



Liebe Leserinnen, liebe Leser,

wir freuen uns, Ihnen nach einer längeren, Corona-bedingten Pause eine neue Ausgabe des Forschungsmagazins ‚Blick in die Wissenschaft‘ in der Ausgabe 44/45 präsentieren zu können.

Die Corona-Pandemie hat auch die Universität Regensburg und alle ihre Mitglieder vor große Herausforderungen gestellt. Dennoch konnten zentrale Zukunftsprojekte weitergeführt und umgesetzt werden. So stellt vor allem die Gründung unserer neuen Fakultät für Informatik und Data Science (FIDS) einen wahren Meilenstein in der Geschichte und Entwicklung der Universität Regensburg dar. Als größtes Strukturprojekt seit der Gründung der Fakultät für Medizin vor 30 Jahren ist unsere Informatikfakultät ein Zukunftsprojekt von weitreichenden Dimensionen. Mit der neuen strategischen Schwerpunktsetzung im Bereich Informatik und Data Science und vor allem auch der Querschnittsorientierung der neuen Fakultät sieht sich die Universität Regensburg sehr gut gerüstet, ihre bisherigen Stärken in diesen Bereichen zu bündeln, weiter auszubauen und zu

vertiefen. Schließlich sind *Digital Transformations* als eines der vier Gestaltungsfelder und Zukunftsthemen in unserem *Hochschulentwicklungsplan 2025* fest verankert. Dieses Gestaltungsfeld adressiert die neue Fakultät ebenso wie den Bereich *Integrated Sciences in Life, Health, and Disease* als ein weiteres Schwerpunktgebiet unserer Universität.

Die Grundsatzbeschlüsse in den Gremien der Universität Regensburg im Sommer und Herbst 2019 zur Einrichtung der neuen Fakultät erfolgten nach einer vorhergehenden Phase intensiver Planungen dann letztlich fast zeitgleich mit der Regierungserklärung des Bayerischen Ministerpräsidenten Dr. Markus Söder am 10. Oktober 2019 und der Verkündung der Hightech Agenda Bayern. Unterstützt und beschleunigt durch die Mittel der Hightech Agenda Bayern konnte der Auf- und Ausbau der Fakultät für Informatik und Data Science zügiger umgesetzt werden, nachdem die neue Fakultät im März 2020 formal gegründet und im Laufe des WS 2021/22 aus sich heraus handlungs- und

funktionsfähig wurde. Im Mai 2022 konnten wir gemeinsam mit Ministerpräsident Dr. Markus Söder und Staatsminister für Wissenschaft und Kunst Markus Blume den offiziellen Kickoff für die Fakultät begehen. Dass dieser komplexe Prozess im Kontext der Herausforderungen der Corona-Pandemie vollzogen und abgeschlossen werden konnte, ist ein Zeichen für die Bedeutung dieser gesamtuniversitär-strategischen Maßnahme und für den Rückhalt für das Großprojekt in der universitären Gemeinschaft.

Im Laufe des Gründungsprozesses ist es gelungen, die verschiedenen Informatiknahen und -interessierten Kräfte der Universität an einen Tisch zu bringen und gemeinsam ein zukunftsorientiertes Konzept für die Fakultät zu entwickeln. Ein externes Gutachten mit hochrangiger Expertise skizzierte und evaluierte 2019 wesentliche inhaltliche Schwerpunkte und Strukturierungen für die neue Fakultät, an denen sich in den Jahren 2019-2021 die von Vizepräsident Prof. Dr. Nikolaus Korber geleitete Gründungskommission in der

konkreten Arbeit zum Aufbau der Fakultät orientierte. In insgesamt 15 Berufungsverfahren wurden die ersten neuen Professuren in der Fakultät zügig besetzt – ein Prozess, der in Kürze abgeschlossen sein wird. Im Besetzungsprozess hat sich vor allem auch gezeigt, wie attraktiv die Neugründung einer Fakultät und die Möglichkeiten zur Mitgestaltung und zum Aufbau neuer Strukturen für Wissenschaftlerinnen und Wissenschaftler sind und wie viel Zukunftspotential von unserer neuen Fakultät ausgeht. So konnten wir zum Wintersemester 2023/24 130 Studierende für den B.Sc. Informatik und den B.Sc. Data Science begrüßen.

