

Interdisziplinäre Überlegungen zu Erweiterten DNA-Analysen¹

Im Jahr 2017 und erneut im Jahr 2019 wurden in Deutschland Gesetzesänderungen vorgelegt, welche die Analyse-Möglichkeiten für DNA-Spuren in Ermittlungsverfahren erweitern sollen: Es geht in erster Linie um die Vorhersage von Haut-, Haar- und Augenfarbe, die durch die Technologie des sogenannten „Forensic DNA Phenotyping“ (FDP) möglich sein soll. Die ebenfalls diskutierte Vorhersage der sogenannten „biogeografischen Herkunft“ (auf Englisch „biogeographical ancestry“, hier: BGA) war im Gesetzesentwurf von 2019 nicht mehr vorgesehen, begründet mit dem Argument, dass dies zu Diskriminierungen führen könnte; eine Annahme, die weiter unten diskutiert wird.

Im Dezember wurde §81 StPO tatsächlich geändert, so daß Ermittelnden der Einsatz des Forensic DNA Phenotyping nunmehr erlaubt ist. Der Einsatz von BGA-Analysen ist gesetzlich derzeit (September 2020) nicht möglich. Vor der zweiten Lesung des Gesetzes im Bundestag im Oktober 2019 hatte das Bundesjustizministerium zwar um Stellungnahmen von Expertinnen und Experten gebeten, diese jedoch erst nach der Lesung publiziert.²

Dieser Beitrag diskutiert die Verlässlichkeit, Nützlichkeit und Legitimität³ von Erweiterten DNA-Technologien und damit die Rahmenbedingungen ihres Einsatzes in Deutschland. Von den Argumenten, die seit 2017 für die Gesetzesänderung angebracht wurden, seien hier drei genannt, auf die unsere Argumentation eingehen wird: (1) Die Technologien seien ausgereift und einsatzbereit; (2) es gelte, potenzielle Opfer zu schützen bzw. die Rechte der Opfer und Opfer-Angehörigen zu stärken; (3) die Technologien könnten Minderheiten von ungerechtfertigten Verdächtigungen entlasten. Unsere Position sieht die Technologien deutlich kritischer, stellt sich aber nicht prinzipiell gegen deren Anwendung. Vielmehr sollte die konstruktive Frage gestellt werden, wie und unter welchen Bedingungen ein verantwortungsbewusster Einsatz möglich und denkbar wäre. Einzelne Fallbeispiele aus dem europäischen Ausland, in denen ein möglichst verantwortungsvoller Einsatz angestrebt wurde, dienen dabei als Richtschnur.

Eines unserer Hauptargumente lautet, dass ein verantwortungsbewusster Einsatz der Methoden hohe gesellschaftliche Sensibilität erfordert. Dies vor allem aus einem Grund, den einschlägige Expertinnen und Experten schon vor Jahren benannt haben: Die Ergebnisverwertung bietet sich für die Ermittlungsarbeit vor allem dann an, wenn das Analyse-Ergebnis auf ein Individuum verweist, das einer gesellschaftlichen Minderheit angehört, deren Vorfahren von einem anderen Kontinent stammen. So schreiben die Forensischen Genetiker Manfred Kayser und Peter de Knijf (2011): Zu den regulatorischen Fragen, die man bedenken müsse, gehöre „non-discrimination (particularly salient as FDP is most useful for tracing suspects from minority

¹ Dieser Beitrag beruht auf Ergebnissen und Diskussionsbeiträgen der Wissenschaftlichen Initiative „WIE-DNA“ (siehe www.WIE-DNA.de). Im Wesentlichen bezieht er sich auf den Referentenentwurf des Bundesministeriums für Justiz und Verbraucherschutz (BMJV), „Entwurf eines Gesetzes zur Modernisierung des Strafverfahrens“ vom 08.08.2019. Auf unserer Website (www.WIE-DNA.de [11.09.2020]) stehen Links zu den Gesetzesentwürfen, dem ursprünglichen und dem geänderten Gesetzestext, unsere Stellungnahmen (z.B. Lipphardt et al. 2017) und Öffentlichkeitsarbeit, umfangreiche Literaturhinweise sowie weitere Materialien und Informationen zur Verfügung.

² Vgl.

https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/DE/Modernisierung_Strafverfahren.html [11.09.2020]; darunter die Stellungnahme unserer Initiative, auch abrufbar unter <http://www.wie-dna.de/statement-stpo/> [11.09.2020].

³ Vgl. Wienroth 2020b.

groups)“.⁴ Weshalb und inwiefern sich die Methoden in einem solchen Fall besonders gut eignen und welche Folgerungen sich daraus für den gesellschaftlichen Zusammenhalt ergeben, wird dieser Beitrag erläutern. Dabei werden wir mit den wissenschaftlich-technischen Details beginnen, mit der Einsatzlogistik fortfahren und schließlich gesellschaftliche Aspekte aufzeigen.⁵

Anschließend widmen wir uns einem zweiten Argument: Schlüsselvoraussetzung für einen sensiblen Umgang mit diesen Technologien ist eine verantwortungsvolle, gut durchdachte Kommunikation aller Beteiligten mit und in der Öffentlichkeit, mit Anwenderinnen und Anwendern und mit den voraussichtlich am stärksten betroffenen gesellschaftlichen Gruppen. Alle Beteiligten sollten eine realistische, transparente und möglichst situationsbezogene Darstellung der Potenziale, Grenzen und Risiken der Technologien anstreben, anstatt Einsatzmöglichkeiten abstrakt und idealisiert darzustellen. Dazu gehört auch eine realistische Einschätzung der tatsächlich erreichbaren Anwendungsqualität in einem konkreten, realen Kontext.

Wenn eine solche Kommunikation ausbleibt, ergeben sich falsche oder übertriebene Vorstellungen davon, wofür die Technologie sinnvoll verwendet werden kann; Vorstellungen, die die konkrete Technologie-Anwendung prägen werden. Eine solche Kommunikation sollte prinzipiell nicht erst mit den ersten Einsätzen beginnen, sondern bereits in dem Moment, in dem eine Gesetzesänderung angestrebt wird. (Für die BGA-Analysen ist es also noch nicht zu spät).

Dieser Beitrag wird im dritten Kapitel im Rückblick auf das Jahr 2017 diskutieren, inwiefern die damalige öffentliche Kommunikation das Verständnis der Einsatzmöglichkeiten bis heute geprägt hat. Abschließend wird der Beitrag konstruktive Vorschläge für einen verantwortungsvollen Einsatz der Technologien machen.

Als Quellen verwenden wir Gesetzesentwürfe mitsamt den sie begleitenden Begründungstexten, öffentlich zugängliche Aussagen von Befürworterinnen und Befürwortern, wissenschaftliche Publikationen und Tagungsdokumentationen, amtliche Dokumente sowie Medienberichte. Methodisch basiert dieser Beitrag auf diskurs- und dokumentenanalytischen Arbeitsweisen sowie Methoden der Wissenschaftsforschung, vor allem aber auf einer Vorgehensweise, die die interdisziplinäre Verständigung in den Mittelpunkt stellt: Um über realistische Einsatzmöglichkeiten zu diskutieren, bedarf es einer interdisziplinären Verständigung über sämtliche Rahmenbedingungen eines solchen Technologie-Einsatzes.

Aus unserer Sicht sind auch noch jetzt, nach der Änderung des §81 StPO, zahlreiche offene Fragen zu klären, wie etwa die folgenden:

- In welchen Fallkonstellationen ist ein Einsatz der Erweiterten DNA-Analysen sinnvoll, in welchen nicht? Welche Anforderungen ergeben sich daraus? Welche Sachkompetenz ist dafür angemessen?
- Welche Fehlerquellen gilt es zu berücksichtigen?
- Wer wird die DNA-Analysen durchführen, interpretieren, anwenden; wie werden die

⁴ Kayser/Knijff (2011) zählen zudem weitere ethische Risiken auf: „[...] privacy and data protection, the ‘right not to know’, and preventing ‘slippery slopes’. These are serious issues, but their particular relevance to FDP should not be overestimated“ (Kayser/Knijff 2011, 183). Die Autoren verweisen auf Ossorio 2006.

⁵ Wir stützen uns auf hauptsächlich englischsprachige Literatur, die ansonsten in Deutschland wenig rezipiert wurde, darunter: M’charek 2008; Toom et al 2016; Scudder / McNevin 2018; Wienroth 2018, 2020b; Williams / Wienroth 2017; M’charek / Toom / de Jong 2020.

Anwenderinnen und Anwender geschult und auf ihre anspruchsvolle Aufgabe vorbereitet?

- Welche unabhängige Institution wird in Zukunft die wichtige Aufgabe der Kontrolle und Regulierung polizeilicher DNA-Datenerhebung und -nutzung übernehmen?
- Wie steht es um die molekulargenetischen, statistischen und gesellschaftswissenschaftlichen Hintergrundkenntnisse von Ermittlerinnen und Ermittlern, Richterinnen und Richtern, Staatsanwältinnen und Staatsanwälten sowie Verteidigerinnen und Verteidigern?
- Welche „checks and balances“, welche Schulungen gibt es bereits und welche neuen müssten aufgrund der Gesetzesänderung eingeführt werden, um etwa sicherzustellen, dass Fehlerquellen berücksichtigt wurden oder dass Voreingenommenheit gegenüber Minderheiten nicht zu einseitigen Interpretationen und vorschnellen Festlegungen führt?
- Welche Kosten entstehen durch die Sachverständigen-Gutachten, die mit der gebotenen Sachkompetenz zu erstellen sind?
- Welche Aufwendungen entstehen jenseits der Kosten für die technische Ausstattung – etwa für Schulungen und für die notwendigen Regulierungsinstitutionen?
- Wie können Ermittlungsmaßnahmen, die auf Erweiterte DNA-Analysen folgen – insbesondere DNA-Reihenuntersuchungen und Öffentlichkeitsfahndungen –, so gestaltet werden, dass dadurch Minderheiten nicht unter Generalverdacht geraten oder stigmatisiert werden?
- Welche Bedenken haben in Deutschland lebende Menschen bezüglich der Speicherung ihrer DNA-Daten (in Forschungsdatenbanken, im Gesundheitswesen, im Polizeiwesen); und wie kann diesen Bedenken im Rahmen rechtlicher Regulierungen Rechnung getragen werden?
- Wie kann einem möglichen Missbrauch der DNA-Daten(banken) vorgebeugt werden?

1. Validität: Wissenschaftliche Grenzen der Technologien

Zunächst ist zu klären, wie einsatzbereit, wie exakt, wie verlässlich die beiden Technologien derzeit tatsächlich sind. Expertinnen und Experten aus dem Bereich der Forensischen Genetik sind sich in diesem Punkt nicht unbedingt einig.⁶ Die pauschal hohen Wahrscheinlichkeitsangaben, mit denen die Gesetzesentwürfe des Jahres 2017 aufwarteten, und die sich immer noch in den Referenzen des Gesetzesentwurfs von 2019 finden, sind nach unserem Wissensstand nicht haltbar. Dies sei im Folgenden für FDP und BGA dargestellt, weil ein Verständnis der Fehlerquellen eine Voraussetzung für einen gesellschaftlich sensiblen Umgang mit den Technologien ist.

1.1 Forensic DNA Phenotyping (FDP)

Die hohen Wahrscheinlichkeitswerte, die in der Öffentlichkeit und unter BefürworterInnen zirkulieren, sind keine Vorhersagegenauigkeiten. Auch der Gesetzesentwurf von 2019 irrt in diesem wichtigen Punkt.⁷ Vielmehr handelt es sich im Falle des FDP um Angaben zur sogenannten „Area under the Curve“, ein Wert, der die Performance einer Methode unter Laborbedingungen beschreibt. Die für die Ermittlerinnen und Ermittler interessanten Vorhersagegenauigkeiten lassen sich am ehesten mit den

⁶ Ausführlich zu den divergierenden Ansichten siehe Lipphardt 2018. Der Text diskutiert außerdem die Beteiligung Forensischer Genetikerinnen und Genetiker am öffentlichen Diskurs sowie unterschiedliche Vorstellungen von verantwortungsvollem Handeln in der Politikberatung.

