

Forensische Genetik

Beweiskraft von DNA-Spuren: Was macht das Y-Chromosom so besonders?

AMKE CALIEBE, MICHAEL KRAWCZAK
INSTITUT FÜR MEDIZINISCHE INFORMATIK UND STATISTIK, UNIVERSITÄT ZU KIEL,
UNIVERSITÄTSKLINIKUM SCHLESWIG-HOLSTEIN, KIEL

Many crime cases require clarification whether or not a trace was left by a suspect. This is achieved by comparing their DNA profiles, a match indicating identity of suspect and donor. In sexual crimes, the trace is often a mixture of male and female material. This renders the analysis of the Y chromosome useful, but the interpretation of the genetic results is complicated by the specific mode of inheritance. We review some of the arising challenges as well as recently suggested solutions.

DOI: 10.1007/s12268-020-1399-x
© Die Autoren 2020

■ Bei der Verfolgung von Straftaten geht es oftmals um die Beantwortung der Frage, ob eine am Tatort gefundene biologische Spur von einer bestimmten Person stammt – in der Regel einem Verdächtigen. Zu diesem Zweck wird das Erbmaterial, das heißt die DNA, die in der Spur enthalten ist, mit der DNA der fraglichen Person verglichen. In der Praxis konzentriert sich diese Analyse auf zehn bis 30 kurze Abschnitte der DNA, die als Mikrosatelliten (*short tandem repeats*, STRs) bezeichnet werden und in der Bevölkerung eine hohe Variabilität aufweisen (Abb. 1). Mit Ausnahme von monozygoten Zwillingen tragen Menschen jeweils höchst individuelle STR-Profile. Wie sehr eine Übereinstimmung zweier STR-Profile dafür spricht, dass der Verdächtige der Verursacher der Spur (das heißt der „Spurleger“) ist, hängt neben der Quantität und Qualität der DNA maßgeblich von der Häufigkeit des STR-Profiles in der Bevölkerung ab.

In der Forensik wird die Beweiskraft genetischer Daten, D , üblicherweise durch ein Chancenverhältnis (*likelihood ratio*, LR) quantifiziert. Zur Erklärung einer Übereinstimmung lassen sich nämlich zwei Hypothesen aufstellen: (1) die Hypothese H_p der Anklage (*prosecution*), wonach der Verdächtige der Spurleger ist, und (2) die Hypothese H_D der Verteidigung (*defence*), wonach die

Spur von einer anderen Person stammt. Das Chancenverhältnis stellt die Vereinbarkeit beider Hypothesen mit den genetischen Daten einander gegenüber, das heißt

$$LR = \frac{P(D|H_p)}{P(D|H_D)}$$

wobei $P(D|H_p)$ bzw. $P(D|H_D)$ die Wahrscheinlichkeit bezeichnet, mit der sich die genetischen Daten unter der jeweiligen Hypothese ergäben. Die LR gibt also an, um wie viel wahrscheinlicher die Daten D unter H_p sind als unter H_D .

Weisen Spur und verdächtige Person das gleiche STR-Profil auf, und geht man vom einfachen Fall aus, dass die Spur nur von einer einzelnen Person stammt, dann gilt

A ... AGTCGTTATGATGGAGTCATTAATAA...ACTCTTAATTATAATCTGAACT
GGTACAGTAAGCCATTAGTAGACATAGAAACGGAAGCTTTAAAGATTTCAGA
GTTCCGCGAGTCTATCTATCTATCTATCTATCTATCTATCTACTGAGTGAAGT
GTAGTGCCAAACCTCCGGTATTACAGGCAAAGACAAAACAGAACCCATAAA
TGACCTGAAGACTATGTCTGCAGCCCTGAAGATACCTCATTTTGGTTA ...