Im vorliegenden Heft von ‚Blick in die Wissenschaft‘ möchten wir Ihnen nunmehr vor allem die Forschungsaktivitäten der Fakultät für Informatik und Data Science näher vorstellen. Dabei beglückwünsche ich die Fakultät, dass sie bereits eineinhalb Jahre nach ihrer vollständigen Handlungs- und Funktionsfähigkeit und während der weiteren Planungen zum Aufbau und der Ausarbeitung ihrer Studiengänge insbe-

sondere im Master-Bereich ein so vielfältiges Themenheft zu ihren aktuellen Forschungsarbeiten vorlegen konnte.

Das facettenreiche und vielfältige Themenspektrum dieses Sonderhefts illustriert, wie die Fakultät für Informatik und Data Science die an der Universität Regensburg bisher vorhandenen IT-Kompetenzen erfolgreich bündelt und in die Zukunft gerichtet erweitert. Sie ermöglicht die essentielle interdisziplinäre Vernetzung mit der gesamten Universität, von den Geistes- und Sozialwissenschaften bis zu den Natur- und Lebenswissenschaften. Die Beiträge verdeutlichen, wie interdisziplinäre Forschung das Fundament starker methodischer und fachlicher Grundlagen weiterentwickelt und wie die bisherigen Informatik-Schwerpunkte der Universität Regensburg (Computational Science, Informationswissenschaft, Medieninformatik, Wirtschaftsinformatik) erfolgreich in die neue Fakultät überführt werden konnten und die Querschnittsorientierung unterstützen.

Das vorliegende Heft mit seinem Schwerpunkt auf aktuellen Forschungsar-

beiten begleitet den im Wintersemester 2023/24 erfolgten Start der beiden grundständigen Bachelor-Studiengänge Informatik und Data Science. Ein kurzer Überblicksbeitrag zur Lehre in der neuen Fakultät zeigt anschaulich die bereits gewachsene Vielfalt der Informatikstudiengänge und die intensive und gelebte Verbindung von Forschung und Lehre auch an dieser neuen Fakultät.

Unsere neue Fakultät leistet hervorragende Arbeit und ich bin sicher, dass Ihnen die nachfolgenden Seiten einen spannenden Einblick in die verschiedenen Facetten der FIDS geben werden. Ein ganz besonderes Dankeschön möchte ich an dieser Stelle an das gesamte Dekanat der Fakultät für Informatik und Data Science und insbesondere an Forschungsdekanin Prof. in Dr. Meike Klettke richten, die für diese Sonderausgabe die Koordinationsarbeit der vorliegenden Ausgabe federführend übernommen hat.

Prof. Dr. Udo Hebel
Präsident der Universität Regensburg

**Blick in die Wissenschaft
Forschungsmagazin
der Universität Regensburg**

ISSN 0942-928-X

Heft 44/45

31. Jahrgang

Herausgeber

Prof. Dr. Udo Hebel
Präsident der Universität Regensburg

Redaktionsleitung für diese Ausgabe

Prof.in Dr. Meike Klettke / Fakultät für Informatik und Data Science

Redaktionsbeirat

Prof. Dr. jur. Christoph Althammer
Prof. Dr. rer. nat. Ferdinand Evers
Prof. Dr. rer. nat. Stefan Friedl
Prof. Dr. rer. nat. Mark W. Greenlee
Prof. Dr. theol. Andreas Merkt
Prof. Dr. phil. Omar W. Nasim
Prof. Dr. rer. nat. Klaus Richter
Prof. Dr. rer. pol. Daniel Rösch
Prof. Dr. med. Ernst Tamm
Prof. Dr. paed. Oliver Tepner
Prof. Dr. phil. Christiane Heibach

Universität Regensburg
93040 Regensburg
Telefon +49 941 9432300
Telefax +49 941 9433310

Verlag

Universitätsverlag Regensburg GmbH
Leibnizstraße 13, 93055 Regensburg

Telefon +49 941 78785-0
Telefax +49 941 78785-16

info@univerlag-regensburg.de
www.univerlag-regensburg.de
Geschäftsführer: Dr. Albrecht Weiland,
Felix Weiland M.A.

