

HEIDELBERGER
JAHRBÜCHER
ONLINE
Band 5 (2020)

Gesellschaft der Freunde
Universität Heidelberg e.V.



Entwicklung – Wie aus Prozessen Strukturen werden

Joachim Funke & Michael Wink (Hrsg.)

HEIDELBERG
UNIVERSITY PUBLISHING

Entwicklung – Wie aus Prozessen Strukturen werden

Heidelberger Jahrbücher Online
Herausgegeben von der
Gesellschaft der Freunde Universität Heidelberg e.V.
Band 5

**Entwicklung –
Wie aus Prozessen Strukturen
werden**

Joachim Funke & Michael Wink (Hrsg.)

HEIDELBERG
UNIVERSITY PUBLISHING

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet unter <http://dnb.dnb.de> abrufbar.



Dieses Werk ist unter der Creative Commons-Lizenz 4.0 (CC BY-SA 4.0) veröffentlicht. Die Umschlaggestaltung unterliegt der Creative-Commons-Lizenz CC BY-ND 4.0

Die Online-Version dieser Publikation ist auf den Verlagswebseiten von HEIDELBERG UNIVERSITY PUBLISHING (<https://heiup.uni-heidelberg.de>) unter der Rubrik Campus Media dauerhaft frei verfügbar (open access).

<https://doi.org/10.17885/heiup.hdjbo.2020.0>

Bildnachweis Cover: Alamy Limited, Bild-ID: J8AAMJ (lizenzfrei)

© 2020. Das Copyright der Texte liegt bei den jeweiligen Verfassern.

ISSN 2509-7822 (Print)

ISSN 2509-2464 (Online)

ISBN 978-3-96822-026-0 (Softcover)

ISBN 978-3-96822-027-7 (Hardcover)

ISBN 978-3-96822-025-3 (PDF)

Inhaltsverzeichnis

| | | |
|----------|--|-----------|
| 1 | Vorwort | |
| | Joachim Funke & Michael Wink | 1 |
| 2 | Fehlentwicklungen in der Bibelwissenschaft (Altes Testament): Eine sich kontinuierlich selbst falsifizierende Disziplin | |
| | Bernd J. Diebner | 5 |
| 1 | Kritische Ansätze | 11 |
| 2 | Stichwort „Quellenscheidung“ | 13 |
| 3 | Grundlegende Infragestellung gängiger Hypothesen | 22 |
| 4 | Schlussbemerkungen | 27 |
| 3 | Universalität der Bausteine und Prozesse des Lebens | |
| | Michael Wink | 31 |
| 1 | Einführung | 32 |
| 2 | Ursprung der Erde und des Lebens | 33 |
| 3 | Universalität der Moleküle des Lebens | 34 |
| 4 | RNA-Welt | 36 |
| 5 | Bakterien waren zuerst da | 37 |
| 6 | Von den Prokaryoten zu den Eukaryoten und mehrzelligen Lebewesen | 41 |
| 7 | Entwicklungslinien der Phylogenese | 46 |
| 8 | Ausblick | 48 |

| | | |
|----------|---|------------|
| 4 | Entwicklungen in der Entwicklung – Fortwährende Veränderungen im Fluss der Organismenwelt | |
| | Claudia Erbar & Peter Leins | 53 |
| 1 | Einführung, Begriffliches, Methodenentwicklung | 54 |
| 2 | Ontogenien im Blütenbereich | 62 |
| 3 | Phylogenien | 69 |
| 4 | Hologenie | 76 |
| 5 | Der genetische Ansatz der Blütenbildung | 80 |
| 6 | Ontogenien im unmittelbaren Zusammenhang mit einer Funktion | 85 |
| 7 | Abschließende Bemerkungen | 92 |
| 5 | Cancer in the light of evolution – from cancer genomes towards novel treatment and prevention approaches | |
| | Matthias Kloor & Magnus v. Knebel Doeberitz | 99 |
| 1 | General considerations – cancer as a genetic disease | 100 |
| 2 | Molecular classification of cancers | 101 |
| 3 | Understanding cancer as the result of an evolutionary process – Mutational signatures | 104 |
| 4 | Mismatch repair deficiency | 105 |
| 5 | MMR-deficient cancers – a model of tumor evolution | 107 |
| 6 | Random mutations and selection in MMR-deficient cancers . . . | 107 |
| 7 | The high immunogenicity of MMR-deficient tumors | 109 |
| 8 | Immune surveillance and immunoediting | 112 |
| 9 | Transmissible tumors – a special case of tumor evolution | 114 |
| 10 | Heterogeneity of MMR-deficient cancers | 117 |
| 11 | Host factors – towards early steps of tumor formation | 119 |
| 12 | Hereditary cancer and cancer prevention | 120 |
| 13 | Outlook | 122 |
| 6 | Von Ereignissen zu Strukturen | |
| | Hans J. Pirner | 131 |
| 1 | Feste und flüssige – gasförmige Strukturen | 134 |
| 2 | Den Tanz der Atome stoppen | 137 |
| 3 | Familienstrukturen der Elementarteilchen | 141 |
| 4 | Die lange dunkle Nacht des Universums | 145 |
| 5 | Metastrukturen | 148 |

| | | |
|----------|---|------------|
| 7 | Character formation from a psychological point of view: Search for values, search for “Sinn” | |
| | Joachim Funke | 155 |
| 1 | Introduction | 155 |
| 2 | What is character? | 156 |
| 3 | Classical (narrow) view: Progression through stages | 157 |
| 4 | Broader view: Character formation as personality development | 159 |
| 5 | Moral values: World studies | 160 |
| 6 | Character formation and the genesis of wisdom | 162 |
| 7 | Measuring character and moral values | 163 |
| 8 | Moral dilemmata in experimental research: Trolley experiments | 164 |
| 9 | Conclusion | 164 |
| | | |
| 8 | Vom Datenchaos zum strukturierten Modell – Wie sich Daten zähmen lassen | |
| | Christel Weiß | 171 |
| 1 | Einleitung | 171 |
| 2 | Klinische Studien: Der lange Weg des therapeutischen Fortschritts | 173 |
| 3 | Anwendungsbeispiele | 185 |
| 4 | Schlussfolgerungen | 190 |
| | | |
| 9 | Struktur und Chaos: Kleinskalige Austauschprozesse zwischen Atmosphäre und Meer | |
| | Bernd Jähne | 195 |
| 1 | Einleitung | 195 |
| 2 | Grenzschichten und Transfargeschwindigkeiten | 196 |
| 3 | Gasaustausch und Wind: Feldmessungen | 201 |
| 4 | Der Ozean im Labor | 204 |
| 5 | Bildaufnehmende Messtechnik | 206 |
| 6 | Auflösung der Grenzfläche bei Hurrikan-Windstärke | 212 |
| 7 | Fazit und Ausblick | 213 |
| 8 | Danksagung | 215 |

10 Entwicklung einer Pandemie: Psychologische Aspekte der Corona-Krise

| | |
|--|------------|
| Joachim Funke | 219 |
| 1 Einführung | 219 |
| 2 Situationsfaktoren: Eigenschaften eines komplexen Problems . . | 221 |
| 3 Personfaktoren: Interindividuelle Unterschiede | 226 |
| 4 Risikowahrnehmung und Risikomanagement | 227 |
| 5 Umgang mit Fehlern | 227 |
| 6 Beiträge verschiedener Teildisziplinen der Psychologie | 230 |
| 7 Ethische Aspekte | 235 |
| 8 Ist psychologische Forschung bereit zur Politikberatung? | 236 |
| 9 Abschluss | 238 |

Vorwort

Im vorliegenden fünften Band der *Heidelberger Jahrbücher Online* (HDJBO), den die „Gesellschaft der Freunde Universität Heidelberg e.V.“ unter Federführung der beiden Editoren Joachim Funke und Michael Wink herausgibt, haben sich die Autorinnen und Autoren des Bandes diesmal mit dem Konzept der Entwicklung und Strukturbildung in der Wissenschaft als fächerübergreifender Thematik auseinandergesetzt.

Unter dem Rahmentitel „Entwicklung – Wie aus Prozessen Strukturen werden“ geht es um Entwicklungsprozesse im Kleinen (Zellwachstum, Ontogenese, Phylogenese) und im Großen (vom Big Bang zum Universum), die zu Strukturen werden. Entwicklungen im Bereich der Politik (Entstehen und Vergehen von Abkommen im Bereich Handel, Nuklearwaffen, etc.), der Geographie (Plattentektonik), der Psychologie (Ontogenese), der Medizin (Entwicklung von Organen), der Biologie (Baupläne, Phylogenese, Tree of life), des Aufkommens und Verschwindens von Ideen (Philosophie: Trends und Hypes), der Aufstieg und Niedergang politischer Systeme (Geschichtswissenschaft, Politikwissenschaft), der soziale Auf- und Abstieg (Soziologie), sind allesamt Beispiele für diese strukturbildenden Prozesse. Natürlich könnten auch „Fehl“-Entwicklungen (z. B. in der Psychologie: Klinische Störungen, Entwicklungsstörungen; Medizin: Krebs; Soziologie: Korruption) dazu gerechnet werden.

Diese Thematik beleuchten wir im vorliegenden Band aus der bunten Sicht unserer Volluniversität. Zehn Autorinnen und Autoren aus Geistes-, Kultur-, Naturwissenschaften und der Medizin haben ihr Verständnis von strukturbildenden Entwicklungsprozessen aus unterschiedlichen Gesichtspunkten erörtert. Die durchaus disziplinär angelegten Beiträge thematisieren ganz unterschiedliche Aspekte des Rahmenthemas und erzielen damit am Ende eine interessante Perspektivenvielfalt.

Bernd J. Diebner (Altes Testament) schildert die Rekonstruktionshypothesen über das Zustandekommen des Alten Testaments und macht deutlich, dass zwar viele Hypothesen im Raum stehen, eine Falsifikation dieser Ideen aber kaum möglich scheint. Die teilweise phantasievollen Hypothesen werden verständlich vorgetragen und mit Vergnügen seziert.

Michael Wink (Biologie) betrachtet die Bausteine und universell verbreiteten biochemischen Prozesse des Lebens. Er beschreibt, wie alles Leben auf diesem Planeten auf einer „Urzelle“ beruht (die möglicherweise auch noch von „outer space“ kam). Aus jener Urzelle leiten sich die alle Pro- und Eukaryoten ab, die sich diversifizierten und zu den unterschiedlichen Entwicklungslinien führten. Diese Entwicklungsgeschichte liest sich wie ein Krimi.

Claudia Erbar (Biologie) und *Peter Leins* (Biologie) machen die verschiedenen „-genien“ (Ontogenie, Phylogenie, Hologenie) in der Organismenwelt als strukturbildende Prozesse deutlich. Illustriert mit vielen farbigen Abbildungen lernt man viel über die Entwicklung von Pflanzen, die so herrliche Namen wie z. B. „Johanniskrautblättrige Myrtenheide“ tragen.

Matthias Kloor (Angewandte Tumorbiologie) und *Magnus v. Knebel Doeberitz* (Molekulare Onkologie) machen sich für die evolutionäre Sichtweise auf Krebs stark. Wir lernen dort z. B. dass das Sticker-Sarkom auf eine mehrtausendjährige Geschichte zurückblickt und es einen Hundepatient Null (den „founder dog“) gibt, auf den die Krankheit zurückgeführt werden kann.

Hans J. Pirner (Theoretische Physik) beschreibt strukturbildende Prozesse auf atomarer und sub-atomarer Ebene. Zahlreiche Querverweise auf geistes- und sozialwissenschaftliche Parallelen machen den Text auch für Fachfremde lesenswert.

Joachim Funke (Allgemeine Psychologie) befasst sich mit psychologischen Erkenntnissen zur Entwicklung von Charakter und zum Entstehen von Werten. Trotz kultureller Unterschiede scheint es einen gemeinsamen Kern menschlicher Werte zu geben.

Christel Weiß (Medizin) macht im ersten Teil ihres Beitrags in Form von Empfehlungen deutlich, wie man vom anfänglichen Chaos bei der Suche nach empirischen Regelmäßigkeiten zu geordneten Strukturen fortschreiten kann. Im zweiten Teil gibt sie dann illustrative Beispiele für diesen Übergang.

Bernd Jähne (Umweltphysik) macht deutlich, wie wichtig der Austausch von Luft und Wasser an der Meeresoberfläche für unser Weltklima ist. Diffusionsprozesse und Turbulenzen beeinflussen den Transport von Gasen zwischen Meer

und Atmosphäre - Prozesse, die sowohl unter schwierigen Bedingungen im Feld wie auch unter Laborbedingungen (am Wind-Wellen-Kanal namens „Aeolotron“) untersucht werden.

In einem Kapitel, das ursprünglich nicht geplant war, befasst sich *Joachim Funke* (Allgemeine Psychologie) mit psychologischen Aspekten rund um die aktuelle Corona-Pandemie. Die dortige Entwicklung weist viele Bezüge zu seinem Heimatfach auf, auch wenn der Wissensstand noch eher aus Fragen als aus Antworten besteht.

Die Publikation als e-Book hat sich bislang bewährt: Sie spart Kosten und ermöglicht dank „open access“ eine größere Verbreitung als die Print-Version. Die ersten vier Bände (Stabilität: Wink & Funke, 2016, Citizen Science: 2017; Mobilität: Funke & Wink, 2018, Schönheit: 2019) haben seit ihrem Erscheinen 2016 fast 19.000 Downloads zu verzeichnen. Natürlich sind Download-Zahlen noch kein Indikator für breitere Wirkung, aber das waren die (deutlich niedrigeren!) Verkaufszahlen der alten Print-Ausgaben auch nicht. Auch diesem nun vorliegenden fünften Band der *Heidelberger Jahrbücher Online* wünschen wir daher angemessene Verbreitung!

Wem die digitale Ausgabe nicht genügt und ein Exemplar für seinen Bücher-schrank wünscht: Dank der guten Zusammenarbeit mit „Heidelberg University Publishing (HeiUP)“ kann von allen Bänden für kleines Geld eine Print-Version „on demand“ hergestellt werden. Wir bedanken uns für die wie immer harmonische Zusammenarbeit beim Team der Universitätsbibliothek unter Leitung von Frau Dr. Maria Effinger, aber auch beim Direktor der Universitätsbibliothek, Dr. Veit Probst, der diesen Weg digitaler Informationsverbreitung seit Jahren fördert, ohne die Print-Welt zu vernachlässigen. Ein besonderer Dank ist erneut unser Chefflektorin Julia Karl (M. Sc. Psychologie) auszusprechen, die in gewohnter Qualität und mit hoher Geschwindigkeit den Publikationsprozess befördert hat.

Wir sind gespannt, wie der neue Band ankommt und wie unser Jahrgangsthema aufgenommen wird. Feedback ist wie immer erwünscht!

Heidelberg, im Sommer 2020

Joachim Funke (Psychologie) und Michael Wink (Biologie)

Quellen

- Funke, J., & Wink, M. (Hrsg.). (2018). *Perspektiven der Mobilität* (Heidelberger Jahrbücher Online Band 3). Heidelberg University Publishing. <https://doi.org/10.17885/heiup.hdjbo.2018.0>
- Funke, J., & Wink, M. (Hrsg.). (2019). *Schönheit: Die Sicht der Wissenschaft* (Heidelberger Jahrbücher Online Band 4). Heidelberg University Publishing. <https://doi.org/10.17885/heiup.hdjbo.2019.0>
- Wink, M., & Funke, J. (Hrsg.). (2016). *Stabilität im Wandel* (Heidelberger Jahrbücher Online Band 1). Heidelberg University Publishing. <https://doi.org/10.17885/hdjbo.2016.0>
- Wink, M., & Funke, J. (Hrsg.). (2017). *Wissenschaft für alle: Citizen Science* (Heidelberger Jahrbücher Online Band 2). Heidelberg University Publishing. <https://doi.org/10.17885/heiup.hdjbo.2017.0>

Fehlentwicklungen in der Bibelwissenschaft (Altes Testament): Eine sich kontinuierlich selbst falsifizierende Disziplin

BERND J. DIEBNER

Wissenschaftlich-Theologisches Seminar, Universität Heidelberg

Zusammenfassung

Die Geschichte der historisch-kritischen Erforschung des Alten Testaments in den letzten drei Jahrhunderten bewegt sich im Spannungsfeld von kirchenamtlichen und fundamentalistischen Gemeinde-Dogmen und für ‚ohne Zweifel‘ als ‚sicher‘ geltenden Forschungs-Dogmen, die seit Jahrhunderten von neuen und einander widersprechenden ‚ohne Zweifel gesicherten‘ Forschungsdogmen abgelöst werden. Es ist ein von Hypothesen geprägter *circulus vitiosus academicus*. Falsifizierungen sind kaum von häufigen Paradigmenwechseln zu unterscheiden. Oft tritt dabei der Sprachgestus an die Stelle von (fehlenden) Sach-Argumenten.

Einführung

Die Geschichte der „Historisch-kritischen Erforschung des Alten Testaments“ der christlichen Bibel ist seit ihren methodisch relevanten Anfängen im frühen 18. Jahrhundert (Jh.) die Geschichte eines kontinuierlichen Selbstfalsifizierungs-Prozesses

angeblich ‚gesicherter‘ und als ‚ohne Zweifel‘¹ unumstößlich bezeichneter Ergebnisse². Dabei handelt es sich in der Forschungsgeschichte durchgehend bis heute um eine durch fundamentale hermeneutisch-methodische Neuansätze bedingte resultative Hinterfragung des Forschungsstandes bis jeweils dato, also um einen permanenten Paradigmenwechsel. Dieser bis heute nicht abgeschlossene Prozess lässt sich natürlich auch in anderen und nicht nur kulturwissenschaftlichen Disziplinen beobachten. Er bedeutet zum mindesten ein Fragezeichen hinter den Wissenschaftsanspruch, den die Vertreter der Disziplin behaupten.³

Es gibt m. E. Disziplinen, die durch ihre Bindung an bestimmte vorwissenschaftliche Traditionen besonders belastet sind⁴. Besonders retardierend wirkte und wirkt für einige universitäre Disziplinen die Bindung an die Kirche(n) mit ihren quasi-dogmatisierten Meinungen über die in der Welt herrschenden göttlichen Gesetze, Grundsätze und Wahrheiten, die in der „Heiligen Schrift“ zu finden seien. Zu den hierdurch besonders betroffenen Disziplinen gehört u. a. auch die Astronomie⁵, besonders aber verständlicherweise die Theologie mit ihren historischen Sub-Disziplinen „Kirchengeschichte“ und „Exegese“, d. h. : in ihrer methodisch-kritischen Interpretation (Auslegung) der biblischen Schriften-Korpora Altes und Neues Testament. – Bezüglich des Neuen Testamentes (NT) ist das Konfliktpotenti-

¹ Bis heute üblicher Sprachgestus in einer Disziplin, die von Hypothesen lebt.

² Noch 1961 riet der damals bereits arrivierte Assistent Rudolf Smend (*1932) den Wissenschaftlichen Hilfskräften seines Bonner Chefs Martin Noth (1902–1968): „Werden Sie bloß keine Alttestamentler! Da ist bereits alles erforscht!“.

³ Dabei möchte ich nicht so weit gehen wie Karl Raimund Popper (1902–1994), der Begründer des „kritischen Rationalismus“, der eigentlich nur den naturwissenschaftlichen Disziplinen den Anspruch auf ‚Wissenschaftlichkeit‘ zubilligte. Bei diesen lässt sich das Falsifizierungs-Prinzip methodisch einfacher und somit leichter realisieren. ‚Falsifizierung‘ ist bei Popper gegenüber der viel bequemerem ‚Verifizierung‘ das ‚höhere‘ Forschungsprinzip; denn nur, was sich nicht falsifizieren lässt, darf den Anspruch auf ‚Wissenschaftlichkeit‘ erheben.

⁴ Die Bindung an vorwissenschaftliche Traditionen und an respektvoll behandelte althergebrachte Forschungsmeinungen und Überzeugungen sind Klötze am Bein für jede Forschungs-Disziplin.

⁵ Ich verweise für diese komplexe Disziplin nur auf Nikolaus Kopernikus (1473–1543): Indekongregation 1616 mit dem Verbot seiner Lehre, und auf Galileo Galilei (1574–1641/2), der erstmalig 1616 mit der römischen Kurie zusammenstieß und von Kardinal Bellarmino aufgefordert wurde, seine ‚Irrtümer‘ (typische klerikale Kategorie) aufzugeben, was Galilei auf weiteren Druck der Kirche hin dann 1632 wider besseres Wissen auch tat. Am 2. November 1992 (!) wurde Galilei von der römischen Kirche förmlich-rechtlich rehabilitiert (vgl. *Wikipedia*: Art. „Galileo Galilei: Galilei und die Kirche“).

al zwischen der Kirche und ihren Dogmen und der historisch-kritischen Forschung (HKF) zwar sehr groß⁶, aber für seriöse Datierungsvorschläge bezüglich der einzelnen Schriften des NT gibt es nur wenig Spielraum⁷. – Ganz anders verhält es sich mit den Exegeten-Meinungen zum Alten Testament (AT)⁸. Für diese Schriften wird hypothetisch – zumal im neueren Forschungsspektrum – ein Zeitraum von bis zu 1000 Jahren zwischen der frühesten und der spätesten Niederschrift biblischer Texte angenommen. Grund hierfür ist einmal ein durch die anfängliche Bindung an die kirchliche Tradition bedingter allmählicher Abstraktionsprozess von den kirchlichen, quasi-dogmatischen ‚Wahrheiten‘ im Rahmen einer bis zum heutigen Tage argumentativ nicht zu zügelnden Meinungswillkür. Die vorkritische Meinung der kirchlichen Tradition⁹ darf jeder, der es glaubt, bis heute (mit welchen Argumenten auch immer) und wider alle kritischen Hypothesen vertreten, die diese Meinung seit nunmehr bald 300 Jahren zu falsifizieren suchen. Dies betrifft die sog. „Fundamentalisten“, heute hauptsächlich vertreten durch die „evangelikalen Freikirchen“¹⁰. Dieser „Fundamentalismus“ wird an den Predigerseminaren und

⁶ So ist z. B. die physische (leibliche) Auferstehung Jesu Christi von den Toten zwar ein kirchliches Dogma, das in den Symbolen (Bekenntnissen) der Kirchen fest verankert ist, sich aber nicht argumentativ verplausibilisieren lässt. Wissenschaftlich beurteilt ist die Auferstehung ein religionsphänomenologisch häufig belegter religiöser (überhöht formuliert: theologischer) Mythos.

⁷ Theoretisch setzen die frühesten NT-Schriften den Tod Jesu (um 30 nC) voraus, und spätestens um die Mitte des 2. Jh. nC müsste der Prozess abgeschlossen worden sein. Ab etwa 125 nC setzen die Manuskript-Funde ein, beginnend mit einem Fragment aus dem (späten: um 100 nC) Evangelium des Johannes (P 52), das nach der *communis opinio* der Neutestamentler das jüngste der vier kanonischen Evangelien ist. Eine abweichende Meinung vertritt der Heidelberger NTler Klaus Berger (*1940).

⁸ Das bei den Reformations-Kirchen anders definiert wird als in der römischen Kirche und in den orthodoxen Kirchen. Martin Luther hatte Schriften und Textelemente ausgeschieden, deren hebräischen Text er nicht kannte, und unter dem mehrdeutigen Begriff „Apokryphen“ versammelt. Ein historisches Fehlurteil, wie hebräische Textfunde seit dem 19. Jh. zeigen sollten.

⁹ Etwa dass Moses die fünf mit seinem Namen verbundenen Schriften im 13. Jh. vC eigenhändig niedergeschrieben habe.

¹⁰ In jüngerer Zeit erhielten diese Kirchen und christlichen Gemeinschaften einen weltweiten Impuls durch den bekannten US-amerikanischen Baptisten-Prediger Billy Graham (1918–2018). Ihn lernte ich indirekt seit 1953 kennen, als ich meine schulischen Sommerferien regelmäßig in einer englischen Baptisten-Familie verbrachte, und direkt 1970 als Hamburger Pastor, als Graham in Hamburg evangelisierte. Die Redaktion der damaligen Hamburger Kirchenzeitung bat je einen Pastor „Pro“ bzw. „Contra“ zu schreiben. Ich wurde als ‚linker‘ Pastor um das „Contra“ gebeten.

kirchlichen Hochschulen der Freikirchen vertreten¹¹. Aber auch akademisch angesehene Forscher dürfen ähnliche, in kritischen Sprachgestus gekleidete Meinungen bis heute vertreten¹². Es besteht eine allgemeine, unkontrollierte und unkontrollierbare akademische ‚Meinungsfreiheit‘. Sie hat gegenüber manchen anderen Disziplinen den Vorteil, nicht allgemeingefährdend zu sein. Man könnte es auch eine gewisse Beliebigkeit nennen, von der besonders die arrivierten und etablierten Fachvertreter profitieren¹³.

Wie schon angedeutet, Verifizierung (Gründe *für* etwas zu finden) ist oft leicht. Falsifizierung (stichhaltige Gründe zu finden, die *gegen* etwas sprechen) ist oft sehr schwer. Und – wie auch bereits angesprochen – niemand *muss* sich in der hier zu thematisierenden Disziplin „Erforschung des AT“ einem Falsifikationsversuch anschließen, auch dann nicht, wenn er Anderen plausibel erscheint. Dies bedingt in der HKF des AT schon recht früh eine ‚Deltaisierung‘ des Forschungsstromes. Nachdem sich der Strom teilte, flossen mehrere Arme getrennt in das ‚Meer der *einen* Wahrheit‘.

Um es für die fachfremden Leser/innen nicht zu kompliziert zu machen, beschränke ich mich auf drei maßgebliche Umbrüche in der Interpretation des AT seit dem 18. Jh., die je einen grundsätzlichen Paradigmenwechsel und somit eine (bedingte) Falsifizierung voraufgehender Betrachtungsweisen des AT bedeuten.

Kritische Hypothesen zur Entstehung und Gestaltwerdung der Schriften des AT nehmen traditionellerweise und bis heute meist ihren Ausgang bei der Erforschung

¹¹ Verteufelt wurde in englischen fundamentalistischen Baptistenkreisen die *Bristol Critical Baptist Bible School*, die sich für kritische Exegeten-Meinungen offen zeigte und schon in den 50er Jahren die Texte von Rudolf Bultmann (1884–1976) berücksichtigte.

¹² Wie z. B. der verdienstvolle französische Dominikaner-Pater und Professor Roland Guérin de Vaux (1903–1971), Qumranforscher und Bibelwissenschaftler.

¹³ Anders verhält es sich bei den Studierenden. In meiner langjährigen Tätigkeit als Studienberater der Theologischen Fakultät Heidelberg war meine Standardfrage an ratsuchende Examenskandidat/inn/en: „Wer wird Sie im Alten Testament (oder anderen Fächern) prüfen?“ (Egal ob in Heidelberg oder anderswo.) Konnten die Kandidaten mir einen Namen nennen, so empfahl ich den Kandidaten gezielt, was sie zu lesen hätten, um zu bestehen. Meine Maxime: „Prüfungen sind kein Zeitpunkt für Heldenmut! Prüfungen sind dazu da, bestanden zu werden!“

der fünf Bücher Moses, in der Forschung meist durchgängig als „Pentateuch“ bezeichnet¹⁴. Die Mosesbücher werden wohl als ‚Basis‘ des AT empfunden¹⁵.

Hier die m. E. wesentlichen Umbrüche:

- (1) **Kritische Ansätze im 18. Jh.:** Infragestellung (und damit Falsifizierung) traditioneller kirchlicher Meinungen und Glaubenssätze über die Entstehung des Pentateuch¹⁶.

¹⁴ Griech. *το πεντατευχος* = „das Fünferbuch“.

¹⁵ Möglicherweise hat dies ein jüdisches Erbe mit beeinflusst: die fünf Mosesbücher (hebr. *thorah* = „Lehre, Weisung“ genannt) sind für die Juden (und Samaritaner) mit ihren angeblich 613 Geboten (nachgezählt und mit Kriterien überprüft, was ein ‚Gebot‘ sei, hat wohl noch niemand) „Bibel“ im ersten und kanonisch eigentlichen Sinne. Die „Propheten“ sind nur Predigten für Laien-Vorleser in der Gemeinde (vgl. Lk 4,16–20), der dritte Teil der *Biblia Hebraica et Aramaica* sind nur (weitere religiöse) „Schriften“ (hebr. *kh:tuvim*), von denen in der Antike nur der Psalter in der synagogalen Liturgie eine Rolle spielt (vgl. Lk 24,44).

¹⁶ Der „Pentateuch“ wurde ursprünglich auch als „Septuaginta“ („Buch der 70“) bezeichnet, weil der Legende nach angeblich 72 jüdische Gelehrte im Auftrag des hellenistischen Diadochen-Pharaos, des Ptolemäers Ptolemaios II., um die Mitte des 3. Jh.s vC die hebräische Thorah ins Griechische übersetzt hätten. So der früher in diese Zeit datierte und heute sog. „Aristeasbrief“ / PsArist. Diese Datierung wurde erstmalig 1900 falsifiziert durch Paul Wendland (1864–1915) und später (1929) durch Elias Bi(c)kerman(n) (1897–1981). Wendland beansprucht für die Abfassung des PsArist mit guten Gründen den Zeitraum zwischen 96 und 67 vC. Bickermann ist etwas großzügiger und nimmt die Zeit um 130 vC an. Zitiert wird heute meist Bickermann. Obwohl Bickermanns Meinung gilt, nimmt die HKF noch immer die Mitte des 3. Jh. vC als Zeitraum für die Übersetzung der Thorah ins Griechische an. Das hat – neben Anderem – bis heute Folgen für die Datierung der literarischen Endgestalt des Pentateuch: das Fundament ist weg, aber das Schloss bleibt in der Luft hängen. – Eine deutsche Übersetzung des PsArist ist leicht verfügbar: vgl. N. Meisner (Hrsg.): *Aristeasbrief*. JSHRZ III/I, Gütersloh 1973, S. 35–87. – In der weitschweifigen gelehrten Erzählung kommt die Übersetzung der jüdischen Gesetze aus dem Hebräischen ins Griechische nur gegen Ende kurz zur Sprache. Viel wird über die Vorzüge dieses jüdischen Gesetzes debattiert. Es wird vor dem Ptolemäer-Pharao Ptolemaios II. gepriesen. Dabei kommen nur nomistische Traditionen in Zitaten oder Anspielungen aus den Büchern 2., 3. und 5. Moses - Exodus („Bundesbuch“), Levitikus und Deuteronomium - zur Sprache. Es wird auf keine der für protestantische Exegeten so wichtigen, weil so ‚evangelischen‘ „Väter-Geschichten“ angespielt, auch der so wichtige Gründungsmythos Israels vom „Auszug aus Ägypten“ kommt nicht zur Sprache. Auch keiner der so wichtigen „Väter“-Namen Abraham, Isaak, Jakob und Joseph wird erwähnt. Einzig „Moses“ wird in § 144 als Garant des jüdischen Gesetzes explizit genannt. Es ist im PsArist auch keine Rede davon, dass die 72 jüdischen Gelehrten ‚fünf Bücher‘ (Moses) übersetzt hätten. Das Wort *πεντατευχος* wird nicht erwähnt. Der Autor des PsArist scheint – egal wie wir nun den PsArist datieren (ob unkritisch ins 3. Jh. oder mit Bickermann

- (2) **Stichwort: „Quellenscheidung“:** Der permanente Paradigmenwechsel im 19. Jh, überwiegend bezüglich des Alters, der Zusammensetzung (literarischen Komposition) und der Verfasserschaften des Pentateuchs (aber auch anderer Schriften der *Biblia hebraica et aramaica* (BH) betreffend).
- (3) **Die hermeneutisch betrachtet grundlegende Infragestellung** traditioneller methodischer Zugänge zum Verständnis des AT, besonders auch der Datierung alttestamentlicher Traditionen und Schriften, in mehreren Anläufen datierend seit den 60er Jahren des 20. Jh.

Ich bin hier angesichts eines kaum überschaubaren Forschungsgebietes um Übersichtlichkeit bemüht. Das führt natürlich auch zu Verkürzungen. Dieser Beitrag kann kein forschungsgeschichtliches Kompendium ersetzen. Darum verweise ich hier auf einige neuere Kompendien und kompendienartige unselbständige Übersichten über den gesamten Forschungsbereich und einzelne wesentlichere Aspekte¹⁷. Auf einige Publikationen stütze ich mich stillschweigend, aber nur

und Wendland ins 2. bzw. 1. Jh.) – nur eine rein nomistische Thorah gekannt zu haben, ein reines Gesetzescorpus. Dabei hätten die jüdischen Übersetzer doch so schön die etwa den „Vätern“ Israels gewährte Gnade Gottes vor dem Pharao preisen können. Der PsArist lässt alle bisherigen Pentateuch-Hypothesen zum mindesten als problematisch erscheinen, auch die gern gepflegte Meinung, dass die Gesetzespartien der Thorah erst ‚sekundär‘ in einen älteren narrativen (und evangelisch-erbaulichen) literarischen Kontext eingefügt worden seien.

¹⁷ Einige selbständige Publikationen: H.-J. Kraus: *Geschichte der historisch-kritischen Erforschung des Alten Testaments*. (1. Aufl. 1956; 4. Aufl. Neukirchen 1988 (um eine „Ergänzung“ [§§ 103 – 105] bereicherte, ansonsten unveränd. Fassung der 3. Aufl. 1982; in der Ergänzung berücksichtigt Kraus einige Neuansätze seit der 2. Hälfte des 20. Jh., aber keine fundamentalen; zitiert: Kraus / 1988); H. Graf Reventlow: *Episoden der Bibelauslegung*. München, Bd. 1 1990; Bd. 2 1994; Bd. 3 1997; Bd. 4 2001; W. H. Schmidt, W. Thiel, R. Hanhart: *Das Alte Testament*. UTB 421. Stuttgart 1989; R. Smend: *Deutsche Alttestamentler in drei Jahrhunderten*. Göttingen 1989 (zitiert: Smend 1989); ders.: *Epochen der Bibelkritik*. Ges. Studien Bd. 3. München 1991. – Einige unselbständige Publikationen: E. Ruprecht: „Die Religion der Väter: Hauptlinien der Forschungsgeschichte“, in: *DBAT* 11 (1976), S. 2–29; E. Otto: „Stehen wir in einem Umbruch in der Pentateuchkritik?“, in: *VF* 22 (1977), S. 82–97; B. J. Diebner: „Neue Ansätze in der Pentateuchforschung“, in: *DBAT* 13 (1978), S. 2–13 (zitiert: Diebner 1978); ders.: „Neu erforschter Pentateuch“, in: *LM* 18 (1979), S. 372–373; Kl. Spronk: „Biblische Theologie in den Niederlanden“, 2009; leichtest zugänglich bei *Wikipedia sub voce*: „Klaas Spronk“ (stellt die „Amsterdamse School“ mit ihrem Hauptakteur Karel A. Deurloo [1936–2019] vor; dazu in Abschnitt 3). – Enzyklopädische Artikel: *TRE* Bd. VI, Lfg. 1 / 2; 3 / 4, Berlin, New York 1980, Art. „Bibelwissenschaft: I. Altes Testament:

wörtliche Zitate merke ich als solche in Fußnoten an. Ansonsten beschränke ich mich auf Verweise („vgl.“).

1 Kritische Ansätze

Es wäre unangemessen zu behaupten, es habe vor dem Zeitalter der Aufklärung (*illumination*) keine ‚kritische‘ Betrachtung der Bibeltexte gegeben. Nur wurden die Bibeltexte nicht unter *historisch-kritischem* Aspekt untersucht. Es wurden also keine Hypothesen erstellt über die mögliche Entstehungsgeschichte der Schriften. Es wurden die Etiketten der Schriften nicht in Frage gestellt. Es wurde nicht gefragt: „Stammen die Mosesbücher tatsächlich von einem als historisch angenommenen Moses? Stammt das Buch Jesaja wirklich oder wesentlich von dem für das 8. Jh. vC angenommenen historischen jüdischen Propheten?“ – Rogerson rechnet zurecht mit „Wegbereitern der modernen Bibelwissenschaft 1577–1770“¹⁸. Dabei lässt sich die „Textkritik“, also die Frage nach ‚ursprünglichen‘ Lesarten, angesichts der Fülle (oft nur minimal) voneinander abweichender Manuskripte (abgekürzt: MSS) bis in die Antike zurückverfolgen¹⁹. Die protestantische Orthodoxie (spätes 16. bis zum 18. Jh.) verhinderte mit ihrer rigiden göttlichen Inspirationslehre (Verbalinspiration) jede kritische Hinterfragung des Textbestandes. Ja, das ging soweit, dass sie diese (gelehrt war man ja) auch auf die interlinearen Punktierungen²⁰ des hebräischen Textes bezog. Diese waren ein Produkt jüdischer Gelehrter des 4. bis 9. Jh. nC, um die tradierte Aussprache des Hebräischen in einer Zeit zu bewahren, zu der es längst nur noch Literatursprache

1 / 2: Geschichte und Methoden“; J. W. Rogerson: „Alte Kirche“ bis „Vergleichende Kultur- und Religionsgeschichte“, S. 346–361 (= Rogerson 1980); B. J. Diebner: „Entwicklungen und Tendenzen in der jüngsten Vergangenheit“, S. 361–374 (= Diebner 1980); *RGG* 4. Aufl., Bd 1, Tübingen 1998, Art. „Bibelwissenschaft“; E. Otto: *I. Altes Testament*, Sp. 1517–1527. – Zur knappen Gesamtübersicht über den Forschungsstand des *mainstreams* der AT-Forschung gegen Ende des 20. Jh. vgl. B. Lang: *Die Bibel*. UTB 1594, Paderborn 1990 (mit „Trends in der gegenwärtigen Bibelwissenschaft“); Chr. Levin: *Das Alte Testament*. 4. Aufl. München 2010.

¹⁸ Rogerson 1980, S. 350.

¹⁹ Origenes (ca. 185–253/4 nC) stellte aus ihm verfügbaren MSS verschiedener Versionen (Übersetzungen) eine sechs- und eine vierspaltige Edition zusammen: „Hexapla“ bzw. „Tetrapla“ (nur durch Zitate bei Kirchenschriftstellern fragmentarisch rekonstruierbar).

²⁰ Eine sekundäre Hilfsvokalisation, auch Betonungszeichen und Satztrenner supra- und infralinear, weil man in den ‚heiligen‘ Konsonantentext nicht eingreifen wollte.

war. Ausgerechnet oder besser: verständlicherweise ein katholischer Priester, Richard Simon (1638–1712)²¹, stellte mit seiner Schrift *Histoire critique du Vieux Testament* (1678) das protestantische *sola scriptura* wegen der Varianten und Schreibfehler in den tradierten Bibel-MSS in Frage. Simon rechnete auch mit Schreiberschulen, womit er die eigenhändige Autorenschaft von Moses, Josua usw. bestritt. Für die Geschichtsbücher rechnete er damit, dass sie lange nach den geschilderten Ereignissen verfasst worden seien. Man darf Simon daher mit guten Gründen als einen „(Mit-)Begründer²² der historisch-kritischen Methode“²³ betrachten. Allerdings wurden viele kritische Schriften Simons indiziert (kirchlich verboten). – Man kann sagen, mit der neuzeitlichen Aufklärung begann in der Erforschung des AT die erste Falsifikationswelle, und zwar mit der Infragestellung des Verbalinspirations-Dogmas und der Bestreitung der ‚Echtheit‘²⁴ angegebener Verfassernamen biblischer Schriften. Verstärkt wurde diese Hinterfragung im 18. Jh. noch durch die ersten definierten Hypothesen über die Entstehung des Pentateuchs. Einige Forschungsgeschichtler setzen erst hier – mit der Beschreibung der ersten Quellen-Hypothesen – den Beginn der historisch-kritischen Erforschung des Alten Testaments an. Im Grunde bedeuten aber schon – wie gesagt – die ersten Ansätze kritischer Betrachtung des Bibeltextes einen ersten Falsifikationsschritt, auch wenn dieser vermutlich mit Rücksicht auf die mächtige kirchliche Dogmen-Tradition als solcher nie *expressis verbis* formuliert wurde: nämlich die faktische Aufgabe des Dogmas der Verbalinspiration der biblischen Texte bereits durch Simons Beobachtung der Manuskript-Varianten; denn welche konkrete Variante sollte im Gegensatz zu den anderen die göttlich inspirierte sein?²⁵

²¹ Vgl. zu Simon: Kraus 1988, S. 65–70.

²² Hier ist vor Anderen auch noch Johann le Clerc (1657–1712; mehrere Schreibweisen) zu erwähnen. – Es herrscht bei den Forschungsgeschichtlern keine Einigkeit, wann, wo und mit wem sie die Geschichte der HKF beginnen lassen sollen; vgl. auch Smend 1991, S. 253. Bei Smend finden sich auch sehr detailliert weitere ‚Väter‘ der historisch-kritischen Methoden (HKM).

²³ *Der Große Brockhaus* Bd. 10, Wiesbaden 1956, S. 722 r; vgl. auch *Meyers Enzyklopädisches Lexikon* Bd. 22, Mannheim usw. 1977, S. 742 l.

²⁴ Ein urmythischer Terminus in der HKF bis heute.

²⁵ Eine eigenartige Variante des Verbalinspirationsdogmas stellte der erkonservative Hamburger Theologe Pastor Dr. Dr. habil. Helmut Echternach (1907–1988) 1936 auf: die Lehre von der Verbalinspiration der Lutherübersetzung letzter Hand (1545).

2 Stichwort „Quellenscheidung“

Seit dem Beginn des 18. Jahrhunderts und bis weit ins 20. Jahrhundert²⁶ bestimmten im Blick auf den Pentateuch²⁷, der (wie gesagt) stets Ausgangstext für Umbrüche in der Erforschung des Alten Testaments war und ist, „Quellen-“, „Urkunden-“ oder „Dokumenten-Hypothesen“ das Forschungsbild. Dabei ist zu sagen, dass (historisch gesehen) der Begründer einer Pentateuch-Quellenhypothese erst 1924 wiederentdeckt wurde. Es ist der Hildesheimer (evangelische) Pfarrer Henning Bernhard Witter (1682–1715)²⁸. Witter wurde von der frühaufklärerischen Bibelkritik beeinflusst, u. a. von Ioannes Clericus. 1711 publizierte Witter ein Opus mit einem acht Zeilen langen barocken lateinischen Titel, den ich nur teilweise zitiere²⁹. Witter gab zwar grundsätzlich die Autorenschaft Moses nicht auf, meinte aber, dass Moses zwei Quellenschriften verschiedener Autoren zusammengearbeitet habe, die er durch eine unterschiedliche Gottesbezeichnung unterschied. Dies führte er nur an Gen 1–3 aus, den Schöpfungs- und der Paradiesgeschichte(n). Hier unterschied er Gen 1,1 – 2,3 von Gen 2,4 – 3,24³⁰. Witter rechnete auch bereits mit einer mündlichen Tradition – *traditio oralis* – vor der Verschriftlichung der frühesten Dokumente. Auch damit war er seiner Zeit um

²⁶ Und – weil man keine Forschungsmeinung verbieten kann – von einigen Fachvertretern auch noch danach.

²⁷ „Pentateuch“, das „Fünferbuch“, hebräisch die *Thorâh* (= „Lehre“, „Weisung“), deutsch die „Fünf Bücher Moses“.

²⁸ Vgl. A. Lods: „Un précurseur allemand de Jean Astruc: Henning Bernhard Witter“, in: ZAW 43 (1925), S. 134f.

²⁹ Vgl. H. B. Witter: *Iura Israelitarum in Palaestinem terram Chananaeam Commentatione in Genesin perpetua sic demonstrata ...* (auf den ganzen Pentateuch angelegt, aber fragmentarisch: veröffentlicht nur die Prolegomena und der Kommentar zu Gen 1–17); vgl. J. Chr. Gertz: Art. „Witter“, in: RGG (4. Aufl.) Bd. 8, Tübingen 2005, Sp. 1668.

³⁰ Diese Abgrenzung wurde von den Alttestamentlern grundsätzlich bis heute beibehalten, nur dass gegenwärtige Forscher Gen 1,1 – 2,4a von Gen 2,4b – 3,24 abgrenzen, Den halben Paragraphen 2,4a also zum voraufgehenden Text rechnen (weil wie Gen 1,1 – 23 der „Priesterschrift“ (P) zugerechnet): fälschlicherweise, wie als Erster K. A. Deurloo (1936–2019; Amsterdam) zeigte. Vers 2,4a ist – wie Analogien belegen – die Überschrift zum folgenden Text. Bereits Witter, der „P“ noch nicht kannte, schien damit gerechnet zu haben.

rund 200 Jahre voraus³¹. Witter gab die quasi-dogmatisierte Lehre von der Verbalinspiration zwar nicht auf, schrieb die Inspiration jedoch dem Quellensammler Moses zu. Dies brachte ihm Ärger mit der Wittenberger (lutherischen) Orthodoxie.

Witters Werk wurde im 18. Jahrhundert viel diskutiert, geriet dann aber unter den Schatten des Leibarztes Ludwigs XIV., Jean Astruc (1684–1766), und damit in Vergessenheit. Astruc entwickelte eine sehr ähnliche Quellenhypothese wie Witter. Man könnte sagen: er entwickelte Witters Hypothese fort, allerdings, ohne diesen und sein Werk zu kennen. Neben vielen medizinischen Publikationen brachte Astruc im Jahre 1753 sein bibelwissenschaftliches Hauptwerk heraus³². Astrucs Hypothese kann hier nur grob umrissen werden. Astruc möchte grundsätzlich an Moses Verfasserschaft des Pentateuch festhalten. Wie schon Witter unterschied Astruc für Gen 1–3 am Gottesnamen „Elohim“ (Gen 1) bzw. „Jehova“ (Gen 2–3) – damals übliche und fälschliche³³ Lesung des Tetragramms YHWH – zwei *mémoires* („Gedenkschriften“), die er *mémoire A* bzw. *mémoire B* nennt. Allerdings kann er unter diesen beiden Siglen nicht alle Genesis-Texte unterbringen. Darum sammelt er diesen in sich heterogenen Rest als *mémoire C* zusammen und verteilt ihn auf acht³⁴ Kolumnen. Damit, meint Astruc, habe er die gesamte vormosaische Überlieferung des Buches Genesis erfasst. Spätere unbekanntere Redaktionen hätten die Kolumnen ineinander geschoben, was die Chronologie der Schrift durcheinander gebracht habe. Astruc möchte die mosaische Verfasserschaft (erst nur) des Buches Genesis (1. Moses) gegen prominente Bestreiter wie Hobbes und Spinoza verteidigen, wenn auch nicht als Autor, so doch als Redakteur.

Astrucs Arbeit wurde anfangs nur verhalten, später aber mit Interesse aufgenommen. Besonders eingewirkt hat Astruc auf Theologen³⁵ wie Johann David Mi-

³¹ Erst im 20. Jahrhundert wurde die sog. „Überlieferungsgeschichte“ als exegetische Subdisziplin entwickelt, die sich mit der hypothetischen vorschriftlichen Tradition angeblich später verschrifteter biblischer Texte befasste.

³² Vgl. J. Astruc: *Conjectures sur les mémoires originaux dont il paroît que Moÿse s'est servi pour composer le Livre de Gènesè* (in deutscher Übersetzung 1783 erschienen).

³³ Von der christlichen Randgruppe „Jehovas Zeugen“ bis heute beibehalten, aus evangelischen Gesangbüchern aber schon vor einiger Zeit getilgt.

³⁴ So Kraus 1988, S. 96; vier Kolumnen nach J. Chr. Gertz: Art. „Astruc“, in: *RGG* (4. Aufl.) Bd. 1, Tübingen 1998, Sp. 863.

³⁵ Im 18. und 19. Jahrhundert vertraten die Theologieprofessoren stets mehrere Disziplinen wie Philosophie, Orientalistik und Bibelwissenschaft. Noch im 19. Jahrhundert unterschied man noch nicht Alt- und Neutestamentler. Man befasste sich mit der gesamten Bibel. Erst im 20. Jahrhundert

chaelis (1717–1791), Johann Salomo Semler (1725–1791) und besonders Johann Gottfried Eichhorn (1752–1827). Eichhorn wandte die sog. „**Ältere Urkundenhypothese**“ auf den gesamten Pentateuch an. Damit gilt er für einige heutige Autoren als der „Begründer der modernen bibl[ischen] Einleitungswiss[enschaft]“³⁶. Sein Entwurf wird auch gern als „**Ältere Fragmenten-Hypothese**“ bezeichnet, weil er mit Textteilen des Pentateuchs rechnet, die er nicht den beiden Hauptquellen zurechnen konnte.

Darüber, wann wer welchen Aspekt der kritischen Bibelwissenschaft begründet habe, gibt es bei den Autor/inn/en, die sich mit der Forschungsgeschichte und ihren historischen Repräsentanten befassen – wie bereits angedeutet – die unterschiedlichsten Meinungen. Man könnte die gesamte bislang beschriebene Entwicklung bereits als den Beginn der historisch-kritischen Bibelforschung betrachten. Für die Heidelberger Professorin für Systematische Theologie und Ökumene Friederike Nüssel beginnt „die moderne hist[orisch]-krit[ische] Exegese“ mit dem theologischen Allround-Wissenschaftler Semler. In seiner *Abhandlung von freier Untersuchung des Kanon* (1771–1775) „kritisiert er die Gleichsetzung von Schrift und Offenbarung bzw. von Schrift und Wort Gottes und konzentriert die exegetische Arbeit auf die Textkritik und die Aussageintention des Autors. „Damit“, meint Nüssel, „legt er den Grundstein für die moderne hist[orisch]-krit[ische] Exegese“³⁷. Nüssel ist insofern zuzustimmen, als der (rationalistische) Neologe Semler den biblischen Kanon einer konsequenten Kritik unterwirft: „Ohne Rücksicht auf das orthodoxe Inspirationsdogma werden die Grundregeln der ‚*critica profana*‘ in die Bibel eingeführt“³⁸. Mit Semler geschieht – nicht im Sinne einer methodischen philosophischen Falsifikation, aber – der rücksichtslose³⁹ Bruch mit der kirchlichen Tradition, was erstaunlicherweise in der späteren Forschung kaum mehr explizit reflektiert, sondern mehr als ‚stille Post‘ aus der Vergangenheit der Disziplin weitergetragen wird.

wurde auf Grund der Forschungs-Diversifikation eine engere Fachspezialisierung üblich, im Rahmen der Theologie u. a. die Unterscheidung von Alt- und Neutestamentlern.

³⁶ Chr. Frevel: Art. „Eichhorn“, in: *RGK* (4. Aufl.) Bd. 2, Tübingen 1999, Sp. 1119.

³⁷ Zitate: Fr. Nüssel: Art. „Semler“, in *RGK* (4. Aufl.) Bd. 7, Tübingen 2004, Sp. 1204.

³⁸ Kraus 1988, S. 107.

³⁹ Im Sinne von „ohne Rücksicht auf traditionelle kirchlich-dogmatische Bindungen“.

Das 19. und das 20. Jahrhundert sind geprägt durch eine ganze Reihe miteinander konkurrierender und somit einander falsifizierenden Pentateuch-Hypothesen⁴⁰. Die Exegese des Alten Testaments ist im 19. Jahrhundert geprägt durch den Ausbau vieler Varianten der „**Urkunden-Hypothese**“ bis hin zu ihrer quasi-dogmatischen Ausformung für etwa 100 Jahre durch den Göttinger Bibelwissenschaftler Julius Wellhausen (1844–1918)⁴¹.

Ein wichtiger Schritt hin zur Entwicklung der „**Neueren**“ und auch der „**Neuesten Urkundenhypothese**“⁴² war eine durchschlagend erfolgreiche Hypothese von Wilhelm Martin Leberecht de Wette (1780–1849)⁴³. De Wette koppelte in seiner Dissertation von 1805⁴⁴ das Buch Deuteronomium (5. Buch Moses) als eigenes ‚Dokument‘ von den Schriften 1.-4. Buch Moses als eine eigene Quellenschrift ab, die in der Forschung nun als Dokument D tradiert wird, und die er jünger datierte, als die für älter gehaltenen Quellen, nämlich in die für Josua angenommene Zeit (12. Jahrhundert v. Chr.). De Wettes Hypothese hatte für die Forschung Folgen bis in das 20. Jahrhundert.⁴⁵ Folgenreich war besonders, dass er den literarischen Stil des Dokuments D auch in weiteren AT-Schriften konstatierte⁴⁶. Damit habe de Wette nach Meinung einiger Forschungshistoriker die Pentateuchkritik des 19. Jahrhunderts eröffnet.

⁴⁰ Eine knappe, aber informierende Übersicht über die Geschichte der Pentateuch-Forschung bietet E. Otto: Art. „Pentateuch“, in: *RGG* (4. Aufl.), Bd. 6, Tübingen 2003, Sp. 1089–1102. Otto blendet neue internationale Entwicklungen aus, die seit der 2. Hälfte des 20. Jahrhunderts publiziert wurden, und ist damit kein Einzelfall. Dies gilt auch für den im Grunde gut informierenden Beitrag: (Anonymos:) Art. „Pentateuch“, in: *wikipedia*; vgl. auch: J. L. Ska: *Introduction to Reading the Pentateuch*. Winona Lake (Ind.) 2006.

⁴¹ Zu J. Wellhausen vgl. bes. R. Smend: „Julius Wellhausen“, in: *Deutsche Alttestamentler in drei Jahrhunderten*. Göttingen 1989 (= Smend 1989), S. 99–113; ders.: „Julius Wellhausen und seine Prolegomena zur Geschichte Israels“, in: *Epochen der Bibelkritik*. München 1991, S. 168–185.

⁴² Zu Otto Eissfeldt und Georg Fohrer weiter unten!

⁴³ Vgl. R. Smend 1989, S. 38–52.

⁴⁴ Vgl. W. M. L. De Wette: *Dissertatio critico-exegetica qua Deuteronomium (diversum) a prioribus Pentateuch libris diversum, alius cuiusdam recentioris (recentioris) auctoris esse (de)monstratur*, Jena 1805. – Der Titel wird leicht unterschiedlich zitiert. – Zu De Wette vgl. bes.: R. Smend: *Wilhelm Martin Leberecht de Wettes Arbeit am Alten und am Neuen Testament*. Basel 1958.

⁴⁵ Darauf komme ich weiter unten (bei M. Noth und H.-W. Wolf) zu sprechen.

⁴⁶ Vgl. weiter unten zu M. Noth und A. Jepsen.

Ich erwähne nur die Haupttypen von nun entwickelten Pentateuch-Hypothesen und charakterisiere sie kurz. Ihre Vielfalt und ihre Unterschiede sind oft auch für Fachleute verwirrend. Meist gehen auch die begrifflich definierten Hypothesen in der Sache ineinander über, so dass die ‚reinen‘ Typen ziemlich fiktiv sind. – Die ‚reine‘ **Quellenschrift** verteilt den gesamten Textbestand des Pentateuch auf definierte „Quellenschriften“. – Die ‚reine‘ **Fragmentenhypothese**⁴⁷ rechnet mit einem ‚Redakteur‘, der eine Unmenge von fragmentarischen Textstücken zu einem Gesamttext zusammengestellt habe. – Die ‚reine‘ **Ergänzungshypothese**⁴⁸ rechnet damit, dass eine Quellenschrift die literarische Grundschrift des Pentateuch sei, die durch eine andere ursprünglich selbständige und parallele Schrift ergänzt worden sei. Ewald gab seine Ergänzungshypothese später zugunsten einer **Urkundenhypothese** auf, bei der zwei Quellenschriften durch einen Redaktor zusammengearbeitet worden seien. – Die ‚reine‘ **Urkundenhypothese**⁴⁹ rechnet⁵⁰ mit einer älteren und einer jüngeren Quelle, die die Gottesbezeichnung „Älohim“ verwendet habe, und die mit einer weiteren Quelle mit dem Gottesnamen YHWH⁵¹ von einem Redaktor zusammengearbeitet und ergänzt worden sei. – Später nannte man nur noch die ältere Älohimquelle „Elohist“, während man die jüngere als „Priesterschrift“ (P) bezeichnete. Dies Modell gilt mitsamt späterer Entwicklungsstufen als „Neuere Urkundenhypothese“.

Gegen Ende des 19. Jh. setzte sich die Quellenschriftshypothese Wellhausens weithin als *communis opinio* zum mindesten bei den germanophonen Forschern durch. Damit waren andere Modelle grundsätzlich überholt. – Wellhausen rechnete mit vier Hauptquellenschriften (also: Autoren) für die literarische Gestaltung des Pentateuch: J = „Jahwist“, der den Gottesnamen YHWH (meist vokalisiert: „Jahwe“) verwendet, angenommen für das 10. Jh. vC; E = „Elohist“, der die Gottesbezeichnung „Ä/Elohim“ verwendet, angenommen für das 9./8. Jh. vC; P = „Priesterschrift“ (nicht durch Gottesbezeichnungen definiert, sondern durch Stil und

⁴⁷ Hauptvertreter der englische Theologe Alexander Geddes (1737–1802).

⁴⁸ Hauptvertreter neben zahlreichen anderen Heinrich Ewald (1803–1875).

⁴⁹ Hauptvertreter Hermann Hupfeld (1796–1866).

⁵⁰ Erst einmal nur für Gen (1. Buch Moses).

⁵¹ Lange Zeit in Gelehrtenstuben seit dem 19. Jahrhundert und noch heute meist in den Medien als „Ja(c)hwe(h)“ vokalisiert. Vermutlich aber wäre der Gottesname „Yahu(h)“ auszusprechen wie noch heute bei dem theophoren Namenselement des israelitischen Premiers: „Nethanyahu“. Die Forscher beschränken sich heute meist auf das konsonantische Namens-Tetragramm „Y/JHWH“.

Tendenz); angenommen für das 6. Jh. vC; und de Wettes in P als literarische Klammer eingearbeitetes Dokument D. Die je für sich fragmentarischen ‚literarischen ‚Quellen‘ J, E, P können, analog zu den Evangelien Mt, Mk, Lk, *cum grano salis* als „Synoptiker“ betrachtet werden, was die zahlreichen Doppelungen und redaktionellen Zusammenarbeitungen erkläre.

Unter dem Einfluss der in der Germanistik beheimateten Volkskunde des 19. Jh. fragte man nun nach den möglichen vorliterarischen Quellen der definierten ‚Quellenschriften‘, als die volkstümliche Sagen und Legenden definiert wurden. Mit Gunkel setzt die „Form-“ und „Gattungsgeschichte“ als Unterdisziplin ein, die nach Typen der mündlichen, vorschriftlichen Überlieferungen fragt. Hier spielte die Überlieferung der *Edda* methodisch eine wichtige Rolle. Von grundsätzlicher Bedeutung für die Erforschung des AT sind hier Arbeiten von Hermann Gunkel (1862–1932)⁵². Gunkel beschrieb des Buch Genesis (1. Moses) als eine „Sammlung von Sagen“. Diese ‚Sagen‘ mussten den Quellenschriftstellern ab dem 10. Jh. vC verfügbar gewesen sein. Diese Annahme fundamentierte nun Albrecht Alt (1883–1956)⁵³ methodisch mit seiner Habilitationsschrift durch eine Hypothese des 20. Jh., die als ‚wichtigste religionsgeschichtliche Hypothese‘ in Bezug auf die Entstehung der Religion ‚Israels‘ bezeichnet wurde.⁵⁴ Ich kann hier nur die Grundzüge der Hypothese anekdotisch zusammenfassen: Die ‚Väter‘ (auch ‚Patriarchen‘ genannt) der Genesis-Sagen, Abraham, Isaak und Jakob, in der Gen auf eine genealogische Folge gebracht, waren Stammesführer ‚vor-israelitischer‘ nomadischer Sippen, die sich auf ihren festgelegten Weiderouten gelegentlich trafen und sich von den sie beschützenden Gottheiten erzählten. Das waren die Götter der jeweiligen Sippengründer-Ahnen, der Stammes-Väter, nämlich der Gott Abrahams bzw. Isaaks bzw. Jakobs. Und während sie so (am Lagerfeuer) ihre jeweiligen Erfahrungen mit ihren Göttern untereinander austauschten, fiel ihnen auf, dass ihre Götter ziemlich identisch waren in ihren Funktionen für die jeweilige Nomadensippe: sie begleiteten die Sippen auf ihren Nomadenzügen,

⁵² Vgl. H. Gunkel: *Genesis*, 1. Aufl. Göttingen 1901; zu Gunkel vgl. Smend 1989, S. 160–172.

⁵³ Vgl. zu A. Alt: Smend 1989, S. 182–207; der Umfang (26 Seiten) dieser biographisch-forschungsgeschichtlichen Studie mit ihren 285 Seiten (ohne Anmerkungsseiten), verteilt auf 18 Alttestamentler des 18.–20. Jh., zeugt für die Bedeutung, die Alt im 20. Jh. beigemessen wurde.

⁵⁴ Vgl. A. Alt: *Der Gott der Väter*. BWANT 3. Folge 12. Stuttgart 1929; jetzt in *Kleine Schriften ...* Bd. I. München 1959, S. 1–78.

waren stets bei ihnen und behüteten sie vor allem Übel, und versorgten sie mit allem Nötigen. So stellten sie fest, das müsse doch ein und derselbe Gott sein und identifizierten ihre Götter. Alts „Gott der Väter“ war geboren, und konsequenterweise vereinigten sie daraufhin auch ihre Sippen zu einer Kultgemeinschaft. Die „Väter“ wurden in eine Genealogie von Großvater bis Enkel arrangiert. Das sei so etwa im 13. Jh. vC geschehen. Aus dieser Kultgemeinschaft wurde später – zur Richter- (im 12. / 11. Jh. vC) und zur staatlichen Zeit (ab dem 10. Jh. vC) das Volk ‚Israel‘. Erst nach dem 2. Weltkrieg – unter dem Einfluss der bedeutenden Altschüler Martin Noth (1902–1968) und Gerhard von Rad (1901–1971) – wurde Alts Vätergott-Hypothese zum unbezweifelten und unbezweifelbaren Forschungsdogma⁵⁵.

Eine vorsichtige, ehrvolle und erst einmal nur modifizierende Kritik an Alts Hypothese setzte in den 70er Jahren des 20. Jh. ein.⁵⁶ 1975 gab es bereits eine detaillierte system-immanente Kritik mit dem Ergebnis: auf Grund des methodischen Forschungsstandes und der materiellen (archäologischen) Kenntnisse um die Mitte der Zwanziger Jahre des 20. Jh. hätte Alt seine Hypothese gar nicht aufstellen dürfen⁵⁷. Heute gilt Alts Hypothese als falsifiziert. Zitiert wird jetzt dafür meist Matthias Köckert⁵⁸.

Das angenommene Alter der von den Quellschriftstellern verarbeiteten vorliterarischen, mündlich tradierten Sagenstoffe über die drei ‚Erzväter‘ kritisierte

⁵⁵ Kein Wunder, dass sich Noth auf der Basis von Alts Datierungen der Vor- und Frühgeschichte ‚Israels‘ gut mit einer US-amerikanischen Fundamentalistenschule verstand. William Foxwell Albright (1891–1971): ein verdienstvoller Archäologe, der vor allem nach Spuren des ‚Davidisch-Salomonischen Königsreiches‘ suchte, John Bright (1908–1995), Albrights Historiker (*A History of Israel*. London 1960), und George Ernest Wright (1909–1974), Albrights theologischer Systematiker. Ein Standardwerk zu Forschungsstand bis dato: M. Noth: *Überlieferungsgeschichte des Pentateuch*. Stuttgart 1948; Nachdruck Darmstadt 1960.

⁵⁶ Vgl. J. Hoftijzer: *Die Verheißung an die drei Erzväter*. Leiden 1956; K. T. Andersen: *Der Gott meines Vaters*. Studia Theologica 16, Århus 1963: vorsichtige Ansätze; H. Vorländer: *Mein Gott*. AOAT 23. Neukirchen 1975; auch J. Van Seters: *Abraham in History and Tradition*. New Haven, London 1975; ders.: *In Search of History*. New Haven, London 1983.

⁵⁷ Vgl. B. Diebner: „Die Götter des Vaters: Eine Kritik der ‚Vätergott‘-Hypothese Albrecht Alts“, in: *DBAT* 9 (1975), S. 21–51; dort im Detail Methoden- und Materialkritik. – W. H. Schmidt (*1935) billigt Alts Darstellung zu „methodisch vorbildlich“ zu sein (vgl. oben Anm. 17: UTB 421, S.13).

⁵⁸ Vgl. M. Köckert: *Vätergötter und Väterverheißungen*. FRLANT 142. Göttingen 1988.

punktuell, aber nachwirkungsvoll der Amerikaner John Van Seters⁵⁹. Er fand heraus, dass die Ehegesetze der Vätererzählungen gut in die Zeit um 500 vC passten. Während seriöse ATler sich noch darüber Gedanken machten, ob ‚Israel‘ durch den Bittersee vor dem Verfolgungsheer des Pharao floh oder bei Ebbe am Strand des Mittelmeeres, wo dann die Ägypter durch die einsetzende Flut überrascht wurden⁶⁰, hatte schon 1967/68 die dänische (damals noch) Studentin Heike Friis (1943–2018) in einer von der Universität von Kopenhagen mit einer Goldmedaille preisgekrönten Studie⁶¹ die These aufgestellt, dass das ‚Volk Israel‘ nicht – wie vom biblischen Mythos behauptet, sich durch Einwanderung über das Ostjordanland und ursprünglich aus Ägypten (Sklaverei, von Moses herausgeführt) konstituierte, sondern ab dem 11. Jh. vC durch soziale Spannungen innerhalb der palästinensischen Bevölkerung (zwischen Stadt- und Landbevölkerung). Friis stützt sich dabei auf archäologische Funde. Ihre These wurde von ihrem damaligen Mitbewerber und Silbermedaillengewinner Niels Peter Lemche (*1945) unter Berücksichtigung der Literatur über die Entstehung Israels seit George Mendenhall (1916–2016) 1962 aufgenommen⁶², allerdings ohne Berücksichtigung der Studie von Friis. Friis wird nur einmal kurz im Vorwort als Diskussionspartnerin erwähnt⁶³. Inzwischen gilt die gängige ‚Landnahme-Hypothese‘ als falsifiziert⁶⁴.

Zuvor (und danach) wurden aber immer wieder Pentateuch-Hypothesen aufgestellt, die mit ‚vor-israelitischem‘ Sagengut rechnen. Als **Neueste Urkundenhypothese** wird von Vielen eine auf den Hallenser ATler Otto Eissfeldt (1887–1973)

⁵⁹ Vgl. J. Van Seters in: *Harvard Theological Review (HTHR)* 1969.

⁶⁰ Warum flohen die ‚Israeliten‘ nicht um den kleinen See herum? Und offenbar hatten die Forscher keine Ahnung über den Höhenunterschied zwischen Ebbe und Flut am Mittelmeer und welche Zeit abfließendes und wieder aufsteigendes Wasser brauchen.

⁶¹ Vgl. H. Friis: *Forudsætninger i og uden for Israel for oprettelsen af Davids imperium*. MS, København 1968.

⁶² Vgl. N. P. Lemche: *Early Israel*. Leiden 1985. – In einer späteren Studie (*The Canaanites and Their Land: The Traditions of the Canaanites*. Sheffield 1991) vertritt Lemche die These, dass die Exilsrückkehrer ab dem 5. Jh. vC mit „Kanaaniter“ im AT die nicht exilierten Bevölkerungen Samariens und Judas bezeichneten.

⁶³ Daraufhin erschien das Opus von H. Friis in deutscher Übersetzung: *Die Bedingungen für die Errichtung des Davidischen Reiches in Israel und seiner Umwelt*. DBAT. B 6. Heidelberg 1986.

⁶⁴ Was nicht heißt, dass nicht noch viele an sie glauben.

zurückgehende Analyse betrachtet⁶⁵. Eissfeldt glaubt an eine kompositorische Fortsetzung der Pentateuch-Quellenschriften bis Ri 2,9. Im Wesentlichen spaltet Eissfeldt den herkömmlichen „Jahwisten“ (J) auf und seziert aus ihm eine „Laienquelle“ (L) heraus. Georg Fohrer (1915–2002) schließt sich ihm im Prinzip an, nur dass er die Quelle L in N („Nomadenquelle“) umbenennt.⁶⁶ L/N sind angeblich älter als J. – Was angenommene mündliche Vorstufen später verschrifteter Texte betrifft, so stützte man sich um die Mitte des 20. Jh. auf die neuere Volkskunde⁶⁷. Neueste Pentateuch-Hypothesen orientieren sich im Titel nicht mehr an den traditionellen Quellen⁶⁸, sondern tragen z. B. die Namen der Wirkstätten ihrer Produzenten. Am bedeutendsten wurde in jüngerer Zeit das sog. „Münsteraner Modell“, das auf den katholischen Forscher⁶⁹ Erich Zenger (1939–2010) zurückgeht⁷⁰.

Dieser Abschnitt darf nicht schließen ohne einen Hinweis auf die Saat von De Wette. Martin Noth hatte das Glück, im 2. Weltkrieg nicht eingezogen zu werden. Und so erarbeitete er in Königsberg seine Hypothese vom „Deuteronomistischen Geschichtswerk“⁷¹ (DtrG). Dieses Geschichtswerk umfasse die Schriften Jos bis

⁶⁵ Vgl. O. Eissfeldt: Hexateuch-Synopse: *Die Erzählung der fünf Bücher Mose und des Buches Josua mit dem Anfange des Richterbuches*. Leipzig 1922; Nachdruck: Darmstadt 1962. – Es gibt auch Oktateuch-Hypothesen die außer den fünf Mosesbüchern noch Josua, Richter, die Samuel- und Königsbücher umfassen. Aber ihnen war in der Forschungsgeschichte keine Zukunft beschieden.

⁶⁶ Was „Laien“ und „Nomaden“ im Alten Orient und in der Antike mit Literatur zu tun haben sollen, bleibt das Geheimnis dieser Forscher. Sie stützen sich wesentlich auf den Erzählstil.

⁶⁷ Vgl. Edward R. Haymes: *Das mündliche Epos: Eine Einführung in die „oral-poetry“-Forschung*. Slg. Metzler. Stuttgart 1977.

⁶⁸ Die Differenzierungen im angenommenen Entwicklungs- / Entstehungs-Prozess des Pentateuch waren zu komplex geworden.

⁶⁹ Bislang waren die historisch-kritisch forschenden AT.ler fast durchweg Protestanten, was auch mit den von der römischen Kirche gesetzten Rahmenbedingungen zu tun hat. Zu den international bekanntesten Ausnahmen gehören Norbert Lohfink SJ (*1928), Frankfurt, und Bernhard Lang (*1946), Paderborn und Aberdeen.

⁷⁰ Eine graphische Schautafel des komplizierten Zenger-Modells findet sich in: *TRE* Bd. 6. Tübingen 2003, Sp. 1099f.

⁷¹ Vgl. M. Noth: *Überlieferungsgeschichtliche Studien: Die sammelnden und bearbeitenden Geschichtswerke im Alten Testament*. Königsberg 1943; Nachdruck Darmstadt 1957. – Eine ähnliche Hypothese beschrieb 1956 der Greifswalder ATler Alfred Jepsen (1900–1979): *Die Quellen des Königsbuches*. Halle 1956.

2. Kön⁷² und sei durch den Sprachstil des 5. Buches Moses (D) geprägt. Noths These setzte sich allgemein durch. Ja, Forscher entdeckten sogar, dass auch andere Schriften des AT stilistisch so geprägt seien. So z. B. der Heidelberger ATler Hans-Walter Wolff (1911–1993) für die Prophetenschrift Amos und ein anderer Heidelberger ATler, Rolf Rendtorff (1925–2014) sogar für das Buch Gen (1. Moses). Man kann schlecht sagen, dass Noths These falsifiziert worden sei; denn Sprach- und Schrifteigentümlichkeiten lassen sich philologisch beschreiben. Man könnte eher sagen: sie löste sich in der Allgemeinheit der AT-Texte auf⁷³.

3 Grundlegende Infragestellung gängiger Hypothesen⁷⁴

Schon in den 70er Jahren des 20. Jh. wurde nach den stichhaltigen Kriterien für eine Frühdatierung von AT-Texten (etwa ins Ende des 2. Jt. oder in die Zeit zwischen 1000 vC und dem 5. Jh. vC) gefragt und die weitere Frage gestellt, ob sich solche Texte nicht mit besseren Gründen in die ‚nachexilische Zeit‘ (ab etwa dem Ende des 6. Jh. vC) datieren ließen⁷⁵. Der hier entworfenen methodischen Skizze entsprach ein Thesenpaket zur ‚Geschichte Israels‘⁷⁶. Hier wurde u. a. die Historizität des ‚Davidisch-Salomonischen Großreiches‘ in Frage gestellt. Dieser Zweifel am Urmythos des Staatsverständnisses unseres modernen Israel wird heute selbst von einigen israelischen Forschern und Archäologen geteilt.

Der Schweizer ATler Hans Heinrich Schmid (1937–2014) bezweifelte die Fundamente der allgemein akzeptierten Neuen Urkundenhypothese durch eine Studie⁷⁷. Er wendet im Grunde die Idee des ‚Deuteronomismus‘ konsequent auf die Pentateuchtexte an und datiert seinen anders definierten ‚Jahwisten‘ ins

⁷² Nicht die 1. und 2. Chr. Diese Schriften werden als *corpus sui generis* betrachtet und laufen unter dem Kürzel „ChrG“ = „Chronistisches Geschichtswerk“.

⁷³ Einige bezeichneten diesen Literaturstil ironisch als „Sprache Kanaans“.

⁷⁴ Der völlige Umbruch in der AT-Forschung. Niels Peter Lemche als ein wesentlicher Akteur wählte den dänischen Begriff: „nybrudde“.

⁷⁵ Vgl. B. Diebner: „Eine methodische Alternative zur gegenwärtigen Erforschung des Alten Testaments: Otto Plöger zum 65. Geburtstag“, in: *DBAT* 10 (1975), S. 48–62.

⁷⁶ B. Diebner, H. Schult: „Thesen zu nachexilischen Entwürfen zur frühen Geschichte Israels im Alten Testament“, in: *DBAT* 10 (1975), S. 41–47; kaum veränderter Nachdruck 1995.

⁷⁷ Vgl. H. H. Schmid: *Der sogenannte Jahwist*. Zürich 1976. Hier nimmt Schmid Anregungen von Diebner und Schult auf.

6. Jh. vC, also weder in die davidisch-salomonische Zeit, noch überhaupt in eine vorexilische Zeit. Zuvor kam ihm allerdings der Heidelberger ATler Rolf Rendtorff mit einer methodischen Studie von 1967⁷⁸.

Die sog. ‚Exilszeit‘⁷⁹ wird für die kommende Zeit zur Literatur-Küche für die Produktion des AT-Schrifttums. Für immer mehr Schriften nimmt ein *mainstream* nun die Entstehung ‚im Exil‘ an⁸⁰. Einige Forscher sind allerdings bemüht, die ‚evangelischen‘ Traditionen der AT-Schriften für die ‚vorexilische‘ Zeit zu retten⁸¹. Für die Mehrheit der gegenwärtigen ATler gilt die **Neu(er)e** und **Neueste Urkundenhypothese** als falsifiziert, was – um darauf noch einmal zu verweisen – nicht ausschließt, dass einige ATler ihr noch heute anhängen. So z. B. der Bielefelder ATler Horst Seebass (1934–2015).

Für die Veränderung des Bildes von der ‚Geschichte Israels‘ in der Forschung ist der italienische Waldenser-Professor Jan Alberto Soggin (1926–2010), Rom, zu erwähnen. In seinen Publikationen zur ‚Geschichte Israels (und Judas)‘ in mehreren Sprachen seit Anfang der 80er Jahre⁸², verzichtete Soggin auf die Annahme des ‚Davidisch-Salomonischen Großreiches‘. Er rechnet mit zwei verwandten,

⁷⁸ Vgl. R. Rendtorff: „Literarkritik und Traditionsgeschichte“, in: *EvTh* 27/3 (1967), S. 138–153; später stellte Rendtorff „die Existenz des Hauptpfeilers der Urkundenhypothese, des ‚Jahwisten‘, in Frage“, nämlich in „Der ‚Jahwist‘ als Theologe? Zum Dilemma der Pentateuchkritik“, in: *VT.S* 28 [1975], S. 158–166. Das Zitat aus R. Rendtorff: *Das überlieferungsgeschichtliche Problem des Pentateuch*. BZAW 147, Berlin, New York 1977, „Vorwort“.

⁷⁹ Angenommen für den Zeitraum 587/86 bis 539/38 vC

⁸⁰ Als hätten die bedauernswerten, aus ihrer israelischen (721vC) bzw. jüdischen (587 vC) Heimat entführten Israeliten und Judäer nichts Anderes zu tun, als über ihre Vergangenheit zu reflektieren und darüber große Literatur zu verfassen.

⁸¹ So, um nur ein Beispiel zu nennen, Frank Crüsemann (*1939) im Blick auf die Zehn Gebote, die er in das Ende des 7. Jh. vC datiert (vgl.: *Der Dekalog: Bewahrung der Freiheit*. München 1983). Hier dürfte noch der klassische Antijudaismus der ATler nachschwingen: Was uns Protestanten am AT teuer und wert ist, geht auf das ‚Alte (vornomistische und vorexilische) Israel‘ zurück – oder gar auf die vor-israelische Nomadenkultur. So skizziert Albrecht Alt mit seinem „Gott der Väter“ theologisch im Grunde die helle Vorderseite des protestantischen Gottesbildes. Dessen dunkle Rückseite ist Martin Luthers *Deus absconditus*, der „verborgene, geheime, unbekannte Gott“, der für alles uns unverständliche Übel verantwortlich sein könnte.

