



Der Genetische Fingerabdruck von heute erklärt - Mit kleinen Übungsaufgaben.

Im unteren Bild sieht man ein Beispiel eines modernen genetischen Fingerabdruckes. Man sieht eine Menge bunter Spitzen mit Nummern. Früher sah das etwas anders aus, da hatte man helle Balken auf einen dunklen Hintergrund. Damals hatte man noch die Gelelektrophorese als Analysemethode genutzt. Man kann das manchmal auch noch in Filmen sehen oder natürlich in Lehrbüchern. Heutzutage nutzt man jedoch die Kapillarelektrophorese, welche den Vorteil hat, die DNS sehr viel feiner aufzutrennen und auch komplexe Mischspuren deutlich darzustellen. Wie man dieses Ergebnis erstellt und liest und was man damit dann anfangen kann, soll im Folgenden erklärt werden.

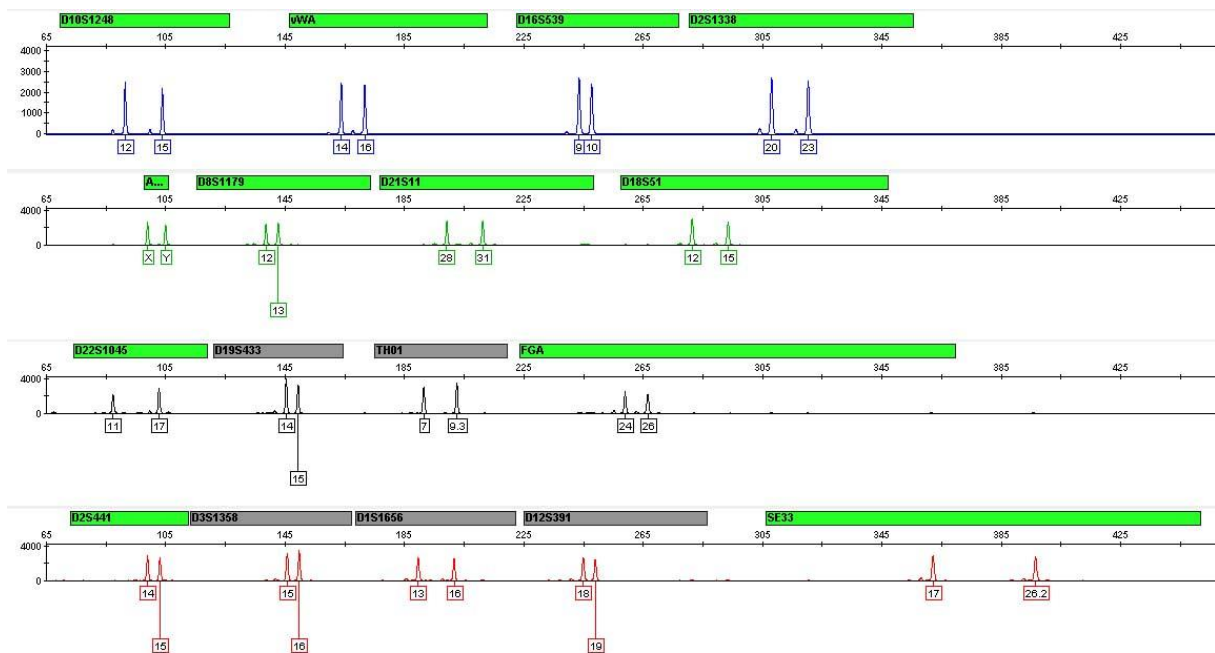


Abbildung 1: Beispiel eines genetischen Fingerabdrucks. Die Abbildung zeigt die Ergebnisse von in gesamt 17 PCR Reaktionen, die gleichzeitig in einer PCR (Multiplex PCR) ablaufen. Diese 17 PCRs vervielfältigen (der molekular Biologe sagt dafür amplifizieren) also 17 kleine Abschnitte unsere DNA. Diese Abschnitte sind für den Körper ohne Funktion und können daher von Mensch zu Mensch sehr unterschiedlich sein. Einer dieser 17 Abschnitte ist das Amelogenin, ein Marker der sich auf den Geschlechtschromosomen befindet und für die Bestimmung des Geschlechtes dient.