Name	Chromosom	STR-Motiv	C	
TPOX	2	AATG	8	9
D3S1358	3	AGAT	14	14
FGA	4	CTTT	24	26
D5S818	5	AGAT	12	12
CSF1PO	5	AGAT	10	11
D7S820	7	GATA	8	10
D8S1179	8	TATC	8	10
TH01	11	AATG	6	8
vWA	12	AGAT	15	18
D13S317	13	GATA	11	11
D16S539	16	GATA	9	11
D18S51	18	GAAA	17	18
D21S11	21	TCTA	31	32

▲ **Abb. 1:** *Short tandem repeats* (STRs). **A**, STRs sind DNA-Abschnitte an festgelegten Positionen im Genom, die aus der tandemartigen Wiederholung von zwei bis vier Nukleotide umfassenden Motiven bestehen. Die Anzahl der Wiederholungen ist höchst variabel, weshalb sich ihre Bestimmung gut für Identifikationszwecke eignet. **B**, Die 13 autosomalen STRs des *Combined DNA Index System* (CODIS) bilden den Kern aller weltweit in der Forensik verwendeten genetischen Analyse-Kits. **C**, Ein typisches mitteleuropäisches STR-Profil für die CODIS-Marker. Die ganzen Zahlen bezeichnen die Anzahl der Motivwiederholungen am jeweiligen STR auf dem väterlichen bzw. mütterlichen Chromosom.

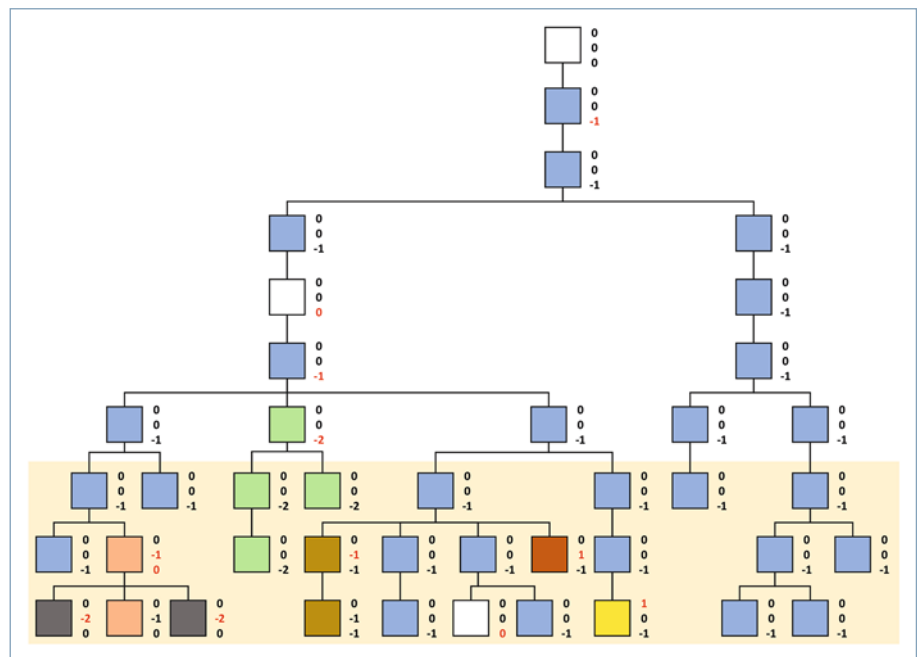
$P(D|H_p) = 1$: Hat die verdächtige Person die Spur gelegt, so findet man auch sicher eine Übereinstimmung. Unter der Hypothese H_D der Verteidigung wird meistens angenommen, dass die Spur von einer zufällig der Bevölkerung entstammenden Person gelegt wurde. In diesem Fall entspricht $P(D|H_D)$ der relativen Häufigkeit, f , des STR-Profiles in der Bevölkerung, und man spricht folgerichtig von f als *random match probability* [1, 2]. Für das Chancenverhältnis gilt in diesem Fall $LR = 1/f$, das heißt, je seltener das STR-Profil von Spur und verdächtigter Person ist, umso mehr spricht die Übereinstimmung für die Hypothese der Anklage.

