

cette réforme, a priori anodine, entraîne toute une série de conséquences dont la mise en œuvre est chaotique, parce que la loi n'a pas été conçue pour parer à cette hypothèse.

Ainsi, on doit désormais reconnaître au prévenu un droit au classement de la procédure par le ministère public, lorsque les conditions d'application d'une clause d'exemption de peine (à l'exception de l'art. 52 CP) sont réalisées au terme de l'instruction. Ce droit et sa justiciabilité devant le Tribunal fédéral découlent du préjudice irréparable que le prévenu subit en raison de l'inscription au casier judiciaire lorsqu'il est condamné mais exempté de peine par le tribunal, alors qu'il échappe à toute inscription lorsque les mêmes conditions matérielles imposaient le prononcé d'un classement en opportunité. Toutefois, la mise en œuvre de ce droit aboutit au constat que seul un recours direct au Tribunal est envisageable contre le refus du ministère public d'y procéder, ce qui n'est assurément pas adéquat.

Plusieurs solutions seraient envisageables. La première, et sans doute la plus simple, consisterait en un revirement de jurisprudence permettant au tribunal saisi de procéder au classement de la procédure, en reprenant une interprétation littérale de l'art. 8 al. 1 et 4 CPP. Ainsi, le refus de classer du ministère public n'emporterait aucun préjudice pour le prévenu qui pourrait encore obtenir un classement devant le juge du fond et donc éviter l'inscription au casier judiciaire. Toutefois, le Tribunal fédéral risque fort de ne pas aller dans cette direction car elle reviendrait à neutraliser la volonté du législateur lorsqu'en adoptant la LCJ, il a voulu que les jugements d'exemption de peine soient inscrits au casier judiciaire. Sur le plan législatif, la LCJ pourrait être modifiée pour revenir à une solution uniforme, à savoir une inscription ou l'absence de toute inscription dans toutes les hypothèses. Enfin, s'il fallait maintenir l'incohérence au sein de la LCJ, alors il faudrait au moins aménager dans le CPP une voie de recours au sens de l'art. 393 CPP contre une décision formelle que le ministère public devrait rendre lorsqu'il refuse un classement en opportunité.

Alex Biedermann*, Lausanne

Joëlle Vuille, Lausanne

Was ist der Beweiswert von DNA-Analyseergebnissen?

2. Teil: Probleme bei der Anwendung der grundlegenden Prinzipien der Befundbewertung

Inhaltsübersicht

- I. Einleitung
- II. Nicht interpretierbare Befunde
- III. Zum Beweiswert in Fällen, in welchen keine (übereinstimmende) DNA festgestellt werden konnte
- IV. Das ungelöste Fehlerproblem
 1. Zum Begriff des Fehlers und zu der Häufigkeit von Fehlern
 2. Warum dem Fehlerpotenzial Rechnung zu tragen ist
 3. Zum Umgang mit dem Fehlerpotenzial
- V. Das Verständnis der Individualisierung als Entscheidungsproblem und die richterliche Überzeugung
- VI. Mischspuren, probabilistische Genotypisierung und die neue Ära des «maschinellen Beweismittels» («Machine Evidence»)
- VII. Schlussfolgerungen

I. Einleitung

Die neuen Bestimmungen des überarbeiteten DNA-Profil-Gesetzes¹ legen das Hauptaugenmerk auf Neuerungen in Bezug auf Zwangsmassnahmen, unter anderem auf die Möglichkeit der Verwandtschaftssuche sowie die forensische Phänotypisierung. Die Methoden der modernen forensischen Genetik präsentieren sich somit einmal mehr als Ermittlungsinstrumente. Dabei steht das Problem des «Findens» im Vordergrund, d. h. der Eruiierung der Person, von welcher eine bestimmte DNA-Spur stammt. Demgegenüber weitgehend unbeachtet bleibt das Problem des «Beweisens» im Rahmen der Tatsachenfeststellung vor Gericht, insbeson-

* Die Autoren danken dem Schweizerischen Nationalfonds für die finanzielle Unterstützung (Projekt Nr. PP00P1_176720) sowie Prof. Dr. Jonathan J. Koehler und Prof. Dr. Niamh Nic Daeid für Literaturhinweise. Der vorliegende Beitrag beruht auf einem Vortrag, welcher am 20. Oktober 2022 an der Juristischen Fakultät der Universität Basel gehalten wurde, auf Einladung von Prof. Dr. Sabine Gless und Prof. Dr. Wolfgang Wohlers. Ihnen sei an dieser Stelle für diese Gelegenheit herzlich gedankt.

¹ Bundesgesetz vom 20. Juni 2003 über die Verwendung von DNA-Profilen im Strafverfahren und zur Identifizierung von unbekanntem oder vermissten Personen (DNA-Profil-Gesetz, SR 363).

dere die Frage, welche Schlüsse berechtigterweise gezogen werden können, wenn eine bestimmte Person mit einer DNA-Spur in Verbindung gebracht wird. Angesichts des investigativen Schwerpunkts der derzeitigen Regulierungsbestrebungen mag der Eindruck entstehen, dass die Lösung des Problems des «Findens» gleichzeitig das Beweisproblem löse.

Im Rahmen unseres vorliegenden, zweiteiligen Beitrags zeigen wir auf, weshalb dieser Auffassung entschieden widersprochen werden muss. Aufgrund einer DNA-Profilübereinstimmung kann nämlich nicht direkt auf die Spurenherkunft geschlossen werden, geschweige denn auf die Täterschaft. Im ersten Teil unseres Beitrags haben wir die hierfür relevanten, für sämtliche Sach- und Indizienbeweise geltenden Prinzipien und Denkgesetze vorgestellt. Im hier vorliegenden zweiten Teil befassen wir uns mit einigen regelmässig auftretenden und gleichzeitig häufig übersehenen Schwierigkeiten, die sich bei der Anwendung der grundlegenden Prinzipien der evaluierenden Befundbewertung ergeben. Wir beginnen mit der Erörterung der Problematik von nicht interpretierbaren Befunden (Abschnitt II.) sowie der Befundbewertung in Fällen, in welchen keine DNA nachgewiesen worden ist, obwohl aufgrund der Umstände des Einzelfalls das Gegenteil zu erwarten gewesen wäre (III.). Anschliessend nehmen wir uns der Frage an, wie das Fehlerpotenzial den Beweiswert beeinflusst und welche argumentativen Konsequenzen und praktischen Handlungsanweisungen sich daraus ergeben (IV.). Im Weiteren befassen wir uns kritisch mit den konzeptuellen Unzulänglichkeiten der in gewissen Sachverständigenkreisen propagierten Haltung, Spuren könnten alleine aufgrund von Ergebnissen von DNA-Profilanalysen eindeutig bestimmten Personen zugeordnet – d. h. individualisiert – werden (V.). Im letzten Abschnitt widmen wir uns der Befundbewertung bei komplexen Spuren(-profilen), wofür mittlerweile moderne, rechenintensive Methoden zur Anwendung kommen (VI.). Insgesamt kommen wir zum Schluss, dass der hohe Regulierungs- und Standardisierungsgrad der DNA-Profilanalyse keineswegs bedeutet, dass die Ergebnisverwendung über alle Zweifel erhaben ist, d. h. zu verhindern vermag, dass übertriebene und realitätsfremde Aussagen in gutachterlichen Berichten Eingang in Gerichtsurteile finden.

II. Nicht interpretierbare Befunde

Im ersten Teil dieses Beitrags haben wir dargelegt, dass der Bewertung von Ergebnissen von DNA-Profilanalysen derzeit (noch) häufig, zumindest in einem ersten Schritt, Hypothesen der Spurenherkunft zugrunde gelegt werden. Wir haben ebenfalls gesehen, dass die Frage der Spurenherkunft für die Gesamtbetrachtung des Einzelfalls jedoch nur von relativ begrenztem Interesse ist, nicht zuletzt wenn eine bestimmte Person, deren DNA-Profil mit dem Profil der Spur übereinstimmt, die Spurenherkunft gar nicht bestreitet.

Von weitaus grösserem Interesse ist hingegen häufig die Frage, ob und inwiefern sich aus den Analyseergebnissen Hinweise ergeben, die es erlauben, zwischen den verschiedenen von den Parteien behaupteten Aktivitäten (Tätigkeiten, Tatabläufen) zu unterscheiden. Diese «Wertsteigerung» der Befundbewertung hat jedoch ihren Preis: Sie erfordert zusätzliche Fachkenntnisse, insbesondere zu Spurenübertragungsmechanismen und zum Verbleib von Spuren auf diversen Spurentägern, unter gleichzeitiger Berücksichtigung der jeweiligen Fallumstände. Diese Anforderung kann sich als Engpass herausstellen, beispielsweise wenn entsprechende Fachkenntnisse nicht existieren, nicht zugänglich sind oder sich der Kompetenz der sachverständigen Person entziehen sollten. Es stellt sich daher die Frage, wie in einem solchen Fall zu verfahren wäre.

