

Hohe Wahrscheinlichkeiten?

Befürworter der forensischen DNA-Analyse werben seit jeher mit hohen Wahrscheinlichkeiten. Die gelten aber bereits bei der bislang zulässigen DNA-Spur nur bei möglichst vollständigen, nicht verunreinigten, vermischten oder zersetzten Spuren. Auch für die erweiterte DNA-Analyse werden oft hohe Wahrscheinlichkeiten angegeben. Die Landesregierungen von Baden-Württemberg und Bayern behaupten bei der sog. biogeographischen Herkunft eine Vorhersagewahrscheinlichkeit von 99,9 Prozent. **Peter Pfaffelhuber** hat nachgerechnet und kommt zu anderen Ergebnissen.

In den Gesetzesanträgen zur Einführung der erweiterten DNA-Analyse (DNA-Phänotypisierung) der Länder Bayern und Baden-Württemberg¹ wird gefordert, die gesetzlichen Grundlagen zur Untersuchung von DNA-fähigem Spurenmaterial dringend an neue wissenschaftliche Erkenntnismöglichkeiten anzupassen. Insbesondere wird in diesen Anträgen über sehr hohe Wahrscheinlichkeiten geredet, mit denen neue wissenschaftliche Methoden aus DNA-Spuren phänotypische Merkmale eines Spurenlegers vorhersagen können. Es ist die Rede davon, Augenfarbe (blau oder braun) mit 90 - 95 Prozent, Haarfarbe (rot, blond, braun oder schwarz) mit 75 bis 90 Prozent, Hautfarbe (hell und dunkel) mit 98 Prozent und die biogeographische kontinentale Herkunft mit 99,9 Prozent Wahrscheinlichkeit korrekt vorhersagen zu können. Diese Wahrscheinlichkeiten sind bei dem Gesetzesantrag zentral, denn es sollen natürlich nur zuverlässige Methoden in Ermittlungen zum Einsatz kommen. Genetische Methoden - allen voran der genetische Fingerabdruck - haben heute den Ruf, zuverlässig zu sein. Doch ist auch die erweiterte DNA-Analyse zuverlässig genug, um in Ermittlungen eingesetzt zu werden? Wir wollen im folgenden die Wahrscheinlichkeiten, die die Zuverlässigkeit der Methode demonstrieren sollen, etwas genauer unter die Lupe nehmen.

Erklärung und Kritik der Methode für die Vorhersage der Augen-, Haut und Haarfarbe

Betrachten wir das Beispiel der Vorhersage der Augenfarbe. Eine bekannte wissenschaftliche Methode ist Irisplex (bzw. HIrisplex zur gleichzeitigen Bestimmung der Haarfarbe)². Hier werden sechs ge-

netische Marker auf fünf verschiedenen Chromosomen betrachtet. Die Marker sind SNPs (Single Nucleotide Polymorphisms), bei denen es zwei verschiedene Ausprägungen auf jedem Chromosom gibt. Für die sechs Marker und den resultierenden $3^6 = 729$ Möglichkeiten (an jedem der sechs Marker kann ein Individuum homozygot für einen der beiden Möglichkeiten, oder aber heterozygot sein) leitet Irisplex eine Vorhersage für die Augenfarben Blau, Braun oder Gemischtfarbig her. Kalibriert wird diese Methode mit Hilfe eines Datensatzes mit Personen, bei denen sowohl die genetischen Marker als auch die phänotypischen Augenfarbe bekannt sind. Die im Gesetzesantrag benannten 90 - 95 Prozent erhält man durch die sogenannte »Area under the Curve« (AUC), wobei eine Kurve gemeint ist, die falsch-positive (also etwa Blauäugige, die aber durch die Methode als nicht-blauäugig klassifiziert werden) gegen richtig-positive (also Blauäugige, die auch als solche klassifiziert werden) aufträgt. Da die Methode wenig falsch-positive, aber viele richtig-positive liefern sollte, ist ein Wert möglichst nahe an 100 Prozent ein guter AUC-Wert, wobei Irisplex eben 90 - 95 Prozent erreicht.

Der AUC-Wert ist in der Statistik durchaus umstritten, was die Vorhersagekraft einer Methode betrifft. Zumindest kann man annähernd sagen, dass ein AUC-Wert von 95 Prozent bedeutet, dass (in der untersuchten Stichprobe) ein Blauäugiger mit 95 Prozent Wahrscheinlichkeit auch als blauäugig klassifiziert wird. Wichtig ist hier, dass es sich um eine *Aussage über die Grundgesamtheit der Blauäugigen* handelt, und von diesen Blauäugigen werden 19 von 20 richtig klassifiziert. In der Ermittlungssituation weiß man ja aber gar nicht, welche Augenfarbe vorliegt, weil

man ja gerade versucht, diese über eine DNA-Analyse herauszufinden. Man muss hier vielmehr von der Grundgesamtheit der von der Methode als blauäugig klassifizierten ausgehen, und zusehen, mit welcher Häufigkeit in der Tat ein blauäugiger Spurenleger vorliegt.