⁷ „Die genannten äußerlich sichtbaren Körpermerkmale können nach aktuellen wissenschaftlichen Erkenntnissen durch Untersuchungen genetischer Informationen mit hinreichender Vorhersagegenauigkeit bestimmt werden“; Referentenentwurf 2019, 27. Der Gesetzesentwurf verweist für diese Information auf eine Stellungnahme der Spurenkommission von 2016, deren Informationen aber inzwischen vielfach relativiert wurden.

sogenannten „positive predictive values“ und „negative predictive values“ angeben. Für zwei Merkmalsausprägungen der Augenfarbe liegen diese Werte öffentlich vor: für blaue und braune Augen. Das Forscherteam um Amke Caliebe hat dafür Menschen aus acht verschiedenen europäischen Ländern untersucht. Für blaue Augen sind die positive predictive values mit 84 %-94 % in allen Ländern relativ hoch.⁸ Für braune Augen ist der positive predictive value allerdings nicht in allen Ländern gleich hoch: in Norwegen erreichte die Bestimmung der braunen Augenfarbe lediglich eine positive predictive value von 65 %, in Großbritannien 67 %, in Holland 68 %, in Estland 69 %, in Italien und Frankreich mittlere 80er Werte und nur in Griechenland und Spanien über 90 %. In Ländern, in denen braune Augen seltener sind, ist die Vorhersage also ungenauer.

Für dunkle bzw. helle Haar- und Hautfarbe dürften die positive predictive values für die meisten Länder deutlich niedriger ausfallen, denn deren Vererbung ist komplexer als die der Augenfarbe. Leider werden diese Werte in der einschlägigen Literatur nicht publiziert, obwohl sie aus den vorhandenen Daten relativ einfach abzuleiten wären. Sämtliche Mischfarben (z.B. grün, grau, hellbraun) werden ohnehin, da sind sich alle Expertinnen und Experten einig, mit deutlich geringerer Genauigkeit vorhergesagt. Das einzige Merkmal, das sich also in verschiedenen europäischen Ländern mit ziemlich hoher Wahrscheinlichkeit präzise voraussagen lässt, ist blaue Augenfarbe.

Die Autorinnen und Autoren einer darauf bezugnehmenden Publikation hielten entsprechend 2017 fest, dass die Technologien sich noch nicht für den Einsatz in „criminal casework“ eignen:

„[...] since the etiological understanding of FDP-relevant appearance phenotypes is still incomplete, so are the prediction models used – and the corresponding predictive values vary to a certain degree. In consequence, these measures have to be determined empirically for each appearance phenotype, prediction model and target population of interest before they can be applied sensibly in criminal casework.“⁹

Wie komplex diese Aufgabe ist, zeigt ein Blick in den Bereich der wissenschaftlichen Vorarbeit für diese Technologien. Die DNA-Daten, an denen die Methoden im Labor entwickelt, trainiert und getestet werden, sind nicht nach dem Zufallsprinzip erhoben, sondern mit einer Samplingstrategie: DNA-Daten einer Person werden nur dann in ein Datenset aufgenommen, wenn ihre Großeltern etwa aus einem bestimmten Land oder einer bestimmten Region stammen. Die Daten für Norwegen sollten also von Angehörigen alteingesessener norwegischer Familien kommen, die in den letzten Generationen keine Migrationserfahrung über die heutigen Landesgrenzen hinweg gemacht haben, und in die keine Person mit nichtnorwegischen Vorfahren eingewandert hat. Dasselbe gilt jeweils für die Daten aus Griechenland, Italien, Großbritannien usw. Diese Bedingungen erfüllt nur ein Teil aller europäischen Familien. Die wechselhafte Geschichte vieler europäischer Länder hat durch Kolonialismus, Umsiedelungen, Industrialisierung, Nahrungsknappheit, Nationalstaatenbildung und Kriege ein reiches Spektrum zwischen erzwungener und freiwilliger Mobilität und Fortpflanzung hervorgebracht.

Deshalb machen Familien, deren Vorfahren alle aus einem einzigen Land oder einer einzigen Bevölkerungsgruppe kommen, nur einen gewissen Teil der Gesamtbevölkerung aus, in manchen Ländern oder Gegenden mehr und in manchen weniger. Und

⁸ Caliebe et al. 2017, 207.

⁹ Caliebe/Krawczak/Kayser 2018, e8; siehe auch folgende Reaktion: Buchanan et al. 2017.

dennoch beschreiben die genannten Werte lediglich Vorhersagegenauigkeiten für genau diesen Teil der Bevölkerung. Für komplexer zusammengesetzte Gesellschaften – z. B. städtische oder postkoloniale – müssten komplizierte Berechnungen angestrengt werden, damit die Referenzdaten und damit auch die Testergebnisse die gesellschaftliche Realität annähernd gut abbilden. Erst dann kann man Vorhersagegenauigkeiten für die reale Bevölkerung eines Ortes (etwa eines Landes, oder eines Tatortes) angeben.

1.2 Biogeografische Herkunft (BGA)

Obwohl das im Dezember 2019 geänderte Gesetz die Bestimmung der BGA nicht vorsieht, soll sie hier diskutiert werden. Denn erstens werden sich die Befürworterinnen und Befürworter weiterhin für die Zulassung der BGA einsetzen, auch wenn viele unter ihnen deren Risiken nicht angemessen zur Kenntnis nehmen. Zweitens lassen sich die Vorhersage der BGA und die der Haar-Haut- und Augenfarben technisch nicht voneinander trennen: wichtige SNP¹⁰-Marker werden in beiden Technologien verwendet. Drittens kann die BGA mit der geforderten technischen Ausrüstung und mit verfügbaren Open-Access-Tools mühelos mitbestimmt werden. Viertens halten viele Forensische Genetikerinnen und Genetiker in Deutschland, die selbst mit den Methoden arbeiten, die BGA-Information für aussagekräftiger als die FDP-Information. Fünftens stellt die Ermittlung der BGA auch eine technisch-wissenschaftliche Herausforderung dar und nicht nur eine gesellschaftliche. Sechstens stellen wir uns nicht prinzipiell gegen die Anwendung von BGA-Analysen, denn wir kennen einzelne Beispiele, in denen damit umsichtig genug umgegangen wurde; wir stellen uns vielmehr gegen eine Anwendung, die die Herausforderungen ignoriert und damit negative Folgen für Betroffene in Kauf nimmt. Das gilt übrigens auch für die FDP-Technologien, die wir – im Gegensatz zum BMJV – bezüglich möglicher Diskriminierungsfolgen nicht als weniger riskant betrachten als die BGA-Analysen: Das Ergebnis „dunkle Hautfarbe“ kann ebenso zu Diskriminierung führen und rassistische Hetze nach sich ziehen wie das einer „afrikanischen BGA“.

Alle Gesetzesentwürfe beziehen sich hinsichtlich der Vorhersagegenauigkeiten auf die bereits erwähnte Stellungnahme der Spurenkommission aus dem Dezember 2016. In diesem Dokument tauchte zum ersten Mal die Angabe von 99,9 % Vorhersagegenauigkeit für kontinentale BGA auf, ohne dass eine wissenschaftliche Publikation als Referenz für diesen Wert angegeben wurde (es gibt allerdings auch keine wissenschaftliche Publikation, in der sich dieser Wert als pauschal erreichbarer Wert genannt wird). Der Vorsitzende der Spurenkommission, Peter Schneider, stets um Differenzierung und eine realitätsnahe Darstellung der Technologien bemüht, stellte später mehrfach öffentlich klar, dass dieser hohe Wert längst nicht in allen Fällen erreicht werden könne. Dennoch machte die Zahl schnell die Runde und hält sich in der politischen Diskussion hartnäckig. Der Verzicht auf diese Technologie wurde vom BMJV öffentlich und im Gesetzesentwurf von 2019 ausschließlich gesellschaftspolitisch begründet: um Diskriminierungen und rassistische Hetze zu vermeiden. Dass es auch wissenschaftlich-technische Fehlerquellen geben könnte, bleibt unerwähnt.

¹⁰ SNP=Single Nucleotide Polymorphisms; diejenigen Stellen im Genom, die nicht bei allen Menschen identisch sind, sondern in mehreren Varianten vorkommen. Manche Varianten kommen in manchen Regionen häufiger vor als andere; manche Varianten sind insgesamt sehr viel seltener als andere. Diese SNP Marker dienen der Analyse der sogenannten BGA.

Unter welchen Umständen, so fragen wir im Folgenden, könnte ein Wert von 99,9 % erreicht werden? Wiederum gilt es, die Vorarbeit in den Blick zu nehmen, die nötig ist, bevor diese Technologie überhaupt entwickelt werden kann: insbesondere die Referenzdatenbanken, die für BGA-Ermittlungen eine wichtige Rolle spielen. Je nachdem, wo die Daten gesammelt wurden und mit welcher Sampling-Strategie, bilden sie bestimmte Bevölkerungsteile ab, andere aber nicht. Eine der wichtigsten Voraussetzungen für eine sehr hohe Vorhersagegenauigkeit lautet: Nur für eine Person, deren Vorfahren alle aus nur einer der Bevölkerungsgruppen kommen, die in der Referenzdatenbank erfasst sind, können die Technologien sehr hohe Werte erreichen.

BGA-Bestimmungen kann man an drei verschiedenen DNA-Systemen durchführen: An den autosomalen Chromosomen, die jeder Mensch von Vater und Mutter erbt und den uniparentalen Systemen, d. h. dem Y-Chromosom, das nur von Vätern an männliche Nachkommen weitervererbt wird sowie der mtDNA, die nur von Müttern an Nachkommen beiderlei Geschlechts vererbt wird. Jede der beiden letztgenannten Informationen gilt also nur für die männliche oder weibliche Vorfahren-Linie *eines* Elternteils. Hat man beide Informationen – was nur bei Männern möglich ist – hat man schon zwei Anhaltspunkte, aber eben auch nicht mehr, und vor allem keine 99,9-prozentige Gewissheit.

1.2.1 Uniparentale BGA-Bestimmungen

Uniparentale BGA-Bestimmungen bieten den Vorteil, dass sie in einer begrenzten Anzahl von Fällen eine recht präzise geografische Information über die Herkunft der Vorfahren einer gesuchten Person liefern können: nämlich dann, wenn eine gesuchte uniparentale DNA-Variante – ein Haplotyp – ganz überwiegend in einem eng umgrenzten Gebiet oder in einer klar abgrenzbaren Gruppe gefunden wurde, aber anderswo oder in anderen Populationen so gut wie nie. Es bedarf also umfassender Vorkenntnisse. Bestenfalls befinden sich in einer entsprechenden Referenzdatenbank DNA-Proben von mehreren Personen mit genau dieser Variante, die alle mit denselben klar umrissenen Herkunftsangaben gespeichert wurden. Dann wird die Probe der gesuchten Person mit sehr großer Sicherheit ebendieser Gruppe zugeordnet. Lebt die gesuchte Person etwa in einer abgeschiedenen Bergregion, die nie vom Tourismus erfasst, aber dafür von Genetikerinnen und Genetiker zum Sampling aufgesucht wurde, kann die Zuordnung sehr gut gelingen. Bei weitverbreiteten Varianten ergibt sich demgegenüber kein weiterführender Hinweis.