Gemäss einer verbreiteten Ansicht zu dieser Frage sollen sich Sachverständige doch «einfach» auf eine Befundbewertung mittels Hypothesen zur Spurenherkunft beschränken, da für eine solche wenig mehr als Angaben zum Seltenheitswert der übereinstimmenden DNA-Merkmal kombination notwendig sei,² und eine solche vergleichsweise einfach vorzunehmen sei. Wie wir nachstehend näher ausführen, kann dieser Sichtweise jedoch nur widersprochen werden.

Zunächst können wir den Richtlinienvermerk Nr. 2 der ENFSI-Richtlinie heranziehen, der sich zu dieser Frage wie folgt äussert: «Wenn es (...) um Fragen betreffend mutmasslicher Aktivitäten geht, ist das Fehlen von Daten oder Expertenwissen zur Spurenübertragung, zum Verbleib oder zum zufälligen Vorhandensein der betreffenden Spurenart keine Rechtfertigung dafür, Hypothesen zu Aktivitäten durch Hypothesen der Spurenherkunft zu ersetzen. Die Wahl zwischen Hypothesen zur Spurenherkunft und Hypothesen zu mutmasslichen Aktivitäten sollte nicht von der Verfügbarkeit von Daten oder Expertenwissen beeinflusst werden, sondern ausschliesslich vom Umstand, dass die Berücksichtigung von Faktoren wie die Spurenübertragung, der Spurenverbleib und das zufällige Auftreten von Spuren, die Aussagekraft der Ergebnisse im Kontext der Fallumstände entscheidend beeinflussen könnte.»³ Die Befundbewertung mit Hypothesen, die sich mit mutmasslichen Aktivitäten befassen, ist somit nicht lediglich eine Option, sondern ein Imperativ, sofern davon auszugehen ist, dass die Empfänger eines Sachverständigengutachtens (Verfahrensleitung, Gericht, Prozessparteien) in Ermangelung der notwendigen Fachkenntnisse selbst nicht in der Lage sein werden, die Bedeutung der Befunde hinsichtlich der mutmasslichen Aktivitäten zu ergründen. Es ist nämlich das Risiko zu vermeiden, dass ein hoher Beweiswert zur Frage der Spurenherkunft in unbegründeter Weise als be-

2 Siehe z. B. Erläuternder Bericht zur Eröffnung des Vernehmlassungsverfahrens, Änderung des DNA-Profil-Gesetzes, Bundesamt für Polizei fedpol, August 2019, 9.

3 S. Willis et al., Guideline for Evaluative Reporting in Forensic Science, Strengthening the Evaluation of Forensic Results across Europe, European Network of Forensic Science Institutes, 2015; abrufbar unter: https://enfsi.eu/wp-content/uploads/2016/09/m1_guideline.pdf (besucht am: 25. 1. 2023), Richtlinienvermerk Nr. 2, 13 (frei übersetzt).

weiskräftig für Hypothesen zu mutmasslichen Aktivitäten «uminterpretiert» wird.⁴ Daraus ergibt sich, dass in Fällen, in welchen ein Gericht mutmassliche Aktivitäten zu beurteilen hat, Sachverständige den Befund hinsichtlich dieser Aktivitäten jedoch nicht zu bewerten vermögen oder gewillt sind, der Befund hinsichtlich der Haupt- bzw. Nichthaupttatsache als *nicht interpretierbar* und somit für die Entscheidungsfindung auch als *nicht verwertbar* zu betrachten ist.⁵

Ungeachtet dieser Bedenken hält sich in gewissen Kreisen dennoch die Auffassung, eine Befundbewertung mit Hypothesen zur Spurenherkunft sei als «Standardoption» ausreichend. Man könne, so die Auffassung, Sachverständige vor Gericht bei Bedarf ja einfach mündlich fragen, welche Hinweise sich aus dem Beweiswert bezüglich der Spurenherkunft zu Fragen mutmasslicher Aktivitäten ergeben würden. Auch dieser Ansicht können wir nur vehement widersprechen.⁶ Die Befundbewertung mittels Hypothesen, die mutmassliche Aktivitäten betreffen, ist in der Regel anspruchsvoll und lässt sich ad hoc nicht fundiert vornehmen. Abhängig von den jeweiligen Fallumständen und den von den beteiligten Parteien vorgebrachten spezifischen Aktivitäten ist nämlich die relevante Fachliteratur zu sichten und abzuklären, ob der wissenschaftliche Kenntnisstand zur Problematik der Spurenübertragung, zu dem Spurenverbleib sowie dem zufälligen Auftreten von Spurenmaterial ausreicht, um eine belastbare Befundbewertung vorzunehmen. Denn sind entsprechende Erfahrungssätze inexistent oder nicht offengelegt, lässt sich nichts über die Beweiskraft sagen,⁷ was den Befund uninterpretierbar und unverwertbar macht.

III. Zum Beweiswert in Fällen, in welchen keine (übereinstimmende) DNA festgestellt werden konnte

Abschnitt V.1 des ersten Teils dieses Beitrags befasste sich mit der Frage, in welchen Fallkonstellationen und weshalb Hypothesen über mutmassliche Aktivitäten gegenüber Hypothesen zur Spurenherkunft vorzuziehen sind. Ein in diesem Zusammenhang wichtiges Argument haben wir bis anhin jedoch noch nicht er-

wähnt: Hypothesen, die sich auf mutmassliche Aktivitäten beziehen, sind insbesondere auch dann zu berücksichtigen, wenn *keine* DNA festgestellt werden konnte. Hypothesen zur Spurenherkunft sind in solchen Fällen nämlich zweifelsohne gegenstandslos: Wo nichts detektiert wurde, stellt sich auch keine Herkunftsfrage.

Anders verhält es sich hingegen mit Fragen betreffend mutmassliche Aktivitäten. Was «abwesende» DNA aber konkret zu beweisen vermag, lässt sich von vornherein – d. h. ohne formale Evaluierung – nicht sagen. Das weit herumgereichte Diktum «Das Fehlen von (passenden) Spuren ist kein Beweis für die Abwesenheit (der beschuldigten Person) am Tatort»⁸ ist nämlich nicht immer zutreffend. Auch liest sich in manchen Strafakten, die Tatsache, dass an einem bestimmten Ort *keine* mit dem Verdächtigen übereinstimmende DNA gefunden wurde, habe einen neutralen Wert.

Solche Beteuerungen bedürfen einer kritischen Betrachtung, wobei den Fallumständen einmal mehr Beachtung zu schenken ist. Ist beispielsweise aufgrund der Fallumstände zu erwarten, dass an einem bestimmten Ort mit der verdächtigen Person übereinstimmende DNA festzustellen sein sollte (unter der Annahme, dass die verdächtige Person die Tat verübt hat), dann jedoch keine DNA gefunden wird, so muss diese Feststellung logischerweise die Überzeugung stärken, dass die verdächtige Person *tatunbeteiligt* ist – dies aufgrund der Erwägung, dass das *Fehlen* von übereinstimmender DNA *besser* mit der Hypothese der Nichttäterschaft vereinbar ist.

Der australische Fall *R. v. Drummond* verdeutlicht diese Überlegung.⁹ In diesem Fall geht es um den schweren tätlichen Angriff (versuchte Entführung) auf eine Jugendliche auf offener Strasse. Dem Angreifer gelang die Flucht, dennoch konnte rasch ein Verdächtiger, D., festgenommen werden. D. konnte ausfindig gemacht werden, weil sich das Opfer das Nummernschild des Fahrzeugs des Angreifers merken konnte. An der Bekleidung von D. sowie an jener des Opfers wurden zeitnah DNA-Proben entnommen. Es konnte jedoch weder in der Probe von D.s Bekleidung noch in jener von der Bekleidung des Opfers übereinstimmende Fremd-DNA¹⁰ festgestellt werden, d. h. DNA, welche mit dem Opfer bzw. mit D. übereinstimmt. D. wurde zunächst verurteilt, weil die Geschworenen der Meinung waren, dass das Fehlen der DNA ihn nicht entlaste. Das Berufungsgericht hingegen sprach ihn frei, da es feststellte, dass das DNA-Untersuchungsergebnis den Geschworenen auf eine Weise präsentiert worden war, welche die Geschworenen hätte irreführen können.

4 Für eine nähere Besprechung der Thematik der unbestrittenen Spurenherkunft siehe auch G. Jackson/A. Biedermann, «Source» or «Activity» What is the level of issue in a criminal trial?, *Significance* 16(2)/2019, 36.

5 A. Biedermann/K. Kotsoglou, (Un-)Interpretability in expert evidence: an inquiry into the frontiers of evidential assessment, *Quaestio facti – International Journal on Evidential Legal Reasoning* 3/2022, 481.

6 J. Vuille, in: Y. Jeanneret/A. Kuhn/C. Perrier Depeursinge (Hrsg.), *Commentaire romand, Code de procédure pénale suisse*, 2. Aufl., Basel 2019, Art. 187 N 5a.

7 Siehe hierzu auch R. Bender/R. Häcker/V. Schwarz, *Tatsachenfeststellung vor Gericht*, 5. Aufl., München 2021, 168 (Rn. 699).