⁸² Nur zwei Titel: J. A. Soggin: *A History of Israel: From the Beginning to the Bar Kochba Revolte, AD 35*. London 1984; ders.: *Einführung in die Geschichte Israels und Judas: von den Ursprüngen bis zum Aufstand Bar Kochbas*. Darmstadt 1991; 3. Aufl. engl.: *An introduction to the History of Israel and Judah*. London 1999.

aber ursprünglich voneinander unabhängigen Kleinstaaten: Israel im Norden („Nordreich“) und Juda (gräzisiert in hellenistischer Zeit: *Ἰουδαία* / Judäa; „Südreich“)⁸³.

Der Prozess der von Traditionalisten sog. „Spätdatierung“ ATlicher Schriften ging indes weiter. Noch immer rechneten Forscher mit dem literarischen Abschluss des hebräischen Pentateuchtextes für die Zeit um 400 vC⁸⁴ oder um 350 vC⁸⁵

⁸³ Hier nahm Soggin Anregungen der sog. „Dielheimer Schule“ auf.

⁸⁴ So der Tübinger ATler Erhard Blum (*1950): *Studien zur Komposition des Pentateuch*. BZAW 189, Berlin, New York 1990. – Hier arbeitet Blum grundsätzliche Impulse seines Lehrers Rolf Rendtorff aus.

⁸⁵ So Fr. Crüsemann: *Die Tora*: 1. Aufl. München 1992; 3. Aufl. Gütersloh 2005. – Es besteht in der AT-Forschung eine Wechselwirkung zwischen dem hypothetischen literarischen Abschluss der Thorah und dem ebenso hypothetischen sog. „Samaritanischen Schisma“ (das historisch wahrscheinlich ein jüdisches war): zwischen 400 und 350 vC. Das Argument: die Samaritaner haben in ihren Synagogen nur die Thorah; späteres Schrifttum bekamen sie nicht mehr mit. Aber auch die Juden haben in ihrem Thorahschrein (daher der Name) nur die Thorah. Der Thorahschrein ist kein Bücherschrank für religiöse Literatur. Die konfessionelle Trennung von Samaritanern und Judäern war wohl ein schleicher Prozess, der sich vom 2. Jh. vC bis zum 2./3. Jh. nC vollzog. Zu den Samaritanern vgl. heute B. Hensel: *Juda und Samaria: Zum Verhältnis zweier nachexilischer Jahwismen*. FAT 110, Tübingen 2016 (hier auch alle belangreiche Literatur). – Dass die Samaritaner übrigens ‚nur die Thorah‘ hätten, trifft nicht zu. Sie hatten auch die Schriften Josua, Richter, 1.2Samuel, 1.2.Könige – nur in ihrer eigenen polemischen (antijerusalemischen) Rezension. Ein Manuskript wurde prominent publiziert von J. Macdonald: *The Samaritan Chronicle No. II*. BZAW 107, Berlin 1969. Ich fand noch keinen ATler, der je hieraus zitiert hätte. Lediglich die Propheten, sei es in narrativen Texten oder in den mit ihren Namen verbundenen Schriften („Hintere Propheten“ des TNK), lehnten die Samaritaner wegen deren überwiegend anti-samaritanischen Tendenzen verständlicherweise als ‚Lügenpropheten‘ ab. Wenn wir an die Samaritaner denken, dann fällt uns das kleine heutige Häufchen ein, für uns ein häretisches und heterogenes ‚Spaltprodukt‘ des ‚Alten Israel‘. Die Samaritaner bildeten in der Antike in Palästina eine größere Gemeinde als die Jerusalem zugewandten Juden. Sie missionierten im Mittelmeerraum ähnlich wie die Judäer. Die berühmte Synagoge von Delos ist wahrscheinlich eine samaritanische. Die Samaritaner waren gegen die oströmische Besatzungsmacht zu frühbyzantinischer Zeit aufständischer als die Juden. Erst Justinian I. dezimierte sie (im Wortsinn) anlässlich eines Aufstandes im 6. Jh. – Für eine kritische Würdigung der Rolle der Samaritaner in der Antike ist unser einseitiger (theologischer) Focus auf Jerusalem und die Jerusalemer Traditionen hinderlich. Wir sind es über 2000 Jahre hinweg durch den jüdisch rezensierten TNK gewöhnt, alles aus Jerusalemer Sicht zu betrachten und zu würdigen, mit der Folge einer abschätzigen parteiischen Sicht auf die Samaritaner. Dazu gehören die historisch völlig unglaubwürdigen negativen und parteiischen theologischen Bewertungen der Könige Israels / Samarias seit Saul. Die ‚Geschichten Judas und Samarias‘ wären aus kritischer und unparteiisch-neutraler Sicht umzuschreiben.

oder (revolutionär) der Jerusalemer Dominikanerpater Étienne Nodet (*1944) um 250 vC. – Ein neuer *mainstream* pendelt sich in der nachexilischen Perserzeit (ab dem mittleren 5. Jh. vC) ein⁸⁶. Einige sind bereits vor der Etablierung dieses *mainstreams* ‚ausgebrochen‘: so der dänische ATler Lemche⁸⁷. Er verlegt den Abschluss des Pentateuchs in die hellenistische Ära (ab dem 3. Jh. vC), nur mit einem methodischen Grundfehler: Wenn auch in hellenistischer Ära abgeschlossen, so ist der Pentateuch doch kein ‚hellenistisches Buch‘, sondern, kulturpolitisch beurteilt, ein dezidiert antihellenistisches Buch, trotz Aufnahme griechisch-hellenistischen Gedankengutes.

Dass griechischer Einfluss auf wichtigste Partien des Pentateuch angenommen werden kann, wurde auch bereits zu zeigen versucht⁸⁸. Es lässt sich annehmen, dass die naturphilosophischen Schriften des Aristoteles (384–322 vC) viel zum Verständnis des Pentateuch beitragen können, nicht nur zur ‚priesterlichen‘ Schöpfungsgeschichte⁸⁹, sondern etwa auch zu den biblischen Speisevorschriften⁹⁰. Man müsste diese Schriften⁹¹ nur lesen⁹².

Angesichts der unzähligen, z.T. grundverschiedenen und einander widersprechenden analytischen Pentateuch-Modelle⁹³ gab es schon im letzten Jahrhundert Ansätze, die auf derartige Analysen verzichteten und einen literarischen Zugang zu den uns biblisch in den alten Manuskripten überlieferten AT-Texten auch au-

⁸⁶ Hauptvertreter dieser Tendenz ist heute u. a. der Lausanner ATler Thomas Römer (*1955).

⁸⁷ Vgl. N. P. Lemche: „The Pentateuch – a Hellenistic Book?“, in: SJOT 7/2 (1993), S. 163–193; leider ist der im Grunde wichtige Beitrag in einem geradezu lächerlichen Englisch abgefasst. Später ließ Lemche seine Artikel und Bücher vom anglophonen Muttersprachler Frederick Cryer (der selber viel zu einem Paradigmen-Wechsel in der AT-Forschung beigetragen hat) in gutes Englisch übersetzen.

⁸⁸ Vgl. B. J. Diebner: „Platonisch-Aristotelisches und frührabbinische Denkstruktur in Gen 1 – 3 : Zur kulturgeschichtlichen Einordnung von Schöpfungs- und Paradies-Erzählung“; Erstpublikation 2010; jetzt in: *Seit wann gibt es „jenes Israel“? Gesammelte Studien zum TNK und zum antiken Judentum*. BVB 17. Berlin 2011 (= Diebner 2011), S. 87–96.

⁸⁹ Vgl. Gen 1,1 – 2,3.

⁹⁰ Vgl. Lev (3. Moses) 11 und Dtn (5. Moses) 14.

⁹¹ Besonders: *De anima*; *De motu animalium*; *De partibus animalium*.

⁹² Was ATler nicht tun, weil die AT-Schriften ja angeblich viel älter sind.

⁹³ Von denen rein statistisch gesehen nur maximal eine Hypothese (einigermaßen) zutreffend sein kann.

ßerhalb des Pentateuch befürworteten⁹⁴. Hierzu gehört auch die „Amsterdamse School“ mit ihrem exegetischen und bibeltheologischen Kopf Karel A. Deurloo (1936–2019). Die „Amsterdamse School“ – die seit 1988 eng mit Prager Theologen um den Alttestamentler Jan Heller (1925–2007) zusammenarbeitet – stellt mit ihren historischen Wurzeln in der jüngeren reformierten Theologie der Niederlande einer ihrer gegenwärtigen Vertreter vor: Klaas Spronk (*1957)⁹⁵. Hier wird nicht eine komplexe Entstehungsgeschichte der AT-Texte verleugnet, aber der tradierte Text als sinnvoll vorausgesetzt und der Versuch unternommen, ihn in seinem wahrscheinlichen historischen Kontext zu interpretieren. Als solcher wird mit guten Gründen das vorchristliche Judentum angenommen. – Hiermit korrespondierend gibt es den Versuch einer synchronen (nicht diachronen) Lektüre des TNK unter konfessionspolitischen Aspekt mit der Frage, wie „Israel“ unter der Berücksichtigung der Spannungen zwischen Samariern (dem historisch alten Israel) und Judäern (die möglicherweise erst im 2. Jh. vC⁹⁶ die Bezeichnung „Israel“ von den Samariern übernahmen) in den verschiedenen Teilen der *Biblia Hebraica*⁹⁷ ekklesiologisch unterschiedlich definiert wird⁹⁸. Dies geschieht in der Thorah „ökumenisch“: Samaria und Judäa sind ekklesiologisch „Israel“, die Kultgemeinde; in den „Propheten“ sind die Samarier nicht exkommuniziert, aber der (prophetische) Bußruf gilt ihnen intensiver als den Judäern; in den religionspolitisch relevanten „Schriften“⁹⁹ sind die Samarier exkommuniziert¹⁰⁰.

⁹⁴ Vgl. die „Uppsala-Schule“ mit ihren Vertretern I. Engnell, H. Ringgren. Am international bekanntesten ist wohl der anglophone „New Literary Criticism“.

⁹⁵ Vgl. oben Anm. 17; dort weitere Literatur zur „Amsterdamse School“, u. a. von Uwe Bauer (*1956), einem deutschen Schüler dieser Richtung. – Anzumerken ist, dass die zahlreichen Publikationen der Repräsentanten der „Amsterdamse School“ von der *mainstream*-Diskussion leider überhaupt nicht zur Kenntnis genommen werden. Man könnte sagen, sie würden nicht einmal ignoriert. Dem „Ignorieren“ läge bereits eine Wahrnehmung und Entscheidung zugrunde.

⁹⁶ Nach der Eroberung Samariens durch den jüdischen Makkabäerfürsten Johannes Hyrkanos I. (134–104 vC).

⁹⁷ „Thora“ (T) = Mosesbücher; „N:vi‘im“ (N) = „Vordere“ (Jos, Ri, 1.2. Sam, 1.2. Kön) und „Hintere Propheten“ (Jes, Jer, Ez, Zwölf-Propheten-Buch), sowie „K:tuvim“ (K) = übrige „Schriften“ des hebr. Kanons.

⁹⁸ Vgl. u. a.: B. J. Diebner: „Seit wann gibt es ‚jenes Israel‘ (Martin Noth)? Anmerkungen zu ‚Israel‘ als ekklesiologische Größe im TNK“, in: Diebner 2011, S. 67–84.

⁹⁹ 1.2.Chronik, Esra, Nehemia.

¹⁰⁰ Dazu mehrere Studien von B. J. Diebner (seit 1990).

Kultgemeinde „Israel“ sind exklusiv die Judäer. Hier – wie auch für das Buch Jesaja¹⁰¹ – wird nach der theologischen Konzeption definierter tradierter Texteinheiten gefragt.

Die Falsifikations-Bemühungen bezüglich tradierter und jeweils etablierter Forschungsmeinungen zu AT-Texten sind noch nicht abgeschlossen. Man könnte statt „Falsifizierungen“ auch den schwächeren Begriff „Paradigmenwechsel“ verwenden. Die Geschichte der Erforschung des AT: seit 300 Jahren ein permanenter Paradigmenwechsel. Es besteht in der Disziplin eine ungebrochene Meinungsfreiheit. Niemand muss sich, durch welche Argumente auch immer, überzeugen und von der eigenen Meinung abbringen lassen.

4 Schlussbemerkungen

Deutlich ist, dass ein Versuch wie der hier vorgelegte nur eklektisch sein kann. Maßgebliche Themen (so z. B. die Monotheismuskussion (seit etwa 1970 [vgl. Bernhard Lang; Manfred Weippert]) und Autor/inn/en können auf einem begrenztem Raum nicht berücksichtigt werden¹⁰². Dazu gehört, dass andere Autoren anders gewichten würden. Die Auswahl dessen, was man im Rahmen einer Fragestellung für belangreich hält, kann nur subjektiv sein.

Ein Manko der meisten kulturwissenschaftlichen Disziplinen ist, dass sich im Grund nichts schlüssig ‚beweisen‘ lässt, obwohl auch manche ATler immer wieder behaupten, etwas ‚bewiesen‘ zu haben, oder dass eine Meinung „außer Zweifel“ / „zweifellos“ richtig sei.¹⁰³ Ich denke dabei immer an den Ausruf Rudolf Smends von 1961: „Werden Sie bloss keine Alttestamentler! Da ist bereits alles erforscht!“

¹⁰¹ Vgl. B. J. Diebner: „Jes 56,1–8 entsprechend Jes 66,18–24 und die prophetische Überbietung der Torah: Yad wa Shem“, in: *Religionsgeschichte des Neuen Testaments: Festschrift für Klaus Berger zum 60. Geburtstag*. Tübingen, Basel 2000, S. 31–42.

¹⁰² Daher in Anm. 17 der Verweis auf die forschungsgeschichtlichen Kompendien und unselbständigen Überblicke.

¹⁰³ Vgl. dazu B. J. Diebner: „‘Es lässt sich nicht beweisen, Tatsache aber ist ...‘ : Sprachfigur statt Methode in der kritischen Erforschung des AT“, in: *DBAT* 18 (1984), S. 128–146. – Eine Spitzenformulierung bietet der berühmte Albrecht Alt: Es „bleibt weithin im Dunkel, und doch wird kein Verständiger leugnen wollen, daß ...“ („Der Gott der Väter“, in: *Kleine Schriften zur Geschichte des Volkes Israel*. 2. Aufl., München 1959, S. 2).

Im selben Jahrzehnt noch begann der bis heute nicht abgeschlossene Umbruch (nybruddet) in der Erforschung des AT.

Ein für Exegeten typisches Manko ist, dass sie oft meinen, besser über die Bibeltexte Bescheid zu wissen als die antiken Autoren letzter Hand selbst. Dies beginnt bei der Textkritik in den wissenschaftlichen Kommentaren, wo Exegeten ‚ursprüngliche‘ Lesungen rekonstruieren, und wird fortgesetzt in der Literarkritik, die Texte zerlegt und verschiedenen Zeitstufen zuordnet, statt nach der möglichen Sinnhaftigkeit der tradierten, vorliegenden Texte zu fragen, diese einfach als Literatur zu lesen. Dabei sollte der hermeneutische Grundsatz gelten: „Nicht die Texte unseren Lesegewohnheiten anzupassen, sondern unsere Leseweisen auf die Texte einzustellen“.

Weil sich im Bereich der Erforschung des AT – anders als in manchen naturwissenschaftlichen Disziplinen – nichts schlüssig beweisen lässt, lässt sich eigentlich auch nichts eindeutig falsifizieren – es müsste denn totaler Unsinn sein. Aber es gibt Forschungspositionen, für die es heute keine guten Gründe und tragenden Argumentationen mehr gibt, obwohl diese Argumente seinerzeit durchaus überzeugend gewesen sein mögen. – Ein Defizit bei den ATlern besteht darin, dass kaum jemand über die seine Methoden und Hypothesen tragenden Prämissen reflektiert, geschweige denn, diese kontrollierbar offenkundig macht. Im Grunde gilt in der AT-Disziplin¹⁰⁴ das Schusterprinzip: Man schaut zu, was und wie es der Meister macht und macht es ihm nach. Originell sind dann einige persönliche Schnörkel.

¹⁰⁴ Wie gewiss auch in manchen anderen – besonders kulturwissenschaftlichen – Disziplinen.

Über den Autor

Bernd Jørg Diebner ist Theologe (Pastor, Bibelwissenschaften, Kirchengeschichte), Archäologe (Frühbyzantinistik) und Orientalist (Koptologie). Promotion 1965 in Heidelberg im Fache Kirchengeschichte (Spezialfach: frühchristlicher Kirchenbau). Lehraufträge für Koptologie am Ägyptologischen Seminar der Universität Hamburg und später am Ägyptologischen Institut der Ruprecht-Karls-Universität Heidelberg. Dort auch Lehrauftrag für Altes Testament seit 1972. Vorlesungen und Seminare über die Pensionierung 2004 hinaus bis 2017/18. – Als Germanist seit 1979 Arbeit, Publikationen und Herausgeber-Tätigkeit im Bereich der niederdeutschen Sprache und Regionalgeschichte. – Seit 1985 in Nebentätigkeit Pastor von Dansk Kirke i Sydslesvig.

Korrespondenz:

Hon.-Prof. Dr. Bernd Jørg Diebner
Universität Heidelberg
Wissenschaftlich-Theologisches Seminar
Kisselgasse 1
69117 Heidelberg
E-Mail: bernd.diebner@wts.uni-heidelberg.de
Homepage:
<https://www.uni-heidelberg.de/fakultaeten/theologie/personen/diebner.html>

Universalität der Bausteine und Prozesse des Lebens

MICHAEL WINK

Institut für Pharmazie und Molekulare Biotechnologie, Universität Heidelberg

Zusammenfassung

Obwohl Bakterien, Pantoffeltierchen, Blütenpflanzen, Insekten oder Säugetiere sehr unterschiedlich aussehen, teilen sie sich viele Gemeinsamkeiten. Die chemischen Bausteine, die für den Aufbau und Unterhalt der Zellen sowie für den Stoffwechsel benötigt werden, sind bei allen Lebewesen nahezu identisch. Daher nimmt man an, dass sich am Anfang der Evolution eine Urzelle bildete (oder von einem anderen System des Universums zu uns transportiert wurde), von der alle Zellen bis heute abstammen. Die Zellen aller lebenden Zellen sind somit in einer kontinuierlichen Linie miteinander verbunden und teilen sich die grundlegenden Bausteine und Prozesse. Zellen differenzierten sich, so dass komplex aufgebaute Gewebe und Organe entstehen konnten. Durch Genomduplikation, Mutation, Hybridisierung, horizontalen Gentransfer und Rekombination wurden die Genome der Organismen im Verlauf der letzten 3,5 Milliarden Jahren immer komplexer. Durch die genetische Variation entstanden neue Phänotypen. Durch natürliche Selektion überlebten die Phänotypen, die am besten an die jeweiligen Umweltbedingungen angepasst waren. Genetische Isolierung der erfolgreichen Genotypen führte zu unterschiedlichen Entwicklungslinien, aus denen die Vielfalt des Lebens entstand.

1 Einführung

Von Theodosius Dobzhansky (1900–1975) stammt der vielzitierte Ausspruch „Nichts in der Biologie ergibt einen Sinn außer im Licht der Evolution.“ Der größte Verdienst von Charles Darwin (1809–1882) bestand darin, dass er eine Evolutionstheorie ausarbeitete (Darwin 1859), die auch heute noch aktuell ist und die biblischen Schöpfungsmythen ablöste. Bis in das 20. Jahrhundert wurde gemäß der biblischen Schöpfungsgeschichte angenommen, dass Gott alles Leben schuf. Nach der biblischen Lehrmeinung hat Gott die Lebewesen so optimal konstruiert, dass sie an die diversen Lebensbedingungen auf unserer Erde bestens angepasst sind und sich erfolgreich fortpflanzen können. Arten galten als konstant und konnten nicht neu entstehen. Darwin entwickelte den revolutionären Gedanken, dass es vor langer Zeit Urformen des Lebens gegeben haben muss, aus denen sich neue Arten kontinuierlich und graduell abgeleitet haben. (Diese Idee hatten aber auch schon andere Wissenschaftler vor Darwin und selbst Johann Wolfgang von Goethe hat schon von der Evolution gesprochen). Darwin nahm folgerichtig an, dass Arten veränderlich sind und sich aus bestehenden Arten ableiten können. Nach Darwin sind alle Lebewesen durch eine gemeinsame Abstammung miteinander verbunden (wir sprechen von Stammesgeschichte oder Phylogenese).

Jedes Lebewesen ist so an seine Umgebung angepasst, dass es sich ernähren, behaupten und fortpflanzen kann. Diese Anpassungen sind nach Darwin (1859) nicht von einem Schöpfergott genial erdacht worden, sondern durch das Prinzip der „natürlichen Selektion“ entstanden. Voraussetzung für die Selektion ist eine hohe Variabilität der meisten Merkmale einer Art und innerhalb einer größeren Population. Durch die natürliche Selektion haben Individuen, die zufällig besser an eine Umweltsituation angepasst sind, bessere Überlebenschancen. Sie können mehr Nachkommen produzieren und dabei diese adaptiven und meist vererbaren Merkmale an ihre Nachkommen weitergeben. Schon zu Darwins Zeiten kannte man das Phänomen der Vererbung, nicht aber deren Mechanismen. Hundert Jahre später haben die Biologen und Genetiker gezeigt, dass unsere Erbinformation in den Chromosomen gespeichert ist und stofflich aus DNA besteht. Gene, von denen es rund 3.0000 im menschlichen Genom gibt, sind die Träger der Erbinformation für einzelne Merkmale (Alberts et al. 2015; Storch et al. 2013). Die moderne Biologie versucht zu erkennen, wie man über Gene den Phänotyp der diversen Lebewesen und die Komplexität der Merkmale, Leistungen und Anpassungen erklären kann.

Die Theorie der „natürlichen Selektion“ hat sich bestens bewährt und Biologen beurteilen die diversen Merkmale eines Lebewesens unter dem Aspekt der evolutionären Anpassung (Adaptionismus). Je besser ein Individuum an seine Umwelt angepasst ist, desto höher (statistisch gesehen) sein Fortpflanzungserfolg und die Chance seine Gene an weitere Generationen weiter zu geben. Aber auch heute noch gibt es viele Menschen, die an die biblische Schöpfungsgeschichte glauben und die Evolutionstheorie ablehnen. Eine Variante des Kreationismus ist die Vorstellung eines „intelligent design“. Danach hat Gott wie ein Uhrmacher alle wichtigen komplexen Einzelemente geschaffen, die im Verlauf der Phylogenie durch Kombination zu neuen angepassten Formen führten (Storch et al. 2013).

In diesem Aufsatz möchte ich erörtern, wann und wie das Leben entstanden sein könnte, wie die ersten Zellen entstanden, wie sie sich differenzierten (Alberts et al. 2015; Storch et al. 2013), wie aus den differenzierten Zellen Gewebe entstanden, die in Organen zusammenarbeiten und in der Gesamtheit den Organismus darstellen. Wir erörtern die Phylogenese von der Urzelle zu Prokaryoten und zu den Eukaryoten. Dann geht es von den Einzellern zu den Vielzellern, also um Pflanzen und Tieren. Und zuletzt machen wir einen kurzen Ausflug in die Phylogenie der Pflanzen und Tiere. Obwohl diese Organismen unterschiedlich aussehen und unterschiedliche Anpassungen zeigen, weisen sie jedoch ähnliche Bausteine, Baupläne und Entwicklungsprozesse auf. Dies ist möglich, da alle Organismen eine gemeinsame Abstammung aufweisen. Außerdem ist die Natur eher konservativ und hat Strukturen und Prozesse, die sich bewährt haben beibehalten. Daher kann man in der äußerst artenreichen Biodiversität eine Universalität der Bausteine und Strukturen erkennen.

2 Ursprung der Erde und des Lebens

Astronomen gehen davon aus, dass unser Universum 13,81 Milliarden Jahre alt ist und durch den Urknall entstand. Aber auch oszillierende Systeme werden diskutiert. Unser Sonnensystem ist Teil des Universums und entstand vor 4,568 Milliarden Jahren, die Erde etwas später von 4,543 Milliarden Jahren. Vor rund 3,9 Milliarden Jahren bildete sich eine stabile Erdkruste aus und die Erde kühlte ab und sammelte Wasser an (Alberts et al. 2015; Storch et al. 2013). Man nimmt an, dass das Leben auf der Erde erst in der Folgezeit entstand. Wie konnte dies geschehen?

Zunächst schauen wir uns die Bausteine der lebenden Materie an und wie daraus die Grundeinheit des Lebens, die Zelle entstand.

3 Universalität der Moleküle des Lebens

Obwohl Bakterien, Pantoffeltierchen, Blütenpflanzen, Insekten oder Säugetiere sehr unterschiedlich aussehen, sind sie alle aus Zellen mit vergleichbarer Biochemie aufgebaut. Alle Zellen bestehen aus vielen kleinen organischen Substanzen und größeren Makromolekülen. Obwohl es theoretisch Milliarden an möglichen kleinen oder großen organischen Molekülen geben könnte, ist die Auswahl bei allen lebenden Organismen übersichtlich. Noch überraschender benutzen alle Organismen, von den einfachen Bakterien bis zu den komplex aufgebauten Tieren, denselben Satz an Bausteinen (Alberts et al. 2015; Storch et al. 2013; Wink 2020). Wir können daher von einer Universalität der Bausteine des Lebens sprechen.

Die wichtigsten Bausteine in allen Zellen sind Zucker (Kohlenhydrate), Fettsäuren, Aminosäuren, organische Säuren und Nucleotide, aus denen die Makromoleküle gebildet werden (Abbildung 1):

Zucker können als Einzelmoleküle auftreten, wie Glucose, die im Energiestoffwechsel aller Lebewesen eine wichtige Rolle spielt. Diese Zucker können auch mit sich selbst oder anderen Zuckermolekülen Dimere oder Multimere bilden; wir sprechen von Disacchariden, Oligosacchariden oder Polysacchariden (wie Stärke

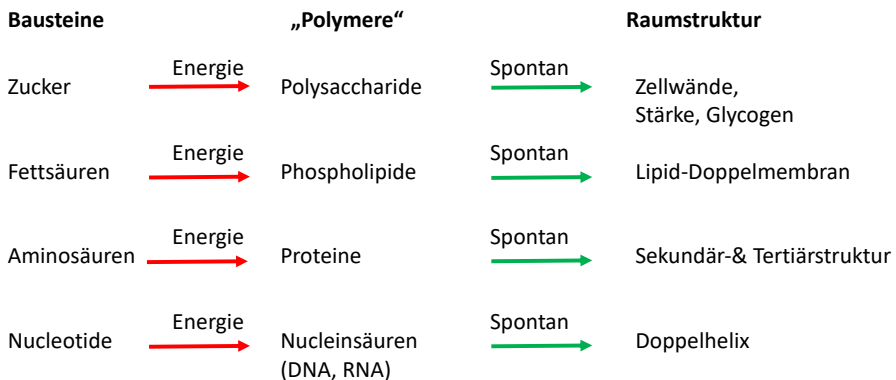


Abbildung 1: Von den Grundbausteinen zu den Makromolekülen und Raumstrukturen.

und Glycogen als Energiespeicher, Zellulose zum Aufbau pflanzlicher Zellwände oder Chitin für pilzliche Zellwände).

Theoretisch gibt es tausende mögliche Aminosäuren, doch sind es nur 20, die zum Aufbau von Proteinen (den wichtigsten Makromolekülen aller Zellen) verwendet werden. Grundsätzlich gibt es bei Aminosäuren zwei mögliche Spiegelbildformen, die als L- oder D-Aminosäuren bezeichnet werden. In den Proteinen unserer Zellen werden ausschließlich L-Aminosäuren eingebaut. Nur in besonderen Spezialmolekülen der Bakterien findet man auch D-Aminosäuren. Die zugehörigen Peptide sind meist resistent gegen Abbau durch Proteasen, die auf Proteine mit L-Aminosäuren ausgerichtet sind.

Auch der Aufbau der Fettsäuren, die zum Aufbau der Membranlipide oder als Speicherfette dienen, ist sehr ähnlich. Viele haben eine Kettenlänge von 16 oder 20 Kohlenstoffatomen. Fettsäuren sind Bestandteile der Phospholipide, die sich in einem wässrigen Milieu spontan zu einer dichten Lipid-Doppelmembran zusammenlagern und die wiederum spontan Kugelform einnehmen.

Nucleotide, wie ATP, GTP, CTP, TTP und UTP sind bei allen Organismen identisch aufgebaut und dienen zum Aufbau der Erbsubstanz, die bei allen Lebewesen aus Nucleinsäuren, meist DNA besteht. Ferner besitzen alle Zellen RNA, die als mRNA, ribosomale RNA und etlichen anderen RNAs für die Prozesse der Transkription und Translation notwendig sind. Nucleotide sind energiereiche Substanzen und ATP wird generell als energieliefernde Substanz in der Zelle genutzt.

Außerdem existieren in unseren Zellen diverse kleinere organische Säuren, wie z. B. Essigsäure oder Milchsäure, die in der Regel eine Rolle im Energiestoffwechsel oder Metabolismus spielen.

Aus den Monomeren, also Zucker, Aminosäuren, Phospholipiden und Nucleotiden werden in allen Zellen nach identischen Regeln die Polysaccharide, Proteine und Nucleinsäuren (DNA, RNA) aufgebaut. Diese Aufbauprozesse erfolgen nicht spontan, sondern benötigen Energie und spezifische Enzyme als Katalysatoren. Als Energielieferant wird der universelle Energieträger ATP (oder andere Trinucleotide) verwendet.

Die so energieaufwändig produzierten Makromoleküle liegen in den Zellen nicht als strukturlose Fäden vor, sondern weisen meist komplexe Raumstrukturen (Konformation) auf. Diese Ausbildung der Raumstrukturen ist durch die Struktur und Reihenfolge der Monomeren bedingt und erfolgt quasi spontan. Wie bereits

erwähnt, bilden die Phospholipide spontan eine Lipiddoppelschicht aus, die sich zu kugelförmigen Bläschen organisieren. Wir können auch von einer automatischen Selbstorganisation der Makromoleküle sprechen (Eigen und Schuster 1979). D.h. die Bauanleitungen für die übergeordneten Strukturen liegen bereits in der Grundstruktur der Moleküle und benötigen keinen eigenen Baumeister.

4 RNA-Welt

Da alle Lebewesen über dieselben Bausteine und Makromoleküle verfügen, nimmt man an, dass das Leben, so wie wir es kennen, nur einmal entstanden sein kann. Wäre es parallel an mehreren Stellen entstanden, müsste es eine größere Variabilität geben, denn aus Sicht der Chemie wären durchaus auch andere Substanzen möglich gewesen.

Alle Lebewesen sind von einer Biomembran, also der Lipid-Doppelschicht umgeben, die sich spontan aus Phospholipid-Molekülen ausbildet. Diese Hülle ist notwendig, um alle die zellulären kleinen und großen Moleküle auf engem Raum zusammenzuhalten. Denn eine Dichte der Grundbausteine und Enzyme ist notwendig zum Aufbau der Makromoleküle und für den Stoffwechsel; in einer verdünnten wässrigen Welt würden sich die Partner niemals finden.

Für den energieaufwändigen Aufbau der Makromoleküle und für den Metabolismus werden heute Enzyme benötigt. Man nimmt jedoch an, dass Proteine in der Ursuppe noch fehlten. Jedoch könnten sich Nucleinsäuren, wie RNA, damals spontan gebildet haben. RNAs sind nicht nur Träger an genetischer Information, sondern haben auch katalytische Eigenschaften. Solche katalytischen RNAs könnten die Synthese von Proteinen und Phospholipiden bewerkstelligt haben. Nach dieser Hypothese startete das Leben demnach mit einer RNA-Welt (Alberts et al., 2015). Erst später kamen die DNA-Welt und die Proteine hinzu.

Aber woher kamen die Bausteine der Makromoleküle, wie Zucker, Fettsäuren, Aminosäuren und Nucleotide? Chemische Experimente, in denen einfache Kohlenstoff- und Stickstoffverbindungen einer thermisch-elektrischen Reaktion ausgesetzt wurden (wie z. B. bei einem Blitzeinschlag), zeigen, dass dabei kleine organische Moleküle entstehen, wie z. B. Aminosäuren oder DNA-Basen. In allen heute lebenden Zellen werden komplexe Biosynthesewege benötigt, um diese Grundbausteine aufzubauen. Daher führen die chemischen Versuche mit der Ursuppe nur bedingt weiter.

Da wir eigentlich nicht so recht wissen, wie man die Universalität der biochemischen Prozesse und ihre Entstehung erklären kann, tendieren einige Wissenschaftler dazu anzunehmen, dass das Leben nicht auf unserem Planeten entstand, sondern dass die Erde vor 3,5 Milliarden Jahren mit Bakterien aus einem anderen erdähnlichen System (von denen es über 10.000 geben soll) mittels Meteoriteneinschlag beimpft wurde (Alberts et al., 2015). Diese Hypothese würde die Einheitlichkeit des Lebens auf unserer Erde auf einen einzigen Ursprung elegant zurückführen. Was jedoch dann bleibt, ist die ungelöste Frage, wie das Leben denn in dem extraterrestrischen System entstand.

5 Bakterien waren zuerst da

Erste Fossilien von einem bakterienähnlichen Lebewesen sind, wie erwähnt, rund 3,5 Milliarden Jahre alt und damit 1 Milliarde Jahre jünger als die Erde. Diese Fossilien zeigen vermutlich Bakterien, die bereits alle grundlegenden Merkmale einer Zelle aufweisen müssten, also Biomembran und vermutlich DNA sowie Proteine. Es ist nicht unwahrscheinlich, dass diese ersten Bakterien schon in der Lage waren, Photosynthese zu betreiben, so wie wir es heute noch von den Cyanobakterien (früher Blaualgen genannt) kennen (Alberts et al. 2015; Storch et al. 2013; Wink 2020). Wenn diese Fossilien korrekt interpretiert wurden, dann war vor 3,5 Milliarden Jahren die Grundlage für die Universalität der Zellstruktur und der Moleküle des Lebens bereits festgelegt.

Unter den Bakterien unterscheidet man die Domänen der Bacteria und Archaea (Alberts et al. 2015; Storch et al. 2013; Wink 2020). Obwohl der Name Archaea impliziert, dass die Bakterien, die an Extremlebensräume wie heiße Quellen oder Salzseen angepasst sind, ursprünglich und alt sein müssten, haben Analysen ihrer Genome gezeigt, dass sie sich von den Bacteria ableiten (Abbildung 2).

Bakterien sind einzellige Lebewesen mit ähnlichem Aufbau (Abbildung 3; Alberts et al. 2015; Storch et al. 2013; Wink 2020). Wie alle Zellen sind sie von einer dichten Biomembran umschlossen, die als Diffusionsbarriere verhindert, dass Moleküle unkontrolliert die Zelle verlassen oder in diese eindringen. Nach außen hin wird die Bakterienzelle durch eine Zellwand geschützt, die verhindert, dass Bakterien aus osmotischen Gründen durch Wasseraufnahme platzen. Außerdem schützt diese Zellwand gegen den Angriff durch andere Bakterien oder Protozoen. Bei gram-positiven Bakterien handelt es sich um eine relative dicke Schicht aus

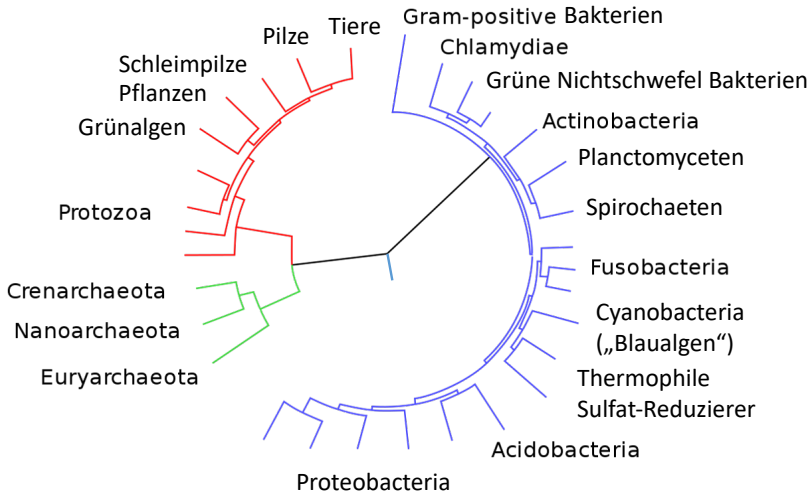


Abbildung 2: Genomischer Stammbaum des Lebens mit den Domänen der Bacteria (blaue Äste), Archaea (grüne Äste) und Eukaryoten (rote Äste).

Murein, während gram-negative Bakterien zusätzlich eine weitere Membran nach außen tragen.

Bakterien haben keine internen Kompartimente, also durch Membranen abgetrennte Reaktionsräume, wie sie typischerweise bei den Eukaryoten auftreten (Abbildung 4). Alle Lebewesen haben Nucleinsäuren (DNA) als Träger der Erbinformation. Die genetische Information liegt bei Bacteria und Archaea als ein ringförmiges Chromosom vor, auf dem zwischen 500 und 8.000 Gene angeordnet sind. Da es keinen Zellkern gibt, zählt man Bacteria und Archaea zu den Prokaryoten. Zusätzliche Erbinformation kann bei Bakterien auf ringförmigen Plasmiden liegen, so z. B. die Gene für Antibiotika-Resistenzen. Diese Plasmide können durch Konjugation zwischen Bakterien derselben Art aber auch nah verwandten Arten ausgetauscht werden. Eine Bakterienzelle pflanzt sich ohne Sexualität fort. Das heißt es gibt keine Keimbahnzellen, sondern jede Zelle teilt sich munter weiter, soweit es ausreichend Nahrung gibt. Bei guter Kondition ist dies alle 20 Minuten der Fall. Vor jeder Zellteilung muss die DNA verdoppelt werden, wir sprechen von der Replikation, deren Grundmechanismus von den frühen Bakterien

bis zum Menschen konserviert wurde und gleichermaßen abläuft (Alberts et al. 2015; Storch et al. 2013; Wink 2020).

Gene kodieren für Proteine. Gene müssen mittels RNA-Polymerase zu mRNA transkribiert werden (daher der Name Transkription für den Prozess). Die mRNA wird in komplex aufgebauten Ribosomen als Matrize genommen, um Proteine aufzubauen. Dem liegt ein Triplet-Kode zugrunde, d. h. für jede der 20 Aminosäuren gibt es ein bis sechs Triplet-Kodes. Auch dieser genetische Kode gilt universell (bis auf kleine unbedeutende Ausnahmen). Ribosomen bestehen aus ribosomaler RNA als Gerüstsubstanz und fast 50 Proteinen, die mRNA binden, Komplexe von tRNA mit Aminosäuren erkennen und Aminosäuren miteinander verknüpfen. Dieser universelle Prozess wird Translation genannt. Ribosomen könnte man auch als Nano-Maschinen betrachten, die ständig abertausende Proteinmoleküle in einer Zelle synthetisieren.

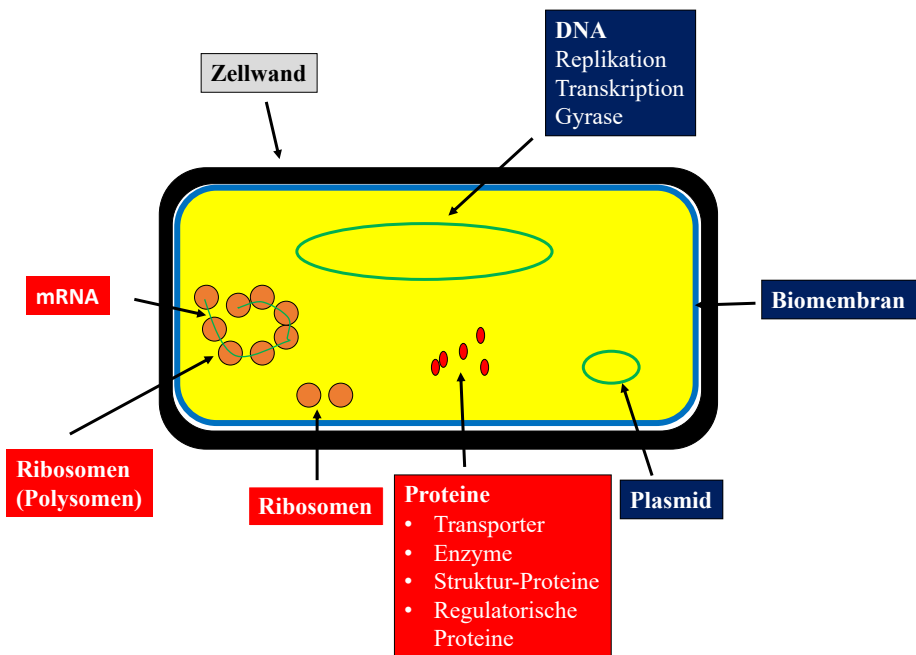


Abbildung 3: Schematischer Aufbau eines gram-positiven Bakteriums.

Proteine sind für das Leben und die Zelle essentiell, denn sie sind die eigentlichen Akteure. Proteine dienen als Enzyme, um Substanzen (kleine Primärstoffe, Makromoleküle) selektiv aufzubauen (Anabolismus) oder aus vorhandener organischer Materie Energie zu gewinnen (Katabolismus). Proteine dienen auch als Membran-Proteine, wie Ionenkanäle, Rezeptoren und Transporter oder bauen das Zytoskelett auf (Aktin, Mikrotubuli).

Eine Zelle kann nur leben, wenn Nahrung zur Verfügung steht, die anaerob zur Milchsäure oder Ethanol oder aerob über den Trikarbonsäure-Zyklus zu CO₂ abgebaut wird. Bei diesen Reaktionen wird ATP hergestellt, das als universelle Energiesubstanz (also quasi wie Benzin für Verbrennungsmotoren) für alle energieverbrauchenden Prozesse der Zelle genutzt wird (Abbildung 1). Wird kein ATP nachgeliefert, so stirbt eine Zelle normalerweise bald ab. Auch die Nutzung von ATP (oder verwandten Nucleotiden) ist bei allen Lebewesen seit wohl 3,5 Milliarden Jahren gleichgeblieben.

Was bisher gesagt wurde, gilt für die Grundversorgung der Bakterien. Nun haben sich Bakterien und Archaea über 3 Milliarden Jahren an die unterschiedlichsten Lebensräume und Lebensbedingungen angepasst, in dem vorhandene Gene verdoppelt und mutiert und danach für neue Funktionen genutzt wurden. Es gibt heute sicher fast keinen Lebensraum der Erde, sei es in aquatischen, marinen oder terrestrischen Ökosystemen oder im Innern der Lebewesen (Mikrobiom), der nicht von hunderten unterschiedlichen Bakterienarten genutzt wird. So leben selbst in unserem Körper im Darm, aber auch auf allen Schleimhäuten und der Haut wohl über 1.000 unterschiedlichen Mikrobenarten, die uns meistens nutzen (als Symbionten produzieren sie z. B. Vitamine), aber auch manchmal schaden (Pathogene). Die Individuenzahl an Mikroben in unserem Körper übersteigt wohl die Zahl der Körperzellen (sie liegt bei 10¹⁴) (Alberts et al. 2015; Storch et al. 2013; Wink 2020).

Die Anzahl der beschriebenen Bakterienarten liegt bei 10.000 Arten; es wird jedoch vermutet, dass die wahre Zahl eher bei 100.000 bis 1 Million Arten liegen müsste. Die Zahl der Archaea ist übersichtlicher und beträgt rund 260 Arten (Alberts et al. 2015; Storch et al. 2013). Aus Sicht der Evolution kann man sagen, dass sich das Bauprinzip der Bakterien als äußerst erfolgreich erwiesen hat und Bakterien zu den individuenreichsten und erfolgreichsten Lebewesen unseres Planeten gemacht hat und das ganz ohne Gehirn und Sex.

6 Von den Prokaryoten zu den Eukaryoten und mehrzelligen Lebewesen

Wie erwähnt, gab es vor rund 3,5 Milliarden Jahren schon Bakterien, die fast 2 Milliarden Jahre lang alleine die Erde bewohnten. Man hat spekuliert, dass vor rund 1,8 Milliarden Jahren der Übergang von den Bakterien zu den Eukaryoten (Abbildung 4) begonnen haben müsste (Alberts et al. 2015; Storch et al. 2013). Wie könnte der Prozess abgelaufen sein?

Evolutionsbiologen postulieren ein Bakterium oder Archaeobakterium als Ausgangszelle. Zunächst wird sich die Zellmembran nach innen hin eingestülpt haben. Diese Schläuche werden vermutlich das zunächst ringförmige Chromosom ein-

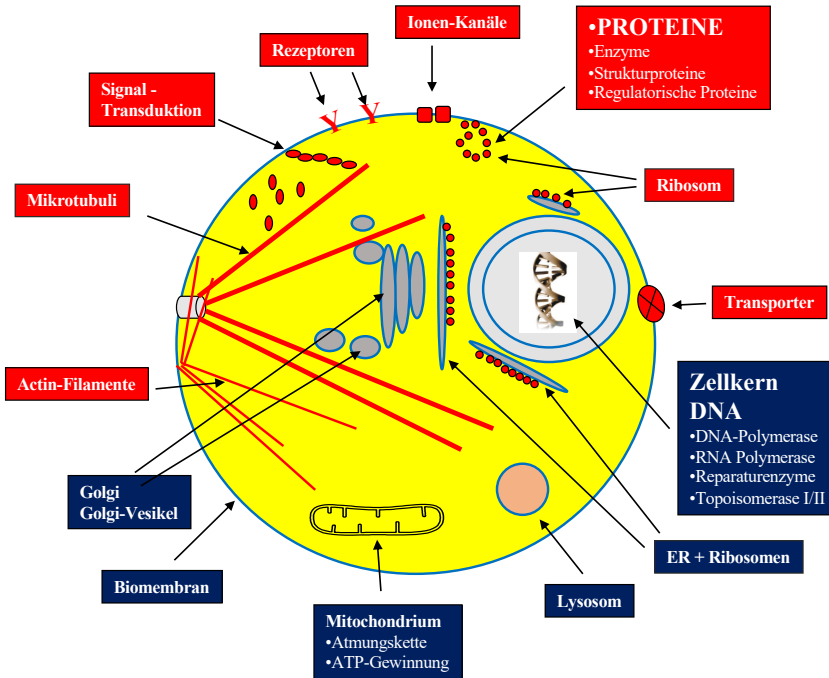


Abbildung 4: Schematische Struktur einer Tierzelle.

gehüllt haben und sich danach von der Zellmembran gelöst haben. So entstand der Zellkern, der bei allen Eukaryoten von einer Kernhülle umgeben und der Sitz der Chromosomen ist. Aus anderen Membran-Abschnürungen entstand das endoplasmatische Retikulum (ER). Auf diese Weise wurden in der frühen Euzyte neue Reaktionsräume geschaffen, die es ermöglichten, dass Biosynthesewege ungestört voneinander ablaufen konnten. Vor rund 1,8 Milliarden Jahren soll eine solche Euzyte sich ein aerobes alpha-Proteobakterium eingefangen und domestiziert haben. Dieses Bakterium wurde beim Eindringen in die Zellen von der Zellmembran der Euzyte umhüllt und gelang so in das Zellinnere. Dort lebte es vermutlich eine Zeitlang als interner Parasit weiter, bis es irgendwann domestiziert wurde. Aus dem Bakterium wurde ein zelluläres Organell, das Mitochondrium, dessen Hauptaufgabe es ist, die Energieversorgung der Zelle, d. h. Bildung von ATP über die Atmungskette, zu gewährleisten. Die frühen Mitochondrien haben einen großen Teil ihrer genetischen Autonomie abgegeben; diese ging nicht verloren, sondern landete im Genom der Euzyte. Mitochondrien haben jedoch ein kleines ringförmiges Chromosom behalten, das aber nur noch über weniger als 50 Gene verfügt. Diese Gene werden repliziert, transkribiert und mittels mitochondrieneigenen Ribosomen in Proteine translatiert. Da die bakterielle Herkunft der Mitochondrien genetisch gesichert ist, sprechen wir auch von der Endosymbionten-Theorie. Es handelt sich zudem um den ersten bekannten „horizontalen Gentransfer (HGT)“ (Alberts et al. 2015; Storch et al. 2013; Wink 2020).

Aus dieser Ur-Eukaryotenzelle entwickelten sich diverse Formen von Einzellern (auch Protozoen genannt), die vor rund 900 Millionen Jahren zu den Vielzellern, also Pilzen und Tieren weiterführten (Abbildung 5; Alberts et al. 2015; Storch et al. 2013; Wink 2020).

Nach Vereinnahmung der Mitochondrien ging die zelluläre Differenzierung weiter, in dem eine weitere Endosymbiose stattfand. Durch Aufnahme von photosynthesetreibenden Cyanobakterien erhielt die frühe Eukaryotenzelle die Fähigkeit zur Photosynthese. Die Mechanismen erfolgten wohl ähnlich wie bei den Mitochondrien. Aus diesen neuen Zellen entwickelten sich dann die photosynthetisch aktiven ein- und mehrzelligen Algen und vor rund 600–500 Millionen Jahren die Landpflanzen (Alberts et al. 2015; Storch et al. 2013; Wink 2020).

Ein wesentliches Merkmal der Eukaryotenzelle (Abbildung 4) ist die Schaffung von diversen Reaktionsräumen durch Kompartimentierung. Im Zellkern sitzen die Chromosomen und hier laufen die wichtigen Prozesse wie Replikation und

Transkription ab. Die mRNA wird über komplex aufgebaute Porenkomplexe aus dem Kern in das Zytoplasma transportiert. Dort verbindet sie sich mit Ribosomen, die als eine abgewandelte Version der bakteriellen Ribosomen vorliegen. Entweder werden die fertigen Proteine direkt in das Zytoplasma entlassen oder aber die Ribosomen binden an das ER und sezernieren die frisch produzierten Proteine in das Lumen oder die Membran des ERs. Nach einigen Modifikationen (Anhängen von Zuckerketten) werden Vesikel aus dem ER abgeschnürt, die sich danach mit dem Golgi-Apparat vereinen. Im Golgi-Apparat erfolgen weitere Modifikationen und letztlich werden Vesikel abgeschnürt, die zur Zellmembran oder zum Lysosom (Ort des Abbaus von zellulärem Abfall) transportiert werden. An den Membranen der Zelle oder des Lysosoms kommt eine zu einer Verschmelzung der Membranen und der Inhalt der Vesikel wird in das Lysosom oder in die Zellumgebung freigesetzt. Die Membranen der Kernhülle, des ERs und des Golgi-Apparat stellen so etwas wie ein kontinuierliches Endomembransystem dar, da diese Membranen miteinander fusionieren können oder sich als Vesikel wieder abschnüren können. Das Zytoplasma wird durch ein Netzwerk von Proteinen des Zytoskeletts (Aktinfilamente, intermediäre Filamente, Mikrotubuli) stabilisiert, die bei der Zellteilung eine essentielle Rolle spielen. Die Biomembran ist wie bei den ursprünglichen Bakterien eine Permeationsschranke und weist viele Membranproteine auf, unter denen Ionenkanäle, Transporter, und Rezeptoren am wichtigsten sind.

Die hier aufgeführten Eigenschaften (Alberts et al. 2015; Storch et al. 2013; Wink 2020) zählen zum Grundrepertoire aller Zellen und sind universell, wiederum ein Hinweis darauf, dass alle Zellen in einer sehr langen Kette von Vorfahrzellen stehen. Schon Richard Virchow (1829–1902) hatte 1855 erkannt, dass Zellen niemals *de novo*, sondern immer durch Teilung aus einer Mutterzelle entstehen können. Er prägte den Leitsatz „Omnis cellula e cellula“ und begründete damit die Zellular-Theorie.

Im Verlauf mehrerer hundert Millionen Jahre der frühen Evolution wurden diverse Zellspezialisierungen erfunden, die man bis heute bei den über 200.000 unterschiedlichen Arten von Einzellern (Protozoen, wie Pantoffeltierchen, Plasmodien oder Algen) beobachten kann. Die Möglichkeiten sind für einen Einzeller jedoch beschränkt, sodass die Evolution der Vielzelligkeit bei Pflanzen und Tieren eine Schlüsselinnovation darstellt (Abbildung 5). Denn ein mehrzelliger Organismus kann mehr als einen Zelltyp mit unterschiedlicher Funktionalität ausbilden. Sind mehrere unterschiedliche Zelltypen vorhanden, kann eine Arbeitsteilung

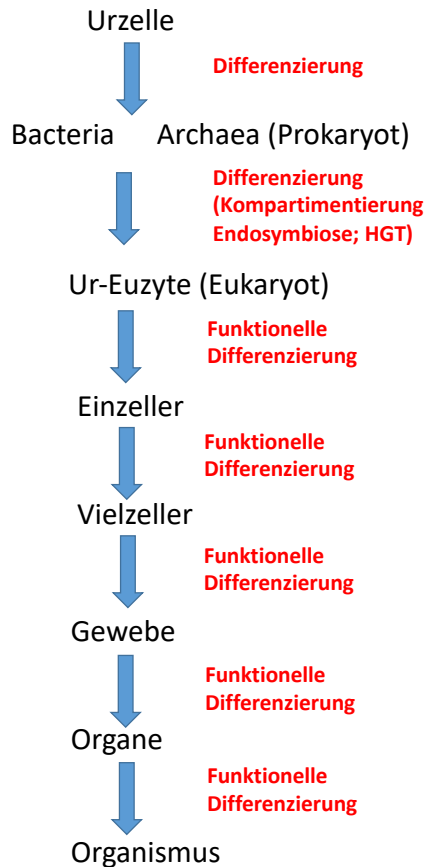


Abbildung 5: Von der Urzelle zum komplexen vielzelligen Organismus.

stattfinden und somit komplexere Aufgaben oder Anpassungen möglich werden. Ähnlich differenzierte Zelle bilden Gewebe (wie zum Beispiel Epithelien), die wiederum mit anderen Zellen sich zu Organen organisieren. Organe mit unterschiedlicher Funktion (Leber, Darm, Magen, Herz, Gehirn, Lunge, Keimdrüsen) arbeiten koordiniert als Organismus zusammen. Die Entwicklung zu differenzierten Zellen und Geweben wird durch Transkriptionsfaktoren (sie sorgen dafür, dass die „richtigen“ Gene aktiviert werden) gesteuert, die modular und hierarchisch

organisiert sind. Auch diese Differenzierungs- und Entwicklungsmoleküle sind universell und funktionieren bei den unterschiedlichen Organismen nach demselben oder ähnlichem Prinzip (Alberts et al. 2015). Zellen kommunizieren über ein Netzwerk von Signalmolekülen und Rezeptoren miteinander, sodass sie „wissen“, wo ihr Platz in einem Gewebe/Organ ist, wann und wie oft sie sich teilen müssen. Wenn dieses Netzwerk gestört wird, entsteht ein ungeordnetes Wachstum, wie man es bei Tumoren vorfindet.

Alle Zellen eines Körpers verfügen über dasselbe Genom. Bei den Wirbeltieren, also auch bei uns Menschen gibt es rund 30.000 Gene, von denen 21.000 für Proteine kodieren, die übrigen sind RNA-Gene (Alberts et al. 2015; Storch et al. 2013; Wink 2020). Die Gameten, also Spermien und Eizellen, gelten als undifferenziert. Nach Vereinigung der Gameten entsteht eine omnipotente Zygote (Stammzelle), aus der sich bei fortschreitender Zellteilung differenzierte Zellen entwickeln. In den differenzierten Zellen sind jedoch nicht mehr alle Gene aktiv, sondern nur eine Teilmenge. Die inaktiven Gene sind durch Methylierung und Modifikation der Histone abgeschaltet. Diese Schaltanweisung wird bei der mitotischen Zellteilung an die Tochterzelle weitergegeben. Dieses Phänomen wird als „epigenetische Vererbung“ bezeichnet. Für das selektive Ablesen von Genen werden spezifische Transkriptionsfaktoren, die bei den Eukaryoten weitgehend konserviert wurden.

Die Zahl von 21.000 proteinkodierenden Genen erscheint bei der Vielzahl der zellulären Differenzierungen, die alle auf Proteinen beruhen, erstaunlich gering. Jedoch haben Zellen einen Trick, um die Anzahl und Funktionalität der Proteine zu erhöhen. Bekanntlich bestehen Gene aus kodierenden Bereichen (Exons) und nicht-kodierenden Sequenzen (Introns). Bei der Transkription entsteht eine primäre mRNA, aus der die Introns herausgeschnitten werden (Spleißen). Da jedes Gen aus mehreren Exons besteht, kann die Zelle Variabilität erreichen, indem nicht alle Exons in die fertige mRNA gelangen, sondern nur eine Auswahl. Da Exons für Proteindomänen kodieren, verändert sich durch den Wegfall einer Domäne auch die Funktionalität des Proteins. In der Tat hat sich gezeigt, dass die mRNA eines spezifischen Gens im Gehirn anders gespleißt wird als zum Beispiel in der Leber. Das heißt durch Kombination von Funktionsmodulen kann funktionale Differenzierung erreicht werden.

Bei jeder Zellteilung wird das Genom durch Replikation fast fehlerfrei verdoppelt, sodass die Genome von den frühen Prokaryoten bis zu den Vertebraten viele Ähnlichkeiten aufweisen (Alberts et al. 2015; Storch et al. 2013). Durch Genom-

und Genduplikationen wuchsen die Genome von rund 5 Millionen Basenpaaren bei den Prokaryoten auf mehrere Milliarden Basenpaaren bei den Pflanzen und Vertebraten. Durch diverse Mutationen und durch Rekombination können sich Gene verändern. Diese Veränderungen des Genotyps können einen Einfluss auf den Phänotyp haben. Sind die Auswirkungen negativ, so wird der Träger dieses Genotyps weniger Überlebens- und Fortpflanzungschancen haben und durch die natürliche Selektion wieder verschwinden. Mutationen mit positiver Wirkung auf die Fitness werden dagegen dazu führen, dass ihr Träger einen besseren Erfolg hat und mit etwas Glück werden seine Gene an die nächsten Generationen weitergegeben. Schon früh in der Evolution wurde die zweigeschlechtliche Fortpflanzung erfunden, die den Vorteil hat, dass die genetische Variabilität gewaltig erhöht wird (Alberts et al. 2015; Storch et al. 2013; Wink 2017). Bei der Reifungsteilung der Gameten (Meiose) und bei der nachfolgenden Vereinigung der Gameten werden die mütterlichen und väterliche Gene kräftig gemischt. Somit erhöhen Mutationen und Sexualität die genetische Variabilität, die offenbar für die Entstehung der Millionen Pflanzen und Tieren notwendig war, die in unterschiedlicher Weise an ihre Umwelt angepasst sind.

7 Entwicklungslinien der Phylogenese

Durch die Genomforschung können wir heute rekonstruieren, wie aus den Einzellern die Entwicklungslinien zu den Pflanzen und Tieren entstanden und wie es durch Differenzierung und Anpassungen zu rund 1,9 Millionen Arten kam, die wir heute kennen (Alberts et al. 2015; Storch et al. 2013).

Aus den aquatischen Grünalgen entstanden die ersten sporenbildenden Landpflanzen, den Leber- und Laubmoosen (Abbildung 6; Storch et al. 2013). Als nächste Innovation wurden Leitgewebe ausgebildet, die es Pflanzen erlaubt, aufrecht zu wachsen. In diesem Zusammenhang entstanden die Farne, Schachtelhalme und Bärlappgewächse, die bis heute überlebt haben. Der nächste Innovationsschritt war die Umstellung der Fortpflanzung durch Sporen durch die Ausbildung von Samen. Zunächst entstanden die Nacktsamer, wie Ginkgo, Nadelbäume und Palmfarne, von denen wir heute noch rund 1.000 Arten existieren. Vor rund 150 Millionen Jahren begann die Entwicklung zu den Blütenpflanzen, die heute mit über 370.000 Arten die vorherrschende Pflanzengruppe darstellen (Storch et al. 2013).

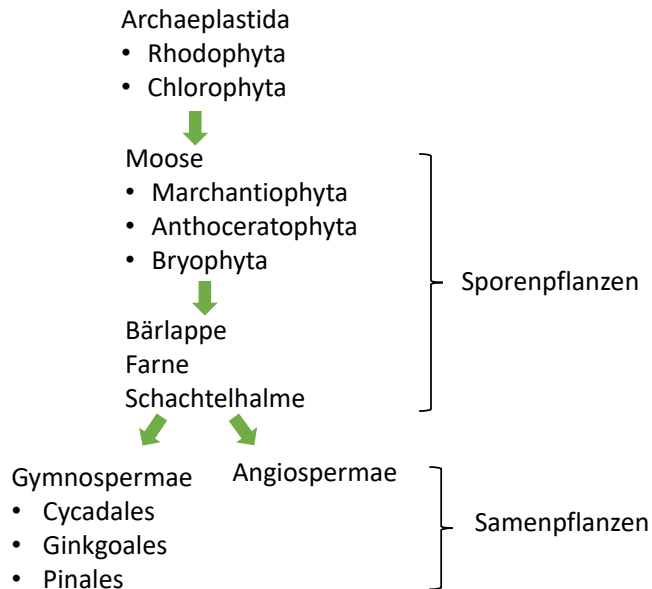


Abbildung 6: Von den Algen zu den Blütenpflanzen.

In der Entwicklung zu den Tieren beginnen wir mit einfachen Choanoflagellaten (Abbildung 7), aus denen sich einfache Tierformen wie Schwämme, Nesseltiere und Quallen entwickelten (Alberts et al. 2015; Storch et al. 2013). Von den radiär aufgebauten Nesseltieren (Radiata) entstanden die bilateral und segmental aufgebauten Tiere (Bilateria). Innerhalb der Bilateria, wurden zwei große Entwicklungslinien zu den Wirbellosen (Protostomia mit den Untergruppen Ecdysozoa und Lophotrochozoa; rund 1,4 Millionen Arten) und den Wirbeltieren (Deuterostomia; rund 80.000 Arten).

Wie entstanden die Arten? Ein wichtiges Moment der Artbildung ist die Trennung von genetischen Entwicklungslinien (Abbildung 8). Im Regelfall war dies eine räumliche Trennung (allopatrische Artbildung), aber auch eine ökologische Trennung ist möglich (parapatrische und sympatrische Artbildung; Storch et al., 2013). Durch die Trennung der Entwicklungslinien können diese sich unabhängig voneinander differenzieren, denn die ständig ablaufenden Mutationen erfolgen unabhängig und werden bei räumlicher Isolierung nicht wieder vermischt. Durch

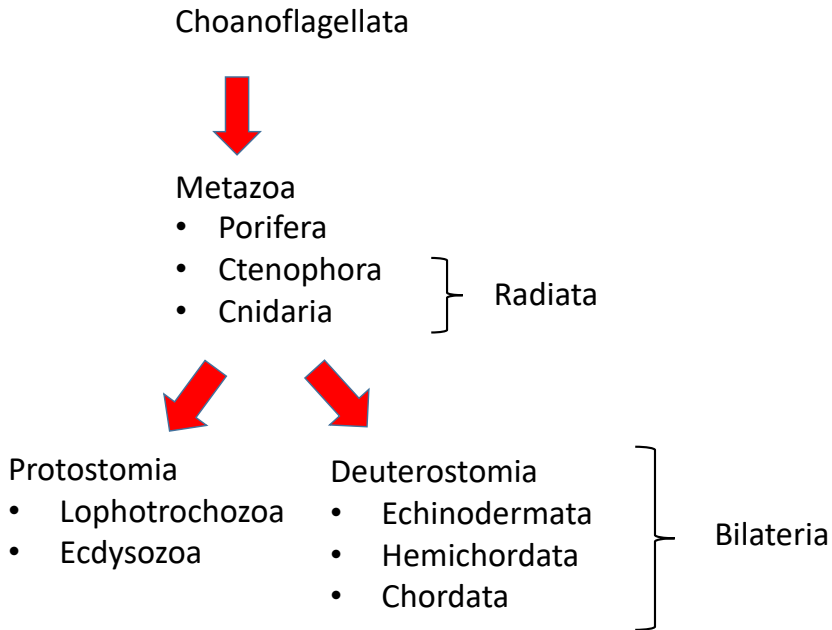


Abbildung 7: Von den Choanoflagellaten zu den Wirbeltieren (=Chordata).

Hybridisierung und horizontalen Gentransfer können etablierte Linien aber lokal vermischt werden, sodass sich die Phylogenese letztlich nicht als ein Stammbaum, sondern als ein Netzwerk darstellt. Durch große klimatische Veränderungen, wie z. B. den regelmäßigen Wechsel von Warm- und Kaltzeiten in den letzten 3 Millionen Jahren, kam zu Arealausweitungen in den Warmzeiten und Arealverkleinerung in den Eiszeiten mit Hybridisierung in den Refugialgebieten.

8 Ausblick

Obwohl Bakterien, Pflanzen, Würmer und Säugetiere sehr unterschiedlich aussehen, sind die Grundstrukturen und Grundprozesse des Lebens universell (d. h. bei allen Lebewesen gleich oder ähnlich) und modular aufgebaut (Abbildung 8).

Grundeinheit der lebenden Natur ist die Zelle, die alleine oder aber im Verbund auftritt. Alle heute existierenden Zellen sind durch Zellteilung aus einer Mutter-

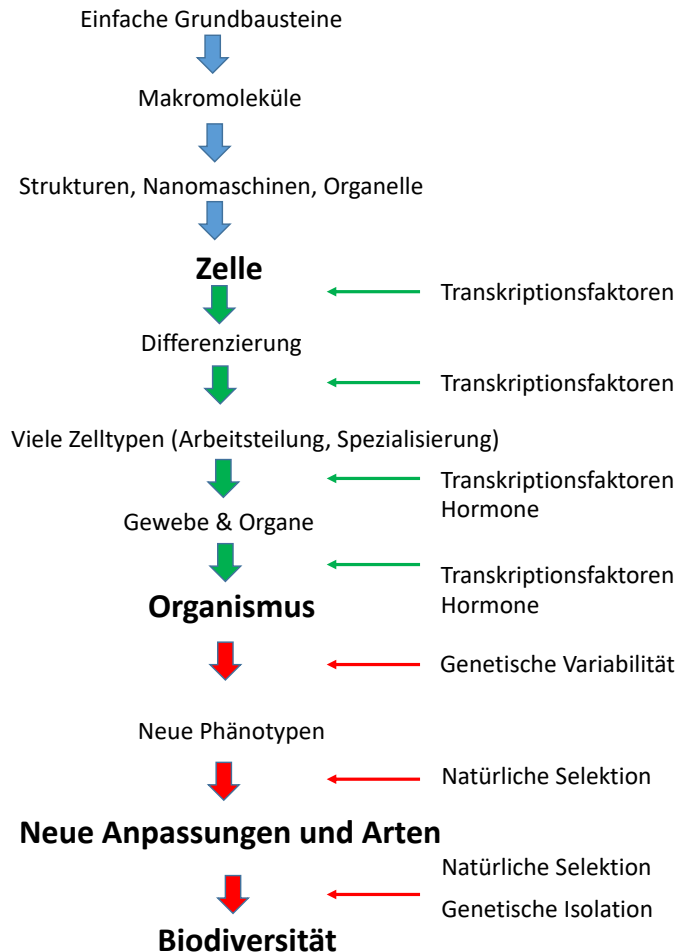


Abbildung 8: Von den einfachen Bausteinen zur Biodiversität.

zelle entstanden, sodass die Zellen in einer kontinuierlichen Linie von heute bis zur Urzelle, die vor rund 3,5 Milliarden Jahre existierte, verbunden sind. Bei den mehrzelligen Organismen differenziert die Zelle zu unterschiedlichen Zelltypen, die quasi als Arbeitsteilung verschiedene Funktionen ausüben können. Zellen bilden Gewebe und diese wiederum komplex aufgebaute Organe mit übergeordneten Funktionen, die das Leben eines höheren Organismus ermöglichen. Alle Prozesse

werden letztlich durch das Genom gesteuert, das seit Beginn einerseits konservativ beibehalten, andererseits durch Mutationen und sexuelle Rekombination ständig abgeändert wird. Der entstehende Phänotyp unterliegt der natürlichen Selektion, sodass die Träger von Merkmalen, die sich nicht bewähren, wieder verschwinden. Durch räumliche und ökologische Trennung entstanden (und entstehen weiter) so die diversen Arten auf der Erde, die in der Regel diverse Anpassungen an ihre Umwelt zeigen, mit denen sie im jeweiligen Lebensraum am besten überleben und fortpflanzen können. Denn der evolutionäre Erfolg eines Lebewesens zeigt sich nicht nur darin zu überleben, sondern seine Gene in nachfolgende Generationen weiterzugeben.

Biochemiker und Evolutionsbiologen können die Entstehung des Lebens und die Phylogenese zwar beschreiben, jedoch lange noch nicht, trotz Erfolge der Genom- und diversen anderen „Omics-Forschungen“, im mechanistischen Detail erklären. Dennoch lassen sich heute schon die Grundprinzipien erkennen, wie Zellen, Organismen und die Biodiversität entstanden sein können.

Literatur

- Alberts, B. et al. (2015) *Molecular Biology of the Cell*, 6. Aufl.; Garland Science, New York, NY, USA.
- Darwin, C. (1859) *On the Origin of Species*. John Murray, London.
- Eigen, M., Schuster, P. (1979) *The Hypercycle – A Principle of Natural Self-Organization*. Springer, Berlin.
- Storch, V., Welsch, U., Wink, M. (2013) *Evolutionsbiologie*. 3. Auflage; Springer, Heidelberg.
- Wink, M. (2017) Sex als Motor der Evolution. Warum es zwei Geschlechter gibt. In *FRAU und MANN; Forschungsmagazin Ruperto Carola*, 10, 32–39.
- Wink, M. (2020) *An Introduction to Molecular Biotechnology. Molecular fundamentals, methods and applications in modern biotechnology*. 3. Ed., Wiley-VCH, Weinheim.

Über den Autor

Prof. Dr. Michael Wink ist Ordinarius für Pharmazeutische Biologie an der Universität Heidelberg, wo er seit 1999 die Abteilung Biologie am Institut für Pharmazie und Molekulare Biotechnologie leitete. Seit 1.10.19 arbeitet er als Seniorprofessor an der Universität Heidelberg. Nach dem Studium der Biologie und Chemie an der Universität Bonn forschte er in Braunschweig, Köln, München und Mainz. Seine Arbeitsgebiete reichen von Phytochemie, Arznei- und Giftpflanzen, Ornithologie und Naturschutz bis zur Systematik, Phylogenie und Evolutionsforschung. Er ist Autor/Co-Autor von mehr als 20 Büchern und über 1.000 Originalarbeiten. Er ist Gastprofessor an Universitäten in China, Thailand, Argentinien und Mexiko, außerdem Mitglied diverser Wissenschaftlicher Beiräte, Herausgeber einiger Zeitschriften und Empfänger mehrerer Auszeichnungen.

Korrespondenz:

Prof. Dr. Michael Wink
Institut für Pharmazie und Molekulare Biotechnologie (IPMB)
Universität Heidelberg
Im Neuenheimer Feld 329
69120 Heidelberg
E-Mail: wink@uni-heidelberg.de
Homepage: <https://www.uni-heidelberg.de/institute/fak14/ipmb/phazb/>
https://www.researchgate.net/profile/Michael_Wink
<https://www.winks-biology.com/>

Entwicklungen in der Entwicklung – Fortwährende Veränderungen im Fluss der Organismenwelt

CLAUDIA ERBAR & PETER LEINS

Centre for Organismal Studies (COS) Heidelberg – Biodiversität und
Pflanzensystematik, Universität Heidelberg

*Unserem lieben Freund, Herrn Prof. Dr. Ernst Jung, zu seinem 88. Geburtstag in
herzlicher Verbundenheit gewidmet.*

Zusammenfassung

„Die reinste Anschauung der Dinge hat, wer sie vom Anbeginn her wachsen sieht“ (Aristoteles 384 v. Chr. – 322 v. Chr.). Dieser Erkenntnis folgend werden Beispiele aus dem Blütenbereich der Angiospermae (Blütenpflanzen) in dem vorliegenden Artikel vorgestellt. Ähnlich aussehende Strukturen, wie etwa die multistaminierten (aus vielen Staubgefäßen bestehenden) Androeceen in den Blüten unterschiedlicher Verwandtschaftsgruppen oder Blumenkronröhren, die insbesondere im oberen Stammbaumniveau zu finden sind, können sich auf vielfältige Weise entwickeln. Durch Vergleich verschieden alter Stadien machen sich die „individuellen“ Entwicklungsweisen beobachtbar; wir nennen sie Ontogenien. Demgegenüber sind Stammbäume nicht beobachtbare Rekonstruktionen der historischen Verwandtschaft von Lebewesen. Eine stammesgeschichtliche Entwicklung, die Phylogenie, ist also stets Theorie behaftet. Um die Plausibilität einer Stammbaumrekonstruktion zu erhöhen, bedient man sich der evolutionären Erkenntnis, d. h. wir können uns auf die (wiederum beobachtbaren) Faktoren der Evolution (Mutation, Rekombination, Migration,

Isolation und Selektion) verlassen. Jede Veränderung in der Phylogenie macht sich in irgendeinem Entwicklungsstadium einer Ontogenie bemerkbar, letztlich durch eine genetische Veränderung verursacht. Zimmermann (1934) bringt in dem Begriff der Hologenie beide Entwicklungen zusammen: Hologenie ist eine Aneinanderreihung von (unzähligen) Ontogenien in einer Phylogenie.

Entwicklungen bedeuten Veränderungen und Bewegungen in einer bestimmten Zeiteinheit. Zum Teil können Entwicklungen unmittelbar in eine Funktion verwickelt sein. Beispiele finden sich sehr häufig in der Pollenportionierung, einem außerordentlich wichtigen ökonomischen Verhalten der Blüten im Umgang mit den Pollenkörnern. Die Mechanismen sind unterschiedlich. Entwicklungen, Veränderungen, Bewegungen sowohl des Lebewesens als auch in demselben und außerhalb seines Körpers, nämlich in den Ökosystemen, halten Alles im Fluss; *panta rhei*, Heraklit von Ephesos, etwa 500 v. Chr.

1 Einführung, Begriffliches, Methodenentwicklung

Jedes lebensfähige Individuum, sei es ein Einzeller, eine vielzellige Alge, ein Moos, ein Farn oder eine Samenpflanze, sei es ein Regenwurm, eine Wespe oder ein Mensch, macht in seinem Leben eine Entwicklung durch, etwa von einer befruchteten Eizelle ausgehend bis zum **adulten** (ausgewachsenen, sich wieder fortpflanzenden) Organismus. Bei vielen pflanzlichen Organismen, beispielsweise bei allen Landpflanzen (Moose, Farne, Samenpflanzen), ist die Individualität zweigeteilt: Es wechseln **haploide** und **diploide Generationen** ab (Trennung der Meiose von der Gametenverschmelzung durch die Entwicklung eines haploiden Organismus aus den Meioseprodukten). So ist etwa das Pollenkorn einer Blütenpflanze keine Geschlechtszelle, sondern ein Zwergmännchen (**Mikrogametophyt**), das zwei Geschlechtszellen bildet und diese über einen sogenannten Pollenschlauch in einen ebenfalls haploiden weiblichen **Megagametophyten** (sitzt in einer Samenanlage) transportiert (siehe Lebenszyklus einer Blütenpflanze, Abb. 1).