Ein weiterer Vorteil der uniparentalen BGA-Bestimmung ist, dass sie aufwandsarm, niedrigschwellig und kostenlos in den beiden Open-Access-Online-Referenzdatenbanken YHRD und EMPOP durchgeführt werden kann. Beide Datenbanken bilden zahlreiche Populationen auf der subkontinentalen Ebene ab, auch solche, die in den Datenbanken für autosomale BGA-Bestimmungen fehlen (siehe unten). Diese Option dürfte den Genetikerinnen und Genetiker an den Landeskriminalämtern (LKÄ) bestens bekannt sein. Sie wird auch von universitären Forensischen Genetikerinnen und Genetikern in den rechtsmedizinischen Instituten in Anspruch genommen, nicht zuletzt, wenn es darum geht, eine unbekannt Leiche zuzuordnen.

Auf Y-chromosomale BGA-Informationen stützen sich eine Handvoll vielzitiert Fälle, in denen BGA-Technologien nach Ansicht von Ermittlerinnen und Ermittlern erfolgreich angewandt wurden: nämlich im „Fall Eva Blanco“, im „Fall Marianne Vaatstra“, im „Fall des Attentat-Bombers von Madrid“ und im „Fall Minka“. Jeweils verwies die erzielte BGA-Information auf eine sehr spezifische Region: nord-

afrikanisch, nordeuropäisch, türkisch. In jedem dieser Fälle stellte sich die Information nach der Erfassung des Täters als korrekt heraus. So entstand der Eindruck, die BGA-Analysen könnten insgesamt viel präzisere Informationen liefern als „nur“ die 99,9 % sichere Zuordnung zu einem Kontinent: Nämlich die sichere Zuordnung zu kleinräumigen Regionen auf der subkontinentalen Ebene. Dabei ist öffentlich nicht bekannt, in wie vielen Fällen die Technologien insgesamt angewandt wurde.

Allerdings bringen die uniparentalen BGA-Bestimmungen auch spezielle, gravierende Probleme mit sich, wenn sie als einzige Informationsquelle herangezogen werden. Die besonders präzise Information einer nur sehr begrenzt verbreiteten uniparentalen DNA kann nämlich falsch sein und in die Irre führen. Das ist der Fall, wenn ein Haplotyp *außerhalb* der gut abgegrenzten Bevölkerungsgruppe auftaucht, für die er eigentlich typisch sein sollte. Abgeschlossene Gesellschaften bringen es nun einmal mit sich, dass es „Abtrünnige“ gibt, d. h. Menschen, die in der Ferne (oder in einem anderen Stadtviertel) ihr Glück suchen; oder Verstoßene, Entführte und Versklavte – die Geschichte ist reich an Beispielen solcher freiwilligen oder erzwungenen Ortswechsel. Selbst wenn es sich bei der Ausgewanderten bzw. dem Ausgewanderten um einen Vorfahren von vor 400 Jahren handelte, kann ein heute lebender Mensch deshalb einen Haplotyp haben, der untypisch für den eigenen Herkunftsort oder die eigene Sozialisation ist. Oft sind diese Familiengeschichten nicht einmal bekannt – nicht jeder bzw. jede weiß, woher die Familien seiner bzw. ihrer (biologischen!) Urgroßeltern kamen. Manche dieser „gewanderten“ Haplotypen tauchen in den Referenzdatenbanken auf, und ihre Diskrepanz zu den Herkunftsangaben kann manchmal historisch erklärt werden; aber wenn sich das Sampling ausschließlich auf Menschen konzentriert, bei denen die Wahrscheinlichkeit dafür geringer ist – weil alle vier Großeltern aus derselben Region kommen – wird nicht die Haplotypenverteilung der Normalbevölkerung dargestellt, sondern nur die eines Ausschnittes. Es wird in solchen Datenbanken dennoch zahlreiche Fälle geben, in denen die subjektiv bekannte Herkunft der Vorfahren von der erzielten BGA-Information abweicht – und außerdem viele Haplotypen, die viel zu weit verbreitet sind, um den Ermittlerinnen und Ermittlern aufschlußreiche Informationen zu liefern.

In den oben genannten Fällen – eine Handvoll veröffentlichter Erfolgsfälle aus ungefähr 20 Jahren Anwendung uniparentaler BGA-Bestimmung – hatten die Ermittlerinnen und Ermittler doppelt Glück: die Täter waren noch in der Bevölkerungsgruppe sozialisiert, die in der Referenzdatenbank mit einer entsprechenden Herkunftsbezeichnung hinterlegt war, und konnten bei einer Anschlußmaßnahme entdeckt werden.

Denkbar sind aber auch Fälle, in denen die BGA-Information zwar auf die Gruppe verweist, in der die Täterin bzw. der Täter sozialisiert ist, dieser bzw. dieser jedoch bei einer Anschlußmaßnahme trotzdem nicht entdeckt werden kann, weil sie bzw. er z. B. nicht vor Ort wohnt. Denkbar sind auch Fälle, in denen die BGA-Information irreführend war (oder die Gruppenbezeichnung in der BGA-Datenbank sich nicht mit einer sozialen Kategorie vor Ort deckte) und die Ermittlerinnen und Ermittler die Täterin bzw. den Täter in der falschen Bevölkerungsgruppe suchten, weil sie oder er (oder ihre bzw. seine ganze Familie) gesellschaftlich einer anderen Gruppe zugehörte. Diese beiden letzteren Fallgruppen lassen sich kaum auseinanderhalten, solange eine Täterin bzw. ein Täter nicht gefasst wird. Vielleicht ist beides bereits in Ermittlungsfällen vorgekommen, aber die Öffentlichkeit hat nichts davon erfahren, weil nur Erfolgsfälle übermittelt wurden.

Beides würde bedeuten, dass eine Bevölkerungsgruppe ungerechtfertigterweise

einem erhöhten Ermittlungsdruck ausgesetzt wird, ohne dass dies zum Aufklärungserfolg beiträgt. Sollte sich im Nachhinein herausstellen, dass eine Person in der BGA-Bestimmung einer Herkunftsregion zugeordnet wurde, die mit ihrer eigenen Biographie nichts zu tun hat, sondern mit der eines Vorfahren, können historisch interessierte Genetikerinnen und Genetiker das Rätsel in manchen Fällen vielleicht lösen und herausfinden, wie es zu diesem untypischen Haplotypen-Auftreten gekommen sein mag. Für die Ermittlerinnen und Ermittler, die sich vorwiegend auf uniparentale BGA Informationen stützen, stellt diese Irrtumsquelle ein handfestes Problem dar.

1.2.2 Autosomale BGA-Bestimmungen

Autosomale BGA-Bestimmungen haben im Vergleich zu den uniparentalen den Vorteil, dass sie die BGA vieler Vorfahren auf einmal abbilden – allerdings um den Preis der Eindeutigkeit. Nur, wenn beide Elternteile ausschließlich Vorfahren aus derselben Region haben und diese Bevölkerung in der Referenzdatenbank gut abgebildet ist, läßt sich ein sehr sicheres Ergebnis erzielen. Deshalb sprechen Expertinnen und Experten diesen Methoden nur für die kontinentale BGA eine gewisse Aussagekraft zu.¹¹

Die Vorhersagegenauigkeit ist abhängig davon, auf welchen Referenzdaten das Marker-Set basiert und nach welchen Kriterien das Marker-Set ausgesucht, trainiert und getestet wird. So schließen viele der gängigen Methoden von vornherein Populationen, die keinem einzelnen Kontinent zugeordnet werden können und daher als „gemischt“¹² bezeichnet werden, aus den Referenzdaten aus (z. B. Menschen aus dem Mittleren Osten, Nordafrika, Südamerika). Entsprechend wird man die BGA von Menschen aus diesen Regionen auch nicht sicher vorhersagen können. Aber nur durch diesen Ausschluß bestimmter Populationen lassen sich hohe Vorhersagegenauigkeiten erzielen. Man darf dabei nie aus den Augen verlieren, dass die sehr hohen Werte nur für Fälle gelten können, in denen die Vorfahren einer Person alle demselben Kontinent entstammen und deren Herkunftspopulation in der Referenzdatenbank gut erfasst ist.¹³

Hat die Person aber einen „gemischten“ Hintergrund, benötigt man für die Interpretation eines BGA-Ergebnisses Fingerspitzengefühl, viel Erfahrung und eine klare Vorstellung von den Grenzen der verwendeten Methode. So geben nicht alle Methoden einen verlässlichen Hinweis darauf, ob ihr Ergebnis wirklich eindeutig ist; und wenn doch, liefert eine Methode nicht unbedingt einen brauchbaren Hinweis darauf, was ein uneindeutiges Ergebnis bedeuten könnte. Es könnte zum Beispiel sein, dass die Eltern einer Person von verschiedenen Kontinenten stammen; oder aber eine Person kommt aus einer zwischen zwei Kontinenten gelegenen Region. Aber welche der beiden „Mischungsformen“ vorliegt – individuell in der letzten Generation, oder in großen Bevölkerungen und schon seit langer Zeit –, kann in manchen Fällen nur durch zusätzliche Methoden herausgefunden werden.

Eine Methode, die Asiatinnen und Asiaten von Europäerinnen und Europäern bzw.

¹¹ Siehe hierzu Cheung 2016. Es gibt auch Ansätze dafür, autosomale Methoden für subkontinentale Unterscheidungen zu entwickeln, was wiederum andere zu bedenkende Fehlerquellen mit sich bringt.

¹² „Gemischt“ ist ein artifizierlicher Ausdruck: Erst, wenn man zwei Populationen als prinzipiell voneinander unterscheidbar bezeichnet hat – z. B. Asien und Europa –, erscheinen Menschen, die in Grenzregionen zwischen Europa und Asien leben, als „Mischung“ der zuvor definierten beiden Populationen. Tatsächlich müsste es sich, wenn man von Ortsansässigkeit ausgeht, um ein Kontinuum handeln; oder, wenn man kontingente Migrationsgeschichten miteinbezieht, von einem mehrfach und komplex durcheinandergebrachten Kontinuum.

¹³ Vgl. Pfaffelhuber et al 2019.

Afrikanerinnen und Afrikanern unterscheiden soll und dies anhand einer Referenzdatenbank, in der es nur Personen gibt, deren Vorfahren entweder alle europäischen oder alle asiatischen oder alle afrikanischen Ursprungs sind, wird also in vielen Fällen Menschen mit je einem Elternteil aus Asien bzw. Afrika und einem aus Europa ebenso falsch klassifizieren (oder nicht sicher klassifizieren können) wie Menschen aus Regionen zwischen den Kontinenten: etwa dem Mittleren Osten oder Nordafrika. Beide Regionen sind in Referenzdatenbanken für autosomale DNA-Daten kaum abgebildet. Methoden, die etwa Mischungsverhältnisse in Prozenten angeben („admixture“), sind in solchen Fällen meist wenig hilfreich, vor allem wenn es sich um komplexere oder länger zurückliegende „Mischungen“ handelt.

Um solche uneindeutigen Fälle zu verstehen und mit der gebotenen Vorsicht zu interpretieren, benötigt man einschlägige Expertise und viel Erfahrung, am besten in einem Land, in dem ungewöhnliche BGAs an der Tagesordnung sind.

Aus all diesen Gründen hält etwa die britische Genetikerin Denise Syndercomb-Court BGA-Bestimmungen nur dann für ausreichend verlässlich, nützlich und weiterverwendbar, wenn sie für alle drei DNA-Systeme dieselbe eindeutige BGA Information erzielt haben.¹⁴

2. Nützlichkeit und Sensibilität: Potenzielle Probleme und Strategien zu ihrer Vermeidung

Aus zahlreichen Gesprächen mit Genetikerinnen und Genetikern im Ausland haben wir gelernt, dass die Technologien nur in wenigen, speziell gelagerten Fällen, manchmal erst nach langwierigen erfolglosen Ermittlungen, sinnvoll und weiterführend zum Einsatz kommen können. Dabei hängt es wesentlich von der gesellschaftlichen Sensibilität der beteiligten Genetikerinnen und Genetikern, Ermittlerinnen und Ermittlern sowie von den gewählten Anschlußmaßnahmen ab, ob ein diskriminierender oder stigmatisierender Effekt vermieden wird. Generell befürworten wir daher einen umsichtigen, mehrstufigen Beratungsprozess und halten es für sinnvoll, die Technologien nur dann einzusetzen, wenn sie eine konkrete Frage möglichst klar zu beantworten helfen. Für einen flächendeckenden, *first-line*, routinemäßigen Einsatz taugen sie unseres Erachtens nicht, wie wir im Folgenden erläutern.