8 Freie Übersetzung der englischen Redewendung «The absence of (matching) evidence is not evidence of absence».

9 *R. v. Adrian Shane Drummond*, Supreme Court of South Australia – (2013) SASFC 135; *R. v. Adrian Shane Drummond* (No. 2), Supreme Court of South Australia – (2015) SASFC 82.

10 Wir sprechen von Fremd-DNA, wenn in einer Probe von einer bestimmten Person DNA festgestellt wird, deren Profil mit dem DNA-Profil dieser Person jedoch nicht übereinstimmt.

Die Fragen-Trias kann uns helfen, die Ergebnisse der DNA-Profilanalyse hinsichtlich der von der Anklage und der Verteidigung vertretenen Ansichten besser zu verstehen. Beginnen wir mit der Sichtweise der Verteidigung und, als Beispiel, dem Untersuchungsergebnis der Probe, welche der Bekleidung von D. entnommen wurde. Unter der Annahme, dass eine andere männliche Person als D. versucht hat, die Jugendliche zu entführen, d. h. D. nichts mit dem Fall zu tun hat, kann D. als beliebige unbeteiligte Person betrachtet werden. Unter dieser Annahme ist es nicht erstaunlich, *keine* mit dem Opfer übereinstimmende Fremd-DNA festgestellt zu haben. Anders ausgedrückt, die Wahrscheinlichkeit des Untersuchungsergebnisses unter der Annahme, dass D. *nicht* der Täter ist, ist sehr hoch. Kontrastieren wir diese Beurteilung nun mit der Sichtweise der Anklage. Unter der Annahme, dass D. der Täter ist, wäre in erster Linie zu erwarten, dass sich in der Probe DNA finden lässt, die mit dem Opfer übereinstimmt; keine übereinstimmende DNA zu finden, ist demzufolge *weniger* wahrscheinlich. Für die Zwecke unserer Diskussion hier ist es nicht notwendig, uns auf genaue Zahlenwerte festzulegen.¹¹ Das Konzept der Grössenordnung ist hingegen wesentlich: Der Befund ist mit dem Standpunkt der Verteidigung besser vereinbar als mit dem Standpunkt der Anklage. Somit spricht das Ergebnis der DNA-Profilanalyse für die Hypothese der Verteidigung, im Vergleich zur Hypothese der Anklage; es ist somit entlastend.

Nun liesse sich natürlich einwenden, dass die soeben ausgeführten Überlegungen nur unter der Annahme haltbar sind, dass ein Fehler ausgeschlossen werden kann. Das heisst, aufgrund irgendeiner technischen Panne sei allfällig vorhandene DNA des Opfers der Detektion entgangen. Somit wäre der Befund «fehlende übereinstimmende DNA», unter der Annahme der Hypothese der Anklage – wonach D. der Täter ist – genauso gut vereinbar mit der Hypothese der Verteidigung, und daher insgesamt nicht mehr entlastend, sondern neutral. Einer solchen Argumentationsweise muss jedoch aus zwei Gründen skeptisch begegnet werden. Erstens ist festzuhalten, dass die Fehlerproblematik erst *post hoc* eingebracht wird, d. h. nachdem das Untersuchungsergebnis vorgelegen hatte. Die Tendenz, Hypothesen im Lichte der Ergebnisse «anzupassen» oder eine «Erklärung»¹² für das Ergebnis zu «konstruieren», wird auch als *post-hoc*-Rationalisierung bezeichnet¹³ und ist zu vermeiden.

11 Für eine vertiefte Besprechung siehe beispielsweise D. Taylor, The Evaluation of Exclusionary DNA Results: A Discussion of Issues in *R v. Drummond*, Law, Probability & Risk 15(3)/2016, 175.

12 Zur Unterscheidung der Begriffe Hypothese und Erklärung siehe auch Abschnitt V.2 des ersten Teils dieses Beitrags.

13 Siehe z. B. G. Jackson/C. Aitken/P. Roberts, Case Assessment and Interpretation of Expert Evidence: Guidance for Judges, Lawyers, Forensic Scientists and Expert Witnesses, Practitioner Guide No. 4, Royal Statistical Society, 2014, abrufbar unter: <https://rss.org.uk/RSS/media/File-library/Publications/rss-case-assessment-interpretation-expert-evidence.pdf> (besucht am: 25. 1. 2023).

Zweitens ist Kohärenz gefordert: Denn nimmt man an, was in der Praxis derzeit dauerlicherweise der Regelfall ist, nämlich dass das Fehlerrisiko vernachlässigbar ist, wenn die DNA den Verdächtigen belastet, so muss man auch annehmen, dass das Fehlerrisiko vernachlässigbar ist, wenn die DNA den Verdächtigen entlastet!

IV. Das ungelöste Fehlerproblem

Für vollständige DNA-Profile ergeben sich bekanntlich leicht Seltenheitswerte – und somit Merkmalswahrscheinlichkeiten – in der Grössenordnung von eins zu mehreren Milliarden. Bei Berücksichtigung von Hypothesen zur Spurenherkunft können daraus, so scheint es, mühelos Beweiswerte im Milliardenbereich resultieren. Diese Darstellung von DNA-Befunden täuscht jedoch über die Tatsache hinweg, dass die weitherum angepriesenen Zahlenwerte im Millionen- und Milliardenbereich, und darüber hinaus, lediglich künstliche Charakterisierungen des Beweiswerts darstellen, weil sie die Möglichkeit eines Fehlers ignorieren. Es drängt sich daher unter anderem die Frage auf, wie sinnvoll es ist, in Sachverständigengutachten Ergebnisse in der erwähnten Grössenordnung auszuweisen. Um dieser Frage nachzugehen, befassen wir uns im vorliegenden Abschnitt zunächst mit einer Begriffseingrenzung und, soweit hierzu Kenntnisse existieren, mit der Häufigkeit von Fehlern. Anschliessend gehen wir darauf ein, wie mit der Fehlerproblematik bei der Befundbewertung umzugehen ist, d. h. wie – aus rationaler Sicht – das Fehlerpotenzial den Beweiswert beeinflusst.

1. Zum Begriff des Fehlers und zu der Häufigkeit von Fehlern

Die Fehlerproblematik ist grundsätzlich keine Unbekannte, sollte sie doch auch der breiten Öffentlichkeit durch hoch mediatisierte Fälle wie das «Phantom von Heilbronn»¹⁴ oder dem australischen Fall «Farah Jama»¹⁵ bekannt sein. Schreib-

14 Siehe z. B. «Der Spiegel», Ermittlungsspanne, «Phantom-Mörderin» ist ein Phantom, 27. 03. 2009, abrufbar unter: <https://www.spiegel.de/panorama/justiz/ermittlungsspanne-phantom-moerderin-ist-ein-phantom-a-615969.html> (besucht am: 25. 1. 2023).

15 Farah Jama wurde 2008 in Australien wegen Vergewaltigung verurteilt aufgrund einer Probe, die im Genitalbereich einer Frau erhoben wurde, die bewusstlos in der Toilette eines Nachtclubs aufgefunden worden war und sich an nichts mehr erinnern konnte. In der Probe wurde DNA festgestellt, die mit Jama übereinstimmte. Jama wurde 2009 im Berufungsverfahren freigesprochen. Eine Untersuchungskommission stellte später fest, dass die Spur, die den Verurteilten belastete, durch eine Übertragung von Jamas DNA auf das Spurensicherungsmaterial entstand, welches am Opfer zur Anwendung kam. Siehe F. H. R. Vincent, Inquiry into the circumstances that led to the conviction of Mr Farah Abdulkadir Jama, Victoria Department of Justice 2010.

fehler, Materialübertragungen zwischen Proben, das Vertauschen von Proben, technische Pannen usw. sind alles in einem gewissen Sinne «Fehler». Dennoch ist die allgemeine Studienlage zum Fehlerproblem derzeit noch sehr dürftig. Beispielsweise befassen sich manche Studien lediglich mit einer breit definierten Fehlerquote, d. h. ohne Unterscheidung zwischen unterschiedlichen Fallkonstellationen wie Materialübertragung zwischen Proben, Verunreinigungen durch Personal am Tatort oder im Labor usw. Solche Vorkommnisse, wenn auch unerwünscht, sind für die Verfahrensleitung und das Gericht jedoch nur von begrenztem Interesse. Eine Verunreinigung durch Laborpersonal lässt sich beispielsweise relativ leicht erkennen, sofern das DNA-Profil der Person im sog. «Staff-Index»¹⁶ gespeichert ist. Die tatsächlich kritischen Vorfälle sind hingegen häufig kaum erkennbar, wenn nicht (un-)glückliche Umstände hinzutreten.

