

Rechtsmedizin 2023 · 33:13–29
<https://doi.org/10.1007/s00194-022-00600-1>
 Angenommen: 23. November 2022
 Online publiziert: 11. Januar 2023
 © Der/die Autor(en) 2023



Einsatz vollkontinuierlicher Modelle zur biostatistischen Bewertung forensischer DNA-analytischer Befunde

Erfahrungen der Projektgruppe „Biostatistische DNA-Berechnungen“

Michael Templin¹ · Peter Zimmermann² · Sebastian Kranz³ · Martin Eckert⁴ · Christoph Leuker⁵ · Sven Razbin⁶ · Andrea Wächter⁷ · Volker Weirich^{8,14} · Katja Anslinger⁹ · Rolf Fimmers¹⁰ · Stefanie Grethe¹¹ · Claus Oppelt¹ · Marielle Vennemann¹² · Peter M. Schneider¹³ · Meinhard Hahn¹

¹Landeskriminalamt Niedersachsen, Hannover, Deutschland; ²Landeskriminalamt Baden-Württemberg, Stuttgart, Deutschland; ³Landeskriminalamt Hamburg, Hamburg, Deutschland; ⁴Bundeskriminalamt, Wiesbaden, Deutschland; ⁵Landeskriminalamt Nordrhein-Westfalen, Düsseldorf, Deutschland; ⁶Landeskriminalamt Bremen, Bremen, Deutschland; ⁷Landeskriminalamt Sachsen-Anhalt, Magdeburg, Deutschland; ⁸Institut für Klinische Chemie, Universitätsklinikum Schleswig-Holstein, Kiel, Deutschland; ⁹Institut für Rechtsmedizin, Ludwig-Maximilians-Universität, München, Deutschland; ¹⁰Institut für Forensische Statistik und Qualitätssicherung, St. Augustin, Deutschland; ¹¹Landeskriminalamt Rheinland-Pfalz, Mainz, Deutschland; ¹²Institut für Rechtsmedizin, Westfälische Wilhelms-Universität, Münster, Deutschland; ¹³Institut für Rechtsmedizin, Medizinische Fakultät und Universitätsklinikum, Köln, Deutschland; ¹⁴vormals: Landeskriminalamt Mecklenburg-Vorpommern, Ramepe, Deutschland

Mitglieder der Bund-Länder-Projektgruppe „Biostatistische DNA-Berechnungen“ der Kommission KKWT/ED Kriminalwissenschaft und -technik/Erkennungsdienst: Martin Eckert, Meinhard Hahn, Sebastian Kranz, Christoph Leuker, Sven Razbin, Michael Templin, Andrea Wächter, Volker Weirich, Peter Zimmermann

Mitglieder der Gemeinsamen Spurenkommission der rechtsmedizinischen und kriminaltechnischen Institute: Katja Anslinger, Martin Eckert, Rolf Fimmers, Stefanie Grethe, Carsten Hohoff, Claus Oppelt, Peter M. Schneider, Marielle Vennemann



QR-Code scannen & Beitrag online lesen

Hinführung

Die Bewertung von DNA-Spuren mit biostatistischen Methoden hat in den letzten Jahren enorme Fortschritte erzielt. Unter Verwendung sog. vollkontinuierlicher Modelle (VKM) können forensische DNA-Befunde unter Einbeziehung möglicher Drop-out-Ereignisse bewertet werden. Im deutschsprachigen Raum liegen dazu bisher kaum Erfahrungen vor. Diese Arbeit untersucht, ob sich die aktuell verfügbaren VKM für einen Einsatz bei der Analyse von DNA-Spuren eignen. Sie präsentiert die Ergebnisse verschiedener Testreihen, erläutert die generelle Funktionsweise von VKM und bietet einen Einstieg in die Thematik.

Einleitung

Die zunehmende Anzahl zu analysierender STR-Merkmalssysteme und die gesteigerte Sensitivität heutiger Analysemethoden haben zu einem erheblichen Anstieg unvollständig auswertbarer DNA-Profile bio-

logischer Spuren geführt. Solche Profile weisen zwar häufig genügend Informationen für eine gutachterliche Beurteilung auf, können jedoch mit den derzeit gängigen binären Berechnungsmethoden biostatistisch nur unzureichend bewertet werden.

Mit Stand der Empfehlungen von 2016 [16] wurde es als hinnehmbar erachtet, bestimmte als „Typ C“ klassifizierte Mischspuren [14] unter Ausklammern einzelner, nicht sicher bestimmbarer Merkmalssysteme mittels binärer Modelle zu berechnen. Es ist aber eine Tatsache, dass sich Spuren mit möglichen Drop-in- bzw. Drop-out-Ereignissen mithilfe von binären Modellen nicht adäquat abbilden lassen, da nicht alle zur Verfügung stehenden Informationen der zugrunde liegenden Elektrophorerogramme einbezogen werden.

Seit der Veröffentlichung der gemeinsamen Empfehlungen der Projektgruppe „Biostatistische DNA-Berechnungen“ der Landeskriminalämter und des Bundeskriminalamtes sowie der Spurenkommission der Deutschen Gesellschaft für Rechtsme-

dizin aus 2016 [16] wurden als Alternative zu den bestehenden binären Modellen die sog. probabilistischen Verfahren stetig weiterentwickelt. Während semikontinuierliche Modelle [11] bereits Drop-in-/Drop-out-Ereignisse berücksichtigen, beziehen die vollkontinuierlichen Modelle (VKM) darüber hinaus Signalintensitäten und DNA-Fragmentlängen in die biostatistische Berechnung ein [1, 5, 6, 9].

Von verschiedenen Anbietern wurden auf VKM basierende Programme entwickelt (■ Tab. 1 sowie [2, 3, 12]), die eine biostatistische Bewertung von DNA-analytischen Befunden ermöglichen [1, 5, 6, 9, 15]. Zur Erklärung der gemessenen Signalintensitäten berücksichtigen diese Modelle neben den Allelwerten auch DNA-Fragmentlängen, Stutterpeaks (Vor- und Nachpeaks) sowie Degradationseffekte. Für eine erhebliche Anzahl an DNA-Spuren, die sich bisher lediglich verbal bewerten ließen, ist mithilfe der VKM eine angemessene biostatistische Bewertung möglich [4].

Durch die Verwendung von mathematischen Algorithmen können Aspekte wie Haupt- und Nebenkomponenten, Hetero-/Homozygotie, Signale nahe der Nachweisgrenze etc. standardisiert bewertet werden. Dadurch wird die Bewertung von DNA-Spuren objektiver.

Neben einer Berechnung vollkontinuierlicher Likelihood-Ratio-Werte (LR_{fc}) können mit VKM ggf. auch Mischspuren in ihren Einzelkomponenten dargestellt werden, d. h., die wahrscheinlichen DNA-Profile der zum Mischprofil beitragenden unbekanntenen Personen werden abgeleitet (Deconvolution). Solche DNA-Profile lassen sich bei ausreichender Qualität für Datenbankrecherchen nutzen, woraus sich zusätzliche Ermittlungsansätze ergeben können.

Während Veröffentlichungen in internationalen Fachjournals eindrucksvoll die Eignung und den Einsatz von VKM in der forensischen DNA-Analytik demonstrieren [6, 9], liegen im deutschsprachigen Raum hierzu kaum Erfahrungen vor.

Es war daher das erklärte Ziel der Projektgruppe „Biostatistische DNA-Berechnungen“, solche Modelle hinsichtlich ihrer Funktionsweise und Funktionsfähigkeit zu untersuchen.

Die vorliegende Publikation soll es den Sachverständigen ermöglichen, sich den

Die biostatistische Bewertung DNA-analytischer Befunde unterstützt Gerichte bei der Einschätzung des Beweiswertes einer Spur. In der Praxis werden dabei zunehmend Spuren mit minimaler DNA-Menge und möglichen „Drop-in“- und „Drop-out“-Ereignissen sowie komplexe Mischspuren analysiert. Solche Spuren sind mit einer klassischen „binären“ Berechnung biostatistisch häufig nicht oder nur eingeschränkt bewertbar.

Die Entwicklung vollkontinuierlicher Modelle (VKM) macht eine Vielzahl dieser bisher nicht berechenbaren Spuren einer biostatistischen Bewertung zugänglich. Dabei werden nahezu sämtliche verfügbaren Informationen einer DNA-Spur in die Berechnung einbezogen. Während diese probabilistischen Verfahren international bereits vielfach zum Einsatz kommen, liegen hierzu im deutschsprachigen Raum nur wenige Erfahrungen vor.

Um Funktionsweise, Möglichkeiten und Grenzen von VKM-Berechnungen zu erfassen, wurden Mischspuren bekannter Zusammensetzung mit 4 aktuell verfügbaren VKM-Programmen vergleichend analysiert. Bei der Auswertung wurden zentrale Aspekte betrachtet, wie beispielsweise die Konkordanz von Berechnungsergebnissen, der Einfluss von Drop-in- und Drop-out-Ereignissen auf die berechneten vollkontinuierlichen LR-Werte (LR_{fc}) sowie die Ableitung recherchefähiger DNA-Profile mithilfe wahrscheinlichkeitsbasierter Prognosen (Deconvolution). Die im Rahmen dieser Arbeit gewonnenen Erfahrungen bilden, zusammen mit weiteren bereits international publizierten Studien, eine Basis für Empfehlungen zum Einsatz von VKM-basierter Software bei der biostatistischen Bewertung DNA-analytischer Befunde.

Schlüsselwörter

Probabilistische Genotypisierung · Deconvolution · Likelihood Ratio · DNA-Sachbeweis · Forensische Biostatistik

komplex wirkenden Modellen zu nähern und die Funktionsweise solcher Modelle/Programme anhand von praktischen Erfahrungen aus dem eigenen Arbeitsalltag nachzuvollziehen. Keinesfalls erhebt die vorliegende Publikation den Anspruch, eine Beschäftigung mit einschlägiger Fachliteratur zu ersetzen.

Flankierend zu dieser Publikation veröffentlicht die Projektgruppe „Biostatistische DNA-Berechnungen“ gemeinsam mit der Spurenkommission der Deutschen Gesellschaft für Rechtsmedizin ihre ersten praxisorientierten Empfehlungen zum Einsatz von vollkontinuierlichen Modellen [8]. Diese Empfehlungen sollen die Untersuchungsstellen dabei unterstützen, VKM erfolgreich als Werkzeug im Rahmen von ermittlungsbegleitenden Untersuchungen und für die Berichterstattung vor Gericht einzusetzen.

Binäre, semikontinuierliche und vollkontinuierliche Modelle im Vergleich

Binäre, semikontinuierliche und vollkontinuierliche Modelle verarbeiten Drop-in-

und Drop-out-Ereignisse auf verschiedene Weise. Zur Veranschaulichung dieser Modelle kann es vorteilhaft sein, eine „definierte“ Mischspur, bestehend aus den Merkmalen von 2 bekannten Spurenursachern (2P-Mischspur), zu betrachten.

Eine solche Spur zeigt ■ Abb. 1, wobei das gemessene Elektropherogramm in der ■ Abb. 1a und die 3 zu vergleichenden Modelle in den ■ Abb. 1b–d dargestellt sind. Die bekannten Allele einer Person 1 (*hellgrau*) und einer Person 2 (*dunkelgrau*) sind der besseren Übersichtlichkeit halber jeweils unter den ■ Abb. 1b–d angegeben.

In ■ Abb. 1a sind die gemessenen Signale und zugehörigen Signalintensitäten für die 4 exemplarisch gezeigten Merkmalssysteme schematisch als Peaks dargestellt. Erkennbar sind mehrere „sicher“ bestimmbare Allele, die deutlich oberhalb der Nachweisgrenze liegen. Zudem sind aber weitere Signale unterhalb bzw. in Höhe der Nachweisgrenze sichtbar, bei denen nicht eindeutig ist, ob es sich hierbei um reale Allele, Stutterpeaks oder Drop-ins handelt.

Tab. 1 Übersicht über die getesteten VKM-Programme (Sachstand: Frühjahr 2020)				
Programm	DNA•VIEW Mixture Solution	EuroForMix	GenoProof Mixture	STRmix
Version	dy17-19-8-15	2.1.0	3.1.2	2.6
Hersteller	Charles H. Brenner, DNA•VIEW Software, Oakland, CA, USA	Øyvind Bleka, Peter Gill	qualitype GmbH, Dresden	Institute of Environmental Science & Research Limited (ESR), New Zealand
Internetpräsenz	dna-view.com	www.euroformix.com	www.qualitype.de	www.strmix.com
Open Source/lizenzfrei	Nein	Ja	Nein	Nein
Implementierung mit internen Labordaten	Ja	Nein	Nein	Ja
GeneMapper-Import	Nein	Nein	Ja	Ja
Markov-Chain Monte-Carlo(MCMC)-Verfahren	Nein	Nein	Ja	Ja
Berechnung Replikate gleicher Kits	Nein	Ja	Ja	Ja
Berechnung Replikate unterschiedlicher Kits	Nein	Nein	Ja	Ja
Abschätzung der Anzahl anzunehmender Spurenverursacher	Ja	Nein	Ja	Nein
Import neuer STR-Marker und Analysekits	Ja	Ja	Ja	Ja

Binäre Modelle

Zur Berechnung DNA-analytischer Befunde werden derzeit überwiegend binäre Modelle eingesetzt. **Abb. 1b** zeigt die Verarbeitung eines Spurenbefundes mithilfe eines binären Modellansatzes. In klassischer Weise werden bei binären Modellen Merkmale oberhalb der Nachweisgrenze in die Berechnung einbezogen, wohingegen Merkmale unterhalb der Nachweisgrenze von der Berechnung ausgeschlossen werden. Allerdings werden vor der Bewertung eines Spurenbefundes (Elektropherogramm) oftmals zusätzliche Mechanismen (z.B. Filter für Stutterpeaks) eingesetzt, welche einen Einfluss auf die in die Berechnung eingehenden Befunde ausüben. Solche Mechanismen sind kein Bestandteil eines klassischen binären Berechnungsmodells.

Der Drop-in eines Merkmals im 2. Merkmalssystem (*Raute*) wird daher in die binäre Berechnung einbezogen, während der Drop-out eines weiteren Merkmals im 4. Merkmalssystem (*Dreieck*) aus der Berechnung herausfällt.