2.1 Diskriminierungsgefahr

Fassen wir zusammen, in welchen Fällen die Technologien sichere Ergebnisse erzielen können und in welchen nicht: Blaue Augen lassen sich in vielen europäischen Ländern mit recht hoher Genauigkeit vorhersagen, grüne, graue oder grüngraubraune Augen aber nicht. Dunkle Augen lassen sich in manchen Ländern mit hoher Genauigkeit vorhersagen, in anderen nicht. Für alle anderen Merkmale haben wir nur Anhaltspunkte, weil keine *positive/negative predictive values* vorliegen.

Wir gehen im Folgenden nicht ganz realistisch davon aus, dass dunkle Haare und dunkle Haut gut vorhersagbar wären und ebenso helle Haut und helle Haare. Nehmen wir außerdem an, BGA-Informationen könnten klar zwischen Menschen mit ausschließlich europäischen und ausschließlich afrikanischen oder asiatischen Vorfahren unterscheiden. Dann kann man die nächste Frage stellen: Welche dieser

¹⁴ Mündliche Information, Interview vom 26.2.2018.

Merkmale und welche dieser Merkmalskombinationen sind nützlich für Ermittlungen? Wann erlaubt ein Analyse-Ergebnis eine Fokussierung?

„Gemischte“ Werte sind in allen Methoden unsichere Werte – sie sind also von vornherein nicht verlässlich. Je nachdem, wie viele Menschen in der Nähe eines Tatortes „gemischte“ Werte aufweisen – etwa grüne Augen, mittelbraune Haare, heller bis leicht gebräunter Teint, oder Vorfahren von verschiedenen Kontinenten – wäre der von vornherein erwartbare Nutzen der Anwendung bei einem entsprechenden Ergebnis also eher gering. Am höchsten wäre er, wenn diese Gruppe sehr klein wäre und die Bevölkerung nur aus zwei gut unterscheidbaren Gruppen bestünde: sehr hell pigmentierten und sehr dunkel pigmentierten Menschen, oder einerseits aus Menschen mit Vorfahren nur aus Afrika (bzw. Asien) und andererseits aus Menschen mit Vorfahren nur aus Europa: Dann wäre das Unterscheidungspotenzial besonders groß. Aber wäre das Ergebnis in einer solchen Situation auch in jedem Fall nützlich?

In Bayern wurde kürzlich eine BGA-Bestimmung für eine Serie von Sexualstraftaten durchgeführt und das Ergebnis lautete: Europäer.¹⁵ Die Ermittlerinnen und Ermittler ließen diese Spur fallen, weil sich daraus kein sinnvoller Ermittlungsansatz ergab: Da in Bayern zu viele Europäerinnen und Europäer leben, war die Eingrenzung auf eine zu fokussierende Gruppe schlicht nicht möglich. Dasselbe Problem würde sich bei einer FDP-Analyse für die Ergebnisse „Haare hell, Haut hell, Augen blau“ (etwa in Schleswig-Holstein) und auch für „Haare braun, Haut hell, Augen braun“ (etwa in Süddeutschland) ergeben. Ein solches FDP-Ergebnis, auch wenn ergänzt durch die sichere BGA-Information „Europäer“, kann nicht zu einer Fokussierung führen. Und wie sollte eine solche auch umzusetzen sein, wenn entsprechende Kategorien nicht im Pass verzeichnet sind? Lediglich Haar- und Augenfarbe sind den Meldebehörden bekannt, also könnte man diese höchstens für die Wohnbevölkerung in eine Anschlussmaßnahme übersetzen, d. h. man könnte etwa alle mit der gesuchten Merkmalskombination anschreiben und zu einer freiwilligen DNA-Reihenuntersuchung auffordern. Wir bezweifeln allerdings, dass Ermittlerinnen und Ermittler dies für einen sinnvollen Ansatz halten würden.

Bleibt die Beobachtung und Überprüfung derer, die sich in der Nähe des Tatortes aufhalten oder dort wohnen: Eindeutig helle oder eindeutig dunkle Hautfarbe wäre hierbei zwar ein Anhaltspunkt, aber zumindest für helle Hautfarbe würden Ermittlerinnen und Ermittler in Deutschland dies wahrscheinlich nicht für zielführend halten.

Aus diesen Überlegungen ergibt sich, dass die Analysen nur dann einen brauchbaren Ansatz liefern, wenn sie auf vor Ort relativ seltene Merkmale hinweisen: In Deutschland z. B. dunkle Haut, dunkle Haare und Augen, oder ausschließlich asiatische oder afrikanische Vorfahren. Nützlich sind die Ergebnisse also nur, wenn sie erstens sicher und zweitens selten sind. Damit geraten vor allem Minderheiten in den Blick. Wenn die Methoden wirklich so effektiv eingesetzt werden können, wie die Befürworterinnen und Befürworter meinen, sollte man also eine erhöhte Aufklärungsrate in diesen Gruppen erwarten. Die Aufklärungsrate in der Mehrheitsbevölkerung wird sich hingegen überhaupt nicht verändern, weil Analyse-Ergebnisse in Ermangelung eines Fokussierungsgewinns regelmäßig verworfen würden. Und während sich Minderheitenangehörige einem wiederholten, erhöhten Ermittlungsdruck durch Anschlussmaßnahmen ausgesetzt sähen (z. B. DNA-Reihenuntersuchungen), würde sich der Ermittlungsdruck auf die Mehrheitsbevölkerung nicht erhöhen. Erhöhter Ermittlungsdruck kann zu höheren Aufklärungsquoten führen, was wünschenswert

¹⁵ Vgl. Rath 2019.

ist. Wenn dies aber nur auf bestimmte Gruppen zutrifft, würde dies Folgen für die Kriminalstatistik haben: Der Anteil der Minderheitenangehörigen unter allen Tatverdächtigen und auch unter allen überführten Täterinnen und Tätern wäre überrepräsentiert; und letztlich auch in den DNA-Datenbanken, in denen die aus Reihenuntersuchungen gewonnenen DNA-Daten unschuldiger Bürgerinnen und Bürger zeitweise aufbewahrt werden.

Es sei ausdrücklich betont: Dieser Diskriminierungseffekt ergibt sich aus der begrenzten Nützlichkeit spezifischer Analyse-Ergebnisse, nicht etwa aus rassistischen oder stereotypen Einstellungen der Ermittlerinnen und Ermittler. Auch wenn alle Ermittlerinnen und Ermittler von vornherein anti-rassistisch eingestellt sind und zusätzlich Anti-Bias-Trainings durchlaufen, würde die hier skizzierte Anwendung der Technologien diesen Effekt zeitigen.

Es bleibt zudem eine logistische und gesellschaftlich heikle Herausforderung für die Ermittlerinnen und Ermittler, diejenigen, die die gesuchten Merkmale aufweisen, auffindig zu machen. Die obengenannte Überlegung, Menschen in der Nähe des Tatortes zu überprüfen, zeigt dies besonders deutlich. Analyse-Ergebnisse würden hier nur weiterhelfen, wenn sie ein äußerlich klar sichtbares Merkmal ergeben, das nicht allzu häufig vorkommt. Weil solche Merkmale – z. B. dunkle Haut – in Deutschland ohnehin als auffällig gelten, hätte dies einen zusätzlich stigmatisierenden Effekt, wenn diese Maßnahme öffentlich erfolgt.

Man könnte eine Abfrage bei der Meldebehörde zu Geburtsort oder Nationalität in Betracht ziehen, was jedoch diejenigen nicht erfassen würde, die in Deutschland geboren wurden oder die deutsche Staatsangehörigkeit haben. Sich auf Sammelunterkünfte oder Gruppentreffpunkte zu fokussieren, wo sich mehrere Menschen mit den gesuchten Merkmalen aufhalten, mag naheliegend erscheinen; dieses Vorgehen wäre aber nicht logisch begründbar, würde nur einen spezifischen Teil der zu fokussierenden Gruppe erfassen und von den Betroffenen womöglich als stigmatisierend empfunden. Die Frage lautet also, wie die Ermittler konkret vorgehen müssten, um dies zu vermeiden.

Manche Ermittlerinnen und Ermittler, die sich in der deutschsprachigen Debatte zu den Technologien geäußert haben, scheinen den uniparentalen BGA-Markern die größte Aussagekraft beizumessen: So würde etwa ein Y-Chromosom, wie es vorwiegend „am Hindukusch“ vorkommt, für aussagekräftig gehalten.¹⁶ Damit lässt sich eine weitere grundlegende Problematik der Anwendung aufzeigen: Die von den Genetikerinnen und Genetikern erzielten Ergebnisse werden in Kategorien ausgegeben, die die Daten in genetischen Datenbanken strukturieren, die aber nicht unbedingt mit den sozialen, administrativen und/ oder gesellschaftlichen Kategorien am Tatort übereinstimmen. Wie übersetzt man „am Hindukusch“ (wo es laut der Referenzdatenbank YHRD (Y-STR Haplotype Reference Database) einen bestimmten Y-chromosomalen Haplotyp besonders häufig gibt) in eine Kategorie, mit der die Polizei etwas anfangen kann, ohne auf stereotype Assoziationen zurückgreifen zu müssen? Bedeutet dies, die Täterin bzw. der Täter ist entweder Afghanin bzw. Afghane oder Pakistani? Mitnichten! Dort ist der Haplotyp in populationsgenetischen Untersuchungen lediglich am häufigsten aufgefunden worden. Der dazugehörige Mensch kann eine vollkommen andere Nationalität haben und tausende von Kilometern entfernt geboren worden sein. Andersherum kann ein in Westeuropa (oder auch in Ostasien) am häufigsten anzutreffender Haplotyp bei

¹⁶ Vgl. LKA-Molekularbiologe Andreas Stenger im Interview in einem themenbezogenen Feature des DLF: Geuther / Kazmierczak 2017.

einem Menschen gefunden werden, der am Hindukusch geboren wurde und afghanische Staatsbürgerin bzw. afghanischer Staatsbürger ist.

Was für manche Genetikerinnen und Genetiker „Herkunft“ ist, bedeutet anderen Menschen nichts; aber dafür verstehen andere unter „Herkunft“ etwas ganz anderes als manche Genetikerinnen und Genetiker: etwas kulturelles, familiäres, soziales, lokales, zufälliges oder historisches. Wie sich die Freiburger Stadtgesellschaft in verschiedene soziale Gruppen einteilen lässt, sieht zudem in den Köpfen von Sozialarbeiterinnen und Sozialarbeitern anders aus als in denen von Zugewanderten, Kriminalbeamtinnen und Kriminalbeamten, Mitarbeiterinnen und Mitarbeitern der Meldebehörde, Studierenden oder Supermarktleitenden. Selbst wenn das Labor-Analyse-Ergebnis technisch sehr sicher ist – z.B. hundertprozentig eine Asiatin bzw. ein Asiate: Wie übersetzt sich diese Angabe in all die Kategorien, die andere Beteiligte verwenden?