Diese Erfahrung machte Lukis Anderson, der zum Tatzeitpunkt nachweislich im Spital in Behandlung war, nachdem er zuvor in der Innenstadt von San Jose betrunken umgefallen war.¹⁷ Unter den Fingernägeln eines Mordopfers – das in derselben Nacht starb – wurde ein DNA-Profil festgestellt, das mit dem DNA-Profil von Anderson übereinstimmte. Anschliessend konnte rekonstruiert werden, dass dieselben Sanitäter, die Anderson ins Spital überführten, einige Stunden später versuchten, am Tatort das Opfer wiederzubeleben, und dabei Andersons DNA auf das Opfer übertrugen.

Angesichts der höchst idiosynkratischen Natur von Einzelfällen scheint es daher weder möglich noch sinnvoll, eine allgemeine und abstrakte Fehlerquote, d. h. bezogen auf den gesamten Fachbereich der forensischen Genetik, zu erheben. Doch selbst wenn sich ein berichterstattendes Labor nur auf das Fehlerpotenzial seiner eigenen internen Prozesse konzentriert, ergeben sich Schwierigkeiten. Denn abgesehen von Qualitätsunterschieden zwischen Laboren, kann das Qualitätsniveau im Laufe der Zeit innerhalb ein und desselben Labors variieren. Das Fehlerpotenzial ist somit permanent zu überwachen, was mit erheblichem Aufwand verbunden ist. Laboren sollte dies jedoch nicht als Vorwand dienen, sich der Fehlerproblematik insgesamt zu verweigern.

Erwähnenswert in diesem Zusammenhang ist eine von der Abteilung für biologische Spuren des niederländischen forensischen Instituts (NFI) durchge-

16 Der «Staff-Index» bezeichnet eine zu Qualitätssicherungszwecken, vom DNA-Informationssystem getrennt geführte Sammlung von DNA-Profilen von tatortberechtigten Personen im Strafverfahren (Art. 11 der DNA-Profil-Verordnung vom 3. Dezember 2004, SR 363.1). Zu einer Studie zum Auftreten von Verunreinigungen durch Polizei- und Laborpersonal in der Schweiz siehe P. Basset/V. Castella, *Lessons from a Study of DNA Contaminations From Police Services and Forensic Laboratories in Switzerland*, *Forensic Science International: Genetics* 33/2018, 147.

17 Siehe u. a. The Marshall Project, <https://www.themarshallproject.org/2018/04/19/framed-for-murder-by-his-own-dna> (besucht am: 25. 1. 2023).

führte Studie. Die Arbeit befasst sich mit den im Zeitraum der Jahre 2008 bis 2012 beim NFI festgestellten Fehlern, einschliesslich deren tatsächlichen und potenziellen Auswirkungen, sowie dem Umgang mit Fehlerzahlen in der Kommunikation mit der Justiz.¹⁸ Es handelt sich im Übrigen um eine der wenigen Studien, in welcher die Autoren die Fehlerhäufigkeit nicht nur quantitativ zu erfassen versuchten, sondern die Ergebnisse anschliessend auch veröffentlicht haben. Die dabei ausgewiesene Fehlerquote liegt zwischen 0,3 und 0,4%.¹⁹ Zu einem ähnlichen Ergebnis kommt ein aktueller Bericht des Home Office des Vereinigten Königreichs, in welchem, je nach Fehlerursache, die Rate erkannter Fehler mit bis zu 0,2% beziffert wird.²⁰

2. Warum dem Fehlerpotenzial Rechnung zu tragen ist

Die Fachliteratur, insbesondere die der forensischen Statistik, hat wiederholt dargelegt,²¹ dass das Fehlerpotenzial den Beweiswert wie folgt beeinflusst: je kleiner die Wahrscheinlichkeit der zufälligen Merkmalsübereinstimmung, im Vergleich zur Wahrscheinlichkeit eines Fehlers, desto geringer der Einfluss des Seltenheitswerts auf den Beweiswert. Das heisst, vereinfacht gesagt, die Fehlerwahrscheinlichkeit rückt in den Vordergrund.

Konkret ergibt sich daraus, dass in Fällen, in welchen die Fehlerwahrscheinlichkeit in der Grössenordnung von maximal eins zu ein paar Tausend (oder Zehntausend) angesetzt werden muss – denn realistischlicherweise kann sie kaum im Be-

18 A. D. Kloosterman/M. Sjerps/A. Quak, *Error Rates in Forensic DNA Analysis: Definition, Numbers, Impact and Communication*, *Forensic Science International: Genetics* 12/2014, 77.

19 Kloosterman/Sjerps/Quak (Fn. 18), 80. Es ist anzumerken, dass sich diese Fehlerrate auf alle im System erfassten Fehlermeldungen bezieht; daraus folgt nicht, dass jeder dieser Fälle negative Auswirkungen auf das betreffende Strafverfahren hatte oder ein konkretes Risiko eines Justizirrtums darstellte. Die Autoren halten ausserdem fest, dass die ermittelte Fehlerrate zum Zeitpunkt ihrer Studie mit den Fehlerraten verschiedener medizinischer und/oder genetischer Labore vergleichbar ist.

20 *Home Office/National Police Chiefs' Council*, *Forensic Information Databases, Strategy Board, Annual Report 2020–2021*, April 2022, Table 4, 25.

21 Interessierte Leser finden eine zugängliche Darstellung der mathematischen Erschliessung der Fehlerproblematik in W. C. Thompson/F. Taroni/C. Aitken, *How the Probability of a False Positive Affects the Value of DNA Evidence*, *Journal of Forensic Sciences* 48/2003, 47. Siehe auch D. A. Berry, *DNA Fingerprinting: A Review of the Controversy: Comment*, *Statistical Science*, 9(2)/1994, 252; J. J. Koehler, *Why DNA Likelihood Ratios Should Account for Error (Even When a National Research Council Report Says They Should Not)*, *Jurimetrics* 37/1997, 425; J. J. Koehler/A. Chia/S. Lindsey, *The Random Match Probability in DNA Evidence: Irrelevant and Prejudicial?* *Jurimetrics* 35/1995, 201.

reich von eins zu einer Milliarde oder kleiner liegen²² – auch der Beweiswert nur höchstens im Bereich von ein paar Tausend (oder Zehntausend) liegen kann. Zabell resümiert daher treffend, «dass letztlich die Beweiskraft durch die Möglichkeit einer irrtümlichen Assoziierung begrenzt wird, sei es aufgrund eines Schreibfehlers, eines versehentlichen Austauschs von Proben oder was auch immer» und konstatiert: «Leider kann die sirenenhafte Anziehungskraft vermeintlich astronomisch kleiner Wahrscheinlichkeiten oft dazu führen, dass man ihre praktischen Grenzen übersieht.»²³

Es ist daher keineswegs so, dass nicht bekannt wäre, wie das Fehlerpotenzial mit der Merkmalswahrscheinlichkeit zu verknüpfen ist und welche Bedeutung dem Fehlerpotenzial in der Befundbewertung zukommt. Das Fehlerpotenzial ist sozusagen «der Elefant im Raum».²⁴ Dennoch beschränkt sich die forensische Genetik derzeit noch weitgehend auf die vergleichsweise einfach zu quantifizierende – in der Befundbewertung jedoch sekundäre – Merkmalswahrscheinlichkeit. Die einschlägig in Sachverständigengutachten angegebenen Ergebnisse sind somit lediglich als künstliche, sprich realitätsfremde Charakterisierungen des Beweiswertes anzusehen. Für eine realistische Befundbewertung ist es daher unerlässlich, sich der Fehlerproblematik nicht länger zu verschliessen.

Die Fehlerproblematik im Zusammenhang mit forensischen DNA-Profilanalysen findet trotz deren nicht von der Hand zu weisenden Wichtigkeit, abgesehen von einigen wenigen akademischen Beiträgen, in der Praxis so gut wie keine Beachtung. Sogar die Empfehlungen der DNA-Kommission der *International Society for Forensic Genetics* geben dies mit folgendem Hinweis offen zu: «Es wurde davon ausgegangen, dass der DNA-Prozess fehlerfrei war, was, wie wir wissen, nicht der Fall ist. Dieser Aspekt ist nicht ideal und macht es umso wichtiger, zu betonen, dass wir nur über den Wert der Ergebnisse unter unseren Annahmen berichten.»²⁵

Nach unserem Dafürhalten sollte die Fehlerproblematik *proaktiv* von jenen Kreisen angegangen werden, von welchen die Problematik in erster Linie ausgeht: den mit der Untersuchung und der Analyse betrauten Dienststellen und -leistern (DNA-Analyselaboren). Das Argument, es stehe jeder verfahrensbeteiligten Partei

22 Das Problem ist empirischer Natur, denn kein Labor kann behaupten, es habe bereits Milliarden von Analysen unter kontrollierten Bedingungen vorgenommen und dabei die Fehlerquote erfasst. Beispielsweise wird im oben erwähnten Bericht des Home Office (Fn. 20) die Fehlerquote der forensischen Dienstleister bei der Interpretation von DNA-Profilen von Verdächtigen mit ca. 0,004% angegeben.

23 S. L. Zabell, Book Review: Statistical DNA Forensics: Theory, Methods and Computation, by Wing Kam Fung and Yue-Qing Hu, Law, Probability and Risk 11/2012, 105, 109 (frei übersetzt).