Abonnementsservice

bestellung@univerlag-regensburg.de

Anzeigenleitung

Larissa Nevecny
MME-Marquardt
info@mme-marquardt.de

Herstellung

Universitätsverlag Regensburg GmbH
info@univerlag-regensburg.de

**Einzelpreis € 7,00
Doppelheft € 14,00**

Jahresabonnement

bei zwei Ausgaben pro Jahr

€ 10,00 / ermäßigt € 9,00

Für Schüler, Studierende und Akademiker/innen im Vorbereitungsdienst (inkl. 7% MwSt.) zzgl. Versandkostenpauschale € 1,64 je Ausgabe. Bestellung beim Verlag. Für **Mitglieder des Vereins der Ehemaligen Studierenden der Universität Regensburg e.V.**, des **Vereins der Freunde der Universität Regensburg e.V.** und des **Vereins ehemaliger Zahnmedizinstudenten Regensburg e.V.** ist der Bezug des Forschungsmagazins im Mitgliedsbeitrag enthalten.

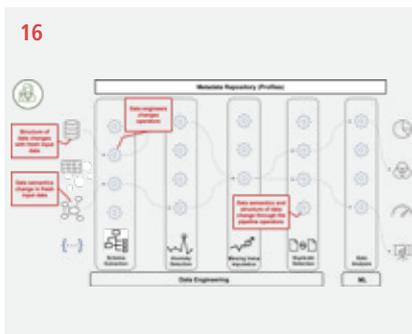
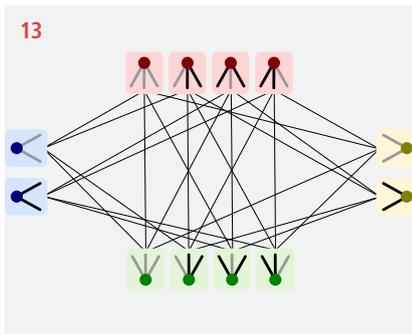
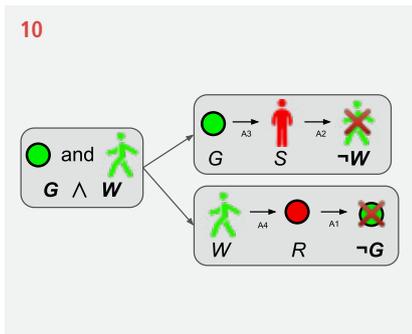
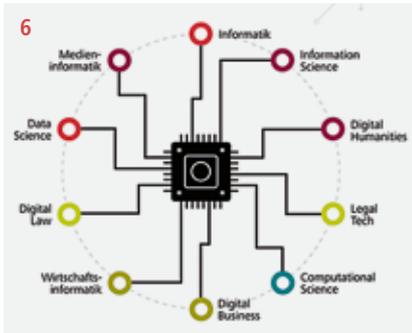


Rohstoffe
Transporte
Produktion

g CO₂e
492
Pro Produkt

CO₂-Emissionen
ausgeglichen

Inhalt



Einleitung 5
Florian Erhard, Bernd Heinrich, Meike Klettke, Christian Wolff

Lehre an der Fakultät für Informatik und Data Science 6
Florian Erhard, Udo Kruschwitz, Bernd Heinrich, Christian Wolff

Automatisches Beweisen: Methoden und Anwendungen 10
Julie Cailler, Philipp Rümmer

Algorithmen und Komplexitätstheorie 13
Radu Curticapean

Evolution in Datenbanken und Data Engineering Workflows 16
Meike Klettke

IoT-basiertes Prozessmanagement – Mobile Benutzerführung in der digitalen Fabrik 19
Stefan Schönig

Cyber Threat Intelligence: Gemeinschaftliche IT-Sicherheit durch den Austausch von Informationen 23
Johannes Grill, Daniel Schlette, Günther Pernul

Kann man den Entscheidungen Künstlicher Intelligenz trauen? Zu den Auswirkungen unsicherer Daten auf die Entscheidungen Neuronaler Netze 26
Thomas Krapf, Bernd Heinrich

Mensch vs. Maschine: Wettbewerb und Kooperation mit künstlicher Intelligenz in digitalen Märkten 30
Andreas Schauer, Daniel Schnurr

Notfallpläne für den Ernstfall testen 34
Maria Leitner

Maschinelles Lernen mit Anwendungen in den Naturwissenschaften 37
Merle Behr, Markus Schmitt

Automatisierte, KI-basierte Analyse von Bilddaten:

Der Lehrstuhl für Bildverarbeitung

Dorit Merhoff

40

Die Genome des Menschen – Forschungsschwerpunkte der Arbeitsgruppe für Algorithmische Bioinformatik