2.2 Stigmatisierungsgefahr und mangelnde Sensibilität: Fallbeispiel Freiburg

Der Freiburger Mordfall, so haben es viele Expertinnen und Experten öffentlich bestätigt, habe sich von vornherein nicht für die Anwendung der Technologien geeignet, weil im urbanen, weltweiten Kontext die Ergebnisse wenig verlässlich und/ oder kaum nützlich gewesen wären. Zwar hätten bestimmte Ergebnisse für konkrete Fragestellungen eine verlässliche Information liefern können; aber ob sie auch in einer Stadt wie Freiburg, die auf Jahrzehnte gelungener Integration zurückblickt, nützlich für eine Fokussierung gewesen wäre, um die Gruppe der Merkmalsträgerinnen und Merkmalsträger sinnvoll einzuschränken, ist zweifelhaft. Eine ungeschickte Fokussierung hätte stigmatisierende Erfahrungen der betroffenen Gruppe zur Folge gehabt. Diese könnten sich vermeiden lassen, wenn die Aspekte der Verlässlichkeit und Nützlichkeit von vornherein mit Expertinnen und Experten sowie Insiderinnen und Insidern gut durchdacht würden.

Im Freiburger Mordfall äußerten Ermittlerinnen und Ermittler den Wunsch, sie hätten BGA-Bestimmungen durchführen können, um zwei konkrete Aufgaben zu lösen. Der allererste Verdacht fiel auf eine Gruppe Freiburger Stadtbewohner, die aus Gambia kamen, an bestimmten Orten in der Stadt als Gruppe auffielen und von vielen Stadtbewohnerinnen und Stadtbewohnern als besonders kriminell erachtet wurden, vor allem im Bereich der Drogendelikte. Angehörige dieser Gruppe – wie auch viele andere Personen, die nicht als Gruppen betrachtet werden, etwa Studierende oder Touristinnen und Touristen – hielten sich in den Sommermonaten mitunter am Ufer des Flusses auf, in einiger Entfernung zum ebenfalls am Fluß gelegenen Tatort. Die Ermittlerinnen und Ermittler entschieden sich sehr bald für einen DNA-Reihentest an Gambiern, der kein Ergebnis erzielte. Hätte man einen BGA-Test durchführen können, so ein Ermittler später, hätte man sich und den Gambiern diesen Reihentest ersparen können.

Wenn das Testergebnis ganz eindeutig nicht auf Afrika verwiesen hätte – was in diesem Fall wahrscheinlich ist – wäre dies der Entlastung einer Minderheit von einem Verdacht gleichgekommen und hätte Geld sowie Arbeitszeit eingespart. Ob auch erwogen wurde, wie man den Gambiern, die ohnehin öfters als andere Gruppen in Freiburg mit der Polizei zu tun haben, von vornherein eine weitere stigmatisierende Erfahrung hätte ersparen oder diese wenigstens hätte abmildern können, ist nicht bekannt.

Man kann aber auch fragen, ob der Verdacht von vornherein so berechtigt und so

dringend war, dass er eine Maßnahme rechtfertigte, die eine Minderheit belastete und stigmatisierte. Es gab ja laut öffentlich zugänglichen Informationen keinerlei Hinweise auf die Gambier, außer, dass einige von ihnen (wie viele andere Freiburgerinnen und Freiburger auch) im Sommer gerne am Flussufer in der Sonne saßen. Anders als alle anderen Flussufergenießer fielen sie der Polizei jedoch als Gruppe auf. Es ist keine Zeuginnen- oder Zeugenaussage bekannt, die bekundet hätte, dass sich in der Mordnacht im kühlen Herbstmonat Oktober verdächtige Gambier oder Menschen mit dunkler Hautfarbe am Flussufer aufhielten. Die Technologien hätten bei einer solchen Verdachtslage also eher dazu gedient, vagen Verdachtsmomenten der Ermittlerinnen und Ermittler so früh wie möglich die Grundlage zu entziehen. Dies kann nur mit ganz eindeutigen Analyse-Ergebnissen funktionieren, nicht aber mit unsicheren, denn dann wäre der vage Verdacht nicht ausgeräumt. Und es funktioniert auch nur, wenn die DNA-Spur nach einem Analyse-Ergebnis, das dem ursprünglichen Verdacht eindeutig widerspricht, noch als eindeutig vom Täter stammend anerkannt wird). Beides hängt davon ab, ob die Ermittlerinnen bereit sind, ihren Verdacht zu revidieren.

Im vorliegenden Fall konnte es also nur darum gehen, kaum begründete, vage Verdachtsmomente gegen Menschen mit dunkler Hautfarbe oder afrikanischer BGA gleich zu Beginn der Ermittlungen zu entkräften. Dafür sind Erweiterte DNA-Analysen sicherlich nicht das geeignetste und auch nicht das kostengünstigste Mittel. Die Unterscheidung, wann ein Verdachtsmoment begründet ist und wann nicht, ist für Ermittlerinnen und Ermittler oftmals ausgesprochen schwierig. Hier wären Schulungen und fallbezogene Beratungen durch ein interdisziplinäres Expertinnen- und Expertengremium einschließlich internationaler Expertise vermutlich hilfreicher als DNA-Analysen.

Die aktuellen Diskussionen um strukturellen Rassismus und Vorbehalte gegen Minderheiten in staatlichen Institutionen werfen außerdem die Frage auf, ob der Wunsch nach der Entlastung von Minderheiten tatsächlich das Hauptmotiv der Befürwortenden war. Eine solche Entlastung ist dann möglich, wenn sie von den Ermittlerinnen gewünscht ist und angestrebt wird. Selbst aus einem eindeutig entlastenden Analyse-Ergebnis wird sich eine solche Entlastung nicht automatisch ergeben.

Die zweite konkrete Aufgabe, die der Mitarbeiter des badenwürttembergischen Landeskriminalamts Andreas Stenger mithilfe der Technologien bearbeiten wollte, bestand in einer Eingrenzung innerhalb einer zweiten DNA-Reihenuntersuchung: Ein Mantrailer-Hund hatte die Ermittlerinnen und Ermittler zu einem Hörsaal geführt, in dem genau zu diesem Zeitpunkt gut 300 Studierende saßen. Es wurde umgehend ein DNA-Reihentest angeordnet und durchgeführt – die notwendigen Utensilien hatte man durch einen erstaunlichen Zufall einfach so dabei. Hätte man damals „gewusst“, so Stenger, dass der Täter zum Beispiel „vom Hindukusch“ kommt, hätte man nicht alle männlichen Studierenden zum DNA-Test gebeten, sondern nur die mit dieser Herkunft.¹⁷

Die Information „vom Hindukusch“ kann nur einer Y-chromosomalen Analyse entstammen, denn autosomale Analysen erzielen eine solche Angabe nicht. Stenger ging also selbstverständlich davon aus, dass ein Mensch mit einem solchen Haplotyp aus derjenigen Region kommen muss, in der Genetikerinnen und Genetiker den Haplotyp am häufigsten gefunden haben. Er nahm außerdem an, dass man diese Herkunft an äußeren Merkmalen erkennen könne, denn wie sonst hätte man die

¹⁷ Geuther / Kazmierczak 2017.

entsprechenden Studierenden aussuchen wollen. Nichts davon ist stichhaltig; zahlreiche Äußerungen und Publikationen von Genetikerinnen und Genetikern belegen, dass dies nicht dem Stand der Forschung entspricht. Zudem ging es Stenger offensichtlich vor allem um Geld- und Zeitersparnis; gesellschaftliche Sensibilität wurde nicht priorisiert. Dass ein solches Vorgehen für Studierende mit dunklen Haaren, dunklen Augen und nicht ganz hellem Teint eine stigmatisierende Erfahrung bedeutet hätte, liegt auf der Hand.

Es geht hier ausdrücklich nicht darum, die Ermittlerinnen und Ermittler als rassistisch oder diskriminierend zu verunglimpfen. Bias, Voreingenommenheit, Stereotype und Vorurteile gehören zum Alltagsgeschehen und prägen die sozialen Interaktionen aller Menschen.¹⁸ Man kann sie nicht einfach ausschalten, aber man kann für ihre Auswirkungen, wie z. B. Stigmatisierung, sensibilisiert werden. Auch eine höhere Empathie für Minderheitenangehörige, die lebenslang pauschalen Verdächtigungen ausgesetzt sind, obwohl sich die meisten von ihnen gesetzestreu verhalten, könnte zu dieser Sensibilität beitragen.

Die Folgen von Stigmatisierung sind vielfältig: Zum einen ist es für eine Unschuldige bzw. einen Unschuldigen, eine Unbeteiligte bzw. einen Unbeteiligten schwer zu ertragen, als Schwerverbrecherin bzw. Schwerverbrecher verdächtigt zu werden, nur weil man derselben Minderheit zugerechnet wird wie die Täterin bzw. der Täter. Wenn dies wiederholt vorkommt, kann es das Verhältnis zwischen Minderheitsangehörigen und Ermittlungsbehörden dauerhaft belasten. Zum anderen werden sich Mehrheitsangehörige, die sich als ein potenzielles Opfer sehen, vor Angehörigen der fokussierten Minderheit zu fürchten beginnen, wenn die Fokussierung auf diese Minderheit in der Öffentlichkeit bekannt wird. All dies kann den sozialen Frieden gefährden.

Um den möglichen Schaden einer öffentlich sichtbaren Fokussierung gegen ihren Nutzen abzuwägen, müsste man nicht nur fragen: In welcher Bevölkerungsgruppe kommt dieses oder jenes Merkmal am häufigsten vor? Sondern auch: In welchen Bevölkerungsgruppen kommt es außerdem noch vor? Wie groß ist also die Chance, durch eine Fokussierung auf eine bestimmte Gruppe die Täterin bzw. den Täter dort auch tatsächlich zu finden? Wenn diese Chance nicht groß ist, weil das Merkmal auch anderswo vorkommt, ist es dann gerechtfertigt, diese Gruppe zu fokussieren und einem erhöhten Ermittlungsdruck auszusetzen? Kann die Anschlußmaßnahme so gestaltet werden, daß eine Stigmatisierung vermieden wird?

Diese Problematik kann man am Fall des Heilbronner Polizistinnen- und Polizistenmordes aufzeigen.¹⁹ Zu den vier vorab in Betracht gezogenen Gruppen von Verdächtigen zählten die Ermittlerinnen und Ermittler etwa „Landfahrer“ (damit waren Sinti und Roma gemeint), Prostituierte und Drückerkolonnen. Eine uniparentale mtDNA-BGA-Analyse ergab – so berichtet es der zuständige Innsbrucker Genetiker Walter Parson²⁰ –, dass der gesuchte Haplotyp häufig in Osteuropa vorkomme. Die Ermittlerinnen und Ermittler hielten fest, die gesuchte Person habe ihre Wurzeln vermutlich in einer Landfahrerfamilie. Man fokussierte sich auf weibliche Angehörige der Sinti und Roma und nahm hunderte DNA-Proben. Sollte es sich bei dieser Fokussierung tatsächlich um eine Übersetzung der BGA-Angabe „Osteuropa“ in die sozialen Kategorien „Sinti und Roma“ oder „Landfahrer“ gehandelt haben, dann war

¹⁸ Solche Phänomene sind in der Sozialpsychologie gut untersucht. Empfohlen sei etwa der Implicit-Association-Test, der online auf der Website der Universität Harvard zu finden ist.

¹⁹ Die folgenden Passagen stützen sich auf Lipphardt, Anna 2019.

²⁰ Vgl. Parson 2014, 153.

dies einen folgenschwerer Irrtum – folgenschwer vor allem für die betreffende Gruppe, denn in der Öffentlichkeit suchten die Ermittlerinnen und Ermittler nun gezielt und explizit nach „Zigeunerinnen“ und Personen aus dem „Zigeunermilieu“. In der Bevölkerung führte diese Nachricht zu wachsender Angst vor Personen, die dieser Gruppe zugeordnet wurden: Immerhin war die Gesuchte als skrupellose, brutal mordende Verbrecherin beschrieben worden.

