den Patienten direkt verwertbaren Resultate vorweisen. Um trotzdem eine Wirkung für den Patienten zu erzielen, gehen wir historisch vor: Wir sammeln große Mengen von klinischen Daten über das Therapieansprechen von spezifischen HIV Varianten (weit über 100.000 Datensätze). Auf der Grundlage der plausiblen Hypothese, dass eine spezifische Virusvariante in der Zukunft

so reagieren wird, wie sie es in der Vergangenheit getan hat, können wir aus solchen Daten Vorhersagen über die virale Resistenz bei individuellen Patienten ableiten.

Diese Vorgehensweise ist die Grundlage für unsere erfolgreichen Softwaresysteme zur Analyse von viralen Resistenzen bei HIV, dem Hepatitis B Virus und dem Hepatitis C Virus (siehe Seite 56). Diese Softwaresysteme wurden in substantiellen nationalen und internationalen Konsortien entwickelt (siehe Seite 59). Die von denselben Konsortien gesammelten Datenmengen bilden auch einen reichen Fundus für die Analyse der epidemiologischen Prozesse bei der HIV-Infektion. Auch in diesem Bereich betreibt das ZBI Forschung (siehe Seite 58).

Die hohe evolutive Dynamik von HIV ist auch der Grund, warum bis heute keine Impfstoffe gegen dieses Virus existieren. Das Virus bietet ein zu bewegliches Ziel für jeden einzelnen Impfstoff. Darum arbeitet man heute daran, Impfstoffe zu entwickeln, die eine große Breite von verschiedenen HIV Varianten abdecken. Auch hier sind wieder molekulare Analysen der Schlüssel für einen Forschungserfolg (siehe Seite 57).

Im Vergleich zu Viren haben Bakterien ein wesentlich komplexeres Genom. Das ermöglicht ihnen auf vielfältigere Art, sich an einen Wirt oder an eine medikamentöse Therapie anzupassen. Aus diesem Grund folgt der Fortschritt im Bereich der Forschung über bakterielle Infektionen dem über virale Infektionen. Die Resistenzentwicklung bei Bakterien ist dabei jedem im Zusammenhang mit der zurückgehenden Effektivität von Antibiotika bewusst. Im Zentrum untersuchen wir solche Resistenzphänomene und nutzen bakterielle Eigenschaften aus, um neue Medikamente zu entwickeln (siehe Seite 60). Schließlich forscht das Zentrum an einem groß angelegten Projekt, das das Ziel hat, den bereits erfolgreichen Zugang zur bioinformatischen Analyse von viraler Resistenz auf Bakterien zu übertragen (siehe Seite 61).



Infektionskrankheiten

Bioinformatische Unterstützung von HIV-Therapie	56
Vakzine und potente Antikörper gegen HIV	57
Transmissionsnetze und Epidemiologie.....	58
Konsortien	59
Entdeckung neuer Antibiotika	60
Resistenz von Bakterien und Entwicklung neuer Wirkstoffe.....	61

Bioinformatische Unterstützung von HIV-Therapie

Gefördert durch das Bundesministerium für Gesundheit und das Bundesforschungsministerium

56

Die Vermeidung und Kontrolle von viraler Resistenz ist das zentrale Ziel bei der Auswahl von medikamentösen Therapien gegen virale Infektionen. HIV, zum Beispiel, gibt es in Millionen von Varianten. Das HI Virus integriert sein Genom in das Genom der befallenen menschlichen Zelle. Darum ist es bisher aus dem Körper des infizierten Menschen nicht zu eliminieren. Der einzige therapeutische Ansatz ist, die Vervielfältigung der Viren im Körper des Patienten zu unterbinden. Zu diesem Zweck werden Kombinationen von Medikamenten aus einer Grundmenge von über zwei Dutzend Medikamenten zusammengestellt (Abbildung 20). Die wesentliche Information für die Auswahl der Medikamente ist der virale Genotyp, also die Variante des Genoms des im Patienten vorkommenden Virus, der mit Methoden der Genomsequenzierung aus Blutproben des Patienten ermittelt werden kann. Aus dem Genotyp kann mit bioinformatischen Methoden der virale Resistenz-Phänotyp, also die Reaktion des Virus auf das Medikament, ermittelt werden. Hierbei gibt es zwei Ansätze. Der erste besteht darin, von Experten manuell bestimmte Regeln zur Ermittlung des Phänotyps aus dem Genotypen in rechnergestützten Expertensystemen zur Anwendung kommen zu lassen. Expertensysteme sind Sammlungen von Experten manuell zusammengestellter Regeln, mit denen aus dem Vorkommen oder der Abwesenheit von Mutationen im Virusgenom auf die virale Resistenz geschlossen wird. Der zweite, systematischere Ansatz besteht in der Ermittlung des viralen Phänotyps aus einer geeigneten Menge klinischer Daten über die virale Resistenz mittels statistischer Methoden, die der Bioinformatik zuzuordnen sind.

Am Zentrum für Bioinformatik verfolgen wir den zweiten Ansatz. Unsere Arbeit der letzten zwölf Jahre hat zu dem geno2pheno System geführt, das im Internet unter www.geno2pheno.org frei verfügbar ist und in Deutschland, Europa und darüber hinaus zur Behandlung von HIV-Patienten eingesetzt wird. Analyseangebote des geno2pheno Servers haben Eingang in die europäischen Richtlinien zur Behandlung von HIV Patienten mit bestimmten Wirkstoffen gefunden. Im Jahr 2010 wurde die Arbeit an geno2pheno mit dem AIDS Forschungspreis der Heinz-Ansmann Stiftung gewürdigt. Die Klasse der Analyseangebote von geno2pheno, die in die klinische Praxis Eingang gefunden haben, kann man als virtuelle Phänotypen bezeichnen.

Unter einem virtuellen Phänotyp versteht man eine bioinformatische, also rechnergestützte Prozedur, die das Ergebnis eines Laborexperiments schätzt, das für die Behandlung des

Patienten informativ ist, also als Begleitdiagnostik bei der Medikamentenauswahl dienen kann:

Das Laborexperiment selbst ist dabei in der Regel in der Klinik nicht einsetzbar, sei es, weil es zu teuer, zu langwierig, oder weil es unzugänglich ist. Deshalb wird ein solches Experiment nur in einem begrenzten Forschungsszenario durchgeführt, um eine ausreichende Datenmenge zu erhalten, auf deren Basis der virtuelle Phänotyp entwickelt werden kann. Geno2pheno stellt diverse virtuelle Phänotypen zur Verfügung, die als Begleitdiagnostik für die Verabreichung von bestimmten AIDS-Medikamenten dienen.

Neben diversen Resistenz Phänotypen von HIV bietet der geno2pheno Server auch Resistenzanalysen für das Hepatitis B und das Hepatitis C Virus an. Insbesondere für Hepatitis C wird derzeit eine Vielzahl von vielversprechenden Medikamenten entwickelt. Während gegenwärtig ein gewisser Optimismus vorherrscht, dass diese Medikamente das Hepatitis C Virus aus dem Körper entfernen können, bevor sich Resistenzen entwickeln, zeigt die langfristige Erfahrung, dass nach einer gewissen Zeit des Einsatzes solcher Medikamente Resistenzen trotzdem entstehen werden. Weil für eine statistische Analyse bei Hepatitis B und Hepatitis C Viren bisher noch die Daten fehlen, basieren diese Resistenzanalysen auf Expertenregeln. Alle Serverangebote werden kontinuierlich aktualisiert, weiter entwickelt und durch neue Analyseangebote ergänzt. Alle Analysen sind kostenfrei verfügbar.

Die molekulare Basis von viraler Resistenz ist ein komplexes Mysterium, zu dem wir mit Fallstudien sowie durch Grundlagenforschung beitragen (siehe Seite 37).

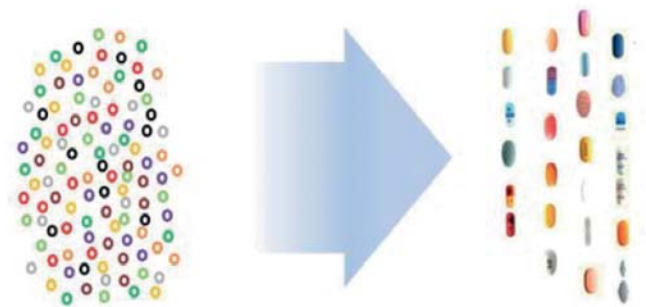


Abbildung 20 // Millionen von HIV-Varianten (links) stehen hunderte von Kombinationstherapien aus über zwei Dutzend Wirkstoffen (rechts) gegenüber. Der Rechner hilft bei der Auswahl der Medikamentenkombination.

Referenzen

- [1] T. Lengauer, N. Pfeifer and R. Kaiser. **Personalized HIV therapy to control drug resistance.** *Drug Discov Today Technol* 11: 57-64. 2014



Prof. Dr. Dr. Thomas Lengauer

**Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik**

Telefon +49 681 9325 3000
lengauer@mpi-inf.mpg.de

Vakzine und potente Antikörper gegen HIV

Gefördert durch das Bundesministerium für Bildung und Forschung (BMBF)

Da eine HIV-Infektion bisher nicht heilbar ist, konzentriert sich ein Großteil der HIV-Forschung darauf, die Symptome der infizierten Patienten mit Hilfe von antiretroviralen Medikamenten zu lindern und deren Leben zu verlängern. Die weltweite HIV-Epidemie hat jedoch Ausmaße erreicht, die darüber hinaus effektivere Vorsorgemaßnahmen erfordern. Denn nach einem HIV-Impfstoff wird seit langem gesucht. Die hohe Variabilität des Virus macht dieses Problem zu einer großen Herausforderung. In den letzten Jahren ist es jedoch zu einigen wichtigen Fortschritten gekommen. So gelang es zum Beispiel Walker und Kollegen, Antikörper zu isolieren, die sehr viele verschiedene HIV-Stämme binden und damit neutralisieren können.

Vorhersage der Resistenz verschiedener HIV-Varianten gegen passive Antikörpertherapie

Vor kurzem wurde bei mit Immunsystemzellen menschlichen Ursprungs ausgestatteten, sogenannten humanisierten Mäusen sowie Affen gezeigt, dass eine medikamentöse Therapie mit einer Kombination von breit neutralisierenden Antikörpern (bNAbs) eine Behandlungsoption sein kann. Eine klinische Studie, bei der einer dieser bNAbs bei Menschen getestet wurde, hat bereits erste positive Ergebnisse erzielt. Dies legt nahe, dass eine HIV-Therapie mit bNAbs in naher Zukunft möglich sein könnte.

Um über die Therapie für einen HIV-infizierten Patienten zu entscheiden, werden in Deutschland zunächst die wichtigen Bereiche des HIV-Genoms der Virusvarianten der Patienten sequenziert. Wichtige Bereiche sind solche, die darüber Auskunft geben, welche antiviralen Medikamente noch aktiv sind, d. h., gegenüber welchen Medikamenten das Virus noch keine Resistenzen entwickelt hat.

Wenn die Antikörpertherapie mit bNAbs bei HIV-infizierten Patienten zugelassen wird, dann müsste man mit einer ähnlichen Begleitdiagnostik entscheiden können, welche Antikörper noch gegen die jeweiligen Virusvarianten aktiv sind. Zu diesem Zweck haben wir bioinformatische Vorhersagemodelle entwickelt, die hochgenau abschätzen können, ob eine bestimmte Virusvariante resistent gegen einen bNAb ist oder nicht [1]. Zusätzlich haben wir Methoden entwickelt, welche die Resultate solcher Vorhersagen besser verständlich machen (Abbildung 21).



Abbildung 21 // Motif Logo für die Klassifikation eines Testbeispiels mit unbekanntem Resistenzpotential gegen einen bestimmten bNAb. Das Logo zeigt die Positionen an, die für die Testsequenz eines viralen Proteins am informativsten für die Vorhersage waren. In diesem Beispiel sind die wichtigsten 5% aller Positionen berücksichtigt. Blaue Großbuchstaben repräsentieren Aminosäuren in der Testsequenz, die stark mit einer positiven Klassifikation assoziiert sind (HIV-Variante ist nicht gegen bNAb resistent). Orangefarbene Kleinbuchstaben repräsentieren Aminosäuren in der Testsequenz, die für eine Klassifikation als resistent sprechen.

Auch wenn die Anwendung einer Antikörpertherapie für HIV-infizierte Patienten noch ein paar Jahre in der Zukunft liegt, können unsere Verfahren schon jetzt eingesetzt werden, um die Planung zukünftiger klinischer Studien zu unterstützen.

Analyse verschiedener Antikörper gegen HIV

Ein wichtiger Schritt in Richtung eines universellen Impfstoffes gegen HIV ist die Untersuchung der Eigenschaften von Antikörpern, die das Virus neutralisieren können. Um die Wirksamkeit potentieller Kandidaten zu untersuchen, werden im Labor Tests durchgeführt, die zeigen, welche HIV-Stämme von einem bestimmten Antikörper neutralisiert werden können. Hierzu werden Virenstämme verwendet, welche die Diversität der verschiedenen in der Welt vorkommenden Varianten so gut wie möglich abdecken. Unsere Analysen bilden die Grundlage für eine genauere Bewertung von experimentellen Ergebnissen über die Qualität von Vakzinen [2]. Die Forschungsergebnisse können dazu beitragen, die Antikörpertests zu verbessern und sind damit ein wichtiger Schritt auf dem Weg zur Entwicklung eines universell einsetzbaren Impfstoffes gegen HIV.

Referenzen

- [1] A. Feldmann and N. Pfeifer. **From Predicting to analyzing HIV-1 resistance to broadly neutralizing antibodies.** *German Conference on Bioinformatics 2015*. 2015
- [2] N. Pfeifer, H. Walter and T. Lengauer. **Association between HIV-1 coreceptor usage and resistance to broadly neutralizing antibodies.** *J Acquir Immune Defic Syndr* 67 (2): 107-12. 2014



Anna Feldmann

**Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik**

Telefon + 49 681 9325 3027
feldmann@mpi-inf.mpg.de



Dr. Nico Pfeifer

**Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik**

Telefon + 49 681 9325 3019
npfeifer@mpi-inf.mpg.de

Transmissionsnetze und Epidemiologie

58

Im Bereich der Epidemiologie sind auf Netzwerken basierende Modelle von fundamentaler Bedeutung. Sie ermöglichen es, die Verbreitung einer übertragbaren Krankheit in der Bevölkerung abzubilden. Diese Erkenntnisse sind wichtig, um die Ausbreitung von Krankheiten überwachen und kontrollieren zu können. Unsere Arbeitsgruppe erforscht diesen Bereich sowohl mit praktischen als auch mit theoretischen Untersuchungen.

Um die Dynamik der HIV-Epidemie zu verstehen und mögliche Vorsorge gegen Neuinfektionen zu treffen, ist es wichtig, die Übertragungsmuster des Virus identifizieren zu können. Das Virus verändert sich etwa so schnell, wie es von Patient zu Patient weitergereicht wird, d.h. die DNA-Sequenz eines Virus kann Hinweise auf dessen Übertragungsweg geben. Wir haben diesbezüglich eine Studie mit 21844 HIV-1-Sequenzen (Subtyp B) aus dem EuResist Konsortium durchgeführt. Basierend auf der evolutionären Distanz zwischen den DNA-Sequenzen wurde ein Übertragungsnetzwerk erstellt (siehe Abbildung 22). 4775 Sequenzen hatten eine hohe genetische Ähnlichkeit zu mindestens einer anderen Sequenz des Datensatzes. In den meisten Fällen kamen beide ähnlichen Sequenzen aus demselben Land, was darauf hindeutet, dass HIV-1-Übertragungsnetzwerke sehr spezifisch für die einzelnen Länder sind. Diese aktuell größte europaweite Studie lieferte auch kongruente Ergebnisse zu anderen Studien, die sich auf einzelne Länder, wie Deutschland, Italien und die Schweiz, beziehen [1].

Ein epidemiologischer Ansatz hilft uns außerdem, eine wichtige, noch offene Frage in der Netzwerktheorie zu beantworten. Frühe Arbeiten in der Netzwerkforschung beschäftigten sich damit, basierend auf der Topologie des Netzwerks dessen wichtigsten Knoten zu identifizieren. Z. B. ist der Knoten mit der größten Anzahl an Verbindungen oft der wichtigste eines Netzwerkes. Allerdings stellte man im letzten Jahrzehnt zunehmend fest, dass dies nicht auf den Rest des Netzwerkes übertragbar ist, d. h. die Anzahl an Verbindungen gibt im Allgemeinen keinen Aufschluss auf die Bedeutung des Knotens für das gesamte Netzwerk.

Wir messen den Einfluss eines Knotens mittels der erwarteten Anzahl an Infektionswegen, die die lokale Nachbarschaft des Knotens verlassen. Dieses Maß wird *expected force* genannt, da es die erwartete Kraft einer Infektion beschreibt, die von einer Epidemie mit Ursprung in diesem Knoten ausgehen würde. Umfassende Tests zeigen, dass mit Hilfe der *expected force* der Ausgang einer Epidemie für unterschiedliche Netzwerktopologien und Ausbreitungsprozesse sehr genau vorhergesagt werden kann [1]. In unserer jüngsten Studie

evaluierten wir die Vorhersagekraft der Methode bzgl. einer pandemischen Verbreitung über das weltweite Netzwerk von Flugbewegungen. Um zu simulieren, wie sich eine Krankheit global ausbreitet, wurden hier realistische Modelle verwendet, die die lokale Bevölkerungsdynamik ebenso berücksichtigen wie die Nutzungsmuster für bestimmte Reiselinien. Die *expected force* des Flughafens, an dem die Krankheit ausbrach, korrelierte zu 90% mit der Virulenz, die der Virus benötigte, um sich global auszubreiten und zu 87% mit der Geschwindigkeit, mit der die Krankheit den Status einer Pandemie erreichte. [2].

Simulationen zeigen die Verzögerung vom Ausbruch der influenza-artigen Krankheit bis sie eine Pandemie wird. Die *expected force* des Flughafens in der Stadt des Krankheitsausbruchs korreliert zu 85% mit dieser Zeitspanne. In einer Welt, die zunehmend durch Netzwerke definiert ist, erlauben uns Netzwerkmodelle für Krankheiten daher eine verbesserte Gesundheitsvorsorge.

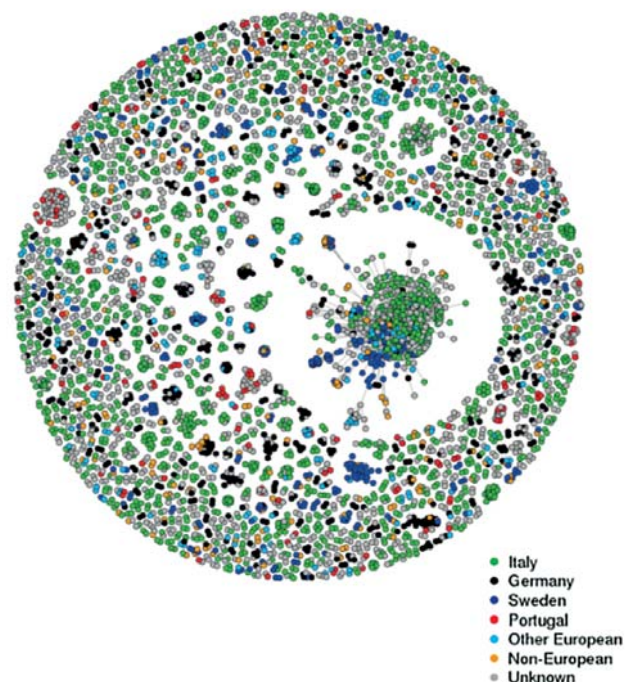


Abbildung 22 // Europaweites HIV-Übertragungsnetzwerk: Ein Übertragungsnetzwerk konstruiert mittels ausgewählter Grenzwerte für die evolutionäre Distanz. Die Knotenfarbe zeigt das Herkunftsland des infizierten Patienten.

Referenzen

- [1] D. Frenzt, A. M. Wensing, J. Albert et al. **Limited cross-border infections in patients newly diagnosed with HIV in Europe.** *Retrovirology* 10: 36. 2013
- [2] G. Lawyer. **Understanding the influence of all nodes in a network.** *Sci Rep* 5: 8665. 2015



Dr. Glenn Lawyer

**Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik**

Telefon +49 681 9325 3007

lawyer@mpi-inf.mpg.de

Konsortien

Gefördert durch Bundesministerium für Bildung und Forschung (BMBF), Bundesministerium für Gesundheit (BMG), Deutsche Forschungsgemeinschaft (DFG) und Europäische Union (EU)

Insbesondere im Bereich der Bioinformatik der Infektionskrankheiten hat sich das Zentrum für Bioinformatik zum Ziel gesetzt, über die Grundlagenforschung hinauszugehen und einige seiner Ergebnisse in die klinische Anwendung zu bringen. Im Bereich der Medikamentenentwicklung ist es aber kaum möglich, dies in einem überschaubaren Zeitrahmen zu erreichen. Die Medikamentenentwicklung, von der Suche nach der ersten Leitstruktur bis zu den klinischen Studien vor Zulassung und Markteinführung, findet in einem Zeitrahmen von deutlich über einer Dekade statt und verschlingt mehrere 100 Millionen Euro pro neuem Medikament. Hier kann die Grundlagenforschung nur den ersten Anstoß geben, auf den ein langer, teurer und risikoreicher Weg der Entwicklung, Validierung und Zertifizierung folgt. Im Falle von Infektionskrankheiten, bei denen der Erreger auf eine applizierte medikamentöse Therapie mit der Entwicklung von Resistenz gegen das Medikament reagiert, ist die Situation günstiger. Hier ist im besten Falle ein ansehnliches Wirkstoff-Arsenal verfügbar. Die Therapie wird individuell auf den Patienten zugeschnitten, indem aus den verfügbaren Wirkstoffen eine geeignete Kombination ausgewählt wird. Die Information, die für die Auswahl der Medikamente benötigt wird, betrifft den Resistenz-Phänotyp des Erregers, d.h. die zu erwartende Reaktion des Erregers auf das Medikament. Diese Information wird mit heutigen Verfahren durch Analyse der Mutationen im Virusgenom gewonnen. Zusätzlich verwendet man zur Wirkstoffauswahl Informationen über Charakteristika des Patienten und seiner Krankheitsgeschichte. Bei dieser Methode sind alle infrage kommenden Medikamente bereits zugelassen. Für die Auswahl der Kombination selbst gibt es einen Prozess, der vergleichsweise schnell an den Patienten herangeführt werden kann.

Ist ein ausreichend großer Satz klinischer Daten über die Resistenz des Erregers vorhanden, sind unsere statistischen Ansätze für die Vorhersage der Resistenz beim Patienten den regelbasierten Ansätzen überlegen. Die Erhebung solcher Datensätze ist ein aufwändiges Unterfangen, an dem wir im Rahmen von nationalen und internationalen Konsortien teilnehmen. Die Konsortien haben eine dezidiert klinische Ausrichtung und bestehen aus Klinikern, Virologen und Bioinformatikern.

Der Erfolg einer bioinformatischen Software zur Vorhersage von Wirkstoffresistenz beruht jedoch nicht allein auf der Entwicklung und Validierung effektiver Software. Eine tiefgehende Kenntnis der zu Grunde liegenden Daten, wie sie Bioinformatiker letztlich nur im Dialog mit Klinikern und Virologen erzielen können, ist unbedingte Voraussetzung. Ferner muss die Validierung der Analyseverfahren von Bioinformatikern in Kooperation mit Virologen und Klinikern stattfinden. Schließlich stehen Biologen, Kliniker und Bioinformatiker in einem ständigen Gedankenaustausch und geben so den Anstoß für weitere Entwicklungen.

Im Zentrum für Bioinformatik schätzen wir uns glücklich, zwei höchst effektiven Konsortien dieser Art zum Thema der HIV-Resistenz angehören zu dürfen. Auf deutscher Ebene ist dies das Arevir-Konsortium [2], das sich im Jahre 2000 mit dem Ziel der Entwicklung des *geno2pheno* Servers zusammengefunden hat. Seine initiale Phase im Rahmen eines Schwerpunktprogramms der Deutschen Forschungsgemeinschaft wurde gefördert. Institutioneller Rahmen für das Arevir Konsortium ist die Gesellschaft für nachhaltige Forschung (*genafor*) (www.genafor.org). Auf europäischer Ebene hat sich darüber hinaus im Jahr 2004 das *EuResist* Konsortium (www.euresist.org) gebildet [1], dessen Kernteam aus Virologen, Klinikern und Bioinformatikern aus Italien, Schweden und Deutschland besteht. Auch das *EuResist* Team hat eine Organisationsform für sich gefunden, die Europäische Wirtschaftliche Interessenvereinigung (EWIV). Für beide Konsortien bezeichnend ist, dass sie nach Auslaufen der ersten Förderung, die jetzt zehn bzw. sieben Jahre zurückliegt, weiter bestehen und weder an Energie noch an Intensität verloren haben. Bezeichnend dafür ist, dass *EuResist* mehrfach als Einheit in nachfolgenden EU-Anträgen und -Projekten aufgetreten ist. *Arevir* veranstaltet alljährlich ein immer mehr beachtetes internationales Symposium, das „*Arevir Meeting*“, das sich inzwischen zu einer Marke entwickelt hat.

Die partnerschaftliche und effektive Kooperation in diesen beiden Konsortien ist die Grundlage für den Erfolg unserer bioinformatischen Software im Bereich der viralen Resistenzanalyse. In diesem Rahmen sind wir auch Mitglied weiterer Konsortien zur medizinischen Datensammlung und Dateninterpretation, etwa des europäischen *EUCOHIV* Konsortiums (www.eucohiv.org).

Referenzen

- [1] M. Zazzi, F. Incardona, M. Rosen-Zvi, M. Prosperi, T. Lengauer, A. Altmann, A. Sonnerborg, T. Lavee, E. Schuster and R. Kaiser. **Predicting response to antiretroviral treatment by machine learning: the *EuResist* project.** *Intervirology* 55 (2): 123-7. 2012
- [2] K. Roomp, N. Beerenwinkel, T. Sing, E. Schülter, J. Büch, S. Sierra-Aragon, M. Däumer, D. Hoffmann, R. Kaiser, T. Lengauer and J. Selbig. **Arevir: A secure platform for designing personalized antiretroviral therapies against HIV.** *Lecture Notes in Computer Science: Third International Workshop on Data Integration in the Life Sciences (DILS 2006)* 4075: 185-194. 2006



Prof. Dr. Dr. Thomas Lengauer

Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik

Telefon +49 681 9325 3000
lengauer@mpi-inf.mpg.de

Entdeckung neuer Antibiotika

Gefördert durch das BMBF im Rahmen des Deutschen Zentrums für Infektionsforschung

60

Bakterielle Krankheitserreger, wie zum Beispiel *Enterococcus faecium*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumannii*, *Pseudomonas aeruginosa* und *Enterobacter spp.*, stellen eine bedeutende Gefahr für die Gesundheit dar. Insbesondere sogenannte Krankenhauskeime, die vor allem ältere und immungeschwächte Patienten gefährden, sind oft resistent gegen gängige Antibiotika. Der traditionelle Ansatz, indem große Wirkstoffbibliotheken gegen ein bekanntes Zielprotein gescreent werden, wird im Bereich der Antibiotikaforschung oft zu kostspielig und zu zeitintensiv, weswegen Alternativen gesucht werden. Von Mikroorganismen produzierte Naturstoffe sind hierbei ein sehr attraktiver Ansatzpunkt. Eine Gruppe von Bodenbakterien, die sogenannten Myxobakterien, produzieren sehr viele solcher Naturstoffe, wobei einige davon eine starke antimikrobielle Aktivität und einzigartige Wirkmechanismen aufweisen. In unserem Projekt werden experimentelle und computergestützte Methoden miteinander kombiniert, um antibakterielle Wirkmechanismen von myxobakteriellen Naturstoffen zu untersuchen. In einer Fallstudie wurde Disciformycin, ein neuartiges Antibiotikum aus dem Myxobakterium *Pyxidiccoccus fallax* erforscht, welches gegen den weit verbreiteten Krankenhauskeim *Staphylococcus aureus* wirkt [1]. Wir haben zunächst Disciformycin-resistente *S. aureus* Stämme im Labor erzeugt. Dann wurden die kompletten Genome der resistenten Staphylokokken und das des ursprünglichen sensitiven *S. aureus* Stammes sequenziert. Schließlich wurde nach Mutationen gesucht, d.h. nach einzelnen Basenpaaren, die im Bakteriengenom ausgetauscht wurden, und die womöglich zur Entstehung von Antibiotikaresistenz führen.

In unserem Fall konnten wir zeigen, dass die Mutationen, die zu einem Aminosäureaustausch innerhalb der Proteinsequenz führen, in genau zwei Genen liegen. Diese kodieren wiederum für Untereinheiten desselben Enzyms, der DNA-abhängigen RNA-Polymerase. Sie wandelt DNA-Gensequenzen in mRNA-Moleküle um und übernimmt somit einen elementaren Schritt der Proteinbiosynthese. Durch rechnergestützte Modellierung dieser Mutationen in ein dreidimensionales Modell der RNA-Polymerase, konnten wir zeigen, dass die Veränderungen am Enzym an der Interaktionsfläche zwischen den beiden Untereinheiten – dem sogenannten „Brücken-Helix“ Motiv – auftreten (Abbildung 23).

Da Antibiotikaresistenz-assoziierte Mutationen häufig dazu führen, dass der Wirkstoff nicht mehr an sein Zielprotein binden kann, liegt die Vermutung nahe, dass Disciformycin in der oben beschriebenen Region bindet und dadurch die Funktion der RNA-Polymerase stört. Weitere Modelle von Wild-



Abbildung 23 // Dreidimensionales Strukturmodell des RNA-Polymerase-Komplexes mit Disciformycin-assoziierten Resistenzmutationen (Violett) und „Brücken-Helix“ (Rot).

typ- und mutierter RNA-Polymerase mit gebundenem Disciformycin unterstützten diese Theorie. Des Weiteren haben wir unsere Hypothese experimentell getestet und konnten bestätigen, dass Disciformycin tatsächlich die RNA Synthese in Wildtyp-*S. aureus* inhibiert, jedoch in den mutierten Stämmen unwirksam ist. Außerdem konnte gezeigt werden, dass es zu keiner Kreuzresistenz mit anderen häufig verwendeten Antibiotika, z. B. Rifampicin, kommt. In weiteren in vitro Studien konnten wir nachweisen, dass Disciformycin die RNA-Polymerase direkt inhibiert, was letztlich den vorgeschlagenen Wirkmechanismus eindeutig bestätigte. Aktuelle Untersuchungen beschäftigen sich vor allem mit der experimentellen Darstellung der dreidimensionalen Struktur von Disciformycin im Komplex mit RNA-Polymerase über Kristallisation und Röntgenstrukturanalyse.

In Zukunft möchten wir den beschriebenen Ansatz weiter ausbauen: Es sollen große klinische Datensammlungen zu Antibiotikaresistenz-assoziierten Mutationen als Ausgangspunkt für eine automatisierte Modellierung und funktionelle Charakterisierung verwendet werden (siehe Seite 61).



Prof. Dr. Rolf Müller
Helmholtz-Institut für Pharmazeutische
Forschung Saarland (HIPS)
Universität des Saarlandes
Telefon +49 681 98806 3000
rom@mx.uni-saarland.de



Dr. Olga Kalinina
Abt. Bioinformatik und Angewandte Algorithmik
Max-Planck-Institut für Informatik
Telefon +49 681 9325 3004
kalinina@mpi-inf.mpg.de

Referenzen

- [1] F. Surup, K. Viehrig, K. I. Mohr, J. Herrmann, R. Jansen and R. Müller. **Disciformycins A and B: 12-membered macrolide glycoside antibiotics from the myxobacterium *Pyxidiccoccus fallax* active against multiresistant staphylococci.** *Angew Chem Int Ed Engl* 53 (49): 13588-91. 2014

Resistenz von Bakterien und Entwicklung neuer Wirkstoffe

Gefördert durch Bundesministerium für Bildung und Forschung (BMBF) und Deutsches Zentrum für Infektionsforschung

Die ständig zunehmende Resistenz von Bakterien gegen verfügbare antibakterielle Wirkstoffe stellt eines der größten Probleme in den Gesundheitssystemen in der gesamten Welt dar. Der vielfache und teilweise unkritische Einsatz von Antibiotika hat dazu geführt, dass viele Bakterien inzwischen Multiresistenzen evolviert haben und auch miteinander austauschen. Parallel dazu hat zudem in den letzten 30 Jahren die Entwicklung und Zulassung neuer Antibiotika signifikant abgenommen. Ziel muss es daher sein, aus der momentan schlechten Ausgangslage mittel- bis langfristig das beste Ergebnis für Patienten zu erzielen. Dazu setzt das Zentrum für Bioinformatik an zwei Punkten an.

Zu diesem Zweck werden neue genetische Tests mit dem Ziel erarbeitet, eine schnellere Analyse von Resistenzen in infektiösen Pathogenen zu ermöglichen (siehe Abbildung 24). Bei vielen Krankheiten, die auf Infektionen mit Bakterien zurückgehen, kann ein solch schneller Test entscheidend sein. Eines der wohl besten Beispiele ist die bakterielle Sepsis. Hier ist das richtige Behandeln in den ersten Stunden oft entscheidend für das Überleben der Patienten. Gemeinsam mit Siemens Healthcare, dem Helmholtz Institut für Pharmazeutische Forschung (HIPS) und der Medizinischen Mikrobiologie der UdS arbeitet das Zentrum für Bioinformatik daran, die genetischen Ursachen der Resistenz von Bakterien besser zu verstehen. Dieses Wissen wird in Zukunft die Grundlage einer verbesserten Diagnostik bilden.

Um das bestmögliche Ergebnis zu erzielen, ist es essenziell, die verschiedenen Kompetenzen am Zentrum für Bioinformatik zu bündeln. Insgesamt arbeiten sechs Gruppen am Zentrum gemeinsam an diesem Thema. Geleitet wird die Forschung vom Lehrstuhl für Klinische Bioinformatik (Prof. Keller). Zwei Gruppen am Max-Planck-Institut (Dr. Kalinina und Dr. Pfeifer) sowie Juniorprofessor Tobias Marschall erforschen spezifische Aspekte der genetischen Ursachen für Resistenzen. Die klinische Expertise wird vom Zentrum für Mikrobiologie und Hygiene (Prof. Hermann) beigesteuert. In dieser Kooperation steht die Erforschung des in Krankenhäusern verbreiteten gefährlichen Bakteriums MRSA (multiresistenter *Staphylococcus aureus*) im Vordergrund. Die biochemische Analyse und Validierung der Ergebnisse erfolgt am HIPS unter Leitung von Prof. Müller.