Birte Kehr

43

Algorithmen zum Entschlüsseln der Genregulation

Francisca Rojas Ringeling, Stefan Canzar

46

Mit Hilfe von Daten Immunprozesse entschlüsseln:

Der Lehrstuhl Computational Immunology

Florian Erhard

49

Maschinelles Lernen enthüllt den verborgenen Prozess der Tumorentstehung

Linda Hu, Andreas Lösch, Rainer Spang

52

Allgegenwärtige Mensch-Maschine-Interaktion: Entwicklung, Forschung und Infrastruktur der Medieninformatik

Raphael Wimmer, Johanna Bogon, Niels Henze, Christian Wolff

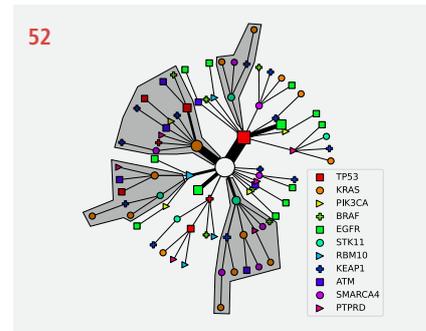
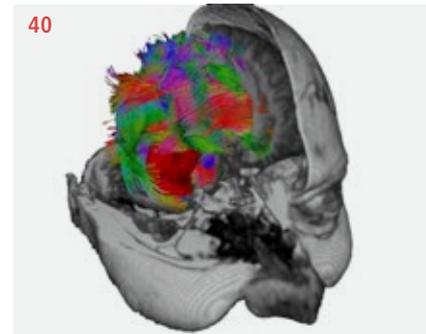
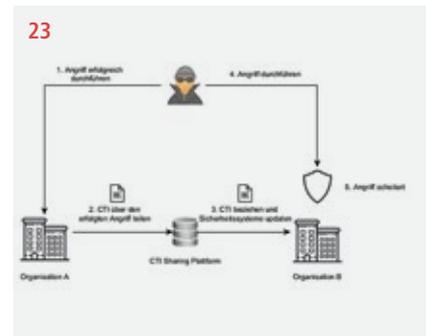
54

Wissen aus dem Internet – Genug, genau, geprüft und geeignet?

Informationswissenschaft in Regensburg

Udo Kruschwitz, Bernd Ludwig, David Elsweiler

60



SOFTWAREENTWICKLER (M/W/D) GESUCHT

Voll- oder Teilzeit (30 -40 Std.) | Alle Fachbereiche

Hallo, **wir sind mb Support.**

Seit 20 Jahren unterstützen wir die Versicherungswirtschaft mit unseren Lösungen, u.a. Konzerte und Kunstwerke weltweit zu versichern. Als inhabergeführtes **Familienunternehmen** arbeiten wir auf Augenhöhe in einer familiären Atmosphäre. Als stark wachsendes und profitables Unternehmen bieten wir gleichzeitig **attraktive Benefits**.

Wer bist du? Jedes unserer rund 60 Teammitglieder bringt seine eigene Persönlichkeit, Geschichte und Perspektive mit. Für uns zählt, was dich interessiert, was dich antreibt, wie du bist. Du entscheidest, wie du bei uns mitgestalten möchtest. **Wir freuen uns auf dich.**

Die gesamte Stellenausschreibung findest du auf www.mbsupport.de/karriere.

www.mbsupport.de | Friedenstraße 18 | 93053 Regensburg | +49 941 942 60 0 | mb Support GmbH



Mit Hilfe von Daten Immunprozesse entschlüsseln

Der Lehrstuhl Computational Immunology

Prof. Dr. Florian Erhard

Unser Immunsystem ist zweifellos eines der erstaunlichsten Wunderwerke der Natur. Es ist ein hochkomplexes Netzwerk aus Zellen, Proteinen, Organen und Geweben, das ständig daran arbeitet, uns vor einer schier endlosen Vielfalt von Krankheitserregern und Gefahren zu schützen. Doch wie schafft es das Immunsystem, diese komplexe Aufgabe zu bewältigen, ohne uns selbst zu schaden? Zu verstehen, wie diese Balance zwischen der Abwehr von Krankheiten und der Vermeidung einer übermäßigen Reaktion gegen den eigenen Körper funktioniert, ist eine der faszinierendsten Herausforderungen in der heutigen Biomedizin.