2009 flog die sogenannte „Wattestäbchen-Affäre“ auf: Die Polizei hatte an vielen Tatorten mit Wattestäbchen gearbeitet, die im Herstellungsprozess von derselben Verpackungsarbeiterin kontaminiert worden waren; die Herstellerfirma trug daran ebenso wenig Schuld wie die Arbeiterin, deren DNA auf den Wattestäbchen gefunden worden war, denn die Wattestäbchen waren nicht korrekt bestellt worden. Es gab mehr Verliererinnen und Verlierer als Gewinnerinnen und Gewinner: Die Arbeiterin wurde durch eine Indiskretion in den Medien namentlich genannt und stigmatisiert. Die Ermittlungsinstitutionen mußten eingestehen, große Mengen an Steuergeldern und Überstunden in einen peinlichen Irrtum gesteckt zu haben. Der Genetiker aber, der die BGA-Analyse durchgeführt und das Ermittlerteam beraten hatte, feierte das Ergebnis als Erfolg: Die Familie der Fabrikarbeiterin kam nämlich aus Osteuropa, sein BGA-Ergebnis stimmte also – zumindest seiner Interpretation zufolge. Allerdings: Einen Haplotyp zu haben, der besonders häufig in Osteuropa vorkommt, dürfte auf sehr viele Menschen in Deutschland zutreffen, auch auf viele, die von ihrer osteuropäischen Familiengeschichte mütterlicherseits nicht die geringste Ahnung haben. Eine Fokussierung, gar auf Sinti und Roma, hätte ein solches Ergebnis also nicht erlaubt.

Ein wissenschaftlicher „Erfolg“ der BGA-Analysen ist also keineswegs gleichbedeutend mit einem Ermittlungserfolg. Für die Ermittlerinnen und Ermittler war der Irrtum ein weiteres peinliches Detail in einem insgesamt ebenso peinlichen wie großen Skandal. Der baden-württembergische Innenminister entschuldigte sich beim Zentralrat der Sinti und Roma für den hohen Ermittlungsdruck, der kaum gerechtfertigt gewesen war. Die Minderheit der Sinti und Roma war somit zwar vom Verdacht, dem sie monatelang öffentlich ausgesetzt gewesen war, entlastet, aber die Stigmatisierung und deren subjektive Erfahrung dürften noch lange angehalten haben. Zudem forderten Ermittlende auch noch in den Wochen nach der Enttarnung des „Heilbronner Phantoms“, als die Entlastung längst hätte greifen müssen, von diesen Minderheitsangehörigen DNA-Proben ein (Lipphardt, Anna 2019).

2.3 Denkbare Einsatzszenarien und ihre Rahmenbedingungen

Haben die Ermittlerinnen und Ermittler einen guten Anhaltspunkt dafür, dass der Technologie-Einsatz eine konkrete Frage klar beantworten kann und gehen sie im Folgenden mit großer Umsicht und Sensibilität vor, dann kann der Einsatz gelingen. Einige hypothetische Fallbeispiele seien hier angeführt. Es geht dabei nicht um strenge Regeln, die in jedem Fall um jeden Preis eingehalten werden müssen, sondern vielmehr darum, Eckpunkte dafür aufzuzeigen, wie intensive, fallbezogene, interdisziplinäre Vorgehensweisen aussehen könnten, um einen verantwortungsvollen Umgang mit den Technologien zu gewährleisten.

Angenommen, zwei Frauen, die Opfer desselben Sexualverbrechers wurden, machen widersprüchliche Angaben zu dessen vermutlicher Herkunft, die sie von seinen äußerlichen Merkmalen ableiten. Die Ermittlerinnen und Ermittler lassen durch eine BGA-Analyse eine der beiden Vermutungen bestätigen; sie erscheint um sehr viel wahrscheinlicher als die andere. Von einem Beratungsteam mit Kommunikationsex-

peritise unterstützt, entscheiden sie sich für eine vorsichtige, diskrete Ermittlung in der entsprechenden Community, wo der Täter auch gefunden wird.

In einem anderen hypothetischen Fall findet die Polizei ein totes Neugeborenes in der Nähe einer Siedlung, in der ausschließlich Zugewanderte aus einer sehr weit entfernten Region leben, die fast gar keinen Kontakt mit der einheimischen Bevölkerung haben. Von einer Genetikerin oder einem Genetiker lassen die Ermittlerinnen und Ermittler die BGA analysieren: Alle drei DNA Systeme weisen auf dieselbe weit entfernte Region hin, aus der die Zugewanderten kommen; also sind die Eltern eher nicht in der Mehrheitsbevölkerung zu finden. Die Ermittlerinnen und Ermittler beraten stundenlang mit der Genetikerin bzw. dem Genetiker, was die Ergebnisse bedeuten, welche Vorsicht man bei ihrer Interpretation walten lassen muss und wie sie in diesem Fall weiterverfahren sollen. Sie entscheiden sich für ein vertrauliches Gespräch mit einem Mitglied der Community und finden die Eltern des Kindes.

In einem tatsächlichen Fall gelang es Ermittlerinnen und Ermittlern in den Niederlanden, einen Sexualmord mithilfe der Minderheits-Community aufzuklären, aus der der Täter stammte. Die Y-chromosomale BGA-Analyse hatte einen seltenen Haplotyp aus einer Region in der Türkei angezeigt; eine Community aus ebendieser Region lebte in der Nähe des Tatortes und pflegte wenige Kontakte mit der Mehrheitsbevölkerung. Bei der Kontaktaufnahme und Durchführung der Anschlußmaßnahmen wurden die Ermittlerinnen und Ermittler von einem Beratungsteam mit Kommunikationsexpertise unterstützt, was die Akzeptanz der Maßnahme deutlich erhöhte; der Täter konnte so gefunden werden. Wäre der Täter nicht in dieser Community gefunden worden, hätte dies an einem „abgewanderten“ Y-chromosomalen Haplotypen gelegen haben können, oder daran, dass der Täter nur auf der Durchreise war. Weitere Merkmale – autosomale BGA, Haar-, Haut- und Augenfarbe – hätten nur dann eine Fokussierung erlaubt, wenn außer den Angehörigen dieser Community keine anderen Menschen in der Nähe diese Merkmale aufgewiesen hätten.

An einem weiteren hypothetischen Fall lässt sich aufzeigen, welche Fehlerquellen es geben kann, selbst wenn das Ergebnis eindeutig ist und die Gruppen klar voneinander getrennt sind.²¹ Angenommen, in einem Hochhaus in Spanien geschieht ein Mord. In dem Haus leben 1010 alleinstehende Männer, und zwar 1000 mit ausschließlich europäischen Vorfahren und 10 mit ausschließlich ostasiatischen Vorfahren. Letztere sind erst seit einem Jahr im Land. Alle haben dunkle Haare, dunkle Augen und leicht gebräunte Haut, weshalb ein DNA-Phenotyping wenig sinnvoll erscheint. Eine Tatort-DNA-Spur wird analysiert: Die beiden uniparentalen Haplotypen erlauben keine eindeutige Zuordnung zu einer Region, sie kommen weltweit vor.

Aber die autosomale BGA-Analyse ergibt, dass der Täter mit 99-prozentiger Wahrscheinlichkeit ein Mensch mit ostasiatischen Vorfahren sein muß. Die Fehlerrate beträgt also 1 %. Das Ergebnis scheint somit auf die 10 Menschen mit ostasiatischen Vorfahren hinzuweisen. Aber die Wahrscheinlichkeit, dass Menschen mit europäischen Vorfahren fälschlicherweise als Ostasiaten klassifiziert wird, beträgt wegen der testinvarianten Fehlerrate ebenfalls 1 %: Das sind bei 1000 Einwohnern mit europäischen Vorfahren 10 fehlklassifizierte Fälle. Die Chance, die DNA von einem der 10 Ostasiaten korrekt klassifiziert zu haben, steht dem Risiko gegenüber, 10 von 1000 Spaniern irrtümlich als Ostasiaten zu klassifizieren. Die

²¹ Wir danken Fabian Staubach für die Ausformulierung eines Fallbeispiels, das dem hier formulierten Fallbeispiel als Vorlage diente.

Wahrscheinlichkeit, dass die Analyse korrekterweise auf einen Ostasiaten verweist, beträgt also 50 %.

Eines haben alle Fallbeispiele gemeinsam: Die Technologien können am besten eingesetzt werden, wenn die gesellschaftliche Integration von Minderheiten schwach ist, wenn sich also Gruppen ohne viele Kontakte gegenüberstehen. Was aus der Sicht der Gesellschaft meist ein bedauernswerter oder zumindest problematischer Zustand ist, erscheint aus Ermittlungs- und Technologieperspektive als ein Vorteil. Dies bedeutet aber auch, dass die Technologien besonders gut in Situationen angewendet werden können, in denen der gesellschaftliche Frieden ohnehin schon gefährdet sein könnte und die Gefahr für gegenseitige Anfeindungen groß ist. In einer solchen Situation müssen Ermittlerinnen und Ermittler besonders vorsichtig und verantwortungsbewusst vorgehen, um bereits bestehende Spannungen nicht zu verschärfen.

2.4 Anwendbarkeit und Anwendungsrealität

In wie vielen bzw. in welchen Fällen kann man die Technologien nutzbringend anwenden?

Nicht jeder Fall eignet sich für einen Einsatz der Technologien; tatsächlich schränken schon rein ermittlungstaktische Gesichtspunkte die Anzahl der potenziellen Anwendungsfälle stark ein:

- Die DNA-Spur sollte eindeutig dem Täter oder der Täterin zuzuordnen sein. Dies ist bei Sexualverbrechen öfters der Fall als z. B. bei Wohnungseinbrüchen (DNA-Spuren etwa am äußeren Fensterkreuz können auch vom Reinigungspersonal stammen).
- Die Bevölkerungsstruktur sollte hinsichtlich des Analyse-Ergebnisses deutlich in mindestens zwei Gruppen zu unterscheiden sein: z. B. mit BGA von verschiedenen Kontinenten, oder mit entweder heller oder dunkler Hautfarbe.
- Das Analyse-Ergebnis sollte nicht auf „Mischwerte“ bei der Pigmentierung oder auf eine „gemischte“ BGA hinweisen, denn dann ist es wenig verlässlich.
- Die unterscheidbaren Gruppen dürfen miteinander vor Ort nicht zu viele soziale Beziehungen haben, weil sonst „Mischungen“ zu erwarten sind, was zu weniger verlässlichen Ergebnissen führt. In städtischen Bevölkerungen mit einer langen kosmopolitischen Tradition, wie etwa Freiburg, ist der Einsatz kaum aussichtsreich.
- Für die Interpretation und Weiterverwendung von anspruchsvollen, komplizierten Analyse-Ergebnissen bedarf es spezieller Expertisen und somit erhöhter Personalressourcen.

Innerhalb einzelner Tatcategorias sind keine großen Durchbrüche zu erwarten; die meisten Sexualmorde etwa sind Beziehungstaten, die relativ zügig mit herkömmlichen Ermittlungsmethoden aufgeklärt werden können. Für sogenannte „cold cases“ bei denen keinerlei brauchbarer Ermittlungsansatz gefunden wurde, und bei denen kein unmittelbarer Zeitdruck den vorsichtigen und verantwortungsvollen Umgang beeinträchtigt, erscheint die Entscheidung für einen Einsatz gerechtfertigt – wenn es sich um ein Schwerverbrechen handelt.

All das schränkt die Anzahl der Anwendungsfälle, in denen eine ausreichend hohe wissenschaftlich-technische Qualität und gleichzeitig gesellschaftliche Sensibilität erreicht werden können, stark ein.

Können Opferangehörige sich von den Technologien einen Ermittlungserfolg erhoffen?