Entwicklung ist immer auch Veränderung und Vergrößerung eines Organismus beziehungsweise eines Teils von diesem. Oberflächliche Veränderungen beispielsweise bei der Entfaltung einer Blüte oder eines Sprosses aus einer gerade aufbrechenden Knospe lassen sich einfach durch fotografische Zeitraffung sichtbar machen. Wir beobachten dabei ein und denselben Bereich ein und desselben Organismus. Eine derartige Zeitraffer-Analyse findet natürlich seine Grenzen, wenn wir ins Detail gehen wollen. Bei der Verfolgung winziger Strukturen müssen wir verschiedene Stadien der Entwicklung an jeweils verschiedenen Organis-

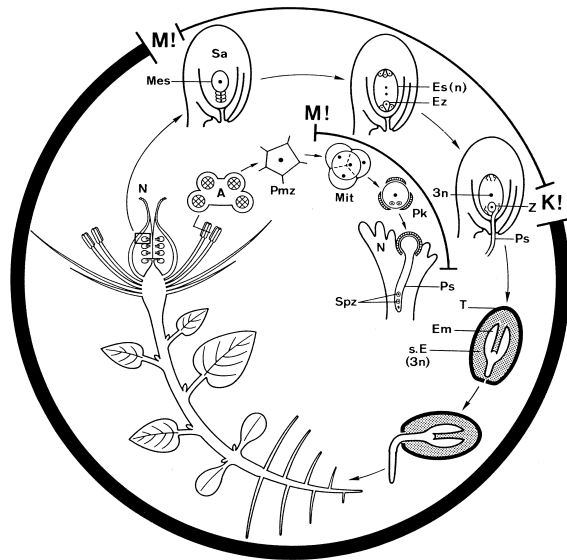


Abbildung 1: Lebenszyklus einer Blütenpflanze in verkürzter Darstellung. A = Anthere (Staubbeutel), Em = Embryo, Es = Embryosack (= weiblicher Gametophyt), Ez = Eizelle, K! = Gametenkopulation, M! = Meiose, Mes = Megaspore, Mit = Mikrosporentetrade, N = Narbe, n = haploid, 3n = triploid, Pk = Pollenkorn, Pmz = Pollenmutterzelle (Mikrosporenmutterzelle), Ps = Pollenschlauch, Sa = Samenanlage, s. E = sekundäres Endosperm, Spz = Spermazellen, T = Testa (Samenschale), Z = Zygote; die diploide Phase (2n) dick, die haploide (n) dünn ausgezeichnet. Aus Leins & Erbar 2008.

men (der gleichen Art) herausgreifen und aneinanderreihen; streng genommen verlassen wir also die Individualität des Untersuchungsobjektes. Dennoch sprechen wir von einer individuellen Entwicklung, der sogenannten **Ontogenie**. Im Gegensatz zu „beobachtbaren“ (mit der genannten Einschränkung) Ontogenien sind stammesgeschichtliche Entwicklungen, die sogenannten **Phylogenie** nicht beobachtbar, sondern auf der Grundlage (beobachtbarer) evolutionärer Prozesse (Mutation, Rekombination, Migration, Isolation und Selektion) nach dem derzeitigen Wissensstand gemachte plausible Rekonstruktionen. Es ist gleichgültig, ob wir unsere phylogenetischen Folgerungen nur auf einzelne Merkmale beziehen (**Merkmalsphylogenie**) oder auf ganze Sippen (**Sippenphylogenie**), die Stam-

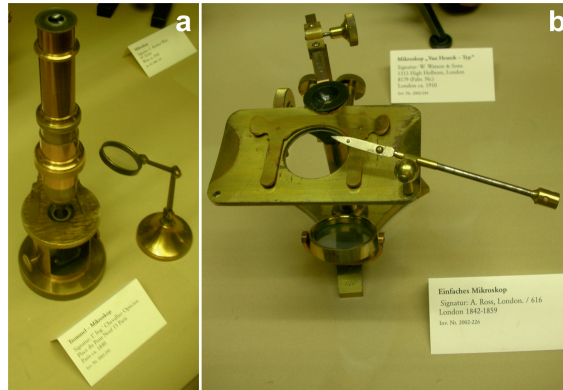


Abbildung 2: Frühe Mikroskope. – a, Trommel-Mikroskop, Paris ca. 1840. – b, Einfaches Mikroskop, London 1842–1859. Eigene Aufnahmen aus dem Deutschen Museum, München.

mesgeschichte wird unter Hinzuziehung aller nur denkbaren Disziplinen und deren Untersuchungsergebnisse eine unerreichbare, wenngleich zunehmend plausibel erscheinende Synthese bleiben (diese Äußerung geht auf Herrmann Merxmüller in den 70er Jahren des vorigen Jahrhunderts zurück; Merxmüller 1972: „Systematic botany – an unachieved synthesis“).

Man darf sehr wohl von einer Sensation sprechen, als 1857 von dem großen französischen Forscher Jean-Baptiste Payer (1818–1860) ein monumentales Werk zur Ontogenie von Blüten erschien („*Traité d’organogénie comparée de la fleur*“). Es grenzt geradezu an ein Wunder, mit welcher Hingabe zur Beobachtung und einer kaum zu überbietenden Exaktheit z. T. winzige Entwicklungsstadien mit dem damals noch eher bescheidenen Instrumentarium der Mikroskopie (Abb. 2) zur Darstellung (Lithographien) kamen (als Kostprobe aus Payer’s Blütenatlas haben wir die Tafel 16 – Blütenentwicklung der Kapuzinerkresse – ausgewählt; Abb. 3).

Inzwischen haben die Lichtmikroskopie und ihr Zubehör (z. B. Zeichenapparate und vieles andere mehr) dieser räumlichen Beobachtungsweise immense Fortschritte eingebracht, so dass die Bildergebnisse immer verlässlicher wurden (Abb. 4).

1973 gelang unserem Freund Rolf Sattler eine Präparationsmethode, Blütenknospen oder Teile von diesen (mit zellulärem Oberflächenmuster) fotografisch festzuhalten (Abb. 5).

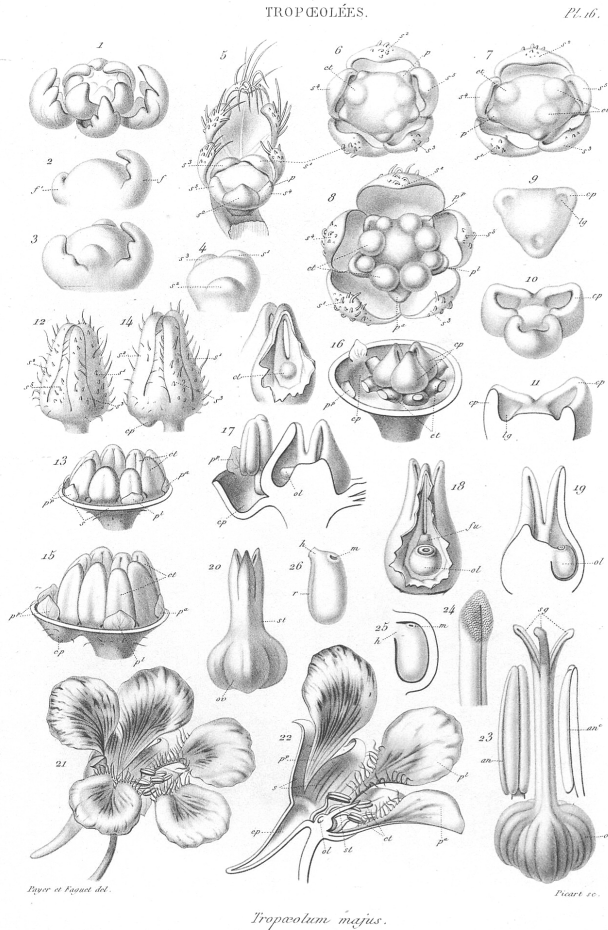


Abbildung 3: Blütenentwicklung der Großen Kapuzinerkresse *Tropaeolum majus* (Kapuzinerkressengewächse, Tropaeolaceae). Tafel 16 von 154 Tafeln aus dem bis heute unübertroffenen Werk „Traité d’organogénie comparée de la fleur“ von Jean-Baptiste Payer (1857), in dem für viele Blütenpflanzenfamilien entwicklungsgeschichtliche Studien präsentiert sind. Payer standen für seine Studien Mikroskope zur Verfügung wie sie in Abb. 2 zu sehen sind. Link für die Tafel: <https://digi.ub.uni-heidelberg.de/diglit/payer1857bd2/0028>. Das gesamte 2-bändige Werk ist von der Universitätsbibliothek Heidelberg digitalisiert worden und ist für jedermann zugänglich unter: Heidelberger historische Bestände – digital <https://www.ub.uni-heidelberg.de/helios/digi/digilit.html>.

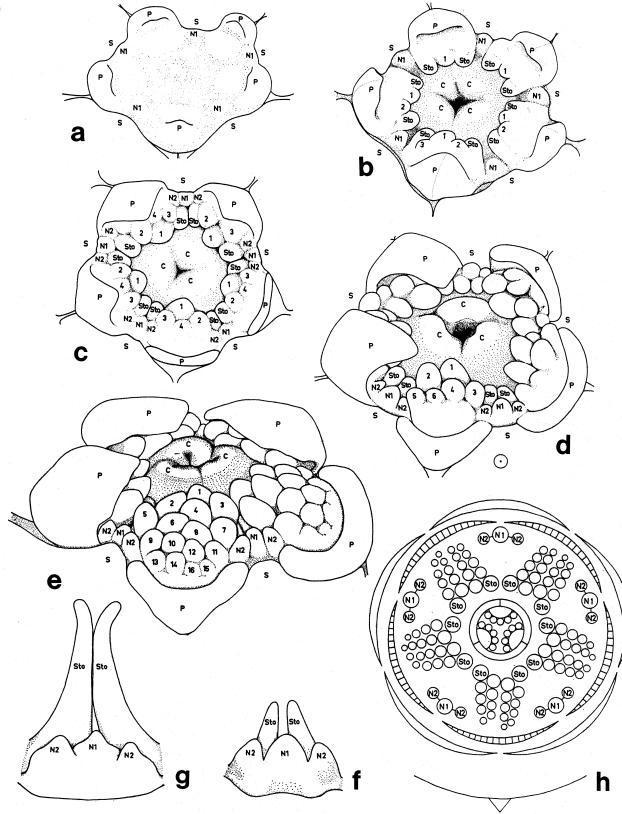


Abbildung 4: Entwicklung des Staminodien-Komplexes bei *Loasa triphylla* (Loasaceae). – a-e, Aufeinanderfolgende Stadien der Androcealentwicklung. Die Zahlen geben die Anlagenfolge der fertilen Staubgefäße wieder. Diese stehen in epipetalen Gruppen (d. h. vor den Kronblättern). Zu diesen gehören zwei abgewandelte sterile Staubgefäße (Staminodien, Sto), die die Androcealentwicklung dieser Gruppen einleiten. Zuvor jedoch wird über den Kelchblättern das mittlere Primordium (N₁) einer dreiteiligen Nektarschuppe angelegt. Etwas später (c) treten zwei laterale Primordien der Nektarschuppe in Erscheinung (N₂). – f-g, Zwei ältere Entwicklungsstadien des episepalen Komplexes (über den Kelchblättern), die die pfriemlichen Staminodien umschließt; letztere werden immer mehr von den sich entwickelnden Staubgefäßen von der Seite gedrängt. – h, Entwicklungsdiagramm. – C = Karpell (Fruchtblatt), P = Petalum (Kronblatt), S = Sepalum (Kelchblatt). Aus Leins & Winhard 1973 zusammengestellt.

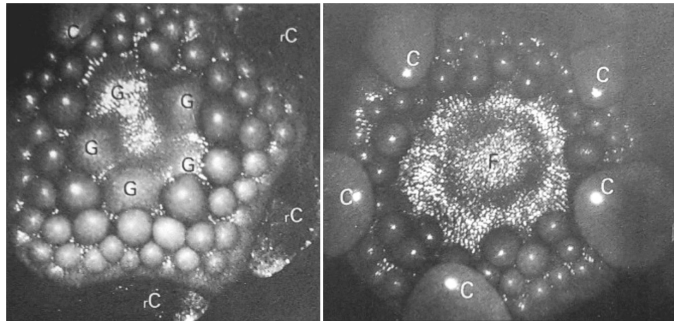


Abbildung 5: Aufsichten auf präparierte Blütenknospen der australischen *Hibbertia scandens* („Kletterndes Münzgold“, Rosenapfelgewächse, Dilleniaceae) mit vielen Staubgefäßen angelegt zwischen Kron- (C) Fruchtblättern (G); F = Blütenvegetationskegel. Die Blütenknospen wurden eingetaucht in 100 % Äthanol fotografiert (aus Sattler 1973).



Abbildung 6: Unterständiger Fruchtknoten. – a–c, Längsschnitte durch Blütenknospen verschiedenen Alters vom Liebstöckel *Levisticum officinale* (a–b) und der Sterndolde *Astrantia major* (c), Apiaceae. – a, Anlage der beiden Fruchtblätter am tellerförmigen Blütenboden. – b, Aufwölbung des gemeinsamen Septums (Pfeil). Das interkalare Wachstum im Blütenboden hat bereits eingesetzt und verschiebt die Basis der Dorsalfanken der Karpelle (im Schnitt median getroffen) nach oben. – c, Ein weit fortgeschrittenes Stadium des Unterständigwerdens. Die Wandung des Bechers besteht aus einem sich kontinuierlich verlängernden diffusen Meristem. – C = Karpell (= Fruchtblatt), P = Petalum (Kronblatt), R = Karpellrand, S = Sepalum (Kelchblatt, reduziert), Sa = Samenanlage, Se = Septum (Trennwand), St = Stamen (Staubgefäß). Der Messbalken entspricht 100 μ m. Aus Leins & Erbar 1985. – d, Blütenstand der Großen Sterndolde (*Astrantia major*, Apiaceae). Der Pfeil weist auf den unterständigen Fruchtknoten einer zwittrigen Blüte. Im Blütenstand sind zwittrige Blüten von funktionell männlichen Blüten (σ) umgeben.

Zur Ermittlung der inneren Vorgänge während der Entwicklung eines Organs, also auf der Zell- beziehungsweise Gewebeebene, bedient man sich Schnittserien durch verschieden alte Stadien. Hierzu stehen heutzutage hochmoderne Mikrotome zur Verfügung. Aufschlussreich sind anatomische (histologische) Schnittbilder beispielsweise für Untersuchungen zum Zustandekommen von unterständigen Fruchtknoten, die, betrachtet man eine Blüte von außen unterhalb der Blütenhülle und der Staubgefäße, als Verdickung (oberhalb eines Blütenstiels) zu erkennen sind (Abb. 6d). Die Unterständigkeit eines Fruchtknotens, der zum Beispiel bei den Doldenblütlern (Apiaceae, Abb. 6d) aus zwei an ihrer Basis gemeinsam aus-

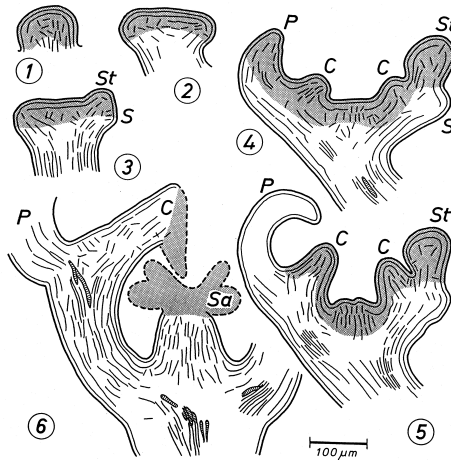


Abbildung 7: Unterständiger Fruchtknoten. Längsschnitte durch verschieden alte Blütenknospen vom Liebstöckel *Levisticum officinale* (Apiaceae). Zur Veranschaulichung der Hauptrichtungen der Zellvermehrung während der Entwicklung sind in den Blütenknospenlängsschnitten (genetische) Zellreihen durch Linien angedeutet. Der Prozess der Becherbildung wird schon sehr früh eingeleitet und beginnt damit, dass die Zellvermehrungsrichtung, die im anfänglich gewölbten Blütenvegetationskegel unter der zweiten Zellschicht hauptsächlich parallel zur Längsachse der Blütenanlage verläuft (1), sich dergestalt ändert, dass sich die Zellen bevorzugt senkrecht zur Längsachse teilen, wobei es zur scheibenförmigen Ausbildung kommt (2). Unter dem Scheibenrand wechselt dann die Zellvermehrungsrichtung in die Diagonale, von der Längsachse nach oben divergierend, wobei, während sich die Organprimordien aufwölben, der Blütenboden eine tellerförmige Gestalt annimmt (3–4). Durch weiteres entsprechendes Zellteilungswachstum, in der Außenflanke etwas intensiver, sowie durch differentielles Streckungswachstum (5–6) erhält er die Becherform. In den Schnitten ist das embryonale Gewebe punktiert. C = Karpell (Fruchtblatt), P = Petalum (Kronblatt), S = Sepalum (Kelchblatt, reduziert), Sa = Samenanlage, St = Stamen (Staubgefäß). Aus Leins & Erbar 1985.

wachsenden Fruchtblättern besteht (Abb. 6a–c), kommt ganz einfach dadurch zustande, dass unterhalb der Blütenhüll-, Staubgefäß- und Fruchtblattanlagen entsprechende Zellteilungen stattfinden, die zu immer länger werdenden Zellreihen (Abb. 7) und insgesamt dadurch zu einer becher- bis krugförmigen Bildung führen (Blütenbecher = becherförmiger Blütenboden = becherförmige Blütenachse = gynoeciales Hypanthium). Dieser einfache Prozess ändert die Proportionen an den Fruchtblättern und mag manchen Betrachter mangels räumlichen Vorstellungsvermögens dazu verführen, die Fruchtblattrüben mit einem krugförmigen Blütenboden verwachsen zu sehen. Eine solche fast esoterisch anmutende Sichtweise ist in manchen heutigen Lehrbüchern leider noch immer präsent. Aber schon Matthias Schleiden, der Begründer der Zellenlehre, formulierte 1846: „Es kann hier für den, der nur einige Entwicklungen in der Natur verfolgt hat, keinem Zweifel unterliegen, dass hier der ganze unterständige Fruchtknoten nur aus einer becherförmig gebildeten Axe entwickelt wird.“



Abbildung 8: Rasterelektronenmikroskop (REM) Leitz 1200B (Baujahr 1982; Standort: COS-Biodiversität und Pflanzensystematik). Im REM wird ein Elektronenstrahl zeilenförmig über das vergrößert abzubildende Objekt geführt („gerastert“; siehe den linken, grünlichen Bildschirm, auf dem die schwach aufleuchtende Zeile die aktuelle Bildinformation darstellt), und Wechselwirkungen der Elektronen mit dem Objekt werden zur Erzeugung eines Bildes genutzt. Die vom Objekt rückgestreuten Elektronen werden von einem Kollektorsystem erfasst und verstärkt. Das so erhaltene Signal wird zur Helligkeitsmodulation einer Bildröhre eingesetzt. Die erzeugten Bilder sind dreidimensionale Abbildungen der Objektoberfläche. Früher wurden über einen parallel geschalteten Bildschirm mit einer analogen Kamera Negativ-Filme belichtet. Heute werden direkt digitale Bilder erzeugt (Bildschirm ganz rechts, auf dem eine präparierte Blütenknospe zu erkennen ist).

„Die reinste Anschauung der Dinge hat, wer sie vom Anbeginn her wachsen sieht.“ Aristoteles (384 v. Chr.–322 v. Chr.)

Eine noch wesentlich mehr ins Detail gehende exomorphologische Betrachtung von jungen Entwicklungsstadien erlaubt uns seit Mitte der 60er Jahre des vorigen Jahrhunderts beziehungsweise der 70er Jahre (ab dann waren sie für Universitätsinstitute finanzierbar) das Rasterelektronenmikroskop (REM; Abb. 8). So gewähren etwa die beiden Stadien der Fruchtknotenentwicklung in Abb. 9 beim Liebstöckl (Doldenblütler, Apiaceae) Einblick in das Unterständigwerden des Fruchtknotens, indem auf der Innenfläche des Blütenbodens durch ein ringförmiges Wachstum desselben die Insertion der beiden Fruchtblätter steiler gestellt wird.

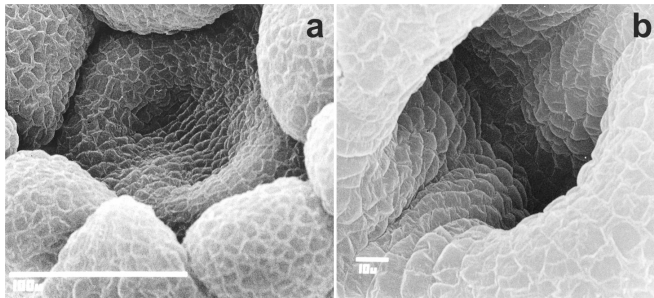


Abbildung 9: Junge Entwicklungsstadien des zweikarpelligen Fruchtknotens vom Liebstöckel *Levisticum officinale* (Apiaceae). – a, Zwei hufeisenförmige Karpell-(Fruchtblatt-)anlagen treffen sich in der Blütenmitte, wo sie das gemeinsame Septum (= Trennwand) bilden. – b, Die beiden Karpellanlagen sind durch das Blütenbodenwachstum bereits steiler gestellt. Aus Leins & Erbar 2008.

2 Ontogenien im Blütenbereich

Zum besseren Verständnis der folgenden Ausführungen benutzen wir ein Blütenmodell (Abb. 10), das nichts Anderes darstellt als ein didaktisches Hilfsmittel um notwendige Begriffe deutlich zu machen. Es ist weder eine „Urblüte“ noch sind ihre Teile „Blattmetamorphosen“, was wiederum aus dem Vokabular idealistischer Morphologie stammt und vielfach auch heute noch in Blütendefinitionen herumgeistert. Sehr oft können wir in einer Blüte eine Blütenhülle (häufig in Kelch und Blumenkrone differenziert), Pollen produzierende Staubgefäße (in

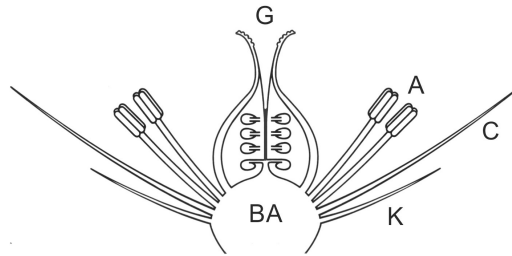


Abbildung 10: Modell einer Blüte der Bedecktsamer (Angiospermen). BA = Blütenachse, K = Kelch, C = Krone (Corolla), A = Androeceum (Staubgefäßbereich), G = Gynoeceum (Fruchtblattbereich). Aus Leins & Erbar 2018.

ihrer Gesamtheit als Androeceum bezeichnet) und Samenanlagen tragende Fruchtblätter (in ihrer Gesamtheit Gynoeceum genannt) erkennen; diese Organe sitzen „normalerweise“ an der sogenannten Blütenachse (= Blütenboden).

Wenn wir den exakten räumlich-zeitlichen Aufbau eines Organbereiches ermitteln wollen, müssen wir, Aristoteles folgend, wiederum verschiedene Entwicklungsstadien zu Rate ziehen. Ein geradezu klassisches Beispiel bieten uns die auf den ersten Blick alle im Prinzip gleichaussehenden sogenannten polymeren (aus vielen Staubgefäßen bestehenden) Androeceen. Es ist schon erstaunlich, wie unterschiedlich ihre Entwicklungsmuster sein können. In der oberen Reihe in der Zusammenstellung der Abb. 11 sind adulte Blüten folgender Pflanzen abgebildet: *Magnolia denudata* (Yulan-Magnolie, Magnoliaceae), *Hypericum hookerianum* (Hookers Johanniskraut, Hypericaceae), *Capparis spinosa* (Kapernstrauch, Capparaceae), *Melaleuca hypericifolia* (Johanniskrautblättrige Myrtenheide, Myrtaceae).

Den Entwicklungsstadien im rasterelektronenmikroskopischen (REM-) Bild zufolge entstehen die jungen Anlagen der späteren Staubgefäße entweder direkt am Blütenvegetationskegel von unten nach oben in spiralförmiger Sequenz (vgl. auch Erbar & Leins 2019), wie bei der Magnolie, oder sie bilden sich auf Primärprimordien (Primärhöckern), wie beim Johanniskraut oder der Myrtenheide, oder aber die Primäranlagen bilden einen gemeinsamen Ringwall, auf dem sich die Einzelanlagen ausgliedern, wie bei der Kapernblüte.

Die Ausgliederungsfolge auf den Primäranlagen vollzieht sich in Bezug auf den Blütenmittelpunkt in zentrifugaler Richtung (von innen nach außen), nämlich bei

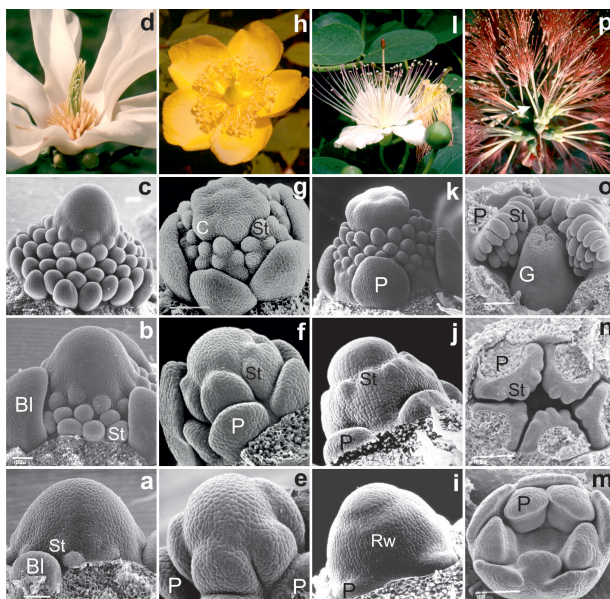


Abbildung 11: Entwicklungsmuster in polymeren Androeceen. Die obere Reihe zeigt die adulten Blüten, in denen das Anlagemuster der zahlreichen Staubgefäße sich nicht mehr unmittelbar erschließt. In den Reihen darunter sind jeweils drei charakteristische Stadien im rasterelektronenmikroskopischen Bild dargestellt. – a–d, Yulan-Magnolie (*Magnolia denudata*, Magnoliaceae). Spirale Ausgliederung zahlreicher einzelner Staubgefäße direkt am konischen Blütenvegetationskegel. – e–h, Hookers Johanniskraut (*Hypericum hookerianum*, Hypericaceae). – e, Vor den fünf Kronblättern werden Primäranlagen gebildet. – f, Die Primäranlagen beginnen sich von der Spitze her in Staubgefäße aufzugliedern. – g, In Bezug auf den Blütenmittelpunkt erfolgt die Aufgliederung der Primäranlagen von innen nach außen (zentrifugal). – h, Die Bündelung macht sich auch in adulten Blüten noch schwach bemerkbar. – i–l, Kapernstrauch (*Capparis spinosa*, Capparaceae). – i, Sichtbarwerden eines Ringwalls (Rw) im späteren Staubgefäßbereich. – j, Entstehung der obersten (innersten) Staubgefäßanlagen. – k, Zur Basis des Ringwalls fortschreitende weitere (zentrifugale) Ausgliederung von Staubgefäßen. – m–p, Johanniskrautblättrige Myrtenheide (*Melaleuca hypericifolia*, Myrtaceae). – m, Die vor den Kronblattanlagen befindlichen Primäranlagen sind noch ungegliedert. – n, Beginn der Aufgliederung der Primäranlagen in einzelne Staubgefäßanlagen (Kronblattanlagen abpräpariert). – o, Die Aufgliederung erfolgt an der Primäranlage in Bezug auf den Blütenmittelpunkt von außen nach innen (zentripetal); Blütenknospe längs aufgeschnitten (Kronblattanlagen abpräpariert). – d, In adulten Blüten haben sich die Primäranlagen stielartig verlängert (blütenökologischer Pinselblumen-Typ; Pfeil weist auf einzelnes Staubgefäßbündel). Bl = Blütenhüllblatt, C = Karpell (Fruchtblatt), G = Gynoceum (Fruchtknoten), P = Petalum (Kronblatt), St = Staubgefäß. Zusammengestellt aus Leins & Erbar 2008.

Hypericum und *Capparis*¹, oder umgekehrt, wie bei *Melaleuca*, in zentripetaler Richtung.

Ohne solche gründlichen Analysen der Entwicklungsgeschichte polymerer Androeceen wären phylogenetische Folgerungen, wie wir sie noch anstellen wollen (siehe Kapitel 3), nicht möglich. Ziehen wir einen weiteren Organbereich in Betracht, etwa die Blumenkrone (Corolla) einer zweikreisigen (in Kelch und Krone differenzierten) Blütenhülle, entdecken wir, dass in vielen Fällen die Kronblätter einer mehr oder weniger langen Röhre aufsitzen; man spricht auch von einer Kronröhre (oder sympetalen Krone; Petalum = Kronblatt). Diese wiederum ist sehr oft in zwei übereinanderliegende Teile differenziert, die eigentliche Kronröhre, der sogenannte Corollentubus, und darunter die „Staubgefäßkronröhre“ (sogenannter Stamen-Corollentubus; Abb. 12). Ein Stamen-Corollentubus ist eine späte, meist kurz vor dem Aufblühen sich entwickelnde Bildung, die oft im Zusammenhang mit einer wichtigen Blütenfunktion steht (siehe weiter unten unter sekundärer Pollenpräsentation, Kapitel 6). Sie kommt dadurch zustande, dass unter dem

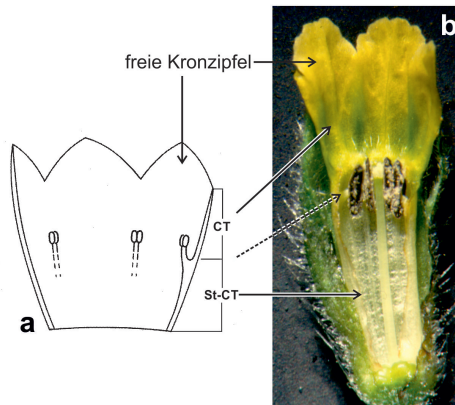


Abbildung 12: Kronröhre im weiteren Sinn aus Corollentubus (CT, Kronröhre im engeren Sinn) und Stamen-Corollentubus (St-CT) bestehend (Röhre aufgeschnitten). – a, Schema. – b, Blüte des Gelben Mönchskrauts (*Nonea lutea*, Boraginaceae) längs aufgeschnitten.

¹ Übrigens: Was wir bei den Kapern benutzen, sind die in ihrer Entwicklung weit fortgeschrittenen (in Salz eingelegten) Blütenknospen, die von den Kelchblättern geschützt beim Aufpräparieren im Innern alle Blütenorgane schon deutlich erkennen lassen.

zuvor gebildeten Corollentubus ein eingeschobenes (interkalares) Zellteilungs- und Zellstreckungswachstum in einer ringförmigen Zone stattfindet, das auch die Ansätze der Staubgefäße erfasst und diese auf der Innenseite des Tubus mit hochzieht und somit jeweils eine mehr oder weniger prominente Leiste erkennen lässt; oberhalb dieser leistenartigen Ansätze befindet sich der Corollentubus, der für unsere Betrachtung von besonderem Interesse ist, weil seine Genese recht unterschiedlich verlaufen und – wie wir noch sehen werden – bedeutende Hinweise in der Verwandtschaftsforschung (Systematik) liefern kann.

Kronröhren (im weiteren Sinn aus Corollen- und Stamen-Corollentubus bestehend) dienen vielfach dem Schutz der Staubgefäße und ermöglichen das Verbergen des Nektars an deren Grund. So gesehen sind Kronröhren häufig das Resultat einer gekoppelten Evolution (Coevolution) mit der Ausgestaltung der Insekten-Mundwerkzeuge. Coevolution führt also zur Coadaptation, einem gemeinsamen Angepasstsein, das in unüberschaubarer Variation, was die Form der Kronröhren und mancher Anhangsgebilde betrifft, mündet. Ein kleiner Ausschnitt aus dem Kaleidoskop sich ergebender ästhetischer Blumengestalten präsentiert der Blumenkorb der Abb. 13 (Schönheit durch Zufall und Zweck!; Erbar & Leins 2019).



Abbildung 13: Sympetale Blüten aus dem oberen Stammbaumniveau (Euasteriden).

Immer noch wird in der Literatur, was die Sympetalie betrifft, der Begriff „Verwachsung“ benutzt (was praktisch, aber nicht richtig ist), und alle Kronröhren werden als ähnlich betrachtet, nämlich als typisches Beispiel für eine sogenannte „kongenitale Verwachsung“. Ontogenetische Studien aber haben gezeigt, dass sympetale Kronen, wie oben bereits angedeutet, sich auf ganz unterschiedliche Weise entwickeln. Ontogenetisch kann eine Kronröhre extrem früh gebildet werden, nämlich bevor die Kronzipfel erscheinen. „Frühsympetal“ nennen wir solche Blumenkronen, bei denen sich zunächst ein Ringwall bildet, auf dem sich erst später die Kronblattzipfel (vermehrte Zellteilungen an fünf oder gelegentlich vier Stellen) ausgliedern (Abb. 14, 17a). Aus dem Ringwall geht durch kontinuierliches Wachstum ein Corollentubus hervor. Die Staubgefäßanlagen entstehen alternierend zu den Kronzipfeln, abgerückt von deren Verbindungsschultern (Abb. 14).

Eine Kronröhrenbildung kann aber auch relativ spät einsetzen: Die Kronblätter werden zunächst frei angelegt und bleiben mehr oder weniger lange Zeit voneinander getrennt (Abb. 15–16, 17b–c). Erst eine geraume Zeit nach der Anlage der Staubgefäße dehnen sich die Basen der Kronblattanlagen seitlich aus, bis sie sich hinter dem dorsalen Staubgefäßansatz treffen, das heißt, ihre Randmeristeme

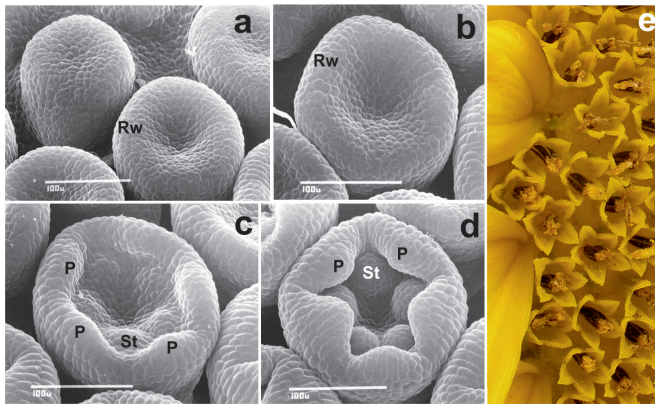


Abbildung 14: Kronentwicklung bei der Ringelblume *Calendula officinalis* (Asteraceae). – a–b, Vorwölbung der Kronröhrenanlage als Ringwall (Rw). – c–d, Auf den Kronröhrenringwall werden die Primordien (Anlagen) der Kronzipfel „aufgesetzt“. Zur Zeit der Staubgefäßentstehung ist die Kronröhre schon deutlich ausgeprägt („frühe Sympetalie“). – P = Petalum (Kronblatt = Kronzipfel), St = Stamen (Staubgefäß). Aus Erbar 1991. – e, Ausschnitt aus dem Köpfchen der Sonnenblume (*Helianthus annuus*).

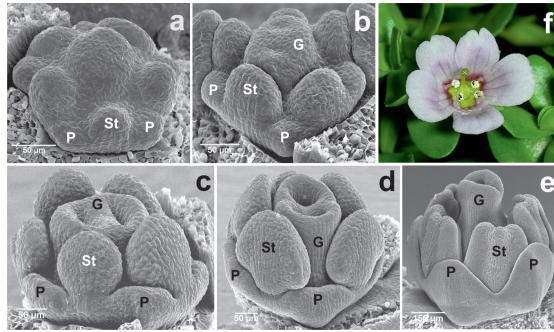


Abbildung 15: Kronentwicklung beim Kleinen Fettblatt (*Bacopa monnieri*, Plantaginaceae). – a, Die Kronblätter sind als völlig separate Primordien angelegt worden. Alternierend zu den Kronblättern sind auch schon die jungen Staubgefäße angegliedert, die an Größe die Kronblattanlagen „überholen“. – b, Etwas älteres Stadium, in dem die Kronblattanlagen immer noch frei sind. – c, Die Kronblattanlagen dehnen sich langsam hinter den Staubgefäßen an ihrer Insertion aus. – d, Basal treffen sich die Ränder der Kronblätter. – e, Jetzt erst können durch die Meristemfusion die Kronblätter gemeinsam hochwachsen und eine Röhre bilden („späte Sympetalie“). Man beachte die Differenzierung der Staubgefäße in Staubfaden und Staubbeutel (c–e). REM-Bilder: Staatsexamensarbeit Julia Cronauer (2007). – P = Petalum (Kronblatt), St = Stamen (Staubgefäß). – f, Adulte Blüte.

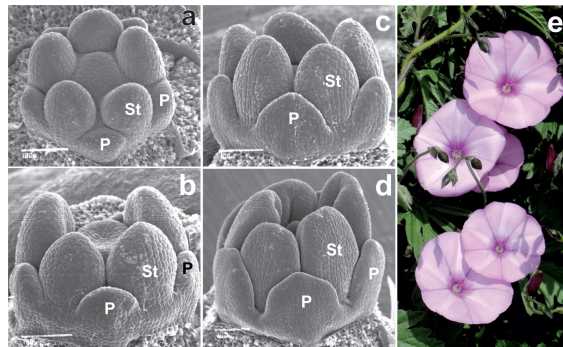


Abbildung 16: Kronentwicklung bei der Eibischblättrigen Winde (*Convolvulus althaeoides*, Convolvulaceae). – a, Die in der Entwicklung gegenüber den Staubgefäßeanlagen im Wachstum verlangsamt Kronblattanlagen sind noch völlig frei. – b–c, Meristembrückenbildung zwischen den Kronblattanlagen („späte Sympetalie“). – d, Die zunächst eben verlaufende interprimordiale Verbindung wird mehr und mehr durch die sich basal verbreiternden Kronzipfel zu einer U-förmigen Bucht verändert. – e, Adulte Blüten. – P = Petalum (Kronblatt), St = Stamen (Staubgefäß). – e, Fertige Blüten.

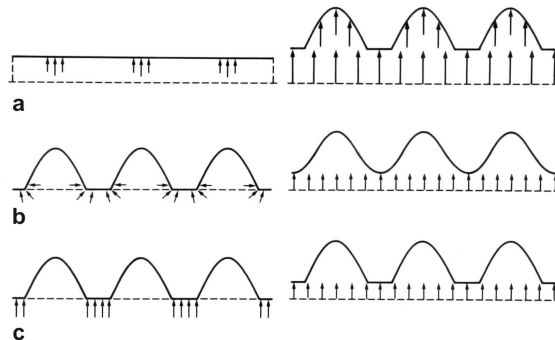


Abbildung 17: Entwicklungsmöglichkeiten von Kronröhren (im engeren Sinne) bei den Tetra-cyclischen Sympetalen. Frühe Sympetalie (a): Die Kronzipfel entstehen auf einem einheitlichen Ringwall. – Späte Sympetalie (b + c). – b, Zwischen frei angelegten Petalumprimordien entstehen interprimordiale Verbindungen durch langsame Ausdehnung der seitlichen Insertionsbereiche der Petalen bis zur Fusion der Meristeme. – c, Zwischen frei angelegten Petalumprimordien entstehen interprimordiale Verbindungen durch „Brücken“ (schlagartig entstehende Leisten zwischen den Petalumprimordien). – Verändert aus Erbar & Leins 2011.

sind jetzt miteinander verbunden (Abb. 15). Diese Verbindungsstellen wachsen gemeinsam mit den Kronzipfeln hoch. Wir haben es hier also mit einer „späten Sympetalie“ zu tun. Das Initialstadium einer spät gebildeten Kronröhre kann aber auch dadurch zustande kommen, dass mehr oder weniger schlagartig zwischen separat angelegten Kronblattanlagen Meristembrücken entstehen (Abb. 16).

Einige wenige Fälle können wir als „Übergang zwischen früher und später Sympetalie“ bezeichnen, nämlich dann, wenn die Kronblattanlagen am Rand eines Plateaus entstehen, oder die Ausdehnung und Verbindung der Kronblattbasen zeitlich mit der Anlage der Staubgefäße zusammenfällt.

3 Phylogenie

Stammesgeschichtlich, beziehungsweise merkmalsphylogenetisch, gesehen, repräsentieren unter den oben gezeigten polymeren Androecealtypen die Magnolien

ein mehr ursprüngliches Muster², während die Aufgliederung einzelner Anlagen oder eines Ringwalles als sekundär, also als abgeleitet, interpretiert werden kann. Eine während der Evolution aus einfachen Androceen mit zwei- oder einkreisig angeordneten Staubgefäßen (oft 10 oder 5 in einer Blüte) sich durch eine Aufgliederung (Fiederung) ihrer jungen Anlagen erreichte, oft starke Vervielfachung mag mit dem hohen Bedarf an Pollenkörnern seitens der aus grabwespenähnlichen Hautflüglern entstandenen Bienen als Anpassung im Zusammenhang stehen. Eine in dieser Richtung vollzogene Interpretation steht im Einklang mit einem modernen, auf der Grundlage von molekularen Daten propagierten Stammbaum der Blütenpflanzen (Abb. 18), in dem die Entwicklungsmuster polymerer Androceen eingetragen sind. Sekundär (zentrifugal oder zentripetal) vermehrte Staubgefäße kommen im mittleren Evolutionsniveau vor, wo auch die zwei- und einkreisigen Androceen zu Hause sind. Aus dem Stammbaum geht außerdem hervor, dass „gebündelte“ polymere Androceen, ob zentrifugal oder zentripetal oder auch nur am Rand von Primärhöckern „aufgespalten“, im Verlauf der Evolution mehrfach entstanden sind (Parallelismus).

Greifen wir bei unseren phylogenetischen Überlegungen im Folgenden nochmals die unterschiedlichen Entwicklungsmuster von Kronröhren auf. Dabei bewegen wir uns ins obere Stammbaumniveau der Blütenpflanzen, nämlich zu den derzeit Euasteriden genannte, recht artenreiche Blütenpflanzen-Gruppe, die in der Vergangenheit als „Sympetalae Tetracyclae“ bezeichnet wurde; eine treffende Bezeichnung, da fast alle Vertreter durch eine Kronröhre gekennzeichnet sind (d. h. sie sind sympetal), und die Blütenorgane sind auf vier Kreise („Zyklen“) verteilt: jeweils ein Kreis von Kelch- und Kronblättern, ein Staubgefäßkreis (alternierend mit den Kronblättern) und ein Fruchtblattkreis (Fruchtknoten meist aus zwei Fruchtblättern). Projiziert man die Entwicklungsmuster auf die Familien dieser Großgruppe (Abb. 19–20), so ergibt sich, dass basierend auf breiten ontogenetischen Studien und im Vergleich mit anderen Merkmalen (u. a. Daten aus

² Die Magnolien (Magnoliaceae) gehören zwar nach heutigem Verständnis nicht zu den basalen Gruppen der Blütenpflanzen (Angiospermen), der sogenannten ANITA-Gruppe (Amborellaceae – Nymphaeales – Illiciaceae – Trimeniaceae – Austrobaileyaceae) oder ANA-Gruppe (Amborellales – Nymphaeales – Austrobaileyales), sondern zum nächsthöheren Clade (Stammbaumast), teilen aber eine Reihe von Merkmalen mit den basalen Gruppen: neben sekundären Pflanzeninhaltsstoffen eben auch die spiralförmige Anlage der Blütenorgane (siehe für den Sternanis *Illicium* Abb. 14 in Erbar & Leins 2019).

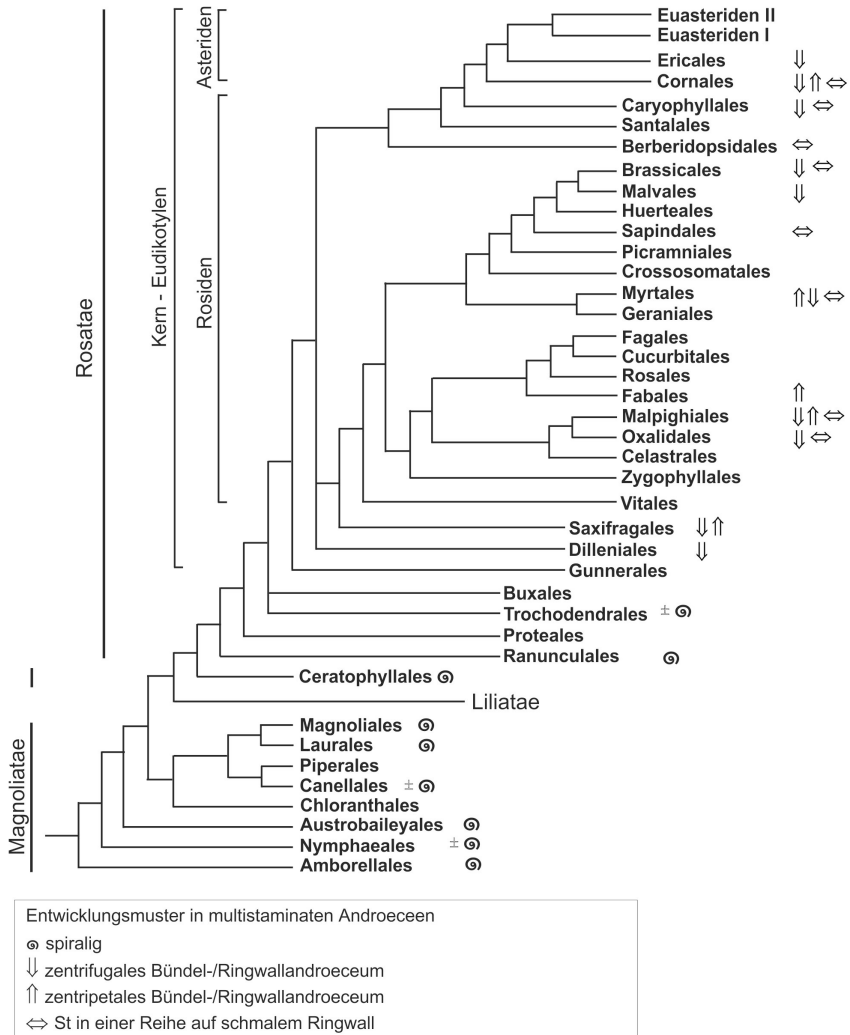


Abbildung 18: Vorkommen von multistaminaten spiralgigen und Bündel-/Ringwallandroeceen in den Ordnungen der Blütenpflanzen; berücksichtigt sind auch (nicht-wirtelige) Androeceen mit moderater Staubgefäßzahl; zur Umschreibung der Ordnungen (mit Angabe der Entwicklungsmuster in den einzelnen Familien) siehe systematischen Anhang in Leins & Erbar 2010.

DNA-Sequenzanalysen, auf denen der Stammbaum erstellt wurde) die entwicklungsgeschichtlichen Merkmale „frühe“ und „späte Sympetalie“ als gute Marker für systematische Überlegungen zur Verfügung stehen.

Die „Tetracyclischen Sympetalen“ oder Kern-Asteriden (Euasteriden) werden heute in zwei Untergruppen gegliedert, die Euasteriden I (= Lamiiden) und die Euasteriden II (= Campanuliden). Bei den Euasteriden I, zu denen unter vielen anderen die Raublattgewächse, Enziangewächse, Nachtschattengewächse, Rachenblütler, Wegerichgewächse und Lippenblütler gehören, überwiegt späte Sympetalie (Abb. 19). Weitere Merkmale sind die meist oberständigen Fruchtknoten, die sich oft zu Kapsel Früchten entwickeln. Das Nektar absondernde Gewebe, das Nektarium, sitzt an der Fruchtknotenaußenseite oder an der Blütenachse. Bei den Euasteriden II, zu denen die Kardengewächse, Glockenblumengewächse und die artenreichste Familie der Köpfchenblütler gehören, findet sich, soweit untersucht, durchgehend die frühe Sympetalie (Abb. 20). Weiterhin ist diese Gruppe charakterisiert durch oft unterständige Fruchtknoten, die sich in der Regel zu Schließfrüchten entwickeln. Das Nektarium sitzt meist am Dach des unterständigen Fruchtknotens.

An zwei Beispielen soll der taxonomische Wert des Merkmals „Zustandekommen der Sympetalie“ demonstriert werden. Eines der überraschendsten Ergebnisse, das mit den molekularen Daten erhalten wurde, ist die neue Stellung der Fieberkleegegewächse (Menyanthaceae) in der Nähe der Köpfchenblütler (Asteraceae), eine Verbindung, die vorher niemals in Betracht gezogen wurde. Die Familie wurde bevorzugt in die Nähe der Enziangewächse (Gentianaceae) gestellt, wobei die Anordnung der Samenanlagen an der Innenwand des Fruchtknotens (parietale Plazentation) und das Vorkommen von Bitterstoffen (Iridoide) als wesentliche Merkmale angesehen wurden. Das ontogenetische Merkmal „frühe Sympetalie“ (Abb. 21) aber unterstützt eine Position der Fieberkleegegewächse nahe den Köpfchenblütlern (Übereinstimmungen finden sich außerdem im Vorkommen von Inulin als weiteren Inhaltsstoff sowie in der Embryoentwicklung im Samen).

Eine Verbindung der Doldenblütler-Verwandtschaft mit den Köpfchenblütlern wurde schon in der „vor-molekularen“ Zeit auf Grund vieler übereinstimmender Merkmale aus den Bereichen sekundäre Inhaltsstoffe, Pollenkunde, Embryoentwicklung und auch der Blütenmorphologie hergestellt (siehe Erbar & Leins 1988). Mit der Doldenblütler-Verwandtschaft sind die Familien der Ordnung Araliales (Efeugewächse im weiteren Sinne) gemeint, in die die Doldenblütler im engeren Sinne als Familie Apiaceae platziert werden. Die einzige Crux einer verwandt-

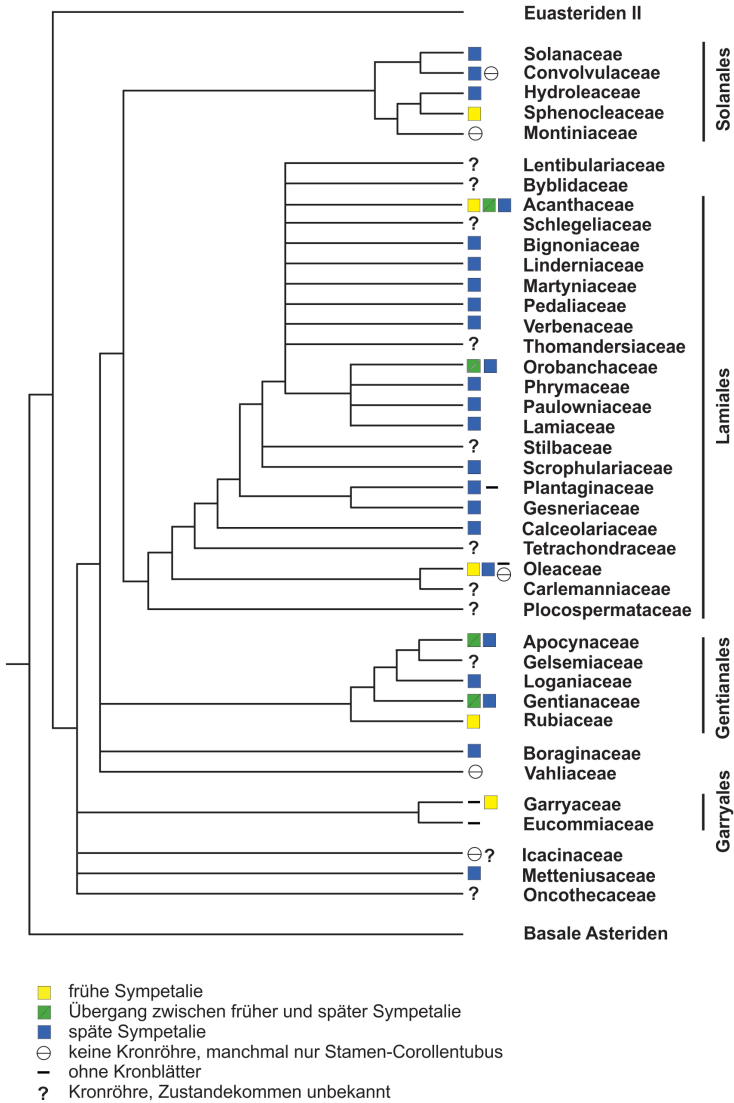


Abbildung 19: Ein vereinfachter Stammbaum der Euasteriden I, in den das entwicklungsgeschichtliche Merkmal des Zustandekommens der Kronröhre (i. e. S.) eingetragen ist. Verändert aus Erbar & Leins 2011.

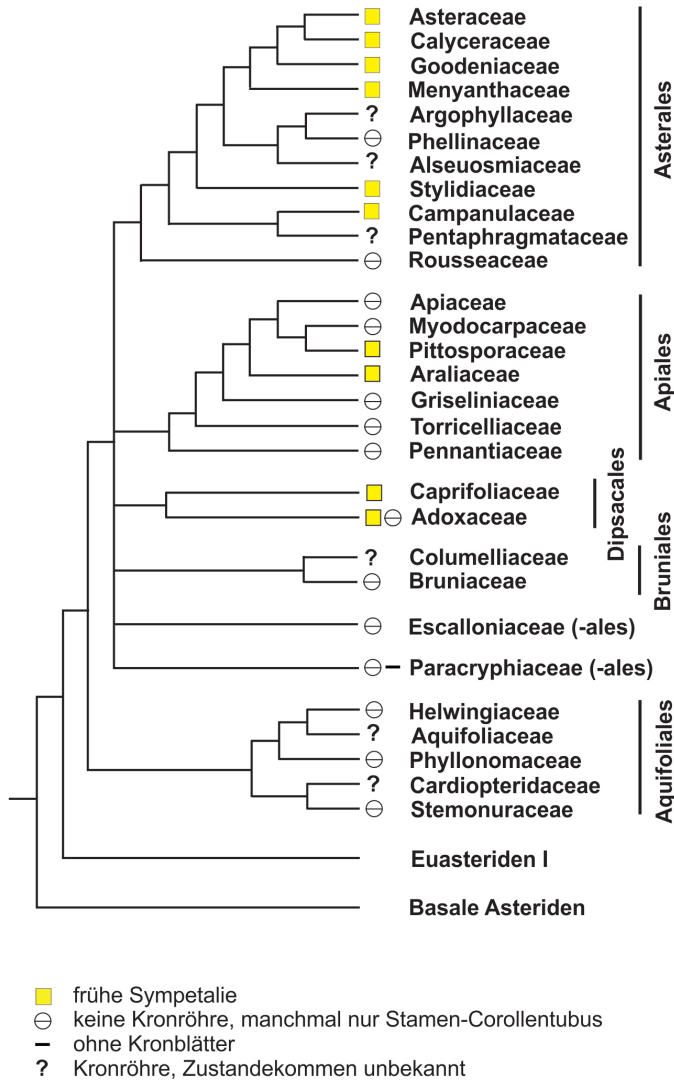


Abbildung 20: Ein vereinfachter Stammbaum der Euasteriden II, in den das entwicklungsgeschichtliche Merkmal des Zustandekommens der Kronröhre (i. e. S.) eingetragen ist. Verändert aus Erbar & Leins 2011.

schaftlichen Zusammenführung ist das Fehlen einer Kronröhre in ihren tetrazyklischen Blüten. Eine Brücke kann jedoch hergestellt werden, wenn wir die frühe Ontogenie der Blumenkrone der verwandten Efeugewächse (Araliaceae) zu Rate ziehen (Erbar & Leins 1988, 2004). Beim Efeu (*Hedera helix*, Abb. 22) entsteht nach der Kelchanlage nämlich ein niedriger Ringwall, auf dem fast gleichzeitig mit diesem die Kronblattprimordien hervortreten (Abb. 22a–b). Das Ringprimordium, das die Kronblattprimordien verbindet (frühe Sympetalie!), wächst jedoch nicht zu einer Röhre aus (Abb. 22c).

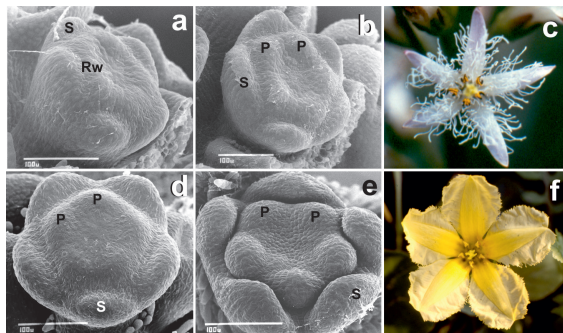


Abbildung 21: „Frühe Sympetalie“ bei Vertretern der Fieberkleegeewächse (Menyanthaceae): Fieberklee (*Menyanthes trifoliata*; a–c) und Europäische Seekanne (*Nymphoides peltata*; d–f). Auf einem Ringwall entstehen bald nach dessen Bildung die Anlagen für die Kronblätter. P = Petalum (Kronblatt), Rw = Kronringwall, S = Sepalum (Kelchblatt). Verändert aus Erbar 1977.

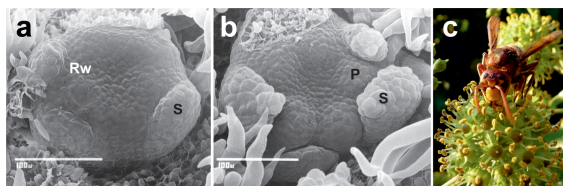


Abbildung 22: Frühe Kronentwicklung beim Efeu (*Hedera helix*, Araliaceae). Obwohl die Blüten später keine Kronröhre aufweisen (c; Ausschnitt aus dem Blütenstand mit einer Hornisse als Bestäuber), beginnt die Kronentwicklung mit einem niedrigen Ringwall (a), auf dem sich die Kronblätter entwickeln (b). Der frühe Kronringwall wächst nicht zu einer Kronröhre aus. P = Petalum (Kronblatt), Rw = Ringwall. Verändert aus Erbar & Leins 1988.

4 Hologenie

Mögen auch, wie schon erwähnt, Stammbäume Theorie behaftet, an manchen Stellen sogar stark hypothetisch sein, so bleibt immer wieder festzuhalten, klammert man die Rekonstruktionen aus und beschränkt den Evolutionsbegriff auf die nachvollziehbaren Faktoren (Abb. 23), die die Evolution antreiben, dass man sie mit Fug und Recht auf die Stufe einer wissenschaftlichen Erkenntnis stellen kann. Tatsache ist: Evolution findet statt!

Mutationen (spontane Änderungen an der Basensequenz der Doppelhelix – DNS – oder allgemein am Genom) können nur dann in der Nachkommenschaft wirksam werden, wenn diese in der Keimbahn stattfinden und schließlich nach der Meiose in die Geschlechtszellen oder aber – bei Pflanzen – in die haploiden Meiosporen (aus denen sich durch Mitosen Gametophyten entwickeln) münden. Eine Veränderung an den Erbanlagen kann in vielen Fällen zu einer Formveränderung führen, die sich in einem bestimmten Entwicklungsstadium des Organismus durch einen abgeänderten Ablauf der weiteren Entwicklung bemerkbar macht. Die Einbeziehung der unzähligen Ontogenien in die Phylogenie aller Lebewesen ist die ganzheitliche Sicht der Entstehung der gesamten Diversität der (sowohl ausgestorbenen als auch lebenden) Sippen. Diese bereits von einigen Pionieren der Evolutionsforschung angedachte holistische Sichtweise hat später Walter Zimmermann (1934, 1961, 1966, 1968), einer der wohl bedeutendsten Phylogenetiker des

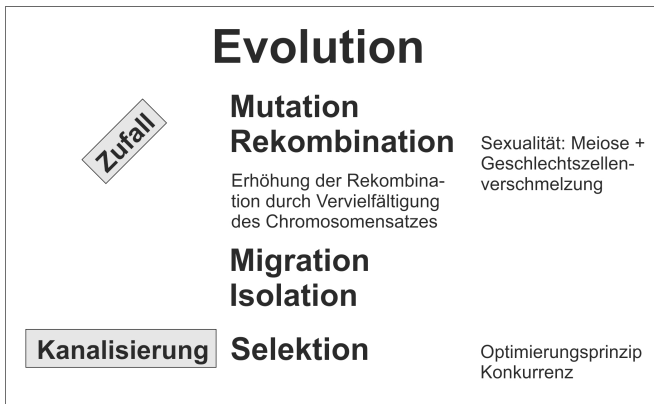


Abbildung 23: Die Faktoren der Evolution. Aus Leins & Erbar 2018.

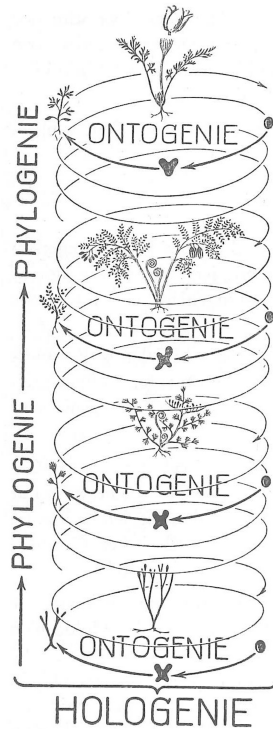


Abbildung 24: Hologenie von den ersten Landpflanzen (Psilophytatae) über verschiedene Farngruppen zu den Blütenpflanzen (Angiospermen). Aus Zimmermann 1961, 1968.

20. Jahrhunderts, als Hologenie definiert. In seinem Schema (Abb. 24) bildet er in einer Kette von Ontogenien (letztere als Spiralen verdeutlicht) über weit mehr als 400 Millionen Jahre hinweg die Entstehung der Blütenpflanzen beginnend mit den Urfarnen (Telomstände) des Silurs über die Wedelfarne und Samenfarne jeweils mit vier Ontogenie-Stadien ab. Aus dem äußerst kompakten Schema geht in einfacher Weise hervor, um was es geht. Würde man sämtliche Ontogenien aller Arten (und Individuen) in einem solch langen Zeitraum berücksichtigen, wäre eine Hologenie wegen der nicht mehr zählbaren Ontogenien und deren Abänderungen natürlich niemals darstellbar. Wir können unsere hologenetischen Konstruktionen nur in kleinen Geltungsbereichen durchführen und diese als Puzzleteile für größere Zusammenhänge benutzen. Folgt man Zimmermanns Konzept der kettenförmigen

Reihung von Ontogenien (Abb. 24), so ergäbe sich eine Hologenie von Androeceen und Kronröhren bei den Blüten, wie sie nach unserer Konstruktion in Abb. 25 wiedergegeben ist.

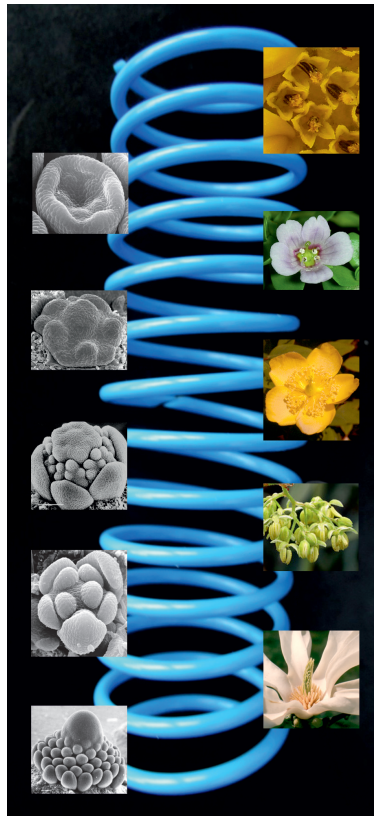


Abbildung 25: Ontogenien multistaminater Androeceen, ausgehend von einem spiralgigen über ein einfaches zum zentrifugalen komplexen Androeceum sowie der späten Sympetalie und frühen Sympetalie im oberen Stammbaum-Niveau (hier finden wir hauptsächlich einfache – einkreisige – Androeceen) in Zimmermanns Hologenie-Schema übertragen. Die Ontogenien sind jeweils nur durch ein Knospen- und ein Adultstadium repräsentiert. Beispielhaft (!) sind von unten nach oben die Verhältnisse bei der Magnolie (*Magnolia denudata*; vgl. Abb. 11a–d), dem Hopfen (*Humulus lupulus*; vgl. Abb. 26), dem Johanniskraut (*Hypericum hookerianum*; vgl. Abb. 11 e–h), beim Fettblatt (*Bacopa monnieri*; vgl. Abb. 15) und der Ringelblume (*Calendula officinalis*; vgl. Abb. 14).

Aufschlussreicher ist jedoch eine Darstellung, die einem Stammbaum mit V-förmig sich verzweigenden Ästen entspricht. Allerdings werden in der Darstellung aus Platzgründen die Zweige (als oben offenes Rechteck) um 90° geknickt (es ergibt sich spöttisch ein „Gabelsystem“). Die divergierende Weiterentwicklung eines bestimmten Entwicklungszustandes ist für das einfache und gebündelte Androeceum in Abb. 26 gezeigt. Man stelle sich vor, welche noch unüberschaubarere Gabelnetze sich dabei ergeben (eine unerreichbare Synthese!).

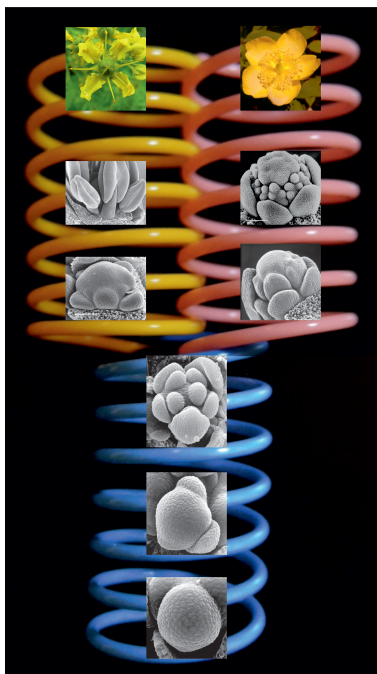


Abbildung 26: Divergenz einfacher (ein- bis zweikreisiger) und multistaminater (z. B. zentrifugaler) komplexer Androeceen auf eine einer Stammbaumverzweigung entsprechende Hologenie-Darstellung projiziert. Ontogenien stärker gedehnt als in Abb. 25. Blaue Spirale beinhaltet die frühe Ontogenie der Blütenhüll- und Staubgefäßenanlage (5 Stamina vor 5 spiralig gebildeten Blütenhüllblättern) in den männlichen Blüten beim Hopfen (*Humulus lupulus*); rosa Spirale führt zur Aufgliederung der Androecealanlagen vor den Kronblättern (gebündeltes Androeceum; *Hypericum hookerianum*, vgl. Abb. 11 e–h); gelbe Spirale zeigt die Entwicklung eines zweikreisigen Androeceums am Beispiel der Weinraute (*Ruta graveolens*, vgl. Abb. 30).

5 Der genetische Ansatz der Blütenbildung

Bevor Blüten gebildet werden können, müssen die entsprechenden Gewebe vom vegetativen Wachstum zur Blütenbildung umgesteuert werden. Auslöser können innere (endogene) oder äußere (exogene) Faktoren sein. Die wichtigsten Faktoren sind dabei die Tageslänge (Kurztag- und Langtagpflanzen) und Kälteperioden (Vernalisation). Zu den inneren Faktoren gehören die Aktivitäten bestimmter Gene und Pflanzenhormone. Nachdem ein Meristem (Bildungsgewebe) vom vegetativen (Blätter bildendes) in ein generatives (Blütenorgane bildendes) Meristem umgewandelt wurde, erfolgt durch eine Vielzahl von Regulator-Genen, die miteinander agieren, die eigentliche Blütenbildung.

Eine vollständige Blüte der Bedecktsamer (Angiospermen, d. h. die Samenanlagen sind von einem Fruchtblatt beziehungsweise Fruchtknoten eingeschlossen) mit all ihren Organen haben wir ja schon in der Abb. 10 kennengelernt. Die ersten erkennbaren Entwicklungsprozesse der Blüten bestehen in der Aufwölbung eines Blütenvegetationskegels (der späteren Blütenachse), an dem sich in der Regel von unten nach oben bzw. außen nach innen (akropetal bzw. zentripetal) die Primordien (Anlagen) der Blütenorgane ausgliedern (Abb. 27a). Unabhängig von der späteren Form und Funktion der einzelnen Blütenorgane in der adulten Blüte beginnen alle Blütenorgane ihre Entwicklung in Form kleiner, oft halbkugelförmiger Vorwölbungen, indem sich embryonale Zellen periklin teilen.

Die Frage war aber, wie diese undifferenzierten Zellen und ihre Abkömmlinge ihre Position innerhalb der sich entwickelnden Blüte finden und sich in die ent-

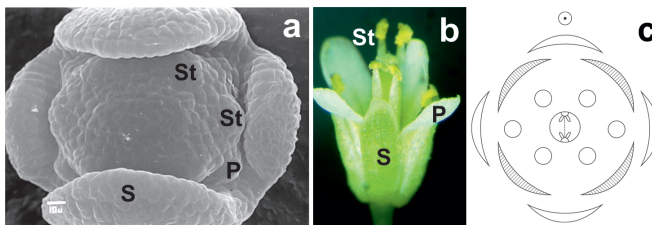


Abbildung 27: Blütenverhältnisse bei den Kreuzblütlern (Brassicaceae). – a, Junge Blütenknospe der Pfeilkresse (*Lepidium draba*) im rasterelektronenmikroskopischen Bild. Vier Kelchblätter, vier alternierende Kronblätter und sechs Staubgefäße in zwei Kreisen (2 St im äußeren, 4 St im inneren Kreis) sind ausgegliedert. – b, Blüte der Acker-Schmalwand in der Seitenansicht. – c, Blütendiagramm. – P = Petalum (Kronblatt), S = Sepalum (Kelchblatt), St = Stamen (Staubgefäß).

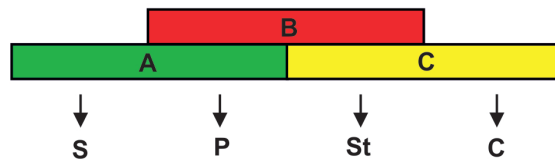


Abbildung 28: Das klassische A-B-C-Genklassenmodell in Anlehnung an Coen & Meyerowitz (1991). A, B, C: Wirkungsorte der Gene, C = Karpell (Fruchtblatt), P = Petalum (Kronblatt), S = Sepalum (Kelchblatt), St = Stamen (Staubgefäß).

sprechenden Zelltypen differenzieren. Genetische Untersuchungen an der Blüte der Acker-Schmalwand *Arabidopsis thaliana* haben zu dem sogenannten A-B-C-Modell der Organdetermination geführt. Die Blüten von *Arabidopsis* (Abb. 27b) besitzen nach Kreuzblütlermanier vier Kelchblätter, vier Kronblätter, zwei äußere kürzere und vier innere längere Staubgefäße sowie einen für die Familie typischen aus zwei Karpellen aufgebauten Fruchtknoten (Diagramm Abb. 27c).

Für die Determination dieser vier Organkategorien sind sogenannte homöotische³ Gene verantwortlich, die sich auf drei Genklassen⁴ verteilen, die als A, B und C bezeichnet werden. Die Organidentität wird nun entweder durch die Funktion von nur einer Genklasse oder durch Expressionsüberlagerung zweier Klassen determiniert; letzteres trifft für die Blumenkrone und die Staubgefäße zu: Genklasse A ist für die Kelchblätter bestimmend, die Genklassen A und B zusammen für die Kronblätter, die Genklassen B und C für die Staubgefäße und die Genklasse C allein für die Fruchtblätter (Abb. 28).

Da sich herausstellte, dass die Samenanlagen als fünfte Organkategorie aufzufassen sind, die durch die Funktion einer zusätzlichen Genklasse D determiniert wird, musste das Modell auf ein A-B-C-D-Modell erweitert werden. Die ABCD-Gene alleine sind aber nicht ausreichend für die Bestimmung der Organidentität. Daher wurde das 1991 aufgestellte klassische A-B-C-Modell im Jahr 2000 zum A-B-C-D-E-Modell erweitert, wobei die E-Funktion in allen Organkategorien für die Bestimmung der Identität erforderlich ist (Abb. 29a; für ausführliche Literaturangaben siehe Erbar 2007, 2010). Die Gene A, B, C, D und E codieren

³ Der Begriff leitet sich von „Homöosis“ ab. Darunter wird das Auftreten einer Organkategorie an einer falschen Stelle verstanden (z. B. die Bildung von Fruchtblättern anstelle von Staubgefäßen).

⁴ Eine Genklasse ist eine Gruppe von Genen mit dergleichen codierenden Funktion.

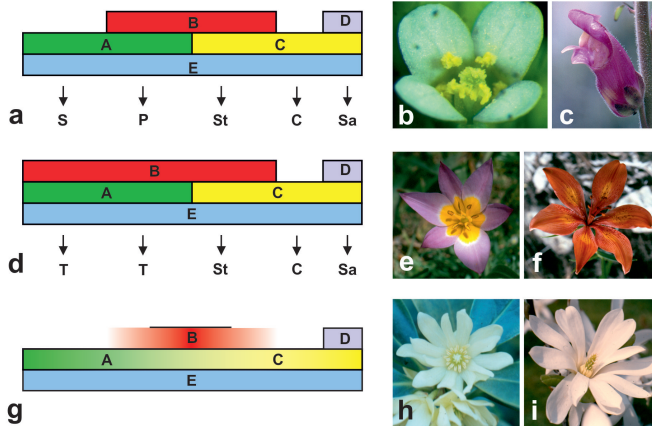


Abbildung 29: Das genetische Modell zur Blütenorgandeterminierung. – a, A-B-C-D-E-Modell der Blütenorgan-Spezifizierung bei den Kern-Eudikotylen, z. B. Acker-Schmalwand (*Arabidopsis thaliana*; b) und Löwenmäulchen (*Antirrhinum majus*; c). – d, A-B-C-D-E-Modell Modifikation bei den Einkeimblättrigen, z. B. bei Tulpe (*Tulipa bakeri*; e) und Lilie (*Lilium bulbiferum*; f): „Modell der sich verschiebenden Grenzen“ („sliding-boundary-model“). – g, A-B-C-D-E-Modell für basale Angiospermen, z. B. bei Sternanis (*Illicium anisatum*) und Stern-Magnolia (*Magnolia stellata*): „Modell der verwischten Grenzen“ („fading-borders-model“). A, B, C, D, E: Wirkungsorte der Gene, C = Karpell (Fruchtblatt), P = Petalum (Kronblatt), S = Sepalum (Kelchblatt), Sa = Samenanlage, St = Stamen (Staubgefäß), T = Tepalum (undifferenziertes Blütenblatt). Verändert aus Erbar 2010.

sogenannte Transkriptionsfaktoren, d. h. Proteine, die spezifische DNA-Regionen anderer Gene erkennen und deren Transkription (Ablesen und Vervielfältigen der Informationen auf dem Gen) beeinflussen.

Das vorgestellte Modell beschreibt aus genetischer Sicht, wie die Organidentität reguliert wird. Bestimmte Mutanten von *Arabidopsis thaliana* zeigen, dass unabhängig von der Organidentität die Organorte am Blütenvegetationskegel fixiert sind. Beispielsweise können vier zusätzliche Staubgefäße die Orte der vier Kronblätter einnehmen oder sämtliche Staubgefäße sind durch sechs zusätzliche Fruchtblätter ersetzt. Wir können daraus schließen, dass die Stellungsverhältnisse in der Blüte auf einer anderen genetischen Ebene als auf der der Organidentität reguliert werden. Ungeklärt ist auch noch die genetische Basis für die polymeren Androeceen (siehe Kapitel 2). Auf welcher Hierarchie-Ebene des regulatorischen Netzwerks der Genfunktionen findet das Umschalten zu den Entwicklungsmustern spiralig oder wirtelig oder zentrifugal bzw. zentripetal auf Bündelanlagen statt?

Da bei den bisher untersuchten Blütenpflanzen jeweils homologe A-, B-, C-, D- und E-Gene gefunden wurden, darf angenommen werden, dass die Festlegung der Blütenorganidentität nach weitgehend konservierten Grundprinzipien erfolgt. Das bisher vorgestellte A-B-C-D-E-Modell erfasst die Verhältnisse in Blüten wie sie bei den meisten sogenannten Eudikotylen zu finden sind. Ungefähr 74 % aller Blütenpflanzen-Arten gehören zu dieser Gruppe, bei der die Blütenorgane meist in fixierter Zahl in fünf (z. B. *Arabidopsis*, Abb. 29b) oder vier Kreisen (z. B. Löwenmäulchen *Antirrhinum*, Abb. 29c) angeordnet sind. Im Gegensatz dazu sind bei den Einkeimblättrigen Blütenpflanzen (Monokotyle), die ungefähr 22 % aller Blütenpflanzen-Arten stellen, die Blütenorgane typischerweise in fünf dreizähligen Kreisen angeordnet, und die Blütenhüllblätter sind in vielen Fällen gleich gestaltet (z. B. Tulpe *Tulipa*, Lilie *Lilium*, Abb. 29e–f). Wie muss das A-B-C-D-E-Modell „abgewandelt“ werden? In der Tulpe – wie wahrscheinlich bei vielen anderen Monokotylen – werden die B-Gene in beiden Blütenhüllblattkreisen (und in den Staubgefäßen) exprimiert, d. h. es kommt verglichen mit dem Eudikotylen-Modell zu einer Ausdehnung der B-Funktion zur Peripherie der Blüte hin (Abb. 29d). Diese Modifikation lässt sich als „Modell der sich verschiebenden Grenzen“ („sliding-boundary model“) im Wirkungsort der B-Gene beschreiben, da die in beiden Kreisen einheitliche Blütenhülle (Perigon) durch eine Verschiebung der B-Funktion nach außen (Expansion) erhalten wird.⁵

Die verbleibenden ca. 4 % der Blütenpflanzen-Arten gehören zu den Basalen Blütenpflanzen, die neben beispielsweise dem Sternanis *Illicium* (Abb. 29h) hauptsächlich den „Verwandtschaftsbereich der Magnolien“ (Abb. 29i) umfassen (siehe Fußnote 2). Diese Gruppe zeigt trotz der geringen Artenzahl eine beträchtliche Diversität in den Blütenformen und -strukturen. Die Blüten variieren in Größe, Organzahl und -stellung. Gene der B-Klasse werden sowohl in spiralig oder wirrtelig angeordneten Blütenhüllblättern als auch in den Staubgefäßen exprimiert. Allerdings scheint das Expressionsmuster in der Blütenhüllentwicklung weder einheitlich noch konstant zu sein (in den Eudikotylen ist die B-Wirkung in allen Entwicklungsstadien der Kronblätter und Staubgefäße gleichmäßig). In ähnlicher Weise ist die Expression der A- und C-Gene in einem größeren Bereich

⁵ Eine andere Möglichkeit ist beim Sauer-Ampfer (*Rumex acetosa*, Polygonaceae) verwirklicht. Die einkreisige Blütenhülle wird ohne die Mitwirkung der B-Funktion realisiert, d. h. die B-Funktion wird nach innen verschoben, erfährt also eine Kontraktion.

des Blütenvegetationskegels zu beobachten. Diesen Befunden wird im „Modell der verwischten Grenzen“ („fading borders model“) der Wirkungsorte der Gene A, B und C Rechnung getragen (Abb. 29g), das postuliert, dass in den Basalen Blütenpflanzen die Organidentität durch eine breite und überlappende Expression der Blütengene reguliert wird, wobei die Expression zu den Grenzen der Wirkungsweise hin schwächer wird.