24 P. Gill, Misleading DNA Evidence, Reasons for Miscarriages of Justice, London 2014, 177.

25 P. Gill et al., Assessing the Value of Forensic Biological Evidence – Guidelines Highlighting the Importance of Propositions: Part I: Evaluation of DNA Profiling Comparisons Given (Sub-)Source Propositions, Forensic Science International Genetics 36/2018, 189, 201 (frei übersetzt).

doch grundsätzlich offen, vor Gericht die Fehlerproblematik bei Bedarf aufzugreifen, überzeugt hingegen nicht. Es genügt, sich vor Augen zu führen, dass vielen Prozessbeteiligten die Materie schlicht zu wenig bekannt ist. Zudem fehlen ihnen häufig die für die Beurteilung der Integrität des Untersuchungsprozesses notwendigen Hintergrundinformationen, entweder weil sie unzulänglich dokumentiert wurden oder nicht ohne Weiteres zugänglich sind. Der Status quo ist auch insofern unbefriedigend, als dass durch die Ausklammerung der Fehlerproblematik der falsche Eindruck entstehen kann, die Wahrscheinlichkeit eines Fehlers sei insgesamt vernachlässigbar gering.²⁶

Die Rechtspraxis in der Schweiz ist von einem offenen Umgang mit Fehlern leider noch weit entfernt. Obwohl akkreditierte Labore bereits jetzt Konformitätsverletzungen dokumentieren müssen, sind den in Strafverfahren involvierten Parteien sowie der breiteren Öffentlichkeit so gut wie keine Informationen zur Fehlerproblematik in den jeweiligen DNA-Analyselaboren zugänglich.²⁷ Es lässt sich daher auch nur schwerlich oder gar nicht beurteilen, wie viel Glauben einem bestimmten Untersuchungsergebnis zu schenken ist.²⁸ Insofern ist es bedauerlich, dass der Gesetzgeber es in der laufenden Reform des DNA-Profilgesetzes einmal mehr versäumt hat, sich der Fehlerproblematik anzunehmen.

3. Zum Umgang mit dem Fehlerpotenzial

Die Fehlerproblematik sollte nicht dahingehend verstanden werden, dass es ausreiche, vor Gericht zu behaupten, es bestünde ein allgemeines Fehlerpotenzial, aufgrund dessen Ergebnisse von DNA-Profilanalysen per se als gegenstandslos zu betrachten seien.²⁹ Vielmehr muss fallbezogen argumentiert werden, d. h., die Möglichkeit eines Fehlers muss sich substantiieren lassen. Parteien, welche die Fehlerproblematik aufzuwerfen gedenken, sind daher gut beraten, kritischen Fragen vorzugreifen und die für eine Konkretisierung des Fehlerpotenzials im Kontext des vorliegenden Falles notwendigen Argumente bereitzustellen.

26 Um vernachlässigbar zu sein, müsste die Fehlerwahrscheinlichkeit rein rechnerisch *kleiner* sein als die Merkmalswahrscheinlichkeit, was gemäss dem derzeitigen Kenntnisstand nicht empirisch dargelegt werden kann und somit als nicht plausibel anzusehen ist.

27 Das Gleiche gilt für sämtliche anderen forensischen Untersuchungsbereiche (z. B. Daktyloskopie, Werkzeugspuren, Handschriftenuntersuchung, digitale/elektronische Spuren usw.).

28 A. Biedermann/J. Vuille, Bewertung von DNA-Untersuchungsergebnissen aus der Sicht von Gerichten und Sachverständigen: Wie viel von unserer Wahrnehmung können wir «für wahr nehmen»? ZStrR 2011, 278 ff.

29 Siehe z. B. A. Biedermann/J. Vuille/F. Taroni, Apprécier le risque d'erreur lors d'une analyse ADN: de la nécessité d'être concret, AJP 2013, 1217.

Um der Möglichkeit eines Fehlers Rechnung zu tragen, ist es hingegen nicht notwendig, Fehlerwahrscheinlichkeiten exakt zu quantifizieren. Wie im Unterabschnitt IV.2 erwähnt ist es ausreichend, die Grössenordnung der Wahrscheinlichkeit eines Fehlers im Vergleich zum Seltenheitswert der Merkmalsübereinstimmung zu beurteilen. Dies ist umso leichter, je kleinere – und daher empirisch ohnehin immer schwieriger zu rechtfertigende – Merkmalswahrscheinlichkeiten im Sachverständigengutachten proklamiert werden.³⁰ Anschliessend ergeben sich zwei Optionen:

1. Man kann sich auf keine Fehlerwahrscheinlichkeit festlegen, auch nicht in einer qualitativen Grössenordnung: In einem solchen Fall kann nichts über den Beweiswert gesagt werden. Damit müsste das Ergebnis der DNA-Profilanalyse als uninterpretierbar und somit als nicht verwertbar bezeichnet werden.
2. Man ist gewillt, für die Fehlerwahrscheinlichkeit zumindest eine Grössenordnung anzunehmen: In einem solchen Fall lässt sich der Beweiswert zu einem gewissen Grad ausloten. Wie wir gesehen haben, kann realistischerweise jedoch nur schwerlich argumentiert werden, die Fehlerwahrscheinlichkeit sei bedeutend kleiner als etwa eins zu 10 000; dies nicht zuletzt aufgrund der Tatsache, dass der Analyseprozess auf fehlbaren menschlichen Arbeitsvorgängen beruht. Aus dieser Annahme ergibt sich, dass sich der Beweiswert absehbar auf ungefähr den Kehrwert der Fehlerwahrscheinlichkeit reduziert, d. h. die Grössenordnung von 10 000, sofern die Wahrscheinlichkeit der zufälligen Merkmalsübereinstimmung im Vergleich dazu (verschwindend) klein ist.³¹

Zusammenfassend können wir somit festhalten, dass nicht die «scheinbar astronomisch kleinen»³² Merkmalswahrscheinlichkeiten den Beweiswert bestimmen, sondern die Fehlerwahrscheinlichkeit. Das Gegenteil zu behaupten, ohne die Fehlerproblematik zuerst ausgelotet zu haben, kann daher zu übertriebenen und irreführenden Schlussfolgerungen bezüglich der Beweiskraft führen. Bei den derzeit standardmässig erstellten Gutachten ist dies regelmässig der Fall. Die Schlussfolgerungen in solchen Gutachten sind somit insofern als *hypothetisch* zu betrachten, als dass sie sich dazu äussern, was der Beweiswert *wäre*, sofern der Untersuchungsprozess immer fehlerfrei verlaufen *würde*. Gerade davon darf aber nicht ohne Weiteres ausgegangen werden, weil Fehler grundsätzlich unvermeidbar sind.

30 Siehe z. B. den Wert von 1:256 Milliarden in BGH 1 StR 722/08, 2.

31 Siehe beispielsweise *Thompson/Taroni/Aitken* (Fn. 21).

32 *Zabell* (Fn. 23), 109.

V. Das Verständnis der Individualisierung als Entscheidungsproblem und die richterliche Überzeugung

Sowohl in der Praxis als auch im Schrifttum hält sich hartnäckig die Auffassung, forensische DNA-Profile seien «individualspezifisch»³³ bzw. «einzigartig»³⁴ und somit individualisierend, d. h. einer einzigen Person zuweisbar.³⁵ Diese Sichtweise ist jedoch irreführend, denn wenn forensische DNA-Profile einzigartig und individualisierend wären, dann stellt sich die Frage, weshalb irgendwelche Berechnungen des Beweiswertes hinsichtlich Hypothesen der Spurenherkunft in Sachverständigengutachten überhaupt notwendig sind.

Rufen wir uns in Erinnerung: DNA-Profile bestehen aus einer Kombination von Merkmalsausprägungen an verschiedenen Positionen der menschlichen DNA. Der Seltenheitswert eines DNA-Profiles wird dabei nicht direkt bestimmt, d. h., es wird nicht überprüft, wie häufig ein DNA-Profil tatsächlich in einer bestimmten Population vorkommt. Es werden lediglich die Merkmalswahrscheinlichkeiten für die einzelnen Genorte bestimmt und anschliessend mithilfe eines populationsgenetischen Modells miteinander verknüpft. Dadurch entstehen die sog. DNA-Profilwahrscheinlichkeiten. Die weitläufig gepriesene Seltenheit von forensischen DNA-Profilen beruht somit in erster Linie auf statistischen Überlegungen und nicht auf einer direkten empirischen Grundlage. Doch genau hier liegt das fundamentale Problem: Mit Statistik (alleine) lässt sich nicht individualisieren!³⁶ Das heisst, die angebliche Einzigartigkeit eines DNA-Profiles wird nicht bewiesen; bestenfalls kann man sich von der Einzigartigkeit überzeugen angesichts errechneter zunehmend kleinerer Wahrscheinlichkeiten der festgestellten Merkmalskombination (DNA-Profil). Im Weiteren gilt es zu beachten, dass das Endergebnis solcher Berechnungen, entgegen weitläufiger Meinungen, keineswegs objektiv ist: wie Evett und Weir treffend festhalten, ist «Objektivität (...) eine Illusion, weil sie nur innerhalb eines Rahmens von Annahmen existiert».³⁷ Diese Annahmen sind, wie wir in den vorhergehenden Abschnitten dargelegt haben, vielfältig, angefangen bei der unzutreffenden Auffassung, dass der gesamte Untersuchungsprozess immer fehlerfrei verläuft.