Das Besondere an Y-chromosomalen STR-Profilen

Für ihre Verwendung in der Forensik werden STR-Marker auf den Autosomen (das heißt den humanen Chromosomen 1 bis 22) so ausgewählt, dass zwischen ihren Ausprägungen weitestgehende stochastische Unabhängigkeit besteht. Man erreicht dies z. B. dadurch, dass nur STRs von unterschiedlichen Chromosomen oder mit hinreichend großem Abstand voneinander verwendet werden. Ist die Unabhängigkeit gewährleistet, so ergibt sich die Häufigkeit f des Gesamtprofils durch Multiplikation der Häufigkeiten der einzelnen STR-Ausprägungen (üblicherweise im einstelligen oder unteren zweistelligen Prozentbereich), was die Ermittlung selbst kleinster Werte von f möglich macht.

Die Nutzung autosomaler STR-Profile ist in der Forensik weitverbreitet und gut etabliert. In einigen Fällen helfen sie den Ermittlern aber nicht weiter, und es muss gezielt auf STRs vom für Männer spezifischen Y-Chromosom zurückgegriffen werden. Y-chromosomale Marker eignen sich besonders gut, wenn es sich bei der Spur um eine Mischung von weiblichem und männlichem biologischen Material handelt. Dies ist regelmäßig bei Sexualdelikten der Fall, wo häufig ein sehr hoher Anteil der DNA vom weiblichen Opfer und nur ein geringer Teil vom männlichen Täter stammt. Diesen männlichen Anteil mithilfe autosomaler STRs eindeutig zu identifizieren, ist wegen der Überlagerung durch das weibliche Profil meist schwierig. Eine Lösung des Problems besteht in der Erzeugung eines Y-chromosomalen STR-Profiles der Spur, das dann nur vom männlichen Täter stammen kann.

Wir haben bereits gesehen, dass die Beweiskraft einer Übereinstimmung für ein autosomales STR-Profil dem Kehrwert seiner



▲ **Abb. 2:** Simulierte männliche Abstammungslinien eines Zehn-Generationen-Stammbaums (nach [6]). Neben jedem Mann (Quadrat) ist sein Profil für drei hypothetische Y-chromosomale STRs angegeben, wobei die Ausprägungen durch Subtraktion der Wiederholungen im Profil des letzten gemeinsamen Vorfahrens (LGV) skaliert wurden (das heißt 0, 1, -1 usw. statt der jeweils absoluten Anzahl der Wiederholungen; 0 entspricht somit der Ausprägung beim LGV). Y-chromosomale STR-Profile ändern sich nur durch Mutationen, die nahezu ausschließlich im Hinzutreten oder Wegfallen einer Motivwiederholung bestehen (rot markiert). Die in der Simulation durch Mutation aus dem Profil des LGV (weiß) neu entstandenen Profile sind farblich kennzeichnet. Unter realistischen Bedingungen könnten die letzten drei Generationen (gelb hinterlegt) den Anteil des Stammbaums an der aktuellen Bevölkerung ausmachen.

Häufigkeit in der Bevölkerung entspricht. Je seltener das Profil ist, umso mehr spricht die Übereinstimmung gegen den Verdächtigten und somit für die Hypothese der Anklage. Im Prinzip lässt sich mit Y-chromosomalen STR-Profilen ebenso verfahren, allerdings mit einer Einschränkung: Da das Y-Chromosom – mit Ausnahme der pseudoautosomalen Regionen an dessen Enden – während seiner Vererbung keine Rekombination des großelterlichen Erbmaterials durchläuft (wie dies bei autosomalen Markern der Fall ist), besteht keine stochastische Unabhängigkeit zwischen den Markern. Daher lässt sich die Häufigkeit des Gesamtprofils nicht durch einfache Multiplikation der Häufigkeiten der STR-Ausprägungen errechnen (die von ähnlicher Größenordnung sind wie bei autosomalen STRs), sondern muss auf andere Weise ermittelt werden.