Das A-B-C-D-E-Modell mit festgelegten Grenzen der Wirkungsorte der Gene (Abb. 29a) trifft also nur auf die Kern-Eudikotylen zu und dürfte den Endpunkt einer Entwicklungsreihe mit fließenden und überlappenden Wirkungsorten der Gene (Abb. 29g) darstellen. Mit anderen Worten: Das Muster der breiten Genexpression der homologen A-, B- und C-Gene ist ursprünglich und das A-B-C-Modell mit den definierten Grenzen der Wirkungsorte der Gene ist abgeleitet.

Bereits bei den Nacktsamern (Gymnospermen), also den Vorfahren der Blütenpflanzen (Bedecktsamer, Angiospermen), finden sich Homologe der floralen homöotischen Gene der Klassen B, C und D, wobei die ursprüngliche Funktion der B-Gene in der Differenzierung von männlichen (B-Gen-Expression „an“) bzw. weiblichen Reproduktionsorganen (B-Gen-Expression „aus“) liegt.

Aus phylogenetischer Sicht ergeben sich vier Schlussfolgerungen:

1. Das A-B-C-D-E-System der Blütenorganidentität lässt sich von einem älteren B-C-D-System ableiten.
2. Bei der Angiospermen-Evolution muss also eine Veränderung in der B-Funktion von der reinen Kontrolle der Geschlechtsbestimmung vielleicht über schon gefärbte Blütenblätter zur eindeutigen Spezifizierung der Kronblätter stattgefunden haben.
3. A- und E-Gene sind wohl in der Samenpflanzen-Evolution erst später als die B-, C- und D-Gene entstanden (wobei die Rolle der A-Gene noch nicht für alle Großgruppen gänzlich geklärt ist).
4. Da speziell die Gene der Klasse E für die eindeutige Bestimmung der Organidentität in allen Organkategorien nötig sind, scheint diesen eine Schlüsselrolle bei der Entstehung der angiospermen Zwitterblüten, die durch den Einschluss der Samenanlagen in Fruchtblättern charakterisiert sind, zuzukommen.

Die Möglichkeiten zu großen Veränderungen ergeben sich sicherlich aus den immer wieder stattgefundenen Genom-Duplikationen (Verdopplung des gesamten Erbguts), denen für die Differenzierung von Genfunktionen eine große evolutionäre Bedeutung zukommt (Jiao et al. 2011, Landis et al. 2018). Falls nach einer Genomverdoppelung beide Genkopien erhalten bleiben, gibt es mehrere Möglichkeiten für die Funktionsausprägung eines Gens, z. B. Neofunktionalisierung. Doppelt vorhandene Genkopien erlauben es der Pflanze, vorteilhafte Mutationen zu übernehmen und gleichzeitig eine ursprüngliche Kopie wichtiger Gene beizubehalten. Das ergibt eine „Spielwiese“ für neue Möglichkeiten und somit sind Genomduplikationen Biodiversitätsbeschleuniger!

Unsere heutigen Erklärungen werfen aber weitere Fragen auf für weitere Forschungen: Was löst Genom-Duplikationen aus?

6 Ontogenien im unmittelbaren Zusammenhang mit einer Funktion

Schauen wir uns auf unseren Spaziergängen durch „Wald und Flur“ die Blüten der gerade blühenden Pflanzen mit Aufmerksamkeit (entspannter Hingabe!) an, wird uns ein Phänomen von weitreichender Bedeutung begegnen: Im Androeceum einer Blüte öffnen sich die Staubbeutel (Antheren), die normalerweise einem sogenannten Staubfaden (Filament) aufsitzen, nicht gleichzeitig, sondern nacheinander. Der Blütenstaub (die Pollenkörner) einer Blüte wird demnach portioniert abgegeben – ein ökonomisches Verhalten in der Freigabe der Pollenkörner für die Bestäubung etwa durch Insekten. Die portionierte Freigabe kann allgemein im Zusammenhang mit einer optimalen Anpassung an unterschiedlichste Witterungsverhältnisse und an nicht immer verlässliche Bestäuber interpretiert werden. Die Schädigung offenen dargebotener Pollenkörner durch Regen (z. B. Platzen der Pollenkörner aus osmotischen Gründen) wird, betrachtet man die gesamte Pollenproduktion einer Blüte, minimiert; Pollenkörner, die sich noch in den geschlossenen Staubbeuteln befinden, werden weitgehend geschützt.

Besonders beeindruckend ist der Blühverlauf bei der Weinraute *Ruta graveolens* (Rutaceae). Die Blüten sind vormännlich (proterandrisch); auf eine männliche Blühphase, die unter normalen Bedingungen etwas mehr als zwei Tage dauert – bei schlechten Wetterbedingungen verzögert sie sich –, folgt eine weibliche Phase, indem sich der Griffel des Fruchtknotens streckt und am Ende eine reife Narbe bildet (erst jetzt können die auf ihr landenden Pollenkörner einen Schlauch bil-

den, um die Geschlechtszellen des männlichen Gametophyten zur Eizelle des weiblichen Gametophyten in der Samenanlage im Fruchtknoten zu transportieren). Die meisten Blüten sind aus fünf vierzähligen Kreisen aufgebaut (4 Kelch-, 4 Kronblätter, 4+4 Staubgefäße, 4-karpelliger Fruchtknoten). Beim Aufblühen (Beginn der männlichen Blühphase) sind alle acht Staubgefäße in der Blüte nach außen gespreizt. Das Portionierungsmuster startet sofort damit, dass durch einseitiges Wachstum der Staubfaden- (= Filament-) Basis eines über einem Kelchblatt positioniertes Staubgefäß aufgerichtet wird; die sich öffnende Anthere befindet sich jetzt über dem in Entwicklung begriffenen (also noch sterilen) Griffel. Dann erheben sich die drei weiteren über den Kelchblättern stehenden Staubgefäße sukzessive, während die jeweils vorausgegangenen durch ein umgekehrtes einseitiges Filamentwachstum wieder in Spreizlage versetzt werden. Durch die gegenläufig (einmal auf der Außenseite, einmal auf der Innenseite) sich abspielenden Wachs-

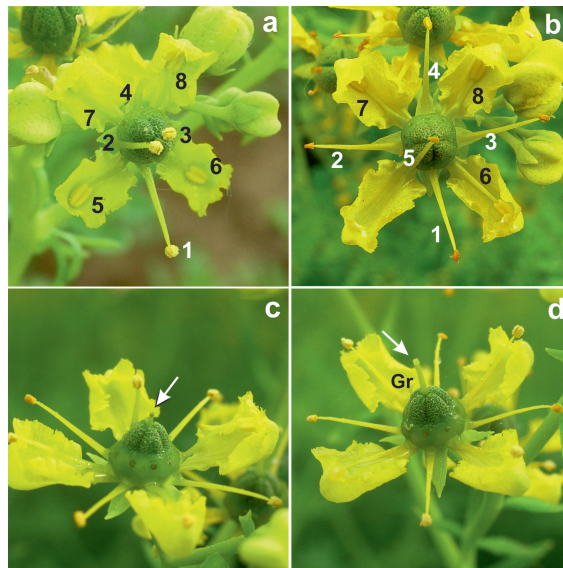


Abbildung 30: Sukzessive Staubbeutel-Präsentation bei der Weinraute (*Ruta graveolens*). Die Nummerierung der Staubgefäße entspricht der Sequenz der Bewegung bzw. Öffnung – a, Staubbeutel no. 2 kurz vor dem Zurückbiegen; Staubbeutel no. 3 kurz vor dem Erreichen der Pollen präsentierenden Position. – b, Staubbeutel no. 5 in Pollen präsentierender Position. – c, Blüte nach der männlichen Blühphase; der Griffel (Pfeil) beginnt sich zu strecken. – d, Blüte in der weiblichen Blühphase mit verlängertem Griffel (Gr) und empfangnisbereiter Narbe (Pfeil). Verändert aus Erbar et al. 2012.

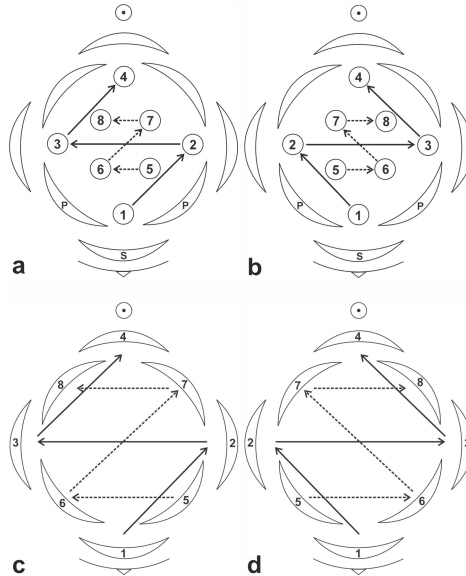


Abbildung 31: Entwicklungsdiagramme vierzähliger *Ruta*-Blüten. – a-b, Zickzack-Muster im Androeceum: rechtsherum (a) und linksherum beginnend (b). Durchgezogene Pfeile: Sequenz der über den Kelchblättern stehenden Staubgefäße; gestrichelte Pfeile markieren die Sequenz der über den Kronblättern befindlichen Staubgefäße. – c–d, Die entsprechenden Sequenzen in Kelch und Blumenkrone. P = Petalum (Kronblatt), S = Sepalum (Kelchblatt). Aus Erbar et al. 2012.

tumsprozesse an der Filamentbasis verlängert sich das Filament insgesamt. Das Portionierungsmuster gehorcht stets einer Zickzackfolge, die entweder links- oder rechtsherum verlaufen kann (siehe Abb. 30–31). Die über den Kronblättern stehenden Staubgefäße zeigen ebenfalls eine entsprechende Zickzackfolge und zwar gegenläufig zu jenen über den Kelchblättern.

Reiht man jüngste Entwicklungsstadien der Weinrautenblüte aneinander, so wird klar, dass nicht nur das Zickzackmuster in der Folge der Organanlagen im Androeceum, sondern auch in Kelch und Krone bereits manifestiert ist (siehe Abb. 3–10 in Erbar et al. 2012).

Nachdem alle acht Staubgefäße ihre zweite Spreizlage erreicht haben, beginnt die weibliche Blühphase (Abb. 30d). Es lässt sich immer wieder beobachten, dass am Ende der weiblichen Phase die Staubgefäße sich nochmals aufrichten; meist ist bereits der Fruchtknoten schon vergrößert und schickt sich an, in die Entwick-

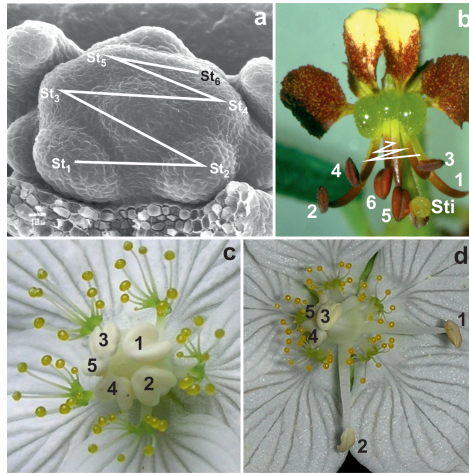


Abbildung 32: Sukzessive Staubbeutel-Präsentation. – a–b, Violette Spinnenpflanze *Cleome violacea* (Cleomaceae). Nummerierung der Staubgefäße entsprechend der Anlagen-Sequenz in der Blütenknospe (a) und der Antherenöffnungsfolge in einer adulten Blüte (b); zusätzlich ist das Zickzack-Muster durch eine entsprechende Linie verdeutlicht. – c–d, Studentenröschen *Parnassia palustris* (Parnassiaceae). – a, Blüte kurz vor dem Beginn der männlichen Blühphase; Nummerierung der Staubbeutel entsprechend der Größenabnahme und der vermuteten Öffnungsreihenfolge. – d, Staubbeutel no. 3 in der Pollen präsentierenden Position. Verändert aus Erbar et al. 2012.

lung zur Frucht überzugehen. Vielleicht versteckt sich hinter diesem Verhalten eine eventuelle Notfallsituation, nämlich ein letzter Versuch, falls eine Fremdbestäubung durch ein Insekt fehlgeschlagen ist, wenigstens eine Selbstbestäubung herbeizuführen, wenn diese auch weniger einträglich sein kann (Inzuchtdepression).

Zickzackmuster sowohl in der Anlagenfolge als auch im portionierten Verstäubungsverhalten finden sich auch in anderen Verwandtschaftsbereichen, wie etwa bei den mit den Kaperngewächsen (Capparaceae) verwandten Spinnenpflanzengewächsen (Cleomaceae) sowie den Herzblattgewächsen (Parnassiaceae; siehe Abb. 32).

Allgemein kann festgehalten werden, dass die „Reifungsfolge“ der Staubgefäße in der Regel ihrer Anlagenfolge entspricht. So kann beispielsweise in den hochpolymeren spiraligen Androeceen die männliche Phase durch viele Portionierungsschritte die Pollenverteilung raum-zeitlich gedehnt beziehungsweise aufgefächert werden, was letztlich zu einer Erhöhung der Rekombinationsrate führen kann.

Ein sparsamer Umgang mit den Pollenkörnern (Pollenkörner sind für die Pflanze „teure“ Produkte, da sie einerseits für die sexuelle Fortpflanzung benötigt werden und andererseits viele stickstoffreiche Proteine enthalten) gehört zu den wichtigsten Faktoren der Bestäubungsoptimierung. Er kann in vielen Fällen durch andere Mechanismen erreicht werden, nämlich durch die sogenannte **sekundäre Pollenpräsentation**. Immer wieder im System der Blütenpflanzen vorkommend, sind sie im obersten Stammbaumniveau, den Asterales, zur Regel geworden (Leins & Erbar 2006). Man versteht darunter, dass die Pollenkörner nicht direkt aus ihrem Bildungsort, den Staubbeuteln (Antheren), sondern über Hilfsstrukturen den bestäubenden Tieren zum Transfer dargeboten werden. Im Falle der Köpfchenblütler (Asteraceae) ist es der Griffel, der für die Pollenpräsentation in verschiedener Weise genutzt wird. Unabdingbar für die in den Köpfchenblütlern vorkommenden Mechanismen ist die für die Familie so charakteristische Antherenröhre, die aus fünf seitlich miteinander verklebten Staubbeuteln besteht. Die beiden häufigsten Mechanismen der sekundären Pollenpräsentation in den Asteraceae sind der Bürsten- und der Pumpmechanismus, die sich in der Anordnung von Haaren am Griffel und in der Position der Griffelspitze zum Zeitpunkt der Öffnung der Staubbeutel in der noch geschlossenen Blütenknospe unterscheiden. Der Bürstenmechanismus (Abb. 33a) ist typisch für die Unterfamilie Cichorioideae, zu denen etwa die Wegwarte (*Cichorium intybus*) und der Kopfsalat (*Lactuca sativa*) gehören. Für den einfachen aber effizienten Bürstenmechanismus wird die Antherenröhre vor dem Öffnen der Blüte durch Wachstum der Staubfäden bzw. der Staubgefäß-Kronblatt-Röhre (Stamen-Corollentubus) auf das gleiche Niveau wie

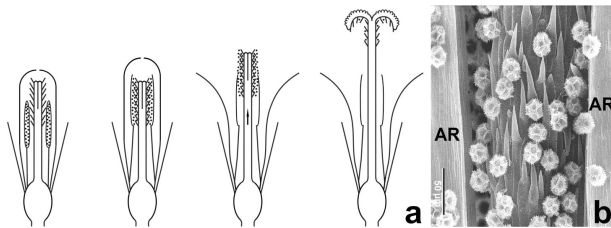


Abbildung 33: Der Bürstenmechanismus der sekundären Pollenpräsentation bei den Asteraceae. – a, In stark schematisierten Längsschnitten; von links nach rechts jeweils verschiedene Stadien von der noch geschlossenen Knospe bis zur weiblichen Blühphase. – b, Rainkohl *Lapsana communis*. Ausschnitt aus der Griffelbürste in der aufpräparierten Staubbeutelröhre (= Antherenröhre AR; rasterelektronenmikroskopisches Bild).

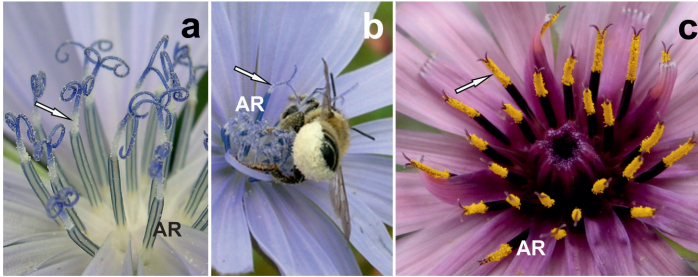


Abbildung 34: Das Pollenbürsten bei den Cichorieae (Asteraceae, Cichorioideae). – a–b, Wegwarte *Cichorium intybus*. – c, Roter Bocksbart *Tragopogon porrifolius*. Der Pollen wird außen am Griffel (Pfeile) über der Antherenröhre (AR) präsentiert.

die „Fegehaarzone“ des Griffels gebracht, so dass diese in ihrer gesamten Länge den Innenraum der Röhre ausfüllt (Abb. 33b). Nach Öffnen der Antheren (nach innen zu) beginnt der Griffel sofort sich zu verlängern. Er bürstet dadurch die Pollenkörner aus der Röhre heraus, die dann in der männlichen Blühphase außen am Griffel präsentiert werden (Abb. 34). Durch Narbenspreizung wechselt später die Blüte in die weibliche Phase über.

Der größere Teil der Köpfchenblütler, die Asteroideae, zu denen Gänseblümchen (*Bellis*) und Sonnenblume (*Helianthus*) gehören, weist überwiegend einen sogenannten Pumpmechanismus auf (Abb. 35a). Vor der Anthese wird durch Verlängerung der Filamente bzw. des Stamen-Corollentubus die Antherenröhre über die Griffelspitze gehoben (Abb. 35b–c). Zu Beginn der Anthese verschließt somit die Griffelspitze die Antherenröhre nach unten zu, so dass die in die Röhre entlassenen Pollenkörner nicht herausfallen können. Als Kolben wirkend schiebt der sich verlängernde Griffel, der nur an seinen beiden Narbenspitzen Fegehaare trägt (Abb. 35b–c), den Pollen vor sich her. Letzterer tritt am oberen Röhrende aus (Abb. 36) und kann nun Portion für Portion von den Blütenbesuchern aufgenommen werden. Zu einer feineren Portionierung kommt es zusätzlich dadurch, dass die für die gesamte Familie der Köpfchenblütler charakteristischen dreieckigen Anhängsel an den Antherenspitzen (Konnektivanhängsel) so zusammenneigen, dass die Pollenkörner nur an den fünf Schlitzen zwischen diesen austreten können: Kleine Sterne leuchten dem Betrachter entgegen (Abb. 36a–b).

Es gibt auch die Kombination von Pumpen und Bürsten, etwa bei den Sonnenblumen (*Helianthus*) oder den Dahlien (*Dahlia*). Lange, behaarte Griffelanhängsel

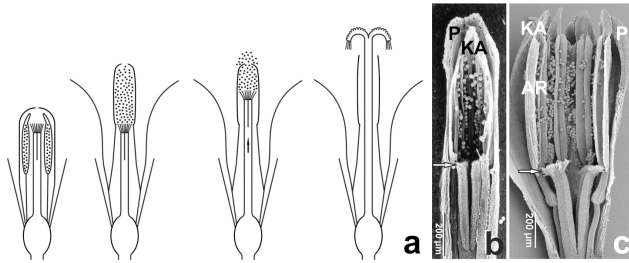


Abbildung 35: Der Pumpmechanismus der sekundären Pollenpräsentation bei den Asteraceae. – a, In stark schematisierten Längsschnitten; von links nach rechts jeweils verschiedene Stadien von der noch geschlossenen Knospe bis zur weiblichen Blühphase. – b, Lage der Griffelspitze kurz vor Beginn des Pollenpumpens beim Greiskraut *Senecio vulgaris*; die Pollenkörner sind schon in die Antherenröhre abgegeben. Der Pfeil weist auf die Griffelspitze, die sich in Höhe der Basis der Antherenröhre befindet. In der aufgeschnittenen Antherenröhre finden sich präparationsbedingt nur noch wenige Pollenkörner (rasterelektronenmikroskopisches Bild). – c, Etwas jüngeres Stadium beim Frühlings-Greiskraut *Senecio vernalis*; die Pollensäcke beginnen sich gerade erst zu öffnen. Ein Staubgefäß wurde entfernt und die Antherenröhre (AR) aufgeklappt, um die Öffnung der Pollensäcke zu demonstrieren. Griffelschenkel mit endständigen Haaren sowie die dreieckigen Anhängsel an den Antherenspitzen (Konnektivanhängsel, KA) sind wichtig für einen reinen Pumpmechanismus (die Konnektivanhängsel dienen hier einer Feinportionierung).

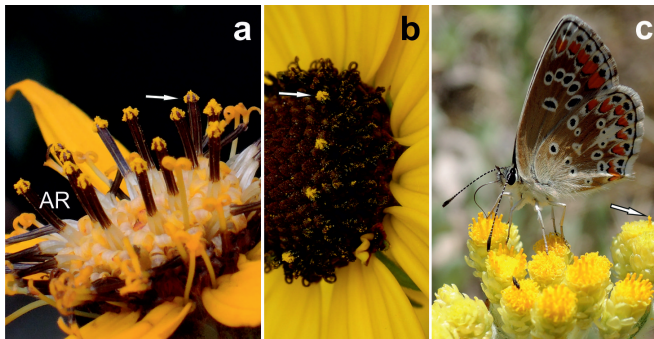


Abbildung 36: Das Pollenpumpen bei den Asteroideae (Asteraceae). – a, Gezählter Goldkolben *Ligularia dentata*. – b, Schwachfilziger Sonnenhut *Rudbeckia subtomentosa*. – c, Sand-Strohblume *Helichrysum arenarium*. Pfeile weisen auf Pollenhäufchen an der Spitze der Antherenröhren (AR). Durch die zusammenneigenden apikalen dreieckigen Konnektivanhängsel entstehen fünf Spalten, durch die die Pollenkörner „sternförmig“ wie aus einer Garniertülle herausgedrückt und dadurch feinportioniert den Bestäubern präsentiert werden.

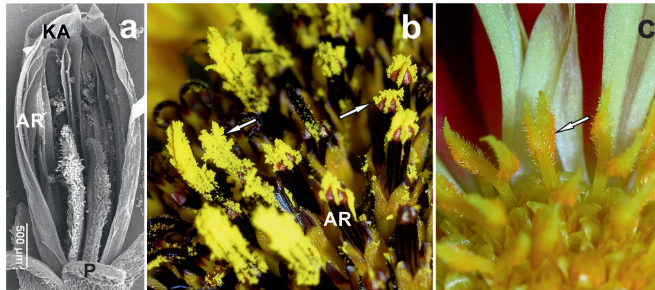


Abbildung 37: Die Kombination von Pump- und Bürstenmechanismus bei den Asteroideae (Asteraceae). – a, Gewöhnliche Sonnenblume *Helianthus annuus*. – b, Knollige Sonnenblume oder Topinambur *Helianthus tuberosus*. – c, Garten-Dahlie *Dahlia hortensis*. – a, Zum Zeitpunkt der Antherenöffnung befindet sich der behaarte Teil des Griffels im unteren Teil der Antherenröhre. Ein Teil der Pollenkörner wird in den oberen Teil der Antherenröhre abgegeben. – b, Durch Griffelwachstum werden die Pollenkörner aus dem oberen Teil der Antherenröhre, wie beim Pumpmechanismus, zwischen den dreieckigen Konnektivanhängseln „sternförmig“ aus der Antherenröhre herausgedrückt (Pfeil im rechten Teil des Bildes). In älteren Stadien der männlichen Blühphase werden die Pollenkörner, die von den Griffelhaaren aus den Antheren herausgebürstet wurden, außen am Griffel präsentiert (Pfeil in linker Bildseite). – c, Pfeil weist auf Pollenkörner am haarigen Griffelabschnitt. AR = Antherenröhre, KA = Konnektivanhängsel, P = Petalum (Kronzipfel).

ragen zum Zeitpunkt der Staubbeutelöffnung in die Antherenröhre hinein. Beim anschließenden Griffelwachstum wird ein Teil der Pollenkörner aus der Antherenröhre gepumpt, ein Teil wird gebürstet (Abb. 37).

Die Größe der Pollenportion ist u. a. mit dem Bestäuberspektrum korreliert. Ein großes Spektrum an Bestäubern wie Käfer, Bienen, Wespen, Fliegen und Schmetterlinge, die auf den offenen Blütenständen der Köpfchenblütler häufig anzutreffen sind, könnte zu einem beträchtlichen Pollenverlust führen, da einige dieser Insekten unzuverlässig sind, was den wiederholten Besuch der gleichen Pflanzenart betrifft. Ein Pollenverlust wird jedoch durch eine Limitierung der Pollenmenge, die eine einzelne Blüte gerade anbietet, reduziert.

7 Abschließende Bemerkungen

Strukturelle Entwicklungen von und in Organismen gründen sich auf differenzierende Wachstumsprozesse, zeichnen sich also durch Veränderungen in einem bestimmten zeitlichen Ablauf aus. Wenn die beiden Autoren dieses Artikels sich

auf die Formulierung des Textes einlassen, strapazieren sie verschiedene Areale ihrer linken Großhirnrinde, die die Sprachverarbeitung steuern: Wortwahl, Satzbau, Semantik, Kontrolle. Es sind die Synapsen, die Schaltstellen zweier Nervenzellen – bestehend aus einer Präsynapse (Senderhälfte) und einer Postsynapse (Empfängerhälfte) –, in denen richtungsorientiert Erregungen in die Nachbarzelle (Empfängerhälfte der Synapse = Postsynapse) in Form von mit Neurotransmittern gefüllten Vesikeln sich in den Zwischenraum der beiden Synapsenhälften (synaptischer Spalt) ergießen und die Neurotransmitter zum Andocken an Rezeptoren der Nachbarzelle (Postsynapse) präsentieren. Dieser Vorgang (Exocytose) dauert weniger als eine Millisekunde. Das dürfte wohl der schnellste strukturelle Prozess sein, der in unserem Körper stattfindet. Unser Körper selbst, umgeben von dem größten Organ, nämlich der Haut, die alle Innereien von der Außenwelt abschirmt und deren äußere Schicht (die Hornhaut oder Epidermis) sich durchschnittlich alle 28 Tage durch Zellteilungen im Stratum germinativum erneuert (nach außen zu werden die Abkömmlinge nach und nach verhornt und abgeschilfert), macht in seiner Gesamtheit eine über Jahre hinweg dauernde Entwicklung mit, von der Embryo-, Säuglings-, und Kindheitsphase, über die Pubertät zur Erwachsenen- (geschlechtsreifen) Phase, die schließlich in die Altersphase mündet. All diese Phasen sind begleitet von äußerst komplexen hinter- und ineinander geschachtelten Prozessen.

„Wenn wir ... die Erde von einem Kirchhofe nehmen, so nehmen wir die Theile, welche Menschen ausgemacht haben, und von Menschen in Erde verwandelt sind; führen wir dieselbe auf unsre Kohlgärten, und setzen Kohlpflanzen hinein, so bekommen wir Kohlköpfe statt Menschenköpfe; werden diese Kohlköpfe wiederum von Menschen gegessen, so verwandeln sie sich in Theile derselben. Auf diese Art essen wir unsere Todten, und sie bekommen uns wohl; ich meines Theils aber gestehe gern, daß ich, wenn ich es wüste, daß ich solchergestalt meinen oder eines anderen Großvater essen sollte, an dergleichen Kohl nicht gerne gehen würde, ich müste dann sehr starken Appetit haben.“ Linné 1745 (Übersetzung von 1765)

Diese Zeilen finden wir bei Carl von Linné im Bericht zu seiner Reise im Jahre 1741 durch Westgotland. Er stellt damit den Bezug der toten Organismen zu einem „Quasi-Gleichgewicht“ zwischen Produzenten (grüne Pflanzen), Konsumenten und

Destruenten (z. B. Mikroorganismen) her, das wir heute als Ökosystem bezeichnen. Mit den Ökosystemen haben wir wieder ineinander geschachtelte, über **längere** Zeiträume veränderliche (etwa durch Klimawandel), also nur vorübergehend abgrenzbare Komplexe vor uns. Sie ordnen sich in die unzähligen Veränderlichkeiten von Allem ein.

Panta rhei – Alles fließt – eine Aussage, die dem antiken Philosophen Heraklit von Ephesos (um 500 v. Chr.), über dessen Denkweise nur Fragmente überliefert sind, zugeordnet wird, verwickelt uns in eine hoch spannende Frage: Steckt hinter der ständigen Veränderung ein unbewegter Bewegter oder ist die Bewegtheit selbst das wahre Sein, das Zufälle nach einem ökonomischen „Schöpfungsprinzip“, nämlich der Optimierung hinsichtlich des Zwecks und der Funktionalität durch Konkurrenzdruck, selektiert?

Die Frage bewegt die Köpfe von Philosophen bis zum heutigen Tag und wird sie, solange es Köpfe gibt, weiterhin bewegen, auch wenn letztere zuweilen nur ein Sandkorn in der Wüste in Bewegung zu setzen vermögen. Allemal ist es vergnüglich, sich „dem lebendigen Geist“⁶ unter dem Motto „Semper Apertus“⁷ hinzugeben.

Literatur

- Coen, E.S. & Meyerowitz, E.M. 1991: The war of the whorls: genetic interactions controlling flower development. – *Nature* **353**: 31–37.
- Cronauer, J. 2007: Blütenontogenetische Untersuchungen an den Vertretern der ehemaligen Scrophulariaceae: *Gratiola* und *Bacopa*. – Staatsexamensarbeit Universität Heidelberg, unpubl.
- Erbar, C. 1991: Symptety – a systematic character? – *Bot. Jahrb. Syst.* **112**: 417–451.
- Erbar, C. 1997: Fieberklee und Seekanne - Enzian- oder Aster-verwandt? Zur Blütenentwicklung und systematischen Stellung der Menyanthaceae. – *Bot. Jahrb. Syst.* **119**: 115–135.

⁶ Von Friedrich Gundolf formulierte Widmung des Neuen Universitätsgebäudes, 1930 errichtet durch amerikanische Spenden auf Initiative des US-Botschafters Jacob Gould Schurmann, ehemaliger Student in Heidelberg.

⁷ Motto der Universität Heidelberg.

- Erbar, C. 2007: Current opinions in flower development and the evo-devo approach in plant phylogeny. – *Pl. Syst. Evol.* **269**: 107–132.
- Erbar, C. 2010: Floral organ determination and ontogenetical patterns during Angiosperm evolution. – *Int. J. Plant Dev. Biol.* **4** (Special Issue): 1–16.
- Erbar, C. & Leins, P. 1988: Blütenentwicklungsgeschichtliche Studien an *Aralia* und *Hedera* (Araliaceae). – *Flora* **180**: 391–406.
- Erbar, C. & Leins, P. 2004: Sympetaly in Apiales (Apiaceae, Araliaceae, Pittosporaceae). – *S. Afr. J. Bot.* **70**: 459–468.
- Erbar, C. & Leins, P. 2011: Synopsis of some important, non-DNA character states in the asterids with special reference to sympetaly. – *Plant Div. Evol.* **129**: 93–123.
- Erbar, C. & Leins, P. 2019: Ästhetische Ökonomie in der Gestalt der Pflanzen – HDJBO **4**: 39–90. doi:10.17885/heiup.hdjbo.2019.0.24006.
- Erbar, C., Franz, C. & Leins, P. 2012: Recent investigations on the pattern of pollen portioning in *Ruta graveolens* (Rutaceae). – *Plant Div. Evol.* **130**: 159–177.
- Jiao, Y., Wickett, N.J., Ayyampalayam, S., Chanderbali, A.S., Landherr, L., Ralph, P.E., Tomsho, L.P., Hu, Y., Liang, H., Soltis, P.S., Soltis, D.E., Clifton, S.W., Schlarbaum, S.E., Schuster, S.C., Ma, H., Leebens-Mack, J. & dePamphilis, C.W. 2011: Ancestral polyploidy in seed plants and angiosperms. – *Nature* **473**: 97–100. doi:10.1038/nature09916.
- Landis, J.B., Soltis, D.E., Li, Z., Marx, H.E., Barker, M.S., Tank, D.C. & Soltis, P.S. 2018: Impact of whole-genome duplication events on diversification rates in angiosperms. – *Amer. J. Bot.* **105**: 348–363.
- Leins, P. & Erbar, C. 1985: Zur frühen Entwicklungsgeschichte des Apiaceen-Gynoeceums. Ein Beitrag zur Entmythologisierung des unterständigen Fruchtknotens. – *Bot. Jahrb. Syst.* **106**: 53–60.
- Leins, P. & Erbar, C. 2006: Secondary pollen presentation syndromes of the Asterales – a phylogenetic perspective. – *Bot. Jahrb. Syst.* **127**: 83–103.
- Leins, P. & Erbar, C. 2008: Blüte und Frucht. Morphologie, Entwicklungsgeschichte, Phylogenie, Funktion und Ökologie. 2. Aufl. – Stuttgart: Schweizerbart'sche Verlagsbuchhandlung.
- Leins, P. & Erbar, C. 2010: Flower and Fruit. Morphology, Ontogeny, Phylogeny, Function and Ecology. – Stuttgart: Schweizerbart Science Publishers.
- Leins, P. & Erbar, C. 2018: Bäume und Sträucher in Frühjahr und Sommer erkennen. Bebilderte Steckbriefe und allerlei Begleitgeschichten aus Biologie, Mythologie und Verwendbarkeit. – Stuttgart: Schweizerbart'sche Verlagsbuchhandlung.
- Linné, C. von 1745: Reisen durch Oeland und Westgothland. – Aus dem Schwedischen übersetzt von D.C.D. Schreber 1765 (Teil: Westgothland). – Halle: Curts.

- Merxmüller, H. 1972: Systematic botany – an unachieved synthesis. – Biol. J. Linn. Soc. **4**: 311–321.
- Payer, J.-B. 1857: Traité d'organogénie comparée de la fleur. – Paris: Masson.
- Sattler, R. 1973: Organogenesis of flowers: a photographic text atlas. – Toronto & Buffalo: Univ. of Toronto Press.
- Schleiden, M.J. 1846: Grundzüge der wissenschaftliche Botanik 2. Teil. 2. Aufl. – Leipzig: W. Engelmann.
- Zimmermann, W. 1934: Genetische Untersuchungen an *Pulsatilla* I-III. – Flora **129**: 158–234.
- Zimmermann, W. 1961: Zur Phylogenie der Urlandpflanzen. – Ber. Geol. Ges. **6**: 348–367.
- Zimmermann, W. 1966: Kritische Beiträge zu einigen biologischen Problemen. VII. Die Hologenie. – Z. Pflanzenphysiol. **54**: 125–144.
- Zimmermann, W. 1968: Evolution und Naturphilosophie. – Berlin: Duncker & Humblot.

Über die Autoren

Prof. Dr. **Claudia Erbar** wurde nach ihrem Biologie- und Chemiestudium, das sie mit dem 1. Staatsexamen in beiden Fächern abschloss, an der Universität Bonn zum Dr. rer.nat. promoviert. Für ihre Staatsexamensarbeit erhielt sie den Preis der Konrad-Adenauer-Stiftung für Studierende der Botanik. Seit 1983 ist sie Wissenschaftliche Mitarbeiterin an der Universität Heidelberg. Nach der Habilitation im Fach Botanik 1993 hat sie seit 2000 eine apl. Professur und ist heute Forschungsgruppenleiterin für das Gebiet „Blütenbiologie und Evolution“ am Centre for Organismal Studies (COS) Heidelberg. Ihre Forschungsschwerpunkte liegen in den Bereichen der Blütenentwicklungsgeschichte vor dem Hintergrund morphologisch-funktioneller Fragestellungen, Blütenökologie (Blütenfunktionen bei Bestäubung und Befruchtung und der Interaktion mit Insekten) und Verwandtschaft und Evolution der Blütenpflanzen. In der Lehre vertritt sie auch die Pflanzengeographie.

Prof. Dr. **Peter Leins** studierte Biologie, Chemie und Philosophie an den Universitäten Tübingen, Freiburg und München. In München wurde er zum Dr. rer.nat. promoviert. Anschließend war er Assistent, später Oberassistent, am Institut für Systematische Botanik bei Prof. Dr. Hermann Merxmüller. Zwischendurch erlernte er die Methodik der Pollenkunde bei Prof. Dr. Gunnar Erdtman in Stockholm. Er habilitierte sich im Fach Botanik mit einer pollensystematischen Forschungsarbeit an einer Compositengruppe. Danach erhielt er einen Ruf auf eine Professur am Botanischen Institut der Universität Bonn. Drei weitere Rufe folgten: FU Berlin, Uni Heidelberg, Uni Kiel (FU Berlin und Uni Kiel nicht angenommen). An der Universität Heidelberg war er zunächst Direktor des Instituts für Systematische Botanik und Pflanzengeographie und des Botanischen Gartens (später Abteilung Biodiversität und Pflanzensystematik des neu gegründeten Heidelberger Instituts für Pflanzenwissenschaften). Seit 2002 ist er im forschenden und lehrenden „Ruhestand“. Seine Forschungsschwerpunkte liegen in den Bereichen der Blütenmorphologie und -entwicklungsgeschichte, Blütenökologie, Blütenpflanzen-systematik, Pollenkunde, Ausbreitungsbiologie, Evolutionsbiologie und Biophilosophie.

Korrespondenz:

Prof. Dr. Claudia Erbar, Prof. Dr. Peter Leins

Universität Heidelberg

COS-Biodiversität und Pflanzensystematik

Im Neuenheimer Feld 345

69120 Heidelberg

E-Mail: erle@urz.uni-heidelberg.de

Homepage Erbar: <https://www.cos.uni-heidelberg.de/index.php/c.erbar>

Homepage Leins: <https://www.cos.uni-heidelberg.de/index.php/p.leins>

Cancer in the light of evolution – from cancer genomes towards novel treatment and prevention approaches

MATTHIAS KLOOR & MAGNUS V. KNEBEL DOEBERITZ

Abteilung für Angewandte Tumorbiologie, Institut für Pathologie,
Universitätsklinikum Heidelberg

Abstract

Cancer is one of the deadliest diseases and a major health burden for mankind. In the past, our conceptual understanding of cancer has largely been determined by the medical need to classify distinct stages of the disease to select the most appropriate therapy. This view created only a very limited understanding of the true processes governing the disease, as it only perceives a snapshot at the time point of operation or the initial therapy. Cancer, though, is the result of a long-lasting evolutionary process following exponential expansion rules. It usually takes years or even decades until an initially hidden transformed cell in one of our organs is being noticed through symptoms as a clinically detectable lesion. During this time of inappreciation, multiple genomic alterations accumulate from generation to generation of replicating cells, that are all governed by the nature of the initial molecular hit that was responsible to push normal somatic cell into the state of a true cancer precursor. To develop better tools to diagnose, treat and prevent cancer, it becomes more and more clear that it is essential to understand the dynamic genomic processes that allow pre-cancer cells to expand and continuously change their biological behavior. Recent advances in DNA sequencing techniques allowed to obtain an unprecedented amount of high-resolution data on genomic changes in cancer cells. This increasing knowledge

helps to reconstruct the genomic history of individual cancer cells. However, it also indicates that pre-cancerous lesions in all of us seem to be substantially more common than previously anticipated, suggesting that we are well armed with a broad range of genetic and immunological weapons to control or defeat the vast majority of initial cancer cells clones. We will discuss the current understanding of these evolutionary processes using the paradigmatic example of DNA mismatch repair-deficient, microsatellite-unstable cancer. We will outline how a better understanding of cancer evolution can guide the development of tailored tumor diagnostics and treatments and retain the clues to effective cancer prevention.

1 General considerations – cancer as a genetic disease

The evolution of life is driven by replication of genomes, and by genetic variation during replication. Whereas increased cell division and proliferation capacity represents an evolutionary advantage in unicellular organisms, the evolution of multi-cellular organisms required coordination and governance mechanisms to control and limit the proliferative activity of individual cells within an organism. Cell replication and continuous renewal of tissues has emerged as a powerful concept of maintaining the homeostasis of multicellular organisms; however, the price is the risk of developing cancer. Therefore, increasingly potent and multi-layered tumor-suppressive mechanisms evolved in parallel with the increasing complexity and size of multicellular organisms (1).

Similar to the evolution of organisms within a population and a given ecosystem, tumors are the result of continuous evolution of somatic cells within an organism. The principle of variability and selection postulated by Charles Darwin (2) shapes tumor development at all stages from initial pre-cancerous stages up to late stages of invasive cancer, metastasis formation and development of resistance under systemic treatment.

Paramount to the understanding of cancer as the result of an evolutionary process is the concept that cancer is caused by alterations of the genome. In pioneering work of 1914, Theodor Boveri proposed abnormal chromosomal aggregation as a cause of cancer (3). Following this theory and the experimental observations by others, Karl Heinrich Bauer, later Head of the Surgery Clinic at Heidelberg University and one of the founders of the DKFZ, in 1928 published a seminal book titled “Mutationstheorie der Geschwulst-Entstehung—Übergang von Körperzellen

in Geschwulstzellen durch Gen-Änderungen (Mutation Theory of the Origin of Tumors—Transition of Somatic Cells into Tumor Cells by Gene Changes)” (4). In his publication, Bauer provides profound considerations on genome changes on the cancer cell phenotype and clinical implications such as invasive growth and metastasis. Together with other landmark publications (5) outlining the impact of genome-damaging agents on cancer risk, the foundation for understanding cancer as a disease of the genome was laid, long before the molecular structure of DNA and the sequence of specific genes were discovered (6).

With new information from epidemiological, cell biology, and gene sequencing studies becoming available, the multistep hypothesis of cancer was more widely accepted and acquired an increasingly well-defined shape (7). In 1990, Fearon and Vogelstein proposed a genetic model for colorectal tumorigenesis (8). This model has been the paradigm of genetic changes promoting the development of colorectal cancer; moreover, it has described the multi-step nature of solid tumor development in an exemplary way. The concept that solid tumor development is the result of stepwise progression from normal cells over pre-cancerous cells towards invasive cancer, a process driven by defined genetic alterations, has been summarized in the “Three Strikes” hypothesis proposed by Vogelstein in 2015 (9).

Although this model is of tremendous importance for understanding the genetic alterations giving rise to human cancers, it only develops its full power if it is embedded in a broader evolutionary context, accounting for the fact that most, i.e. almost all, precancerous cells acquiring somatic mutations are prone to cell death or senescence and will never grow out to manifest cancers. A schematic illustration of cancer evolution as a process of random mutation and subsequent selection is provided in Figure 1.

2 Molecular classification of cancers

Therapy selection for cancer patients has been based historically entirely on the organ, from which a tumor originated, the tumor stage, i.e. its local and systemic spread, and tumor histology as seen in light microscopy. Although still the majority of therapeutic decision making is based on this type of information, molecular information has become increasingly important for the selection of appropriate tumor therapies.

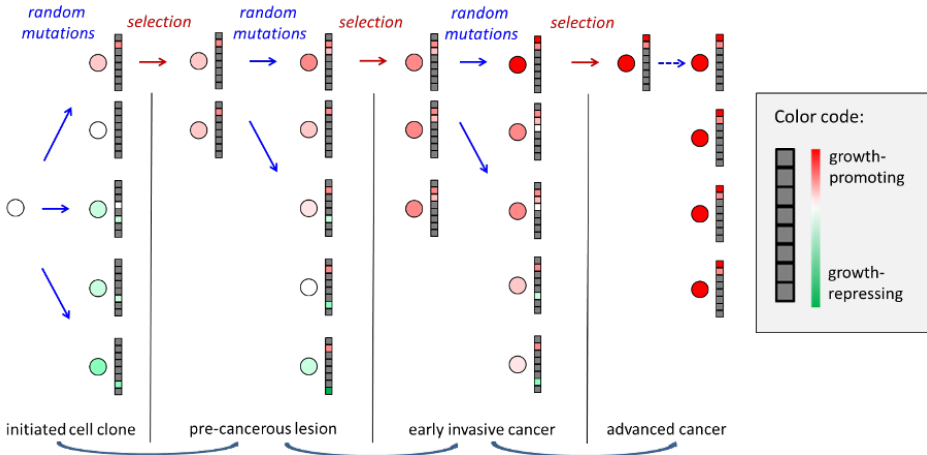


Figure 1: Schematic illustration of cancer evolution. Cancer evolution is a stepwise process that through the accumulation of somatic mutations leads to the malignant transformation of normal body cells (white circle on the right). This process, however, is not sequential and directed, but determined by the principles of random mutations and the subsequent selection of cell clones equipped to survive and grow out in the given environment. In contrast, after an initiation event, which crucially determines the subsequent evolutionary process (for example deficiency of the DNA mismatch repair system), random mutations occur and enhance the diversity in the cell pool, generating cells with enhanced (red) or reduced (green) proliferative capacity. Subsequently, diversity is reduced by selection, and survival of cells with high proliferative capacity is favored. For most cancer types, at least three crucial mutation events (driver mutations) are required to cause malignant transformation. The model implies that the majority of cells derived from the initiated cell will die and not grow out to the cancerous stage, because they acquired the wrong type or combinations of mutations. Conversely, cell clones in manifest cancers share characteristic combinations of mutations that occur significantly more frequently than expected by chance. The prevalence of a certain mutation in a group of manifest cancers is therefore a simple, but very helpful criterion for predicting its functional significance.

Therapeutic decisions based on molecular information were initially restricted to a few advanced stage cancers and a very limited number of specific gene mutations with well-characterized functional consequence. For example, the admission of targeted therapies using antibodies specifically blocking the epidermal growth factor receptor (EGFR) in colorectal cancer by the European Agency for the Evaluation of Medicinal Products (EMEA) (since 2009 European Medicines Agency, EMA) required the molecular determination of the presence or absence of a specific mutation activating the *RAS* (Rat sarcoma) oncogenes. The human

RAS gene family encompasses the *KRAS*, *NRAS*, and *HRAS* genes; activating mutations, which affect very specific positions of these genes, lead to a constitutive activation of the growth-promoting pathways that are usually triggered by binding of EGFR to activating ligands. Therefore, antibody blockade of the EGFR protein is ineffective in *RAS*-mutant tumors (10).

Similar to *RAS* mutation status, mutations of the *Breast cancer type 1 or 2* genes (*BRCA1* or *BRCA2*), which both code for repair enzymes involved in homologous recombination, are clinically relevant therapy predictors. Because homologous recombination is required for repairing double strand breaks in DNA, platinum salt-based chemotherapeutics that introduce such breaks, are effective in *BRCA1/2*-mutated cancers. Moreover, *BRCA1/2* mutations predict the effectiveness of novel tumor treatment modalities targeting the Poly-ADP-ribose polymerase (PARP inhibitors). PARP inhibitors impair tumor cells' capability of repairing treatment-induced DNA damage; therefore, they show activity in cells with homologous recombination deficiency such as *BRCA1/2*-mutant cancer cells (11).

The two examples of *RAS* mutations (as negative predictors) and *BRCA1/2* mutations (as positive predictors) illustrate a general principle, more and more molecular classifiers are used to guide clinical decision making in oncology. Interestingly, molecular characteristics such as *BRCA1/2* mutations often predict therapy response irrespective of the organ, from which the cancer has developed, and hence have significantly re-shaped the process of clinical decision making in oncology.

However, technically, these molecular tumor classifications are still end point classifications not directly accounting for the evolutionary history of a cancer. Nevertheless, they opened the path toward a profound reconsideration of the ways that can be used to classify tumors for clinical decision making: Although every individual cancer has a mutational landscape of its own, reflecting the randomness of single mutational events, computational neural network-based approaches were able to identify a few distinct common characteristics typical of certain evolutionary processes influencing tumor development, i.e. their *mutational signatures*.

3 Understanding cancer as the result of an evolutionary process – Mutational signatures

Mutational signatures mark a milestone on the road from an organ-based, phenotypical tumor classification to a more profound and comprehensive understanding of cancer as a dynamic process (12,13). This conceptual change represents a major paradigm shift, potentially comparable to the introduction of Darwinian evolution, which replaced the perception of species as discontinuous, constant entities by a non-discontinuous model of evolution, enabling the concept of phylogenesis.

A mutational signature is defined as a characteristic combination of specific types of mutations, which are associated with specific mutagenic processes underlying cancer formation. Such processes can be, as we will discuss in more depth, deficiencies of DNA repair systems or exogenous factors that cause certain types of DNA damage, e.g. UV light. For example, it has been known long before whole-genome sequencing techniques became available that certain tumor types were enriched for certain mutational base exchanges in the DNA: Smoking-induced lung cancer typically shows C/G to A/T transversions, whereas C/G to T/A transitions are commonly found in UV light-related skin tumors (14,15). The earliest alterations of a somatic cell's genome, if causing a deficiency of the machineries physiologically ensuring the fidelity of DNA replication, hence can define the mutational signature of the later evolving cancer cell clone. Its distinct molecular nature decides how the genome of this cell is being altered during the carcinogenic process and it critically also defines how this cell population will respond to individual therapy strategies.

For example, a deficiency of *BRCA1/2*, often observed in breast, prostate and ovarian cancers, leads to typical alterations, favoring certain types of basepair exchanges. Using computational approaches for detecting signatures of *BRCA1/2* deficiency, recently, a classifier was developed that with high accuracy was able to identify tumors as *BRCA1/2*-mutant without having knowledge on the actual *BRCA1/2* mutational status (16). Notably, a substantial proportion of tumors, though not harboring detectable *BRCA1/2* mutations, showed mutational signatures indicative of *BRCA1/2* deficiency and responded to platinum salt-based chemotherapy and PARP inhibitors, just as expected for *BRCA1/2*-mutant tumors.

This observation clearly indicates the power and superiority of using molecular processes of tumor evolution rather than actual snapshot mutational data for tumor classification. In terminology, the paradigm shift from mutational classification

to an evolution-based classification is represented by the introduction of the term “BRCAness”, which instead of referring to the presence or absence of *BRCA1/2* mutations underlines that *BRCA1/2*-related homologous recombination was impaired during tumor evolution, leading to the mutational signature of *BRCA1/2* deficiency.

Mutational signatures as fingerprints of evolutionary forces at work during tumorigenesis not only help in therapeutic decisions and drug selection in oncology. They also can be extremely helpful in the detective work, which is often required for the clarification of tumor etiology and the identification of cancers that arise in the context of a hereditary syndrome. Certain, otherwise rare base exchanges found in a tumor cell’s genome can suggest possible culprits, for example deficiency of the base excision repair system, thus guiding human genetics diagnostic approaches.

4 Mismatch repair deficiency

The strength of using mutational signatures or, in general, tumor evolution-related parameters for therapy selection instead of anatomical tumor location or the mere presence of a certain mutation within a tumor, is strongly underlined by the great success of introducing DNA mismatch repair deficiency as a predictor of immune therapy response.

The DNA mismatch repair (MMR) machinery represents a very important system for maintaining genome integrity during the division of somatic cells. Therefore, MMR enzymes are highly conserved across different species. Tumors that develop as a consequence of MMR deficiency accumulate tremendous amounts of somatic mutations (17). The phenotype of MMR deficiency has first been detected by chance in 1993. When actually searching for larger chromosomal alterations in tumors, Ionov and coworkers noticed that some human tumors showed an unexpected pattern; instead of chromosomal losses or deletions of larger parts of genomic material, these tumors presented with additional, aberrant genome fragments indicating a previously unknown type of genomic instability in human cancer (18).

The high mutational load commonly found in MMR-deficient tumors causes a high load of mutational antigens. These antigens are highly immunogenic, due to mechanisms we will discuss in more detail below, and can therefore readily

be recognized by the host's immune system. Therefore, MMR-deficient cancer patients show pronounced responses towards treatment with immune checkpoint modulators (19). Immune checkpoints physiologically represent "brakes" that prevent overshooting immune responses and autoimmune reactions detrimental for the host. In the scenario of cancer, however, immune checkpoints can trigger tolerance of the immune system towards otherwise highly immunogenic tumors such as MMR-deficient cancers. Here, the application of antibodies targeting immune checkpoints can reactivate the host's immune response against the tumor and lead to immune-mediated tumor cell killing. This approach is only effective if a tumor presents sufficient amounts of sufficiently immunogenic antigens to the immune system.

Notably, a substantial proportion of patients with MMR-deficient cancers resistant to all other available treatment modalities such as chemotherapy, radiation, or oncogene-targeted therapies, showed pronounced tumor shrinkage and sometimes even complete tumor elimination upon immune checkpoint blockade. The overwhelming clinical success of immune checkpoint blockade led to the FDA approval of immune checkpoint-blocking antibodies for the treatment of metastasized MMR-deficient cancer. Interestingly, this was the first admission of an oncological drug entirely based on a molecular classification, accounting only for molecular mechanisms at play during tumor evolution, but irrespective of the tissue origin.

The 2018 Nobel Prize in Physiology or Medicine has been awarded jointly to James P. Allison and Tasuku Honjo, who pioneered research on immune checkpoints and identified two major targets of current immune checkpoint therapies, CTLA-4 (cytotoxic T-lymphocyte-associated protein 4) and PD-1 (programmed cell death protein 1) (20–22).

Notably, the remarkable success of immune checkpoint blockade has only become possible in an era that has already started embracing principles of tumor evolution for clinical decision making.

In the following, we will focus on mismatch repair-deficient cancers to highlight general principles, driving forces, and clinical implications of the evolutionary processes responsible for cancer formation.

5 Mismatch repair-deficient cancers – a model of tumor evolution

The human genome consists of more than 3.2 billion base pairs, every single one of which theoretically can be affected by multiple types of alterations, including point mutations, large chromosomal rearrangements, and epigenetic changes. In order to understand key principles of tumor evolution and identify specific forces that shape tumors over time, it is helpful to reduce complexity by focusing on certain tumor types with defined mechanisms of genomic instability.

Mismatch repair-deficient cancers represent a highly valuable model in this regard. They are driven mainly by one specific type of somatic alteration that affects the genome, namely single base pair insertion or deletion mutations (indels). MMR deficiency-related indel mutations affect a limited number of well-defined regions of the human genome, short repetitive sequences termed “microsatellites”. In the scientific literature, MMR-deficient tumors are therefore also referred to as microsatellite-unstable (MSI) tumors, although the two terms are not exactly synonymous, the former referring to the forces shaping the evolutionary process and the latter to the resulting phenotype (17).

The likelihood of acquiring a mutation during the division of an MMR-deficient cell is elevated up to 100 times or more compared to other tumor cells. MMR-deficient tumors therefore not only represent a tumorigenesis model with a very well-defined type of somatic mutations, but also an extremely high mutational load, exposing the process of tumor evolution, in a figurative sense, under a magnifying glass.

6 Random mutations and selection in MMR-deficient cancers

As discussed above, the major type of mutation favored by MMR deficiency are indels, single nucleotide insertions or deletions. Indel mutations can affect all microsatellite sequences in the human genome. Most microsatellites are located in non-coding regions, and mutations affecting such non-coding microsatellites are mostly expected to have little or no effect on the phenotype or fitness of the cell. However, protein-encoding regions, which only represent about 1.5% of the entire human genome, are not entirely free from repetitive sequences. In fact, more than 1000 human genes encompass a repetitive microsatellite sequence of 8 base pairs or longer (for example an “AAAAAAAA” or A₈ sequence in the gene ACVR2A, or an “AAAAAAAAAA” or A₁₀ sequence in the gene TGFBR2) (23).

Whenever indel mutations strike at such coding microsatellite sequences, the effects on the protein level can be dramatic. The insertion or deletion of a single base pair can disrupt the translation of a gene into a protein by shifting the translational reading frame and disrupting the functionality of the protein product.

Our group and others have studied hundreds of microsatellites in thousands of MMR-deficient cancers. When we look at all microsatellites in the human genome to identify parameters that can predict the likelihood, with which a certain microsatellite displays a mutation in MMR-deficient cancers, a clear pattern emerges: the longer a microsatellite, i.e. the more repetitive units it contains (for example 10 in *TGFBR2* and 8 in *ACVR2A*), the higher its average mutation frequency in MMR-deficient tumors (24). This correlation can be approximated by a sigmoid function. It is rooted in the fact that with increasing numbers of repetitive units the correct replication gets increasingly difficult for the enzymes involved in DNA replication.

The sigmoid function describing the correlation between microsatellite length and microsatellite mutation frequency sets a very useful baseline for studying MMR-deficient tumor evolution, as it provides information about average microsatellite mutation frequencies depending on microsatellite length. For example, a microsatellite consisting of 8 basepairs, such as the repeat in the coding region of the *ACVR2A* gene, has a predicted mutation frequency in MMR-deficient colorectal cancer of 9.7% (www.seltarbase.org) (25). In other words, if we examine 100 MMR-deficient tumors for mutations at a randomly selected microsatellite of 8 basepairs length, we expect a total of approximately 10 mutations.

The wealth of microsatellite mutation data collected for all kinds of microsatellites in the human genome does not only inform us about average mutation frequencies, it also provides a range of expected mutation frequencies. For example, 95% of microsatellites consisting of 8 basepairs have a mutation frequency lower than 30.5% in MMR-deficient colorectal cancer. Considering these numbers, how can a microsatellite such as the A8 repeat in the *ACVR2A* gene display an actual mutation frequency of more than 80%, according to some studies even more than 90%, “much higher than expected”?

Although limited information is available about the influence of certain other parameters such as the sequence context, in which the microsatellite appears, there is clear evidence that the observed mutation frequencies in MMR-deficient cancers mainly reflect one factor: selection during tumor evolution. We have good reasons

to assume that microsatellite length determines the likelihood of an indel mutation to form during the division of an MMR-deficient cell. However, microsatellite length is completely unrelated to the likelihood, with which a cell affected by the respective mutation will survive. A microsatellite mutation that inactivates a gene essential for cell survival (illustrated as green boxes in Figure 1) will induce cell death in mutant cells and therefore disappear from the pool of precancerous cells. On the other hand, a microsatellite mutation inactivating a gene coding for a growth-suppressive protein (red boxes in Figure 1) may enable survival and even favor proliferation of the affected cell, ultimately providing it with the means to overgrow other competing precancerous cell clones.

If this hypothesis is true one would expect a significant enrichment of microsatellite mutations inactivating growth-suppressive genes in MMR-deficient cancers. The genome-wide maps of microsatellite mutations faithfully follow the prediction of random mutations (determined by microsatellite length) and consecutive selection, showing microsatellites located in tumor suppressor genes significantly enriched among those with high observed mutation frequencies (overrepresentation of red boxes in early invasive and advanced cancers, Figure 1) (26). Taking a closer look, commonly detected microsatellite mutations can be made responsible for the functional breakdown of a broad spectrum of genes, indicating how a single mutational mechanism enabled by the deficiency of the DNA MMR system provides the framework for the step-wise acquisition of all the hallmarks required for a cell to become a transformed malignant tumor cell (27).

7 The high immunogenicity of MMR-deficient tumors

The history of cancer immunology in the modern era even goes back to the late 19th century, when the American surgeon William B. Coley observed several cases of end-stage tumor patients who showed dramatic improvement of their tumor condition upon severe bacterial infection (28). Similar to complete remission of some MMR-deficient tumor patients upon immune checkpoint blockade, Coley observed complete tumor elimination in a patient with recurrent and inoperable sarcoma shortly after contracting erysipelas, a severe bacterial infection affecting the skin and the underlying lymphatic tissue. He later developed a mixture of killed bacteria, “Coley’s toxin”, to treat patients with malignant tumors. Similar clinical histories were reported by the German physicians Friedrich Fehleisen and

Wilhelm Busch who independently from each other reported significant tumor shrinkage in patients suffering from erysipelas (29,30).

These clinical observations very early indicated that the immune system in principle is capable of targeting tumor cells and, under certain conditions, even eliminating them. However, anti-tumor immune responses and the success of modern-day immune therapy, seemed to be restricted to certain tumor types.

As mentioned above, patients with MMR-deficient cancers respond very well to treatment with immune checkpoint inhibitors. Pronounced local immune responses were among the first characteristics that were noted when this new molecular subtype of human cancer was discovered (31). In MMR-deficient cancers, we often encounter extreme amounts of lymphocytes, immune cells of the adaptive immune system, which are trained to recognize and fight specific structures that indicate a danger to the integrity of the organism. The phenotypical observation of massive lymphocyte infiltration in MMR-deficient cancers suggests that the immune system reacts particularly strongly against emerging MMR-deficient cell clones during their evolution. It has long been known that the strength of the host's local immune response influences prognosis, the more lymphocytes at the tumor site, the longer the patients' average survival time (32). The pronounced immune response of the host represents the most likely reason for the comparatively good prognosis associated with MMR-deficient cancers (33).

Why are particularly MMR-deficient cancers such an attractive target for the immune system? Why would immune cells be more capable of recognizing and attacking tumors that exhibit MMR deficiency than other tumors?

The answer is only partly related to the high number of mutations accumulating during the evolution of MMR-deficient cancers. More importantly, the type of mutation plays a key role: Indel mutations responsible for the malignant transformation of MMR-deficient cells are frameshift mutations, this means that the deletion or insertion of a single base pair within a gene causes the wrong translation of the entire part of the gene that follows this mutation Figure 2. Whereas point mutations, such as the ones activating the *RAS* oncogenes described above, generally only lead to one wrong amino acid, indel mutations in MMR-deficient cancers can trigger the generation of long non-functional protein stretches that are completely unknown to the host's immune system. In this sense, immune responses against frameshift peptides (also termed "neo"-peptides) can mimic immune

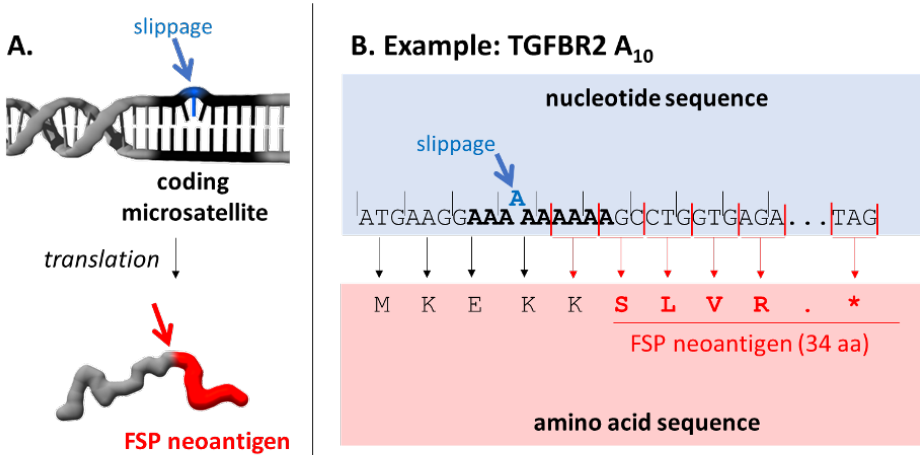


Figure 2: Mutations give rise to antigens that can be detected by the immune system. **A.** In MMR-deficient cells, indel mutations (i.e. losses or gains of single nucleotides, blue arrow) occur at repetitive microsatellite sequences located in genes that code for proteins. The translation of mutant nucleotide sequences results in the generation of a translational frameshift and protein sequences encompassing entirely novel amino acid stretches downstream of the mutation (red lines framing the shifted, novel nucleotide triplets in **B**). Key to this process is that the alteration of a single nucleotide, by causing a frameshift, affects the translation of the entire protein area that follows the mutation (red letters in **B**). Thus, a small alteration on the nucleotide level (blue) has big effects on the protein level (red) and for the immunological visibility of tumor cells.

responses directed against pathogen-related peptides, as there is no tolerance of the immune system.

In a simplified picture, MMR-deficient cancers are composed of cells, which are full of non-self proteins that are exposed to the immune system. As the capacity of the human immune system to recognize tumor cells as foreign increases with the number of antigenic proteins present in tumor cells, MMR deficiency with the dramatically increased frequency of indel mutations and the resulting translational frameshifts are a prototype of immunogenic tumors.

8 Immune surveillance and immunoediting

Patients with MMR-deficient cancer show immune responses against frameshift peptides generated by tumor cells. Our group could demonstrate that lymphocytes specifically recognizing these neo-peptide targets were present not only in the peripheral blood of these patients, but also locally, infiltrating the tumor (34).

This raises a critical question: How can MMR-deficient cancers survive and even prosper in the presence of potentially hostile immune cells?

In principle, two explanations are possible: On the one hand, the immune system may transiently or permanently lose the ability of controlling the outgrowth of MMR-deficient cancer cells, either by exhaustion or by suppression of immune cell functionality. On the other hand, tumor cells may become invisible to the immune system by “hiding” immunogenic antigens to immune cells. Molecular studies of manifest MMR-deficient cancers indicate that the latter mechanism appears to be more frequent, as impairments of the cellular antigen presentation machinery can be found in up to 70% of MMR-deficient cancers, at least those originating in the colorectum.

Physiologically, all nucleated cells of the body continuously present their spectrum of intracellular proteins to the immune system. This process is mediated by the major histocompatibility complexes (MHC), which in humans are also referred to as human leukocyte antigens (HLA). The HLA system serves several purposes, including identifying a cell as “self”, i.e. part of the organism. This is achieved by a high diversity of HLA genes between and among human populations. HLA diversity induces a high barrier between individuals; therefore, malignant tumor cells are not infectious and will not grow out to manifest lesions even if they are transferred to another host.

In addition to allowing for distinguishing self and foreign cells, HLA molecules also present intracellular protein fragments to circulating immune cells. This feature of the adaptive immune system, which apparently has evolved even before the separation of jawed and jawless vertebrates, approximately 500 million years ago (35), ensures that information about alterations of the intracellular protein content, which can indicate imminent danger e.g. in the context of a viral infection, is communicated to the immune system.

Similar to the situation in virally infected somatic cells, emerging tumor cell clones may accumulate non-self proteins that may be visible to the immune system upon HLA-mediated antigen presentation. Although rarely as prominent as in

MMR-deficient cancer, immune surveillance plays an important role in protecting humans from the development of many different types of tumors. The process of immune-mediated elimination of pre-cancerous cell clones and the outgrowth of cell clones capable of immune escape is often referred to as “immunoediting” (36).

The high frameshift mutation load and immunogenicity of MMR-deficient cell clones poses a serious threat for MMR-deficient cells to be eliminated by immune cells. Therefore, manifest MMR-deficient cancers show a wide spectrum of alterations impairing or completely abolishing the function of the physiological antigen presentation machinery.

The antigen presentation component most commonly inactivated by mutation in MMR-deficient cancer is *Beta2-microglobulin (B2M)*, an essential part of the most important cellular HLA complex (“HLA class I”), if no B2M protein is present, HLA class I breaks down, and no antigens can be presented to the immune system.

Again, as almost all alterations shaping the phenotype of MMR-deficient cancers, MMR deficiency is the enabling mechanism: the *B2M* gene encompasses four short microsatellite sequences (a (CT)₄ dinucleotide repeat in exon 1 and three five base pair repeats in exon 2), which represent mutational hot spots in MMR-deficient cancer (37). Using our model to estimate the average likelihood of mutations at a microsatellite of five base pairs length, we receive a prediction of 0.37%. In reality, however, more than 25% of MMR-deficient cancers harbor such hot spot mutations in *B2M*, underlining the massive positive selection pressure supporting the outgrowth of precancerous cell clones that have acquired one of these rare mutations.

The power of this selection pressure is supported by two additional observations: First, *B2M* mutations and other alterations that cause a breakdown of antigen presentation are mutually exclusive. This indicates that *B2M*-mutant cell clones only outcompete other cell clones if other clones still have an intact antigen presentation machinery and may therefore face the risk of elimination (38). Second, *B2M* mutations are significantly associated with the strength of the host’s local immune response; if local infiltration with activating immune cells is absent, no *B2M* mutations are detected in MMR-deficient cancers (39).

The example of *B2M* mutations offers another insight into fundamental principles of tumor evolution: Surprisingly, patients affected by *B2M*-mutant cancers have a very good prognosis and long survival. In fact, *B2M*-mutant tumors very rarely

metastasize, so they can in almost all cases be cured by surgery. This first seems counterintuitive, regarding the fact that *B2M* mutations abrogate HLA-mediated presentation of tumor antigens to the immune system and thereby stops potential tumor cell elimination by lymphocytes, the main effector cells of the adaptive immune system.

What could be the reasons for an improved survival of patients with *B2M*-mutant cancer? Although data are not yet definitive, it seems reasonable to assume that *B2M* mutation-triggered loss of HLA molecules on the cell surface deprives tumor cells of their ability to identify themselves as “self”, a process for which HLA molecules are essential. Loss of HLA on the surface may expose *B2M*-mutant tumor cells towards natural killer cells, effectors of the innate immune system, the most ancient defense system of multicellular organisms.

In an oversimplified analogy, emerging MMR-deficient cancer cells may face a situation resembling Odysseus between Scylla and Charybdis: either they fall prey to effectors of the adaptive immune system and die, or they through random mutation of *B2M* obtain the chance of sailing closer to Scylla, sacrificing the potential of metastatic spread by becoming targets of the innate immune system, but surviving for the time being.

The example of *B2M* mutations in MSI cancer that at the same time enable local tumor outgrowth and inhibit metastasis formation illustrates a principle common to the evolution of tumors and the evolution of species. Evolutionary adaptation allowing survival or increasing fitness under certain conditions may be unfavorable under other conditions or become unfavorable when environmental conditions change. This further extends the analogy between tumor and species evolution, underlining that tumor evolution may require massive adaptation under sudden environmental changes, including outgrowth in a completely different tissue context in the process of metastasis formation or when systemic cancer treatment is applied, often resulting in the generation of new tumor phenotypes mimicking new speciation (40).

9 Transmissible tumors – a special case of tumor evolution

Transmissible tumors represent a highly interesting group of malignant lesions that open a window into processes of cancer evolution and immune evasion that remain closed in clinically diagnosed human tumors. In contrast to virally induced

tumors, such as those caused by human papillomaviruses, transmissible tumors are caused by transmission of entire tumor cells from one host to another.

Conventionally, cancer cells are depending on the survival of the host, and the host's death marks an evolutionary dead-end. As described above, the diversity of MHC molecules between individuals of a species represent a strong inter-individual barrier that tumor cells cannot cross.

On some rare occasion, however, tumor cells have been documented to be transmitted from the original host to a recipient organism. Such clonally transmissible tumors are extremely rare in the human population. Tumor transmission in humans requires severe impairment of the immune system of the recipient organism, for example in immunodeficient individuals or after organ transplantation (41).

In animals, two examples of transmissible tumors have gained broader attention, the so-called “Devil facial tumour disease” (DFTD) in Tasmanian devils, and a transmissible venereal tumor (TVT, also termed Sticker sarcoma) in dogs. As immunosuppression is obviously not the reason for inter-individual spread of these tumors, the question is what is?

In Tasmanian devils', the genetic diversity is very low, probably due to rapidly declining population sizes during the last century (42). A low genetic diversity, accompanied by low diversity of the MHC gene loci, naturally favors transmission of tumor cells from one affected animal to another. On top of that, Tasmanian devils are involved in fighting, during which they bite each other, thereby supporting the spread of cells; in fact, aggressive male devils are more affected by transmissible tumors.

Obviously, low population diversity cannot explain the astonishing spread of TVT, a canine transmissible venereal tumor that affects dogs all over the world. However, recent genomic sequencing studies of more than 500 TVT lesions sampled over the whole world have revealed highly interesting facts about the early years of TVT evolution. At the time of its origination in the “founder dog”, presumably between 4000 and 8500 years ago in Asia, the originating population was characterized by a very low genetic diversity, similar to the situation in Tasmanian devils. The fingerprint of this low genetic diversity, which has been conserved over millennia, is a low degree of heterozygosity at single nucleotide polymorphisms, indicating that the founder dog was considerably inbred (43).

According to genomic data, spreading of TVT around the world started around 2000 years ago, so at least hundreds of years after its origination. Hence, we can

assume that TVT could evolve over a long time span in the context of a largely inbred population. During this time, critical mutations could emerge that provided tumor cells with the prerequisites to survive under more “hostile” conditions in genetically further remote hosts, thus enabling world-wide propagation of TVT.

Notably, transmissible tumors in Tasmanian devils and canines, despite their ability to jump the barrier between individuals, both up to today have retained intact MHC genes. Instead, they downregulate MHC expression to very low levels. Assuming that there would have been ample time to irreversibly get rid of MHC molecules by random mutation, the retention of MHC further underlines that a complete loss of MHC is disadvantageous to tumor cells as exemplified by *B2M*-mutant MMR-deficient cancer, thus supporting the analogy to Scylla and Charybdis discussed above.

Due to the long history of transmissible tumors that in the case of TVT can be traced back thousands of years, these tumors represent a fascinating treasure chest that we have only started to explore. Current TVT tumors carry the memory of mutational signatures and processes at work during the past millennia. For example, mutations compatible with UV light signatures were found to be strongly associated with the country of sampling, showing an enrichment in areas with high sun exposure such as Mauritius (43). Notably, a previously unknown mutational signature (“signature A”, characterized by CC>TC mutations) was identified that dominated the early evolution of TVT, an observation that may coincide with the accumulation of similar C>T mutations in the human germline postulated for a similar time span (44). Possibly, environmental factors that later vanished left their fingerprints in the human germline and in the mutational profile of TVT. This example demonstrates that mutational signatures and transmissible tumors hold great potential, providing a genomic analog to stratigraphy in archeology.

Some recent speculative publications take the idea of tumor cells abandoning the host and developing a life on their own even further. The SCANDAL (“speciated by cancer development animals”) hypothesis postulates that some simple animals may have originated from tumors that had developed in more complex organisms (45). Conclusive proof for this intriguing idea is outstanding, although it may be in line with observations about the biochemistry and pathophysiology of tumor cells, such as their dramatically increased glucose consumption even under aerobic conditions (Warburg effect (46)) and the sensitivity of some tumor cells to antibiotic substances (47).

In comparison with transmissible tumors, the life span of “conventional”, non-transmissible tumors is limited. This is reflected by the fact that the number of somatic mutations of TVT is more than hundred times higher than the average mutation frequency in human cancer. MMR-deficient tumors represent an exception among non-transmissible tumors, as the extremely high rate of mutations during their “accelerated” evolution compensates for the comparatively short life span. This accelerated evolution has direct implications on the clinical and histopathology presentation of these tumors, which will be discussed below.

10 Heterogeneity of MMR-deficient cancers

Histopathology classification is usually performed to inform clinical decision making and to estimate the prognosis of patients affected by cancer. In colorectal cancer, routine histopathology classification implies determining the differentiation state of the tumor, i.e. their “distance” from normal colon epithelial tissue, and describing the phenotype of cell growth, for example using categories such as “medullar” for tumors growing as a solid, cell-rich mass, or “mucinous” for tumors accumulating extracellular mucus. Typically, tumors diagnosed in pathology are characterized homogeneously by a dominant histopathology differentiation, i.e. they are, following this example, either medullar or mucinous.

However, one characteristic feature of MMR-deficient cancer is the simultaneous presence of several distinct histopathology differentiation patterns within the same tumor (48). This intra-tumoral phenotypic diversity reflects mutational diversity, which itself is a direct consequence of the high mutation likelihood under MMR-deficient conditions Figure 3.

Hence, MMR-deficient tumors nicely illustrate through their morphology one of the most striking features of tumors, which result from their evolution: heterogeneity. Most solid human tumors show mutational heterogeneity (49), although rarely as obvious for the naked eye as MMR-deficient cancers. Heterogeneity is one of the most important obstacles for the success of tumor treatment using targeted drugs. Such drugs in theory can attack tumor cells with a high degree of specificity, sparing normal cells that lack the respective target. However, in practice, disease relapses under targeted therapy are very common, because tumor cells either develop resistance after being exposed to treatment, or, considered more likely for most tumors, subclones of the tumor have been resistant upfront before

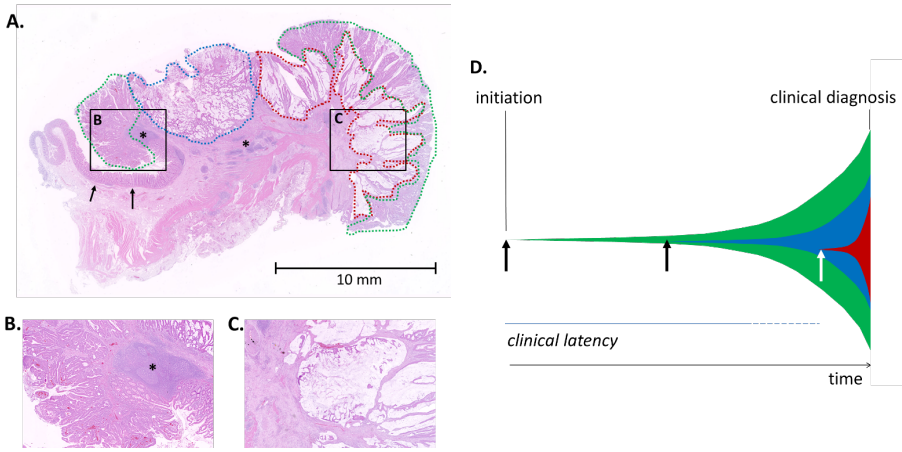


Figure 3: Cancer evolution and histology. **A.** Overview of a formalin-fixed, paraffin-embedded section (2 μ m thickness) through an MMR-deficient colon cancer. The section is stained by hematoxylin (blue) and eosin (pink) to visualize cell nuclei and cytoplasm, respectively. Tumor areas are circled by dashed lines, normal colon mucosa is indicated by arrows. Different colors indicate well-differentiated tumor areas still resembling the normal tissue architecture (*green*, see **B** for higher magnification), and poorly differentiated tumor areas (*blue* – an area with mixed growth patterns encompassing so-called cribriform, i.e. sieve-like, areas; *red* – areas with pronounced extracellular mucus secretion, see **C** for higher magnification). In the center of the section and the transition from normal mucosa to cancer (marked by *asterisks*), signs of a pronounced reaction of the immune system are visible, represented by aggregation of lymphatic cells appearing as small bluish clouds (see also activated lymph follicles in **B**). **D.** Simplified schematic illustration of cancer evolution and its effect on tumor heterogeneity. After an initial transformation event, which enables subsequent steps of tumor evolution, precancerous cell clones may expand slowly, remaining clinically undetectable for a long time. Additional key mutation events (*black and white arrows*) can promote tumor growth by accelerating the proliferative activity of affected cells. Moreover, they can alter their biology and significantly shape the appearance of the manifest tumor in histopathology. In the given example, the clinically manifest tumor after surgical removal will consist of different parts schematically illustrated as *green*, *blue*, and *red*, in analogy to the histopathology section provided in **A**.

treatment. After treatment-induced shrinkage of the tumor mass, such resistant subclones can re-grow to large masses that are then completely therapy-resistant (50).