Ein genetischer Fingerabdruck hat das Ziel anhand der DNS eines biologischen Materials (z.B. Speichel, Blutanhaftungen, Hautzellen etc.) eine Person zu identifizieren. Das heißt, wir müssen ein genetisches Muster finden das für jede Person einzigartig ist. Und in der Tat ist es so, dass jede Person DNA-Abschnitte besitzt, die sie von anderen Menschen unterscheidet. Die DNA-Abschnitte, die in irgendeiner Weise eine Information besitzen, werden in der Genetik Marker genannt. In der Forensischen Genetik betrachten wir jedoch nur Marker im nicht codierenden Bereich, d.h. außerhalb von Genen, also an Orten die für unseren Phänotyp (Aussehen, Gesundheit...) unwichtig sind. Es gibt sehr viele verschiedene Varianten an Markern, für den genetischen Fingerabdruck nutzt man jedoch heute nur noch die sogenannten short tandem repeat, kurz STRs. Short tandem repeats sind kurze Sequenzwiederholungen von nur wenigen Basenpaaren. Ein Beispiel dafür wäre der in der Abbildung 2 dargestellte Marker D8S1179. Die Bezeichnung des Markers setzt sich wie folgt zusammen: D=DNA, 8=der Marker befindet sich auf dem Chromosom 8, S = single copy (es gibt ihn also nur einmal im gesamten Genom), 1179 = fortlaufende Nummer. In Abbildung 2 erkennt man,



dass der Marker aus dem Sequenzmotiv TCTA besteht, das sich hier genau 12mal wiederholt. Personen können sich anhand dieser Wiederholungen unterscheiden. Für den Marker D8S1179 sind bis jetzt 8 bis 18 mögliche Anzahlen an Wiederholungen bekannt. Es gibt also 11 verschiedene Allele. Jetzt muss man noch berücksichtigen, dass jeder Mensch einen doppelten Chromosomensatz besitzt (also einen Satz von der Mutter und einen vom Vater). Dadurch besitzt jeder Mensch für einen STR Marker zwei Allele (Ausnahme bildeten hierbei das Y-Chromosom und das X-Chromosom beim Mann, sowie die mitochondrielle DNS). Mit zwei Allelen ergeben sich also für den Marker D8S1179 insgesamt 66^1 mögliche Kombinationen. Eingeschlossen sind dabei natürlich auch die Fälle, in dem die Person genau das gleiche Allel von Mutter und Vater geerbt hat, also homozygot ist. Auf der Internetseite <https://strider.online/frequencies> kann man sich für alle in der Forensik genutzten Markersysteme die Allele sowie deren Häufigkeit in den verschiedenen Ländern anschauen.

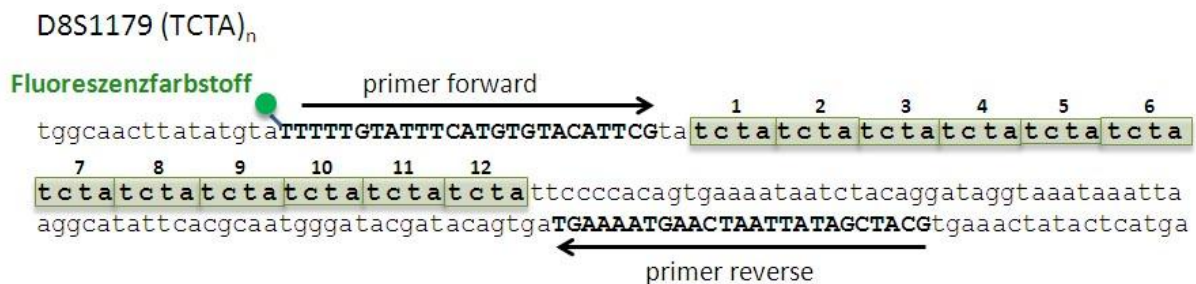


Abbildung 2: DNA Sequenz des STR Marker D8S1179 mit 12 Wiederholungen des Motives TCTA.²