Einschub: die Rolle der Verwandtschaft

Während jedes Autosom eine einzigartige, durch Rekombination entstandene Mischung aus großmütterlichem und großväterlichem Erbmaterial darstellt, werden Y-Chromoso-

men *in toto* von Vätern an ihre Söhne weitergegeben. Neue Y-chromosomale STR-Profile entstehen nur durch (vergleichsweise seltene) Mutationen. Daher weisen selbst Paare entfernter männlicher Verwandter mit relativ hoher Wahrscheinlichkeit identische Y-chromosomale STR-Profile auf [3]. Betrachtet man z. B. die 23 Y-chromosomale STR-Marker, die im PowerPlex-Y23®-Analyse-Kit (Promega) zusammengefasst und – wie bei allen anderen Kits auch – außerhalb der pseudoautosomalen Region lokalisiert sind, dann wird ein zugehöriges Profil mit einer Wahrscheinlichkeit von 92 Prozent unverändert vom Vater auf seinen Sohn vererbt [4]. Für zwei Männer, die über 20 Meiosen miteinander verwandt sind, beträgt die Wahrscheinlichkeit identischer Profile immerhin noch 18 Prozent. Bei 20 trennenden Meiosen handelt es sich um eine sehr entfernte Verwandtschaft, die z. B. dann besteht, wenn zwei Männer vor zehn Generationen den letzten gemeinsamen Vorfahren hatten – von dessen Existenz sie höchstwahrscheinlich gar nichts wüssten. Demgegenüber beträgt die Wahrscheinlichkeit, dass Vater und Sohn für z. B. 15 autosomale STR-Marker das glei-

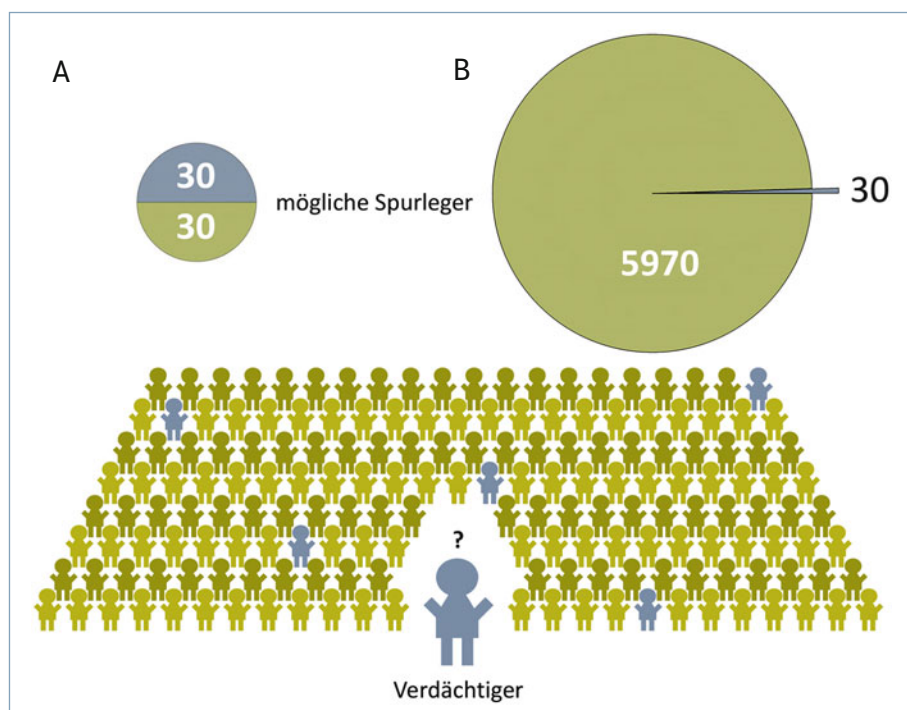
che Profil aufweisen, weniger als ein Milliardstel Prozent; für weitläufigere Verwandtschaften ist diese Wahrscheinlichkeit sogar noch um Größenordnungen geringer.