Compared to targeted therapy, which prolongs survival, but does not cure cancer, immunotherapy, as already reported by Coley, can induce complete and lasting responses in a subset of patients (for exemplary survival curves see (51)). The

direct comparison with oncogene-targeted therapy indicates the benefits of re-activating and supporting the immune system in its fight against cancer cells, using its capacities in tumor control that have evolved and improved in their effectiveness alongside the evolution of the host species.

11 Host factors – towards early steps of tumor formation

Expanding the analogy between the evolution of tumors and the evolution of species, tumors develop within a tissue, which is placed within an individual, similar to the evolution of species in an ecosystem such as an island.

The phenotype of manifest MMR-deficient cancers reflects the influence of the local environment or systemic host factors. MMR-deficient tumors developing in different organs show significant differences of their mutation spectrum. Referring to the examples mentioned above, MMR deficiency-induced mutations inactivating the genes *ACVR2A*, *TGFBR2*, and *B2M*, which are central to the development of MMR-deficient colorectal cancer, are comparatively rare in MMR-deficient endometrial tumors (25).

The reasons for the obvious differences of somatic mutation spectra have not yet been fully figured out. However, growth advantages conferred by somatic mutations depend on the tissue context and the environmental conditions determined by growth factor concentrations, immune cell infiltration and other factors typical of each individual organ site. Other organ-specific differences in the evolutionary process are detectable in the mutational landscape of MMR-deficient tumors: The overall mutation load is higher in MMR-deficient colorectal cancer than in their endometrial counterparts, suggesting that the duration of tumor evolution and the number of cell divisions occurring prior to tumor diagnosis is on average higher in the former.

Reconstructing the time course of tumor evolution by using mutation data as “molecular clocks” (52) provide more precise estimations about the time span of the precancerous evolution, i.e. the phase in which precancerous cells are present, but not yet clinically detectable. For kidney cancer, mathematical modeling based on genome sequencing data revealed the likely presence of hundreds of potential cancer precursor cells in every young adult, although in later life only about 2% of individuals in fact develop cancer, usually with a delay of 20 to 30 years or more (53).

The power of cancer sequence data to inform about the history of an individual cancer sheds light on crucial crossroads, at which key evolutionary milestones were required to further propel tumor growth. From another perspective, we perceive with increasing clarity that cancer development is an exception, and not the cogent result of a sequential progression scenario. This implies that the number of precancerous cell clones that either persist or perish needs to be much higher than the number of manifest tumors. As shown for the example of kidney cancer, it also implies that every human being is host to a multitude of cells with certain genomic changes compatible with our classical concept of cancer. Recent studies using deep sequencing of phenotypically absolutely normal tissue microdissected from different organs including skin, esophagus, endometrium, colorectum and others revealed a surprisingly colorful diversity of mutations, which showed striking resemblance to cancer mutation patterns (54).

Precancerous cell clones also outnumber manifest MMR-deficient cancers by a factor of more than 1000 (55). Here, the observation is visually even more striking, as simple tissue staining techniques used in histopathology can identify potential cancer precursor cells, without the need for sophisticated sequencing data. We will discuss the early pathogenesis of MMR-deficient cancers and the lessons we can learn for cancer prevention in the following.

12 Hereditary cancer and cancer prevention

On some rare occasion, the initial mutation that lays the foundation for later cancer development goes back beyond the birth of an individual. Some individuals inherit cancer-predisposing mutations through the germline and consequently carry them in all body cells. This constellation is called hereditary cancer predisposition.

Hereditary tumor predisposition seems to be responsible for approximately 5 to 10% of the overall tumor burden. The estimation varies depending on the definition; most germline alterations have none or only moderate effects on lifetime cancer risk (polymorphisms or low-risk alleles); others, however, lead to a dramatic increase of the cancer incidence in carriers. The familial adenomatous polyposis (FAP) syndrome is a prominent example for a hereditary cancer syndrome with a tangible and impressive clinical phenotype. Individuals affected by FAP develop hundreds or thousands of precancerous lesions called polyps in the colorectum, and without intervention almost all will develop cancer, often

already in young adulthood. Therefore, preventive surgery with removal of the colorectum is recommended in order to reduce cancer risk. The young age of onset and the severe phenotype is the reason for the relative scarceness of FAP in most populations, as they reduce the likelihood that affected carriers pass on the genetic predisposition to the next generation. This principle leads to an inverse correlation of the frequency and severity of tumor syndromes in the population.

The two most common hereditary cancer predisposition syndromes are hereditary breast and ovarian cancer syndrome (caused by mutations of the *BRCA* genes, see also above) and Lynch syndrome. In Lynch syndrome, named after the visionary pioneer of hereditary cancer research Henry T. Lynch (56), one of four genes coding for proteins (*MLH1*, *MSH2*, *MSH6*, *PMS2*) essential for the functionality of the DNA MMR system is present only as one functional copy or allele. Lynch syndrome mutation carriers have a lifetime risk of developing mismatch repair-deficient cancer between 50 and 80% (57). Current estimations assume that 1/200 to 1/300 individuals carry this genetic predisposition, corresponding to 270,000 to 400,000 people affected in Germany.

The “diploid” nature of our genome is the most powerful factor reducing the likelihood that we develop malignant tumors, as two somatic mutation events are required before a tumor suppressor gene is inactivated (“loss of function”). In Lynch syndrome, one “hit” alone is sufficient for inducing MMR deficiency, because the backup allele is missing in all somatic cells of an affected organism (“two hit hypothesis” formulated by Alfred Knudson (58)). This significantly elevates the lifetime risk for developing MMR-deficient cancer and makes Lynch syndrome a highly valuable condition to test the effectiveness of cancer prevention approaches. Where otherwise huge study populations and observation times would be required to obtain statistically significant results, in Lynch syndrome even moderately effective cancer prevention strategies may show a measurable effect. For example, the cancer-preventive effect of aspirin and other non-steroidal anti-inflammatory drugs could clearly be shown in a limited-size population of Lynch syndrome carriers (59).

Another feature makes Lynch syndrome a prime target for innovative cancer prevention approaches: as outlined above, Darwinian selection during MMR-deficient tumor evolution causes a highly similar landscape of immunogenic antigens (Figures 1 and 2). The rare coincidence of high tumor risk and predictability of tumor antigens provide the setting required to test the feasibility of cancer-preventive

vaccines. We have developed a prototype vaccine consisting of three antigenic peptides occurring in exactly the same form in MMR-deficient cancers; a first clinical trial showed that frameshift peptides can be safely delivered to patients without severe systemic side effects, but consistently inducing immune responses (<https://clinicaltrials.gov/ct2/show/NCT01461148>) (17).

Future clinical trials have to examine the effectiveness of such a vaccination approach for preventing tumors in Lynch syndrome. The innovative approach of cancer-preventive vaccination holds potential also beyond the scenario of Lynch syndrome and hereditary cancer predisposition. In theory, all tumors are associated with a non-infinite number of antigens resulting from genomic alterations. In addition, the amount of genome-wide sequencing data that help to predict these antigens is rapidly growing. Therefore, there is hope that cancer-preventive vaccines may become tangible in the future.

13 Outlook

Modifying Theodosius Dobzhansky's famous dictum one can state that "nothing in *tumor* biology makes sense except in the light of evolution" (60). New technologies providing genome-wide molecular tumor landscapes have helped to overcome traditional classifications of tumors as distinct, unconnected entities. Breaking this "tyranny of the discontinuous mind" (61), that Richard Dawkins blamed as a major obstacle for scientific progress, and instead perceiving tumors as manifestations of an evolutionary continuum are essential pre-requisites to develop improved concepts for tumor therapy and prevention.

However, this paradigm shift, which has been aided significantly by the advent of affordable tools for comprehensive molecular characterization of tumors, per se does not lead to advances in clinical oncology. Making sense of the massive amounts of data about molecular phenotypes of tumors and their evolutionary history is one of the major challenges for cancer research in the next years. The example of MMR-deficient cancer highlights how classification of tumors according to their evolutionary history can provide significant improvements of patients' outcome. Knowledge about "druggable" mutations, i.e. those that give rise to potential therapeutic targets, will help to select promising treatment approaches based on molecular data. Considering the complexity of data and

clinical trial designs required for obtaining high evidence levels, the translation of such approaches into the clinical routine is certainly no small task.

Reconstructing evolutionary trajectories of tumors should also help to develop more effective prevention strategies. If we understand the key events responsible for transforming normal cells into potentially dangerous precancerous cells and at the same time the phenotype and potential “weakness” of these cell populations, we may be able to design new approaches to reduce tumor incidence. Possibly, being able to shift interventions from advanced stage tumor treatment towards elimination of precancerous cell clones is the biggest promise that understanding the principle of tumor evolution holds.

References

1. DeGregori J. Evolved tumor suppression: why are we so good at not getting cancer? *Cancer research* **2011**;71(11):3739-44 doi 10.1158/0008-5472.CAN-11-0342.
2. Darwin C. *On the Origin of Species*. London: John Murray; 1859.
3. Dietel M. Boveri at 100: the life and times of Theodor Boveri. *The Journal of pathology* **2014**;234(2):135-7 doi 10.1002/path.4410.
4. Bauer KH. *Mutationstheorie der Geschwulstentstehung. Übergang von Körperzellen in Geschwulstzellen durch Gen-Änderung*. Berlin: Springer; 1928.
5. McCombs RS, McCombs RP. A Hypothesis on the Causation of Cancer. *Science* **1930**;72(1869):423-4 doi 10.1126/science.72.1869.423.
6. *International journal of epidemiology* **2005**;34(5):1168-70 doi 10.1093/ije/dyi134.
7. Nowell PC. The clonal evolution of tumor cell populations. *Science* **1976**;194(4260):-23-8 doi 10.1126/science.959840.
8. Fearon ER, Vogelstein B. A genetic model for colorectal tumorigenesis. *Cell* **1990**;61(5):759-67 doi 10.1016/0092-8674(90)90186-i.
9. Vogelstein B, Kinzler KW. The Path to Cancer –Three Strikes and You’re Out. *The New England journal of medicine* **2015**;373(20):1895-8 doi 10.1056/NEJMp1508811.
10. Allegra CJ, Rumble RB, Hamilton SR, Mangu PB, Roach N, Hantel A, *et al*. Extended RAS Gene Mutation Testing in Metastatic Colorectal Carcinoma to Predict Response to Anti-Epidermal Growth Factor Receptor Monoclonal Antibody Therapy: American Society of Clinical Oncology Provisional Clinical Opinion Update 2015. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **2016**;34(2):179-85 doi 10.1200/JCO.2015.63.9674.

11. Poggio F, Bruzzone M, Ceppi M, Conte B, Martel S, Maurer C, *et al.* Single-agent PARP inhibitors for the treatment of patients with BRCA-mutated HER2-negative metastatic breast cancer: a systematic review and meta-analysis. *ESMO open* **2018**;3(4):e000361 doi 10.1136/esmoopen-2018-000361.
12. Van Hoeck A, Tjoonk NH, van Boxtel R, Cuppen E. Portrait of a cancer: mutational signature analyses for cancer diagnostics. *BMC cancer* **2019**;19(1):457 doi 10.1186/s12885-019-5677-2.
13. Forbes SA, Beare D, Boutselakis H, Bamford S, Bindal N, Tate J, *et al.* COSMIC: somatic cancer genetics at high-resolution. *Nucleic acids research* **2017**;45(D1):D77-7-D83 doi 10.1093/nar/gkw1121.
14. Maura F, Degasperi A, Nadeu F, Leongamornlert D, Davies H, Moore L, *et al.* A practical guide for mutational signature analysis in hematological malignancies. *Nature communications* **2019**;10(1):2969 doi 10.1038/s41467-019-11037-8.
15. Alexandrov LB, Nik-Zainal S, Wedge DC, Aparicio SA, Behjati S, Biankin AV, *et al.* Signatures of mutational processes in human cancer. *Nature* **2013**;500(7463):415-21 doi 10.1038/nature12477.
16. Davies H, Glodzik D, Morganella S, Yates LR, Staaf J, Zou X, *et al.* HRDetect is a predictor of BRCA1 and BRCA2 deficiency based on mutational signatures. *Nature medicine* **2017**;23(4):517-25 doi 10.1038/nm.4292.
17. Kloor M, von Knebel Doeberitz, M. The immune biology of microsatellite-unstable cancer. *Trends in Cancer* **2016**;2(3):121-31.
18. Ionov Y, Peinado MA, Malkhosyan S, Shibata D, Perucho M. Ubiquitous somatic mutations in simple repeated sequences reveal a new mechanism for colonic carcinogenesis. *Nature* **1993**;363(6429):558-61 doi 10.1038/363558a0.
19. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, *et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *The New England journal of medicine* **2015**;372(26):2509-20 doi 10.1056/NEJMoa1500596.
20. Huang PW, Chang JW. Immune checkpoint inhibitors win the 2018 Nobel Prize. *Biomedical journal* **2019**;42(5):299-306 doi 10.1016/j.bj.2019.09.002.
21. Ishida Y, Agata Y, Shibahara K, Honjo T. Induced expression of PD-1, a novel member of the immunoglobulin gene superfamily, upon programmed cell death. *The EMBO journal* **1992**;11(11):3887-95.
22. Leach DR, Krummel MF, Allison JP. Enhancement of antitumor immunity by CTLA-4 blockade. *Science* **1996**;271(5256):1734-6 doi 10.1126/science.271.5256.1734.

23. Woerner SM, Tosti E, Yuan YP, Kloor M, Bork P, Edelmann W, *et al.* Detection of coding microsatellite frameshift mutations in DNA mismatch repair-deficient mouse intestinal tumors. *Molecular carcinogenesis* **2015**;54(11):1376-86 doi 10.1002/mc.22213.
24. Woerner SM, Benner A, Sutter C, Schiller M, Yuan YP, Keller G, *et al.* Pathogenesis of DNA repair-deficient cancers: a statistical meta-analysis of putative Real Common Target genes. *Oncogene* **2003**;22(15):2226-35 doi 10.1038/sj.onc.1206421.
25. Woerner SM, Yuan YP, Benner A, Korff S, von Knebel Doeberitz M, Bork P. SelTarbase, a database of human mononucleotide-microsatellite mutations and their potential impact to tumorigenesis and immunology. *Nucleic acids research* **2010**;38(Database issue):D682-9 doi 10.1093/nar/gkp839.
26. Jonchere V, Marisa L, Greene M, Virouleau A, Buhard O, Bertrand R, *et al.* Identification of Positively and Negatively Selected Driver Gene Mutations Associated With Colorectal Cancer With Microsatellite Instability. *Cellular and molecular gastroenterology and hepatology* **2018**;6(3):277-300 doi 10.1016/j.jcmgh.2018.06.002.
27. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell* **2011**;144(5):646-74 doi 10.1016/j.cell.2011.02.013.
28. Coley WB. The Treatment of Inoperable Sarcoma by Bacterial Toxins (the Mixed Toxins of the Streptococcus erysipelas and the Bacillus prodigiosus). *Proceedings of the Royal Society of Medicine* **1910**;3(Surg Sect):1-48.
29. Busch W. Aus der Sitzung der medicinischen Section vom 13 November 1867. *Berlin Klin Wochenschr* **1868**;5:137.
30. Fehleisen F. Ueber die Züchtung der Erysipelkokken auf künstlichem Nährboden und ihre Übertragbarkeit auf den Menschen. *Dtsch Med Wochenschr* **1882**;8:553-4.
31. Dolcetti R, Viel A, Doglioni C, Russo A, Guidoboni M, Capozzi E, *et al.* High prevalence of activated intraepithelial cytotoxic T lymphocytes and increased neoplastic cell apoptosis in colorectal carcinomas with microsatellite instability. *The American journal of pathology* **1999**;154(6):1805-13 doi 10.1016/S0002-9440(10)65436-3.
32. Galon J, Costes A, Sanchez-Cabo F, Kirilovsky A, Mlecnik B, Lagorce-Pages C, *et al.* Type, density, and location of immune cells within human colorectal tumors predict clinical outcome. *Science* **2006**;313(5795):1960-4 doi 10.1126/science.1129139.
33. Buckowitz A, Knaebel HP, Benner A, Blaker H, Gebert J, Kienle P, *et al.* Microsatellite instability in colorectal cancer is associated with local lymphocyte infiltration and low frequency of distant metastases. *British journal of cancer* **2005**;92(9):1746-53 doi 10.1038/sj.bjc.6602534.

34. Schwitalle Y, Kloor M, Eiermann S, Linnebacher M, Kienle P, Knaebel HP, *et al.* Immune response against frameshift-induced neopeptides in HNPCC patients and healthy HNPCC mutation carriers. *Gastroenterology* **2008**;134(4):988-97 doi 10.1053/j.gastro.2008.01.015.
35. Hughes AL, Yeager M. Molecular evolution of the vertebrate immune system. *BioEssays : news and reviews in molecular, cellular and developmental biology* **1997**;19(9):777-86 doi 10.1002/bies.950190907.
36. Vesely MD, Schreiber RD. Cancer immunoediting: antigens, mechanisms, and implications to cancer immunotherapy. *Annals of the New York Academy of Sciences* **2013**;1284:1-5 doi 10.1111/nyas.12105.
37. Kloor M, Michel S, Buckowitz B, Ruschoff J, Buttner R, Holinski-Feder E, *et al.* Beta2-microglobulin mutations in microsatellite unstable colorectal tumors. *International journal of cancer Journal international du cancer* **2007**;121(2):454-8 doi 10.1002/ijc.22691.
38. Ozcan M, Janikovits J, von Knebel Doeberitz M, Kloor M. Complex pattern of immune evasion in MSI colorectal cancer. *Oncoimmunology* **2018**;7(7):e1445453 doi 10.1080/2162402X.2018.1445453.
39. Janikovits J, Muller M, Krzykalla J, Korner S, Echterdiek F, Lahrmann B, *et al.* High numbers of PDCD1 (PD-1)-positive T cells and B2M mutations in microsatellite-unstable colorectal cancer. *Oncoimmunology* **2018**;7(2):e1390640 doi 10.1080/2162402X.2017.1390640.
40. Raup DM. Biological extinction in earth history. *Science* **1986**;231:1528-33 doi 10.1126/science.11542058.
41. Morath C, Rohmeiss P, Schwenger V, Waldherr R, Ritz E, Zeier M, *et al.* Transmission of donor-derived small-cell carcinoma cells by a nontumor-bearing allograft. *Transplantation* **2005**;80(4):540-2.
42. Woods GM, Lyons AB, Bettiol SS. A Devil of a Transmissible Cancer. *Tropical medicine and infectious disease* **2020**;5(2) doi 10.3390/tropicalmed5020050.
43. Baez-Ortega A, Gori K, Strakova A, Allen JL, Allum KM, Bansse-Issa L, *et al.* Somatic evolution and global expansion of an ancient transmissible cancer lineage. *Science* **2019**;365(6452) doi 10.1126/science.aau9923.
44. Harris K, Pritchard JK. Rapid evolution of the human mutation spectrum. *eLife* **2017**;6 doi 10.7554/eLife.24284.
45. Panchin AY, Aleoshin VV, Panchin YV. From tumors to species: a SCANDAL hypothesis. *Biology direct* **2019**;14(1):3 doi 10.1186/s13062-019-0233-1.

46. Warburg O, Wind F, Negelein E. The Metabolism of Tumors in the Body. *The Journal of general physiology* **1927**;8(6):519-30 doi 10.1085/jgp.8.6.519.
47. Pestell RG, Rizvanov AA. Antibiotics for cancer therapy. *Oncotarget* 2015;6(5):2587-8 doi 10.18632/oncotarget.3388.
48. Ruschoff J, Roggendorf B, Brasch F, Mathiak M, Aust DE, Plaschke J, *et al.* [Molecular pathology in hereditary colorectal cancer. Recommendations of the Collaborative German Study Group on hereditary colorectal cancer funded by the German Cancer Aid (Deutsche Krebshilfe)]. *Der Pathologe* **2004**;25(3):178-92 doi 10.1007/s00292-003-0641-x.
49. Gerlinger M, Horswell S, Larkin J, Rowan AJ, Salm MP, Varela I, *et al.* Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nature genetics* **2014**;46(3):225-33 doi 10.1038/ng.2891.
50. Geng C, Paganetti H, Grassberger C. Prediction of Treatment Response for Combined Chemo- and Radiation Therapy for Non-Small Cell Lung Cancer Patients Using a Bio-Mathematical Model. *Scientific reports* **2017**;7(1):13542 doi 10.1038/s41598-017-13646-z.
51. Ribas A, Hersey P, Middleton MR, Gogas H, Flaherty KT, Sondak VK, *et al.* New challenges in endpoints for drug development in advanced melanoma. *Clinical cancer research : an official journal of the American Association for Cancer Research* **2012**;18(2):336-41 doi 10.1158/1078-0432.CCR-11-2323.
52. Shibata D, Tavaré S. Counting divisions in a human somatic cell tree: how, what and why? *Cell Cycle* **2006**;5(6):610-4 doi 10.4161/cc.5.6.2570.
53. Mitchell TJ, Turajlic S, Rowan A, Nicol D, Farmery JHR, O'Brien T, *et al.* Timing the Landmark Events in the Evolution of Clear Cell Renal Cell Cancer: TRACERx Renal. *Cell* **2018**;173(3):611-23 e17 doi 10.1016/j.cell.2018.02.020.
54. Gerstung M, Jolly C, Leshchiner I, D'Antonio SC, Gonzalez S, Rosebrock D, *et al.* The evolutionary history of 2,658 cancers. *Nature* **2020**;578(7793):122-8 doi 10.1038/s41586-019-1907-7.
55. Kloor M, Huth C, Voigt AY, Benner A, Schirmacher P, von Knebel Doeberitz M, *et al.* Prevalence of mismatch repair-deficient crypt foci in Lynch syndrome: a pathological study. *The lancet oncology* **2012**;13(6):598-606 doi 10.1016/S1470-2045(12)70109-2.
56. Lynch HT, Snyder CL, Shaw TG, Heinen CD, Hitchins MP. Milestones of Lynch syndrome: 1895-2015. *Nature reviews Cancer* **2015**;15(3):181-94 doi 10.1038/nrc3878.

57. Moller P, Seppala TT, Bernstein I, Holinski-Feder E, Sala P, Gareth Evans D, *et al.* Cancer risk and survival in path_MMR carriers by gene and gender up to 75 years of age: a report from the Prospective Lynch Syndrome Database. *Gut* **2017** doi 10.1136/gutjnl-2017-314057.
58. Knudson AG, Jr. Mutation and cancer: statistical study of retinoblastoma. *Proceedings of the National Academy of Sciences of the United States of America* **1971**;68(4):820-3 doi 10.1073/pnas.68.4.820.
59. Burn J, Gerdes AM, Macrae F, Mecklin JP, Moeslein G, Olschwang S, *et al.* Long-term effect of aspirin on cancer risk in carriers of hereditary colorectal cancer: an analysis from the CAPP2 randomised controlled trial. *Lancet* **2011**;378(9809):2081-7 doi 10.1016/S0140-6736(11)61049-0.
60. Dobzhansky T. Nothing in biology makes sense except in the light of evolution. *American Biology Teacher* **1973**;35(3):125-9.
61. Dawkins R. 2011 Dec 19. Richard Dawkins: The Tyranny of the Discontinuous Mind. In *The New Statesman*. <<https://www.newstatesman.com/blogs/the-staggerers/2011/12/issue-essay-line-dawkins>>. Accessed 2020 April 20.

About the authors

PD Dr. med. Matthias Kloor studied Medicine at the Universities of Heidelberg and Mannheim. Doctoral degree 2002 at Heidelberg University, Habilitation 2012 at Heidelberg University. He is senior scientist and leader of the research group Immune biology of MSI cancer at the Department of Applied Tumor Biology, Institute of Pathology at the Heidelberg University Hospital. His research interests are tumor evolution, tumor immune biology, hereditary cancer, and translational research in oncology. He has been recipient of scientific awards and invited speaker at numerous international conferences on hereditary cancer, tumor immunology, and pathology. He is author and coauthor of more than 150 publications in scientific journals and books.

Magnus von Knebel Doeberitz is Professor of Molecular Oncology and Medical Director of the Department of Applied Tumor Biology, Institute of Pathology of the University of Heidelberg. He is also head of the Clinical Cooperation Unit F210 of the German Cancer Research Center and is affiliated with the Molecular Medicine Partner Unit of the European Molecular Biology Laboratory (EMBL_MMPU). He served as founder and board member of several biotech companies, including

mtm-labs (www.mtm-labs.com) that was acquired by Roche in 2011. Magnus von Knebel Doeberitz spent his post doc period in the laboratory of Nobel laureate Harald zur Hausen at DKFZ. 1995 he was appointed as Professor of Molecular Oncology in the Department of Surgery, University of Heidelberg. 2001 he became Medical Director of the Department of Applied Tumour Biology at the University of Heidelberg. His major scientific interests relate to mechanisms of human carcinogenesis by papillomaviruses and DNA mismatch repair deficiency. He identified novel biomarkers and therapeutic targets and validated them in clinical trials. He co-authored more than 300 peer-reviewed articles. He is Associate Editor of the International Journal of Cancer and member of the Editorial Boards of various other journals in oncology. He received numerous prestigious scientific awards and is inventor of several patent families in his research areas.

Korrespondenz:

PD Dr. med. Matthias Kloor

Abteilung für Angewandte Tumorbologie

Institut für Pathologie

Im Neuenheimer Feld 224

69120 Heidelberg

E-Mail: matthias.kloor@med.uni-heidelberg.de

Von Ereignissen zu Strukturen

HANS J. PIRNER

Institut für Theoretische Physik, Universität Heidelberg

Abstract

A single event is not sufficient, to identify a structure. Only the repetition of similar events can lead to an understanding of the underlying structure. A historical event can never be fully explained by its supposed structure, conversely a structure cannot only be explained by events. In physics, explanation is a necessary condition to understand an event. In this essay, I give examples of structures on various natural scales: In the world around us, we experience matter in solid, liquid or gaseous states, which mirror their structure. At the level of individual atoms, we encounter the slow-down of the thermal dance of the molecules at very low temperature. The elementary particles have intricate family structures that are just as complicated as family structures in anthropology. The universe presents events that can illuminate the long dark night from 1 million years to 1 billion years after the big bang. I show how the researcher inductively derives physical structures from events, but I will also indicate that there are metastructures found by mathematical deduction.

Einleitung

Ein Ereignis ist etwas Singuläres, Einmaliges. Alle Erfahrungswissenschaften beruhen auf dem Studium von Ereignissen, die sie in zweierlei Hinsicht sammeln und ordnen. Vertikal in der Zeit, ist jedes Ereignis eingebettet in einen zeitlichen Ablauf, es ist Teil einer Kette von Ereignissen vorher und nachher. Horizontal in

der Zeit finden parallel zu jedem Ereignis unzählige andere statt, von denen einige die Entwicklung des Ereignisses beeinflussen können.

Die ungestellte Beobachtung der Natur kann nur schwer Zusammenhänge herstellen. Eine Lawine donnert ins Tal und bricht die Ruhe, die danach zurückkehrt. Es hatte viel geschneit und war sehr kalt. Ein einzelnes Ereignis genügt nicht, die Strukturen von Lawinen zu verstehen. Wenn man versucht das singuläre Ereignis zu verstehen, muss man den Prozess studieren, dessen Teil es ist. Nur die Wiederholung ähnlicher oder „gleicher“ Ereignisse hilft das Phänomen „Lawine“ besser zu durchschauen. Im Vergleich kann man dann Schneebrettlawinen mit einem klaren linearen Abrissbereich von Lockerschneelawinen mit einem punktförmigen Anriss und Staublawinen mit viel aufgewirbeltem Schnee unterscheiden. Die Naturwissenschaft erzielte den Durchbruch, als sie anfangs gezielt mit gleichartigen Ereignissen zu experimentieren. Der Experimentator teilt die Mitwelt in den Versuchsaufbau und die Umgebung des Versuchs. Er minimiert den Einfluss der unkontrollierbaren Umgebung auf das Experiment, in dem er den Versuch von ihr isoliert, so gut es geht. Idealerweise kann er allein die Bedingungen des Versuchs bestimmen und die Antwort der so gestellten Natur aufzeichnen. Nachdem er die Daten vieler Ereignisse statistisch aufbereitet hat, kann er sie phänomenologisch analysieren. Oft deutet der Initiator des Versuchs seine Erfahrungen am besten. Er kann sie logisch verbinden, weil er schon mit einer gewissen Vorstellung und Hypothese das Experiment aufgebaut hat. Man spricht dann oft von einem Effekt, den er entdeckt hat.

Rudolf Mößbauer hat 1958 die Resonanzfluoreszenz von Gamma Strahlung in Iridium 191 entdeckt¹ und für diesen „Mößbauereffekt“ drei Jahre später den Nobelpreis erhalten. Bei typischen Kernzerfällen ist die Linienbreite wesentlich kleiner als die Rückstoßenergie des Atomkerns, so dass die Emissions- und Absorptionslinien nicht überlappen. Als Mößbauer seine Probe kühlte, fand er einen großen Anteil von Photonen, bei denen das Kristallgitter den Rückstoß auffing, so dass beide Linien gut überlappten und in Resonanz gingen. Im Allgemeinen muss man die Probe dazu kühlen. In Eisen 57 ist dies schon bei Zimmertemperatur möglich. Durch einen Antrieb kann man das radioaktive Präparat, welches die Gammastrahlung emittiert, mit kleinen Geschwindigkeiten von mm/sec bewegen.

¹ Rudolf L. Mössbauer, Nuclear resonance fluorescence of gamma radiation in ^{191}Ir , Z. Phys. 151 (1958) no. 2, 124–143

Der Dopplereffekt verschiebt die Emissionslinie des bewegten Kerns und tastet damit die Linie im Absorber desselben Kerns ab. Nach dem Durchgang weist man die Intensität der Gamma Linie mit einem Photon-Verstärker nach. Als Quelle wird Co 57 verwendet, das in Fe 57 zerfällt und eine unaufgespaltene Emissionslinie ($\frac{3^-}{2} \rightarrow \frac{1^-}{2}$) aussendet. Wenn man als Absorber natürliches Eisen verwendet, sieht man, dass das Absorptionsspektrum eine *Struktur* aufweist.

Struktur heißt hier eine Regelmäßigkeit, die man verstehen will. Es ergeben sich 6 Linien, bei denen die Zählrate (Counts = Ereignisse) vermindert ist. Diskrete Strukturen sind charakteristisch für die Quantenmechanik, welche die Energieniveaus von Atomen und Kernen quantisiert und folglich auch die Übergänge zwischen verschiedenen Zuständen diskretisiert.

Der magnetische Dipol des Atomkerns will sich parallel zum Magnetfeld der Elektronenhülle im Eisen ausrichten und vermindert dadurch seine Energie. Der Grundzustand $\frac{1^-}{2}$ und der angeregte Zustand $\frac{3^-}{2}$ des Atomkerns spalten durch diese (magnetische Hyperfein-) Wechselwirkung zweifach und vierfach auf. Wegen des Drehimpuls des Photons sind aber nur 6 der $8 = 2 \times 4$ Übergänge möglich, wie man im Experiment sieht (Abb. 1). Die Regelmäßigkeit der Struktur ist somit ein charakteristisches Merkmal der Quantendynamik.

Während in der Geschichte Ereignisse nie gänzlich durch die vorausgesetzten Strukturen erklärbar sind, und man umgekehrt Strukturen nicht nur durch Ereignisse erläutern kann², ist in der Physik die exakte Erklärung eine notwendige Bedingung, um ein Ereignis zu verstehen. Die obige Struktur eines Spektrums ist nur eine von vielen Möglichkeiten wie sich Strukturen in der Physik zeigen können. In diesem Aufsatz will ich Beispiele für Strukturen auf verschiedenen Größenskalen der Natur geben. In der Lebenswelt erfahren wir die Materie in charakteristischen Zuständen, nämlich als fest, flüssig oder gasförmig. Auf dem Niveau der einzelnen Atome begegnen wir dem thermischen Tanz der Moleküle und Atome. Bei den Elementarteilchen treten verzwickte Familienstrukturen auf, die ebenso kompliziert sind wie ethnologische Strukturen. Das Universum präsentiert Ereignisse, welche die lange dunkle Nacht der Astronomen erhellen. Ich werde darstellen, wie der Forscher induktiv aus Ereignissen physikalische Strukturen erschließt, aber es gibt auch Metastrukturen, welche durch mathematische Deduktion gefunden wurden. Sie sind anfänglich nur logische Möglichkeiten, die

² Reinhart Kosellek, Zeitschichten, Frankfurt a/M 2000, S. 329

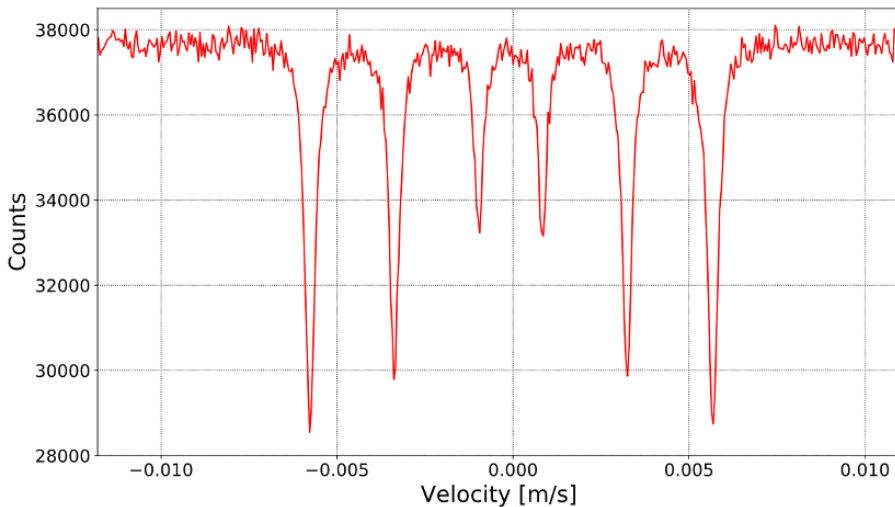


Abbildung 1: Mössbauer Absorptionsspektrum von natürlichem Eisen bei Zimmertemperatur als Funktion der Geschwindigkeit der Quelle in m/s. Die Anzahl der Ereignisse entspricht der Zählrate = Counts. Die Geschwindigkeit 0.001 m/s entspricht einer Energieverschiebung von $4.8 \cdot 10^{-8} \text{ eV}$. Die natürliche Linienbreite ist $4.66 \cdot 10^{-9} \text{ eV}$. Die Abbildung ist von <https://commons.wikimedia.org/wiki/File:M%C3%B6ssbauerSpectrum57Fe.svg> übernommen.

sich jedoch oft später im Experiment bestätigen. Hier findet man z. B. die spezielle und allgemeine Relativitätstheorie, die zu den spektakulärsten Entdeckungen der Physik gehören. Einige andere harren noch ihrem erfolgreichen Nachweis.

1 Feste und flüssige – gasförmige Strukturen

Die Materie besteht aus Atomen, deren Anordnung zu Strukturen führt, die man mit dem bloßen Auge unterscheiden kann. Die Festkörperphysik oder die Physik der kondensierten Materie befasst sich mit Systemen aus extrem vielen Teilchen, die stark miteinander wechselwirken. Diese Systeme können sich in den Phasenzuständen fest, flüssig oder gasförmig anordnen. Der Diamant besteht aus Kohlenstoffatomen und ist ganz verschieden von Graphit, der aus dem gleichen Material aufgebaut ist. Diamant ist sehr hart und hat eine große Brillanz, die ihn zu einem begehrten Schmuckobjekt macht. Graphit ist weich und hinterlässt im Bleistift durch Abreibung eine graue Ablagerung. Diamant kann nur bei sehr

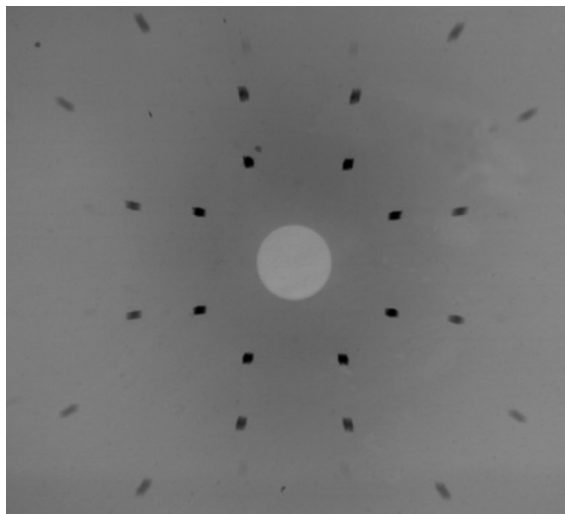


Abbildung 2: „Weiße“ Synchrotronstrahlung generiert diese Laue Aufnahme von Diamant. Die einlaufende Strahlung ist in Richtung einer der vier Symmetrieachsen der Tetraeder Struktur des Diamants ausgerichtet.³

hohen Temperaturen zum Schmelzen gebracht werden ($>3000\text{ °C}$). Welche Struktur macht Diamant so widerstandsfähig? Durch Bestrahlung mit Röntgenlicht findet man wohl definierte (Laue-) Punkte (Abb. 2), die durch die Reflexion an zueinander parallelen Kristallebenen entstehen.

Da das weiße Synchrotron Licht verschiedene Frequenzen enthält, erfüllen mehrere Gitterschichten die Bedingung für konstruktive Überlagerung. Die charakteristische Struktur im Diamant ist eine Tetraeder Struktur, bei der jedes Kohlenstoffatom von vier Nachbaratomen im gleichen Abstand umgeben ist. Man kann sich eine Pyramide vorstellen mit einem C-Atom in der Mitte. Die räumlich kurze Bindung ist die Ursache der extrem harten Struktur des Diamants. Ein besonders spannendes Kapitel der Röntgenstrukturanalyse haben James D. Watson⁴ und Francis Crick geschrieben, als sie mit ihrer Hilfe die Spiralstruktur der DNS entdeckten.

³ Moreton Moore, J. Phys. Cond. Matter, 21 (2009) 364217

⁴ James D. Watson, Die Doppel-Helix, Hamburg 1973

Das weite Meer, der springende Bach, das perlende Wasser aus dem Hahn sind alles Flüssigkeiten. Flüssigkeiten können jede beliebige Form annehmen, je nach dem Behälter, in dem man sie gießt. Sind sie deshalb ohne Struktur? Die molekularen Wechselwirkungen sind in Flüssigkeiten weniger stark ausgeprägt als in Festkörpern. Wenn man versucht Strukturanalyse in Flüssigkeiten zu machen, dann sind die Beugungsbilder verwaschen. Obwohl die Umgebung in der Nähe eines Teilchens noch ziemlich regelmäßig ist, wird sie immer ungeordneter, je weiter man sich vom Teilchen entfernt (siehe Abb. 3). Die Flüssigkeit zeigt, dass die Längenskala auf der man Struktur sucht eine wichtige Rolle spielt. Besondere Probleme treten auf, wenn sich auf verschiedenen Längenskalen die Strukturen ändern.

Wenn man Gase, auch solche mit geladenen Teilchen (Plasmen) mit zu den Flüssigkeiten zählt, dann ist 99 % der sichtbaren Materie in flüssigem Zustand. Ihre Dynamik bildet die Grundlage für viele technische Anwendungen, von der Aerodynamik, Hydraulik, Meteorologie bis zu Verbrennungsprozessen. Bewegte Flüssigkeiten faszinieren durch ihre verschiedenen Erscheinungsformen. Laminarer geordneter Fluss ist charakteristisch für langsam fließende zähe Flüssigkeiten. Turbulenzen setzen ein, wenn die Geschwindigkeit zunimmt. Das Wasser vor dem tiefen Fall mag träge und klar dahinfließen. Nach der Kante am Wasserfall wird es schäumen und zahlreiche Wirbel aufweisen. Wirbel sind eine wichtige Komponente in Turbulenzen. Die hohe kinetische Energie in gewissen Teilen der Flüssigkeit dominiert über die dämpfenden Effekte der niederen Viskosität. Die turbulente, nichtlineare Flüssigkeitsdynamik ist komplex und kann für spezielle Probleme nur in aufwendigen Computerrechnungen simuliert werden.

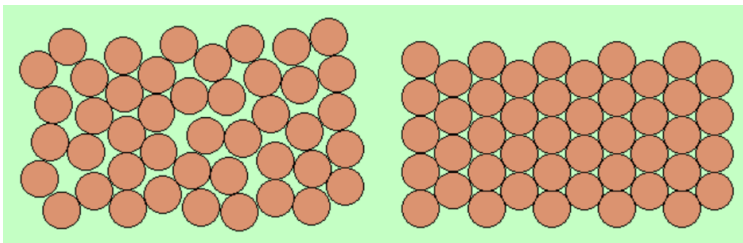


Abbildung 3: Die Nahordnung in Flüssigkeiten (links) im Vergleich zur Kristallstruktur (rechts). https://www.cumschmidt.de/s_aggr_fluessig.htm.

Wenn man bei einem Druck von 1 atm Wasser über 100 °C erhitzt, entsteht Wasserdampf. Wasser ändert seine Struktur, es macht einen Übergang in die Gas Phase. Durch die zugeführte Energie erhöht man die Geschwindigkeit der Moleküle und erlaubt ihnen aus der Flüssigkeit in die gasförmige Phase auszutreten. In einem abgeschlossenen Topf füllt der Dampf das ganze freie Volumen über der Flüssigkeit aus. Wenn der Druck in dem Gas zunimmt, fliegt der Deckel vom Topf. Bei hohen Temperaturen werden die Anziehungskräfte zwischen den Teilchen im Vergleich zu ihrer kinetischen Energie vernachlässigbar. Sie breiten sich aus und zwischen den Teilchen gibt es viel Platz, in dem sich nichts befindet. Oberhalb eines kritischen Punktes mit wohl definiertem Druck und Temperatur ist es unmöglich zwischen Dampf und Flüssigkeit zu unterscheiden.

Die Struktur des Wassers wird durch ein Phasendiagramm beschrieben, das bei ansteigenden Temperaturen zuerst das Schmelzen von Eis in Wasser, und dann das Verdampfen von Wasser in Dampf enthält. Bei niedrigem Druck kann Eis in Dampf sublimieren. Es sind kritischen Ereignisse, wenn eine Struktur in eine andere übergeht. Grenzen trennen die einzelnen Phasen voneinander. Eis kann bei hohen Drucken sogar in acht verschiedenen Phasen existieren.

Helium-4 besitzt bei tiefen Temperaturen eine zweite Flüssigkeitsphase, die mit dem Spin = 0 der He-4 Atome verbunden ist. Das elektronisch gleichartige He-3 (mit nur 3 Nukleonen) und dem Spin = $\frac{1}{2}$ zeigt diese Eigenschaft nicht. Die Bildung des superflüssigen He-4 ähnelt der Kondensation, wenn Flüssigkeitstropfen aus einem Gas gebildet werden. Dieses Kondensat von He-4 wird als Bose Einstein Kondensat bezeichnet, weil es nur bei Quantenteilchen mit ganzzahligem Spin, d. h. Bosonen auftritt. Bosonen können einen gemeinsamen Quantenzustand einnehmen, während Fermionen mit halbzahligem Spin nur jeweils allein einen Quantenzustand besetzen können, wie Fermi und Dirac es theoretisch postuliert haben.

2 Den Tanz der Atome stoppen

Man stelle sich ein Tanzparkett mit Paaren vor, die sich im Rhythmus der Musik über das Parkett bewegen. Wenn die Musik langsamer wird, werden sie ihr Tempo vermindern, um den Takt zu halten. Soweit sehen wir vor unseren Augen eine klassische Szene, wie sie in jedem Film vorkommen kann. Wir würden allerdings sehr überrascht sein, wenn auf einmal in der Mitte der Tanzfläche eine ganze

Gruppe von Tanzenden in der Bewegung erstarrt und sich nicht mehr rührt. Ein solches Verhalten zeigen Atome bei tiefen Temperaturen.

Der Takt der Musik wird durch die Temperatur vorgegeben, die der Experimentator systematisch erniedrigen kann. Dafür gibt es moderne Techniken, wie die Laserkühlung. Gasatome haben bei Zimmertemperatur Geschwindigkeiten von 300 m/s oder 900 km/h wie Flugzeuge. Die Kühlung reduziert ihre Geschwindigkeit auf langsame 1mm/s. Dabei sprechen wir von kälteren Temperaturen als die Temperaturen, bei denen He superflüssig (2 °K) wird, nämlich von Temperaturen im Bereich von Nano Kelvin⁵, d. h. 10^{-9} °K.

Die Quantenmechanik übernimmt das Kommando über den Tanz der Atome, wenn die thermische Wellenlänge λ der Teilchen vergleichbar zum Abstand

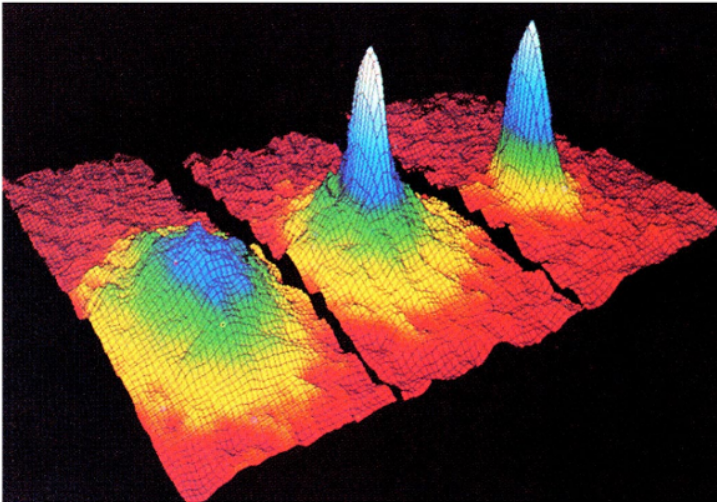


Abbildung 4: Beobachtung eines Bose-Einstein Kondensats in Rubidium durch Cornell und Wiemann. Im ersten Bild, links sieht man die Situation vor der Kondensation, die rot-gelben Atome bewegen sich. Im mittleren und rechten Bild bildet sieht man, wie sich das Kondensat (blau dargestellt) ausbildet und wächst. M. H. Anderson, J. R. Ensher, M. R. Matthews, C. E. Wieman and E. A. Cornell, Science 269, 198 (1995).

⁵ Der Nullpunkt 0 °K der Kelvinskala ist die kleinstmögliche Temperatur. Sie entspricht einer Temperatur von -273.15 °C auf der Celsiuskala.

zwischen den Teilchen wird. Da $\lambda \approx h/\sqrt{2\pi mkT}$, muss die Temperatur umso kleiner sein je größer die Masse m des Atoms ist. 1995 erzeugten Eric A. Cornell und Carl E. Wieman ein reines Kondensat von ca. 2 000 Rubidiumatomen bei 20 nK über dem absoluten Nullpunkt (siehe Abb. 4). Die Atome werden in einem künstlichen Behälter gehalten, der durch elektromagnetische Felder geformt wird, also keine festen Wände hat. Wenn man unterhalb der kritischen Temperatur den Behälter wegnimmt, dann tanzen die angeregten Atome aus der Falle, während die Atome im Kondensat erstarrt sitzen bleibt. Die verbleibenden Atome bewegen sich nicht, weil sie alle den gleichen Quantenzustand mit Impuls = 0 haben. Sie bilden das Bose-Einstein Kondensat (blau in Abb. 4). Wolfgang Ketterle hat ein paar Monate später mittels zwei separater BECs von Natriumatomen gezeigt, dass die beiden Kondensate interferieren können, weil sie jeweils eine quantenmechanische Wellenfunktion besitzen. Wenn die beiden Wellenfunktionen sich auslöschen, gilt die Gleichung „Atom + Atom \rightarrow Nichts“, salopp ausgedrückt. Rührt man das Bose-Einstein Kondensat mit einem Laserstrahl um, bilden sich Wirbel, wie man sie von turbulenten Flüssigkeiten und beim superflüssigen He-4 kennt.

Von Gottfried Wilhelm Leibniz gibt es das Prinzip der Identität des Ununterscheidbaren (*principium identitatis indiscernibilium*): Wenn alles, was von x gilt, auch von y gilt, so ist $x = y$. Die Atome im Bose-Einstein Kondensat befinden sich ununterscheidbar in dem gleichen Quantenzustand, ein Atom x ist ununterscheidbar vom Atom y im Kondensat. Trotzdem sind beide Atome nicht ununterscheidbar gleich, ihre Zahl kann durch Anregung objektiv gemessen werden. In dem Atom Kondensat befinden sich bis zu einer Million Atome. Das Prinzip der Nichtunterscheidbarkeit kann „gerettet“ werden, weil das Bose-Einstein-Kondensat eine gemeinsame Wellenfunktion mit einer gemeinsamen Phase besitzt, insofern ist es eine Einheit und keine Vielheit.

Kalte Quantengase sind extrem dünne Systeme, eine Million Mal dünner als Luft. Während die Atome in Flüssigkeiten und Festkörpern durch ihre Nähe ihre individuellen Eigenschaften verlieren und Bindungen eingehen, sind im Quantengas die Atome von der Größe eines Nanometers weit voneinander getrennt in Abständen von 100 nm. Sie bleiben deshalb gut kontrollierbar und eignen sich als Quantensimulatoren, d. h. Analog-Computer um komplizierte quantenmechanische Rechnungen durchzuführen. Durch raffinierte Lasereinstellungen kann man künstlich eine regelmäßige Anordnung von Fallen herstellen, in denen die Atome nur teilweise gefangen sind, weil sie quantenmechanisch von einer Falle in

die andere tunneln können. Besonders interessant sind Probleme mit Fermionen, weil diese mit konventionellen Monte Carlo Methoden nicht simulierbar sind. Ein bekanntes Modell ist das Hubbard Modell, welches Tunnel Prozesse und eine Coulomb Abstoßung an einem Gitterpunkt einschließt. Dieses Modell beschreibt die Dynamik von Hochtemperatursupraleitern, die sich wie die Kuprate in zweidimensionalen Ebenen anordnen. Damit können sogenannte Quanten Phasen Übergänge studiert werden, die am absoluten Nullpunkt nur auf Grund der Quantenfluktuationen auftreten. Bei den obigen Supraleitern ist es die Veränderung der Dotierung der Elektronen, die den Übergang antreibt.

Ein anderes Problem ist die Zustandsgleichung von stark wechselwirkenden Nucleonen in Kernmaterie. Die Materie in Neutronensternen und kalte Fermi Gase haben ähnliche thermodynamischen Eigenschaften, wenn die Wechselwirkung der Fermionen eine maximale Stärke annimmt. Dadurch wird es mögliche astrophysikalische Phänomene in der kontrollierten Umgebung des Labors mit ultrakalten Atomen zu studieren. Als Funktion der Dichte und Temperatur möchte man die Energie/Teilchen finden.

Man beachte, in beiden obigen Fällen ist die Struktur durch ein theoretisches Modell vorgegeben, das dann mit einem physikalischen Versuchsaufbau simuliert wird. Der Hauptvorteil dieser Methode ist, dass die physikalischen Parameter wie die Spins, die Wechselwirkung, die Potentiale präzise veränderbar sind. Außerdem kann man am physikalischen System einzelne Atome beobachten, d. h. Größen messen, die komplementär zu Beobachtungen in Festkörpern sind. In der klassischen Vielteilchenphysik wird oft von einem idealisierten System mit unendlich vielen Teilchen ausgegangen, mit den kalten Atomen kann man studieren, wie viele einzelne Atome notwendig sind, um ein solches kollektives Verhalten zu sehen.

Im Gegensatz zu induktiven Schlüssen aus dem Studium experimenteller Ereignisse, werden in diesem Zugang die Ereignisse physisch konstruiert. Die physische Rekonstruktion von theoretischen Artefakten unterscheidet sich von der bekannten deduktiven Methode, die mit mathematischen Strukturen (siehe Abschnitt 5) physikalische Probleme löst, wie es z. B. die allgemeine Relativitätstheorie erfolgreich machte. In diesem modernen Zugang zeigt sich eine gewisse Leichtigkeit mit Theorien umzugehen, nämlich kritisch zu hinterfragen, was sie beinhalten. Der Physiker ist nicht mehr damit zufrieden, dass er die „Wirklichkeit“ als Modell

beschreibt, er möchte sich überzeugen, dass die Lösungen des Modells auch im Detail richtig sind.

3 Familienstrukturen der Elementarteilchen

Zu den Anfängen des Strukturalismus gehört die ethnologische Studie von Claude Levi-Strauss über die elementaren Strukturen der Verwandtschaft. Beim Studium primitiver Gesellschaften fand er ein komplexes Schema des Tausches, das die Verheiratung eines Nachkommen regelt. Wenn ein Sohn eine Frau sucht, dann muss der Bruder seiner Mutter ihm eine Tochter als Frau anbieten. Diese Kreuzkusinenheirat soll den Inzest begrenzen und gleichzeitig die gesellschaftliche Ordnung festigen. Der Ethnologe betonte, dass diese Heiratsregeln Teil von „kalten“ Gesellschaften seien, die noch nicht die Komplexität moderner Gesellschaften erreicht haben.

Der Strukturalismus kann als ein Projekt verstanden werden, in dem die Kulturwissenschaften Methoden der Naturwissenschaften nachahmen. Claude Levi-Strauss war nie von einem großen Unterschied zwischen der Natur und der Kultur überzeugt, für ihn waren sie wie das „Rohe und das Gekochte“ nur eine andere Zubereitungsart. Für alle Nichtspezialisten bitte ich also um Geduld, wenn ich jetzt die Strukturen der Verwandtschaften bei den Elementarteilchen darstelle. Sie sind nicht viel komplizierter als die ethnologischen.

Bei den kleinsten Teilchen gibt es bis jetzt drei Generationen. In der bekanntesten Generation befinden sich die Bestandteile der Atomkerne, d. h. die Bestandteile des Neutrons und Protons, die jeweils aus drei Quarks bestehen. Man unterscheidet u-Quarks mit der Ladung $+2/3$ und d-Quarks⁶ mit der Ladung $-1/3$. Ein positiv geladenes Proton hat die Struktur (uud) und ein neutrales Neutron die Struktur (ddu) (siehe Abb. 5). Das Neutron⁷ mit der Masse von 939,5 MeV ist etwas schwerer als das Proton mit der Masse von 938,3 MeV. Die zugehörigen Massen der Quarks spiegeln diese Unterschiede wieder, obwohl die Quarks sehr kleine Massen haben, und die Masse der Nukleonen hauptsächlich durch die Energie der Gluonwolken um die Quarks erzeugt wird. Die up- und down-Quarks haben

⁶ Die u-Quarks werden auch als up-Quarks und die d-Quarks als down-Quarks bezeichnet.

⁷ Wir verwenden Energieeinheiten $1\text{ MeV} = 10^6\text{ eV}$, d. h. wir benutzen die Einstein Formel $E = mc^2$, welche Energie mit der Masse über die Lichtgeschwindigkeit c verbindet (siehe auch Abb. 5).

ungefähr gleiche Massen, dadurch gehören sie zusammen zur gleichen ersten Generation von Quarks. Das freie Neutron lebt eine Zeit von ungefähr 886 s, es zerfällt schwach in ein Proton und ein Elektron und das zugehörige Neutrino.

In der Teilchenphysik sind es zwei mathematische Gruppen, welche die Struktur der schwachen Übergänge in den Generationen und von Generation zu Generation widerspiegeln. Dies sind die Gruppen SU(2) und U(1). Die elektroschwache Wechselwirkung mit der Symmetrie SU(2)×U(1) vereinigt die fundamentale schwache Wechselwirkung und die elektromagnetische Wechselwirkung.

Eine Gruppe G ist ein mathematisches Gebilde, das aus Gruppenelementen {a,b,c,..} und einer Verknüpfung \circ besteht. Es gelten folgende Gesetze:

- (i) Wenn $a, b \in G$ dann ist auch $a \circ b \in G$, d. h. ein Element von G
- (ii) $a \circ (b \circ c) = (a \circ b) \circ c$
- (iii) Es existiert ein Einselement e : $a \circ e = a$ und $e \circ a = a$ für alle $a \in G$
- (iv) Zu jedem a gibt es ein inverses Element $a^{-1} \in G$, so dass $a \circ a^{-1} = e$

Bei den Verwandtschaftsbeziehungen von Levi-Strauss gibt es eine Kleinsche Vierergruppe, welche in kompakter Form die Verhältnisse zusammenfasst. Gruppenstrukturen sind in der Physik immer dann anzutreffen, wenn Symmetrien

| Generation 1 | | | Generation 2 | | | Generation 3 | | |
|--|---------|----------------|---------------|-----------|----------------|--------------|------------|----------------|
| Fermion | Symbol | Weak isospin | Fermion | Symbol | Weak isospin | Fermion | Symbol | Weak isospin |
| Electron neutrino | ν_e | $+\frac{1}{2}$ | Muon neutrino | ν_μ | $+\frac{1}{2}$ | Tau neutrino | ν_τ | $+\frac{1}{2}$ |
| Electron | e^- | $-\frac{1}{2}$ | Muon | μ^- | $-\frac{1}{2}$ | Tau | τ^- | $-\frac{1}{2}$ |
| Up quark | u | $+\frac{1}{2}$ | Charm quark | c | $+\frac{1}{2}$ | Top quark | t | $+\frac{1}{2}$ |
| Down quark | d | $-\frac{1}{2}$ | Strange quark | s | $-\frac{1}{2}$ | Bottom quark | b | $-\frac{1}{2}$ |
| All of the above left-handed (<i>regular</i>) particles have corresponding right-handed <i>anti</i> -particles with equal and opposite weak isospin. | | | | | | | | |
| All right-handed (regular) particles and left-handed antiparticles have weak isospin of 0. | | | | | | | | |

Tabelle 1: Die drei Generationen von Elementarteilchen und der zugehörige schwache Isospin der (linkshändigen) Basiszustände von Quarks und Leptonen (Elektron, Muon, Tau).

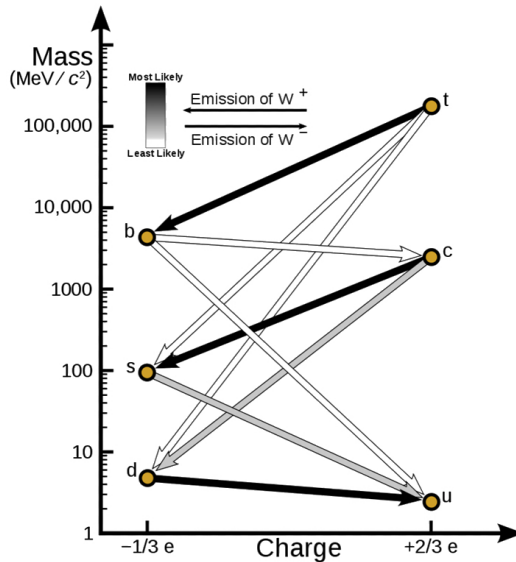


Abbildung 5: Das Bild zeigt die verschiedenen Quarks mit ihren Ladungen auf der horizontalen Achse und ihren Massen auf der vertikalen Achse. Die drei Familien oder Generationen sind (d,u), (s,c) und (b,t), wobei (d,u) für (up,down), (s,c) für (strange,charm) und (b,t) für (bottom,top) Quarks stehen. Die schwachen Zerfälle sind durch Pfeile dargestellt. Die Dicke der Pfeile gibt die Stärke der Übergänge innerhalb und zwischen den Generationen an. [https://en.wikipedia.org/wiki/Weak_interaction#/media/File:Weak_Decay_\(flipped\).svg](https://en.wikipedia.org/wiki/Weak_interaction#/media/File:Weak_Decay_(flipped).svg).

aufzutreten. Die mit der Symmetrie assoziierten Transformationen bilden eine Gruppe. Für das Standardmodell gibt es jeweils zwei Elemente in jeder Generation als Basiszustände der schwachen Isospin (Weak Isospin) Symmetrie $SU(2)$. Diese Verdopplung ist sowohl bei den Quarks als auch bei den Leptonen (Elektron, Muon, Tau und ihren Neutrinos) anzutreffen (siehe Tabelle 1). Zusätzlich besitzen die Teilchen eine schwache Ladung, die mit der Gruppe $U(1)$ verbunden ist.

Beschleuniger können in hochenergetischen Kollisionen Teilchen erzeugen, die es in unserer natürlichen Mitwelt nicht gibt. So findet man in der zweiten Generation z. B. das Λ -Baryon welches aus einem up, down und strange Quark (u,d,s) besteht. Das zugehörige strange Quark s ist schwerer als die Quarks in der ersten Generation und der Übergang ($s \rightarrow u$) ist ungefähr fünfmal schwächer als der Übergang ($d \rightarrow u$) (siehe Abb. 5).

Wenn man einen Übergang von der dritten Generation mit den top und bottom Quarks zu den leichten Quarks betrachtet, dann sieht man, dass der Übergang ($b \rightarrow u$) nur 3 Promille des Übergangs in der ersten Generation beträgt. Universalität besteht jedoch in der Art des Elementareignisses, welches die verschiedenen Materieteilchen miteinander koppelt. Für die schwache geladene Wechselwirkung sind dafür die geladenen W-Bosonen zuständig. Beim schwachen Zerfall des Neutrons (ddu) in ein Proton (duu) wird der Übergang vom down-Quark zum up Quark durch die Emission eines W^- -Bosons erreicht



Am Anfang des Universums bei sehr hohen Temperaturen waren die W-Bosonen, das Photon und das Higgs-Feld alle masselos, die $SU(2) \times U(1)$ Symmetrie war exakt. Erst mit dem Erkalten des Universums ist das Higgs-Feld kondensiert und hat einen endlichen Wert im Vakuum bekommen, durch den die W-Bosonen, das Higgs-Boson und die Materieteilchen eine Masse bekommen haben. Bis jetzt gibt es keine theoretische Erklärung für die Hierarchie der Massen, die sich für die Quarks über 5 Größenordnungen erstreckt. Auch die Größe der schwachen Kopplungen zwischen den Generationen ist nicht erklärt.

Bevor Kepler die Gesetze der Planetenbahnen entdeckt hatte, versuchte er in seinem Buch „Mysterium Cosmographicum“ die Radien der Bahnen zu erklären, indem er sie mit der Oberfläche der platonischen Körper (Würfel, Tetraeder, Oktaeder, Dodekaeder, Ikosaeder) in Beziehung brachte. Neuere Forschungen⁸ deuten darauf hin, dass einige der Planeten wie der Jupiter zuerst in Richtung der Sonne gewandert sind und dann umgekehrt sind, um weiter entfernt, hinter dem Asteroiden Gürtel sich zu stabilisieren. Wie die Geschichte der Verwandtschaften der Elementarteilchen ausgeht ist noch ungewiss, aber gewiss ist es auch eine spannende Geschichte, in der sich die Strukturen und Ereignisse gegenseitig unterstützen müssen.

⁸ Alessandro Morbidelli and Sean N. Raymond, Challenges in planet formation , JGR Planets, Volume 121, 2016, Issue 10, p. 1827–2192, <https://doi.org/10.1002/2016JE005088>

4 Die lange dunkle Nacht des Universums

Ungefähr 400 000 Jahre nach dem Big Bang vereinigen sich die Elektronen und Protonen im heißen Plasma bei einer Temperatur von 3000 °K. Dabei bildet sich neutraler Wasserstoff und Photonen, die uns heute ungestört erreichen, da sie nicht mehr an den Elektronen streuen können. Sie sind von verschiedenen Satelliten sehr präzise gemessen worden. Wegen der Expansion des Universums wird allerdings die ursprüngliche Wellenlänge der Photonen nach längeren Wellenlängen verschoben. Die Ergebnisse zeigen ein ziemlich kaltes Photon Spektrum mit einer Temperatur von 2.7 °K, das in allen Richtungen gleich ist, wenn verschiedene Effekte wie die Bewegung und Lokalisierung unserer Galaxie entfernt werden.

Da diese kosmische Hintergrundstrahlung sich im cm-Bereich befindet, wird sie auch als kosmischer Mikrowellenhintergrund bezeichnet. Das Spektrum entspricht einem thermischen Planck Spektrum von Photonen mit winzigen Temperatur Fluktuationen von $\Delta T/T = 0.00001$. Diese Fluktuation sind im obigen Teil der Abb. 6 farbkodiert wiedergegeben. Zum heutigen Zeitpunkt also $14 \cdot 10^9$ Jahre nach dem Big Bang beobachten wir ein System von Galaxienhaufen, welches im unteren Teil derselben Abb. 6 zu sehen ist. Die hellen Fäden zeigen weiträumige Strukturen von Galaxienhaufen mit Millionen von Galaxien so groß wie unsere Milchstraße, die oft fantasievolle Namen wie die „Große Mauer“ tragen. Dazwischen befinden sich leere Flecken von ebensolcher Größe.

Astrophysiker möchten die Ereignisse finden, welche zwischen diesen beiden Momentaufnahmen unseres Universums liegen. Sie wollen verstehen, wie die großräumigen Strukturen im jetzigen Universum zustande gekommen sind. Wie pflanzen sich die Dichtefluktuationen fort, die sich in den Temperaturfluktuationen der kosmischen Hintergrundstrahlung manifestieren? Wie werden aus ihnen die Galaxien und Sterne, die wir beobachten?

Die Kosmologie benutzt neben der Zeit auch die Rotverschiebung z , um die verschiedenen Zeitalter des Universums zu beschreiben. Der Faktor $(1 + z)$ gibt an um wie viel größer das Universum jetzt ist verglichen mit dem Universum zu jener Zeit. Die Rekombination in der Hintergrundstrahlung mit $z \approx 1200$ besagt, dass das jetzige Universum 1200 Mal größer ist als zur Zeit der Entstehung der Mikrowellenstrahlung.

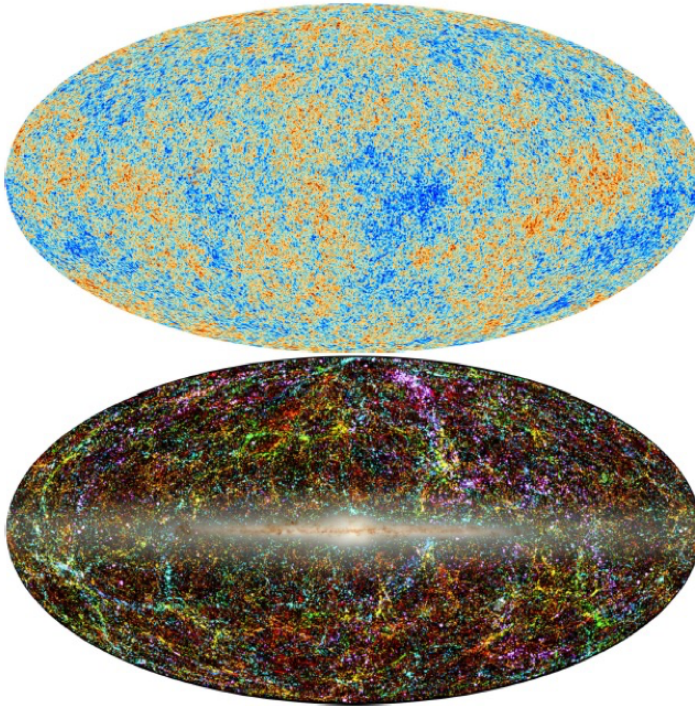


Abbildung 6: Vollständige Abbildung der kosmischen Hintergrundstrahlung durch den Planck Satelliten⁹(oben) und Verteilung der Galaxien im 2-Mass Survey¹⁰(unten). Die Fluktuationen der kosmischen Hintergrundstrahlung im cm-Wellenlängen Bereich sind farbkodiert wiedergegeben, die 2-Mass Untersuchung befindet sich im Infrarot Bereich (Mikrometer). Matthias Bartelmann et al., *Annalen Phys.* 531 (2019) no. 11, 1800446.

Der gleiche Faktor gibt die Verschiebung zu größeren Wellenlängen der beobachteten Strahlung. Mit der Expansion ist die ursprüngliche Strahlung im sichtbaren Licht zu längeren Wellenlängen verschoben worden, wodurch sie für das menschliche Auge unsichtbar wird. Die Astronomen bezeichnen deswegen die Zeit von ungefähr 0.5-1 Million Jahre ($z \approx 1200$) nach dem Big Bang bis zu 1

⁹ Planck- Collaboration, R. Adam et al. *Astronomy & Astrophysics* 594, A9 (2016)

¹⁰ M.E. Skrutskie et al. *Astronomical Journal*, 131, 11632–1183 (2006)

Milliarde Jahre ($z \approx 10$) nach dem Big Bang als „*dark age*“ oder *dunkles Zeitalter des Universums*. Dann nämlich erst formierten sich die ersten Galaxien mit sichtbarem Licht. Da diese Objekte¹¹ sehr weit weg sind, sind sie auch extrem lichtschwach und schwer zu beobachten. Was ist während dieser langen dunklen Zeit passiert?

Im Standard Modell der Kosmologie wird angenommen, dass die ursprünglichen Temperatur- und Dichte-Fluktuationen hauptsächlich durch die dunkle Materie bestimmt sind, die sechsmal häufiger ist als die sichtbare (baryonische) Materie. Es wird angenommen, dass die Gravitation diese Strukturen im Lauf der Zeit verstärkt hat. Wenn man vereinfacht annimmt, dass die Fluktuationen eine Gauss-Verteilung besitzen, dann sind sie nur durch einen Parameter charakterisiert. Man kann dazu die mittlere Fluktuation der Massen $\sigma = \langle \sqrt{(\Delta M/M)^2} \rangle$ verwenden. Kleine Fluktuationen entwickeln sich durch die Gravitation linear, wenn aber die Größe $\sigma > 1$ wird, kollabiert das massereiche System. Fluktuationen bei kleinen Massenskalen sind am stärksten, sie werden sich deshalb als erste von der homogenen Expansion des Weltalls abkoppeln. Am Ende entsteht aus Baryonen und dunkler Materie etwas ungefähr Kugelförmiges, das die Astronomen Halo nennen. Diese Halos wachsen dann an, indem sie mit benachbarten Halos verschmelzen und umliegende Materie aufsammeln. Die treibende Kraft ist dabei die dunkle Materie, die allein der Gravitation unterliegt.

Bei der Entstehung von Galaxien verdichtet sich die baryonische Materie und erhitzt sich dabei, bis der thermische Druck der Schwerkraft die Waage hält. Wenn es dabeigeblichen wäre, wäre die dunkle Nacht nie zu Ende gegangen. Wann oder wie hörte sie auf? Wann sind die ersten Sterne und Galaxien entstanden, die die Verwandlung des Universums vorangetrieben haben. Um den thermischen Druck zu vermindern, muss das heiße Gas abkühlen. Wie wir bei den anderen Betrachtungen gesehen haben, ist für die Entstehung von Strukturen ein kühles Medium wichtig. Wenn Elektronen in diesen Halos aus atomaren Wasserstoff kollidieren, so regen sie den Wasserstoff an oder ionisieren ihn ganz. Diese Epoche wird als Reionisation bezeichnet, weil sie das Plasma aus Protonen und Elektronen wiederherstellt. Die Anregungsenergie wird dann mit großer Wahrscheinlichkeit als Photon abgestrahlt und kann der Gaswolke entkommen und sie kühlen. Ereig-

¹¹ https://en.wikipedia.org/wiki/List_of_the_most_distant_astronomical_objects

nisse¹² im Sloan Digital Sky Survey dokumentieren die Reionisation bei $z \approx 6$. Während man bei $z \approx 6.28$ noch die Absorption der Lyman α - Linie sieht, die vom Übergang des Grundzustands von H in den ersten angeregten Zustand herührt, sieht man bei $z \approx 5.8$ keine solche Absorption. Also muss der neutrale Wasserstoff im intergalaktischen Medium um diese Zeit ionisiert worden sein. Die Details dieser Reionisation sind Gegenstand aktiver Forschungsprojekte, weil sie uns einen Einblick in die Entstehung der Galaxien liefern. Extensive Computersimulationen wie die Illustris TNG Galaxy Formation¹³ Studie dienen dazu diese komplexen Fragestellungen zu untersuchen. Es erscheint wichtig, dass diese Dichtefluktuationen nicht homogen verteilt sind, sondern selbst wieder gehäuft in Gebieten mit Überdichten auftreten. Das Ende der Dunkelheit verlief dann lokalisiert in gewissen weit voneinander liegenden Gebieten, die im Laufe der Zeit anwuchsen. Wie die Reionisation exakt verlief, muss durch weitere Beobachtungen z. B. der Polarisierung der kosmischen Mikrowellenstrahlung und der Beobachtung der 21-cm Hyperfeinstruktur Linie des Wasserstoffs geklärt werden. Das Experiment Edges (Experiment to Detect the Global Epoch of Reionisation Signature) in Australien 2018 hat ein breites Absorptionsband entdeckt. Das Signal stammt möglicherweise vom Licht der ersten Sterne ($z = 17$), welches die lokale Temperatur des Wasserstoffs erniedrigte im Vergleich zur Umgebung. Analytische Rechnungen sind durch die sehr unterschiedlichen räumlichen Skalen und lange Zeitdauer sehr schwierig. Sie sind notwendig, um die numerischen Simulationen besser zu verstehen.

5 Metastrukturen

Die Überschrift drückt aus, dass Strukturen in zwei paradoxe Situationen auftreten können: Einmal kann sich eine Überzahl von Ereignissen schwer auf einfache Strukturen zurückführen lassen (Überbestimmtheit), zum anderen ist es auch möglich, dass eine mathematische Struktur kein physikalisches Ereignis findet (Unterbestimmtheit), das sie als Baustein einer physikalischen Theorie rechtfertigt.

¹² X. Fan et al. *Astronomical Journal* 122, 2833 (2001) und 25, 1649 (2003)

¹³ Simulating Galaxy Formation with the IllustrisTNG Model, Annalisa Pillepich, Volker Springel, Dylan Nelson, Shy Genel, Jill Naiman, Ruediger Pakmor, Lars Hernquist, Paul Torrey, Mark Vogelsberger, Rainer Weinberger, Federico Marinacci, arXiv: 1703.02970

Beide Erscheinungen sind Teile von Metastrukturen, bei denen kein Zusammenhang zwischen Ereignissen und Strukturen hergestellt werden kann.

Lassen Sie mich ein Beispiel für die erste Situation nennen. In einem gemeinsamen Aufsatz¹⁴ haben Physiker und Philosophen untersucht, wie man künstliche Intelligenz benutzen kann, um experimentelle Daten auf Strukturen zurückzuführen. Die Strukturen sollen repräsentativ sein, d. h. sie sollen die notwendigen Schlüsselbegriffe des Systems enthalten. Die Daten des obigen Aufsatzes betreffen zwei Massenpunkte mit den Massen m_1, m_2 und den Ladungen q_1, q_2 , die sich unter dem Einfluss der Schwerkraft und elektrostatischen Wechselwirkung bewegen. Die Daten für verschiedene Kombinationen der Parameter bei gegebenen Anfangswerten werden in verschiedenen Durchgängen ähnlichen neuronalen Netzen (Agenten) vorgelegt. Das gesamte neuronale Netz soll dann herausfinden, dass die Schlüsselbegriffe des Systems die Massen und die Ladungen der Objekte sind. Das neuronale Netz ist ein System von Synapsen, die eingehende Signale mit adjustierbaren Gewichten in ein Ausgangssignal verwandeln. Diese Synapsen sind in verschiedenen Ebenen hintereinandergeschaltet, um dadurch ein adäquates Ausgangssignal zu produzieren. Netzwerke lernen aus Eingangsbeispielen, wie z. B. Bildern von Katzen, Hunden und anderen Haustieren, diese zu kategorisieren. In unserem Beispiel wurde das Netzwerk von verschiedenen Agenten trainiert, die Massen und Ladungen zu identifizieren. Die Erfolgsrate des Netzwerks wurde maximiert, indem die Gewichtung der Synapsen angepasst wurde. Im obigen Beispiel ist es besonders wichtig, dass die Agenten zusammenarbeiten und ihre unterschiedlichen Informationen koordinieren. So mag ein Agent nur Trajektorien von ungeladenen Teilchen, ein anderer einen elastischen Stoß zwischen den Teilchen vorgelegt bekommen. In der Tat finden die Autoren des Projekts, dass es ihren Netzwerken gelang, als richtigen Parameter die Masse m und nicht die Summe von Masse und Ladung als physikalischen Parameter zu identifizieren. Ähnliche Analysen wurden mit stark verrauschten Zeitreihen von neuronalen elektrischen Anregungen (Spikes) gemacht. In der Physik kann man auf diese Weise numerische Simulationen mit Phasenübergängen untersuchen, um die relevanten Ordnungsparameter zu identifizieren.