Der zweite Forschungsschwerpunkt am ZBI in diesem Bereich ist, die Patientenversorgung bei Infektionskrankheiten durch die Entwicklung neuer antimikrobieller Wirkstoffe zu verbessern. Unsere Arbeiten haben das Ziel, innovative Zielmoleküle zu entdecken, welche nachfolgend mit Wirkstoffen so adressiert werden können, dass diese einen antibakteriellen oder auch einen der Pathogenität der Bakterien unterbindenden Effekt haben. Dabei spielt die genetische Analyse der Selbstresistenz von mikrobiellen Produzenten von Antibiotika – also der Fähigkeit des Bakteriums, dem von ihm selbst erzeugten Antibiotikum zu widerstehen – eine ebenso wichtige Rolle wie die der Resistenzbildung in den Pathogenen [1, 2]. Diese Analytik findet mittlerweile praktisch ausschließlich sequenzbasiert und mit Hilfe von bioinformatischen Analysen statt, durch deren Ergebnisse die nachfolgende Experimente im mikrobiologisch/biochemischen Labor priorisiert und konfiguriert werden.

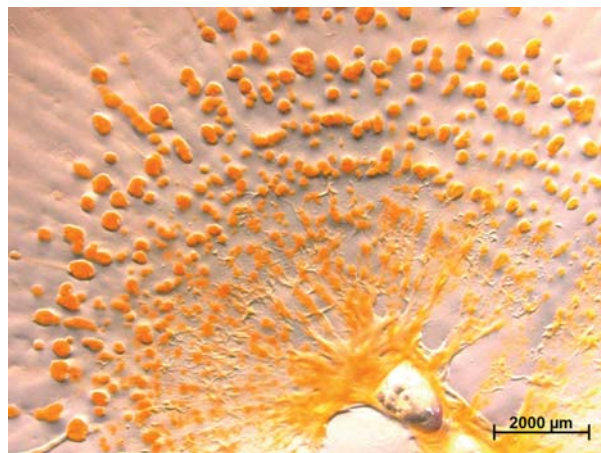


Abbildung 24 // Myxobakterien aus dem Genus *Sorangium*, welche aus Bodenproben mit verrottendem Pflanzenmaterial isoliert werden und herausragende Produzenten von Antibiotika sind. Die Abbildung zeigt kreisförmig angeordnete Fruchtkörper einer Bakterienspezies auf Agarmedium. Darin enthalten sind dicht gepackte sowie hitze- und trocknungsresistente Myxosporen als Dauerformen.

Referenzen

- [1] A. Kling, P. Lukat, D. V. Almeida, A. Bauer et al. **Antibiotics. Targeting DnaN for tuberculosis therapy using novel griselimycins.** *Science* 348 (6239): 1106-12. 2015
- [2] S. Baumann, J. Herrmann, R. Raju, H. Steinmetz, K. I. Mohr, S. Huttel, K. Harmrolfs, M. Stadler and R. Müller. **Cystobactamids: myxobacterial topoisomerase inhibitors exhibiting potent antibacterial activity.** *Angew Chem Int Ed Engl* 53 (52): 14605-9. 2014



Prof. Dr. Rolf Müller
Helmholtz-Institut für Pharmazeutische
Forschung Saarland (HIPS)
Universität des Saarlandes
Telefon +49 681 98806 3000
rom@mx.uni-saarland.de



Prof. Dr. Andreas Keller
Klinische Bioinformatik
Universität des Saarlandes
Telefon +49 681 302 68611
andreas.keller@ccb.uni-saarland.de

Krebs

Trotz der beachtlichen Fortschritte der Onkologie in den letzten Jahrzehnten ist Krebs auch heute noch für 13% aller Todesfälle weltweit verantwortlich und damit direkt nach kardiovaskulären Erkrankungen eine der Haupttodesursachen. Im Saarland werden bis zum Alter von 74 Jahren 41.1% aller Männer und 27.9% aller Frauen an Krebs erkranken. Bei Männern stellen Prostata-, Lungen-, Darm- und Blasenkrebs und bei Frauen Brust-, Darm-, Lungen- und Zervix-Karzinome die häufigsten Krebserkrankungen dar. Aufgrund des demographischen Wandels, mit einer immer älter werdenden Bevölkerung, wird in Zukunft die absolute Anzahl von Krebserkrankungen noch zunehmen, da Krebs eine typische Erkrankung des älteren Menschen ist.

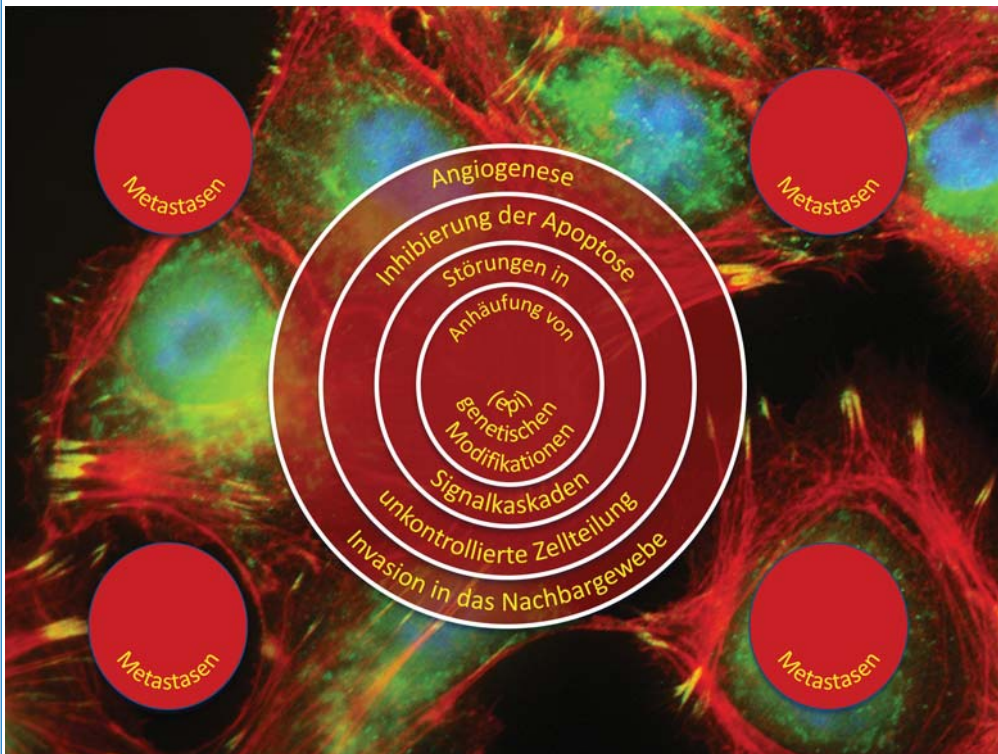
Die Ursachen von Krebs sind multifaktoriell. Sowohl genetische Aberrationen als auch Umwelteinflüsse und Lebensstil tragen zur Entstehung von Krebs bei. Heute beruht die Behandlung immer noch weitgehend auf Operation, Bestrahlung und Chemotherapie. Obwohl die meisten Patienten auf diese Behandlungen ansprechen, werden nicht alle geheilt. Ein Hauptgrund hierfür ist die Heterogenität von Tumoren, die vor allem auf eine in Tumoren ablaufende „zelluläre Evolution“ zurückzuführen ist. Diese wird wiederum durch hohe Zellteilungs- und Mutationsraten und ein Versagen der zellulären Reparaturmechanismen ausgelöst. Vereinfacht dargestellt bedeutet dies, dass ein Tumor sich in der Regel aus genetisch unterschiedlichen Tumorzellen zusammensetzt. Viele dieser Zellen können durch eine Therapie getötet werden, andere Tumorzellen wiederum sind entweder bereits resistent gegen die Therapie oder aus ihnen können sich durch das Auftreten neuer Mutationen resistente Zellen bilden.

Die Heterogenität von Tumoren hat auch zur Folge, dass Tumore vom gleichen Typ sich in ihrem Geno- und Phänotyp sehr stark unterscheiden können. Dies bedeutet auch, dass jeder Tumor eine eigens auf ihn abgestimmte und optimierte Therapie benötigt, die auch gegebenenfalls nach dem Auftreten Resistenz verursachender Mutationen wieder adaptiert werden muss. Diese Art von personalisierter Medizin basiert auf einer detaillierten genetischen und molekularbiologischen Charakterisierung der Tumorzellen. Eine solche Typisierung erfordert wiederum eine integrative bioinformatische Analyse verschiedener hochdimensionaler und extrem umfangreicher (epi-)genetischer und molekularbiologischer Hochdurchsatzdaten. Eine detaillierte molekulare Typisierung wird zu einer feineren Unterteilung der meisten Tumortypen in Subtypen führen, wie z.B. bei Brustkrebs schon geschehen, mit der Folge, dass dann die verschiedenen Subtypen des Tumors

unterschiedlich, entsprechend ihrem molekulargenetischen Profil behandelt werden können. Obwohl die intensive Krebsforschung der letzten Jahrzehnte unser Wissen über die Entstehung und Entwicklung von Krebs enorm vergrößert hat, sind wir jedoch noch weit davon entfernt, die komplexen Zusammenhänge zwischen den genetischen Modifikationen und ihren Effekten in Form von Änderungen der molekularen Signalflüsse und der biochemischen Prozesse in ihrer gesamten Komplexität zu verstehen.

Wir sprechen in diesem Zusammenhang von deregulierten Signalkaskaden und deregulierten molekularbiologischen Prozessen. Wie oben bereits diskutiert, führt eine detaillierte molekulare Typisierung nicht nur zu einem besseren Verständnis der pathogenen Prozesse, sondern sie liefert auch neue Ansätze für die Diagnose, Prognose und Therapie von Tumorerkrankungen. Das Zentrum für Bioinformatik hat eine Reihe von innovativen Methoden und Softwaretools entwickelt, die es ermöglichen, sehr detaillierte genetische und molekulare Typisierungen von Tumoren anhand von deregulierten Signalkaskaden und biologischen Prozessen zu erstellen (siehe Seite 64). Die beteiligten Arbeitsgruppen des ZBI haben mit Hilfe dieser Tools signifikante wissenschaftliche Beiträge zum besseren Verständnis der Entstehung und Entwicklung verschiedener Krebstypen geleistet, darunter Lungentumore, Glioblastome, Meningeome, Wilms Tumore etc. Da diese Tools aber auch von vielen anderen Gruppen weltweit zur Untersuchung von Tumoren (siehe Seite 42) eingesetzt wurden, haben diese bioinformatischen Werkzeuge auf vielfältige Art und Weise zur Aufklärung pathogener Mechanismen verschiedenster Krebsarten beigetragen.

Die Heilungschancen bei vielen Krebsarten ist immer noch stark davon abhängig, in welchem Entwicklungsstadium die Erkrankung diagnostiziert wird. Hier gilt immer noch die einfache Aussage: Je früher, desto besser. Bei der Früherkennung spielen neben bildgebenden Verfahren nicht-invasive, auf Biomarkern in Körperflüssigkeiten basierende Diagnoseansätze eine zentrale Rolle. Arbeitsgruppen des ZBI haben für eine Reihe von Krebserkrankungen neue, auf hochdimensionalen Profilen von Tumorantigenen und MicroRNAs basierende, diagnostische Ansätze entwickelt. Diese besitzen eine höhere Genauigkeit als die zurzeit zugelassenen Diagnoseverfahren. Neben zahlreichen Publikationen sind aus diesen Projekten auch eine Reihe von Patenten hervorgegangen (siehe Seite 65). Für ein Verfahren werden bereits die für eine Kommerzialisierung notwendigen Validierungsstudien durchgeführt. Sie zeigen bis dato sehr vielversprechende Ergebnisse.



Prozesse in der Tumorentwicklung

Eine weitere gute Nachricht im Kampf gegen Krebs ist, dass bereits mehr als 100 Anti-Tumor-Medikamente verfügbar sind und dass die Zahl der verfügbaren Medikamente weiterhin stetig steigt. Jedoch ist die Bestimmung einer optimalen Therapie, unter anderem aufgrund der in den vorhergehenden Abschnitten skizzierten Heterogenität der Tumore, immer noch eine große Herausforderung. Während bioinformatische Werkzeuge zur molekulargenetischen Typisierung von Tumoren heute bereits den Onkologen bei der Auswahl einer geeigneten Therapie unterstützen, werden sie in Zukunft die entscheidenden Werkzeuge bei der Therapieoptimierung sein, da sie nach Auswertung aller genetischen und molekularbiologischen Daten des Tumors optimale Therapieoptionen liefern werden. Auch in diesem Bereich ist die Forschung des ZBI international hochsichtbar (siehe Seite 66).

Krebs

Wie entsteht Krebs?	64
Biomarker und Diagnose.....	65
Therapieoptimierung.....	66

Wie entsteht Krebs?

Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) (Schwerpunktprogramm Scalable Visual Analytics 1335)

64

Ein tiefgehendes Verständnis der Auswirkungen der in Tumorzellen vorliegenden Mutationen und genetischen Modifikationen auf die molekulare Signalübertragung und die dadurch modifizierten biologische Prozesse eröffnet, wie im vorgehenden Abschnitt diskutiert, neue Optionen für diagnostische, prognostische und therapeutische Ansätze. Die Arbeitsgruppen des ZBI haben hier zu verschiedenen wissenschaftlichen Fragestellungen wesentliche Beiträge geleistet. Ein zentraler Fragenkomplex beschäftigt sich mit der Suche nach genetischen Aberrationen, die Krebs auslösen (siehe Abbildung 25). Hierzu werden weltweit zahlreiche Genome von Krebszellen und von gesunden Zellen sequenziert und verglichen, was heutzutage in der Regel die Auswertung von vielen Terrabyte an Sequenz-Daten erforderlich macht. Neben der Herausforderung durch die schiere Datenmenge wird das Problem dadurch erschwert, dass in einem Tumor neben den sogenannten „Driver Mutations“ auch noch viele „Passenger Mutations“ vorliegen können, die keinen wesentlichen Einfluss auf die Entwicklung des Tumors haben. Auch zu diesem Fragenkomplex haben die Arbeitsgruppen des ZBI international vielbeachtete wissenschaftliche Beiträge geleistet. Zum Beispiel haben Arbeitsgruppen des ZBI gemeinsam mit einem internationalen Forscherteam vor kurzem in der hochrangigen Zeitschrift „Cancer Cell“ neue Mutationsprofile von Blastemreichen Wilms Tumoren vorgestellt, welche Erklärungen für die Malignität dieses Subtyps liefern [1].

Ein weiterer Fragenkomplex beschäftigt sich mit den Auswirkungen der genetischen Aberrationen auf die Signalübertragung und die nachgeschalteten biologischen Prozesse in Krebszellen. Die Detektion von solchen deregulierten biologischen Prozessen, Kategorien und Netzwerken ist ein Forschungsschwerpunkt des ZBI. Wie auf Seite 42 ausführlich dargestellt, haben viele Forschergruppen weltweit die am ZBI entwickelten bioinformatischen Werkzeuge verwendet, um für verschiedenste Erkrankungen deregulierte Prozesse zu identifizieren und zu studieren, unter anderem auch für eine Reihe von Krebsarten.

Natürlich wurden diese von uns entwickelten innovativen bioinformatischen Werkzeuge auch in internen Projekten eingesetzt. Unter anderem haben Arbeitsgruppen des ZBI in einem gemeinsamen Projekt mit der Klinik für Innere Medizin V – Pneumologie (Direktor Prof. Dr. Dr. Robert Bals) die Auswir-

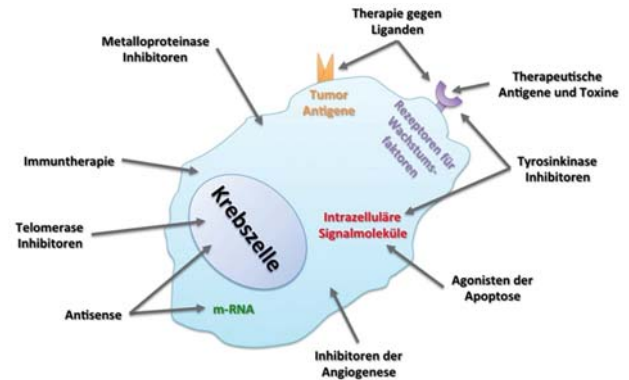


Abbildung 25 // Mögliche Ursachen der Krebsentstehung und dagegen gerichtete Therapieoptionen. Abbildung angelehnt und erweitert aus: Figure 1, Banik K.: A review on targeted Drug Therapy for cancer (pharmatutor.org).

kung von durch myeloide Zellen produziertem Protein RelA/p65 auf das Wachstum und die Malignität von Lungentumoren untersucht. Die Transkriptom- und Netzwerkanalyse hat hierbei gezeigt, dass das von myeloiden Zellen produzierte Protein RelA/p65, welches eine zentrale Rolle im Immunsystem und bei Entzündungsprozessen spielt, einen fundamentalen Einfluss auf verschiedene Wachstumssignalkaskaden der Krebszellen hat, unter anderem auch auf den Wnt/ beta-Catenin-Pathway [2]. Ferner wurde aufgedeckt, dass RelA/p65 der molekulare Faktor ist, der durch Rauchen induzierte Entzündungsprozesse mit der Aktivierung bestimmter Wachstumsprozesse in Lungentumorzellen verlinkt.

Referenzen

- [1] J. Wegert, N. Ishaque, R. Vardapour, C. Georg et al. **Mutations in the SIX1/2 pathway and the DROSHA/DGCR8 miRNA microprocessor complex underlie high-risk blastemal type Wilms tumors.** *Cancer Cell* 27 (2): 298-311. 2015
- [2] D. Li, C. Beisswenger, C. Herr, J. Hellberg et al. **Myeloid cell RelA/p65 promotes lung cancer proliferation through Wnt/beta-catenin signaling in murine and human tumor cells.** *Oncogene* 33 (10): 1239-48. 2014



Prof. Dr. Norbert Graf

Klinik für Pädiatrische Onkologie und Hämatologie
Universitätsklinikum Saarland

Telefon +49 68 41 162 83 97
graf@uks.eu



Prof. Dr. Eckart Meese

Institut für Humangenetik
Universität des Saarlandes

Telefon +49 68 41 162 60 38
eckart.meese@uks.eu

Biomarker und Diagnose

Gefördert durch Comprehensive Biomarker Center (Heidelberg) und Siemens Healthcare

Bei vielen Krebsarten ist eine frühzeitige Diagnose der sicherste Weg zu einer Heilung. Daher wird fieberhaft an entsprechenden Tests geforscht, die eine Frühdiagnose erlauben und gegebenenfalls sogar die erfolgversprechendste Therapie vorschlagen. Das Zentrum für Bioinformatik befasst sich seit mehr als einem Jahrzehnt mit der Entwicklung solcher Diagnostika.

Grundlegend folgt die Forschung in Saarbrücken von Beginn an drei Paradigmen:

1. Entsprechende Tests sollten minimal- oder nicht-invasiv sein.
2. Anstelle von einzelnen Biomarkern sollen Kombinationen von Markern verwendet werden.
3. Die Marker-Profile sollen nicht nur für Patienten, die an einer bestimmten Krankheit leiden, und Kontrollen gemessen werden, sondern die Verteilung über verschiedene Krankheitsbilder hinweg soll untersucht und verstanden werden.

Tests, die diese drei Paradigmen erfüllen, haben entscheidende Vorteile. Durch das Messen von Körperflüssigkeiten wie Blut oder Serum ist kein chirurgischer Eingriff notwendig, der nicht nur mit hohen Kosten, sondern darüber hinaus mit erheblichen Risiken verbunden wäre. Durch die Kombination von Biomarkern kann die Genauigkeit der Diagnose deutlich verbessert und die Zahl falsch-positiver oder -negativer Diagnosen verringert werden. Weil mehrere Erkrankungen gleichzeitig untersucht werden, kann die Spezifität der Tests erhöht werden. So ist es möglich, Signaturen zu finden, die gleichzeitig für Lungen-, Brust- oder Prostata-Tumore verwendet werden können.

Die Forschung des ZBI auf dem Gebiet der minimal-invasiven Tumordiagnostik startete mit der Untersuchung von Autoantikörper-Profilen im Blutserum, die Zeichen einer durch die Krankheit induzierten humoralen Immunantwort sind. Eine erste Publikation zum Thema Diagnostik wurde 2005 in der renommierten Fachzeitschrift „Proceedings of the National Academy of Science“ (PNAS) veröffentlicht. Ab 2008 hat sich das ZBI darüber hinaus mit der Erforschung von MicroRNAs befasst: kleinen RNA-Bausteinen, die vom Erbgut des Menschen abgelesen, aber nicht in Proteine übersetzt werden.

Anstelle von Serum wurden die Profile aus Blutzellen, hauptsächlich weißen Blutkörperchen, extrahiert. Ein Manuskript, das die spezifischen Profile von mehr als einem Dutzend Erkrankungen beschreibt, wurde in „Nature Methods“ [1] veröffentlicht. Die entsprechenden Signaturen sind inzwischen für eine Vielzahl von Tumorerkrankungen validiert worden. Mit am vielversprechendsten sehen die Ergebnisse zur Frühdiagnose bei Lungentumoren aus. Erwähnenswert ist, dass entsprechende Tests über die Onkologie hinaus entwickelt werden. Besonders erwähnenswert sind hier Profile zur Detektion von Multipler Sklerose oder Alzheimer. Neben zahlreichen Publikationen in Fachzeitschriften wurden entsprechende Muster mit verschiedenen Firmen, unter anderem dem Comprehensive Biomarker Center (Heidelberg) oder Siemens Healthcare (Erlangen) zum Patent angemeldet. Sowohl in Europa als auch den USA sind die ersten Patente erteilt worden. Eine klinische Erprobung der Marker-Profile wird gerade vorbereitet.

Referenzen

- [1] A. Keller, P. Leidinger, A. Bauer, A. Elsharawy et al. **Toward the blood-borne miRNome of human diseases.** *Nat Methods* 8 (10): 841-3. 2011



Prof. Dr. Andreas Keller

**Klinische Bioinformatik
Universität des Saarlandes**

Telefon +49 681 302 68611
andreas.keller@ccb.uni-saarland.de



Prof. Dr. Eckart Meese

**Institut für Humangenetik
Universität des Saarlandes**

Telefon +49 68 41 162 60 38
eckart.meese@uks.eu



Prof. Dr. Hans-Peter Lenhof

**Zentrum für Bioinformatik
Universität des Saarlandes**

Telefon +49 681 302 647 01
lenhof@bioinf.uni-sb.de

Therapieoptimierung

Gefördert durch das EU-Projekt p-Medicine

66

In der Medizin zeichnet sich derzeit ein Paradigmenwechsel von einer bisher überwiegend reaktiven zu einer präventiven, prädiktiven und personalisierten medizinischen Versorgung ab. Dabei wird eine phänotypische Charakterisierung von Krankheiten zunehmend durch eine genotypische ersetzt. Als ‚personalisierte Medizin‘ bezeichnet man die systematische Nutzung von Informationen über einen einzelnen Patienten und dessen Krankheit, um seine individuelle Diagnose und Therapie optimieren zu können. Neue und stetig wachsende Erkenntnisse aus den Grundlagenwissenschaften führen zu einem immensen Anstieg biologischer, genetischer und molekularer Erkenntnisse. Mittels moderner Informationstechnologie könnten z.B. alle Krankheitsdaten eines Patienten mit den molekulargenetischen Daten seines Tumors und weiteren Forschungsdaten unmittelbar verzahnt und ausgewertet werden. Bioinformatische Tools und Modellierungen von Krankheiten und deren Ansprechen auf verschiedene Therapien im Computer (in silico), erlauben dann individuelle Behandlungsmöglichkeiten für Patienten zu eröffnen. Bei Krebserkrankungen sind dadurch höhere Heilraten durch Therapien mit weniger Nebenwirkungen zu erwarten, da diese gezielt auf den einzelnen Patienten abgestimmt sind.

Das Hauptziel des EU-Projekts „p-Medicine – From data sharing and integration via VPH (Virtual Physiological Human) models to personalized medicine“ war es, die Errungenschaften und den Nutzen der personalisierten Medizin allen Krebspatienten zukommen zu lassen. Das Projekt wurde von Februar 2011 bis Juli 2015 durch die Europäische Kommission im Rahmen des siebten Rahmenprogramms gefördert. In diesem Projekt haben 19 Partner aus neun europäischen Ländern und Japan zusammengearbeitet, um neues Wissen und innovative Technologien zur Lösung aktueller Probleme der klinischen Forschung zu gewinnen und den Weg hin zur personalisierten Medizin zu bahnen (siehe Abbildung 26).

Das Projekt wurde vom Direktor der Klinik für Kinderheilkunde, Pädiatrische Onkologie und Hämatologie Prof. Dr. Norbert Graf koordiniert. Die integrierte Software-Plattform, die im Rahmen dieses Projekts geschaffen wurde, bietet eine extrem umfangreiche Funktionalität, die es nicht nur ermöglicht, effizient und unter Beachtung des Datenschutzes große personalisierte klinische und molekularbiologische Datensätze zu handhaben, sondern sie ermöglicht auch integrative Analysen der hochdimensionalen Datensätze (In-Silico-Onkologie), welche die Onkologen bei der Therapiestratifizierung im Sinne der personalisierten Medizin unterstützt [1].

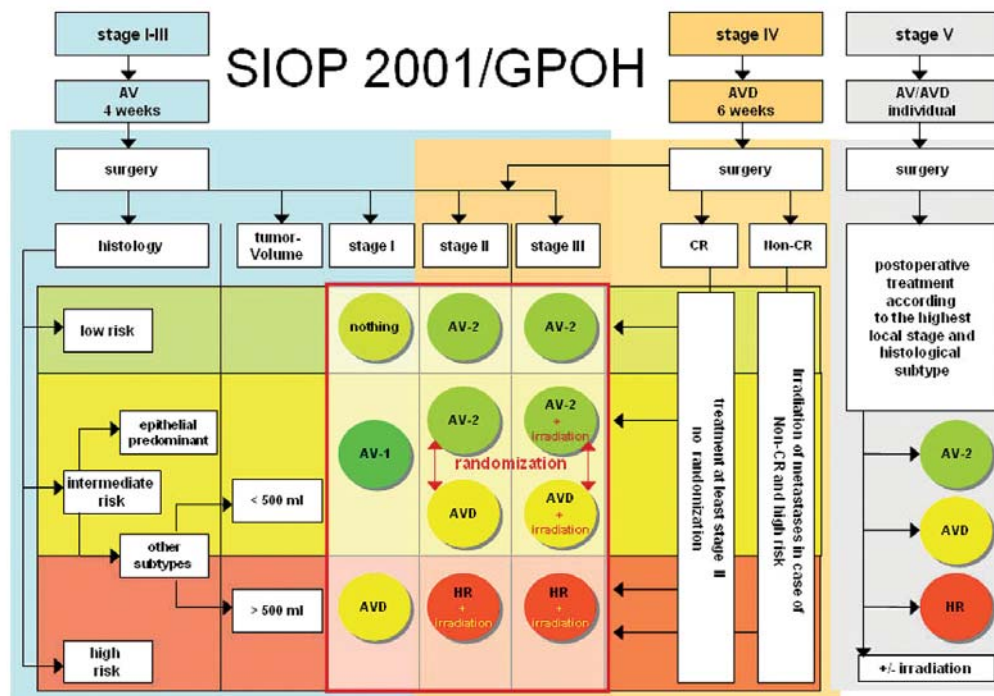


Abbildung 26 // Therapieoptimierungsplan für das Nephroblastom entsprechend der Studie SIOP 2001.

Referenzen

- [1] K. Pritchard-Jones, C. Bergeron, B. de Camargo et al. Omission of doxorubicin from the treatment of stage II-III, intermediate-risk Wilms' tumour (SIOP WT 2001): an open-label, non-inferiority, randomised controlled trial. *Lancet* 386 (9999): 1156-64. 2015



Prof. Dr. Norbert Graf

Klinik für Pädiatrische Onkologie und Hämatologie
Universitätsklinikum Saarland

Telefon +49 68 41 162 83 97

graf@uks.eu

Weitere medizinische Anwendungen

Der Schwerpunkt des Zentrums für Bioinformatik auf medizinische Themen zeigt sich auch in weiteren Forschungsprojekten, die am Zentrum durchgeführt werden. Die vier Forschungsprojekte, die in diesem Abschnitt beschrieben werden, sind in der Grundlagenforschung verortet, haben aber einen hohen translationalen Anspruch. Allen gemeinsam ist die Zielsetzung, die Diagnostik durch neuartige bioinformatische Methoden und Softwareangebote, die zum Teil mit neuartigen diagnostischen Messmethoden verbunden sind, voranzubringen.

MicroRNA sind kurze Ribonukleinsäuren, die erst vor etwa zwei Jahrzehnten entdeckt worden sind und in die Zellregulation eingreifen. Ihre Funktionen sind heute noch nicht vollständig aufgeklärt, ihr Vorkommen kann aber bei vielen Krankheiten diagnostisch genutzt werden. Die Arbeitsgruppe von Professor Andreas Keller führt entsprechende Forschung durch.

Die Phenomizer App ist eine von Dr. Marcel Schulz in Kooperation mit der Charité Berlin entwickelte Software, die bei der Diagnose von genetisch bedingten Krankheiten hilft, indem sie das Wissen über diese Krankheiten zusammenfasst und statistisch mit eingegebenen Anfragen abgleicht.

In einem weiteren am Zentrum für Bioinformatik durchgeführten Projekt analysierte Professor Jan Baumbach das molekulare Profil von kleinen Molekülen in der Atemluft von Patienten und Probanden. Dieses Profil ist in hohem Maße aussagekräftig bezüglich des körperlichen Zustandes und insbesondere bezüglich verschiedener Krankheiten. Die entwickelte Software wird in Zusammenhang mit einem entsprechenden Messgerät verwendet, in das der Patient hineinbläst.

Schließlich wird über ein Projekt berichtet, bei dem das Genom der Gletschermumie Ötzi sequenziert und auf krankheitsbedingte genomische Varianten untersucht wurde. In dem Projekt wurde festgestellt, dass schon Ötzi eine genomische Risikovariante für Arteriosklerose besaß und diese Krankheit auch hatte.

Die Stärke des Zentrums für Bioinformatik liegt sowohl in der Breite der medizinisch relevanten Themen, die bioinformatisch bearbeitet werden, als auch in der Tiefe der entsprechenden Forschung.



Weitere medizinische Anwendungen

MicroRNA Diagnostik	70
Computergestützte Diagnose genetischer Krankheiten.	71
Atemluft	72
Medizinische Genomik bei der Gletschermumie Ötzi	73

MicroRNA Diagnostik

Förderung durch Comprehensive Biomarker Center (Heidelberg), Siemens Healthcare, Deutsche Forschungsgemeinschaft (DFG) und Deutsche Krebshilfe

In der ersten Phase, der sich nach 1950 schnell entwickelnden Molekularbiologie, standen zuerst Gene im Mittelpunkt der Forschung. Gene, die Proteine kodieren, machen allerdings nur einen kleinen Teil der DNA aus (weniger als 3%). Die Bedeutung der restlichen DNA, die zunächst abwertend als „Abfall“ (Junk)-DNA bezeichnet wurde, ist erst in den letzten zwei Jahrzehnten zunehmend erkannt worden – nicht zuletzt in Folge der kompletten Sequenzierung des menschlichen Genoms. Die erste nahezu vollständige DNA-Sequenz eines Menschen wurde im Jahr 2001 veröffentlicht. Die fälschlicherweise als Junk bezeichnete DNA, die hauptsächlich aus sich wiederholenden (repetitiven) Sequenzen besteht, spielt – wie wir heute wissen – bei der Regulation der Aktivität von Genen eine zentrale Rolle.

In diesem Zusammenhang sind besonders kurze DNA-Bereiche interessant, die zwar keine Gene in dem Sinne darstellen, dass sie Informationen zum Bau von Proteinen enthalten, aber kurze RNA-Moleküle kodieren und deshalb als nicht-Protein-kodierende RNA-Moleküle (non coding RNAs) bezeichnet werden. Unter diesen RNAs gibt es eine große Gruppe sehr kurzer Moleküle, die nur eine Länge von 17-25 Nucleotiden – die Bausteine, aus denen sich DNA und RNA zusammensetzen – haben, und deshalb als MicroRNAs bezeichnet werden. Das gesamte Genom hat demgegenüber eine Länge von 3 Milliarden Nucleotiden. Viele dieser MicroRNAs haben zentrale regulatorische Funktionen in der Zelle und können die zelluläre Entstehung hunderter Proteine gleichzeitig beeinflussen. Die Erkenntnisse der letzten 15 Jahre lassen darauf schließen, dass MicroRNAs in nahezu allen zellulären Prozessen eine sehr wichtige Rolle spielen – und damit auch bei der Entstehung von Krankheiten.

Das Zentrum für Bioinformatik hat in enger Zusammenarbeit mit dem Institut für Humangenetik der UdS unter anderem Bioinformatik-Methoden für die Untersuchung der Interaktionen zwischen MicroRNAs und Genen, die durch sie reguliert werden (die sogenannten Zielgene), entwickelt. Diese Methoden sind auf Seite 40 dargestellt. Darüber hinaus wurden in diesen kooperativen Arbeiten MicroRNAs identifiziert, die bei bestimmten Erkrankungen eine zentrale Rolle spielen [1].

MicroRNAs finden sich nicht nur in Zellen von Organen, sondern auch in Körperflüssigkeiten des Menschen. Dieser Umstand eröffnet die Möglichkeit, MicroRNAs zur minimal-invasiven Diagnostik von Krankheiten zu nutzen, d.h. zu einer Diagnostik, die nicht die Entnahme eines Gewebestücks

(Biopsie) voraussetzt, sondern die z.B. anhand von Blut durchgeführt werden kann. Eine weitere Eigenschaft, die MicroRNAs als entsprechende Biomarker interessant macht, ist ihre große Stabilität, denn viele Moleküle, wie Proteine oder andere RNAs, sind sehr instabil und unterliegen Zerfallsprozessen, sobald sie ihren ursprünglichen Ort (ein bestimmtes Gewebe, eine bestimmte Zelle) im Körper verlassen haben.

Einer unserer Forschungsschwerpunkte liegt auf der Entwicklung von Bluttests, mit denen über den Nachweis bestimmter MicroRNAs Hinweise für das Auftreten einer Erkrankung gewonnen werden können. Der Anwendungsbereich solcher Tests erstreckt sich von der Frühdiagnostik, durch die der Behandlungserfolg ggf. vergrößert werden kann, über eine verbesserte Differentialdiagnostik, mit der ähnliche Krankheitsbilder unterschieden werden können, bis hin zur Prognose von Krankheitsentwicklungen. Durch die Zusammenarbeit des Zentrums für Bioinformatik und des Instituts für Humangenetik konnten bereits zahlreiche sogenannte MicroRNA-Signaturen, d.h. Biomarker, die sich aus mehreren MicroRNAs zusammensetzen, für verschiedenste Krankheitsbilder identifiziert werden, darunter Multiple Sklerose, chronisch obstruktive Lungenerkrankungen und Herzinfarkte. Hervorzuheben ist, dass sich solche MicroRNA-Signaturen auch für verschiedene Tumorerkrankungen finden lassen. Wir konnten MicroRNA-Biomarker im Blut von Patienten mit Lungenkarzinomen, Hirntumoren, Prostatakarzinomen, Melanomen sowie Tumoren, die spezifisch bei Kindern auftreten, identifizieren. Wie bereits auf Seite 65 dargelegt, haben diese Arbeiten zu über 45 Veröffentlichungen in Fachzeitschriften geführt; darüber hinaus wurden mit verschiedenen Firmen, unter anderem dem Comprehensive Biomarker Center (Heidelberg) und Siemens Healthcare (Erlangen), entsprechende Patente angemeldet.