Bei Anwendung der binären Methode unter Berücksichtigung aller Merkmalssysteme ist eine Berechnung der Spur nicht sinnvoll möglich, da das Fehlen von Merkmalen (Drop-out) nicht abgebildet werden kann. Eine binäre Berechnung des LR-Wertes wäre demnach nur unter Ausklammern von Befunden auf Basis der beobachte-

ten Spurendaten ohne Berücksichtigung des Genotyps der POI möglich (vgl. Abschnitt 2.3 „Grenzfälle“ [16]).

Für eine Berechnung mit binären Modellen werden ausschließlich die nachgewiesenen Allele einer DNA-Spur sowie deren Populationsfrequenzen verwendet. Ebenfalls bei der Messung anfallende Informationen zur Signalintensität und zur DNA-Fragmentlänge von Allelen bleiben dagegen gänzlich unberücksichtigt. Aus diesem Grund sind in binären Modellen keine Aussagen zu den Mischungsverhältnissen und den Merkmalsmustern beteiligter Spurenverursacher möglich. Jegliche Rückschlüsse auf die Merkmalsmuster von beteiligten Spurenverursachern beruhen somit ausschließlich auf der Einschätzung des auswertenden Sachverständigen.

Im Allgemeinen sind binäre Modelle sehr gut für die Berechnung von eindeutigen bzw. hochqualitativen DNA-Spuren geeignet. Aufgrund der wenigen in das Modell einfließenden Informationen sind diese Modelle sehr schnell und liefern bei guter Spurenlage entsprechende LR-Werte. Angesichts des reduzierten Informationsgehalts ermöglichen sie jedoch keine angemessene Bewertung von DNA-Spuren mit möglichen Drop-in- bzw. Drop-out-Ereignissen.

Semikontinuierliche Modelle

Semikontinuierliche Modelle (**Abb. 1c**) weisen gegenüber binären Modellen einen Vorteil auf, indem sie mögliche Drop-in- und Drop-out-Ereignisse grundsätzlich in die Berechnung einbeziehen. Dabei stellen sie aber keinen konkreten Bezug zu den gemessenen Signalintensitäten her, sondern verwenden zugewiesene, errechnete oder experimentell bestimmte Drop-in- und Drop-out-Wahrscheinlichkeiten. Entsprechend werden Allelen oberhalb der Nachweisgrenze Drop-in-Wahrscheinlichkeiten und Allelen unterhalb der Nachweisgrenze Drop-out-Wahrscheinlichkeiten zugewiesen.

Semikontinuierliche Modelle weisen gegenüber VKM den Vorteil auf, dass LR-Werte mit relativ geringem Rechenaufwand schnell zu berechnen sind.

Vollkontinuierliche Modelle

Vom methodischen Ansatz her beziehen vollkontinuierliche Modelle neben den Allelen und deren Frequenzen zusätzlich die gemessenen Signalintensitäten und DNA-Fragmentlängen sowie Stutterpeaks in die Berechnungen von DNA-analytischen Befunden ein. Dabei werden die Informationen eines Elektropherogramms umfassend berücksichtigt (**Abb. 1d**), was prinzipiell im Vergleich zu semikontinuierlichen und binären Modellen eine detailgetreue-

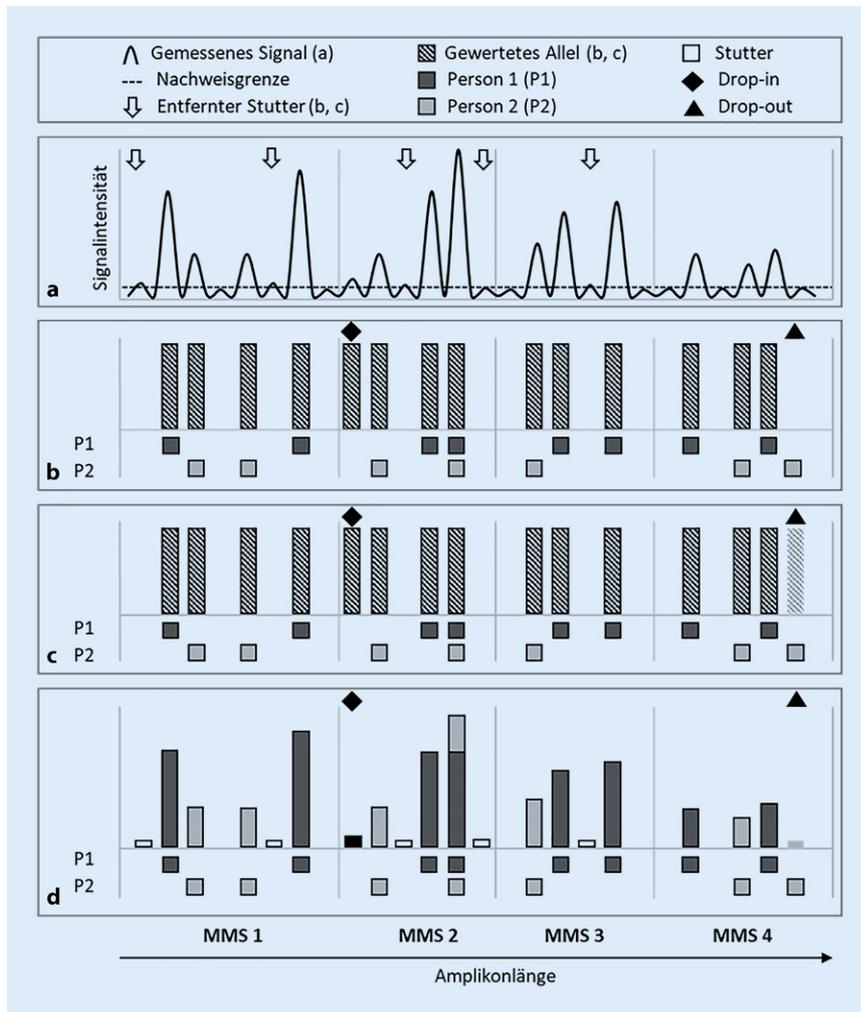


Abb. 1 ▲ Binäre, semi- und vollkontinuierliche Modelle im schematischen Vergleich. **a** Schematische Darstellung eines Elektropherogramms für eine „definierte“ Mischspur, bestehend aus den Merkmalen von 2 bekannten Spurenverursachern (Personen P1 und P2). Stellvertretend für das Gesamtprofil werden hier 4 Merkmalssysteme (MMS) dargestellt. **b** Binäre Bewertung nach Anwendung von Stutter-Filtern. **c** Semikontinuierliche Bewertung nach Anwendung von Stutter-Filtern. **d** Vollkontinuierliche Bewertung unter Berücksichtigung der Signalintensitäten

re Modellierung von DNA-Spuren ermöglicht.

Aufgrund der Einbeziehung all dieser Informationen erlauben vollkontinuierliche Modelle eine sachgerechtere Bewertung einer Spur hinsichtlich der DNA-Menge und -Qualität. Zudem ermöglichen es VKM, das Mischungsverhältnis der Komponenten einer DNA-Spur rechnerisch abzuschätzen.

Aus diesen Größen lässt sich letztlich eine wahrscheinlichkeitsbasierte Prognose über die zur Erklärung eines DNA-Befundes geeigneten Genotypen erstellen. Dieser methodische Ansatz wird als Probabilistic Genotyping oder Deconvolution bezeichnet (siehe Abschnitt Ableitung recher-

fähiger Datenbankmuster aus Mischspuren mithilfe vollkontinuierlicher Modelle [Deconvolution]).

Im oben gezeigten Beispiel einer 2P-Mischspur führt die „Deconvolution“ zur Ableitung zweier Profile, die den beiden (unbekannten) Personen (UP1 und UP2) zugeordnet werden können (■ Abb. 9).

Unter Umständen kann die Berechnung für einen Genotypen eine derart hohe Evidenz ergeben, dass dieser in der DNA-Analyse-Datei recherchiert bzw. erfasst werden kann.

Funktionsweise vollkontinuierlicher Modelle

VKM beruhen auf dem Berechnungsansatz der Likelihood Ratio und benötigen ebenso wie semikontinuierliche und binäre Modelle eine Vorgabe zur Anzahl der Spurenverursacher des zu interpretierenden Spurenprofils.

Es existieren unterschiedliche VKM, die jeweils über einen eigenen spezifischen Berechnungsalgorithmus verfügen. Jedes Programm modelliert Parameter wie z. B. DNA-Menge, Degradation und Amplifikationseffizienz und berücksichtigt auch Stutterpeaks sowie die Möglichkeit evtl. Drop-in- und Drop-out-Ereignisse. Nicht jedes Modell verwendet zwangsläufig die gleichen Parameter.

Die Algorithmen der VKM sind in der Regel in Fachjournalen publiziert. Mit höherer Anzahl zu modellierender Parameter steigen Intensität und Dauer einer Berechnung zumeist erheblich an. Zur Modellierung dieser Parameter stellt das Modell bestimmte Annahmen und Erwartungen zum Verhalten der Signale (Peaks) zueinander auf.

Ein einfaches Beispiel zur Verdeutlichung: Gegeben sei im Merkmalssystem vWA ein Spurenbefund mit den Allelen 17 (Peakhöhe von 3500 rfu) und 18 (Peakhöhe von 3450 rfu).

Unter Annahme einer Person als Spurenverursacher wird jedes VKM ebenso wie jeder Sachverständige hier mit sehr hoher Wahrscheinlichkeit die Genotypkonstellation 17/18 eines Spurenverursachers ableiten. Andere Konstellationen sind zwar grundsätzlich möglich, werden aber als sehr viel unwahrscheinlicher eingeschätzt bzw. berechnet.

Hätte nun das Allel 17 eine Peakhöhe von lediglich 350 rfu, wäre die Konstellation 17/18 eines Spurenverursachers nicht mehr die wahrscheinlichste Erklärung. In diesem Fall wäre das Allel 17 eher als Stutterpeak zu erklären. Sowohl VKM als auch der Sachverständige würden somit die Genotypkonstellation 18/18 des Spurenverursachers als die wahrscheinlichste Erklärung ableiten.

Die Gewichtung der für die Entstehung der Spur wahrscheinlichsten Genotypen erfolgt bei VKM unter Einbeziehung sämtlicher verfügbarer Informa-

tionen (Allelwerte, Signalintensitäten, Amplikonlängen) und unter Annahme von Grundvoraussetzungen wie z. B. der Spurenlegeranzahl.

Das vollkontinuierliche Modell erzeugt bei der Berechnung fortwährend neue Kombinationen von Genotypen und Parametern und berechnet dann in einem 2. Schritt, mit welcher Wahrscheinlichkeit eine Kombination bestimmter Parameter ein zu interpretierendes Spurenprofil (Elektropherogramm) erklären kann. Der Parameterraum wird im Rahmen eines Optimierungsprozesses durchlaufen, bis die bestmögliche Übereinstimmung zwischen modelliertem Profil und gemessenem Elektropherogramm erreicht ist.

Für eine Deconvolution können während des Optimierungsprozesses die einzelnen Kombinationen aus Genotypen und weiteren Parametern gemäß ihrer Wahrscheinlichkeit, das vorhandene Spurenprofil zu erklären, gewichtet werden. Dazu werden die Einzelwahrscheinlichkeiten aller getesteten Genotypkombinationen addiert und die Gesamtwahrscheinlichkeit aller möglichen Genotypen auf 1 normiert. Hieraus lässt sich dann die Gewichtung für jede einzelne getestete Genotypkombination errechnen.

Im Idealfall wird einem Genotyp einer Spurenkomponente (Person) eine Gewichtung von nahezu 1 zugewiesen. In solchen Fällen hat das VKM lediglich eine einzige, mathematisch realistische Lösungsmöglichkeit finden können, die die vorhandene Spurenkomponente erklären kann. Die Gewichtungen aller anderen Genotypen müssen dann gegen Null tendieren.

Bei Spurenprofilen ohne bekannte Spurenverursacher lassen sich also u. U. geeignete Einzelprofile für eine Recherche in einer Datenbank ableiten [2, 13]. Diese Eindeutigkeit ist bei vielen Mischspuren nicht gegeben, sodass ein VKM ebenso wie ein Sachverständiger nicht in der Lage sein wird, hier eine Prognose über die Genotypkonstellationen der einzelnen Spurenverursacher abzugeben.

Material und Methoden – Vorgehen der Projektgruppe

Software

Vor Beginn der Testphase von VKM wurden zunächst bekannte Entwickler und Firmen mittels eines umfangreichen Fragebogens gebeten, ihre Programme zu beschreiben. Unter den Rückmeldungen wurden diejenigen VKM-Programme ausgewählt, für die seitens der Entwickler/Firmen kostenfreie Testversionen zur Verfügung gestellt wurden. In diesem Rahmen war es möglich, die Programme „DNA•VIEW Mixture Solution“, „EuroForMix“ [2], „GenoProof Mixture“ [12] und „STRmix“ [3] hinsichtlich ihrer Funktionsfähigkeit und Funktionsweise zu testen (■ Tab. 1). Bei den in der Tabelle angegebenen Funktionalitäten ist zu beachten, dass die einzelnen Programme seit Abschluss der Testphase im Frühjahr 2020 weiterentwickelt wurden und somit davon auszugehen ist, dass Veränderungen erfolgt sind.

Anfertigung von DNA-Mischspuren

Zur Anfertigung von DNA-Mischspuren wurde freiwilligen Probanden mit deren schriftlichen Einverständnis Blut abgenommen. Für die Berechnungen wurden standardisierte Blutmischungen von 2 bis 4 Personen mit je 1 bis 5 Mischungsanteilen hergestellt. Hierzu wurde für 5 gespendete EDTA-Blut-Proben (bezeichnet als A bis E) jeweils die Leukozytenzahlen bestimmt, anhand derer eine Normalisierung der Konzentration kernhaltiger Zellen möglich war. Im Anschluss standen einzelne Blutverdünnungen zur Verfügung, die nach einem vorher festgelegten Schema gemischt wurden, um die gewünschten Kombinationen und Mischungsverhältnisse zu erreichen. Die hierbei eingesetzten Vollblutmengen richteten sich nach der Leukozytenzahl, dem theoretischen DNA-Gehalt und der prognostizierten Extraktionseffizienz. Anschließend erfolgten die DNA-Extraktion und Quantifizierung. Die für die nachfolgenden Amplifikationen eingesetzten DNA-Mengen wurden so gewählt, dass für jeden einzelnen Spurenanteil ca. 12 pg DNA im PCR-Ansatz zu 25 µl vorlagen.