So genannte »omics« Technologien haben in den letzten Jahren einen atemberaubenden Fortschritt erlebt. Diese Technologien können zwar einen riesigen Berg an Daten über alle möglichen biologischen Vorgänge generieren, die Art und Weise wie aus den Daten Informationen gewonnen werden müssen stellt uns aber vor große Herausforderungen: Stellen Sie sich ein Puzzle mit Millionen von Teilen vor. Sie kennen das Motiv nicht, das das fertige Puzzle zeigen wird. Zudem fehlen viele Teile, viele Teile sind leider identisch und oft sind Teile schwer beschädigt. Ein solches Puzzle gilt es nun zu lösen, denn das fertige Puzzlemotiv würde uns eine neue Erkenntnis über einen Aspekt des Immunsystems liefern.

Unser Forschungsansatz ist multidisziplinär und tief in den Schnittstellen zwischen Biologie, Informatik und Statistik verwurzelt. Wir entwickeln innovative computergestützte Methoden, um spezielle Arten von solchen Puzzles lösbar zu machen,

d.h. Informationen aus »omics«-Daten zu ziehen und daraus Erkenntnisse zu gewinnen. Doch dies ist nur ein Teil unserer Mission. Unser zentrales Interesse gilt dem Immunsystem und seiner Funktionsweise. Unsere Forschung reicht von der Identifizierung von zellulären Erkennungsmerkmalen, mit denen das Immunsystem kranke von gesunden Zellen unterscheiden kann, bis zur Untersuchung der molekularen und dynamischen Wechselwirkungen zwischen Proteinen und Genen, die bei der primären Immunabwehr eine Rolle spielen.

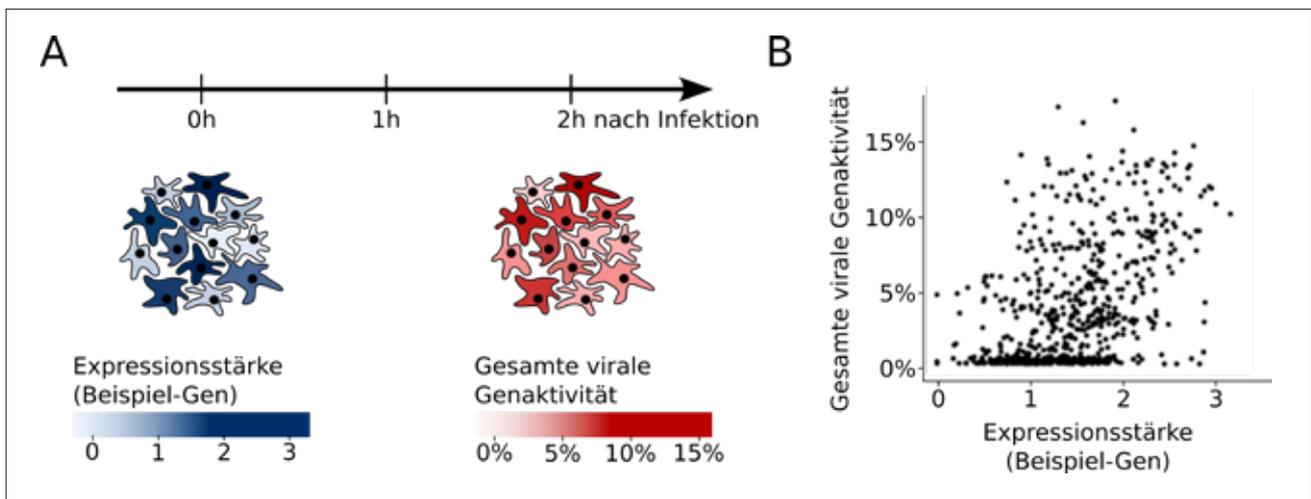
Wie funktioniert das zellintrinsic Immunsystem

Beinahe jede unserer Zellen im Körper hat erstaunliche Fähigkeiten, sich und andere Zellen gegen Krankheitserreger wie Viren oder Bakterien zu verteidigen: Nach Infektion erkennen sie spezielle Muster in diesen Erregern und setzen daraufhin komplexe Abwehrmechanismen in Gang. Dazu gehören Entzündungsreaktionen, das Ausschütten des Botenstoffs »Interferon«, welcher Nachbarzellen alarmiert, und die darauf folgende Produktion von hundert von verschiedenen Effektorproteinen, die auf viele unterschiedliche Weisen Viren und Bakterien abwehren. Zu diesen Effektorproteinen gehören zum Beispiel Enzyme, die virales Genmaterial zerschneiden können, Adhäsionsmoleküle, die die Freisetzung neuer Viruspartikel verhindern, oder Proteine, die zum Zelltod führen.