Prof. Dr. Andreas Keller

**Klinische Bioinformatik
Universität des Saarlandes**

Telefon +49 681 302 68611
andreas.keller@ccb.uni-saarland.de



Prof. Dr. Eckart Meese

**Institut für Humangenetik
Universität des Saarlandes**

Telefon +49 68 41 162 60 38
eckart.meese@uks.eu



Prof. Dr. Hans-Peter Lenhof

**Zentrum für Bioinformatik
Universität des Saarlandes**

Telefon +49 681 302 647 01
lenhof@bioinf.uni-sb.de

Referenzen

- [1] A. Keller, P. Leidinger, A. Bauer, A. Elsharawy et al. **Toward the blood-borne miRNome of human diseases.** *Nat Methods* 8 (10): 841-3. 2011

Computergestützte Diagnose genetischer Krankheiten

Phänotypische Merkmale erlauben die computergestützte klinische Diagnose von genetischen Krankheiten. Eine korrekte, differentielle Diagnose für Patienten mit genetischen Krankheiten zu erstellen, ist oft eine Herausforderung. Diese Differentialdiagnose wird erschwert durch die große Anzahl von genetischen Erkrankungen (über 8.000 sind heute bekannt), von denen jede durch eine Vielzahl von klinischen Symptomen charakterisiert ist. Zusätzlich überlappen die Merkmale vieler Krankheiten und nicht alle Patienten weisen die bekannten klinischen Symptome auf, die mit einer bestimmten Krankheit assoziiert sind.

Eine Krankheit schnell richtig zu diagnostizieren, erspart aber unnötige Diagnoseverfahren, reduziert die Belastung für die Patienten und die Kosten der medizinischen Versorgung. Aufgrund der Komplexität bei seltenen genetischen Krankheiten warten Patienten manchmal Jahre bis eine korrekte Diagnose gestellt wird. Auch wenn eine Heilung noch nicht für alle genetischen Krankheiten möglich ist, so hilft die Diagnose den Betroffenen und deren Familien mit der Krankheit besser zurechtzukommen.

In seiner Zeit als Postdoc in Pittsburgh hat Marcel Schulz neue Algorithmen und statistische Ansätze entwickelt, um eine computergestützte Differentialdiagnose zu ermöglichen [1]. Diese Arbeit geschah in Kooperation mit der Gruppe von Prof. Peter Robinson an der Charité in Berlin. Die Ansätze nutzen eine graphenbasierte, hierarchisch organisierte Datenbank aller Krankheitsmerkmale, genannt die Human Phenotype Ontology (HPO), die in Berlin entwickelt wurde. Am ZBI entwickelt die Gruppe von Dr. Schulz neue Software, um die Differentialdiagnose im klinischen Alltag leichter zugänglich zu machen.

Der Phenomizer Webserver erlaubt computergestützte klinische Diagnosen.

Im Phenomizer Webserver haben wir unsere neuen Ansätze implementiert und für Ärzte auf der ganzen Welt frei verfügbar gemacht. Der Benutzer gibt die klinischen Symptome des Patienten durch standardisierte HPO Begriffe ein und der Phenomizer liefert eine Rangliste der am besten passenden Differentialdiagnosen unter Verwendung aller genetischen Erkrankungen, die in der Datenbank annotiert sind.

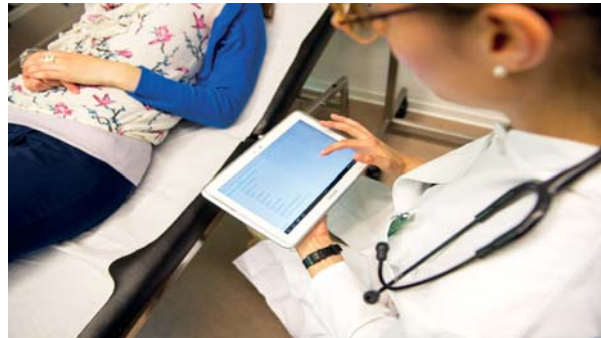


Abbildung 27 // Die Phenomizer App vereinfacht die computergestützte Diagnose von genetischen Krankheiten im Klinikalltag.

Derzeit kennt der Phenomizer mehr als 10.000 Krankheitssymptome und ihre Ähnlichkeit zueinander. Diese werden über 7.500 Krankheiten zugeordnet.

Der Phenomizer berechnet mit Hilfe schneller Algorithmen, die wir entwickelt haben, für jedes Suchergebnis die geschätzte Wahrscheinlichkeit, das gleiche Ergebnis durch eine Zufallssuche zu erhalten. Innerhalb von Sekunden erhält der Arzt eine Liste mit den wahrscheinlichsten Ergebnissen, so dass er nicht mehrere Stunden in Datenbanken oder Büchern suchen muss. Die Liste dient einer schnellen Erfassung der richtigen Krankheit oder der Planung weiterer Untersuchungen, um zwischen genetischen Krankheiten zu unterscheiden, die als gleichwahrscheinlich für die Anfrage gelten.

Mobilität für die klinische Praxis

Um eine schnelle Diagnose in der klinischen Praxis zu ermöglichen und aufgrund der weitverbreiteten Verfügbarkeit von Smartphones und Tablets, haben wir in Saarbrücken die Phenomizer App für iPhone und Android-basierte Systeme entwickelt (Abbildung 27). Die App hat die gleichen Funktionen wie der Phenomizer Webserver und unterstützt zusätzlich das Speichern von Anfragen und den Austausch von solchen.

Der Phenomizer Webserver, lokalisiert in Berlin, wurde über 50.000 Mal in den letzten drei Jahren mit Anfragen ausgeführt und über 1.000 Nutzer haben die App installiert. Am häufigsten wird der Phenomizer in den USA, Deutschland, Großbritannien, Frankreich und Italien benutzt.

Referenzen

- [1] M. H. Schulz, S. Köhler, S. Bauer and P. N. Robinson. **Exact score distribution computation for ontological similarity searches.** *BMC Bioinformatics* 12: 441. 2011



Dr. Marcel Schulz
High-throughput Genomics & Systems Biology
Universität des Saarlandes /
Max-Planck-Institut für Informatik
Telefon +49 681 9325 3115
mschulz@mmci.uni-saarland.de

Atemluft

72

Das Volatolom ist die Summe aller flüchtigen organischen Komponenten, die von allen lebenden Organismen, Zellen und Geweben abgegeben werden. Unser Ziel ist es, potentielle Biomarker im Volatolom zu finden, die Auskunft über den Gesundheitszustand eines Patienten geben können. Die Atemluftanalyse als nicht invasive Alternative der klinischen Diagnostik birgt die Chance der Krankheitsfrüherkennung und damit die Hoffnung, das therapeutische Fenster vorzuverlegen und Heilungschancen zu erhöhen.

Während tragbare Geräte zur Analyse der Komponenten in der Atemluft bereits existieren, stehen wir vor ähnlichen Herausforderungen, wie in der traditionelleren Biomarker-Forschung, z. B. in Blut oder Urin: Ein Mangel an Robustheit behindert die nächsten Schritte zum Einsatz in der klinischen Praxis (Abbildung 28).

Um von der Entdeckung erster Biomarker zu deren Validierung und der Vorhersage von Krankheiten zu gelangen, nutzen und entwickeln wir verschiedene Methoden der Bioinformatik und passen diese den neuen Anforderungen in der Atemluftanalyse an.

Diese Kombination aus analytischen und informatischen Methoden hat das Potential für eine moderne nicht invasive Methode der klinischen Diagnostik, die eine schnelle und günstige Detektion von Krankheiten ermöglicht und die klassischen Diagnosewerkzeuge wie Urin, Blut und Gewebeproben hervorragend ergänzt.



Abbildung 28 // Die Kombination aus analytischen und informatischen Methoden birgt das Potential für eine moderne nicht invasive klinische Diagnostik. Unser Fokus liegt sowohl auf der Automatisierung der Analyse der Atemluftproben, wie zum Beispiel der Peak-Erkennung, als auch auf der Entwicklung von maschinellen Lern- und Vorhersageverfahren zur direkten Unterstützung der klinischen Diagnostik.

Referenzen

- [1] A. C. Hauschild, T. Frisch, J. I. Baumbach, and J. Baumbach. **Carotta: Revealing hidden confounder markers in metabolic breath profiles.** *Metabolites*, 5(2):344–363, 2015
- [2] A.C. Hauschild, D. Kopczynski, M. D'Addario, J. I. Baumbach, S. Rahmann and J. Baumbach. **Peak detection method evaluation for ion mobility spectrometry by using machine learning approaches.** *Metabolites*, 3(2):277–293, 2013



Prof. Dr. Jan Baumbach

**Dept. of Mathematics and Computer Science
University of Southern Denmark**

Telefon +45 65 50 23 09

jan.baumbach@imada.sdu.dk

Medizinische Genomik bei der Gletschermumie Ötzi

Bereits seit 2010 arbeitet das Zentrum für Bioinformatik gemeinsam mit der Europäischen Akademie in Bozen daran, molekulare Muster der Gletschermumie „Ötzi“ zu erforschen (Abbildung 29 zeigt die Mumie). Zunächst stand das Genom im Fokus der Forscher. Mittels NGS-Sequenzieretechnologie ist es dem Team um Prof. Keller gelungen, über 95% des Erbgutes der 3.500 Jahre alten Mumie, die 1991 in den Ötztaler Alpen gefunden wurde, zu entschlüsseln [1]. Im Jahre 2012 veröffentlicht, zählt das Genom zu einem der ersten öffentlich verfügbaren und bis heute zu einem der best-charakterisier-testen Genome einer Mumie.

Nach der Entschlüsselung des Genoms wurden die Zusammensetzung der Proteine in Zellen von Ötzi untersucht [2]. Dabei wurden massenspektrometrische Verfahren eingesetzt, um das Proteom zweier Biopsien aus dem Schädel der Gletschermumie zu vermessen (Abbildung 30 zeigt CT Aufnahmen des Schädels mit den Stellen der Probeentnahmen). Den Forschern ist es bei der Untersuchung gelungen, Varianten (Mutationen) im Genom, die zum Austausch einer Aminosäure eines Proteins führen, auch in der Untersuchung mit dem Massenspektrometer in den entsprechenden Proteinen nachzuweisen. Neben dem Genom und dem Proteom bieten sich aber auch noch viele weitere Untersuchungen bei der Mumie an. Dabei ist die gute Konservierung der Mumie von entscheidendem Vorteil.



Abbildung 29 // Die 3.500 Jahre alte Gletschermumie „Ötzi“.

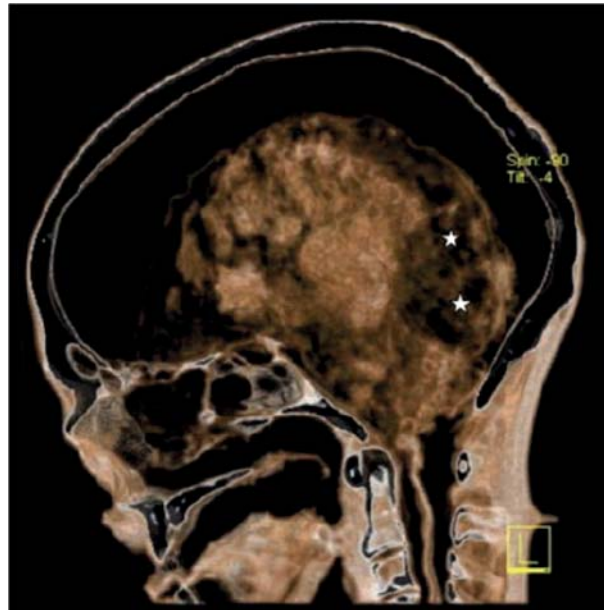


Abbildung 30 // CT Aufnahme des „Ötzi“ Schädels mit den Entnahmestellen (*).

Von uns gefundene genetische Variationen im Genom von Ötzi lassen auf eine Arteriosklerose schließen. Er hatte also eine genetische Veranlagung für diese Erkrankung. Entsprechend sind auf CT Aufnahmen auch Verkalkungen in der Mumie zu erkennen.

Die Herstellung eines Bezuges zwischen dem Genom von Ötzi und bei ihm festgestellten Krankheitsphänotypen birgt potenziell wichtige Einsichten in die Entstehung und für die Behandlung heute verbreiteter Krankheiten. Damit kann Ötzi sogar einen Beitrag zur heutigen Diskussion über genomische Medizin leisten – ein Aspekt, der für das Saarbrücker Team sehr wichtig ist.

Referenzen

- [1] A. Keller, A. Graefen, M. Ball, M. Matzas, V. Boisguerin et al. **New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing.** *Nat Commun* 3: 698. 2012
- [2] F. Maixner, T. Overath, D. Linke, M. Janko, G. Guerriero, G. et al. **Paleoproteomic study of the Iceman's brain tissue.** *Cell Mol Life Sci* 70 (19): 3709-22. 2013



Prof. Dr. Andreas Keller

**Klinische Bioinformatik
Universität des Saarlandes**

Telefon +49 681 302 68611
andreas.keller@ccb.uni-saarland.de



Prof. Dr. Eckart Meese

**Institut für Humangenetik
Universität des Saarlandes**

Telefon +49 68 41 162 60 38
eckart.meese@uks.eu

Biotechnologische Anwendungen

Schon seit Jahrtausenden setzen Menschen bei der Herstellung von Brot, Bier und Wein auf Mikroorganismen. Die von ihnen ausgelösten chemischen Prozesse, etwa bei der Fermentierung, waren lange ein unverstandenes aber willkommenes „Geschenk der Götter“. Heute ist die Biotechnologie eine der Schlüsseltechnologien des 21. Jahrhunderts. Mikroorganismen synthetisieren eine Fülle von Produkten im Wert von jährlich fast 500 Milliarden Euro. Durch die Nutzung nachwachsender Rohstoffe als Ausgangsmaterialien für biotechnische Synthesen lassen sich industrielle Herstellungsverfahren ressourcen- und umweltschonend gestalten.

Dies ist ein zukunftsweisender Vorteil gegenüber herkömmlichen Prozessen, die auf schwindenden, fossilen Rohstoffen basieren und mit teils hohen Energiekosten und der Emission von klimaschädlichen Treibhausgasen verbunden sind. Die Produktion von morgen nutzt biosynthetische Verfahren. Natürliche Zusätze für Nahrung und Futtermittel, Kunststoffe, Düngemittel, Treibstoffe, Farben, Aromen sowie Verpackungen sind nur einige Beispiele einer breiten Produktpalette. Mikroorganismen können dabei aus einfachen Grundbausteinen komplizierte Naturstoffe mit spezifischen dreidimensionalen Strukturen erstellen, die für den Chemiker im Labor oft nur schwer, mitunter sogar überhaupt nicht zugänglich sind.

So werden auch Moleküle mit therapeutischer Wirkung zugänglich, mit denen sich schwerwiegende Infektionen oder Krebs bekämpfen lassen. Im Zentrum der neuartigen Biotechnologieprozesse und damit auch im Fokus der Forschung steht die mikrobielle Zelle. Ihr kommt für die Umsetzung der Ausgangsstoffe in die Zielprodukte eine entscheidende Rolle zu. Da Mikroorganismen natürlicherweise nur geringe Mengen der gewünschten Produkte synthetisieren, ist deren Entwicklung in industrielle Zellfabriken eines der Schlüsselfelder der Biotechnologie. Mit dem Aufkommen der Gentechnik vor etwa 50 Jahren eröffnete sich die Möglichkeit, Zellen zu modifizieren, um ihre Produktionsleistung zu verbessern. Allerdings sind die für den Experten – durch biochemische und molekulare Kenntnisse – sozusagen auf der Hand liegenden gentechnischen Optimierungen selten ausreichend für eine wirtschaftliche Produktion. Jahrzehnte der Forschung haben gezeigt, dass ein sehr viel tieferes Verständnis der komplexen, zellulären Vorgänge erforderlich ist und man die Zellen gezielt in ganz verschiedenen Bereichen ihres Stoffwechsels verändern muss, um eine erfolgreiche Umsetzung zu erhalten.

In diesem Zukunftsbereich arbeitet das Team um Professor Christoph Wittmann. Am Institut für Systembiotechnologie der Universität des Saarlandes erforschen die Wissenschaftler das komplexe Innenleben von Mikroorganismen.

In den letzten fünfzehn Jahren hat die Arbeitsgruppe an der UdS, und zuvor an der Universität Münster und der TU Braunschweig, systembiologische Methoden entwickelt, mit denen sie diese zellulären Vorgänge gezielt erfassen können. So lässt sich durch Verfütterung von Nährstoffen, die an ausgewählten Atompositionen spezifische Isotope tragen, und anschließende modellgestützte Auswertung des Einbaus der Isotope in die zahlreichen Metabolite der Zellen der „molekulare Verkehr“ durch das komplexe Stoffwechselnetzwerk der Zellen erfassen. Darüber hinaus verfügt die Arbeitsgruppe über Technologien zur Messung zellulärer Komponenten wie Gen-Transkripte oder Metabolite.

Das Ziel der Integration solcher umfangreicher Datensätze mit Hilfe mathematischer und bioinformatischer Ansätze ist ein grundlegendes Verständnis der zugrundeliegenden Stoffwechselforgänge sowie das Entdecken der richtigen Stellschrauben zur gezielten Veränderung der Stoffwechsellistung mit Hilfe der synthetischen Biologie. So lassen sich unerwünschte biochemische Reaktionen durch das Entfernen der entsprechenden Genabschnitte im Erbgut der Zelle eliminieren oder gewünschte Aktivitäten durch den Einbau entsprechend aktiver Promotorsequenzen verstärken. Auch lassen sich neue Reaktionen und sogar ganze Module einbringen oder auch komplexe Änderungen der Regulationsvorgänge der Zelle verwirklichen. Auf diesem Wege werden Zellen zu zellulären Minifabriken optimiert, um Chemikalien, Treibstoffe und Arzneimittel herzustellen. In geeigneten Bioreaktoren wird dann eine geeignete verfahrenstechnische Umgebung geschaffen, um mit Hilfe der Zellfabriken nachwachsende Rohstoffe effektiv und möglichst umfassend in die gewünschten Wertprodukte umzusetzen.



Biotechnologische Anwendungen

Biodübel: Vom nachwachsenden Rohstoff
zum Industrieprodukt 76

Das Geheimnis guter Schokolade 77

Biodübel: Vom nachwachsenden Rohstoff zum Industrieprodukt

Förderung durch BMBF in einem Verbund der Initiative „Bioindustrie21“ mit BASF SE, Daimler AG, Robert-Bosch-GmbH und Fischerwerke GmbH sowie im Rahmen des VIP-Projektes „Bio2Nylon“

76

Die Verknappung von petrochemischen Rohstoffen und der durch die Kohlendioxidkonzentration in der Atmosphäre bedingte Wärmeanstieg der Erde sind treibende Kräfte für die verstärkte Diskussion um die biotechnologische Verwendung von nachwachsenden Rohstoffen. Denn auch mit ihnen kann man Chemikalien herstellen. Steigende Ölpreise und die globale Erwärmung der Atmosphäre haben die Ökonomie- und Ökologie-Betrachtung der Biotechnologie entscheidend verändert. Polymere stellen einen wesentlichen Anteil der jährlich aus petrochemischen Rohstoffen hergestellten Industrieprodukte dar.

Biotechnologische Innovationen bieten nun die Möglichkeit, alternative Herstellungsrouten für Polymere zu etablieren. Sie eröffnen eine weitgehende Unabhängigkeit von fossilen Rohstoffen und basieren auf der im Jahresrhythmus verfügbaren Biomasse. Aufbauend auf früheren Forschungsaktivitäten der Arbeitsgruppe an der Technischen Universität Braunschweig forscht das Team um Prof. Wittmann nun am Institut für Systembiotechnologie gemeinsam mit den Industriepartnern BASF SE, Daimler AG, Fischerwerke GmbH und Robert-Bosch-GmbH daran, biobasiertes Nylon (Bionylon) auf biologischer Basis herzustellen und dabei die gesamte Wertschöpfungskette vom nachwachsenden Rohstoff bis zum Industrieprodukt abzudecken. Im Fokus steht dabei die Optimierung des Bodenbakteriums *Corynebacterium glutamicum* zu einem effizienten Produktionsstamm für einen der Schlüsselbausteine des neuen Polymers: das fünf Kohlenstoffe umfassende Diaminopentan.

Durch Ansätze der Systembiotechnologie und der synthetischen Biologie konnte das Team einen Produktionsstamm entwickeln, der aufgrund seines systemweit optimierten Stoffwechsels im Bioreaktor bis zu 90 Gramm pro Liter des wertvollen, biobasierten Diaminopentan akkumulieren kann [1].



Abbildung 31 // Innovative biobasierte Verfahren ermöglichen eine ressourcenschonende Herstellung vieler Produkte des täglichen Lebens. Maßgeschneiderte Zellfabriken wandeln dabei nachwachsende Rohstoffe effizient in Chemikalien, Materialien und Treibstoffe um. Damit lassen sich auch hochwertige bio-basierte Kunststoffe herstellen. Erste überwiegend bio-basierte Materialien sind als Industrieprodukte einer neuen Ära biosynthetischer Produktion bereits auf dem Markt.

Über Polymerisierung mit Sebazinsäure, einem Bestandteil von Pflanzenöl, gelang daraus erstmals die Herstellung des Biokunststoffes Nylon 5.10, einem Polymer mit ausgezeichneten Materialeigenschaften. Weiterführende Materialien, wie das verwandte Nylon 6.10, haben mittlerweile sogar Marktreife erlangt: Sie sind u. a. als Bauteile im Automobilbereich oder als Dübel im Baubereich im Einsatz.

In aktuellen Arbeiten wollen die Forscher Abfallstoffe zur Herstellung von Bionylon aus nachwachsenden Rohstoffen nutzen. Bei dem Vorhaben setzen Bakterien verarbeitete Holzabfälle zu einer Vorstufe der organischen Säure Adipinsäure um, dem Hauptausgangstoff für die weltweit im Millionen Tonnen Maßstab eingesetzten Industriepolymere Nylon 6 und Nylon 6.6. Erste Analysen haben gezeigt, dass das neue Verfahren sehr umweltschonend und wirtschaftlich rentabel sein kann.

Referenzen

- [1] S. Kind, S. Neubauer, J. Becker, M. Yamamoto, M. Völkert, et al. **From zero to hero – Production of bio-based nylon from renewable resources using engineered *Corynebacterium glutamicum*.** *Metab Eng* 25: 113-23. 2014



Prof. Dr. Christoph Wittmann

**Systembiotechnologie
Universität des Saarlandes**

Telefon +49 681 302 71970
christoph.wittmann@uni-saarland.de

Das Geheimnis guter Schokolade

Förderung durch Nestlé im Rahmen einer Industriekooperation



77

Abbildung 32 // Wohlgeschmeckende Schokolade ist das Ergebnis komplexer biotechnologischer Prozesse.

Als Praline, Riegel, Brotaufstrich oder Eis kommt sie in den Handel: Schokolade. Bis sich der zarte Schmelz und das volle Aroma der Kakaobohnen auf unserer Zunge entfalten, ist es aber ein langer Weg mit vielen Verarbeitungsschritten. Was dabei genau passiert, haben Biotechnologen um Professor Christoph Wittmann von der Saar-Uni herausgefunden. Bei der Fermentation arbeiten bestimmte Mikroorganismen eng zusammen. Nur dadurch wird der Grundstein für das charakteristische Schokoladenaroma gelegt. Bei der Fermentation setzen natürliche Mikroorganismen die Nährstoffe aus dem Kakao so um, dass daraus Vorstufen des Schokoladenaromas entstehen. Was dabei genau vonstattengeht, haben Wittmann und sein Team zunächst an der Technischen Universität Braunschweig und mit dem Wechsel der Arbeitsgruppe dann an der Universität des Saarlandes mit Forschern des Nestlé Research Centers in Lausanne herausgefunden [1].

Sie haben dabei das Innenleben der beteiligten Mikroorganismen bei der Arbeit beobachtet und verfolgt, wie sie einzelne Nährstoffe nutzen und sich an der Kakaofermentation beteiligen. Durch die sogenannte metabolische Flussanalyse konnte das Team einen Blick in den molekularen Verkehr durch die Stoffwechselwege der Zellen werfen. Faszinierenderweise kommt es unter den Mikroorganismen zu einer richtigen Zusammenarbeit, um die Schlüsselsubstanzen für die spätere Aromaentwicklung herzustellen. Fehlt bei diesen Vorgängen eine bestimmte Substanz oder auch einer der Mikroorganismen, kann später das gewünschte Aroma nicht entfaltet werden. Für gute Schokolade bedeutet dies, dass eine ausgewogene Zusammensetzung der richtigen Mikroorganismen nötig ist. Durch den Einsatz von natürlichen Starterkulturen, die die richtige Kombination der Mikroorganismen enthalten, könnten auf diesem Wege zukünftig Kakaobauern ihren Ertrag verbessern.

Referenzen

- [1] P. Adler, L. J. Frey, A. Berger, C. J. Bolten, C. E. Hansen and C. Wittmann. **The key to acetate: Metabolic fluxes of acetic acid bacteria under cocoa pulp fermentation simulating conditions.** *Appl Environ Microbiol* 80 (15): 4702-16. 2014



Prof. Dr. Christoph Wittmann

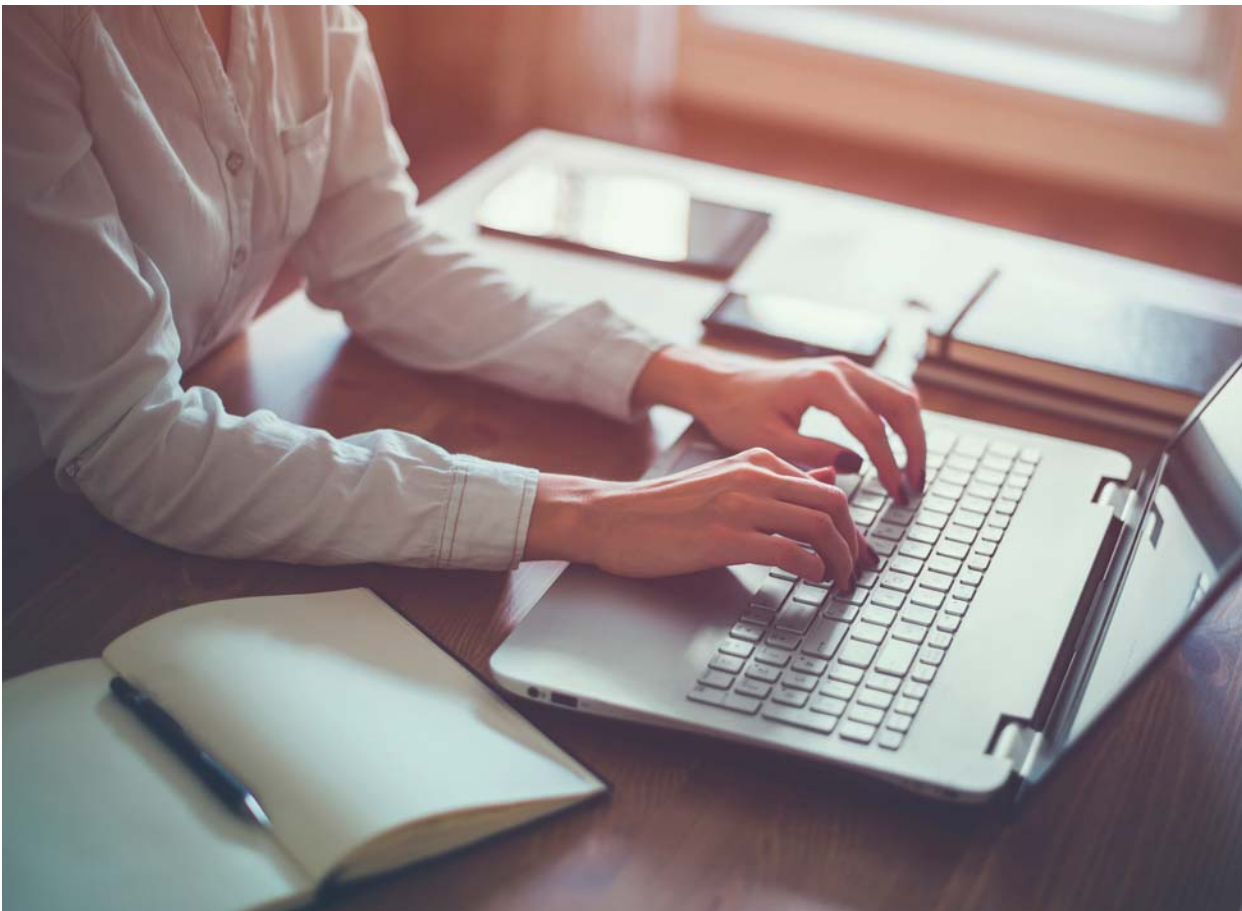
**Systembiotechnologie
Universität des Saarlandes**

Telefon +49 681 302 71970

christoph.wittmann@uni-saarland.de

4 Aktuelles und Zahlen

4 Aktuelles und Zahlen



Aussagen von Absolventen	80
Nachgefragt bei... ..	91
Aktivitäten des Zentrums	96
Preise und Auszeichnungen.....	103
Kolloquien und Vorträge.....	108
Drittmittel in Zahlen.....	114
Patente und Erfindungen	115
Kooperationen	116
Graduierungen	117
Rufe auf Professuren	121

Aussagen von Absolventen



Jan Christoph

Welche Zeit haben Sie am ZBI verbracht?

_ WS 03/04 bis Ende WS 08/09, abzüglich eines Auslandssemesters in den USA.

Was waren die Schwerpunkte Ihres Studiums bzw. Ihrer Forschung?

_ Der Schwerpunkt lag eher auf der theoretisch algorithmischen als auf der anwenderorientierten Seite (Bachelor mit CMB-Ausrichtung).

Was war das Thema Ihrer Abschlussarbeit?

_ Hm, für den genauen Wortlaut nachgesehen: „*Analysis of membrane transporters – focused on amino acid transporters and their pore region by threading and similarity search*“.

Im Kern ging es darum, für eine Chemiefirma herauszufinden, ob deren zur Produktion bestimmter Aminosäuren verwendeten Bakterien bzgl. der Transporterproteine in der Membrantechnisch zur Erhöhung der Produktionsmenge verändert werden können.

Wie war Ihr weiterer Werdegang?

_ Nach dem Masterabschluss 2009 wechselte ich für die Promotion zum Lehrstuhl für medizinische Informatik nach Erlangen und absolvierte dort einen zweiten Master „Medical Process Management“. Wegen der Geburt dreier Kinder, jeweils daran folgender Elternzeitreisen sowie der sehr guten Arbeitsbedingungen weile ich hier noch immer forschend im Bereich der translationalen und personalisierten Medizin bzw. der sich inzwischen dort entwickelnden klinischen Bioinformatik.

Was können Sie uns über ihr Studium berichten?

_ Hatte ich nach dem Abitur erst andere Universitäten für ein Studium der Bioinformatik im Blick, kam durch die sehr gute Bewertung des ZBI im CHE-Ranking die Universität des Saarlandes in die nähere Auswahl, welche ich mir im Mai 2003 zwei Tage lang anschaute. Dabei überzeugten mich neben der modern erscheinenden Infrastruktur vor allem die Gespräche mit Vertretern der Urgesteine des Fachschaftsrats (Tiffy&Co) sowie mit Prof. Lengauer und Prof. Lenhof, die sich dankbarerweise jeweils die Zeit dafür genommen hatten.

Die Bioinformatik in Saarbrücken bot im Vergleich zu vielen anderen Universitäten eine bemerkenswert erfreuliche Vielfalt und Freiheit bei der Durchführung des Studiums. Bei der Lehre gefiel neben den angenehmen Arbeitsbedingungen mit generell kleinen Gruppen, modernen CIP-Pools sowie z. B. den hilfreichen Mathematik-Tutorien die grundsätzliche Gestaltung mit den Wahlmöglichkeiten aus relativ vielen angebotenen Vorlesungsmodulen sowie die Übungen, bei denen das Gehörte/Gelernte angewandt und vertieft wurde.

Das ermöglichte ein fachlich erfreulich solides Fundament, wie sich in den Jahren nach dem Abschluss herausstellte. Als besonders intensiv und positiv erlebte Zeiten für Bildung und Ausbildung in Erinnerung geblieben sind vor allem die Blockveranstaltungen, z. B. das legendäre Software-und-Design-Praktikum SoDePra, die Möglichkeiten der studentischen Mitwirkung im Fachschafts- und Zentrumsrat sowie das Förder- und Mentorprogramm. Das alles schuf eine sehr angenehme Atmosphäre der kurzen Wege und des Miteinanders, sowohl unter den Studierenden als auch mit den Lehrenden. Privat bestehen einige tolle Kontakte aus der damaligen Zeit weiterhin, wobei sich inzwischen auch fachlich zunehmend Vernetzungen ergeben, gemeinsame Anträge für Forschungsanträge gestellt oder einfach nur in miteinander sich ergänzenden Themengebieten gearbeitet wird, sodass die Verbindung mit der UdS nicht nur emotional sondern auch fachlich gehalten wird.

Verbesserungspotential

_ Zumindest aus damaliger Sicht fand ich die Prüfungsordnung in einigen Punkten bemerkenswert fortschrittlich (z. B. Möglichkeit des Teilzeitstudiums) bzw. die Freiheit erhaltend (z. B. durch die Optionen, bei freier zeitlicher Kapazitäten zum Semesterende hin, an Prüfungen zusätzlicher Vorlesungen teilzunehmen). Die Änderungen im Laufe der Zeit schienen mir aber beides eher einzuschränken. Fachlich hätte ich mir damals einen wesentlich stärkeren Einbezug der Industrie gewünscht, weshalb ich Prof. Helms sehr dankbar um die Möglichkeit der Masterarbeit in einem Unternehmen war, da ansonsten während des Studiums nur sehr wenige Einblicke in die Wirtschaft gegeben waren.

Fazit:

_ Für die Zeit des Bioinformatikstudiums in Saarbrücken bin ich sehr dankbar und ich würde rückblickend in jedem Falle dort wieder studieren.



Sarah Diehl

Wann waren Sie am ZBI?

_ 2003 – 2010

Womit haben Sie sich wissenschaftlich beschäftigt? Was war das Thema Ihrer Abschlussarbeit?

_ Meine Arbeit hat sich meistens um die Kernthemen Datenbanken und Netzwerke/Interaktionen gedreht.

Der Titel meiner Bachelor Thesis:

„*The biochemical database webinterface*“

Der Titel meiner Master Thesis: „*Analysis and reconstruction of protein complex topology. Danach wissenschaftl. Mitarbeiter am MPII: analysis of results from RNA interference screens*“

Haben Sie noch Kontakte aus dieser Zeit?

_ Da die Welt der Bioinformatik in Deutschland klein ist, bin ich ein paar meiner alten Studienkollegen später wieder über den Weg gelaufen. Ständigen Kontakt habe ich nur noch zu ein-zwei Leuten.