Obwohl der AUC-Wert im Ermittlungsfall zu wenig informativ ist, werden in wissenschaftlichen Publikationen nur sehr selten die informativeren »prädiktiven Werte« genannt. Diese für den Ermittler relevanten Werte erklären sich wie folgt: Nehmen wir weiter das Beispiel einer blauen Augenfarbe, so ist der positive prädiktive Wert genau die Häufigkeit der Blauäugigen in der Grundgesamtheit der als blauäugig Klassifizierten. Diese Größe lässt sich genau wie die AUC aus der untersuchten Stichprobe herausfinden, denn man kennt dort ja auch die Grundgesamtheit der als blauäugig Klassifizierten, genau wie ihre Augenfarbe. Analog ist der negative prädiktive Wert die Häufigkeit der nicht-Blauäugigen unter allen, die als nicht blauäugig klassifiziert werden. In *Caliebe et al*³ erklären die Autoren nicht nur genau die Sachverhalte, die dazu führen, dass nur die prädiktiven Werte die im Ermittlungsfall relevanten Größen sind, sondern berichten zumindest für die Augenfarbe solche Werte. Es finden sich für blaue Augen positiv-prädiktive Werte zwischen 84 und 94 Prozent, und negativ-prädiktive Werte zwischen 78 und 98 Prozent (bei einem AUC zwischen 89 und 98 Prozent). Für braune Augen sind die positiv-prädiktiven Werte zwischen 65 und 96 Prozent, und die negativ-prädiktiven zwischen 87 und 98 Prozent (bei einem AUC zwischen 89 und 96 Prozent). Wenn also über eine Vorhersagewahrscheinlichkeit geredet wird, sollten allenfalls diese

gerade beschriebenen prädiktiven Werte im Gesetzesantrag genannt werden.

Höchstwahrscheinlich gibt es eine Vielzahl von genetischen Markern, aber auch Umwelteinflüsse, die für die Pigmentierung von Augen und Haut verantwortlich sind. Da Irisplex mit nur sechs genetischen Markern arbeitet, kann es also keine perfekte Vorhersage liefern. So erklärt es sich auch, warum Irisplex, je nachdem, aus welchem Land - und damit mit welchem genetischen Hintergrund - eine Stichprobe kommt, verschiedene prädiktive Werte erhält. In der gerade erwähnten Veröffentlichung waren die prädiktiven Werte in insgesamt acht Studien aus europäischen Ländern (Norwegen, Estland, UK, Frankreich, Italien, Griechenland, Spanien, Niederlande) untersucht worden. Übertragen wir diese Erkenntnisse auf den Ermittlungsfall, so muss sichergestellt sein, dass die Stichprobe, mit der eine Methode (etwa Irisplex) kalibriert worden ist, die Population, aus die der Spurenleger kommt, repräsentiert. Dies ist eigentlich nur dann möglich, wenn man zahlreiche Studien mit verschiedensten Populationen durchführt und immer wieder zum gleich Schluss kommt, dass Irisplex unabhängig von der Referenzpopulation hohe Vorhersagewahrscheinlichkeiten liefert. Da es momentan sehr wenige Studien gibt, die prädiktive Werte angeben - mir ist nur *Caliebe et al (2017)* bekannt - scheint die Präzision der Methode noch nicht hinreichend geklärt.

Die biogeographische Herkunft

Der Antrag Bayerns (BR-Drucksache 117/1/17) unterscheidet sich vom vorherigen Gesetzesantrag Baden-Württembergs (BR-Drucksache 117/17) an einer entscheidenden Stelle: Es wird beantragt, die biogeographische Herkunft eines Spurenlegers ermitteln zu dürfen. Dieses Merkmal ist prinzipiell verschieden von den anderen in den Gesetzesanträgen genannten. Zum einen ist fraglich, ob es sich wirklich um ein »äußeres Merkmal« handelt. Zum anderen werden hier viele SNPs auch aus nicht-kodierenden genetischen Regionen (d.h. die DNA zwischen den Genen) verwendet. Aufgrund dieser Unterschiede und da gerade bei diesem Merkmal eine äußerst hohen Vorhersagewahrscheinlichkeit von 99,9 Prozent genannt wird, möchte ich gesondert auf die biogeographische Herkunft im Rahmen der erweiterten DNA-Analysen eingehen.