Durch Kombination und Variation der 5 Blutproben wurden insgesamt 19 DNA-Mischspuren erzeugt, deren Zusammensetzungen und DNA-Anteile ein breites Spektrum in der täglichen Fallbearbeitung abdeckten. Hierbei wurden bewusst geringe DNA-Mengen eingesetzt, die das Auftreten von Befunden im Bereich der Nachweisgrenze (mit Drop-out-Ereignissen) wahrscheinlich machten. Die DNA-Mischspuren wurden keinen künstlichen Degradationsmaßnahmen unterzogen.

Die Typisierungen erfolgten mit den Analysekits NGM Detect™ PCR Amplification Kit (Fa. ThermoFisher Scientific, Waltham, MA, USA) und Powerplex® ESI 17 Fast System (Fa. Promega, Madison, WI, USA) auf einem 3500xL Genetic Analyzer (Fa. Applied Biosystems, Waltham, MA, USA). Aufgrund der z. T. geringen Mischungsanteile der analysierten Spuren zeigten die Befunde erwartungsgemäß z. T. starke Peakimbilanzen, Drop-in- und Drop-out-Ereignisse sowie mischungs- bzw. proportionsbedingte Signalunterschiede von teils erheblichen Ausmaßen. Die Amplifikationen der nachfolgend gezeigten Mischspuren wurden exemplarisch mit dem NGM Detect™ PCR Amplification Kit durchgeführt.

Berechnungen mit den verfügbaren VKM-Programmen

Um Fehlerquellen zu vermeiden und eine möglichst einheitliche Datengrundlage zu bilden, wurde ein makrobasiertes Konvertierungstool für den Import von Daten in die VKM-Programme entwickelt. Somit standen für alle Programme identische Ausgangsdaten zur Verfügung. Die bei der Installation der jeweiligen Programme empfohlenen oder bereits als Standard mitgelieferten Einstellungen von internen Parametern wurden nicht verändert.

Anhand der 19 angefertigten Blutmischungen wurden insgesamt 60 Hypothesenpaare mit entsprechenden Fragestellungen nach jeder einzelnen der 5 Personen A bis E als „person of interest“ (POI) aufgestellt. Für alle hier gezeigten Berechnungen wurde die zur Erklärung der Spuren minimal notwendige Zahl an Spurenverursachern eingestellt. Die weiteren zur Erklärung einer DNA-Spur benötigten Spurenverursacher wurden jeweils durchnum-

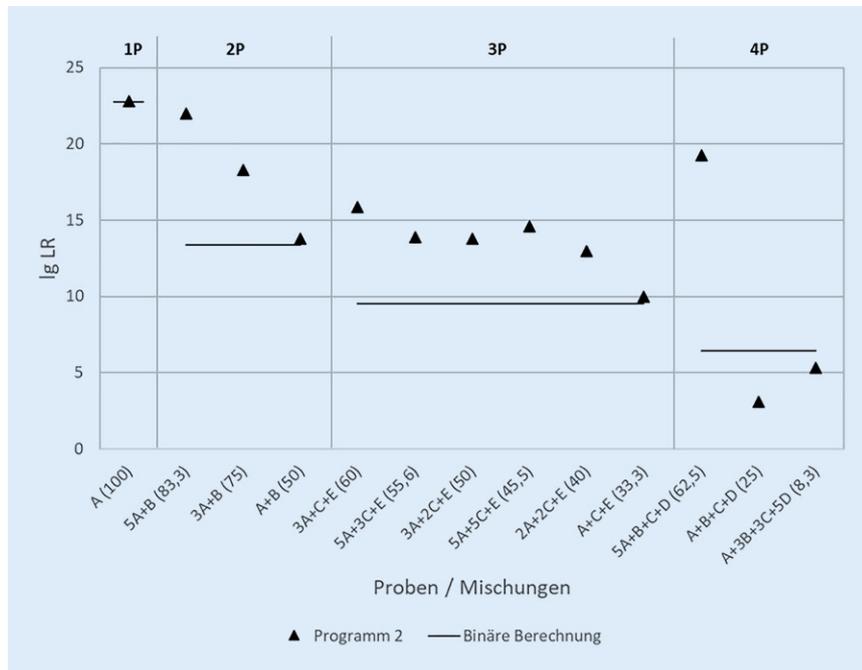


Abb. 2 ▲ Vergleich von vollkontinuierlichen LR_{fc} -Werten mit theoretisch erreichbaren binären LR_{bin} -Werten. Dargestellt sind die $lg-LR_{fc}$ -Werte einer Testreihe von 1P- bis 4P-(Misch-)Spuren, die aus unterschiedlichen Mischungsanteilen der benannten Personen A bis E bestehen und von links nach rechts nach Spurenlegeranzahl und Mischungsanteil der POI (Person A) aufgetragen wurden. Für jede (Misch-)Spur sind jeweils das Mischungsverhältnis der beteiligten Personen sowie der Mischungsanteil der Person A (Prozentwert in Klammern) angegeben. Um den errechneten $lg-LR_{fc}$ -Werten jeweils den theoretisch erreichbaren binären $lg-LR_{bin}$ -Wert gegenüberstellen zu können, wurden für Letztere sämtliche DNA-Merkmale der bekannten Spurenverursacher der jeweiligen Mischspur zu einem theoretischen DNA-Profil zusammengesetzt und als Typ-A-Mischspur binär berechnet (Linien). Für jede angegebene Mischspur wurden parallel Berechnungen mit den 4 getesteten Programmen durchgeführt. Repräsentativ für diese werden hier nur die errechneten $lg-LR_{fc}$ -Werte eines Programms gezeigt

meriert und als unbekannte Personen (UP) bezeichnet.

Exemplarisches Hypothesenpaar für eine 3P-Mischspur

Zusammensetzung der Spur:

Person B + Person D + Person E
(Mischungsverhältnis 5 : 3 : 1)

H1 : Person B(POI) + UP1 + UP2

H2 : UP1 + UP2 + UP3

Jedes Hypothesenpaar wurde einzeln mit den Ergebnissen der NGM-Detect™-Amplifikation und den Ergebnissen der Powerplex®-ESI 17-Fast-Amplifikation berechnet. Zusätzlich wurden bei den Programmen, die eine Kombination verschiedener Analysekits ermöglichten, bei gleicher Hypothesenstellung die Ergebnisse beider Kits in Kombination berechnet. Darüber hinaus wurden die Befunde der NGM-Detect™-Amplifikation

in Triplikaten berechnet, um die Varianz von Einzelberechnungen, welche bei der Berechnung mittels vollkontinuierlicher Modelle auftreten können, zu erfassen. Die Berechnungen der hier gezeigten Mischspuren erfolgten ausschließlich bei einer analytischen Nachweisgrenze von 50 rfu. Die Daten werden auf begründete Nachfrage zur Verfügung gestellt.

Ergebnisse

Vergleiche von binären und vollkontinuierlichen LR_{fc} -Berechnungen

Für den Vergleich von binären LR_{bin} mit vollkontinuierlichen LR_{fc} -Werten (LR_{fc} „fully continuous“) wurden Spuren aus Anteilen von 1 bis 4 Personen mit bekannten Mischungsverhältnissen analysiert. Ziel war es, für die vollkontinuierlichen Berechnungen Spurenserien mit

schwachen Merkmalsausprägungen und teilweise auch Drop-outs zu generieren, die gemäß den gemeinsamen Empfehlungen der Projektgruppe „Biostatistische DNA-Berechnungen“ und der Spurenkommission von 2016 für eine binäre Berechnung eher ungeeignet sind. Für die Berechnung von theoretisch erreichbaren binären LR_{bin} als Vergleichswerte wurden sämtliche DNA-Merkmale der bekannten Spurenverursacher der jeweiligen Mischspur zu einem theoretischen DNA-Profil zusammengesetzt (ohne Drop-outs; entsprechend einer Typ-A-Mischspur [14]).

Ein Spurenverursacher

Obwohl die analysierte Einzelspur (Spur A) lediglich einen DNA-Gehalt von 12 pg und daher z.T. deutliche Imbalancen (jedoch keine Drop-outs) aufwies, lieferten die Berechnungen der LR_{fc} mit allen getesteten vollkontinuierlichen Programmen sowie auch die Berechnung des theoretisch erreichbaren binären LR_{bin} nahezu identische Werte (Abb. 2).

Zwei Spurenverursacher

In Abb. 2 ist erkennbar, dass sich für die Einzelspur (Spur A) als auch für die Typ-B-Mischspur [14] mit deutlich abtrennbarer Hauptkomponente (Spur 5A + B) beinahe gleichwertige LR_{fc} -Werte errechnen. Mit abnehmendem Mischungsanteil der POI „A“ (z.B. bei den Spuren 3A + B und A + B) sinken auch die errechneten LR_{fc} -Werte, bis sie das Niveau einer ausgeglichenen A + B-Mischung (Mischspur Typ B) erreichen und sich dem theoretisch erreichbaren binären LR_{bin} -Wert der 2P-Mischspur (A + B) annähern.

Drei und vier Spurenverursacher

Gleiches trifft grundsätzlich auch für 3P-Mischspuren (A + B + C) und 4P-Mischspuren (A + B + C + D) zu. Dominiert eine Komponente also erheblich, so nähert sich der LR_{fc} -Wert wieder dem theoretisch erreichbaren LR_{bin} -Wert der Einzelspur an. Bei ausgeglichenen Mischungsanteilen (A + B, A + B + C oder A + B + C + D) fällt der LR_{fc} -Wert dagegen bis auf den theoretisch erreichbaren LR_{bin} -Wert der 3P-Mischspur ab oder sogar darunter, sofern für die POI A Merkmale in der Spur nicht nachgewiesen wurden (Drop-outs).

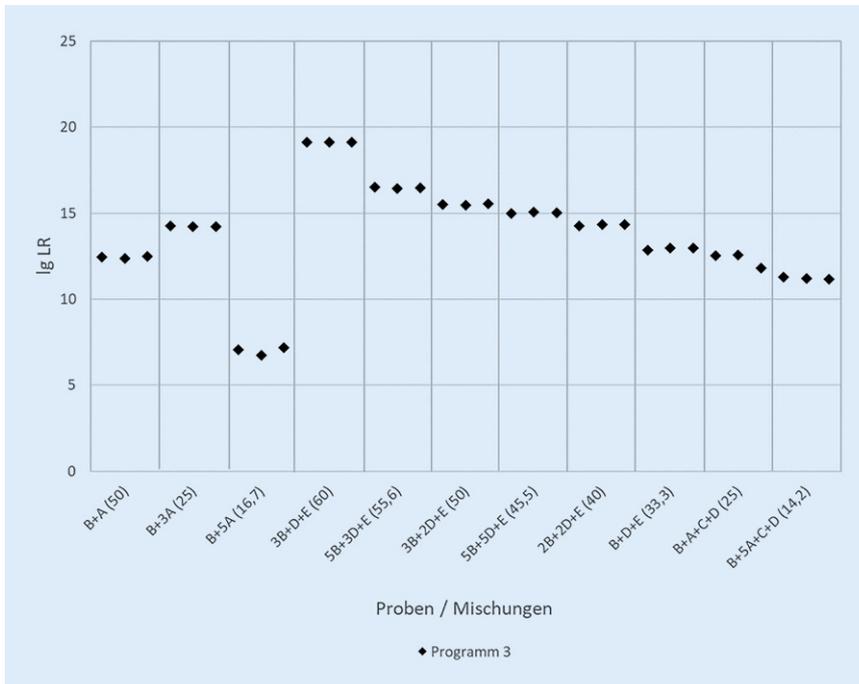


Abb. 3 ▲ Konkordanz innerhalb eines Programmes. Dargestellt sind die $\lg\text{-LR}_{fc}$ -Werte einer Testreihe von 2P- bis 4P-(Misch-)Spuren, die aus unterschiedlichen Mischungsanteilen der benannten Personen A bis E bestehen und von *links nach rechts* nach Spurenlegeranzahl und Mischungsanteil der POI (Person B) aufgetragen wurden. Für jede Mischspur sind jeweils das Mischungsverhältnis der beteiligten Personen sowie der Mischungsanteil der Person B (Prozentwert in Klammern) angegeben. Für jede angegebene Mischspur wurden parallel 3 Berechnungen mit den 4 getesteten Programmen durchgeführt. Repräsentativ für diese werden hier nur die $\lg\text{-LR}_{fc}$ -Werte des Programms 3 gezeigt. Für jede Mischspur sind jeweils die 3 einzelnen berechneten $\lg\text{-LR}_{fc}$ -Werte dargestellt

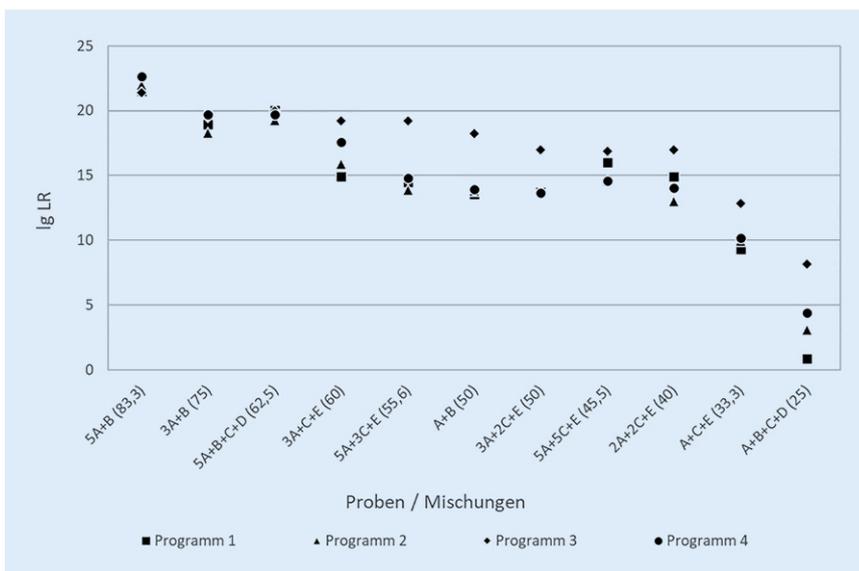


Abb. 4 ▲ Konkordanz zwischen den Programmen. Dargestellt sind die $\lg\text{-LR}_{fc}$ -Werte einer Testreihe von 1P- bis 4P-(Misch-)Spuren, die aus unterschiedlichen Mischungsanteilen der benannten Personen A bis E bestehen und von *links nach rechts* nach Mischungsanteil der POI (Person A) aufgetragen wurden. Für jede (Misch-)Spur sind jeweils das Mischungsverhältnis der beteiligten Personen sowie der Mischungsanteil der Person A (Prozentwert in Klammern) angegeben. Für jede angegebene Mischspur wurden parallel Berechnungen mit den 4 getesteten Programmen durchgeführt

In den 3P-Mischspuren unterschreiten in der hier gezeigten Testreihe die für die POI A errechneten LR_{fc} -Werte allerdings in keinem Fall den theoretisch erreichbaren LR_{bin} -Wert. Dieses Beispiel belegt eindrucksvoll, welche enorme Indizienkraft auch eine 3P-Mischspur mit ungünstigen Mischungsverhältnissen, die für eine binäre Berechnung nicht bzw. nur sehr eingeschränkt geeignet wäre, noch aufweisen kann. Zu interpretierende Peakimbilanzen bis hin zu Drop-outs für eine zu betrachtende POI können aber dazu führen, dass der LR_{fc} -Wert unter den theoretisch erreichbaren LR_{bin} -Wert abfällt:

So unterschreitet bei den gezeigten 4P-Mischspuren A + B + C + D und A + 3B + 3C + 5D die vollkontinuierliche Bewertung mit den getesteten Programmen den theoretisch erreichbaren binären LR_{bin} -Wert. Hier ist der Unterschied zu einer Mischspur mit einem Hauptanteil erwartungsgemäß besonders hoch.