Was fällt Ihnen als Erstes ein, wenn Sie an Ihre Zeit am ZBI zurückdenken?

_ Samstagmorgende beim Programmierung I Tutorial von Andreas Augustin

Was hat Ihnen am besten gefallen?

_ Dass sich gut um die Studenten gekümmert wurde und der Dekan, Prof. Helms, immer ein offenes Ohr hatte.

Was hat Sie gestört?

_ Dass man nicht über die Möglichkeiten nach dem Studium informiert wurde.

Wie war Ihr weiterer Werdegang?

_ Von 2010 bis Ende letzten Jahres war ich Data Manager und Bioinformatiker in der Bioinformatics and Deep Sequencing Unit des Max-Planck-Instituts für Immunbiologie und Epigenetik in Freiburg im Breisgau. Dort habe ich in enger Zusammenarbeit mit den Biologen hauptsächlich Next-Generation Sequencing Daten von unserer Illumina Maschine analysiert. Seit Anfang diesen Jahres arbeite ich im Bioinformatics Core des Luxembourg Centre for Systems Biomedicine und bin Teil des HPC Management Teams der Uni Luxembourg.



When have you been at CBI?

_ January, 2011 – August, 2011

What were your study foci and/or research interests?

_ Biological function of non-coding small RNAs in molecular biology.

What was the topic of your thesis?

_ „Prediction of binding affinity of aptamer towards a protein binding domain using a PWM based scoring function“

Can you characterize the study and research environment at CBI a few sentences?

_ CBI promotes independent research with a supportive group. Students are given the resource and a relaxed work environment to be productive and creative while the professor and supervises provide guidance and valuable ideas and suggestions.

Do you still have contacts from that time?

_ Yes, some fellow students.

What is the first thing that comes to your mind, when you think of your time at CBI?

_ Room for creativity and supportive group.

What did you like the most?

_ A friendly environment and mentoring-ship.

What was bothering you?

_ The distinctive study topic of my thesis

What was your further career?

_ PhD at the institute of clinical molecular biology of Kiel as a bioinformatic researcher. The projects focus on miRNA profile – disease association studies based on NGS.

As a update, I am now working in Berlin at the honey bee breeding institute. The project is genetic research of honey bees for sub species and traits studies. I find it very interesting and quite passionate about it.

Thank you for all the support over the years ever since I studied in CBI.



Dr. Christoph Hartmann

Welche Zeit haben Sie am ZBI verbracht?

_ 1.1.2005 – 31.3.2008

Was waren die Schwerpunkte Ihres Studiums bzw. Ihrer Forschung? Was war das Thema Ihrer Abschlussarbeit?

_ Protein-Ligand-Docking.
 „Modeling of Flexible Side Chains for Protein-Ligand Docking“

Können Sie in wenigen Sätzen für Sie Wesentliches über die Studien- und Forschungsumgebung am ZBI skizzieren?

_ Die Arbeit in einer internationalen, interdisziplinären Arbeitsgruppe war sehr anregend, meine Kollegen haben mir ganz neue Sichtweisen auf meine Fragestellungen ermöglicht. Die Rahmenbedingungen wie Räumlichkeiten, technische Ausstattung, Zugang zu Literatur, Zeit für die Forschung und auch die Entlohnung waren ausgezeichnet.

Haben Sie noch Kontakte aus dieser Zeit?

_ Kaum

Was fällt Ihnen als Erstes ein, wenn Sie an Ihre Zeit am ZBI zurückdenken?

_ Ein Grillabend, den wir 2005 mit der Arbeitsgruppe hinter dem Institut hatten.

Was hat Ihnen am besten gefallen?

_ Der Grillabend :-)
 Aber auch: Der wöchentliche Austausch und die Kurzvorträge, die wir in der Arbeitsgruppe gehalten haben.
 Die Ruhe, man konnte wirklich konzentriert arbeiten.
 Der Austausch mit anderen Forschungsgruppen weltweit, u.a. mit Professor Shoichet und John Irwin aus SF.

Was hat Sie gestört?

_ Ich habe mir in der Anfangszeit gewünscht, noch enger mit einigen Doktoranden an einem konkreten Projekt zu arbeiten. Das kam dann erst 2007 mit der HCV-Forschung mit Christoph Welsch zustande.

Die Lage in Saarbrücken war ungünstig, da ich so für 3 Jahre von Freunden und Familie im Rheinland räumlich getrennt war.

Die schlechten Berufsaussichten in der Bioinformatik im Rheinland: Eine akademische Laufbahn erschien mir wenig geeignet für eine Familie, wegen der damit verbundenen Standortwechsel. Eine Laufbahn in der Wirtschaft erschien mir wenig attraktiv, da um das Jahr 2006 zahlreiche Pharmafirmen ihre Forschungsabteilungen in Deutschland verkleinerten oder verlagerten, und auch die Softwareentwicklung für die Bioinformatik maßgeblich im Ausland getätigt wurde.

Wie war Ihr weiterer Werdegang?

2008-2009 IT-Berater tätig bei T-Mobile, Bonn
 2009-2014 Softwareentwickler bei CommaSoft AG, Bonn
 seit 2015 Softwareentwickler und Business-Consultant bei PatentSight GmbH, Bonn
 seit 2010 Ausbildung zum Ständigen Diakon mit Zivilberuf

seit 10 Jahren verheiratet und 3 Kinder:
 Dominik 8, Johanna 6, Martin 2



Dr. Alexander Rurainski

Wann waren Sie am ZBI?

_ Von 2004 bis 2010. Zuerst als Diplomand, dann als Assistent von Prof. Dr. Lenhof. In letzterem Rahmen habe ich meine Doktorarbeit angefertigt und meine Promotion abgeschlossen.

Womit haben Sie sich wissenschaftlich beschäftigt? Was war das Thema Ihrer Abschlussarbeit?

Schwerpunkte meiner Arbeit waren mathematische Modellierung, sowie numerische und algorithmische Optimierung. Diese für die Bioinformatik so wichtigen Gebiete sind unverzichtbar in Bereichen wie beispielsweise molekulare Kraftfelder, maschinelle Lernverfahren und regulatorische Netzwerke, in denen ich tätig war. Schon meine Diplomarbeit beschäftigte sich mit einer speziellen Klasse von Problemen: „*Semidefinite Programmierung in der Bioinformatik*“. Meine anschließenden Forschungen am Zentrum für Bioinformatik mündeten in meiner Dissertation, die unter dem Titel „*Optimization in Bioinformatics*“ zusammengefasst wurden.

Haben Sie noch Kontakte aus dieser Zeit?

_ Ja, insbesondere arbeite ich mit Personen aus dieser Zeit in der eigenen Firma täglich zusammen.

Was fällt Ihnen als Erstes ein, wenn Sie an Ihre Zeit am ZBI zurückdenken?

_ Ich habe den Umzug in das jetzige Bioinformatik Gebäude selbst miterlebt und z.B. Rechner abgebaut, getragen und wieder aufgebaut. Unabhängig ob vom „alten“ oder vom „neuen“ Gebäude die Rede ist, die Raumausstattungen boten/bieten in jedem Fall optimale Arbeitsbedingungen.

Was hat Ihnen am besten gefallen?

_ Teamgeist und Zusammenhalt

Was hat Sie gestört?

_ Es war nicht immer ganz einfach, als Assistent den Vorlesungsbetrieb mit den eigenen Forschungen zeitlich zu koordinieren.

Wie war Ihr weiterer Werdegang?

_ Nach meiner Promotion habe ich mit Freunden aus dem Zentrum für Bioinformatik eine eigene Firma für Biocomputing gegründet, die sich auf die Analyse und Simulation biologischer Daten und pharmazeutischer Eigenschaften spezialisiert hat: Scientific Consilience (www.scientific-consilience.com).

Die Scientific Consilience GmbH blickt nun als von Universitäten unabhängige Vollzeitfirma auf mehr als 5 Jahre wissenschaftlichen und marktwirtschaftlichen Erfolg zurück.



Dr. Nadine Schaadt

Welche Zeit haben Sie am ZBI verbracht?

_ WS 2005 – 31.12.2013

Was waren die Schwerpunkte Ihres Studiums bzw. Ihrer Forschung?

_ Forschung: Mathematische Modellierung + Machine Learning im Bereich Systems Biology
Studium (wichtige Vorlesungen): Bioinfo 3, Image Analysis and Computer Vision, Membrane Bioinformatics, Computational Immunology, Statistical Learning

Was war das Thema Ihrer Abschlussarbeit?

_ Bachelorarbeit: „*Quantitative Analysis of Quorum Sensing in Vibrio fischeri*“; in AG Prof. Helms
Masterarbeit: „*Substrate Prediction of Membrane Transporters in Arabidopsis thaliana*“; in AG Prof. Helms
Doktorarbeit: „*Computational Systems Biology Methods for Functional Prediction and Classification of Membrane Proteins and Modeling Quorum Sensing in Pseudomonas aeruginosa*“; in AG Prof. Helms

Können Sie in wenigen Sätzen für Sie Wesentliches über die Studien- und Forschungsumgebung am ZBI skizzieren? Was fällt Ihnen als Erstes ein, wenn Sie an Ihre Zeit am ZBI zurückdenken?

_ Sehr gute Betreuung (vor allem während Abschlussarbeiten); DANKE!
Gute Strukturierung + Organisation des Studiums; (sehr) gute Ausstattung. Info- + Spezialvorlesungen sehr hilfreich

Was hat Sie gestört?

_ Qualität der Bio-Vorlesungen

Haben Sie noch Kontakte aus dieser Zeit?

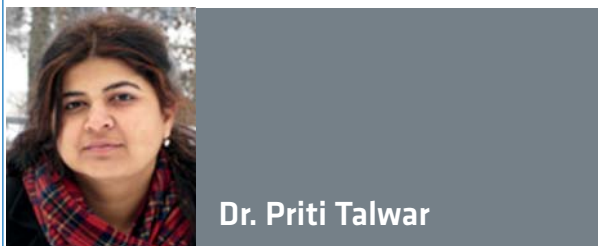
_ Ja

Was hat Ihnen am besten gefallen?

_ Meine Arbeit in der Forschung

Wie war Ihr weiterer Werdegang?

_ Derzeit PostDoc in Digital Pathology an der Hannover Medical School



Dr. Priti Talwar

When have you been at CBI?

_ 2002 – 2007

What were your study foci and/or research interests?

What was the topic of your thesis?

_ Computational Systems Biology, Metabolomics
Thesis title: „Development of computational methods for metabolic network analysis using metabolomics data”

Can you characterize the study and research environment at CBI a few sentences?

_ CBI is a very supportive environment. During my period of stay, I found it to be absolutely fantastic in terms of scientific culture and work ethics.

Do you still have contacts from that time?

_ Yes I have contacts with many of my fellow doctoral students, master’s students and post-doctoral senior scientists.

What is the first thing that comes to your mind, when you think of your time at CBI?

_ One thing that is always there in my mind is ever smiling face of my guide my mentor Prof. Dr. Thomas Lengauer. There has not been a single day when Dr. Lengauer went angry on me. Since I came from a biological background, he was always patient in explaining me basics of graph theory, statistics, network biology etc. I still vividly remember my Thursday morning one to one weekly progress meeting with Thomas Sir. My journey with the group was a super exciting one. I learnt a lot during my stay.

I also remember ever supportive, man of few words, Prof. Dr. Hans-Peter Lenhof from ZBI. He always encouraged me to work with his students who were working in the same area. Another sweet memory I have is of my Ph.D. defence, when I was able to answer his question though I was shaking in fear. I also remember all my peers and seniors who were always supportive whether it was in the research or working out the correct mix for the super coffee machine on first floor or trekking in Scotland after ISMB conference.

A special word of gratitude goes for Mrs. Ruth Schneppen-Christmann. She was very cooperative at each step during my doctoral studies. Thank you Madam.

CBI was a big family and whatever I have achieved today is due to kind guidance of my Sir Prof. Lengauer.

What did you like the most?

_ Diversity of researchers.

What was bothering you?

_ Nothing.

What was your further career?

_ After my thesis submission, I got postdoctoral offers from MPI Dresden, INRIA postdoctoral fellowship, and Postdoctoral fellowship from Sanford Burnham Medical Research Institute (SBMRI) in USA.

During 2007-2009, I did Postdoctoral research at SBMRI and University of California San Diego, USA.

In 2010, I returned to my home country India and joined as Assistant Professor (Senior) in School of Biosciences and Technology, VIT University, Vellore, India.

I am a recipient of Annual Research Award for consecutive last 3 years (2012-2014) conferred by VIT University, Vellore. In July 2015, I have received the DST FastTrack Young Scientist Start-up Grant Award from SERB, Department of Science and Technology, Govt. of India.



Dr. Laura Tolosi

When have you been at CBI?

_ 2006 - 2012

What were your study foci and/or research interests?

What was the topic of your thesis?

_ I was (and still am) interested in applied machine learning. Stepping into the field of computational biology was a great opportunity, because I knew I would be highly motivated by the applications. I studied patterns of chromosomal aberrations in cancer. Beyond the natural peaks and valleys of the everyday work, there was the strong driving force: I need keep trying, to help sick people.

Can you characterize the study and research environment at CBI a few sentences?

_ The courses that I took early on in my Master and PhD were deep enough to establish a strong basis for a successful thesis. Mentors and colleagues were equally inspiring, supportive, helpful and easy to communicate with. I never felt financial pressure (funding was not an issue), which set no limitations to my research, I felt free to explore and deepen into any idea. Rather, the guidance came from group discussions, which eventually would single out the most promising directions. Interactions with hospitals and medics also played a major role in my professional development: it taught me to abstract from technicalities and deliver simple messages. It is not important to demonstrate that I can implement complicated methods, but to show that the results are meaningful.

Do you still have contacts from that time?

_ Professional contacts, no. But I definitely made very good friends. We keep seeing each other, even if we need to fly for that. I also met my husband there.

What is the first thing that comes to your mind, when you think of your time at CBI?

_ Great people, passionate about research. Freedom of ideas. Diversity, I made friends from many different countries.

What did you like the most?

_ The professional honesty of the people. Research results were presented as objectively as possible, leaving out biases and exaggerated claims. And when it was hard to be objective with one's work, colleagues were always there to challenge with questions and alternative points of view. I don't find this honesty in industry.

What was bothering you?

_ Few things. Maybe transience, the fact that one needs to move on after a few years.

What was your further career?

_ I work in industry, in the research and development department of a small Bulgarian company that does machine learning for natural language processing. I went on developing as a machine learning expert, but left behind computational biology.



Dr. Junming Yin

When have you been at CBI?

_ 04/2004 – 03/2005

What were your study foci and/or research interests?

What was the topic of your thesis?

_ Machine Learning and Computational Biology. The title of my master thesis is „*Model selection for mixtures of mutagenetic trees*“.

Can you characterize the study and research environment at CBI a few sentences?

_ It is a great place to study bioinformatics.

Do you still have contacts from that time?

_ I had a few contacts with Niko Beerenwinkel.

What is the first thing that comes to your mind, when you think of your time at CBI?

_ Wonderful experience.

What did you like the most?

_ Friendly people.

What was bothering you?

_ None.

What was your further career?

_ I am currently an assistant professor in the Eller College of Management at the University of Arizona.



Dr. Siti Azma Yusoh

When have you been at CBI?

_ May 2006 – Dec 2010

What were your study foci and/or research interests?

_ Computational Biology and Bioinformatics

What was the topic of your thesis?

_ „Molecular modeling of the transmembrane domain of envelope glycoproteins from flaviviridae viruses”

Can you characterize the study and research environment at CBI a few sentences?

_ CBI is a great place to study bioinformatics, as CBI have expertise of experience and skilled scientist. CBI have produced many graduate students (MSc and PhD) with excellent research skills that meet the current demand in the area bioinformatics/computational biology and industry.

Do you still have contacts from that time?

_ Yes, I am still in contact with Professor Helms, my former advisor and several alumni from CBI.

What is the first thing that comes to your mind, when you think of your time at CBI?

_ CBI is a great place to study bioinformatics! Graduates from CBI are able to work independently in an encouraging educational environment.

What did you like the most?

_ CBI building is pretty impressive

What was bothering you?

_ Computational resources is limited compared to most of the US dedicated bioinformatics/computational research centre.

What was your further career?

_ I am now a senior scientist and lecturer in Universiti Teknologi MARA Malaysia in the area of Computational Drug Discovery and Biology



When have you been at CBI?

_ I studied for my PhD from Oct, 2003 to Jan, 2009 in the MPI for Informatics.

What were your study foci and/or research interests?

What was the topic of your thesis?

_ My PhD research was mainly in the field of structural bioinformatics, specifically, about the study of protein-protein interaction. My PhD dissertation titles „*Characterization, Classification and Alignment of Protein-Protein Interfaces*“.

Can you characterize the study and research environment at CBI a few sentences?

_ The researchers and students at CBI clearly exhibited solid and professional knowledge and expertise. Colleagues were from different yet related background. The research topics were diverse and all were very curious. Synergy was thus easy to set up. Additionally, there had been great chances to exchange ideas with external researches via invited talks and participation of conferences. Research resources were more than adequate and easy to acquire. Work environment was very pleasant.

Do you still have contacts from that time?

_ Yes, I still have contacts with previous colleagues. But these are mainly personal, rather than work-related.

What is the first thing that comes to your mind, when you think of your time at CBI?

_ An international team of talented researchers and students, all very kind, smart and extremely creative.

What did you like the most?

_ Impressive expertise of the colleagues. Whenever you have a question in your work, you could almost always find some experts in the group or the institute to discuss with. Colleagues were also very kind and willing to share their knowledge.

What was bothering you?

_ I don't recall any of that.

What was your further career?

_ I continued doing research in structural bioinformatics and related fields.

Nachgefragt bei...

Prof. Dr. Rudi Balling

Interdisziplinär und International. Daran denke ich, wenn das Gespräch auf das Zentrum für Bioinformatik in Saarbrücken kommt. Bioinformatik ist vom Grundansatz her interdisziplinär. Sie baut Brücken, z. B. zwischen der Biologie und den Computerwissenschaften oder der Biologie und der Medizin. Das ZBI war dabei schon immer Vorreiter und besonders erfolgreich. Hier wird nicht nur über eine enge Zusammenarbeit von universitärer und außeruniversitärer Forschung geredet. Hier wird sie gelebt. Das gilt auch für die Brücke zwischen dem Saarland und Luxemburg. Interessanterweise führt der Weg zwischen beiden über Schengen. ZBI und LCSB sind dabei wichtige Brückenpfeiler. Ich freue mich auf die weitere Zusammenarbeit und wünsche dem ZBI alles Gute!



Prof. Dr. Rudi Balling ist Gründungsdirektor des Luxembourg Center for Systems Biomedicine (LCSB) in Belval, Luxemburg und Professor an der Mehrsprachigen Forschungsuniversität in Luxemburg. Von 2001 bis 2009 war er Wissenschaftlicher Direktor des Helmholtz Zentrums für Infektionsforschung in Braunschweig, und davor Leiter des Institute für Säugetiergenetik am GSF Forschungszentrum in München (1993 – 2000). Er erhielt den Friedrich Wilhelm Preis der Universität Aachen im Jahr 1992. Seit 1999 ist er Ehrenmitglied der Assoziation für Amerikanische Anatomen und der Japanischen Gesellschaft für vererbte Stoffwechselkrankheiten. Und seit 2002 Mitglied der Berlin-Brandenburgischen Akademie der Wissenschaften. Seit 2007 ist er Honorarprofessor am Nationalen Institut für Genetik in Mishima, Japan. Seit 2007 ist er Präsident der Verbandes für Biologie, Biowissenschaften und Biomedizin in Deutschland (VBIO). Von 2003 bis 2007 war er Präsident des Verbandes der biowissenschaftlichen und biomedizinischen Fachgesellschaften (Vbbm), von 2002 bis 2004 Präsident der Deutschen Gesellschaft für Genetik (GfG) und von 2001 bis 2002 Präsident der International Mammalian Genome Society (IMGS).

Dr. Rolf Kaiser

Das Zentrum für Bioinformatik Saar, insbesondere die Abteilung von Thomas Lengauer, Max-Planck-Institut für Informatik, ist seit seiner Entstehung mein wesentlicher bioinformatischer Kooperationspartner. Ende der neunziger Jahre suchte ich erstmal die Kooperation mit seiner Forschungsgruppe. Diese Kooperation, die wir sowohl auf nationalem als auch auf europäischem Niveau in aktive interdisziplinäre Konsortien mit Klinikern und anderen Virologen einbetten konnten, war der Schlüssel für äußerst effektive transnationale Forschung im Bereich der Resistenzanalyse von pathogenen Viren. Dabei hat das Zentrum für Bioinformatik essenzielle Expertise in Methoden des statistischen Lernens einbringen sowie hohe interdisziplinäre Qualitäten im Bereich der Modellierung der entsprechenden medizinischen Probleme zeigen können.

Die aus der Kooperation resultierende Analysesoftware des geno2pheno Systems deckt heute einen breiten Bereich von Resistenzanalysen bei HIV, Hepatitis B und Hepatitis C ab. Sie war die erste Software ihrer Art, die auf statistischer Datenanalyse im Kontrast zu von Experten manuell zusammengestellten Regeln basierte. Mit dieser neuen Methodik hat sie das Gebiet der viralen Resistenzanalyse einen wesentlichen Schritt vorangebracht, und zwar nicht nur in der Forschung sondern auch in der klinischen Praxis, ein Tatbestand, der auch durch die Verleihung des AIDS Forschungspreis ist der Heinz-Ansmann Stiftung im Jahr 2010 gewürdigt wurde.

Solche Erfolge sind nur in einem Umfeld möglich, wie es das ZBI bietet: hier kommt methodische Kompetenz zusammen mit interdisziplinärer Qualität und einem hohen Maß an Einsatz, nicht nur für die Forschung sondern auch für deren Umsetzung in die Praxis. Die gegenseitige Freundschaft und das echte Interesse an den wissenschaftlichen Problemen des Partners, die wir in dieser Kooperation leben können, waren und sind für mich immer ein Quell der Inspiration.



Dr. Rolf Kaiser studierte Biologie an der Universität Bonn und promovierte dort im Fach Humangenetik über Molekularbiologische Analysen der Cystische Fibrose. Seit 1990 arbeitet er im Gebiet der Virologie, bis 1998 an der Universitätsklinik Bonn und seit 1999 an der Universitätsklinik Köln. Neben der allgemeinen Virusdiagnostik umfasst diese Tätigkeit auch Forschungsaktivität auf den Gebieten HIV, Hepatitis B Virus (HBV), Hepatitis C Virus (HCV) und Cytomegalovirus; sowie Aktivität in verschiedenen Fachausschüssen nationaler und internationaler Fachgesellschaften für Virologie und Mikrobiologie. In der Forschung spezialisiert sich Herr Kaiser auf die virologischen und medizinischen Aspekte der Evolution von Viren und des Ansprechens von Viren auf Wirkstoffe. Herr Kaiser war Initiator oder wesentlicher Gestalter von mehreren interdisziplinären Projekten in diesem Bereich mit nationalen und internationalen Partnern.

Prof. Dr.-Ing. Oliver Kohlbacher

Das Zentrum für Bioinformatik Saar hatte – und hat – einen ganz entscheidenden Einfluss auf die Bioinformatik-Landschaft in Deutschland. Ausgehend von den wissenschaftlichen Vorarbeiten am Fachbereich Informatik und dem MPI für Informatik, hat das Zentrum schon früh wesentliche Teilgebiete der Bioinformatik abgedeckt und mit geformt. Die Gründung des ZBI als eines der fünf Exzellenzzentren aus der DFG-Initiative Bioinformatik hat diese vorhandene Expertise dann gebündelt und Saarbrücken in der Bioinformatikforschung national und international auf die Landkarte gebracht. Forscher des ZBI waren und sind an internationalen Projekten beteiligt, bis hin zur historischen Entschlüsselung des menschlichen Genoms.

Die in Saarbrücken entwickelten Ausbildungskonzepte im Rahmen der Bioinformatikstudiengänge waren schon früher als an anderen Hochschulen breit aufgestellt und zwischen Methodik und Anwendung gut balanciert. Sie hatten damit Vorbildcharakter auch für die anderen Studiengänge, die in den Folgejahren an anderen Standorten entstanden und haben dadurch auch in der Ausbildung das Fach Bioinformatik maßgeblich mitdefiniert. Die Qualität dieser Ausbildung sieht man nicht zuletzt auch an der großen Zahl von Saarbrücker Absolventen, die heute leitende Stellungen in der Industrie oder auch Professuren in der Bioinformatik innehaben.



Prof. Dr.-Ing. Oliver Kohlbacher studierte an der Universität des Saarlandes Chemie und Informatik. Nach der Promotion bei Prof. Hans-Peter Lenhof leitete er am Zentrum für Bioinformatik eine Nachwuchsgruppe für Protein Docking. Nach einer Postdoc Zeit bei Prof. Eugene Myers (Celera Genomics, Rockville, MD), wurde er 2003 auf die Professur for Simulation biologischer Systeme (jetzt: Angewandte Bioinformatik) an der Universität Tübingen berufen. Seit 2012 ist er dort auch Direktor des Zentrums für Quantitative Biologie. 2015 wurde er zum Fellow am Max-Planck-Institut für Entwicklungsbiologie ernannt. Er ist stellvertretender Sprecher des Zentrums für personalisierte Medizin und stellvertretender Sprecher des Universitätsrates der Universität Tübingen. Schwerpunkte seiner Forschung liegen derzeit in der Entwicklung von Algorithmen zur Analyse biologischer Hochdurchsatzdaten (Proteomik, Metabolomik, Transkriptomik, Genomik) der Strukturbioinformatik und der Immunoinformatik.

Prof. Dr. Gene Myers

Schon seit seiner Gründung ist das Bioinformatikzentrum Saar ein starker international sichtbarer Brennpunkt für Forschung an der Grenze zwischen Informatik und Molekularbiologie bzw. Medizin. Das Zentrum hat viele der heutigen jungen Stars in diesem wichtigen Gebiet hervorgebracht und wird von seinen leitenden Wissenschaftlern mit klarer und markanter Vision im Hinblick auf sozial relevante Probleme und der Humangenetik, der Behandlung von Krankheiten beim Menschen und der Entwicklung biotechnologischer Lösungen geführt.



Prof. Dr. Eugene Myers ist Direktor am Institut für Molekulare Zellbiologie und Genetik in Dresden und Klaus Tschira Gründungsdirektor des dortigen Zentrums für Systembiologie. Von 2005 bis 2012 war Dr. Myers Gruppenleiter am Howard Hughes Medical Institute in Janelia Farms, Virginia, USA, und davor Professor für Informatik an der Universität von Kalifornien in Berkeley, USA (2002-2005). Von 1998 bis 2002 war er Vizepräsident für Informatikforschung bei Celera Genomics in Rockville, Maryland, USA, wo er die bioinformatische Basis zur Sequenzierung des Human Genoms gelegt hat. Davor war er 17 Jahre lang Professor für Informatik an der Universität von Arizona in Tucson. Seine Forschungsinteressen umfassen den Entwurf und die Analyse von Algorithmen in der Bioinformatik, derzeit vor allem in der Bildanalyse und Lichtmikroskopie mit einem Schwerpunkt auf Modellbildung von zellulären Systemen auf der Basis von Bilddaten.

Dr. Myers ist vor allem bekannt für die Entwicklung von BLAST – dem meistgenutzten bioinformatischen Programm – sowie für das Protokoll für die Paired-end Whole Genome Shotgun Sequenzierung und den Genomassembler, den er bei Celera Genomics entwickelt hat. Dr. Myers hat diese revolutionäre Technik im Jahr 1996 vorgeschlagen und konnte mit ihr in drei Jahren die Genome von Fliege, Mensch und Maus bestimmen.

Dr. Myers erhielt den IEEE 3rd Millennium Achievement Award in 2000, den Newcomb Cleveland Best Paper in Science Award in 2001 und den ACM Kanellakis Prize in 2002. Im Jahr 2004 erhielt er den Internationalen Max-Planck Forschungspreis. Das Genome Technology Magazine hat ihn zum einflussreichsten Bioinformatiker im Jahr 2001 bestimmt. Er hat ein Ehrendoktorat von der Eidgenössischen Technischen Hochschule in Zürich. Er ist Mitglied der National Academy of Engineering der USA und der Nationalen Akademie für Wissenschaften Leopoldina.

Prof. Dr. Martin Vingron

Vor kurzem promovierte bei mir (mit Auszeichnung) eine Studentin, die aus Saarbrücken zu uns nach Berlin kam. Da ich mit der Bioinformatik-Ausbildung in Saarbrücken gut vertraut bin, konnten wir von Anfang an auf diesem exzellenten Grundstock aufbauen und es fiel ihr nicht schwer, zwischen biologischen und mathematischen Themen hin- und herzuspringen.

Vor einigen Jahren ging auch ein früherer Doktorand von mir, der in den USA Postdoc gemacht hatte, als Gruppenleiter an das Saarbrückener Zentrum für Bioinformatik. Wir sind wiederum im Deutschen Epigenomprojekt miteinander verbunden und es macht Freude zu verfolgen, was er publiziert und erarbeitet. Selbst meine Praktikumsstudenten mussten schon zum Vergleich naive Algorithmen aus biologischen Publikationen und ein effizientes, in Saarbrücken entwickeltes Verfahren für das gleiche Problem implementieren, um zu erfahren, welchen Unterschied effiziente Algorithmen in der Praxis machen können. Jedenfalls sind die Kollegen des ZBI und ihre Arbeiten aus dem Zusammenhang der deutschen Bioinformatik nicht wegzudenken und tragen wesentlich zu ihrer internationalen Sichtbarkeit bei.



Prof. Dr. Martin Vingron studierte Mathematik in Wien und promovierte 1991 in Heidelberg an EMBL und Universität Heidelberg. Es schlossen sich zwei Postdoc-Aufenthalte zuerst an der University of Southern California in Los Angeles und dann an der GMD – Forschungszentrum Informationstechnik in Bonn an. Von 1995 bis 2000 war er Abteilungsleiter der Abteilung Theoretische Bioinformatik am Deutschen Krebsforschungszentrum in Heidelberg, bevor er einem Ruf als Direktor an das Max-Planck-Institut für molekulare Genetik in Berlin folgte.

Dort leitet er seither die Abteilung Bioinformatik. Von 2006 bis 2015 war Herr Vingron daneben Direktor am Partnerinstitut für Bioinformatik (PICB) der Max-Planck-Gesellschaft und der Chinesischen Akademie der Wissenschaften in Shanghai, China. Von 2008 bis 2014 leitete er das Steuerremium für die Internationale Bioinformatik-Konferenzserie ‚Research in Computational Biology‘ (RECOMB). Er erhielt 2004 den Max-Planck-Forschungspreis für Bioinformatik gemeinsam mit Gene Myers. Er ist Fellow der International Society for Computational Biology, sowie Mitglied der Akademie deutscher Naturwissenschaftler Leopoldina und der Academia Europaea.

Aktivitäten des Zentrums

SchülerUni

Zentrum für Bioinformatik präsentiert sich jährlich bei Schülerinnen und Schülern

Schülerinnen und Schüler ab Klasse 10, die Lust auf eine Entdeckungsreise in die Welt der Mathematik, Informatik und Naturwissenschaften haben, können an der Saarbrücker Schüler-Uni teilnehmen. Dieses Angebot besteht bereits seit mehreren Jahren und die Professoren des Zentrums für Bioinformatik haben in vielen Vorträgen und Besichtigungen das Zentrum präsentiert. Dabei wurden die Führungen sehr gut angenommen und die Möglichkeit, selbst am Computer zu arbeiten und dadurch die Inhalte des Studienfachs Bioinformatik kennen zu lernen, hat den Schülerinnen und Schülern viel Freude gemacht.



SchülerAkademie

Die Bioinformatik stellt sich vor

Das Zentrum für Bioinformatik hat sich mehrfach an der jährlichen SchülerAkademie beteiligt und in einem 14-tägigen Block die Bioinformatik an der Universität des Saarlandes vorgestellt. Wissenschaftler des Zentrums für Bioinformatik stellten ihr Know-How zur Verfügung und gaben zwei Wochen lang Einblicke in spannende Themen der Bioinformatik. Die saarländische SchülerAkademie findet in den beiden ersten Wochen der Sommerferien statt und bietet besonders begabten Schülerinnen und Schülern, die von ihren Schulen vorgeschlagen werden, die Möglichkeit, bestimmte Themengebiete vertieft kennenzulernen. Die Einführung der Teilnehmer in ein bestimmtes Themengebiet erfolgt durch Lehrkräfte, Vertreter der Hochschulen und durch andere, außerschulische Experten.

Jeder Kurs wird von zwei Kursleitern/Kursleiterinnen geleitet. Konkret werden folgende Ziele angestrebt:

- Erweiterung und Vertiefung vorhandener Kenntnisse
- Kennenlernen neuer Gebiete
- Wissenschaftliches Arbeiten
- Verbesserung der Technik des mündlichen Vortrags
- Schriftliche Darstellung unter Beachtung wissenschaftlicher Standards
- Selbständiger Wissenserwerb und selbständige Wissensverarbeitung
- Arbeit im Team und wechselseitige Nutzung der jeweiligen individuellen Stärken
- Abklärung persönlicher Begabungen und Fähigkeiten

Das Zentrum für Bioinformatik hat bei den Schülerinnen und Schülern großes Interesse geweckt und sehr gute Studierende gewinnen können.

Schülerinformationstag Bioinformatik 2015

Schülerinnen und Schüler eroberten das Zentrum für Bioinformatik

Alle saarländischen Gymnasien mit Schwerpunkt in den naturwissenschaftlich-mathematischen Fächern waren am 25.02.2015 eingeladen, sich vor Ort am Zentrum für Bioinformatik an der Universität des Saarlandes über das Fach Bioinformatik zu informieren. Zahlreiche Schulen waren vertreten. Ihre Schüler gewannen erste Eindrücke, als Peter Ebert in Vertretung von Prof. Lengauer über „Medizin aus dem Computer – Was kann die Bioinformatik“ sprach oder Lara Schneider und Florian Schmidt, die selbst Bioinformatik in Saarbrücken studierten und nun an ihrer Promotion arbeiten, über „Bioinformatik – ein Fach mit Zukunft“ referierten.

Die Schülerinnen und Schüler lauschten danach aufmerksam dem Vortrag von Prof. Dr. Andreas Keller über ‚die molekularen Geheimnisse des Ötzi‘. Die Entschlüsselung des Erbgutes der Gletschermumie „Ötzi“ führt zu spannenden Fragen: Wie die Krankheitsgeschichte von Ötzi wohl aussah? Und was kann man heute daraus lernen?

So regte auch der Vortrag von Dr. Marcel Schulz mit dem Thema „Dr. House 2.0: Diagnose genetischer Krankheiten mit dem Tablet“ zu vielen Fragen an.