¹⁴ Hendrik Poulsen Nautrup et al., Operationally meaningful representations of physical systems in neural networks, arXiv:2001.00593 vom 2. Januar 2020

Ich nenne diese Methode „Dekonstruktion“ in Analogie zu dem Begriff, der in den Geisteswissenschaften verwendet wird. Dort bezeichnet er den Versuch eine philosophische Position mit den Begriffen zu hinterfragen, die dem System eigen sind. Die Dekonstruktion entkleidet das System und seine Ereignisse von allen unnötigen Details. Sie ist der umgekehrte Prozess zu der im Kap. 2 beschriebenen Methode, mit Hilfe von kalten Quantengasen ein wohldefiniertes quantenmechanisches Energiefunktional (Hamiltonsfunktion) nachzubauen. Im Allgemeinen muss der Physiker mit seiner Intuition erraten, wie er die Fülle der Informationen in den Daten reduzieren und die relevanten Aspekte des Problems isolieren kann. Ein externes Analyseinstrument hat den Vorteil, die persönlichen Vorlieben des Forschers zu unterdrücken und vorurteilsfrei Strukturen erkennen.

Die Struktur ist Teil einer Zeichensprache, welche in den physikalischen Wissenschaften die Mathematik ist. Im Idealfall erklärt die mathematische Theorie die Daten des Ereignisses. Die minimale Erklärung enthält keine zusätzlichen Begriffe außerhalb der Daten, z. B. keine Informationen über die Temperatur des Objekts, wenn Orte und Geschwindigkeiten des Objekts als Funktion der Zeit gegeben sind.

Eine Triade aus Ereignis, Messdaten und mathematischer Struktur definiert gut die verschiedenen Aspekte im Prozess der Strukturbildung. Ein Ereignis ist der Auslöser für die Betrachtung eines unbekanntes Phänomens. Durch ein Experiment versucht der Physiker die charakteristischen Eigenschaften dieses Ereignisses nachzubilden. Mathematische Konzepte komprimieren die experimentelle Erfahrung in einer theoretischen Struktur (Abb. 7). Diese Dreierheit hat eine gewisse

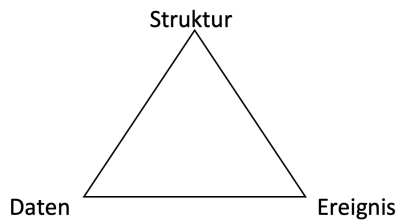


Abbildung 7: Die Triade aus Ereignis, Messdaten und mathematischer Struktur.

Ähnlichkeit mit dem „Zeichen“ von Charles S. Peirce¹⁵, das die Grundlage des Strukturalismus in den Geisteswissenschaften bildet.

Allgemein gilt, dass die Messung der experimentellen Daten der Begründung der Struktur vorausgeht. Aber die Reihenfolge kann auch umgekehrt sein. Ein mathematisches Modell überzeugt besonders, wenn es Daten vorhersagen kann. Mit der Superstringtheorie ist in den letzten 35 Jahren eine vielfältige Theorie entwickelt worden, die bis jetzt wenig mit physikalisch messbaren Ereignissen in Verbindung gebracht werden kann. Hier sind also die Theoretiker auf der Suche nach Ereignissen, die ihre Theorie unterstützen. Der String stellt eine Verallgemeinerung des quantisierten punktförmigen Teilchens dar. Ein Teilchen ist ein „Punkt“, der sich in der Raum-Zeit auf einer Weltlinie bewegt. Der String ist ein eindimensionaler „Faden“, der in der Raum-Zeit eine teppichartige Fläche beschreibt. Nichtkritische Strings beschreiben in der Quantenchromodynamik die starke Attraktion zwischen Quarks und Antiquarks, die durch ein linear anwachsendes Potential daran gehindert werden, sich voneinander zu trennen. Die Quarks sind in den Mesonen gefangen, keine isolierten freien Quarks wurden je gesehen. In den letzten Jahren ist durch die Vermutung von Juan Maldacena die Korrespondenz einer speziellen Stringtheorie mit Anti-de-Sitter-Geometrie und einer Yang-Mills-Theorie postuliert worden, welche der Quantendynamik der Quarks, d. h. hadronischer Elementarteilchen, sehr ähnlich ist.¹⁶ Diese Theorie hat sich nicht induktiv, sondern deduktiv aus rein mathematischen Prinzipien entwickelt und zu einigen schönen Anwendungen geführt.¹⁷ Die eigentliche kritische Superstringtheorie lebt in 10 Dimensionen und zielt darauf ab, die Gravitation mit den Wechselwirkungen des Standardmodells zu vereinigen. Sie enthält zu allen Teilchen des Standardmodells Superpartner, d. h. Partner deren Spin sich um $\frac{1}{2}$ von dem der Teilchen des Standardmodells unterscheidet. Fermionen mit halbzahligen Spin haben supersymmetrische Partner mit ganzzahligem Spin, zu Bosonen mit ganzzahligem Spin existieren supersymmetrische Teilchen mit halbzahligen

¹⁵ Charles Sanders Peirce, Syllabus of Certain Topics of Logic, dt. Phänomen und Logik der Zeichen, hg. Und übersetzt von Helmut Pape, Frankfurt am Main 1983

¹⁶ Juan Martin Maldacena, The Large N limit of Superconformal field theories and supergravity, 1997, Adv.Theor.Math.Phys.2:231-252,1998, Int.J.Theor.Phys.38:1113-1133,1999, e-Print: hep-th/9711200.

¹⁷ E. Megias and H.J. Pirner, QCD thermodynamics using five-dimensional gravity, Phys.Rev. D83 (2011) 056003, arXiv:1009.2953

Spin. Die Teilchenphysiker hofften am Large Hadron Collider in Genf solche Superteilchen zu finden. Unter diesen gäbe es auch solche, welche Kandidaten für die bis jetzt unbekannte dunkle Materie wären. Bis jetzt ist allerdings keiner dieser Superpartner gesehen worden.

Ursprünglich wandten sich die Philosophen dagegen, die Natur mathematisch zu erklären. Besonders Aristoteles argumentierte, dass die Bewegung eine Veränderung der Natur darstellt und deshalb inkompatibel mit den statischen Strukturen der Mathematik sei. Er widersprach damit Platon, der die Elemente Feuer, Luft und Wasser aus den idealen mathematischen Formen des Tetraeders, Oktaeders und Ikosaeders konstruierte. Die Differentialrechnung von Leibniz und Newton machte es jedoch möglich die zeitliche Veränderung der Bewegung mathematisch zu formulieren. Neu in der mathematischen Physik ist, dass man das Problem der Evolution modelliert. Man kann berechnen, wie sich die elementaren Bausteine verändern, wenn man Quantenfluktuationen auf kleineren Skalen summiert. Aus einer Theorie von Quarks auf mikroskopischer Skala, wird auf einer größeren Längenskala eine Theorie von Nukleonen. Simulationen in der Kosmologie oder kondensierten Materiephysik rekonstruieren komplexe zeitliche Entwicklungen. Der Exzellenzcluster STRUKTUREN in Heidelberg untersucht solche Themen, wie Struktur, kollektive Phänomene und Komplexität aus den Grundgesetzen der Physik entstehen. Für Details zu diesem Projekt verweise ich auf den Artikel von Manfred Salmhofer im gleichen Jahrbuch.

Über den Autor

Hans J. Pirner ist theoretischer Physiker. Er hat in USA an der State University of New York at Stony Brook promoviert. Nach Aufhalten am Niels Bohr Institut in Kopenhagen, im CEA Saclay in Paris and am CERN in Genf wurde er 1988 Professor für Physik in Heidelberg mit den Fachgebieten Teilchen-und Kernphysik. Im Marsilius-Kolleg, dem Zentrum für interdisziplinäre Forschung, befasste er sich mit Problemen der Unsicherheit, Unbestimmtheit und Vagheit. Er ist Autor/Coautor von über 200 wissenschaftliche Veröffentlichungen in theoretischer Physik und hat drei Bücher publiziert. Das letzte Buch befasste sich mit der Untersuchung physikalischer Welten im Vergleich zu möglichen Welten. In seinem gegenwärtigen Projekt versucht er zu zeigen, dass Ereignisse die fundamentalen Bausteine der Wirklichkeit sind.

Korrespondenz:

Prof. Dr. Hans J. Pirner

Institut für Theoretische Physik

Universität Heidelberg

Philosophenweg 19

69120 Heidelberg

E-Mail: H.J.Pirner@tphys.uni-heidelberg.de

Homepage: <http://www.tphys.uni-heidelberg.de/~pir/>

Character formation from a psychological point of view: Search for values, search for “Sinn”

JOACHIM FUNKE¹

Department of Psychology, Heidelberg University

Abstract

Character formation has a lot to do with the development of moral values. Classical as well as modern conceptions of moral development will be presented here. Additionally, a world view on values shows that there is a significant convergence on human values in different cultures. Also, the connection between moral values and wisdom is discussed. Measurement approaches to moral values are listed and commented on. Finally, some results from recent “trolley dilemmata” will be presented and evaluated concerning issues of moral judgments.

1 Introduction

Have you ever asked yourself where your values comes from, where your character stems from? The easy answer might be: from your parents and from the culture

¹ This chapter is intended for a book entitled “The impact of education on character formation, ethics, and the communication of values in late modern pluralistic societies,” edited by Stephen Pickard (Canberra, Australia) and Michael Welker (Heidelberg, Germany). I thank the editors for allowing me to pre-publish the chapter within our “Jahrbuch”.

you are living in. However, this answer is a bit superficial and needs a more in-depth look. From a psychological perspective, human development can be seen as a steady process starting from the first moment of conception to the last moment in the hour of death. Developmental psychology analyzes how all human attributes unfold during the stream of life. Part of this psychological perspective is an analysis of the lifespan development of character formation. The special subdiscipline is called “moral development” and deals with questions like the following: How do moral values develop? Where do they come from? How can they be changed if once developed? This “lifespan perspective” of human development has become the standard paradigm in developmental psychology within the last 80 years. It started with the book “Der menschliche Lebenslauf als psychologisches Problem” (engl.: “The human course of life as a psychological problem”) from Charlotte Bühler (1933) and culminated in the book “Life-span developmental psychology” from Baltes, Reese, and Nesselrode (1977).

2 What is character?

Two potential understandings of what constitutes a character will be presented: A narrow one that connects character to moral attitudes and sees character formation as moral development. This view is compared to a broader one that connects character formation to personality development.

The narrow view sees character formation as the construction of a kind of psychological “faculty” (like the faculties of cognition, emotion, motivation, language, etc.) that develops over time, with discrete levels. This view is represented mainly by the ideas of Lawrence Kohlberg, who proposed a famous model of moral “stages” (levels) that are passed through childhood until adolescence. Depending on your course of development, persons end up in one of six postulated levels. Below, I will present his approach in more detail.

A broader view takes a different stance: Character is embedded in a larger view of personality. According to Wrzus and Roberts (2017), “personality constitutes characteristics that reflect relatively enduring patterns of typical cognition, emotion, motivation, and behavior in which individuals differ from others of the same culture or subpopulation.”

Within research on personality, the term “character” sounds old-fashioned and represents a trait approach to personality. Modern views follow an interactionist

view (Endler & Edwards, 1986; Magnusson, 1980): personality is a mixture of (more permanent, stable) *traits* that characterize a person and (more transient, variable) *states* that vary across situations.

3 Classical (narrow) view: Progression through stages

The classical (narrow) view on moral development was established by the Swiss psychologist Jean Piaget (1896–1980) and his American colleague Lawrence Kohlberg (1927–1987). They saw the process of moral development as an ordered progression through developmental stages. During each stage, a specific rule describes the respective word behavior. In the classical view, stages are connected to certain age periods.

To assess the actual stage of moral development, Kohlberg used vignettes like the famous “Heinz dilemma” (Kohlberg, 1981, p. 12):

In Europe, a woman was near death from a very bad disease, a special kind of cancer. There was one drug that the doctors thought might save her. It was a form of radium that a druggist in the same town had recently discovered. The drug was expensive to make, but the druggist was charging ten times what the drug cost him to make. He paid \$200 for the radium and charged \$2,000 for a small dose of the drug. The sick woman’s husband, Heinz, went to everyone he knew to borrow the money, but he could get together only about \$1,000, which was half of what it costs. He told the druggist that his wife was dying and asked him to sell it cheaper or let him pay later. But the druggist said: “No, I discovered the drug and I’m going to make money from it.” Heinz got desperate and broke into the man’s store to steal the drug for his wife.

Should the husband have done that? Was it right or wrong? Is your decision that it is right (or wrong) objectively right, is it morally universal, or is it your personal opinion?

Subjects had to read the vignettes and write down their answer to this question.

Based on the answers to these questions, subjects were classified into one of three levels: Level 1, Pre-Conventional Morality, contains no personal code

of morality. Instead, the moral code is shaped by the standards of adults. The consequences of following or breaking their rules are most important.

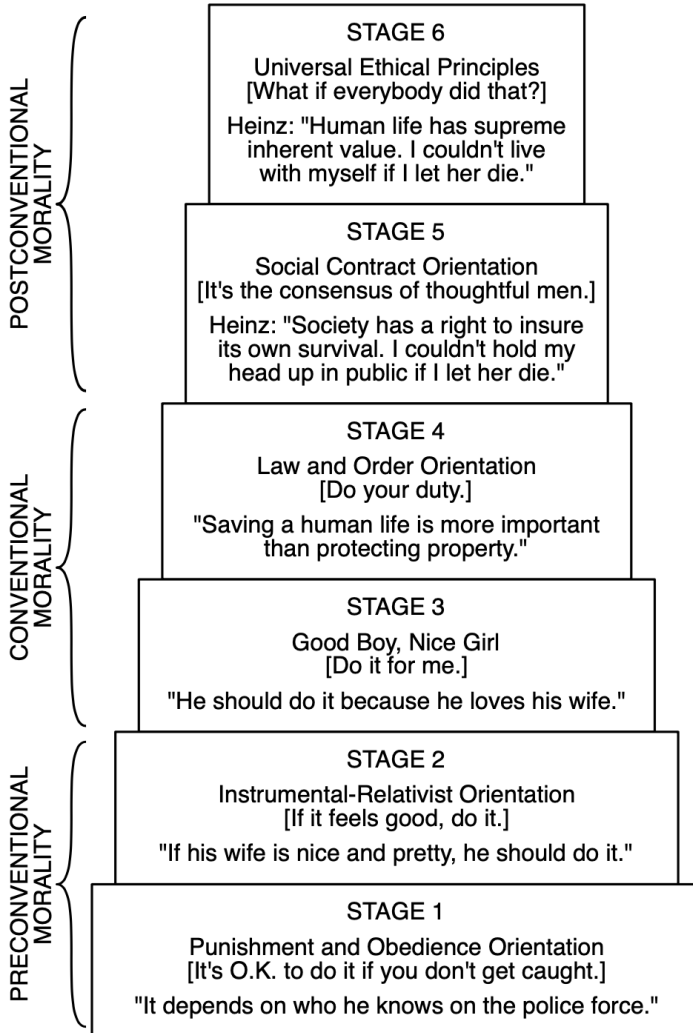


Figure 1: The six stages of moral development, according to Lawrence Kohlberg (Source: http://commons.wikimedia.org/wiki/File:Kohlberg_Model_of_Moral_Development.png).

On Level 2, Conventional Morality, the moral standards of valued adult role models are internalized. Also, authority is internalized but not questioned. Moral reasoning is based on the norms of the group to which the person belongs.

Level 3, Post-Conventional Morality: The individual judgment of moral dilemmata relies on self-chosen principles. Moral reasoning is based on individual rights and justice. According to Kohlberg, this level of moral reasoning is as far as most people get.

Following McLeod (2013), there are some known problems with the Kohlberg approach to the development of moral reasoning. First, dilemmata lack ecological validity (i.e., they are, in a sense, artificial). The “Heinz” dilemma (stealing a drug to save the life of his wife) does not reflect the life experience of 10–16-year old subjects. Second, one might criticize the hypothetical settings: no real consequences will follow the decisions (low-stake instead of high-stake testing, see Sackett et al., 2008). Third, there was biased sampling: according to Gilligan (1977), Kohlberg’s samples were purely male subjects and represented an “androcentric” definition of morality (most important are the principles of law and justice instead of compassion and care). Fourth, a cross-sectional study design was used instead of longitudinal designs, which would better reflect the developmental process. Nevertheless, Kohlberg’s model serves as a reference in many approaches and should be seen as a starting point for improvements.

4 Broader view: Character formation as personality development

To change moral values, we have to change the core of a person, that means: we have to change her or his personality. To change personality is not easy, but it occurs. One simple way is aging: whereas younger people – according to Wrzus and Roberts (2017) – increase their Big Five scores in agreeableness, conscientiousness, emotional stability, and social dominance (Lucas & Donnellan, 2011; Roberts et al., 2006; Roberts & Mroczek, 2008), “older people show a reversed pattern with longitudinal decreases in agreeableness, conscientiousness, emotional stability, and openness” (Berg & Johansson, 2014; Kandler et al., 2015; Lucas & Donnellan, 2011; Möttus et al., 2012). Other opportunities for trait changes often occur together with significant life events (life transitions, changes in personal relationships, work experiences). However, a direct influence on these types of events is not possible. In this broader view of character formation,

education seems to be most important (Engelen et al., 2018). No wonder that in modern times even computer games are seen as training instruments for moral sensitivity (Katsarov et al., 2019).

5 Moral values: World studies

How are moral values distributed on Earth? In an exciting study run by a group of anthropologists, Curry, Mullins, and Whitehouse (2019) checked the prevalence of seven forms of cooperative behaviors (helping kin, helping your group, reciprocating, being brave, deferring to superiors, dividing disputed resources, and respecting prior possession) in 60 societies from all over the world. The background idea was to test their theory of “morality-as-cooperation.” This theory says “that morality consists of a collection of biological and cultural solutions to the problems of cooperation recurrent in human social life” (Curry et al., 2019, p. 48) and is based on assumptions from evolutionary biology and game theory. The theory predicts cooperative behavior in seven domains and postulates that these seven moral values will be universal. To test their predictions, Curry et al. (2019) made a “content analysis of the ethnographic record of 60 societies” distributed over the world (see Figure 2). Data come from the six regions of the globe (Sub-Saharan Africa, Circum-Mediterranean, East Eurasia, Insular Pacific, North America, South America).

They carefully rated data from a huge archive (1200 selected pages from the digital version of the Human Relations Area Files). This archive contains thousands of full-text ethnographies. According to a codebook, the authors selected examples for the following seven areas of morality: (1) family (helping family members), (2) group (helping group members), (3) reciprocity (engaging in reciprocal cooperation), (4) bravery (being brave), (5) respect (respecting your superiors), (6) fairness (sharing or dividing a disputed resource), and (7) property (respecting other’s property). As a result of this analysis, 961 out of 962 relevant text paragraphs valenced these cooperative behaviors positively. The one exception was a negative evaluation that came from the Chuuk, Central America, and concerns property issues: “to steal openly from others is admirable in that it shows a person’s dominance and demonstrates that he is not intimidated by the aggressive powers of others” (Curry et al., 2019, p. 54). The conclusion of the authors (p. 55): “As such, these results provide strong support for the theory of morality-as-cooperation, and



Figure 2: The 60 societies analyzed in the study from Curry, Mullins, and Whitehouse (from Curry et al., 2019, p. 53).

no support for the more extreme versions of moral relativism.” The cross-cultural survey contains a lot of interesting points, but people in western industrialized countries might follow different moral principles than people in foraging societies.

Similarly, Schwartz (1992) explores the universality of value systems by drawing samples from 20 countries, mostly consisting of school teachers and university students, based on the assumption that 11 basic value types can be found universally, all over the world within all cultural regions. These basic (“universal”) values are self-direction, stimulation, hedonism, achievement, power, security, conformity, tradition, spirituality, benevolence, and universalism; see Figure 3. Interestingly, he separates instrumental values (the “means” in a means-end relation) from terminal ones (the “end states”).

His approach comes in the tradition of Hofstede (2001) and does not contain a developmental perspective. Thus, it does not contribute much to the question of character formation.

6 Character formation and the genesis of wisdom

One of the long-term results of character formation can be seen in the development of wisdom. As Barbara Tuchman (1984, p. 21) defined wisdom: “the exercise of judgment acting on experience, common sense and available information.” Is wisdom the result of successful character formation? What do we know about the connection?

In her recent review, Glück (2019, Table 16.1, p. 310) presents twelve definitions of wisdom. Only one of them mentions “values” explicitly, namely the “balance theory of wisdom” from Sternberg (1998). According to that theory, wise people know – besides other competencies – that different people can have different values. This idea of “value relativism” in wise persons is also one of the five criteria for wisdom within the Berlin Wisdom Paradigm (see, e.g., Baltes & Staudinger, 2000). However, to know that there are different perspectives on dilemmata does not imply that one has clear moral values – it is a kind of meta-knowledge, free of any special content.

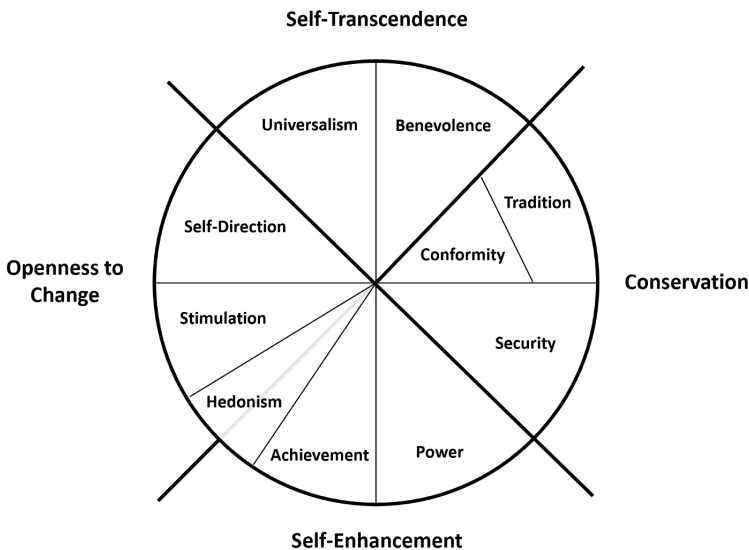


Figure 3: The basic values, according to Schwartz, sorted into four main classes (Source: <https://medium.com/bits-and-behavior/measuring-values-and-culture-264205035c87>).

Similarly, Fischer (2015) argues for a context-free view of wisdom and sees it as “independent of one’s values and context.” On the other hand: Fischer has collected 12 propositions that were commonly known to wise men from four different cultures (Socrates, Jesus, Confucius, Buddha). Those four wise persons show huge parallels concerning certain wise content (e.g., Proposition 10: “Good people (and children) make good company”). Once again, there is no idea about the acquisition of these pieces of wisdom. We all know that reading alone those “wise” propositions will not make us a wise person instantaneously.

7 Measuring character and moral values

Psychologists are known for their expertise to measure dispositions (for a critical position, see Gould, 1996). So, they also developed ideas on how to measure morale and character.

Based on the “narrow” perspective, the “Heinz dilemma” (presented earlier) represents an item from the “Moral Judgements Scale” (MJS) that was developed and used by Kohlberg. It allows subjects to write open answers. A bit more standardized is the “Defining Issues Test” (DIT) that also presents moral dilemmata (like the “Heinz dilemma) but requires a categorical answer instead of free text. As Giammarco (2016) reports, there exist also dilemma-free assessments (“Ethics Position Questionnaire,” Forsyth, 1980; “Visions of Morality Scale,” Shelton & McAdams, 1990) and self-reports (“Moral Foundations Questionnaire”, Graham et al., 2009; “Moral Justification Scale,” Gump et al., 2000; “Measure of Moral Orientation,” Liddell et al., 1992; “Moral Orientation Scale,” Yacker & Weinberg, 1990).

Coming from the broader view (presented above), the measurement of character implies the measurement of personality. The “Big Five” inventories (e.g., BFI; HEXACO-PI-R; NEO-PI-R) measure the following personality attributes (character) via questionnaires: (1) Extraversion (the degree to which one is active, assertive, talkative, etc.), (2) Neuroticism (the degree to which one is anxious, depressed, irritable, etc.), (3) Agreeableness (whether one is generous, gentle, kind, etc.), (4) Conscientiousness (whether one is dutiful, organized, reliable, etc.), and (5) Openness to Experience (whether one is creative, imaginative, introspective, etc.). The current state of affairs concerning this trait approach can be found in a recent review by Costa, McCrae, and (2019).

8 Moral dilemmata in experimental research: Trolley experiments

In recent years, the analysis of moral dilemmata has shown interesting results in moral decision-making. Bostyn et al. (2018, p. 1084) describe the “trolley-style” dilemmata as follows: “In their archetypal formulation, these dilemmas require participants to imagine a runaway trolley train on a deadly collision course with a group of unsuspecting victims. Participants are asked whether they would consider it morally appropriate to save the group but sacrifice a single innocent bystander by pulling a lever to divert the trolley to another track, where it would kill only the single bystander.” For some time, it was an open question whether these hypothetical moral judgments have anything to do with real-life moral decision making. Nevertheless, with the advent of self-driving autonomous cars, these hypothetical situations have become very realistic. Artificial moral agents have already been developed by computer scientists (for a review, see Cervantes et al., 2020)

What are the fundamental insights from experiments with trolley-style dilemmata? It seems that subjects follow a utilitarian perspective to save the most lives possible (see, e.g., Greene et al., 2008). What can we learn about character formation from these studies? Due to the highly artificial situation that has to be imagined (i.e., that is not real), I have doubts about the validity of these studies. I do not believe that we can learn a lot from these highly unnatural settings about human character. It is a bit like insights from Milgram’s famous experiment on obedience: Recent interpretations of the experimental situation argue that the experimental setup (volunteer teacher subjects had to electrically shock bogus learner subjects for errors on a memory task with increasing degrees of shock until deadly volumes) says more about the willingness of the subjects to fulfill the requests of a demanding experimenter (“engaged followership”) than about obedience (Haslam et al., 2016). Likewise, trolley dilemmata might tell us something different than moral decisions.

9 Conclusion

Character formation is a complex process that is not easily accessible to psychological measurement. One thing seems to be sure: “Moral reasoning is developmental” (Killen & Mulvey, 2018, p. 112). Moreover, humans are always searching for Sinn – even in senseless written words, we try to find a message. We see things that are

not present (visual illusion); we hear things that are not spoken (phonological gap); we feel things that are not there (rubber hand illusion); we remember stories that were not told (Frederic Bartlett, schema). We, as human beings, are not robots that require error-free programming and need perfect input; we construct the world around us in such a way that it makes sense to us even if the input is ambiguous. It is argued that humans develop “cargo cults” (Feynman, 1974) when they do not understand the deeper meaning of certain rituals. Humans search for sense, but at the same time, humans search for values: What action is good and should be done more often, what actions are bad and should be reduced in their frequency? It will be an endless story – but one that sharply discriminates humans from machines. The search for values and the search for “Sinn”: it makes us human.

Acknowledgements Thanks to Julia Karl for her help in preparation of this manuscript; thanks to Dr. Marlene Endepohls for carefully reading and commenting my manuscript.

References

- Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life-span developmental psychology: Introduction to research methods*. Lawrence Erlbaum.
- Baltes, P. B., & Staudinger, U. M. (2000). Wisdom: A metaheuristic (pragmatic) to orchestrate mind and virtue toward excellence. *American Psychologist, 55*, 122–136.
- Berg, A. I., & Johansson, B. (2014). Personality change in the oldest-old: Is it a matter of compromised health and functioning? *Journal of Personality, 82*(1), 25–31. <https://doi.org/10.1111/jopy.12030>
- Bostyn, D. H., Sevenhant, S., & Roets, A. (2018). Of mice, men, and trolleys: Hypothetical judgment versus real-life behavior in trolley-style moral dilemmas. *Psychological Science, 29*(7), 1084–1093. <https://doi.org/10.1177/0956797617752640>
- Bühler, C. (1933). *Der menschliche Lebenslauf als psychologisches Problem*. Hirzel.
- Cervantes, J.-A., López, S., Rodríguez, L.-F., Cervantes, S., Cervantes, F., & Ramos, F. (2020). Artificial moral agents: A survey of the current status. *Science and Engineering Ethics, 26*(2), 501–532. <https://doi.org/10.1007/s11948-019-00151-x>
- Costa, P. T., McCrae, R. R., & Löckenhoff, C. E. (2019). Personality across the life span. *Annual Review of Psychology, 70*, 423–448. <https://doi.org/10.1146/annurev-psych-010418-103244>

- Curry, O. S., Mullins, D. A., & Whitehouse, H. (2019). Is it good to cooperate? Testing the theory of morality-as-cooperation in 60 societies. *Current Anthropology*, *60*(1), 47–69. <https://doi.org/10.1086/701478>
- Endler, N. S., & Edwards, J. M. (1986). Interactionism in personality in the twentieth century. *Personality and Individual Differences*, *7*(3), 379–384. <https://doi.org/10.1037/h0036598>
- Engelen, B., Thomas, A., Archer, A., & van de Ven, N. (2018). Exemplars and nudges: Combining two strategies for moral education. *Journal of Moral Education*, *47*(3), 346–365. <https://doi.org/10.1080/03057240.2017.1396966>
- Feynman, R. (1974). Cargo cult science. *Engineering and Science*, *37*(7), 10–13.
- Fischer, A. (2015). Wisdom -The answer to all the questions really worth asking. *International Journal of Humanities and Social Science*, *5*(9), 73–83.
- Forsyth, D. R. (1980). A taxonomy of ethical ideologies. *Journal of Personality and Social Psychology*, *39*(1), 175–184.
- Giammarco, E. A. (2016). The measurement of individual differences in morality. *Personality and Individual Differences*, *88*, 26–34. <https://doi.org/10.1016/j.paid.2015.08.039>
- Gilligan, C. (1977). In a different voice: Women’s conceptions of self and of morality. *Harvard Educational Review*, *47*, 481–517.
- Glück, J. (2019). Wisdom. In R. J. Sternberg & J. Funke (Eds.), *The psychology of human thought* (pp. 307–326). Heidelberg University Publishing. <https://books.ub.uni-heidelberg.de/index.php/heiup/catalog/book/470>
- Gould, S. J. (1996). *The mismeasure of man* (revised and expanded). W. W. Norton.
- Graham, J., Haidt, J., & Nosek, B. A. (2009). Liberals and conservatives rely on different sets of moral foundations. *Journal of Personality and Social Psychology*, *96*(5), 1029–1046. <https://doi.org/10.1037/a0015141>
- Greene, J. D., Morelli, S. A., Lowenberg, K., Nystrom, L. E., & Cohen, J. D. (2008). Cognitive load selectively interferes with utilitarian moral judgment. *Cognition*, *107*(3), 1144–1154. <https://doi.org/10.1016/j.cognition.2007.11.004>
- Gump, L. S., Baker, R. C., & Roll, S. (2000). The moral justification scale: Reliability and validity of a new measure of care and justice orientations. *Adolescence*, *35*, 67–76.
- Haslam, S. A., Reicher, S. D., & Birney, M. E. (2016). Questioning authority: New perspectives on Milgram’s ‘obedience’ research and its implications for intergroup relations. *Current Opinion in Psychology*, *11*, 6–9. <https://doi.org/10.1016/j.copsy.2016.03.007>
- Hofstede, G. (2001). *Culture’s consequences: Comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Sage Publications.

- Kandler, C., Kornadt, A. E., Hagemeyer, B., & Neyer, F. J. (2015). Patterns and sources of personality development in old age. *Journal of Personality and Social Psychology, 109*(1), 175–191. <https://doi.org/10.1037/pspp0000028>
- Katsarov, J., Christen, M., Mauerhofer, R., Schmocker, D., & Tanner, C. (2019). Training moral sensitivity through video games: A review of suitable game mechanisms. *Games and Culture, 14*(4), 344–366. <https://doi.org/10.1177/1555412017719344>
- Killen, M., & Mulvey, K. L. (2018). Challenging a dual-process approach to moral reasoning: Adolescents and adults evaluations of trolley car situations. *Monographs of the Society for Research in Child Development, 83*(3), 110–123. <https://doi.org/10.1111/mono.12380>
- Kohlberg, L. (1981). *Essays on moral development*. Harper & Row.
- Liddell, D. L., Halpin, C., & Halpin, W. G. (1992). The measure of moral orientation: Measuring the ethics of care and justice. *Journal of College Student Development, 33*, 325–330.
- Lucas, R. E., & Donnellan, M. B. (2011). Personality development across the life span: Longitudinal analyses with a national sample from Germany. *Journal of Personality and Social Psychology, 101*(4), 847–861. <https://doi.org/10.1037/a0024298>
- Magnusson, D. (1980). Personality in an interactional paradigm of research. *Zeitschrift Für Differentielle Und Diagnostische Psychologie, 1*(1), 17–34.
- McLeod, S. (2013). Kohlberg's stages of moral development. <https://www.simplypsychology.org/kohlberg.html>
- Möttus, R., Johnson, W., Starr, J. M., & Deary, I. J. (2012). Correlates of personality trait levels and their changes in very old age: The Lothian Birth Cohort 1921. *Journal of Research in Personality, 46*(3), 271–278. <https://doi.org/10.1016/j.jrp.2012.02.004>
- Roberts, B. W., & Mroczek, D. K. (2008). Personality trait change in adulthood. *Current Directions in Psychological Science, 17*, 31–35.
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132*(1), 1–25. <https://doi.org/10.1037/0033-2909.132.1.1>
- Sackett, P. R., Bornemann, M. J., & Connelly, B. S. (2008). High-stakes testing in higher education and employment: Appraising the evidence for validity and fairness. *American Psychologist, 63*, 215–227.
- Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. *Advances in Experimental Social Psychology, 25*, 1–65. [https://doi.org/10.1016/S0065-2601\(08\)60281-6](https://doi.org/10.1016/S0065-2601(08)60281-6)
- Shelton, C. M., & McAdams, D. P. (1990). In search of an everyday morality: The development of a measure. *Adolescence, 25*, 923–943.

- Sternberg, R. J. (1998). A balance theory of wisdom. *Review of General Psychology*, 2(4), 347–365.
- Tuchman, B. W. (1984). *The march of folly. From Troy to Vietnam*. Ballantine Books.
- Wrzus, C., & Roberts, B. W. (2017). Processes of personality development in adulthood: The TESSERA framework. *Personality and Social Psychology Review*, 21(3), 253–277. <https://doi.org/10.1177/1088868316652279>
- Yacker, N., & Weinberg, S. L. (1990). Care and justice moral orientation: A scale for its assessment. *Journal of Personality Assessment*, 55(1–2), 18–27. <https://doi.org/10.1080/00223891.1990.9674043>

About the author

Prof. Dr. Dr. h.c. Joachim Funke studied philosophy, German culture, and psychology at the universities of Düsseldorf, Basel, and Trier. Doctoral degree 1984 at Bonn University, Habilitation 1990 at Bonn University. Honorary doctorate 2015 at Szeged University, Hungary. From 1997 until his retirement in 2019, he was full professor for Cognitive and Experimental Psychology at Heidelberg University. His research interests are thinking and intelligence, creativity and problem solving. He is author and coauthor, editor and coeditor of more than 15 books and has authored and coauthored many journal publications. From 2010 until 2019, he was spokesman of the Academic Senate of Heidelberg University. From 2009 until 2014, he was chairman of the international “PISA Problem Solving Expert Group” from the OECD at Paris.

Correspondence:

Prof. Dr. Dr. h.c. Joachim Funke (ORCID ID: 0000-0001-9129-2659)
University of Heidelberg
Department of Psychology
Hauptstr. 47-51
69117 Heidelberg, Germany
E-Mail: Joachim.funke@psychologie.uni-heidelberg.de
Homepage: <https://funke.uni-hd.de>

Vom Datenchaos zum strukturierten Modell – Wie sich Daten zähmen lassen

CHRISTEL WEISS

Medizinische Fakultät Mannheim der Universität Heidelberg

Zusammenfassung

Dieser Beitrag stellt sich der Herausforderung, die beiden Begriffe „Chaos“ und „Struktur“ zu einem harmonischen Ganzen zu verbinden. Es wird dargelegt, wie aus dem Gedankenchaos im Kopf eines Forschers ein klares Konzept entsteht, wie Chaos bereits bei der Studienplanung vermieden werden kann und wie es mittels statistischer Methoden gelingt, einem unstrukturierten Datenchaos ein strukturiertes Profil zu verleihen. Das Entdecken dieser Strukturen ermöglicht es, Zusammenhänge zu erkennen und neues Wissen zu erlangen. Diese Vorgehensweise wird anhand ausgewählter Therapiestudien aus zwei Jahrhunderten erläutert.

1 Einleitung

Chaos und Struktur: Diese beiden Begriffe wecken höchst unterschiedliche Assoziationen. Das Chaos ist nach der griechischen Mythologie der Urstoff, aus dem die Welt hervorgegangen ist. Ursprünglich bedeutet das altgriechische Wort *Χαος* „gähnende Leere“. Nach alttestamentarischen Überlieferungen erschuf Gott aus dem Chaos die Welt. Im übertragenen Sinne ist der Begriff „Chaos“ negativ konnotiert: Er steht für Unordnung, Ineffizienz oder eine verfahrenere Situation.

Dagegen symbolisiert der Begriff „Struktur“ Eigenschaften wie Ordnung, Klarheit oder Übersichtlichkeit. Er leitet sich ab vom lateinischen „structura“ (Bauart,

Aufbau) und beschreibt die Art, wie die Teile eines komplexen Gefüges miteinander in Beziehung stehen. In zahlreichen Naturwissenschaften ist „Struktur“ zu einem Grundbegriff avanciert: Man denke etwa an Zellstrukturen oder an chemische Strukturen, die den Aufbau einer Zelle, eines Atoms oder Moleküls beschreiben, oder an mathematische Strukturen wie beispielsweise eine Boolesche Algebra, die die Verknüpfungen innerhalb einer Menge definiert. Auch in den Sozial- und Wirtschaftswissenschaften ist „Struktur“ ein Begriff von zentraler Bedeutung. Eine Gemeinschaft von menschlichen Individuen ist durch Sozialstrukturen gekennzeichnet. Ein Staat wird charakterisiert durch seine Bevölkerungsstruktur sowie seine ökonomische, Agrar- und Infrastruktur. In kleineren Sozialgebilden wie Familien, Kollegien und Vereinen regeln Strukturen die Beziehungen der Mitglieder untereinander, deren Funktionen und Verantwortung innerhalb der Gemeinschaft. Auch im Leben eines Individuums sind Strukturen unerlässlich: Niemand ist gerne über einen längeren Zeitraum dem totalen Chaos ausgeliefert. Deshalb ist jeder Mensch bemüht, dem Chaos ein gewisses Maß an Ordnung abzuringen – ansonsten würde er daran zugrunde gehen. Strukturen gewähren Orientierung, sie bieten Sicherheit und Stabilität.

Chaos per se ist ein unbefriedigender Zustand. Andererseits ist er die Voraussetzung dafür, dass eine Struktur entstehen kann [6]. Ein schlichtes Nebeneinander von Elementarteilchen eines Atoms, Organismen einer Zelle oder einzelnen Personen in einem sozialen Verbund scheint nicht sinnvoll zu sein. Erst die Offenlegung der Struktur, also der Art, wie Objekte oder Menschen miteinander agieren, zeigt die Komplexität der Beziehungsgeflechte auf und ermöglicht tiefer gehende Erkenntnisse.

Wie lässt sich eine Struktur im Chaos aufspüren? Um Ordnungsstrukturen zu erkennen, müssen Regelmäßigkeiten gefunden werden. Dies gilt für die hohe Wissenschaft ebenso wie für die Niederungen des Alltags. Die Entwirrung eines Chaos beginnt mit der Beschreibung der einzelnen Elemente. Durch die Entwicklung der Elektronenmikroskopie wurde es möglich, die Bestandteile einer Zelle (Membran, Plasma, Organellen u. a.) zu erkennen und zu benennen. Zu Beginn des 20. Jahrhunderts entdeckten Wissenschaftler wie Ernest Rutherford (1871–1937) oder Niels Bohr (1885–1962), dass ein Atom aus einem Kern und Elektronen besteht. Soziale Wesen lassen sich durch Merkmale wie Alter, Bildung oder sozio-ökonomischer Status beschreiben. Das Chaos auf einem Schreibtisch ist nichts

weiter als eine Ansammlung von Büchern, Zetteln, Büroutensilien, wichtigen Dokumenten, unerledigten Aufgaben etc.

Freilich erklären einfache Beschreibungen nicht die Funktionsweise der einzelnen Elemente und deren Interaktionen untereinander. Gleichwohl ist dies ein erster und wichtiger Schritt, um Licht ins Chaos zu bringen. Die Beschreibung regt die Kreativität an und verleitet dazu, zu ordnen und zu sortieren, Fragen zu stellen, zu handeln und zu forschen. Welchem Zweck dienen die Gegenstände auf einem Schreibtisch? Welche Beziehungen verbinden die Mitglieder einer Gesellschaft untereinander? Wie sind die Elemente in einem Atom oder einer Zelle angeordnet? Welche Funktion haben sie? In nachfolgenden Schritten wird man sich bemühen, Antworten zu finden, Strukturen im Chaos zu entdecken und darin verborgene Geheimnisse zu ergründen.

Was haben diese Überlegungen mit Studien, Daten und statistischen Analysen zu tun? In empirischen Wissenschaften wie beispielsweise der Medizin oder der Psychologie ist die Statistik als wissenschaftsübergreifende Disziplin unerlässlich, um Studien zu planen und durchzuführen, Daten zu analysieren und die Ergebnisse zu interpretieren. Chaos macht sich zuweilen breit: Noch vor Studienbeginn im Kopf eines Wissenschaftlers (etwa eines Kliniklers), der seine Gedanken ordnen muss, um eine konkrete Hypothese zu formulieren und eine Studie planen zu können. Später wird der Statistiker mit einem Chaos konfrontiert sein, denn Daten im Urzustand stellen im wahrsten Sinne des Wortes ein Chaos dar. Das Studienziel besteht im Wesentlichen darin, das Chaos aufzulösen, um eine Antwort auf die der Studie zugrundeliegende Fragestellung zu finden und so zu neuen Erkenntnissen zu gelangen.

2 Klinische Studien: Der lange Weg des therapeutischen Fortschritts

2.1 Vom Gedankenchaos zum klaren Konzept

Bei einer empirischen Studie sind in aller Regel Wissenschaftler aus mehreren Fachrichtungen beteiligt. Neben einem Vertreter aus dem Wissenschaftsgebiet, in dem die Studie durchgeführt wird (etwa eines Kliniklers), bedarf es eines Statistikers, der bei der Studienplanung, der Datenanalyse sowie der Interpretation der Ergebnisse behilflich ist. Dies sei am Beispiel einer Therapiestudie erläutert (wenngleich andere Studientypen wie Risiko-, Diagnose- oder Präventionsstudien grundsätzlich ähnliche Überlegungen erfordern).

Was steht am Beginn einer solchen Studie? Ein Forscher entwickelt eine Idee für ein neues therapeutisches Konzept – sei es aufgrund der Kenntnis der Krankheitsmechanismen oder einer zufälligen Entdeckung, aufgrund einer spontanen Idee oder nach langjährigen Erfahrungen. Er möchte diese neue Therapie überprüfen und – falls sie sich als tauglich erweisen sollte – dafür Sorge tragen, dass sie später bei der Patientenversorgung Anwendung findet. Ein neues Arzneimittel, ein neu entwickeltes chirurgisches Verfahren oder die Optimierung einer bereits zugelassenen Therapie soll letztlich den Patienten zugutekommen. Dieser Nutzen ist im Rahmen einer Studie nachzuweisen. Doch wie soll der Nutzen gemessen werden? Eine Antwort auf diese Frage ist keineswegs trivial. Es ist anzunehmen, dass dem Initiator einer Studie ein Wirrwarr von Ideen durch den Kopf huscht, ehe er einen konkreten Plan erarbeitet. Tatsächlich sind die Wirkungen einer Therapie, die Vorgänge und Wechselwirkungen im menschlichen Körper so komplex, dass sie sich einer genauen Beschreibung entziehen. Deshalb fällt es zunächst schwer, eine präzise Fragestellung zu formulieren. Chaos ist also quasi vorprogrammiert. Um dieses Gedankenchaos aufzulösen, sollte der Forscher strukturiert vorgehen und folgende Fragen klären:

- **Gegenstand der Forschung:** Zunächst ist die Forschungsthematik inhaltlich einzugrenzen. Handelt es sich bei der zu evaluierenden Therapie um ein Medikament, einen chirurgischen Eingriff, ein Medizingerät, eine Psychotherapie oder eine andere Therapieart? Wird ein neues Produkt getestet oder ist die zu prüfende Therapie bereits auf dem Markt?
- **Stand der Forschung:** Der Studienverantwortliche sollte rekapitulieren, welche Voruntersuchungen bezüglich des Forschungsvorhabens durchgeführt wurden und welche Erkenntnisse vorliegen. Sodann ist zu klären: Welche Phase des Entwicklungsprozesses steht an? Zudem sollte mittels einer umfassenden Literaturrecherche der aktuelle Stand der Wissenschaft bezüglich dieser Thematik eruiert werden.
- **Studienziel:** Welches Ziel steht im Vordergrund? Die Beantwortung dieser Frage hängt wesentlich von der Entwicklungsphase der Therapie ab. Bei der Entwicklung eines Medikaments müssen diverse Phasen durchlaufen werden, ehe die Zulassung beantragt werden kann [3]. In der vorklinischen Phase wird das Arzneimittel mittels Labor- oder Tierexperimenten bezüglich Wirkmechanismus und Toxikologie untersucht. In Phase I wird es an

gesunden Probanden getestet, um Informationen bezüglich Pharmakodynamik (Wirkung des Arzneimittels im Organismus) und Pharmakokinetik (Verteilung und Verstoffwechslung im Körper) zu gewinnen. In Phase II wird das Medikament erstmals Patienten verabreicht, um die therapeutische Wirksamkeit zu prüfen und unerwünschte Nebenwirkungen zu erfassen. In Phase III wird das neue Medikament mit einer anderen Therapie verglichen, um das Wirkungsvermögen unter Studienbedingungen zu prüfen („Efficacy“). Wenn all diese Phasen erfolgreich abgeschlossen sind, kann die Zulassung beantragt werden. Danach beginnt die Phase IV, in der die Compliance („Einnahme-Treue“) unter Alltagsbedingungen („Effectiveness“), seltene Nebenwirkungen und die Bewährung des Medikaments bei speziellen Patientengruppen erforscht sowie Untersuchungen zu Therapieoptimierung durchgeführt werden. – Soweit in aller Kürze der Prozess bei der Entwicklung eines Arzneimittels, der im Schnitt insgesamt – von der Idee bis zur Marktreife – mehr als zehn Jahre andauert. Bei anderen Therapieformen wie beispielsweise Medizinprodukten (dazu zählen Implantate, Instrumente, Diagnostika) sind die Vorgaben weniger streng. Aber auch sie sind klinisch zu bewerten, ehe sie routinemäßig angewandt werden können.

- **Zielgrößen:** Die Wahl relevanter Zielgrößen ergibt sich aus dem Studienziel und hängt wie dieses von der Entwicklungsphase des Produkts und der zu therapierenden Erkrankung ab. Studienärzte sprechen von „Endpunkten“ oder verwenden den englischen Begriff „Outcomes“. In der vorklinischen Phase und in der Phase I werden Endpunkte verwendet, die sich mit molekularen oder physiologischen Termini beschreiben lassen. In der Phase II werden häufig Surrogatendpunkte erhoben, um die Wirkung der Therapie zu messen [2]. Diese Ersatzgrößen sind hilfreich, da sie meist schon nach kurzer Zeit gemessen werden können, während sich der eigentlich angestrebte Nutzen einer Behandlung oft erst nach Jahren oder Jahrzehnten erweist. So stellen zum Beispiel Biomarker (z. B. Blutdruck- oder Cholesterinwerte) Surrogatparameter dar bei Therapieoptionen, die auf eine Verlängerung der Lebenszeit oder Vermeidung von Herzinfarkten zielen. Dagegen bevorzugt man in den Phasen III und IV klinische Endpunkte: Das sind Symptome oder Zustände, die ein Patient unmittelbar an sich wahrnehmen kann (etwa das Auftreten eines Symptoms oder die Linderung von Schmerzen) oder die Zeitdauer, die vergeht, bis ein bestimmtes Ereignis eintritt. Man unter-

scheidet zwischen „harten“ Endpunkten, die auf objektiven Fakten basieren und präzise gemessen oder erfasst werden können (etwa Laborwerte, das Auftreten einer Nebenwirkung oder eines Rezidivs), und „weichen“ Endpunkten, die weniger gut objektivierbar sind (dazu zählen beispielsweise die Merkmale „Schmerzintensivität“ oder „Lebensqualität“) [3].

In aller Regel kommen bei einer Studie mehrere Endpunkte in Betracht. So wird bei Krebserkrankungen bevorzugt die progressionsfreie Zeit oder die Zeit des Überlebens als „harter“ Endpunkt gewählt, wenngleich auch der „weiche“ Endpunkt „Lebensqualität“ relevant erscheint – insbesondere dann, wenn die Lebenserwartung der Patienten nur noch wenige Monate beträgt. In jedem Fall sollte der Studienleiter einen primären Endpunkt festlegen – das ist der Parameter, der das wichtigste Ziel der Studie am besten beschreibt. Gleichwohl verhindert diese Festlegung nicht, dass weitere, sogenannte sekundäre Endpunkte in Augenschein genommen werden.

- **Einflussgrößen:** Die wichtigste Einflussgröße bei Therapiestudien ist die Art der Therapie. Darüber hinaus bietet es sich an, weitere potenzielle Einflussparameter zu erfassen, die eventuell dazu beitragen, die primäre Zielgröße zu erklären und Unterschiede zwischen Subgruppen zu erkennen (z. B. das Alter der Patienten, Risikofaktoren oder Komorbiditäten). Insbesondere bei nicht randomisierten Studien sollte man darauf achten, Störgrößen zu erheben (sogenannte Confounder), die das Ergebnis einer Studie verzerren und zu unzulässigen Schlussfolgerungen verleiten könnten, um deren Effekte später herausrechnen zu können.
- **Studiendesign:** Studien der präklinischen Phase, der Phasen I und II werden meist als prospektive Studien ohne Vergleichsgruppe durchgeführt. Dagegen werden in Studien der Phase III (zuweilen schon in der Phase II) Patienten, die die neue Therapie erhalten, mit einer Gruppe von Patienten verglichen, die anders therapiert werden (z. B. mit der bisher verwendeten Standardtherapie, „treatment as usual“, TAU). Ein solcher Vergleich ist nicht nur bei Arzneimittelstudien vorgeschrieben, sondern erscheint auch bei anderen Behandlungsoptionen wie bei Medizinprodukten geboten. Um faire Vergleiche zu ermöglichen, sollten die Gruppen strukturgleich sein. Dies wird am ehesten durch Randomisierung, also durch die zufällige Zuteilung der Studienteilnehmer zu einer Therapiegruppe, erreicht. Deshalb werden

Phase-III-Studien im Sinne der Evidenzbasierten Medizin vorzugsweise als „Randomisierte Studie“ (RCT = randomized clinical trial) durchgeführt. Weitere Kriterien stellen die Behandlungs- und die Beobachtungsgleichheit dar: Die Patienten beider Gruppen sollten in gleicher Weise behandelt (mit Ausnahme der Therapien) und beobachtet werden. Um dies zu gewährleisten, eignen sich doppelblinde Studien, bei denen weder der Patient noch der behandelnde Arzt im Einzelfall die spezielle Therapie kennt. Allerdings sind diese Qualitätskriterien nicht in jedem Fall praktisch umsetzbar (sei es aus ethischen oder aus organisatorischen Gründen). Als Alternative zu RCTs bieten sich prospektive Beobachtungsstudien an, bei denen die Zuteilung der Patienten zu einer Therapie nach anderen Kriterien erfolgt (beispielsweise durch den behandelnden Arzt oder aufgrund der Wünsche des Patienten). Studien in der Phase IV lassen sich als randomisierte Studien oder als Anwendungsbeobachtungen durchführen. Letztere sind nicht-interventionelle Studien, die dazu dienen, Informationen über bereits zugelassene Therapien zu erheben. Auch retrospektive Studien sind denkbar, bei denen die Daten beispielsweise aus Patientenakten oder einem Register rekrutiert werden.

- **Studienpopulation:** Bei Studien der Phasen II und III werden die teilnehmenden Patienten im Allgemeinen durch vorab festgelegte Ein- und Ausschlusskriterien definiert. Während die Einschlusskriterien definieren, welche Patienten von der zu testenden Therapie profitieren könnten, beschreiben die Ausschlusskriterien, welche Patienten von der Studienteilnahme auszuschließen sind (etwa weil sie Risikofaktoren aufweisen). Je strenger die Auswahlkriterien, desto homogener ist die Studienpopulation und desto besser lässt sich der therapeutische Effekt nachweisen. In jedem Fall sollte der Studienleiter darüber nachdenken, wie viele Probanden oder Patienten in seiner Klinik oder seinem Institut in absehbarer Zeit rekrutierbar sind, oder ob eventuell eine multizentrische Studie in Erwägung gezogen werden muss. Bei Studien der Phase IV gelten keine einschränkenden Auswahlkriterien. Bei retrospektiven Studien ist die Studienpopulation durch die gegebenen Daten bereits vorgegeben. In diesen Fällen muss der Forscher evaluieren, ob diese Population für seine Fragestellung geeignet ist.

Diese Darlegungen zeigen: Bei der Planung einer Studie sind mannigfaltige Überlegungen anzustellen und weitreichende Entscheidungen zu treffen, die auch

nachfolgende Entwicklungsphasen betreffen. Eine strukturierte Vorgehensweise zahlt sich hier aus. Sie hilft, die Gedanken zu ordnen und Klarheit über das eigene Forschungsvorhaben zu gewinnen. Das Forschungsvorhaben beginnt, Struktur anzunehmen.

2.2 Chaosvorbeugung durch Kompromisse

Zur konkreten Studienplanung sollte ein Statistiker herangezogen werden. Des- sen Know-How kann sehr viel dazu beitragen, das Datenchaos in Grenzen zu halten (noch ehe die Daten erhoben werden). Dies betrifft vor allem das Studien- design und den Stichprobenumfang. Bei einer prospektiven Studie sind folgende Überlegungen anzustellen:

- **Subgruppen:** Wird nur eine Gruppe untersucht (etwa bei Phase-I-Studien) oder werden mehrere Subgruppen verglichen (wie bei Phase-III-Studien)? Weisen diese Subgruppen etwa die gleiche Gruppengröße auf? Falls nein: In welchem Verhältnis stehen die Gruppengrößen?
- **Primärer Endpunkt:** Welche Eigenschaften hat dieser Endpunkt: stetig, binär, ordinal skaliert oder quantitativ abgestuft? Wie sind die Daten verteilt? Steht eine zuverlässige und validierte Messmethodik zur Verfügung?
- **Fragestellung:** Welche Vergleiche bezüglich des primären Endpunkts sollen durchgeführt werden (etwa zwischen zwei Gruppen oder Vorher-Nachher-Vergleiche)? Soll ein Überlegenheitsnachweis erbracht werden (zum Bei- spiel, dass eine neue Therapie bezüglich ihrer Wirkung einem Placebo oder der bisher verwendeten Standardtherapie überlegen ist), oder soll ein Un- terschied nachgewiesen werden? In manchen Situationen wird nicht die Überlegenheit angestrebt: Es genügt zu zeigen, dass die neue Therapie der Vergleichstherapie nicht unterlegen ist (wenn sie dafür in anderen Bereichen Vorteile aufweist, etwa weniger Nebenwirkungen oder geringere Kosten). Bei der Untersuchung von Generika ist der Nachweis der Äquivalenz zweier Therapien gefragt. Daran entscheidet sich, welcher statistische Test geeignet und welche Fallzahl erstrebenswert ist.
- **Fallzahlplanung:** Dies ist ein bedeutender Aspekt der Planungsphase, da die erforderliche Fallzahl darüber entscheidet, wie lange die Rekrutierungs- phase dauern wird und wann mit ersten Ergebnissen zu rechnen ist. Dieser

Schritt stellt eine große Herausforderung an den Statistiker dar: Die Fallzahl sollte einerseits ausreichend hoch sein, um einen relevanten Unterschied nachweisen zu können. Andererseits sollte sie nicht zu hoch sein, da dadurch Patienten unnötigerweise belastet würden und sich die Studiendauer in die Länge ziehen würde. Beide Szenarien wären aus ethischer Sicht problematisch. Um eine adäquate Fallzahl zu ermitteln, muss der Statistiker wissen, welche Ergebnisse erwartet werden und ab welcher Größenordnung ein Effekt klinisch relevant ist.

- **Poweranalyse:** Die Situation stellt sich anders dar, wenn im Rahmen von retrospektiven Studien auf bereits vorhandenes Datenmaterial zugegriffen wird. Anstelle einer Fallzahlplanung ist eine Poweranalyse erforderlich, um herauszufinden, ob und mit welchen Einschränkungen die der Studie zugrunde liegende Fragestellung beantwortbar ist. Außerdem sollte der Statistiker die Qualität des Datenmaterials dahingehend sichten, ob die Daten korrekt und vollständig erfasst sind.
- **Bias und Confounder:** Bei einem Bias handelt es sich um einen systematischen Fehler, der das Ergebnis verzerren und zu unzulässigen Schlussfolgerungen führen kann [4]. Ein Selektionsbias entsteht, wenn die Gruppe der Studienteilnehmer nicht repräsentativ für die Zielpopulation ist. Dieser Bias ist zu befürchten, wenn nur spezielle Patienten an einer Studie teilnehmen (beispielsweise aufgrund strenger Ausschlusskriterien) oder wenn viele Patienten während der laufenden Studie ausscheiden und nicht mehr für die Endanalyse zur Verfügung stehen. In diesen Fällen wären die Ergebnisse der Studie nicht verallgemeinerbar. Ein Selektionsbias liegt auch vor, wenn zwei zu vergleichende Gruppen a priori nicht strukturgleich sind. Ein Informationsbias entsteht durch nicht korrekte, lückenhafte oder ungenaue Datenerhebung oder -erfassung. Dieser Bias kann durch Mess- oder Dokumentationsfehler bedingt sein; er kann auch durch fehlerhaft ausgefüllte Fragebögen, durch falsche Angaben bei Interviews oder durch mangelhaftes Erinnerungsvermögen von Probanden entstehen. Insbesondere retrospektive Studien, bei denen die interessierenden Ergebnisse bereits stattgefunden haben, sind anfällig für Informationsbias. – Generell ist ein Bias bei keinem Studiendesign auszuschließen. Deshalb sollte man Vorsorgemaßnahmen treffen, um dessen Einfluss zu minimieren. Bei nichtrandomisierten Studien

ist darauf zu achten, dass mögliche Confounder erfasst werden, um faire Gruppenvergleiche zu ermöglichen.

Es liegt also nicht zuletzt im Verantwortungsbereich des Statistikers, dafür Sorge zu tragen, dass Chaos bei der Datenerfassung, der Analyse und der Interpretation der Ergebnisse vermieden wird.

Bei der Studienplanung sind zuweilen Kompromisse einzugehen. Die Randomisierung gilt zwar als Goldstandard bei Therapiestudien, weil damit ein Selektionsbias vermieden wird. Dieses Design lässt sich jedoch nicht erzwingen (wenn etwa Patienten Vorbehalte äußern, an einer solchen Studie teilzunehmen). Doppelverblindung wird wegen der Beobachtungsgleichheit empfohlen; dieses Design ist jedoch nicht in jedem Fall praktikabel. Eine hohe Fallzahl ermöglicht präzise Schätzungen; andererseits bereitet diese Forderung in der klinischen Forschung mitunter erhebliche Probleme (insbesondere bei selten auftretenden Krankheiten). Als primärer Endpunkt eignet sich unter statistischen Aspekten eine stetige, normalverteilte Größe; diese ist jedoch nicht unbedingt klinisch relevant. Harte Endpunkte lassen sich präzise messen oder objektiv erfassen; andererseits sind weiche Endpunkte wie beispielsweise die Lebensqualität oft aussagekräftiger. Biomarker und andere Surrogatparameter lassen sich bereits kurz nach Therapiebeginn messen; sie stellen jedoch nur Ersatzgrößen für die eigentlich interessierenden Studienziele dar (z. B. Überlebenszeit) [2]. Homogene Patientengruppen lassen sich zwar effizienter analysieren als heterogene Gruppen. Allerdings gelten die so erhaltenen Ergebnisse nur eingeschränkt; sie sind nicht ohne Weiteres auf die Zielpopulation übertragbar. Die Studienpopulation sollte repräsentativ für die Zielpopulation sein; jedoch unterliegt die Studienteilnahme einschränkenden Bedingungen. Aus Gründen der Sicherheit mag es sinnvoll sein, Risikopatienten (z. B. ältere Menschen mit Komorbiditäten) von der Studienteilnahme auszuschließen; jedoch sind es häufig genau diese Patienten, die von der Therapie profitieren würden. Eine lang andauernde Studie hat den Vorteil, dass sich Langzeitverläufe beobachten lassen. Andererseits stellt dies hohe Anforderungen an die Patienten (z. B. häufige Wahrnehmung von Untersuchungsterminen), und es vergeht entsprechend viel Zeit, ehe valide Ergebnisse vorliegen. Die zu vergleichenden Subgruppen mögen hinreichend groß und strukturgleich sein. Es ist jedoch keineswegs gewährleistet, dass dieser Zustand so lange aufrechterhalten werden kann, bis die relevanten Endpunkte erhoben werden können.

Diese Widersprüche werden sich in den seltensten Fällen zu jedermanns Zufriedenheit auflösen lassen. Es hängt vor allem von der Phase des Entwicklungsprozesses und vom konkreten Studienziel ab, welche Einschränkungen akzeptabel erscheinen. Die jeweiligen Entscheidungen sollten vom Fachvertreter und vom Statistiker gemeinsam im Hinblick auf das eigentliche Studienziel und das Wohl der Patienten getroffen werden. Jeder an einer Studie beteiligte Wissenschaftler sollte das Seine dazu beitragen, dass die Studie gut geplant wird, und dass aus Chaos Ordnung entsteht.

2.3 Vom Datenchaos zum statistischen Modell

Der Statistiker sieht sich häufig mit einer Excel-Tabelle konfrontiert, die einen schier unübersehbaren Datenwust ohne erkennbare Zusammenhänge enthält. Diese liefert (zumindest auf den ersten Blick) keinen hohen Erkenntnisgewinn. Eine strukturierte Vorgehensweise erweist sich als nützlich, um schrittweise das Dunkel zu lichten.

- **Tabelleninhalt:** Eine Beschreibung des Tabelleninhalts trägt bereits zur Veranschaulichung bei: Die Zeilen und Spalten einer Tabelle repräsentieren die Studienteilnehmer bzw. die erfassten Merkmale; einzelne Tabellenblätter könnten sich auf aufeinanderfolgende Untersuchungszeitpunkte beziehen. Ein kritischer Blick auf die Excel-Tabelle lässt erahnen, ob die Daten möglichst vollständig oder eher lückenhaft erfasst sind, ob gravierende Fehler erkennbar sind und ob unnütze Daten, die für die Analyse nicht relevant sind, enthalten sind. Außerdem sollte geprüft werden, ob alle relevanten Ziel- und Einflussgrößen erfasst sind.
- **Plausibilitätsprüfung:** Zur statistischen Analyse empfiehlt sich die Verwendung einer validierten Statistiksoftware (z. B. R, SAS oder SPSS). Um eine hohe Datenqualität zu gewährleisten, sollten nach dem Datenimport Plausibilitätsprüfungen durchgeführt werden. Ein Blick auf die Häufigkeitsverteilung eines jeden Merkmals zeigt, ob die aufgelisteten Werte oder Klassifikationen realistisch sind. Grobe Fehler bei der Dateneingabe werden dabei offensichtlich.
- **Ausreißer:** Dies sind extrem hohe oder extrem niedrige Werte eines quantitativen Merkmals, bei denen fraglich ist, ob sie unter denselben Bedingun-

gen wie die anderen Werte der Datenreihe erhoben wurden. Durch einen Blick auf das Minimum oder Maximum sind Ausreißer leicht auffindbar. Die Einstufung eines Wertes als Ausreißer ist in erster Linie inhaltlich motiviert. Zunächst ist – am besten mit einem klinischen Fachvertreter – zu klären, ob es sich um einen falschen Eintrag (aufgrund eines Mess-, Dokumentations- oder Rechenfehlers), um eine pathologische Besonderheit oder um eine physiologische Rarität handelt. Fehlerhafte Einträge, die sich nicht korrigieren lassen, sind freilich von der Analyse auszuschließen. Wenn sachlogische Gründe dies erfordern (etwa bei Non-Konformität mit den Ein- und Ausschlusskriterien), müssen gar die dazugehörigen Datensätze komplett eliminiert werden. Ansonsten sind Ausreißer bei der statistischen Analyse durch die Wahl geeigneter Methoden und bei der Interpretation der Ergebnisse in angemessener Weise zu berücksichtigen.

- **Deskriptive Statistik:** Mit diesen Methoden werden charakteristische Eigenschaften eines einzelnen Merkmals beschrieben. Wenn zwei Gruppen verglichen werden, sollten beide Subgruppen separat beschrieben werden. Bei qualitativen Merkmalen werden absolute und relative Häufigkeiten ermittelt; für quantitative Variablen lassen sich Lagemaße, Streuungsmaße und Formmaße berechnen. Mit einem Assoziationsmaß (z. B. Korrelationskoeffizient oder *Odds Ratio*) lässt sich die Stärke des Zusammenhangs zwischen zwei Merkmalen quantifizieren. Auch graphische Darstellungen (Balkendiagramme, Histogramme oder *Box-and-Whisker-Plots*) eignen sich, um Besonderheiten einer Datenreihe aufzuspüren. Das geübte Auge eines Statistikers wird wertvolle Informationen aus diesen Darstellungen extrahieren und erkennen, inwieweit diese Ergebnisse vielversprechend in Hinblick auf das Studienziel sind und worauf er in den folgenden Analyseschritten zu achten hat. Das anfängliche Chaos hat sich zu diesem Zeitpunkt bereits erkennbar reduziert.
- **Gruppenvergleiche:** Der Vergleich ist ein wichtiges Element in der klinischen und der epidemiologischen Forschung: Bei Therapiestudien werden meist zwei Therapieformen, bei Fall-Kontroll-Studien erkrankte Patienten mit gesunden Kontrollen verglichen. Dichotomisierungen sind allseits beliebt: Sie machen das Leben einfach und bringen Struktur ins Chaos. Der Statistiker wird sich bemühen, vorhandene (oder nicht vorhandene) Unter-

schiede zwischen den Gruppen mittels statistischer Tests zu belegen. So kann er beurteilen, ob die beiden Gruppen zu Studienbeginn strukturgleich sind und ob sich die Gruppen bezüglich des Endpunkts unterscheiden.

- **Interpretation der Ergebnisse:** Der Statistiker ist vor die Aufgabe gestellt, aus der Vielzahl von statistischen Testverfahren jeweils einen adäquaten Test auszuwählen und diese Wahl zu begründen. Das Ergebnis eines jeden Tests ist der sogenannte p-Wert. Dieser Wert quantifiziert den sogenannten α -Fehler (auch Fehler 1. Art genannt). Dies ist Wahrscheinlichkeit dafür, dass das erhaltene Testergebnis allein aufgrund des Zufalls zustande gekommen ist. Gemeinhin bezeichnet man ein Testergebnis mit einem p-Wert, der geringer ist als 0,05, als „statistisch signifikant“. Zusätzlich zum p-Wert sollte eine Effektgröße angegeben werden (etwa die Differenz zwischen zwei Mittelwerten oder eine *Odds Ratio*), am besten zusammen mit einem Konfidenzintervall. Nur mit diesen Angaben lässt sich beurteilen, ob ein statistisch signifikantes Ergebnis klinisch relevant ist und wie verlässlich die Schätzung ist. Ein nicht-signifikantes Ergebnis bedeutet keineswegs, dass definitiv kein Unterschied zwischen zwei Gruppen vorhanden ist. Es könnte auch aufgrund eines zu geringen Stichprobenumfangs zustande kommen (dies wäre ein β -Fehler oder Fehler 2. Art). Auch in diesem Fall sind der p-Wert und ein Konfidenzintervall nützlich, um das Testergebnis angemessen zu interpretieren.
- **Analyse des Endpunkts:** Meist werden bei einer Therapiestudie neben der Therapieform weitere potenzielle Einflussgrößen erhoben. Es ist nützlich zu evaluieren, ob und in welchem Maße jede dieser Einflussgrößen mit dem Endpunkt assoziiert ist. So lässt sich einschätzen, welche Einflussgrößen wichtig oder weniger wichtig sind, und bei welchen am ehesten Kausalitäten zu vermuten sind. Häufig wird sich zeigen, dass eine Zielgröße nicht nur mit einer, sondern mit mehreren Einflussgrößen statistisch assoziiert ist. Dann lohnt es sich, eine multiple Analyse durchzuführen, bei der simultan mehrere Einflussgrößen evaluiert werden [11]. Ein multiples Verfahren weist gegenüber einfachen statistischen Tests mehrere Vorteile auf: Individuelle Charakteristika der Studienteilnehmer können im Rahmen eines multiplen Modells angemessen berücksichtigt werden. Bei nicht randomi-

sierten Studien helfen multiple Methoden, nach Confoundern zu adjustieren und Verzerrungen zu vermeiden [4,11].

Die Konstruktion eines adäquaten statistischen Modells verlangt sehr viel Know-How und detailliertes Fachwissen. Anders als man es bei Lösungen von mathematischen Problemstellungen gemeinhin erwartet, ist das finale Modell keineswegs eindeutig bestimmt. Auch bei einer gut geplanten Studie ergeben sich an vielen Stellen der statistischen Analyse Fragen, die zu Kompromissen zwingen. Das beginnt bei der Wahl der Testverfahren: Einerseits werden Tests mit einer hohen Power bevorzugt; diese stellen jedoch Anforderungen an die Daten, die mitunter nicht gegeben oder nicht überprüfbar sind. Andere Fragen betreffen den Umgang mit fehlenden Werten: Sollte man diese durch Imputationsverfahren schätzen (und damit Ungenauigkeiten in Kauf nehmen), oder sollte man die jeweiligen Datensätze bei der Analyse nicht berücksichtigen (was einen Verlust an Power mit sich bringt)? Wie handhabt man Ausreißer? Welche Variablen sollten in ein multiples Modell aufgenommen werden [11]? Was besagen die Ergebnisse der statistischen Analyse bezüglich Kausalitäten? Besonders bei dieser Frage sollte man sehr vorsichtig argumentieren, um nicht zu irreführenden Schlussfolgerungen zu gelangen.

Nichtsdestotrotz bleibt festzuhalten: Leistungsstarke Rechner und effiziente Analysemethoden ermöglichen es, basierend auf den Rohdaten eine Regressionsgleichung zu erstellen, die die interessierende Zielgröße (z. B. die Wirkung einer Therapie oder die Wahrscheinlichkeit für das Eintreten eines Ereignisses) in Abhängigkeit von mehreren relevanten Einflussfaktoren in präziser Formelsprache beschreibt. Damit lassen sich Phänomene erklären, Zusammenhänge erkennen, Kausalitäten vermuten und Prognosen wagen. Eine solche Gleichung ist faszinierend! Sie versinnbildlicht: Aus dem Chaos ist ein strukturiertes Modell entstanden, das die Wirklichkeit zwar nicht vollständig, aber zumindest teilweise abbildet. Mehr noch: Das unstrukturierte Gedankenchaos im Hirn eines Forschers und das Chaos der Rohdaten haben sich aufgelöst und Geheimnisse preisgegeben. Daraus sind neue Erkenntnisse hervorgegangen.

3 Anwendungsbeispiele

3.1 Historische Beispiele

Die in Abschnitt 2 dargelegten Empfehlungen, die dazu dienen, ein Gedankenchaos zu ordnen und ein Datenchaos zu strukturieren, sind in der klinischen und epidemiologischen Forschung noch nicht allzu lange etabliert. Heute herrscht zwar allgemeiner Konsens darüber, dass statistische Analysen für die medizinische Forschung unverzichtbar sind. Jedoch hatten Ärzte und Patienten bis weit ins 20. Jahrhundert hinein massive Vorbehalte bezüglich Studien in der klinischen Forschung. Erst im Jahre 1964 wurden auf der 18. Generalversammlung des Weltärztebundes in Helsinki Richtlinien für die Forschung am Menschen als „Deklaration von Helsinki“ erarbeitet. Diese Deklaration wurde seither mehrfach revidiert und neuen Erkenntnissen angepasst; die letzte Revision wurde im Jahre 2013 in Fortaleza (Brasilien) verabschiedet. Die technischen Voraussetzungen, die mathematischen Verfahren und adäquate Softwareprodukte, die erforderlich sind, um große Datenmengen effizient zu bearbeiten und aus einem Datenchaos ein statistisches Modell zu generieren, entstanden erst mit dem Aufkommen leistungsfähiger Computer und benutzerfreundlicher Software in den 1980er Jahren.

Dennoch haben auch in früheren Zeiten, lange vor der Deklaration von Helsinki und dem EDV-Zeitalter, Wissenschaftler versucht, Daten zu generieren und auszuwerten, um Antworten auf Fragen zu finden, die sie allein mit ihrem Fachwissen nicht beantworten konnten. Als Beispiel seien die Forschungsarbeiten des Gynäkologen *Ignaz Philipp Semmelweis* (1818–1865) erwähnt. Semmelweis arbeitete um die Mitte des 19. Jahrhunderts in einer geburtshilflichen Abteilung des Wiener Allgemeinen Krankenhauses. In dieser Abteilung betreuten Ärzte und Medizinstudenten die Frauen bei der Entbindung, während in einer zweiten Abteilung ausschließlich Hebammen als Geburtshelferinnen fungierten. In Semmelweis' Abteilung trat ein gravierendes Problem zutage: Dort starben sehr viele Mütter an Kindbettfieber, wesentlich mehr als in der Abteilung, die von Hebammen geführt wurde. Die Fachwelt stand vor einem Rätsel. Es wurden zahlreiche, meist abstruse Theorien aufgestellt, als mögliche Gründe für die hohe Mortalität wurden diskutiert: Stau des Wochenflusses; Blutunreinheiten, die während der Schwangerschaft entstehen; Wettereinflüsse; Angst und Scham der Frauen; Verunreinigung der Muttermilch oder verunreinigte Luft [8]. Freilich waren all diese Theorien rein spekulativ und für einen kritischen Forscher nicht überzeugend.

Semmelweis war sehr daran gelegen, die Krankheitsursachen zu ergründen. So seziierte er, seine Kollegen und Studenten Leichen von Frauen, die der Krankheit zum Opfer gefallen waren. Dabei ging Semmelweis strukturiert vor: Er notierte seine Befunde und versuchte, Muster zu erkennen. Die Mortalitätsrate sank jedoch nicht. Das Phänomen blieb mysteriös, auch nachdem Tausende von Leichen untersucht worden waren.

Semmelweis entschloss sich, weitere Informationen einzuholen. Er setzte er sich mit Kollegen in Verbindung, um deren Erfahrungen einzuholen. Er las sämtliche Bücher und Zeitschriften, die sich dieser Thematik widmeten und setzte sich mit den Ansichten seiner Kollegen kritisch auseinander. Mit diesem Hintergrundwissen stellte er alle bis dahin bekannten Fakten zusammen. Dabei konnte er sich auch auf seine eigenen, mannigfaltigen Beobachtungen am Krankenbett und am Seziertisch und auf seine sorgfältigen Dokumentationen stützen. Über einen Zeitraum von mehreren Jahren hatte er festgehalten, wie viele Frauen an welchen Tagen in seiner Abteilung entbunden wurden, wie viele erkrankten und wie viele verstarben. Er stellte fest [8]:

- In beiden Abteilungen des Wiener Krankenhauses wurden pro Jahr etwa gleich viele Entbindungen verzeichnet (jeweils 3000–3500). Die Anzahl der Todesfälle durch Kindbettfieber belief sich auf 600–680 in der Abteilung, in der Ärzte arbeiteten. Dagegen verstarben in der Hebammenabteilung nur etwa 60 Frauen pro Jahr.
- Die Sterblichkeitsraten bei Hausgeburten und bei privat praktizierenden Gynäkologen waren wesentlich geringer als die Raten in Krankenhäusern, in denen Ärzte bei der Entbindung zugegen waren. Dies galt nicht nur in Wien, sondern auch in anderen Städten, die geburtshilfliche Abteilungen in Krankenhäusern unterhielten (u. a. in Paris, London oder Dresden).
- Das Phänomen war keineswegs neu. Bereits im Jahre 1746 (also mehr als 100 Jahre vor Semmelweis' Zeit) wurde über eine Kindbettfieber-Epidemie am Pariser Hospital Hôtel-Dieu berichtet.
- Unter dem ehemaligen Leiter der Wiener Geburtsklinik Dr. Johann Boër (1751–1835), der sanfte Methoden bei der Geburtshilfe empfahl und nur im Notfall Instrumente einsetzte, lag die durch Kindbettfieber bedingte Mortalität über 30 Jahre lang bei etwa einem Prozent. Erst unter Boërs

Nachfolger Johann Klein (1788–1856), der auch Semmelweis' Chef war, stieg die Mortalität sprunghaft an.

- Je traumatischer die Geburt war, desto höher war die Wahrscheinlichkeit zu erkranken.
- Die Erkrankungsrate (Inzidenz) war unabhängig vom Wetter.
- Mit einer Frau erkrankte meist auch deren Kind.