Es ist klar, dass 66 Möglichkeiten noch nicht ausreichen, um eine Person eindeutig zu identifizieren. Bei nur 66 Varianten werden viele Personen genau die gleiche Kombination an Allelen haben. Eine Individualisierung erreicht man, indem man viele solcher STR Marker untersucht und kombiniert, das nennt man dann ein DNA-Profil oder auch genetischer Fingerabdruck. Momentan gelten 15 STR Marker als Europäischer Standard. In Deutschland kommt noch ein 16tes System dazu. Mit der Untersuchung von 16 STR Markern erreicht man eine Einzigartigkeit von ca. Eins zu Einer Trillion bis Quadrillionen. Das in Abbildung 1 gezeigte Profil hat eine Häufigkeit von 1 in 4 Quadrillionen, damit ist es sehr unwahrscheinlich, dass eine zweite Person genau dieses Muster besitzt.

Da man bei der Untersuchung von Spuren oft nicht sehr viel DNA Material zur Verfügung hat, ist es notwendig alle STR Marker gleichzeitig in einer einzigen Analyse zu untersuchen. Wie funktioniert das?

Zunächst einmal muss man die Sequenzen der STR Marker, die man untersuchen möchte, vermehren, man sagt dazu amplifizieren. Dies geschieht durch die Polymerasen-Ketten-Reaktion (<https://de.wikipedia.org/wiki/Polymerase-Kettenreaktion>). Dafür werden Primer vor und hinter den

¹ Die Anzahl der Möglichkeiten kann mit der Formel $(n+k-1)!/(n-1)!k!$ berechnet werden. Dabei ist n=Anzahl der möglichen Allele (11) und k = Anzahl der zusammenzutretenden Elemente (2, also ein mütterliches und ein väterliches Allel). Die Formel ist aus der Kombinatorik für Kombinationen mit Wiederholungen, d.h. die Reihenfolge ist ohne Belang bzw. mit Zurücklegen. Dies trifft auch auf den genetischen Fingerabdruck zu, da es irrelevant ist ob wir z.B. die Allele 11, 12 oder 12, 11 angeben.

² Primersequenzen wurden aus der Publikation Barber, M.D. and B.H. Parkin (Sequence analysis and allelic designation of the two short tandem repeat loci D18S51 and D8S1179. Int J Legal Med, 1996. 109(2):p.62-5) entnommen.



Marker gelegt und mit einem Enzym, der Polymerase, verlängert. In einer einzigen PCR Reaktion kann man sehr viele Sequenzabschnitte (also Marker) gleichzeitig amplifizieren, indem man einfach alle Primer-Paare in eine Reaktion gibt. Als Ergebnis dieser sogenannten Multiplex-PCR bekommt man 16 verschiedene Sequenzen (PCR-Produkte) mit einer genau definierten Länge, welche durch die Primer bestimmt werden. Im Beispiel des D8S1179 Markers aus den Abbildungen 2 und 3, ergibt sich eine Länge von genau 177bp: 12 Wiederholungen je 4 Basen = 48 Basenpaare (bp), plus die Sequenzen zwischen den Marker und den Primern, plus die Sequenzen der Primer selbst.

Für die Analyse in einer Kapillarelektrophorese muss einer der Primer mit einem Fluoreszenzfarbstoff markiert sein. Später in der Kapillarelektrophorese wird der Farbstoff mit einem Laser zum Leuchten gebracht. Dieses Lichtsignal wird dann detektiert und erscheint in der Auswertung als Spitze (Peak) im Elektropherogramm. Benutzt man verschiedene Farbstoffe und designt die entstehenden PCR Produkte zu verschiedenen Längen, ist man in der Lage mehrere STR Marker gleichzeitig, in einer einzigen Reaktion zu analysieren. Als Ausgangsmaterialien braucht man dazu nur ein paar Zellen.