Abi-Messen

Das Zentrum für Bioinformatik stellt sich vor

Das Zentrum für Bioinformatik hat sich in den Jahren 2008 bis 2014 an der Informationsmesse „Abi – was dann?“ beteiligt. Diese Messe, zu der alle Schulen im Saarland und den angrenzenden Bundesländern eingeladen werden, findet alle zwei Jahre statt. In der Congresshalle Saarbrücken gab es Stände mit Informationsmaterialien und Broschüren sowie Vorträge über die Inhalte der Bioinformatik. Hier fanden Abiturientinnen und Abiturienten jede Menge Infos, Tipps und Beratung. Darüber hinaus haben Studierende der Bioinformatik an Messen in Karlsruhe, Köln, Mainz und Homburg teilgenommen und den Studiengang ‚Bioinformatik‘ sowie die Software BALLView präsentiert.



Back-to-school-Programm

Das Zentrum für Bioinformatik besucht Schulen

Bioinformatik-Studierende im Zentrum für Bioinformatik haben ein Back-to-School Programm entwickelt, um in ihrer ehemaligen Schule das Fach Bioinformatik und das ZBI vorzustellen. Die Studierenden melden sich in ihrer ehemaligen Schule an und halten einen 30-minütigen Vortrag über ihr Studienfach Bioinformatik. Dabei geht es nicht nur um Inhalte des Studienfaches, sondern auch um persönliche Eindrücke.

Im allgemeinen Teil geht es um den Studienstandort Saarbrücken und die Forschung im Zentrum für Bioinformatik. Die Schülerinnen und Schüler sollten ein minimales Vorwissen im Bereich Biologie mitbringen, um konkrete Inhalte wie Medikamentenentwicklung oder Genomsequenzierung verstehen zu können. Hierdurch wird ein großes Interesse an der Bioinformatik geweckt und die Lust auf ein Studium in Saarbrücken gefördert.

BIOINFORMATICS STUDENTS GO...
... BACK TO SCHOOL

FAHRE AN DEINE ALTE SCHULE UND MACH WERBUNG FÜR DEIN STUDIENFACH

ZBI
 ZENTRUM FÜR
 BIOINFORMATIK

DIE FAHRTKOSTEN IN DIE HEIMAT WERDEN NATÜRLICH ÜBERNOMMEN UND DU BEKOMMST EIN HONORAR
 Bei Interesse und für weitere Informationen, melde dich bei uns!

back2school@bioinf.uni-sb.de

Unicamp für Mädchen

Schülerinnen interessieren sich für Bioinformatik

Beim Unicamp, das jedes Jahr im Sommer stattfindet, lernen Schülerinnen der Klassenstufen 8 und 9 die naturwissenschaftlich-technischen Studiengänge der Universität des Saarlandes kennen und erfahren, welche Berufe sie mit einem solchen Studium ergreifen können. In Vorträgen, Workshops und praktischen Übungen erleben die Mädchen Forschung und Technik zum Anfassen.

Das Zentrum für Bioinformatik hat bisher in jedem Jahr am UniCamp teilgenommen. Im Rechnerraum des ZBI wurden Vorträge gehalten und die Schülerinnen hatten die Möglichkeit, selbst am Rechner zu programmieren. Sie wurden unter Anleitung von Bioinformatik-Studierenden an bioinformatisches Arbeiten herangeführt.



Girls' Day

Naturwissenschaft und Technik zum Anfassen

Für Schülerinnen ab Klassenstufe 8 findet jedes Jahr im April der bundesweite Girls' Day statt. Das Motto des Tages lautet „Naturwissenschaft und Technik zum Anfassen“. Wissenschaftlerinnen der Uni und des Zentrums für Bioinformatik geben in kurzen Vorträgen mit zahlreichen Anschauungsbeispielen, Laborführungen und Experimenten Einblicke in die verschiedenen Gebiete der Naturwissenschaft und Bioinformatik.

Doktorandentag am ZBI

Poster Session der PhD-Studenten

In Zusammenarbeit mit der Graduiertenschule für Informatik veranstaltete das Zentrum für Bioinformatik einen ‚Doktorandentag‘. Hier haben alle Doktoranden der verschiedenen Gruppen im Zentrum für Bioinformatik die Möglichkeit ihre Forschungsthemen vorzustellen. Die Doktoranden präsentierten ihre Forschungsschwerpunkte und diskutierten darüber. Die Idee war es, den Doktoranden die Möglichkeit zur Erweiterung ihres Wissens zu geben und den Austausch im Bereich der Forschung zu fördern.



Hochschulinformationsbesuche (HIB)

Abiturienten informieren sich über das Studienfach Bioinformatik

Bei den Hochschulinformationsbesuchen stellen auch Bioinformatik- und Informatik-Professoren jährlich an verschiedenen Nachmittagen im Februar und März ihre Studiengänge vor und geben so die Gelegenheit, sich über Inhalte, Anforderungen und Abschlüsse sowie Bewerbungsverfahren und Berufsperspektiven der einzelnen Studiengänge zu informieren. Das Zentrum für Bioinformatik war dabei immer beteiligt.

So hat unter anderem Prof. Volkhard Helms im Jahr 2014 einen Vortrag gehalten mit dem Titel: „Zentrum für Bioinformatik Saar – Bioinformatik für pharmazeutische und medizinische Probleme.“ Prof. Andreas Keller referierte 2015 über das Thema: „Vom Bluttest zur Genom-Sequenzierung. Die Bedeutung der Informatik für das Gesundheitswesen.“



Prof. Volkhard Helms



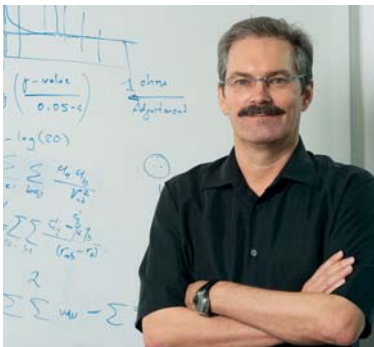
Prof. Andreas Keller

Tag der offenen Tür

Das Zentrum für Bioinformatik präsentiert sich

Alljährlich lädt die Universität des Saarlandes zu einem Tag der offenen Tür ein und somit zu einer Reise über den Campus Saarbrücken. Auch das Zentrum für Bioinformatik präsentiert sich: Hier werden Einblicke gegeben, wie unsere Gene gesteuert werden und Bioinformatiker zeigen, wie Moleküle auf dem Computer dargestellt werden können. Auch Technikfans kommen auf ihre Kosten. Weiterhin gibt es zahlreiche Informations-Broschüren und ganztägig Information und Beratung am Stand der Bioinformatik.

Prof. Hans-Peter Lenhof informierte in den Jahren 2012 bis 2015 in seinem Vortrag „Vom Genom zum Medikament“ wie die Bioinformatik zur Medikamentenentwicklung beitragen kann.



Das Gebäude E21 der Bioinformatik auf dem Campus Saarbrücken.

MS-Wissenschaft – neue Wege in der Medizin

Zentrum für Bioinformatik präsentiert BALLView – Medikamentenentwurf am Computer

Die Wissenschaftsausstellung an Bord der MS Wissenschaft zeigte im Jahr 2011, woran Wissenschaftlerinnen und Wissenschaftler arbeiten, um Krankheiten zu erkennen, zu heilen und die Lebensqualität der Menschen zu steigern.

Im „Wissenschaftsjahr 2011 – Forschung für unsere Gesundheit“ war an Bord eine Ausstellung zu Themen der Gesundheitsforschung zu sehen, die faszinierende Einblicke in die aktuelle Forschung bot.

Zu sehen war auch ein Ausstellungsstück des Zentrums für Bioinformatik. Die Software BALLView ist ein hilfreiches Instrument für die Entwicklung von Medikamenten am Computer. Moleküle können in dreidimensionaler Form sichtbar gemacht werden und man erkennt, wie Medikamente an ihre Zielproteine binden.



60. Geburtstag Thomas Lengauer

Sprecher des Zentrums feiert runden Geburtstag

Zum Anlass des 60. Geburtstages von Thomas Lengauer lud das Zentrum für Bioinformatik zusammen mit dem Max-Planck-Institut für Informatik zum Festkolloquium in den Günther-Hotz-Hörsaal ein. Die Professoren Kurt Mehlhorn, Max-Planck-Institut für Informatik, Burkhard Rost, TU München und Volkhard Helms, Universität des Saarlandes/ Zentrum für Bioinformatik, lobten als Laudatoren Thomas Lengauer für seine wissenschaftliche Arbeit und seinen Schaffensdrang. Sie stellten heraus, dass der Jubilar in der Lage sei, schneller als die meisten anderen Dinge anzuschieben und voran zu treiben. Alle zehn Jahre habe er ein neues Thema innerhalb der Wissenschaft begonnen und seine Beiträge darin zur Weltspitze geführt.

Sechs Gastredner steckten mit ihren Vorträgen das Gebiet der Bioinformatik ab. Drei von ihnen, Niko Beerenwinkel, Christoph Bock und Matthias Rarey, waren ehemalige Doktoranden von Thomas Lengauer, die inzwischen selbst Lehrstühle besetzen; drei weitere Fachkollegen, Eugene Myers, Burkhard Rost und Martin Vingron, komplettierten das Sextett. Das Auditorium hörte in komprimierter Form eine breite Auswahl von Themen rund um die Bioinformatik auf internationalem Spitzenniveau.



Zehn Jahre Zentrum für Bioinformatik an der Universität des Saarlandes

Hochkarätige Wissenschaftler würdigen Zentrum für Bioinformatik

Im Jahr 2011 beging das Zentrum für Bioinformatik sein zehnjähriges Jubiläum. In diesen zehn Jahren haben die Wissenschaftler biologische Prozesse im Körper und das menschliche Erbgut durchleuchtet, um zum Beispiel Krankheiten besser heilen zu können. Am 23.03.2011 feierte das Zentrum für Bioinformatik sein zehnjähriges Bestehen mit einem internationalen Kolloquium.



Der Präsident der Universität des Saarlandes, Prof. Dr. Volker Linneweber, begrüßte fast 300 Gäste zur 10-Jahresfeier des Zentrums für Bioinformatik

Der Sprecher des Zentrums für Bioinformatik, Prof. Dr. Dr. Thomas Lengauer, Direktor am Max-Planck-Institut für Informatik, eröffnete das Kolloquium, begrüßte die weit angereisten Gäste und stellte die hochkarätigen Redner des Kolloquiums vor. Er zählt zu den Gründern des noch jungen Fachgebietes der Bioinformatik.



Als einer der eingeladenen prominenten Redner begeisterte Gene Myers vom Howard Hughes Medical Institute (USA) das Publikum. Er hatte zehn Jahre davor maßgeblich dazu beigetragen, das menschliche Genom zu entschlüsseln – auch mit Unterstützung von Saarbrücker Bioinformatikern. Er hielt am Nachmittag einen sehr interessanten Vortrag zum Thema „*Image-based Informatics for Molecular Biology*“, der großen Applaus erntete.



Bei diesem internationalen Kolloquium präsentierten auch die Nachwuchsforscher des Zentrums für Bioinformatik ihre aktuellen Forschungsergebnisse. Dr. Andreas Hildebrandt zeigte die Software BALLView, mit der man neue Wirkstoffe am Bildschirm entwerfen und visualisieren kann. Damit können sich jetzt auch internationale Forscherteams über das Internet vernetzen, um gleichzeitig dreidimensionale Modelle von Molekülen zu bearbeiten. Dr. Alice McHardy hat sich mit der Frage befasst, wie man große Gendatenbanken sinnvoll analysiert und Dr. Mario Albrecht stellte vor, wie die Bioinformatik dabei helfen kann, Virusinfektionen zu bekämpfen.

Am Nachmittag präsentierten dann führende Bioinformatiker aus Deutschland, Dänemark, Großbritannien und den USA ihre Forschungsprojekte, vor allem zur Molekularbiologie.

Prof. Dr. Janet Thornton, Direktorin des European Bioinformatics Institute in England sprach über Computeranalysen von molekularen Strukturen und enzymatischen Wechselwirkungen. Prof. Dr. Richard Karp aus Berkeley, einer der weltweit bekanntesten Informatiker, der unter anderem den Begriff NP-Vollständigkeit eingeführt hat, zeigte ein zentrales Konzept in der theoretischen Informatik auf. Weitere Redner waren Prof. Dr. Søren Brunak von der Technischen Universität Lyngby in Dänemark und Prof. Dr. Helmut Grubmüller vom Max-Planck-Institut für biophysikalische Chemie in Göttingen.

Preise und Auszeichnungen

Grant Fertility Innovation (GFI) 2015 für Andreas Keller

Andreas Keller, Professor für Klinische Bioinformatik der Universität des Saarlandes, wurde für sein Projekt „Improved diagnostic of in-vitro-fertilization using miRNome and microvesicles from early stage embryos“ mit dem „Grant Fertility Innovation“ ausgezeichnet. Keller, der am Zentrum für Bioinformatik forscht, wurde als Preisträger des GFI-Förderprogramms anlässlich der Jahrestagung ESHRE in Lisbon (Portugal) bekanntgegeben. Der Preis wird jährlich an Forscher vergeben, die innovative Forschungsprojekte im Bereich Fertilität zu konkreten Gesundheitslösungen weiterentwickeln, um die Erfolgsrate der Reproduktionstechnologie zu verbessern. Bei der aktuellen Vergabe wurden sechs Projekte aus fünf Ländern mit Fördergeldern von insgesamt 1,2 Millionen Euro bedacht.



Hector Wissenschaftspreis 2015 für Thomas Lengauer

Thomas Lengauer, Sprecher des Zentrums für Bioinformatik und Direktor am Max-Planck-Institut für Informatik in Saarbrücken, wurde mit dem Hector-Wissenschaftspreis 2015 ausgezeichnet. Die Hector Stiftung II würdigt damit seine Leistungen auf dem Gebiet der Bioinformatik und sein Engagement in der Hochschullehre.

Der Preis wird jährlich an herausragende Forscher deutscher Universitäten vergeben und ist mit 150.000 Euro dotiert. Zu dem Preis gehört auch die Ernennung zum Fellow der Hector Fellow Academy, die interdisziplinäre Projekte und akademische Netzwerke fördert. „Ich freue mich sehr auf die fachübergreifenden Diskussionen in der Hector Fellow Akademie“ betonte Prof. Lengauer, „denn an den Schnittstellen zwischen den Disziplinen liegt der Schlüssel für viele künftige Durchbrüche in der Wissenschaft.“

Verdienstkreuz am Bande für Norbert Graf

Prof. Dr. Norbert Graf, Direktor der Klinik für Pädiatrische Onkologie und Hämatologie an den Universitätskliniken und Mitglied im Zentrum für Bioinformatik, wurde 2011 vom damaligen Minister für Wirtschaft und Wissenschaft, Dr. Christoph Hartmann, das Verdienstkreuz am Bande des Verdienstordens der Bundesrepublik Deutschland verliehen.

Neben seiner Tätigkeit in der Krankenversorgung arbeitet er auch in mehreren wissenschaftlichen Studien mit. So ist er auch an der internationalen Nephroblastomstudie (Wilmstumorstudie) beteiligt. Über das gleiche Thema kooperiert er auch mit den Professoren Meese und Lenhof (beide Mitglied im Zentrum für Bioinformatik).



Heinz-Ansmann-Preis für Aidsforschung

Professor Thomas Lengauer ist gemeinsam mit Dr. Rolf Kaiser (Virologisches Institut der Universität zu Köln) und Dr. Mark Oette (Krankenhaus der Augustinerinnen Köln) im Jahr 2011 mit dem Heinz-Ansmann-Preis für AIDS-Forschung ausgezeichnet worden. Das Team hat über mehr als zehn Jahre bioinformatische Methoden erarbeitet, um Resistenzen des HI-Virus gegen verschiedene Wirkstoffe zu untersuchen. Die auf der Basis der gesammelten Daten entwickelte Software ist frei im Internet verfügbar (www.geno2pheno.org).

Der Heinz-Ansmann-Preis für AIDS-Forschung ist mit 15.000 Euro dotiert. Er wird seit den Achtzigerjahren alle zwei Jahre vergeben. Stifter ist die Düsseldorfer Heinz-Ansmann-Stiftung für AIDS-Forschung.



PD Dr. Rolf Kaiser, Professor Thomas Lengauer und PD Dr. Mark Oette



Professor Rolf Müller mehrfach ausgezeichnet

Für seine Forschungsarbeit wurde Professor Rolf Müller, langjähriges Mitglied im Zentrum für Bioinformatik, bereits zweimal mit dem Phoenix-Pharmazie Wissenschaftspreis ausgezeichnet (2001, 2007) und erhielt den DECHEMA Nachwuchswissenschaftler-Preis für Naturstoff-Forschung (2002), den BioFuture-Preis des Bundesministeriums für Bildung und Forschung (2003) sowie den DECHEMA-Preis der Max-Buchner-Forschungstiftung (2010).

2012 wurde Rolf Müller in die Deutsche Akademie der Technikwissenschaften (acatech) gewählt.

Mehrfache Auszeichnung für Professor Lehr

Prof. Dr. Claus-Michael Lehr, von Anfang an Mitglied im Zentrum für Bioinformatik, wurde mehrfach ausgezeichnet. Er erhielt für die Entwicklung tierversuchsfreier Methoden zur Untersuchung von Darmentzündungen 2011 den Forschungspreis des Bundesministeriums für Ernährung, Landwirtschaft und Verbraucherschutz „zur Förderung methodischer Arbeiten mit dem Ziel der Einschränkung und des Ersatzes von Tierversuchen“. Das Land Rheinland-Pfalz verlieh ihm 2011 den Forschungspreis „Ersatz- und Ergänzungsmethoden zum Tierversuch“ für sein Projekt „A three-dimensional coculture of enterocytes, monocytes and dendritic cells to model inflamed intestinal mucosa in vitro“.

2010 erhielt er den Dorothy Hegarty Award für das beste Paper in ATLA (Alternatives to Laboratory Animals) und wurde zum AAPS-Fellow (American Association of Pharmaceutical Scientists) ernannt.



ERC Starting Grant 2016

Dr. Christoph Bock erhält aktuell den ERC Starting Grant (2016-2021) für geplante Forschungsarbeiten im Bereich der funktionellen Epigenomik und ihrer Anwendung bei Krebs. Die öffentliche Ankündigung erfolgte im Dezember 2015.



Otto-Hahn-Medaille

Dr. Christoph Bock wurde für seine Dissertation im Bereich ‚Computational Epigenetics‘ 2009 mit der Otto-Hahn-Medaille ausgezeichnet. Sie ist ein Ansporn für eine Karriere in der Wissenschaft und wurde ihm für besondere Leistungen im Rahmen seiner wissenschaftlichen Arbeit verliehen.



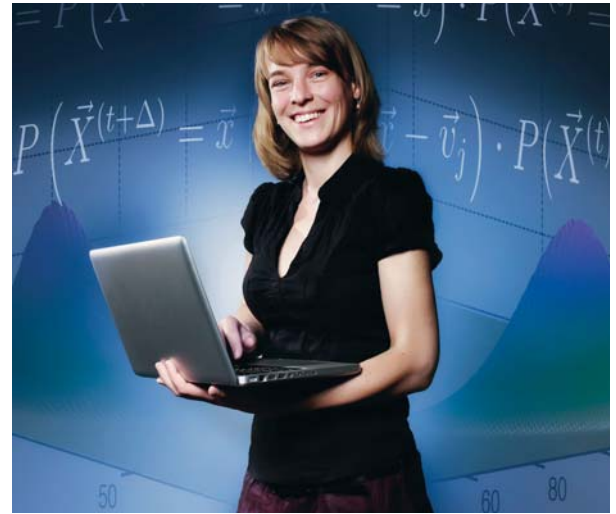
Eduard-Martin-Preis

Dieser Dissertationspreis wird jedes Jahr an die beste Doktorandin oder den besten Doktoranden aus den einzelnen Fakultäten der Universität des Saarlandes vergeben. Der Preis wurde nach Dr. Eduard Martin benannt, dem langjährigen Präsidenten der Vereinigung der Freunde der Universität des Saarlandes. Dr. Christoph Bock (2010) und Dr. Sven-Eric Schelhorn (2015) haben diesen Preis erhalten.

Die Preisverleihung im Jahr 2015 fand im neuen Graduate Centre der Universität statt. Bei der Festveranstaltung der Universitätsgesellschaft des Saarlandes erhielten die Preisträger einen Geldpreis und – mit einem leichten Augenzwinkern – dazu eine sympathisch bunte Eule.

Preis für Innovatoren unter 35

Im Rahmen des Nachwuchswettbewerbs „Innovatoren unter 35“ des Wissenschaftsmagazins „Technology Review“ wurde die Informatik-Professorin Verena Wolf 2013 für ihre Forschungen ausgezeichnet. Sie entwickelte einen Algorithmus, der es erlaubt, die Vorgänge in Zellen mit statistischen Methoden zu berechnen.



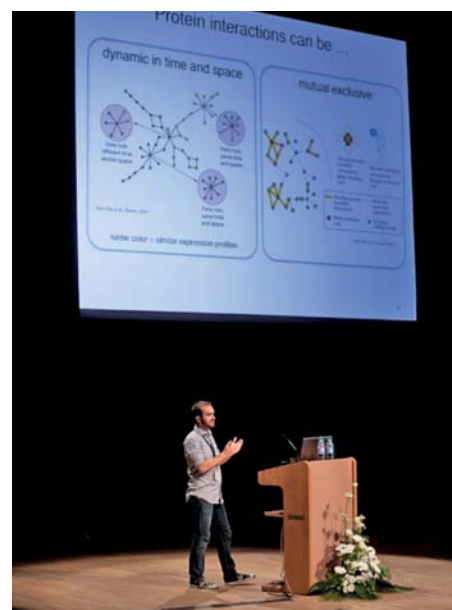
HUPO Early Career Investigator

Die internationale Wissenschaftsorganisation „Human Proteome Organization“ (HUPO) repräsentiert weltweit den Bereich „Proteomics“ und verlieh im Jahr 2010 Dr. Mario Albrecht den Early Career Investigator-Preis für seine Forschung im Bereich „Human Disease“.

Auszeichnung auf der ECCB 2014 (European Conference on Computational Biology)

Thorsten Will, wissenschaftlicher Mitarbeiter in der Gruppe von Professor Helms, erhielt auf der ECCB 2014 die Auszeichnung „Honorable Proceedings Papers“ für sein dort vorgestelltes Paper.

Außerdem wurde die Masterarbeit von Thorsten Will in die Buchreihe BestMasters aufgenommen, mit der der Springer-Verlag die besten Masterarbeiten auszeichnet, die an renommierten Hochschulen in Deutschland, Österreich und der Schweiz entstanden sind.



RCSB PDB Poster Preis für Matthias Dietzen

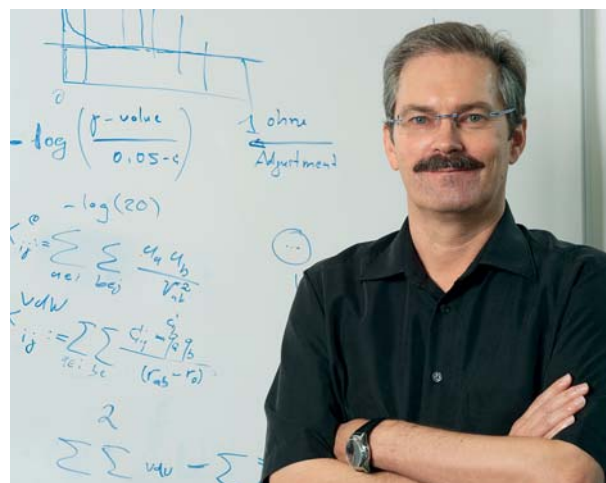
Im Rahmen gemeinsamer Konferenzen, der 21. Conference on Intelligent Systems for Molecular Biology (ISMB) und der 12. European Conference on Computational Biology (ECCB), ist Matthias Dietzen, Doktorand in der Arbeitsgruppe von Professor Lengauer, im Jahr 2013 in Berlin mit dem RCSB PDB Posterpreis ausgezeichnet worden. Der Preis wird traditionell für die beste Posterpräsentation im Bereich „Protein Structure and Function Prediction and Analysis“ vergeben und Matthias Dietzen setzte sich in diesem Jahr unter 106 Postern durch.



Prof. Dr. Hans-Peter Lenhof erhielt 2010 und 2013 vom Fachschaftsrat der Informatikstudiengänge den Preis für herausragende Leistungen in der Lehre für seine Vorlesung „Bioinformatik 1“.

Preise für herausragende Leistungen in der Lehre

Professor Thomas Lengauer erhielt im Wintersemester 2013/2014 sowie im Wintersemester 2014/2015 vom Fachschaftsrat der Informatikstudiengänge den Preis für herausragende Leistungen in der Lehre für seine Vorlesung „The Elements of Statistical Learning“.



Kolloquien und Vorträge

108

Kolloquien (Distinguished Speaker Series): Seit 2009 veranstaltet das ZBI ein ‚Distinguished Speaker Colloquium‘. Hier trägt während der Vorlesungszeit monatlich eine herausragende Forscherpersönlichkeit zu ihrem Themenbereich vor.

Der Vortrag richtet sich an Studierende und Forscher der Bioinformatik am Standort Saarbrücken. Neben den fachlichen Inhalten trägt die Serie dazu bei, durch Vorstellung von Vorbildern der Bioinformatik die Nachwuchswissenschaftler zu wissenschaftlichen Karrieren in diesem Gebiet zu inspirieren.

Datum	Name	Institution / Firma	Thema
16.12.2015	Prof. Dr. Torsten Schwede	<i>Universität Basel, Biozentrum und Schweizer Institut für Bioinformatik, Schweiz</i>	Modeling Protein Structures and Complexes using Evolutionary Information
11.11.2015	Prof. Dr. Uwe Ohler	<i>Max-Delbrück-Centrum für Molekulare Medizin, Berlin</i>	Predictive Models for Eukaryotic Gene Regulation: from Transcription to Translation
08.07.2015	Prof. Dr. Eric Westhof	<i>University of Strasbourg, Institute for Molecular and Cell Biology of CNRS, Frankreich</i>	The Detection of Architectural Modules and of Protein Binding Residues in RNA Sequences – Progress in RNA-Puzzles Contest
10.06.2015	Prof. Dr. Ines Thiele	<i>University of Luxembourg, Luxembourg Centre for Systems Biomedicine</i>	Computational Modeling of Host Metabolism and Its Interaction with the Gut Microbiome
27.05.2015	Prof. Dr. Paul de Bakker	<i>Utrecht University, Medical Center, Niederlande</i>	Inherited and De Novo Variation in Dutch Genomes
15.04.2015	Prof. Dr. Ulrich Mansmann	<i>Ludwig-Maximilians-Universität München, Institut für Medizinische Informationsverarbeitung, Biometrie und Epidemiologie</i>	The Molecular Tumor Board – Clinical Information Provided by Validated Bioinformatics Tools
04.02.2015	Prof. Dr. Michael Sternberg	<i>Imperial College London, Center for Bioinformatics, UK</i>	Phyre: Modelling Protein Structures and their Association with Disease
14.01.2015	Prof. Dr. Markus Loeffler	<i>Universität Leipzig, Institut für Medizinische Informatik, Statistik und Epidemiologie</i>	Clinical Trial Concepts in Personalized Medicine
10.12.2014	Prof. Dr. Mihaela Zavolan	<i>Universität Basel, Biozentrum, Schweiz</i>	Prediction, Identification and Functional Characterization of miRNA-Target Interactions
12.11.2014	Dr. Janet Kelso	<i>Max-Planck-Institut für evolutionäre Anthropologie, Leipzig</i>	Reconstructing Ancient Human Genomes
22.10.2014	Dr. Jasmin Fisher	<i>Microsoft Research, Cambridge, UK</i>	Computing Cancer
02.07.2014	Dr. Wolfgang Huber	<i>Europäisches Laboratorium für Molekularbiologie (EMBL), Heidelberg</i>	Gene-Gene and Gene-Drug Interactions
11.06.2014	Prof. Dr. Jörg Stelling	<i>Eidgenössische Technische Hochschule (ETH) Zürich, Abteilung Biosysteme, Zürich, Schweiz</i>	Automated Generation of Predictive Models for Systems Biology

Datum	Name	Institution / Firma	Thema
21.05.2014	Prof. Dr. Daniel Huson	<i>Eberhard Karls Universität Tübingen, Institut für Informatik</i>	Computational Metagenome Analysis
30.03.2014	Prof. Dr. Ursula Kummer	<i>Ruprecht-Karls-Universität Heidelberg, Institut für Biowissenschaften</i>	From Cellular Processes to Whole-Body Distributions – Integrating Different Scales in Computational Models
05.02.2014	Prof. Dr. Gunnar Rätsch	<i>Memorial Sloan-Kettering Cancer Center, New York City, USA</i>	Understanding and Modeling Transcriptome Variability
15.01.2014	Prof. Dr. Martin Fussenegger	<i>Eidgenössische Technische Hochschule (ETH) Zürich, Abteilung Biosysteme, Schweiz</i>	Prosthetic Networks – a New Treatment Strategy for Metabolic Disorders
11.12.2013	Prof. Dr. Nikolaus Rajewski	<i>Max-Delbrück-Centrum für Molekulare Medizin, Berlin</i>	Circular and Micro RNAs
20.11.2013	Prof. Dr. Knut Reinert	<i>Freie Universität Berlin, Institut für Informatik</i>	Bridging the Gap: Enabling Top Research in Translational Research
16.10.2013	Prof. Dr. Yves Moreau	<i>SymBioSys Center, University of Leuven, Belgien</i>	Prioritizing Disease-Causing Variation by Genomic Data Fusion
10.07.2013	Prof. Dr. Christine Orengo	<i>University College London, School of Life and Medical Sciences, London, UK</i>	Exploiting Protein Evolution to Predict Protein Functions and Functional Networks
20.06.2013	Prof. Dr. Johannes Söding	<i>Ludwig-Maximilians-Universität München, Genzentrum München und Institut für Biochemie</i>	Four Core Promoter Classes with Characteristic Regulatory Properties
22.05.2013	Prof. Dr. Luis Serrano Pubull	<i>Centre Regulació Genòmica (CRG), Barcelona, Spanien</i>	A Quantitative Systems Biology on a Model Bacterium
17.04.2013	Prof. Dr. Hermann-Georg Holzhütter	<i>Humboldt-Universität zu Berlin, Charité, Institut für Biochemie</i>	Mathematical Modeling of the Cellular Metabolism
06.02.2013	Prof. Dr. Henk Stunnenberg	<i>Nijmegen Centre for Molecular Life Sciences (NCMLS), Niederlande</i>	The Epigenetic Landscape of Ground State Pluripotent Stem Cell
09.01.2013	Prof. Dr. Uli Sattler	<i>Information Management Group (IMG), University of Manchester, UK</i>	OWL Ontologies – Magic Dust, Snake Oil, or the Best Thing since Sliced Bread?
14.11.2012	Prof. Dr. Michele Vedruscolo	<i>University of Cambridge, Department of Chemistry, UK</i>	Determination of the Structure and Dynamics of Proteins using NMR Chemical Shifts
11.07.2012	Prof. Dr. Robert Russell	<i>Universität Heidelberg, Biochemie-Zentrum</i>	Predicting Protein-Protein, Protein-Peptide and Protein-Chemical Interactions within Biological Networks
20.06.2012	Prof. Dr. Uwe Sauer	<i>Eidgenössische Technische Hochschule Zürich, Institut für Molekulare Systembiologie, Schweiz</i>	Identification of Allosteric and Transcriptional Regulation that Controls Metabolic Function

Datum	Name	Institution / Firma	Thema
16.05.2012	Prof. Dr. Marie-France Sagot	<i>Université de Lyon, INRIA Grenoble, Frankreich</i>	Towards an Algorithmic and Mathematical Exploration of Symbiosis
18.04.2012	Prof. Dr. Sarah Teichmann	<i>MRC Laboratory of Molecular Biology, Cambridge, UK</i>	From Protein Interactions to Gene Expression Distributions
08.02.2012	Prof. Dr. Ron Shamir	<i>Tel Aviv University, School of Computer Science, Israel</i>	Computational Analysis of Gene Regulation, Disease Classification and Protein Networks
11.01.2012	Prof. Dr. Ewan Birney	<i>European Bioinformatics Institute (EBI), Hinxton, Cambridge, UK</i>	Understanding Biology using Individual Differences
14.12.2011	Prof. Dr. Paul Flicek	<i>European Bioinformatics Institute (EBI), Hinxton, Cambridge, UK</i>	Insights into Vertebrate Genome Evolution from Comparative Functional Genomics
30.11.2011	Prof. Dr. Amos Tanay	<i>Weizmann Institute of Science, Rehovot, Israel</i>	The Evolution of Normal and Aberrant Epigenomes
19.10.2011	Prof. Dr. Martin Zacharias	<i>Technische Universität München Physik-Department</i>	Studying Nucleic Acid Flexibility Using Molecular Dynamics and Advanced Sampling Methods
13.07.2011	Prof. Dr. Thomas Höfer	<i>Deutsches Krebsforschungszentrum (DKFZ), Heidelberg</i>	Dynamic Regulatory Networks of Gene Expression
11.05.2011	Prof. Dr. Jens Timmer	<i>Institute for Advanced Studies (FRIAS)</i>	Systems Biology of the Erythropoietin Receptor
13.04.2011	Prof. Dr. Burkhard Rost	<i>Technische Universität München, Institut für Informatik; Columbia University, New York, USA</i>	Evolution Teaches Protein Prediction
09.02.2011	Prof. Dr. Gisbert Schneider	<i>Eidgenössische Technische Hochschule Zürich (ETH), Institut für Pharmazeutische Wissenschaften, Schweiz</i>	From Protein Pockets to Ligands and Back Again
12.01.2011	Prof. Dr. Wolfgang Wiechert	<i>Forschungszentrum Jülich</i>	Tools and Algorithms for Biochemical Network Modeling
17.11.2010	Prof. Dr. Arne Elofsson	<i>Stockholm University, Department of Biochemistry & Biophysics, Schweden</i>	Multidomain Protein Evolution and Orphans
20.10.2010	Prof. Dr. Peter Stadler	<i>Universität Leipzig, Institut für Informatik</i>	The Long and the Short of RNAs: Complex Processing of ncRNAs
07.10.2010	Prof. Dr. Christian von Mering	<i>Universität Zürich, Institut für Molekulare Biologie, Schweiz</i>	A Global Network of Co-existence Relationships among Microbes in the Wild
09.06.2010	Prof. Dr. Martin Vingron	<i>Max-Planck-Institut für Molekulare Genetik, Berlin</i>	Transcription Factor Binding Sites and Chromatin Modifications Point at Two Classes of Human Promoters
19.05.2010	Dr. Eran Segal	<i>The Weizmann Institute of Science, Department of Computer Science and Applied Mathematics, Rehovot, Israel</i>	Transcriptional Lego: Predictable Control of Gene Expression by Manipulating Promoter Building Blocks