Aufgrund der speziellen Vererbung des Y-Chromosoms bilden sich in einer Bevölkerung im Verlauf von Generationen Cluster identischer oder sehr ähnlicher Y-chromosomaler STR-Profile heraus, die im Wesentlichen (entfernte) Verwandte umfassen (**Abb. 2**). Computersimulationen ergaben, dass dabei von der sehr großen Zahl theoretisch möglicher Profile nur eine geringe Zahl auch tatsächlich vorliegt. Beim häufig verwendeten Analyse-Kit PowerPlex Y (bestehend aus 12 Markern) sind theoretisch rund 50 Billionen verschiedene Profile denkbar; in simulierten Populationen von jeweils 1,5 Millionen Männern waren aber im Mittel nur rund 70.000 vorhanden, das heißt ein Ausschritt von wenigen Zehnmillionstel Prozent aller Möglichkeiten [5].

Die Beweiskraft einer Übereinstimmung bei Y-chromosomalen STR-Profilen

Wegen der bereits erwähnten Besonderheiten Y-chromosomaler STR-Profile ergeben sich für die Bewertung einer Übereinstimmung zwischen Spur und Verdächtigem zwei Probleme. Zum einen erschweren das Fehlen stochastischer Unabhängigkeit und das Vorliegen familiärer Profil-Cluster die Quantifizierung der Häufigkeit eines STR-Profils in der Bevölkerung. Darüber hinaus kann davon ausgegangen werden, dass männliche Verwandte eines Verdächtigen mit höherer Wahrscheinlichkeit dem Kreis möglicher Spurleger angehören als unverwandte Männer (siehe unten). Beides führt dazu, dass sich herkömmliche Methoden zur Berechnung des Chancenverhältnisses LR auf Y-chromosomale STR-Profile nicht sinnvoll anwenden lassen.

So kann etwa der klassische Ansatz, die Häufigkeit eines STR-Profils aus einer Referenzstichprobe oder einer Datenbank zu schätzen, nicht ohne Weiteres auf Y-chromosomale Marker übertragen werden. Vielmehr spricht die Tatsache, dass das fragliche Profil bereits in der Spur beobachtet wurde (und unter H_D sogar mindestens ein zweites Mal in der Bevölkerung, nämlich beim unbeteiligten Verdächtigen), stark dafür, dass das Profil Teil eines Clusters ist, dessen Umfang durch die Stichprobe nicht hinreichend abgebildet ist. Forensische Datenbanken setzen sich aus Einzelstudien zusammen, in denen



▲ **Abb. 3:** Die Beweiskraft einer Übereinstimmung zwischen den Y-chromosomalen STR-Profilen einer Spur und eines Verdächtigen hängt von der Häufigkeit des Profils (blaugrau) bei möglichen Spurlegern ab. Ist dieser „Kreis der Verdächtigen“ klein (**A**), so ergibt sich ein niedriger Wert für das Chancenverhältnis (hier: $LR = 2$) und die Beweiskraft gegen den Verdächtigen ist gering. Sind Träger des Profils hingegen Teil einer großen Gruppe möglicher Spurleger (**B**), so ist das Chancenverhältnis groß (hier: $LR = 200$) und die Beweiskraft gegen den Verdächtigen entsprechend hoch.

oftmals gezielt die Profile unverwandter Personen eingeschlossen werden. Der dänische Mathematiker Mikkel Meyer Andersen und sein englischer Kollege David Balding entwickelten daher 2017 einen gänzlich neuen methodischen Ansatz zur Bewertung der Beweiskraft einer Übereinstimmung zwischen Y-chromosomalen STR-Profilen [6]. Sie stützten sich dabei auf Simulationen der Historie einer großen männlichen Bevölkerung und ermittelten, wie häufig ein tatsächlich beobachtetes Profil in der Regel in der Bevölkerung auftritt. Für PowerPlex Y23 ergab sich daraus, dass im Mittel weltweit 20 bis 30 Männer das gleiche Profil aufweisen; bei weniger als fünf Prozent der Simulationen kamen die Autoren auf mehr als 100 Männer. Zudem waren alle Männer mit gleichem Profil über väterliche Linien miteinander verwandt, und die meisten Paare trennten nur wenige Meiosen (z. B. Onkel-Neffe oder Cousins). Andererseits wiesen aber viele Paare Abstände zwischen fünf und 20 Meiosen auf – zu weit entfernt, um von dieser Verwandtschaft verlässlich Kenntnis zu haben.