Konkordanz der Programme

Bei binären Berechnungen innerhalb des gleichen Programms ergeben sich bei Verwendung gleicher Einstellungen und identischer Populationsdaten exakt gleiche LR-Werte; zwischen verschiedenen Programmen entstehen lediglich kleine Unterschiede durch Rundungen. Demgegenüber können sich für die gleiche Berechnung mit vollkontinuierlichen Programmen aufgrund spezifischer probabilistischer Modelle sowohl innerhalb des gleichen Programms als auch zwischen verschiedenen Programmen unterschiedliche LR_{fc} -Werte ergeben.

Im Rahmen der Testphase der Projektgruppe wurden sowohl die Unterschiede bei Mehrfachberechnungen mit dem gleichen Programm als auch Unterschiede bei Berechnungen mit verschiedenen Programmen untersucht.

Mehrfachberechnungen mit dem gleichen Programm

Aus den Mehrfachberechnungen identischer Hypothesenpaare erschließt sich, dass innerhalb der einzelnen verwendeten Programme in der Regel lediglich geringe Unterschiede bei den erhaltenen LR_{fc} -Werten zu erwarten sind. Die Unterschiede zwischen dem höchsten und

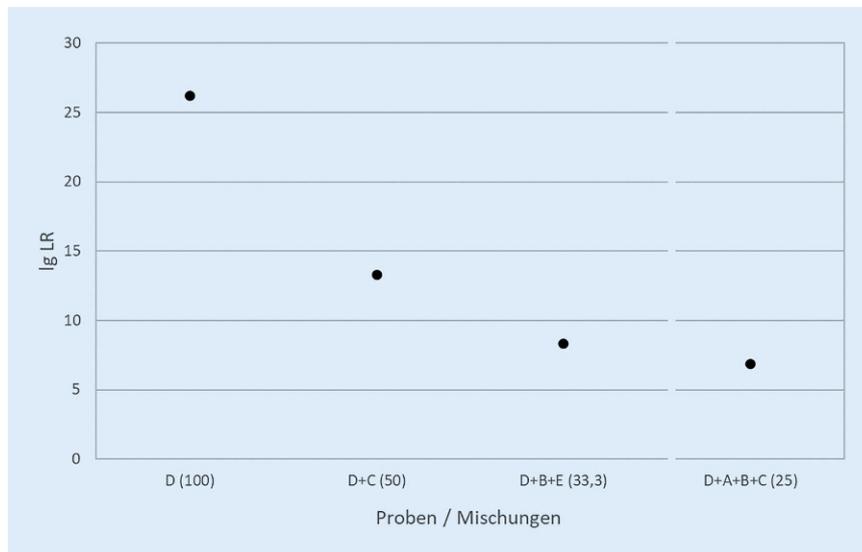


Abb. 5 ▲ Einfluss der Spurenlegeranzahl auf die LR_{fc} -Werte. Dargestellt sind die $lg-LR_{fc}$ -Werte einer Testreihe von 1P- bis 4P-(Misch-)Spuren, die jeweils aus gleichen ausgewogenen Mischungsanteilen der jeweils benannten Personen A bis E bestehen und von links nach rechts nach Spurenlegeranzahl aufgetragen wurden. Für jede (Misch-)Spur ist jeweils der Mischungsanteil der POI (Person D) angegeben (Prozentwert in Klammern). Für jede angegebene Mischspur wurden parallel Berechnungen mit den 4 getesteten Programmen durchgeführt. Repräsentativ für diese werden hier nur die errechneten $lg-LR_{fc}$ -Werte eines Programms gezeigt

niedrigsten berechneten LR_{fc} -Wert lagen bei den meisten Berechnungen unterhalb einer Zehnerpotenz (■ Abb. 3).

Divergieren die Ergebnisse von Mehrfachberechnungen mit einem Programm erheblich, kann dies ein Hinweis auf falsche Parametereinstellungen oder auf einen ungeeigneten DNA-analytischen Spurenbefund darstellen. Mehrfachberechnungen und die Höhe der festgestellten Abweichungen können daher als Hinweis für die Validität einer durchgeführten Berechnung verwendet werden.¹

Konkordanz zwischen mehreren Programmen

Studien zur Konkordanz verschiedener Programme belegen, dass es bei der Analyse von DNA-Spuren modellbedingt z.T. zu Abweichungen bei den errechneten LR_{fc} -Werten kommen kann. Diese Abweichungen fallen bei 3P- und 4P-Mischspuren deutlich stärker als bei Ein-

zelspuren oder 2P-Mischspuren aus [5]. Auch in unseren Testreihen erbrachten Berechnungen von 3P- bzw. 4P-Mischspuren mit unterschiedlichen Programmen z.T. größere Abweichungen (■ Abb. 4). So lagen die Differenzen zwischen dem niedrigsten und dem höchsten errechneten LR_{fc} -Wert in vielen Fällen bei mehr als 3 Zehnerpotenzen, im Extremfall sogar bei über 6 Zehnerpotenzen (4P-Mischspur A + B + C + D).

Die Unterschiede in den erhaltenen LR_{fc} -Werten sind schwierig zu bewerten, da zum einen eines der 4 Programme relativ konstant höhere LR_{fc} -Werte produziert hat und zum anderen die Unterschiede bei den LR_{fc} -Werten der 4 Programme auch in Relation zu den absoluten LR_{fc} -Werten zu sehen sind.

Beispielsweise wäre ein bei einer Berechnung mit 2 Programmen festgestellter Unterschied von 3 bis 6 Zehnerpotenzen bei LR_{fc} -Werten im Größenbereich bis zur 23. Zehnerpotenz zwar erheblich, jedoch verändern diese Abweichungen nicht die prinzipielle Kernaussage zur betrachteten Spurenlegerschaft. Denn beide errechneten LR_{fc} -Werte befinden sich weit jenseits der Grenze, bei der noch berechnete Zweifel an einer Spurenlegerschaft angebracht wären.

Dagegen wäre eine Abweichung von mehr als 6 Zehnerpotenzen bei geringen errechneten LR_{fc} -Werten als äußerst kritisch zu bewerten.

Betrachtet man die z.T. erheblichen Unterschiede in den berechneten LR_{fc} -Werten der unterschiedlichen Programme, führt dies zwangsläufig zu der Frage, ob diese valide Ergebnisse ausweisen.

Tatsächlich ist für den Anwender nicht festzustellen, wie nun der „wahre“ LR_{fc} -Wert einer Berechnung ausfallen müsste. Bei VKM machen es die komplexeren Rechenalgorithmen nahezu unmöglich, die Ergebnisse solcher Berechnungen „per Hand“ nachvollziehen zu können.

Was die Programme voneinander unterscheidet, ist die Anzahl der Berechnungsparameter, mit denen jeweils versucht wird, die Wirklichkeit abzubilden. Darüber hinaus sind auch mathematische Näherungsverfahren, die in keinem Fall identische Zahlenwerte produzieren, Bestandteil der LR_{fc} -Berechnungen.

Es ist also gar nicht so verwunderlich, dass die mithilfe von mehreren Programmen errechneten LR_{fc} -Werte in einem gewissen Maße streuten. Dies traf besonders für solche Spuren zu, bei denen es sich um 2P- bis 4P-Mischspuren mit einem nur geringen Mischungsanteil der zu betrachtenden POI handelte.

Vor dem routinemäßigen Einsatz von VKM wird es daher die zentrale Aufgabe einer Validierung sein, die Grenzen für die Anwendbarkeit der Methode auf Spuren und die Verwertbarkeit von erzielten Berechnungsergebnissen festzulegen (nähere Informationen hierzu sind den Empfehlungen der Begleitpublikation zu entnehmen [8]).

Anzahl der Spurenverursacher und Mischungsverhältnisse

Die Anzahl der anzunehmenden Spurenverursacher ist ein wesentlicher Bestandteil biostatistischer Berechnungen. Wie bei allen Berechnungsmethoden sinken auch bei vollkontinuierlichen Programmen grundsätzlich (Mischungen mit gleichen Mischungsanteilen vorausgesetzt) die LR_{fc} -Werte bei zunehmender Spurenlegeranzahl (■ Abb. 5). Dies ist in erster Linie der größeren Anzahl der Kombinationsmöglichkeiten der Allelwerte

¹ Anders als bei den weiteren getesteten Programmen wurden die Berechnungen im Programm EuroForMix mithilfe einer Maximum Likelihood Estimation (MLE) durchgeführt, bei der grundsätzlich identische LR -Werte generiert werden und kein MCMC-Algorithmus angewandt wird.

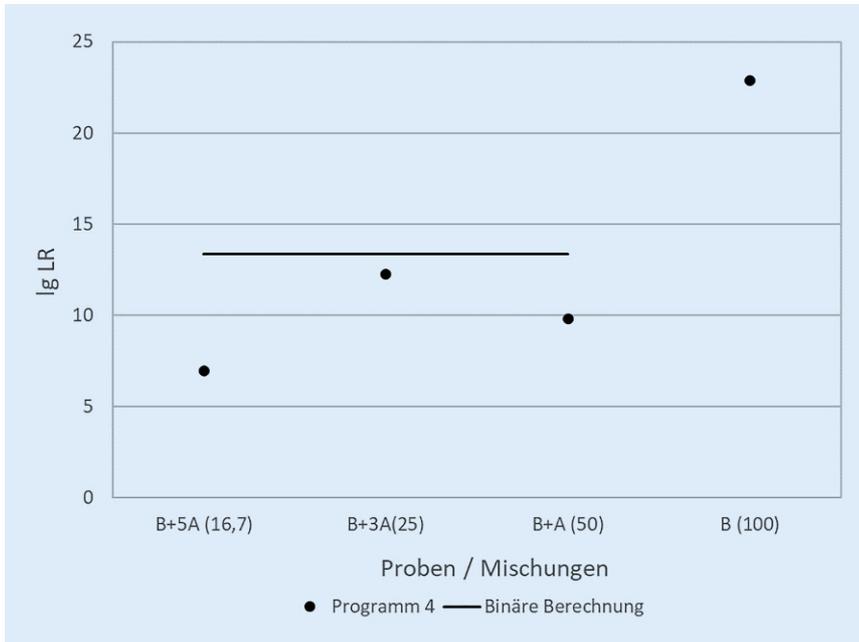


Abb. 6 ▲ Einfluss des Mischungsanteils auf die LR_{fc} -Werte. Dargestellt sind die $\lg-LR_{fc}$ -Werte einer Testreihe von 1P- und 2P-(Misch-)Spuren, die aus unterschiedlichen Mischungsanteilen der benannten Personen A und B bestehen und von links nach rechts nach Mischungsanteil der POI (Person B) aufgetragen wurden. Für jede (Misch-)Spur ist das Mischungsverhältnis der beteiligten Personen sowie der Mischungsanteil der Person B (Prozentwert in Klammern) angegeben. Um die errechneten $\lg-LR_{fc}$ -Werte der 2P-Mischspuren (B + A) einem theoretisch erreichbaren $\lg-LR_{bin}$ -Wert gegenüberstellen zu können, wurden für Letzteren die Allelwerte der Spurenverursacher A und B zu einem theoretischen DNA-Profil zusammengesetzt und als Typ A-Mischspur binär berechnet (Linie). Für die Einzelspur sind LR_{fc} und LR_{bin} identisch. Für jede angegebene Mischspur wurden parallel Berechnungen mit den 4 getesteten Programmen durchgeführt. Repräsentativ für diese werden hier nur die errechneten $\lg-LR_{fc}$ -Werte eines Programms gezeigt

geschuldet. Wie im Abschnitt „Vergleiche von binären und vollkontinuierlichen LR-Berechnungen“ beschrieben, unterscheiden sich die errechneten LR_{bin} - bzw. LR_{fc} -Werte für eine DNA-Spur mit gleichen Mischungsanteilen nicht wesentlich in ihren Größenordnungen. Binäre oder semikontinuierliche Modelle berücksichtigen keine Signalintensitäten und betrachten somit jedwede Mischspur als ausgewogen.