Datum	Name	Institution / Firma	Thema
13.01.2010	Prof. Dr. Matthias Reuss	<i>Universität Stuttgart, Forschungszentrum für Systembiologie</i>	Dynamic Modeling of the Central Metabolism of E.coli - Linking Metabolite and Regulatory Networks
09.12.2009	Prof. Dr. Matthias Rarey	<i>Universität Hamburg, Zentrum für Bioinformatik</i>	Computing in Life Science: From Virtual Screening to Lead Optimization
11.11.2009	Dr. Rebecca Wade	<i>Europäisches Laboratorium für Molekularbiologie (EMBL), Heidelberg</i>	Bridging from Molecular Simulation to Biochemical Networks
14.10.2009	Dr. Yanay Ofran	<i>Bar Ilan University, Laboratory of Systems Biology and Functional Genomics, Ramat Gan, Israel</i>	From a Single Residue to Disease, a Multilevel Approach to Function Prediction

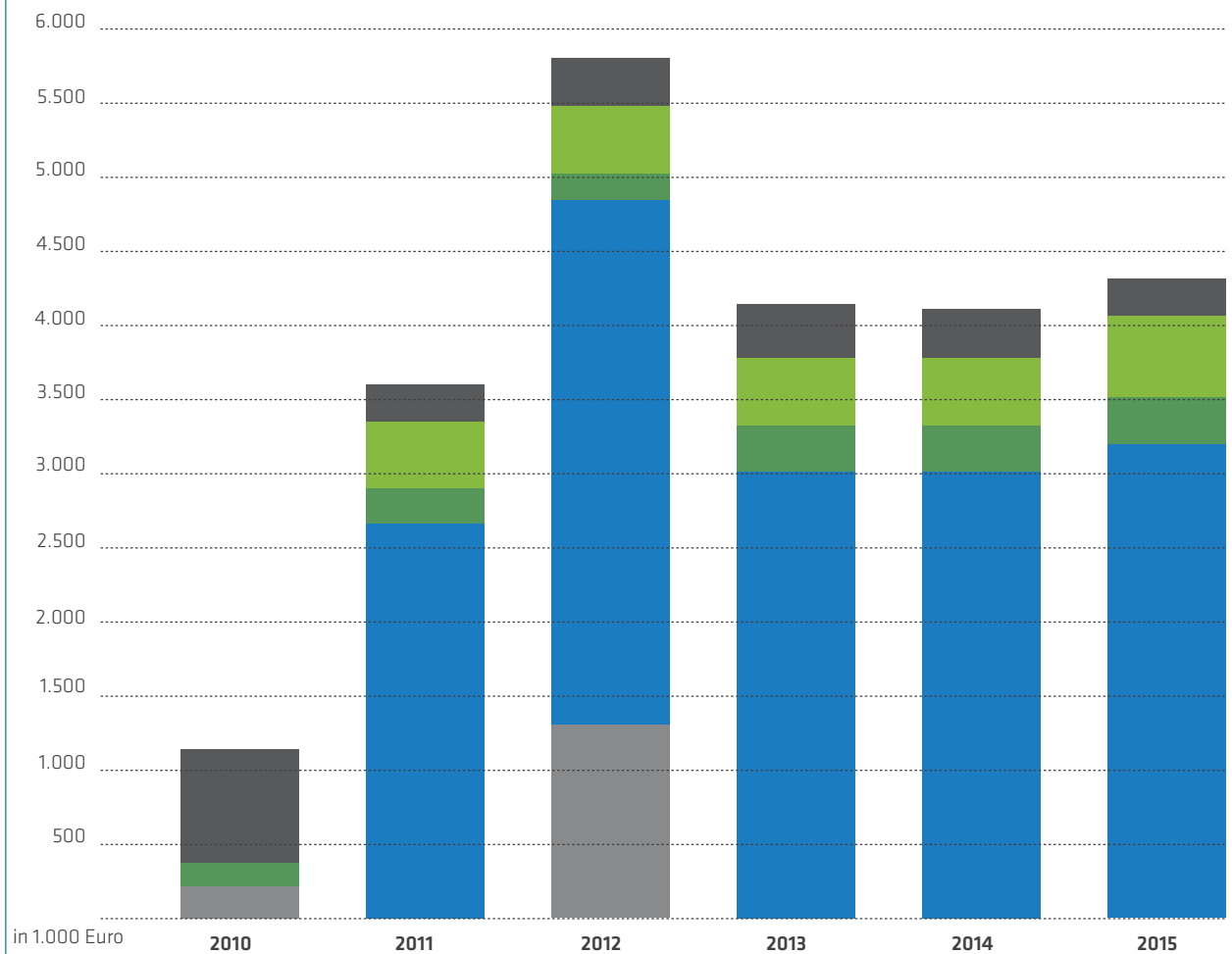
Die folgende Liste umfasst weitere zentrumsweit bekannt gegebene eingeladene Vorträge:

Datum	Name	Institution / Firma	Thema
12.01.2014	Prof. Dr. Andreas Keller	Universität des Saarlandes, Lehrstuhl für klinische Bioinformatik	Vom Bluttest zur Genom-Sequenzierung – Die Bedeutung der Informatik für das Gesundheitswesen (Vortrag im Festsaal des Saarbrücker Rathauses)
04.12.2013	Prof. Dr. Andreas Keller	Universität des Saarlandes, Lehrstuhl für klinische Bioinformatik	Bioinformatics on the Way to Clinical Routine (Antrittsvorlesung)
14.08.2013	Dr. Tobias Marschall	Centrum Wiskunde & Informatica, Amsterdam, Niederlande	Next-Generation Tools for Next-Generation Sequencing Data
14.08.2013	Dr. Benjamin Georgi	University of Pennsylvania, Department of Genetics, Philadelphia, USA	Complex Genetic Architecture of Bipolar Disorder in a Genetic Isolate
14.08.2013	Dr. André Altmann	Stanford University, School of Medicine, USA	Approaching Neuropsychiatric Disorders with Neuroimaging and Molecular Genetics
16.11.2012	Prof. Dr. Martin Vingron	Max-Planck-Institut für Molekulare Genetik, Berlin	Computational Regulatory Genomics (Vortrag anlässlich des 60. Geburtstages von Prof. Dr. Dr. Thomas Lengauer)
16.11.2012	Prof. Dr. Burkhard Rost	Technische Universität München, Institut für Informatik	Beaming from Observed SNPs to Individual Health (Vortrag anlässlich des 60. Geburtstages von Prof. Dr. Dr. Thomas Lengauer)
16.11.2012	Dr. Eugene Myers	MPI für Molekulare Zellbiologie und Genetik, Dresden	Progressive Merging Strategies for Bioimage Analysis (Vortrag anlässlich des 60. Geburtstages von Prof. Dr. Dr. Thomas Lengauer)
16.11.2012	Prof. Dr. Matthias Rarey	Universität Hamburg, Zentrum für Bioinformatik	Cheminformatics – Computer Science for the Chemical Bench (Vortrag anlässlich des 60. Geburtstages von Prof. Dr. Dr. Thomas Lengauer)
16.11.2012	Dr. Christoph Bock	Forschungszentrum für Molekulare Medizin der Österreichischen Akademie der Wissenschaften, Wien, Österreich	Remembering the Past, Predicting the Future – How the Epigenome Defines Cellular Identity (Vortrag anlässlich des 60. Geburtstages von Prof. Dr. Dr. Thomas Lengauer)
16.11.2012	Prof. Dr. Niko Beerenwinkel	Eidgenössische Technische Hochschule (ETH) Zürich, Abteilung Biosysteme, Zürich, Schweiz	Evolution towards Disease (Vortrag anlässlich des 60. Geburtstages von Prof. Dr. Dr. Thomas Lengauer)
30.05.2012	Prof. Dr. Manfred Schmitt	Universität des Saarlandes, Molekular- und Zellbiologie, Zentrum für Human- und Molekularbiologie (ZHMB)	Microbial A/B Toxins – Strategies and Mechanisms of Cell Entry & Retrograde Transport (Vortrag anlässlich der Aufnahme ins ZBI)
23.03.2012	Dr. Aurelie Tomczak	Molecular Signaling Section, Laboratory of Molecular Immunology, NIAID, National Institutes of Health, Bethesda, MD, USA	Chemokines by a Computational 3D Profile-based Approach
20.02.2012	Dr. Simon Anders	EMBL, Heidelberg	Design and Analysis of Quantitative RNA-Seq Experiments
23.01.2012	Dr. Julia Engelmann	Universität Regensburg, Institut für funktionelle Genomik	Predicting Causal Effects of Stromal Cell Signals on Hepatic Cancer Cells
23.01.2012	Dr. Christof Winter	Translational Oncogenomics Lab, Lund University, Schweden	How Structural Protein Interactions and Regulatory Networks Predict Drug Targets and Survival in Pancreatic Cancer

Datum	Name	Institution / Firma	Thema
23.01.2012	Dr. Nicholas Furnham	<i>European Bioinformatics Institute, Cambridge</i>	Exploring the Evolution of Novel Protein Functions within Structurally Defined Superfamilies
23.01.2012	Dr. Mareike Fischer	<i>Center for Integrative Bioinformatics, Wien</i>	Maximum Parsimony – New Insights into an Old Method
23.01.2012	Dr. Barbara Hutter	<i>Computational Oncology-Gruppe, German Cancer Research Center, Heidelberg</i>	Analysis of New Generation Sequencing Data in Computational Oncology
23.01.2012	Dr. Dominic Rose	<i>Chair for Bioinformatics, Universität Freiburg</i>	Computational Biology of Non-Coding RNAs
07.12.2011	Prof. Dr. Dietrich Volmer	<i>Universität des Saarlandes, Institut für Bioanalytische Chemie</i>	Differential Measurements of Small Molecule Metabolites (Vortrag anlässlich der Aufnahme ins ZBI)
09.11.2011	Prof. Dr. Norbert Graf	<i>Universität des Saarlandes, Campus Homburg, Klinik für Pädiatrische Onkologie und Hämatologie</i>	On the Way to Personalized Medicine – A Clinical View (Vortrag anlässlich der Aufnahme ins ZBI)
23.03.2011	Prof. Dr. Richard Karp	<i>University of California, Berkeley, USA</i>	Effective Heuristics for NP-Hard Problems Arising in Molecular Biology
23.03.2011	Prof. Dr. Helmut Grubmüller	<i>Max-Planck-Institut für Biophysische Chemie, Göttingen</i>	Mechanoenzymatics: Atomistic Simulation of Bio-molecular Nanomachines
23.03.2011	Prof. Dr. Janet Thornton	<i>European Bioinformatics Institute, Hinxton, UK</i>	The Evolution of Enzyme Mechanisms and Functional Diversity
23.03.2011	Prof. Dr. Søren Brunak	<i>Technical University of Denmark, Lyngby, Dänemark</i>	Linking Phenotypes Extracted from Electronic Patient Records onto Systems Biology Frameworks
23.03.2011	Prof. Dr. Eugene Myers	<i>Howard Hughes Medical Institute (HHMI), Janelia Farms Research Campus, USA</i>	Image-based Informatics for Molecular Biology
25.11.2009	Herrn Dr. Hauke Busch	<i>Freiburg Institute for Advanced Studies</i>	Using Time-Scale Separation of Cellular Processes to Untangle Biological Complexity
28.10.2009	Dr. Jürgen Pahle	<i>University of Manchester</i>	How to Quantify Information Transfer in Signaling Systems
24.06.2009	Prof. Dr. Eugene Myers	<i>Howard Hughes Medical Institute (HHMI), Janelia Farms Research Campus, USA</i>	Machine-Vision and Molecular Biology
24.06.2009	Dr. Christoph Bock	<i>Max-Planck-Institut für Informatik, Saarbrücken</i>	Epigenome Analysis with Bioinformatic Methods and Applications to Cancer Biomarker Discovery
27.05.2009	Dr. Alice McHardy	<i>Max-Planck-Institut für Informatik, Saarbrücken</i>	Computational Methods for the Analysis of Metagenome Sequence Samples
06.05.2009	Dr. Mario Albrecht	<i>Universität Greifswald</i>	Utilizing Network Data for Biological Research

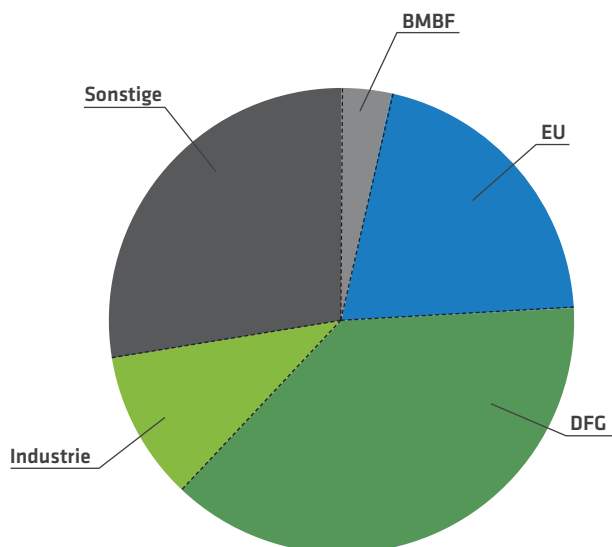
Drittmittel in Zahlen

114



Drittmittelprojekte 2010 bis 2015 // Einnahmen

■ BMBF ■ EU ■ DFG ■ Industrie ■ Sonstige



Drittmittelprojekte 2010 bis 2015 // Anzahl und Verteilung



Pia Scherer-Geiss

**Zentrum für Bioinformatik (ZBI)
Geschäftsführung**

Tel.: +49 681 302 68605

Email psg@bioinf.uni-sb.de

Patente und Erfindungen

Patente und Erfindungen der Professoren des ZBI 2010 – 2015

115

Prof. Dr. Andreas Keller

- 26 Erfindungen im Bereich „Bakterielle Resistenzen“ (alle in Anspruch genommen)
- 25 in Kooperation mit Industriepartner und an Industriepartner übertragen/verwertet oder aktuell zu übertragen
- 1 Inanspruchnahme (Verwertung steht bevor)
- 10 Erfindungen im Bereich „nicht-invasive Biomarker“
- 6 in Kooperation mit Industriepartner und an Industriepartner übertragen/verwertet (3 mit Professor Meese)
- 1 an Verwertungspartner übertragen (mit Professor Meese)
- 3 freigegebene Erfindungen (1 mit Professor Meese)

Prof. Dr. Eckart Meese

- 21 Erfindungsmeldungen im Bereich „nicht-invasive Biomarker“ (davon 20 Inanspruchnahmen)
- 19 in Kooperation mit Industriepartner und an Industriepartner übertragen (davon 3 mit Professor Keller)
- 1 an Verwertungspartner übertragen (mit Professor Keller)
- 1 freigegebene Erfindung (mit Professor Keller)

Prof. Dr. Christoph Wittmann

- 3 Patentanmeldungen im Bereich „System-Biotechnologie“ davon 1 Gemeinschaftserfindung mit dem HZI eine bereits in Braunschweig entstanden

Prof. Dr. Jörn Walter

- 3 Erfindungsmeldungen (alle in Anspruch genommen) im Bereich „Epigenetische Diagnostik“
- 1 Patentanmeldung – wurde inzwischen fallen gelassen
- 1 Inanspruchnahme
- 1 Erfindungsmeldung in Bewertung

Kooperationen

116

Ausgewählte nationale und internationale Kooperationen

Prof. Dr. Dr. Thomas Lengauer

- _ Universität Nijmegen (*Niederlande*)
- _ Institut Pasteur (*Paris, Frankreich*)
- _ Deutsches Krebsforschungszentrum (*Heidelberg, Deutschland*)
- _ Max-Planck-Institut für Molekulare Genetik (*Berlin, Deutschland*)
- _ Max-Delbrück-Zentrum (*Berlin, Deutschland*)
- _ Max-Planck-Institut für Immunbiologie und Epigenetik (*Freiburg, Deutschland*)
- _ Robert-Koch Institut (*Berlin, Deutschland*)
- _ Karolinska Institut (*Stockholm, Schweden*)
- _ Universität Siena (*Siena, Italien*)
- _ Informa s.r.l. (*Rom, Italien*)
- _ Universität zu Köln (*Deutschland*)
- _ Universitätsklinikum Düsseldorf (*Deutschland*)
- _ Medizinische Universität Hannover (*Deutschland*)
- _ BioSolveIT GmbH (*Sankt Augustin, Deutschland*)
- _ Institut für Immunologie und Genetik (*Kaiserslautern, Deutschland*)
- _ Universitätsklinikum Frankfurt (*Deutschland*)

Prof. Dr. Hans-Peter Lenhof

- _ Universität Salzburg (*Österreich*)
- _ Universität Tübingen (*Deutschland*)
- _ Universität Würzburg (*Deutschland*)
- _ Universität Leipzig (*Deutschland*)
- _ Roche Diagnostics GmbH (*Penzberg, Deutschland*)

Prof. Dr. Volkhard Helms

- _ Universität zu Köln (*Deutschland*)
- _ Konsortium Mario Negri Sud (*Santa Maria Imbaro, Italien*)
- _ Shanghai Institut für Physik (*Shanghai, China*)
- _ Pacific Northwest National Laboratory (PNNL) (*Richland, USA*)

Prof. Dr. Andreas Keller

- _ Universität Luxemburg (*Luxemburg*)
- _ Christian-Albrechts-Universität Kiel (IKMB) (*Deutschland*)
- _ Universität Würzburg (*Deutschland*)
- _ Universität Heidelberg (*Deutschland*)
- _ Deutsches Krebsforschungszentrum Heidelberg (DKFZ) (*Deutschland*)
- _ EURAC Bozen (*Südtirol, Italien*)
- _ Siemens Healthcare (*Erlangen, Deutschland*)

- _ Merck Serono (*Darmstadt, Deutschland*)
- _ Comprehensive Biomarker Center (CBC) (*Heidelberg, Deutschland*)
- _ Center for Genomics and Transcriptomics (CeGAT) (*Tübingen, Deutschland*)
- _ BioVendor Laboratory (*Karasek, Tschechien*)

Dr. Olga Kalinina

- _ Eötvös Loránd Universität (*Budapest, Ungarn*)
- _ Max-Planck-Institut für biophysikalische Chemie (*Göttingen, Deutschland*)
- _ UCLA (*Los Angeles, USA*)
- _ Universität zu Köln (*Deutschland*)
- _ Universität Duisburg-Essen (*Essen, Deutschland*)
- _ Medizinisches Infektiologiezentrum (MVZ) (*Berlin, Deutschland*)

Dr. Nico Pfeifer

- _ Microsoft Research (*Los Angeles, USA*)
- _ Rockefeller University (*New York, USA*)
- _ Universität Leuven (KU) (*Leuven, Belgien*)
- _ Universität zu Köln (*Deutschland*)
- _ Universität Duisburg-Essen (*Essen, Deutschland*)
- _ Medizinisches Infektiologiezentrum (MVZ) (*Berlin, Deutschland*)
- _ Heinrich-Pette-Institut & UKE (*Hamburg, Deutschland*)
- _ Unisversität Basel (*Basel, Schweiz*)

Juniorprofessor Dr. Tobias Marschall

- _ Centrum Wiskunde & Informatica (*Amsterdam, Niederlande*)
- _ Robert-Koch-Institut (*Berlin, Deutschland*)
- _ European Research Institute for the Biology of Ageing, Universität Groningen (*Niederlande*)
- _ Genome Institute, Washington University, St. Louis (*Missouri, USA*)
- _ Universität Pisa (*Italien*)

Dr. Marcel Schulz

- _ Freie Universität Berlin (*Deutschland*)
- _ Charité (*Berlin, Deutschland*)
- _ Carnegie Mellon University (*Pittsburgh, USA*)
- _ Roche, Basel (*Basel, Schweiz*)
- _ Universität Luxemburg (*Luxemburg*)
- _ Universität Pierre und Marie Curie (UPMC) (*Paris, Frankreich*)

Graduierungen

Seit Bestehen der Studiengänge hat das Zentrum für Bioinformatik 181 Bachelor-Absolventen und 156 Master-Absolventen hervorgebracht. Das summiert sich zu 337 Graduierungen.

Promotionen

Insgesamt 61 Promotionen wurden im ZBI erfolgreich abgeschlossen.

43 DoktorandInnen gingen nach ihrer Promotion in die Industrie und 18 in die Wissenschaft. 20 der promovierten WissenschaftlerInnen sind heute im Ausland tätig



Übergabe der Urkunde an eine Bioinformatik-Absolventin

117

Doktorandinnen und Doktoranden

<i>Name, Vorname</i>	<i>Titel der Dissertation</i>	<i>Jahr</i>
Kohlbacher, Oliver P.	New Approaches to Protein Docking	2001
Burkhardt, Stefan	Filter Algorithms for Approximate String Matching	2002
Beerenwinkel, Niko	Computational Analysis of HIV Drug Resistance Data	2004
Hildebrandt, Andreas	Biomolecules in a Structured Solvent - A Novel Formulation of Nonlocal Electrostatics and its Numerical Solution	2005
Frigato, Tomaso	Computational Study of Charged Species Solvation: Proton Transfer in Aqueous Systems and Microhydration of Neural and Anionic Thymine	2006
Albrecht, Mario	Combining Protein Structure Prediction with Experiments and Functional Information	2006
Ansari ,Sam	Computational Analysis of Protein-Protein Interactions	2007
Moll, Andreas	BALL-View - A Molecular Viewer and Modeling Tool	2007
Park, Yungki	Development of Computational Methods for Predicting Structural Characteristics of Helical Membrane Proteins	2007
Steffen, Andreas	Computational Approaches in a Supramolecular Chemistry with a Special Focus on Virtual Screening	2008
Gu, Wei	Computer Simulation of Biomolecular Solvation, Recognition and Proton Transfer Equilibria	2008
Sing, Tobias	Model-Based Anti-HIV Therapy	2008
Küntzer, Jan	BN++-Abiological Information System	2008
Hartmann, Christoph	Modeling of Flexible Side Chains for Protein-Ligand Docking	2008
Bock, Christoph	Computational and Epigenetics-Bioinformatic Methods for Epigenome Prediction, DANN Methylation Mapping and Cancer Epigenetics	2008
Talwar, Priti	Development of Computational Methodes for Metabolic Network Ananalysis Based on Metabolomics Data	2008
Maydt, Jochen	Analysis of Recombination in Molecular Sequence Data	2008

Doktorandinnen und Doktoranden

<i>Name, Vorname</i>	<i>Titel der Dissertation</i>	<i>Jahr</i>
Sander, Oliver	Structural Descriptors for the Analysis of Protein Structure, Function and Evolution	2008
Keller, Andreas	Bioinformatikmethode zur Diagnose und Prognose von Krebserkrankungen	2009
Fuhrmann, Jan	Gradient Based Optimazation in Ligand-Receptor Docking	2009
Eyrisch, Susanne	Computational Design and Analysis of Binding Pockets at Protein-Protein Interaction Interfaces	2009
Hutter, Barbara	Bioinformatics Analyses of Genomic Imprinting	2009
Möhl, Mathias	Dynamic Programming Based RNA Pseudoknot Alignment	2010
Rurainski, Alexander	Optimization in Bioinformatics	2010
Altmann, André	Bioinformatical Approaches to Ranking of Anti-HIV Combination Therapies and Planning of Treatment Schedules	2010
Zhu, Hongbo	Characterization, Classification and Alignment of Protein-Protein-Interfaces	2010
Backes, Christina	Bioinformatics Approaches for Cancer Research	2010
Schlicker, Andreas	Ontology-Based Similarity Measures and their Application in Bioinformatics	2010
Hayat, Sikander	Sequence Based Methods for the Prediction and Analysis of the Structural Topology of Transmembrane Beta Barrel Proteins	2010
Hussong, René	Efficient and Accurate Processing of Biological Mass Spectrometry Data	2010
Jusoh, Siti Azma	Molecular Modeling of the Transmembrane Domain of Envelope Glycoproteins from Flaviviridae Viruses	2010
Emig, Dorothea	Novel Analysis Approaches to Context-Dependent Molecular Networks	2011
Kunert, Lars	Maximal Common Subgraph DAGs: Theory and Application to Virtual Screening in Drug Development	2011
Weggler, Sophie	Correlation Induced Electrostatics Effects in Biomolecular Systems	2011
Thielen, Alexander	Genotypic Analysis of HIV-1 Coreceptor Usage	2011
Rump, Kirsten	Computational Immunology: Analyses of Viral Escape, Epitope Binding and T Cell Receptor Recognition	2011
Bozek, Kasia	Analysis of HIV-Host Interaction on Different Scales	2011
Bogojeska, Jasmina	Statistical Learning Methods for Bias-aware HIV Therapy Screening	2011
Walter, Peter	Storing and Analysing Biomolecular Contacts	2012
Dehof, Anna Katharina	Novel Approaches for Bond Assignment and NMR Shift Prediction	2012
Ramirez, Fidel	Novel Approaches to the Integration and Analysis of Systems Biology Data	2012
Tolosi, Laura Maria	Finding Regions of Abberant DNACopy Number Associated with Tumor Phenotype	2012

<i>Name, Vorname</i>	<i>Titel der Dissertation</i>	<i>Jahr</i>
Ahmad, Mazen	Mechanisms of Protein-Protein Association	2012
Patil, Kaustubh	Genome Signature-Based Sequence Comparison for Taxonomic Assignment and Tree Inference	2013
Schaadt, Nadine	Computational Systems Biology Methods for Functional Classification of Membrane Proteins and Modeling of Quorum Sensing in <i>Pseudomonas Aeruginosa</i>	2013
Feuerbach, Lars	Evolutionary Epigenomics - Identifying Functional Genome Elements by Epigenetic Footprints in the DNA	2014
Schelhorn, Sven-Eric	Going Viral - an Integrated View on Virological Data Analysis from Basic Research to Clinical Applications	2014
Blankenberg, Hagen	Computational Methods for Integrating and Analyzing Human Systems Biology Data	2014
Röttger, Richard	Active Transitivity Clustering of Large-Scale Biomedical Datasets	2014
Assenov, Yassen	Identification and Prioritization of Genomic Loci with Disease-Specific DANN Methylation	2014
Halachev, Konstantin	Exploratory Visualizations and Statistical Analysis of Large, Heterogeneous Epigenetic Datasets	2014
Beggel, Bastian	Determining and Utilizing the Quasispecies of the Hepatitis B Virus in Clinical Applications	2014
Lee, Po-Hsien	Computational Approaches to Investigate Structural and Functional Properties of Transmembrane Proteins	2014
Dietzen, Matthias	Modeling Protein Interactions in Protein Binding Sites and Oligomeric Protein Complexes	2014
Spieler, David	Numerical Analysis of Long-Run Properties for Markov Population Models	2014
Shanak, Siba	Dynamics of Epigenetic Reader Proteins and their Interplay with Expression in Development	2015
Ibragimov, Rashid	Exact and Heuristic Algorithms for Network Alignment Using Graph Edit Distance Models	2015
Ulucan, Özlem	Molecular Modeling of Biomolecular Associations and Quantifying Allosteric Effects	2015
Hamed, Mohamed	Integrative Computational Approaches for Studying Stem Cell Differentiation and Complex Diseases	2015
Spaniol, Christian	Integrative Analysis and Application Pipelines for Complex Disease Data	2015
Barghash, Ahmad	Identifying Biological Associations from High-Throughput Datasets	2015

Habilitationen

Obwohl die Habilitation heute von der Politik nicht mehr durchgängig unterstützt wird, ist sie in manchen wissenschaftlichen Disziplinen weiter erwünscht. Das trifft teilweise auch auf die avisierten Forschungsbereiche unserer BioinformatikerInnen zu. Daher wurden am ZBI auch eine Reihe Habilitationen durchgeführt.

<i>Name, Vorname</i>	<i>Fakultät</i>	<i>Titel der Habilitation</i>	<i>Jahr</i>
Rahmenführer, Jörg	6	Statistical Methods for the Biological Interpretation of Genome-Wide Measurements	2006
Hutter, Michael	8	Vorhersage von Moleküleigenschaften mit computergestützten Methoden	2007
Antes, Iris	6	Computational Methods for the Investigation of Protein-Ligand Complexes	2008
Sommer, Ingolf	6	From Protein Structure to Function	2010
Keller, Andreas	2	Multi-Parameter Genomics and Transcriptomics Diagnostics in Human Genetics	2012
Geyer, Tihamér	8	Coarse-Grained Modeling of Biological Systems	2013

Rufe auf Professuren

Aus dem ZBI sind sechzehn Professuren hervorgegangen, wie die Tabelle zeigt. Neun Absolventen haben am ZBI promoviert (und vier auch habilitiert) und danach einen Ruf erhalten.

Fünf davon waren als Postdocs am ZBI und haben danach Rufe angenommen.

121

DoktorandInnen des ZBI, die Rufe erhalten haben

Name, Vorname	Graduierungen (UdS)	Erhaltener Titel	Affiliation
Albrecht, Mario	Promotion (Bioinformatik): 2006	Professor	Technische Universität Graz, Bioinformatik
Beerenwinkel, Niko	Promotion (Bioinformatik): 2004	Professor	ETH Zürich, Department of Biosystems Science and Engineering
Hildebrandt, Andreas	Promotion (Bioinformatik): 2005	Professor	Johannes Gutenberg Universität Mainz, Scientific Computing and Bioinformatics
Keller, Andreas	Bachelor (Bioinformatik): 2005 Master (Bioinformatik): 2005 Promotion (Bioinformatik): 2009 Habilitation (Bioinformatik): 2012	Professor	Universität des Saarlandes, ZBI, Klinische Bioinformatik
Kohlbacher, Oliver	Diplom (Physikalische Chemie): 1996 Promotion (Bioinformatik): 2001	Professor	Universität Tübingen, Applied Bioinformatics Group
Park, Yungki	Master (Bioinformatik): 2006 Promotion (Bioinformatik): 2007	Assistant Professor	University at Buffalo New York, Department of Biochemistry
Röttger, Richard	Promotion (Universität des Saarlandes): 2014, verlässt ZBI 2014	Assistant Professor	University of Southern Denmark, Computational Biology Group
Shanak, Siba	Master (Bioinformatik): 2009 Promotion (Bioinformatik): 2015	Assistant Professor	Arabisch-Amerikanische Universität Jenin, Biology and Biotechnology Department
Talwar, Priti	Promotion (Bioinformatik): 2008	Associate Professor	VIT University Vellore

Postdocs / Nachwuchsgruppenleiter des ZBI, die Rufe erhalten haben

Name, Vorname	Graduierungen	Erhaltener Titel	Affiliation
Antes, Iris	Promotion (Universität Zürich): 2001 Habilitation (Bioinformatik): 2009 verlässt ZBI 2008	Professor	Technische Universität München, Theoretical Chemical Biology and Protein Modelling Group
Baumbach, Jan	Promotion (Universität Bielefeld): 2008 verlässt ZBI 2012	Associate Professor	University of Southern Denmark, Computational Biology Group
Böckmann, Rainer	Promotion (Universität Göttingen): 2002 verlässt ZBI 2009	Professor	Friedrich-Alexander-Universität Erlangen-Nürnberg, Lehrstuhl Biotechnik (Computational Biology)
Hellmuth, Marc	Promotion (Universität Leipzig): 2010 verlässt ZBI 2015	Juniorprofessor	Ernst Moritz Arndt Universität Greifswald
McHardy, Alice	Promotion (Universität Bielefeld): 2004 verlässt ZBI 2011	Professor	Technische Universität Braunschweig, Helmholtz-Zentrum für Infektionsforschung, Leiterin der Abt. Bioinformatik der Infektionsforschung
Rahmenführer, Jörg	Promotion (Heinrich-Heine-Universität Düsseldorf): 1999 Habilitation (Bioinformatik): 2006 verlässt ZBI 2007	Professor	Technische Universität Dortmund, Statistische Methoden in der Genetik und Chemometrie
Saigo, Hiroto	Promotion (Universität Kyoto, Japan): 2006 verlässt ZBI 2010	Associate Professor	Department of Bioscience and Bioinformatics, Kyushu Institute of Technology, Japan

5 Lehre

5 Lehre



123

Studienprogramm	124
Ziele und Inhalte der Bioinformatik-Studiengänge	125
Berufsfelder für Bioinformatiker	126
Informationen für Studierende und Bewerber	127

Studienprogramm

124

Die Bioinformatik ist ein interdisziplinäres und noch junges Studienfach. Aus der modernen Biologie und der molekularen Medizin ist sie nicht mehr wegzudenken. Bioinformatiker entwickeln Rechenverfahren und Software, um medizinische und biochemische Prozesse zu simulieren und Daten aus Laboren oder Krankenhäusern zu analysieren. So werden mit Hilfe der Bioinformatik z. B. neue Medikamente entwickelt, individuelle Therapien für einzelne Patienten erstellt oder Tumorerkrankungen erforscht, damit sie früher erkannt und behandelt werden können.

Die Universität des Saarlandes bietet seit dem Wintersemester 2001/2002 die Möglichkeit eines Bioinformatik-Studiums an, das durch das Zentrum für Bioinformatik organisiert wird. Hierbei arbeiten die Fakultäten 2, 6 und 8 der UdS mit dem Max-Planck-Institut für Informatik, dem Deutschen Forschungszentrum für Künstliche Intelligenz und dem Fraunhofer-Institut für Biomedizinische Technik in St. Ingbert zusammen. Es gibt einen eigenen Prüfungsausschuss für die Bioinformatik-Studiengänge und einen Studiendekan des Zentrums für Bioinformatik.

Der sechs-semestrige Bachelorstudiengang und der vier-semestrige forschungsorientierte Masterstudiengang Bioinformatik wurden von Anfang an gemäß den Vorgaben des Bologna-Prozesses als konsekutive Studiengänge mit studienbegleitenden Prüfungen, einem Leistungspunktesystem auf der Basis des ECTS Systems und einer durchgehenden Modularisierung konzipiert. Im Schnitt haben in den letzten Jahren etwa 40 Studierende das Bachelorstudium begonnen. Der Zulassung zum Masterstudiengang geht ein Auswahlverfahren voraus. Hier nahmen in den letzten Jahren im Schnitt 40 Studierende pro Jahr ihr Studium auf. Der Bachelorstudiengang richtet sich vorwiegend an inländische Studierende. Daher ist die Unterrichtssprache des Bachelorstudiums überwiegend Deutsch. Der Masterstudiengang ist deutlich internationaler ausgerichtet. Hier ist die Unterrichtssprache Englisch.

Bioinformatiker haben beste Berufsaussichten und sind begehrte Experten an Universitäten, Forschungseinrichtungen oder in Pharmaunternehmen. Wer sein Studium fortsetzen möchte, kann nach dem Bachelorstudium oder dem Masterstudium eine Promotion in Bioinformatik anschließen.

Besonders talentierte Bioinformatikstudierende können direkt nach dem Bachelor in der Graduiertenschule mit ihrem Promotionsstudium beginnen. Studierende aus aller Welt werden hier in einer Einführungsphase an das wissenschaftliche Arbeiten herangeführt. Die Studierenden werden dabei in drei Semestern mit Vorlesungen und Seminaren auf die wissenschaftliche Arbeit vorbereitet. In dieser Vorbereitungsphase werden Studierende im Rahmen der Graduiertenschule durch Stipendien oder Hilfskraftverträge unterstützt. Sie müssen danach eine Qualifizierungsprüfung bestehen, der dann die Forschungs- und Dissertationsphase folgt. Alternativ werden auch wie bisher Studierende nach dem Master-Abschluss aufgenommen, die ebenfalls individuell auf ihre Forschungsphase vorbereitet werden und eine Qualifizierungsprüfung ablegen müssen.