Damit konnte Semmelweis das Phänomen der hohen Mortalität zwar nicht erklären. Er konnte aber einige der gängigen Theorien widerlegen und kam seinem Ziel, die Ursache des Kindbettfiebers zu ergründen, ein wenig näher. Letztlich kam ihn der Zufall zu Hilfe: Nachdem sich sein Kollege Jakob Kolletschka (1803–1847) beim Sezieren einer Leiche mit dem Messer verletzt hatte und an Wundfieber erkrankte, erkannte Semmelweis die Parallelität der beiden Krankheitsbilder. Daraus schloss er messerscharf, dass den Frauen durch die Ärzte und Studenten, die in seiner Abteilung arbeiteten, „Kadaverteilchen“ übertragen wurden, die mit bloßem Auge nicht sichtbar waren. Als Konsequenz ordnete an, dass sich fortan alle Ärzte, Studenten und Hebammen mit Chlorkalk die Hände desinfizieren sollten, ehe sie den Frauen bei der Entbindung beistanden. Die Mortalität sank daraufhin dramatisch auf etwa 1,3 % im Jahre 1848. Semmelweis war es gelungen, durch strukturiertes Vorgehen und durch statistische Analysen die Krankheitsursache ausfindig zu machen und erfolgreich zu bekämpfen. Er gilt heute als einer der Väter der Evidenzbasierten Medizin [1].

Es lassen sich weitere Beispiele für medizinische Studien aus der Zeit des 19. Jahrhunderts aufführen [10]. Als im Jahre 1854 im Londoner Stadtteil Soho eine Cholera-Epidemie wütete, führte der Chirurg *John Snow* (1813–1858) eine systematische Untersuchung durch, indem er zahlreiche Erkrankte und Nicht-Erkrankte befragte. Mittels diesen Zwei-Gruppen-Vergleichs konnte er nachweisen, dass die Krankheit auf verunreinigtes Trinkwasser, das einem bestimmten Brunnen entnommen wurde, zurückzuführen war. Nachdem er den Brunnen außer Betrieb gesetzt hatte, nahm die Zahl der Neuerkrankungen drastisch ab. Damit konnte Snow die damals gängige Miasmen-Theorie widerlegen, adäquate Maßnahmen treffen und auf diese Weise viele Menschenleben retten. – Der französische Arzt *Pierre Alexandre Louis* (1787–1872) führte am Pariser Krankenhaus „Hôtel-Dieu“ eine retrospektive Studie durch, indem er zwei Gruppen von lungenkranken

Patienten verglich: Die Patienten der einen Gruppe waren zur Ader gelassen worden, während die anderen mit Bettruhe therapiert worden waren. Louis konnte nachweisen, dass die Behandlung mit Aderlass nutzlos oder gar schädlich war. Heute gilt er als der Gründer der „numerischen Methode“. – Zuletzt sei der Augustinermönch *Johann Gregor Mendel* (1822–1884) genannt, der seine Vererbungsgesetze mit statistischen Methoden verifizierte.

Leider war es diesen Forschern nicht vergönnt, den Ruhm für ihre Entdeckungen zu Lebzeiten zu ernten. Von ihren Kollegen, die sich in ihrer Ehre verletzt fühlten, ernteten sie vielfach Hohn und Spott. Die Methodik, Daten zu analysieren, um zu neuem Wissen zu gelangen, stieß überwiegend auf Ablehnung. Dennoch haben sich ihre Ideen langfristig durchgesetzt und die klinische Praxis maßgeblich beeinflusst. Auch wenn damals systematische Studien unbekannt waren und die angewandten statistischen Methoden trivial erscheinen, sollten wir anerkennend festhalten: Sie waren Pioniere auf dem Gebiet der klinischen Forschung. Es gelang ihnen durch eine strukturierte Vorgehensweise, ihr Gedankenchaos zu sortieren, unsinnige Theorien zu widerlegen, Auffälligkeiten in ihrem Zahlenmaterial zu entdecken und auf diese Weise, teilweise gegen heftige Widerstände, der Wissenschaft und zahllosen Patienten einen wertvollen Dienst zu erweisen.

3.2 Aktuelle Studien

Bleiben wir bei Entbindungen: Im Jahre 2016 wurden die Ergebnisse einer randomisierten Studie von Kirscht et al. publiziert, in der die Auswirkungen von mechanischer Dilatation der Zervix bei Frauen, die durch einen Kaiserschnitt entbunden wurden, untersucht wurde [5]. Insgesamt wurden 447 Frauen zufällig einer Gruppe (mit oder ohne Dilatation) zugewiesen. Die primäre Zielgröße war das Auftreten einer postpartalen Hämorrhagie (Blutung) innerhalb von sechs Wochen nach der Geburt. Dieser Endpunkt wurde gewählt, weil die postpartale Hämorrhagie eine lebensbedrohliche Komplikation darstellt, und weil noch keine randomisierte Studie zu dieser Thematik vorlag. Vergleiche der beiden Gruppen bezüglich demographischer Merkmale und Baseline-Charakteristika (Alter, BMI, Schwangerschaftsdauer, Gravidität, Parität, Mehrlingsschwangerschaft, Geburtsgewicht des Kindes u. a.) zeigten, dass die zu vergleichenden Gruppen weitgehend strukturgleich waren – wie man das bei einer randomisierten Studie dieses Umfangs erwartet –, und dass das Endergebnis nicht durch Confounder verzerrt sein würde.

Mit einem einfachen Gruppenvergleich ergab sich kein signifikanter Unterschied bezüglich des primären Endpunkts zwischen den Gruppen mit und ohne Dilatation: Die Inzidenzen betragen 2,4 % bzw. 1,2 % ($p = 0,479$). Allerdings gab es einen hoch signifikanten Unterschied bezüglich des Anteils der Frauen, bei denen Teile der Schwangerschaftsanlage im Uterus verblieben (0 % versus 6,2 %; $p < 0,0001$). Bezüglich weiterer sekundärer Endpunkte (Operationsdauer, Notwendigkeit einer Bluttransfusion, Auftreten von Kindbettfieber u. a.) waren keine signifikanten Unterschiede nachweisbar. Bei der Analyse des primären Endpunkts ergab sich, dass ein erhöhter BMI und eine Mehrlingsschwangerschaft mit einem höheren Risiko für das Auftreten einer Hämorrhagie assoziiert waren. Schließlich wurde ein multiples Modell erstellt, das es ermöglicht, die Wahrscheinlichkeit für das Auftreten einer postpartalen Hämorrhagie in Abhängigkeit der Einflussfaktoren Dilatation (ja/nein), BMI und Mehrlingsschwangerschaft (ja/nein) zu schätzen.

Als zweites Beispiel betrachten wir eine Studie der Adipositaschirurgie [7], bei der zwei Operationstechniken verglichen wurden, die eine Gewichtsreduktion bewirken sollten: ein Magenbypass, bei dem der Magen unterhalb des Mageneingangs abgetrennt wird, und eine Schlauchmagenbildung, bei der das Organ irreversibel verkleinert wird. Lange Zeit wurde der Nutzen dieser Techniken kontrovers diskutiert. Es herrschte Unklarheit, ob eine und gegebenenfalls welche Operationstechnik der anderen überlegen war. In diesem Fall schien eine randomisierte Studie aus ethischen Gründen problematisch zu sein, da die Entscheidung für eine bestimmte OP-Technik der behandelnde Chirurg üblicherweise zusammen mit dem Patienten traf. Also entschloss man sich, eine prospektive Beobachtungsstudie durchzuführen. 127 partizipierende Patienten erhielten einen Bypass, nur 46 Patienten einen Schlauchmagen. Die zu vergleichenden Gruppen unterschieden sich a priori bezüglich ihres BMIs: Bei den Bypass-Patienten betrug er im Mittel ($45,6 \pm 5,7$) kg/m^2 , während er bei den Patienten, die einen Schlauchmagen erhielten, mit durchschnittlich ($55,9 \pm 7,8$) kg/m^2 wesentlich höher war. Als primäre Zielgröße wurde der Exzess Weight Loss (EWL) gewählt. Diese Variable quantifiziert, um wieviel Prozent das Übergewicht durch die chirurgische Maßnahme ein Jahr nach dem Eingriff reduziert war. Beim direkten Vergleich ergab sich, dass der Bypass eine höhere Wirksamkeit aufwies ($(62,9 \pm 18) \%$ versus $(52,3 \pm 15) \%$; $p = 0,0024$). Dieses Ergebnis war jedoch wegen mangelnder Strukturgleichheit mit einem Selektionsbias behaftet. Um dies zu kompensieren, wurde eine Kovarianzanalyse als multiples Verfahren durchgeführt, in der nach dem Baseline-BMI

adjustiert wurde. Demnach beeinflusste zwar der BMI das Endergebnis ($p = 0,01$); der Unterschied bezüglich des EWL war aber nicht mehr signifikant. Am Ende wurde ein statistisches Modell entwickelt, womit der zu erwartende Wert für den EWL im Einzelfall mittels einer linearen Regressionsgleichung prognostiziert werden kann. Darüber hinaus wurden noch zahlreiche sekundäre Endpunkte analysiert (z. B. die Änderung von Stoffwechselfparametern). Dabei zeigte sich, dass der BMI vor der Operation in jedem Fall die wichtigste Einflussgröße war.

Es ist nicht anzunehmen, dass eine dieser beiden Studien die klinische Praxis nachhaltig beeinflusst hat. Trotzdem führten sie zu neuen Erkenntnissen.

4 Schlussfolgerungen

Der Vergleich von Studien aus zwei Jahrhunderten macht deutlich: Die medizinische Forschung hat sich in den vergangenen 200 Jahren dramatisch entwickelt [9]. Semmelweis und Louis befassten sich mit Krankheitsbildern, die zu ihrer Zeit lebensbedrohlich waren, weil keine wirksame Therapie zur Verfügung stand. Also war man bemüht, Risikofaktoren zu evaluieren (Semmelweis) oder die Nutzlosigkeit einer Maßnahme nachzuweisen (Louis), um das Entstehen bzw. die Progression einer Krankheit zu verhindern. Die beobachteten Effekte waren so überzeugend, dass es keiner aufwändigen statistischen Analysen bedurfte, um diese abzusichern. Die kausalen Zusammenhänge waren offensichtlich, auch ohne ein ausgeklügeltes Studiendesign. Um ethische Belange der Patienten machte sich damals niemand Gedanken. – Anders die heutige Situation: Meist steht ein Krankheitsbild im Fokus des Interesses, das nicht lebensbedrohlich ist und für das eine effiziente Therapie bereits verfügbar ist (wenngleich auch in der heutigen Zeit Krankheiten erforscht werden, die das weitere Leben der Patienten massiv beeinträchtigen, etwa in der Onkologie oder in der Neonatologie). Die Voraussetzungen zum Durchführen einer Therapiestudie sind in mehreren Gesetzesvorlagen verankert, unter anderem im Arzneimittelgesetz (AMG), im Medizinproduktegesetz (MPG) und in den Leitlinien der GCP (Good Clinical Practice). Vor Studienbeginn ist ein Votum der zuständigen Ethikkommission einzuholen. Dadurch soll gewährleistet werden, dass dem Wohle des Patienten absolute Priorität eingeräumt wird.

Vieles ist einfacher geworden. Eine Literaturrecherche und der Zugriff auf relevante Informationen ist dank des Internets quasi zu jeder Zeit und an je-

dem beliebigen Ort möglich. Es existieren eine Fülle von Fachzeitschriften und entsprechend viele Publikationen, sodass sich jeder Forscher umfassend über den aktuellen Stand seines Fachgebiets informieren kann. Die Analyse großer Datenmengen stellt – zumindest in technischer Hinsicht – keine allzu große Herausforderung dar. Anderes ist jedoch komplizierter geworden: Heutzutage werden ungleich mehr Studien durchgeführt als in früheren Zeiten, woraus entsprechend viele Publikationen resultieren. Die Datenerfassung ist weniger aufwändig, was mitunter zu einer enormen Datenmenge führt. Da fällt es mitunter schwer, die Spreu vom Weizen zu trennen.

Jeder Fortschritt entsteht aus einem Chaos – unabhängig von Zeit und Ort. Um dieses Chaos aufzulösen, bedarf es eines kritischen Geistes und einer strukturierten Vorgehensweise. Forscher – egal in welchem Jahrhundert sie leben – sollten sich zunächst umfassend über den aktuellen Stand ihres Forschungsgebietes informieren. Dazu gehört die Fähigkeit zu beurteilen, welche Erkenntnisse relevant und nachvollziehbar sind. Semmelweis musste evaluieren, welche der damals sich im Umlauf befindlichen Theorien zur Entstehung von Kindbettfieber sinnvoll und welche zu verwerfen waren. In der heutigen Zeit steht ein Forscher vor anderen Herausforderungen: Er muss aus der Fülle von Publikationen zu seinem Fachgebiet die wirklich relevanten aussortieren, lesen und bewerten. Bereits Semmelweis erkannte die Wichtigkeit von systematischen Untersuchungen, ohne die keine aussagekräftigen Informationen zu gewinnen sind. Daran hat sich bis heute nichts geändert. Während sich in früheren Zeiten die Datenanalyse im Wesentlichen auf das zeitraubende Zählen von Ereignissen und Vergleichen von Häufigkeiten beschränkte, stehen heute eine Fülle von effizienten Analysetechniken zur Verfügung, die in Sekundenschnelle Ergebnisse liefern. Dennoch ist es keineswegs trivial, eine schier unüberschaubare Datenmenge zu strukturieren und aus der Vielzahl von statistischen Techniken geeignete Verfahren auszuwählen. Mitunter verleitet dies dazu, ohne Bedacht eine Vielzahl von Tests durchzuführen in der Hoffnung, irgendein signifikantes Ergebnis zu erhalten. Dies ist jedoch keine empfehlenswerte Strategie, um dem Chaos zu entfliehen!

Der Umgang mit Daten lässt sich zwischen zwei Polen einordnen: Auf der einen Seite das Chaotische, Unstrukturierte, Unüberschaubare, Unsichere – versinnbildlicht durch die Rohdaten. Auf der anderen Seite eine Struktur, dargestellt durch ein reduktionistisches Modell, das die Komplexität des untersuchten Sachverhalts in mathematisch präziser Formelsprache beschreibt. Ein solches Modell

zu generieren, erfordert Neugierde und Leidenschaft und den unbedingten Willen des Forschenden, den Dingen auf den Grund zu gehen.

Das Reduzieren von Chaos besteht im Erkennen von Regelmäßigkeiten. Dafür sind eine gut durchdachte Planung, sorgfältig durchgeführte Untersuchungen und eine effiziente Datenanalyse unerlässlich. In letzter Zeit spielen zunehmend häufiger Begriffe wie *Big Data*, *Data Science* oder Künstliche Intelligenz (KI), die vielfach synonym verwendet werden, eine wichtige Rolle. Mit diesen Techniken sind hohe Erwartungen verknüpft: Die immensen Datenmengen bergen ein ungeheures Potential in sich. Möglicherweise lassen sich daraus Algorithmen herleiten, die bei der Früherkennung von Krankheiten, zur Vorhersage von akuten Ereignissen und bei der Entwicklung diagnostischer Tests dereinst eingesetzt werden könnten. Andererseits besteht gerade mit Big Data die Gefahr, im Chaos zu versinken!

Allerdings kann das Verharren an neuen Erkenntnissen weitere Forschung behindern [6]. Es birgt die Gefahr in sich, dass man – ohne die Besonderheiten des Einzelfalls gebührend zu berücksichtigen – Entscheidungen trifft, die sich im Nachhinein als wenig sinnvoll erweisen, oder dass man andere Aspekte außer Acht lässt. Wenn beispielsweise in einer randomisierten Studie die Überlegenheit eines neu entwickelten Arzneimittels im Vergleich zu einer Standardtherapie nachgewiesen wurde, ist man als behandelnder Arzt mitunter versucht, jedem Patienten die vermeintlich wirkungsvollere Therapie zu verordnen ohne einen Gedanken daran zu verschwenden, ob sie im konkreten Einzelfall wirklich erfolgversprechend ist. Man konzentriert sich auf eine einzige Zielgröße – etwa die Wirksamkeit – und lässt besondere Charakteristiken eines Patienten oder andere Eigenschaften der Therapie (z. B. Nebenwirkungen) außer Acht. Damit wird der Patient nicht mehr als ein Individuum wahrgenommen, sondern als eine Beobachtungseinheit, die durch eine mathematische Gleichung berechenbar ist.

Chaos und Struktur – das ist kein Widerspruch! Chaos kann Verwirrung stiften und gleichzeitig die Fantasie anregen. Strukturen können einerseits einengen und andererseits Sicherheit bieten. Fazit: Die Mischung macht's! Jeder Mensch (und jeder Wissenschaftler) braucht ein gewisses Maß an Chaos – sonst wäre das Leben langweilig, und jegliche Kreativität wäre im Keime erstickt. Im Chaos zu verharren, ist jedoch auf Dauer unbefriedigend. Deshalb werden Maßnahmen ergriffen, um das Chaos aufzulösen und ein Modell zu generieren, das die komplexen Zusammenhänge zumindest partiell beschreibt. Aber ein solches Modell ist

niemals perfekt! Deshalb verlangt auch dieser Zustand nach einiger Zeit Anpassung an neue Gegebenheiten. Immer wieder werden sich neue Herausforderungen ergeben – und der Kreislauf beginnt von vorn. Der Philosoph Manfred Hinrich (1926–2015) hat dies mit folgendem Aphorismus treffend zum Ausdruck gebracht: „Chaosbewältigung ist das halbe Leben, Chaosschaffen die andere Hälfte“.

Literatur

1. Aigner F, Grafenholder D: Leben retten, mit Chlorkalk und Statistik. <https://www.naklar.at/content/features/semmelweis/>
2. Atkinton AJ Jr, Colburn WA, De Gruttola VG, DeMets DL, Dwoning GJ, Hoth DF, Oates JA, Peck CC, Schooley RT, Spilker BA, Woodcock J, Zeger SL: Biomarkers and Surrogate Endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther* 69 (3): 89–95; 2001
3. Gierschik P: Die Phaseneinteilung klinischer Studien. In: Lenk C, Duttge G, Fange-rau H (Hrsg.): *Handbuch Ethik und Recht der Forschung am Menschen*. Springer-Verlag 2014
4. Hammer GP, Prel du JP, Blettner M: Vermeidung verzerrter Ergebnisse in Beobachtungsstudien. *Dtsch Arztebl Int* 1061 (41): 664–668; 2009
5. Kirscht J, Weiß C, Nickol J, Berlit S, Tuschy B, Hoch B, Trebin AV, Große-Steffen T, Sütterlin M, Kehl S: Dilatation or no dilatation of the cervix during cesarean section (Dondi Trial): a randomized clinical trial. *Arch Gynecol Obstet* 295: 39–43; 2017
6. Kriz J: Chaos, Angst und Ordnung. *Ärztliche Praxis. Neurologie Psychiatrie* 1; 24–27; 1999
7. Otto M, Elrefai M, Krammer J, Weiß C, Kienle P, Hasenberg T: Sleeve Gastrectomy and Roux-en-Y Gastric Bypass Lead to Comparable Changes in Body Composition after Adjustment for Initial Body Mass Index. *Obes Surg* 26(3): 479–485; 2016
8. Nuland SB: *Ignaz Semmelweis. Arzt und großer Entdecker*. Piper Verlag München 2008
9. Tröhler U: Die therapeutische „Erfahrung“ – Geschichte ihrer Bewertung zwischen subjektiv sicherem Wissen und objektiv wahrscheinlichen Kenntnissen. In: Köbberling J: *Die Wissenschaft in der Medizin. Selbstverständnis und Stellenwert in der Gesellschaft*. Schattauer-Verlag 1992

10. Weiß C: Entwicklung der Medizinischen Statistik in Deutschland. Der lange Weg dahin. GMSD Med Inform Biom Epidemiol; 2005
11. Weiß C: Regressionsanalysen. In: Basiswissen Medizinische Statistik, Springer-Verlag, 7. Auflage 2019

Über die Autorin

Christel Weiß ist Professorin für Biomathematik und Epidemiologie an der Medizinischen Fakultät Mannheim der Universität Heidelberg. Zu ihren Tätigkeiten zählen Lehrveranstaltungen für Studierende der Medizin und Masterkurs-Absolventen, das Durchführen von Seminaren sowie die Beratung von Ärzten, wissenschaftlichen Mitarbeitern und Doktoranden bei der Planung und Durchführung von klinischen und epidemiologischen Studien. Frau Weiß ist Autorin des Lehrbuchs „Basiswissen Medizinische Statistik“ (erschieden im Springer-Verlag, 7. Auflage), des Ratgebers „Promotion. Die medizinische Doktorarbeit – von der Themensuche bis zur Dissertation“ (zusammen mit Prof. Dr. Axel Bauer, erschienen im Thieme-Verlag, 4. Auflage) sowie Koautorin zahlreicher Papers und Buchbeiträge.

Korrespondenz:

Prof. Dr. Christel Weiß

Medizinische Fakultät Mannheim der Universität Heidelberg

Abteilung für Medizinische Statistik und Biomathematik

Theodor-Kutzer-Ufer 1

68167 Mannheim, Germany

E-Mail: Christel.Weiss@medma.uni-heidelberg.de

Struktur und Chaos: Kleinskalige Austauschprozesse zwischen Atmosphäre und Meer

BERND JÄHNE

Heidelberg Collaboratory for Image Processing (HCI) am IWR
und Institut für Umweltphysik
Universität Heidelberg

Zusammenfassung

Austauschprozesse zwischen Atmosphäre und Meer bestimmen die Verteilung von Energie, Wärme und Stoffen auf unserem Planeten. In Bezug auf den Klimawandel ist dabei der Transport von Kohlendioxid von der Atmosphäre ins Meer von eminenter Bedeutung. Die Geschwindigkeit des Austauschs ergibt sich aus dem Zusammenspiel zwischen molekularer Diffusion und Turbulenz unmittelbar an der Meeresoberfläche. An dem bisher am wenigsten verstandenen Prozess des Austauschs von Spurengasen wird gezeigt wieso dies so ist und dass damit die Bildung von raumzeitlichen Strukturen verbunden ist, die wohl geordnet bis chaotisch sein können.

1 Einleitung

Gut 70 % der Oberfläche des Planeten Erde bedecken die Weltmeere. Damit stellt die Meeresoberfläche nicht nur die größte Grenzfläche der Erde dar, sondern über sie tauschen sich die beiden großen fluiden Kompartimente Ozean und Atmosphäre aus und bestimmen damit entscheidend, wie sich was auf der

Erde verteilt. Der ständig wehende Wind ist einer der treibenden Kräfte der großräumigen Ozeanströmungen. Durch Temperaturunterschiede zwischen Luft und Wasser wird Wärme ausgetauscht und die Verdunstung an der Wasseroberfläche kühlt die Ozeanoberfläche (latenter Wärmeaustausch) und führt gleichzeitig der Atmosphäre Wasserdampf zu.

Schließlich werden Gase und flüchtige Stoffe ausgetauscht. Darunter sind viele klima- und umweltrelevante Stoffe. Das bekannteste Beispiel ist Kohlendioxid. So hat der Ozean etwa 40 % des durch menschliche Aktivitäten — im wesentlichen durch die Verbrennung fossiler Energieträger — in die Atmosphäre gelangte Kohlendioxid aufgenommen. Ohne diesen Austausch wäre der Kohlendioxidgehalt in der Atmosphäre entsprechend höher. Weitere wichtige Beispiele sind Dimethylsulfid (DMS), das von Plankton produziert in die Atmosphäre gelangt, dort oxidiert wird und dann als Partikel die Wolkenbildung fördert, und natürlich Sauerstoff, dessen Konzentration im Wasser entscheidend für die Biologie der Gewässer ist.

In diesem Beitrag wird zuerst erläutert, dass diese Austauschprozesse inhärent nichtlinearer Natur sind, erst durch ein Wechselspiel zwischen turbulenter Strömung und molekularer Diffusion zustande kommen und erheblich durch die gleichzeitige Generierung von Wasserwellen durch den Wind und die Chemie oberflächenaktiver Substanzen beeinflusst werden. Da man diese Prozesse mit bloßem Auge nicht sieht, war es notwendig, diese durch eine Reihe von geeigneten bildaufnehmenden Messtechniken sichtbar zu machen. Mit diesen konnte dann gezeigt werden, dass sich abhängig von den vorherrschenden Mechanismen sehr unterschiedliche Strukturen in Raum und Zeit herausbilden, die charakteristische Muster aufweisen, aber auch völlig chaotisch sein können.

2 Grenzschichten und Transfargeschwindigkeiten

Da die Strömungen nahe der Meeresoberfläche parallel zu dieser verlaufen, kann diese keinen vertikalen Austausch bewirken. Daher kommt den zufälligen Schwankungen, sprich der Turbulenz, die entscheidende Bedeutung für den vertikalen Austausch zu. Ohne diese, also in einer sogenannten laminaren Strömung, gäbe es keinen durch die Strömung induzierten vertikalen Austausch. Turbulente Fluktua-

tionen wiederum gibt es nur, weil die grundlegende Gleichung für Strömungen mit Reibung, die Navier-Stokes-Gleichung, nicht linear ist.¹

Wovon die Stärke dieser turbulenten Durchmischung abhängt, hatte Prandtl schon vor knapp 100 Jahren herausgefunden. Man kann den Prozess wie die molekulare Diffusion betrachten, die auf der zufälligen thermischen Bewegung der Moleküle beruht. Diese bewegen sich geradlinig, bis sie mit einem anderen Molekül zusammenstoßen und sich danach in eine andere Richtung weiter bewegen. Ohne mathematische Ableitung kann man intuitiv verstehen, dass die Durchmischung um so schneller geht, je länger die geradlinige Bewegung andauert (mittlere freie Weglänge) und je größer die Geschwindigkeit ist. Das Produkt der beiden Größen geht in die Diffusionskonstante D ein.

Während beim Transport in der Hauptströmungsrichtung parallel zur Meeresoberfläche die Transportstrecke (Advektion) proportional mit der Zeit wächst, ist dies wegen der zufälligen Bewegung in alle Richtungen bei der molekularen Diffusion nicht der Fall. Zum einen folgt die Ausbreitung in alle Richtungen gleichmäßig, zum anderen wächst die Verbreiterung nur mit der Wurzel aus der Zeit. Mit anderen Worten: Diffusionsprozesse sind nur auf kurze Distanzen effektiv. In der Atmosphäre werden für die Verbreiterung um 1 mm etwa 30 ms benötigt, für 1 cm schon 3 s und für 1 m lange 9 h.

Ein weiteres Charakteristikum ist für die molekulare Diffusion entscheidend. In der Atmosphäre ist die mittlere freie Weglänge knapp unter einem Zentel Mikrometer. Daher „sieht“ man nichts von diesen Fluktuationen auf den Skalen, die für die Austauschprozesse von Bedeutung sind. Molekulare Diffusion kann damit Konzentrationsunterschiede nur glätten, aber nicht zur Bildung irgendwelcher Strukturen führen.

Wie kann nun turbulente Diffusion überhaupt effektiv sein, wenn diese genauso beschrieben werden kann wie die molekulare? Das Entscheidende ist, dass die freie Weglänge ungleich größer ist. Das Prandtl'sche Mischungsweg-Konzept besagt, dass turbulente Fluktuationen eine Weglänge haben können, die so groß ist wie die Entfernung z zur Grenzfläche. Eine turbulente Fluktuation (Wirbel) kann daher, unabhängig von ihrer Entfernung, bis an die Meeresoberfläche reichen. Damit ist in

¹ Die mittleren turbulenten Austauschtermine resultieren unmittelbar aus dem nichtlinearen Term in dieser Gleichung, wenn diese in mittlere und fluktuierende Terme aufgetrennt wird (Reynoldsgleichung).

der Luft die mittlere freie Weglänge in 10 m Entfernung um acht Größenordnungen höher als für die molekulare Diffusion. Die turbulenten Geschwindigkeitsfluktuationen sind mit 0,05–0,5 m/s allerdings drei bis vier Größenordnungen kleiner als die Molekülgeschwindigkeiten mit etwa 500 m/s. Insgesamt dominiert aber trotzdem die große freie Weglänge und damit ist die turbulente Durchmischung in 10 m Entfernung von der Grenzfläche 10 000–100 000-mal effektiver als die molekulare. Aus diesen Grundtatsachen lassen sich leicht eine Reihe wichtiger Schlussfolgerungen ziehen.

Der vertikale turbulente Transport wird mit Annäherung an die Meeresoberfläche immer weniger effektiv. Irgendwann werden die Fluktuationen durch die Viskosität (Reibung) zusätzlich vermindert und fallen daher zur Grenzfläche noch schneller als linear ab. An der Grenzfläche selbst ist die turbulente Diffusionskonstante Null. Überraschend: Gäbe es keine molekulare Diffusion, so wäre deswegen keinerlei Austausch über die Wasseroberfläche möglich.

Bedingt durch diesem Effekt bilden sich auf beiden Seiten der Meeresoberfläche Grenzschichten aus (Abb. 1a), in denen die molekulare Diffusion größer ist als die turbulente. Dort ist der Widerstand für den Transport am größten und damit der Konzentrationsgradient am steilsten. Daraus resultiert zum Beispiel, dass der Wind in 10 m Höhe erst an der Kante der viskosen Grenzschicht, also Millimeter über der Wasseroberfläche, auf die Hälfte abgefallen ist. Das gleiche gilt für den Verlauf von vertikalen Temperatur- und Stoffkonzentrationsgradienten.

Da die molekulare Diffusion der Geschwindigkeit (physikalisch korrekter ausgedrückt, der Impulsdichte), Wärme und Stoffen in Luft etwa gleich groß ist, haben die Grenzschichten für alle drei Größen die gleiche Dicke. Die Geschwindigkeit für den vertikalen luftseitigen Austausch von der Meeresoberfläche bis in 10 m Höhe beträgt etwa 0,1–0,2 % der Windgeschwindigkeit und liegt größenordnungsmäßig im Bereich von Zentimetern pro Sekunde. Diese vertikale Transportgeschwindigkeit, Transfargeschwindigkeit genannt, kann man sich anschaulich als die Geschwindigkeit vorstellen, mit der die transportierte Größe von der Meeresoberfläche in die Atmosphäre geschoben wird. Damit dauert es Minuten, um zum Beispiel verdunstetes Wasser auf 10 m Höhe zu transportieren.

Auf der Wasserseite sind die Verhältnisse wesentlich komplizierter, da die Diffusion von Impuls, Wärme und Stoffen nicht mehr gleich schnell ist. Wärme wird etwa zehnmals und Stoffe und gelöste Gase sogar tausendmal langsamer molekular transportiert als Impuls. Letzteres liegt daran, dass Impuls und Wärme

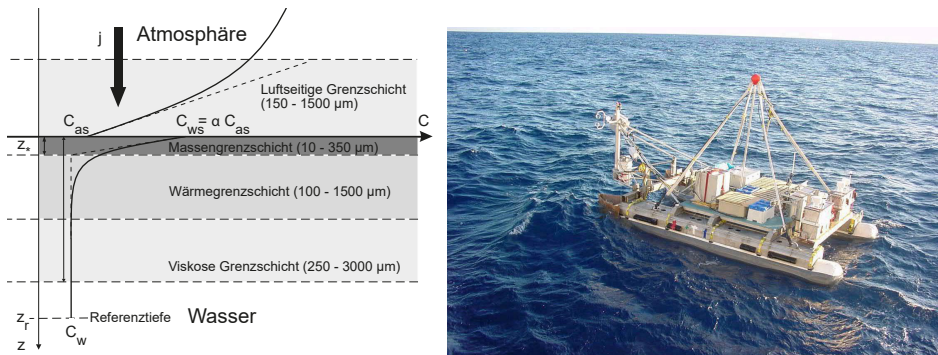


Abbildung 1: (a) Grenzschichten für Austauschprozesse über die Meeresoberfläche, nach Kunz und Jähne, 2018 (links, nähere Erläuterung im Text); (b) Forschungskatamaran zur Untersuchung kleinskaliger Austauschprozesse an der Meeresoberfläche gebaut von Dr. Erik Bock (1961–2001), zuerst am Woods Hole Ozeanographischen Institut, dann am Institut für Umweltphysik, Universität Heidelberg tätig (rechts, Foto Nachlass Erik Bock).

durch Stöße übertragen werden, beim Stoffaustausch aber das entsprechende Molekül sich selbst bewegen muss. Dazu muss Energie aufgebracht werden, um Anziehungskräfte zwischen Wassermolekülen zu überwinden.

In Wasser ist daher die Wärmegrenzschicht nur halb so dick wie die viskose für die Geschwindigkeit, die Massengrenzschicht sogar etwa zehnmals dünner (Abb. 1a). Das bedeutet für den Stoffaustausch: Innerhalb der viskosen Grenzschicht ist zwar die Reibung groß genug, dass molekularer Transport für Geschwindigkeit gegenüber dem turbulenten Transport dominiert. Die kleinen Restturbulenzen reichen aber noch aus, um effektiv Stoffe zu transportieren. Erst viel näher an der Grenzfläche sind diese soweit reduziert, dass schließlich auch für den Stofftransport die molekulare Diffusion überwiegt.

Die geringe Diffusionskonstante für Stoffe in Wasser bewirkt, dass sich fast der ganze Widerstand für den Transport auf eine submillimeter dünne Grenzschicht konzentriert und dass die Stärke der kleinen Restturbulenz in der viskosen Grenzschicht die Dicke der Grenzschicht und damit die Geschwindigkeit für den Austausch bestimmt. Zu weiteren Konzentrationsgradienten im Wasser kommt es nur dann, wenn stabile Schichtungen vorliegen, die den turbulenten Transport unterdrücken. Dies ist vor allem an der Temperatursprungschicht im Ozean zwischen warmem Oberflächenwasser und kaltem Tiefenwasser in ca. 50–200 m Tiefe der Fall.

Die Austauschgeschwindigkeit für Stoffe auf der Wasserseite ist wegen der geringen Diffusion viel kleiner als auf der Luftseite. Sie liegt — unter Ausschluss sehr hoher Windgeschwindigkeiten — bei 2–70 cm/h und ist damit ca. tausendmal langsamer als auf der Luftseite. Für den Transport von Wärme und Impuls ist sie nur grob hundert- beziehungsweise vierzigmal langsamer.

Beim Vergleich zwischen luftseitigem und wasserseitigem Transport ist noch ein weiterer Parameter von entscheidender Bedeutung: Wie viel kann pro Volumen in den beiden Medien transportiert werden? Das Verhältnis im thermodynamischen Gleichgewicht heißt Partitionskoeffizient oder dimensionslose Löslichkeit. Nur wenn pro Volumeneinheit in beiden Medien die gleiche Menge transportiert werden kann, könnte man die Transfergeschwindigkeiten direkt vergleichen. Generell muss aber die wasserseitige Transfergeschwindigkeit mit der Löslichkeit multipliziert werden, um sie mit der luftseitigen zu vergleichen. Damit ist bei einer Löslichkeit von etwa tausend die wasserseitige Transfergeschwindigkeit für den Stofftransport so groß wie die luftseitige und die Transportwiderstände liegen jeweils zur Hälfte in der Atmosphäre und dem Wasser.

Das bedeutet, dass für die schon genannten Beispiele Kohlendioxid, Methan, Dimethylsulfid und Sauerstoff mit Löslichkeiten von etwa 0,83, 0,035, 12,7 und 0,031 bei 20 Grad Celsius der wasserseitige Austauschprozess der bestimmende ist. Es gibt auch eine Reihe flüchtiger organischer Verbindungen, deren Löslichkeit in Wasser so hoch ist, dass der Transportwiderstand im Wasser gering ist. Dazu zählen Alkohole wie Methanol, Ketone und diverse Stickstoffverbindungen (Amine, Anilin, Pyridin). Ebenso kann eine schnelle chemische Reaktion den wasserseitigen Widerstand kurzschließen. Das bekannteste Beispiel ist Schwefeldioxid. Im folgenden beschränken wir uns auf den am wenigsten verstandenen Austauschprozess mit Gasen niedriger Löslichkeit, bei dem der Flaschenhals für den Transport die außerordentlich dünne wasserseitige Massengrenzschicht an der Meeresoberfläche ist.

Interessant ist es, auch den Partitionskoeffizienten für Impuls und Wärme zu betrachten. Dieser ist bei Impuls einfach durch das Dichteverhältnis zwischen Wasser und Luft gegeben, also etwa tausend. Für Wärme ergibt sich der Partitionskoeffizienten aus dem Verhältnis der volumenbezogenen Wärmekapazität also dem Produkt aus Dichte und spezifischer Wärme. Da die spezifische Wärme von Wasser etwa viermal größer ist als die der Luft, ergibt sich für Wärme ein Partitionskoeffizient von etwa viertausend. Das Ergebnis ist für Impuls und Wärme das

gleiche: Der Hauptwiderstand liegt wegen Partitionskoeffizienten, die wesentlich größer sind als das Verhältnis der Transfersgeschwindigkeit, auf der Luftseite. Das Ergebnis ist für den Impuls Allgemeinwissen: Winderzeugte Wasseroberflächengeschwindigkeiten sind nur ein kleiner Bruchteil der Windgeschwindigkeit.

Wie kann nun die Austauschgeschwindigkeit über die Meeresoberfläche gemessen werden? Der Schlüssel dazu ist, dass der Transport von Stoffen — im Gegensatz zu der turbulenten Strömung selbst — ein linearer Prozess ist. Daher ergibt sich die vertikale Austauschgeschwindigkeit, die schon eingeführte Transfersgeschwindigkeit k direkt aus dem Verhältnis der Flussdichte j , also der Stoffmenge, die pro Fläche und Zeiteinheit durch die Grenzschicht transportiert wird, und der Konzentrationsdifferenz Δc über die Transportstrecke:

$$k = \frac{j}{\Delta c}, \quad z_* = \frac{D}{k}, \quad t_* = \frac{z_*}{k} = \frac{D}{k^2} \quad (9.1)$$

Nur mit der Annahme, dass der turbulente Transport an der Meeresoberfläche Null ist, lässt sich die Transfersgeschwindigkeit über die Diffusionskonstante D mit der Grenzschichtdicke in Beziehung setzen. Das folgt aus dem ersten Fickschen Gesetz, wonach die Flussdichte proportional zur Diffusionskonstanten und dem Konzentrationsgradienten ist:

$$j = -D \left. \frac{\partial c}{\partial z} \right|_{z=0} \rightarrow j = D \frac{\Delta c}{z_*} \rightarrow k = \frac{D}{z_*} \quad (9.2)$$

Die Zeit, die der Transport durch die Grenzschicht mit der Dicke z_* benötigt, ergibt sich schließlich aus dem Verhältnis von Strecke und Geschwindigkeit. Damit konnte mit wenig Aufwand gezeigt werden, dass sich die Transfersgeschwindigkeit entweder durch die Messung von Stromdichten und Konzentrationsdifferenz messen lässt, durch die direkte Bestimmung der Grenzschichtdicke oder der Zeitkonstante t_* . Die Größenordnung der Grenzschichtdicke (Abb. 1a) und der Transfersgeschwindigkeiten (2–70 cm/h) wurden schon genannt. Die Zeitkonstante t_* liegt im Bereich von 0,06–60 s.

3 Gasaustausch und Wind: Feldmessungen

Der Blick auf die Meeresoberfläche (Abb. 1b) zeigt, dass die Austauschprozesse dem bloßen Auge verborgen bleiben. Wir sehen weder das turbulente Strömungsfeld über dem Wasser, dass dieses an der Wasseroberfläche eine Scherströmung

induziert, noch die resultierenden Austauschprozesse wie die Verdunstung von Wasser oder den Austausch von Wärme und Spurengasen. Es ist auch offensichtlich, dass Messungen in einer submillimeter dicken Massengrenzschicht an einer wellenbewegten Wasseroberfläche im Feld sehr schwierig sind.

Daher wurde die Transfersgeschwindigkeit auf dem Ozean bisher ausschließlich aufgrund der Messung der Flussdichte j und der Konzentrationsdifferenz Δc bestimmt. Dazu kann man entweder Spurenstoffe in die Warmwasserschicht geben und messen, wie deren Konzentration sich zeitlich ändert. Mit einem Spurenstoff geht dies nicht, da die Konzentration sich auch durch die horizontale Ausbreitung verringert. Nimmt man aber zwei Spurenstoffe („Dual-Tracer-Methode“) mit deutlich unterschiedlicher Diffusionskonstante wie Helium-3 und Schwefelhexafluorid, so kann aus dem schnelleren Konzentrationsabfall von Helium-3 durch den erhöhten Gasaustausch dieser von dem Mischungseffekt unterschieden werden.

Die zweite gängige Methode ist es, die Flussdichte j direkt aus den turbulenten Fluktuationen in der Atmosphäre zu messen (Eddykorrelationsmethode) und aus der gegebenen Konzentrationsdifferenz zwischen Atmosphäre und oberflächennahem Wasser die Transfersgeschwindigkeit zu bestimmen.

Die Arbeitshypothese für die Feldmessungen hat sich in den letzten fünfzig Jahren nicht grundlegend geändert: Dadurch dass — wie geschildert — der Impuls- und Stofftransport den gleichen Mechanismen unterliegen und der Wind das ganze System antreibt, sollte es einen direkten Zusammenhang zwischen der Transfersgeschwindigkeit für Gase und der Windgeschwindigkeit geben. Eine Zusammenstellung von vielen Feldmessungen, die in den letzten Jahren durchgeführt wurden, zeigt allerdings ein komplexeres Bild und Lücken (Abb. 2).

Ein Anstieg des Gasaustausches (Transfersgeschwindigkeit) mit dem Wind ist klar zu erkennen, aber ebenso die große Streuung der Einzelmessungen. Feldmessungen sind oft mit hohen relativen Fehlern verbunden. Eine Regression der Daten mit einer empirischen Windgeschwindigkeitsabhängigkeit reduziert die Unsicherheit im Windgeschwindigkeitsbereich zwischen 4 und 20 m/s deutlich, nun treten aber signifikante Differenzen zwischen den drei Methoden bzw. Spurengasen auf, für die es bis heute keine widerspruchsfreie Erklärung gibt.

Dass es keine Messdaten über 20 m/s gibt, ist nicht verwunderlich, da bei rauer See Messungen schwierig und gefährlich sind. Warum gibt es aber so gut wie keine Messungen bei niedrigen Windgeschwindigkeiten unter 4 m/s? Das liegt

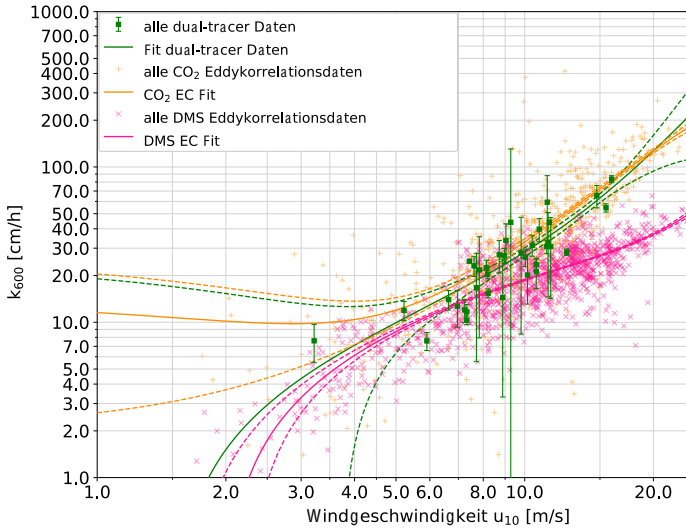


Abbildung 2: Zusammenstellung von Transferrgeschwindigkeiten k_{600} (korrigiert auf die Diffusionskonstante von Kohlendioxid bei 20 Grad Celcius) gemessen auf dem Ozean mit der „Dual-Tracer-Methode“ und aus direkten Flussmessungen für Kohlendioxid und Dimethylsulfid (DMS) mit Hilfe der Eddykorrelationsmethode und aufgetragen gegen die Windgeschwindigkeit in 10 m Höhe.

daran, dass der Gasaustausch ein so langsamer Prozess ist. Eddykorrelationsmessungen leiden unter der geringen Flusssdichte und weisen deswegen einen hohen statistischen Fehler auf. Die „Dual-Tracer-Methode“ versagt noch früher, da die gasaustausch-bedingten Konzentrationsänderungen so langsam sind, dass sie kaum zu messen sind und die Windgeschwindigkeit über so lange Zeiträume nicht konstant bleibt.

Das ist eine ernüchternde Bilanz aus einem halben Jahrhundert Feldmessungen mit essentiell unveränderten Methoden. In einem begrenzten Windgeschwindigkeitsbereich von ca. 4–20 m/s konnte eine empirische Windgeschwindigkeitsbeziehung aufgestellt werden, die allerdings noch signifikante systematische Unsicherheiten aufweist. Bei höheren oder niedrigeren Windgeschwindigkeiten sind sichere Voraussagen nicht möglich. Und es gibt weder sichere Aussagen auf Basis der Feldmessung alleine, ob es außer der Windgeschwindigkeit noch weitere

Parameter gibt, die den Gasaustausch über die Meeresoberfläche beeinflussen, noch Hinweise, ob und welche Strukturen in Raum und Zeit auftreten könnten.

4 Der Ozean im Labor

Labormessungen haben gegenüber Feldmessungen den großen Vorteil, dass gezielte und systematische Messungen durchgeführt werden können. In der Tat haben Labormessungen schon sehr früh wesentliche Erkenntnisse über die Mechanismen des Gasaustausches beigetragen. Das hat vor allem mit etwas zu tun, was in Abb. 1b und Abb. 3 sofort auffällt, in die bisherigen Überlegungen aber noch nicht eingeflossen ist: Der Wind erzeugt nicht nur turbulente Strömungen und Grenzschichten beiderseits der Ozeanoberfläche, sondern auch Wellen.

Da Wellen keinen gravierenden Einfluss auf den luftseitig kontrollierten Stoff- (Verdunstung) und Wärmeaustausch haben, ging man davon aus, dass dies auch für den wasserseitig kontrollierten Gasaustausch der Fall sein sollte. Weit gefehlt. Der Autor konnte schon 1980 und 1984 zeigen, dass bei gleicher Windgeschwindigkeit (genauer: Windschub an der Wasseroberfläche) die Transfergeschwindigkeit mit Wellen bis zu fünfmal höher sein kann als ohne Wellen.

Damit stellt sich die Frage, wieso in einem Wind-Wellen-Kanal bei gleichem Wind einmal Wellen erzeugt werden und einmal nicht. An diesem Punkt kommt die Chemie ins Spiel. Geringste Konzentrationen von sogenannten oberflächenaktiven Substanzen, deren Moleküle einen wasserabstoßenden und wasseranziehenden Teil haben, sammeln sich an der Wasseroberfläche und bilden dort einen monomolekularen Film. Dieser kann effektiv Wellen dämpfen und damit offensichtlich auch den Gasaustausch reduzieren.

Nun stellen sich zwei entscheidende Fragen bezüglich der Labormessungen. Tauchen die im Labor gefundenen Effekte auch auf dem Ozean auf und, wenn ja, kann man diese im Labor adäquat simulieren? Nur bei positiver Antwort auf beide Fragen können Laborexperimente dazu beitragen, die Mechanismen des Gasaustausches zu entschlüsseln und den Gasaustausch an der Meeresoberfläche quantitativ aus den vorliegenden Bedingungen zu bestimmen.

Die positive Antwort auf die erste Frage ist heute unumstritten. Die Erzeugung oberflächenaktiver Substanzen im Meer ist eng mit biologischen Aktivitäten verknüpft. Damit sind oberflächenaktive Substanzen überall im Meer zu finden. Es ist nachgewiesen, dass schon geringste Konzentrationen, die noch nicht einmal



Abbildung 3: Im ringförmigen Heidelberger Aeolotron wurden während des *SOPRAN-Projekts* 2014 Gasaustauschmessungen mit Meerwasser durchgeführt. Im Gegensatz zu linearen Kanälen können durch die unendliche Wechselwirkungszeit zwischen Wind und Wellen sich viel größere Wellen aufbauen als in linearen Kanälen (Foto B. Jähne). *CampusTV-Bericht*.

die Oberflächenspannung signifikant reduzieren, Wellen und den Gasaustausch dämpfen können. Unter dem Schlagwort „Surface Micro Layer“ (SML) zeichnet sich die Entwicklung eines neuen interdisziplinären Forschungsschwerpunkts in der Ozeanographie ab.

Die Antwort auf die zweite Frage war das *Heidelberger Aeolotron*² und Entwicklung bildaufnehmender Messtechniken am Institut für Umweltphysik. Dies war ein schwieriger Spagat. Zum einen musste die Versuchseinrichtung groß genug sein, um Wellen zu erzeugen, die den Bedingungen an der Meeresoberfläche entsprechen, zum anderen sollte es noch möglich sein, bildaufnehmende Messtechniken einzusetzen, die die dünne Massengrenzschicht auflösen.

Das Ergebnis war der Bau des weltweit einmaligen Heidelberger Aeolotrons mit 10 m Durchmesser, benannt nach Aeolus, dem griechischen Gott der Winde (Abb. 3). Vier Axialventilatoren erzeugen den Wind in einer 0,6 m breiten und 2,4 m hohen ringförmigen Rinne, die bis zu 1 m mit Wasser gefüllt ist. Aufgrund dieser Geometrie bauen sich die Wellen so lange auf, bis sie, wie auf dem Ozean, mit dem Wind im Gleichgewicht stehen. Das dauert viele Minuten und es können sehr große Wellen erzeugt werden (Abb. 3). In einem herkömmlichen linearen Windkanal ist das nicht möglich, dort erreicht das Wellenfeld nie einen Gleichge-

² Fett und kursiv gesetzte Textteile führen zu Links

wichtszustand. Durch den Aufbau des Wellenfeldes mit der Überstreichlänge des Windes ist es außerdem inhomogen.

Jetzt werden nur noch Messtechniken benötigt, mit denen lokal und instantan Wellen und Gasaustausch gemessen werden können. Bildaufnehmende optische Messtechniken machen dies möglich. Sie erlauben auch erstmals Messungen beim Anfahren und Abflauen des Winds und damit Untersuchungen, wie sich die Transfergeschwindigkeit für Gase beim Auf- und Abbau des Wellenfeldes, so wie es natürlich auf dem Ozean vorkommt, verändert. Und schließlich liefern diese Techniken endlich den gewünschten Einblick in die räumlichen und zeitlichen Strukturen der Austauschprozesse und damit deren Mechanismen.

5 Bildaufnehmende Messtechnik

5.1 Windwellen

Da Windwellen offensichtlich eine zentrale Rolle für die wasserseitigen Austauschprozesse spielen, wurde zuerst eine bildgebende Wellenmesstechnik entwickelt. Das Prinzip ist ähnlich der als „Gestalt aus Schattierung“ bekannten Methode, mit dem das menschliche Sehvermögen dreidimensionale Strukturen erkennt. Objektflächen erscheinen unterschiedlich hell, je nach dem welche Neigung sie haben.

Auf diese Weise sehen wir übrigens auch Wellen auf der Wasseroberfläche (Abb. 1b). Der Helligkeitsunterschied wird hierbei durch die Änderung des Reflexionskoeffizienten mit dem Winkel zwischen Oberflächennormale und Sehrichtung bewirkt. In der Ozeanographie ist dieses Verfahren als Stilwell-Fotografie bekannt, war aber wegen diverser praktischer Probleme wenig erfolgreich.

Da Wasser aber nur wenige Prozente des einfallenden Lichts reflektiert, ist es zur Sichtbarmachung der Wellenneigung besser, die Brechung des Lichts zu nutzen. Dabei schaut eine Kamera aus großer Entfernung senkrecht von oben auf die Wasseroberfläche. Ist diese durch eine Welle geneigt, so kommt das Licht, das die Kamera sieht, abhängig von der Neigung der Welle aus einer bestimmten Richtung. Jetzt ist es nur noch notwendig eine Lichtquelle zu bauen, deren Helligkeit von der Richtung der Strahlen abhängt.

Eine solches System hat unsere Arbeitsgruppe erstmals Ende der 1980er bis Anfang der 1990er Jahre in den großen linearen Wind-Wasser-Kanälen der Universität Marseille und von Delft Hydraulics in den Niederlanden eingesetzt. Damals

war es noch nicht möglich, Bildsequenzen in einem Rechner zu digitalisieren. Daher wurden diese analog auf einem Videorekorder mit 60 Bildern pro Sekunde aufgezeichnet. Trotz dieser technischen Beschränkungen ergab sich gleich ein überraschendes Ergebnis, das die historischen Bilder von vor 30 Jahren in Abb. 4 zeigen.

Schon bei niedrigen Windgeschwindigkeiten können Wellen brechen. Dieser Prozess ist mit bloßem Auge kaum zu sehen, da er innerhalb eines Bruchteils einer Sekunde geschieht. Man sieht kein Überschlagen der Wellen mit Blaseneintrag ins Wasser, wie das jeder von Wellen am Strand kennt, sondern hier wird eine Welle so steil, dass die regelmäßige Wellenbewegung in eine turbulente Wirbelströmung umschlägt, die an der unregelmäßigen Oberflächenstruktur zu erkennen ist. Dieser Prozess wird „micro-scale breaking“ genannt und ist ein schönes mit den Austauschprozessen verbundenes Beispiel für den Übergang einer regelmäßigen Struktur, der Wellenbewegung, in eine ungeordnete Turbulenz.

Damit wird schlagartig ein Mechanismus klar, mit dem Wellen den Gasaustausch erhöhen: Steile Wellen können durch solche Ereignisse intensive Turbulenz unmittelbar an der Wasseroberfläche erzeugen. Inzwischen ist bekannt, dass die Häufigkeit des kleinskaligen Brechens mit der Windgeschwindigkeit zunimmt bis schließlich die ganze Wasseroberfläche damit bedeckt ist und es keine regelmäßigen kleinskaligen Wellen mehr gibt.

5.2 Sichtbarmachung von Grenzschichten

Die spannende Frage ist nun, wie sich dieser Prozesses auf die Ausbildung der Grenzschichten auswirkt. Bisher wurden diese nur als eindimensionaler vertikaler Transportprozess betrachtet (Abb. 1a). Es bedarf Verfahren zur Sichtbarmachung der Grenzschichtdicke oder der Konzentrationen in der Grenzschicht, um die Auswirkungen zu sehen und weitere Prozesse zu finden, die man nicht schon an den winderzeugten Wellen erkennen kann.

Ein Verfahren stammt aus den Biowissenschaften und wurde benutzt, um den Sauerstoffgehalt in Zellen zu bestimmen. Die Fluoreszenzintensität gewisser Farbstoffe kann durch im Wasser gelösten Sauerstoff vermindert werden. Diese Methode wurde eingesetzt, um konvektionsinduzierten Austausch ohne Wind zu untersuchen (Abb. 5). Dazu wurde ein flacher Tank mit sauerstofffreiem Wasser und einem geeigneten Fluoreszenzfarbstoff von unten mit einer konstanten Wärmestromdichte erhitzt und von oben betrachtet.

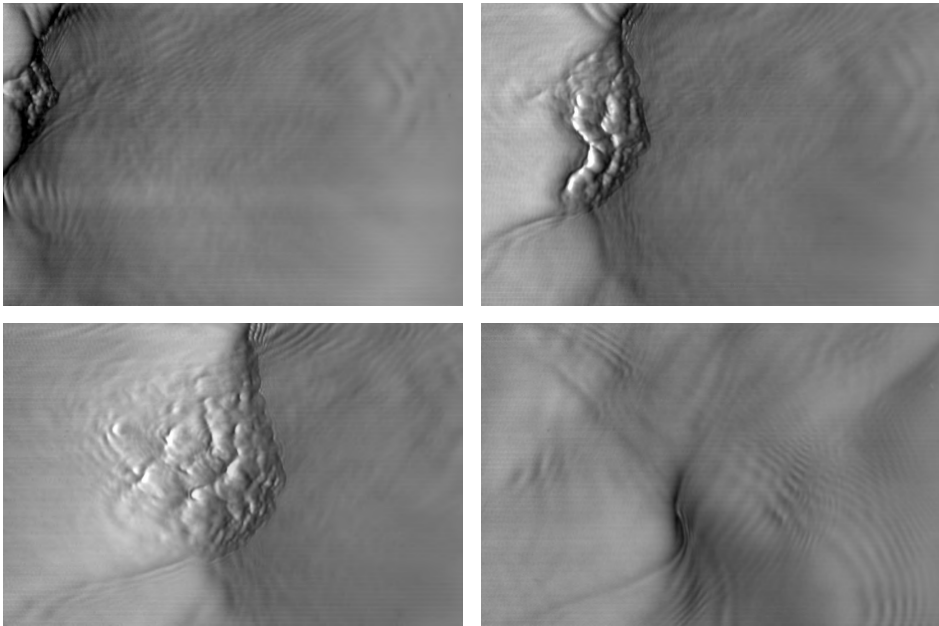


Abbildung 4: Entwicklung einer kleinen brechenden Welle ohne Blaseneintrag („micro-scale breaking“) bei nur 3,6 m/s Windgeschwindigkeit im 100 m Wind-Wellen-Kanal von Delft Hydraulics am 05.10.1990. Bildausschnitt 33 cm \times 25 cm, zeitlicher Abstand 0,1 s. Das Bild rechts unten wurde bei nur 2,5 m/s Windgeschwindigkeit aufgenommen. Alle Bilder zeigen als Grauwert die Wellenneigung in Windrichtung. Der Wind weht von links nach rechts.

Entlang schmaler Linien an der Oberfläche taucht abgekühltes Wasser, das Sauerstoff enthält und deswegen dunkel ist, nach unten ab, während dazwischen sauerstofffreies, wärmeres Wasser gleichmäßig nach oben strömt. Die gleichen Muster sieht man in Wärmebildern von der Wasseroberfläche (Abb. 5, rechts). Wegen der etwa hundertmal höheren Diffusionskonstante von Wärme in Wasser gegenüber gelöstem Sauerstoff sind die Muster viel verwaschener.

Die Fluoreszenzmethode mit Sauerstoff eignet sich auch sehr gut, um hochaufgelöste vertikale Sauerstoffkonzentrationsprofile bis an die Wasseroberfläche in einem kleinen Wind-Wasser-Kanal aufzunehmen. Dabei wurde Sauerstoff in sauerstoffärmeres Wasser transportiert. Die Zeitserien in Abb. 6 zeigen eine ganze Reihe von interessanten Eigenschaften:

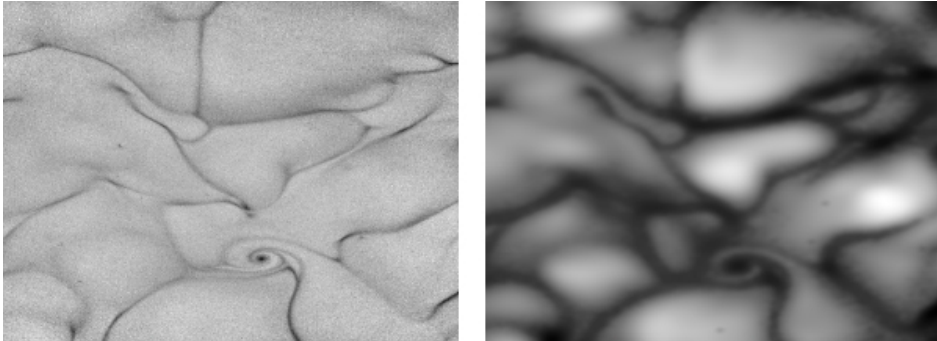


Abbildung 5: Strukturen beim konvektionsinduzierten Austausch über die Wasseroberfläche: links Sauerstoff, rechts Wärme (Bachelorarbeit Jakob Kunz, 2011).

- Direkter Nachweis, dass der Hauptwiderstand für den Transport in der dünnen Massengrenzschicht an der Wasseroberfläche liegt. Mit steigendem Wind wird die Massengrenzschicht deutlich dünner.
- In dem Wechselspiel zwischen Eindringen von Sauerstoff durch gleichmäßige molekulare Diffusion an der Wasseroberfläche und Transport in die Tiefe durch die turbulente Strömungen treten starke Fluktuationen auf.
- Die Zeitskalen für die Fluktuationen liegen, wie mit Gl. (1.1) abgeschätzt im Sekundenbereich.

Räumliche Strukturen lassen sich mit dieser Methode leider nicht bestimmen. Dazu bedarf es einer anderen Methode. Deren Grundidee ist einfach: Ein alkalisches Gas, beispielsweise Ammoniak, wird in geringer Konzentration in das Aeolotron gegeben, das Wasser im Kanal angesäuert und ein Indikator hinzugegeben, der nur im alkalischen Bereich fluoresziert. Aufgrund der Absorption des alkalischen Gases Ammoniak wird ein von dessen Konzentration in der Luft abhängiger Teil der Grenzschicht alkalisch. Infolgedessen leuchtet eine dünne Schicht an der Wasseroberfläche mit einer Helligkeit, die proportional zur momentanen Grenzschichtdicke ist.

Nach vielen Versuchen ist es schließlich gelungen, dieses Verfahren mit einer hohen Bildqualität und bis zu 100 Bildern pro Sekunde zeitliche Auflösung am Aeolotron zu realisieren. Das schwache Fluoreszenzlicht konnte nur in Kombination mit einem neuen Fluoreszenzindikator, einer sehr hellen gepulsten LED-

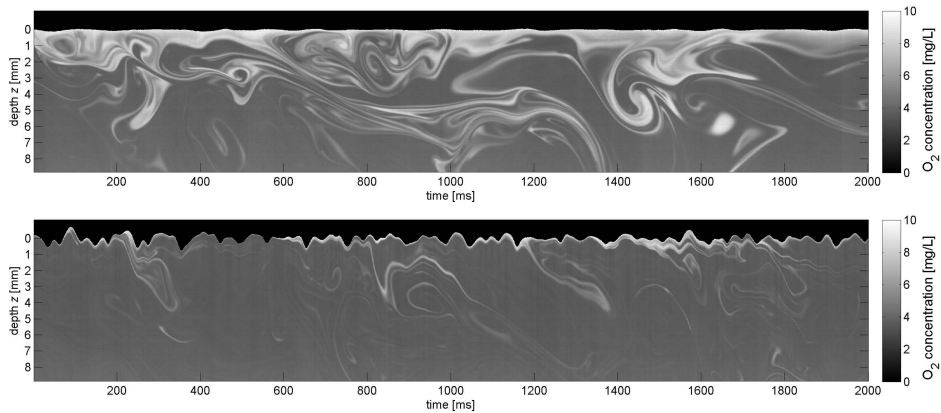


Abbildung 6: Zeitserien vertikaler Sauerstoffprofile unmittelbar unter der Wasseroberfläche in einem kleinen linearen Wind-Wellen-Kanal bei 2,8 m/s (oben) und bei 4,3 m/s (Dissertation Felix Friedl, 2013).

Lichtquelle und mit hochempfindlichen Kameras gemessen werden. Bisher ist es nur am Aeolotron als weltweit einzigem großen Wind-Wellen-Kanal gelungen, die Massengrenzschicht sichtbar zu machen.

In den Grenzschichtbildern zeigen sich neue Strukturen, die in den Wellenbildern nicht zu sehen waren (Abb. 7). Parallel zum Wind bilden sich Streifenstrukturen aus, die ein klarer Hinweis sind auf ausgedehnte kleinskalige turbulente Wirbelstrukturen mit Achsen parallel zur Windgeschwindigkeit und Abständen von wenigen Zentimetern. Wenn eine Welle kleinskalig bricht (eingekreiste Bereiche in Abb. 7), dann werden diese Strukturen durch die intensive Turbulenz zerstört und in die Tiefe gespült und müssen sich erst langsam wieder aufbauen, nachdem diese Turbulenz abgeklungen ist. Das sieht man auch gut in dem Video zu Abb. 8 links.

Spuren oberflächenaktiver Stoffe unterbinden nicht generell die Wellenbildung. Die größeren Wellen sind nach wie vor vorhanden, lediglich die kleinskaligen Kapillarwellen sind verschwunden. Dadurch wirken die Wellen glatt und gleichzeitig unterbinden die oberflächenaktiven Substanzen effektiv das kleinskalige Wellenbrechen (Abb. 8, rechts) und erniedrigen damit auch die Transfargeschwindigkeit für Gase.

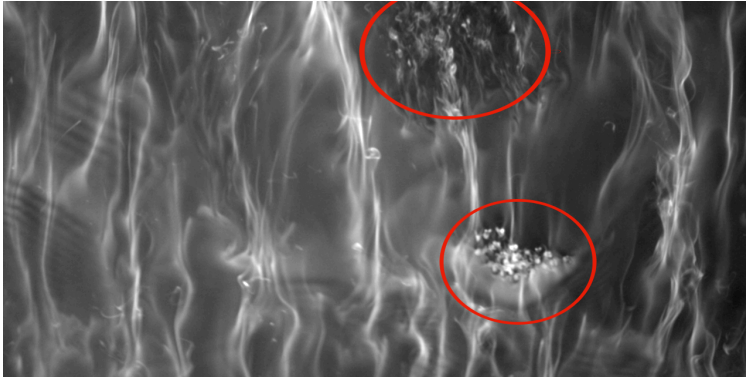


Abbildung 7: Massengrenzschichtbilder bei niedriger Windgeschwindigkeit mit zwei kleinskalig brechenden Wellen (rot eingekreist); Bildausschnitt $30\text{ cm} \times 15\text{ cm}$, Belichtungszeit: 1 ms, Windrichtung von oben nach unten (Dissertation Darya Trofimova, 2015).

Von großem Interesse ist schließlich die Frage, ob der Wärmetransfer im Wasser mit der bis zu zehnmal dickeren Grenzschicht (Abb. 1a) den gleichen Mechanismen unterliegt und damit die gleichen Strukturen aufweist. Das lässt sich mit einer Technik untersuchen, die aktive Thermographie genannt wird. Man zwingt dem System dadurch einen fest vorgegebenen Wärmefluss auf, dass man mit einem CO_2 -Laser großflächig direkt die Wasseroberfläche erwärmt und dann die resultierende Temperaturerhöhung mit einer Wärmebildkamera misst. Da man die Wärmeflussdichte j kennt, kann aus der Temperaturerhöhung nach Gl. (1.1) direkt die Transfergeschwindigkeit berechnet werden.

Im direkten Vergleich von simultan aufgenommenen Grenzschichtdicken- und Thermografiebildern sind leichte Unterschiede zu erkennen (Abb. 9). Die Thermografiebilder sind wegen der hundertmal höheren Diffusionskonstante für Wärme deutlich verwaschener. Die Bilder zeigen außerdem nicht das Gleiche. In den Grenzschichtbildern ist die Helligkeit proportional zur momentanen Grenzschichtdicke, die Thermografiebilder zeigen die Temperatur direkt an der Wasseroberfläche. Dennoch ist unverkennbar, dass die grundlegenden Strukturen die gleichen sind und damit Wärme- und Stoffaustausch trotz der großen Unterschiede in den Diffusionskonstanten den gleichen Mechanismen unterliegen.

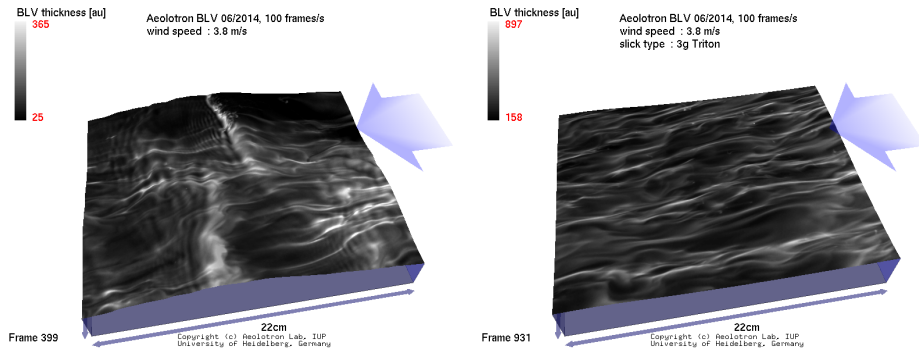


Abbildung 8: Bilder der Dicke der Massengrenzschicht projiziert auf gleichzeitig aufgenommene Wellendaten bei 3,8 m/s Windgeschwindigkeit im Aeolotron mit *sauberer Oberfläche* (links) und mit $0.26 \mu\text{mol/L}$ Triton X-100. Die Konzentration der oberflächenaktiven Chemikalie war so gering, dass die Erniedrigung der Oberflächenspannung unwesentlich ist. Belichtungszeit: 1 ms (Dissertation Christine Kräuter, 2015).

6 Auflösung der Grenzfläche bei Hurrikan-Windstärke

Erst seit kurzen ist der Bereich extrem hoher Windgeschwindigkeiten, wie sie in Hurrikanen und anderen tropischen Wirbelstürmen auftreten, Experimenten in Laboreinrichtungen zugänglich. Meine Arbeitsgruppe hatte die Gelegenheit, als Erste umfangreiche Gasaustauschmessungen mit vielen Gasen gleichzeitig in zwei Hochgeschwindigkeits-Wind-Wasser-Kanälen durchzuführen: 2015 an der Universität Kyoto und 2017 an der Rosenstiel School of Marine and Atmospheric Science (RSMAS) der Universität Miami (SUSTAIN Facility).

Die gängige Meinung war, dass bei hohen Windgeschwindigkeiten der Gasaustausch durch Blasen dominiert wird, die durch brechende Wellen ins Wasser geschlagen werden. Blasen-induzierter Gasaustausch ist für Gase mit moderater Löslichkeit wenig bedeutend. Das liegt daran, dass die große Menge im Wasser gelösten Gases um die Blase bewirkt, dass diese schnell im Gleichgewicht mit dem umgebenden Wasser ist und dann eine tote Oberfläche bildet, an der kein Austausch mehr stattfindet. Daher wurde erwartet, dass die Transfergeschwindigkeit von niedrig löslichen Gasen wie Helium, Schwefelhexafluorid und Sauerstoff stark, während moderat lösliche Gase wie Kohlendioxid und vor allem Dimethylsulfid nur langsam mit dem Wind ansteigen sollten.

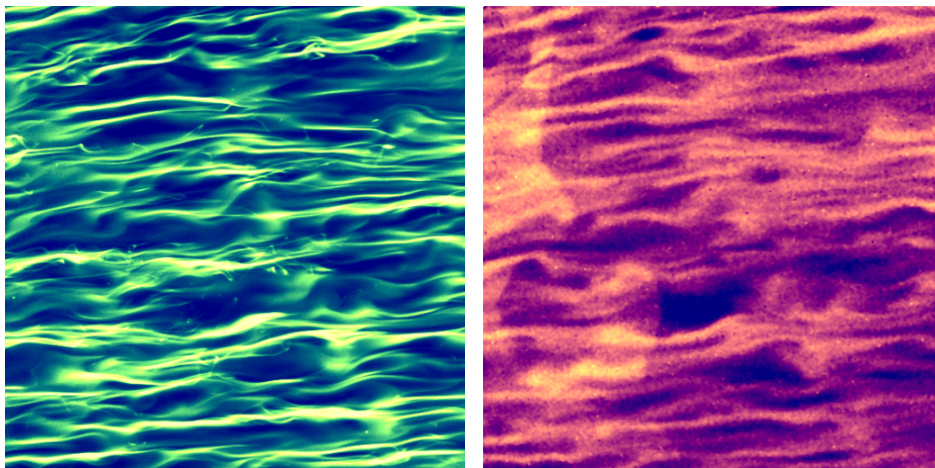


Abbildung 9: Simultan aufgenommene Grenzschicht- (links) und Thermographiebilder (rechts) zeigen die gleichen Strukturen (Dissertation Angelika Klein, 2019).

Unsere Messungen ergaben, dass die Transfergeschwindigkeiten *aller* Gase ab etwa 36 m/s Windgeschwindigkeit stärker als kubisch mit der Windgeschwindigkeit anstiegen, die leicht löslichen in Meerwasser noch etwas stärker. Damit ist offensichtlich, dass sich mit diesen hohen Windgeschwindigkeiten ein neues Regime ausbildet. Ein Blick in die Kanäle bei diesen hohen Windgeschwindigkeiten macht deutlich, was passiert (Abb. 10). Der starke Wind bläst die Wellenkämme weg und eine Schicht mit dichtem Spray bildet sich über der Wasseroberfläche. Die Wasseroberfläche wird durch eine mehr oder weniger ausgedehnte chaotische Zwei-Phasen-Strömung mit einem Gemisch aus Luft und Wasser ersetzt. Wellenkämme werden zerrissen, schnelle Tropfen schlagen im Wasser ein und platzende Blasen zerreißen die Wasseroberfläche (Abb. 11). Damit vergrößert sich die Austauschfläche signifikant und die turbulente Durchmischung der Oberfläche wird deutlich stärker. Welcher der beiden Effekte wie viel beiträgt, bedarf noch weiterer Untersuchungen.

7 Fazit und Ausblick

Es liegt essentiell in der Natur der turbulenten Strömung an Grenzflächen, dem Prandtlschen Mischungsweg-Konzept, das besagt, dass turbulente Fluktuationen



Abbildung 10: Wind-Wellen-Kanäle für hohe Windgeschwindigkeiten bis zu Hurrikanwindstärke (> 200 km/h): Universität Kyoto (links) und *SUSTAIN*, RSMAS, Universität Miami (rechts). Videos vom Anfahren des Windes: *Kyoto* und *SUSTAIN*.

Impuls, Wärme und Stoff bis an die Grenzfläche transportieren kann, und der Nichtlinearität winderzeugter Wasserwellen, dass die gezeigten vielfältigen Strukturen für die kleinskaligen Austauschprozesse an der Meeresoberfläche entstehen.

Es bleibt noch einiges zu tun, diese Strukturen besser zu verstehen und daraus ein physikalisch-fundiertes Modell für die Transfargeschwindigkeiten an der Ozeanoberfläche zu entwickeln. Experimentell ist noch die dreidimensionale Sichtbarmachung der Grenzschichten zu lösen. Hier verfolgen wir neuartige lichtfeld-basierte Ansätze, die am Heidelberg Collaboratory for Image Processing (HCI) des IWR entwickelt wurden.

Und schließlich könnte sich der Kreis von Prozessen zu Strukturen wieder zurück von Strukturen zu den Prozessen schließen mit einem neuen Ansatz, den wir zur Zeit verfolgen. Wärmebildkameras können, sobald es einen Netto-Wärmefluss über die Ozeanoberfläche oder andere Gewässerflächen gibt, ohne weitere Hilfsmittel gesehen werden. Abb. 9 zeigt, dass die Strukturen der Wärmebilder denen der nicht im Feld aufzunehmenden Grenzschichtbilder sehr ähnlich sind. Dieser Artikel hat darüber hinaus gezeigt, wie unterschiedlich die Strukturen an der Wasseroberfläche sind, je nachdem welche Verhältnisse vorliegen. Daher sollte es möglich sein, alleine aus den raumzeitlichen Strukturen an der Wasseroberfläche eindeutig die Gastransfargeschwindigkeit zu bestimmen, zusammen mit den vorliegenden Mechanismen. Gelänge dieser neue Ansatz, dann würde ein denkbar einfaches Instrument zur Verfügung stehen, mit dem sich der Gasaustausch lokal

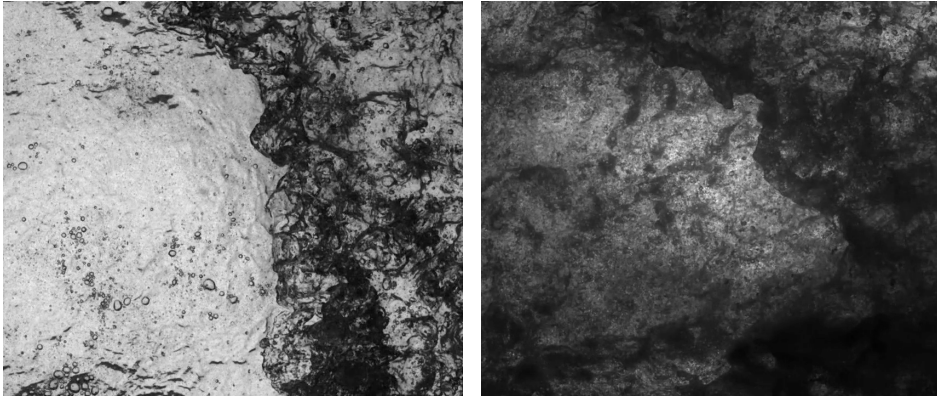


Abbildung 11: Bilder von der Wasseroberfläche von oben betrachtet mit einer gleichmäßig ausgeleuchteten Lichtquelle unter der Wasserrinne. Aufnahmen 10 000 Bildern pro Sekunde von der RSMAS SUSTAIN Facility bei 36 m/s (links) und 66 m/s (rechts) Windgeschwindigkeit. Bildausschnitt ca. $28\text{ cm} \times 24\text{ cm}$. Die Videos zeigen einen Zeitausschnitt von nur 0.36 s .

und instantan im Feld messen ließe. Dann könnten mit einer Messfahrt mehr Daten gewonnen werden als mit allen bisherigen Feldmessungen zusammen.

Generell bietet das Heidelberger Aeolotron zusammen mit den Bildverarbeitungsaktivitäten am HCI des IWR die idealen Voraussetzungen, die gesetzten Ziele zu erreichen.

8 Danksagung

Die gezeigten Resultate sind Ergebnis interdisziplinärer Teamarbeit innerhalb meiner Arbeitsgruppen am Institut für Umwelphysik und dem HCI, in der über die Jahre so viele Bachelor- und Masterstudenten, Diplomanden, Doktoranden und Postdoktoranden mitgearbeitet haben, dass ich sie hier nicht alle persönlich aufzählen kann. Ihre Namen sind dem Publikationsverzeichnis meiner Arbeitsgruppe zu entnehmen und allen gilt mein herzlicher Dank.

Ebenso möchte ich den vielen externen Forschungsgruppen danken, deren Einrichtungen wir im Laufe von vierzig Jahren in enger Zusammenarbeit nutzen konnten. Hier seien nur die wichtigsten genannt: die Universität Marseille, das Institut für Meereskunde der Universität Hamburg, Delft Hydraulics Niederlande, das Scripps Institution of Oceanography der Universität von Kalifornien in

San Diego, die Universität Kyoto und die Rosenstiel School of Marine and Atmospheric Science der Universität Miami.