Im vollkontinuierlichen Modell ist der errechnete LR_{fc} -Wert zusätzlich von den jeweiligen Mischungsanteilen in der Spur abhängig: Je mehr sich die Mischungsanteile unterscheiden, desto stärker überschreitet der LR_{fc} -Wert für die fragliche Hauptkomponente den LR_{bin} -Wert oder den semikontinuierlichen LR-Wert der Mischspur [5].

Hebt sich der Mischungsanteil der fraglichen Hauptkomponente sehr deutlich von den restlichen Nebenkomponten ab (z.B. Mischungsanteil von 5:1), so nähert sich der zu berechnende LR_{fc} -Wert

sogar zunehmend dem LR_{bin} -Wert der Einzelspur der POI an (Abb. 2).

Für Nebenkomponten trifft dies ebenfalls zu, jedoch nur in eingeschränktem Maße. Denn mit sinkendem DNA-Gehalt bzw. Mischungsanteil greifen zunehmend Effekte wie Maskierungen, Heterozygotenimbalance sowie Drop-in- und Drop-out-Ereignisse und führen dadurch wieder zu einer Verringerung des LR_{fc} -Wertes [5].

Auch in unseren Testreihen waren diese Trends deutlich ablesbar (Abb. 6 exemplarisch für die POI B gezeigt). Allerdings erbrachten hier Berechnungen der analysierten 2P-Mischspuren zumeist LR_{fc} -Werte, die unterhalb des LR_{bin} -Wertes einer idealisierten 2P-Mischspur lagen. Dies ist vermutlich auf den Entwurf und die Anfertigung unserer Testreihe bzw. die sehr geringen DNA-Gehalte von nur 12 pg für die jeweils fragliche POI zurückzuführen. Aufgrund des geringen DNA-Gehaltes kam es in den analysier-

ten 2P-Mischspuren für die fragliche POI z.T. zu erheblichen Peakimbalance und damit zu entsprechenden Ergebnisverschlechterungen in den errechneten LR_{fc} -Werten.

Einfluss von Drop-outs

Eine häufige Problematik bei der Interpretation von DNA-Befunden stellen Mischspuren dar, bei denen aufgrund der niedrigen DNA-Menge des entsprechenden Mischungsanteils nicht sämtliche Merkmale vollständig nachgewiesen werden können. Bei vielen derartigen Spuren kommt der Sachverständige bei Betrachtung aller vorliegenden Befunde zu dem Ergebnis, dass die POI unter Annahme von Drop-out-Ereignissen die betreffende Mischspur mitverursacht haben kann.

Mit den derzeit verwendeten binären Berechnungsmethoden kann diese Aussage jedoch nicht bzw. nur unter eingeschränkter Einbeziehung geeigneter Merkmalsysteme biostatistisch bewertet werden. Ein wesentliches Ziel für die Anwendung von VKM ist die sachgerechte biostatistische Bewertung von Befunden hinsichtlich einer fraglichen Beteiligung einer POI an einer DNA-Spur unter Einbeziehung aller untersuchten Merkmalssysteme.

Um die verfügbaren VKM hinsichtlich ihrer Fähigkeit zur Berechnung von DNA-Spuren mit Drop-outs zu testen, wurden von der Projektgruppe DNA-Spuren mit so geringen Mischungsanteilen hergestellt, dass hierbei auch Befunde im Bereich der Nachweisgrenze und mit Drop-out-Ereignissen generiert wurden.

Ein wesentliches Ergebnis der VKM-Berechnungen dieser Spuren war, dass die VKM-Ergebnisse trotz des Vorliegens von einem bis 3 Drop-out-Ereignissen die korrekte Hypothese nach der gefragten und bei der Herstellung der DNA-Spur beteiligten POI unterstützten (Abb. 7).

Die erhaltenen LR_{fc} -Werte fielen bei DNA-Spuren mit Drop-outs signifikant niedriger aus als bei Mischspuren, in denen durchgängig die DNA-Merkmale der POI nachgewiesen wurden. So fielen diese LR_{fc} -Werte häufig unter den LR_{bin} -Wert einer hypothetischen Mischspur, die alle Merkmale der beteiligten Personen enthielt. Hierbei führten mehrere Drop-outs erwartungsgemäß zu kontinuier-

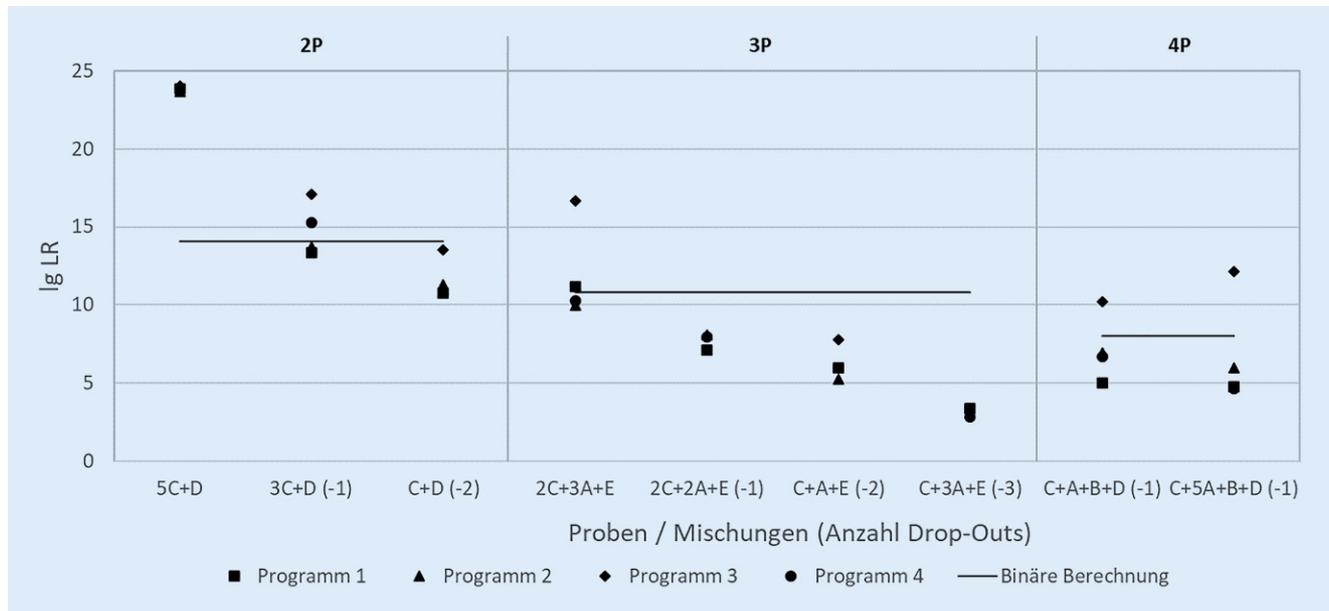


Abb. 7 ▲ Einfluss von Drop-out-Ereignissen auf die LR_{fc} -Werte. Dargestellt sind die $\lg-LR_{fc}$ -Werte einer Testreihe von 2P- bis 4P-Mischspuren, die aus unterschiedlichen Mischungsanteilen der benannten Personen A bis E bestehen und von *links nach rechts* geordnet nach Spurenlegeranzahl und dem Mischungsanteil der POI (Person C) aufgetragen wurden. Für jede Mischspur sind jeweils das Mischungsverhältnis der beteiligten Personen sowie die Anzahl der gemessenen Drop-out-Ereignisse der Person C (in Klammern) angegeben. Um die errechneten LR_{fc} -Werte den LR_{bin} -Werten gegenüberstellen zu können, wurden die Allelwerte der jeweils benannten Spurenverursacher A, B, C, D bzw. E zu hypothetischen Modellmischspuren zusammengesetzt. Auf diese Weise ließen sich die LR_{bin} -Werte für die entsprechenden 1P- bis 4P-(Misch-)Spuren berechnen (*Linien*). Für jede angegebene Mischspur wurden parallel Berechnungen mit den 4 getesteten Programmen durchgeführt

lich abnehmenden LR_{fc} -Werten. Für die 3P-Mischspur (C + 3A + E; POI Person C) führte die Berechnung der DNA-Spur mit 3 Drop-outs bei allen 4 Programmen zu LR_{fc} -Werten unter 10^4 .

Im Gegensatz dazu wiesen DNA-Spuren ohne Drop-out-Ereignisse in der Regel deutlich höhere LR_{fc} -Werte im Vergleich zu den errechneten zugehörigen LR_{bin} -Werten auf (Abb. 2).

Demnach führt die Berücksichtigung von Peakhöhen in einem VKM zu höheren LR_{fc} -Werten im Vergleich zu den nur hypothetisch berechenbaren LR_{bin} -Werten, gleichzeitig senkt aber jedes zur Erklärung einer DNA-Spur anzunehmende Drop-out-Ereignis diesen Wert wiederum ab. Beide Effekte beeinflussen das Ergebnis somit gegensätzlich.

Einfluss von Drop-out-Ereignissen auf die locuspezifischen LR_{fc} -Werte
Der Einfluss eines einzelnen Drop-out auf die LR_{fc} -Berechnung war mit den hergestellten Mischspuren nur qualitativ und nicht quantitativ zu erfassen, da sich die Spuren unabhängig vom beobachteten Drop-out auch sonst unterschieden.

Neben dem gewünschten Drop-out wurden auch immer zusätzliche Effekte wie Unterschiede in Signalintensitäten, Heterozygotenungleichheiten, Drop-in- sowie weitere Drop-out-Ereignisse generiert, die den LR_{fc} -Wert der Gesamtspur ebenfalls beeinflussten und bei der Berechnung mit einem VKM miterfasst wurden.

Für die systematische Untersuchung des Einflusses von Drop-outs auf die LR_{fc} -Werte müssten aber idealerweise 2 Spuren betrachtet werden, die in der Zusammensetzung der Allelwerte sowie in den Signalintensitäten gleich wären und sich nur um den Drop-out unterscheiden. Durch Mischen sind solche DNA-Spuren kaum zu erhalten. Um zumindest eine Vorstellung davon zu erlangen, in welchem Maße ein Drop-out den LR_{fc} -Wert einer DNA-Spur beeinflusst, wurden daher die Effekte von Drop-outs anhand künstlich generierter DNA-Spuren simuliert.

Zur Demonstration wurde eine DNA-Spur mit erkennbaren Imbalancen, jedoch ohne Drop-outs ausgewählt. Hierbei handelte es sich um eine 2P-Mischspur (A + B 1:1; Gesamt-DNA-Gehalt 24 pg). Für die zu betrachtende POI Person A errechnete sich

für die gesamte DNA-Spur ein $\lg-LR_{fc}$ -Wert von 14,52.

Ausgehend von dieser realen Mischspur wurden nun insgesamt 14 künstliche Mischspuren mit Drop-out-Ereignissen konstruiert. Hierzu wurden 14 Allele, welche die Person A nicht mit der Person B in der Mischspur teilte, einzeln und der Reihe nach aus den Rohdaten entfernt. Die betroffenen Allele wurden in ihren Signalintensitäten auf 0 rfu gesetzt und die relevanten Stutterpeaks entsprechend korrigiert. Um den Grad des Einflusses eines künstlich gesetzten Drop-out zu messen, wurde nun der locuspezifische Einfluss auf das jeweils betroffene Merkmalssystem berechnet. Die Auswirkung eines Drop-outs errechnete sich aus dem Quotienten des locuspezifischen LR_{fc} -Werts ohne Drop-out dividiert durch den locuspezifischen LR_{fc} -Wert mit Drop-out.

In Abb. 8 wurden die Einflusswerte der 14 Drop-outs halblogarithmisch gegen die Allelfrequenz des jeweils entfernten Allels aufgetragen. Es zeigte sich, dass das Entfernen eines Allels mit einer geringeren Verbreitungshäufigkeit generell einen höheren Einfluss auf den locusspe-

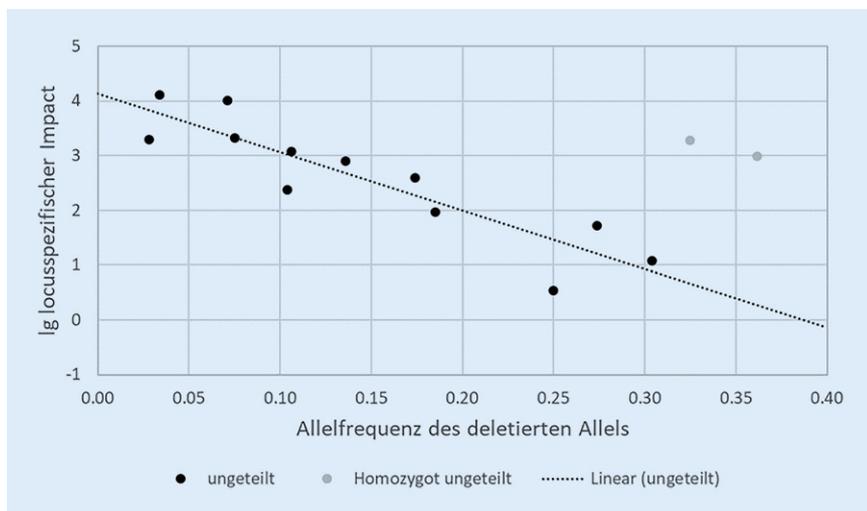


Abb. 8 ▲ Einfluss von Drop-out-Ereignissen auf die locuspezifischen LR_{fc} -Werte in Abhängigkeit von der Allelhäufigkeit. Ausgehend von einer realen 2P-Mischspur ($A + B$ 1:1) wurden insgesamt 14 künstliche Mischspuren mit je einem Drop-out-Ereignis generiert. Hierfür wurde jeweils ein ungeteiltes Allel der Person A gewählt, welches nicht mit der Person B im entsprechenden Locus geteilt wurde. Zur Einführung des Drop-out-Ereignisses wurde die Signalintensität des jeweiligen Allels zunächst auf 0 rfu reduziert. Um ein Maß für den Einfluss des jeweiligen Drop-outs auf den locuspezifischen LR_{fc} -Wert zu erhalten, wurden die LR_{fc} -Werte der originären und der künstlich veränderten Spur mit Drop-out-Ereignis im entsprechenden Locus miteinander verglichen. Dazu wurden zunächst alle locuspezifischen LR_{fc} -Werte beider Mischspuren mithilfe eines VKM berechnet. Aus den ursprünglichen LR_{fc} -Werten dividiert durch die entsprechenden LR_{fc} -Werte nach Setzen eines Drop-out-Ereignisses ergibt sich als Quotient der Einfluss des jeweiligen Drop-out auf die locuspezifischen LR_{fc} -Werte („impact“). Dazu wurden die Impact-Werte der 14 künstlich gesetzten Drop-out-Ereignisse halblogarithmisch gegen die Allelhäufigkeit aufgetragen. Ungeteilte heterozygote Allele sind in *schwarz*, ungeteilte homozygote Allele in *grau* dargestellt. Mittels linearer Regression wurde die Trendlinie ermittelt und in der Abbildung dargestellt

zifischen LR-Wert des betroffenen Merkmalssystems hat als das Entfernen eines Allels mit größerer Verbreitungshäufigkeit. Im Ergebnis führten Drop-out-Ereignisse in heterozygoten Allelen (*schwarz*) in der Regel zu Verringerungen der errechneten LR_{fc} -Werte. Bei der gezeigten 2P-Mischspur waren Reduktionen der locuspezifischen LR_{fc} -Werte um einen Faktor von bis zu 10^4 feststellbar. Drop-out-Ereignisse in homozygoten Allelen (*grau*) führten dagegen zu deutlich stärker verringerten locuspezifischen LR_{fc} -Werten, die oberhalb der eingezeichneten Regressionsgeraden lagen. Dieser Effekt ist nachvollziehbar, da ein VKM dem Drop-out eines homozygoten Allels bei der Bewertung ein höheres Gewicht als dem Drop-out eines heterozygoten Allels beimisst.