Die Promotionsphase wird über Stipendien oder Mitarbeiterstellen finanziert. Für die Graduiertenschule gibt es ein kompetitives Zulassungsverfahren. Alle Studierenden werden intensiv betreut und ihre Fortschritte werden jedes Semester gemeinsam von allen Professoren bewertet. Daneben werden ihnen sogenannte Soft Skills vermittelt, etwa zum Thema „Wissenschaftliches Schreiben und Präsentieren“. Durch gemeinsame Büroräume in der Vorbereitungsphase soll der Teamgeist unter den Studierenden darüber hinaus gefördert werden.

Neben der Naturwissenschaftlich-Technischen Fakultät I beteiligen sich das Max-Planck-Institut für Informatik, das Max-Planck-Institut für Softwaresysteme sowie das Deutsche Forschungszentrum für Künstliche Intelligenz (DFKI) an Forschung und Lehre im Rahmen der Graduiertenschule. Die Bioinformatik ist eine der Säulen der Graduiertenschule.

Generell gelten Absolventen der Bioinformatik, die in Saarbrücken studiert haben, als sehr gut ausgebildet und haben daher auch gute Chancen auf dem Arbeitsmarkt.

Ziele und Inhalte der Bioinformatik-Studiengänge

Das generelle Ziel beider Studiengänge ist ein wissenschaftlich fundiertes, grundlagenorientiertes Studium. Auf der Grundlage einer breiten und in ausgewählten Teilgebieten vertieften fachlichen Basis wird den Studierenden die Fähigkeit zu einer grundlagen- oder anwendungsorientierten Forschung auf dem Gebiet der Bioinformatik vermittelt. Dazu erlernen sie analytische, kreative und konstruktive Fähigkeiten zur Neu- und Weiterentwicklung von bioinformatischer Software und deren Anwendung in komplexen Fragestellungen. Zu den Ausbildungszielen gehören darüber hinaus allgemeines Anwendungswissen (insbesondere vernetztes Denken), Sozialkompetenz (Teamfähigkeit, Führungs- und Kommunikationskompetenz) und der Erwerb von Schlüsselqualifikationen.

Der Bachelorstudiengang vermittelt ein breites Spektrum an Fachwissen und die für den Einstieg in die berufliche Praxis notwendigen Grundlagen. Insbesondere erlernen die Absolventinnen und Absolventen, wissenschaftliche Erkenntnisse und Problemlösungskonzepte in den Anwendungsfeldern einzusetzen. Die Ausbildung im Bachelorstudiengang vermittelt den Studierenden eine wissenschaftliche Grundqualifizierung sowie grundlegende Fachkenntnisse und Fertigkeiten. Sie vermittelt auch die Fähigkeit zur Erschließung neuer Gebiete und zur selbständigen Weiterbildung.

Neben der (Kern-)Bioinformatik umfasst das Fächerangebot im Bachelorstudiengang Mathematik, Informatik, Chemie, Biologie und Pharmazie. Die Studierenden können bei der Zusammenstellung ihres individuellen Studienprogramms gemäß einer eher informatischen und einer eher biologischen Vertiefung eine Gewichtung vornehmen. Durch die Modularisierung des Studiums können die Studierenden gemäß den beiden Vertiefungen ihr Studium nach ihren Fähigkeiten und Interessen ausgestalten. Inhalte des Bioinformatik-Studiums sind insbesondere die theoretischen Grundlagen und Methoden des Faches Bioinformatik. Als Grundlagen werden dafür Grundkenntnisse und Methoden aus den benachbarten Disziplinen Mathematik und Informatik sowie die theoretischen Grundlagen ausgewählter Bereiche der Lebenswissenschaften (allgemeine, organische und physikalische Chemie, Biochemie, Molekular- und Mikrobiologie, Pharmazie und medizinische Chemie sowie Biophysik) vermittelt.

Ausbildungsziel im Kernbereich Bioinformatik: Im dritten bis sechsten Semester des Bachelorstudiengangs wird eine breit angelegte, methodische Grundausbildung in den Gebieten algorithmische Sequenzanalyse, algorithmische Strukturbiologie, Analyse integrierter biologischer Netzwerke sowie in computergestützter Chemie angeboten. Der Vorlesungsstoff wird jeweils in theoretischen und programmiertechnischen Übungen vertieft.

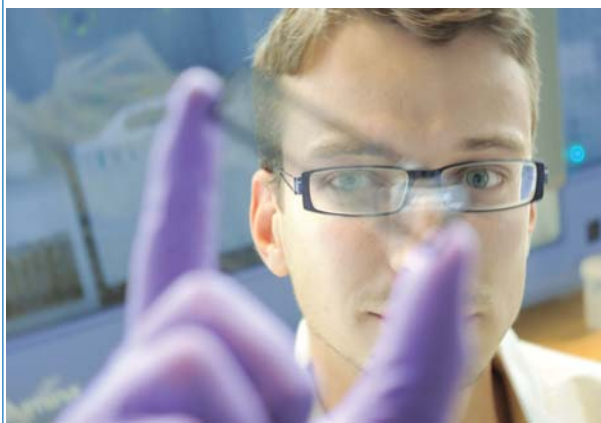
Ein wichtiges Element im Bachelorstudiengang ist der Einsatz des theoretisch Gelernten in Praktika. Im biowissenschaftlichen Bereich führen die Studierenden im Allgemeinen nach dem dritten und nach dem fünften Semester zweiwöchige ganztägige Grund- und Aufbaupraktika durch. Im Informatik-Bereich wird in einem achtwöchigen Softwarepraktikum in Teamarbeit ein größeres Softwaredesignprojekt umgesetzt. Als vorgeschriebene größere Projekte sollten neben dem angeführten Softwarepraktikum vor allem auch die Bachelor- und Masterabschlussarbeiten angesehen werden. Eine externe Durchführung in Firmen ist möglich. Wir ermutigen unsere Studierende, ein achtwöchiges Industrie- oder Auslandspraktikum bei einem/einer einschlägig im Bereich der Bioinformatik arbeitenden Unternehmen oder Arbeitsgruppe durchzuführen. Die Betreuung erfolgt in diesem Fall in den betreffenden Unternehmen bzw. Arbeitsgruppen. Die Studierenden sind selbst für die Organisation des Praktikums verantwortlich (Herstellung des Kontaktes, Organisation der Unterbringung vor Ort etc.). Sie erwerben dadurch Selbständigkeit und Selbstvertrauen.

Der Masterstudiengang ist forschungsorientiert und vermittelt, ergänzend zum vorhergehenden Bachelorstudiengang, tiefergehendes Fachwissen. Die Studierenden werden befähigt, wissenschaftliche Methoden und Erkenntnisse bei komplexen Problemstellungen in der Praxis einzusetzen. Ziel des Masterstudiengangs ist es, auf eine anspruchsvolle Forschungs- und Entwicklungstätigkeit im Bereich der Bioinformatik vorzubereiten, bzw. den Studierenden zu ermöglichen, ihre Laufbahn im akademischen Sektor der Bioinformatik fortzusetzen. Dazu sollen insbesondere die methodischen Fähigkeiten der Studierenden gefördert werden, um sie in die Lage zu versetzen, innovative Methoden zu entwickeln.

Berufsfelder für Bioinformatiker

126

Der Bachelorstudiengang soll die Absolventinnen und Absolventen auf ihre berufliche Praxis im Bereich der Bioinformatik vorbereiten. Sie müssen die biowissenschaftlichen Probleme und Fragestellungen verstehen können, sie mathematisch modellieren sowie wissenschaftliche Methoden und Erkenntnisse der Bioinformatik auf diese Probleme anwenden können. Ein Teil der Absolventinnen und Absolventen des Bachelorstudiengangs wird im späteren Berufsalltag existieren-



Forscher Martin Kircher (ehemals ZBI) beim Prüfen einer Illumina GAIi flow cell. © Frank Vinken

de Bioinformatik-Software und Tools anwenden lernen, um neues biologisches Wissen zu gewinnen (zum Beispiel: Identifizierung neuer Zielmoleküle für die Therapie von bestimmten Krankheiten, Identifizierung neuer Leitstrukturen für die Entwicklung von Medikamenten, Optimierung von Therapien usw.). Ein anderer Teil der Absolventinnen und Absolventen wird neue Bioinformatik-Methoden und -software entwickeln und implementieren. Wir unterscheiden hier also zwischen dem Berufsbild des „Bioinformatik-Anwenders“ und dem Berufsbild des „Bioinformatik-Entwicklers“. Der Anwender benötigt für den Berufsalltag ein umfangreiches Wissen im Bereich der Lebenswissenschaften und der Bioinformatik-Tools, wohingegen die Entwicklung neuer Bioinformatik-Software tiefere Kenntnisse der Informatik-Methoden und der Mathematik erfordert.

Um den Studierenden eine zu diesen Berufsbildern passende berufsqualifizierende Ausbildung zu garantieren und ihnen mehr Wahlmöglichkeiten innerhalb des Bachelorstudiengangs einzuräumen, erlaubt der Studiengang den Studenten während der zweiten Hälfte ihres Bachelorstudiums entweder mehr Leistungspunkte aus den biologischen Kategorien oder aus den informatischen Kategorien einzubringen. Die erste Profilbildung entspricht dem Berufsbild des „Bioinformatik-Anwenders“, während die zweite Profilbildung dem Berufsbild des „Bioinformatik-Entwicklers“ entspricht. Das letztere mehr theorieorientierte Profil bereitet auch auf eine akademische Laufbahn im Bereich Bioinformatik sowie auf Leitungsfunktionen in der Industrie vor und ist deshalb für die Zulassung zum Masterstudiengang Bioinformatik erforderlich. Bachelor-Absolventinnen und -Absolventen, die zunächst mehr Gewicht auf die biologische Seite der Bioinformatik legen möchten und dennoch später einen Master-Abschluss erwerben möchten, empfehlen wir, zum Beispiel den Masterstudiengang „Biotechnologie“ oder „Drug Design“ zu wählen. Durch diese zwei unterschiedlichen Profile wird nicht nur die zentrale Forderung berücksichtigt, dass die Ausbildung in Bachelorstudiengängen berufsqualifizierend sein soll, sondern es werden den Studierenden auch mehr Wahlmöglichkeiten eingeräumt und es wird ihnen ermöglicht, sich entsprechend ihrer Interessen (Lebenswissenschaften oder IT) und Fähigkeiten zu spezialisieren.



Bioinformatiker bei der Analyse von Expressionsdaten.

Informationen für Studierende und Bewerber

An der Lehre in den sehr interdisziplinären Bioinformatik-Studiengängen sind Lehrende dreier Fakultäten und mehrerer An-Institute beteiligt. Für die eigentliche Bioinformatik-Lehre sind zuständig der Lehrstuhl Professor Helms (*Biologie*), der Lehrstuhl Professor Lenhof (*Informatik*), der Lehrstuhl Professor Keller (*Medizin*), Juniorprofessor Marschall (*Informatik*) sowie die Abteilung von Professor Lengauer am Max-Planck-Institut für Informatik. An der Lehre im Kernbereich Bioinformatik sind derzeit weiterhin beteiligt Professor Verena Wolf (*Informatik*) und Dr. Marcel Schulz (*Informatik*).

Wer Bioinformatik studieren will, sollte ein besonderes Interesse für Mathematik, die Naturwissenschaften – Biologie im Besonderen – und Informatik mitbringen. Das sechsemestrige Bachelorstudium vermittelt wissenschaftliche Grundlagen aus Biowissenschaften und Informatik. Aber auch Fächer wie Mathematik oder Chemie stehen auf dem Stundenplan. Zudem sind grundlegende Englischkenntnisse wichtig, da Fachliteratur meist nur in englischer Sprache angeboten wird und auch fast alle Lehrveranstaltungen des anschließenden Masterstudienganges in englischer Sprache stattfinden.

Für Studienanfänger zum Wintersemester bietet die Fachrichtung Informatik einen kostenlosen vierwöchigen Auffrischkurs Mathematik an, um Lücken zu schließen und eine gemeinsame Grundlage zu Studienbeginn zu schaffen.

Jeweils in der Woche vor Vorlesungsbeginn veranstaltet die Fachschaft Informatik und Bioinformatik in Zusammenarbeit mit der Studienkoordinatorin der Informatik Einführungsveranstaltungen, in denen nicht nur der Studienablauf und die Prüfungsordnungen erklärt werden, sondern auch Gäste anderer Einrichtungen der Universität ihre Arbeit vorstellen.

Das ZBI verfügt über einen eigenen Rechner-Pool für die Studierenden der Bioinformatik, der mit leistungsfähigen PCs und Flachbildschirmen ausgestattet ist. Der CIP-Pool ist rund um die Uhr für die Studierenden per Magnetkarte zugänglich.

Studierende, die herausragende Studienleistungen erbringen und die ihr Studium schneller als in der Regelzeit abschließen wollen, können in ein Förderprogramm aufgenommen werden. Dort werden sie von einem Mentor unterstützt.

Die Studierenden werden regelmäßig und flächendeckend mittels Fragebogen zu ihrer Einschätzung der Lehrveranstaltungen befragt. Die Befragungen finden in der Regel zur Mitte des Semesters statt. Die anonymisierten Fragebögen werden zentral auf automatisiertem Weg ausgewertet. Die Lehrenden erhalten gegen Ende des Semesters eine detaillierte Rückmeldung über die Ergebnisse der Evaluation und nutzen diese zur Optimierung ihrer Lehrveranstaltungen.

Die folgenden Beispielstundenpläne enthalten nähere Angaben für einen zweckmäßigen Aufbau des Studiums für beide Bachelor-Vertiefungsrichtungen sowie für den Masterstudiengang.

Beispielstundenplan für den Bachelor (BI)

Sem	Informatik	Mathematik	Bioinformatik	Grundvorlesungen der Chemie und Biowissenschaften
1	Programmierung I (9 CPs)	Mathematik für Informatiker I (9 CPs)	Ringvorlesung: Einführung in die Bioinformatik (3 CPs)	Allgemeine Chemie (4 CPs) Organische Chemie und Biochemie (5 CPs)
2	Programmierung II (9 CPs)	Mathematik für Informatiker II (9 CPs)		Molekularbiologie (3 CPs)
3	Grundzüge von Datenstrukturen und Algorithmen (6 CPs)		Bioinformatik I (9 CPs)	
4			Bioinformatik II (9 CPs)	Physikalische Chemie (4 CPs)
5			Spezial-Vorlesung Bioinformatik (5 CPs)	
6			Spezial-Vorlesung Bioinformatik (5 CPs) Computational Chemistry (6 CPs) Bachelorarbeit (12 CPs)	

Sem	Vorlesungen der Biowissenschaften	Seminare	Praktika	Schlüsselqualifikationen	ECTS
1				Effizientes Lernen (1 CP)	31 CPs
2			Software-Praktikum (9 CPs)		30 CPs
3	Einführung in die Biotechnologie (3 CPs)		Softwarewerkzeuge (9 CPs) Grund-Praktikum Biowissenschaften (4 CPs)		31 CPs
4	Molekulare Mikrobiologie (3 CPs)	Proseminar (5 CPs)	Tutortätigkeit (4 CPs)	Projekt-Management (1 CP)	30 CPs
5	Einführung in die Zellbiologie (5 CPs) Grundlagen der Genetik (6 CPs) Medizinische Chemie und Drug Design (5 CPs) Biopharmazie und Drug Delivery (5 CPs)		Aufbaupraktikum Biowissenschaften (4 CPs)		30 CPs
6		Bachelor-Seminar (9 CPs)			32 CPs

Beispielstundenplan für den Bachelor (CMB)

Sem	Informatik	Mathematik	Bioinformatik	Grundvorlesungen der Chemie und Biowissenschaften
1	Programmierung I (9 CPs)	Mathematik für Informatiker I (9 CPs)	Ringvorlesung: Einführung in die Bioinformatik (3 CPs)	Allgemeine Chemie (4 CPs) Organische Chemie und Biochemie (5 CPs)
2	Programmierung II (9 CPs)	Mathematik für Informatiker II (9 CPs)		Molekularbiologie (3 CPs)
3	Grundzüge von Datenstrukturen und Algorithmen (6 CPs)	Mathematik für Informatiker III (9 CPs)	Bioinformatik I (9 CPs)	
4			Bioinformatik II (9 CPs) Spezial-Vorlesung Bioinformatik (5 CPs)	Physikalische Chemie (4 CPs)
5			Spezial-Vorlesung Bioinformatik (5 CPs)	
6			Computational Chemistry (6 CPs) Bachelorarbeit (12 CPs)	

Sem	Vorlesungen der Biowissenschaften	Seminare	Praktika	Schlüsselqualifikationen	ECTS
1				Effizientes Lernen (1 CP)	31 CPs
2			Software-Praktikum (9 CPs)		30 CPs
3	Einführung in die Biotechnologie (3 CPs)		Grund-Praktikum Biowissenschaften (4 CPs)		31 CPs
4	Molekulare Mikrobiologie (3 CPs)	Proseminar (5 CPs)	Tutortätigkeit (4 CPs)		30 CPs
5	Einführung in die Zellbiologie (5 CPs) Grundlagen der Genetik (6 CPs) Medizinische Chemie und Drug Design (5 CPs) Biopharmazie und Drug Delivery (5 CPs)		Aufbaupraktikum Biowissenschaften (4 CPs)		30 CPs
6		Bachelor-Seminar (9 CPs)		Organisation wiss. Forschung (1 CP)	28 CPs

Beispielstundenplan für den Master

Sem	Stammvorlesungen Informatik	Bioinformatik	Fortgeschrittenen- Vorlesungen Biowissenschaften	Seminare	Praktika	Schlüssel- qualifikati- onen	ECTS
1	Data Structures and Algorithms (9 CPs)	Bioinformatik III (9 CPs)	Spezial-Vorlesung Biowissenschaften (5 CPs)	Seminar (Bioinformatik) (7 CPs)		Wiss. Publizieren (1 CPs)	31 CPs
2	Artificial Intelligence (9 CPs)	Spezial-Vorlesung Bioinformatik (5 CPs) Spezial-Vorlesung Bioinformatik (5 CPs)	Molekulare Biotechnologie II (3 CPs) Das menschliche Genom (3 CPs) Systems Toxicology (3 CPs)			Organisation wiss. Forschung (1 CPs)	29 CPs
3		Spezial-Vorlesung Bioinformatik (5 CPs) Spezial-Vorlesung Bioinformatik (5 CPs)		Masterseminar (12 CPs)	Fortge- schrittenen- Praktikum Biowissen- schaften (8 CPs)		30 CPs
4	Masterarbeit (30 CPs)						30 CPs

Bewerbung und Zulassungsvoraussetzungen

Beide Studiengänge unterliegen Zulassungsbeschränkungen. Notwendige Voraussetzungen für die Bewerbung stehen in der Prüfungsordnung des Studienfaches Bioinformatik.

(<http://www.zbi.uni-saarland.de/de/prüfungsausschuss/studien-und-prüfungsordnungen.html>)

Zugangsberechtigt zum Masterstudiengang ist, wer das Zeugnis über die bestandene Bachelor-Prüfung im Studiengang Bioinformatik oder einem vergleichbaren Studiengang an einer deutschen Hochschule hat bzw. vergleichbare Leistungen oder das Zeugnis eines gleichwertigen Abschlusses an einer ausländischen Hochschule erworben hat bzw. sonstige gleichwertige Leistungen und ausreichende Sprachkenntnisse für das Fachstudium sowie die besondere Eignung nachweist.

Bewerbungsfrist und -unterlagen

Im Bachelor- und Masterstudiengang ist der Bewerbungsschluss der 15. Juli (für das Wintersemester) und der 15. Januar (für das Sommersemester).

Die Bewerbung zum Masterstudiengang erfolgt über ein **Online-Formular** (<http://www.zbi.uni-saarland.de/de/informationen-für-studieninteressierte/bewerbung.html>)

Die notwendigen Unterlagen (Antrag auf Zulassung, Lebenslauf, Zeugniskopien, Auflistung der besuchten Lehrveranstaltungen mit entsprechender Wochenstundenzahl, Leistungspunkten sowie Noten) können dort hochgeladen werden.

Kontakt



Prof. Dr. Volkhard Helms

Studiendekan

Telefon +49 681 302 70701
volkhard.helms@bioinformatik.uni-saarland.de



Frau Karin Jostock

Prüfungssekretariat Bioinformatik

Telefon +49 681 302 68602
jostock@bioinf.uni-sb.de



Prof. Dr. Hans-Peter Lenhof

Vorsitzender des Prüfungsausschusses

Telefon +49 681 302 64701
lenhof@bioinf.uni-sb.de



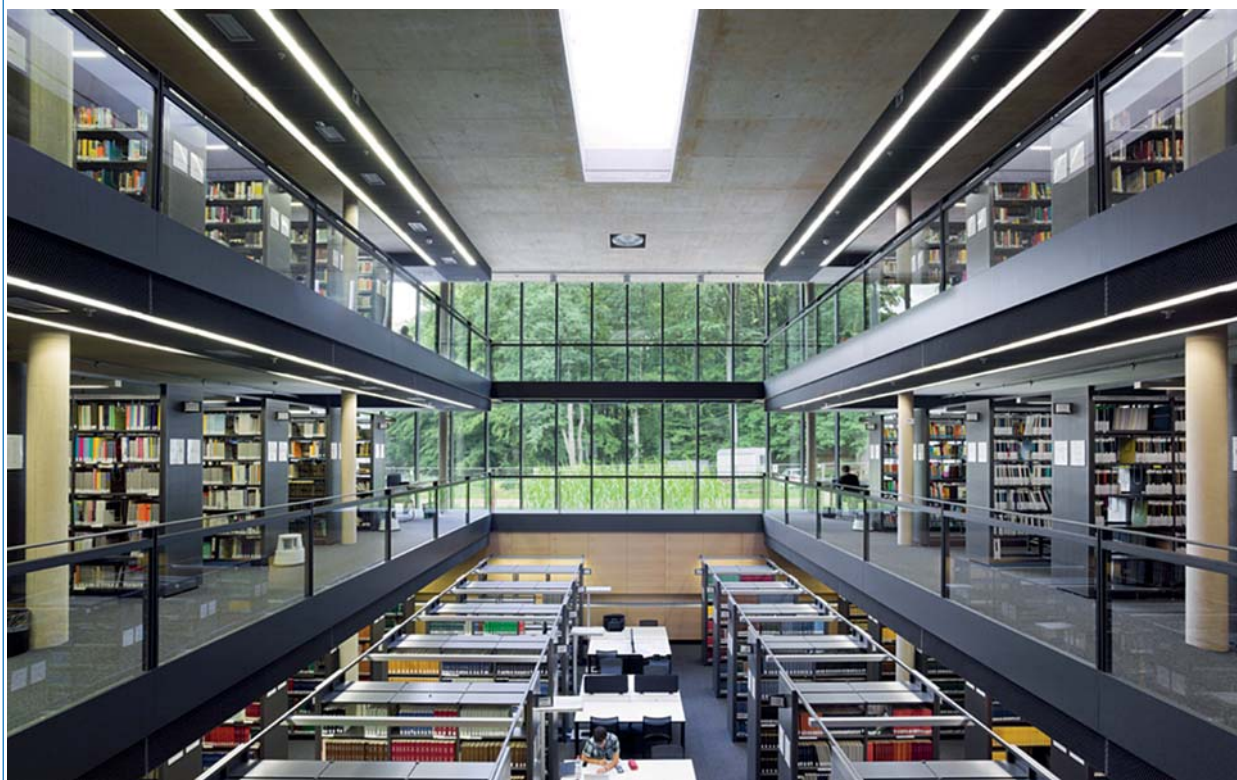
Frau Pia Scherer-Geiss

Geschäftsführerin des Zentrums für Bioinformatik

Telefon +49 681 302 68605
psg@bioinf.uni-sb.de

6 Publikationen

6 Ausgewählte Publikationen



133

P. Adler, L. J. Frey, A. Berger, C. J. Bolten, C. E. Hansen and C. Wittmann. **The key to acetate: metabolic fluxes of acetic acid bacteria under cocoa pulp fermentation-simulating conditions.** *Appl Environ Microbiol* 80 (15): 4702-16. 2014.

M. Ahmad, W. Gu, T. Geyer and V. Helms. **Adhesive water networks facilitate binding of protein interfaces.** *Nat Commun* 2: 261. 2011.

M. Ahmad, V. Helms, T. Lengauer and O. V. Kalinina. **How Molecular Conformational Changes Affect Changes in Free Energy.** *J Chem Theory Comput* 11 (7): 2945-57. 2015.

M. Ahmad, V. Helms, T. Lengauer and O. V. Kalinina. **Enthalpy-Entropy Compensation upon Molecular Conformational Changes.** *J Chem Theory Comput* 11 (4): 1410-8. 2015.

R. Akulenko and V. Helms. **DNA co-methylation analysis suggests novel functional associations between gene pairs in breast cancer samples.** *Hum Mol Genet* 22 (15): 3016-22. 2013.

G. Amabile, A. Di Ruscio, F. Muller, R. S. Welner, H. Yang, A. K. Ebralidze, H. Zhang, E. Levantini, L. Qi, G. Martinelli, T. Brummelkamp, M. M. Le Beau, M. E. Figueroa, C. Bock and D. G. Tenen. **Dissecting the role of aberrant DNA methylation in human leukaemia.** *Nat Commun* 6: 7091. 2015.

J. Arand, D. Spieler, T. Karius, M. R. Branco, D. Meilinger, A. Meissner, T. Jenuwein, G. Xu, H. Leonhardt, V. Wolf and J. Walter. **In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases.** *PLoS Genet* 8 (6): e1002750. 2012.

Y. Assenov, F. Muller, P. Lutsik, J. Walter, T. Lengauer and C. Bock. **Comprehensive analysis of DNA methylation data with RnBeads.** *Nat Methods* 11 (11): 1138-40. 2014.

C. Backes, J. Haas, P. Leidinger, K. Frese, T. Grossmann, K. Rupprecht, B. Meder, E. Meese and A. Keller. **miRFrame: analysis and visualization of miRNA sequencing data in neurological disorders.** *J Transl Med* 13: 224. 2015.

C. Backes and A. Keller. **Reanalysis of 3,707 novel human microRNA candidates.** *Proc Natl Acad Sci U S A* 112 (22): E2849-50. 2015.

C. Backes, P. Leidinger, G. Altmann, M. Wuerstle, B. Meder, V. Galata, S. C. Mueller, D. Sickert, C. Stahler, E. Meese and A. Keller. **Influence of next-generation sequencing and storage conditions on miRNA patterns generated from PAXgene blood.** *Anal Chem* 87 (17): 8910-6. 2015.

- C. Backes, E. Meese, H. P. Lenhof and A. Keller. **A dictionary on microRNAs and their putative target pathways.** *Nucleic Acids Res* 38 (13): 4476-86. 2010.
- C. Backes, A. Rurainski, G. W. Klau, O. Müller, D. Stockel, A. Gerasch, J. Kuntzer, D. Maisel, N. Ludwig, M. Hein, A. Keller, H. Burtscher, M. Kaufmann, E. Meese and H. P. Lenhof. **An integer linear programming approach for finding deregulated sub-graphs in regulatory networks.** *Nucleic Acids Res* 40 (6): e43. 2012.
- T. Barbier, F. Collard, A. Zuniga-Ripa, I. Moriyon, T. Godard, J. Becker, C. Wittmann, E. Van Schaftingen and J. J. Letesson. **Erythritol feeds the pentose phosphate pathway via three new isomerases leading to D-erythrose-4-phosphate in Brucella.** *Proc Natl Acad Sci U S A* 111 (50): 17815-20. 2014.
- T. Bastys, N. T. Doncheva, H. Walter, R. Kaiser, M. Albrecht and O. V. Kalinina. **Molecular Mechanisms of HIV-1 Protease Resistance and Resensitization towards Lopinavir and Saquinavir upon L76V Mutation.** *International Workshop on Antiviral Drug Resistance: Meeting the Global Challenge.* 2014.
- S. Baumann, J. Herrmann, R. Raju, H. Steinmetz, K. I. Mohr, S. Huttel, K. Harmrolfs, M. Stadler and R. Müller. **Cystobactamids: myxobacterial topoisomerase inhibitors exhibiting potent antibacterial activity.** *Angew Chem Int Ed Engl* 53 (52): 14605-9. 2014.
- D. Becker, P. Lutsik, P. Ebert, C. Bock, T. Lengauer and J. Walter. **BiQ Analyzer HiMod: an interactive software tool for high-throughput locus-specific analysis of 5-methylcytosine and its oxidized derivatives.** *Nucleic Acids Res* 42 (Web Server issue): W501-7. 2014.
- J. Becker, A. Lange, J. Fabarius, C. Wittmann. **Top value platform chemicals: bio-based production of organic acids.** *Curr Opin Biotechnol.* 6:168-75. 2015.
- J. Becker and C. Wittmann. **Advanced biotechnology: metabolically engineered cells for the bio-based production of chemicals and fuels, materials, and health-care products.** *Angew Chem Int Ed Engl* 54 (11): 3328-50. 2015.
- G. V. Beznoussenko, S. Parashuraman, R. Rizzo, R. Polishchuk, D. Martella, D. Di Giandomenico, A. Fusella, A. Spaar, M. Salles, M. G. Capestrano, M. Pavelka, M. R. Vos, Y. G. Rikers, V. Helms, A. A. Mironov and A. Luini. **Transport of soluble proteins through the Golgi occurs by diffusion via continuities across cisternae.** *Elife* 3. 2014.
- P. Bielecki, P. Lukat, K. Husecken, A. Dotsch, H. Steinmetz, R. W. Hartmann, R. Müller and S. Haussler. **Mutation in elongation factor G confers resistance to the antibiotic argyrimycin in the opportunistic pathogen *Pseudomonas aeruginosa*.** *Chembiochem* 13 (16): 2339-45. 2012.
- C. Bock and T. Lengauer. **Managing drug resistance in cancer: lessons from HIV therapy.** *Nat Rev Cancer* 12 (7): 494-501. 2012.
- K. Bozek, M. Eckhardt, S. Sierra, M. Anders, R. Kaiser, H. G. Krausslich, B. Muller and T. Lengauer. **An expanded model of HIV cell entry phenotype based on multi-parameter single-cell data.** *Retrovirology* 9: 60. 2012.
- K. Bozek, T. Lengauer, S. Sierra, R. Kaiser and F. S. Domingues. **Analysis of physicochemical and structural properties determining HIV-1 coreceptor usage.** *PLoS Comput Biol* 9 (3): e1002977. 2013.
- R. Bücker, A. K. Heroven, J. Becker, P. Dersch and C. Wittmann. **The pyruvate-tricarboxylic acid cycle node: a focal point of virulence control in the enteric pathogen *Yersinia pseudotuberculosis*.** *J Biol Chem* 289 (43): 30114-32. 2014.
- S. Caprari, S. Metzler, T. Lengauer and O. V. Kalinina. **Sequence and Structure Analysis of Distantly-Related Viruses Reveals Extensive Gene Transfer between Viruses and Hosts and among Viruses.** *Viruses* 7 (10): 5388-409. 2015.
- M. Cheaib, A. Dehghani Amirabad, K. J. Nordstrom, M. H. Schulz and M. Simon. **Epigenetic regulation of serotype expression antagonizes transcriptome dynamics in *Paramecium tetraurelia*.** *DNA Res* 22 (4): 293-305. 2015.
- J. Degac, U. Winter and V. Helms. **Graph-Based Clustering of Predicted Ligand-Binding Pockets on Protein Surfaces.** *J Chem Inf Model* 55 (9): 1944-52. 2015.
- M. Dietzen, O. V. Kalinina, K. Taškova, B. Kneissl, A. K. Hildebrandt, E. Jaenicke, H. Decker, T. Lengauer and A. Hildebrandt. **Large oligomeric complex structures can be computationally assembled by efficiently combining docked interfaces.** *Proteins* 83 (10): 1887-99. 2015.
- M. Dietzen, E. Zotenko, A. Hildebrandt and T. Lengauer. **On the applicability of elastic network normal modes in small-molecule docking.** *J Chem Inf Model* 52 (3): 844-56. 2012.
- M. Dietzen, E. Zotenko, A. Hildebrandt and T. Lengauer. **Correction to On the applicability of elastic network normal modes in small-molecule docking.** *J Chem Inf Model* 54 (12): 3453-3453. 2014.