All diese Prozesse geschehen nicht isoliert – sie interagieren miteinander und mit anderen zellulären Faktoren, was den Grad

der Abwehr bestimmt. Zudem ist das Timing wichtig: Wann welche Prozesse wie stark aktiviert werden, ist entscheidend, ob Krankheitserreger erfolgreich bekämpft werden können oder die Oberhand gewinnen. Für normale Körperfunktionen müssen Zellen außerdem nach erfolgreicher Bekämpfung der Infektion aus diesem alarmierten Zustand in den Normalzustand zurückkehren können. Dieses ganze System ist in seiner Komplexität genau so ausbalanciert, um auf der einen Seite eine rasche und starke Immunantwort hervorzurufen, auf der anderen Seite aber eine andauernde Überreaktion zu verhindern. Viren und Bakterien haben im Lauf ihrer Evolution gelernt, diese komplexen Vorgänge zu unterlaufen, und können dieses System aus der Balance bringen. Man weiß zum Beispiel, dass schwere Verläufe von COVID-19 mit einer verstärkten Entzündungsreaktion einhergehen, aber gleichzeitig die Interferon-Antwort durch das Virus unterdrückt wird.

Um solche zeitlichen Verläufe in einzelnen Zellen erstmals messbar zu machen, haben wir zusammen mit experimentellen Gruppen ein neues Verfahren entwickelt: Dieses Verfahren basiert auf herkömmlicher Einzelzellsequenzierung, mit deren Hilfe die Aktivität tausender Gene in tausenden bis zehntausenden von Zellen festgestellt werden kann. Anders als herkömmliche Einzelzellsequenzierung, die nur Momentaufnahmen liefert und die Zellen dabei zerstört, liefert unser *scSLAM-seq* [2] Verfahren zusätzlich Information über zeitliche Dynamiken. Diese Zusatzinformation äußert sich zum Beispiel darin, dass wir damit den genauen Zustand aller



1 Das Konzept »Heterogeneity sequencing« **A** Bei einem Experiment werden viele Zellen (hier dargestellt: 13 Zellen) zum Zeitpunkt 0 mit Viren infiziert. In einzelnen Zellen weist die Expressionsstärke jedes Gens eine natürliche Heterogenität auf, d. h. das Gen ist unterschiedlich stark aktiv in jeder Zelle (in unterschiedlichen Blautönen dargestellt). Zwei Stunden (2h) später hat das Virus mehr oder weniger erfolgreich in den einzelnen Zellen dafür gesorgt, dass seine eigenen Gene aktiviert wurden (in unterschiedlichen Rottönen dargestellt). Gene, für die die Expression zum Zeitpunkt 0 und der Erfolg der viralen Genexpression korreliert sind, sind Kandidaten für pro- oder antivirale Faktoren: Ihre Expression vor Infektion begünstigt (positive Korrelation) oder vermindert die Infektion (negative Korrelation). Um solche Korrelationen bestimmen zu können, benötigt man Messungen von Expressionsstärken vor **und** nach Infektion. **B** Graphische Darstellung der Expressionsstärke zum Zeitpunkt 0 und der viralen Genaktivität für hunderte von Zellen (Punkte) eines potenziellen pro-viralen Faktors.

Zellen vor **und** nach der Infektion messen können anstatt nur vorher **oder** nachher [2]. Mit dieser Art von Daten sind gänzlich neue Analysekonzepte möglich (Abb. 1).

Der Knackpunkt an scSLAM-seq ist die Datenanalyse [3]. Die Art und Weise, wie die Informationen über die zeitliche Dynamik in den Daten hinterlegt sind, macht es zum Beispiel nötig, nicht nur feste Werte aus den Daten abzuleiten (z. B. »welcher Anteil der gesamten Aktivierung stammt aus der letzten Stunde«), sondern die Mess-Unsicherheit abzubilden (»die Daten belegen, dass dieser Anteil im Bereich 10-30% liegt«).