Ableitung recherchefähiger Datenbankmuster aus Mischspuren mithilfe vollkontinuierlicher Modelle (Deconvolution)

Klassische Mischspurinterpretationen basieren überwiegend auf der beruflichen Erfahrung des jeweiligen Sachverständigen. Aus den gemessenen Daten werden auch Rückschlüsse auf die wahrscheinliche Spurenlegeranzahl und die Mischungsanteile gezogen. Sofern möglich und erforderlich, erfolgt im Anschluss „per Auge“ die Ableitung der zugehörigen spurenbeteiligten Genotypen. Hierzu werden neben den gemessenen Signalen auch die Signalintensitäten und Amplitudenlängen berücksichtigt. Zudem werden analysesektorspezifische Eigenschaften wie z. B. die Amplifikations-effizienz oder das Auftreten von Stutterpeaks ebenso wie eine evtl. Degradation des Spurenmaterials, evtl. Peakimbilanzen oder das mögliche Auftreten von Drop-in- und Drop-out-Ereignissen in die Auswertung einbezogen.

Im Regelfall unterscheiden sich die Ergebnisse derartiger Bewertungen zwischen verschiedenen Sachverständigen qualitativ nicht. Da dieses Verfahren jedoch subjektiven Einflüssen unterliegt, sind bei problematischen Spurenlagen im Detail auch Abweichungen möglich. So kann es beispielsweise vorkommen, dass ein Sachverständiger einer spurbeteiligten Person mehr gemessene Allele zuordnet als ein anderer.

Mit VKM lassen sich bei entsprechender Spurenlage Prognosen über spurenbeteiligte Genotypen aufstellen. Diese Auflösung in einzelne Spurenkomponenten wird im VKM als Deconvolution bezeichnet. Im Gegensatz zum klassischen visuellen Ansatz basieren die Ergebnisse einer VKM-basierten Deconvolution auf mathematischen Algorithmen und sind unabhängig vom Sachverständigen. Darüber hinaus werden die prognostizierten Genotypen mit ihren errechneten Wahrscheinlichkeiten angegeben.

Zur Verdeutlichung des Prinzips der Deconvolution ist in **Abb. 9** schematisch eine problematische Mischspur dargestellt. Im Elektropherogramm (**Abb. 9a**) ist erkennbar, dass hier eine Mischung von mindestens 2 Personen vorliegt. Anhand des MMS 1 ließe sich ggf. visuell die Arbeitshypothese aufstellen, dass diese Mischspur von 2 Personen stammt, von denen eine Person einen größeren Mischungsanteil als die zweite Person beiträgt. Diese Arbeitshypothese lässt sich aber mit den Befunden in den MMS 3 und 4 nicht in Einklang bringen. Es bleibt unklar, welche Genotypen den beiden Personen UP1 und UP2 zuzuordnen sind.

Im Gegensatz zu diesem visuellen Ansatz zeigt **Abb. 9b**, wie die einzelnen Peaks im VKM in ihre Komponenten quantitativ aufgetrennt und so den Personen UP1 und UP2 oder aber Stutterpeaks zugeordnet werden können.

Ist diese wahrscheinlichkeitsbasierte Prognose des Genotyps eines Spurenlegers plausibel und überschreitet diese einen festgelegten Mindestwahrscheinlichkeitswert, kann der entsprechende Genotyp, sofern er die weiteren Anforderungen der DAD erfüllt, für eine Datenbankrecherche verwendet werden (**Abb. 9c, d**). Auch in einem VKM ist diese

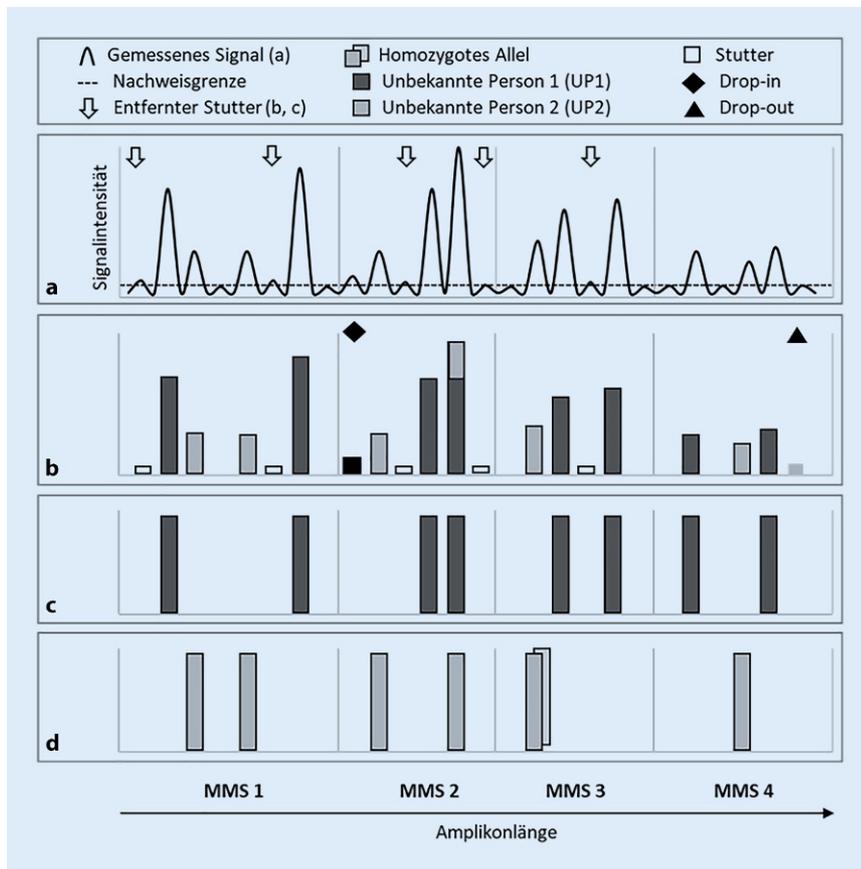


Abb. 9 ▲ Schematische Darstellung einer Deconvolution mit 2 unbekannt Personen (UP1 und UP2). **a** Schematische Darstellung eines Elektropherogramms. Stellvertretend für das Gesamtprofil werden hier lediglich 4 Merkmalssysteme (MMS) dargestellt. *Gestrichelte Linie:* Nachweisgrenze. **b** Darstellung der quantitativen Signalintensitäten als Summation der individuellen Anteile. Die in weiß dargestellten Signale werden weder der UP1 noch der UP2 zugeordnet. Es handelt es sich hierbei um Stutterpeaks. **c, d** Exemplarisch abgeleitete Profile für die Personen UP1 und UP2. Für die Person UP2 liegt im MMS 3 eine Homozygotie für das abgeleitete Allel vor

Sicherheit für den im MMS 4 vorliegenden Drop-out der Person UP2 nicht gegeben.

Testreihe zur Deconvolution von 2P-Mischspuren

Um die Funktionsweise und die technischen Grenzen der vollkontinuierlichen Deconvolution auszuloten, wurden seitens der Projektgruppe zahlreiche Berechnungen durchgeführt. Unter anderem wurden hierzu auch mehrere 2P-Mischspuren bekannter Zusammensetzung analysiert. **Abb. 10** zeigt Ausschnitte aus den Elektropherogrammen (lediglich 4 MMS) von 3 exemplarisch analysierten Mischspuren, die sich aus Anteilen von 2 bekannten Spurenverursachern A und B in verschiedenen Mischungsverhältnissen (5:1, 3:1 und 1:1) zusammensetzen. Im Folgenden werden ausschließlich Ergebnisse der De-

convolution für den Spurenverursacher A betrachtet.

2P-Mischspur (A + B 5:1)

Abb. 10a und **Tab. 2** zeigen eine 2P-Mischspur mit den Spurenverursachern A und B im Mischungsverhältnis 5:1. Die deutlich unterschiedlichen Mischungsanteile erlauben eine relativ klare Abtrennung von Haupt- und Nebenkomponente. Auch mit der visuellen Methode kann hier in allen 16 Merkmalssystemen ein nachvollziehbarer Genotyp mit einer relativ hohen Sicherheit abgeleitet werden. Vermutlich würden in diesem Fall sowohl die visuelle Methode als auch die vollkontinuierliche Deconvolution identische und vollständige Genotypen in 16 Merkmalssystemen erbringen (**Tab. 2**).

2P-Mischspur (A + B 3:1)

Bei Mischspuren mit ausgewogeneren Mischungsanteilen ist der visuelle Ansatz der vollkontinuierlichen Methode erheblich unterlegen. Bereits bei einem Mischungsverhältnis von 3 zu 1 ist in der exemplarisch analysierten Mischspur die Hauptkomponente mit dem Auge nur noch unzureichend und nicht mehr mit erforderlicher Sicherheit bestimmbar (**Abb. 10b** und **Tab. 3**). In mehreren STR-Systemen ist der Genotyp des Spurenverursachers A auf diese Weise nicht mehr zuverlässig ableitbar (z.B. D16S539 und D18S51).

Die VKM-basierte Deconvolution leitet dagegen auch für diese Mischspur einen reproduzierbaren und statistisch gewichteten Genotyp für den Spurenverursacher A ab. Hierbei wurden für den prognostizierten Genotypen im System D18S51 eine Einzelwahrscheinlichkeit von über 95% und für die Genotypen in den Systemen TH01 und D12S391 Einzelwahrscheinlichkeiten von jeweils über 99% errechnet (siehe **Tab. 3**). Dagegen errechnete sich für den prognostizierten Genotyp im System D16S539 lediglich eine Einzelwahrscheinlichkeit von weniger als 95%.

Ob die mithilfe eines VKM abgeleiteten Merkmale für eine Recherche in der DAD verwendbar sind, richtet sich maßgeblich nach der festzulegenden Mindestwahrscheinlichkeit für das Zutreffen eines in einem Merkmalssystem prognostizierten Genotyps. Dabei beeinflusst die Höhe dieses Grenzwertes für die Mindestwahrscheinlichkeit die im zu recherchierenden Genotyp zu berücksichtigende Allelzahl.

Legt man beispielsweise die Mindestwahrscheinlichkeit auf einen relativ niedrigen Wert fest (z.B. 95%), so können für die in der **Abb. 10b** gezeigte 2P-Mischspur (A + B 3:1) zwar die abgeleiteten Allele aus 11 untersuchten Merkmalssystemen für eine Recherche verwendet werden, gleichzeitig errechnet sich aber aus den jeweiligen Einzelwahrscheinlichkeiten für das Zutreffen der Prognose (je Merk-