33 BBl 2021 44, 10; im gleichen Sinne siehe auch Art. 2 Abs. 1 nDNA-Gesetz.

34 Siehe z. B. DNA-Analysen im Strafverfahren, BBl 2019 7149.

35 Siehe z. B. auch den erläuternden Bericht zur Eröffnung des Vernehmlassungsverfahrens (Fn. 2): «Zeigt sich beim Vergleich (...) eine Übereinstimmung, so ist die Spurengerbin oder der Spurengerbe dadurch identifiziert.» (S. 9).

36 D. A. Stoney, What Made Us Ever Think We Could Individualize Using Statistics? *Journal of the Forensic Science Society* 31(2)/1991, 197.

37 I. W. Evett/B. S. Weir, *Interpreting DNA Evidence*, Sunderland (Massachusetts) 1998, 240.

Die offene Frage bleibt daher: Wie kann man eine Spur unbekannter Herkunft einer bestimmten Person argumentativ stichhaltig zuordnen? Im einschlägigen Schrifttum hat sich nunmehr die Erkenntnis durchgesetzt, dass die Individualisierung – d. h. die Zuordnung einer Spur zu einer bestimmten Person – eine *persönliche Entscheidung* ist.³⁸ Das heisst konkret, es geht um die grundsätzliche Frage, wie sich eine Entscheidungsinstanz von einer bestimmten Hypothese der Spurenherkunft auf der Grundlage von Daten (hier: forensisch-wissenschaftlichen und nicht wissenschaftlichen Informationen) überzeugen kann, welche die besagte Hypothese insgesamt nicht schlüssig unterstützen. Dabei ist gleichzeitig zu beachten, dass sich die Konsequenzen der Entscheidung als faktisch richtig oder falsch herausstellen können.

Die logische Struktur dieses Entscheidungsproblems entspricht weitgehend jener, mit welcher ein Gericht konfrontiert wird, das über die Schuldfrage einer bestimmten beklagten Person befinden muss. Die Anatomie dieser Kategorie von Entscheidungsproblemen lässt sich mit der normativen (d. h. rationalen) Entscheidungstheorie untersuchen bzw. darlegen.³⁹ Wir können die wesentlichen Eckpunkte des entscheidungstheoretischen Ansatzes, angewandt auf die Individualisierungsproblematik, wie folgt zusammenfassen:⁴⁰ Die *Entscheidung*, eine bestimmte Hypothese der Spurenherkunft (z. B.: Die Hypothese «die Tatortspur stammt von X.») *als erwiesen anzusehen*, hängt von zwei Faktoren ab:

- erstens vom Grad, mit welchem die Entscheidungsinstanz von der Hypothese der Spurenherkunft, unter der Berücksichtigung aller zum Zeitpunkt der Entscheidungsfindung vorliegenden relevanten Informationen (einschliesslich des DNA-Analyseergebnisses), überzeugt ist;
- zweitens von der Einordnung der Entscheidungsfolgen nach deren Vorzuehenswürdigkeit, d. h. den Präferenzen der Entscheidungsinstanz hinsichtlich der Konsequenzen der einzelnen Handlungsalternativen. Von besonderem Interesse ist hierbei, vereinfacht ausgedrückt, wie «schwerwiegend» eine falsche Individualisierung ist im Vergleich zu einer falschen Nichtindividualisierung.

38 S. A. Cole/A. Biedermann, How Can a Forensic Result Be a «Decision»? A Critical Analysis of Ongoing Reforms of Forensic Reporting Formats For Federal Examiners, *Houston Law Review*, 57/2020, 551; A. Biedermann/S. Bozza/F. Taroni, Decision Theoretic Properties of Forensic Identification: Underlying Logic and Argumentative Implications, *Forensic Science International* 177 (2–3)/2008, 120.

39 J. Kaplan, Decision Theory and the Factfinding Process, *Stanford Law Review* 20(6)/1968, 1065–1092; D. H. Kaye, Clarifying the Burden of Persuasion: What Bayesian Decision Rules Do and Do Not Do, *International Journal of Evidence and Proof* 3/1999, 1.

40 Siehe auch A. Biedermann/J. Vuille, Understanding the logic of forensic identification decisions (without numbers), *sui generis* 2018, 397.

Als entscheidungstheoretische Handlungsanweisung ergibt sich anschliessend, dass je schwerwiegender (d. h. unerwünschter) eine irrtümliche Individualisierung (d. h., falsche Spurenzusammenfassung) ist im Vergleich zu einer irrtümlichen Nichtindividualisierung, umso *stärker* sollte die Entscheidungsinstanz von der Hypothese überzeugt sein, dass die Spur von X. stammt.

Dieses Ergebnis mag banal erscheinen, gibt aber einen grundlegenden Aufschluss über das Problem der Individualisierung. Für eine rationale Begründung einer Individualisierungsentscheidung ist demzufolge weder *zwingend absolute Gewissheit notwendig*, noch geht es lediglich nur um den individuellen Überzeugungsgrad; vielmehr geht es letztlich auch um Werturteile. Diese beiden Aspekte – der individuelle Grad der Gewissheit und Werturteile – entziehen sich jedoch dem Kompetenzbereich des Sachverständigengutachtens.⁴¹ Die weit verbreitete Auffassung, Sachverständigengutachten könnten selbstständig Spuren unbekannter Herkunft *eindeutig* bestimmten Personen zuordnen, d. h., man könnte bei der Individualisierungsentscheidung alleine auf das Sachverständigengutachten abstellen oder sogar die Entscheidungshoheit an Sachverständige insgesamt delegieren, stellt daher ein fundamentales Missverständnis des Wesens der Entscheidung unter Unsicherheit sowie der Aufgabenteilung zwischen Sachverständigen und Gerichten dar. Anders ausgedrückt, Sachverständige, die vorgeben, sie könnten kategorische Schlussfolgerungen zu Hypothesen der Spurenherkunft ziehen, überspannen den Bogen des wissenschaftlich Leistbaren und entlarven sich selbst als unwissenschaftlich.⁴²

Ein notorisches Beispiel für die Missachtung der konzeptuellen Grundlagen der Individualisierung sowie der Zuweisung der Entscheidungskompetenz in Strafverfahren sind die Empfehlungen der deutschen Projektgruppe «Biostatistische DNA-Berechnungen» und der Spurenkommission.⁴³ Sowohl in deren aktuellen als auch in der vorherigen Fassung wird ein «Schwellenwert zur statistisch begründbaren Zuordnung von Spuren» ausgewiesen – konkret: ein Likelihood-Quotient von 30 Milliarden. Wenn dieser Wert überschritten werde, könne «eine Befundinterpretation im Sinne von *«Aus gutachterlicher Sicht besteht somit kein begründeter Zwei-*

41 K. Kotsoglou/A. Biedermann, Inroads into the Ultimate Issue Rule? Structural Elements of Communication Between Experts and Fact Finders, *The Journal of Criminal Law* 86/2022, 223.

42 Siehe auch D. A. Stoney, Discussion on the Paper by Neumann, Evett and Skerrett, *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 175/2012, 399.

43 W. Ulbrich et al., Gemeinsame Empfehlungen der Projektgruppe «Biostatistische DNA-Berechnungen» und der Spurenkommission zur biostatistischen Bewertung von DNA-analytischen Befunden, *Rechtsmedizin*, 2016, 291; M. Hahn et al., Gemeinsame Empfehlungen der Projektgruppe «Biostatistische DNA-Berechnungen» und der Spurenkommission zur biostatistischen Bewertung forensischer DNA-analytischer Befunde mit vollkontinuierlichen Modellen (VKM), *Rechtsmedizin* 2022; <https://doi.org/10.1007/s00194-022-00599-5>.

fel, dass die Merkmale der Spur von Person X (ggf. und . . .) stammen», folgen»,⁴⁴ d. h. eine Individualisierung. Die Wortwahl in dieser Schlussfolgerungsempfehlung ist höchst problematisch – und dies aus mehreren Gründen:

Erstens ist festzuhalten, dass sich die Aussage nicht auf die Befunde beschränkt, sondern – entgegen den Prinzipien der Befundbewertung⁴⁵ – sich direkt auf eine Hypothese (hier: der Spurenherkunft) bezieht.