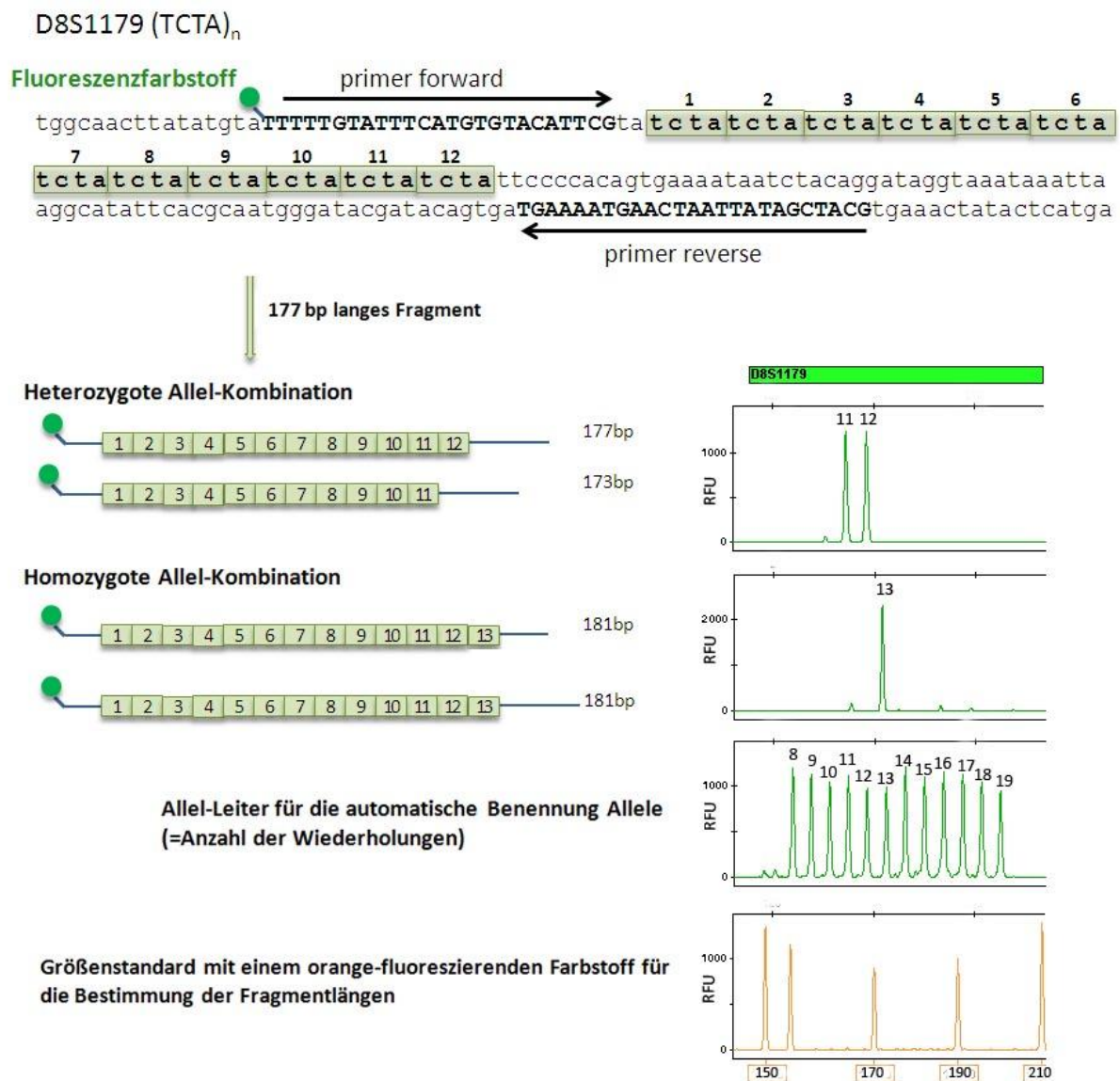


Abbildung 3: STR-Analyse mit der Kapillarelektrophorese.