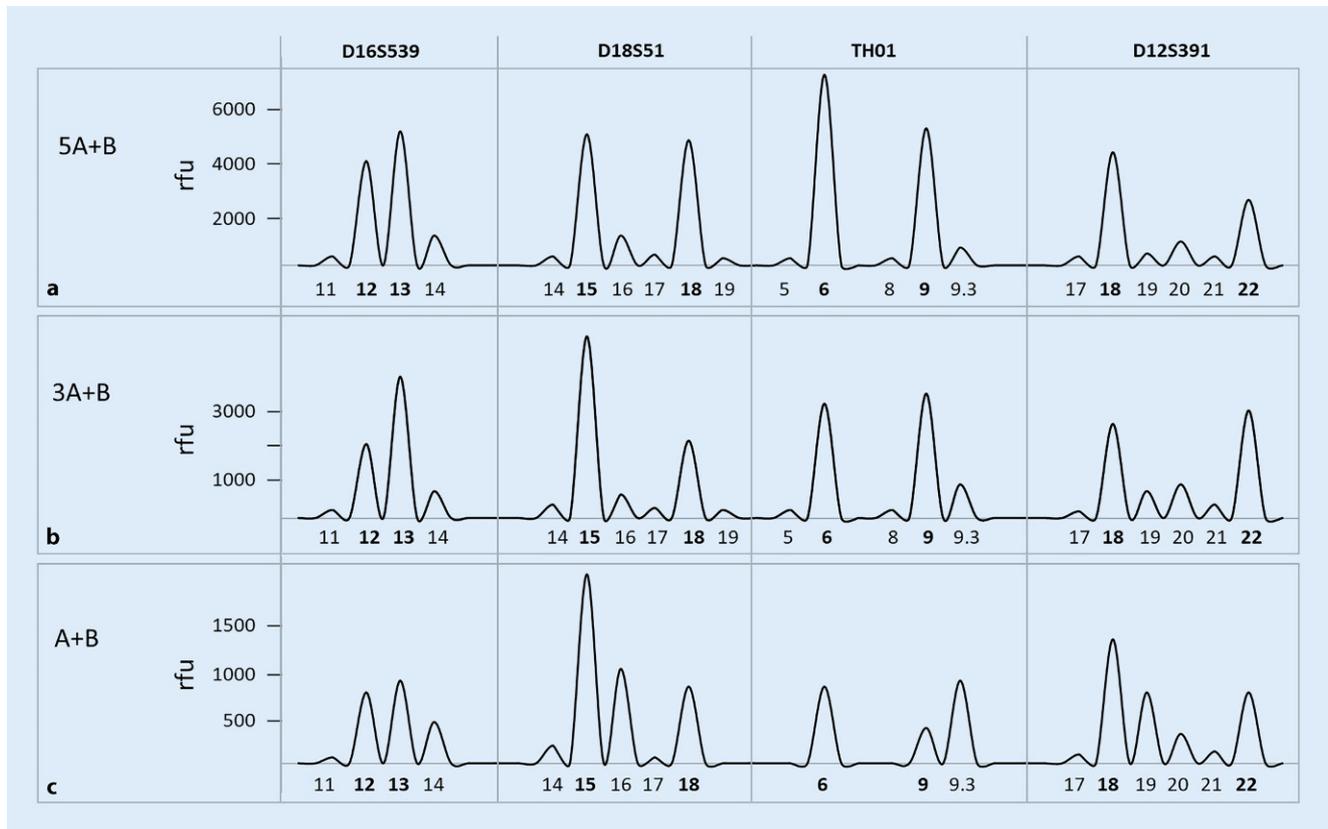


Abb. 10 ▲ Elektropherogrammausschnitte aus 2P-Mischspuren (A + B). Gezeigt sind Elektropherogrammausschnitte von 2P-Mischspuren (A + B) mit den Mischungsverhältnissen 5:1 (a); 3:1 (b) und 1:1 (c). Ausschnittsweise werden stellvertretend für die 16 analysierten Merkmalsysteme (MMS) lediglich die Ergebnisse von 4 Merkmalsystemen gezeigt. Die Merkmale der POI (Person A) sind in **Fettdruck** dargestellt. Die Berechnungen zur Deconvolution der POI wurden stellvertretend für die anderen VKM-Programme exemplarisch nur mit einem Programm unter identischen Einstellungen durchgeführt. Die Ergebnisse der Deconvolutionen in den 4 gezeigten Merkmalsystemen sind den **Tab. 2, 3 und 4** zu entnehmen

malssystem) mit ca. 12,64% eine relativ hohe Unsicherheit² für dieses Profil.

Legt man die Mindestwahrscheinlichkeit dagegen auf einen höheren Wert fest (z. B. 99%), sinkt die Anzahl verwendbarer Allele auf lediglich 5 Merkmalsysteme ab. Gleichzeitig sinkt aber auch die Unsicherheit für den in 5 Merkmalsystemen abgeleiteten Genotypen auf nur noch 1,41% ab.

2P-Mischspur (A + B 1:1)

Bei ausgewogenen Spurenanteilen ist erwartungsgemäß auch mit VKM keine Ableitung einzelner Genotypen mehr möglich (**Abb. 10c** und **Tab. 4**). Zwar benennt das hier verwendete Programm noch einen wahrscheinlichsten Geno-

typen, aber die Wahrscheinlichkeiten für das Zutreffen dieser Prognosen sind viel zu gering, um hieraus mit hinreichender Sicherheit ein Merkmalsmuster für die Person A abzuleiten (**Tab. 4**). In keinem der 16 analysierten Merkmalsysteme wird ein Wahrscheinlichkeitswert von mehr als 95% für das Zutreffen der Prognose erreicht.

Komplexere Mischspuren

Im Idealfall kann mithilfe der vollkontinuierlichen Deconvolution neben einer Hauptkomponente ggf. auch der Genotyp einer Nebenkompente abgeleitet werden. Zudem sind auch Deconvolutionen von komplexeren Mischspuren (z. B. 3P- oder 4P-Mischspuren) bei entsprechenden Spurenlagen möglich.

In Abstimmung mit der Spurenkommission hat die Projektgruppe „Biostatistische DNA-Berechnungen“ auch an der VKM-Berechnung im Rahmen des GEDNAP-Ring-

versuches 2019 (GEDNAP 56-LR₁) teilgenommen. In der Ringversuchsaufgabe sollte eine 3P-Mischspur („Opfer“, „Tatverdächtiger“ und eine weitere Person) mit den folgenden Hypothesen berechnet werden:

H1 : Opfer + Tatverdächtiger + UP1

H2 : Opfer + UP1 + UP2

Im Rahmen des 39. GEDNAP-Spurenworkshops der DGRM wurden neben weiteren Resultaten auch die mit den 4 Programmen erzielten Ergebnisse zur Deconvolution (Opfer + UP1 + UP2) präsentiert und diskutiert. Wenn auch mit allen Programmen jeweils keine vollständigen Befunde erzielt wurden, waren die Ergebnisse der Deconvolution durchgängig konkordant (**Abb. 11**).

Dieses Ergebnis demonstriert exemplarisch die Leistungsfähigkeit der 4 getesteten Programme, denn bei der Beschaffenheit der analysierten GEDNAP 56-LR₁-

² Die Unsicherheit bezeichnet hier die Wahrscheinlichkeit, dass der in n Merkmalsystemen abgeleitete Genotyp mindestens in einem Allelwert falsch prognostiziert wird.

Tab. 2 Ausschnitt aus der Deconvolution einer 2P-Mischspur (A + B 5:1)

Locus	Genotyp des Spurenverursachers A	Prognose des wahrscheinlichsten Genotyps für Person A in der Mischspur	Wahrscheinlichkeit für den prognostizierten Genotyp	Recherchefähiges Muster bei einer Mindestwahrscheinlichkeit von	
				> 95 %	> 99 %
D16S539	12/13	12/13	1	12/13	12/13
D18S51	15/18	15/18	1	15/18	15/18
TH01	6/9	6/9	1	6/9	6/9
D12S391	18/22	18/22	1	18/22	18/22
Gesamtanzahl abgeleiteter Systeme in 16 Systemen	–	16 Systeme	–	16 Systeme	16 Systeme
Unsicherheit für das Gesamtprofil (mindestens eine Abweichung)	–	0,02 %	–	~ 0 %	~ 0 %

Die gezeigten Berechnungen erfolgten für die POI Person A stellvertretend für die anderen VKM-Programme exemplarisch unter identischen Einstellungen mit nur einem Programm. Ausschnittsweise werden stellvertretend für die 16 analysierten Merkmalsysteme (MMS) lediglich die Ergebnisse von 4 Merkmalsystemen gezeigt. Der prognostizierte wahrscheinlichste Genotyp ist in **Fettdruck** dargestellt

Tab. 3 Ausschnitt aus der Deconvolution einer 2P-Mischspur (A + B 3:1)

Locus	Genotyp der Person A	Prognose des wahrscheinlichsten Genotyps für Person A in der Mischspur	Wahrscheinlichkeit für den prognostizierten Genotyp	Recherchefähiges Muster bei einer Mindestwahrscheinlichkeit von	
				> 95 %	> 99 %
D16S539	12/13	12/13	0,8175	–	–
D18S51	15/18	15/18	0,9712	15/18	–
TH01	6/9	6/9	0,9994	6/9	6/9
D12S391	18/22	18/22	0,9976	18/22	18/22
Gesamtanzahl abgeleiteter Systeme	–	16 Systeme	–	11 Systeme	5 Systeme
Unsicherheit für das Gesamtprofil (mindestens eine Abweichung)	–	78,40 %	–	12,64 %	1,41 %

Die gezeigten Berechnungen erfolgten für die POI Person A stellvertretend für die anderen VKM-Programme exemplarisch unter identischen Einstellungen mit nur einem Programm. Ausschnittsweise werden stellvertretend für die 16 analysierten Merkmalsysteme (MMS) lediglich die Ergebnisse von 4 Merkmalsystemen gezeigt. Der prognostizierte wahrscheinlichste Genotyp ist in **Fettdruck** dargestellt

Tab. 4 Ausschnitt aus der Deconvolution einer 2P-Mischspur (A + B 1:1)

Locus	Genotyp des Spurenverursachers A	Prognose des wahrscheinlichsten Genotyps für Person A in der Mischspur	Wahrscheinlichkeit für den prognostizierten Genotyp	Recherchefähiges Muster bei einer Mindestwahrscheinlichkeit von	
				> 95 %	> 99 %
D16S539	12/13	12/13	0,3354	–	–
D18S51	15/18	15/16*	0,3318	–	–
TH01	6/9	6/9	0,3323	–	–
D12S391	18/22	18/22	0,1943	–	–
Gesamtanzahl abgeleiteter Systeme	–	16 Systeme	–	0 Systeme	0 Systeme
Unsicherheit für das Gesamtprofil (mindestens eine Abweichung)	–	~ 100 %	–	–	–

Die gezeigten Berechnungen erfolgten für die POI Person A stellvertretend für die anderen VKM-Programme exemplarisch unter identischen Einstellungen mit nur einem Programm. Ausschnittsweise werden stellvertretend für die 16 analysierten Merkmalsysteme (MMS) lediglich die Ergebnisse von 4 Merkmalsystemen gezeigt. Der prognostizierte wahrscheinlichste Genotyp ist in **Fettdruck** dargestellt. Eine fehlerhafte Prognose des verwendeten Programms im System D18S51 wurde mit einem **Asteriskus** gekennzeichnet

	Programm 1 ■	Programm 2 ▲	Programm 3 ◆	Programm 4 ●
D1S1656	11 / 15	11 / 15	11 / 15	11 / 15
D2S441	10 / 13	10 / 13	10 / 13	10 / 13
D10S1248	13 / -	13 / -	13 / 14	13 / 14
D12S391	16 / 19	16 / 19	16 / 19	16 / 19
D22S1045	16 / -	16 / -	16 / 16	16 / -
D16S539	12 / -	12 / -	10 / 12	10 / 12
D2S1338	17 / 25	17 / 25	17 / 25	17 / 25
D19S433	13 / -	13 / -	- / -	13 / -
SE33	14 / -	14 / 15	14 / 15	14 / 15
D21S11	29 / 31.2	29 / 31.2	29 / 31.2	29 / 31.2
VWA	17 / -	17 / -	16 / 17	17 / -
TH01	9 / 9.3	9 / 9.3	9 / 9.3	9 / 9.3
FGA	23 / -	23 / -	- / -	23 / -
D3S1358	15 / 17	15 / 17	15 / 17	15 / 17
D8S1179	13 / 16	- / -	- / -	16 / -
D18S51	17 / 20	17 / 20	17 / 20	17 / 20

vollständig
 unvollständig
 fehlerhaft

Abb. 11 ▲ Deconvolution GEDNAP-56 LR_{ic}-Mischspur. Dargestellt sind die Ergebnisse von Deconvolutionen der GEDNAP 56-LR_{ic}-Mischspur mit den getesteten Programmen DNA•VIEW Mixture Solution, STRmix, GenoProof Mixture und EuroForMix. Die Mindestwahrscheinlichkeit für akzeptierte Allele wurde für diese Analyse auf 99% festgesetzt. Vollständig und korrekt deduzierte Systeme sind in weiß, unvollständig deduzierte Systeme in hellgrau dargestellt. In allen 16 analysierten Merkmalssystemen erbrachten die Deconvolutionen der getesteten 4 Programme korrekte Prognosen

Mischspur wäre eine visuelle Ableitung nur sehr unvollständig möglich gewesen.

Fazit

Vorteile von VKM

VKM ermöglichen unter Berücksichtigung aller vorhandenen Analysedaten eine Abschätzung, wie wahrscheinlich unter gegebenen Bedingungen ein gemessenes Analyseergebnis zu erwarten ist. Dabei besteht das gemeinsame Prinzip aller Programme mit VKM darin, die bestmögliche Erklärung für die Entstehung einer Spur zu finden. Dieser probabilistische Ansatz vollkontinuierlicher Modelle erlaubt es, eine weitestgehend von subjektiven Einflüssen bereinigte Bewertung von DNA-Spuren unter Einbeziehung aller untersuchten Merkmalssysteme vorzunehmen.

Mit dem klassischen binären Berechnungsansatz konnte eine Vielzahl von Mischspuren und schwachen DNA-Spuren nicht oder nur unzureichend biostatistisch bewertet werden. Solche Spuren waren bisher zumeist lediglich in verbaler Form bewertbar. Im Gegensatz dazu lassen

sich solche Spuren bis zu einem gewissen Grad mithilfe von VKM unter Angabe eines LR_{ic}-Wertes bewerten. Der unvollständige Nachweis einiger DNA-Merkmale einer POI führt in diesen Modellen zu keinem Ausschluss.

Die Auftrennung von Mischspuren in ihre Einzelkomponenten (Deconvolution) ist ein integraler Bestandteil von VKM. Sie kann im Sinne einer Objektivierung bei der Ableitung recherchefähiger Muster genutzt werden. Unserer Erfahrung nach unterschieden sich bei einer gut ausgeprägten Hauptspurenkomponente (z. B. bei einem Mischungsverhältnis von 5:1) die gutachterlichen Interpretationen per Auge nicht wesentlich von den Ergebnissen der Deconvolution vollkontinuierlicher Modelle. Bei ungünstigeren Mischungsverhältnissen (z. B. 3:1) offenbarten VKM ihre besondere Stärke. Hier waren die Prognosen zu den an einer DNA-Spur beteiligten Genotypen im Vergleich zu denen erfahrener Sachverständiger präziser und mathematisch nachvollziehbar. Zum Teil konnten auch in Systemen, in denen Peakhöhenverhältnisse von weniger als 4 zu 1 vorlagen [14], Auftrennungen

der einzelnen beteiligten Genotypen mit hoher Sicherheit erreicht werden.

Es ist davon auszugehen, dass der Einsatz von VKM zu einer steigenden Anzahl erstellbarer Meldebögen bei gleichzeitiger Verringerung von Abweichungen/Fehlern führen wird. Infolgedessen steigt auch die Chance auf einen Datenbanktreffer.

Zudem ist durch den Einsatz von VKM auch eine weitere Stufe einer Automatisierung in der Datenbewertung denkbar, bei der alle anfallenden Spuren zunächst eine programmgestützte Deconvolution durchlaufen und die Ergebnisse erst dann von einem Sachverständigen bewertet werden [13]. Insbesondere bei der routinemäßigen Fallbearbeitung können VKM für den Sachverständigen eine enorme Hilfe darstellen.