Zweitens ist zu beachten, dass der Ausdruck «kein begründeter Zweifel» jeglicher argumentativer Grundlage entbehrt: er enthält nämlich *keine* Begründung, sondern stellt vielmehr eine reine Behauptung dar. Wie wir an anderer Stelle dargestellt haben,⁴⁶ handelt es sich bei Individualisierungen (Spurenuordnungen) um *Entscheidungen*, deren Konsequenzen nicht direkt den Entscheidungsträger selbst betreffen, sondern *Dritte*. Was ein «begründeter Zweifel» angesichts solcher Entscheidungskonsequenzen darstellt, ist keine wissenschaftliche Frage, sondern erfordert ein Werturteil, was sich dem Kompetenzbereich von Sachverständigen klar entzieht.⁴⁷ Angesichts der Schwierigkeiten, welche die Rechtslehre und die Praxis gegenüber dem Heiligen Gral der richterlichen Überzeugungsbildung – dem Beweismass – bekunden, vor allem dessen Quantifizierung mittels Wahrscheinlichkeiten, wäre es im Übrigen erstaunlich, wenn die forensische Genetik tatsächlich den Schwellenwert für den Grad der Gewissheit «entdeckt» hätte, auf dessen alleiniger Grundlage eine Individualisierungsentscheidung (Spurenuweisung) begründet werden könnte.

Abschliessend ist noch hervorzuheben, dass die Formulierung «statistisch begründbare Zuordnung» schlicht irreführend ist, weil die Statistik – wie wir bereits erwähnt haben – die Individualisierung genau *nicht* begründet.⁴⁸ Die Verwendung des Begriffs «Statistik» erfolgt somit missbräuchlich und verkommt zu nichts anderem als der Anmassung, dem Gericht unter dem Deckmantel angeblicher Wissenschaftlichkeit mittels eines willkürlich festgelegten Schwellenwerts eine Schlussfolgerung aufzuzwingen.

44 Ulbrich et al. (Fn. 43), 296.

45 Siehe 1. Teil, Abschnitt III.

46 Biedermann/Vuille (Fn. 40).

47 Selbst in den Vereinigten Staaten, gemeinhin bekannt für markante Sachverständigenaussagen, untersagt ein Memorandum des Generalstaatsanwalts seit 2016 die Verwendung von Formulierungen wie «angemessene wissenschaftliche Gewissheit» oder «angemessene [forensisch gutachterliche] Gewissheit» auf Bundesebene (Office of the Attorney General, Memorandum for Heads of Department Components, September 6, 2016, abrufbar unter: <https://www.justice.gov/opa/file/891366/download>, [besucht am: 31. 1. 2023]). Im gleichen Sinne siehe auch U.S. Department of Justice, Uniform Language for Testimony and Reports for Forensic Autosomal DNA Examinations Using Probabilistic Genotyping Systems, vers. 09/13/2022, abrufbar unter: <https://www.justice.gov/media/969526/dl?inline> (besucht am: 22. 1. 2023).

48 Stoney (Fn. 36).

VI. Mischspuren, probabilistische Genotypisierung und die neue Ära des «maschinellen Beweismittels» («Machine Evidence»)

Die Bedeutung der im vorhergehenden Unterabschnitt behandelten Frage der kategorischen Spurenuordnung, die sog. Individualisierung, ist tendenziell abnehmend, nicht zuletzt wegen der zunehmenden Anzahl von (komplexen) DNA-Mischspuren, d. h. Spuren, in welchen sich DNA von mehr als einer Person befindet, in zuweilen unterschiedlicher Menge und variabler Qualität. Denn anders als bei herkömmlichen «einfachen» Spuren, die DNA von nur einer Person enthalten und ein vollständiges DNA-Profil ergeben, sind gewisse (unvollständige) DNA-Mischprofile häufig sogar für Sachverständige nur sehr schwierig, wenn überhaupt, interpretierbar.

Bei komplexen DNA-Profilen ergeben sich gleich mehrere praktische Herausforderungen. Eine weit verbreitete Schwierigkeit besteht beispielsweise nur schon darin, Elektropherogramme – die grafische Darstellung von DNA-Analyseergebnissen – «richtig» zu lesen: denn wenn DNA nur in sehr geringer Menge vorhanden ist, kann die grafische Darstellung der gemessenen Signalintensitäten, in der Form von sog. «Peaks», sehr schwach oder gänzlich ausfallen; umgekehrt können zufällige Artefakte auftreten, die fälschlicherweise für detektierte Merkmale angesehen werden können. Häufig besteht auch Unsicherheit darüber, von wie vielen Personen überhaupt DNA vorliegt und in welcher Menge (das sog. Mischverhältnis). Schliesslich ist auch zu bedenken, dass die detektierbaren DNA-Merkmale aller Personen, von welchen DNA in der Spur vorhanden sind, in *demselben* Elektropherogramm erscheinen, weshalb sich in der Regel gar nicht sagen lässt, welches Merkmal welcher Person zuzuordnen ist.⁴⁹

Angesichts dieser Schwierigkeiten ist von der klassischen Individualisierung in der Praxis gar nicht mehr zu sprechen. Der Bewertung komplexer DNA-Mischprofile werden nämlich weitaus diffusere Hypothesen zugrunde gelegt: Es geht dabei nicht mehr um die eng umrissene Frage, ob die detektierte DNA von einer bestimmten Person stammt, sondern lediglich darum, ob eine bestimmte Person zum *Kreis* der Personen gehört, von welchen DNA detektiert wurde. Grundsätzlich ist es zwar möglich, den Beweiswert des Ergebnisses der DNA-Profilanalyse hinsichtlich solcher Hypothesen ebenfalls in der Form eines Likelihood-Quotienten zu charakterisieren. Die hierzu notwendigen Berechnungen erweisen sich in der Praxis jedoch als so komplex, dass sie sich von Hand kaum mehr zuverlässig vornehmen lassen. Dies hat zur Entwicklung computergestützter, sog. probabilistischer Geno-

49 Wie im nächsten Absatz erwähnt existieren mittlerweile rechenintensive, computergestützte Verfahren für die «Entflechtung» (engl. «Deconvolution») von DNA-Mischprofilen. Doch auch die Ergebnisse solcher Verfahren sind bestenfalls nur Wahrscheinlichkeitsaussagen.

typisierungssysteme⁵⁰ (PG, «probabilistic genotyping») geführt, allen voran das System *STRmix* des neuseeländischen Crown Research Instituts und des Labors Forensic Science South Australia.

Der Einsatz solcher Systeme wirft eine Reihe komplizierter Fragen auf, die derzeit vor allem in den USA heftig diskutiert werden.⁵¹ Ein kritischer Punkt bei der Verwendung eines PG-Systems besteht darin, dass Sachverständige praktisch die *gesamte* Beweiswertberechnung einer Maschine anvertrauen. Im Endeffekt reduziert sich der im Sachverständigenbericht vorgelegte Beweiswert somit auf die direkte Wiedergabe eines maschinellen Rechenergebnisses. Vereinfacht gesagt ist die gutachterliche Aussage im Prinzip eine Computerausgabe, was sich im Begriff «Machine testimony»⁵² widerspiegelt. Das ist insofern problematisch, als dass sich die komplexen Berechnungen von PG-Systemen dem Verständnis der meisten Sachverständigen entziehen dürfte.

In einschlägigen Debatten wird zwar zuweilen gefordert, der Quellcode von PG-Systemen müsse den Strafbehörden und Parteien offengelegt und zugänglich gemacht werden, wogegen sich kommerzielle Anbieter solcher Systeme – wenig überraschend – vehement wehren.⁵³ Diese Debatte greift jedoch zu kurz, denn auch wenn der Quellcode zur Verfügung stünde, bedeutet das noch lange nicht, dass er

50 Probabilistische Genotypisierung bezeichnet die kombinierte Anwendung von biologischen Modellen, Statistik und Berechnungsmethoden, um die Bewertung komplexer DNA-Profilierungsergebnisse zu unterstützen. Siehe Scientific Working Group on DNA Analysis Methods (SWGDM), Guidelines for the Validation of Probabilistic Genotyping Systems, 2015, abrufbar unter: <https://www.swgdam.org/publications> (besucht am: 25. 1. 2023).

51 Siehe z. B. United States vs. Gissantaner, 2019 U.S. Dist. LEXIS 178848, 2019 WL 5692183 (W. D. Mich. Oct. 16, 2019); C. Chessman, A «source» of error: computer code, criminal defendants, and the constitution, California Law Review 105/2017, 101–193; B. Stiffelman, No Longer the Gold Standard: Probabilistic Genotyping is Changing the Nature of DNA Evidence in Criminal Trials, Berkeley Journal of Criminal Law 24/2019, 110; J. Buckleton/B. Robertson/J. Curran/C. Berger/D. Taylor/J. A. Bright/T. Hicks/S. Gittelson/I. Evett/S. Pugh/G. Jackson/H. Kelly/T. Kalafut/F. R. Bieber, A Review of Likelihood Ratios in Forensic Science Based on a Critique of Stiffelman «No longer the Gold standard: Probabilistic genotyping is changing the nature of DNA evidence in criminal trials», Forensic Science International 310/2020, 110251.