Bleibt nun nur noch die letzte Frage: Woher weiß man welche Peaks wie viele Sequenzwiederholungen haben? Man errechnet die Anzahl der Wiederholungen über die Länge der Fragmente. In der Abbildung 3 ist es am Beispiel von 12 und 13 Wiederholungen demonstriert. Die Sequenz mit 12 Wiederholungen ergibt ein Produkt mit 177bp, die Sequenz mit 13 Wiederholungen ist genau 4 bp länger und ergibt damit einen neuen Peak. Bei einer Homozygoten Konstellation erhält man natürlich nur einen größeren Peak. In der Kapillarelektrophorese wird die Länge eines Fragments durch einen internen Größenstandard bestimmt, der bei jeder Probe parallel mitläuft. Ein Größenstandard besteht aus Fragmenten mit genau definierter Länge, welche durch einen weiteren Fluoreszenzfarbstoff (hier orange) markiert sind. Anhand der Laufgeschwindigkeit des Standards, kann eine Software anschließend die Länge der Fragmente der Proben bestimmen. Dabei wandern die Fragmente umso langsamer je länger sie sind. Eine sogenannte Allelleiter, welche alle möglichen Fragmentlängen der vorkommenden Allele eines Markers enthält, ermöglicht eine automatische Benennung der Peaks. Das heißt, der Wissenschaftler muss später im Elektropherogram die Allele (= Anzahl der Wiederholungen) nur noch ablesen.

Für Fortgeschrittene: Berechnung der Häufigkeit

Um eine Aussage über die Wahrscheinlichkeit treffen zu können, dass ein ermitteltes DNA-Muster einzigartig ist bzw. kein zweites Mal auftritt, berechnet man dessen Häufigkeit. Wie schon erwähnt, nimmt die Häufigkeit mit größer werdender Anzahl an STR Markern ab. Also je mehr STR Marker man für ein DNA-Profil von einer Person untersucht, umso kleiner ist das Vorkommen (Häufigkeit) dieses Profils, bzw. umso unwahrscheinlicher ist es, dass eine zweite Person genau dieses Profil besitzt. Die Häufigkeit eines DNA-Profiles wird mit Hilfe des Hardy-Weinberg Gesetzes ($p^2 + 2pq + q^2 = 1$, mit **p** und **q** sind die Frequenzen der Allele) berechnet. Die Häufigkeit einer homozygoten Allelkombination lässt sich demnach mit p^2 und einer heterozygoten Allelkombination mit $2pq$ ausrechnen. Anschließend werden dann die berechneten Häufigkeiten der einzelnen Marker miteinander multipliziert. Voraussetzungen für die Anwendung des Hardy-Weinberg Gesetzes sind, dass alle Marker unabhängig voneinander vererbt werden und dass die Frequenzen der einzelnen Allele in einer Population über Generationen hinweg stabil bleibt. Für die in Abbildung 1 dargestellten genetischen Fingerabdruck benutzen Marker liegen fast alle auf unterschiedlichen Chromosomen oder sind sehr weit voneinander entfernt, so dass gezeigt werden konnte, dass hier eine unabhängige Vererbung und damit eine ungerichtete Vermischung der Merkmale gewährleistet ist. In unabhängigen Populationsstudien wurden die Allelfrequenzen der verschiedenen Marker bestimmt und bestätigt und damit auch die zweite Voraussetzung des Hardy Weinberg Gesetzes erfüllt. Abbildung 4 zeigt ein Beispiel für die Berechnung der Häufigkeit des in Abbildung 1 dargestellten DNA-Profiles. Die Häufigkeit liegt bei 1 in 4 Quadrillionen.

D10S1248	VWA	D16S539	D2S1338	AMEL	D8S1179	D21S11	D18S51	D22S1045	D19S433	TH01	FGA	D2S441	D3S1358	D1S1656	D12S391	SE33
12,15	14,16	9,10	20,23	XY	12,13	28,31	12,15	11,17	14,15	7,9,3	24,26	14,15	15,16	13,16	18,19	17,26,2



Die Häufigkeit des angegebenen Profils ist ca. 1 in 4 Quadrillionen ($f = 2.475e-25$) basierend auf der Populationsstichprobe 'ENFSI2015'.