Allerdings sind auch mit VKM nicht alle DNA-Spuren sinnvoll bewertbar. Sehr geringe DNA-Mengen, eine große Anzahl an Spurenverursachern oder ungünstige (weil z. B. ähnliche) Mischungsanteile bringen die VKM an ihre Grenzen.

VKM in der Fallarbeit

Schulung

Um die Theorie von VKM sowie die Konzepte, Funktionalitäten und Grenzen von VKM-Programmen zu verstehen und einen sachgerechten Gebrauch dieser Programme zu gewährleisten, ist eine Schulung von Anwendern unabdingbar. Weitreichendere mathematische Kenntnisse sind jedoch nicht erforderlich.

Hardware

VKM-Berechnungen können sehr komplex sein und erhebliche Hardwareressourcen beanspruchen. So wurden bei der Berechnung von 3P- oder 4P-Mischspuren im Rahmen der Testphase auch Rechenzeiten von mehr als 24 h bis hin zu Programmabbrüchen beobachtet. Da dies bei typischen Bürorechnern mit einer normalerweise völlig ausreichenden Konfiguration (z. B. i5-Prozessor mit 4 Kernen und 8 GB RAM) dazu führen würde, dass parallel kein weiteres Arbeiten möglich ist, sollte über eine Auslagerung von VKM-Berechnungen auf speziell für diese Zwecke ausgerüstete Rechner/Server nachgedacht werden.

Laborroutine

Für eine vollkontinuierliche Auswertung sind Änderungen in der Laborroutine notwendig. Im Rahmen der bisherigen Praxis erfahren die bei der STR-Analyse gewonnenen Rohdaten im Verlauf ihrer Auswertung und vor der Bewertung bereits Modifikationen. So werden mithilfe von Grundeinstellungen beispielsweise Artefakte und Stutterpeaks entfernt; Analysesignale werden oberhalb der Nachweisgrenze gewertet. Für eine vollkontinuierliche Auswertung ist es in der Regel dagegen nicht erforderlich, Stutterpeaks zu entfernen. Die Analyseprogramme müssen daher in der Lage sein, sowohl binäre als auch vollkontinuierliche Auswertungen zu ermöglichen, zu exportieren und zu speichern.

Dabei ist die Generierung einer geeigneten Schnittstelle eine nicht zu unterschätzende Aufgabe. So müssen Informationen wie Analysekit, Peakhöhe, DNA-Fragmentlänge und Allelbezeichnung aus den initial vorliegenden Datenformaten in die jeweiligen VKM-Programme sicher importiert werden. Gegebenenfalls muss hierzu auf externe Unterstützung zurückgegriffen werden, wobei entsprechende Kosten entstehen können.

Implementierung und Validierung

Je nach verwendetem VKM erfordern die Programme teils unterschiedliche Anpassungen von einstellbaren laborspezifischen Parametern für die Berechnung. Einige Programme machen detaillierte Vorgaben bezüglich der durchzuführenden Testreihen für die Implementierung, bei anderen Programmen gibt es Voreinstellungen, die überprüft und ggf. angepasst werden sollten.

Eine Validierung der Programme [7] durch die Analyse von Spuren bekannter Zusammensetzung ist vor einem Einsatz in der Fallarbeit unabdingbar. Dabei sollte die Auswahl der Testspuren die in der Praxis vorkommende Spurenvelfalt nach Möglichkeit abdecken. Hierbei ist zu beachten, dass die Programme auf artifizielle, von der Realität abweichende Bedingungen z. T. sehr empfindlich reagieren. So zeigten unsere Versuchsreihen beispielsweise, dass künstliche Mischspuren aus einzelnen DNA-Isolaten im Vergleich zu Mischspuren ausgehend von Zellproben weniger zur Validierung geeignet sind [10]. Eine DNA-

Spur sollte sinnvollerweise im selben Labor gemessen und mithilfe eines in diesem Labor implementierten VKM-Programmes berechnet werden.

Diskussion zu den getesteten Programmen

Die Modelle der 4 getesteten Programme sind in unterschiedlichem Umfang in Veröffentlichungen beschrieben. Von uns analysierte Testreihen mit z. T. extremen Mischspurenzusammensetzungen belegen die Funktions- und Leistungsfähigkeit aller vier Programme.

Trotz feststellbarer Funktionalitätsunterschiede war jedes Programm in der Lage, die angefertigten Testspuren zu berechnen. Dabei zeigten die erzielten LR_{IC} -Werte je nach verwendeten VKM und programminternen Parametern teilweise aber Unterschiede. Hierzu könnten die fehlenden Validierungen der einzelnen Programme beigetragen haben. Zudem ergaben sich bei den Berechnungen mitunter erheblich abweichende Rechenzeiten zwischen den Programmen.

Die Ergebnisse der Deconvolution waren bei jedem Programm korrekt, hinsichtlich der Anzahl abgeleiteter Allele und deren Prognosewahrscheinlichkeiten unterschieden sich jedoch die Ergebnisse.

In den Abbildungen und Diskussionen wurden die zugrunde liegenden Programme bewusst nicht ausgewiesen. Hiermit sollte der Eindruck vermieden werden, dass ein Programm aufgrund bestimmter Eigenschaften zu empfehlen sei.

Empfehlungen zur Arbeit mit VKM

Um den Sachverständigen beim Gebrauch vollkontinuierlicher Programme Handlungssicherheit zu geben und eine Harmonisierung der berichteten LR_{IC} -Werte zu gewährleisten, wurden von der Projektgruppe „Biostatistische DNA-Berechnungen“ und der Spurenkommission begleitende Empfehlungen zur Anwendung von VKM in der Fallarbeit ausgearbeitet (s. Gemeinsame Empfehlungen der Projektgruppe „Biostatistische DNA-Berechnungen“ und der Spurenkommission zur biostatistischen Bewertung DNA-analytischer Befunde unter besonderer

Berücksichtigung vollkontinuierlicher Modelle (VKM) [8]).

Die erarbeiteten Empfehlungen greifen die im Rahmen dieser Veröffentlichung diskutierten Aspekte auf und versuchen, diese in konkrete Handlungsempfehlungen für den Anwender umzusetzen. Bei entsprechender Berücksichtigung werden die Sachverständigen beim Gebrauch von VKM für investigative und evaluative Untersuchungen unterstützt. Gleichzeitig sollen sie dazu beitragen, den Umgang mit VKM-Software in der Fallarbeit auch für die Bedarfsträger aus der Justiz zu erläutern. Die Empfehlungen bieten zudem Handlungssicherheit bei der Berechnung, Bewertung und Darstellung von LR_{IC} -Werten.

Korrespondenzadresse

Prof. Dr. Katja Anslinger
Institut für Rechtsmedizin, Ludwig-Maximilians-Universität
Nußbaumstr. 26, 80336 München, Deutschland
katja.anslinger@med.uni-muenchen.de

Danksagung. Wir danken Dr. Klaus Bender (Rechtsmedizin Mainz) und seinem Team für die Voruntersuchung der Blutproben sowie Eva Weissenberger (BKA) für ihre große Hilfsbereitschaft bei der Probenbearbeitung, ebenso den ehemaligen Mitgliedern der Projektgruppe für ihre wertvolle Mitarbeit. Weiterhin gilt unser Dank den Entwicklern und Firmen der VKM-Programme DNA•VIEW, EuroForMix, GenoProof Mixture und STRmix für die unverbindliche Bereitstellung von Testversionen und die vertrauensvolle Zusammenarbeit.

Funding. Open Access funding enabled and organized by Projekt DEAL.

Einhaltung ethischer Richtlinien

Interessenkonflikt. M. Templin, P. Zimmermann, S. Kranz, M. Eckert, C. Leuker, S. Razbin, A. Wächter, V. Weirich, K. Anslinger, R. Fimmers, S. Grethe, C. Oppelt, M. Vennemann, P.M. Schneider und M. Hahn geben an, dass kein Interessenkonflikt besteht.

Für diesen Beitrag wurden von den Autor/-innen keine Studien an Menschen oder Tieren durchgeführt. Für die aufgeführten Studien gelten die jeweils dort angegebenen ethischen Richtlinien.

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>.

Literatur

- Alladio E, Omedei M, Cisana S, D'Amico G, Caneparo D, Vincenti M, Garofano P (2018) DNA mixtures interpretation—A proof-of-concept multi-software comparison highlighting different probabilistic methods' performances on challenging samples. *Forensic Sci Int Genet* 37:143–150
- Bleka Ø, Storvik G, Gill P (2016) EuroForMix: An open source software based on a continuous model to evaluate STR DNA profiles from a mixture of contributors with artefacts. *Forensic Sci Int Genet* 21:35–44
- Bright JA, Taylor D, McGovern CE, Cooper S, Russell L, Abarno D, Buckleton J (2016) Developmental validation of STRmix™, expert software for the interpretation of forensic DNA profiles. *Forensic Sci Int Genet* 23:226–239
- Buckleton JS, Pugh SN, Bright JA, Taylor DA, Curran JM, Kruijver M, Gill P, Budowle B, Cheng K (2020) Are low LR's reliable? *Forensic Sci Int Genet* 49:102350
- Cheng K, Bleka Ø, Gill P, Curran J, Bright JA, Taylor D, Buckleton J (2021) A comparison of likelihood ratios obtained from EuroForMix and STRmix™. *J Forensic Sci* 66(6):2138–2155
- Coble MD, Bright JA (2019) Probabilistic genotyping software: an overview. *Forensic Sci Int Genet* 38:219–224
- Coble MD, Buckleton J, Butler JM, Egeland T, Fimmers R, Gill P, Gusmão L, Guttman B, Krawczak M, Morling N, Parson W, Pinto N, Schneider PM, Sherry ST, Willuweit S, Prinz M (2016) DNA Commission of the International Society for Forensic Genetics: Recommendations on the validation of software programs performing biostatistical calculations for forensic genetics applications. *Forensic Sci Int Genet* 25:191–197
- Hahn M, Anslinger K, Eckert M, Fimmers R, Grethe S, Hohoff C, Kranz S, Leuker C, Oppelt C, Razbin S, Rothämel T, Schneider H, Templin M, Vennemann M, Wächter A, Weirich V, Zimmermann P, Schneider PM (2022) Gemeinsame Empfehlungen der Projektgruppe „Biostatistische DNA-Berechnungen“ und der Spurenkommission zur biostatistischen Bewertung DNA-analytischer Befunde mit vollkontinuierlichen Modellen (VKM). *Rechtsmedizin*. <https://doi.org/10.1007/s00194-022-00599-5>
- Gill P, Benschop C, Buckleton J, Bleka Ø, Taylor D (2021) A review of probabilistic genotyping systems: EuroForMix, DNASTatX and STRmix™. *Genes* 12(10):1559
- Gill P, Curran J, Elliot K (2005) A graphical simulation model of the entire DNA process associated with the analysis of short tandem repeat loci. *Nucl Acids Res* 33(2):632–643
- Gill P, Gusmão L, Haned H, Mayr WR, Morling N, Parson W, Prieto L, Prinz M, Schneider H, Schneider PM, Weir BS (2012) DNA commission of the International Society of Forensic Genetics: Recommendations on the evaluation of STR typing results that may include Drop-out and/or Drop-in using probabilistic methods. *Forensic Sci Int Genet* 6(6):679–688
- Götz F, Schönborn H, Bordorf V, Pflugbeil A-M, Labudde D (2017) GenoProof Mixture 3—New software and process to resolve complex DNA mixtures. *Forensic Sci Int Genet Suppl Ser* 6:e549–e551
- Schmidt M, Schiller R, Anslinger K, Wiegand P, Weirich V (2021) Statistefix4.0: a novel probabilistic software tool. *Forensic Sci Int Genet* 55:102570
- Schneider PM, Fimmers R, Keil W, Molsberger G, Patzelt D, Pflug W, Rothämel T, Schmitter H, Schneider H, Brinkmann B (2006) Allgemeine Empfehlungen der Spurenkommission zur Bewertung von DNA-Mischspuren. *Rechtsmedizin* 16:401–404
- Taylor D, Bright JA, Buckleton J (2013) The interpretation of single source and mixed DNA profiles. *Forensic Sci Int Genet* 7(5):516–528
- Ulbrich W, Anslinger K, Bäßler G, Eckert M, Fimmers R, Hohoff C, Kraft M, Leuker C, Molsberger G, Pich U, Razbin S, Schneider H, Templin M, Wächter A, Weirich V, Zierdt H, Schneider PM (2016) Gemeinsame Empfehlungen der Projektgruppe „Biostatistische DNA-Berechnungen“ und der Spurenkommission zur biostatistischen Bewertung von DNA analytischen Befunden. *Rechtsmedizin* 26:291–298

Use of fully continuous models for biostatistical evaluation of forensic DNA analytical findings. Experiences of the project group “Biostatistical DNA calculations”

The biostatistical evaluation of DNA profiles supports the courts in the assessment of the evidential value of a DNA stain. In practice, an increasing number of stains characterized by minimal amounts of DNA and possible allelic drop-in and drop-out as well as complex mixtures are analyzed. Often such complex DNA profiles cannot, or only to a limited extent, be evaluated biostatistically with a traditional binary approach of interpretation.

The development of fully continuous models (FCM) renders a large number of these previously incalculable DNA profiles accessible for biostatistical evaluation. Almost all available information of the DNA profile is included into this calculation. While these probabilistic genotyping methods are widely used internationally, only few experiences have become known in German-speaking countries.

In order to understand the functionality, possibilities and limitations of FCM calculations, mixed DNA stains of known composition were comparatively analyzed with four currently available software products. The evaluation examined central aspects of FCM, such as concordance of calculation results, influence of drop-in and drop-out events on the calculated fully continuous likelihood ratios (LR_{FC}) values, as well as the derivation of database-suitable DNA profiles derived from probability-based predictions (deconvolution).

The experiences gained in this study, together with other already published data, provide a basis for recommendations on the use of fully continuous probabilistic genotyping software in the biostatistical evaluation of DNA profiles.

Keywords

Probabilistic genotyping · Deconvolution · Likelihood ratio · DNA evidence · Forensic biostatistics