An die Schätzung, wie viele Männer das gleiche Y-chromosomale STR-Profil tragen wie Spur und Verdächtiger, schließt sich

unmittelbar die Frage an, inwieweit diese Männer auch zum Kreis der Verdächtigen gehören. Die Charakterisierung möglicher Spurleger ist für die Berechnung des Chancenverhältnisses gerade bei Y-chromosomalen Profilen von zentraler Bedeutung. Ein Verdächtiger entstammt in der Regel nicht „zufällig“ der Gesamtbevölkerung, sondern gehört zu einer deutlich kleineren, fallspezifischen Gruppe von Personen mit ähnlichen demographischen, geographischen, ethnischen und sozialen Merkmalen (*suspect population* [7]), der dann auch männliche Verwandte des Verdächtigen bevorzugt angehören. Diese Einschränkung hat erhebliche Auswirkungen auf das Chancenverhältnis: Gehören z. B. alle 30 Profilträger einem kleinen Kreis von 60 möglichen Spurlegern an, so gilt $LR = 1/0,5 = 2$. Ist der Kreis der Verdächtigen jedoch sehr groß, z. B. 6.000 wie bei einem Fundort der Spur mit überregionalem Milieu (Autobahnraststätte, Bahnhof, Flughafen), so beträgt die LR schon 200 (**Abb. 3**).

In der Realität ist es unwahrscheinlich, dass alle Träger des STR-Profils einer Spur zum Kreis der Verdächtigen gehören (viele werden schon aufgrund ihres Alters nicht

infrage kommen). Die Träger haben aber eine gemeinsame väterliche Abstammung und werden sich daher mit erhöhter Wahrscheinlichkeit hinsichtlich des sozialen Status, Wohnorts, der Ethnizität, Religion und des Aussehens ähneln [8]. Deshalb ist davon auszugehen, dass das Y-chromosomale STR-Profil einer Spur im Kreis der möglichen Spurleger angereichert ist [4, 6]. Daher sollte in jedem einzelnen Fall der Versuch unternommen werden, unter Berücksichtigung seiner besonderen Umstände die Größe des Kreises möglicher Spurleger und der darin zu erwartenden Übereinstimmungen abzuschätzen, um so einen realistischen Anhalt für das tatsächliche Chancenverhältnis LR zu bekommen.

Schlussfolgerung

Y-chromosomale STR-Profile weisen wegen fehlender Rekombination statistische Besonderheiten auf, deretwegen sich Verfahren zur Quantifizierung der Beweiskraft von Übereinstimmungen zwischen Spur und Verdächtigem nicht direkt von autosomalen Profilen übertragen lassen. Profilhäufigkeiten in Datenbanken oder Referenzstichproben sind nur bedingt aussagekräftig und entsprechen nicht der Wahrscheinlichkeit einer Übereinstimmung unter der Hypothese H_D der Verteidigung („Der Verdächtige ist nicht der Spurleger“). Eine solche Gleichsetzung wäre für Y-chromosomale STR-Profile zudem anti-konservativ in dem Sinne, dass es die Hypothese H_p der Anklage begünstigte. Stattdessen müsste eine faire Quantifizierung der Beweiskraft einer Übereinstimmung für Y-chromosomale Profile auch die familiäre Struktur der Bevölkerung und die Beschaffenheit des Kreises möglicher Spurleger

berücksichtigen. Dass dies in der Praxis nicht einfach und standardisiert möglich ist, schließt nach Ansicht der Autoren die Notwendigkeit einer kritischen Auseinandersetzung mit der zugehörigen Problematik nicht aus.