D10S1248	VWA	D16S539	D2S1338	D8S1179	D21S11	D18S51	D22S1045	D19S433	TH01	FGA	D2S441	D3S1358	D1S1656	D12S391	SE33
12, 16	14, 16	9, 10	20, 23	12, 13	28, 31	12, 18	11, 17	14, 16	7, 9,3	24, 26	14, 16	15, 16	13, 16	18, 19	17, 20,2

RECHENWEG

```

D10S1248: 2pq = 2 * 2,83893e-02 * 2,05098e-01 = 1,16452e-02
VWA: 2pq = 2 * 1,03345e-01 * 2,08720e-01 = 4,31403e-02
D16S539: 2pq = 2 * 1,10564e-01 * 5,69718e-02 = 1,25981e-02
D2S1338: 2pq = 2 * 1,44928e-01 * 1,05894e-01 = 3,06938e-02
D8S1179: 2pq = 2 * 1,38857e-01 * 3,27153e-01 = 9,08549e-02
D21S11: 2pq = 2 * 1,53051e-01 * 7,18424e-02 = 2,19912e-02
D18S51: 2pq = 2 * 1,31569e-01 * 6,85858e-02 = 1,80475e-02
D22S1045: 2pq = 2 * 1,37215e-01 * 8,87408e-02 = 2,43532e-02
D19S433: 2pq = 2 * 3,48494e-01 * 1,60873e-01 = 1,12126e-01
TH01: 2pq = 2 * 1,63480e-01 * 2,99730e-01 = 9,79997e-02
FGA: 2pq = 2 * 1,38357e-01 * 2,70531e-02 = 7,48601e-03
D2S441: 2pq = 2 * 2,94941e-01 * 4,33481e-02 = 2,55703e-02
D3S1358: 2pq = 2 * 2,56616e-01 * 2,50145e-01 = 1,28382e-01
D1S1656: 2pq = 2 * 6,83530e-02 * 1,16818e-01 = 1,59697e-02
D12S391: 2pq = 2 * 1,86716e-01 * 1,13922e-01 = 4,25419e-02
SE33: 2pq = 2 * 7,31062e-02 * 5,40853e-02 = 7,90795e-03
-----
= 2,47488e-25
=====