- N. T. Doncheva, Y. Assenov, F. S. Domingues and M. Albrecht. **Topological analysis and interactive visualization of biological networks and protein structures.** *Nat Protoc* 7 (4): 670-85. 2012.
- J. Dudek, S. Pfeffer, P. H. Lee, M. Jung, A. Cavalie, V. Helms, F. Forster and R. Zimmermann. **Protein transport into the human endoplasmic reticulum.** *J Mol Biol* 427 (6 Pt A): 1159-75. 2015.
- P. Ebert, F. Müller, K. Nordstrom, T. Lengauer and M. H. Schulz. **A general concept for consistent documentation of computational analyses.** *Database (Oxford)* 2015: bav050. 2015.
- S. P. Eckel, J. Baumbach and A. C. Hauschild. **On the importance of statistics in breath analysis—hope or curse?** *J Breath Res* 8 (1): 012001. 2014.
- F. Erdmann, N. Schauble, S. Lang, M. Jung, A. Honigmann, M. Ahmad, J. Dudek, J. Benedix, A. Harsman, A. Kopp, V. Helms, A. Cavalie, R. Wagner and R. Zimmermann. **Interaction of calmodulin with Sec61alpha limits Ca²⁺ leakage from the endoplasmic reticulum.** *EMBO J* 30 (1): 17-31. 2011.
- S. Eyrisch, J. L. Medina-Franco and V. Helms. **Transient pockets on XIAP-BIR2: towards the characterization of putative binding sites of small-molecule XIAP inhibitors.** *Journal of Molecular Modeling*, Vol. 18, p. 2031-2042. 2012.
- A. Feldmann and N. Pfeifer. **From Predicting to analyzing HIV-1 resistance to broadly neutralizing antibodies.** *German Conference on Bioinformatics* 2015. 2015.
- A. F. Fernandez, Y. Assenov, J. I. Martin-Subero, B. Balint, R. Siebert, H. Taniguchi, H. Yamamoto, M. Hidalgo, A. C. Tan, O. Galm, I. Ferrer, M. Sanchez-Cespedes, A. Villanueva, J. Carmona, J. V. Sanchez-Mut, M. Berdasco, V. Moreno, G. Capella, D. Monk, E. Ballestar, S. Roper, R. Martinez, M. Sanchez-Carbayo, F. Prosper, X. Agirre, M. F. Fraga, O. Grana, L. Perez-Jurado, J. Mora, S. Puig, J. Prat, L. Badimon, A. A. Puca, S. J. Meltzer, T. Lengauer, J. Bridgewater, C. Bock and M. Esteller. **A DNA methylation fingerprint of 1628 human samples.** *Genome Res* 22 (2): 407-19. 2012.
- U. Fischer, C. Backes, A. Raslan, A. Keller, C. Meier and E. Meese. **Gene amplification during differentiation of mammalian neural stem cells in vitro and in vivo.** *Oncotarget* 6 (9): 7023-39. 2015.
- U. Fischer, N. Ludwig, A. Keller, C. Backes and E. Meese. **Genome-wide copy number profiling of mouse neural stem cells during differentiation.** *Genom Data* 5: 3-6. 2015.
- D. Frentz, A. M. Wensing, J. Albert, D. Paraskevis, A. B. Abe-casis, O. Hamouda, L. B. Jorgensen, C. Kucherer, D. Struck, J. C. Schmit, B. Asjo, C. Balotta, D. Beshkov, R. J. Camacho, B. Clotet, S. Coughlan, S. De Wit, A. Griskevicius, Z. Grossman, A. Horban, T. Kolupajeva, K. Korn, L. G. Kostrikis, K. Liitsola, M. Linka, C. Nielsen, D. Otelea, R. Paredes, M. Poljak, E. Puchhammer-Stockl, A. Sonnerborg, D. Stanekova, M. Stanojevic, A. M. Vandamme, C. A. Boucher, D. A. Van de Vijver and S. Programme. **Limited cross-border infections in patients newly diagnosed with HIV in Europe.** *Retrovirology* 10: 36. 2013.
- K. S. Frese, B. Meder, A. Keller, S. Just, J. Haas, B. Vogel, S. Fischer, C. Backes, M. Matzas, D. Kohler, V. Benes, H. A. Katus and W. Rottbauer. **RNA splicing regulated by RBFOX1 is essential for cardiac function in zebrafish.** *J Cell Sci* 128 (16): 3030-40. 2015.
- C. Grimm, L. Chavez, M. Vilardell, A. L. Farrall, S. Tierling, J. W. Bohm, P. Grote, M. Lienhard, J. Dietrich, B. Timmermann, J. Walter, M. R. Schweiger, H. Lehrach, R. Herwig, B. G. Herrmann and M. Morkel. **DNA-methylome analysis of mouse intestinal adenoma identifies a tumour-specific signature that is partly conserved in human colon cancer.** *PLoS Genet* 9 (2): e1003250. 2013.
- W. Gu, B. Zhou, T. Geyer, M. Hutter, H. Fang and V. Helms. **Design of a gated molecular proton channel.** *Angew Chem Int Ed Engl* 50 (3): 768-71. 2011.
- K. Halachev, H. Bast, F. Albrecht, T. Lengauer and C. Bock. **Epi-Explorer: live exploration and global analysis of large epigenomic datasets.** *Genome Biol* 13 (10): R96. 2012.
- M. Hamed, C. Spaniol, M. Nazarieh and V. Helms. **TFmiR: a web server for constructing and analyzing disease-specific transcription factor and miRNA co-regulatory networks.** *Nucleic Acids Res* 43 (W1): W283-8. 2015.
- M. Hamed, C. Spaniol, A. Zapp and V. Helms. **Integrative network-based approach identifies key genetic elements in breast invasive carcinoma.** *BMC Genomics* 16 Suppl 5: S2. 2015.
- A. C. Hauschild, T. Frisch, J. I. Baumbach and J. Baumbach. **Carotta: Revealing Hidden Confounder Markers in Metabolic Breath Profiles.** *Metabolites* 5 (2): 344-63. 2015.
- A. C. Hauschild, D. Kopczynski, M. D'Addario, J. I. Baumbach, S. Rahmann and J. Baumbach. **Peak detection method evaluation for ion mobility spectrometry by using machine learning approaches.** *Metabolites* 3 (2): 277-93. 2013.
- X. He, A. E. Cicek, Y. Wang, M. H. Schulz, H. S. Le and Z. Bar-Joseph. **De novo ChIP-seq analysis.** *Genome Biol* 16: 205. 2015.

- S. Hein, M. Bur, T. Kolb, B. Muellinger, U. F. Schaefer and C. M. Lehr. **The Pharmaceutical Aerosol Deposition Device on Cell Cultures (PADDCC) in vitro system: design and experimental protocol.** *Altern Lab Anim* 38 (4): 285-95. 2010.
- M. Hellmuth, N. Wieseke, M. Lechner, H. P. Lenhof, M. Midden-dorf and P. F. Stadler. **Phylogenomics with paralogs.** *Proc Natl Acad Sci U S A* 112 (7): 2058-63. 2015.
- A. K. Hildebrandt, M. Dietzen, T. Lengauer, H. P. Lenhof, E. Althaus and A. Hildebrandt. **Efficient computation of root mean square deviations under rigid transformations.** *J Comput Chem* 35 (10): 765-71. 2014.
- J. Hochscherf, D. Lindenblatt, M. Steinkruger, E. Yoo, O. Ulu-can, S. Herzig, O. G. Issinger, V. Helms, C. Gotz, I. Neundorf, K. Niefind and M. Pietsch. **Development of a high-throughput screening-compatible assay to identify inhibitors of the CK-2alpha/CK2beta interaction.** *Anal Biochem* 468C: 4-14. 2014.
- M. C. Hutter, C. Brengel, M. Negri, C. Henn, C. Zimmer, R. W. Hartmann, M. Empting and A. Steinbach. **Mechanistic details for anthraniloyl transfer in PqsD: the initial step in HHQ biosynthesis.** *J Mol Model* 20 (6): 2255. 2014.
- A. Jalali and N. Pfeifer. **Interpretable per case weighted ensemble method for cancer associations.** *Lecture Notes in Bioinformatics: Proceedings of WABI 2014.* 2014.
- O. V. Kalinina, N. Pfeifer and T. Lengauer. **Modelling binding between CCR5 and CXCR4 receptors and their ligands suggests the surface electrostatic potential of the co-receptor to be a key player in the HIV-1 tropism.** *Retrovirology* 10: 130. 2013.
- A. Kappel, C. Backes, Y. Huang, S. Zafari, P. Leidinger, B. Meder, H. Schwarz, W. Gumbrecht, E. Meese, C. F. Staehler and A. Keller. **MicroRNA in vitro diagnostics using immunoassay analyzers.** *Clin Chem* 61 (4): 600-7. 2015.
- E. Kayvanpour, T. Mansi, F. Sedaghat-Hamedani, A. Amr, D. Neumann, B. Georgescu, P. Seegerer, A. Kamen, J. Haas, K. S. Frese, M. Irawati, E. Wirsz, V. King, S. Buss, D. Mereles, E. Zitron, A. Keller, H. A. Katus, D. Comaniciu and B. Meder. **Towards Personalized Cardiology: Multi-Scale Modeling of the Failing Heart.** *PLoS One* 10 (7): e0134869. 2015.
- A. Keller, A. Graefen, M. Ball, M. Matzas, V. Boisguerin, F. Maixner, P. Leidinger, C. Backes, R. Khairat, M. Forster, B. Stade, A. Franke, J. Mayer, J. Spangler, S. McLaughlin, M. Shah, C. Lee, T. T. Harkins, A. Sartori, A. Moreno-Estrada, B. Henn, M. Sikora, O. Semino, J. Chiaroni, S. Roots, N. M. Myres, V. M. Cabrera, P. A. Underhill, C. D. Bustamante, E. E. Vigil, M. Samadelli, G. Cipollini, J. Haas, H. Katus, B. D. O'Connor, M. R. Carlson, B. Meder, N. Blin, E. Meese, C. M. Pusch and A. Zink. **New insights into the Tyrolean Iceman's origin and phenotype as inferred by whole-genome sequencing.** *Nat Commun* 3: 698. 2012.
- A. Keller, P. Leidinger, A. Bauer, A. Elsharawy, J. Haas, C. Backes, A. Wendschlag, N. Giese, C. Tjaden, K. Ott, J. Werner, T. Hackert, K. Ruprecht, H. Huwer, J. Huebers, G. Jacobs, P. Rosenstiel, H. Dommisch, A. Schaefer, J. Muller-Quernheim, B. Wullich, B. Keck, N. Graf, J. Reichrath, B. Vogel, A. Nebel, S. U. Jager, P. Staehler, I. Amarantos, V. Boisguerin, C. Staehler, M. Beier, M. Scheffler, M. W. Buchler, J. Wischhusen, S. F. Haeusler, J. Dietl, S. Hofmann, H. P. Lenhof, S. Schreiber, H. A. Katus, W. Rottbauer, B. Meder, J. D. Hoheisel, A. Franke and E. Meese. **Toward the blood-borne miRNome of human diseases.** *Nat Methods* 8 (10): 841-3. 2011.
- J. R. Kelsen, N. Dawany, C. J. Moran, B. S. Petersen, M. Sarmady, A. Sasson, H. Pauly-Hubbard, A. Martinez, K. Maurer, J. Soong, E. Rappaport, A. Franke, A. Keller, H. S. Winter, P. Mamula, D. Piccoli, D. Artis, G. F. Sonnenberg, M. Daly, K. E. Sullivan, R. N. Baldassano and M. Devoto. **Exome Sequencing Analysis Reveals Variants in Primary Immunodeficiency Genes in Patients With Very Early Onset Inflammatory Bowel Disease.** *Gastroenterology* 149 (6): 1415-24. 2015.
- S. M. Kessler, A. Barghash, S. Laggai, V. Helms and A. K. Kiemer. **Hepatic hepcidin expression is decreased in cirrhosis and HCC.** *J Hepatol* 62 (4): 977-9. 2015.
- S. M. Kessler, S. Laggai, A. Barghash, V. Helms and A. K. Kiemer. **Lipid metabolism signatures in NASH-associated HCC.-letter.** *Cancer Res* 74 (10): 2903-4. 2014.
- S. M. Kessler, S. Laggai, A. Barghash, C. S. Schultheiss, E. Lederer, M. Artl, V. Helms, J. Haybaeck and A. K. Kiemer. **IMP2/p62 induces genomic instability and an aggressive hepatocellular carcinoma phenotype.** *Cell Death Dis* 6: e1894. 2015.
- Y. Khatri, F. Hannemann, M. Girhard, R. Kappl, M. Hutter, V. B. Urlacher and R. Bernhardt. **A natural heme-signature variant of CYP267A1 from *Sorangium cellulosum* So ce56 executes diverse omega-hydroxylation.** *FEBS J* 282 (1): 74-88. 2015.
- S. Kind, S. Neubauer, J. Becker, M. Yamamoto, M. Volkert, G. Abendroth, O. Zelder and C. Wittmann. **From zero to hero – production of bio-based nylon from renewable resources using engineered *Corynebacterium glutamicum*.** *Metab Eng* 25: 113-23. 2014.
- A. Kling, P. Lukat, D. V. Almeida, A. Bauer, E. Fontaine, S. Sordello, N. Zaborannyi, J. Herrmann, S. C. Wenzel, C. Konig, N. C. Ammerman, M. B. Barrio, K. Borchers, F. Bordon-Pallier, M. Bronstrup, G. Courtemanche, M. Gerlitz, M. Geslin, P. Hammann, D. W. Heinz, H. Hoffmann, S. Klieber, M. Kohlmann, M. Kurz, C. Lair, H. Matter, E. Nuermberger, S. Tyagi, L. Fraisse, J. H. Grosset, S. Lagrange and R. Müller. **Antibiotics. Targeting DnaN for tuberculosis therapy using novel griselimycins.** *Science* 348 (6239): 1106-12. 2015.

- M. Kohlstedt, P. K. Sappa, H. Meyer, S. Maass, A. Zaprasis, T. Hoffmann, J. Becker, L. Steil, M. Hecker, J. M. van Dijl, M. Lalk, U. Mader, J. Stulke, E. Bremer, U. Volker and C. Wittmann. **Adaptation of *Bacillus subtilis* carbon core metabolism to simultaneous nutrient limitation and osmotic challenge: a multi-omics perspective.** *Environ Microbiol* 16 (6): 1898-917. 2014.
- G. Lawyer. **Understanding the influence of all nodes in a network.** *Sci Rep* 5: 8665. 2015.
- G. Lawyer. **Stochasticity in pandemic spread over the World Airline Network explaining by local flight connections.** arXiv 2015 1504.05743. 2015.
- P. Leidinger, V. Galata, C. Backes, C. Stahler, S. Rheinheimer, H. Huwer, E. Meese and A. Keller. **Longitudinal study on circulating miRNAs in patients after lung cancer resection.** *Oncotarget* 6 (18): 16674-85. 2015.
- P. Leidinger, A. Keller, L. Milchram, C. Harz, M. Hart, A. Werth, H. P. Lenhof, A. Weinhausel, B. Keck, B. Wullich, N. Ludwig and E. Meese. **Combination of Autoantibody Signature with PSA Level Enables a Highly Accurate Blood-Based Differentiation of Prostate Cancer Patients from Patients with Benign Prostatic Hyperplasia.** *PLoS One* 10 (6): e0128235. 2015.
- T. Lengauer. **Stellenwert der Bioinformatik für die personalisierte Medizin.** *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz* 56 (11): 1489-1494. 2013.
- T. Lengauer, N. Pfeifer and R. Kaiser. **Personalized HIV therapy to control drug resistance.** *Drug Discov Today Technol* 11: 57-64. 2014.
- U. Lesnik, T. Lukezic, A. Podgorsek, J. Horvat, T. Polak, M. Sala, B. Jenko, K. Harmrolfs, A. Ocampo-Sosa, L. Martinez-Martinez, P. R. Herron, S. Fujs, G. Kosec, I. S. Hunter, R. Müller and H. Petkovic. **Construction of a new class of tetracycline lead structures with potent antibacterial activity through bio-synthetic engineering.** *Angew Chem Int Ed Engl* 54 (13): 3937-40. 2015.
- W. Y. Leung, T. Marschall, Y. Paudel, L. Falquet, H. Mei, A. Schönhuth and T. Y. Maoz Moss. **SV-AUTOPILOT: optimized, automated construction of structural variation discovery and benchmarking pipelines.** *BMC Genomics* 16: 238. 2015.
- D. Li, C. Beisswenger, C. Herr, J. Hellberg, G. Han, T. Zakharkina, M. Voss, R. Wiewrodt, R. M. Bohle, M. D. Menger, R. M. Schmid, D. Stockel, H. P. Lenhof and R. Bals. **Myeloid cell RelA/p65 promotes lung cancer proliferation through Wnt/ beta-catenin signaling in murine and human tumor cells.** *Oncogene* 33 (10): 1239-48. 2014.
- N. Ludwig, Y. J. Kim, S. C. Mueller, C. Backes, T. V. Werner, V. Galata, E. Sartorius, R. M. Bohle, A. Keller and E. Meese. **Post-transcriptional deregulation of signaling pathways in meningioma subtypes by differential expression of miRNAs.** *Neuro Oncol* 17 (9): 1250-60. 2015.
- N. Ludwig, N. Nourkami-Tutdibi, C. Backes, H. P. Lenhof, N. Graf, A. Keller and E. Meese. **Circulating serum miRNAs as potential biomarkers for nephroblastoma.** *Pediatr Blood Cancer* 62 (8): 1360-7. 2015.
- P. Lutsik, L. Feuerbach, J. Arand, T. Lengauer, J. Walter and C. Bock. **BiQ Analyzer HT: locus-specific analysis of DNA methylation by high-throughput bisulfite sequencing.** *Nucleic Acids Res* 39 (Web Server issue): W551-6. 2011.
- F. Maixner, T. Overath, D. Linke, M. Janko, G. Guerriero, B. H. van den Berg, B. Stade, P. Leidinger, C. Backes, M. Jaremek, B. Kneissl, B. Meder, A. Franke, E. Egarter-Vigl, E. Meese, A. Schwarz, A. Tholey, A. Zink and A. Keller. **Paleoproteomic study of the Iceman's brain tissue.** *Cell Mol Life Sci* 70 (19): 3709-22. 2013.
- F. Maurer, A. C. Hauschild, K. Eisinger, J. Baumbach and A. Mayor. **MIMA – a software for analyte identification in mcc/ims chromatograms by mapping accompanying gc/ms measurement.** *International Journal for Ion Mobility Spectrometry* 17 (2): 95-101. 2014.
- National Research Council (US) Committee in Metagenomics. **The New Science of Metagenomics: Revealing the Secrets of Our Microbial Planet** The National Academies Collection: Reports funded by the National Institutes of Health. 2007.
- S. Metzler and O. V. Kalinina. **Detection of atypical genes in virus families using a one-class SVM.** *BMC Genomics* 15: 913. 2014.
- A. M. Miederer, D. Alansary, G. Schwar, P. H. Lee, M. Jung, V. Helms and B. A. Niemeyer. **A STIM2 splice variant negatively regulates store-operated calcium entry.** *Nat Commun* 6: 6899. 2015.
- L. Mikeev and V. Wolf. **Parameter estimation for stochastic hybrid models of biochemical reaction networks.** *Proceedings of the 15th ACM international conference on Hybrid Systems: Computation and Control*: 155-166. 2012.
- A. Mosa, M. C. Hutter, J. Zapp, R. Bernhardt and F. Hannemann. **Regioselective Acetylation of C21 Hydroxysteroids by the Bacterial Chloramphenicol Acetyltransferase I.** *Chembiochem* 16 (11): 1670-9. 2015.

- S. C. Mueller, C. Backes, J. Haas, I. S. Group, H. A. Katus, B. Meder, E. Meese and A. Keller. **Pathogenicity prediction of non-synonymous single nucleotide variants in dilated cardiomyopathy.** *Brief Bioinform* 16 (5): 769-79. 2015.
- S. C. Mueller, C. Backes, O. V. Kalinina, B. Meder, D. Stockel, H. P. Lenhof, E. Meese and A. Keller. **BALL-SNP: combining genetic and structural information to identify candidate non-synonymous single nucleotide polymorphisms.** *Genome Med* 7 (1): 65. 2015.
- D. Nguyen, V. Helms and P. H. Lee. **PRIMSIPLR: prediction of inner-membrane situated pore-lining residues for alpha-helical transmembrane proteins.** *Proteins* 82 (7): 1503-11. 2014.
- K. R. Patil, P. Haider, P. B. Pope, P. J. Turnbaugh, M. Morrison, T. Scheffer and A. C. McHardy. **Taxonomic metagenome sequence assignment with structured output models.** *Nat Methods* 8 (3): 191-2. 2011.
- M. Patterson, T. Marschall, N. Pisanti, L. van Iersel, L. Stougie, G. W. Klau and A. Schönhuth. **WhatsHap: Weighted Haplotype Assembly for Future-Generation Sequencing Reads.** *J Comput Biol* 22 (6): 498-509. 2015.
- N. Pfeifer and T. Lengauer. **Improving HIV coreceptor usage prediction in the clinic using hints from next-generation sequencing data.** *Bioinformatics* 28 (18): i589-i595. 2012.
- N. Pfeifer, H. Walter and T. Lengauer. **Association between HIV-1 coreceptor usage and resistance to broadly neutralizing antibodies.** *J Acquir Immune Defic Syndr* 67 (2): 107-12. 2014.
- A. Pironti, S. Sierra, R. Kaiser, T. Lengauer and N. Pfeifer. **Effects of sequence alterations on results from genotypic tropism testing.** *J Clin Virol* 65: 68-73. 2015.
- K. Pritchard-Jones, C. Bergeron, B. de Camargo, M. M. van den Heuvel-Eibrink, T. Acha, J. Godzinski, F. Oldenburger, L. Boccon-Gibod, I. Leuschner, G. Vujanic, B. Sandstedt, J. de Kraker, H. van Tinteren, N. Graf and S. R. T. S. Group. **Omission of doxorubicin from the treatment of stage II-III, intermediate-risk Wilms' tumour (SIOP WT 2001): an open-label, non-inferiority, randomised controlled trial.** *Lancet* 386 (9999): 1156-64. 2015.
- A. L. Rodrigues, J. Becker, A. O. de Souza Lima, L. M. Porto, C. Wittmann. **Systems metabolic engineering of Escherichia coli for gram scale production of the antitumor drug deoxyviolacein from glycerol.** *Biotechnol Bioeng*. 111:2280-9. 2014.
- K. Roomp, N. Beerenwinkel, T. Sing, E. Schülter, J. Büch, S. Sierra-Aragon, M. Däumer, D. Hoffmann, R. Kaiser, T. Lengauer and J. Selbig. **Arevir: a secure platform for designing personalized antiretroviral therapies against HIV.** *Lecture Notes in Computer Science: Third International Workshop on Data Integration in the Life Sciences (DILS 2006)* 4075: 185-194. 2006.
- R. Röttger, P. Kalaghatgi, P. Sun, C. Soares Sde, V. Azevedo, T. Wittkop and J. Baumbach. **Density parameter estimation for finding clusters of homologous proteins—tracing actinobacterial pathogenicity lifestyles.** *Bioinformatics* 29 (2): 215-22. 2013.
- N. Schäuble, S. Lang, M. Jung, S. Cappel, S. Schorr, O. Ulucan, J. Linxweiler, J. Dudek, R. Blum, V. Helms, A. W. Paton, J. C. Paton, A. Cavalie and R. Zimmermann. **BiP-mediated closing of the Sec61 channel limits Ca²⁺ leakage from the ER.** *EMBO J* 31 (15): 3282-96. 2012.
- S. E. Schelhorn, M. Fischer, L. Tolosi, J. Altmüller, P. Nurnberg, H. Pfister, T. Lengauer and F. Berthold. **Sensitive detection of viral transcripts in human tumor transcriptomes.** *PLoS Comput Biol* 9 (10): e1003228. 2013.
- S. E. Schelhorn, J. Mestre, M. Albrecht and E. Zotenko. **Infering physical protein contacts from large-scale purification data of protein complexes.** *Mol Cell Proteomics* 10 (6): M110004929. 2011.
- T. Schneider, A. C. Hauschild, J. I. Baumbach and J. Baumbach. **An integrative clinical database and diagnostics platform for biomarker identification and analysis in ion mobility spectra of human exhaled air.** *J Integr Bioinform* 10 (2): 218. 2013.
- M. H. Schulz, D. Weese, M. Holtgrewe, V. Dimitrova, S. Niu, K. Reinert and H. Richard. **Fiona: a parallel and automatic strategy for read error correction.** *Bioinformatics* 30 (17): i356-63. 2014.
- F. Sedaghat-Hamedani, E. Kayvanpour, L. Frankenstein, D. Mereles, A. Amr, S. Buss, A. Keller, A. Giannitsis, K. Jensen, H. Katus and B. Meder. **Biomarker changes after strenuous exercise can mimic pulmonary embolism and cardiac injury – a meta-analysis on 45 studies.** *Clinical Chemistry* 61 (10): 1246-1255. 2015.

- S. Shanak and V. Helms. **Hydration properties of natural and synthetic DNA sequences with methylated adenine or cytosine bases in the R.DpnI target and BDNF promoter studied by molecular dynamics simulations.** *J Chem Phys* 141 (22): 22D512. 2014.
- M. Sikora, M. L. Carpenter, A. Moreno-Estrada, B. M. Henn, P. A. Underhill, F. Sanchez-Quinto, I. Zara, M. Pitzalis, C. Sidore, F. Busonero, A. Maschio, A. Angius, C. Jones, J. Mendoza-Revilla, G. Nekhrizov, D. Dimitrova, N. Theodossiev, T. T. Harkins, A. Keller, F. Maixner, A. Zink, G. Abecasis, S. Sanna, F. Cucca and C. D. Bustamante. **Population genomic analysis of ancient and modern genomes yields new insights into the genetic ancestry of the Tyrolean Iceman and the genetic structure of Europe.** *PLoS Genet* 10 (5): e1004353. 2014.
- A. Smolinska, A. C. Hauschild, R. R. Fijten, J. W. Dallinga, J. Baumbach and F. J. van Schooten. **Current breathomics-a review on data pre-processing techniques and machine learning in metabolomics breath analysis.** *J Breath Res* 8 (2): 027105. 2014.
- Souren et al. **Adult monozygotic twins discordant for intra-uterine growth have indistinguishable genome-wide DNA methylation profiles.** *Genome Biol.* (2013): 14:R44 doi:10.1186/gb-2013-14-5-r44.
- N. K. Speicher and N. Pfeifer. **Integrating different data types by regularized unsupervised multiple kernel learning with application to cancer subtype discovery.** *Bioinformatics* 31 (12): i268-75. 2015.
- C. G. Spruijt, F. Gnerlich, A. H. Smits, T. Pfaffeneder, P. W. Jansen, C. Bauer, M. Munzel, M. Wagner, M. Muller, F. Khan, H. C. Eberl, A. Mensinga, A. B. Brinkman, K. Lephikov, U. Muller, J. Walter, R. Boelens, H. van Ingen, H. Leonhardt, T. Carell and M. Vermeulen. **Dynamic readers for 5-(hydroxy)methylcytosine and its oxidized derivatives.** *Cell* 152 (5): 1146-59. 2013.
- P. Sun, N. K. Speicher, R. Rottger, J. Guo and J. Baumbach. **Bi-Force: large-scale bicluster editing and its application to gene expression data biclustering.** *Nucleic Acids Res* 42 (9): e78. 2014.
- F. Surup, K. Viehrig, K. I. Mohr, J. Herrmann, R. Jansen and R. Müller. **Disciformycins A and B: 12-membered macrolide glycoside antibiotics from the myxobacterium *Pyxidicoccus fallax* active against multiresistant staphylococci.** *Angew Chem Int Ed Engl* 53 (49): 13588-91. 2014.
- L. Tolosi, J. Theissen, K. Halachev, B. Hero, F. Berthold and T. Lengauer. **A method for finding consensus breakpoints in the cancer genome from copy number data.** *Bioinformatics* 29 (14): 1793-800. 2013.
- O. Ulucan and V. Helms. **How Hydrophilic Proteins Form Non-specific Complexes.** *J Phys Chem B* 119 (33): 10524-30. 2015.
- O. Ulucan, T. Jaitly and V. Helms. **Energetics of Hydrophilic Protein-Protein Association and the Role of Water.** *J Chem Theory Comput* 10 (8): 3512-24. 2014.
- J. Wegert, N. Ishaque, R. Vardapour, C. Georg, Z. Gu, M. Bieg, B. Ziegler, S. Bausenwein, N. Nourkani, N. Ludwig, A. Keller, C. Grimm, S. Kneitz, R. D. Williams, T. Chagtai, K. Pritchard-Jones, P. van Sluis, R. Volckmann, J. Koster, R. Versteeg, T. Acha, M. J. O'Sullivan, P. K. Bode, F. Niggli, G. A. Tytgat, H. van Tinteren, M. M. van den Heuvel-Eibrink, E. Meese, C. Vokuhl, I. Leuschner, N. Graf, R. Eils, S. M. Pfister, M. Kool and M. Gessler. **Mutations in the SIX1/2 pathway and the DROSHA/DGCR8 miRNA microprocessor complex underlie high-risk blastemal type Wilms tumors.** *Cancer Cell* 27 (2): 298-311. 2015.
- K. J. Weissman and R. Müller. **Myxobacterial secondary metabolites: bioactivities and modes-of-action.** *Nat Prod Rep* 27 (9): 1276-95. 2010.
- T. Will and V. Helms. **Identifying transcription factor complexes and their roles.** *Bioinformatics* 30 (17): i415-21. 2014.
- T. Will and V. Helms. **PPIXpress: construction of condition-specific protein interaction networks based on transcript expression.** *Bioinformatics*, Vol. 32, P. 571-578, 2016.
- R. Wittler, T. Marschall, A. Schönhuth and V. Mäkinen. **Repeat- and error-aware comparison of deletions.** *Bioinformatics* 31 (18): 2947-54. 2015.
- S. Zafari, C. Backes, P. Leidinger, E. Meese and A. Keller. **Regulatory microRNA networks: complex patterns of target pathways for disease-related and housekeeping microRNAs.** *Genomics Proteomics Bioinformatics* 13 (3): 159-68. 2015.
- S. Zafari, C. Backes, E. Meese and A. Keller. **Circulating Biomarker Panels in Alzheimer's Disease.** *Gerontology* 61 (6): 497-503. 2015.
- M. Zazzi, F. Incardona, M. Rosen-Zvi, M. Prosperi, T. Lengauer, A. Altmann, A. Sonnerborg, T. Lavee, E. Schulter and R. Kaiser. **Predicting response to antiretroviral treatment by machine learning: the EuResist project.** *Intervirology* 55 (2): 123-7. 2012.

A. Zeke, T. Bastys, A. Alexa, A. Garai, B. Meszaros, K. Kirsch, Z. Dosztanyi, O. V. Kalinina and A. Remenyi. **Systematic discovery of linear binding motifs targeting an ancient protein interaction surface on MAP kinases.** *Mol Syst Biol* 11 (11): 837. 2015.

A. Zink, L. S. Wann, R. C. Thompson, A. Keller, F. Maixner, A. H. Allam, C. E. Finch, B. Frohlich, H. Kaplan, G. P. Lombardi, M. L. Sutherland, J. D. Sutherland, L. Watson, S. L. Cox, M. I. Miyamoto, J. Narula, A. F. Stewart, G. S. Thomas and J. Krause. **Genomic correlates of atherosclerosis in ancient humans.** *Glob Heart* 9 (2): 203-9. 2014.

7 Wege & Anfahrt

7 Wege & Anfahrt

Das Zentrum für Bioinformatik, Gebäude E2 1 befindet sich auf dem Campus der Universität des Saarlandes etwa fünf Kilometer nordöstlich vom Zentrum der Stadt Saarbrücken im Wald nahe bei Dudweiler.

Saarbrücken besitzt einen eigenen Flughafen (Saarbrücken-Ensheim) und ist mit Auto-, Shuttle-Bus- und Zugverbindungen an die Flughäfen Frankfurt und Luxemburg angebunden. Zugstrecken verbinden Saarbrücken im Stundentakt mit zahlreichen Städten innerhalb Deutschlands. Regelmäßig verkehrende Züge schaffen eine Anbindung an die Städte Metz, Nancy und Paris.

Autobahnen führen nach Mannheim/Frankfurt, Luxemburg/Trier/Köln, Straßburg und Metz/Nancy/Paris.

Sie erreichen den Campus.....

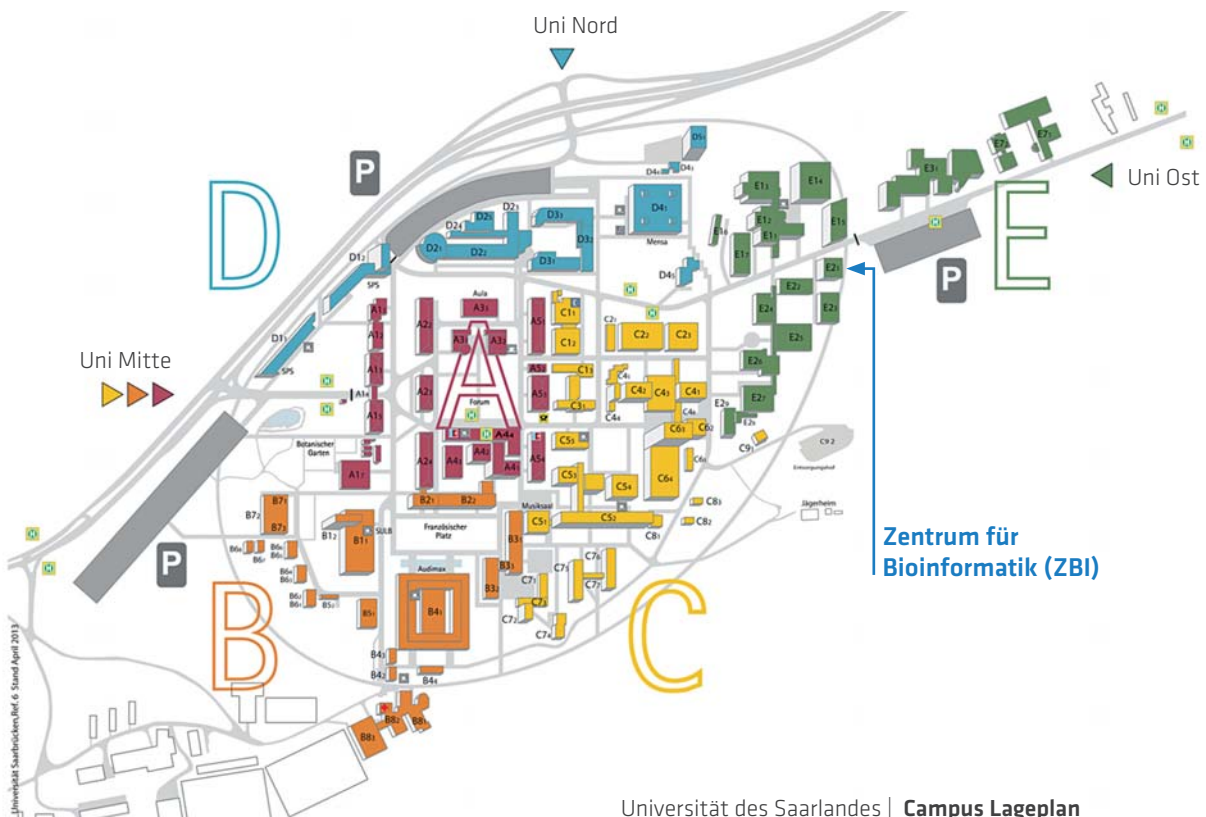
von Saarbrücken-Ensheim, Flughafen
mit dem Taxi in ungefähr 20 Minuten

von Saarbrücken, Hauptbahnhof
mit dem Taxi in ungefähr 15 Minuten
mit dem Bus in ungefähr 20 Minuten
Richtung „Dudweiler-Dudoplatz“ oder „UniversitätCampus“
Ausstieg „Universität Mensa“
alternativ Ausstieg „Universität Busterminal“

von Frankfurt oder Mannheim über die Autobahn A6
Abfahrt „St. Ingbert-West“
den weißen Schildern „Universität“ zum Campus folgend

von Paris über die Autobahn A4
Abfahrt „St. Ingbert-West“
den weißen Schildern „Universität“ zum Campus folgend

143



Impressum

Herausgeber

Zentrum für Bioinformatik
Pia Scherer-Geiss, Geschäftsführerin
Campus E2 1
D-66123 Saarbrücken

Redaktion & Koordination

Nicole Comtesse
Volkhard Helms
Karin Jostock
Olga Kalinina
Andreas Keller
Thomas Lengauer
Hans-Peter Lenhof
Tobias Marschall
Friederike Meyer zu Tittingdorf
Eckart Meese
Nico Pfeifer
Pia Scherer-Geiss
Ruth Schnepfen-Christmann
Marcel Schulz
Jörn Walter
Christoph Wittmann

Lektorat

Kamila Kolesniczenko

Kontakt

Zentrum für Bioinformatik
Telefon +49 681 302 686 05
Telefax +49 681 302 647 19
Email psg@bioinf.uni-sb.de
Internet <http://www.zbi.uni-saarland.de>

Berichtszeitraum

1. Januar 2010 bis 31. Dezember 2015

Gestaltung

Behr Design | Saarbrücken
Marcus Pirron
Olga Puhl

Druck

O/D DIE DRUCKEREI | Ottweiler
Gedruckt auf FSC-Mix zertifiziertem Papier



Zentrum für Bioinformatik
Campus E2 1
Universität des Saarlandes
D-66123 Saarbrücken
Telefon +49 681 302 6 86 05
Telefax +49 681 302 6 47 19

<http://www.zbi.uni-saarland.de>



UNIVERSITÄT
DES
SAARLANDES