Zunächst einmal ist es recht schwierig, eine genaue Definition zu geben, was überhaupt unter der biogeographischen Herkunft einer Person gemeint ist. Der

Begriff will einerseits geographische Orte und andererseits genetische (und damit biologische) Aspekte der Herkunft einer Person verbinden. Das ist nicht unplausibel, da genetische Marker in verschiedenen Regionen der Welt unterschiedlich oft vorkommen. (Man denke etwa an einen Marker für helle Hautfarbe in Europa.) Die im Gesetzesantrag genannten 99,9 Prozent Vorhersagewahrscheinlichkeit jedoch beziehen sich auf die kontinentale Herkunft, also den Geburtskontinent einer Person. Eine wichtige Einschränkung hierbei ist es, dass aus Nordamerika nur Ureinwohner als Personen mit kontinentaler Herkunft Nordamerika gezählt werden. Die Mehrzahl der Nordamerikaner hat europäische Wurzeln und wäre deshalb nicht von einer europäischen biogeographischen Herkunft zu unterscheiden. Weiterhin ist erwähnenswert, dass die beantragte Gesetzesänderung die biogeographische Herkunft allgemein ermöglichen soll, und damit auch feinere Auflösungen als die kontinentale möglich werden soll, obwohl nur bei der groben kontinentalen Auflösung gute Vorhersagewahrscheinlichkeiten angegeben werden.

Ein bekanntes Verfahren für die Bestimmung der biogeographischen Herkunft ist etwa EUROFORGENI⁴. Hier werden deutlich mehr (nämlich 128) genetische Marker verwendet als bei der Bestimmung der Augenfarbe. Um die Methode zu kalibrieren, benötigt man große weltweite Stichproben mit vielen genetischen Markern. Hierzu greift man auf Daten aus Großprojekten zurück, die heute unter den Namen »1.000 Genomes Project« und »Human Genome Diversity Project« (HGDP) bekannt sind. Der Leser kann sich sogar selbst ein Bild dieser Daten machen, da sie frei zugänglich sind.⁵

Bei der Erstellung dieser genetischen Datenbanken war es ein Ziel, die menschliche genetische Vielfalt zu erkunden. Kalibriert man die Methode zur Vorhersage der biogeographischen Herkunft mit Hilfe der »1.000 Genomes«-Daten und verwendet man die so entstehende Methode auf die HGDP-Daten an, so entstehen sehr gute Vorhersagewahrscheinlichkeiten von nahezu 100 Prozent.⁶

Die Interpretation dieser hohen Vorhersagewahrscheinlichkeiten im Ermittlungsfall setzt voraus, dass die Population, aus die der Spurenleger kommt, in etwa genauso zusammengesetzt ist wie die in der »1.000 Genomes«-Datenbank. Aus zwei Gründen ist dies eine unrealistische Annahme. Erstens bilden die »1.000 Genomes« zwar eine weltweite Stichprobe, jedoch gibt es große blinde Flecken auf der Landkarte

(etwa der gesamte Nahe Osten), von denen keine Stichproben enthalten sind. Zweitens - und vermutlich im Ermittlungsfall deutlich wichtiger - bilden die Stichproben in den Großprojekten nicht die genetische Vielfalt von Individuen ab, die durch Mischung genetischen Materials verschiedener Kontinente entstanden sein kann. Schon die Definition der biogeographischen Herkunft eines in Europa geborenen Kindes einer Asiatin und eines Afrikaners stellt hier eine fast unüberwindliche Hürde dar. Ein Ausweg wäre es, Methoden wie EUROFORGENI in ethnisch gemischten Populationen anzuwenden und die entsprechenden prädiktiven Werte zu ermitteln. Dies ist jedoch bis heute nicht geschehen.

Ziel der Gesetzesinitiativen zu erweiterten DNA-Analysen ist es, die Voraussetzungen für den Einsatz neuer wissenschaftlicher Erkenntnismöglichkeiten zu schaffen. Doch sind die wissenschaftlichen Fortschritte wirklich schon einsatzbereit? Gerade bei der Vorhersage der biogeographischen Herkunft darf angezweifelt werden, dass die Datenbanken, mit denen die erweiterte DNA-Analyse kalibriert wird, für den Ermittlungsfall repräsentativ ist. Bei Augen- Haut- und Haarfarbe bedarf es außerdem weiterer Studien, die die Grenzen der prädiktiven Werte in allen Bereichen der DNA-Phänotypisierung in verschiedenen Populationen untersuchen.

Prof. Dr. Peter Pfaffelhuber lehrt Statistik an der Abteilung für Mathematische Stochastik der Universität Freiburg und ist beteiligt an der Initiative von Wissenschaftler*innen zur erweiterten DNA-Analyse.

1 : BR-Drucksachen 117/17 v. 3.2.2017 und 117/1/17 v. 28.3.2017

2 : siehe etwa *Walsh et al. 2011. Fate of Transgenic DNA from Orally Administered Bt MON810 Maize and Effects on Immune Response and Growth in Pigs. PLoS ONE 6(11): e27177; Walsh et al 2013*

3 : *Caliebe, Walsh et al. 2017. Likelihood ratio and posterior odds in forensic genetics: Two sides of the same coin. Forensic Science International: Genetics 28: 203–10.*

4 : siehe z.B. *Eduardoff et al, 2016*

5 : unter <http://www.internationalgenome.org/> und <http://www.hagsc.org/hgdp/>. Siehe auch den Browser unter <http://spsmart.cesga.es/>

6 : *Cheung, Gahan, McNevin. 2016. Prediction of biogeographical ancestry from genotype: a comparison of classifiers. International Journal of Legal Medicine.*