```

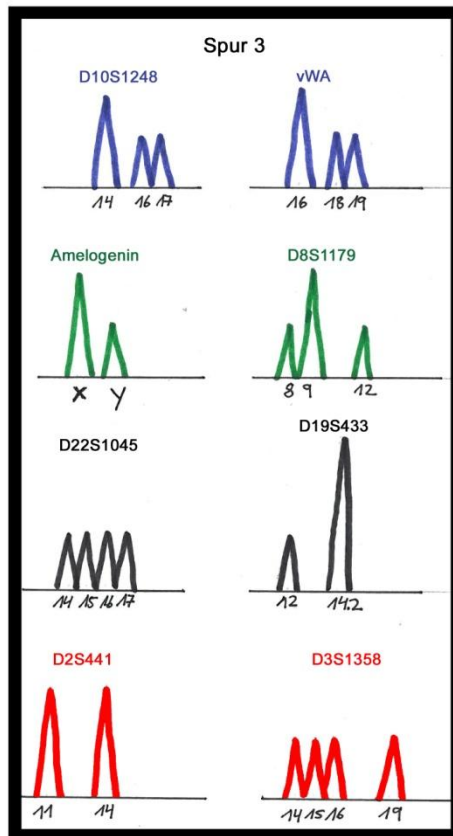
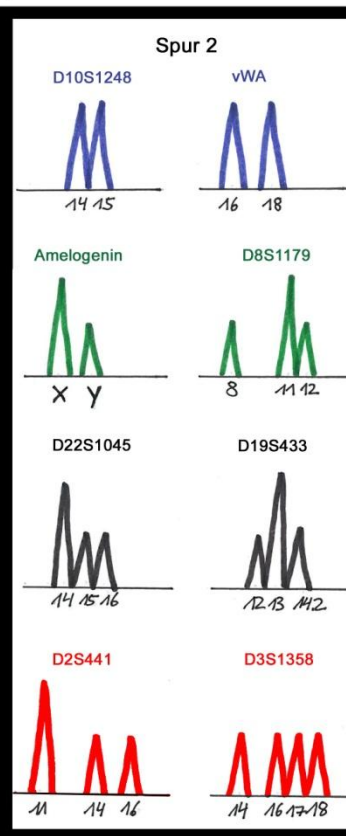
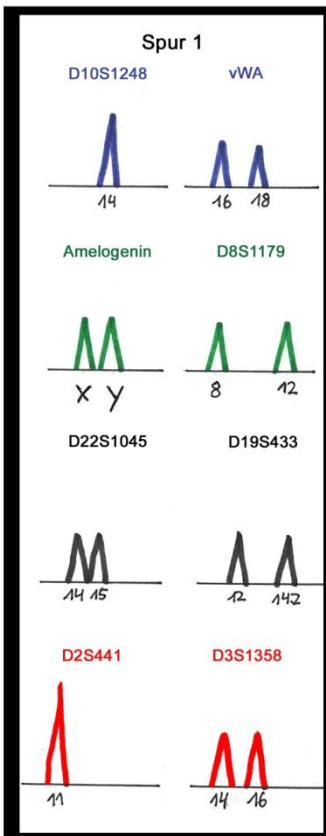
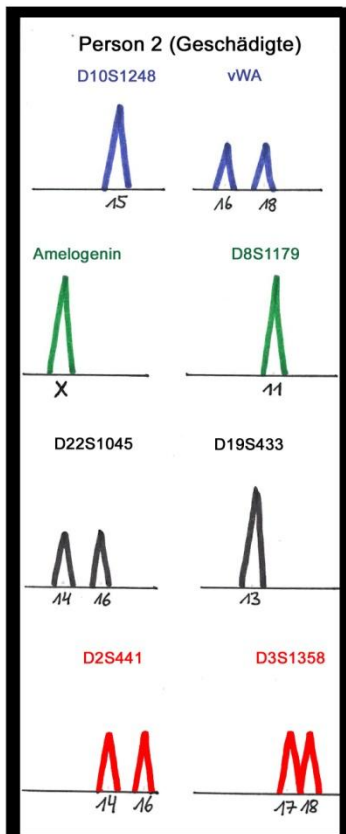
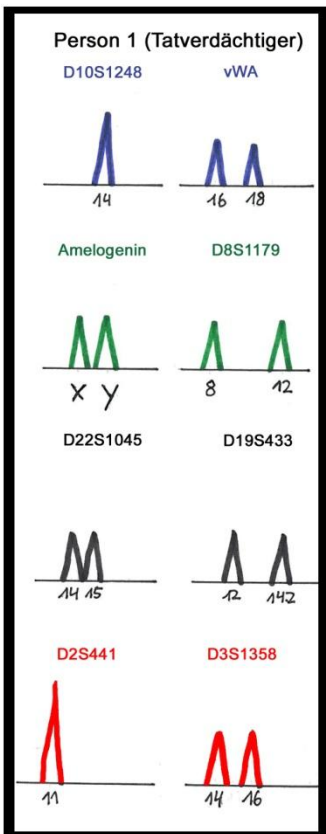
Abbildung 4: Berechnung der Häufigkeit eines DNA-Profiles.

Eine weitere Voraussetzung für die Anwendung des Hardy Weinberg Gesetzes sind unverwandte Personen. Gibt es zum Beispiel zwei verwandte Tatverdächtige, dann ist die Häufigkeit der übereinstimmenden Profile der Tatverdächtigen in den Spuren höher, als die mit der nach dem Hardy Weinberg Gesetzes formulierten Häufigkeit, da verwandte Personen mehr Allele gemeinsam haben als unverwandte Personen. Eine weitere Besonderheit bilden Populationen aus stark isolierten Bereichen, wie zum Beispiel Inseln oder abgeschnittenen Bergregionen. Besonders bedeutsam werden diese Ausnahmen vor allem dann, wenn nur Teilprofile vorliegen. Eine weitere Besonderheit, die jedoch sehr häufig auftritt, sind Mischspuren. Auch in diesem Fall ist die Anwendung des Hardy-Weinberg Gesetzes nicht möglich. Um für Mischspuren eine Aussage zu treffen zu können, wie wahrscheinlich es ist, dass ein DNA-Muster einer Person auftritt und diese Person somit Mitverursacher dieser Spur ist, nutzt man komplizierte Likely-Hood Berechnungen.