52 A. Roth, Machine Testimony, Yale Law Journal, 126/2017, 1972. Für eine rechtsvergleichende Perspektive siehe S. Gless, AI in the Courtroom: A Comparative Analysis of Machine Evidence in Criminal Trials, Georgetown Journal of International Law 51 (2020), 195.

53 Für eine detaillierte Diskussion dieser Problematik siehe z. B. E. J. Imwinkelried, Computer Source Code: A Source of the Growing Controversy Over the Reliability of Automated Forensic Techniques, De Paul Law Review, 66(1)/2017, 97. Anzumerken ist, dass gewisse Anbieter Einsicht in den Quellcode im Rahmen einer NDA-Vereinbarung («non-disclosure agreement») zulassen (P. Gill/C. Benschop/J. Buckleton/Ø. Bleka/D. Taylor, A Review of Probabilistic Genotyping Systems: EuroForMix, DNASTatistX and STRmix™, Genes, 12/2021, 1559, 32).

sich den Parteien einfach erschliesst. Ausser den Entwicklern des jeweiligen PG-Systems, in der Regel ein Nischenprodukt, dürften sich weltweit nur schwerlich Sachkundige finden lassen, welche in der Lage und gewillt wären, den Quellcode kompetent zu durchleuchten. Nicht zuletzt auch wegen der Zeit, die eine solche Prüfung in Anspruch nehmen würde, entzieht sich den Parteien daher die Möglichkeit, Rechenergebnisse von PG-Systemen grundlegend zu hinterfragen und somit zu konfrontieren.

Angesichts dieser Herausforderungen konzentriert sich die wissenschaftliche Forensik derzeit darauf, PG-Systeme zumindest empirisch zu testen und untereinander zu vergleichen. Entsprechende Studien liefern zwar wertvolle Hinweise zur *allgemeinen* Leistungsfähigkeit von bestimmten PG-Systemen unter kontrollierten Bedingungen, lösen jedoch das Kernproblem des «maschinellen Beweises» nicht einmal im Ansatz. Denn Sachverständige, die in Gutachten Ergebnisse von PG-Systemen als Beweiswerte anzuführen gedenken, müssen – ungeachtet der allgemeinen Studienlage – Validierungsdaten über die Verwendung eines bestimmten Systems auf der Stufe des einzelnen Labors vorlegen können.⁵⁴ Wie ein unlängst vom *National Institute for Standards and Technology* der Vereinigten Staaten publizierter, umfangreicher Berichtsentwurf jedoch festhält, existieren derzeit schlicht «nicht genügend öffentlich zugängliche Daten, um die Zuverlässigkeit der Praktiken der DNA-Mischspureninterpretation, einschliesslich der Verwendung von probabilistischen Genotypisierungssoftware-Systemen (PGS), extern und unabhängig zu beurteilen».⁵⁵

Obschon wir hier bei Weitem keine erschöpfende Abhandlung konzeptueller und praktischer Schwierigkeiten bieten, sehen wir, dass die Idee einer «maschinengestützten» Interpretation derzeit mindestens noch ebenso viele Probleme generiert, wie sie vorgibt zu lösen. Sachverständige, welche über kein validiertes computergestütztes Befundbewertungsverfahren verfügen, müssen komplexe DNA-Profile daher konsequenterweise als nicht interpretierbar bezeichnen und von der weiteren Betrachtung ausschliessen.

54 Hahn et al., (Fn. 43), 9; *European Network of Forensic Science Institutes*, Best Practice Manual for the Internal Validation of Probabilistic Software to Undertake DNA Mixture Interpretation 2017, abrufbar unter: <https://enfsi.eu/wp-content/uploads/2017/09/Best-Practice-Manual-for-the-internal-validation-of-probabilistic-software-to-undertake-DNA-mixture-interpretation-v1.docx.pdf> (besucht am: 25. 1. 2023).

55 J. Butler/H. Iyer/R. Press/M. K. Taylor/P. M. Vallone/S. Willis, DNA Mixture Interpretation: A NIST Scientific Foundation Review, NISTIR 8351-DRAFT, National Institute of Standards and Technology, U. S. Department of Commerce, 75, abrufbar unter: <https://doi.org/10.6028/NIST.IR.8351-draft> (besucht am: 25. 1. 2023).

VII. Schlussfolgerungen

Bei der Frage des Beweiswerts unterscheiden sich Ergebnisse von forensischen DNA-Profilanalysen grundsätzlich kaum von anderen forensisch-wissenschaftlichen Untersuchungsergebnissen. So ist weitgehend unbestritten, dass ein rationales Vorgehen bei der Befundbewertung eine Auseinandersetzung mit der Fragen-Trias⁵⁶ erfordert. Wie wir im vorliegenden Beitrag aufgezeigt haben, ergibt sich bei der konkreten Anwendung der grundlegenden Prinzipien der Befundbewertung jedoch eine Reihe von Schwierigkeiten, sodass Ergebnisse von DNA-Profilanalysen weder als ein unfehlbares noch als ein unanfechtbares «Allheilmittel» für die Tatsachenfeststellung vor Gericht anzusehen sind.

In Anbetracht der zahlreichen technischen Aspekte und der vielschichtigen konzeptuellen Analyse, die jede fallbezogene Interpretation von DNA-Untersuchungsergebnissen erfordert, ist es verständlich, aber leider auch bedauerlich, dass Prozessbeteiligte davon absehen, Sachverständigengutachten einer näheren Überprüfung zu unterziehen. In Ermangelung ausreichender eigener Fachkenntnisse sowie schlüsselfertiger Instrumente, um Sachverständigenberichte kompetent zu durchleuchten, scheint es angezeigt, unabhängige (private) Experten zu konsultieren. Diese Option kann sich jedoch aus Gründen der Verfügbarkeit von solchen Experten sowie fehlender Zeit und Ressourcen als wenig praktikabel herausstellen. Für die Strafverteidigung ergeben sich daher nach wie vor unverhältnismässig grosse Anstrengungen, um eine kritische Überprüfung forensisch-wissenschaftlicher Untersuchungsergebnisse anzustossen. Und selbst wenn dieses Unterfangen gelingen sollte, wird es von hiesigen Gerichten erfahrungsgemäss noch häufig skeptisch, wenn nicht tendenziell abschlägig behandelt.

Engagierte Kreise der Anwalt- und Richterschaft sollten sich jedoch nicht davon abhalten lassen, Sachverständigengutachten zumindest einer rudimentären Prüfung zu unterziehen. Die ENFSI-Richtlinie bietet hierfür praktische Hilfestellung in der Form einer Audit-Vorlage.⁵⁷ Ihr Zweck besteht darin, Empfängern von befundbewertenden Berichten (Sachverständigengutachten) ein Evaluierungsrastrer an die Hand zu geben, mit welchem sie rasch eine erste Beurteilung vornehmen können, ob der vorliegende Bericht den minimalen Anforderungen der ENFSI-Richtlinien entspricht. Unserer Erfahrung nach weisen viele Sachverständigengutachten in den verschiedensten Bereichen der forensischen Wissenschaft immer noch zahlreiche Unzulänglichkeiten auf, selbst was die grundlegenden Prinzipien der Befundbewertung betrifft, sodass es sich lohnt, sich kritisch mit dieser Materie auseinanderzusetzen.

56 Siehe 1. Teil, Abschnitt III.

57 Willis et al. (Fn. 3), 122 f.

In welche Richtung sollte sich die Verwendung von Ergebnissen von DNA-Profilanalysen im Rahmen der Tatsachenfeststellung vor Gericht in Zukunft entwickeln? Nach dem, was wir im vorliegenden Beitrag aufgezeigt haben, plädieren wir vor allem für zwei Punkte: Erstens sind die bei der befundbewertenden Berichterstattung zu berücksichtigenden Hypothesen kritisch zu hinterfragen, und sie sollten sich den Bedürfnissen der Empfänger von Sachverständigengutachten anpassen. Es kann nicht sein, dass Sachverständige der Befundbewertung standardmässig Hypothesen der Spurenherkunft unterlegen, nur weil sich damit die Befundbewertung am leichtesten bewerkstelligen lässt. Zweitens ist überfällig, dass Sachverständige von offensichtlich realitätsfremden Annahmen, beispielsweise die Annahme eines stets fehlerfreien Untersuchungsprozesses und damit unrealistischer Beweiswerte, Abstand nehmen. Damit liesse sich vermeiden, dass vermeintlich «astronomische» Zahlenwerte Eingang in Verfahren nehmen, ohne dass das Gericht, die Prozessparteien oder die Öffentlichkeit in der Lage wären, die in Gutachten genannten Zahlen und Ergebnisse rational einzuordnen, geschweige denn deren Unzulänglichkeiten zu verstehen. Dem Einwand, es sei ein Verlust, von Werten im Milliarden- und Trilliardenbereich abzusehen, können wir nur widersprechen, denn – wie de Finetti treffend bemerkte – «Nichts ist verloren, was nur eine Illusion war.»⁵⁸

58 B. de Finetti, Bayesianism: Its Unifying Role for Both the Foundations and Applications of Statistics, *International Statistical Review* 42 (1974), 117, 121 (frei übersetzt).