Literatur

- [1] Caliebe A, Krawczak M (2016) Probability and likelihood. In: Amorim A, Budowle B (Hrsg) Handbook of Forensic Genetics: Biodiversity and Heredity in Civil and Criminal Investigation. Imperial College Press, London, S 61–80
- [2] Caliebe A (2020) Statistische Methoden in der forensischen Genetik: Y-chromosomale Match-Probabilities und Forensic-DNA-Phenotyping. Habilitationsschrift, Christian-Albrechts-Universität zu Kiel
- [3] Brenner CH (2010) Fundamental problem of forensic mathematics – the evidential value of a rare haplotype. *Forensic Sci Int Genet* 4:281–291
- [4] Caliebe A, Krawczak M (2018) Match probabilities for Y-chromosomal profiles: a paradigm shift. *Forensic Sci Int Genet* 37:200–203
- [5] Andersen MM, Caliebe A, Kirkeby K et al. (2020) Estimation of Y haplotype frequencies with lower order dependencies. *Forensic Sci Int Genet* 46:102214
- [6] Andersen MM, Balding DJ (2017) How convincing is a matching Y-chromosome profile? *PLOS Genet* 13:e1007028
- [7] Balding DJ, Nichols RA (1994) DNA profile match probability calculation: how to allow for population stratification, relatedness, database selection and single bands. *Forensic Sci Int* 64:125–140
- [8] Roewer L (2009) Y chromosome STR typing in crime casework. *Forensic Sci Med Pathol* 5:77–84

relatedness, database selection and single bands. *Forensic Sci Int* 64:125–140

[8] Roewer L (2009) Y chromosome STR typing in crime casework. *Forensic Sci Med Pathol* 5:77–84

Funding: Open Access funding provided by Projekt DEAL.

Open Access: Dieser Artikel wird unter der Creative Commons Namensnennung 4.0 International Lizenz veröffentlicht, welche die Nutzung, Vervielfältigung, Bearbeitung, Verbreitung und Wiedergabe in jeglichem Medium und Format erlaubt, sofern Sie den/die ursprünglichen Autor(en) und die Quelle ordnungsgemäß nennen, einen Link zur Creative Commons Lizenz beifügen und angeben, ob Änderungen vorgenommen wurden. Die in diesem Artikel enthaltenen Bilder und sonstiges Drittmaterial unterliegen ebenfalls der genannten Creative Commons Lizenz, sofern sich aus der Abbildungslegende nichts anderes ergibt. Sofern das betreffende Material nicht unter der genannten Creative Commons Lizenz steht und die betreffende Handlung nicht nach gesetzlichen Vorschriften erlaubt ist, ist für die oben aufgeführten Weiterverwendungen des Materials die Einwilligung des jeweiligen Rechteinhabers einzuholen. Weitere Details zur Lizenz entnehmen Sie bitte der Lizenzinformation auf <http://creativecommons.org/licenses/by/4.0/deed.de>.

Korrespondenzadresse:

PD Dr. Amke Caliebe
Institut für Medizinische Informatik und Statistik
Universitätsklinikum Schleswig-Holstein,
Campus Kiel
Arnold-Heller-Straße 3, Haus V40
D-24 105 Kiel
caliebe@medinfo.uni-kiel.de
www.uni-kiel.de/medinfo/mitarbeiter/caliebe

AUTOREN



Amke Caliebe

Physik- und Mathematikstudium (Diplom) an den Universitäten Oldenburg und zu Kiel. 1999 Promotion in Stochastik und Statistik an der Universität zu Kiel; dort 2020 Habilitation in medizinischer Biometrie. Seit 2003 wissenschaftliche Angestellte und seit 2020 Privatdozentin am Institut für Medizinische Informatik und Statistik der Universität zu Kiel unter der Leitung von Prof. Dr. M. Krawczak.



Michael Krawczak

Mathematik- und Volkswirtschaftslehrestudium (Diplom) in Göttingen. Dort nach einem Aufbaustudium 1989 Promotion im Fach Entwicklungsbiologie, 1994 Habilitation für Humangenetik an der Medizinischen Hochschule Hannover. 1996–2001 DFG-Heisenbergstipendiat und Professor für Mathematische Genetik am University of Wales College of Medicine, Cardiff, UK. Seit 2001 Inhaber des Lehrstuhls für Medizinische Informatik und Statistik an der Universität zu Kiel und Direktor des gleichnamigen Instituts.