Das gerade Gesagte bedeutet auch: Bei weitem nicht jeder bzw. jede Opferangehörige kann sich vom Einsatz der Technologien die Aufklärung eines Verbrechens erhoffen; in großen Städten wäre dies z.B. eine trügerische Hoffnung. Umgekehrt werden Opferangehörige wenig Verständnis dafür haben, wenn die Technologien, die ihnen öffentlich als hochpotente Ermittlungsinstrumente vorgestellt wurden, gerade in ihrem Fall nicht den versprochenen Durchbruch liefern können.

Wer soll die Technologien anwenden, und wie oft?

In den Gesetzesentwürfen von 2017 und 2019 wurde eine technische Ausrüstung beschrieben, die vermuten lässt, dass es nicht um Einzelfallanwendungen geht, sondern eher um Routine-Anwendungen, und zwar durchgeführt an den LKÄ, für die die Geräte gefordert werden. Massenhafte Routine-Anwendungen lassen sich aber nicht in der erforderlichen Qualität durchführen, weil eine sachkompetente Auswertung und Interpretation vor allem Investitionen in speziell ausgebildetes Personal und verstärkte Forschungsaktivitäten erfordert. Höhere Personalkosten für eine fachgerechte Auswertung, für Schulungen und Beratung sind jedoch vom Gesetzgeber nicht vorgesehen, auch keine Investitionen in Forschung und universitäre Forensische Genetik. Wer in einem solchen Szenario die Anwendungsqualität garantieren soll, bleibt unklar. Löhnen würde sich eine solche Anschaffung eher, wenn man die NGS-Sequencing-Maschinen für andere Zwecke verwenden wollte.

Können die Technologien mehr Sicherheit garantieren?

Das Potenzial der Technologien, potenzielle Opfer zu schützen, bleibt auf jene Fälle beschränkt, in denen eine Serientäterin bzw. ein Serientäter von weiteren Gewalttaten abgehalten werden kann. Ohne Zweifel wecken diese Fälle starkes öffentliches Interesse; sie mögen das Sicherheitsgefühl massiv beeinträchtigen, aber sie sind sehr selten. Bei Wohnungseinbrüchen, die ebenfalls das Sicherheitsgefühl massiv beeinträchtigen, wird sich nur selten eine eindeutig der Täterin bzw. dem Täter zuzuordnende DNA-Spur auffinden lassen. Die Technologien werden also keinen flächendeckenden Sicherheitseffekt haben.

Belegen die bisherigen Erfolge, wie nützlich die Technologien sind?

Dem legalen Einsatz der Technologien in anderen Ländern werden nur wenige Erfolge zugeschrieben: Expertinnen und Experten verweisen öffentlich auf ca. fünf bis sechs Fälle in den letzten 20 Jahren, in denen BGA-Analysen (aber nicht FDP-Analysen) als hilfreich bezeichnet wurden. Zu den Einsätzen insgesamt gibt es keine Statistik, keine Begleitforschung, keine retrospektiven Fallanalysen: Es bleibt also selbst in den Ländern, in denen der Einsatz rechtlich möglich ist, unbekannt, wie oft die DNA-Analysen *keinen* brauchbaren Hinweis lieferten, oder in wie vielen Fällen und mit welchen Methoden soziale Gruppen in den Fokus genommen wurden. Die wenigen Erfolgsgeschichten stechen deshalb umso stärker hervor. Wären sämtliche Einsatzgeschichten, auch die ohne nutzbringende Anwendung, mit derselben Offenheit erzählt worden, hätte man in der Öffentlichkeit kaum so wirksam für die Technologien werben können.

3. Die öffentliche Darstellung der Technologien seit 2016/ 2017

Nach dem Freiburger Mordfall wurden die Technologien nicht nur vehement als Ermittlungsinstrument gefordert, ihre Aussagekraft und Zuverlässigkeit wurde auch stark übertrieben.²² Der Freiburger Mordfall wurde einige Monate lang von allen Befürwortenden als *der* Beispielfall schlechthin dargestellt, weil er die Notwendigkeit der Technologien eindringlich vor Augen führe. Erst, als diese unsachgemäße Instrumentalisierung des Mordfalls für die Forderung nach der Gesetzesänderung öffentlich kritisiert wurde, rückten die Befürwortenden allmählich davon ab, auf den Freiburger Mordfall Bezug zu nehmen, und gaben öffentlich zu, daß die Technologien in diesem urbanen Kontext kaum sinnvoll hätten eingesetzt werden können. Denn wie hätte man in Freiburg nach all jenen Männern suchen können, in deren männlicher Vorfahrenreihe irgendwann – z.B. im 17. Jahrhundert – jemand „vom Hindukusch“ vorgekommen war?

Wie man auf der Basis unseres eigens erstellten Medienarchivs zeigen kann,²³ anhand dessen sich die Debatte detailliert nachvollziehen lässt, war dies nicht die erste öffentliche Debatte um diese Technologien. Als der Einsatz Erweiterter DNA-Analysen erstmals 2007–2008 während der Ermittlungen zum Heilbronner Polizistinnen- und Polizistenmord in der deutschsprachigen Medienlandschaft gefordert wurde, hatten sich die Expertinnen und Experten nicht etwa einheitlich befürwortend, sondern differenziert geäußert, zum Teil skeptisch und auf die Grenzen der Technologien verweisend. 2017 wurden zunächst keine differenzierten und skeptischen Stimmen hörbar. Erst öffentliche Proteste gegen die übertrieben positive Darstellung bewirkten einen Wandel – erst danach und allmählich ergab sich eine differenzierte und vielschichtige Debatte.

Die übertrieben positiven Einschätzungen der Technologien gingen einher mit der Forderung, man solle Minderheiten nicht aus Gründen der *political correctness* mit Ermittlungen verschonen; wenn die Täterin bzw. der Täter eine Ausländerin bzw. ein Ausländer sei, müsse auch entsprechend ermittelt werden. Dass die Forderungen nach einer Gesetzesänderung im Zusammenhang mit einem Fall laut wurden, in dem schon sehr früh der vage Verdacht aufkam, ein Ausländer könne der Täter sein, verlieh den Forderungen starken Aufwind – zumal im Wahljahr 2016, als hunderttausende Menschen nach Deutschland kamen, um Asyl zu beantragen. Die Überzeugung, die Zugewanderten seien für steigende Kriminalitätsraten verantwortlich, passte gut zu der Forderung, man dürfe diese in Ermittlungen nicht verschonen. Im Kontext der Migrationsdebatte ließ sich das Sicherheitsbedürfnis der Mehrheitsbevölkerung somit erfolgreich mobilisieren, auch wenn die Technologien de facto keinen großen Zuwachs an Sicherheit bieten können.

Weiter oben haben wir gezeigt, dass die Analyse-Ergebnisse sich aus logistischen Gründen nur für Ermittlungen eignen, wenn sie auf klar eingrenzbar, möglichst kleine Gruppen verweisen. Man stelle sich vor, es wäre andersherum: Die Analysen würden sich aus irgendeinem technischen oder logistischen Grund ausschließlich für Ermittlungen in der deutschen Mehrheitsbevölkerung eignen. Man stelle sich ferner vor, dies hätte zu einem enormen Anstieg des Ermittlungsdrucks und der Aufklärungsrate in der Mehrheitsbevölkerung geführt, die Minderheiten wären aber aus rein technischen Gründen von beidem verschont geblieben. Dann hätten sich die Befürworterinnen und Befürworter wohl kaum mit derselben Vehemenz für eine Gesetzesänderung eingesetzt. Angenommen, eine BGA-Analyse hätte zu der

²² Vgl. Lipphardt 2018.

²³ Vgl. Weitz/Buchanan 2017. Vgl. auch die von Anna Lipphardt zusammengestellte Chronologie der Debatten-Anfangsphase nach dem Freiburger Mordfall auf unserer Website (<http://www.wie-dna.de/mediale-debatte/>, [11.09.2020]).

Erkenntnis geführt: „Der Freiburger Täter kann definitiv kein Ausländer sein“, so hätte sich kaum politischer Druck für eine Gesetzesänderung zu einem so sensiblen Thema mobilisieren lassen. Auch für Ermittlerinnen und Ermittler wäre ein solches Ergebnis uninteressant gewesen, weil es keine Fokussierung ermöglicht hätte. Um eine Entlastung von Minderheiten ging es in Deutschland damals jedenfalls nur sehr wenigen Befürworterinnen und Befürwortern.

Von Relevanz ist hier ein ausgewogenes Erwartungsmanagement: Wie wird öffentliche Akzeptanz hergestellt? Wie wird über die Technologien kommuniziert? Allzu große Versprechungen über ihre Leistungsfähigkeit werden einen unrealistisch hohen Erwartungsdruck erzeugen. Wenn diese Versprechungen dann nicht eingelöst werden können, weil nur in einem Bruchteil der Fälle ein brauchbares und verlässliches Ergebnis erzielt werden kann, wird das enttäuschte Publikum sein Vertrauen in die staatlichen Institutionen und die beteiligten Wissenschaftlerinnen und Wissenschaftler in Frage stellen: Es hatte doch geheißen, man könne in allen Fällen mit fast hundertprozentiger Sicherheit die Herkunft eines Täters herausbekommen; wieso kommt dann in so wenigen Fällen etwas Klares heraus? Ist hier *political correctness* am Werk? Wird der Öffentlichkeit etwas vorenthalten? In vielen Fällen dürfte dieses Mißtrauen auf die Forensischen Genetikerinnen und Genetiker zurückfallen, die sich bei der Interpretation ihrer Ergebnisse verantwortungsbewußt verhalten und keine Verlässlichkeit suggerieren, wo es diese nicht gibt.

Eine bedenkenswerte Folge der allzu positiven und undifferenzierten Darstellung der Technologien im Jahr 2016 war, dass viele Bürgerinnen und Bürger sowie auch Anwenderinnen und Anwender das Potenzial der Technologie missverstanden hatten und überzeugt waren, dass „BGA“ dasselbe sei wie „Rasse“, dass man die „Rasse“ einer Person ohne weitere Umschweife aus ihrer DNA „herauslesen“ könne, und dass gegenteilige (wissenschaftliche) Argumente lediglich auf *political correctness* zurückzuführen seien. Das ist jedoch falsch, und zwar nach den Regeln der *scientific correctness*. Es kann für tatsächliche Anwendungen fatale Folgen haben, wenn Ermittlungsbeteiligte und Laien solche wissenschaftlich falschen Vorstellungen hegen, wenn die Anwendungen nicht umsichtig und verantwortungsvoll begleitet und beraten werden, und wenn öffentlich die Gleichsetzung von „BGA“ und „Rasse“ nicht einheitlich und explizit als unwissenschaftlich ablehnt wird.²⁴

Schon vor dem ersten Einsatz, d. h. bereits in einer Debatte um Gesetzesänderungen, besteht also ethische Verantwortung bei der öffentlichen Kommunikation: Sie sollte auf eine allgemeine Sensibilisierung für alle Fehlerquellen und negativen Begleiterscheinungen abzielen. Expertinnen und Experten in anderen Ländern, etwa in Großbritannien und Australien, ist die vorsichtige Öffentlichkeitskommunikation ein sehr wichtiges Anliegen, wie auch die Wahrung des Datenschutzes.²⁵

²⁴ Dazu gehört auch der Hinweis, dass „race“ und „Rasse“ nicht dasselbe bedeuten; vgl. Lipphardt et al. 2018. Zum wissenschaftlichen Konsens, dass sich die genetische Vielfalt beim Homo sapiens nicht in Rassen aufteilen läßt, sei hier auf einen von zahlreichen einschlägigen Experten verwiesen: vgl. Stoneking 2017, 103 ff. Vgl. auch die „Jenaer Erklärung“, https://www.uni-jena.de/190910_JenaerErklaerung [11.09.2020].