Anwendungsaufgaben:



Einer Frau wurde auf offener Straße ihre Handtasche geraubt. Die Handtasche wird kurz darauf in einer Hecke wieder aufgefunden und zur genetischen Analyse ins Labor geschickt. Dort werden mit Watteträgern vom Henkel, dem Zipper und dem Öffnungsbereich drei Proben (Spur 1-3) entnommen. Dabei werden die Watteträger befeuchtet und mehrmals, kräftig über die entsprechenden Stellen gerieben. Da eine hohe Wahrscheinlichkeit besteht, dass man auch genetisches Material der Besitzerin der Handtasche findet, wird von der Polizei eine Vergleichsprobe der Geschädigten entnommen (Person 2). Gleichzeitig gibt es über eine Überwachungskamera auch Hinweise auf einen möglichen Täter. Auch von dieser Person gibt es eine Probe (Person 1). Die untere Abbildung zeigt einen vereinfachten genetischen Fingerabdruck der zwei Vergleichsproben sowie der genommenen Spuren.





Lösungen:

Aufgabe 1:

Person	D10S1248		vWA		D8S1179		D22S1045		D19S433		D2S441		D3S1358	
Person 1	14	14	16	18	8	12	14	15	12	14.2	11	11	14	16
Person 2	15	15	16	18	11	11	14	16	13	13	14	16	17	18

Aufgabe 2:

Bei dem Amelogenin Marker handelt es sich um einen Sequenzabschnitt, der sowohl auf dem X sowie auch auf dem Y Chromosom vorkommt. Der Abschnitt des Y-Chromosoms ist jedoch ein paar Basen länger. Handelt es sich um einen Mann werden in der PCR somit zwei Fragmente erzeugt, der kürzere Abschnitt des X-Chromosom und der etwas längere Abschnitt des Y-Chromosoms. Im Elektropherogramm werden die beiden Peaks ganz einfach X und Y genannt. Erscheint nur der X Peak handelt es sich um eine Frau.

Aufgabe 3

Spur	Spurenart	Begründung
Spur 1	Einzelspur	Nicht mehr als 2 Allele pro Marker
Spur 2	Mischspur	Einige Marker zeigen mehr als 2 Allele
Spur 3	Mischspur	Einige Marker zeigen mehr als 2 Allele

Aufgabe 4

Person 1 tritt in den Spuren 1, 2 und 3 auf.

Person 2 tritt in der Spur 2 auf.



Aufgabe 5

Die Häufigkeit des angegebenen Profils ist ca. 1 in 39 Milliarden ($f = 2.572e-11$)

D10S1248	VWA	D8S1179	D22S1045	D19S433	D2S441	D3S1358
14, 14	16, 18	8, 12	14, 15	12, 14.2	11, 11	14, 16

Rechenweg

D10S1248: $q^2 = 3.13345e-01 * 3.13345e-01 = 9.81850e-02$
 VWA: $2pq = 2 * 2.08720e-01 * 2.13844e-01 = 8.92670e-02$
 D8S1179: $2pq = 2 * 1.24565e-02 * 1.38857e-01 = 3.45935e-03$
 D22S1045: $2pq = 2 * 4.75087e-02 * 3.59212e-01 = 3.41314e-02$
 D19S433: $2pq = 2 * 8.44921e-02 * 2.38509e-02 = 4.03043e-03$
 D2S441: $q^2 = 3.35489e-01 * 3.35489e-01 = 1.12553e-01$
 D3S1358: $2pq = 2 * 1.09523e-01 * 2.50145e-01 = 5.47932e-02$

 = 2.57229e-11
 =====

Aufgabe 6

Person	D10S1248		vWA		D8S1179		D22S1045		D19S433		D2S441		D3S1358	
Unbekannte Person 3	16	17	16	19	9	9	16	17	14.2	14.2	14	14	15	19