Erst diese von uns entwickelten Analysemethoden für scSLAM-seq machen es möglich, die zeitliche Dynamik dieser komplexen Vorgänge der zellintrinsic Immunabwehr vollständig zu erfassen. Aufbauend auf solchen Messungen werden wir Modelle entwickeln, mit dem wir die genauen zeitlichen Abläufe vorhersagen können: Wann werden in welchen Zellen welche Signalwege aktiv, was bewirken im Einzelnen die aktivierten Faktoren und wie beeinflussen sich unterschiedliche Zellen in ihrer Nachbarschaft.

Aufklären der Funktionen und Mechanismen von kryptischen Peptiden

Die meisten Zellen haben neben der zellintrinsic Immunabwehr außerdem die Fähigkeit, einen anderen Teil des Immunsystems in Gang zu setzen: Sie tragen sogenannte humane Leukozytenantigene Klasse I (HLA-I) auf ihrer Zelloberfläche. Diese Strukturen zeichnen sich verantwortlich, Peptide, das sind Fragmente von Proteinen, die in der Zelle hergestellt werden, nach außen zu präsentieren. Spezialisierte Immunzellen namens T-Zellen können unterscheiden, ob es sich bei einem präsentierten Peptid um ein körpereigenes oder fremdes handelt. »Fremd« heißt zum Beispiel, dass das Peptid aus einem viralen Protein stammt, oder aufgrund einer in einer Tumorzelle entstandenen Mutation als nicht mehr körpereigen erkannt wird. Eine T-Zelle, die ein fremdes Peptid erkennt, kann die präsentierende Zelle daraufhin abtöten und so Infektionskrankheiten oder das Entstehen von Tumoren eindämmen oder verhindern.

Mit Hilfe einer Kombination mehrerer »omics«-Technologien und neuen Datenanalysemethoden [4,5] haben wir vor kurzer Zeit entdeckt, dass bis zu 10% aller präsentierten Peptide nicht etwa Fragmente von bekannten Proteinen, sondern das Produkt von sogenannter *kryptischer*

Translation ist: Bekannte Proteine üben im Allgemeinen wichtige molekulare Funktionen aus und überdauern für mehrere Stunden oder sogar Tage in unseren Zellen bevor sie einem natürlichen Abbauprozess unterliegen. Kryptische Translation hingegen stellt (zum Teil extrem kleine) Produkte her, die unmittelbar nach ihrer Herstellung sofort abgebaut werden. Lange stellte sich die Frage, warum Zellen die Energie dafür aufwenden, kryptische Peptide herzustellen.

Eine Antwort darauf ist, dass sie den T-Zellen als erweiterte Erkennungsmerkmale zur Verfügung stehen. Aus dieser Entdeckung ergeben sich mehrere Fragestellungen, die wir in Zukunft erforschen werden:

- Zusammen mit experimentellen Kollaborationspartnern evaluieren wir den Nutzen von kryptischen Peptiden für immuntherapeutische Ansätze.
- Mit Hilfe weiterentwickelter »omics«-Messungen und unseren Analyse-Methoden werden wir dem Zusammenhang zwischen kryptischer Translation und der Präsentation kryptischer Peptide quantitativ untersuchen
- Wir werden erforschen, warum so viel kryptische Translation stattfindet, obwohl unsere Zellen eine ganze Reihe von Kontrollmechanismen besitzen, die eine erfolgreiche Herstellung von funktionellen Proteinen sicherstellen.

Wir leben zweifellos in aufregenden Zeiten, in denen die Synergie zwischen fortschrittlicher Rechenleistung, innovativen KI-Methoden und modernen Technologien zur Datengewinnung die Immunologie revolutioniert. Diese multidisziplinären Fortschritte ermöglichen es uns, die komplexen Prozesse des Immunsystems tiefer zu durchdringen als je zuvor. Während wir in der Grundlagenforschung immer weiter voranschreiten, ist es ermutigend zu sehen, wie erste Entwicklungen bereits den Weg zu den Patienten finden und einen positiven Einfluss auf die Medizin haben.