Schließlich möchte ich nicht unerwähnt lassen, wie ich als Umweltphysiker Anfang der 1980er Jahren zur Bildverarbeitung kam. Ich habe dabei eine Erfahrung gemacht, von der viele andere Wissenschaftler auch berichten, nämlich dass sie entscheidende Fortschritte für ihre Arbeit Träumen zu verdanken haben. Ich arbeitete damals an einer neuen optischen Technik, um die Neigung von Wasserwellen durch Brechung eines Laserstrahls an der Wasseroberfläche an einem Punkt zu messen, und war von der Technik sehr überzeugt. Dann träumte ich zu meinem großen Erstaunen, „dass diese Methode nur unscharfe Resultate liefert und ich daran ging Bilder aufzunehmen.“ Die Korrektheit dieser Traumaussage bestätigte die weitere Analyse des Laserverfahrens und ich begann daraufhin mich mit Bildverarbeitung zu beschäftigen und konnte dann die hier beschriebene bildaufnehmende Messtechnik für Wellen entwickeln. Und dies war für mich der Ausgangspunkt, nach meiner Habilitation 1985 Forschung und Lehre in der Bildverarbeitung an der Universität Heidelberg aufzubauen.

Literatur

- Friedl, F. (2013). Investigating the Transfer of Oxygen at the Wavy Air-Water Interface under Wind-Induced Turbulence. *Dissertation, Fakultät für Physik und Astronomie, Universität Heidelberg*, doi:10.11588/heidok.00014582.
- Jähne, B. (1980). Zur Parametrisierung des Gasaustauschs mit Hilfe von Laborexperimenten. *Dissertation, Fakultät für Physik und Astronomie, Universität Heidelberg*, doi:10.11588/heidok.00016796.
- Jähne, B. (2019). Air-Sea Gas Exchange, in *Encyclopedia of Ocean Sciences*, J. K. Cochran, H. J. Bokuniewicz, and P. L. Yager (eds.), Academic Press, Vol. 6, 1–13, doi:10.1016/B978-0-12-409548-9.11613-6.
- Jähne, B., Huber, W. A., Dutzi, A., Wais, T. and Ilmberger, J. (1984). Wind/wave-tunnel experiments on the Schmidt number and wave field dependence of air-water gas exchange, in *Gas transfer at water surfaces*. Reidel. 303–309, doi:10.1007/978-94-017-1660-4_28.
- Klein, A. (2019). The Fetch Dependency of Small-Scale Air-Sea Interaction Processes at Low to Moderate Wind Speeds. *Dissertation, Fakultät für Physik und Astronomie, Universität Heidelberg*, doi:10.11588/heidok.00026559.

- Krall, K. E., Smith, A. W., Takagaki, N. and Jähne, B. (2019). Air–sea gas exchange at wind speeds up to 85 m/s. *Ocean Science*. 15, 1783–1799, doi:10.5194/os-15-1783-2019.
- Kräuter, C. (2015). Visualization of air-water gas exchange. *Dissertation, Fakultät für Physik und Astronomie, Universität Heidelberg*, doi:10.11588/heidok.00018209.
- Kunz, J. (2011). Visualisierung der wasserseitigen Massengrenzschicht beim konvektionsgetriebenen Gasaustausch mithilfe einer Lumineszenzmethode und Thermografie. *Bachelorarbeit, Fakultät für Physik und Astronomie, Universität Heidelberg*.
- Kunz, J. and Jähne, B. (2018). Investigating small scale air-sea exchange processes via thermography, *Front. Mech. Eng.*, 4:4, doi:10.3389/fmech.2018.00004.
- Trofimova, D. (2015). Towards Four Dimensional Visualization of Air-Water Gas Exchange. *Dissertation, Fakultät für Physik und Astronomie, Universität Heidelberg*, doi:10.11588/heidok.00019066.
- Eine Sammlung von Arbeiten, Berichten, Präsentationen und Videos der Arbeitsgruppe findet sich auch auf der *Small-Scale Air-Sea Interaction Community* des CERN, <https://zenodo.org/communities/asi>

Über den Autor

Bernd Jähne studierte Physik in Saarbrücken und Heidelberg, promovierte am Institut für Umweltphysik der Universität Heidelberg 1980 und habilitierte sich dort 1985 in Physik und 1992 in angewandter Informatik an der TU Hamburg-Harburg. Von 1988 bis 2000 war er Forschungsprofessor am Scripps Institut für Ozeanographie der Universität von Kalifornien in San Diego, von 1994 – 2018 Professor am Interdisziplinären Zentrum für Wissenschaftliches Rechnen (IWR) und dem Institut für Umweltphysik der Universität Heidelberg und ist seitdem dort Seniorprofessor. Er leitet seit 1995 das Heidelberger Bildverarbeitungsforum, koordinierte von 2008 bis 2017 das Heidelberg Collaboratory for Image Processing (HCI) am IWR, ein „Industry on Campus“ Projekt der Universität Heidelberg, und ist seit 2015 ehrenamtlich einer der Direktoren der European Machine Vision Association (EMVA). Seine Forschungsinteressen umfassen kleinskalige Austauschvorgänge zwischen Atmosphäre und Meer, neuartige Bildaufnahmesysteme insbesondere mit Lichtfeldern, Bildverarbeitung und deren Anwendungen in Wissenschaft und Industrie.

Korrespondenz:

Prof. Dr. Bernd Jähne

HCI am IWR

Universität Heidelberg

Berliner Straße 43

69120 Heidelberg

E-Mail: bernd.jaehne@uni-heidelberg.de

Homepage: <https://hci.iwr.uni-heidelberg.de/dip>,

<https://www.iup.uni-heidelberg.de/institut/mitarbeiter/bernd-jaehne>

Entwicklung einer Pandemie: Psychologische Aspekte der Corona-Krise

JOACHIM FUNKE

Psychologisches Institut, Universität Heidelberg

Zusammenfassung

Der Ausbruch der Corona-Krise zu Beginn des Jahres 2020 hat die gesamte Welt tiefgreifend beeinflusst. Volkswirtschaften wurden in den „lockdown“ versetzt, die Bevölkerung in ihren bürgerlichen Freiheiten beschränkt, zahlreiche Gesundheitssysteme von Nationalstaaten standen am Rande des Zusammenbruchs. In diesem Beitrag geht es um die verschiedenen psychologischen Aspekte bei der Entwicklung der Corona-Krise, wobei zum derzeitigen Stand vor allem Fragen formuliert und weniger Antworten darauf gegeben werden können. Deutlich wird die breite Beteiligung psychosozialer Faktoren am Infektionsgeschehen und an der Bekämpfung der Pandemie. Offen bleibt die Frage, inwiefern psychologische Erkenntnisse robust genug für evidenzbasierte Politikberatung sind. Zumindest basale (theoretisch fundierte) Konzepte erweisen sich als tauglich.

1 Einführung

Die aktuelle Krisensituation der Covid-19-Pandemie (am 12.3.2020 von der Weltgesundheitsbehörde WHO als solche ausgerufen) ist für die meisten Menschen belastend. Als Wissenschaftler betrachte ich das ablaufende Geschehen immer auch aus meiner psychologischen Forscher-Brille. Dabei gibt es im aktuellen

Pandemie-Geschehen viele Bezüge zur Psychologie, denn es geht ja um menschliches Denken, Handeln und Fühlen. Und es geht um das Thema unseres Jahrbuchs: Entwicklung. Ich möchte dieses Thema zunächst einmal aus meinem Spezialgebiet heraus betrachten, der Perspektive des Problemlösens.¹ Aber selbstverständlich werde ich auch versuchen, die gesamte Bandbreite psychologischer Bezüge zu Corona-relevanten Themen aufzuzeigen (und da gibt es viele, wie man sehen wird).

Was sagt man als Problemlöse-Forscher zu dieser problematischen Pandemie-Situation? Es ist nicht die erste Pandemie, die die Menschheit überkommt (eine kurze Geschichte der Pandemien findet man bei McMillen, 2016), dennoch haben wir damit ein riesiges Problem. Fangen wir vorne an. Zunächst einmal geht es darum zu verstehen, was ein Problem zu einem Problem macht. Dabei wird rasch deutlich, dass es sinnvoll ist, verschiedene Arten von Problemen zu unterscheiden. Hier begnügen wir uns mit der Abgrenzung einfacher (normaler) Probleme von sogenannten komplexen Problemen.

Ein „normales“ Problem ist dadurch definiert, dass wir ein Ziel haben und nicht genau wissen, wie wir es erreichen können. Es besteht z. B. Unklarheit über die zur Zielerreichung einsetzbaren Mittel, wenn man einkaufen möchte und hört, dass die Straßenbahnen und Busse nicht mehr fahren. Für „normale“ Probleme haben wir in aller Regel Ideen („Mittel“ zum Zweck der Zielerreichung) zu dessen Lösung (Funke, 2003), im konkreten Fall etwa: zu Fuß gehen, Fahrrad nutzen, sich mit autobesitzenden Nachbarn absprechen, oder Lieferdienste einschalten. Die typischen Phasen eines normalen Problems verlaufen in mehreren Schritten: Erkennen eines Problems – Erzeugung eines mentalen Abbilds – Wissen über das Problem aktivieren – Entwickeln einer Lösungsstrategie – Bereitstellung von

¹ Zahlreiche andere Corona-relevante Hilfestellungen der Psychologie finden sich z. B. hier: American Psychological Association, APA, <https://www.apa.org/topics/covid-19>; American Psychological Science, APS, <https://www.psychologicalscience.org/covid-19-information>; Berufsverband Deutscher Psychologinnen und Psychologen, BDP, <https://www.bdp-verband.de/aktuelles/2020/corona/aktuelle-informationen-zum-coronavirus-sars-cov-2.html>; neueste Forschung zu Corona allgemein: arXiv, <https://arxiv.org/search/?query=psychology>; und ZPID mit COSMO-Monitoring, <https://www.psycharchives.org/handle/20.500.12034/2397>. Empfehlenswert ist auch das Cochrane-Netzwerk: <https://www.cochrane.org/coronavirus-covid-19-cochrane-resources-and-news>

mentalenen Ressourcen zur Lösung – Überwachen der Zielannäherung – Bewerten der Lösungsqualität (nach Pretz et al., 2003, p. 3).

„Komplexe“ Probleme sind dagegen strukturell (von ihrem Aufbau her gesehen) und prozessual (vom ablaufenden Prozess her gesehen) von anderer Art: Schon die zu erreichenden Ziele sind alles andere als klar definiert (Was bedeutet etwa „Abflachen der Infektionskurve“ genau? Wann ist die Kapazität unseres Gesundheitssystems ausreichend zur Behandlung aller schweren Infektionsfälle? Wie wichtig ist soziale Distanzierung im Vergleich zur Sicherstellung einer elementaren Wirtschaft? Woran erkenne ich eine Verbesserung der Infektionslage? usw.). Die zur Lösung in Frage kommenden Mittel sind unbekannt, man muss sich zunächst ein Bild der Lage verschaffen. Lösungsrelevantes Wissen muss erst einmal gesammelt werden. Eine komplexe Problemsituation wie diese zwingt uns zum Nachdenken über das für diese Situation nicht mehr ausreichende Routine-Handeln: Fast sämtliche Routinen des täglichen Lebens werden in Frage gestellt, neue Lösungen für die anstehenden Probleme werden dringend gesucht.

Aus psychologischer Sicht erkennt man Einflussfaktoren auf Seiten der *Situation* (z. B. die anfänglich erschreckend hohe Ausbreitungsdynamik des Virus), zum anderen solche auf Seiten des *Individuums* / der *Person* (z. B. das sehr unterschiedliche Aushalten von Ungewissheit). Diese Aufteilung entspricht der Feldtheorie von Kurt Lewin (1935), wonach Verhalten aus der Wechselwirkung von Person und Umgebung hervorgeht. Im nachfolgenden Teil werden zunächst die situativen Merkmale beleuchtet, bevor es dann mit personalen Merkmalen weitergeht.

2 Situationsfaktoren: Eigenschaften eines komplexen Problems

Komplexe Probleme (alternative Bezeichnungen: „ill-defined problems“, „wicked problems“, „clumsy problems“; Hartmann, 2012; Simon, 1973; Verweij & Thompson, 2006) sind durch fünf Merkmale charakterisiert: Komplexität, Vernetztheit, Dynamik, Intransparenz und Vielzieligkeit („Polytelie“). Die Corona-Krise weist alle diese Merkmale eines komplexen Problems auf. Sie werden nachfolgend genauer geschildert.

- (1) *Komplexität*: Die Komplexität des ablaufenden Geschehens ist hoch – sehr viele Baustellen müssen gleichzeitig im Blick behalten werden. Neben der weltweit bedrohten Gesundheit sehen wir eine bedrohte Weltwirtschaft.

Aber hinter diesen globalen Großproblemen verbergen sich Tausende Detailprobleme. Das überfordert unsere begrenzte menschliche Verarbeitungskapazität. Was ist die Konsequenz? Vereinfachung! Gerade Verschwörungsmymthen², die in diesen Zeiten Hochkonjunktur haben, bieten wunderbare Vereinfachungen an, mal ist Bill Gates der Schuldige, mal die chinesische Staatsregierung (zum Umgang mit Verschwörungsmymthen ist das „debunking handbook“ sehr zu empfehlen: Lewandowsky & Cook, 2020). Die vom Nobelpreisträger Herbert Simon postulierte „bounded rationality“ (siehe z. B. Gigerenzer & Selten, 2001) führt auch dazu, dass wir nicht viele komplexe Probleme *gleichzeitig* im Blick behalten können, worauf „Fridays for Future“ mit ihrer Aktion #FightEveryCrisis am 24.4.2020 hingewiesen haben. Dörner und Schaub (1995) beschreiben diesen Sachverhalt auch als „Überwertigkeit des aktuellen Motivs“.

- (2) *Vernetztheit*: Früher sagte man leichthin „Was kümmert es mich, wenn in China ein Sack Reis umfällt“ – heute weiß man, dass uns das etwas angehen könnte. Der berühmte Flügelschlag eines Schmetterlings, der einen Tornado auslöst, oder das Sandkorn, das ein Erdbeben in Gang setzt (so der Titel des lesenswerten Buches von Buchanan, 2001), zeigen die Bedeutsamkeit scheinbar weit entfernt liegender Ereignisse für unseren unmittelbaren Alltag. Die Vernetztheit unserer globalen Welt zeigt Nebenwirkungen an Stellen, an denen wir nicht damit gerechnet haben. Ein schönes Beispiel für Vernetztheit liefert die Corona-Infektionsquelle im Südtiroler Skiort Ischgl, von wo aus das Virus vermutlich durch Skiurlauber in verschiedene Orte in der BRD verschleppt wurde, aber auch in Island auftauchte.³
- (3) *Dynamik*: „Wir haben es mit einem dynamischen Ausbruchsgeschehen zu tun“, sagte Kanzlerin Merkel am 12.3.2020 auf einer Pressekonferenz - gemeint ist damit, dass sich die Situation (und damit auch die Lage-Bewertung) rasch ändern kann. Ständig wird nachjustiert. Im Hintergrund läuft nämlich ein nicht-linearer Prozess der Virenverbreitung ab, der unser Vorstellungsver-

² Der häufig benutzte Begriff „Verschwörungstheorien“ suggeriert, dass eine Widerlegung möglich sei. Dem ist nicht so.

³ <https://www.welt.de/wissenschaft/article206592389/Coronavirus-Wie-Skiort-Ischgl-zur-europaeischen-Brutstaette-wurde.html>

mögen überschreitet. Das berühmte Reiskorn, das man auf den 64 Feldern eines Schachbretts von Feld zu Feld verdoppeln soll: es fängt mit 1-2-4-8-16 ganz harmlos an, aber schon bei 64-128-256-512-1024 merken wir, dass die Zahlen rapide wachsen. Wenn wir also am Tag X 12.000 Infizierte haben, sind das in X plus vier weiteren Tagen (12 - 24 - 48 - 96) schon 96.000 und nach weiteren vier Tagen 1.5 Mio (192 - 384 - 768 - 1.536). Was für eine Dynamik! Natürlich hat dieses exponentielle Wachstum eine Obergrenze und ist irgendwann vorbei...⁴

Das aus derartigen Überlegungen resultierende Programm „flattening the curve“ (siehe Abbildung 1) versucht, das ab einem gewissen Zeitpunkt explosive Wachstum zu dämpfen und damit die massiven Belastungen des Gesundheitssystems bewältigbar zu machen. In Ländern wie Italien, Spanien oder auch den USA ist dies misslungen und der Kollaps des Gesundheitssystems hat zusätzliche Menschenleben gekostet.

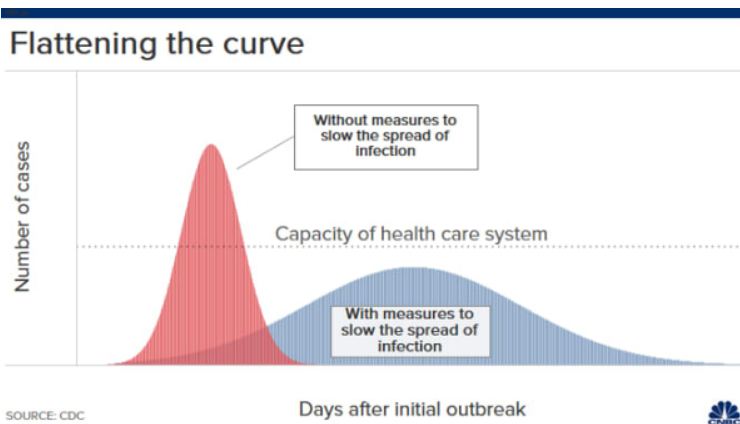


Abbildung 1: „Flattening the curve“ (Source: CDC; <https://www.cnbc.com/2020/03/22/these-charts-show-how-fast-coronavirus-cases-are-spreading.html>).

⁴ Eine logistische Funktion (genauer gesagt: eine Sigmoidfunktion) beschreibt diesen Prozess angemessener, da sie Sättigungsprozesse miteinbezieht (wieviele Personen können überhaupt noch angesteckt werden?).

- (4) *Intransparenz*: Neuartige Situationen wie die Corona-Pandemie sind extrem undurchsichtig und werfen viele Fragen auf (Wann ist das exponentielle Wachstum vorbei? Was sind die wichtigen Indikatoren?). Diese Undurchsichtigkeit weckt einen Informationsbedarf, dessen Grenze unklar ist: Wann habe ich genügend Information, um handeln zu können? Welche Informationen sind wichtig und verlässlich, welche irreführend oder überflüssig? Einerseits ist es eine Sternstunde für die Wissenschaften, die nun gefragt sind (und Antworten geben wie z. B. hier: Leopoldina, 2020), aber natürlich auch nicht auf alle Fragen eine Antwort kennen (seriöse Info beim Robert-Koch-Institut, Berlin, oder bei der Johns Hopkins Universität, Baltimore, ML)⁵, andererseits ist es auch eine Hoch-Zeit von Halbwahrheiten und Falschmeldungen (Nocun & Lamberty, 2020).⁶ Nichts Genaueres zu wissen verursacht Angst – und Angst ist ein schlechter Ratgeber für Problemlösungen. Hieß es anfangs von Seiten der Politik, bei einer Reproduktionszahl unterhalb von Eins könne man „lockern“, wurde später auch die Zahl der Neuinfektionen herangezogen. Diese Zahl wiederum hängt ab von der Anzahl an Tests, die man durchführt. Auch die Zahl der Betten auf Intensivstationen sollte man kennen. Die diffuse Gemengelage schafft Unsicherheit durch Unklarheit – aber Transparenz und Klarheit lassen sich leider bei einem komplexen Problem nicht leicht herstellen.
- (5) *Vielzieligkeit* („Polytelie“): Auch wenn es ein halbwegs klares Oberziel gibt (die Ausbreitung des Virus verhindern und das Gesundheitssystem funktionsfähig erhalten, um Menschen gesund bleiben oder im Falle ihrer Infektion nicht sterben zu lassen), gibt es viele Neben-Ziele: die Wirtschaft, wenigstens in Grundbereichen, am Laufen halten; soziale Kontakte aufrecht-erhalten; bürgerliche Grundrechte wahren; etc. Diese Konflikte zwischen verschiedenen Zielen versuchen Politiker mit ihren Entscheidungen zu lösen (radikale Ausgangssperren versus moderate Ausgangsbeschränkung). Kompromisse sind dabei unvermeidlich, Grundrechte können manchmal unlösbar aufeinanderprallen. Klare Wertorientierungen sind bei polytelischen Problemlagen sehr hilfreich, weil sie helfen Prioritäten zu setzen. Zusätzlich

⁵ RKI: <https://www.rki.de>, JHU: <https://www.jhu.edu>.

⁶ Übrigens: MedWatch.de warnt vor unseriösen Heilsversprechen.

überlagert wird die Vielzieligkeit der Pandemie durch Nebenziele der deutschen Politiker, die etwa auf bessere Chancen bei einer Kanzlerkandidatur hoffen, wenn sie sich als souveräne Krisenmanager präsentieren können (in den USA ähnlich wegen anstehender Präsidentschaftswahlen).

Soviel zu den fünf Merkmalen eines komplexen Problems. Mehr dazu findet man z. B. bei Dörner (1980, 1996) oder auch bei Funke (2003, 2012). Die „Logik des Misslingens“ von Dietrich Dörner (1989) ist eine auch für Laien gut lesbare Einführung in diese Thematik.

In der Erforschung komplexer Probleme sind in den letzten Jahren kontroverse Diskussionen über den Erkenntniswert von computersimulierten Szenarien geführt worden (für einen Überblick siehe Dörner & Funke, 2017). Im Rahmen der internationalen Schulleistungsstudie PISA wurden in der Erhebung 2012 die Problemlöseleistungen 15jähriger Schülerinnen und Schüler mittels computersimulierter Szenarien erhoben – Deutschland und Amerika lagen dabei im Mittelfeld, asiatische Länder an der Spitze (Csapó & Funke, 2017; Dossey & Funke, 2016).

Zeitverzögerungen. Eine der unangenehmen Eigenschaften dynamischer Prozesse sind Zeitverzögerungen in ihren verschiedenen Spielarten bei der Information über die Lage des Systems (z. B. als Totzeit zwischen dem Eingriff in ein System und der darauf folgenden Reaktion: Man klickt eine App an, muss sich aber beim Ladevorgang gedulden; z. B. als verzögertes Feedback bei der Anzeige von Systemzuständen: Der Temperaturanzeiger beim Auto meldet die tatsächliche Temperatur verzögert). Derartige Verzögerungselemente in komplexen Systemen erschweren die Regelbarkeit eines Systems. Zeitverzögerungen als Eigenschaften komplexer Systeme werden hier als Stichwort separat neben den weiter oben geschilderten fünf Merkmalen komplexer Probleme aufgeführt, da sie an verschiedenen Stellen im Problemraum auftreten. Zeitverzögerungen gibt es z. B. bei der Ausbreitung des Virus (Inkubationszeit): Wer sich heute ansteckt, wird erst in 10–14 Tagen krank. Zeitverzögerungen gibt es bei der Meldung von Fallzahlen (nach Wochenenden stieg in der BRD die Fallzahl mehrfach außergewöhnlich stark an, da die Meldungen der Wochenend-Fallzahlen erst mit Verspätung eintrafen). Die Fallzahlen von heute zeigen dabei die Lage von 14 Tagen zuvor an. Politische Entscheidungen (z. B. über die Entwicklung einer *Tracing-App*) wurden früh getroffen, aber erst spät umgesetzt. Die Entwicklung eines Impfstoffes wird sehnlich erwartet, braucht aber seine Zeit. Über die nachteiligen Auswirkungen von verzögertem Feedback hat die Gruppe um den schwedischen Psychologen

Berndt Brehmer viel geforscht (z. B. Brehmer, 1992, 1995; Brehmer & Allard, 1991), daneben ist John Stermans Idee der „misperception of feedback“ (Sterman, 1989) sicher erwähnenswert.

3 Personfaktoren: Interindividuelle Unterschiede

Persönlichkeitsfaktoren spielen natürlich im Umgang mit Risiken und mit deren Bewertung eine wichtige Rolle. Bei den Ministerpräsidenten ist von „Dränglern“ (Laschet) und „Zauderern“ (Söder) die Rede – tatsächlich gibt es starke Persönlichkeitsunterschiede in Bezug auf „represser“ bzw. „sensitizer“, zwei Extrempole für den Umgang mit angstauslösenden Situationen. Während der eine Pol auf Abwehr und Vermeidung (repression) setzt, widmet der andere Pol den beängstigenden Reizen erhöhte Aufmerksamkeit und Zuwendung (detailliertere Hintergründe und Erläuterungen zum Konstrukt bei Stemmler et al., 2016, Kap. 9.2). Die Corona-Pandemie lässt gut erkennen, zu welchem Pol sich ein Individuum hingezogen fühlt.

Die „Terror Management Theory“ (mehr zu Theorie und Empirie findet man bei Hayes et al., 2010) beruht auf Vorstellungen von Ernest Becker (1973, 1975) über die Bedeutung des Todes für unsere Spezies. Sie beschreibt, wie Menschen Todesgedanken und daraus resultierende Ängste verdrängen durch (a) kulturelle Weltansichten (die Welt ist geordnet und stabil; religiöse Weltansichten behaupten ein Leben nach dem Tod) sowie durch (b) Selbstwertgefühle (man erlebt sich als bedeutsam und wichtig). So unterscheiden sich Personen darin, für wie unverletzlich sie sich halten. Eine Reihe von führenden Politikern sind dabei „ertappt“ worden, dass sie vorgeschriebene Schutzmaßnahmen selbst nicht beachten – ein Hinweis auf hohes Selbstwertgefühl und Unverletzlichkeitsannahmen?

Von grundsätzlicher Bedeutung bei den interindividuellen Unterschieden dürfte das jeweilige Ausmaß an Ängstlichkeit sein. Wie Stemmler et al. (2016, S. 494) betonen, „lehrt die Fachliteratur, dass die psychologisch-wissenschaftliche Beschäftigung mit den Phänomenen der Angst und Ängstlichkeit in den letzten Jahrzehnten eine Vielzahl verschiedener theoretischer und empirischer Zugänge zu diesem Bereich gewählt hat und dass mit den Begriffen Angst, Ängstlichkeit und verwandten Bezeichnungen eine komplexe Vielfalt von Erlebnisweisen, Verhaltenskomponenten und äußeren Bedingungen angesprochen wird, die wissenschaftlich schwer zu integrieren ist.“ Diese Vielzahl an Zugängen hat sicher

mit der großen Bedeutung dieser Emotion zu tun, die unser ganzes Leben dominieren kann und mit der typologischen Dimension „Neurotizismus“ einhergeht, immerhin eine der Big-Five-Dimensionen (1: Extraversion, 2: Verträglichkeit, 3: Gewissenhaftigkeit, 4: Emotionale Stabilität vs. Neurotizismus und 5: Offenheit für Erfahrungen), mit denen differentielle Psychologen Unterschiede wie auch Ähnlichkeiten zwischen Menschen beschreiben. Bezogen auf die Corona-Situation bedeutet diese Differenzierung: Je nach Kombination dieser Dimensionen wird das Geschehen unterschiedlich beurteilt. Gewissenhaftigkeit dürfte das Befolgen von Verordnungen beeinflussen, Neurotizismus steigert die Ängstlichkeit, Verträglichkeit bestimmt den Grad der Kooperationswilligkeit. Noch fehlen empirische Daten hierzu, aber die Erwartungen gehen in die genannten Richtungen.

4 Risikowahrnehmung und Risikomanagement

Von Politikern erwartet man, dass sie das Risiko einer Bedrohung durch das Virus frühzeitig erkennen und wirksame Gegenmaßnahmen ergreifen. Verschiedene Staatslenker (z. B. USA: Donald Trump; Brasilien: Jair Bolsonaro; Großbritannien: Boris Johnson; Ungarn: Viktor Orban; Russland: Wladimir Putin) haben anfangs versucht, die Bedrohung herunterzuspielen („it's like a flu“). Diese Strategie der Risiko-Ignoranz hat sich jedoch nicht ausgezahlt. Gefährliches Gruppendenken droht (Janis, 1982), wenn nur Berater aus einer Richtung gehört werden.

Risikomanagement: das ist ein kontinuierlicher Prozess von Planung, Umsetzung, Überwachung und Verbesserung der Risikobewertung. In Zeiten der Corona-Pandemie gehört die tägliche Überwachung pandemie-spezifischer Kennzahlen zu diesem Management. Dabei spielt der psychologische Aspekt der Risikowahrnehmung eine entscheidende Rolle. In der Bewertung von Risiken kommen typische Verzerrungen („biases“) zum Tragen, die gut untersucht sind (Gigerenzer & Brighton, 2009; Renn, 2014; Tversky et al., 1974).

5 Umgang mit Fehlern

Bei einer komplexen, intransparenten, dynamischen Situation bleibt es nicht aus, dass Fehler gemacht werden („errare humanum est“). Entscheidend ist daher die Fehlerkultur, d.h. die Beantwortung der Frage, wie mit Fehlern umgegangen

wird (siehe z. B. Mandl, 2017). Die naheliegende Strategie der Vertuschung oder Attribution auf eine andere Ursache gelingt nicht immer, besser ist ein offenes Eingeständnis verbunden mit dem Bemühen, den gemachten Fehler nicht noch einmal zu machen.

Dörner (2019) hat eine Liste von Torheiten (die sogenannte T-Liste) erstellt, die bei politischen Akteuren im Lauf der Geschichte immer wieder zu beobachten

Tabelle 1: Vereinfachung, Aktionismus und Abwehr von Kritik: Drei Klassen von Torheiten mitsamt Unterpunkten und deren Charakterisierung („T-Liste“ von Dörner, 2019; Abdruck mit Genehmigung des Autors).

| a) Vereinfachung: Konstruktion einer einfachen Weltsicht | |
|--|--|
| a1 Reduktionismus | Man macht sich ein Bild von der Welt, in dem nur wenige oder gar nur ein einziger Faktor eine zentrale Rolle spielt. |
| a2 „Rumpelstilzchenplanung“ | „Heute back’ ich, morgen brau’ ich, übermorgen hol’ ich der Königin ihr Kind!“ Man hält die Bedingungen, die für bestimmte Tätigkeiten gegeben sein müssen, für vorhanden. Denn „backen“ kann ich heute nur, wenn genügend Feuerholz da ist (und außerdem genügend Mehl, Salz, Zucker usw.). Und „brauen“ kann ich morgen nur, wenn ich nicht das Feuerholz schon für das Backen verbraucht habe. Usw. |
| a3 Vernachlässigung von Neben- und Fernwirkungen | In komplexen Realitäten sind (unerwünschte) Neben- und Fernwirkungen die Regel und nicht die Ausnahme. |
| a4 „Leerplanung“, „Ziele ohne Maßnahmen“ | Pläne allein nur durch die Festsetzung von Zielen, ohne Berücksichtigung der Aktionen. |
| a5 Übergeneralisierung | Man schließt von einem Einzelfall auf alle Fälle des gleichen Typs. |
| a6 Methodismus | Man macht das Gleiche immer wieder. Clausewitz und auch schon Platon empfehlen, in jeder Situation immer wieder neu nachzudenken. |
| a7 Lineare Extrapolation | Das Morgen ist ein Abbild des Heute, allenfalls etwas vergrößert oder verkleinert. |

Fortsetzung Tabelle 1: Vereinfachung, Aktionismus und Abwehr von Kritik: Drei Klassen von Torheiten mitsamt Unterpunkten und deren Charakterisierung („T-Liste“ von Dörner, 2019).

| | |
|--|--|
| b) Aktionismus: Handeln hauptsächlich, um sich selbst und anderen zu zeigen, dass man handlungsfähig ist | |
| b1 Ausweichen ins Machbare | Man tut das, was man machen kann und nicht das, was man machen soll. |
| b2 Die „Große Tat“ | Allein die Absicht, eine „große Tat“ zu begehen, hebt das Selbstgefühl ganz ungeheuer und damit steigt dann auch der Optimismus. Und daraus ergibt sich Selbstüberschätzung. Man traut sich etwas zu, was man sich nicht zutrauen sollte. Und meist misslingt es dann deshalb! |
| b3 Symbolisches Handeln | Man verspricht etwas für die Zukunft. Oder man beiratsucht sich am Ruhm vergangener Taten. |
| b4 Methodismus | Man tut das, was sich in der Vergangenheit bewährt hat. |
| c) Abwehr von Kritik (Einkapselung) | |
| c1 Unterlassung von Selbstkritik | Bei sich keine Fehler sehen. |
| c2 „Ballistisches Handeln“ | Die Effekte des Handelns werden nicht überprüft (ganz viele Handlungen auf einmal ausführen, so viele Kanonen wie möglich abfeuern, egal ob sie treffen). |
| c3 Wahl „loyaler“ Berater | Sich mit Gleichgesinnten umgeben, kritische Stimmen „entfernen“. |
| c4 „Gruppendenken“ (Janis, 1972) | „Wenn alle das Gleiche denken, dann muss das ja wohl richtig sein!“ |
| c5 Abwehr von Schuld | Die Schuld immer erst einmal bei anderen suchen. |
| c6 Sturheit | Man hält Sturheit für „Standhaftigkeit“; Sturheit ist unreflektierte Standhaftigkeit. |

sind (siehe auch die Sammlung historischer Fälle bei Tuchman, 1984). Die drei

Hauptkategorien an Torheiten (Vereinfachung, Aktionismus und Abwehr von Kritik) haben jeweils mehrere Unterpunkte (siehe Tabelle 1).

Ein Blick in die aktuelle Politik liefert auf Anhieb Anschauungsbeispiele für die verschiedenen Kategorien und Subkategorien. Der amerikanische Präsident liefert mit seinen Entscheidungen und mit seinen Tweets Stoff für gleich mehrere Kategorien.

Was sind Ursachen von Fehlern? Schaub (2016, S. 41f.) unterscheidet kognitive und motivationale Ursachen. Zu den kognitiven Ursachen zählt er die „begrenzte Verarbeitungskapazität des Denkens und begrenzte Kapazität des Gedächtnisses“, zu den motivationalen Ursachen dagegen „Überwertigkeit des aktuellen Motivs und Schutz des eigenen Kompetenzzempfindens“. Diese Ursachen sind so allgemein in der menschlichen Natur verankert, dass man sie kaum vermeiden kann. Weiterführende Literatur zum Thema „Fehler“ findet man bei Hofinger (2013) oder bei Reason (1990, 2000).

6 Beiträge verschiedener Teildisziplinen der Psychologie

Auch wenn in diesem Beitrag die denkpsychologischen Aspekte im Vordergrund stehen: Natürlich sind durch die Corona-Pandemie viele weitere Teildisziplinen der Psychologie angesprochen. Eine kurze Übersicht soll das verdeutlichen. Ich folge hierbei der Struktur der Fachgruppen, die die bedeutendste deutschsprachige Fachgesellschaft, die „Deutsche Gesellschaft für Psychologie“ (DGPs), ihrer Organisationstruktur zugrunde legt. Siebzehn Fachgruppen sind dort alphabetisch gelistet. Ich folge dieser Aufzählung.⁷

Allgemeine Psychologie mit ihren umfangreichen Spezialgebieten Wahrnehmung, Aufmerksamkeit, Denken, Sprache, Lernen, Gedächtnis, Motivation und Emotion. Zahlreiche Bezüge sind erkennbar. Die *Aufmerksamkeit*, mit der wir Nachrichten und neue Entwicklungen der Pandemie verfolgen: Die Spannweite reicht vom minütlichen Checken der Fallzahlen bis hin zum „homo ignorans“ (siehe dazu Hertwig & Engel, 2016). Die zum *Denken* gehörenden Urteilsprozesse, die Inferenzen, die man zieht, die Problemlöseprozesse sind bereits weiter vorne angesprochen worden. *Sprache* spielt natürlich eine zentrale Rolle (das

⁷ Auch die angeführten Beschreibungen der Aufgaben der siebzehn Fachgruppen stammen zu einem großen Teil von der Webseite <https://www.dgps.de/index.php?id=48>.

„framing“ etwa, siehe Christmann & Göhring, 2016; Lakoff & Wehling, 2012; Wehling, 2016), aber auch im Kontext der Interpretation von Vorschriften und Verordnungen. So ist z. B. der Mannheimer „Luisenpark“ nach den Anfang Mai verkündeten Lockerungen zunächst nicht geöffnet worden,⁸ da die Zuordnung zur (öffnungsfähigen) Kategorie „Park“ unklar bleibt wegen zahlreicher Freizeitanlagen auf dem Gelände, die für einen „normalen“ Park untypisch sind (z. B. Gondoletta-Boote, Trampolinanlage). Ein klassischer Fall unscharfer Begriffe (*fuzzy concepts*, siehe Zadeh, 1965). *Lernen*: Wie schnell haben wir Nies- und Hustenetikette oder regelmäßiges Händewaschen gelernt? *Gedächtnis*: Wann haben wir zum letzten Mal die Großeltern besucht? Wieviel (Halb-)Wissen über das Virus haben wir inzwischen erworben? Kenne ich die aktuelle Vorschrift über das Ausgehen in meiner Kommune? Auch Idealisierungen der Vergangenheit kommen in der Erinnerung zustande („wie toll war es vor Corona“). *Motivation*: Warum sollte man Mund-Nasenschutz tragen? Warum sollte man den einschränkenden Vorschriften folgen? Woher nimmt man die Kraft für eine weitere Videokonferenz, woher die Kraft zum Arbeiten? *Emotion*: Überkommt einen die Empörung, wenn man jemanden gegen die Vorschriften verstoßen sieht oder bleibt man gelassen? Steigt die Angst, wenn in meiner Umgebung jemand positiv getestet wurde? Wie groß ist die Traurigkeit darüber, die Enkelkinder nicht besuchen zu dürfen? Über Themen aus dem breiten Feld der Allgemeinen Psychologie findet man Vertiefendes in den einschlägigen Lehrbüchern (z. B. Müsseler & Rieger, 2017; Spada & Kiesel, 2018; Strobach & Wendt, 2019).

Arbeits-, Organisations- und Wirtschaftspsychologie. Der Bezug dieser Fachgruppe zur Pandemie dürfte unstrittig sein: Themen wie „home-office“, „Kurzarbeit“ oder „Arbeitslosigkeit“, Überlastung und „burn-out“ bei Pflegekräften und medizinischem Personal sind wichtig. Aber auch Themen wie „bedingungsloses Grundeinkommen“ (z. B. Werner et al., 2012) oder Fragen eines Systemwechsels von der Wegwerf-Wirtschaft zur Kreislaufwirtschaft (z. B. Raworth, 2018) stehen dort im Raum, da sie mit menschlichen Bedürfnissen und deren Befriedigung zu tun haben.

Biologische Psychologie und Neuropsychologie. Während sich die Biopsychologie um anatomische und physiologische Grundlagen menschlichen Verhaltens und Erlebens sowie physiologische Effekte psychologischer Prozesse kümmert,

⁸ So ein Bericht in der „Rhein-Neckar-Zeitung“ vom 4.5.2020.

beschäftigt sich die Neuropsychologie mit den neuronalen Bedingungen psychologischer Prozesse. Das bio-psycho-soziale Modell (siehe z. B. Gilbert, 2019) betont die gegenseitigen Wechselwirkungen zwischen der biologischen, psychologischen und sozialen Dimension – die Auswirkungen der Pandemie auf den Körper scheinen evident zu sein.

Differentielle Psychologie, Persönlichkeitspsychologie und psychologische Diagnostik. Die gleiche Situation der Pandemie wird von verschiedenen Individuen äußerst unterschiedlich verarbeitet. Weiter vorne im Beitrag wurde dies schon anhand der Persönlichkeitsdimensionen „Big Five“ ausgeführt. Diagnostische Instrumente zum Wissen über Corona und zum Umgang mit der Situation liegen bereits vor (siehe z. B. das Material zum „Covid-19 Snapshot Monitoring“, COSMO, von Cornelia Betsch und Mitarbeitenden an der Universität Erfurt⁹).

Entwicklungspsychologie. Wie Kinder mit einer Pandemie umgehen und wie sie sie verstehen, ist hier ebenso Thema wie die in der Gerontopsychologie diskutierte potentielle Diskriminierung älterer Menschen („Risikogruppe“). Es geht unter anderem um intergenerationale Solidarität (Ayalon et al., 2020), aber auch um Fragen des Kindeswohls (Kray, 2019; Lohaus & Vierhaus, 2015; Siegler et al., 2016).

Geschichte der Psychologie. Diese Fachgruppe, die sich mit der historischen Entwicklung der Psychologie als eigenständige Wissenschaft befasst, wird vermutlich erst in einigen Jahren auf diese Zeit zurückschauen und sich fragen, welchen Einfluss die Corona-Pandemie auf die Fachentwicklung genommen hat – wie sich abgesagte Konferenzen und Tagungen ausgewirkt haben oder welche (negativen) Effekte die Krise auf die Karrierewege junger Wissenschaftlerinnen und Wissenschaftler genommen hat, die weitgehend prekär (zeitlich befristet) arbeiten.

Gesundheitspsychologie. Das Thema „mental health“ liegt geradezu in der Luft – wie sorgt man in dieser Krisenzeit für ein gutes und glückliches Leben, wenn der Arbeitsplatz in Gefahr ist und das soziale Stützsystem einbricht? Positive Psychologie (z. B. Seligman & Csikszentmihalyi, 2000) befasst sich mit den Bedingungen für ein glückliches Leben. Die Widerstandskraft (Resilienz), die aus dem Bestehen schwieriger Herausforderungen resultiert (siehe z. B. Kruse, 2015, über Johann Sebastian Bach), gehört dazu und wird in Zeiten einer Pandemie benötigt. Gesundheitsschutz ist nicht nur eine Frage der Medizin und Hygiene, sondern auch eine

⁹ <https://projekte.uni-erfurt.de/cosmo2020/cosmo-analysis.html>

des Verhaltens: Wie vermeidet man etwa ungewollte Gesichtsberührungen mit der Hand? Das ist ein Thema der Verhaltenspsychologie, nicht der Medizin (und zugleich muss man gegen das Präventionsparadox angehen: „there is no glory in prevention“: es passiert nichts Schlimmes, *weil* man vorbeugend gehandelt hat – also erscheinen die Präventionsmaßnahmen als übertrieben).

Ingenieurpsychologie. In dieser Gruppe steht das Erleben und Verhalten von Menschen in der Interaktion mit technischen Systemen im Vordergrund. Das Krankenhauspersonal wird gerade in der Handhabung von Beatmungsgeräten geschult, an Schulen und Hochschulen werden Videosysteme eingesetzt. Der Umgang mit diesen technischen Systemen ist nicht unproblematisch und bedarf ingenieurpsychologischer Unterstützung.

Klinische Psychologie und Psychotherapie. Natürlich ist diese (größte) Teildisziplin an ganz vielen Corona-Themen beteiligt. Nicht nur die Frage, wie Psychotherapie ohne physische Begegnung digital funktionieren kann, steht im Raum, sondern auch die Auswirkungen der Corona-bedingten Beschränkungen auf das Entstehen etwa einer Depression oder einer Essstörung. Krankmachende Bedingungen wie drohende Arbeitslosigkeit, Einsamkeit, Schlaflosigkeit, gesteigerter Konsum von Suchtmitteln (Alkohol, Schokolade, Spiele), erhöhte Gewalt in Familien auf engem Raum und der Wegfall von außerhäuslichen sozialen Unterstützungssystemen sind natürlich ebenfalls Thema. Dass es nach Corona vermehrt zu posttraumatischen Belastungsstörungen (PTBS) aufgrund schwieriger Erfahrungen kommen könnte, befürchten die Fachleute (zu PTBS siehe Schäfer et al., 2019).

Medienpsychologie. Die in dieser Fachgruppe aufmerksam beobachtete Nutzung von Medien spielt in der Pandemie eine wichtige Rolle. Falschinformationen und Verschwörungsmythen verbreiten sich bevorzugt über die (unkontrollierten) sozialen Medien im World Wide Web. Aber Medien transportieren auch seriöse Informationen, die man für sein Alltagshandeln benötigt (Batinic & Appel, 2008; Trepte & Reinecke, 2013).

Methoden & Evaluation. Diese Teildisziplin ist von besonderer Bedeutung, wenn es um die Interpretation vorgelegter Daten geht. Welche Aussagekraft etwa die Studie des Virologen Hendrik Streeck (Universität Bonn) hat, der die Heinsberger Corona-Karnevalsopfer genauer untersucht hat, muss angemessen bewertet werden. Methodiker helfen dabei, voreilige Fehlschlüsse zu vermeiden. Was ist wirkliche Evidenz, was sind schöne Geschichten, was ist Unsinn?

Pädagogische Psychologie. Die Themen „Schule“ und „Kindergarten“ sind in den Pandemie-Debatten von zentraler Bedeutung. Die Pädagogische Psychologie wird hier gebraucht, um „home-schooling“ (oder eLearning in Firmen und Hochschulen) zu evaluieren und Tipps für eine gute Gestaltung dieser Lernsituationen geben können (Spinath & Brünken, 2016).

Rechtspsychologie. Die Anwendung psychologischer Theorien, Methoden und Erkenntnisse auf Fragestellungen, die sich aus der Gestaltung und Anwendung des Rechts ergeben, spielen natürlich eine Rolle bei der Akzeptanz von Einschränkungen der Grundrechte. Zahlreiche Proteste und Klagen bei Gericht zeigen, dass nicht alle Bürgerinnen und Bürger Verständnis für die (vorübergehende) Einschränkung ihrer Rechte haben. Gerechtigkeit herzustellen ist schwierig, der Glaube an eine gerechte Welt („Belief in a Just World“, siehe Kals & Maes, 2012; Lerner & Montada, 1998) wird oft enttäuscht.

Sozialpsychologie. Der soziale Kontext menschlichen Handelns, Erlebens und Fühlens steht hier im Fokus. Begriffe wie „social distancing“ (besser: „physical distancing“) machen die Bedeutung des sozialen Kontexts für das Pandemiegeschehen evident. Aber auch Phänomene wie Rücksichtnahme, Altruismus (selbstlose Hilfsbereitschaft) sowie deren Gegenteile stehen auf der Agenda (siehe z. B. Fischer et al., 2018; Forgas et al., 2012; Kessler & Fritsche, 2018). Da viele soziale Aktivitäten (Besuche von Freunden und Verwandten; Restaurantbesuche; gemeinsame Ausflüge und Urlaube) zum Erliegen gekommen sind, ist hier ein großes Feld an offenen Fragen entstanden, die auf Antwort warten.

Sportpsychologie. Es geht um menschliches Verhalten, Handeln und Erleben im Sport, in Zeiten von Corona eine gern genutzte Freizeitaktivität. Auf der anderen Seite steht die Corona-bedingte Bewegungseinschränkung, die das Ausüben bestimmter Aktivitäten verhindert (Fußball!).

Umweltpsychologie. Die Einstellungen zur Umwelt und das Umweltbewusstsein, umweltbezogenes Verhalten und die Gestaltung eines ökologisch gesunden Lebensumfeldes sind natürlich Themen in der gegenwärtigen Debatte, geht es doch darum, Gesundheit zu bewahren und Krankheit zu vermeiden. Die finanziellen Rettungsschirme für die Wirtschaft sollen natürlich auch in Richtung auf eine Abschwächung der Klimakrise gelenkt werden, ein schwieriger Spagat.

Verkehrspsychologie. Eine spannende Frage, ob und wie sich das Mobilitätsverhalten nach Corona verändert oder ob es zu Rebound- und Reaktanz-Effekten (Nachhol-Effekten) kommt. Rebound bedeutet Kompensation, z. B. erhöhte Ri-

sikobereitschaft durch bessere Sicherheitstechnik in Fahrzeugen und Anlagen. Die Einführung von ABS (Anti-Blockier-System; zu meiner Fahrschulzeit hieß das schlicht „Stotterbremse“) hat etwa den Gewinn an Sicherheit durch höhere Fahrgeschwindigkeiten wieder kompensiert. Reaktanz bedeutet, dass die Freiheit durch ein und nach einem Verbot auf andere Weise wieder zurückgeholt wird (z. B. Schwarzbrennereien während der Prohibition). In Bezug auf entfallene Urlaubsreisen könnten Rebound- und Reaktanzeffekte zu einem verstärkten Urlaubs-Bedarf („Nachholbedarf“) führen (mehr zur Verkehrspsychologie z. B. bei Vollrath & Krems, 2011).

Wie sich zeigt, gibt es kaum eine Teildisziplin der Psychologie, die nicht hinsichtlich ihrer Forschungsthemen von der Corona-Pandemie betroffen wäre. Und die Aufteilung in Fachgruppen ignoriert den Sachverhalt, dass unser Seelenleben nicht unbedingt den (überlappenden) Kategorien der Fachgesellschaft entspricht, sondern systemisch-ganzheitlich gesehen werden sollte. Zahlreiche Forschungsarbeiten liegen in der Luft (und sind auch schon angestoßen worden). Aus dieser umfangreichen Betroffenheit des Faches Psychologie resultiert zugleich eine enorme Verantwortung: Aufklärung und Beratung sind notwendige Konsequenzen.

7 Ethische Aspekte

Ethische Probleme ergeben sich an vielerlei Stellen. Hier seien ein paar Fragen benannt: Ist es vertretbar, dass ältere Menschen in ihren Bürgerrechten zum Selbstschutz stärker eingeschränkt werden als jüngere Menschen? Wenn Beatmungsmöglichkeiten knapp werden: Wer erhält die bevorzugte Behandlung? Muss man anderen Menschen zuliebe Mund-Nase-Schutz tragen, auch wenn es für das maskentragende Individuum selbst kaum Schutz bedeutet? Müssen wir Nachbarstaaten in Not mit unseren Ressourcen aushelfen (Stichwort: EU-Rettungsschirm; Stichwort: Italien-Hilfe)?

Zu ethischen Aspekten zählt auch die Frage, ob Personen in „systemrelevanten Bereichen“ adäquat bezahlt werden. Dass das Jahresgehalt des derzeitigen SAP-Vorstandsvorsitzenden Christian Klein nach Ausscheiden seiner Co-Vorsitzenden Jennifer Morgan auf über 34 Mio Euro steigern könnte,¹⁰ während eine examinierte Krankenschwester (TVöD Entgeltgruppe 7, Stufe 2: 2.830 € monatliches

¹⁰ Rhein-Neckar-Zeitung vom 23.4.2020.

Brutto) auf ein Jahresgehalt von etwa 34 T€ kommt, zeigt eine erhebliche Unterscheidung in der Wertschätzung (in diesem Fall um das Tausendfache!). Natürlich sind die Aufgabenbereiche eines SAP-Vorstands nicht mit denjenigen einer Krankenschwester zu vergleichen, aber der Faktor 1.000 in der Gehaltszahlung scheint doch übertrieben. Mutterarbeit wird immerhin einmal im Jahr am Muttertag (seit 1914 am zweiten Sonntag im Mai) öffentlich wertgeschätzt...

Die unterschiedliche Honorierung von Erwerbstätigkeit und Sorgetätigkeit (unbezahlte Haushalts-, Erziehungs- oder Pflegetätigkeiten) entspricht nur bedingt deren Systemrelevanz, wie man in Krisenzeiten merkt. Wenn es ums Überleben geht, werden plötzlich ganz andere Dinge und ganz andere Tätigkeiten relevant.

8 Ist psychologische Forschung bereit zur Politikberatung?

Wann ist Forschung so gut validiert, dass sie sich zur (gut begründbaren) Politikberatung eignet? Die Forscherinnen und Forscher der Leopoldina (2020) haben sich (wie schon erwähnt) mit ihren Ratschlägen an die Politik aus dem Fenster des Elfenbeinturms gelehnt und geben konkrete Empfehlungen wie z. B. diese hier (S. 3): „Da kleinere Kinder sich nicht an die Distanzregeln und Schutzmaßnahmen halten können, gleichzeitig aber die Infektion weitergeben können, sollte der Betrieb in Kindertagesstätten nur sehr eingeschränkt wiederaufgenommen werden.“ Das scheint mir eine gut vertretbare und verantwortungsvolle Empfehlung zu sein, die allerdings mehr auf gesundem Menschenverstand als auf wissenschaftlichen Befunden über die Infektiosität kleiner Kinder beruht (das wissen wir nämlich nicht genau). Eine mehr aus der Psychologie stammende Empfehlung lautet (S. 10): „Bei den psychischen Folgen und gravierenden Überlastungen müssen sozioökonomische Aspekte und der Mangel an sozialer Einbettung dringend berücksichtigt werden. Zu den besonderen Risikogruppen gehören Alleinerziehende, Migrantinnen und Migranten ohne Sprachkenntnisse, alleinlebende Ältere, psychisch Erkrankte, Pflegefälle und Arbeitslose. In ärmeren und eher bildungsfernen Schichten fehlen tendenziell materielle, psychische und soziale Ressourcen.“ Auch hier fehlen wissenschaftliche Belege und die Empfehlung fällt mit gesundem Menschenverstand zusammen.

Skeptiker wie IJzerman et al. (2020) bezweifeln, dass psychologische Erkenntnisse – auch angesichts der noch nicht überwundenen Replikationskrise (Ioannidis, 2005), die auf viele empirisch vorgehende Wissenschaftsdisziplinen (nicht nur die

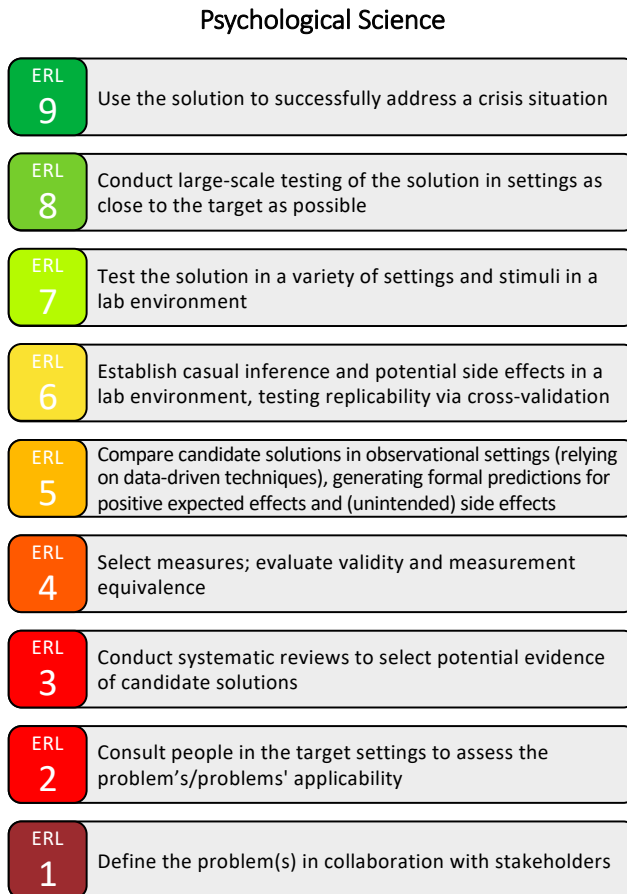


Abbildung 2: Definition von neun verschiedenen Evidenzgraden („Evidence Readiness Levels“, ERL; Source: IJzerman et al., 2020, p. 10; nach dem Vorbild der NASA¹¹).

Psychologie!) zutrifft – zur evidenzbasierten Politikberatung geeignet seien. Die Autoren schlagen vor, neun verschiedene Stufen von Evidenz (siehe Abbildung 2) zu unterscheiden. Nur die höchste Stufe (ERL9) sollte bei Entscheidungen über Leben und Tod herangezogen werden, aber psychologische Forschung liegt nach Meinung der Kritiker auf deutlich niedrigeren Stufen.

Das Potential der Psychologie zur Politikberatung wird bei IJzerman et al. (2020) nicht bestritten, allerdings kommt eine hohe Skepsis über den gegenwärtigen Evidenzstatus psychologischer Forschung zum Ausdruck. Zu beachten ist dabei, dass sich diese kritische Einschätzung auf die Validität empirischer Befunde bezieht. Daran arbeitet unser Fach nachdrücklich und in geradezu vorbildlicher Weise (so z. B. zu lesen bei Vazire, 2018). Das Potential unseres Faches beschränkt sich aber nicht nur auf die Sammlung verlässlicher Empirie, sondern enthält auch gut bewährte konzeptuelle Differenzierungen. Da zeigen sich die Stärken einer „Theoretischen“ Psychologie (Farrell & Lewandowsky, 2018; Fiedler, 2017, 2018; Gigerenzer, 2011, 2017; Oberauer & Lewandowsky, 2019), die wir bei aller Freude über gute Empirie nicht unterschätzen sollten.

9 Abschluss

Was können wir aus bisheriger Forschung zum Lösen komplexer Probleme lernen? Eine ganze Menge, wie ich finde! Hier ein paar „highlights“: (1) Es gibt (leider!?) keine Patentrezepte. (2) Niemand weiß, was die beste Lösung des Problems ist. (3) Man sollte trotz Ungewissheit und Unsicherheit Ruhe bewahren und nicht die Übersicht verlieren. (4) Fehler werden gemacht werden – es gilt, sie nicht vertuschen, sondern zu versuchen aus ihnen zu lernen. (5) Zeitverzögerungen erschweren die Erstellung eines adäquaten Lagebilds.

Hoffnungsvoll stimmen nicht nur Empathie und Hilfsbereitschaft großer Teile der Bevölkerung (als Gegengewicht gegen Egoismus und Verantwortungslosigkeit, Stichwort „Corona-Parties“), sondern auch zahlreiche kreative Lösungsversuche. Ein Beispiel: die schwäbische Textilfirma Trigema etwa produziert in diesen Zeiten keine T-Shirts, sondern Mundschutz. Automobilhersteller könnten bei der Produktion von Medizintechnik (Beatmungsgeräte) helfen.

Wir müssen wohl ertragen, was da kommt (und auch wieder vorübergeht – fragt sich nur, wann...). Die Entwicklung der Pandemie geht weiter (weitere Infektionswellen werden erwartet), strukturbildende Prinzipien (Abstandswahrung, Maskenpflicht, Hygiene) sind gut erkennbar. Wir sind potentiell Virentragende (und damit durch unser Verhalten Mitverantwortliche!) und gleichzeitig Leidtra-

¹¹ Die NASA spricht von „Technology Readiness Levels“, siehe https://www.nasa.gov/directorates/heo/scan/engineering/technology/txt_accordion1.html

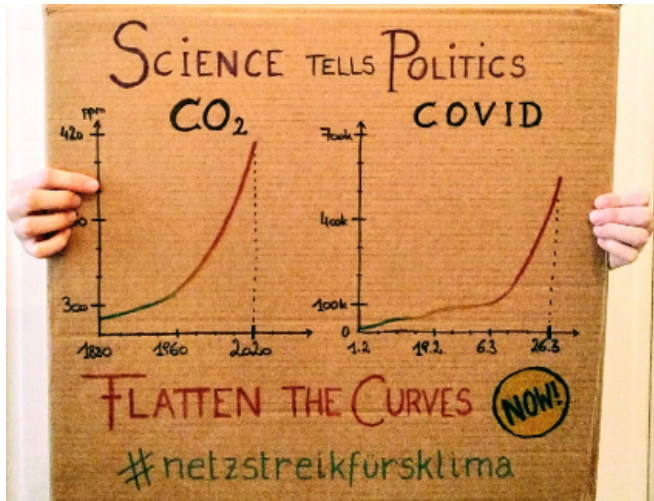


Abbildung 3: Parallelität von Corona- und Klima-Krise (Source: Twitter-Account #FloSchle, Florian Schlederer, 27.3.2020, on the occasion of #ClimateStrikeOnline).

gende dieser Krise, aber die gute Botschaft ist: Wir sind auch diejenigen, die etwas tun können! Und im Hintergrund wartet mit der Klimakrise eine weitere katastrophale Entwicklung, die zwar ganz ähnlich, aber wesentlich langsamer verläuft (siehe Abbildung 3).

Was auf jeden Fall weiterhilft, ist die Kompetenz zum „kritischen Denken“ (eines der besten Bücher zu diesem Thema: Halpern, 2013) und zum Denken in Systemen (siehe dazu auch Funke et al., 2018). Die weltweiten Geschehnisse zu verstehen ist sehr schwierig, aber reduktionistisches Denken, also die Rückführung komplexer Ereignisketten auf einzelne Ursachen hilft nicht weiter und führt in die Irre.

Im Fall der Pandemie ist Nichts-Tun (Stillhalten, also *physical distancing* - nicht *social distancing*)¹² sogar ausgesprochen vernünftiges Handeln („Stay safe at home“), im Fall der Klimakrise müssen wir rasch handeln! Aber natürlich kann jeder Einzelne auch mit seinen Kompetenzen zum Lösen von Problemen beitragen: Allein der Hackathon der Bundesregierung „WirVsVirus“ (= Wir gegen

¹² Dank an Lenelis Kruse (Waldwimmersbach) für den Hinweis auf diesen wichtigen Unterschied!

das Virus) hat vom 20.–22.3.2020 fast 30.000 Personen auf der Suche nach Problemen und deren Lösungen zusammengebracht. Rund 3.000 Mentorinnen und Mentoren haben den teilnehmenden Personen geholfen, knapp 2.000 Probleme zu identifizieren und zu bearbeiten.¹³ Motto: „Wir sind sektorübergreifend zusammengekommen, um gemeinsam einen digitalen Raum zu schaffen, in dem an Lösungen für die Herausforderungen durch COVID-19 gearbeitet werden kann“. Das sind eindrucksvolle Zeichen solidarischen Handelns, die deutlich machen: Menschen sind dem Virus nicht hilflos ausgesetzt, sondern setzen sich so gut es geht zur Wehr. So etwas macht Mut und lässt auf bessere Zeiten hoffen! Die nächste Krise, das nächste komplexe Problem kommt bestimmt!

Danksagung Ich bedanke mich bei Dr. Marlene Endepohls (Heidelberg) für wertvolle Kommentare zu einer Vorfassung dieses Beitrags.

Literatur

- Ayalon, L., Chasteen, A., Diehl, M., Levy, B., Neupert, S. D., Rothermund, K., Tesch-Römer, C., & Wahl, H.-W. (2020). Aging in times of the COVID-19 pandemic: Avoiding ageism and fostering intergenerational solidarity. *The Journals of Gerontology: Series B*, 75(5). <https://doi.org/10.1093/geronb/gbaa051>
- Batinic, B., & Appel, M. (2008). *Medienpsychologie*. Springer.
- Becker, E. (1973). *The denial of death*. Free Press.
- Becker, E. (1975). *Escape from evil*. Free Press.
- Brehmer, B. (1992). Dynamic decision making: Human control of complex systems. *Acta Psychologica*, 81(3), 211–241.
- Brehmer, B. (1995). Feedback delays in complex dynamic decision tasks. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 103–130). Lawrence Erlbaum Associates.
- Brehmer, B., & Allard, R. (1991). Dynamic decision making: The effects of task complexity and feedback delay. In J. Rasmussen, B. Brehmer, & J. Leplat (Eds.), *Distributed decision making: Cognitive models for cooperative work* (pp. 319–334). Wiley.

¹³ Nach Angaben der Veranstalter auf der Webseite <https://wirvsvirushackathon.org/>.

- Buchanan, M. (2001). *Das Sandkorn, das die Erde zum Beben bringt: Dem Gesetz der Katastrophen auf der Spur oder warum die Welt einfacher ist, als wir denken*. Campus.
- Christmann, U., & Göhring, A.-L. (2016). A German-language replication study analysing the role of figurative speech in reasoning. *Scientific Data*, 3(October), 160098. <https://doi.org/10.1038/sdata.2016.98>
- Csapó, B., & Funke, J. (Eds.). (2017). The nature of problem solving. Using research to inspire 21st century learning. OECD Publishing. <https://doi.org/10.1787/9789264273955-en>
- Dörner, D. (1980). On the difficulties people have in dealing with complexity. *Simulation & Gaming*, 11(1), 87–106. <https://doi.org/10.1177/104687818001100108>
- Dörner, D. (1989). *Die Logik des Mißlingens. Strategisches Denken in komplexen Situationen*. Rowohlt.
- Dörner, D. (1996). *The logic of failure. Recognizing and avoiding error in complex situations*. Basic Books.
- Dörner, D. (2019). *Die Psychologie politischen Denkens* [Unveröffentlichter DFG-Antrag].
- Dörner, D., & Funke, J. (2017). Complex problem solving: What it is and what it is not. *Frontiers in Psychology*, 8(1153), 1–11. <https://doi.org/10.3389/fpsyg.2017.01153>
- Dörner, D., & Schaub, H. (1995). Handeln in Unbestimmtheit und Komplexität. *Organisationsentwicklung*, 3, 34–47.
- Dossey, J. A., & Funke, J. (2016). Canadian and United States students' performances on the OECD's PISA 2012 problem-solving assessment. *Canadian Journal of Science, Mathematics and Technology Education*, 16(1), 92–108. <https://doi.org/10.1080/14926156.2015.1119332>
- Farrell, S., & Lewandowsky, S. (2018). *Computational modeling of cognition and behavior*: Cambridge University Press. <https://doi.org/10.1017/CBO9781316272503>
- Fiedler, K. (2017). What constitutes strong psychological science? The (neglected) role of diagnosticity and a priori theorizing. *Perspectives on Psychological Science*, 12(1), 46–61. <https://doi.org/10.1177/1745691616654458>
- Fiedler, K. (2018). The creative cycle and the growth of psychological science. *Perspectives on Psychological Science*, 19(6), 433–438. <https://doi.org/10.1177/1745691617745651>
- Fischer, P., Jander, K., & Krueger, J. (2018). *Sozialpsychologie für Bachelor*. Springer. <https://doi.org/10.1007/978-3-662-56739-5>
- Forgas, J. P., Fiedler, K., & Sedikides, C. (Eds.). (2012). *Social thinking and interpersonal behavior*. Psychology Press.

- Funke, J. (2003). *Problemlösendes Denken*. Kohlhammer. <https://doi.org/10.1024/1016-264X.15.4.313>
- Funke, J. (2012). Complex problem solving. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (Vol. 38, pp. 682–685). Springer. https://doi.org/10.1007/978-1-4419-1428-6_685
- Funke, J., Fischer, A., & Holt, D. V. (2018). Competencies for complexity: Problem solving in the twenty-first century. In E. Care, P. Griffin, & M. Wilson (Eds.), *Assessment and teaching of 21st century skills. Research and applications* (pp. 41–53). Springer. https://doi.org/10.1007/978-3-319-65368-6_3
- Gigerenzer, G. (2011). Personal reflections on theory and psychology. *Theory & Psychology, 20*(6), 733–743. <https://doi.org/10.1177/0959354310378184>
- Gigerenzer, G. (2017). A theory integration program. *Decision, 4*(3), 133–145. <https://doi.org/10.1037/dec0000082>
- Gigerenzer, G., & Brighton, H. (2009). Homo Heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science, 1*(1), 107–143. <https://doi.org/10.1111/j.1756-8765.2008.01006.x>
- Gigerenzer, G., & Selten, R. (Eds.). (2001). *Bounded rationality: The adaptive toolbox*. MIT Press.
- Gilbert, P. (2019). Psychotherapy for the 21st century: An integrative, evolutionary, contextual, biopsychosocial approach. *Psychology and Psychotherapy: Theory, Research and Practice, 92*(2), 164–189. <https://doi.org/10.1111/papt.12226>
- Halpern, D. F. (2013). *Thought and knowledge: An introduction to critical thinking* (5th ed.). Psychology Press.
- Hartmann, T. (2012). Wicked problems and clumsy solutions: Planning as expectation management. *Planning Theory, 11*(3), 242–256. <https://doi.org/10.1177/14730952124440427>
- Hayes, J., Schimel, J., Arndt, J., & Faucher, E. H. (2010). A theoretical and empirical review of the death-thought accessibility concept in terror management research. *Psychological Bulletin, 136*(5), 699–739. <https://doi.org/10.1037/a0020524>
- Hertwig, R., & Engel, C. (2016). Homo ignorans: Deliberately choosing not to know. *Perspectives on Psychological Science, 11*(3), 359–372. <https://doi.org/10.1177/1745691616635594>
- Hofinger, G. (2013). Entscheiden in komplexen Situationen – Anforderungen und Fehler. In R. Heimann, S. Strohschneider, & H. Schaub (Eds.), *Entscheiden in kritischen Situationen: Neue Perspektiven und Erkenntnisse* (Issue 2013, pp. 3–21). Verlag für Polizeiwissenschaft.

- IJzerman, H., Lewis, N. A., Weinstein, N., DeBruine, L. M., Ritchie, S. J., Vazire, S., Forscher, P. S., Morey, R. D., Ivory, J. D., Anvari, F., & Przybylski, A. K. (2020). Psychological science is not yet a crisis-ready discipline. *PsyArXiv*. <https://doi.org/10.31234/osf.io/whds4>
- Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8), 696–701. <https://doi.org/10.1371/journal.pmed.0020124>
- Janis, I. L. (1982). *Groupthink. Psychological studies of policy decisions and fiascoes. Revised and enlarged edition*. Houghton Mifflin.
- Kals, E., & Maes, J. (Eds.). (2012). *Justice and conflicts*. Springer. <https://doi.org/10.1007/978-3-642-19035-3>
- Kessler, T., & Fritsche, I. (2018). *Sozialpsychologie*. Springer. <https://doi.org/10.1007/978-3-531-93436-5>
- Kray, J. (2019). *Entwicklungspsychologie: Ein Überblick für Psychologiestudierende und -interessierte*. Springer. <https://doi.org/10.1007/978-3-662-57761-5>
- Kruse, A. (2015). *Resilienz bis ins hohe Alter – was wir von Johann Sebastian Bach lernen können. Für alle Interessierten*. Springer.
- Lakoff, G., & Wehling, E. (2012). *The little blue book: The essential guide to thinking and talking Democratic*. Simon & Schuster.
- Leopoldina. (2020). *Dritte Ad-hoc-Stellungnahme: Coronavirus-Pandemie – Die Krise nachhaltig überwinden*. https://www.leopoldina.org/uploads/tx_leopublication/2020_04_13_Coronavirus-Pandemie-Die_Krise_nachhaltig_%C3%BCberwinden_final.pdf
- Lerner, M. J., & Montada, L. (1998). An overview: Advances in Belief in a Just World. Theory and methods. In L. Montada & M. J. Lerner (Eds.), *Responses to victimizations and belief in a just world* (pp. 1–8). Plenum Press.
- Lewandowsky, S., & Cook, J. (2020). *Das Handbuch über Verschwörungsmythen*. <http://sks.to/conspiracy>
- Lewin, K. (1935). *A dynamic theory of personality (selected papers)*. McGraw-Hill.
- Lohaus, A., & Vierhaus, M. (2015). *Entwicklungspsychologie des Kindes- und Jugendalters für Bachelor*. Springer. <https://doi.org/10.1007/978-3-662-45529-6>
- Mandl, C. (2017). *Vom Fehler zum Erfolg. Effektives Failure Management für Innovation und Corporate Entrepreneurship*. Springer. <https://doi.org/10.1007/978-3-658-18261-8>
- McMillen, C. W. (2016). *Pandemics: A very short introduction*. Oxford University Press.
- Müsseler, J., & Rieger, M. (Eds.). (2017). *Allgemeine Psychologie* (3. Auflage). Springer.

- Nocun, K., & Lamberty, P. (2020). *Fake Facts: Wie Verschwörungstheorien unser Denken bestimmen*. Quadriga.
- Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*. <https://doi.org/10.3758/s13423-019-01645-2>
- Pretz, J. E., Naples, A. J., & Sternberg, R. J. (2003). Recognizing, defining, and representing problems. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (1st ed., pp. 3–30). Cambridge University Press. <https://doi.org/10.1017/CBO9780511615771.002>
- Raworth, K. (2018). *Doughnut economics. Seven ways to think like a 21st-century economist*. Random House.
- Reason, J. (1990). *Human error*. Cambridge University Press.
- Reason, J. (2000). Human error: Models and management. *British Medical Journal*, 320(6), 768–770.
- Renn, O. (2014). *Das Risikoparadox – Warum wir uns vor dem Falschen fürchten*. Fischer.
- Schäfer, I., Gast, U., Hofmann, A., Knaevelsrud, C., Lampe, A., Liebermann, P., Lotzin, A., Maercker, A., Rosner, R., & Wöller, W. (Eds.). (2019). *S3-Leitlinie Posttraumatische Belastungsstörung*. Springer. <https://doi.org/10.1007/978-3-662-59783-5>
- Schaub, H. (2016). Der Mensch als Problem und Problemlöser. Systemische Analyse zur Unterstützung von Entscheidern im Krisenmanagement. In T. Jäger, A. Daun, & D. Freudenberg (Eds.), *Politisches Krisenmanagement* (pp. 33–45). Springer. <https://doi.org/10.1007/978-3-658-09223-8>
- Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist*, 55(1), 5–14. <https://doi.org/10.1037/0003-066X.55.1.5>
- Siegler, R., Eisenberg, N., DeLoache, J., & Saffran, J. (2016). *Entwicklungspsychologie im Kindes- und Jugendalter* (S. Pauen, Trans.). Springer. <https://doi.org/10.1007/978-3-662-47028-2>
- Simon, H. A. (1973). The structure of ill structured problems. *Artificial Intelligence*, 4(3–4), 181–201. [https://doi.org/10.1016/0004-3702\(73\)90011-8](https://doi.org/10.1016/0004-3702(73)90011-8)
- Spada, H., & Kiesel, A. (Eds.). (2018). *Lehrbuch Allgemeine Psychologie* (4., vollständig überarbeitete und erweiterte Auflage). Hogrefe.
- Spinath, B., & Brünken, R. (2016). *Pädagogische Psychologie – Diagnostik, Evaluation und Beratung*. Hogrefe. <https://doi.org/10.1026/02222-000>
- Stemmler, G., Hagemann, D., Amelang, M., & Spinath, F. M. (2016). *Differentielle Psychologie und Persönlichkeitsforschung* (8., überarbeitete Auflage). Kohlhammer.

- Sterman, J. D. (1989). Misperceptions of feedback in dynamic decision making. *Organizational Behavior and Human Decision Processes*, 43(3), 301–335. [https://doi.org/10.1016/0749-5978\(89\)90041-1](https://doi.org/10.1016/0749-5978(89)90041-1)
- Strobach, T., & Wendt, M. (2019). *Allgemeine Psychologie: Ein Überblick für Psychologiestudierende und -interessierte*. Springer.
- Trepte, S., & Reinecke, L. (2013). *Medienpsychologie*. Kohlhammer.
- Tuchman, B. W. (1984). *The march of folly. From Troy to Vietnam*. Ballantine Books.
- Tversky, A., Kahneman, D., Slovic, P., & Tversky, A. (1974). Judgement under uncertainty. Heuristics and biases. *Science*, 185, 1124–1131.
- Vazire, S. (2018). Implications of the credibility revolution for productivity, creativity, and progress. *Perspectives on Psychological Science*, 19(6), 7. <https://doi.org/10.1177/1745691617751884>
- Verweij, M., & Thompson, M. (Eds.). (2006). *Clumsy solutions for a complex world. Governance, politics and plural perceptions*. Palgrave Macmillan.
- Vollrath, M., & Krems, J. F. (2011). *Verkehrspsychologie: Ein Lehrbuch für Psychologen, Ingenieure und Informatiker*. Kohlhammer.
- Wehling, E. (2016). *Politisches Framing. Wie eine Nation sich ihr Denken einredet – und daraus Politik macht*. Halem.
- Werner, G. W., Eichhorn, W., & Friedrich, L. (Eds.). (2012). *Das Grundeinkommen*. KIT Scientific Publishing.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8, 338–353.

Über den Autor

Joachim Funke (geb. 1953 in Düsseldorf) ist seit 1997 Professor für Allgemeine und Theoretische Psychologie am Psychologischen Institut der Universität Heidelberg. Promoviert wurde er 1984 an der Universität Trier. Im Jahr 1990 habilitierte er sich an der Universität Bonn. Funke war Gastprofessor an verschiedenen Universitäten, darunter Fribourg (Schweiz), Melbourne (Australien), Nanjing (China) und Szeged (Ungarn). Seine primären Forschungsinteressen liegen im Bereich von Denken, Kreativität und Problemlösen. Seine Forschungsideen wurden unter anderem von der Deutschen Forschungsgemeinschaft, von verschiedenen Bundesministerien und von der Europäischen Union finanziell unterstützt. Funke hat zahlreiche Artikel in internationalen Fachzeitschriften veröffentlicht, Kapitel zu Fachbüchern beigetragen und eigene Bücher herausgegeben und veröffentlicht.

Von 2010 bis 2014 war er Vorsitzender der Internationalen Expertenkommission für Problemlösen im Rahmen der weltweiten PISA-Studien der OECD. Auf ihn geht ein Wechsel im Verständnis von Problemlösen zurück, der die Perspektive von statischem zu dynamischen Problemlöseaktivitäten verschiebt. Von der ungarischen Universität Szeged wurde ihm 2015 für seine Verdienste um die computerbasierte Erfassung von Problemlöseprozessen der Titel eines Ehrendoktors verliehen. Von Oktober 2011 bis März 2019 war Funke Sprecher des Akademischen Senats der Universität. Im April 2019 hat sein Ruhestand begonnen.

Korrespondenz:

Prof. Dr. Dr. h.c. Joachim Funke (ORCID ID: 0000-0001-9129-2659)

Universität Heidelberg

Psychologisches Institut

Hauptstr. 47–51

69117 Heidelberg, Germany

E-Mail: Joachim.funke@psychologie.uni-heidelberg.de

Homepage: <https://funke.uni-hd.de>