²⁵ Vgl. Scudder et al. 2018; BGA-Analysen können zu Datenschutzverletzungen führen; vgl. auch Kayser 2015, 45. Auch in FDP-Technologien kommen SNP-Marker zum Einsatz, die mit datenschutzrelevanten phänotypischen Merkmalen (z. B. Krankheitsbildern) korrelieren. Um hier die Rechte der Betroffenen zu wahren, müssten diese SNPs aus den Markersets für FDP und BGA entfernt werden. Siehe Bradbury/Köttgen/Staubach 2019.

4. Unsere Vorschläge für eine verantwortungsvolle Regulierung in Deutschland

Dieser Beitrag hat aufgezeigt, wie kompliziert, fehleranfällig, voraussetzungsreich, anspruchsvoll und heikel der Einsatz von Erweiterten DNA-Technologien in Ermittlungen sein kann, wenn nicht entsprechende Vorkehrungen getroffen werden, und wenn sie nicht auf umsichtig begleitete Einzelfälle beschränkt bleiben. Für einen verantwortungsbewussten Einsatz der Technologien bedarf es mehr als nur eines Gesetzes, das es allein den Ermittlungsbehörden anheimstellt, wie oft, auf welche Weise, mit welcher Expertise und mit welchem Grad an Transparenz die Technologien eingesetzt werden sollen. Nicht einmal Richterinnen- bzw. Richtervorbehalt und Fallbeschränkung haben es in das verabschiedete Gesetz geschafft. Dabei sollten wissenschaftlich höchste Qualität (im interdisziplinären Sinn), gesellschaftliche Sensibilität und verantwortungsbewusste Kommunikation zu den drei Standards gehören, an denen jeder geplante Einsatz Erweiterter DNA-Analysen gemessen wird.

Folgende Maßnahmen schlagen wir für die Anwendung und Regulation der Technologien vor:²⁶

- Einrichtung einer interdisziplinär besetzten Nationalen DNA-Ethikkommission für forensische, medizinische und kommerzielle DNA-Anwendungen, unter Einbeziehung aller relevanten Expertisen (einschließlich Minderheitenvertreterinnen und -vertreter) sowie verschiedener Positionen und einschließlich internationaler Expertinnen und Experten. Diese Kommission sollte auch eine beratende Funktion für weitere Gesetzesänderungsanträge haben.
- Bildung eines interdisziplinär besetzten Unterausschusses dieser Kommission für „cold cases“ und aktuelle Einzelfälle.²⁷
- Fallbezogene Kommunikationsberatung durch interdisziplinäre Teams.
- Ausbildung, Sensibilisierung, Anti-Bias-Training für Anwenderinnen und Anwender.
- Durchführung und/oder Begleitung der Analysen durch universitäre Expertinnen und Experten (Universitäre Forensische Genetik, Humangenetik).
- Förderung interdisziplinärer Forschungsprojekte.
- Aufnahme des Richterinnen- bzw. Richtervorbehalts in den Gesetzestext.
- Öffentliche Berichterstattung über bisherige Anwendungserfahrungen.
- Bereitstellung von Datenschutzressourcen für DNA-Analysen.
- Regelmäßige und unabhängige Prüfung von DNA-Datenbanken und DNA-Analyseverfahren sowie von FDP- und BGA Technologie-Einsätzen (v. a. in den LKÄ). Dazu gehört auch die Prüfung der verwendeten Referenzdatenbanken.
- Klärung des Validierungsprozederes forensischer Technologien.

Literaturverzeichnis

Bradbury, Cedric / Köttgen, Anna / Staubach, Fabian (2019): Off-target Phenotypes in Forensic DNA Phenotyping and Biogeographic Ancestry Inference: A Resource. In: *Forensic Science International: Genetics* 38, 93–104.

²⁶ In Anlehnung an Lipphardt et al. 2017.

Buchanan, Nicholas / Staubach, Fabian / Wienroth, Matthias / Pfaffelhuber, Peter / Surdu, Mihai / Lipphardt, Anna / Köttgen, Anna / Syndercombe-Court, Denise / Lipphardt, Veronika (2018): Forensic DNA Phenotyping Legislation cannot be Based on „Ideal FDP“ – A Response to Caliebe, Krawczak and Kayser (2017). In: *Forensic Science International: Genetics* 34, e13–e14.

Caliebe, Amke / Walsh, Susan / Liu, Fan / Kayser, Manfred / Krawczak, Michael (2017): Likelihood Ratio and Posterior Odds in Forensic Genetics: Two Sides of the same Coin. In: *Forensic Science International: Genetics* 28, 203–210.

Caliebe, Amke / Krawczak, Michael / Kayser, Manfred (2018): Predictive Values in Forensic DNA Phenotyping are not Necessarily Prevalence-dependent. In: *Forensic Science International: Genetics* 33, e7–e8.

Cheung, Elaine Y. Y. / Gahan, Michelle E. / McNevin, Dennis (2016): Prediction of Biogeographical Ancestry from Genotype: A Comparison of Classifiers. In: *International Journal for Legal Medicine* 131(4), 901–912.

Geuther, Gudula / Kazmierczak, Ludger (2017). Den Tätern auf der Spur. In: *Deutschlandfunk*, vom 20.06.2017. http://www.deutschlandfunk.de/erweiterte-dna-analyse-den-taetern-auf-der-spur.724.de.html?dram:article_id=389187 [28.06.2018].

Kayser, Manfred / Knijff, Peter de. (2011): Improving Human Forensics through Advances in Genetics, Genomics and Molecular Biology. In: *Nature Reviews Genetics* 12, 179–192.

Kayser, Manfred (2015): Forensic DNA Phenotyping: Predicting Human Appearance from Crime Scene Material for Investigative Purposes. In: *Forensic Science International: Genetics* 18, 33–48.

Lipphardt, Veronika / Lipphardt, Anna / Wienroth, Matthias / Staubach, Fabian / Buchanan, Nicholas / Momsen, Carsten / Mupepele, Anne-Christine / Pfaffelhuber, Peter / Plümecke, Tino / Lemke, Thomas / Surdu, Mihai / Köttgen, Anna / Bradbury, Cedric / Weitz, Sarah (2017): Stellungnahme zu drei aktuellen Gesetzesinitiativen für eine Ausweitung von DNA-Analysen in kriminalpolizeilichen Ermittlungen (Änderungen im § 81 StPO), 02.06.2017. <http://www.wie-dna.de/stellungnahme-zu-drei-aktuellen-gesetzesinitiativen-fuer-eine-aus-weitung-von-dna-analysen-in-kriminalpolizeilichen-ermittlungen-aenderungen-im-81-stpo/> [29.07.2020].

Lipphardt, Veronika / Lipphardt, Anna / M'charek, Amade / Momsen, Carsten / Pfaffelhuber, Peter / Mupepele, Anne-Christine / Plümecke, Tino / Reardon, Jenny / Schredelseker, Theresa / Surdu, Mihai / Syndercombe-Court, Denise / Wienroth, Matthias (2018): Lost in Translation. In: *Süddeutsche Zeitung* 113 (17.05.2018), 18.

Lipphardt, Veronika (2018): Vertane Chancen? Die aktuelle politische Debatte um Erweiterte DNA-Analysen in Ermittlungsverfahren. In: *Berichte zur Deutschen Wissenschaftsgeschichte* 41(3), 279–301.

Lipphardt, Anna (2019): Die Erfindung des „Heilbronner Phantoms“. Kulturanthropologische Annäherungen an den NSU-Komplex. In: *Zeitschrift für Volkskunde* 1, 50–70.

M'charek, Amade (2008): Silent Witness, Articulate Collective: DNA Evidence and the Inference of Visible Traits. In: *Bioethics* 22(9), 519–528.

M'charek, Amade / Toom, Victor / Jong, Lisette (2020): The Trouble with Race in Forensic Identification. In: *Science, Technology and Human Values* 45(5), 804–828.

- Ossorio, Pilar N. (2006): About Face: Forensic Genetic Testing for Race and Visible Traits. In: *The Journal of Law, Medicine & Ethics* 34(2), 277–292
- Parson, Walther (2014): *Irgendwann kommt alles ans Licht*. Elsbethen: Ecowin.
- Pfaffelhuber, Peter / Grundner-Culemann, Franziska / Lipphardt, Veronika / Baumdicker, Franz (2020), 'How to choose sets of ancestry informative markers: A supervised feature selection approach'. In: *Forensic Science International: Genetics* 46.
- Rath, Christian (2019): Erweiterte DNA-Fahndung. Hat auch nichts gebracht. In: *Die Tageszeitung*, vom 01.08.2019.
- Referentenentwurf des Bundesministeriums der Justiz und für Verbraucherschutz (2019): *Entwurf eines Gesetzes zur Modernisierung des Strafverfahrens*, vom 08.08.2019.
https://www.bmjv.de/SharedDocs/Gesetzgebungsverfahren/DE/Modernisierung_Strafverfahren.html [30.07.2020].
- Schneider, Peter M. / Autoren-Kollektiv STS@Freiburg (2017): Pro und Contra: Steckbrief per DNA-Spur? In: *Recht und Politik* 53(2), 220–221.
- Schneider, Peter M. (2016): Stellungnahme der Spurenkommission zu den Möglichkeiten und Grenzen der DNA-gestützten Vorhersage äußerer Körpermerkmale, der biogeographischen Herkunft und des Alters unbekannter Personen anhand von Tatortspuren im Rahmen polizeilicher Ermittlungen, vom 14.12.2016, S. 1. http://www.gednap.org/wp-content/uploads/2016/12/Stellungnahme_DNA-Vorhersage_Spurenkommission_2016-12-141.pdf [18.08.2019].
- Scudder, Nathan / McNevin, Dennis (2018): Massively parallel sequencing and the emergence of forensic genomics: Defining the policy and legal issues for law enforcement. In: *Science and Justice*, 58 (2), 153-158.
- Staubach, Fabian / Buchanan, Nicholas / Köttgen, Anna / Lipphardt, Anna / Lipphardt, Veronika / Mupepele, Anne-Christine / Pfaffelhuber, Peter / Surdu, Mihai / Wienroth, Matthias (2017): Germany: Note Limitations of DNA Legislation. Correspondence. In: *Nature* 545, 30.
- Toom, Victor / Wienroth, Matthias / M'charek, Amade / Prainsack, Barbara / Williams, Robin / Duster, Troy / Heinemann, Torsten / Kruse, Corinna / Machado, Helena / Murphy, Erin (2016): Approaching Ethical, Legal and Social Issues of Emerging Forensic DNA Phenotyping (FDP) Technologies Comprehensively: Reply to 'Forensic DNA Phenotyping: Predicting Human Appearance from Crime Scene Material for Investigative Purposes' by Manfred Kayser. In: *Forensic Science International: Genetics* 22, e1–e4.
- Weitz, Sarah / Buchanan, Nicholas (2017): Eine Technologie der Angstkultur. In: *Freispruch* 11, 13–17.
- Wienroth, Matthias (2018): Socio-technical Disagreements as Ethical Fora: Paragon NaonLab's Forensic DNA Snapshot™ Service at the Intersection of Discourse around Robust Science, Technology Validation, and Commerce. In: *BioSocieties* 15, 28–45.
- Wienroth, Matthias (2020a): Governing Anticipatory Technology Practices. Forensic DNA Phenotyping and the Forensic Genetics Community in Europe. In: *New Genetics and Society* 37, 137–152.

Wienroth, Matthias (2020b) Value beyond scientific validity: Let's RULE (Reliability, Utility, LEgitimacy), *Journal of Responsible Innovation*

Williams, Robin / Wienroth, Matthias (2017): Social and Ethical Aspects of Forensic Genetics: A Critical Review. In: *Forensic Science Review* 29(2), 145–169.

Zieger, Martin / Roewer, Lutz (2019): Plädoyer für eine nationale Ethikkommission für die erweiterte Forensische DNA-Analyse. In: *Rechtsmedizin* 29, 415–418.