Literatur

- [1] Erhard, F. *et al.* scSLAM-seq reveals core features of transcription dynamics in single cells. *Nature* **571**, 419–423 (2019).
- [2] Erhard, F. *et al.* Time-resolved single-cell RNA-seq using metabolic RNA labelling. *Nat Rev Methods Primers* **2**, 1–18 (2022).
- [3] Jürges, C., Dölken, L. & Erhard, F. Dissecting newly transcribed and old RNA using GRAND-SLAM. *Bioinformatics* **34**, i218–i226 (2018).
- [4] Erhard, F. *et al.* Improved Ribo-seq enables identification of cryptic translation events. *Nature Methods* **15**, 363–366 (2018).
- [5] Erhard, F., Dölken, L., Schilling, B. & Schlosser, A. Identification of the Cryptic HLA-I Immunopeptidome. *Cancer Immunol Res* **8**, 1018–1026 (2020).

DREI FRAGEN ...

Welches Fachbuch lesen Sie gerade?

Finite Mixture Models von Geoffrey McLachlan und David Peel. Diese Klasse von Modellen ist unglaublich wichtig für verschiedenste biologische Daten, und dieses Buch bietet einen hervorragenden Überblick über die theoretischen Grundlagen, Algorithmen für Parameterschätzung, Probleme der Identifizierbarkeit von Parametern und illustriert die Anwendung dieser Modelle in verschiedenen Bereichen.

Welche Entwicklung im Bereich Data Science hat Sie im letzten Jahr am meisten überrascht?

Maschinelle Lernverfahren werden seit jeher in der Bioinformatik für verschiedenste Probleme eingesetzt. Auf Konferenzen hat man aus meiner Sicht immer sehr schnell erkannt, ob die oder der Vortragende aus der Ecke »Machine learning« kam und aufregende neue Methoden auf biologische Daten angewendet hat, oder ob der Vortrag aus der Biologie-Ecke kam, d. h. ein spannendes biologisches Problem angegangen wurde und dazu maschinellen Lernen verwendet wurde. Auf der kürz-

lich stattgefundenen ISMB, der wichtigsten Bioinformatik-Konferenz, war schön zu sehen, dass diese beiden Lager nicht (mehr) existieren. Es wurden ausschließlich relevante und hochspannende biologische Fragestellungen präsentiert, die mit modernen und auf die Problemstellung passenden Methoden des maschinellen Lernens gelöst wurden.

Was ist Ihre größte fachliche Herausforderung?

Jede Datenanalyse-Methode von »omics«-Daten trifft gewisse Annahmen. Viele dieser Annahmen sind vereinfachend und damit bei genauer Betrachtung schlichtweg falsch. Bei manchen Annahmen hat das keine oder sehr geringe Auswirkungen auf Ergebnisse, die Verletzung anderer Annahmen kann aber Ergebnisse drastisch verfälschen. Welche Annahmen zu welcher von diesen beiden Kategorien gehören, erfordert oft ein sehr genaues Verständnis der Biologie und der Messmethode sowie der explorativen Analyse vieler Datensätze. Dieses Einschätzen von vereinfachenden Annahmen halte ich für meine größte fachliche Herausforderung.



Foto © Margit Scheid

Prof. Dr. Florian Erhard hat in München Bioinformatik studiert und 2014 dort am Lehrstuhl von Prof. Dr. Ralf Zimmer promoviert. Nach einer kurzen Postdoc-Phase wurde er im Jahr 2016 zum Juniorprofessor für Systemvirologie an der Julius-Maximilians-Universität Würzburg berufen. Dort forschte er an der Genomik von Herpesviren, an kryptischen Translationsmechanismen und der Präsentation von Produkten kryptischer Translation via HLA-I, sowie an der Dynamik von genregulatorischen Prozessen unter akuten Perturbationen von Zellpopulationen oder auch einzelnen Zellen. In seiner Arbeitsgruppe wurden zu diesem Zweck mehrere maßgeschneiderte bioinformatische Methoden neu entwickelt, die auch breite Anwendung außerhalb seiner eigenen Arbeit finden. Florian Erhard ist darüber hinaus Mitglied in mehreren Forschungsverbänden wie dem durch die EU geförderten VIROINF, dem bayerischen Corona-Konsortium FOR-COVID, den DFG Forschungsgruppen DEEP-DV und „Fortschrittliche Konzepte in der zellulären Immunkontrolle von Zytomegalieviren“, und den Sonderforschungsbereichen DECIDE und „Cardioimmune interfaces“. Seit März 2023 ist er nun Inhaber des Lehrstuhls Computational Immunology der Universität Regensburg.